

Constructie van kinetische modellen met behulp van chemoinformatica

Kinetic Model Construction Using Chemoinformatics

Nick Vandewiele

Promotoren: prof. dr. ir. G. B. Marin, prof. dr. M.-F. Reyniers
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen: Chemische Technologie

Vakgroep Chemische Proceskunde en Technische Chemie
Voorzitter: prof. dr. ir. G. B. Marin
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2013 - 2014



ISBN 978-90-8578-680-1
NUR 952
Wettelijk depot: D/2014/10.500/26

Promotoren:

prof. dr. ir. Guy B. Marin

Laboratorium voor Chemische Technologie

Vakgroep Chemische Proceskunde en Technische Chemie

Universiteit Gent

prof. dr. Marie-Françoise Reyniers

Laboratorium voor Chemische Technologie

Vakgroep Chemische Proceskunde en Technische Chemie

Universiteit Gent



Fonds Wetenschappelijk Onderzoek
Research Foundation – Flanders



De auteur genoot tijdens de onderzoeksactiviteiten de steun van het Bijzonder Onderzoeksfonds (B/10681/02 –BOF09/DOC/374) en een aspirant mandaat van het Fonds voor Wetenschappelijk Onderzoek (FWO).

EXAMENCOMMISSIE

Leescommissie

Prof. dr. Marie-Françoise Reyniers [promotor]

Laboratorium voor Chemische Technologie

Vakgroep Chemische Proceskunde en Technische Chemie

Universiteit Gent

Prof. dr. ir. Kevin Van Geem [secretaris]

Vakgroep Chemische proceskunde en Chemische Technologie

Faculteit Ingenieurswetenschappen en Architectuur

UGent

Prof. dr. Frederique Battin-Leclercq

Département de Chimie-Physique des Réactions

ENSIC-Nancy

Prof. dr. ir. Sebastian Verhelst

Vakgroep Mechanica van Stroming, Warmte en Verbranding

Faculteit Ingenieurswetenschappen en Architectuur

UGent

Prof. dr. ir. Vassily Hatzimanikatis

Laboratory of Computational Systems Biotechnology

Institut des sciences et ingénierie chimiques

Ecole Polytechnique Fédérale de Lausanne

Andere leden

Prof. dr. ir. Guy B. Marin [promotor]

Laboratorium voor Chemische Technologie

Vakgroep Chemische Proceskunde en Technische Chemie

Universiteit Gent

Prof. dr. ir. Joris Thybaut

Laboratorium voor Chemische Technologie

Vakgroep Chemische Proceskunde en Technische Chemie

Universiteit Gent

Prof. dr. ir. Jan Van Campenhout [voorzitter]

Vakgroep Elektronica en Informatiesystemen

Universiteit Gent

voor mijn ouders

Acknowledgements

In de eerste plaats wil ik mijn promotoren, prof. Marin and prof. Reyniers bedanken voor de kansen die ze mij gegeven hebben. Bedankt ook voor de zorgvuldige revisie van mijn doctoraat en publicaties. Een woordje van dank voor prof. Van Geem kan hier ook niet ontbreken. Kevin, bedankt voor de raad en begeleiding, en niet in het minst om mijn ruwe versies van manuscripten in “menselijker” taal om te zetten.

Many thanks go out to Dr. Stolle, for the fruitful collaboration on terpene rearrangements, and to Prof. Green of the Chemical Engineering Department of the Massachusetts Institute of Technology for giving me the opportunity to spend nine months in his research group. JP-10 was certainly not the most straight-forward choice for modeling. Luckily, my partners in crime, Gregory Magoon, Luwi Oluwole, Jorge Aguilera, Robin Edwards, and later on Connie Gao and Nathan Wa-Wai Yee helped me weather many storms in that regard. Greg, I am glad that you finally caught the white whale of JP-10.

Veel dank ook voor mijn twee thesisstudenten, Maarten Eestermans, and Ruben Van de Vijver, die rechtstreeks tot dit werk hebben bijgedragen. Maarten, zonder jou en je McGyver trucs zou die peristaltische pomp nog altijd gelect hebben. Ruben, zonder jou zou het symmetriegetal van methaan nog altijd 24 geweest zijn.

Acknowledgements

Ook zou ik mijn dank willen uitdrukken voor alle technici die altijd klaarstonden om me te helpen bij de experimenten op de bench-scale reactor. Michaël, Hans, Erwin en Bert, en alle andere technici, bedankt voor al het werk dat jullie verricht hebben. Ook Georges kan hier niet ontbreken: zonder jou geen servers om mijn computer escapades uit te testen. De technische omkadering van het labo zou ook niet mogelijk geweest zijn zonder de steun van de langdurige en structurele financiering van de Vlaamse Overheid in de vorm van de “Multiscale Modeling and Design of Reactions and Reactors (M2dcR2)”, en “Multiscale Analysis and Design for Process Intensification and Innovation (MADPII)” projecten.

The developers of RMG cannot be forgotten neither, in particular Richard West, Michael Harper, Josh Allen, Connie Gao, Greg Magoon, Amrit Jalan, Shamel Merchant. It’s hard to imagine that none of you had a formal training in software development and still were able to maintain and extend such a complex software project.

The incredibly talented people involved in open-source chemoinformatics, in particular Egon Willighagen, John May, Gilleain Torrance, Nina Jeliaskova, and Noel O’Boyle. Among many others, they have helped me solve all sorts of chemoinformatics and programming issues, big and small. Without their support and vision of open-source, open standards, and open data, most of this thesis would not have been possible.

Een woord van dank is ook op zijn plaats voor de vele collega’s van het LCT waarmee ik de 4.5 jaar heb doorgebracht. In het bijzonder voor diegenen die door dezelfde vier muren (of flinterdunne wanden) omgeven waren, en die willens nillens mijn raaskalderij moesten aanhoren: Steven, Carl, Thomas, Ruben, Bart, en Maarten. Anderen werden gespaard tijdens de werkuren, maar kregen de volle lading tijdens de koffiepauzes: Aäron, Andres, Amit, Ezgi, Gonzalo, Hans-Heinrich, Hassan, Jelena, Kaustav, Kostas, Maria, Marko, Mike, Panos, Natalia, Pieter, Yu, en vele anderen. Let’s not forget the folks at M.I.T that made my time in the land of hamburgers a real pleasure: Barb, Yuko, Caleb, Ray, Yu-Shi, Zan, Beat, ...

Graag wil ik ook mijn vrienden bedanken, die mij van het nodige vertier voorzien hebben. Bedankt Ward, Brecht, Imran, Annelies, Stefanie, Sébastien, Kristel, Bart, Quinten, Tom, Dries,

Acknowledgements

Dave, Andres en alle anderen. Ten slotte wil ik mijn familie bedanken, in't bijzonder mijn ouders en broer, die mij door dik en dun gesteund hebben. Finalement, je tiens à remercier Noémie, mon autre moitié, mon refuge, qui peut me garder en équilibre comme aucune autre.

Nick Vandewiele

Gent 2014

Acknowledgements

Contents

Contents.....	i
Notation.....	v
List of structures.....	xiii
Samenvatting.....	xvii
Summary.....	xxiii
Glossary.....	xxix
Chapter 1: Introduction.....	1
1.1 Automated kinetic model generation.....	5
1.1.1 Reaction network generation.....	5
1.1.1.1 Molecule representation.....	5
1.1.1.2 Reaction families.....	8
1.1.1.3 Network size control.....	11
1.1.2 Thermochemistry.....	15
1.1.3 Kinetics.....	18
1.2 Chemoinformatics.....	23
1.2.1 Species uniqueness and canonicalization.....	24
1.2.2 Graph and sub-graph isomorphism.....	27
1.2.3 Stereochemistry and 3D structure.....	32
1.3 Objective and thesis overview.....	35
1.4 References.....	37
Chapter 2: Thermal rearrangement of 2-pinanol.....	43
2.1 Abstract.....	44
2.2 Introduction.....	44
2.3 Experimental Procedures.....	47
2.3.1 Materials.....	47
2.3.2 Experimental reactor setup.....	47
2.3.3 Reactor Modeling.....	50
2.3.4 Regression of experimental data.....	50
2.4 Thermochemical Properties.....	51
2.5 Results and Discussion.....	53
2.5.1 Cis- and trans-2-pinanol rearrangement mechanism.....	53
2.5.2 Pre-exponential factors.....	61

2.5.3	Activation energies	62
2.5.4	2-pinanol reactivity - cyclobutane fragmentation	63
2.5.5	Product selectivity - biradical rearrangements.....	66
2.6	Conclusions	68
2.7	References	70
Chapter 3: Thermal decomposition of Jet Propellant-10		73
3.1	Abstract	73
3.2	Introduction	74
3.3	Experimental Procedures.....	75
3.3.1	Materials	75
3.3.2	Experimental reactor setup.....	77
3.3.3	Conditions in the pyrolysis zone.....	78
3.3.4	Kinetic models	78
3.4	Results and Discussion.....	79
3.4.1	JP-10 conversion as a function of temperature	79
3.4.2	Product distribution.....	80
3.4.2.1	Major pyrolysis products	82
3.4.2.2	Primary decomposition products	84
3.4.2.3	Secondary products.....	88
3.4.2.4	Effect of dilution.....	91
3.4.3	Kinetic model validation.....	92
3.5	Conclusions	99
3.6	References	99
Chapter 4: Kinetic modeling of Jet Propellant-10 pyrolysis		103
4.1	Abstract	103
4.2	Introduction	104
4.3	Computational methods.....	105
4.3.1	Kinetic model construction	105
4.3.2	Thermochemistry	107
4.3.3	Quantum-chemical calculations.....	109
4.4	Reactor modeling and experimental data	110
4.5	Results and Discussion.....	112
4.5.1	Kinetic model validation.....	113
4.5.2	Primary decomposition chemistry	125

4.5.2.1	Exo-TCD initiation	125
4.5.2.2	Tricyclodecyl decomposition reactions	130
4.5.3	Main reaction pathways	133
4.5.3.1	Major decomposition pathways of exo-TCD	135
4.5.3.2	Secondary chemistry and aromatics formation	137
4.6	Conclusions	139
4.7	References	140
Chapter 5: Genesys: basic concepts and functionality		143
5.1	Abstract	144
5.2	Introduction	144
5.3	Methodology and algorithms	146
5.3.1	Input	148
5.3.2	Processing	152
5.3.2.1	Network generation	154
5.3.2.2	Thermochemical properties assignment	157
5.3.2.3	Arrhenius parameter assignment	160
5.3.3	Output	165
5.4	Case studies	165
5.4.1	Case 1: calculation of thermochemical data	165
5.4.2	Case 2: Arrhenius parameters via group additivity	169
5.4.3	Case 3: reaction network generation: thermal decomposition of hexamethyl phosphoric triamide	172
5.5	Conclusions	177
5.6	References	178
Chapter 6: Symmetry calculation for reactants and transition states		183
6.1	Abstract	183
6.2	Introduction	184
6.3	Theory	186
6.4	Results and discussion	195
6.4.1	Validation	204
6.4.2	Application	205
6.5	Conclusions	207
6.6	References	208
Chapter 7: Thermal decomposition of diethylsulfide		211

7.1	Abstract	211
7.2	Introduction	212
7.3	Methodology	214
7.3.1	Model generation	214
7.3.1.1	Thermochemistry	218
7.3.1.2	Kinetics	218
7.3.2	Experimental data and reactor modeling	224
7.4	Results and discussion.....	224
7.5	Conclusions	235
7.6	References	236
Chapter 8: Implementation of stereochemistry in kinetic model generation		239
8.1	Abstract	239
8.2	Introduction	240
8.3	Methodology	241
8.3.1	Stereochemistry of molecules	245
8.3.1.1	Stereocenter detection.....	246
8.3.1.2	Stereoisomer generation	249
8.3.1.3	Steric relation detection	254
8.3.1.4	Chemical constraints.....	258
8.4	Application to pinanol thermal rearrangements	260
8.4.1	Reactants	262
8.4.2	Reaction families	263
8.4.3	Thermochemistry and Kinetics	266
8.4.4	Kinetic model generation	270
8.4.5	Kinetic model validation.....	274
8.5	Conclusions	277
8.6	References	278
Chapter 9: Conclusions and Perspectives.....		281
Appendix A: Experimental procedures for JP-10 pyrolysis experiments		287

Notation

Roman symbols

\tilde{A}	Single-event pre-exponential factor	$\text{m}^3 \text{mol}^{-1} \text{s}^{-1}$ or s^{-1}
A	Pre-exponential factor	$\text{m}^3 \text{mol}^{-1} \text{s}^{-1}$ or s^{-1}
$Aut(G)$	Automorphism group of graph G	-
$ Aut(G) $	Order of the automorphism group	-
\mathbf{b}	Vector of parameter estimates	-
c_p	Constant pressure heat capacity	$\text{J K}^{-1} \text{mol}^{-1}$
E_a	Activation energy	kJ mol^{-1}
E_0	Evans-Polanyi intrinsic barrier	kJ mol^{-1}
$\Delta E(0K)$	Electronic barrier at 0K	kJ mol^{-1}
F	Molar flow rate	mol s^{-1}
F_{calc}	Calculated F-value	-

F_{tab}	Tabulated F-value	-
h	Planck constant	$6.62 \cdot 10^{-34} \text{ J s}^{-1}$
H	Enthalpy	kJ mol^{-1}
i_L	Label-stereocenter	-
k	Rate coefficient	$\text{m}^3 \text{ mol}^{-1} \text{ s}^{-1}$ or s^{-1}
k_B	Boltzmann constant	$1.38 \cdot 10^{-23} \text{ J K}^{-1}$
k_{obs}	Observable rate coefficient	$\text{m}^3 \text{ mol}^{-1} \text{ s}^{-1}$ or s^{-1}
$k_{\infty}(T)$	High-pressure limit rate coefficient	$\text{m}^3 \text{ mol}^{-1} \text{ s}^{-1}$ or s^{-1}
m	Number of equivalent atoms	-
n_{label}	Total number of label-stereocenters	-
n_{pot}	Number of potential stereocenters	-
n_{real}	Number of real stereocenters	-
n_d	Number of unique configurations of the equivalent atoms	-
n_e	Number of single events	-
P	Pressure	Pa
P_{xy}	Atom permutation between atoms x and y	-
q	Partition function per unit volume	-
q_{rot}	Rotational partition function	-
R	Universal gas constant	$8.314 \text{ J mol}^{-1} \text{ K}^{-1}$
r	Reaction rate	$\text{mol m}^{-3} \text{ s}^{-1}$
(R)	“Rectus”, absolute configuration	-

R_{ij}	Residuals for experiment i and response variable j	-
S	Entropy	J K ⁻¹ mol ⁻¹
$s(\bar{\beta})$	Objective function	-
(S)	“ <i>Sinister</i> ”, absolute configuration	-
S^P	Signed permutation associated with permutation P	-
t	Time	S
t_{tab}	Tabulated t-value	-
T	Temperature	K
$\hat{V}(b)$	Estimate for the covariance matrix of the parameters b	-
v_i	Configuration vector of a molecule	-
V_E	Equivalent reactor volume	m ³
V_p	Parameter in Blowers-Masel correlation	-
V_R	Equivalent reactor volume	m ³
X	Conversion	-
y_{ij}	Experimental value for experiment i and response variable j	-
\hat{y}_{ij}	Simulated value for experiment i and response variable j	-
\bar{y}_j	Arithmetic average of experimental response variable j over the complete set of	-

experiments

Greek symbols

α	Evans-Polanyi transfer coefficient	-
$\Delta_f H$	Enthalpy of formation	kJ mol^{-1}
$\Delta_r H$	Reaction enthalpy	kJ mol^{-1}
$\Delta_r S$	Reaction entropy	$\text{J K}^{-1} \text{mol}^{-1}$
ε	Tolerance for rate-based kinetic model construction	-
κ	Tunneling coefficient	-
σ	Rotational symmetry number	-
$\sigma_{\text{reactants}}$	Rotational symmetry number of the reactants of an elementary reaction	-
σ^\ddagger	Rotational symmetry number of transition state structure of an elementary reaction	-
σ_{ext}	External rotational symmetry number	-
$\sigma_{\text{int},i}$	Internal symmetry number accounting for rotor “i”	-
τ	Residence time	s

Sub- and superscripts

\ddagger	Transition state
<i>calc</i>	Calculated
<i>eq</i>	Equilibrium
<i>exo</i>	Exothermic
<i>exp</i>	Experimental
<i>ext</i>	External
<i>f</i>	Forward
<i>o</i>	Standard state
<i>opt</i>	Optical isomers
<i>R</i>	Reference
<i>r</i>	Reverse
<i>res</i>	Resonance
<i>rot</i>	Rotational
<i>tab</i>	Tabulated
<i>tot</i>	Total

Acronyms

BDE	Bond Dissociation Energy
BE	Bond-Electron (BE-Matrix)
CBS	Complete Basis Set

CDK	Chemistry Development Kit
CIP	Cahn, Ingold And Prelog
CPD	1,3-Cyclopentadiene
CPU	Central Processing Unit
CST	Crude Sulfate Turpentine
DES	Diethyl Sulfide
DFT	Density Functional Theory
DU	Dungundji-Ugi (DU-Formalism)
FID	Flame Ionization Detector
GA	Group Additivity
GAV	Group Additivity Value
GC	Gas Chromatography
HBI	Hydrogen Bond Increment
HMPA	Hexamethyl Phosphoric Triamide
InChI	International Chemical Identifier
JP-10	Jet Propellant-10
MS	Mass Spectrometer
NNI	Non-Nearest Neighbor Interaction
ODE	Ordinary Differential Equation
RMG	Reaction Mechanism Generator
RRHO	Rigid-Rotor, Harmonic Oscillator
RSC	Ring Strain Correction
SEMK	Single-Event Microkinetic

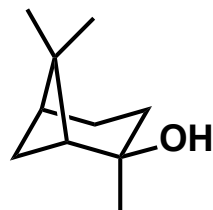
SMARTS	SMiles ARbitrary Target Specification
TCD	Tricyclo[5.2.1.0 ^{2,6}]decane
TS	Transition State
TST	Transition State Theory

List of structures

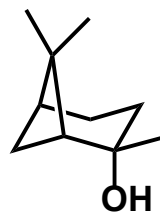
2-norbornene



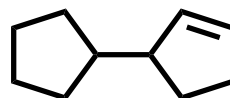
trans-2-pinanol



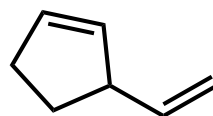
cis-2-pinanol



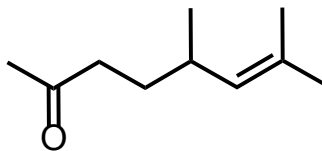
3-cyclopentylcyclopentene



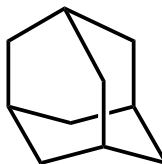
3-ethenylcyclopentene



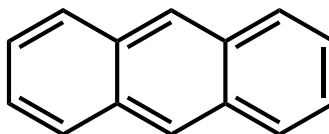
5,7-dimethyloct-6-ene-2-one



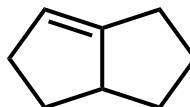
adamantane



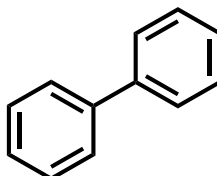
anthracene



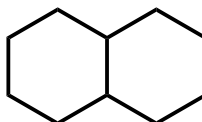
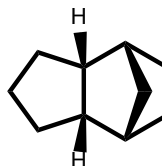
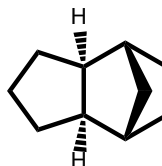
bicyclo[3.3.0]octene



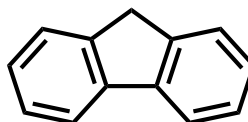
biphenyl



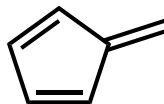
decalin

endo-tricyclo[5.2.1.0^{2,6}]decaneexo-tricyclo[5.2.1.0^{2,6}]decane

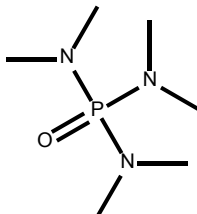
fluorene



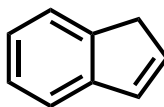
fulvene



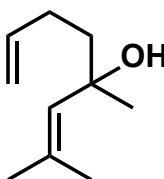
hexamethyl phosphoric triamide



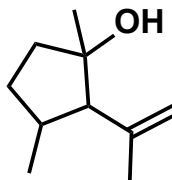
indene



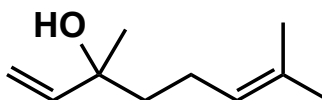
isolinalool



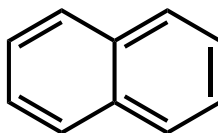
isoplinol



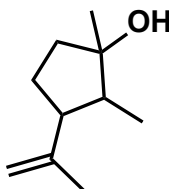
linalool



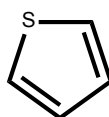
naphthalene

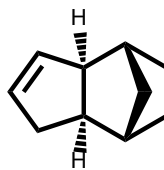
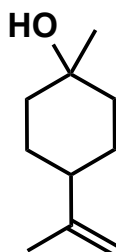


plinol



thiophene



tricyclo[5.2.1.0^{2,6}]dec-4-ene β -terpineol

Samenvatting

Gedetailleerde kinetische modellen zijn niet enkel een krachtig hulpmiddel voor het ontwikkelen en optimaliseren van chemische processen, maar helpen ook bij de opheldering van de onderliggende scheikunde die het proces beheerst. De handmatige opbouw van gedetailleerde kinetische modellen is steeds moeilijker te bewerkstelligen voor complexe scheikundige processen. Computers zijn daarom een interessant alternatief die de opbouw van kinetische modellen zouden vergemakkelijken en automatiseren, op voorwaarde dat geschikte methoden en algoritmes beschikbaar zijn voor de verschillende taken in de geautomatiseerde opbouw van kinetische modellen.

Het doel van dit onderzoek was methoden te ontwikkelen en te gebruiken voor de geautomatiseerde opbouw van kinetische modellen, die bovendien ook in staat zijn om de driedimensionale structuur van moleculen in rekening te brengen. Om dit doel te bereiken, werden technieken uit de chemoinformatica voor de eerste maal aangewend.

In het eerste deel van dit proefschrift werd het verband bestudeerd tussen de drie-dimensionale structuur van moleculen en hun reactiviteit voor de thermische omlegging van twee diastereomeren van 2-pinanol. Een gedetailleerd, kinetisch model werd opgebouwd dat goed overeenkomt met de experimenteel waargenomen trends en dat de waargenomen verschillen in reactiviteit en selectiviteit van beide diastereomeren verklaart. De reactiviteits- en selectiviteitsverschillen tussen de twee diastereomere reactanten werden toegeschreven aan een verschil in activeringsenergie van de koolstof-koolstof scissie-reactie die de cyclobutaanring van de reactanten opent met vorming van een 1,4-digesubstitueerde cyclohexanolbiradicaal. De hoge selectiviteit voor het reactiekanaal dat aanleiding geeft tot linalool ten opzichte van de parallelle weg naar isolinalool werd toegeschreven aan de grote verschillen in activeringsenergie van de cyclobutaan ring-openingsreacties die twee soorten biradicalen vormen: een 1,2-digesubstitueerd en een 1,4-digesubstitueerd cyclohexanolbiradicaal.

In een tweede deel van dit proefschrift werd de thermische ontbinding bestudeerd van Jet-Propellant 10 (JP-10), een geavanceerde vliegtuig- en raketbrandstof die voornamelijk bestaat uit een tricyclische $C_{10}H_{16}$ koolwaterstof. JP-10 pyrolyseexperimenten werden uitgevoerd op de proefbank-schaalreactoropstelling van het Laboratorium voor Chemische Technologie, met behulp van een combinatie van “online” twee-dimensionale gaschromatografie, “time-of-flight” massaspectrometrie, en een vlamioniseringsdetector die de identificatie en kwantificatie van meer dan zeventig producten mogelijk maakten. De experimenten droegen bij tot een beter begrip van de initiële ontbindingschemie van JP-10, niet in het minst door de kwantificering van moleculen zoals tricyclo[5.2.1.0^{2,6}]dec-4-een die niet eerder gerapporteerd waren in de literatuur. Er werd

ook aangetoond dat aromatische verbindingen zoals benzeen, toluen en styreen en belangrijke fractie vormen van de ontbindingsproducten van JP-10.

Een nieuw kinetisch model voor de JP-10 pyrolyse werd opgebouwd met behulp van reactiesnelheidsgebaseerde netwerkgenereringstechnieken, computationele-chemieberekeningen voor thermochemische eigenschappen, en een beperkt aantal *ab initio* berekeningen voor de primaire ontbindingsreacties van JP-10, berekend in samenwerking met de Green group aan het M.I.T.. Een belangrijk deel van het model omvat secundaire conversieroutes naar aromaten zoals benzeen, toluen, styreen en polyaromatische koolwaterstoffen wat kan gebruikt worden bij de beoordeling van de neiging tot afzettingen in brandstofrijke gebieden in endotherme brandstoftoepassingen. Een analyse van vormingssnelheden van de ontbinding van JP-10 toonde het belang aan van vier dominante primaire ontbindingskanalen. Het kinetisch model werd vervolgens gevalideerd over een breed scala van experimentele gegevens en een opmerkelijke overeenkomst tussen model en experiment werd vastgesteld, vooral gezien het feit dat geen van de modelparameters aangepast werd aan de experimentele gegevens.

Het derde deel van dit proefschrift bestaat uit een beschrijving van een nieuwe kinetisch-modelgenerator, genaamd GENESYS, waarin een aantal van de tekortkomingen van vorige programma's voor modelconstructie werden aangepakt. Een abstracte graafvoorstelling van moleculen en reacties werd gebruikt doorheen GENESYS. De grafenvoorstelling maakt de toepassing mogelijk van graaf-theoretische algoritmes en werd gebruikt in cruciale aspecten van reactienetwerkgenerering zoals het vaststellen van de uniciteit van moleculen en het zoeken naar sub-moleculaire patronen in moleculen. Daarenboven werden met relatief gemak

groepscontributiemethodes geïmplementeerd die dienen voor de berekening van thermochemische eigenschappen van moleculen en snelheidscoëfficiënten van elementaire reacties, en waarvan in het verleden aangetoond werd dat die nauwkeurige resultaten opleveren wanneer die gebaseerd zijn op accurate *ab initio* berekeningen. Groepentheorie vormt de basis voor de berekening van de symmetrie van moleculen en reacties in GENESYS. Het is gebleken dat het symmetriegetal van een groot aantal moleculen kan berekend worden op basis van de orde van de automorfismegroep van de overeenkomstige graaf en gecorrigeerd voor de asymmetrie die ontstaat door onderscheid te maken tussen identieke substituenten in de molecule. Deze waarneming werd geformaliseerd in een nieuw concept van label-stereoisomeren, dit wil zeggen stereoisomeren die ontstaan na het labelen van homomorfe liganden in de oorspronkelijke molecule zodat ze te onderscheiden zijn van elkaar. De vooropgestelde aanpak maakte de berekening mogelijk van zowel het rotationeel-symmetriegetal van moleculen evenals de reactiepaddegenerering van elementaire reacties door het algoritme toe te passen op zowel moleculen als op transitietoestandsstructuren. Het gebruik van een topologische weergave van moleculen en transitietoestandsstructuren vermeerde de noodzaak voor driedimensionale atomaire coördinaten, zodat het snel scannen van grote databases van moleculen om symmetrie te detecteren een haalbare optie wordt.

De functionaliteit van GENESYS werd geïllustreerd voor de opbouw van een kinetisch model voor de thermische ontbinding van diethylsulfide. Snelheidscoëfficiënten van elementaire reacties en thermochemische eigenschappen van moleculen werden berekend aan de hand van groepsadditieve methoden, waarvan de parameters enkel afkomstig waren van quantumchemische berekening uit de literatuur. De goede overeenkomst tussen de

modelvoorspellingen en experimenten was niet in het minst mogelijk door de beschikbaarheid van deskundige kennis over de desbetreffende chemie en de beschikbaarheid van gegevens voor thermochemie en kinetiek afkomstig van *ab initio* berekeningen uit literatuur.

Tenslotte werd aangetoond hoe kinetisch-modelgeneratoren kunnen worden uitgebreid om kinetische modellen waarin stereochemie belangrijk is, te construeren. Hiertoe werd een zogenaamde 2.5D weergave van moleculen geïntroduceerd, waarin de aanwezigheid van stereocentra en bijbehorende stereoconfiguratie van de stereocentra in rekening wordt gebracht. Een nieuw algoritme voor de detectie van sterische relaties tussen substituenten maakt de automatische toewijzing van snelheidscoëfficiënten van stereoselectieve reacties mogelijk. De uitbreiding van GENESYS voor stereochemie werd geïllustreerd voor de opbouw van kinetische modellen voor de thermische omlegging van 2-pinanol waarin rekening werd gehouden met de stereochemische kenmerken van de desbetreffende moleculen en reacties in de modellen.

Een pleiade aan computergestuurde methoden werden ontwikkeld en gebruikt om op een automatisch manier kinetische modellen te bouwen. Het gebruik van technieken uit de chemoinformatica in GENESYS maakte de constructie van kinetische modellen mogelijk, ongeacht de aard van de betrokken chemische elementen. Ook stereochemische aspecten van het kinetisch model werden in acht genomen werden, wat voor de eerste keer toegepast is voor de automatische constructie van kinetische modellen. Bijzondere aandacht werd besteed aan de relatie tussen de 3D-structuur van moleculen en de bijbehorende reactiviteit, waarna dit werd toegepast op de kinetische modellering van de thermische ontbinding van 2-pinanol en JP-10. Ondanks de huidige mogelijkheden van computers, is het gebleken dat de beschikbaarheid

van relevante scheikundige kennis en hoogwaardige scheikundige gegevens nog steeds essentiële ingrediënten vormen voor de opbouw van performante kinetische modellen.

Summary

Detailed kinetic models are not only a powerful means to develop and optimize chemical processes, but also provide guidance in the elucidation of the underlying chemistry that governs the process. The manual construction of detailed kinetic models is increasingly prohibitive for very complex chemical processes. Computers are therefore an interesting alternative that would facilitate and automate the construction of kinetic models provided viable methodologies and algorithms are available to accomplish the various tasks involved in automated kinetic model construction.

The objective of this research was to develop and use computer-aided methods for the automated construction of kinetic models that account for the three-dimensional structure of molecules. In order to achieve this goal, techniques from chemoinformatics were employed for the first time.

In the first part of this thesis, the relation between the three-dimensional structure of molecules and their reactivity was studied for the thermal rearrangement of the two diastereomers of 2-

pinanol. A detailed kinetic model was constructed that agrees well with the experimentally observed trends. The reactivity and selectivity differences between the two reactant diastereomers was attributed to a difference in activation energies of the carbon-carbon scission reaction that opens the cyclobutane ring of the reactant molecules and forms the 1,4-disubstituted cyclohexanol biradical. The high selectivity for the linalool reaction channel compared to the parallel pathway to isolinalool was attributed to the large differences in activation energies for the cyclobutane rupture reactions that form two types of biradicals, a 1,2-disubstituted and a 1,4-disubstituted cyclohexanol biradical.

In a second part of this thesis, the thermal decomposition of Jet-Propellant 10 (JP-10), an advanced jet and missile fuel that primarily consists of a tricyclic, strained $C_{10}H_{16}$ hydrocarbon was studied. JP-10 pyrolysis experiments were carried out on the bench-scale reactor setup of the Laboratory of Chemical Technology using a combination of on-line two-dimensional gas chromatography, time-of-flight mass spectrometry and a flame ionization detector for the identification and quantification of more than seventy products. The experiments contributed to a better understanding of the initial decomposition chemistry of JP-10 pyrolysis, not in the least by the quantification of species such as tricyclo[5.2.1.0^{2,6}]dec-4-ene that were not reported before in literature. It was also shown that aromatics such as benzene, toluene and styrene represent an important fraction of the decomposition products of JP-10.

A new, dedicated model for JP-10 pyrolysis was constructed using a combination of rate-based model generation techniques, computational chemistry calculations for thermochemical properties, and makes use of a limited number of high-level *ab initio* calculations for primary

decomposition reactions of JP-10, in collaboration with the Green group at M.I.T. A significant part of the model comprises secondary conversion routes to aromatics such as benzene, toluene, styrene and polyaromatic hydrocarbons (PAHs) and could thus be used in assessing the tendency for deposit formation in fuel rich zones of endothermic fuel applications. Rate-of-production analysis showed the importance of four dominant primary decomposition channels of JP-10. The model was subsequently validated by comparison to experimental data sets from various literature sources next to our own extensive data set. The simulation results showed remarkable agreement, especially given that none of the model parameters were fitted to the data.

The third part of this thesis consists of a description of a new kinetic model generator, called Genesys, in which some of the shortcomings of previous model generation programs were addressed. An abstract graph representation of molecules and reactions was used throughout Genesys. The graph representation enables the application of graph-theoretic algorithms and was used in crucial aspects of reaction network generation such as determination of species uniqueness, and substructure searching. Moreover, group contribution models were implemented for the estimation of thermochemical properties of molecules and rate coefficients of elementary reactions that have shown to yield accurate results when based on high level *ab initio* calculations, were implemented with relative ease. Group theory forms the basis for the calculation of symmetry of molecules and reactions in Genesys. It was shown that the symmetry number of a molecule can be calculated based on the automorphism group order of the associated graph, corrected for the asymmetry induced by distinguishing between identical substituents. This observation was formalized by the introduction of a novel concept of label-stereoisomers, i.e. stereoisomers that arise after labeling homomorphic ligands in the original molecule so that they

become distinguishable. The proposed approach allowed the calculation of the rotational symmetry number of molecules as well as the reaction path degeneracy of elementary reaction, by applying the same algorithm to molecules and transition state structures. The use of a topological representation of molecules and transition state structures avoided the need for three dimensional atomic coordinates so that a fast scanning of large database of molecules to detect symmetry becomes a viable option.

The functionality of Genesys was illustrated by the construction of a kinetic model for the thermal decomposition of diethylsulfide. Rate coefficients of elementary reactions and thermochemical properties of molecules were estimated through group additive methods, with parameters solely derived from quantum-chemical calculations from literature. The good agreement between model predictions and experiments was not in the least possible by the availability of expert knowledge on the chemistry and high level *ab initio* data for thermochemistry and kinetics.

Finally, it was shown how kinetic model generators can be extended to construct kinetic models in which stereochemistry is important. To that purpose a so-called 2.5D representation of molecules that accounts for the presence of stereocenters and that keeps track of the associated stereoconfiguration of the stereocenters was introduced. A novel algorithm for the detection of steric relations between substituents allows the automated assignment of rate coefficients to stereoselective reactions. The extension of Genesys for stereochemistry was illustrated by the construction of kinetic models for the thermal rearrangement of 2-pinanol that fully account for

stereochemical features in molecules and reactions. Good agreement with experimental data from literature was obtained.

A variety of computer-aided approaches were developed and used to automatically construct kinetic models. The use of techniques from chemoinformatics in Genesys allowed the construction of kinetic models regardless of the involved chemical elements and accounting for stereochemistry, and is the first kinetic model generator in its kind to do so. Special attention was paid to the relation between 3D structure of molecules and the associated reactivity and applied to the kinetic modeling of the thermal decomposition of 2-pinanol and of JP-10. Despite the current capabilities of computers, it was shown that the availability of relevant chemical knowledge and high-quality chemical data are still essential ingredients for the construction of performant kinetic models.

Glossary

Ab initio	Latin term for “from first principles”. It refers to the fact that the results are obtained by applying the established laws of nature without assumptions, special models or experimental input. Ab initio methods determine the energy of a molecule or transition state by solving the Schrödinger equation.
Arrhenius activation energy	The coefficient E_a catching the temperature dependency of the rate constant $k = A \exp(-E_a/RT)$ with A the temperature independent pre-exponential factor.
Arrhenius pre-exponential factor	See Arrhenius activation energy.
Atom permutation	A bijection from the set of atoms in a molecule onto itself.
Automorphism group	The automorphism group $\text{Aut}(G)$ of a weighed graph

	defined by its vertices V , edges E , and respective weights w and λ , is a permutation group in which all the elements g fulfill two conditions: 1) For every vertices u and v : $g(u)$ and $g(v)$ are adjacent if and only if u is adjacent to v and 2) for edge e : $\lambda(g(e)) = \lambda(e)$, i.e. The weights of the edges in the permuted graph should be identical to the weights of the same edges the original graph.
Automorphism permutation	An atom permutation that belongs to the automorphism group of the associated graph of the molecule.
Bond dissociation energy	The bond dissociation energy (BDE) is the enthalpy change when cleaving a bond by homolysis.
Clique	A graph with every two vertices connected by an edge.
Configuration vector	One row of the stereo parity matrix, representing a possible combination of parities for each stereocenter.
Diastereomers	Diastereomers are stereoisomers that are not enantiomers.
Elementary reaction	A chemical reaction in which one or more species react to products in a single step and with a single transition state.
Enantiomer	An enantiomer is one of two stereoisomers that are mirror images of each other that are non-superposable.
Enthalpy	The enthalpy H is a thermodynamic quantity and is calculated from the internal energy U as $H = U + PV$, with P the pressure and V the volume of the system.

Entropy	The entropy S is a thermodynamic property that is related to the disorder of the system. A system with a larger number of states that can be occupied, will therefore have a higher entropy.
Graph invariant	A number or sequence of numbers, computed from the graph topology that is not dependent on the graph labeling.
Graph isomorphism	The graph-theoretic problem of determining whether there is a one-to-one mapping between the vertices of one graph to the vertices of a second graph.
Group additivity method	A group additivity method is a technique that allows to predict properties from molecular structures. For example, within Benson's group additivity method a property can be written as a sum of contributions arising from its constituent groups.
Group contribution method	See group additivity method.
Homomorphic	Identical, when considered in isolation
Label-stereocenter	A label-stereocenter is a center that becomes a stereocenter after distinguishing between its homomorphic ligands.
Label-stereoisomer enumeration	Counting the number of possible label-stereoisomers.
Label-stereoisomers	Stereoisomers that originate by labeling homomorphic ligands in a molecule.
Monoterpenoids	Oligomers derived from two isoprene subunits, containing functional groups that may contain heteroelements such as

	oxygen
Order of the automorphism group	The number of elements (permutations) that are part of the automorphism group.
Para-stereocenter	Stereocenter that does not necessarily lead to 2 stereoisomers. A necessary condition for para stereocenters is topological symmetry in the molecule.
Parsing	The process of analyzing a string of symbols in a language, converting the string into a more manipulable representation.
Partition function	The partition function encodes the statistical properties of a system. For a canonical ensemble, the partition function is the Boltzmann sum over the different microstates, the system can occupy.
Prostereogenic element	An element (center, axis, plane) that becomes a stereocenter if exactly two homomorphic ligands are made different.
Pyrolysis	The uncatalysed decomposition of organic components resulting from exposure to high temperature, in the absence of molecular oxygen.
Reaction family	A class of reactions that are characterized by the same pattern of electron rearrangement steps.
Reaction path degeneracy	The number of energetically equivalent paths that reactants can follow to be converted into products.
Tolerance for rate-based kinetic model enlargement (ε)	The tolerance is a user-defined constant that is multiplied by

	the characteristic rate R_{char} of the kinetic model. $\varepsilon \cdot R_{\text{char}}$ is used to determine whether a new species is added to the enlarged model based on comparison of the rate of production of that species and $\varepsilon \cdot R_{\text{char}}$.
Signed permutation	The row-product of permutation matrix with its corresponding stereo index vector.
Single-event microkinetic model	A kinetic model that consists of elementary reactions and accounts for all energetically equivalent reaction paths, i.e. Single-events, to determine each reaction rate.
Single-events, number of	Cf. Reaction path degeneracy.
Single-event pre-exponential factor	The pre-exponential factor excluding the number of single-events of the reaction.
Standard enthalpy of formation	The standard enthalpy of formation is the change of enthalpy that accompanies the formation of 1 mole of a substance in its standard state from its constituent elements in their standard state.
Steam cracking	A petrochemical process in which saturated hydrocarbons are converted into small unsaturated hydrocarbons by exposure to high temperature in the presence of steam.
Stereocenter	A locus of stereoisomerism in a molecule such that interchange of two ligands attached to an atom in such a molecule leads to a stereoisomer.

Stereoconfiguration	The possible arrangements of ligands around a stereocenter.
Stereoindex vector	A binary vector (-1/+1) that shows the behavior of the stereocenter parities under a given automorphism permutation.
Stereoisomer generation	The process of creating all possible stereoisomers given a structure with unspecified stereocenters.
Stereoparity	Binary value (-1 / +1) of the configuration of stereocenters
Stereoselectivity	Preferential formation of one or the other of two (or more) stereoisomers from a single molecule with a prostereogenic element.
Stereospecificity	The difference in reaction rates of two stereochemically different molecules, i.e. diastereomers or enantiomers.
Steric relation	The three-dimensional arrangement of substituents with respect to each other.
Sub-graph isomorphism	The graph-theoretic problem of determining whether a graph G contains a subgraph that is isomorphic to another graph H.
Topological symmetry	Symmetry induced solely by connectivity patterns in the associated graph of the structure.
Topology	The topology of a graph refers to the connectivity of the vertices, and makes abstraction of the specific spatial positioning of the vertices.
Transition state	The transition state of an elementary reaction is that set of states in which an assembly of atoms, when randomly placed

	there, would have equal probability to form the reactants or products of that elementary reaction.
Transition state structure	Saddle point on the potential energy surface along the minimum energy path. A normal mode analysis on the TS structure yields one imaginary frequency.
Transition state theory	Theory that allows calculating rate coefficients assuming quasiequilibrium between the reactant and transition state.
Tunneling coefficient	The tunneling coefficient is a correction factor to the rate coefficient accounting for quantum effects, mainly tunneling of particles through the reaction barrier.

Chapter 1: Introduction

The prediction ability has always been a central theme in chemical engineering. Not only does it confer an alternative to ad-hoc and costly experiments in the field of chemical process optimization, it also embodies an accelerator for innovation and process development. The 20th century has proven to be a fruitful century in the prediction of physical properties through chemical thermodynamics, heat transfer and mass transport phenomena giving birth to a generation of chemical process simulation software. The latter presented a milestone in the advancement of chemical engineering science and is now an essential part of a chemical engineer's toolbox.

In contrast to the significant advances made in the aforementioned fields of chemical engineering, the prediction and modeling of reactive processes remains one of the major hurdles to be taken on in the 21st century. The capability to reliably predict the behavior of reactive chemical processes would allow rational *a priori* design of chemical reaction processes and provide a more adequate response to societal challenges [1]. Kinetic models are one of the cornerstones of the field of chemical reaction engineering. They provide quantitative insights in how species react and have the capacity to reliably predict the behavior of chemical reactions over a broad range of operating conditions of the chemical process. In recent years, it has become possible to accurately predict the behavior of some rather complex gas phase chemical processes such as the chemistry occurring in steam cracking and pyrolysis of hydrocarbons [2], or the combustion of fuels in

engines [3]. Even in biological reactive processes such as cells, very detailed, quantitative models for the analysis of metabolic pathways have been applied [4, 5].

Kinetic models can be constructed following two contrasting philosophies [6]. The inductive approach departs from a kinetic model with a minimal number of species and reactions that allows explaining the observed trends in experimental data. As new experiments become available, and new experimental trends are observed in the experimental data that cannot be explained by the previous model, the latter is expanded with the aim of describing the new experimental data. As such, a feedback loop occurs between newly gained experimental insights and the gradual increase in complexity of the kinetic model. The opposite philosophy, also called the deductive approach, departs from a kinetic model that reflects the underlying chemistry as much as possible by incorporating all relevant elementary pathways and molecules. Deductive kinetic models are not constructed with the aim to explain a specific experimental data set, but rather use experimental data to validate the model. The deductive approach is used throughout this thesis. The downside of the high level of detail of deductive models is that the complexity and size of the models also steadily increase, so that the models eventually become too difficult to be manually constructed, cf. Figure 1-1 [7].

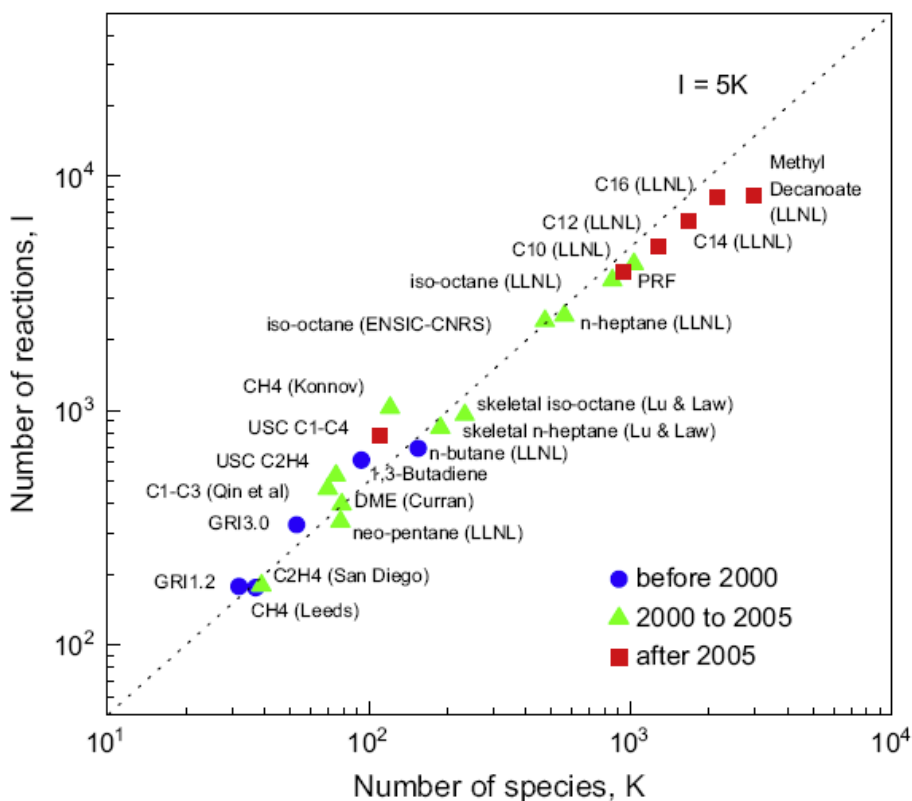


Figure 1-1: Size of selected detailed and skeletal networks for hydrocarbon fuels, together with the approximate years when the networks were compiled. From Lu and Law (2009) [7].

With the advent of affordable computer power in the last decades, computers have more and more been used in a variety of domains of science and engineering. In the fields of chemistry and biology, computers complemented bench-scale experiments in providing efficient methods to store, retrieve, organize and analyze data leading to the well-established fields of chemo- and bioinformatics [8]. Especially bio-informatics has bloomed and lead to ground-breaking successes such as protein structure elucidation [9], sequence alignment [10] and *in silico* drug design [11].

Computers represent a very powerful resource for kinetic modeling purposes as well, since they can be programmed to automate the construction of kinetic models. Despite the potential use of computers, it remains a partially untapped resource for kinetic modeling so far. Many kinetic models are still built by hand, despite the tediousness and the error-proneness of the effort. One of the reasons for this may lie in the fact that until now there have not been game-changing ways

to deploy computers in generating networks that significantly surpasses the performance of hand-built ones. Computers are often employed in only one segment of the process of kinetic model construction or simply to speed-up the procedure, without providing truly innovative capabilities, as is shown in subsequent sections. Moreover, in contrast to bio-informatics, large-scale publicly available chemical data is not easily available for chemical reaction engineering applications. Experimental techniques are constrained to measure only a fraction of the characteristics of the model, accompanied by significant error bars, and regularly in the form of a variable such as an ignition delay that only indirectly reflects the chemistry. More recently, the use of high level quantum-chemical calculations has provided a boost for kinetic modeling since, at least in principle, every parameter of the kinetic model can be derived. However, *ab initio* calculations still require a large amount of expert know-how and computational resources. Moreover, in contrast to the major providers of life science data such as the National Center for Biotechnology Information and European Bioinformatics Institute [12, 13], joint efforts from academia or industry to standardize and centralize on a large scale chemical data for kinetic modeling purposes remain rather limited [14].

This thesis focuses on kinetic modeling of gas phase thermal decomposition (pyrolysis) processes that typically proceed through free-radical mechanisms. The complexity of these models arises due to the generation and propagation of chain carriers, leading to models that require hundreds of species and thousands of reactions. Due to the absence of a catalyst, or a surrounding medium such as a solvent, the thermochemistry of molecules in gas phase pyrolysis processes is often described by treating them as ideal gases, which was also done in this thesis.

Since this thesis is intended to contribute to the general research objective to develop techniques for computer-aided kinetic modeling of chemical processes, the remainder of this chapter consists of an overview of the state-of-art of the automatic generation of kinetic models and an overview of some advances of chemoinformatics that are essential in the development of network generation tools. Finally, the objectives of this thesis are formulated.

1.1 Automated kinetic model generation

A kinetic model consists of a reaction network, rate coefficients, and optionally thermochemical parameters. The reaction network defines the relationships among reactants, intermediates and products. To simulate the product distribution as a function of operating conditions in a reactor every reaction in the reaction network needs a rate coefficient that is used to calculate the reaction rate based on the concentration of the reactants. If an energy equation is to be simultaneously solved next to the continuity equations, or if the rate coefficients of reversible reactions are calculated by accounting for thermodynamic consistency, then each molecule and intermediate considered in the reaction network requires also the knowledge of the species' thermochemical properties. Computer-aided methods for the construction of detailed kinetic models therefore need to provide solutions for two sub-problems: 1) the automatic generation of the reaction network and 2) the automated calculation of thermochemical data and reaction rate coefficients. These topics are discussed in the following sections and have been extensively reviewed, e.g. in refs. [1, 15, 16].

1.1.1 Reaction network generation

Reaction network generation consists of a recursive procedure to construct all possible products from an initial set of reactants. Three principle features are required for automatic reaction network generation:

1. a computer-manipulable representation of the chemical structure of molecules,
2. the identification and execution of reaction possibilities,
3. a rationale that controls the size of the reaction network and prevents the network generator from endless continuation.

These three aspects are discussed in the following sections.

1.1.1.1 Molecule representation


A molecule representation refers to the computer representation of constituent atoms and their chemical environment in the molecule. The level of detail of molecule representations is often a function of the intended application. Molecules may simply be treated as a collection of

functional groups, for which the number of occurrences of each functional group is stored [17, 18], cf. Figure 1-2. This molecule representation is useful when the exact connectivity of atoms is of less importance, and can be used to model complex, large molecules such as in petroleum vacuum residua [19].

		A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	IH	AA	NS	RS	AN	NN	RN	NO	RO	KO
2,3,5 Trimethyl hexane		0	0	0	0	0	0	0	0	0	9	3	0	1	0	0	0	0	0	0	0	0	0
Benzene		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Naphthalene		1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Phenanthrene		1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pyrene		1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Carbazol		2	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0
Benzofuran		1	0	0	0	0	0	1	0	0	0	0	0	-1	0	0	0	0	0	0	1	0	0

Figure 1-2: Vector representation that reports the number of functional groups in the molecule. From Jaffe et al. [19].

When the exact atom connectivity is important, a graph representation of the molecule is very useful. Bond-electron (BE, binary, boolean) matrices are often chosen as the data structure to store the connectivity of the graph, and often referred to as the Dungundji-Ugi (DU) formalism [20]. In BE-matrices, off-diagonal elements (i,j) refer to the presence or absence of a bond between atoms i and j , while diagonal elements (i,i) correspond to the number of unpaired electrons for that atom. For example, Clymans and Froment [21] used BE-matrices to represent the connectivity between carbon atoms for normal and branched paraffins, cf. Figure 1-3.



	1	2	3	4	5	6	7	8	9	10
1	0	1	0	0	0	0	0	0	0	0
2	1	0	1	0	0	0	0	0	0	0
3	0	1	0	1	0	0	0	0	0	0
4	0	0	1	0	1	0	0	0	0	0
5	0	0	0	1	0	1	0	0	0	0
6	0	0	0	0	1	0	1	0	0	0
7	0	0	0	0	0	1	0	1	0	0
8	0	0	0	0	0	0	1	0	1	0
9	0	0	0	0	0	0	0	1	0	1
10	0	0	0	0	0	0	0	0	1	0

Figure 1-3: A BE-matrix representation of n-decane as used by Clymans and Froment [21].

Supplementary arrays may augment the matrix, to account for heteroelements, electron deficiencies, etc..., and illustrate the variations of the DU-formalism that emerged in applications such as hydrocarbon oxidation and pyrolysis [22-25] and catalytic transformation of hydrocarbons [26, 27]. The construction of a BE-matrix requires the assignment of an index to each of the carbon atoms in the molecule that corresponds to the row and column number in the matrix. Since many equally valid atom numbering schemes can be envisioned, multiple distinct matrices can be constructed that represent the same molecule. This is problematic for crucial operations in the reaction network procedure. For example, the uniqueness of one molecule cannot be accomplished by comparing the identity of the corresponding matrices. Determining the uniqueness of a molecule is a problem that has been addressed in the field of chemoinformatics by a class of algorithms, called canonicalization, algorithms that intend to uniquely number the atoms in the molecule. Canonicalization plays a major role throughout this thesis, and hence is further discussed in Section 1.2.1 of this chapter.

BE-matrices were very popular molecule representations in an era of reaction network generation in which multi-dimensional arrays were the main data structures of programming languages [25, 28, 29]. With the advent of object-oriented languages, more complex and flexible data structures, also called objects, became available and allowed new organizations of the chemical information in the molecule representation. Object-oriented languages provided frameworks for more advanced species representations covering an increasingly large space of chemical elements and

possible structures. Moreover, species representations were no longer created for a specific application, but rather became an end-product as such. The popularity of freely available toolkits [30-32] in the chemical community that offer “off-the-shelf” species representations, illustrates the need for this product.

The field of chemoinformatics played a pioneering role in the advancement of species representations, not in the least to incorporate information on the three-dimensional arrangement of atoms in the molecule. Since the relation between chemical reactivity and 3D structure was further explored in this thesis, the representation of 3D structure of molecules is further discussed in Section 2.3 of this chapter.

1.1.1.2 Reaction families

Reaction families are the direct extension of the notion that sub-molecular patterns and more specifically functional groups inside a chemical species govern the reactivity of a molecule rather than the molecule in its entirety. Without reaction families, every chemical bond and every atom should *a priori* be eligible to undergo a chemical reaction leading to a combinatorial explosion of species and reactions and a “non-sense” reaction network. The predefinition of reaction families, also called “formal” network generation [33], results in algorithms that generate reactions whose structure is directly defined by one of the reaction family templates. The nature of the created molecules in the network is only indirectly constrained by the definition of reaction families.

A critical part of the definition of a given reaction family is the creation of products from reactants. The way products are derived from reactants is tightly linked with the employed molecule representation. Within the DU-formalism, chemical reactions are described by transformations of the BE-matrices. NETGEN [34] added a so-called reaction matrix (R-matrix) to the BE-matrix, resulting in a new product matrix. For example, the homolytic C-C scission reaction in n-decane between the 6th and 7th carbon atom is accomplished by adding an R-matrix with elements equal to zero, except for the 6th and 7th columns, cf. Figure 1-4. The resulting product matrix is then split in two pieces, with the sub-matrices corresponding to the formed n-hexyl and n-butyl radical.

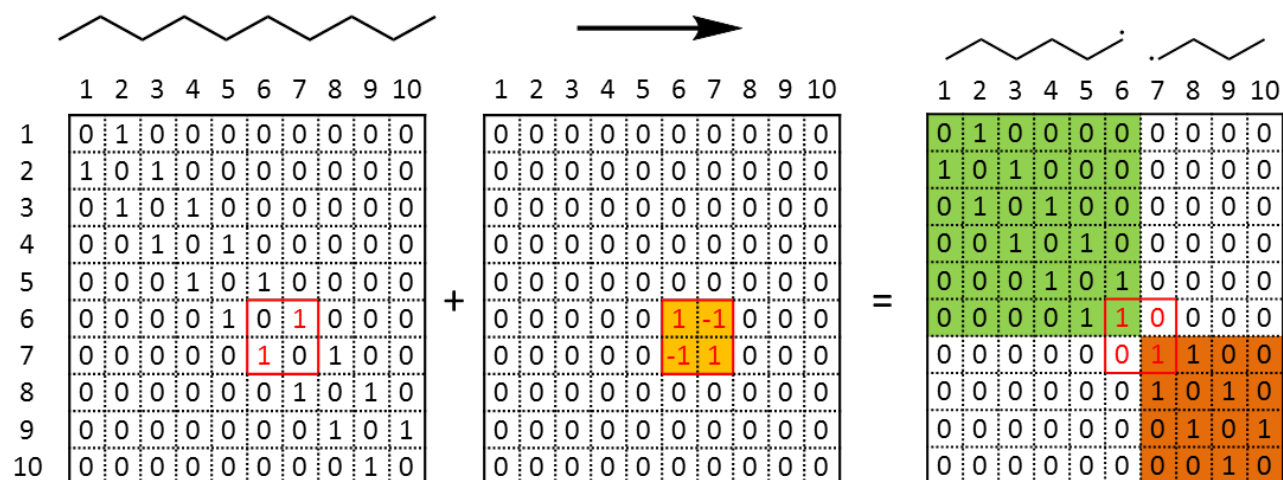


Figure 1-4: Reaction execution for the homolytic C-C scission of the 6th and 7th carbon atom by adding a reaction matrix to the reactant matrix of n-decane, used in NETGEN [34].

Matrices can also be manipulated using techniques from linear algebra to uncover sub-molecular characteristics of molecules. This was done in the program PRIM [21, 35] for the generation of reaction networks for steam cracking. For example, the sum of the non-diagonal elements in one row indicates whether a carbon atom is primary, secondary, tertiary or quaternary. Taking the square of the BE-matrix reveals the β -position of atoms, cf. Figure 1-5. This relation can subsequently be used in the β -scission reaction family of radicals to determine the location of the free electron and the double bond in the product structures.

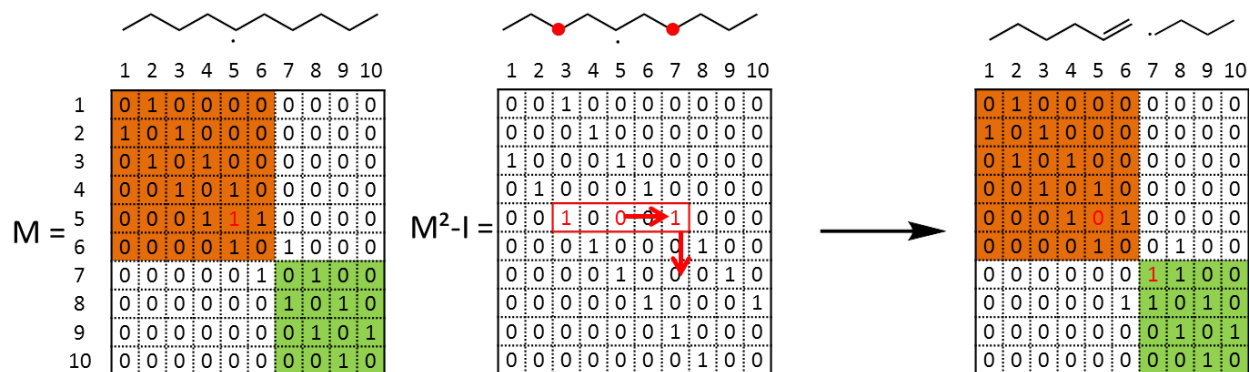
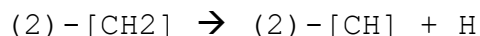


Figure 1-5: Multiplication of the matrix of 5-decyl radical by itself leads to the identification of beta-positions between atoms. They are used to locate the free electron in the product matrix of 1-butyl and the location of the double bond in 1-hexene in the program PRIM [21, 35].

Although the use of matrices for the execution of reactions worked well within the intended scope of applications, the procedure is not straightforward for more complex reaction families. For example, for pericyclic reactions with multiple electron rearrangement steps, the derivation of the product structure by matrix operations is far from trivial. Moreover, the methodology to retrieve the structure of the product molecule(s) is difficult to separate from the program itself, implying that users are forced to alter the program when a new reaction family is considered. To simplify and standardize defining reaction families, more recent approaches used techniques based on graph theory to execute reactions. In the reaction network generation program COMGEN [36], the atoms that undergo the bond and electron rearrangement processes are described using a domain specific language. For example, the C-H bond scission reaction of secondary carbon atoms is described by the following sequence of characters in COMGEN:



The arrow symbol divides the expression in two pieces; the left part referring to the specifications of the reactant $((2) - [\text{CH}_2])$, the right part $((2) - [\text{CH}] + \text{H})$ referring to the obtained product structures. Element symbols such as C and H in combination with integer numbers give information on how many atoms of each kind are expected. For example, CH₂ refers to a carbon atom surrounded by two hydrogen atoms. The square brackets “[]” denote the “free valence” locations of the carbon, allowing various, undefined combinations of substituents. For example, [CH] refers to a carbon atom bonded to a hydrogen atom that can be surrounded by two non-specified substituents. The reactant pattern is translated into a graph representation and used in a sub-graph search algorithm [37] that matches the query substructure within the graph structure of the molecule. Language interfaces were illustrated in several network generation programs [38-41], and further elaborated in the program RING [42]. The reaction transformation in the latter was described by a recipe, a list of the necessary bond and electron rearrangement steps that transform the reactant graph(s) into products. For example, Figure 1-6 shows the recipe for the adsorption of ketones on acidic sites that consists of a reactant pattern section and a transformation operations section. In the reactant pattern section, the reactant atom pattern is described and reactive center atoms are labeled. The next section defines the transformations using the reactive center labels from the reactant pattern section.

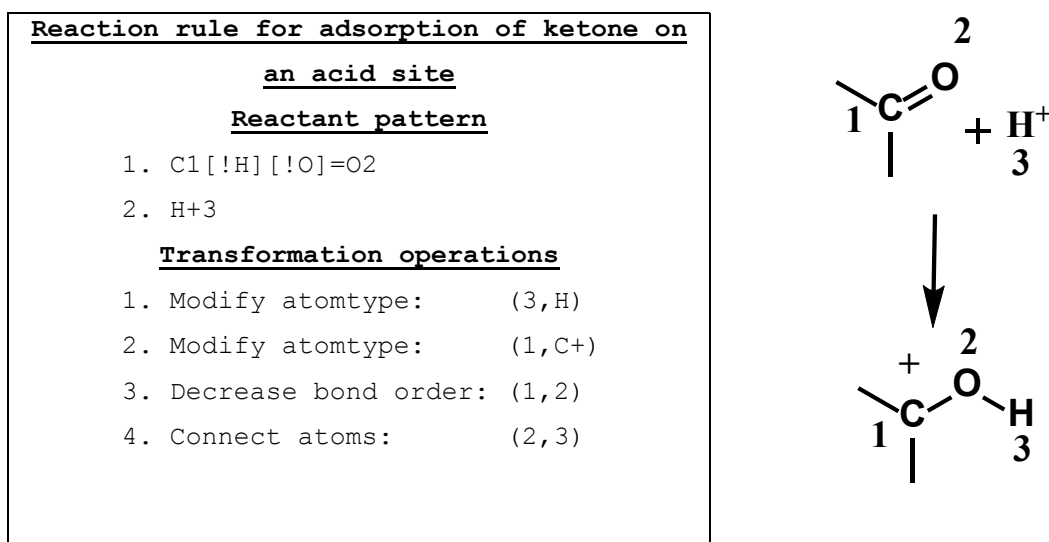


Figure 1-6: Example of reaction rule in RING [42-44]. Reactive atoms are numbered 1-3.

Similar recipe-like approaches were used in other network generation programs, e.g. RMG [45] because of the simplicity and flexibility in defining reaction families. Only a handful elementary transformations, e.g. “Decrease bond order”, “Connect atoms”, “Modify atomtype”, etc... of the reactive centers allow the definition of any kind of reaction family. An extension of the recipe approach is the use of language expressions. For example, RING has developed a specific language interface [44], based on the use of SILVER, an attribute grammar specification language [46]. The English-like language expressions are subsequently translated into C++ program that accesses the actual network generation library of RING. Since the syntax of the language closely resembles chemistry parlance, it facilitates the definition of reaction families for non-expert users of RING. Furthermore, it also enables catching chemistry-specific inconsistencies and allows improving the speed of execution through domain-specific optimization of the input instructions. The reader is referred to Section 5.3.1 of Chapter 5 for more information on the implementation of reaction families used in this thesis.

1.1.1.3 Network size control

A side-effect of the tediousness of the manual construction of kinetic models was the careful selection of the reactions and species that are incorporated in a reaction network by the expert. This invariably lead to small models of a manageable size for which the corresponding set of reactor model equations could easily be solved numerically. With the advent of computer programs that automatically generate models, the coupling between tediousness and expert

scrutiny is no longer present. This is regrettable, since larger models do not necessarily imply superior model performances. Moreover, the large number of species present in the network are not only nonlinearly coupled, but frequently have vastly different time scales, leading to simulations that are exceedingly impractical to solve.

The size of automatically generated kinetic models can be controlled in two ways, either by reducing the model *a posteriori*, or by *a priori* control of the model generation process. Several methodologies were developed for network reduction [7]. Networks can be reduced to so-called skeletal networks by removing kinetically unimportant species and reactions from the detailed networks, e.g. by means of sensitivity analysis [47-49]. Specific lumping strategies [17, 18, 50, 51] can also be a powerful model reduction technique, especially for species that have similar reactivity characteristics. Furthermore, the analysis of time scales can identify short living species or fast chemical reactions, leading to approximations such as the quasi-steady state approximation (QSSA), or the partial equilibrium approximation. For example, the network generation program PRIM [35] applies the QSSA to all μ -radicals [52], i.e. large radicals for which bimolecular reactions can be neglected, that appear in the primary decomposition schemes of hydrocarbon molecules. Concentrations of these μ -radicals are subsequently derived by solving the set of algebraic equations that correspond to the QSSA, and substituted into the smaller, reduced system of ordinary differential equations corresponding to the species for which the QSSA is not applied.

The *a priori* control of the network generation procedure refers to the measures that are taken to inhibit the creation of kinetically irrelevant species and reaction pathways. Curran et al. [53, 54] divided the elementary reaction families relevant for the low-temperature oxidation of n-heptane and iso-octane, into 25 reaction families. The same principle was used in the network generator REACTION [55]. Furthermore, the creation of new molecules in the network generation algorithm can be terminated by counting the number of atoms in the molecule [34, 56], introducing a so-called carbon-count termination criterion. Alternatively, it can be determined by the rank of the molecule, i.e. the minimum number of steps required for that molecule to form from any of the initial reactants [57]. The definition of a set of possible products can also avoid the combinatorial explosion of reactions in the network [58, 59]. Constraints, i.e. restrictions on

the presence of molecular and sub-molecular characteristics in a molecule as is done in the program RING [42, 43], is another option to limit the reaction network size. For example, a global constraint can prevent the creation of consecutive double bonds in the molecules that are created by the network generation procedure. In general, the above examples can be seen as derivations of rules that reflect the chemical knowledge of the user expert on the process of interest. Therefore, this approach is referred to as rule-based network generation [16]. However, the performance of the model also heavily depends on the knowledge of the chemical process, which can often be insufficient due to the lack of experimental or theoretical investigations on the subject. Rate-based algorithms for the generation of kinetic models tried to address this issue [60]. In rate-based network generation methodologies, species are added to the model when the rate of production of candidate species is higher than a threshold value, cf. Figure 1-7.

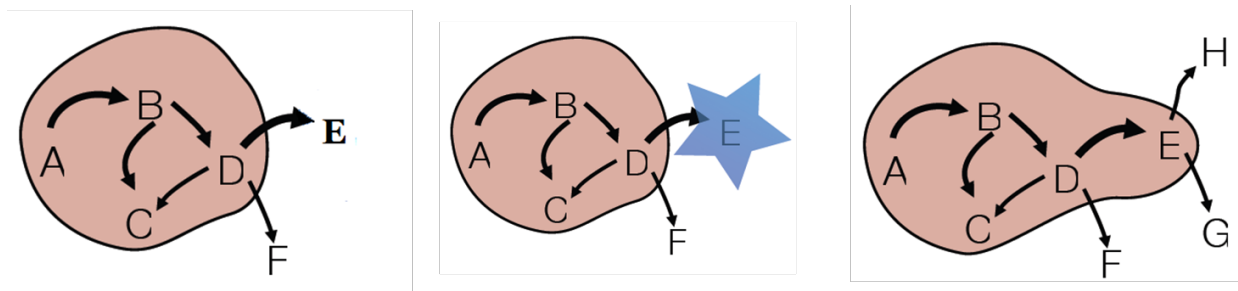


Figure 1-7: Rate based model enlargement by evaluation of the rates of formation of species (i.e. E, and F) that are not yet part of the model in comparison with a characteristic reaction rate derived from the species that are part of the model (i.e. A,B,C, and D) [61].

The minimum rate of production R_{min} is determined by a user-defined tolerance ε , cf. Eq. 1-1, and a characteristic rate R_{char} of the kinetic model, calculated based on the rates of production of species R_j that are already part of the kinetic model, cf. Eq. 1-2.

$$R_{min} \geq \varepsilon \cdot R_{char} \quad \text{Eq. 1-1}$$

$$R_{char} = \sqrt{\sum_j R_j^2} \quad \text{Eq. 1-2}$$

XMG (Exxon-Mobil Mechanism Generator) [62] was the first network generation program that applied the principle of rate-based enlargement, and formed the basis for the more recent Reaction Mechanism Generator (RMG) [45]. RMG integrates the system of ordinary differential equations (ODEs) for the balances of the considered components in order to evaluate the rates of production at a constant temperature and pressure that is representative to the application.

Hence, the dependency on a thorough understanding of the underlying chemistry for rule-based generation approaches is partially addressed by rate-based approaches. Instead, the availability and quality of thermochemical and kinetic data is of more importance, since these data are used to evaluate reaction rates. Inaccurate reaction rates can result in incomplete or skewed models that are unable to grasp critical features of the underlying chemistry. Despite the evaluation of reaction rates as the main criterion to include species in the model, it should be noted that rate-based approaches still require defining the reaction families that are used to generate the reactions.

Rate-based approaches represent more technical challenges for the management of computer memory and processors than rule-based approaches. In rate-based approaches, each iteration of the network generation procedure generates a large number of species that are not yet part of the model (“edge species”), and that eventually may never be included in the model. The number of edge species is generally one to two orders of magnitude larger than the number of species in the model and can represent memory problems for large reaction networks. Rate-based approaches also require solving a system of ODEs each time a new species is added to the model. As a result, a large fraction of the computer processor time in rate-based approaches is dedicated to solving the system of ODEs, which can be prohibitive for very large or stiff models.

In this thesis, a network generation program is developed based on a rule-based approach. This choice follows the same spirit of previous network generation programs developed at the Laboratory of Chemical Technology. This paradigm is chosen due to the reliability of the generated models. They are the straight-forward and consistent product of the chemical knowledge of the user. In other words, when sufficient chemical knowledge expertise is available on a specific chemical process, then rule-based approaches are the best guarantee for the construction of adequate models of this process.

Since access to accurate thermochemistry and kinetics is important for all network generation tools, it is discussed more in detail in the next sections.

1.1.2 Thermochemistry

Complex networks can contain up to hundreds of species, whose thermochemical properties need to be known. The experimental determination of thermochemical properties is, in many cases, not *a priori* trivial, demanding, or in some cases even impossible, e.g. for unstable species such as radicals. During the last decades, the use of *ab initio* quantum mechanics has become a viable alternative to experiments to obtain thermochemical properties. Popular methods such as CBS-QB3 have shown to be a good trade-off between accuracy and computational expense for many reactions, with errors for heats of formation not exceeding the chemical accuracy ($< 4 \text{ kJ mol}^{-1}$) for sufficiently small molecules [63]. The corpus of high-level *ab initio* calculations is constantly expanding, especially with the arrival of high performance computing capacity, which allows the use of increasingly more accurate methods. *Ab initio* methods, however, often have bad scaling characteristics (e.g. CBS-QB3 $\sim N(7)$), leading to prohibitive calculations for too large molecules. Also, errors of *ab initio* methods tend to increase with size of the molecule unless specific corrections are incorporated, that are determined by fitting for a set of molecules. Finally, obtaining accurate results through *ab initio* methods still requires significant expert involvement, and know-how, currently making automated, mass-scale calculations still largely infeasible. One may thus wonder whether obtaining *ab initio* values for every single thermochemical and kinetic parameter in the model is a good investment of human and computational resources.

Group contribution, or group additivity, methods complement *ab initio* methods by addressing some of the aforementioned shortcomings [64, 65]. Group contribution methods are based on the empirical observation that many physical properties can be estimated through contributions of various constituent parts of the molecule. For the prediction of enthalpies of formation, entropies and heat capacities, the group additivity method devised by Benson and coworkers [65, 66] is arguably the most renowned. Within this additivity scheme a group is defined as a polyvalent atom with all of its neighboring ligands. Groups are labeled $X-(A)_i(B)_j(C)_k(D)_l$ with X the central atom having i A ligands, j B ligands, etc. For each group, the numerical contribution to a specific physical property, e.g. the standard enthalpy of formation at 298K, is tabulated. Besides the groups, additional corrections need to be introduced in the additivity scheme to account for interactions between parts of the molecule whose proximity is not explicitly expressed by a bond.

These corrections include non-nearest neighbor interactions (NNI), e.g. cis-corrections in alkenes and gauche-corrections in alkanes, ring strain corrections (RSC) in cyclic molecules and resonance corrections (RES) that account for the stabilization effect of electron interactions in a molecule. The physical property, e.g. the standard enthalpy of formation at 298K, can then be calculated as:

$$\Delta_f H^\circ = \sum_{j=1}^n GAV(X_j) + \sum_{j=1}^m NNI_j + \sum_{j=1}^q RSC_j + \sum_{j=1}^p RES_j \quad \text{Eq. 1-3}$$

For radicals, the group additivity method of Benson requires radical-centered groups and groups adjacent to radical sites, leading to multiple groups that account for the effect of the radical. To address this, Lay et al. [67] developed the hydrogen-atom bond increment (HBI) method that combined the influence of the radical site on the thermochemical properties in a single “hydrogen-bond increment”. Within the HBI method, properties of the radical R^\bullet , are obtained by adding a hydrogen bond increment to the properties of the corresponding parent molecule RH, in which the unpaired electron is replaced by a hydrogen bond. The reader is referred to Section 5.3.2 of Chapter 5 for more information on the methodology for the calculation of thermochemical properties of species used in this thesis.

Group contribution methods are very powerful since they provide data for molecules whose thermochemical properties are *a priori* unavailable. They are easy to use, and are straightforward to implement under the condition that the sub-molecular groups in a molecule can be algorithmically identified. This resulted in several programs that implement Benson’s method [45, 68-70]. The numerical values of group contributions are determined as parameters originating from the regression of a training set of molecules, and can therefore be combined with high level *ab initio* calculations. This was successfully done for hydrocarbons [71, 72], but also for oxygen [73, 74], sulfur [75], phosphorous [76] and nitrogen [77] containing components. Despite its capabilities, group contribution methods should still be applied with care. Group contribution methods are limited to the components for which numerical values exist for the groups that are present in the molecule [64], and therefore can be seen more as an instrument of interpolation rather than extrapolation. In principle, the accuracy of the use of group additive values is only ensured in combination with groups that are derived from the same statistical regression. When molecule properties are estimated by groups that originate from different sources, upper limits for

the introduced error can no longer be established based on the reported maximum deviations of the individual groups. Although group contribution methods dramatically reduce the number of parameters needed to determine thermochemistry of molecules, the number of required groups still scales in a combinatorial way with respect to the number of atom types. The combination of even a limited set of atom types leads to a large number of possible groups, many of them have not been determined yet. Furthermore, for a subspace of molecules, the basic premise of group contribution methods to predict molecular properties based on additive schemes of sub-molecular moieties may no longer hold. For example, for polycyclic strained molecules, the ring strain may only be accurately assessed if the entire structure is accounted for. Applying ring corrections based on single ring structures in polycyclic structures may not be adequate. To cope with these limitations of group contribution methods for thermochemistry, network generation programs attempted to incorporate on-the-fly computational chemistry calculations [78, 79].

The use of on-the-fly computational chemistry calculations in network generation algorithms requires the conversion of the graph representation of the molecule into a geometry in terms of 3D coordinates for each of the atoms in the structure. Next, the obtained 3D geometry is then used as the starting point for computational chemistry methods that determine the thermochemical properties of the molecule. The most recent program [45] that implements this strategy of on-the-fly generation of thermochemical properties uses semi-empirical methods such as PM3 [80, 81] and force-field methods such as MM4 [82]. Magoon and Green validated the accuracy for 43 polycyclic hydrocarbons [83] and reported a mean absolute error of 29 kJ mol⁻¹ for the enthalpy of formation using PM3 compared to experimentally measured values. This was significantly more accurate than the value of 167 kJ mol⁻¹ obtained by the group additivity approach [79], and could be reduced to 5.4 kJ mol⁻¹ using MM4 if only a subset of 23 molecules was validated for which suitable MM4 parameters were available. Despite these promising results, on-the-fly generation of thermochemical properties is not a panacea. The use of *ab initio* or even DFT methods in on-the-fly estimation of thermochemical properties has not been demonstrated yet. Moreover, the variability of the mean absolute deviations for the computationally cheaper semi-empirical and force-field methods illustrate that accuracy of the obtained results is heavily dependent on the type of molecules that is studied.

1.1.3 Kinetics

Complex reaction networks can easily contain up to $O(10^4)$ reactions, especially when only elementary reaction paths are included. Rate coefficients for each of these reactions need to be determined in order to be used in numerical simulations. Moreover, rate coefficients are heavily temperature dependent, and are often expressed through the Arrhenius equation, cf. Eq. 1-4.

$$k = A \exp\left(\frac{-E_a}{RT}\right) \quad \text{Eq. 1-4}$$

With A the pre-exponential factor, E_a the activation energy, R the universal gas constant, and T the temperature. Experimental values of rate coefficients for individual elementary steps are challenging and hence scarce, partially because of the difficulty in detecting and measuring unstable intermediates such as radicals. Theoretical derivation of rate coefficients is often based on transition state theory, cf. Eq. 1-5.

$$k(T) = \frac{k_B T}{h} \cdot \frac{q^\ddagger}{\prod_r q_r} \exp\left(-\frac{\Delta E(0K)}{RT}\right) \quad \text{Eq. 1-5}$$

With q the partition function per unit volume, and $\Delta E(0K)$ the electronic barrier. Here, r stands for reactants and \ddagger for the transition state of the reaction. This formulation is particularly useful since all quantities of Eq. 1-5 can be calculated with quantum-chemical methods and statistical mechanics. Partition functions are calculated from vibrational frequencies and moments of inertia while the electronic barrier is calculated as the energy difference between reactants and the transition state. Despite the potential of *ab initio* methods for the calculation of rate coefficients, its use in kinetic model generation tools remains limited. Similarly to the calculation of thermochemical properties of molecules, the use of *ab initio* methods for rate coefficients remains a time-consuming and CPU-intensive task that still requires a substantial manual intervention. As a consequence, the need for predictive methodologies for the estimation of rate coefficients is as critical as for thermochemical properties.

The starting point of all predictive methodologies that aim at estimating rate coefficients is the categorization of elementary reactions in reaction families, in which the same rearrangements of

atoms occur. One of the simplest methods to parameterize rate variations between members of the same reaction family is the use of linear free-energy relationships (LFERs), cf. Eq. 1-6.

$$\ln k_i = \ln k_0 + m(x_i - x_0) \quad \text{Eq. 1-6}$$

Where k is the rate coefficient of reaction i and the reference reaction 0 , m is characteristic of the reaction family, x_i is some property of the species involved in the reaction i , and x_0 corresponds to the reference reaction that defines the reaction family. Examples of LFERs include the Evans-Polanyi correlation [84] that describes a linear relation between the activation energy and the reaction enthalpy for a series of homologous reactions, cf. Eq. 1-7.

$$E_a^i = E_0 + \alpha \cdot \Delta_r H_i^\circ \quad \text{Eq. 1-7}$$

with E_a^i the activation energy of reaction i , and $\Delta_r H_i^\circ$ the reaction enthalpy. E_0 and α are the intrinsic barrier and transfer coefficient respectively. The pre-exponential factor for the reactions belonging to the same reaction family is taken constant, assuming that differences in entropic changes during the reaction are zero, or proportional to enthalpic changes for all members of the same reaction family.

Based on successes for the accurate estimation of thermochemical properties, group contribution methods were also applied to rate coefficients. Willems and Froment [85, 86] applied a structural group contribution method to model the reactions involved during the steam cracking of hydrocarbons, and derived numerical values for the groups contributing to the activation energy from experimental data. With the advent of quantum-chemical methods many group contribution methods were proposed that involve the calculation of transition state properties [87]. Sumathi et al. [87-90] calculated the thermochemistry of the transition state using group additivity. They split the transition state into unreactive moieties that were also present in the reactants and reactive moieties; also called supergroups. Supergroups contrast the original definition by Benson [65] because they comprised more than one central atom. Moreover, the application of this concept to bimolecular reaction families such as H-abstraction resulted in a large number of supergroups since they contained atoms that were both part of the abstracting radical and of the attacked hydrocarbon. More recently, group additive methodologies were devised in which the transition state is split up into groups consisting of only a single central atom, similarly to the Benson approach for thermochemical properties of species. Hence, this drastically reduced the

number of groups required to cover the reactions belonging to a reaction family. Saeys et al. [91, 92] proposed a method following this approach to model the activation energy of addition and H-abstraction reactions while others [93-99] extended the formalism to model both activation energies and pre-exponential factors. The method is illustrated for radical addition reactions to alkenes, cf. Figure 1-8.

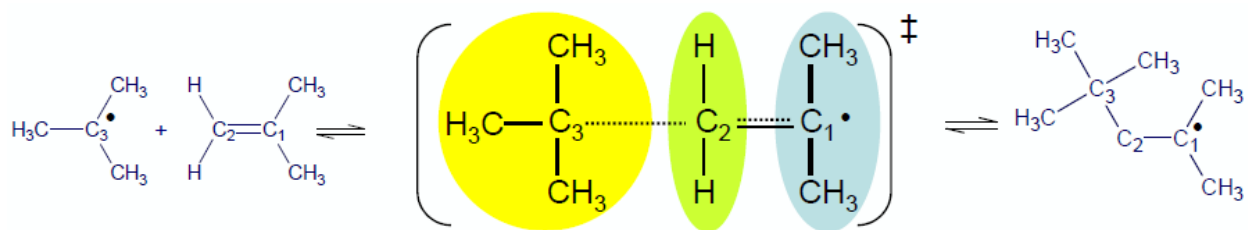


Figure 1-8: Identification of the three groups in the transition state for the addition of t-butyl to isobutene [100].

Arrhenius parameters can be written as perturbations of a reference reaction, cf. Eq. 1-8 and Eq. 1-9.

$$E_a(T) = E_{a,ref}(T) + \sum_i \Delta GAV_{E_a}^0(C_i) \quad \text{Eq. 1-8}$$

$$\log_{10} A(T) = \log_{10} \tilde{A}_{ref}(T) + \sum_i \Delta GAV_{\tilde{A}}^0(C_i) + \log_{10} n_e \quad \text{Eq. 1-9}$$

$E_{a,ref}$ and \tilde{A}_{ref} are the activation energy and pre-exponential factor of the reference reaction of the family and ΔGAV^0 the group additive values accounting for the contribution of structural differences between a member of the family and the reference reaction. For the addition reaction, three groups are identified that correspond to the attacking radical (C_3), the atom onto which the attacking radical adds (C_2) and the atom that receives an unpaired electron after the addition (C_1). The advantage of introducing a reference reaction in Eq. 1-8 and Eq. 1-9 is that temperature dependencies can be captured within $E_{a,ref}(T)$ and $\log_{10} \tilde{A}_{ref}(T)$ leaving the ΔGAV^0 s almost temperature independent. The term n_e refers to the number of single events of the target reaction and needs to be determined separately. For reactions in which cross-resonance and/or hyperconjugative effects in the transition state [101] are observed, the group additivity model needs to be expanded with correction factors $\Delta E_{a,res}^0$ and $\log_{10} \Delta \tilde{A}_{res}^0$, cf. Eq. 1-10 and Eq. 1-11.

Since resonance corrections account for cross-effects across the entire transition state structure, originating from the simultaneous presence of multiple groups, they cannot be incorporated in either of the groups, and lead to contributions that incorporate multiple central atoms.

$$E_a(T) = E_{a,ref}(T) + \sum_i \Delta GAV_{Ea}^0(X_i) + \Delta E_{a,res}^0 \quad \text{Eq. 1-10}$$

$$\log_{10} A = \log_{10} \tilde{A}_{ref} + \sum_i \Delta GAV_{\tilde{A}}^0(X_i) + \log_{10} \tilde{\Delta A}_{res}^0 + \log_{10} n_e \quad \text{Eq. 1-11}$$

The derivation of numerical values for the contributions in this group additivity approach based on high level *ab initio* methods has been shown to provide reliable estimates for H-abstraction reactions [94, 97, 99], radical additions [91, 95], and homolytic substitution reactions [98].

Despite the potential of group contribution methods and LFERs, they have been applied to a very limited number of reaction families, often for only a few possibilities within a single family. The development of methods for the *a priori* calculation of rate coefficients are often based on strict assumptions that may not be fulfilled for other types of reaction families. Moreover, when methods make use of *ab initio* calculations, the performance of these methods for a set of reactions is dependent on the accuracy of the level of theory, and the size of the training set of reactions that are used to estimate the model parameters. To overcome the discrepancy between the available kinetic data and the large number of rate coefficients in kinetic models, more generally applicable methods should supplement these efforts, but accompanied by the inevitable loss of accuracy. In RMG, for example, the database of Arrhenius parameters for a reaction family is structured using hierarchical trees for each of the reactants in the reaction family [102, 103], cf. Figure 1-9. Details of the molecular structure of a reactant are systematically increased from the root node, i.e. the most generic representation of the reactant, to leaf nodes, i.e. representations of specific molecules.

When Arrhenius parameters cannot be found for a specific reaction, the hierarchical tree is used to determine the most appropriate parameters for the reaction, based on the proximity and similarity of other nodes in the tree for which kinetics are available. This ensures that to each reaction that is created during the network generation procedure Arrhenius parameters are

assigned, which, in the worst case, correspond to the rate rules of the root node of the tree of the reaction family.

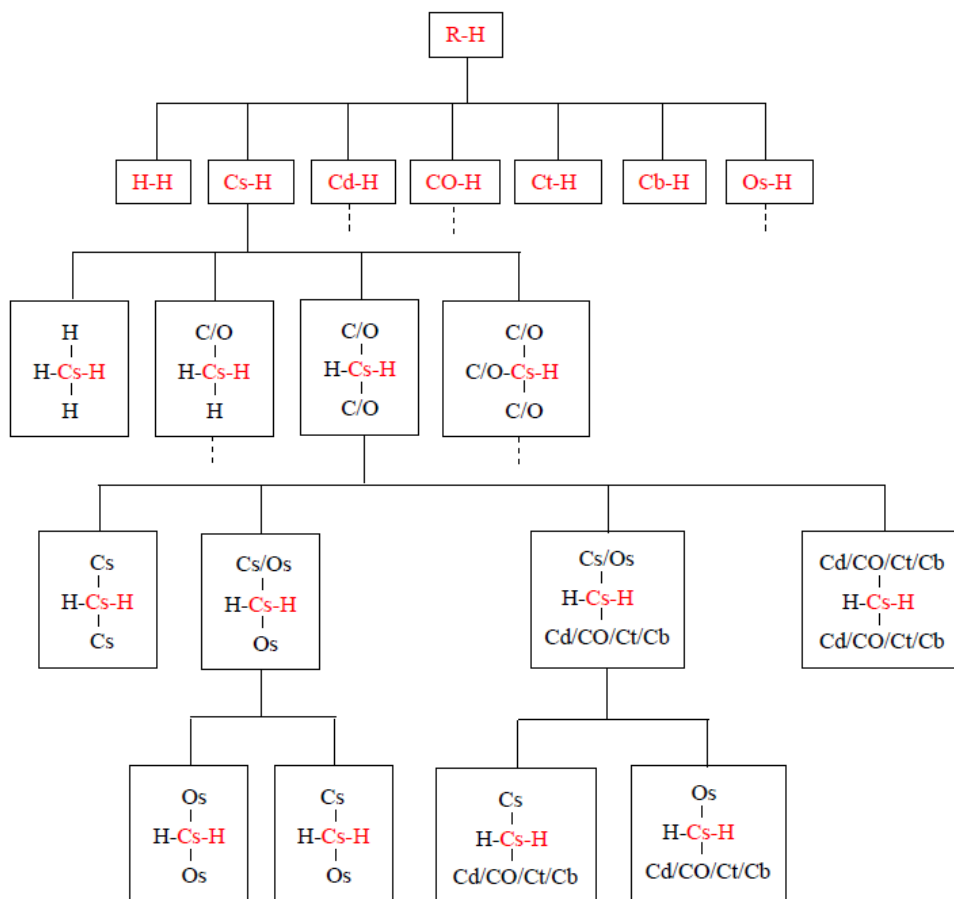


Figure 1-9: Hierarchy tree structure of the kinetics database of a reaction family in RMG [102, 103].

Finally, automated methods for locating transition states can also be considered as a route for the automated calculation of rate coefficients, similarly to on-the-fly generation of thermochemical properties. Several algorithms are implemented in existing quantum-chemical software packages [104, 105] but are often unsuccessful, and require initial guess geometries very close to the saddle point on the potential energy surface. In reaction network generation programs, geometries associated with the transition state are *a priori* unavailable and need to be derived from the graph representation of reactants and/or products. Automated calculations of transition states using computational chemistry calculations is still in its initial stages of development. Very recently, a proof of concept was given by Bhoorasingh and West [106] in which the transition state geometries of a set of H-abstraction reactions were automatically localized. The reader is referred

to Section 5.3.2 of Chapter 5 for more information on the methodology for the calculation of rate coefficients of elementary reactions used in this thesis.

The presented overview illustrates that, despite the wide diversity of existing automated kinetic model generation approaches, all of them require a solution to a limited number of key issues: the verification of the uniqueness of a molecule, the identification of sub-molecular patterns in the molecule, and a molecule representation that is suitable for the intended application. A recurring phenomenon among many of the presented network generation techniques is that solutions to these three key issues are tailored to the intended application. The strength of this strategy of tailored solutions is that it reduces the short-term investment of human capital for the development of the program, and hence allows spending more time on studying the actual chemical process at hand. The other side of the coin is that the offered solutions are often not suitable for other, even related applications. The resulting inefficiencies in communal efforts could be avoided by focusing on extensible, reusable, and standardized solutions. This principle was acknowledged in pharmaceutical sciences and industry and led to the field of chem(o)-informatics. In what follows, a brief overview is given of the solutions that are provided in the field of chemoinformatics to the three aforementioned issues of automated network generation.

1.2 Chemoinformatics

Chemoinformatics is a relatively young field at the cross section of computer science, mathematics, and chemistry with the aim of handling chemical information. With the advent of tremendous computer power in the last decades, algorithms and tools can now be employed to search and manipulate large amounts of chemical data. Advances in chemoinformatics were used in molecular modeling, chemometrics, building databases of chemical structures and reactions, quantitative structure activity/property relationships, computer-assisted synthesis design, and automatic structure elucidation. Moreover, the use of computers significantly increased the pace of innovation, especially in pharmaceutical sciences. Historically, lead components that bind to a target receptor were discovered by costly *in situ* drug discovery experiments. With the advent of chemoinformatics, virtual screening, i.e. the *in silico* selection of molecules likely to show a

desired bioactivity from a large database, has become a prominent tool in the computational drug discovery toolbox.

Many of the algorithms developed in chemoinformatics use graphs as a fundamental abstraction for diverse chemical concepts. A graph $G(V,E)$, or network is a mathematical structure consisting of a set of vertices V connected by a set of edges E [6, 107]. Every edge has two endpoints in the set of vertices, and is said to connect or join the two endpoints. Many concepts that are important in kinetic models can be represented by variations of graphs [108]. In this thesis, weighted graphs are often used, i.e. graphs in which weights or labels are associated to the vertices and edges of the graph. For example, a kinetic model can be represented by a weighted graph: molecules represent the vertices and are connected by reactions (edges). The weights of the vertices and edges can refer to the thermochemical properties of molecules and rate coefficients of elementary reactions, respectively. Molecules can also be represented by weighted graphs. Atoms and bonds are represented by vertices and edges, respectively while atom and bond attributes such as elemental information, electrons, bond orders can be considered as weights of atoms and bonds. Through the use transition state theory, an elementary reaction can be linked to a unique transition state structure. The latter can also be represented by a weighted graph, in which changes in electronic structure that occur during the course of a reaction can be represented by weights of the atoms and bonds that take part in the reaction.

1.2.1 Species uniqueness and canonicalization

Canonicalization refers to the unique and unambiguous numbering of atoms in a molecule. If a molecule can be canonicalized, a canonical representation can be used to compare molecules, store and retrieve chemical information for the molecule in external databases, and generate unique name (identifier) for the molecule. The historically most important canonicalization algorithm is the algorithm by Morgan [109]. The procedure starts by partitioning the atoms based on graph invariants. A graph invariant is a number, or a sequence of numbers computed from the graph topology, and independent of the indices of the atoms in the graph. Examples of graph invariants are the degree of an atom, the atomic number of the atom, etc. Next, the graph invariant associated to each atom is recursively updated by taking the sum of the invariants of its

neighbors computed in the previous step, cf. Figure 1-10. The recursive procedure stops when the largest number of distinct invariants is obtained.

