Systems biology of inborn errors of metabolism

Swagatika Sahoo

Faculty of Faculty of Life and Environmental Sciences
University of Iceland
2013

SYSTEMS BIOLOGY OF INBORN ERRORS OF METABOLISM

Swagatika Sahoo

Dissertation submitted in partial fulfillment of
Philosophiae Doctor degree in Biology

Advisor
Professor Ines Thiele

Thesis Committee
Professor Ólafur S. Andrésson
Professor Ines Thiele
Professor Jón J. Jónsson

Opponents
Professor Hermann-Georg Holzhütter
Professor Barbara Bakker

Life and Environmental Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, September 2013

Systems biology of inborn errors of metabolism
Systems biology : inborn errors of metabolism
Dissertation submitted in a partial fulfillment of a Ph.D. degree in Biology

Life and Environmental Sciences
School of Engineering and Natural Sciences
University of Iceland
Sturlugata 8
101, Reykjavik, Reykjavik
Iceland

Telephone: 525 4000

Bibliographic information:
Swagatika Sahoo, 2013, Systems biology of inborn errors of metabolism, Ph.D. thesis, Faculty of Life and Environmental Sciences, University of Iceland.

ISBN 978-9935-9164-5-7

Printing: Háskólaprent, Fálkagata 2, 107 Reykjavik
Reykjavik, Iceland, September 2013

Abstract

Inborn errors of metabolism (IEMs) are the hereditary metabolic disorders often leading to life threatening conditions when left un-treated. IEMs not only demand better diagnostic methods and efficient therapeutic regimen, but also, a high level understanding of the precise biochemical pathology involved. Constraint-based metabolic network reconstruction and modeling is the core systems biology methods to analyze the complex interactions between cellular components that maintain cellular homeostasis. Simultaneously, the global human metabolic networks, Recon 1 and Recon 2, are landmarks in this regard. The work presented in this thesis discusses the combined bottom-up and module approaches to expand and refine the mother networks. It also highlights the importance of transport proteins in network reconstruction, their associated properties, and their involvement in various IEMs. Thereafter, we focused on expansion of Recon 1 with acylcarnitine metabolism, via the Ac/FAO module, which enabled mapping of newborn screening data, facilitating the use of *in silico* models in IEMs diagnosis. Additionally, a compendium of IEMs was assembled for interrogation of 235 IEMs using this expanded network. IEMs are mainly treated by specific diet and medication. The impact of diet and IEMs on cellular metabolism was explored by manually building a metabolic network of small intestinal epithelial cells (sIEC), taking into account the correct representation of the biochemical, anatomical and physiological attributes of the human small intestine. Limiting constituent of specific diets that deranged the metabolic capability of sIEC was revealed. Subjecting the sIEC to particular diet and IEM, lead to further understanding of the biochemical relatedness to the clinical picture of IEMs, identification of novel comorbid patterns, as well as, studying cellular adaptive mechanisms employed by cells to bypass the metabolic block. The role of medications was analyzed through a drug module; further expanding Recon 2 to account for the metabolism of commonly used drugs. Administering a defined combination of diet and drug under IEMs, their effect on drug metabolism and elimination was studied. Further, statin-associated myopathy in mitochondrial energy disorders was found to be linked to hindrance of statin metabolism. The statin-cyclosporine interaction in IEMs cases was predicted to be related to common metabolic and transport proteins involved in their metabolism. Finally, we observed that the cellular energy level being compromised in order to enable elimination of drugs, like antihypertensives, analgesics and midazolam. This signifies the potential of metabolic modeling for biomedical applications, and the presented work can be extended to predict better therapy in terms of diet and medications, as well as their combination to combat various enzyme deficiencies.

Útdráttur

Meðfæddir efnaskiptagallar eru genagallar sem oft leiða til lífshættulegs ástands séu þeir ekki meðhöndlaðir. Meðfæddir efnaskiptagallar krefjast betri greininga og áhrifaríkrari meðferða, auk ríks skilnings á þeim lífefnafræðilegu meinafræðum sem þeir eiga aðild að. Endursmíði og líkanagerðir efnaskiptaneta byggðar á takmörkunum eru grundvallaraðferðir kerfislíffræðinnar til að greina flókin samskipti á milli þeirra frumuhluta sem viðhalda samvægi í frumunni. Efnaskiptanetin Recon 1 og Recon 2 eru kennileiti að þessu leyti. Hér verða ræddar grasrótar- og einingaaðferðir til að auka þekkingu og hreinsa til í þessum móðurnetum. Einnig verður fjallað um mikilvægi flutningspróteina í smíði efnaskiptaneta, eiginleika þeirra og aðild að ýmsum meðfæddum efnaskiptagöllum. Þar á eftir jukum við út Recon 1 með asýlkarnitín efnaskiptum í gegn um Ac/FAO eininguna. Þetta gerði okkur kleift að kortlegga skimunargögn úr nýfæddum börnum og leiddi til auðveldunar á notkun *in silico* reiknilíkana við greiningu á meðfæddum efnaskiptagöllum. Að lokum settum við saman yfirlit yfir meðfædda efnaskiptagalla, til rannsókna á 235 göllum með notkun þessa stækkaða nets. Meðfæddir efnaskiptagallar eru meðhöndlaðir með sérstöku mataræði og lyfjameðferð. Áhrif mataræðis og meðfæddra efnaskiptagalla á efnaskipti voru könnuð með því búa handvirkt til efnaskiptanet af þekjufrumum í smáþörmunum (sIEC). Líffæra-, lífefna- og lífeðlisfræðilegir eiginleikar smáþarma í mönnum voru tekin til greina við byggingu líkansins.Í ljós kom takmarkandi þáttur sérstaks mataræðis sem truflaði efnaskiptahæfni sIEC. Líkanið var prófað með ákveðnu mataræði og meðfæddum efnaskiptagallum, og leiddi það til frekari skilnings á lífefnafræðilegum tengslum við klíníska mynd meðfæddra efnaskiptagalla, greiningu á nýjum samsláandi mynstrum, auk rannsókna á aðlagandi ferlum sem frumurnar beita til að komast framhjá efnaskiptatálmum. Hlutverk lyfja var greint með lyfjaeiningu, og var Recon 2 þar með stækkað til að innihalda efnaskipti algengra lyfja. Áhrif mataræðis og meðfæddra efnaskiptagalla á efnaskipti og brottnám lyfja voru rannsökuð með því að gefa inn skilgreinda blöndu mataræðis og lyfja. Að auki var statín-tengdur vöðvakvilli í hvatberaorku röskunum tengdur við truflun á statín efnaskiptum. Statín-sýklósporín samskipti í meðfæddum efnaskipta tilfellum voru spáð fyrir að vera tengd algengum efnaskipta og flutningsprótínum sem koma við sögu í efnaskiptum þeirra. Að lokum sáum við að orkustig frumunnar var skert svo fruman gæti virkjað brotthvarf lyfja; háþrýstingslyfja, verkjalyfja og midazolam. Allt þetta styður mögleika efnaskiptalíkana í líflæknisfræði. Þessa rannsókn má útvíkka til að spá betur fyrir um meðferðir vegan skorts á ýmsum ensímum.

To Dr. Arjun Sahoo, Jayanti Sahoo, Niharika Senapati, Sagarika Sahoo, Madhulika Sahoo, Aditya Senapati, Rudransh Muduli, Ratna Prakash, and Saiwal Topno

Contents

List of Figures

List of Tables

List of abbreviations

Abbreviation	Full Name
IEM	Inborn errors of metabolism
NBS	Newborn screening
ACMG	American College of Medical Genetics
COBRA	Constraint-based reconstruction and analysis
GPR	Gene-protein-reaction association
GENRE	Genome scale network reconstruction
FBA	Flux balance analysis
FVA	Flux variability analysis
MS/MS	Tandem mass spectrometry
RBC	Red blood cell
ECG	Electrocardiogram
EEG	Electroencephalogram
Ac/FAO	Acylcarnitine/fatty acid oxidation module
GI	Gastrointestinal
sIEC	Small Intestinal epithelial cells
GABA	Gamma-Aminobutyric acid
TCA	Tricarboxylic acid cycle
HMP	Hexose monophosphate shunt
OMIM	Online mendelian inheritance in man
GeneID	Official gene identification number as per EntrezGene database
BiGG	Biochemical genetic and genomic knowledgebase
E.C	Enzyme commission number
SGLT	Sodium glucose co-transporters
GLUT	Facilitated glucose transporters
SLC	Solute carrier genes
ABC	ATP binding cassette transporter
CNT	Concentrative nucleoside transporter
ENT	Equilibrative nucleoside transporter
FSV	Fat soluble vitamin
OAT	Organic anion transporter
OATP	Organic anion transporter polypeptide
PKU	Phenylketonuria
AKU	Alkaptonuria
IBD	Inflammatory bowel disease

MMA	Methylmalonicacidemia
SLOS	Smith-Lemli-Opitz syndrome
DRI	Dietary reference intake
UDP	Uridine-di-phosphate
UGT	UDP-glucuronyltransferase
ADME	Absorption, distribution, metabolism, and excretion
CYP	Cytochrome-P450

Acknowledgements

I thank the Icelandic Research Fund (No.100406022) that funded my doctorate studies at the Center for Systems Biology, University of Iceland.

My sincere gratitude to Dr. Ines Thiele, my advisor, who gave me the vision to study and practice science with a completely new dimension. Her timely suggestion, the belief in me, and words of motivation helped me explore new frontiers in scientific research.

The scientific environment at the Center, and each person attached to the lab, has contributed in my growth both as a researcher and as an individual. The shared laughter, coffee breaks, and cake time gave me a platform to share and discuss the group's viewpoints, as well as their inputs in further improvements. My special thanks to Jon, Leifur, Steina, and Sigi for their constant support and encouragement.

Above all, I thank my Family for being there for me and sending me cross-borders blessings to make this thesis happen.

1 Introduction

1.1 Short introduction on IEMs, classification, data acquisition

Metabolism remains a common factor for a myriad of human diseases. While a whole majority of them are related to diet and lifestyle patterns (e.g., diabetes, obesity, atherosclerosis), there exists very high proportion that are caused due to specific mutations in the genome. Such disorders that have a genetic base are called inborn errors of metabolism (IEMs). An error or defect in a metabolic pathway can be either due to missing enzyme(s) or deficiency of the enzyme function(s). However even if the enzyme/protein is present at optimum concentrations, the transport system might fail, or the regulation of the metabolic pathway may be disrupted, all of which may be associated with an IEM. IEMs were first described by Sir Archibald Garrod in 1908, during the famous Croonian lectures on the inborn errors of metabolism, i.e., alkaptonuira, cystinuria, albinism and pentosuria [1]. Following this, the world of IEMs has been ever expanding, both in terms of diagnosing new affected phenotypes and in their therapeutic measures. They have also been coined different names, e.g., inherited metabolic disorders, hereditary metabolic disorders, or congenital metabolic diseases. For simplicity we shall use IEMs.

Although there exist various forms of classification scheme for IEMs, they can also be classified according to pathophysiology as disorders (i) giving rise to intoxication, (ii) involving energy metabolism, (iii) involving complex molecules [2] (Figure 1.1A). The first group of disorders results in accumulation of the toxic metabolite, proximal to the metabolic block, and include aminoacidopathies (phenylketonuria, maple syrup urine disease etc.), organic acidemias (methylmalonic-acidemia, isovaleric acidemia etc..), urea cycle disorders, porphyrias, mineral metabolism disorders (Menkes disease, Wilsons disease etc.), sugar intolerances (galactosemias, fructose intolerance etc.), as well as synthesis defects of neurotransmitters (GABA, glycine etc.) and amino acids (serine, glutamine, proline etc.). The diagnosis of this class of IEMs includes simple tests, such as blood and urine amino acid, organic acid and acylcarnitine profiling. Additionally, this group is often amenable to treatment either by specific dietary patterns or medications. The second group comprises mitochondrial and cytoplasmic energy defects, e.g., pyruvate dehydrogenase deficiency, alpha-ketoglutarate dehydrogenase deficiency, mitochondrial electron trans-

port chain disorders, etc.. The cytoplasmic energy defects as compared to the mitochondrial ones are less severe, and include pyruvate kinase deficiency, glucose-6-phosphate dehydrogenase deficiency, transketolase deficiency, etc.. Mostly untreatable, this group of disorders is identified by enzyme analysis, tissue biopsies or molecular testing. Interestingly, the first group of IEMs does not interfere with embryo-fetal development, in comparison to the second group that causes congenital malformations and dysmorphism. The third group is the most diverse and includes lysosomal storage disorders, peroxisomal biogenesis disorders, congenital disorders of glycosylation, cholesterol synthesis defects etc.. Usually the diagnosis involves molecular testing methods. Majority of the disorders of the third group are difficult to treat, and mostly rely on enzyme replacement therapy (lysosomal storage disorders) [2].

'Being individually rare, but collectively numerous' [2], the current incidence of IEMs stand at 1:800 live births [3]. They can be present at any age, from infancy (e.g., peroxisome biogenesis disorders) to adult (e.g., late onset forms of pyruvate dehydrogenase deficiency), and can affect either individual (for e.g. primary hyperoxaluria type 2 without end-stage renal disease affects only kidney) or multiple organs (e.g., lysinuric protein intolerance affects liver, spleen, muscle, kidney, pancreas, gastrointestinal system, respiratory system, nervous system, and immune system). However, in ~ 50% of patients with IEMs, the symptom free period is >1 year that extends into adulthood [2]. Depending on the severity of the disorder, they usually affect several organs. The various diagnostic tests are shown in Figure 1.1B, where the newborn screening (NBS) program is routinely performed. NBS depends on estimation of specific analyte or their ratios, usually called biomarkers, and thereby, aims for an early and rapid detection of IEMs, in order to reduce infant mortality and morbidity. With the advent of tandem mass spectrometry (MS/MS), NBS saw a tremendous rise in the detection of IEMs in infants, as this could rapidly detect several analytes. NBS includes the process of sample collection, identification of an affected infant, commencement of treatment procedures, and finally the overall outcome [2]. Usually, blood spots (>1 ml of whole blood sample) are collected when the infant is between 24-48 hours old, by a heel-prick test and then sent for MS/MS analysis, where the precision of the quantitative results depends highly on the sample used (i.e., paper sample less precise than the plasma or serum sample) [2]. NBS currently covers estimation of amino acids, amino acid ratios, acylcarnitines, and acylcarnitine ratios for 64 different conditions (20 primary conditions in the ACMG panel, 21 secondary, eight maternal and 15 other conditions) that is usually adopted worldwide via collaborative projects [4].

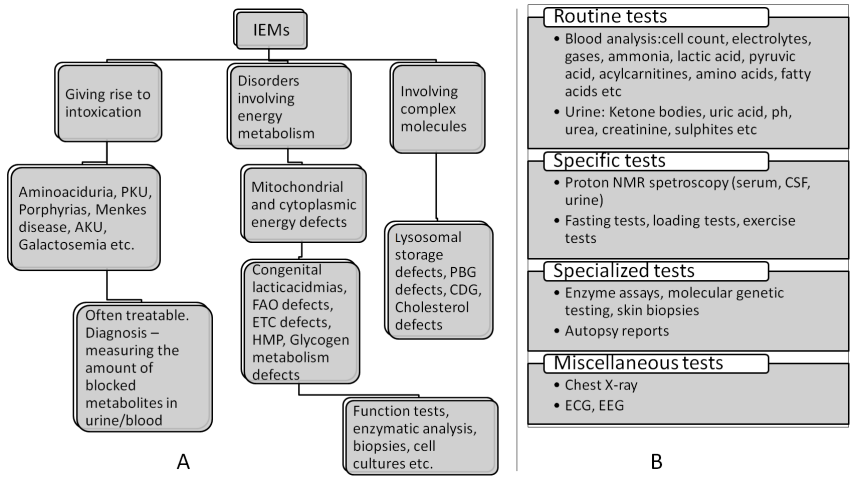


Figure 1.1: IEMs classification and diagnostics. A. Classification scheme of IEMs. B. Diagnostic tests for IEMs. (Both the figures have been adapted and modified from [2]).

1.2 Systems biology and constraint-based modeling

Living cells comprise different metabolic (e.g., metabolites) and non-metabolic (e.g., genes and proteins) components (Figure 1.2). Systems biology deals with studying the complex interactions or links between the different cellular components, mainly with a holistic view. However, contrary to the component view, the systems view takes into account of the state of the whole system, rather than of individual component [5]. The publication of whole genome sequence of the bacterium *Haemophilus influenzae Rd.* [6], and later the first whole human genome sequence [7], made it possible to identify all the genes and their gene products, ultimately meaning that important physiological processes involving these components can now be studied, and the wealth of information be harnessed.

A cell is typically under various biological constraints, i.e., the need to grow and divide, interacting with the environment, running metabolic processes etc. [5]. The intracellular environment further controls the functional state of the biological network [5]. Hence, the concept of constraints existed since the beginning of life forms. The various constraints can be categorized into: (i) physiochemical constraints that are hard constraints, and include conservation of mass, energy, momentum, free energy change etc., (ii) topological constraints that constrain the structural entity of cell, e.g., definitive arrangement of DNA within the nucleus, (iii) environmental constraints those are time dependent and include nutrient availability of a cell, pH control, temperature, osmolarity maintenance etc., and (iv) regulatory constraints

that mainly involve evolutionary changes [5]. The regulatory constraints are otherwise the non-metabolic ones, controlling ultimately the metabolic events (Figure 1.2B). These constraints together aid in proper functioning of a cell type.

The mathematical formulation of biological constraints is mainly in the form of balances and bounds. Balances are equality constraints and bounds are inequalities [5]. A typical example for balance constraint is the conservation of mass. This particularly hold true for a steady state, where there is no accumulation or depletion of metabolites/compounds, indicating that rate of synthesis equals the rate of degradation for each metabolite in the network [5]. Mathematical representation is as $S.v = 0$, where S is the stoichiometric matrix ($m \times n$, m is the number of metabolites, and n is the number of reactions in the network) and v is the flux through each reaction in the network). Such balances can also be formulated for osmotic pressure, electroneutrality, and free energy around biochemical loops [5]. In contrast, in case of assigning bounds (usually done for concentration, flux or kinetic constants), a range is given. Typical example include those done for flux values for every single reaction in a metabolic network, such that $V_{min} \leq V \leq V_{max}$, irreversible reactions being constrained as $V_{min} = 0$ and $V_{max} = \infty$ [5].

Application of these constraints will result in an underdetermined systems of linear equations [8] necessitating the use of constraint-based analysis methods to decipher the network properties. There exist a number of such methods discussed in [9] and [8]. These methods otherwise known as constraint-based reconstruction and analysis (COBRA) methods can be categorized as (i) biased methods and (ii) unbiased methods. Biased methods depend on an objective function (usually denoted by z , Figure 1.2C). In case of microbial cells, the objective function is usually the biomass reaction (i.e., fractional contribution of all known biomass constituents to the overall cellular biomass, unit being $mmol/gDW/hr$ [10]. The biased methods that are most explicitly used are flux balance analysis (FBA) [11], flux variability analysis (FVA) [12], single gene deletion [13], and phenotypic phase plane analysis (PhPP analysis) [5, 14]. The unbiased methods do not require an objective function, and include extreme pathway analysis (ExPa analysis) [15], elementary mode analysis (ElMo analysis) [16], and random sampling [17]. The majority of biased approaches, have been devised as a variation of FBA (e.g., FVA and single gene deletion analysis). Once reconstructed, biochemical networks are effectively converted into a stoichiometric matrix, $S = m \times n$ (as discussed above). Each row in S is a metabolite and each column is a reaction, and each entry in S is coefficient of the participating metabolite, where the negative coefficients represent consumption and positive means production of the metabolite in the corresponding reaction/column. Under a steady state with the given constraints, FBA seeks to either minimize or maximize an objective function Z,subject to $S.v = 0$, and $V_{min} \leq V \leq V_{max}$. While there are alternative optimal solutions, FBA outputs one of the alternative optima. On the other hand, FVA uses FBA to maximize and minimize each reaction in the

network, thereby giving a range of numerical values through which the flux can vary (i.e., $V_{i,min}$ and $V_{i,max}$). Single gene deletion analysis constrains the reactions that are associated with the specified gene to be inactive, i.e., $lb = ub = 0$, and then uses FBA to compute optimal flux values for the wild type and knock out strains/models. However, only those reactions are turned off that are associated with either single gene or with 'and' relationship in the gene-protein-reaction associations (GPRs), as described below.

1.3 Human metabolic network reconstruction in general

Genome scale metabolic network reconstruction (GENRE) is the heart of systems biology methods to elucidate physiological properties of an organism. While >100 organism specific GENREs have been published, most notably the human specific GENREs shall be the focus point of our discussion. The first global reconstruction of the human metabolic network, Recon 1 [18] was published in 2007. It contained 3744 reactions, 2744 metabolites distributed over seven intracellular (i.e., cytoplasm, mitochondrion, golgi apparatus, endoplasmic reticulum, lysosome, peroxisome, and nucleus) and one extracellular space, and 1496 genes. This network was a generic reconstruction representing a biochemical transformation that could occur in any cell of the human body (i.e., not cell/tissue specific). Since it was based on >50 years of cumulated biochemical scientific literature and the human genome annotation data, it effectively captured the sub-cellular localization of the metabolic and transport reactions. Due to extensive manual curation of the bibliome, this network became one of the most successfully validated network (where 288 known metabolic functions of a human cell were tested and accomplished), and hence, served as a platform for a series of studies that benefit not only the human community but also provided useful insights about lower organisms, e.g., the mouse metabolic network (Table 1.1).

Table 1.1: A few of the important biomedical applications of Recon 1

Subsequent models that used Recon 1	Study area	References
Biomarker validation and identification	A Red blood cell model was used to validate metabolite concentration changes, and Recon 1 used for large scale prediction of biomarkers for IEMs	[19]

Subsequent models that used Recon 1	Study area	References
Neuron-specific metabolic models	Systems level understanding of Alzheimer's disease association with disturbed cholinergic and glutamatergic neurotransmission	[20]
Liver-specific metabolic network, Hepatonet1	Characterized liver specific functions (bile acid synthesis, ammonia detoxification etc.), and liver homeostasis	[21]
Kidney-specific metabolic model	Off-target effects of torcetrapib	[22]
Red blood cell model	Analyzed morbid SNPs and drug targets in the network	[23]
Joint model of human alveolar macrophage-*Mycobacterium tuberculosis*	Investigated drug targets for different stages of bacterial infections	[24]
Mouse model	Homolog gene mapping used to create a metabolic network of five mammals including mouse	[25]
Multi-tissue integration of hepatocyte, myocyte & adipocyte	Metabolic profile of diabetic and normal obese patients	[26]
Generic cancer model	Synthetic lethal pairs in renal cancer	[27]
Heart-specific model	Epistatic interactions in heart, and identified potential biomarkers for cardiovascular diseases	[28]
Compendium of IEMs	Ac/FAO module to map newborn screening data. Compendium of IEMs, topological analysis and metabolic hotspots	[29]
Small intestine-specific enterocyte metabolic network	Effect of diet and IEMs on enterocyte network, explained cellular adaptive mechanisms under IEMs, identified novel co-morbidities	[30]
Recon 2	Predicted 54 biomarkers for 49 IEMs with high accuracy, generated 65 cell-specific metabolic models using protein expression data	[31]

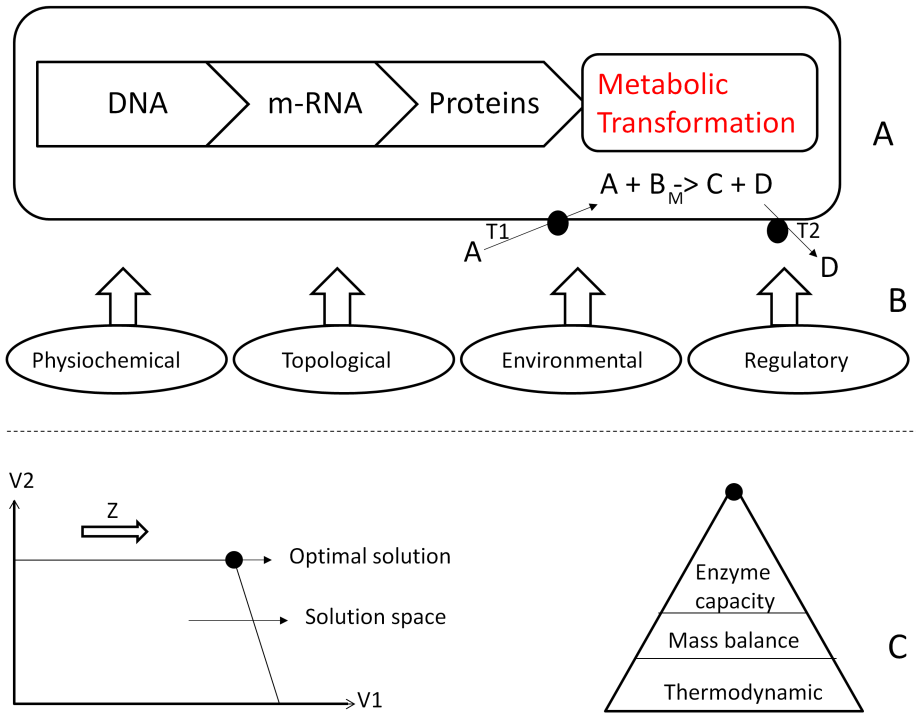


Figure 1.2: Overview of cellular components, biological constraints and formulation of mathematical constraints. A. Components of cells. The metabolic components of cell include the metabolic and transport reactions, and the non-metabolic ones are DNA, m-RNA and proteins. The central dogma of life is the flow of information from DNA to m-RNA to proteins, which catalyze the various metabolic and transport reactions operating in a cell. Building of a genome scale metabolic network reconstruction takes into account of the metabolic components. B. various constraints governing a cell. Adapted from [5]. These biological constraints are transformed into mathematical ones as bounds and balances, and applied as model constraints to obtain a solution space. C. Application of constraints to obtain a solution space, and finally the optimal solution. Sequential administration of thermodynamic, mass balance and enzyme capacity constraints, gives in a solution space. Under these constraints, an objective (z) is maximized or minimized to obtain an optimal solution point.

Later another extensive human metabolic network was reconstructed, i.e., the Edinburgh human metabolic network [32] that contained 2823 reactions, 2671 metabolites and 2322 genes, which was subsequently compartmentalized [33]. Simultaneously, liver-specific metabolic network was available, i.e., Hepatonet1 [21] that used Recon 1 as a starting point, and contained 2539 reactions and 777 metabolites. Very recently community-driven global reconstruction of human metabolism, Recon 2 [31] was created. Not only did it integrate the above mentioned networks, but also a human small intestinal erythrocytes reconstruction [30], an acylcarnitine/fatty acid oxidation module (Ac/FAO module) [29], and a transport/absorption module, all together comprising 7440 reactions, 5063 metabolites and 1789 genes (Figure 1.3A). With reaction content doubled, and 62% of the Recon 1 pathways expanded, the increase in the pathway knowledge has been evident specifically for xenobiotic metabolism, cholesterol metabolism and intracellular transporters (endoplasmic reticulum, mitochondrial, peroxisomal) to name a few. Additionally, nine new pathways were added into Recon 2 [31]. A dead-end metabolite can only be produced or consumed in a network, and a blocked reaction (usually associated with a dead-end metabolite) is the one that cannot carry any non-zero flux under steady state. Recon 2 resolved 307 dead-end metabolites and 827 blocked reactions that were existent in Recon 1 [31]. Moreover, it successfully fulfilled 354 defined metabolic tasks [31]. Hence, this recent reconstruction is an attempt to capture the most up-to date knowledge on human metabolism. Although Recon 2 is extensive and in-detailed representation of almost all the major biochemical transformations occurring in man, there are certain knowledge gaps (Figure 1.3B), which paves the road towards future investigations.

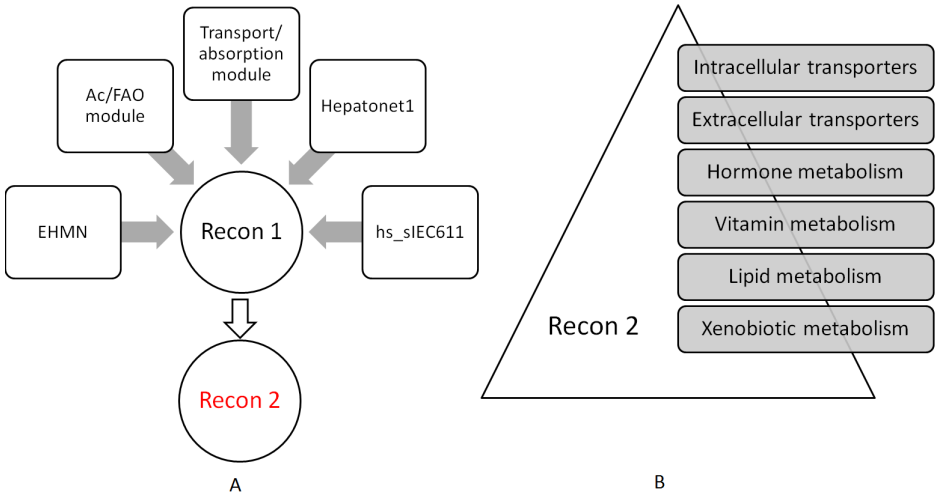


Figure 1.3: Building of Recon 2, and remaining gaps. A. Recon 2 [31] was built from Recon 1 [18], and combined all the large reconstructions available, along with small scale modules (refer text). B. Metabolic pathways in Recon 2 that require much more work for the different targeted applications. Any additions and updates should always feedback to Recon 2 to profit the community directly and immediately.

1.3.1 Manual reconstruction

The process of building a high quality GENRE is essentially iterative, laying its foundation on genome annotation data and scientific literature [10]. The basic steps remain:

(i) Obtaining a draft reconstruction using genome annotation data: Collect the list of reactions that are evident from the annotation of the corresponding metabolic genes. Usually, organism specific databases and text books can be followed simultaneously. Enzyme databases like BRENDA [34] and pathway databases like KEGG [35] are also helpful.

(ii) Refinement of the draft reconstruction: This is done by verifying the substrate and co-factor usage by the various enzymes, identifying the correct reaction stoichiometry and directionality by manually curating the scientific literature. Further, other information such as sub-cellular localization, as well as neutral and charged states of the participating metabolites needs to be added. A very unique feature contained in GENREs is the gene-protein-reaction (GPR) associations and the assignment of confidence scores to the reactions. GPRs are Boolean logics that uses 'and' association between the genes when the gene products form an enzyme complex

and all of them are necessary for the reaction to proceed (e.g., fatty acid synthase complex containing seven enzyme activities catalyzes synthesis of fatty acids from malonyl-CoA) . On the contrary 'or' association is used to represent an isozyme, where either of gene products can catalyze a single reaction (e.g., glucokinase and hexokinase can catalyze phosphorylation of glucose to glucose-6-phosphate). Confidence scoring system represents the level of evidence that is used to re-construct a corresponding reaction. This ranges from 4 to 0, where 4 means highest level of biochemical evidence exists (e.g. protein has been crystallized and its functions have been characterized), 3 means genetic information is available (e.g., knockout studies performed), 2 indicates availability of sequence data or physiological evidence of the reaction occurrence (e.g., secretion pattern in biofluids) and 0/1 indicates addition of reaction for modeling purpose only. This kind of scoring system becomes extremely useful when one has to know the weaker areas of the network, hence, furthering future research.

(iii) Conversion of the reconstruction into a computable format: This involves conversion into stoichiometric matrix, usually done within the MATLAB environment using COBRA toolbox [13]. rBioNet [36], which is a reconstruction tool can be used as it ensures quality control along with various other add-on features (e.g., dead-end metabolite identification, visualization of metabolite connectivity and plotting neighboring reactions). Thereafter, systems boundaries are defined (addition of exchange reactions), demand and sink reactions added (if required), and the appropriate constraints are set. Exchange reactions are different from transport reactions, in that they ensure availability of medium components, secretion and uptake patterns of metabolites. Demand reactions are added when the metabolic fate is unknown/out of the scope of a metabolic reconstruction, e.g., demand of nucleotide-triphosphates for DNA synthesis in rapidly growing cells. Sink reactions are used when source or synthesis reactions are unknown, but presence of the metabolite has been reported, e.g., synthesis of unsaturated fatty acids, with unknown position of double bonds. While exchange reactions can be either reversible or irreversible (depending on the constraint set), demands by rule are irreversible, and sinks are reversible.

(iv) Network evaluation: This step analyzes the functional properties of the network. Analysis of network gaps, checking for the production of biomass precursors, and comparing with other experimental data (if available) are extensively carried out. These steps have also been shown in Figure 1.4B, and further details can be followed in [10].

Notably, at every step, there is continuous iteration, improving the metabolic quality of the reconstruction. It is this iterative nature that makes the final product (i.e., a functional condition specific model) highly valuable in terms of knowledge content. Although generation of draft reconstruction is semi-automated, its refinement essentially involves an in-depth manual curation of the bibliome. The network de-

bugging also warrants revisiting the scientific literature. What is required is some prior information about the organism or the cell/tissue type of interest (Figure 1.4A).

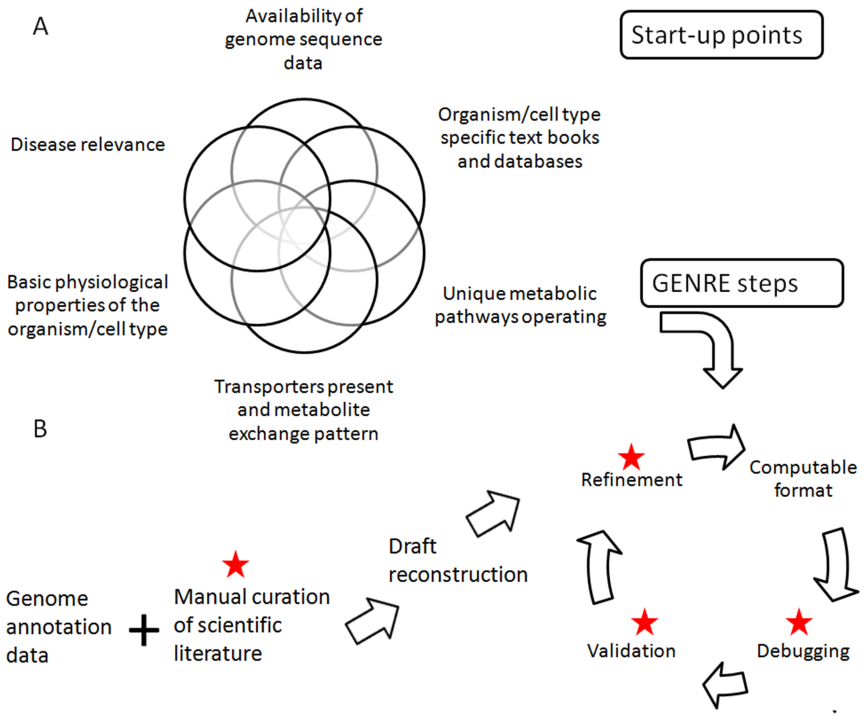


Figure 1.4: Metabolic network reconstruction as a continuously iterative process. Need of manual curation of scientific literature at crucial stages. A. Certain start-up points for an effective GENRE building. B. Basic steps of GENRE building that re continuously iterated. Note the use of manual curation in almost every step on GENRE building.

1.3.2 Check list, what are the startup points

It's very important to gather some basic knowledge about the organism/cell type of interest, before getting started with the reconstruction. Since, the reconstruction is based on genome annotation data, one should have an idea as to how much annotated information is approximately there. This is an approximation of the network size, as well as its metabolic content, e.g., a search query for metabolic genes annotated for humans, in NCBI EntrezGene [37], using the term 'human' and 'metabolic' returns ~21000 hits; and the number of unique gene contained in Recon 2 [31] is 1789, and these numbers are fairly comparable. In-depth knowledge of the major and minor physiological processes operating or maintained are related to the underlying

metabolic pathways operating within the cell, e.g., small intestine provides ~25% of body's glucose [38] and cholesterol [39] supply indicating that gluconeogenesis and denovo cholesterol synthesis pathways might be functional in this organ. Additionally, considering its role in digestion and absorption of food, the presence of specific transport proteins are indicated. Additionally, in case of building a cell type specific model these unique processes contribute to the maintenance of whole body homeostasis. This information is relatively easy to obtain. The metabolic pathways that are uniquely operating within the organism/cell type heavily depends on the transport proteins present. The transporters impart the specific metabolite exchange pattern to the organism/cell type, e.g., abundance of amino acids transporters in the small intestinal epithelial cells, the enterocytes, makes them unique in the choice of their energy source, where glutamine is the most preferred one [38]. Hence, knowledge on transport proteins is equally important. However, organism/cell type specific books serve as one of the best source of information, necessitating their rapid usage during the GENRE building, as well in its validation and elucidation of the resulting network properties. A very intriguing feature that not only helps in enhancing the applicability of the final model, but also in the initial building of the GENRE is its association with known pathological processes. Disease specific information pinpoints exactly what genes, and gene products are present in the organism/cell type, and also highlights important research work done in this regard, e.g., acute GI disturbances are caused in acute intermittent porphyria (OMIM: 176000), concluding the presence of heme metabolism in the small intestine [40]. This will further enhance the understanding of its metabolic content. But one needs to be careful while cell type specific GENRE building, when the clinical symptoms are primary (i.e., caused solely because of metabolic derangement in that particular cell type) or secondary. Such information is indispensable for building a high quality GENRE.

1.3.3 Pit falls

The reconstruction methods described above is usually called the bottom-up reconstruction approach. This is rigorous and labor intensive that requires significant manual curation of the scientific literature. Therefore, this ensures (i) reliability for its continuous usage for various biomedical applications as well it future expansion, (ii) validation and correctness of its metabolic content, and (iii) obtaining meaningful and physiologically relevant predictions by the final model. However, the amount of time invested ranges from six months to one year in generating and validating such network. Alternatively, various automated methods have been described (Model SEED, use of 'omics' data etc.) that help generate a GENRE in a considerable less time. However, such models again require significant manual curation for their validation. Additionally, these automated methods suffer from various limita-

tions, and certain steps have not been fully automated (detailed in next chapter).

In case of manual reconstruction, most often, reaction directionality is assumed to be bi-directional and localization as cytosol, whenever sufficient information isn't available. In case of transport reactions, non-availability of conclusive evidence of transport mechanism, as well ambiguity in substrate coverage by a transport protein, may occur. In such cases, based on physiological evidence of the presence of the metabolite in the exometabolome (e.g., biofluids like urine, blood CSF etc. in case of human cells), a diffusion reaction is usually added. Correct assignment of confidence score keeps a quality-control check on this, highlighting the weaker areas of a GENRE. However, there are organisms that are less studied, similarly, various cell types for which enough information isn't yet available. Under such conditions, protein and gene expression data can serve as good starting points. Such limitations in GENRE building needs to be addressed by high level biochemical experimentations, to aid in manual reconstruction for GENRE building.

However, one of the major considerations is that, one needs to be careful during manual evaluation of the scientific literature, and transferring the information, i.e., distinguishing between *in vitro* experimental data and physiologically relevant information that can vary most often. A database search should always accompany in thorough reading on the cited articles. Additionally, the reactions and pathways most often mentioned in databases, biochemistry text books or primary literature may not account for proton and water molecules, and at times co-factors and important co-substrates may be missing, which should be evaluated carefully. Nevertheless, books, primary literature, and review papers are one of the best sources for achieving finesse in assembling organism/cell specific GENREs.

1.3.4 Need for transporters

Transporters contribute significantly in deciding the metabolic profile of a cell/organism. Additionally, they (i) aid in nutrient sensing in case of enteroendocrine cells [41], (ii) regulation of the metabolic pathways, by either providing starting points (in case external source, e.g., phenylalanine being an essential amino acids, needs to be supplied via a LAT3 transport protein for the synthesis of tyrosine and its subsequent products, i.e., norepinephrine and epinephrine that have various roles in the body depending on the site of action), or co-factors (e.g., vitamins of the B-complex chain are essential co-factors for various biochemical reactions, and depend on specific transporters for their entry into cell), and (iii) generating concentration and membrane potential. Additionally, various diseases are caused due to faulty transport systems (e.g., cystinuria caused due to defect in the intestinal dibasic amino acid transporter, OMIM.220100). Knowledge on presence of specific transport pro-

teins is therefore crucial not only in GENRE building but also in making meaningful predictions by the final model. However, transporter content has been one of the highly neglected areas of most automated and manually curated GENREs, particularly for the intracellular transport proteins, e.g., models generated using Model SEED approach [42], adds reactions only into cytosol and extracellular compartments, hence, precise work is required via thorough manual curation in identifying these transporters, and adding in the missing knowledge.

1.4 Preview of the thesis

- Chapter 1: This chapter gives an introduction to IEMs (their classification, and diagnostics), constraint based modeling, and bottom-up reconstruction.

- Chapter 2: : This chapter describes the iterative nature of metabolic network reconstruction, and the need for manual curation.

- Chapter 3: This chapter highlights the need of plasma membrane transporters in metabolic network reconstruction. Assesses the transporter content of major metabolite class in recently published global human metabolic network, Recon 2, in comparison to Recon 1. Additionally, transport protein associated IEMs have been discussed. A transport module has been presented.

- Chapter 4: This chapter present the acylcarnitine/fatty acid oxidation module, mapping of newborn screening data onto the combined model, as well provides a compendium of IEMs that has been mapped onto the first global human metabolic network, Recon 1.

- Chapter 5: Diet plays key role in deciding cellular metabolic profile. Simultaneously, enzymopathies dictate the phenotype of a cell. The effect of diet and enzymopathies have been analyzed on a metabolic network of small intestinal enterocytes. The network assembly, dietary influences and effect of IEMs have been presented.

- Chapter 6: Cells call in for extra expenses in terms of energy cost, in order to metabolize drugs. Reconstruction and modeling of commonly prescribed drugs, their cellular effects, interaction profiles as well as metabolic differences under different dietary patterns and altered genetic backgrounds (i.e., IEMs) on a combined network of Recon 2-Drug module have been analyzed.

- Chapter 7: Conclusion and future perspective.

- The text of chapter 4 in full is a reprint of the material as it appears in Sahoo, S., Franzson, L., Jonsson, J. J. and Thiele, I. (2012) A compendium of inborn errors of metabolism mapped onto the human metabolic network. Molecular BioSystems, 8, 2545-2558. I was the primary author of this publication, and the co-authors participated in research guidance.

- The text of chapter 5 in full is a reprint of the material as it appears in Sahoo, S. and Thiele, I. (2013) Predicting the impact of diet and enzymopathies on human small intestinal epithelial cells. Human Molecular Genetics, 22, 2705-2722. I was the primary author of this publication, and the co-author participated in research guidance.

2 Human metabolic reconstruction as an iterative process

2.1 Introduction

The understanding of the oceanic knowledge about human metabolism remains and shall continue to fascinate cell biologists, more because of its integral connection with a number of human diseases, such as diabetes, obesity, IEMs, and cancer. To capture these intriguing facts in its fullest has been a herculean task, mainly in the form of writing traditional biochemistry books, primary scientific literature or review papers, thereafter, bestowing it to its readers. A genome scale metabolic reconstruction (GENRE) is usually based on such legacy data. Therefore, GENREs are knowledgebases and not just databases. The basic components forming the metabolic network are metabolites and reactions (also referred to as nodes and links of the network, respectively). A GENRE built for an organism is as specific as its genome [10], and intends to capture its essential genomic, physiological, and biochemical knowledge [5]. Recon 1, the first global human specific GENRE published in 2007 [18], used genome annotation data and bibliomic data (i.e., scientific literature) to extract the various biochemical pathways occurring in any human cell. The extensive manual curation of literature resulted in 3744 reactions, 2744 metabolites and 1496 genes. Five years later, the most current community-driven global reconstruction of human metabolism, Recon 2 [31] was published that not only expanded the existing metabolic content, but also filled in the knowledge gaps, which existed in its predecessor, in the form of 7440 reactions, 5063 metabolites and 1789 genes. These metabolic networks built through a bottom-up approach (i.e., manual curation efforts) are generic networks and comprise reactions occurring throughout the body, and hence, not specific to any cell/tissue type.

2.2 Biochemical knowledge is in-complete demanding daily updates

The scientific community is growing at a dynamic rate. There has been significant developments since the human genome project began. Not only new gene

functions been characterized, there are several disease mechanisms that have been explored. Due to tremendous advancements in clinical diagnostics, mainly through MS/MS, new IEMs have been reported, and to name a few, MEGDEL syndrome (OMIM:614739) [43], phosphoglucomutase 1 deficiency (OMIM:612934) [44], peroxisome biogenesis disorder caused due mutation in PEX14B (OMIM:614920) [45], and primary hyperoxaluria type III (OMIM:613616) [46]. Furthermore, well known IEMs like phenylketonuria (OMIM:261600), lysosomal storage disorders, and urea cycle disorders demand new biomarkers as well as dietary measures as treatment strategy. Hence, there is a continuous need not only to update our knowledge on a daily basis, but also, to plan distinct biomarker identification, and effective therapeutic measures.

With Recon 2 being published with significant advancements, it also points us towards certain areas of human metabolism that are yet to be explored with great details. In terms of pathways lipids, hormones, vitamins, transporters (intracellular, extracellular) needs to be curated and expanded. Recon 2 can definitely serve as a good starting point to expand these pathways for the different targeted applications.

This means that we are just at the tip of iceberg with respect to human diseases and human metabolism, and that these areas need significant work, in order to understand the precise pathology. We therefore, need a module approach (explained below) to fine tune Recon 2, so as to make it amenable for various biomedical applications.

2.2.1 Bottom-up reconstruction: iterative process, requires correction, refinement, or addition

The bottom-up reconstruction process is continuously iterative, and comprises of 96 steps detailed protocol [10]. These steps can be broadly classified into 4 basic steps : (i) obtaining the draft reconstruction, (ii) its refinement, (iii) conversion of the reconstruction into a mathematical format, and (iv) network debugging and evaluation of its functional properties. Although step two and four require much more continuous checks and frequent updates, as compared to other steps, these steps can also ask for the corrections to be made to the initial draft (e.g., gene functions in the genome annotation might have been updated and needs to be incorporated) as well as during conversion to the mathematical form (i.e., the systems boundaries, objective functions, and applied constraints may require corrections). This implies that these steps are inter-related and improvement at any step, shall ultimately improve the predictions of the final model. Recon 1 [18] was iteratively built from the genome annotation and >50 years of biochemical scientific literature data. Due to the intensive manual curation of its reaction content, it has been used as starting point

for a plethora of biomedical usages (Table 1.1), helped in gaining new biological insights, and has received >100 citations. The recently published human GENRE, Recon 2 [31], iteratively improved and resolved the dead-end metabolites as well as blocked reactions that exited in Recon 1 [18]. Additionally, it expanded 62% of the metabolic content (i.e., existing pathways were expanded) and nine pathways were added, which resulted in 77% accurate predictions for 54 metabolite biomarkers in 49 different IEMs. Additionally, by mapping of proteomics data 65 cell types specific models were generated that hold potential for various biomedical applications. Apart from the human GENREs, below we vividly explain the iterative improvements of various other metabolic networks.

An iterative process

The iterative nature of the GENRE makes it highly valuable with regards to its knowledge content. The soil bacterium, *Acinetobacter baylyi* ADP1 [47] was iteratively refined for its genes, reactions, gene-protein-reaction association (GPRs), and biomass precursors by comparison with three experimental data sets. Hence, the improved model correctly accounted for the growth phenotypes, corrected for inconsistent predictions, and was further used to infer knowledge about existence of alternative metabolic pathways. In case of the *Bacillus subtilis* network [48], network gap analysis and comparison to the phenotype data resulted in addition of missing reactions that corrected model predictions and also revealed 80 novel enzyme essentiality, which could not be provided by the genome annotation data alone. Furthermore, the draft metabolic network of the yeast, *Scheffersomyces stipitis* [49] was refined and curated sequentially using the Biolog phenotyping data, leading to addition of new genes and reactions that provided insights into its mitochondrial oxidative phosphorylation mechanisms. The continuous updates into archaeal methanogen *Methanosarcina barkeri* network [50] improved growth predictions by mutant strains and furthered our knowledge on possible regulatory events. The most recent yeast consensus reconstruction, Yeast 5 [51] refined its reaction content (particularly for sphingolipid metabolism) and GPRs that enhanced its predictive accuracy as compared to the previous versions. Finally, decade long refinements of the bacterium Escherichia coli K-12 network [52] has resulted in discovery of eight new open reading frames and new mutant strains phenotypes, novel antimicrobials aiding in drug discovery, as well as in production of important biotechnological compounds (e.g., 1,4-Butanediol).

Since, the manually built reconstruction is based on the available literature data, it requires continuous updates in the form of (i) correction, (ii) refinements, and (iii) additions. As new information becomes handy (mainly biochemical investigations) by the time the network is assembled and made available, it becomes incompetent for specific usage, yet remains important for broader and general usages. These modifications needs to be periodically incorporated, by a community effort, mainly extracting the updated knowledge from domain experts. This calls for an approach which is very precise and goes via a component by component basis, mainly serving two purposes. Firstly, it should significantly improve our knowledge level, and secondly, should feed back to the original model for the benefit of the scientific community. This shall also ensure continuous usage of the mother GENREs. We use the term 'module' to define this method (Figure 2.1). A functional module captures a specific part of the mother GENRE (mostly the network gaps), and adds in new reactions and expands existing pathways (e.g., adding newly identified gene functions). It also corrects pre-existing ones (e.g., correction in enzyme localization data, GPRs or substrate or co-factor utilization) as the knowledge becomes available. This way not only the original model is retained, but also advancements are kept abreast. Hence, both the general and specific scope of a reconstruction are fulfilled.

2.2.2 Module approach : adding missing parts sequentially, piece-by-piece

Very recently the acylcarnitine/fatty acid oxidation module, Ac/FAO module [29] was released that captured the lipid metabolism of Recon 1. The lipid metabolism was expanded in order to account for the acylcarnitines that are used as biomarkers in newborn screening programs worldwide. Hence, this module captured alpha, beta, and omega oxidation pathways of fatty acids, in the form of 352 reactions, 139 metabolites and 14 genes that were not only new, but also compatible with Recon 1. After combining with the mother network, resulting in Recon1_Ac/FAO module, enabled the mapping of newborn screening data onto this network. Similarly, an intestinal transport/absorption module was assembled using manual curation of scientific literature, resulting in 371 transport and exchange reactions, 27 metabolites and 29 genes [53]. This module has been combined with Recon 2 and well as recently built enterocyte reconstruction [30] expanding its transporter content. However, the importance on extracellular transport proteins (plasma membrane transport system) has been realized, and significant progress has been made in this regard, which led to the formulation of yet another transport module comprising of 78 reactions, 87

metabolites, and 33 new genes, which can be added to Recon 2 if desired. This highlights that such module approach adds in the missing parts sequentially to the mother network, and is essential for frequent updates and refinements.

2.3 Need for tissue/cell specific models

As the published human GENREs are generic, one needs a tissue specific network to analyze local effects on a tissue/cell type in response to various metabolic perturbations. Additionally, generic networks capture the whole body metabolism, accounting for various alternate pathways. However, the need to run a particular pathway, depends heavily on the specific conditions or specific cell type, e.g., under fasting condition, due to the absence of the enzyme CoA-transferase (E.C. 2.8.3.5), liver cannot utilize ketone bodies as an energy source, in comparison to extra-hepatic tissues like brain and skeletal muscles. In order to account for typical cell specific effects building a cell/tissue specific network in desirable. There exists a number of such studies that have used Recon 1 [18] to map various 'omics' data to generate cell-type specific metabolic reconstructions for cardiomyocyte [28, 54], hepatocyte [26, 55], alveolar macrophage [24], red blood cell [23], renal cell [22], and different cancer cells [27, 56, 57]. Additionally, interaction between different cell-types at a metabolic level, such as between brain cells [20], and between hepatocytes, myocytes, and adipocytes [26] has been modeled. However, in comparison to these that used either transcriptomics or proteomics data as starting point, the metabolic network of the liver, Hepatonet 1 [21] and that of the enterocytes of the small intestine, hs_sIEC611 [30] used the available scientific literature, in addition to Recon 1, as the starting points. Furthermore, protein expression data was used as an evidence to validate their metabolic network in part.

2.3.1 Computational methods are currently limited

In contrast to the bottom-up reconstruction approach that involves through manual curation of the available literature to build a GENRE, there exists various automated methods, where genuine data types help understand the metabolic architecture of the cell. This approach is called a top-down approach. Usually 'omics' data, namely transcriptomic, proteomic, metabolomic help specify transcripts, proteins or metabolites that make up the corresponding cell type (Figure 2.2). Although a number of other 'omics' data has been developed (i.e., genomics, glycomics, lipidomics, localizomics, epigenomics, immunomics, cytomics, fluxomics to name a few [58, 59]), we shall focus on the aforementioned three data types, due to their high usage. Each of these data types represent a separate entity of the cell, and

can be used to infer respective cellular attributes. But, significant effort has been put into using them together, in an integrated fashion, in order to acquire maximum possible as well as more meaningful information at a systems level [58]. However, these data are often noisy with technical artifacts (i.e., print tip effects, scratches, blotches, cross hybridization in case of transcriptomics [60]) making them difficult to interpret. Additionally, transcriptomic data suffer from lack of consistency due to lab to lab differences [61]. In case of proteomics, particularly the shotgun approach, peptides-proteins connectivity is lost, hence, information about the number and properties on the proteins present in the sample becomes non-available [62]. Additionally, distinction between protein isoforms becomes difficult. Artificial redundancies between protein sequences, as well as inconsistencies between assembling peptides into proteins, incompleteness of the reference database, and inaccurate characterization of a protein's mature form are related issues [62]. Similarly, various problems persist in case of interpreting metabolomics data, e.g., identification of all the metabolites in a cell/system cannot be done with any one analytical method due to chemical heterogeneity. Further, appropriate standards have not been characterized yet for all the compounds, and finally to differentiate significant metabolites from the noise [63]. However, in spite of these associated problems, 'omics' data remain the method of choice to build a cell specific reconstruction, as these generate a huge amount of information as well as within limited time span to build a reconstruction, as compared to the manual one.

2.3.2 Need for manual reconstruction

As highlighted above the usage 'omics' data presents various problems that demands their interpretation with caution. Additionally, there are various steps in the GENRE building that cannot be automated, and needs through manual curation. Particularly, the assignment of GPRs, confidence scoring system, and identification of organism or cell specific pathways, has to very precise in order to make physiologically relevant predictions, as well as to pinpoint the weak and strong part of the network. Additionally, inconsistent and incorrect information in the genome annotation data, missing enzyme functionalities, and transporter specificities require detailed manual curation [64]. Recently developed Model Seed [42] platform can automatically generate a draft reconstruction based on the genome sequence data. However, addition of intracellular transport reactions has not been automated yet. Moreover, it requires extensive manual curation of the generated biomass reaction, electron transport chain reactions so as to ensure correct electron sources. Identification of futile cycles, cofactor utilization, reaction localization, and validation of the model output with experimental data again needs to be done thoroughly via manual curation.

Furthermore, manual curation has resulted in correcting and refining the final metabolic models in many cases. During building of Recon 1 [18], genome annotation data resulted in 1134 genes, out of which 731 genes were discarded, due to wrong information. Additionally, literature mining revealed 362 genes to be added to the final network. In case of Hepatonet1 [21], the model that was obtained solely from transcriptomic data alone, was able to achieve a steady state flux distribution in only 71 of 119 simulations, thereby necessitating a labor-intensive manual curation effort to validate 442 metabolic objectives that was performed by this *in silico* network as well as by human hepatocytes in general. Furthermore, during the enterocyte reconstruction [30], while evaluating for a functional biomass reaction, literature mining led to the addition of phosphatidyl-choline biosynthetic pathway that was overlooked during initial steps. A series of enzyme localization and reaction directionality corrections were made to the initial set of reactions of lipid metabolism obtained from Recon 1 (e.g., enzymes involved in the cholesterol synthesis pathway), as well as in-depth manual curation of the relevant literature revealed set of pathways that should be operating (e.g., urea cycle, gluconeogenesis from glutamine, fructose metabolism, re-synthesis of triacylglycerol etc.), partially operative (e.g., glycine, serine branched chain amino acid metabolism etc.) and should not be operating (lysine, threonine metabolism etc.) in the small intestinal enterocytes (details in [30]). Additionally, prior to model building of the small intestine, an in-depth study was performed on the anatomy of the organ and major blood supply to the organ that permitted us to set the exchange reactions or system boundaries accordingly. As mentioned in the introduction chapter, regarding the start-up points, maximum information was obtained for the various transport proteins present in this cell type, which helped in setting specific metabolite exchange pattern and also served as one of the basis for including the operational metabolic pathways.

The in silicon models generated via 'omics' data also demands in-depth manual curation of scientific literature for model validation. This highlights the importance of manual curation in GENRE and condition specific model building. There is no alternative to an in-depth manual curation of scientific literature. Therefore, the need of the time is to build a junction between bottom-up and top-down approaches, so as to build a genuine and reliable network within limited time period. Additionally, manual curation can not only improve but also act as a quality control for the 'omics' data and its interpretation (Figure 2.2). One should also keep in mind that the various 'omics' data are a snapshot of the cellular metabolism under specified conditions, and does not necessarily capture the whole cell metabolism. Contrastingly, manual curation although laborious shall help acquire knowledge about the entire tissue type in general. An amalgam between automated and manual methods guarantees continuous usage of metabolic networks, e.g., Recon 1 and Hepatonet 1 have served start up points for various biomedical studies [65–67](in addition to the above).

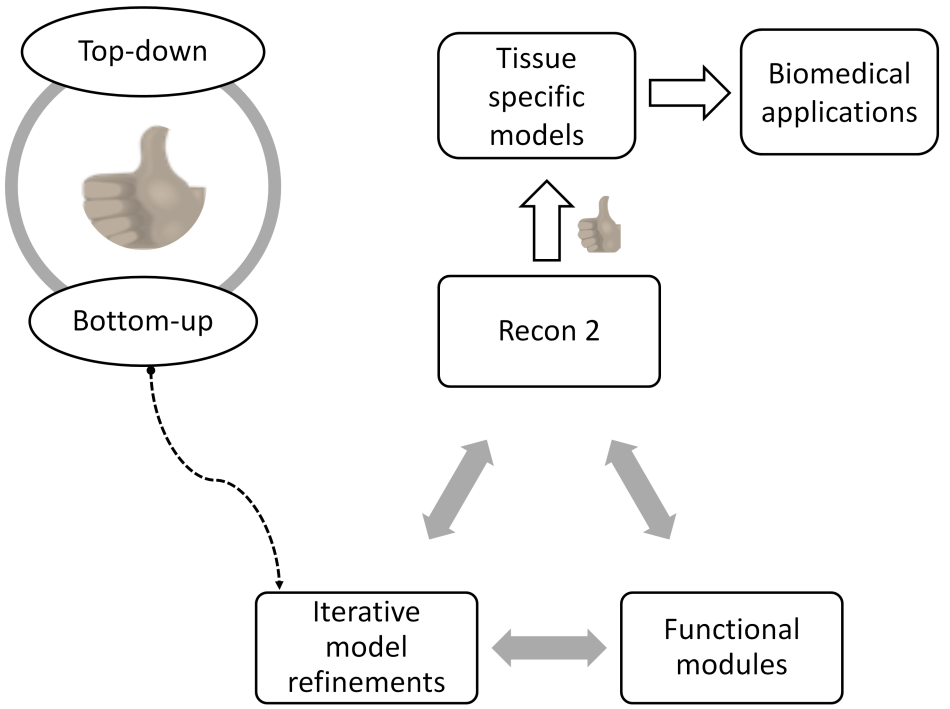


Figure 2.1: Reconciliation between top-down and bottom-up reconstruction approaches could help in generation of physiologically and functionally relevant tissue specific models, which could be generated from a mother reconstruction, like Recon 2. These tissue-specific models should aid in various biomedical applications. Further, iterative model refinements capturing additions, corrections and updates should result in functional modules that should refer back to the mother network, and needs to be periodically incorporated. These modules should capture the weaker sections of the mother network, and curate and expand them either via bottom-up approach or reconciliation between bottom-up and top-down approaches. The reconciliations are signified by a thumbs-up sign.

