CONCEPTUAL DESIGN OF BIOREFINERIES THROUGH THE
SYNTHESIS OF OPTIMAL CHEMICAL-REACTION PATHWAYS

A Thesis
by
ERIC JAMES PENNAZ

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2011

Major Subject: Chemical Engineering
CONCEPTUAL DESIGN OF BIOREFINERIES THROUGH THE
SYNTHESIS OF OPTIMAL CHEMICAL-REACTION PATHWAYS

A Thesis

by

ERIC JAMES PENNAZ

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee, Mahmoud El-Halwagi
Committee Members, Carl Laird
Guy Curry
Head of Department, Michael Pishko

August 2011

Major Subject: Chemical Engineering
ABSTRACT

Conceptual Design of Biorefineries Through the
Synthesis of Optimal Chemical-Reaction Pathways.

(August 2011)

Eric James Pennaz, B.S., University of Chicago;
M.A., City College of the City of New York
Chair of Advisory Committee: Dr. Mahmoud El-Halwagi

Decreasing fossil fuel reserves and environmental concerns necessitate a shift
toward biofuels. However, the chemistry of many biomass to fuel conversion pathways
remains to be thoroughly studied. The future of biorefineries thus depends on developing
new pathways while optimizing existing ones. Here, potential chemicals are added to
create a superstructure, then an algorithm is run to enumerate every feasible reaction
stoichiometry through a mixed integer linear program (MILP). An optimal chemical
reaction pathway, taking into account thermodynamic, safety, and economic constraints
is then found through reaction network flux analysis (RNFA). The RNFA is first
formulated as a linear programming problem (LP) and later recast as an MILP in order to
solve multiple alternate optima through integer cuts. A graphical method is also
developed in order to show a shortcut method based on thermodynamics as opposed to
the reaction stoichiometry enumeration and RNFA methods. A hypothetical case study,
based on the conversion of woody biomass to liquid fuels, is presented at the end of the work along with a more detailed look at the glucose and xylose to 2-methyltetrahydrofuran (MTHF) biofuel production pathway.
DEDICATION

To:

My Parents,

My Brother,

From Hawaii, with Love
ACKNOWLEDGEMENTS

I would like to thank Dr. Mahmoud El-Halwagi for his very generous support and guidance of my research. I would also like to thank the members of my advisory committee, Dr. Carl Laird and Dr. Guy Curry, for their very important feedback on the direction of my research. Special thanks also go out to my office mates, Kerron Gabriel and Mohamed Noureldin, for their very valuable feedback at every step of the research process. They have all been truly supportive.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Reaction Network Synthesis</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Overview of Biomass Conversions to Chemicals</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Overview of the Thesis</td>
<td>2</td>
</tr>
<tr>
<td>1.4 Motivation</td>
<td>4</td>
</tr>
<tr>
<td>2. LITERATURE REVIEW</td>
<td>7</td>
</tr>
<tr>
<td>3. PROBLEM STATEMENT</td>
<td>11</td>
</tr>
<tr>
<td>4. METHODOLOGY AND APPROACH</td>
<td>13</td>
</tr>
<tr>
<td>4.1 Overview of the Approach</td>
<td>13</td>
</tr>
<tr>
<td>4.2 Compound Identification (Classify All Potential Molecules)</td>
<td>15</td>
</tr>
<tr>
<td>4.3 Reactions Optimization Model (Generate Feasible Reactions)</td>
<td>17</td>
</tr>
<tr>
<td>4.4 RNFA Optimization Model (Screen Reaction Paths)</td>
<td>24</td>
</tr>
<tr>
<td>4.5 Graphical Targeting Approach</td>
<td>29</td>
</tr>
<tr>
<td>5. CASE STUDY</td>
<td>35</td>
</tr>
<tr>
<td>5.1 Building a Starting Point</td>
<td>35</td>
</tr>
<tr>
<td>5.2 More Detailed Reactant to Product: Biomass to MTHF</td>
<td>39</td>
</tr>
<tr>
<td>6. RESULTS AND DISCUSSIONS</td>
<td>46</td>
</tr>
<tr>
<td>7. CONCLUSIONS</td>
<td>47</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>48</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>52</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>80</td>
</tr>
<tr>
<td>APPENDIX C</td>
<td>81</td>
</tr>
<tr>
<td>APPENDIX D</td>
<td>87</td>
</tr>
<tr>
<td>VITA</td>
<td>90</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>The motivation pyramid</td>
<td>5</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>An interlocking of the processes</td>
<td>6</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>The reaction pathways</td>
<td>11</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>The stepwise method</td>
<td>14</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>The compound hierarchy</td>
<td>16</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Superstructure of compounds</td>
<td>17</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Superstructure of compounds with reactions</td>
<td>23</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Superstructure of compounds with the best reaction</td>
<td>29</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>Example pathway targeting and reactions</td>
<td>33</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Reaction Synthesis Tree, based on Huber</td>
<td>37</td>
</tr>
<tr>
<td>Figure 5.2</td>
<td>The enumerated pathways from glucose and xylose to MTHF</td>
<td>40</td>
</tr>
<tr>
<td>Figure 5.3</td>
<td>The targeting of optimal pathways</td>
<td>43</td>
</tr>
<tr>
<td>Figure 5.4</td>
<td>The targeting of two alternate pathways</td>
<td>44</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

1.1 Reaction Network Synthesis

Optimizing biorefinery production through chemical reactions involves finding the best chemical route from biomass to bio-products. While some methods to predict chemical routes use mathematical models and others use qualitative techniques, they both achieve the same objective: to find the best route from reactants to products. Reaction network synthesis is a combinatorial approach that finds all of the reaction routes from reactants to products based on inputted compounds and ultimately selects the best pathway. While there may be many reaction steps, or there may be multiple inputs or outputs for the reactants, intermediates, and products, there is a best pathway that can be generated based on mathematical rules.

The topic has been extensively studied for many years, even before the advent of the computer and logic based methods. The increase in available computing power has, however, allowed for an exhaustive search of many more compounds than was previously available. Where there once was a preference for qualitative methods, there is now a preference for more logic based, mathematical rule sets. Reaction network synthesis can therefore be formatted as an optimization problem and solved based on rigorous mathematical targets and formulations.

This thesis follows the style of Clean Technologies and Environmental Policy.
The preferred method involves finding a superstructure, which is a total amalgamation of all of the reactants, intermediates, and products in the system. From this superstructure then the reactions are enumerated and then a best pathway through the system is chosen based on a number of factors, such as thermodynamic, economic, safety, and heuristics.

1.2 Overview of Biomass Conversions to Chemicals

With increasing demand for fuel in the future and a dwindling supply of nonrenewable fuels, demand will soon outstrip supply. Biomass conversion into renewable fuels therefore needs to be considered when it comes to selecting a replacement for current nonrenewable chemical and fuel sources.

Biomass includes any organic material that grows through photosynthesis. Typical biomass includes corn stover, algae, and woody biomass such as trees. Fuels include octane, kerosene, and 2-methyltetrahydrofuran (MTHF) which is a fuel additive.

There are many available methods to convert biomass into fuels and chemicals, with gasification, pyrolysis, and catalytic upgrading being the most popular methods.

1.3 Overview of the Thesis

The objective of this work is to find new and potentially more favorable chemical reaction pathways in biorefineries. While the search space has contained the molecules
selected to those which would be most suited to biorefineries, there is applicability of the model across all chemical domains, especially in biomass to fuel studies. If there were another chemical system that required a best chemical pathway from reactants to products, this model would be applicable as well, there would only have to be changes in the constraints and starting species indicated in the superstructure.

The thesis starts with an introduction and literature review, and then discusses the problem statement and methodology. It then moves into a case study, results, and conclusion. The case study involves the production of fuels from glucose derived from biomass. A mixed integer linear program (MILP) model is formulated in the reaction enumeration stage to solve for the optimal reactions in a biorefinery. Integer cuts are then be done to the MILP to enumerate all of the reactions in a biorefinery. The reaction network flux analysis and optimization can be formulated as a linear program (LP) and solved for a global solution. The alternate optimal pathways can then be solved by converting the LP to a MILP and solving multiple times via integer cuts. A graphical method, used as a shortcut tool rather than the more rigorous methods of chemical reaction enumeration and reaction network flux analysis (RNFA) to determine the best reaction pathway is then proposed based on the results of the first two optimization problems of reaction enumeration and RNFA.
1.4 Motivation

When a chemical reaction is set for a certain chemical process, it is very difficult if not impossible to change the chemistry for a process. Therefore, it becomes imperative to set a chemical process that is favorable from the beginning, as the chance to have the largest impact on chemical processes occurs during the initial planning and development stages of the industrial plant. From inception, the chemical reactions that occur in the plant determine what temperatures and pressures the chemical plant is operated at, what separations have to be done, and what the theoretical target for the chemical process can be. It also allows for the largest impact on the economics of a process. For example, designing a chemical reaction that minimizes waste has a chance to have the largest impact on the chemical process through prevention. This method beats process minimization, recycling, energy recovery in the form of heat exchanger networks, and disposal of waste products as shown in Figure 1.1.
Figure 1.1: The motivation pyramid. The pyramid shows how chemical reactions have the ability to have the largest impact on chemical process design, that of prevention. The arrow on the left shows the relative value of each process.

The chemical reaction pathways have a large impact on the economic potential of the process design. Setting theoretical maximum targets for the benefit of a chemical process is the realm of chemical reaction pathways.
Figure 1.2: An interlocking of the processes. The chemical reaction pathways influence the industrial process and the economic potential both ways.

The necessity of finding better chemical reaction pathways for biorefineries can be seen in the economics of the current processes as well as the industrial processes and chemical reaction pathways shown interlined in Figure 1.2. For renewable fuels produced in biorefineries to become competitive, better chemical reaction pathways would certainly help to bring the economics into a more competitive realm.
2. LITERATURE REVIEW

Dwindling natural resources have led to a need for alternate fuels and alternate chemical products (Shafiee and Topal 2009). Biorefineries present an opportunity to take a natural renewable resource such as plant or waste material and convert it into a useful product, such as a fuel or an added value chemical (Pokoo-Aikins, Heath, Mentzer, Mannan, Rogers et al. 2010). The future directions of biorefineries and technologies rely on streamlining the current production pathways and finding more integrated methods to produce products (Fernando, Adhikari, Chandrapal and Murali 2006). The optimization of biorefineries has thus been an ongoing process, through both hierarchal and economic modeling (Sammons Jr, Yuan, Eden, Aksoy and Cullinan 2008; Ng, Pham, El-Halwagi, Jiménez-Gutiérrez and Spriggs 2009). Development of sustainability has been key in many papers. Selecting the appropriate chemical reactants, intermediates and products is essential to this optimization (Kohse-Höinghaus, Oßwald, Cool, Kasper, Hansen et al. 2010). Determining these reactants, intermediates, and products is the realm of chemical reaction pathway synthesis.

Developing reaction pathways are the first step in determining the proper routes for a chemical process given a certain reactant and expecting a certain product. Selection of chemical reaction pathways involves selecting the best route from reactants to products while meeting certain criteria in-between. These criteria can be divided into a number of categories, such as economic, thermodynamic, safety, process tasks, separations, and handling constraints. While research on chemical reaction networks has
been ongoing for decades, Huber’s work was seminal in the field in representing biorefinery chemical systems (Ugi and Gillespie 1971). Siirola had also previously tied in reaction networks with process tasks (Siirola and Rudd 1971). Historically, two approaches have been taken for the synthesis of chemical reaction pathways (Agnihotri and Motard 1980):

(I) The Information Centered or Direct-Associative, and

(II) Logic Centered methods.

Information centered approaches rely on using data to bring together subunits of chemical reactions that are already known and then to synthesize a pathway. This approach may be limited in its scope. The second approach, logic centered methods relies on multiple intermediates that form a synthetic tree and more abstract representation (Nishida, Stephanopoulos and Westerberg 1981; Rotstein, Resasco and Stephanopoulos 1982).

The synthesis of chemical reaction paths was later expanded to cover (ΔG, T) space in a primitive synthesis procedure (Rotstein, Resasco and Stephanopoulos 1982). Beard then extended the use of thermodynamics in complex metabolic networks while taking into account thermodynamic considerations with Energy Balance Analysis (EBA), also applicable to any chemical system (Beard, Liang and Qian 2002).

Fornari expanded the synthesis of chemical reaction paths to include two degrees of freedom (Fornari, Rotstein and Stephanopoulos 1989). Retrosynthetic analysis and complete reactant to product production was improved upon by Johnson. through the LHASA program (Johnson and Marshall 1992). Adding group contribution methods
and economic considerations into the reaction paths was done by Fornari and Stephanopoulos 1994a and 1994b (Fornari and Stephanopoulos 1994a; Fornari and Stephanopoulos 1994b). Methods for environmental impact minimization (MEIM) became the prime focus in a process route a chemical route after incidents, and was later expanded to solvent design and reaction paths (Crabtree and El-Halwagi 1994; Stefanis, Buxton, Livingston and Pistikopoulos 1996).

Reducing the size of the problem became a main focus for environmental impact minimization later on, with a focus on co-material design and structural restrictions that allowed only certain reactions to progress. Computer aided molecular design (CAMD) was pivotal in this role and was based on this co-material design approach (Buxton, Livingston and Pistikopoulos 1997).

A complete algorithmic solution to superstructures of chemicals and reactions through an MILP method, including thermodynamic, economic, safety, and chemical plant was well reviewed by Buxton, Hugo, Livingston and Pistikopoulos 2002a and 2002b with many examples given (Buxton, Hugo, Livingston and Pistikopoulos 2002a; Buxton, Hugo, Livingston and Pistikopoulos 2002b). Various optimization approaches to chemical reaction networks have also been proposed (Li, Hu, Li and Shen 2000; Majumdar and Mitra 2004). Thermodynamic descriptions and kinetic modeling have also been proposed (Hatzimanikatis, Li, Ionita and Broadbelt 2004; Hatzimanikatis, Li, Ionita, Henry, Jankowski et al. 2005).

Using metabolic networks as a foundation to find the optimal output of a reaction network, the LP of RNFA was turned into an MILP in order to find multiple alternate
optima (Lee, Phalakornkule, Domach and Grossmann 2000). The methods proposed by Lee and others are the foundation of using multiple alternate optima of RNFA in chemical networks (Beard, Liang and Qian 2002; Stelling, Klamt, Bettenbrock, Schuster and Gilles 2002; Lee, Yun, Park and Lee 2003; Sauer 2006). Alternate improved formulations have also been proposed in literature (Murabito, Simeonidis, Smallbone and Swinton 2009). RNFA was also later applied by Besler to chemical synthesis (Besler, Harwardt and Marquardt 2009). Hechinger then expanded this to simultaneous product and process design using CAMD and quantitative modeling of biofuel products (Hechinger, Voll and Marquardt 2010).

An intelligent way to break down complicated biomass structures and direct the subsequent molecules into products may be a way to circumvent costly or unnecessary thermodynamic and oxidative changes. As Szmant points out, there are six criteria for a successful organic chemical plant (Szmant 1989). The favorable demand, reliable supply, technological know-how, profitability, diversification potential, and merchandizing potential are all essential to the successful implementation of organic chemical plant, and by extension, biorefineries. For example, being able to skip current energy wasting steps that are present in a biorefinery may be useful in improving the economics or conversion yields of a process. While the catalytic or enzymatic tools to accomplish this are left to future research, the reaction path synthesis roadmap from biomass to liquid fuels is laid out in the following steps.
3. PROBLEM STATEMENT

The problem is stated as follows: Given a desired product or set of products (e.g., fuels, specialty chemicals), a set of biomass reactants / feedstocks, and possible intermediates, it is desired to develop systematic procedures for the synthesis of optimal reaction pathways from reactant(s) to product(s) through any number of steps and intermediates as shown in Figure 3.1.

**Figure 3.1:** The reaction pathways. The reactants and products of a process are known. The question is what intermediates to use and how many steps there are between each of the reactant to intermediates and intermediates to products.
The objective is to aid the engineers during the early stage of conceptual design of a biorefinery by generating promising reaction pathways upon which flowsheet alternatives can be constructed. The objective of the pathways synthesis is to maximize a given criterion, such as profit or yield, in a biomass to renewable fuel/chemical biorefinery. Then, based on the combined criteria of the building a chemical pathway, screening and optimization, it is desired to find a way to represent the best reaction pathways in a simplified graphical form. Thermodynamic, technical, and economic criteria are to be used in the screening of alternatives.
4. METHODOLOGY AND APPROACH

4.1 Overview of the Approach

Starting with the reactants, intermediates, and products, one identifies the compounds to be used in the chemical reaction pathways. The method involves first developing a superstructure of molecules that can be found from a literature search, existing chemical facilities, chemical databases such as DIPPR, or heuristic searches.