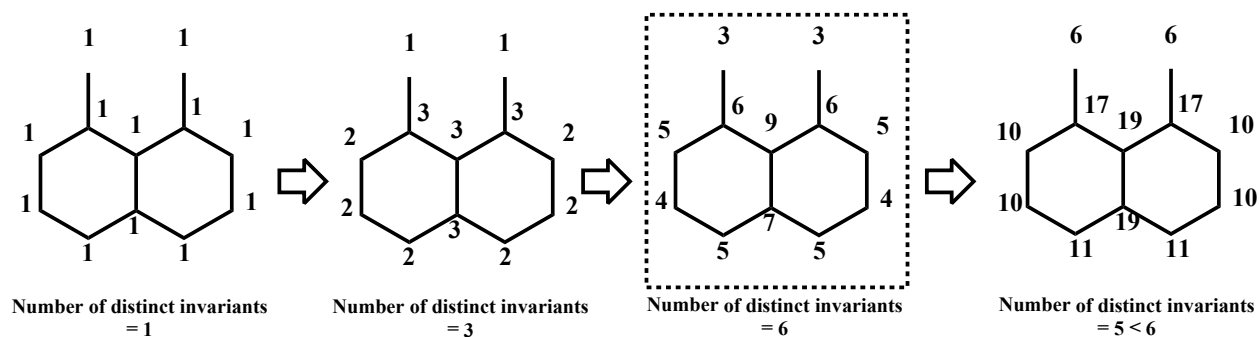
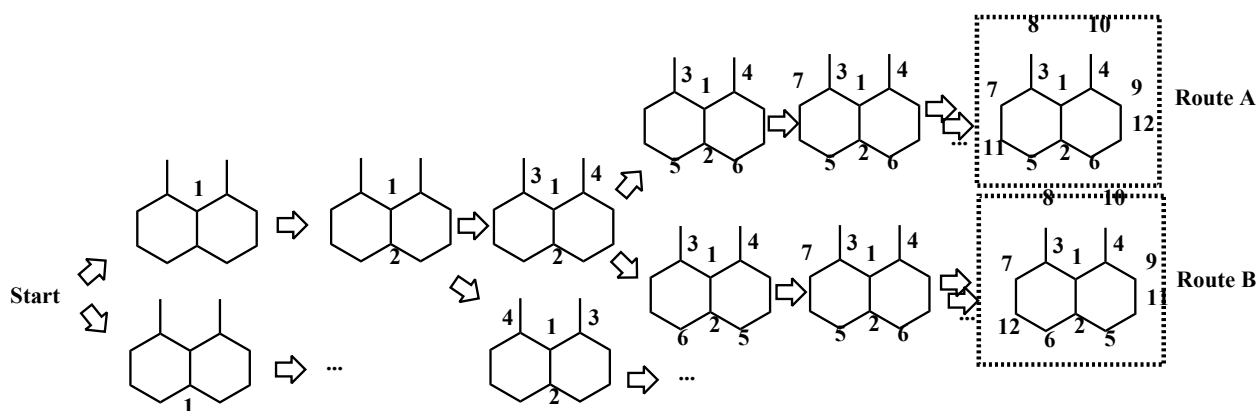


Figure 1-10: Recursive calculation of graph invariants for the atoms in dimethyldecalin. Initial invariant chosen to be 1. From ref. [110].

Indices are iteratively assigned to each of the atoms, in decreasing order of the graph invariant, calculated in the previous step, cf. Figure 1-11. If two atoms have the same graph invariant, both possibilities are generated.



Route A: [1,2][1,3][1,4][2,5][2,6][3,7][3,8][4,9][4,10][5,11][6,12][7,11][9,12]

Route B: [1,2][1,3][1,4][2,5][2,6][3,7][3,8][4,9][4,10][5,11][6,12][7,12][9,11]

Figure 1-11: Exhaustive atom numbering dimethyldecalin in decreasing order of the graph invariant, cf. Figure 1-10. Subsequent enumeration of the bonds, indicated by the atom indices for each possible numbering, cf. route A, and route B. From ref. [110].

Finally, the bonds of the molecules are enumerated in ascending order of the atom indices that were assigned and for each of the possible numbering scenarios that exist. Next, one ordering of bonds is chosen and corresponds to the “canonical” indices of the atoms in the molecule. It is

unimportant which bond ordering is chosen as long as the same bond ordering is always selected for the same molecule. Since it can be shown that the set of bond orderings is finite for any molecule composed of a finite number of atoms, it is possible to select the unique atom numbering by generating all bond orders, lexicographically ordering them based on the indices involved in the description, and subsequently selecting the first member of the resulting list as the unique atom indices. For example, routes A and B in Figure 1-11 differ only in the two last bonds of the description. Route A is taken as the atom indices of dimethyldecalin since the index 11 of bond [7,11] of route A precedes the corresponding index 12 in bond [7,12] of route B.

Several variations [111-113] of the canonicalization principle exist and differ in the choice of initial graph invariant, or in the recursive procedure of invariant generation. Morgan [109] developed the algorithm for generating a unique machine description for chemical structures at the Chemical Abstracts Service, the service that provides the widely used CAS-registry numbers for chemical structures. Since then, several so-called canonical molecule identifiers have been developed. One identifier that is systematically used in this thesis is InChI (International Chemical Identifier) developed by IUPAC and NIST [114, 115]. Figure 1-12 shows an example of an InChI of (R)-glutamic acid. Every InChI starts with the string “InChI” followed by the version number, and the molecular formula. The remaining information is structured as a sequence of layers, with each layer providing one specific type of information. For example, the layer that starts with ‘c’ describes the atom connectivity. For example, atom 6 is connected to atom 3. Atom 3 is connected to atom 1 and to a side branch 5 (9) 10 starting with atom 5. Bond order information is not stored explicitly, but rather derived from the ‘h’ layer that contains the information on the connectivity of hydrogen atoms. For example, atom 3 bears only one hydrogen atom (expressed as 3H), while atoms 1, 2, and 6 all bear 2 hydrogen atoms (1-2, 6H2).

InChI=1S/C5H9NO4/c6-3(5(9)10)1-2-4(7)8/h3H,1-2,6H2,(H,7,8)(H,9,10)/p+1/t3-/m1/s1

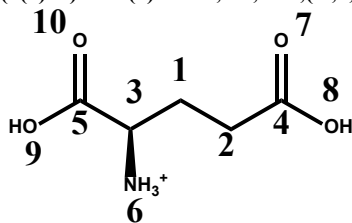
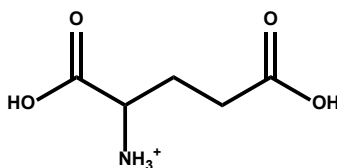


Figure 1-12: InChI and structure of (R)-glutamic acid. Indices refer to the indices used in the connectivity layer (starts with /c) of the InChI. From ref. [110].

The layer starting with ‘p’ represents the net charge of the substance. Information for double bond stereochemistry, and sp^3 tetrahedral (‘t’) stereochemistry is stored in subsequent layers. Stereolayers can be omitted; in that case the absolute stereoconfiguration of stereocenters is unspecified. As a consequence, molecules can be compared by including or excluding the stereolayers, whether or not taking into account the configuration of stereocenters.

InChI=1S/C5H9NO4/c6-3(5(9)10)1-2-4(7)8/h3H,1-2,6H2,(H,7,8)(H,9,10)/p+1



The power of canonical identifiers such as InChI is that there is a one-on-one relation between the identifier and the molecular structure. Each molecule is represented by a unique canonical identifier, and each canonical identifier leads to only one molecule. Canonical identifiers are primarily intended as computer-readable notations, and not designed for the human readability. Other, more human-friendly identifiers, such as the ubiquitous SMILES notation [116], or the IUPAC name can be used to specify the structure of a molecule. For example, a SMILES string that represents (R)-glutamic acid is [NH3+][C@@H](C(O)=O)CCC(O)=O. Another valid possibility is [NH3+][C@@H](C(O)=O)CCC(=O)O, and illustrates that the classical SMILES notation is non-canonical.

1.2.2 Graph and sub-graph isomorphism

As shown in the previous section, the identity of two molecules can be verified by canonicalizing the structures, and a subsequent string comparison of the identifier derived from the canonical atom labels. Next to canonicalization, molecules can also be compared through isomorphism

algorithms. Two graphs are isomorphic when there is a one-to-one mapping from the vertices of one graph to the vertices of the second graph, such that the edge connections are respected. The problem definition can be generalized into a sub-graph isomorphism problem, i.e. the computational task in which two graphs G and H are given as input, and one must determine whether G contains a sub-graph that is isomorphic to H . The output of sub-graph isomorphism algorithms is a mapping between the matched vertices of graph G and H . If a sub-graph could not be found in a graph, then the set of atom mappings will be empty.

For network generation applications, sub-graph isomorphism algorithms can be applied in various tasks. The eligibility of a molecule to undergo a reaction family can be tested by verifying whether the pattern of atoms that is characteristic to the reaction family, is present in the molecule. For the application of reaction family constraints, the reverse relation can be tested: a molecule is forbidden to undergo a specific reaction if a particular sub-graph, i.e. constraint, is found in the molecule. In group contribution methods, the presence of a specific group from a database of available groups can be verified by using sub-graph isomorphism algorithms. In the context of network generation, the sub-graph isomorphism problem is formulated as follows: a pattern of atoms, also called query pattern, is searched for in a molecule, also called target graph. For each of the atoms in the query graph, a corresponding matching atom in the target graph needs to be present, in order to have a positive match between the query graph and the target graph, cf. Figure 1-13.

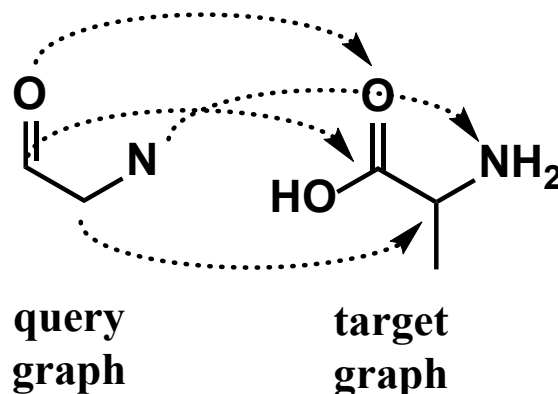


Figure 1-13: A match between each of the atoms of the query graph (left) and an atom in the target graph (right) is searched in the formulation of the sub-graph isomorphism problem in network generation applications.

Finding a solution to a sub-graph isomorphism problem is known to be computationally expensive [117], to the extent that it becomes infeasible for large enough graphs. One strategy for solving sub-graph isomorphism problems is to translate them into the detection of the maximum clique in the compatibility graph (modular product graph, association graph, correspondence graph). A clique is a graph such that every two vertices are connected by an edge. Each vertex in the compatibility graph - also called an *a*-vertex – corresponds to a matching between a vertex of the query graph and a vertex of the target graph. An *a*-edge between the *a*-vertices (q,m) and (p, n) is created if $q \neq p$ and $m \neq n$ and if 1) there does not exist an edge (q, p) in the query graph nor an edge (m, n) in the target graph or 2) the edge (q, p) matches the edge (m, n). The example in Figure 1-14 shows the compatibility graph of the query graph and the target graph corresponding to alanine. A vertex (1,1') exists in the compatibility graph because both atoms are oxygen atoms. An edge between (1,1') and (5,3') exists because atoms 1 and 5 in alanine are not connected, nor is this the case for atoms 1' and 3' in the query graph.

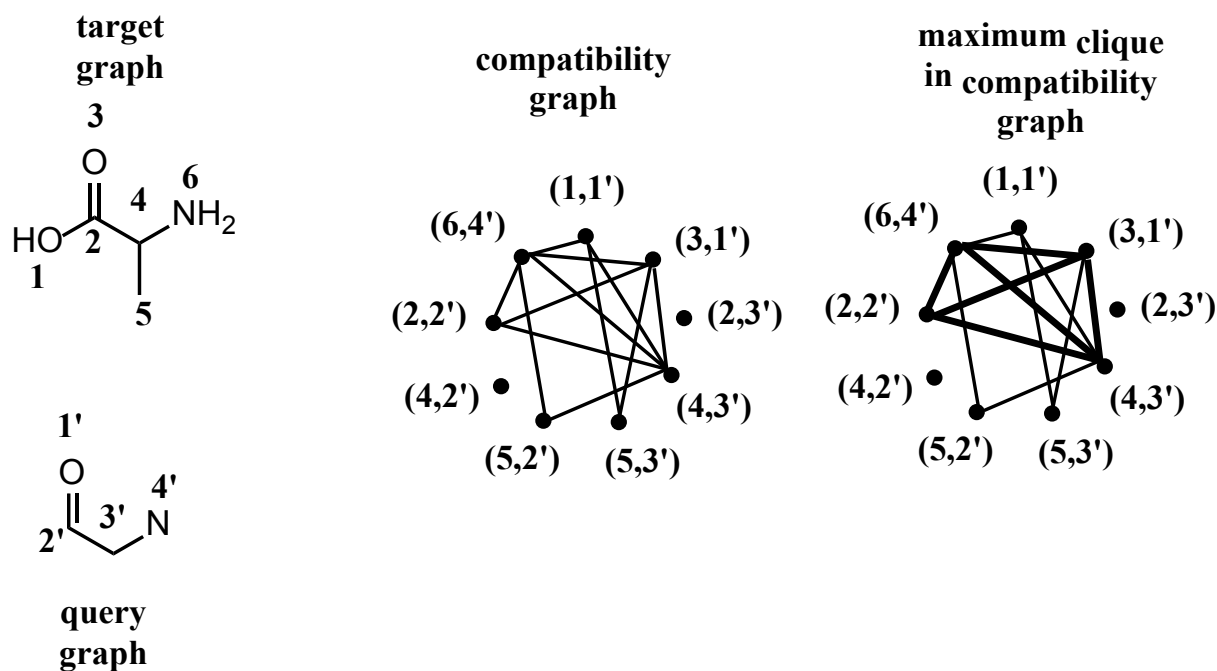


Figure 1-14: the target graph (alanine), and a query graph lead to the corresponding compatibility graph. A vertex in the compatibility graph exists when the atoms in the query and target graph belong to the same chemical element, e.g. oxygen atoms 1, and 3 in the target graph and atom 1' in the query graph. An edge in the compatibility graph exists when there is no corresponding edge in the query nor target graph OR an edge in the query graph matches an edge in the target graph. E.g. vertices (1,1') and (6,4') in the compatibility graph are connected by an edge because vertex 1 and 6 are not connected in the target graph, and vertex 1' and 4' are not connected in the query graph. The sub-graph highlighted in bold is the maximum clique that is present in the compatibility graph.

The next step involves the detection of the maximal cliques in the compatibility graph. If the clique with the highest number of vertices contains the same number of vertices as the query pattern, the query graph is found in the molecule, and vice versa. For example, the queried graph in Figure 1-14 contains 4 atoms and so does the maximum clique detected in the compatibility graph. Maximum clique detection is a well-studied problem for which many solutions exist [118-122].

Translating the substructure matching problem into the more expensive problem of finding the maximum common sub-graph of the query graph and the target graph is not ideal. Therefore, several efforts were made to create faster substructure matching algorithms [37, 123-126], with

multiple implementations tailored to chemical graphs [127, 128]. These algorithms are very similar to the manual methods that chemists use to compare two chemical structures. The atom mappings of atom matches in both structures are iteratively expanded by pairwise comparisons of the neighboring atoms of a previous successful atom match and backtracking to the last successful state whenever the atom comparison fails. Several variations of this theme exist such as the algorithm by Ullmann [37] and Vento-Foggia (VF) [124, 125] and differ in the effectiveness in pruning off fruitless branches of a search tree.

Next to the development of algorithms for the detection of sub-graphs or chemical patterns in molecules, languages were developed with the aim to unambiguously describe these chemical sub-graphs. Pattern languages express sub-molecular patterns in a condensed string so that they can be exchanged or stored in databases. For example, external databases containing Benson group additive values could be constructed by mapping sub-molecular patterns (groups), expressed by the syntax of the pattern language, onto a set of associated values (group additive values). The use of a pattern description language provides a basis for the standardization of chemical patterns, regardless of the application the pattern is used for. The language used in this thesis is the SMILES Arbitrary Target Specification (SMARTS) [129]. It extends the rules of SMILES. For example, the string [c,N] denotes a SMARTS atom that can be either an aromatic carbon or an aliphatic nitrogen; the logical OR operator is expressed by the colon. SMARTS strings allow both the description of very general or very specific patterns, cf. Table 1-1.

Table 1-1: Examples of SMARTS strings and their description.

String	Description
<chem>[CX4]</chem>	Alkyl Carbon connected to four neighbors
<chem>[NX3][CX3]=[CX3]</chem>	Enamine
<chem>[\$([cX2+](::*)::*)]</chem>	sp2 cationic carbon
<chem>[#7;X2v4+0]</chem>	Nitrogen Free-Radical
<chem>[nX2r5]</chem>	N in 5-sided aromatic ring

The reader is referred to Section 5.3.1 of Chapter 5 for more information on the species representation and on the algorithms for (sub)-graph isomorphism applied in this thesis.

Next to the development of pattern languages, methods were developed to uniquely represent sub- molecular parts, e.g. atoms, in a molecule [130, 131]. Similar to unique line identifiers for molecules, these sub-molecular segments identifiers, called “hash-codes” provide a unique way of expressing the environment of a particular atom in a molecule. By combining the hash codes of the different atoms in a molecule, molecule hash codes can be constructed, resulting in similar canonical representations of molecules as the InChI identifier.

1.2.3 Stereochemistry and 3D structure

As discussed previously, the graph representations of molecules allow graph-theoretic algorithms to be implemented for network generation. However, the pure graph-oriented representation of molecules is not always sufficient to understand the interactions of molecules with their environments, and requires information on the three-dimensional arrangement of atoms in some cases as well. This is especially true for stereoisomers, i.e. isomers of identical constitution but differing in the arrangement of their atoms in space [132]. To account for the 3D nature of molecules, more advanced molecule representations are required.

The most powerful representation that overcomes the limitations of a graph representation is a 3D (topographical) molecular representation that departs from the geometry of the molecule, cf. Figure 1-15.

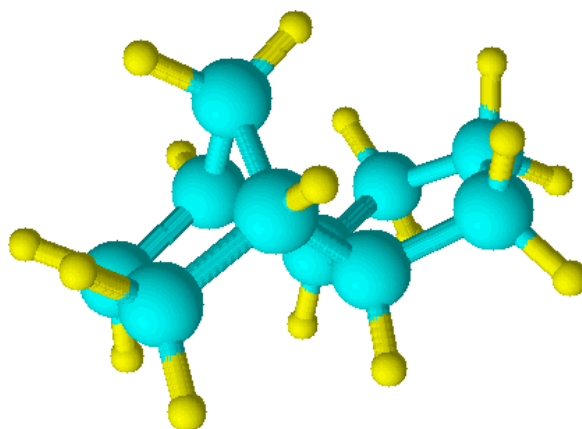


Figure 1-15: 3D geometry of exo-tricyclo[5.2.1.0^{2,6}]decane. Depicted using ACDLabs’ ACD/3D Viewer v11.

The geometry of a molecule is, ultimately, governed by the distribution of electrons in the molecule, and thus can only be obtained from the fundamental laws of quantum-mechanics. Nevertheless, more practical solutions were provided in the form of 3D structure generation

algorithms [133]. These 3D models are very powerful because they do not contain any generalizations on the geometry of the molecule. Moreover, they can be used as input to computational chemistry tools, e.g. for the calculation of thermochemical properties as described in a previous section. Explicit atom coordinates also allow the calculation of dihedral angles, and Euclidean distances between atoms and allow the detection of molecular properties. For example, chiral components are characterized by the lack of improper rotational (rotation-reflection) axes S_n . Therefore, a chirality detection algorithm based on 3D atomic coordinates, such as in ref. [134], consists of the execution of rotation-reflection operation S_n aligned along one of the principal axes of inertia of the molecule. Next, a scalar f is calculated that reflects the similarity between the original and transformed molecule, based on the Euclidean distances r_i between the original and transformed atoms i , cf. Eq. 1-12.

$$f(S_n) = \sum_i r_i^2 \quad \text{Eq. 1-12}$$

If a molecule is achiral, then at least one improper rotational symmetry element will be found, leading to $f(S_n) = 0$.

3D models, however, have a number of disadvantages. First of all, they describe fixed positions of the atoms in the molecule and inherently are unable to deal with conformational flexibility. Similarly to graphs, they represent an abstraction of the true geometry of molecules. More importantly, 3D coordinates are usually *a priori* unavailable and therefore need to be generated, often starting from the graph representation of the molecule. Although the 3D structure generation has made many advances during the last decades [8, 133], a generally accepted tool is not available yet. Many 3D structure generation programs are specifically parameterized for only a subspace of molecules. Outside that domain, the algorithm might not be able to converge, or at least generate doubtful and inaccurate 3D coordinates. For molecules with rotatable bonds, the 3D structure generation algorithm might be unable in successfully identifying the lowest-energy conformation.

An alternative to 3D models is the augmented graph representation, or 2.5D model, for the description of stereochemistry. The augmented graph representation starts from the notion that stereochemistry in molecules can be described by focusing on sub-molecular stereocenters.

Stereocenters are atoms, bonds, planes, axes, etc. in a molecule such that interchange of two ligands attached to the stereocenter leads to another stereoisomer [135]. In this framework, special purpose attributes are added to the graph representation to denote mobility restrictions implied by the stereoconfiguration. For example, in graphical depictions of stereoisomers, this can be reflected by adding filled and dashed wedges to bonds to describe whether the bond is pointing upwards or downwards relative to the plane of the 2D molecular diagram, cf. Figure 1-16.

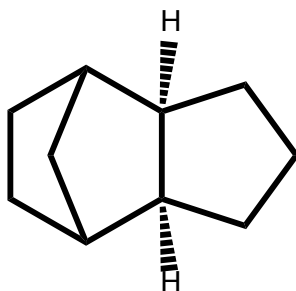


Figure 1-16: Graphical depiction of exo-tricyclo[5.2.1.0^{2,6}]decane using wedged bonds based on the augmented graph representation of the molecule. (“2.5D representation”)

Although the wedged bond representation in 2D diagrams of molecules is often clear for human interpretation, it leads to ambiguous interpretation by computers. Molecules can be depicted from different view angles, implying that the wedge up/down attributes of bonds cannot be considered to allow computers to distinguish between stereoisomers. Unambiguous identification and description of the chemical structure of a component requires systematic naming schemes that incorporate stereodescriptors in the names. Stereodescriptors are prefixes that specify the absolute configuration of stereocenters in a molecule. Many different stereodescriptors were devised, e.g. based on optical activity (+/-) or by relating to glyceraldehyde (D/L), etc... Probably the most widely used method is based on the priority rules of Cahn, Ingold and Prelog (CIP) [136]. The CIP method algorithmically associates labels, e.g. R/S and E/Z, to atoms and bonds based on the prioritization of the substituents of the stereocenters. For example, one of the best known rules is that higher atomic numbers precede lower ones, and allows to prioritize the four halogen atoms in (R)-bromochlorofluoroiodomethane, as follows: I > Br > Cl > F. The stereodescriptor “(R)” denotes that the three highest prioritized ligands rotate in clockwise sense around the C–F axis, cf. Figure 1-17.

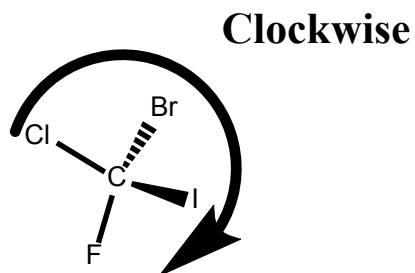


Figure 1-17: Assignment of the (R) configuration to an enantiomer of bromochlorofluoriodomethane after ranking the four ligands attached to the chiral center according to the CIP rules. ($I > Br > Cl > F$)

The CIP method illustrates that an unambiguous interconversion between a graph representation augmented with the CIP stereodescriptors and the absolute configuration of stereocenters in a molecule is possible. It is also observed that for most common forms of stereoisomerism, e.g. tetrahedral, double bond, axial chiral stereoisomerism, a binary stereodescriptor suffices in describing the absolute configuration of the stereocenter, leading to the notion of stereoparity, to emphasize the binary character of most stereocenters. Computer algorithms may use any type of stereoparity (+1/-1, 1/2, etc...) as long as it is possible to reconstruct the molecule unambiguously from its code, as exemplified by the stereoparities in the canonical identifier InChI and other stereochemically unique naming algorithms [137].

The reader is referred to Chapter 8 for more information on the implementation of stereochemistry used in this thesis.

1.3 Objective and thesis overview

The contributions of this thesis cover a range of topics, but can be broadly gathered under the heading of the use of automated methods to construct kinetic models. In particular, the need to account for the three-dimensional arrangement of atoms in molecules of kinetic models was investigated. To this end, the thermal rearrangement of two monoterpenoids, *cis*- and *trans*-2-pinanol, was studied and a kinetic model that accounts for the differences in reactivity and selectivity of both pinanol diastereomers was proposed. Secondly, the thermal decomposition of the fuel Jet-Propellant 10 (JP-10) was studied experimentally. JP-10 mainly consists of a tricyclic, strained $C_{10}H_{16}$ hydrocarbon; accounting for the three-dimensional nature of this molecule is

essential to understand its decomposition behavior. A large set of JP-10 pyrolysis experiments were carried out at the bench-scale reactor setup of the Laboratory of Chemical Technology of Ghent University and allowed gaining new insights in the decomposition model of this fuel. Because existing models for the thermal decomposition of JP-10 have proven to be largely inadequate in modeling these in-house experiments, a new kinetic model was constructed and features the use of state-of-the-art techniques for the automated estimation of thermochemistry of polycyclic molecules.

Furthermore, a new program was developed for the automated rule-based construction of kinetic models and takes advantage of the advances made in the field of chemoinformatics to tackle fundamental issues of previous approaches. The program was programmed in the programming language JAVA. The latter is one of the main industry standards for object-oriented programming. JAVA is easy to learn, possesses powerful features, and is very popular in the chemoinformatics community, with many free, open-source implementations of important algorithms. Novel algorithms were developed for important aspects of automated construction of kinetic models: the estimation of thermochemistry and kinetics using predictive structure-property methods (Section 5.3.2 of Chapter 5), and the calculation of symmetry of molecules and reactions (Chapter 6). Genesys was used for the kinetic modeling of diethylsulfide pyrolysis (Chapter 7). The application illustrates the possibility to automatically generate a kinetic model involving heteroelements, with thermochemical data and kinetic data in the form of GAVs, and ΔGAV° s respectively, obtained from *ab initio* data from literature. Finally, the incorporation of stereochemistry in kinetic models is discussed in Chapter 8 and is illustrated by the automatic construction of a kinetic model for the thermal rearrangement of 2-pinanol.

In this thesis, Genesys is solely used to study non-catalytic, free radical processes in the gas phase, primarily because of the availability of accurate methods to determine thermochemical and kinetic properties of molecules and reactions. Hence, the methods for the calculation of these properties are thus tailored to the aforementioned types of chemical processes. Nevertheless, the algorithms for reaction network generation are in principle also valid for other types of processes, e.g. involving ions instead of free radicals.

This thesis consists of a compilation of published journal papers and manuscripts that are to be submitted for publication in the near future. Every paper/manuscript describes a distinct aspect of automated kinetic model generation and is preceded by a short introduction describing the context. First, the work on the thermal rearrangement of 2-pinanol is presented (Chapter 2). The thermal decomposition of JP-10 is studied in two consecutive chapters. In Chapter 3, the JP-10 pyrolysis experiments are described and used to validate available kinetic models of the thermal decomposition of JP-10. The new dedicated model for JP-10 pyrolysis is introduced and validated in Chapter 4. The functionality of Genesys is described in the subsequent chapters starting from a global overview of Genesys (Chapter 5). This is followed by a description of the algorithm for the calculation of symmetry of molecules and transition states (Chapter 6). The application of Genesys to the thermal decomposition of diethylsulfide is illustrated in Chapter 7. The incorporation of stereochemistry in kinetic models is discussed and validated in Chapter 8. Finally, conclusions and perspectives are given in the last chapter (Chapter 9).

1.4 References

- [1] W.H. Green, Predictive Kinetics: A New Approach for the 21st Century, *Advances in Chemical Engineering*, 32 (2007) 1-50.
- [2] M.K. Sabbe, K.M. Van Geem, M.-F. Reyniers, G.B. Marin, First Principle-Based Simulation of Ethane Steam Cracking, *AIChE J.*, 57 (2011) 482-496.
- [3] C.K. Westbrook, W.J. Pitz, O. Herbinet, H.J. Curran, E.J. Silke, A comprehensive detailed chemical kinetic reaction mechanism for combustion of n-alkane hydrocarbons from n-octane to n-hexadecane, *Combustion and Flame*, 156 (2009) 181-199.
- [4] C.S. Henry, L.J. Broadbelt, V. Hatzimanikatis, Thermodynamics-based metabolic flux analysis, *Biophysical Journal*, 92 (2007) 1792-1805.
- [5] S. Klamt, J. Stelling, Two approaches for metabolic pathway analysis?, *Trends in Biotechnology*, 21 (2003) 64-69.
- [6] G.B. Marin, G.S. Yablonsky, *Kinetics of Chemical Reactions*, Wiley-VCH, Weinheim, 2011.
- [7] T.F. Lu, C.K. Law, Toward accommodating realistic fuel chemistry in large-scale computations, *Progress in Energy and Combustion Science*, 35 (2009) 192-215.
- [8] J. Gasteiger, *Handbook of Chemoinformatics*, Wiley-VCH Verlag GmbH & Co., Weinheim, 2003.
- [9] M. Levitt, The birth of computational structural biology, *Nature Structural Biology*, 8 (2001) 392-393.
- [10] D.W. Mount, *Sequence and genome analysis*, *Bioinformatics: Cold Spring Harbour Laboratory Press: Cold Spring Harbour*, 2 (2004).
- [11] T.I. Oprea, *Chemoinformatics and the quest for leads in drug discovery*, *Handbook of Chemoinformatics: From Data to Knowledge in 4 Volumes*, (2003) 1508-1531.
- [12] A. Gaulton, L.J. Bellis, A.P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani, ChEMBL: a large-scale bioactivity database for drug discovery, *Nucleic acids research*, 40 (2012) D1100-D1107.
- [13] T. Barrett, D.B. Troup, S.E. Wilhite, P. Ledoux, D. Rudnev, C. Evangelista, I.F. Kim, A. Soboleva, M. Tomashevsky, R. Edgar, NCBI GEO: mining tens of millions of expression profiles--database and tools update, *Nucleic Acids Res*, 35 (2007) D760-765.

- [14] R.H. West, G.R. Magoon, Letter to the editor: Electronic data formats for open storage of published chemical kinetic data, *International Journal of Chemical Kinetics*, 42 (2010) 200-201.
- [15] F. Battin-Leclerc, E. Blurock, R. Bounaceur, R. Fournet, P.A. Glaude, O. Herbinet, B. Sirjean, V. Warth, Towards cleaner combustion engines through groundbreaking detailed chemical kinetic models, *Chemical Society Reviews*, 40 (2011) 4762-4782.
- [16] L.J. Broadbelt, J. Pfaendtner, Lexicography of kinetic modeling of complex reaction networks, *AIChE Journal*, 51 (2005) 2112-2121.
- [17] R.J. Quann, S.B. Jaffe, Building useful models of complex reaction systems in petroleum refining, *Chemical Engineering Science*, 51 (1996) 1615-1631.
- [18] R.J. Quann, S.B. Jaffe, Structure-oriented lumping - Describing the chemistry of complex hydrocarbon mixtures, *Industrial & Engineering Chemistry Research*, 31 (1992) 2483-2497.
- [19] S.B. Jaffe, H. Freund, W.N. Olmstead, Extension of structure-oriented lumping to vacuum residua, *Industrial & Engineering Chemistry Research*, 44 (2005) 9840-9852.
- [20] I. Ugi, J. Bauer, J. Brandt, J. Friedrich, J. Gasteiger, C. Jochum, W. Schubert, New applications of computers in chemistry, *Angew. Chem.-Int. Edit. Engl.*, 18 (1979) 111-123.
- [21] P.J. Clymans, G.F. Froment, Computer-generation of reaction paths and rate equations in the thermal cracking of normal and branched paraffins, *Computers & Chemical Engineering*, 8 (1984) 137-142.
- [22] M. Dente, E. Ranzi, A.G. Goossens, Detailed prediction of olefin yields from hydrocarbon pyrolysis through a fundamental simulation model (SPYRO), *Computers & Chemical Engineering*, 3 (1979) 61-75.
- [23] G.G. Martens, G.B. Marin, Kinetics for hydrocracking based on structural classes: Model development and application, *AIChE Journal*, 47 (2001) 1607-1622.
- [24] S. Wauters, G.B. Marin, Computer generation of a network of elementary steps for coke formation during the thermal cracking of hydrocarbons, *Chemical Engineering Journal*, 82 (2001) 267-279.
- [25] Y. Yoneda, Chemograph, a computer-program package for chemical logic. GRACE. 1. Generation of elementary reaction network in radical reactions. 2. Computer-program package for the analysis, creation, and estimation of generalized reactions, *Bull. Chem. Soc. Jpn.*, 52 (1979) 8-14.
- [26] W. Feng, E. Vynckier, G.F. Froment, Single-event kinetics of catalytic cracking, *Industrial & Engineering Chemistry Research*, 32 (1993) 2997-3005.
- [27] G.D. Svoboda, E. Vynckier, B. Debrabandere, G.F. Froment, Single-event rate parameters for paraffin hydrocracking oil on a Pt/US-Y zeolite, *Industrial & Engineering Chemistry Research*, 34 (1995) 3793-3800.
- [28] F.P. Dimaio, P.G. Lignola, KING, a kinetic network generator, *Chemical Engineering Science*, 47 (1992) 2713-2718.
- [29] M.A. Baltanas, G.F. Froment, Computer-Generation of Reaction Networks and Calculation of Product Distributions in the Hydroisomerization and Hydrocracking of Paraffins on Pt-Containing Bifunctional Catalysts, *Computers & Chemical Engineering*, 9 (1985) 71-81.
- [30] C. Steinbeck, Y. Han, S. Kuhn, O. Horlacher, E. Luttmann, E. Willighagen, The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics, *J Chem Inf Comput Sci*, 43 (2003) 493 - 500.
- [31] N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, Open Babel: An open chemical toolbox, *Journal of Cheminformatics*, 3 (2011).
- [32] G. Landrum, RDKit: Open-source cheminformatics; <http://rdkit.sourceforge.net>, (2006).
- [33] I. Ugi, J. Bauer, K. Bley, A. Dengler, A. Dietz, E. Fontain, B. Gruber, R. Herges, M. Knauer, K. Reitsam, N. Stein, Computer-assisted solution of chemical problems - the historical development and the present state-of-the-art of a new discipline of chemistry, *Angew. Chem.-Int. Edit. Engl.*, 32 (1993) 201-227.
- [34] L.J. Broadbelt, S.M. Stark, M.T. Klein, Computer-generated pyrolysis modeling on-the-fly generation of species, reactions, and rates, *Industrial & Engineering Chemistry Research*, 33 (1994) 790-799.
- [35] L.P. Hillewaert, J.L. Dierickx, G.F. Froment, Computer-generation of reaction schemes and rate equations for thermal cracking, *AIChE Journal*, 34 (1988) 17-24.
- [36] A. Ratkiewicz, T.N. Truong, Application of chemical graph theory for automated mechanism generation, *Journal of Chemical Information and Computer Sciences*, 43 (2003) 36-44.
- [37] J.R. Ullmann, Algorithm for subgraph isomorphism, *Journal of the ACM*, 23 (1976) 31-42.
- [38] S.E. Prickett, M.L. Mavrovouniotis, Construction of complex reaction systems .1. Reaction description language, *Computers & Chemical Engineering*, 21 (1997) 1219-1235.
- [39] S.-H. Hsu, B. Krishnamurthy, P. Rao, C. Zhao, S. Jagannathan, V. Venkatasubramanian, A domain-specific compiler theory based framework for automated reaction network generation, *Computers & Chemical Engineering*, 32 (2008) 2455-2470.

- [40] S.E. Prickett, M.L. Mavrovouniotis, Construction of complex reaction systems .2. Molecule manipulation and reaction application algorithms, *Computers & Chemical Engineering*, 21 (1997) 1237-1254.
- [41] S.E. Prickett, M.L. Mavrovouniotis, Construction of complex reaction systems - III. An example: alkylation of olefins, *Computers & Chemical Engineering*, 21 (1997) 1325-1337.
- [42] S. Rangarajan, A. Bhan, P. Daoutidis, Rule-Based Generation of Thermochemical Routes to Biomass Conversion, *Industrial & Engineering Chemistry Research*, 49 (2010) 10459-10470.
- [43] S. Rangarajan, A. Bhan, P. Daoutidis, Language-oriented rule-based reaction network generation and analysis: Applications of RING, *Computers & Chemical Engineering*, 46 (2012) 141-152.
- [44] S. Rangarajan, A. Bhan, P. Daoutidis, Language-oriented rule-based reaction network generation and analysis: Description of RING, *Computers & Chemical Engineering*, 45 (2012) 114-123.
- [45] W.H. Green, J.W. Allen, R.W. Ashcraft, G.J. Beran, C.A. Class, C. Gao, C.F. Goldsmith, M.R. Harper, A. Jalan, G.R. Magoon, D.M. Matheu, S.S. Merchant, J.D. Mo, S. Petway, S. Raman, S. Sharma, J. Song, K.M. Van Geem, J. Wen, R.H. West , A. Wong, H.S. Wong, P.E. Yelvington, J. Yu, RMG - Reaction Mechanism Generator v4.0, (2013).
- [46] E. Van Wyk, D. Bodin, J. Gao, L. Krishnan, Silver: An extensible attribute grammar system, *Science of Computer Programming*, 75 (2010) 39-54.
- [47] A.S. Tomlin, M.J. Pilling, J.H. Merkin, J. Brindley, N. Burgess, A. Gough, Reduced mechanisms for propane pyrolysis, *Industrial & Engineering Chemistry Research*, 34 (1995) 3749-3760.
- [48] A.S. Tomlin, M.J. Pilling, T. Turanyi, J.H. Merkin, J. Brindley, Mechanism reduction for the oscillatory oxidation of hydrogen - sensitivity and quasi-steady-state analyses, *Combustion and Flame*, 91 (1992) 107-130.
- [49] L.E. Whitehouse, A.S. Tomlin, M.J. Pilling, Systematic reduction of complex tropospheric chemical mechanisms, Part I: sensitivity and time-scale analyses, *Atmospheric Chemistry and Physics*, 4 (2004) 2025-2056.
- [50] S. Pierucci, E. Ranzi, A review of features in current automatic generation software for hydrocarbon oxidation mechanisms, *Computers & Chemical Engineering*, 32 (2008) 805-826.
- [51] E. Ranzi, M. Dente, A. Goldaniga, G. Bozzano, T. Faravelli, Lumping procedures in detailed kinetic modeling of gasification, pyrolysis, partial oxidation and combustion of hydrocarbon mixtures, *Progress in Energy and Combustion Science*, 27 (2001) 99-139.
- [52] K.J. Laidler, *Chemical Kinetics*, 3rd ed., Harper & Row, New York, 1987.
- [53] H.J. Curran, P. Gaffuri, W.J. Pitz, C.K. Westbrook, A comprehensive modeling study of iso-octane oxidation, *Combustion and Flame*, 129 (2002) 253-280.
- [54] H.J. Curran, P. Gaffuri, W.J. Pitz, C.K. Westbrook, A comprehensive modeling study of n-heptane oxidation, *Combustion and Flame*, 114 (1998) 149-177.
- [55] E.S. Blurock, Reaction: System for Modeling Chemical Reactions, *Journal of Chemical Information and Computer Sciences*, 35 (1995) 607-616.
- [56] M.K. Sabbe, K.M. Van Geem, M.F. Reyniers, G.B. Marin, First Principle-Based Simulation of Ethane Steam Cracking, *AIChE Journal*, 57 (2011) 482-496.
- [57] L.J. Broadbelt, S.M. Stark, M.T. Klein, Termination of computer-generated reaction-mechanisms - species rank - based convergence criterion, *Industrial & Engineering Chemistry Research*, 34 (1995) 2566-2573.
- [58] I. Nemes, T. Vidoczy, L. Botar, D. Gal, Possible construction of a complex chemical reaction network. 1. Definitions and procedure for construction, *Theor. Chim. Acta*, 45 (1977) 215-223.
- [59] I. Nemes, T. Vidoczy, L. Botar, D. Gal, Possible construction of a complex chemical reaction network. 2. Applications, *Theor. Chim. Acta*, 45 (1977) 225-233.
- [60] R.G. Susnow, A.M. Dean, W.H. Green, P. Peczak, L.J. Broadbelt, Rate-based construction of kinetic models for complex systems, *Journal of Physical Chemistry A*, 101 (1997) 3731-3740.
- [61] R.H. West , W.H. Green, Reaction Mechanism Generator: Chemoinformatics in Kinetic Modeling, in: ACS Meeting, Boston, 2010.
- [62] J.M. Grenda, J.W. Bozzelli, A.M. Deant, Automated methods of treating chemically activated reactions in kinetics mechanism generation., Eighth international conference on numerical combustion, CP4, SIAM. (2000).
- [63] J.A. Montgomery, M.J. Frisch, J.W. Ochterski, G.A. Petersson, A complete basis set model chemistry. VI. Use of density functional geometries and frequencies, *J. Chem. Phys.*, 110 (1999) 2822-2827.
- [64] V. Van Speybroeck, R. Gani, R.J. Meier, The calculation of thermodynamic properties of molecules, *Chem Soc Rev*, 39 (2010) 1764-1779.
- [65] S.W. Benson, *Thermochemical Kinetics*, John Wiley & Sons, New York, 1976.
- [66] S.W. Benson, N. Cohen, Current Status of Group Additivity, in: *Computational Thermochemistry*, American Chemical Society, 1998, pp. 20-46.

- [67] T.H. Lay, J.W. Bozzelli, A.M. Dean, E.R. Ritter, Hydrogen-atom bond increments for calculation of thermodynamic properties of hydrocarbon radical species, *Journal of Physical Chemistry*, 99 (1995) 14514-14527.
- [68] E.R. Ritter, J.W. Bozzelli, THERM - Thermodynamic property estimation for gas-phase radicals and molecules, *International Journal of Chemical Kinetics*, 23 (1991) 767-778.
- [69] E.S. Blurock, V. Warth, X. Grandmougin, R. Bounaceur, P.-A. Glaude, F. Battin-Leclerc, JTHERGAS: Thermodynamic estimation from 2D graphical representations of molecules, *Energy*, 43 (2012) 161-171.
- [70] C. Muller, V. Michel, G. Scacchi, G.M. Côme, 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*, 92 (1995) 1154-1178.
- [71] M.K. Sabbe, F. De Vleeschouwer, M.F. Reyniers, M. Waroquier, G.B. Marin, 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*, 112 (2008) 12235-12251.
- [72] M.K. Sabbe, M. Saeys, M.F. Reyniers, G.B. Marin, V. Van Speybroeck, M. Waroquier, Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals, *Journal of Physical Chemistry A*, 109 (2005) 7466-7480.
- [73] S.S. Khan, X. Yu, J.R. Wade, R.D. Malmgren, L.J. Broadbelt, Thermochemistry of Radicals and Molecules Relevant to Atmospheric Chemistry: Determination of Group Additivity Values using G3//B3LYP Theory, *Journal of Physical Chemistry A*, 113 (2009) 5176-5194.
- [74] P.D. Paraskevas, M.K. Sabbe, M.-F. Reyniers, N. Papayannakos, G.B. Marin, Group Additive Values for the Gas-Phase Standard Enthalpy of Formation, Entropy and Heat Capacity of Oxygenates, *Chemistry – A European Journal*, 19 (2013) 16431-16452.
- [75] A.G. Vandeputte, M.K. Sabbe, M.-F. Reyniers, G.B. Marin, Modeling the Gas-Phase Thermochemistry of Organosulfur Compounds, *Chemistry-a European Journal*, 17 (2011) 7656-7673.
- [76] O.V. Dorofeeva, O.N. Ryzhova, N.F. Moiseeva, The quantum-chemical determination of group contributions to the thermodynamic properties of organophosphorus compounds, *Russian Journal of Physical Chemistry A*, 82 (2008) 933-937.
- [77] R.W. Ashcraft, W.H. Green, Thermochemical properties and group values for nitrogen-containing molecules, *J Phys Chem A*, 112 (2008) 9144-9152.
- [78] L.J. Broadbelt, S.M. Stark, M.T. Klein, Computer generated reaction networks: On-the-fly calculation of species properties using computational quantum chemistry, *Chemical Engineering Science*, 49 (1994) 4991-5010.
- [79] G.R. Magoon, W.H. Green, 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*, 52 (2013) 35-45.
- [80] J.J. Stewart, Optimization of parameters for semiempirical methods I. Method, *J. Comput. Chem.*, 10 (1989) 209-220.
- [81] J.J. Stewart, Optimization of parameters for semiempirical methods II. Applications, *J. Comput. Chem.*, 10 (1989) 221-264.
- [82] N. Nevins, K.S. Chen, N.L. Allinger, Molecular mechanics (MM4) calculations on alkenes, *J. Comput. Chem.*, 17 (1996) 669-694.
- [83] A. Osmont, I. Gokalp, L. Catoire, Evaluating missile fuels, *Propellants Explosives Pyrotechnics*, 31 (2006) 343-354.
- [84] M.G. Evans, M. Polanyi, Inertia and driving force of chemical reactions, *Transactions of the Faraday Society*, 34 (1938) 0011-0023.
- [85] P.A. Willems, G.F. Froment, Kinetic modeling of thermal-cracking of hydrocarbons. 1. Calculation of frequency factors, *Industrial & Engineering Chemistry Research*, 27 (1988) 1959-1966.
- [86] P.A. Willems, G.F. Froment, Kinetic modeling of thermal-cracking of hydrocarbons.2. Calculation of activation-energies, *Industrial & Engineering Chemistry Research*, 27 (1988) 1966-1971.
- [87] R. Sumathi, W.H. Green, A priori rate constants for kinetic modeling, *Theoretical Chemistry Accounts*, 108 (2002) 187-213.
- [88] R. Sumathi, H.H. Carstensen, W.H. Green, Reaction rate predictions via group additivity. Part 3: Effect of substituents with CH₂ as the mediator, *Journal of Physical Chemistry A*, 106 (2002) 5474-5489.
- [89] R. Sumathi, H.H. Carstensen, W.H. Green, 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*, 105 (2001) 8969-8984.

- [90] R. Sumathi, H.H. Carstensen, W.H. Green, Reaction rate prediction via group additivity Part 1: H abstraction from alkanes by H and CH₃, *Journal of Physical Chemistry A*, 105 (2001) 6910-6925.
- [91] M. Saeys, M.F. Reyniers, G.B. Marin, V. Van Speybroeck, M. Waroquier, Ab initio group contribution method for activation energies for radical additions, *AIChE Journal*, 50 (2004) 426-444.
- [92] M. Saeys, M.F. Reyniers, V. Van Speybroeck, M. Waroquier, G.B. Marin, Ab initio group contribution method for activation energies of hydrogen abstraction reactions, *Chemphyschem*, 7 (2006) 188-199.
- [93] M.K. Sabbe, M.F. Reyniers, V. Van Speybroeck, M. Waroquier, G.B. Marin, Carbon-centered radical addition and beta-scission reactions: Modeling of activation energies and pre-exponential factors, *Chemphyschem*, 9 (2008) 124-140.
- [94] M.K. Sabbe, A.G. Vandeputte, M.-F. Reyniers, M. Waroquier, G.B. Marin, Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions, *Phys. Chem. Chem. Phys.*, 12 (2010) 1278-1298.
- [95] M.K. Sabbe, M.-F. Reyniers, M. Waroquier, G.B. Marin, Hydrogen Radical Additions to Unsaturated Hydrocarbons and the Reverse beta-Scission Reactions: Modeling of Activation Energies and Pre-Exponential Factors, *Chemphyschem*, 11 (2010) 195-210.
- [96] M.K. Sabbe, M.-F. Reyniers, G.B. Marin, H-abstraction reactions by a hydrogen radical from hydrocarbons, Personal Communication.
- [97] A.G. Vandeputte, M.K. Sabbe, M.-F. Reyniers, G.B. Marin, Kinetics of alpha hydrogen abstractions from thiols, sulfides and thiocarbonyl compounds, *Phys. Chem. Chem. Phys.*, 14 (2012) 12773-12793.
- [98] A.G. Vandeputte, M.-F. Reyniers, G.B. Marin, Kinetics of Homolytic Substitutions by Hydrogen Atoms at Thiols and Sulfides, *Chemphyschem*, 14 (2013) 1703-1722.
- [99] A.G. Vandeputte, M.-F. Reyniers, G.B. Marin, Kinetic Modeling of Hydrogen Abstractions Involving Sulfur Radicals, *Chemphyschem*, 14 (2013) 3751-3771.
- [100] M.K. Sabbe, Ab Initio Based Kinetic Modeling for the Simulation of Industrial Chemical Processes, *Chemische Proceskunde en Technische Chemie*, Gent University (2009).
- [101] M.K. Sabbe, A.G. Vandeputte, M.F. Reyniers, M. Waroquier, G.B. Marin, Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions, *Phys. Chem. Chem. Phys.*, 12 (2010) 1278-1298.
- [102] J. Song, Building Robust Chemical Reaction Mechanisms: Next Generation of Automatic Model Construction Software. Ph.D thesis, MIT (2004).
- [103] J. Song, S. Raman, J. Yu, C.D. Wijaya, G. Stephanopoulos, W.H. Green, Development of automatic chemical reaction mechanism generation software using object-oriented technology., *Abstracts of Papers of the American Chemical Society*, 226 (2003) U530-U531.
- [104] G. Mills, H. Jonsson, G.K. Schenter, Reversible work transition state theory - Application to dissociative adsorption of hydrogen, *Surface Science*, 324 (1995) 305-337.
- [105] B. Peters, A. Heyden, A.T. Bell, A. Chakraborty, A growing string method for determining transition states: Comparison to the nudged elastic band and string methods, *J. Chem. Phys.*, 120 (2004) 7877-7886.
- [106] P.L. Bhoorasingh, R.H. West, Automatic Transition State Searches for On-The-Fly Kinetic Calculations, *AIChE Annual Meeting*, (2013).
- [107] J.A. Bondy, U.S.R. Murty, *Graph theory with applications*, Macmillan London, 1976.
- [108] L.T. Fan, B. Bertok, F. Friedler, A graph-theoretic method to identify candidate mechanisms for deriving the rate law of a catalytic reaction, *Computers & Chemistry*, 26 (2002) 265-292.
- [109] H.L. Morgan, The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service, *Journal of Chemical Documentation*, 5 (1965) 107-113.
- [110] J.-L. Faulon, A. Bender, *Handbook of chemoinformatics algorithms*, Chapman & Hall/CRC, 2010.
- [111] B.D. McKay, Practical Graph Isomorphism, *Congressus Numerantium*, 30 (1981) 45-87.
- [112] D. Weininger, A. Weininger, J.L. Weininger, SMILES .2. Algorithm for generation of unique SMILES notation, *Journal of Chemical Information and Computer Sciences*, 29 (1989) 97-101.
- [113] J.L. Faulon, M.J. Collins, R.D. Carr, The signature molecular descriptor. 4. Canonizing molecules using extended valence sequences, *Journal of Chemical Information and Computer Sciences*, 44 (2004) 427-436.
- [114] S. Heller, S. Stein, IUPAC InChI, Google Tech Talk, (2006).
- [115] S.R. Heller, S.E. Stein, D.V. Tchekhovskoi, InChI: Open access/open source and the IUPAC international chemical identifier, *Abstracts of Papers of the American Chemical Society*, 230 (2005) 60-CINF.
- [116] D. Weininger, SMILES, a chemical language and information-system. 1. Introduction to methodology and encoding rules, *Journal of Chemical Information and Computer Sciences*, 28 (1988) 31-36.
- [117] M.R. Garey, D.S. Johnson, *Computers and intractability. A guide to the theory of NP-completeness*, *Computers and intractability. A guide to the theory of NP-completeness*, (1979) vii+338.

- [118] C. Bron, J. Kerbosch, Finding all cliques of an undirected graph H, *Communications of the Acm*, 16 (1973) 575-577.
- [119] E. Balas, C.S. Yu, Finding a Maximum Clique in an Arbitrary Graph, *Siam Journal on Computing*, 15 (1986) 1054-1068.
- [120] R. Carraghan, P.M. Pardalos, An exact algorithm for the maximum clique problem, *Operations Research Letters*, 9 (1990) 375-382.
- [121] P.M. Pardalos, J. Xue, The maximum clique problem, *Journal of Global Optimization*, 4 (1994) 301-328.
- [122] M. Shindo, E. Tomita, A simple algorithm for finding a maximum clique and its worst-case time complexity, *Systems and Computers in Japan*, 21 (1990) 1-13.
- [123] C. Tonnelier, P. Jauffret, T. Hanser, G. Kaufmann, Machine learning of generic reactions: 3. an efficient algorithm for maximal common substructure determination, *Tetrahedron Computer Methodology*, 3 (1990) 351-358.
- [124] L.P. Cordella, P. Foggia, C. Sansone, M. Vento, A (sub)graph isomorphism algorithm for matching large graphs, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26 (2004) 1367-1372.
- [125] L.P. Cordella, P. Foggia, C. Sansone, M. Vento, An Improved Algorithm for Matching Large Graphs, (2008) 8.
- [126] C. Bron, J. Kerbosch, Finding all the cliques in an undirected graph, *Communication of the ACM*, 16 (1973) 189 - 201.
- [127] S. Rahman, M. Bashton, G. Holliday, R. Schrader, J. Thornton, Small Molecule Subgraph Detector (SMSD) toolkit, *Journal of Cheminformatics*, 1 (2009) 12.
- [128] M. Rijnbeek, C. Steinbeck, OrChem - An open source chemistry search engine for Oracle(R), *Journal of Cheminformatics*, 1 (2009) 17.
- [129] Daylight Chemical Information Systems, Daylight Theory Manual, 2010 <http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>.
- [130] W.T. Wipke, S. Krishnan, G.I. Ouchi, Hash Functions for Rapid Storage and Retrieval of Chemical Structures, *Journal of Chemical Information and Computer Sciences*, 18 (1978) 32-37.
- [131] W.D. Ihlenfeldt, J. Gasteiger, Hash Codes for the Identification and Classification of Molecular-Structure Elements, *J. Comput. Chem.*, 15 (1994) 793-813.
- [132] E.L. Eliel, S.H. Wilen, *Stereochemistry of organic compounds*, Wiley-VCH, Weinheim, 2008.
- [133] J. Sadowski, J. Gasteiger, From atoms and bonds to 3-dimensional atomic coordinates - Automatic model builders, *Chemical Reviews*, 93 (1993) 2567-2581.
- [134] V. Kuz'Min, I. Stel'Makh, M. Bekker, D. Pozigun, Quantitative aspects of chirality. I. Method of dissymmetry function, *Journal of physical organic chemistry*, 5 (1992) 295-298.
- [135] E. Eliel, S. Wilen, L. Mander, *Stereochemistry of organic Compounds*, *Stereochemistry of Organic Compounds*, (1994).
- [136] R.S. Cahn, C. Ingold, V. Prelog, Specification of Molecular Chirality, *Angewandte Chemie International Edition in English*, 5 (1966) 385-415.
- [137] W.T. Wipke, T.M. Dyott, Stereochemically Unique Naming Algorithm, *J. Am. Chem. Soc.*, 96 (1974) 4834-4842.

Chapter 2: Thermal rearrangement of 2-pinanol

This chapter includes the following paper:

Vandewiele, N. M.; Van Geem, K. M.; Reyniers, M.-F.; Marin, G. B., “Kinetic study of the thermal rearrangement of *cis*- and *trans*-2-pinanol.” *Journal of Analytical and Applied Pyrolysis* 2011, 90 (2), 187-196.

2.1 Abstract

Pyrolysis of *cis*- and *trans*-2-pinanol was performed in a temperature range from 753 K to 843 K and conversions ranging from 10 to 96%. Kinetic models for the pyrolysis of both diastereomers are presented. Good agreement between simulated and experimental data was achieved in all cases. The reactivity and selectivity differences between the two reactant diastereomers was attributed to a difference in activation energies of the carbon-carbon scission reaction that opens the cyclobutane ring of the reactant molecules and forms the 1,4-disubstituted cyclohexanol biradical: 206 kJ mol⁻¹ in the case of *cis*-2-pinanol and 211 kJ mol⁻¹ for *trans*-2-pinanol. The high selectivity for the linalool reaction channel compared to the parallel pathway to isolinalool was attributed to the large differences in activation energies, 19 kJ mol⁻¹ for *cis*-2-pinanol and 14 kJ mol⁻¹ for *trans*-2-pinanol, for the cyclobutane rupture reactions that form two types of biradicals, a 1,2-disubstituted and a 1,4-disubstituted cyclohexanol. All other Arrhenius parameters of the reactions in the kinetic models of *cis*- and *trans*-2-pinanol were not found to be significantly different for both diastereomers.

Keywords: linalool, 2-pinanol, thermal rearrangement, terpenes, kinetic model

2.2 Introduction

Terpenes are a diverse class of isoprene oligomers and have proven to be valuable fine chemicals used in the fragrance, flavor and pharmaceutical industry [1-6] and could even be applied as diesel fuel additive [7]. Also, the atmospheric emissions of terpenes and terpenoids as biogenic volatile organic components are recently gaining attention in the scientific community [8]. The gas-phase rearrangement of *cis*- and *trans*-2-pinanol is an important production route for (+)-linalool and (-)-linalool ((S)- and (R)-3,7-dimethylocta-1,6-dien-3-ol) [9] (cf. Figure 2-1), used in perfumes and as a precursor for vitamin A and E [4]. This thermal isomerization is industrially relevant due to the high linalool selectivity, up to 90% [10]. The main undesired side-products of this thermal rearrangement is pinol [11]. Besides pinol, Ohloff et al. [9] identified β -terpineol (< 3%) and a methyl-ketone (<5%) as byproducts of the pyrolysis of *cis*- and *trans*-2-pinanol at 873K. The structure of the methyl-ketone was later elucidated as 5,7-dimethyloct-6-ene-2-one by

Coxon et al. [12] Erman et al. [13] reported isolinalool as a minor product during the pyrolysis of pinanols. Chemical structures of the aforementioned components are depicted in Figure 2-1.

2-Pinanol can be synthesized from α -pinene and β -pinene [14] which are found in selected coniferous biomass [15] or in crude sulfate turpentine (CST) [16], a by-product of the paper industry. The catalytic hydrogenation of α -pinene yields mixtures of the diastereomers *cis*-pinane and *trans*-pinane with high stereoselectivity to the *cis*-hydrogenation product [17]. The subsequent oxidation of both pinanes with air or oxygen produces 2-pinane hydroperoxide [18] with *cis*-pinane being the more reactive diastereomers [19]. 2-Pinanol isomers are obtained industrially by catalytic hydrogenation of the respective 2-pinane hydroperoxydes [18, 20-22].

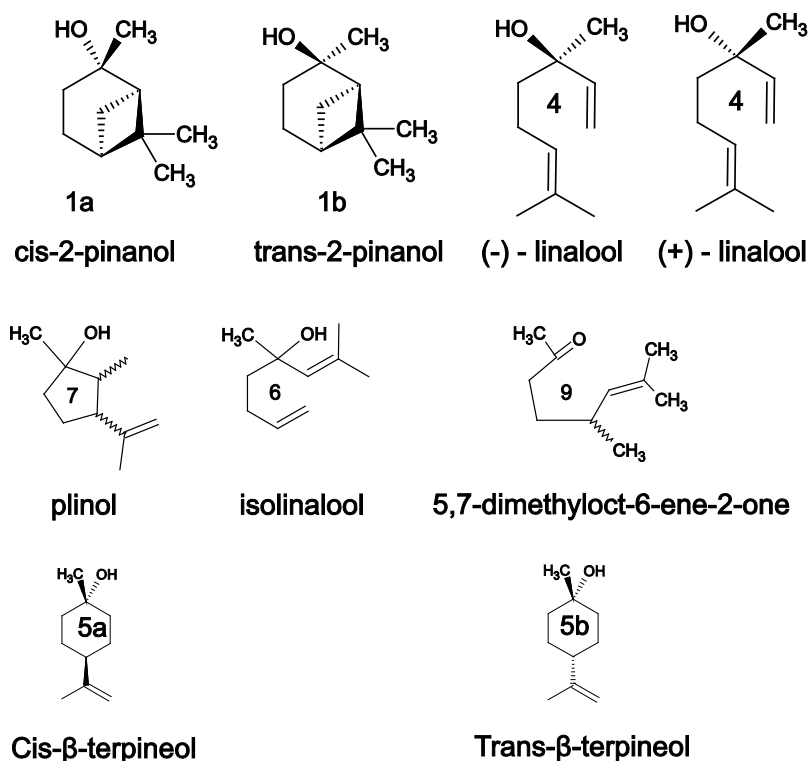


Figure 2-1: Chemical structures of molecular components involved in the thermal rearrangement of 2-pinanol (the corresponding number used in the current chapter is mentioned)

The reaction mechanism for the thermal isomerization of pinane- and pinene-type components has been extensively discussed in literature and it is generally agreed upon that the fragmentation of the four-membered ring is the initial step in the thermal isomerization of bicyclic

monoterpenes consisting of a bicyclo[3.1.1]heptane system [23-33]. Two types of cyclobutane rupture mechanisms were proposed to explain the formation of the different primary acyclic and monocyclic pyrolysis products: a concerted mechanism of the cyclobutane fragmentation containing a 4- π -electron anti-aromatic transition state and a stepwise mechanism with biradical intermediates. One of the main objections against the concerted mechanism is that, according to the Woodward-Hoffmann rules, this retro-[2+2]-cycloaddition is a thermally forbidden reaction [34, 35]. The stepwise mechanism first proposed by Burwell in 1952 is therefore the prevailing and generally accepted mechanism [36-38].

Kinetic data for thermal rearrangements of bicyclic monoterpenes and more particular for 2-pinanol isomerization are scarce. Stolle et al. investigated the mechanism and kinetics of the gas-phase isomerization of monoterpenes [23, 28] and showed that there are mechanistic analogies in the isomerization of pinane- and pinene-type structures. Moreover, the authors were able to model reactivity differences between *cis*-pinane and *trans*-pinane. Dervan et al. [39, 40] examined the kinetics of butadiyl biradical rearrangements and concluded that their further cleavage to ethylene occurs faster than their recombination to cyclic structures. Gajewski et al. [41] investigated the pyrolysis of deuterium labeled α -pinene and concluded that the monocyclic racemic limonenes were formed through a [1,5]-H-shift of the biradical intermediate. Semikolenov et al. [10, 18, 42] studied the isomerization kinetics of mixtures of *cis*- and *trans*-2-pinanol in a fixed-bed reactor type with a heat conducting carbon monolith catalyst and reported a higher reactivity for the *cis*-isomer relative to the *trans*-isomer. The addition of small amounts of pyridine (< 2 mol%) significantly increased the linalool selectivity. Buddoo et al. [43] studied the *cis*-2-pinanol pyrolysis kinetics in a multichannel microreactor system and concluded that these systems offer advantages in 2-pinanol conversion and linalool selectivity over conventional tubular reactors. Pickenhagen et al. [44] investigated the kinetics of pericyclic ene-cyclization reactions with linalyltrimethylsilyl ether, which is analogous to linalool, and reported the formation of plinol-type products, 1,2-dimethyl-3-isopropenyl-1-cyclopentyl trimethylsilyl ethers. It was shown that the 2,3-*cis* products, with the C(2)-methyl- and isopropenyl-group in *cis*-position, were favored over the 2,3-*trans* products. The preference for cyclic products with *cis*-oriented substituents was confirmed by Roy et al. [45] using quantum-chemical calculations.

Although kinetic studies of 2-pinanol thermal isomerizations are available in literature [10, 18, 26, 42, 43], none of them has addressed this topic in terms of elementary reactions i.e. reactions in which a single transition state connects the reactants to the products of the reaction. This implies that the cyclobutane fragmentation is treated as a stepwise mechanism, with two distinct biradical species as intermediates originating from two parallel cyclobutane fragmentation routes. The present chapter presents thermodynamically consistent kinetic models involving only elementary reactions for the thermal rearrangement of *cis*- and *trans*-2-pinanol. Thermochemical properties of the reactive species are estimated using Benson's group additivity method [46]. Pre-exponential Arrhenius factors were retrieved from literature while activation energies were fitted based on regression of experimental data. The kinetic model presented in this work also enables to describe the rate of formation of the products thereby allowing to interpret reactivity and selectivity differences between *cis*- and *trans*-2-pinanol in terms of the kinetics of the elementary steps..

2.3 Experimental Procedures

2.3.1 Materials

Both *cis*-2-pinanol (97.039%) and *trans*-2-pinanol (98.362%) diastereomers were provided by Millenium Specialty Chemicals, formerly known as Glidco Organics and were used without further purification. The enantiomeric purity of the reactant isomers is unknown.

2.3.2 Experimental reactor setup

Experimental studies of the pyrolysis of the 2-pinanol isomers were conducted in the bench-scale set-up of the Laboratory for Chemical Technology (LCT) of Ghent University. A schematic of the experimental apparatus is shown in Figure 2-2. The reactor is a 1.475-m long, 6-mm internal diameter tube made of Incoloy 800HT (Ni, 30–35; Cr, 19 – 23; and Fe > 39.5 wt %). In line with previous work [47], wall surface effects were negligible.

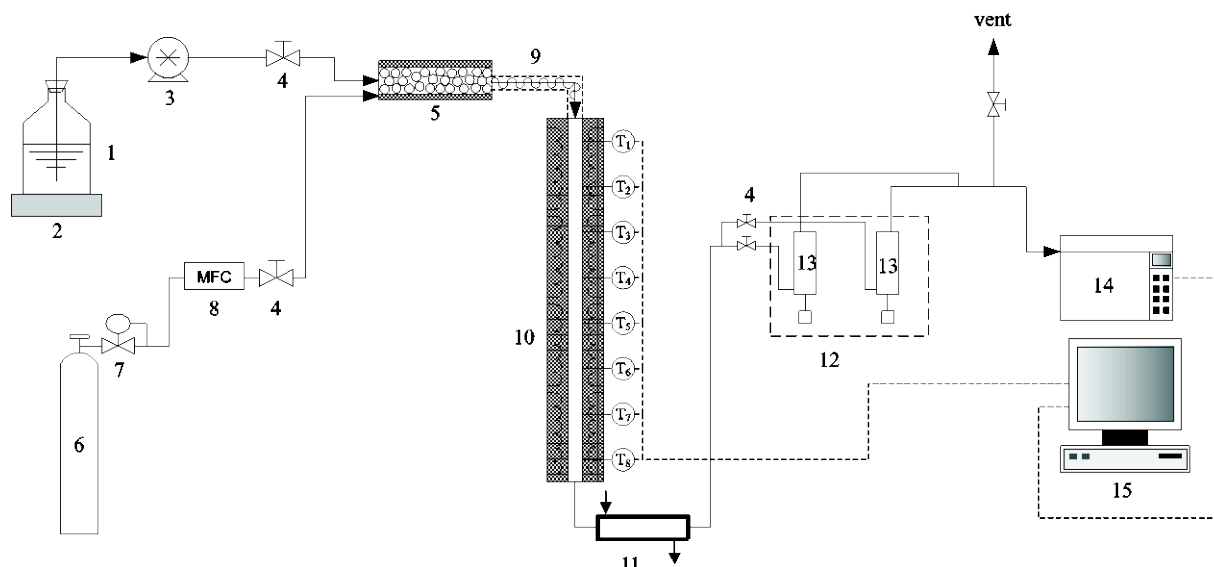


Figure 2-2: Schematic of the experimental pyrolysis set-up. 1: 2-pinanol vessel, 2: electronic balance, 3: pump, 4: valve, 5: evaporator/mixer, 6: methane, 7: pressure regulator, 8: mass flow controller, 9: heater, 10: reactor, 11: oil cooler, 12: water cooler, 13: condenser, 14: GC, 15: data acquisition

An overview of the experimental conditions of *cis*- and *trans*-2-pinanol experiments can be found in Table 2-1. A condenser is used to collect the condensed pyrolysis products.

Table 2-1: Range of experimental conditions used in the pyrolysis of *cis*- and *trans*-2-pinanol

	Cis-2-pinanol	Trans-2-pinanol
Temperature range (K)	753 – 843	753 – 838
Pressure levels (10^5 Pa)	1.15 / 4.4	1.15 / 4.4
2-pinanol mass flow rates (10^{-2} g s $^{-1}$)	1.56 / 2.97 / 5.28	1.80 / 3.47
Dilution levels (mol CH $_4$ / mol 2-pinanol)	1.8 / 8.5	1.8 / 8.5
Conversion range (%)	10 - 95	11 - 95

A total number of 35 experiments were carried out with *cis*-2-pinanol. The reactor outlet pressure was set at two distinct levels: $4.4 \cdot 10^5$ Pa and $1.15 \cdot 10^5$ Pa. Methane was selected as the diluent and aims to avoid excessive secondary, bimolecular reactions between 2-pinanol reactants. Nitrogen was also considered as a diluent; however, nitrogen cannot be detected by the used Packard GC-FID and thus makes an analysis on an absolute basis impossible. Dilution with methane was set at two levels: 0.18 and 0.89 kg methane per kg of feed, equivalent to 1.8 and 8.5 mol CH $_4$ per mol of 2-pinanol. Experiments can thus be classified into four different categories according to the

pressure and dilution. *Cis*-2-pinanol mass flow rates were set at three levels in all experiments: 0.0156 g s⁻¹, 0.0297 g s⁻¹ and 0.0528 g s⁻¹, corresponding to residence times of orders of magnitude of 10⁻¹ s to 10¹ s. Temperature levels of the experiments ranged from 753K to 843K. Per experimental condition, at least 6 samples were taken and analyzed by GC. Conversion of *cis*-2-pinanol varied from 9.6% to 95.59%. Dehydration reactions of *cis*-2-pinanol to pinenes and of β -terpineol to limonene were found to be negligible by GC analysis of the liquid fraction of the pyrolysis products. Mass balances attained 99% on average and were not normalized to 100%. Production of light components was 0.6 mol% on average.

A total number of 24 experiments were carried out with *trans*-2-pinanol. Similar to the experiments with *cis*-2-pinanol two different dilution levels (1.8 and 8.5 mol CH₄ per mol of hydrocarbons) and two pressure levels (4.4 10⁵ Pa and 1.15 10⁵ Pa) were set resulting in four categories of experiments. *Trans*-2-pinanol mass flow rates were set at two levels in all experiments: ca. 0.0180 g s⁻¹ and ca. 0.0347 g s⁻¹. *Trans*-2-pinanol conversion ranged from 11.6% to 95.3% and temperature levels varied from 753K to 838K. Dehydration reactions of *trans*-2-pinanol to pinenes and of β -terpineol to limonene were found to be negligible by GC analysis of the liquid fraction of the pyrolysis products. Mass balances attained 98% on average and were not normalized to 100%. Production of light components was 0.8 mol% on average.

The light components, identified as the components with molecular weight lower than 2-pinanol, are assumed to be the cracking products of linalool, *cis*- and *trans*-terpineol, isolinalool, plinol and component 9. The molecular weight of these light components is assumed to be half of the molecular weight of linalool, although there is no detailed identification of these components. The mole fractions of light components is less than 4 mol% based on the feed for all experiments with *cis*- and *trans*-2-pinanol.

For the analysis of the pyrolysis products a Packard model 724 gas chromatograph equipped with a flame ionization detector (FID) is used. A high performance capillary column Agilent Carbowax HP 20M (50 m \times 0.2 mm \times 0.2 μ m) is used for the separation of the products from *cis*-2-pinanol pyrolysis. For the separation of the products from *trans*-2-pinanol pyrolysis an Agilent Ultra 2 capillary column is added for the separation of *trans*-2-pinanol from linalool. The gas and

the liquid fraction of the reactor effluent are analyzed separately. The gas sample is taken directly from the reactor after the condenser and analyzed by means of an on-line sampling system.

2.3.3 Reactor Modeling

The tubular pyrolysis reactor was modeled using the Chemkin software package [32]. Plug-flow was assumed and radial temperature gradients were neglected, as was assumed previously by Harper et al. [47] and Chen and Froment [48]. The experimentally measured temperature profile was supplied, as was the pressure set by the backpressure controller. The default absolute ($1.0 \cdot 10^{-8}$) and relative ($1.0 \cdot 10^{-6}$) solver tolerances were sufficient for convergence.

2.3.4 Regression of experimental data

The activation energies of the forward reaction rates are estimated by fitting model predictions to experimental data. The parameter fitting procedure involved a weighted non-linear least-squares regression applying the Rosenbrock direct-search minimization algorithm [49]. The objective function $S(\bar{\beta})$ to be minimized is the sum of squared residuals R_{ij} over all n experiments and v response variables:

$$S(\bar{\beta}) = \sum_{i=1}^n \sum_{j=1}^v R_{ij}^2 \quad \text{Eq. 2-1}$$

The residuals are the weighted differences between the mass fractions of the simulated \hat{y}_{ij} values and the experimental values y_{ij} :

$$R_{ij} = \frac{\left(\hat{y}_{ij} - y_{ij} \right)}{\bar{y}_j}, \quad i = 1..n, j = 1..v \quad \text{Eq. 2-2}$$

Weights \bar{y}_j for the response variables are set equal to the arithmetic averages of each experimental response variable over the complete set of experiments:

$$\bar{y}_j = \frac{\sum_{i=1}^n y_{ij}}{n}, \quad j = 1..v \quad \text{Eq. 2-3}$$

The significance of the regression is assessed applying the F-test:

$$F_{calc} = \frac{\sum_{i=1}^n \sum_{j=1}^v \left(\frac{\hat{y}_{ij}}{\hat{y}_j} \right)^2 / p}{\sum_{i=1}^n \sum_{j=1}^v R_{ij}^2 / (n \cdot v - p)} \quad \text{Eq. 2-4}$$

With p the number of fitted parameters.

The calculated F-value is compared with the tabulated one for a probability level of $1-\alpha$. When $F_{calc} > F_{tab}(p, n \cdot v - p, 1-\alpha)$ the regression is considered to be significant. For all statistical tests the probability is chosen as 95%, i.e. $\alpha = 0.05$. The individual confidence interval of model parameter b_t delimits the range in which the real parameter β_t is located with a selected probability level of $1-\alpha$ and is given by:

$$b_t - t_{tab}(n - p, 1 - \alpha / 2) \sqrt{\hat{V}(b)_t} \leq \beta_t \leq b_t + t_{tab}(n - p, 1 - \alpha / 2) \sqrt{\hat{V}(b)_t} \quad \text{Eq. 2-5}$$

With b the vector of parameter estimates, t_{tab} the tabulated t-value for a probability level of $1 - \alpha$, and $\hat{V}(b)$ an estimate for the covariance matrix of the parameters.

2.4 Thermochemical Properties

It is opted to construct a microkinetic model considering both forward and reverse reactions of any elementary step between species in the model. Rather than verifying the thermodynamic consistency of the forward and reverse reaction rate coefficients *a posteriori*, the approach taken in the present chapter incorporates this principle *a priori* by deriving the reverse reaction rate coefficient from the forward reaction rate coefficient and the thermochemical properties of the reacting species using the following expression:

$$\frac{k_f}{k_r} = K_{eq} = e^{\frac{-\Delta_r G^\circ}{RT}} \cdot \left(\frac{RT}{P} \right)^{-\Delta n} \quad \text{Eq. 2-6}$$

With R the universal gas constant, T temperature, P pressure and Δn the change in the number of moles between products and reactants. Assuming an Arrhenius temperature dependency for the rate coefficients,

$$k = Ae^{\frac{E_a}{RT}} \quad \text{Eq. 2-7}$$

And $\Delta n = 0$ for unimolecular elementary steps (cf. Eq. 2-6), the Arrhenius parameters of the reverse reaction are calculated using Eq. 2-8 and Eq. 2-9 with:

$$E_a^r = E_a^f - \Delta_r H^\circ \quad \text{Eq. 2-8}$$

$$A^r = \frac{A^f}{e^{\frac{\Delta_r S^\circ}{R}}} \quad \text{Eq. 2-9}$$

Where E_a^f, E_a^r are the activation energies of the forward, reverse reaction, A^f, A^r the pre-exponential factors of the forward, reverse reaction and $\Delta_r H^\circ, \Delta_r S^\circ$ the reaction enthalpy and entropy of the forward reaction.

Arrhenius parameters of the forward reaction rate coefficients (E_a^f, A_a^f) are regressed to experimental data. The reaction enthalpy and entropy of the forward reaction ($\Delta_r H^\circ, \Delta_r S^\circ$) is calculated from the enthalpy of formation and the molar entropy of reactants and products, and therefore relies on the availability of thermochemical data for all reacting species.

The current approach has several advantages: first of all, reaction rate coefficients do not infringe thermodynamic consistency. Secondly, the number of Arrhenius parameters to be estimated is reduced since only the kinetics of the forward reaction need to be estimated. Thirdly, this approach prevents the insertion of enthalpic inconsistencies in the energy conservation equation that could translate into incorrect predictions of global variables for reactor design such as heat exchange, conversion and selectivity. Group additivity schemes like Benson's method provide a quick way of obtaining thermochemical properties of species that do not appear in databases [46]. Finally, fairly reliable estimations of thermochemical properties can be achieved through *ab initio* quantum chemistry calculations with ever diminishing computational costs [50, 51].

In this work, thermochemical properties are estimated using Benson's group additivity method [46]. Group additive methods enable the calculation of standard enthalpies of formation, heat capacities and entropies based on the contributions, so-called group-additive values (GAVs), of groups of atoms in the molecule. Additionally, interactions between non-adjacent atoms or

groups of atoms were quantified in the form of non-next-nearest neighbor interactions (NNIs) and ring strain corrections (RSCs). The current chapter employs GAVs, RSCs and NNIs reported by Poling et al. [52] and Sabbe et al. [50, 53]. A number of species have identical estimated thermochemical properties: linalool and isolinalool; all four biradicals; pinol and component **8**; *cis*- β -terpineol and *trans*- β -terpineol.

2.5 Results and Discussion

Table 2-2 shows the Arrhenius parameters of the reactions of the kinetic model of *cis*- and *trans*-2-pinanol thermal decomposition.

Table 2-2: Pre-exponential factors and fitted activation energies for the two possible cyclobutane rupture reactions (A and B), for the biradical rearrangements (C, D and E), for the ene cyclizations (F and G) and the retro-ene reaction (H) used in the kinetic models of *cis*- and *trans*-2-pinanol (cf. Scheme 2-1 and Scheme 2-2).

F-test value for the global significance of the global kinetic model = 2622.

Reaction	Cis-2-pinanol		Trans-2-pinanol	
	Log ₁₀ A (s ⁻¹)	E _a (kJ mol ⁻¹) ^a	Log ₁₀ A (s ⁻¹)	E _a (kJ mol ⁻¹) ^a
A	14.0	206 ± 2	14.0	211 ± 2
B	14.0	225 ± 2	14.0	225 ± 2
C	13.8	88 ± 6	13.8	88 ± 6
D	11.2	67 ± 6	11.2	67 ± 6
E	13.8	88 ^b	13.8	88 ^b
	Log ₁₀ A (s ⁻¹)		E _a (kJ mol ⁻¹)	
F	8.0		122 ± 2	
G	8.0		122 ± 2	
H	11.3		100 ^c	

^a) 95% confidence level

^b) Arrhenius parameters of the subsequent biradical rearrangement E were taken equal to the fitted parameters of the analogous rearrangement reaction C

^c) Arrhenius parameter not estimated

2.5.1 Cis- and trans-2-pinanol rearrangement mechanism

Analysis of the experimental data shows that *cis*-2-pinanol conversion varied between 8% to 95% and *trans*-2-pinanol conversion between 11% and 95%. Table 2-3 shows a comparison of simulated and experimental mass fractions for two experiments.

Table 2-3: Comparison between simulated and experimental product mass fractions for *cis*- and *trans*-2-pinanol pyrolysis.

	Cis-2-pinanol		Trans-2-pinanol	
Conditions				
$F_{0,2\text{-pinanol}}$ (10^{-2} g s ⁻¹)	1.568		1.837	
$F_{0,\text{methane}}$ (10^{-2} g s ⁻¹)	1.397		1.636	
T _{avg} (K)	773		773	
P (10 ⁵ Pa)	4.4		4.4	
Wt %	Sim.	Expt.	Sim.	Expt.
Linalool	20.7	20.2	11.3	11.1
Cis/trans-β-terpineol	1.7	1.6	0.9	0.7
Isolinalool	1.3	1.3	1.5	1.4
Plinol	6.2	5.8	2.9	2.8
Component 9	0.5	0.4	0.5	0.4
X _{2-pinanol} (%)	57	55	32	32

Components encountered in the reactor effluent were methane, *cis*-2-pinanol, *trans*-2-pinanol, linalool, *cis*- β -terpineol, *trans*- β -terpineol, isolinalool, plinol and 5,7-dimethyloct-6-ene-2-one . Although the isomers of plinol occurred as separate GC peaks, they are lumped together and treated as one component. Figure 2-3 shows the mass fraction and selectivity of linalool as a function of 2-pinanol conversion.

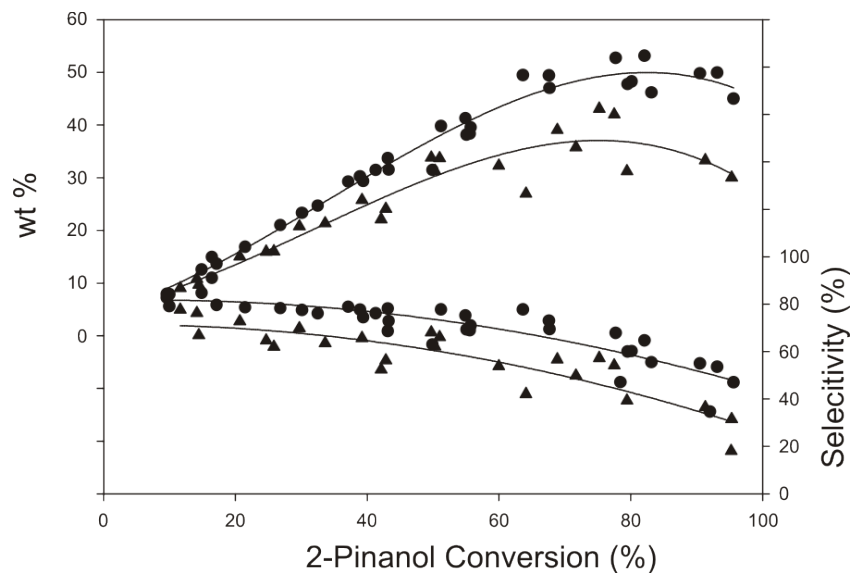


Figure 2-3: Mass fraction and selectivity of linalool as a function of 2-pinanol conversion. Experimental values: ●: *cis*-2-pinanol / ▲: *trans*-2-pinanol. —: model calculations using the Arrhenius parameters of Table 2-2.

Linalool mass fractions and selectivities at equal conversion are significantly higher for *cis*-2-pinanol than for *trans*-2-pinanol. The preference of *cis*-2-pinanol for the fragmentation route via biradical **2** is also confirmed by the lower mass fractions of isolinalool and component **9** compared to *trans*-2-pinanol experiments, as depicted in Figure 2-4.

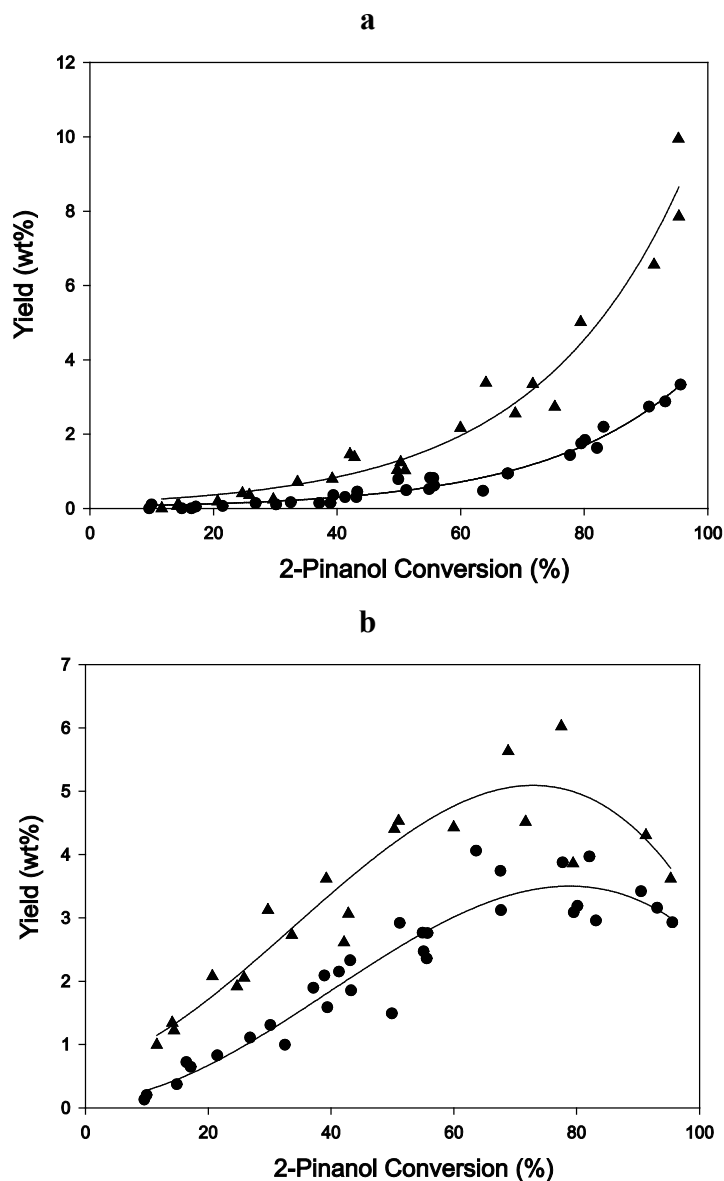


Figure 2-4: Mass fractions as a function of 2-pinanol conversion of a) component 9 and b) isolinalool. Experimental values: ●: *cis*-2-pinanol/ ▲: *trans*-2-pinanol as reactant. —: model calculations using the Arrhenius parameters of Table 2-2.

The undesired cyclization product of linalool, pinol, attains significant mass fractions at higher conversions of 2-pinanol. The components *cis*- β -terpineol and *trans*- β -terpineol are produced in modest amounts in *cis*- and *trans*-2-pinanol experiments respectively and their mass fractions increase linearly with 2-pinanol conversion, cf. Figure 2-5.

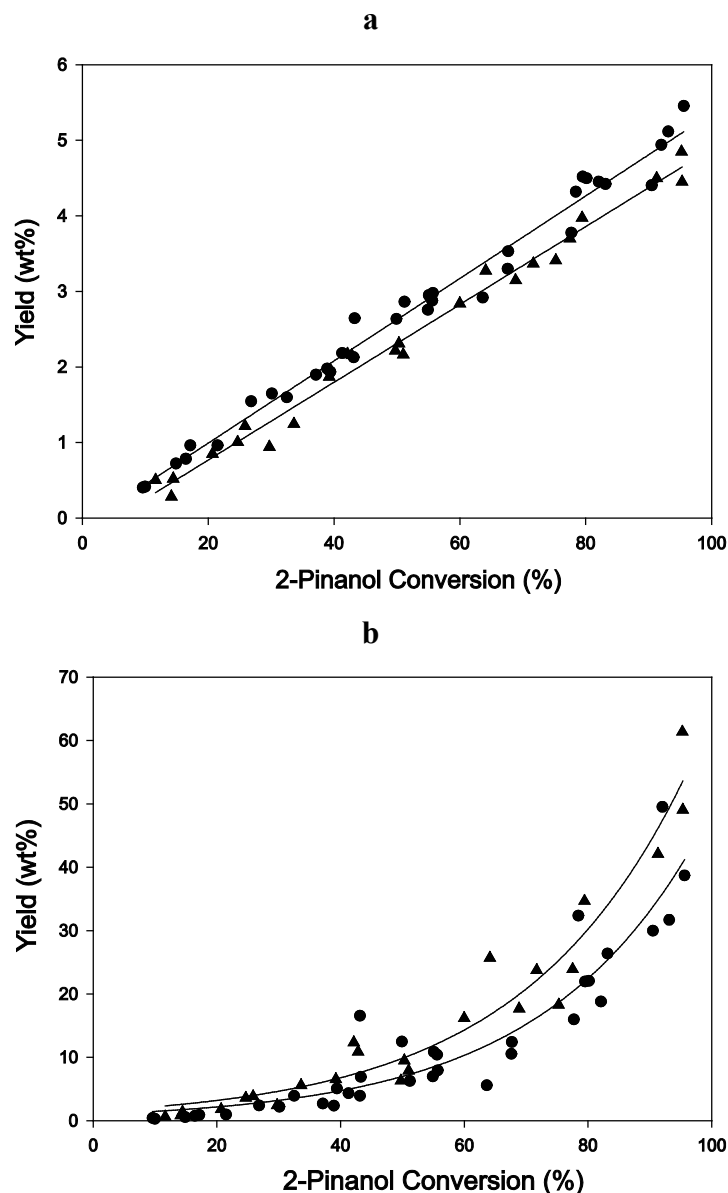
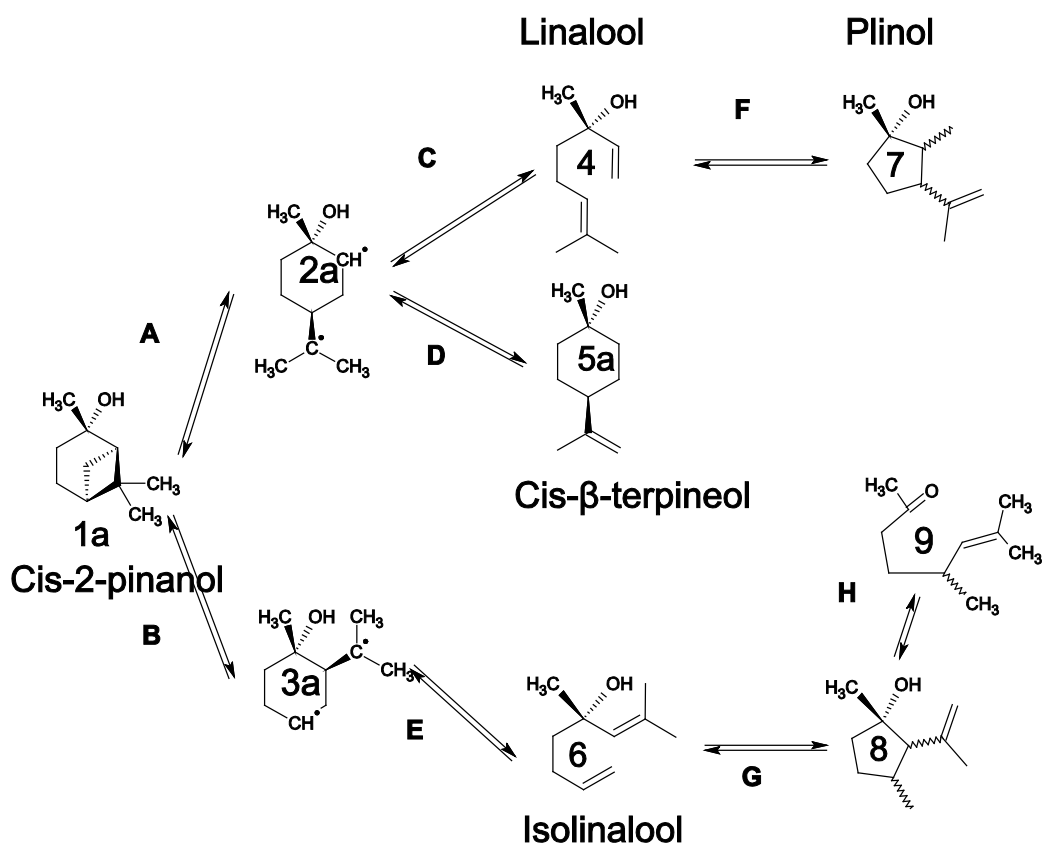


Figure 2-5: Mass fractions as a function of 2-pinanol conversion of a) *cis*-β-terpineol / *trans*-β-terpineol and b) pinol. Experimental values: ●: *cis*-2-pinanol ▲: *trans*-2-pinanol as reactant. —: model calculations using the Arrhenius parameters of Table 2-2.

