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-mthyltetrahydrofuran (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

Page

LIST OF FIGURES

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-methyltetrahydrofuan (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.

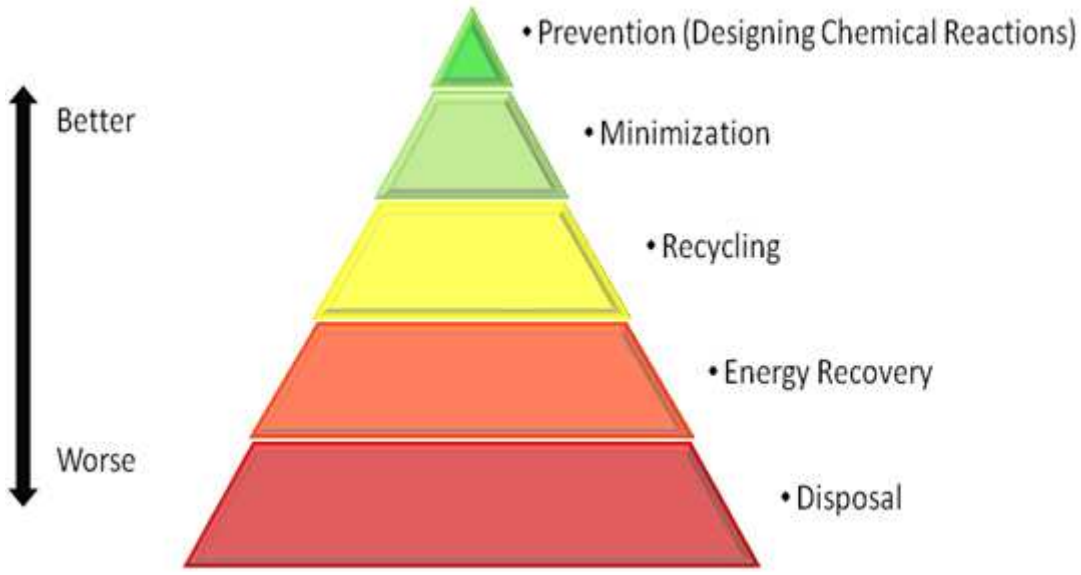

Figure 1.1: The motivation pyramid. The pyramid shows how chemical reactions have the ability to have the largest impact on chemical process design, that of prevention. The arrow on the left shows the relative value of each process.

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.

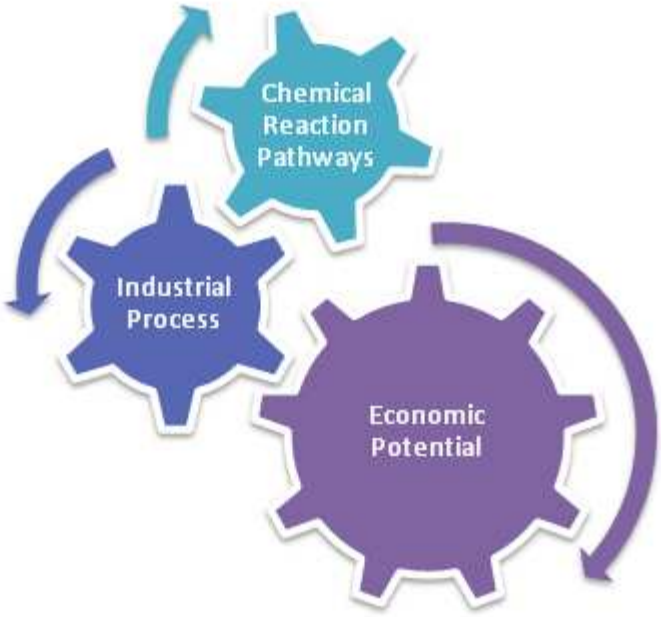

Figure 1.2: An interlocking of the processes. The chemical reaction pathways influence the industrial process and the economic potential both ways.

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.

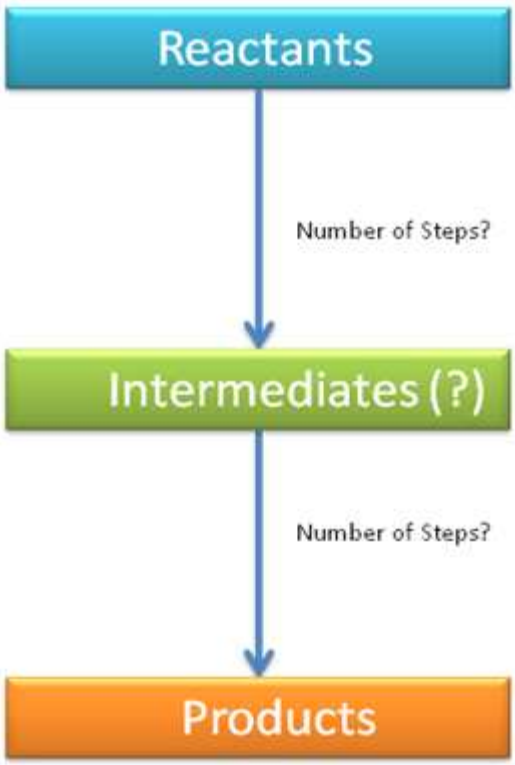

Figure 3.1: The reaction pathways. The reactants and products of a process are known. The question is what intermediates to use and how many steps there are between each of the reactant to intermediates and intermediates to products.

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.

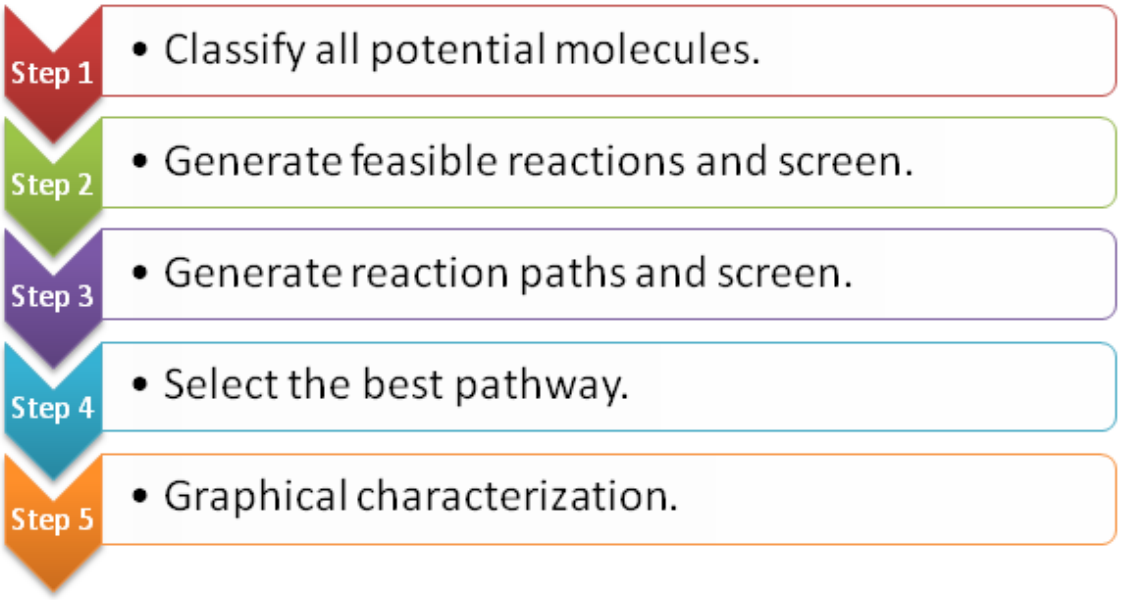

Figure 4.1: The stepwise method. The method involves these five steps in order to generate the best and most feasible chemical reaction pathways in a biorefinery.

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.

