Metabolic Modeling of Inborn Errors of Metabolism: Carnitine Palmitoyltransferase II Deficiency and Respiratory Chain Complex I Deficiency

Citation

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:24078365

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
Metabolic Modeling of Inborn Errors of Metabolism:

Carnitine Palmitoyltransferase II Deficiency

and Respiratory Chain Complex I Deficiency

Judy Brewer

A Thesis in the Field of Biotechnology

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

November 2015
Abstract

The research goal was to assess the current capabilities of a metabolic modeling environment to support exploration of inborn errors of metabolism (IEMs); and to assess whether, drawing on evidence from published studies of EMs, the current capabilities of this modeling environment correlate with clinical measures of energy production, fatty acid oxidation, accumulation of toxic by-products of defective metabolism, and mitigation via therapeutic agents.

IEMs comprise several hundred disorders of energy production, often with significant impact on morbidity and mortality. Despite advances in genomic medicine, currently the majority of therapeutic options for IEMs are supportive only, and most only weakly evidenced. Metabolic modeling could potentially offer an in silico alternative for exploring therapeutic possibilities.

This research established models of two inborn errors of metabolism (IEMs), carnitine palmitoyltransferase (CPT) II deficiency and respiratory chain complex I deficiency, allowing exploration of combinations of IEMs at different degrees of enzyme deficiency. It utilized a modified version of the human metabolic network reconstruction, Recon 2, which includes known metabolic reactions and metabolites in human cells, and which allows constraint-based modeling within a computational and mathematical representation of human metabolism. It utilized the Matlab-based COBRA (Constraint-based Reconstruction and Analysis) Toolbox 2.0, and a customized suite of functions, to model ATP production, long-chain fatty acid oxidation (LCFA), and acylcarnitine accumulation in response to varying defect levels, inputs and a simulated candidate therapy.
Following significant curation of the metabolic network reconstruction and customization of COBRA/Matlab functions, this study demonstrated that ATP production and LCFA oxidation were within expected ranges, and correlated with clinical data for enzyme deficiencies, while acylcarnitine accumulation inversely correlated with the degree of enzyme deficiency; and that it was possible to simulate upregulation of enzyme activity with a therapeutic agent.

Results of the curation effort contributed to development of an updated version of the metabolic reconstruction Recon 2. Customization of modeling approaches resulted in a suite of re-usable Matlab functions and scripts usable with COBRA Toolbox methods available for further exploration of IEMs.

While this research points to potentially greater suitability of kinetic modeling for some aspects of metabolic modeling of IEMs, it helps to demonstrate potential viability of constraint-based steady state modeling as a means to explore some clinically relevant measures of metabolic function for single and combined inborn errors of metabolism.
Acknowledgements

Many people have encouraged me throughout this work. For inspiring me on this adventure of exploring metabolic pathways as a window into inborn errors of metabolism, my deep thanks to Melissa Moore, PhD, director of the Moore laboratory at the University of Massachusetts Medical School.

Thank you to Nathan Lewis, PhD, from University of California at San Diego for discussions about Recon 2 and the COBRA Toolbox; Ines Thiele, PhD at the University of Luxemburg for her work on Recon 2 and inspiring discussions about metabolic modeling and inborn errors of metabolism; Steven Denkin, PhD and Maura McGlame of the Biotechnology Program at Harvard Extension School; and friends and extended family for their encouragement.

Above all I would like to thank my thesis director, Neil Swainston, PhD, Senior Experimental Officer at the Manchester Institute of Biotechnology at the University of Manchester in the United Kingdom, and one of the original authors of Recon 2. His in-depth knowledge of Recon 2, and insight, dedication, and patient guidance throughout the development of this thesis, including on model curation and development of custom Matlab/COBRA functions, has been invaluable. His interest in potential applications of metabolic modeling to help support an understanding of inborn errors of metabolism—and his willingness to take on a master’s degree student from across the pond—are very deeply appreciated.
# Table of Contents

Acknowledgements ...........................v

List of Tables .................................................. viii

List of Figures ................................................... ix

I.  Introduction .................................................. 1

  Relevance of Metabolic Modeling to Researching IEMs .....................2
  Modeling Carnitine Palmitoyltransferase II Deficiency .........................6
  Modeling Respiratory Chain Complex I Deficiency .............................9
  Modeling Combinations of CPT II and RCCI Deficiencies ..................10
  Modeling Bezafibrate as Therapeutic Intervention for IEMs .................11
  Establishing a Constraint-Based Modeling Environment ......................13
  Goals, Hypothesis, and Research Aims ..............................................16
  Known Limitations .............................................................................17

II.  Materials and Methods ........................................... 19

  Modifications to the Metabolic Network Reconstruction Recon 2 ..........19
  Publicly Available Data for Model Corroboration ...............................20
  Protocol for Preparing Model, Performing Tests, and Analyzing Data .......21
  Customization of the Constraint-Based Modeling Environment .........24

III. Results .................................................................. 28

  Tests in a Physiologically Normal Model .........................................29
  Tests within CPT II Deficiency Models ............................................31
  Tests within Respiratory Chain Complex I Deficiency Models ............36
  Simulating Bezafibrate Treatment ....................................................37
  Combined IEM Modeling ...............................................................39
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV. Discussion</td>
<td>41</td>
</tr>
<tr>
<td>Observations of Methods and Materials</td>
<td>41</td>
</tr>
<tr>
<td>Significance of Results</td>
<td>46</td>
</tr>
<tr>
<td>Potential Further Investigation of IEMs through Metabolic Modeling</td>
<td>555</td>
</tr>
<tr>
<td>Conclusion</td>
<td>57</td>
</tr>
<tr>
<td>Appendix 1. Publicly Available IEM Reference Data</td>
<td>58</td>
</tr>
<tr>
<td>Appendix 2. Research Protocol and Computation Environment Components</td>
<td>60</td>
</tr>
<tr>
<td>Appendix 3: Test Harness and Custom COBRA Methods</td>
<td>62</td>
</tr>
<tr>
<td>References</td>
<td>68</td>
</tr>
</tbody>
</table>
List of Tables

Table 1. Data representative of select CPT II genotypes and enzyme activity levels 58
Table 2. Data for modeling bezafibrate treatment of CTP II deficiency .................... 58
Table 3. Data for exploring models of CPT II plus respiratory chain disorders ........ 59
Table 4. Data for exploring models of bezafibrate treatment of Complex I disorder 59
Table 5. Components of modeling environment.........................................................60
Table 6. Preparation of materials and model ..............................................................60
Table 7. Research aims, tasks, expected results, comparison data .............................61
List of Figures

Figure 1. Carnitine Shuttle showing transport of long-chain fatty acids (LCFA) across the inner mitochondrial membrane via the carnitine shuttle ......................... 6

Figure 2. Respiratory Chain Complex I .................................................................. 10

Figure 3. Predicted vs Modeled ATP in CPT II deficiency ................................. 32

Figure 4. Reduced ATP production proportional to reduced CPT II enzyme activity, when fed palmitate ................................................................. 33

Figure 5. Fatty acid oxidation pathway ................................................................. 34

Figure 6. Increased palmitoylcarnitine accumulation with reduced CPT II enzyme activity, when fed palmitate ................................................................. 35

Figure 7. ATP production in CPT II deficiencies with palmitate vs. palmitate + glucose ......................................................................................... 36

Figure 8. Limited reduction of ATP production with reduced RC Complex I enzyme activity, when fed palmitate ................................................................. 37

Figure 9. Upregulation of ATP production by bezafibrate in CPT II deficiency ...... 38

Figure 10. ATP production with combined enzyme disorders, without and with bezafibrate ......................................................................................... 39
Chapter I
Introduction

This research establishes models of two inborn errors of metabolism (IEMs), carnitine palmitoyltransferase (CPT) II deficiency and respiratory chain complex I deficiency at different levels of modeled enzyme defects, and combinations of both. It utilizes a modified version of the human metabolic network reconstruction, Recon 2, (Thiele et al., 2013), allowing constraint-based modeling to be performed within a computational and mathematical representation of human metabolism. It utilizes the Matlab-based COBRA (Constraint-based Reconstruction and Analysis) Toolbox 2.0 (Schellenberger et al., 2011), with modified COBRA functions and customized Matlab functions to simulate changes to ATP production, long-chain fatty acid oxidation, and acylcarnitine accumulations.

The study tested robustness of these models for simulating clinically relevant measures of metabolic function for single and combined inborn errors of metabolism. Tests were conducted first in a simulated healthy state; then by modeling physiological data for varying degrees of CPT II and respiratory chain complex I deficiencies; then by modeling experimental data for bezafibrate therapy for a single deficiency; and finally by modeling combinations of deficiencies.

Results are provided for both IEMs singly, jointly, and in response to simulated bezafibrate therapy. A suite of re-usable Matlab COBRA functions and a test harness is provided that is adaptable for modeling other inborn errors of metabolism. Discussion includes significance of results, advantages and limitations of materials, methods and the modeling environment used to achieve the goals; potential applications of this modeling approach to investigation and/or management of
therapies for inborn errors of metabolism; and improvements to the modeling environment that could better support investigation of IEMs.

Relevance of Metabolic Modeling to Researching IEMs

Metabolic modeling is currently in a rapid stage of evolution, with expanded biological databases and new modeling tools appearing frequently. Current metabolic models are primarily either knowledge-based or inference-based. Knowledge-based reconstructions use a constraint-based approach based on biochemical and genetic data (Hyduke, Lewis, & Palsson, 2013). Inferential approaches are based on statistical analyses of high-throughput data (Oberhardt, Palsson, & Papin, 2009). They have the capacity to identify interactions on a genome-wide scale, but lack the capacity for stoichiometrically valid modeling of therapeutic or toxic interventions. Hybrid approaches such as PathWave (Schramm et al., 2010) combine the strengths of reconstructive and inferential approaches, allowing insights from constraint-based modeling to inform high-throughput data analysis (Lewis et al., 2010).

A knowledge-based reconstruction for researching IEMs should ideally include comprehensive biological data including all genes, reactions and metabolites for the modeled tissue. It should allow the creation of IEM models representing different degrees of severity of enzyme defects. Ideally it would include a mechanism to reflect potential transcriptional regulatory effects and complex network effects of IEMs under investigation; and the ability to integrate biochemical information for drugs under investigation. A constraint-based approach for modeling would allow investigation of physiologically normal states, various deficiency levels, rates of metabolic catabolism, synthesis, and accumulation; and interactions with a variety of drugs with different mechanics of interaction.
Potentially Relevant Characteristics of IEMs for Modeling

Metabolic modeling provides a potentially useful avenue for investigating IEMs, which can have significantly impact on morbidity and mortality, yet which because of their diversity and other characteristics can be challenging to research.

IEMs comprise several hundred disorders of energy production (Campeau, Scriver, & Mitchell, 2008). Broadly characterized, IEMs can include disorders of metabolism of carbohydrates, amino acids, urea cycle, organic acids, fatty acid oxidation, oxidative phosphorylation, porphyrin, purine and pyrimidine, steroids, peroxisomal function, and lysosomal storage. Mitochondrial disorders, a subset of IEMs, are encoded by mitochondrial DNA (mtDNA), or by nuclear DNA expressed within mitochondria.

IEMs are individually rare and phenotypically diverse, and frequently have multi-system, multi-organ impact (Lieber et al., 2013; Sahoo, Franzson, Jonsson, & Thiele, 2012). Individuals with IEMs are at risk for a variety of adverse events, ranging from routine activities and events such as overnight fasting (Derks et al., 2007), to excessive physical exertion (Ørngreen, Ejstrup, & Vissing, 2003), and infections (Edmonds et al., 2002). Individuals with IEMs are also at risk from medications such as statins (Fernandez, Spatz, Jablecki, & Phillips, 2011; Vladutiu, 2008; Vladutiu et al., 2006); anesthetics (Amacher, 2005); and antibiotics and antiretrovirals (Zhang et al., 2009). The general inability to reliably predict which individuals are at what degree of risk from different events can leave individuals with IEMs unprepared to exercise appropriate caution, and consequently more vulnerable to complications that can adversely impact health.

Some IEMs have a low correspondence of genotype to phenotype, although detailed mutational profiles are increasingly leading to phenotypically clearer
associations particularly for mtDNA-based disorders (DiMauro & Schon, 2008).

Bénit notes the extent of interindividual genetic variability due to a variety of factors in nuclear and mitochondrial genetic background that may influence the phenotype of mitochondrial disorders (Bénit, El-Khoury, Schiff, Sainsard-Chanet, & Rustin, 2010). Mootha and Hirschhorn confirm that genome-wide analysis shows significant interindividual variation in ratios of plasma metabolites (2010).