The next step is to use an established chemical synthesis route, or a few typical chemical process industry routes to build this superstructure of chemical compounds. The current steps in an established or industrial chemical reaction pathway are used as a base case scenario from which all other subsequent materials are enumerated. Starting with this base case, one adds a number of molecules that take the place of the black boxes in this process and that can also serve as alternative feasible reactions.

Additional molecules are added in order to provide the opportunity for alternative synthetic routes. These additional molecules are selected based on chemical factors, heuristics and economic knowledge. An overall superstructure is then generated which allows for the identification of chemical reactions and routes.

Optimal reactions are then enumerated through the use of an algorithm and optimization model based on the given constraints that takes into account stoichiometric mass balances. The reactions are generated and screened out based on thermodynamics and heuristics. The next step, RNFA, involves finding the optimal way to go from a
given reactant such as a biomass to a given liquid fuel, and takes into account economic, yield, safety, and thermodynamics constraints. Both the reaction optimization and the RNFA optimization models are formulated in the LINGO® mathematical formulation system. Finally, a graphical method is shown for the chemical potential versus the reaction step number in order to show a shortcut method in the process.

The five step process is shown below as in Figure 4.1.

**Figure 4.1:** The stepwise method. The method involves these five steps in order to generate the best and most feasible chemical reaction pathways in a biorefinery.

The thermodynamic data comes from a variety of sources, Yaw’s handbook, the DIPPR database, and the group contribution methods of Gani and Marrero for any compounds that have are not in the databases (Marrero and Gani 2001). The cost data is
taken from the cost of the industrial chemicals ICIS database, 2006. Only starting reactants and products price are taken into account, and the overall profit of the reaction pathway taken as the basis of the sum of the products minus the sum of the reactants.

4.2 Compound Identification (Classify All Potential Molecules)

Through a combination of heuristics, biomass to liquid fuel literature papers, and databases, the compounds have been identified. Different types of searches for compounds to be used in the overall superstructure of molecules leads to different numbers of total compounds that can be identified. If rules of thumb and heuristics are used, there may be a relatively small number of compounds identified, around 25. If one were to then look in the literature of all biorefinery compounds that have been considered as potential molecules in chemical reaction paths, there are around 100 compounds that readily become available. Expanding further and looking in an organic compound database such as the DIPPR database, the number of compounds increases to the limit of the number of compounds that are in the database. The DIPPR database contains approximately 1500 compounds that can be considered for addition to the superstructure. If the database search space is limited to the carbon, hydrogen, and oxygen search space, then there are fewer, around 1100 compounds that could be used in the search. It must be noted that some of the compounds identified in the literature and heuristics may not be available in the DIPPR database and vice versa. Therefore it is useful to have an overlapping search space.
Finally, if one were to move to carbon, hydrogen, and oxygen molecular generation, computer aided molecular design (CAMD), to obtain all molecules that are between certain molecular weights, or certain number of atoms, such as 2 to 200 atoms, then 10000 plus compounds become available. In this work, a combination of the first three approaches, heuristics, literature, and databases has been used to select the molecules for addition to the superstructure as shown in Figure 4.2.

**Figure 4.2:** The compound hierarchy. The number of compounds that can be found in the Carbon, Hydrogen, and Oxygen search space varies with the method you choose, from 10000 plus compounds for a certain molecular generation approach, to approximately 25 compounds for rules of thumb.

Once the compounds have been identified, they are put into a superstructure and the feasible chemical reactions can then be enumerated as shown in Figure 4.3.
Figure 4.3: Superstructure of compounds. All potential molecules are now classified as reactants, intermediates, or products.

It should be noted that the intermediates here are not arranged in a stepwise fashion with layers of intermediates, but rather form a superstructure of all materials that can react with each other are formed.

4.3 Reactions Optimization Model (Generate Feasible Reactions)

This is an algorithm that is run over the course of all possible stoichiometric combination reactions and the feasible solutions are enumerated that pass the thermodynamics, process, economic, heuristic, and safety factors. The data for each reaction stoichiometry and associated terms of thermodynamics, process economics, etc., is stored and later used in the RNFA step. Notably, no reaction in this step is explicitly
refused based on thermodynamics. The individual reaction step may be thermodynamically infeasible, but the overall reaction may be thermodynamically feasible.

The objective function is to enumerate the total number of reactions from least number of reactants and products to most:

\[
\text{Min } i_s + ii_s, \quad \forall s \in S
\]  

(1)

The \(i_s\) variable represents the product flag for a compound. The \(ii_s\) variable represents the reactant flag for a compound. For example, the binary variable \(i\) is the product flag which is set to 1 if the species is a product, and set to 0 if the species is either a reactant or will not participate in the reaction. The binary variable \(ii\) is the reactant flag which is set to 1 if the species is a reactant, and set to 0 if the species is a product or does not participate in the reaction. The first reactions enumerated will be isomerizations and later steps will involve, for example, one reactant to two products, or two reactants to one product, and continue on from there. As this model is an MILP, integer cuts are added in order to enumerate multiple solutions.

The first constraint is to add an atomic balance:

\[
\epsilon_s V_s = 0, \quad \forall s \in S
\]

(2)

The atom balance is the starting point in any chemical reaction. \(\epsilon_s\) is the \(E*S\) atomic matrix and \(V_s\) is the \(S*1\) column vector of stoichiometric coefficients \(v_s\). The atom balance is necessary in order to follow mass balance principles. There must, for example, be the same number of total carbons, hydrogen’s, and oxygen’s in the reactants and the products.
Whole number stoichiometries products constraints:

\[ v_p \geq 1, \]  \hspace{1cm} (3)

This says that the stoichiometric numbers of the coefficients are greater than or equal to one. This also forces the reactants to be whole number stoichiometries based on further constraints that will be imposed. It is also necessary to define a dummy variable to prevent the formation of irrational numbers in the calculation of the stoichiometry coefficients:

\[ x_s \]  \hspace{1cm} (4)

Set \( x_s \) as a dummy variable which is positive and continuous.

The stoichiometric coefficients are related to the dummy variables as follows:

\[ v_s = x_s - 2 \times x_s \times i_{is}, \quad \forall \ s \in S \]

This equation can be recast to avoid the bilinear term of \( x_s \times i_{is} \), using a transformation of the dummy variables:

\[ v_s = x_s - 2 \times y_s, \quad \forall \ s \in S \]  \hspace{1cm} (5)

\[ y_s - v_{max} \times i_{is} \leq 0, \quad \forall \ s \in S \]  \hspace{1cm} (6)

\[ x_s + v_{max} \times (i_{is} - 1) - y_s \leq 0, \quad \forall \ s \in S \]  \hspace{1cm} (7)

\[ y_s - x_s \leq 0, \quad \forall \ s \in S \]  \hspace{1cm} (8)

The dummy variable \( y \) has to be related to the reactant flag:

\[ y_s \geq i_{is}, \quad \forall \ s \in S \]  \hspace{1cm} (9)

The dummy variables and equations are provided for two reasons, that the stoichiometric coefficients are larger than one for the product, and that only rational numbers and non-irrational numbers are present in the stoichiometric coefficients.
Dummy variable constraints:

\[ x_s - y_s - v_{\text{max}} \ast I_s \leq 0, \quad \forall \ s \in S \]  \hspace{1cm} (10)

\[ x_s - y_s - (v_{\text{max}} + 1) \ast I_s + v_{\text{max}} \leq 0, \quad \forall \ s \in S \]  \hspace{1cm} (11)

where \( v_{\text{max}} \) is set to 50. Equations allow for the dummy variables to be smaller than some \( v_{\text{max}} \), or maximum value. This means that the stoichiometric coefficients don’t go above a value of approximately 50 and become unwieldy.

Role Specification Constraint:

\[ i_s + ii_s + iii_s = 1, \quad \forall \ s \in S \]  \hspace{1cm} (12)

The binary variable \( iii \) is the participation flag. If the chemical participates, meaning it is either a reactant or a product, the \( iii \) is set to 0. If the chemical does not participate, the value of \( iii \) is set to 1. A chemical cannot be both a reactant and a product. While it may occur that some chemicals can be both reactants and products in chemical reactions, for example \( \text{H}_2\text{O} \) in pyrolysis, the overall value of the chemical is taken as the flag for that chemical. Meaning that if there is, for example, stoichiometrically more \( \text{H}_2\text{O} \) on the reactants than products side, then \( \text{H}_2\text{O} \) is taken as a reactant and vice versa.

Constraints limiting total number of reactants and products:

\[ \sum_s i_s \leq N_{p_{\text{max}}}, \quad \forall \ s \in S \]  \hspace{1cm} (13)

Equation (13) serves to limit the number of different species that can be products. Most reactions are limited to a total of three for the \( N_{p_{\text{max}}} \).

\[ \sum_s ii_s \leq N_{r_{\text{max}}}, \quad \forall \ s \in S \]  \hspace{1cm} (14)
Equation (14) serves to limit the number of different species that can be reactants. Most reactions are limited to a total of three for the $N_r^{\text{max}}$.

\[ \sum_s (i_s + ii_s) \leq N_{\text{spe}}^{\text{max}}, \quad \forall \, s \in S \]  

(15)

\[ v_s \leq v_s^{\text{max}}, \quad \forall \, s \in S \]  

(16)

These equations are presented so that the number of reactants and products in a given reaction are set to some maximum. $N_r^{\text{max}}$ and $N_p^{\text{max}}$, 3 as stated before, while $N_{\text{spe}}^{\text{max}}$ is usually set to 6. This means that there can at most be 3 reactants and 3 products in a given chemical reaction. Constraints 13, 14 and 15 are not necessary together, either 13 and 14, or just 15 will usually suffice. Using constraints 13 and 14 is more specific than using just constraint 15. The $v_s^{\text{max}}$ is normally set to what the system requires, meaning the stoichiometric coefficient for a balanced reaction is usually sufficient to be less than 50.

Certain chemistry reaction or product composition constraints are also possible (For example: species a and b must not react together):

\[ ii_a + ii_b \leq 1 \]  

(17)

The thermodynamics of the chemical reaction system is also taken into account. The Gibbs energy, enthalpy, enthalpy and operating temperatures for the system are based on either the data from the DIPPR database or the functional group approximations of the Marrero and Gani method (Marrero and Gani 2001). The following thermodynamics constraints are part of the functional group approximation methods used in the Gani and Marrero methods which allow for the approximation of thermodynamic properties.

Enthalpy of Formation:
\[ \Delta H_f^o \, (\text{J/mol}) \]  
(18)

Entropy of Formation:
\[ \Delta S_f^o \, (\text{J/mol*K}) \]  
(19)

Gibbs free energy of Formation, either found explicitly through the DIPPR database, Gani method, or enumerated through the other parameters with the following equation:
\[ \Delta G_f^o = \Delta H_f^o - T_{\text{oper}} \ast \Delta S_f^o \, (\text{J/mol}) \]  
(20)

The operating temperature is thus restricted:
\[ 300 \leq T_{\text{oper}} \leq 1000 \]  
(21)

The operating temperature for a chemical reaction must be researched in the relevant literature. If no data is available, the constraint must be either approximated or ignored.

Toxicity is also included as the safety factor for each compound:
\[ -\log (\text{LC50}_s), \quad \forall \ s \in S \]  
(22)

These factors are all tabulated and screened after the reactions are generated, then the economics are also included in the following equation:
\[ \sum_s iii_s \ast (i_s - ii_s) \ast C_s \ast v_s, \quad \forall \ s \in S \]  
(23)

where \( C_s \) is the cost of species \( s \). This constraint takes the cost of the compounds participating in the reactions times the product minus the reactant times the cost of species \( s \) times the stoichiometric number for that species. The cost is then tabulated.

The above MILP may generate isomerization as the first viable stoichiometry looking at the values for the stoichiometry variables where negatives are reactants and positives are products. Upon implementation of integer cuts to the objective function, more reactions become available and are enumerated from the least number of reactants.
and products, 1 of each, to the most, up to a total of 3 reactants and 3 products per reaction.

The thermodynamics of the system are tabulated as the enthalpy of reaction and the Gibbs free energy of reaction. Unfeasible reactions are screened through an excel program that eliminates any reactions above a 10 kcal/mole Gibbs energy of reaction based on previous literature assertions for this as a viable industrial value (May and Rudd 1976; Clausen and Mattson 1978; Agnihotri and Motard 1980).

After the reactions have been enumerated and screened, the reactants, intermediates, and products are placed into a reaction tree and show all of the feasible reactions in a system as shown in Figure 4.4.

**Figure 4.4:** Superstructure of compounds with reactions. The reactants, intermediates, and products are now arranged into a chemical reaction tree after the reactions are generated and the reactions are screened based on the given criteria.
It is possible to form multiple sets of these trees depending on where the intermediates are placed based on the reactions forming. In this case, it is useful to use established chemical routes. There is also the possibility of not arranging the superstructure into a tree and instead letting the reactions formed over all feasible reactions become the tree instead. However, through heuristics and typical progressions in chemical reactions in industry, it is possible to get a close to optimal approximate arrangement of the tree to where logical chemical progressive steps are formed, and let other molecules take the place of current ones used in industry in a chemical pathway.

4.4 RNFA Optimization Model (Screen Reaction Paths)

The reaction network flux analysis model (RNFA) serves as a way to decide the best “flow” through a system (Beard, Liang and Qian 2002; Stelling, Klamt, Bettenbrock, Schuster and Gilles 2002; Lee, Yun, Park and Lee 2003; Sauer 2006). Originally developed for systems biology, the RNFA model can be adapted to any chemical system. Here, it is used to find the path of least resistance depending on certain constraints to flow. The constraints come in the form of thermodynamics, economics, safety factors, and heuristics to chemical reaction synthesis steps. The compounds are tabulated into a reaction network from the previous steps of compound identification and reaction generation and screening. The pathway is then comparable to a synthesis pathway that might be seen in systems biology such as glucose flow through a metabolic network. Once all of the compounds are arranged into a hierarchal pathway, the
comparable flows can be enumerated through a linear programming (LP) problem to
determine the best objective for flow, and for the second and later iterations, converted
into a mixed integer linear programming problem (MILP) in order to enumerate multiple
solutions with integer cuts (Lee, Phalakornkule, Domach and Grossmann 2000).

The first step is to determine the component mass balances for flow:

\[ S_s \ast v_j = 0 \quad \forall \ s \in S, \forall \ j \in J \tag{24} \]

where \( S_s \) is the stoichiometric matrix of reactions solved for in the reactions optimization
model. \( v_j \) is the flux vector of each reaction in the network. Constraint (24) is the
fundamental mass balance equation for flow through a network.

Limits can also be placed on the total flux for each flux vector:

\[ v_j^L \leq v_j \leq v_j^U , \quad \forall \ j \in J \tag{25} \]

The lower bound for the flux is normally set to a small number, such as 0.1, while the
upper flux will be limited to 100. These fluxes are a combination of theoretical and
heuristic knowledge as to the best target (please see the constraints 25a-25g for example
below). Additional constraints, based on thermodynamics, economics, and safety factors
are also applied.

As before, the enthalpy of formation of the pathway is tabulated:

\[ \Delta H_f^o \ (J/mol) \tag{25a} \]

Entropy of formation:

\[ \Delta S_f^o \ (J/mol*K) \tag{25b} \]
Gibbs free energy of formation for the pathway found explicitly through the DIPPR database, Gani method, or enumerated through the other parameters with the following equation and tabulated for all of the chemical reactions:

$$
\Delta G_f^o = \Delta H_f^o - T_{oper} \cdot \Delta S_f^o \text{ (J/mol)}
$$

(25c)

The operating temperature is again restricted:

$$
300 \leq T_{oper} \leq 1000, \quad \forall \ s \in S
$$

(25d)

Toxicity is also again included as the safety factor for each pathway:

$$
-\log(\text{LC50}_s), \quad \forall \ s \in S
$$

(25e)

These factors are again all tabulated and screened after the reactions paths are generated, the economics are also included in the following equation (where the binaries are taken from the previous example of reaction enumeration):

$$
\sum_s \ ii_s \cdot (i_s - ii_s) \cdot C_s \cdot v_s, \quad \forall \ s \in S
$$

(25f)

Flow constraints allow for the reaction to be set up in a superstructure that was only theoretically arranged in the reactions enumeration step. When the flows are balanced around a node, it shows a mass balance around that location and allows for the reaction pathway to be set up.