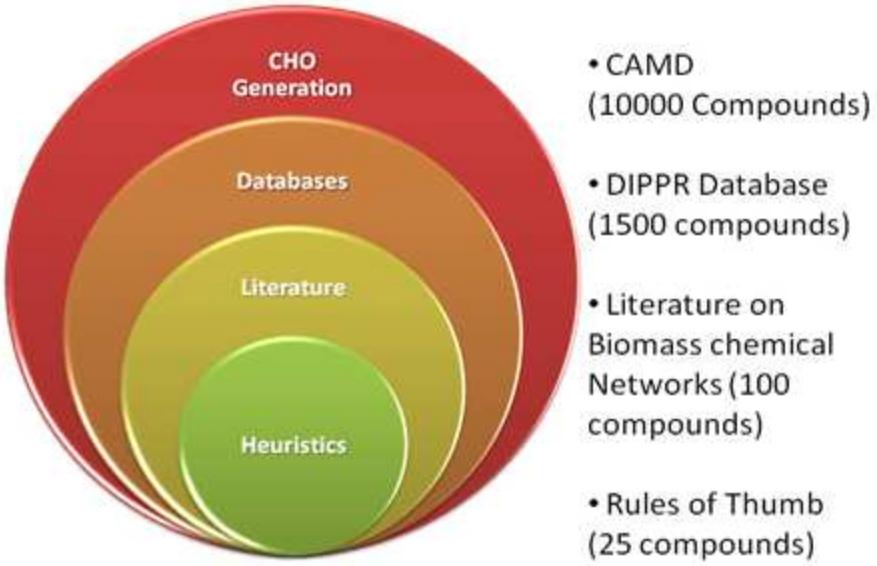

Figure 4.2: The compound hierarchy. The number of compounds that can be found in the Carbon, Hydrogen, and Oxygen search space varies with the method you choose, from 10000 plus compounds for a certain molecular generation approach, to approximately 25 compounds for rules of thumb.

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:

$$\text{Min } i_s + ii_s, \qquad \forall s \in S \tag{1}$$

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:

$$\varepsilon_s V_s = 0, \qquad \forall s \in S \tag{2}$$

The atom balance is the starting point in any chemical reaction. ε_s is the E*S atomic matrix and V_s is the S*1 column vector of stoichiometric coefficients ν_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:

$$v_p \geq 1, \tag{3}$$

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:

$$x_s \tag{4}$$

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

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

$$v_s = x_s - 2 * x_s * ii_s, \qquad \forall\ s \in S$$

This equation can be recast to avoid the bilinear term of $x_s * ii_s$, using a transformation of the dummy variables:

$$v_s = x_s - 2 * y_s, \qquad \forall\ s \in S \tag{5}$$

$$y_s - v_{max} * ii_s \leq 0, \qquad \forall\ s \in S \tag{6}$$

$$x_s + v_{max} * (ii_s - 1) - y_s \leq 0, \qquad \forall\ s \in S \tag{7}$$

$$y_s - x_s \leq 0, \qquad \forall\ s \in S \tag{8}$$

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

$$y_s \geq ii_s, \qquad \forall\ s \in S \tag{9}$$

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:

$$x_s - y_s - v_{max} * I_s \leq 0, \qquad \forall s \in S \qquad (10)$$

$$x_s - y_s - (v_{max} + 1) * I_s + v_{max} \leq 0, \qquad \forall s \in S \qquad (11)$$

where v_{max} is set to 50. Equations allow for the dummy variables to be smaller than some v_{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:

$$i_s + ii_s + iii_s = 1, \qquad \forall s \in S \qquad (12)$$

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 H_20 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 H_20 on the reactants than products side, then H_20 is taken as a reactant and vice versa.

Constraints limiting total number of reactants and products:

$$\sum_s \; i_s \leq N_p^{max}, \qquad \forall s \in S \qquad (13)$$

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^{max}.

$$\sum_s \; ii_s \leq N_r^{max}, \qquad \forall s \in S \qquad (14)$$

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^{max}.

$$\sum_s (i_s + ii_s) \leq N_{spe}^{max}, \qquad \forall s \in S \tag{15}$$

$$v_s \leq v_s^{max}, \qquad \forall s \in S \tag{16}$$

These equations are presented so that the number of reactants and products in a given reaction are set to some maximum. N_r^{max} and N_p^{max}, 3 as stated before, while N_{spe}^{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^{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):

$$ii_a + ii_b \leq 1 \tag{17}$$

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:

$$\Delta H_f^o \text{ (J/mol)} \tag{18}$$

Entropy of Formation:

$$\Delta S_f^o \text{ (J/mol*K)} \tag{19}$$

Gibbs free energy of Formation, either found explicitly through the DIPPR database, Gani method, or enumerated through the other parameters with the following equation:

$$\Delta G_f^o = \Delta H_f^o - T_{oper} * \Delta S_f^o \text{ (J/mol)} \tag{20}$$

The operating temperature is thus restricted:

$$300 \leq T_{oper} \leq 1000 \tag{21}$$

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:

$$-\log(LC50_s), \qquad \forall\, s \in S \tag{22}$$

These factors are all tabulated and screened after the reactions are generated, then the economics are also included in the following equation:

$$\sum_s iii_s * (i_s - ii_s) * C_s * v_s, \qquad \forall\, s \in S \tag{23}$$

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.

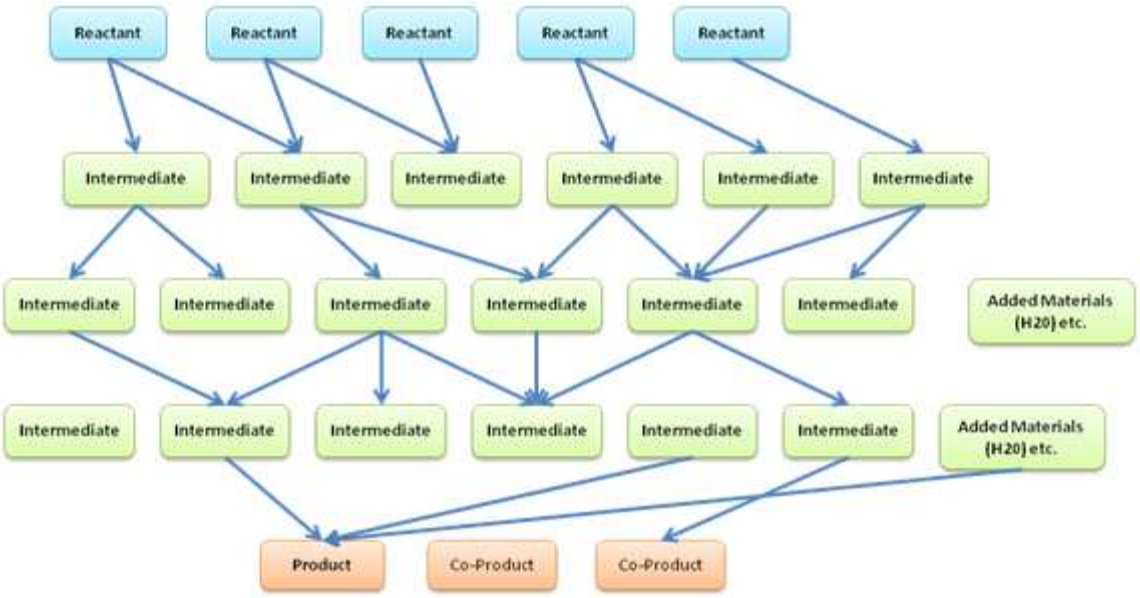

Figure 4.4: Superstructure of compounds with reactions. The reactants, intermediates, and products are now arranged into a chemical reaction tree after the reactions are generated and the reactions are screened based on the given criteria.

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:

$$S_s * v_j = 0 \qquad \forall\, s \in S, \forall\, j \in J \tag{24}$$

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:

$$v_j^L \leq v_j \leq v_j^U, \qquad \forall\, j \in J \tag{25}$$

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:

$$\Delta H_f^o \text{ (J/mol)} \tag{25a}$$

Entropy of formation:

$$\Delta S_f^o \text{ (J/mol*K)} \tag{25b}$$

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:

$$\Delta G_f^o = \Delta H_f^o - T_{oper} * \Delta S_f^o \text{ (J/mol)} \qquad (25c)$$

The operating temperature is again restricted:

$$300 \leq T_{oper\,s} \leq 1000, \qquad \forall\, s \in S \qquad (25d)$$

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

$$-\log(LC50_s), \qquad \forall\, s \in S \qquad (25e)$$

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):

$$\sum_s iii_s * (i_s - ii_s) * C_s * v_s, \qquad \forall\, s \in S \qquad (25f)$$

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:

$$Flow_{j1} - Flow_{j2} = 0 \qquad \forall\, j \in J \qquad (25g)$$

$$Flow_{j4} - Flow_{j5} - Flow_{j6} = 0 \qquad \forall\, j \in J\,.$$

Flux Balance Analysis 1st iteration LP problem:

$$\max f(v) \tag{26}$$

$$f: R^n \rightarrow R$$

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:

$$w_j^0 = 1 \leftrightarrow v_j = 0, \qquad \forall\, j \in J \tag{27}$$

Every flux variable is changed into three different variables:

$$v_j = v_j^0 + v_j^- + v_j^+, \qquad \forall\, j \in J \tag{28}$$

For each flux, three binary variables are introduced to recast the LP into an MILP:

$$w_j^0 + w_j^- + w_j^+ = 1, \qquad \forall\, j \in J \tag{29}$$

Coupling the w and v variables:

$$v_j^L(1 - w_j^0) \leq v_j \leq v_j^U\ (1 - w_j^0)\ , \qquad \forall\ j \in J \qquad (30)$$

$$-\varepsilon w_j^0 < v_j^0 < \varepsilon w_j^0, \qquad \forall\ j \in J \qquad (31)$$

$$v_j^L w_j^- \leq v_j^- \leq \varepsilon w_j^-, \qquad \forall\ j \in J \qquad (32)$$

$$\varepsilon w_j^+ \leq v_j^+ \leq w_j + v_j^U, \qquad \forall\ j \in J \qquad (33)$$

$$\varepsilon = 1*10^{-6}.$$

The new formulation the flux balance analysis for the 2nd and later iterations as an MILP is then written as follows:

$$\max f(v) \qquad (34)$$

$$f: R^n \rightarrow R$$

s.t. (24-25), (27-33).