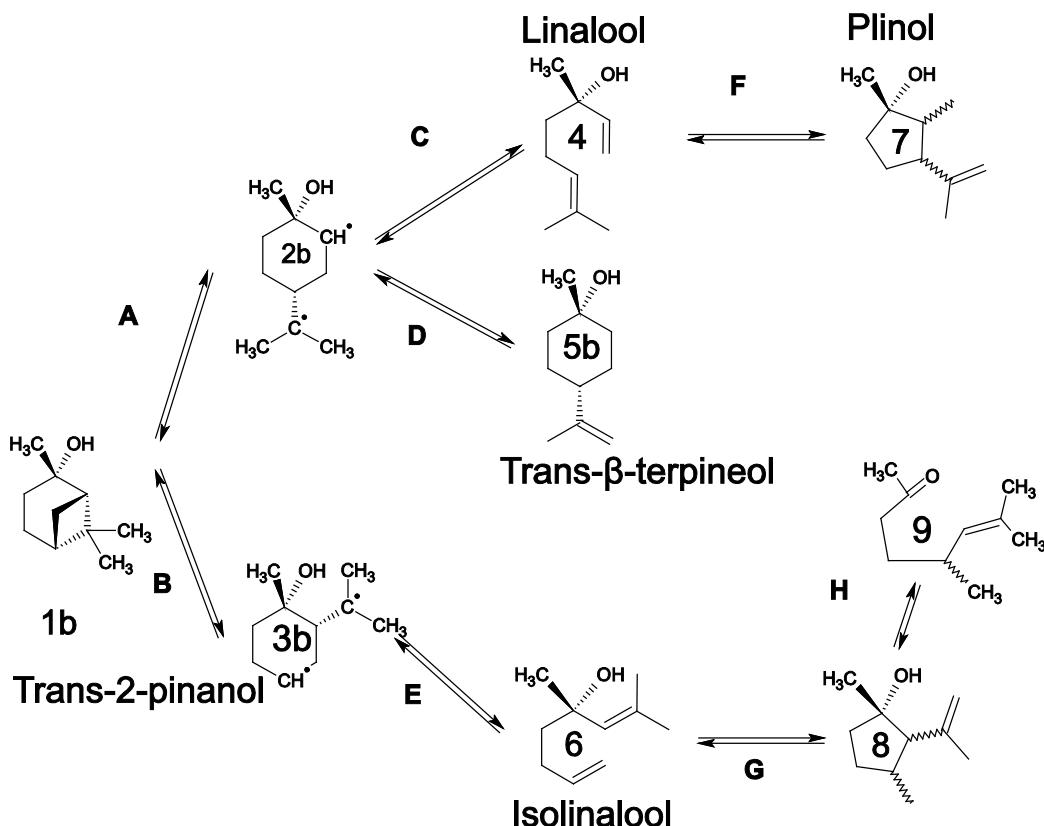
Kinetic models were constructed for the rearrangement of *cis*- and *trans*-2-pinanol which are able to predict the reactant reactivities, product mass fractions and selectivities. Both kinetic models are based on the mechanism for thermal rearrangement of bicyclic monoterpenes containing a [3.1.1]bicycloheptane system that was reported in literature [23, 36]. A stepwise fragmentation of the cyclobutane ring with biradical intermediates is generally accepted as the initial

decomposition of the reactant molecules. The reactions leading to the formation of products formed following the initial cyclobutane fragmentation were also retrieved from literature on the rearrangement of 2-pinanol [44] or similar species, e.g. pinane [23]. Only elementary reactions are taken into account. The rate coefficient of each elementary reaction obeys an Arrhenius temperature dependency, incorporating a pre-exponential factor A , and an activation energy E_a .

As can be observed from Scheme 2-1 (*cis*-2-pinanol) and Scheme 2-2 (*trans*-pinanol), the stepwise mechanism starts with opening of the cyclobutane ring of the 2-pinanol reactant (**1**) by a carbon-carbon scission leading to cyclohexanol biradicals with the methyl and isopropenyl substituent in 1,4 position (component **2**) and in 1,2 position (component **3**) in the cyclohexanol ring.



Scheme 2-1: Reaction mechanism for *cis*-2-pinanol (**1a**) isomerization.



Scheme 2-2: Reaction mechanism for *trans*-2-pinanol (1b) isomerization.

The two remaining C-C scission possibilities of the cyclobutane ring were not considered as they produce unstable biradicals containing primary carbon radicals. Moreover, Kinzel et al. [33] showed through quantum chemical calculations that transition state structures for these remaining ring opening reactions in the thermal rearrangement of *cis*-pinane and *trans*-pinane are energetically disfavored due to the boat or twisted conformation of the cyclohexane ring in contrast to the chair conformation of the cyclohexane ring in the transition state structures of the ring opening reactions leading to biradicals with tertiary radical carbon atoms. The bond dissociation energy (BDE) of the C-C scission producing primary carbon biradicals was found to be 15 kJ mol^{-1} higher compared to the C-C scission producing biradicals **2** and **3** at a temperature of 793K. The biradicals **2** and **3** are subsequently subjected to scissions of the weakened carbon-carbon bond in β -position of the two radical carbon centers, leading to acyclic substituted 1,6-octadienes linalool and isolinalool. A second reaction possibility of the biradical intermediate **2** results in a monocyclic p-menthene-type component **5**, β -terpineol, via a sigmatropic [1,5]-hydrogen shift. It was reported that the acyclic primary pyrolysis products linalool and isolinalool

can further undergo pericyclic ene-cyclizations via six-membered transition states resulting in cyclopentanol derivatives plinol and component **8** [54]. However, cyclopentanol derivative **8** was not detected in the reactor effluents of the experiments. Due to the relative position of the hydroxyl and the isopropenyl groups in component **8**, the latter can rearrange into an acyclic δ,ϵ -unsaturated ketone component **9** via a retro-ene reaction [55], in which a γ -hydrogen atom is transferred to an unsaturated center via a six-electron cyclic transition state [56]. The inability of the analysis apparatus to detect component **8** was excluded since the plinol isomers (**7**) formed in the parallel isomerization route were detected by the apparatus. It was therefore concluded that the disappearance rate of **8** forming **9** is significantly higher than the ene-cyclization **G** and **8** is considered to be a fast reacting intermediate. Monocyclic products formed by a [1,5]H-shift originating from biradical **3** were not detected but neither reported in literature.

Although *cis*- and *trans*-2-pinanol rearrangements are similar from a mechanistic point of view, the stereo-configurations of the species in each mechanism differ: *cis*- and *trans*-2-pinanol have different stereo-configurations of the carbon atom in α -position of the bridgehead atom. Ohloff and Klein [9] already showed that both *cis*- and *trans*-2-pinanol, each isomer consisting of 2 enantiomers, formed one of the linalool enantiomers resulting in four reaction channels. The present chapter only takes diastereomeric differences between species into account. Due to the conservation of the stereo-configuration of the carbon atom in α -position of the bridgehead atom of both 2-pinanol isomers, each of the C-C scissions of the cyclobutane ring of *cis*- and *trans*-2-pinanol gives rise to diastereomeric biradicals (**2a** and **2b**, **3a** and **3b**). In the subsequent biradical rearrangements **C** and **E**, the stereocenter carrying the isopropenyl ligand is lost and linalool and isolinalool are formed. Conversely, the sigmatropic [1,5]-hydrogen shift reaction turning the biradicals **2a** and **2b** into monocyclic p-menthene-type components does preserve the respective stereo-configurations of the biradicals and forms diastereomers of β -terpineol, **5a** and **5b**. The (3,4)-intramolecular ene-cyclization of plinol and component **8** yielding cyclopentanol derivatives creates two additional stereocenters, yielding four distinct stereo-isomers for both components as reported by Coxon et al. [12] The subsequent retro-ene reaction of **8** in which the cyclopentanol ring is opened, yielding **9**, eliminates one stereo-center. Thus, differences in stereo-configurations of species or reactions between the isomerization mechanism of *cis*- and *trans*-2-pinanol are limited to reactions involving the biradicals, (reactions **A**, **B**, **C**, **D**, **E**). The

rearrangements of linalool and isolinalool and the subsequent retro-ene reaction of component **8** (reactions **F**, **G**, **H**) do not exhibit differences in stereo-aspects between the *cis*- and *trans*-2-pinanol model.

2.5.2 Pre-exponential factors

Due to the high number of parameters to be estimated relative to the number of available experiments, pre-exponential factors of all forward reactions are fixed and set to the same values in the *cis*- and *trans*-2-pinanol model. Although small differences could exist between the pre-exponential factors of the competing initial fragmentation routes of both isomers, the assumption of equal values for the pre-exponential factors eliminates possible compensation effects when comparing the corresponding activation energies and allows interpreting differences in reactivity and selectivity between *cis*- and *trans*-2-pinanol in terms of differences in activation energies. Pre-exponential factors for the forward reactions are adopted from the analogous pinane rearrangement, reported by Stolle et al. [26]. (cf. Table 2-2) This choice is justified by the similarities in mechanistic aspects between pinane and 2-pinanol rearrangements. Indeed, the relative differences between the pre-exponential factors of the corresponding reactions of pinane and 2-pinanol can be expected to be insignificantly small because the transition states of the corresponding reactions of pinane and 2-pinanol involve the same atoms and bonds. Only small entropic and enthalpic changes between the reactants and their transition states are expected since the only difference embodies a stretched carbon-carbon bond that is eventually ruptured, i.e. an early transition state is expected with the transition state structure resembling reactants rather than products. Moreover, the hydroxyl group present in 2-pinanol is not directly involved in the transition states of initial fragmentations and biradical rearrangements. Furthermore, other studies of the analogous α -pinene resulted in very similar pre-exponential factors for the corresponding reactions [41, 57], even more corroborating the mechanistic and kinetic similarities of bicyclic monoterpenes. Therefore, the values used for the pre-exponential factors are $1.0 \cdot 10^{14} \text{ s}^{-1}$ for the cyclobutane fragmentations (**A**, **B**) and $0.6 \cdot 10^{14} \text{ s}^{-1}$ for the subsequent rearrangement of the biradicals (**C**, **E**). The pre-exponential factor for the sigmatropic [1,5]-hydrogen shift reaction forming β -terpineol isomers (**D**) was given a value of $1.6 \cdot 10^{11} \text{ s}^{-1}$, in agreement with values reported for similar systems[23]. Pre-exponential factors for the pericyclic (3,4)ene-reactions

(**F,G**) were set to $1.0 \cdot 10^8 \text{ s}^{-1}$. Similar values were reported for the reaction of linalool to plinol [26]. The tight transition states are cited as the reason for the low A-factors and were also found by Pickenhagen et al. [44] through experiments. Arrhenius parameters for the thermal intramolecular retro-ene reaction **H** were not found in literature. Smith and Yates [58] reported a pre-exponential factor of $2.0 \cdot 10^{11} \text{ s}^{-1}$ for the retro-ene reaction of 4-hydroxy-4-methyl-1-pentene producing propylene and acetone. Egger et al. [59] found similar values for other types of components that are able to undergo retro-ene rearrangements. This value was adopted in the current kinetic models.

2.5.3 Activation energies

Biradical intermediates **2** and **3** could not be detected with the current analysis apparatus, causing the kinetics of reactions **B** and **E** to be experimentally indistinguishable. It was assumed that the cyclobutane rupture reaction **B** leading to the 1,2-disubstituted cyclohexanol biradical was the rate-determining reaction for the formation of isolinalool. Arrhenius parameters of the subsequent biradical rearrangement **E** were taken equal to the fitted parameters of the analogous rearrangement reaction **C**, since both reactions exhibit great similarities. Since the subsequent ene cyclizations of the acyclic primary pyrolysis products linalool and isolinalool into cyclopentanol derivatives **7** and **8** are common in the models for the two 2-pinanol isomers, the Arrhenius parameters of these reactions are regressed based on the entire set of both *cis*- and *trans*-2-pinanol experiments. For the thermal intramolecular retro-ene reaction (**H**) Smith and Yates [58] reported an activation energy of 170 kJ mol^{-1} for the retro-ene reaction of 4-hydroxy-4-methyl-1-pentene producing propylene and acetone. It was found that this activation energy resulted in retro-ene kinetics that underestimated mass fractions of component **9** and produced significant amounts of component **8**. It was therefore opted to lower the activation energy of reaction **H** to 100 kJ mol^{-1} , resulting in predicted mass fractions of component **9** that are insensitive to the kinetics of reaction **H**.

The regression to the experimental data of *cis*- and *trans*-2-pinanol experiments, in which a total number of 10 parameters are fitted, yielded a kinetic model for the isomerization of *cis*- and *trans*-2-pinanol that is globally significant (F-test value = 2622). Model predictions for mass

fractions of response variable as a function of reactant conversion are depicted in Figure 2-3 to Figure 2-5.

2.5.4 2-pinanol reactivity - cyclobutane fragmentation

The direct comparison of the conversion of *cis*- and *trans*-2-pinanol based on experimental data is impossible due to the non-identical operating conditions or mass flow rates set in the experiments of the two isomers. In order to compare the conversion of both isomers at similar conditions, the equivalent reactor volume V_E concept [60] is used. V_E is defined as the reactor volume, which at a reference temperature, T_R , would give the same conversion as the actual reactor volume. Figure 2-6 shows the conversion of *cis*- and *trans*-2-pinanol as a function of the “equivalent space time” V_E/F_{tot} with F_{tot} the total volumetric flow rate at the inlet of the reactor and at 763K.

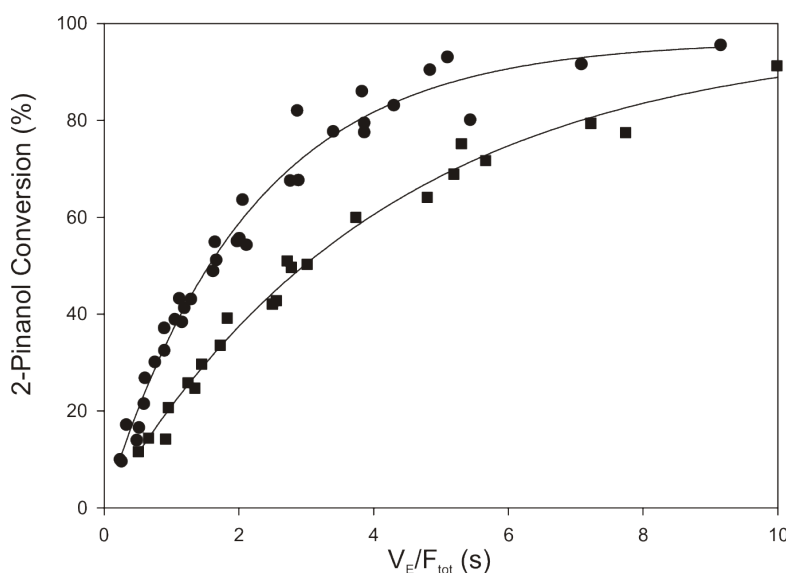


Figure 2-6: Conversion of 2-pinanol as a function of reactor space time, V_E/F_{tot} , with V_E the equivalent reactor volume and F_{tot} the total volumetric flow rate at the inlet of the reactor at reference temperature of 763K. Experimental values: ●: *cis*-2-pinanol and ■: *trans*-2-pinanol. —: model calculations using the Arrhenius parameters of Table 2-2.

It can be observed that for equal equivalent space times the conversion of *cis*-2-pinanol is higher, as much as 20% (absolute) at most. Model predictions for identical flow rates of reactants, diluents and equal operating conditions indicated a temperature for equal conversions of *cis*- and

trans-2-pinanol shifted about 20K to higher values for *trans*-2-pinanol relative to the *cis* isomer. Temperatures corresponding to complete conversion are 838K and 858K for *cis*- and *trans*-2-pinanol respectively, see Figure 2-7.

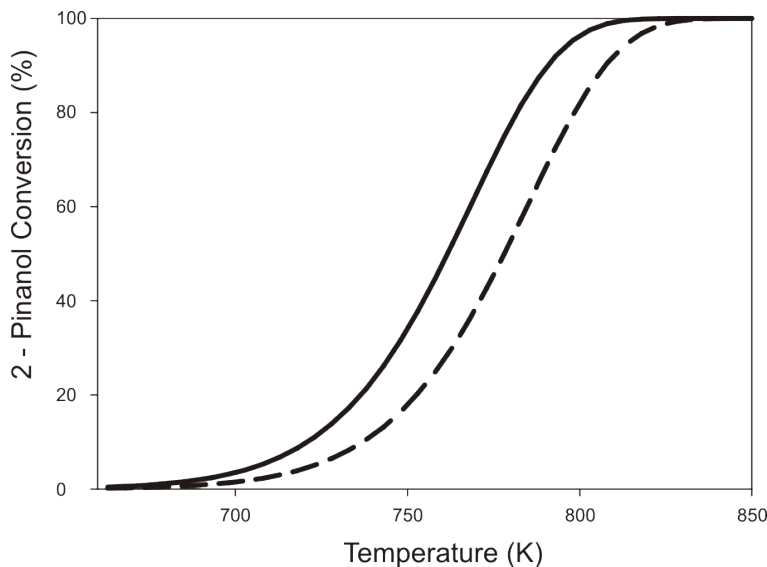


Figure 2-7: 2-Pinanol conversion as a function of temperature. Model calculations for —: *cis*-2-pinanol and - - : *trans*-2-pinanol. Conditions: 2-pinanol mass flow rate: $1.568 \cdot 10^{-2} \text{ g s}^{-1}$, methane mass flow rate: $0.3147 \cdot 10^{-2} \text{ g s}^{-1}$, $P = 1.15 \cdot 10^5 \text{ Pa}$.

The disappearance of both 2-pinanol diastereomers depends on the two cyclobutane fragmentation reactions. The two pairs of activation energies of these two possible reactions **A** and **B**, 206 kJ mol^{-1} and 225 kJ mol^{-1} for *cis*-2-pinanol and 211 kJ mol^{-1} and 225 kJ mol^{-1} for *trans*-2-pinanol. (see Table 2-2). BDEs of the C-C bonds that are broken during the cyclobutane opening are 210 kJ mol^{-1} for the bond in *cis*-2-pinanol and 216 kJ mol^{-1} in *trans*-2-pinanol is found, calculated from standard enthalpy difference between the reactant isomers and the formed biradicals. The similar values of the obtained activation energies and the calculated BDEs confirm our assumption of the stepwise mechanism in favor of the concerted pathway for the cyclobutane fragmentation. The activation energies for reaction **A** yielding biradical **2** are significantly lower than the activation energies for reaction **B** yielding biradical **3** for both reactant isomers; a difference of 19 kJ mol^{-1} for *cis*-2-pinanol and 14 kJ mol^{-1} for *trans*-2-pinanol. It is observed that the activation energies for reaction **B** are not significantly different in both kinetic models. For reaction **A**, a difference of 5 kJ mol^{-1} between both activation energies is found. When using the activation energy of reaction **A** of the *cis*-2-pinanol model in the *trans*-2-

pinanol and vice versa, the global significance of the model drops significantly (F-value: 505 and 1213 respectively), implying that both kinetic models are very sensitive to the value of this activation energy. Comparison of activation energies with kinetic data from literature is biased because the disappearance of the reactant isomers in the current models was represented by two reactions instead of one global reaction as is frequently done in literature. Stolle et al. [26] reported activation energies for the disappearance of *cis*-pinane and *trans*-pinane of 201.1 and 213.0 kJ mol⁻¹ respectively and evidenced the higher reactivity of *cis*-pinane relative to *trans*-pinane by the higher temperature required for complete conversion for *trans*-pinane (873K) compared to *cis*-pinane (823K) resulting in a temperature difference of 50K. This was explained by a difference in activation energies of the reactant disappearance rates of 11.9 kJ mol⁻¹. Il'Ina et al. [42] reported activation energies for the disappearance of *cis*- and *trans*-2-pinanol of 215.1 kJ mol⁻¹ and 190.3 kJ mol⁻¹ respectively but observed a higher reactivity of *cis*-2-pinanol relative to the *trans*-isomer. This could be attributed to a higher pre-exponential factor of the kinetic coefficient in case of *cis*-2-pinanol compared to *trans*-2-pinanol.

Interpreting the differences in activation energies of the cyclobutane fragmentations in terms of steric differences in *cis*- and *trans*-2-pinanol is less straight-forward. Il'Ina et al. attributed the higher instability of *cis*-2-pinanol to the “weakening of the C-C bonds in the pinane ring because of the repulsion of the bulky CH₃-groups oriented in the same direction in the *cis*-2-pinanol molecule” and to the stabilization effect of the hydrogen-bridge bond between a hydrogen of the CH₃ group on the cyclobutane ring and the oxygen of the hydroxyl group in *trans*-2-pinanol [42]. Stolle et al. [31] showed, through a 1D NMR-NOE spectroscopy study, of *cis*- and *trans*-pinane that attractive interactions existed in *trans*-pinane between the protons in the methyl group in α position of the bridgehead atom and the protons of the methylene group in the cyclobutane. These interactions were not found in *cis*-pinane and it was therefore concluded that *trans*-pinane was thermally the more stable molecule. Furthermore, Weigand and Schneider reported a higher ring strain energy in *cis*-pinane than in *trans*-pinane [61].

2.5.5 Product selectivity - biradical rearrangements

The presence of linalool and the two β -terpineol diastereomers in the reactor effluent provides information on the competitive rearrangement reactions **C** and **D** of biradical **2** formed through the cyclobutane rupture reaction **A**. The corresponding activation energies are presented in Table 2-2. For both types of biradical rearrangements very low activation energies are obtained: 88 kJ mol⁻¹ for reaction **C** leading to acyclic products and 67 kJ mol⁻¹ for reaction **D**. The activation energies for the rearrangement reactions were found to be equal for the *cis*- and *trans*-2-pinanol model. It was therefore concluded that the differences in stereo-configurations of reactants biradicals **2a** and **2b** and for products such as **5a** and **5b** did not lead to distinguishable differences in activation barriers. The low activation energy for the C-C scission reaction leading to the acyclic diene linalool is explained by the weakening of the C-C bond in allylic position of the two radical centers in biradical **2**. Additionally, the sp² hybridized carbon radical that is part of the cyclohexane ring, induces ring strain and could also contribute to the low activation energy.

Stolle et al reported that sigmatropic [1,5]-H shift reactions of molecular components leading to monocyclic products have low activation energies (~100 – 160 kJ mol⁻¹) [23]. Activation energies for the reactions of biradical **2** producing β -terpineol diastereomers **5a** and **5b** were estimated to be 67 kJ mol⁻¹. Since the current chapter employs similar pre-exponential factors as for molecular sigmatropic [1,5]-H-shift reactions (~10¹¹s⁻¹) and significantly lower activation energies were found, it is concluded that hydrogen shifts involving biradicals occur faster than molecular hydrogen shifts. The activation energies of the cyclobutane rupture reaction (**A**) and the consecutive biradical rearrangements (**C** and **D**) show that the rate of formation of the linalool and β -terpineol is governed by the cyclobutane rupture reaction **A** while the product selectivity is determined by the two competing biradical rearrangements **C** and **D**.

As mentioned earlier, the activation energy for the rearrangement reaction **E** of biradical **3** to **6** is not fitted due to the impossibility of estimating this parameter from the current experimental data. It is therefore copied from the analogous isomerization of biradical **2** to linalool. The good agreement between experimental data and model predictions for isolinalool (Figure 2-4) supports the assumption that the biradical rearrangements **C** and **E** to acyclic diene products obey the same

kinetics. Moreover, the hypothesis that reaction **B**, i.e. the rupture of the cyclobutane ring, is the rate-determining step for the two consecutive reactions **B** and **E** leading to isolinalool is also justified by the good agreement. This hypothesis was required since only one response, isolinalool, originates from these reactions.

Moreover, the parallel fragmentation route via the cyclobutane ring opening reaction **A** and the consecutive biradical rearrangements **C** and **D** yields two responses: linalool and β -terpineol. A hypothesis on the choice of the rate-determining step was not required for this route since the responses of linalool and β -terpineol allowed the estimation of the Arrhenius parameters of the consecutive biradical rearrangements **C** and **D**. A comparison of the kinetic coefficients k of reactions **A**, **C** and **D** show that the kinetic coefficient of reaction **A**, i.e. the parallel rupture reaction of the cyclobutane ring, is several orders of magnitude lower than the consecutive biradical rearrangements **C** and **D**, at the mean temperature throughout the experiments (793K) (cf. Table 2-4).

Table 2-4: Rate coefficients for reactions A, C, D at the mean temperature throughout the experiments (793K).

	Rate coefficient k (s ⁻¹)	
	Cis-2-pinanol model	Trans-2-pinanol model
Reaction A	2.69	1.26
Reaction C	$1.01 \cdot 10^8$	$1.01 \cdot 10^8$
Reaction D	$6.12 \cdot 10^6$	$6.12 \cdot 10^6$

The pericyclic ene cyclizations **F** and **G**, which convert the acyclic primary pyrolysis products linalool and isolinalool to cyclopentanol derivatives **7** and **8**, were fitted based on the complete set of *cis*- and *trans*-2-pinanol pyrolysis experiments. Table 2-2 presents the obtained activation energies for both cyclizations. The value of 122 kJ mol⁻¹ is identical to the linalool cyclization kinetics reported by Stolle et al. [26].

2.6 Conclusions

A thermodynamically consistent reaction network was constructed incorporating only elementary reactions for the initial isomerization of the diastereomers *cis*- and *trans*-2-pinanol.

Thermochemical properties of the reacting species were estimated using Benson's group additivity method.

Kinetic models for the isomerization reactions were constructed with all pre-exponential Arrhenius factors fixed to values found in literature for similar reactions. A non-linear least-squares regression of the experimental data allowed to obtain statistically significant estimates of the activation energies and to describe adequately the experimental data over the entire range of operating conditions.

Cyclobutane fragmentation reactions are rate limiting for the formation of primary isomerization products such as linalool, isolinalool and β -terpineol. Since both linalool and β -terpineol originate from the same cyclobutane fragmentation route, the subsequent competing biradical rearrangements of the 1,4-disubstituted biradical determines the selectivity towards the acyclic (linalool) and cyclic products (β -terpineol).

The high selectivity for the linalool reaction channel compared to the parallel pathway yielding isolinalool was attributed to the large differences in activation energies for the cyclobutane rupture reactions that form a 1,4-disubstituted and a 1,2-disubstituted cyclohexanol biradicals, a difference of 19 kJ mol^{-1} and 14 kJ mol^{-1} for *cis*- and *trans*-2-pinanol respectively. The higher reactivity of *cis*-2-pinanol compared to *trans* isomer was observed as well as the higher selectivity of *cis*-2-pinanol relative to *trans* isomer for the linalool reaction pathway. The reactivity and selectivity differences between the two reactant diastereomers was attributed to a difference of 5 kJ mol^{-1} in activation energies of the carbon-carbon scission reaction that opens the cyclobutane ring of the reactant molecules and forms the 1,4-disubstituted cyclohexanol biradical, all other Arrhenius parameters of the reactions in the kinetic models of *cis*- and *trans*-2-pinanol being identical.

The kinetic model presented in this chapter will be validated in Chapter 8 of this thesis using experimental data from literature.

2.7 References

- [1] K.A.D. Swift, Catalytic transformations of the major terpene feedstocks, *Topics in Catalysis*, 27 (2004) 143-155.
- [2] G.W. Huber, S. Iborra, A. Corma, Synthesis of transportation fuels from biomass: Chemistry, catalysts, and engineering, *Chemical Reviews*, 106 (2006) 4044-4098.
- [3] K. Fahlbusch, Flavors and Fragrances, in: *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH, Weinheim, 2005.
- [4] C. Mercier, P. Chabardes, Organometallic Chemistry in Industrial Vitamin-A and Vitamin-E Synthesis, *Pure Appl. Chem.*, 66 (1994) 1509-1518.
- [5] J. Nowicki, Claisen, Cope and related rearrangements in the synthesis of flavour and fragrance compounds, *Molecules*, 5 (2000) 1033-1050.
- [6] A. Corma, S. Iborra, A. Velty, Chemical routes for the transformation of biomass into chemicals, *Chemical Reviews*, 107 (2007) 2411-2502.
- [7] N.I. Tracy, D.C. Chen, D.W. Crunkleton, G.L. Price, Hydrogenated monoterpenes as diesel fuel additives, *Fuel*, 88 (2009) 2238-2240.
- [8] A. Guenther, C.N. Hewitt, D. Erickson, R. Fall, C. Geron, T. Graedel, P. Harley, L. Klinger, M. Lerdau, W.A. McKay, T. Pierce, B. Scholes, R. Steinbrecher, R. Tallamraju, J. Taylor, P. Zimmerman, A Global-Model of Natural Volatile Organic-Compound Emissions, *Journal of Geophysical Research-Atmospheres*, 100 (1995) 8873-8892.
- [9] G. Ohloff, E. Klein, Die Absolute Konfiguration des Linalools durch Verknupfung mit dem Pinansystem, *Tetrahedron*, 18 (1962) 37-42.
- [10] V.A. Semikolenov, Ilyna, II, I.L. Simakova, Effect of heterogeneous and homogeneous pathways on selectivity of pinane-2-ol to linalool isomerization, *Journal of Molecular Catalysis a-Chemical*, 182 (2002) 383-393.
- [11] R. Sercheli, A.L.B. Ferreira, L.H.B. Baptistella, U. Schuchardt, Transition-metal catalyzed autoxidation of cis- and trans-pinane to a mixture of diastereoisomeric pinanols, *J. Agric. Food Chem.*, 45 (1997) 1361-1364.
- [12] J.M. Coxon, M.P. Hartshorne, R.P. Garland, Pyrolysis of Pinanes.5. Pyrolysis of 10 alpha-pinane-2-ols and 10beta-pinane-2-ols and 2-hydroxy-10beta-pin-3-ene, *Australian Journal of Chemistry*, 25 (1972) 353-360.
- [13] M.B. Erman, B.J. Kane, Chemistry around pinene and pinane: A facile synthesis of Cyclobutanes and oxatricyclo-derivative of pinane from cis- and trans-pinane, *Chemistry & Biodiversity*, 5 (2008) 910-919.
- [14] D. Banthorpe, D. Whittaker, Preparation and Stereochemistry of Pinane Derivatives, *Chemical Reviews*, 66 (1966) 643-654.
- [15] D.H. Grayson, Monoterpenoids, *Natural Product Reports*, 14 (1997) 477-522.
- [16] M. Eggersdorfer, Terpenes, in: *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH, Weinheim, 2000, pp. 653-671.
- [17] T. Jenke, G. Sussfink, Kinetic Enantiomeric Differentiation in Catalytic Hydration of non-functionalized terpene olefins with chirally modified ruthenium clusters, *J. Organomet. Chem.*, 405 (1991) 383-391.
- [18] V.A. Semikolenov, Ilyna, II, I.L. Simakova, Linalool synthesis from alpha-pinene: kinetic peculiarities of catalytic steps, *Applied Catalysis a-General*, 211 (2001) 91-107.
- [19] G.S. Fisher, J.S. Stinson, R.N. Moore, L.A. Goldblatt, Peroxides from Turpentine - Production of Technical Grade Pinane Hydroperoxide, *Industrial and Engineering Chemistry*, 47 (1955) 1368-1373.
- [20] L.A. Shutikova, Cherkaev, V.G., Erzhanova, M.S., Alifanova, A.V., *Maslo-zhirovaya promyshlennost.*, 8 (1978) 23.
- [21] B.J. Kane, G. Marcelin, S.G. Traynor, Chemistry of the 2-Pinanols. 1. Pinanoxide Basicities, *Journal of Organic Chemistry*, 45 (1980) 895-900.
- [22] S.G. Traynor, B.J. Kane, J.B. Coleman, C.G. Cardenas, Chemistry of the 2-Pinanols. 2. Investigation of the Utility of the Enhanced Basicities of the Pinanoxide Bases, *Journal of Organic Chemistry*, 45 (1980) 900-906.
- [23] A. Stolle, B. Ondruschka, H. Hopf, Thermal Rearrangements of Monoterpenes and Monoterpenoids, *Helvetica Chimica Acta*, 92 (2009).
- [24] A. Stolle, Study on the Pyrolysis Behaviour of some monoterpenes and monoterpenoids: a Mechanistic and Kinetic Overview, *Friedrich-Schiller Universität Jena* (2008).
- [25] A. Stolle, W. Bonrath, B. Ondruschka, Kinetic and mechanistic aspects of myrcene production via thermal-induced beta-pinene rearrangement, *Journal of Analytical and Applied Pyrolysis*, 83 (2008) 26-36.

- [26] A. Stolle, W. Bonrath, B. Ondruschka, D. Kinzel, L. Gonzalez, Kinetic model for the thermal rearrangement of cis- and trans-pinane, *Journal of Physical Chemistry A*, 112 (2008) 5885-5892.
- [27] A. Stolle, C. Brauns, M. Nuchter, B. Ondruschka, W. Bonrath, M. Findeisen, Thermal behaviour of selected C₁₀H₁₆ monoterpenes, *European Journal of Organic Chemistry*, (2006) 3317-3325.
- [28] A. Stolle, B. Ondruschka, An effort to generalize the thermal isomerization of 6,6-dimethylbicyclo[3.1.1]heptanes and 6,6-dimethylbicyclo[3.1.1]heptenes: Comparative pyrolysis of pinane, alpha-pinene, and beta-pinene, *Journal of Analytical and Applied Pyrolysis*, 85 (2009) 252-259.
- [29] A. Stolle, B. Ondruschka, Synthesis of myrcene by pyrolysis of beta-pinene: Analysis of decomposition reactions, *Journal of Analytical and Applied Pyrolysis*, 81 (2008) 136-138.
- [30] A. Stolle, B. Ondruschka, W. Bonrath, Comprehensive kinetic and mechanistic considerations for the gas-phase behaviour of pinane-type compounds, *European Journal of Organic Chemistry*, (2007) 2310-2317.
- [31] A. Stolle, B. Ondruschka, W. Bonrath, T. Netscher, M. Findeisen, M.M. Hoffmann, Thermal isomerization of (+)-cis- and (-)-trans-pinane leading to (-)-beta-citronellene and (+)-isocitronellene, *Chemistry-a European Journal*, 14 (2008) 6805-6814.
- [32] A. Stolle, B. Ondruschka, M. Findeisen, Mechanistic and Kinetic Insights into the Thermally Induced Rearrangement of alpha-Pinene, *Journal of Organic Chemistry*, 73 (2008) 8228-8235.
- [33] D. Kinzel, A. Stolle, B. Ondruschka, L. Gonzalez, Quantum chemical investigation of the thermal rearrangement of cis- and trans-pinane, *Phys. Chem. Chem. Phys.*, 12 9884-9892.
- [34] R.B. Woodward, R. Hoffmann, Conservation of Orbital Symmetry, *Angewandte Chemie-International Edition*, 8 (1969) 781-852.
- [35] R. Hoffmann, R.B. Woodward, Orbital Symmetry Control of Chemical Reactions, *Science*, 167 (1970) 825-831.
- [36] R.L. Burwell, The mechanism of the pyrolysis of pinenes, *J. Am. Chem. Soc.*, 73 (1951) 4461-4462.
- [37] B.H. Northrop, K.N. Houk, Vinylcyclobutane-cyclohexene rearrangement: Theoretical exploration of mechanism and relationship to the Diels-Alder potential surface, *Journal of Organic Chemistry*, 71 (2006) 3-13.
- [38] J.E. Baldwin, R.C. Burrell, Kinetics of thermal gas-phase isomerizations and fragmentations of cis- and trans-1-(E)-propenyl-2-methylecyclobutanes at 275 degrees C, *Journal of Organic Chemistry*, 67 (2002) 3249-3256.
- [39] P.B. Dervan, D.S. Santilli, Synthesis and Thermal Decomposition of Cis-3,4,5,6-Tetrahydropyridazine-3,4-D₂ - Relative Rates of Rotation, Cleavage, and Closure for Tetramethylene, *J. Am. Chem. Soc.*, 102 (1980) 3863-3870.
- [40] D.S. Santilli, P.B. Dervan, Thermal Decomposition of Cis-Tetrahydropyridazine-3,4-D₂- Relative Rates of Rotation, Cleavage, and Closure for Tetramethylene, *J. Am. Chem. Soc.*, 101 (1979) 3663-3664.
- [41] J.J. Gajewski, C.M. Hawkins, Gas-Phase Pyrolysis of Isotopically and Stereochemically Labeled Alpha-Pinene - Evidence for a Nonrandomized Intermediate, *J. Am. Chem. Soc.*, 108 (1986) 838-839.
- [42] Il'ina, II, I.L. Simakova, V.A. Semikolenov, Kinetics of 2-pinanol isomerization to linalool on the monolith carbon-containing catalyst, *Kinetics and Catalysis*, 42 (2001) 686-692.
- [43] S. Buddoo, N. Siyakatshana, B. Zeelie, J. Dudas, Study of the pyrolysis of 2-pinanol in tubular and microreactor systems with reaction kinetics and modelling, *Chemical Engineering and Processing: Process Intensification*, 48 (2009) 1419-1426.
- [44] W. Pickenhagen, G. Ohloff, R.K. Russel, W.D. Roth, Intra-Molecular Ene Reaction - Thermal Rearrangement of Linalool and 1,2-Dehydrolinalool, *Helvetica Chimica Acta*, 61 (1978) 2249-2253.
- [45] S. Roy, K. Chakrabarty, G.K. Das, Comparative study on the transition structures of (3,4) and (3,5) ene cyclizations: A theoretical approach, *Journal of Molecular Structure-Theochem*, 820 (2007) 112-117.
- [46] S.W. Benson, *Thermochemical Kinetics*, John Wiley & Sons, New York, 1976.
- [47] M.R. Harper, K.M. Van Geem, S.P. Pyl, G.B. Marin, W.H. Green, Comprehensive reaction mechanism for n-butanol pyrolysis and combustion, *Combustion and Flame*, 158 (2011) 16-41.
- [48] Q. Chen, G.F. Froment, Thermal-Cracking of Substituted Aromatic-Hydrocarbons.1. Kinetic-Study of the Thermal-Cracking of I-Propylbenzene, *Journal of Analytical and Applied Pyrolysis*, 21 (1991) 27-50.
- [49] H.H. Rosenbrock, An Automatic Method for Finding the Greatest or Least Value of a Function, *The Computer Journal*, 3 (1960) 175-184.
- [50] M.K. Sabbe, M. Saeys, M.F. Reyniers, G.B. Marin, V. Van Speybroeck, M. Waroquier, Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals, *Journal of Physical Chemistry A*, 109 (2005) 7466-7480.
- [51] M.K. Sabbe, F. De Vleeschouwer, M.F. Reyniers, M. Waroquier, G.B. Marin, 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*, 112 (2008) 12235-12251.
- [52] B.E. Poling, J.M. Prausnitz, J.P. O'Connell, *The properties of gases and liquids*, McGraw-Hill, New York, 2001.

- [53] M.K. Sabbe, F. De Vleeschouwer, M.F. Reyniers, M. Waroquier, G.B. Marin, First principles based Group additive values for the gas phase standard entropy and heat capacity of hydrocarbons and hydrocarbon radicals, *J.Phys.Chem.A*, 112 (2008) 12235-12251.
- [54] W. Oppolzer, V. Snieckus, Intra-Molecular Ene Reactions in Organic Synthesis, *Angew. Chem.-Int. Edit. Engl.*, 17 (1978) 476-486.
- [55] J.M. Coxon, R.P. Garland, M.P. Harthshorn, Pyrolysis of Nopinol, *Journal of the Chemical Society D-Chemical Communications*, (1970) 1709-1710.
- [56] J.L. Ripoll, Y. Vallee, Synthetic Applications of the Retro-Ene Reaction, *Synthesis*, (1993) 659-677.
- [57] J.J. Gajewski, I. Kuchuk, C. Hawkins, R. Stine, The kinetics, stereochemistry, and deuterium isotope effects in the alpha-pinene pyrolysis. Evidence for incursion of multiple conformations of a diradical, *Tetrahedron*, 58 (2002) 6943-6950.
- [58] G.G. Smith, B.L. Yates, Pyrolysis Studies. 16. A Mechanistic Study of Pyrolysis of Beta-Hydroxy Olefins, *Journal of the Chemical Society*, (1965) 7242-7246.
- [59] K.W. Egger, P. Vitins, Thermochemical kinetics of the retro-ene reactions of molecules with the general structure (allyl)XYH in the gas phase. 6. Concerted unimolecular decomposition of hepta-1,6-diene, *J. Am. Chem. Soc.*, 96 (1974) 2714-2719.
- [60] G.F. Froment, K.B. Bischoff, *Chemical reactor design and analysis*, J. Wiley, New York, 1990.
- [61] E.F. Weigand, H.J. Schneider, C-13 NMR Spectroscopic and Stereochemical Investigations. 23. Conformation, C-13 NMR Shifts and C-13 H-1 Couplings in Pinanes, *Organic Magnetic Resonance*, 12 (1979) 637-644.

Chapter 3: Thermal decomposition of Jet Propellant-10

3.1 Abstract

An experimental study of Jet-Propellant-10 (JP-10) pyrolysis is performed in a continuous flow tubular reactor at atmospheric pressure (1.7×10^5 Pa) in the temperature range of 930 – 1080K, residence times of 0.5s and two dilution levels of 7 and 10 mol % JP-10 in nitrogen. JP-10 starts to react at 900K and is fully converted at 1100K under the investigated conditions. Primary and secondary pyrolysis products were identified and quantified by a combination of on-line two-dimensional gas chromatography, time-of-flight mass spectrometry and a flame ionization detector. Over seventy species with mass-to-charge ratio up to 178 ($C_{14}H_{10}$) were identified and quantified with minimum mole fractions of 0.01 mol %. Decomposition products, such as tricyclo[5.2.1.0^{2,6}]dec-4-ene indicated the importance of bimolecular H-abstraction reactions in the consumption of JP-10. Benzene, toluene and styrene comprise up to 70 mol % of the aromatic hydrocarbons fraction. Polycyclic aromatic hydrocarbon formation starts at higher JP-10 conversion with the consumption of 1,3-cyclopentadiene into bicyclic aromatic species such as indene and naphthalene. Validation of the experimental data with the combustion model of Magoon et al. [1] shows that the model predictions of JP-10 agree reasonably well given the uncertainty of many of the parameters used in the model.

Keywords: JP-10; pyrolysis; thermal decomposition; PAH formation; GC×GC

3.2 Introduction

Twenty-first century advanced aviation applications impose ever increasing demands on fuel performance [2-5]. At present, the fuel “Jet Propellant-10, JP-10” is the only air-breathing missile fuel in operational use [4]. It consists almost entirely of a tricyclic hydrocarbon, exo-tricyclo[5.2.1.0^{2,6}]decane (exo-TCD) that has superior qualities to other aviation fuels in terms of volumetric energy density (39.4 GJ/m³ [6-8]), freezing point (as low as 194K [4]) and thermal stability. Thermal decomposition becomes noticeable at temperatures higher than 900 K at atmospheric pressure [7, 8]. A quantitative understanding of the decomposition chemistry is important for high altitude/high speed applications of JP-10 such as in ramjets and scramjets. Chemical decomposition of the “endothermic” fuel to absorb heat prior to combustion can be helpful when severe demands are placed on the structural integrity and the thermal management of the propulsion technology [9-13].

A lot of information is available about the thermochemical and thermophysical properties of exo-TCD [14-16] but less is known about its pyrolysis and combustion chemistry. Davidson [17] used a high-speed UV absorption kinetic spectrograph to measure exo-TCD pyrolysis products at temperatures between 1100K and 1700K and atmospheric pressure under very diluted conditions. Wohlwend [18] and Striebich [19] studied the thermal stability and decomposition of exo-TCD in a flow reactor under supercritical conditions at a pressure of 34 10⁵ Pa and temperatures ranging from 373 to 923K. Besides the major products, cyclopentene and 1,3-cyclopentadiene (CPD), significant amounts of substituted cyclopentenenes, substituted cyclopentadienes and aromatic hydrocarbons such as benzene, toluene and naphthalene were detected. Nakra et al. [20] investigated exo-TCD pyrolysis on the millisecond timescale in a micro flow tube at temperatures between 900K and 1700K and pressures between 300 and 400 Pa. Mass spectrometry with electron impact ionization and chemical ionization was used to identify products up to C₆. Flow experiments in an annular tubular reactor at atmospheric pressure between 903K and 968K performed by Rao and Kunzru [21] showed that CPD was the most important product at temperatures between 903K and 968K for residence times varying between 0.7s and 6.4s. Other important products included ethene, propene, benzene, toluene and cyclopentene. Herbinet et al. [22] studied the decomposition of helium diluted exo-TCD at atmospheric pressure between temperatures of 848K and 933K in a jet-stirred reactor and identified 3-cyclopentylcyclopentene as an important primary product using mass

spectrometry. Xing et al. [23] investigated the thermal cracking of exo-TCD in a batch reactor at sub- and supercritical conditions and observed that the product spectrum under both regimes was very similar.

A fair number of studies quantified the major products of exo-TCD pyrolysis, but minor intermediate species were seldom reported although their characterizations could provide additional insights in relevant decomposition pathways. In this work a comprehensive experimental dataset for the decomposition of exo-TCD obtained in a bench scale pyrolysis set-up is presented. The on-line application of GC×GC ToF-MS and GC×GC FID allows the identification and quantification of more than 70 pyrolysis products. Several of these products are for the first time identified and quantified. The reported data is discussed in terms of important initial reactions and the chemistry leading to PAH production, and is used to validate the presently available decomposition models that were proposed for JP-10.

3.3 Experimental Procedures

3.3.1 Materials

JP-10 is provided by Sichuan Zhongbang Technical Developing Limited and is used without further purification. JP-10 feedstock was analyzed off-line using a GC×GC–FID/ToF-MS. The latter was described in more detail in Appendix A of this thesis. In the present study the JP-10 feed contains three major components, cf. Figure 3-1. Based on the mass spectra, these were identified as endo-tricyclo[5.2.1.0^{2,6}]decane (endo-TCD), adamantane and decalin. GC×GC/FID analysis reveals that the JP-10 sample consists of 98.43 mol% exo-TCD (in line with fuel specifications [24]), 0.66 mol% endo-TCD, 0.14 mol% adamantane, and 0.06 mol% decalin. Note also that many components with significantly lower mole fractions were also observed with *m/z* ranging from 124 to 136 Da and collectively comprise ca. 0.70 mol % of the feedstock. They are formed as by-products in the production of exo-TCD during the acid-catalyzed isomerization of endo-TCD to exo-TCD [25]. Bruno et al. [16] reported similar levels of these components and provided tentative assignments of those components.

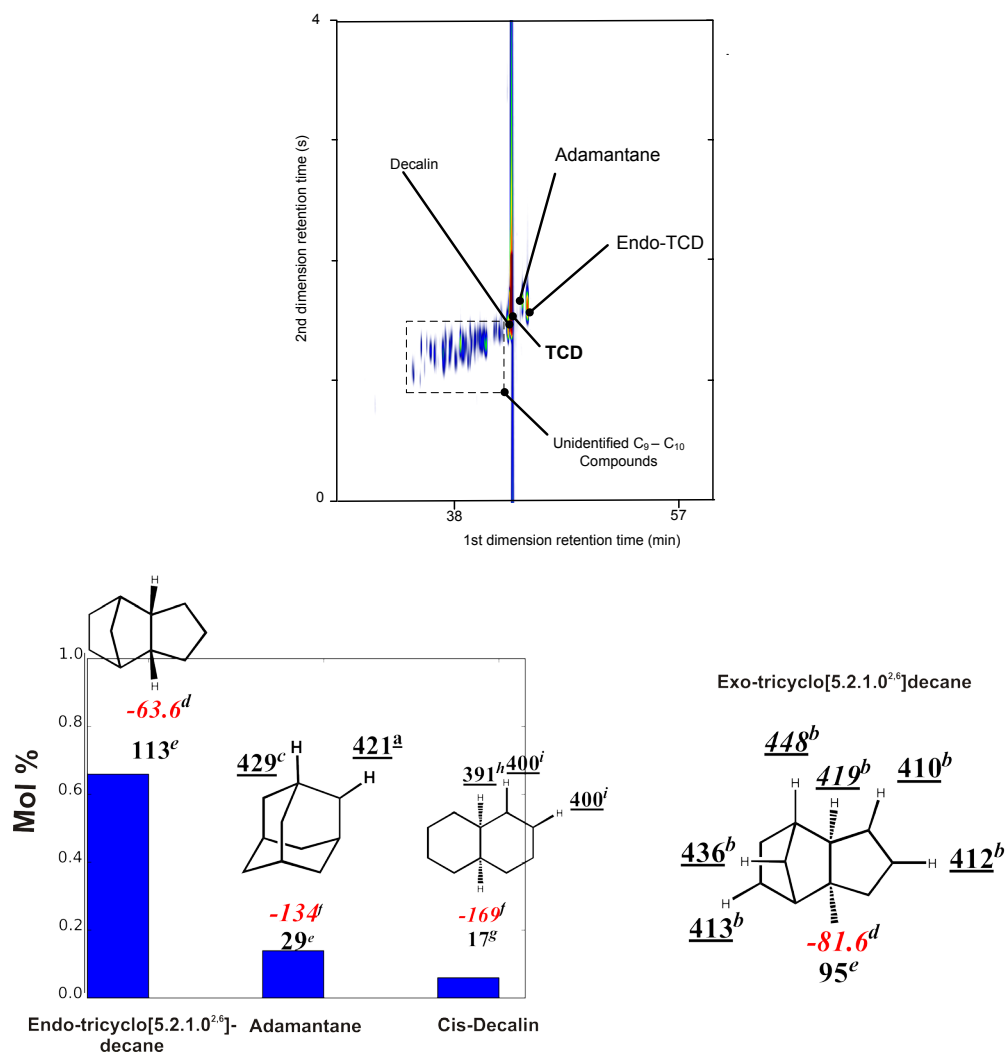


Figure 3-1: Top: GC×GC chromatogram of JP-10 feedstock containing exo-TCD, cis-decalin, adamantane, and endo-TCD and C₉-C₁₀ components. Bottom left: mole fractions (%) of JP-10 feedstock components. Bottom right: structure of TCD. Bond dissociation energies of C-H bonds (underlined, kJ mol⁻¹), standard enthalpies of formation at 298K (in italics, kJ mol⁻¹) and component ring strain energies reported in literature (in bold face, kJ mol⁻¹). aRef [26] b Ref. [27] c Ref. [28] d Ref. [27] e Ref. [29] f Ref. [30] g Ref. [31] h Ref. [32] i taken equal to BDE in cyclohexane.

Analytical gases (N₂, He, CO₂) were obtained from Air Liquide at a minimum purity of 99.999%.

3.3.2 Experimental reactor setup

The JP-10 pyrolysis experiments were conducted in a bench-scale set-up. This reactor set-up has been discussed extensively [33-35] and therefore only a brief description is given below. A schematic of the experimental apparatus is provided in Appendix A of this thesis.

The stream exiting the reactor (623K) is sent to online C_{5+} analysis to a GC \times GC (Thermo Scientific Interscience Belgium) equipped with a Time of Flight Mass Spectrometer (ToF-MS) allowing online identification of components in the effluent. Moreover, the sample may be sent to a Flame Ionization Detector (FID) for the quantification of peaks on the GC \times GC chromatogram. The C_4 -hydrocarbon analysis section consists of a Refinery Gas Analyzer (RGA) equipped with two thermal conductivity detectors for the quantification of hydrogen and other permanent gases respectively, and one FID for the quantification of hydrocarbons up to C_4 .

The inert stream of nitrogen not only serves as a diluent but also as the primary internal standard to determine the absolute flow rates of the other components in the reactor effluent [33]. The flow rates of hydrogen, methane, and hydrocarbons up to C_2 were calculated using the peak surface areas of the RGA thermal conductivity detectors, the experimentally determined relative response factors for each of the detectors, and the known flow rate of the primary internal standard. Quantification of the peaks of the hydrocarbons with higher molecular masses on the RGA-FID and GC \times GC-FID was accomplished by considering ethene as a secondary internal standard for the detectors.

Overall mass balances taking into account all identified products attained on average 100 ± 5 wt%. Elemental balances for carbon closed within 1% and were then normalized. The formation of coke deposits was monitored by the pressure drop along the reactor. The pressure drop over the reactor was negligible under normal operating conditions. Typically less than 0.01 wt% of the feed is converted into cokes during the pyrolysis of hydrocarbons. The current experiments were carried out in presence of a diluent. As coke formation under pyrolysis conditions scales to the third order with hydrocarbon partial pressure [36], the formation of cokes was negligible in comparison with the mole fractions of the reported product species.

3.3.3 Conditions in the pyrolysis zone

A total of 28 experiments were carried out. An overview is given in Table 3-1. Product mole fractions were reproducible within 5%.

Table 3-1: Range of experimental conditions.

Temperature range (K)	936 – 1083
Pressure (10^5 Pa)	1.70
JP-10 mass flow rate (10^{-2} g s $^{-1}$)	2.33
Dilution levels of JP-10 in N $_2$ (mol %)	7 / 10
Conversion range (%)	4–94

Two sets of dilution levels were set by increasing the nitrogen flow rate while keeping the JP-10 flow rate constant. This results in JP-10 dilution levels of 7 mol % and 10 mol % JP-10 in nitrogen. Also, this implies that in both sets of experiments the JP-10 space time remains constant.

3.3.4 Kinetic models

The current data set of JP-10 pyrolysis experiments was used for validating available kinetic models of the thermal decomposition of JP-10, cf. Table 3-2. Li et al. [37] developed the first kinetic model for JP-10 combustion, hereafter referred to as the “San Diego model”. The manually assembled model included 174 reactions among 36 species, primarily consisting of an earlier developed C $_1$ -C $_3$ combustion model [38] (147 reactions between 33 species) while the remaining reactions comprised global pathways from exo-TCD to cyclopentene, 1,3-butadiene and lighter species. Herbinet et al. [22] manually assembled a pyrolysis model, the so-called “NANCY model” containing 898 species and 2623 reactions, but this model is not publicly available. The most recent kinetic model for JP-10 was built by Magoon et al. [39] using the Reaction Mechanism Generator code (RMG) [40]. This model is referred to as the “RMG model”. The latter comprised over 300 species and more than 7000 reactions.

Table 3-2: Summary of published kinetic models for thermal decomposition of JP-10.

Model	San Diego model [37]	Herbinet et al. [22]	RMG model [39]
Primary application domain	Combustion	Pyrolysis	Combustion
Number of species	36	898	320
Number of reactions	174	2623	7740
Mechanism generator	Manually	Manually	RMG [1]

The San Diego and models by Magoon et al. [39] were validated using the current experimental data. Reactor simulations were systematically conducted with the plug flow reactor model of the Chemkin-Pro package [41].

3.4 Results and Discussion

3.4.1 JP-10 conversion as a function of temperature

Exo-TCD conversion as a function of the mean reactor temperature is shown in Figure 3-2, along with the conversion of endo-TCD and adamantane.

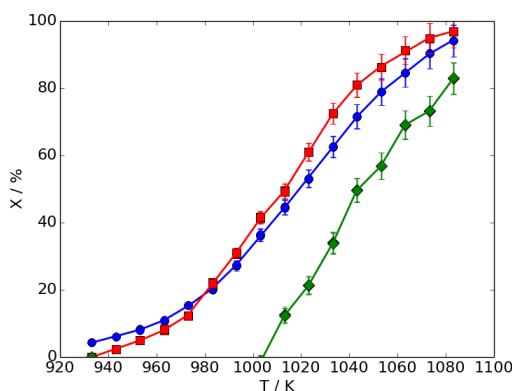


Figure 3-2: Conversion as a function of mean reactor temperature of ●: exo-TCD, ■: endo-TCD, and ♦: adamantane. Conditions: 10 mol % exo-TCD in N_2 at reactor inlet. $P = 1.7 \cdot 10^5$ Pa. $\pm 5\%$ error bars are indicated.

The very small mole fractions of decalin were not reported because of the level of uncertainty associated with them. Essentially full exo-TCD conversion for the experiments with 10 mol % exo-TCD at the reactor inlet occurs at 1100K. Conversion of endo-TCD is slightly higher than for

exo-TCD, while adamantane becomes reactive at temperatures above 1000K. Adamantane, and decalin, are believed to have negligible impact on the initiation of exo-TCD. Only endo-TCD has a higher global ring strain energy than exo-TCD [27] while the other components are more stable. Therefore, endo-TCD may be important for the initiation of exo-TCD, despite the low mole fractions found in the feedstock. Carbon-hydrogen bond dissociation energies (BDE) in decalin are the lowest of all C-H bonds in the four components, cf. Figure 3-1. Radical H-abstraction reactions can therefore preferentially attack decalin C-H bonds but given the low amounts of this component its impact is believed to be limited.

3.4.2 Product distribution

A total of 75 species were identified and quantified. Qualitative differences in the chromatograms of the product effluent at low and high conversion can be observed from Figure 3-3. A summary of the product mole fractions for selected components of an experiment at low and high conversion at both dilution levels is given in Table 3-3.

Table 3-3: Summary of experimental product mole fractions of selected components at various process conditions.

Conditions		10		7	
Dilution JP-10 (mol %)		963	1083	987	1086
Temperature (K)		963	1083	987	1086
Exo-TCD Conversion (%)		10.8	94.1	12.8	89.4
Mole fractions (%)					
Dihydrogen		5.1	27.7	7.1	26.7
C₄		8.6	44.6	10.8	43.5
Methane		1.1	12.9	1.4	10.7
Ethene		4.1	21.7	5.2	21.5
Propene		2.9	6.8	3.3	7.2
Cyclo C₅ derivatives		6.0	8.3	8.0	10.0
Cyclopentene		1.5	0.2	1.7	0.3
1,3-Cyclopentadiene (CPD)		3.6	7.9	5.2	9.5
3-Ethenyl-2-cyclopentene		0.7	0.1	0.8	0.1
Fulvene		0.2	0.1	0.3	0.1
Aromatics		2	16	4.2	15.6
Benzene		0.8	9.4	1.1	8.7
Toluene		0.4	2.8	0.5	2.5
Indene		0.1	0.9	<0.1	0.8
Naphthalene		<0.1	1.2	<0.1	0.9
C ₁₁₊ aromatics		<0.1	0.6	<0.1	0.8
Tricyclo[5.2.1.0 ^{2,6}] dec-4-ene		0.3	<0.1	0.3	<0.1

Since sufficient resolution was available for the adequate quantification of lower molecular mass components, modulation was only initiated after the elution of the C₄ components. The comprehensive GC×GC allows distinguishing between components based on their carbon numbers using the first dimension retention time of the chromatogram. The group of C₉-C₁₀ species observed at low conversion experiments completely disappeared on the chromatogram at 1083K and instead substituted monocyclic aromatic hydrocarbons such as ethylbenzene, styrene, xylenes, propylbenzene, and indene were observed. The separation power of the second dimension column makes that heavier aromatics elute in the band indicated. To the right of exo-TCD, fused ring aromatic components such as methylindenes, naphthalene, methylnaphthalenes, acenaphthylene, and fluorene were detected. Hydrocarbons containing up to fourteen carbons, such as anthracene, were found in non-negligible quantities at conversions close to 100%. The latter also shows the value of the experimental data for these soot precursors.

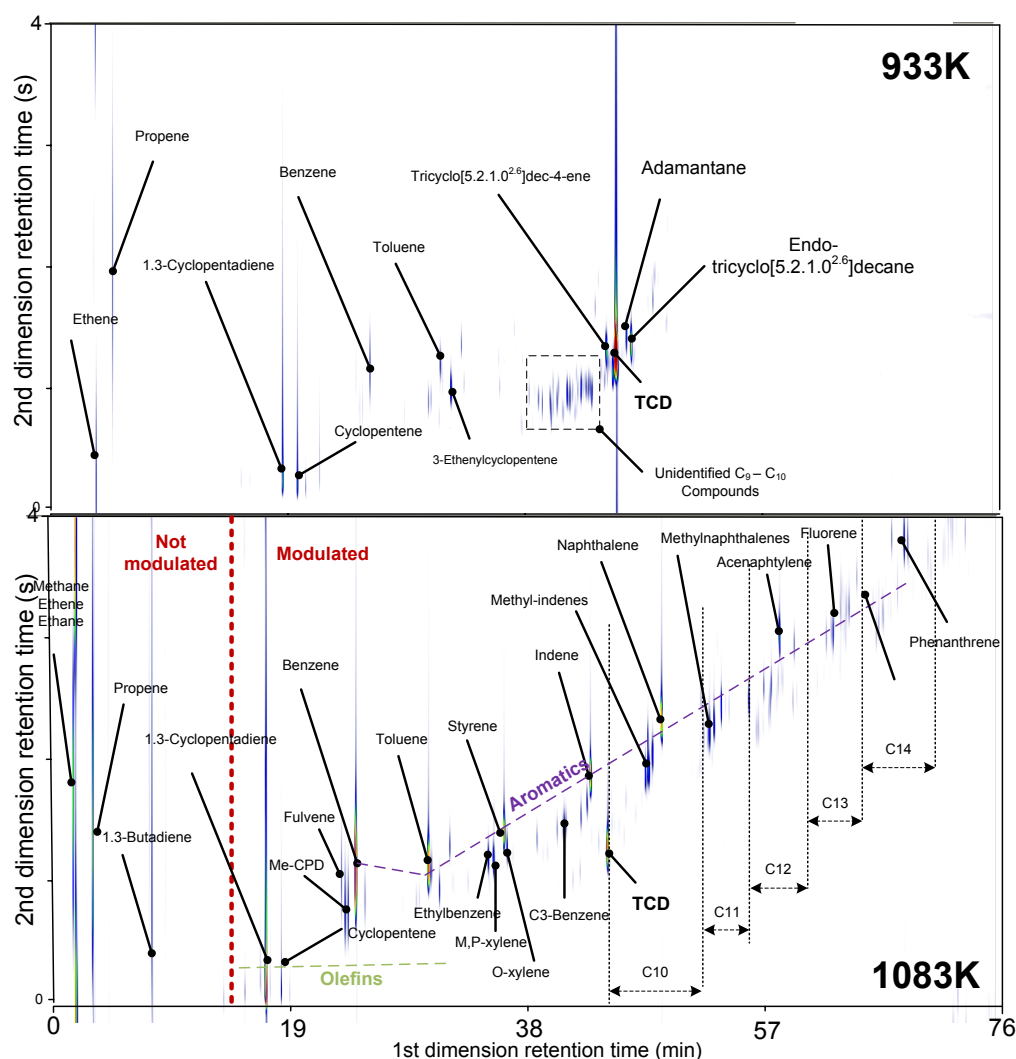


Figure 3-3: GC×GC-FID chromatogram of pyrolysis products at reactor temperatures 933K (upper) and 1083K (lower), corresponding to 4% and 90% exo-TCD conversion respectively for 10% exo-TCD at the reactor inlet.

3.4.2.1 Major pyrolysis products

In Figure 3-4 profiles of components with the highest mole fractions throughout the experiments are depicted: dihydrogen, methane, ethene, propene, CPD, cyclopentene, benzene and toluene.

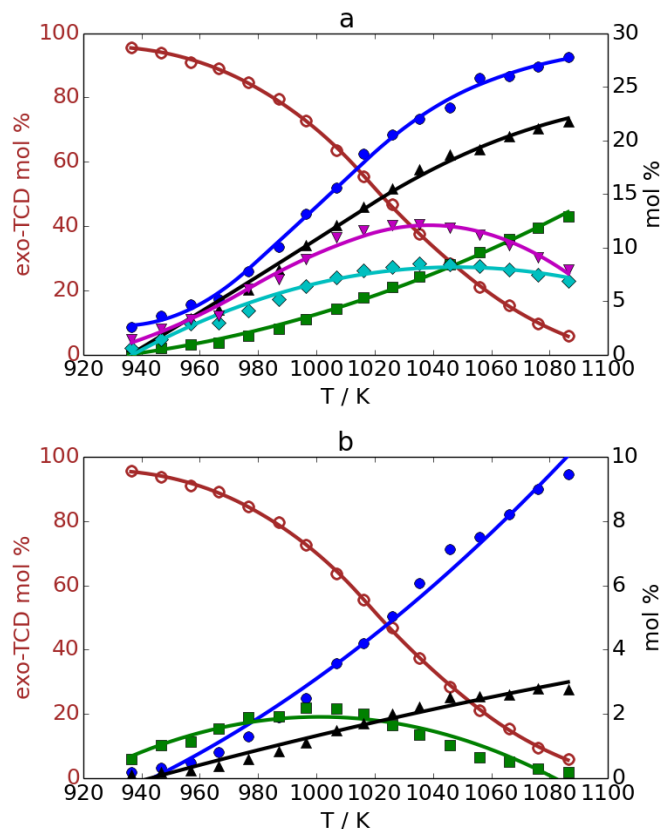


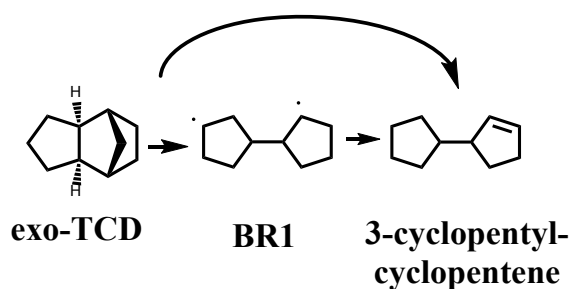
Figure 3-4: Normalized (excluding N₂) mole fractions as a function of temperature of a) ○: exo-TCD (left axis), ●: dihydrogen, ■: methane, ▲: ethene, ◆: propene, ▼: 1,3-cyclopentadiene b) ○ exo-TCD (left axis), ■: cyclopentene, ●: benzene, and ▲: toluene. Conditions: 10 mol % exo-TCD in N₂ at reactor inlet. $P = 1.7 \cdot 10^5$ Pa. Lines are spline interpolations.

Dihydrogen attains the highest mole fraction over the entire temperature range, up to 28 mol % at 1090K. Of the major products dihydrogen, ethene, CPD, and propene mole fractions are significantly higher at lower exo-TCD conversions than methane, benzene and toluene. Mole fractions of propene, CPD and cyclopentene attain maxima of 8%, 12%, and 2% at temperatures of 1045 K, 1040 K, and 990K respectively. The current results agree well with other studies [18, 20-22] that also confirmed the importance of these components as major products of exo-TCD pyrolysis.

3.4.2.2 Primary decomposition products

The 933K chromatogram shows the peak of *exo*-TCD along with feedstock components such as adamantane and *endo*-TCD, cf. Figure 3-3. A cloud of small peaks (C_9 - C_{10}) at the lower left of *exo*-TCD is visible. In this region both feedstock components and primary decomposition products were found with mole fractions of less than 1 mol %. These small quantities did not warrant further investigations. Lower molecular mass products such as ethene, propene, CPD, cyclopentene, benzene, and toluene are also distinguished.

The consumption of *exo*-TCD may occur by two types of reaction families: either unimolecular scission reactions of one of the carbon-carbon bonds of *exo*-TCD that opens one of the rings, or either by H-abstraction reactions by radicals resulting in tricyclodecyl monoradicals ($C_{10}H_{15}$) in which the original tricyclo[5.2.1.0^{2,6}]decane ring structure is preserved. The unimolecular scission of C – H bonds was assumed to be negligible given its significantly higher bond dissociation energy relative to the C – C bonds. Experiments with less than 4 mol % *exo*-TCD in argon by Herbinet et al. [22], observed only one $C_{10}H_{16}$ component, 3-cyclopentylcyclopentene, as a major primary decomposition product. This should indicate the importance of the C–C scission leading to the biradical containing the ring system of two non-fused five-membered rings, cf. R 3-1.



R 3-1

Recent quantum-chemical calculations suggest that R 3-1 may also proceed as one concerted step parallel to the reaction via a distinct biradical intermediate [42]. In the current study, 3-cyclopentylcyclopentene ($C_{10}H_{16}$), with retention time in between *exo*-TCD and *endo*-TCD [22], is not detected. Since experiments in the current study are performed at higher mole fractions of *exo*-TCD as compared to the experiments by Herbinet et al., bimolecular pathways such as hydrogen abstractions from *exo*-TCD gain importance and could explain the lesser formation of unimolecular decomposition products such as 3-cyclopentylcyclopentene. Adamantane is the

only observed component with retention time between exo-TCD and endo-TCD, cf. Figure 3-3A. Instead, significant quantities of a component with molecular formula $C_{10}H_{14}$ with a maximum mole fraction of 0.35 mol % at 975K are observed, cf. Figure 3-6A. The latter, retention time of 43.8 minutes, was tentatively identified as tricyclo[5.2.1.0^{2,6}]dec-4-ene ($C_{10}H_{14}$), based on the highest degree of similarity between the measured mass spectra and the library entry (NIST Mass Spectral Library v2.6), cf. Figure 3-5.

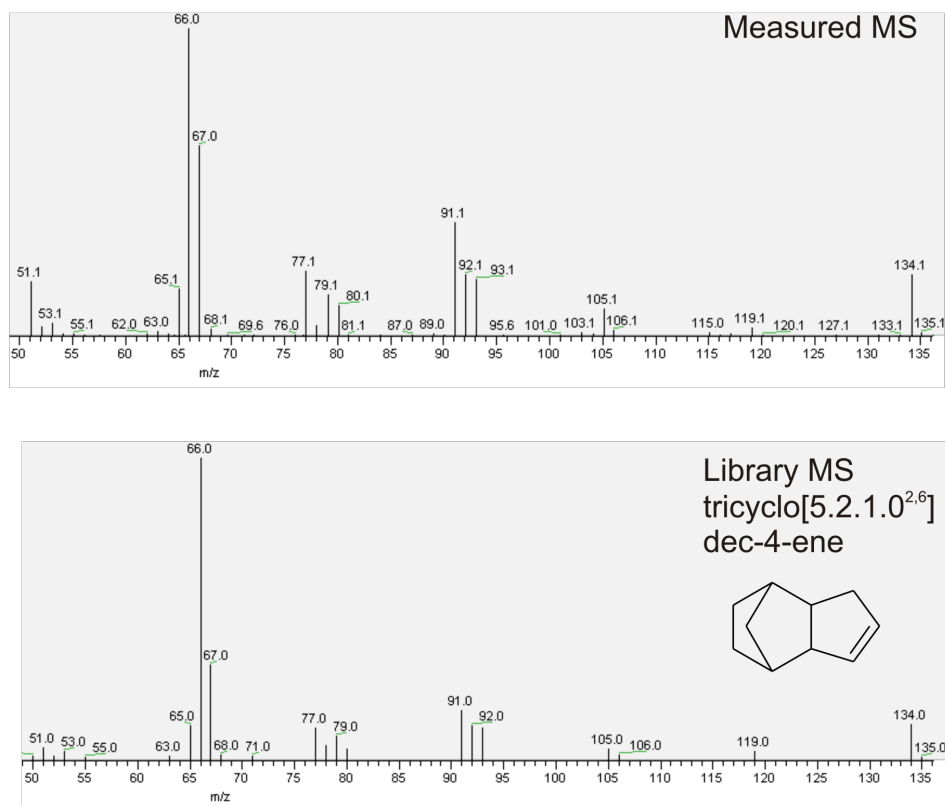
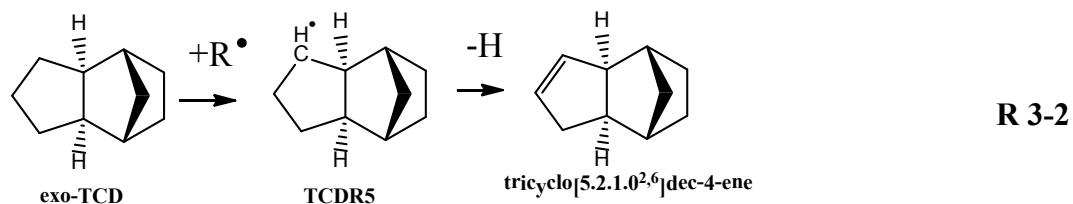


Figure 3-5: Measured and library mass spectra for tricyclo[5.2.1.0^{2,6}]dec-4-ene.

Both spectra showed molecular ions of 134 Da, with major signal peaks at $m/z = 66$ and 67 corresponding to C_5H_6 and C_5H_7 . Lower intensities were observed for $m/z = 91$, 92 and correspond to the norbornane-type (C_7H_7 , C_7H_8) ion fragments.

Computed C-H bond strengths of exo-TCD identified the C-H bond in the non-bridged cyclopentane ring in α -position to the bridgehead atom (cf. Figure 3-1) as the weakest C-H bond of exo-TCD ($BDE = 410 \text{ kJ mol}^{-1}$) [27]. The formation of tricyclo[5.2.1.0^{2,6}]dec-4-ene through

TCDR5 by hydrogen abstraction reactions followed by a β C-H scission reaction, cf. (R2) thus seems plausible.



Besides tricyclo[5.2.1.0^{2,6}]dec-4-ene, several products were quantified that give more information on the initial decomposition chemistry of exo-TCD, cf. Figure 3-6. The most important components that were tentatively identified are: 2-norbornene, bicyclo[3.3.0]octene, cyclopentene, 1-methyl-1,3-cyclopentadiene, fulvene, pentenes, 3-ethenylcyclopentene, and 1,3-butadiene.

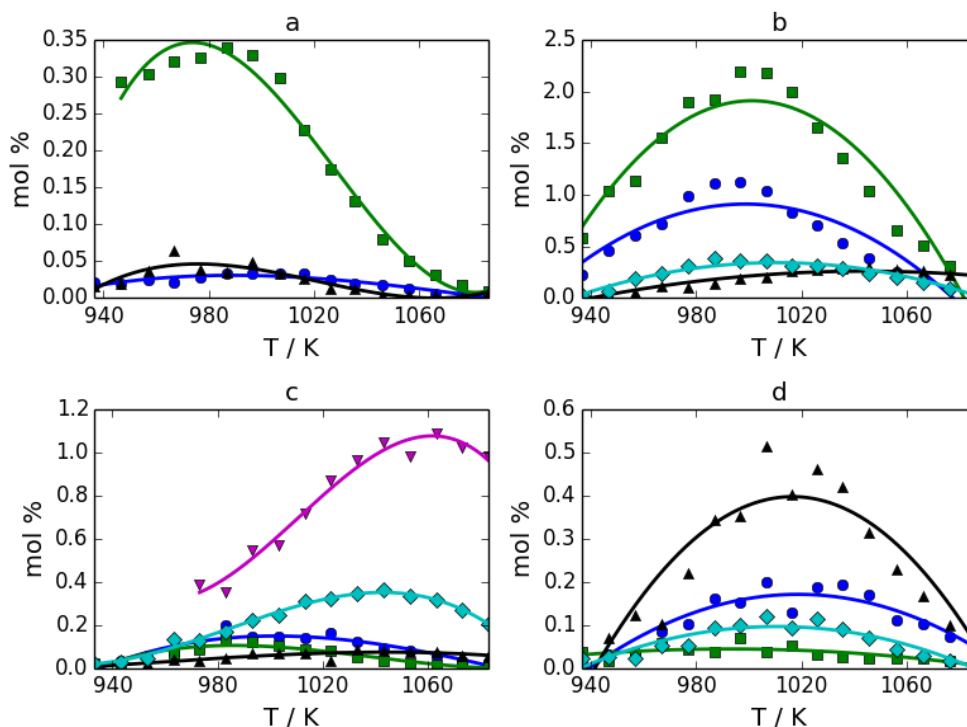


Figure 3-6: Normalized (excluding N_2) mole fractions as a function of mean reactor temperature of a) ●: bicyclo[3.3.0]octene, ■: tricyclo[5.2.1.0^{2,6}]-dec-4-ene, ▲: 2-norbornene b) ●: 3-ethenylcyclopentene, ■: cyclopentene, ▲: 1-methyl-1,3-cyclopentadiene, ◆: fulvene c) ●: 1-pentene, ■: 2-methyl-1-butene, ▲: 2-methyl-1,3-butadiene, ◆: hexadiene, ▼: 1,3-butadiene d) ●: 1,4-cyclohexadiene, ■: methylcyclohexadiene, ▲: methylenecyclohexadiene, ◆: cycloheptadiene. Conditions: 10 mol % exo-TCD in N_2 at reactor inlet. $P = 1.7 \times 10^5$ Pa. Experimental mole fractions of 1,3-butadiene were below the reproducible detection limit for temperatures between 930 – 970K. Lines are spline interpolations.

2-norbornene (C_7H_{10}) was detected in quantities of 0.05 mol % with a maximum at 980K, cf. Figure 3-6A. A component with molecular formula C_8H_{12} containing a bicyclo[3.3.0]-octane ring system was identified. Mole fractions of this component do not exceed 0.04 mol %, cf. Figure 3-6A. Besides CPD, other five-membered ring components were identified, cf. Figure 3-6B. Cyclopentene reaches a mole fraction of maximum 2 mol % at 990K. 1-Pentene was observed with mole fractions around 0.15 mol % while several pentadienes (C_5H_8) have similar mole fractions. A peak (retention time 31.4 min) to the right of toluene was identified as 3-ethenylcyclopentene (C_7H_{10}). Its mass spectra shows a molecular ion with $m/z = 94$, an intense

signal at $m/z = 79$ and signals at $m/z = 77, 80, 91$ and 93 with intensities similar to the library mass spectra, cf. Figure 3-7. 3-Ethenylcyclopentene (C_7H_{10}) attains mole fractions of 1 mol %.

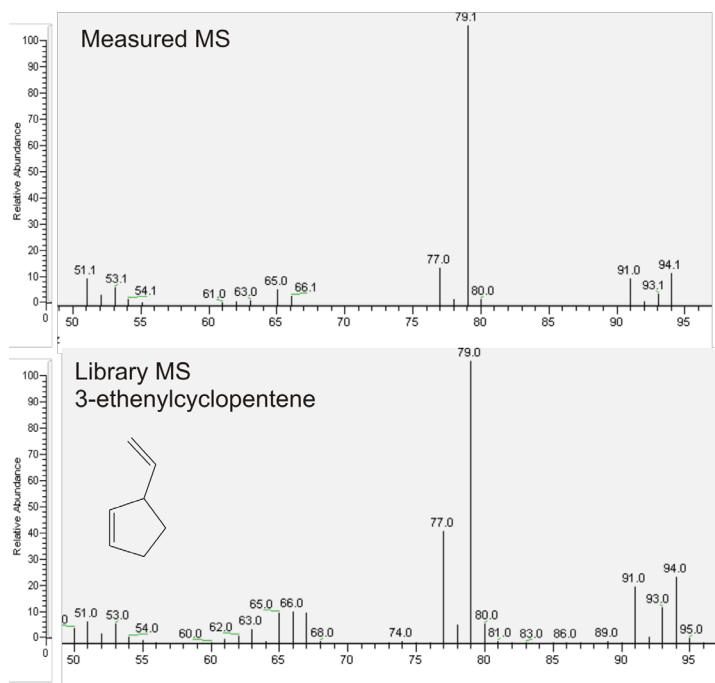


Figure 3-7: Measured and library mass spectra for 3-ethenylcyclopentene.

1,3-Butadiene is the most prominent C_4 hydrocarbon and attains mole fractions of up to 1 mol % at 1060K (cf. Figure 3-6C).

3.4.2.3 Secondary products

The 1083K chromatogram shows both the appearance of peaks between CPD and exo-TCD and an additional trail of higher molecular mass components with elevated second dimension retention times that elute later than exo-TCD. The sum of all components containing at least one aromatic ring amounts to 18 mol% at 1080K, cf. Figure 3-9B. Among the aromatic hydrocarbons, a pronounced selectivity towards benzene, toluene, and styrene was observed with mole fractions at the highest temperature of 10%, 3% and 0.8% respectively. These mono-aromatic hydrocarbons comprise at least 70 mol % of the entire aromatic hydrocarbons fraction throughout the entire conversion range. Ethylbenzene exhibits a maximum mole fraction at 1040K of 0.15 mol %. Ethylbenzene consumption is explained by the activation of the C-H in α -position of the aromatic ring that becomes subject to H-abstraction reactions forming a benzylic radical that

undergoes consecutive β -scission to styrene. Xylenes were present in quantities comparable to ethylbenzene with mole fractions leveling off at higher exo-TCD conversions too.

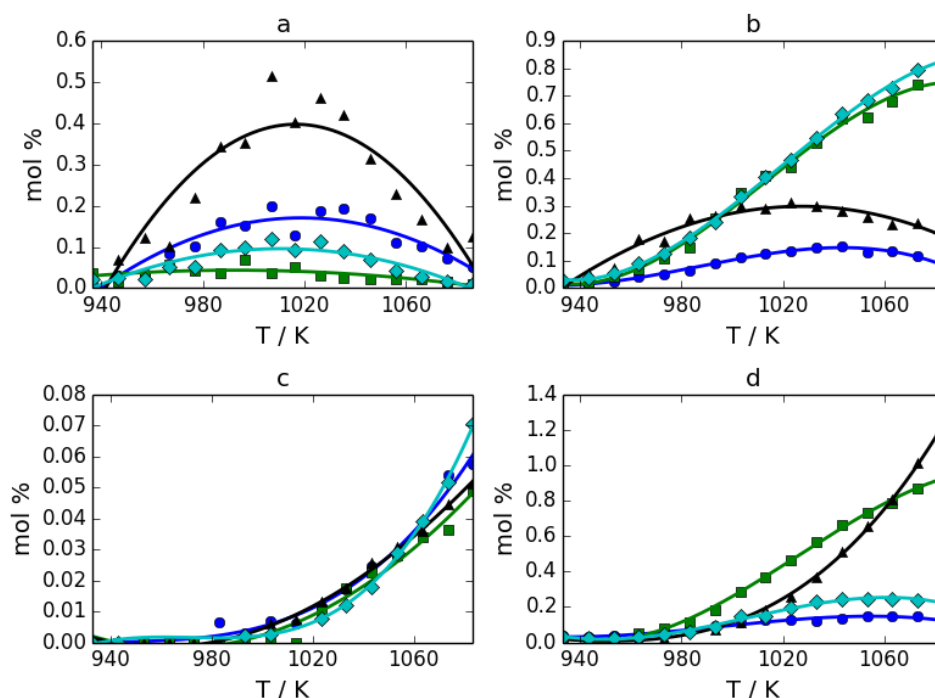
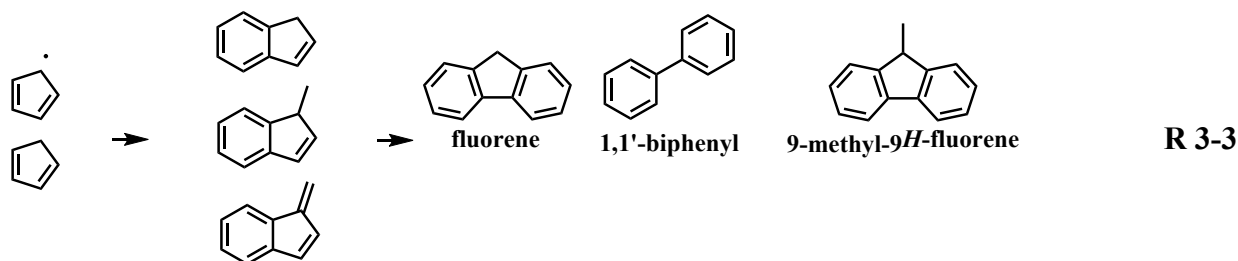


Figure 3-8: Normalized (excluding N₂) mole fractions as a function of mean reactor temperature of a) ●: 1,4-cyclohexadiene, ■: methylhexadiene, ▲: methylidenecyclohexadiene, ◆: cycloheptadiene b) ●: ethylbenzene, ■: styrene, ▲: C₃-benzenes, ◆: the entire aromatics lump divided by 20, c) ●: Biphenyl, ■: methylfluorenes, ▲: fluorene, ◆: anthracene d) ●: methylindenes, ■: indene, ▲: naphthalene, ◆: benzofulvene. Conditions: 10 mol % exo-TCD in N₂ at reactor inlet. $P = 1.7 \cdot 10^5$ Pa. Lines are spline interpolations.

A number of non-aromatic intermediate components containing a six-membered ring, such as methyl- and methylene substituted cyclohexadienes, and 1,4-cyclohexadiene were quantified as well, cf. Figure 3-9A. All of them exhibit maxima in the profiles of mole fractions around 1020 K in the range of 0.1 - 0.5 mol %. In addition to the lower molecular weight species formed by the decomposition pathways discussed above, more than 20 individual polycyclic aromatic hydrocarbons (PAHs) with up to fourteen carbons were detected and quantified. Observable quantities of PAHs were detected at temperatures starting from 980 K. A total cumulative mole fraction of less than 3 mol % at 1090 K is measured, with indene, and naphthalene the most prominently present.

The significant decline in CPD mole fractions occurs at a temperature where the PAHs are formed, suggesting that the formation of the observed PAHs is governed by successive growth chemistry initiated from abundant five-membered rings. Indene and naphthalene attain effluent mole fractions of 1 mol % at 1090K, cf. Figure 3-9D. Other bicyclic PAHs such as methyl-substituted indenenes, methyl-substituted naphthalenes, and biphenyl were found with mole fractions not exceeding 0.5 mol%. The relative position of methyl substituents on the indene- and naphthalene-ring systems could not be determined with certainty. Indene-derived structures are formed at slightly lower temperatures than naphthalene. The sum of all methylindenenes reaches a maximum mole fraction of 0.5 mol% at 1065K while the total mole fraction of all methylnaphthalenes monotonously increases. A crossover of indene and naphthalene mole fractions was observed at a reactor temperature of 1070K. A similar crossover temperature was reported in CPD pyrolysis experiments and predicted by computational studies of CPD pyrolysis [43, 44]. The CPD derived 1,3-cyclopentadienyl radical is known to play a significant role in the formation of polycyclic aromatic hydrocarbons (PAHs) and soot [45-50]. For example, Butler and Glassmann found that naphthalene comprises almost 80% of the product spectrum during CPD pyrolysis at 1200K and atmospheric pressure [51], supporting this hypothesis. Computational studies also indicate the importance of 1,3-cyclopentadienyl self-recombination [52] and 1,3-cyclopentadienyl addition to CPD [53, 54] as pathways to naphthalene. Indene was also observed as a major product of CPD pyrolysis [51]. Many routes from lower molecular mass moieties were suggested such as the 1,3-cyclopentadienyl addition to CPD with methyl elimination (e.g. [53, 55]), and successive ethyne additions to CPD through seven-membered cyclic moieties [56-58], or via 1,3-cyclopentadienyl self-recombination [59].

Tricyclic PAHs such as biphenyl, anthracene, fluorene, and methyl-substituted fluorenes were detected in quantities of less than 0.1 mol %, cf. Figure 3-9C. The exo-TCD conversion corresponding to the initial appearance of these tricyclic PAHs is slightly higher than for bicyclic PAHs and coincides with the emergence of ethyne with comparable mole fractions. It is plausible that they are formed through successive growth of pre-existing bicyclic aromatic moieties such as indene derived radicals by lower molecular mass species as suggested by many PAH growth and soot formation mechanisms [47, 60], cf. **R 3-3**.



3.4.2.4 Effect of dilution

Figure 3-10 shows the effect of dilution on product mole fractions as a function of exo-TCD conversion is shown for a selection of components. The components that were not depicted have no observable differences for the two sets of experiments.

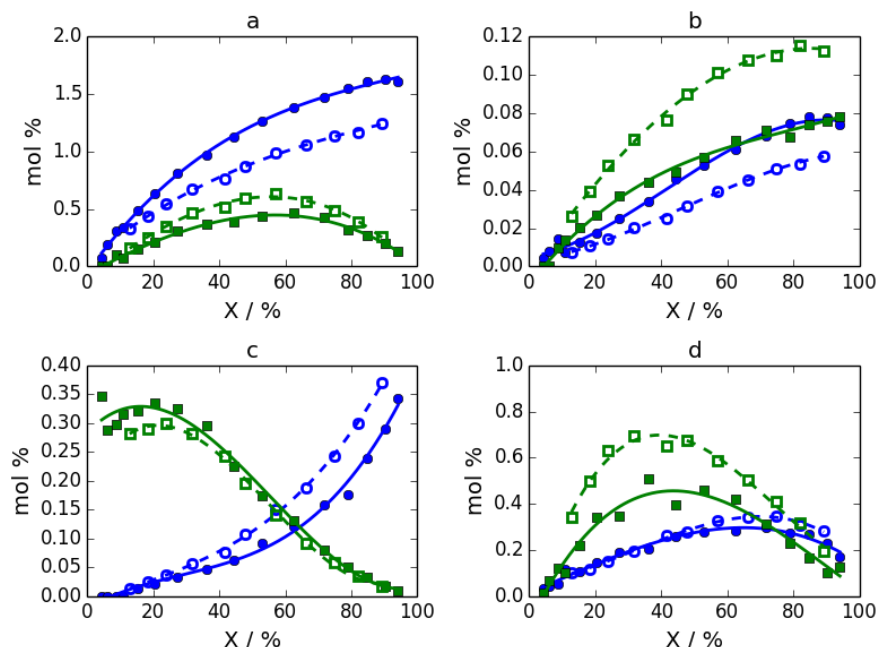


Figure 3-9: Normalized (excluding N_2) mole fractions as a function of exo-TCD conversion of a) ●: ethane, ■: 1-butene b) ●: propane, ■: propadiene c) ●: ethyne, ■: tricyclo[5.2.1.0^{2,6}]-dec-4-ene d) ●: methyl-1,3-cyclopentadiene, ■: methylenecyclohexadiene. Conditions: 10 mol % exo-TCD (—, full symbols) and 7 mol % exo-TCD (- -, hollow symbols) in N_2 at reactor inlet. $P = 1.7 \cdot 10^5$ Pa. Lines are spline interpolations.

Saturated species such as ethane and propane attain higher selectivities at lower dilutions. The latter imply higher radical concentrations and thus elevated rates for bimolecular reactions. Inverse trends, i.e. higher selectivities were observed at the higher dilution level, were observed

for species such as ethyne, and propadiene. The higher selectivity for tricyclo[5.2.1.0^{2,6}]dec-4-ene for lower levels of dilutions is significant at lower exo-TCD conversion (cf. Figure 3-10C) and shows the favorable effect of higher hydrocarbon partial pressures on bimolecular exo-TCD conversions compared to the competing unimolecular C–C scission reactions leading to biradicals such as BR1.

Mole fractions for the toluene precursor, methylenecyclohexadiene cf. Figure 3-10D, show significant differences as a function of the level of dilution, cf. Figure 3-10D. Since the exo-TCD decomposition pathway to methylenecyclohexadiene involves saturated cycloalkane rings, the formation of this component must proceed through a series of both H-abstraction reactions and C–H β -scission reactions.

3.4.3 Kinetic model validation

Figure 3-12 shows the exo-TCD conversion as a function of the mean reactor temperature of the current experimental data versus the predictions of the San Diego model [37]. Exo-TCD conversion is significantly underestimated, which is not surprising given that the model was built for combustion applications at elevated temperatures of 1500K and above. Noticeable JP-10 conversion only occurs at temperatures higher than 1000K which does not agree with experiment.

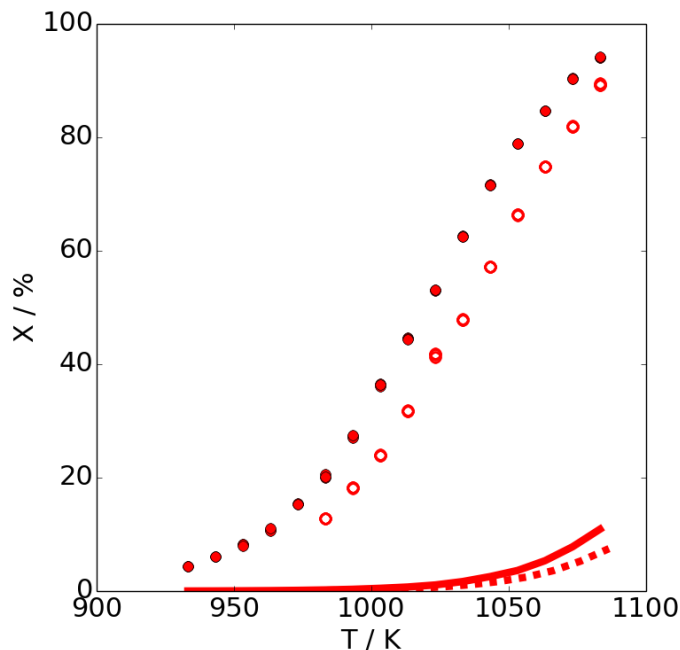
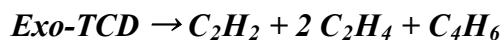


Figure 3-10: JP-10 conversion as a function of temperature for two dilution levels of JP-10 in N_2 . Experimental data: ●: 10 mol%, and ○: 7 mol % JP-10 in N_2 . Predictions of San Diego model [37]: —: 10 mol%, and - -: 7 mol % JP-10 in N_2 . $P = 1.7 \cdot 10^5$ Pa.

Figure 3-12 shows the mole fractions as a function of the mean reactor temperature of some major products such as dihydrogen, ethene and propene of the current experimental data versus the predictions of the San Diego model [37]. Major products such as dihydrogen, ethene and propene are significantly underpredicted at the operating conditions of the experimental data, which is attributed to the underprediction of conversion of *exo*-TCD. At temperatures above 1000K the mole fraction of ethene increases rapidly. Similar trends are seen for ethyne, 1,3-butadiene, and propyne (Figure 3-12b), while model predictions significantly overpredict experimentally observed values. The pronounced selectivity for species such as ethyne, ethene, and 1,3-butadiene is attributed to the presence of global decomposition reactions of *exo*-TCD towards these species that appear in the model, e.g. **R 3-4**:



R 3-4

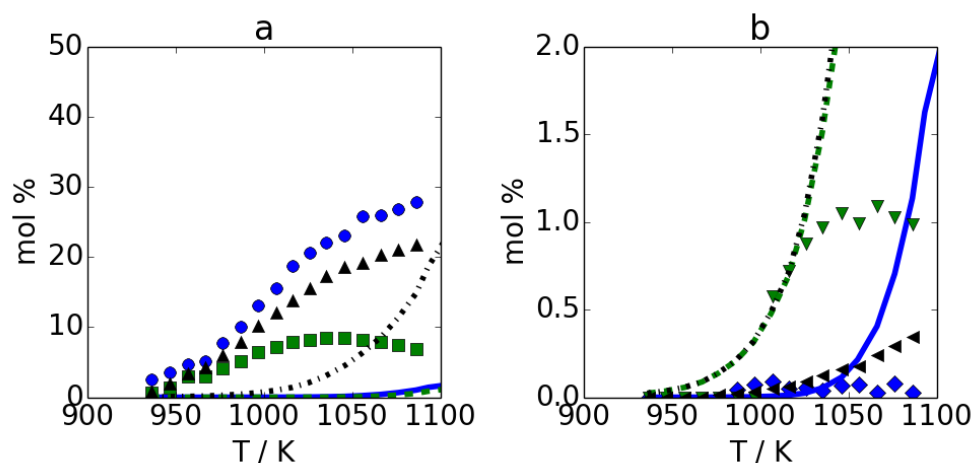


Figure 3-11: Mole fractions as a function of temperature for a) experimental: ●: dihydrogen, ▲: ethene, ■: propene, predictions of San Diego [37] model: —: dihydrogen, - -: ethene, - -: propene, b) experimental: ♦: propyne, ▼: 1,3-butadiene, ◄: ethyne, predictions of San Diego [37] model: —: propyne, - -: 1,3-butadiene, - -: ethyne. Conditions: 10 mol % JP-10 in N₂ at reactor inlet. $P = 1.7 \cdot 10^5$ Pa.

