Norwegian University
of Life Sciences

Master's Thesis 2018 60 ECTS
Fakultet for biovitenskap
Jon Olav Vik

Reconstruction and modelling of amino acid synthesis in *Salmo salar*

Filip Rotnes
Chemistry and Biotechnology
Fakultet for biovitenskap / Fakultet for kjemi, bioteknologi og matvitenskap

Aknowledgements The work behind this thesis has been exciting, challenging and filled with new experiences. I would like to thank my supervisors Jon Olav Vik and Maksim Zakhartsev for their support, patience and dedication. Thanks to Arne Bjørke Gjuvsland for getting me started on this adventure, to Liv Torunn Mydland for sharing her extensive knowledge about salmon, thanks to Fabian Grammes and Yang Jin for inspiring me. Being part of the DigiSal project has been a joy, and sparked my interest in science, especially metabolic models. Special thanks also to Amalie Nordberg Juel for motivational support.

Filip Rotnes

Reconstruction and modelling of amino acid Synthesis in Salmo salar

Filip Rotnes

May 15, 2018

Abstract

This thesis documents the process behind the reconstruction of the Salmo Salar Amino Acid Synthesis model (SalSAA). The model is a mathematical description of amino acid biosynthesis and protein polymerisation in Atlantic salmon. Flux balance analysis was performed to study the effect of the amino acid composition of fish meal versus soyben meal on protein polymerisation. As indicated by earlier reseach, fish meal has a composition of amino acids promoting higher growth rate for protein than soybean meal. The model also correctly predicts the most beneficial amino acid additions to soybean meal to promote growth.

Sammendrag

Dette dokumentet dokumenterer prosessen bak rekonstruksjonen av den støkiometriske modellen for aminosyresyntese og proteinpolymerisering i atlanterhavslaks (SalSAA). Modellen er en matematisk beskrivelse av aminosyresyntese og proteinpolymerisering i Atlanterhavslaks. Det ble gjort simuleringer for å studere effekten av protein fra fiskemel og soyamel på proteinsyntese. Tidligere forskning tilsier at fiskemel har en komposisjon av aminosyrer som fremmer vekst bedre enn soyamel, hvilket ble bekreftet av modellen. Simuleringene forutså også korrekt hvilke aminosyretilsetninger som forbedrer vekst på soyamel best.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

The aquaculture industry is expanding, now providing close to half of the fish eaten in the world [3]. Consumption of fish instead of meat is indicated to be beneficial for cardiovascular health. An increase of fish availbility may therefore contribute to better health world wide. However, as the industry is growing, so is it's demand for resources. With diminishing resources in the oceans, parts of feed for aquaculture are now based on plants. This leads to competition for resources with other types of food production, and complicates the process of formulating feeds with suitable nutritional composition. In order to meet nutritional needs while maintaining ecological and economical sustainability, formulas for aquaculture feeds should be adaptable to the availability of ingredients. Mathematical tools for modelling the effects of different feed formulations may move the way feeds are formulated from a reactive to a predictive way. Constraint based stoichiometric modelling is such a way of mathematical exploration. A benefit of this method is it's ability to make predictions based on the network of metabolic reactions without the need of kinetic data for every single reaction.

1.2 Outline of problem

This thesis documents the process behind the reconstruction of SalsAA – the **Sal***mo* **s***alar* **A**mino **A**cid synthesis model – a stoichiometric model of the import and biosynthesis of amino acids and their consumption from protein polymerization for an average salmon cell. The model is intended to comprise the reactions and metabolites most central in amino acid synthesis and import, and is therefore encompassed by an artificial boundary, over which metabolites flow freely, but import of amino acids is taken from the apical membrane of the intestinal epithelium.

The model should be capable of protein polymerization under steady state, a state where all reactions happen at a constant rate and every metabolite is consumed at the same rate as it is produced, meaning no accumulation of metabolites. Amino acid composition of two foods will be used as constraints for flux balance analysis, and further investigation will be done to determine which amino acids supplements would be beneficial. Salmon are known to grow faster when fed marine food sources such as fish and crustaceans than plant based food sources such as soybeans. Certain amino acids are shown to further promote
The efficiency of plant based feeds are shown to be higher when certain limiting amino acids are added. This will be used to validate the model.

Assumptions/scope of model

1. Proteins are fully digested to amino acids in the gut
2. Amino acids can be fully absorbed by the gut
3. Absorbed amino acids are prioritized for protein synthesis
4. Glycolysis and TCA are not limiting factors
5. Energy (ATP) and redox potential are not limiting factors

1.3 Fish farming

The aquaculture industry is in rapid growth, [8] is a significant part of the global food market and one of Norway's most important exports. The transformation of salmon from a luxury food to an everyday food item has made omega 3-fatty acids more available. The omega3-fatty acids EPA and DHA are highly represented in fatty fishes like herring, mackerel and salmon. Substitution of meat in the diet with fish replaces saturated fat in favor of unsaturated fat, as advised by the World Health Organization [20], and is considered to be beneficial to cardiovascular health.

Salmon is a carnivorous fish, largely feeding on crustaceans and fishes. The earliest feeds for salmon farming were therefore based on these ingredients. However, with an expanding industry and diminishing fish stocks, increasing the consumption of fish based feeds is not sustainable, environmentally nor economically. Protein is the most expensive part of salmon feed [2], and the processes involved in amino acid synthesis are well known, so this is a good starting point for metabolic modelling.

Plant based fish meals demand less of the world's resources, but there are challenges of changing the diet of a carnivore to vegetarian. Plants are composed of a higher amount of starch, different amino acid ratios and different fats. About 75% of protein in modern fish fees are derived from plants. These feeds are less effective than the fish derived feeds, but addition of extra amino acids to the plant based diet has shown to raise efficiency [12]. Research on feedstuffs that are not competing with human food is being done, animal plankton [5], yeast grown on substrate from trees [16], fly larvae grown on marine seaweed [14]. Atlantic salmon can grow just as well on other feeds than those it has evolved for, if they are balanced correctly.

Table 1: Total value of slaughtered fish from Norwegian aquaculture industry in millions NOK. Based on report from Directory of fisheries [8]

2016	2015	2014	2013	2012	2011	2010	2009	2008	2007
64 014	46 834	44 319	40 466	30 028	28 926	30 615	22 443	17 447	17 509

1.4 Mathematical models

Mathematical models are abstractions of reality designed to explain specific phenomena. A mathematical model is not all explanatory, but is a useful tool for prediction. When the model has been shown to predict correct results for known conditions, the model can be used to predict outcomes of conditions not earlier investigated. This can be used to find which trials would be infeasible and to predict the outcomes of feasible ones, generating testable hypotheses. A highly trusted computational model may even be used in place of physical experiments (in silico experiments). In the perspective of salmon feed production, it could for instance be used to compose the most suitable feed based on ingredients availability of and price fluctuations.
Models describe systems within boundaries (the thermodynamic systems):

- isolated systems – exchanges nothing with it's environment
- closed systems – exchange energy with it's environment
- open systems – exchange energy and matter with environment.

Biological systems are open systems, and can be considered to be parts of larger systems – ecosystems, in turn open systems. The perspective and boundaries of the model depends on the goal.

1.5 Metabolic modelling

Biological organisms are complex systems, and the differences between them is the product of their genes and the environment surrounding them. Scientists have gained large insight to the components and their interactions through molecular biology. However, many traits are the result of multiple factors, and as metabolites often take part in several processes, the connectivity between the different parts of metabolism is high.

Genes encoding enzymes are responsible for the reactions an organism can perform. Combined, these form the pathways, which again forms the metabolism of the organism. The types of metabolic reactions in an organism greatly outnumber the species of participating metabolites. With high connectivity follows that the balance can be skewed by small changes in the network, as silencing of a gene, or a change in available metabolites. Systems biology embraces the possibilities of modern sequencing technologies as well as the extensive knowledge bases that are available today to study biological systems through an integrative approach.