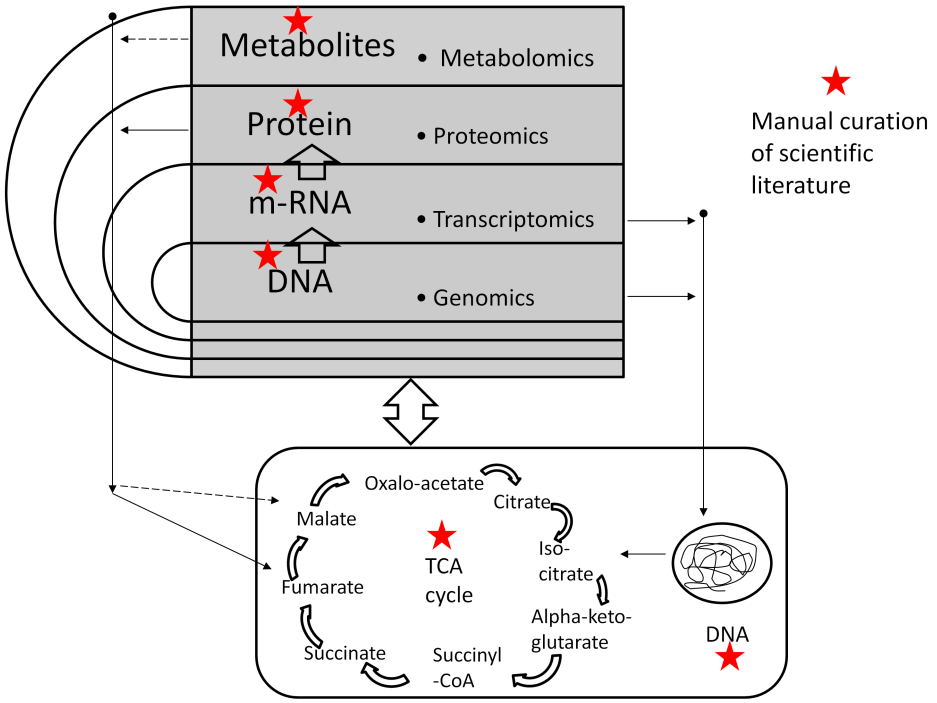


Figure 2.2: Need for manual curation in usage of 'omics' or automated methods. The various 'omics' data types have significantly improved our knowledge about underlying cellular processes. However, usage and interpretation of these data types demands extensive manual curation when used to build a tissue-specific GENRE.

3 An assessment of membrane transporters in human metabolism: Coverage in the human genome-scale metabolic knowledgebase and implications in diseases

3.1 Abstract

Membrane transporters enable efficient operation of cellular metabolism, aid in nutrient sensing, and have been associated with various metabolic diseases, such as obesity, inborn errors of metabolism, and cancer. Genome scale metabolic network reconstructions capture genomic, physiological, and biochemical knowledge of the target organism, along with detailed representation of the cellular metabolite transport mechanisms. Since the first reconstruction of human metabolism, Recon 1, was published in 2007, progress has been made in the field of metabolite transport. Recently, we published an updated reconstruction, Recon 2, which significantly improved the metabolic coverage and functionality. The human metabolic reconstructions have been used to investigate metabolism in the disease context and to predict biomarkers and potential drug targets. Given the importance of cellular transport systems in understanding human metabolism in health and disease as well as for the generation of cell-type specific metabolic models, we analyze the coverage of transport systems for various metabolite classes in these two metabolic reconstructions. Moreover, we discuss the current knowledge about transporters for these metabolite classes based on extensive literature review, assess missing coverage, and propose modifications and additions. This includes an overview of the preferred substrates, transport mechanisms, metabolic relevance, organ distribution, and disease association. This information will be valuable to further refine the human metabolic reconstruction and to provide starting points for further experiments by highlighting incomplete knowledge. This review represents the first comprehensive overview of transporters and their transport mechanism involved in central metabolism and thus can serve as a compendium of metabolite transporters.

3.2 Introduction

Membrane transporters mediate transport of solutes from the outside to the inside of a cell and vice versa. Transport processes generate concentration gradients and membrane potential as well as contribute to regulation of biochemical pathways by maintaining cellular levels of substrates and products (e.g., GLUT proteins regulate availability of glucose). It has been approximated that 2000 genes encoding for transporters or transport related proteins are present in the human genome [68]. Defective metabolite transport processes have been associated with various pathological conditions, including inborn errors of metabolism (IEMs) [69–71], obesity and related abnormalities [72], and cancer [73, 74]. Hence, knowledge on cellular transport systems is foundational in understanding human metabolism.

Genome-scale metabolic reconstructions (GENREs) integrate the genomic, physiological, and biochemical knowledge of a target organism [5]. They are knowledge-bases of metabolites, biochemical transformations/reactions, enzymes catalyzing the reactions, and genes that encode these enzymes. GENREs are available for more than 100 organisms, including human [18, 31]. GENREs can be easily converted into mathematical models and used for constraint-based modeling and analysis (COBRA), e.g., flux balance analysis [11]. Details on the procedures for GENRE and COBRA modeling are discussed elsewhere [5,10,13]. The in-depth biochemical knowledge captured by GENREs is mostly attributed to the gene-protein-reaction associations (GPRs), where individual metabolic/transport reactions are represented along with the genes that encode enzyme/protein catalyzing the reactions. These GPRs are Boolean relationships between genes/transcripts and reactions with 'and' and 'or' association. 'And' indicates that the expression of all genes, is necessary for the reaction(s) to be active (e.g., multi-enzyme complexes catalyzing a single reaction). An 'or' relationship implies that any of the genes or gene products can catalyze the reaction (e.g., isozymes catalyzing the same reaction).

The first GENRE for human metabolism, Recon 1 [18], captures biochemical transformations occurring in any human cell in a stoichiometrically accurate manner, which are distributed over seven intracellular compartments (i.e., cytoplasm, mitochondrion, Golgi apparatus, endoplasmic reticulum, lysosome, peroxisome, and nucleus). Additionally, Recon 1 includes a representation of the extracellular space to account for the exchanges and transport systems connecting the extracellular space with the cytoplasm. The most recent community driven global reconstruction of human metabolism, Recon 2 [31], is a substantial expansion over Recon 1 and includes more than 370 additional transport and exchange reactions. These global reconstructions do not represent the metabolic capability of a single cell or tissue but rather are blue prints for all human cells, similar to the human genome, which encodes for all cellular functions that may be active in one or more cell-types and

conditions.

A reconstruction represents the metabolic repertoire of an organism or cell in a condition-independent manner and can give rise to multiple condition-specific metabolic models. Consequently, recent COBRA modeling efforts have focused on generating cell-type specific metabolic models using cell-type and condition-specific data (e.g., transcriptomic and metabolomic data). Cell-type specific reconstructions have been assembled for cardiomyocyte [28, 54], hepatocyte [21, 26, 55], alveolar macrophage [24], red blood cell [23], renal cell [22], enterocytes of the small intestine [30], and different cancer cells [56, 57]. Additionally, metabolic interactions between different cell-types, such as between brain cells [20], and between hepatocytes, myocytes, and adipocytes [26] have been modeled.

The generation of cell- and tissue-specific reconstructions requires extensive knowledge about metabolites that can be transported across the plasma membrane. Transporters connect cells and tissues with their immediate environment and thus, can be used for defining, which intracellular metabolic pathways must be active to fulfill the chief functions of the cell or tissue. At the same time, transport reactions are amongst the less well studied reactions captured in Recon 1 [18]. This property is retained in Recon 2 despite of substantial efforts to include more transport information. This is mostly because precise function and mechanism cannot always be predicted from sequence alone for transporter encoding genes. As transporters across the plasma membrane are generally better studied than intracellular transporter, we focus this review on plasma membrane transporters.

The aim of the present review is to highlight, how extensive our current knowledge about plasma membrane transporters is and how well it is captured in Recon 2. We therefore introduce first general transport mechanisms and then discuss the transport of five major metabolite classes (i.e., sugar, amino acids, lipids, vitamins, and others). Ions are also discussed while not being metabolites as they are important co-substrates for many transport systems. At the end, we discuss the importance of transporter in different metabolic diseases. This review is accompanied with a transport module containing 70 new transport reactions, which can be added to Recon 2. The module does not account for modifications of GPRs that we report to be necessary to capture current knowledge. This is because new genes need to be added to the GPRs of existing, mechanistically correct transport reactions of Recon 2.

This review does not represent an update of Recon 2 but rather an expansion of its coverage of plasma membrane transport. The module-based approach permits researchers to actively contribute to the expansion of human metabolism, while maintaining Recon 2 as core GENRE.

3.3 General transport mechanisms

The cell membrane separates a cell from the extracellular environment. There are two basic modes of cellular transport, i.e., active and passive. These basic mechanisms can operate: 1) without a carrier protein (simple diffusion), 2) involving a carrier protein (facilitated diffusion), and 3) with expenditure of energy (primary and secondary active). The various modes of transport are shown in Figure 3.1.

The cell membrane's lipophilicity defines metabolites that can freely move in and out of the cell, a process called "simple diffusion". While hydrophobic substances can easily cross the membrane, hydrophilic cannot [75]. Simple diffusion is directed from a region of high to a region of low solute concentration. Various factors determine the net rate of diffusion, including concentration difference of the solute, pressure difference between the cell and the environment, membrane electric potential, and osmosis [75].

Transport processes allowing the passage of a single solute at a time are referred to as uniport. "Facilitated diffusion" is an example for uniport [76]. In facilitated diffusion (also called carrier diffusion), a cargo molecule itself causes a conformational change in the carrier protein, which opens a channel for the cargo to cross the cell membrane. The capacity of this transport mechanism is thereby limited by the time needed to change the conformation back and forth. Facilitated diffusion occurs in both directions [75].

Active transport is required to move molecules against their concentration and electrochemical gradient. It requires energy in form of ATP or other high energy phosphate bond [75], which is either directly connected to the transport (primary active transport) or derived from an electrochemical gradient (secondary active transport) [77, 78]. When a secondary active transport process is further coupled to another distinct exchange mechanism, it is referred to as tertiary active transport. One example is the coupling of amino acid transport system A (SNAT2) with system L (LAT1/4f2hc) for uptake of leucine. At the same time, SNAT2 utilizes the electrochemical gradient established by the Na^+/K^+ ATPase pump to drive its substrate inside the cell [79].

Symport is the transport of multiple solutes across the cell membrane at once. If inward transport of one solute is connected to the outward transport of another solute or vice versa, the process is referred to as antiport [77].

3.4 Overview of extracellular transport reactions in human GENREs

Recon 1 and Recon 2 are based on manually assembled biochemical knowledge, and their reactions are extensively annotated with literature evidence (Figure 3.2A). They contain 537 and 1537 extracellular transport reactions, respectively. The majority (89%) of these reactions in Recon 2 was supported by literature evidence. However, the availability of literature coverage differs among metabolites, which were grouped into ten classes. The 'blood' group comprises mostly of glycolipids. The group 'others' contains metabolites, such as urea, water, heme, bilirubin, bile acids, and protein (e.g., albumin and fibrinogen), which could not be grouped into the other classes as well as intermediate metabolites of amino acid, carbohydrate, and lipid metabolism (e.g., creatine, glycerate, propionate, and acetate). The amino acid class has the highest number of transport reactions associated, many of which have supporting evidence associated (Figure 3.2B). In contrast, literature support for transport reactions in the group 'others', 'xenobiotics', and 'hormone' was low. Such confidence gaps arise during the reconstruction process when information on the transport mechanism is not available, yet physiological evidence for transport of a metabolite across the cell membrane has been reported or suggested (e.g., by body fluid or exo-metabolomic data). In such cases, corresponding diffusion reactions are added to the reconstruction [10]. This lack of information regarding carrier proteins and mechanisms explains the high number of diffusion reactions for lipophilic metabolites falling into the group deemed 'others', 'hormones', and 'lipids' in Recon 2 (Figure 3.2C). These gaps need to be filled as the corresponding knowledge becomes available.

3.4.1 Transport of sugars

Carbohydrates form a major part of the human diet. Polysaccharides, such as starch, are broken down into simple sugars in the intestinal lumen. Glucose, galactose, and fructose are the chief monosaccharides absorbed by the enterocytes. From the enterocytes, sugars are released into the portal blood. Two main groups of sugar transporters exist, sodium/glucose co-transporters (SGLTs) and facilitated glucose transporters (GLUTs). Both groups are encoded by solute carrier (SLC) genes, where 'SLC' is the initial for their official gene symbols (see Table 3.1). The SGLT (SGLT-1 to SGLT-6) family of transporters transport sugars coupled with sodium ions (secondary active transport). In contrast, GLUT (GLUT-1 to GLUT-14) transporters mediate facilitated diffusion [80–82]. A number of additional cell types express both transporters for uptake and secretion of simple sugars (see below).

SGLT transporters

SGLT-1 (*SLC5A1*, GeneID: 6523) is located at the apical side of enterocytes and renal tubules, and mediates influx of glucose and galactose via two sodium ions coupled symport [81, 83]. The transporter is energized by the sodium gradient established by the Na/K-ATPase located at basolateral surface [84]. However, under conditions of decreased luminal pH, proton coupled glucose transport by SGLT-1 has been observed [85] but the affinity of the transporter for glucose is reduced under these circumstances [85]. SGLT-2 (*SLC5A2*, GeneID: 6524) mediated sugar re-absorption has been described for renal cells. SGLT-2 showed low affinity but high capacity for glucose and galactose transport [86]. SGLT-3, also called SAAT1 (*SLC5A4*, GeneID: 6527), is expressed in cholinergic neurons, small intestinal and skeletal muscle cells. So far, the only confirmed function of SGLT-3 in humans is that it acts as glucose sensor by depolarization of the membrane in presence of high glucose [87]. However, in C. elegans, a single amino acid substitution (from glutamate to glutamine) enabled the protein to behave as glucose transporter with transport properties similar to SGLT-1 [88]. Additionally, recent reports claim SGLT-3 to be a glucose-stimulated Na^+ transporter [89]. SGLT-4 (*SLC5A9*, GeneID: 200010) is a sodium-dependent mannose transporter, which also showed affinity for 1,5-anhydro-glucitol and fructose [90]. It is highly expressed in small intestine and kidney, but moderately in liver [90]. SGLT-5 (*SLC5A10*, GeneID: 125206) mediates sodium-dependent uptake of sugars with highest affinity for mannose, followed by fructose, and very low affinity for glucose and galactose [91]. It is expressed in the kidney. SGLT-6 (*SLC5A11*, GeneID: 115584) is a glucose and myo-inositol transporter. Compared to other tissues, SGLT-6 is most highly expressed in the brain [92]. Sodium myo-inositol co-transporter, SMIT (*SLC5A3*, GeneID: 6526) is a myo-inositol transporter, which is ubiquitously and most prominently expressed in cells lining blood vessels, the kidney, and the thyroid gland [92]. For a more elaborate tissue distribution of all SGLTs, see recent review by Wright et al [93].

GLUT transporters

The GLUT transporters have traditionally been divided into three families on the basis of sequence similarity [94, 95]. Wilson-O'Brien et al. proposed a division of the mammalian GLUT transporters into five distinct classes, subdividing class three proteins into three new classes (i.e., GLUT-6 & GLUT-8, GLUT-10 & GLUT-12, and H+/myo-inositol transporter (HMIT) [96]. Most of the GLUT transporters have been shown to transport glucose, with GLUT-2, GLUT-5, and GLUT-11 also transporting fructose, while HMIT also transports myo-inositol (Table 3.1). Transport of non-carbohydrate substrates by GLUT transporters has also been reported.

For example, uric acid transport by GLUT-9 [97] and dehydroascorbate transport by GLUT-1 to GLUT-4 as well as GLUT-8 [98, 99]. Only GLUT-4, GLUT-8, and GLUT-12 have been shown so far to exhibit insulin sensitivity [81]. Most of the GLUT transporters are expressed in the brain, which depends largely on glucose as energy source [100, 101].

Sugar transporters and Recon 2

Transport of sugars is well studied and generally well captured by Recon 2. In total, 80% (60/75 reactions) of the sugar transport reactions in Recon 2 are supported by literature evidence (Figure 3.2B). While the reactions for the SGLT-1 mediated transport were added in Recon 2 [83], transport functions of the other SGLTs are still missing (Table 3.1). For instance, while SGLT-2 mediated glucose transport is captured in Recon 2, but its galactose transport capability [86] is not accounted for. Moreover, SGLT-5 is only associated with glucose transport, but not its additional substrates mannose, fructose, and galactose [91]. Finally, SGLT-6 is only associated with inositol but not glucose transport [92]. The transport of galactose, mannose, fructose, and glucose already exist in Recon 2 with correct transport mechanism; hence, only new genes have to be added to the GPRs of the corresponding reactions.

3.4.2 Transport of amino acids

Ingested proteins represent the main source of amino acids and peptides. Usually multiple amino acids have the same transport protein (Table 3.2). Transport systems of both amino acids and peptides are discussed in this section.

Amino acid transporters

Eleven different SLC families are known to be involved in transport of amino acids either via antiport or symport [102]. There is considerable overlap in their substrate specificity (Table 3.2). Transporters of the SLC3 and SLC7 family form heteromeric protein complexes comprising of heavy and light sub-units (where SLC3 genes encode the heavy sub-unit and SLC7 genes encode the light subunit of the transport protein) that interact through disulphide bridges. The heavy sub-unit is a glycosylated membrane protein; hence, these transporters are also called glyco-protein associated amino acid transporters. The light sub-unit is required to confer specific amino acid transport activity [103]. While the light sub-unit is fully functional even in the absence of the heavy sub-unit, these heteromeric amino acid transporters

are usually amino acid exchangers (antiports) [103]. However, the alanine-serine-cysteine transporter (a.k.a. Asc-type amino acid transporter 1), a heteromeric amino acid transporter whose heavy sub-unit is encoded by *SLC3A2* (GeneID: 6520), and the light sub-unit by *SLC7A10* (GeneID: 56301), transports glycine, L- and D-alanine, L- and D-serine, threonine, and cysteine. This transport can be mediated either via facilitated diffusion or antiport. The antiport is the predominating mechanism [104]. The transporter acts Na^+/Cl^- independent and is found in brain, heart, placenta, skeletal muscle, and kidney [105]. In mice, this transporter was additionally identified in cells of the lung and small intestine [105].

Other members of the SLC7 family (SLC7A1-A4) are non-heteromeric proteins and are cationic amino acid transporters. The amino acids arginine, lysine, and ornithine are transported in a sodium independent manner [106]. While SLC7A1 is ubiquitously expressed, SLC7A2 transporter (*SLC7A2*, GeneID: 6542) is found in skeletal muscle, placenta, ovary, liver, pancreas, kidney, and heart [107]. SLC7A3 transporter (*SLC7A3*, GeneID: 84889) is expressed in brain, thymus, uterus, testis, and mammary gland [108]. SLC7A4 transport protein (*SLC7A4*, GeneID: 6545) has no confirmed transporter activity, however, binding to other sub-units in order to confer transport activity has been postulated [109].

Peptide transporters

Four peptide transport proteins have been identified, being PEPT-1 (*SLC15A1*, GeneID: 6564), PEPT-2 (*SLC15A2*, GeneID: 6565), hPHT1 (*SLC15A4*, GeneID: 121260), and hPHT2 (*SLC15A3*, GeneID: 51296). PEPT-1 and PEPT-2 are well characterized proton symporters transporting 400 distinct di-peptides, 8,000 tri-peptides, and synthetically formulated drugs [110], for which they have overlapping substrate specificities. Generally, PEPT-2 exhibits higher affinity for di- and tri-peptides than PEPT-1 [111, 112].

Peptide transporters are usually expressed on the apical side of polarized cells. While PEPT-1 is highly expressed in small intestine, PEPT-2 in renal cells [111]. All four peptide transport proteins were identified in the nasal epithelium [113]. Both PEPT-1 and PEPT-2 work via secondary active transport coupled to Na^+/H^+ exchange, where sodium ions are exported out of the cell via basolateral Na^+/K^+ ATPase pump [114] in order to maintain extra-cellular sodium concentration. The entire process is supported by intra-cellular peptide hydrolyzing enzymes. Peptides that escape hydrolysis are transported out of the cell via an uncharacterized basolateral peptide transporter [115] that can mediate either facilitative transport [116] or proton mediated transport [117].

PEPT-1 is regulated both by hormones and its substrates. Besides peptides, it is activated by various amino acids, including lysine, arginine, and phenylalanine. Some hormones, such as insulin, can activate the basolateral peptide transporter, while others, such as leptin, epidermal growth factor, and the thyroid hormone, inhibit the apical uptake of peptides by this transporter [110]. On the other hand, PEPT-2 activity was enhanced by minerals, such as copper, zinc, and manganese [114]. Interestingly, PEPT-1, when expressed in enteroendocrine cells, is involved in hormone secretion and thus, participates in nutrient sensing [41]. Additionally, increased expression of this transport protein has been associated with various inflammatory conditions, such as ulcerative colitis and Crohn's disease, as it is not expressed in colonocytes under normal physiological conditions [118].

Amino acid and peptide transport systems and Recon 2

Amino acid and peptide transport systems are well described in the literature and in Recon 2. The 667 amino acid transporter reactions make up the largest group of extracellular transport reactions in Recon 2, and 98% of them are supported by literature evidence (Figure 3.2B). Recon 2 already covers recent additions and modification to the amino acid transport systems, which were identified during the reconstruction of the small intestinal epithelial cell [30] and of the liver [21].

However, Recon 2 still need to be extended to account for current knowledge (Table 3.2). i) Missing transported substrates need to be added and our module provides the corresponding transport reactions. For example, the renal cell specific transport systems SNAT4 (*SLC38A4*, GeneID: 55089) also transports cysteine and methionine [102, 119]. Also, the transporters B(0)AT2, B(0)AT3, and PAT4 are missing along with the transport reactions for their substrates. ii) GPRs of existing transport reactions need to be expanded. The function of PAT2 can be added by expanding the GPRs of the appropriate reactions in Recon 2. iii) In the case of PEPT-2 transporter, Recon 2 captures correctly its transport of the di-peptide Cys-Gly but other substrates are missing such as the tri-peptide Trp-Trp-Trp [114], However, the metabolic fate of these missing substrates is currently not captured in Recon 2 and thus the addition of the transport reactions would create gaps. Hence, we did not include these reactions in the transport module. For a detailed list on endogenous and xenobiotic substrates for PEPT-2, refer to [112].

3.4.3 Transport systems for lipids

Lipids are essential for many biological processes. Major dietary lipids are triacylglycerol, phospholipids, and sterols. These dietary lipids are broken down to free fatty acids, mono-acylglycerols, and cholesterol, which are subsequently absorbed by cells [83]. Cholesterol and phospholipids are essential bio-membrane constituents. Additionally, phospholipids and glycolipids form lung surfactants. Fat is stored within cells as triacylglycerols, which, when broken down to fatty acids via beta-oxidation, provide energy for various cellular processes. Lipids also act as precursors for second messengers, and cholesterol specifically acts as precursor for steroid hormones (with sterol nucleus in their chemical structure, e.g., progesterone and testosterone) and bile acids [120]. Eicosanoids are biologically active components that include prostaglandins, thromboxane, leukortienes, and lipoxins.

Due to their hydrophobic properties, the majority of lipids can freely diffuse across the cell membrane. Nevertheless, a number of alternative transport mechanisms exist: 1) fatty acid transport proteins: FATP1 (*SLC27A1*, GeneID: 376497), FATP2 (*SLC27A2*, GeneID: 11001), FATP3 (*SLC27A3*, GeneID: 11000), FATP4 (*SLC27A4*, GeneID: 10999), FATP5 (*SLC27A5*, GeneID: 10998), FATP6 (*SLC27A6*, GeneID: 28965), 2) membrane associated fatty acid transporters, FABPpm, (*GOT2*, GeneID: 2806) and fatty acid translocase, FAT, (*CD36*, GeneID: 948), 3) ATP binding cassette transporters, 4) various lipoproteins (i.e., chylomicrons, very low density lipoprotein, low density lipoprotein, and high density lipoprotein), and 5) intracellular lipid transporters, such as non-specific lipid-transfer protein, (*SCP2*, GeneID: 6342), acyl CoA binding protein (*DBI*, GeneID: 1622), fatty acid binding proteins/ cytoplasmic fatty acid binding proteins, i.e., FABPc (FABP1-9) [121, 122], and various proteins that aid in membrane turnover via insertion of new lipids into pre-existing membranes (i.e., flippase, floppase, scramblase, and flip-flop) [123, 124]. The presence of multiple additional transport mechanisms besides diffusion is explained by the essential role of lipids in the cell and the need to control their transport and distribution. Additionally, structural differences of fatty acids, mono-acylglycerol, and cholesterol necessitate distinct transport systems

Fatty acid transport

Fatty acid transport proteins (FATPs) are a family of six transporters that mediate the influx of long chain fatty acids (>10 carbons in chain length), usually associated with a long chain fatty acid activating enzyme present on the membrane (acyl-CoA synthetases, E.C. 6.2.1.3) [125]. FATPs have also been suggested to possess inherent fatty acid activating properties [126] and also bear an AMP binding mo-

tif [127]. The membrane associated fatty acid transporters (FABPpm) also transport long chain fatty acids, although, compared to FATPs, FABPpms have a higher affinity towards long chain poly-unsaturated fatty acids and essential fatty acids [128]. The FABPpm transport mechanism is slightly different from the one of FATP. Once brought inside the cell by FABPpm, fatty acids bind to the cytoplasmic counterpart (FABPc) and undergo activation [127]. They might then be transported to other sub-cellular compartments by FABPc [129]. The fatty acid translocase protein (FAT, *CD36*, GeneID: 948) is usually expressed at low fatty acid concentrations. It binds and concentrates fatty acids at the cell surface and enhances their diffusion across the membrane. Alternatively, FAT can also deliver fatty acids to FABPpm [127]. It has a wide substrate coverage including low and high density lipoproteins and phospholipids [130, 131]. Interestingly, individual transport capacities (i.e., without any concertation) of FAT and FABPpm (including FATP) have been shown in the rat skeletal muscles for palmitate (C16:0), and that these transport proteins also play a significant role in fatty acid oxidation [132]. Therefore, [133] proposed that the need for association between FAT and FABPpm would arise during conditions of increased fatty acid oxidation in order to meet increased substrate demands.

Various ABC transport proteins transport fatty acids, cholesterol, phospholipids, and cholesterol derivatives (bile acids). ABC transporters generally conduct primary active transport, act as ion channels for chloride, or regulate the function of ATP-sensitive potassium channels [134]. These transport proteins have a wide substrate spectrum, including drugs, lipid metabolites, hormones, heme, iron, peptides, nucleosides, and vitamins. A number of the ABC transport proteins are functional monomers, while most other transport proteins require dimerization or binding of other proteins to gain complete functionality (e.g., ABCB2/TAP1, ABCB3/TAP2, four transporters of ABCD sub-family, and five transporters of ABCG sub-family) [134]. The group of ATP binding cassette transporters comprises of 48 transport proteins, categorized into six different families. Of the total number of transport proteins, 32 are located on the plasma membrane, 13 are intracellular transport proteins (ABCA2, ABCB2, ABCB7-10, ABCC6, ABCD1-4, ABCG1, and ABCG4), while only three transport proteins act at both plasma membrane and intracellular locations. These are the ATP-binding cassette sub-family B member 6, (*ABCB6*, GeneID: 10058) located in the plasma membrane, in the golgi apparatus, and in lysosomes; the ATP-binding cassette sub-family A member 1, (*ABCA1*, GeneID: 19) located in the plasma membrane and in the golgi apparatus,; and the ATP-binding cassette sub-family B member 5 (*ABCB5*, GeneID: 340273), whose location remains to be identified [135, 136].

Lipid digestion products, such as triacylglycerols, cholesterol, cholesterol-ester, and phospholipids, are hydrophobic and need to be transported within lipoproteins that represent specialized transport systems for lipids, distinct from membrane transporters. Lipoproteins are spherical components, containing a hydrophobic lipid core and hydrophilic amino acid side chains on the surface [137]. Lipoproteins vary in their apolipoprotein (Apo) content, density, and lipid components. Chylomicrons are the largest lipoproteins, have the least density (i.e., <1.006 g/ml), and carry the highest fraction of triacylglycerols [137]. They are formed in the endoplasmic reticulum of the small intestine, carrying the lipid components of the diet into the lymph from where they enter the blood via the left subclavian vein [137]. When passing through the blood capillaries, lipoprotein lipase (LPL, GeneID: 4023, E.C. 3.1.1.34) extracts the free fatty acids and releases them into muscle and adipose tissues. The liver takes up the remnant chylomicrons, where the excess fatty acids may be used to synthesize triacylglycerols, which are further transported into tissues via very low density lipoprotein (VLDL). Post removal of triglycerides, the unused VLDL or VLDL remnants, which are intermediate density lipoproteins, are then either reabsorbed into the liver or form low density lipoprotein (LDL). The small intestine and liver also form precursors for high density lipoprotein (HDL) and release them into circulation. The HDL transport is also called reverse cholesterol transport. The major components transported by the four lipoprotein classes are i) triacylglycerol by chylomicrons, ii) phospholipids and triacylglycerol by VLDL, iii and iv) cholesteryl esters and phospholipids by LDL and HDL [137]. Additionally, all fat soluble vitamins (vitamin A, D, E, and K) are also transported within chylomicrons, passing from the cells into the lymph [138]. The cellular uptake of cholesterol is also mediated by Niemann-Pick C1-like protein 1 (*NPC1L1*, GeneID: 29881) and SRB-I (*SCARB1*, GeneID: 949), where the latter takes up cholesterol from the HDL protein [138, 139]. In contrast, in the case of polarized cells, luminal efflux occurs through ATP-binding cassette sub-family G member 5/ ATP-binding cassette sub-family G member 8, ABCG5/ABCG8 (*ABCG5, ABCG8*, GeneID: 64240 & 64241), and basolateral efflux is done by ATP-binding cassette sub-family A member 1, ABCA1 (*ABCA1*, GeneID: 19) [139]. The substrate specificities of these proteins have not been entirely resolved and the exact transport mechanism needs further experimental support. According to current understanding, NPC1L1 is a uniport and is recycled through the endocytic route [138]. Bi-directional transport has been indicated for SRB-I [138, 139]. Still, for the majority of lipid transporters (including fat soluble vitamins), the precise transport mechanism with respect to directionality, coupled ions or other compounds, and substrate stoichiometry remains uncertain [138].

In Recon 2, the majority of reactions associated with lipid transport are simple diffusion reactions (91 of 183 reactions, Figure 3.2C) and 11% of reactions are not supported by literature evidence (Figure 3.2B). The substrate coverage of the existing FATPs was increased in Recon 2 with the addition of long chain fatty acid transport reactions (Table 3.3). However, FAT (*CD36*, GeneID: 948) and FABPpm (*GOT2*, GeneID: 2806) are still missing in Recon 2, and are captured in transport module.

Of all ABC transport proteins located at the plasma membrane, only seven transporters are captured in Recon 2. Moreover, a number of substrates are missing, which include poly-unsaturated fatty acids, xenobiotics, nucleosides, nucleotides, and ions. The transport module accounts for the missing transporters (ABCA3, ABCA4, ABCA8, ABCC11, and ABCG2). Additionally, it expands the substrate coverage for ABCA1 transporter with phsophatidyl-choline and phosphatidyl-serine. Further, the cholesterol transport proteins ABCG5 and ABCG8 are missing in Recon 2 and can be added by expanding GPRs of corresponding reactions in Recon 2.

3.4.4 Transport system for nucleosides

The liver is the major organ for the de novo synthesis of all four nucleotides [140]. Another resource of nucleic acids is ingestion and digestion. Nucleosides and nucleic acid bases are transported across biological membranes by concentrative nucleoside transporters (CNT) [141], equilibrative nucleoside transporters (ENT) [141], and transporters of the ABC transport family [142, 143].

Three CNTs exist, each with distinct substrate specificity: CNT1 (*SLC28A1*, GeneID: 9154) shows high affinity for pyrimidine nucleosides (e.g., cytidine, thymidine), CNT2 (*SLC28A2*, GeneID: 9153) prefers purine nucleosides (e.g., adenosine, guanosine), and CNT3 (*SLC28A3*, GeneID: 64078) exhibits wide substrate specificity [141, 144]. CNTs mediate sodium coupled secondary active symport. Recent findings revealed coupling of CNT3 either with Na^+ or H^+ [145]. The four existing ENTs belong to the SLC29A gene family and exhibit wide substrate specificity [141, 144], including nucleic acid bases. ENT1-ENT3 mediate facilitated diffusion or uniport in a sodium independent manner. ENT4 (*SLC29A4*, GeneID: 222962) works by secondary active transport coupled to H^+ [145]. In the case of enterocytes and renal cells, the CNTs are located at the apical surface mediating uptake of nucleosides, while ENTs mediate both influx and efflux at the basolateral side [144]. Both

CNTs and ENTs transport a wide range of pharmacologically important drugs [145]. The transporters of the ABC transport family that transport nucleotides and nucleosides are multidrug resistance-associated protein 4 (*ABCC4*, GeneID: 10257), multidrug resistance-associated protein 5 (*ABCC5*, GeneID: 10057), ATP-binding cassette sub-family C member 11 (*ABCC11*, GeneID: 85320), and ATP-binding cassette sub-family G member 2 (*ABCG2*, GeneID: 9429) [142, 143].

Nucleoside transporters and Recon 2

Nucleoside transport is well established. Accordingly, 88% of nucleoside transport reactions in Recon 2 are supported by literature evidence (Figure 3.2B). While only the reactions of the CNTs and ENT1-ENT3 are present in Recon 2, ENT4 (*SLC29A4*, GeneID: 222962) [146] can be added with the transport module. Transport via CNT3 is associated with either sodium or H^+ gradients [145]. However, only sodium coupled secondary active transport is captured in Recon 2. The CNT3 mediated H+ coupled transport is covered in the transport module. Additionally, gene information for ABC transport proteins, i.e., ATP-binding cassette sub-family C member 11 (*ABCC11*, GeneID: 85320) and ATP-binding cassette sub-family G member 2 (*ABCG2*, GeneID: 9429) need to be added to the corresponding Recon 2 reactions. Another useful, future addition with respect to disease directed research and application of human GENREs could be transport of nucleotide-derived drugs. Refer to [145](Molina-Arcas et al., 2008) for a list of drugs that are transported by these transporters.

3.4.5 Transport system for vitamins

Vitamins are not synthesized by the human body and are therefore essential components of the human diet. Some vitamins, such as biotin, are also synthesized by the commensal gut microflora [120]. Niacin can be synthesized in the body from amino acid tryptophan [120]. Vitamins have traditionally been divided into two groups, 1) fat soluble vitamins comprising of vitamin A, D, E, and K, and 2) water soluble vitamins comprising of vitamins of the B complex, being thiamin (B1), riboflavin (B2), niacin (B3), pyridoxal (B6), folic acid, cobalamin (B12), pantothenic acid, biotin, and vitamin C. Vitamins play a major role in a variety of biochemical processes. Vitamin A is involved in the visual cycle, vitamin D in calcium metabolism, vitamin E in free radical scavenging, vitamin K in blood clotting, B1 in carbohydrate metabolism and nerve conduction, B2 and B3 in redox reactions, B6 in transamination reactions, folic acid and cobalamin in one carbon metabolism, pantothenic acid in fatty acid metabolism, biotin in carboxylation reactions and vita-

min C in hydroxylation reactions [120]. Deficiency of any of the 13 vitamins is thus associated with various diseases [120]. The body´s inability to synthesize vitamins in combination with their pivotal role in metabolic processes necessitates transport mechanisms other than simple diffusion for their import into a particular cell type as well as efflux for utilization by other cell types. Epithelial cells of the small intestine, kidney, and liver express the vast majority of ABC and SLC transporters, since these organs play chief role in absorption and secretion of endogenous metabolites and xenobiotics [68, 147]. Epithelial cells are among the best models to study transport processes. Hence, we focus on transport proteins present in the enterocytes of the small intestine.

Transport of fat soluble vitamins (FSV)

FSVs were believed to enter enterocytes from the intestinal lumen via passive diffusion. However, transport proteins for vitamins A, D, and E have been identified, and energy dependent transport has been suggested for vitamin K in rats [148].

Vitamin A (bound to retinol binding protein) uptake into the enterocyte is mediated by retinoic acid gene 6 protein homolog (STRA6) protein (*STRA6*, GeneID: 64220) [149]. Additionally, transporter of ABC family, i.e., retinal-specific ATP-binding cassette transporter (*ABCA4*, GeneID: 24), is involved in the transport of retinaldehyde, a form of vitamin A, into retinal photoreceptors [150]. Retinol-binding protein 1 (*RBP1*, GeneID: 5947) and retinol-binding protein 2 (*RBP2*, GeneID: 5948) aid in apical uptake, esterification, and secretion of retinol [151]. So far, no basolateral transport protein for vitamin A has been identified [138]. Basolateral efflux of retinol (in the form of retinyl-esters) has been assumed to primarily occur via chylomicrons.

Vitamin D transport mainly occurs in combination with vitamin D-binding protein, which has higher affinity for 25-hydroxy vitamin D than for vitamin D3 [152]. Once bound, the complex is recruited by megalin (*LRP2*, GeneID: 4036) and cubilin (*CUBN*, GeneID: 8029) either for hydroxylation or efflux [153]. Additionally, SR-BI (*SCARB1*, GeneID: 949), FAT (*CD36*, GeneID: 948) and NPC1L1 (*NPC1L1*, GeneID: 29881) are involved in the intestinal uptake of vitamin D [154]. The SR-BI protein has been shown to play a role both in apical uptake and basolateral efflux of vitamin D and E in caco-2 cells [138]. Moreover, vitamin D is transported within chylomicrons into the lymph [152].

Vitamin E uptake appears to be similar to that of cholesterol. Although passive diffusion has been observed, additional receptor mediated transport exists, mediated by SRB-I and NPC1L1 [155,156]. Basolateral efflux occurs via ABC family protein

ABCA1 (*ABCA1*, GeneID: 19) [157].

Transport of water soluble vitamins

Since the low concentration of water soluble vitamins make simple diffusion highly unlikely, distinctive carrier dependent transporters exist at both apical and basolateral sides of the enterocyte to mediate vitamin exchange

Vitamin B1 Three transport proteins are associated with transport of vitamin B1, and its structural analogs, i.e., ThTr1 (*SLC19A2*, GeneID: 10560), ThTr2 (*SLC19A3*, GeneID: 80704), and RFT (*SLC19A1*, GeneID: 6573). While ThTr1 and ThTr2 can transport free thiamin, RFT transports mono- and di-phosphate forms of thiamin [158, 159]. Sub-cellular locations vary between the transporters. ThTr1 is located at both apical and basolateral membrane. In contrast, ThTr2 and RFT are localized only at the basolateral membrane. The transport of vitamin B1 occurs against both concentration and outwardly directed H+ gradient, and appears to be sodium independent, electroneutral, and pH dependent [160]. The directionality or reversibility of these transport processes, however, remains to be elucidated.

Vitamin B2 The uptake and secretion of vitamin B2 from the enterocyte involves primary active transport [161–163] mediated by three transport proteins, RFT1 (*SLC52A1*, GeneID: 55065), RFT2 (*SLC52A3*, GeneID: 113278), and RFT3 (*SLC52A2*, GeneID: 79581). RFT1 is expressed at the basolateral membrane and RFT2 at the apical membrane, leading to vitamin B2 efflux and uptake respectively [163]. Both transporters are expressed in the small intestine [164]. RFT3 is specifically expressed in the brain [164].

Vitamin B3 Cellular uptake of niacin (a.k.a. vitamin B3 or nicotinic acid) can occur by simple diffusion [152]. Additionally, niacin is taken up via sodium independent, temperature and acidic pH dependent facilitated diffusion [165]. No specific carrier protein has been so far identified for vitamin B3. Yet, SMCT1 (*SLC5A8*, GeneID: 160728), principally an iodide transporter, has been suggested to mediate sodium coupled niacin transport [166]. The mechanism of basolateral efflux of niacin and the carrier protein involved is unknown [159].

Vitamin B6 Vitamin B6 freely diffuses across the cell membrane [152]. However, carrier dependent transport (sodium independent but pH, energy, and temper-

ature dependent) has also been suggested [167]. No specific transport protein has been characterized at the molecular level.

Folate (B9) Folate plays a role in various biochemical processes (e.g., DNA synthesis, one carbon metabolism, and amino acid metabolism), in the prevention of congenital abnormalities (e.g., neural tube defects, urogenital abnormalities, cardiovascular malformations, cleft lip, and palate) and in the treatment of cardiovascular diseases [168–171]. Given its general importance, it is not surprising that multiple folate transporters exist. Traditionally, two folate carriers, i.e., the reduced folate carrier (*SLC19A1*, GeneID: 6573) and the proton-coupled folate transporter (*SLC46A1*, GeneID: 113235) exist. Additionally, three high affinity folate receptors, a number of ABC transporters, and members of the solute carrier organic anion transporter family have been associated with the transport of folate or its derivatives. Each of which will be discussed in the following paragraphs.

The reduced folate carrier is an organic anion antiporter that utilizes high transmembrane organic phosphate gradient. It is expressed at the distal part of the small intestine, where it operates at neutral pH [159]. The proton-coupled folate transporter mediates the transport of folic acid, 5-methyl- and formyl- tetrahydrofolates, is localized at the proximal part of the small intestine, operates at acidic pH [159, 172], and has been shown to also transport heme [172].

The three high affinity folate receptors are FR*a* (*FOLR1*, GeneID: 2348), FR*b* (*FOLR2*, GeneID: 2350), and FR*g* (*FOLR3*, GeneID: 2352) mediate unidirectional influx of folate, whereby the entire folate-receptor complex is internalized [152, 172]. A reversible transport has also been suggested [158].

The basolateral folate transporter has not yet been characterized at the molecular level. However, presence of a specific carrier protein mediating sodium independent but pH dependent folic acid transport has been shown in rats [173]. In case of humans, 5-methyl-tetrahydrofolate has been identified in the portal blood [152], but the corresponding transport mechanisms and proteins need to be elucidated.

Seven ABC transporters expressed at the plasma membrane in different epithelial and non-epithelial cells show affinity towards folate and its derivatives [142, 172, 174]. Multidrug resistance-associated protein 1 (*ABCC1*, GeneID: 4363) and multidrug resistance-associated protein 5 (*ABCC5*, GeneID: 10057) at the basolateral side and canalicular multispecific organic anion transporter 1 (*ABCC2*, GeneID: 1244) at apical side of enterocytes. ABCC5 and ABCC2 are further expressed at the basolateral side of liver. The remaining four ABC transporters are ATP-binding cassette sub-family G member 2 (*ABCG2*, GeneID: 9429), multidrug resistance

protein 3 (*ABCB4*, GeneID: 5244), ATP-binding cassette sub-family C member 11 (*ABCC11*, GeneID: 85320), and multidrug resistance-associated protein 4 (*ABCC4*, GeneID: 10257).

Numerous solute carrier organic anion transporters (OAT) transport methotrexate (a structural analog of folic acid) and are also relevant folate transporters in liver and kidney [174, 175]. These transporters include OATP-2 (*SLCO1B1*, GeneID: 10599) in the liver (basolateral side) and OATP-1 (*SLCO1A2*, GeneID: 6579), OAT1 (*SLC22A6*, GeneID: 9356), OAT3 (*SLC22A8*, GeneID: 9376), OAT2 (*SLC22A7*, GeneID: 10864), and OAT4 (*SLC22A11*, GeneID: 55867) in the kidney [172, 176, 177]. One may refer to [172] for the precise apical/basolateral localizations of these folate transporters.

Vitamin B12 Vitamin B12 is the precursor for two coenzymes, adenosylcobalamin and methylcobalamin. Adenosylcobalamin is required for the activity of methylmalonyl CoA-mutase (E.C. 5.4.99.2), which catalyzes the conversion of methyl malonyl-CoA to succinyl-CoA. Methylcobalamin is required for the activity of methionine synthase (E.C. 2.1.1.13), which catalyzes the methylation of homocysteine to methionine [120, 178]. The absorption of cobalamin by simple diffusion along the entire small intestine accounts for 1-3% of dietary vitamin B12. Interestingly, this vitamin depends on a carrier mediated transport when administered at pharmacological doses [152].

Cobalamin is transported into the cell by cubilin-mediated absorption. Thereby, cobalamin binds to intrinsic factor (*GIF*, GeneID: 2694) building the intrinsic factor-cobalamin complex (IF-Cbl) and to two proteins called cubilin (*CUBN*, GeneID: 8029) and amnionless (*AMN*, GeneID: 81693). The latter serve as anchors for the receptor and aid cobalamin uptake. Additionally, proteins, such as megalin and receptor-associated protein, can interact with CUBN. If and how the binding of additional proteins plays a role in CUBN-mediated absorption of IF-Cbl has not been determined [179]. The protein responsible for basolateral efflux of cobalamin has not been experimentally validated. However, multidrug resistance-associated protein 1 (*ABCC1*, GeneID: 4363) has been shown to possess transport capacity for cobalamin in prokaryotes and eukaryotes in addition to mouse [180].

Pantothenic acid and biotin Pantothenic acid and biotin are absorbed at the apical membrane by a common sodium coupled symporter, SMVT (*SLC5A6*, GeneID: 8884) [152]. Additionally, SMVT transports lipoic acid and is called a multi vitamin transporter. Basolateral release of biotin is mediated by an yet uncharacterized carrier protein in sodium independent manner [181]. The mechanism

of basolateral pantothenate efflux remains to be elucidated [159].

Vitamin C Two transport proteins, the apically located SVCT1 (*SLC23A1*, GeneID: 9963) and the basolaterally located SVCT2 (*SLC23A2*, GeneID: 9962), mediate the vitamin C (a.k.a. ascorbate or L-ascorbic acid) uptake. The membrane location has been confirmed in rats [182]. SVCT2 is more ubiquitously expressed (except in lung and skeletal muscle), whereas SVCT1 is confined to intestine, liver, kidney, colon, ovary, and prostrate [183]. Both transport proteins exhibit Na^+ coupled secondary active symport (coupling ratio 2:1) [183], energized by Na^+/K^+ ATPase [152]. Ascorbate export has been assumed via volume sensitive anion channels [184]. Alternatively, intracellular vitamin C can be oxidized to dehydroascorbate, which can freely diffuse into the blood stream [83]. Three transporters of the GLUT family (GLUT-1, GLUT-3, and GLUT-4) also mediate the uptake of dehydroascorbate at the basolateral side [184]. In astrocytes, GLUT-1 facilitates entry and GLUT-3 the efflux of dehydroascorbate [185].

Vitamin transporters and Recon 2

Transport systems for water soluble vitamins have been more elaborately investigated than FSVs [138]. Consistently, FSVs transport is not well represented in Recon 2, but the transport of water soluble vitamins is fairly well captured. Overall, 74% of the vitamin transport reactions are supported by literature evidence (Figure 3.2B). However, the genes encoding for proteins transporting fat soluble vitamins, including those discussed for vitamin A, D, and E (i.e., *STRA6, ABCA4, RBP1, RBP2, LRP2, CUBN, SR-BI, and NPC1L1*), are absent in Recon 2. The transport protein encoded by the gene ABCA1 is so far only associated with cholesterol, but not with vitamin E transport (Table 3.3). The transport module accounts for the vitamin A transport by ABCA4, while the other missing genes can be added by expanding the GPRs of the respective transport reactions.

Recon 2 includes the vitamin B2 transporters, RFT1 and 2, but not RTF3, which can be added by expanding the corresponding GPRs . Recon 2 also accounts for the substrate specificity of PCFT, FOLR1, and FOLR3. The transport of folate by FOLR2, and of vitamin B3 by SMCT1 can be accounted for by expanding the GPR of the corresponding reaction. OAT1 - OAT4 mediated transport can be added, via the module, to completely capture current knowledge about folate transporters. The vitamin B12 transport proteins (i.e., intrinsic factor, cubilin, and amnionless) as well as the ATP costs of SVCT1/SVCT2 transport are already accounted for in Recon 2.

3.4.6 Transport of water, heme, and other special compounds

Water moves across biological membranes via different mechanisms. Apart from diffusing through the lipid bilayer, co-transporters in form of protein channels exist in the membrane, through which water can diffuse. The movement of water molecules through such a channel, called aquaporin, is driven by osmosis [186]. Water is also a substrate for co-transporters, such as excitatory amino acid transporter 1 EAAT1 (*SLC1A3*, GeneID: 6507), which is expressed in the brain and moves both urea and water along with glutamate [187], and for the sodium glucose co-transporter, SGLT1, which transports sodium and glucose, while causing water influx [188]. For details on the various water co-transporters, specifically those operating in the brain, one may refer to [186].

Aquaporins

Aquaporins are a family of membrane channel proteins that allow passage of water molecules, neutral molecules (e.g., urea and glycerol) and other small solutes [189]. In total, 13 members of this family have been characterized at the molecular level and they are expressed in a wide variety of tissues (abundantly in the epithelial layer of kidney, intestine, lungs, and brain) [190]. While a majority of these are expressed at the plasma membrane, aquaporin-6 (*AQP6*, GeneID: 363) and aquaporin-2 (*AQP2*, GeneID: 359) have also been localized in intracellular vesicles [191, 192]. Interestingly, these proteins have been associated with various cellular functions, including skin hydration [193], neural signal transduction [194, 195], and cell volume regulation [196]. Aquaporins are further believed to hold therapeutic potential for congestive heart failure, hypertension, glaucoma, brain swelling, epilepsy, obesity, and cancer [190, 192, 197].

Heme

Heme forms the prosthetic group of hemoglobin and other heme containing proteins, such as myoglobin, cytochrome C, tryptophan pyrrolase, and catalase [120]. Additionally, heme degradation serves as source for the essential micronutrient iron [198]. Two transport proteins have been identified for heme, the proton-coupled folate transporter (*SLC46A1*, GeneID: 113235, discussed above) and the feline leukemia virus subgroup C receptor-related protein 1 (*FLVCR1*, GeneID: 28982). These transport proteins directly transfer extracellular heme into the cell. While the proton-coupled folate transporter acts at the apical surface, the feline leukemia virus subgroup C receptor-related protein 1 is believed to have an active transport mech-

anism [199] and is localized at the basolateral surface of polarized cells [200]. The hemopexin protein directly interacts with feline leukemia virus subgroup C receptor-related protein 1, hence, increasing heme efflux, which is perceived as cellular protection against heme toxicity [201]. Heme transport can also occur via receptor mediated endocytosis, by prolow-density lipoprotein receptor-related protein 1 (*LRP1*, GeneID: 4035), which has been proposed to play a role in inflammation [202]. The ABC transporter ATP-binding cassette sub-family G member 2 (*ABCG2*, GeneID: 9429) can also transport heme [203].

Transport of conditionally essential nutrients

In addition to essential nutrients, there are certain other conditionally essential nutrients (CEN), which are usually synthesized by the body in almost sufficient amounts. However, under conditions of increased need, such as tissue injury or neonatal conditions, these nutrients need to be derived from the diet. CEN includes compounds, such as arginine, Coq10, carnitine, propionyl carnitine, taurine, lipoic acid, betaine, ribose, cysteine, chondroitin sulphate, and glutamine [204, 205]. In this section, we focus only on carnitine, taurine and betaine, as the transport of arginine, cysteine, glutamine, and ribose has already been discussed in the relevant sections above.

Carnitine Carnitine transports fatty acyl-CoAs (i.e., activated fatty acids) into mitochondria, via the carnitine shuttle system [120]. A positive effect of carnitine supplementation has been demonstrated upon neuro-regeneration in rats [206], liver cirrhosis in children [207], in obesity and associated metabolic disorders [208], in congestive heart failure [209], and various other diseases, which are reviewed in [210]. Carnitine further exerts protective effects in corneal epithelial cells preventing deleterious effects of dry eye syndrome [211]. While synthesis of carnitine occurs from methionine and lysine in liver and kidney [210], exogenous carnitine needs transporters to reach the target cells. There are two membrane transport proteins for this purpose, OCTN1 (*SLC22A4*, GeneID: 6583) and OCTN2 (*SLC22A5*, GeneID: 6584). OCTN1 transports organic cations as well as carnitine (in zwitter ion form) in a pH dependent and sodium independent manner, chiefly behaving as a proton/organic cation antiporter at the apical surface of polarized cells [212]. OCTN2 is a sodium dependent carnitine transporter that is also localized at the apical membrane mediating organic cation/carnitine exchange [213]. Additional carnitine transport proteins are the amino acid transporter $ATB^{0,+}$ (*SLC6A14*, GeneID: 11254), which we discussed above (Hatanaka et al., 2004), and CT2 (*SLC22A16*, GeneID: 85413), which is exclusively found in testis and functions in a sodium independent manner [214]. $ATB^{0,+}$ operates when OCNT2 is defective [215].

Taurine One of the end products of methionine and cysteine metabolism is taurine, which plays an important role in a number of tissues. In the brain, taurine acts as neuromodulator, neurotransmitter, and membrane stabilizer [216]. High taurine concentrations in heart and muscle tissue support contractile function and osmoregulation, and also exert antioxidant action by neutralizing hypocholorous acid and regulating mitochondrial protein synthesis in these tissues [217]. Additional evidence for the importance of this amino acid in human health suggests positive effect on growth in low birth weight infants, promotion of biliary flow, and prevention of cholestasis [218, 219]. Disruption of taurine transport causes retinal degeneration in mice [220]. Two taurine transporter exist, TAUT (*SLC6A6*, GeneID: 6533) and PAT1 (*SLC36A1*, GeneID: 206358). TAUT (*SLC6A6*, GeneID: 6533) mediates sodium and chloride ion coupled secondary active transport. The stoichiometry is 1 taurine: 2 sodium: 1 chloride, but limited transport activity has also been observed without chloride [216]. Although the transport directionality remains to be confirmed, movement of taurine through the blood-brain barrier was shown to occur from the blood into the brain [216]. The second taurine transporter PAT1 (*SLC36A1*, GeneID: 206358) operates via H+/taurine symport. This high capacity but low affinity transporter, which also transports beta-alanine in addition to taurine, is highly expressed at the apical membrane of enterocytes [221].

Betaine Betaine is another important molecule involved in methionine metabolism. Once it is synthesized from choline, it donates its methyl group to regenerate methionine from homocysteine and helps in conserving cellular methionine levels [222]. Additionally, betaine acts as an osmolyte, particularly helpful for normal physiological functions of kidney, intestinal epithelium, red blood cells, and skeletal muscles. Moreover, its protective role has been observed in heart and liver cells [222]. Na^+/Cl^- dependent secondary active betaine transport is mediated by BGT-1 (*SLC6A12*, GeneID: 6539) [223]. Another study reported Na^+ independent, passive transport in rats [222]. An alternate substrate of BGT-1 is gamma-aminobutyric acid. Details regarding the directionality of BGT-1 mediated transport remain unknown. The amino acid transporter, imino (*SLC6A20*, GeneID: 54716) also transports betaine [119].

Transport of water, heme and other special compounds in Recon 2

Recon 2 contain the aquaporin-8 (*AQP8*, GeneID: 343) and aquaporin-9 (*AQP9*, GeneID: 366) for the transport of water, urea, and lactate. Extracellular water transport also occurs in Recon 2 through simple diffusion ('H2Ot') and co-transport via SGLT-1 protein ('UREAt5'). The other aquaporins (AQP0, AQP1- AQP5, AQP7, and AQP10) need to be added. The transport module adds reactions and genes for

AQP3, AQP7, and AQP10. The remainder of the aquaporins can be accounted for by expanding the GPRs of the corresponding reactions. Recon 2 lacks the heme transporter FLVCR1 as additional biochemical experiments still need to clarify the precise transport mechanism. LRP1 can be added by GPR modification. All of the above discussed carnitine transport proteins, but CT2, are present in Recon 2. The function of CT2 is captured in the transport module. Additionally, carnitine transport mediated by amino acid transporter $ATB^{0,+}$ is missing in Recon 2 but can be accounted for through the transport module. Efficient taurine transport, via TAUT, coupled to both Na^+ and Cl^- ions is present in Recon 2. The transport reactions catalyzed by BGT-1 (i.e., 'ABUTt4(2)r' for betaine, and 'GLYBt4(2)r' for GABA) need to be corrected for the requirement of both sodium and chloride ions. Therefore, the transport module contains the improved reactions for the $ATB^{0,+}$ and BGT-1 transporters.

3.5 Transport reactions module

The transport module was assembled according to the established reconstruction process [10] using rBioNet as reconstruction tool [36]. The functionality of reactions in the module, in conjunction with Recon 2, was subsequently tested. All discussed modifications and additions are provided through a transport module, which comprises of 71 metabolites, 70 reactions, and 41 genes (including 19 newly added genes). These additional transport reactions are for amino acids (27 reactions), lipids (16 reactions), nucleotides (6 reactions), vitamins and minerals (8 reactions), hormones (6 reactions), and others (7 reactions). Additionally, 24 Recon 2 reactions need to be updated with respect to their gene-protein-reaction associations. Details of this module can be found at http://humanmetabolism.org. This module is an extension to Recon 2, which can be added to the existing reconstruction if desired.

3.6 Transport proteins associated with diseases

Transporters fulfill a broad range of functions, which go far beyond the sole movement of metabolites. In our discussion on the transport of distinct metabolite classes, many of these functions have been mentioned. Targeting specific transport proteins in order to combat disease conditions, such as cholestatic conditions [224], neurodegenerative disorders [225], cystic fibrosis [226], cancer [227, 228], cerebral ischemia [229], diabetes, and secretory diarrhoea [230], have gained considerable attention in recent years. Herein, we will focus on the discussion of transport proteins in disease groups concerning IEMs.

3.6.1 Transport proteins associated with metabolic diseases

Metabolic disorders are associated with disturbed cellular metabolism. IEMs are the hereditary metabolic disorders, caused by specific mutations in genes encoding metabolic enzymes/transporters. We have recently mapped more than 200 IEMs onto the two human GENREs [29, 31]. Of those, 14% (Table 3.4) were caused by faulty transport systems, including 22 IEMs causative genes associated with ABC transporters and 21 IEMs causative genes associated with SLC transporters. The most commonly occurring among these diseases is cystic fibrosis, with an incidence of 1 in 2,500 to 3,500 Caucasian newborns, caused by mutations in the *CFTR* (GeneID: 1080), which encodes an ABC transporter and acts as a chloride channel. Affected individuals usually exhibit respiratory and digestive problems, and the condition becomes fatal when left untreated [231]. Recon 1 captured 25 of 57 transport protein-associated IEMs genes. Recon 2 captured six additional diseases/genes. The remaining 26 transport protein associated IEMs could not be mapped onto Recon 2 due to missing genes (see Table 3.4 for details). The IEMs of ABC class of transporters account for 68% of these missing IEMs. The non-inclusion of the ABC transport proteins into human GENREs is due to the fact that they have been shown to transport majorly medicinally important drugs and their derivatives (e.g., *ABCB1*, GeneID: 5243, *ABCG2*, GeneID: 9429) or that insufficient information on preferred substrates is available (e.g., for *ABCD4*, GeneID: 5826). On the other hand, another major proportion of missed IEMs/genes concerns fat soluble vitamins and lipids (e.g., SRB-I, Niemann-Pick C1-like protein 1), for which transport mechanisms have only partially been resolved (see relevant sections above). Therefore, we would like to emphasize the need to have sufficient information regarding the preferred substrates, associated cofactors/ions, substrate:ion stoichiometry, transport kinetics, and sub-cellular localization of transport proteins for building a high quality reconstruction of transport reactions/pathways.

3.7 Conclusion

A great deal of work in the field of constraint-based modeling has focused on the generation of highly curated GENREs and their usage for the generation of tissue specific metabolic models for biomedical applications.