Flow Constraints:

$$
\text{Flow}_{j1} - \text{Flow}_{j2} = 0 \quad \forall \ j \in J
$$

(25g)

$$
\text{Flow}_{j4} - \text{Flow}_{j5} - \text{Flow}_{j6} = 0 \quad \forall \ j \in J
$$

Flux Balance Analysis 1st iteration LP problem:
\[
\max f(v) \tag{26}
\]
\[
f: \mathbb{R}^n \to \mathbb{R}
\]
\[
s.t. \ (24-25)
\]

where \( f \) is the objective function to maximize the flow of the product in the reaction network, subject to the conservation of mass and flux vector limits (24-25). It is worthwhile to note that the inclusion of the cost \((25f)\) would turn the LP into an MILP, so it may not be included.

Each acceptable reaction is subject to a thermodynamic constraint. The values for Gibbs free energy are tabulated for each reaction. A multi-step reaction may be thermodynamically viable while individual steps are not viable, for example a Solvay cluster.

An algorithm from Murabito will be used in order to enumerate the reactions, flagging each reaction as active or inactive and limiting backflow in a network (Murabito, Simeonidis, Smallbone and Swinton 2009).

Flagging the reaction as active or inactive:
\[
w_j^0 = 1 \leftrightarrow v_j = 0, \quad \forall j \in J \tag{27}
\]

Every flux variable is changed into three different variables:
\[
v_j = v_j^0 + v_j^- + v_j^+, \quad \forall j \in J \tag{28}
\]

For each flux, three binary variables are introduced to recast the LP into an MILP:
\[
w_j^0 + w_j^- + w_j^+ = 1, \quad \forall j \in J \tag{29}
\]

Coupling the \( w \) and \( v \) variables:
The new formulation the flux balance analysis for the 2nd and later iterations as an MILP is then written as follows:

\[
\begin{align*}
    v_j^L(1 - w_j^0) \leq v_j \leq v^U_j (1 - w_j^0), & \quad \forall j \in J \quad (30) \\
    -\varepsilon w_j^0 < v_j^0 < \varepsilon w_j^0, & \quad \forall j \in J \quad (31) \\
    v_j^L w_j^- \leq v_j^- \leq \varepsilon w_j^-, & \quad \forall j \in J \quad (32) \\
    \varepsilon w_j^+ \leq v_j^+ \leq w_j^+ v_j^U, & \quad \forall j \in J \quad (33) \\
    \varepsilon = 1 \times 10^{-6}.
\end{align*}
\]

The model is first solved as a LP with constraints (24-25), and then turned into a MILP with the addition of constraints (27-33) and run iteratively with integer cuts to determine the multiple alternate optima. The resulting reaction pathways are then put forward as new potential chemical synthesis routes based on the optimal economic, safety, and thermodynamic pathways. The optimal pathway is shown in red in Figure 4.5:
Figure 4.5: Superstructure of compounds with the best reaction. The best pathway is now shown in red after the chemical reaction pathways have been compared based on the criteria given.

The program is usually run with the setting for five or more integer cuts to generate multiple solutions with the same optimal value. As soon as the objective value for the flow through the system decreases, the program stops running.

4.5 Graphical Targeting Approach

In chemical systems, it is often necessary to develop a way to mark the changes in energy between different compounds in a reaction pathway. A proposed approach is to take the energy of each compound at a given point in a reaction pathway and plot it against the chemical potential of the compound reaction. The point in the chemical
pathway is based on the fraction of path length, allowing chemical pathways of different numbers of intermediates to be compared against each other. This method serves as a shortcut method to the more rigorous chemical stoichiometric enumeration and reaction network flux analysis. It allows for a map of chemical thermodynamic efficiency through a reaction progression.

The chemical potential is a part of the fundamental energy balance equation (Callen and Bridgman 1960) (Falk, Herrmann and Schmid 1983):

$$dE = T \, dS - p \, dV + \mu \, dn + \phi \, dQ + v \, dp + \psi \, dm + \cdots$$

(35)

The equation relates the energy of a system to the temperature (T), entropy (S), pressure (P), volume (V), chemical potential (µ), charge (Q), velocity (v), electric potential (ψ), and other fundamental quantities.

The chemical potential is also the derivative of the energy with respect to the number of particles:

$$\left(\frac{\partial E}{\partial n}\right)_{S,V,...} = \mu_i$$

(36)

An overall definition of the chemical potential can then be given as a potential of the chemical system to undergo a change in its chemical composition, or move from a gradient of higher chemical potential to one of lower chemical potential. The values for the chemical potential are tabulated in from a number of references (Herrmann and Job 1996; Job and Herrmann 2006; Rüffler and Job 2009).

For a chemical reaction pathway, the first way of looking at the chemical potential is to use the simplest reaction, an isomerization:
Then show how chemical potential changes can be used to predict the direction of the transformation or chemical equilibrium.

When: \( \mu_A > \mu_B \) (38)
compound A will turn into compound B, or a compound will flow from A to B.

When: \( \mu_A = \mu_B \) (39)
compound A is in equilibrium with compound B, or there will be equilibrium between location A and B.

When: \( \mu_A < \mu_B \) (40)
compound B will turn into compound A, or a compound will flow from B to A.

Here the Gibbs energy is measured against the chemical potential to show which chemical pathway is the best based on a graphical targeting approach. The Gibbs energy is a function of volume, pressure, entropy, temperature, chemical potential, and the amount of the species:

\[
dG = Vdp - SdT + \sum \mu_i dN_i - \ldots
\] (41)

where \( \mu_i \) is in Joules/mol. The compounds are tabulated for standard temperatures and pressures. The chemical potential can also be related to the actual temperature and pressure of the industrial reactions by the following formula:

\[
\mu = \mu^o + RT \ln \left( \frac{p}{p_o} \right)
\] (42)

where \( \mu^o \) is the standard state 298K, 1 atm, R is the universal gas constant, T is the actual temperature, p is the actual pressure and \( p_o \) is the standard state pressure.
The first step to develop a targeting approach is to find the best overall reaction and then compare other reaction pathways based on the belief that the lowest relative changes in chemical potential and by extension the Gibbs free energy will lead to the most favorably efficient reaction. The chemical potential will be plotted against the reaction steps in order to show a progression over the reaction pathway from the reactants to the products and the relative chemical changes that lead to an increase in the chemical potential in this biorefinery system (Hatzimanikatis, Li, Ionita and Broadbelt 2004; Hatzimanikatis, Li, Ionita, Henry, Jankowski et al. 2005).

The steps to developing the graphical methods are as follows:

1) Find the best line for a reaction pathway between a given reactant and a given product for targeting purposes. This may or may not involve adding molecular hydrogen to the reactant and removing molecular oxygen from the product for biorefineries. The theoretical targeting line corresponds to the, in a biorefinery, addition of molecular hydrogen and the removal of molecular oxygen. For example in the conversion of glucose to octane the equation might be as follows:

\[ 8 \text{C}_6\text{H}_{12}\text{O}_6 + 6 \text{H}_2 \rightarrow 6 \text{C}_8\text{H}_{18} + 24 \text{O}_2 \]  

(43)

2) Determine the reaction pathways and products of the first reaction, and the subsequent reactants and products to following the approach given in the reaction enumeration and reaction network synthesis parts.

3) Fill in the thermodynamic chemical potentials over the course of the reaction pathway on the same graph.
When the graph is put together, they will show a line for each of the reaction pathways enumerated, with chemical potential on the y-axis and fraction of the path length on the x-axis (Finley, Broadbelt and Hatzimanikatis 2009) as shown in Figure 4.6.

**Figure 4.6**: Example pathway targeting and reactions. The targeting pathway is in blue, while the three enumerated pathways based on chemical potential are show in red, green, and purple respectively. Each node corresponds to a compound intermediate.
When the chemical potentials are arranged in such a way, they show the thermodynamic feasibility of the reaction pathways, and the efficiency of each chemical reaction. It should be noted, however, that there is no particular way to assign a certain pathway as the best case based solely on the graphical method. A number of heuristic factors have to be taken into account in order to classify a pathway as the best case. There should also be an agreement between the graphical methods and the more rigorous optimization formulations presented earlier.

Large changes in chemical potentials should be avoided. There is an efficiency that can be compared between pathways based on relative changes in the thermodynamic chemical potentials. This efficiency has to do with the large changes in chemical potentials.
5. CASE STUDY

5.1 Building a Starting Point

Modeling feasible chemical pathways from biomass to liquid fuels while exploring alternate pathways that may potentially prove useful in the future requires many different factors to be taken into account. Here, the chemical reaction network is modeled from the work of Huber and relies on finding feasible chemical steps from biomass to liquid fuels, and is then extended to novel reactions beyond what is found in the literature (Huber, Iborra and Corma 2006).

When the compound identification step was performed, a total of 85 compounds were found by using a combination of chemical knowledge, heuristics, safety, economics, and thermodynamic factors. Please see APPENDIX B for the full compound list and APPENDIX C for the thermodynamic properties of the compounds. Reactants were classified including woody biomass, cellulose, hemi-cellulose, and lignin and other starting reactant materials such as glucose, fructose, syngas (CO and H$_2$), bio-oil, the products of gasification, pyrolysis or hydrolysis of biomass and its component molecules. Intermediates and co-materials such as C$_2$H$_6$, CH$_4$, CH$_2$O, H$_2$O, H$_2$, CO, CO$_2$ were included as reacting intermediates based on current chemical conversion methods. Liquid fuel products such as 2-methyltetrahydrofuran (MTHF), alkanes, and ethanol were selected as the products. Once all of the molecules have been added to the superstructure, the total reacting system is complete.
The 85 compounds were then grouped into 15 feasible reaction clusters so that they could be enumerated in the reaction stoichiometric enumeration step, based on known and unknown chemical reactions. Please see APPENDIX A for the LINGO formulation.

Catalysts, kinetics, and enzymes were not taken into account. The reactions were first enumerated for each reaction step, for example: syngas (composed of CO and H₂) to methanol synthesis, and later combined into a reaction network. The difficulty of chemically modeling bio-oil meant that the reactions including bio-oil were not enumerated, but rather taken as a one reaction stoichiometry and combined into the reaction model.

Once a chemical reaction was enumerated, it was looked up in the literature to see if there was an analogous reaction mechanism proposed and a way to carry out that reaction. So, the reactions were first enumerated, and then potentially feasible pathways were compared to literature studies as to the applicability of the chemical models, please see APPENDIX D for an example of the reaction enumeration data in Excel.

When the pathways are put into a synthesis tree, the following detailed synthesis map is shown in Figure 5.1.
Figure 5.1: Reaction Synthesis Tree, based on Huber. The chemical reaction pathways have been shown to follow this basic shape. Additional reactions are created through stoichiometric enumeration. Not all chemical intermediates are shown.
When the reactions were enumerated, 280 reactions were found to be feasible and combined into Figure 5.1. The starting point for reaction synthesis occurs at the C6 sugars, C5 sugars, lignin, bio-oils, and syngas level. The thermodynamic and economic factors were then tabulated for each reaction. The network was then built using the 280 feasible reactions and solved through RNFA and optimized based on the model developed for the problem.

Stoichiometrically enumerating all of the potential feasible reactions in the 85 compound system between glucose and MTHF leads to a reaction structure that includes several different intermediates between glucose and MTHF. The reaction formulation is set to loop over all potential target molecules, with each reactant then being able to become a product. Once these are all enumerated, the pathway is setup based on heuristics and the RNFA step can be run. The biomass is broken down into two parts of xylose and glucose, and the following reaction tree was enumerated, the red arrows are the optimized pathway based on RNFA, while the blue arrows are part of the overall stoichiometric reactions:

It is necessary to create all of the iterations between the different compounds of reactants to intermediates and intermediates to products. The first step is to identify the pretreatment steps as converting biomass into reactants. The actual case study starts when the chemical compounds of high molecular weight biomass are broken down into their simpler chemical forms such as glucose, fructose, coumaryl alcohols and related forms.
Given biomass, these are the most favorable chemical pathways based on the included constraint factors. If a different starting material was used, most likely a different set of reactions would be enumerated, leading to a different RNFA model and a different optimal solution as to the best fuel products.

So the entire chemical reaction process was mapped out by solving the feasible chemical reaction pathways from glucose to liquid fuels. The following example shows a part of the synthesis tree and how it can be applied in taking glucose and xylose to a product.

5.2 More Detailed Reactant to Product: Biomass to MTHF

The example given here is the conversion of glucose (C$_6$H$_{12}$O$_6$) or xylose (C$_5$H$_{10}$O$_5$) to the alternative fuel MTHF, also known as 2-Methyltetrahydrofuran (C$_5$H$_{10}$O). The first step is to look at the pathway from biomass to the starting reactants, such as glucose and xylose. Since this initial biomass to glucose pathway is not explicitly modeled, the yields from a particular biomass could be used. The starting materials of glucose and xylose are then used and can be enumerated to their product. The reaction pathway is shown in Figure 5.2.
Figure 5.2: The enumerated pathways from glucose and xylose to MTHF.
The results of Figure 5.2 are a result of the compound identification, reaction enumeration, placing the compounds into a chemical pathway and running the RNFA model. When the reactions are placed into a flux network, intermediates such as 1,2-dihydroxypentanal versus angelica lactone are compared as to which intermediate pathway is better. The mathematical LINGO® steps compare the thermodynamics, economics, safety, and heuristics then arrive at a best pathway.

When the reactions are placed into a graphical form, the first step is to draw the direct line for the reduction of glucose to produce octane. This corresponds to the following unbalanced reaction of glucose with elemental hydrogen with elemental oxygen as the product:

\[
C_6H_{12}O_6 + H_2 \rightarrow C_5H_{10}O + O_2
\]  

Balancing the equation gives:

\[
5C_6H_{12}O_6 \rightarrow 6C_5H_{10}O + 12O_2
\]  

And the following unbalanced reaction of xylose with elemental hydrogen with elemental oxygen as the product:

\[
C_5H_{10}O_5 + H_2 \rightarrow C_5H_{10}O + O_2
\]  

Balancing the equation gives:

\[
C_5H_{10}O_5 \rightarrow C_5H_{10}O + 2O_2
\]  

These reactions serve as the best targeting for the direct conversion of reactants into products. These reactions may be potentially possible in the future through mechanisms such as reversing the photosynthesis process or with advanced catalysts or enzymes.
This best case occurs because of the chemical potentials of the elements in their elemental form is zero.

The best we may be able to do realistically is not on this line, but for targeting purposes, it serves as the best theoretical target to a pathway. Once the entire problem is enumerated, all solutions will start at the reactant point and end at the product point based on a reaction step graph.

Now, take the reactions generated in the first two steps of the reactions optimization model, and the reaction network flux analysis, and plot them on the graph. The chemical potentials are plotted on the Y-axis, and the reaction steps are plotted on the X-axis. The number of reaction steps will be determined from the reaction pathway that takes the most number of steps.

While there may be multiple different numbers of reaction steps between the different reaction paths, there is a quick and convenient way to represent them, at standard temperatures and pressures is in Figure 5.3.
Figure 5.3: The targeting of optimal pathways. This illustrates the differences between different intermediates between glucose and MTHF. The relative changes are differences in chemical potential.

Two alternative chemical pathways are also presented in Figure 5.4, through different intermediates of levulinic esters and 1,2-dihydroxypentanal, all at standard temperatures and pressures.
Figure 5.4: The targeting of two alternate pathways. This illustrates the differences between different intermediates between glucose and MTHF. The relative changes are differences in chemical potential.

The graph highlights the large chemical potential changes through the use of methanol as an intermediate in the production of alkanes. Large differences in chemical potentials are sought to be avoided. There is a theoretical industrial limit of approximately +10 kCal/mol for a reaction step (41.8 kJ/mol) which has been suggested by various sources (May and Rudd 1976; Clausen and Mattson 1978; Agnihotri and Motard 1980). This means that endothermic reactions that are above this limit are not practically feasible in an industrial setting without enzyme or catalyst development.
Forming this product from biomass shows that there are thermodynamic benefits and drawbacks to going through the levulinic ester intermediate as opposed to 1,2-dihydroxypentanal intermediate.

The previous example has shown the difference between chemical pathways to achieve a product from a glucose reactant to a MTHF fuel product. The reaction enumeration and reaction network flux analysis and corresponding graphical representation show feasible comparable pathways.
6. RESULTS AND DISCUSSIONS

The compound identification first allows the setting of the superstructure for the problem. The compounds are found through a search of different methods, from the DIPPR database to literature searches. In the case study, the 85 compounds were found to be the most applicable to the current biorefinery biomass to fuel concept and placed into the superstructure. These compounds were then run through the reactions optimization model and the best reaction pathway was selected through the RNFA optimization problem. In the case study, the pathway selected was shown in Figure 5.2. Once the case study pathways were enumerated, they were then shown graphically on the chemical potential graph of Figure 5.3 and the alternate pathways on Figure 5.4. The graph shows that the pathway between glucose and MTHF can go through 3 different intermediates.

From the graphical method it would appear that the intermediate 1,2-dihydroxypentanal deserves to be looked at as opposed to using the angelica lactone intermediate in glucose to MTHF synthesis on a thermodynamic chemical potential basis.