The stoichiometric network representation emphasizes the reactions the organism of study is capable of performing, and makes the network possible to study without knowledge of the activity and regulation of every enzyme involvedl The properties of orthologs may vary between species, but the stoichiometry of the reactions they catalyze is conserved. This makes stoichiometric models powerful and the functions transferable between models for organisms with orthologous genes. Knowledge of enzyme activity and regulation of genes in the organism of study may be applied as constraints to refine the model.

1.5.1 Genome scale reconstructions

The connectivity of the metabolic network calls for representation of the system on a global scale. An in-depth protocol for genome scale reconstruction has been published by Thiele and Palsson [18].

1.5.2 Multiple compartments

Modelling eukaryotic systems is more complex than prokaryotes. In addition to larger genomes, they also consist of several comparments, each with different functions. A model of an eukaryotic cell therefore needs to take transport of compounds between compartments into account. Multicellular and multi tissue organisms lead to further complications.

1.5.3 Partial models

Subsets of the reaction networks may also be of interest to study which genes are central to perform certain metabolic tasks. To build such an 'imaginary compartment', the parts of the network to be omitted may be replaced by transport reactions, to create an artificial boundary. Such models are simpler to build, and can provide insight to subsets of the larger network.

1.6 Types of reactions

Metabolic reactions come in different flavors:

$$\begin{array}{rcll} A & \rightarrow & B & isomerization \\ A & \rightarrow & B+C & dissociation \\ A+B & \rightarrow & C & addition \\ A+B & \rightarrow & C+D & metasthesis \end{array}$$

Transport reactions are necessary for a balanced system:

$$\begin{array}{rcll} A_1 & \rightarrow & A_2 & compartmental\ \ transport \\ A & \rightarrow & 0 & export \\ 0 & \rightarrow & A & import \end{array}$$

The concentration of compounds external to the system is not taken into account of the model (limited amounts of external compounds can be reflected by constraints to the system) Therefore, mass balance, etc may be disregarded in terms of transport to and from the external compartment. Transport between the compartments that are part of the model does need to be balanced for biological relevance, but from the perspective of the compartments, the concept is the same.

1.6.1 Transformers

Transformer reactions represent chemical reactions; isomerization, dissociation, addition and metathesis reactions.
The above transformer reactions can be represented mathematically by vectors collected in a matrix:

$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} r_{iso} & r_{dis} & r_{add} & r_{met} \end{array} \\ \begin{bmatrix} -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Where each column is a vector representing a reaction. There are several prerequsites such a vector need to fulfill in order to be a valid representation of a chemical reaction within the model. It needs to be balanced in terms of stoichiometry and charge.

1.6.2 Transporters

Cells consume compounds in their environment and export compounds that are in excess. In a multi compartment model there are two types of transport reaction. Between compartments and across the boundaries of the system. The latter can be divided into import and export reactions and is required in order to achieve a balanced system, and the prior in order to achieve balanced compartments. The compartmental transport can be treated as a transformation between two identical species in different compartments. As only the inside of the system is of interest for the simulation, the boundary transport is represented by consumption or generation of the species in question, and elemental and charge balance is abandoned.

$$\begin{array}{c} \\ A_c \\ A_m \end{array} \begin{array}{c} \begin{array}{ccc} T_{internal} & T_{export} & T_{import} \end{array} \\ \begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{array}$$

1.6.3 Polymerization

The multitude of different proteins present in an organism is vast, and proteins and other polymers are only interesting for the model in one respect, it functions as a sink for metabolites. It would be infeasible to include the stoichiometry of all possible polymerization reactions, therefore, the protein polymerization of the model is creating only one, average protein. As gene expression differs between tissue types and under different conditions, it would be sensible to take quantitative gene expression data into account for each case.

1.7 Linear algebra

Linear algebra is the branch of mathematics where properties of matrices, vector spaces and transformations between these are studied. A system of linear equations can be structured as a matrix, and the *space of solutions* is the possible combinations of the variables. Considering *m* to be the number of equations and *n* the number of variables between the equations, the matrix has *dimensions* $m \times n$. The number of independent variables is referred to as the *rank*. A *basis* for a space is the smallest set of vectors (equations) that span a solution space. Bases may appear in different forms, the typical for computational applications being ortho-normal basis vectors, unitary vectors orthogonal to each other. In description of metabolic networks, a *convex basis*, consisting of positive combinations of positive variables, is more interpretable. The *column space* is the solution space for a set of equations where each column corresponds to a variable in it's corresponding row. The *null space* is the space of solutions yielding 0 as the answer. Stoichiometric models result in large sets of linear equations, often composed of a larger number of variables than equations, meaning the solution is not a single number, but a *space* of solutions. Optimization algorithms such as linear programming may be applied to explore such spaces.

2 Methods

2.1 Tools

2.1.1 KEGG pathway maps

The Kyoto Encyclopedia of Genes and Genomes (KEGG) [11] contains reference maps for a range of metabolic pathways. Individual maps are generated for various organisms based on their genomes.

KEGG metabolic pathways are generalized wiring diagrams for well known biological processes. Based on the notion that metabolism is highly conserved across all domains of life, the same diagrams are used to display the metabolic capabilities of the organisms represented in the database. These *reference pathways* consist of enzymes, represented by their EC[1] numbers, and the metabolites that are further modified in the pathway of each reaction. Although orderly, this representation does not display the stoichiometry of the reaction and lacks participating metabolites. Each node is linked to the KEGG Reaction database entry for the reaction, where stoichiometry is accounted for.

KEGG Reaction entries are linked to the KEGG Compound database, with composition of metabolites. The specificity of metabolic reactions varies, and some of the compounds are generalizations of polymers or classes of compounds. In a modelling context, the composition of compounds must be defined precisely in terms of atomic composition, charge and mass.

The organism specific pathway diagrams are colored according to which enzymes are identified in the genome. The maps thus highlight which of the predefined paths are available for the organism. Organism specific pathway diagrams are accounted for by KGML[2]-files, specifying which enzyme nodes to highlight, the KEGG reaction identifier and what genes are associated in this organism.

KEGG pathway diagrams differ from models in several manners. They do indicate connections of large parts of the network, but do not account for the full connectivity of the system. Boundaries is an integral part of a constraint based model, but are not accounted for in the pw diagrams. The representation from KEGG does not ensure balance in terms of mass and charge.

[1] Enzyme Commision

[2] KeGg Markup Language, an XML format containing the species specific part of a KEGG pathway diagram

KEGG maps are very clear, perhaps clearer than the phenomena they illustrate. In many cases, in most cases, only one reactant and one product is shown in the reactions of the charts, which may lead to omission of reactions that should have been included.

2.1.2 Databases and tools for reference

MetaCyc contains databases for several organisms. The yeast database is used as reference for directionality and charge.

Pathway tools is used to access the MetaCyc database (part of the BioCyc collection) to use reactions from S.cerevisae as reference for directionality.

Virtual Metabolic Human is a database containing the reactions and metabolites of the Recon 3D [7] human reconstruction. This extensive database is used mainly as a reference for directionality of the reactions it has in common with this model.

2.1.3 InSilico Discovery

Is a computational tool for building cellular models and in silico experiments. In addition to powerful capabilities for analysis, it conveniently provides a visual approach to model construction. This overview allows for an intuitive process, and the built-in functions for identification of dead ends, unused reactions and parallel sequences of reactions is convenient.

2.2 Collection of data

Reconstruction of a metabolic network for an organism may be approached in several ways. The process can be based on a previous reconstruction for a related species, an assembly of several such networks or based entirely on literature.

Lien et al. published the genome of Salmo salar in 2016 [13]. The genome is represented in the KEGG pathway database, so this database is used as the main reference for this reconstruction process. Reactions in KEGG maps are only displaying the 'main' reactant and product, so in some cases, the part interest of a reactions may be hidden.

182 reactions were collected based on their presence in the amino acid metabolism pathways in KEGG. In addition to stoichiometrical and charge balanced chemical equations, EC numbers, KEGG Reaction identifier, NCBI gene-entries and (when stated) compartment were collected. All reactions were compared to the metacyc database for directionality for the reactions represented in metacyc.

2.3 Assembly of network

A plethora of tools for reconstruction of metabolic networks exist, and the reconstruction process can be performed either manually or partially automatic. The Salmo salar Amino Acid synthesis network was assembled manually to only include reactions directly involved in amino acid biosynthesis.