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:

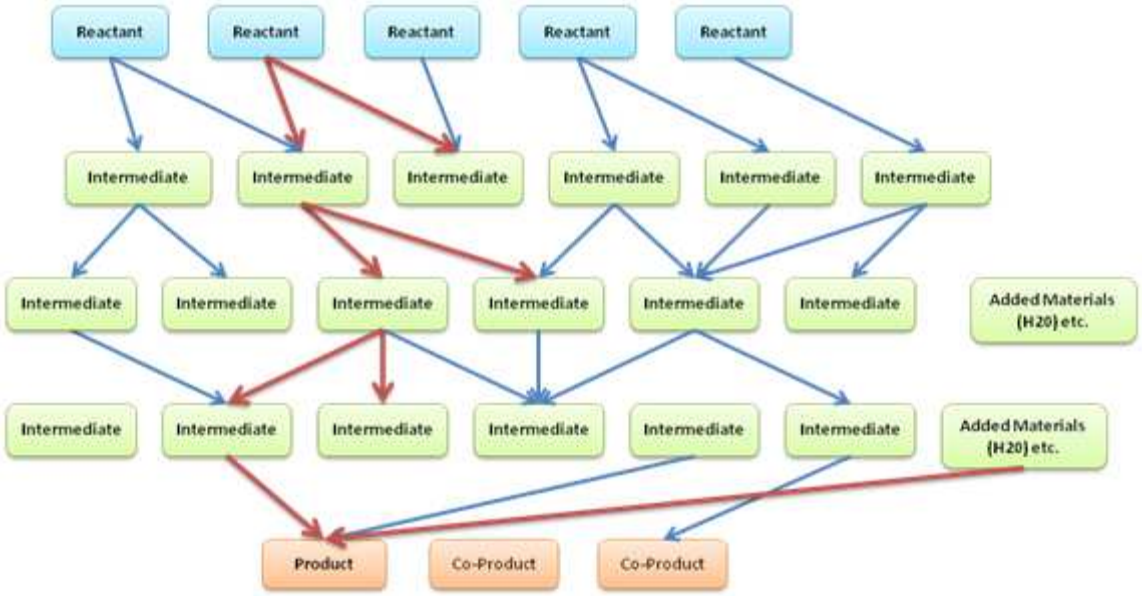

Figure 4.5: Superstructure of compounds with the best reaction. The best pathway is now shown in red after the chemical reaction pathways have been compared based on the criteria given.

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):

$$\mathrm{d}E = T\,\mathrm{d}S - p\,\mathrm{d}V + \mu\,\mathrm{d}n + \phi\,\mathrm{d}Q + \mathbf{v}\,\mathrm{d}\mathbf{p} + \psi\,\mathrm{d}m + \cdots \tag{35}$$

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:

$$(\partial \mathrm{E} / \partial \mathrm{n})_{\mathrm{S,V\ldots}} = \mu_{\mathrm{i}} \tag{36}$$

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:

$$A \leftrightarrow B \tag{37}$$

Then show how chemical potential changes can be used to predict the direction of the transformation or chemical equilibrium.

$$\text{When:} \quad \mu_A > \mu_B \tag{38}$$

compound A will turn into compound B, or a compound will flow from A to B.

$$\text{When:} \quad \mu_A = \mu_B \tag{39}$$

compound A is in equilibrium with compound B, or there will be equilibrium between location A and B.

$$\text{When:} \quad \mu_A < \mu_B \tag{40}$$

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:

$$dG = Vdp - SdT + \sum \mu_i \, dN_i - \ldots \tag{41}$$

where μ_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:

$$\mu = \mu^o + RT \ln (p/p_o) \tag{42}$$

where μ^o is the standard sate 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 sate 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:

$$8\ C_6H_{12}O_6 + 6\ H_2 \rightarrow 6\ C_8H_{18} + 24\ O_2 \quad (43)$$

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.

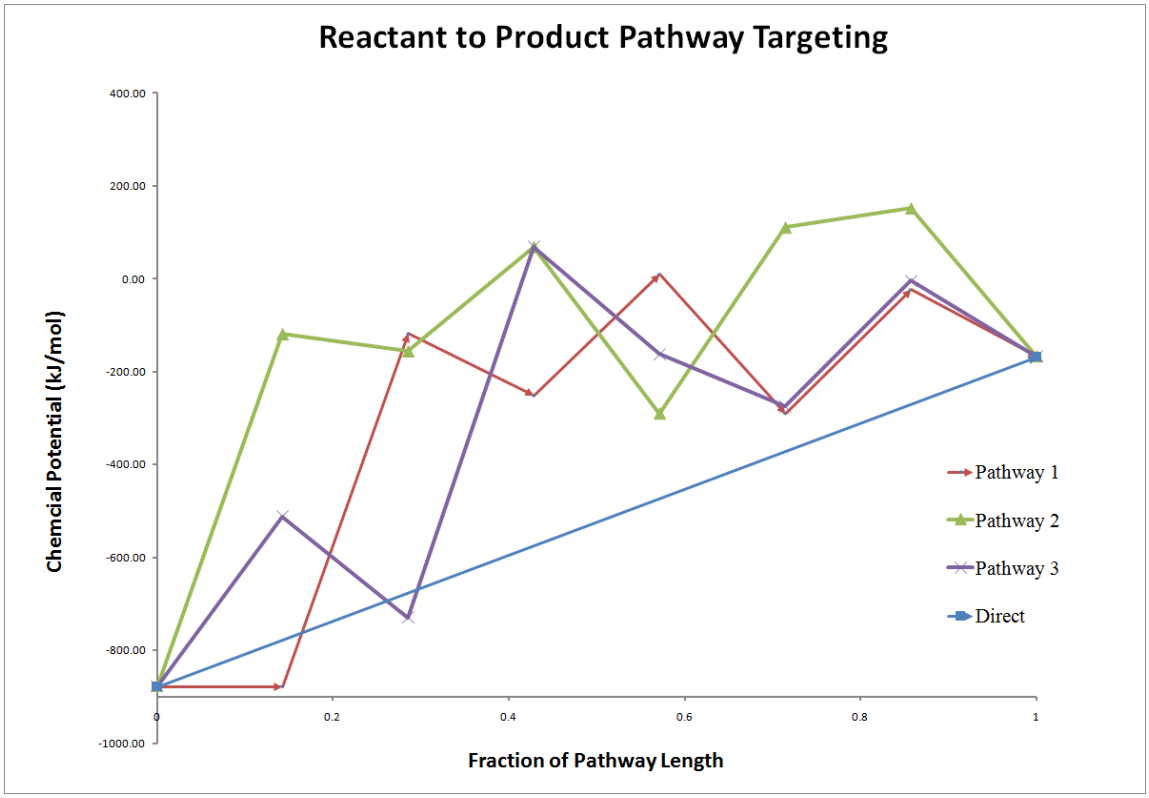

Figure 4.6: Example pathway targeting and reactions. The targeting pathway is in blue, while the three enumerated pathways based on chemical potential are show in red, green, and purple respectively. Each node corresponds to a compound intermediate.

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 preformed, 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_2H_6, CH_4, CH_2O, H_20, H_2, CO, CO_2 were included as reacting intermediates based on current chemical conversion methods. Liquid fuel products such as 2-methyltetrohydrofuran (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 H_2) 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.