Transporters do not only maintain the connectivity of metabolites across different cell types but also determine the uptake and secretion profile of individual cells. The metabolite exchanges of individual cell types with the corresponding extracellular compartment are inevitably connected to its internal biochemical pathways and cell functions. Inclusion of the cell type specific transporters is important for enabling

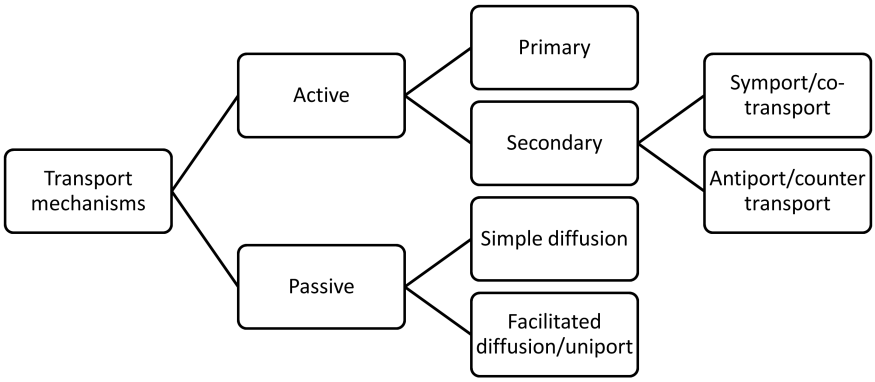


Figure 3.1: Overview of transport mechanisms. The basic modes of metabolite transport across the plasma membrane are shown. Based on the energy association, transport processes can be categorized into active and passive modes. The active mode can be further classified into primary and secondary mechanisms, while the metabolites can also be transported without an expense of energy, mainly via simple diffusion or facilitated diffusion. Refer to the text for further explanation of these transport processes. Specialized transport mechanisms, e.g., receptor mediated endocytosis and tertiary active process are not shown.

the usage of the human metabolic reconstruction as a template for the generation of more accurate and physiologically relevant cell type specific sub-networks, and ultimately of functional, representative models. Moreover, information regarding distinct transporter utilization at different locations as it is the case in polarized cells, is crucial for such effort and has thus been pointed out throughout this review. We identified numerous gaps through our literature review, many of which could be filled and are provided in the accompanying transport module. However, some gaps remain as knowledge gaps, as responsible transporter or transport mechanisms are unknown. Such knowledge update of the GENRE needs to be done periodically, as it has important implications on their predictive potential [232] and thus, on biomedical applications.

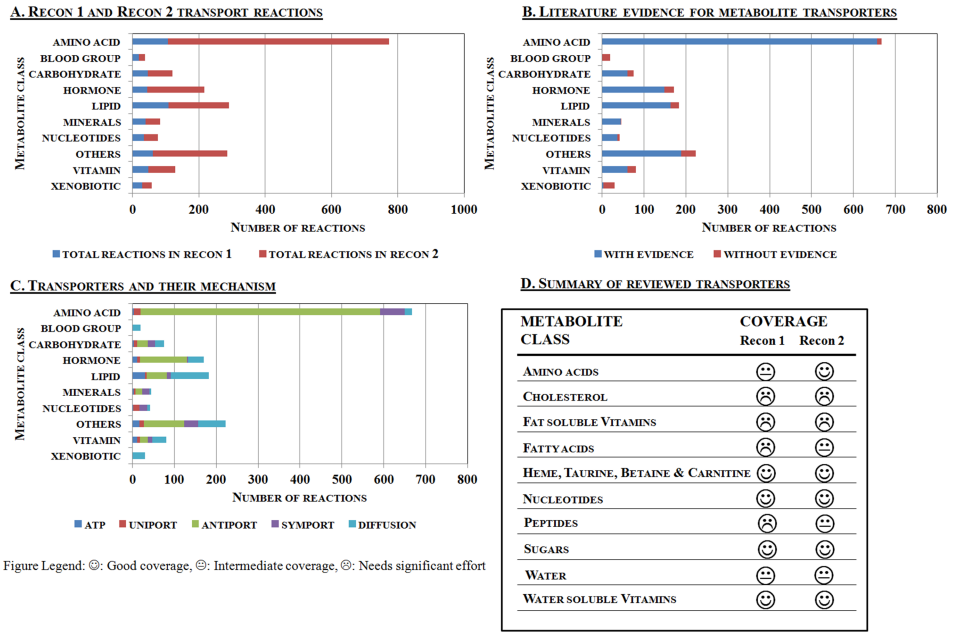


Figure Legend: ☺: Good coverage, 😐: Intermediate coverage, ☹: Needs significant effort

Figure 3.2: Overview of transport reactions captured in the human metabolic reconstructions. A. Quantitative assessment of the transport reactions present in Recon 1 [18] and Recon 2 [31]. B. Literature support for the transport reactions present in Recon 2. B. Metabolites in the ten metabolite classes are shown with their major transport mechanisms, as captured in Recon 2. D. Comparison of transport reactions present in Recon 1 and Recon 2. The gain in knowledge and expanded scope of Recon 2, over Recon 1, resulted in better transporter coverage for the amino acid and peptide class, while significant work is needed for cholesterol and fat soluble vitamins.

Table 3.1: Suagar transporters. Data was assembled from [81, 95, 233]. Yellow shading indicates genes encoding for either absent transport proteins or present transport protein but with limited substrate specificity in Recon 2. Blue shading indicates improvement of the transporter (either addition of the protein and its associated reactions, or expansion of its substrates, or modification of the GPRs) in Recon 2 over Recon 1.

Encoding genes (Gene symbols)	Encoding genes (EntrezGene IDs)	Transport proteins	Substrates
GLUTs/class			
SLC2A1	6513	GLUT-1 /1	Glucose, dehydro-ascorbic acid
SLC2A2	6514	GLUT-2 /1	Glucose, fructose, glucosamine
SLC2A3	6515	GLUT-3 /1	Glucose, dehydro-ascorbic acid
SLC2A4	6517	GLUT-4 /1	Glucose, dehydro-ascorbic acid
SLC2A5	6518	GLUT-5 /2	Fructose, glucose (very low affinity)
SLC2A6	11182	GLUT-6 /3	Glucose
SLC2A7	155184	GLUT-7 /2	Glucose, fructose
SLC2A8	29988	GLUT-8 /3	Glucose (glucose transporter activity is inhibited by fructose and galactose)
SLC2A9	56606	GLUT-9 /2	Glucose
SLC2A10	81031	GLUT-10 /3	Glucose
SLC2A11	66035	GLUT-11 /2	Glucose, fructose
SLC2A12	154091	GLUT-12 /3	Glucose
SLC2A14	144195	GLUT-14 /1	Glucose
SLC2A13	114134	HMIT /3	Myo-inositol
SGLTs			
SLC5A9	200010	SGLT4	Mannose, 1,5-anhydro-glucitol, fructose
SLC5A3	6526	SMIT	Myo-inositol
SLC5A1	6523	SGLT1	Glucose, galactose
SLC5A2	6524	SGLT2	Glucose, galactose

Encoding genes (Gene symbols)	Encoding genes (EntrezGene IDs)	Transport proteins	Substrates
SLC5A4	6527	SGLT3 (under amino acid substitution)	Glucose
SLC5A10	125206	SGLT5	Mannose, fructose, glucose, galactose
SLC5A11	115584	SGLT6	Glucose, myo-inositol

Table 3.2: Amino acids transport systems. Important amino acid derived compounds and peptides. Color coding is the same as in Table 3.1.

Encoding genes (Gene symbols)	Encoding genes (Entrez-Gene IDs)	Transport systems	Substrate	Mechanism
SLC1A1	6505	EAAT3	Cystine, glutamic acid, aspartic acid	Symport and antiport
SLC1A2	6506	EAAT2	Aspartic acid, glutamic acid	Symport and antiport
SLC1A3	6507	EAAT1	Glutamic acid, aspartic acid	Symport and antiport
SLC1A4	6509	ASCT1	Alanine, serine, cysteine	Symport and antiport
SLC1A5	6510	ASCT2	Alanine, serine, cysteine, threonine, glutamine	Antiport
SLC1A6	6511	EAAT4	Glutamate, aspartate	Symport and antiport
SLC6A1	6529	GAT1	Gamma-aminobutyric acid	Symport

Encoding genes (Gene symbols)	Encoding genes (Entrez-Gene IDs)	Transport systems	Substrate	Mechanism
SLC6A2	6530	NET	Dopamine, norepinephrine	Symport
SLC6A3	6531	DA trans-porter	Dopamine	Symport
SLC6A4	6532	SERT	Serotonin	Symport
SLC6A5	9152	GLYT2	Glycine	Symport
SLC6A7	6534	PROT	Proline	Symport
SLC6A8	6535	CT1	Creatine	Symport
SLC6A11	6538	GAT3	GABA	Symport
SLC6A13	6540	GAT2	GABA	Symport
SLC7A2	6542	CAT-2	Arginine, lysine, ornithine	Uniport
SLC7A3	84889	CAT-3	Homoarginine, arginine, lysine, ornithine	Uniport
SLC7A10/ SLC3A2	56301/ 6520	Asc-1/4f2hc	Glycine, alanine, serine, cysteine, threonine	Antiport
SLC7A11/ SLC3A2	23657/6520	XCT/4f2hc	Aspartic acid, glutamic acid, cysteine	Antiport
SLC16A10	117247	TATA1	Tryptophan, tyrosine, phenylalanine	Uniport
SLC38A1	81539	SNAT1	Glycine, alanine, asparagine, cysteine, glutamine, histidine, methionine	Symport
SLC38A3	10991	SNAT3	Glutamine, asparagine, histidine	Symport and antiport
SLC43A2	124935	LAT4	Leucine, isoleucine, methionine, phenylalanine, valine	Uniport
SLC6A6	6533	TautT	Taurine, beta-alanine	Symport
SLC6A14	11254	ATB0,+	All neutral and cationic amino acids	Symport

Encoding genes (Gene symbols)	Encoding genes (Entrez-Gene IDs)	Transport systems	Substrate	Mechanism
SLC6A20	54716	IMINO	Proline, hydroxy-proline, betaine	Symport
SLC7A1	6541	Y+ (CAT-1)	Lysine, arginine, ornithine, histidine	Uniport
SLC7A5/ SLC3A2	8140/6520	LAT1/4f2hc	Histidine, methionine, leucine, isoleucine, valine, phenylalanine, tryptophan	Antiport
SLC7A6/ SLC3A2	9057/6520	y+LAT2/4f2	Lysine, arginine, glutamine, histidine, methionine, leucine	Antiport
SLC7A7/ SLC3A2	9056/6520	y+LAT1/4f2	Lysine, arginine, glutamine, histidine, methionine, leucine, alanine, cysteine	Antiport
SLC7A8/ SLC3A2	23428/6520	LAT2/4f2hc	All neutral amino acids except proline	Antiport
SLC7A9/ SLC3A1	11136/6519	b0,+AT	Arginine, lysine, cystine, ornithine	Antiport
SLC15A1	6564	PEPT1	Peptides	Symport
SLC36A1	206358	PAT1	Glycine, proline, alanine	Symport
SLC38A2	54407	SNAT2	Glycine, proline, alanine, serine, cysteine, glutamine, asparagine, histidine, methionine	Symport
SLC38A5	92745	SNAT5	Glutamine, asparagine, histidine, alanine	Symport
SLC43A1	8501	LAT3	Leucine, isoleucine, methionine, phenylalanine, valine	Uniport

Encoding genes (Gene symbols)	Encoding genes (Entrez-Gene IDs)	Transport systems	Substrate	Mechanism
SLC6A15	55117	B(0)AT2	Proline, leucine, valine, isoleucine, methionine	Symport
SLC6A18	348932	B(0)AT3	Glycine, alanine, methionine, serine, cysteine	Symport
SLC6A19	340024	B(0)AT1	All neutral amino acids	Symport
SLC7A4	6545	CAT-4	Arginine	Requires additional factors for function as amino acid transporter
SLC15A2	6565	PEPT2	Di-peptides, tri-peptides, drugs	Symport
SLC36A2	153201	PAT2	Glycine, alanine, proline	Symport
SLC36A4	120103	PAT4	Proline, tryptophan, alanine	Facilitated diffusion
SLC38A4	55089	SNAT4	Glycine, alanine, serine, cysteine, glutamine, asparagine, methionine	Symport

Table 3.3: Lipid and vitamin transporters. Color coding is the same as in Table 3.1.

Encoding genes (Gene symbols)	Encoding genes (Entrez-Gene IDs)	Transporters	Substrate
Lipid transporters			

Encoding genes (Gene symbols)	Encoding genes (Entrez-Gene IDs)	Transporters	Substrate
ABCA1	19	ATP-binding cassette sub-family A member 1	Cholesterol
SLC27A1	376497	FATP-1	Broad specificity for both long and very long chain fatty acids, C18:1, C20:4, C16:0, C24:0
SLC27A2	11001	FATP-2	C16:0, C24:0, bile acids, other long chain fatty acids
SLC27A3	11000	FATP-3	Long chain fatty acids
SLC27A4	10999	FATP-4	Long chain fatty acids, C18:1, C20:4
SLC27A5	10998	FATP-5	Long chain fatty acids
SLC27A6	28965	FATP-6	Long chain fatty acids, C16:0, C18:0 (not active for chain length shorter than C10)
ABCG5	64240	ATP-binding cassette sub-family G member 5	Cholesterol
ABCG8	64241	ATP-binding cassette sub-family G member 8	Cholesterol
CD36	948	FAT	Very long chain fatty acids, HDL, VLDL, LDL, phospholipids, advanced glycation end products, GHRP, hexarelin, and a derivative with no growth hormone releasing properties, EP 80317, vitamin D
GOT2	2806	FABPpm	Long chain fatty acids

Encoding genes (Gene symbols)	Encoding genes (Entrez-Gene IDs)	Transporters	Substrate
NPC1L1	29881	Niemann-Pick C1-like protein 1	Cholesterol, cholestanol, campesterol, sitosterol, vitamin E, vitamin D
SCARB1	949	Scavenger receptor class B, member 1	HDL-cholesterol
Vitamin transporters			
SLC5A6	8884	SMVT	Pantothenic acid, biotin
SLC19A1	6573	RFT / Reduced folate carrier (RFC)	5-methyl THF, thiamin-mono- and di-phosphates, but not free thiamine
SLC19A2	10560	ThTr1	Thiamin, thiamin-mono- and di-phosphate
SLC19A3	80704	ThTr2	Thiamin, thiamin-mono- and di-phosphate
CUBN	8029	Vitamin D transporters	Vitamin D
FOLR1, FOLR3, FOLR2	2348, 2352, 2350	Folate receptor (FOLR1, FOLR3)	5-Methyltetrahydrofolate
GIF, CUBN, AMN	2694, 8029, 81693	Cobalamin transporters	B12 (cobalamin)
SLC23A1	9963	SVCT1	L-ascorbic acid
SLC23A2	9962	SVCT2	L-ascorbic acid
SLC46A1	113235	PCFT	Heme, folate
SLC52A1	55065	RFT1	Riboflavin
SLC52A3	113278	RFT2	Riboflavin
GC, LRP2	2638, 4036	Vitamin D transporters	Vitamin D

Encoding genes (Gene symbols)	Encoding genes (Entrez-Gene IDs)	Transporters	Substrate
OATP-2 (SLCO1B1), OATP-1 (SLCO1A2), OAT1 (SLC22A6), OAT3 (SLC22A8), OAT2 (SLC22A7), OAT4 (SLC22A11)	10599, 6579, 9356, 9376, 10864, 55867	Folate transporters	Folate
RBP4, STRA6, RBP1, RBP2	5950, 64220, 5947, 5948	Vitamin A transporters	Retinol
SCARB1, NPC1L1, ABCA1	949, 29881, 19	Vitamin E transporters	Vitamin E
SLC5A8	160728	SMCT1	Iodide, lactate, short-chain fatty acids, niacin
SLC52A2	79581	RFT3	Riboflavin

Table 3.4: Metabolic diseases associated with transport proteins. Color coding is the same as in Table 3.1.

EntrezGene ID	Transporter name	Diseases with defect in transport proteins	References
19	ATP-binding cassette sub-family A member 1	Tangier disease	OMIM: 205400

EntrezGene ID	Transporter name	Diseases with defect in transport proteins	References
5244	MRP3	Progressive familial intrahepatic cholestasis, biliary gallstone disease, intrahepatic cholestasis of pregnancy	OMIM: 171060
6342	SCP2	Zellweger syndrome and Leukoen-cephalopathy with dystonia and motor neuropathy	OMIM: 184755
6505	EAAT3	Dicarboxylic aminoaciduria	OMIM: 222730, [234]
6513	GLUT1	GLUT1 deficiency syndrome	OMIM: 606777
6514	GLUT2	Fanconi-Bickel syndrome	OMIM: 227810
6519	rBAT	Cystinuria	OMIM: 220100
6523	SGLT1	Glucose-galactose malabsorption	OMIM: 606824
6524	SGLT2	Renal glucosuria	OMIM: 233100
6555	Ileal sodium/bile acid cotransporter	Primary bile acid malabsorption	OMIM: 613291
6584	OCTN2	Systemic carnitine deficiency	OMIM: 212140
8647	Bile salt export pump	Benign recurrent intrahepatic cholestasis-2	OMIM: 603201
9056	y+LAT-1/4f2hc	Lysinuric protein intolerance	OMIM: 222700
9152	GLYT2	Hyperekplexia	OMIM: 614618, [235]
10165	CTLN2	Type II citrullinemia	OMIM: 603471
10166	ORNT 1	HHH syndrome	OMIM: 238970

EntrezGene ID	Transporter name	Diseases with defect in transport proteins	References
10560	ThTr1	Thiamine-responsive megaloblastic anemia syndrome	OMIM: 249270
10599	OATP-C	Rotor type hyperbilirubinemia	OMIM: 604843
10999	FATP-4	Ichthyosis prematurity syndrome (IPS)	OMIM: 608649
11001	FATP-2	Milder variant of X-linked adrenoleukodystrophy	OMIM: 603247
11136	b0,+AT	Cystinuria	OMIM: 220100
26503	Sialin	Sialuria	OMIM: 269920
55315	ENT3	Familial Histiocytosis Syndrome (Faisalabad Histiocytosis) and Familial Rosai-Dorfman Disease	OMIM: 602782, [236, 237]
56606	GLUT9	Renal hypouricemia-2	OMIM: 612076
113235	PCFT	Hereditary folate malabsorption	OMIM: 229050, [238]
215	ATP-binding cassette sub-family D member 1	X-linked adrenoleukodystrophy	OMIM: 300100
225	ATP-binding cassette sub-family D member 2	X-linked adrenoleukodystrophy	OMIM: 300100
1244	MRP2	Dubin-Johnson syndrome	OMIM: 237500, [239]
5825	ATP-binding cassette sub-family D member 3	Zellweger syndrome	OMIM: 170995

EntrezGene ID	Transporter name	Diseases with defect in transport proteins	References
8029	Cubilin	Hereditary megaloblastic anemia 1 (MGA1)	OMIM: 261100, [240]
54716	Sodium- and chloride-dependent transporter XTRP3	Iminoglycinuria	OMIM: 242600
20	ATP-binding cassette sub-family A member 2	Alzheimer disease (early onset)	OMIM: 600047
21	ATP-binding cassette sub-family A member 3	Pulmonary surfactant metabolism dysfunction type 3 (SMDP3)	OMIM: 610921, [241]
22	ATP-binding cassette sub-family B member 7, mitochondrial	Mutations in this gene have been implicated in X-linked sideroblastic anemia with ataxia.	OMIM: 300135
24	Retinal-specific ATP-binding cassette transporter	Stargardt disease, a form of juvenile-onset macular degeneration, retinitis pigmentosa-19, cone-rod dystrophy type 3, early-onset severe retinal dystrophy, fundus flavimaculatus, and macular degeneration age-related 2	OMIM: 248200
359	AQP2	Nephrogenic diabetes insipidus	OMIM: 125800, OMIM: 304800, [242]

EntrezGene ID	Transporter name	Diseases with defect in transport proteins	References
364	AQP7	Several metabolic alterations including obesity, type 2 diabetes	OMIM: 125853, [72]
368	MRP6	Pseudoxanthoma elasticum	OMIM: 603234
949	SRB-I	Familial hyperc-holesterolemia	OMIM: 601040
1080	CFTR	Cystic fibrosis	OMIM: 219700, [231]
2638	DBP	Graves disease	OMIM: 139200
4036	Megalin	Donnai-Barrow syndrome	OMIM: 222448, [243, 244]
5243	MRP1	Crohn's disease	OMIM: 171050
5826	ATP-binding cassette sub-family D member 4	Adrenoleukodystrophy	OMIM: 603214
6833	ATP-binding cassette sub-family C member 8	Hyperinsulinemic hypoglycemia of infancy, non-insulin-dependent diabetes mellitus type II	OMIM: 256450, OMIM: 125853, [245]
6890	TAP 1	Ankylosing spondylitis, insulin-dependent diabetes mellitus, and celiac disease	OMIM: 222100, [246, 247]
6891	TAP 2	Ankylosing spondylitis, insulin-dependent diabetes mellitus, and celiac disease	OMIM: 170261
9429	ATP-binding cassette sub-family G member 2	Gout	OMIM: 138900

EntrezGene ID	Transporter name	Diseases with defect in transport proteins	References
10058	ATP-binding cassette sub-family B member 6, mitochondrial	This gene is considered a candidate gene for lethal neonatal metabolic syndrome, a disorder of mitochondrial function.	OMIM: 605452
10060	ATP-binding cassette sub-family C member 9	Dilated cardiomyopathy	OMIM: 601439
26154	ATP-binding cassette sub-family A member 12	Lamellar ichthyosis type 2, Harlequin ichthyosis	OMIM: 607800, [248]
29881	Niemann-Pick C1-like protein 1	Multiple lipid transport defects	OMIM: 608010, [249]
64240	ATP-binding cassette sub-family G member 5	Sitosterolemia	OMIM: 210250
64241	ATP-binding cassette sub-family G member 8	Sitosterolemia	OMIM: 210250
116085	URAT1	Renal hypouricemia-1	OMIM: 220150
153201	Proton-coupled amino acid transporter 2	Iminoglycinuria and hyperglycinuria	OMIM: 242600
340024	B(0)AT1	Hartnup disorder	OMIM: 234500

4 A compendium of inborn errors of metabolism mapped onto the human metabolic network

4.1 Abstract

Inborn errors of metabolism (IEMs) are hereditary metabolic defects, which are encountered in almost all major metabolic pathways occurring in man. Many IEMs are screened for in neonates through metabolomic analysis of dried blood spot samples. To enable the mapping of these metabolomic data onto the published human metabolic reconstruction, we added missing reactions and pathways involved in acylcarnitine (Ac) and fatty acid oxidation (FAO) metabolism. Using literary data, we reconstructed an Ac/FAO module consisting of 352 reactions and 139 metabolites. When this module was combined with the human metabolic reconstruction, the synthesis of 39 acylcarnitines and 22 amino acids, which are routinely measured, was captured and 235 distinct IEMs could be mapped. We collected phenotypic and clinical features for each IEM enabling comprehensive classification. We found that carbohydrate, amino acid, and lipid metabolism were most affected by the IEMs, while the brain was the most commonly affected organ. Furthermore, we analyzed the IEMs in the context of metabolic network topology to gain insight into common features between metabolically connected IEMs. While many known examples were identified, we discovered some surprising IEM pairs that shared reactions as well as clinical features but not necessarily causal genes. Moreover, we could also re-confirm that acetyl-CoA acts as a central metabolite. This network based analysis lead to further insight of hot spots in human metabolism with respect to IEMs. The presented comprehensive knowledgebase of IEMs will provide a valuable tool in studying metabolic changes involved in inherited metabolic diseases.

4.2 Introduction

Inborn errors of metabolism (IEMs) are individually rare but collectively numerous, affecting the metabolism of many human organs. These metabolic defects are congenital and represent single or multiple enzyme deficiencies, which, if left untreated

can lead to life threatening conditions, with a current incidence rate of 1:800 live births [3]. IEMs can be grouped into three diagnostically meaningful groups [250]: i) disorders giving rise to intoxication, via accumulation of intracellular compounds over time, ii) disorders involving energy metabolism, and iii) disorders involving metabolism of complex molecules (Figure S4.1). Moreover, IEMs can be present at any age - from fetal life to old age and the symptoms vary between the groups as well as within the groups. Not all IEMs can be easily diagnosed and treatment does not exist for many of them, although many disorders of the first group can now be treated by changing the patient's diet [250]. In recent years, numerous reviews have been published analyzing IEMs from various aspects, including newborn screening programs and systems biology [251–253], highlighting their importance and the general interest in understanding the molecular basis of IEMs.

The use of tandem mass spectrometry (MS/MS) has enabled improvements in metabolic screening methods to more reliably detect IEMs among newborns with a low false positive rate [254]. Newborn screening is mandatory in most countries; however, the measured biomarkers differ significantly between them, with 42 (20 primary and 22 secondary disorders) metabolic conditions screened for newborn screening in some states within US and countries participating in the Stork program (Iceland being one of the participating country) [4, 255]. The substantial increase in the number of patients diagnosed with IEMs [252, 256], can be majorly attributed to improved and standardized measurements of amino acids and acylcarnitine concentrations [4, 257]. For the diagnosis of some IEMs, ratios of biomarker provide more insight about pathological states than absolute biomarker concentrations (e.g., tyrosine to phenylalanine ratio for the diagnosis of phenylketonuria, OMIM: 261600).

Fatty acid oxidation defects are highly prevalent IEMs with a collective incidence rate of 1:9000 [258]. In the case of infants, fatty acid oxidation serves as energy source within the first 12 hours of fasting; thus, they are more likely to manifest fatty acid oxidation defects [259]. Acylcarnitines tend to rise in blood and urine on the first day of the birth and then decline gradually [254]. By profiling numerous acylcarnitines, MS/MS technology has expanded the range of detection for fatty acid disorders (e.g., of medium chain acyl-CoA dehydrogenase deficiency) and for organic acid disorders (e.g., propionic academia and glutaric aciduria type-I).

Metabolic network reconstructions are assembled in a bottom-up approach based on genome annotation, biochemical, and physiological data [5]. They summarize current knowledge about the target organism in a structured, stoichiometric manner. In many cases, manual curation including in-depth literature search is performed to ensure high quality and coverage of the reconstruction [10]. Once reconstructed, metabolic networks can be queried for their biochemical information, be employed to map various data types (e.g., transcriptomic data or metabolomic data), or serve as starting point for computational modelling (see [260]for a recent review).

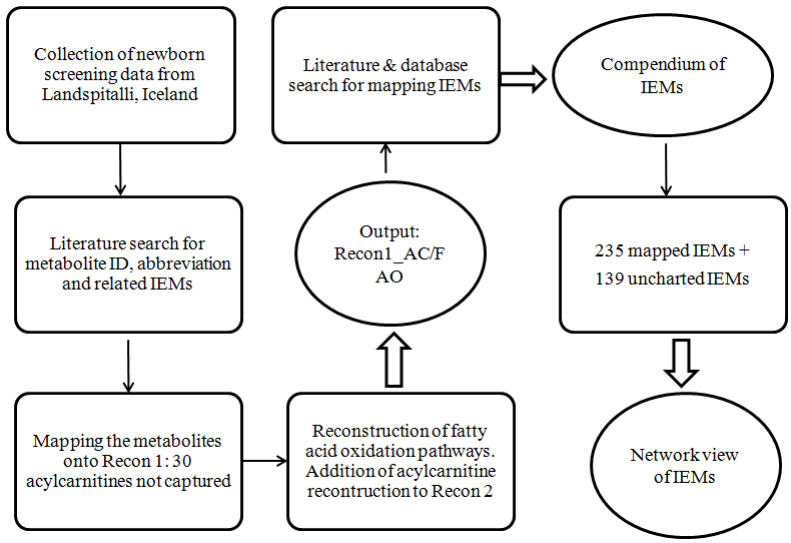


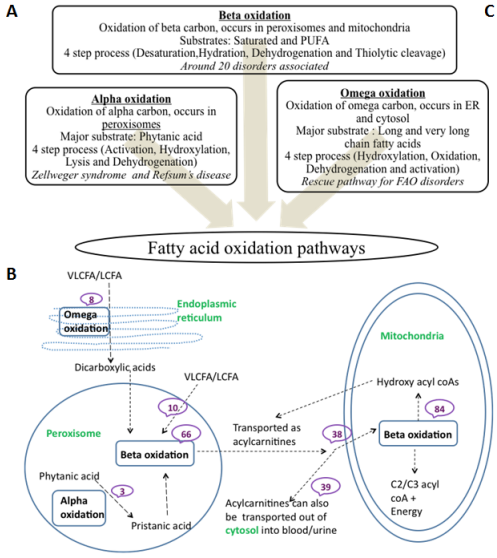
Figure 4.1: Overview of the workflow that we used in this study. The metabolites, which are measured in the newborn screening program, were identified along with their chemical names, chemical formula, charge, normal physiological concentration, and related IEMs. These metabolites were then mapped onto Recon1, using name and formula, which led to the mapping of 20 amino acids and six acylcarnitines. For the remaining metabolites (i.e., 33 acylcarnitines), which could not be mapped, we reconstructed their respective metabolic pathways based on the literature, and added the corresponding reactions to the Ac/FAO module. Subsequently, we combined the Ac/FAO module with Recon 1, resulting in Recon1_Ac/FAO. Thereafter, we employed the information retrieved from the literature to map the IEMs onto this extended reconstruction. A total of 235 IEMs could be mapped and subsequently classified along with their different physiological and clinical features. Another set of 139 IEMs were identified, termed as future IEMs, which could not be mapped onto Recon1_Ac/FAO due to missing genes in the reconstruction. These 374 IEMs constitute the IEM compendium, which we mapped onto the human metabolic network to obtain a holistic network perspective of the captured IEMs.

4.3 Results and discussion

The aim of this study was to compile a compendium of IEMs that map onto the available reconstruction of human metabolism [18] to enable the analysis of IEMs in a metabolic modelling context (Figure 1). Therefore, we first reconstructed the fatty acid oxidation pathways to enable the mapping of routinely measured acylcarnitines, which were missing in the published reconstruction of human metabolism [18]. Subsequently, we mapped IEMs onto the expanded human metabolic network based on known affected genes, and assembled important phenotypic and clinical features of each IEM. Finally, we investigated the IEMs in the context of the expanded human metabolic reconstruction by identifying IEMs, which shared reactions and metabolites, therefore, representing hot spots in human metabolism.

4.3.1 Detailed reconstruction of acylcarnitine metabolism

In this study, we employed the published human metabolic reconstruction, Recon 1 [18], as a starting point for the acylcarnitine and fatty acid oxidation (Ac/FAO) reconstruction module. The human metabolic reconstruction accounts for metabolic reactions occurring in any human cell, as defined by the human genome. As such, the reactions present in Recon 1 may not occur in all cells, but a particular cell type may only express a subset of metabolic enzymes captured by Recon 1. For example, the liver completely detoxifies ammonia to urea via urea cycle, however, in kidney the urea cycle is only partially active, i.e., kidney accounts for arginine synthesis. The majority of the reactions are mass and charge balanced, and many of them are assigned with structured Boolean relationships between genes, proteins, and reactions. Recon 1 captures all amino acids (except methylhistidine) measured in the newborn screening but it only accounts for the metabolism of six of the 39 acylcarnitines. Moreover, Recon 1 accounts for most of the relevant reactions in beta fatty acid oxidation, while alpha and omega fatty acid oxidation reactions are only partially captured. After collecting supporting evidence from more than 150 peer reviewed articles and books, we reconstructed an Ac/FAO module that accounts for 352 reactions, 139 metabolites, and 14 genes, with the reactions distributed within the endoplasmic reticulum, peroxisomes, mitochondria, and cytosol (Figure 4.2B). The majority of the reactions are involved in fatty acid metabolism (Figure 4.2A). Note that this reconstruction module is only functional in conjunction with Recon 1 [18] or its successor [31].



C

Components	Recon 1	Recon1_ AC/FAO
Genes	1496	1510
Transcripts	1905	1919
Compartments	8	8
Reactions	3743	4095
Exchange reactions	405	444
Unique metabolites	1509	1648
Citations	1587	+150**

*Figure 4.2: General properties of the Ac/FAO module. (A) Overview of the fatty acid oxidation pathways, along with their definitions, substrate preferences, key steps involved, and associated disorders. (B) Schematic map of the key reactions in the Ac/FAO module and their sub-cellular location. The numbers indicate the number of reactions per metabolic pathway. The substrates for the omega oxidation pathway are very long chain fatty acids and long chain fatty acids, while the products are dicarboxylic acids, which further undergo beta oxidation in the peroxisome. Phytanic acid, which is a substrate for the alpha fatty acid oxidation, produces pristanic acid. This product is further beta-oxidized in the proxisomes before finally entering the mitochondrion. The fatty acyl-CoA molecules are transported between the peroxisomes and mitochondria in the form of acylcarnitines, and once inside the mitochondria they may either undergo complete beta oxidation generating acetyl-CoA (in the case of even chain fatty acid) or propionyl-CoA (in the case of odd chain fatty acid) and finally produce energy through the oxidative phosphorylation in the mitochondria. However, if the beta oxidation is incomplete, it produces hydroxy fatty acids, which are then shuttled out of the mitochondrion as acylcarnitines. Substrates for peroxisomal beta oxidation are long chain and vary long chain fatty acids entering the peroxisome via an ATP dependent transport system. Short, medium, and long chain fatty acids enter mitochondria as acylcarnitines. (C) Key characteristics of Recon 1 and Recon1_Ac/FAO. A reversible version of the Recon 1 reactions, i.e., 64 reactions of the carnitine shuttle system, was also added (not shown here). *Taken from ref. 103. **Corresponds to primary literature, peer review articles, and books that were used only for the acylcarnitine reconstruction. Abbreviations used: FA – fatty acid; PUFA – poly unsaturated fatty acid; ER – endoplasmic reticulum; FAO – fatty acid oxidation; VLCFA – very long chain fatty acid; LCFA – long chain fatty acid; MCFA – medium chain fatty acid; SCFA – short chain fatty acid; C2 – acetyl-CoA; C3 – propionyl-CoA.*

Acylcarnitines in the Ac/FAO module

The Ac/FAO module accounts for 39 acylcarnitines routinely measured in newborn screening, which include dicarboxylic-acylcarnitines, hydroxy-acylcarnitines, and other short chain acylcarnitines (Figure 4.2B). For instance, succinylcarnitine, a four carbon fatty acid is a dicarboxylic-acylcarnitine and generated via the valine, isoleucine and methionine pathway. Hydroxy-isovalerylcarnitine is a five carbon fatty acid attached to a carnitine moiety, has a hydroxy group at the third carbon atom and is generated via the leucine pathway. These two dicarboxylic-acylcarnitines are used as biomarkers for organic acidemias and more recently also for biotin deficiency [261–263]. A total of five reactions were added describing the synthesis of succinylcarnitine and hydroxy-isovalerylcarnitine, and subsequent processing by the carnitine palmitoyltransferase 1 enzyme (CPT-1, E.C. 2.3.1.21), which is encoded by three gene isoforms *CPT1A* (GeneID: 1374), *CPT1B* (GeneID: 1375), and *CPT1C* (GeneID: 126129). Another dicarboxylic-acylcarnitine is glutarylcarnitine, which is an important biomarker in glutaric aciduria [264, 265], generated as an intermediate in lysine and tryptophan metabolism. Likewise, 3-hydroxybutyrylcarnitine, a hydroxy-acylcarnitines and biomarker for short-chain-acyl-CoA-dehydrogenase deficiency [266], is generated via butanoate metabolism. Hydroxyacylcarnitines generally serve as biomarkers of several mitochondrial fatty acid beta oxidation disorders [267].

Other short chain acylcarnitines include acetylcarnitine, tiglylcanitine, and isovalerylcarnitine. Acetylcarnitine is produced from acetyl-CoA, which is an important endproduct of fatty acid oxidation, ketogenic amino acid metabolism, and a precursor of cholesterol metabolism. An increased level of acetylcarnitine may reflect CPT-2 deficiency [268]. Tiglylcarnitine and isovalerylcarnitine are intermediates of isoleucine and leucine metabolism, respectively, and are formed from their respective precursor fatty acyl-CoAs, tiglyl-CoA and isovaleryl-CoA, by the action of the CPT-1 enzyme. These short chain acylcarnitines are biomarkers for methylacetoacetyl-CoA deficiency, isovaleric acidemia, and short-chain acyl-CoA dehydrogenase deficiency [269–271]. The metabolic reactions of these important short chain acylcarnitines were added to the Ac/FAO module.

Addition of carnitine shuttle to mitochondria and ABC transport reactions to the peroxisome

Specific transporters are required to deliver acyl-CoAs to their destined sub-cellular compartments for their oxidation. The three step carnitine shuttle system transports acyl-CoAs into the mitochondrion and this transport is reversible [272] involves re-

actions catalyzed by the CPT-1 enzyme, the carnitine/acylcarnitine translocase protein (*SLC25A20*, GeneID: 788) and the carnitine O-palmitoyltransferase 2 (*CPT2*, GeneID: 1376, E.C. 2.3.1.21). The transport is affected in some IEMs, such as the carnitine uptake defect, CPT-1 deficiency, and CPT-2 deficiency [272]. The Ac/FAO module contains a refined carnitine shuttle system with 38 reactions accounting for the reversible transport of a wide range of metabolites. Peroxisomes contain ATP-dependent transporters for the import of very long chain acyl-CoAs (Figure 4.2B) [273]. A total of ten ABC transport reactions and their associated genes were identified, including three genes of the ATP-binding cassette superfamily (*ABCD1*, GeneID: 215; *ABCD2*, GeneID: 225; *ABCD3*, GeneID: 5825). The human metabolic reconstruction accounts for the transport of 4,8-dimethylnonanoyl-CoA by the carnitine O-octanoyltransferase (*CROT*, GeneID: 54677, E.C. 2.3.1.137). The same protein also transports butyryl-CoA and hexanoyl-CoA into the peroxisome [272]. We included this extended substrate specificity in the Ac/FAO module.

Addition of new fatty acid oxidation genes

To capture all fatty acid oxidation related IEMs, 84 mitochondrial beta-oxidation reactions and 63 metabolites were added to the Ac/FAO module. We also added the necessary genes and their respective gene products, including hydroxyacyl-CoA dehydrogenase (*HADH*, GeneID: 3033, E.C. 1.1.1.35) and the acetyl-CoA acyltransferase 2 (*ACAA2*, GeneID: 10449, E.C. 2.3.1.16), which have substrate specificity for short and medium chain acyl-CoAs and catalyze the dehydrogenation and thiolytic cleavage step, respectively [274]. The 2,4-dienoyl-CoA reductase 1 (*DECR1*, GeneID: 1666, E.C. 1.3.1.34) was another auxiliary enzyme of the mitochondrial beta-oxidation of unsaturated acyl-CoAs that was missing in the human metabolic reconstruction. The addition of these three genes enabled the mapping of further five IEMs. Moreover, four additional genes were added to the Ac/FAO module, whose products are mainly involved in activation and transport of fatty acids [275, 276]. These were fatty acid transporter (*SLC27A6*, GeneID: 28965), bubblegum (*ACSBG1*, GeneID: 23205, E.C. 6.2.1.3), bubblegum related protein (*ACSBG2*, GeneID: 81616, E.C. 6.2.1.3), and diazepam binding inhibitor (*DBI*, GeneID: 1622).

Long and very long chain acyl-CoAs are oxidized in peroxisome before entering the mitochondrion [277, 278]. Therefore, 66 peroxisomal beta-oxidation reactions and 59 metabolites were included in the Ac/FAO module (Figure 4.2B) permitting the mapping of X-linked adrenoleukodystrophy (OMIM: 300100), acyl-CoA oxidase-1 (OMIM: 264470), and D-bifunctional-protein deficiencies (OMIM: 261515). The peroxisome contains the complete enzymatic machinery for handling unsaturated fatty acids [279, 280]. In the human metabolic reconstruction, important auxiliary

proteins, reactions, and their respective genes were missing and were thus included in the Ac/FAO module. This addition comprises the functions of the peroxisomal 2,4-dienoyl-CoA reductase 2 (*DECR2*, GeneID: 26063, E.C. 1.3.1.34) and the enoyl-CoA delta isomerase 2 (*PECI*, GeneID: 10455, E.C. 5.3.3.8).

Alpha oxidation of phytanic acid was also reconstructed in detail in the Ac/FAO module (Figure 4.2B). Two new genes and the reactions catalyzed by the respective gene product were added: fatty acid transporter (*SLC27A2*, GeneID: 11001, E.C. 6.2.1.-) and phytanoyl-CoA 2-hydroxylase 1 (*HACL1*, GeneID: 26061, E.C. 4.1.-.-). Reduced activity of the SLC27A2 transporter protein has been reported to be responsible for the biochemical pathology in X-linked adrenoleukodystrophy in mouse models [281]. Omega oxidation is important for the generation of dicarboxylic acids; therefore, eight reactions were added to Ac/FAO (Figure 4.2B).

Detailed reconstruction of fatty acid oxidation reactions

The long chain fatty acids were only partially captured in the human metabolic reconstruction. Gene-protein-reaction (GPR) associations for three acyl-CoA dehydrogenases: C-2 to C-3 short chain (*ACADS*, GeneID: 35, E.C. 1.3.99.2), C-4 to C-12 straight chain (*ACADM*, GeneID: 34, E.C. 1.3.99.3), and very long chain (*ACADVL*, GeneID: 37, E.C. 1.3.99.-), were incorporated according to their substrate specificities. We also added the reactions catalyzed by the trifunctional hydroxyacyl-CoA dehydrogenase (*HADHA*, GeneID: 3030 and *HADHB*, GeneID: 3032; E.C. 4.2.1.17/1.1.1.211/ 2.3.1.16) [274], which catalyses the steps after desaturation of long acyl-CoAs (>14 carbon units). The mitochondrial short chain enoyl-CoA hydratase 1 (*ECHS1*, GeneID: 1892, E.C. 4.2.1.17) catalyzes the hydration step for medium and short chain fatty acyl-CoAs, but in Recon 1, the gene product catalyzes only reactions involved in tryptophan and beta-alanine metabolism. Hence, the hydration steps needed to be added to account for the relevant fatty acid oxidation reactions. We also expanded the substrate specificity of the enoyl-CoA delta isomerase 1 (*DCI*, GeneID: 1632, E.C. 5.3.3.8), which has only one reaction associated in Recon 1. These comprehensive additions to the fatty acid oxidation pathway enabled the mapping of numerous IEMs, including peroxisomal acyl-CoA oxidase deficiency/ adrenoleukodystrophy pseudoneonatal (OMIM: 264470), Zellweger syndrome, neonatal adrenoleukodystrophy and infantile Refsum disease, D-bifunctional protein deficiency (OMIM: 261515), and pseudo-Zellweger syndrome.

All enzymes involved in omega oxidation were present in Recon 1, but catalyzed only reactions involving either xenobiotic or eicosanoid metabolism, while reactions of the omega oxidation of fatty acids were missing. The Ac/FAO module accounts for all known reactions of the omega oxidation pathway [282–286] along with the

relevant GPR association, linking the corresponding genes to these reactions. The participating genes include the cytochrome P450 family genes (*CYP4F2*, GeneID: 8529, E.C. 1.14.13.30; *CYP4F3*, GeneID: 4051, E.C. 1.14.13.30), the aldehyde dehydrogenase 3 family (*ALDH3A2*, GeneID: 224, E.C. 1.2.1.3), the alcohol dehydrogenase (*ADH5*, GeneID: 128, E.C. 1.1.1.1), and the acyl-CoA synthetase (*ACSL5*, GeneID: 51703, E.C. 6.2.1.3).

4.3.2 Mapping of inborn errors of metabolism onto the human metabolic reconstruction

A major motivation of this study was to create a comprehensive compendium of IEMs that maps onto the human metabolic reconstruction. We therefore joined our Ac/FAO module with the published human metabolic reconstruction (Recon 1) [18] and called this extended network Recon1_Ac/FAO. To map defective genes (i.e., gene mutations that cause an IEM) reported in literature and databases for the various IEMs to reactions, we employed the GPR associations defined for the 1510 metabolic genes accounted for in Recon1_Ac/FAO. After extensive literature search, we identified 235 IEMs, which had all causal genes captured in Recon1_Ac/FAO. In 139 cases, not all genes known to cause an IEM were present in the reconstruction and we collected them in a separate list, which represents a starting point for future extension and gap filling of functionalities included in the human metabolic reconstruction [232, 287]. The following information was retrieved for each IEM from literature and databases: affected organs, mode of inheritance, biomarkers, biomarker concentrations, and phenotype. Below we illustrate some interesting cases below.

Pathway classification

The 235 mapped IEMs could be grouped into 16 different central metabolic pathways (Figure 4.3A). The highest number of IEMs was found to be in carbohydrate metabolism (65), followed by amino acid metabolism (54), and lipid metabolism (51). Carbohydrate metabolism includes a wide range of metabolic pathways, more so, since we counted also glycoproteins and glycolipids metabolism to this subsystem. Moreover, this class of inherited disorders is well studied [288]. The 235 IEMs were caused by defects in 250 unique metabolic genes, which encode enzymes catalyzing 1067 reactions in total (this includes 750 unique and 317 shared metabolic reactions). Similar distribution pattern was observed for the affected genes, when categorizing them according to their biochemical pathway. The high number of reactions relative to the number of genes was caused by the broad substrate specificity

of many enzymes, which was particularly true for enzymes involved in the fatty acid oxidation metabolism. This analysis highlights that a systems biology modelling approach may facilitate significantly the assessment of an IEM's contribution to its global metabolic phenotype.

Mode of inheritance and phenotype analysis

IEMs follow a particular mode of inheritance [289]. We grouped the IEMs based on five different modes of inheritance: i) autosomal recessive, ii) autosomal dominant, iii) autosomal recessive or autosomal dominant (as seen in hypophosphatasia, OMIM: 241500, OMIM: 241510, and OMIM: 146300), iv) X-linked or autosomal dominant (as seen in familial gynecomastia, OMIM: 139300 and OMIM: 107910), and v) X-linked pattern (X-linked recessive and X-linked dominant IEMs treated as one mode, i.e., X-linked pattern). Inheritance patterns were identified for 199 (61%) of the IEMs, as the majority of the remaining IEMs have not been studied sufficiently or have been only identified in very few individuals so that the pattern of inheritance could not be established. This latter group includes acetyl-CoA carboxylase deficiency (OMIM: 200350), AICA-ribosiduria (OMIM: 608688), and hepatic lipase deficiency (OMIM: 151670). The majority of the IEMs with reported inheritance pattern were autosomal recessive (163 IEMs, 82%).

IEMs can also be classified according to the number of phenotypic patterns observed. Different phenotypes arise due to modifier genes, allelic variations, complex genetic and environmental interactions, or environmental factors. Therefore, specific effects may be seen at different age (i.e., appearance of specific morphological features), or may lead to difference in the severity of the disorder [290, 291]. We were able to collect such information for 90 IEMs (38%). The majority of these IEMs present two phenotypic forms. For example, a classic form and a mild form (e.g., in medium chain acyl-CoA dehydrogenase deficiency, OMIM: 201450), an acute and a chronic form (e.g., in isovaleric academia, OMIM: 243500), an early and a delayed onset (e.g., in carbamoyl phosphate synthetase I deficiency, OMIM: 237300), a severe and a moderate form (e.g., in glucose-6-phosphate dehydrogenase deficiency, OMIM: 305900), a type 1 and a type 2 form (e.g., in 3-methylcrotonyl-CoA carboxylase deficiency, OMIM: 210200 and OMIM: 210210). Other IEMs have more complex phenotypic pattern.

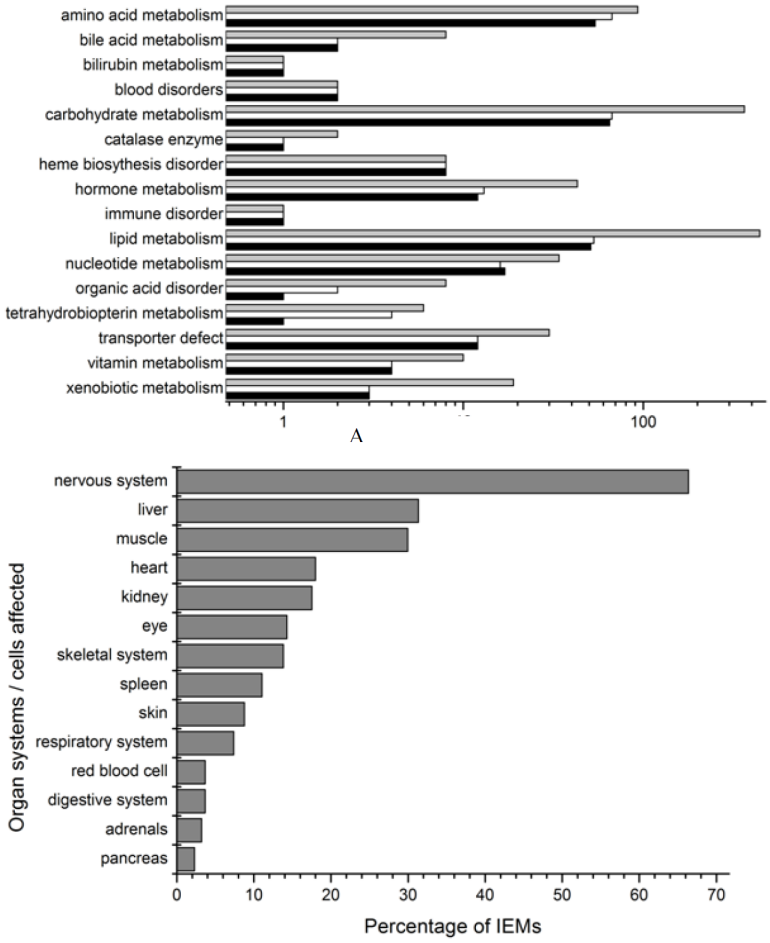


Figure 4.3: Distribution of IEMs based on various properties. (A) The 235 IEMs mapped onto Recon1_Ac/FAO were categorized into 16 different metabolic categories. The comparison between the number of IEMs (black), number of genes (white), and number of reactions (grey) per IEM category is plotted on a logarithmic scale. A total of 250 unique genes encode for gene products catalyzing 750 unique metabolic reactions, which are affected in at least one IEM. (B) Percentage of IEMs that affect different organ systems. For simplicity we only show highly affected organs.

Clinical manifestations observed in affected patients with particular IEMs were also noted. A total of 28 organ systems were affected in 217 distinct IEMs. Most of the remaining IEMs, such as essential pentosuria (OMIM: 260800), erythrocyte AMP deaminase deficiency (OMIM: 612874), and acatalasemia (OMIM: 115500), present no apparent clinical dysfunction. As one may expect, the organ systems were not equally affected by IEMs (Figure 4.3B). A criterion for calling an organ 'affected' was either i) when a clinical sign was reported by a clinical practitioner (e.g., tongue in mucopolysaccharidosis VI, OMIM: 253200), ii) when a symptom was described by the patient, iii) an identified pathological report, iv) an established biochemical defect (e.g., kidneys in renal glucosuria, OMIM: 233100, due to inability of the renal tubules to reabsorb glucose), or v) a combination of these criteria. For example, cherry red spot signs on the eye have been reported as abnormal feature in Sandhoff disease (OMIM: 268800) and Tay-Sach's disease (OMIM: 272800); thus, the eye was noted as being affected. For some IEMs, organs can be affected at a later stage, as it is the case for the CPT-2 deficiency (lethal neonatal form, OMIM: 608836), where postmortem examination showed diffuse lipid accumulation in the liver, heart, kidney, adrenal cortex, skeletal muscle, and lungs. In these cases, the affected organs were noted based on this data. Certain organs, such as liver, pancreas, bile canaliculi, and gall bladder were deliberately treated separately from the digestive system since these organs are specifically affected in certain disorders without involvement of the digestive system as a whole. For example, glycogen storage disease type 6/ Hers disease affects only the liver [292], while the HMG-CoA lyase deficiency (OMIM: 246450) affects only liver and pancreas and the familial hypercholanemia (OMIM: 607748) affects liver and bile canaliculi. The category digestive system mainly represents small and large intestine in our classification scheme. The nervous system includes central nervous system and peripheral nervous system. The nervous system was also deemed to be affected in IEMs with hypotonia as a clinical sign (e.g., in 3-methylcrotonyl-CoA carboxylase deficiency, propionic academia, and aromatic L-amino acid decarboxylase deficiency [293]) since hypotonia, or reduced muscle tone, is related to motor nerve dysfunction, which is controlled by the brain. Using this classification scheme, we found the nervous system to be the most commonly affected organ by IEMs. This result can be explained in part on the basis that in almost all aminoacidopathies (second most prevalent category of IEMs, Figure 4.3A), blood concentration of one or more amino acid is increased, which then competitively inhibits the transport of essential amino acids across the blood-brain-barrier [294]. Amino acids and their important derivative metabolites also serve as neurotransmitters, which are serine (produces the choline part of acetylcholine), tyrosine (produces L-3,4-dihydroxyphenylalanine, dopamine, and catecholamines), tryptophan (produces seretonin), glycine, glutamate (produces gama-aminobutyric acid), and histidine (produces histamine). In

many aminoacidurias, there may be observed deficits of neurotransmitters, particularly of N-methyl-D-aspartate, and of their receptors [294]. Furthermore, low levels of N-acetylaspartate and high levels of lactate in brain, besides, high levels of branched chain amino acids in both brain and blood, is observed during metabolic de-compensation, and indicates mitochondrial dysfunction leading to brain energy failure [295,296]. These examples illustrate how a defective amino acid metabolism can dramatically affect the nervous system. The category 'muscles' includes skeletal and smooth muscles but not cardiac muscles, since we considered the heart as a separate organ category. Carbohydrates and lipids provide energy to muscles while amino acids help in forming muscle proteins, which explains why these metabolic pathway defects affect muscle function. Liver is the central metabolic unit of the human body accounting for majority of the biochemical transformations. A disturbance in any metabolic pathway will eventually affect the liver's function. Energy supply of the heart is mostly provided by fatty acids; however, glucose, lactate, ketone bodies, and amino acids are also used at different proportions [297].

Biomarker classification

Enzyme assays, which test for enzyme activity, and identification of specific mutations, are frequently used as diagnostic tools for IEMs2. The diagnosis of IEMs experienced a tremendous rise with the advent of high-throughput detection techniques, such as mass spectrometry [298]. Metabolic biomarkers arise due to the presence of an IEM and are extensively used for diagnosis. For example, a disease either causes accumulation of a biomarker (e.g., elevated triglycerides in Tangier disease, OMIM: 205400) or the decrease/absence of a specific metabolite (e.g., carnitine in the carnitine uptake defect, OMIM: 212140). In fact, 17 different metabolites are used as biomarkers for the diagnosis of 130 of the 235 IEMs (Figure S4.2). Metabolic biomarkers are, of course, acylcarnitines and amino acids, which also include amino acid derivatives, such as 5-oxoproline, 3-methoxytyrosine, and 5-hydroxytryptophan, but also include the other classes of compounds. For example, carbohydrates account for a variety of compounds, most notably dolichol, oligosaccharides, galactitol, galactose, glucose, xylulose, arabitol, ribitol, glycogen, and glycan molecules. Cholesterol includes free and esterified cholesterol and its steroid derivatives (lipoproteins like HDL-cholesterol). Fatty acids include both straight and branched chain fatty acids as well as their derivatives, such as 2-ethyl-3-ketohexanoic acid and 3-hydroxyisovaleric acid. The lipids class includes, for example, triglycerides, prostagladins (leukotrienes), and glycerol. The organic acids class represents a wide range of compounds, such as glycolic acid, glyoxylic acid, glyceric acid, and oxalic acid.

To assess the use of different metabolic biomarkers and enzyme assays employed for

diagnosis of IEMs, we compared the reported biomarkers across the mapped IEMs. We found that enzyme assays are used for the diagnosis of 34% of the mapped IEMs (Figure S4.2). Moreover, we found that the measurement of amino acid and acylcarnitine concentrations are primary laboratory tests, followed by tests for carbohydrates, organic acids, and fatty acids. In fact, acylcarnitine profiling is one of the emerging trends [299] due to the numerous roles of carnitine in the body [300]. Abnormal concentrations of acylcarnitines may indicate aminoacidopathies, fatty acid oxidation defects, and organic acid disorders. However, the absence of acylcarnitine abnormalities does not rule out disorders and needs to be confirmed by enzyme assays [299]. Moreover, all diagnostic criteria need confirmation for the presence of known mutations in the corresponding gene(s) [299, 301, 302]. The DNA analysis may serves as sole diagnostic test for some IEMs, including infantile neuroaxonal dystrophy (OMIM: 256600) and leber congenital amaurosis type 3 (OMIM: 612712).

Types of therapies

Six different types of therapies are used to treat 158 of the 235 mapped IEMs (67%). The different therapeutic measures include dietary interventions, medications, organ transplantation, enzyme replacement therapy, and gene therapy. Enzyme replacement therapy is used to treat, for example, adenosine deaminase deficiency (OMIM: 102700), Fabry disease (OMIM: 301500), Pompe disease (OMIM: 232300), and mucopolysaccharidosis type VI/ Maroteaux-Lamy syndrome (OMIM: 253200). Organ transplantation usually involves either bone marrow transplant or a specific organ, and is generally employed in cases of mucopolysaccharidosis type I (OMIM: 607014) and methylmalonic academia [303]. So far, enzyme replacement therapy has been promising, e.g., for lysosomal storage disorders, while organ transplantation has only shown limited success due to graft rejection [250]. Additionally, gene therapy depends on the possible extent and duration of the gene expression without leading to toxicity [250]. On the other hand, dietary measures and specific medications have been preferred methods for the treatment of many IEMs. Therefore, we considered diet and medications as separate categories for IEMs, which are treated solely by these measures.

4.3.3 Functional network view of the IEM compendium

The combined Recon1_Ac/FAO metabolic network accounted for 3,032 metabolites, 4,095 reactions, and 1919 transcripts, corresponding to 1510 unique genes, with metabolites and reactions distributed over eight compartments (Figure 4.2C).

We investigated the 235 mapped IEMs from a functional network topology perspective. First, we determined the number of reactions associated with each IEM (Figure 4.4A). Interestingly, six IEMs were associated with more than 50 reactions each in Recon1_Ac/FAO, including CPT-1 deficiency and Zellweger syndrome (both having 71 reactions associated with their respective defective genes). These IEMs with high reaction association represent hubs in the metabolic network, which are expected to have a great impact on the network functionality when perturbed. In fact, many of these IEMs are known to have a severe phenotype. While 97 (41%) of the IEMs had only one reaction associated, we were interested in whether there were any IEMs, which had multiple associated reactions, that shared reactions and exhibited similar or overlapping phenotype. We identified 53 IEM pairs, consisting of 52 unique IEMs, which share at least one metabolic reaction in Recon1_Ac/FAO. The reaction overlap, and potential similarities of disease phenotypes, may also hamper the diagnosis of these IEMs. At the same time, this network driven view on IEMs permits us to correlate diseases that may not be associated with each other otherwise. We were particularly interested in those IEMs that had all reactions in common since these IEMs and knowledge about their phenotypes and biomarkers could be used in subsequent studies to assess the model's predictive capabilities and may also underline the importance of tissue- and cell-type specific reconstructions for metabolic modelling.

Furthermore, we identified seven reaction pairs in Recon1_Ac/FAO that shared all reactions (Table 4.2). Interestingly, these IEM pairs also shared clinical features. For example, the Tay-Sachs disease (OMIM: 272800) and the Sandhoff disease (OMIM: 268800) are caused by a defect in the beta-hexosaminidase (E.C. 3.2.1.52) [304]. The beta-hexosaminidase catalyzes the glycosyl bond hydrolysis, thus, removing amino sugars from the non-reducing ends of oligosaccharides [305]. Three forms of the enzyme have been described (A, B, and S) with distinct structures and subunit compositions [306]. The gene *HEXA* (Gene ID: 3073) encodes the alpha subunit, whereas the gene *HEXB* (Gene ID: 3074) encodes the beta subunit of the enzyme. Though different mutations in these two genes lead to Tay-Sachs disease and Sandhoff disease, these diseases are clinically indistinguishable (OMIM: 268800) [307]. Nonetheless, a mouse model study has reported differences in phenotypes [308]. Another interesting case was the IEM pair of sialidosis (OMIM: 256550) and Morquio syndrome (OMIM: 253000), which are both defects in carbohydrate metabolism and exhibit similar neurologic abnormalities [309] (Table 4.2).