The boundaries of the network are essential to take into account in order to convert the reconstructed network into a mathematical model.
As the sole purpose of the SalsAA model is to model the amino acid synthesis and protein polymerization, pathways like glycolysis, and the TCA cycle are not included although they closely connect to amino acid synthesis.
This means the model can be considered to be of an imaginary compartment with part of it's membrane in contact with the exterior of the cell, part with cytosol and a part sectioning mitochondria. Transport of compounds other than amino acids is not to be taken in account by this reconstruction, so other metabolites are free to diffuse over the membranes if necessary.

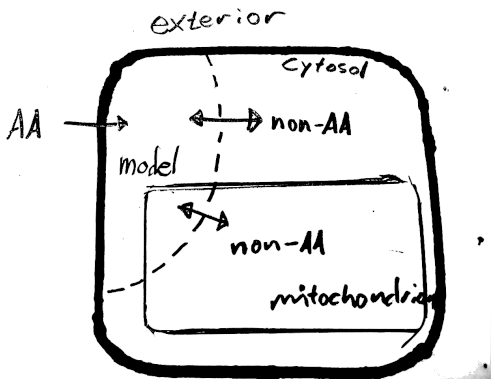


Figure 1: The artificial compartment is in contact with cytosol, exterior and mitochondria

The high connectivity of the network means it is hard to choose a subset of reactions. Under steady state, each metabolite in the network is either produced and consumed at the same rate or not produced or consumed at all. From this follows that a compound that is produced by a reaction but not further degraded or exported will force the reaction it participates in to a zero flux. These metabolites are referred to as dead-end metabolites.
The reactions are based on the end and starting points of the system. The end point of the system is the protein polymerization reaction. Some obvious starting points are the import of amino acids, but the roots of the biosynthesis of amino acids opens for several possibilities. Most likely, most of the reactions could be considered to affect each other. The roots chosen for this network are 3-phosphohydroxypyruvate(3PG), oxaloacetate, 2-oxoglutarate and pyruvate, as these are common precursors for most of biosynthesized amino acids except those descended from essential amino acids. Reactions connecting 3PG and pyruvate to amino acids are included in the network, and metabolites needed for these reactions to happen are provided to the network by virtual transport reactions. These 'virtual reactions' correct for the limited scope of the network, and allows for treating the network as a cell or compartment.

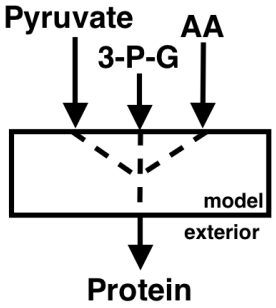


Figure 2: The reactions are chosen to connect the end points of the network, here illustrated by four of the end points

2.4 Protein composition

The composition of protein is important to take into account in order to get a realistic consumption of amino acids, as the role of the protein polymerization in the model is as a metabolite sink. As proteins are the catalysts of metabolic reactions, the protein composition of a cell is closely connected to the metabolic state of the cell at any time point. Proteins come in different sizes and with different composition of amino acids, so the consumption of amino acids may vary quantitatively and qualitatively between different tissues and cell types. To model this realistically, transcriptomics (or even proteomics) should be taken in account to find the balance and magnitude of amino acid consumption by the protein polymerization reaction.To avoid having to describe the stoichiometry of each synthesized protein, the polymerization reaction is based on statistical data. This may be adapted based on the tissue or cell type the model is used to study.
The protein polymerization was based on amino acid content for Atlantic salmon from Wilson and Cowey [19], as displayed in table 2. Techniques for analysis of amino acid composition do not differentiate between asparagine and aspartate or glutamine and glutamate. With no other indications of the balance between these pairs, asparagine and aspartate were assumed to be represented equally by weight in salmon protein, and vice versa for glutamine and glutamate. The protein polymerization reaction is illustrated in figure 3.

Table 2: Amino acid composition for whole body Atlantic salmon, grams per 100 grams protein. [19]

AA	$\frac{g}{g\ Protein}$
Ala	6.52
Arg	6.61
Asn/Asp	9.92
Cys	0.95
Gln/Glu	14.31
Gly	7.41
His	3.02
Ile	4.41
Leu	7.72
Lys	9.28
Met	1.83
Phe	4.36
Pro	4.64
Ser	4.61
Thr	4.95
Trp	0.93
Tyr	3.50
Val	5.09

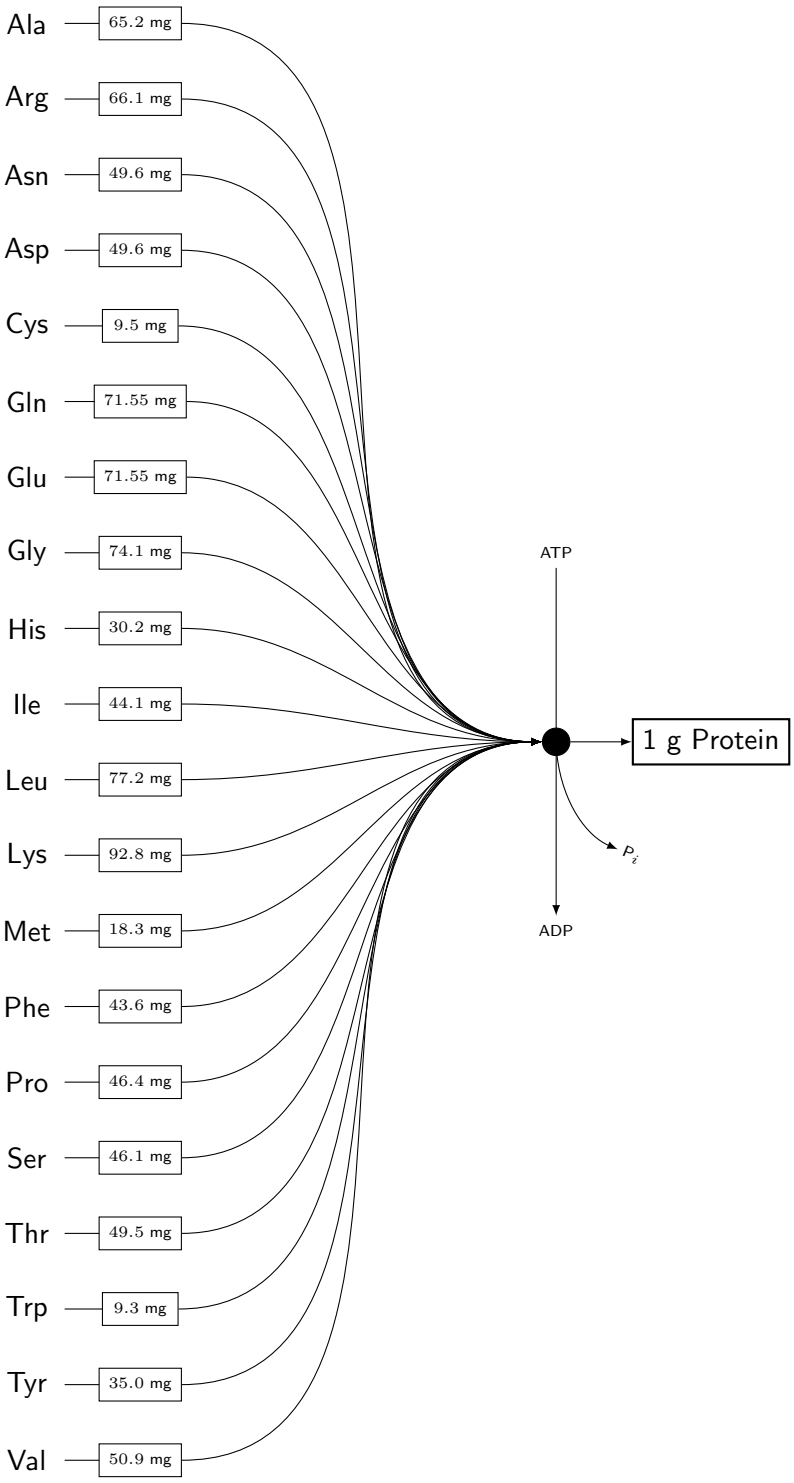


Figure 3: Protein polymerization reaction.