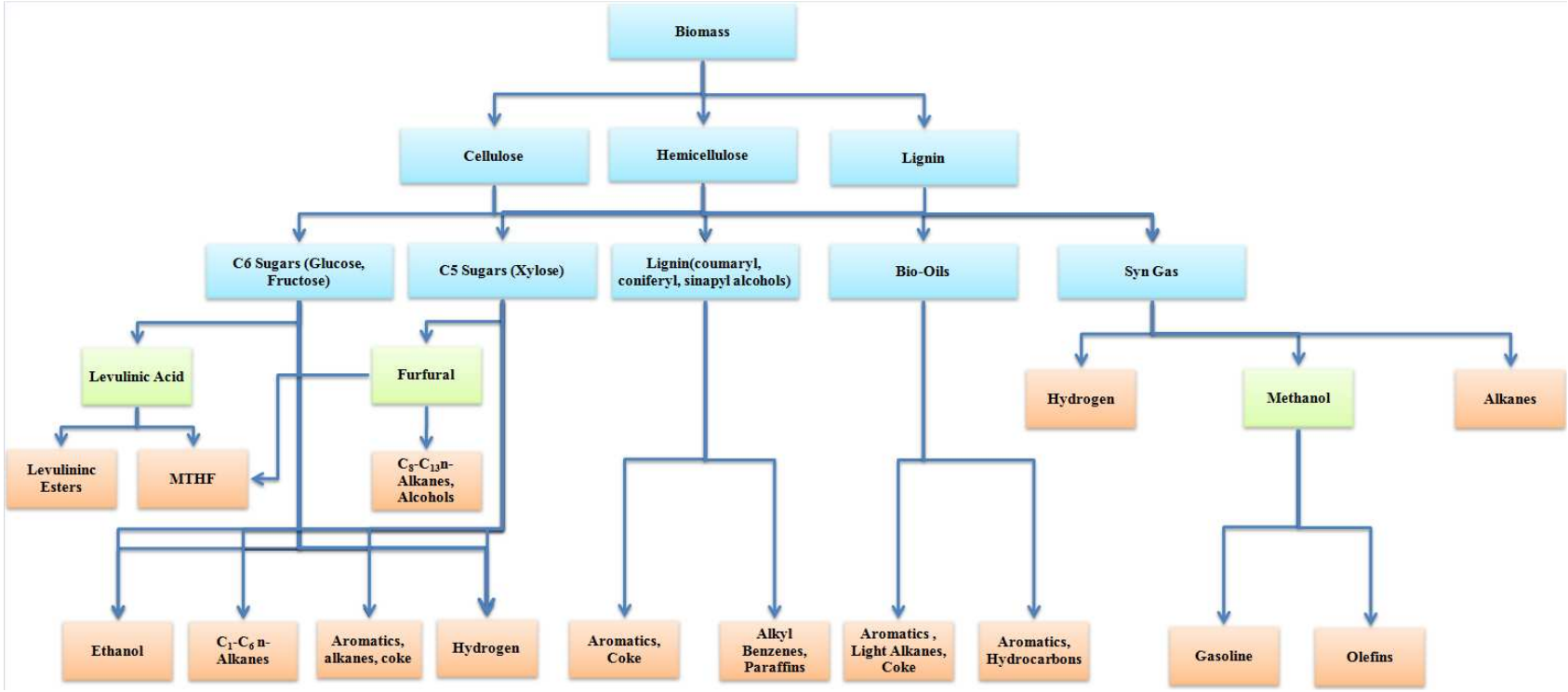

Figure 5.1: Reaction Synthesis Tree, based on Huber. The chemical reaction pathways have been shown to follow this basic shape. Additional reactions are created through stoichiometric enumeration. Not all chemical intermediates are shown.

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 ($C_6H_{12}O_6$) or xylose ($C_5H_{10}O_5$) to the alternative fuel MTHF, also known as 2-Methyltetrahydrofuran ($C_5H_{10}O$). 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.

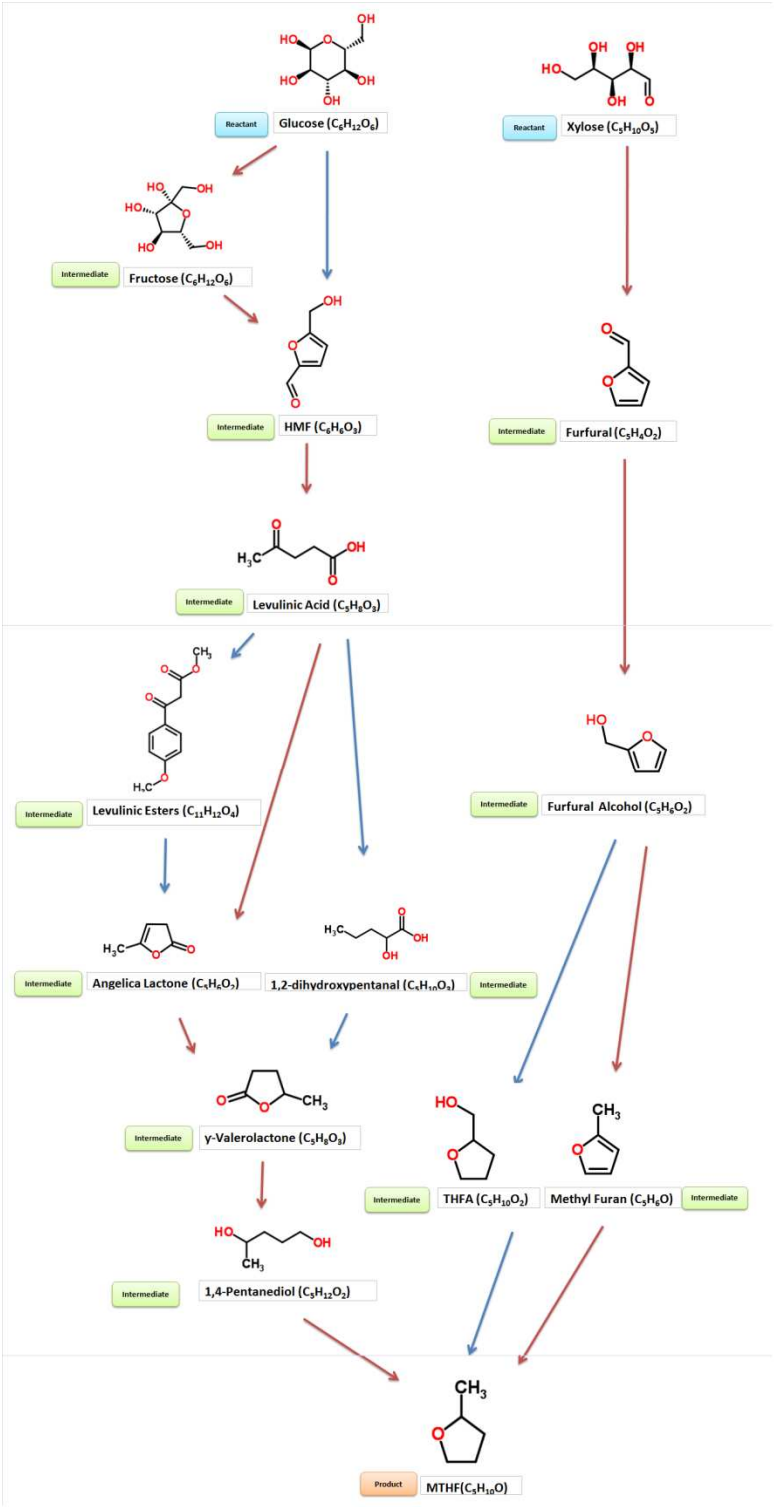

Figure 5.2: The enumerated pathways from glucose and xylose to MTHF.

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:

$$C_6H_{12}O_6 + H_2 \rightarrow C_5H_{10}O + O_2 \quad (44).$$

Balancing the equation gives:

$$5\ C_6H_{12}O_6 \rightarrow 6\ C_5H_{10}O + 12\ O_2 \quad (45).$$

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

$$C_5H_{10}O_5 + H_2 \rightarrow C_5H_{10}O + O_2 \quad (46).$$

Balancing the equation gives:

$$C_5H_{10}O_5 \rightarrow C_5H_{10}O + 2O_2 \quad (47).$$

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.