The network view of IEM compendium does not only allow us to determine reactions that are shared between IEMs but also assists in the systematic identification of IEMs, which are connected through common metabolites. This analysis was triggered by the observation that many IEMs were adjacent on the metabolic map (Figure 4.5). Metabolic connectivities between IEMs may also highlight co-

morbidities [310]. When excluding highly connected metabolites, such as adenosine triphosphate, water, and protons, we found that acetyl-CoA in particular acted as a common link between the metabolic pathways, which is a well-known fact in biochemistry. Acetyl-CoA is i) used to acetylate various metabolites, including polyamines, ii) generated as an end-product of leucine and isoleucine catabolism, iii) a precursor for cholesterol and fatty acid synthesis, and iv) required to form N-acetyl-glutamate, which is an allosteric activator of carbamoyl-phosphate-synthase-I (E.C. 6.3.4.16) catalyzing the first reaction of the urea cycle [120]. Under normal physiological conditions, acetyl-CoA enters the TCA cycle after being produced by beta fatty acid oxidation. However, under prolonged fasting conditions, acetyl-CoA enters ketone body synthesis, and they appear in urine [120]. Our network-based analysis further revealed that acetyl-CoA serves a hub linking different important IEMs (Figure 4.4B), whereas the metabolic fate of acetyl-CoA varies depending on pathways and pathophysiological conditions.

Table 4.1: Properties of mapped IEMs explained using classical phenylketonuria (PKU, OMIM: 261600) as an example

Pathway classification	IEMs were classified according to their metabolic pathways.	Amino acid metabolism
Affected genes	Mutations in the gene lead to (i) non-functional protein, (ii) mutated protein, or (iii) no protein synthesis. All these instances are giving rise to an IEM.	Phenylalanine hydroxylase (PAH, GeneID: 5053, E.C. 1.14.16.1)
Affected reactions	Genes encode enzymes that catalyze reactions. IEMs are seen due to enzyme deficiencies, transporter, or co-factor defects; thereby, reactions are either blocked or follow abnormal pathways.	O2+ phenylalanine + tetrahydrobiopterin -> tyrosine+ dihydrobiopterin + H2O (PHETHPTOX2 in Recon 1). Accumulated phenylalanine produces phenylacetate and phenyllactate as abnormal metabolites.
Mode of inheritance	IEMs are inherited in a dominant or recessive pattern. Majority of the IEMs are autosomal recessive.	Phenylketonuria follows an autosomal recessive pattern of inheritance.

Pathway classification	IEMs were classified according to their metabolic pathways.	Amino acid metabolism
Phenotypes	A single metabolic defect may present various phenotypic patterns, which may range from two to seven different forms (see text).	Three major types: phenylketonuria (PKU), non-PKU hyperphenylalaninemia and variant PKU.
Affected organs	The severity of the IEMs is assessed by their ability to affect organ systems. While some of the IEMs may be un-noticed, many of them have profound effects on different organ systems.	Brain
Biomarkers	Biomarkers enable easy and early diagnosis of IEMs.	In plasma: increased phenylalanine and phenylalanine/tyrosine ratio. In urine: increased excretion of 2-hydroxyphenylacetic acid, phenyl lactic acid, and phenyl pyruvic acid.
Therapies available	IEMs, when diagnosed at a young age, are often treatable. Various therapeutic measures are available, including dietary interventions, medications, gene therapy, enzyme replacement, and organ transplantation.	Phenylalanine is added to the diet. In fact, most disorders of amino acid metabolism are controlled by dietary interventions.

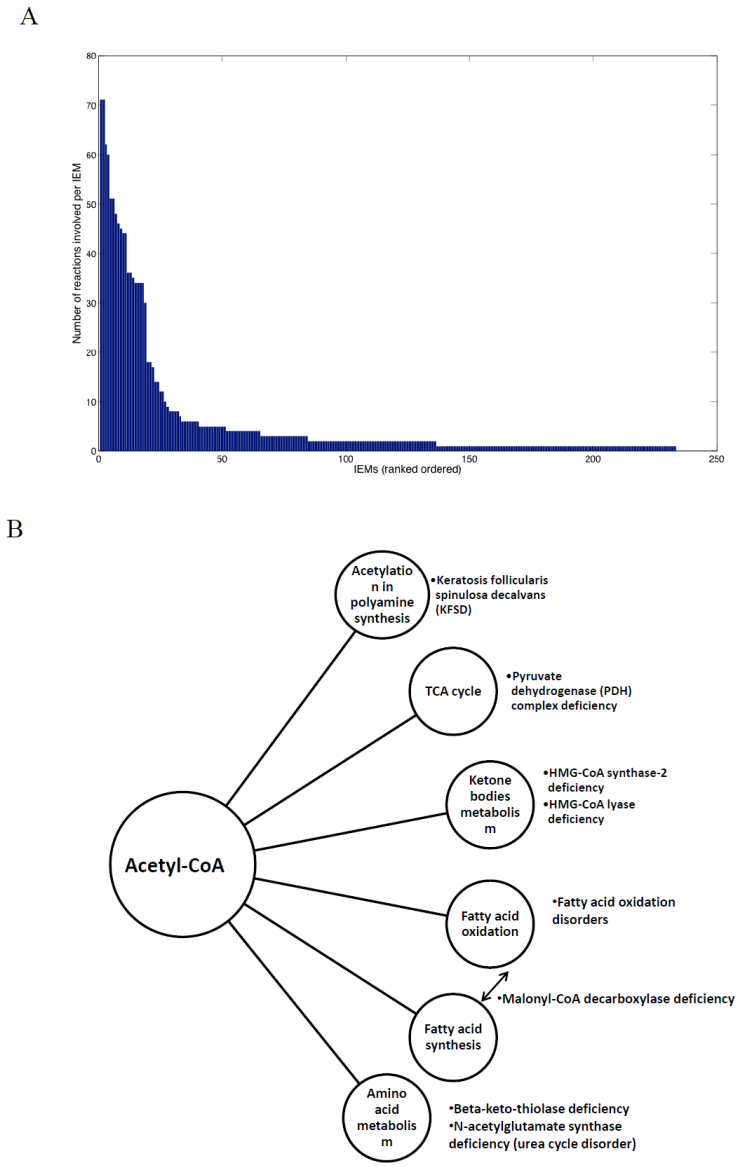


Figure 4.4: Network properties of the IEM compendium. (A) The graphic shows the number of reactions affected per IEM (rank ordered). We used the gene-protein-reaction associations present in Recon1_Ac/FAO to determine the number of reactions catalyzed by enzymes in Recon1_AC/ FAO that is known to be deficient in IEMs. (B) Many IEMs are connected in the metabolic reconstruction through shared metabolites. We use the highly connected, important central metabolite acetyl-CoA to illustrate the metabolic connectivity of the IEMs (shown adjacent to the respective metabolisms) within the human metabolic network. Malonyl-CoA, which is formed from acetyl-CoA via the action of acetyl-CoA carboxylase (E.C. 6.4.1.2), participates in fatty acid synthesis, and a potent inhibitor of CPT-I enzyme (E.C. 2.3.1.21). Malonyl-CoA levels are regulated by malonyl-CoA decarboxylase (E.C. 4.1.1.9). All together, these enzymes maintain fine-tuning between fatty acid breakdown and synthesis.

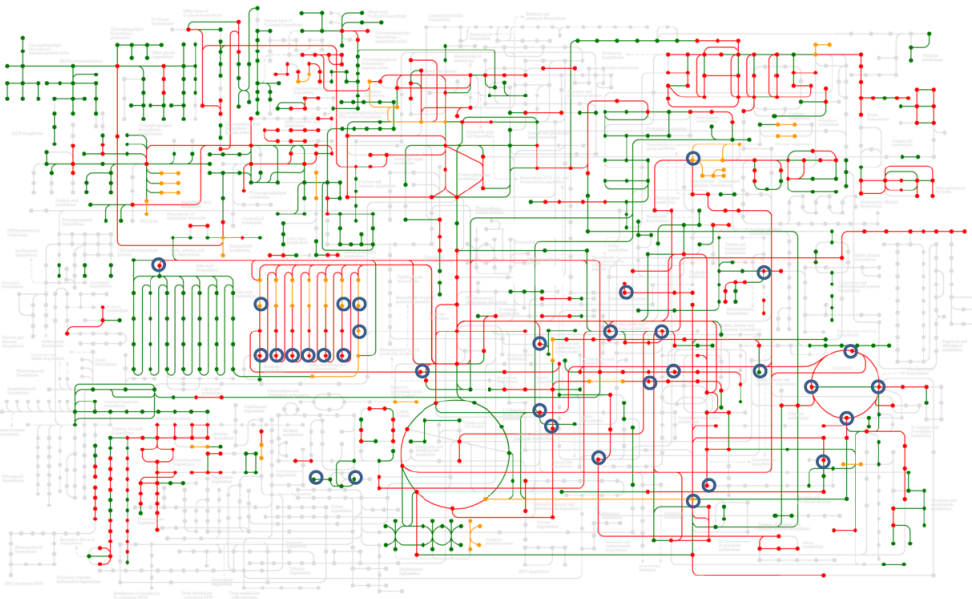


Figure 4.5: Visualization of 375 IEMs in our IEM compendium on the human metabolic map. This metabolic map, obtained from Kegg,104 gives an overview of current knowledge about human metabolism based on the human genome and known biochemistry. We map the genes from Recon1_Ac/FAO (green), the genes underlying the 235 mapped IEMs (red), and the genes corresponding to the 139 uncharted IEMs (orange). Note that 245 Recon1_Ac/FAO genes and 55 uncharted IEM genes could not be mapped. The blue circles represent the different amino acids and acylcarnitines, which are routinely measured in the newborn screening program. This map represents only a subset of the metabolic pathways captured in Recon1_Ac/FAO; we could only map 14 acyl-CoAs and 17 amino acids.

4.3.4 Simulation of phenylketonuria with the metabolic model of Recon1_Ac/FAO

Phenylketonuria (PKU, OMIM: 261600) is caused by mutations in *PAH* (GeneID: 5053) that may lead to complete or partial loss of the phenylalanine hydroxylase activity (E.C. 1.14.16.1). The enzyme catalyzes hydroxylation of phenylalanine to tyrosine (Table 4.1). We chose PKU to demonstrate how a model derived from Recon1_Ac/FAO can be used to simulate enzymopathies and to predict potential changes in biofluid concentrations. Therefore, we deleted the model reactions associated with the PAH gene, as defined through the GPRs, to simulate the consequences of complete loss of enzyme activity. The model was allowed the uptake and secretion of all metabolites with defined exchange reactions. We then performed flux balance analysis [11], where the exchange reaction for L-tyrosine or L-phenylalanine was used as an objective function. As expected, L-tyrosine secretion rate was zero in the PKU model, while L-tyrosine secretion was possible in the model corresponding to the healthy state. This result is consistent as PKU patients usually have low plasma tyrosine levels (OMIM: 261600), which is most often seen due to low enzyme activity. In contrast, no change in L-phenylalanine uptake or secretion was calculated in the PKU model compared to the healthy model. Thus, the model could not accurately represent the accumulation of phenylalanine (in form of phenylalanine secretion), which has been observed in plasma, urine, and cerebrospinal fluid of PKU patients (OMIM: 261600). The accumulation is due to decreased utilization of phenylalanine by the phenylalanine hydroxylase, while continuous supply of dietary phenylalanine. As we do not explicitly model the dietary intake and human cells cannot synthesize phenylalanine, we did not calculate increase in net phenylalanine secretion.

4.4 Conclusion

In this study, we generated an extended metabolic reconstruction, Recon1_Ac/FAO, which accounts for all metabolic and transport reactions necessary to enable the mapping of biomarkers measured in newborn screening programs [4,255]. Furthermore, we compiled a comprehensive compendium of 235 IEMs that have all causal genes captured in Recon1_Ac/FAO, containing also important phenotypic, genetic, and clinical features for each IEM.

A challenge that we encountered during the reconstruction process was the correct identification of some acylcarnitines, as MS/MS measurements do not always provide sufficient chemical details about the unsaturated fatty acids (i.e., positional isomers). To enable the inclusion of more metabolites and corresponding reactions

to the reconstruction, more precise measurements are needed.

The IEMs of the compendium were categorized based on different criteria to assess to which extent they affect whole body and the global metabolism. For instance, the organ classification will be of additional value when tissue- and cell type specific reconstructions and models are employed rather than a generic human metabolic network. A current challenge in computational modelling of metabolism is how one can generate in a (semi-) automated manner tissue- and cell-type specific reconstructions based on 'omics' data (e.g., transcriptomics) [311]. The organ classification of IEMs could be used to further refine tissue-specific metabolic reconstruction. Moreover, as the simulation of PKU highlighted, the use of a tissue-specific model rather than the generic model derived from Recon1_Ac/FAO and more defined boundary constraints will further increase the predictive potential of the effects of IEMs on the overall metabolism and metabolite biomarkers.

In addition to the mapped 235 IEMs, we identified 139 'uncharted' IEMs, which could not be mapped onto Recon1_Ac/FAO due to missing genes. These IEMs were classified under metabolic and non-metabolic IEMs (e.g., causing a defect in regulatory or signalling pathways), depending on the chief mechanism involved. The uncharted metabolic IEMs will be a good starting point for further extension and refinement of the human metabolic reconstruction. For instance, the mapping of inherited disorders of lipid metabolism will require further extension of the current human metabolic reconstruction as only 51 IEMs of the lipid metabolism could be mapped. Numerous new disorders have been identified [259] and many of them are fatal, therefore, this class of IEM will be an important future extension of the human metabolic reconstruction. In contrast, the inclusion of the non-metabolic IEM will require the expansion of the metabolic reconstruction for further cellular processes, such as signalling [312–314], macromolecular synthesis [64, 315, 316], and transcriptional regulation [64, 317, 318].

The presented extension to the current human metabolic reconstruction is not only important and relevant to map and analyze newborn screening data but also to map large-scale metabolomic data sets of plasma and urine from defined patient groups. For instance, serum metabolite concentration data for more than 1000 participants have been recently published, along with their genetic variants, from the KORA population [319]. Thanks to our extension, a majority of these metabolite measurements could be mapped and analyzed within the metabolic network context.

Metabolism has been found to be abnormal in many diseases. In this study, we identified potential IEM hubs in the metabolic network, which are connected by interlinking metabolites. The use of comprehensive mathematical metabolic models for the analysis of IEMs is promising, especially, for those IEMs, which have currently no diagnosis. Moreover, computational modelling may be particularly valuable for

those IEMs that are very rare and have currently no promising diagnostic and therapeutic regime.

4.5 Materials and methods

4.5.1 Routinely measured biomarkers in newborn screening data

We used information obtained from the Landspitali (Icelandic National University Hospital) for routinely measured biomarkers, which include 18 standard and four non-standard amino acids, (i.e., argininosuccinate, methylhistidine, ornithine and citrulline), 36 acylcarnitines, and 23 ratios (of amino acids and acylcarnitines). Furthermore, we considered three additional acylcarnitines (C14:1-OH, C16:2-OH, C14:2-OH) to be mapped onto the human reconstruction. Note that measured methylhistidine concentrations cannot be mapped onto the human reconstruction. We retrieved the molecular formulae and chemical structure for each metabolite from literature and/or databases.

4.5.2 Human metabolic reconstruction

The published human metabolic reconstruction, Recon 1 [18], was obtained from the BiGG database [320] in SBML format [321] and loaded into our reconstruction tool, rBioNet [36], using the COBRA toolbox [13]. Recon 1 consists of 3743 reactions, 2766 metabolites, and 1905 transcripts (corresponding to 1495 unique genes). Genes are connected to their respective reactions based on Boolean logic by defining gene-protein-reaction (GPR) associations for each network reaction. AND signifies that two gene products are required to carry out one reaction (enzyme complex), while OR stands for either gene product is sufficient for the reaction (isozyme).

4.5.3 Ac/FAO reconstruction module

The reconstruction was done using an established protocol [10] and the rBioNet reconstruction extension [36] for the COBRA toolbox [13]. Beside information retrieved from more than 150 primary publications and books, we used numerous databases. Reaction directionality information was obtained from literature and/or from thermodynamic considerations [322–324]. We employed the COBRA toolbox to combine the Ac/FAO module with Recon 1, and named the resulting expanded

reconstruction Recon1_Ac/FAO.

4.5.4 IEM collection and analysis

Candidate IEMs were mapped onto Recon1_Ac/FAO using GPR associations in Recon1_Ac/FAO and reported affected gene(s) of an IEM. Phenotypic and clinical information for each IEM was retrieved from primary literature, books, and databases.

4.5.5 Network based analysis of IEMs

Using the GPR associations defined in Recon1_Ac/FAO, we mapped all 235 IEMs onto the network. The reconstruction was converted into a mathematical model using the COBRA toolbox. The resulting stoichiometric matrix, S, has m rows, one for each metabolite, and n columns, one for each reaction. If a metabolite i participates in a reaction j, then the cell S(i,j) has a non-zero entry, where the number corresponds to the stoichiometric coefficient of the corresponding reaction j. By definition, substrates have a negative coefficient, while products have a positive coefficient. Similarly, the model of Recon1_Ac/FAO contains a matrix describing the links between genes and reactions, deemed G. Using G, we mapped the IEMs from genes to reactions, resulting in a new matrix M, in which each row corresponds to an IEM and each column corresponds to a reaction. Using a binary version of M (Mbin), we multiplied Mbin with its transpose. Each diagonal element of this adjacency matrix of Mbin corresponds to the number of reactions that are connected with each IEM. The off-diagonal elements corresponding to the number of reactions shared between two IEMs, which allowed us to identify reaction pairs. Similarly, by multiplying the Mbin with the transpose of the binary version of S, Sbin, we could determine metabolites shared between two IEMs. For the visualization of all IEMs, we employed the pathway viewer from the KEGG database [35].

4.5.6 Simulation of PKU

Recon1_Ac/FAO was converted into a computational model using the COBRA toolbox [13]. The reaction list of the reconstruction was therefore converted into a so-called stoichiometric matrix, S, consisting of m rows (one row for each metabolite) and n columns (one for each reaction in the network). If a metabolite i participates in a reaction j, then the entry Si,j contains the corresponding non-zero stoichiometric coefficient, which is a negative number for substrates and a positive number for

reaction products. Constraint-based modelling assumes that the modelled system is in a quasi-steady state, i.e., the change of concentration of a metabolite i is zero over time t:

$di/dt = S.v \equiv 0$,where v is the flux vector containing a flux value vi for each reaction i. In flux balance analysis, we generally optimize for an objective reaction, thus, identifying one possible flux vector with optimal value in the objective function that is consistent with the applied constraints [9, 11]. Constraints can be applied to any reaction i in the model, such that $v_{i,min} \leq v_i \leq v_{i,max}$.To simulate PKU, we set the constraints of the reaction PHETHPTOX2, catalyzed by PAH, to $V_{PHETHPTOX2,min} = V_{PHETHPTOX2,max} = 0$ mmol.gDW-1.hr-1, while the constraints through this reaction in the healthy model were unchanged. Additionally, for the healthy and PKU models, we set the lower bounds (vi,min) of the exchange reactions for water, carbon dioxide, proton, bicarbonate, ammonium, inorganic phosphate and sulphate to -100 mmol.gDW-1.hr-1, and the lower bound for oxygen uptake was set to be -40 mmol.gDW-1.hr-1. The lower bounds of the remaining exchange reactions were set to be -1 mmol.gDW-1.hr-1. The upper bounds of all exchange reactions were unconstrained, i.e., $v_{i,max} = 1000$ mmol.gDW-1.hr-1. In both models, we selected the exchange reaction for L-tyrosine and L-phenylalanine separately as objective reactions for maximisation and minimisation. The calculations were carried out using Matlab (Mathworks, Inc.) as programming environment and TomLab (TomOpt, Inc) as linear programming solver.

Table 4.2: Metabolically linked IEMs with known overlapping phenotypes. The IEM pairs are grouped together

IEM pairs	Genes involved	Pathway involved	Enzyme (E.C. number)	No. of shared reactions	Remarks
Tay-Sach's disease	HEXA (GeneID: 3073)	Carbohydrate metabolism	Beta-N-acetyl-hexosamini-dase (E.C. 3.2.1.52)	51	HEXA gene encodes the alpha subunit, whereas HEXB encodes the beta subunit of the enzyme. Even though different mutations in these two genes lead to Tay-Sachs disease and Sandhoff disease, these diseases are clinically indistinguishable [307].
Sandhoff disease	HEXB (GeneID: 3074)	Carbohydrate metabolism	Beta-N-acetyl-hexosamini-dase (E.C. 3.2.1.52)	51	
Sialidosis	NEU1 (GeneID: 4758)	Carbohydrate metabolism	Sialidase (3.2.1.18)	34	Common clinical features in both, i.e., coarse features and neurologic abnormalities. Additionally, the structural changes in Morquio syndrome may lead to nervous symptoms, which are also seen in sialidosis [309].

IEM pairs	Genes involved	Pathway involved	Enzyme (E.C. number)	No. of shared reactions	Remarks
Mucopolysaccharidosis type IV type A / Morquio syndrome	GALNS (GeneID: 2588)	Carbohydrate metabolism	Chondroitinase (E.C. 3.1.6.4)	34	
NICCD (Neonatal intrahepatic cholestasis caused by citrin deficiency)	SLC25A13 (GeneID: 10165)	Transport	Calcium-binding mitochondrial carrier protein Aralar2	3	Same gene is affected in NICCD and citrullinemia type -2. Citrullinemia type 2 may develop after several years of being asymtopmatic in NICCD patients [70, 325].
Type II citrullinemia	SLC25A13 (GeneID: 10165)	Amino acid metabolism	Calcium-binding mitochondrial carrier protein Aralar2	3	

IEM pairs	Genes involved	Pathway involved	Enzyme (E.C. number)	No. of shared reactions	Remarks
Erythrocyte AMP deaminase deficiency	AMPD3 (GeneID: 272)	Nucleotide metabolism	AMP deaminase (3.5.4.6)	1	The enzyme is encoded by different isoforms, which depends on tissue specific expression. Low plasma uric acid level but no apparent clinical signs are observed in cases of Erythrocyte AMP deaminase deficiency [326]. The defect in the muscle form is one of the most common types of muscle enzyme defects. It exhibits myalgia [327] and is linked to obesity and diabetes [328].
Myoadenylate deaminase deficiency	AMPD1 (GeneID: 270)	Nucleotide metabolism	AMP deaminase (3.5.4.6)	1	
Liver glycogen synthase deficiency	GYS2 (GeneID: 2998)	Carbohydrate metabolism	Glycogen (starch) synthase (E.C. 2.4.1.11)	2	Glycogen synthase has two isozymes (liver and muscle isoform). Liver glycogen synthase deficiency is comparatively rare and associated with hypoglycemia, mental retardation and seizures [329–331]. The muscle form is associated with cardiomyopathy and exercise intolerance [332].

IEM pairs	Genes involved	Pathway involved	Enzyme (E.C. number)	No. of shared reactions	Remarks
Muscle glycogen synthase deficiency	GYS1 (GeneID: 2997)	Carbohydrate metabolism	Glycogen (starch) synthase (E.C. 2.4.1.11)	2	
Distal renal tubular acidosis type 1	SLC4A1 (GeneID: 6521)	Transport	Band 3 anion transport protein	1	Band 3 protein is a multi-spanning membrane protein in the red blood cell [333]. Genetic mutations in the SLC4A1 gene result in two different types of defect, i.e. hereditary spherocytosis and distal renal tubular acidosis type 1. These diseases may co-occur in an affected individual [334]. Specific reasons for this co-occurrence is still not known but very frequently reported [334, 335].
Hereditary spherocytosis-4	SLC4A1 (GeneID: 6521)	Transport	Band 3 anion transport protein	1	

IEM pairs	Genes involved	Pathway involved	Enzyme (E.C. number)	No. of shared reactions	Remarks
Renal tubular acidosis, distal, autosomal recessive	ATP6V0A4 (GeneID: 50617)	Transport	V-type proton ATPase 116 kDa subunit a isoform 4	1	The affected genes encode for different subunits of the V-ATPase pump. This pump mediates intracellular acidification and consists of a V0 (A, B, C, D, E, F, G and H subunits) and a V1 domain (a, c, c', c", and d subunits). The ATP6V0A4 gene encodes the a-subunit of the V1 domain, whereas the ATP6V1B1 gene encodes B-subunit of the V0 domain (Gene ID 50617 and 525). Similar clinical features are observed in renal tubular acidosis (distal, autosomal recessive) and RTA.
Renal tubular acidosis (RTA)	ATP6V1B1 (GeneID: 525)	Transport	V-type proton ATPase subunit B, kidney isoform	1	

5 Predicting the impact of diet and enzymopathies on human small intestinal epithelial cells

5.1 Abstract

Small intestinal epithelial cells (sIECs) have a significant share in whole body metabolism as they perform enzymatic digestion and absorptions of nutrients. Furthermore, diet plays a key role in number of complex diseases including obesity and diabetes. The impact of diet and altered genetic backgrounds on human metabolism may be studied using computational modeling. A metabolic reconstruction of human sIEC was manually assembled using literature. The resulting sIEC model was subjected to two different diets to obtain condition-specific metabolic models. Fifty defined metabolic tasks evaluated the functionalities of these models, along with the respective secretion profiles, which distinguished between impacts of different dietary regimes. Under the average American diet, the sIEC model resulted in higher secretion flux for metabolites implicated in metabolic syndrome. In addition, enzymopathies were analyzed in the context of the sIEC metabolism. Computed results were compared with reported gastrointestinal pathologies and biochemical defects as well as with biomarker pattern used in their diagnosis. Based on our simulations, we propose that i) sIEC metabolism is perturbed by numerous enzymopathies, which can be used to study cellular adaptive mechanisms specific for such disorders, and in identification of novel co-morbidities, ii) porphyrias are associated with both heme synthesis and degradation, and iii) disturbed intestinal gamma-aminobutyric acid synthesis may be linked to neurological manifestations of various enzymopathies. Taken together, the sIEC model represents a comprehensive, biochemically accurate platform for studying the function of sIEC and their role in whole body metabolism.

5.2 Introduction

The major purpose of the human digestive system is to process food to provide the body with essential nutrients and energy [83]. After being partially digested at the

level of mouth and stomach, the food components (i.e., 60-70% of complex carbohydrates, 80-90% of dietary proteins, and approximately 85% fat) reach the duodenum, the proximal part of the small intestine, for complete enzymatic digestion and absorption [336]. The small intestine is unique in its variety of cell types, including enterocytes (i.e., columnar small intestinal epithelial cells for digestion and absorption of nutrients), goblet cells (for mucus secretion), entero-endocrine cells (for hormone secretion), and paneth cells (i.e., stem cells that differentiate to form enterocytes). The large surface area, provided by the presence of hundreds of enterocytes, contributes to the maximal absorptive capacity of the small intestine.Small intestinal enterocytes account for the majority of the enzymatic digestion and nutrient absorption [83]. In fact, enterocytes are highly metabolically active cells and provide approximately 25% of the endogenous glucose and cholesterol [38, 39]. As such, they contribute significantly to the metabolism of the whole body, acting as a gateway for nutrients. They channel essential nutrients to the liver (upon hormonal influence) as well as account for first pass drug metabolism [337]. The liver and the small intestine are anatomically in close proximity, connected through the portal vein, and they have related physiological and metabolic functions. The small intestine provides approximately 75% of blood flow to the liver via the hepatic portal system [338] and the liver, in turn, supplies biliary constituents through the common bile duct to the duodenum. The liver synthesizes bile acids from cholesterol, which are then delivered to the intestinal lumen to aid in the digestion and absorption of fat with their emulsifying properties. The bile acids are then actively absorbed by the enterocytes, but half of the bile acids can also diffuse through the lumen into the enterocytes. Bile acids are then sent back to the liver via the portal circulation, permitting their extensive recycling through this entero-hepatic circulation before they are finally excreted in feces [75].

Inborn errors of metabolism (IEMs) are the hereditary metabolic defects that are encountered in all major metabolic pathways occurring in man [250]. IEMs have a myriad of pathological effects, affecting multiple organ systems that may lead to fatal phenotypes. IEMs can arise due to mutations in single genes [339] or multiple genes [340], which add another level of complexity to their diagnosis. Mass spectrometric analysis of whole blood samples from infants is the usual diagnostic method (e.g., looking for concentration changes of specific biomarkers). However, there exists a series of other tests, including molecular genetic testing for specific mutations, enzyme assays, and further biochemical tests (e.g., urine tests and blood tests for blood gases and electrolytes), which together are employed to confirm the presence of an IEM [341]. There exist multiple classification systems for these disorders depending on the clinical phenotypes, affected organs, mode of inheritance, occurrence in a specific metabolic pathway, and other factors. While IEMs of the amino acid metabolism are usually treatable, IEMs involving biosynthesis of complex molecules (e.g., lysosomal storage or peroxisomal biogenesis disorders) have generally no specific treatment available [250]. Use of special dietary formulations

and medications usually alleviates clinical symptoms of the aminoacidopathies, and hence, is widely used as treatment strategy [250].

Genome-scale metabolic network reconstructions capture genomic, physiological, and biochemical accurate data for target organisms [5]. More than 100 organism specific metabolic reconstructions have been published, including for human [18] and mouse [25]. Moreover, cell-type specific reconstructions have been published for hepatocytes [21], for kidney cells [22], for alveolar macrophages [24], for red blood cells [23], and for cardiomyocytes [28]. Metabolic reconstructions serve as a knowledge-base for a target organism and target cell-type and can be converted into mathematical models [5]. Besides describing the exact stoichiometry for metabolic reactions occurring in a cell, these reconstructions also contain information about the enzymes catalyzing the reactions and the corresponding encoding genes through gene-protein-reaction (GPR) associations. These GPR associations are expressed using Boolean rules ('and' and 'or' relationships), where 'and' represents the requirement of the gene products for a reaction. Conversely, the 'or' relationship is employed to represent the relationship between genes encoding for isozymes, which catalyzes the same reaction. These GPRs permit to connect the genotype to the phenotype and enable prediction of phenotypic changes associated with genetic alterations, e.g., enzyme deficiencies [19, 25, 29, 53] or single nucleotide polymorphism [342]. An increasing number of studies have become available highlighting the value of computational modeling for understanding metabolism in health and disease. These studies include the analysis of co-morbidity [310], the prediction off-target drug effects [22], the prediction of specific drug targets [27, 343], the assessment of metabolic changes occurring during diabetes and ischemia [344], and the prediction of the metabolic consequences of genetic and epigenetic properties, such as dosage effect [345].

The aim of the present study was to model the digestion and absorption of dietary components by the small intestine and to analyze changes in the intestinal metabolism under different genetic backgrounds. We therefore reconstructed a metabolic network specific for small intestinal epithelial cells and subjected it to two different diets to obtain condition specific metabolic models. Subsequently, we analyzed the effect of various IEMs on the models' metabolic capabilities. For IEMs that highly affected the enterocyte's metabolism, we compared the computed results with known intestinal features, such as phenotypes exhibiting extensive intestinal pathology associated with these disorders as well as with biomarker pattern used in their diagnosis.

5.3 Results

In this study, we assembled manually a comprehensive, predictive metabolic network for small intestinal enterocytes, which captures known, essential physiology and biochemistry of this cell type. We then employed the metabolic model to assess the impact of diet on the absorption and secretion capabilities of enterocytes. Finally, we investigated the consequences of enzymopathies on the enterocyte's metabolic functionalities.

5.3.1 Reconstruction of the sIEC specific metabolic network

We reconstructed the first small intestinal epithelial cells (sIECs) specific metabolic network in a bottom-up fashion through extensive literature review. We identified metabolic pathways known to occur within these cells, including sIEC-specific pathways, such as citrulline and proline metabolism, the conversion of fructose to glucose, and the re-synthesis of tri-glyceride and cholesterol-ester. We collected information on the various transport systems specific for sIECs (Figure 5.1B), resulting in the addition of 38% new enzymatic and transport reactions, which were not captured by a previously published global human metabolic reconstruction [18] (Figure 5.1E). The final human small intestine specific enterocyte metabolic reconstruction was deemed hs_sIEC611, where 'hs' stands for homo sapiens, sIEC for small intestinal epithelial cell, and 611 for the number of included genes. hs_sIEC611 accounts for 1282 reactions and 433 unique metabolites distributed over five intracellular (cytosol, mitochondria, nucleus, peroxisome, and endoplasmic reticulum) and two extra-cellular compartments (extracellular space and lumen) (Figure 5.1A). Information from more than 250 peer reviewed articles and books were reviewed and provide supporting evidence for the metabolic content of hs_sIEC611.

When comparing the number of reactions associated with the different metabolic pathways in sIECs (Figure 5.1B), we found that the transport subsystem had the highest number of reactions associated (538; Figure 5.1D and 5.1E), since a chief function of the enterocyte is the absorption of dietary nutrients from the lumen (apical uptake). Additionally, enterocytes take up nutrients from the arterial blood (basolateral uptake) depending on their own metabolic need. Numerous metabolites, including bilirubin, nitric oxide, and cholesterol, are also secreted into the lumen. Nutrient transporters are normally located either at the luminal or basolateral side, depending on the nature of the transporter. An exception is the GLUT-2 transporter, which is usually present at the basolateral side, but which can also migrate onto the luminal side to aid in glucose absorption under fully fed state [346]. Moreover, the lipid metabolic pathway accounted for most metabolic reactions (134 reactions),

followed by nucleotide metabolism (113 reactions), and carbohydrate metabolism (70 reactions).

To evaluate the predictive potential of a condition-specific model that can be derived from the sIEC-specific reconstruction, we tested for the model's capability to produce all defined biomass precursors under rich medium conditions (i.e., all metabolites, for which exchange reactions were defined, were allowed to be taken up) as well as under two defined dietary conditions, being an average American diet and a balanced diet. Moreover, we performed an extensive literature review to collect metabolic capabilities that can be accomplished by sIECs and formulated 50 different 'metabolic tasks', which sIECs, and thus our model, should be able to fulfill (Table 5.1). We then tested the model's capability to carry a non-zero flux for each metabolic task under different simulation conditions using flux balance analysis [11]. We also defined a sIEC-specific secretion profile based on thorough literature review. A total of 267 exchange reactions were constrained accordingly for luminal uptake-secretion and basolateral uptake-secretion pattern. We subsequently tested the model's capability to carry flux through the exchange reactions using flux variability analysis [12] and compared the results with the defined secretion profile.

5.3.2 Effects of different diets on the metabolic tasks of sIECs

We were interested in investigating the consequences of different diets on the sIEC metabolism. Therefore, we simulated an average American [347] and a balanced diet [83, 348, 349] and tested the model's capability to perform the 50 different metabolic tasks (Table 5.1). When comparing the results for the maximal possible flux values of the different metabolic tasks, we found increased flux for the synthesis of important products, such as glucose, ornithine, and citrulline, from glutamine in the average American diet compared with the balanced diet (Table 5.1). Simultaneously, the glutamine oxidation flux to carbon dioxide was higher under the average American dietary regime, reflecting the higher glutamine availability in this diet. The conversion of fructose to glucose was not possible in the balanced diet due to the absence of fructose in this diet [83]. The synthesis flux of 5-methyltetrahydrofolate (5-methyl-THF) was higher in the balanced diet compared to the average American diet due to the influx of its precursor folate in the simulation of the balanced diet. 5-methyl-THF is an important one carbon substituted folate and participates in amino acid and purine metabolism. Further, when compared to a balanced diet, the average American diet caused a higher flux of luminal inputs, which was particularly the case for fat (i.e., triglycerides and cholesterol) and for carbohydrates. This increased input flux led to a 45% higher flux through the biomass reaction in the average American diet. Another noteworthy observation was the failure to synthesize alanine from glutamine in absence of glucose, which we observed for both diets

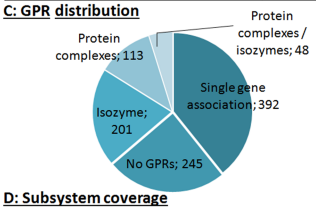
A: Reconstruction statistics

Total number of reactions	1282
Metabolic and transport	999
Exchange/demand/sink/biomass	267/14/1/1
Unique metabolites	433
Unique Genes	611
Intra-cellular compartments	5
Extra-cellular compartments	2
Metabolic objective functions	50
Secretion profile (basolateral secretion)	103
References	275*

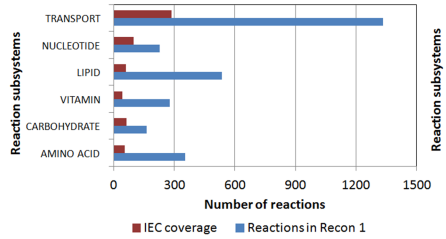
B: Pathway coverage

Metabolic category	Pathways occurring within the small intestinal enterocyte reconstruction
Carbohydrate	Glycolysis, TCA, HMP, gluconeogenesis, fructose metabolism.
Dietary fibers	Luminal binding to bile salts and their subsequent excretion.
Amino acid	Glutamine, aspartate, arginine, ornithine, methionine, cysteine, histidine, proline, urea cycle.
Di and tri peptides	Digestion in the lumen, transport of released amino acids into the enterocyte.
Lipid	Fatty acid beta oxidation, tri-acylglycerol synthesis, choleterol ester synthesis, phospholipid synthesis, cholesterol synthesis and excretion.
Vitamins	A, D, B1, B2, B3, B6, folic acid, C, pantothenic acid.
Nucleotide	Purine salvage, de-novo synthesis and catabolism. Pyrimidine synthesis and catabolism.
ETC	Oxidative phosphorylation.
Heme	Heme synthesis and degradation.
Others	Re-absorption and secretion of bile acids, luminal digestion of macronutrients, ubiquinone, carboxylic acid dissociation, nitric oxide synthase.

C: GPR distribution



D: Subsystem coverage



E: Addition of novel reactions per subsystem

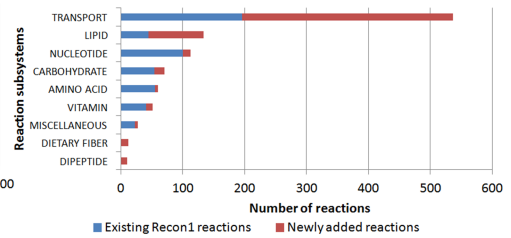


*Figure 5.1: Properties of the human metabolic reconstruction for small intestine epithelial cells. A: Overview of the sIEC reconstruction. *peer-reviewed journal articles, primary literature and books. B: Metabolic pathways that have been identified through thorough literature review to be present in sIECs. C: Overview of the GPR associations in the enterocyte reconstruction. D: Coverage of metabolic and transport reactions in the enterocyte reconstruction with respect to Recon 1 (exchanges, luminal adjustments and novel metabolic and transport reactions existing in the enterocyte network have not been shown). E: Addition of novel reactions to the enterocyte reconstruction, di-peptides, and dietary fibers were two subsystems that were absent in the previous human metabolic reconstruction, Recon 1 [18].*

(when glucose was removed under this particular simulation condition). Literature review revealed that the major source of pyruvate, which serves as the precursor for lactate and alanine production, is glucose rather than glutamine in sIECs. This conversion route has been demonstrated by 13C-nuclear magnetic resonance studies [38].

To evaluate potential effects of intestinal nutrient absorption on whole body metabolism, we analyzed secretion patterns obtained from the two diets. We performed flux variability analysis for the models corresponding to the two diets and compared the computed maximum flux value of each exchange reaction on the basolateral side. A total of 80 metabolites showed differences in their secretion profiles between the two diets. For instance, higher flux through the secretion reactions for vitamin E (two fold), choline (1.4 fold), and retinol (two fold) was observed when the balanced diet was compared to the average American diet. The balanced diet is designed to provide essential vitamins (both water and fat soluble ones) and minerals, which should be principally also provided in an average American diet. However, the model failed to secrete components, such as biotin, pantothenic acid, and chloride, since these components were missing in the average American diet chart [347]. Interestingly, we observed a twofold elevation in maximal possible flux values for the secretion of cholesterol esters, and triglycerides when the model was provided with the average American diet compared to the balanced diet. Additionally, the maximal possible secretion flux values for cholesterol and glucose were four and three fold higher, respectively, in the average American diet than in the balanced diet. When we compared the essential fatty acids secretion pattern, linoleic acid (C18:2) secretion was twofold higher in the average American diet, whereas linolenic acid (C18:3) secretion was twofold higher in the balanced diet.

Table 5.1: Effect of diet on the metabolic tasks defined for sIECs. D1 = Average American diet. D2 = Balanced diet. ✓ *- represents corresponding task was performed and the flux value obtained for both the diets were similar,* ↑ *represents higher flux values obtained in American diet when compared to balanced diet, and* × *represents task could not be performed by either diet*

Metabolic task	Physiological function	D1 vs D2
Secretion of lactate from glucose uptake	Glucose utilization for energy production.	↑
Glutamine to glucose conversion	Endogenous glucose production.	↑
Glutamine to proline conversion	Delivery to the liver.	↑
Glutamine to ornithine conversion	For spermine synthesis.	↑

Metabolic task	Physiological function	D1 vs D2
Glutamine to citrulline conversion	For arginine synthesis. Biomarker of active enterocyte mass [350]. This metabolic task is exclusive to enterocytes [351]	↑
Glutamine to lactate conversion	Utilization of glutamine, sparing glucose.	✓
Glutamine to aspartate conversion	For nucleotide synthesis.	✓
Glutamine to carbon dioxide conversion	Utilization of glutamine, sparing glucose.	↑
Glutamine to ammonia conversion	Glutaminase activity to generate glutamate.	✓
Putrescine to methionine conversion	Regeneration of methionine.	↑
Basolateral secretion of alanine	Delivery to the liver.	↑
Basolateral secretion of lactate	Delivery to the liver.	↑
Synthesis of arginine	Semi-essential amino acid.	↑
Synthesis of proline	To be sent to liver. The pyroline-5-carboxylate synthase is specific to the small intestine [352].	✓
Synthesis of alanine from glutamine	To be sent to the liver.	✗
Basolateral secretion of proline	To be sent to the liver.	↑
Basolateral secretion of arginine	To be sent to the liver.	↑
Basolateral secretion of ornithine	To be sent to the liver.	↑
Synthesis of spermine	Protective effect in small intestine.	↑
Synthesis of spermidine	To be sent to the liver.	↑
Synthesis of nitric oxide	Aids in GI motility.	↑
Synthesis of cholesterol	Membrane constituent.	↑
De novo purine synthesis	Energy intermediate.	↑
Salvage of purine bases	Energy intermediate.	↑
Purine catabolism	Energy intermediate.	↑
Pyrimidine synthesis	Maintenance of purine level and important by-product release.	↑
Pyrimidine degradation (uracil)	Maintenance of purine and pyrimidine levels and important by-product release.	↑
Fructose to glucose conversion	Fructose utilization.	↑

Metabolic task	Physiological function	D1 vs D2
Uptake and secretion of cholic acid	Enterohepatic circulation.	↑
Uptake and secretion of glyco-cholic acid	Enterohepatic circulation.	↑
Uptake and secretion of tauro-cholic acid	Enterohepatic circulation.	↑
Hexose mono-phosphate shunt pathway	Generation of reducing equivalents and precursor for nucleotide synthesis.	↑
Malate to pyruvate conversion	Maintenance of cytosolic pyruvate pool	↑
Synthesis of urea	Elimination of ammonia.	↑
Cysteine to pyruvate conversion	Maintenance of cytosolic pyruvate pool.	↑
Methionine to cysteine conversion	Synthesis of cysteine.	↑
Synthesis of triacylglycerol	To be transported in chylomicrons for body needs.	↑
Phosphatidylcholine synthesis in mitochondria	Membrane constituent.	↑
Binding of bile acids to dietary fibers	Cholesterol lowering effect of dietary fibers.	✓
Synthesis of FMN	Conversion of riboflavin to active form.	↑
Synthesis of FAD	Fatty acid oxidation.	↑
Synthesis of 5-methyl-tetrahydrofolate	Generated from folic acid, an important co-factor in one carbon metabolism.	↓
Putrescine to GABA conversion	Exerts protective effect in small intestine.	↑
Superoxide dismutase activity	Free radical scavenger.	✓
Carbonic anhydrase activity	For availability of bicarbonate.	↑
TCA cycle flux	Energy generation, anaplerosis, and amphibolic role.	↑
Histidine to form-imino-glutamic acid conversion	One-carbon metabolism.	↑
Heme synthesis	Prosthetic part of hemoglobin.	↑
Heme degradation	For availability of iron.	↑
Cell maintenance	Cell viability.	↑

5.3.3 Impact of enzymopathies on the enterocyte metabolism

To study the difference in the sIEC metabolism during healthy and diseased states, we analyzed the effect of IEMs on the enterocyte metabolism in the context of the average American diet. hs_sIEC611 captured 109 IEMs out of the 235 IEMs that we have recently mapped onto the global human metabolic reconstruction [29]. Of these, we considered for further analysis 103 IEMs, which are caused by mutations in a single gene, meaning that they had one gene-one reaction (66 IEMs) or one gene-many reaction (37 IEMs) associations. Additionally, we also analyzed the in silico effect of cystinuria (OMIM: 220100), which is caused by mutation in either solute carrier family 3 (*SLC3A1*, GeneID: 6519) or solute carrier family 7 (*SLC7A9*, GeneID: 11136). Cystinuria was included in the analysis as the intestinal dibasic amino acid transport system is defective in these patients [353], and thus evidence of the expression of both genes in human sIECs exists. For each IEM, we deleted the corresponding metabolic reaction(s) and computed the maximal possible flux value for each metabolic task. We found that 76% (79/104) IEMs affected at least one of the defined metabolic tasks in hs_sIEC611, when simulated on an average American diet (Figure 5.3). Thirty-six of these IEMs (45%) resulted in intestinal pathology or biochemical defect and correspond to reported true positive cases (Table 5.2). When compared to the reported biomarker patterns used for diagnosis of the simulated IEMs, the model failed to predict the elevated biomarker patterns for ten IEMs, including hyperprolinemia and glycogen storage disease XI. We also investigated the basolateral secretion profiles in the different IEM conditions (Figure 5.4). The model corresponding to carbamoyl phosphate synthetase I deficiency (OMIM: 237300) showed a reduced secretion of citrulline, which is consistent with the known reduced citrulline level in these patients. The methylmalonic acidemia model predicted reduced secretion flux for cysteine, which is in agreement with reduced cysteine levels seen in methylmalonic acidemia patients [354]. Analysis of the model's metabolic capability revealed that this reduced cysteine secretion flux is due to the model's inability to convert methionine to cysteine. The computational analysis thus provides a mechanistic explanation for the observed biomarker, which represents a novel testable hypothesis.

We then categorized the IEMs according to the metabolic pathways they occurred in. The highest number of IEMs was associated with carbohydrate metabolism disorders (23 IEMs), followed by amino acid metabolism disorders (19 IEMs), lipid metabolism disorders (18 IEMs), nucleotide metabolism disorders (nine IEMs), heme biosynthesis defects (eight IEMs), and two further IEMs (see also Figure 5.2). We then analyzed the computed effects of the IEMs on the defined metabolic tasks, which were either flux reduction (drop) or abolishment (block) (Figure 5.3). A total of eight IEMs affected only one metabolic task. The hexose monophosphate shunt (HMP) metabolic task had the highest reduction in flux value in 27 IEMs (Figure

5.3). Twelve IEMs caused reduction and five IEMs caused block in the cell maintenance/biomass function. Out of these 17 IEMs that affected the cell maintenance function, eight were found to be associated with intestinal pathology or intestinal biochemical defect (Table 5.2). Interestingly, reduction in the flux through the TCA cycle, synthesis of gamma-amino-butyric acid, and gluconeogenesis from glutamine were the other common effects observed in the simulation results for the IEMs affecting the cell maintenance task. Moreover, we predicted various IEMs to affect the transport function of the sIECs by causing reduction in alanine (10 IEMs), lactate (7 IEMs), proline (20 IEMs), arginine and ornithine secretion (6 IEMs each), or a block in arginine secretion (7 IEMs). The metabolic task for synthesis of phosphatidylcholine was affected by six IEMs, with all of them causing blocked flux (Figure 5.3). Interestingly, three of these IEMs have been associated with small intestinal pathology (Table 5.2) but no changes in phosphatidylcholine level have been reported.

Our model also predicted disabled heme synthesis in eight different porphyrias (Figure 5.3), which corresponds to the metabolic task to be eliminated by the highest number of IEMs. hs_sIEC611 also predicted a block in heme degradation in porphyrias, since; there is no dietary supply of heme in form of heme-proteins in the simulated average American and balanced diets. Additionally, we also observed a block in secretion of carbon monoxide and a reduction in urea synthesis and in urea secretion in the porphyria simulations (Figure 5.4).

When simulating the average American diet, glucose-galactose malabsorption (OMIM: 606824) affected in silico the highest number of metabolic tasks (i.e., 22 out of 50) (Figure 5.3), including reduced fluxes through gluconeogenesis from glutamine and the synthesis of gamma-aminobutyric-acid, i.e., GABA from putrescine. The GABA synthesis in particular was affected in 23 IEMs (Figure 5.3), with 17 of them exhibiting variable extents of neurological symptoms. Two IEMs, gyrate atrophy of the choroid and retina (OMIM: 258870) and ornithine translocase deficiency (OMIM: 238970), showed the highest reduction in flux value through synthetic reaction for GABA (17% of the flux value of the healthy model).

Under an average American diet, the model for type I citrullinemia (OMIM: 215700), which is defective in *ASS1* (GeneID: 445, E.C.6.3.4.5), led to increased flux through the citrulline basolateral secretion reaction and a single fumarate producing reaction ('ADSL1') and decreased flux through the fumarase ('FUM'). These predictions were consistent with the reported build-up of citrulline in patients [355]. Moreover, the model predicted increased flux through the luminal uptake reactions of two amino acid transporters, ASCT2 (*SLC1A5*, GeneID: 6510) and B(0)AT1 (*SLC6A19*, GeneID: 340024), and through the alanine transaminase (E.C.2.6.1.2.). The latter converts alanine to pyruvate, which subsequently enters gluconeogenesis providing endogenous glucose. Furthermore, the type I citrullinemia model was able to uti-

lize partially the urea cycle by synthesizing ornithine, which was then converted either to glutamate by the ornithine transaminase (E.C.2.6.1.13) or to citrulline by the ornithine carbamoyltransferase (E.C. 2.1.3.3).

Smith-Lemli-Opitz syndrome (OMIM: 270400) is a cholesterol synthesis disorder caused by a mutation in *DHCR7* (GeneID: 1717, E.C.1.3.1.21). Under the average American diet, this enzymopathy led to increased flux through 'nadp' and 'nad' utilizing reactions (i.e., 'G6PDH2c', 'GNDc', 'AKGDm', 'SUCD1m', and 'ICDHxm') and decreased flux through the basolateral cholesterol secretion reaction ('EX_chsterol(e)') in the corresponding model. Additionally, we predicted increased fluxes through the reaction of electron transport chain, gluconeogenesis, and purine salvage pathway reactions. These predicted changes were consistent with the reported low cholesterol level in patients [356].

Does the effect of IEMs on the metabolic capabilities of sIECs depend on diet? To address this question, we repeated the computations while simulating the balanced diet. We overall found the metabolic capabilities similar in enzyme-deficient sIECS under the balanced diet. Fifity-four of the IEMs resulted in altered fluxes for various metabolic tasks. For instance, while the biomass reaction was blocked in the cystinuria model under the average American diet, the flux value was identical to the healthy model under the balanced diet. In contrast, modeling the ribose 5-phosphate isomerase deficiency resulted in a block of the flux through the biomass reaction under balanced diet only. Moreover, IEMs of the amino acid metabolism (i.e., propionic acidemia, methyl malonic academia (MMA), cystathioninuria, and MMA type 3) and of the lipid metabolism (i.e., CACT deficiency and CPT-1 deficiency) showed the same flux value through the biomass reaction under balanced diet as the healthy model but had a lower flux value under the average American diet. The computed block or reduction in biomass reaction flux in these IEMs was due to limited uptake of methionine and proline in the case of cystinuria, and of alanine, asparagine, leucine, and phosphatidyl-ethanolamine for the IEMs of amino acid and lipid metabolism. The uptake limitation was a result of missing chloride ions in the in silico average American diet, as they were not listed in the corresponding diet chart [347]. Consequently, one of the intestinal amino acid transporter (ATB0, *SLC6A14*, GeneID: 11254), which has broad substrate specificity for both neutral and cationic amino acids [357] and requires the presence of sodium and chloride for its activity, could not be used in this diet. In contrast, the limiting component in the simulated balanced diet was ribose. Additionally, blocked flux through HMP, reduced fluxes through reactions involved in pyrimidine synthesis, in the conversion of glutamine to carbon dioxide, and in the purine catabolism pathway contributed to a block in the biomass reaction for cystinuria.

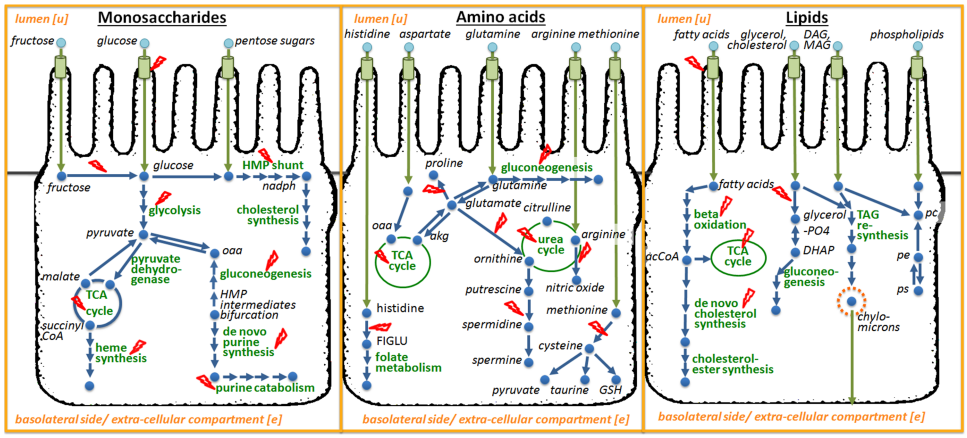


Figure 5.2: Overview of enterocyte metabolism and its associated IEMs (in red). Only pathways that are associated with IEMs are shown. A: Major pathways operating under carbohydrate metabolism. B: Amino acid metabolism. C: Lipid metabolism. The metabolic pathways appear in green and the metabolites in blue.

Figure 5.3: Metabolic tasks affected by the different IEMs, under average American diet. Black bars where the metabolic task could not carry any non-zero flux (flux block). Orange bars represent reduced flux through the metabolic task. White bars represent no change.

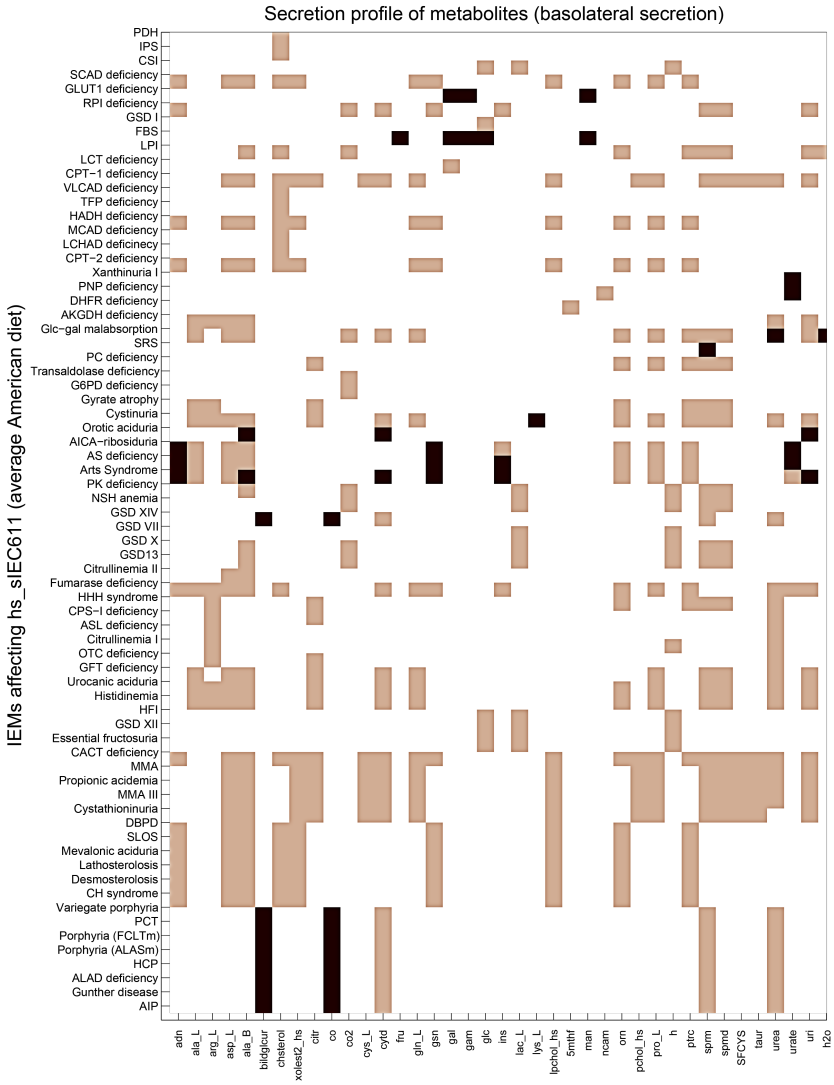


Figure 5.4: Basolateral secretion profile obtained for different IEMs. Only highly affected secretion reactions are reported. Black bars where the secretion reaction could not carry any flux (flux block). Pink bars represent reduced flux through the secretion reaction. White bars represent no change.

5.4 Discussion

In this study, we presented a comprehensive, predictive metabolic network for small intestinal enterocytes, which captures comprehensively its physiology and biochemistry. The key results include: i) the average American diet resulted in higher secretion flux for metabolites involved in metabolic syndrome, ii) enterocyte metabolism was perturbed by a large number of IEMs, iii) porphyrias may not only perturb heme synthesis but also its degradation, iv) a disturbed intestinal GABA synthesis may be linked to the neurological dysfunction observed in ornithine translocase deficient and gyrate atrophy patients, acting through the gut-brain axis, and v) identification of known and novel co-morbidities.

An objective of this work was to access the role of diet under physiological and pathological conditions. We chose to compare an average American diet with a balanced diet. The balanced diet modeled in this study represents to an ideal diet composition [83], which may not be achieved by most of us on a daily basis. When we applied these dietary patterns, we observed similar overall metabolic performance of the sIECs under both diets, but we also observed flux changes of metabolic tasks for metabolites that have been shown to have implications in metabolic syndrome [358, 359]. For instance, the average American diet led to significant increase of the fructose and glucose secretion flux into the extra-cellular compartment. The high consumption of fructose from added sugars has been linked to hypertension [360] and its increased levels in serum and urine has also been also found in patients with diabetes type 2 [361]. Additionally, it has been suggested that high glucose intake, mainly from sweetened beverages, could lead to weight gain [359]. The absorption pathway of dietary lipid components is distinct from water-soluble nutrients as they are secreted into the lymph from the enterocytes. Once absorbed into the lymphatics, which drain into the left subclavian vein, dietary lipids reach finally the bloodstream [137]. In the model, the computed higher secretion fluxes for cholesterol, cholesterol ester, and triglycerides corresponded to the increased availability of dietary lipids in the average American diet compared to the balanced diet. It has been recommended to reduce the level of lipid in the average American diet [362] due to its close association to diseases, such as obesity and type-2 diabetes [363].

The epithelial cell surface of the small intestine consists of enterocytes and immune systems cells, which are covered by a mucus layer, and represent a protective barrier between the gut microbiota and the host. Numerous diseases exhibit GI signs, even though the small intestine may not be recognized as the primary site of disease. We investigated the potential consequences of IEMs on the sIEC metabolism and correlated them with known clinical features (Table 5.2). Surprisingly, one-third of the IEMs that are currently captured by the global human metabolic reconstruc-

tion [29] mapped onto hs_sIEC611. Even more astonishingly, almost two thirds of the simulated IEMs perturbed significantly the enterocyte metabolism by deranging the function of at least one of the 50 metabolic tasks (Figure 5.3). Almost half of the perturbing IEMs are known to result in intestinal pathology and thus, correspond to true positive predictions (Table 5.2), underlining again the value of computational modeling for elucidating genotype-phenotype relationships. The model's failure to correctly predict some known elevated reported biomarkers, e.g., citrulline and glycogen highlights that enterocyte metabolism is not the only cause for biomarkers but that other tissues, such as liver and muscle, also contribute significantly to plasma biomarkers. Our simulations represent isolated small intestinal enterocytes, but more accurate IEM modeling will require a comprehensive, multi-tissue model. Additionally, our study presents a qualitative but not a quantitative comparison between the changes in flux values for the basolateral secretion reactions and reported biomarker pattern. The presented work is a good starting point to investigate quantitative predictions for the flux ranges, once sufficient kinetic parameters for the reactions are known.