Figure 3-14 shows the JP-10 conversion as a function of temperature of the current experimental data versus the predictions of the RMG model [1]. The RMG model results in better model predictions for the exo-TCD conversion in the current pyrolysis experiments than the San Diego model, but still slightly underestimates the conversion.

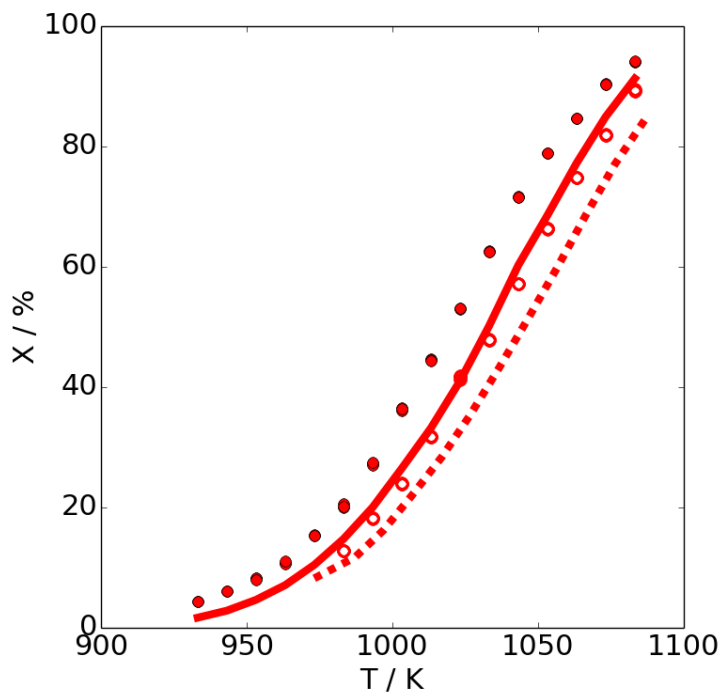


Figure 3-12: JP-10 conversion as a function of temperature for two dilution levels of JP-10 in N_2 . Experimental data: \bullet : 10 mol%, and \circ : 7 mol % JP-10 in N_2 . Predictions of the RMG model [1]: —: 10 mol%, and - -: 7 mol % JP-10 in N_2 . $P = 1.7 \cdot 10^5$ Pa.

Especially at elevated conversions many of the major products, especially ethene, are underpredicted compared to the observed mole fractions. For CPD, the maximum at 1030K is not present in the model calculations because CPD decomposition reactions are missing in the model. This also explains the absence of important aromatics in the model such as benzene, indene and naphthalene, because they are primarily formed by the consumption of CPD. Also, many of the detected intermediate products, such as tricyclo[5.2.1.0^{2,6}]dec-4-ene, are not present in the model.

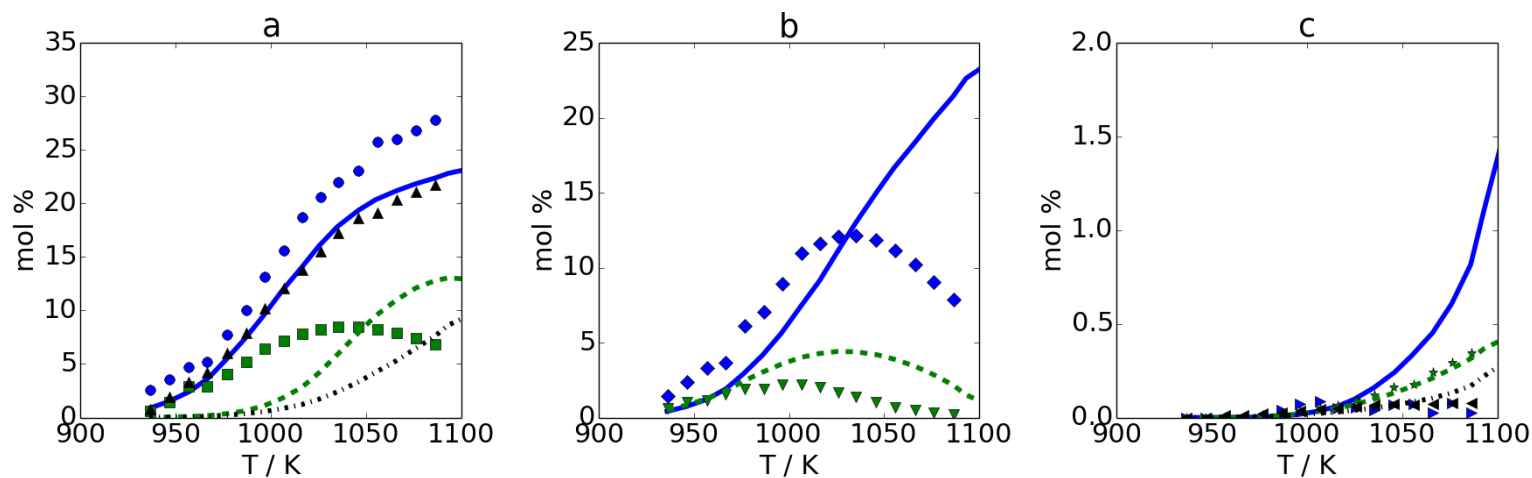
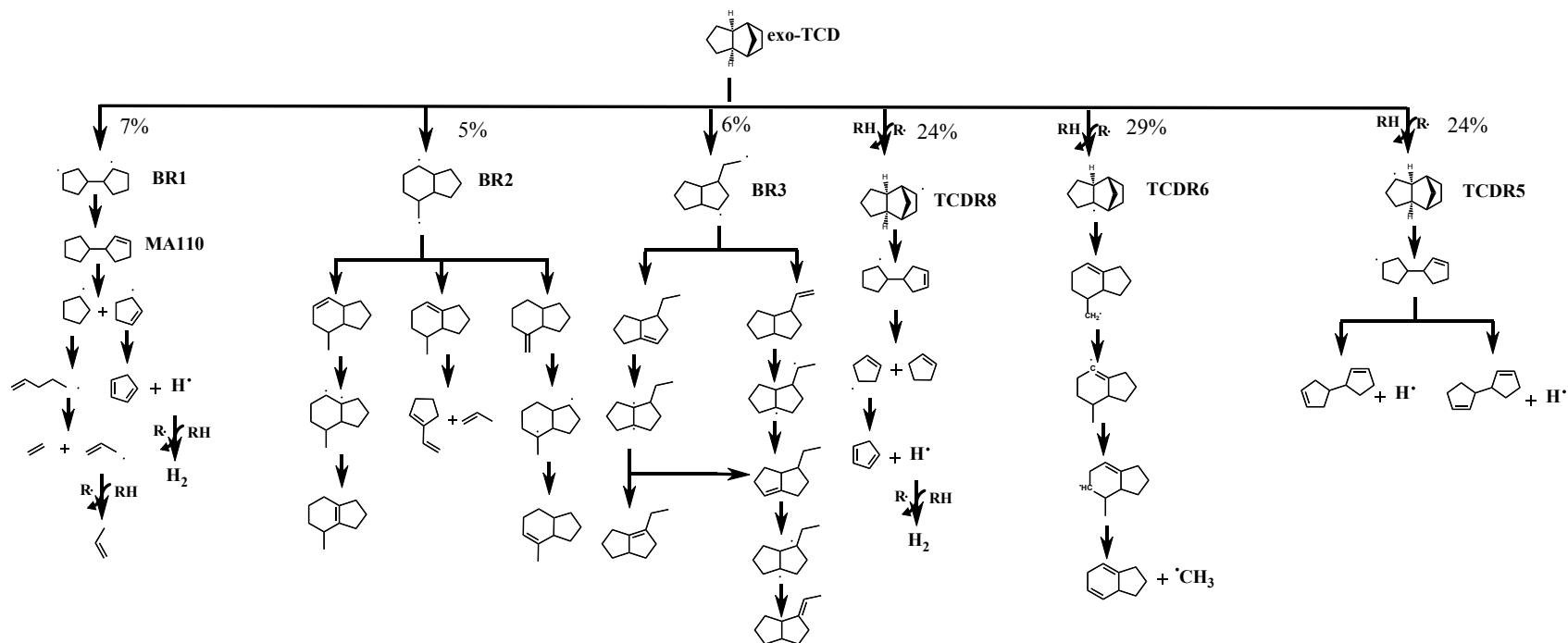


Figure 3-13: Mole fractions as a function of temperature for a) experimental: \bullet : dihydrogen, \blacktriangle : ethene, \blacksquare : propene, predictions of RMG model [1]: —: dihydrogen, - -: ethene, - -: propene, b) experimental: \blacklozenge : 1,3-cyclopentadiene, \blacktriangledown : cyclopentene, predictions of RMG model [1]: —: 1,3-cyclopentadiene, - -: cyclopentene, c) experimental: \blacktriangleright : propyne, \blacktriangleleft : propadiene, $*$: ethyne, predictions of RMG model [1]: —: propyne, - -: propadiene, - -: ethyne. Conditions: 10 mol % JP-10 in N_2 at reactor inlet. $P = 1.7 \cdot 10^5$ Pa.

A rate of production plot is shown in Scheme 3-1 for the RMG model [1], and corresponds to conditions of the current experimental data. At a conversion of exo-TCD of 17%, exo-TCD is mainly consumed by three H-abstraction reactions, whereas the fraction of exo-TCD converted through unimolecular initiation routes only amounts to 18%. Because very little information is present on the kinetics and thermochemistry of reactions and species relevant in exo-TCD decomposition, significant uncertainties can be associated to the parameters that were used to build the model. For example, increasing the rate of the 3 most important H-abstraction reactions by H from exo-TCD by a factor of two, leads to good agreement between simulated and experimental exo-TCD conversions.

Furthermore, the reactant in both models was represented as 100% exo-TCD, thus neglecting the 5% endo-TCD, because the endo isomer was not present in the model. Since the endo-isomer is more unstable compared to the exo-isomer, it is possible that the initiation of exo-TCD is influenced by radicals originating from endo-TCD. Future models will need to investigate this in more detail.



Scheme 3-1: Reaction pathway analysis for the decomposition of exo-TCD with the RMG model [1]. The reported percentages represent the reaction rate relative to the total exo-TCD decomposition rate. Conditions: $T=1100\text{K}$, $P=1.7 \cdot 10^5 \text{ Pa}$, $\text{Conversion}_{\text{exo-TCD}} = 17\%$, 10 mol% exo-TCD in N_2 at the reactor inlet corresponding to conditions of the experimental data of this work. Only major exo-TCD decomposition pathways are shown.

3.5 Conclusions

New experimental data was presented in this work for exo-TCD pyrolysis under mildly diluted conditions. The data shows that exo-TCD starts to decompose at 900K and attains full conversion at 1100K at the investigated conditions. Hydrogen, ethene, cyclopentadiene and benzene were among the major decomposition products of exo-TCD in line with literature and comprised ca. 75 mol % of the product effluent at the 1080K. The presence of tricyclo[5.2.1.0^{2,6}]dec-4ene as a primary product suggested that bimolecular exo-TCD conversion pathways become important and contrast earlier findings of 3-cyclopentylcyclopentene as the main initial decomposition product. A number of initial decomposition products, that were tentatively identified as 3-ethenylcyclopentene, 2-norbornene, and a bicyclo[3.3.0]octene isomer were quantified.

Important pathways to PAHs were identified to start from CPD. These reactions result in the formation of naphthalene, indene, and substituted derivatives of these bicyclic aromatic components. Indene derived moieties are further converted into tricyclic PAHs such as fluorenes and confirmed the propensity of strained-ring hydrocarbons to soot, and similar pyrolytic deposition in hypersonic engine designs. C₁-substituted 1,3-cyclopentadienes such as methyl-1,3-cyclopentadiene and fulvene further indicated the importance of the rearrangements of five-membered rings into aromatic moieties.

Comparison of the current experimental data against kinetic models from literature of the thermal decomposition of JP-10 showed that none of them adequately describes the product distribution trends. A new dedicated kinetic model for the breakdown of JP-10 under pyrolysis conditions is therefore imperative and is discussed in a Chapter 4 of this thesis.

3.6 References

- [1] G.R. Magoon, J. Aguilera-Iparraguirre, W.H. Green, J.J. Lutz, P. Piecuch, H.W. Wong, O.O. Oluwole, Detailed chemical kinetic modeling of JP-10 (exo-tetrahydrodicyclopentadiene) high-temperature oxidation: Exploring the role of biradical species in initial decomposition steps, *International Journal of Chemical Kinetics*, 44 (2012) 179-193.
- [2] L.Q. Maurice, H. Lander, T. Edwards, W.E. Harrison, Advanced aviation fuels: a look ahead via a historical perspective, *Fuel*, 80 (2001) 747-756.
- [3] T. Edwards, USAF Supercritical Hydrocarbon Fuels Interests, 31 st Aerospace Sciences Meeting & Exhibit, (1993).

- [4] T. Edwards, Advancements in Gas Turbine Fuels From 1943 to 2005, *Journal of Engineering for Gas Turbines and Power*, 129 (2007) 13-20.
- [5] H.S. Chung, C.S.H. Chen, R.A. Kremer, J.R. Boulton, G.W. Burdette, Recent developments in high-energy density liquid hydrocarbon fuels, *Energy & Fuels*, 13 (1999) 641-649.
- [6] B. Van Devener, S.L. Anderson, Breakdown and combustion of JP-10 fuel catalyzed by nanoparticulate CeO₂ and Fe₂O₃, *Energy & Fuels*, 20 (2006) 1886-1894.
- [7] A. Osmont, I. Gokalp, L. Catoire, Evaluating missile fuels, *Propellants Explosives Pyrotechnics*, 31 (2006) 343-354.
- [8] H.A. Meylemans, R.L. Quintana, B.G. Harvey, Efficient conversion of pure and mixed terpene feedstocks to high density fuels, *Fuel*, 97 (2012) 560-568.
- [9] M. Cooper, J.E. Shepherd, Thermal and Catalytic Cracking of JP-10 for Pulse Detonation Engine Applications, (2002)
- [10] S.P. Heneghan, W.E. Harrison, JP-8+100: The Development of High Thermal Stability Jet Fuel, 6th International Conference on Stability and Handling of Liquid Fuels, (1997).
- [11] H. Huang, D.R. Sobel, L.J. Spadaccini, Endothermic Heat-Sink of Jet Fuels for Scramjet Cooling, 38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, (2002).
- [12] Y. Xing, D. Li, W.J. Xie, W.J. Fang, Y.S. Guo, R.S. Lin, Catalytic cracking of tricyclo [5.2.1.0(2,6)] decane over HZSM-5 molecular sieves, *Fuel*, 89 (2010) 1422-1428.
- [13] L.-y. Hou, N. Dong, D.-p. Sun, Heat transfer and thermal cracking behavior of hydrocarbon fuel, *Fuel*, 103 (2013) 1132-1137.
- [14] H.A. Meylemans, L.C. Baldwin, B.G. Harvey, Low-Temperature Properties of Renewable High-Density Fuel Blends, *Energy & Fuels*, 27 (2013) 883-888.
- [15] W. Xie, W. Fang, D. Li, Y. Xing, Y. Guo, R. Lin, Coking of Model Hydrocarbon Fuels under Supercritical Condition, *Energy & Fuels*, 23 (2009) 2997-3001.
- [16] T.J.H. Bruno, M.L.; Laesecke, A.; Lemmon, E.W.; Perkins, R.A., Thermochemical and Thermophysical Properties of JP-10, Report Prepared for: Fuels Branch, Turbine Engine Division Propulsion Directorate, AFRL Wright Patterson Air Force Base, Ohio, (2006)
- [17] D.F. Davidson, D.C. Horning, M.A. Oehlschlaeger, R.K. Hanson, The Decomposition Products of JP-10, *American Institute of Aeronautics and Astronautics Journal*, 3707 (2001).
- [18] K. Wohlwend, L.Q. Maurice, T. Edwards, R.C. Striebich, M. Vangsness, A.S. Hill, Thermal stability of energetic hydrocarbon fuels for use in combined cycle engines, *J. Propul. Power*, 17 (2001) 1258-1262.
- [19] R.C. Striebich, J. Lawrence, Thermal decomposition of high-energy density materials at high pressure and temperature, *Journal of Analytical and Applied Pyrolysis*, 70 (2003) 339-352.
- [20] S. Nakra, R.J. Green, S.L. Anderson, Thermal decomposition of JP-10 studied by micro-flowtube pyrolysis-mass spectrometry, *Combustion and Flame*, 144 (2006) 662-674.
- [21] P.N. Rao, D. Kunzru, Thermal cracking of JP-10: Kinetics and product distribution, *Journal of Analytical and Applied Pyrolysis*, 76 (2006) 154-160.
- [22] O. Herbinet, B. Sirjean, R. Bounaceur, R. Fournet, F. Battin-Leclerc, G. Scacchi, P.M. Marquaire, Primary mechanism of the thermal decomposition of tricyclodecane, *Journal of Physical Chemistry A*, 110 (2006) 11298-11314.
- [23] Y. Xing, W.J. Fang, W.J. Xie, Y.S. Guo, R.S. Lin, Thermal Cracking of JP-10 under Pressure, *Industrial & Engineering Chemistry Research*, 47 (2008) 10034-10040.
- [24] C.R. Council, Handbook of Aviation Fuel Properties, 1983.
- [25] M.-Y. Huang, J.-C. Wu, F.-S. Shieu, J.-J. Lin, Preparation of high energy fuel JP-10 by acidity-adjustable chloroaluminate ionic liquid catalyst, *Fuel*, 90 (2011) 1012-1017.
- [26] J.L.M. Abboud, O. Castano, J.Z. Davalos, R. Gomperts, The standard enthalpies of formation of 1-and 2-Adamantyl cations and radicals. An ab initio study, *Chem. Phys. Lett.*, 337 (2001) 327-330.
- [27] J.M. Hudzik, R. Asatryan, J.W. Bozzelli, Thermochemical Properties of exo-Tricyclo 5.2.1.0(2,6) decane (JP-10 Jet Fuel) and Derived Tricyclodecyl Radicals, *Journal of Physical Chemistry A*, 114 (2010) 9545-9553.
- [28] A. Fattahi, L. Lis, Z.A. Tehrani, S.S. Marimanikkuppam, S.R. Kass, Experimental and Computational Bridgehead C-H Bond Dissociation Enthalpies, *Journal of Organic Chemistry*, 77 (2012) 1909-1914.
- [29] J. Gasteiger, O. Dammer, Automatic Estimation of Ring Strain Energies, *Tetrahedron*, 34 (1978) 2939-2945.
- [30] R.H. Boyd, S.N. Sanwal, Sharyteh.S, D. McNally, Thermochemistry, Thermodynamic Functions, and Molecular Structures of some Cyclic Hydrocarbons, *Journal of Physical Chemistry*, 75 (1971) 1264-1271.

- [31] S. Chang, R.H. Boyd, D. McNally, Sharyteh.S, M.J. Hickey, Heats of Combustion and Strain Energies of Bicyclo[N.M.0]-Alkanes, *J. Am. Chem. Soc.*, 92 (1970) 3109-3122.
- [32] Y.-R. Luo, *Handbook of Bond Dissociation Energies in Organic Compounds*, Taylor & Francis, 2002.
- [33] M.R. Harper, K.M. Van Geem, S.P. Pyl, G.B. Marin, W.H. Green, Comprehensive reaction mechanism for n-butanol pyrolysis and combustion, *Combustion and Flame*, 158 (2011) 16-41.
- [34] K.M. Van Geem, S.P. Pyl, G.B. Marin, M.R. Harper, W.H. Green, Accurate High-Temperature Reaction Networks for Alternative Fuels: Butanol Isomers, *Industrial & Engineering Chemistry Research*, 49 (2010) 10399-10420.
- [35] K.M. Van Geem, A. Cuoci, A. Frassoldati, S.P. Pyl, G.B. Marin, E. Ranzi, An Experimental and Kinetic Modeling Study of Pyrolysis and Combustion of Acetone-Butanol-Ethanol (ABE) Mixtures, *Combustion Science and Technology*, 184 (2012) 942-955.
- [36] G.C. Reyniers, G.F. Froment, F.D. Kopinke, G. Zimmermann, Coke formation in the Thermal Cracking of Hydrocarbons. 4. Modeling of coke formation in naphtha cracking, *Industrial & Engineering Chemistry Research*, 33 (1994) 2584-2590.
- [37] S.C. Li, B. Varatharajan, F.A. Williams, Chemistry of JP-10 ignition, *American Institute of Aeronautics and Astronautics J.*, 39 (2001) 2351-2356.
- [38] B. Varatharajan, F.A. Williams, Ethylene ignition and detonation chemistry, Part 1: Detailed modeling and experimental comparison, *J. Propul. Power*, 18 (2002) 344-351.
- [39] G.R. Magoon, W.H. Green, 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*, 52 (2013) 35-45.
- [40] J. Song, S. Raman, J. Yu, C.D. Wijaya, G. Stephanopoulos, W.H. Green, Development of automatic chemical reaction mechanism generation software using object-oriented technology., *Abstracts of Papers of the American Chemical Society*, 226 (2003) U530-U531.
- [41] CHEMKIN-PRO v150101, Reaction Design, San Diego, (2009).
- [42] G.R. Magoon, J. Aguilera-Iparraguirre, W.H. Green, J.J. Lutz, P. Piecuch, H.-W. Wong, O.O. Oluwale, Detailed chemical kinetic modeling of JP-10 (exo-tetrahydrodicyclopentadiene) high-temperature oxidation: Exploring the role of biradical species in initial decomposition steps, *International Journal of Chemical Kinetics*, 44 (2012) 179-193.
- [43] D.H. Kim, J.A. Mulholland, D. Wang, A. Violi, Pyrolytic Hydrocarbon Growth from Cyclopentadiene, *Journal of Physical Chemistry A*, 114 (2010) 12411-12416.
- [44] M.M. Lu, J.A. Mulholland, PAH growth from the pyrolysis of CPD, indene and naphthalene mixture, *Chemosphere*, 55 (2004) 605-610.
- [45] A. D'Anna, A. Violi, Detailed modeling of the molecular growth process in aromatic and aliphatic premixed flames, *Energy & Fuels*, 19 (2005) 79-86.
- [46] M.S. Skjoth-Rasmussen, P. Glarborg, M. Ostberg, J.T. Johannessen, H. Livbjerg, A.D. Jensen, T.S. Christensen, Formation of polycyclic aromatic hydrocarbons and soot in fuel-rich oxidation of methane in a laminar flow reactor, *Combustion and Flame*, 136 (2004) 91-128.
- [47] H. Richter, J.B. Howard, Formation of polycyclic aromatic hydrocarbons and their growth to soot - a review of chemical reaction pathways, *Progress in Energy and Combustion Science*, 26 (2000) 565-608.
- [48] P. Lindstedt, L. Maurice, M. Meyer, Thermodynamic and kinetic issues in the formation and oxidation of aromatic species, *Faraday Discussions*, 119 (2001) 409-432.
- [49] N.M. Marinov, W.J. Pitz, C.K. Westbrook, A.M. Vincitore, M.J. Castaldi, S.M. Senkan, C.F. Melius, Aromatic and polycyclic aromatic hydrocarbon formation in a laminar premixed n-butane flame, *Combustion and Flame*, 114 (1998) 192-213.
- [50] A. Lamprecht, B. Atakan, K. Kohse-Hoinghaus, Fuel-rich flame chemistry in low-pressure cyclopentene flames, *Proceedings of the Combustion Institute*, 28 (2000) 1817-1824.
- [51] R.G. Butler, I. Glassman, Cyclopentadiene combustion in a plug flow reactor near 1150 K, *Proc. Combust. Inst.*, 32 (2009) 395-402.
- [52] C.F. Melius, M.E. Colvin, N.M. Marinov, W.J. Pitz, S.M. Senkan, Reaction mechanisms in aromatic hydrocarbon formation involving the C₅H₅ cyclopentadienyl moiety, 26th International Symposium on Combustion, Naples, Italy, 1996.
- [53] D. Wang, A. Violi, D.H. Kim, J.A. Mulholland, Formation of naphthalene, indene, and benzene from cyclopentadiene pyrolysis: A DFT study, *Journal of Physical Chemistry A*, 110 (2006) 4719-4725.

- [54] R.W. Alder, S.P. East, J.N. Harvey, M.T. Oakley, The azulene-to-naphthalene rearrangement revisited: a DFT study of intramolecular and radical-promoted mechanisms, *J. Am. Chem. Soc.*, 125 (2003) 5375-5387.
- [55] V.V. Kislov, A.M. Mebel, An ab initio G3-type/statistical theory study of the formation of indene in combustion flames. II. The pathways originating from reactions of cyclic C-5 species - Cyclopentadiene and cyclopentadienyl radicals, *Journal of Physical Chemistry A*, 112 (2008) 700-716.
- [56] C. Cavallotti, S. Mancarella, R. Rota, S. Carra Conversion of C5 into C6 Cyclic Species through the Formation of C7 Intermediates *The Journal of Physical Chemistry A*, 111 (2007) 3959-3969.
- [57] S. Fascella, C. Cavallotti, R. Rota, S. Carra, The peculiar kinetics of the reaction between acetylene and the cyclopentadienyl radical, *Journal of Physical Chemistry A*, 109 (2005) 7546-7557.
- [58] S. Fascella, C. Cavallotti, R. Rota, S. Carra Quantum Chemistry Investigation of Key Reactions Involved in the Formation of Naphthalene and Indene, *The Journal of Physical Chemistry A*, 108 (2004) 3829-3843.
- [59] R.K. Robinson, R.P. Lindstedt, On the chemical kinetics of cyclopentadiene oxidation, *Combustion and Flame*, 158 (2011) 666-686.
- [60] H. Wang, M. Frenklach, A detailed kinetic modeling study of aromatics formation in laminar premixed acetylene and ethylene flames, *Combustion and Flame*, 110 (1997) 173-221.

Chapter 4: Kinetic modeling of Jet Propellant-10 pyrolysis

4.1 Abstract

The thermal decomposition of Jet-Propellant 10 (JP-10), consisting of tricyclic $C_{10}H_{16}$ exo-tricyclo[5.2.1.0^{2,6}]decane (exo-TCD), is modeled using the automated mechanism generator RMG. The kinetic model contains 5261 reactions between 384 species. Rate coefficients for important unimolecular initiation routes of exo-TCD are calculated using the multireference method CAS-PT2, while rate coefficients for the various primary decompositions of the exo-TCD derived monoradicals are obtained using CBS-QB3. The model predictions agree well with five independent experimental data sets of JP-10 pyrolysis that cover a wide range of operating conditions ($T = 300\text{--}1500\text{K}$, $P = 300\text{ Pa} - 37 \times 10^5\text{ Pa}$, dilution = 0.7 – 100% JP-10, conversion = 0–100 %). A significant part of the model comprises secondary conversion routes to aromatics such as benzene, toluene, styrene and polyaromatic hydrocarbons (PAHs) and could thus be used in assessing the tendency for deposit formation in fuel rich zones of endothermic fuel applications. Rate-of-production analysis showed the importance of four dominant primary decomposition channels of JP-10. Sensitivity analysis and rate-of-production analysis was used to obtain a skeletal model of 1098 reactions and 320 species containing the dominant reactions towards the main products.

Keywords: JP-10, pyrolysis; thermal decomposition; polyaromatic hydrocarbon formation; kinetic modeling

4.2 Introduction

Mono- and polycyclic hydrocarbons represent an interesting class of molecules that can be used as fuels in advanced aviation applications [1]. Many of them have superior physical properties such as high energy content, low viscosity at low temperatures, elevated flash point, and higher density compared to n-alkanes [2, 3]. Cyclic alkanes also have higher thermal stability limits and a larger heat sink capacity than their acyclic counterparts (cf. n-decane vs. decalin in ref. [4]), which are important properties in the propulsion technology. Endothermic cracking of the fuel prior to the combustion chamber [5-7] not only amplifies the heat sink, it can also reduce ignition delays as illustrated by Colket et al. [8]. Hence, a thorough understanding of the decomposition chemistry of cyclic hydrocarbons is not only important for the design and optimization of the endothermic reaction technology itself, but also for the correct design of the combustor that converts the products of the endothermic decomposition.

This study focuses on the kinetic modeling of the thermal decomposition of the fuel Jet Propellant-10, “JP-10”, the only air-breathing missile fuel in use [9]. JP-10 consists of a tricyclic hydrocarbon, exo-tricyclo[5.2.1.0^{2,6}]decane (exo-TCD). Although a fair number of experimental JP-10 pyrolysis studies covering a wide range of operating conditions identified and quantified light products containing up to five carbons [10-15], less is known about the importance of heavier primary products or intermediates. Several studies noted the significant formation of aromatic components such as benzene, and toluene under pyrolysis conditions [10, 12-16], which was also observed in Chapter 3 of this thesis. Even in the presence of an oxidizer, aromatic components represented an important share of the products as the comparative study between the flame structure of n-decane and JP-10 flames [17] indicated. In Chapter 3 of this thesis, a new data set of JP-10 pyrolysis experiments was obtained using GC×GC-FID/ToF-MS in which over seventy species with mass-to-charge ratio up to 178 (C₁₄H₁₀) were identified and quantified. Validation of two available kinetic models of JP-10 thermal decomposition, the San Diego model [18] and the RMG model v0.19 [19], revealed that they were largely inadequate in accurately predicting product distribution trends of the experiments of Chapter 3. This is not surprising since

these kinetic models were designed for higher temperature combustion applications and not for the pyrolysis conditions that were investigated.

A kinetic model for the purpose of modeling JP-10 pyrolysis is missing in the literature. Therefore, a new dedicated pyrolysis model was generated with the help of the Reaction Mechanism Generator (RMG). Five experimental data sets that encompass a wide range of operating conditions were used for the validation of the model. In order to further elucidate the initial decomposition of *exo*-TCD, a large number of electronic structure calculations for reactions, for which accurate reaction rate coefficients are missing, were carried out. The predictive kinetic model was reduced to 1098 reactions and 320 species for engineering applications using a combination of rate-of-production and sensitivity analysis.

4.3 Computational methods

4.3.1 Kinetic model construction

Fully automated kinetic model generation is not possible at present. A lot of user involvement is still required to obtain an accurate model. Fortunately, automatic network generation programs can reduce part of the burden if the relevant reaction families are known and thermochemical and kinetic data is available. Reaction Mechanism Generator (RMG) [20, 21] is an open-source software package that automatically constructs kinetic models based on the evaluation of reaction rates and species concentrations. The kinetic model is expanded by successive addition of kinetically significant species [22] to the kinetic model, as was described in Section 1.1.1.3 of Chapter 1 of this thesis. The general functionality of RMG was also extensively discussed elsewhere (e.g. [23-25]), and therefore only elements specific to the pyrolysis modeling of polycyclic hydrocarbons are highlighted here. Temperature, pressure, and initial *exo*-TCD dilution, which are used to evaluate rates of production in RMG, are set to 1100K, $1 \cdot 10^5$ Pa and 10 mol % respectively. The time at which the integration of the ODEs is stopped, and tolerance ϵ for the network generation (cf. Eq. 1.1 in Section 1.1.1.3 of Chapter 1), which are the main parameters determining the size of the generated reaction network, were set to 0.5s and 0.001, respectively. For these conditions the most important products observed in the experiments are included in the reaction network. The use of tighter tolerances significantly increases the size of

the model, but does not affect the model predictions. The effect of both parameters on the size of the reaction network for steam cracking of n-hexane was illustrated by Van Geem et al. [25].



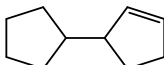
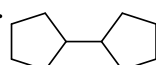
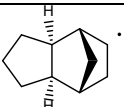
The pyrolysis of exo-TCD is described by three elementary reaction families: intermolecular hydrogen abstraction reactions by radicals, intra- and intermolecular radical addition reactions, and intermolecular radical recombination reactions. Details of the rate rules of these reaction families can be found elsewhere [20, 21]. Note that the reactions belonging to the reverse reaction families, i.e. β -scission reactions and unimolecular scission reactions, are allowed too; rate coefficients of these reactions are calculated based on thermodynamic consistency.

In addition to built-in databases for reaction families, a number of reactions from the literature sources containing chemistry relevant to exo-TCD pyrolysis were added. First of all, two sub-models were added that contain chemistry of species with two carbon atoms or less using the seed mechanism option in RMG. The Leeds methane oxidation model v1.5 [26], stripped of all oxygen chemistry, was used as a base model for high temperature pyrolysis reactions, resulting in 34 reactions between 22 species. An ethane steam cracking model by Sabbe et al. [24], stripped of all reactions containing four or more carbons, was used as a base model for the low temperature pyrolysis and contains 277 reactions between 35 species. Next, a number of literature sub-models were added to RMG through the so-called “Reaction Library” option of RMG. The reactions in these sub-models are considered during the iterative enlargement procedure and are added to the final model based on the kinetic significance of the species of these reactions. These literature sub-models can be divided into two categories. First of all, sources of cyclic C_5 – C_6 – C_7 chemistry [27–31] were included, as it is believed that they are crucial in bridging the gap between the exo-TCD decomposition products and secondary products such as aromatics. Secondly, sources were included that contain pathways for the growth of polycyclic aromatic hydrocarbons such as indene and naphthalene [32, 33]. Finally, thirty-three exo-TCD decomposition reactions were added, primarily originating from new transition state theory (TST) calculations performed in this study. These exo-TCD decomposition reactions are further discussed in the next section.

4.3.2 Thermochemistry

A detailed decomposition model necessitates the inclusion of thermochemical parameters for all intermediate species; This can be achieved on the basis of group and bond additivity methods [34] but the accuracy of these estimates may be doubtful for species with structures that fall beyond the scope of the regressed group additive parameters. More specifically, the influence of ring strain on the thermochemistry estimates of polycyclic intermediates remains a point of concern. Magoon et al. [35, 36] addressed this issue by incorporating on-the-fly thermochemical calculations for polycyclic intermediates using three-dimensional atomic coordinates in molecules [19]. In the present work, several important modifications in the calculation of thermochemical parameters for polycyclic intermediates were incorporated. First of all, a database of polycyclic ring strain corrections (RSC) was added containing ca. one hundred polycyclic structures [37-41]. If an appropriate polycyclic RSC is found for a component, it is used in the estimation of the thermochemical parameters. If not, the molecule is passed-on to the framework that generates 3D atomic coordinates on-the-fly using a distance-geometry algorithm implemented in the chemoinformatics toolkit “RDKit” [42] to convert the chemical graph representation into a 3D geometry. Secondly, the database of hydrogen bond increments (HBIs), a method developed by Lay et al. [43] to estimate the thermochemistry of radicals, was expanded with groups for cyclic and polycyclic radicals found in literature [44-47]. The radical atom in strained species cannot adopt the energetically favored planar geometry because of the constriction of the rigid carbon frame and therefore significantly alters the numerical values of the HBI in comparison with the acyclic analog. Table 4-1 shows a comparison of the thermochemistry estimates for a selected set of polycyclic strained species relevant for exo-TCD decomposition. Three values were reported: the Benson group additivity method including polycyclic ring strain corrections (RMG-GA), the semi-empirical PM3 method using 3D atomic coordinates that are generated on-the-fly (RMG-PM3) and finally the best available benchmark value. It was observed that neither RMG-GA nor RMG-PM3 reproduces all benchmark data well. In general, RMG-GA scores better for enthalpies of formation than for entropies or heat capacities because more data is available in literature. In the worst case, S and C_p contributions of polycyclic RSCs were estimated as a sum of contributions of smaller rings. Note that the thermochemistry of species in the model is always based on the best available estimates, e.g. CBS-QB3 values for the six tricyclodecyl $C_{10}H_{15}$ radicals from ref. [48].

Table 4-1: Comparison of the thermochemical properties for selected polycyclic species from the following methods: 1. PM3 method implemented through RMG with on-the-fly generation of atomic coordinates (RMG-PM3); 2. Benson group additivity method implemented in RMG including polycyclic ring strain corrections (RMG-GA); 3. benchmark values either from *ab initio* calculations or experimental data.

Component		$\Delta_f H^\circ$ (298K)	S° (298K)	C_p° (298K)	C_p° (800K)	C_p° (1500K)
		kJ mol^{-1}		$\text{J K}^{-1} \text{mol}^{-1}$		
 2-Norbornene	Benchmark	80[49]	N/A	112.7[50]	268.2[50]	N/A
	RMG-PM3	92	306	100	260	343
	RMG-GA	80	188	121	272	352
 Norbornane	NIST	-54.9±1.1[51]	N/A	120[50]	296[50]	N/A
	RMG-PM3	-59	306	105	289	381
	RMG-GA	-50	193	134	301	389
 3-cyclopentyl-cyclopentene (MA110)	THERGAS	-4	431	167	406	536
	RMG-PM3	-54	422	163	406	532
	RMG-GA	-13	423	163	402	527
 BR1	UB3LYP/cbsb7 [10]	243	410	180	410	536
	RMG-PM3	218	467	159	398	527
	RMG-GA	251	481	155	389	523
 tricyclodecyl radical (TCDR8)	CBS-QB3[48]	113	364	155	389	511
	RMG-PM3	105	377	142	385	511
	RMG-GA	134	377	142	385	511

4.3.3 Quantum-chemical calculations

High level thermochemical data for species involved in the JP-10 thermal decomposition is very scarce. Rate coefficients are even less available, in particular regarding the initial decomposition of exo-TCD. Therefore in this work new high-level ab-initio calculations were carried out on a set of reactions that are believed to be important for the decomposition of exo-TCD. Three types of reactions were calculated.

Unimolecular C-C scission reactions of exo-TCD leading to biradicals and subsequent intramolecular H-abstraction reactions of the biradicals were explored using density functional theory (DFT) and multireference methods. In the DFT methodology, geometries were optimized using UB3LYP (unrestricted B3LYP) and energies were calculated using CBS-QB3 to explore the PES and obtain initial refinements of kinetic parameters. Using the DFT geometries as initial guess, a multireference methodology CASPT2 [52, 53], a second-order perturbation theory, multireference counterpart of the MP2 method, as implemented in MOLPRO [54], was applied. A four-electron, four-orbital active space, generally corresponding to the electrons and orbitals of interest for these reactive transformations, as described in [19], was employed. In contrast to the previous study [19], which relied on CASSCF reaction path geometries for performing higher-level single-point calculations, the saddle points and minima considered here were obtained directly on the higher-level, in this case CASPT2. A 6-311G(d,p) basis set, as used by Sirjean et al in CASSCF studies of similar reactions [55, 56], was applied here. Hessian/frequency calculations at the minima and saddle points were performed using finite differencing of analytic CASPT2 gradients, as calculated by MOLPRO. Partition functions were computed with rigid-rotor, harmonic oscillator (RRHO) approximation and Arrhenius parameters were obtained using conventional transition-state theory (TST).

A third type of reaction is the unimolecular C-C and C-H β -scission of tricyclodecyl ($C_{10}H_{15}$) monoradicals. These reactions were studied on the CBS-QB3 level of theory using Gaussian03 and Gaussian09. Tricyclodecyl reactant structures were optimized with “verytight” convergence criteria; correspondingly, a very fine integration grid was used (“int=ultrafine”). Vibrational frequencies, calculated at B3LYP/CBSB7 level of theory, confirmed that the transition state structures were saddle points of first order have a single imaginary frequency. Geometries were

visually inspected to confirm that the structure was intermediate between the intended reactant and product(s). Conventional transition-state theory (TST) calculations were performed using results for the saddle point and corresponding tricyclodecyl reactant. Quantum tunneling coefficients were calculated by using the symmetric Eckart tunneling scheme, using energetic information from tricyclodecyl reactant and the saddle point [57]. Frequencies were corrected by a scaling factor of 0.99. These TST calculations were performed using the software package CanTherm [58]. Rate coefficients were computed at ten different temperatures ranging from 298K to 3000K. Based on these sampled points, a three-parameter least squares fit was performed by CanTherm [58] to determine the parameters for a modified Arrhenius rate coefficient expression.

4.4 Reactor modeling and experimental data

A wide variety of experimental data on the thermal decomposition of exo-TCD published in open literature were used for validating the model. The kinetic model was compared to the product distribution data obtained by Herbinet et al. [10] (hereafter referred to as the “Nancy data”), Nakra et al. [13] (“Nakra data”), Van Devener and Anderson [14] (“Van Devener data”), Rao and Kunzru (“Rao data”). [12], Xing et al. [15] (“Xing data”) and the data from Chapter 3, referred to as “Ghent data” in this chapter. Table 4-2 shows an overview of the operating conditions of each experimental data set.

Table 4-2: Summary of operating conditions of the experimental data sets used for model validation.

	Nakra data	Vandevener data	Nancy data	Rao data	Xing data	Ghent
	[13]	[14]	[10]	[12]	[15]	†
Pressure / Pa	300-400	1500 - 1900	$1 \cdot 10^5$	$1 \cdot 10^5$	$1\text{-}37 \cdot 10^5$	$1.7 \cdot 10^5$
Residence time* / ms	2 – 10	3 – 10	500-6000	700-6400	480-26400	300-500
mol % exo-TCD in inert	2.6;Ar	3.9;Ar	0.7-4; He	100 [#]	100 [#]	7/10; N ₂
Temperature range / K	300- 1500	300 - 1400	848-933	903-968	823-903	930-1080
Conversion range / %	0-100	0-100	0.01-25	10-60	0-25	4-94
Reactor type	Plug flow	Plug flow	CSTR	Plug flow	Plug flow	Plug flow
Endo-TCD content / %	N/R [°]	N/R [°]	2.5	3.6 ⁺	0.23	0.6

[#] No diluent present

* Residence time based on the total (including diluent) molar flow rate at the entrance of the reactor, at the mean temperature inside the reactor.

† based on reported 96.4% purity of reactant

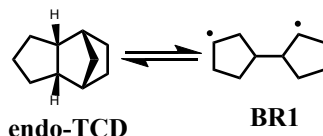
° Not reported

† Experimental data from Chapter 3 of this thesis

Reactor simulations were systematically conducted with the Chemkin Pro package [59]. The experiments of Herbinet et al. [10] in a jet-stirred reactor were modeled with a perfectly mixed reactor model, similar to what was done in their comparison with the Nancy model. The steady-state reactor model was solved using a reactor volume of 90 cm³ and a given temperature and pressure. Integration was stopped when the reported space time was reached. The data from Nakra et al. [13] and Van Devener and Anderson [14] were collected at the millisecond time scale under reduced pressures. Both reactors were modeled as plug flow reactors specifying the given diameter and length of the hot zone in the reactor. Isothermal and isobaric conditions with the pressure based on the provided hot zone midpoint pressure were used. In the continuous flow tubular reactor of Rao and Kunzru [12] the axial temperature profile along the reactor was not reported. Instead, the equivalent reactor temperature, i.e. the temperature at which the same exo-TCD conversion would be attained as in the non-isothermal case was provided. Therefore, the reactor was modeled using isothermal, isobaric boundary conditions. Integration was stopped when the reported space time was reached. Xing et al. [15] reported JP-10 pyrolysis data up to pressures of $37 \cdot 10^5$ Pa. Because species fractions were reported for the gaseous and liquid phase separately, without any indication of the relative ratio of the two phases, only exo-TCD conversion was compared. The flow reactor was modeled as a plug flow reactor using isothermal, isobaric boundary conditions. The integration was stopped at the reported space time. In Chapter 3, experiments in a continuous flow tubular reactor were reported. Radial temperature and

concentration gradients were shown to be negligible using two-dimensional reactor models for this reactor [23, 60].

JP-10 feedstock also contains a number of impurities, such as endo-tricyclo[5.2.1.0^{2,6}]decane (endo-TCD), adamantane, and decalin in decreasing order of importance. While the cumulative content of these impurities typically does not exceed 3.5% of JP-10 feedstock [61], the presence of endo-TCD may be of importance for the reactivity of exo-TCD. Hudzik et al. [48] reported that the standard enthalpy of formation of endo-TCD is 17 kJ mol⁻¹ higher than the value for exo-TCD, suggesting that the endo-isomer may be less stable than the exo-isomer. Information on the initial decomposition of endo-TCD is non-existing in literature; therefore, only one reaction was added that converts endo-TCD into BR1, similar to the initiation of exo-TCD.



The activation energy of this reaction is lowered by 17 kJ mol⁻¹, relative to the value of the reaction with exo-TCD, to reflect the high enthalpy of formation of the endo-isomer. The JP-10 feedstock of the modeled experiments was treated as a mixture of exo- and endo-TCD whenever information on the content of endo-TCD was available, cf. Table 4-2. If not, JP-10 was considered as 100% exo isomer.

4.5 Results and Discussion

The final RMG-generated, kinetic model for the pyrolysis of exo-TCD contains 5261 reactions between 384 species, i.e. 234 radicals and 150 molecular species. Thermochemistry of 18% of the cyclic species in the model was estimated via PM3 calculations, 19% from literature while the remainder was estimated via group contribution methods. The model consists of 4308 H-abstraction reactions, 316 radical addition reactions to molecules with double or triple bonds, 158 radical recombination reactions, the remaining 479 reactions originated from specific libraries, as was discussed in Section 4.3.1 of this chapter. A combination of sensitivity analysis and rate of production analysis was used to reduce the generated model to its most important reactions. The reduced model consists of 1098 reactions and 320 species and was able to capture the main trends of the full model.

4.5.1 Kinetic model validation

The generated kinetic model was validated against five data sets of exo-TCD pyrolysis experiments. None of the rate coefficients or thermochemical parameters was adjusted by regression of the experimental data.

Figure 4-1 shows the exo-TCD conversion as a function of the mean reactor temperature for the Ghent data. It can be seen that good agreement is found for all temperatures, and both levels of dilution.

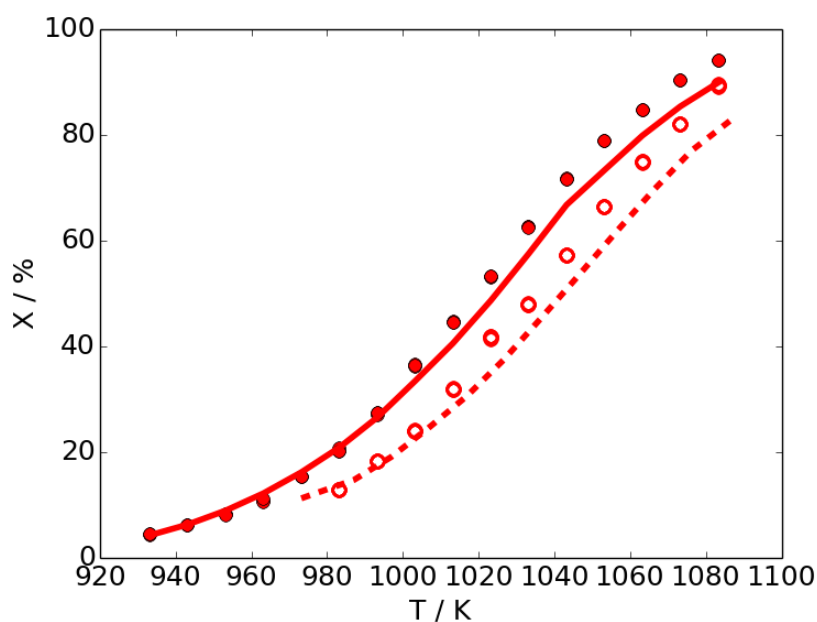


Figure 4-1: JP-10 conversion as a function of temperature for two dilution levels of JP-10 in N_2 at reactor inlet. Experiments from Chapter 3: \bullet : 10 mol%, and \circ : 7 mol % JP-10 in N_2 . Model predictions: —: 10 mol%, and - -: 7 mol % JP-10 in N_2 . $P = 1.7 \cdot 10^5$ Pa.

Figure 4-2 shows the exo-TCD conversion as a function of the equivalent space time for the Rao data. The good agreement between experiments and model predictions indicates that the generated kinetic model performs well for high exo-TCD partial pressures.

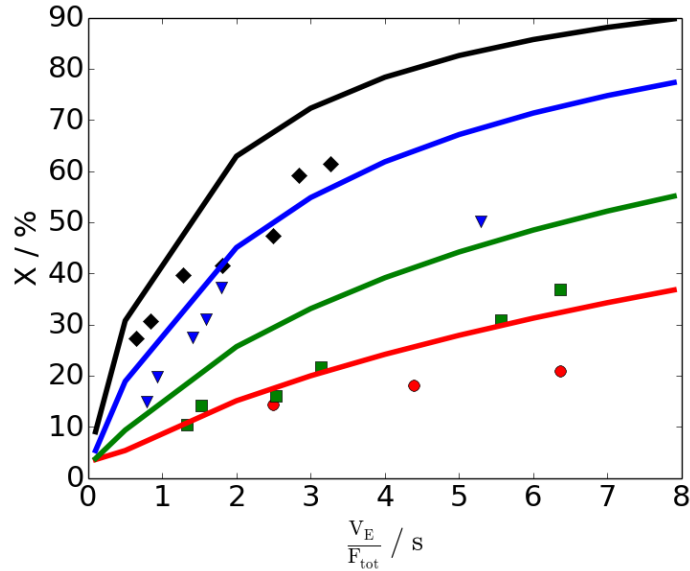


Figure 4-2: JP-10 conversion as a function of equivalent space time at different temperatures: ♦: 968K, ▼: 948K, ■: 923K, ●: 903K.. Experiments from Rao and Kunzru [12]. Lines are model predictions. Conditions: 100% JP-10 at the reactor inlet, $P = 1 \cdot 10^5$ Pa, $T = 903\text{--}968\text{K}$.

Figure 4-3 shows the exo-TCD conversion as a function of the reactor pressure for the Xing data, the only dataset at elevated pressures up to $37 \cdot 10^5$ Pa[15]. While exo-TCD conversion is predicted to increase with increasing pressure by the model, the experiments show little influence of pressure on exo-TCD conversion at a given temperature. In the experiments of Xing et al. [15], the reported space time increases by a factor of 50 from the lowest reported pressure relative to the highest pressure, and explains why the model predictions for exo-TCD conversion at a given temperature increase as a function of pressure. The model predictions are in good agreement at pressures up to $1.8 \cdot 10^5$ Pa and tend to slightly deviate at more elevated pressures.

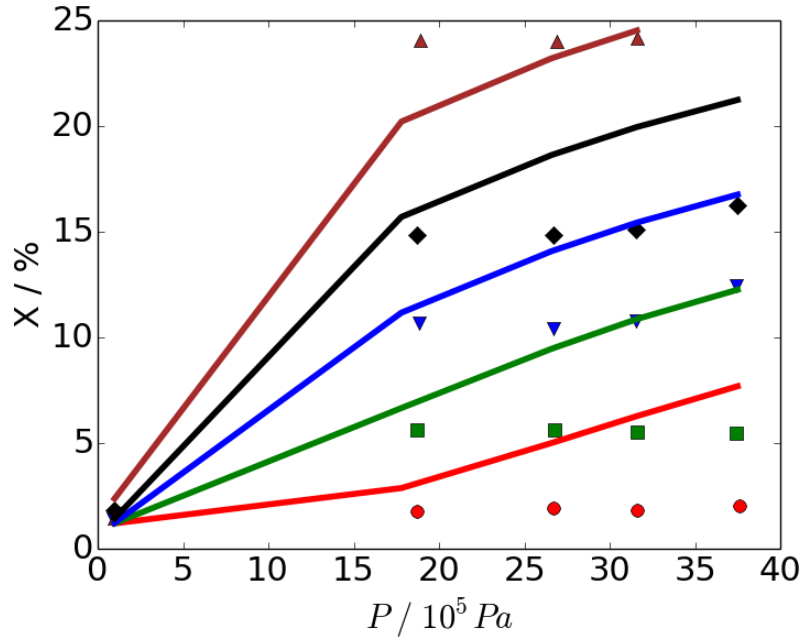


Figure 4-3: JP-10 conversion as a function of reactor pressure at different temperatures: \blacktriangle : 903K, \blacklozenge : 883K, \blacktriangledown : 863K, \blacksquare : 843K, \bullet : 823K. Experiments from Xing et al. [15]. Lines are model predictions. Conditions: 100% JP-10 at the reactor inlet, $T = 823\text{-}903\text{K}$.

Figure 4-4 shows the exo-TCD conversion as a function of the residence time for the Nancy data. Good agreement is found between model predictions and experiments, indicating the validity of the model at low conversions under diluted conditions.

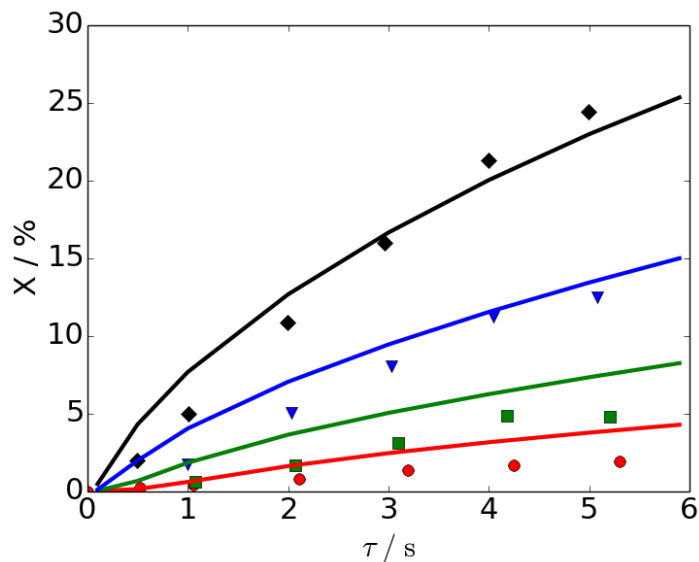


Figure 4-4: JP-10 conversion as a function of residence time at different temperatures: \blacklozenge : 933K, \blacktriangledown : 913K, \blacksquare : 893K, \bullet : 873K. Experiments from Herbinet et al. [10]. Lines are model predictions. Conditions: 4 mol% exo-TCD in He at the reactor inlet, $T = 848\text{--}933\text{K}$, $P = 1 \cdot 10^5 \text{ Pa}$, conversion = 0.01-25%.

Figure 4-5 shows the mole fractions of the major products as a function of the mean reactor temperature for the Ghent data. Model predictions of hydrogen, propene, CPD and cyclopentene are in good agreement with the data. The predicted maxima of CPD concentrations are at slightly higher temperatures than experimentally observed. Ethene is consistently underpredicted by a factor of two for the data which is analyzed further in Section 4.5.3 of this chapter.

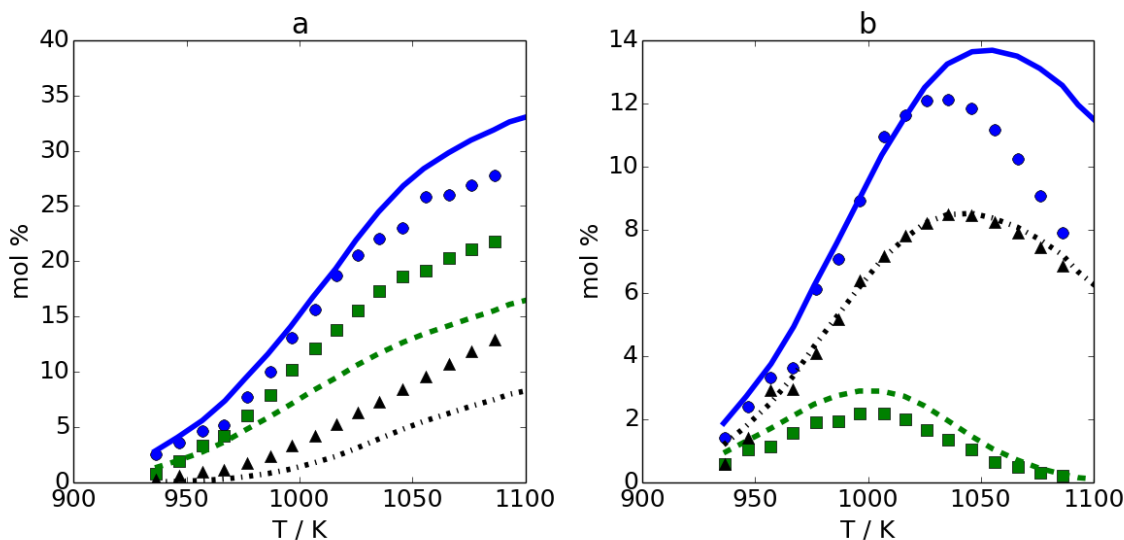


Figure 4-5: Mole fractions as a function of temperature of a) experimental: ●: dihydrogen, ▲: methane, ■: ethene, model predictions: —: dihydrogen, - -: methane, - -: ethene, b) experimental: ●: 1,3-cyclopentadiene, ▲: propene, ■: cyclopentene, model predictions : —: 1,3-cyclopentadiene, - -: propene, - -: cyclopentene. Conditions: 10 mol % JP-10 in N₂ at reactor inlet. $P = 1.7 \cdot 10^5$ Pa. Experimental data from Chapter 3 of this thesis.

Figure 4-6 shows mass fractions of the major products as a function of the equivalent space time for the Rao data. Similar conclusions for the agreement between model predictions and experiments are drawn as for the Ghent data, although model predictions of methane agree with experiments, while methane was slightly underpredicted for the Ghent experiments.

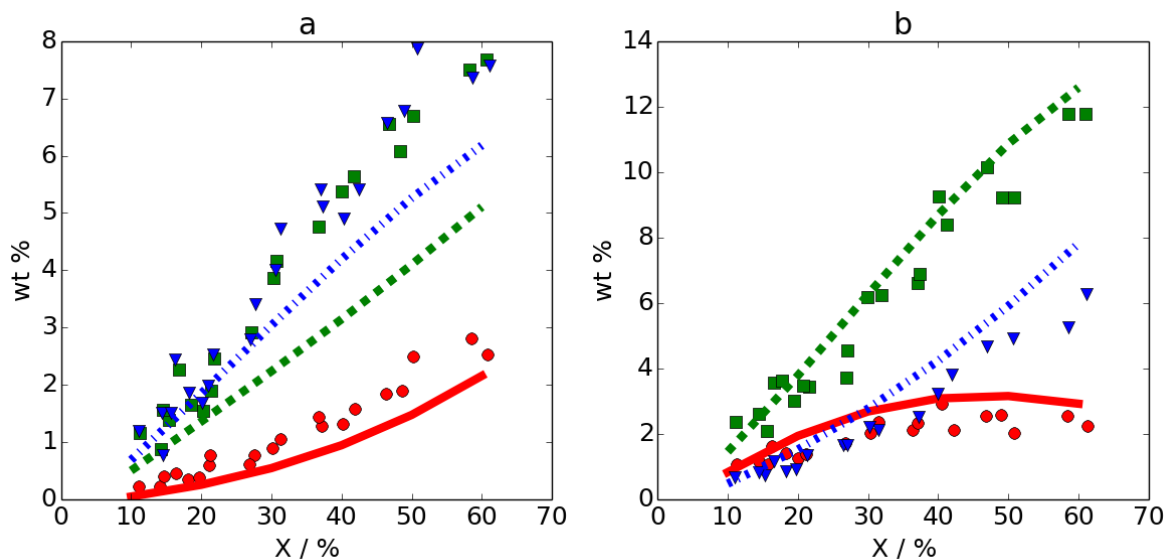


Figure 4-6: Mass fractions as a function of exo-TCD conversion of a) experimental: ●: methane, ▲: propene, ■: ethene, model predictions: —: methane, - -: propene, - · -: ethene, b) experimental: ●: cyclopentene, ▲: benzene, ■: 1,3-cyclopentadiene, model predictions : —: cyclopentene, - -: benzene, - · -: 1,3-cyclopentadiene. Conditions: 100 mol % JP-10 at reactor inlet, $T=903-968\text{K}$, $P = 1 \cdot 10^5 \text{ Pa}$. Experimental data from Rao and Kunzru [12].

Figure 4-7 shows the mole fractions of the major products as a function of the mean reactor temperature for the Nancy data. Again, good agreement is found for the majority of the major products, and confirms the trends seen in Figure 4-5.

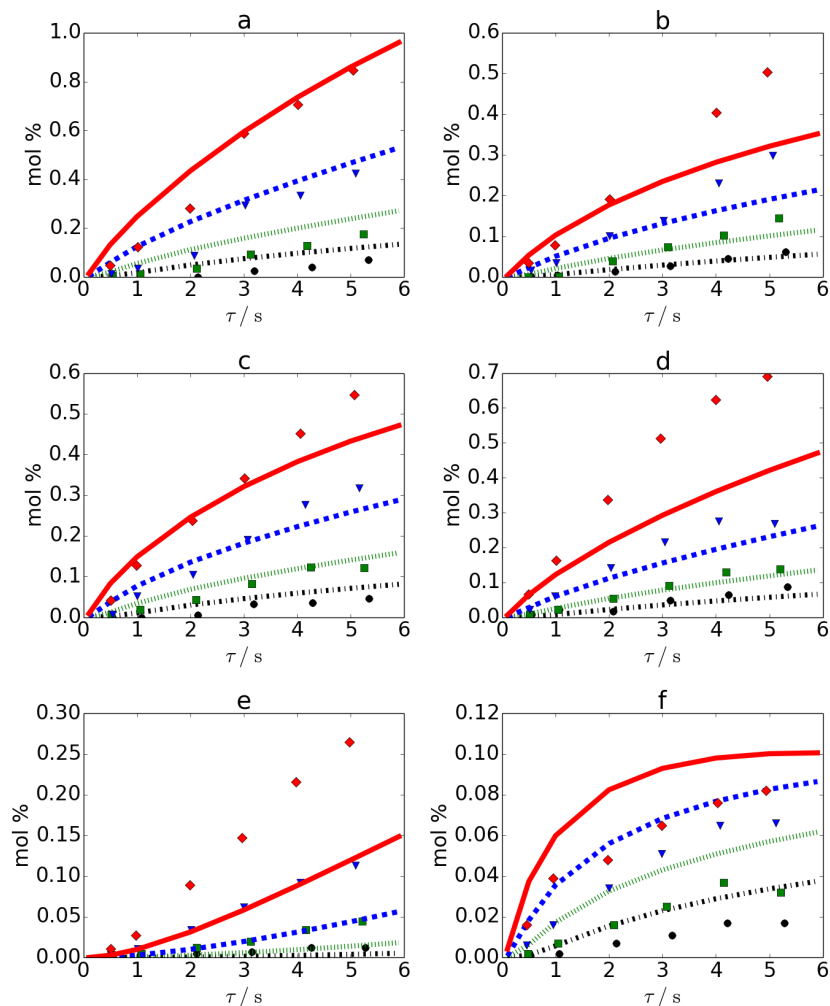


Figure 4-7: Mole fractions as a function of residence time of a) dihydrogen, b) propene, c) 1,3-cyclopentadiene, d) ethene, e) methane, f) cyclopentene at four temperatures: \blacklozenge : 933K, \blacktriangledown : 913K, \blacksquare : 893K, \bullet : 873K. Lines are model predictions, symbols are experimental data from Herbinet et al. [10]. Conditions: 4 mol% exo-TCD in He at the reactor inlet, $T = 848-933\text{K}$, $P = 1 \cdot 10^5 \text{ Pa}$, conversion = 0.01-25%.

Mole fractions of the most important formed aromatic species as a function of the mean reactor temperature for the Ghent data are shown in Figure 4-8. Besides benzene, the four most significantly formed aromatic species are toluene, styrene, indene and naphthalene. Benzene, and toluene are slightly underpredicted at temperatures below 1000K. Indene and naphthalene are overpredicted by a factor of two over the entire temperature range, and could suggest that channels to indene and naphthalene are too much favored over the channels to monoaromatics.

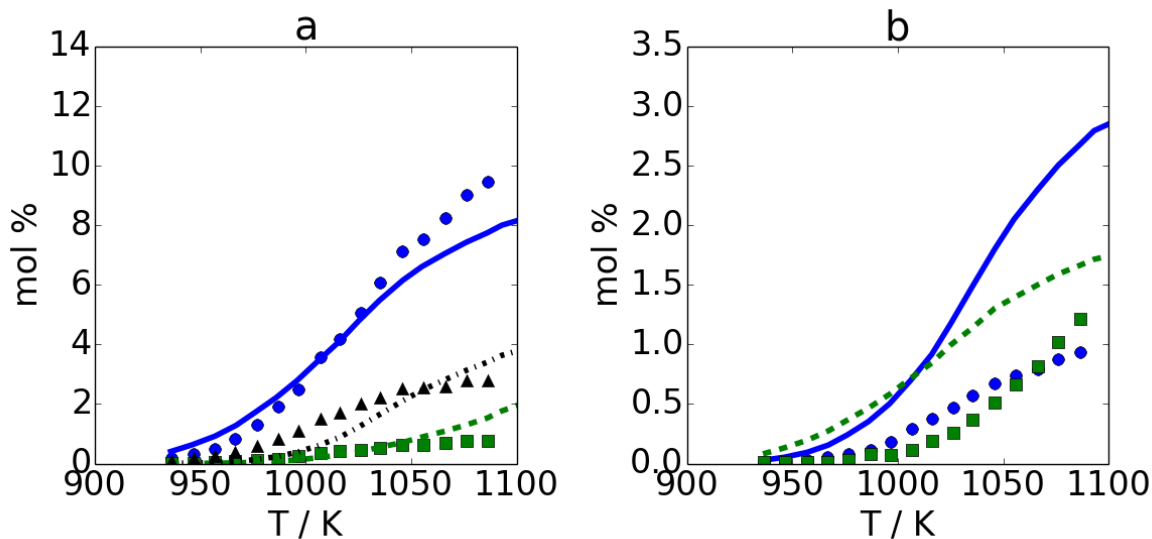


Figure 4-8: Mole fractions as a function of temperature for a) experimental: ●: benzene, ▲: styrene, ■: toluene, model predictions: —: benzene, - -: styrene, - -: toluene, b) experimental: ●: indene, ■: naphthalene, model predictions: —: indene, - -: naphthalene. Conditions: 10 mol % JP-10 in N_2 at reactor inlet. $P = 1.7 \cdot 10^5$ Pa.

The same trends are observed in Figure 4-9 that shows mole fractions of benzene and toluene as a function of the mean reactor temperature for the Nancy experiments.

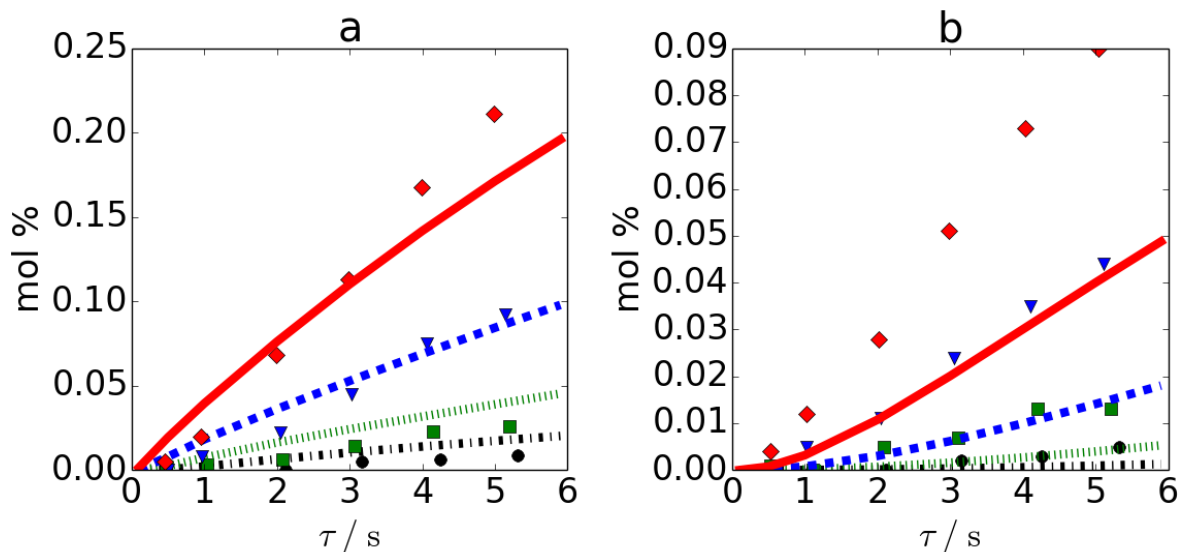


Figure 4-9: Mole fractions as a function of residence time of a) benzene, b) toluene at four temperatures: \blacklozenge : 933K, \blacktriangledown : 913K, \blacksquare : 893K, \bullet : 873K. Lines are model predictions, symbols are experimental data from Herbinet et al. [10]. Conditions: 4 mol% exo-TCD in He at the reactor inlet, $T = 848\text{--}933\text{K}$, $P = 1 \cdot 10^5 \text{ Pa}$, conversion = 0.01-25%.

Figure 4-10 shows the model predictions versus the Nakra and Van Devener data as a function of temperature. Overall, good agreement between model and experiment indicates that the current model performs well at low pressures and elevated temperatures despite being generated with RMG at significantly different conditions: $T = 1100\text{K}$, $P = 1 \cdot 10^5 \text{ Pa}$, 10 mol% exo-TCD in N_2 . Good agreement is observed for exo-TCD conversion. Although both data sets were recorded at similar operating conditions, CPD and ethyne concentrations are underpredicted in the Van Devener data and overpredicted in the Nakra data while reverse trends are seen for ethene. Less discordant trends are found for benzene. The latter is systematically underpredicted. Benzene model predictions also show a small decrease in benzene mole fractions starting at $T > 1500\text{K}$ which was not observed in the experimental data.

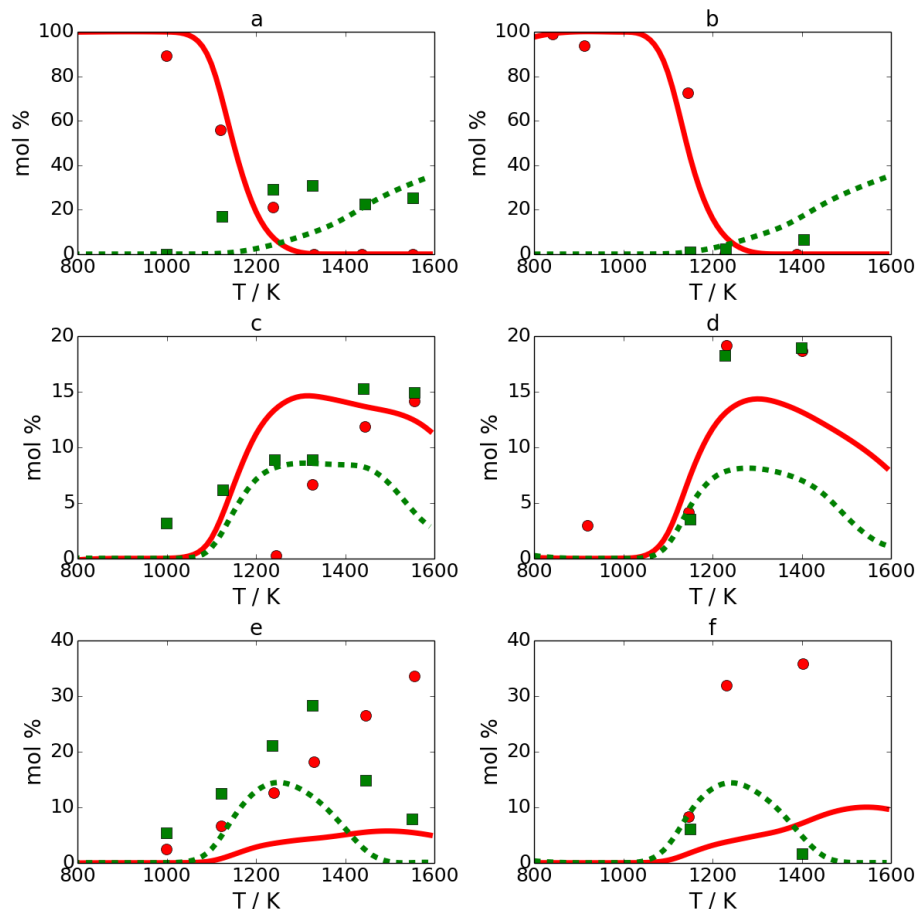


Figure 4-10: Mole fractions as a function of reactor temperature for model predictions versus experimental data of Nakra et al. [13] (left column) and Van Devener and Anderson [14] (right column). a) and b) experiments: ●: exo-TCD, ■: ethyne, model predictions: —: exo-TCD, - -: ethyne, c) and d) experiments: ●: ethene, ■: propyne, model predictions: —: ethene, - -: propyne, e) and f) experiments: ●: benzene, ■: 1,3-cyclopentadiene, model predictions: —: benzene, - -: 1,3-cyclopentadiene. Conditions for the Nakra et al. [13] and Van Devener and Anderson [14] respectively: 2.6 / 3.9 mol% JP-10 in Argon, $P = 300\text{-}400 / 1500\text{-}1900$ Pa, conversion = 0-100%.

The Ghent and Nancy experiments allowed to detect a number of species with concentrations of less than 2 mol%. First of all, the Ghent experiments isolated a number of products containing a five-membered ring, such as fulvene (C_6H_6), methylcyclopentadiene (C_6H_8), and ethenylcyclopentene (C_7H_{10}), cf. Chapter 3. Figure 4-11 shows the mole fractions of these products as a function of the mean reactor temperature for the Ghent experiments. Since the position of the methyl or ethenyl substituent in methylcyclopentadiene and ethenylcyclopentene could not be determined, they are compared to model predictions of 1-methyl-1,3-

cyclopentadiene and 3-ethenylcyclopentene respectively. Good agreement between experiments and model predictions is found for the three components, although the maximum concentration of fulvene is overestimated by a factor of 1.5.

Tricyclo[5.2.1.0^{2,6}]-dec-4-ene is another higher molecular weight component that was identified as an important primary decomposition product, and which also appears in the current model. As can be seen on Figure 4-11, the temperature of peak concentration is well predicted, but concentrations are overpredicted by a factor of three. A number of hydrocarbons up to four carbons were detected in experiments as well, cf. Figure 4-11. Ethyne is very well predicted whereas propadiene and propyne are slightly overpredicted. 1,3-Butadiene shows significant underestimations of the model, too. The wrong prediction 1,3-butadiene is a recurring problem in many other fuel models. It is suspected that the model is missing an important pathway for 1,3-butadiene; future work is needed to resolve this discrepancy.

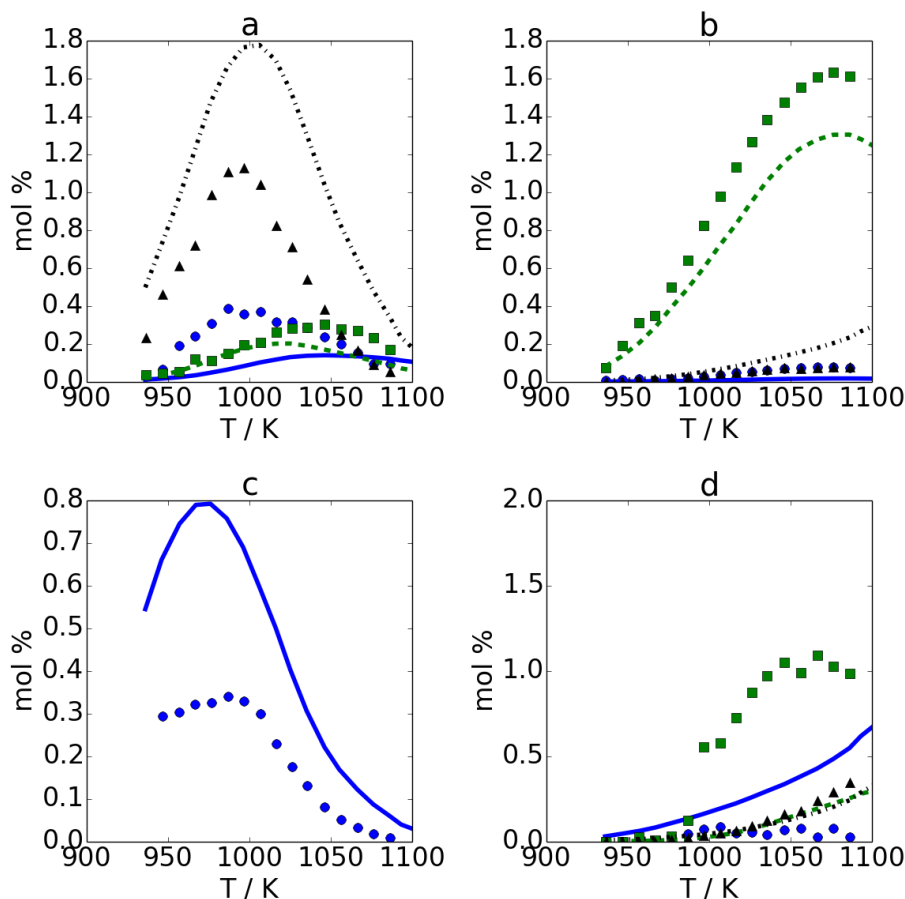


Figure 4-11: Mole fractions as a function of temperature for a) experimental: ●: fulvene, ▲: 3-ethenylcyclopentene, ■: 1-methyl-1,3-cyclopentadiene, model predictions: —: fulvene, - -: 3-ethenylcyclopentene, - -: 1-methyl-1,3-cyclopentadiene, b) experimental: ●: propane, ▲: propadiene, ■: ethane, model predictions: —: propane, - -: propadiene, - -: ethane, c) experimental: ●: tricyclo(5.2.1.0^[2,6])dec-4-ene, model predictions: —: tricyclo(5.2.1.0^[2,6])dec-4-ene, d) experimental: ●: propyne, ▲: ethyne, ■: 1,3-butadiene, model predictions: —: propyne, - -: ethyne, - -: 1,3-butadiene. Conditions: 10 mol % JP-10 in N₂ at reactor inlet. $P = 1.7 \cdot 10^5$ Pa.

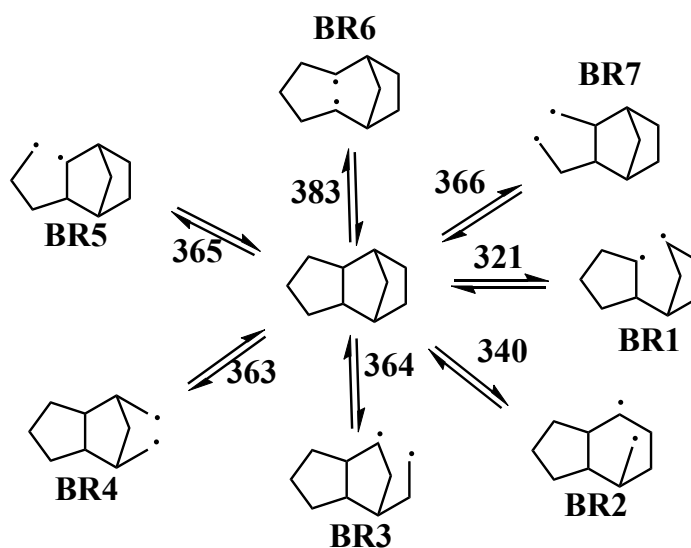
The overall good agreement between the model predictions and experimentally measured concentrations for a large number of experimental data sets indicates the validity of the current model. None of the model parameters were fitted, despite the many uncertainties in kinetics of reactions and thermochemistry of species. This gives an indication of the overall quality of the kinetic model. In comparison with previous models tested against the Ghent data (cf. Section

3.4.3 of Chapter 3), the current model provides better estimates for major products, and contains predictions of many minor products that are not present in previous models.

4.5.2 Primary decomposition chemistry

4.5.2.1 Exo-TCD initiation

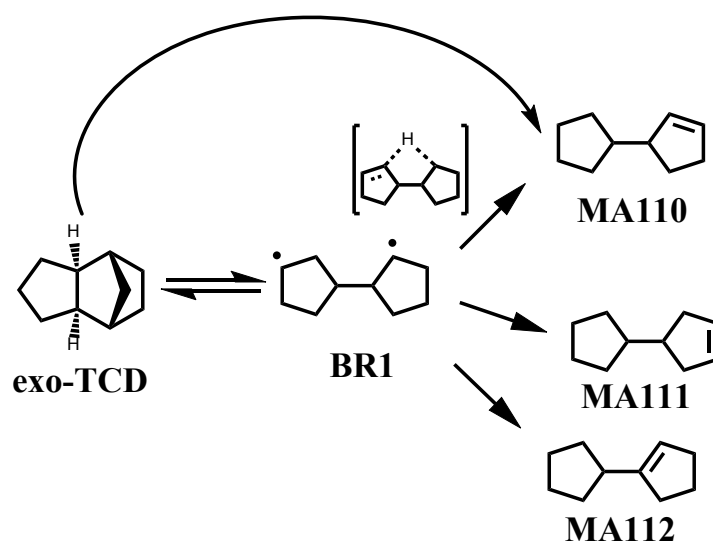
It is generally assumed that the pyrolysis of hydrocarbons such as exo-TCD proceeds through a free radical mechanism. The initiation of exo-TCD involves the creation of radicals through unimolecular bond scission reactions of exo-TCD. Although the rupture of C–H bonds is possible as an exo-TCD initiation route, these bonds are stronger than C–C bonds and were consequently neglected. The unimolecular initiation of exo-TCD can proceed through unimolecular C–C bond scission reactions resulting in the creation of seven distinct biradicals named BR_x, x=1..7, cf. Scheme 4-1.



Scheme 4-1: Exo-TCD initiation routes through unimolecular C–C scission reactions resulting in seven biradicals, BR_x, x=1..7. Nomenclature for the biradicals is adopted from Herbinet et al. [10]. Values represent standard reaction enthalpies (kJ mol⁻¹) at 298K, based on thermochemistry used in this model.

The resulting highly unstable biradicals can further isomerize to stable molecules via intramolecular H-abstraction reactions (disproportionation reactions), proposed by Herbinet et al. [10] upon detecting 3-cyclopentylcyclopentene, further referred to as “MA110”, as a major primary decomposition product.

Intramolecular H-abstraction reactions are thought to proceed via multiple distinct transition states that differ by the polycyclic entity that is formed through the course of the reaction, cf. Scheme 4-2.



Scheme 4-2: The unimolecular C-C scission of exo-TCD yielding biradical BR1, and subsequent intramolecular H-abstraction reactions yielding MA110, MA111, and MA112, as well as the direct, concerted conversion of exo-TCD to MA110.

Herbinet et al. [10] were among the first to estimate both the ring opening reactions as well as the intramolecular H-abstraction reactions via semi-empirical correlations based on observations by O’Neal and Benson [62] in which the ring strain of reactants, products and transition states played a central role. These qualitative estimations were combined with thermochemistry estimated via Benson group additivity, and showed that the initiation route via BR1 to MA110 was significantly faster than the other competing initiation routes. More recently, Magoon et al. [19] discussed the difficulty in accurately quantifying reaction barriers for the intramolecular disproportionation reactions; the highest level *ab initio* calculations performed by Magoon et al. [19] suggested there may be a very small barrier or even no barrier at all. So, Magoon et al. [19]

suggested that initiation of exo-TCD towards MA110 may proceed as a concerted step rather than via a distinct biradical intermediate [19], cf. Scheme 4-2. In the present work, rate coefficients for 15 ring-opening and intramolecular H-abstraction reactions were calculated. The latter are particularly useful in the re-evaluation of the relative importance of the different exo-TCD initiation routes. Table 4-3 shows values for the Arrhenius parameters for these reactions. Activation energies for the ring-opening reactions are very similar to those reported by Herbinet et al. [10], who calculated the activation barriers based on the difference in ring strain. The good agreement between the *ab initio* values and value obtained via the semi-empirical correlation also shows that the estimation of activation barriers based on the difference in ring strain between the reactant and product is acceptable as a reasonable first approximation. Values for the initiation of exo-TCD leading to the formation of BR2 are the exception where the current calculations indicate an activation barrier that is comparable to the other scission possibilities while the Herbinet et al. estimates [10] are ca. 30 kJ mol⁻¹ lower. Good agreement for the ring-opening reactions between the CBS-QB3 values and the CAS-PT2 values for activation energies show the usefulness of the CBS-QB3 method for these kinds of calculations.

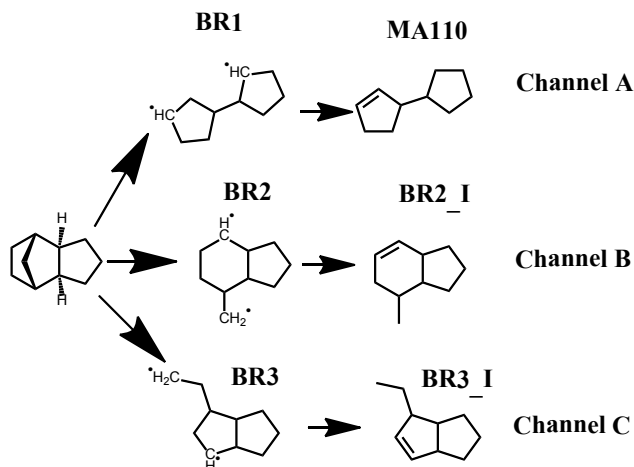
Table 4-3: Rate coefficients for exo-TCD initiation reactions and intramolecular H-abstraction reactions. Rate coefficient k is expressed as $k = AT^n \exp\left(-\frac{E_a}{RT}\right)$. Arrhenius parameters are fitted for a temperature range from 298 - 3000K. Unit for A is s^{-1} , for E_a $kJ\ mol^{-1}$.

Reaction	A	n	E_a	Method used
<u>Exo-TCD initiation reactions</u>				
Exo-TCD = BR1	$1.80\ 10^{13}$	0.83	326	CAS-PT2*
Exo-TCD = BR1	$7.09\ 10^{12}$	0.97	319	UB3LYP/CBS-QB3
Exo-TCD = BR1	$5\ 10^{15}$	0	322	Herbinet et al. [10]
Exo-TCD = BR2c [#]	$8.32\ 10^{10}$	1.87	333	CAS-PT2*
Exo-TCD = BR2d [#]	$1.09\ 10^{13}$	0.80	325	CAS-PT2*
Exo-TCD = BR2c [#]	$5.06\ 10^{12}$	0.96	328	UB3LYP/CBS-QB3
Exo-TCD = BR2d [#]	$3.87\ 10^{12}$	0.95	321	UB3LYP/CBS-QB3
Exo-TCD = BR2	$5\ 10^{15}$	0	297	Herbinet et al. [10]
Exo-TCD = BR3	$6.09\ 10^{12}$	0.92	335	CAS-PT2*
Exo-TCD = BR3	$3.40\ 10^{12}$	1.09	334	UB3LYP/CBS-QB3
Exo-TCD = BR3	$5\ 10^{15}$	0	322	Herbinet et al. [10]
Exo-TCD = BR4	$8.05\ 10^{12}$	1.24	334	UB3LYP/CBS-QB3*
Exo-TCD = BR4	$5\ 10^{15}$	0	322	Herbinet et al. [10]
Exo-TCD = MA110	$5.76\ 10^{10}$	6.88	416	UB3LYP/CBS-QB3*
exo-TCD = MA111	$1.85\ 10^{11}$	1.23	336	UB3LYP/CBS-QB3*
<u>Intramolecular H-abstraction reactions</u>				
BR1 = MA110	$8.64\ 10^{11}$	0.08	-1	CAS-PT2*
BR1 = MA110	$2.90\ 10^{10}$	0.44	15	UB3LYP/CBS-QB3
BR1 = MA110	$1.9\ 10^{10}$	1	32	Herbinet et al. [10]
BR2 = BR2_I	$4.5\ 10^9$	0.65	20	UB3LYP/CBS-QB3*
BR3 = BR3_I	$2.87\ 10^{10}$	0.40	23	UB3LYP/CBS-QB3*

[#]For BR2, two reactions are identified corresponding to two different BR2 product conformers.
 * denotes the rate coefficient used in the current model.

The relative importance of the unimolecular decomposition channels of exo-TCD through the biradical BR_x to the isomerization products MA110, BR2_I, and BR3_I, cf. Scheme 4-3, was evaluated by comparing the global rate coefficient k_g , given by Eq. 4-1.

$$k_g = \frac{k_1 \cdot k_2}{k_{-1} + k_2} \quad \text{Eq. 4-1}$$



Scheme 4-3: Three unimolecular initiation pathways of exo-TCD and subsequent intramolecular H-abstractions.

The latter was calculated assuming quasi-stationary state for the biradicals of exo-TCD and uses the rate coefficients of the ring opening and reverse ring closing reaction (k_1 and k_{-1}) and the rate coefficient of the rearrangement reaction of the biradical k_2 .

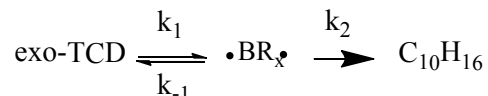


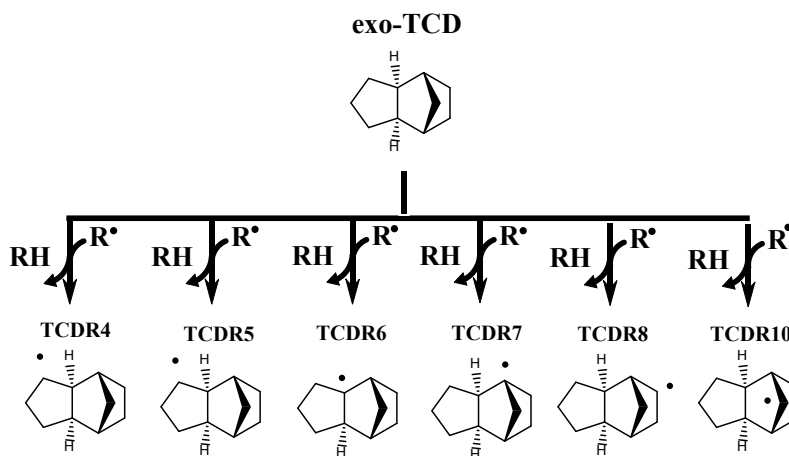
Table 4-4: Comparison of the global rate coefficient k_g at 1000K for three exo-TCD initiation routes based on Arrhenius parameters from this study. Unit for k_g is s^{-1} .

		$k_g = \frac{k_1 \cdot k_2}{k_{-1} + k_2}$
Channel A	Exo-TCD = BR1 = MA110	$5.2 \cdot 10^{-2}$
Channel B	Exo-TCD = BR2 = BR2_I	$2.1 \cdot 10^{-4}$
Channel C	Exo-TCD = BR3 = BR3_I	$2.4 \cdot 10^{-4}$

Table 4-4 shows that the initiation route via BR1 to MA110 is two orders of magnitude faster than the competing channels via BR2 and BR3 using the rate coefficients calculated in this work, implying that this channel is the principal unimolecular initiation route for exo-TCD at temperatures around 1000K. This is indeed in line with conclusions by Herbinet et al. [10]. The concerted pathway from exo-TCD to MA110, using the kinetics by Magoon et al. [19] is significantly slower than the route via biradical BR1.

4.5.2.2 Tricyclodecyl decomposition reactions

It is expected that a large part of *exo*-TCD conversion is attributed to H-abstraction reactions by radicals. The abstraction of a hydrogen atom from *exo*-TCD yields six distinct $C_{10}H_{15}$ tricyclodecyl radicals, cf. Scheme 4-4. The kinetics of these H-abstractions not only greatly affect the *exo*-TCD conversion, but they also determine selectivity of the primary products since each tricyclodecyl radical leads to distinct primary products. The activation energy of the H-abstraction reactions by H atoms from *exo*-TCD were calculated using an Evans-Polanyi relationship proposed by Dean and Bozzelli [63]. The instability induced by the strained nature of the resulting tricyclodecyl radicals was thus taken into account in the kinetics of these hydrogen abstraction reactions.



Scheme 4-4: Six possible $C_{10}H_{15}$ tricyclodecyl radicals created by H-abstraction reactions from *exo*-TCD. Nomenclature for the tricyclodecyl monoradicals is adopted from Hudzik et al. [48].

The tricyclodecyl radicals subsequently react via various C-H and C-C β -scission reactions leading to either $C_{10}H_{14}$ molecules or to new $C_{10}H_{15}$ radicals. A single monoradical can react through multiple parallel decomposition channels. The kinetics for these elementary reactions were not available at present and therefore they were calculated.

Literature providing kinetic data for cycloalkyl decomposition reactions is very scarce; a comparison of the calculated rate coefficients with previously reported data for these reactions is impossible. Instead, the question arises whether the obtained rate coefficients are comparable to

the decomposition rates of the smaller units that are present in the tricyclodecyl radicals. Exo-TCD can be regarded as the joining of a cyclopentane ring and a norbornane bicyclic ring. Cyclopentyl, derived from cyclopentane, has two decomposition routes, one C-C β -scission reaction leading to 1-penten-5-yl and one C-H β -scission reaction leading to cyclopentene and H. Norbornane has three different types of carbon atoms, leading to three possible norbornyl radicals, each of these norbornyl radicals possess C-C β -scission possibilities. Sirjean et al. provided kinetic data at the CBS-QB3 level of theory for these reactions [64, 65] and can thus be used in comparing the rate coefficients with the tricyclodecyl decomposition reactions.

Rate coefficients for tricyclodecyl reactions with the analogous reactions of cyclopentyl (Figure 4-12a) and norbornyl radicals (Figure 4-12b-e) are compared. In some cases, the joining of the cyclopentane and norbornane ring in tricyclodecyl radicals lead to reactions whose TS structure is not comparable to TS's of the smaller units. Therefore, Figure 4-12 only compares tricyclodecyl reactions whose TS's are only located in one of the two rings. Rate coefficients of the tricyclodecyl decomposition reactions are smaller than for the analogous reactions of the smaller ring units, up to one order of magnitude of difference between β -scission reactions of TCDR4 and cyclopentyl, cf. Figure 4-12-a. Presumably, the presence of the adjacent ring leads to tighter cyclic TS structures and, hence, lower pre-exponential factors.