Currently the majority of therapeutic options for IEMs are supportive only, and most are only weakly evidenced (Horvath, Gorman, & Chinnery, 2008; Kerr, 2010; Jerry Vockley, Chapman, & Arnold, 2013). The rarity of individual IEMs makes it difficult to gather sufficient candidates for clinical trials for potential therapies for IEMs (Kerr, 2010; J. Vockley & C. Vockley, 2010). Emerging therapies based on genomic medicine utilize increasingly diverse approaches, but to date there has been little generalization from one IEM to another (Schiff, Bénit, Jacobs, Vockley, & Rustin, 2012).

Metabolic modeling could potentially support in silico exploration of personalized therapeutic approaches for individuals with rare, super-rare and ultra-rare IEMs.

Modeling Compound and Double Heterozygosity in IEMs

Combinations of IEMs represent one form of interindividual variation that may lend itself to exploration through metabolic modeling. Vockley hypothesized that combined disorders could manifest atypical phenotypes, potentially with earlier onset or more severe symptoms than expected for either partial deficiency alone. He suggested that a synergistic gating effect might operate on affected metabolic
pathways, even where individuals have residual enzyme activity levels that might otherwise not result in manifestation of symptoms (Vockley et al., 2000).

Research with mouse models demonstrated synergistic effects from combined partial enzyme disorders (Schuler et al., 2005). Dipple suggested that more highly connected nodes of metabolic networks may be more vulnerable to synergistic heterozygosity (Dipple, Phelan, & McCabe, 2001).

One challenge in investigating potentially synergistically heterozygous effects is that data on compound and doubly heterozygous IEMs is limited. Even where patients are known to have a combination of disorders, their data is sometimes excluded from studies of IEMs in order to avoid confounding the analysis of monogenic disorders (Anichini et al., 2011). Yet Vockley suggested that double heterozygosity may be more common than homozygous IEMs, and noted the potential relevance of systems biology approaches to support better understanding of these disorders (J. Vockley, 2008).

Until recently, diagnostic workup for IEMs primarily tested only for those disorders for which there was a high index of suspicion based on clinical observation; and of those, often only common mutations would be screened. This approach leaves open the possibility that individuals with unexpected phenotypes of presumed monogenic disorders could instead have undetected concurrent disorders. Initial IEM detection is now often through newborn screening (NBS), which is more likely to identify multiple genetic defects, if present.

The effects of findings of multiple genetic defects through NBS, however, is not yet well understood. In some cases, mutations detected through NBS may be benign, while in other cases, co-occurring defects may have medical significance (Bennett et al., 2012). Modeling could provide a way to explore biochemical
consequences of different combinations of defects to assess their potential for clinical significance, and potentially better predict which combinations of disorders may entail medical risks over time.

Modeling Carnitine Palmitoyltransferase II Deficiency

The primary IEM modeled in this study, carnitine palmitoyltransferase (CPT) II deficiency, is the most common fatty acid oxidation disorder affecting skeletal muscle tissue (Wieser, 2004). CPT II removes carnitine from long-chain fatty acids (LCFA) after helping transport them across the inner mitochondrial membrane, while also preparing LCFAs for beta oxidation (see Figure 1). In the presence of certain CPT2 mutations, long-chain acylcarnitines (LCFAs without the carnitine palmitoyltransferase removed) may accumulate in the inner mitochondrial matrix, or migrate out to damage various organs ("CPT II deficiency - Genetics Home"
Reference”); while a back-up of un-transported LCFAs may accumulate in the cytosol.

Three main phenotypes of CPT II deficiency have been observed. These include a lethal neonatal form; a severe infantile hepatocardiomuscular form; and a myopathic form. The latter is often referred to as the “adult” version, though it may have onset anywhere from early childhood on (Wieser, 2011).

Mutations that truncate CPT2 protein or that compromise its catalytic activity generally correlate with the lowest levels of enzyme activity, and with the infantile and neonatal forms of CPT II deficiency. Mutations that instead affect protein folding or substrate binding generally correlate with higher enzyme activity levels and with the myopathic form of CPT II. In some CPT II cases these folding or binding mutations may be associated with a cascade of thermosensitive conformational changes, for instance in the presence of febrile or exertional overheating. Olpin has hypothesized a degradation of CPT II crystalline structure under these conditions, and subsequent impairment of its ability to bind substrate, as a potential explanation for a triggerable myopathic CPT II phenotype (Olpin et al., 2003).

Compound heterozygosity, involving a combination of mild and severe mutations on different alleles, generally correlates with intermediate levels of enzyme activity and with a relatively severe form of the myopathic phenotype. Conversely, heterozygotes may be mildly symptomatic (“manifesting carriers”) or asymptomatic (Anichini et al., 2011; Bonnefont et al., 2004; Isackson, Bennett, & Vladutiu, 2006; Olpin et al., 2003; Thuillier et al., 2003).

Selected common mutational profiles of CPT II deficiencies modeled in this study, drawn from publically available patient data (Appendix 1, Table 1), include a symptomatic heterozygous form with enzyme activity clustering at 50%; another
Symptomatic heterozygous form enzyme activity clustering at 25%; the S113L homozygous form with enzyme activity clustering at 15%; a compound heterozygous level clustering at 10%; and the R151Q homozygous level with enzyme activity levels clustering at 5%.

Corroboration of a metabolic model requires comparing model behavior with clinical or experimental evidence (Thiele & Palsson, 2010). This already poses a challenge in the case of CPT II deficiency, given the complexity of CPT II genotype-phenotype relationships, and the scarcity of significant clusters of patient data for any given mutation. This challenge of data for model corroboration is compounded by inconsistent utilization of biochemical measures across patient studies. Publicly available patient data may be drawn from different tissue types, with corresponding differences in pathways and reactions. It may have been analyzed using substantially different methodologies, using clinical practices that are not standardized from one lab to the next, or that even within the same lab may exhibit inter-assay variability (Yao et al., 2015).

Nevertheless, CPT II deficiency is the most common fatty acid oxidation (FAO) disorder, and relative to other FAO disorders, significant amounts of public data exist in the form of aggregated patient data from different studies. When comparing data on specific mutational profiles, in combination with specific biochemical data from multiple studies, some patterns that may be relevant for corroboration of metabolic models have been identified.

One such pattern that may be relevant to explore through metabolic modeling is the contrast that several authors have noted (J. P. Bonnefont et al., 2004; Olpin et al., 2003; Yao et al., 2015) between the very low rates of fatty acid oxidation observed for the neonatal and infantile CPT II phenotypes on the one hand, consistent
with low levels of enzyme activity, versus the relatively high levels of fatty acid oxidation observed across diverse mutational profiles and enzyme activity levels for the myopathic CPT II phenotype. This discontinuous clustering of fatty acid oxidation rates according to phenotype has been postulated to be due to thermal instability and abbreviated half-lives of CPT II protein, allowing for the detection of a given level of enzyme, but not a corresponding level of successful beta oxidation (J.P. Bonnefont et al., 2004; Yao et al., 2015). An interesting corroboration of a modeling environment would be if this same separation of measurement ranges were observed in a metabolic model.

Another pattern to examine through modeling could be the ratios of acylcarnitine accumulation. Gempel has proposed that the ratio of certain acylcarnitines—specifically palmitoylcarnitine (C16:0) and oleoylcarnitine (C18:1) to acetylcarnitine (C2)—is indicative of CPT II deficiency (Gempel et al., 2002).

Modeling Respiratory Chain Complex I Deficiency

Mitochondrial respiratory chain complex I deficiency (see Figure 2) is the most common respiratory chain disorder, causing a wide range of mitochondrial disorders including Leigh syndrome and Alper Disease, severe pediatric neurological disorders (Swalwell et al., 2011).
With 45 subunits, some encoded by mtDNA and some by nuclear DNA, many Complex I mutations have been identified. Similarly to CPT II deficiency, there is a general lack of effective therapies for respiratory chain complex I (DiMauro & Rustin, 2009).

Though a wide range of enzyme activity levels have been observed for respiratory chain complex I mutations, many of these mutations are not associated with energy metabolism disorders. It has been suggested that only enzyme activity levels below 20% of the control mean be considered as primary respiratory chain complex I disorders, and that activity levels between 20% to 30% of the control mean be considered indicative of secondary disorders (Hui, Kirby, Thorburn, & Boneh, 2006); though some studies note symptomatic effects up to 50% of complex I enzyme activity. This study models 20% and 50% respiratory chain complex I enzyme activity levels.

Modeling Combinations of CPT II and RCCI Deficiencies

Several clusters of doubly or multiply heterozygous deficiencies of CPT II were documented by Vockley (2000) and Olpin (2003). Vockley’s review included patients who were doubly or multiply heterozygous with CPT II deficiency and...
glycogen storage disease type V, CPT II and myoadenylate deaminase deficiency (AMPD1), and CPT II and respiratory chain (RC) disorders. Olpin’s data on patients with CPT II deficiency included several with concurrent AMPD1 deficiency. Modeling AMPD1 deficiency may not be revealing as it is frequently asymptomatic; its relevance may be in increasing susceptibility to CPT II deficiency. To explore a doubly heterozygous disorder, this study models CPT II deficiency plus respiratory chain complex I disorder. Publically available patient data is available in Appendix 1, Table 2 for reference.

Modeling Bezafibrate as Therapeutic Intervention for IEMs

Even though evidence is limited, there is a broad range of potentially therapeutic interventions which could be suitable for modeling across the spectrum of IEMs. One potential therapeutic agent for long chain fatty acid oxidation disorders that has received research focus, mostly in France, is bezafibrate.

Bezafibrate is a hypolipidemic agent, but has the added action of upregulating some enzymes involved in long chain and very-long chain fatty acid oxidation. Specifically its action on CPT2 appears to be as a peroxisome proliferator activated receptor pan-agonist (PPAR α, β/δ and γ), by stimulation of PPAR δ, correlating with a reduction in CPT II symptoms (Bonnefont et al., 2010). Those authors offer evidence that bezafibrate can stimulate a significant increase (averaging 65% but ranging to over 100%) in CPT II enzyme activity levels and bring enzyme activity back within the control range.

Available experimental data for CPT II deficiency therapy includes evidence of substantial increases in availability of enzymes following administration of bezafibrate to tissue from individuals with CPT II deficiency. This amount of increase
could potentially rescue missense mutations, though not improving severe mutations (Bastin, Aubey, Rotig, Munnich, & Djouadi, 2008). This study simulates the use of bezafibrate with data representing varying CPT II activity levels. Data for comparison is available in Appendix 1, Table 3.

Though much less well studied, recent research shows that bezafibrate may have independent stimulatory effects on respiratory chain complex I, III and IV (Bastin et al., 2008; Ioannou et al., 2010; Noe et al., 2012), however while conversely exhibiting potential mitochondrial toxicity through other effects on respiratory chain complex I (Yamada et al., 2013).

Bastin hypothesized that bezafibrate acts on respiratory chain complexes through a more complex mechanism: bezafibrate activation of peroxisome proliferator activated receptor-gamma (PPAR γ), which stimulates expression of PPAR γ Co-Activator 1α (PGC-1α), in turn regulating nuclear respiratory factors (NRF) 1, NRF2 and mitochondria transcription factor A (Tfam) expression, which upregulate expression of mitochondrial respiratory chain complexes.

Such a complex chain of regulatory processes is beyond the current capabilities of the model, but nevertheless is an interesting reflection of the complexity that would eventually need to be modeled to more closely represent candidate therapeutic interventions. Data for exploring bezafibrate use with respiratory chain Complex I is in Appendix 1, Table 4.

If bezafibrate is modeled, but these aspects of complexity are not fully accounted for, the potentially independent effects of bezafibrate could complicate the evaluation of modeled outcomes of bezafibrate as a therapy for CPT II deficiency, particularly if a respiratory chain disorder might be present in addition to a LCFA oxidation disorder. Any modeling of bezafibrate as a treatment for the combined
disorders of CPT II and respiratory chain complex I in this study is exploratory only, as there is no clinical data available for comparison.

Establishing a Constraint-Based Modeling Environment

Metabolic modeling is currently in a rapid stage of evolution, with expanding biological databases and new modeling tools appearing frequently. Current metabolic models are primarily either knowledge-based or inference-based. Knowledge-based reconstructions use a constraint-based approach, based on biochemical and genetic data (Hyduke et al., 2013), and offer the capacity for stoichiometric validity.

Constraint-based modeling relies on flux-balance analysis, which uses a mathematical representation of metabolic reactions and constraints, with linear equations defined by mass balance, to solve an objective function and calculate the minimum or maximum flux through a given reaction (Orth, Thiele, & Palsson, 2010).

Inferential approaches are based on statistical analysis of high-throughput data (Oberhardt et al., 2009). They have the capacity to identify interactions on a genome-wide scale, but lack the capacity for stoichiometrically valid modeling of therapeutic or toxic interventions. Hybrid approaches such as PathWave (Schramm et al., 2010) combine the strengths of reconstructive and inferential approaches, allowing insights from constraint-based modeling to inform high-throughput data analysis (Lewis et al., 2010).