Heme is essential for life and is often referred to as 'pigment of life' [364]. It is synthesized from succinyl-CoA and glycine. Enzyme deficiencies in any of the steps along the biosynthetic pathway cause different forms of porphyrias [364]. Consistently, the metabolic task 'heme synthesis' was blocked in the porphyria models (Figure 5.3). Usually porphyrias show an accumulation of porphyrins, excreted in urine and feces, which is used for biochemical diagnosis [250]. Possible gastro-intestinal (GI) signs include ileal infarction in acute intermittent porphyria and villous atrophy in variegate porphyria [365,366]. These clinical signs may be attributed to the lack of heme in the diet, which can lead to non-optimal bilirubin concentrations within the enterocyte. Since neither of the analyzed diets contained heme supplementation, the porphyrias also resulted in a block in the metabolic task 'heme degradation', which produces bilirubin. This compound has potential anti-oxidant activity, comparable to that of beta-carotene and tocopherol, under low oxygen condition [367] and, at very low concentrations, bilirubin tends to exert protective effects in the small intestine [368]. The porphyria models were not able to secrete carbon monoxide (Figure 5.4), which has various roles in the GI tract, including regulation of gastro-intestinal motility, activation of guanylate cyclase (suggested role in inflammation), function as a neurotransmitter, contribution to membrane potential gradient, and exerting protective role during gastro-intestinal injury [369, 370]. Particularly, this latter role suggests this compound may contribute to the GI inflammation observed in chronic inflammatory bowel disease (IBD). Interestingly, IBD patients are prone to develop acute intermittent porphyria, i.e., AIP [371]. The clinical manifestations of AIP patients also include neurological dysfunction [353]. The predicted reduced urea secretion flux (Figure 5.4) in the porphyria models was due to a reduced urea cycle, indicating ammonia toxicity. For instance, clinical features of hereditary coproporphyria and variegate porphyria include neurological

dysfunction [353]. The computed altered urea secretion flux for porphyrias could thus indicate that ammonia toxicity may also affect the brain of patients with types of other porphyrias. These computational results provide a possible mechanistic explanation of the extensive clinical manifestations observed in porphyria patients.

Gyrate atrophy (OMIM: 258870) and ornithine translocase deficiency (OMIM: - 238970) are caused by mutations in ornithine aminotransferase (*OAT*, GeneID: 4942, E.C. 2.6.1.13) and mitochondrial ornithine transporter (*SLC25A15*, GeneID: 10166), respectively. While gyrate atrophy patients usually have vision loss (OMIM: - 258870), ornithine translocase deficiency is associated with neurocognitive deficits, liver dysfunctions, and lethargy [69]. Interestingly, both disorders exhibit variable degree of neurological manifestations. Gyrate atrophy patients have degenerative changes in the white matter and brain atrophy [372], whereas ornithine translocase deficiency patients usually show developmental delay, ataxia, spasticity, and learning disabilities [69]. Although a disturbed creatine and phosphocreatine metabolism in gyrate atrophy and intermittent hyper-ammonimia in ornithine translocase deficiency have been suggested to be partly responsible for the neurological symptoms, further clarification is needed [69,372]. This is particularly true for the class of patients of ornithine translocase deficiency, who have deficiencies in executive function but never show symptoms of hyper-ammonimia [373]. Interestingly, these two IEMs showed the highest reduction in flux through the synthesis of GABA, an inhibitory neurotransmitter [374]. It has been proposed that the gut and the brain interact via the enteric nervous system through exchange of signaling molecules, which include cytokines and certain metabolites [375]. This brain-gut axis is bidirectional and may have implications in various diseases [376]. Recently, modulation of the GABA receptors by the probiotics was shown to interact via the brain-gut axis [377]. To our knowledge, there have not been any studies analyzing these IEMs from the aspects of brain-gut axis interactions. Based on our computational results, one may speculate that neurological dysfunctions in gyrate atrophy and ornithine translocase deficient patients may be caused by reduced GABA synthesis in the small intestine.

Metabolic networks have been used previously to study co-morbidity, i.e., the co-occurrence of diseases [310]. Our simulations revealed a flux block in the conversion of methionine to cysteine for methylmalonicacidemias (MMA), propionic academia, and cystathionuria (Figure 5.3). This incapability agrees with reports of low cysteine levels observed in these IEMs [354, 378, 379], which differ from one another with respect to the deficient enzyme involved (Table 5.2). Interestingly, co-existence of methylmalonic acidemias and cystathionuria in patients has been reported in literature [353]. Additionally, MMA and propionic academia share common clinical features with respect to movement disorders, gastroenteritis, and neurologic abnormalities (particularly vulnerable basal ganglia) [303]. Our in silico analysis also predicted a failure to synthesize phosphatidylcholine from phosphatidyl-

ethanolamine in the latter two IEMs. This inability may also contribute to their similar clinical features, since phosphatidylcholine is an important membrane constituent and has been shown to exert anti-inflammatory roles, most highly beneficial in ulcerative colitis [380]. Additionally, phosphatidylcholine acts in the mucus as a barrier against many inflammatory conditions [381]. This example highlights again the promise of the computational systems biology approach to identify and investigate co-morbidities.

Enzyme-deficient cells can bypass the missing functionality by employing alternate metabolic pathways [382]. The sIEC model may be used to investigate potential effect of such alternate pathways and may explain the variability of clinical phenotypes within the same IEM. For instance, broad phenotypic variability is particularly evident for type I citrullinemia presenting the classic neonatal form, which causes severe neurological deficit or death without early treatment, the mild adult-onset form, which may not exhibit any symptoms, and a form usually seen in women during pregnancy or post partum [355]. The predicted flux changes through citrulline utilizing, fumarate producing and amino acids related reactions suggest that sIECs could switch to alternative pathways for production of energy precursors and to increased amino acid metabolism in order to maintain the biomass flux and energy state. Any mutation along these pathways could alter the observed phenotype and may contribute to the broad clinical features observed in type I citrullinemia patients. Another example is Smith-Lemli-Opitz syndrome (SLOS), which also has a broad clinical spectrum. While the severe end of the spectrum presents with life threatening conditions (growth retardation, intellectual disability and multiple major and minor malformations), the milder forms have only minor physical abnormalities [356]. The SLOS model utilized at a higher rate the nadh generating reactions of TCA cycle and reactions of electron transport chain to meet the energy demands of the cell. Additionally, reactions of the HMP were utilized to generate nadp, which was channeled into proline synthesis, and 5-phospho-alpha-D-ribose 1-diphosphate, which was used for purine salvage pathway. Interestingly, key gluconeogenic reactions (i.e., 'PCm', 'FBP', and 'G6PPer') also had a higher flux in the SLOS model, indicating that when faced with enzyme deficiencies, sIECs could switch their dependency from amino acids to carbohydrate sources. These examples demonstrate how computational modeling of IEMs could provide biochemical explanations for adaptive mutations in alternative pathways and for phenotypic variability within the same IEM groups.

In conclusion, we want to emphasize that small intestinal enterocytes play a key role in whole body metabolism due to their involvement in enzymatic digestion and absorptions of nutrients as well as inter-organ metabolism. Investigating whole body metabolism in silico will require a comprehensive organ- and tissue resolved metabolic model, which also includes small intestinal enterocyte and hepatocyte metabolism. Such model could be employed to predict metabolic changes upon al-

teration of nutrition, genetic variations, and their impact onto human health. Since metabolic models connect genes with metabolic function, they are promising platforms for personalized systems biomedicine.

5.5 Materials and methods

5.5.1 Flux balance analysis (FBA) and flux variability analysis

The constraint based modeling approach assumes a steady state for the biochemical transformations to occur [11]. Under this condition, the sum of input fluxes equals the sum of output fluxes since the change of metabolite concentration dx over time (dt) is zero. The steady state equation can be written as $S.v = 0$, where S represents the stoichiometric matrix of size m×n (with 'm' being the number of metabolites and 'n' being the number of reactions in the network). v is a flux vector of size $nx1$ containing a flux value v_i for each reactions i in the network. FBA aims to maximize a given objective function under a particular simulation condition as defined by the constraints, which are applied to the model. In this study, metabolic tasks were defined (see below) and employed individually as objective functions under different simulation conditions (e.g., dietary intake or IEM conditions). In flux variability analysis (FVA), the minimum and maximal flux values is computed for each reaction in the model, by choosing each reaction individually as objective function and performing FBA for that reaction [12, 383].

5.5.2 Reconstructing a sIEC-specific metabolic network

A thorough literature search was performed to obtain an in-depth knowledge of all the major metabolic pathways known to occur in an epithelial cell of the small intestine. We then retrieved the corresponding reactions and genes from the human metabolic reconstruction [18], which is accessible through the BiGG database [320], to compile a draft reconstruction. Missing transport and metabolic reactions for peptides and for dietary fibers were added to the initial draft reconstruction upon detailed manual gap analysis and further review of the corresponding literature. Genome annotations from the EntrezGene database [37] as well as protein information from the UniProt [384] and BRENDA database [34] were used in addition to the information retrieved from the scientific literature to assign GPR associations to the reactions not present in Recon 1. For the reactions that were extracted from the Recon 1, GPR associations were kept as reported in Recon 1, since no comprehensive transcriptomic data are available for sIECs. The sIEC metabolic reconstruction

was assembled and converted to a mathematical model using rBioNet as reconstruction environment [36] and an established protocol [10]. We used the global human metabolic network, Recon 1 [18], as reaction database, but adjusted sub-cellular and extra-cellular location, reaction stoichiometry, and directionality according to literature evidence. Only those reactions and pathways with literature evidence for their occurrence in human small intestinal enterocytes were incorporated into hs_sIEC611 from the global human metabolic reconstruction Recon 1, which captures metabolic capabilities known to occur in any human cell. Moreover, we added 262 transport and 50 metabolic reactions, which were not present in Recon 1, but for which supporting information for their presence in sIECs could be found in the scientific literature (Figure 5.1E). These reactions included many transport systems specific for enterocytes and metabolic pathways for sulfo-cysteine metabolism, dietary fiber metabolism, di- and tri-peptide degradation, and cholesterol-ester synthesis (Figure 5.1E). In addition to these reactions, 73 reactions were added from our recently published acylcarnitine/fatty acid oxidation module for the human metabolic reconstruction [29]. We added further 95 reactions, which were present in Recon 1, but for which the compartment was adjusted by placing them into the lumen compartment. The stoichiometry of the reactions catalyzed by the glucose 6-phosphate dehydrogenase (E.C. 1.1.1.49), the 6-phosphogluconolactonase (E.C. 3.1.1.31), and the phosphogluconate dehydrogenase (E.C. 1.1.1.44) was changed to three, since, they were required to generate three molecules each of 6-phospho-D-glucono-1,5-lactone, 6-Phospho-D-gluconate and ribulose-5-phosphate simultaneously [120]. The directionality of ATP requiring fatty acid activation reactions catalyzed by the fatty acyl-CoA ligase (E.C. 6.2.1.3) was changed in agreement with a recent report [120]. Also, the cofactor requirement and sub-cellular localization of reactions included in the cholesterol synthesis pathway that are catalyzed by the desmosterol reductase (E.C. 1.3.1.72) and HMG-CoA reductase (E.C. 2.3.3.10) reactions, were updated in accordance with the current literature evidence [120].

Cellular compartments

The sIEC reconstruction accounts for the following cellular compartments: cytoplasm, mitochondrion, endoplasmic reticulum, peroxisome, and nucleus. Due to its specific anatomy of being a columnar cell, the sIEC possesses an apex and a base (see also Figure 5.2). Both of these structural entities are specific with respect to their substrate uptake and product secretion and exhibit very distinct transporters. Therefore, an additional extra-cellular compartment (‘u’, lumen) was added to the sIEC reconstruction, which represents the gut lumen. Moreover, metabolites are released from the sIEC into either portal venous blood (non-fat nutrients) or lymph (fat nutrients). sIECs take up nutrients from the arterial supply (i.e., superior and inferior pancreatico-duodenal and superior mesenteric arteries) [385]. These three

biofluids are treated as one extra-cellular compartment ('e') in our metabolic reconstruction. Metabolic and transport reactions occurring at the lumen and at the basal side were added based on available literature enabling the assignment of the transporter location, and thus, the luminal uptake and basal release of specific metabolites.

Formulation of biomass reaction

Metabolic reconstructions generally account for a biomass reaction, which captures all known precursors required for physiological and functional maintenance of the cell [10]. We adopted the biomass reaction from Recon 1 [18] while making sIEC-specific modifications. For instance, phosphatidylinositol, phosphatidylglycerol, cardiolipin, phosphatidylserine, and sphingomyelin were removed because the exact metabolic pathways involved in the synthesis and degradation of these molecules within the enterocytes are not known. However, the synthesis of phosphatidylethanolamine from phosphatidylserine in rat small intestine has been suggested [386]. Deoxy-nucleotides have been also removed from the Recon 1 biomass reactions as sIECs cannot replicate [387]. The sIEC biomass reaction thus comprises 20 amino acids, four nucleotides, cholesterol, phosphatidylcholine, phosphatidylethanolamine, folic acid, and water. In order to account for the energy required for the cell maintenance, ATP hydrolysis was also included in the biomass reaction.

GPR associations in the enterocyte reconstruction

The present metabolic reconstruction for human sIECs was assembled in a bottom-up manner based on available literature data. A recent study has reported the proteomic composition of lipid droplets of large intestinal enterocytes, but it focused mainly on proteins participating in non-metabolic reactions [388]. Since no comprehensive transcriptomic and proteomic data set is currently available for human sIECs, GPR associations for all those reactions derived from the generic human metabolic network reactions in hs_sIEC611 were kept as reported in Recon 1. For the reactions that were derived from the IEM module, the GPR associations were retained as described in the module. For the newly added metabolic and transport reactions, the literature was thoroughly reviewed to derive the respective GPR associations. Subsequently, the number of included metabolic genes (611) may be higher than the actual number of metabolic genes expressed in sIECs. However, only 201 reactions are associated with isozymes, while additional 48 reactions are associated with protein complexes as well as isozymes (Figure 5.1C). Thus, 19% (249/1282) of the reactions of hs_sIEC611 may be affected by the current lack of comprehensive

omics data for human sIECs, which corresponds to 396 unique genes. Further 19% (245/1282) of the reactions have no gene associated, which includes 61 diffusion and 16 spontaneous and non-enzymatic reactions. The remaining 168 reactions are orphan reactions.

5.5.3 Debugging of the sIEC model

Dead-end metabolites can only be produced or consumed within a metabolic reconstruction. A key step of the reconstruction process is to evaluate and eliminate these dead-end metabolites depending on available experimental support in the literature [10]. Dead-end metabolites were identified using the COBRA toolbox [13]. Blocked reactions, which cannot carry any flux value under the given simulation constraints, were identified using FVA [12]. Dead-end metabolites and blocked reactions were evaluated manually and supporting evidence from the literature was assembled to connect them to the remainder of the network. Furthermore, we tested for the model's ability to produce each biomass component, as described elsewhere [10].

5.5.4 Secretion profile of the enterocyte model

Based on data derived from literature we constrained 267 exchange reactions for luminal uptake-secretion and basolateral uptake-secretion. Dietary fibers, triglycerides, and other partially digested products cannot enter enterocytes and remain in the lumen, unless they are broken down into simpler units. For example, polysaccharides and disaccharides are hydrolyzed into monomeric forms prior to absorption. The bounds on the corresponding exchange reactions were kept unchanged.

5.5.5 Collection of dietary composition

Diet information was collected from the United States Department of Agriculture, Agricultural research service [347]. All nutrients were converted to a common unit of grams per day. An average American diet, for both males and females (2 years and older), was collected [347]. A balanced diet representing the recommended dietary reference intakes (DRI) of micronutrients as well as macronutrients [83, 348] was also assembled. The DRI reflects the minimum requirements advised to be taken daily to maintain a healthy lifestyle. Values were taken for males between 30-50 years of age. The total fat and cholesterol intake for balanced diet was also represented [349, 389]. We did not represent the energy value in food (i.e., caloric

load) due to disparity between various sources and also due to lack of specific information. The uptake values of the various dietary nutrients in both diets were set as lower bounds on the corresponding luminal exchange reactions. The total carbohydrate content in the diet was converted to starch and dextrin uptake, since, polysaccharides are most abundant carbohydrate component present in our diet [83]. In case of total sugars, disaccharides formed 60% and monosaccharides formed 40%, since, free monosaccharides do not form significant part of the diet [83]. Under the given simulation conditions, we assumed all the available nutrients are completely taken up by the enterocytes.

5.5.6 Conversion of the dietary components into exchange reaction constraints

All dietary components were converted to moles/human/day. To obtain the number of enterocytes per small intestine in a human we used data reported by Chougule et al. [390]. The authors reported that a 10 cm segment of the small intestine contained 7.5×10^6cells, of which 92% were enterocytes. The wet weight of the small intestine was derived from their median regional wet weights (79g for duodenum, 411g for jejunum, and 319g for ileum) [391]. We assumed that the average length of the human small intestine was 20 foot (609.6cm) [385] and that the enterocytes consist of 70% water [75]. The dry weight of the small intestine enterocytes per human was thus $809\times0.3\times0.552 = 133.9704$ $gDW_{enterocytes}$. The conversion factor for all the dietary inputs from $moles/human/day$ to $mmol/gDW_{enterocytes}/hr$ was thus $R = 24\times133.9704\times10^{-3} = 3.2152896$.

5.5.7 Mapping of IEMs onto the enterocyte metabolic network

The IEM information, i.e., affected genes, reactions, organs, and identified biomarkers, was taken from our previous work [29]. For the simulations of each IEM condition on an average American diet, we only considered those IEMs with one gene-one reaction and one gene-multiple reaction associations. We constrained the corresponding reactions to be inactive (i.e., lower and upper reaction bounds were set to be zero) and performed FBA for each of the metabolic tasks.

5.5.8 Capturing adaptive mechanisms under IEM conditions

The affected reaction(s) for an IEM was set as objective function in the healthy model and maximized using FBA. Then, 25% of this maximized value was used

to assign the lower bound to this reaction in the healthy model and FVA was performed to obtain the minimal and maximal flux through each model reaction $i - (FVAmin_i and FVAmax_i)$. The IEM condition was simulated by constraining the lower and the upper bounds of the affected reaction(s) to zero and FVA was performed. The outputs obtained for the healthy model were compared with the output of the corresponding IEM model. The ratio of flux span (FSr) was calculated, using the following formula $FSr = abs(FVAmax_{i,Healthy} - FVAmin_{i,Healthy})/$ $abs(FVAmax_{i,IEM} - FVAmin_{i,IEM})$, where $FVAmax$ is the mximal flux value and $FVAmin$ is the minimal flux value under healthy and IEM conditions. The candidate reactions were identified using a 5% cutoff on the FSr (i.e., reactions with at least 5% difference in flux span between healthy and IEM were chosen).

All fluxes were computed and given in $mmol/gDW_{enterocytes}/hr$. All simulations were carried out with MatLab (MathWork Inc.) as programming environment, using Tomlab (TomOpt Inc.) as linear programming solver, and the COBRA toolbox [13].

Table 5.2: IEMs that showed intestinal metabolic dysfunction during in silico simulation and their relation to small intestinal pathology and biochemical defects as reported in literature

IEMs	Deficient enzyme and specific gene mutation	Symptoms/ organs affected	Relation to small intestine
Porphyrias			
Porphyria ('FCLTm')/ erythropoietic protoporphyria	Ferrochelatase (*FECH*, Gene ID: 2235, E.C. 4.99.1.1).	Acute photosensitivity and liver failure, polyneuropathy [392].	One retrospective case study revealed black blood stained fluid in small intestine [393]. Mouse model showed massive bile duct proliferation and biliary fibrosis [394].
Porphyria ('ALASm')/ sideroblastic anemia	5-aminolevulinate synthase (*ALAS2*, Gene ID: 212, E.C. 2.3.1.37).	Fatigue, dizziness, a rapid heartbeat, pale skin, and an enlarged liver and spleen (OMIM: 300751).	No direct link between small intestine malfunction and sideroblastic anemia could be found. However, erythropoietic protoporphyria patients have been reported to have small intestinal signs and have also been reported to develop eventually sideroblastic anemia [395].
Acute intermittent porphyria	Hydroxymethylbilane synthase (*HMBS*, Gene ID: 3145, E.C. 2.5.1.61)	Gastro-intestinal system, nervous system (OMIM: 176000).	In one case, trans-mural infarction of distal ileum and the existence of sickle cell anemia caused death [365].
Variegate porphyria	Protoporphyrinogen oxidase (*PPOX*, Gene ID: 5498, E.C. 1.3.3.4)	Skin, peripheral nervous system, liver (OMIM: 176200).	Two cases of co-existence of variegate porphyria and celiac disease have been reported. Coincidental existence of porphyrias and celiac disease also exists, villous atrophy has been seen [366].

IEMs	Deficient enzyme and specific gene mutation	Symptoms/ organs affected	Relation to small intestine
Delta-amino-levulinate dehydratase deficiency	Delta-aminolevulinate dehydratase (*ALAD*, Gene ID: 210, E.C. 4.2.1.24)	Described as one form of acute hepatic porphyria (OMIM: 125270)	Six cases have been reported. Severe abdominal pain may cause nervous symptoms [364].
Hereditary coproporphyria	Coproporphyrinogen-III oxidase (*CPOX*, Gene ID: 1371, E.C. 1.3.3.3)	Nervous system, skin (OMIM: 121300)	Delay in transit time in small intestine has been observed [396].
Nucleotide disorders			
Adenosine deaminase (ADA) deficiency	Adenosine deaminase (*ADA*, Gene ID: 100, E.C. 3.5.4.4)	Immune system (OMIM: 102700)	ADA deficient mice died at the third day of birth and have been reported to show pathological signs of small intestinal cell death [397].
Orotic aciduria	Uridine 5'-monophosphate synthase (*UMPS*, Gene ID: 7372, E.C. 2.4.2.10 and E.C. 4.1.1.23)	Heart, muscle, nervous system (OMIM: 258900)	Orotic aciduria has been reported in short bowel syndrome [398].
Carbohydrate disorders			
Glycogen storage disease XI	L-lactate dehydrogenase (*LDHA*, Gene ID: 3939, E.C. 1.1.1.27)	Kidney, muscle, skin (OMIM: 612933)	Patients exhibit intestinal mal-absorption and diarrhoea [399].

IEMs	Deficient enzyme and specific gene mutation	Symptoms/ organs affected	Relation to small intestine
Hereditary fructose intolerance	Fructose-bisphosphate aldolase B (*ALDOB*, Gene ID: 229, E.C. 4.1.2.13)	Liver (OMIM: 229600)	Gastro-intestinal symptoms include nausea, vomiting, abdominal pain, and meteorism (i.e., excessive gas accumulation) [400].
Glucose-galactose malabsorption	Sodium/glucose cotransporter 1 (*SLC5A1*, Gene ID: 6523)	Gastro-intestinal system (OMIM: 606824)	The transport protein, located at the apical surface of enterocytes, actively absorbs glucose and galactose into the cells. Defective transporter leads to primary gastro-intestinal disturbances, i.e., chronic diarrhoea as seen in this disorder [401].
Pyruvate kinase deficiency	Pyruvate kinase 1 (*PKLR*, Gene ID: 5313, E.C. 2.7.1.40)	Liver, spleen, gall bladder (OMIM: 609712 and OMIM: 266200)	No direct link reported but an increased abundance of the pyruvate kinase M2 form has been used as a inflammatory marker in patients with inflammatory bowel disease [402].
Congenital lactase defi-ciency/congenital alactasia	Lactase-phlorizin hydrolase (*LCT*, GeneID: 3938, E.C. 3.2.1.108 & 3.2.1.62)	Digestive system (OMIM:223000)	Congenital lactase deficiency is a severe gastrointestinal disorder characterized by watery diarrhoea in infants fed with breast milk or other lactose-containing formulas. Other clinical signs include vomiting, failure to thrive, dehydration, disaccharidurıa including lactosuria, renal tubular acidosis, and amino aciduria. In a clinical study of 16 patients, identifed that mean height of the epithelial cells was reduced in all cases [329, 403].

IEMs	Deficient enzyme and specific gene mutation	Symptoms/ organs affected	Relation to small intestine
Glycogen storage disease I (GSD I)	Glucose-6-phosphatase and Glucose-6-phosphate translocase (*G6PC* or *SLC37A4*, GeneID: 2538, 2542, E.C. 3.1.3.9). GSD type 1 a is casued due to mutations in G6PC and type 1 b due to defect in SLC37A.	liver, kidney, digestive system, respiratory system, skeletal system (OMIM:232200)	In both of these disorders there is frequently occurring chronic diarrhoea. Additionally, In GSD Ib, loss of mucosal barrier function due to inflammation seems to be the main cause of diarrhoea [404]. There is disordered intestinal function in both [405], while untreated type 1b can lead to intestinal mucosal ulcers [406].
Fanconi-Bickel syndrome (FBS)	GLUT2 (*SLC2A2*, GeneID: 6514)	nervous system, skeletal system, liver, kidney, pancreas (OMIM:227810)	FBS patients exhibit intestinal malabsoprtion and diarrhoea [407]. Microvillus inclusion disease (MVID) is a rare congenital enteropathy associated with brush border atrophy and reduced expression of enzymes at the enterocyte‘s apical surface. MVID is associated with mutations in the MYO5B gene, which is expressed in all epithelial tissues. Patients with MVID were seen to develop FBS while receiving total parenteral nutrition [408].
Congenital sucrase-isomaltase deficiency (CSI)	Sucrase-isomaltase (*SI*, Gene ID:6476, E.C. 3.2.1.48 and 3.2.1.10)	Digestive system	CSI is caused by a deficiency in sucrase-isomaltase, an integral brush border membrane protein of the small intestine (OMIM:222900). It is charaterized by fermentative diarrhoea, abdominal pain, and cramps [409].

IEMs	Deficient enzyme and specific gene mutation	Symptoms/ organs affected	Relation to small intestine
Aminoacidopathies			
Methylmalonic acidemia	Methylmalonyl-CoA mutase (*MUT*, Gene ID: 4594, E.C. 5.4.99.2)	Brain, kidney, pancreas	Co-occurance of methylmalonic and propionic acidemia has been reported in patients, who also had gastroenteritis [303].
Methylmalonic aciduria III	Methylmalonyl-CoA epimerase (*MCEE*, Gene ID: 84693, E.C. 5.1.99.1)	Nervous system (OMIM: 251120)	Co-occurance of methylmalonic and propionic acidemia has been reported in patients, who also had gastroenteritis [303].
Propionic acidemia	Propionyl-CoA carboxylase (*PCCA* or *PCCB*, Gene ID: 5095, Gene ID: 5096, E.C. 6.4.1.3)	Nervous system, muscle, heart (OMIM: 606054)	Co-occurrence of methylmalonic and propionic acidemia has been reported in patients, who also had gastroenteritis [303].
Arginase deficiency	Arginase (*ARG1*, Gene ID: 383, E.C. 3.5.3.1)	Nervous system, RBC (OMIM: 207800)	No small intestine disorder reporter but a mouse model of arginase deficiency exhibited urea cycle perturbations in liver [410].
Cystinuria	b0,+AT transporter (*SLC3A1* or *SLC7A9*, Gene ID: 6519, Gene ID: 11136)	Kidney, urinary bladder (OMIM.220100)	Cystinuria is caused by a defect in the intestinal dibasic amino acid transporter. However, no associated gastro-intestinal pathology has been reported [411].

IEMs	Deficient enzyme and specific gene mutation	Symptoms/ organs affected	Relation to small intestine
Histidinemia	Histidine ammonia-lyase (*HAL*, Gene ID: 3034, E.C. 4.3.1.3)	Nervous system (OMIM: 235800)	One case reported with gastro-intestinal disorder [412].
Ornithine transcarbamylase deficiency	Ornithine carbamoyltransferase, mitochondrial (*OTC*, Gene ID: 5009, E.C. 2.1.3.3)	Nervous system, liver, skin, hair (OMIM: 311250)	Intestinal ornithine transcarbamylase /lactase ratio indicates degree of mucosal damage and atrophy [413].
Ornithine translocase deficiency/HHH (hyperornithinemia, hyperammonemia, homocitrullinuria) syndrome	Mitochondrial ornithine transporter 1 (*SLC25A15*, Gene ID: 10166)	Nervous system, liver (OMIM: 238970)	Gastro-intestinal illness is a secondary complication in urea cycle disorders [373].
Snyder-Robinson syndrome	Spermine synthase (*SMS*, Gene ID: 6611, E.C. 2.5.1.22)	Nervous system, skeletal system (OMIM: 309583)	Targeted deletion of the gene in mouse showed no pathology. However, deletion of the downstream enzyme to spermine synthase in polyamine metabolism, i.e., diamine acetyltransferase 1 (*SAT1*, 6303, E.C.2.3.1.57), in mouse showed gastro-intestinal carcinogenesis [414].
Citrullinemia	Argininosuccinate synthase or Citrin (*ASS1* or *SLC25A13*, Gene ID: 445, Gene ID: 10165, E.C. 6.3.4.5)	Nervous system, liver (OMIM: 215700, OMIM: 603471)	Contrastingly low levels of plasma citrulline is an indicator for villous atrophy and short bowel syndrome [350].

IEMs	Deficient enzyme and specific gene mutation	Symptoms/ organs affected	Relation to small intestine
Lysinuric protein intolerance (LPI)	Y+L amino acid transporter 1 (*SLC7A7*, GeneID: 9056)	liver, spleen, skeletal system, muscle, kidney, respiratory system, nervous system, immune system, pancreas (OMIM:222700)	LPI is caused by defective cationic amino acid transport at the basolateral membrane of epithelial cells in kidney and intestine. LPI should be considered for the differential diagnosis of conditions associated with intestinal villous atrophy (blunting of villi with elongation of crypts) [415].
Cholesterol synthesis disorders			
Conradi-Hunermann syndrome	3-beta-hydroxysteroid-Delta8,Delta7-isomerase (*EBP*, Gene ID: 10682, E.C. 5.3.3.5)	Skeletal system (OMIM: 302960)	Case report of a newborn male revealed significant delay in motility in small intestine and developmental delay [416].
Mevalonic aciduria	Mevalonate kinase (*MVK*, Gene ID: 4598, E.C. 2.7.1.36)	Nervous system, liver, spleen, muscles (OMIM: 251170)	One case reported complete jejunal and ileal obstruction and failure to thrive in mevalonic aciduria [417].
Smith-Lemli-Opitz syndrome (SLOS)	7-dehydrocholesterol reductase (*DHCR7*, Gene ID: 1717, E.C. 1.3.1.21)	Nervous system, muscles (OMIM: 270400)	Infants with SLOS usually present gastrointestinal problems that include dysmotility, hypomotility, gastrointestinal reflux, constipation, formula intolerance and developmental delay [356].
Other lipid disorders			
Medium chain acyl CoA dehydrogenase deficiency	Medium chain acyl CoA dehydrogenase (*ACADM*, Gene ID: 34, E.C. 1.3.99.3)	Liver, nervous system (OMIM: 201450)	Clinical manifestation includes acute intestinal intussusceptions, i.e., enfolding of segment of intestine causing intestinal obstruction [418].

IEMs	Deficient enzyme and specific gene mutation	Symptoms/ organs affected	Relation to small intestine
Trifunctional protein deficiency	Trifunctional enzyme subunit alpha, mitochondrial (*HADHA*, Gene ID: 3030, E.C. 4.2.1.17 and E.C. 1.1.1.211)	Muscle, liver, eye, nervous system (OMIM: 609016)	Intestinal pseudo-obstruction reported in mitochondrial tri-functional protein deficiency [419].
Long chain acyl-CoA dehydrogenase deficiency (LCHAD)	Trifunctional enzyme subunit beta, mitochondrial (*HADHB*, Gene ID: 3032, E.C. 2.3.1.16)	Heart, muscle, respiratory system (OMIM: 143450)	Echogenic bowel syndrome reported in LCHAD patient [420].
Ichthyosis prematurity syndrome (IPS)	Solute carrier family 27 (fatty acid transporter) (*SLC27A4*, GeneID: 10999)	Skin, respiratory system	IPS is caused by a mutation in the gene encoding FATP4, a major fatty acid transporter in the enterocytes (OMIM: 608649, OMIM: 604194).
Miscellaneous disorders			
Dihydrofolate reductase (DHFR) deficiency	Dihydrofolate reductase (*DHFR*, Gene ID: 1719, E.C. 1.5.1.3)	Nervous system (OMIM: 126060)	DHFR deficiency results in megaloblastic anemia (OMIM: 613839). Megaloblastic anemia leads to gastro-intestinal disturbances like malabsorption and diarrhoea [421].
Glycerol kinase deficiency	Glycerol kinase (*GK*, GeneID: 2710, E.C. 2.7.1.30)	Digestive system, nervous system	Gastrointestinal symptoms (including gastro-enteritis and decreased oral intake) are the initial features of this disease (OMIM: 307030). [422].

6 Assessing the metabolic cost of commonly used drugs

6.1 Abstract

Metabolism contributes significantly towards the pharmacokinetic and pharmacodynamic properties of a drug molecule. The recently published, human metabolic reconstruction, Recon 2 represents the most current knowledgebase of human biochemistry. On the other hand, diet and genetics have profound effect on health and disease. The presented work combines manually assembled reconstruction of the metabolism of 5 most highly prescribed drug groups, i.e., the stains, antihypertensives, immunosuppressants, analgesics, and miscellaneous groups. Having combined this with Recon 2, in order to obtain a condition and functional drug module, Recon2_DM1796, which includes 2803 metabolites, 8161 reactions distributed over eight sub-cellular compartments, and 1796 genes. The efficiency of the module was tested under three simulated diets, i.e., average American, balanced and vegetarian diets, against 50 drug objectives that captured the overall drug metabolism. The role of diseases was studied in the context of inborn errors of metabolism. Our important results include: (i) diet and disease significantly affect the metabolism of statins and acetaminophen, (ii) drug interactions between stains and cyclosporine can be caused due to common metabolic and transport pathways, in addition to the previously established CYP3A4 connection, (iii) cellular energy levels are compromised in order to operate efficient drug metabolism and its subsequent elimination. Hence, this work hold the potential for aiding further research on adverse drug reactions as well as designing patient specific therapy.

6.2 Introduction

Drugs are known to undergo two distinct phases of metabolic fate for complete metabolism and elimination, i.e., phase I: functionalization reactions, where the functional group of the parent drug molecule is exposed for subsequent transformations, often resulting in its activation/inactivation (e.g., oxidation, reduction, and hydrolysis reactions), and phase II: conjugation reactions, where the metabolized

derivative combines with glucuronic acid (via UDP-glucuronyltransferases, UGTs), sulphate (via sulfotransferases), or acetyl groups (via acetyltransferases), resulting in the formation of highly polar compounds, which can be readily excreted in feces or urine [423]. Drug molecules require specific membrane transport proteins. For example, ATP binding cassettes are usually involved in drug export and organic anion transporting polypeptides in import [68, 424]. Studying a drug should involve both its pharmacokinetics (absorption from the gastrointestinal tract in case of oral drugs and distribution to various organs, metabolism for activation and degradation, and elimination, ADME) and pharmacodynamics (interaction of the drug with the target molecule) that initiates biochemical and physiological responses in the body [423]. Various factors can affect drug action, such as age, gender, genetics, diet, disease, and hormonal influence, as these can either potentiate or inhibit the required drug effect. For instance, interaction of drugs and certain dietary components (e.g., grape fruit juice, watercress, vitamin C, and cruciferous vegetables) as well as fully fed and starvation states is well documented [425, 426]. Additionally, disease states have been known to interfere with the drug action, particularly liver diseases [423]. Reports of drugs have shown further precipitating inborn errors of metabolism (IEMs) in patients, either on single drug or under multiple drug administration [427–429]. IEMs are the metabolic disorders, characterized by a specific mutation in the genome, leading to a metabolic block in the downstream metabolic pathways.

A genome scale metabolic reconstruction (GENRE) captures genomic, physiological and biochemical knowledge of the target organism [5]. More than 100 GENREs have been published, including for humans [31]. The most recent human metabolic reconstruction, Recon 2, is a significant expansion both in content and knowledge, over its previous version [18], comprising 7440 reactions, 2626 unique metabolites, distributed over eight sub-compartments/organelles, i.e., cytoplasm, mitochondrion, nucleus, endoplasmic reticulum, peroxisome, lysosome, golgi apparatus, and extra-cellular space, and 1789 unique genes. These reactions represent biochemical transformation occurring in any human cell, making it highly compliant for studying drug metabolism, which occurs in both hepatic and extrahepatic tissues (e.g., small intestine, kidney, lung, and heart) [430]. The human metabolic reconstruction, and its predecessor, has been used to investigate the role of IEMs [19, 31, 431] and the off-target effect of drugs [22] as well as to predict novel drug targets [20, 27, 56, 65, 343, 432–434]. Furthermore, the liver specific metabolic network, HepatoNet1 [21], has been combined with whole-body physiology based pharmacokinetic model to study the toxic effects of drugs [66]. So far, no study investigated the metabolic load of drug metabolism itself using a genome scale human metabolic reconstruction.

Here, we investigate *in silico* whether the genetic background of an individual and/or the dietary intake alter the efficiency, with which a drug is converted into

its active form or cleared from the body. Therefore, we assembled a comprehensive drug metabolic reconstruction based on literature permitting the analysis of the dietary composition and of inherited metabolic disorders on drug metabolism and drug-drug interactions.

6.3 Results

The current study presents a manually assembled metabolic reconstruction of five most highly prescribed drug groups and captures in detail our current understanding of their absorption, intra-cellular distribution, metabolism, and elimination. We then combined this drug reconstruction with the recently published community driven human metabolic reconstruction, Recon 2, and obtained a functional drug module to study: (i) the role of diet in metabolizing drugs, (ii) the effect of previously occurring genetic deficiencies on drug metabolic patterns as well as (iii) on drug-drug metabolic interactions, and (iv) the metabolic cost of drug metabolism.

6.3.1 Reconstruction of drug metabolism

We present the first extensive metabolic reconstruction of five most commonly used drug groups, built through extensive manual curation of available scientific literature (Figure 6.1). These drug groups comprise of 18 drug compounds: (i) statins group: eight anti-hypercholesterolemic compounds, (ii) anti-hypertensives group: losartan, torasemide, and nifedipine; (iii) immunosuppressant group: cyclosporine and tacrolimus; (iv) analgesic group: ibuprofen and acetaminophen; and (v) miscellaneous group: anti-diabetic/gliclazide, sedative/midazolam, and anti-hyperuricemic/allopurinol. Information from more than 250 peer reviewed research articles and books has been used to capture in great detail the known metabolic fate of these drugs in the small intestine and the liver (Table 6.1). In the following sections, we provide a short summary of the various drugs metabolic and transport pathways included in the reconstruction. Refer to Figure 6.1 for a depiction of possible elimination routes.

Table 6.1: Statistics of drug reconstruction and drug module after being combined with Recon 2

Components	Drug Reconstruction	Drug Reconstruction + Recon 2: Recon2_DM1796
Total number of metabolic reactions	187	4,489
Total number of transport reactions	386	2,782
Number of exchange/demand reactions	148	890
Total number of unique metabolites	210	2,803
Total number of unique genes	57	1,796
Number of compartments	5	8
References	259*	>1,500*
*peer reviewed journal articles, primary literature, and books		

Statin group

Statins are a group of drugs that target key enzyme of de novo cholesterol biosynthetic pathway, i.e., HMG-CoA reductase (E.C.1.1.1.88), and are principally comprised of eight drug compounds (Table 6.2). They are mainly used to treat hypercholesterolemia and coronary atherosclerosis, with liver being the target organ [435]. Most of the currently used statins show to some extend adverse drug action. For instance, rhabdomyolosis is a common side effect of cerivastatin, which led to its withdrawal in 2001 [436]. While all statins are administered in their active form, lovastatin and simvastatin are given as pro-drugs [435]. All statins exhibit two chemically different forms, an acid and a lactone form, with the acid forms being biologically more potent [437, 438]. The chief metabolic reactions/pathways, in which these forms participate, are: (i) oxidation, (ii) beta oxidation (pravastatin, fluvastatin), (iii) glucuronidation, (iv) interconversion between acid and lactone forms, (v) isomerization, (vi) aromatization (pravastatin, fluvastatin), and (vi) demethylation (cerivastatin, and rosuvastatin). These reactions are majorly catalyzed by enzymes encoded by genes of CYP, UGT families, as well as fatty acid oxidation genes (Table 6.2). The drug module captures all the known metabolic pathways of each of these statin drugs in the form of 94 metabolic reactions, leading to formation of eight metabolites by lovastatin, 12 by simvastatin, ten by atorvastatin, 13 by pravastatin, 15 by fluvastatin, seven by cerivastatin, three by rosuvastatin, and five

by pitavastatin. Metabolism of these statins is important for their conversion into active forms to act on the target enzyme as well as for their effective elimination. The major part of statin conversion into their derivative metabolites occurs in the liver. Small intestinal enterocytes also contribute in part to the oxidation, interconversion between acid and lactone forms, and isomerization of lovastatin, atorvastatin, and pravastatin. There exist distinct membrane transporters for all discussed statins, including for lovastatin and simvastatin, which are lipophilic statins that can cross cell membrane with ease.

The transporters are from the organic anion transporter polypeptide family (OATPs), organic anion transporters (OATs), and ATP binding cassette family (ABCs), where different representatives are known to transport different statins (Table 6.2). The various drug derived compounds can follow either the biliary (e.g., simvastatin, pitavastatin metabolites), the renal (pravastatin metabolites), or both (lovastatin, atorvastatin, fluvastatin, cerivastatin, rosuvastatin metabolites) routes. Generally, the transporters of the renal route are OATPs and of the biliary route are from the ABCs family [68]. Alternatively, a few statins and their metabolites can also undergo entero-hepatic re-circulation (e.g., pravastatin [439]). The drug reconstruction captures 103 extra-cellular transport reactions for the transport of statins and their metabolites between cytosol and extracellular space.

Anti-hypertensive group

The anti-hypertensives were majorly represented by three drugs (Table 6.2) and are mainly used in the treatment of hypertension, congestive heart failure, and angina. Losartan inhibits the binding of angiotensin II, a peptide hormone, to angiotensin II type 1 (AT1) receptor. Torasemide inhibits the $Na^+/K^+/2Cl^-$transport system in the kidney and nifedipine blocks L-type calcium channels. Chief metabolic pathways contained in the drug reconstruction are: (i) oxidation, (ii) glucuronidation (losartan), (iii) tolyl methylhydroxylation (torasemide), (iv) locatonization (nifedipine), and (v) demethylation (nifedipine), represented by 14 metabolic reactions, catalyzed by enzymes encoded by CYP2C and CYP3A gene families (Table 6.2). These reactions lead to the formation of seven metabolites for losartan, three metabolites for torasemide, and four metabolites for nifedipine. The oxidation of the parent forms of the drugs, such as losartan and torasemide, produces active metabolites (losartan-E3174/ losartan-M6 [440], torasemide-M3 [441]), which are equally active against their respective targets. The three anti-hypertensives and their derived metabolites are majorly excreted via renal route. However, losartan in the parent form has also been recovered in bile, transported by ABCB1 transporter (GeneID: 5243) [424]. Although majority of the efflux transporters remain unknown, the hepatic uptake transporter has been identified as being the organic anion transporter,

OAT1 (*SLC22A6*, GeneID: 9356) [442]. A total of 24 extra-cellular transport reactions are captured for this group in the drug reconstruction.

Immunosuppressant group

Cyclosporine A and tacrolimus are the considered representatives of the immunosuppressant group in the drug reconstruction. Both the drugs are administered for inhibiting inter-leukin 2 transcription and T-lymphocyte activation. Cyclosporine A acts via binding to cyclophilin, a peptidyl prolyl isomerase [443, 444], while tacrolimus binds to the tacrolimus-binding protein. These drug-receptor complexes inhibit calcineurin, a protein phosphatase, thus, restraining a number of transcriptional events [445]. Immunosuppressants are used in rheumatoid arthritis, severe psoriasis, and against graft rejection. Metabolic pathways for these drugs include: (i) oxidation (cyclosporine, tacrolimus), (ii) demethylation (cyclosporine, tacrolimus), and (iii) sulphation and glucuronidation (cyclosporine), catalyzed by enzymes encoded by *CYP3A4* (GeneID: 1576) and *CYP3A5* (GeneID: 1577) for both the drugs [446, 447], leading to the formation of 14 metabolites by cyclosporine and eight metabolites by tacrolimus. These pathways are described with 27 metabolic reactions in the drug reconstruction. Precise transporters for cyclosporine and tacrolimus have not been identified, however, both these drugs are relatively lipophilic, having high permeability, and their metabolites follow both renal and biliary routes for excretion [448, 449]. Consequently, 44 extracellular transport reactions were defined. See Figure 6.1 for the elimination routes.

Analgesic group

Analgesics are used for pain relief. Ibuprofen and acetaminophen are the most commonly used drugs of this group. While ibuprofen inhibits both cyclooxygenase 1 and 2 (E.C.1.14.99.1) [450], acetaminophen selectively inhibits cyclooxygenase 2 [451] and is therefore widely used for fever and pain. Ibuprofen mainly undergoes (i) chiral inversion, (ii) glucuronidation, and (iii) oxidation [452, 453]. These transformations are catalyzed by racemase (E.C.5.1.99.4) [454], UDP-glucuronsyltransferases (E.C.2.4.1.17) [455], and mono-oxygenase (E.C.1.14.14.1) [456]. These enzymes are encoded by multiple genes of the ACOT, UGT, and CYP families, respectively (Table 6.2). The metabolism of acetaminophen can be broadly classified as (i) non-oxidative and (ii) oxidative. The majority of the drug (~ 80%) undergoes non-oxidative metabolism [457], which consist of sulphation, catalyzed by sulpho-transferase 1 (*SULT1*, GeneID: 6817, E.C.2.8.2.1) [458], and glucuronidation, catalyzed by UDP-glucuronsyltransferases, encoded by various UGTs (Ta-

ble 6.2) [459]. On the other hand, the oxidative metabolism is catalyzed via cytochrome P450 family proteins (i.e., CYP1, CYP2 and CYP3 families) (Table 6.2). Accordingly, the drug reconstruction captures 32 metabolic reactions for the analgesic group, comprising of 16 different metabolites each from ibuprofen and acetaminophen. With respect to the elimination route, ibuprofen and its metabolites follow majorly the renal route, while <1% of the parent drug follows the biliary route [453]. Although ibuprofen can diffuse into intestinal cells, organic anion transporter, OAT1 (*SLC22A6*, GeneID: 9356) mediates its transport in the kidney [460] and in the liver [461]. In case of acetaminophen, both passive transport process [457] and active transport via various ABC transporters (Table 6.2) have been reported [424, 457]. The majority of acetaminophen metabolites follow renal route, while very few follow the biliary route (e.g., sulphate conjugates). Additionally, acetaminophen-glutathione-conjugates undergo entero-hepatic re-circulation [457]. In total, 38 extracellular transport reactions are shown for this group in the drug reconstruction. See Figure 6.1 for the elimination routes.

Miscellaneous group

This group consisted of three drug compounds that are used as (i) oral hypoglycemic agent/gliclazide targets K^+/ATP channel and helps in controlling basal glucose level via acting on glycogen synthase (E.C.2.4.1.11) and fructose-2,6 bis-phosphatase (E.C. 2.7.1.105) [462], (ii) midazolam is used as a sedative, as it produces antegrade amnesia by causing gamma-amino-butyrate accumulation [463], and (iii) allopurinol is used in the treatment of hyper-uricemic conditions as it targets xanthine oxidase (E.C.1.17.3.2), an urate producing enzyme [464]. These drugs majorly undergo (i) oxidation and (ii) conjugation with either UDP-ribose (allopurinol) or UDP-glucuronate (gliclazide, midazolam), which is catalyzed by (i) oxidoreductases for allopurinol [464] and (ii) monooxygenase (E.C. 1.14.14.1) encoded by genes of *CYP2C/2D* family for gliclazide [465, 466] and by *CYP3A* family for midazolam [467] (see Table 6.2). Various UGTs catalyze the conjugation reactions (Table 6.2). The metabolism of this group in the drug reconstruction has been covered in the form of 20 reactions, involving ten metabolites from gliclaizde, six from midazolam, and three from allopurinol. The elimination routes for gliclazide and its metabolites are both renal (60-70%) and biliary (10-20%) [468], but exclusively renal route for midazolam, allopurinol, and their metabolites [464, 469]. While gliclazide and midazolam are highly lipophilic [462, 470], allopurinol transporters have been suggested as concentrative nucleoside transporters (SLC family) [471]. All the transport mechanisms have been represented with 37 reactions in the drug reconstruction.

Table 6.2: Drugs with their corresponding metabolic and transport genes

Drug group	Drug common name	ATC code	Genes encoding metabolic and transport proteins	References
Statins	Lovastatin	C10AA02	*CYP3A4* (GeneID: 1576), *CYP3A5* (GeneID: 1577), *CYP2C8* (GeneID: 1558), *ABCB1* (GeneID:5243)	[354, 424, 472, 473]
	Simvastatin	C10AA01	*CYP3A4* (GeneID: 1576), *CYP3A5* (GeneID: 1577), *UGT1A1* (GeneID:54658), *UGT1A3* (GeneID:54659), *ABCB1* (GeneID:5243)	[474, 475]
	Atorvastatin	C10AA05	*CYP3A4* (GeneID: 1576), *CYP3A5* (GeneID: 1577), *CYP2C8* (GeneID: 1558), *UGT1A3* (GeneID: 54659), *UGT1A4* (GeneID: 54657), *UGT1A1* (GeneID: 54658), *UGT1A8* (GeneID: 54576), *UGT2B7* (GeneID: 7364), *UGT1A9* (GeneID: 54600), *SLCO1B1* (GeneID: 10599), *SLCO2B1* (GeneID: 11309), *ABCB1* (GeneID: 5243)	[473, 475–477]
	Pravastatin	C10AA03	*CYP3A4* (GeneID: 1576), *CYP3A5* (GeneID: 1577), *SLCO1B1* (GeneID: 10599), *SLCO1B1* (GeneID: 10599), *SLCO2B1* (GeneID: 11309), *ABCB11* (GeneID: 8647), *ABCB1* (GeneID: 5243), *ABCC2* (GeneID: 1244), *ABCG2* (GeneID: 9429)	[424, 472, 478]
	Fluvastatin	C10AA04	*CYP3A4* (GeneID: 1576), *CYP3A5* (GeneID: 1577), *CYP2C8* (GeneID: 1558), *CYP2C9* (GeneID: 1559), *CYP2D6* (GeneID: 1565), *CYP2C19* (GeneID: 1557), *SLCO1B3* (GeneID: 28234), *SLCO2B1* (GeneID: 11309), *ABCB11* (GeneID: 8647)	[473, 479, 480]

Drug group	Drug common name	ATC code	Genes encoding metabolic and transport proteins	References
	Cerivastatin	C10AA06	*CYP3A4* (GeneID: 1576), *CYP3A5* (GeneID: 1577), *CYP2C8* (GeneID: 1558), *UGT1A1* (GeneID: 54658), *UGT1A3* (GeneID: 54659), *SLCO1B1* (GeneID: 10599), *ABCB11* (GeneID: 8647), *ABCB1* (GeneID: 5243), *ABCC2* (GeneID: 1244), *ABCG2* (GeneID: 9429)	[424, 473, 481]
	Pitavastatin	C10AA08	*CYP3A4* (GeneID: 1576), *CYP3A5* (GeneID: 1577), *CYP2C8* (GeneID: 1558), *CYP2D6* (GeneID: 1565), *UGT1A1* (GeneID: 54658), *UGT1A3* (GeneID: 54659), *UGT2B7* (GeneID: 7364), *UGT1A4* (GeneID: 54657), *UGT1A6* (GeneID: 54578), *SLCO1B3* (GeneID: 28234), *SLCO2B1* (GeneID: 11309), *ABCB11* (GeneID: 8647), *ABCG2* (GeneID: 9429)	[424, 482, 483]
	Rosuvastatin	C10AA07	*CYP2C9* (GeneID: 1559), *CYP2C19* (GeneID: 1557), *SLCO1B1* (GeneID: 10599), *SLCO1B3* (GeneID: 28234), *SLCO2B1* (GeneID: 11309), *ABCG2* (GeneID: 9429)	[424, 484, 485]
Anti-hypertensive	Losartan	C09CA01	*CYP3A4* (GeneID: 1576), *CYP2C9* (GeneID: 1559), *UGT1A3* (GeneID:54659), *UGT2B7* (GeneID:7364), *UGT1A10* (GeneID: 54575), *UGT1A1* (GeneID: 54658), *UGT1A4* (GeneID: 54657), *UGT1A7* (GeneID: 54577), *UGT1A8* (GeneID: 54576), *UGT1A9* (GeneID: 54600), *UGT2B17* (GeneID: 7367), *ABCB1* (GeneID: 5243)	[424, 486, 487]

Drug group	Drug common name	ATC code	Genes encoding metabolic and transport proteins	References
	Torasemide	C03CA04, C03CA01	*CYP2C9* (GeneID: 1559), *SLCO1B1* (GeneID: 10599), *SLC22A6* (GeneID: 9356)	[424, 441, 488]
	Nifedipine	C08CA05	*CYP3A4* (GeneID: 1576), *CYP3A5* (GeneID: 1577)	[489, 490]
Immuno-suppressants	Cyclosporine A	L04AD01, S01XA18	*CYP3A4* (GeneID: 1576) and *CYP3A5* (GeneID: 1577), *ABCB1* (GeneID: 5243)	[424, 446]
	Tacrolimus	D11AX14, L04AA05	*CYP3A4* (GeneID: 1576) and *CYP3A5* (GeneID: 1577), *ABCB1* (GeneID: 5243)	[424, 447]
Analgesics	Ibuprofen	C01EB16	*ACOT2* (GeneID: 10965), *ACOT4* (GeneID: 122970), *ACOT11* (GeneID: 26027), *ACOT8* (GeneID: 10005), *AMACR* (GeneID: 23600), *UGT1A3* (GeneID: 54659), *UGT1A9* (GeneID: 54600), *UGT2B7* (GeneID: 7364), *CYP2C9* (GeneID: 1559), *CYP2C8* (GeneID: 1558), *CYP3A4* (GeneID: 1576), *CYP2C19* (GeneID: 1557), *SLC22A6*, (GeneID: 9356)	[450, 452, 454, 460, 491]
	Acetamino-phen	N02BE01	*SULT1* (GeneID: 6817), *UGT1A1* (GeneID: 54658), *UGT1A6* (GeneID: 54578), *UGT1A9* (GeneID: 54600), *UGT1A10* (GeneID: 54575), *CYP2E1* (GeneID: 1571), *CYP1A2* (GeneID: 1544), *CYP2A6* (GeneID: 1548), *CYP2D6* (GeneID: 1565), *CYP3A4* (GeneID: 1576), *ABCC3* (GeneID: 8714), *MRP2* (GeneID: 1244), *BCRP* (GeneID: 9429), *ABCB1* (GeneID: 5243)	[424, 457–459, 492]

Drug group	Drug common name	ATC code	Genes encoding metabolic and transport proteins	References
Miscellaneous	Gliclazide	A10BB09	*CYP2C9* (GeneID: 1559), *CYP2C18* (GeneID: 1562), *CYP2C19* (GeneID: 1557), *CYP2C8* (GeneID: 1558), *CYP2D6* (GeneID: 1565)	[465, 466]
	Midazolam	N05CD08	*CYP3A4* (GeneID: 1576), *CYP3A5* (GeneID: 1577), *UGT1A4* (GeneID: 54657), *UGT2B4* (GeneID: 7363), *UGT2B7* (GeneID: 7364), *ABCB1* (GeneID: 5243)	[467]
	Allopurinol	M04AA01	*XDH* (GeneID: 7498), *AOX1* (GeneID: 316)	[464]

6.3.2 Modeling drug metabolism under different dietary regimes

The drug reconstruction module was combined with Recon 2 [31], yielding in Recon2_DM1796, where DM stands for drug module, followed by the total number of unique genes contained in the combined metabolic network (Table 6.1). This combined reconstruction was used to interrogate role of diet and genetics on drug metabolism.

Effect of the diet composition on the synthesis and clearance of different drug molecules

We were interested in evaluating the role of diet on the metabolism of drugs. Therefore, we simulated three different dietary regimes, i.e., an average American diet [347], a balanced diet [83, 348], and a vegetarian diet [493]. We defined 50 drug objectives representing the synthesis or breakdown of the various drug metabolites and their derivatives (Table 6.3). In order to avoid any metabolic interaction between drugs, only one drug was given at a time. Flux balance analysis [11] was performed on Recon2_DM1796. While the majority of the drug objectives (44/50, 88%) had the same maximal flux values under all three diet conditions, considerable difference was found for six objectives. When compared to the average American diet, 81% flux reduction in vegetarian diet and 79% flux reduction in balanced diet was observed for the glucuronidation reactions of cerivastatin-M1, rosuvastatin, and for the spontaneous conversion of glucuronide to lactone form of pitavastatin. More-

over, 90% flux reduction for the conversion of pro-drug to active drug was predicted for lovastatin and simvastatin, under both vegetarian and balanced diet. Additionally, the sulphation of acetaminophen had a low flux value only under vegetarian diet, i.e., 62% reduction when compared to the average American and balanced diet (Table 6.3). These modeling results suggest that diet does influence the capabilities of the human body to eliminate these drugs.

Effect of the diet composition on drug excretion and key cellular pathways

While it has been known previously that diet influences drug metabolism, we can now use the model to investigate how drug metabolism may affect cellular metabolism. We therefore performed flux variability analysis (fastFVA [383]), which computes for each reaction of interest its possible flux range (minimally and maximally feasible flux value), given some simulation constraints. Changes in flux range upon altered simulation conditions (e.g., diet composition or drug uptake) are one way to identify how the models compensate for the alterations. We computed the secretion pattern of Recon2_DM1796 for various drug metabolites (124) under the three different diet conditions. Each drug was given individually in a particular diet, allowing the comparison of diet and drug specific models. As expected, all drug metabolites known to be excreted were found to be secreted under all three diet conditions. In most cases, the diet composition did not alter the excretion flux. However, under vegetarian diet compared to the average American diet, 6 out of 124 (5%) drug metabolites had a reduced flux (ranging from 62% to 81 % flux reduction) through the secretion reactions and corresponded to acetaminophen metabolites. These observations are consistent with our aforementioned results.

We then analyzed the flux range through some important cellular metabolic reactions, including reactions of the electron transport chain, urea cycle, lactate production via anaerobic glycolysis, and the hexose monophosphate shunt. The administration of drugs did not cause any change in the overall metabolic network, when intra-diet conditions were compared between presence and absence of a drug. However, when we compared the flux ranges through these reactions between the dietary conditions, average American diet showed highest flux values for all the above pathways (except lactate production), where the 'ATP' synthase reaction ('ATPS4m') had the largest increase in maximally feasible flux value (90%), when compared to balanced and vegetarian diets. This result is independent of the presence of a drug and indicates that the average American diet provokes a high energy state of the human metabolism, consistent with increased occurrence of obesity and diabetes in individuals consuming this diet [494].

To study the relationship between drug metabolism in healthy and disease conditions, we performed single gene deletion [13] for all metabolic genes captured in Recon2_DM1796, and computed the maximally feasible flux through each of the drug objectives, under the average American diet. An objectives was deemed 'affected' by a single gene deletion when its flux was at least 20% reduced compared to the healthy model. Overall, we found that the majority of the single gene deletions did not affect the different drug objectives. However, 73 out of the 1796 (4%) gene deletions altered the flux through at least one out of 34 unique drug objectives (Figure 6.2). For instance, the deletion of the gene encoding the cytidine monophosphate kinase (*CMPK1*, GeneID: 51727, E.C. 2.7.4.14) affected the highest number of drug objectives (28), highlighting the importance of this reaction for the metabolism of numerous drug molecules. In fact, CMPK1 phosphorylates uridine-monophosphate to uridine-diphosphate (UDP), which is consequently converted into uridine-triphosphate (UTP) by the nucleoside-diphosphate kinase. UTP is required for the synthesis of UDP-glucuronate, an important co-factor in glucuronidation reactions for the majority of the modeled drugs (Figure 6.3A). The highest number of single gene deletions (54) altered the flux through drug objectives for the glucuronidation of cerivastatin-M1 and of rosuvastatin as well as for the spontaneous conversion of glucuronide to lactone form of pitavastatin.