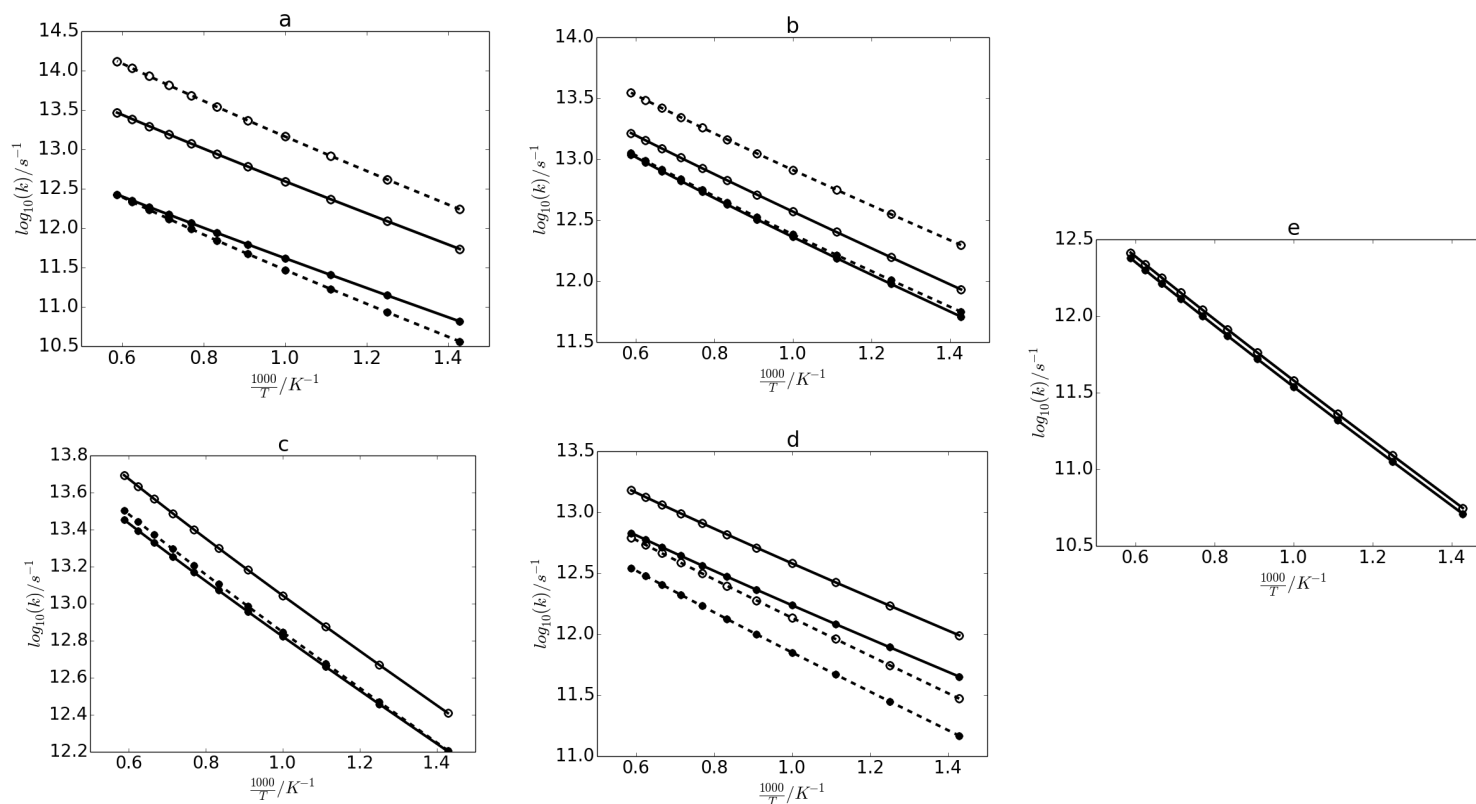
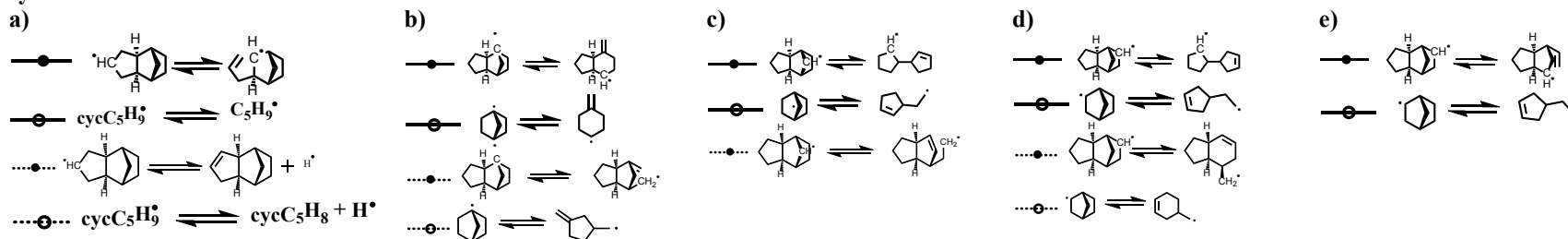
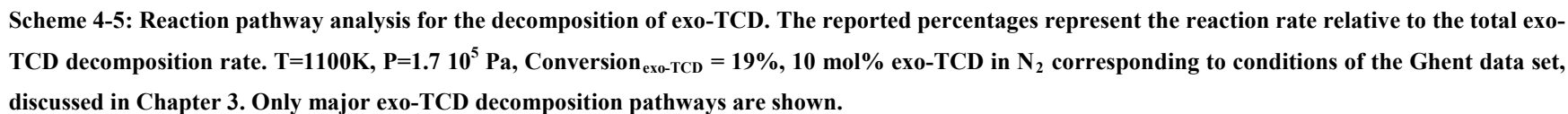


Figure 4-12: Rate coefficients as a function of temperature. Cyclopentyl and norbornyl decomposition reaction rate coefficients from Sirjean et al. [64, 65]. Tricyclodecyl decomposition reaction rate coefficients calculated in this work. Tricyclodecyl decomposition reactions are represented by full symbols.



4.5.3 Main reaction pathways

A schematic representation of the main decomposition pathways of *exo*-TCD can be found in Scheme 4-5. The rates of the important reactions are shown as a percentage relative to the total decomposition rate of *exo*-TCD.



4.5.3.1 Major decomposition pathways of *exo*-TCD

While the thermal decomposition of *exo*-TCD is initiated by the scission of a C-C bond yielding BR1, 85% of the decomposition rate of *exo*-TCD is accounted for by H-abstraction reactions by H at the conditions corresponding to Scheme 4-5. The main decomposition channels involve TCDR5, TCDR8, and TCDR4 with contributions of 33%, 26% and 14% respectively. Decomposition pathways through TCDR7 and TCDR10 are negligible due to the significantly higher bond dissociation energies of the corresponding C-H bonds [48]. The important C₁₀ radicals lead to different primary products: TCDR8 leads primarily to cyclopentene and 1,3-cyclopentadiene, with a minor pathway forming benzene and ethyl, TCDR5 leads to 1,4-pentadiene, 1,3-cyclopentadiene, 3-ethenylcyclopentene, and propene. TCDR4 leads to ethene and 1,3-cyclopentadiene via a retro-Diels-Alder reaction of 2-norbornene. A secondary decomposition channel of TCDR4 leads to tricyclo[5.2.1.0^{2,6}]dec-4-ene and accounts for 45% of the decomposition rate of TCDR4. The channel through TCDR4 is the only significant source of tricyclo[5.2.1.0^{2,6}]dec-4-ene. One of the major uncertainties affecting tricyclo[5.2.1.0^{2,6}]dec-4-ene predictions is the ring strain present in unsaturated species and derived radicals. For example, H-abstraction reactions from tricyclo[5.2.1.0^{2,6}]dec-4-ene result in resonance-stabilized C₁₀H₁₃ species containing rings whose ring strain is unknown.

Ethene is initially formed via various primary decomposition routes starting from BR1, TCDR4, TCDR6, and TCDR8. Once propene attains high enough concentrations, the reaction of propene with H atoms resulting in methyl and ethene forms another major ethene formation channel. Propene is formed via H-abstraction reactions by the allyl radical. The latter is formed through four out of five of the major *exo*-TCD decomposition channels.

The observation that ethene mole fractions are underestimated over most of the data sets at lower temperatures raises questions on the source of the discrepancy. One potential source is the lack of pathways leading towards ethene, e.g. because of the exclusion of reaction families such as 1,4- and 1,5-H-isomerizations. It is possible that the tricyclodecyl or other derived radicals that do not lead to ethene may isomerize to other C₁₀ radicals that do eventually decompose into ethene. Secondly the accuracy of rate coefficients can be questioned. Sensitivity analysis was used to quantitatively assess the importance of rate coefficients. Normalized sensitivity coefficients S_{ij}

for pre-exponential factors A_j of reaction j on the mole fraction of species X_i were calculated using Chemkin-Pro and defined as in Eq. 4-2:

$$S_{ij} = \frac{A_j}{X_i} \cdot \frac{\partial X_i}{\partial A_j} = \frac{\partial \ln X_i}{\partial \ln A_j} \quad \text{Eq. 4-2}$$

Figure 4-13 shows large positive sensitivity coefficients for the major primary decomposition channels through BR1. Smaller positive coefficients are calculated for decomposition reactions of the channels through TCDR6 and TCDR8 since these channels form ethene. 1,3-Cyclopentadiene is slightly overpredicted; it is possible that the deviating ethene concentration can be attributed to small errors on the rate coefficients for the H-abstractions by H from exo-TCD. Moreover, since reverse rate coefficients are calculated through thermodynamic consistency, uncertain estimates of thermochemistry may play a role too.

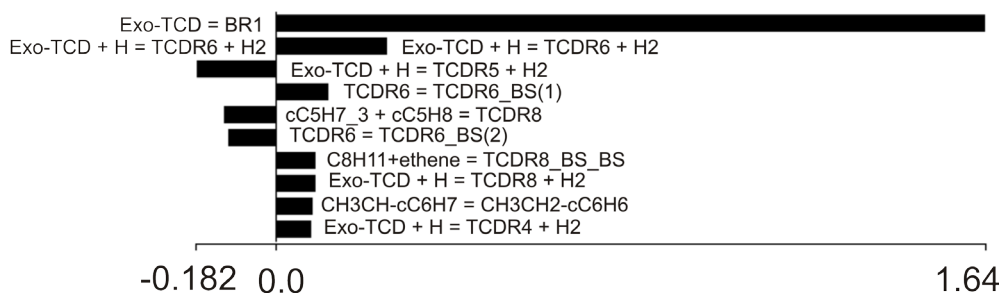
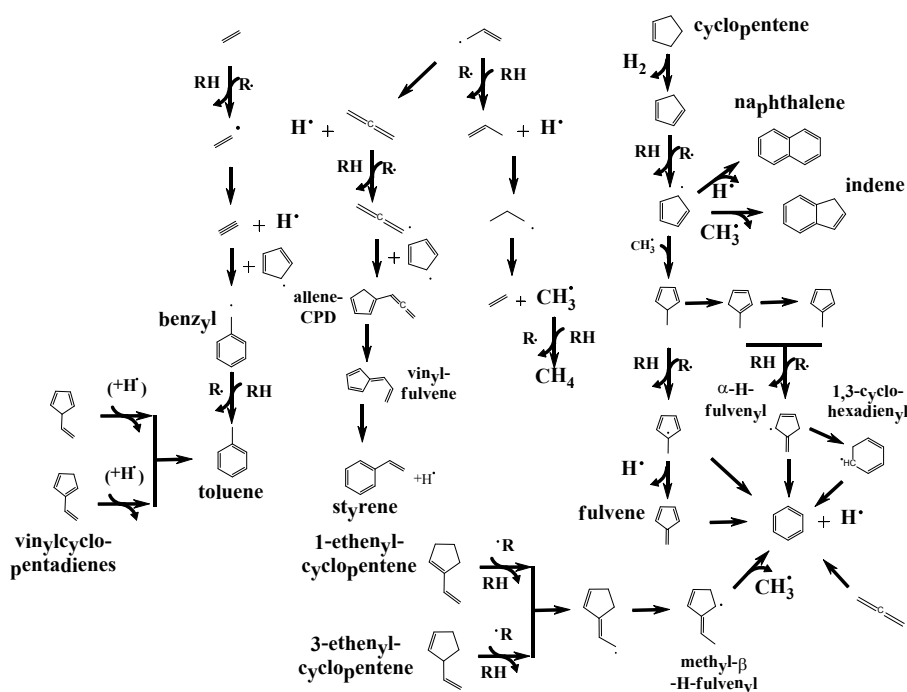


Figure 4-13: Normalized sensitivity coefficients for ethene at $T = 1100\text{K}$, $P = 1.5 \cdot 10^5 \text{ Pa}$, corresponding to an exo-TCD conversion = 16%. “BS” indicates that the species represents is formed through a β -scission reaction of the reactant.

Furthermore, the overprediction of tricyclo[5.2.1.0^{2,6}]-dec-4-ene is correlated with the underprediction of ethene, as less TCDR4 is decomposing through the competing channel that produces ethene. A major factor of uncertainty is the ring strain in the unsaturated tricyclo[5.2.1.0^{2,6}]decene rings. At the conditions in Scheme 4-5, benzene is primarily formed via two channels of primary decomposition of exo-TCD, which account for 90% of the total rate of production of benzene. 46% of benzene is formed through the primary decomposition channel starting from TCDR6 through α -H-fulvenyl, cf. Scheme 4-5. The second important benzene formation pathway starts from TCDR8, and proceeds via a series of rearrangements of C_8H_{11} , eventually resulting in benzene and ethyl.

4.5.3.2 Secondary chemistry and aromatics formation

Many studies of JP-10 pyrolysis indicated the presence of significant quantities of aromatic components, with benzene, toluene, styrene, indene and naphthalene as the most prominent species [10, 12], which was also confirmed in Chapter 3 of this work. The current model suggests that major exo-TCD decomposition pathways lead to hydrogen, ethene, propene, CPD, cyclopentene, and ethenylcyclopentenenes. These primary products represent the hydrocarbon pool from which secondary products such as aromatics are formed. CPD and the derived CPDyl radical play a central role in the formation of aromatic species. A schematic overview of important secondary pathways is depicted in Scheme 4-6.



Scheme 4-6: Pathways of secondary conversion of primary products relevant at $T=1100\text{K}$, $P=1.7 \cdot 10^5 \text{ Pa}$.

At higher exo-TCD conversion H-abstraction reactions convert CPD into CPDyl with allyl and H the most important abstracting radicals. Routes from CPD/CPDyl to indene and naphthalene form the two most important ways of consumption of CPD, accounting both for 30% of the total decomposition rate of CPDyl. Naphthalene formation kinetics were based on Lindstedt et al. [66] It involves a two-step sequence in which CPDyl radicals recombine to the stabilized $\text{C}_5\text{H}_5\text{-C}_5\text{H}_4$ moiety, which rearranges to naphthalene. Kinetics of the indene formation route by addition of CPDyl to CPD originate from Cavallotti et al. [32] reported at $P=1.1 \cdot 10^5 \text{ Pa}$.

Secondary conversion of cyclopentene leads to 1,3-cyclopentadiene, with the direct dehydrogenation pathway accounting for ca. 80% of the total cyclopentene decomposition rate. The conversion of propene produces methyl radical and is the second major source for methyl radicals. The latter recombine with CPDyl yielding 5-methylcyclopentadiene. This C_6H_8 molecule can rearrange into the other methylcyclopentadiene isomers or can undergo H-abstraction resulting in methyl-1,3-cyclopentadien-5-yl and isomeric H-fulvenyl radicals. The latter rearrange via various exit channels to fulvene, 1,3-cyclohexadienyl, and eventually benzene. The kinetics and thermochemistry for these reactions and species were taken from Sharma and Green [27].

Another important aspect of the kinetic model at high *exo*-TCD conversion, is the secondary conversion of ethenylcyclopentenenes. The latter produce fulvene and represents an important pathway to α -H-fulvenyl, cf. Scheme 4-6. Next to the primary decomposition channels from TCDR6, and TCDR8, the hydrogen assisted fulvene isomerization and the decomposition of 5-methyl-1,3-cyclopentadien-5-yl [27] become importance sources of benzene too. A sensitivity analysis of benzene predictions at 1100K is shown in Figure 4-14 and illustrates again the importance of the unimolecular ring opening reaction of *exo*-TCD yielding BR1.

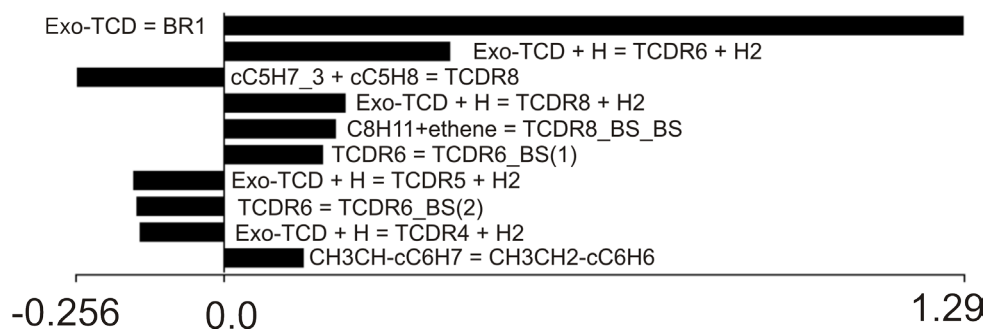


Figure 4-14: Normalized sensitivity coefficients for benzene at $T = 1100K$, $P = 1.7 \cdot 10^5$ Pa, corresponding to an *exo*-TCD conversion = 36%. “BS” indicates that the species represents is formed through a β -scission reaction of the reactant.

Furthermore, the parallel primary decomposition channels of *exo*-TCD are important too since they represent the largest contribution to benzene yields. The majority of the reactions shown in Figure 4-14 were based on estimations from RMG’s databases. Rate coefficients for these reactions based on *ab initio* techniques may therefore be valuable.

4.6 Conclusions

A detailed kinetic model was developed to simulate JP-10 pyrolysis using the Reaction Mechanism Generator program. New elements of this model compared to previously developed JP-10 combustion models are the inclusion of more accurate kinetics for initial exo-TCD decomposition chemistry and the incorporation of many secondary conversion routes to aromatics. A smaller, skeletal model of 1098 reactions was derived from the generated model and can be used as a basis to improve the understanding of the thermal decomposition behavior of strained hydrocarbons in potential advanced fuel applications.

CBS-QB3 calculations of 21 tricyclodecyl radical C-C and C-H β -scissions reactions were carried out to reveal the importance of the various decomposition possibilities of the tricyclodecyl radicals. Significant differences exist between the calculated rate coefficients of the tricyclodecyl radical decomposition reactions with literature values for the rate coefficients of the smaller analogous units of norbornane and cyclopentyl, illustrating that the use of rate coefficients from analogous reactions to model JP-10 decomposition reactions could lead to inaccurate results. CAS-PT2 calculations for the exo-TCD unimolecular C-C scissions confirmed the importance of the initiation route via biradical BR1 and 3-cyclopentylcyclopentene.

The good agreement between model predictions and five independent sets of experimental data was surprising given the uncertainty in many parameters in the model and without adjustment of any of the model parameters to experimental data. It not only indicated the usefulness and validity of the kinetic model over a wide range of conditions but also confirmed the value of RMG as a tool for the kinetic modeling of complex decomposition chemistry of polycyclic hydrocarbons.

Analysis of the kinetic model showed the importance of four important H-abstraction reactions from exo-TCD besides the unimolecular initiation route leading to 3-cyclopentylcyclopentene and indicated that each of these routes lead to primary products such as hydrogen, ethene, propene, 1,3-cyclopentadiene, cyclopentenenes, ethenylcyclopentenenes and benzene. Next to the primary decomposition reactions, a large part of the model involves detailed kinetics for the secondary conversion of primary products towards aromatics such as benzene, toluene and

styrene and PAHs such as indene and naphthalene. Monoaromatics such as benzene, toluene and styrene are formed through the secondary conversion of unsaturated five-membered rings such as ethenylcyclopentenes and through the recombination of methyl with 1,3-cyclopentadienyl. Indene and naphthalene are formed by molecular growth reactions of cyclopentadienyl/cyclopentadiene moieties rather than direct dehydrogenation routes from exo-TCD.

4.7 References

- [1] H.S. Chung, C.S.H. Chen, R.A. Kremer, J.R. Boulton, G.W. Burdette, Recent developments in high-energy density liquid hydrocarbon fuels, *Energy & Fuels*, 13 (1999) 641-649.
- [2] H.A. Meylemans, L.C. Baldwin, B.G. Harvey, Low-Temperature Properties of Renewable High-Density Fuel Blends, *Energy & Fuels*, 27 (2013) 883-888.
- [3] S.L. Outcalt, A. Laesecke, Measurements of Density and Speed of Sound of JP-10 and a Comparison to Rocket Propellants and Jet Fuels, *Energy & Fuels*, 25 (2011) 1132-1139.
- [4] F. Billaud, P. Chaverot, E. Freund, Cracking of Decalin and Tetralin in the Presence of Mixtures of N-decane and Steam at about 810-degrees-C, *Journal of Analytical and Applied Pyrolysis*, 11 (1987) 39-53.
- [5] M. Cooper, J.E. Shepherd, Thermal and Catalytic Cracking of JP-10 for Pulse Detonation Engine Applications, (2002)
- [6] S.P. Heneghan, W.E. Harrison, JP-8+100: The Development of High Thermal Stability Jet Fuel, 6th International Conference on Stability and Handling of Liquid Fuels, (1997).
- [7] H. Huang, D.R. Sobel, L.J. Spadaccini, Endothermic Heat-Sink of Jet Fuels for Scramjet Cooling, 38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, (2002).
- [8] M.B. Colket, L.J. Spadaccini, Scramjet fuels autoignition study, *J. Propul. Power*, 17 (2001) 315-323.
- [9] T. Edwards, Advancements in Gas Turbine Fuels From 1943 to 2005, *Journal of Engineering for Gas Turbines and Power*, 129 (2007) 13-20.
- [10] O. Herbinet, B. Sirjean, R. Bounaceur, R. Fournet, F. Battin-Leclerc, G. Scacchi, P.M. Marquaire, Primary mechanism of the thermal decomposition of tricyclodecane, *Journal of Physical Chemistry A*, 110 (2006) 11298-11314.
- [11] D.F. Davidson, D.C. Horning, M.A. Oehlschlaeger, R.K. Hanson, The Decomposition Products of JP-10, *American Institute of Aeronautics and Astronautics Journal*, 3707 (2001).
- [12] P.N. Rao, D. Kunzru, Thermal cracking of JP-10: Kinetics and product distribution, *Journal of Analytical and Applied Pyrolysis*, 76 (2006) 154-160.
- [13] S. Nakra, R.J. Green, S.L. Anderson, Thermal decomposition of JP-10 studied by micro-flowtube pyrolysis-mass spectrometry, *Combustion and Flame*, 144 (2006) 662-674.
- [14] B. Van Devener, S.L. Anderson, Breakdown and combustion of JP-10 fuel catalyzed by nanoparticulate CeO₂ and Fe₂O₃, *Energy & Fuels*, 20 (2006) 1886-1894.
- [15] Y. Xing, W.J. Fang, W.J. Xie, Y.S. Guo, R.S. Lin, Thermal Cracking of JP-10 under Pressure, *Industrial & Engineering Chemistry Research*, 47 (2008) 10034-10040.
- [16] K. Wohlwend, L.Q. Maurice, T. Edwards, R.C. Striebich, M. Vangsness, A.S. Hill, Thermal stability of energetic hydrocarbon fuels for use in combined cycle engines, *J. Propul. Power*, 17 (2001) 1258-1262.
- [17] R. Seiser, U. Niemann, K. Seshadri, Experimental study of combustion of n-decane and JP-10 in non-premixed flows, *Proceedings of the Combustion Institute*, 33 (2011) 1045-1052.
- [18] S.C. Li, B. Varatharajan, F.A. Williams, Chemistry of JP-10 ignition, *American Institute of Aeronautics and Astronautics J.*, 39 (2001) 2351-2356.
- [19] G.R. Magoon, J. Aguilera-Iparraguirre, W.H. Green, J.J. Lutz, P. Piecuch, H.W. Wong, O.O. Oluwole, Detailed chemical kinetic modeling of JP-10 (exo-tetrahydrodicyclopentadiene) high-temperature oxidation: Exploring the role of biradical species in initial decomposition steps, *International Journal of Chemical Kinetics*, 44 (2012) 179-193.

- [20] J. Song, S. Raman, J. Yu, C.D. Wijaya, G. Stephanopoulos, W.H. Green, Development of automatic chemical reaction mechanism generation software using object-oriented technology., Abstracts of Papers of the American Chemical Society, 226 (2003) U530-U531.
- [21] J. Song, Building Robust Chemical Reaction Mechanisms: Next Generation of Automatic Model Construction Software. Ph.D thesis, MIT (2004).
- [22] R.G. Susnow, A.M. Dean, W.H. Green, P. Peczak, L.J. Broadbelt, Rate-based construction of kinetic models for complex systems, *Journal of Physical Chemistry A*, 101 (1997) 3731-3740.
- [23] M.R. Harper, K.M. Van Geem, S.P. Pyl, G.B. Marin, W.H. Green, Comprehensive reaction mechanism for n-butanol pyrolysis and combustion, *Combustion and Flame*, 158 (2011) 16-41.
- [24] M.K. Sabbe, K.M. Van Geem, M.-F. Reyniers, G.B. Marin, First Principle-Based Simulation of Ethane Steam Cracking, *AIChE J.*, 57 (2011) 482-496.
- [25] K.M. Van Geem, M.F. Reyniers, G.B. Marin, J. Song, D.M. Matheu, W.H. Green, Automatic Reaction Network Generation using RMG for Steam Cracking of n-Hexane, *AIChE J.*, 52 (2006) 718-730.
- [26] K.J. Hughes, T. Turanyi, A.R. Clague, M.J. Pilling, Development and testing of a comprehensive chemical mechanism for the oxidation of methane, *International Journal of Chemical Kinetics*, 33 (2001) 513-538.
- [27] S. Sharma, W.H. Green, Computed Rate Coefficients and Product Yields for $c\text{-C}_5\text{H}_5 + \text{CH}_3 \rightarrow \text{Products}$, *Journal of Physical Chemistry A*, 113 (2009) 8871-8882.
- [28] S. Sharma, M.R. Harper, W.H. Green, Modeling of 1,3-hexadiene, 2,4-hexadiene and 1,4-hexadiene-doped methane flames: Flame modeling, benzene and styrene formation, *Combustion and Flame*, 157 (2010) 1331-1345.
- [29] M.B. Colket, D.J. Seery, Reaction mechanisms for toluene pyrolysis, Symposium (International) on Combustion, 25 (1994) 883-891.
- [30] J.L. Emdee, K. Brezinsky, I. Glassman, A Kinetic Model for the Oxidation of Toluene near 1200-K, *Journal of Physical Chemistry*, 96 (1992) 2151-2161.
- [31] J.A. Miller, S.J. Klippenstein, The recombination of propargyl radicals and other reactions on a C_6H_6 potential, *Journal of Physical Chemistry A*, 107 (2003) 7783-7799.
- [32] C. Cavallotti, D. Polino, A. Frassoldati, E. Ranzi, Analysis of Some Reaction Pathways Active during Cyclopentadiene Pyrolysis, *Journal of Physical Chemistry A*, 116 (2012) 3313-3324.
- [33] L. Zhang, J. Cai, T. Zhang, F. Qi, Kinetic modeling study of toluene pyrolysis at low pressure, *Combustion and Flame*, 157 (2010) 1686-1697.
- [34] S.W. Benson, *Thermochemical Kinetics*, John Wiley & Sons, New York, 1976.
- [35] G.R. Magoon, W.H. Green, 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*, 52 (2013) 35-45.
- [36] G. Magoon, W.H. Green, O.O. Oluwole, H.W. Wong, S.E. Albo, D.K. Lewis, 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, (2010).
- [37] B.E. Poling, J.M. Prausnitz, J.P. O'Connell, *The properties of gases and liquids*, McGraw-Hill, New York, 2001.
- [38] K.B. Wiberg, The Concept of Strain in Organic-Chemistry, *Angew. Chem.-Int. Edit. Engl.*, 25 (1986) 312-322.
- [39] J. Gasteiger, O. Dammer, Automatic Estimation of Ring Strain Energies, *Tetrahedron*, 34 (1978) 2939-2945.
- [40] S.P. Verevkin, V.N. Emel'yanenko, A.A. Pimerzin, E.E. Vishnevskaya, Thermodynamic Analysis of Strain in Heteroatom Derivatives of Indene, *The Journal of Physical Chemistry A*, 115 (2011) 12271-12279.
- [41] S.P. Verevkin, V.N. Emel'yanenko, A.A. Pimerzin, E.E. Vishnevskaya, Thermodynamic Analysis of Strain in the Five-Membered Oxygen and Nitrogen Heterocyclic Compounds, *Journal of Physical Chemistry A*, 115 (2011) 1992-2004.
- [42] G. Landrum, RDKit: Open-source cheminformatics; <http://rdkit.sourceforge.net>, (2006).
- [43] T.H. Lay, J.W. Bozzelli, A.M. Dean, E.R. Ritter, Hydrogen-atom bond increments for calculation of thermodynamic properties of hydrocarbon radical species, *Journal of Physical Chemistry*, 99 (1995) 14514-14527.
- [44] P.M. Nunes, S.G. Estacio, G.T. Lopes, B.J. Costa Cabral, R.M. Borges dos Santos, J.A. Martinho Simoes, C-H Bond Dissociation Enthalpies in Norbornane. An Experimental and Computational Study, *Organic Letters*, 10 (2008) 1613-1616.
- [45] D.F. Mcmillen, D.M. Golden, Hydrocarbon Bond-Dissociation Energies, *Annual Review of Physical Chemistry*, 33 (1982) 493-532.
- [46] Z.X. Tian, A. Fattahi, L. Lis, S.R. Kass, Cycloalkane and cycloalkene C-H bond dissociation energies, *J. Am. Chem. Soc.*, 128 (2006) 17087-17092.

- [47] Y. Feng, L. Liu, J.T. Wang, S.W. Zhao, Q.X. Guo, Homolytic C-H and N-H bond dissociation energies of strained organic compounds, *Journal of Organic Chemistry*, 69 (2004) 3129-3138.
- [48] J.M. Hudzik, R. Asatryan, J.W. Bozzelli, Thermochemical Properties of *exo*-Tricyclo 5.2.1.0(2,6) decane (JP-10 Jet Fuel) and Derived Tricyclodecyl Radicals, *Journal of Physical Chemistry A*, 114 (2010) 9545-9553.
- [49] O. Castano, R. Notario, J.L.M. Abboud, R. Gomperts, R. Palmeiro, L.M. Frutos, Organic thermochemistry at high ab initio levels. 2. Meeting the challenge: Standard heats of formation of gaseous norbornane, 2-norbornene, 2,5-norbornadiene, cubane, and adamantane at the G2 level, *Journal of Organic Chemistry*, 64 (1999) 9015-9018.
- [50] R. Walsh, J.M. Wells, Enthalpy of Formation of Bicyclo[2.2.1]hepta-2,5-diene - Thermodynamic Functions of Bicyclo[2.2.1]heptane and Bicyclo[2.2.1]hepta-2,5-diene, *Journal of Chemical Thermodynamics*, 7 (1975) 149-154.
- [51] X.W. An, H.P. Zhu, R.H. Hu, Heats of Combustion and Formation of Norbornane, *Thermochimica Acta*, 121 (1987) 473-477.
- [52] P. Celani, H.J. Werner, Multireference perturbation theory for large restricted and selected active space reference wave functions, *J. Chem. Phys.*, 112 (2000) 5546-5557.
- [53] H.J. Werner, Third-order multireference perturbation theory - The CASPT3 method, *Molecular Physics*, 89 (1996) 645-661.
- [54] H.-J. Werner, P.J. Knowles, G. Knizia, F.R. Manby, M. Schutz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K.R. Shamasundar, T.B. Adler, R.D. Amos, A. Bernhardsson, A. Berning, D.L. Cooper, M.J.O. Deegan, A.J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Koppl, Y. Liu, A.W. Lloyd, R.A. Mata, A.J. May, S.J. McNicholas, W. Meyer, M.E. Mura, A. Nicklass, D.P. O'Neill, P. Palmieri, D. Peng, K. Pfluger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A.J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, MOLPRO, version 2012.1, a package of ab initio programs, (2012)
- [55] B. Sirjean, P.A. Glaude, M.F. Ruiz-Lopez, R. Fournet, Detailed kinetic study of the ring opening of cycloalkanes by CBS-QB3 calculations, *Journal of Physical Chemistry A*, 110 (2006) 12693-12704.
- [56] B. Sirjean, R. Fournet, P.-A. Glaude, M.F. Ruiz-López, Extension of the composite CBS-QB3 method to singlet diradical calculations, *Chem. Phys. Lett.*, 435 (2007) 152-156.
- [57] C. Eckart, The penetration of a potential barrier by electrons, *Physical Review*, 35 (1930) 1303-1309.
- [58] S. Sharma, M.R. Harper, W.H. Green, CANTHERM, (2010).
- [59] CHEMKIN-PRO v150101, Reaction Design, San Diego, (2009).
- [60] K.M. Van Geem, S.P. Pyl, G.B. Marin, M.R. Harper, W.H. Green, Accurate High-Temperature Reaction Networks for Alternative Fuels: Butanol Isomers, *Industrial & Engineering Chemistry Research*, 49 (2010) 10399-10420.
- [61] T.J.H. Bruno, M.L.; Laesecke, A.; Lemmon, E.W.; Perkins, R.A., Thermochemical and Thermophysical Properties of JP-10, Report Prepared for: Fuels Branch, Turbine Engine Division Propulsion Directorate, AFRL Wright Patterson Air Force Base, Ohio, (2006)
- [62] H.E. O'Neal, S.W. Benson, Biradical Mechanism in Small Ring Compound Reactions, *Journal of Physical Chemistry*, 72 (1968) 1866-1887.
- [63] A.M. Dean, J.W. Bozzelli, Combustion chemistry of nitrogen, in: *Gas-phase combustion chemistry*, ed. by W. Gardiner, Springer, 2000, pp. 125-341.
- [64] B. Sirjean, O. Herbinet, P.-A. Glaude, M. Ruiz-Lopez, R. Fournet, Theoretical Study of the Thermal Decomposition of a Jet Fuel Surrogate, WSS/CI Spring 2008 Meeting, (2007).
- [65] B. Sirjean, P.A. Glaude, M.F. Ruiz-Lopez, R. Fournet, Theoretical Kinetic Study of Thermal Unimolecular Decomposition of Cyclic Alkyl Radicals, *Journal of Physical Chemistry A*, 112 (2008) 11598-11610.
- [66] P. Lindstedt, L. Maurice, M. Meyer, Thermodynamic and kinetic issues in the formation and oxidation of aromatic species, *Faraday Discussions*, 119 (2001) 409-432.

Chapter 5: Genesys: basic concepts and functionality

This chapter is based on the following paper:

Vandewiele, N. M.; Van Geem, K. M.; Reyniers, M.-F.; Marin, G. B., “Genesys: Kinetic model construction using chemoinformatics.” *Chemical Engineering Journal* 2012, 207–208 (0), 526-538.

5.1 Abstract

A new tool for the automatic generation of kinetic models called Genesys is presented. A rule-based network generation methodology creates dedicated kinetic models based on the chemical knowledge of the end-user by iterating over all user-defined reaction families. The latter are constrained as much as possible to avoid the creation of unimportant species and insignificant reactions. Through integration with existing open-source chemoinformatics libraries the methodology is not limited to specific chemical elements or to specific chemistries. Chemical data is stored in databases, and is fully separated from the network generation program. This was accomplished by adopting a sub-molecular pattern language called SMARTS, and enables the use of quantitative structure property relationships such as group contribution methods. To demonstrate this, a Benson group additivity method for the estimation of ideal gas phase thermochemical properties is described and compared to a benchmark database. A group additivity scheme for the estimation of Arrhenius parameters was implemented and validated.

Keywords: Kinetic model; reaction network generation; chemoinformatics

5.2 Introduction

Kinetic models that not only provide information on the types of molecules and reactions involved in a chemical process, but also on the product distribution as a function of reaction conditions are indispensable in the design and optimization of chemical applications. For chemical processes involving a large number of species and reactions such as combustion processes, atmospheric chemistry, soot formation, etc... these detailed kinetic models are characterized by a complexity that makes the manual construction of them sometimes impossible to achieve and certainly tedious and error-prone. In the majority of the kinetic model construction programs graph theory, an abstraction of chemical species into mathematical graphs is applied to represent chemical species [1]. In most of the programs graphs are used with atoms represented as vertices and chemical bonds as edges, thus retaining the connectivity information between individual atoms.

The exponential growth of computational power[2] and the ever-growing amount of available chemical data in digital databases [3] enabled the field of chem(o)informatics and bioinformatics [4] to benefit from an advanced species representation and algorithms in areas such as structure-based virtual screening [5], pharmacophore discovery [6], and protein structure elucidation [7] to mention only a few. Despite the maturity of chemoinformatics, with several available open-source chemoinformatics libraries [8-11], the use of generally applicable chemoinformatics algorithms and species representation in the construction of kinetic models remains limited.

The lack of stand-alone chemoinformatics libraries and algorithms in the past implied that researchers, when developing programs for the generation of kinetic models [12-18] and the estimation of thermochemical and kinetic data [19-21] needed to develop their own species representation and accompanying algorithms for species isomorphism, reaction identification, and reactant transformations, etc.... As a consequence, the species representation and accompanying algorithms were designed for a specific class of components of interest e.g. hydrocarbons in PRIM [14], and made it difficult to extend the use of the program to other molecules outside the scope of the original class. Often, the chemical data used for the estimation of molecule properties was interwoven with the program code, making it difficult to expand or revise the used chemical data without inspecting the code itself [22].

This chapter presents a new kinetic model generation tool, called Genesys, which creates detailed chemical kinetic models. The latter solely consist of elementary reactions, i.e. reactions characterized by a single transition state. The primary goal is to discuss the architecture of this tool including its integration with state-of-the-art chemoinformatics libraries. More specifically, it will be explained how kinetic models can be automatically generated fully based on the user's expertise and how all chemical information is organized in databases that are not interwoven with the network generation program itself. These principles will be applied to quantitative structure-property relationships that are helpful in reaction engineering: the determination of ideal gas phase thermochemical properties, and the calculation of rate coefficients. Finally, the generation of detailed kinetic models will be analyzed.

5.3 Methodology and algorithms

The automated construction of kinetic models is divided into three sequential parts, cf. Figure 5-1 : 1) the input of required information 2) the processing step consisting of the generation of the actual reaction network and the assignment of thermochemical and Arrhenius parameters to species and reactions and 3) the post-processing step producing human readable information on the kinetic model as well as formats that allow the integration with third-party tools for the analysis or further use of the kinetic model.

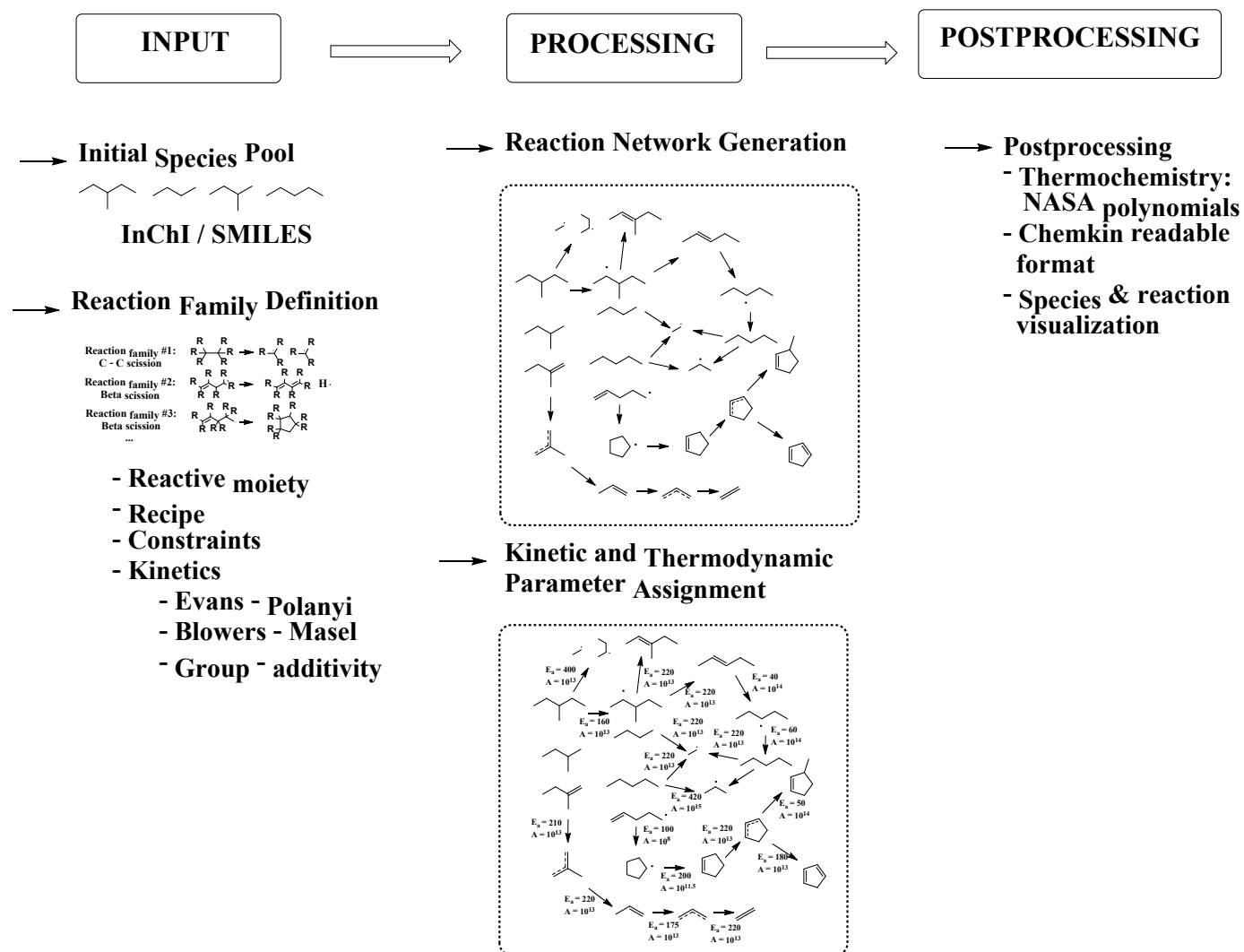


Figure 5-1: Architecture of Genesys divided into 3 sequential parts: input, processing and output.

Genesys creates kinetic models based on the widely adopted strategy of condensing the relevant chemistry that molecules can undergo in reaction families [23]. The reaction family (class, rule) concept is the direct extension of the idea that sub-molecular patterns and more specifically functional groups inside a chemical species govern the reactivity of a molecule rather than the molecule in its entirety. Without reaction families, every chemical bond and every atom should *a priori* be eligible to undergo a chemical reaction leading to a combinatorial explosion of species and reactions and a “non-sense” reaction network. Moreover, if one considers that the chemical environment to the reactive moiety, i.e. a set of atoms that are required for this reaction family to be executed, has a diminishing impact with distance on the actual kinetics of the reaction, then empirical and semi-empirical kinetic expressions can be employed once high-quality data is available for a few members of the reaction family. As a corollary, small molecules, of which the chemical structure as a whole influences the reaction rates, should be treated separately. Hence, network generation programs typically start with small chemistry models such as the “C₀-C₁-C₂” reaction base in EXGAS [24] or “seed mechanisms” in the Reaction Mechanism Generator (RMG) [18] as a base for the construction of the larger models.

Genesys is written in Java and uses the chemoinformatics library “Chemistry Development Kit” v1.4.5 (CDK) [8, 9] for the internal representation of chemical objects. Besides CDK’s support of nearly all of the chemical elements it provides easy access to routines converting line identifiers such as InChI [25, 26] or SMILES [27, 28], to the internal representation and vice versa, and algorithms required for reaction network generation such as graph isomorphism, substructure matching, and aromaticity identification. Moreover, CDK is compatible with a cloud of open source chemoinformatics projects called the “Blue Obelisk Movement” [29], making it easy to integrate it with third party tools, e.g. for visualization of structures or reactions.

5.3.1 Input

An initial set of species is the starting point for the creation of new products and the gradual expansion of the reaction network by the application of the user-defined reaction families. These species can represent the feedstock of the chemical process that is modeled or species that are likely not formed during the network generation but should be still be included because they play an important role. Both InChI and SMILES line identifiers can be used to specify these species.

The definition of all reaction families is separated from the code itself in user-supplied libraries, cf. Figure 5-2. This externalization of not only chemical data, but also of the creation of products gives full control, flexibility and responsibility to the user to allow only the chemistry of interest. Moreover, the way reactants are identified and converted into products by the definition of the reaction family is non-specific to the type of molecule, chemical element or types of transformation. The definition of a reaction family includes: 1) a description of the reactive moiety required inside potential reactant molecules, and of the atoms constituting the reactive moiety, 2) a recipe-like scheme where keywords specifying the type of transformation are combined with symbols representing the atoms of the reactive moiety, 3) information on how Arrhenius parameters of a reaction belonging to this reaction family are calculated.

[-] [e] config	
[e] inp-temperature	1000
[-] [e] inp-reaction-family	
[a] name	h abstractions to a carbon centered group by a carbon centered radical - Phys. Chem. Chem. Phys., 2010, 12, 1278–1298
[e] bimolecular	true
[-] [e] inp-recipe	
[-] [e] inp-transformation	
[a] type	FORM_BOND
[a] centers	A,B
[+] [e] inp-transformation	
[+] [e] inp-transformation	
[+] [e] inp-transformation	
[-] [e] inp-reactant	
[a] value	1
[a] smarts	[C;v3]
[-] [e] inp-reactive-center	
[a] value	A
[a] smarts	[\$([C;v3])]
[-] [e] inp-molecule-constraint	
[a] type	SINGLEELECTRONCOUNT
[a] limit	MAX
[a] value	1
[+] [e] inp-molecule-constraint	
[+] [e] inp-molecule-constraint	
[+] [e] inp-reactant	
[-] [e] inp-kinetics	
[a] type	GROUP_ADDITIVITY
[a] path	h abstractions to a carbon centered group by a carbon centered radical - Phys. Chem. Chem. Phys., 2010, 12, 1278–1298
[-] [e] inp-tunneling	
[a] expression	$1 + (162/T)^3 * E_a + (2.71E-6) * 2.7^{-(T-300)/26} * (E_a^4)$
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	
[+] [e] inp-reaction-family	

Figure 5-2: Excerpt from a reaction family definition in xml-format for the H-abstraction from a hydrocarbon by a carbon-centered radical.

All species that are eligible to undergo a specific reaction family are characterized by a common reactive moiety, and some of the atoms will undergo the actual transformations in terms of atom connectivity, order of connected bonds, etc... For example, the β -scission of a carbon-hydrogen bond in a γ,δ -alkenyl radical requires a set of 5 atoms, 4 carbons and 1 hydrogen, cf. Figure 5-3.

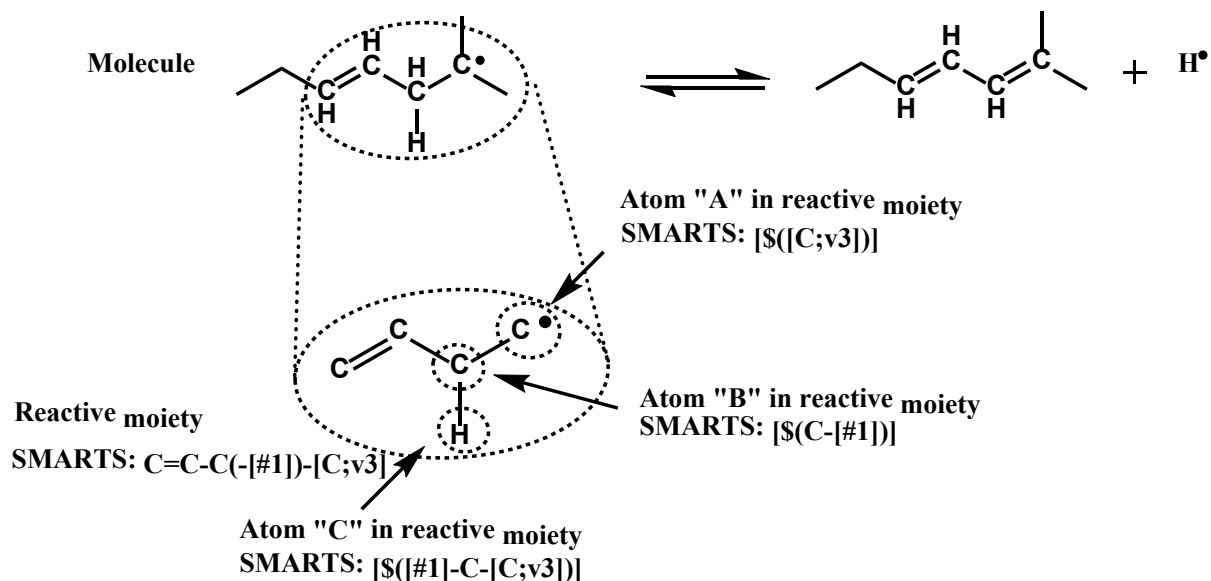


Figure 5-3: Reactive moiety atoms required for the β -scission of a carbon-hydrogen bond in a γ,δ -alkenyl radical.

In order to indicate which reactive moiety is needed for a particular reaction family without hard coding these rules, a language that expresses atom environment patterns is needed. Genesys uses the SMARTS[®] language [30], originally developed and maintained by Daylight Inc., that enables to specify both very general or very specific sub-molecular features and patterns. While previous reaction network generation programs [31, 32] tailored the SMARTS vocabulary with custom symbols and notations, Genesys avoids balkanization into SMARTS dialects, by exclusively adopting SMARTS specifications by Daylight Inc. This is made possible thanks to the SMARTS interpreter in CDK that supports almost the complete functionality of the pattern language. First, all possible configurations of atoms inside the target molecule that comply with the description of the SMARTS query are identified. Next, the atoms of the reactive moiety that take part in the transformation from reactants to molecules are also defined by a SMARTS pattern description, cf. Figure 5-3. Once a reactive moiety has been detected in a molecule, each of the atoms of the reactive moiety is then identified again with the given SMARTS description.

To describe the transformations the atoms of the reactive moiety undergo through the course of a reaction, a set of atom, bond and electronic changes is created by the user, called a “recipe” by Song et al.[33] . With the help of a number of pre-defined transformations, the user is able to describe any modification in atom connectivity, bond orders, presence of charge or unpaired electrons between reactants and products. E.g. the unimolecular retro-ene reaction converts an acyclic 1-alkene into smaller alkenes via a six-membered transition state, cf. Figure 5-4. A total of six atoms (reactive center atoms) undergo six transformations defined in the recipe of the retro-ene reaction family.

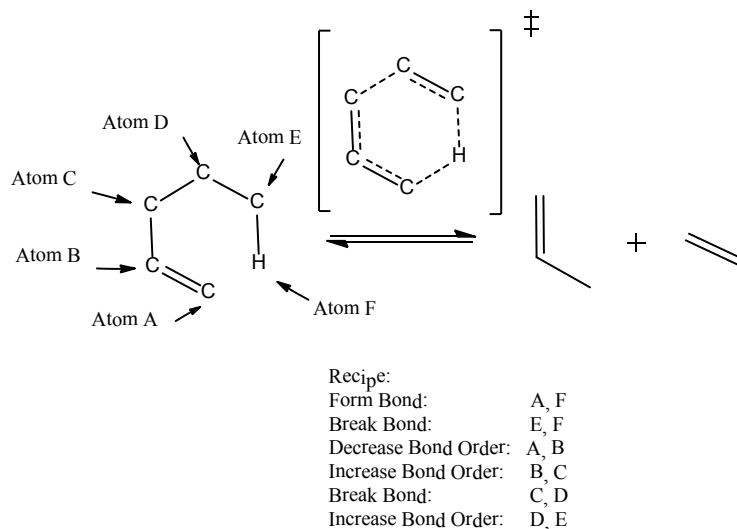


Figure 5-4: Example of the execution of the reaction family recipe with the application of six transformations on six atoms in the reactive moiety: the transformation of a 1-alkene via the pericyclic retro-ene reaction via a six-membered transition state denoted as \ddagger .

5.3.2 Processing

The processing section of Genesys involves the actual generation of the reaction network, consisting of a list of unique species and a list of unique elementary reactions between those species, and the creation of a kinetic model, with thermochemical and Arrhenius parameters assigned to species and reactions. In both steps, there is a need to perceive sub-molecular patterns of atoms inside a molecule, where an atom-to-atom mapping between the sought pattern and the atoms in the target molecule is required.

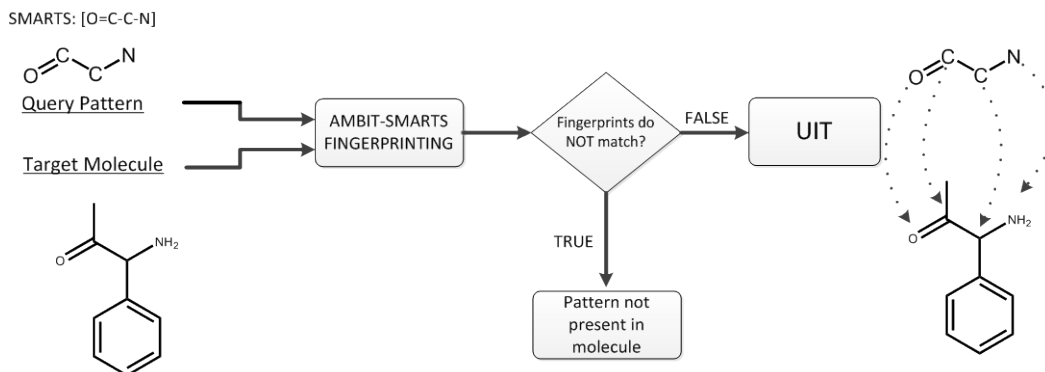


Figure 5-5: Substructure matching strategy including the pre-screening fingerprint generation step via AMBIT-SMARTS prior to the computationally expensive sub-graph isomorphism algorithm (UIT, Universal Isomorphism Tester).

These substructure searches are accomplished by sub-graph isomorphism algorithms, known to be a non-polynomial-complete (NP-complete) [34] problem. This implies that substructure searches for increasingly large molecules will, given a worst case scenario, eventually become computationally infeasible. Moreover, large databases with sub-molecular patterns and associated values need to be queried when quantitative structure-property relationships for the estimation of properties of species and reactions are used. A dependency between the size of the databases and the computation time is expected, since the database will be iterated over until a match is found between a sub-molecular pattern and the atoms in a target molecule. In a worst case scenario of not finding an apt database entry for a particular atom, this leads to testing the entire database of sub-molecular patterns. Structuring databases into hierarchical trees [33] drastically diminish the database size dependency of the searching algorithm from $O(N)$ to $O(\log N)$ where N is the size of the database. Genesys uses a substructure searching algorithm developed by Tonnelier et al. [35], implemented in CDK and known as “UniversalIsomorphismTester, (UIT)”. Prior to the detailed substructure mapping, a pre-screening stage is performed using AMBIT-SMARTS v2 [36], cf. Figure 5-5. This open source software package performs a pre-screening step, generating structural fingerprints [4] of the target structure and the query pattern. Genesys only retains those query patterns for the rigorous sub-graph mapping algorithm that pass the pre-screening step by AMBIT-SMARTS.

5.3.2.1 Network generation

Genesys adopts a rule-based termination criterion to prevent endless generation of new reactions and species in similar ways to what was done in network generation programs such as COMGEN [31] and RING [32]. In a rank-based criterion [37] the maximum number of reaction steps between the initial species pool and a given product species is applied. In a rate-based criterion [38] the rate and species concentration information is used to select kinetically significant reactions. In contrast a rule-based control scheme applies constraints on the applicability of a reaction family and thus adopts an a priori network reduction paradigm based on chemical principles.

Constraints (rules) prevent the reaction family description from generating chemical reactions that are kinetically insignificant or containing chemical species with insignificant concentrations and reflect the chemical knowledge of the user expert regarding the chemical process in the operating conditions at hand. Two types of constraints are envisioned for each reaction family: molecular and atomic constraints. Molecular constraints concern global molecular features of a candidate species. For example, in steam cracking of hydrocarbons acyclic radicals with more than five carbon atoms are considered μ radicals [14, 39], following the nomenclature of $\beta\mu$ rules of Goldfinger-Letort-Niclause [40], i.e. only unimolecular reactions are considered significant, cf. Figure 5-6. Thus, bimolecular reaction families defined in the framework of steam cracking would thus require a molecular constraint specifying the maximum number of carbon atoms present in a candidate chemical species.

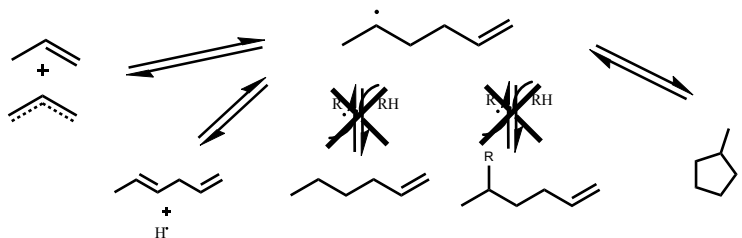


Figure 5-6: Example of a μ radical (5-hexenyl) radical for which only unimolecular reactions are kinetically significant.

Atomic constraints feature restrictions on a particular atom, e.g. the electronic state, the hybridization. For example, in pyrolysis conditions carbon atoms that are part of an aromatic nucleus are relatively unreactive compared to aliphatic carbon atoms. Hence, it can be specified

that they do not undergo the reaction families the atom would undergo if the atom possessed an identical electronic configuration. Therefore, the addition of a radical atom to a double bonded carbon, would be restricted to non-aromatic carbon atoms. As stated previously, in Genesys constraints can be specified per reaction family, i.e. every reaction family has its own molecular, atomic constraints assigned to it, cf. Figure 5-2. Constraints can be added by the user in several ways: through the definition of the SMARTS string by making the SMART string narrower, e.g. by using the logical NOT operator (represented by the ‘!’ symbol) allowing to exclude specific atoms, bonds, atom types, or molecular features, and by pre-defined keywords, as proposed by Rangarajan et al. [32]. The latter is particularly useful for constraints such as “the total number of carbon atoms ” that are more difficult to include in the SMARTS string describing the reactive moiety.

The algorithm that generates and systematically enlarges the reaction network starts with a pool of initial species, cf. Figure 5-7.

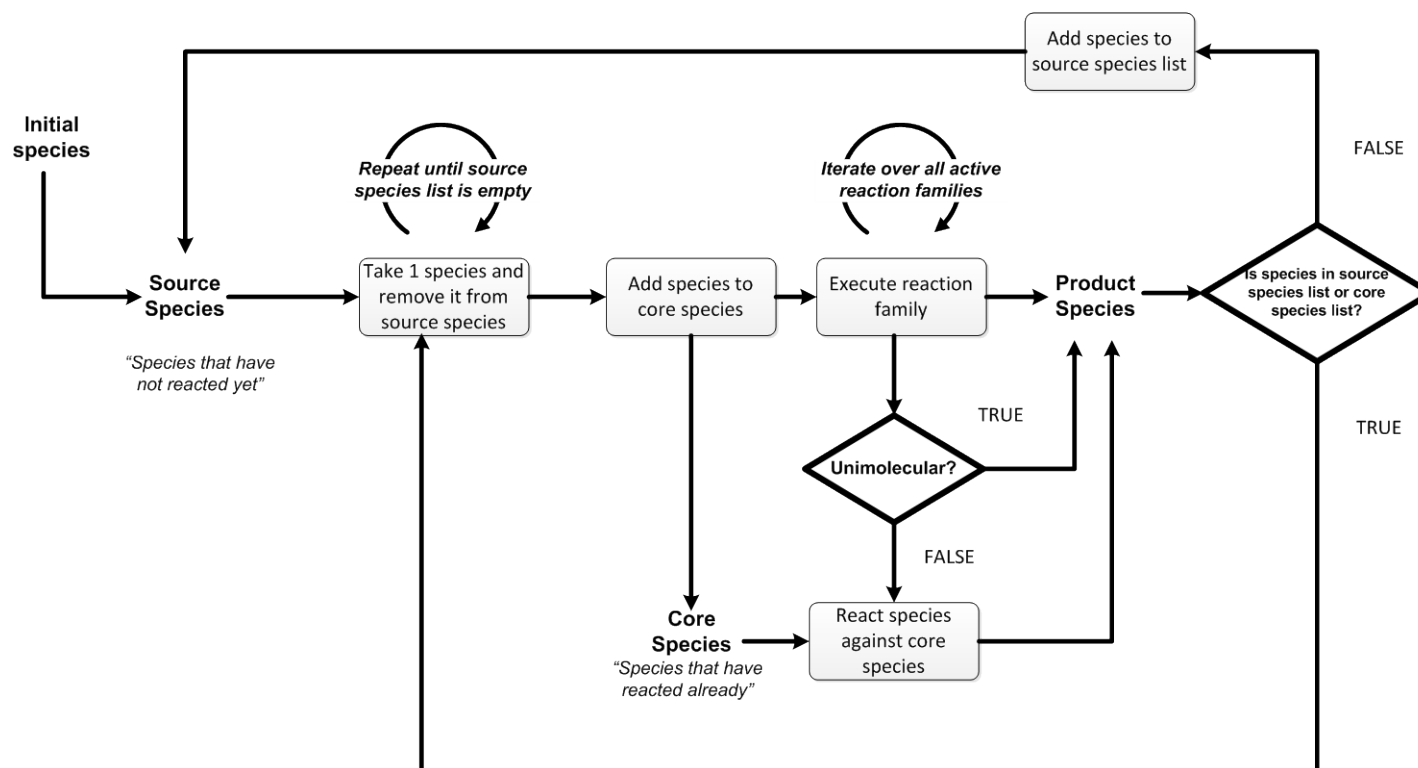


Figure 5-7: Network generation algorithm which gradually expands the reaction network by iteratively reacting unreacted species according to the defined reaction families.

Next, a species from this pool of molecules that has not yet been subjected to the set of defined reaction families (“unreacted species”) is iteratively tested whether it is eligible to undergo a reaction family. Molecular constraints are verified before the substructure algorithm scans the species for a particular reactive moiety. When product species emerge after the execution of the reaction family they are added to the list of unreacted species under the condition that they did not appear in the list of unreacted species or in the list of reacted species. In order to determine the uniqueness of species in a kinetic model, a method of comparing chemical structures needs to be devised. In the present work a canonical string identifier is generated for each molecule and stored (cached) inside the chemical object representing that molecule. In this way, the computationally expensive InChI generation is limited to only newly created species. The reader is referred to Section 1.2.1 of Chapter 1 for further information on canonical identifiers. Using graph isomorphism algorithms to verify species uniqueness is an alternative and has the advantage that rapid pre-screening techniques such as fingerprinting can be applied prior to the actual graph isomorphism.

A list of unique reactions in the reaction network is gradually expanded with new reactions. The reaction network enlargement phase finalizes with an empty pool of unreacted species. The uniqueness of reactions is determined by comparison of the transition state (TS) graph stored for every elementary reaction. The transition state graph is derived from the reactant graphs by “coloring” the bonds that undergo the transformations, with “transition-state”-labels. A fingerprint with MACSS keys [41] is generated for every TS graph and used in a pre-screening stage prior to the graph isomorphism step.

5.3.2.2 Thermochemical properties assignment

In the current method of determining the thermochemistry of a molecule, it is first verified whether thermochemical properties of the molecule are present in one of the user-defined databases by InChI comparison of the candidate molecule and the database entries. If this is not the case, estimation techniques are applied to obtain thermochemical properties for that species. Resonance structures are generated for a candidate molecule, and thermochemical properties for every resonance structure will be generated. Since the different resonance structures may lead to

the identification of different GAVs or HBIs this may lead to different resulting values for the thermochemical properties of the resonance structure. Therefore, the resonance structure with the lowest, calculated standard enthalpy of formation at the designated temperature will be used to provide the thermochemistry of the molecule, following the guideline by Sabbe et al. [42].

A Benson group additivity scheme for ideal gas phase thermochemical properties [43] is implemented. This predictive method calculates these properties of the molecule by adding contributions of sub-molecular groups present in the molecule and is generally applicable regardless of the involved chemical elements, as will be shown in the case study further on. These contributions are derived by regression of experimentally obtained data or *ab initio* calculations of a training set of species. Four types of Benson contributions are considered: group additive values (GAVs) for polyvalent (ligancy > 1) atom-centered groups that take into account the nature of the nearest-neighbor ligands, ring-strain corrections (RSCs) that account for the presence of rings compared to the acyclic counterparts. Non-nearest neighbor interactions (NNIs) that model interactions between non-bonded sub-molecular fragments such as alkane 1,4-gauche, 1,5-dimethyl repulsion, and cis interactions, finally resonance corrections (RESs) that account for the stabilization effect of electron interactions in a molecule.

Next to the Benson group additivity scheme, the hydrogen-atom bond increment (HBI) method of Lay et al. [44] is implemented for the estimation of radical species. This method determines the thermochemistry of radical species by adding an increment to the parent molecule, which has the same structure of the radical except that an extra hydrogen atom is attached to the atom that contains the radical site.

Benson groups and HBIs are represented by a SMARTS description that maps the sub-molecular group associated with the SMARTS onto a vector containing values for the standard enthalpy of formation, and entropy at 298K, and heat capacity at 7 distinct temperatures.

The algorithm that iteratively assigns GAVs to atoms in a candidate species is depicted in Figure 5-8. For each of the GAV entries in the database defined by the SMARTS description, a substructure matching query is performed on the atoms of the molecule. Atoms that match the

SMARTS description of the GAV are flagged as ‘visited’ and the contribution to the global thermochemical properties of the molecule is stored. This process is repeated for every entry in the GAV database until all non-hydrogen atoms have been assigned a GAV.

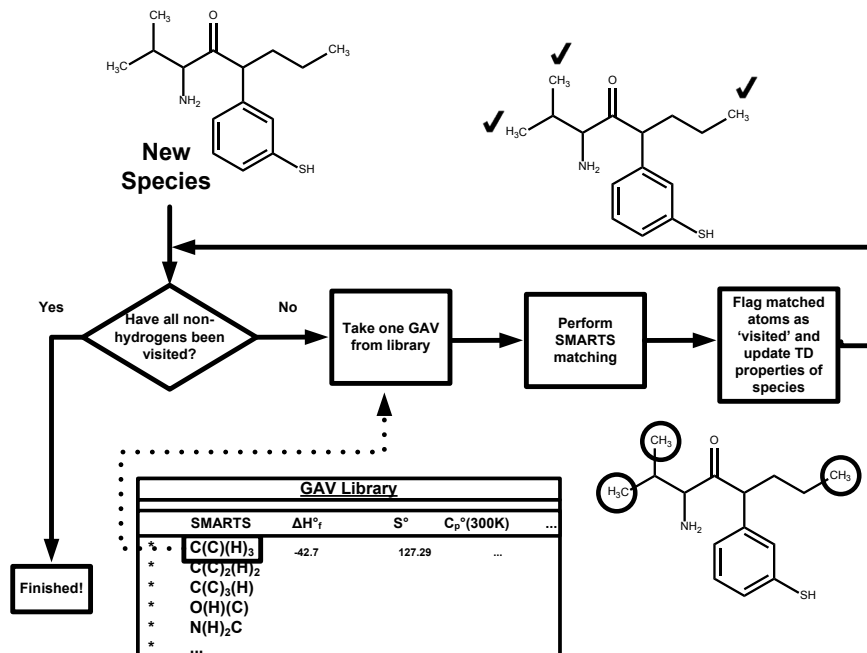


Figure 5-8: Algorithm for the assignment of Benson group additive values to polyvalent atoms in a candidate species.

For the identification of rings in a molecule, a Smallest Set of Smallest Rings (SSSR)-algorithm implementation in CDK based on an algorithm by Berger et al. [45] is used to identify the smallest set of smallest rings. Then for each ring subunit, the RSC database is iterated over until a match is found between the RSC SMARTS pattern and the ring atoms. For the HBI method, a hierarchy in the HBI databases should be respected, since different HBIs can correspond to the same radical site. For example, as pointed out by Sabbe et al. [42], the carbon with the unpaired electron in the 1,3-pentadien-1-yl radical matches both the $C=CC=CJ$ and $C=CJ$ HBI. However, since the former HBI incorporates the conjugative effects of the two double bonds, this HBI should be used rather than the simple vinylic $C=CJ$ HBI. The most specific HBIs are to be found at the top of the database since the current implementation in Genesys iteratively checks every element in the HBI database from top to bottom.

The group additivity concept is not able to account for molecular, non-local properties such as molecular symmetry σ_{glob} and the number energetically equivalent optical isomers n_{opt} and thus yields a symmetry-independent entropy term S_{int} . The total entropy S of the molecule is calculated as follows (Eq. 5-1):

$$S = S_{int} - R \ln \sigma_{glob} \quad \text{Eq. 5-1}$$

Genesys employs a generally applicable algorithm based on the concept of graph automorphisms [46] and stereoisomer enumeration [47] to calculate the total symmetry number σ_{glob} and the number of optical isomers n_{opt} without the need for explicit three-dimensional coordinates of the molecule. A detailed discussion of the algorithm for the calculation of σ_{glob} and n_{opt} will be given in Chapter 6.

5.3.2.3 Arrhenius parameter assignment

An Arrhenius equation is used to model the temperature dependency of the reaction rate coefficient of each elementary reaction, giving rise to a pre-exponential factor A , and an activation energy E_a . Assigning Arrhenius parameters to elementary reactions is the final step before a kinetic model is obtained that can be used in reactor simulations and produce quantitative results. For large detailed models containing hundreds of species and thousands of elementary reactions, this proves problematic. Experimental data on elementary reaction kinetics is often missing. Moreover, measuring the kinetics of specific reactions requires probing the concentration of short-lived intermediate species which often demands advanced analytical techniques. Calculating chemically accurate Arrhenius parameters using quantum-chemical methods is still not feasible on a large scale and suffers from the bad scaling characteristics. E.g. CPU time for the coupled cluster methods with single and double and perturbative triple excitations (CCSD(T) [48]) scale to the seventh power with respect to the number of basis functions and thus present an inescapable burden for the computation of large molecules.

In order to circumvent the lack of kinetic data for the majority of the reactions in the reaction network, engineering approximations were developed based on a limited number of high-quality data and applied to a larger set of related reactions. E.g. the Evans-Polanyi principle [49] and the

more recent Blowers and Masel correlation [50] demonstrated the usefulness of thermochemical data of reactants and products for the calculation of the reaction enthalpy $\Delta_r H_i^0$, in the estimation of reaction barriers. Others extended Benson's group additivity concept to transition states. E.g. for H-abstraction reactions several categories can be envisioned based on the nature (i.e. primary, secondary, tertiary) of the abstracting radical and of the hydrogen abstracted. Originally, values for the rate rules were obtained via regression to experimental data [51-53]. With the advent of affordable and accurate *ab initio* methods, direct quantitative information on the transition states could be incorporated into the group additive values and lead to many variations of this methodology [54-60].

In Genesys several of the existing methodologies to calculate Arrhenius parameters for a large number of reactions belonging to the same reaction family are included.

- the Evans-Polanyi correlation [49], cf. Eq. 5-2:

$$E_a^i = E_0 + \alpha \cdot \Delta_r H_i^0 \quad \text{Eq. 5-2}$$

with E_a^i the activation energy of reaction i, and $\Delta_r H_i^0$ the reaction enthalpy. E_0 and α are the intrinsic barrier and transfer coefficient respectively and need to be provided as input.

- the Blowers & Masel correlation [50], cf. Eq. 5-3:

$$E_a^i = \left(\frac{w_b + w_f + \Delta_r H_i^0}{2} \right) \frac{(V_p - (w_b + w_f) + \Delta_r H_i^0)^2}{V_p^2 - (w_b + w_f)^2 + (\Delta_r H_i^0)^2} \quad \text{Eq. 5-3}$$

With w_b and w_f the bond dissociation energies of the bonds being broken and formed during the reaction. Estimates for $(w_b + w_f)$ and V_p need to be provided per reaction family.

A group additivity method for the estimation of Arrhenius parameters is provided and is based on the methodology proposed by Saeys et al. [54, 60]. In this method the rate coefficient is expressed as, cf. Eq. 5-4:

$$k = \kappa \cdot n_e \cdot \tilde{k}_{GA} = \kappa n_e \tilde{A} \exp\left(\frac{-E_a}{RT}\right) \quad \text{Eq. 5-4}$$

With κ the tunneling coefficient, n_e accounting for the reaction path degeneracy and \tilde{k}_{GA} the rate coefficient estimated via group additivity.

The Arrhenius parameters E_a , A for a given reaction can be written as perturbations $\Delta GAV_{Ea}^0(C_i)$, $\Delta GAV_{\tilde{A}}^0$ to the single event Arrhenius parameters, i.e. excluding the number of single events n_e , of a reference reaction $E_{a,ref}$, \tilde{A}_{ref} that relate to the structural differences between the transition states in both reactions. Only primary contributions, i.e. group additive values centered around atoms that change in atom type during the reaction, are taken into account, yielding Eq. 5-5 and Eq. 5-6:

$$E_a = E_{a,ref} + \sum_i \Delta GAV_{Ea}^0(C_i) \quad \text{Eq. 5-5}$$

$$\log_{10} A = \log_{10} \tilde{A}_{ref} + \sum_i \Delta GAV_{\tilde{A}}^0(C_i) + \log_{10} n_e \quad \text{Eq. 5-6}$$

Databases containing SMARTS descriptions pointing to the corresponding ΔGAV° s are provided for the polyvalent central atoms C_i of the transition state. Next, a substructure search is initiated onto each of the reactive centers until a match is found between the pattern defined by the SMARTS string and the atom of the reactive moiety.

Similar to the Benson method for the estimation of thermochemistry, this additivity for Arrhenius parameters is generally applicable regardless of the type of reaction family or chemical nature of the reactants or products. For example, Sabbe et al. [60] provided ΔGAV° s for the hydrogen abstraction from a carbon by a carbon radical.



This reaction family requires an abstracting reactant containing a trivalent carbon radical atom and a reactant containing a carbon-hydrogen bond from which a hydrogen atom will be abstracted. For each of the reactive centers in an elementary reaction, e.g. the abstraction of hydrogen from 2-methylpropane by ethyl forming ethane and t-butyl, the database with ΔGAV° s will be iterated over until a match is found between the SMARTS string in the database and the atom in the molecule.

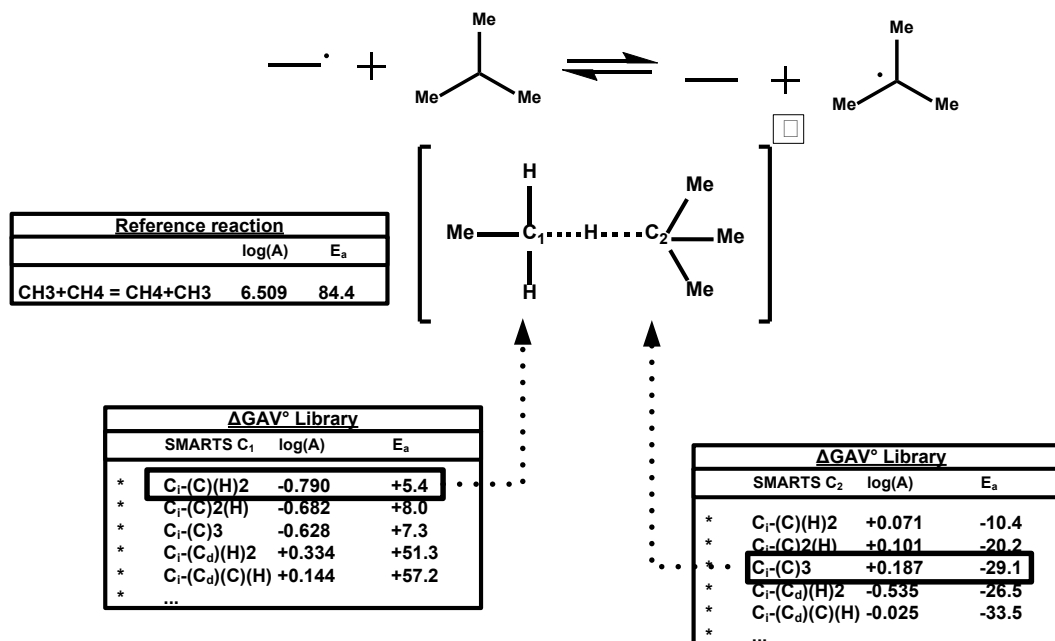


Figure 5-9: Assignment of ΔGAV° s for the H-abstraction reaction by ethyl from the tertiary hydrogen atom of 2-methylpropane forming ethane and t-butyl. Databases for both reactive centers (C_1 & C_2) are matched against the atoms in the transition state. The reference reaction for this reaction family is $\bullet CH_3 + CH_4 \leftrightarrow CH_4 + \bullet CH_3$.

For this reaction, a match for the carbon radical atom of ethyl will be found for the database entry $C_1-(C)(H)2$, and for the tertiary carbon of 2-methylpropane, a match will be found with the database entry $C_2-(C)3$, cf. Figure 5-9. Finally, Arrhenius parameters will be constructed based on the matched ΔGAV° s in the database and the Arrhenius parameters for the reference reaction, i.e. the abstraction by methyl radical from methane, with rate coefficient $k =$

$3.23 \cdot 10^6 e^{\frac{-84.4 \text{ kJ mol}^{-1}}{RT}} \text{ m}^3 \text{ mol}^{-1} \text{ s}^{-1}$. Thus, the resulting Arrhenius parameters are:

$$E_a = 84.4 + 8.0 - 29.1 = 63.3 \text{ kJ mol}^{-1}$$

$$\log_{10} A = 6.509 - 0.682 + 0.187 + \log_{10}(1) = 6.014 \rightarrow A = 1.03 \cdot 10^6 \text{ m}^3 \text{ mol}^{-1} \text{ s}^{-1}$$

The factor accounting for reaction path degeneracy n_e (also known as statistical factors, number of single events) accounts for the number of distinguishable but energetically equivalent reaction pathways from reactants to products. In Genesys, n_e is calculated as a function of the symmetry numbers of reactants σ_j and transition state σ^\ddagger [61, 62], cf. Eq. 5-7:

$$n_e = \frac{\prod_j \sigma_j}{\sigma^\ddagger} \quad \text{Eq. 5-7}$$

The calculation of the symmetry number of the transition state uses the same graph algorithm as for molecules by treating the transition state graph in the same way as the reactant graphs. For example: the addition of the hydrogen radical to trans-2-butene yielding 2-butyl has two transition-state bonds: the C-H transition bond between the added hydrogen radical and transition bond between the two carbon atoms that were previously connected by a double bond, cf. Figure 5-10.

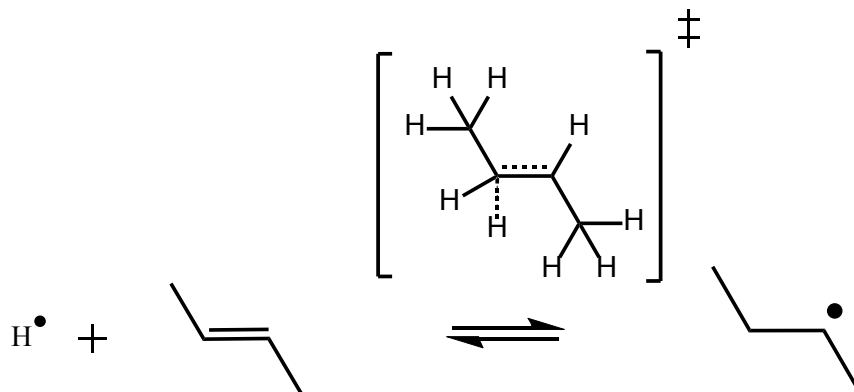


Figure 5-10: Graph representation of the transition state of the addition of a hydrogen radical to trans-2-butene with the two “transition” bonds colored in red.

The tunneling coefficient κ is a factor superimposed onto the group additivity rate coefficient and cannot be calculated based on structural contributions of the transition state. Since the calculation of tunneling coefficients based on tunneling potentials require information (imaginary frequencies etc.) which is not available without *ab initio* calculations, an empirical method is used to obtain an estimation of κ as a function of temperature and the net electronic barrier. E.g. Sabbe et al. [60] described κ for hydrocarbon H-abstraction reactions as a fourth-order polynomial, cf. Eq. 5-8:

$$\kappa = 1 + \left(\frac{162}{T} \right)^3 (E_{a,exo}) + 2.71 \cdot 10^{-6} \exp\left(-\frac{T-300}{26} \right) (E_{a,exo})^4 \quad \text{Eq. 5-8}$$

Genesys allows the user to define a correlation for the temperature dependent tunneling coefficient based on the calculated activation energy of the particular elementary reaction.

Reaction families can be defined that are the reverse direction of another reaction family, e.g. β -scission reaction family and the intermolecular addition reaction family. In case Arrhenius

parameters are assigned to both these reaction families, conflicting situations due to the thermodynamic consistency principle can arise for the Arrhenius parameters of forward and reverse reactions. Therefore, a particular reaction family can be defined as a “reverse” reaction family in the input meaning that explicit Arrhenius parameters for these reactions will not be present in the kinetic model, but will follow by the application of thermodynamic consistency.

5.3.3 Output

The calculated thermochemical properties of all molecules are stored as NASA polynomials, and the kinetic model formatted in such a way that the integration with chemical reactor simulation packages such as Chemkin [63] is possible. The generated model could be subject to postprocessing steps, such as Arrhenius parameter fitting based on experimental data, uncertainty quantification and error propagation methods, although these techniques are not present in Genesys [64-66]. Species are stored in formats such as the MDL-MOL and SDF format allowing the visualization and inspection of them through third-party tools. Elementary reactions are visualized through the JChemPaint package [67], cf. Figure 5-11.

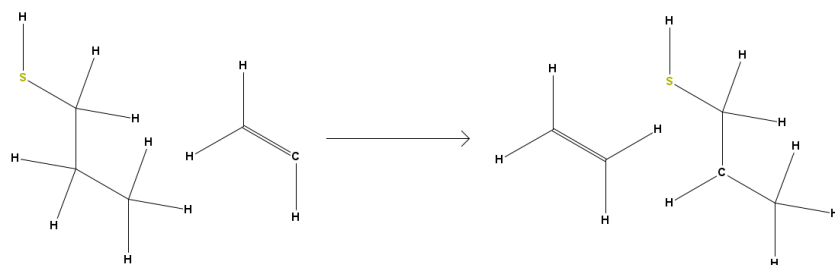


Figure 5-11: reaction visualization generated using JChemPaint integrated with CDK [67].

5.4 Case studies

5.4.1 Case 1: calculation of thermochemical data

A total number of ca. 700 unique entries for the Benson group database were collected from various literature sources [42, 68-70], primarily consisting of GAVs. Mainly carbon centered

groups are present but a limited subset groups centered on hetero-elements such as sulfur, nitrogen, oxygen, phosphorous, boron, silicon and halides. Next to the GAVs 69 HBIs, 61 RSCs, 3 RESs and 5 NNIs are present in the database. Recursive SMARTS strings, which allows to match a particular atom in an atomic environment, are used to represent both monoatomic, e.g. C-(C_d)₃(H), and polyatomic centered groups, such as CO-(C)(C) and allows to define atomic environments rather entire substructures. E.g. [$\$(C(-C)(-C)(-C)-O)$] refers to a carbon atom that is surrounded by three single bonded aliphatic carbons and one single bonded aliphatic oxygen atom, C_s-(C_s)₃(O_s). The use of logical operators in SMARTS expressions allows introducing approximations. E.g. Sabbe et al. [42] reduced the number of GAVs to be estimated for the prediction of standard enthalpies of formation for radical species from 125 groups to 41 by approximating radical-adjacent GAVs by the value of the corresponding hydrocarbon group. The SMARTS expression for a C-(C)₃(H) GAV with both radical and non-radical carbon ligands is as follows: [$\$(C(-[C;v3,v4])(-[C;v3,v4])(-[C;v3,v4])-H)$] where [C;v3,v4] denotes that the valency of the carbon ligand is either 3 or 4.

In order to assess the performance of the current Benson group additivity database, the Third Millennium Ideal Gas and Condensed Phase Thermochemical Database for Combustion [71] database was used for comparison. From this database, species with attached CAS numbers and standard enthalpy of formation at 298K and the standard entropy at 298K are retained. CAS numbers are converted into InChI's using NIH's Chemical Identifier Resolver beta 4 [72], and a set of 574 species out of 1310 entries is retained which can be used for benchmarking purposes. Genesys was able to estimate the enthalpy of formation and the entropy at 298K for 266 species, constituted from a total 9 chemical elements (H, C, O, N, I, Cl, Br, F, S). On a single linux core containing 8 CPUs the simulation took 440 seconds in total. Without the pre-screening step it took ca. 163000 seconds on the same machine, showing the advantage of using preliminary fingerprinting. The ability to determine only a small fraction of the benchmark database of species demonstrates the limitations of the current database due to the many missing GAVs.

Nevertheless, Figure 5-12 shows that a wide range of species can be determined, from halogenated hydrocarbons with $\Delta_f H^\circ(298\text{K})$ round $-3500 \text{ kJ mol}^{-1}$ to ethyne hydrocarbon radicals with $\Delta_f H^\circ(298\text{K})$ round 1000 kJ mol^{-1} . The largest absolute deviations between the Benson group additivity estimate and the benchmark database exist for species containing rings, for which adequate RSCs are missing in the Genesys databases and lead to underestimated heats of formation. The higher number of deviations of the entropy is attributed to a smaller number of GAV database entries for entropies than for contributions to the enthalpy of formation. Also, some outliers were attributed to errors in the benchmark database rather than in Genesys' database, and show the capabilities of the current tool for external database validation.

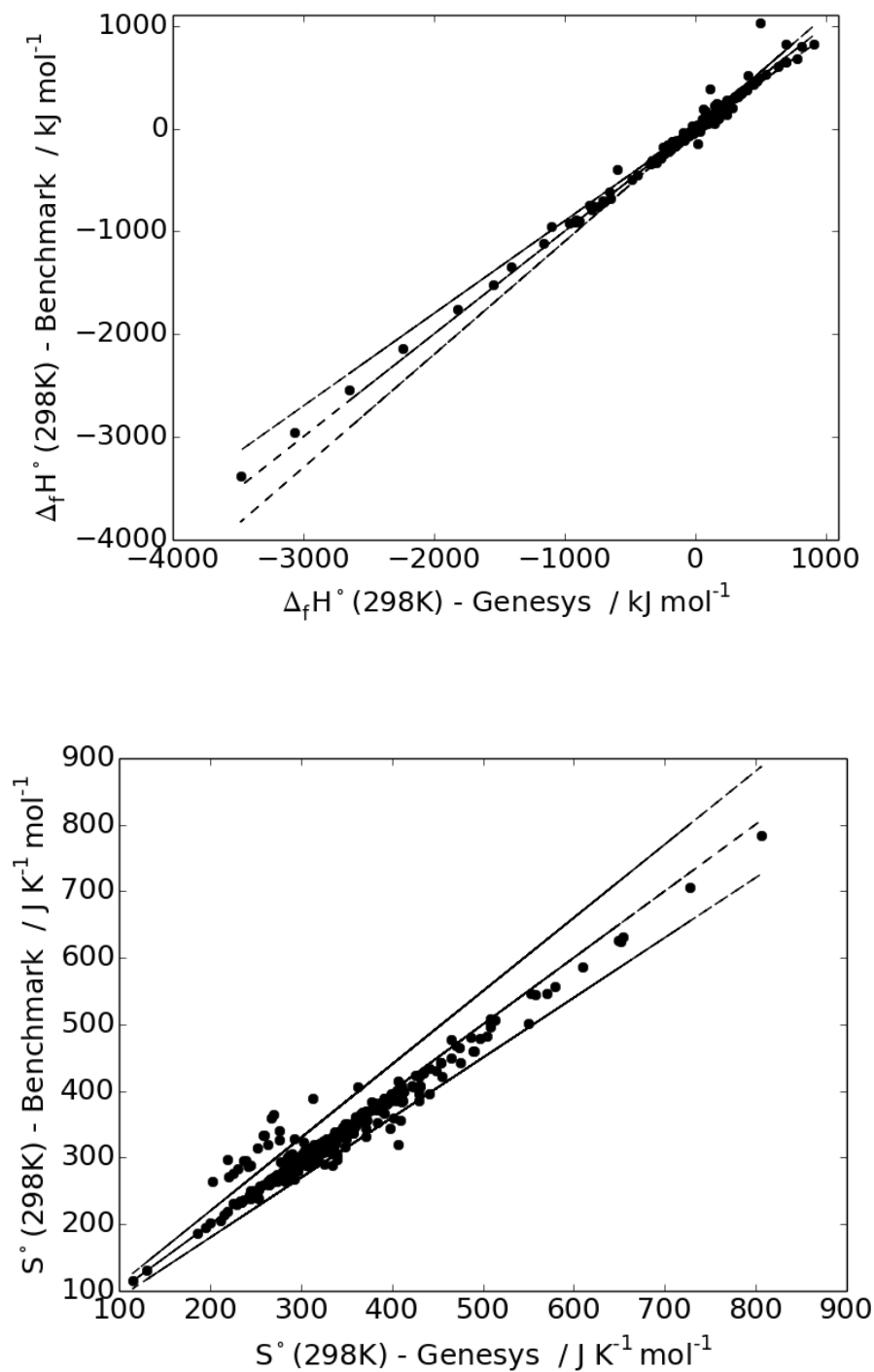
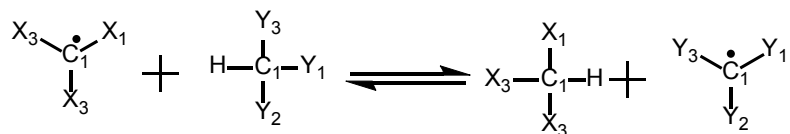


Figure 5-12: Parity plot for: Genesys estimated standard enthalpy of formation at 298K (upper) and standard entropy at 298K (lower) against the benchmark database, $\pm 10\%$ relative deviation is indicated.

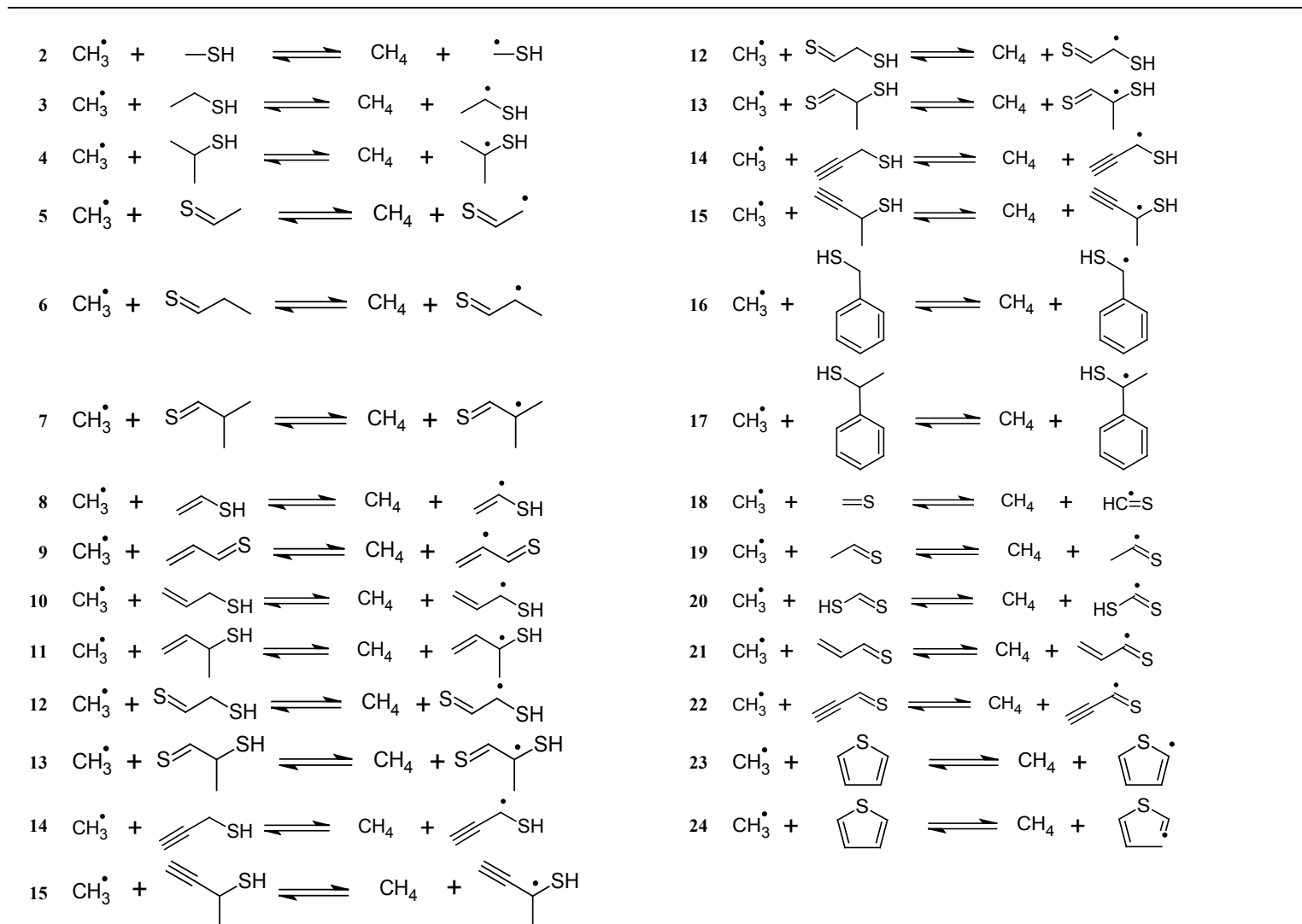
5.4.2 Case 2: Arrhenius parameters via group additivity

As to showcase Genesys' capabilities in applying the group additivity methodology, ΔGAV° s at 1000K calculated by Sabbe et al. [60] and Vandeputte et al. [73] were used to populate a database for α -hydrogen abstraction reactions by carbon centered radicals from organosulfur components.