An environment for modeling IEMs would ideally be a functional metabolic model with stoichiometric validity. It should include comprehensive biological data reflecting transcriptional regulatory and other network effects of IEMs under investigation, and the ability to integrate biochemical information for drugs under investigation. It should allow modeling of normal physiological states and generation
of output simulating ATP and metabolite production. A modeling environment should allow the creation of IEM models representing different degrees of severity of enzyme defects, and testing of drug interactions within the models.

Recon 2

Components of the modeling environment used for this study included a modified version of Recon 2, written in Systems Biology Markup Language (SBML) using an Extensible Markup Language (XML)-based data structure that is extensible if needed to represent additional data types. The most comprehensive consensus-based metabolic network reconstruction at the time of this study, Recon 2 contains 1,789 enzyme-encoding genes, 7440 reactions, and 2626 unique metabolites. It was developed through a series of community jamborees (Thiele et al., 2013). Advantages of Recon 2 over Recon 1 for this study include its support for acylcarnitine and fatty acid oxidation due to incorporation of Sahoo’s Recon1 AC/FAO module (Sahoo et al., 2012).

Recon 2 organizes metabolites into tissue-specific compartments. Gene-to-protein relationships are organized so that one gene can drive multiple enzymes, but also that one enzyme can catalyze multiple reactions, such as is the case with the gene CPT2 and the enzyme it generates, CPT II. The complexity is compounded by reactions sometimes consisting of single steps; sometimes of a few steps on a pathway; and sometimes of the entire pathway. Conversely, a reaction catalyzed by a complex multi-function protein made up of subunits assembled from over 40 different genes, can be described in a single reaction, such as is the case with respiratory chain complex I.
Protocol for Developing a Metabolic Model

Thiele and Palsson define a four-stage, 96-step protocol for developing network reconstructions (Thiele & Palsson, 2010). Previous studies have demonstrated the ability of metabolic modeling to accurately represent physiological states including metabolic disorders. Lewis corroborated models of brain metabolism based on Recon 1, accurately predicting ATP production rates to within 8% of average published values (Lewis et al., 2010). Sahoo developed a model of phenylketonuria (PKU) using the Recon1 acylcarnitine/fatty acid oxidation (AC/FAO) module, performing flux balance analysis on a simulation of the complete loss of phenylalanine hydroxylase (PAH). The model successfully predicted an L-tyrosine secretion of zero, but could not accurately predict L-phenylalanine accumulation in plasma, urine, or cerebrospinal fluid, as it had not simulated dietary intake of phenylalanine (Sahoo et al., 2012). Thiele’s introduction to the capabilities of Recon 2 accurately predicted 77% of IEM biomarkers (Thiele et al., 2013).

Constraint-Based Modeling Tools

This study used a constraint-based modeling approach, with capabilities largely provided by the COBRA (Constraint-based Reconstruction and Analysis) Toolbox 2.0. Schellenberger describes a protocol for using COBRA’s primary functions, and lists several modeling studies already performed with COBRA methods (Schellenberger et al., 2011).

Matlab served as a numerical computing environment and programming language, with Gurobi linear solvers and optimizers. Constraint-based modeling tools are provided by COBRA Toolbox 2.0, which includes pre-scripted Matlab functions to handle many of the modeling tasks used for this study. Key tasks performed with
COBRA Toolbox 2.0 functions include optimization of energy production via flux balance analysis.

A custom test harness was scripted with COBRA/Matlab functions for this study to ensure that tests were performed and could be repeated within consistent parameters and could be used for future modeling of IEMs. Components of the modeling environment including compatible software versions are listed in Appendix 2, Table 5. Additional software used for supporting tasks included XML Viewer, which allowed exploration of detailed information embedded in Recon 2.

Goals, Hypothesis, and Research Aims

The research goal was to assess the current capabilities of the metabolic network reconstruction Recon 2 in a modeling environment based on Matlab and COBRA Toolbox 2.0, in order to support exploration of IEMs, by focusing on specific questions around CPT II deficiency in combination with an additional IEM and a single drug.

Drawing on evidence from multiple CPT II studies, and allowing for the current capabilities of the modeling environment, this project modeled skeletal muscle CPT II activity levels approximating specific CPT2 mutational profiles. It sought positive associations but not precise correlations for different levels of CPT II enzyme activity, by modeling ATP production rates, long chain fatty acid oxidation rates, and simulated accumulation of acylcarnitine.

A primary question under investigation was whether a model of CPT II deficiency, when configured to represent specific levels of enzyme deficiency, correlates with observed reductions in ATP production and LCFA oxidation, and with observed increases in acylcarnitine accumulations. A further question was how an
IEM model would respond with the addition of a simulated therapeutic intervention, or with a modeled transcriptional regulatory reaction (such as the upregulation of CPT II by bezafibrate) or a series of such reactions (such as the previously described more complex upregulation of respiratory chain complex I by bezafibrate).

Known Limitations

Some limitations of this modeling environment relevant to the focus of this study were known in advance. These included that Recon 2, despite extensive community curating, had not been fully curated and might not include all genes, reactions and metabolites needed; therefore the time potentially needed for curation was unknown.

The model does not a priori include supplements and pharmaceuticals that may be interesting from a perspective of therapeutic investigation of IEMs, though there is the capability to add these to the model.

While a stoichiometric metabolic reconstruction is unable to fully represent the complexity of metabolic networks, nor the particular interindividual variation inherent in human metabolism, the capability to represent these network effects was not needed to demonstrate the general capabilities of the model.

The reference data for corroboration of CPT II deficiencies and respiratory chain complex I deficiencies is known to be inconsistently measured and incomplete, yet should be sufficient for indicating general correlations between physiological data and modeling results.

The modeling environment current lacks a mechanism for transcriptional regulatory modeling necessary for biochemically valid representation of the action of the target drug in this study. However, the average upregulatory effect on CPT II
enzyme activity can be simulated as needed. The environment cannot currently model more complex thermosensitive conformational changes potentially precipitating drops in enzyme activity levels observed in some CPT II mutation types, which may require the capability to reflect the crystalline structure of enzymes.

The model in its current state should provide a basis to demonstrate in silico production, catabolism, and accumulation of key indicators of IEMs; and changes in rates of these in response to several types of potentially therapeutic interventions; plus the interaction of combined disorders. Detail on materials, methods and results in the next sections is followed by reflections on potential IEM research in this modeling environment.
Chapter II

Materials and Methods

Materials utilized in this study and described below include a modified version of the metabolic network reconstruction Recon 2, and publicly available clinical and experimental data for model corroboration. Methods include a Matlab-based programming environment with COBRA Toolbox 2 for constraint-based modeling, customized COBRA functions, and a Matlab-based test harness; and a research protocol for preparing the model and performing tests. Tables of clinical data, programming components and research protocol, and function code, are available in Appendices 1, 2, and 3 respectively.

Modifications to the Metabolic Network Reconstruction Recon 2

For the metabolic network reconstruction, a modified version of Recon 2 was used. Recon 2.1.1 (Swainston, 2014), utilized the most well-tested parts of Recon 1, while retaining improvements to fatty acid oxidation and acylcarnitine reactions that had been integrated during the development of Recon 2.0. Further modifications to the model were developed to block reactions with anomalous flux patterns when testing for ATP production; to establish a separate compartment for the inner mitochondrial matrix; and to add a proton shuttle between cytosol and the inner mitochondrial matrix. As with Recon 2, Recon 2.1.1 was written in SBML using an XML-based data structure.

Recon 2.1.1 is further modified during performance of tests using the test harness in the prepModel subroutine to more closely approximate the predominance
of mitochondrial fatty acid beta oxidation for long chain fatty acids, vs the less-efficient peroxisomal fatty acid oxidation.

A baseline version of the modified model with no IEMs was first tested to ensure that ATP production corroborated with physiological norms; then tested with step-wise reductions in enzyme activity levels to confirm corresponding reductions in ATP production levels; and then with ATP production levels for several common CPT mutations.

Publicly Available Data for Model Corroboration

For publicly available reference data for model corroboration, detailed data and sources are available in Appendix 1. The CPT II enzyme activity levels modeled included a 5% level representing a homozygous R151Q neonatal or infantile phenotype; a 10% level representing a compound heterozygote; a 15% level representing the most common homozygous myopathic genotype, S113L; and 25% and 50% levels representing different clusters of symptomatic heterozygous patients, as shown in Appendix 1 Table 1 (Anichini et al., 2011; J.-P. Bonnefont et al., 2004; Isackson et al., 2006; Olpin et al., 2003; Thuillier et al., 2003).

The respiratory chain complex I enzyme activity levels modeled include a 20% level and a 50% level, representing symptomatic clusters of Complex I deficiencies. The bezafibrate reference data (Table 3) modeled showed an average of 65% increase in CPT II mRNA availability following in vitro administration of bezafibrate. Data and sources are available in Table 4 (Bastin et al., 2008). While these in vitro bezafibrate values are neither definitive, nor representative of in vivo clinical tests, they provide general reference points to help corroborate model capabilities.
Protocol for Preparing Model, Performing Tests, and Analyzing Data

The initial model preparation and planned research protocol are available in Tables 6 and 7, respectively, of Appendix 2. Development of the model required additional model curation, and experimentation with COBRA/Matlab function development. The general sequence of tasks to address the three research aims remained unchanged:

1. Established individual models of two IEMs at different levels of enzyme activity, and compared these models to physiological data;
2. Tested each model with a single drug, and compared those to physiological data;
3. Explored a combined model of two IEMs with and without the drug.

Details of the modified protocol follow.

Preparation of the Model

Preparation of the model included steps to confirm that the baseline model, before configuring the model to represent any enzyme defects, produced results corresponding to physiological norms. Tasks included:

1. confirmed data against which to compare IEM models;
2. confirmed the presence of basic nutritional media for metabolic processes, and oxygen;
3. confirmed theoretical maximums for ATP production in the baseline model with the two planned substrates, palmitate and glucose;
4. established choices of carbon sources, including a common dietary fatty acid, palmitic acid (palmitate), and a common sugar (glucose), as well as a choice that combined both of these;
5. scripted a test harness to obtain user choices for test substrates;
6. modified a script for measuring ATP production by optimizing the model for maximum theoretical ATP production, since physiological norms can vary;
7. compared a candidate series of baseline levels of ATP production with theoretical maximums, in a series of candidate modified Recon 2 models;
8. curated selected pathways according to evaluation of expected flux values generated when testing ATP production, including modification of models to simulated expected availability of tissue-specific reactions;
9. developed a script to measure oxidation of long-chain fatty acids by capturing flux through the FAOXC160 reaction, a “lumped” reaction with which Recon 2 aggregates the complete sequence of beta-oxidation steps;
10. developed a script to measure acylcarnitine accumulation by setting a sink reaction on palmitoylcarnitine, after supplying carnitine as a substrate for the accumulation reaction.
11. re-performed planned tests for the baseline model until curation yielded expected results.

Establishing Individual Models of Two IEMs

Establishment of IEM models included steps to replicate multiple documented enzyme activity levels for two different IEMs within the modeling environment, and steps to test whether the model generated data corresponding to three corresponding biological markers. Tasks included:

1. aggregated all CPT II reactions driven by the CPT2 gene into a global CPT II reaction;
2. established user-selectable downregulation of CPT II enzyme activity through the test harness by constraining the aggregated CPT II enzyme reaction;

3. established user-selectable downregulation of respiratory chain complex I enzyme activity by constraining the NADH -> FADH2 reaction (NADH2_u10m);

4. measured ATP production, LCFA oxidation, and acylcarnitine accumulation in CPT II deficiency models at even steps of reduced enzyme activity, then at representative levels for common CPT2 mutations;

5. measured ATP production, LCFA oxidation, and acylcarnitine accumulation in respiratory chain complex I deficiency models at even steps of reduced enzyme activity, and then at a representative level for symptomatic respiratory chain complex I deficiency.

Testing each Model with a Single Drug

The prescribed action of the drug to be tested, bezafibrate, is as a hypolipidemic agent for treatment of hyperlipidemia. An additional effect of bezafibrate is to upregulate the expression of CPT2. Tasks included:

1. developed a script to simulate upregulation of CPT II activity by a drug, bezafibrate, by allowing a user-selectable factoring in of the average up-regulatory effect of bezafibrate on CPT II enzyme activity on the aggregated CPT II enzyme reaction;

2. added option for bezafibrate treatment choice to user selections in test harness.
Exploring a Combined Model of Two IEMs with and without a Drug

The model was subsequently explored with different levels of enzyme defects for combinations of deficiency levels for both IEMS. Tests included ATP production, long chain fatty acid oxidation, acylcarnitine accumulation, and the optional addition of bezafibrate. Tasks included:

1. established user-selectable independent downregulation of two different IEM enzymes at once through the test harness.