We analyzed the relevance of these single gene deletions to known diseases by mapping the recently published compendium of inborn errors of metabolism [29]. We found that 25 of 73 (34%) gene deletions correspond to 12 unique IEMs (Figure 6.2). Interestingly, 80% of these affected genes (20/25) encoding for proteins of energy metabolism (electron transport chain, inosine-monophosphate synthesis, pyrimidine metabolism, TCA cycle, and glycolysis) and are predicted to alter particularly the statin drug metabolism. Further 56% of the IEM associated gene deletions (14/25) encode for complex I & II proteins of the mitochondrial electron transport chain (Figure 6.2).

Table 6.3: Performance of drug metabolic objectives on different diets. D1: average American diet, D2: balanced diet, D3: vegetarian diet

Drug group	Drug metabolic objectives	D1	D2 v/s D1	D3 v/s D1
Statins	Conversion to active hydroxyacid lovastatin	✓	↓	↓
	Oxidation to 6-beta-hydroxy-lovastatin	✓	✓	✓
	Oxidation 6-exomethylene-lovastatin	✓	✓	✓

Drug group	Drug metabolic objectives	D1	D2 v/s D1	D3 v/s D1
	Conversion to active simvastatin	✓	↓	↓
	Conversion of hydroxy simvastatin lactone to acid form	✓	✓	✓
	Conversion of methyl simvastatin lactone to acid form	✓	✓	✓
	Conversion of carboxy simvastatin lactone to acid form	✓	✓	✓
	Oxidation to 4-hydroxy-atorvastatin-acid	✓	✓	✓
	Oxidation to 2-hydroxy-atorvastatin-acid	✓	✓	✓
	Isomerization 3-alpha-iso-pravastatin	✓	✓	✓
	Isomerization to 6-epi-pravastatin	✓	✓	✓
	Beta oxidation to tetranor-CoA-pravastatin	✓	✓	✓
	Oxidation to 5-hydroxy-fluvastatin	✓	✓	✓
	Oxidation to 6-hydroxy-fluvastatin	✓	✓	✓
	Conversion to des-isoproylpropionic-acid-fluvastatin-CoA	✓	✓	✓
	Hydroxylation of cerivastatin to M23	✓	✓	✓
	Hydroxylation of cerivastatin-M1 to M24	✓	✓	✓
	Glucuronidation of cerivastatin-M1	✓	↓	↓
	Glucuronidation of rosuvastatin	✓	↓	↓
	Secretion of N-desmethyl-rosuvastatin	✓	✓	✓
	Conversion of glucuronide to pitavastatin-lactone	✓	↓	↓
Anti-hypertensives	Oxidation to active metabolite EXP3174 (losartan)	✓	✓	✓
	Glucuronidation to losartan-M7	✓	✓	✓
	Formation of losartan-N1-glucuronide	✓	✓	✓
	Tolyl methylhydroxylation of torasemide	✓	✓	✓
	Oxidation of torasemide-M1 to M5	✓	✓	✓
	Oxidation of nifedipine to nitropyridine metabolite	✓	✓	✓
	Oxidation of acid metabolite of nifedipine	✓	✓	✓
Immuno-suppressants	Oxidation of cyclosporine to AM1	✓	✓	✓
	Oxidation of cyclosporine to AM9	✓	✓	✓
	Glucuronidation of AM1c-cyclosporine	✓	✓	✓

Drug group	Drug metabolic objectives	D1	D2 v/s D1	D3 v/s D1
	Sulphation of cyclosporine for excretion	✓	✓	✓
	Demthylation to 13-O-desmethyl tacrolimus	✓	✓	✓
	Demethylation to 15-O-desmethyl tacrolimus	✓	✓	✓
	Oxidation of tacrolimus	✓	✓	✓
Analgesics	Chiral conversion of iburpofen	✓	✓	✓
	Oxidation of S-3-hydroxy ibuprofen	✓	✓	✓
	Secretion of 1-hydroxy-ibuprofen	✓	✓	✓
	Secretion of 1-hydroxy S-ibuprofen-glucuronide	✓	✓	✓
	Glucuronidation of acetaminophen	✓	✓	✓
	Sulphation of acetaminophen	✓	✓	↓
Miscellaneous	Oxidation to methyl-hydroxy-gliclazide	✓	✓	✓
	Oxidation to carboxy-gliclazide	✓	✓	✓
	Secretion of 6-beta-OH-gliclazide-glucuronide	✓	✓	✓
	Oxidation to 1-OH-midazolam	✓	✓	✓
	Oxidation to 4-OH-midazolam	✓	✓	✓
	Glucuronidation of 1-OH-midazolam	✓	✓	✓
	Oxidation of allopurinol	✓	✓	✓
	Conjugation to 1-riboside	✓	✓	✓
	Conjugation to 7-riboside	✓	✓	✓

6.3.3 Atorvastatin and cyclosporine drug-drug interaction: single gene deletion & secretion profile

One of the most fascinating aspect of drug metabolism is in the drug's ability to interact with each other, when administered together. In order to study the ability of Recon2_DM1796 to capture the metabolic interaction of known drugs, we chose to study the interaction between statins and immunosuppressants. The joint administration of atorvastatin and cyclosporine was simulated, under the average American diet. We predicted that 21 drug metabolites can be secreted, seven derived from atorvastatin, and eleven from cyclosporine. No change in maximally feasible flux rate was predicted for these secretion exchange reactions, when given together or individually. We also investigated the effect of the genotype on the drug interac-

tion by comparing simulations with and without gene deficiency. We computed the maximally feasible flux through all 80 associated drug reactions and found that 48% of the reactions (38/80) were affected by at least one of 58 gene deletions (Figure 6.4). Interestingly, 46 of the 58 (79%) affected genes encoded for proteins of the mitochondrial electron transport chain, resulted in reduced flux through reactions of atorvastatin metabolism, and of the transport of the two drugs. Additionally, two of the gene deletions, being cytidine monophosphate kinase (GeneID: 51727) and UDP-glucose 6-dehydrogenase (GeneID: 7358), affected metabolic reactions of both drugs but only elimination reactions of atorvastatin.

6.3.4 Metabolic cost of drug metabolism

One of the chief aims of this study was to analyze the effect of drug metabolism on cellular energy levels. Beside the energy utilization for biomass synthesis, cells require energy to meet maintenance functions (e.g., turgor pressure and membrane potential). This is usually modeled by an ATP hydrolysis reaction placed in the cytosol (atp + h2o $\rightarrow$ adp + pi + h, DM_ATP). Each drug was administered *in silico* individually and its effect on the ATP hydrolysis reaction was computed under the three diets. While majority of the drugs caused reduction in cellular eneergy levels, only two drugs (gliclazide and nifedipine) did not alter the maximal possible flux through the ATP hydrolysis reaction (Table 6.4). Fluvastatin and pravastatin increased the ATP yield by synthesizing 29 milimoles and 9.5 milimoles per milimole of drug administered, under the average American and 20.3 milimoles under the balanced diet and 5.4 milimoles under the vegetarian diet, respectively. This increase was seen due to their ability to undergo beta-oxidation, hence contributing to cellular energy levels. The drugs that caused the highest reduction in cellular energy levels were torasemide, acetaminophen, ibuprofen, losartan, and midazolam under all the three diets. Overall, the effect of drug metabolism on ATP hydrolysis capacity was more profound under average American diet, which represent high energy diet.

To assess the susceptibility of drug metabolism for single gene disorders, we performed single gene deletion of the drug module, maximizing for the ATP_demand reaction, under the three diets. Using a 20% cutoff for the ratio between healthy and single gene deficiency flux values, in total 97 out of 1796 gene deletions (5.4%) caused a reduction in the flux values through ATP_demand reaction (Figure 6.5). Among these, 48 gene deletions were common between the three diets, with four specific to average American, 45 unique to balanced diet, and 44 unique to vegetarian diets. When analyzed for association with IEMs, 37 of these (38%) were found to be associated with 26 unique IEMs.

Table 6.4: Metabolic cost of drug metabolism. Effect of drugs and diets on ATP hydrolysis capacity

Drugs	Milimole ATP per milimole drug in average American diet	Milimole ATP per milimole drug in balanced diet	Milimole ATP per milimole drug in vegetarian diet
Lovastatin	3.5	0.8	0.8
Simvastatin	2.5	-0.3	-0.3
Atorvastatin	-2	0.8	0.8
Pravastatin	-9.5	-5.4	-5.4
Fluvastatin	-30	-20.3	-20.3
Cerivastatin	5.3	2.3	2.3
Rosuvastatin	3.8	2.5	2.5
Pitavastatin	3	3.5	3.5
Losartan	11.5	6.5	6.5
Torasemide	12.7	9.9	9.9
Nifedipine	0	0	0
Cyclosporine	2.1	0.8	0.8
Tacrolimus	2.1	0.8	0.8
Ibuprofen	11.8	6.8	6.8
Acetaminophen	12.5	7.5	7.5
Gliclazide	0	0	0
Midazolam	11.5	6.5	6.5
Allopurinol	2.5	-0.3	-0.3

6.4 Discussion

The presented drug module represents a comprehensive, manually assembled collection of metabolic reactions for the 18 most commonly used drugs. In conjunction with the most recent human metabolic reconstruction, we investigated the role of dietary regime and genetic background on drug metabolism. Our key results include: (i) diet plays a crucial role in the metabolism and elimination of acetaminophen and statins, (ii) disturbed statin metabolism may contribute to the phenotypic effects of mitochondrial energy disorders, (iii) metabolic interactions between atorvastatin and cyclosporine in individuals with inherited metabolic diseases may occur due to common metabolic and transport reactions, using common metabolites and transport proteins, and (iv) energy is drained into metabolizing drugs.

Diet has been known to play both a direct and indirect role in drug metabolism, ei-

ther by inducing/inhibiting phase-I enzymes (e.g., while garlic inhibits CYP3A4 activity [495], caffeine induces CYP1A1/1A2 activity [496]) or by affecting the drug transport systems (e.g., fruit juices and herbal supplements induce the P-glycoprotein activity [425,497]). Additionally, gut microbiota modulates drug metabolism either alone or in combination with ingested dietary ingredients [498]. In the current study, we analyzed the effect of three different dietary regimes on various drug metabolic pathways. The diets included an average American diet [347], a balanced diet that constitutes the necessary inputs required to maintain a healthy lifestyle [348], and a vegetarian diet based on a mixture of corn, lima beans, snap beans, green peas, and carrots [493]. The metabolism of acetaminophen and statins was predicted to be greatly altered depending on the diet. For instance, the simulated vegetarian diet was low in L-cystin, a sulfur containing amino acid, which resulted in a reduced excretion capability for six acetaminophen metabolites as well as reduced flux through the acetaminophen sulphation reaction, when compared to the two other diets (Table 6.3). The sulfur contained in such amino acids enters various metabolic pathways, including the biosynthesis of phosphoadenylyl sulfate (paps). Paps is an important cofactor in the sulphation reaction of acetaminophen (Figure 6.3B), in which the sulphate conjugate of acetaminophen is formed by the sulpho-transferase 1 (E.C.2.8.2.1). Of an administered acetaminophen dosage, about 30-50% is recovered in urine in the form of sulphate conjugates [457]. Sulfur-amino acids as well as sulfur deficient diets have been experimentally shown to cause reduced acetaminophen elimination [499], hepatic necrosis, and liver injury in rats [500]. Furthermore, consistent with our predictions, apiaceous vegetable diets, consisting of carrots, parsnips, celery, dill, and parsley, which are generally low in sulfur containing compounds, have been reported to cause reduced activity of acetaminophen metabolic enzymes [501]. The average American diet utilized more of carbohydrates (e.g., starch, maltose, lactose, sucrose) and lipids, such as diacylglycerol and fatty acids, while the balanced and vegeterian diets utilized more of amino acids (i.e., glycine, alanine, leucine, valine, isoleucine) . The carbohydrates are directly fed into energy production, via their conversion into glucose. Lipids yielded ATP via lipolysis of diacylglycerol, followed by beta-oxidation of fatty acids . Glycine enters glycine cleavage system, yielding 'nadh', another energy precurssor. Alternatively, glycine can also produce glycoaldehyde via a five step conversion, which ultimately produces 'nadh' (via a dehydrogenase). On the other hand, alanine, leucine, valine, and isoleucine are utilized via their respective transamination reactions. Due to this shift in favourable energy precurssors, which is dictated by the presence of a particular diet, there was a marked difference in efficacy of statin metabolism. These statins require energy for their elimination via ABC transporters and for UDP synthesis, which in turn is needed glucuronidation of the statins.

The cellular energy state influences the drug metabolism, which has been particularly reported for antipyrine and aminopyrine drugs [502] as well as for phenazone and theophylline drugs [426], where an energy deficiency resulted in low clearance

and elimination of these drugs. We observed in average a 80% flux reduction for the drug objectives of numerous statins under balanced and vegetarian diets (Table 6.3). When the general cellular reactions were analyzed, average American diet had much higher flux through ATP synthase, complex I -IV reactions of electron transport chain, and biomass reaction, implying that a higher cellular energy level is available for growth and non-growth associated functions under this diet. Hence, we suggest that a diet rendering energy rich condition to the cell, is favorable for an efficient drug metabolism and ultimately their excretion, as seen in case of the statins.

Myopathy is associated with 0.1% patient population on statin monotherapy, which may lead to rhabdomyolosis (i.e., severe muscle damage, myoglobinuria, acute renal failure and death) when combined with other drugs, such as cyclosporine, digoxin, and erythromycin. This adverse drug reaction may be due to complex interactions between the drugs and disease [503]. Atorvastatin and cyclosporine are commonly prescribed after organ transplantation and rhabdomyolysis due to their concomitant usage has been reported in renal [504] and cardiac [505] transplant patients. Various mechanisms of statin-cyclosporine interactions have been proposed, namely via shared metabolic pathways (as both the drugs are CYP3A4 substrates) as well as common elimination routes via P-glycoprotein [506]. Based on our simulations, we propose that additional metabolic routes, such as UDP-glucose 6-dehydrogenase (E.C. 1.1.1.22) and UMP-CMP kinase (E.C. 2.7.4.14), play a significant role in the drug-drug interaction. The UMP-CMP kinase provides UDP for synthesis of UDP-glucuronate that is required for the glucuronidation of these drugs. Statin induced myopathy may also be caused due to genetic predispositions, particularly seen in individuals with IEMs causing muscle disease/exercise intolerance, e.g., McArdle disease (OMIM: 232600), CPT-II deficiency (OMIM: 255110), myoadenylate deaminase deficiency (OMIM:102770) [507], as well as a number of neuromuscular disorders [508]. Simultaneously, genetic polymorphism in a number of CYP, OATP, and APOE family genes have also been linked to myopathic outcomes by statins [509]. We predicted that a major fraction of the gene knockouts (79%), encoding for proteins of the mitochondrial electron transport chain, affected metabolic as well as transport reactions of both atorvastatin and cyclosporine, indicating their dependence on cellular energy for elimination. Hence, we propose that specific mutations in these genes (Figure 6.4) could alter the interaction profile of these drugs leading to energy deficiency, which could further contribute to their adverse drug reactions.

Disease conditions have been known to influence drug metabolism. Such conditions include liver diseases, endocrine disorders, and various types of infections [423, 510]. Only a limited number of studies have analyzed the effect of drugs on individuals with particular inherited metabolic diseases. The examples include i) statins and lactic acidosis (OMIM: 245400) [428], ii) antiepileptic drugs and di-

hydrolipoamide dehydrogenase deficiency (OMIM: 238331) [429], iii) the drugs methotrexate, trimethoprin, and niacin, and hyperhomocysteinemia (OMIM: 603174), iv) drugs known to induce hemolytic conditions (e.g., antimalarials, sulfonamides, antipyretics) and G6PD deficiency (OMIM: 305900) [250], v) the drugs sulfanol, aminopyrine, and pyrazinamide, and hepatic porphyria [427], and vi) thiopurine drugs and thiopurine S-methyltransferase deficiency (OMIM:610460). We showed that the metabolism of statins was most affected by the deletion of various genes, most of which encode for proteins involved in energy metabolism. In fact, 46 of these genes encode for the NADH and the succinate dehydrogenases (i.e., complex I and II of electron transport chain). In support of these computational results, statins are contraindicated in various mitochondrial disorders and a Coq10 therapy is generally advised [511]. This is due to the fact that several studies have suggested that a prolong and excessive usage of statins may cause depletion of Coq10 pools, which may consequently lead to statin induced myopathy (reviewed in [512]) and to mitochondrial dysfunction. However, conflicting results on Coq10 supplementation exists, suggesting further research on statin induced mitochondrial dysfunction is required [513]. The predicted reduction of statin metabolism in lower energy diets occurred because, all statins and most of their derivatives require ATP for their elimination, which occurs via ABC transport proteins. Further energy is required for the activation of fluvastatin and pravastatin, and also for the interconversion between acid and lactone forms of atorvastatin, simvastatin, and lovastatin. At the same time, even though our simulations of the enzyme deficiencies were carried out under the average American diet, which is rather energy rich, gene deletions in energy metabolism led to a reduced energetic metabolism of the corresponding models. We therefore propose that the observed adverse effect of statins on individuals with mitochondrial disorder may be due to reduced cellular ATP availability.

One way of determining the energy availability of the metabolic model is by computing the maximal possible flux through the ATP hydrolysis reaction. Besides the other drugs, two of three antihypertensive drugs, analgesics and midazolam, significantly reduced the cellular energy availability (Table 6.4), i.e., caused reduction by >10 milimoles of ATP/milimole of drug. A key reason is the that these drugs are dependant on ATP dependant transport processes for biliary excretion of parent and derivative compounds (i.e., losartan, ibuprofen, and acetaminophen). Additionally, losartan, acetaminophen, and midazolam are P-gp (ABCB1, GeneID: 5243, E.C. 3.6.3.44) substrates, which requires ATP hydrolysis for the biliary/luminal excretion. In this regard, the module contains six ABC transporters that dedicatedly transport various drug compounds. Apart from this, ten SLC group transporters require co-substrates that are intermediates of cellular energy metabolism. For example, OAT1 that aids in efflux of torasemide operates in antiport mechanism, exchanging with alpha-ketoglutarate that is intermediate of TCA cycle [488]. On the other hand, knockouts of the genes encoding for proteins of the mitochondrial electron transport chain caused significant reduction in the flux through ATP hydrolysis

reaction under diet and drug administration, indicating the susceptibility of drug metabolism for a defect in energy metabolism.

However, we also encountered conflicting results in case of three drugs, i.e., atorvastatin, simvastatin and allopurinol, when the ATP cost was compard for the thee diets. While there was ATP production by simvastatin and allopurinol in balanced and vegetarian diets, under the average American diet, there was ATP expenditure. The reverse was seen for atorvastatin under the three diets (Table 6.4). A possible reason for such a discrepency could be due to the presence of futile cycles that lead to the energy production by these drugs. Hence, these values needs to be interpreted with caution.

In conclusion, we want to highlight the need for a computational modeling to analyze the effects of diets and enzymopathies on drug metabolism. Furthermore, the need of additional energy precursors to account for efficient operation of drug metabolic pathways, particularly under IEMs of energy metabolism. We present here a drug module, which can be readily combined with the human metabolic reconstruction, to investigate the metabolism of drugs on the overall cellular metabolism. Moreover, tissue specific metabolism of drugs can also be derived from current module (provided that the precise uptake and secretion transporters as well as the metabolic enzymes are sufficiently known), thereby aiding in a more individualized and personalized medication design.

6.5 Materials and methods

6.5.1 Reconstruction of drug metabolism

Drugs were chosen qualifying the following criteria: (i) majorly administered via the oral route, (ii) most commonly used/highly prescribed by the clinicians, (iii) undergo extensive metabolism (including first pass extra-hepatic metabolism) for excretion, and (iv) exhibit significant interactions with other drugs. A thorough manual curation of the scientific literature was performed to obtain information on absorption, distribution, biotransformation, elimination, therapeutic usage, clinical dosage, adverse drug reactions, and associated transporters of the covered drugs. Additional information was obtained from PharmGKB [514], DrugBank [515], Human Metabolome Database [516], Chemical Entities of Biological Interest (ChEBI) database [517], and PubChem database [518]. All the drug compounds were drawn using MarvinSketch 5.7.0, 2011, from ChemAxon (http://www.chemaxon.com). The elemental formula, charge, and chemical structure was thoroughly checked against the aforementioned databases and scientific literature. Genome annotation

was obtained from EntrezGene [37], and protein information from Uniprot [519] and BRENDA [34] databases, and used along with scientific literature for assigning gene-protein-reaction associations (GPR) to metabolic and transport reactions. Reconstruction methodology was kept in accordance with the reconstruction protocol [10].

6.5.2 Collection of dietary constituents

Diet information was collected from the US Department of Agriculture, Agricultural research service [347]. All nutrients were converted to a common unit of grams per day. An average American diet, for both males and females ($\geq$2 years), was collected [347]. A balanced diet representing the recommended dietary reference intakes (DRIs) of micronutrients as well as macronutrients [348], [83], was also assembled. The DRI reflects the minimum requirements advised to be taken daily to maintain a healthy lifestyle. The values were taken for males between 30 and 50 years of age. The total fat and cholesterol intake for balanced diet was also represented [349, 389] . The vegetarian diet information was collected from [493], and calculated for three cups, as suggested by centers for disease control and prevention (http://www.cdc.gov/nutrition/index.html). The vegetarian diet consisted of mixture of corn, lima beans, snap beans, green peas, and carrots (with salt). We did not represent the energy value in food (i.e., caloric load) due to disparity between various sources and also due to lack of specific information. The total carbohydrate content in the diet was converted to starch and dextrin uptake, since polysaccharides are most abundant carbohydrate component present in our diet [83]. In the case of total sugars, disaccharides represent 60% and monosaccharides represent 40%, since free monosaccharides do not represent significant part of the diet [83].

6.5.3 Combining the drug module with the human metabolic reconstruction

The drug reconstruction was assembled and converted into a mathematical model using rBioNet [36] as a reconstruction environment, and an established protocol [10]. All the reactions were checked for mass and charge balance, and manually corrected. Thereafter, it was combined with recently published community driven global reconstruction of human metabolism, Recon 2 [31]. Both the matrices were merged, and analyzed for dead end metabolites and network gaps, using the COBRA toolbox [13]. Gaps were filled by re-visiting scientific literature. Furthermore, the biomass reaction was tested for carrying a non-zero, positive flux value, and futile cycles (only on exchange reactions) were checked.

6.5.4 Simulations

The drug and dietary ingredients dosage values were converted to grams per day. Since, the small intestine acts as a gateway for passage of all ingested components into the body, the model inputs were formulated as gram dry weight of small intestinal enterocytes, i.e., in $mmol \cdot g_{DW,enterocytes} \cdot hr^{-1}$, described in [30]. The absorption rates of the inputs are unknown in most cases; hence, we assumed them to be sub-optimally. The minimum uptake rate was fixed at 10% of maximal possible flux rate. The lower and upper bounds of the corresponding exchange reactions were set accordingly. We obtained diet and drug specific models, and tested them for the fifty drug metabolic objectives, using flux balance analysis [11]. Consequently, the models were also tested for a functional biomass under each of these conditions.

Drug objectives and secretion reactions

Drugs objectives were chosen such that they represent the overall metabolism of each drug, including their transport reactions. However, in cases where it was known that a major part of drugs undergoes a metabolic pathway, e.g., 80% of acetaminophen undergoes non-oxidative route, this one was chosen over the oxidative route. All drug metabolites derived from the parent form of the drug were constrained to be 'only secreted' into the extracellular space, i.e., the corresponding exchange reactions were set to have a lower bound $lb = 0$.

Modeling the effect of drug metabolism on the ATP demand reaction

The biomass reaction was constrained to carry a minimal flux, i.e., the lower bound on this reaction was set to be 10% of the maximal possible flux value. The flux through the cytosolic ATP demand reaction ('DM_ATP') was then maximized. This computation was performed for each drug independently under all three diets. The exchange reaction for each parent drug compounds was set to $lb = ub = -1\, mmol \cdot g^{-1}{}_{DW,enterocytes} \cdot hr^{-1}$. Simultaneously, single gene deletion was performed optimizing for ATP_demand reaction.

Flux balance analysis (FBA)

Assuming steady state condition, where sum of input fluxes equal the sum of the output fluxes, FBA follows $S.v = 0$, where S represents the stoichiometric matrix

of size $m \times n$ (with m being the number of metabolites and n being the number of reactions in the network). v is a flux vector of size $n \times 1$ containing a flux value v_i, for each reaction, i in the network. Under the given simulation conditions, we used FBA to maximize a given objective function, as defined by the constraints, which were applied to the model.

Fast flux variability analysis (fast FVA)

Flux variability analysis [12], uses FBA to maximize and minimize each reaction in the model. The fastFVA [383] implementation was used to perform the FVA.

Single gene deletion analysis

Generally, reactions in a metabolic networks are represented with GPR associations, which are Boolean relationship between the genes that encode the enzymes/proteins that catalyze the reactions, making use of 'or' when isozymes are involved, and 'and', when multimeric enzyme complexes are involved. In a single gene deletion [13], those reactions associated through an 'and' relationship or those encoded by a single gene are inactived by setting the corresponding reaction bounds to zero ($lb = ub = 0\, mmol \cdot g_{DW,enterocytes} \cdot hr^{-1}$), and uses FBA to maximize the given objective function. This procedure was done for all network genes.

All fluxes were computed and given in $mmol \cdot g_{DW,enterocytes} \cdot hr^{-1}$. All simulations were carried out using MatLab (MathWork, Inc.) as programming environment, Tomlab (TomOpt Inc.) as a linear programming solver, and the COBRA toolbox [13]

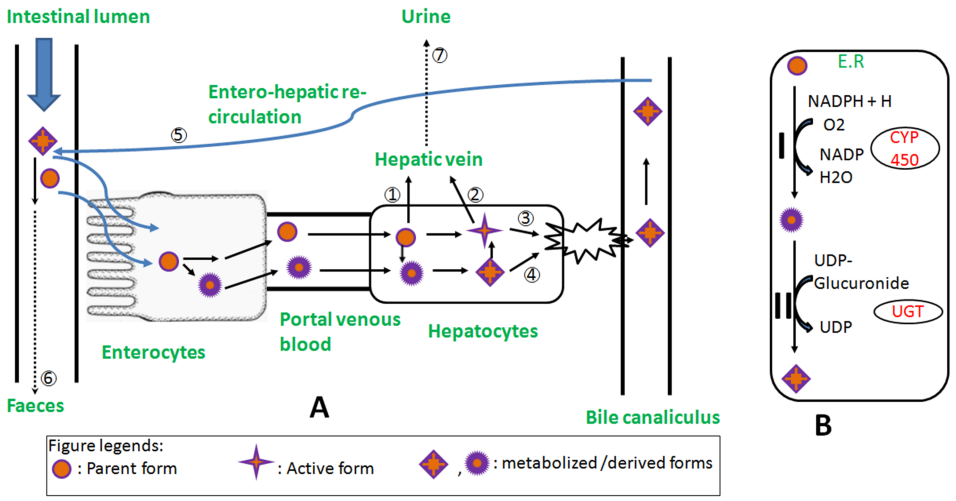


Figure 6.1: Simplified overview of drug metabolism and elimination routes. ***A.*** *Orally administered drugs, are primarily absorbed via the enterocytes of the small intestine, via specific transporters (e.g., SLCO2B1, GeneID: 11309). Once inside the enterocyte, phase I and phase II drugs metabolizing enzymes may act on the parent form of the drug. Alternatively, the parent form may enter into the portal venous blood. Specific hepatic uptake transporters (e.g., SLCO1B1, GeneID: 10599) take up the drug and its derivatives for further metabolism. Majority of the drugs undergo activation from the pro-drug to the active form in liver, and can then either exit into the hepatic vein to be exposed to systemic circulation for targeted action, or can enter the biliary route. Drugs can 1) enter the hepatic vein also in their parent form or 2) in their activated form, which can also enter the 3) bile. On the other hand, they can undergo extensive metabolism in the liver to be released into both hepatic vein and 4) bile. Once inside the bile canaliculus, the drugs can either enter into 5) entero-hepatic re-circulation or be excreted in 6) the feces. Drugs exposed to the systemic circulation usually exit through 7) the renal route, via urine.* ***B.*** *Drugs undergo biotransformation from lipophilic to polar state, which further aids in easy excretion. Shown are phase I reaction, i.e., oxidation of a drug molecule that is catalyzed by enzymes of the cytochrome P450 family (CYP450), followed by phase II reaction, i.e., glucuronidation catalyzed by UDP-glucuronyltransferases (UGT). Both the reactions occur in the endoplasmic reticulum (E.R) of either enterocytes or hepatocytes.*

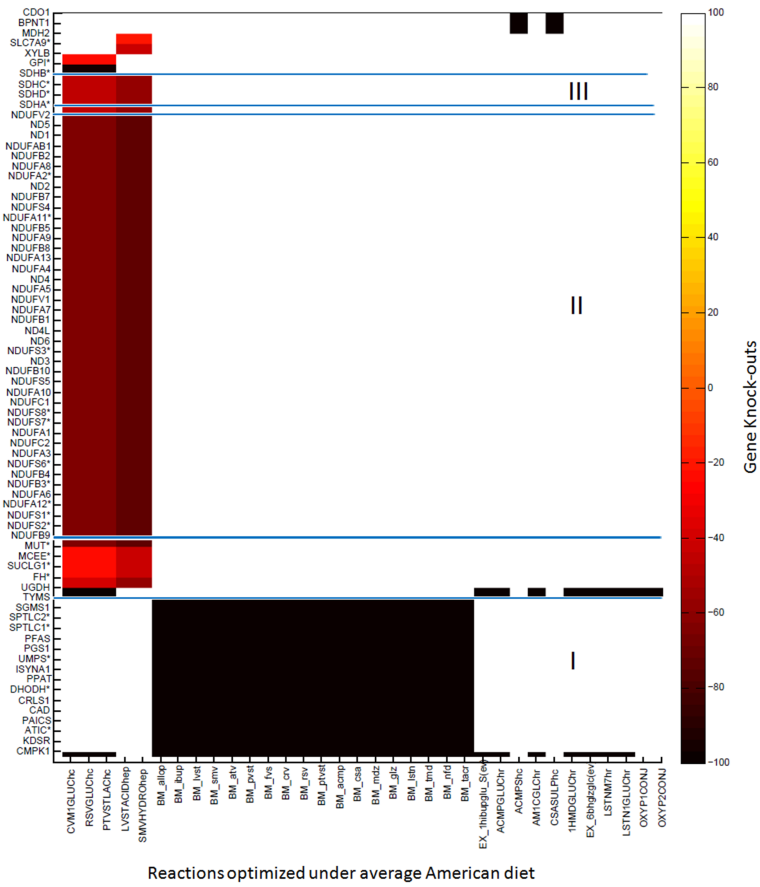


*Figure 6.2: Effect of single gene deletion on the drug objectives under average American diet. Black bars indicate zeros flux, red bars indicate reduced flux, and white bars indicate no change in flux values compared to the healthy model (percentage reduction in the flux values through the mentioned objective reactions are shown). Gene symbols with * are the ones associated with IEMs. The plot has been divided into three areas, where area I consist of 16 biomass essential genes under average American diet constraints. These genes are essential for the biomass reaction to carry a non-zero flux, under the applied simulation constraints. Area II consists of 42 genes encoding for the NADH dehydrogenase, i.e., the complex I. Area III consists of 4 genes encoding for the succinate dehydrogenase, i.e., complex II reactions of the mitochondrial electron transport chain.*

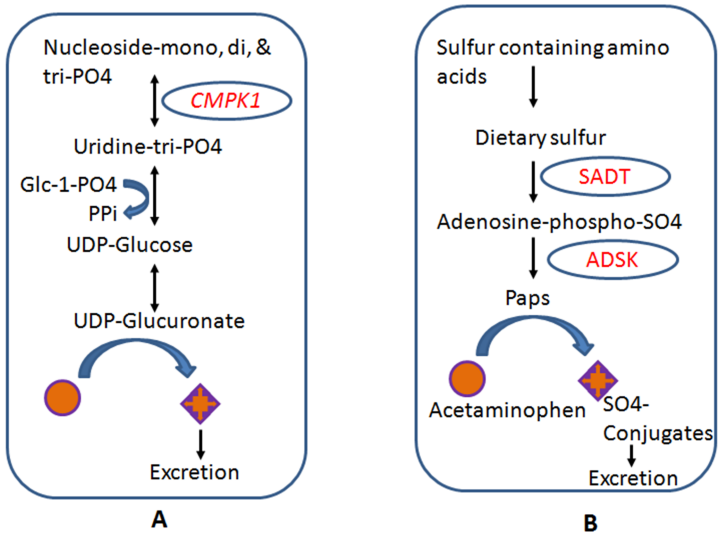


Figure 6.3: Involvement of endogenous metabolites in xenobiotic metabolism: ***A****: UDP-glucuronate is an important co-factor for glucuronidation reactions for a majority of drugs, including statin group, analgesic group, cyclosporine, losartan, midazolam, and gliclazide. Knockout of CMPK1 results in non-availability of uridine di- and tri-phosphates that are required for synthesis of UDP-glucuronate. Hence, a major fraction of drug objectives are affected by CMPK1 knock out.* ***B****. Sulfur containing amino acids are source for endogenous paps (i.e., 3'-phosphonato-5'-adenylyl sulfate) that is involved in sulphation of acetaminophen metabolites, ultimately leading to their excretion. Abbvr: SADT: sulfate adenylyltransferase (E.C. 2.7.7.4), ADSK: adenylyl-sulfate kinase (E.C. 2.7.1.25).*

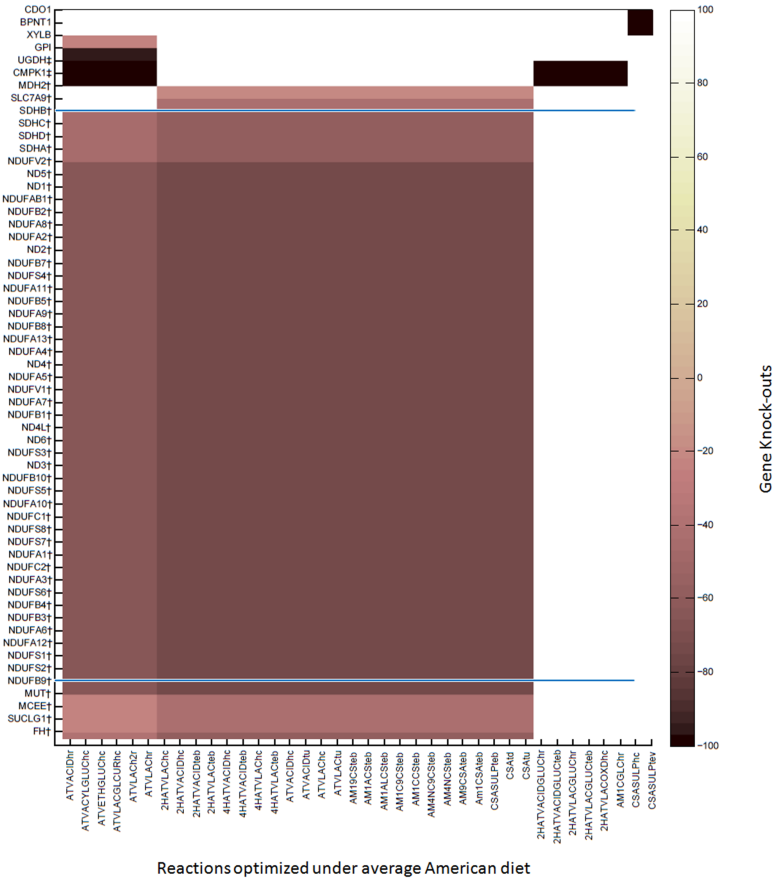


Figure 6.4: Drug interactions between atorvastatin and cyclosporine. Gene symbols with ‡ and † correspond to those that are altered by both drugs. Genes with ‡ represent the ones affecting metabolic reactions of atorvastatin and cyclosporine, but, transport reactions of only atorvastatin. Genes with † are the ones affecting transport reactions of atorvastatin and cyclosporine, but, metabolic reaction of only atrovastatin The area between dashed lines highlights the genes encoding the proteins of mitochondrial electron transport chain. Black bars represent drug objectives carrying zero flux, and pink bars are the ones carrying reduced flux (percentage reduction in the flux values through the mentioned objective reactions are shown).

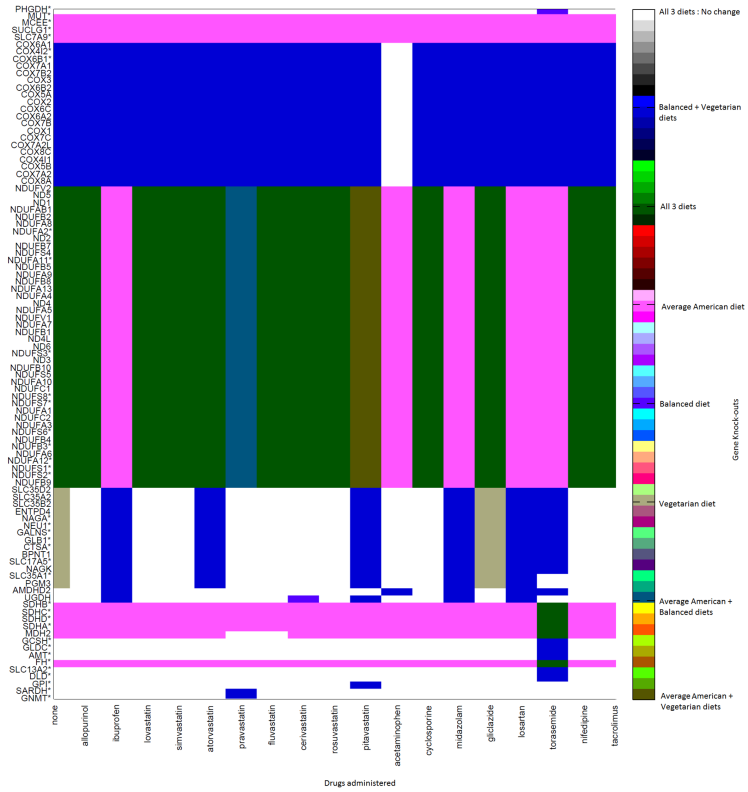


*Figure 6.5: Effect of drug metabolism on cellular ATP maintenance under three diets. Drugs have been shown on the X-axis and all the gene knockouts on the Y-axis. The ATP_demand reaction was optimized under different diets and single gene deletion performed. Gene knockouts that caused a reduction in the flux value through ATP_demand has only been shown. Gene symbols with * are the ones associated with IEMs. Different color codes represent the reduction in the flux values through the ATP-demand, under different diet conditions, and white bars represent no change.*

7 Conclusion and perspective

7.1 Human metabolism

Human metabolism is integrated, wherein the major metabolic pathways for carbohydrates, lipids, and amino acids, come together, in order to extract various ingredients from the diet, and provide the body with energy and cell maintenance essentials. Additionally, vitamins and minerals provide the required co-factors, and meanwhile, the hormones mediate the pathway regulation (condition dependent hormonal cascades, e.g., fully fed and starvation states). Each organ then runs specific pathways depending upon the dietary inputs received and presence of hormonal influence ensuring the body homeostasis. This balanced mechanism can be disturbed under metabolic and non-metabolic diseases. Hence, a system analysis of not only individual component, but also, their complex interaction can aid in restoring back normalcy. The work presented in this thesis, aimed to understand the mechanisms of cellular homeostasis and its perturbation under various IEMs.

7.2 Genome scale network reconstruction (GENRE) for capturing cellular metabolism

Building metabolic GENREs remains one of the fundamental methods of systems biology to study cellular biochemistry, physiology, and genetics. Throughout this thesis, metabolic network reconstructions have been done using the manual or bottom-up approach. This method uses the genome annotation data to start with, and involves intensive study of the relevant scientific literature to accumulate information on the operating metabolic and transport pathways, the genes, proteins and their relationship. Subsequently, when the reconstruction is converted into a mathematical model, the COBRA methods, e.g., FBA, FVA etc. were used to interrogate the network properties, and test the functionality. Although, laborious and time consuming, such approach continues to be the most realistic and reliable method of GENRE building that is more close to the metabolic and physiological processes occurring within that organism or cell type. The two human specific GENREs, i.e., the first global human metabolic network Recon 1 [18] and the most recent Recon 2 [31] represent milestones in exploring and capturing the various biochemical

mechanisms responsible for attaining cellular homeostasis. These GENREs built through extensive manual curation efforts, lay their foundation on human genome annotation and scientific literature. Hence, these mother GENREs attempted to capture all the known metabolic pathways functioning in a human body, accounting for their correct sub-cellular localization, substrate and cofactor utilization, stoichiometry, and directionality. Additionally, they provide catalogue for genes that encode the proteins catalyzing the metabolic and transport reactions, through establishing the GPR associations. A plethora of studies that subsequently utilized Recon 1 to not only build cell/tissue specific networks, disease specific networks, but also other mammalian networks, which all together furthered biomedical prospects of systems biology in general. However, as the scientific community continues to grow at a dynamic rate, so does our understanding. Therefore, frequent updates, refinements, and corrections needs to be periodically incorporated into the mother GENREs, which shall retains their continuous usage as well as provide the most current knowledge (Chapter 2).

7.3 Present status of IEM diagnosis and treatment : Use of specific diet and medications

Inborn errors of metabolism (IEMs) are the metabolic disorders, caused due to specific mutation in the genome. Although individually rare, but, collectively numerous, they have an incidence rate of 1:800 live births [3]. Among these, fatty acid oxidation disorders are the most highly prevalent ones, with a collective incidence rate of 1:9000 [258] live births. IEMs in general exhibit a myriad of clinical symptoms, and affect multiple organ systems, with nervous system being the most affected one. These disorders follow various inheritance (with the majority being autosomal recessive) and a range of phenotypic patterns. Typically, they are associated with either a mild form or a moderate to severe form. The diagnosis at an early stage is done via newborn screening program that measures amino acids, acylcarnitines and their ratios, respectively. While the aminoacidopathies and organic acidemias (the first group of IEMs that give rise to intoxication in general) are usually treatable, others are non-treatable. The preferable mode of treatment remain use of specific diet and medications.

7.4 Summary of research methodology used

In order to understand the pathology of metabolic disorders, a sound knowledge on human metabolism is integral. The use of bottom-up reconstruction as it is based on

conclusive biochemical data, and subsequent constraint based modelling approach, provides mechanistic explanation for complex biological interactions. Additionally, occurrence of higher number of network redundancies necessitates assembly of a cell/tissue specific network. However, reconstructions are not complete and needs to be updated frequently as new knowledge becomes available. Recon 1 [18] and Recon 2 [31] were used as starting points. Expansion of these networks was done in order to account for missing transporter content (Chapter 3) and detailed lipid metabolism (Chapter 4). We mainly explored the impact of specific dietary patterns (Chapter 5) and commonly prescribed medications (Chapter 6) on IEMs and cellular biochemistry. Therefore, the work presented in this thesis involved:

- A review of literature to investigate the plasma membrane transporter content of global human metabolic networks, Recon 1 and Recon 2, for major metabolite classes (i.e., sugar, amino acids, peptides, lipids, nucleosides, vitamins, and others comprising water, heme, carnitine, taurine and betaine). This resulted in expansion of substrate coverage for the various transport proteins as well as addition of new proteins and improved reactions for their transport mechanism. The proposed new transport reactions were combined into a functional 'transport module' for its ready usage. The transport module comprised of 71 metabolites, 70 reactions, and 41 genes. Faulty transport mechanism have been reported in various metabolic disorders, hence, these IEMs were highlighted along with their extent of coverage (in terms of their associated genes) in the human GENREs.

- Expanding the global human metabolic network in order to account for all the metabolites used in newborn screening (i.e., amino acids and acylcarnitines). The lipid metabolism in Recon 1 was extended with 139 metabolites, 352 reactions, and 14 genes, in order to include complete alpha, omega and beta-oxidation pathways that included synthesis of these acylcarnitines, resulting in an acylcarnitine/fatty acid oxidation module (Ac/FAO module). This was built by manual curation of relevant scientific literature, and combined with Recon 1 to form a functional Recon1_Ac/FAO network. Thereafter a compendium of IEMs was assembled and mapped onto this extended network that comprised of mapped (235 IEMs) and un-charted (139 IEMs) ones. These IEMs were characterized based on their mode of inheritance, phenotypic presentations, biomarkers/metabolites used in their identification, as well as the various organs/organ systems affected. Topological analysis of Recon1_Ac/FAO network was used to infer the network properties.

- The previous work revealed that IEMs are majorly treated by specific dietary measures and medications. The small intestine is the chief organ that performs majority of digestion and absorption of food components. Reconstruction of the first human small intestinal enterocyte specific metabolic network

(hs_sIEC611) was built in a bottom-up fashion, via an in-depth manual curation of available scientific literature. Information on various metabolic and transport reactions specific for small intestinal enterocytes was collected, and the corresponding reactions were extracted from Recon 1. Additionally, new reactions were added particularly for peptides and dietary fiber metabolism, all together comprising 433 unique metabolites, 1282 reactions (distributed over five intracellular compartments, i.e., cytosol, mitochondria, nucleus, peroxisome and endoplasmic reticulum, and two extracellular compartments, i.e., extracellular space and lumen) and 611 genes.

- The dietary patterns for an average American diet and balanced diet were used as model constraints, and metabolic tasks (metabolic reactions performed by a small intestinal enterocyte as inferred from literature) were set as model objectives, and tested using flux balance analysis. Additionally, flux variability was used to analyze secretion profile as well as whole network changes. hs_sIEC611 captured 109/235 IEMs, hence, was used to study the effect of diet and enzymopathies on cellular metabolism. Furthermore, the metabolic link for the clinical symptoms observed in certain IEMs (porphyrias, ornithine translocase deficient and gyrate atrophy) was investigated, as well as novel comorbidities (between methylmalonic acidemia, propionic acidemia and cystathionuria) and probable cellular adaptive mechanisms (for type I citrullinemia and Smith-Lemli-Opitz syndrome) were proposed.

- The role of diet in metabolizing drugs has been widely explored. However, there exists limited number of studies that analyzed the effect of genetic perturbations/genetic disorders on drug metabolism. We chose to study the metabolism of commonly prescribed drugs (i.e., statin group, anti-hypertensives, immunosuppressants, analgesics, anti-diabetic, sedative, and anti-hyperuricimic drugs) by manually assembling a drug reconstruction, comprising 210 metabolites, 721 reactions, distributed over four intracellular (cytosol, endoplasmic reticulum, mitochondria, and peroxisome, and one extracellular compartment) and 57 genes. This was combined with Recon 2 to form a functional Recon 2-drug module (Recon2_DM1796).

- In addition to the average American and balanced diets, formulated in the previous study, we formulated a vegetarian diet. These were used as model constraints, along with the drugs constraints (e.g., secondary drug metabolites derived from the parent form, were constrained to be only secreted, as lb = 0, ub = 1000), and the drug objective were set as model objectives. Drug and diet specific models were obtained from Recon2_DM1796 and the role of diet analyzed. Moreover, single gene deletion analysis was performed to account for the effect of IEMs.

7.5 Main findings of the research performed

IEMs have long been segregated as 'rare diseases', confined to a particular geographical location. However, as MS/MS technology is becoming more precise, targeted, as well as sensitive, new IEMs are being identified. This demands a more elaborate and systems analysis of their specific pathological mechanisms. Therefore, one of the best platforms to study the various abnormal processes caused due to an IEM and predicting possible adaptive mechanisms occurring in such cells, depends heavily on building IEM specific *in silico* metabolic models.

7.5.1 Biomedical prospects : IEMs, biomarkers, biomedicine

When the newborn screening data was mapped onto Recon 1, the acylcarnitines could not be captured. Hence, the existing lipid metabolism (alpha, omega, and beta oxidation pathways) was expanded to account for the newborn screening data, which could aid in IEM identification within a metabolic network context [29]. Furthermore, topological analysis of this network resulted in seven IEM pairs that shared all reactions (i.e., affected reactions) and the literature study revealed that they do also have overlapping clinical symptoms. This signifies a systems approach to identify and compare IEM pairs that may not be associated otherwise. Additionally, acetyl-CoA was found to be a common link, connecting various IEMs on the metabolic map, revealing the metabolic 'hotspots'.

A significant part of this work involved the assembling of the IEM compendium (comprising the mapped and uncharted IEMs) [29], which paved the way for further study. Since, diet and medications are the most commonly used treatment strategy, the next level of work focused on these areas.

Enterocytes of the small intestine are majorly involved in digestion and absorption of food components. Hence, we built hs_sIEC611 [30], human small intestinal enterocyte metabolic network, in a bottom-up approach, which essentially captured the physiology and biochemistry of this cell type. The building of a cell type specific metabolic network was necessitated to overcome the generic network's metabolic redundancy, as well as to observe the local effects of specific dietary patterns and IEMs specifically on this cell type. Dietary patterns not only dictated the model predictions under healthy state, but also, IEMs inflicted models. The derangement of the enterocyte metabolic functions by IEMs, as well as successful prediction for metabolic biomarkers for known IEMs, projects the potential of computational modelling for studying genotype-phenotype correlations and IEMs identification. Although, this study was a qualitative one, but, if adequate kinetic parameters are

known this can be used as starting point to make quantitative predictions. One of the biggest advantage that *in silico* modelling provided was inferring mechanistic explanations for the clinical symptoms observed and their biochemical connection. This was achieved for eight different porphyrias, gyrate atrophy, and ornithine translocase deficiency, again showing possible applications of computational modelling to decipher the precise IEM pathology.

Can we investigate why multiple IEMs co-occur in a patient? Why do such IEMs have overlapping clinical features? This was studied for methylmalonicacidemia (MMA), propionic acidemia, and cystathionuria. Patients positive for both methylmalonic acidemias and cystathionuria has been reported in the literature [353]. Additionally, MMA and propionic acidemia share common clinical features. The computational modelling predicted a failure to synthesize phosphatidylcholine (phosphatidylcholine being an important membrane constituent exerting anti-inflammatory role [380, 381]), in both MMA and propionic academia, and a flux block in methionine to cysteine conversion common to all three IEMs. These findings suggest that co-morbid patterns as well as overlapping phenotypes may be linked to underlying biochemical defects.

The next question that we tried to answer was; what are the special alternative mechanisms possessed within cells that help them bypass the deleterious effects of IEMs. This was analyzed for type I citrullinemia and Smith-Lemli-Opitz syndrome, where in we were able to identify potential adaptive mechanisms. Further, any form of metabolic disturbance of these target reactions could be responsible for the broad phenotypic spectrum observed in these IEMs.

There have been very few reports that analyzed the effect of genetic background on efficacy of drug metabolism. Additionally, few reports claimed that co-administration of multiple drugs to IEM patients caused an aggravation of clinical symptoms, further worsening the pathological consequences (e.g., prolonged use of statins and cyclosporine in cases of IEMs of energy metabolism caused increased muscle problems, leading to rhabdomyolosis) [507]. Previously we studied the effect of popular diets and enzymopathies on cellular metabolism, we next chose to study their effects on xenobiotic metabolism, namely for the most commonly prescribed drugs. We saw that cells had to bear the burden of extra metabolic cost, in terms of ATP expenditure, when faced with IEMs, in order to carry out an efficient drug metabolism. This was highly reflected in case of IEMs of mitochondrial electron transport chain disorders (e.g., Leigh syndrome, complex I deficiency, succinate dehydrogenase deficiency, etc.) that caused marked reduction in the flux through the statin metabolism. Hence, a disturbed statin metabolism may be responsible in part for the phenotypic effects of mitochondrial energy disorders, when such patients are put onto statin therapy. Additionally, we also observed that the choice of diet significantly altered the efficacy of drug metabolism and ultimately their elimination

(e.g., a low sulfur containing amino acid diet contained in the simulated vegetarian diet caused reduced flux through metabolism and elimination of acetaminophen metabolites). Hence, with this study we showed that diet and IEMs significantly altered the xenobiotic metabolism, which may cause severe phenotypes, as seen for IEMs of energy metabolism. In case, of co-administration of drugs, i.e., for atorvastatin and cyclosporine, both shard not only common transport proteins but also metabolic reactions.

7.6 Limitations

The key enzymes of metabolic pathways are strictly regulated (either by hormones or other forms of external stimuli), e.g., hyperglycemic hormones (glucagon, glucocorticoids, and thyroid) and hypoglycemic hormone (insulin) reciprocally regulate the glycolysis and gluconeogenesis pathways by monitoring transcription and translation of pyruvate kinase, phosphofructokinase, and the bi-functional enzymes [520]. However, metabolic networks generally doesn't involve any regulatory aspect of metabolism, mainly because all the necessary regulatory or signalling components largely remain unknown. The components forming the regulatory network are DNA-binding proteins and their target promoters and for that of a signalling network are signalling receptors, protein kinases, and protein phosphatases [5]. Integration of metabolic network with transcriptional regulatory network has been done for lower organisms like *E.coli* strain iMC1010 [521], yeasts e.g., *S.cerevisiae* strain iMH805/775 [522]. Additionally, integration of metabolic and transcription and translation networks for E.coli strain iAF1260 [316, 523] has also been accomplished, as well as, various other algorithms developed to aid in this regard [524, 525]. This highlights the scope of systems biology methods to not only build component specific networks (i.e., metabolic and non-metabolic components of a cell), but also integrate these to unravel phenotypic features, network properties, biological discovery, and biotechnological applications [52]. However, such integrated networks for specific human cell types are yet to be constructed. Moreover, this shall also result in combined networks of enormous size, which shall depend heavily on intensive computational power and relatively faster computational algorithms to study the network properties. Such combinatorial approach holds potential to study complex disease pathologies and suggest patient specific medicare. Therefore, building of a combinatorial network blending all the basic cellular components, i.e., both the metabolic and non-metabolic components would be highly desirable for making highly meaningful predictions by the finished model.

One of the major concerns often raised for steady state modelling is that the classical COBRA modelling approach using flux balance analysis could be used for

determining metabolic flux distribution, but not for metabolite concentration and dynamic behaviour of the metabolic fluxes [11, 526], which can be studied using the dynamic flux balance analysis [526]. Apart from steady state modelling, the other most widely used modelling technique remain kinetic modelling. In general, kinetic modelling is not amenable for large genome scale networks, as most of the kinetic parameters remain unknown. However, there exists kinetic metabolic models of human red blood cell [527], human mesenchymal stem cell [528], as well as for Escherichia coli [529]. The choice of a specific modelling approach depends on the questions one tries to answer, e.g., mechanistic explanations for the effects of network perturbations is essentially done with steady state modelling, and analysis of dynamic behaviour of the network by kinetic modelling.

7.7 Possible areas for future research

As highlighted in the previous chapters 1 and 2, there exists a number of IEMs without any treatment procedures. Additionally, most of the well-known IEMs demand better diagnostic criteria. The work presented in this thesis holds promising potential for various future applications of systems biology methods in the field of IEMs, particularly for their diagnosis and treatment. The detailed discussion on transport proteins highlighted the role of transporters in metabolic reconstruction, and can be used to refine the reconstruction methodology of metabolic networks. Moreover, it also highlighted the IEMs caused due to faulty transport system, and this information could be handy while modelling such diseases. The compendium of IEMs can be expanded to include newly diagnosed ones. This shall not only help as a global catalogue to study IEMs, but also highlight the need to expand or refine the mother generic metabolic network in order to capture such IEMs. Subsequently, cell specific models can be generated from the mother networks to study the local effects of IEMs. One such attempt was modelling the metabolism of human small intestinal enterocytes. Although, we had two specific dietary patterns as model constraints, more elaborate and patient specific diets can be formulated and applied. Here in we also predicted certain novel metabolic biomarkers, as well as validated the use of current ones. Hence, this work can be expanded by building combined organ models in order to capture those biomarkers missed out due to single cell type modelling. We also provided a platform for modelling drug metabolism, with the Recon2-drug module. This can be used to predict better combination therapy of drugs, as well as synergistic effects of either diets or drugs. One of the basic aim to model commonly used drugs, was to come up with hypothesis if these easily available drugs be used as IEM medications. This can be applied to subsequent cell specific networks, in order to derive more meaningful results. Such targeted approaches will aim to deliver better diagnostic and individualized and patient centric therapeutic measures.

7.7.1 Perspective

The world of IEMs is ever expanding, necessitating the need for cataloguing and studying them, as well as finding easy and reliable identification strategy, and formulating better medications for them. Although, individually rare, significant proportion of these can be fatal, when left untreated. System biology aims to analyze a system or cell with a holistic approach, yet taking into account of its individual component and their complex interaction. Genome scale metabolic network reconstruction and COBRA modelling methods have proven to be useful in understanding these interactions, as well providing mechanistic explanation for the effects and causes of various perturbations caused to the network. Therefore, it is worthwhile to use these approaches to explore the world of IEMs. The work presented here aimed at understanding the precise pathologies of IEMs as well as predict the effect of these on healthy cellular metabolism, additionally, combining these with specific diet and drug administration. Hence, this work can be used as a starting point to hypothesize effective treatment strategies and novel biomarkers for IEMs.

Bibliography

1. Scriver, C. R. (2008). Garrod's Croonian Lectures (1908) and the charter 'Inborn Errors of Metabolism': albinism, alkaptonuria, cystinuria, and pentosuria at age 100 in 2008. Journal of Inherited Metabolic Disease *31*, 580–598.

2. Saudubray, J. M., Van den Berghe, G., and Walter, J. H. (2012). Inborn metabolic diseases: diagnosis and treatment. (Springerverlag Berlin Heidelberg).

3. Pampols, T. (2010). Inherited metabolic rare disease. Advances in Experimental Medicine and Biology *686*, 397–431.

4. McHugh, D. M., Cameron, C. A., Abdenur, J. E., Abdulrahman, M., Adair, O., Al Nuaimi, S. A., Ahlman, H., Allen, J. J., Antonozzi, I., Archer, S., et al. (2011). Clinical validation of cutoff target ranges in newborn screening of metabolic disorders by tandem mass spectrometry: A worldwide collaborative project. Genetics in Medicine *13*, 230–254.

5. Palsson, B. O. (2006). Systems Biology- Properties of Reconstructed Networks. (Cambridge University Press).

6. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J.-F., Dougherty, B. A., Merrick, J. M., et al. (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science *269*, 496–512.

7. Schmutz, J., Wheeler, J., Grimwood, J., Dickson, M., Yang, J., Caoile, C., Bajorek, E., Black, S., Chan, Y. M., Denys, M., et al. (2004). Quality assessment of the human genome sequence. Nature *429*, 365–368.

8. Lewis, N. E., Jamshidi, N., Thiele, I., and Palsson, B. O. (2009). Metabolic systems biology: a constraint-based approach. Encyclopedia of Complexity and Systems Science, Robert A Meyers (ed) pp. 5535.

9. Price, N. D., Reed, J. L., and Palsson, B. O. (2004). Genome-scale models of microbial cells: evaluating the consequences of constraints. Nature reviews Microbiology *2*, 886–897.

10. Thiele, I. and Palsson, B. O. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. Nature Protocols *5*, 93–121.

11. Orth, J. D., Thiele, I., and Palsson, B. O. (2010). What is flux balance analysis? Nature Biotechnology *28*, 245–248.

12. Mahadevan, R. and Schilling, C. H. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metabolic Engineering *5*, 264–276.

13. Schellenberger, J., Que, R., Fleming, R. M., Thiele, I., Orth, J. D., Feist, A. M., Zielinski, D. C., Bordbar, A., Lewis, N. E., Rahmanian, S., et al. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nature Protocols *6*, 1290–1307.

14. Ibarra, R. U., Edwards, J. S., and Palsson, B. O. (2002). Escherichia coli k-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. Nature *420*, 186–189.

15. Papin, J. A., Price, N. D., and Palsson, B. O. (2002). Extreme Pathway Lengths and Reaction Participation in Genome-Scale Metabolic Networks. Genome Research *12*, 1889–1900.

16. Gagneur, J. and Klamt, S. (2004). Computation of elementary modes: a unifying framework and the new binary approach. BMC Bioinformatics *5*, 175.

17. Price, N. D., Schellenberger, J., and Palsson, B. O. (2004). Uniform sampling of steady-state flux spaces: means to design experiments and to interpret enzymopathies. Biophysical Journal *87*, 2172–2186.