They comprise of a total of 44 ΔGAV° s for carbon radicals. A test set of 24 H-abstraction reactions by methyl from organosulfur components is taken for which both the forward and reverse rate coefficient is calculated through group additivity.

Table 5-1: 24 α -hydrogen abstraction reactions (from ref.[73]) by methyl from organosulfur components automatically calculated with the group additivity implementation for Arrhenius parameters in Genesys.



The ratio of the forward and reverse rate coefficients k_f/k_r is compared to the equilibrium coefficient K_{eq} of the reaction. The latter is calculated based on the thermochemistry of reactants and products, using Genesys.

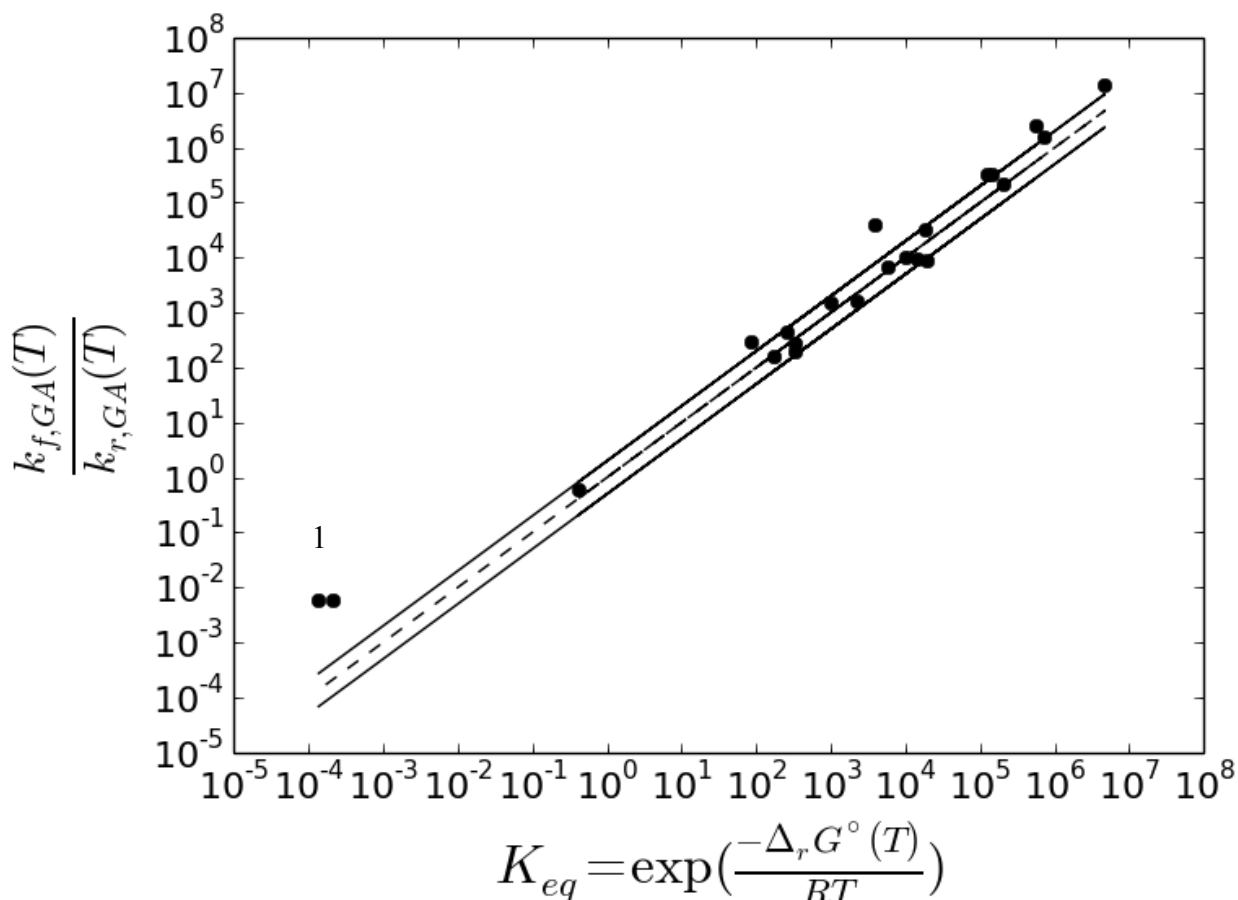


Figure 5-13: The ratio of forward and reverse rate coefficient of 22 α -H-abstraction reactions by methyl on an organosulfur component as a function of the equilibrium coefficient of the reaction calculated through the thermochemistry of reactants and products. $\pm 10\%$ relative deviation is indicated.

As can be observed from Figure 5-13, excellent agreement is found between both methods to calculate the reaction equilibrium coefficient. Most deviations remain within a factor 2 and can be attributed to the accuracy of the ΔGAV° s, HBIs, and GAVs. Significant deviations exist for reactions 23 and 24, with K_{eq} close to 10^{-4} and involve thiophene. The error is attributed to the

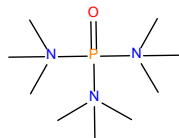
thermochemistry estimates for the thiophenyl radicals because of an incorrect assignment of a GAV to the radical carbon atom. CDK treats the thiophenyl radical as aromatic and therefore treats all the atoms and bonds in the thiophenyl radical as aromatic as well. Hence, a Benson group was assigned the radical atom that corresponds to a phenylic carbon radical atom, instead of the radical carbon in a thiophenic ring. Although this error is easily addressed by limiting the use of the phenylic radical GAV to only benzene-type components and not to any aromatic carbon (e.g. thiophene, etc...) it exemplifies the accuracy and precision that is required when defining the sub-molecular pattern associated to a given Benson group.

Currently, the database of ΔGAV° s has been populated with the following data:

- 96 ΔGAV° s for H-abstraction reactions by a carbon centered radical to a carbon centered group. 50 ΔGAV° s were reported for pure hydrocarbon abstraction reactions [60] and 44 ΔGAV° s for α -hydrogen abstraction reactions onto organosulfur components [73].
- 47 ΔGAV° s for carbon-centered radicals [74] and 34 ΔGAV° s for hydrogen radical addition [59] reactions to unsaturated hydrocarbons and the reverse β -scission reactions.
- 25 ΔGAV° s for the H-abstraction reactions by a hydrogen radical from hydrocarbons. [75]
- 53 ΔGAV° s for H-abstractions to a sulfur centered group by a carbon centered radical. [76]

5.4.3 Case 3: reaction network generation: thermal decomposition of hexamethyl phosphoric triamide

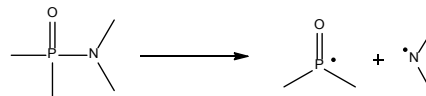
To illustrate the advantage of using a chemoinformatics library with support of many chemical elements, oxidation states and atom hybridizations, a second example is included in which a reaction network is generated for the thermal decomposition of hexamethyl phosphoric triamide (HMPA), a coke inhibiting additive used in steam cracking applications.



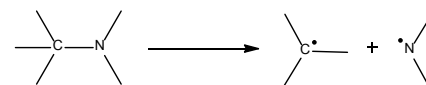
A limited number of reaction families are used as shown in Table 5-2, which could be considered relevant in the kinetic modeling of the decomposition of HMPA.

Table 5-2: Reaction families used in modeling the thermal decomposition of hexamethyl phosphoric triamide

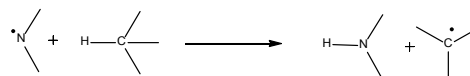
Homolytic P-N scission with phosphorous in P=O bond and P connected to 4 neighbours



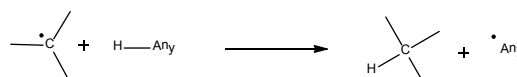
Homolytic N-C scission with sp³ Carbon



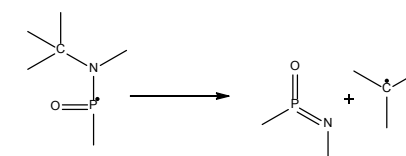
H abstractions to a carbon centered group by a nitrogen centered radical



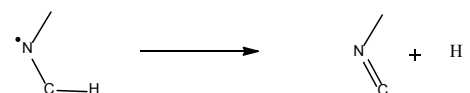
H-abstraction reactions by carbon centered radicals



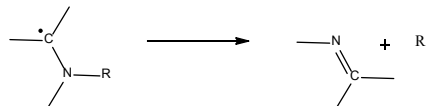
Carbon-Centered β -Scission onto a phosphorus-centered radical next to a nitrogen



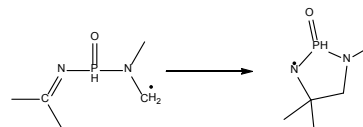
Hydrogen-Centered β -Scission onto a nitrogen-centered radical.



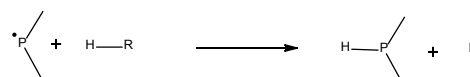
Beta-Scission onto a carbon-centered radical next to a nitrogen with radical of generic element



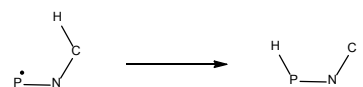
Endo-intramolecular addition to five-membered ring, with a carbon radical adding onto a C=N bond next to a -P=O fragment, with a nitrogen radical as a result



H abstractions to a carbon centered group by a bivalent phosphorus centered radical with two neighbours



Intramolecular 1,3-H shift between P-radical and γ -hydrogen



Intramolecular 1,4-H shift between carbon-hydrogen and double bonded oxygen to tetravalent P radical

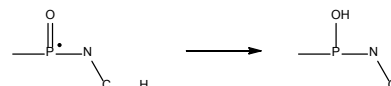
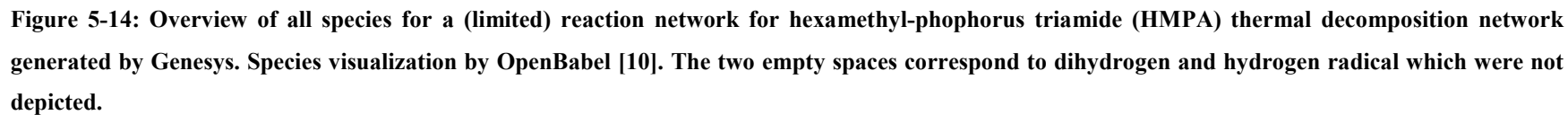


Figure 5-14 gives an overview of the product species created in the reaction network formed by the limited number of implemented reaction families.



The generated decomposition products of HMPA are again a straightforward result of the defined reaction families and the constraints applied on them. For example, among the small decomposition products ammonia, amine-type components, and methyl are found, while small oxygen containing radicals such as hydroxyl, and phosphorus-containing fragments such as PH_3 , and $\dot{\text{P}}\text{H}_2$ are missing. The reason for this is the absence of a reaction family that ruptures the C-O bond.

5.5 Conclusions

A new automatic reaction network generation tool, called Genesys, was presented for the purpose of constructing and applying kinetic models in reaction engineering. Through the use of open source chemoinformatics framework (the Chemistry Development Kit) for the internal representation of chemical objects Genesys was created with relative little effort (total number of lines of code < 20000) and generally applicable regardless of the involved chemical elements or chemistry.

Genesys is structured in a sequential way with the generation of the reaction network decoupled from the assignment of quantitative information to species and reactions. The reaction network is expanded by iterating over a number of user-defined reaction families that verify the eligibility of a candidate species to undergo that reaction family and that transform the reactants into products by executing a set of transformations onto designated reactive center atoms. The identification of key sub-molecular patterns required for a particular reaction family was accomplished by the use of the unambiguous pattern language called SMARTS. The applicability of the reaction families is constrained as much as possible by the use of atom and molecule constraints and reflect the significance of a particular reaction in the model.

SMARTS also enabled the use of quantitative structure property relationships through group contribution methods. It was shown that ideal gas phase thermochemical properties can be estimated for a wide range of species and chemical elements by implementing Benson's group additivity scheme. Discrepancies between the estimations and the validation test set were attributed to the still missing entries in the databases of group additive values and supplementary

corrections and exemplify the ever growing demand for chemical data for the construction of kinetic models. Arrhenius parameters for elementary reactions were estimated through a priori rate constant estimation techniques such as group additivity methods for the estimation of Arrhenius parameters.

Considerable efficiency improvements in the kinetic model generation were accomplished by adopting various techniques. Substructure searches were optimized using the AMBIT-SMARTS library which combines a series of pre-screening steps to filter out negative SMARTS query matches before the actual rigorous substructure search algorithm. Reaction isomorphism was accomplished by transition state graph comparison and preceded by a fingerprint matching step, in a similar way to the substructure query technique. Species comparison was performed through cached InChI comparison.

By externalizing all chemical data from the actual network generation code, the user has full control and responsibility over the creation of product species and reactions. The resulting kinetic models are a straight-forward result of the chemical knowledge of the user and are reflected by the user-defined set of reaction families controlling reaction network size and generation time.

5.6 References

- [1] S. Pierucci, E. Ranzi, A review of features in current automatic generation software for hydrocarbon oxidation mechanisms, *Computers & Chemical Engineering*, 32 (2008) 805-826.
- [2] G. Moore, Cramming more components onto integrated circuits, *Electronics*, 38 (1965).
- [3] T.I. Oprea, A. Tropsha, Target, chemical and bioactivity databases – integration is key, *Drug Discovery Today: Technologies*, 3 (2006) 357-365.
- [4] J. Gasteiger, *Handbook of Chemoinformatics*, Wiley-VCH Verlag GmbH & Co., Weinheim, 2003.
- [5] A.S. Reddy, S.P. Pati, P.P. Kumar, H.N. Pradeep, G.N. Sastry, Virtual screening in drug discovery - A computational perspective, *Current Protein & Peptide Science*, 8 (2007) 329-351.
- [6] J. Van Drie, Pharmacophore Discovery, in: *Computational Medicinal Chemistry for Drug Discovery*, CRC Press, 2003.
- [7] R. Hoof, C. Spronk, S. Nabuurs, E. Krieger, G. Vriend, Protein Structures, in: *Computational Medicinal Chemistry for Drug Discovery*, CRC Press, 2003.
- [8] C. Steinbeck, Y. Han, S. Kuhn, O. Horlacher, E. Luttmann, E. Willighagen, The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics, *J Chem Inf Comput Sci*, 43 (2003) 493 - 500.
- [9] C. Steinbeck, C. Hoppe, S. Kuhn, M. Floris, R. Guha, E.L. Willighagen, Recent developments of the Chemistry Development Kit (CDK) - An open-source Java library for chemo- and bioinformatics, *Current Pharmaceutical Design*, 12 (2006) 2111-2120.
- [10] N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, Open Babel: An open chemical toolbox, *Journal of Cheminformatics*, 3 (2011).
- [11] G. Landrum, RDKit: Open-source cheminformatics; <http://rdkit.sourceforge.net>, (2006).

- [12] V. Warth, F. Battin-Leclerc, R. Fournet, P.A. Glaude, G.M. Côme, G. Scacchi, Computer based generation of reaction mechanisms for gas-phase oxidation, *Computers & Chemistry*, 24 (2000) 541-560.
- [13] L.J. Broadbelt, S.M. Stark, M.T. Klein, Computer-generated pyrolysis modeling on-the-fly generation of species, reactions, and rates, *Industrial & Engineering Chemistry Research*, 33 (1994) 790-799.
- [14] P.J. Clymans, G.F. Froment, Computer-generation of reaction paths and rate-equations in the thermal cracking of normal and branched paraffins, *Computers & Chemical Engineering*, 8 (1984) 137-142.
- [15] M. Dente, E. Ranzi, A.G. Goossens, Detailed prediction of olefin yields from hydrocarbon pyrolysis through a fundamental simulation-model (SPYRO), *Computers & Chemical Engineering*, 3 (1979) 61-75.
- [16] F. Battin-Leclerc, E. Blurock, R. Bounaceur, R. Fournet, P.-A. Glaude, O. Herbinet, B. Sirjean, V. Warth, Towards cleaner combustion engines through groundbreaking detailed chemical kinetic models, *Chemical Society Reviews*, 40 (2011) 4762-4782.
- [17] S. Wauters, G.B. Marin, Computer generation of a network of elementary steps for coke formation during the thermal cracking of hydrocarbons, *Chemical Engineering Journal*, 82 (2001) 267-279.
- [18] W.H. Green, J.W. Allen, R.W. Ashcraft, G.J. Beran, C.A. Class, C. Gao, C.F. Goldsmith, M.R. Harper, A. Janan, G.R. Magoon, D.M. Matheu, S.S. Merchant, J.D. Mo, S. Petway, S. Raman, S. Sharma, J. Song, K.M. Van Geem, J. Wen, R.H. West, A. Wong, H.S. Wong, P.E. Yelvington, J. Yu, RMG - Reaction Mechanism Generator v4.0, (2013).
- [19] E.R. Ritter, J.W. Bozzelli, THERM - Thermodynamic property estimation for gas-phase radicals and molecules, *International Journal of Chemical Kinetics*, 23 (1991) 767-778.
- [20] C. Muller, V. Michel, G. Scacchi, G.M. Côme, 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*, 92 (1995) 1154-1178.
- [21] B.K. Harrison, CHETAH, a program for convenient thermochemistry applications, *Abstracts of Papers of the American Chemical Society*, 212 (1996) 174-COMP.
- [22] E.S. Blurock, F. Battin-Leclerc, T. Faravelli, W.H. Green, United mechanism generator framework for detailed chemical kinetic mechanism models, *COST Network - 1st Annual Meeting*, (2010).
- [23] E.S. Blurock, Reaction - System for modeling chemical reactions, *Journal of Chemical Information and Computer Sciences*, 35 (1995).
- [24] V. Warth, N. Stef, P.A. Glaude, F. Battin-Leclerc, G. Scacchi, G.M. Côme, Computer-aided derivation of gas-phase oxidation mechanisms: Application to the modeling of the oxidation of n-butane, *Combustion and Flame*, 114 (1998) 81-102.
- [25] S.R. Heller, S.E. Stein, D.V. Tchekhovskoi, InChI: Open access/open source and the IUPAC international chemical identifier, *Abstracts of Papers of the American Chemical Society*, 230 (2005) 60-CINF.
- [26] S. Heller, S. Stein, IUPAC InChI, Google Tech Talk, (2006).
- [27] D. Weininger, SMILES, a chemical language and information-system. 1. Introduction to methodology and encoding rules, *Journal of Chemical Information and Computer Sciences*, 28 (1988) 31-36.
- [28] D. Weininger, A. Weininger, J.L. Weininger, SMILES .2. Algorithm for generation of unique SMILES notation, *Journal of Chemical Information and Computer Sciences*, 29 (1989) 97-101.
- [29] R. Guha, M.T. Howard, G.R. Hutchison, P. Murray-Rust, H. Rzepa, C. Steinbeck, J.r. Wegner, E.L. Willighagen, The Blue Obelisk Interoperability in Chemical Informatics, *Journal of Chemical Information and Modeling*, 46 (2006) 991-998.
- [30] Daylight Chemical Information Systems, Daylight Theory Manual, 2010 <http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>.
- [31] A. Ratkiewicz, T.N. Truong, Application of chemical graph theory for automated mechanism generation, *Journal of Chemical Information and Computer Sciences*, 43 (2003) 36-44.
- [32] S. Rangarajan, A. Bhan, P. Daoutidis, Rule-Based Generation of Thermochemical Routes to Biomass Conversion, *Industrial & Engineering Chemistry Research*, 49 (2010) 10459-10470.
- [33] J. Song, S. Raman, J. Yu, C.D. Wijaya, G. Stephanopoulos, W.H. Green, Development of automatic chemical reaction mechanism generation software using object-oriented technology., *Abstracts of Papers of the American Chemical Society*, 226 (2003) U530-U531.
- [34] M.R. Garey, D.S. Johnson, *Computers and intractability*, Freeman New York, 1979.
- [35] C. Tonnelier, P. Jauffret, T. Hanser, G. Kaufmann, Machine learning of generic reactions: 3. an efficient algorithm for maximal common substructure determination, *Tetrahedron Computer Methodology*, 3 (1990) 351-358.
- [36] N. Jeliaskova, N. Kochev, AMBIT-SMARTS: Efficient Searching of Chemical Structures and Fragments, *Molecular Informatics*, 30 (2011) 707-720.

- [37] L.J. Broadbelt, S.M. Stark, M.T. Klein, Termination of computer-generated reaction-mechanisms - species rank - based convergence criterion, *Industrial & Engineering Chemistry Research*, 34 (1995) 2566-2573.
- [38] R.G. Susnow, A.M. Dean, W.H. Green, P. Peczak, L.J. Broadbelt, Rate-based construction of kinetic models for complex systems, *Journal of Physical Chemistry A*, 101 (1997) 3731-3740.
- [39] E. Ranzi, M. Dente, S. Plerucci, G. Biardi, Initial product distributions from pyrolysis of normal and branched paraffins, *Industrial & Engineering Chemistry Fundamentals*, 22 (1983) 132-139.
- [40] K.J. Laidler, *Chemical Kinetics*, 3rd ed., Harper & Row, New York, 1987.
- [41] J.L. Durant, B.A. Leland, D.R. Henry, J.G. Nourse, Reoptimization of MDL Keys for Use in Drug Discovery, *Journal of Chemical Information and Computer Sciences*, 42 (2002) 1273-1280.
- [42] M.K. Sabbe, M. Saeys, M.F. Reyniers, G.B. Marin, V. Van Speybroeck, M. Waroquier, Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals, *Journal of Physical Chemistry A*, 109 (2005) 7466-7480.
- [43] S.W. Benson, *Thermochemical Kinetics*, John Wiley & Sons, New York, 1976.
- [44] T.H. Lay, J.W. Bozzelli, A.M. Dean, E.R. Ritter, Hydrogen-atom bond increments for calculation of thermodynamic properties of hydrocarbon radical species, *Journal of Physical Chemistry*, 99 (1995) 14514-14527.
- [45] F. Berger, P. Gritzmann, S. de Vries, Minimum cycle bases for network graphs, *Algorithmica*, 40 (2004) 51-62.
- [46] L.D. Kreher, R.D. Stinson, *Combinatorial algorithms: generation, enumeration, and search*, CRC press LTC, Boca Raton, Florida, USA, 1998.
- [47] M. Razinger, K. Balasubramanian, M. Perdih, M.E. Munk, Stereoisomer generation in computer-enhanced structure elucidation, *Journal of Chemical Information and Computer Sciences*, 33 (1993) 812-825.
- [48] K. Raghavachari, G.W. Trucks, J.A. Pople, M. Headgordon, A 5-th order perturbation comparison of electron correlation theories, *Chem. Phys. Lett.*, 157 (1989) 479-483.
- [49] M.G. Evans, M. Polanyi, Inertia and driving force of chemical reactions, *Transactions of the Faraday Society*, 34 (1938) 0011-0023.
- [50] P. Blowers, R. Masel, Engineering approximations for activation energies in hydrogen transfer reactions, *AIChE Journal*, 46 (2000) 2041-2052.
- [51] P.A. Willems, G.F. Froment, Kinetic modeling of thermal-cracking of hydrocarbons. 1. Calculation of frequency factors, *Industrial & Engineering Chemistry Research*, 27 (1988) 1959-1966.
- [52] P.A. Willems, G.F. Froment, Kinetic modeling of thermal-cracking of hydrocarbons. 2. Calculation of activation-energies, *Industrial & Engineering Chemistry Research*, 27 (1988) 1966-1971.
- [53] E. Ranzi, M. Dente, T. Faravelli, G. Pennati, Prediction of kinetic-parameters for hydrogen abstraction reactions, *Combustion Science and Technology*, 95 (1994) 1-50.
- [54] M. Saeys, M.F. Reyniers, G.B. Marin, V. Van Speybroeck, M. Waroquier, Ab initio group contribution method for activation energies for radical additions, *AIChE Journal*, 50 (2004) 426-444.
- [55] R. Sumathi, H.H. Carstensen, W.H. Green, Reaction rate predictions via group additivity. Part 3: Effect of substituents with CH₂ as the mediator, *Journal of Physical Chemistry A*, 106 (2002) 5474-5489.
- [56] R. Sumathi, H.H. Carstensen, W.H. Green, 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*, 105 (2001) 8969-8984.
- [57] R. Sumathi, H.H. Carstensen, W.H. Green, Reaction rate prediction via group additivity Part 1: H abstraction from alkanes by H and CH₃, *Journal of Physical Chemistry A*, 105 (2001) 6910-6925.
- [58] T.N. Truong, Reaction class transition state theory: Hydrogen abstraction reactions by hydrogen atoms as test cases, *J. Chem. Phys.*, 113 (2000) 4957-4964.
- [59] M.K. Sabbe, M.-F. Reyniers, M. Waroquier, G.B. Marin, Hydrogen Radical Additions to Unsaturated Hydrocarbons and the Reverse beta-Scission Reactions: Modeling of Activation Energies and Pre-Exponential Factors, *Chemphyschem*, 11 (2010) 195-210.
- [60] M.K. Sabbe, A.G. Vandeputte, M.-F. Reyniers, M. Waroquier, G.B. Marin, Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions, *Phys. Chem. Chem. Phys.*, 12 (2010) 1278-1298.
- [61] E. Pollak, P. Pechukas, Symmetry numbers, not statistical factors, should be used in absolute rate theory and in bronsted relations, *J. Am. Chem. Soc.*, 100 (1978) 2984-2991.
- [62] D.R. Coulson, Statistical factors in reaction-rate theories, *J. Am. Chem. Soc.*, 100 (1978) 2992-2996.
- [63] R.J. Kee, F.M. Rupley, J.A. Miller, M.E. Coltrin, J.F. Grcar, E. Meeks, H.K. Moffat, G. Lutz, A.E. Dixon-Lewis, M.D. Smooke, J. Warnatz, G.H. Evans, R.S. Larson, R.E. Mitchell, L.R. Petzhold, W.C. Reynolds, M. Caracotsios, W.E. Stewart, P. Glarborg, C. Wang, O. Adigun, W.G. Houf, C.P. Chou, S.F. Miller, P. Ho, D.J. Young, *CHEMKIN Release 4.1.1*, (2007).

- [64] T. Turányi, T. Nagy, I.G. Zsély, M. Cserhádi, T. Varga, B.T. Szabó, I. Sedyó, P.T. Kiss, A. Zempléni, H.J. Curran, Determination of rate parameters based on both direct and indirect measurements, *International Journal of Chemical Kinetics*, 44 (2012) 284-302.
- [65] J.C. Pinto, M.W. Lobao, A.L. Alberton, M. Schwaab, M. Embirucu, S.V. de Melo, Critical Analysis of Kinetic Modeling Procedures, *International Journal of Chemical Reactor Engineering*, 9 (2011).
- [66] G.P. Yoga, M.G. Coelho, G.M. de Lima, J.C. Belchior, Experimental and Theoretical Studies of the Thermal Behavior of Titanium Dioxide-SnO₂ Based Composites, *Journal of Physical Chemistry A*, 115 (2011) 2719-2726.
- [67] S. Krause, E. Willighagen, C. Steinbeck, JChemPaint - Using the collaborative forces of the Internet to develop a free editor for 2D chemical structures, *Molecules*, 5 (2000) 93-98.
- [68] B.E. Poling, J.M. Prausnitz, J.P. O'Connell, *The properties of gases and liquids*, McGraw-Hill, New York, 2001.
- [69] N. Cohen, S.W. Benson, Estimation of heats of formation of organic-compounds by additivity methods, *Chemical Reviews*, 93 (1993) 2419-2438.
- [70] A.G. Vandeputte, M.K. Sabbe, M.-F. Reyniers, G.B. Marin, Modeling the Gas-Phase Thermochemistry of Organosulfur Compounds, *Chemistry-a European Journal*, 17 (2011) 7656-7673.
- [71] E. Goos, A. Burcat, B. Ruscic, Third Millennium Ideal Gas and Condensed Phase Thermochemical Database for Combustion with updates from Active Thermochemical Tables, ANL-05/20 and TAE 960 Technion-IIT, (2010).
- [72] CADD Group Chemoinformatics Tools and User Services: Chemical Identifier Resolver beta 4, 2012 <http://cactus.nci.nih.gov/>.
- [73] A.G. Vandeputte, M.K. Sabbe, M.-F. Reyniers, G.B. Marin, Kinetics of alpha hydrogen abstractions from thiols, sulfides and thiocarbonyl compounds, *Phys. Chem. Chem. Phys.*, 14 (2012) 12773-12793.
- [74] M.K. Sabbe, M.-F. Reyniers, V. Van Speybroeck, M. Waroquier, G.B. Marin, Carbon-centered radical addition and beta-scission reactions: Modeling of activation energies and pre-exponential factors, *Chemphyschem*, 9 (2008) 124-140.
- [75] M.K. Sabbe, M.-F. Reyniers, G.B. Marin, H-abstraction reactions by a hydrogen radical from hydrocarbons, Personal Communication.
- [76] A.G. Vandeputte, M.-F. Reyniers, G.B. Marin, Kinetic Modeling of Hydrogen Abstractions Involving Sulfur Radicals, *Chemphyschem*, 14 (2013) 3751-3771.

Chapter 6: Symmetry calculation for reactants and transition states

6.1 Abstract

A novel concept of label-stereoisomers, i.e. stereoisomers that arise after labeling homomorphic ligands in the original molecule so that they become distinguishable, is introduced. It allows the calculation of the rotational symmetry number of molecules and the reaction path degeneracy of elementary reactions. It uses the automorphism group order of the associated graph and corrects this by accounting for the number of label-stereoisomers. A topological representation of chemical objects was applied such that the atomic coordinates in the molecule were not required. The developed algorithm differs from previous approaches for determining topological symmetry by avoiding the use of heuristic rules that limit the generic nature of the latter. Mesomerism, isotopic labeling, electron-deficient atoms, and stereochemistry are handled. The proposed algorithm is able to accurately calculate symmetry numbers and reaction path degeneracies for a large and diverse set of structures, ranging from asymmetric molecules such as fluorochlorobromomethane to highly symmetric structures such as 2,2-dimethylpropane.

Keywords: topological symmetry; automorphism; rotational symmetry number; reaction path degeneracy

6.2 Introduction

The notion of symmetry is an essential part of molecules. Molecular symmetry refers to the indistinguishable orientations of a molecule [1]. This is macroscopically quantified as a decrease of the entropy S by a term $-R \cdot \ln(\sigma)$ [2] with R the universal gas constant and σ the global symmetry number, corresponding to the number of indistinguishable orientations of the molecule. The analogue of the symmetry number of molecules for reactions is the reaction path degeneracy [3], also known as statistical factors [4], or the number of single events n_e [5]. It refers to the number of energetically equivalent paths that reactants can follow to be converted into products [6]. Symmetry of molecules is important in many areas including chemistry, biology and physics. Symmetry numbers have been correlated with physical properties such as melting points and melting entropies [7, 8]. Identifying symmetry elements in molecules allows to assess physical properties such as the presence of a dipole moment or optical activity [9] or to interpret NMR or IR spectroscopy data [10-12]. In *in silico* drug design, knowledge of the symmetry of a drug-like molecule can be used to predict the binding affinity to a target receptor [13] or to speed up conformational searches by reducing the parameter space [14]. The reaction path degeneracy is important in analyzing the rates of reactions: an experimentally observable rate coefficient of a reaction can be interpreted as the product of the reaction path degeneracy and a symmetry-independent rate coefficient.

The symmetry of molecules and reactions can be manually determined through visualization of the three-dimensional structure of the relevant entity. This becomes cumbersome and error-prone in applications where this procedure needs to be repeated many times. For example, in chemical engineering applications such as the automated construction of detailed kinetic models of complex chemical processes [15] thermochemical properties of hundreds of species, and the rate coefficients of thousands of reactions need to be evaluated. Moreover, group contribution methods such as Benson group additivity [16] that are devised to provide fast and reliable estimates of thermochemical properties [16, 17] or rate coefficients [18-25] cannot determine non-local properties such as symmetry. It therefore always needs to be accounted for in a separate way. Many symmetry identification algorithms have borne on the representation of molecules as mathematical graphs [9, 26-29]. This connectivity-based representation greatly facilitated the

introduction of concepts and solutions from graph theory, but also circumvented the need of the availability of three dimensional (3D) atomic coordinates. If these coordinates are available, then brute force techniques that determine point groups such as in [30] and implemented in many *ab initio* packages, offer a way to determine symmetry numbers. In many cases, however, information on the atomic coordinates of a molecule is unavailable *a priori*. Automatic geometry optimization techniques are not straightforward, and inherently suffer from the trade-off between accuracy and computational cost. Moreover, computational geometry optimization methods may lead to small errors in the atomic coordinates; the subsequent identification of a symmetry element based on 3D coordinates always needs an arbitrary tolerance threshold that accounts for deviations of the transformed coordinates relative to the original ones. For supramolecular structures such as proteins, macromolecules, nanostructures where this effect can be even more pronounced, distance functions such as the continuous symmetry measure [31-33] may be more useful.

Graph based symmetry identification algorithms lead to a symmetry number based on topology. Information on bond lengths and torsion angles is thus lost and not taken into account. A fair number of attempts [27, 28, 34, 35] were made to automate the calculation of topological symmetry numbers, many of them created in the framework of automatic generation of kinetic models. In the Reaction Mechanism Generator (RMG) program [36], for example, σ is composed as a product of contributions of three types of centers of symmetry: atoms, bonds, and axes [37], cf. Eq. 6-1.

$$\sigma = \prod_i \sigma_{atom,i} \cdot \prod_j \sigma_{bond,j} \cdot \prod_k \sigma_{axis,k} \quad \text{Eq. 6-1}$$

A value is assigned to each contribution, depending on the nature of the surrounding constituents. Other approaches, e.g. in [27, 28, 34, 35], differ in the nature of the sub-molecular fragments that contribute to the overall symmetry but have in common that the algorithm is based on a limited set of heuristic rules. Heuristics lead to more efficient symmetry calculations, e.g. [9], but often limit the applicability of the algorithm to a subspace of classes of structures for which the rules were created. When applied outside the intended domain, errors in the calculated symmetry numbers may completely penalize the achieved accuracy of previously mentioned group contribution methods. Furthermore, often algorithms with completely different rules would be

developed for the calculation of reaction path degeneracies compared to the calculation of symmetry numbers [28, 38].

This work presents a novel algorithm that unifies both the evaluation of symmetry numbers of molecules as well as the reaction path degeneracy of reactions into a single methodology, by applying the same methodology for the calculation of symmetry numbers for molecules as well as for transition state structures. It does so by combining a topological representation of chemical structures along with the use of mathematical group theory to obtain an accurate value for a wide range of different components and reactions.

6.3 Theory

The symmetry number σ originally represented the external rotational symmetry number σ_{ext} , i.e. the number of indistinguishable positions that can be obtained by rotating a molecule as a rigid body about its center of mass [1], i.e. assuming that the internal coordinates of the molecule do not vary. In many cases, however, internal rotors exist, i.e. sub-molecular fragments of the molecule that rotate around single bonds. In that case, an internal symmetry number $\sigma_{\text{int},i}$ accounting for that rotor “ i ” is multiplied with σ_{ext} , cf. Eq. 6-2.

$$\sigma = \prod_i \sigma_{\text{int},i} \cdot \sigma_{\text{ext}} \quad \text{Eq. 6-2}$$

The idea of internal rotors arises from the thermal activation of internal degrees of freedom at temperatures higher than 0 K. Moreover, because the behavior of internal rotors is a function of the potential energy of the conformationally flexible species, its contribution to the symmetry number may be ill-defined. As was noted by Gilson et al. [39], one can argue that other sub-molecular processes such as small-amplitude atomic vibrations, torsions, and low-barrier inversions make the symmetry number rather a mathematical property than a physical property due to the flexible nature of a molecule. The proposition of a definition and use of a continuous symmetry number in contrast to the integral number is therefore not surprising [32]. According to that proposition, the symmetry number of ammonia would gradually change from three to six, as the angle between free electron pair, the nitrogen and hydrogen varies from 107 degrees to 90.

In classical transition state theory the rate coefficient $k_{\infty}(T)$ of a chemical reaction is expressed as Eq. 1-5 in Section 1.1.3 of Chapter 1. Rotational symmetry numbers of transition state and reactants can be factored out of the respective rotational partition functions leading to a reaction path degeneracy [3] n_e and partition functions q_{rot} in which all identical particles are treated as distinguishable, cf. Eq. 6-3.

$$\frac{q^{\ddagger}}{\prod_r q_r} = n_e \cdot \frac{q^{\ddagger}_{\text{no symmetry}}}{\prod_r q_{\text{no symmetry},r}} \quad \text{Eq. 6-3}$$

Rotational symmetry numbers appear in the denominator of the partition function and thus lead to the expression for n_e , cf. Eq. 6-4.

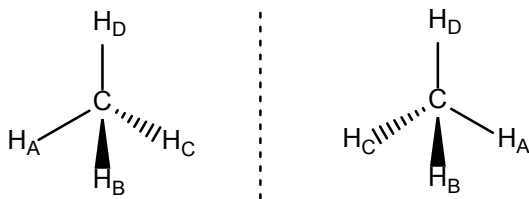
$$n_e = \frac{\prod_r \sigma_r}{\sigma^{\ddagger}} \quad \text{Eq. 6-4}$$

Eq. 6-4 provides a convenient way of calculating the reaction path degeneracy of reactions because it translates the abstract problem of the determination of reaction path degeneracy into the more tangible problem of the determination of rotational, global symmetry numbers of reactants r and transition state \ddagger . By doing so, a method that allows calculating the symmetry number of chemical graphs, regardless whether they are molecules or transition states can then serve as a unified way of calculating symmetry numbers and reaction path degeneracies.

The method that was used in this work for the calculation of the symmetry number σ continues the method proposed by Fernandez-Ramos et al. [40], cf. Eq. 6-5.

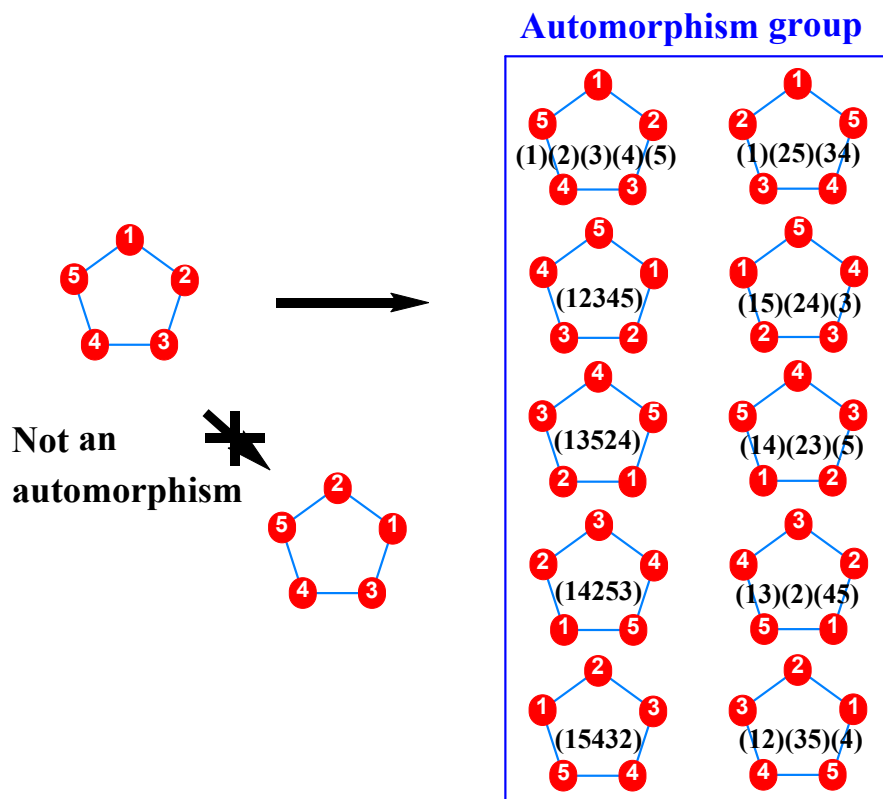
$$\sigma = \frac{m!}{n_d} \quad \text{Eq. 6-5}$$

m represents the number of equivalent atoms and n_d the number of unique configurations of the equivalent atoms. For example, methane has four equivalent hydrogen atoms leading to a numerator equal to twenty-four; labeling each of the equivalent hydrogen atoms so that they become distinguishable leads to a denominator equal to two, cf. Structure 6-1.



Structure 6-1: Labeling identical hydrogen atoms in methane results in two distinguishable configurations

The problem of quantifying the numerator of Eq. 6-5, i.e. the detection and enumeration of equivalent atoms in more complex molecules was addressed through the automorphism group order of the associated graph of the molecule. The automorphism group $Aut(G)$ of a weighed graph $G = (V, E, w, \lambda)$ defined by its vertices V , edges E , and respective weights w and λ , is a permutation group in which all the elements g fulfill two conditions [41]: 1) For every vertices u and v : $g(u)$ and $g(v)$ are adjacent if and only if u is adjacent to v and 2) for every vertex v and edge e : $w(g(v)) = w(v)$ and $\lambda(g(e)) = \lambda(e)$, i.e. the weights of the edges in the permuted graph should be identical to the weights of the same edges the original graph. In a chemical context, the weights of the vertices and edges of a graph refer to the chemical elements of the atoms, and the bond orders of the bonds in a molecule. A permutation of the atoms in a molecule belongs to the automorphism group if the way the individual atoms are connected to each other does not change. The order of the automorphism group $|Aut(G)|$ corresponds to the number of permutations belonging to the group. For example, the automorphism group order of a five-membered cyclic graph with identical nodes is ten because there are ten permutations of the nodes that comply with the definition of an automorphism permutation, cf. Structure 6-2.

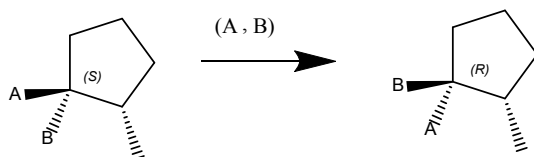


Structure 6-2: Ten permutations of a five-membered cyclic graph with identical nodes belonging to automorphism group. The permutations are represented in the cycle notation. In cycle notation, elements within parentheses are permuted in the order shown from left to right. E.g. (1 2) (3) denotes a permutation of elements 1 and 2, whereas element 3 is not permuted.

The calculation of the automorphism group can be performed for various types of graphs by means of well-established algorithms such as *nauty* [42], and *saucy* [43], and have found many applications in disciplines from chemistry, combinatorics and computer science [44-46]. Many authors acknowledged the similarities between the automorphism group of a graph and the topological symmetry of the molecule which lead to several algorithms for the construction of the automorphism group tailored for chemical graphs [26, 29, 41, 47-50]. It is beyond the scope of this work to try to improve the efficiency of the previous approaches to obtain the automorphism group. Rather, the automorphism group was used in a general methodology to calculate the rotational symmetry number of a chemical graph, cf. Eq. 6-6.

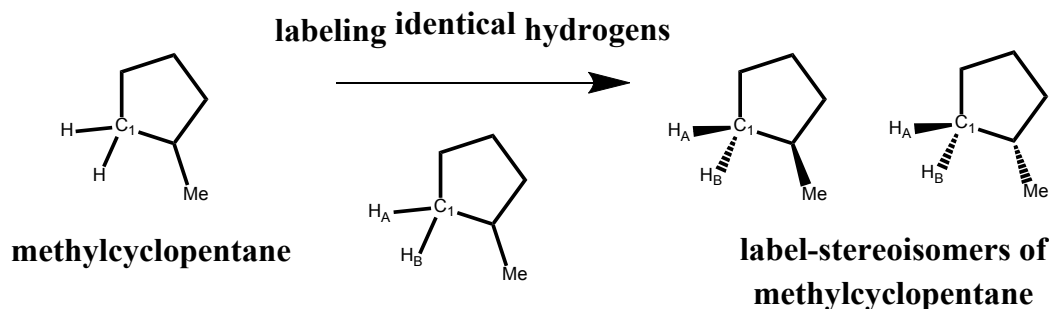
$$\sigma = \frac{|Aut(G)|}{n_d} \quad \text{Eq. 6-6}$$

The factor n_d in Eq. 6-6 refers to the number of distinguishable structures that arise when labeling all identical ligands in a molecule. This factor emerges because the automorphism group is a graph-theoretic concept, and thus cannot inherently deal with features of chemical objects that are not represented by graphs. These features revolve around the relative spatial arrangement of atoms, and represent a reduction of the overall symmetry of 3D chemical entities compared to mathematical graphs. As illustrated in Structure 6-3, a permutation of substituents A and B leads to a new graph that is identical to the original graph in terms of connectivity between the atoms.



Structure 6-3: An automorphism permutation of substituents A and B that results in the inversion of the absolute configuration of the stereocenter bonded to the two substituents.

If substituents A and B are homomorphic ligands (i.e. identical when considered in isolation), then this permutation belongs to the automorphism group of permutations of that molecule, and contributes a factor two to the total order of the automorphism group. If substituents A and B are heteromorphic ligands, then the carbon connected to A and B becomes asymmetric and the permutation of A and B reverses the absolute configuration of that carbon atom. To quantify the symmetry of the molecule, and not merely of the graph associated with the molecule, the automorphism group order should be reduced by a factor of two since this permutation does not generate a molecule that is identical to the original. The two possible enantiomers that arise in Structure 6-3 after making the two ligands A and B distinguishable is a specific example of the more general concept of *label-stereoisomers*, i.e. stereoisomers that originate by labeling of all homomorphic ligands in a molecule. The notion of label-stereoisomers is the key element of the approach taken in this thesis, to determine the symmetry number. The creation of label-stereoisomers is illustrated in Structure 6-4. The carbon atom with label C_1 in methylcyclopentane that bears two identical hydrogen atoms is a label-stereocenter, because two label-stereoisomers are created after labeling each of the two identical hydrogen atoms, i.e. H_A and H_B , connected to C_1 . The two label-stereoisomers differ in the relative arrangement of the methyl group with respect to the two distinct H_A and H_B hydrogen atoms.



Structure 6-4: The creation of label-stereoisomers after labeling identical atoms in methylcyclopentane.

Correctly counting the number of label-stereoisomers, referred to as label-stereoisomer enumeration, is crucial in accurately determining the symmetry number. The approach taken in this work is based on what is done in traditional stereochemistry: the stereochemical properties of a molecule are described by sub-molecular stereocenters such as chiral carbon atoms or cis/trans double bonds. In other words, the stereochemistry of a molecule is factorized into separate stereocenters that comprise the stereochemical information of the molecule. Similarly, the problem of label-stereoisomer enumeration was tackled by factorizing it in the enumeration of separate sub-molecular label-stereocenters. A label-stereocenter is a center that becomes a stereocenter after distinguishing between its homomorphic ligands. A molecule or transition state may thus consist of one or more label-stereocenters. If each label-stereocenter “i” can adopt 2 configurations, the total number of label-stereoisomers n_d is then calculated according to Eq. 6-7.

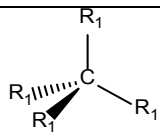
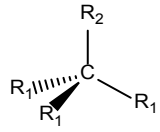
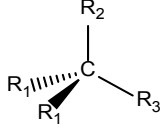
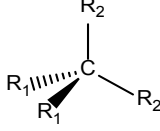
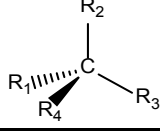
$$n_d = 2^{n_{label}} \quad \text{Eq. 6-7}$$

with n_{label} the total number of label-stereocenters i_L in the molecule or transition state. The correct determination of the number of label-stereocenters in a molecule depends on the types of label-stereocenters that are accounted for. This study addresses two types of label-stereocenters: label-asymmetric carbon atoms and label-asymmetric double bonds, which, because of their definition, bear many similarities to chiral carbon atoms, and cis/trans double bonds respectively.

Label-asymmetric carbon atoms are surrounded by four substituents. Depending on whether or not the substituents are homomorphic, they can adopt five possible combinations surrounding the central carbon, cf. Table 6-1. For four out of five possible configurations of substituents, the automorphism group order overestimates the symmetry number by a factor of two because one pair of label-stereoisomers exists in each case. Only for the configuration with four distinct

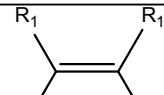

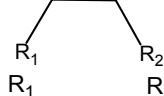
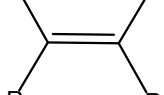
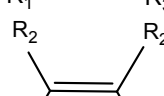
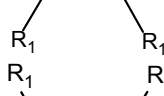
substituents, i.e. a chiral carbon, there exists only one label-stereoisomer, because there are no permutations of homomorphic substituents.

Table 6-1: Possible environments of a tetrahedral carbon atom along with the automorphism group order, the rotational symmetry number σ , and the number of label-stereoisomers n_d .

		$ Aut(G) $	n_d	σ
4 equal		24	2	12
3 equal		6	2	3
2 equal		2	2	1
2x2 equal		4	2	2
4 distinct		1	1	1

The same principle is applied to label-asymmetric double bonds; all possible combinations of the four substituents on both sides of the double bond are depicted in Table 6-2. Six, instead of five configurations are now possible, because the double bond discriminates between the *cis* and *trans* side relative to a particular substituent. Likewise to the case of tetrahedral atoms, the combinations of substituents that give rise to *cis/trans* stereoisomerism have one label-stereoisomer, whereas the cases with combinations of substituents that do not result in *cis/trans* stereoisomerism (cases A, B, C, cf. Table 6-2), have two possible label-stereoisomers each.

Table 6-2: Possible environments of a double bond along with the automorphism group order, the rotational symmetry number σ , and the number of label-stereoisomers n_d

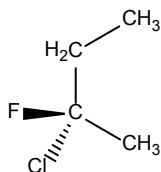
			$ Aut(G) $	n_d	σ
A	4 equal		8	2	4
B	2 equal on each side		4	2	2
C	2 equal on one side		2	2	1
D	2x2 distinct groups, cis		2	1	2
E	2x2 distinct groups, trans		2	1	2
F	4 distinct groups		1	1	1

From the analysis above, it can be concluded that a structural feature such as a tetrahedral carbon or a carbon-carbon double bond is either a label-stereocenter, or a real stereocenter; they cannot be a real stereocenter and label-stereocenter at the same time. Real stereocenters are surrounded by distinct (i.e. non-homomorphic) ligands, whereas label-stereocenters must bear at least two identical ligands. This relation between real stereocenters and label-stereocenters is useful to enumerate the number of label-stereocenters n_{label} , because it can now be expressed as the difference between the number of potential stereocenters n_{pot} and the number of real stereocenters n_{real} , cf. Eq. 6-8.

$$n_{label} = n_{pot} - n_{real}$$

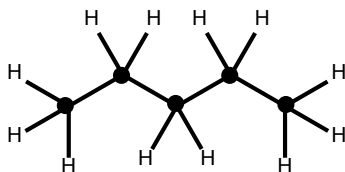
Eq. 6-8

For example, 2-chloro-2-fluorobutane counts four tetrahedral carbons ($n_{pot} = 4$), and one chiral carbon ($n_{real} = 1$) resulting in three label-stereocenters ($n_{label} = 4 - 1 = 3$).



Structure 6-5: 2-Chloro-2-fluorobutane counts four tetrahedral carbons, and one chiral carbon resulting in three label-asymmetric tetrahedral carbons.

The enumeration of the number of potential stereocenters n_{pot} is trivial and merely requires the summation of easily detectable moieties, e.g. tetrahedral carbons, double bonds. For example, pentane contains five potential stereocenters ($n_{pot} = 5$), corresponding to the five tetracoordinate carbon atoms, each bearing either two or three identical hydrogen atoms, cf. Structure 6-6.



● potential stereocenters

Structure 6-6: n-pentane, bearing five potential stereocenters, corresponding to the five carbon atoms bearing identical hydrogen atoms

Hence, the problem of the label-stereoisomer enumeration is translated into the problem of stereocenter detection and enumeration, for which solutions are far more abundant than for the former problem. Stereocenters are detected if a formal definition of a stereocenter is provided. For example, a formal definition of stereogenic carbon atoms is formulated as “tetracoordinate carbon atoms bearing four constitutionally different substituents”. A formal definition for stereogenic double bonds is formulated as “sets of two or more contiguously-joined, doubly-bonded carbon atoms, each terminus of which bears constitutionally different substituents”. Once stereocenters are formally characterized, computer algorithms can search for these stereocenters in molecules based on the provided definition.

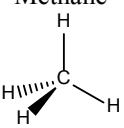
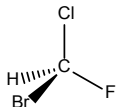
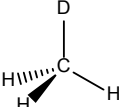
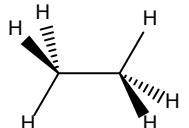
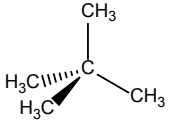
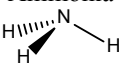
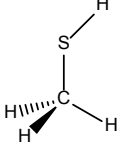
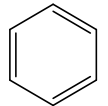
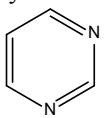
6.4 Results and discussion

A code of the current algorithm was programmed in Java, and was given the name SIGMA. The symmetry calculation algorithm is part of a framework with the purpose to automatically construct detailed kinetic models, called Genesys [51] and introduced in Chapter 5. A code that implements the stereocenter detection algorithm of Razinger et al. [52] was found in OpenBabel [53] and used as such in the present work. An algorithm and code by Torrance [54] was used as such for the identification of the automorphism group of a graph. The Chemistry Development Kit (CDK) [55, 56], an open-source chemoinformatics toolkit, was used for the internal representation of molecules in the code of the proposed algorithm. CDK provides a very flexible and powerful data structure for molecules, which supports many heteroelements such as sulfur, nitrogen, oxygen, etc.... CDK also includes algorithms that enable the conversion of line identifiers such as InChI [57] and SMILES [58] into the CDK molecule representation. More information on the integration of CDK in Genesys can be found in Chapter 5 of this thesis. The remainder of the algorithmic elements for the calculation of symmetry numbers and reaction path degeneracy were implemented as explained in the present chapter.

To illustrate the functionality of the current algorithm, Table 6-3 shows a selection of molecules along with the symmetry number calculated by the program, and the different contributions that lead to that result. All of the calculated symmetry numbers correspond to the expected value. The automorphism group algorithm distinguishes atoms based on elemental and isotope information: 1-deuterio-methane is assigned a symmetry number of three instead of twelve. The examples of ethane, and neopentane with respective calculated symmetry numbers of 18 and 972, indicate that the current algorithm calculates the total symmetry number, in which internal rotors around single bonds are always included, regardless of potential energy considerations. For 5-methyl-1,3-cyclopentadiene with a combination of a two-fold and three-fold symmetry, this leads to a symmetry number of six. The example of methanethiol shows that not only hydrocarbons are correctly characterized, but also other heteroelements such as sulfur, nitrogen, oxygen, etc... thanks to the use of CDK. For atoms surrounded by a single lone electron pair, the latter is regarded as a fourth substituent. For example, for ammonia a symmetry number of three is calculated and not six. The aromatic components benzene and pyrimidine illustrate the treatment

of aromatics. First, bonds that are part of an aromatic ring are identified and converted into a new uniform aromatic bond type. Next, this aromatized structure is used in determining the automorphism group, which in the case of benzene leads to an automorphism group order of twelve. CDK's algorithm "HueckelAromaticityPerceiver" was used for aromaticity detection such that structures like pyrimidine are also correctly perceived as aromatic.

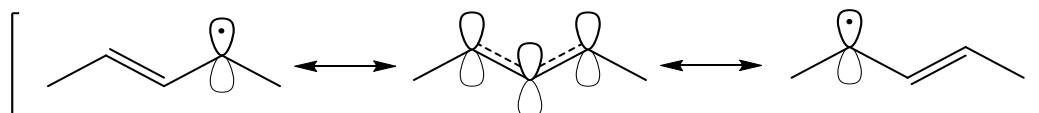
Table 6-3: Selected molecules and their respective symmetry number, automorphism group order, number of label-stereoisomers, and the contributions to n_d .

Molecule	σ	$ Aut(G) $	n_d	n_{pot}	n_{real}	n_{label}
Methane 	12	24	2	1	0	1
Fluoro-chloro-bromomethane 	1	1	1	1	1	0
Deuteriomethane 	3	6	2	1	0	1
Ethane 	18	72	4	2	0	2
2,2-dimethylpropane 	972	31104	32	5	0	5
Ammonia 	3	6	2	1	0	1
Methanethiol 	3	6	2	1	0	1
Benzene 	12	12	1	0	0	0
Pyrimidine 	2	2	1	0	0	0

5-methyl-1,3-cyclopentadiene 	6	192	32	5	0	5
Allyl radical 	2	8	4	2	0	2
pent-3-en-2-yl radical 	18	72	4	4	2	2
1-phenylethyl cation 	3	12	4	2	0	2
3-ethenyl-1,4-pentadiene 	3	48	16	4	0	4
tert-butyl radical 	162	1296	8	3	0	3
S ₈ 	16	16	1	0	0	0
C ₉ H ₁₈ O ₆ 	$3^6 \cdot 2^3 \cdot 6$	$(3!)^6 \cdot 2^3 \cdot 6$	2^9	9	0	9

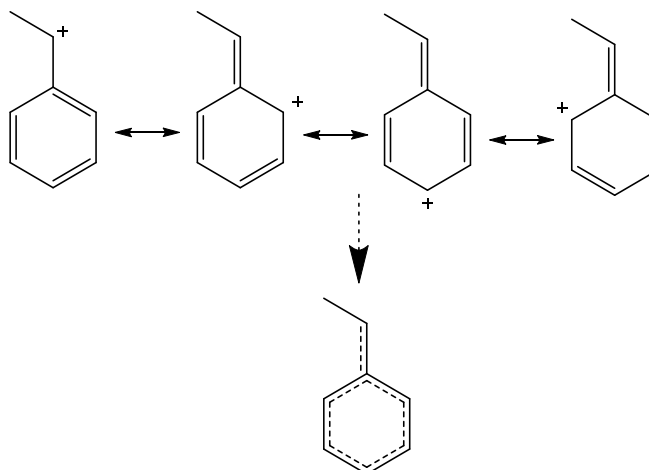
Molecules and reactions not only differ from mathematical graphs by the spatial arrangement of atoms, but they also possess sub-molecular entities. More specifically, the presence of electrons and orbitals, e.g. unpaired electrons in radical species, lone electron pairs, or the lack or excess of electrons in charged species, can lead to electron delocalization. The latter can result in differences in the geometry of the resonance hybrid relative to the resonance structures, with bond lengths that deviate from the default bond length of the resonance structures. Electron delocalization can increase or decrease the overall symmetry of the molecule, e.g. pent-3-en-2-yl, cf. Structure 6-7. The delocalization of the three electrons in the p-orbitals leads to a resonance

hybrid with an external symmetry number of two rather than one, as expected for the resonance structure.



Structure 6-7: Resonance structures of pent-3-en-2-yl (left and right) and symmetric resonance hybrid (center).

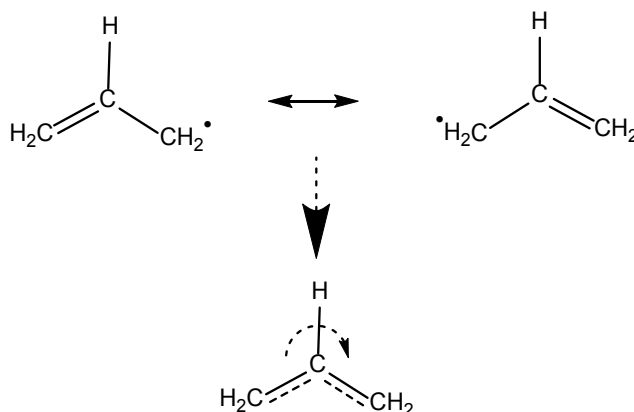
Electron delocalization also affects the rotor behavior of sub-molecular fragments. An example of this is the 1-phenylethyl cation (Structure 6-8). The presence of the exocyclic electron hole leads to resonance with the aromatic ring. The sp^2 -hybridization of the exocyclic double bond prevents the aromatic ring from rotating; the internal benzene rotor contribution to σ should therefore be excluded.



Structure 6-8: Resonance of 1-phenylethyl cation.

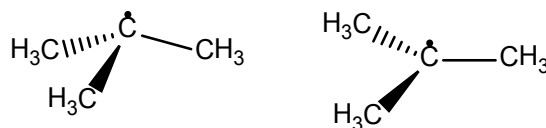
The influence of electron delocalization on the rotational symmetry number was quantified using the resonance hybrid representation of the structure with delocalized electrons. The resonance hybrid is constructed by converting the bond orders of the resonance bonds into new bond orders that are intermediate between the bond orders in the resonance structures. The resonance structures were generated using an algorithm available in CDK under the name “StructureResonanceGenerator” [56]. The resonance hybrid graph representation containing the new intermediate bond orders was subsequently used in determining the automorphism group order. The example of allyl, cf. Structure 6-9, illustrates the procedure. The creation of the two intermediate bond orders, creates an additional C_2 symmetry element along the axis of the central

carbon and the adjacent hydrogen atom (cf. Structure 6-9), but it also prevents the bond in the original resonance structure between the central carbon atom and the electron deficient carbon atom from rotating, since the intermediate bond order in the resonance hybrid is associated with a non-rotatable bond.



Structure 6-9: Resonance hybrid graph for allyl constructed by replacing the resonating single (in green) and double (in blue) bonds with a new type of intermediate bond order type (in red).

For t-butyl radical a symmetry number of 162 is calculated and corresponds to a planar configuration of the methyl groups around the central carbon, Structure 6-10.



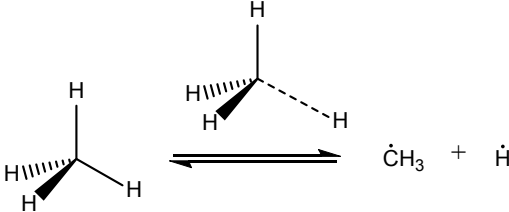
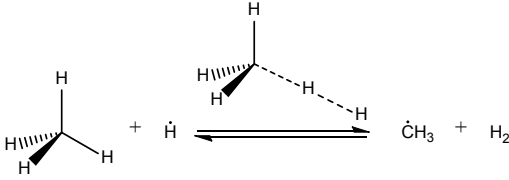
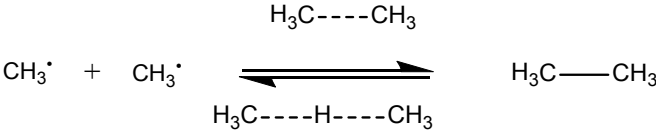
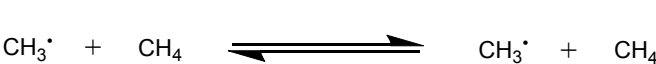
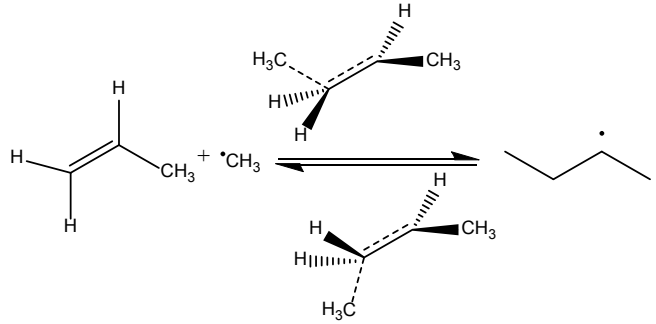
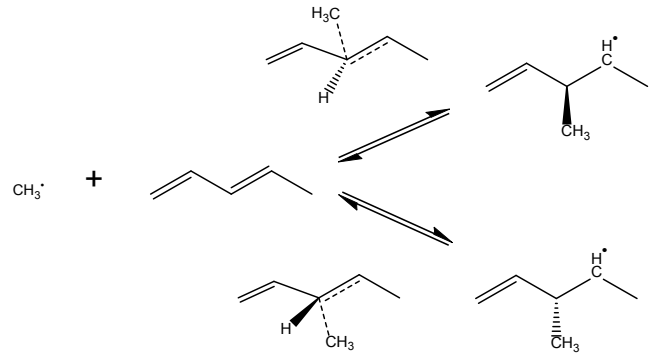
Structure 6-10: Pyramidal (left) vs. trigonal (right) structure of tert-butyl radical.

At room temperature conditions, t-butyl radical is considered to be pyramidal, such that the three axes with a two-fold symmetry element along each of the C-C bonds no longer contribute to the overall symmetry, and lead to a global symmetry number of 81 instead of the 162 calculated through the current algorithm. Since the inversion barrier height in t-butyl radical is low (3 kJ mol⁻¹ [59]), the global symmetry number of 162 is also justifiable and corresponds to a time-averaged value of the molecule. More generally, the trigonal or pyramidal orientation of the three substituents relative to the central carbon are a function of the bulkiness and the electronegativity of the substituents. In contrast to ammonia, where the lone pair was considered as a substituent that plays a similar role as an atom, it was opted to consider the most symmetric configuration for structures with unpaired electrons or electron deficient structures.

For cyclic molecules in which the ring atoms have significant conformational freedom, the symmetry number of the most symmetric conformation was calculated, although this does not necessarily comply with the minimum energy conformation of the molecule. For example, the octatomic S₈ ring adopts a crown-like minimum energy conformation reducing the symmetry from sixteen to eight. In the C₉H₁₈O₆ component of Table 6-3 a supposed C₃ rotation axis is absent because of the asymmetric 3D spatial arrangement of the O-O groups. However, similarly to the case of t-butyl, at high enough temperatures, these structures adopt more symmetric configurations at a large enough time-scale of observation. Therefore, the calculated symmetry number may be considered as a time-averaged, upper bound to the symmetry number of individual conformers.

The algorithm for the calculation of σ presented in this work is based on graph-theoretic fundamentals, and thus is equally applicable to the calculation of σ^\ddagger as it is for $\sigma_{\text{reactants}}$ provided that a graph representation of the transition state is available. In Genesys [51], reaction family templates contain a stepwise description of the bonds that are being broken or formed. Because of this mapping between reactant and product atoms, the graph corresponding to the transition state structure is therefore available and can be used for the determination of σ^\ddagger . Table 6-4 shows a selection of reactions, and their transition states along with the calculated symmetry number, and the different contributions that lead to that result.

Table 6-4: Selected reactions and their respective reaction path degeneracy, symmetry contributions from reactants and transition state.

Reaction	n_e	$\prod_r \sigma_r$	σ^\ddagger
R 6-1 	4	12	3
R 6-2 	4	12	3
R 6-3 	4	36	18
R 6-4 	8	72	18
R 6-5 	2	18	9
R 6-6 	2	18	9

Transition state structures differ from molecules because they contain a number of bonds that are being formed or broken. These intermediate bond orders have to be explicitly taken into account because they are distinguishable from conventional bonds such as single or double bonds and thus reduces the overall symmetry of the transition state. For example, R 6-1 shows the homolytic C–H scission reaction of a methane molecule forming a methyl and hydrogen. The length of the bond being broken is longer than the original C–H bond in methane and hence lowers the rotational symmetry number of the transition state structure to three. This results in a reaction path degeneracy of four. A similar situation occurs in the H-abstraction of hydrogen from methane, cf. R 6-2. In case reactants are identical the mutual permutation of the reactants needs to be accounted for in the reaction path degeneracy. For example, the reaction path degeneracy of R 6-3 is calculated as four, corresponding to a value of two obtained by Eq. 6-4, multiplied by a factor of two. R 6-4 shows the abstraction of a hydrogen atom from methane by methyl. A reaction path degeneracy of eight is calculated. The current approach assumes that the two transitioning bonds have the same bond order, and thus the same bond length. This results in a transition state in which both carbons and their corresponding substituents can be permuted, leading to an additional factor of two in the automorphism group order. However, this twofold symmetry axis in the transition state is not present along the reaction path. Any symmetry element that is present in the transition state but not along the reaction path has to be removed from the transition state partition function to avoid imposing this symmetry element on the entire reaction path, as was pointed out by Pollak and Pechukas [3]. When Genesys encounters a symmetric reaction, i.e. when reactants are found to be identical to products, the reaction path degeneracy of symmetric reaction is divided by a factor of two. The addition of methyl to propene leading to 2-butyl (R 6-5) is a reaction in which attacking methyl can approach the plane of the double bond from two sides. Because the formed transition states are not superimposable, they are described as molecular chiral transition states, and represent two distinguishable iso-energetic pathways to yield 2-butyl. The calculated reaction path degeneracy is four, the product of the value obtained by Eq. 6-4 and a factor two accounting for the chirality of the transition states. In case the two chiral transition states lead to two distinguishable products, e.g. enantiomers, then the reaction path degeneracy of the reaction that leads to either one of the products does not need the factor two. This is the case for the addition of methyl to trans-1,3-pentadiene's third carbon leading to a chiral 3-methyl-pent-1-en-4-yl (R 6-6). The reaction path

degeneracy for each channel is calculated as two. In general, in this thesis, enantiomers are treated as distinct species, i.e. rate expressions for enantiomers are not lumped together, as was done elsewhere [5, 21, 60]. This implies that if a pair of enantiomers undergoes the same reaction to distinct enantiomer products, connected by only one transition state, reaction path degeneracies of these reactions are treated separately.

6.4.1 Validation

Assessing the accuracy of SIGMA requires the comparison of the calculated symmetry number or reaction path degeneracy to an expected value. Large and diverse sets of molecules and elementary reactions annotated with the expected rotational symmetry number or reaction path degeneracy are difficult to find in literature, because the symmetry contribution is often incorporated in experimentally observable quantities such as entropy, or rate coefficients. Next to the comparison of the current algorithm with expected values, the results of SIGMA were also compared to the calculated values produced by the RMG algorithm. While a fair number of proposed algorithms were found in literature, the algorithm implemented in the RMG software [61] is one of the few recent, available programs that are able to calculate symmetry numbers of molecules solely based on topological information. The RMG algorithm is tailored to yield accurate results for species containing C,H and O atoms and should therefore perform well for molecules within its scope. The test set consists of 193 hydrocarbons, the symmetry numbers of which were published as part of *ab initio* calculations [62].

Hundred and six of them are hydrocarbon radicals bearing one unpaired electron while the remaining 87 are closed-shell species. Of the latter, symmetry numbers of 85 molecules calculated by SIGMA agreed with the expected values. RMG calculated identical results, except for 16 cases that contained phenyl fragments. In RMG these fragments are not considered as internal rotors; leading to a factor of two difference for structures with this feature. More differences between expected and calculated values existed for the 106 hydrocarbon radicals, and overestimated the symmetry number by a factor two or three. The origin for this relates to the aforementioned ambiguities of the planarity or pyramidalty of tri-substituted radicals.

Nevertheless, the accuracy of both algorithms was very similar and agreed well with the expected symmetry number.

A second database, the Third Millennium Ideal Gas and Condensed Phase Thermochemical Database for Combustion [63], was used to compare the calculated symmetry numbers of the current algorithm. This database differs from the first database because of its diversity in elemental composition, electronic and molecular structure. A total of 445 entries of expected symmetry numbers were obtained from it. The current program was able to calculate symmetry numbers in all cases and lead to the same value for the symmetry number as to what was expected in 92% of the cases. A large fraction of the remaining deviations originated from the ambiguity related to the (non-) planarity and the associated asymmetry of ring structures such as the S_8 ring.

6.4.2 Application

Next to applications in automatic mechanism generation, cf. Chapter 7 and Chapter 8, where this algorithm originally was developed for, an example of another possible application domain is illustrated here. The current algorithm was used to calculate the symmetry number for drug-like molecules in an assay for drug discovery campaigns. The publicly available Maybridge Screening Collection was used for this purpose and consists of a highly diverse set of 52160 organic components [64]. Obtaining information on the degree of symmetry, expressed by the symmetry number could be useful in the selection process of suitable candidates for pharmacophore or docking algorithms. Because this database does not contain 3D coordinates of these components, the current topological algorithm becomes critical for the quick and efficient determination of the symmetry measure. Because this collection of molecules did not have expected symmetry numbers to compare with, it was verified for how many molecules the symmetry number could be calculated. This was possible in 99% of all cases. The failing 1% was due to molecules such as salts for which the algorithm likely gives erroneous results. Figure 6-1 shows the frequency of symmetry number values encountered in the database.

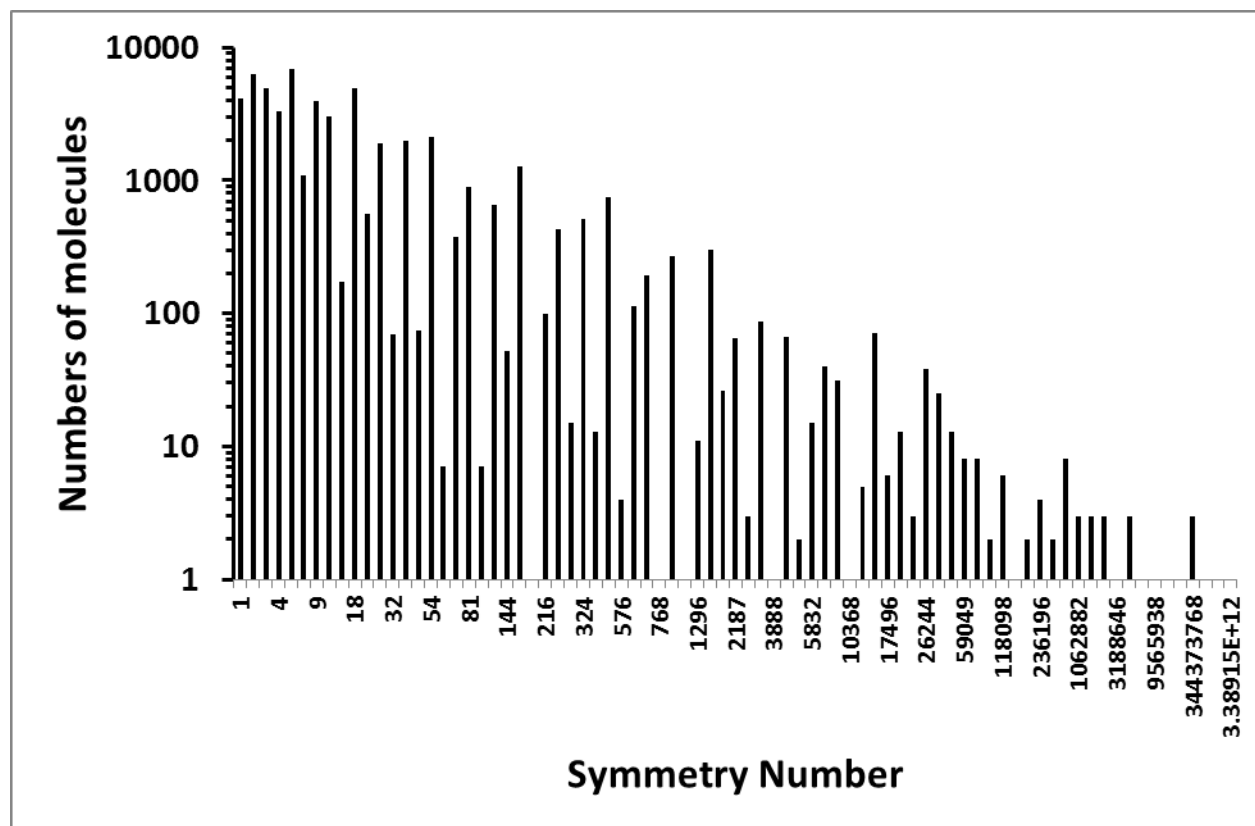
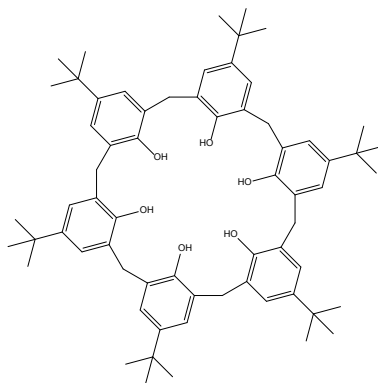


Figure 6-1: Histogram of the frequencies of the calculated symmetry number for the Maybridge data set.

The range of the symmetry number values is vast and reaches up to $O(10^{12})$, illustrating the wide variety of symmetry present in the database. Only 8% of the components do not exhibit any symmetry, indicated by a value of one. The remaining of these molecules do possess some level of symmetry, e.g. structures with a high symmetry number, do indeed possess a high degree of symmetry, cf. Structure 6-11. The frequency of the encountered symmetry numbers rapidly decreases with 50% of the structures having a symmetry number of eight or less.



Structure 6-11: Molecule in the Maybridge database identified as the structure with the highest degree of symmetry (symmetry number equal to 3389154437772).

The calculation of the symmetry numbers for the Maybridge data set took 10h on a dual core Intel Centrino 3GHz computer, with individual calculation times per molecule heavily depending on the size and the symmetry of the molecule. The largest fraction of the CPU time is spent on the automorphism group construction.

6.5 Conclusions

The introduction of the concept of label-stereoisomers allowed calculating the symmetry number of molecules and reaction path degeneracy of elementary reactions. Label-stereoisomers are stereoisomers that arise after labeling identical ligands in a structure. A unified approach for both symmetry numbers and reaction path degeneracies was presented by making use of a chemical graph representation of the transition state structure associated to the elementary reaction. The algorithm builds further upon the use of the automorphism group order to represent topological symmetry and calculates total symmetry numbers, i.e. also accounting for internal rotations.

The label-stereoisomer enumeration problem was subsequently solved by identifying stereocenters, using available stereocenter identification algorithms from literature. The current graph-theoretic approach is chemistry-ignorant, implying that it does not assume anything on the molecular structure *a priori*. Comparison with existing databases of symmetry numbers showed that the calculated symmetry numbers agree well with expected values, indicating that the used

graph representation of molecules and transition states, augmented with non-topological attributes suffices in describing the rotational symmetry. In cases with ambiguity regarding the value of the symmetry number, the calculated symmetry number should be regarded as a time-averaged upper limit. The use of a topological representation of molecules and transition states also avoided the need for three dimensional atomic coordinates, so that quickly scanning large database of molecules to detect symmetry becomes a viable option.

6.6 References

- [1] P. Atkins, J. De Paula, Atkins' physical chemistry, Oxford University Press, 2010.
- [2] S.-K. Lin, Correlation of Entropy with Similarity and Symmetry, *Journal of Chemical Information and Computer Sciences*, 36 (1996) 367-376.
- [3] E. Pollak, P. Pechukas, Symmetry numbers, not statistical factors, should be used in absolute rate theory and in bronsted relations, *J. Am. Chem. Soc.*, 100 (1978) 2984-2991.
- [4] D.M. Bishop, K.J. Laidler, Symmetry Numbers and Statistical Factors in Rate Theory, *J. Chem. Phys.*, 42 (1965) 1688-1693.
- [5] M.A. Baltanas, K.K. Vanraemdonck, G.F. Froment, S.R. Mohedas, Fundamental Kinetic Modeling of Hydroisomerization and Hydrocracking on Noble-Metal-Loaded Faujasites .1. Rate Parameters for Hydroisomerization, *Industrial & Engineering Chemistry Research*, 28 (1989) 899-910.
- [6] A.J. Karas, R.G. Gilbert, M.A. Collins, Rigorous Derivation of Reaction-Path Degeneracy in Transition-State Theory, *Chem. Phys. Lett.*, 193 (1992) 181-184.
- [7] S.H. Yalkowsky, R.M. Dannenfelser, P. Myrdal, P. Simamora, D. Mishra, Unified Physical Property Estimation Relationships (Upper), *Chemosphere*, 28 (1994) 1657-1673.
- [8] S.H. Yalkowsky, P. Myrdal, R.M. Dannenfelser, P. Simamora, Upper .2. Calculation of Physical-Properties of the Chlorobenzenes, *Chemosphere*, 28 (1994) 1675-1688.
- [9] M.L. Contreras, J. Alvarez, D. Guajardo, R. Rozas, Understanding topological symmetry: A heuristic approach to its determination, *J. Comput. Chem.*, 29 (2008) 588-600.
- [10] H.H. Adler, P.F. Kerr, Infrared spectra, symmetry and structure relations of some carbonate minerals, *American Mineralogist*, 48 (1963) 839-853.
- [11] H.H. Adler, P.F. Kerr, Variations in infrared spectra molecular symmetry and site symmetry and of sulfate minerals., *American Mineralogist*, 50 (1965) 132-147.
- [12] D.C. Harris, M.D. Bertolucci, *Symmetry and Spectroscopy: An Introduction to Vibrational and Electronic Spectroscopy*, DoverPublications. com, 1978.
- [13] M.K. Gilson, J.A. Given, B.L. Bush, J.A. McCammon, The statistical-thermodynamic basis for computation of binding affinities: A critical review, *Biophysical Journal*, 72 (1997) 1047-1069.
- [14] M. Hechinger, K. Leonhard, W. Marquardt, What is Wrong with Quantitative Structure-Property Relations Models Based on Three-Dimensional Descriptors?, *Journal of Chemical Information and Modeling*, 52 (2012) 1984-1993.
- [15] F. Battin-Leclerc, E. Blurock, R. Bounaceur, R. Fournet, P.-A. Glaude, O. Herbinet, B. Sirjean, V. Warth, Towards cleaner combustion engines through groundbreaking detailed chemical kinetic models, *Chemical Society Reviews*, 40 (2011) 4762-4782.
- [16] S.W. Benson, *Thermochemical Kinetics*, John Wiley & Sons, New York, 1976.
- [17] M.K. Sabbe, M. Saeys, M.F. Reyniers, G.B. Marin, V. Van Speybroeck, M. Waroquier, Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals, *Journal of Physical Chemistry A*, 109 (2005) 7466-7480.
- [18] M. Saeys, M.F. Reyniers, G.B. Marin, V. Van Speybroeck, M. Waroquier, Ab initio group contribution method for activation energies for radical additions, *AIChE Journal*, 50 (2004) 426-444.
- [19] R. Sumathi, W.H. Green, A priori rate constants for kinetic modeling, *Theoretical Chemistry Accounts*, 108 (2002) 187-213.

- [20] M.K. Sabbe, M.-F. Reyniers, V. Van Speybroeck, M. Waroquier, G.B. Marin, Carbon-centered radical addition and beta-scission reactions: Modeling of activation energies and pre-exponential factors, *Chemphyschem*, 9 (2008) 124-140.
- [21] M.K. Sabbe, A.G. Vandeputte, M.-F. Reyniers, M. Waroquier, G.B. Marin, Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions, *Phys. Chem. Chem. Phys.*, 12 (2010) 1278-1298.
- [22] M.K. Sabbe, M.-F. Reyniers, M. Waroquier, G.B. Marin, Hydrogen Radical Additions to Unsaturated Hydrocarbons and the Reverse beta-Scission Reactions: Modeling of Activation Energies and Pre-Exponential Factors, *Chemphyschem*, 11 (2010) 195-210.
- [23] A.G. Vandeputte, M.-F. Reyniers, G.B. Marin, Kinetics of Homolytic Substitutions by Hydrogen Atoms at Thiols and Sulfides, *Chemphyschem*, 14 (2013) 1703-1722.
- [24] A.G. Vandeputte, M.-F. Reyniers, G.B. Marin, Kinetic Modeling of Hydrogen Abstractions Involving Sulfur Radicals, *Chemphyschem*, 14 (2013) 3751-3771.
- [25] A.G. Vandeputte, M.K. Sabbe, M.-F. Reyniers, G.B. Marin, Kinetics of alpha hydrogen abstractions from thiols, sulfides and thiocarbonyl compounds, *Phys. Chem. Chem. Phys.*, 14 (2012) 12773-12793.
- [26] J. Ivanov, G. Schuurmann, Simple algorithms for determining the molecular symmetry, *Journal of Chemical Information and Computer Sciences*, 39 (1999) 728-737.
- [27] W.P. Walters, S.H. Yalkowsky, ESCHER - A computer program for the determination of external rotational symmetry numbers from molecular topology, *Journal of Chemical Information and Computer Sciences*, 36 (1996) 1015-1017.
- [28] C. Muller, G. Scacchi, G.M. Côme, A topological method for determining the external symmetry number of molecules, *Computers & Chemistry*, 15 (1991) 17-27.
- [29] W. Chen, J. Huang, M.K. Gilson, Identification of symmetries in molecules and complexes, *Journal of Chemical Information and Computer Sciences*, 44 (2004) 1301-1313.
- [30] S. Patchkovskii, SYMMETRY - Brute force symmetry analyzer - <http://www.cobalt.chem.ucalgary.ca/ps/symmetry/>, (2003).
- [31] C. Dryzun, A. Zait, D. Avnir, Quantitative Symmetry and Chirality-A Fast Computational Algorithm for Large Structures: Proteins, Macromolecules, Nanotubes, and Unit Cells, *J. Comput. Chem.*, 32 (2011) 2526-2538.
- [32] E. Estrada, D. Avnir, Continuous symmetry numbers and entropy, *J. Am. Chem. Soc.*, 125 (2003) 4368-4375.
- [33] H. Zabrodsky, S. Peleg, D. Avnir, Continuous Symmetry Measures, *J. Am. Chem. Soc.*, 114 (1992) 7843-7851.
- [34] C. Muller, V. Michel, G. Scacchi, G.M. Côme, 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*, 92 (1995) 1154-1178.
- [35] O.D.G. Ytalo, M.M.C. Jorge, Graph-Based Method for the Automated Calculation of Thermochemical Properties of Components and Intermediate Species in the Hydroprocessing of Light Oil Fractions, *Industrial & Engineering Chemistry Research*, 50 (2011) 12774-12783.
- [36] W.H. Green, J.W. Allen, R.W. Ashcraft, G.J. Beran, C.A. Class, C. Gao, C.F. Goldsmith, M.R. Harper, A. Jalan, G.R. Magoon, D.M. Matheu, S.S. Merchant, J.D. Mo, S. Petway, S. Raman, S. Sharma, J. Song, K.M. Van Geem, J. Wen, R.H. West, A. Wong, H.S. Wong, P.E. Yelvington, J. Yu, RMG - Reaction Mechanism Generator v4.0, (2013).
- [37] J. Yu, Estimation method for the thermochemical properties of polycyclic aromatic molecules, Department of Chemical Engineering, Massachusetts Institute of Technology (2005).
- [38] V. Warth, N. Stef, P.A. Glaude, F. Battin-Leclerc, G. Scacchi, G.M. Côme, Computer-aided derivation of gas-phase oxidation mechanisms: Application to the modeling of the oxidation of n-butane, *Combustion and Flame*, 114 (1998) 81-102.
- [39] M.K. Gilson, K.K. Irikura, Symmetry Numbers for Rigid, Flexible, and Fluxional Molecules: Theory and Applications, *J. Phys. Chem. B*, 114 16304-16317.
- [40] A. Fernandez-Ramos, B.A. Ellingson, R. Meana-Paneda, J.M.C. Marques, D.G. Truhlar, Symmetry numbers and chemical reaction rates, *Theoretical Chemistry Accounts*, 118 (2007) 813-826.
- [41] A.R. Ashrafi, On symmetry properties of molecules, *Chem. Phys. Lett.*, 406 (2005) 75-80.
- [42] B.D. McKay, Practical Graph Isomorphism, *Congressus Numerantium*, 30 (1981) 45-87.
- [43] P.T. Darga, K.A. Sakallah, I.L. Markov, Faster symmetry discovery using sparsity of symmetries, *Proceedings of the 45th annual Design Automation Conference*, (2008) 149-154.
- [44] F.A. Aloul, A. Ramani, I.L. Markov, K.A. Sakallah, Solving difficult instances of boolean satisfiability in the presence of symmetry, *Computer-Aided Design of Integrated Circuits and Systems*, *IEEE Transactions on*, 22 (2003) 1117-1137.

- [45] G. Brinkmann, Isomorphism rejection in structure generation programs, *Discrete mathematical chemistry*, 51 (2000) 25-38.
- [46] T. Junttila, On the symmetry reduction method for Petri Nets and similar formalisms, Helsinki University of Technology, 2003.
- [47] A.R. Ashrafi, M.R. Ahmadi, New computer program to calculate the symmetry of molecules, *Cent. Eur. J. Chem*, 3 (2005) 647-657.
- [48] M. Randic, Discerning Symmetry Properties of Graphs, *Chem. Phys. Lett.*, 42 (1976) 283-287.
- [49] M. Randic, Recognition of Identical Graphs Representing Molecular Topology, *J. Chem. Phys.*, 60 (1974) 3920-3928.
- [50] J. Ivanov, Molecular symmetry perception, *Journal of Chemical Information and Computer Sciences*, 44 (2004) 596-600.
- [51] N.M. Vandewiele, K.M. Van Geem, M.F. Reyniers, G.B. Marin, Genesys: Kinetic model construction using chemoinformatics, *Chemical Engineering Journal*, 207 (2012) 526-538.
- [52] M. Razinger, K. Balasubramanian, M. Perdih, M.E. Munk, Stereoisomer generation in computer-enhanced structure elucidation, *Journal of Chemical Information and Computer Sciences*, 33 (1993) 812-825.
- [53] N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, Open Babel: An open chemical toolbox, *Journal of Cheminformatics*, 3 (2011).
- [54] G. Torrance, https://github.com/gilleain/cdk_signature, (2012)
- [55] C. Steinbeck, Y. Han, S. Kuhn, O. Horlacher, E. Luttmann, E. Willighagen, The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics, *J Chem Inf Comput Sci*, 43 (2003) 493 - 500.
- [56] C. Steinbeck, C. Hoppe, S. Kuhn, M. Floris, R. Guha, E.L. Willighagen, Recent developments of the Chemistry Development Kit (CDK) - An open-source Java library for chemo- and bioinformatics, *Current Pharmaceutical Design*, 12 (2006) 2111-2120.
- [57] S.R. Heller, S.E. Stein, D.V. Tchekhovskoi, InChI: Open access/open source and the IUPAC international chemical identifier, *Abstracts of Papers of the American Chemical Society*, 230 (2005) 60-CINF.
- [58] D. Weininger, SMILES, a chemical language and information-system. 1. Introduction to methodology and encoding rules, *Journal of Chemical Information and Computer Sciences*, 28 (1988) 31-36.
- [59] A.Y. Sokolov, S. Mittal, A.C. Simmonett, H.F. Schaefer, Characterization of the t-Butyl Radical and Its Elusive Anion, *J. Chem. Theory Comput.*, 8 (2012) 4323-4329.
- [60] M.K. Sabbe, M.F. Reyniers, V. Van Speybroeck, M. Waroquier, G.B. Marin, Carbon-centered radical addition and beta-scission reactions: Modeling of activation energies and pre-exponential factors, *Chemphyschem*, 9 (2008) 124-140.
- [61] J.W. Allen, W.H. Green, R.W. Ashcraft, G.J. Beran, C.A. Class, C. Gao, C.F. Goldsmith, M.R. Harper, A. Jalan, G.R. Magoon, D.M. Matheu, S.S. Merchant, J.D. Mo, S. Petway, S. Raman, S. Sharma, J. Song, K.M.V. Geem, J. Wen, R.H. West, A. Wong, H.-W. Wong, P.E. Yelvington, J. Yu, RMG - Reaction Mechanism Generator (2012).
- [62] M.K. Sabbe, F. De Vleeschouwer, M.F. Reyniers, M. Waroquier, G.B. Marin, 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*, 112 (2008) 12235-12251.
- [63] E. Goos, A. Burcat, B. Ruscic, Third Millennium Ideal Gas and Condensed Phase Thermochemical Database for Combustion with updates from Active Thermochemical Tables, ANL-05/20 and TAE 960 Technion-IIT, (2010).
- [64] M.J. McGregor, P.V. Pallai, Clustering of Large Databases of Compounds: Using the MDL "Keys" as Structural Descriptors, *Journal of Chemical Information and Computer Sciences*, 37 (1997) 443-448.

Chapter 7: Thermal decomposition of diethylsulfide

7.1 Abstract

The use of a rule-based automated kinetic model builder, Genesys, is illustrated for the thermal decomposition of diethylsulfide. The chemistry that is considered relevant for the process was translated into three reaction families that were used to iteratively expand the model with molecules and reactions. Rate coefficients of elementary reactions and thermochemical properties of molecules were estimated through group additive methods, with parameters solely derived from high level quantum-chemical calculations. Constraints per reaction family limit the model size and exclude species and reactions that are considered irrelevant. The generated model, consisting of 445 reactions between 66 species, was validated using experimental data and showed that measured and predicted species mole fractions are in good agreement.

Keywords: automatic kinetic model construction; alkylsulfide pyrolysis

7.2 Introduction

Kinetic models of complex chemical processes can easily contain thousands of reactions between hundreds of species. To cope with the level of complexity, the automated construction of such models represents an interesting alternative to the manual derivation of them. Many groups have looked into the possibility of automated construction of kinetic models [1-5]. For historical reasons, these programs are often constrained to species involving only a limited number of elements, such as carbon, and hydrogen, making them less suitable for applications in which heteroelements such as sulfur are involved. The extension of existing programs to new chemical elements, or new types of atoms is not straightforward. For example, differences in electron valency, the number of free electron pairs in heteroelements may induce the creation of structures that are not prevalent for hydrocarbons. Obviously, the presence of heteroelements may lead to new types of reactions that do not occur for hydrocarbons. Often, the incorporated algorithms in the program are tailored for a specific subspace of molecules, and not based on principles that would make them more generally applicable. Finally, the associated chemical data of the program to determine model parameters is not sufficient to adequately predict thermochemistry and kinetics that involve other heteroelements.