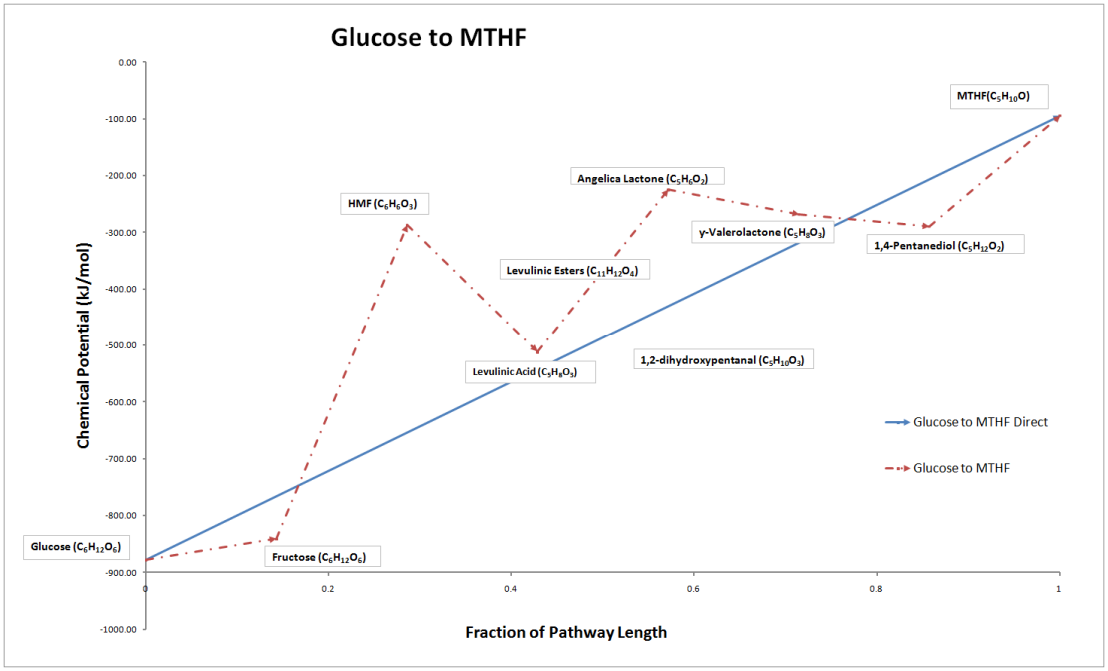

Figure 5.3: The targeting of optimal pathways. This illustrates the differences between different intermediates between glucose and MTHF. The relative changes are differences in chemical potential.

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.

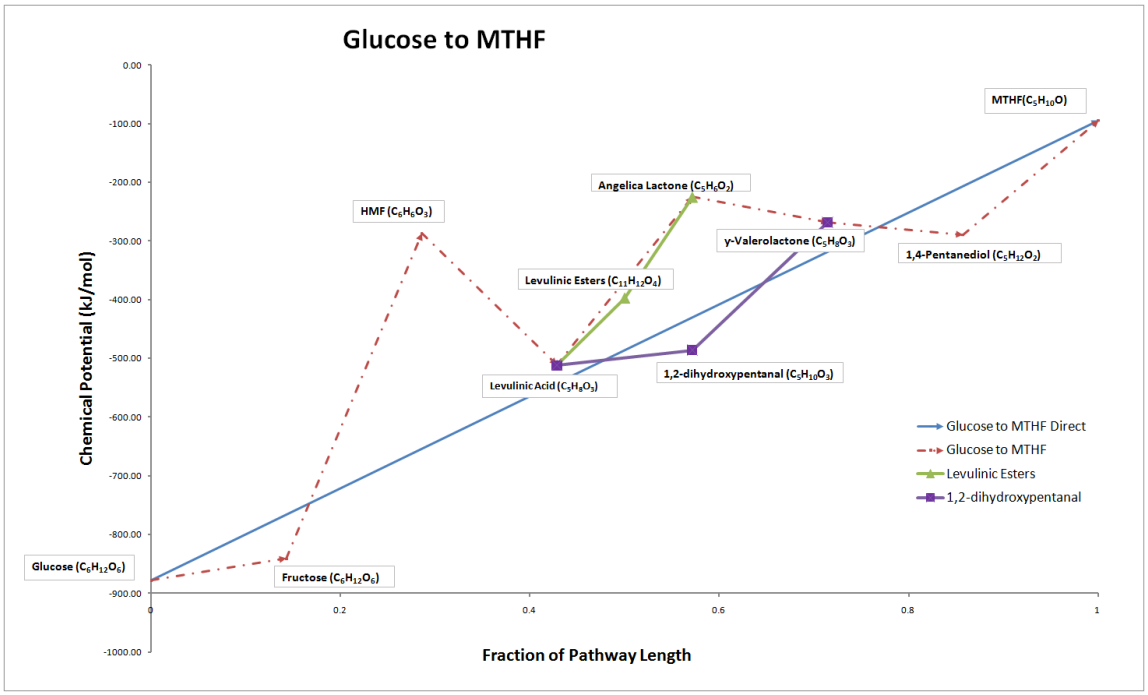

Figure 5.4: The targeting of two alternate pathways. This illustrates the differences between different intermediates between glucose and MTHF. The relative changes are differences in chemical potential.

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

Agnihotri RB, Motard RL (1980) Reaction path synthesis in industrial chemistry. Comp App Chem Eng, American Chemical Society. 124: 193-206

Beard DA, Liang S, Qian H (2002) Energy balance for analysis of complex metabolic networks. Biophysical Journal 83(1): 79-86

Besler A, Harwardt A, Marquardt W (2009) Reaction networks - a rapid screening method. Computer Aided Chemical Engineering 26: 243-248

Buxton A, Hugo A, Livingston AG, Pistikopoulos EN (2002a) Chapter 7: Identification of multistep reaction stoichiometries: CAMD problem formulation. Comp Aid Chem Eng. 12: 167-209

Buxton A, Hugo A, Livingston AG, Pistikopoulos EN (2002b) Chapter 14: Case study in identification of multistep reaction stoichiometries. Comp Aid Chem Eng. 12: 319-327

Buxton A, Livingston AG, Pistikopoulos EN (1997) Reaction path synthesis for environmental impact minimization. Comp Chem Eng 21(Supplement 1): S959-S964

Callen HB, Bridgman P (1960) Thermodynamics, an introduction to the physical theories of equilibrium thermostatics and irreversible thermodynamics. New York, Wiley.

Clausen CA, Mattson G (1978) Principles of industrial chemistry. The University of Michigan, Wiley.

Crabtree E, El-Halwagi M (1994) Synthesis of environmentally acceptable reactions. New York, AIChE

Falk G, Herrmann F, Schmid GB (1983) Energy forms or energy carriers. Am. J. Phys 51(12): 1074-1077

Fernando S, Adhikari S, Chandrapal C, Murali N (2006) Biorefineries: current status, challenges, and future direction. Energy & Fuels 20(4): 1727-1737

Finley SD, Broadbelt LJ, Hatzimanikatis V (2009) Computational framework for predictive biodegradation. Biotech Bioeng 104(6): 1086-1097

Fornari T, Stephanopoulos G (1994a) Synthesis of chemical reaction paths: The scope of group contribution methods. Chem Eng Comm 129: 135-157

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

Fornari T, Rotstein E, Stephanopoulos G (1989) Studies on the synthesis of chemical reaction paths--II: reaction schemes with two degrees of freedom. Chem Eng Sci 44(7): 1569-1579

Hatzimanikatis V, Li C, Ionita JA, Broadbelt LJ (2004) Metabolic networks: enzyme function and metabolite structure. Cur Op Struc Bio 14(3): 300-306

Hatzimanikatis V, Li C, Ionita JA, Henry CS, Jankowski MD, et al (2005) Exploring the diversity of complex metabolic networks. Bioinformatics 21(8): 1603-1609

Hechinger M, Voll , Marquardt W (2010) Towards an integrated design of biofuels and their production pathways. Comp Chem Eng 34(12): 1909-1918

Herrmann F, Job G (1996) The historical burden on scientific knowledge. European J Phys 17: 159-163

Huber GW, Iborra S, Corma A (2006) Synthesis of transportation fuels from biomass: chemistry, catalysts, and engineering. Chemical Reviews 106(9): 4044-4098

Job G, Herrmann F (2006) Chemical potential—a quantity in search of recognition. European J Phys 27: 353-370

Johnson AP, Marshall C (1992) Starting material oriented retrosynthetic analysis in the LHASA program: 1. General description. J Chem Inform Comp Sci 32(5): 411-417

Kohse-Höinghaus KP, Oßwald P, Cool TA, Kasper T, Hansen N, et al (2010) Biofuel combustion chemistry: from ethanol to biodiesel. Angewandte Chemie Inter Ed 49(21): 3572-3597

Lee DY, Yun H, Park S, Lee SY (2003) MetaFluxNet: the management of metabolic reaction information and quantitative metabolic flux analysis. Bioinformatics 19(16): 214-2146