Approaches for Debugging

Because of the complexity of the modeling environment, a variety of approaches were required to debug the functions during development. These included:

- Confirmed the reaction:
  - confirmed biochemistry of target reaction;
  - confirmed that there are sufficient appropriate substrates and nutritional media available to fuel the reaction;
  - captured the flux value upon testing, and confirmed that it is as expected for the circumstances;
  - as needed also checked flux through precursor reactions;
  - rechecked whether a single reaction is adequate for modeling targeted enzyme activity, or whether parallel reactions may be diverting flux.

Customization of the Constraint-Based Modeling Environment

For the computational environment, Matlab R2012A was initially used, and eventually replaced with Matlab R2014B for greater stability. The capabilities of
Matlab were extended to support flux balance analysis by adding libSBML (Bornstein, Keating, Jouraku, & Hucka, 2008) for an SBML parser, and Gurobi linear solvers (Gurobi Optimization, Inc., 2015). Data visualization was provided by Microsoft Excel.

For the function library, constraint-based modeling capabilities were provided by the COBRA Toolbox 2.0 (Schellenberger et al., 2011). COBRA is a library of Matlab scripts and functions for key modeling tasks of flux balance analysis such as linear optimization of maximum or minimum objective functions (Reed & Palsson, 2003); as well as for more mundane yet essential tasks such as adding missing reactions to a metabolic network reconstruction, or changing the upper and/or lower bounds of exchange reactions to simulate the supply of substrates to a model.

Installation, configuration, and troubleshooting of these components variously used online installation instructions; and, where these were conflicting or outdated, also product tech support and guidance from the thesis director.

Two types of custom COBRA/Matlab scripts and functions were developed. Custom COBRA scripts and functions were developed to accommodate changes in model syntax due to splitting of bi-directional exchange reactions into separate import and export reactions, which had resulted in an expanded nomenclature for reactions; and to accommodate addition of the inner mitochondrial membrane compartment.

Custom COBRA/Matlab functions were developed to address modeling tasks specific to this study. A Matlab test harness was developed to ensure that these custom functions and scripts could be performed within replicable parameters when testing and exploring the effects of different configurations of enzyme activity, substrate input, and treatment. Subroutines for the test harness were developed to load and prepare the model; to obtain user input on options for enzyme defect levels,
substrates, and treatments to be tested in the model; to perform tests including measurements of ATP production, long-chain fatty-acid oxidation, and acylcarnitine accumulation; and to log and display test results.

A synopsis of custom Matlab scripts and custom COBRA functions is provided here in workflow sequence. Full function code is available in Appendix 3.