A new automated kinetic model construction program, i.e. Genesys, was introduced in Chapters 5-7 of this thesis. Through the use of chemoinformatics libraries, graph theoretic algorithms were applied allowing the construction of kinetic models, and thus regardless of the type of chemical elements involved. The same graph algorithms allowed the use of predictive group contribution methods for the estimation of thermochemical properties and rate coefficients. To illustrate the functionality of Genesys, a kinetic model for the thermal decomposition of diethylsulfide (DES), was constructed. This process is chosen to demonstrate the kinetic model generation procedure for a process containing the heteroelement sulfur. Also, sufficient knowledge of the relevant chemistry and thermochemical and kinetic data for the use in group contribution methods is

available allowing the construction of a quantitative model that can be used in numerical simulations.

Sulfur components are important substances in many industrial processes [6]. Steam cracking feedstocks can contain up to 1000 ppm of sulfur components such as thiols, sulfides and thiophenes [7]. They can influence the product distribution and are thought to play a role in the inhibition of CO formation and in coke formation [8-12]. In order to improve understanding of the role of sulfur components in steam cracking, the decomposition of several sulfur components and the pyrolysis of several mixtures containing hydrocarbons and sulfur components have been studied experimentally [12-17]. An alternative to conducting experiments is the construction of a kinetic model to gain a better understanding of the reactive behavior of alkylsulfides. A detailed kinetic model provides a starting point to further elucidate the underlying chemistry by comparing model predictions with experimental data. Recently, it was shown that the use of quantum-chemical methods such as CBS-QB3 can be used for the accurate calculation of thermochemistry and rate coefficients for sulfur components [18]. Therefore, the incorporation of model parameters derived from *ab initio* calculations rather than from fitting to experiments, becomes a viable option for kinetic model construction.

In this work, the thermal decomposition of diethylsulfide (DES) is studied. A kinetic model was automatically constructed using Genesys with reaction rate coefficients based on ΔGAV° s from *ab initio* calculations from literature. Simulations were carried out of the experimental dataset of Zheng et al. [19] and the resulting model predictions were compared to the measured product mole fractions to validate the generated kinetic model. An existing kinetic model for DES pyrolysis [18], automatically constructed by the Reaction Mechanism Generator [20] (RMG), is used for comparison. Differences in the main decomposition pathways between the current model and the RMG generated model are highlighted.

7.3 Methodology

7.3.1 Model generation

The aim of a kinetic model generation program is to generate an accurate and detailed kinetic model, while minimizing the number of unimportant species and reactions. Genesys starts from two fundamental principles to construct kinetic models, i.e. a thorough understanding of the relevant chemistry of the process, and the availability of thermochemical data for the calculation of the species properties and the reaction rate coefficients. Guided by those principles, a number of reaction family templates were created that are necessary to automatically construct a kinetic model by an iterative procedure. First, a starting pool of molecules that are known as important intermediates species in the thermal decomposition of DES based on literature data is provided by the user. Next, the available reaction family templates are repeatedly applied, and the eligibility of the pool of initial species to react according to one of the reaction families, is verified, cf. Figure 7-1. New product molecules arise from the application of the matched reaction family and serve as reactants for the next iteration in the model building process. Finally, the expansion of the kinetic model terminates when the molecules in the model no longer give rise to new product structures.

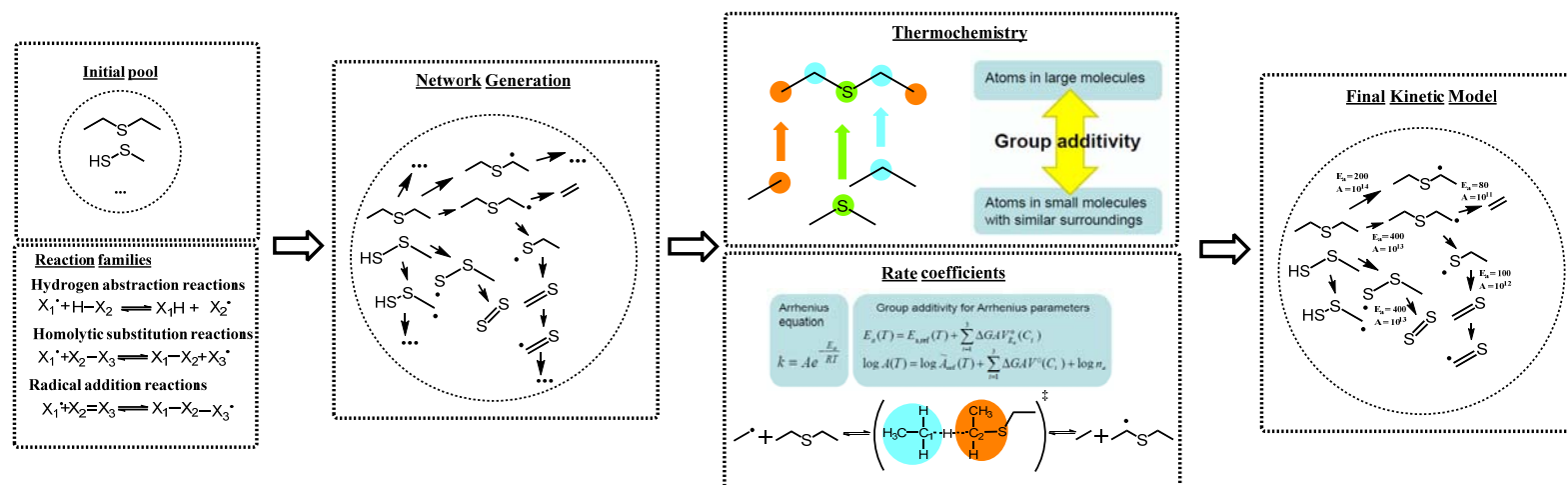


Figure 7-1: The methodology used by Genesys to create kinetic models.

Recently, the Reaction Mechanism Generator (RMG) [20] was extended to support sulfur chemistry, and used the same thermochemical and kinetic data of Vandeputte et al. [18] to generate kinetic model for DES. RMG generates kinetic models in an entirely different way than Genesys. RMG uses a rate-based algorithm [21] to construct the kinetic model, as opposed to the rule-based approach of Genesys. Rate-based model enlargement implies that the kinetic model is successively expanded by the addition of kinetically significant species to the kinetic model, as was described in Section 1.1.1.3 of Chapter 1 of this thesis. A more detailed description of RMG can also be found elsewhere, e.g. refs. [22-24].

Table 7-1 shows the considered reaction families for the model construction of DES pyrolysis using Genesys and using RMG [20]. In Genesys, three reaction families are used to systematically enlarge the reaction network: intermolecular H-abstraction reactions, intermolecular radical addition reactions, and intermolecular homolytic substitution reactions.

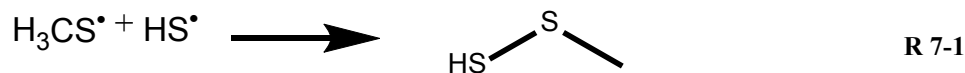
Table 7-1: Reaction families considered in the current model generated using Genesys and the RMG model.

Reaction family	Current model	RMG model
intermolecular H-abstraction reactions	✓	✓
intermolecular radical addition / β -scission reactions	✓	✓
intermolecular homolytic substitution	✓	✓
homolytic bond-scission / intermolecular radical recombination reactions	✗	✓
intramolecular H-abstraction reactions	✗	✓

In principle, other reaction families, such as homolytic bond-scission reactions, could also be included to enlarge the reaction network. However, the inclusion of an additional reaction family in the network generation scheme also requires the systematic assignment of rate coefficients to all of the generated reactions resulting from that reaction family. Arrhenius parameters of reactions originating from the reaction families, are obtained using group contribution methods [25], as will be further explained in detail in Section 8.1.3.2 of this chapter. Unfortunately, these types of group contribution methods are not available for every type of reaction family relevant

for DES decomposition. For reactions such as homolytic bond-scission reactions, intermolecular radical recombination reactions and intramolecular H-abstraction reactions, Arrhenius parameters cannot be determined using the group contribution methodology because there are no group additive values available for these reaction families. In contrast to the model built by Genesys, the RMG model systematically includes intermolecular radical recombination reactions and intramolecular H-abstraction reactions. The rate coefficients for these reaction families, however, are not derived from group additivity, but originate from hierarchical, tree-structured databases in RMG. The reader is referred to Section 1.1.3 of Chapter 1 for more information on the RMG databases.

Important reactions that do not belong to one of the three considered reaction families used by Genesys (cf. Table 7-1) are added after the network generation procedure. For example, disulfanylmethane (H_3CSSH) can be formed through a recombination reaction of methylthio radicals ($\text{H}_3\text{CS}^\bullet$) and thiyl radicals (HS^\bullet), cf. R 7-1.



Because radical recombination reactions are not included as a reaction family, this reaction is manually added, once the network generation procedure is finished. However, since H_3CSSH can only be formed through this reaction during the thermal decomposition of DES, further decomposition reactions of H_3CSSH are not included, unless H_3CSSH is added to the initial pool of molecules, next to the reactant DES, cf. Figure 7-1.

Genesys allows the inclusion of so-called constraints for each reaction family. They constrain the application domain of a reaction family to avoid the addition of irrelevant species and reactions to the model, control the model size, and ultimately prevent the model construction procedure from continuing indefinitely. For example, the experiments by Zheng et al. [19, 26] did not report molecules larger than DES. Therefore, reaction families that can lead to larger chains of carbon or sulfur atoms were constrained so that they cannot lead to molecules with more than five carbon or sulfur atoms. Species containing consecutive double bonds such as 1,2-butadiene were not observed in the experiments neither, with the exception of CS_2 . Therefore, β -scission reactions that could create such molecules were restricted to only C-H bonds that do not lead to

species containing consecutive double bonds. More information on the functionality of the reaction family constraints of Genesys can be found in Chapter 5 of this thesis.

7.3.1.1 Thermochemistry

Two types of databases were used to estimate the thermochemistry of species in the reaction network. First of all, Genesys disposes of a large database of thermochemistry values for hydrocarbon molecules, radicals and sulfur containing components. These values primarily originate from *ab initio* studies such as refs. [27-29] or from aggregate databases such as the Third Millennium Ideal Gas and Condensed Phase Thermochemical Database for Combustion [30]. Secondly, Benson's group additivity method was used to estimate the enthalpy of formation, entropy and heat capacity for species that could not be found in one of these species thermochemistry databases. In the Benson group additivity framework [31] each of the polyvalent atoms in the molecule are interpreted as a group surrounded by a number of ligands. The sum of the group additive contributions ($GAV(X_i)$), along with contributions accounting for interactions between non-adjacent groups (NNI_j), yield estimates of the enthalpy of formation, entropy and heat capacity of the entire molecule.

$$\begin{cases} \Delta_f H^\circ \\ S^\circ \\ C_p^\circ \end{cases} = \sum_{i=1}^n GAV(X_i) + \sum_{j=1}^m NNI_j \quad \text{Eq. 7-1}$$

With $GAV(X_i)$ the group additive value of the group centered on the atom X_i and where the summation index i goes from 1 to n , the total number of groups the component consist of. For the estimation of thermochemistry of radicals, the hydrogen-atom bond increment (HBI) method of Lay et al. [32] was used. This method determines the thermochemistry of radical species R^\bullet by adding an increment to the parent molecule RH , which has the same structure of the radical except that an extra hydrogen atom is attached to the atom that contains the radical site. In this work, a consistent set of *ab initio* derived group additive values, and hydrogen bond increments was used to estimate the thermochemistry of the species in the reaction network [27-29].

7.3.1.2 Kinetics

The kinetics of the reaction families were estimated using a group additive method for Arrhenius parameters. The group additive method, as developed by Saeys et al. [33] for activation energies,

and extended by others for both pre-exponential factors and activation energies [34-39], allows calculating the pre-exponential factor and activation energy of a reaction as in Eq. 7-2 and Eq. 7-3.

$$E_a = E_{a,ref} + \sum_i \Delta GAV_{E_a}^0(X_i) + \Delta E_{a,res}^0 \quad \text{Eq. 7-2}$$

$$\log_{10} A = \log_{10} \tilde{A}_{ref} + \sum_i \Delta GAV_{\tilde{A}}^0(X_i) + \log_{10} \Delta \tilde{A}_{res}^0 + \log_{10} n_e \quad \text{Eq. 7-3}$$

$E_{a,ref}$ and \tilde{A}_{ref} are the activation energy and pre-exponential factor of a reference reaction and ΔGAV^0 the group additive values accounting for the contribution of structural differences between the target and reference reaction. $\Delta E_{a,res}^0$ and $\Delta \tilde{A}_{res}^0$ are the correction terms accounting for resonance in the transition state while the number of single events, n_e , accounts for the reaction path degeneracy of the target reaction. X_i refers to the central atoms in the transition state that undergo changes in terms of bonding, cf. Figure 7-2. For example, H-abstraction reactions have two central atoms that surround the hydrogen atom that is being exchanged. The ligands of the central atom define the group additive value while atoms further away from the central atoms have proven to be of no significant influence for the types of reactions studied in this work [35-37, 40].

Group additive methods offer several advantages, both for thermochemistry and kinetics. They provide a fast and predictive means to obtain rate coefficients for which *ab initio* values are unavailable. Group additive estimates are reliable if they are used in the application domain they were devised for. The used ΔGAV^0 s are also almost temperature independent, since the activation energy of the reference reaction captures most of the temperature dependency.

Genesys contains databases of ΔGAV^0 s for every reaction family considered during the automated construction of the kinetic model of the thermal decomposition of diethylsulfide. Each entry in the database comprises a description of the atomic environment of the central atom, using the SMARTS [41] pattern language, along with the numerical values for contributions to the Arrhenius parameters. The pattern description is converted into a graph representation and

used in a substructure searching algorithm provided by the Chemistry Development Kit (CDK) [42, 43], the chemoinformatics library that is used in Genesys. For each of the reactive centers in an elementary reaction, the database with ΔGAV° s is iterated over until a match is found between the SMARTS string in the database and the atom in the transition state, cf. Figure 7-2. A similar approach was followed for resonance corrections (RESs), although the SMARTS description now match a cluster of atoms rather than a single central atom. ΔGAV° s and RESs used in this work, originate from refs. [34-39] and have shown to reproduce theoretical rate coefficients very well. Values at 1000K were used.

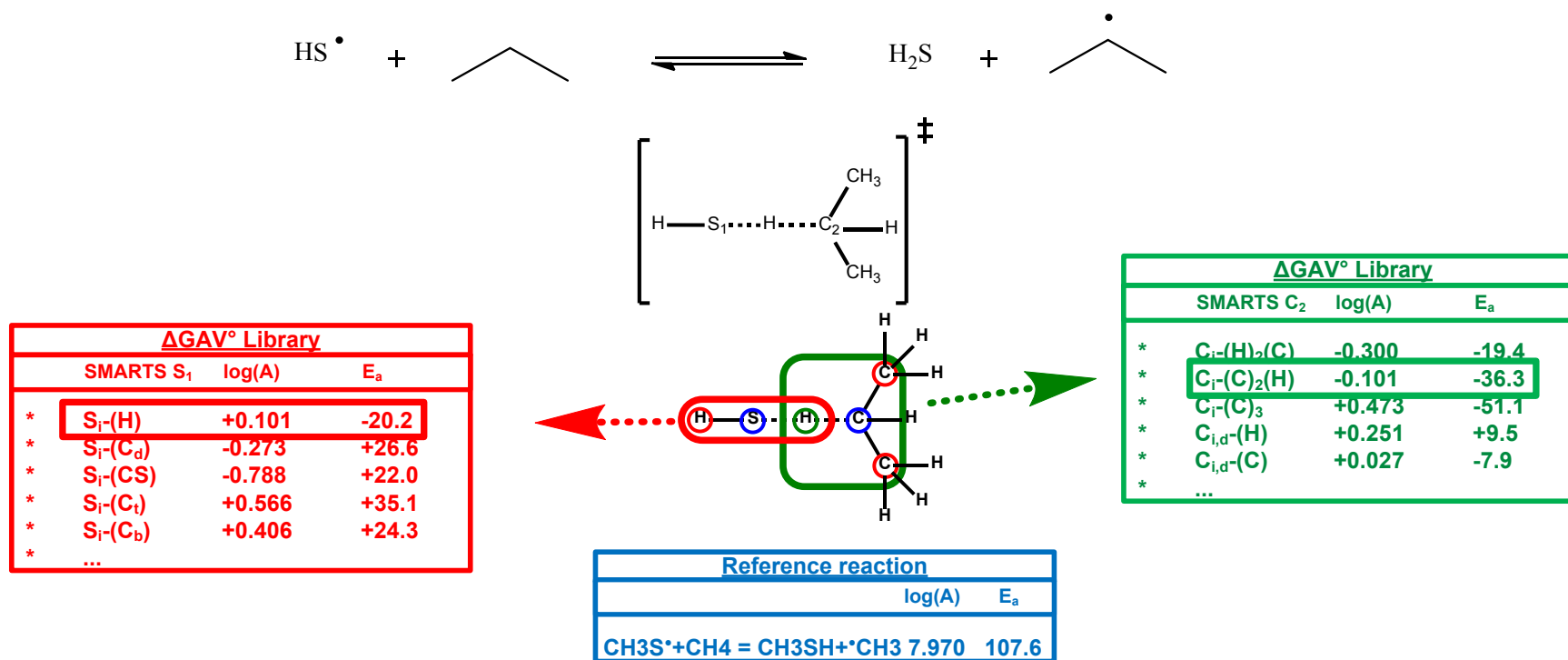


Figure 7-2: Assignment of ΔGAV° s for the H-abstraction reaction by thyl from the secondary hydrogen of propane. Databases for both reactive centers (S_1 & C_2) are matched against the atoms in the transition state. The reference reaction for this reaction family is $\text{CH}_3\text{S}^\bullet + \text{CH}_4 \leftrightarrow \text{CH}_3\text{SH} + \text{CH}_3^\bullet$

Tunneling contribution can have a profound effect on the calculated rate coefficients, for example for H-abstraction reactions at lower temperatures [44-46]. Since the group additive values used in this work explicitly excluded tunneling contributions, they need to be accounted for in a separate way. Several authors, e.g. [35, 40] illustrated that a polynomial expression using only the activation energy of the reaction and the temperature as the main factors can be employed for the calculation of tunneling coefficients, e.g. Eq. 5-8 in Section 5.3.2 of Chapter 5. This type of expression was used in Genesys to evaluate tunneling coefficients on-the-fly, using a temperature of 1000K. The reaction enthalpy at the provided temperature was calculated using the thermochemistry of reactants and products and subsequently used to determine the exothermic direction of a particular reaction. At the temperatures relevant for thermal decomposition of DES, values of the tunneling coefficient for the reactions of the considered reaction families are situated between 1.0 and 1.4.

Besides the reaction families mentioned in Table 7-1, 30 reactions were included that were deemed important for alkylsulfide pyrolysis, cf. Table 7-2. These reactions have rate coefficients for which predictive kinetic expressions such as group additivity were unavailable, and therefore needed to be included individually. Rate coefficients were taken from a theoretical study on the thermal decomposition of dimethyl disulfide [47].

Table 7-2: Arrhenius parameters for the reactions that were manually added to the model. Units: $\text{m}^3 \text{mol}^{-1} \text{s}^{-1}$ for A, kJ mol^{-1} for E_a . Arrhenius parameters originate from [47].

Reaction	$\text{Log}_{10}(\text{A})$	n	E_a	Reaction	$\text{Log}_{10}(\text{A})$	n	E_a	Reaction	$\text{Log}_{10}(\text{A})$	n	E_a
$\text{CH}_3\text{SSCH}_3 = \bullet\text{CH}_2\text{S} + \text{CH}_3\text{S}\bullet$	8.5	0	73	$\text{CH}_3\text{CH}_2\text{S}\bullet = \text{CH}_2\text{S} + \bullet\text{CH}_3$	9.6	0	174	$\text{H}\bullet + \text{CH}_3\text{S}\bullet\text{C}_2\text{H}_5 = \text{H}_2 + \text{CH}_3\text{SCH}_2\bullet\text{CH}_2$	-6.6	4.34	9
$\text{SCHCH}_2\text{S}\bullet = \text{CH}_2\text{S} + \bullet\text{CHS}$	8.8	0	154	$\text{C}_2\text{H}_4\text{S} + \text{H}\bullet = \text{CH}_3\text{CH}_2\text{S}\bullet$	2.7	1.7	2	$\text{H}\bullet + \text{CH}_3\text{S}\bullet\text{C}_2\text{H}_5 = \text{H}_2 + \text{CH}_3\text{S}\bullet\text{CHCH}_3$	-6.1	4.34	24
$\text{CH}_3\text{SCH}_2\bullet\text{CH}_2 + \text{H}\bullet = \text{CH}_3\text{SC}_2\text{H}_5$	6.9	0.23	-2	$\text{C}_2\text{H}_4\text{S} + \text{C}_2\text{H}_5 = \text{CH}_3\text{CH}_2\text{SC}\bullet\text{HCH}_3$	-3.1	2.98	0.1	$\text{H}\bullet + \text{DES} = \text{CH}_3\text{CH}_2\text{SH} + \text{C}_2\text{H}_5$	2.6	1.5	14
$\text{CH}_3\text{S}\bullet\text{CHCH}_3 + \text{H}\bullet = \text{CH}_3\text{SC}_2\text{H}_5$	7.7	0.16	0	$\text{C}_2\text{H}_4 + \text{CH}_3\text{CH}_2\text{S}\bullet = \text{CH}_3\text{CH}_2\text{SCH}_2\text{C}\bullet\text{H}_2$	-3.8	3.1	5	$\text{CH}_2\text{S} + \text{H}\bullet = \bullet\text{CHS} + \text{H}_2$	-1	2.72	15
$\text{C}_3\text{H}_7\text{S}\bullet + \text{H}\bullet = \text{CH}_3\text{SC}_2\text{H}_5$	5.7	0.65	-2	$\text{H}\bullet + \text{CH}_2\text{S} = \text{CH}_3\text{S}\bullet$	3	1.7	-3	$\text{H}\bullet + \text{CH}_3\text{S}\bullet\text{C}_2\text{H}_5 = \text{H}_2 + \text{C}_3\text{H}_7\text{S}\bullet$	-6.8	4.34	15
$\text{HS}\bullet + \bullet\text{CHS} = \text{CH}_2\text{S}_2$	6.9	0.23	-2	$\text{H}\bullet + \text{CH}_2\text{S} = \bullet\text{CH}_2\text{SH}$	2.7	1.7	3	$\text{CH}_3\text{CH}_2\text{S}\bullet = \bullet\text{CH}_2\text{CH}_2\text{SH}$	-4.1	3.04	49
$\text{C}_2\text{H}_5 + \text{CH}_3\text{S}\bullet = \text{CH}_3\text{SC}_2\text{H}_5$	2.7	1.54	-6	$\text{C}_2\text{H}_4 + \text{CH}_3\text{S}\bullet = \text{CH}_3\text{S}\bullet\text{CHCH}_3$	-3.8	3.1	5	$\text{H}\bullet + \text{CH}_3\text{S}\bullet\text{C}_2\text{H}_5 = \text{H}_2 + \text{CH}_3\text{SCH}_2\bullet\text{CH}_2$	-6.6	4.34	9
$\text{CH}_3\text{CH}_2\text{S}\bullet + \bullet\text{CH}_3 = \text{CH}_3\text{SC}_2\text{H}_5$	2	1.54	-6	$\text{CH}_2\text{S} + \text{C}_2\text{H}_5 = \text{C}_3\text{H}_7\text{S}\bullet$	-3.8	3	-4	$\text{H}\bullet + \text{CH}_3\text{S}\bullet\text{C}_2\text{H}_5 = \text{H}_2 + \text{CH}_3\text{S}\bullet\text{CHCH}_3$	-6.1	4.34	24
$\text{DES} = \text{C}_2\text{H}_5 + \text{CH}_3\text{CH}_2\text{S}\bullet$	8.1	0	250	$\text{C}_2\text{H}_4\text{S} + \text{H}\bullet = \text{CH}_3\text{CH}_2\text{S}\bullet$	2.7	1.7	3	$\text{H}\bullet + \text{DES} = \text{H}_2 + \text{CH}_3\text{CH}_2\text{SCH}_2\text{C}\bullet\text{H}_2$	-5.8	4.34	24
$\text{H}\bullet + \text{DES} = \text{H}_2 + \text{CH}_3\text{CH}_2\text{SC}\bullet\text{HCH}_3$	-6.3	4.34	9	$\text{C}_2\text{H}_4\text{S} + \bullet\text{CH}_3 = \text{CH}_3\text{SCH}_2\bullet\text{CH}_2$	-2.3	2.8	7	$\text{CH}_3 + \text{CH}_3\text{S}\bullet\text{C}_2\text{H}_5 = \text{C}_2\text{H}_5 + \text{CH}_3\text{CH}_2\text{SH}$	-2.1	2.69	45

7.3.2 Experimental data and reactor modeling

The experimental data reported by Zheng et al.[19, 26], was measured in a tubular reactor of 1m long with an internal diameter of 45 mm. The experimental conditions are described in Table 7-3. The mole fractions were measured at four distinct positions in the reactor by withdrawing samples through a quartz probe.

Table 7-3: Experimental conditions in the experiments by Zheng et al. [19, 26].

Temperature (K)	903-1013
Pressure (10^5 Pa)	1
DES conversion (%)	10-80
Residence time (s) [‡]	0.07-0.1
DES inlet mole fraction (ppm)	150
N ₂ mol flow rate (10^{-3} mol s ⁻¹)	0.54

[‡] Residence time based on the total molar flow rate at the entrance of the reactor, at the mean temperature inside the reactor.

The generated kinetic model was used to predict product composition by conducting reactor simulations with the Chemkin software version 4.1 [48]. The plug flow reactor model with isothermal and isobaric boundary conditions was used to model the experimental reactor setup. The temperature was set to the reported mean temperature over the reaction section [19, 26], because temperature profiles in axial direction were not reported.

7.4 Results and discussion

The different approaches of RMG and Genesys lead to kinetic models that are significantly different from one another. The model constructed in this work consists of only 445 reactions between 66 species while the larger RMG model consists of 3125 reactions between 135 species. The experimental data from Zheng et al. [19] was used to validate the predictions of both models. Figure 7-3 shows the mole fractions as a function of reactor residence time of DES, and ethene, the main product, for the model predictions of the Genesys and RMG model versus the experimental mole fractions of Zheng et al. [19]. Model predictions for DES are very similar for

both models and agree well with experiments at all four temperatures, although the Genesys model tends to slightly underestimate ethene.

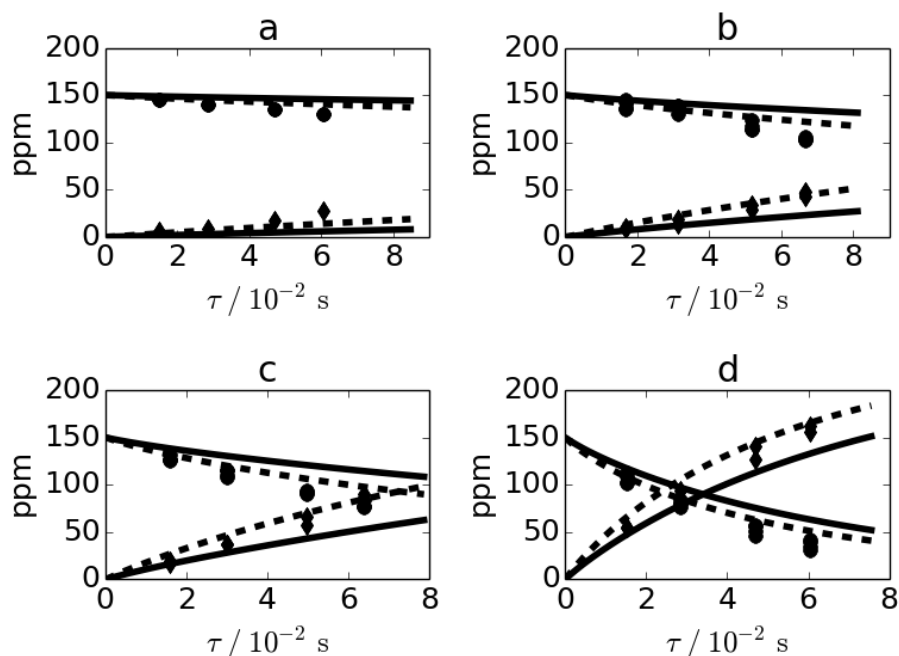
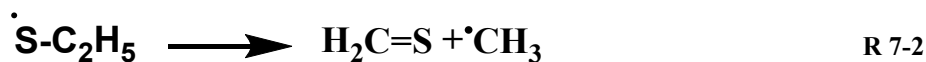


Figure 7-3: Mole fractions as a function of residence time at a) 903K, b) 943K, c) 973K, d) 1013K. Model predictions of DES (blue) and ethene (red) by Genesys model (—) and RMG model (- -) versus experiments from Zheng et al. [19]: ●: DES, ♦: ethene. Conditions: initial DES mole fraction in $\text{N}_2 = 150 \text{ ppm}$, $P = 1 \cdot 10^5 \text{ Pa}$. Figure 7-4 shows the mole fractions of methane and ethane as a function of residence time. Methane mole fractions predicted by the Genesys model are overpredicted while ethane is underpredicted compared to the experimental observations. Predictions of the RMG model agree with the model by Genesys, except at the highest temperature where significant differences were observed between both models. Mole fractions of methane and ethane are correlated since both species are formed through methyl radicals. Small errors in the rate constants of H-abstraction reactions by methyl or a too slow methyl-methyl recombination reaction may explain the observed deviations. Higher mole fractions of methyl in the Genesys model compared to the RMG model can explain the differences in model predictions. Methyl is primarily formed through the β -scission reaction of $\bullet\text{S-C}_2\text{H}_5$ forming thioformaldehyde ($\text{H}_2\text{C}=\text{S}$).



The reaction rate of this reaction is significantly higher in the Genesys model, compared to the model by RMG. Mole fractions of the thioformaldehyde are also slightly higher in the model generated by Genesys.

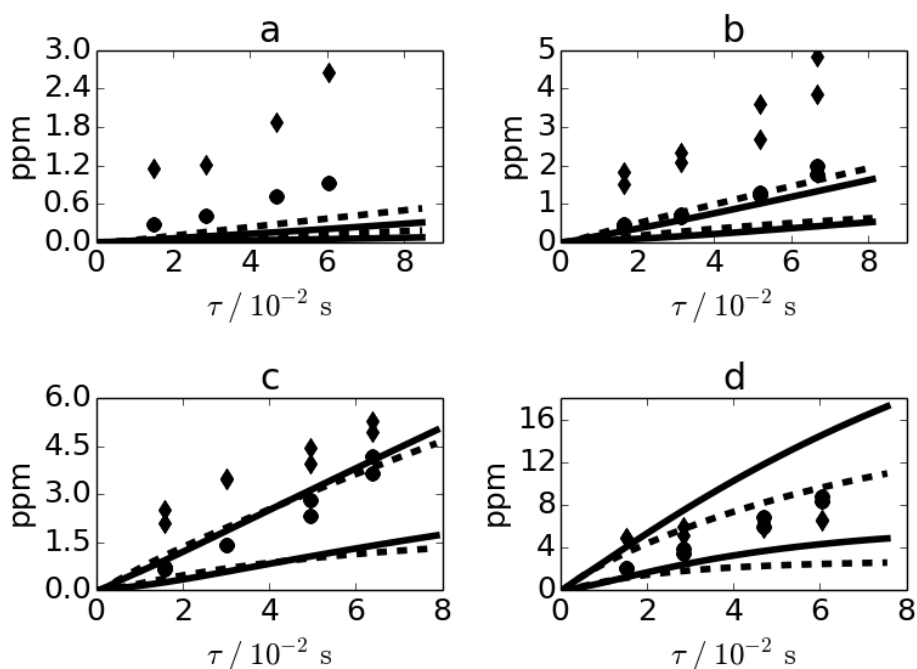


Figure 7-4: Mole fractions as a function of residence time at a) 903K, b) 943K, c) 973K, d) 1013K. Model predictions of methane (blue) and ethane (red) by Genesys model (—) and RMG model (- -) versus experiments from Zheng et al. [19]: ●: methane, ◆: ethane. Conditions: initial DES mole fraction in $\text{N}_2 = 150 \text{ ppm}$, $P = 1 \cdot 10^5 \text{ Pa}$.

Figure 7-5 shows the mole fractions of ethyne and CS_2 as a function of residence time. Predicted CS_2 mole fractions by both models of the latter agree well with experiments. Ethyne mole fractions are slightly underestimated and can be correlated with the small underestimations of ethene mole fractions. CS_2 is formed by the decomposition of dithioformic acid ($\text{S}=\text{CH}-\text{SH}$) while ethyne is primarily formed via a β -scission reaction of vinyl radical that is formed by H-abstraction reactions by radicals from ethene.

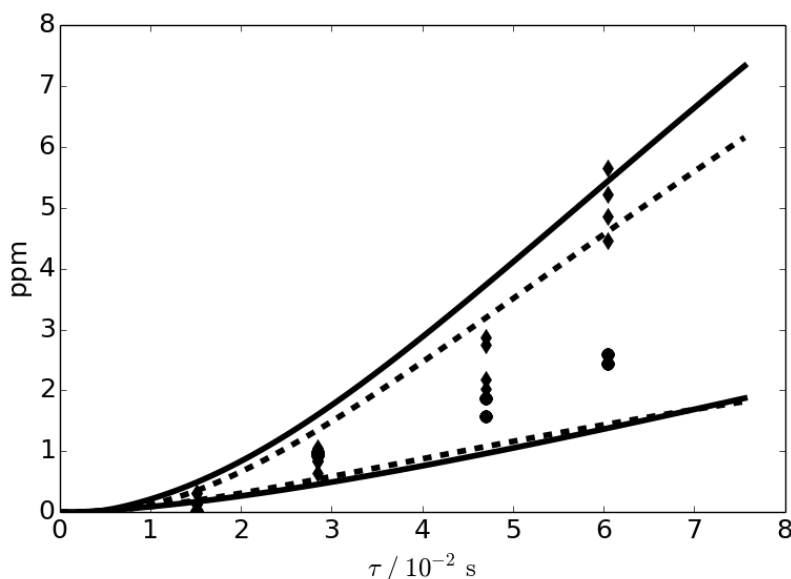


Figure 7-5: Mole fractions as a function of residence time. Model predictions of C_2H_2 (blue) and CS_2 (red) by Genesys model (—) and RMG model (- -) versus experiments from Zheng et al. [19]: \bullet : C_2H_2 , \blacklozenge : CS_2 . Conditions: initial DES mole fraction in $\text{N}_2 = 150 \text{ ppm}$, $T = 1013\text{K}$, $P = 1 \cdot 10^5 \text{ Pa}$.

Only a small number of products were quantified in the experiments by Zheng et al. [19]. Therefore, Figure 7-6 shows the mole fractions of important decomposition products as a function of residence time predicted by both models, without experimental data that can be used for validation. New experimental data in which these species can be quantified is required to further validate the currently existing models of DES pyrolysis.

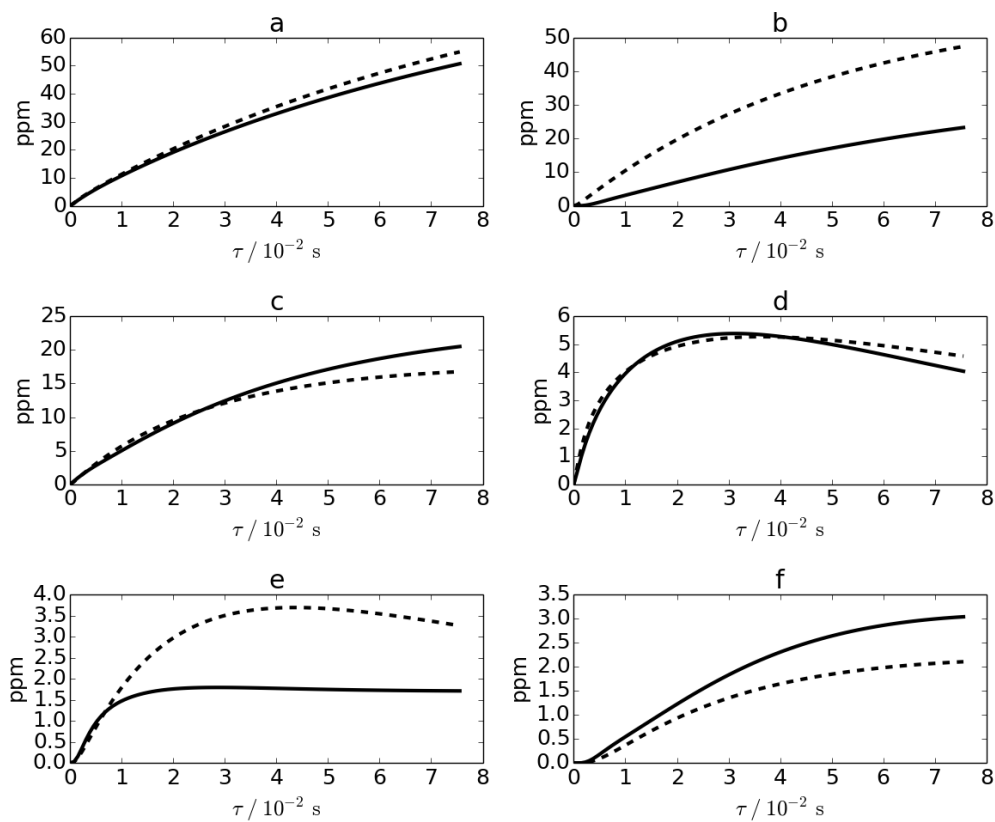


Figure 7-6: Mole fractions as a function of residence time of a) H_2 , b) H_2S , c) $H_2C=S$, d) $H_3C-CH=S$, e) CH_3SH , f) $S=CH-SH$. Model predictions by Genesys model (—) and RMG model (---). Conditions: initial DES mole fraction in $N_2 = 150$ ppm, $T = 1013$ K, $P = 1 \times 10^5$ Pa.

Predicted mole fractions of H_2 agree well between both models. H_2 is formed by H atoms subsequently abstract H atoms from molecules such as DES to form H_2 . H atoms primarily originate from the β -scission reactions of ethyl which also produces ethene. More significant deviations exist for hydrogensulfide. Since the latter is primarily formed through H-abstraction reactions by thiyl ($\bullet SH$), the higher mole fractions of H_2S can be attributed to the higher number of H-abstraction reactions included in the RMG model compared to the Genesys model. Mole fractions of thioformaldehyde ($S=CH_2$) and dithioformic acid agree well between both models while slightly higher deviations exist for minor products such as methanethiol (CH_3SH), and ethanethial ($CH_3CH=S$).

Comparison of the model predictions of the RMG and Genesys model with the experimental data indicated that the model built by Genesys is capable of accurately predicting the experimental data, with very similar results compared to the much larger RMG model. Nevertheless, the RMG generated model shows a slightly better agreement for the species that were quantified in the experiments of Zheng et al. [19]. The differences between both models were also observed when comparing mole fractions of components that were not measured experimentally. A number of reasons can be cited to explain the differences in model predictions.

First of all, the number of species and reactions differs greatly between both models. The number of reactions belonging to a particular reaction family for both models is reported in Table 7-4.

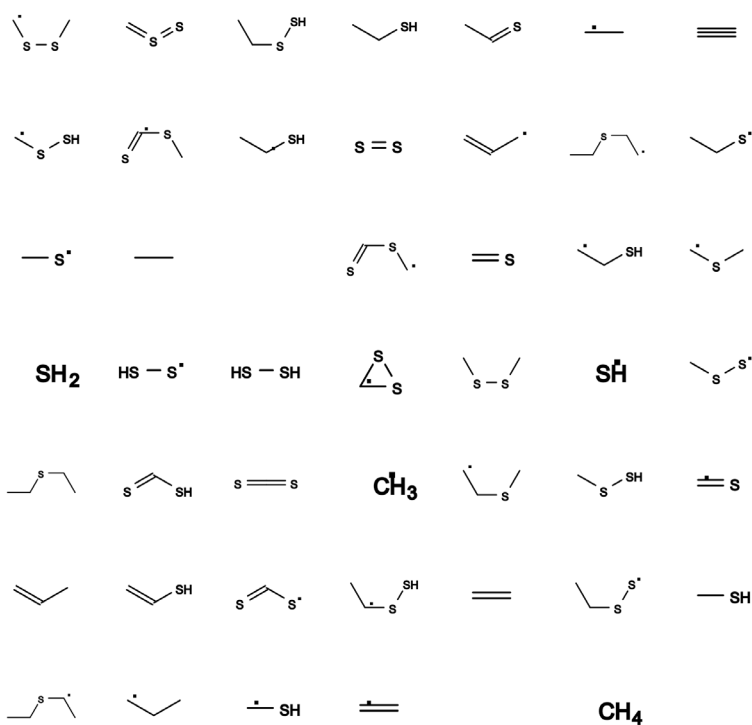
Table 7-4: The number of reactions for each reaction family for the Genesys and RMG generated kinetic model.

	Genesys	RMG
Number of species	66	142
Number of species that only appear in one model	18	91
Number of reactions	445	3125
Reactions / species	6.6	23.1
H-Abstraction reactions	395	2590
Addition/β-scission reactions	22	117
Scission/recombination reactions	9	164
Substitution reactions	6	254

The presence of reaction family constraints in Genesys significantly reduces the number of species but also the number reactions per species. RMG does not use reaction family constraints; as a result, all possible reactions between a species that is being added to the model and the species that were already part of the model are included in the model. Table 7-5 shows the common and unique species of the two models. The Genesys model includes 19 branched alkylthiols and alkyl species that are not present in the RMG model. Many cyclic species appear in the RMG model and are formed through cyclization reaction families such as the intramolecular addition. Cyclization reaction families were not used for the construction of the model by Genesys and therefore explain the absence of cyclic species in the model. The RMG

model also contains many biradical species, disulfides and disulfide derived radicals, whereas the formation of these species was systematically avoided in the Genesys model by the use of reaction family constraints.

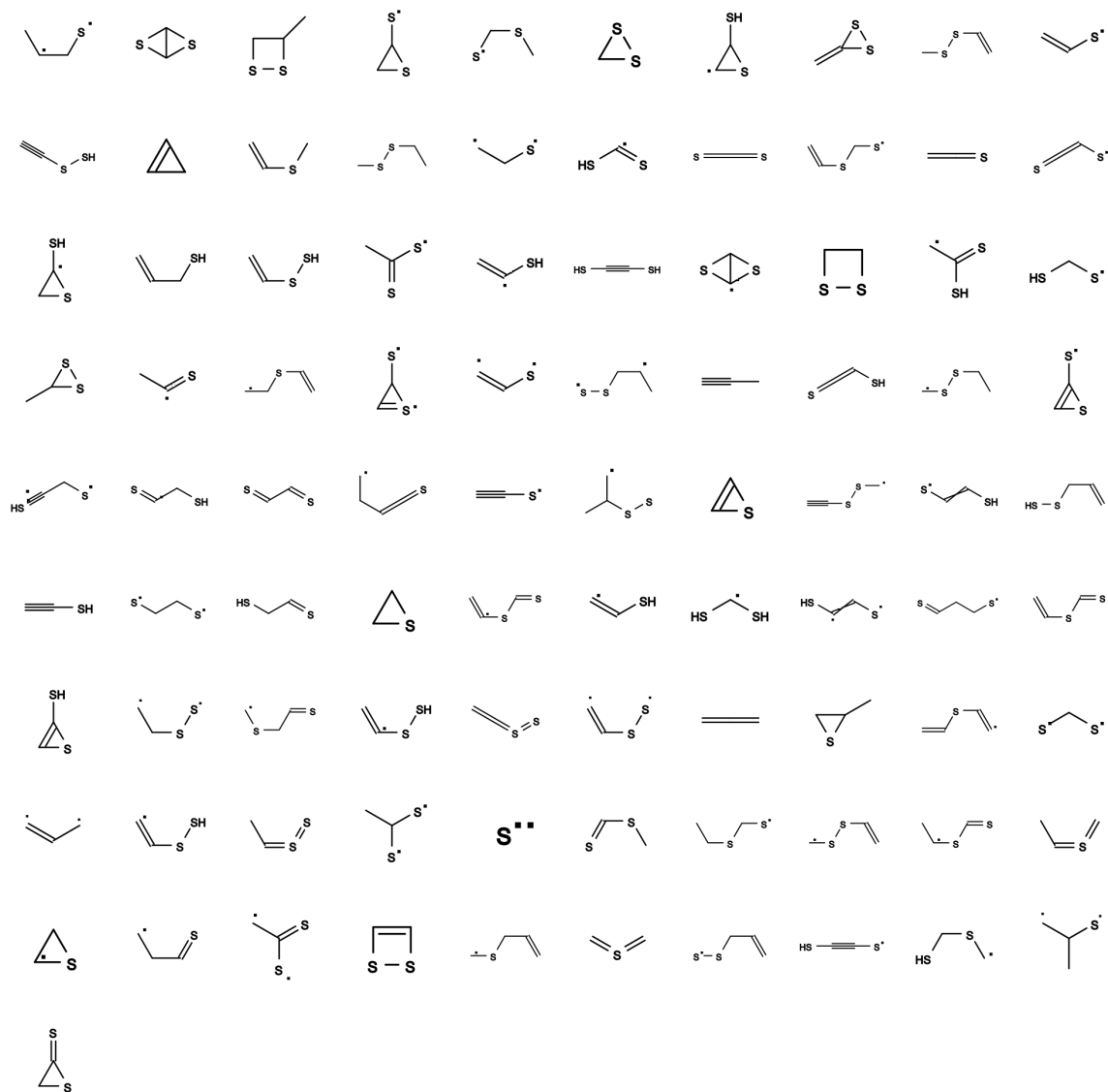
Species that are common to both models



The image displays 15 skeletal chemical structures arranged in a 3x5 grid. The structures are as follows:

- Row 1:
 - 1. A skeletal structure of a thiol, $\text{CH}_3\text{CH}_2\text{S}^\bullet$, with a radical dot on the carbon atom adjacent to the sulfur.
 - 2. A skeletal structure of an alkene, $\text{CH}_2=\text{CHCH}_3$.
 - 3. A skeletal structure of an alkane, $\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_3$, with a radical dot on the second carbon atom.
 - 4. A skeletal structure of a thiol, $\text{CH}_3\text{CH}_2\text{SH}$, with a radical dot on the carbon atom adjacent to the sulfur.
 - 5. A skeletal structure of an alkane, $\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_3$, with a radical dot on the second carbon atom.
- Row 2:
 - 1. A skeletal structure of an alkane, $\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_3$.
 - 2. A skeletal structure of an alkene, $\text{CH}_2=\text{CHCH}_3$, with a radical dot on the carbon atom adjacent to the double bond.
 - 3. A skeletal structure of a thiol, $\text{CH}_3\text{CH}_2\text{S}^\bullet$, with a radical dot on the carbon atom adjacent to the sulfur.
 - 4. A skeletal structure of an alkene, $\text{CH}_2=\text{CHCH}_3$.
 - 5. A skeletal structure of an alkane, $\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_3$, with a radical dot on the second carbon atom.
- Row 3:
 - 1. A skeletal structure of an alkene, $\text{CH}_2=\text{CHCH}_3$, with a radical dot on the carbon atom adjacent to the double bond.
 - 2. A skeletal structure of an alkane, $\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_3$.
 - 3. A skeletal structure of an alkane, $\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_3$, with a radical dot on the second carbon atom.
 - 4. A skeletal structure of an alkene, $\text{CH}_2=\text{CHCH}_3$.
 - 5. A skeletal structure of a thiol, $\text{CH}_3\text{CH}_2\text{SH}$.

Species that only appear in the RMG model



Secondly, the reaction rates for some of the most important pathways of DES consumption are slightly different. Figure 7-7 shows an overview of the main reaction pathways for the thermal decomposition of DES along with reaction rates for some of the most important pathways in the model relative to the rate of consumption of DES. In the Genesys model more than 50% of DES is consumed via the homolytic C-S scission reaction, while the initiation via H-abstraction reactions and homolytic substitutions account for 38% and 9% respectively. The contribution of the homolytic C-S scission reaction is smaller in the RMG model with respect to the Genesys model, but is replaced by a larger fraction of DES molecules that are consumed by homolytic substitution reactions and H-abstraction reactions. The difference in importance of the various decomposition channels in both models can also be attributed to the use of reaction family constraints for substitution reactions and H-abstraction reactions in Genesys. The RMG model incorporates more reactions belonging to these two reaction families. As a result, a larger fraction of DES is converted via these types of reactions.

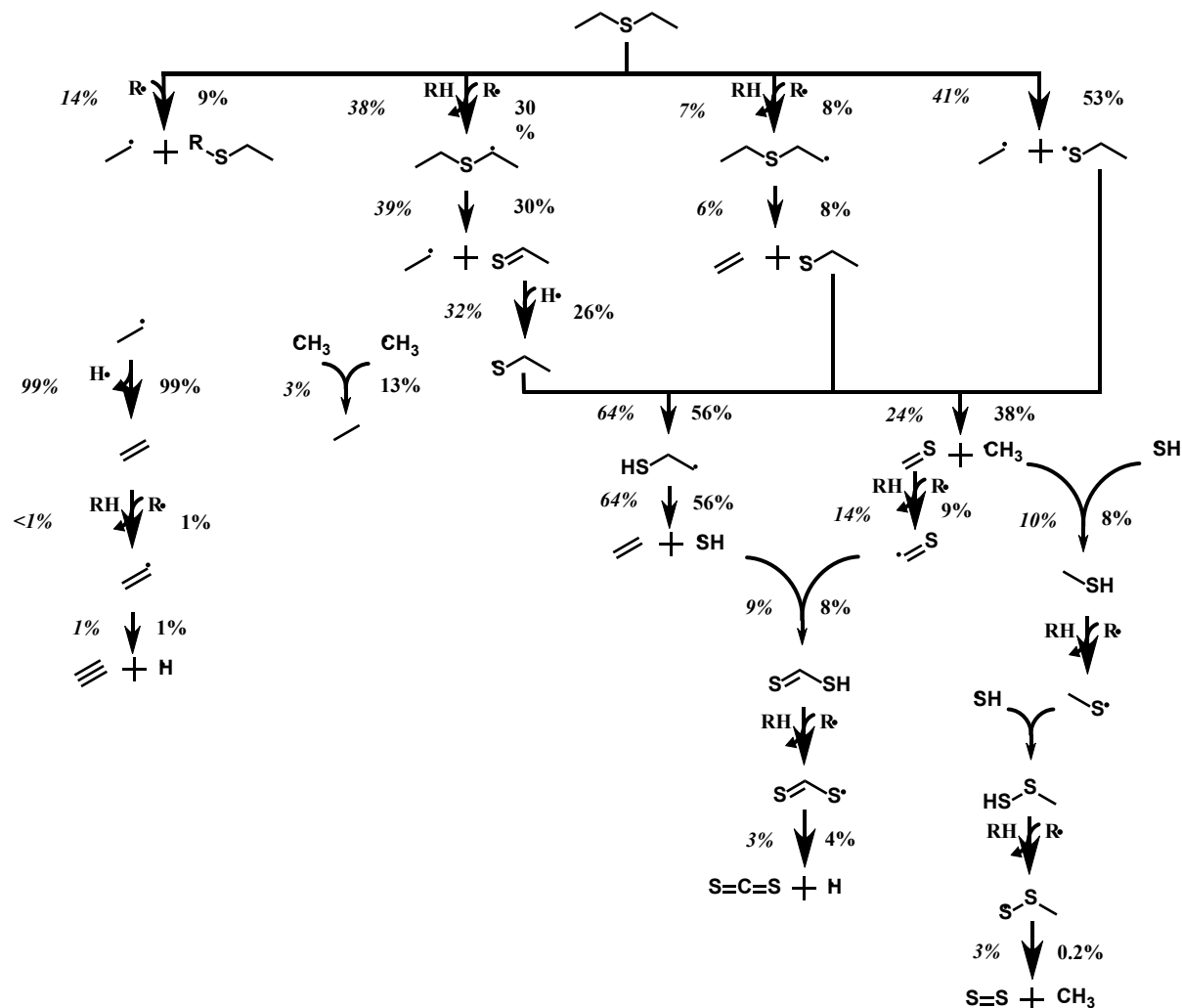


Figure 7-7: Reaction path analysis for the pyrolysis of DES calculated at a temperature of 1013K, a residence time of 20 ms, initial DES mole fraction = 150 ppm. Percentages report the reaction rates relative to the total decomposition rate of DES. Bold percentages originate from the model by Genesys; italic percentages originate from the RMG model.

A third reason that may explain the differences between models is the systematic inclusion of reaction families in the RMG model that are not present in the Genesys model. For example, intramolecular H-abstraction reactions and bimolecular radical recombination reactions are systematically present in the RMG model, cf. Table 7-1. The incorporation of these reaction families in Genesys requires values for the corresponding ΔGAV° s, which are currently unavailable. Future work may therefore focus on the derivation of ΔGAV° s for these reaction families.

7.5 Conclusions

In this chapter Genesys was used to automatically build a detailed microkinetic model for modeling the thermal decomposition of diethylsulfide. Knowledge of the relevant chemistry was translated into a number of reaction families that were used to iteratively expand the kinetic model. Reaction family constraints limit the application domain of the reaction families and control the model construction procedure in only allowing the creation of relevant species and reactions. For the thermal decomposition of diethylsulfide the obtained kinetic model consists only of 445 reactions between 66 species. The calculation of the rate coefficients of elementary reactions and thermochemical properties of molecules was accomplished through predictive group contribution methodologies that drastically reduced the number of parameters needed to construct the kinetic model. Graph-theoretic algorithms and a chemical graph representation of molecules originating from the domain of chemoinformatics provided a flexible way to construct kinetic models regardless of specific nature of the chemical process.

Validation of the obtained model with experimental data showed a good agreement, which was remarkable given that none of the model parameters were fitted to the experiments. The model generated by Genesys is significantly smaller than the model generated by RMG in terms of number of species and elementary reactions, but still showed good agreement between model and experiment. Mutual differences between both models are attributed to the use of reaction family constraints by Genesys that limit the creation of certain species and reactions. The performance of the Genesys model could be improved by incorporating additional reaction families. For that matter, ΔGAV° s need to be derived for these reaction families, which could be done in future work through the use of high quality quantum-chemistry calculations.

Overall, it was shown that the combination of user expertise on the relevant chemistry and a model generation framework that is not tailored to specific applications is a powerful asset for the quantitative investigation of chemical processes.

7.6 References

- [1] V. Warth, F. Battin-Leclerc, R. Fournet, P.A. Glaude, G.M. Côme, G. Scacchi, Computer based generation of reaction mechanisms for gas-phase oxidation, *Computers & Chemistry*, 24 (2000) 541-560.
- [2] M.R. Dente, E. Mathematical modeling of hydrocarbon pyrolysis reactions., in: *Pyrolysis: Theory and industrial practice*, Academic Press, New York, 1983, pp. 133 - 175.
- [3] E.S. Blurock, Detailed Mechanism Generation. 1. Generalized Reactive Properties as Reaction Class Substructures, *Journal of Chemical Information and Computer Sciences*, 44 (2004) 1336-1347.
- [4] S. Rangarajan, A. Bhan, P. Daoutidis, Language-oriented rule-based reaction network generation and analysis: Applications of RING, *Computers & Chemical Engineering*, 46 (2012) 141-152.
- [5] J. Song, S. Raman, J. Yu, C.D. Wijaya, G. Stephanopoulos, W.H. Green, Development of automatic chemical reaction mechanism generation software using object-oriented technology., *Abstracts of Papers of the American Chemical Society*, 226 (2003) U530-U531.
- [6] S. Patai, Z. Rappoport, *The chemistry of sulphur-containing functional groups*, Wiley & Sons, 1993.
- [7] C. Venduvre, R. Ruiz-Guerrero, F. Bertoncini, L. Duval, D. Thiebaut, Comprehensive two-dimensional gas chromatography for detailed characterisation of petroleum products, *Oil & Gas Science and Technology-Revue D IFP Energies Nouvelles*, 62 (2007) 43-55.
- [8] X. Ma, M. Sprague, C. Song, Deep Desulfurization of Gasoline by Selective Adsorption over Nickel-Based Adsorbent for Fuel Cell Applications, *Industrial & Engineering Chemistry Research*, 44 (2005) 5768-5775.
- [9] M. Bajus, J. Baxa, Coke formation during the pyrolysis of hydrocarbons in the presence of sulfur compounds, *Collection of Czechoslovak Chemical Communications*, 50 (1985) 2903-2909.
- [10] M. Bajus, Sulfur Compounds in Hydrocarbon Pyrolysis, *Sulfur reports*, 9 (1989) 25-66.
- [11] G.C. Reyniers, G.F. Froment, F.D. Kopinke, G. Zimmermann, Coke formation in the Thermal Cracking of Hydrocarbons. 4. Modeling of coke formation in naphtha cracking, *Industrial & Engineering Chemistry Research*, 33 (1994) 2584-2590.
- [12] M.-F. Reyniers, G.F. Froment, Influence of Metal-Surface and Sulfur Addition on Coke Deposition in the Thermal-Cracking of Hydrocarbons, *Industrial & Engineering Chemistry Research*, 34 (1995) 773-785.
- [13] J.H. Kolts, Heterogeneous and homogeneous effects of hydrogen sulfide on light-hydrocarbon pyrolysis, *Industrial & Engineering Chemistry Fundamentals*, 25 (1986) 265-269.
- [14] M. Bajus, J. Baxa, P.A. Leclercq, J.A. Rijks, Steam cracking of hydrocarbons. 6. Effect of dibenzyl sulfide and dibenzyl disulfide on reaction kinetics and coking, *Industrial & Engineering Chemistry Product Research and Development*, 22 (1983) 335-343.
- [15] M. Bajus, V. Vesely, J. Baxa, P.A. Leclercq, J.A. Rijks, Steam cracking of hydrocarbons. 5. Effect of thiophene on reaction kinetics and coking, *Industrial & Engineering Chemistry Product Research and Development*, 20 (1981) 741-745.
- [16] J. Wang, M.-F. Reyniers, G.B. Marin, Influence of dimethyl disulfide on coke formation during steam cracking of hydrocarbons, *Industrial & Engineering Chemistry Research*, 46 (2007) 4134-4148.
- [17] I. Dhuyvetter, M.F. Reyniers, G.F. Froment, G.B. Marin, D. Viennet, The influence of dimethyl disulfide on naphtha steam cracking, *Industrial & Engineering Chemistry Research*, 40 (2001) 4353-4362.
- [18] A.G. Vandeputte, *The thermochemistry and Decomposition Mechanism of Organosulfur and Organophosphorus Compounds*; Ph.D Dissertation, Universiteit Gent (2012).
- [19] X. Zheng, E.M. Fisher, F.C. Gouldin, L. Zhu, J.W. Bozzelli, Experimental and computational study of diethyl sulfide pyrolysis and mechanism, *Proceedings of the Combustion Institute*, 32 (2009) 469-476.
- [20] W.H. Green, J.W. Allen, R.W. Ashcraft, G.J. Beran, C.A. Class, C. Gao, C.F. Goldsmith, M.R. Harper, A. Jalan, G.R. Magoon, D.M. Matheu, S.S. Merchant, J.D. Mo, S. Petway, S. Raman, S. Sharma, J. Song, K.M. Van Geem, J. Wen, R.H. West, A. Wong, H.S. Wong, P.E. Yelvington, J. Yu, RMG - Reaction Mechanism Generator v4.0, (2013).
- [21] R.G. Susnow, A.M. Dean, W.H. Green, P. Peczak, L.J. Broadbelt, Rate-based construction of kinetic models for complex systems, *Journal of Physical Chemistry A*, 101 (1997) 3731-3740.
- [22] M.R. Harper, K.M. Van Geem, S.P. Pyl, G.B. Marin, W.H. Green, Comprehensive reaction mechanism for n-butanol pyrolysis and combustion, *Combustion and Flame*, 158 (2011) 16-41.
- [23] M.K. Sabbe, K.M. Van Geem, M.-F. Reyniers, G.B. Marin, First Principle-Based Simulation of Ethane Steam Cracking, *AIChE J.*, 57 (2011) 482-496.
- [24] K.M. Van Geem, M.F. Reyniers, G.B. Marin, J. Song, D.M. Matheu, W.H. Green, Automatic Reaction Network Generation using RMG for Steam Cracking of n-Hexane, *AIChE J.*, 52 (2006) 718-730.

- [25] V. Van Speybroeck, R. Gani, R.J. Meier, The calculation of thermodynamic properties of molecules, *Chem Soc Rev*, 39 (2010) 1764-1779.
- [26] X. Zheng, E.M. Fisher, F.C. Gouldin, J.W. Bozzelli, Pyrolysis and oxidation of ethyl methyl sulfide in a flow reactor, *Combustion and Flame*, 158 (2011) 1049-1058.
- [27] M.K. Sabbe, F. De Vleeschouwer, M.F. Reyniers, M. Waroquier, G.B. Marin, 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*, 112 (2008) 12235-12251.
- [28] M.K. Sabbe, M. Saeys, M.F. Reyniers, G.B. Marin, V. Van Speybroeck, M. Waroquier, Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals, *Journal of Physical Chemistry A*, 109 (2005) 7466-7480.
- [29] A.G. Vandeputte, M.K. Sabbe, M.-F. Reyniers, G.B. Marin, Modeling the Gas-Phase Thermochemistry of Organosulfur Compounds, *Chemistry-a European Journal*, 17 (2011) 7656-7673.
- [30] E. Goos, A. Burcat, B. Ruscic, Third Millennium Ideal Gas and Condensed Phase Thermochemical Database for Combustion with updates from Active Thermochemical Tables, ANL-05/20 and TAE 960 Technion-IIT, (2010).
- [31] S.W. Benson, *Thermochemical Kinetics*, John Wiley & Sons, New York, 1976.
- [32] T.H. Lay, J.W. Bozzelli, A.M. Dean, E.R. Ritter, Hydrogen-atom bond increments for calculation of thermodynamic properties of hydrocarbon radical species, *Journal of Physical Chemistry*, 99 (1995) 14514-14527.
- [33] M. Saeys, M.F. Reyniers, G.B. Marin, V. Van Speybroeck, M. Waroquier, Ab initio group contribution method for activation energies for radical additions, *AIChE Journal*, 50 (2004) 426-444.
- [34] M.K. Sabbe, A.G. Vandeputte, M.F. Reyniers, M. Waroquier, G.B. Marin, Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions, *Phys. Chem. Chem. Phys.*, 12 (2010) 1278-1298.
- [35] A.G. Vandeputte, M.K. Sabbe, M.-F. Reyniers, G.B. Marin, Kinetics of alpha hydrogen abstractions from thiols, sulfides and thiocarbonyl compounds, *Phys. Chem. Chem. Phys.*, 14 (2012) 12773-12793.
- [36] A.G. Vandeputte, M.-F. Reyniers, G.B. Marin, Kinetics of Homolytic Substitutions by Hydrogen Atoms at Thiols and Sulfides, *Chemphyschem*, 14 (2013) 1703-1722.
- [37] A.G. Vandeputte, M.-F. Reyniers, G.B. Marin, Kinetic Modeling of Hydrogen Abstractions Involving Sulfur Radicals, *Chemphyschem*, 14 (2013) 3751-3771.
- [38] M.K. Sabbe, M.-F. Reyniers, V. Van Speybroeck, M. Waroquier, G.B. Marin, Carbon-centered radical addition and beta-scission reactions: Modeling of activation energies and pre-exponential factors, *Chemphyschem*, 9 (2008) 124-140.
- [39] M.K. Sabbe, M.-F. Reyniers, M. Waroquier, G.B. Marin, Hydrogen Radical Additions to Unsaturated Hydrocarbons and the Reverse beta-Scission Reactions: Modeling of Activation Energies and Pre-Exponential Factors, *Chemphyschem*, 11 (2010) 195-210.
- [40] M.K. Sabbe, A.G. Vandeputte, M.-F. Reyniers, M. Waroquier, G.B. Marin, Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions, *Phys. Chem. Chem. Phys.*, 12 (2010) 1278-1298.
- [41] Daylight Chemical Information Systems, Daylight Theory Manual, 2010 <http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>.
- [42] C. Steinbeck, Y. Han, S. Kuhn, O. Horlacher, E. Luttmann, E. Willighagen, The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics, *J Chem Inf Comput Sci*, 43 (2003) 493 - 500.
- [43] C. Steinbeck, C. Hoppe, S. Kuhn, M. Floris, R. Guha, E.L. Willighagen, Recent Developments of the Chemistry Development Kit (CDK) - An Open-Source Java Library for Chemo- and Bioinformatics, *Current Pharmaceutical Design*, 12 (2006) 2111-2120.
- [44] A.G. Vandeputte, M.K. Sabbe, M.-F. Reyniers, V. Van Speybroeck, M. Waroquier, G.B. Marin, Theoretical study of the thermodynamics and kinetics of hydrogen abstractions from hydrocarbons, *Journal of Physical Chemistry A*, 111 (2007) 11771-11786.
- [45] T.N. Truong, W.T. Duncan, M. Tirtowidjojo, A reaction class approach for modeling gas phase reaction rates, *Phys. Chem. Chem. Phys.*, 1 (1999) 1061-1065.
- [46] R.J. Leroy, H. Murai, F. Williams, Tunneling model for hydrogen abstraction reactions in low-temperature solids. Applications to reactions in alcohol glasses and acetonitrile crystals, *J. Am. Chem. Soc.*, 102 (1980) 2325-2334.
- [47] A.G. Vandeputte, M.F. Reyniers, G.B. Marin, Theoretical Study of the Thermal Decomposition of Dimethyl Disulfide, *Journal of Physical Chemistry A*, 114 (2010) 10531-10549.
- [48] R.J. Kee, F.M. Rupley, J.A. Miller, M.E. Coltrin, J.F. Grcar, E. Meeks, H.K. Moffat, G. Lutz, A.E. Dixon-Lewis, M.D. Smooke, J. Warnatz, G.H. Evans, R.S. Larson, R.E. Mitchell, L.R. Petzhold, W.C. Reynolds, M. Caracotsios,

W.E. Stewart, P. Glarborg, C. Wang, O. Adigun, W.G. Houf, C.P. Chou, S.F. Miller, P. Ho, D.J. Young, CHEMKIN Release 4.1.1, (2007).

Chapter 8: Implementation of stereochemistry in kinetic model generation

8.1 Abstract

This chapter discusses the extension of an existing network generation tool, presented in Chapter 5, by accounting for stereochemistry in kinetic models. Genesys, see Chapter 5, was extended by the addition of algorithms for the detection of stereocenters, and for the identification of stereoisomers. To that purpose a so-called 2.5D representation of molecules that accounts for the presence of stereocenters and that keeps track of the associated stereoconfiguration of the stereocenters was introduced. A novel algorithm for the detection of steric relations between substituents allows the automated assignment of rate coefficients to stereoselective reactions. The functionality of the tool is illustrated for the thermal rearrangement of the monoterpenoid 2-pinanol, in which accounting for the stereochemistry is crucial to explain the reactive behavior of 2-pinanol and its products. Good agreement with experimental data from literature was obtained.

Keywords: kinetic model generation, stereochemistry, terpenoid chemistry; kinetics; monoterpenes

8.2 Introduction

During the last decades, new applications involving stereochemistry, such as asymmetric synthesis of pharmaceuticals and agrochemicals [1], and conversion processes of natural products such as terpenes [2] emerged for which the characterization of stereochemistry is required to understand the reactivity of molecules. In this context, computer-aided modeling of these chemical processes requires an adequate representation of stereochemistry in order to capture the essential characteristics of the chemical process. A valuable tool to study chemical processes are kinetic models, which give quantitative insights in the product distribution as a function of reaction conditions. For many chemical processes, the complexity of kinetic models surpasses the ability to manually construct them. To cope with this complexity, computers were programmed to automatically generate kinetic models [3-5]. Many tools have been proposed to automatically create kinetic models [6-8], but none of them account for stereochemistry. The reason for the lack of support for stereochemistry in kinetic models generated by these tools was the underlying representation of molecules that is used in the programs. Molecules were often represented as mathematical graphs, in which the nodes of the graph represent atoms, and the edges represent bonds. By doing so, a plethora of graph-theoretic algorithms and solutions could be implemented to accomplish tasks such as species comparison, substructure searching, greatly facilitating the efforts of extracting knowledge from chemical data by means of computers. Unfortunately, the graph representation also implied that the characterization and manipulation of stereochemical features of molecules and reactions became impossible since graphs only represent the connectivity between atoms, and not the arrangement in three-dimensional (3D) space.

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

8.3 Methodology

Genesys [7] is a tool for the automatic generation of kinetic models, consisting of a reaction network containing molecules and elementary reactions. The reaction network is generated by the iterative application of a limited set of reaction families that convert reactant molecules into product molecules. A pool of species, e.g. the reactants of a chemical process, is used as the start of the network generation procedure and is continuously expanded with molecules that arise as products of elementary reactions. A detailed description of the functionality of Genesys can be found in Chapter 5.

For kinetic models involving stereochemistry, a number of additional aspects need to be considered, as shown in Figure 8-1. First of all, the data structures designed to represent molecules inside the network generation program should allow distinguishing between stereoisomers.

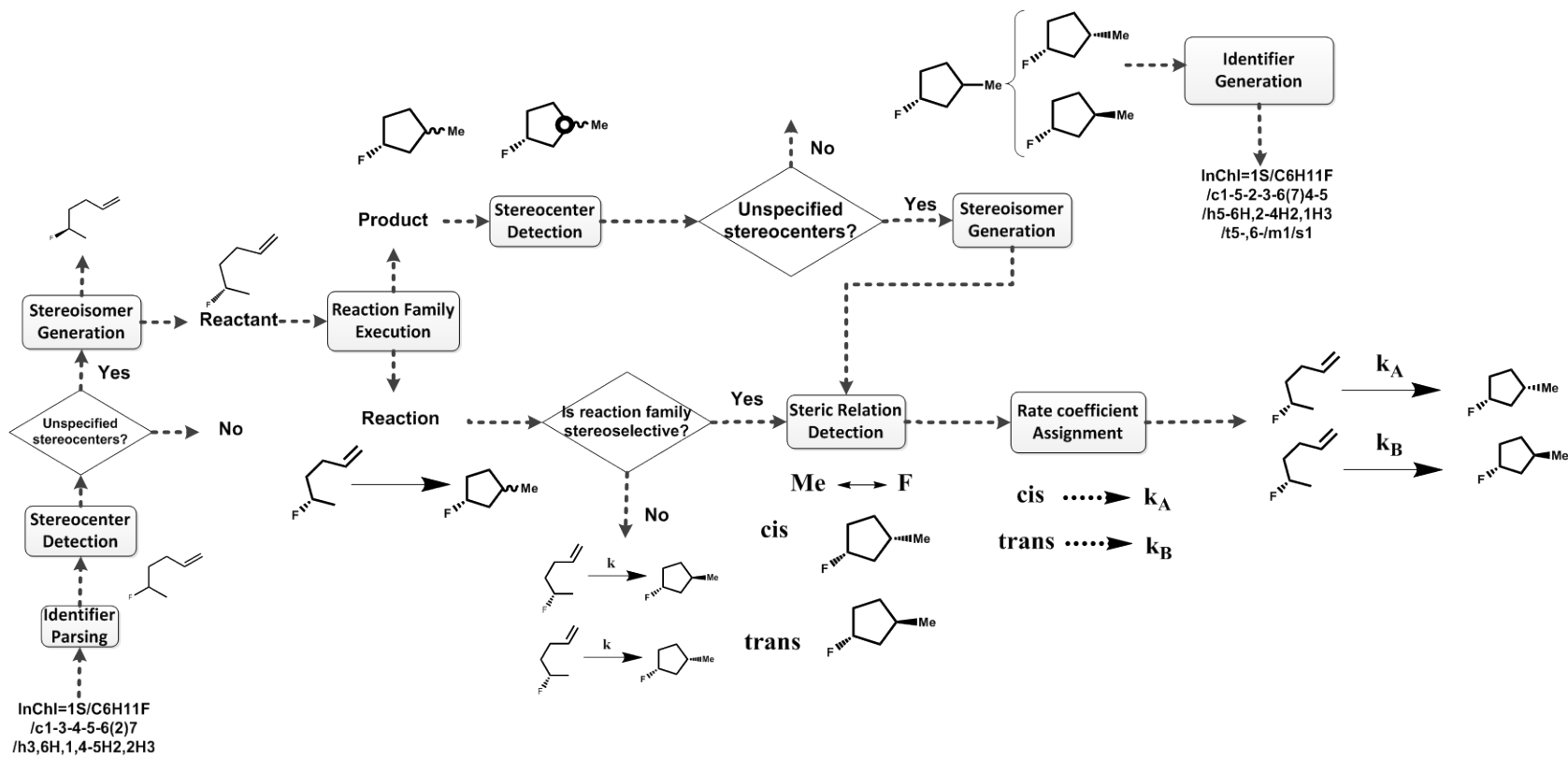


Figure 8-1: Overview of the algorithms in Genesys that deal with stereochemical aspects of molecules and reactions, through the example of the hypothetical reaction of (S)-5-fluorohex-1-ene yielding *cis*- and *trans*-1-fluoro-3-methylcyclopentane.

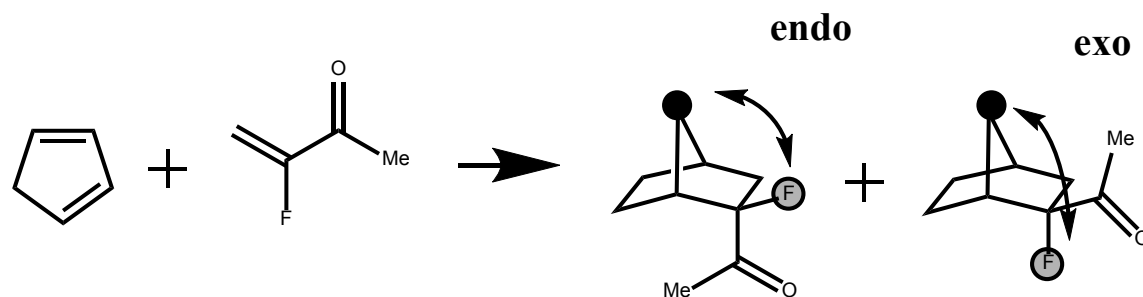
Information on stereochemical aspects of molecules needs to be correctly converted from the user interface into an internal molecule representation, and vice versa. Alternatively, if the user specifies a reactant in which the absolute configuration of the stereocenters is not fully specified, Genesys identifies these stereocenters through an adequate stereocenter detection algorithm, and exhaustively generates all possible stereoisomers using a so-called stereoisomer generation algorithm.

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

Finally, reactions may be stereoselective or stereospecific. Stereoselectivity refers to the preferential formation of one or the other of two (or more) stereoisomers from a single molecule with a prostereogenic element [9]. The latter refers to an element that can be converted from non-stereocenter to stereocenter in a single step [10]. Stereospecificity refers to the difference in reaction rates of two stereochemically different molecules, i.e. diastereomers or enantiomers [9]. Since stereospecificity emphasizes the role of the reactants in describing the differences in reaction rates, it is often described as reactant selectivity, whereas stereoselectivity focuses on the products and is called product selectivity. The stereospecificity of a reaction is sometimes linked to the absolute stereoconfiguration of stereocenters and is explained by the lock-and-key specificity of catalysts such as enzymes. Stereoselectivity depends on the difference in free energies of the respective transition states and is explained by stereoelectronic and steric factors among others.

Stereoselectivity and stereospecificity of reactions is often correlated with the three-dimensional arrangement between substituents in the stereoisomer, called the “steric relation” of substituents hereafter. An example of a steric relation is expressed by the *cis* and *trans* configuration of the fluor and methyl substituents in 1-fluoro-3-methylcyclopentane, cf. Figure 8-1. Prefixes such as *cis/trans*, and *exo/endo* used in combination with systematic names in IUPAC nomenclature refer

to the steric relation of two substituents with the highest priority surrounding the stereocenter. Atoms with higher atom numbers have higher priorities. For example, the prefix *cis* in *cis*-1,2-dichloro-1,2-difluorocyclopropane, refers to the two chlorine substituents in the cyclopropane ring, because the chlorine atoms have a higher atom number than the respective fluorine atoms. In the context of stereoselectivity of reactions, the priority of two substituents is less important. Rather, the choice of the substituents for which a steric relation is expressed, depends upon the way that the user wants to describe the stereoselectivity of a reaction. Scheme 8-1 illustrates this principle for the (hypothetical) Diels-Alder reaction of 1,3-cyclopentadiene with 3-fluorobut-3-en-2-one. The steric relation between the methylene bridge of the norbornane structure and the fluorine atom, rather than the acetyl group may be used as a discriminant to describe the stereoselectivity of the Diels-Alder reaction. In that case, the prefixes *endo* and *exo* refer to the relative arrangement of the methylene bridge and the fluorine atom.



Scheme 8-1: The (hypothetical) Diels-Alder reaction of 1,3-cyclopentadiene with 3-fluorobut-3-en-2-one in which the steric relation between the methylene bridge and the fluorine atom is expressed as *endo* and *exo*.

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

associated with a particular steric relation is retrieved from the reaction family definition, provided by the user.

The following sections explain the methodology used to represent stereochemistry and discuss the functionality of the introduced algorithms.

8.3.1 Stereochemistry of molecules

Kinetic model generation tools have historically represented molecules as graphs. The representation of stereochemistry in a molecule requires the introduction of stereocenters as part of the representation of the molecule. A stereocenter is a locus of stereoisomerism in a molecule so that the interchange of two ligands attached to an atom in such a molecule leads to another stereoisomer [11]. The specific configuration that is adopted by a stereocenter is represented by a binary stereodescriptor. Systematic naming conventions of molecules included stereodescriptors based on optical activity (+/-) or by relating to glyceraldehyde (D/L), etc... and allowed translating a chemical name into a depiction in 3D space, and vice versa. Especially the R/S and E/Z stereodescriptors [12] based on the Cahn–Ingold–Prelog (CIP) priority rules [13, 14] are useful for computer applications, since the assignment of them is solely based on a set of rules that can be implemented in computer algorithms, as explained in Section 1.2.3 of Chapter 1. Stereodescriptors in computer applications are stored as special attributes to atoms and bonds in the graph representation of the molecule, and are called stereoparities [15]. Similarly to the stereodescriptor, the stereoparity adopts two integer values; e.g. +1 and -1. By convention, the parity +1 is assigned to a stereocenter if a traversal of the prioritized ligands is clockwise, and -1 if it is counter-clockwise. In the framework of Genesys, the (R)-stereodescriptor is reserved for the +1 stereoparity, while (S) is reserved for the -1 stereoparity. The combination of the connectivity information of the graph representation with the stereochemical information stored in the stereoparity of a stereocenter is sometimes called a “2.5D representation” of the molecule [16]. The 2.5D representation is used throughout this chapter. Genesys uses the Chemistry Development Kit [17, 18] (CDK), an open-source chemoinformatics toolkit written in the programming language Java, for the internal representation of molecules. This toolkit not only provides a comprehensive data structure for molecules that incorporate stereoparities for atoms and bonds, CDK also includes algorithms that enable the conversion, i.e. parsing and generation,

of popular line identifiers that may contain stereochemical information such as InChI [19] and SMILES [20] into the CDK molecule representation. As a result of the integration of this chemoinformatics platform in Genesys, stereoisomers are treated as distinct species during the network generation procedure.

Transition state structures that are used for the calculation of the reaction path degeneracy of an elementary reaction, cf. Chapter 6, are represented by the same internal representation of CDK as molecules. The use of a 2.5D representation for transition states allows the correct calculation of the reaction path degeneracy for reactions that involve molecular chiral transition states, e.g. reaction R6-5 and R6-6 in Section 6.4 of Chapter 6.

8.3.1.1 Stereocenter detection

Besides the presence of a suitable data structure for stereoisomers, a definition of a stereocenter needs to be proposed that allows the unambiguous identification of stereocenters. A classical definition of a stereocenter starts from the constitution of the ligands surrounding the stereocenter. Stereocenters that comply with this definition are called true-stereocenters. For example, a tetrahedral, sp^3 hybridized stereocenter such as a chiral carbon atom is surrounded by four constitutionally distinct substituents. However, in the presence of topological symmetry, a molecule can contain stereocenters that do not satisfy this definition, but still determine the number and kinds of stereoisomers of a given component. For example, 4-fluoro-3-difluoroethylpentan-2-ol contains 5 stereocenters and exists as eight stereoisomers, cf. Figure 8-2.



Figure 8-2: 4-fluoro-3-difluoroethylpentan-2-ol (left) and 5-methylcyclohexane-1,3-diol (right) containing stereocenters that do not comply with the classical definition of a stereocenter.

Three of the four substituents of the central quaternary carbon atom are identical, and influence the structure and number of the stereoisomers of 4-fluoro-3-difluoroethylpentan-2-ol. In 5-

methylcyclohexane-1,3-diol, which exists as four stereoisomers, two of the four substituents of the tertiary carbon atom that bears the methyl group are identical, cf. Figure 8-2.

Stereocenters that do not comply with the classical definition of being surrounded by constitutionally distinct substituents, are called *para*-stereocenters, following the nomenclature proposed in Razinger et al. [21]. These authors defined three rules to identify *para*-stereocenters [21]. Consider for example, the carbon atom bearing the methyl group in 5-methylcyclohexane-1,3-diol, cf. Figure 8-3. Two substituents of this carbon atom are constitutionally identical so that this atom cannot be considered a true-stereocenter. However, as one of the rules of Razinger et al [21] dictates, the carbon atom is considered a *para*-stereocenter because the atom is connected to a pair of identical substituents, each containing at least one true-stereocenter and part of the same ring.

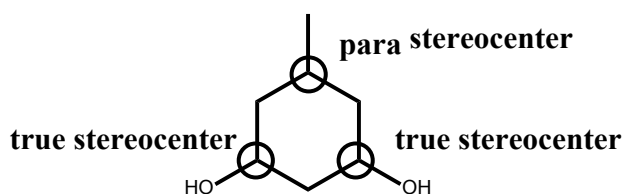


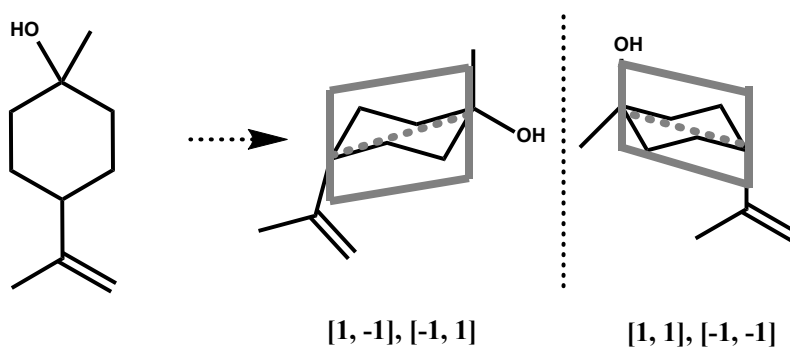
Figure 8-3: 5-Methylcyclohexane-1,3-diol bearing two true-stereocenters and one *para*-stereocenter. The *para*-stereocenter is connected to two topologically identical substituents each containing a true stereocenter. The two topologically identical substituents differ from each other because the true stereocenter of the substituent can adopt two different, absolute configurations (R/S).

Although these rules are heuristic in nature, they are applicable to a large and diverse set of components.

A formal definition of stereocenters allows the detection of stereocenters in molecules. Stereocenter detection is a necessary ingredient in network generation procedures since reactions can create or destroy stereocenters. Hence, the structure of the products that arise after the completion of a reaction needs to be searched for stereocenters. An implementation of a stereocenter detection algorithm based on the Razinger et al. rules is available in the open-source chemoinformatics toolkit OpenBabel [22]. The OpenBabel implementation is used throughout Genesys to detect stereocenters in stereoisomers. Stereocenters that arise in molecular chiral transition states may not necessarily comply with the formal definition of stereocenters in molecules. For example, the transition state of reaction R6-5 of Chapter 6 contains a bond that is

transitioning from a double to a single bond order. One terminal carbon atom of this transitioning double bond is surrounded by three ligands instead of two that is expected for a double bond in a molecule. Therefore, besides the formal definition of stereocenters in molecules, the definition of stereocenters in molecular chiral transition states is expanded with a limited number of additional rules that allow the detection of stereocenters in transition state structures.

When all stereocenters are identified by means of a stereocenter detection algorithm, the stereoparities of the identified stereocenters are *a priori* undefined. In principle, each stereocenter can adopt two possible configurations, leading to 2^n possible combinations, with n the total number of stereocenters. The combination of the stereoparities of the stereocenters in a molecule is stored in a so-called configuration vector [21], with each element of the vector corresponding to a specific stereocenter. If a molecule contains topological symmetry, some of the combinations of stereoparities may represent the same stereoisomer, leading to a total number of stereoisomers less than 2^n . In other words, some of the 2^n configuration vectors represent the same stereoisomer. These stereoisomers, i.e. meso-components, are characterized by the presence of a symmetry plane that superimposes the mirror image of the structure onto the original. For example, the twofold rotational symmetry axis in β -terpineol reduces the number of unique stereoisomers from four to only two, cf. Scheme 8-2. The symmetry plane of the two distinct meso-components of β -terpineol passes through the tertiary and quaternary carbon atom and is perpendicular to the cyclohexane plane.

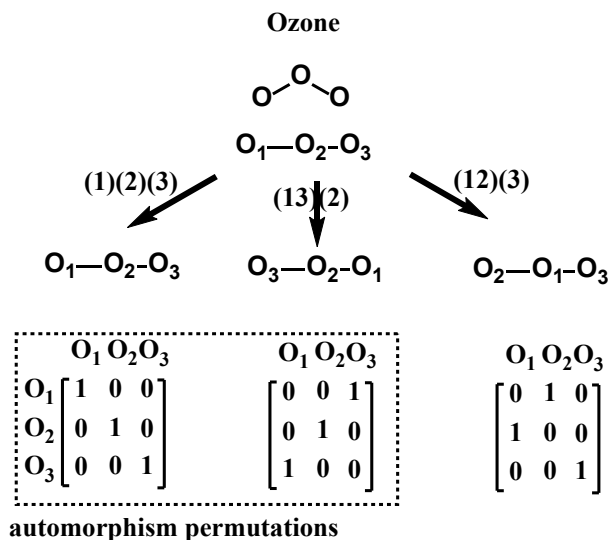


Scheme 8-2: β -terpineol and its two unique stereoisomers, each of them a meso-component due to the presence of a symmetry plane (in gray). The corresponding set of configuration vectors is depicted below each stereoisomer.

8.3.1.2 Stereoisomer generation

The generation of all stereoisomers that are associated with a structure with unspecified stereocenters is executed when a reaction creates product structures with newly created stereocenters. A stereoisomer generation algorithm aims at identifying the configuration vectors that represent the same stereoisomer. The algorithm is also used when the user defines stereoisomers as reactants without specifying the stereoconfiguration of the stereocenters of that stereoisomer. A number of stereoisomer generation algorithms exist in literature [23, 21, 24-26], but free and publicly available programs based on them were not found. The methodology used in this chapter is based on the algorithm by Razinger et al. [21] and follows the lines of the work of Nourse [27]. The key principle of the stereoisomer generation algorithm is the detection of topological symmetry in the molecule by means of atom permutations, and the subsequent determination of the impact of the atom permutations on the stereoparities of the stereocenters.

Topological symmetry in molecules can be detected using atom permutations. An atom permutation is a bijection of the set of atoms in a molecule to the same set of atoms. For a molecule with n atoms there are $n!$ possible atom permutations. For molecules that exhibit symmetry, some of the atom permutations preserve the connectivity between the atoms of the original molecule. These atom permutations are called automorphism permutations. The concept of automorphism permutations was also used to calculate the symmetry number, as explained in Chapter 6. For example, in the case of ozone, two permutations are automorphism permutations, i.e. the trivial identity permutation (1) (2) (3) and the permutation in which the oxygen atoms labeled with index 1 and 3 are permuted (13) (2), cf. Scheme 8-3. A third permutation that permutes atom 1 and 2 (12) (3) is not an automorphism permutation because the original bond between oxygen atom 2 and 3 is absent in the resulting molecule.

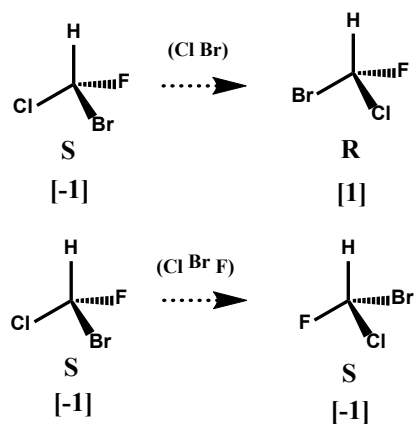


Scheme 8-3: Three atom permutations of oxygen atoms in ozone with two of them being automorphism permutations. (1 2) (3) is the cycle notation of the permutation that permutes elements 1 and 2; elements within parentheses are permuted in the order shown from left to right. Permutation matrix representation are shown below each permutation: diagonal elements equal to one indicate that the atom is not permuted. Off-diagonal elements at row *i* and column *j* indicate the bijection between atom *i* and atom *j*.

An important property of the set of automorphism permutations is that they form a mathematical group, satisfying the axioms of closure, associativity, identity and inversion [28]. In the case of ozone, the six permutations of the three oxygen atoms can be partitioned into three automorphism groups, each consisting of two permutations. Many authors acknowledged the similarities between the automorphism group of a graph and the topological symmetry of the molecule which lead to several algorithms for the efficient construction of the automorphism group tailored for chemical graphs [29-35]. This chapter uses an existing algorithm [36] to obtain the automorphism group of a chemical graph.

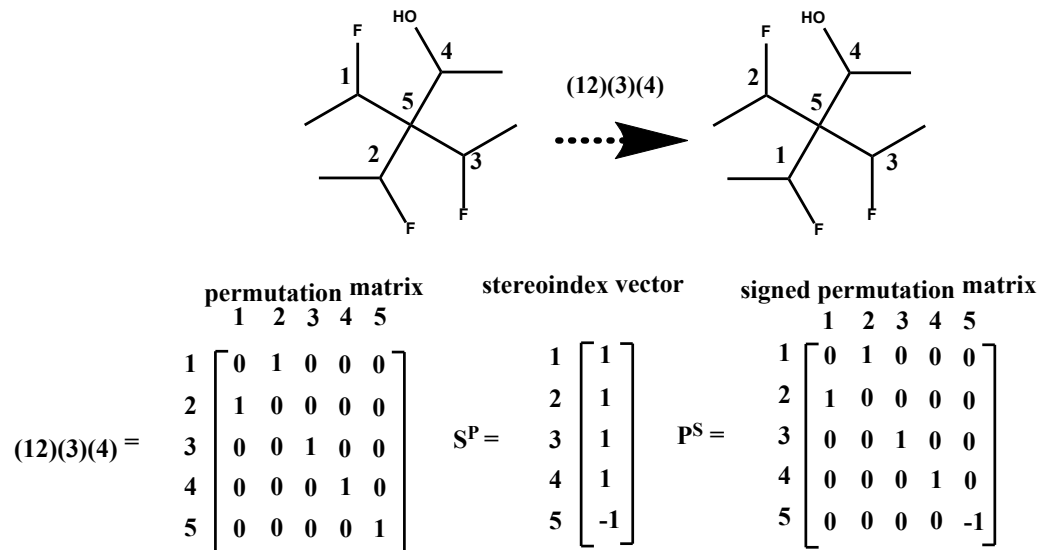
Atom permutations change the absolute configuration of stereocenters, if the sign of the permutation is negative. The sign of an atom permutation refers to the number of pairwise atom changes into which an atom permutation can be decomposed. The sign is negative is when the atom permutation consists of an odd number of pairwise atom changes. This is illustrated in Scheme 8-4. A single swap between the chlorine and bromine atom in (S)-bromo-

chlorofluoromethane results in the (R)-enantiomer, whereas applying two pairwise atom changes results in the (S)-enantiomer.



Scheme 8-4: The effect of two atom permutations of (S)-bromo-chloro-fluoromethane on the absolute configuration of the chiral carbon, and the corresponding configuration vectors shown below each of the stereoisomers.

The stereoisomer generation methodology is illustrated for 4-fluoro-3-difluoroethylpentan-2-ol. The latter contains five stereocenters, hence leading to 32 possible combinations of stereocenters. Four out of five stereocenters are surrounded by four distinct substituents, and are described as true-stereocenters. The central, quaternary carbon atom bears three constitutionally identical substituents, although the absolute configuration of each of the true-stereocenters may differ. Therefore, this central atom is a *para*-stereocenter, and its presence may reduce the number of unique stereoisomers with this constitution. The impact of an atom permutation on the stereoconfiguration of each of the stereocenters is represented by a vector of binary values, called the stereoindex vector. A value of +1 in the stereoindex vector represents “no change” in stereoconfiguration and -1 the opposite. An automorphism permutation (12)(3)(4) of the atoms of 4-fluoro-3-difluoroethylpentan-2-ol interchanges the substituents with labels 1 and 2, while the third identical substituent with index 3 remains unchanged.



Scheme 8-5: Construction of the stereoindex vector S^P and the signed permutation P^S associated with the permutation (12)(3)(4) for 4-fluoro-3-difluoroethylpentan-2-ol. The five stereocenters in 4-fluoro-3-difluoroethylpentan-2-ol are labeled 1 – 5. The permutation swaps stereocenter 1 and 2 leading to a value of -1 for the element of stereocenter 5 in the stereoindex vector. The signed permutation matrix is obtained by multiplying each row of (12)(3)(4) with the corresponding element in S^P .