As a shortcut method, the graphical approach seeks to unify the three other methods of compound identification, reaction enumeration, and flux analysis. Quickly being able to screen the best pathway without having to do the detailed analysis shows how a common thread runs through the other three methods. The graphical method could have also been used independently of the other two methods.
7. CONCLUSIONS

In this thesis, a method was developed to enumerate all of the feasible reaction pathways from reactants to products and the optimal solution picked based on thermodynamic, economic, safety, and heuristic factors. A hypothetical case study was presented at the end of the work that involved the use of wood biomass and its conversion to liquid fuels. Based on the work of Huber, the optimal chemical pathways have been found and extended with additional compounds. In the case study, a chemical superstructure of 85 compounds was first made from common biomass to liquid fuel routes and chemical heuristics were also used to add new molecules. Every feasible reaction was enumerated through a stoichiometric mass balance algorithm, and the thermodynamic and economic properties tabulated. RNFA with integer cuts was then used in order to find the optimal yield and economic reaction pathways from starting reactants to products for various fuels.

The graphical method shows the use of chemical potential versus the reaction step in order to show how the changes in chemical potential over the course of the reaction can be used as a rough guide to the more rigorous reaction stoichiometry and RNFA enumeration models. When the graphs are used, they allow for a quick comparison between different chemical intermediates.
REFERENCES


Fornari T, Stephanopoulos G (1994b) Synthesis of chemical reaction paths: economic and specification constraints. Chem Eng Comm 129: 159-182


APPENDIX A

A.1 Stoichiometric code formulation in LINGO

!Model by E.Pennaz 2010-2011;

!This model is a feasible reaction enumeration and chemistry constraint model;

!It solves reactions in "reaction clusters" following simple iteration rules and using
integer cuts, then outputs the reaction data to excel;

!The reaction system used is based on "Biomass to Liquid Fuels"

MODEL:

!For each time you run the program for an iteration or reacting system, change:

1) Iteration # for which reacting system is calculated (in K-Best solutions sub model)

2) Send solutions to excel output data 'Stoichiometry1' for reacting system 1 etc, 2 for
2, etc.

3) Number of K Best Solutions wanted

4) Number Q = ? For each species target

5) Number Z = ? for the number of different product species targets;

DATA:

!Number of K-Best Solutions Wanted;

\[ K = 25; \]

ENDDATA
SETS:

SPECIES / A1..A85 / : Carbon, Hydrogen, Oxygen ;

Coefficients / v1..v85 / : Stoic;

Product_Flag / iA1 .. iA85 / : Product;

Reactant_Flag / iiA1 .. iiA85 / : Reactant;

Participation_Flag / iiiA1 .. iiiA85 / : Participation;

Dummy_VariableX / X1..X85 / : DummyX;

Dummy_VariableY / Y1..Y85 / : DummyY;

KSOLUTIONS /1..K/: OBJ, RHS, Molecule;

KXI( Coefficients, KSOLUTIONS): CUT, INCLUDE2;

ActiveMolecules /AC1..AC85/ !Stoic Species which have a target; : Points;

Final( ActiveMolecules, Coefficients, KSOLUTIONS ): INCLUDE3;

Total /T1..T14 / : Target;

ENDSETS

DATA:

!set members = 85 molecules

A1 = Carbon_Monoxide,  A2 = Hydrogen,  A3 = Formic_Acid, A4 = Acetic_Acid, A5 = Propanoic_Acid, A6 = Acetone, A7 = Formaldehyde,  A8 = Acetaldehyde,  A9 = Oxaldehyde,  A10 = Glycolaldehyde, A11 = Acetal,  A12 = Cellulose_Acetate,  A13 = Celllobiose, A14 = Fructose,  A15 = Glucose, A16 = DME(dimethyl ether),A17 = HMF,
|---------|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
Carbon =  1  0  1  2  3  3  1  2  2  6  76  12  6  6  2  6
  5  6  10  10  8  5  5  5  5  5  5  5  5  5  5  1  1
  2  3  4  8  2  3  4  4  6  7  8  8  8  7  7  7  9  10  9
 10  11  12  13  5  2  5  6  1  0  3  0  1  1  1  1  1  6  5
 10  1  9  10  11  11  6  3  4  6  6  6  20  5 ; 

Hydrogen =  0  2  2  4  6  6  2  4  2  4  14  114 22  12 12  6  3
  2  1  12  12  2  10  4  6  6  10  8  6  8  12  12  12  4  4
  6  8  10  18  4  6  8  8  6  8  10  10  10  8  8  8  10  12
 20  22  24  26  28  10  6  12  14  0  2  8  0  1.6  2.59 0.72 0.54 10
  8  12  1.6 10  12  14  20  12  8 10  12  10  42 12 ;

Oxygen =  1  0  2  2  2  1  1  1  2  2  2  49 11  6  6  1  3
  2  1  2  2  2  5  2  2  1  2  3  2  2  2  1  1  1  0
  0  0  0  0  0  0  0  0  0  0  0  0  1  1  1  0  0  0
  0  0  0  0  1  1  0  0  2  1  3  2 0.68 1.05 0.22 1.3  5  4
  3 0.68 2  3  4  3  3  1  1  0  0  0  1 ;

ENDDATA

SUBMODEL Base:

!Objective Function to minimize the number of reactants, products, and coproducts in
a given system ;
[R_OBJ] \textbf{MIN} = @SUM (\text{Product\_Flag (I)} : \text{Product (I)}) + @SUM(\text{Reactant\_Flag (I)} : \\
\text{Reactant (I)}) ;

!atom mass balance;

[Carbon\_Balance] @SUM (\text{SPECIES (I)} : \text{Carbon (I)}* \text{Stoic (I)}) = 0;

[Hydrogen\_Balance] @SUM (\text{SPECIES (I)} : \text{Hydrogen (I)}* \text{Stoic (I)}) = 0;

[Oxygen\_Balance] @SUM (\text{SPECIES (I)} : \text{Oxygen (I)}* \text{Stoic (I)}) = 0;

!The following lines takes care of for stoeiometric constraints;

@FOR (\text{Dummy\_VariableX (I)} : @GIN (\text{DummyX (I)}));

@FOR( \text{Dummy\_VariableX (I)} : @BND( 0, \text{DummyX(I)}, 1000));

!Integer Values of Product (21-24);

@FOR (\text{Coefficients (I)} : \text{Stoic (I)} = \text{DummyX (I)} - 2 * \text{DummyY (I)});

@FOR (\text{Dummy\_VariableY (I)} : \text{DummyY (I)} - \text{Vmax} * \text{Reactant (I)} <= 0);

@FOR (\text{Dummy\_VariableX (I)} : \text{DummyX (I)} + \text{Vmax} * (\text{Reactant (I)} - 1) - \text{DummyY (I)} <= 0);

@FOR (\text{Dummy\_VariableY (I)} : \text{DummyY (I)} - \text{DummyX (I)} <= 0);

!Bounded stoeiometric coefficient;

\text{Vmax} = 50;

!Additional constraint for non zeros when only when species s is a reactant ;

@FOR (\text{Dummy\_VariableY (I)} : \text{DummyY (I)} >= \text{Reactant (I)});

!Role Specification Constraints, Raw Material and Product identifiers ;
@FOR (Dummy_VariableX (I) : DummyX(I) - DummyY(I) - Vmax * Product(I) <= 0);

@FOR (Dummy_VariableX (I) : DummyX(I) - DummyY(I) - (Vmax + 1) * Product(I) + Vmax >= 0);

!Role Specification Constraints, i = product flag, ii = reactant flag, iii = participation flag

@FOR (Product_Flag (I) : Product(I) + Reactant(I) + Participation(I) = 1);

!Limit on the number of reacting species and product species;

@SUM (Reactant_Flag (I) : Reactant(I) ) <= 3;

@SUM (Product_Flag (I) : Product(I) ) <= 3;

!Limit on the stoichiometric maximum of all species;

@FOR (Coefficients (I) : Stoic(I) <= Vmax );

@FOR (Coefficients (I) : @FREE (Stoic(I)));

@FOR (Coefficients: @BND(-50, Stoic, 50));

!Binary values of the variables is, iis, iiis;

@FOR( Product_Flag: @BIN(Product));

@FOR( Reactant_Flag: @BIN(Reactant));

@FOR( Participation_Flag: @BIN(Participation));

ENDSUBMODEL
SUBMODEL Iteration4: !C6Sugars to produce ethanol, C1-C6 alkanes, aromatics, alkanes, hydrogen; !K = 25;

!Chemistry Constraints, only certain species may react with each other;

!Only one alkane product produced per run;

Product(37) + Product(38) + Product (39) + Product (40) + Product (41) + Product (42) + Product (43) <= 1;

!Cannot produce n-octane directly from glucose;

Reactant(15) + Product(39) <= 1;

!Reactant, Product, or Not in Reaction, Flag: Role Specification Constraints, and active reactions;

Stoic(1) = 0; Stoic(2) >= 0; Stoic(3) = 0; Stoic(4) = 0; Stoic(5) = 0; Stoic(6) = 0; Stoic(7) = 0; Stoic(8) = 0; Stoic(9) = 0; Stoic(10) = 0;

Stoic(11) = 0; Stoic(12) = 0; Stoic(13) = 0; Stoic(14) = 0; Stoic(15) <= -1; Stoic(16) = 0; Stoic(17) = 0; Stoic(18) = 0; Stoic(19) = 0; Stoic(20) = 0;

Stoic(21) = 0; Stoic(22) = 0; Stoic(23) = 0; Stoic(24) = 0; Stoic(25) = 0; Stoic(26) = 0; Stoic(27) = 0; Stoic(28) = 0; Stoic(29) = 0; Stoic(30) = 0;

Stoic(31) = 0; Stoic(32) = 0; Stoic(33) = 0; Stoic(34) = 0; Stoic(35) >= 0; Stoic(36) >= 0; Stoic(37) >= 0; Stoic(38) >= 0; Stoic(39) >= 0; Stoic(40) >= 0; Stoic(41) >= 0; Stoic(42) >= 0; Stoic(43) >= 0; Stoic(44) >= 0; Stoic(45) >= 0; Stoic(46) >= 0; Stoic(47) >= 0; Stoic(48) >= 0; Stoic(49) >= 0; Stoic(50) >= 0;
Stoic(51) >= 0; Stoic(52) >= 0; Stoic(53) = 0; Stoic(54) = 0; Stoic(55) = 0; Stoic(56) = 0; Stoic(57) = 0; Stoic(58) = 0; Stoic(59) = 0; Stoic(60) >= 0;

Stoic(61) = 0; Stoic(62) = 0; Stoic(63) >= 0; Stoic(64) >= 0; Stoic(65) = 0; Stoic(66) = 0; Stoic(67) = 0; Stoic(68) = 0; Stoic(69) = 0; Stoic(70) = 0;

Stoic(71) = 0; Stoic(72) = 0; Stoic(73) = 0; Stoic(74) = 0; Stoic(75) = 0; Stoic(76) = 0; Stoic(77) = 0; Stoic(78) = 0; Stoic(79) = 0; Stoic(80) = 0;

Stoic(81) = 0; Stoic(82) = 0; Stoic(83) = 0; Stoic(84) = 0; Stoic(85) = 0;

ENDSUBMODEL

SUBMODEL Iteration5: !C5Sugars to produce Ethanol, C1-C6 alkanes, aromatics, hydrogen; !K = 25;

!Chemistry Constraints, only certain species may react with each other;

!Only one alkane product produced per run;

Product(37) + Product(38) + Product (39) + Product (40) + Product (41) +

Product (42) + Product (43) <= 1;

!Cannot produce n-octane directly from xylose;

Reactant(23) + Product(39) <= 1;

!Reactant, Product, or Not in Reaction, Flag: Role Specification Constraints, and active reactions;

Stoic(1) = 0; Stoic(2) >= 0; Stoic(3) = 0; Stoic(4) = 0; Stoic(5) = 0; Stoic(6) = 0;

Stoic(7) = 0; Stoic(8) = 0; Stoic(9) = 0; Stoic(10) = 0;
Stoic(11) = 0; Stoic(12) = 0; Stoic(13) = 0; Stoic(14) = 0; Stoic(15) = 0; Stoic(16) = 0; Stoic(17) = 0; Stoic(18) = 0; Stoic(19) = 0; Stoic(20) = 0; Stoic(21) = 0; Stoic(22) = 0; Stoic(23) <= -1; Stoic(24) = 0; Stoic(25) = 0; Stoic(26) = 0; Stoic(27) = 0; Stoic(28) = 0; Stoic(29) = 0; Stoic(30) = 0; Stoic(31) = 0; Stoic(32) = 0; Stoic(33) = 0; Stoic(34) = 0; Stoic(35) >= 0; Stoic(36) >= 0; Stoic(37) >= 0; Stoic(38) >= 0; Stoic(39) >= 0; Stoic(40) >= 0; Stoic(41) >= 0; Stoic(42) >= 0; Stoic(43) >= 0; Stoic(44) >= 0; Stoic(45) >= 0; Stoic(46) >= 0; Stoic(47) >= 0; Stoic(48) >= 0; Stoic(49) >= 0; Stoic(50) >= 0; Stoic(51) >= 0; Stoic(52) >= 0; Stoic(53) = 0; Stoic(54) = 0; Stoic(55) = 0; Stoic(56) = 0; Stoic(57) = 0; Stoic(58) = 0; Stoic(59) = 0; Stoic(60) >= 0; Stoic(61) = 0; Stoic(62) = 0; Stoic(63) >= 0; Stoic(64) >= 0; Stoic(65) = 0; Stoic(66) = 0; Stoic(67) = 0; Stoic(68) = 0; Stoic(69) = 0; Stoic(70) = 0; Stoic(71) = 0; Stoic(72) = 0; Stoic(73) = 0; Stoic(74) = 0; Stoic(75) = 0; Stoic(76) = 0; Stoic(77) = 0; Stoic(78) = 0; Stoic(79) = 0; Stoic(80) = 0; Stoic(81) = 0; Stoic(82) = 0; Stoic(83) = 0; Stoic(84) = 0; Stoic(85) = 0; END SUBMODEL

SUBMODEL Iteration6: !C6 Sugars to via 5-HMF to produce Levulinic Acid, H2 use is ACID(H+ and H+); !K = 25;

!Chemistry Constraints, only certain species may react with each other;

!Only use glucose or fructose, not both;
Reactant (14) + Reactant(15) = 1;

!Reactant, Product, or Not in Reaction, Flag: Role Specification Constraints, and active reactions;

Stoic(1) = 0; !Stoic(2) = 0; Stoic(3) >= 0; Stoic(4) = 0; Stoic(5) = 0; Stoic(6) = 0; Stoic(7) = 0; Stoic(8) = 0; Stoic(9) = 0; Stoic(10) = 0; Stoic(11) = 0; Stoic(12) = 0; Stoic(13) = 0; Stoic(14) <= 0; Stoic(15) <= 0; Stoic(16) = 0; Stoic(17) = 0; Stoic(18) = 0; Stoic(19) = 0; Stoic(20) = 0; Stoic(21) = 0; Stoic(22) = 0; Stoic(23) = 0; Stoic(24) = 0; Stoic(25) = 0; Stoic(26) = 0; Stoic(27) = 0; Stoic(28) >= 0; Stoic(29) = 0; Stoic(30) = 0; Stoic(31) = 0; Stoic(32) = 0; Stoic(33) = 0; Stoic(34) = 0; Stoic(35) = 0; Stoic(36) = 0; Stoic(37) = 0; Stoic(38) = 0; Stoic(39) = 0; Stoic(40) = 0; Stoic(41) = 0; Stoic(42) = 0; Stoic(43) = 0; Stoic(44) = 0; Stoic(45) = 0; Stoic(46) = 0; Stoic(47) = 0; Stoic(48) = 0; Stoic(49) = 0; Stoic(50) = 0; Stoic(51) = 0; Stoic(52) = 0; Stoic(53) = 0; Stoic(54) = 0; Stoic(55) = 0; Stoic(56) = 0; Stoic(57) = 0; Stoic(58) = 0; Stoic(59) = 0; Stoic(60) = 0; Stoic(61) = 0; Stoic(62) = 0; Stoic(63) = 0; !Stoic(64) = 0; Stoic(65) = 0; Stoic(66) = 0; Stoic(67) = 0; Stoic(68) = 0; Stoic(69) = 0; Stoic(70) = 0; Stoic(71) = 0; Stoic(72) = 0; Stoic(73) = 0; Stoic(74) = 0; Stoic(75) = 0; Stoic(76) = 0; Stoic(77) = 0; Stoic(78) = 0; Stoic(79) = 0; Stoic(80) = 0; Stoic(81) = 0; Stoic(82) = 0; Stoic(83) = 0; Stoic(84) = 0; Stoic(85) = 0;
SUBMODEL Iteration 7: !Levulinic Acid to produce Levulinic Esters, MTHF; !K = 25;