Lee S, Phalakornkule C, Domach MM, Grossmann IE (2000) Recursive MILP model for finding all the alternate optima in LP models for metabolic networks. Comp Chem Eng 24(2-7): 711-716

Li M, Hu S, Li Y, Shen JZ (2000) A hierarchical optimization method for reaction path synthesis. Ind Eng Chem Res 39(11): 4315-4319

Majumdar S, Mitra K (2004) Modeling of a reaction network and its optimization by genetic algorithm. Chem Eng J 100(1-3): 109-118

Marrero J, Gani R (2001) Group-contribution based estimation of pure component properties. Fluid Phase Equilibria 183: 183-208

May D, Rudd DF (1976) Development of solvay clusters of chemical reactions. Chem Eng Sci 31(1): 59-69

Murabito E, Simeonidis E, Smallbone K, Swinton J (2009) Capturing the essence of a metabolic network: a flux balance analysis approach. J Theor Bio 260(3): 445-452

Ng, D, Pham V, El-Halwagi M, Jimenez-Gutierrez A, Spriggs HD (2009) A hierarchical approach to the synthesis and analysis of integrated biorefineries. Design for Energy and the Environment, CRC Press

Nishida N, Stephanopoulos G, Westerberg A (1981) A review of process synthesis. AIChE Journal 27(3): 321-351

Pokoo- Aikins G, Heath A, Mentzer RA, Mannan MS, Rogers WJ, et al (2010) A multi-criteria approach to screening alternatives for converting sewage sludge to biodiesel. J Loss Preven Pro Ind 23(3): 412-420

Rotstein E, Resasco D, Stephanopoulos G (1982) Studies on the synthesis of chemical reaction paths--I : Reaction characteristics in the ([Delta]G, T) space and a primitive synthesis procedure. Chem Eng Sci 37(9): 1337-1352

Rüffler R, Job G (2009) Chemical potential from the beginning. Chem Ed ICT Age, Netherlands, Springer

Sammons Jr NE, Yuan W, Eden MR, Aksoy B, Cullinan HT (2008) Optimal biorefinery product allocation by combining process and economic modeling. Chemical Engineering Research and Design 86(7): 800-808

Sauer, U (2006) Metabolic networks in motion: 13C-based flux analysis. Mol Sys Bio 2: 62

Shafiee S, Topal E (2009) When will fossil fuel reserves be diminished? Eng Pol 37(1): 181-189

Siirola JJ, Rudd DF (1971) Computer-aided synthesis of chemical process designs: from reaction path data to the process task network. Ind Eng Chem Fund 10(3): 353-362

Stefanis SK, Buxton A, Livingston AG, Pistikopoulos EN (1996) A methodology for environmental impact minimization: solvent design and reaction path synthesis issues. Comp Chem Eng 20(Supplement 2): S1419-S1424

Stelling J, Klamt S, Bettenbrock K, Schuster S, Gilles ED (2002) Metabolic network structure determines key aspects of functionality and regulation. Nature 420(6912): 190-193

Szmant HH (1989) Organic building blocks of the chemical industry. Canada, Wiley-Interscience

Ugi I, Gillespie P (1971) Representation of chemical systems and interconversions by be matrices and their transformation properties. Angewandte Chem Intern Ed in English 10(12): 914-915

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 fun 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 =
Cellobiose, A14 = Fructose,  A15 = Glucose, A16 = DME(dimethyl ether),A17 = HMF,
```

A18 = Furfural, A19 = Phenol, A20 = Isoeugenol, A21 = Eugenol, A22 = Methyl_Guaiacol, A23 = Xylose, A24 = Furfural, A25 = Furfuryl_Alcohol, A26 = Methyl_Furan, A27 = Tetrahydrofurfuryl, A28 = Levulinic_Acid, A29 = Angelica_Lactone, A30 = Gamma_valerolactone, A31 = 1,4-pentanediol, A32 = 1,2-dihydroxypentanal, A33 = 1-pentanol, A34 = Methanol, A35 = Methane, A36 = Ethane, A37 = Propane, A38 = n-Butane, A39 = n-Octane, A40 = Ethene,A41 = Propene, A42 = But-1-ene, A43 = But-2-ene, A44 = Benzene, A45 = Toluene, A46 = p-Xylene, A47 = m-Xylene, A48 = o-Xylene, A49 = o-Cresol, A50 = p-Cresol, A51= m-Cresol, A52 = Indane, A53 = Tetralin, A54 = n-Nonane, A55 = n-Decane,A56 = u-Decane, A57 =n-Dodecane, A58 = n-Tridecane, A59 = 2-MTHF, A60 = Ethanol, A61 = n-Pentane, A62 = n-Hexane, A63 = Carbn_Dioxide, A64 = Water, A65 = Glycerol A66 = Oxygen, A67 = Wood, A68 = Bio_Oil, A69 = Char, A70 = Pyrolysis_Gas A71 = Cellulose , A72 = Hemicellulose, A73 = Lignin, A74 = Biomass, A75 = Coumaryl Alcohol A76 = Coniferyl_Alcohol, A77 = Sinapyl_Alochol, A78 = Levulinic_Ester, A79 = 1-Hexene, A80 = Propanol, A81 = Butanol, A82 = Cyclohexane, A83 = Cyclohexene, A84 = Parrafin, A85 = MTBE;

! SPECIES A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12 A13 A14 A15 A16 A17 A18 A19 A20 A21 A22 A23 A24 A25 A26 A27 A28 A29 A30 A31 A32 A33 A34 A35 A36 A37 A38 A39 A40 A41 A42 A43 A44 A45 A46 A47 A48 A49 A50 A51 A52 A53 A54 A55 A56 A57 A58 A59 A60 A61 A62 A63 A64 A65 A66 A67 A68 A69 A70 A71 A72 A73 A74 A75 A76 A77 A78 A79 A80 A81 A82 A83 A84 A85 ;

Carbon = 1 0 1 2 3 3 1 2 2 2 6 76 12 6 6 2 6
5 6 10 10 8 5 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 6 5
10 1 9 10 11 11 6 3 4 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 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] MIN = @SUM ( Product_Flag ( I) : Product ( I)) + @SUM(Reactant_Flag ( I) :
Reactant ( I)) ;

!atom mass balance;

[Carbon_Balance] @SUM (SPECIES (I) : Carbon( I)* Stoic ( I)) = 0;

[Hydrogen_Balance] @SUM (SPECIES (I) : Hydrogen( I)* Stoic ( I)) = 0;

[Oxygen_Balance] @SUM (SPECIES (I) : Oxygen( I)* Stoic ( I)) = 0;

!The following lines takes care of  for stoiciometric constraints;

@FOR (Dummy_VariableX (I) : @GIN (DummyX( I)));

@FOR( Dummy_VariableX (I) : @BND( 0, DummyX(I), 1000));

!Integer Values of Product (21-24);

@FOR (Coefficients ( I) : Stoic( I) =  DummyX( I) - 2 *  DummyY( I)) ;

@FOR (Dummy_VariableY (I) : DummyY( I) - Vmax *  Reactant ( I) <= 0);

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

@FOR (Dummy_VariableY (I) : DummyY( I)  - DummyX( I) <= 0);

!Bounded stoiciometric coefficient;

Vmax = 50;

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

@FOR (Dummy_VariableY (I) : DummyY( I) >= 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 numer of reacting species and product species;

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

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

!Limit on the stoiciometric 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;

ENDSUBMODEL

SUBMODEL Iteration6: !C6Sugars 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;

ENDSUBMODEL

SUBMODEL Iteration7: !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,

parrafins; !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;

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) <= -1; 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 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 (176) + Flow (177) +Flow (178) +Flow (179) +Flow (180) +Flow (181) +

Flow (182) + Flow (183) +Flow (184) +Flow (185) +Flow (186) +Flow (187) +

Flow (188) + Flow (189) +Flow (190) +Flow (191) +Flow (192) +Flow (193) +

Flow (194) + Flow (195) +Flow (196) +Flow (197) +Flow (198) +Flow (199) +

Flow (200) - (Flow(252) + Flow(257)) = 0 ;

!Flow Constraints around Syngas;

Flow (201) + Flow (202) + Flow (203) + Flow (204) + Flow (205) +

Flow (206) + Flow (207) + Flow (208) + Flow (209) + Flow (210) +

Flow (211) + Flow (212) + Flow (213) + Flow (214) + Flow (215) +

Flow (216) + Flow (217) + Flow (218) + Flow (219) + Flow (220) +

Flow (221) + Flow(263) - (Flow(255) + Flow(251)) = 0;

!Flow Constraints around C5 Sugars;

Flow(26) + Flow (27) + Flow (28) + Flow(29) + Flow(30) +

Flow(31) + Flow (32) + Flow (33) + Flow(34) + Flow(35) +

Flow(36) + Flow (37) + Flow (38) + Flow(39) + Flow(40) +

Flow(41) + Flow (42) + Flow (43) + Flow(44) + Flow(45) +

Flow(46) + Flow (47) + Flow (48) + Flow(49) + Flow(50)

- (Flow(254)) = 0 ;

!Flow Constraints around C6 Sugars;

Flow(1) + Flow (2) + Flow (3) + Flow(4) + Flow(5) +

Flow(6) + Flow (7) + Flow (8) + Flow(9) + Flow(10) +

Flow(11) + Flow (12) + Flow (13) + Flow(14) + Flow(15) +