**iemTestHarness**

```matlab
function iemTestHarness()

% Query-based interface for managing the exploration of combinations of inborn errors of metabolism (IEMs) in a metabolic model; modeling with different substrates and potential treatments; testing ATP production, long chain fatty acid oxidation, and accumulation of acylcarnitines; and logging and displaying results.

getChoices

function [ CPT2_degree, RCCI_degree, substrateChoice, treatmentChoice ] = getChoices()

%getChoices() offers choices of degree of IEM defects, substrates, and treatments

prepModel

function model = prepModel()

% prepModel.m loads and prepares a model of inborn errors of metabolism for tests of ATP production, long-chain fatty acid oxidation, flux measurements through specific reactions, and detection of accumulation of specific metabolites.

setChoices

function model = setChoices(model, CPT2_degree, RCCI_degree, ...
substrateChoice, treatmentChoice)

% setChoices(model, substrateChoice, treatmentchoice, CPT2_degree, RCCI_degree) sets choices for substrate and treatments in the model; gets baseline flux for key reactions; sets enzyme activity level for one or more IEMs; reblocks carbon export for substrates; and returns the updated model to the test harness.
getBaseFlux

function [ CPT2baseFlux, RCCIbaseFlux ] = getBaseFlux( model )

% getBaseFlux(model) gets baseline flux values for CPT2 and RCCI, without
% returning the modified model

constrainReactions

function model = constrainReactions(model, reaction_names, degreeDefect)

% Adds constraints of multiple reactions, e.g of the type
% v1 + v7 <= 16
% Method from Neil Swainston

runTests

function [ ATPprod, nadh2_flux, LCFAox_flux, pmtcrn_flux ] = runTests(model)

% runTests(model) tests ATP production, RC Complex I flux, long chain
% fatty acid oxidation, and palmitoylCarnitine flux in an IEM model

% Get FBA solution
% set objective to ATP demand reaction
model = changeObjective(model, 'DM_atp_c_');

logTests

function logTests(count, CPT2_degree, RCCI_degree, substrateChoice,
treatmentChoice, ATPprod, nadh2_flux, LCFAox_flux, pmtcrn_flux )

% logTests(...) logs input of test parameters and test results in Excel
Chapter III

Results

In a series of tests of models with different levels of CPT II enzyme activity, results showed suppressed rates of ATP production and LCFA oxidation, and elevated levels of acylcarnitines indicative of CPT II deficiency, generally consistent with expected values.

Addition of a simulated bezafibrate treatment to CPT II deficiency models showed partially restored ATP production and LCFA oxidation, and suppressed acylcarnitine accumulation, as expected. Tests of models of respiratory chain complex I deficiency showed partially suppressed ATP production, no change in LCFA oxidation, and no accumulation of acylcarnitines, also as expected. Tests of models of combined partial deficiencies of CPT II and respiratory chain complex I deficiency showed suppressed ATP production at the lower levels of respiratory chain complex I deficiencies; and with the addition of a simulated bezafibrate therapy, partial restoration of ATP production.

The CPT II deficiency model did not show discontinuities in LCFA oxidation rates relative to enzyme activity levels as might have been expected based on test data noted earlier. This showed low LCFA oxidation rates corresponding with very low CPT II enzyme activity levels for the most severe CPT II phenotypes, but relatively high LCFA oxidation rates for low-to-mid CPT II enzyme activity levels. As noted earlier this discontinuity in fatty acid oxidation rates may potentially be due to different enzyme half-lives between phenotypes.

Initially, development of a baseline model and specific enzyme deficit models did not yield expected values for ATP production, nor subsequently for LCFA
oxidation or acylcarnitine accumulation. Since the goal for a baseline model was to
demonstrate the capability to generate three types of indicators—production,
catabolism, and accumulation in IEMs—in response to reduced levels of enzyme
activity, this necessitated extensive curation, and identification and addressing of
problems in a series of candidate models. These are described in the Materials and
Methods section. The curation, and the scripts and functions developed to capture
these indicators, can serve as the basis for additional metabolic modeling of IEMs in
the future. They included the development of customized scripts for measuring long
chain fatty acid oxidation and acylcarnitine accumulation, and enumeration of various
approaches for working with the model. These improvements also contributed to the
development of an updated Recon 2.2 human metabolic reconstruction (Swainston et
al., n.d.). Detailed results follow in text and charts.

Tests in a Physiologically Normal Model

In a physiologically normal metabolic model, the practical expected yield for
one mole of palmitic acid should be 104 ATP (Salway, 2004), one mole of oxidized
long-chain fatty acid, and no accumulation of acylcarnitine. One mole of glucose
should yield 31 ATP (Salway, 2004); but no long-chain fatty acids or accumulated
acylcarnitine. A combination of one mole of glucose and one mole of palmitic acid
should therefore yield approximately 135 ATP. Given that glucose is not metabolized
through fatty acid oxidation, there should be no additive effect for LCFA oxidation or
acylcarnitine accumulation; the latter values should be comparable to those obtained
from one mole of palmitic acid.

Initially, a series of candidate baseline models demonstrated problems ranging
from no ATP production, to unlimited ATP production, to only negligible reductions
in ATP despite near total reduction of CPT II enzyme activity levels. Models tested included: Recon 1.0 (Duarte et al., 2007; Rolfsson, Palsson, & Thiele, 2011); Recon 2.0 (Thiele et al., 2013); Recon 2.01, Recon 2.02, Recon 2.03 (“Recon 2 - Models,” n.d.); a skeletal muscle model automatically generated from Recon 2.0, and an updated version of a skeletal muscle model generated from Recon 2.0 (Thiele et al., 2013). Different problems were encountered in various versions of these models. An earlier version (Recon 1.0), while more stable, had incomplete acylcarnitine metabolites; some included a set of erroneous reactions; and others required the blockage of certain reactions in order to more accurately model tissue-specificity.

The Recon 2.1.1 model (Swainston, 2014), developed at the University of Manchester and based on Recon 2.1, addressed the majority of these issues. It provided a sufficient foundation from which to make customized modifications, including blockage of peroxisomal fatty acid oxidation to simulate the less efficient conditions with which peroxisomal oxidation is used for long chain fatty acid oxidation, as opposed to for very long chain fatty acid oxidation.

This model differed from most of the other candidate models in that it split the exchange reactions into an inbound and an outbound version requiring modification of several Cobra Toolbox functions utilized in this study. This initially appeared to afford an advantage in increasing precision of handling of bi-directional reactions for measurement of ATP production. However, through later investigation by Neil Swainston and Kieran Smallbone, this proved to be unnecessary. A subsequent version, Recon 2.2, re-collapsed the exchange reactions after this study was largely completed.

Following selection of the Recon 2.1.1 model, with modifications to the models and the methods as described, yields were as either as predicted, or close to
predictions, on all three measures. This enabled general corroboration of the model, and helped establish a baseline model for measuring ATP production, long chain fatty-acid oxidation, and acylcarnitine accumulation in IEMs.

Tests within CPT II Deficiency Models

Baseline testing was initially performed at evenly spaced enzyme activity levels of 1.0, 0.75, 0.50, and 0.25 to facilitate corroboration of model behavior at regularly decreasing levels of enzyme activity. Testing at enzyme activity levels representing common CPT II defects, based on aggregated patient data, followed.

Tests of ATP Production

Tests of ATP production showed a linear reduction in yield nearly proportional to the degree of CPT II enzyme activity. As already observed above, with full CPT II enzyme activity (e.g., no enzyme defect), one mole of palmitic acid (C:16) yields 104 ATP. Based on this, reductions in ATP yield would be expected to be proportional, e.g., 78 ATP at 75% of CPT II enzyme activity level; 52 ATP as 50% of enzyme activity level; and 26 ATP at 25% of enzyme activity levels. Instead, ATP yield in the model at these respective levels was marginally lower than predicted: 77.375 ATP at 75% CPT II enzyme activity level; 50.75 ATP at 50% CPT II activity; and 24.125 ATP at 25% CPT II activity level, as shown in Figure 3.
When establishing the baseline model, discrepancies in ATP yields had consistently shown higher ATP yields than expected, and had been traceable to effective optimization of solutions by COBRA methods due to utilization of alternative pathways, even when primary CPT II enzyme activity was partially or fully blocked. That the ATP values were in this case lower than predicted, rather than higher, suggested a different cause.

Supplementation with carnitine partially normalized the values, delivering a yield of 77.5 ATP, 51 ATP, and 24.5 ATP respectively. Additional substrates and cofactors may explain the remaining discrepancy. Figure 4 shows ATP production at five representative levels of CPT II deficiency discussed in Materials and Methods.
Tests of Long-chain Fatty Acid Oxidation

Values for tests of long chain fatty acid oxidation would theoretically be expected to be proportionally related to the degree of enzyme defect: for one mole of palmitic acid with no enzyme defect, the model should oxidize one mole of palmitic acid; at a 0.75 level of enzyme activity it should oxidize 0.75 mole of palmitic acid; at 0.50 it should oxidize 0.50 mole of palmitic acid; and at .25 it should oxidize .25 of palmitic acid. These were indeed the results returned by the model.

Since fatty acids are oxidized through a sequence of beta oxidation cycles, a “lumped” reaction encompassing multiple fatty acid oxidation steps in one (FAOXC160) was chosen to measure total oxidation of 1 mole of palmitic acid. Figure 5 shows the many beta-oxidation cycles of the fatty acid oxidation pathway; so in this case it is useful to be able to specify a reaction, and run various tests on, a reaction that encompasses all the reactions shown.
Tests of Acylcarnitine Accumulation

The modeling approach used for detecting acylcarnitine accumulation was to establish a palmitoylcarnitine sink reaction, then to measure the flux through that reaction. This very roughly parallels the physiological diversion of carnitine in CPT II deficiency from its normal recycling path via the carnitine shuttle, from the mitochondria back to the cytosol. However in so doing, it does not track the fate of palmitoylcarnitine, though this is of interest in modeling IEMs with pathological bioaccumulation of metabolic intermediates. In steady-state constraint-based modeling, a sink reaction allows one to capture a snapshot of the activity through a specific reaction; while in more dynamic modeling that flux may move through a sequence of reactions.
Tests of acylcarnitine were expected to yield accumulations inversely correlating with the level of CPT II enzyme activity. Accumulation of these metabolites remained at zero through a number of trials, even at only 5% of normal CPT II enzyme activity levels, until the supply of unlimited carnitine to the model, at which point the sink reaction approach was successful.

At 100% CPT II enzyme activity, ATP production was maximized while palmitoylcarnitine was 0%. But once provided with an unlimited supply of carnitine, ATP production declined proportionally to CPT II enzyme activity levels, while palmitoylcarnitine rose, reaching 95% when CPT II activity was at 5%. Figure 6 shows this inverse pattern of accumulation.

![Figure 6. Increased palmitoylcarnitine accumulation with reduced CTP enzyme activity, when fed palmitate](image)

To demonstrate the metabolic resilience available when more than a single carbon source is present, the CPT II deficiency model was fed a combination of palmitic acid and glucose at various levels of enzyme deficiency. At only mild levels of enzyme deficiency, glucose-derived ATP appeared to provide a reserve of energy.
that might potentially buffer against energy demands such as physical activity or illness. At more severe levels of CPT II enzyme deficiency, ATP produced from non-fatty-acid carbon sources including glucose would become critically important for sustaining basic metabolic processes, even without periodic metabolic stress. See figure 7.

![Figure 7. ATP production in CPT II deficiencies with palmitate, vs. palmitate and glucose](image)

Tests within Respiratory Chain Complex I Deficiency Models

Tests of palmitate and glucose within a series of respiratory chain complex I deficiency models showed partially suppressed ATP production, with no change in LCFA oxidation, and no accumulation of palmitoylcarnitine.

At common respiratory chain complex I deficiency levels of 50% and 20% enzyme activity, flux through the NADH2 (complex I) reaction drops from 31 to 15.5 6.2, proportional to the reduction in complex I enzyme activity; while ATP production in the model shows only a mild reduction, from 104 to 88.5 to 79.20.
An incomplete reduction of ATP production is to be expected for complex I because ATP can alternatively be produced via succinate dehydrogenase input through complex II (Dröse, 2013). Figure 8 shows that partial blockages of complex I only result in mild reductions of ATP production, even with nearly complete reductions in flux through the complex I reaction. While it is unclear whether the very mild reductions in ATP production are consistent with enzyme kinetics, it again demonstrates the model’s optimization approach and success in finding alternative pathways.

![Figure 8. Limited reduction of ATP production with reduced RC Complex I enzyme activity, when fed palmitate](image)

Simulating Bezafibrate Treatment

The primary treatment modeled in this study was bezafibrate. Expected results were data simulating partially restored rates of ATP production and LCFA oxidation for milder cases of CPT II deficiency, and possibly partial reduction of palmitoylcarnitine accumulation.
The action of bezafibrate in CPT II treatment is to upregulate CPT II enzyme activity levels. Given the complexity of the upregulatory action, and the lack of transcriptional regulatory modeling approaches within the model, this action of bezafibrate was simulated by factoring the average increase in CPT II activity levels from in vitro testing of bezafibrate into the enzyme activity setting in the test harness described in materials and methods.

A value of 1.65 was used to simulate upregulation for both enzymes. Data is available in Appendix 1, Table 4.

![Figure 9. Upregulation of ATP production by bezafibrate in CPT II deficiency](image)

With full CPT II enzyme activity, administration of bezafibrate had no effect, since with enzyme activity at 100% and ATP production was already at a maximum of 104. With 75% CPT II enzyme activity, the simulated bezafibrate treatment restored ATP production activity to the maximum. At each of the 50% and 25% levels, bezafibrate simulated a restoration of ATP production to a high enough level that, if valid in vivo, could be nearly asymptomatic. However, at 5%, the lowest level of enzyme activity modeled in this study, the simulated bezafibrate treatment triggered negligible increase in ATP production. See Figure 9.
Combined IEM Modeling

A subsequent series of tests explored combinations of two enzyme deficiencies, CPT II deficiency and respiratory chain complex I deficiency, with and without simulated in-model treatment. Due to the absence of comparison data, there were no specific expectations with regard to results, though correlations of the directionality of changes were expected.

When modeled with several different combinations of two IEMs at different levels of enzyme deficiency, the models showed no synergistic effect on ATP production; and it showed no additive effect at moderate levels of respiratory chain complex I deficiency. A combination of 50% CPT II enzyme activity level and full respiratory chain complex I enzyme activity yielded ATP production of 51, a little less than half of the normally expected 104. ATP production levels remained unchanged at 75% and 50% complex I enzyme activity. Dropping complex I enzyme activity to 25%, ATP production reduced to 43.5; and with complex I enzyme activity of zero, ATP production reduced further to 35.75. These incremental reductions of ATP output roughly matched the incremental reductions in ATP output from single deficiency data for respiratory chain complex I, and showed a combined effect of the deficiencies at lower levels of enzyme activity. It is possible that additional curation of alternative pathways for respiratory chain complex I could result in more uniform additive effects across combined deficiency values.