!Chemistry Constraints, only certain species may react with each other;

!Reactant, Product, or Not in Reaction, Flag: Role Specification Constraints, and active reactions;

Stoic(1) = 0; Stoic(2) <= 0; Stoic(3) = 0; Stoic(4) = 0; Stoic(5) = 0; Stoic(6) = 0;
Stoic(7) = 0; Stoic(8) = 0; Stoic(9) = 0; Stoic(10) = 0;
Stoic(11) = 0; Stoic(12) = 0; Stoic(13) = 0; Stoic(14) = 0; Stoic(15) = 0; Stoic(16) = 0;
Stoic(17) = 0; Stoic(18) = 0; Stoic(19) = 0; Stoic(20) = 0;
Stoic(21) = 0; Stoic(22) = 0; Stoic(23) = 0; Stoic(24) = 0; Stoic(25) = 0; Stoic(26) = 0;
Stoic(27) = 0; Stoic(28) <= 0; Stoic(29) = 0; Stoic(30) = 0;
Stoic(31) = 0; Stoic(32) = 0; Stoic(33) = 0; Stoic(34) = 0; Stoic(35) = 0; Stoic(36) = 0;
Stoic(37) = 0; Stoic(38) = 0; Stoic(39) = 0; Stoic(40) = 0;
Stoic(41) = 0; Stoic(42) = 0; Stoic(43) = 0; Stoic(44) = 0; Stoic(45) = 0; Stoic(46) = 0;
Stoic(47) = 0; Stoic(48) = 0; Stoic(49) = 0; Stoic(50) = 0;
Stoic(51) = 0; Stoic(52) = 0; Stoic(53) = 0; Stoic(54) = 0; Stoic(55) = 0; Stoic(56) = 0;
Stoic(57) = 0; Stoic(58) = 0; Stoic(59) >= 0; Stoic(60) = 0;
Stoic(61) = 0; Stoic(62) = 0; Stoic(63) = 0; Stoic(64) >= 0; Stoic(65) = 0; Stoic(66) = 0;
Stoic(67) = 0; Stoic(68) = 0; Stoic(69) = 0; Stoic(70) = 0;
Stoic(71) = 0; Stoic(72) = 0; Stoic(73) = 0; Stoic(74) = 0; Stoic(75) = 0; Stoic(76) = 0;
Stoic(77) = 0; Stoic(78) >= 0; Stoic(79) = 0; Stoic(80) = 0;
Stoic(81) = 0; Stoic(82) <= 0; Stoic(83) = 0; Stoic(84) = 0; Stoic(85) = 0;
!Stoic(Q) >= 1;
ENDSUBMODEL

SUBMODEL Iteration8: !C5 Sugars to produce Furfural; !K= 25;

!Chemistry Constraints, only certain species may react with each other;

!Reactant, Product, or Not in Reaction, Flag: Role Specification Constraints, and active reactions;

Stoic(1) = 0; Stoic(2) = 0; Stoic(3) = 0; Stoic(4) = 0; Stoic(5) = 0; Stoic(6) = 0;
Stoic(7) = 0; Stoic(8) = 0; Stoic(9) = 0; Stoic(10) = 0;
Stoic(11) = 0; Stoic(12) = 0; Stoic(13) = 0; Stoic(14) = 0; Stoic(15) = 0; Stoic(16) = 0;
Stoic(17) = 0; Stoic(18) = 0; Stoic(19) = 0; Stoic(20) = 0;
Stoic(21) = 0; Stoic(22) = 0; Stoic(23) <= -1; Stoic(24) >= 1; Stoic(25) = 0; Stoic(26) = 0; Stoic(27) = 0; Stoic(28) = 0; Stoic(29) = 0; Stoic(30) = 0;
Stoic(31) = 0; Stoic(32) = 0; Stoic(33) = 0; Stoic(34) = 0; Stoic(35) = 0; Stoic(36) = 0;
Stoic(37) = 0; Stoic(38) = 0; Stoic(39) = 0; Stoic(40) = 0;
Stoic(41) = 0; Stoic(42) = 0; Stoic(43) = 0; Stoic(44) = 0; Stoic(45) = 0; Stoic(46) = 0;
Stoic(47) = 0; Stoic(48) = 0; Stoic(49) = 0; Stoic(50) = 0;
Stoic(51) = 0; Stoic(52) = 0; Stoic(53) = 0; Stoic(54) = 0; Stoic(55) = 0; Stoic(56) = 0;
Stoic(57) = 0; Stoic(58) = 0; Stoic(59) = 0; Stoic(60) = 0;
Stoic(61) = 0; Stoic(62) = 0; Stoic(63) = 0; Stoic(64) >= 1; Stoic(65) = 0; Stoic(66) = 0;
Stoic(67) = 0; Stoic(68) = 0; Stoic(69) = 0; Stoic(70) = 0;
Stoic(71) = 0; Stoic(72) = 0; Stoic(73) = 0; Stoic(74) = 0; Stoic(75) = 0; Stoic(76) = 0;
Stoic(77) = 0; Stoic(78) = 0; Stoic(79) = 0; Stoic(80) = 0;
Stoic(81) = 0; Stoic(82) = 0; Stoic(83) = 0; Stoic(84) = 0; Stoic(85) = 0;
!Stoic(Q) >= 1;

ENDSUBMODEL

SUBMODEL Iteration9: !Furfural to produce MTHF, C8-C13 alkanes, alcohols; !K= 25;

!Chemistry Constraints, only certain species may react with each other;

!Reactant, Product, or Not in Reaction, Flag: Role Specification Constraints, and active reactions;

Stoic(1) = 0; Stoic(2) <= -1; Stoic(3) = 0; Stoic(4) = 0; Stoic(5) = 0; Stoic(6) = 0;
Stoic(7) = 0; Stoic(8) = 0; Stoic(9) = 0; Stoic(10) = 0;
Stoic(11) = 0; Stoic(12) = 0; Stoic(13) = 0; Stoic(14) = 0; Stoic(15) = 0; Stoic(16) = 0;
Stoic(17) = 0; Stoic(18) = 0; Stoic(19) = 0; Stoic(20) = 0;
Stoic(21) = 0; Stoic(22) = 0; Stoic(23) = 0; Stoic(24) <= -1; Stoic(25) = 0; Stoic(26) >= 0; Stoic(27) = 0; Stoic(28) = 0; Stoic(29) = 0; Stoic(30) = 0;
Stoic(31) = 0; Stoic(32) = 0; Stoic(33) = 0; Stoic(34) = 0; Stoic(35) = 0; Stoic(36) = 0;
Stoic(37) = 0; Stoic(38) = 0; Stoic(39) >= 0; Stoic(40) = 0;
Stoic(41) = 0; Stoic(42) = 0; Stoic(43) = 0; Stoic(44) = 0; Stoic(45) = 0; Stoic(46) = 0;
Stoic(47) = 0; Stoic(48) = 0; Stoic(49) = 0; Stoic(50) = 0;
Stoic(51) = 0; Stoic(52) = 0; Stoic(53) = 0; Stoic(54) >= 0; Stoic(55) >= 0; Stoic(56) >=
0; Stoic(57) = 0; Stoic(58) >= 0; Stoic(59) = 0; Stoic(60) = 0;
Stoic(61) = 0; Stoic(62) = 0; Stoic(63) = 0; Stoic(64) >= 1; Stoic(65) = 0; Stoic(66) = 0;
Stoic(67) = 0; Stoic(68) = 0; Stoic(69) = 0; Stoic(70) = 0;
Stoic(71) = 0; Stoic(72) = 0; Stoic(73) = 0; Stoic(74) = 0; Stoic(75) = 0; Stoic(76) = 0;
Stoic(77) = 0; Stoic(78) = 0; Stoic(79) = 0; Stoic(80) = 0;
Stoic(81) = 0; Stoic(82) = 0; Stoic(83) = 0; Stoic(84) = 0; Stoic(85) = 0;
!Stoic(Q) >= 1;
ENDSUBMODEL

SUBMODEL Iteration10: !Lignin alcohols to produce aromatics, alkyl benzenes,
paraffins; !K = 25;

!Chemistry Constraints, only certain species may react with each other;

!Cannot use more than one type of alcohol reactant;

Reactant (75) + Reactant (76) + Reactant (77) <= 1;

!Reactant, Product, or Not in Reaction, Flag: Role Specification Constraints, and active reactions;
Stoic(1) <= 0;  Stoic(2) = 0;  Stoic(3) = 0;  Stoic(4) = 0;  Stoic(5) = 0;  Stoic(6) = 0;
Stoic(7) = 0;  Stoic(8) = 0;  Stoic(9) = 0;  Stoic(10) = 0;
Stoic(11) = 0;  Stoic(12) = 0;  Stoic(13) = 0;  Stoic(14) = 0;  Stoic(15) = 0;  Stoic(16) = 0;
Stoic(17) = 0;  Stoic(18) = 0;  Stoic(19) = 0;  Stoic(20) = 0;
Stoic(21) = 0;  Stoic(22) = 0;  Stoic(23) = 0;  Stoic(24) = 0;  Stoic(25) = 0;  Stoic(26) = 0;
Stoic(27) = 0;  Stoic(28) = 0;  Stoic(29) = 0;  Stoic(30) = 0;
Stoic(31) = 0;  Stoic(32) = 0;  Stoic(33) = 0;  Stoic(34) = 0;  Stoic(35) = 0;  Stoic(36) = 0;
Stoic(37) = 0;  Stoic(38) = 0;  Stoic(39) = 0;  Stoic(40) = 0;
Stoic(41) = 0;  Stoic(42) = 0;  Stoic(43) = 0;  Stoic(44) >= 0;  Stoic(45) >= 0;  Stoic(46) >=
0;  Stoic(47) >= 0;  Stoic(48) >= 0;  Stoic(49) >= 0;  Stoic(50) >= 0;
Stoic(51) >= 0;  Stoic(52) >= 0;  Stoic(53) >= 0;  Stoic(54) = 0;  Stoic(55) = 0;  Stoic(56) =
0;  Stoic(57) = 0;  Stoic(58) = 0;  Stoic(59) = 0;  Stoic(60) = 0;
Stoic(61) = 0;  Stoic(62) = 0;  Stoic(63) >= 0;  Stoic(64) >= 0;  Stoic(65) = 0;  Stoic(66) =
0;  Stoic(67) = 0;  Stoic(68) = 0;  Stoic(69) = 0;  Stoic(70) = 0;
Stoic(71) = 0;  Stoic(72) = 0;  Stoic(73) = 0;  Stoic(74) = 0;  Stoic(75) <= 0;  Stoic(76) <=
0;  Stoic(77) <= 0;  Stoic(78) = 0;  Stoic(79) = 0;  Stoic(80) = 0;
Stoic(81) = 0;  Stoic(82) = 0;  Stoic(83) = 0;  Stoic(84) >= 0;  Stoic(85) = 0;
ENDSUBMODEL

SUBMODEL Iteration11:  !Bio-Oil to produce aromatics;  !K = 25;

!Chemistry Constraints, only certain species may react with each other;
Reactant, Product, or Not in Reaction, Flag: Role Specification Constraints, and active reactions;

\[ \text{Stoic}(1) = 0; \text{Stoic}(2) \leq 0; \text{Stoic}(3) = 0; \text{Stoic}(4) = 0; \text{Stoic}(5) = 0; \text{Stoic}(6) = 0; \]
\[ \text{Stoic}(7) = 0; \text{Stoic}(8) = 0; \text{Stoic}(9) = 0; \text{Stoic}(10) = 0; \]
\[ \text{Stoic}(11) = 0; \text{Stoic}(12) = 0; \text{Stoic}(13) = 0; \text{Stoic}(14) = 0; \text{Stoic}(15) = 0; \text{Stoic}(16) = 0; \]
\[ \text{Stoic}(17) = 0; \text{Stoic}(18) = 0; \text{Stoic}(19) = 0; \text{Stoic}(20) = 0; \]
\[ \text{Stoic}(21) = 0; \text{Stoic}(22) = 0; \text{Stoic}(23) = 0; \text{Stoic}(24) = 0; \text{Stoic}(25) = 0; \text{Stoic}(26) = 0; \]
\[ \text{Stoic}(27) = 0; \text{Stoic}(28) = 0; \text{Stoic}(29) = 0; \text{Stoic}(30) = 0; \]
\[ \text{Stoic}(31) = 0; \text{Stoic}(32) = 0; \text{Stoic}(33) = 0; \text{Stoic}(34) = 0; \text{Stoic}(35) \geq 0; \text{Stoic}(36) \geq 0; \text{Stoic}(37) \geq 0; \text{Stoic}(38) \geq 0; \text{Stoic}(39) \geq 0; \text{Stoic}(40) = 0; \]
\[ \text{Stoic}(41) = 0; \text{Stoic}(42) = 0; \text{Stoic}(43) = 0; \text{Stoic}(44) \geq 0; \text{Stoic}(45) \geq 0; \text{Stoic}(46) \geq 0; \text{Stoic}(47) \geq 0; \text{Stoic}(48) \geq 0; \text{Stoic}(49) \geq 0; \text{Stoic}(50) \geq 0; \]
\[ \text{Stoic}(51) \geq 0; \text{Stoic}(52) = 0; \text{Stoic}(53) = 0; \text{Stoic}(54) = 0; \text{Stoic}(55) = 0; \text{Stoic}(56) = 0; \]
\[ \text{Stoic}(57) = 0; \text{Stoic}(58) = 0; \text{Stoic}(59) = 0; \text{Stoic}(60) = 0; \]
\[ \text{Stoic}(61) = 0; \text{Stoic}(62) = 0; \text{Stoic}(63) = 0; \text{Stoic}(64) \geq 0; \text{Stoic}(65) = 0; \text{Stoic}(66) = 0; \]
\[ \text{Stoic}(67) = 0; \text{Stoic}(68) \leq -1; \text{Stoic}(69) = 0; \text{Stoic}(70) = 0; \]
\[ \text{Stoic}(71) = 0; \text{Stoic}(72) = 0; \text{Stoic}(73) = 0; \text{Stoic}(74) = 0; \text{Stoic}(75) = 0; \text{Stoic}(76) = 0; \]
\[ \text{Stoic}(77) = 0; \text{Stoic}(78) = 0; \text{Stoic}(79) = 0; \text{Stoic}(80) = 0; \]
\[ \text{Stoic}(81) = 0; \text{Stoic}(82) = 0; \text{Stoic}(83) = 0; \text{Stoic}(84) = 0; \text{Stoic}(85) = 0; \]

ENDSUBMODEL
SUBMODEL Iteration12: !Syngas to produce Hydrogen, Methanol, Alkanes; !K = 25;

!Chemistry Constraints, only certain species may react with each other;

!Produce only one type of alkane with each reaction;

Product (35) + Product (36) + Product(37) + Product(38) + Product(39) +
Product(54) + Product (55) + Product(56) + Product(57) + Product (58) <= 1;

!Reactant, Product, or Not in Reaction, Flag: Role Specification Constraints, and active reactions;

Stoic(1) <= -1; Stoic(2) <= -1; Stoic(3) = 0; Stoic(4) = 0; Stoic(5) = 0; Stoic(6) = 0;
Stoic(7) = 0; Stoic(8) = 0; Stoic(9) = 0; Stoic(10) = 0;
Stoic(11) = 0; Stoic(12) = 0; Stoic(13) = 0; Stoic(14) = 0; Stoic(15) = 0; Stoic(16) = 0;
Stoic(17) = 0; Stoic(18) = 0; Stoic(19) = 0; Stoic(20) = 0;
Stoic(21) = 0; Stoic(22) = 0; Stoic(23) = 0; Stoic(24) = 0; Stoic(25) = 0; Stoic(26) = 0;
Stoic(27) = 0; Stoic(28) = 0; Stoic(29) = 0; Stoic(30) = 0;
Stoic(31) = 0; Stoic(32) = 0; Stoic(33) = 0; Stoic(34) >= 0; Stoic(35) >= 0; Stoic(36) >=
0; Stoic(37) >= 0; Stoic(38) >= 0; Stoic(39) >= 0; Stoic(40) = 0;
Stoic(41) = 0; Stoic(42) = 0; Stoic(43) = 0; Stoic(44) = 0; Stoic(45) = 0; Stoic(46) = 0;
Stoic(47) = 0; Stoic(48) = 0; Stoic(49) = 0; Stoic(50) = 0;
Stoic(51) = 0; Stoic(52) = 0; Stoic(53) = 0; Stoic(54) >= 0; Stoic(55) >= 0; Stoic(56) >=
0; Stoic(57) >= 0; Stoic(58) >= 0; Stoic(59) = 0; Stoic(60) = 0;
Stoic(61) = 0; Stoic(62) = 0; Stoic(63) = 0; Stoic(64) >= 0; Stoic(65) = 0; Stoic(66) = 0;
Stoic(67) = 0; Stoic(68) = 0; Stoic(69) = 0; Stoic(70) = 0;
Stoic(71) = 0; Stoic(72) = 0; Stoic(73) = 0; Stoic(74) = 0; Stoic(75) = 0; Stoic(76) = 0;
Stoic(77) = 0; Stoic(78) = 0; Stoic(79) = 0; Stoic(80) = 0;
Stoic(81) = 0; Stoic(82) = 0; Stoic(83) = 0; Stoic(84) = 0; Stoic(85) = 0;