2.5 Transport reactions

The boundaries of the network

Amino acids are absorbed from the lumen of the gastrointestinal tract by the epithelial cells over the apical membrane and exported to the bloodstream over the basolateral membrane. Similarly to glucose transport, this import is largely in cotransport with sodium, thereby driven by the sodium export of the sodium potassium pump. Since the SalsAA model is a representation an artificial compartment of an average cell, internal transport reactions cancel out, and only interactions with the environment are taken into account. The transport reactions considered here are therefore only over the apical membrane of intestinal epithelial cells.The mechanisms of transport reactions come in different flavors. In figure 5, the three basic transport mechanisms are exemplified by uniport of AA, AA in symport with Na^+ and antiport of two AAs. Symport and antiport may also occur in combination. The following sections describe the import of amino acids over the epithelial membrane in intestine. The charge of the amino acids refer to the charge of the amino acid side chain. Amino acids were assigned to the three categories of transport based on their charge according to Bröer [6].

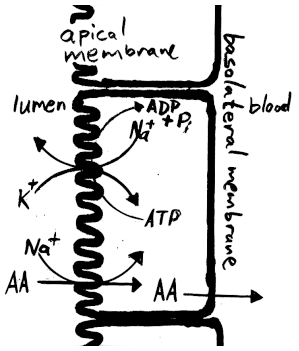


Figure 4: Amino acid transport in epithelial cells is mostly driven by the sodium potassium pump

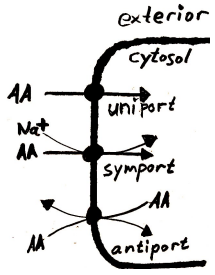


Figure 5: Three classes of transport mechanisms

2.5.1 Which reactions are relevant for the model + simplifications

Some transport reactions exchange amino acids for amino acids by an antiport mechanism.

$$\begin{matrix} & {\scriptstyle T_{antiport}} \\ {\scriptstyle aa_a^0} & \begin{bmatrix} 1 \end{bmatrix} \\ {\scriptstyle aa_b^0} & \begin{bmatrix} -1 \end{bmatrix} \end{matrix}$$

This does not change the net content of AAs, but the proportions between the individual types In the case that aa_a^0 and aa_b^0 have the same import mechanism, an antiport reaction of this form does not affect the balance of ions:

$$\begin{matrix} & {\scriptstyle T_a} & {\scriptstyle T_b} \\ {\scriptstyle aa_a^0} & 1 & 0 \\ {\scriptstyle aa_b^0} & 0 & 1 \\ {\scriptstyle Na^+} & 1 & 1 \end{matrix} \rightarrow \begin{matrix} {\scriptstyle T_a + T_{antiport}} \\ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \end{matrix} = T_b$$

However, if the antiport exchanges a biosynthesized amino acid for one that otherwise would be imported in symport with an ion, this does make a difference. Thus, antiport reactions should be included either for all combinations of amino acids that may be exchanged or carefully picked for the cases where this makes a difference.

2.5.2 Neutral transporters

According to Bröer [6], neutral amino acid transport over the apical membrane is performed by SLC38A2, SLC6A191 and SLC1A5.

SLC38A2 The Sodium-coupled neutral amino acid transporter 2 is the main transporter for neutral amino acids and is coded for by the SLC38A2 gene.

SLC6A19 is a sodium-dependent neutral amino acid transporter also referred to as B(0)AT1 and coded for by the gene SLC6A19.

SLC1A5 is a high-affinity glutamate and neutral amino acid transporter. SLC1A5 is Na^+ dependent and exchanges neutral amino acids against each other through an antiport reaction, possibly useful to adjust the balance of neutral amino acids. It transports alanine, cysteine, glycine, serine and threonine. All of these are neutral, and are otherwise transported the same way, threonine and, under some conditions, cysteine and glycine are essential amino acids, so exchange reactions between these and alanine and serine might be beneficial to add. According to Bröer,SLC38A2 and SLC6A19 both catalyze a symport reaction of one neutral amino acid and one Na^+. Hence, the main transport of neutral amino acids is considered to be through a symport reaction:

$$\begin{matrix} & T^0 \\ aa_C^0 & \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ Na_C^+ & \end{matrix}$$

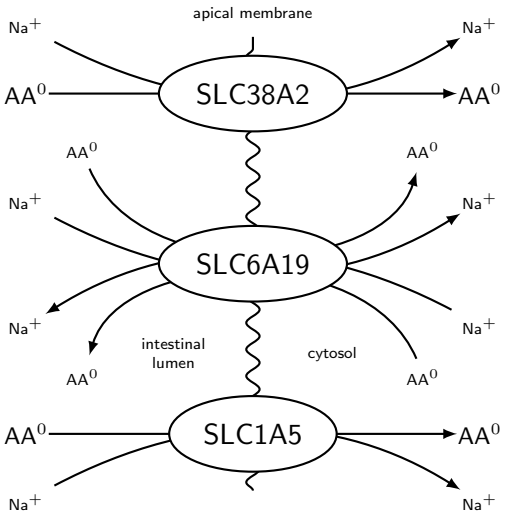


Figure 6: Neutral amino acid transporters in intestinal apical membrane

2.5.3 Transport of cationic amino acids

Cationic amino acids are mainly transported across the apical membrane of intestine according to Bröer by the system SLC3A1/SLC7A9.
neutral and basic amino acid transporter and L-type amino acid transporter
The system imports cationic amino acids in antiport with neutral amino acids and in symport with cystine::

$$\begin{array}{c} \\ aa_C^+ \\ aa_C^0 \\ cystine_C \end{array} \begin{array}{c} rBAT/b^0AT \\ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \end{array}$$

Cystine is not used in any other reactions of this reconstructed network, and is therefore disregarded:

$$\begin{array}{c} \\ aa_C^+ \\ aa_C^0 \end{array} \begin{array}{c} T^+ \\ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{array}$$

Using the transport reaction for neutral amino acids and the addition method, making reactions for each combination of cationic amino acid and neutral amino acid imported in symport with Na^+ may be omitted:

$$\begin{array}{c} \\ aa_C^+ \\ aa_C^0 \\ Na_C^+ \end{array} \begin{array}{c} t^+ \quad t^0 \\ \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} \end{array} \rightarrow \begin{array}{c} T^+ + T^0 \\ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \end{array}$$

The above method is, however not valid if the neutral amino acid is synthesized by the system, so transport reactions for each such combination should be included in the model.

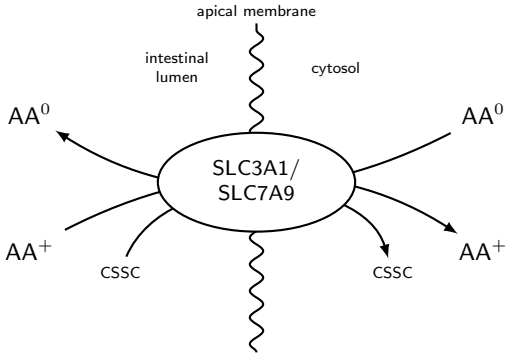


Figure 7: Cationic amino acid transport over intestinal apical membrane

2.5.4 Anionic amino acid transport

The main transport reaction of anionic amino acids across the apical membrane in intestine seems to be via the neuronal/epithelial high affinity glutamate transporter SLC1A1. This transporter has the following reaction:

$$\begin{array}{c} \\ aa_C^- \\ Na_C^+ \\ H_C^+ \\ K_C^+ \end{array} \begin{array}{c} EEAT3 \\ \left[\begin{array}{c} 1 \\ 3 \\ 1 \\ -1 \end{array}\right] \end{array}$$

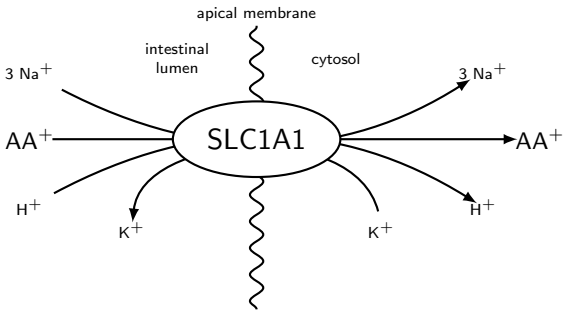


Figure 8: Anionic amino acid transport over intestinal apical membrane

2.6 Topological analysis

The topological analysis identifies the properties of the network (null space and degrees of freedom) without taking constraints like the directionality into account. It returns the degrees of freedom divided between outer and inner degrees of freedom. Inner degrees of freedom is the number of solutions of 'parallel routes and reaction cycles'. These are solutions with the same phenotypic behavior. Total degrees of freedom is the number of basis vectors of the null space of the stoichiometric matrix. Outer degrees of freedom is thereby the solutions of the network due to variation in the transport reactions. 'Compute null space' returns the basis vectors of the null space as chemical equations and visualizes the paths. However, the directionality of the reactions and constraints of the network are not taken into account. In spite of the limited insight this gives to the actual biological functions, the metabolic functions expected of the network should be reflected in the basis vectors.