Additional tests with combined CPT II and respiratory chain complex I deficiencies, without and then with simulated bezafibrate treatment, showed significantly upregulated ATP yield consistent with the additive programmed effects on each disorder. Figure 10 shows the result of simulated bezafibrate treatment on
combined disorders of 50% CPT II and 50% respiratory chain complex I enzyme activity; and also with both disorders modeled at 25% levels of enzyme activity.

![Graph showing ATP production with combined enzyme disorders, without and with bezafibrate](image)

Figure 10. ATP production with combined enzyme disorders, without and with bezafibrate

At the mid-level (50%) of enzyme activity for both IEMs, the proportional independent upregulatory effect of bezafibrate on both CPT II deficiency and respiratory chain deficiencies could, theoretically, substantially increase ATP production above a potentially symptomatic level for either disorder.
Chapter IV
Discussion

All primary research aims were achieved: initial establishment of a physiologically normal skeletal muscle model; development of models with different degrees of CPT II deficiencies; testing models with a simulated treatment; and exploring models with a combination of two IEMs. In addition, the study resulted in improvements to the modeling environment to support further modeling of IEMs; and the study results contributed to an update to the metabolic reconstruction Recon 2.

The following discussion addresses advantages and limitations of the materials, methods and model; the significance of results in relation to original goals; and potential avenues for further investigation.

Observations of Methods and Materials

Materials and methods included publicly available reference data for CPT II and respiratory chain complex I deficiencies, and the modeling environment including the metabolic network reconstruction and the constraint-based modeling tools.

Suitability of Reference Data

The availability of suitable data for comparison of CPT 2 genotype-phenotype biomarkers with results of metabolic modeling is complicated by studies that typically report only a few biochemical measures from among many available types. These may include measures of the amount of mRNA, amount of CPT2 protein, enzyme
activity levels, LCFA oxidation rates, acylcarnitine accumulation, ratios between accumulated acylcarnitines, and creatine phosphokinase levels (CPK).

Additionally, different methods of assessment of enzyme activity levels are used across the field; application of these methods is not standardized from one laboratory to another; and even inter-assay reliability can vary within a given lab. Sampled tissue-types vary from one study to another—in some cases skeletal muscle, in others lymphocytes, in yet others fibrocytes—and all have different levels of reliability. The publicly available data referenced in this study was therefore used only to seek positive associations, but not precise correlations, with data representing specific genotype-phenotype combinations from studies cited in Appendix 1, Table 1.

Published data on individuals with combined IEMs is limited, and clusters of individuals who are doubly or multiply heterozygous with the same or similar conditions are even more limited. For exploration of combinations of IEMs, this study modeled only a very small cluster of individuals with both CPT II deficiency and a respiratory chain complex I disorder, even where there was evidence of other concurrent respiratory chain disorders. However, development of a mechanism for modeling combined disorders may be useful in further modeling of combined disorders. As more data becomes available from metabolic profiling studies, for instance studies of plasma biomarkers of individuals with respiratory chain disorders (Shaham et al., 2010), it may be possible to obtain high-quality data sets with which to better corroborate metabolic models of IEMs.

Suitability of Modeling Environment

The stoichiometric specificity of the Recon 2 model, and the extensive library of constraint-based functions of the COBRA Toolbox 2, together provided a modeling
environment that generally supported the aims of the study. However, several challenges of this modeling environment became evident during the course of the study, as described below.

**Comprehensiveness of model.** Recon 2 currently contains 1,789 enzyme-encoding genes, 7440 reactions, and 2626 unique metabolites; but may also still contain gaps, blocked reactions and dead ends, any of which can impair modeling of specific pathways. The updated model utilized, Recon 2.1.1, which with modifications later led to Recon 2.2, included most reactions and metabolites needed for this study.

**Choice of model versions.** Over the course of experimenting with model versions more or less suitable for the particular aims of this study, there were challenges when porting sample code from model to model. The model version used in this study split out the exchange reactions used for import and export of metabolites into the system in order to better control energy production. This splitting of exchange reactions resulted in altered reaction syntax, which in turn required modification of the syntax of the COBRA functions to implement user selection of substrates and treatments and for up- and down-regulation of enzyme activity levels. This therefore limited the re-usability of this suite of functions for the standard Recon 2 model syntax. Standardization of reaction syntax would increase usability of COBRA methods among metabolic models and their variations.

**Carbon leakage.** In initial trials, the modeling environment demonstrated carbon leakage resulting in seemingly unlimited energy production. This was indicative of a problem in mass-energy balancing in the model, and counterproductive to a key aim of measuring ATP production in the presence of various levels of enzyme defects. Causes included the effectiveness of COBRA methods at finding alternative pathways
through a network of reactions despite partial or complete blockage of key reactions. Curation required establishing an additional mitochondrial intermembrane compartment to capture the proton gradient and better define the electron transport chain through more precise modeling of mass/charge balancing.

Since the model identifies and optimizes flux through all available alternative pathways without regard to enzyme kinetics, it initially gave misleading results by significantly over-utilizing peroxisomal fatty acid oxidation pathways, which while metabolically efficient for very long chain fatty acid oxidation, are less so for long chain fatty acid oxidation. Reddy and Hashimoto note that in the first step of peroxisomal beta-oxidation of long chain fatty acids, energy is dissipated as heat, whereas in mitochondrial beta-oxidation it would be conserved to undergo oxidative phosphorylation (Reddy & Hashimoto, 2001). The model was therefore curated to suppress peroxisomal fatty acid oxidation, better simulating the enzyme kinetics of skeletal muscle tissue and yielding ATP production closer to expected levels.

Limitations with regard to metabolic complexity. Current metabolic models can only approximate a limited number of factors that may influence bioenergetics and resulting phenotypes. For instance, the proposed modeling approach did not, nor was it expected to, represent interindividual genetic variation affecting metabolism outside of the specific conditions established in the study, nor the presence of potential susceptibility variants such as AMPD1 (Olpin et al., 2003), nor specific biochemical processes such as thermosensitive conformational changes which have been hypothesized to explain certain CPT II deficiency phenotypes (Olpin et al., 2003). Furthermore, since the model currently lacks a method for modeling the transcriptional regulatory action of the drug in question, its impact on CPT II activity instead required simulation by factoring in an approximation of experimentally
observed upregulation of enzyme activity. The focus in this study on combining two partial IEMs may improve understanding of one potential source of CPT II genotype-phenotype complexity.

Advantages and Limitations of the COBRA Toolbox 2

Existing COBRA Toolbox 2.0 included many pre-scripted functions useful in developing the test harness. In some cases existing functions needed minor modifications. In other cases customized scripts were developed to suit the particular needs of this study. This study has demonstrated the capacity to accumulate a specific metabolite often indicative of CPT II deficiency, despite previous doubts (Sahoo et al., 2012).

Complexity of the modeling environment. The modeling environment is complex and requires a different knowledge base and skillset than is needed for general research into IEMs, even for users with some coding experience. It requires installing and fine-tuning a complex modeling environment; knowledge of Matlab, and an understanding of how to perform modeling tasks with a variety of COBRA scripts and functions, and how to customize those tasks as needed.

The description of a protocol for developing a model described by Thiele (Thiele & Palsson, 2010) is helpful in mapping out the protocol for establishing a model; and the COBRA tutorial by Jeff Orth (Orth et al., 2010) was indispensable in providing a walk-through of basic modeling tasks, though these exercises were based on an e-coli model, not the more complex Recon 2 human metabolic model. The online community forum supporting Cobra Toolbox was useful (“The openCOBRA Project,” n.d.), as well as the guidance of the thesis director.
Additional tutorials and reference modeling methods, and up-to-date installation instructions, could make the model more approachable to a broader range of users. For instance, a tutorial on use of the modeling environment for human metabolic modeling could be useful, with a repository of functions useful for modeling IEMs.

Significance of Results

The goal of this study was to investigate the extent to which a reconstructive metabolic model could support exploration of IEMs as a foundation for later exploration of questions such as potential therapeutic efficacy and mitigation of toxin risks. It was expected that metabolic models currently have limitations, but also opportunities for improvement, to fulfill their potential as predictive environments. Significance of results across each of the specific research aims is examined below.

Modeling CPT II Deficiency

Three primary measures were chosen for modeling different levels of CPT II deficiency: the rates of ATP production, long-chain fatty-acid oxidation, and palmitoylcarnitine accumulation when fed a substrate of palmitate and/or glucose, in order to contrast results via active pathways.

Tests of ATP production with CPT II deficiency: The primary result expected for measurement of energy production for different degrees of CPT II deficiency was a positive correlation between the degree of enzyme defect and the ATP yield when fed a fatty acid substrate, palmitic acid (palmitate; C:16). As detailed in Results, a continuous linear relationship was demonstrated, with full CPT II enzyme activity yielding 104 ATP, and zero CPT II enzyme activity yielding zero ATP.
From a biochemical perspective, this correlation is to be expected if one restricts metabolic inputs to a single carbon source, and that carbon source is a fatty acid relying on a defective enzyme for transport into the inner mitochondrial matrix for beta oxidation. If CPT II enzyme activity is entirely blocked, and no other pathways in skeletal muscle tissue are available to utilize palmitic acid, then palmitic acid cannot undergo beta oxidation and no ATP can be produced.

The clinical presentation of CPT II, however, shows a discontinuous relationship between enzyme activity level and illness. CPT II deficiency infrequently presents as symptomatic for enzyme activity levels above 70% of the normal range (Olpin et al., 2003). In contrast, below 10% of the normal range of CPT II activity, CPT II deficiency may continuously present with severe symptoms.

**Potential additional ATP tests in this CPT II model.** An avenue that could be further explored through this model is diversity of dietary intake. Normally, individuals consume more than one type of carbon source (such as palmitic acid) at a time; instead they consume a variety of carbon sources: sugars, amino acids, and fatty acids; and of these, a variety of each. Consequently, even with defective fatty acid oxidation, alternative fuel sources are typically available to produce a baseline level of ATP through other metabolic pathways. Mild deficiencies in fatty acid oxidation can sometimes thus be partially compensated for by energy production from normal consumption of other carbon sources.

A deficit in energy supply may only become clinically evident when ATP production from a single fuel source is severely impaired, or when energy demand is high and a compromised metabolic pathway cannot quickly replenish available ATP, or when energy reserves have become depleted for instance from overnight fasting when reliance on stored fat may increase.
Even a two-carbon-source model is greatly simplified since an ordinary diet is composed of a variety of fuels from all three types of carbon sources. An interesting further line of inquiry could be what combinations of alternate fuel sources, maximizing use of alternative pathways, could provide optimal energy supply for specific IEMs and/or combinations of IEMs.

Two such alternative fatty acid fuel sources could be medium chain triglycerides (MCT oil: C8 and C10), which can pass through to the inner mitochondrial membrane to undergo beta-oxidation without the assistance of CPT II; and the synthetic 7-carbon chain triheptanoin, which acts as an odd-chain anaplerotic substrate for the TCA (Roe et al, 2002). Triheptanoin can produce 5-carbon chain ketones for direct uptake to the brain, and is under investigation as a treatment for CPT II deficiency (C. R. Roe et al., 2008). The model would need to be more thoroughly curated to support precise modeling of ATP production and other measures from alternative fatty acids, as well as from additional sugars and amino acids.

Such questions of dynamic demand and supply of energy production in the context of different combinations of energy metabolism disorders could be investigated more thoroughly through kinetic modeling, which allows capacities of enzymes, and competition for such enzymes between metabolites, to be modeled (van Eunen et al., 2013).

**Potential additional ATP tests in a kinetic model.** A possible additional factor explaining the discontinuous clinical presentation of CPT II deficiency relative to ATP production is that CPT II is potentially available in “overabundance.” Therefore, even when the CPT II protein produced by an individual is not fully functional, a
partial reduction in enzyme supply or reduced efficacy of protein binding or folding may have limited clinical impact. The ability to model concentrations of metabolites may be critical to understanding the discontinuities between ATP production and LCFA oxidation, or between ATP production and the clinical picture.

If the lower range of CPT II enzyme activity is considered from a strictly biochemical perspective, one might expect a continuous linear correlation between CPT II enzyme activity levels and ATP production. But if one considers the functional performance of CPT II protein that is driven by different types of CPT2 mutations, and the potential effect on enzyme kinetics and supply, a different picture emerges.

The average enzyme activity level of 5%, representing the R151Q:R151Q genotype, yields 2.83 ATP per mole of palmitic acid without carnitine (3.27 ATP with carnitine), and is associated with the lethal neonatal phenotype for CPT II deficiency. Similarly, the average enzyme activity level of 15%, representing the S113L:S113L genotype, yields 13.48 to 13.90 ATP per mole of palmitic acid, and is associated with the hepatocardio infantile phenotype for CPT II deficiency. Olpin has noted that these two phenotypes are associated with mutations that truncate or compromise the catalytic activity of CPT II. It is be possible that no amount of alternative fuel sources, and no amount of bezafibrate upregulation, can restore an “adequate” level of CPT II enzyme activity.