!Stoic(Q) >= 1;

ENDSUBMODEL

SUBMODEL Iteration13: !Gasoline and Olefins from Methanol; !K = 25;

!Chemistry Constraints, only certain species may react with each other;

!Methanol cannot react with olefins:

Reactant(34) + Reactant(42) <= 1;

Reactant(34) + Reactant(43) <= 1;

!Reactant, Product, or Not in Reaction, Flag: Role Specification Constraints, and active reactions:

Stoic(1) = 0; Stoic(2) <= 0; Stoic(3) = 0; Stoic(4) = 0; Stoic(5) = 0; Stoic(6) = 0;
Stoic(7) = 0; Stoic(8) = 0; Stoic(9) = 0; Stoic(10) = 0;
Stoic(11) = 0; Stoic(12) = 0; Stoic(13) = 0; Stoic(14) = 0; Stoic(15) = 0; Stoic(16) = 0;
Stoic(17) = 0; Stoic(18) = 0; Stoic(19) = 0; Stoic(20) = 0;
Stoic(21) = 0; Stoic(22) = 0; Stoic(23) = 0; Stoic(24) = 0; Stoic(25) = 0; Stoic(26) = 0;
Stoic(27) = 0; Stoic(28) = 0; Stoic(29) = 0; Stoic(30) = 0;
Stoic(31) = 0; Stoic(32) = 0; Stoic(33) = 0; Stoic(34) <= 0; Stoic(35) = 0; Stoic(36) = 0; 
Stoic(37) = 0; Stoic(38) >= 0; Stoic(39) >= 0; Stoic(40) >= 0; 
Stoic(41) >= 0; Stoic(42) >= 0; Stoic(43) >= 0; Stoic(44) = 0; Stoic(45) = 0; Stoic(46) = 0; 
Stoic(47) = 0; Stoic(48) = 0; Stoic(49) = 0; Stoic(50) = 0; 
Stoic(51) = 0; Stoic(52) = 0; Stoic(53) = 0; Stoic(54) = 0; Stoic(55) = 0; Stoic(56) = 0; 
Stoic(57) = 0; Stoic(58) = 0; Stoic(59) = 0; Stoic(60) = 0; 
Stoic(61) = 0; Stoic(62) = 0; Stoic(63) = 0; Stoic(64) >= 0; Stoic(65) = 0; Stoic(66) = 0; 
Stoic(67) = 0; Stoic(68) = 0; Stoic(69) = 0; Stoic(70) = 0; 
Stoic(71) = 0; Stoic(72) = 0; Stoic(73) = 0; Stoic(74) = 0; Stoic(75) = 0; Stoic(76) = 0; 
Stoic(77) = 0; Stoic(78) = 0; Stoic(79) = 0; Stoic(80) = 0; 
Stoic(81) = 0; Stoic(82) = 0; Stoic(83) = 0; Stoic(84) = 0; Stoic(85) = 0; 

!Stoic(Q) >= 1; 
ENDSUBMODEL 

!Iteration 14 and 15 are reaction stoichiometry mass balances based on 1 known reaction 
each; 

SUBMODEL KBESTCUTS: !Allows for integer cuts to each stoichiometry model; 

@FOR( KSOLUTIONS( I2) | I2 #LE# I: 

[R_CUT] @SUM( Coefficients( J): CUT( J, I2) * (Product( J)+ Reactant ( J))) >= 
RHS( I2) 
); 
ENDSUBMODEL
CALC:

! Set some parameters;

@SET ('DEFAULT');

@SET ( 'TERSEO', 2);

! Resolve the model with K cut solutions for total reactions;

Q = 0; ! sets the target stoichiometry molecule;

Z = 0; ! first iteration of the moveable target molecule;

@WHILE( Z #LT# 1: ! loops over different target molecules;

I = 0;

ISTATUS = 0;

@WHILE( I #LT# K #AND# ISTATUS #EQ# 0: ! loops over each reaction stoichiometry;

! @GEN( Balances, KBESTCUTS);

! Solve current base model;

@SOLVE( Base, KBESTCUTS, Iteration10);

ISTATUS = @STATUS();

! Generate next cut to cutoff current solution;

I = I + 1;

RHS( I) = 1;

@FOR( Coefficients( J):

@IFC( ((Product( J)+ Reactant ( J)))#LT# .1:

CUT( J, I) = 1;
@ELSE

CUT( J, I) = -1;
RHS( I) = RHS( I) - 1;

);;
)

! Save current K cut solution;

OBJ( I) = R_OBJ;

@FOR( Coefficients( J): INCLUDE2( J, I) = Stoic( J));

);;

Z = Z + 1; !updates the loop to again solve for a new target molecule;
Q = Q + 1; !updates the target molecule;

!Save each molecule target current solution;

Points( Z) = R_OBJ;

@FOR( KXI( L, N): INCLUDE3(Z, L, N) = INCLUDE2( L, N));

);

! Send solutions to Excel;

@SET('FILOUT', 1);

@OLE('C:\Users\ejp7950\Desktop\Flash Drive\GasStoicRxns.xlsx',
'Stoichiometry10') = INCLUDE3;

ENDCALC

END
A.2 RNFA code formulation in LINGO

MODEL:

DATA:

! Number of K-Best Solutions Wanted;

    K = 15;

ENDDATA

SETS:

Compounds / v1..v85 / : Amount;

Reactions / T1..T277/ : Total;

MatrixA (Compounds, Reactions): Stoichiometry;

MolarFlow / Rxn1..Rxn277 / : Flow, Flowlower, Flowupper, Flowzero, FlowMinus,
Flowplus, Wzero, Wminus, Wplus, CARD;

Inputs (MolarFlow) / Rxn277/:  ;

KSOLUTIONS /1..K/: OBJ, RHS;

KXI( Reactions, KSOLUTIONS): CUT, INCLUDE2;

ENDSETS

DATA:

! Import the data from Excel;
Stoichiometry = @OLE('C:\Users\ejp7950\Desktop\Flash Drive\GasStoicRxns.XLSX', 'Input');

ENDDATA

SUBMODEL GF:

!Objective Function;

MAX = Flow(264) ; !Max production of Octane;

!Subject to Ax = 0;

@FOR(COMPounds(J): @SUM (Reactions(L): Stoichiometry(J,L) *Flow(L)) = 0);

!Subject to  100 >= x >=0;

@FOR (MolarFlow (I) : Flowupper(I) = 100);

@FOR (MolarFlow (I) : Flowlower(I) = 0.1);

!Reactant Flow in order to avoid infinite reaction scenarios;

!Biomass input;

!Flow(268) = 1;

!Every other reactant input = 0;

!Flow Constraints around Biomass;

!Flow Constraints around Cellulose;

!Flow Constraints around Hemicellulose;

!Flow Constraints around Lignin;

!Flow Constraints around Biooil;

!Flow Constraints around Syngas;


!Flow Constraints around C5 Sugars;


!Flow Constraints around C6 Sugars;

Flow(51) + Flow (52) + Flow (53) + Flow(54) + Flow(55) +
Flow(56) + Flow (57) + Flow (58) + Flow(59) + Flow(60) +
Flow(61) + Flow (62) + Flow (63) + Flow(64) + Flow(65) +
Flow(66) + Flow (67) + Flow (68) + Flow(69) + Flow(70) +
Flow(71) + Flow (72) + Flow (73) + Flow(74) + Flow(75) - Flow(256) + Flow(258) ) = 0;

!Flow Constraints around C5 Sugars;
ENDSUBMODEL

SUBMODEL NF:

!NF.1 Constraints, for iterations k = 2 and larger;

!Condition (5a) splitting up the flux variables;
@FOR( MolarFlow(I): @Free( Flowzero(I)));
@FOR( MolarFlow(I): @Free( Flowminus(I)));
@FOR( MolarFlow(I): @Free( Flowplus(I)));
@FOR (MolarFlow(I): Flow(I) = Flowzero(I) + FlowMinus(I) + Flowplus (I));

!Condition (5b) introducing binary variables;

@FOR( MolarFlow(I): @BIN( Wzero(I)));
@FOR( MolarFlow(I): @BIN( Wminus(I)));
@FOR( MolarFlow(I): @BIN( Wplus(I)));
@FOR (MolarFlow(I): Wzero(I) + Wminus(I) + Wplus(I) = 1 );

!Coupling the binary variables (5c);

Eta = 0.000001; ! Eta = 1*10^-6;

@FOR (MolarFlow (I) : (Flowlower(I)*(1-Wzero(I))) <= Flow(I));

@FOR (MolarFlow (I) : Flow(I) <= Flowupper(I)*(1-Wzero(I)));

!Coupling the binary variables (5d);

@FOR (MolarFlow (I) : -Eta*Wzero(I) <= Flowzero(I));

@FOR (MolarFlow (I) : Flowzero(I) <= Eta*Wzero(I));

!Coupling the binary variables (5e);

@FOR (MolarFlow (I) : Flowlower(I)*Wminus(I) <= Flowminus(I));

@FOR (MolarFlow (I) : Flowminus(I) <= Eta*Wminus(I));

!Coupling the binary variables (5f);

@FOR (MolarFlow (I) : Eta*Wplus(I) <= Flowplus(I));

@FOR (MolarFlow (I) : Flowplus(I) <= Wplus(I)*Flowupper(I));

!Cardinality limit on active flows;

CARD1 = 250;

[R_Obj] CARD1 = @SUM (MolarFlow(I): WZero (I)) ;

ENDSUBMODEL

SUBMODEL KBESTCUTS:
@FOR( KSOLUTIONS( I2) | I2 #LE# I:

[R_CUT] @SUM( MolarFlow( J): CUT( J, I2) * (Wzero(J)))  >= RHS( I2)
);

ENDSUBMODEL

CALC:

!Set some parameters;

@SET( 'DEFAULT');

@SET( 'TERSEO', 2);

I = 0;

ISTATUS = 0;

@WHILE( I #LT# K #AND# ISTATUS #EQ# 0:

!Solve current base model;

@SOLVE( GF, NF, KBESTCUTS);

ISTATUS = @STATUS();

!Generate next cut to cutoff current solution;

I = I + 1;

RHS( I) = 1;

@FOR( MolarFlow( J):

@IFC( (Wzero(J))#LT# .1:

CUT( J, I) = 1;

@ELSE

CUT( J, I) = -1;
RHS( I) = RHS( I) - 1;

);;

);;

!Save current K cut solution;

OBJ( I) = R_OBJ;

@FOR( Reactions( J): INCLUDE2( J, I) = Flow( J));

);

! Send solutions to Excel;

@SET( 'FILOUT', 1);

@OLE( 'C:\Users\ejp7950\Desktop\Flash Drive\GasStoicRxns.XLSX', 'Output') = Include2;

ENDCALC

END
### APPENDIX B

#### B.1 Compounds used in the case study

Compounds used in the case study:

<table>
<thead>
<tr>
<th>1. Carbon Monoxide</th>
<th>31. 1,4-pentanediol</th>
<th>63. Carbon Dioxide</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Hydrogen</td>
<td>32. 1,2-</td>
<td>64. Water</td>
</tr>
<tr>
<td>3. Formic Acid</td>
<td>dihydroxypentana</td>
<td>65. Glycerol</td>
</tr>
<tr>
<td>4. Acetic Acid</td>
<td>33. 1-pentanol</td>
<td>66. Oxygen</td>
</tr>
<tr>
<td>5. Propanoic Acid</td>
<td>34. Methanol</td>
<td>67. Woody Biomass</td>
</tr>
<tr>
<td>7. Formaldehyde</td>
<td>36. Ethane</td>
<td></td>
</tr>
<tr>
<td>8. Acetaldehyde</td>
<td>37. Propane</td>
<td>69. Char (from Woody Biomass)</td>
</tr>
<tr>
<td>9. Oxaldehyde</td>
<td>38. n-Butane</td>
<td>70. Pyrolysis Gas</td>
</tr>
<tr>
<td>10. Glycol-aldehyde</td>
<td>39. n-Octane</td>
<td>(from Woody Biomass)</td>
</tr>
<tr>
<td>11. Acetal</td>
<td>40. Ethene</td>
<td></td>
</tr>
<tr>
<td>12. Cellulose Acetate</td>
<td>41. Propene</td>
<td>71. Cellulose</td>
</tr>
<tr>
<td>13. Cellobiose</td>
<td>42. But-1-ene</td>
<td>72. Hemicellulose</td>
</tr>
<tr>
<td>14. Fructose</td>
<td>43. But-2-ene</td>
<td>73. Lignin</td>
</tr>
<tr>
<td>15. Glucose</td>
<td>44. Benzene</td>
<td>74. Biomass</td>
</tr>
<tr>
<td>16. DME(Dimethyl Ether)</td>
<td>45. Toluene</td>
<td>(Woody)</td>
</tr>
<tr>
<td>17. HMF</td>
<td>46. p-Xylene</td>
<td>75. Coumaryl</td>
</tr>
<tr>
<td>18. Furfural</td>
<td>47. m-Xylene</td>
<td>Alcohol</td>
</tr>
<tr>
<td>19. Phenol</td>
<td>48. o-Xylene</td>
<td>76. Coniferyl_Alcohol</td>
</tr>
<tr>
<td>20. Isoeugenol</td>
<td>49. o-Cresol</td>
<td>1</td>
</tr>
<tr>
<td>21. Eugenol</td>
<td>50. p-Cresol</td>
<td>77. Sinapyl_Alochol</td>
</tr>
<tr>
<td>22. Methyl Guaiacol</td>
<td>51. m-Cresol</td>
<td>78. Levulínico Ester (from 1-Hexene)</td>
</tr>
<tr>
<td>23. Xylose</td>
<td>52. Indane</td>
<td>79. 1-Hexene</td>
</tr>
<tr>
<td>24. Furfural</td>
<td>53. Tetralin</td>
<td>80. Propanol</td>
</tr>
<tr>
<td>25. Furfuryl Alcohol</td>
<td>54. n-Nonane</td>
<td>81. Butanol</td>
</tr>
<tr>
<td>26. Methyl Furan</td>
<td>55. n-Decane</td>
<td>82. Cyclohexene</td>
</tr>
<tr>
<td>27. Tetrahydrofurfury</td>
<td>56. u-Decane</td>
<td>83. Cyclohexene</td>
</tr>
<tr>
<td></td>
<td>57. n-Dodecane</td>
<td>84. Parrafin</td>
</tr>
<tr>
<td></td>
<td>58. n-Tridecane</td>
<td>85. MTBE</td>
</tr>
<tr>
<td>28. Levulinic Acid</td>
<td>59. 2-MTHF</td>
<td></td>
</tr>
<tr>
<td>29. Angelica Lactone</td>
<td>60. Ethanol</td>
<td></td>
</tr>
<tr>
<td>30. Gamma</td>
<td>61. n-Pentane</td>
<td></td>
</tr>
<tr>
<td>Valerolactone</td>
<td>62. n-Hexane</td>
<td></td>
</tr>
</tbody>
</table>
# APPENDIX C