2.7 Flux Balance Analysis

Flux Balance Analysis (FBA) is a mathematical method for optimization of metabolic networks at steady state within defined constraints, using an *objective function* as criterion. The method finds a solution in the null space of the stoichiometric matrix
The nullspace of the stoichiometric matrix can be considered to be the space of phenotypes available of the network. In this context, each point in the flux space represents a phenotype. (there may be several ways to reach a certain point in the space, as the connectivity is high – connected to inner degrees of freedom – the phenotype can be illustrated by as a 'black box' – is is a result of the inner workings, but the the inputs and outputs is what matters...) The algorithm is fast, and has a relatively low demand for data input. Constraints are defined as limits or fixed values for one or a combination of reaction fluxes (reaction rate). A reversible reaction is by default considered to be unbounded, meaning the flux through it can be anywhere from negative to positive infinity. This is of course unlikely, but it does not necessarily affect the network. Other constraints may be imposed on the network, these constraints are either defined by equations or inequalities, and are refereed to as balances and bounds correspondingly. [17]

The objective function is a mathematical representation of the objectives of the system, meaning the desired outputs. Such an objective may be to maximize growth of biomass or to maximize the production of a certain metabolite. Increasing the number of producers is often as effective as increasing the rate in which they produce, so objective function is often defined as a sum of biomass growth and production rate of the compound(s) specified, often resulting in a (hyper)plane of solutions. FBA is finding a solution in the null space that maximizes the objective function while taking the defined constraints into account.

2.7.1 Constraints

To constrain the network, limits can be set for each of the reactions but as each reaction is reliant on connected reactions, the steady state assumption limits the need of such constraints Consider the series of reactions from metabolite A to metabolite D:

$$A \rightarrow B \rightarrow C \rightarrow D$$

The rates of these reactions would be equal under steady state, meaning that constraining one would constrain all of them.

DNA (cytosine-5-)-methyltransferase (EC 2.1.1.37, id *dnmt1*) involves methylation of DNA. However, DNA is not represented in the database, nor is methyl-DNA. DNA and methyl-DNA are substituted by tetrahydrofolate and 5-methyl-tetrahydrofolate in the model, but as they also form a cycle connected by EC 2.1.1.10 (*mmum*), EC 2.1.2.1 (*gly1*) and 1.5.1.20 (*mhfred*), the transport of tetrahydrofolate and 5-methyl-tetrahydrofolate are constrained to equal the rate of EC 2.1.1.37:

$$dnmt + T.5mthf = 0$$

$$dnmt - T.tfol = 0$$

Cationic amino acids are transported in exchange for neutral amino acids, increasing the complexity of the network by creating parallel routes for import of arginine and lysine. The complexity is reduced by the substitution of Na^+ for essential amino acids,
The exchange of cationic amino acids against synthetic amino acids is included in the model, but for the initial analyzes the reactions are blocked to prevent the system from output of amino acids.

Amino acid composition of two feed types was obtained from Ajinomoto Eurolysine. The growth rate of salmon fed fish protein is shown to be higher than salmon fed plant protein. To validate the model, the amino acid composition of fish meal (FM) and soybean meal (SBM) will be used as constraints for flux balance analysis.

Table 3: Amino acid ratios of fish meal (FM) and soybean meal (SBM) from Ajinomoto Eurolysine [1]

AA	FM	SBM
Ala	6.8	4.5
Arg	6.3	7.4
Asn/Asp	10.0	11.6
Cys	0.9	1.4
Gln/Glu	14.1	18.4
Gly	6.6	4.4
His	2.6	2.6
Ile	4.7	4.7
Leu	8.1	7.9
Lys	8.7	6.3
Met	3.1	1.4
Phe	4.4	5.3
Pro	4.3	5.0
Ser	4.4	5.1
Thr	4.6	4.0
Trp	1.3	1.4
Tyr	3.6	3.8
Val	5.5	4.9

Limiting amino acids are identified by comparing the constraints and the FBA solution. Under steady state, the import and biosynthesis of amino acids is constrained by the ratio of consumption by the protein polymerization reaction. The restraint is mutual, and as for all linked reactions, the rate of consumption of a substrate cannot exceed its maximal rate of provision.

Amino acid import is constrained by feed composition, the limiting amino acid often has a maximized rate of import. Essential amino acids that may be metabolized to other amino acids may be maximized without being limiting, such as cysteine and methionine, arginine and proline, glycine and threonine or phenylalanine and tyrosine. Some amino acids are fully consumed by the system, while others may not be required at all in order for maximum yield of protein.

2.7.2 Feed efficiency with added supplements

To estimate feed efficiency, FBA results were compared to feed composition
The consumption of each amino acid per 100g of polymerized protein (PP) was estimated from the FBA results, using the rates of the import of amino acids and export of protein as $\frac{\frac{mol}{s}AA}{\frac{mol}{s}Protein} = \frac{molAA}{molProtein}$.

$$\text{Consumption} = \frac{\frac{gAA}{mol} \times molAA^3}{\frac{gProtein}{mol} \times PP} \tag{1}$$

The needed input of feed per gram protein was estimated by scaling the ratio between polymerized protein and feed protein (FP) by the limiting amino acid (AA_{lim}.

$$\frac{gFP}{gPP} = \frac{gAA_{lim}^{FBA}}{100gPP} \times \frac{100gFP}{gAA_{lim}^{Feed}} = \text{Input of total feed protein} \tag{2}$$

$$\frac{gFP}{gPP} \times \frac{gAA}{100gFP} = \frac{gAA}{100gPP} = \text{Input of AA} \tag{3}$$

The surplus of amino acids can be calculated

$$\text{Surplus of AA} = \text{Input of AA} - \text{Consumption of AA} \tag{4}$$

FBA simulations where the limiting amino acids are unconstrained will return negative values. This is the minimal extra addition to achieve the

$$\text{Minimal addition} = \frac{|\text{negative surplus of AA}|}{100 + |\text{negative surplus of AA}|} \tag{5}$$

3 Results

3.1 Properties of the network

The resulting network of reactions contains 33 transformer reactions, 87 transport reactions and 1 polymerization reaction. 40 Degrees of freedom, whereas 18 are inner[4], and 22 are outer degrees of freedom.

3.1.1 Included reactions

The transformer reactions included in the model are displayed in table 4.