```
Flow(16) + Flow (17) + Flow (18) + Flow(19) + Flow(20) +
Flow(21) + Flow (22) + Flow (23) + Flow(24) + Flow(25) +
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:

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

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

Compound Name	C	H	O	Gf[298K]	Hf[298K]	METHOD
carbon monoxide	1	0	1	-137.15	-110.53	DIPPR
hydrogen	0	2	0	0	0	DIPPR
formic acid	1	2	2	-278.83	-301.3	MG
acetic acid	2	4	2	-389	-484.5	DIPPR
propanoic acid	3	6	2	-384.6	-510.866	DIPPR
acetone	3	6	1	-155.4	-248.1	DIPPR
formaldehyde	1	2	1	-102.6	-108.6	DIPPR
acetaldehyde	2	4	1	-133.3	-166.4	DIPPR
oxaldehyde	2	2	2	-192.2	-244.4	DIPPR
glycol aldehyde	2	4	2	-289	-404.2	DIPPR
acetal	6	14	2	-251.8	-491.41	DIPPR
cellulose acetate	76	114	49	N/A	N/A	MG
cellobiose	12	22	11	-1477.51	-1895	MG
fructose	6	12	6	-840.72	-1118.76	MG
glucose	6	12	6	-877.39	-1127.36	MG
dimethyl ether	2	6	1	-112.8	-184.1	DIPPR
HMF (5-Hydroxymethyl-2-furaldehyde)	6	6	3	-287.7	-368.47	MG
furan-2-carbaldehyde	5	4	2	-102.56	-179.2	MG
phenol	6	6	1	-50.41	-165.1	DIPPR
isoeugenol	10	12	2	-41.82	-193.58	MG
eugenol	10	12	2	-41.82	-193.58	MG
cresol	8	10	2	-33.94	-130.92	MG
xylose	5	10	5	-729.71	-940.54	MG
furfural	5	4	2	-119	-201.6	DIPPR
furfuryl alcohol	5	6	2	-154.1	-276.2	DIPPR
methyl furan	5	6	1	10.06	-93.62	MG
tetrahydrofurfural	5	10	2	-200.17	-311.31	MG
levulinic acid	5	8	3	-512.2	-697.054	DIPPR

Compound Name	C	H	O	Gf[298K]	Hf[298K]	METHOD
angelica lactone	5	6	2	-225.22	-309.64	MG
gamma valerolactone	5	8	2	-268.49	-383.93	MG
1,4-pentanediol	5	12	2	-289.92	-476.88	MG
1,2-dihydroxypentanal	5	12	1	-155.8	-351.9	DIPPR
1-pentanol	5	12	1	-169.1	-366.5	DIPPR
methanol	1	4	1	-166.9	-239.1	DIPPR
methane	1	4	0	-50.49	-74.52	DIPPR
ethane	2	6	0	-31.92	-83.82	DIPPR
propane	3	8	0	-24.39	-104.68	DIPPR
n-butane	4	10	0	-16.7	-125.79	DIPPR
n-octane	8	18	0	6.587	-249.78	DIPPR
ethene	2	4	0	68.44	52.51	DIPPR
propene	3	6	0	62.64	20.23	DIPPR
but-1-ene	4	8	0	70.41	-0.5	DIPPR
but-2-ene	4	8	0	-222.3	-401.6	DIPPR
benzene	6	6	0	124.4	48.95	DIPPR
toluene	7	8	0	113.8	12.01	DIPPR
p-xylene	8	10	0	111.4	-24.35	DIPPR
m-xylene	8	10	0	107.3	-25.36	DIPPR
o-xylene	8	10	0	110.6	-24.35	DIPPR
o-cresol	7	8	1	-55.66	-204.6	DIPPR
p-cresol	7	8	1	-50.901	-199.28	DIPPR
m-cresol	7	8	1	-59.1	-194	DIPPR
indane	9	10	0	152	11.7	DIPPR
tetralin	10	12	0	163.99	39.15	MG
n-nonane	9	20	0	12.647	-274.68	DIPPR
n-decane	10	22	0	17.74	-300.62	DIPPR
n-undecane	11	24	0	22.78	-326.6	DIPPR
n-dodecane	12	26	0	28.203	-352.13	DIPPR

Compound Name	C	H	O	Gf[298K]	Hf[298K]	METHOD
n-tridecane	13	28	0	33.681	-377.69	DIPPR
2-methyltetrahydrofuan	5	10	1	-94.37	-221.33	MG
ethanol	2	6	1	-173.86	-276.98	DIPPR
n-pentane	5	12	0	-9.928	-173.51	DIPPR
n-hexane	6	14	0	-4.154	-198.66	DIPPR
carbon dioxide	1	0	2	-394.37	-393.51	DIPPR
water	0	2	1	-237.214	-285.83	DIPPR
glycerol	3	8	3	-478.6	-669.6	DIPPR
Oxygen	0	0	2	0	0	DIPPR
Wood	1	1.6	0.68	N/A	N/A	MG
BioOil	1	2.59	1.05	N/A	N/A	MG
Char	1	0.72	0.22	N/A	N/A	MG
Pyrolysis Gas	1	0.54	1.3	N/A	N/A	MG
cellulose	6	10	5	N/A	N/A	MG
hemicellulose	5	8	4	N/A	N/A	MG
lignin	10	12	3	N/A	N/A	MG
biomass(woody)	1	1.6	0.68	N/A	N/A	MG
coumaryl alcohol	9	10	2	-82.66	-217.92	MG
coniferyl alcohol	10	12	3	-193.46	-364.45	MG
sinapyl alcohol	11	14	4	-296.08	-505.39	MG
Ethyl 4-oxo-5-phenylpentanoate	11	20	3	-274.97	-513.53	MG
1-Hexene	6	12	0	83.6	-72.24	DIPPR
Propan-1-ol	3	8	1	-167	-300.8	DIPPR
n-Butanol	4	10	1	-161.4	-326.4	DIPPR
CycloHexane	6	12	0	26.77	-156.15	DIPPR
Cyclohexene	6	10	0	103	-38.2	DIPPR
Parrafin	20	42	0	N/A	N/A	MG
MTBE (tert-Butyl methyl ether)	5	12	1	-104.72	-271.82	MG

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:

Compound	**S/L/G**	**μ (kG)**	**Temp Coff. G/K**
C	gaseous	669.58	-157.99
C diamond	solid	2.90	-2.38
C graphite	solid	0.00	-5.74
C graphite	solid	0.00	-5.69
CH	gaseous	560.75	-182.92
CHO_2^- formiate	aq. sol.	-351.04	-92.05
CH_2	gaseous	371.87	-181.04
CH_2 polyethylene	solid	4.40	-25.34
CH_2O formaldehyde	gaseous	-112.97	-218.66
CH_2O formaldehyde	gaseous	-109.90	-218.66
CH_2O formaldehyde	gaseous	-109.93	-218.80
CH_2O_2 formic acid	gaseous	-350.03	-251.60
CH_2O_2 formic acid	liquid	-359.57	-129.00
CH_2O_2 formic acid	aq. sol.	-372.38	-163.18
CH_3	gaseous	147.92	-194.05
CH_4	gaseous	-50.75	-186.15
CH_4 methane	gaseous	-50.81	-186.10
CH_4 methane	gaseous	-50.89	-186.19
CH_4O methanol	gaseous	-162.52	-239.70
CH_4O methanol	liquid	-166.36	-126.78
CH_4O methanol	liquid	-166.34	-126.70
CO	gaseous	-137.15	-197.56
CO	gaseous	-137.16	-197.53
CO_2	gaseous	-394.36	-213.64
CO_2	gaseous	-394.40	-213.68
CO_2	aq. sol.	-385.99	117.53
CO_2 in air	solution	-4114.32	-281.21
CO_3^{2-} carbonate ion	aq. sol.	-527.90	56.90
CO_3H^- hydrogen- carbonate ion	aq. sol.	-586.85	-91.21
CO_3H_2	aq. sol.	-608.25	N/A
C_2H_2 ethyne	gaseous	209.20	-200.83
C_2H_2 ethyne	gaseous	209.17	-200.85
C_2H_2 ethyne	gaseous	209.24	-200.80