Analysis of modeling results in relation to the clinical picture may likewise be interesting to explore. The measurement of available energy beyond requirements for sustaining basal metabolism is a frequent matter of investigation in metabolic modeling of experimental and synthetic organisms, as it is useful in determining energy available for biomass growth. In contrast, determination of basal metabolism
requirements in humans is multi-factorial and includes many uncertainties. This study
did not attempt to determine the basal metabolic rate.

As noted in the Methods and Materials section, ATP production in this study
was measured as a maximum theoretical value, rather than by maximizing excess
production beyond basal metabolic requirements. For individuals unable to
significantly utilize fatty acids, the most energy-rich of the three primary nutrients,
maintenance of baseline metabolic processes especially in times of metabolic stress
could be difficult even with significant dietary intake of amino acids and sugars.
Different patterns of morbidity and mortality observed with the two most severe
phenotypes could be indirectly indicative of actual basal metabolic requirements.
However, the pathological impact of insufficient energy to sustain basal metabolic
requirements cannot be considered apart from other factors such as the impact of
increased accumulation of toxic by-products from incomplete catabolism of fatty
acids at lower levels of enzyme activity. At the lowest level of enzyme activity,
increased accumulation of toxic by-products might correlate more directly with organ
damage sometimes seen in the more severe CPT II phenotypes than it would with a
threshold of basal metabolic requirements.

Modeling long chain fatty acid oxidation in CPT II deficiency. The result from
modeling long chain fatty acid oxidation was a positive, continuous correlation
between the degree of enzyme activity and the rate of LCFA oxidation. In theory a
continuous correlation would seem logical since the more palmitic acid is transported
through the inner mitochondrial matrix to undergo beta oxidation, the more long chain
fatty acid should be oxidized in the matrix.

However, a different pattern is observed in clinical data, where a positive yet
discontinuous correlation between the degree of enzyme activity and the rate of long
chain fatty acid oxidation has been observed (Olpin et al., 2003; Thuillier et al., 2003). While enzyme activity levels measured across different CPT2 genotypes form an overlapping continuum, measurements of long chain fatty acid oxidation instead show a clustering of oxidation rates, with a distinct gap between the lethal neonatal and the infantile CPT II phenotypes on the one hand, and the musculoskeletal form on the other hand.

Several factors could be contributing to this discontinuity. As Olpin has noted, mutations that truncate CPT2 protein, or that compromise catalytic activity, generally correlate with the lowest levels of enzyme activity; while mutations that affect protein folding or substrate binding generally correlate with higher enzyme activity levels and with the myopathic form of CPT II. Measurement of the extent of enzyme activity may therefore not be a sufficient predictor of successful LCFA acid oxidation, absent information on the stability and catalytic activity resulting from differing mutations. As has been noted (Yao et al., 2015), mutations that result in short half-lives of enzymes may affect the availability of enzymes to do any useful work in the mitochondrial matrix, despite their apparent availability.

Modeling these types of factors would require not only an approach that can take into account fluctuations in the supply and concentration of enzymes, but would ideally also require the capability to associate the genotype of particular CPT II deficiencies with the resulting crystalline structure of CPT II enzyme, and then in turn with data about the expected kinetic activity of those enzymes. It would be interesting to see whether this would lead to more realistic correlations between modeled ATP yields and rates of LCFA oxidation in metabolic modeling vs clinical observations.

Acylcarnitine accumulation in CPT II deficiency. A third avenue through which to explore potential relevance of this model of CPT II deficiency is accumulation of
acylcarnitines at different levels of CPT II enzyme activity. Accumulation of specific long-chain acylcarnitines is an indicator of the failure to remove carnitine from long-chain fatty acids after transport across the inner mitochondrial membrane (Houten & Wanders, 2010). The expected association between enzyme activity level and acylcarnitine accumulation should therefore be inverse, in that reduced levels of CPT II enzyme activity should result in elevated levels of the corresponding long-chain acylcarnitines.

Indeed, palmitoylcarnitine, the acylcarnitine expected to most directly accumulate when fed palmitic acid, was shown to accumulate at levels inversely proportional to the extent of CPT II activity. While this appears roughly as expected—that the toxic metabolite of an incompletely catabolized fatty acid should accumulate in greater quantities as the activity level of the enzyme declines—there is insufficient patient data to verify what the extent of accumulation should be. Furthermore, the model environment is unable to capture the cascade of dynamically accumulating metabolites.

Modeling Respiratory Chain Complex I Deficiency

Modeling mitochondrial respiratory chain complex I deficiency differs from modeling CPT II deficiency both in methods and results, as detailed above. While complex I is a protein cluster, it only drives one reaction in the electron transport chain. Therefore, simulating complex I enzyme activity does not require an aggregation of the flux values of multiple reactions, as is the case with reduction of CPT II enzyme activity, but rather can be modeled by constraining the bounds on a single reaction.
Modeling complex I deficiencies also differs from modeling CPT II deficiencies in that reduction of enzyme activity to zero should not result in a reduction of ATP production to zero. Since uptake of ATP precursors can be either through complex I or complex II, or both, some output of ATP should be expected even with complete blockage of complex I. The partial but not complete reduction in ATP levels observed for complex I deficiency appears to support this.

Also as expected for complex I deficiency, no changes were observed to fatty acid oxidation rates or accumulation of palmitoylcarnitine. This is consistent with expectations since the respiratory chain is multiple steps downstream from long chain fatty acid oxidation and any potential acylcarnitine accumulation. An interesting question for further exploration would be whether there may be any detectable accumulation of other by-products from incomplete metabolism of NADH, when incompletely metabolized in the complex I reaction.

**Modeling bezafibrate therapy for CPT II deficiency and Complex I deficiency.** The in vitro action of bezafibrate on CPT II deficiency is to upregulate CPT II enzyme activity; in other words, to drive a proliferation of enzyme production, regardless of the quality of the enzymes produced. Since a transcriptional regulatory capability is not currently supported in this modeling environment, this upregulatory action was simulated by factoring in a simple multiplier on the enzyme activity level, as described in the materials and methods section. The results were, not surprisingly, precisely reflective of the level of the bezafibrate “bump” of CPT II activity expected.

The significance is therefore not in the values obtained, since these were directly factored in, but in development of the function for connecting the up- or down-regulatory effect of a therapeutic or toxic influence on a modeled enzyme
reaction. Additionally, it shows the need for more realistic transcriptional regulatory capability for investigating one possible route of drug action for IEMs.

Modeling a Combination of Inborn Errors of Metabolism

The modeling environment in this study was designed not only to test individual IEMs, but also combinations of IEMs including potential synergy between partial defects in activity of multiple defective enzymes. For this study, a combination of the most common fatty acid oxidation disorder and the most common respiratory chain disorder were chosen to model potential combinations of defects.

While there has been some evidence of synergistic effects in mouse models (Schuler et al., 2005), observed results in this in silico human metabolic model only showed additive effects. Tests of combined disorders measured the same indicators as for single disorders: ATP production, LCFA oxidation, palmitoylcarnitine accumulation, with and without bezafibrate therapy.

For ATP production, the results showed additive of the combined enzyme defects at lower levels of respiratory chain complex I deficiency, though not synergistic effects. For bezafibrate therapy, the increase in enzyme activity level showed the additive effect of bezafibrate-induced upregulation of CPT II enzyme activity plus independent bezafibrate-induced upregulation of Complex I enzyme activity.

It is possible that a kinetic model would yield different results than the constraint-based modeling used here, since it could better simulate reductions in concentration of available enzymes in combined disorders, and/or potential competition between enzymes.
Potential Further Investigation of IEMs through Metabolic Modeling

The establishment of a baseline model for testing IEMs and combinations of IEMs, and identification of some advantages and shortcomings of this modeling environment, open possibilities for later investigation of IEMs. A number of observations from this study have highlighted the potential benefits of kinetic modeling to further investigation of IEMs through metabolic modeling. The capabilities to model changes in enzyme kinetics in response to thermal variation, and the ability to model not only concentrations of enzymes, but also binding of metabolites and turnover rates, could be valuable in investigating a variety of questions with IEMs. Several of these are detailed below.

Customizing the “Mitochondrial Cocktail”

Currently many people with IEMs are prescribed a combination of supplements, sometimes up to a dozen or more types. But with the great variety of IEMs, the diversity of presentations, and the rare and super-rare status of some of these disorders, there is limited opportunity for clinical trials. The ability to customize the settings for one or more partial IEMs through the test harness could offer the opportunity of exploring the biochemistry of personalized mitochondrial cocktails. Constraint-based modeling could provide a steady-state view of the mitochondrial cocktail, while kinetic modeling could potentially reveal more about the reactions between different aspects of the treatment.

While metabolic modeling can by no means replace clinical testing, it could be useful to explore interactions between a variety of supplements in an in silico environment. With further refinements to the modeling environment, observations
from such exploration could potentially augment the current approach, which in some cases is still individualized trial and error.

Improved Understanding of Accumulation of Toxic Metabolites

The capability to model the accumulation of potentially toxic metabolic by-products is of critical importance in investigating IEMs, as it is frequently the cause of pathological progression of disorders, for instance in phenylketonuria, and in maple syrup urine disease, where accumulation of toxic intermediaries from incomplete amino acid catabolism can result in early and severe neurological damage.

While the modeling environment used in this study captured simulated accumulation via a single sink reaction, it is not well suited to capture a series of dynamically accumulating metabolites. A better understanding of the accumulation of different metabolic by-products, and how that accumulation might be mitigated, could help inform different approaches for slowing progression of disorders.

Exploration of Potential Synergistic Heterozygosity in IEMs

Use of a kinetic approach to metabolic modeling could potentially yield more revealing results on the competitive interactions and potential synergies between combinations of partial IEMs. The availability of aggregated newborn screening (NBS) data identifying multiple IEMs, and the recent availability of more data on expression profiling and plasma biomarkers for repositories of IEM data, could improve the quality of reference data for model corroboration.
Exploration of Thermosensitive Variants of IEMs

Use of a kinetic model could provide an avenue with which to explore thermosensitive conformational changes in enzymes, and their potential relationship to triggerable myopathic phenotypes, including for CPT II deficiency.

Conclusion

The modeling environment used in this study, based on Recon 2.1.1 and COBRA Toolbox 2, provided an effective environment in which to study energy production, fatty acid oxidation, and the accumulation of toxic metabolites at different levels of enzyme deficiency in two inborn errors of metabolism. The need for detailed curation of fatty acid and respiratory chain pathways at the current state of Recon 2’s development was challenging at times, but helped focus attention on key metabolic interactions for each of the two IEMs modeled. Additionally, the curation led to some improvements which are being captured in the updated model, Recon 2.2. Development of the COBRA/Matlab test harness resulted in a suite of Matlab scripts that could be re-usable or adaptable for modeling other IEMs.

The experience of modeling IEMs in this environment could be portable to other modeling approaches, and could provide a basis for further investigation of IEMs, including combined IEMs.
### Table 1: Data representative of select CPT II genotypes and enzyme activity levels

<table>
<thead>
<tr>
<th>Genotype or zygosity</th>
<th>Source</th>
<th>Patients</th>
<th>CPT II activity, % of controls</th>
<th>Fatty acid oxidation % of controls</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>R151Q:R151Q</td>
<td>Olpin 2003</td>
<td>pt. 3</td>
<td>5%</td>
<td>4%</td>
<td>liver disease, hypotonia, died at 3 mo.</td>
</tr>
<tr>
<td></td>
<td>Thuillier 2003</td>
<td>4 patients</td>
<td>avg. 6%</td>
<td>&lt;10%</td>
<td>infantile</td>
</tr>
<tr>
<td>compound heterozygotes</td>
<td>Isackson 2006</td>
<td>8 patients</td>
<td>10.2 +/- 4.5</td>
<td></td>
<td>adult onset myopathic</td>
</tr>
<tr>
<td></td>
<td>Thuillier 2003</td>
<td>2 patients</td>
<td>10 – 12%</td>
<td>45 – 70%</td>
<td>adult myopathic</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genotype or zygosity</th>
<th>Source</th>
<th>Patients</th>
<th>CPT II activity, % of controls</th>
<th>Fatty acid oxidation % of controls</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>S113L:S113L</td>
<td>Olpin 2003</td>
<td>pt. 14</td>
<td>13%</td>
<td>67%</td>
<td>myalgia, rhabdomyolysis, infection and fasting triggers</td>
</tr>
<tr>
<td></td>
<td>Isackson 2006</td>
<td>8 patients</td>
<td>17.5 +/- 10.5</td>
<td></td>
<td>adult onset myopathic</td>
</tr>
<tr>
<td></td>
<td>Thuillier 2003</td>
<td>2 patients</td>
<td>10 – 12%</td>
<td>45 – 70%</td>
<td>adult myopathic</td>
</tr>
<tr>
<td>heterozygotes</td>
<td>Isackson 2006</td>
<td>9 patients</td>
<td>26.0 +/- 8.6</td>
<td></td>
<td>adult onset myopathic</td>
</tr>
<tr>
<td></td>
<td>Olpin 2003</td>
<td>pt. 24</td>
<td>47%</td>
<td>93%</td>
<td>myalgia, rhabdomyolysis</td>
</tr>
</tbody>
</table>

* Olpin data is from Table 1, Table 2, and text
** Isackson data is from Table 1
*** Thuillier enzyme activity levels are from lymphocytes and fibrocytes

### Table 2: Data for exploring models of CPT II plus respiratory chain disorders

<table>
<thead>
<tr>
<th>Source</th>
<th>Patients</th>
<th>CPT2, enzyme%</th>
<th>RC, enzyme%</th>
<th>Additional IEMs</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vockley 2000 (out of 8 patients, those with CPT II muscle enzyme data plus RC defects)</td>
<td>Pt. 1</td>
<td>CPT II 46%</td>
<td>RC Complex I 26%; citrate synthase 48%; fumarase 46%;</td>
<td>carrier for myoadenylate deaminase deficiency</td>
<td>myalgia, myoglobinuria</td>
</tr>
<tr>
<td></td>
<td>Pt. 2**</td>
<td>CPT II 10%; R503C heterozygous</td>
<td>RC Complex I, 22%; RC Complex III, 49%; RC Complex IV, 34%</td>
<td>β-oxidation enzymes 29-66%;</td>
<td>exercise intolerance; myalgia, weakness</td>
</tr>
<tr>
<td></td>
<td>Pt. 5</td>
<td>CPT II 51%;</td>
<td>RC CI 49%; RC I–III 40%; RC II–III 34%</td>
<td></td>
<td>exercise intolerance; myalgia</td>
</tr>
</tbody>
</table>

* data on respiratory chain complexes from text of Vockley 2000; data in table is incomplete
** RC Complex enzyme levels for Patient 2 are from Vladutiu 2002
### Table 3: Data for modeling bezafibrate treatment of CTP II deficiency

<table>