[4] 14 are due to the parallel routes of lysine and arginine exchange with neutral amino acids

Table 4: The included transformer reactions in the model. The Abbreviations: THF = tetrahydrofolate

ID	equation	
cytosol:		
gludeh_c	$Glutamate + NADP^+ + H_2O \xleftrightarrow[gludeh\_c]{1.4.1.4} 2-Oxoglutarate + NH4^+ + NADPH + H^+$	
gln1	$Glutamate + NH4^+ + ATP \xrightarrow[gln1]{6.3.1.2} Glutamine + ADP + P_i$	
ser1_1	$3\text{-}Phosphohydroxypyruvate + Glutamate \xrightarrow[ser1\_1]{2.6.1.52} 2\text{-}Oxoglutarate + Phospho\text{-}Serine$	
ser2	$H_2O + Phospho\text{-}Serine \xrightarrow[ser2]{3.1.3.3} NP_i + Serine$	
serpyr	$Serine \xrightarrow[serpyr]{4.3.1.19} Pyruvate + NH_4$	
cys1	$Serine + H_2S \xrightarrow[cys1]{4.2.1.22} Cysteine + H_2O$	
metk	$Methionine + ATP + H_2O \xrightarrow[metk]{2.5.1.6} S\text{-}Adenosyl\text{-}Methionine + PP_i + P_i + H^+$	
cys2	$Serine + Homocysteine \xrightarrow[cys2]{4.2.1.22} Cystathionine + H_2O$	
cys3	$H_2O + Cystathionine \xrightarrow[cys3]{4.4.1.1} Cysteine + NH_4 + 2\text{-}Oxobutyrate$	
gluphos	$ATP + Glutamate \xrightarrow[gluphos]{2.7.2.11} ADP + Glutamate\text{-}5\text{-}Phosphate$	
gluro	$Glutamate\text{-}5\text{-}Phosphate + NADPH + H^+ \xrightarrow[gluro]{1.2.1.41} Glutamate\text{-}5\text{-}Semialdehyde + P_i + NADP^+$	
gshyd_c	$Glutamate\text{-}5\text{-}Semialdehyde \xleftrightarrow[gshyd\_c]{spontaneous} 1\text{-}Pyrroline\text{-}5\text{-}Carboxylate + H_2O + H^+$	
proox_c	$1\text{-}Pyrroline\text{-}5\text{-}Carboxylate + NADPH + H^+ \xrightarrow[proox\_c]{1.5.1.2} Proline + NADP^+$	
gly1	$THF + Serine \xleftrightarrow[gly1]{2.1.2.1} Glycine 5,10\text{-}Methylenete\text{-}THF + H_2O$	
thr1	$Threonine \xrightarrow[thr1]{4.1.2.48} Glycine + Acetaldehyde$	
aspox	$Aspartate + H_2O + O_2 \xrightarrow[aspox]{1.4.3.2} Oxaloacetate + NH_4 + H_2O_2$	
gluasp_c	$Aspartate + 2\text{-}Oxoglutarate \xrightarrow[gluasp\_c]{2.6.1.1} Oxaloacetate + Glutamate$	
asn1	$Aspartate + ATP + Glutamine + H_2O \xrightarrow[asn1]{6.3.5.4} Glutamate + Asparagine + AMP + PP_i + 2H^+$	
phetyr	$Tetrahydrobiopterin + Phenylalanine + O_2 \xrightarrow[phetyr]{1.14.16.1} 6,7\text{-}Dihydrobiopterin + Tyrosine + H_2O$	
metk	$Methionine + ATP + H_2O \xrightarrow[metk]{2.5.1.6} S\text{-}Adenosyl\text{-}Methionine + PP_i + P_i + H^+$	
mmum	$5\text{-}Methyl\text{-}THF + Homocysteine \xrightarrow[mmum]{2.1.1.10} THF + Methionine$	
dnmt1	$S\text{-}Adenosyl\text{-}Methionine + THF \xrightarrow[dnmt1]{2.1.1.37} S\text{-}Adenosyl\text{-}Homocysteine + 5\text{-}Methyl\text{-}THF + H^+$	
mhfred	$5\text{-}Methyl\text{-}THF + NADP^+ \xleftrightarrow[mhfred]{1.5.1.20} 5,10\text{-}Methylene\text{-}THF + NADPH + H^+$	
ahc	$S\text{-}Adenosyl\text{-}Homocysteine + H_2O \xrightarrow[ahc]{3.3.1.1} Adenosine + Homocysteine$	
mitochondrion:		
gludeh_m	$Glutamate + NADP^+ + H_2O \xleftrightarrow[gludeh\_m]{1.4.1.4} 2-Oxoglutarate + NH4^+ + NADPH + H^+$	
glnase_m	$Glutamine + H_2O \xrightarrow[glnase\_m]{3.5.1.2} Glutamate + NH_4^+ + H^+$	
glu1_m	$Glutamate\text{-}5\text{-}semialdehyde + NAD^+ + H_2O \xrightarrow[glu1\_m]{1.2.1.88} Glutamate + NADH + 2H^+$	
gshyd_m	$Glutamate\text{-}5\text{-}Semialdehyde \xleftrightarrow[gshyd\_m]{spontaneous} 1\text{-}Pyrroline\text{-}5\text{-}Carboxylate + H_2O + H^+$	
proox_m	$1\text{-}Pyrroline\text{-}5\text{-}Carboxylate + NADPH + H^+ \xrightarrow[proox\_m]{1.5.1.2} Proline + NADP^+$	
sergly	$Serine + Glyoxylate \xleftrightarrow[sergly]{2.6.1.45} 3\text{-}Hydroxypyruvate + Glycine$	
gluasp_m	$Aspartate + 2\text{-}Oxoglutarate \xrightarrow[gluasp\_m]{2.6.1.1} Oxaloacetate + Glutamate$	
argna_m	$Arginine + H_20 + 2H^+ \xrightarrow[argna\_m]{3.5.3.1} Ornithine + Urea$	
ornt_m	$Ornithine + 2\text{-}Oxoglutarate \xleftrightarrow[ornt\_m]{2.6.1.13} Glutamate\text{-}5\text{-}Semialdehyde + Glutamate$	

3.2 Flux balance analysis

3.2.1 FBA with constraints on all AA-transporters

FBA was performed with amino acid composition[5] from fish meal (FM) and soybean meal (SBM) as constraints in order to determine their efficiency, to identify the limiting amino acids and study how amino acid supplements may improve the efficiency of these feeds in terms of protein production. As expected, the efficiency of FM is higher than SBM.

Figure 9 summarizes the consumption of each metabolite per synthesized protein in the model under the constraints defined by soybean meal (SBM_0). The arrows indicate consumption or production[6], are labeled with the amount in milligrams they are consumed/produced per gram synthesized protein.
Table 5 displays how many milligrams of each amino acid was consumed per gram synthesized protein for fishmeal (FM_0) and soybean meal (SBM_0). Note that the collected mass of amino acids do not sum to 1 gram. This is because the supply of metabolites from glycolysis and TCA cycle, as well as ATP and redox coenzymes are unconstrained. This follows from assumptions 2 and 3. Table 6 displays the surplus of feed under the different conditions.

The total amount of feed protein needed per gram synthesized protein is the sum of table 5 and 6, 1160.41 mg fishmeal-protein and 1473.39 mg soybean meal-protein for the simulations with constraints on all amino acids.

[5] Displayed in table 3

[6] Consumption of metabolite A: $A \rightarrow Cell$, production of metabolite B: $Cell \rightarrow B$