Compound	**S/L/G**	**μ (kG)**	**Temp Coff. G/K**
$C_2H_3O_2^-$ acetic acid anion	aq. sol.	-369.41	-86.61
C_2H_4 ethene, ethylene	gaseous	68.12	-219.45
C_2H_4 ethene, ethylene	gaseous	68.36	-219.22
C_2H_4 ethene, ethylene	gaseous	68.16	-219.40
C_2H_4O acetaldehyde	gaseous	-132.92	-264.20
C_2H_4O epoxyethane	gaseous	-11.84	-243.70
$C_2H_4O_2$ acetic acid	gaseous	-378.95	-282.50
$C_2H_4O_2$ acetic acid	liquid	-389.95	-159.83
$C_2H_4O_2$ acetic acid	liquid	-390.28	-159.80
$C_2H_4O_2$ acetic acid	aq. sol.	-396.56	-178.66
$C_2H_4O_4$ formic acid, dimer	gaseous	-713.17	-348.70
C_2H_6 ethane	gaseous	-32.89	-229.49
C_2H_6 ethane	gaseous	-32.62	-229.50
C_2H_6O dimethyl ether	gaseous	-114.07	-266.60
C_2H_6O ethanol	gaseous	-168.57	-282.00
C_2H_6O ethanol	liquid	-174.89	-160.67
C_2H_6O ethanol	liquid	174.72	-160.70
$C_2H_6O_2$ ethane-1,2-diol, ethylene glycol	liquid	-327.07	-179.50
C_3H_4 propadiene	gaseous	202.38	-234.90
C_3H_4 propyne	gaseous	194.16	-248.10
C_3H_6 propene	gaseous	74.66	-226.90
C_3H_6 cyclopropane	gaseous	104.11	-237.90
C_3H_6O propanone, acetone	gaseous	-151.82	-294.90
C_3H_6O propanone, acetone	liquid	-154.83	-200.00
C_3H_8 propane	gaseous	-23.43	-269.90
C_4H_6 but-1-ene	gaseous	72.03	-307.40
C_4H_8 but-2-ene cis	gaseous	67.20	-300.80
C_4H_8 but-2-ene trans	gaseous	64.16	-296.50
$C_4H_8O_2$ ethyl acetate	liquid	-323.19	-259.00
C_4H_{10} butane	gaseous	-15.62	-310.00

Compound	**S/L/G**	**μ (kG)**	**Temp Coff. G/K**
C_4H_{10} 2-methylpropane	gaseous	-17.92	-294.60
C_5H_{10} cyclopentane	gaseous	38.67	-292.90
C_5H_{10} cyclopentane	liquid	36.49	-204.10
C_5H_{12} pentane	gaseous	-8.11	-348.40
C_5H_{12} pentane	liquid	-9.21	-262.70
C_5H_{12} 2,2-dimethylpropane	gaseous	-15.18	-306.40
C_6H_6 benzene	gaseous	129.73	-269.20
C_6H_6O phenol	solid	-47.50	-142.00
$C_6H_{10}O_5$ glycogen	aq. sol.	-662.49	
C_6H_{12} cyclohexane	gaseous	31.75	-298.20
C_6H_{12} cyclohexane	liquid	26.83	-204.10
$C_6H_{12}O_6$ fruit sugar, fructose	aq. sol.	-915.59	N/A
$C_6H_{12}O_6$ grape sugar, D-glucose	solid	-910.49	N/A
$C_6H_{12}O_6$ grape sugar, D-glucose	aq. sol.	-917.44	N/A
C_6H_{14} hexane	gaseous	0.30	-386.80
C_6H_{14} hexane	liquid	-4.26	-296.00
C_7H_8 methylbenzene, toluene	gaseous	122.39	-319.70
C_7H_8 methylbenzene, toluene	liquid	110.61	-219.00
C_8H_{18} octane	gaseous	17.44	-463.70
C_8H_{18} octane	liquid	6.41	-361.20
C_8H_{18} 2,2,4-trimethylpentane	gaseous	13.09	-425.20
C_8H_{18} 2,2,4-trimethylpentane	liquid	6.32	-330.00
$C_{12}H_{22}O_{11}$ cane sugar	solid	-1543.52	-360.00
$C_{12}H_{22}O_{11}$ cane sugar	aq. sol.	-1550.63	N/A
$C_{12}H_{22}O_{11}$ cane sugar	aq. sol.	-1552.22	N/A

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:

Compound	All Rxns	Rxn1	Rxn2
carbon monoxide	v1	0	0
hydrogen	v2	0	0
formic acid	v3	0	0
acetic acid	v4	0	0
propanoic acid	v5	0	0
acetone	v6	0	0
formaldehyde	v7	0	0
acetaldehyde	v8	0	0
oxaldehyde	v9	0	0
glycol aldehyde	v10	0	0
acetal	v11	0	0
cellulose acetate	v12	0	0
cellobiose	v13	0	0
fructose	v14	0	0
glucose	v15	-1	-1
DME(dimethyl ether)	v16	0	0
HMF (5-Hydroxymethyl-2-furaldehyde)	v17	0	0
furan-2-carbaldehyde	v18	0	0
phenol	v19	0	0
isoeugenol	v20	0	0
eugenol	v21	0	0
cresol	v22	0	0
xylose	v23	0	0
furfural	v24	0	0
furfuryl alcohol	v25	0	0
methyl furan	v26	0	0
tetrahydrofurfural	v27	0	0
levulinic acid	v28	0	0
angelica lactone	v29	0	0

Compound	All Rxns	Rxn1	Rxn2
gamma valerolactone	v30	0	0
1,4-pentanediol	v31	0	0
1,2-dihydroxypentanal	v32	0	0
1-pentanol	v33	0	0
methanol	v34	0	0
methane	v35	0	3
ethane	v36	0	0
propane	v37	0	0
n-butane	v38	0	0
n-octane	v39	0	0
ethene	v40	0	0
propene	v41	0	0
but-1-ene	v42	0	0
but-2-ene	v43	0	0
benzene	v44	0	0
toluene	v45	0	0
p-xylene	v46	0	0
m-xylene	v47	0	0
o-xylene	v48	0	0
o-cresol	v49	0	0
p-cresol	v50	0	0
m-cresol	v51	0	0
indane	v52	0	0
tetralin	v53	0	0
n-nonane	v54	0	0
n-decane	v55	0	0
n-undecane	v56	0	0
n-dodecane	v57	0	0
n-tridecane	v58	0	0
2-methyltetrahydrofuan	v59	0	0
ethanol	v60	2	0
n-pentane	v61	0	0
n-hexane	v62	0	0
carbon dioxide	v63	2	3
water	v64	0	0
glycerol	v65	0	0
Oxygen	v66	0	0
Wood	v67	0	0

Compound	All Rxns	Rxn1	Rxn2
BioOil	v68	0	0
Char	v69	0	0
Pyrolysis Gas	v70	0	0
cellulose	v71	0	0
hemicellulose	v72	0	0
lignin	v73	0	0
biomass(woody)	v74	0	0
coumaryl alcohol	v75	0	0
coniferyl alcohol	v76	0	0
sinapyl alcohol	v77	0	0
Ethyl 4-oxo-5-phenylpentanoate	v78	0	0
1-Hexene	v79	0	0
Propan-1-ol	v80	0	0
n-Butanol	v81	0	0
CycloHexane	v82	0	0
Cyclohexene	v83	0	0
Parrafin	v84	0	0
MTBE (tert-Butyl methyl ether)	v85	0	0
	Delta G Rxn	-259.07	-457.1
	Delta H Rxn	-213.62	-276.7
	Delta $ Rxn	5	0
	Ox. State Reac	0	0
	Ox State Prod.	4	0
	G Reactants	-877.39	-877.3
	G Products	-1136.46	-1334.5

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

Reactions	Delta G_f Rxn	Delta H_f Rxn	$Economics (+ = Profit)	Reaction Number
1*glucose ----------> 2*ethanol + 2*carbon dioxide	-259.07	-213.62	5	Rxn 1
1*glucose ----------> 3*methane + 3*carbon dioxide	-457.1	-276.7	0	Rxn 2

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