18. Duarte, N. C., Becker, S. A., Jamshidi, N., Thiele, I., Mo, M. L., Vo, T. D., Srivas, R., and Palsson, B. O. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proceedings of the National Academy of Sciences of the United States of America of the United States of America *104*, 1777–1782.

19. Shlomi, T., Cabili, M. N., and Ruppin, E. (2009). Predicting metabolic biomarkers of human inborn errors of metabolism. Molecular Systems Biology *5*, 263.

20. Lewis, N. E., Schramm, G., Bordbar, A., Schellenberger, J., Andersen, M. P., Cheng, J. K., Patel, N., Yee, A., Lewis, R. A., Eils, R., et al. (2010). Large-

scale in silico modeling of metabolic interactions between cell types in the human brain. Nature Biotechnology *28*, 1279–1285.

21. Gille, C., Bolling, C., Hoppe, A., Bulik, S., Hoffmann, S., Hubner, K., Karlstadt, A., Ganeshan, R., Konig, M., Rother, K., et al. (2010). HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. Molecular Systems Biology *6*, 411.

22. Chang, R. L., Xie, L., Bourne, P. E., and Palsson, B. O. (2010). Drug off-target effects predicted using structural analysis in the context of a metabolic network model. PLoS Computational Biology *6*, e1000938.

23. Bordbar, A., Jamshidi, N., and Palsson, B. O. (2011). iAB-RBC-283: A proteomically derived knowledge-base of erythrocyte metabolism that can be used to simulate its physiological and patho-physiological states. BMC Systems Biology *5*, 110.

24. Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B. O., and Jamshidi, N. (2010). Insight into human alveolar macrophage and M. tuberculosis interactions via metabolic reconstructions. Molecular Systems Biology *6*, 422.

25. Sigurdsson, M., Jamshidi, N., Steingrimsson, E., Thiele, I., and Palsson, B. O. (2010). A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1. BMC Systems Biology *4*, 140.

26. Bordbar, A., Feist, A. M., Usaite-Black, R., Woodcock, J., Palsson, B. O., and Famili, I. (2011). A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. BMC Systems Biology *5*, 180.

27. Folger, O., Jerby, L., Frezza, C., Gottlieb, E., Ruppin, E., and Shlomi, T. (2011). Predicting selective drug targets in cancer through metabolic networks. Molecular Systems Biology *7*, 501.

28. Zhao, Y. and Huang, J. (2011). Reconstruction and analysis of human heart-specific metabolic network based on transcriptome and proteome data. Biochemical and Biophysical Research Communications *415*, 450–454.

29. Sahoo, S., Franzson, L., Jonsson, J. J., and Thiele, I. (2012). A compendium of inborn errors of metabolism mapped onto the human metabolic network. Molecular BioSystems.

30. Sahoo, S. and Thiele, I. (2013). Predicting the impact of diet and enzymopathies on human small intestinal epithelial cells. Human Molecular Ge-

netics.

31. Thiele, I., Swainston, N., Fleming, R. M. T., Hoppe, A., Sahoo, S., Aurich, M. K., Haraldsdottir, H., Mo, M. L., Rolfsson, O., Stobbe, M. D., et al. (2013). A community-driven global reconstruction of human metabolism. Nature Biotechnology *31*, 419–425.

32. Ma, H., Sorokin, A., Mazein, A., Selkov, A., Selkov, E., Demin, O., and Goryanin, I. (2007). The Edinburgh human metabolic network reconstruction and its functional analysis. Molecular Systems Biology *3*, 135.

33. Hao, T., Ma, H. W., Zhao, X. M., and Goryanin, I. (2010). Compartmentalization of the Edinburgh Human Metabolic Network. BMC Bioinformatics *11*, 393.

34. Scheer, M., Grote, A., Chang, A., Schomburg, I., Munaretto, C., Rother, M., Sohngen, C., Stelzer, M., Thiele, J., and Schomburg, D. (2011). BRENDA, the enzyme information system in 2011. Nucleic Acids Research *39*, D670–D676.

35. Okuda, S., Yamada, T., Hamajima, M., Itoh, M., Katayama, T., Bork, P., Goto, S., and Kanehisa, M. (2008). KEGG Atlas mapping for global analysis of metabolic pathways. Nucleic Acids Research *36*, W423–6.

36. Thorleifsson, S. G. and Thiele, I. (2011). rBioNet: A COBRA toolbox extension for reconstructing high-quality biochemical networks. Bioinformatics *27*, 2009–2010.

37. Maglott, D., Ostell, J., Pruitt, K. D., and Tatusova, T. (2011). Entrez Gene: gene-centered information at NCBI. Nucleic Acids Research *39*, D52–D57.

38. Mithieux, G. (2001). New data and concepts on glutamine and glucose metabolism in the gut. Current Opinion in Clinical Nutrition & Metabolic Care *4*, 267–271.

39. Dietschy, J. M. and Siperstein, M. D. (1965). Cholesterol synthesis by the gastrointestinal tract: localization and mechanisms of control. The Journal of Clinical Investigation *44*, 1311–1327.

40. Hartmann, F. and Bissell, D. M. (1982). Metabolism of heme and bilirubin in rat and human small intestinal mucosa. The Journal of Clinical Investigation *70*, 23–29.

41. Miguel-Aliaga, I. (2012). Nerveless and gutsy: intestinal nutrient sensing from invertebrates to humans. Seminars in Cell & Developmental Biology.

42. Henry, C. S., DeJongh, M., Best, A. A., Frybarger, P. M., Linsay, B., and Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. Nature Biotechnology *28*, 977–982.

43. Wortmann, S. B., Vaz, F. M., Gardeitchik, T., Vissers, L. E. L. M., Renkema, G. H., Schuurs-Hoeijmakers, J. H. M., Kulik, W., Lammens, M., Christin, C., Kluijtmans, L. A. J., et al. (2012). Mutations in the phospholipid remodeling gene SERAC1 impair mitochondrial function and intracellular cholesterol trafficking and cause dystonia and deafness. Nature Genetics *44*, 797–802.

44. Stojkovic, T., Vissing, J., Petit, F., Piraud, M., Orngreen, M. C., Andersen, G., Claeys, K. G., Wary, C., Hogrel, J. Y., and Laforet, P. (2009). Muscle glycogenosis due to phosphoglucomutase 1 deficiency. New England Journal of Medicine *361*, 425–427.

45. Ebberink, M. S., Koster, J., Visser, G., van Spronsen, F., Stolte-Dijkstra, I., Smit, G. P. A., Fock, J. M., Kemp, S., Wanders, R. J. A., and Waterham, H. R. (2012). A novel defect of peroxisome division due to a homozygous non-sense mutation in the PEX11beta gene. Journal of Medical Genetics *49*, 307–313.

46. Beck, B. B., Baasner, A., Buescher, A., Habbig, S., Reintjes, N., Kemper, M. J., Sikora, P., Mache, C., Pohl, M., Stahl, M., et al. (2013). Novel findings in patients with primary hyperoxaluria type III and implications for advanced molecular testing strategies. European Journal of Human Genetics *21*, 162–172.

47. Durot, M., Le Fevre, F., de Berardinis, V., Kreimeyer, A., Vallenet, D., Combe, C., Smidtas, S., Salanoubat, M., Weissenbach, J., and Schachter, V. (2008). Iterative reconstruction of a global metabolic model of Acinetobacter baylyi ADP1 using high-throughput growth phenotype and gene essentiality data. BMC Systems Biology *2*, 85.

48. Oh, Y.-K., Palsson, B. O., Park, S. M., Schilling, C. H., and Mahadevan, R. (2007). Genome-scale reconstruction of metabolic network in Bacillus subtilis based on high-throughput phenotyping and gene essentiality data. Journal of Biological Chemistry *282*, 28791–28799.

49. Balagurunathan, B., Jonnalagadda, S., Tan, L., and Srinivasan, R. (2012). Reconstruction and analysis of a genome-scale metabolic model for scheffersomyces stipitis. Microbial Cell Factories *11*, 27.

50. Feist, A. M., Scholten, J. C. M., Palsson, B. O., Brockman, F. J., and Ideker, T. (2006). Modeling methanogenesis with a genome-scale metabolic reconstruction of Methanosarcina barkeri. Molecular Systems Biology *2*.

51. Heavner, B. D., Smallbone, K., Barker, B., Mendes, P., and Walker, L. P. (2012). Yeast 5–an expanded reconstruction of the Saccharomyces cerevisiae metabolic network. BMC Systems Biology *6*, 55.

52. McCloskey, D., Palsson, B. O., and Feist, A. M. (2013). Basic and applied uses of genome-scale metabolic network reconstructions of Escherichia coli. Molecular Systems Biology *9*.

53. Heinken, A., Sahoo, S., Fleming, R. M., and Thiele, I. (2013). Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. Gut Microbes *4*, 28–40.

54. Karlstaedt, A., Fliegner, D., Kararigas, G., Ruderisch, H. S., Regitz-Zagrosek, V., and Holzhutter, H. G. (2012). CardioNet: A human metabolic network suited for the study of cardiomyocyte metabolism. BMC Systems Biology *6*, 114.

55. Jerby, L., Shlomi, T., and Ruppin, E. (2010). Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. Molecular Systems Biology *6*, 401.

56. Agren, R., Bordel, S., Mardinoglu, A., Pornputtapong, N., Nookaew, I., and Nielsen, J. (2012). Reconstruction of Genome-Scale Active Metabolic Networks for 69 Human Cell Types and 16 Cancer Types Using INIT. PLoS Computational Biology *8*, e1002518.

57. Jerby, L., Wolf, L., Denkert, C., Stein, G. Y., Hilvo, M., Oresic, M., Geiger, T., and Ruppin, E. (2012). Metabolic associations of reduced proliferation and oxidative stress in advanced breast cancer. Cancer Research.

58. Joyce, A. R. and Palsson, B. O. (2006). The model organism as a system: integrating'omics' data sets. Nature Reviews Molecular Cell Biology *7*, 198–210.

59. Martin, L., Anguita, A., Maojo, V., and Crespo, J. (2010). Integration of omics data for Cancer Researchearch. In An Omics Perspective on Cancer Researchearch In An Omics Perspective on Cancer Researchearch. (Springer).

60. Kluger, Y., Yu, H., Qian, J., and Gerstein, M. (2003). Relationship between

gene co-expression and probe localization on microarray slides. BMC Genomics *4*, 49.

61. Shields, R. (2006). MIAME, we have a problem. Trends in Genetics *22*, 65–66.

62. Nesvizhskii, A. I. and Aebersold, R. (2005). Interpretation of shotgun proteomic data the protein inference problem. Molecular & Cellular Proteomics *4*, 1419–1440.

63. Booth, S. C., Weljie, A., and Turner, R. J. (2013). Computational Tools for the Secondary Analysis of Metabolomics Experiments. Computational and Structural Biotechnology Journal *4*.

64. Feist, A. M., Herrgard, M. J., Thiele, I., Reed, J. L., and Palsson, B. O. (2009). Reconstruction of biochemical networks in microorganisms. Nature Reviews Microbiology *7*, 129–143.

65. Mardinoglu, A., Agren, R., Kampf, C., Asplund, A., Nookaew, I., Jacobson, P., Walley, A. J., Froguel, P., Carlsson, L. M., Uhlen, M., et al. (2013). Integration of clinical data with a genome-scale metabolic model of the human adipocyte. Molecular Systems Biology *9*.

66. Krauss, M., Schaller, S., Borchers, S., Findeisen, R., Lippert, J., and Kuepfer, L. (2012). Integrating cellular metabolism into a multiscale whole-body model. PLoS Computational Biology *8*, e1002750.

67. Bazzani, S., Hoppe, A., and Holzhutter, H.-G. (2012). Network-based assessment of the selectivity of metabolic drug targets in Plasmodium falciparum with respect to human liver metabolism. BMC Systems Biology *6*, 118.

68. Brunton, L., Lazo, J., and Parker, K. (2006). Goodman and Gilman's : The pharmacological basis of therapeutics volume 1157. (The McGraw-Hill companies.).

69. Camacho, J. and Rioseco-Camacho, N. (1993). Hyperornithinemia-Hyperammonemia-Homocitrullinuria Syndrome.

70. Kobayashi, K., Saheki, T., and Song, Y. Z. (1993). Citrin Deficiency.

71. Seow, H. F., Broer, S., Broer, A., Bailey, C. G., Potter, S. J., Cavanaugh, J. A., and Rasko, J. E. (2004). Hartnup disorder is caused by mutations in the gene encoding the neutral amino acid transporter SLC6A19. Nature Genetics *36*,

1003–1007.

72. Prudente, S., Flex, E., Morini, E., Turchi, F., Capponi, D., De Cosmo, S., Tassi, V., Guida, V., Avogaro, A., Folli, F., et al. (2007). A functional variant of the adipocyte glycerol channel aquaporin 7 gene is associated with obesity and related metabolic abnormalities. Diabetes *56*, 1468–1474.

73. Cooper, R., Sarioglu, S., Sokmen, S., Fuzun, M., Kupelioglu, A., Valentine, H., Gorken, I. B., Airley, R., and West, C. (2003). Glucose transporter-1 (GLUT-1): a potential marker of prognosis in rectal carcinoma? British Journal of Cancer *89*, 870–876.

74. Macheda, M. L., Rogers, S., and Best, J. D. (2005). Molecular and cellular regulation of glucose transporter (GLUT) proteins in cancer. Journal of Cellular Physiology *202*, 654–662.

75. Guyton, A. C. and Hall, J. E. (2000). Textbook of Medical Physiology. (W.B.Saunders company).

76. Lodish, H., Berk, A., and Zipursky, S. L. (2000). Molecular Cell Biology : Uniporter-Catalyzed Transport. (New York: W. H. Freeman).

77. Alberts, B., Johnson, A., and Lewis, J. (2002). Molecular Biology of the Cell. (New York: Garland Science).

78. Forrest, L. R., Kramer, R., and Ziegler, C. (2011). The structural basis of secondary active transport mechanisms. Biochimica et Biophysica Acta (BBA) - Bioenergetics *1807*, 167–188.

79. Baird, F. E., Bett, K. J., MacLean, C., Tee, A. R., Hundal, H. S., and Taylor, P. M. (2009). Tertiary active transport of amino acids reconstituted by coexpression of System A and L transporters in Xenopus oocytes. American Journal of Physiology - Endocrinology and Metabolism *297*, E822–9.

80. Wu, X. and Freeze, H. H. (2002). GLUT14, a duplicon of GLUT3, is specifically expressed in testis as alternative splice forms. Genomics *80*, 553–557.

81. Wood, I. S. and Trayhurn, P. (2003). Glucose transporters (GLUT and SGLT): expanded families of sugar transport proteins. British Journal of Nutrition *89*, 3–9.

82. Augustin, R. (2010). The protein family of glucose transport facilitators: It's not only about glucose after all. IUBMB Life *62*, 315–333.

83. Gropper, S. S., Smith, J. L., and Groff, J. L. (2009). Advanced nutrition and human metabolism. (Australia; United States: Wadsworth/Cengage Learning).

84. Hediger, M. A., Kanai, Y., You, G., and Nussberger, S. (1995). Mammalian ion-coupled solute transporters. The Journal of Physiology *482*, 7S–17S.

85. Thwaites, D. T. and Anderson, C. M. H. (2007). H+-coupled nutrient, micronutrient and drug transporters in the mammalian small intestine. Experimental Physiology *92*, 603–619.

86. Hummel, C. S., Lu, C., Loo, D. D., Hirayama, B. A., Voss, A. A., and Wright, E. M. (2011). Glucose transport by human renal Na+/D-glucose cotransporters SGLT1 and SGLT2. American Journal of Physiology - Cell Physiology *300*, C14–C21.

87. Diez-Sampedro, A., Hirayama, B. A., Osswald, C., Gorboulev, V., Baumgarten, K., Volk, C., Wright, E. M., and Koepsell, H. (2003). A glucose sensor hiding in a family of transporters. Proceedings of the National Academy of Sciences of the United States of America of the United States of America *100*, 11753–11758.

88. Bianchi, L. and Diez-Sampedro, A. (2010). A single amino acid change converts the sugar sensor SGLT3 into a sugar transporter. PLoS ONE *5*, e10241.

89. Kothinti, R. K., Blodgett, A. B., North, P. E., Roman, R. J., and Tabatabai, N. M. (2012). A novel SGLT is expressed in the human kidney. European Journal of Pharmacology.

90. Tazawa, S., Yamato, T., Fujikura, H., Hiratochi, M., Itoh, F., Tomae, M., Takemura, Y., Maruyama, H., Sugiyama, T., Wakamatsu, A., et al. (2005). SLC5A9/SGLT4, a new Na+-dependent glucose transporter, is an essential transporter for mannose, 1,5-anhydro-D-glucitol, and fructose. Life Sciences *76*, 1039–1050.

91. Grempler, R., Augustin, R., Froehner, S., Hildebrandt, T., Simon, E., Mark, M., and Eickelmann, P. (2012). Functional characterisation of human SGLT-5 as a novel kidney-specific sodium-dependent sugar transporter. FEBS Lettersers *586*, 248–253.

92. Chen, J., Williams, S., Ho, S., Loraine, H., Hagan, D., Whaley, J. M., and Feder, J. N. (2010). Quantitative PCR tissue expression profiling of the human SGLT2 gene and related family members. Diabetes Therapy *1*, 57–92.

93. Wright, E. M., Loo, D. D., and Hirayama, B. A. (2011). Biology of human sodium glucose transporters. Physiological Reviews *91*, 733–794.

94. Joost, H. G. and Thorens, B. (2001). The extended GLUT-family of sugar/polyol transport facilitators: nomenclature, sequence characteristics, and potential function of its novel members (review). Molecular Membrane Biology *18*, 247–256.

95. Scheepers, A., Joost, H. G., and Schurmann, A. (2004). The glucose transporter families SGLT and GLUT: molecular basis of normal and aberrant function. Journal of Parenteral and Enteral Nutrition *28*, 364–371.

96. Wilson-O'Brien, A. L., Patron, N., and Rogers, S. (2010). Evolutionary ancestry and novel functions of the mammalian glucose transporter (GLUT) family. BMC Evolutionary Biology *10*, 152.

97. Doblado, M. and Moley, K. H. (2009). Facilitative glucose transporter 9, a unique hexose and urate transporter. American Journal of Physiology - Endocrinology and Metabolism *297*, E831–E835.

98. Cura, A. J. and Carruthers, A. (2012). Role of monosaccharide transport proteins in carbohydrate assimilation, Distribution, metabolism, and homeostasis. Comprehensive Physiology.

99. Corpe, C. P., Eck, P., Wang, J., Al-Hasani, H., and Levine, M. (2013). Intestinal Dehydroascorbic acid (DHA) transport mediated by the facilitative sugar transporters, GLUT2 and GLUT8. Journal of Biological Chemistry *288*, 9092–9101.

100. Vannucci, S. J., Clark, R. R., Koehler-Stec, E., Li, K., Smith, C. B., Davies, P., Maher, F., and Simpson, I. A. (1998). Glucose transporter expression in brain: relationship to cerebral glucose utilization. Developmental Neuroscience *20*, 369–379.

101. Bakirtzi, K., Belfort, G., Lopez-Coviella, I., Kuruppu, D., Cao, L., Abel, E. D., Brownell, A. L., and Kandror, K. V. (2009). Cerebellar neurons possess a vesicular compartment structurally and functionally similar to Glut4-storage vesicles from peripheral insulin-sensitive tissues. The Journal of Neuroscience *29*, 5193–5201.

102. Broer, S. and Palacin, M. (2011). The role of amino acid transporters in inherited and acquired diseases. Biochemical Journal *436*, 193–211.

103. Palacin, M., Nunes, V., Font-Llitjos, M., Jimenez-Vidal, M., Fort, J., Gasol, E., Pineda, M., Feliubadalo, L., Chillaron, J., and Zorzano, A. (2005). The genetics of heteromeric amino acid transporters. Physiology (Bethesda) *20*, 112–124.

104. Fukasawa, Y., Segawa, H., Kim, J. Y., Chairoungdua, A., Kim, D. K., Matsuo, H., Cha, S. H., Endou, H., and Kanai, Y. (2000). Identification and characterization of a Na(+)-independent neutral amino acid transporter that associates with the 4F2 heavy chaiN. and exhibits substrate selectivity for small neutral D- and L-amino acids. The Journal of Biological Chemistry *275*, 9690–9698.

105. Nakauchi, J., Matsuo, H., Kim, D. K., Goto, A., Chairoungdua, A., Cha, S. H., Inatomi, J., Shiokawa, Y., Yamaguchi, K., Saito, I., et al. (2000). Cloning and characterization of a human brain Na(+)-independent transporter for small neutral amino acids that transports D-serine with high affinity. Neuroscience Letters *287*, 231–235.

106. Closs, E. I., Boissel, J. P., Habermeier, A., and Rotmann, A. (2006). Structure and function of cationic amino acid transporters (CATs). The Journal of Membrane Biology *213*, 67–77.

107. Hoshide, R., Ikeda, Y., Karashima, S., Matsuura, T., Komaki, S., Kishino, T., Niikawa, N., Endo, F., and Matsuda, I. (1996). Molecular cloning, tissue distribution, and chromosomal localization of human cationic amino acid transporter 2 (HCAT2). Genomics *38*, 174–178.

108. Vekony, N., Wolf, S., Boissel, J. P., Gnauert, K., and Closs, E. I. (2001). Human cationic amino acid transporter hCAT-3 is preferentially expressed in peripheral tissues. Biochemistry *40*, 12387–12394.

109. Wolf, S., Janzen, A., Vekony, N., Martine, U., Strand, D., and Closs, E. I. (2002). Expression of solute carrier 7A4 (SLC7A4) in the plasma membrane is not sufficient to mediate amino acid transport activity. Biochemical Journal *364*, 767–775.

110. Adibi, S. A. (2003). Regulation of expression of the intestinal oligopeptide transporter (Pept-1) in health and disease. American Journal of Physiology - Gastrointestinal and Liver Physiology *285*, G779–88.

111. Shu, C., Shen, H., Hopfer, U., and Smith, D. E. (2001). Mechanism of intestinal absorption and renal reabsorption of an orally active ace inhibitor: uptake and transport of fosinopril in cell cultures. Drug Metabolism and Disposition *29*, 1307–1315.

112. Biegel, A., Knutter, I., Hartrodt, B., Gebauer, S., Theis, S., Luckner, P., Kottra, G., Rastetter, M., Zebisch, K., Thondorf, I., et al. (2006). The renal type H+/peptide symporter PEPT2: structure-affinity relationships. Amino Acids *31*, 137–156.

113. Agu, R., Cowley, E., Shao, D., Macdonald, C., Kirkpatrick, D., Renton, K., and Massoud, E. (2011). Proton-coupled oligopeptide transporter (POT) family expression in human nasal epithelium and their drug transport potential. Molecular Pharmaceutics *8*, 664–672.

114. Leibach, F. H. and Ganapathy, V. (1996). Peptide transporters in the intestine and the kidney. Annual Review of Nutrition *16*, 99–119.

115. Pieri, M., Christian, H. C., Wilkins, R. J., Boyd, C. A., and Meredith, D. (2010). The apical (hPepT1) and basolateral peptide transport systems of Caco-2 cells are regulated by AMP-activated protein kinase. American Journal of Physiology - Gastrointestinal and Liver Physiology *299*, G136–G143.

116. Terada, T., Sawada, K., Saito, H., Hashimoto, Y., and Inui, K. (1999). Functional characteristics of basolateral peptide transporter in the human intestinal cell line Caco-2. The American Journal of Physiology *276*, G1435–G1441.

117. Thwaites, D. T., Brown, C. D., Hirst, B. H., and Simmons, N. L. (1993). Transepithelial glycylsarcosine transport in intestinal Caco-2 cells mediated by expression of H(+)-coupled carriers at both apical and basal membranes. The Journal of Biological Chemistry *268*, 7640–7642.

118. Charrier, L. and Merlin, D. (2006). The oligopeptide transporter hPepT1: gateway to the innate immune response. Laboratory Investigation *86*, 538–546.

119. Broer, S. (2008). Amino acid transport across mammalian intestinal and renal epithelia. Physiological Reviews *88*, 249–286.

120. Murray, R. K., Bender, D. A., Botham, K. M., Kennelly, P. J., Rodwell, V. W., and Weil, P. A. (2009). A lange medical book: Harper's illustrated Biochemistry. (Mc Graw Hill Medical).

121. Gossett, R. E., Frolov, A. A., Roths, J. B., Behnke, W. D., Kier, A. B., and Schroeder, F. (1996). Acyl-CoA binding proteins: multiplicity and function. Lipids *31*, 895–918.

122. Furuhashi, M. and Hotamisligil, G. S. (2008). Fatty acid-binding proteins:

role in metabolic diseases and potential as drug targets. Nature Reviews Drug Discovery *7*, 489–503.

123. Devaux, P. F., Herrmann, A., Ohlwein, N., and Kozlov, M. M. (2008). How lipid flippases can modulate membrane structure. Biochimica et Biophysica Acta (BBA) *1778*, 1591–1600.

124. Sanyal, S. and Menon, A. K. (2009). Flipping lipids: why an' what's the reason for? ACS Chemical Biology *4*, 895–909.

125. Jia, Z., Pei, Z., Maiguel, D., Toomer, C. J., and Watkins, P. A. (2007). The fatty acid transport protein (FATP) family: very long chain acyl-CoA synthetases or solute carriers? Journal of Molecular Neuroscience *33*, 25–31.

126. Stahl, A. (2004). A current review of fatty acid transport proteins (SLC27). Pflugers Archiv - European Journal of Physiology *447*, 722–727.

127. Glatz, J. F., Luiken, J. J., and Bonen, A. (2010). Membrane Fatty Acid Transporters as Regulators of Lipid Metabolism: Implications for Metabolic Disease. Physiological Reviews *90*, 367–417.

128. Dutta-Roy, A. K. (2000). Cellular uptake of long-chain fatty acids: role of membrane-associated fatty-acid-binding/transport proteins. Cellular and Molecular Life Sciences *57*, 1360–1372.

129. Stewart, J. M. (2000). The cytoplasmic fatty-acid-binding proteins: thirty years and counting. Cellular and Molecular Life Sciences *57*, 1345–1359.

130. Stahl, A., Gimeno, R. E., Tartaglia, L. A., and Lodish, H. F. (2001). Fatty acid transport proteins: a current view of a growing family. Trends in Endocrinology and Metabolism *12*, 266–273.

131. Febbraio, M. and Silverstein, R. L. (2007). CD36: implications in cardiovascular disease. The International Journal of Biochemistry & Cell Biology *39*, 2012–2030.

132. Nickerson, J. G., Alkhateeb, H., Benton, C. R., Lally, J., Nickerson, J., Han, X. X., Wilson, M. H., Jain, S. S., Snook, L. A., Glatz, J. F., et al. (2009). Greater transport efficiencies of the membrane fatty acid transporters FAT/CD36 and FATP4 compared with FABPpm and FATP1 and differential effects on fatty acid esterification and oxidation in rat skeletal muscle. The Journal of Biological Chemistry *284*, 16522–16530.

133. Chabowski, A., Gorski, J., Luiken, J. J., Glatz, J. F., and Bonen, A. (2007). Evidence for concerted action of FAT/CD36 and FABPpm to increase fatty acid transport across the plasma membrane. Prostaglandins, Leukotrienes and Essential Fatty Acids *77*, 345–353.

134. Glavinas, H., Krajcsi, P., Cserepes, J., and Sarkadi, B. (2004). The role of ABC transporters in drug resistance, metabolism and toxicity. Current Drug Delivery *1*, 27–42.

135. Orso, E., Broccardo, C., Kaminski, W. E., Bottcher, A., Liebisch, G., Drobnik, W., Gotz, A., Chambenoit, O., Diederich, W., Langmann, T., et al. (2000). Transport of lipids from golgi to plasma membrane is defective in tangier disease patients and Abc1-deficient mice. Nature Genetics *24*, 192–196.

136. Kiss, K., Brozik, A., Kucsma, N., Toth, A., Gera, M., Berry, L., Vallentin, A., Vial, H., Vidal, M., and Szakacs, G. (2012). Shifting the paradigm: the putative mitochondrial protein ABCB6 resides in the lysosomes of cells and in the plasma membrane of erythrocytes. PLoS ONE *7*, e37378.

137. Nelson, D. L. and Cox, M. M. (2000). Lehninger principles of Biochemsitry. (Worth publication) 3rd editio edition.

138. Reboul, E. and Borel, P. (2011). Proteins involved in uptake, intracellular transport and basolateral secretion of fat-soluble vitamins and carotenoids by mammalian enterocytes. Progress in Lipid Research *50*, 388–402.

139. Ikonen, E. (2008). Cellular cholesterol trafficking and compartmentalization. Nature Reviews Molecular and Cellular Biology *9*, 125–138.

140. Fustin, J. M., Doi, M., Yamada, H., Komatsu, R., Shimba, S., and Okamura, H. (2012). Rhythmic nucleotide synthesis in the liver: temporal segregation of metabolites. Cell Reports *1*, 341–349.

141. Marce, S., Molina-Arcas, M., Villamor, N., Casado, F. J., Campo, E., Pastor-Anglada, M., and Colomer, D. (2006). Expression of human equilibrative nucleoside transporter 1 (hENT1) and its correlation with gemcitabine uptake and cytotoxicity in mantle cell lymphoma. Haematologica *91*, 895–902.

142. Toyoda, Y. and Ishikawa, T. (2010). Pharmacogenomics of human ABC transporter ABCC11 (MRP8): potential risk of breast cancer and chemotherapy failure. Anti-Cancer Agents in Medicinal Chemistry *10*, 617–624.

143. Fukuda, Y. and Schuetz, J. D. (2012). ABC transporters and their role in

nucleoside and nucleotide drug resistance. Biochemical Pharmacology *83*, 1073–1083.

144. Pastor-Anglada, M., Errasti-Murugarren, E., Aymerich, I., and Casado, F. J. (2007). Concentrative nucleoside transporters (CNTs) in epithelia: from absorption to Cellular Signallinging. Journal of Physiology and Biochemistry *63*, 97–110.

145. Molina-Arcas, M., Trigueros-Motos, L., Casado, F. J., and Pastor-Anglada, M. (2008). Physiological and pharmacological roles of nucleoside transporter proteins. Nucleosides Nucleotides Nucleic Acids *27*, 769–778.

146. Baldwin, S. A., Beal, P. R., Yao, S. Y., King, A. E., Cass, C. E., and Young, J. D. (2004). The equilibrative nucleoside transporter family, SLC29. Pflugers Archiv - European Journal of Physiology *447*, 735–743.

147. Klaassen, C. D. and Aleksunes, L. M. (2010). Xenobiotic, bile acid, and cholesterol transporters: function and regulation. Pharmacological Reviews *62*, 1–96.

148. Hollander, D., Rim, E., and Muralidhara, K. S. (1977). Vitamin K1 intestinal absorption in vivo: influence of luminal contents on transport. American Journal of Physiology - Endocrinology and Metabolism *232*, E69.

149. Berry, D. C., Jin, H., Majumdar, A., and Noy, N. (2011). Signaling by vitamin A and retinol-binding protein regulates gene expression to inhibit insulin responses. Proceedings of the National Academy of Sciences of the United States of America *108*, 4340–4345.

150. Burke, T. R. and Tsang, S. H. (2011). Allelic and phenotypic heterogeneity in ABCA4 mutations. Ophthalmic Genetics *32*, 165–174.

151. Harrison, E. H. (2005). Mechanisms of digestion and absorption of dietary vitamin A. Annual Review of Nutrition *25*, 87–103.

152. Ball, G. F. M. (2006). Vitamins : Their role in the human body. (Blackwell publishing).

153. Dusso, A. S., Brown, A. J., and Slatopolsky, E. (2005). Vitamin D. American Journal of Physiology *289*, F8–F28.

154. Reboul, E., Goncalves, A., Comera, C., Bott, R., Nowicki, M., Landrier, J. F., Jourdheuil-Rahmani, D., Dufour, C., Collet, X., and Borel, P. (2011). Vitamin

D intestinal absorption is not a simple passive diffusion: evidences for involvement of cholesterol transporters. Molecular Nutrition & Food Research *55*, 691–702.

155. Reboul, E., Klein, A., Bietrix, F., Gleize, B., Malezet-Desmoulins, C., Schneider, M., Margotat, A., Lagrost, L., Collet, X., and Borel, P. (2006). Scavenger receptor class B type I (SR-BI) is involved in vitamin E transport across the enterocyte. Journal of Biological Chemistry *281*, 4739–4745.

156. Narushima, K., Takada, T., Yamanashi, Y., and Suzuki, H. (2008). Niemann-Pick C1-like 1 mediates alpha-tocopherol transport. Molecular Pharmacology *74*, 42–49.

157. Rigotti, A. (2007). Absorption, transport, and tissue delivery of vitamin E. Molecular Aspects of Medicine *28*, 423–436.

158. Zempleni, J., Rucker, R. B., Mccormick, D. B., and Suttie, J. W. (2007). Handbook of vitamins. (CRC press, Taylor & Francis Group).

159. Said, H. M. (2011). Intestinal absorption of water-soluble vitamins in health and disease. Biochemical Journal *437*, 357–372.

160. Said, H. M. (2004). Recent advances in carrier-mediated intestinal absorption of water-soluble vitamins. Annual Review of Physiology *66*, 419–446.

161. Said, H. M., Hollander, D., and Mohammadkhani, R. (1993). Uptake of riboflavin by intestinal basolateral membrane vesicles: a specialized carrier-mediated process. Biochimica et Biophysica Acta (BBA) - Biomembranes *1148*, 263–268.

162. Bates, C. J. (1997). Bioavailability of riboflavin. European Journal of Clinical Nutrition *51*, S38.

163. Subramanian, V. S., Subramanya, S. B., Rapp, L., Marchant, J. S., Ma, T. Y., and Said, H. M. (2011). Differential expression of human riboflavin transporters -1, -2, and -3 in polarized epithelia: a key role for hRFT-2 in intestinal riboflavin uptake. Biochimica et Biophysica Acta (BBA) *1808*, 3016–3021.

164. Yao, Y., Yonezawa, A., Yoshimatsu, H., Masuda, S., Katsura, T., and Inui, K.-i. (2010). Identification and comparative functional characterization of a new human riboflavin transporter hRFT3 expressed in the brain. The Journal of Nutrition *140*, 1220–1226.

165. Nabokina, S. M., Kashyap, M. L., and Said, H. M. (2005). Mechanism and regulation of human intestinal niacin uptake. American Journal of Physiology - Cell Physiology *289*, C97–C103.

166. Gopal, S., E. ND Miyauchi, Martin, P. M., Ananth, S., Roon, P., Smith, S. B., and Ganapathy, V. (2007). Transport of nicotinate and structurally related compounds by human SMCT1 (SLC5A8) and its relevance to drug transport in the mammalian intestinal tract. Pharmaceutical Research *24*, 575–584.

167. Said, H. M., Ortiz, A., and Ma, T. Y. (2003). A carrier-mediated mechanism for pyridoxine uptake by human intestinal epithelial Caco-2 cells: regulation by a PKA-mediated pathway. American Journal of Physiology - Cell Physiology *285*, C1219–C1225.

168. Tolarova, M. M. (1982). Periconceptional supplementation with vitamins and folic acid to prevent recurrence of cleft lip. The The Lancet *320*, 217.

169. Czeizel, A. E. and Dudas, I. (1992). Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. New England Journal of Medicine *327*, 1832–1835.

170. Czeizel, A. E. (1996). Reduction of urinary tract and cardiovascular defects by periconceptional multivitamin supplementation. American Journal of Medical Genetics *62*, 179–183.

171. Tian, R. and Ingwall, J. S. (2008). How does folic acid cure heart attacks? Circulation *117*, 1772–1774.

172. Zhao, R., Diop-Bove, N., Visentin, M., and Goldman, I. D. (2011). Mechanisms of membrane transport of folates into cells and across epithelia. Annual Review of Nutrition *31*, 177–201.

173. Hamid, A., Kiran, M., Rana, S., and Kaur, J. (2009). Low folate transport across intestinal basolateral surface is associated with down-regulation of reduced folate carrier in in vivo model of folate malabsorption. IUBMB life *61*, 236–243.

174. Matherly, L. H. and Goldman, D. I. (2003). Membrane transport of folates. Vitamins & Hormones *66*, 403–456.

175. Zhao, R. and Goldman, I. D. (2003). Resistance to antifolates. Oncogene *22*, 7431–7457.

176. Takeda, M., Khamdang, S., Narikawa, S., Kimura, H., Hosoyamada, M., Cha, S. H., Sekine, T., and Endou, H. (2002). Characterization of methotrexate transport and its drug interactions with human organic anion transporters. Journal of Pharmacology and Experimental Therapeutics *302*, 666–671.

177. Badagnani, I., Castro, R. A., Taylor, T. R., Brett, C. M., Huang, C. C., Stryke, D., Kawamoto, M., Johns, S. J., Ferrin, T. E., Carlson, E. J., et al. (2006). Interaction of methotrexate with organic-anion transporting polypeptide 1A2 and its genetic variants. Journal of Pharmacology and Experimental Therapeutics *318*, 521–529.

178. Watkins, D. and Rosenblatt, D. S. (2011). Inborn errors of cobalamin absorption and metabolism. In American Journal of Medical Genetics Part C: Seminars in Medical Genetics volume 157 Wiley Online Library pp. 33–44.

179. Quadros, E. V. (2010). Advances in the understanding of cobalamin assimilation and metabolism. British Journal of Haematology *148*, 195–204.

180. Green, R. (2010). Ins and outs of cellular cobalamin transport. Blood *115*, 1476–1477.

181. Said, H. M. (1999). Cellular uptake of biotin: mechanisms and regulation. The Journal of Nutrition *129*, 490S–493S.

182. Boyer, J. C., Campbell, C. E., Sigurdson, W. J., and Kuo, S. M. (2005). Polarized localization of vitamin C transporters, SVCT1 and SVCT2, in epithelial cells. Biochemical and Biophysical Research Communications *334*, 150–156.

183. Liang, W. J., Johnson, D., and Jarvis, S. M. (2001). Vitamin C transport systems of mammalian cells. Molecular Membrane Biology *18*, 87–95.

184. Wilson, J. X. (2005). Regulation of vitamin C transport. Annual Review of Nutrition *25*, 105–125.

185. Hediger, M. A. (2002). New view at C. Nature Medicine *5*, 445–446.

186. MacAulay, N., Hamann, S., and Zeuthen, T. (2004). Water transport in the brain: role of cotransporters. Neuroscience *129*, 1031–1044.

187. Vandenberg, R. J., Handford, C. A., Campbell, E. M., Ryan, R. M., and Yool, A. J. (2011). Water and urea permeation pathways of the human excitatory amino acid transporter EAAT1. Biochemical Journal *439*, 333–340.

188. Zeuthen, T., Meinild, A. K., Loo, D. D., Wright, E. M., and Klaerke, D. A. (2001). Isotonic transport by the Na+-glucose cotransporter SGLT1 from humans and rabbit. The Journal of Physiology *531*, 631–644.

189. Zardoya, R. and Villalba, S. (2001). A phylogenetic framework for the aquaporin family in eukaryotes. Journal of Molecular Evolution *52*, 391–404.

190. Verkman, A. S. (2005). More than just water channels: unexpected cellular roles of aquaporins. Journal of Cell Science *118*, 3225–3232.

191. Yasui, M., Kwon, T. H., Knepper, M. A., Nielsen, S., and Agre, P. (1999). Aquaporin-6: An intracellular vesicle water channel protein in renal epithelia. Proceedings of the National Academy of Sciences of the United States of America of the United States of America *96*, 5808–5813.

192. Verkman, A. S. (2012). Aquaporins in clinical medicine. Annual Review of Medicine *63*, 303–316.

193. Dumas, M., Sadick, N. S., Noblesse, E., Juan, M., Lachmann-Weber, N., Boury-Jamot, M., Sougrat, R., Verbavatz, J. M., Schnebert, S., and Bonte, F. (2007). Hydrating skin by stimulating biosynthesis of aquaporins. Journal of Drugs in Dermatology *6*, s20–4.

194. Amiry-Moghaddam, M., Williamson, A., Palomba, M., Eid, T., de Lanerolle, N. C., Nagelhus, E. A., Adams, M. E., Froehner, S. C., Agre, P., and Ottersen, O. P. (2003). Delayed K+ clearance associated with aquaporin-4 mislocalization: phenotypic defects in brains of alpha-syntrophin-null mice. Proceedings of the National Academy of Sciences of the United States of America of the United States of America *100*, 13615–13620.

195. Yang, M., Gao, F., Liu, H., Yu, W. H., Zhuo, F., Qiu, G. P., Ran, J. H., and Sun, S. Q. (2013). Hyperosmotic induction of aquaporin expression in rat astrocytes through a different MAPK pathway. Journal of Cellular Biochemistry *114*, 111–119.

196. Hansen, A. K. and Galtung, H. K. (2007). Aquaporin expression and cell volume regulation in the SV40 immortalized rat submandibular acinar cell line. Pflugers Archiv - European Journal of Physiology *453*, 787–796.

197. Tradtrantip, L., Tajima, M., Li, L., and Verkman, A. S. (2009). Aquaporin water channels in transepithelial fluid transport. Journal of Medical Investigation *56 Suppl*, 179–184.

198. Iannotti, L. L., Tielsch, J. M., Black, M. M., and Black, R. E. (2006). Iron supplementation in early childhood: health benefits and risks. The American Journal of Clinical Nutrition *84*, 1261–1276.

199. Uc, A., Stokes, J. B., and Britigan, B. E. (2004). Heme transport exhibits polarity in Caco-2 cells: evidence for an active and membrane protein-mediated process. American Journal of Physiology - Gastrointestinal and Liver Physiology *287*, G1150–G1157.

200. West, A. R. and Oates, P. S. (2008). Mechanisms of heme iron absorption: current questions and controversies. World Journal of Gastroenterology *14*, 4101–4110.

201. Yang, Z., Philips, J. D., Doty, R. T., Giraudi, P., Ostrow, J. D., Tiribelli, C., Smith, A., and Abkowitz, J. L. (2010). Kinetics and specificity of feline leukemia virus subgroup C receptor (FLVCR) export function and its dependence on hemopexin. The Journal of Biological Chemistry *285*, 28874–28882.

202. Hvidberg, V., Maniecki, M. B., Jacobsen, C., Hojrup, P., Moller, H. J., and Moestrup, S. K. (2005). Identification of the receptor scavenging hemopexin-heme complexes. Blood *106*, 2572–2579.

203. Krishnamurthy, P., Ross, D. D., Nakanishi, T., Bailey-Dell, K., Zhou, S., Mercer, K. E., Sarkadi, B., Sorrentino, B. P., and Schuetz, J. D. (2004). The stem cell marker Bcrp/ABCG2 enhances hypoxic cell survival through interactions with heme. The Journal of Biological Chemistry *279*, 24218–24225.

204. Kendler, B. S. (2006). Supplemental conditionally essential nutrients in cardiovascular disease therapy. The Journal of Cardiovascular Nursing *21*, 9–16.

205. Soghier, L. M. and Brion, L. P. (2006). Cysteine, cystine or N-acetylcysteine supplementation in parenterally fed neonates. The Cochrane Database of Systematic Reviews pp. CD004869.

206. McKay Hart, A., Wiberg, M., and Terenghi, G. (2002). Pharmacological enhancement of peripheral nerve regeneration in the rat by systemic acetyl-L-carnitine treatment. Neuroscience Letters *334*, 181–185.

207. Selimoglu, M. A., Aydogdu, S., Yagci, R. V., and Huseyinov, A. (2001). Plasma and liver carnitine status of children with chronic liver disease and cirrhosis. Pediatrics International *43*, 391–395.

208. Amin, K. A. and Nagy, M. A. (2009). Effect of Carnitine and herbal mixture

extract on obesity induced by high fat diet in rats. Diabetology & Metabolic Syndrome *1*, 17.

209. Kobayashi, A., Masumura, Y., and Yamazaki, N. (1992). L-carnitine treatment for congestive heart failure–experimental and clinical study. Japanese Circulation Journal. The Official Journal Of The Japanese Circulation Society *56*, 86–94.

210. Flanagan, J. L., Simmons, P. A., Vehige, J., Willcox, M. D., and Garrett, Q. (2010). Role of carnitine in disease. Nutrition & Metabolism *7*, 30.

211. Xu, S., Flanagan, J. L., Simmons, P. A., Vehige, J., Willcox, M. D., and Garrett, Q. (2010). Transport of L-carnitine in human corneal and conjunctival epithelial cells. Molecular Vision *16*, 1823–1831.

212. Yabuuchi, H., Tamai, I., Nezu, J., Sakamoto, K., Oku, A., Shimane, M., Sai, Y., and Tsuji, A. (1999). Novel membrane transporter OCTN1 mediates multispecific, bidirectional, and pH-dependent transport of organic cations. Journal of Pharmacology and Experimental Therapeutics *289*, 768–773.

213. Ohashi, R., Tamai, I., Nezu Ji, J., Nikaido, H., Hashimoto, N., Oku, A., Sai, Y., Shimane, M., and Tsuji, A. (2001). Molecular and physiological evidence for multifunctionality of carnitine/organic cation transporter OCTN2. Molecular Pharmacology *59*, 358–366.

214. Enomoto, A., Wempe, M. F., Tsuchida, H., Shin, H. J., Cha, S. H., Anzai, N., Goto, A., Sakamoto, A., Niwa, T., Kanai, Y., et al. (2002). Molecular identification of a novel carnitine transporter specific to human testis. Insights into the mechanism of carnitine recognition. The Journal of Biological Chemistry *277*, 36262–36271.

215. Srinivas, S. R., Prasad, P. D., Umapathy, N. S., Ganapathy, V., and Shekhawat, P. S. (2007). Transport of butyryl-L-carnitine, a potential prodrug, via the carnitine transporter OCTN2 and the amino acid transporter ATB(0,+). American Journal of Physiology - Gastrointestinal and Liver Physiology *293*, G1046–G1053.

216. Tamai, I., Senmaru, M., Terasaki, T., and Tsuji, A. (1995). Na(+)- and Cl(-)-dependent transport of taurine at the blood-brain barrier. Biochemical Pharmacology *50*, 1783–1793.

217. Schaffer, S. W., Jong, C. J., Ramila, K. C., and Azuma, J. (2010). Physiological roles of taurine in heart and muscle. Journal of Biomedical Science *17*

Suppl 1, S2.

218. Guertin, F., Roy, C. C., Lepage, G., Perea, A., Giguere, R., Yousef, I., and Tuchweber, B. (1991). Effect of taurine on total parenteral nutrition-associated cholestasis. Journal of Parenteral and Enteral Nutrition *15*, 247–251.

219. Stapleton, P. P., Charles, R. P., Redmond, H. P., and Bouchier-Hayes, D. J. (1997). Taurine and human nutrition. Clinical Nutrition *16*, 103–108.

220. Heller-Stilb, B., van Roeyen, C., Rascher, K., Hartwig, H. G., Huth, A., Seeliger, M. W., Warskulat, U., and Haussinger, D. (2002). Disruption of the taurine transporter gene (taut) leads to retinal degeneration in mice. FASEB Journal *16*, 231–233.

221. Anderson, C. M., Howard, A., Walters, J. R., Ganapathy, V., and Thwaites, D. T. (2009). Taurine uptake across the human intestinal brush-border membrane is via two transporters: H+-coupled PAT1 (SLC36A1) and Na+- and Cl(-)-dependent TauT (SLC6A6). The Journal of Physiology *587*, 731–744.

222. Craig, S. A. (2004). Betaine in human nutrition. The American Journal of Clinical Nutrition *80*, 539–549.

223. Yamauchi, A., Uchida, S., Kwon, H. M., Preston, A. S., Robey, R. B., Garcia-Perez, A., Burg, M. B., and Handler, J. S. (1992). Cloning of a Na(+)- and Cl(-)-dependent betaine transporter that is regulated by hypertonicity. The Journal of Biological Chemistry *267*, 649–652.

224. Wagner, M. and Trauner, M. (2005). Transcriptional regulation of hepatobiliary transport systems in health and disease: implications for a rationale approach to the treatment of intrahepatic cholestasis. Annals of hepatology *4*, 77–99.

225. Hinoi, E., Takarada, T., Tsuchihashi, Y., and Yoneda, Y. (2005). Glutamate transporters as drug targets. Current Drug Targets - CNS & Neurological Disorders *4*, 211–220.

226. Amaral, M. D. and Kunzelmann, K. (2007). Molecular targeting of CFTR as a therapeutic approach to cystic fibrosis. Trends in Pharmacological Sciences *28*, 334–341.

227. Lo, M., Wang, Y. Z., and Gout, P. W. (2008). The x(c)- cystine/glutamate antiporter: a potential target for therapy of cancer and other diseases. Journal of Cellular Physiology *215*, 593–602.

228. Ganapathy, V., Thangaraju, M., and Prasad, P. D. (2009). Nutrient transporters in cancer: relevance to Warburg hypothesis and beyond. Pharmacology & Therapeutics *121*, 29–40.

229. Kimelberg, H. K. (2005). Astrocytic swelling in cerebral ischemia as a possible cause of injury and target for therapy. Glia *50*, 389–397.

230. Wright, E. M., Hirayama, B. A., and Loo, D. F. (2007). Active sugar transport in health and disease. Journal of Internal Medicine *261*, 32–43.

231. Moskowitz, S. M., Chmiel, J. F., Sternen, D. L., Cheng, E., and Cutting, G. R. (1993). CFTR-Related Disorders.

232. Thiele, I. and Palsson, B. O. (2010). Reconstruction annotation jamborees: a community approach to systems biology. Molecular Systems Biology *6*.

233. Thorens, B. and Mueckler, M. (2010). Glucose transporters in the 21st Century. American Journal of Physiology - Endocrinology and Metabolism *298*, E141–E145.

234. Teijema, H. L., van Gelderen, H. H., Giesberts, M. A., and Laurent de Angulo, M. S. (1974). Dicarboxylic aminoaciduria: an inborn error of glutamate and aspartate transport with metabolic implications, in combination with a hyperprolinemia. Metabolism *23*, 115–123.

235. de Koning-Tijssen, M. A. J. and Rees, M. I. (1993). Hyperekplexia.

236. Barnes, K., Dobrzynski, H., Foppolo, S., Beal, P. R., Ismat, F., Scullion, E. R., Sun, L., Tellez, J., Ritzel, M. W., Claycomb, W. C., et al. (2006). Distribution and functional characterization of equilibrative nucleoside transporter-4, a novel cardiac adenosine transporter activated at acidic pH. Circulation Research *99*, 510–519.

237. Morgan, N. V., Morris, M. R., Cangul, H., Gleeson, D., Straatman-Iwanowska, A., Davies, N., Keenan, S., Pasha, S., Rahman, F., Gentle, D., et al. (2010). Mutations in SLC29A3, encoding an equilibrative nucleoside transporter ENT3, cause a familial histiocytosis syndrome (Faisalabad histiocytosis) and familial Rosai-Dorfman disease. PLoS Genetics *6*, e1000833.

238. Diop-Bove, N., Kronn, D., and Goldman, I. D. (1993). Hereditary Folate Malabsorption.

239. Kanda, D., Takagi, H., Kawahara, Y., Yata, Y., Takakusagi, T., Hatanaka, T.,

Yoshinaga, T., Iesaki, K., Kashiwabara, K., Higuchi, T., et al. (2009). Novel large-scale deletion (whole exon 7) in the ABCC2 gene in a patient with the Dubin-Johnson syndrome. Drug Metabolism and Pharmacokinetics *24*, 464–468.

240. Kristiansen, M., Aminoff, M., Jacobsen, C., de La Chapelle, A., Krahe, R., Verroust, P. J., and Moestrup, S. K. (2000). Cubilin P1297L mutation associated with hereditary megaloblastic anemia 1 causes impaired recognition of intrinsic factor-vitamin B(12) by cubilin. Blood *96*, 405–409.

241. Wert, S. E., Whitsett, J. A., and Nogee, L. M. (2009). Genetic disorders of surfactant dysfunction. Pediatric and Developmental Pathology *12*, 253–274.

242. Knoers, N. (1993). Nephrogenic Diabetes Insipidus.

243. Kantarci, S., Donnai, D., Noonan, K. M., and Pober, B. R. (1993). Donnai-Barrow Syndrome.

244. Pober, B. R., Longoni, M., and Noonan, K. M. (2009). A review of Donnai-Barrow and facio-oculo-acoustico-renal (DB/FOAR) syndrome: clinical features and differential diagnosis. Birth defects research. Part A, Clinical and molecular teratology *85*, 76–81.

245. Glaser, B. (1993). Familial Hyperinsulinism.

246. Ghanem, E., Fritzsche, S., Al-Balushi, M., Hashem, J., Ghuneim, L., Thomer, L., Kalbacher, H., van Endert, P., Wiertz, E., Tampe, R., et al. (2010). The transporter associated with antigen processing (TAP) is active in a post-ER compartment. Journal of Cell Science *123*, 4271–4279.

247. Linton, K. J. and Holland, I. B. (2011). The ABC transporters of human physiology and disease: Genetics and biochemistry of ATP binding cassette transporters. (World scientific publishing co. pte. ltd.).

248. Mitsutake, S., Suzuki, C., Akiyama, M., Tsuji, K., Yanagi, T., Shimizu, H., and Igarashi, Y. (2010). ABCA12 dysfunction causes a disorder in glucosylceramide accumulation during keratinocyte differentiation. Journal of Dermatological Science *60*, 128–129.

249. Polisecki, E., Peter, I., Simon, J. S., Hegele, R. A., Robertson, M., Ford, I., Shepherd, J., Packard, C., Jukema, J. W., de Craen, A. J., et al. (2010). Genetic variation at the NPC1L1 gene locus, plasma lipoproteins, and heart disease risk in the elderly. Journal of Lipid Research *51*, 1201–1207.

250. Fernandes, J., Saudubray, J. M., van den Berghe, G., and Walter, J. H. (2006). Inborn metabolic diseases: diagnosis and treatment. (Springer publication).

251. Mamas, M., Dunn, W., Neyses, L., and Goodacre, R. (2011). The role of metabolites and metabolomics in clinically applicable biomarkers of disease. Archives of Toxicology *85*, 5–17.

252. Leonard, J. V. and Dezateux, C. (2011). Newborn screening for inborn errors of metabolism: principles, policies and weighing the evidence. Paediatrics and Child Health *21*, 56–60.

253. McCabe, E. R. (2010). Inborn Errors of Metabolism: The metabolome is our world. Presidential address for the 11th International Congress of Inborn Errors of Metabolism (ICIEM). Molecular Genetics and Metabolism *100*, 1–5.

254. Chace, D. H. and Kalas, T. A. (2005). A biochemical perspective on the use of tandem mass spectrometry for newborn screening and clinical testing. Clinical Biochemistry *38*, 296–309.

255. Howell, R. R. (2011). Quality improvement of newborn screening in real time. Genetics in Medicine *13*, 205.

256. Levy, H. L. (2010). Newborn screening conditions: What we know, what we do not know, and how we will know it. Genetics in Medicine *12*, S213–S214.

257. Chalcraft, K. R. and Britz-McKibbin, P. (2009). Newborn screening of inborn errors of metabolism by capillary electrophoresis-electrospray ionization-mass spectrometry: a second-tier method with improved specificity and sensitivity. Analytical Chemistry *81*, 307–314.

258. Lindner, M., Hoffmann, G. F., and Matern, D. (2010). Newborn screening for disorders of fatty-acid oxidation: experience and recommendations from an expert meeting. Journal of Inherited Metabolic Disease *33*, 521–526.

259. Rector, R. S. and Ibdah, J. A. (2010). Fatty acid oxidation disorders: maternal health and neonatal outcomes. Seminars in Fetal & Neonatal Medicine *15*, 122–128.

260. Oberhardt, M. A., Palsson, B. O., and Papin, J. A. (2009). Applications of genome-scale metabolic reconstructions. Molecular Systems Biology *5*, 320.

261. Valayannopoulos, V., Haudry, C., Serre, V., Barth, M., Boddaert, N., Arnoux, J. B., Cormier-Daire, V., Rio, M., Rabier, D., Vassault, A., et al. (2010). New

SUCLG1 patients expanding the phenotypic spectrum of this rare cause of mild methylmalonic aciduria. Mitochondrion *10*, 335–341.

262. Horvath, T. D., Stratton, S. L., Bogusiewicz, A., Pack, L., Moran, J., and Mock, D. M. (2010). Quantitative measurement of plasma 3-hydroxyisovaleryl carnitine by LC-MS/MS as a novel biomarker of biotin status in humans. Analytical Chemistry *82*, 4140–4144.

263. Eminoglu, F. T., Ozcelik, A. A., Okur, I., Tumer, L., Biberoglu, G., Demir, E., Hasanoglu, A., and Baumgartner, M. R. (2009). 3-Methylcrotonyl-CoA carboxylase deficiency: phenotypic variability in a family. Journal of Child Neurology *24*, 478–481.

264. Popek, M., Walter, M., Fernando, M., Lindner, M., Schwab, K. O., and Sass, J. O. (2010). Two inborn errors of metabolism in a newborn: glutaric aciduria type I combined with isobutyrylglycinuria. Clinica Chimica Acta; International Journal of Clinical Chemistry and Diagnostic Laboratory Medicine; International Journal of Clinical Chemistry and Diagnostic Laboratory Medicine *411*, 2087–2091.

265. Sauer, S. W., Opp, S., Hoffmann, G. F., Koeller, D. M., Okun, J. G., and Kolker, S. (2011). Therapeutic modulation of cerebral L-lysine metabolism in a mouse model for glutaric aciduria type I. Brain *134*, 157–170.

266. Blau, N., Duran, M., and Gibson, K. M. (2008). Laboratory guide to the methods in biochemical genetics. (Springer).

267. Jones, P. M., Quinn, R., Fennessey, P. V., Tjoa, S., Goodman, S. I., Fiore, S., Burlina, A. B., Rinaldo, P., Boriack, R. L., and Bennett, M. J. (2000). Improved stable isotope dilution-gas chromatography-mass spectrometry method for serum or plasma free 3-hydroxy-fatty acidS. and its utility for the study of disorders of mitochondrial fatty acid beta-oxidation. Clinical Chemistry *46*, 149–155.

268. Hori, T., Fukao, T., Kobayashi, H., Teramoto, T., Takayanagi, M., Hasegawa, Y., Yasuno, T., Yamaguchi, S., and Kondo, N. (2010). Carnitine palmitoyltransferase 2 deficiency: the time-course of blood and urinary acylcarnitine levels during initial L-carnitine supplementation. The Tohoku Journal of Experimental Medicine *221*, 191–195.

269. Fontaine, M., Briand, G., Ser, N., Armelin, I., Rolland, M. O., Degand, P., and Vamecq, J. (1996). Metabolic studies in twin brothers with 2-methylacetoacetyl-CoA thiolase deficiency. Clinica Chimica Acta; Interna-

tional Journal of Clinical Chemistry and Diagnostic Laboratory Medicine; International Journal of Clinical Chemistry and Diagnostic Laboratory Medicine *255*, 67–83.

270. Castelnovi, C., Moseley, K., and Yano, S. (2010). Maternal isovaleric acidemia: observation of distinctive changes in plasma amino acidS. and carnitine profiles during pregnancy. Clinica Chimica Acta; International Journal of Clinical Chemistry and Diagnostic Laboratory Medicine; International Journal of Clinical Chemistry and Diagnostic Laboratory Medicine *411*, 2101–2103.

271. Forni, S., Fu, X., Palmer, S. E., and Sweetman, L. (2010). Rapid determination of C4-acylcarnitine and C5-acylcarnitine isomers in plasma and dried blood spots by UPLC-MS/MS as a second tier test following flow-injection MS/MS acylcarnitine profile analysis. Molecular Genetics and Metabolism *101*, 25–32.

272. Ramsay, R. R., Gandour, R. D., and van der Leij, F. R. (2001). Molecular enzymology of carnitine transfer and transport. Biochimica et Biophysica Acta (BBA) *1546*, 21–43.

273. Wanders, R. J., Visser, W. F., van Roermund, C. W., Kemp, S., and Waterham, H. R. (2007). The peroxisomal ABC transporter family. Pflugers Archiv - European Journal of Physiology *453*, 719–734.

274. Wanders, R. J., Ruiter, J. P., L, I., Waterham, H. R., and Houten, S. M. (2010). The enzymology of mitochondrial fatty acid beta-oxidation and its application to follow-up analysis of positive neonatal screening results. Journal of Inherited Metabolic Disease *33*, 479–494.