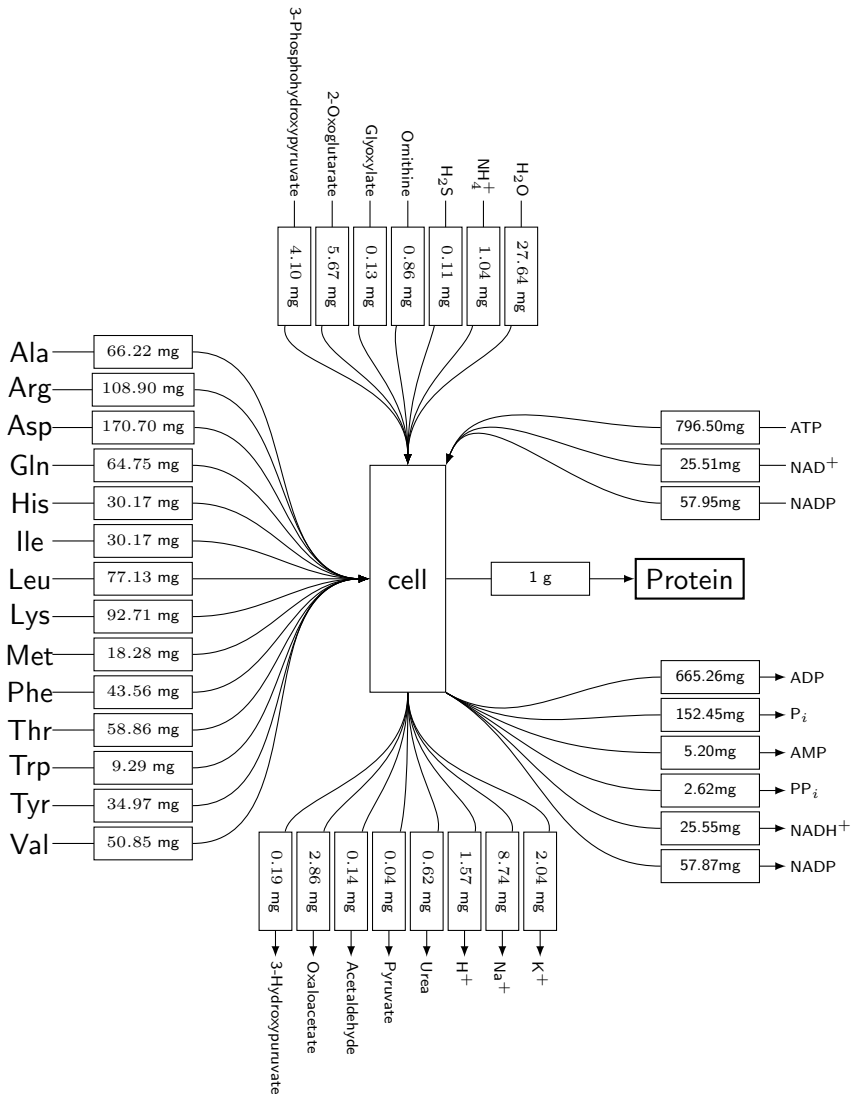


Figure 9: The import and output for the FBA solution under constraints SBM_0

3.2.2 Constraints were removed in order to identify the limiting amino acids.

The limiting amino acid was identified and the corresponding constraint removed before running FBA again. Five iterations were done for both FM and SBM, leading identification of the top five limiting amino acids for both feeds.

Limiting AA for FM:

1. Histidine
2. Threonine
3. Lysine
4. Arginine
5. Phenylalanine

Limiting AA for SBM:

1. Lysine
2. Methionine
3. Threonine
4. Histidine
5. Valine

FBA was performed, removing constraints for limiting amino acids subsequently to identify the next limiting amino acid. The resulting ratios of amino acid consumption per gram polymerized protein are rendered in table 5. Based on the consumption of the limiting amino acid in each case, the amount of feed protein required to synthesize one gram of protein was calculated. The consumed amounts of amino acids were subtracted from the feed requirement, resulting in the surplus of amino acids represented in table 6. Note that the 'surplus' values for the limiting amino acids are negative. These values represent the additional amount required of that specific amino acid in order for the system to reach the FBA result.

Table 5: Amino acid consumption from FBA with feed amino acid composition as constraints. All units are $\frac{\text{milligram AA}}{\text{gram protein}}$. Six analyzes were performed per feed type, FM_0 constraining all amino acids based on the composition of fish meal, FM_1 constraining all amino acids except the first limiting, FM_2 constraining all amino acids except the first and second limiting and so on. Vice versa for soybean meal (SBM)

AA	FM_0	FM_1	FM_2	FM_3	FM_4	FM_5	SBM_0	SBM_1	SBM_2	SBM_3	SBM_4	SBM_5
Ala	78.91	73.10	72.46	71.28	67.33	64.75	66.22	58.76	55.63	52.22	46.70	43.93
Arg	73.11	66.04	66.04	66.04	**66.04**	**66.04**	108.90	96.63	91.48	85.87	76.79	72.24
Asn	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Asp	116.04	107.50	106.56	104.82	99.00	95.22	170.70	151.48	143.41	134.61	120.38	113.25
Cys	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gln	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Glu	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gly	76.59	70.95	70.33	69.18	65.34	62.84	64.75	57.46	54.40	51.06	45.66	43.96
His	30.17	**30.17**	**30.17**	**30.17**	**30.17**	**30.17**	30.17	30.17	30.17	30.17	**30.17**	**30.17**
Ile	44.06	44.06	44.06	44.06	44.06	44.06	44.06	44.06	44.06	44.06	44.06	44.06
Leu	77.12	77.12	77.12	77.12	77.12	77.12	77.12	77.12	77.12	77.12	77.12	77.12
Lys	92.71	92.71	92.71	**92.71**	**92.71**	**92.71**	92.71	**92.71**	**92.71**	**92.71**	**92.71**	**92.71**
Met	18.28	18.28	18.28	18.28	18.28	18.28	18.28	18.28	**18.28**	**18.28**	**18.28**	**18.28**
Phe	43.56	43.56	43.56	43.56	43.56	**43.56**	43.56	43.56	43.56	43.56	43.56	43.56
Pro	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ser	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Thr	53.38	49.45	**49.45**	**49.45**	**49.45**	**49.45**	58.86	49.45	49.45	**49.45**	**49.45**	**49.45**
Trp	9.29	9.29	9.29	9.29	9.29	9.29	9.29	9.29	9.29	9.29	9.29	9.29
Tyr	34.97	34.97	34.97	34.97	34.97	34.97	34.97	34.97	34.97	34.97	34.97	34.97
Val	50.85	50.85	50.85	50.85	50.85	50.85	50.85	50.85	50.85	50.85	50.85	**50.85**
total	799.03	768.05	765.85	762.77	748.14	736.2	870.44	814.78	795.38	774.21	739.99	722.84

Table 6: Surplus of amino acids in $\frac{\text{milligram AA}}{\text{gram Protein}}$

AA	FM_0	FM_1	FM_2	FM_3	SBM_0	SBM_1	SBM_2	SBM_3
Ala	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Arg	0.00	1.69	1.10	0.00	0.00	0.00	0.00	0.00
Asn	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Asp	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cys	10.44	9.68	9.59	9.43	20.60	18.28	17.31	16.25
Gln	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Glu	163.62	151.58	150.25	147.79	270.77	240.27	227.48	213.51
Gly	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
His	0.00	**-2.22**	**-2.46**	**-2.92**	8.09	3.78	1.97	0.00
Ile	10.48	6.47	6.03	5.21	25.11	17.32	14.05	10.48
Leu	16.87	9.95	9.19	7.78	39.13	26.04	20.54	14.55
Lys	8.52	0.82	0.00	**-1.52**	0.00	**-10.44**	**-14.82**	**-19.60**
Met	17.69	15.04	14.75	14.21	2.32	0.00	**-0.97**	**-2.04**
Phe	7.50	3.74	3.33	2.56	34.44	25.65	21.96	17.94
Pro	49.90	46.23	45.82	45.07	73.58	65.29	61.81	58.02
Ser	51.06	47.30	46.89	46.12	75.05	66.60	63.05	59.18
Thr	0.00	0.00	**-0.43**	**-1.23**	0.00	2.78	0.00	**-3.04**
Trp	5.79	4.68	4.56	4.34	11.31	8.99	8.02	6.95
Tyr	6.81	3.74	3.40	2.77	20.95	14.66	12.01	9.13
Val	12.97	8.28	7.76	6.80	21.26	13.14	9.73	6.01
total	361.38	309.20	302.67	292.09	602.61	502.79	457.93	412.03

Table 7: Additions in milligrams per gram feed protein saves % feed protein

	FM				SBM		
saves / AA	3.83%	4.31%	5.09%	saves / AA	6.23%	9.03%	11.89%
His	0.17	0.19	0.23	Lys	0.69	1.01	1.37
Thr	0.00	0.03	0.10	Met	0.00	0.10	0.20
Lys	0.00	0.00	0.12	Thr	0.00	0.00	0.21

4 Discussion

4.1 SalsAA confirms the limiting amino acids of soybean meal

Lysine, methionine and threonine were indicated to be the three most limiting amino acids in soybean meal protein, and they are indeed considered to be limiting amino acids in plant based feeds [2] [12] [4] [10]. Methionine and lysine are usually the most limiting amino acids in feeding trails [12] [9], but threonine is also known to benefit the diet, due to it's limited abundance in plant protein. Factors that would lead to a larger demand for methionine include synthesis of taurine, which too is deficient in plant protein, and constraining the H_2S dependent cysteine biosynthesis, through limiting reaction *cys1* or import of H_2S, leading to methionine dependent synthesis of cysteine. Dietary cysteine is shown to decrease the demand for methionine in several species [9], and constraining H_2S dependent synthesis of cysteine would correspond to this.