## C.1 Compounds Thermodynamic Data

**Table C1:** Compounds used in the case study and their thermodynamic data and method to obtain that data

<table>
<thead>
<tr>
<th>Compound Name</th>
<th>C</th>
<th>H</th>
<th>O</th>
<th>Gf[298K]</th>
<th>Hf[298K]</th>
<th>METHOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>carbon monoxide</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>-137.15</td>
<td>-110.53</td>
<td>DIPPR</td>
</tr>
<tr>
<td>hydrogen</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>DIPPR</td>
</tr>
<tr>
<td>formic acid</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>-278.83</td>
<td>-301.3</td>
<td>MG</td>
</tr>
<tr>
<td>acetic acid</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>-389</td>
<td>-484.5</td>
<td>DIPPR</td>
</tr>
<tr>
<td>propanoic acid</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>-384.6</td>
<td>-510.866</td>
<td>DIPPR</td>
</tr>
<tr>
<td>acetone</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>-155.4</td>
<td>-248.1</td>
<td>DIPPR</td>
</tr>
<tr>
<td>formaldehyde</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>-102.6</td>
<td>-108.6</td>
<td>DIPPR</td>
</tr>
<tr>
<td>acetaldehyde</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>-133.3</td>
<td>-166.4</td>
<td>DIPPR</td>
</tr>
<tr>
<td>oxaldehyde</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>-192.2</td>
<td>-244.4</td>
<td>DIPPR</td>
</tr>
<tr>
<td>glycol aldehyde</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>-289</td>
<td>-404.2</td>
<td>DIPPR</td>
</tr>
<tr>
<td>acetal</td>
<td>6</td>
<td>14</td>
<td>2</td>
<td>-251.8</td>
<td>-491.41</td>
<td>DIPPR</td>
</tr>
<tr>
<td>cellulose acetate</td>
<td>76</td>
<td>114</td>
<td>49</td>
<td>N/A</td>
<td>N/A</td>
<td>MG</td>
</tr>
<tr>
<td>cellobiose</td>
<td>12</td>
<td>22</td>
<td>11</td>
<td>-1477.51</td>
<td>-1895</td>
<td>MG</td>
</tr>
<tr>
<td>fructose</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>-840.72</td>
<td>-1118.76</td>
<td>MG</td>
</tr>
<tr>
<td>glucose</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>-877.39</td>
<td>-1127.36</td>
<td>MG</td>
</tr>
<tr>
<td>dimethyl ether</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>-112.8</td>
<td>-184.1</td>
<td>DIPPR</td>
</tr>
<tr>
<td>HMF (5-Hydroxymethyl-2-furaldehyde)</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>-287.7</td>
<td>-368.47</td>
<td>MG</td>
</tr>
<tr>
<td>furan-2-carbaldehyde</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>-102.56</td>
<td>-179.2</td>
<td>MG</td>
</tr>
<tr>
<td>phenol</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>-50.41</td>
<td>-165.1</td>
<td>DIPPR</td>
</tr>
<tr>
<td>isoeugenol</td>
<td>10</td>
<td>12</td>
<td>2</td>
<td>-41.82</td>
<td>-193.58</td>
<td>MG</td>
</tr>
<tr>
<td>eugenol</td>
<td>10</td>
<td>12</td>
<td>2</td>
<td>-41.82</td>
<td>-193.58</td>
<td>MG</td>
</tr>
<tr>
<td>cresol</td>
<td>8</td>
<td>10</td>
<td>2</td>
<td>-33.94</td>
<td>-130.92</td>
<td>MG</td>
</tr>
<tr>
<td>xylose</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>-729.71</td>
<td>-940.54</td>
<td>MG</td>
</tr>
<tr>
<td>furfural</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>-119</td>
<td>-201.6</td>
<td>DIPPR</td>
</tr>
<tr>
<td>furfuryl alcohol</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>-154.1</td>
<td>-276.2</td>
<td>DIPPR</td>
</tr>
<tr>
<td>methyl furan</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>10.06</td>
<td>-93.62</td>
<td>MG</td>
</tr>
<tr>
<td>tetrahydrofurfural</td>
<td>5</td>
<td>10</td>
<td>2</td>
<td>-200.17</td>
<td>-311.31</td>
<td>MG</td>
</tr>
<tr>
<td>levulinic acid</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td>-512.2</td>
<td>-697.054</td>
<td>DIPPR</td>
</tr>
<tr>
<td>Compound Name</td>
<td>C</td>
<td>H</td>
<td>O</td>
<td>Gf[298K]</td>
<td>Hf[298K]</td>
<td>METHOD</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---------</td>
<td>---------</td>
<td>--------</td>
</tr>
<tr>
<td>angelica lactone</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>-225.22</td>
<td>-309.64</td>
<td>MG</td>
</tr>
<tr>
<td>gamma valerolactone</td>
<td>5</td>
<td>8</td>
<td>2</td>
<td>-268.49</td>
<td>-383.93</td>
<td>MG</td>
</tr>
<tr>
<td>1,4-pentanediol</td>
<td>5</td>
<td>12</td>
<td>2</td>
<td>-289.92</td>
<td>-476.88</td>
<td>MG</td>
</tr>
<tr>
<td>1,2-dihydroxypentanal</td>
<td>5</td>
<td>12</td>
<td>1</td>
<td>-155.8</td>
<td>-351.9</td>
<td>DIPPR</td>
</tr>
<tr>
<td>1-pentanol</td>
<td>5</td>
<td>12</td>
<td>1</td>
<td>-169.1</td>
<td>-366.5</td>
<td>DIPPR</td>
</tr>
<tr>
<td>methanol</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>-166.9</td>
<td>-239.1</td>
<td>DIPPR</td>
</tr>
<tr>
<td>methanol</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>-50.49</td>
<td>-74.52</td>
<td>DIPPR</td>
</tr>
<tr>
<td>ethane</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>-31.92</td>
<td>-83.82</td>
<td>DIPPR</td>
</tr>
<tr>
<td>propane</td>
<td>3</td>
<td>8</td>
<td>0</td>
<td>-24.39</td>
<td>-104.68</td>
<td>DIPPR</td>
</tr>
<tr>
<td>n-butane</td>
<td>4</td>
<td>10</td>
<td>0</td>
<td>-16.7</td>
<td>-125.79</td>
<td>DIPPR</td>
</tr>
<tr>
<td>n-octane</td>
<td>8</td>
<td>18</td>
<td>0</td>
<td>6.587</td>
<td>-249.78</td>
<td>DIPPR</td>
</tr>
<tr>
<td>ethene</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>68.44</td>
<td>52.51</td>
<td>DIPPR</td>
</tr>
<tr>
<td>propene</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>62.64</td>
<td>20.23</td>
<td>DIPPR</td>
</tr>
<tr>
<td>but-1-ene</td>
<td>4</td>
<td>8</td>
<td>0</td>
<td>70.41</td>
<td>-0.5</td>
<td>DIPPR</td>
</tr>
<tr>
<td>but-2-ene</td>
<td>4</td>
<td>8</td>
<td>0</td>
<td>-222.3</td>
<td>-401.6</td>
<td>DIPPR</td>
</tr>
<tr>
<td>benzene</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>124.4</td>
<td>48.95</td>
<td>DIPPR</td>
</tr>
<tr>
<td>toluene</td>
<td>7</td>
<td>8</td>
<td>0</td>
<td>113.8</td>
<td>12.01</td>
<td>DIPPR</td>
</tr>
<tr>
<td>p-xylene</td>
<td>8</td>
<td>10</td>
<td>0</td>
<td>111.4</td>
<td>-24.35</td>
<td>DIPPR</td>
</tr>
<tr>
<td>m-xylene</td>
<td>8</td>
<td>10</td>
<td>0</td>
<td>107.3</td>
<td>-25.36</td>
<td>DIPPR</td>
</tr>
<tr>
<td>o-xylene</td>
<td>8</td>
<td>10</td>
<td>0</td>
<td>110.6</td>
<td>-24.35</td>
<td>DIPPR</td>
</tr>
<tr>
<td>o-cresol</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>-55.66</td>
<td>-204.6</td>
<td>DIPPR</td>
</tr>
<tr>
<td>p-cresol</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>-50.901</td>
<td>-199.28</td>
<td>DIPPR</td>
</tr>
<tr>
<td>m-cresol</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>-59.1</td>
<td>-194</td>
<td>DIPPR</td>
</tr>
<tr>
<td>indane</td>
<td>9</td>
<td>10</td>
<td>0</td>
<td>152</td>
<td>11.7</td>
<td>DIPPR</td>
</tr>
<tr>
<td>tetralin</td>
<td>10</td>
<td>12</td>
<td>0</td>
<td>163.99</td>
<td>39.15</td>
<td>MG</td>
</tr>
<tr>
<td>n-nonane</td>
<td>9</td>
<td>20</td>
<td>0</td>
<td>12.647</td>
<td>-274.68</td>
<td>DIPPR</td>
</tr>
<tr>
<td>n-decane</td>
<td>10</td>
<td>22</td>
<td>0</td>
<td>17.74</td>
<td>-300.62</td>
<td>DIPPR</td>
</tr>
<tr>
<td>n-undecane</td>
<td>11</td>
<td>24</td>
<td>0</td>
<td>22.78</td>
<td>-326.6</td>
<td>DIPPR</td>
</tr>
<tr>
<td>n-dodecane</td>
<td>12</td>
<td>26</td>
<td>0</td>
<td>28.203</td>
<td>-352.13</td>
<td>DIPPR</td>
</tr>
<tr>
<td>Compound Name</td>
<td>C</td>
<td>H</td>
<td>O</td>
<td>Gf[298K]</td>
<td>Hf[298K]</td>
<td>METHOD</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----------</td>
<td>----------</td>
<td>------------</td>
</tr>
<tr>
<td>n-tridecane</td>
<td>13</td>
<td>28</td>
<td>0</td>
<td>33.681</td>
<td>-377.69</td>
<td>DIPPR</td>
</tr>
<tr>
<td>2-methyltetrahydrofuran</td>
<td>5</td>
<td>10</td>
<td>1</td>
<td>-94.37</td>
<td>-221.33</td>
<td>MG</td>
</tr>
<tr>
<td>ethanol</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>-173.86</td>
<td>-276.98</td>
<td>DIPPR</td>
</tr>
<tr>
<td>n-pentane</td>
<td>5</td>
<td>12</td>
<td>0</td>
<td>-9.928</td>
<td>-173.51</td>
<td>DIPPR</td>
</tr>
<tr>
<td>n-hexane</td>
<td>6</td>
<td>14</td>
<td>0</td>
<td>-4.154</td>
<td>-198.66</td>
<td>DIPPR</td>
</tr>
<tr>
<td>carbon dioxide</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>-394.37</td>
<td>-393.51</td>
<td>DIPPR</td>
</tr>
<tr>
<td>water</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>-237.214</td>
<td>-285.83</td>
<td>DIPPR</td>
</tr>
<tr>
<td>glycerol</td>
<td>3</td>
<td>8</td>
<td>3</td>
<td>-478.6</td>
<td>-669.6</td>
<td>DIPPR</td>
</tr>
<tr>
<td>Oxygen</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>DIPPR</td>
</tr>
<tr>
<td>Wood</td>
<td>1</td>
<td>1.6</td>
<td>0.68</td>
<td>N/A</td>
<td>N/A</td>
<td>MG</td>
</tr>
<tr>
<td>BioOil</td>
<td>1</td>
<td>2.59</td>
<td>1.05</td>
<td>N/A</td>
<td>N/A</td>
<td>MG</td>
</tr>
<tr>
<td>Char</td>
<td>1</td>
<td>0.72</td>
<td>0.22</td>
<td>N/A</td>
<td>N/A</td>
<td>MG</td>
</tr>
<tr>
<td>Pyrolysis Gas</td>
<td>1</td>
<td>0.54</td>
<td>1.3</td>
<td>N/A</td>
<td>N/A</td>
<td>MG</td>
</tr>
<tr>
<td>cellulose</td>
<td>6</td>
<td>10</td>
<td>5</td>
<td>N/A</td>
<td>N/A</td>
<td>MG</td>
</tr>
<tr>
<td>hemicellulose</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>N/A</td>
<td>N/A</td>
<td>MG</td>
</tr>
<tr>
<td>lignin</td>
<td>10</td>
<td>12</td>
<td>3</td>
<td>N/A</td>
<td>N/A</td>
<td>MG</td>
</tr>
<tr>
<td>biomass(woody)</td>
<td>1</td>
<td>1.6</td>
<td>0.68</td>
<td>N/A</td>
<td>N/A</td>
<td>MG</td>
</tr>
<tr>
<td>coumaryl alcohol</td>
<td>9</td>
<td>10</td>
<td>2</td>
<td>-82.66</td>
<td>-217.92</td>
<td>MG</td>
</tr>
<tr>
<td>coniferyl alcohol</td>
<td>10</td>
<td>12</td>
<td>3</td>
<td>-193.46</td>
<td>-364.45</td>
<td>MG</td>
</tr>
<tr>
<td>sinapyl alcohol</td>
<td>11</td>
<td>14</td>
<td>4</td>
<td>-296.08</td>
<td>-505.39</td>
<td>MG</td>
</tr>
<tr>
<td>Ethyl 4-oxo-5-phenylpentanoate</td>
<td>11</td>
<td>20</td>
<td>3</td>
<td>-274.97</td>
<td>-513.53</td>
<td>MG</td>
</tr>
<tr>
<td>1-Hexene</td>
<td>6</td>
<td>12</td>
<td>0</td>
<td>83.6</td>
<td>-72.24</td>
<td>DIPPR</td>
</tr>
<tr>
<td>Propan-1-ol</td>
<td>3</td>
<td>8</td>
<td>1</td>
<td>-167</td>
<td>-300.8</td>
<td>DIPPR</td>
</tr>
<tr>
<td>n-Butanol</td>
<td>4</td>
<td>10</td>
<td>1</td>
<td>-161.4</td>
<td>-326.4</td>
<td>DIPPR</td>
</tr>
<tr>
<td>CycloHexane</td>
<td>6</td>
<td>12</td>
<td>0</td>
<td>26.77</td>
<td>-156.15</td>
<td>DIPPR</td>
</tr>
<tr>
<td>Cyclohexene</td>
<td>6</td>
<td>10</td>
<td>0</td>
<td>103</td>
<td>-38.2</td>
<td>DIPPR</td>
</tr>
<tr>
<td>Parrafin</td>
<td>20</td>
<td>42</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>MG</td>
</tr>
<tr>
<td>MTBE (tert-Butyl methyl ether)</td>
<td>5</td>
<td>12</td>
<td>1</td>
<td>-104.72</td>
<td>-271.82</td>
<td>MG</td>
</tr>
</tbody>
</table>

Note: DIPPR refers to the DIPPR database values, while MG refers to the Marrero and Gani method of property prediction for a molecule.
## C.2 Compounds Chemical Potential Data