<thead>
<tr>
<th>Source</th>
<th>Patient</th>
<th>Genotype or zygosity</th>
<th>% increase enzyme activity</th>
<th>% increase palmitoyl CoA oxidation</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Bonnefont et al., 2010)</td>
<td>pt. 1</td>
<td>compound heterozygote</td>
<td>+87%</td>
<td></td>
<td>pre-treatment: myalgia, rhabdomyolysis; post-treatment: decrease in myalgia and rhabdomyolysis; increase in exercise tolerance</td>
</tr>
<tr>
<td></td>
<td>pt. 2</td>
<td>compound heterozygote</td>
<td>+71%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pt. 3</td>
<td>compound heterozygote</td>
<td>+39%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pt. 4</td>
<td>S113L:S113L</td>
<td>+70%*</td>
<td>+58%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pt. 5</td>
<td>compound heterozygote</td>
<td>+86%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pt. 6</td>
<td>S113L:S113L</td>
<td>+130%*</td>
<td>+206%</td>
<td></td>
</tr>
</tbody>
</table>

* enzyme activity percentages extrapolated from histograms

### Table 4: Data for exploring models of bezafibrate treatment of RC complex I disorder

<table>
<thead>
<tr>
<th>Source</th>
<th>Patient</th>
<th>RC Complex I mutation</th>
<th>% increase mRNA*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Bastin et al., 2008)</td>
<td>1</td>
<td>CI NDUFS1</td>
<td>~ +40%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>CI NDUFS3</td>
<td>~ +70%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>CI NDUFS4</td>
<td>~ +120%</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>CI NDUFV1</td>
<td>~ +40%</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>CI NDUFV2</td>
<td>~ +55%</td>
</tr>
</tbody>
</table>

* percentages extrapolated from bar chart, Bastin et al, 2008, figure 4
Appendix 2. Research Protocol and Computation Environment Components

### Table 5: Components of modeling environment

<table>
<thead>
<tr>
<th>Item</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fujitsu U772, i5 3427U, 1.8 GHz machine</td>
<td></td>
</tr>
<tr>
<td>Matlab R2012a 64-bit; Java</td>
<td>numerical computing program language; Java virtual machine</td>
</tr>
<tr>
<td>Matlab Toolboxes, COBRA Toolbox 2.0</td>
<td>bioinformatics, statistical, SBML; constraint-based modeling toolbox</td>
</tr>
<tr>
<td>Recon 2</td>
<td>SBML-based human metabolic network reconstruction</td>
</tr>
<tr>
<td>libSBML 5.1.0b0 64</td>
<td>programming library for systems biology markup language</td>
</tr>
<tr>
<td>Solvers and optimizers</td>
<td>Gurobi 5.5 64-bit (LP, QP, QCP, MILP)</td>
</tr>
<tr>
<td>Visual C++, Win7SDK, .Net Framework</td>
<td>C++ compiler; software developers’ kit; programming model for XML</td>
</tr>
<tr>
<td>XML, XML Schema</td>
<td>for extending SBML data types</td>
</tr>
</tbody>
</table>

### Table 6: Preparation of materials and model

<table>
<thead>
<tr>
<th>Task</th>
<th>Approach</th>
<th>Notes</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluate potential metabolic models</td>
<td>Analyze network reconstruction, toolkits, discuss w/ developers</td>
<td>Selected Recon2 with Matlab/COBRA Toolbox 2.0</td>
<td>DONE</td>
</tr>
<tr>
<td>Install modeling environment</td>
<td>Arrange licenses, debug installation</td>
<td>See table 5</td>
<td>DONE</td>
</tr>
<tr>
<td>Confirm data for model corroboration</td>
<td>Document relevant publicly available physiological data</td>
<td>See tables 1 – 4</td>
<td>DONE</td>
</tr>
<tr>
<td>Determine enzyme activity ranges to model</td>
<td>Review genotype-biomarker-phenotype associations</td>
<td>See page 10</td>
<td>DONE</td>
</tr>
<tr>
<td>Develop tissue-specific subset as needed</td>
<td>Review metabolites and reactions for redundancies and unused reactions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulate regulatory effects of drug</td>
<td>Partially restore enzyme flux unless stoichiometrically valid method becomes available</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-test modeling environment</td>
<td>Test relevant COBRA functions, script modifications as needed</td>
<td></td>
<td>In progress</td>
</tr>
<tr>
<td>Write test harness</td>
<td>Script Matlab functions to create IEM models and perform test series</td>
<td></td>
<td>In progress</td>
</tr>
</tbody>
</table>
Table 7: Research aims, tasks, expected results, comparison data

**Research Aim #1: Establish models of CPT II deficiency with different levels of enzyme defect**

<table>
<thead>
<tr>
<th>Task</th>
<th>Expected results from model</th>
<th>Comparison data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A. Test model with no IEM.</td>
<td>Baseline ATP production, and possibly LCFA oxidation and a normal acylcarnitine profile.</td>
<td>(Bangsbo, et al., 2001; Wibom &amp; Hultman, 1990)</td>
</tr>
<tr>
<td>1B. Test models with CPT II enzyme deficiencies.</td>
<td>Suppressed ATP production, and possibly suppressed LCFA oxidation and indicative acylcarnitine profile.</td>
<td>Table 1 (Isackson et al., 2006; Olpin et al., 2003; Thuillier et al., 2003)</td>
</tr>
</tbody>
</table>

**Research Aim #2: Test models of CPT II deficiency with bezafibrate**

<table>
<thead>
<tr>
<th>Task</th>
<th>Expected results from model</th>
<th>Comparison data</th>
</tr>
</thead>
<tbody>
<tr>
<td>2A. Test CPT II deficiency models with bezafibrate.</td>
<td>Partially restored ATP production, and possibly partially restored LCFA oxidation and partially normalized acylcarnitine profile.</td>
<td>Table 2 (Bonnefont et al., 2010)</td>
</tr>
</tbody>
</table>

**Research Aim #3: Explore model with CPT II deficiency plus an additional IEM**

<table>
<thead>
<tr>
<th>Task</th>
<th>Preparation</th>
<th>Expected results from model</th>
</tr>
</thead>
<tbody>
<tr>
<td>3A. Explore model with CPT II &amp; RC CI disorder.</td>
<td>Suppression of ATP, possibly with synergistic effect.</td>
<td>Table 3 (J Vockley et al., 2000)</td>
</tr>
<tr>
<td>3B. Explore model with RC CI + bezafibrate, but no CPT II.</td>
<td>Partially restored ATP production, or no response because regulatory factors are too complex.</td>
<td>Table 4 (Bastin et al., 2008)</td>
</tr>
<tr>
<td>3C. Explore model with CPT II, RCCI, + bezafibrate.</td>
<td>Same as above.</td>
<td>Hypothetical.</td>
</tr>
</tbody>
</table>
Appendix 3: Test Harness and Custom COBRA Methods

iemTestHarness

function iemTestHarness()

% Query-based interface for managing the exploration of combinations of
% inborn errors of metabolism (IEMs) in a metabolic model; modeling with
% different substrates and potential treatments; testing ATP production,
% long chain fatty acid oxidation, and accumulation of acylcarnitines;
% and logging and displaying results.

% iemTestHarness calls the following local functions:
% getChoices
% prepModel
% setChoices
% which calls getBaseFlux, strfndw, constrainReactionsNewer
% runTests
% calls writeResult, strfndw
% logTests

% Get choices of test options
[CPT2_degree, RCCI_degree, substrateChoice, treatmentChoice ]... = getChoices();

% Ensure that solvers are set
initCobraToolbox;

% Load or read-in model, do basic model prep
model = prepModel();

% Set custom choices in model
model = setChoices( model, CPT2_degree, RCCI_degree, substrateChoice, ...
% treatmentChoice );

% Test ATP production, LCFA oxidation, acylcarnitine panel
[ATPprod, nadh2_flux, LCFAox_flux, pmtcrn_flux] = runTests(model);

% Log test parameters and results, offer to display
% Log test parameters and results
count = 1; % incomplete code
logTests(count, CPT2_degree, RCCI_degree, substrateChoice, ... 
 treatmentChoice, ATPprod, nadh2_flux, LCFAox_flux, pmtcrn_flux)

% Read back results into a table
% Specify sheet and range to read from, then convert format
% from online help forum
sheet = 1;
filename = 'tests.xlsx';
range = 'A1:12';
 [~,~,raw] = xlsread(filename, range);
containsNumbers = cellfun(@isnumeric,raw);
raw(containsNumbers)= ... cellfun(@num2str,raw(containsNumbers));

% Clear model modifications
 clear model
end
getChoices

function [ CPT2_degree, RCCI_degree, substrateChoice, treatmentChoice ]...
  = getChoices();
%getChoices() offers choices of degree of IEM defects, substrates, and treatments

  % Get degree of defect for each enzyme deficiency
  CPT2_degree = input('Enter degree of CPT2 activity, from 0 (no activity) to 1.00 (full activity): 
');
  RCCI_degree = input('Enter degree of RCCI activity: from 0 (no activity) to 1.00 (full activity): 
');
  % Get choice of substrate
  substrateChoice = input('Specify choice(s) of substrate: glucose, fructose, palmitic_acid, glc+hdca: 
');
  % Get choice of treatment
  treatmentChoice = input('Specify treatment choice: bezafibrate, carnitine, nothing: 
');
end

prepModel

function model = prepModel()

% prepModel.m loads and prepares a model of inborn errors of metabolism for tests of ATP production, long-chain fatty acid oxidation, flux measurements through specific reactions, and detection of accumulation of specific metabolites.

  % Load model [currently without choice of reading in a different file]
  load model;
  % Initially block all uptake reactions
  in_reactions = strfindw(model.rxns,'*_IN|*)in');
  model = changeRxnBounds(model, model.rxns(in_reactions), 0, 'b');

  % Block diffusion of fatty acids from extracellular space to cytosol
  model = changeRxnBounds(model, 'VACCT', 0, 'b');

  % Block all peroxisomal FAOX reactions (from NS)
  x_mets_indices = cellfun(@(x)~isempty(x)),
  [~,j,~] = find(model.S(x_mets_indices,:));
  x_rxns = model.rxns(unique(j));
  faox_rxns = model.rxns(cellfun(@(x)~isempty(x)),
  [~,faox_x_rxns = intersect(x_rxns, faox_rxns);]
  model = changeRxnBounds(model, faox_x_rxns, 0, 'b');

  % Allow unlimited oxygen
  model = changeRxnBounds(model, 'EX_o2(e)in', 1000, 'u');
% Blocked phosphate shuttle around Complex I
  model = changeRxnBounds(model, 'P5CRm', 0, 'b');
  model = changeRxnBounds(model, 'PROD2m', 0, 'b');

% Allow unlimited pi in mitochondria:
  model = addReaction(model, 'pi_transport_c_m', 'pi[c] <=> pi[m]');

% Allow unlimited carnitine
  model = changeRxnBounds(model, 'EX_crn(e)in', 1000, 'u');

% Add a sink reaction to detect accumulation of palmitoylcarnitine
  model = addReaction_mod(model, 'pmtcrn_accumulation', 'pmtcrn[c] ->');
end

setChoices

function model = setChoices(model, CPT2_degree, RCCI_degree, ...
  substrateChoice, treatmentChoice)

% setChoices(model, substrateChoice, treatmentChoice, CPT2_degree, ...
% RCCI_degree) sets choices for substrate and treatments in the model;
% gets baseline flux for key reactions; sets enzyme activity level for
% one or more IEMs; reblocks carbon export for substrates; and returns the
% updated model to the test harness.

% Set substrate choice(s)
  switch substrateChoice
    case 'glucose'
      model = changeRxnBounds(model, 'EX_glc(e)in', 1, 'b');
    case 'fructose'
      model = changeRxnBounds(model, 'EX_fru(e)in', 1, 'b');
    case 'palmitic_acid'
      % Import palmitic acid, and block its export
      model = changeRxnBounds(model, 'EX_hdca(e)in', 1, 'b');
      model = changeRxnBounds(model, 'EX_hdca(e)ex', 0, 'b');
    case 'glc+hdca'
      model = changeRxnBounds(model, 'EX_glc(e)in', 1, 'b');
      model = changeRxnBounds(model, 'EX_hdca(e)in', 1, 'b');
      model = changeRxnBounds(model, 'EX_hdca(e)ex', 0, 'b');
  end

% Set treatment choice(s)
  switch treatmentChoice
    case 'bezafibrate'
      % Set parameters for bezafibrate upregulation treatment
      CPT2bezUpreg = 1.65;
      RCCIbezUpreg = 1.00;
    case 'nothing'
      CPT2bezUpreg = 1;
      RCCIbezUpreg = 1;
  end

% Get baseline fluxes before any reductions to enzyme activity
  [CPT2baseFlux, RCCIbaseFlux] = getBaseFlux(model);

% Set degree of defect for CPT2 deficiency
  % Create variable for all cpt2 -driven reactions
  cpt2_rxn_ids = strfndw(model.grRules, '1376.1');
  cpt2_rxn_names = model.rxns(cpt2_rxn_ids);
  % Constrain the sum of all CPT2 reactions according to degree of Defect
% Include CPT2baseFlux factor in case multiple substrates
model = constrainReactions(model, cpt2_rxn_names, CPT2bezUpreg * CPT2_degree * CPT2baseFlux);

% Set degree of defect for all RCCI (NADH2_u10m -driven) reactions
% Set degree of RCCI enzyme activity
model = changeRxnBounds(model, 'NADH2_u10m', RCCIbezUpreg ... * RCCI_degree * RCCIbaseFlux, 'u');

% Reblock export reactions to prevent carbon export for any substrate added
switch substrateChoice
    case 'glucose'
        model = changeRxnBounds(model, 'EX_glc(e)ex', 0, 'b');
    case 'fructose'
        model = changeRxnBounds(model, 'EX_fru(e)ex', 0, 'b');
    case 'palmitic_acid'
        model = changeRxnBounds(model, 'EX_hdca(e)ex', 0, 'b');
    case 'glc+hdca'
        model = changeRxnBounds(model, 'EX_glc(e)ex', 0, 'b');
        model = changeRxnBounds(model, 'EX_hdca(e)ex', 0, 'b');
end
end

getBaseFlux

function [ CPT2baseFlux, RCCIbaseFlux ] = getBaseFlux( model )

% getBaseFlux(model) gets baseline flux values for CPT2 and RCCI, without
% returning the modified model

% Set ATP demand reaction with no IEM, and optimize model
model = changeObjective(model, 'DM_atp_c_');
FBAsolution = optimizeCbModel(model, 'max', 'one');

% Get baseline flux value for CPT2 with no IEM defect
% Supersedes all previous code.
cpt2_rxns = strfndw(model.grRules, '1376.1');
CPT2baseFlux = sum(FBAsolution.x(cpt2_rxns));

% Get baseline flux value for NADH2_u10m (RCCI) with no IEM defect
% Specify RCCI reaction; get ID; run FBA; pull out flux for that
reaction name = 'NADH2_u10m';
reaction_index = strfndw(model.rxns, reaction_name);
% FBAsolution = optimizeCbModel(model, 'max', 'one');
RCCIbaseFlux = FBAsolution.x(reaction_index);
end

constrainReactions

function model = constrainReactions(model, reaction_names, degreeDefect)
% Adds constraints of the multiple reactions, e.g of the type v1 + v7 <= 16
% Method from NS

% Concatenate row representing reaction names:
model.S = [model.S; ismembernew(model.rxns, reaction_names)'];

% Concatenate empty column:
model.S = [model.S zeros(size(model.S, 1),1)];

model.S(end, end) = -1;
model.lb = [model.lb; 0];
model.ub = [model.ub; degreeDefect];
model.b = [model.b; 0];
model.c = [model.c; 0];
model.mets = [model.mets; {'dummy[c]'}];
model.rxns = [model.rxns; {'dummy'}];
model.rxnNames = [model.rxnNames; {'dummy'}];
model.subSystems = [model.subSystems; {'dummy'}];
model.rev = [model.rev; 0];

end

runTests

function [ ATPprod, nadh2_flux, LCFAox_flux, pmtcrn_flux ] = runTests(model)

% runTests(model) tests ATP production, RC Complex I flux, long chain fatty acid oxidation, and palmitoylCarnitine flux in an IEM model

% Get FBA solution
% set objective to ATP demand reaction
model = changeObjective(model, 'DM_atp_c_');

% perform FBA on model
FBAsolution = optimizeCbModel(model, 'max', 'one');

% Get specific test values from FBA solution

% If no solution found, display "0" for ATP and pmtcrn flux values
if FBAsolution.f == 0
    disp('No solution found')
    ATPprod = 0;
    pmtcrn_flux = 0;

    % Write results to Excel file
    else
    filename = 'fluxTableNew.xls';
    writeResult(model, FBAsolution, filename);

    % Get ATP yield
    ATPprod = FBAsolution.f;

    % Get RCCI flux value (and enable checking via debug)
end

66
reaction_name = 'NADH2_u10m';
reaction_index = strfind(model.rxns, reaction_name);
nadh2_flux = FBAsolution.x(reaction_index);

% Get LCFAox flux (and enable checking via debug)
% Try checking output of single reaction first
reaction_name = 'FAOXC160';
reaction_index = strfind(model.rxns, reaction_name);
LCFAox_flux = FBAsolution.x(reaction_index);

% Get flux for acylcarnitines (initially just via palmitoylcarnitine sink)
% capture flux value for palmitoylcarnitine from earlier FBA
% not currently accumulating in proportion to extent of CPT2 defect
reaction_name = 'pmtcrn_accumulation';
reaction_index = strfind(model.rxns, reaction_name);
pmtcrn_flux = FBAsolution.x(reaction_index);
end
end

logTests

function logTests(count, CPT2_degree, RCCI_degree, substrateChoice,
treatmentChoice, ATPprod, nadh2_flux, LCFAox_flux, pmtcrn_flux)

% logTests(...) logs input of test parameters and test results in Excel
filename = 'tests.xlsx';

% Create array of test parameters and results
if count == 1 % not using "count" yet
    A = {'Count', 'CPT2', 'RCCI', 'Substrate', 'Treatment', 'ATP',
         'NADH2', 'LCFAox', 'pmtcarn'; count, CPT2_degree, RCCI_degree,
         substrateChoice, treatmentChoice, ATPprod, nadh2_flux, LCFAox_flux,
         pmtcrn_flux};
elseif count >1
    A = {count, CPT2_degree, RCCI_degree, substrateChoice,
treatmentChoice, ATPprod, nadh2_flux, LCFAox_flux, pmtcrn_flux};
end
% Use xlswrite to generate Excel file that is openable by xlsread
xlswrite(filename, A)
end
References


Isackson, P. J., Bennett, M. J., & Vladutiu, G. D. (2006). Identification of 16 new disease-causing mutations in the CPT2 gene resulting in carnitine...


http://doi.org/10.1007/s13340-012-0104-9

http://doi.org/10.1371/journal.pone.0119936