If a permutation alters the order of the substituents of a stereocenter, the value of -1 is assigned to the element in the stereoindex vector. For permutation (12)(3)(4), only the order of the substituents surrounding the stereocenter with index 5 changes, but not of the other stereocenters with indices 1 – 4. Hence, the stereoindex vector S^P consists of four +1 values and only one -1 value. Multiplying every row of the matrix associated with the permutation with the corresponding element in the stereoindex vector creates a new matrix that combines information on the permutation, and information on the behavior of stereocenters under a given permutation. It is called the signed permutation matrix P^S , and is the central entity that allows the detection of equivalent configuration vectors.

The multiplication of a signed permutation with a configuration vector returns the stereoparities of the corresponding stereocenters after they underwent the permutation, and after accounting for the impact of the permutation on the stereoparity of the stereocenters. If the product of a signed

permutation P^S with a configuration vector v_i yields another distinct configuration vector v_j , it implies that both configuration vectors represent one and the same stereoisomer, cf. Eq. 8-1 [21]:

$$v_i \equiv v_j \Leftrightarrow \exists P^S : P^S \cdot v_i = v_j \quad \text{Eq. 8-1}$$

In the example of Scheme 8-5, the multiplication of the signed permutation associated with (12)(3)(4) with a configuration vector with elements (1, 1, 1, 1, 1), yields a configuration vector with elements (1, 1, 1, 1, -1). Thus, vectors (1, 1, 1, 1, 1) and (1, 1, 1, 1, -1) represent the same stereoisomer. Hence, the equivalence of two configuration vectors v_i and v_j is tested by verifying whether a signed permutation P^S from the set of automorphism permutations exists that converts one configuration vector into the other. The set of 2^n configuration vectors is divided into subsets of equivalent configuration vectors. Each subset represents a distinct stereoisomer, obtained by iterative verification of the equivalence of a configuration vector for a given signed permutation. The application of the above stereoisomer generation procedure for 4-fluoro-3-difluoroethylpentan-2-ol leads to eight distinct stereoisomers, cf. Figure 8-4.

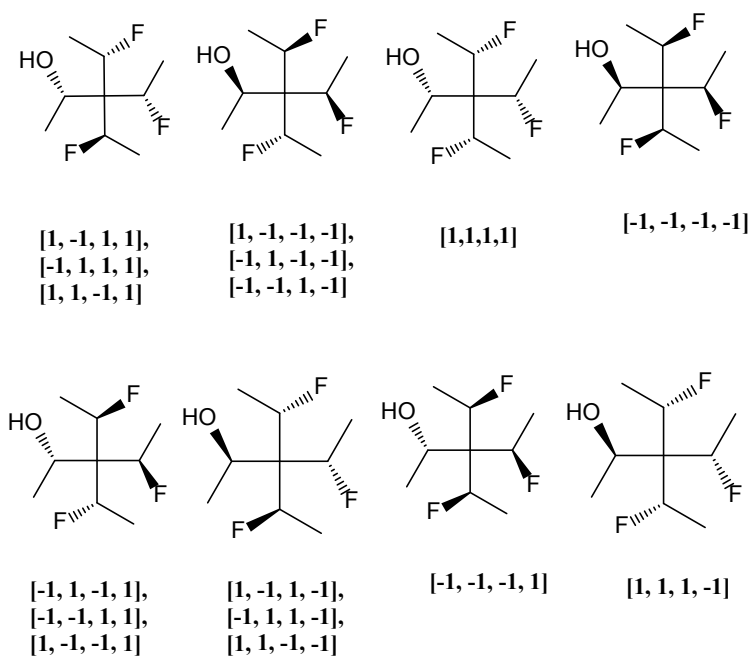
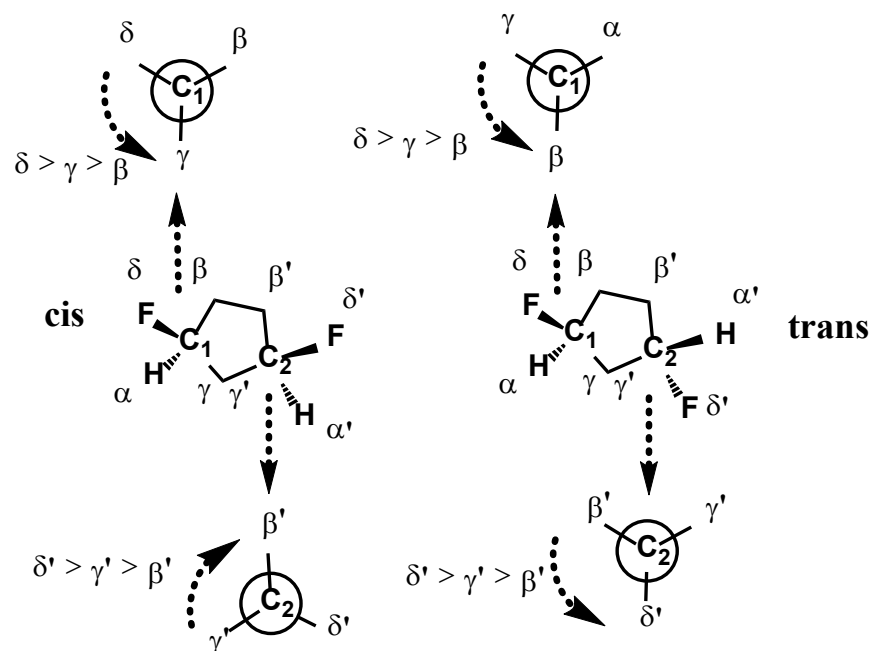


Figure 8-4: Eight stereoisomers generated from 4-fluoro-3-difluoroethylpentan-2-ol by the application of the stereoisomer generation algorithm. The set of corresponding configuration vectors is shown below each structure.

8.3.1.3 Steric relation detection

The detection of steric relations between two designated substituents is not explicitly stored in the molecule representation and therefore needs to be determined *ad hoc*. A novel algorithm was designed for the detection of the steric relation of two designated substituents of two stereocenters, and is discussed below. For reasons of simplicity, the principle of steric relation detection is illustrated for the 1,3-difluorocyclopentane in Scheme 8-6. The two stereocenter atoms are labeled C_1 and C_2 . First, the non-hydrogen substituents of the bridgehead atom C_1 , i.e. the three branches starting from bridgehead atom C_1 are labeled β , γ , and δ . When the substituents are visited in the order $\delta \rightarrow \gamma \rightarrow \beta$ and the viewer is positioned at the hydrogen atom of the C_1 bridgehead atom and views this ordered array of substituents along the H- C_2 bond, these substituents are visited in counterclockwise sense for the isomer with fluor substituents in *cis*-position with respect to the cyclopentane ring ("*cis*-isomer") and in clockwise sense for the isomer with fluor-substituents in *trans*-position ("*trans*-isomer"). If the viewer changes its point of view and is positioned at the hydrogen atom of the C_1 bridgehead atom and views the same ordered array of substituents along the H- C_2 bond, then the substituents in both isomers are visited in counterclockwise sense. Hence, it can be seen that the rotation direction around the two stereocenters is in opposite direction for the *cis*-isomer structure (counterclockwise and clockwise), while in the *trans*-isomer the rotation direction around both stereocenters is the same (counterclockwise).



Scheme 8-6: Principle of steric relation detection illustrated for 1,3-difluorocyclopentane. Identical substituents of both stereocenters C_1 and C_2 are labeled with indices α , β , γ and δ . Newman projections show the projection along the $H \rightarrow C_1$ axis (above) and along the $H \rightarrow C_2$ axis (below). For the isomer with fluor substituents in *cis* position relative to the cyclopentane ring, the rotation of the substituents that are visited in the same order, i.e. $\delta \rightarrow \gamma \rightarrow \beta$, is in opposite direction for both stereocenters, while the rotation is in the same direction for the isomer with fluor substituents in *trans* position relative to the cyclopentane ring.

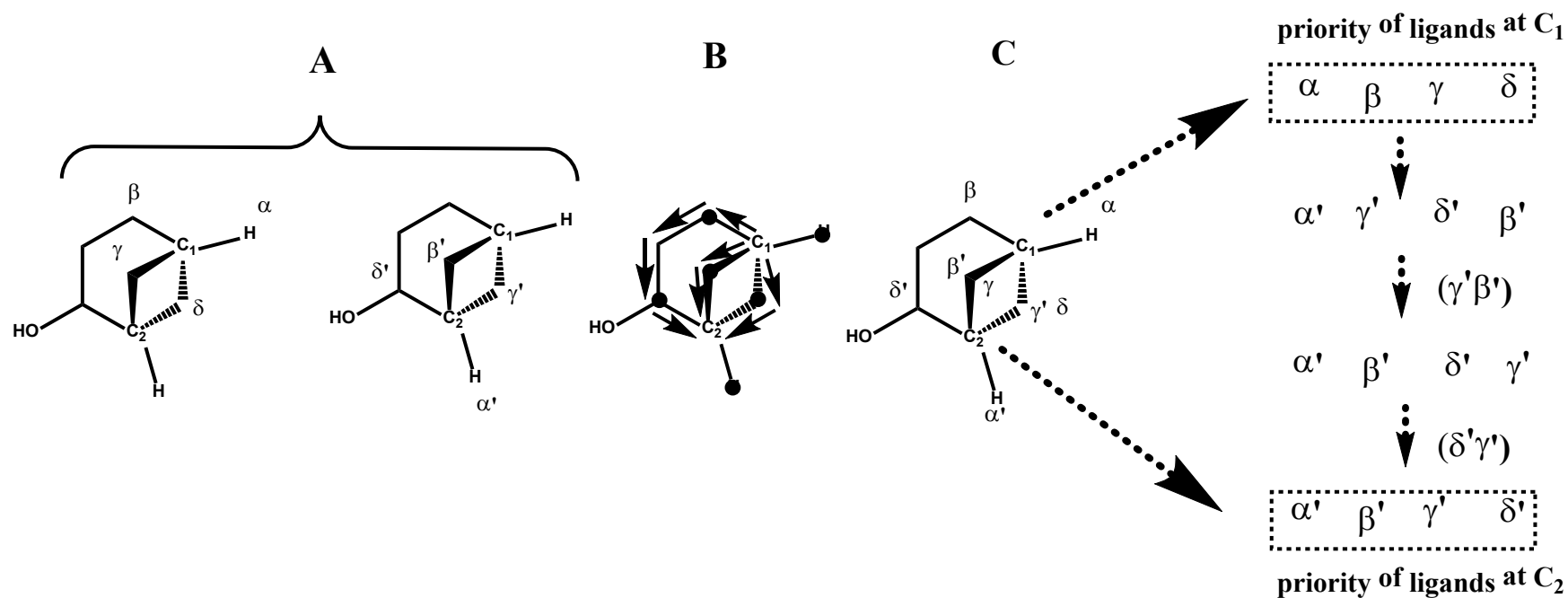
In the example of 1,3-difluorocyclopentane, the relative rotation around both stereocenters is visually determined by iterating over the same sequence of identical substituents, i.e. the fluor atom, and the two branches of the cyclopentane ring. A computer representation does not *a priori* have information on the rotation direction around both stereocenters and can only use the stereoparity values of both stereocenters to determine this.

The methodology to obtain information on the relative rotation direction of two stereocenters is illustrated for bicyclic hydrocarbon, bicyclo[3.1.1]heptane, and shown in Scheme 8-7. When the bridgehead atoms in the bridged ring systems are identified as stereocenters C_1 and C_2 , the stored stereoparity of one bridgehead, e.g. C_1 , is retrieved. The next step involves the prioritization of the ligands around the stereocenter C_1 using the CIP rules [13, 14]. Hydrogen is assigned the lowest priority (α). The branch containing hydroxyl group has the second lowest priority (β) since

the carbon in β -position of the bridgehead is a secondary carbon atom while for the other two branches it is a tertiary carbon atom. Finally, assuming that the stereoparity of the stereocenter C_1 is +1, the methylene bridge pointing away from the observer has the highest priority (δ), over the methylene bridge (γ) pointing towards the observer, given that stereoparity value of +1 refers to a clockwise orientation of the prioritized ligands. The resulting order is $\alpha < \beta < \gamma < \delta$, with $\alpha - \delta$ referring to the substituents of C_1 . The same steps are performed for the other bridgehead atom C_2 . The stereoparity of stereocenter C_2 is retrieved, e.g. -1, and the substituents surrounding C_2 are prioritized: $\alpha' < \beta' < \gamma' < \delta'$.

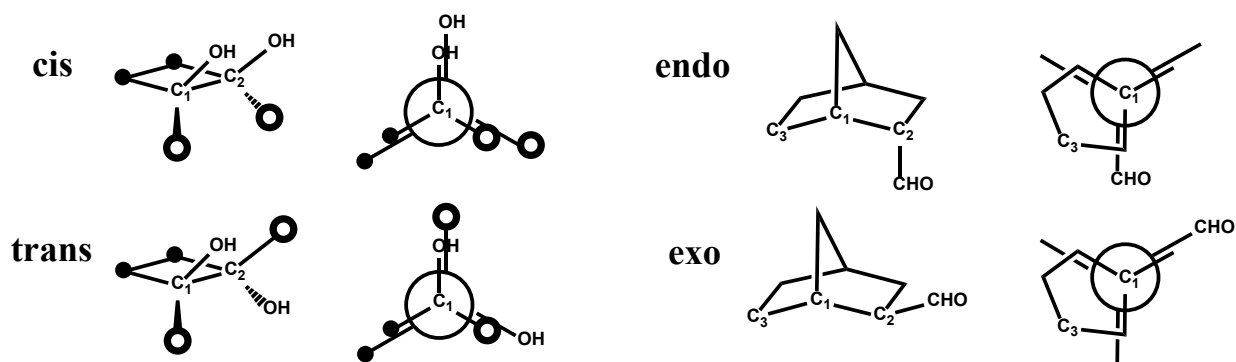
Since the branches between both bridgehead atoms C_1 and C_2 may consist of more than one atom, the bridgehead atoms are not necessarily connected to the same atoms. Therefore, each branch of atoms is walked from one bridgehead atom to the other to determine which substituent at bridgehead C_1 corresponds to the substituent at bridgehead C_2 . Using this operation to associate the atom substituents of both bridgehead atoms, the prioritized array of substituents of C_1 is translated into the order of the substituents surrounding C_2 . This leads to the following order around stereocenter C_1 : $\alpha' < \gamma' < \delta' < \beta'$. Next, the sign of the permutation that converts the order of the substituents of C_1 ($\alpha' < \gamma' < \delta' < \beta'$) into the order of the substituents of C_2 ($\alpha' < \beta' < \gamma' < \delta'$) is determined. Two pairwise permutations are needed to convert $\alpha' < \gamma' < \delta' < \beta'$ into $\alpha' < \beta' < \gamma' < \delta'$. Each pair-wise permutation contributes a factor -1 to the sign of the global permutation. Hence, the sign of the permutation that converts $\alpha' < \gamma' < \delta' < \beta'$ into $\alpha' < \beta' < \gamma' < \delta'$ equals $(-1)^2 = +1$.

Finally, the relative parity of both bridgehead atoms is determined. This is done by multiplying the stereoparity of stereocenter C_1 with the sign of the permutation that obtains the order the same substituents around stereocenter C_2 . In this example, this leads to a parity of $(+1) \cdot (-1)^2 = +1$ for stereocenter C_1 . This value of stereocenter C_1 is compared to the stereoparity of stereocenter C_2 , i.e. -1. As a result, the relative parity between both stereocenters is “odd”, referring to the inequality between the parity of C_1 and the parity of C_2 . Hence, this combination of stereoparities, i.e. +1 for C_1 and -1 for C_2 , leads to *cis*-linked methylene bridge in bicyclo[3.1.1]heptane and thus represents the chemically meaningful stereoisomer.



Scheme 8-7: Methodology for detection of orientation of the methylene bridge in bicyclo[3.1.1]heptane. A: Prioritization of the ligands of both bridgehead atoms. B: Determining the correspondence between both sets of prioritized ligands. (the colors red, green orange and blue indicate the pairs of corresponding branch atoms.) C: Determining the relative parity of both bridgehead atoms, by comparing the stored stereoparities.

The methodology described above can also be used to detect steric relations in other types of molecules. Scheme 8-8 illustrates the detection of the steric relation of two hydroxyl substituents connected to two stereocenters that are part of the same cyclobutane ring. For hydroxyl substituents in *cis* position relative to the cyclobutane ring, the direction of rotation of the three substituents of the C_1 carbon atom viewed along the C_1 - C_2 axis is opposite to the direction of rotation of the three substituents of the C_2 carbon atom if the substituents are visited in the same order.



Scheme 8-8: The steric relations *cis* and *trans* of the two hydroxyl substituents in 1,2-dihydroxycyclobutane (left) and the steric relation *exo* – and *endo* configuration of the aldehyde group with respect to the norbornane ring (right). Newman projections show the projection along the C_1 - C_2 axis.

Other steric relations can be detected along the same lines. *Exo-endo* isomerism refers to the position of a substituent with respect to the bridge. The prefix *endo* is reserved for the isomer with the substituent located closest to the longest bridge, i.e. the bridge containing the largest number of atoms. Scheme 8-8 illustrates this for *endo* and *exo*-norbornylcarbaldehyde. For the *endo* isomer, the longest bridge is in *cis* position to the carbonyl substituent if the cyclopentane ring is considered as the reference plane. For the *exo*-isomer, the longest bridge and the carbonyl substituent are in *trans*-position relative to the reference plane. Hence, the *exo* and *endo* configuration can be related to the *trans* and *cis* position of the substituent with respect to the atom belonging to the longest bridge respectively.

8.3.1.4 Chemical constraints

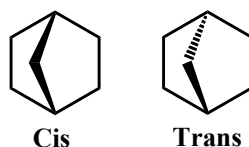
The stereoisomer generation algorithm identifies the equivalence of configuration vectors based on symmetry considerations. The algorithm does, however, not verify whether the combination of

stereoparities of the stereocenters leads to a chemically meaningful structure. As a consequence, the stereoisomer structures that are generated by the algorithm need to be tested on their physical meaningfulness. Two features are verified: double bond stereocenters in rings, and bridged rings in polycyclic structures.

Double bond stereocenters that are part of rings with only a few atoms cannot adopt both stereoconfigurations; the *trans*-isomer is considered chemically meaningless when the ring is too small. Cyclooctene is the smallest cycloalkene in which the *Z*-isomer is stable at room temperature [37]. Therefore, double bond stereocenters that are part of cycles with less than eight atoms, are not considered stereocenters, and are not used in the stereoisomer generation procedure. Hence, only *cis*-cyclohexene is retained, while *trans*-cyclohexene is considered chemically impossible.

The bridgehead atoms in bridged systems in polycyclic hydrocarbons such as bicyclo[2.2.1]heptane are two *para*-stereocenters according to the rules of Razinger et al. [21]. Application of the stereoisomer generation algorithm for bicyclo[2.2.1]heptane leads to two distinct stereoisomers because of the symmetry plane in the molecule. The detection of the steric relation of two substituents of stereocenters is also useful in describing chemically meaningless arrangements of atoms, as is illustrated for the bicyclo[2.2.1]heptane, cf. Scheme 8-9.

bicyclo[2.2.1]heptane



Scheme 8-9: *Cis*- and *trans*-linked stereoisomers of bicyclo[2.2.1]heptane that are generated by the application of the stereoisomer generation algorithm. Only the *cis*-linked form (norbornane) is chemically meaningful.

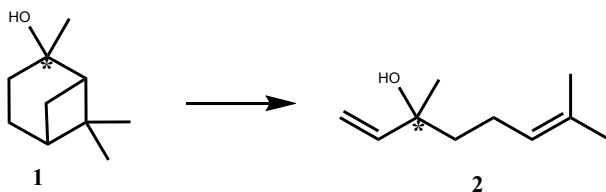
For bicyclo[2.2.1]heptane, the steric relation is expressed for the methylene bridge which is the common substituent of both bridgehead atoms. In one stereoisomer of bicyclo[2.2.1]heptane, the methylene bridge is “*cis*”-linked, and corresponds to norbornane. The other “*trans*”-linked stereoisomer is not chemically meaningful because it is too strained to exist under normal conditions. Potentially problematic bridged species are characterized by a combination of three elements: 1) the bridge connecting at least two rings 2) the bridge contains two bridgehead atoms

that are both stereocenters and 3) each of the branches connecting both bridgehead atoms consists of at least one bridge atom. When a potentially problematic bridge is identified, the steric relation algorithm is used to only retain the *cis*-linked bridges that arise after the stereoisomer generation procedure.

The exclusion of chemically meaningless stereoisomers that can be generated by assigning values to the stereoparities of the created stereocenters bears many similarities to the more generally applicable rules in organic chemistry. Rules such as Bredt's rule [38] for the location of double bonds in bridged rings, and the Woodward-Hoffmann rules [39, 40] for "allowed/forbidden" pericyclic reactions provide empirical guidelines in predicting the physical meaningfulness of proposed chemical structures, or proposed reactions. Hence, besides the two constraints that were implemented in Genesys, other constraints such as Bredt's rule or the Woodward-Hoffmann rules may be implemented in the future.

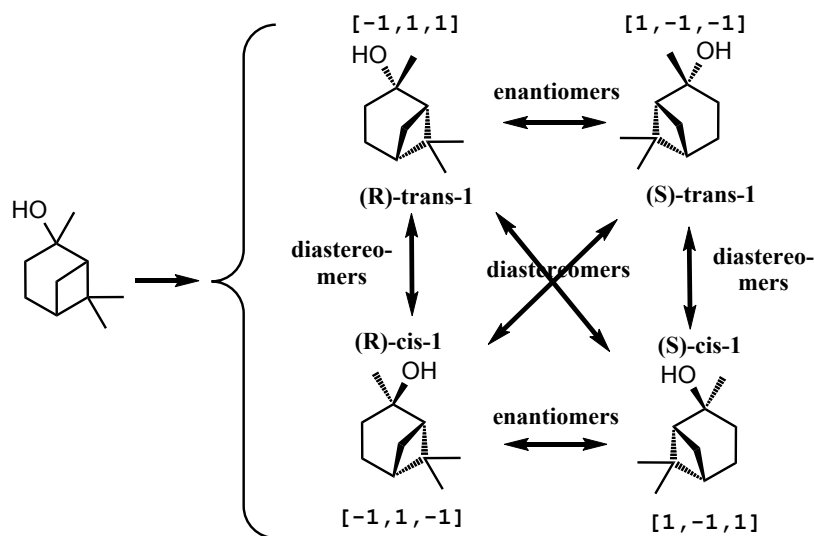
8.4 Application to pinanol thermal rearrangements

The method and implementation of kinetic model generation incorporating stereochemistry is validated for the thermal rearrangements of the monoterpene 2-pinanol (**1**). **1** belongs to the terpenoids, a highly diverse and functional class of biomass-derived isoprene oligomers, and is heavily used in the fragrance, flavor and pharmaceutical industry [41-45, 2]. The thermal rearrangement of **1** is an important industrial production route for linalool (**2**) (3,7-dimethylocta-1,6-dien-3-ol) [46], used in perfumes and as a precursor for vitamins A and E [41], cf. **R 8-1**.



R 8-1

There are four stereoisomers of **1**. They relate to each other as two pairs of diastereomers, *cis*-**1** and *trans*-**1**, depending on the spatial arrangement of the dimethyl bridge relative to the hydroxyl group, cf. Scheme 8-10.



Scheme 8-10: Four stereoisomers and the associated configuration vectors of 2-pinanol: two *cis*-2-pinanol enantiomers (left) and two *trans*-2-pinanol enantiomers (right).

Each pair of diastereomers of **1** relates to each other as enantiomers. Chapter 2 of this thesis showed that significant differences in reactivity and selectivity towards major products could be observed between the two diastereomeric pairs of **1** during the gas-phase thermal decomposition. There are two enantiomers of **2**, (R)-**2** (licareol), and (S)-**2** (coriandrol), depending on the absolute configuration of the chiral carbon atom bearing the hydroxyl substituent. Distinguishing between both enantiomers of **2** is important because each enantiomer of **2** has distinct odor characteristics and thresholds [46-48]. The structures of the product species arising during the pyrolysis of **1** are shown in Figure 8-5.

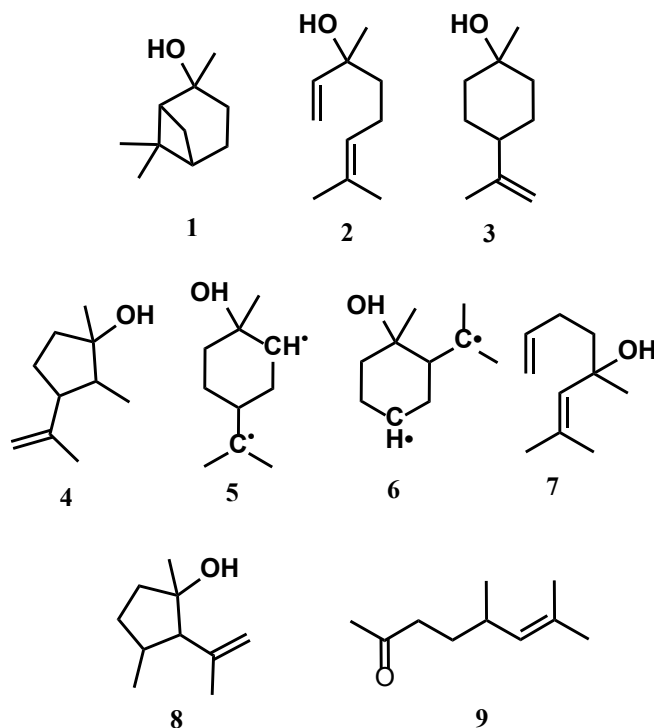


Figure 8-5: Chemical structures and their respective labels of components involved in the thermal rearrangement of 2-pinanol. 1: 2-pinanol, 2: linalool, 3: β -terpineol, 4: plinol, 5: 1,4-disubstituted cyclohexanol biradical, 6: 1,2-disubstituted cyclohexanol biradical, 7: isolinalool, 8: isoplinol,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

The generation of an adequate kinetic model using Genesys requires the knowledge of the relevant reactions that occur during the studied chemical process. This chemical knowledge is translated into a number of reaction families. The reaction network is generated by iterating over these reaction families, and the eligibility of the reactant to react according to one of the reaction family templates, is verified. New product molecules arise from the execution of the reaction family and serve as reactants for the next round in the model building process. Once the reaction network is constructed, thermochemical and kinetic parameters are assigned to the species and elementary reactions of the network, leading to the kinetic model that can be used in reactor simulations.

8.4.1 Reactants

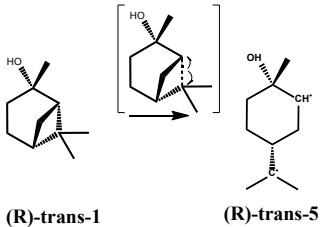
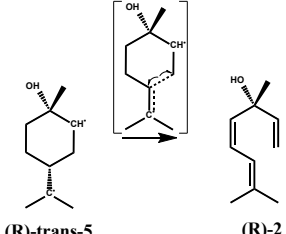
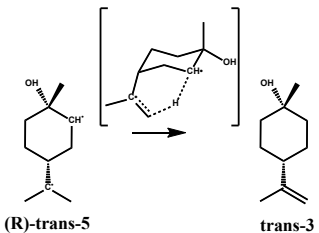
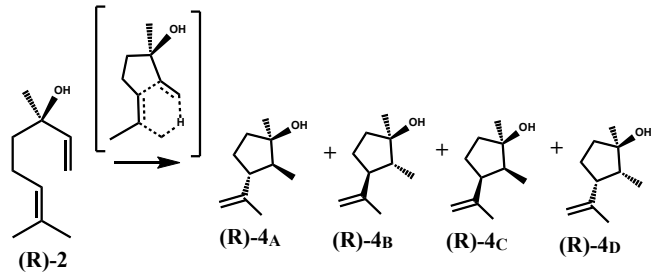
The procedure of the automated kinetic model construction proceeds as follows: **1** is entered as a reactant using a species identifier such as InChI, without specifying the absolute configurations

of the stereocenters. The stereoisomer generation algorithm generates eight distinct structures corresponding to eight distinct configuration vectors, and excludes four stereoisomers of **1** with *trans*-linked bridges. The four stereoisomers of **1** relate to each other as two pairs of enantiomers, cf. Scheme 8-10. Enantiomeric and diastereomeric relations between stereoisomers are identified using the associated configuration vectors. The “mirror image”, i.e. enantiomer, of a stereoisomer is identified by multiplying the associated configuration vector by a factor -1. For example, the configuration vector $[-1, 1, 1]$ of (R)-*trans*-**1**, is converted into $[1, -1, -1]$, which represents (S)-*trans*-**1**. Two stereoisomers are diastereomers with respect to each other if their respective configuration vectors are not the opposite of each other.

8.4.2 Reaction families

The reaction pathways that are relevant for the thermal decomposition of **1** were described in detail in Chapter 2. The chemical knowledge on the relevant pathways involved in the thermal rearrangement of **1** is translated into reaction families. Four reaction families are defined: homolytic C-C scission reactions, biradical β -scission reactions, sigmatropic [1,5]-H-shift reactions and ene-cyclization reactions, cf. Table 8-1. The fragmentation of the four-membered ring by homolytic C-C scission reactions is the initial step in the thermal isomerization of bicyclic monoterpenes consisting of a bicyclo[3.1.1]heptane system. This scission reaction leads to four 1,4-disubstituted cyclohexanol biradicals **5**. The 1,4-disubstituted biradicals quickly rearrange into linalool enantiomers through C-C β -scission reactions involving the bond in β -position of the two carbon atoms bearing the unpaired electron. An alternative, minor pathway leads to p-menthene type diastereomers of β -terpineol (**3**), via sigmatropic [1,5]-H-shift reactions. Linalool further isomerizes into plinol products via pericyclic ene-cyclization reactions.

Table 8-1: Reaction families and corresponding Arrhenius parameters considered in the current model generated using Genesys. Unit for single-event pre-exponential factor \tilde{A} s⁻¹, for activation energy E_a kJ mol⁻¹.

Reaction family	Example reaction	\tilde{A}	E_a
Homolytic C-C scission	 (R)-trans-1 → (R)-trans-5	Cf. Table 8-2	
Biradical β -scission reactions	 (R)-trans-5 → (R)-2	$6.30 \cdot 10^{13}$	88
Sigmatropic [1,5]-H-shift reactions	 (R)-trans-5 → trans-3	$0.26 \cdot 10^{11}$	67
Ene-cyclization reactions	 (R)-2 → (R)-4A + (R)-4B + (R)-4C + (R)-4D	Cf. Table 8-4	

The application range of a reaction family needs to be constrained as much as possible to only allow molecules that may undergo this reaction type. For example, the homolytic C-C scission reaction family models the initial fragmentation of the cyclobutane ring in **1**. It should be applied to all stereoisomers of **1**, while excluding other molecules that also contain a carbon-carbon single bond, such as **2**. Furthermore, the C-C scission reaction family should only target the weakest of the four bonds of the cyclobutane ring in **1**, resulting in **5**. Imposing restrictions on the applicability of a reaction family is achieved by specifying the pattern of atoms that should be present in each candidate reactant molecule. The required sub-molecular pattern for all four

reaction families is defined by the user through the use of the pattern language SMARTS [49]. For example, the SMARTS string CC1(O)CCC2C(C)(C)C1C2 describes the pattern of atoms depicted in Figure 8-6.

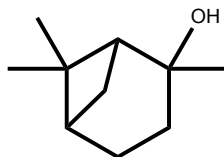


Figure 8-6: The 2-pinanol frame that is used as the pattern of atoms required for candidate molecules for the homolytic C-C scission reaction family.

Next to the identification of candidate reactants for a reaction family, the latter also includes a recipe-like description of the necessary steps (“BREAK BOND”, “FORM BOND”, etc...) that convert the reactant structure into a product structure. For example, the ene-cyclization reaction family requires six electron rearrangement steps to convert reactants into products, cf. Figure 8-7. Each of the atoms that are part of the electron rearrangement step are denoted by a letter in the recipe, enabling the unambiguous definition of the atoms that are involved in one of the electron rearrangement steps of the recipe. Further information on the definition of reaction families can be found in Section 5.3.1 of Chapter 5.

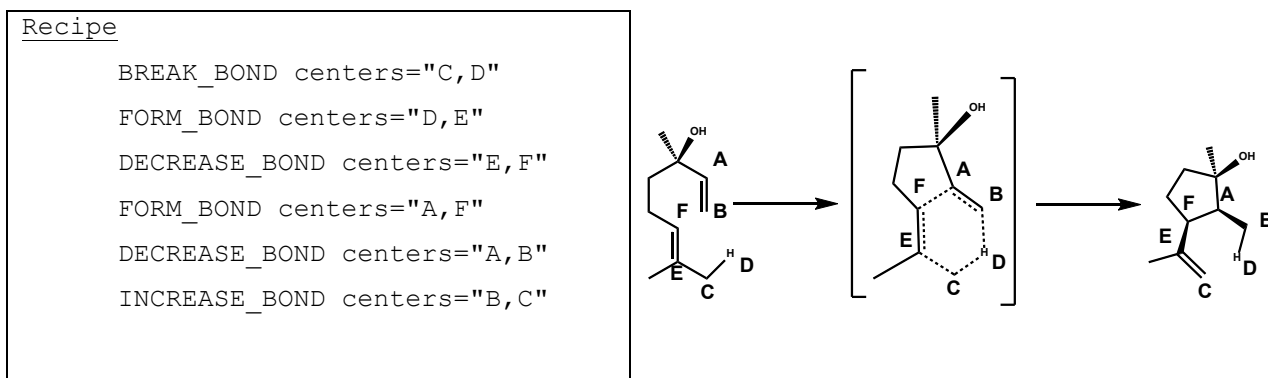
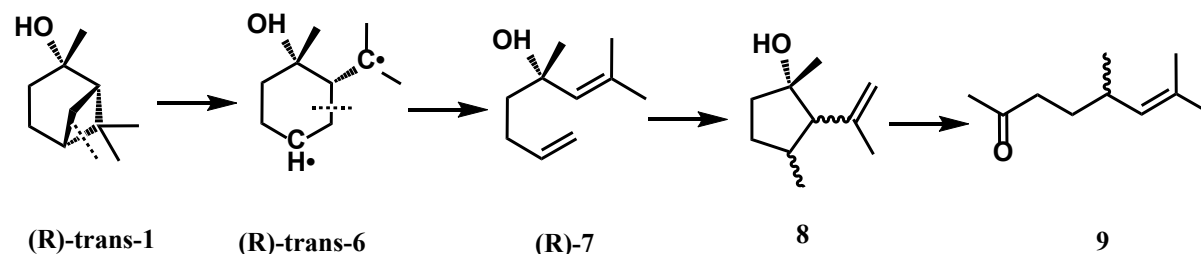


Figure 8-7: Reaction family recipe for the ene-cyclization. Letters A to F denote the reactive center atoms that undergo one of the electron rearrangement operations.

Besides the C-C scission of **1** that leads to **5**, another minor decomposition pathway exists, initiated by a C-C scission of **1** leading to 1,2-disubstituted biradicals **6**.



Scheme 8-11: Minor decomposition pathway of 1 via 1,2-disubstituted biradicals 6.

Similar to the parallel pathway leading to **5**, **6** quickly rearranges into a constitutional isomer of linalool, called isolinalool (**7**). The latter further undergoes similar ene-cyclizations leading to constitutional isomers of plinol, called isoplinol (**8**). However, **8** was never observed experimentally; instead an acyclic δ,ϵ -unsaturated ketone component (**9**) was detected [46, 50]. The latter was assumed to be formed through a fast retro-ene type reaction in which a γ -hydrogen atom is transferred to an unsaturated center via a six-electron cyclic transition state [51]. The parallel initiation of **1** via **6** leading to **8** and **9** represents only a minor consumption pathway of **1** with selectivities of **9** less than 5% at highest conversion, as was shown in Chapter 2. Since the types of reactions of this decomposition channel of **1** are very similar to the main path via **5**, they are further neglected in the discussion below for reasons of conciseness of this chapter.

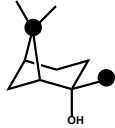
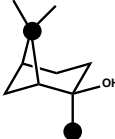
8.4.3 Thermochemistry and Kinetics

The calculation of thermochemical properties of the species in the model follows the same methodology as described in Section 5.3.2 of Chapter 5, i.e. a combination of group contribution methods such as Benson group additivity and the Hydrogen Bond Increment method of Lay et al. [52]. Symmetry numbers are calculated based on the algorithm described in Chapter 6. The ring strain in isomers of **1** is modeled using the ring strain correction corresponding to cyclobutane. The implemented methods for the calculation of thermochemical properties are approximations, depending on the availability of numerical values for the groups that are present in the molecule. For example, some stereoisomers of **4** relate to each other as diastereomers. Nevertheless, they are assigned the same thermochemical properties, due to the lack of appropriate groups that describe the interaction between the substituents of the cyclopentane ring. The same principle applies for constitutional isomers such as **2** and **7**. Despite the limitations of the used methods, it is expected that the calculated thermochemical properties are good approximations of the true values for the molecules in the current kinetic model.

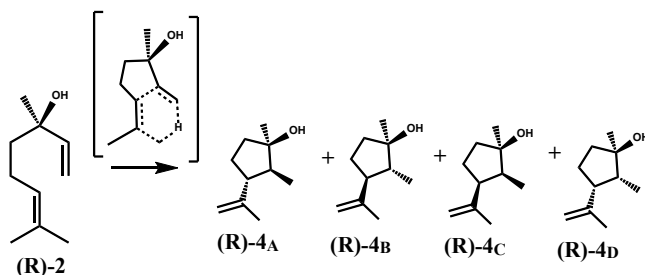
The assignment of the rate coefficients to each elementary reaction occurs on the level of reaction families in Genesys. Ideally, ΔGAV° s obtained from *ab initio* calculations are used to calculate values for the rate coefficients of the generated reactions. Unfortunately, values for ΔGAV° s for the reaction families required for the thermal decomposition of **1** are unavailable. Instead, fixed Arrhenius parameters were used for the generated reactions of each reaction family and were derived from experiments, as was described in Chapter 2, cf. Table 8-1. The provided pre-exponential factor per reaction family is a single-event pre-exponential factor that needs to be multiplied by the number of single-events of the particular reaction. The number of single-events is calculated following the algorithm described in Chapter 6.

As was shown in Chapter 2, the activation energy of the rate coefficient of the homolytic C-C scission reactions of **1** is 5 kJ mol^{-1} lower when the methyl group of the chiral carbon atom is in *cis* position relative to the dimethyl bridge of the bicyclo[3.1.1]heptane system, compared to *trans*-**1** in which the methyl group is in *trans* position relative to the dimethyl bridge. For the homolytic C-C scission reactions, the influence of the arrangement of the methyl group next to the chiral carbon with respect to the dimethyl bridge on the rate coefficients is added as a further specification to the reaction family, cf. Table 8-2. One rate coefficient is assigned to the C-C scission reactions with stereoisomers of **1** with the methyl group and dimethyl bridge atom in *cis* position with respect to each other. Another rate coefficient is assigned to C-C scission reactions with isomers of **1** in which the methyl group and dimethyl bridge atom in *trans* position with respect to each other. The steric relation detection algorithm enables the automatic assignment of the appropriate rate coefficient for the corresponding stereoisomer of **1**.

Table 8-2: Arrhenius parameters for the homolytic C-C scission reaction family, depending on the steric relation of the substituents circled. Unit for single-event pre-exponential factor \tilde{A} s⁻¹, for activation energy E_a kJ mol⁻¹.

	Steric relation	\tilde{A}	E_a
	Cis	$1.0 \cdot 10^{14}$	206
	Trans	$1.0 \cdot 10^{14}$	211

In Chapter 2, rate coefficients were derived for the biradical β -scission reactions. No significant difference in rates were reported between the *cis*- and *trans*-1,4-disubstituted biradicals as reactants. Therefore, the same rate coefficient was used for all reactions of this reaction family, cf. Table 8-1. Similarly to the biradical β -scission reactions involving the 1,4-disubstituted biradicals, rate coefficients of the sigmatropic [1,5]-H-shift reactions were not found to be significantly different between the *cis*- and *trans*-1,4-disubstituted biradicals [46], as can be read in Chapter 2. Therefore, the same rate coefficient was used, cf. Table 8-1. Each enantiomer of **2** forms four distinct diastereomers of **4**, depending on the orientation of the isopropenyl and adjacent methyl substituents of the cyclopentane ring, cf. R 8-2.

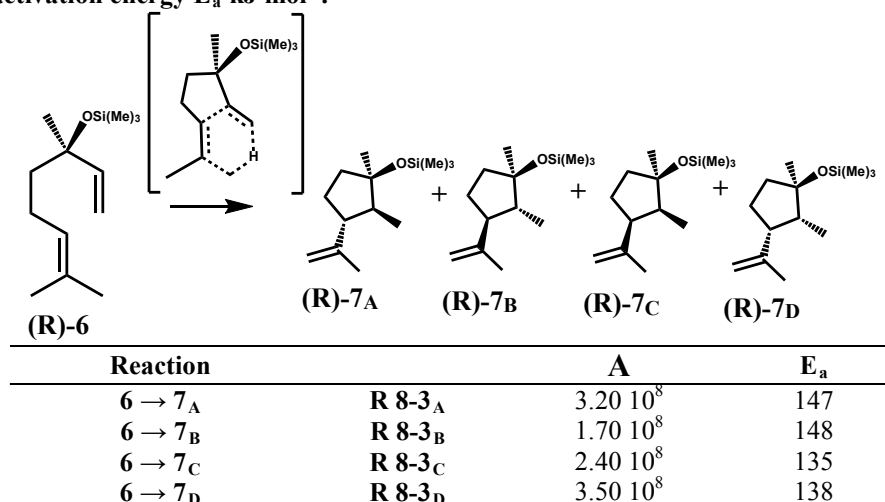


R 8-2

The ene-cyclization of **2** not only produces multiple distinct stereoisomers, it is also characterized by differences in stereoselectivity. Many studies indicated the selectivity of the ene-cyclization of **2** towards the isomers with the isopropenyl and adjacent methyl substituent in *cis*-position [53-56], i.e. a stereoselectivity towards isomers **4C** and **4D**. In Chapter 2, an average reaction rate was obtained for the ene-cyclization reaction of **2** based on the sum of the formed isomers of **4**. Pickenhagen et al. [55] derived four rate coefficients for the ene-cyclization reactions of the silyl

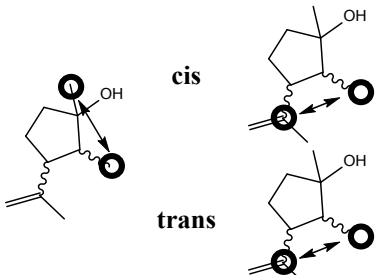
ether analogon of **2** (**6**), cf. Table 8-3, and noticed not only the influence of the relative position of the adjacent methyl and isopropenyl substituents, but also of the relative arrangement of the two adjacent methyl substituents on the rate of formation.

Table 8-3: Four ene-cyclizations of the silyl ether analogon of (**R**)-**2** together with the Arrhenius parameters derived by Pickenhagen et al. [55]. Unit for pre-exponential factor A (accounting for the number of single-events) s^{-1} , for activation energy E_a kJ mol^{-1} .



A similar procedure was designed for the assignment of rate coefficient for the stereoselective ene-cyclization reaction family as for the stereospecific homolytic C-C scission reactions of diastereomers of **1**. The stereoselectivity of the ene-cyclization towards the different isomers of **4** is encoded using two specifications: 1) the steric relation of the isopropenyl and the vicinal methyl group of **4** and 2) the steric relation of the two adjacent methyl groups of **4**. This leads to four rate coefficients that can be assigned to a given ene-cyclization reaction of **2** depending on the relative position of the designated substituents. Table 8-4 shows the values for the four ene-cyclization possibilities.

Table 8-4: Arrhenius parameters for the ene-cyclization reaction family, depending on the steric relation of the substituents circled. Unit for single-event pre-exponential factor \tilde{A} s⁻¹, for activation energy E_a kJ mol⁻¹.

			\tilde{A}	E_a
	cis	cis	$1.30 \cdot 10^7$	131
		trans	$0.70 \cdot 10^7$	132
		cis	$0.97 \cdot 10^7$	119
	trans	trans	$1.43 \cdot 10^7$	122

The rate coefficients for each of the ene-cyclization reactions of **2** yielding a stereoisomer of **4** are derived from the average value reported in Chapter 2, weighted by a factor derived from the stereoselective rate coefficients of the ene-cyclization of **6**, reported by Pickenhagen et al. [55]. The values of the stereoselective factor are derived by comparison of the relative rate coefficients of each of the four ene-cyclization reactions of **6** into a stereoisomer of **7** with respect to the mean rate of the global reaction of **6** into **7**.

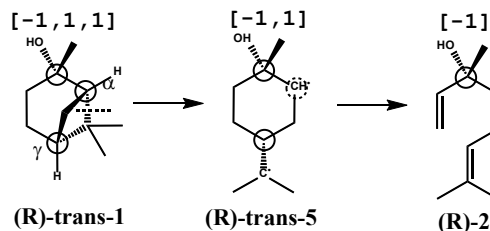
8.4.4 Kinetic model generation

Table 8-5 shows the generated kinetic model for the thermal rearrangement of **1**, consisting of 20 elementary reactions between 20 species. The presented kinetic model differs from the kinetic model presented in Chapter 2 in several aspects. The current model is automatically constructed by Genesys, as opposed to the previous model that was constructed by hand. The new kinetic model distinguishes between enantiomers while the hand-built model lumps them together. Finally, the model presented in this chapter accounts for the stereoselectivity of the ene-cyclization of **2**, by incorporating distinct reactions to each of the stereoisomers of **4**. The model in Chapter 2 only provides the lumped reaction in which a global reaction rate is provided towards the lump of isomers of **4**.

Table 8-5: Reactions and Arrhenius parameters in the kinetic model for the thermal isomerization of 2-pinanol isomers generated by Genesys. Units for pre-exponential factor A (accounting for the number of single-events) are s⁻¹, for E_a kJ mol⁻¹.

Reaction	A	E _a	Reaction family	Reference
(R)- <i>cis</i> -1 ↔ (R)- <i>cis</i> -5	1.0 10 ¹⁴	206	Homolytic C-C scission	Chapter 2
(S)- <i>cis</i> -1 ↔ (S)- <i>cis</i> -5	1.0 10 ¹⁴	206	Homolytic C-C scission	Chapter 2
(R)- <i>trans</i> -1 ↔ (R)- <i>trans</i> -5	1.0 10 ¹⁴	211	Homolytic C-C scission	Chapter 2
(S)- <i>trans</i> -1 ↔ (S)- <i>trans</i> -5	1.0 10 ¹⁴	211	Homolytic C-C scission	Chapter 2
(R)- <i>cis</i> -5 ↔ (S)-2	6.3 10 ¹³	88	Biradical double β-scission	Chapter 2
(S)- <i>cis</i> -5 ↔ (R)-2	6.3 10 ¹³	88	Biradical double β-scission	Chapter 2
(R)- <i>trans</i> -5 ↔ (R)-2	6.3 10 ¹³	88	Biradical double β-scission	Chapter 2
(S)- <i>trans</i> -5 ↔ (S)-2	6.3 10 ¹³	88	Biradical double β-scission	Chapter 2
(S)- <i>cis</i> -5 ↔ <i>cis</i> -3	1.6 10 ¹¹	67	[1,5]-Sigmatropic H-shift	Chapter 2
(R)- <i>cis</i> -5 ↔ <i>cis</i> -3	1.6 10 ¹¹	67	[1,5]-Sigmatropic H-shift	Chapter 2
(R)- <i>trans</i> -5 ↔ <i>trans</i> -3	1.6 10 ¹¹	67	[1,5]-Sigmatropic H-shift	Chapter 2
(R)- <i>cis</i> -5 ↔ <i>trans</i> -3	1.6 10 ¹¹	67	[1,5]-Sigmatropic H-shift	Chapter 2
(R)-2 ↔ (R)-4 _A	3.9 10 ⁷	131	Ene-cyclization	This chapter
(R)-2 ↔ (R)-4 _B	2.1 10 ⁷	132	Ene-cyclization	This chapter
(R)-2 ↔ (R)-4 _C	2.9 10 ⁷	119	Ene-cyclization	This chapter
(R)-2 ↔ (R)-4 _D	4.3 10 ⁷	122	Ene-cyclization	This chapter
(S)-2 ↔ (S)-4 _A	3.9 10 ⁷	131	Ene-cyclization	This chapter
(S)-2 ↔ (S)-4 _B	2.1 10 ⁷	132	Ene-cyclization	This chapter
(S)-2 ↔ (S)-4 _C	2.9 10 ⁷	119	Ene-cyclization	This chapter
(S)-2 ↔ (S)-4 _D	4.3 10 ⁷	122	Ene-cyclization	This chapter

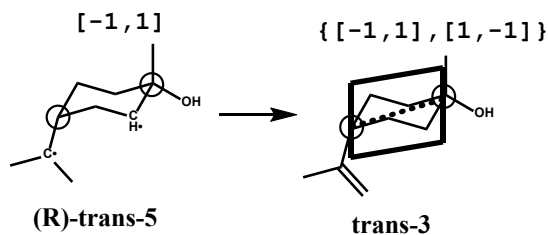
The four reaction families are used to generate all possible reactions and products and the appropriate rate coefficients are assigned to each reaction. The pool of reactant species, i.e. the four isomers of **1**, reacts to the homolytic C-C scission reaction family, leading to four cyclohexanol biradicals following the methodology described in Figure 8-1. After the reaction family recipe produces the cyclohexanol biradical structure, the fate of the three original stereocenters of the reactant molecule is investigated. Since the C-C scission reaction does not affect the true-stereocenter, i.e. the chiral carbon atom bearing the hydroxyl group, the absolute configuration of the chiral carbon in **5** is adopted from the corresponding chiral carbon atom in the reactant isomer of **1**, cf. Scheme 8-12.



Scheme 8-12: Carbon-carbon scission of a 2-pinanol isomer leading to cyclohexanol biradicals with the methyl and isopropenyl substituent in 1,4 position. Stereocenters are circled. Associated configuration vectors are depicted above each structure.

The situation is different for the two *para*-stereocenters, corresponding to the two bridgehead atoms of the bicyclo[3.1.1]heptane system. The *para*-stereocenter in α -position of the true-stereocenter is destroyed since the carbon atom in the cyclohexanol biradical only contains three substituents and hence is no longer a stereocenter. The *para*-stereocenter in γ -position of the true-stereocenter is converted into a true-stereocenter because its four substituents are now constitutionally different. However, the stereoparity of its *para*-stereocenter predecessor in **1** is preserved and adopted in the newly created true-stereocenter since the orientation of the isopropenyl ligand relative to the hydroxyl substituent in the biradical product adopts the same orientation as in the parent reactant **1**. The four isomers of **1** each lead to a distinct isomer of **5**.

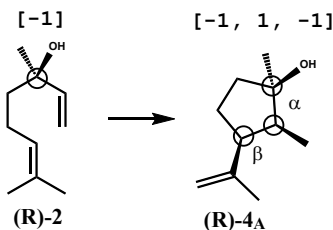
The four isomers of **5** undergo biradical β -scission reactions that destroy the cyclohexane ring, leading to **2**. The true-stereocenter of **5** bearing the isopropenyl substituents no longer exists and is converted into a non-stereogenic sp^2 -hybridized unsaturated carbon atom. On the other hand, the chiral carbon atom bearing the hydroxyl group is not affected by the biradical β -scission reaction. Thus, similarly to the initial C-C scission of **1**, the stereoparity of the true-stereocenter corresponding to the chiral carbon atom is preserved. Hence, the biradical β -scission reactions convert the four isomers of **5** into two enantiomers of **2**. The absolute configuration of the preserved chiral carbon atom reveals which isomer of **5** is linked to which isomer of **2**. Another reaction family that is applicable to the isomers of **5** consists of sigmatropic [1,5]-H-shift reactions. The asymmetry in the cyclohexane ring that is present in **5** due to the endocyclic carbon atom bearing the unpaired electron is removed during this reaction. Instead, β -terpineol isomers (**3**) emerge. As was already noted in Section 8.3.1.1 of this chapter, these molecules are characterized by a symmetry plane going through the tertiary and quaternary carbon atoms, and perpendicular to the cyclohexane plane, cf. Scheme 8-13.



Scheme 8-13: Sigmatropic [1,5]-H-shift reaction of (R)-trans-5 yielding *trans*-3. True-stereocenters are transformed into *para*-stereocenters. The symmetry plane in *trans*-3 is depicted. Associated configuration vectors are depicted above each structure.

Neither of the two stereocenters in the reactant **5** are affected by the sigmatropic [1,5]-H-shift reaction. Therefore, isomers of **5** with the isopropyl group in *cis*-position with respect to the methyl group of the chiral carbon lead to structures of **3** with the same relative positioning of both substituents. This principle is reflected during the reaction network generation procedure by adopting the same stereoparities of the stereocenters in the product structure of **3** as they are present in the respective reactant structure of **5**. The symmetry in the isomers of **3** leads to the creation of two *para*-stereocenters, corresponding to the atoms that were true-stereocenters in **3**. The four sigmatropic [1,5]-H-shift reactions that convert the four stereoisomers of **5** yield only two distinct stereoisomers of **3**, due to the symmetry plane in **3**. Therefore, the four configuration vectors representing the stereoparities of the two stereocenters of **3** can be clustered into two sets of two vectors, each set representing a diastereomer of **3**. For example, *trans*-**3** is represented by the set of configuration vectors $[-1, 1]$ and $[1, -1]$. The equivalence of two configuration vectors is verified through the construction of signed automorphism permutations as described by the algorithm for stereoisomer generation, cf. Eq. 8-1. As such, both *cis*-**1** enantiomers form the same product, i.e. *cis*-**3**, whereas both *trans*-**1** enantiomers form *trans*-**3**.

2 is further converted to **4** via the ene-cyclization reaction family. The chiral carbon atom of **2** is not affected by the cyclization reaction; hence the absolute configuration of this stereocenter is preserved in the product structures of **4**. In Genesys, this is reflected by copying the stereoparity of the stereocenter in **2** to the stereocenter in the product **4**. The emergence of a cyclopentane ring creates two additional stereocenters, in α and β position of the already existing stereocenter, cf. Scheme 8-14.



Scheme 8-14: The ene-cyclization reaction family creates two additional true-stereocenters in α and β of the already existing stereocenter. Associated configuration vectors are depicted above each structure.

Since the two newly created stereocenters in **4** did not exist in the reactant **2**, their respective stereoconfigurations are *a priori* undefined. The presented stereoisomer generation algorithm inspects the product structure for symmetry and generates all possible stereoisomers that arise from defining the stereoconfiguration of each of the previously undefined stereocenters. Each of the created stereocenters are true-stereocenters, and the structure of **4** does not possess any forms of symmetry. Therefore, the stereoisomer generation procedures leads to 2^2 distinct stereoisomer structures per linalool enantiomer, cf. R 8-2.

8.4.5 Kinetic model validation

The predictions of the generated kinetic model are compared against experimental data by Leiner et al. [57]. The reactor in the experiments consists of a 200 mm long quartz tube with a diameter of $15 \cdot 10^{-3}$ m. A more extended description of the reactor can be found in ref. [58]. Experimental conditions are summarized in Table 8-6.

Table 8-6: Experimental conditions in the experiments by Leiner et al. [57].

Temperature (K)	623-873
Pressure (10^5 Pa)	1.01
Conversion of feed (%)	0-100
Feed mole fraction at reactor inlet (ppm)	5000
N ₂ molar flow rate (10^{-3} mol s ⁻¹)	0.55

Temperatures vary between 623-873 K while the pressure is kept constant at $1.01 \cdot 10^5$ Pa. Temperature and pressure were quasi-uniform along the axial and radial reactor coordinate. Nitrogen is used as a diluent and dilution levels at the reactor inlet of 0.5 mol% of the feed are used. Calculated residence times are based on the reactor volume and the volumetric flow rates of the feed and the diluent at the reactor inlet, and range from 0.6-0.8 seconds. Species at the reactor outlet are identified and quantified using GC-FID and GC-MS. Components **1**, **2**, and **4** are identified and quantified, next to two dehydration products, with very low selectivities. *cis*-**1**,

trans-**1** and **2** were used as feeds in separate experiments. In the case of **1** as the feed, the latter consisted of a mixture of both enantiomers. In the case of **2**, the feed consisted of pure (R)-**2**. Impurities in the terpene feedstock typically comprised less than 5 mol% of the feed while their absolute and relative concentrations remained constant throughout the experiments. They are therefore neglected. Designation of the relative orientation of the substituents in **4** was determined by comparison of their retention times to previous published experiments with isomers of **4** [46, 59]. Thanks to the use of enantiomerically pure **2** in the experiments, the absolute configuration of the four isomers of **4** was established, assuming that the absolute configuration of the original chiral carbon in **2** remains unchanged.

Simulations were carried using the plug flow reactor model of the Chemkin 4.1 package [60], using the reported reactor dimensions. The feed composition at the reactor inlet was taken equal to the reported 0.5 mol% terpenes in the diluent nitrogen. In both the experiments with *cis*-**1** and *trans*-**1**, the feed of **1** was considered as a mixture of enantiomers, with a ratio of 63:37 for the (+)-enantiomer relative to the (-)-enantiomer of **1**, and originates from the reported ratio in ref. [57]. Flowrates of the terpene feed and diluent were taken equal to the reported values. Isothermal and isobaric operating conditions were used in the reactor simulations with values equal to the temperature and pressure reported for the experiments.

Figure 8-8 shows the mole fractions as a function of the reactor temperature for **1**, **2**, and **4** for experiments with *cis*- and *trans*-**1** as the feed. Since the feed was not enantiomerically pure, the mole fractions of enantiomeric species in the model were lumped together and the sum of the mole fractions of both enantiomers for the experiments and the model predictions was compared instead. Also, measured concentrations of minor products such as **3** were not reported, and thus could be used for comparison. The conversion of both diastereomers of **1** is slightly overestimated, resulting in a small overestimation of the mole fractions of **2**. Overall, good agreement was found between model and experiment, especially given that none of the parameters in the model were fitted to the experimental data.

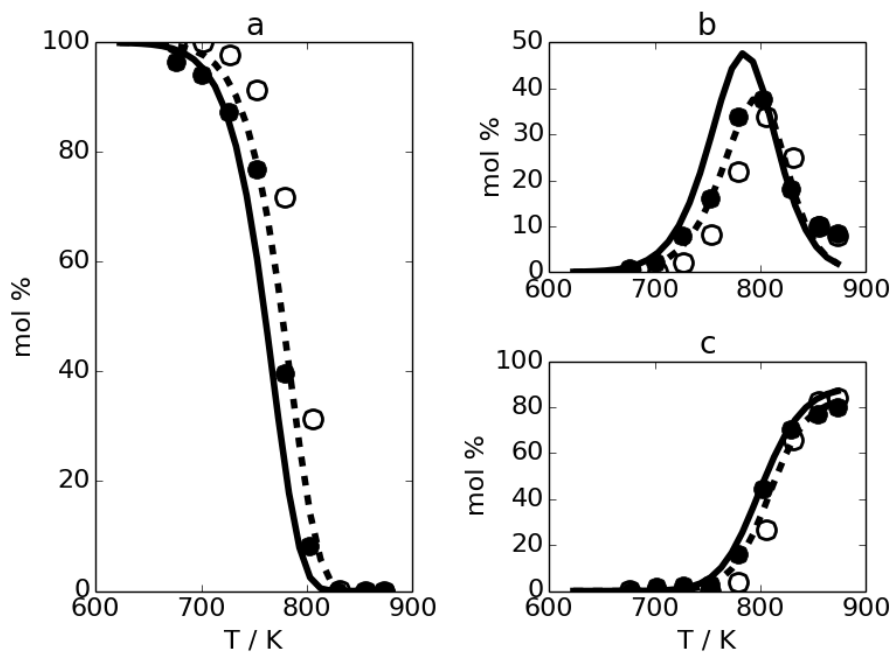


Figure 8-8: Mole fractions as a function of reactor temperature of a) 2-pinanol , b) linalool, c) sum of isomers of plinol. Model predictions: —: *cis*-2-pinanol as feed, - -: *trans*-2-pinanol as feed. Experiments from Leiner et al. [57]: ●: *cis*-2-pinanol as feed, ○: *trans*-2-pinanol as feed. Conditions: 5000 ppm feed at reactor inlet, carrier gas: N₂, total flow rate: $0.55 \cdot 10^{-3} \text{ mol s}^{-1}$.

Leiner et al. [57] also studied the thermal rearrangement of (R)-**2** and achieved the complete separation of the four isomers of **4**. As can be seen in Figure 8-9, good agreement was obtained between the model predictions and the measured mole fractions of the different isomers of **4**. At the highest conversions, formation of (R)-**4_D** is slightly overpredicted, at the expense of (R)-**4_A** and (R)-**4_B**. Nevertheless, the overall good agreement illustrates that the constructed kinetic model is capable of accurately reproducing the observed stereoselectivity of the ene-cyclization reaction of **2**.

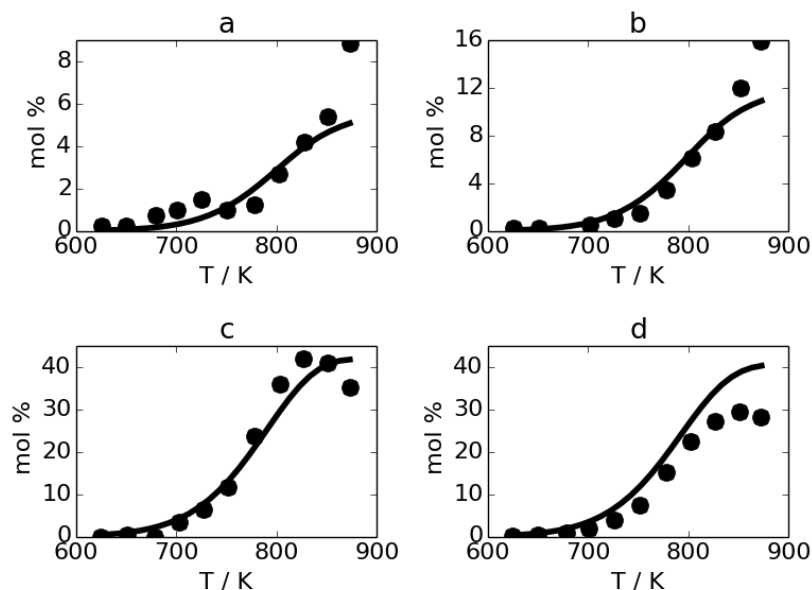


Figure 8-9: Mole fractions as a function of reactor temperature of a) (R)-**4_A** b) (R)-**4_B** c) (R)-**4_C** d) (R)-**4_D** with (R)-**2** as feed. •: experiments from Leiner et al. [57]. —: model predictions, Conditions: 5000 ppm feed at reactor inlet, carrier gas: N₂, total flow rate: $0.55 \cdot 10^{-3} \text{ mol s}^{-1}$, conversion: 0-100%.

8.5 Conclusions

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

The presence of topological symmetry in molecules may reduce the number of unique stereoisomers that arise from a structure with unspecified stereoparities. The detection of the equivalence of stereoisomers was addressed by an algorithm that determines the impact on the stereoparity of atom permutations that belong to the automorphism groups. Steric relations between substituents of stereocenters were determined by establishing the ordering of substituents around the central stereocenter. The conversion of the stereoparity into an array of ordered substituents was accomplished using the Cahn-Ingold-Prelog priority rules. This allowed the detection of chemically unrealistic stereoisomers that were created by the stereoisomer generation algorithm, and the introduction of stereoselectivity in kinetic model construction tools.

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

8.6 References

- [1] R. Noyori, Asymmetric Catalysis: Science and Opportunities (Nobel Lecture), *Angewandte Chemie International Edition*, 41 (2002) 2008-2022.
- [2] A. Corma, S. Iborra, A. Velty, Chemical routes for the transformation of biomass into chemicals, *Chemical Reviews*, 107 (2007) 2411-2502.
- [3] A. Ratkiewicz, T.N. Truong, Automated mechanism generation: From symbolic calculation to complex chemistry, *International Journal of Quantum Chemistry*, 106 (2006) 244-255.
- [4] S. Pierucci, E. Ranzi, A review of features in current automatic generation software for hydrocarbon oxidation mechanisms, *Computers & Chemical Engineering*, 32 (2008) 805-826.
- [5] F. Battin-Leclerc, E. Blurock, R. Bounaceur, R. Fournet, P.A. Glaude, O. Herbinet, B. Sirjean, V. Warth, Towards cleaner combustion engines through groundbreaking detailed chemical kinetic models, *Chemical Society Reviews*, 40 (2011) 4762-4782.
- [6] V. Warth, F. Battin-Leclerc, R. Fournet, P.A. Glaude, G.M. Côme, G. Scacchi, Computer based generation of reaction mechanisms for gas-phase oxidation, *Computers & Chemistry*, 24 (2000) 541-560.

- [7] N.M. Vandewiele, K.M. Van Geem, M.F. Reyniers, G.B. Marin, Genesys: Kinetic model construction using chemoinformatics, *Chemical Engineering Journal*, 207 (2012) 526-538.
- [8] W.H. Green, J.W. Allen, R.W. Ashcraft, G.J. Beran, C.A. Class, C. Gao, C.F. Goldsmith, M.R. Harper, A. Jalan, G.R. Magoon, D.M. Matheu, S.S. Merchant, J.D. Mo, S. Petway, S. Raman, S. Sharma, J. Song, K.M. Van Geem, J. Wen, R.H. West, A. Wong, H.S. Wong, P.E. Yelvington, J. Yu, RMG - Reaction Mechanism Generator v4.0, (2013).
- [9] E. Eliel, *Stereochemistry of Carbon Compounds*, 1962, New York, (1962) 188.
- [10] A.D. McNaught, A. Wilkinson, *Compendium of chemical terminology*, Blackwell Science Oxford, 1997.
- [11] E. Eliel, S. Wilen, L. Mander, *Stereochemistry of organic Compounds*, *Stereochemistry of Organic Compounds*, (1994).
- [12] B. Testa, *Principles of organic stereochemistry*, Dekker, New York, 1979.
- [13] R.S. Cahn, C. Ingold, V. Prelog, *Specification of Molecular Chirality*, *Angewandte Chemie International Edition in English*, 5 (1966) 385-415.
- [14] M. Perdih, M. Razinger, *Stereochemistry and Sequence Rules - A Proposal for Modification of Cahn-Ingold-Prelog System*, *Tetrahedron: Asymmetry*, 5 (1994) 835-861.
- [15] J. Gasteiger, *Handbook of Chemoinformatics*, Wiley-VCH Verlag GmbH & Co., Weinheim, 2003.
- [16] B. Rohde, *Representation and Manipulation of Stereochemistry*, *Handbook of Chemoinformatics: From Data to Knowledge in 4 Volumes*, (2003) 206-230.
- [17] C. Steinbeck, Y. Han, S. Kuhn, O. Horlacher, E. Luttmann, E. Willighagen, The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics, *J Chem Inf Comput Sci*, 43 (2003) 493 - 500.
- [18] C. Steinbeck, C. Hoppe, S. Kuhn, M. Floris, R. Guha, E.L. Willighagen, Recent developments of the Chemistry Development Kit (CDK) - An open-source Java library for chemo- and bioinformatics, *Current Pharmaceutical Design*, 12 (2006) 2111-2120.
- [19] S.R. Heller, S.E. Stein, D.V. Tchekhovskoi, InChI: Open access/open source and the IUPAC international chemical identifier, *Abstracts of Papers of the American Chemical Society*, 230 (2005) 60-CINF.
- [20] D. Weininger, SMILES, a chemical language and information-system. 1. Introduction to methodology and encoding rules, *Journal of Chemical Information and Computer Sciences*, 28 (1988) 31-36.
- [21] M. Razinger, K. Balasubramanian, M. Perdih, M.E. Munk, Stereoisomer generation in computer-enhanced structure elucidation, *Journal of Chemical Information and Computer Sciences*, 33 (1993) 812-825.
- [22] N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, Open Babel: An open chemical toolbox, *Journal of Cheminformatics*, 3 (2011).
- [23] H. Abe, H. Hayasaka, Y. Miyashita, S. Sasaki, Generation of Stereoisomeric Structures Using Topological Information Alone, *Journal of Chemical Information and Computer Sciences*, 24 (1984) 216-219.
- [24] C. Benecke, R. Grund, R. Hohberger, A. Kerber, R. Laue, T. Wieland, MOLGEN+, a generator of connectivity isomers and stereoisomers for molecular structure elucidation, *Analytica Chimica Acta*, 314 (1995) 141-147.
- [25] T. Wieland, A. Kerber, R. Laue, Principles of the generation of constitutional and configurational isomers, *Journal of Chemical Information and Computer Sciences*, 36 (1996) 413-419.
- [26] M.L. Contreras, J. Alvarez, D. Guajardo, R. Rozas, Algorithm for exhaustive and nonredundant organic stereoisomer generation, *Journal of Chemical Information and Modeling*, 46 (2006) 2288-2298.
- [27] J.G. Nourse, Applications of artificial intelligence for chemical inference. 28. The configuration symmetry group and its application to stereoisomer generation, specification, and enumeration, *J. Am. Chem. Soc.*, 101 (1979) 1210-1216.
- [28] R.G. Busacker, T.L. Saaty, *Finite Graphs and Networks: an introduction with applications*, McGraw-Hill New York, 1965.
- [29] M. Randic, Recognition of Identical Graphs Representing Molecular Topology, *J. Chem. Phys.*, 60 (1974) 3920-3928.
- [30] M. Randic, Discerning Symmetry Properties of Graphs, *Chem. Phys. Lett.*, 42 (1976) 283-287.
- [31] J. Ivanov, G. Schuurmann, Simple algorithms for determining the molecular symmetry, *Journal of Chemical Information and Computer Sciences*, 39 (1999) 728-737.
- [32] W. Chen, J. Huang, M.K. Gilson, Identification of symmetries in molecules and complexes, *Journal of Chemical Information and Computer Sciences*, 44 (2004) 1301-1313.
- [33] J. Ivanov, Molecular symmetry perception, *Journal of Chemical Information and Computer Sciences*, 44 (2004) 596-600.
- [34] A.R. Ashrafi, On symmetry properties of molecules, *Chem. Phys. Lett.*, 406 (2005) 75-80.
- [35] A.R. Ashrafi, M.R. Ahmadi, New computer program to calculate the symmetry of molecules, *Cent. Eur. J. Chem*, 3 (2005) 647-657.

- [36] G. Torrance, https://github.com/gilleain/cdk_signature, (2012)
- [37] A.C. Cope, R.D. Bach, trans-Cyclooctene, *Organic Syntheses*, (1973) 39-39.
- [38] J. Bredt, Über sterische Hinderung in Brückenringen (Bredtsche Regel) und über die meso-trans-Stellung in kondensierten Ringsystemen des Hexamethylens, *Justus Liebigs Annalen der Chemie*, 437 (1924) 1-13.
- [39] R.B. Woodward, R. Hoffmann, Conservation of Orbital Symmetry, *Angewandte Chemie-International Edition*, 8 (1969) 781-852.
- [40] R. Hoffmann, R.B. Woodward, Orbital Symmetry Control of Chemical Reactions, *Science*, 167 (1970) 825-831.
- [41] C. Mercier, P. Chabardes, Organometallic Chemistry in Industrial Vitamin-A and Vitamin-E Synthesis, *Pure Appl. Chem.*, 66 (1994) 1509-1518.
- [42] J. Nowicki, Claisen, Cope and related rearrangements in the synthesis of flavour and fragrance compounds, *Molecules*, 5 (2000) 1033-1050.
- [43] K.A.D. Swift, Catalytic transformations of the major terpene feedstocks, *Topics in Catalysis*, 27 (2004) 143-155.
- [44] K. Fahlbusch, Flavors and Fragrances, in: *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH, Weinheim, 2005.
- [45] G.W. Huber, S. Iborra, A. Corma, Synthesis of transportation fuels from biomass: Chemistry, catalysts, and engineering, *Chemical Reviews*, 106 (2006) 4044-4098.
- [46] G. Ohloff, E. Klein, Die Absolute Konfiguration des Linalools durch Verknüpfung mit dem Pinansystem, *Tetrahedron*, 18 (1962) 37-42.
- [47] H.T. Fritsch, P. Schieberle, Identification based on quantitative measurements and aroma recombination of the character impact odorants in a Bavarian Pilsner-type beer, *J. Agric. Food Chem.*, 53 (2005) 7544-7551.
- [48] H.-G. Schmarr, K.-H. Engel, Analysis and stereodifferentiation of linalool in Theobroma cacao and cocoa products using enantioselective multidimensional gas chromatography, *European Food Research and Technology*, 235 (2012) 827-834.
- [49] Daylight Chemical Information Systems, Daylight Theory Manual, 2010 <http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>.
- [50] N.M. Vandewiele, K.M. Van Geem, M.-F. Reyniers, G.B. Marin, Kinetic study of the thermal rearrangement of cis- and trans-2-pinanol, *Journal of Analytical and Applied Pyrolysis*, 90 (2011) 187-196.
- [51] J.L. Ripoll, Y. Vallee, Synthetic Applications of the Retro-Ene Reaction, *Synthesis*, (1993) 659-677.
- [52] T.H. Lay, J.W. Bozzelli, A.M. Dean, E.R. Ritter, Hydrogen-atom bond increments for calculation of thermodynamic properties of hydrocarbon radical species, *Journal of Physical Chemistry*, 99 (1995) 14514-14527.
- [53] W.D. Huntsman, V.C. Solomon, D. Eros, Reactions of Diolefins at High Temperatures. 2. The Cyclization of 1,6-octadiene and 7-Methyl-1,6-octadiene, *J. Am. Chem. Soc.*, 80 (1958) 5455-5458.
- [54] W. Oppolzer, V. Snieckus, Intra-Molecular Ene Reactions in Organic Synthesis, *Angew. Chem.-Int. Edit. Engl.*, 17 (1978) 476-486.
- [55] W. Pickenhagen, G. Ohloff, R.K. Russel, W.D. Roth, Intra-Molecular Ene Reaction - Thermal Rearrangement of Linalool and 1,2-Dehydrolinalool, *Helvetica Chimica Acta*, 61 (1978) 2249-2253.
- [56] S. Roy, K. Chakrabarty, G.K. Das, Comparative study on the transition structures of (3,4) and (3,5) ene cyclizations: A theoretical approach, *Journal of Molecular Structure-Theochem*, 820 (2007) 112-117.
- [57] J. Leiner, A. Stolle, B. Ondruschka, T. Netscher, W. Bonrath, Thermal Behavior of Pinan-2-ol and Linalool, *Molecules*, 18 (2013) 8358-8375.
- [58] A. Stolle, C. Brauns, M. Nuchter, B. Ondruschka, W. Bonrath, M. Findeisen, Thermal behaviour of selected C₁₀H₁₆ monoterpenes, *European Journal of Organic Chemistry*, (2006) 3317-3325.
- [59] H. Strickler, G. Ohloff, E.S. Kovats, Zur Kenntnis der Atherissschen Ole. 6. Die Thermische Cyclisation des (-)-(R)-Linalools. Die Struktur der Plinole und Einiger Derivate mit Iridane-gerüst, *Helvetica Chimica Acta*, 50 (1967) 759-797.
- [60] R.J. Kee, F.M. Rupley, J.A. Miller, M.E. Coltrin, J.F. Grcar, E. Meeks, H.K. Moffat, G. Lutz, A.E. Dixon-Lewis, M.D. Smooke, J. Warnatz, G.H. Evans, R.S. Larson, R.E. Mitchell, L.R. Petzhold, W.C. Reynolds, M. Caracotsios, W.E. Stewart, P. Glarborg, C. Wang, O. Adigun, W.G. Houf, C.P. Chou, S.F. Miller, P. Ho, D.J. Young, CHEMKIN Release 4.1.1, (2007).

Chapter 9: Conclusions and Perspectives

In this thesis, a range of computer-aided approaches were used and developed with the aim to facilitate the construction of kinetic models of complex chemical processes. The computer was used as an amplifier of human intelligence, rather than a creator of artificial, creative intelligence. As demonstrated in this thesis, the automated construction of kinetic models with a broad application range be it *via* rule-based or rate-based approaches, still requires a lot of expert involvement. Both Genesys and RMG also require the availability of high-quality chemical data for the construction of powerful kinetic models, and to this end experiments and *ab initio* calculations still remain a critical ingredient for the elucidation of the underlying chemistry of chemical processes.

One of the main focal points of this thesis was the relation between the three-dimensional structure of molecules and their respective reactivity. First, the study on the thermal rearrangement of the two diastereomers of 2-pinanol showed that reactivity differences of the reactants and selectivity towards desired products could be pinpointed to a single elementary reaction, i.e. the initial cyclobutane ring opening. For the first time, a detailed, microkinetic model was constructed for this process which was able to predict the product trends observed in the experiments. However, the kinetic model was constructed by hand not only because the limited complexity of the mechanism allowed it, but also because the available mechanism generation programs were unable to account for stereochemistry.

In the two subsequent chapters, a second process in which accounting for the polycyclic structure was critically important, was the thermal decomposition of Jet Propellant-10 (JP-10). The JP-10 pyrolysis experiments, carried out on the bench-scale reactor setup, were very useful in two aspects. They provided a large experimental data set of experimentally measured concentrations for over seventy species, based on state-of-the-art analytical techniques, which can be used for the validation of proposed kinetic models of JP-10 thermal decomposition. Moreover, besides the identification and quantification of already well-known products, a number of species such as tricyclo[5.2.1.0^{2,6}]dec-4-ene, unreported in literature, were quantified for the first time which contributed to improving the scientific community's understanding of the initial decomposition chemistry of JP-10 pyrolysis. The combination of rate-based mechanism generation techniques, on-the-fly computational chemistry calculations for thermochemical properties, and a limited number of high-level *ab initio* calculations for primary decomposition reactions of JP-10 derived radicals has shown to be a powerful methodology to build mechanisms for JP-10 pyrolysis. The newly created JP-10 pyrolysis model is the first, publicly available pyrolysis model that was validated over a range of experimental data from literature.

In the subsequent chapters, a novel mechanism generation tool, called Genesys, was developed, implemented, and validated. The premises of general applicability of the tool regardless of the intended application and the separation between chemical libraries and code prompted the use of techniques from chemoinformatics. Also, the modular approach adopted by Genesys, i.e. the segregation of chemical data libraries from the code, and the use of generic and generally applicable algorithms such as for symmetry identification, is a powerful technique that allows the integration of new modules with new functionalities, or to replace obsolete algorithms with more efficient and accurate solutions.

It was shown that a chemical graph representation of molecules and reactions together with graph-theoretic algorithms for species comparison, and substructure searching is a powerful combination for the generation of reaction networks, the calculation of thermochemical properties of molecules and rate coefficients of elementary reactions. With the advent of computationally feasible *ab initio* methods, the corpus of reliable chemical data derived from computational chemistry methods will continuously grow. The use of canonical, computer-

readable identifiers such as InChI, and well-accepted pattern languages such as SMARTS provided an efficient means for collecting, and disseminating chemical data for microkinetic modeling. As an example, the size of Genesys' databases containing thermochemical properties of molecules and Benson group additive values are both already in the order of $O(10^3)$. Moreover, new thermochemical data can be incorporated in the existing libraries with relative ease.

Group theory forms the basis for the calculation of entities related to symmetry and stereochemistry in this thesis. Not surprisingly, asymmetry and stereochemistry are often used in the same context. It was shown that the symmetry number of a molecule can be calculated based on the automorphism group order of the associated graph, corrected for the asymmetry induced by distinguishing between identical substituents. By using the graph associated with the transition state structure, the calculation of the reaction path degeneracy was retranslated into that of the symmetry of reactants and transition state structure. This formulation of the symmetry of molecules and reactions is fundamentally sound, and provides a solution for the problem of symmetry identification that is independent of the nature of the studied chemical structures.

The ultimate goal of creating compact, but detailed and predictive kinetic models with Genesys regardless of the involved chemical elements was illustrated for the thermal decomposition of diethylsulfide. For this application, rate coefficients of the employed reaction families were calculated solely through the use of ΔGAV° s. The good agreement between model predictions and experimental data could only be accomplished by expert knowledge on the chemistry and the availability of high level *ab initio* data for thermochemistry and kinetics.

In the next chapter, the capabilities of Genesys were extended to construct kinetic models that involved stereochemistry. In this context, automorphism groups found a second application domain, since they were used for the detection of equivalent stereoisomers. The example of the thermal rearrangement of 2-pinanol served as another application of Genesys and illustrated how concepts such as absolute configuration of stereocenters, and stereoselectivity of reactions could be accounted for by computer-manipulable algorithms, without the need for three-dimensional atomic coordinates.

In this thesis, it was illustrated how the use of computer-aided methods could be employed in the automated construction of kinetic models. However, several additions to the presented work can further strengthen the capabilities of Genesys. On the one hand, the current databases for the estimation of thermochemical properties of species and rate coefficients of elementary reactions can be improved and extended. On the other hand, new algorithms can be further developed for the use in mechanism generation applications. The suggestions below are a few examples that illustrate these two directions for improvement.

1. The quantitative structure-property relation (QSPR) methods that are presently available in Genesys are merely an illustration of the ease with which these techniques can be implemented. For kinetic modeling purposes many other QSPRs are relevant. For example, bond-additivity methods for the prediction of thermochemical properties can show superior performance compared to group-additive methods such as Benson in some cases. Other physical properties, such as critical temperature and pressure were estimated through group-additivity approaches as well, and could subsequently be used in determining Lennard-Jones potential well parameters. The latter can be used for problems in which transport phenomena compete with intrinsic reaction kinetics and hence, require the evaluation of physical properties such as viscosities and diffusion coefficients to model the system.
2. The performance of the JP-10 pyrolysis model can also be further improved. Intramolecular H-abstraction reactions were not systematically added in the model due to the lack of accurate kinetics for strained polycyclic hydrocarbons. It is possible that this reaction family can significantly impact the selectivity of the major pathways of the primary decomposition of JP-10. Ethenylcyclopentenenes are important primary products in the current model, while the chemistry describing the decomposition of these components is mostly missing. A better understanding of the secondary conversion of ethenylcyclopentenenes and its products may further improve the model as well.
3. The methodology for the calculation of symmetry of molecules and reactions assumes that stereocenters can adopt exactly two configurations, correcting the automorphism group order with a factor of two for each label-stereocenter. While this is true for the majority of stereocenters such as tetrahedral carbons and double bonds, there are specific

stereocenters that can adopt more than two configurations. Examples are octahedral molecular geometries such as in sulfurhexafluoride SF₆, or trigonal bipyramidal geometries for components such as phosphorus pentafluoride PF₅. For these molecules, symmetry numbers are likely to be erroneous; the methodology for these types of stereocenters therefore needs to be revised. As an example, instead of the binary configuration of ligands around a single chiral carbon atom, the six fluor atoms around the central sulfur atom in SF₆ can adopt 30 unique configurations.

Besides the thermal decomposition of diethylsulfide, the present library of group additive values in Genesys also allows studying other radical schemes involving hydrocarbons and organosulfur components. Future work can therefore focus on validating the influence of various sulfur additives on industrial processes such as the cracking of hydrocarbons. Ideally, the models developed in this thesis will assist in the development of new or in the optimization of existing industrial applications.

The kinetic model for the thermal decomposition of diethylsulfide generated by Genesys did not systematically include reaction families involving carbon radicals such as intramolecular H-abstraction reactions, intramolecular radical addition reactions and homolytic C-C scission reactions because values for ΔGAV° s for these reaction families were unavailable. Even more so, predictive methodologies for these reaction families are not yet available; the estimates by RMG for these types of reactions should also be carefully scrutinized. It is quite possible that the performance of the Genesys model can be improved by systematically including these reaction families as well. As a consequence, the derivation of ΔGAV° s for these reaction families based on high-quality quantum-chemistry calculations is an interesting opportunity for many applications in gas-phase free-radical chemistry.

A 2.5D-representation of molecules and transition states was used throughout Genesys and allowed the use of very generic graph-theoretic algorithms, but also of the identification of very specific steric relations. This 2.5D-representation of molecules can be enhanced even more, for example by distinguishing between singlet and triplet states of biradicals. Nevertheless, it was shown that regardless of the advanced internal representation of chemical structures, the exclusion of chemically meaningless structures such as *trans*-bicyclo[2.2.1]heptane after the

execution of the stereoisomer generation algorithm is still required. Additional chemical constraints could be implemented such as the Woodward-Hoffman rules for the conservation of orbital symmetry and may aid in improving the predictive capabilities of Genesys even more.

Appendix A: Experimental procedures for JP-10 pyrolysis experiments

Reactor set-up

The reactor set-up was extensively discussed elsewhere [1-3]. A schematic overview is given in Figure A-1.

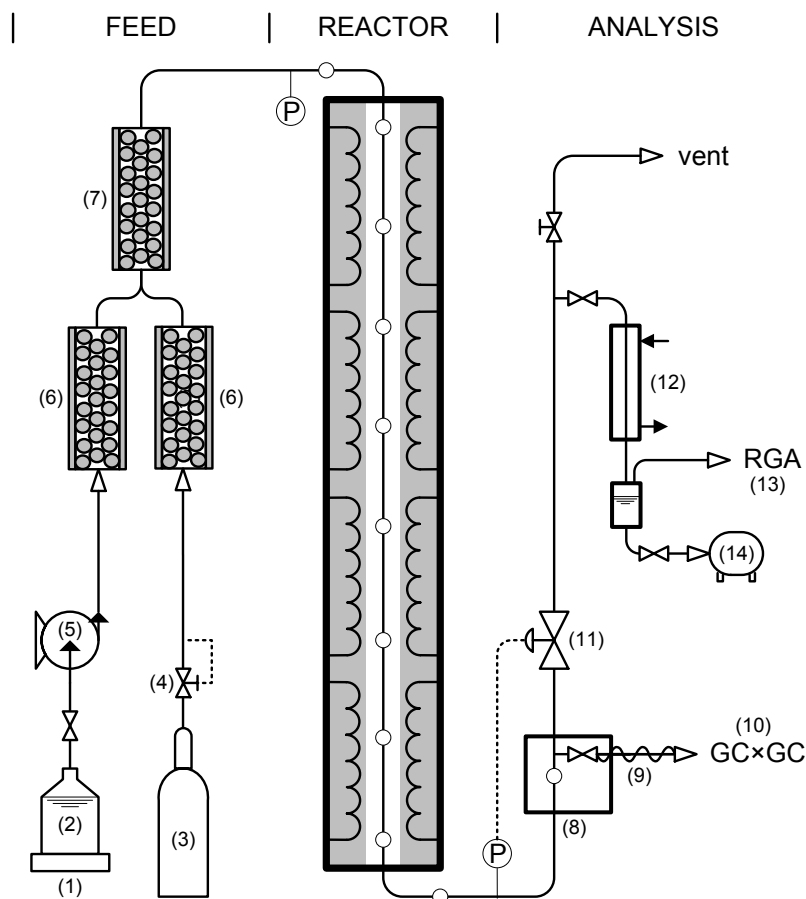


Figure A-1: Bench-scale reactor setup at the Laboratory for Chemical Technology

1: electronic balance, 2: JP-10 container, 3: gaseous diluent (N₂), 4: coriolis mass flow controller, 5: peristaltic pump, 6: evaporator, 7: quartz beads filled mixer, 8: heated sampling oven, 9: heated transfer lines, 10: GC×GC, 11: pressure regulator, 12: condenser, 13: refinery bas analyzer, 14: condensate drum.

JP-10 was fed by a peristaltic pump (Heidolph PD 5201, Germany) and the flow rate was determined by an electronic balance monitoring the mass of the JP-10 container. The liquid reactant (exo-TCD: boiling point 457K [4]) flows through a vaporizer kept at 523K filled with quartz pellets to allow a smooth evaporation. Gaseous nitrogen was heated parallel to JP-10 and mixed with it downstream of the vaporizer. The mass flow rate of nitrogen was controlled by a coriolis mass flow controller (Gefran - Bronkhorst Cori-Tech, mini CORI-FLOW, M1x). To ensure homogeneous gas-gas mixing, a mixing volume filled with quartz pellets was placed upstream of the reactor. The tubes from the evaporator/heater to the mixing unit were heated to prevent condensation. The reactor pressure was controlled by a valve which is placed

downstream of the reactor. The pressure drop over the reactor was negligible under normal operating conditions. Eight thermocouples were placed at the various positions along the reactor length to measure the process gas temperature, cf. Figure A-2.

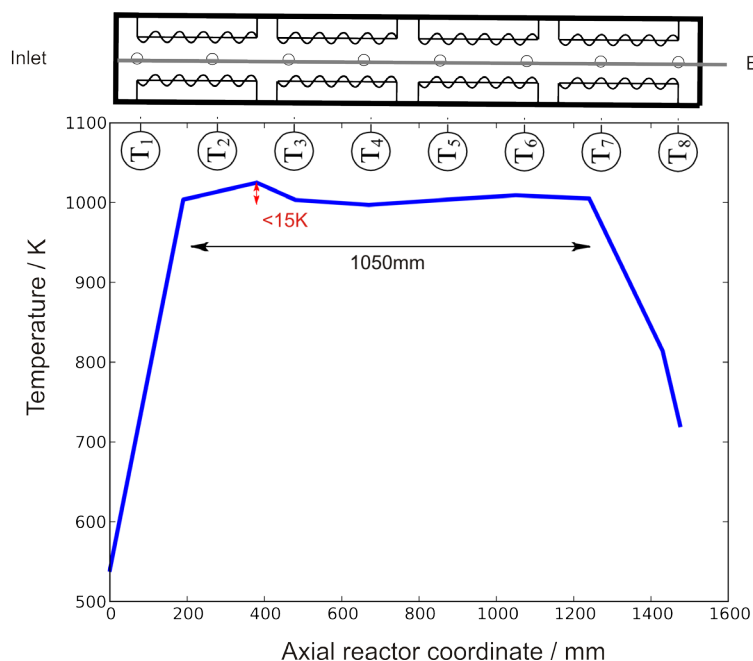


Figure A-2: Process gas temperature profile along the axial reactor coordinate measured by the eight thermocouples corresponding to a set temperature in the isothermal section of 1003 K.

The furnace is divided into four separate sections, controlled by four thermocouples. The feed mixture enters the reactor and was heated to the set temperature in a section of 0.200 m. A quasi-isothermal zone of 1.050 m is the main pyrolysis zone. The heating of the inlet gas causes a small temperature overshoot of the set temperature of less than 15 K in the first 0.100 m downstream of the heating section. Temperature variations along the remaining part of the isothermal pyrolysis zone remain within ± 5 K. The isothermal pyrolysis zone was followed by a zone of 0.200 m with a steep temperature decrease to 623 K maintained at the sampling box.

Analysis set-up

Comprehensive two-dimensional gas chromatography has shown to be a powerful tool for the on-line analysis of complex hydrocarbon mixtures [5]. The combination of an apolar (Rtx-1 PONA, dimethyl polysiloxane, 50m \times 0.25mm \times 0.5 μ m, Restek) and a more polar (BPX-50, 50% phenyl polysilphenylene-siloxane, 2m \times 0.15mm \times 0.15 μ m, SGE) gives rise to an enhanced separation resolution for structurally related components. Parallel to the effluent analysis at high temperature,

the reactor effluent was further cooled down to below 150 K by passing it through a water cooled condenser and cyclone to remove liquids and tars.

A TEMPUS ToF-MS (Thermo Scientific, Interscience) was used. Electron impact (EI) ionization was performed at 70 eV, a detector voltage of 1700V was applied and the acquisition frequency was set at 30 spectra s⁻¹ in a mass range of 35-400 amu. The data obtained with ToF-MS was acquired using thermo Scientific's XCalibur software. The raw GC×GC data files were processed using HyperChrom, i.e. the Chrom-Card extension for GC×GC data handling that enables 3D representation as well as the common color plot representation of the data. HyperChrom also allows automatic 3D peak quantification and identification. The latter was accomplished by cross referencing the measured mass spectra to the spectra in the available MS libraries.

Calibration factors for many important larger hydrocarbons (cyclopentene, benzene, toluene, ethylbenzene, xylenes, styrene, indene, naphthalene and exo-TCD) were experimentally determined by injecting known mixtures of these components. Calibration factors for other minor hydrocarbons were calculated using the effective carbon number concept [6].

References

- [1] M.R. Harper, K.M. Van Geem, S.P. Pyl, G.B. Marin, W.H. Green, Comprehensive reaction mechanism for n-butanol pyrolysis and combustion, *Combustion and Flame*, 158 (2011) 16-41.
- [2] K.M. Van Geem, S.P. Pyl, G.B. Marin, M.R. Harper, W.H. Green, Accurate High-Temperature Reaction Networks for Alternative Fuels: Butanol Isomers, *Industrial & Engineering Chemistry Research*, 49 (2010) 10399-10420.
- [3] K.M. Van Geem, A. Cuoci, A. Frassoldati, S.P. Pyl, G.B. Marin, E. Ranzi, An Experimental and Kinetic Modeling Study of Pyrolysis and Combustion of Acetone–Butanol–Ethanol (ABE) Mixtures, *Combustion Science and Technology*, 184 (2012) 942-955.
- [4] Y. Xing, W.J. Fang, W.J. Xie, Y.S. Guo, R.S. Lin, Thermal Cracking of JP-10 under Pressure, *Industrial & Engineering Chemistry Research*, 47 (2008) 10034-10040.
- [5] K.M. Van Geem, S.P. Pyl, M.F. Reyniers, J. Vercammen, J. Beens, G.B. Marin, On-line analysis of complex hydrocarbon mixtures using comprehensive two-dimensional gas chromatography, *J. Chromatogr. A*, 1217 6623-6633.
- [6] J. Beens, H. Boelens, R. Tijssen, J. Blomberg, Quantitative aspects of comprehensive two-dimensional gas chromatography (GC x GC), *Journal of High Resolution Chromatography*, 21 (1998) 47-54.