**Table C2:** Chemical potentials of some common compounds:

<table>
<thead>
<tr>
<th>Compound</th>
<th>S/L/G</th>
<th>$\mu$ (kJ)</th>
<th>Temp Coff. G/K</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>gaseous</td>
<td>669.58</td>
<td>-157.99</td>
</tr>
<tr>
<td>C diamond</td>
<td>solid</td>
<td>2.90</td>
<td>-2.38</td>
</tr>
<tr>
<td>C graphite</td>
<td>solid</td>
<td>0.00</td>
<td>-5.74</td>
</tr>
<tr>
<td>C graphite</td>
<td>solid</td>
<td>0.00</td>
<td>-5.69</td>
</tr>
<tr>
<td>CH</td>
<td>gaseous</td>
<td>560.75</td>
<td>-182.92</td>
</tr>
<tr>
<td>CHO$_2^-$ formiate</td>
<td>aq. sol.</td>
<td>-351.04</td>
<td>-92.05</td>
</tr>
<tr>
<td>CH$_2$</td>
<td>gaseous</td>
<td>371.87</td>
<td>-181.04</td>
</tr>
<tr>
<td>CH$_2$ polyethylene</td>
<td>solid</td>
<td>4.40</td>
<td>-25.34</td>
</tr>
<tr>
<td>CH$_3$O formaldehyde</td>
<td>gaseous</td>
<td>-112.97</td>
<td>-218.66</td>
</tr>
<tr>
<td>CH$_2$O formaldehyde</td>
<td>gaseous</td>
<td>-109.90</td>
<td>-218.66</td>
</tr>
<tr>
<td>CH$_2$O formaldehyde</td>
<td>gaseous</td>
<td>-109.93</td>
<td>-218.80</td>
</tr>
<tr>
<td>CH$_2$O$_2$ formic acid</td>
<td>gaseous</td>
<td>-350.03</td>
<td>-251.60</td>
</tr>
<tr>
<td>CH$_2$O$_2$ formic acid</td>
<td>liquid</td>
<td>-359.57</td>
<td>-129.00</td>
</tr>
<tr>
<td>CH$_2$O$_2$ formic acid</td>
<td>aq. sol.</td>
<td>-372.38</td>
<td>-163.18</td>
</tr>
<tr>
<td>CH$_3$</td>
<td>gaseous</td>
<td>147.92</td>
<td>-194.05</td>
</tr>
<tr>
<td>CH$_4$</td>
<td>gaseous</td>
<td>-50.75</td>
<td>-186.15</td>
</tr>
<tr>
<td>CH$_4$ methane</td>
<td>gaseous</td>
<td>-50.81</td>
<td>-186.10</td>
</tr>
<tr>
<td>CH$_3$ methane</td>
<td>gaseous</td>
<td>-50.89</td>
<td>-186.19</td>
</tr>
<tr>
<td>CH$_3$O methanol</td>
<td>gaseous</td>
<td>-162.52</td>
<td>-239.70</td>
</tr>
<tr>
<td>CH$_2$O methanol</td>
<td>liquid</td>
<td>-166.36</td>
<td>-126.78</td>
</tr>
<tr>
<td>CH$_3$O methanol</td>
<td>liquid</td>
<td>-166.34</td>
<td>-126.70</td>
</tr>
<tr>
<td>CO</td>
<td>gaseous</td>
<td>-137.15</td>
<td>-197.56</td>
</tr>
<tr>
<td>CO</td>
<td>gaseous</td>
<td>-137.16</td>
<td>-197.53</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>gaseous</td>
<td>-394.36</td>
<td>-213.64</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>gaseous</td>
<td>-394.40</td>
<td>-213.68</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>aq. sol.</td>
<td>-385.99</td>
<td>117.53</td>
</tr>
<tr>
<td>CO$_2$ in air</td>
<td>solution</td>
<td>-4114.32</td>
<td>-281.21</td>
</tr>
<tr>
<td>CO$_3^{2-}$ carbonate ion</td>
<td>aq. sol.</td>
<td>-527.90</td>
<td>56.90</td>
</tr>
<tr>
<td>CO$_3$H$^+$ hydrogen-carbonate ion</td>
<td>aq. sol.</td>
<td>-586.85</td>
<td>-91.21</td>
</tr>
<tr>
<td>CO$_2$H$_2$</td>
<td>aq. sol.</td>
<td>-608.25</td>
<td>N/A</td>
</tr>
<tr>
<td>C$_2$H$_2$ ethyne</td>
<td>gaseous</td>
<td>209.20</td>
<td>-200.83</td>
</tr>
<tr>
<td>C$_2$H$_2$ ethyne</td>
<td>gaseous</td>
<td>209.17</td>
<td>-200.85</td>
</tr>
<tr>
<td>C$_2$H$_2$ ethyne</td>
<td>gaseous</td>
<td>209.24</td>
<td>-200.80</td>
</tr>
<tr>
<td>Compound</td>
<td>S/L/G</td>
<td>μ (kG)</td>
<td>Temp Coff. G/K</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------</td>
<td>----------</td>
<td>----------------</td>
</tr>
<tr>
<td>C₂H₃O₂⁻ acetic acid anion</td>
<td>aq. sol.</td>
<td>-369.41</td>
<td>-86.61</td>
</tr>
<tr>
<td>C₂H₄ ethene, ethylene</td>
<td>gaseous</td>
<td>68.12</td>
<td>-219.45</td>
</tr>
<tr>
<td>C₂H₄ ethene, ethylene</td>
<td>gaseous</td>
<td>68.36</td>
<td>-219.22</td>
</tr>
<tr>
<td>C₂H₄ ethene, ethylene</td>
<td>gaseous</td>
<td>68.16</td>
<td>-219.40</td>
</tr>
<tr>
<td>C₂H₄O acetaldehyde</td>
<td>gaseous</td>
<td>-132.92</td>
<td>-264.20</td>
</tr>
<tr>
<td>C₂H₄O epoxideethane</td>
<td>gaseous</td>
<td>-11.84</td>
<td>-243.70</td>
</tr>
<tr>
<td>C₂H₄O₂ acetic acid</td>
<td>gaseous</td>
<td>-378.95</td>
<td>-282.50</td>
</tr>
<tr>
<td>C₂H₄O₂ acetic acid</td>
<td>liquid</td>
<td>-389.95</td>
<td>-159.83</td>
</tr>
<tr>
<td>C₂H₄O₂ acetic acid</td>
<td>liquid</td>
<td>-390.28</td>
<td>-159.80</td>
</tr>
<tr>
<td>C₂H₄O₂ acetic acid</td>
<td>aq. sol.</td>
<td>-396.56</td>
<td>-178.66</td>
</tr>
<tr>
<td>C₂H₄O₄ formic acid, dimer</td>
<td>gaseous</td>
<td>-713.17</td>
<td>-348.70</td>
</tr>
<tr>
<td>C₂H₆ ethane</td>
<td>gaseous</td>
<td>-32.89</td>
<td>-229.49</td>
</tr>
<tr>
<td>C₂H₆ ethane</td>
<td>gaseous</td>
<td>-32.62</td>
<td>-229.50</td>
</tr>
<tr>
<td>C₂H₆O dimethyl ether</td>
<td>gaseous</td>
<td>-114.07</td>
<td>-266.60</td>
</tr>
<tr>
<td>C₂H₆O ethanol</td>
<td>gaseous</td>
<td>-168.57</td>
<td>-282.00</td>
</tr>
<tr>
<td>C₂H₆O ethanol</td>
<td>liquid</td>
<td>-174.89</td>
<td>-160.67</td>
</tr>
<tr>
<td>C₂H₆O ethanol</td>
<td>liquid</td>
<td>174.72</td>
<td>-160.70</td>
</tr>
<tr>
<td>C₂H₆O₂ ethane-1,2-diol, ethylene glycol</td>
<td>liquid</td>
<td>-327.07</td>
<td>-179.50</td>
</tr>
<tr>
<td>C₃H₄ propadiene</td>
<td>gaseous</td>
<td>202.38</td>
<td>-234.90</td>
</tr>
<tr>
<td>C₃H₄ propyne</td>
<td>gaseous</td>
<td>194.16</td>
<td>-248.10</td>
</tr>
<tr>
<td>C₃H₆ propene</td>
<td>gaseous</td>
<td>74.66</td>
<td>-226.90</td>
</tr>
<tr>
<td>C₃H₆ cyclopropane</td>
<td>gaseous</td>
<td>104.11</td>
<td>-237.90</td>
</tr>
<tr>
<td>C₃H₆O propanone, acetone</td>
<td>gaseous</td>
<td>-151.82</td>
<td>-294.90</td>
</tr>
<tr>
<td>C₃H₆O propanone, acetone</td>
<td>liquid</td>
<td>-154.83</td>
<td>-200.00</td>
</tr>
<tr>
<td>C₄H₈ propane</td>
<td>gaseous</td>
<td>-23.43</td>
<td>-269.90</td>
</tr>
<tr>
<td>C₄H₆ but-1-ene</td>
<td>gaseous</td>
<td>72.03</td>
<td>-307.40</td>
</tr>
<tr>
<td>C₄H₈ but-2-ene cis</td>
<td>gaseous</td>
<td>67.20</td>
<td>-300.80</td>
</tr>
<tr>
<td>C₄H₈ but-2-ene trans</td>
<td>gaseous</td>
<td>64.16</td>
<td>-296.50</td>
</tr>
<tr>
<td>C₄H₈O₂ ethyl acetate</td>
<td>liquid</td>
<td>-323.19</td>
<td>-259.00</td>
</tr>
<tr>
<td>C₄H₁₀ butane</td>
<td>gaseous</td>
<td>-15.62</td>
<td>-310.00</td>
</tr>
<tr>
<td>Compound</td>
<td>S/L/G</td>
<td>μ (kG)</td>
<td>Temp Coff. G/K</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>----------------</td>
</tr>
<tr>
<td>C₅H₁₀ 2-methylpropane</td>
<td>gaseous</td>
<td>-17.92</td>
<td>-294.60</td>
</tr>
<tr>
<td>C₅H₁₀ cyclopentane</td>
<td>gaseous</td>
<td>38.67</td>
<td>-292.90</td>
</tr>
<tr>
<td>C₅H₁₀ cyclopentane</td>
<td>liquid</td>
<td>36.49</td>
<td>-204.10</td>
</tr>
<tr>
<td>C₅H₁₂ pentane</td>
<td>gaseous</td>
<td>-8.11</td>
<td>-348.40</td>
</tr>
<tr>
<td>C₅H₁₂ pentane</td>
<td>liquid</td>
<td>-9.21</td>
<td>-262.70</td>
</tr>
<tr>
<td>C₅H₁₂ 2,2-dimethylpropane</td>
<td>gaseous</td>
<td>-15.18</td>
<td>-306.40</td>
</tr>
<tr>
<td>C₆H₆ benzene</td>
<td>gaseous</td>
<td>129.73</td>
<td>-269.20</td>
</tr>
<tr>
<td>C₆H₆O phenol</td>
<td>solid</td>
<td>-47.50</td>
<td>-142.00</td>
</tr>
<tr>
<td>C₆H₁₀O₅ glycogen</td>
<td>aq. sol.</td>
<td>-662.49</td>
<td></td>
</tr>
<tr>
<td>C₆H₁₂ cyclohexane</td>
<td>gaseous</td>
<td>31.75</td>
<td>-298.20</td>
</tr>
<tr>
<td>C₆H₁₂ cyclohexane</td>
<td>liquid</td>
<td>26.83</td>
<td>-204.10</td>
</tr>
<tr>
<td>C₆H₁₂O₆ fruit sugar, fructose</td>
<td>aq. sol.</td>
<td>-915.59</td>
<td>N/A</td>
</tr>
<tr>
<td>C₆H₁₂O₆ grape sugar, D-glucose</td>
<td>solid</td>
<td>-910.49</td>
<td>N/A</td>
</tr>
<tr>
<td>C₆H₁₂O₆ grape sugar, D-glucose</td>
<td>aq. sol.</td>
<td>-917.44</td>
<td>N/A</td>
</tr>
<tr>
<td>C₆H₁₄ hexane</td>
<td>gaseous</td>
<td>0.30</td>
<td>-386.80</td>
</tr>
<tr>
<td>C₆H₁₄ hexane</td>
<td>liquid</td>
<td>-4.26</td>
<td>-296.00</td>
</tr>
<tr>
<td>C₇H₈ methylbenzene, toluene</td>
<td>gaseous</td>
<td>122.39</td>
<td>-319.70</td>
</tr>
<tr>
<td>C₇H₈ methylbenzene, toluene</td>
<td>liquid</td>
<td>110.61</td>
<td>-219.00</td>
</tr>
<tr>
<td>C₈H₁₈ octane</td>
<td>gaseous</td>
<td>17.44</td>
<td>-463.70</td>
</tr>
<tr>
<td>C₈H₁₈ octane</td>
<td>liquid</td>
<td>6.41</td>
<td>-361.20</td>
</tr>
<tr>
<td>C₈H₁₈ 2,2,4-trimethylpentane</td>
<td>gaseous</td>
<td>13.09</td>
<td>-425.20</td>
</tr>
<tr>
<td>C₈H₁₈ 2,2,4-trimethylpentane</td>
<td>liquid</td>
<td>6.32</td>
<td>-330.00</td>
</tr>
<tr>
<td>C₁₂H₂₂O₁₁ cane sugar</td>
<td>solid</td>
<td>-1543.52</td>
<td>-360.00</td>
</tr>
<tr>
<td>C₁₂H₂₂O₁₁ cane sugar</td>
<td>aq. sol.</td>
<td>-1550.63</td>
<td>N/A</td>
</tr>
<tr>
<td>C₁₂H₂₂O₁₁ cane sugar</td>
<td>aq. sol.</td>
<td>-1552.22</td>
<td>N/A</td>
</tr>
</tbody>
</table>
APPENDIX D

D.1 Reactions Generated Example

**Table D1:** Compounds generated from the compound stoichiometry LINGO program, only the first two of 280 reactions are shown from Excel:

<table>
<thead>
<tr>
<th>Compound</th>
<th>All Rxns</th>
<th>Rxn1</th>
<th>Rxn2</th>
</tr>
</thead>
<tbody>
<tr>
<td>carbon monoxide</td>
<td>v1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>hydrogen</td>
<td>v2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>formic acid</td>
<td>v3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>acetic acid</td>
<td>v4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>propanoic acid</td>
<td>v5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>acetone</td>
<td>v6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>formaldehyde</td>
<td>v7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>acetaldehyde</td>
<td>v8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>oxaldehyde</td>
<td>v9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>glycol aldehyde</td>
<td>v10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>acetal</td>
<td>v11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cellulose acetate</td>
<td>v12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cellobiose</td>
<td>v13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>fructose</td>
<td>v14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>glucose</td>
<td>v15</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>DME(dimethyl ether)</td>
<td>v16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HMF (5-Hydroxymethyl-2-furaldehyde)</td>
<td>v17</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>furan-2-carbaldehyde</td>
<td>v18</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>phenol</td>
<td>v19</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>isoeugenol</td>
<td>v20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>eugenol</td>
<td>v21</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cresol</td>
<td>v22</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>xylose</td>
<td>v23</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>furfural</td>
<td>v24</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>furfuryl alcohol</td>
<td>v25</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>methyl furan</td>
<td>v26</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>tetrahydrofurfural</td>
<td>v27</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>levulinic acid</td>
<td>v28</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>angelica lactone</td>
<td>v29</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Compound</td>
<td>All Rxns</td>
<td>Rxn1</td>
<td>Rxn2</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>gamma valerolactone</td>
<td>v30</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1,4-pentanediol</td>
<td>v31</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1,2-dihydroxypentanal</td>
<td>v32</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1-pentanol</td>
<td>v33</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>methanol</td>
<td>v34</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>methylene</td>
<td>v35</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>ethane</td>
<td>v36</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>propane</td>
<td>v37</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n-butane</td>
<td>v38</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n-octane</td>
<td>v39</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ethene</td>
<td>v40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>propene</td>
<td>v41</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>but-1-ene</td>
<td>v42</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>but-2-ene</td>
<td>v43</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>benzene</td>
<td>v44</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>toluene</td>
<td>v45</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>p-xylene</td>
<td>v46</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>m-xylene</td>
<td>v47</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>o-xylene</td>
<td>v48</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>o-cresol</td>
<td>v49</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>p-cresol</td>
<td>v50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>m-cresol</td>
<td>v51</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>indane</td>
<td>v52</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>tetralin</td>
<td>v53</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n-nonane</td>
<td>v54</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n-decane</td>
<td>v55</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n-undecane</td>
<td>v56</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n-dodecane</td>
<td>v57</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n-tridecane</td>
<td>v58</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2-methyltetrahydrofuran</td>
<td>v59</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ethanol</td>
<td>v60</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>n-pentane</td>
<td>v61</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n-hexane</td>
<td>v62</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>carbon dioxide</td>
<td>v63</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>water</td>
<td>v64</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>glycerol</td>
<td>v65</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oxygen</td>
<td>v66</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Wood</td>
<td>v67</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Compound</td>
<td>All Rxns</td>
<td>Rxn1</td>
<td>Rxn2</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>----------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>BioOil</td>
<td>v68</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Char</td>
<td>v69</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pyrolysis Gas</td>
<td>v70</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cellulose</td>
<td>v71</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>hemicellulose</td>
<td>v72</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>lignin</td>
<td>v73</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>biomass(woody)</td>
<td>v74</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>coumaryl alcohol</td>
<td>v75</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>coniferyl alcohol</td>
<td>v76</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>sinapyl alcohol</td>
<td>v77</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ethyl 4-oxo-5-phenylpentanoate</td>
<td>v78</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1-Hexene</td>
<td>v79</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Propan-1-ol</td>
<td>v80</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n-Butanol</td>
<td>v81</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CycloHexane</td>
<td>v82</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cyclohexene</td>
<td>v83</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Parrafin</td>
<td>v84</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MTBE (tert-Butyl methyl ether )</td>
<td>v85</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Delta G Rxn</th>
<th>Delta H Rxn</th>
<th>Delta $ Rxn</th>
<th>Ox. State Reac</th>
<th>Ox State Prod.</th>
<th>G Reactants</th>
<th>G Products</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-259.07</td>
<td>-213.62</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>-877.39</td>
<td>-1136.46</td>
</tr>
<tr>
<td></td>
<td>-457.1</td>
<td>-276.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-877.3</td>
<td>-1334.5</td>
</tr>
</tbody>
</table>

**Table D2:** Corresponding first two reactions shown in redirected format in Excel:

<table>
<thead>
<tr>
<th>Reactions</th>
<th>Delta G Rxn</th>
<th>Delta H Rxn</th>
<th>$Economics (+ = Profit)</th>
<th>Reaction Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1<em>glucose-------&gt; 2</em>ethanol + 2*carbon dioxide</td>
<td>-259.07</td>
<td>-213.62</td>
<td>5</td>
<td>Rxn 1</td>
</tr>
<tr>
<td>1<em>glucose-------&gt; 3</em>methane + 3*carbon dioxide</td>
<td>-457.1</td>
<td>-276.7</td>
<td>0</td>
<td>Rxn 2</td>
</tr>
</tbody>
</table>
VITA

Name: Eric James Pennaz

Address: Artie McFerrin Department of Chemical Engineering
Texas A&M University
Jack E. Brown Engineering Bldg.
3122 TAMU
College Station, TX 77843-3122

Email Address: eric.pennaz@tamu.edu

Education: M.S., Chemical Engineering, Texas A&M University, 2011
M.A., Education, City College of the City of New York, 2009
B.S., Chemistry, University of Chicago, 2006