275. Watkins, P. A., Maiguel, D., Jia, Z., and Pevsner, J. (2007). Evidence for 26 distinct acyl-coenzyme A synthetase genes in the human genome. Journal of Lipid Research *48*, 2736–2750.

276. Vock, C., Biedasek, K., Boomgaarden, I., Heins, A., Nitz, I., and Doring, F. (2010). ACBP knockdown leads to down-regulation of genes encoding rate-limiting enzymes in cholesterol and fatty acid metabolism. Cellular Physiology and Biochemistry : International Journal of Experimental Cellular Physiology, Biochemistry and Pharmacology *25*, 675–686.

277. Wanders, R. J., van Grunsven, E. G., and Jansen, G. A. (2000). Lipid metabolism in peroxisomes: enzymology, functions and dysfunctions of the fatty acid alpha- and beta-oxidation systems in humans. Biochemical Society

Transactions *28*, 141–149.

278. Wanders, R. J., Ferdinandusse, S., Brites, P., and Kemp, S. (2010). Peroxisomes, lipid metabolism and lipotoxicity. Biochimica et Biophysica Acta (BBA) *1801*, 272–280.

279. Geisbrecht, B. V., Zhang, D., Schulz, H., and Gould, S. J. (1999). Characterization of PECI, a novel monofunctional Delta(3), Delta(2)-enoyl-CoA isomerase of mammalian peroxisomes. The Journal of Biological Chemistry *274*, 21797–21803.

280. Street, J. M., Singh, H., and Poulos, A. (1990). Metabolism of saturated and polyunsaturated very-long-chain fatty acids in fibroblasts from patients with defects in peroxisomal beta-oxidation. Biochemical Journal *269*, 671–677.

281. Heinzer, A. K., Kemp, S., Lu, J. F., Watkins, P. A., and Smith, K. D. (2002). Mouse very long-chain acyl-CoA synthetase in X-linked adrenoleukodystrophy. The Journal of Biological Chemistry *277*, 28765–28773.

282. Pettersen, J. E. (1973). In vitro studies on the metabolism of hexadecanedioic acid and its mono-L-carnitine ester. Biochimica et Biophysica Acta (BBA) *306*, 1–14.

283. Wanders, R. J. and Tager, J. M. (1998). Lipid metabolism in peroxisomes in relation to human disease. Molecular Aspects of Medicine *19*, 69–154.

284. Ferdinandusse, S., Denis, S., Van Roermund, C. W., Wanders, R. J., and Dacremont, G. (2004). Identification of the peroxisomal beta-oxidation enzymes involved in the degradation of long-chain dicarboxylic acids. Journal of Lipid Research *45*, 1104–1111.

285. Sanders, R. J., Ofman, R., Valianpour, F., Kemp, S., and Wanders, R. J. (2005). Evidence for two enzymatic pathways for omega-oxidation of docosanoic acid in rat liver microsomes. Journal of Lipid Research *46*, 1001–1008.

286. Fer, M., Corcos, L., Dreano, Y., Plee-Gautier, E., Salaun, J. P., Berthou, F., and Amet, Y. (2008). Cytochromes P450 from family 4 are the main omega hydroxylating enzymes in humans: CYP4F3B is the prominent player in PUFA metabolism. Journal of Lipid Research *49*, 2379–2389.

287. Rolfsson, O., Palsson, B. O., and Thiele, I. (2011). The human metabolic reconstruction Recon 1 directs hypotheses of novel human metabolic functions. BMC Systems Biology *5*, 155.

288. Freeze, H. H. and Sharma, V. (2010). Metabolic manipulation of glycosylation disorders in humans and animal models. Seminars in Cell & Developmental Biology *21*, 655–662.

289. Scheuner, M. T., Yoon, P. W., and Khoury, M. J. (2004). Contribution of Mendelian disorders to common chronic disease: opportunities for recognition, intervention, and prevention. American Journal of Medical Genetics Part C Seminars in Medical Genetic *125C*, 50–65.

290. Nadeau, J. H. (2001). Modifier genes in mice and humans. Nature Reviews Genetics *2*, 165–174.

291. Weatherall, D. J. (2001). Phenotype-genotype relationships in monogenic disease: lessons from the thalassaemias. Nature Reviews Genetics *2*, 245–55.

292. Dagli, A. I. and Weinstein, D. A. (1993). Glycogen Storage Disease Type VI.

293. Seashore, M. R. (1993). The Organic Acidemias: An Overview.

294. Brady, S., Siegel, G., Albers, R. W., and Price, D. (2006). Basic neurochemistry: molecular, cellular and medical aspects. (Elsevier) 7th edition.

295. Felber, S. R., Sperl, W., Chemelli, A., Murr, C., and Wendel, U. (1993). Maple syrup urine disease: metabolic decompensation monitored by proton magnetic resonance imaging and spectroscopy. Annals of Neurology *33*, 396–401.

296. Jan, W., Zimmerman, R. A., Wang, Z. J., Berry, G. T., Kaplan, P. B., and Kaye, E. M. (2003). MR diffusion imaging and MR spectroscopy of maple syrup urine disease during acute metabolic decompensation. Neuroradiology *45*, 393–399.

297. Beadle, R. M. and Frenneaux, M. (2010). Modification of myocardial substrate utilisation: a new therapeutic paradigm in cardiovascular disease. Heart *96*, 824–830.

298. Piraud, M., Ruet, S., Boyer, S., Acquaviva, C., Clerc-Renaud, P., Cheillan, D., and Vianey-Saban, C. (2011). Amino acid profiling for the diagnosis of inborn errors of metabolism. Methods in Molecular Biology *708*, 25–53.

299. Santra, S. and Hendriksz, C. (2010). How to use acylcarnitine profiles to help diagnose inborn errors of metabolism. Archives of Disease in Childhood - Education and Practice *95*, 151–156.

300. Jones, L. L., McDonald, D. A., and Borum, P. R. (2010). Acylcarnitines: role in brain. Progress in Lipid Research *49*, 61–75.

301. Wilcken, B. and Wiley, V. (2008). Newborn screening. Pathology *40*, 104–115.

302. Boneh, A., Andresen, B. S., Gregersen, N., Ibrahim, M., Tzanakos, N., Peters, H., Yaplito-Lee, J., and Pitt, J. J. (2006). VLCAD deficiency: pitfalls in newborn screening and confirmation of diagnosis by mutation analysis. Molecular Genetics and Metabolism *88*, 166–170.

303. Leonard, J. V. (1995). The management and outcome of propionic and methylmalonic acidaemia. Journal of Inherited Metabolic Disease *18*, 430–434.

304. Proia, R. L., D'Azzo, A., and Neufeld, E. F. (1984). Association of alpha- and beta-subunits during the biosynthesis of beta-hexosaminidase in cultured human fibroblasts. The Journal of Biological Chemistry *259*, 3350–3354.

305. Mark, B. L., Mahuran, D. J., Cherney, M. M., Zhao, D., Knapp, S., and James, M. N. G. (2003). Crystal Structure of Human [beta]-Hexosaminidase B: Understanding the Molecular Basis of Sandhoff and Tay-Sachs Disease. Journal of Molecular Biology *327*, 1093–1109.

306. Norflus, F., Yamanaka, S., and Proia, R. L. (1996). Promoters for the human beta-hexosaminidase genes, HEXA and HEXB. DNA and Cell Biology *15*, 89–97.

307. Branda, K. J., Tomczak, J., and Natowicz, M. R. (2004). Heterozygosity for Tay-Sachs and Sandhoff diseases in non-Jewish Americans with ancestry from Ireland, Great Britain, or Italy. Genetic Testing *8*, 174–180.

308. Sango, K., Yamanaka, S., Hoffmann, A., Okuda, Y., Grinberg, A., Westphal, H., McDonald, M. P., Crawley, J. N., Sandhoff, K., Suzuki, K., et al. (1995). Mouse models of Tay-Sachs and Sandhoff diseases differ in neurologic phenotype and ganglioside metabolism. Nature Genetics *11*, 170–176.

309. Becker, K. L. (2001). Principles and practice of endocrionology and metabolism. (Lippincot williams and wilkins) third edit edition.

310. Lee, D. S., Park, J., Kay, K. A., Christakis, N. A., Oltvai, Z. N., and Barabasi, A. L. (2008). The implications of human metabolic network topology for disease comorbidity. Proceedings of the National Academy of Sciences of the United States of America of the United States of America *105*, 9880–9885.

311. Shlomi, T. (2010). Metabolic network-based interpretation of gene expression data elucidates human cellular metabolism. Biotechnology and Genetic Engineering Reviews *26*, 281–96.

312. Li, L., Rasul, I., Liu, J., Zhao, B., Tang, R., Premont, R. T., and Suo, W. Z. (2009). Augmented axonal defects and synaptic degenerative changes in female GRK5 deficient mice. Brain Research Bulletin *78*, 145–151.

313. Papin, J. A. and Palsson, B. O. (2004). The JAK-STAT signaling network in the human B-cell: an extreme signaling pathway analysis. Biophysical Journal *87*, 37–46.

314. Dasika, M. S., Burgard, A., and Maranas, C. D. (2006). A computational framework for the topological analysis and targeted disruption of signal transduction networks. Biophysical Journal *91*, 382–98.

315. Thiele, I., Fleming, R. M. T., Bordbar, A., Schellenberger, J., and Palsson, B. O. (2010). Functional characterization of alternate optimal solutions of Escherichia coli's transcriptional and translational machinery. Biophysical Journal *98*, 2072–2081.

316. Thiele, I., Jamshidi, N., Fleming, R. M. T., and Palsson, B. O. (2009). Genome-scale reconstruction of Escherichia coli's transcriptional and translational machinery: a knowledge base, its mathematical formulation, and its functional characterization. PLoS Computational Biology *5*, e1000312.

317. Gianchandani, E. P., Joyce, A. R., Palsson, B. O., and Papin, J. A. (2009). Functional states of the genome-scale Escherichia coli transcriptional regulatory system. PLoS Computational Biology *5*, e1000403.

318. Gianchandani, E. P., Papin, J. A., Price, N. D., Joyce, A. R., and Palsson, B. O. (2006). Matrix formalism to describe functional states of transcriptional regulatory systems. PLoS Computational Biology *2*, e101.

319. Illig, T., Gieger, C., Zhai, G., Romisch-Margl, W., Wang-Sattler, R., Prehn, C., Altmaier, E., Kastenmuller, G., Kato, B. S., Mewes, H.-W., et al. (2010). A genome-wide perspective of genetic variation in human metabolism. Nature Genetics *42*, 137–41.

320. Schellenberger, J., Park, J. O., Conrad, T. M., and Palsson, B. O. (2010). BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. BMC Bioinformatics *11*, 213.

321. Hucka, M., Finney, A., Sauro, H. M., Bolouri, H., Doyle, J. C., Kitano, H., Arkin, A. P., Bornstein, B. J., Bray, D., Cornish-Bowden, A., et al. (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics *19*, 524–31.

322. Fleming, R. M. and Thiele, I. (2011). von Bertalanffy 1.0: a COBRA toolbox extension to thermodynamically constrain metabolic models. Bioinformatics *27*, 142–143.

323. Fleming, R. M., Thiele, I., and Nasheuer, H. P. (2009). Quantitative assignment of reaction directionality in constraint-based models of metabolism: application to Escherichia coli. Biophysical Chemistry *145*, 47–56.

324. Haraldsdottir, H. S., Thiele, I., and Fleming, R. M. T. (2012). Quantitative assignment of reaction directionality in a multicompartmental human metabolic reconstruction. Biophysical Journal *102*, 1703–11.

325. Saheki, T., Kobayashi, K., Iijima, M., Horiuchi, M., Begum, L., Jalil, M. A., Li, M. X., Lu, Y. B., Ushikai, M., Tabata, A., et al. (2004). Adult-onset type II citrullinemia and idiopathic neonatal hepatitis caused by citrin deficiency: involvement of the aspartate glutamate carrier for urea synthesis and maintenance of the urea cycle. Molecular Genetics and Metabolism *81*, 20–26.

326. Gross, M. (2001). Adenosine monophosphate deaminase deficiency.

327. Gross, M. (1997). Clinical heterogeneity and molecular mechanisms in inborn muscle AMP deaminase deficiency. Journal of Inherited Metabolic Disease *20*, 186–192.

328. Safranow, K., Suchy, J., Jakubowska, K., Olszewska, M., Biaczak-Kuleta, A., Kurzawski, G., Rzeuski, R., Czyaycka, E., Aoniewska, B., Kornacewicz-Jach, Z., et al. (2011). AMPD1gene mutations are associated with obesity and diabetes in Polish patients with cardiovascular diseases. Journal of Applied Genetics *52*, 67–76.

329. Walker, A. (2004). Pediatric gastrointestial disease: pathophysiology, diagnosis, management volume 1. (B.C. Decker Inc) 4th edition.

330. Gitzelmann, R., Spycher, M. A., Feil, G., Muller, J., Seilnacht, B., Stahl, M., and Bosshard, N. U. (1996). Liver glycogen synthase deficiency: A rarely diagnosed entity. European Journal of Pediatrics *155*, 561–567.

331. Weinstein, D. A., Correia, C. E., Saunders, A. C., and Wolfsdorf, J. I. (2006).

Hepatic glycogen synthase deficiency: An infrequently recognized cause of ketotic hypoglycemia. Molecular Genetics and Metabolism *87*, 284–288.

332. Kollberg, G., Tulinius, M., Gilljam, T., Ostman-Smith, I., Forsander, G., Jotorp, P., Oldfors, A., and Holme, E. (2007). Cardiomyopathy and Exercise Intolerance in Muscle Glycogen Storage Disease 0. New England Journal of Medicine *357*, 1507–1514.

333. van den Akker, E., Satchwell, T. J., Williamson, R. C., and Toye, A. M. (2010). Band 3 multiprotein complexes in the red cell membrane; of mice and men. Blood Cells, Molecules, and Diseases *45*, 1–8.

334. Wrong, O., Bruce, L. J., Unwin, R. J., Toye, A. M., and Tanner, M. J. A. (2002). Band 3 mutations, distal renal tubular acidosis, and Southeast Asian ovalocytosis. Kidney International *62*, 10–19.

335. Ribeiro, M. L., Alloisio, N., Almeida, H., Gomes, C., Texier, P., Lemos, C., Mimoso, G., Morle, L., Bey-Cabet, F., Rudigoz, R.-C., et al. (2000). Severe hereditary spherocytosis and distal renal tubular acidosis associated with the total absence of band 3. Blood *96*, 1602–1604.

336. Goodman, B. E. (2010). Insights into digestion and absorption of major nutrients in humans. Advances in Physiology Education *34*, 44–53.

337. Lin, J. H., Chiba, M., and Baillie, T. A. (1999). Is the role of the small intestine in first-pass metabolism overemphasized? Pharmacological Reviews *51*, 135–158.

338. Pang, K. S. (2003). Modeling of intestinal drug absorption: roles of transporters and metabolic enzymes (for the Gillette Review Series). Drug Metabolism and Disposition *31*, 1507–1519.

339. Mootha, V. K. and Hirschhorn, J. N. (2010). Inborn variation in metabolism. Nature Genetics *42*, 97–98.

340. Roy, C. N. and Andrews, N. C. (2001). Recent advances in disorders of iron metabolism: mutations, mechanisms and modifiers. Human Molecular Genetics *10*, 2181–2186.

341. Martins, A. M. (1999). Inborn errors of metabolism: a clinical overview. Sao Paulo Medical Journal *117*, 251–265.

342. Jamshidi, N., Vo, T. D., and Palsson, B. O. (2007). In silico analysis of SNPS

and other high-throughput data. In Cardiac Gene Expression In Cardiac Gene Expression. (Springer).

343. Frezza, C., Zheng, L., Folger, O., Rajagopalan, K. N., MacKenzie, E. D., Jerby, L., Micaroni, M., Chaneton, B., Adam, J., Hedley, A., et al. (2011). Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase. Nature *477*, 225–228.

344. Thiele, I., Price, N. D., Vo, T. D., and Palsson, B. O. (2005). Candidate metabolic network states in human mitochondria. Impact of diabetes, ischemia, and diet. Journal of Biological Chemistry *280*, 11683–11695.

345. Sigurdsson, M. I., Jamshidi, N., Jonsson, J. J., and Palsson, B. O. (2009). Genome-scale network analysis of imprinted human metabolic genes. Epigenetics *4*, 43–46.

346. Le Gall, M., Tobin, V., Stolarczyk, E., Dalet, V., Leturque, A., and Brot-Laroche, E. (2007). Sugar sensing by enterocytes combines polarity, membrane bound detectors and sugar metabolism. Journal of Cellular Physiology *213*, 834–843.

347. U.S. Department of Agriculture, Agricultural Research Service. 2010 (2010). Nutrient Intakes from Food: Mean Amounts Consumed per Individual, by Gender and Age, What We Eat in America.

348. U.S. Department of Agriculture, Agricultural Research Service. 2011 (2011). USDA National Nutrient Database for Standard Reference, Release 24. Nutrient Data Laboratory Home Page.

349. A Report of the Panel on Macronutrients, Subcommittees on Upper Reference Levels of Nutrients Interpretation, Uses of Dietary Reference Intakes, Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. (2005). Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids (Macronutrients). (Washington, DC: The National Academies Press).

350. Crenn, P., Messing, B., and Cynober, L. (2008). Citrulline as a biomarker of intestinal failure due to enterocyte mass reduction. Clinical Nutrition *27*, 328–339.

351. Kohler, E. S., Sankaranarayanan, S., van Ginneken, C. J., van Dijk, P., Vermeulen, J. L., Ruijter, J. M., Lamers, W. H., and Bruder, E. (2008). The human neonatal small intestine has the potential for arginine synthesis; devel-

opmental changes in the expression of arginine-synthesizing and catabolizing enzymes. BMC Developmental Biology *8*, 107.

352. Yamada, E., Wakabayashi, Y., Saito, A., Yoda, K., Tanaka, Y., and Miyazaki, M. (1993). Hyperammonaemia caused by essential aminoacid supplements in patient with short bowel. The Lancet *341*, 1542–1543.

353. Stanbury, J. B. and Wyngaarden, J. B. and Fredrickson, D. S. and Goldstein, J. L. and Brown, M. S. (1983). The metabolic basis of inherited disease. (New York, NY: McGraw-Hill Book Company) 5th edition.

354. Treacy, E., Arbour, L., Chessex, P., Graham, G., Kasprzak, L., Casey, K., Bell, L., Mamer, O., and Scriver, C. R. (1996). Glutathione deficiency as a complication of methylmalonic acidemia: Response to high doses of ascorbate. The Journal of Pediatrics *129*, 445–448.

355. Thoene, J. G. (1993). Citrullinemia Type I.

356. Irons, M. (1993). Smith-Lemli-Opitz Syndrome.

357. Broer, S. (2008). Apical transporters for neutral amino acids: physiology and pathophysiology. Physiology (Bethesda) *23*, 95–103.

358. Grundy, S. M., Brewer Jr., H. B., Cleeman, J. I., Smith Jr., S. C., and Lenfant, C. (2004). Definition of metabolic syndrome: Report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. Circulation *109*, 433–438.

359. Zivkovic, A. M., German, J. B., and Sanyal, A. J. (2007). Comparative review of diets for the metabolic syndrome: implications for nonalcoholic fatty liver disease. The American Journal of Clinical Nutrition *86*, 285–300.

360. Jalal, D. I., Smits, G., Johnson, R. J., and Chonchol, M. (2010). Increased fructose associates with elevated blood pressure. Journal of the American Society of Nephrology *21*, 1543–1549.

361. Kawasaki, T., Akanuma, H., and Yamanouchi, T. (2002). Increased fructose concentrations in blood and urine in patients with diabetes. Diabetes Care *25*, 353–357.

362. Flock, M. R. and Kris-Etherton, P. M. (2011). Dietary Guidelines for Americans 2010: implications for cardiovascular disease. Current Atherosclerosis Reports *13*, 499–507.

363. Grundy, S. M., Cleeman, J. I., Daniels, S. R., Donato, K. A., Eckel, R. H., Franklin, B. A., Gordon, D. J., Krauss, R. M., Savage, P. J., Smith Jr., S. C., et al. (2005). Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Circulation *112*, 2735–2752.

364. Thunell, S. (2000). Porphyrins, porphyrin metabolism and porphyrias. I. Update. Scandinavian Journal of Clinical & Laboratory Investigation *60*, 509–540.

365. Budhoo, M. R., Mitchell, S., Eddlestone, J., and MacLennan, I. (1999). Sickle cell trait and acute intermittent porphyria leading to small bowel infarction. Journal of The Royal College of Surgeons of Edinburgh *44*, 130–131.

366. Twaddle, S., Wassif, W. S., Deacon, A. C., and Peters, T. J. (2001). Celiac Disease in Patients with Variegate Porphyria. Digestive Diseases and Sciences *46*, 1506–1508.

367. Stocker, R., Yamamoto, Y., McDonagh, A. F., Glazer, A. N., and Ames, B. N. (1987). Bilirubin is an antioxidant of possible physiological importance. Science *235*, 1043–1046.

368. Oates, P. S. and West, A. R. (2006). Heme in intestinal epithelial cell turnover, differentiation, detoxification, inflammation, carcinogenesis, absorption and motility. World Journal of Gastroenterology *12*, 4281–4295.

369. Gibbons, S. J. and Farrugia, G. (2004). The role of carbon monoxide in the gastrointestinal tract. The Journal of Physiology *556*, 325–336.

370. Zuckerbraun, B. S., Otterbein, L. E., Boyle, P., Jaffe, R., Upperman, J., Zamora, R., and Ford, H. R. (2005). Carbon monoxide protects against the development of experimental necrotizing enterocolitis. American Journal of Physiology - Gastrointestinal and Liver Physiology *289*, G607–G613.

371. Sieg, I., Beckh, K., Kersten, U., and Doss, M. O. (1991). Manifestation of acute intermittent porphyria in patients with chronic inflammatory bowel disease. Zeitschrift Fur Gastroenterologie *29*, 602–605.

372. Valtonen, M., Nanto-Salonen, K., Jaaskelainen, S., Heinanen, K., Alanen, A., Heinonen, O. J., Lundbom, N., Erkintalo, M., and Simell, O. (1999). Central nervous system involvement in gyrate atrophy of the choroid and retina with hyperornithinaemia. Journal of inherited metabolic disease *22*, 855–866.

373. Lanpher, B. C., Gropman, A., Chapman, K. A., Lichter-Konecki, U., Urea Cycle Disorders, Consortium, and Summar, M. L. (1993). GeneReviews Internet: Urea Cycle Disorders Overview.).

374. Wang, F. Y. U., Zhu, R. M., Maemura, K., Hirata, I., Katsu, K., and Watanabe, M. (2006). Expression of gamma-aminobutyric acid and glutamic acid decarboxylases in rat descending colon and their relation to epithelial differentiation. Chinese Journal of Digestive Diseases *7*, 103–108.

375. Collins, S. M. and Bercik, P. (2009). The relationship between intestinal microbiota and the central nervous system in normal gastrointestinal function and disease. Gastroenterology *136*, 2003–2014.

376. Prins, A. (2011). The brain-gut interaction: the conversation and the implications. South African Journal of Clinical Nutrition *24*, s8–s14.

377. Bravo, J. A., Forsythe, P., Chew, M. V., Escaravage, E., Savignac, H. M., Dinan, T. G., Bienenstock, J., and Cryan, J. F. (2011). Ingestion of Lactobacillus strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. Proceedings of the National Academy of Sciences of the United States of America of the United States of America *108*, 16050–16055.

378. Mani, S., Yang, G., and Wang, R. (2011). A critical life-supporting role for cystathionine gamma-lyase in the absence of dietary cysteine supply. Free Radical Biology & Medicine *50*, 1280–1287.

379. Heil, S. G., Hogeveen, M., Kluijtmans, L. A., van Dijken, P. J., van de Berg, G. B., Blom, H. J., and Morava, E. (2007). Marfanoid features in a child with combined methylmalonic aciduria and homocystinuria (CblC type). Journal of Inherited Metabolic Disease *30*, 811.

380. Stremmel, W., Hanemann, A., Ehehalt, R., Karner, M., and Braun, A. (2010). Phosphatidylcholine (lecithin) and the mucus layer: Evidence of therapeutic efficacy in ulcerative colitis? Digestive Diseases (Basel, Switzerland) *28*, 490–496.

381. Schneider, H., Braun, A., Fullekrug, J., Stremmel, W., and Ehehalt, R. (2010). Lipid based therapy for ulcerative colitis-modulation of intestinal mucus membrane phospholipids as a tool to influence inflammation. International Journal of Molecular Sciences *11*, 4149–4164.

382. Fenton, W. A., Gravel, R. A., and Rosenblatt, D. S. (2001). The metabolic and

molecular bases of inherited disease. (New York: McGraw-Hill Companies).

383. Gudmundsson, S. and Thiele, I. (2010). Computationally efficient flux variability analysis. BMC Bioinformatics *11*, 489.

384. UniProt Consortium. (2012). Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Research *40*, D71–D75.

385. Lockhart, R. D., Hamilton, G. F., and Fyfe, F. W. (1974). Anatomy of the human body. (MacLehose, R. and company limited, London).

386. Wise Jr., E. M. and Elwyn, D. (1965). Rates of reactions involved in phosphtide synthesis in liver and small intestine of intact rats. The Journal of Biological Chemistry *240*, 1537–1548.

387. McDevitt, J., Feighery, C., O'Farrelly, C., Martin, G., Weir, D. G., and Kelleher, D. (1995). Effects of tumour necrosis factor-alpha on BrdU incorporation in cultured human enterocytes. Mediators of Inflammation *4*, 31–37.

388. Bouchoux, J., Beilstein, F., Pauquai, T., Guerrera, I. C., Chateau, D., Ly, N., Alqub, M., Klein, C., Chambaz, J., Rousset, M., et al. (2011). The proteome of cytosolic lipid droplets isolated from differentiated Caco-2/TC7 enterocytes reveals cell-specific characteristics. Biology of the Cell *103*, 499–517.

389. Fernandez, M. L. and Calle, M. (2010). Revisiting dietary cholesterol recommendations: does the evidence support a limit of 300 mg/d? Current Atherosclerosis Reports *12*, 377–383.

390. Chougule, P., Herlenius, G., Hernandez, N. M., Patil, P. B., Xu, B., and Sumitran-Holgersson, S. (2012). Isolation and characterization of human primary enterocytes from small intestine using a novel method. Scandinavian Journal of Gastroenterology *47*, 1334–1343.

391. Paine, M. F., Khalighi, M., Fisher, J. M., Shen, D. D., Kunze, K. L., Marsh, C. L., Perkins, J. D., and Thummel, K. E. (1997). Characterization of interintestinal and intraintestinal variations in human CYP3A-dependent metabolism. Journal of Pharmacology and Experimental Therapeutics *283*, 1552–1562.

392. Lecha, M., Puy, H., and Deybach, J. C. (2009). Erythropoietic protoporphyria. Orphanet Journal of Rare Diseases *4*, 19.

393. Barnes, H. D., Hurworth, E., and Millar, J. H. (1968). Erythropoietic por-

phyrin hepatitis. Journal of Clinical Pathology *21*, 157–159.

394. Meerman, L., Koopen, N. R., Bloks, V., van Goor, H., Havinga, R., Wolthers, B. G., Kramer, W., Stengelin, S., Muller, M., Kuipers, F., et al. (1999). Biliary fibrosis associated with altered bile composition in a mouse model of erythropoietic protoporphyria. Gastroenterology *117*, 696–705.

395. Donovan, A., Roy, C. N., and Andrews, N. C. (2006). The in and outs of iron homeostasis. Physiology *21*, 115–123.

396. Vassallo, M. J., Camilleri, M., Sullivan, S. N., and Thomforde, G. M. (1992). Effects of erythromycin on gut transit in pseudo-obstruction due to hereditary coproporphyria. Journal of Clinical Gastroenterology *14*, 255–259.

397. Migchielsen, A. A., Breuer, M. L., van Roon, M. A., te Riele, H., Zurcher, C., Ossendorp, F., Toutain, S., Hershfield, M. S., Berns, A., and Valerio, D. (1995). Adenosine-deaminase-deficient mice die perinatally and exhibit liver-cell degeneration, atelectasis and small intestinal cell death. Nature Genetics *10*, 279–287.

398. Pita, A. M., Fernandez-Bustos, A., Rodes, M., Arranz, J. A., Fisac, C., Virgili, N., Soler, J., and Wakabayashi, Y. (2004). Orotic aciduria and plasma urea cycle-related amino acid alterations in short bowel syndrome, evoked by an arginine-free diet. Journal of Parenteral and Enteral Nutrition *28*, 315–323.

399. Hendriksz, C. J. and Gissen, P. (2011). Glycogen storage disease. Paediatrics and Child Health *21*, 84–89.

400. Guery, M. J., Douillard, C., Marcelli-Tourvieille, S., Dobbelaere, D., Wemeau, J. L., and Vantyghem, M. C. (2007). Doctor, my son is so tired... about a case of hereditary fructose intolerance. Annales d Endocrinologie *68*, 456–459.

401. Buchman, A. L. (2006). Clinical nutrition in gastrointestinal disease. (SLACK incorporated, USA).

402. Chung-Faye, G., Hayee, B., Maestranzi, S., Donaldson, N., Forgacs, I., and Sherwood, R. (2007). Fecal M2-pyruvate kinase (M2-PK): A novel marker of intestinal inflammation. Inflammatory Bowel Diseases *13*, 1374–1378.

403. Savilahti, E., Launiala, K., and Kuitunen, P. (1983). Congenital lactase deficiency. A clinical study on 16 patients. Archives of Disease in Childhood *58*, 246–252.

404. Visser, G., Rake, J. P., Kokke, F. T., Nikkels, P. G., Sauer, P. J., and Smit, G. P. (2002). Intestinal function in glycogen storage disease type I. Journal of Inherited Metabolic Disease *25*, 261–267.

405. Milla, P. J., Atherton, D. A., Leonard, J. V., Wolff, O. H., and Lake, B. D. (1978). Disordered intestinal function in glycogen storage disease. Journal of Inherited Metabolic Disease *1*, 155–157.

406. Bali, D. S., Chen, Y. T., and Goldstein, J. L. (1993). Glycogen Storage Disease Type I.

407. Pascual, J. M., Wang, D., Lecumberri, B., Yang, H., Mao, X., Yang, R., and De Vivo, D. C. (2004). GLUT1 deficiency and other glucose transporter diseases. European Journal of Endocrinology *150*, 627–633.

408. Golachowska, M. R., van Dael, C. M., Keuning, H., Karrenbeld, A., Hoekstra, D., Gijsbers, C. F., Benninga, M. A., Rings, E. H., and van Ijzendoorn, S. C. (2012). MYO5B mutations in patients with microvillus inclusion disease presenting with transient renal Fanconi syndrome. Journal of Pediatric Gastroenterology and Nutrition *54*, 491–498.

409. Jacob, R., Zimmer, K. P., Schmitz, J., Naim, H. Y., et al. (2000). Congenital sucrase-isomaltase deficiency arising from cleavage and secretion of a mutant form of the enzyme. Journal of Clinical Investigation *106*, 281–287.

410. Iyer, R. K., Yoo, P. K., Kern, R. M., Rozengurt, N., Tsoa, R., O'Brien, W. E., Yu, H., Grody, W. W., and Cederbaum, S. D. (2002). Mouse model for human arginase deficiency. Molecular and Cellular Biology *22*, 4491–4498.

411. Claes, D. J. and Jackson, E. (2012). Cystinuria: mechanisms and management. Pediatric Nephrology *27*, 2031–2038.

412. Wadman, S. K., Sprang, F. J. V., Stekelenburg, G. J. V., and de Bree, P. K. (1967). Three New Cases of Histidinemia. Zeitschrift fur die Gesamte Innere Medizin und Ihre Grenzgebiete *56*, 485–492.

413. Cathelineau, L., Briand, P., Rabier, D., and Navarro, J. (1985). Ornithine transcarbamylase and disaccharidase activities in damaged intestinal mucosa of children–diagnosis of hereditary ornithine transcarbamylase deficiency in mucosa. Journal of Pediatric Gastroenterology and Nutrition *4*, 960–964.

414. Janne, J., Alhonen, L., Keinanen, T. A., Pietila, M., Uimari, A., Pirinen, E., Hyvonen, M. T., and Jarvinen, A. (2005). Animal disease models generated

by genetic engineering of polyamine metabolism. Journal of Cellular and Molecular Medicine *9*, 865–882.

415. Reinoso, M. A., Whitley, C., Jessurun, J., and Schwarzenberg, S. J. (1998). Lysinuric protein intolerance masquerading as celiac disease: a case report. The Journal of Pediatrics *132*, 153–155.

416. Fairbanks, T. and Emil, S. (2005). Colonic perforation in the first few hours of life associated with rhizomelic chondrodysplasia punctata. Pediatric Surgery International *21*, 662–664.

417. Nimubona, L., Laloum, D., Rolland, M. O., Read, M. H., Guillois, B., and Duhamel, J. F. (2002). An intestinal obstruction in an eight-month-old child suffering from mevalonic aciduria. Acta Paediatrica *91*, 714–716.

418. Hoflack, M., Caruba, C., Pitelet, G., Haas, H., Mas, J. C., Paquis, V., and Berard, E. (2010). Infant coma in the emergency department: 2 cases of MCAD deficiency. Archives de Pediatrie *17*, 1074–1077.

419. Gilbert, J. and Ibdah, J. A. (2005). Intestinal pseudo-obstruction as a manifestation of impaired mitochondrial fatty acid oxidation. Medical Hypotheses *64*, 586–589.

420. Lee, M., Cook, C. R., and Wilkins, I. (2007). A new association of second-trimester echogenic bowel and metabolic disease of the neonate. Journal of Ultrasound in Medicine *26*, 1119–1122.

421. Shinton, N. K. (1972). Vitamin B 12 and folate metabolism. British Medical Journal *1*, 556–559.

422. Klein, R. D., Thorland, E. C., Gonzales, P. R., Beck, P. A., Dykas, D. J., McGrath, J. M., and Bale, A. E. (2006). A multiplex assay for the detection and mapping of complex glycerol kinase deficiency. Clinical chemistry *52*, 1864–1870.

423. Caira, M. R. and Corina, I. (2005). Drug metabolism: current concepts volume 7. (Springer).

424. Russel, F. G. M. (2010). Transporters: Importance in Drug Absorption, Distribution, and Removal. In Pang, K. S., Rodrigues, A. D., and Peter, R. M., eds., Enzyme and Transporter Based Drug-Drug Interactions In Pang, K. S., Rodrigues, A. D., and Peter, R. M., eds., Enzyme and Transporter Based Drug-Drug Interactions. (New York: Springer New York).

425. Harris, R. Z., Jang, G. R., and Tsunoda, S. (2003). Dietary effects on drug metabolism and transport. Clinical Pharmacokinetics *42*, 1071–1088.

426. Walter-Sack, I. and Klotz, U. (1996). Influence of diet and nutritional status on drug metabolism. Clinical Pharmacokinetics *31*, 47–64.

427. Granick, S. (1965). Hepatic porphyria and Drug induced or chemical porphyria. Annals of the New York Academy of Sciences *123*, 188–197.

428. Neale, R., Reynolds, T. M., and Saweirs, W. (2004). Statin precipitated lactic acidosis? Journal of Clinical Pathology *57*, 989–990.

429. Mampilly, G. T., Mampilly, T. K., Christopher, R., Chandramohan, N., and Janaki, V. (2013). Challenges in Diagnosing a Metabolic Disorder Error of Pyruvate Metabolism or Drug Induced? Journal of Child Neurology.

430. Jhajra, S., Varkhede, N. R., Ahire, D. S., Naik, B. V., Prasad, B., Paliwal, J., and Singh, S. (2012). Extrahepatic Drug-Metabolizing Enzymes and Their Significance. Encyclopedia of Drug Metabolism and Interactions.

431. Pagliarini, R. and di Bernardo, D. (2013). A Genome-Scale Modeling Approach to Study Inborn Errors of Liver Metabolism: Toward an In Silico Patient. Journal of Computational Biology *20*, 383–397.

432. Kim, H. U., Kim, S. Y., Jeong, H., Kim, T. Y., Kim, J. J., Choy, H. E., Yi, K. Y., Rhee, J. H., and Lee, S. Y. (2011). Integrative genome-scale metabolic analysis of Vibrio vulnificus for drug targeting and discovery. Molecular Systems Biology *7*.

433. Li, L., Zhou, X., Ching, W.-K., and Wang, P. (2010). Predicting enzyme targets for cancer drugs by profiling human metabolic reactions in NCI-60 cell lines. BMC Bioinformatics *11*, 501.

434. Wang, Y., Eddy, J. A., and Price, N. D. (2012). Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE. BMC Systems Biology *6*, 153.

435. Christians, U., Jacobsen, W., and Floren, L. C. (1998). Metabolism and drug interactions of 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors in transplant patients: are the statins mechanistically similar? Pharmacology & Therapeutics *80*, 1–34.

436. Ishikawa, C., Ozaki, H., Nakajima, T., Ishii, T., Kanai, S., Anjo, S., Shirai,

K., and Inoue, I. (2004). A frameshift variant of CYP2C8 was identified in a patient who suffered from rhabdomyolysis after administration of cerivastatin. Journal of Human Genetics *49*, 582–5.

437. Fujino, H., Saito, T., Tsunenari, Y., Kojima, J., and Sakaeda, T. (2004). Metabolic properties of the acid and lactone forms of HMG-CoA reductase inhibitors. Xenobiotica *34*, 961–971.

438. Sakaeda, T., Fujino, H., Komoto, C., Kakumoto, M., Jin, J.-S., Iwaki, K., Nishiguchi, K., Nakamura, T., Okamura, N., and Okumura, K. (2006). Effects of acid and lactone forms of eight HMG-CoA reductase inhibitors on CYP-mediated metabolism and MDR1-mediated transport. Pharmaceutical Research *23*, 506–12.

439. Hatanaka, T. (2000). Clinical Pharmacokinetics of pravastatin. Clinical Pharmacokinetics *39*, 397–412.

440. Stearns, R. A., Chakravarty, P. K., Chen, R., and Chiu, S. H. (1995). Biotransformation of losartan to its active carboxylic acid metabolite in human liver microsomes. Role of cytochrome P4502C and 3A subfamily members. Drug Metabolism and Disposition *23*, 207–215.

441. Knauf, H. and Mutschler, E. (1998). Clinical Pharmacokinetics and pharmacodynamics of torasemide. Clinical Pharmacokinetics *34*, 1–24.

442. Bahn, A., Ebbinghaus, C., Ebbinghaus, D., Ponimaskin, E. G., Fuzesi, L., Burckhardt, G., and Hagos, Y. (2004). Expression studies and functional characterization of renal human organic anion transporter 1 isoforms. Drug Metabolism and Disposition *32*, 424–430.

443. Christians, U. and Sewing, K. F. (1993). Cyclosporin metabolism in transplant patients. Pharmacology & Therapeutics *57*, 291–345.

444. Fairley, J. A. (1990). Intracellular targets of cyclosporine. Journal of the American Academy of Dermatology *23*, 1329–1334.

445. Thomson, A. W., Bonham, C. A., and Zeevi, A. (1995). Mode of action of tacrolimus (FK506): molecular and cellular mechanisms. Therapeutic Drug Monitoring *17*, 584–591.

446. Christians, U. and Sewing, K. F. (1995). Alternative cyclosporine metabolic pathways and toxicity. Clinical Biochemistry *28*, 547–559.

447. Dai, Y., Hebert, M. F., Isoherranen, N., Davis, C. L., Marsh, C., Shen, D. D., and Thummel, K. E. (2006). Effect of CYP3A5 polymorphism on tacrolimus metabolic clearance in vitro. Drug Metabolism and Disposition *34*, 836–847.

448. Drewe, J., Beglinger, C., and Kissel, T. (1992). The absorption site of cyclosporin in the human gastrointestinal tract. British Journal of Clinical Pharmacology *33*, 39–43.

449. Tamura, S., Tokunaga, Y., Ibuki, R., Amidon, G. L., Sezaki, H., and Yamashita, S. (2003). The site-specific transport and metabolism of tacrolimus in rat small intestine. Journal of Pharmacology and Experimental Therapeutics *306*, 310–316.

450. Rainsford, K. D. (2009). Ibuprofen: pharmacology, efficacy and safety. Inflammopharmacology *17*, 275–342.

451. Hinz, B., Cheremina, O., and Brune, K. (2008). Acetaminophen (paracetamol) is a selective cyclooxygenase-2 inhibitor in man. The FASEB Journal *22*, 383–390.

452. Davies, N. M. (2004). Chirality in Drug Design and Development. (CRC Press).

453. Graham, G. and Williams, K. (2004). Metabolism and Pharmacokinetics of Ibuprofen. In Aspirin and Related Drugs In Aspirin and Related Drugs. (CRC Press).

454. Shieh, W. R. and Chen, C. S. (1993). Purification and characterization of novel 2-arylpropionyl-CoA epimerases from rat liver cytosol and mitochondria. Journal of Biological Chemistry *268*, 3487–3493.

455. Tukey, R. H. and Strassburg, C. P. (2000). Human UDP-glucuronosyltransferases: metabolism, expression, and disease. Annual Review of Pharmacology and Toxicology *40*, 581–616.

456. Chang, S. Y., Li, W., Traeger, S. C., Wang, B., Cui, D., Zhang, H., Wen, B. O., and Rodrigues, A. D. (2008). Confirmation that cytochrome P450 2C8 (CYP2C8) plays a minor role in (S)-(+)-and (R)-(-)-ibuprofen hydroxylation in vitro. Drug Metabolism and Disposition *36*, 2513–2522.

457. Graham, G. and Hicks, M. (2004). Pharmacokinetics and Metabolism of Paracetamol (Acetaminophen). In Aspirin and Related Drugs In Aspirin and Related Drugs. (CRC Press).

458. Coughtrie, M. W., Sharp, S., Tan, T. M., Bamforth, K. J., and Wong, K. P. (1990). Liver-specific expression of paracetamol sulphotransferase. Biochemical Society Transactions *18*, 1209.

459. Court, M. H., Duan, S. X., von Moltke, L. L., Greenblatt, D. J., Patten, C. J., Miners, J. O., and Mackenzie, P. I. (2001). Interindividual variability in acetaminophen glucuronidation by human liver microsomes: identification of relevant acetaminophen UDP-glucuronosyltransferase isoforms. Journal of Pharmacology and Experimental Therapeutics *299*, 998–1006.

460. Dobson, P. D. and Kell, D. B. (2008). Carrier-mediated cellular uptake of pharmaceutical drugs: an exception or the rule? Nature Reviews Drug Discovery *7*, 205–220.

461. Cihlar, T., Lin, D. C., Pritchard, J. B., Fuller, M. D., Mendel, D. B., and Sweet, D. H. (1999). The antiviral nucleotide analogs cidofovir and adefovir are novel substrates for human and rat renal organic anion transporter 1. Molecular Pharmacology *56*, 570–580.

462. Campbell, D. B., Lavielle, R., and Nathan, C. (1991). The mode of action and clinical pharmacology of gliclazide: a review. Diabetes Research and Clinical Practice *14*, S21–S36.

463. Reves, J. d., Fragen, R. J., Vinik, H. R., and Greenblatt, D. J. (1985). Midazolam: pharmacology and uses. Anesthesiology *62*, 310–324.

464. Day, R. O., Graham, G. G., Hicks, M., McLachlan, A. J., Stocker, S. L., and Williams, K. M. (2007). Clinical Pharmacokinetics and pharmacodynamics of allopurinol and oxypurinol. Clinical Pharmacokinetics *46*, 623–644.

465. Elliot, D. J., Suharjono, Lewis, B. C., Gillam, E. M., Birkett, D. J., Gross, A. S., and Miners, J. O. (2007). Identification of the human cytochromes P450 catalysing the rate-limiting pathways of gliclazide elimination. British Journal of Clinical Pharmacology *64*, 450–457.

466. Oida, T., Yoshida, K., Kagemoto, A., Sekine, Y., and Higashijima, T. (1985). The metabolism of gliclazide in man. Xenobiotica *15*, 87–96.

467. Paine, M. F., Shen, D. D., Kunze, K. L., Perkins, J. D., Marsh, C. L., McVicar, J. P., Barr, D. M., Gillies, B. S., and Thummel, K. E. (1996). First-pass metabolism of midazolam by the human intestine. Clinical Pharmacology & Therapeutics *60*, 14–24.

468. Sarkar, A., Tiwari, A., Bhasin, P. S., and Mitra, M. (2011). Pharmacological and Pharmaceutical Profile of Gliclazide: A Review. Journal of Applied Pharmaceutical Science *1*, 11–19.

469. Heizmann, P. and Ziegler, W. H. (1981). Excretion and metabolism of 14C-midazolam in humans following oral dosing. Arzneimittel-forschung *31*, 2220.

470. Arellano, C., Philibert, C., Vachoux, C., Woodley, J., and Houin, G. (2007). The metabolism of midazolam and comparison with other CYP enzyme substrates during intestinal absorption: in vitro studies with rat everted gut sacs. Journal of Pharmacy & Pharmaceutical Sciences *10*, 26–36.

471. Pastor-Anglada, M., Errasti-Murugarren, E., Aymerich, I., and Casado, F. J. (2007). Concentrative nucleoside transporters (CNTs) in epithelia: from absorption to Cellular Signallinging. Journal of Physiology and Biochemistry *63*, 97–110.

472. Jacobsen, W., Kirchner, G., Hallensleben, K., Mancinelli, L., Deters, M., Hackbarth, I., Baner, K., Benet, L. Z., Sewing, K. F., and Christians, U. (1999). Small intestinal metabolism of the 3-hydroxy-3-methylglutaryl-coenzyme A reductase inhibitor lovastatin and comparison with pravastatin. The Journal of Pharmacology and Experimental Therapeutics *291*, 131–9.

473. Tornio, A., Pasanen, M. K., Laitila, J., Neuvonen, P. J., and Backman, J. T. (2005). Comparison of 3-Hydroxy-3-methylglutaryl Coenzyme A (HMG-CoA) Reductase Inhibitors (Statins) as Inhibitors of Cytochrome P450 2C8. Basic & Clinical Pharmacology & Toxicology *97*, 104–108.

474. Prueksaritanont, T., Gorham, L. M., Ma, B., Liu, L., Yu, X., Zhao, J. J., Slaughter, D. E., Arison, B. H., and Vyas, K. P. (1997). In vitro metabolism of simvastatin in humans [SBT]identification of metabolizing enzymes and effect of the drug on hepatic P450s. Drug Metabolism and Disposition *25*, 1191–1199.

475. Chen, C., Mireles, R. J., Campbell, S. D., Lin, J., Mills, J. B., Xu, J. J., and Smolarek, T. A. (2005). Differential interaction of 3-hydroxy-3-methylglutaryl-coa reductase inhibitors with ABCB1, ABCC2, and OATP1B1. Drug Metabolism and Disposition *33*, 537–546.

476. Park, K. S., Park, J. H., and Song, Y. W. (2008). Inhibitory NKG2A and activating NKG2 D and NKG2C natural killer cell receptor genes: susceptibility for rheumatoid arthritis. Tissue Antigens *72*, 342–346.

477. Goosen, T. C., Bauman, J. N., Davis, J. A., Yu, C., Hurst, S. I., Williams, J. A., and Loi, C.-M. (2007). Atorvastatin glucuronidation is minimally and nonselectively inhibited by the fibrates gemfibrozil, fenofibrate, and fenofibric acid. Drug Metabolism and Disposition *35*, 1315–1324.

478. Kivisto, K. T. and Niemi, M. (2007). Influence of drug transporter polymorphisms on pravastatin pharmacokinetics in humans. Pharmaceutical Research *24*, 239–247.

479. Scripture, C. D. and Pieper, J. A. (2001). Clinical Pharmacokinetics of fluvastatin. Clinical Pharmacokinetics *40*, 263–281.

480. Beaird, S. L. (1999). HMG-CoA reductase inhibitors: assessing differences in drug interactions and safety profiles. Journal of the American Pharmaceutical Association (Washington, DC: 1996) *40*, 637–644.

481. Boberg, M., Angerbauer, R., Fey, P., Kanhai, W. K., Karl, W., Kern, A., Ploschke, J., and Radtke, M. (1997). Metabolism of Cerivastatin by Human Liver Microsomes In Vitro Characterization of Primary Metabolic Pathways and of Cytochrome P450 Isozymes involved. Drug Metabolism and Disposition *25*, 321–331.

482. Fujino, H., Yamada, I., Shimada, S., Yoneda, M., and Kojima, J. (2003). Metabolic fate of pitavastatin, a new inhibitor of HMG-CoA reductase: human UDP-glucuronosyltransferase enzymes involved in lactonization. Xenobiotica *33*, 27–41.

483. Duggan, S. T. (2012). Pitavastatin: a review of its use in the management of hypercholesterolaemia or mixed dyslipidaemia. Drugs *72*, 565–584.

484. White, C. M. (2002). A review of the pharmacologic and pharmacokinetic aspects of rosuvastatin. The Journal of Clinical Pharmacology *42*, 963–970.

485. Bolego, C., Poli, A., Cignarella, A., Catapano, A. L., and Paoletti, R. (2002). Novel statins: pharmacological and clinical results. Cardiovascular Drugs and Therapy *16*, 251–257.

486. Sica, D. A., Gehr, T. W. B., and Ghosh, S. (2005). Clinical Pharmacokinetics of losartan. Clinical Pharmacokinetics *44*, 797–814.

487. Krieter, P. A., Colletti, A. E., Miller, R. R., and Stearns, R. A. (1995). Absorption and glucuronidation of the angiotensin II receptor antagonist losartan by the rat intestine. Journal of Pharmacology and Experimental Therapeutics

273, 816–822.

488. Vormfelde, S. V., Schirmer, M., Hagos, Y., Toliat, M. R., Engelhardt, S., Meineke, I., Burckhardt, G., Nurnberg, P., and Brockmoller, J. (2006). Torsemide renal clearance and genetic variation in luminal and basolateral organic anion transporters. British Journal of Clinical Pharmacology *62*, 323–335.

489. Zhang, Q. Y., Kaminsky, L. S., Dunbar, D., Zhang, J., and Ding, X. (2007). Role of small intestinal cytochromes P450 in the bioavailability of oral nifedipine. Drug Metabolism and Disposition *35*, 1617–1623.

490. Patki, K. C., von Moltke, L. L., and Greenblatt, D. J. (2003). In vitro metabolism of midazolam, triazolam, nifedipine, and testosterone by human liver microsomes and recombinant cytochromes p450: role of cyp3a4 and cyp3a5. Drug Metabolism and Disposition *31*, 938–944.

491. Hamman, M. A., Thompson, G. A., and Hall, S. D. (1997). Regioselective and stereoselective metabolism of ibuprofen by human cytochrome P450 2C. Biochemical Pharmacology *54*, 33–41.

492. Kalsi, S. S., Wood, D. M., Waring, W. S., and Dargan, P. I. (2011). Does cytochrome P450 liver isoenzyme induction increase the risk of liver toxicity after paracetamol overdose? Open Access Emergency Medicine *3*, 69–76.

493. USDA: Agricultural research service, National agricultural library, Nutrient data laboratory. (2012). USDA National Nutrient Database for Standard Reference Release 25.

494. Mendoza, J. A., Drewnowski, A., and Christakis, D. A. (2007). Dietary energy density is associated with obesity and the metabolic syndrome in US adults. Diabetes Care *30*, 974–979.

495. Foster, B. C., Foster, M. S., Vandenhoek, S., Krantis, A., Budzinski, J. W., Arnason, J. T., Gallicano, K. D., and Choudri, S. (2001). An in vitro evaluation of human cytochrome P450 3A4 and P-glycoprotein inhibition by garlic. Journal of Pharmacy & Pharmaceutical Sciences *4*, 176–184.

496. Goasduff, T., Dreano, Y., Guillois, B., Menez, J. F., and Berthou, F. (1996). Induction of liver and kidney CYP1A1/1A2 by caffeine in rat. Biochemical Pharmacology *52*, 1915–1919.

497. Spahn-Langguth, H. and Langguth, P. (2001). Grapefruit juice enhances in-

testinal absorption of the P-glycoprotein substrate talinolol. European Journal of Pharmaceutical Sciences *12*, 361–367.

498. Saad, R., Rizkallah, M. R., and Aziz, R. K. (2012). Gut pharmacomicrobiomics: the tip of an iceberg of complex interactions between drugs and gut-associated microbes. Gut Pathogens *4*, 16.

499. Gregus, Z., Kim, H. J., Madhu, C., Liu, Y., Rozman, P., and Klaassen, C. D. (1994). Sulfation of acetaminophen and acetaminophen-induced alterations in sulfate and 3'-phosphoadenosine 5'-phosphosulfate homeostasis in rats with deficient dietary intake of sulfur. Drug Metabolism and Disposition *22*, 725–730.

500. Price, V. F. and Jollow, D. J. (1989). Effects of sulfur-amino acid-deficient diets on acetaminophen metabolism and hepatotoxicity in rats. Toxicology and Applied Pharmacology *101*, 356–369.

501. Lampe, J. W., King, I. B., Li, S., Grate, M. T., Barale, K. V., Chen, C., Feng, Z., and Potter, J. D. (2000). Brassica vegetables increase and apiaceous vegetables decrease cytochrome P450 1A2 activity in humans: changes in caffeine metabolite ratios in response to controlled vegetable diets. Carcinogenesis *21*, 1157–1162.

502. Krishnaswamy, K., Kalamegham, R., and Naidu, N. A. (1984). Dietary influences on the kinetics of antipyrine and aminopyrine in human subjects. British Journal of Clinical Pharmacology *17*, 139–146.

503. Paoletti, R., Corsini, A., and Bellosta, S. (2002). Pharmacological interactions of statins. Atherosclerosis Supplements *3*, 35–40.

504. Maltz, H. C., Balog, D. L., and Cheigh, J. S. (1999). Rhabdomyolysis associated with concomitant use of atorvastatin and cyclosporine. The Annals of Pharmacotherapy *33*, 1176–1179.

505. Chin, C., Gamberg, P., Miller, J., Luikart, H., and Bernstein, D. (2002). Efficacy and safety of atorvastatin after pediatric heart transplantation. The Journal of Heart and Lung Transplantation *21*, 1213–1217.

506. Renders, L., Mayer-Kadner, I., Koch, C., Scharffe, S., Burkhardt, K., Veelken, R., Schmieder, R. E., and Hauser, I. A. (2001). Efficacy and drug interactions of the new HMG-CoA reductase inhibitors cerivastatin and atorvastatin in CsA-treated renal transplant recipients. Nephrology Dialysis Transplantation *16*, 141–146.

507. Vladutiu, G. D., Simmons, Z., Isackson, P. J., Tarnopolsky, M., Peltier, W. L., Barboi, A. C., Sripathi, N., Wortmann, R. L., and Phillips, P. S. (2006). Genetic risk factors associated with lipid-lowering drug-induced myopathies. Muscle & Nerve *34*, 153–162.

508. Baker, S. K. and Samjoo, I. A. (2008). A neuromuscular approach to statin-related myotoxicity. The Canadian Journal of Neurological Sciences *35*, 8–21.

509. Vladutiu, G. D. (2008). Genetic predisposition to statin myopathy. Current Opinion in Rheumatology *20*, 648–655.

510. Coleman, M. D. (2010). Human drug metabolism: an Introduction. (Wiley Online Library).

511. Parikh, S., Saneto, R., Falk, M. J., Anselm, I., Cohen, B. H., and Haas, R. (2009). A modern approach to the treatment of mitochondrial disease. Current Treatment Options in Neurology *11*, 414–430.

512. Littarru, G. P. and Langsjoen, P. (2007). Coenzyme q10 and statins: Biochemical and clinical implications. Mitochondrion *7*, S168–S174.

513. Deichmann, R., Lavie, C., and Andrews, S. (2010). Coenzyme q10 and statin-induced mitochondrial dysfunction. The Ochsner Journal *10*, 16–21.

514. Whirl-Carrillo, M., McDonagh, E. M., Hebert, J. M., Gong, L., Sangkuhl, K., Thorn, C. F., Altman, R. B., and Klein, T. E. (2012). Pharmacogenomics knowledge for personalized medicine. Clinical Pharmacology & Therapeutics *92*, 414–417.

515. Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., et al. (2011). DrugBank 3.0: a comprehensive resource for omics research on drugs. Nucleic Acids Research *39*, D1035–D1041.

516. Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., et al. (2007). HMDB: the human metabolome database. Nucleic Acids Research *35*, D521–D526.

517. Hastings, J., de Matos, P., Dekker, A., Ennis, M., Harsha, B., Kale, N., Muthukrishnan, V., Owen, G., Turner, S., Williams, M., et al. (2013). The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. Nucleic Acids Research *41*, D456–D463.

518. Bolton, E. E., Wang, Y., Thiessen, P. A., and Bryant, S. H. (2008). PubChem:

integrated platform of small molecules and biological activities. Annual Reports in Computational Chemistry *4*, 217–241.

519. Apweiler, R., O'onovan, C., Magrane, M., Alam-Faruque, Y., Antunes, R., Bely, B., Bingley, M., Bower, L., Bursteinas, B., Chavali, G., et al. (2012). Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Research *40*, D71–5.

520. Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002). Biochemistry : Gluconeogenesis and Glycolysis Are Reciproally Regulated. (New York: Freeman, W. H.) 5th edition.

521. Covert, M. W., Knight, E. M., Reed, J. L., Herrgard, M. J., and Palsson, B. O. (2004). Integrating high-throughput and computational data elucidates bacterial networks. Nature *429*, 92–96.

522. Herrgard, M. J., Lee, B.-S., Portnoy, V., and Palsson, B. O. (2006). Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in Saccharomyces cerevisiae. Genome Research *16*, 627–635.

523. Thiele, I., Fleming, R. M. T., Que, R., Bordbar, A., Diep, D., and Palsson, B. O. (2012). Multiscale modeling of metabolism and macromolecular synthesis in E. colI. and its application to the evolution of codon usage. PloS ONE *7*, e45635.

524. Barua, D., Kim, J., and Reed, J. L. (2010). An automated phenotype-driven approach (geneforce) for refining metabolic and regulatory models. PLoS computational biology *6*, e1000970.

525. Simeonidis, E., Chandrasekaran, S., and Price, N. D. (2013). A Guide to Integrating Transcriptional Regulatory and Metabolic Networks Using PROM (Probabilistic Regulation of Metabolism). In Systems Metabolic Engineering In Systems Metabolic Engineering. (Springer).

526. Mahadevan, R. and Edwards, J. S. and Doyle, F. J. 3rd. (2002). Dynamic Flux Balance Analysis of Diauxic Growth in Escherichia coli. Biophysical Journal *83*, 1331–1340.

527. Yachie-Kinoshita, A., Nishino, T., Shimo, H., Suematsu, M., and Tomita, M. (2010). A metabolic model of human erythrocytes: practical application of the E-Cell Simulation Environment. BioMed Research International *2010*.

528. Higuera, G., Schop, D., Janssen, F., van Dijkhuizen-Radersma, R., van Box-

tel, T., and van Blitterswijk, C. A. (2009). Quantifying in vitro growth and metabolism kinetics of human mesenchymal stem cells using a mathematical model. Tissue Engineering Part A *15*, 2653–2663.

529. Chakrabarti, A., Miskovic, L., Soh, K. C., and Hatzimanikatis, V. (2013). Towards kinetic modeling of genome-scale metabolic networks without sacrificing stoichiometric, thermodynamic and physiological constraints. Biotechnology Journal.

List of Publications

- Sahoo, S., Franzson, L., Jonsson, J. J. and Thiele, I. (2012) A compendium of inborn errors of metabolism mapped onto the human metabolic network. Molecular BioSystems, 8, 2545-2558

- Sahoo, S. and Thiele, I. (2013) Predicting the impact of diet and enzymopathies on human small intestinal epithelial cells. Human Molecular Genetics, 22, 2705-2722

- Heinken, A., Sahoo, S., Fleming, R. M., Thiele, I. (2013) Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. Gut Microbes, 4, 28-40.

- Thiele, I., Swainston, N., Fleming, R. M., Hoppe, A., Sahoo, S., Aurich, M. K., Haraldsdottir, H., Mo, M. L., Rolfsson, O., Stobbe, M. D., et al. A community-driven global reconstruction of human metabolism. (2013) Nature Biotechnology, 31, 419-425.