H_2S may not be synthesized in salmon, based on the sulphur metabolism pathway diagram in KEGG, so if present at all in the feed, H_2S import should be constrained. Additional FBA was performed to study the effect of blocking H_2S transport, but this did not affect the rate of protein polymerization. However, it did increase cysteine import, but not to a limiting level. The abundance of cysteine in the feed was larger than the demand from the network, and the constraint on cysteine import did not limit the provision of methionine even when blocking the H_2S dependent cysteine synthesis. As the cysteine content of soybean meal is higher than in fish meal, the methionine reserves in fish meal would be more affected than in soymeal, but fishmeal's higher content of methionine compensates for this. Under the conditions of study, H_2S dependent cysteine synthesis does not affect the system much, but this detail may be major for feeds of low cysteine and methionine content.

4.2 SalsAA confirms salmon's preference for fish protein

The FBA simulations indicate fish meal to be a more effective feed for protein polymerisation than soybean meal. This is indicated both by better utilisation of the imported amino acids (table 5) and lower surplus of amino acids (table 6).

4.3 The model predicts additions to improve feed efficiency

The model predicts an addition of 13.7 mg lysine, 2.0 mg methionine and 2.1 mg threonine per gram of soybean meal protein would reduce the feed protein usage by 11.89%. This tendency corresponds with established knowledge, but the numbers do not account for the many functions amino acids have in addition to protein synthesis. Lysine is involved in transport of long chain fatty acids, and methionine is involved in a range of pathways as a carrier for methyl groups. [12] Thus, a whole-genome model would be expected to contain more reactions consuming lysine and methionine, but it would also depend on consumption of other amino acids, so further study is required to conclude on whether the magnitude of the SalsAA model's predicted additives are small or large.

4.4 Limitations of the model

There are several factors that are not accounted for by the SalsAA model. For instance, the digestibility of the feeds may vary, and addition of soluble ingredients to aquafeeds may lead to nutrients getting diluted before they reach their destination. However, in a framework of models accounting for the different aspects and challenges of aquafeed development, the model may still contribute to balance amino acid composition of feeds.

4.5 Further developments

The SalsAA model will be manually checked and used as a reference for the amino acid synthesis in the whole-genome Salarecon model, part of the DigiSal Project. Artificial parts of the network will be removed, and the imaginary boundaries connected to the rest of the network.

SalsAA 1.0 simulates a sub module of an average salmon cell, but it could be converted to simulate specific cell types as well. This would require the transport reactions to be reevaluated, and may require adjustments to the objective function. For instance, a model for intestinal epithelial cell should also account for the transport of amino acids over the basolateral membrane. The model could also be used to study the effects of other feeds based on other sources like insects or yeast.

5 Conclusion

SalsAA 1.0 is validated by the correspondence between predictions of limiting amino acids and earlier research [12] [2]. It also predicts guideline amounts of amino acid supplements that would lead to feed protein savings of up to 11.89% under the assumptions of the model. However, as the model describes a sub module of the network of reactions of an average salmon cell, it does not account for all processes amino acids take part in. SalsAA 1.0 is capable of steady state simulations, and the flux balance analysis correctly predicts fish meal protein to support protein synthesis better than soy protein.

References

[1] Ajinmoto-Eurolysine. Feedstuffs amino acids database. http://ajinomoto-eurolysine.com/feedstuffs-amino-acid-database.html, 2014. Accessed: 2018-05-08.

[2] Andersen, S M, Waagbø, R, Espe, M. Functional amino acids in fish nutrition, health and welfare. *Frontiers in Bioscience, Elite*, 8:143–169, 2016.

[3] Belghit, I, Liland, N S, Waagbø, R, Biancarosa, I, Pelusio, N, Li, Y, Krogdahl, Å, Lock, E-J. Potential of insect-based diets for Atlantic salmon (*Salmo salar*). *Aquaculture*, 491:72–81, 2018.

[4] Berge, G E, Sveier, H, Lied, H. Nutrition of Atlantic salmon (*Salmo salar*); the requirement and metabolic effect of lysine. *Comparative Biochemistry and Physiology Part A*, 120:477–485, 1998.

[5] Bjercke, R. C-feed. http://www.cfeed.no/, 2014. Accessed: 2018-05-14.

[6] Bröer S. Amino Acid Transport Across Mammalian Intestinal and Renal Epithelia. *Physiological Reviews*, 88:249–286, 2008.

[7] Brunk, E, Sahoo, S, Zielinski, D C, Altunkaya, A, Drger, A, Mih, N, Gatto, F, Nilsson, A, Gonzalez, G A P, Aurich, M K, Prli, A, Sastry, A, Danielsdottir, A D, Heinken, A, Noronha, A, Rose, P W, Burley, S K, Fleming, R M T, Nielsen, J, Thiele, I, Palsson, B Ø. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nature Biotechnology*, 36:272281, 2018.

[8] Directorate of fisheries. Total sale of fish in the norwegian aquaculture indystry. https://www.fiskeridir.no/Akvakultur/Statistikk-akvakultur/Akvakulturstatistikk-tidsserier/Totalt-hele-naeringen, 2017. Accessed: 2018-05-07.

[9] Halver,J E, Hardy, R W. *Fish Nutrition, third edition*. Academic Press, 2002.

[10] Helland,S, Grisdale-Helland, B. Dietary threonine requirement of Atlantic salmon smolts. *Aquaculture*, 321:230–236, 2011.

[11] Kanehisa, M, Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28:27–30, 2000.

[12] Li, P, Mai, K, Trushenski, J, Wu, G. New developments in fish amino acid nutrition: Towards functional and environmentally oriented aquafeeds. *Amino Acids*, 37:43–53, 2009.

[13] Lien, S, Koop, B F, Sandve, S R, Miller, J R, Kent, M P, Nome, T, Hvidsten, T R, Leong, J S, Minkley, D R, Zimin, A, Grammes, F, Grove, H Gjuvsland, A, Walenz, B, Hermansen, R A, von Schalburg, K, Rondeau, E B, Di Genova, A, Samy, J K A, Vik, J O, Vigeland, M D, Caler, L, Grimholt, U, Jentoft, S, Våge, de Jong, P, Moen, T, Baranski, M, Palti, Y Smith, D R, Yorke, J A, Nederbragt, A J, Tooming-Klunderud, A, Jakobsen, K S, Jiang X, Fan, D, Hu, Y, Liberles, D A, Vidal, R, Iturra, P, Jones, J M, Jonasen, I, Maass, A, Omholt, S W, Davidson, W S. The Atlantic salmon genome provides insights into rediploidization. *Nature*, 533:200–205, 2016.

[14] Lock, J-E. Aquafly. `https://nifes.hi.no/en/prosjekt/insects-salmon-feed/`, 2014. Accessed: 2018-05-14.

[15] Orth J D, Thiele I, Palsson B Ø. What is flux balance analysis? *Nature Biotechnology*, 28:245–248, 2010.

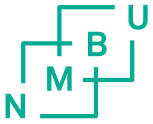
[16] Øverland, M. Biofeed. `https://www.foodsofnorway.net/key-research/biomass/trees`, 2015. Accessed: 2018-05-14.

[17] Palsson, B Ø. *Systems Biology*. Cambridge University Press, 2015.

[18] Thiele I, Palsson B Ø. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, 5:93–121, 2010.

[19] Wilson, R P, Cowey, C B. Amino acid composition of whole body tissue of rainbow trout and Atlantic salmon. *Aquaculture*, 48:373–376, 1985.

[20] World Health Organisation. Fact sheet no 394, healthy diet. `http://web.archive.org/web/20080207010024/`, 2015. Accessed: 2018-02-28.



Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway