UNIVERSIDAD NACIONAL DE COLOMBIA

Identifying proteins and metabolic pathways associated with the neuroprotective response mediated by tibolone in astrocytes under an induced inflammatory model.

Daniel Camilo Osorio Hurtado

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería de Sistemas e Industrial
Bogotá, Colombia
2016

Identifying proteins and metabolic pathways associated to the neuroprotective response mediated by tibolone in astrocytes under an induced inflammatory model.

Daniel Camilo Osorio Hurtado

Tesis presentada como requisito parcial para optar al título de:
Magister en Bioinformática

Directora:
Ph.D. Janneth Gonzalez Santos

Co-Director:
Ph.D. Andrés Mauricio Pinzón Velasco

Línea de Investigación:
Biología de Sistemas - Reconstrucción y Análisis de Redes Metabólicas
Grupos de Investigación:
Bioinformática y Biología de Sistemas Computacional - Universidad Nacional de Colombia
Bioquímica Experimental y Computacional - Pontificia Universidad Javeriana

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería de Sistemas e Industrial
Bogotá, Colombia
2016

Resumen

En este trabajo, se identificaron las proteínas y rutas metabólicas asociadas a la respuesta neuroprotectora mediada por el neuroesteroide sintético tibolona bajo un modelo inflamatorio inducido por palmitato usando análisis de balance de flujo (FBA). Para tal fin, se modelaron tres diferentes escenarios metabólicos ('saludable', 'inflamado' y 'medicado') bajo una reconstrucción metabólica tejido-especifica de astrocitos maduros construida a partir de datos de expresión génica. La reconstrucción metabólica se construyó, validó y limitó usando tres paquetes de software (`'minval'`, `'g2f'` y `'exp2flux'`) liberados a través de los repositorios R CRAN durante el desarrollo de este trabajo. A partir de nuestros análisis, predecimos que la tibolona ejecuta sus acciones neuroprotectoras a través de la reducción de la neurotoxicidad mediada por el L-glutamato en astrocitos, induciendo la activación de varias rutas metabólicas con funciones neuroprotectoras asociadas como: Metabolismo de taurina, gluconeogénesis y rutas de señalización PPAR y mediadas por calcio. Adicional a esto, encontramos un aumento en la tasa de crecimiento asociado a la tibolona que podría estar relacionado con efectos secundarios reportados para los compuestos esteroideos en otros tipos celulares humanos.

Palabras clave: Astrocitos, Tibolona, Neuroproteccion, Inflamación, Analisis de Balance de Flujo.

Abstract

In this work, proteins and metabolic pathways associated with the neuroprotective response mediated by the synthetic neurosteroid tibolone under a palmitate-induced inflammatory model were identified by flux balance analysis (FBA). Three different metabolic scenarios ('healthy', 'inflamed' and 'medicated') were modeled over a gene expression data-driven constructed tissue-specific metabolic reconstruction of mature astrocytes. Astrocyte reconstruction was built, validated and constrained using three open source software packages (`'minval'`, `'g2f'` and `'exp2flux'`) released through CRAN R repositories during the development of this work. From our analysis, we predict that tibolone execute their neuroprotective effects through a reduction of neurotoxicity mediated by L-glutamate in astrocytes, inducing the activation several metabolic pathways with neuroprotective actions associated such as taurine metabolism, gluconeogenesis, calcium and PPAR signaling pathways. Also, we found a tibolone associated increase in growth rate probably in concordance to previously reported side effects of steroid compounds in other human cell types.

Keywords: Astrocytes, Tibolone, Neuroprotection, Inflammation, Flux Balance Analysis.

Content

Figures

Tables

Objectives

Main:

Identify proteins and metabolic pathways involved in the neuroprotective effects of tibolone in human astrocytes based on metabolic scenarios comparison.

Specific

- Build a tissue-specific computational model of astrocytes metabolism using gene expression data integration.
- Evaluate the effects caused by the increase of free fatty acids and tibolone presence in astrocytes metabolism.
- Determine metabolic pathways and relevant functional products in response to steroid tibolone through systems biology approximations.
- Evaluate the importance of proteins and metabolic pathways previously identified on the dynamics of the metabolic model.

1. Building a metabolic reconstruction: Doing a MINimal VALidation of stoichiometric reactions through 'minval' R package.

Original title: minval: An R package for MINimal VALidation of stoichiometric reactions.
Written by: *Daniel Osorio, Janneth Gonzalez and Andrés Pinzón-Velasco*

Abstract

The genome-scale metabolic reconstructions, a compilation of all stoichiometric reactions that can describe the entire cellular metabolism of an organism, have become an indispensable tool for our understanding of biological phenomena, covering fields that range from systems biology to bioengineering. Evaluation of metabolic reconstructions are generally carried through Flux Balance Analysis, an optimization method where the biological sense of optimal solution is sensitive to thermodynamic unbalance, caused by the presence of stoichiometric reactions whose compounds are not produced or consumed in any other reaction (orphan metabolites) and by mass unbalanced stoichiometric reactions. The **minval** package was designed as a tool to identify orphan metabolites and the mass unbalanced reactions in a set of stoichometry reactions, it also permits to extract all reactants, products, metabolite names and compartments from a metabolic reconstruction, moreover specific functions to map compound names associated to the Chemical Entities of Biological Interest (ChEBI) database are also included.

1.1. Introduction

A chemical reaction is a process where a set of chemical compounds called *reactants* are transformed into another compounds called *products* [1]. The accepted way to represent a chemical reaction is called a *stoichiometric reaction*, where reactants are placed on the left and the products on the right separated by an arrow which indicates the direction of the

reaction as is showed in the equation 1-1 [2]. In biochemistry a set of chemical reactions that transform a substrate into a product, after several chemical transformations is called a metabolic pathway [3]. The compilation of all stoichiometric reactions included in all metabolic pathways that can describe the entire cellular metabolism encoded in the genome of a particular organism is known as a *genome-scale metabolic reconstruction* [4] and has become an indispensable tool for studying metabolism of biological entities at the systems level [5].

$$\overbrace{\underbrace{1}_{coefficient}\ \underbrace{cis-aconitic\ acid}_{metabolite\ name}\ \underbrace{[c_a]}_{compartment}\ +1\ water[c_a]}^{reactants}\ \underbrace{\Rightarrow}_{directionallity}\ \overbrace{1\ isocitric\ acid[c_a]}^{products} \quad (1\text{-}1)$$

Reconstruction of genome-scale metabolic models starts with a compilation of all known stoichiometric reactions for a given organism, as evidenced by the presence of enzyme coding genes in its genome. Thus the reactions in which these enzymes are known to participate in, are usually downloaded from specialized databases as KEGG [6], BioCyc [7], Reactome [8], BRENDA [9] and SMPDB [10], however the downloaded stoichiometric reactions are not always mass-charge balanced and don't represent complete pathways to construct a high-quality metabolic reconstruction [5, 11]. The identification and curation of these type of reactions is a time consuming process which the researcher have to complete manually using available literature or experimental data [12].

Genome-scale metabolic reconstructions are usually interrogated through *Flux Balance Analysis* (FBA), an optimization method that allows us to understand the metabolic status of the cell, improve the production capability of a desired product or make a rapid evaluation of cellular physiology at genome-scale [4, 13]. Nevertheless FBA is sensitive to thermodynamic unbalance, so in order to asses the validity of a biological extrapolation (i.e. an optimal solution) from a FBA analysis it is mandatory to avoid this type of unbalancing in mass conservation through all model reactions [14]. Another drawback when determining the validity of a metabolic reconstruction is the presence of reactions with compounds that are not produced or consumed in any other reaction (dead ends), generally known as orphan metabolites [4, 5]. The presence of this type of metabolites can be problematic, since they lead to an artificial cellular accumulation of metabolism products which therefore bias our biological conclusions. Tracking these metabolites is also a time consuming process, which most of the time has to be performed manually or partially automatized by in-house scripting. Given that typical genome-scale metabolic reconstructions account for hundreds or thousands of biochemical reactions, the manually curation of these models is a task that can lead to both, the introduction of new errors and to overlook some others.

The most popular FBA implementations as COBRA and RAVEN includes similar functions

(`checkMassChargeBalance` and `getElementalBalance` respectively) implemented under the commercial MATLAB® environment. These functions identify orphan metabolites and mass unbalanced reaction based in the chemical formula or the IUPAC International Chemical Identifier (InChI) supplied manually by the user for each metabolite included in the genome-scale metabolic reconstruction. With the aim to automatize the identification of orphan metabolites as well as the unbalanced stoichiometric reactions in a genome-scale metabolic reconstruction, we have developed the **minval** package. It includes thirteen functions to evaluate mass balance and extract all reactants, products, orphan metabolites, metabolite names and compartments for a set of stoichiometric reactions, moreover specific functions to map compound names associated to the Chemical Entities of Biological Interest (ChEBI) database are also included.

For this version we use the included "`glugln`" dataset [15], 128 non-exchange/sink stoichiometric reactions from the reconstruction of the glutamate/glutamine cycle constructed in-house using the KEGG database , as an example for each function included in the **minval** package with the aim to show their potential use.

1.2. Installation and functions

`minval` includes 15 functions and is available for download and installation from CRAN, the Comprehensive R Archive Network. To install and load it, just type:

```
> install.packages("minval")
> library(minval)
```

The `minval` package requires R version 2.10 or higher. Development releases of the package are available on the GitHub repository `http://github.com/gibbslab/minval`.

Inputs and syntaxis

The functions included in `minval` package take as input a string list with stoichiometric reactions. The data loading from traditional human-readable spreadsheets can be carried through other CRAN-available packages as `gdata`, `readxl` or `xlsx`. Each reaction string must contain metabolites, with an optional compartment label between square brackets. The metabolites should be separated by a plus symbol (`+`) between two blank spaces and may have just one stoichiometric number before the name. The reactants should be separated of products by an arrow using the following symbol `=>` for irreversible reactions and `<=>` for reversible reactions.

Syntax Validation

Flux Balance Analysis method is implemented in a variety of software under different programming languages. Some of the most popular implementations are `COBRA` [16] and `RAVEN` [17] under matlab language as well as `sybil` and `abcdeFBA` under R language. The `isValidSyntax` function validate the well accepted compartmentalized stoichiometric syntax (Equation 1-1) for several FBA implementations and returns a boolean value TRUE if syntax is correct. In this example we show the stoichiometric syntax for the inter-convertion of malate to fumaric acid and water in astrocytes cytoplasm.

```
> isValidSyntax("(S)-malate(2-)[c_a] <=> fumaric acid[c_a] + water[c_a]")
[1] TRUE
```

Reactants and Products

As defined in introduction, stoichiometric reactions represent the transformation of reactants into products in a chemical reaction. The `reactants` and `products` functions extract and return all reactants and products present in a stoichiometric reaction as a vector. In this example we show the extraction of the reactants (quinone and succinic acid) and products (hydroquinone and fumaric acid) in a reaction that occurs in astrocytes mitochondrias.

```
> reactants("Quinone[m_a] + succinic acid[m_a] => Hydroquinone[m_a] + fumaric acid[m_a]")
[1] "Quinone[m_a]"       "succinic acid[m_a]"
> products("Quinone[m_a] + succinic acid[m_a] => Hydroquinone[m_a] + fumaric acid[m_a]")
[1] "Hydroquinone[m_a]" "fumaric acid[m_a]"
```

Metabolites

Two of the more popular packages that implement FBA analysis such as `COBRA` [16] and `RAVEN` [17] require the complete list of metabolites included in the metabolic reconstruction, in a particular section of the human-readable input file. The `metabolites` function automatically identifies and lists all metabolites (with and without compartments) for a specific or a set of stoichiometric reactions. In this example we show how to extract all metabolites (reactants and products) with and without compartment for the Ubiquinol and FAD production reaction in astrocytes mitochondrias.

```
> metabolites("FADH2[m_a] + ubiquinone-0[m_a] => FAD[m_a] + Ubiquinol[m_a]")
[1] "FADH2[m_a]"        "ubiquinone-0[m_a]" "FAD[m_a]"
[4] "Ubiquinol[m_a]"
```

As was mentioned before, the report option without compartment was added:

```
> metabolites("FADH2[m_a] + ubiquinone-0[m_a] => FAD[m_a] + Ubiquinol[m_a]",
+             woCompartment = TRUE)
[1] "FADH2"        "ubiquinone-0" "FAD"          "Ubiquinol"
```

Orphan Metabolites

Orphan metabolites, compounds that are not produced or consumed in any other reaction are one of the main causes of mass unbalance in metabolic reconstructions. The `orphanReactants` function, identifies compounds that are not produced internally by any other reaction and should be added to the reconstruction as an exchange reaction following the protocol proposed by [5]. In this examples we show how to extract all orphan compounds for all reactions included in the glutamate/glutamine cycle.

```
> data("glugln")
> orphanReactants(glugln)
```

```
 [1] "alpha-D-Glucose 6-phosphate[r_n]"    "water[r_n]"
 [3] "2,3-bisphospho-D-glyceric acid[r_n]" "GTP[c_n]"
 [5] "oxaloacetic acid[m_n]"               "citric acid[c_n]"
 [7] "coenzyme A[c_n]"                     "Quinone[m_n]"
 [9] "D-Glutamine[m_n]"                    "L-Glutamine[m_n]"
[11] "FADH2[m_n]"                          "oxygen atom[m_n]"
[13] "Ferrocytochrome c2[m_n]"             "diphosphate(4-)[m_n]"
[15] "alpha-D-Glucose 6-phosphate[r_a]"    "water[r_a]"
[17] "2,3-bisphospho-D-glyceric acid[r_a]" "GTP[c_a]"
[19] "hydrogencarbonate[m_a]"              "citric acid[c_a]"
[21] "coenzyme A[c_a]"                     "Quinone[m_a]"
[23] "L-glutamic acid[c_a]"                "Ammonia[c_a]"
[25] "FADH2[m_a]"                          "oxygen atom[m_a]"
[27] "Ferrocytochrome c2[m_a]"             "diphosphate(4-)[m_a]"
```

The `orphanProducts` function, identifies compounds that are not consumed internally by any other reaction and should be added to the reconstruction as an sink reaction following the protocol proposed by [5]. In this example we show the option added to `orphan*` functions, that permits to report the orphan metabolites as a list grouped by compartment:

```
> orphanProducts(glugln, byCompartment = TRUE)
```

```
$r_n
[1] "alpha-D-Glucose[r_n]"           "phosphate(3-)[r_n]"
[3] "2-phospho-D-glyceric acid[r_n]"

$c_n
[1] "GDP[c_n]"         "(S)-Lactate[c_n]" "acetyl-CoA[c_n]"

$m_n
[1] "Hydroquinone[m_n]"       "D-glutamic acid[m_n]"
[3] "FAD[m_n]"                "Ferricytochrome c2[m_n]"

$r_a
[1] "alpha-D-Glucose[r_a]"           "phosphate(3-)[r_a]"
[3] "2-phospho-D-glyceric acid[r_a]"
```

```
$c_a
[1] "GDP[c_a]"         "(S)-Lactate[c_a]" "acetyl-CoA[c_a]"  "L-Glutamine[c_a]"

$m_a
[1] "Hydroquinone[m_a]"       "FAD[m_a]"
[3] "Ferricytochrome c2[m_a]"
```

Compartments

As well as in cells, where not all reactions occur in all compartments, stoichiometric reactions in a metabolic reconstruction can be labeled to be restricted for a single compartment during FBA, by the assignment of a compartment label after the stoichiometric coefficient and name of each metabolite. Some FBA implementations require the report of all compartments included in the metabolic reconstruction as an independent part of the human-readable input file. In this example we show how to extrac all compartments for all reactions included in the glutamate/glutamine cycle.

```
> compartments(glugln)

[1] "c_n" "r_n" "m_n" "c_a" "r_a" "m_a"
```

Association with ChEBI

The Chemical Entities of Biological Interest (`ChEBI`) database is a freely available dictionary of molecular entities focused on 'small' chemical compounds involved in biochemical reactions [18]. Amongst other characteristics, the release 136 of ChEBI database contains a set of standardized metabolite names, synonyms and molecular formula for at least 52521 chemical compounds. The use of standardized metabolite names facilitate the sharing process and inter-convertion to another metabolite names or identifiers [19, 20]. The `minval` package contains five functions to check and extract values from a local copy of the ChEBI database release 136. The `is.chebi` function takes a compound name as input, compares it against all the compounds names in ChEBI and returns a logical value TRUE if a match is found. In this next four examples we show the potential use of the functions using as input the acetyl-CoA compound.

```
> is.chebi("acetyl-CoA")

[1] TRUE
```

The `chebi.id` function takes a compound name as input, compares it against all the compounds names in ChEBI and returns the compound identifier if a match is found.

```
> chebi.id("acetyl-CoA")
```

```
[1] "15351"
```

The `chebi.formula` function takes a compound name as input, compares it against all the compounds names in ChEBI and returns the molecular formula if a match is found.

```
> chebi.formula("acetyl-CoA")

[1] "C23H38N7O17P3S"
```

The `chebi.candidates` function takes a compound name as input, compares it against all the compounds synonyms in ChEBI and returns possible compound names if a match is found.

```
> candidates<-chebi.candidates("acetyl-CoA")
> head(candidates)

[1] "acetoacetyl-CoA"                 "acetyl-CoA"
[3] "(1-hydroxycyclohexyl)acetyl-CoA" "cinnamoyl-CoA"
[5] "2-methylacetoacetyl-CoA"         "phenylacetyl-CoA"
```

The `to.ChEBI` function translates the compounds names of a stoichiometric reaction into their corresponding identifier or molecular formula in the ChEBI database. In this example we show how to use the `to.ChEBI` function for the Ubiquinol and FAD production reaction in astrocytes mitochondrias.

```
> toChEBI("FADH2[m_a] + ubiquinone-0[m_a] => FAD[m_a] + Ubiquinol[m_a]")

[1] "1 17877 + 1 27906 => 1 16238 + 1 17976"

> toChEBI("FADH2[m_a] + ubiquinone-0[m_a] => FAD[m_a] + Ubiquinol[m_a]",formula = TRUE)

[1] "1 C27H35N9O15P2 + 1 C9H10O4 => 1 C27H33N9O15P2 + 1 C9H12O4(C5H8)n"
```

Mass Balance Validation

Thermodynamic unbalance of genome-scale metabolic reconstructions can also be promoted by stoichiometric mistakes. In a well balanced stoichiometric reaction according to the Lomonósov-Lavoisier law, the mass comprising the reactants should be the same mass present in the products. The `isBalanced` function converts the metabolites identifiers to molecular formulas, multiplies the atom numbers by their respective stoichiometric coefficient, and establishes if the atomic composition of reactants and products are the same, it returns a logical value TRUE if mass is unbalanced. In this example we show the mass balance evaluation for the first twenty reactions of the glutamate/glutamine cycle.

```
> isBalanced(glugln[1:20])
```

```
 [1] FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[13] FALSE  TRUE FALSE FALSE  TRUE  TRUE  TRUE  TRUE
```

The `isBalanced` function also include an option to show the molecular formula of mass unbalanced formulas through the option `show.formulas`.

```
> isBalanced(glugln[1:20], show.formulas = TRUE)

     [,1]
[1,] "alpha-D-Glucose 6-phosphate[r_n] + water[r_n] => alpha-D-Glucose[r_n] + phos ..."
[2,] "beta-D-fructofuranose 1,6-bisphosphate[c_n] + water[c_n] => beta-D-fructofur ..."
[3,] "D-Glyceraldehyde 3-phosphate[c_n] + phosphate(3-)[c_n] + NAD(+)[c_n] <=> 3-p ..."
[4,] "ATP[c_n] + 3-phosphoglyceric acid[c_n] <=> ADP[c_n] + 3-phosphonato-D-glycer ..."
[5,] "3-phosphonato-D-glyceroyl phosphate(4-)[c_n] => 2,3-bisphospho-D-glyceric ac ..."
[6,] "2,3-bisphospho-D-glyceric acid[c_n] + water[c_n] => 3-phosphoglyceric acid[c ..."
     [,2]
[1,] "1 C6H13O9P + 1 H2O => 1 C6H12O6 + 1 O4P"
[2,] "1 C6H14O12P2 + 1 H2O => 1 C6H13O9P + 1 O4P"
[3,] "1 C3H7O6P + 1 O4P + 1 C21H28N7O14P2 <=> 3 C3H4O10P2 + 1 C21H29N7O14P2 + 1 H"
[4,] "1 C10H16N5O13P3 + 3 C3H7O7P <=> 1 C10H15N5O10P2 + 3 C3H4O10P2"
[5,] "3 C3H4O10P2 => 2 C3H8O10P2"
[6,] "2 C3H8O10P2 + 1 H2O => 3 C3H7O7P + 1 O4P"
```

1.3. Summary

We introduced the `minval` package to evaluate mass balancing correctness of metabolic reconstructions and to extract all reactants, products, orphan metabolites, metabolite names and compartments for a set of stoichiometric reactions. We show step by step the minimal evaluation process of mass balance using the 128 non-exchange reactions included in the glutamate/glutamine cycle included in the "`glugln`" dataset. Also some examples of metabolites names - ChEBI database association was showed.

2. Building a tissue-specific metabolic reconstruction: Finding and filling gaps in metabolic networks through 'g2f' R package.

Original title: g2f: An R package for find and fill gaps in metabolic networks.

Written by: *Kelly Botero, Daniel Osorio, Janneth Gonzalez and Andrés Pinzón-Velasco*

Abstract

During the building of a genome-scale metabolic network reconstruction, several dead-end metabolites which cannot be imported/produced, or that are not used as substrates or released by not any of the reactions incorporated into the network. The presence of these dead-end metabolites can block out the net flux of the objective function when is evaluated through Flux Balance Analysis (FBA), and when it is not blocked, bias in the biological conclusions increase. The refinement to restore the connectivity of the network can be carried out manually or using computational algorithms. The **g2f** package was designed as a tool to find the dead-end metabolites, and fill it from the stoichiometric reactions of a reference, filtering candidate reactions using a weighting function. Also the option to download all the set of gene-associated stoichiometric reactions for a specific organism from the KEGG database is available.

2.1. Introduction

Genome-scale metabolic network reconstructions (GMNR) specify the chemical reactions catalyzed by hundreds of enzymes (registered in enzyme commission – E.C.) and cover the molecular function of a substantial fraction of a genome [21]. The main goal of these network reconstructions is to relate the genome of a given organism with its physiology, incorporating every metabolic transformation that this organism can perform [22, 23]. The GMNR are converted into computational models for simulation of metabolism and gain insight into

the complex interactions that give rise to the metabolic capabilities [24, 25]. The predictive accuracy of a model depends on the comprehensiveness and biochemical fidelity of the reconstruction [26].

The GMNR construction process can be synthesized into two fundamental stages: (1) The generation of a draft network reconstruction, here the reactions associated with the enzymes that participate in the metabolism of a particular organism, are downloaded from specialized genome, biochemical and metabolic databases; and (2) a refinement of the network manually or using computational algorithms. Similar steps are performed during the construction of a tissue-specific metabolic reconstruction, defined as a subset of reactions included in a genome-scale metabolic reconstruction that are highly associated with the metabolism of a specific tissue [27]. They are constructed from measured gene expression or proteomic data and permit characterize or predict the metabolic behavior of a tissue under any physiological condition. Due to only the reactions associated with an enzyme or gene can be mapped from the measured data, the spontaneous reactions, and non-facilitated transport reactions are missing in first stages of a tissue-specific reconstruction.

The refinement stage of the reconstruction is a process to restore the connectivity network, where network gaps in the draft reconstruction are identified, and candidate reactions to fill the gap are find in literature and databases [5, 28]. Since the network reconstructions typically involve thousands of metabolic reactions, the refinement of them can be a very complex task [23].The network gaps can be associated with dead-end metabolites which cannot be imported/produced by any of the reactions in the network; or metabolites that are not used as substrates or released by any of the reactions in the network.When the metabolic network is transformed into a metabolic steady-state model to optimize the distribution of metabolic flux under an objective function, the presence of this type of metabolites can be problematic, due to the flux cannot pass through them due to the incomplete connectivity with the rest of the network [28].

In a high-quality model, all reactions should be able to carry flux if all relevant exchange reactions are available [23]. The lack of flux in dead-end metabolites is propagated downstream/upstream, depending if the metabolites are not produced or not consumed, giving rise to additional metabolites that cannot carry any flux [28]. This can block out the net flux of the objective function and when it is not blocked, bias in the biological conclusions increase. The manual refinement is an iterative process to assemble a higher confidence compendium of organism-specific metabolism in a draft metabolic network reconstruction [29–31]. This type refinement requires time and a labor intensive of use of available literature, databases and experimental data [31, 32]. Given GMNR account for hundreds or thousands of biochemical reactions, this task is very complex and can lead to both, the introduction of new errors and to overlook some others.

Table 2-1.: Description and comparison of the methods used in the available 'gapFill' implemented algorithms under 'Python', 'GAMS' and 'MATLAB' environment.

Algorithm	Package	Environment	Method Description
'SMILEY' [33]	COBRApy	Python	Developed by Reed *et al.* (2006), the 'SMILEY' algorithm identify through an *optimization algorithm* the minimum number of reactions required to allow the model a selected *metabolite* production from a universal database of stoichiometric reactions. The process is carried out one metabolite per time (user defined). It is an open source implementation under an open source environment.
'gapFind' and 'gapFill' [34]	-	GAMS	Developed by Kumar *et al.* (2007), the 'gapFind' and 'gapFill' algorithms identify the metabolites ('gapFind') in the metabolic network reconstruction which cannot be produced under any uptake conditions in both single and multi-compartment, and subsequently, identify the reactions ('gapFill') from a customized multi-organism database that restores the connectivity of these metabolites to the original network using a *optimization based* procedures. In the process, the procedures make several intra *model modifications* such as: (1) modify the directionality of the reactions in the model, (2) add fake external transport mechanisms and (3) add fake intracellular transport reactions in multi-compartment models. It is an open source implementation under an open source environment.

Algorithm	Package	Environment	Method Description
'growMatch' [35]	COBRApy	Python	Developed by Kumar *et al.* (2009), the 'growMatch' algorithm identify through an *optimization algorithm* the minimum number of reactions required to allow the model give flux to a selected *objective function* from a universal database of stoichiometric reactions. The process is carried out one objective function per time (user defined). It is an open source implementation under an open source environment.
'fastGapFill' [36]	openCOBRA	MATLAB©	Developed by Thiele *et al.* (2014), the 'fastGapFill' algorithm identify the blocked reactions through an optimization procedure. It searches candidate reactions to fill the gaps in a universal database of stoichiometric reactions through the 'fastCore' algorithm. This second algorithm computes a compact flux consistent model and uses it to filter and determine the reactions to be added. In the filling process, fake transport reactions between compartments are added. It is an open source implementation under a privative environment.
'gapFill'[37]	g2f	R	The 'gapFill' algorithm identify the orphan metabolites (non-consumed or produced by any other reaction) and traces them in a universal database of stoichiometric reactions used as reference to select candidate reactions to be added. Selected reactions are then filtered by a cost algorithm based in the original metabolites present in the reconstruction, to minimize the number of new metabolites to add into the reconstruction. It is an open source implementation under an open source environment.

The metabolic network gap refinement also can be performed using several algorithms under environments such as Python, GAMS and MATLAB. Implemented algorithms are mainly based in optimization procedures that give flux to a selected objective function or a defined metabolite production.

With the aim of offering an open source tool that facilitates the refinement of drafts network reconstructions and the depuration of metabolic models under the R environment, we introduce the **g2f** R package. It includes four functions to identify and fill gaps, as well as, to calculate the addition cost of a reaction and depurate metabolic networks of blocked reactions (no activated under any scenario).

2.2. Installation and functions

The `g2f` package includes four functions and is available for download and installation from CRAN, the Comprehensive R Archive Network. To install and load it, just type:

```
> install.packages("g2f")
> library(g2f)
```

The `g2f` package requires R version 2.10 or higher. Development releases of the package are available on the GitHub repository `http://github.com/gibbslab/g2f`.

Downloading a reference from the KEGG database

The KEGG database is a collection of databases widely used as a reference in genomics, metagenomics, metabolomics and other omics studies, as well as for modeling and simulation in systems biology [38]. At today, the database includes genomes, biological pathways and its associated stoichiometric reactions for 346 eukaryotes, 3947 bacteria, and 238 archaea. The `getReference` function download from the KEGG database all the KO-associated stoichiometric reactions, and their correspondent E.C. numbers for a customized organism (using KEGG organism ID). Based in the KOs associated to a reaction, their respective GPR is construed as follows: All gene associated to a determined KO are linked by an `AND` operator, after that, when a reaction has more than one KO associated, previously linked genes are now joined by an `OR` operator. As an example, to download all (1392 reactions) stoichiometric reactions associated to *Escherichia coli* just type:

```
> E.coli <- getReference(organism = "eco")
```

Calculating the addition cost

The reactions mapping based on metabolites and posterior addition is a very basic solution for gap fill process which increases the number of dead-end metabolites. As a way to reduce the addition of new dead-end metabolites, the `additionCost` function calculates based on metabolites that constituted the new reaction and those that constitute the stoichiometric reactions present in the metabolic reconstruction a cost (in terms of new metabolites) to be added following the equation 2-1.

$$additionCost = \frac{n(metabolites(\texttt{newReaction}) \notin metabolites(\texttt{reactionList}))}{n(metabolites(\texttt{newReaction}))} \quad (2\text{-}1)$$

As an example, we select a sample of reactions from the downloaded reference for *E. coli* and calculate the addition cost for the remaining reactions (6 first values are showed).

```
> reactionList <- sample(E.coli$reaction,10)
> head(
+   additionCost(reaction = E.coli$reaction,
+                reference = reactionList)
+ )

[1] 0.4000000 0.4000000 0.4000000 0.4000000 0.3333333 0.3333333
```

Performing a gap find and fill

To identify network gaps in a metabolic network and fill it from a reference, the `gapFill` function perform internally several steps: (1) The dead-end metabolites are identified from the stoichiometric matrix using functions included in the `minval` package, (2) the candidate reactions to be added are identified by metabolite mapping, (3) the addition cost of each candidate reaction is calculated, (4) the candidate reactions with an addition cost lower or equal that the user-defined limit are added to the reaction list and finally (5) process return to step 1 until no more original-gaps can be filled under the user-defined limit. Function returns a set of candidate stoichiometric reactions to fill the original-gaps included in the metabolic network.

As an example, we show how to fill the dead-end metabolites included in the previously selected sample using all downloaded stoichiometric reactions from the KEGG database for *E. coli* as the reference.

```
> gapFill(reactionList = reactionList,
+         reference = E.coli$reaction,
+         limit = 1/4
+ )

23 Orphan reactants found
12 Orphan reactants found
```

```
11 Orphan reactants found
11 Orphan products found
 [1] "D-Mannonate + NAD+ <=> D-Fructuronate + NADH + H+"
 [2] "ATP + Thymidine <=> ADP + dTMP"
 [3] "dTTP + H2O <=> dTMP + Diphosphate"
 [4] "(R,R)-Tartaric acid + NAD+ <=> 2-Hydroxy-3-oxosuccinate + NADH + H+"
 [5] "Hypoxanthine + NAD+ + H2O <=> Xanthine + NADH + H+"
 [6] "Ammonia + 3 NAD+ + 2 H2O <=> Nitrite + 3 NADH + 3 H+"
 [7] "L-Glutamine + H2O <=> L-Glutamate + Ammonia"
 [8] "ATP + Deamino-NAD+ + Ammonia <=> AMP + Diphosphate + NAD+"
 [9] "Ammonia + NAD+ + H2O <=> Hydroxylamine + NADH + H+"
[10] "ATP + Glutathione + Spermidine <=> ADP + Orthophosphate + Glutathionylspermidine"
[11] "2 Glutathione + NAD+ <=> Glutathione disulfide + NADH + H+"
[12] "Glutathione + H2O <=> Cys-Gly + L-Glutamate"
[13] "L-Proline + NAD+ <=> (S)-1-Pyrroline-5-carboxylate + NADH + H+"
[14] "L-Glutamate 5-semialdehyde + NAD+ + H2O <=> L-Glutamate + NADH + H+"
[15] "ATP + Pantothenate <=> ADP + D-4'-Phosphopantothenate"
```

Identifying blocked reactions

To identify the blocked reactions included in a metabolic model, the `blockedReactions` function set each one of the reactions included in the model (one by time) as the objective function and optimize through Flux Balance Analysis the model. Reactions that not participate in any possible solution during all evaluations are returned as a blocked reaction.

As an example, we identify the blocked reactions in the *E. coli* core metabolic model included in the '`sybil`' package.

```
> data("Ec_core")
> blockedReactions(Ec_core)

  |======================================================================| 100%
[1] "EX_fru(e)"   "EX_fum(e)"   "EX_mal_L(e)" "FUMt2_2"     "MALt2_2"
```

2.3. Summary

We introduced the `g2f` package to find the dead-end metabolites included in a metabolic reconstruction, and fill it from the stoichiometric reactions of a reference, filtering candidate reactions using a weighting function and a user-defined limit. We show step by step the functionality of each procedure included in the package using a reference downloaded from the KEGG database for *E. coli*, and the core metabolic model included in the '`sybil`' package.

3. Constraining a tissue-specific metabolic reconstruction: Incorporating expression data as FBA limits through 'exp2flux' R package.

Original title: exp2flux: Convert Gene EXPression Data to FBA FLUXes.

Written by: *Daniel Osorio, Kelly Botero, Janneth Gonzalez and Andrés Pinzón-Velasco*

Abstract

Computational simulations of metabolism can help to predict the metabolic phenotype of an organism in response to different stimuli, through constraint-based modeling approaches. To recreate specific metabolic phenotypes and enhance the model predictive accuracy, several methods for the integration of transcriptomics data into constraint-based models have been proposed. The majority of available implemented methods are based on the discretization of data to incorporate constraints into the metabolic models via boolean logic representation, which reduces the accurate of physiological representations. The implemented methods for gene-expression data integration as continuous values are very few. The **exp2flux** package was designed as a tool to incorporate in a continuous way the gene-expression data as FBA flux limit in a metabolic reconstruction. Also, a function to calculate the differences between fluxes in different metabolic scenarios was included.

3.1. Introduction

Metabolism is a cellular system suited for developing studies at the systemic level of the genotype-phenotype relationship and genetic interactions [21]. Genome sequencing projects have been contributed to our understanding of the metabolic capabilities in cellular systems since functional annotation of the gene products (enzymes) allow the reconstruction of genome-scale metabolic networks (GMNs), that summarize these metabolic capabilities consistently and compactly in a stoichiometric matrix [39, 40]. The GMNs can be converted

into computational models to predict the metabolic phenotype of an organism in response to different stimuli, through constraint-based modeling approaches [41, 42]. A widely used approach to perform *in silico* metabolic simulations is the flux balance analysis (FBA), a linear optimization method which uses the imposed mass balance and constraints (that represent genetic or environmental conditions) to define the space of feasible steady-state flux distributions of the network and then identify optimal network states that maximize a defined objective function [21, 43].

Given that steady-state simulations assume enzymatic constant rate and does not consider the real expression of each gene or the subcellular localization of gene products, flux constraints based on different "omics" data (such as transcriptomics, proteomics, and metabolomics) must be integrated into the GMNs, in order to recreate specific metabolic phenotypes and enhance the model predictive accuracy [22, 44–46]. Integrative methods have a powerful potential to describe molecular and biochemical mechanisms of organisms metabolism under specific environmental or genetic conditions, as well as, to allow contextualize high-throughput data [47].

Several algorithms and methods have been developed to integrate experimental data at GMNs [48]. Given the increase of gene expression data, these algorithms have focused on incorporate transcriptomics data into constraint-based models [49–51], and in this way constrain the flux distribution (solution space) in GMNs [41, 52]. Each of the algorithms is based on the assumption that mRNA transcript levels are a strong indicator of the level of protein activity [53]. The algorithms differ mainly in the way to integrate expression data, some of the implemented algorithms incorporate data in a discrete or continuous way, using absolute values for a single condition, or the relative expression levels between different conditions [54]. The most of the methods are based on the discretization of data to incorporate constraints into the metabolic models via boolean logic representation (activation/inactivation flux) [53]. Even when continuous integration could be more accurate for physiological representation of the continuous of the reactions activity gradient [46], the methods of continuous integration are very few [54].

With the aim of offer an open source tool that facilitates the integration of gene-expression data as a continuous constraint (FBA limits) in metabolic models, we introduce the 'exp2flux' R package. The exp2flux package incorporates a previously described but not implemented continuous gene-expression data integration method [55, 56]. The implemented method is based on the association of 'omics' data to the genes included in the Gene-Protein-Reaction (GPR) related to each reaction in a genome-scale metabolic model each reaction [5] and considers different possible biological scenarios that occur during the catalysis of biochemical reactions [55]. Also, a function to calculate the difference in reaction fluxes between simulated scenarios was included.

3.2. Installation and functions

`exp2flux` includes two functions and is available for download and installation from CRAN, the Comprehensive R Archive Network. To install and load it, just type:

```
> install.packages("exp2flux")
> library(exp2flux)
```

The `exp2flux` package requires R version 2.10 or higher. Development releases of the package are available on the GitHub repository `http://github.com/gibbslab/exp2flux`.

Inputs

Functions included in **exp2flux** package takes as input two kind of objects, metabolic models as an object of class `modelorg` for the `'sybil'` R package, and gene-expression data as an object of class `ExpressionSet`, a container for high-throughput assays and experimental metadata described in the `'Biobase'` Bioconductor package.

Converting gene expression data to FBA limits

Gene-protein-reaction (GPR) associations indicate which gene has what function into a genome-scale metabolic network and are represented as boolean relationships between genes. This function calculates and assigns the flux boundaries for each reaction based in their associated GPR. Value is obtained as follows: (1) When two genes are associated with an `AND` operator according to the GPR rule, a minimum function is applied to their associated expression values. In the `AND` case, down-regulated genes alter the reaction acting as the enzyme formation limiting, due both are required to the enzymatic complex formation. In turn, (2) when the genes are associated with an `OR` rule, each one of them can code an entire enzyme to act as a reaction catalyst. In this case, a sum function is applied for their associated expression values. To missing gene expression values, the function do a data imputation and assigns one of: `'min'`, `'1q'`,`'mean'`, `'median'`, `'3q'`, or `'max'` expression value calculated from the genes associated to the same metabolic pathway. Metabolic pathway assignation for each reaction is performed through an organism-specific search in the KEGG database. In the case of not possible pathway assignment to a gene, the value to be assigned is calculated from all gene expression values. The fluxes boundaries of exchange reactions are not modified. To show the potential use of data integration through the `exp2flux` function, in this example, we simulate values to represent gene-expression data and integrate it into the *Escherichia coli* core metabolic model included in the `'sybil'` R package.

Integration of gene-expression data begins loading the `'exp2flux'`, `'sybil'` and `'biobase'` required packages.

```
> library(exp2flux)
> library(sybil)
> library(Biobase)
```

After that, the *E. coli* core metabolic model can be loaded from the 'sybil' package. The model includes 95 biochemical reactions associated with 137 genes in 69 GPR rules.

```
> data("Ec_core")
```

Five different measures to represent the gene-expression data for each gene included in the metabolic model was simulated, generated matrix was converted to an `ExpressionSet`.

```
> geneExpression <- ExpressionSet(assayData = matrix(
+   data = runif(n = 5*length(Ec_core@allGenes), min = 0, max = 1000),
+   nrow = length(Ec_core@allGenes),
+   dimnames = list(c(Ec_core@allGenes))
+   ))
> geneExpression
```

```
ExpressionSet (storageMode: lockedEnvironment)
assayData: 137 features, 5 samples
  element names: exprs
protocolData: none
phenoData: none
featureData: none
experimentData: use 'experimentData(object)'
Annotation:
```

Incorporation of gene-expression data into the metabolic model through `exp2flux` function requires two objects, a metabolic model and an `ExpressionSet` as arguments. The `exp2flux` function returns a constrained metabolic model.

```
> mEc_core <- exp2flux(
+   model = Ec_core,
+   expression = geneExpression
+   )
> mEc_core
```

```
model name:             Ecoli_core_model
number of compartments  2
                        C_c
                        C_e
number of reactions:    95
number of metabolites:  72
number of unique genes: 137
objective function:     +1 Biomass_Ecoli_core_w_GAM
```

To visualize the metabolic changes induced by the incorporation of gene-expression data as FBA limits of the reactions included in the metabolic model, a Flux Variability Analysis was performed using the original *E. coli* core metabolic model and the constrained one.

```
> par(mfcol=c(1,2))
> plot(fluxVar(Ec_core),main="Original Model",ylim=c(-100,1000))

|            :            |            :            | 100 %
|=========================================================| :-)

> plot(fluxVar(mEc_core),main="Modified Model",ylim=c(-100,1000))

|            :            |            :            | 100 %
|=========================================================| :-)
```

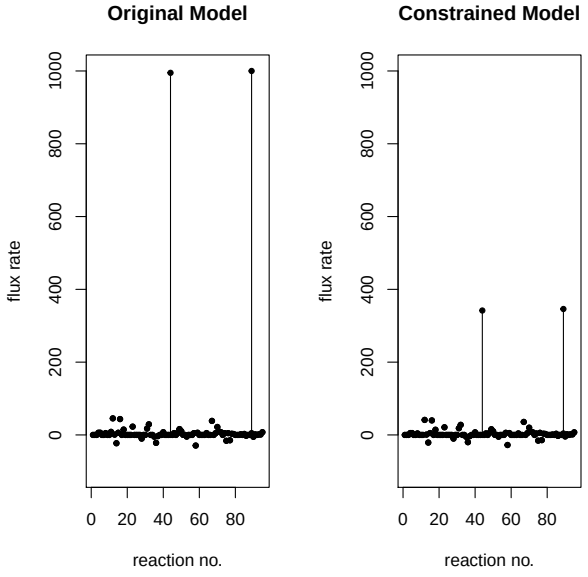

Figure 3-1.: Flux differences between an unconstrained and a constrained model. Constraints were calculated through the `exp2flux` R package using simulated gene expression data.

Identifying flux changes between scenarios

The measurement of flux change for each reaction between metabolic scenarios is a task generally carried out manually and oriented directly to the research objective. However, at system level analysis this process can become laborious. The `fluxDifferences` function calculates the fold change for each common reaction between metabolic scenarios. *Fold change* is a measure that describes how much a quantity changes going from an initial to a final value. Implemented algorithm in the `fluxDifferences` functions are described in equation 4-1, function takes as argument two valid models for the `'sybil'` R package and a customizable

threshold value to filter functions to be reported.

$$foldChange : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

$$(rFluxModel1, rFluxModel2) \mapsto \begin{cases} rFluxModel2, & rFluxModel1 = 0; \\ \dfrac{(rFluxModel2 - rFluxModel1)}{|rFluxModel1|}, & \text{Other cases} \end{cases} \tag{3-1}$$

As an example, we report the fold change of all reactions with an absolute change greatest or equal to 2-fold between the unconstrained and constrained metabolic scenarios simulated previously.

```
> fluxDifferences(
+   model1 = Ec_core,
+   model2 = mEc_core,
+   foldReport = 2
+   )
```

```
           fluxModel1   fluxModel2 foldChange
ADK1     4.547474e-13 1.000444e-11         21
D_LACt2 -6.821210e-13 1.364242e-12          3
LDH_D   -6.821210e-13 1.364242e-12          3
```

3.3. Summary

We introduced the **exp2flux** package, an implementation of a previously described method to integrate gene-expression data in a continuous way into genome-scale metabolic network reconstructions. We show as an example, how the integration of a set of simulated data modifies the behavior of the *E. coli* core metabolic model using Flux Balance Analysis simulations. Also, a example of the measurement of flux change between metabolic scenarios was showed.

4. Exploring the neuroprotective effects of tibolone during astrocytic metabolic inflammation: a flux balance analysis approach

Written by: *Daniel Osorio, Janneth Gonzalez and Andrés Pinzón-Velasco*

Abstract:

Inflammation is a complex biological response to injuries, metabolic disorders or infections and its dysregulation induces many complex diseases through astrocytic dysfunction. The increase of free saturated fatty acid produce a metabolic inflammation response, generally associated with the induction of diverse intracellular stresses, such as mitochondrial oxidative stress, endoplasmic reticulum stress, and autophagy defects. Astrocytes respond to inflammation through a complex reaction called astrogliosis. During astrogliosis, glial cells generally associated with several beneficial activities in the CNS, also act as a source of inflammatory mediators and as generators of ROS that have the potential to damage neurons. In search of compounds with neuroprotective effects that imitate the neuroprotective actions of steroids without their prejudicial side effects; the synthetic neurosteroid Tibolone was identified. Although Tibolone has been shown to exert neuroprotective actions in cultured and under ischemia injury rat neurons, its specific actions on glial cells have received very little attention. Nevertheless, is not well know the effects of tibolone on glial cells that allow its neuroprotective action. In this work, we model and simulate the metabolic inflammation response in mature astrocytes through Flux Balance Analysis (FBA), and explore the neuroprotective effects of tibolone under the inflamed state. We focused on identification of changes in metabolic pathways activation, functional products, gliotransmitter release and the neuroprotective effects mediated by tibolone over inflamed scenario. The generated network consisted of 1262 genes encoding for enzymes performing 2747 reactions distributed across eight compartments, which was studied using a constrained-based modeling approach to recreating three different scenarios in mature astrocytes (healthy, inflamed and medicated), and validated with available experimental evidence. From our analysis, we predict that

tibolone execute their neuroprotective effects through a reduction of neurotoxicity mediated by L-glutamate in astrocytes.

4.1. Introduction

Astrocyte-Neuron Metabolic Relationships

Astrocytes are the most abundant cells in the human brain and play important roles in the central nervous system (CNS) [57]. They are highly associated with several homeostatic functions such as glutamate, ion, and water homeostasis, energy storage in the form of glycogen, synapse formation, and remodeling, defense against oxidative stress, scar formation, tissue repair and modulation of synaptic activity via the release of gliotransmitters [58]. Astrocytes metabolize glucose in the anaerobic way to produce lactate, which is released to neurons through monocarboxylate transporters [59]. Lactate is used in neurons as an energy substrate after its conversion to pyruvate and subsequently to ATP via oxidative phosphorylation [60].

Astrocytes play an important role in glutamate-mediated synaptic activity [61]; according to the astrocyte–neuron lactate shuttle model, astrocytes respond to glutamate-induced activation by increasing their rate of glucose uptake and the release of lactate into the extracellular space, increasing the lactate available to be used by neurons to supply their energetic needs [62]. Glutamate is uptake by astrocytes through the glutamate-aspartate transporter and glial glutamate transporter-1, inducing events that involve the activation of Na^+–K^+-ATPase and maintaining extracellular glutamate at homeostatic levels [63]. Part of incorporated glutamate is converted to glutamine through glutamine synthetase, which is only associated with glial cells and released to neurons using electroneutral systems-N transporters coupled to Na^+ and H^+ [64]. In neurons, glutaminase enzyme converts glutamine back into glutamate which can be used again for neurotransmission or metabolized into the neuronal Krebs cycle [65].

Astrocytes uptake and release many other substances related to synaptic transmission [66]. However D-serine, a neurotransmitter that acts as a co-agonist with glutamate at NMDA receptors is one of the most important [61]. Due in the brain, only glial cells can synthesize serine, all available D-serine at synapsis is associated to be primarily produced and secreted by astrocytes [64]. D-serine is synthesized in astrocytes by serine racemase from L-serine [67]. Serine and glycine are involved in a cycle between astrocytes and neurons similar to the glutamate-glutamine cycle [68]. Additionally, to these energetic and synaptic support associated functions, astrocytes also play an important role in the reduced glutathione (GSH) metabolism of the brain [69]. GSH is the major cellular antioxidant and plays an important neuroprotective role [70]. Cellular GSH levels are closely correlated with cell survival under adverse conditions [71]; it is synthesized from glutamate, cysteine, and glycine and releases

directly from astrocytes through GSH transporters ion-independent in a concentration gradient dependent transport [72].

This strong metabolic cooperation between astrocytes and neurons allows predicting that even a small astrocytic dysfunction might cause and/or contribute neurodegenerative processes [73]. Homeostatic astrocyte function is required for neuronal survival after different brain insults, such as inflammation, glucose deprivation, traumatic brain injury and ischemia [70, 74]. Astrocytes protect neurons of the most important factors that contribute to neuronal cell death such as glutamate-mediated excitotoxicity leading to disturbances in calcium and sodium intracellular metabolism, mitochondrial dysfunction, oxidative stress, cytokines and toxins [57, 58, 63, 75].

Astrocytes response to Inflammation

Inflammation is a complex biological response to injuries, metabolic disorders or infections, and its dysregulation induces many complex diseases through astrocytic dysfunction [70, 76, 77]. In the brain, inflammatory response acts as a defense mechanism against any threat to homeostatic state inducing changes in glucose metabolism and release of pro-inflammatory factors [71]. Inflammation responses in CNS are mediated by glial cells that acquire reactive phenotypes to participate in repair mechanisms [57, 70, 78].

Astrocytes, as glial cells are highly sensitive cells to inflammatory mediators, they respond to inflammation through a complex reaction named astrogliosis [79]. During astrogliosis, glial cells generally associated with several beneficial activities in the CNS, also act as a source of inflammatory mediators and as generators of reactive oxidant species (ROS) that have the potential to damage neurons [80]. Astrogliosis is characterized by a low regulation of mitochondrial dynamics that result in mitochondrial failure [81]. Mitochondrial failure induces the deregulation of Ca^{2+} homeostasis and increased ROS generation, both of which are linked to neurotoxicity [58]. At metabolic level, inflammatory process has been associated with an increase of free saturated fatty acid in comparison with healthy conditions in some brain tissues [82].

The increase of free saturated fatty acid induce metabolic inflammation, a response associated with the induction of diverse intracellular stresses, such as mitochondrial oxidative stress, endoplasmic reticulum stress, and autophagy defects [70]. Lipid excess in metabolic inflammation activates IKKβ and NF-$\kappa\beta$ signaling pathways, which ultimately impairs leptin and insulin hormonal signaling and further triggers the synthesis and release of increased amounts of ROS and pro-inflammatory cytokines (TNF-α and IL-6) from glial cells to sustain the neuroinflammatory state [83]. Enhanced ROS generation by reactive glial cells trigger mitochondria dysfunction in neuron, which induces neuronal apoptosis, the prerequisite for

a diverse number of neurodegenerative conditions [84].

In silico Systems Biology and Inflammation

Inflammatory pathways are evolutionarily conserved, complex, redundant and interconnected [85] . These characteristics difficult each attempt to understand any disease having inflammation as its core using the traditional reductionism-based scientific method and the current regulatory framework [86]. Traditional methods generally focus on single molecules and genes as the targets of study and potential therapy development, nevertheless, mechanistic simulation through a translational systems biology methods allows lead to an understanding of the origin of patterns based on 'omic' data integration in order to facilitate the design of novel therapies [87].

Inflammation is a complex system, which is characterized by sensitivity to initial conditions, positive and negative feedback loops, combined robustness and fragility, and the emergence of nonintuitive behaviors [88]. Translational Systems Biology to inflammation is focused on simulated clinical trials, trying to progress toward personalized diagnostics, personalized medicine, and the rational design of drugs [85].

Tibolone

Drugs as steroids compounds are the most potent and effective agents in controlling chronic inflammatory diseases [89]. However, steroids prescription is limited due to their adverse side effects [90]. Some steroids synthesized in the nervous system, called 'neurosteroids', display beneficial neuroprotective properties, which may be of particular importance in the treatment of diseases where inflammation and neurodegeneration is predominant including age-dependent dementia, stroke, epilepsy, spinal cord injury, Alzheimer's disease (AD) and Parkinson's disease (PD) [91].

Neuroprotective actions of molecules that may imitate the neuroprotective actions of steroids without the prejudicial side effects, such as selective estrogen receptor modulators (SERMs) and selective tissue estrogenic activity regulators (STEARs) have been tested in previous studies [92, 93]. Tibolone is a compound, traditionally used as hormone replacement therapy in post-menopausal women [94], that has been shown neuroprotective effects in cultured and under ischemia injury rat neurons [95].

Tibolone is a synthetic steroid drug with estrogenic, progestogenic, and weak androgenic actions; is metabolized in three compounds, two major active metabolites, 3α-hydroxy tibolone and 3β-hydroxy tibolone acting as potent agonists of the estrogen receptor (ER) and its metabolite Δ4tibolone acting as agonists of the progesterone and androgen receptors

[96]. Tibolone and their metabolites have tissue selective action mechanisms (progestogenic, androgenic and estrogenic) reported in liver, bone, breast and brain according to receptor interaction and activation [92]. Nevertheless, actually, is not well know the effects of tibolone on glial cells that allow its neuroprotective effects [74]. Previous studies have shown that 3-hydroxy-metabolites of tibolone exert agonistic actions on human astrocytes through the activation of estrogen receptors, indicating that astrocytes are a target for tibolone [95].

In this work, we simulate the metabolic inflammatory response in mature astrocytes caused by the increased uptake of palmitate, the most common free saturated fatty acid. We model and simulate the metabolic response using a translational system biology approach called Flux Balance Analysis (FBA) described in methods. We focused on the identification of changes in metabolic pathways activation, functional products, gliotransmitter release and the neuroprotective effects mediated by tibolone in the inflamed scenario.

4.2. Material and Methods

Tissue Specific Model Construction

The tissue specific model construction process started with the identification of all enzyme-coding genes expressed over the mean in at least 50 % of samples for healthy human astrocytes indexed in the GEO database [97] as GSE73721 [98]. Gene identifier conversion from GeneCards[99] to ENTREZ [100] was performed through 'UniProt.ws' R Package [101]. Reactions associated with the identified genes were mapped from the Human Genome-Scale Metabolic Reconstruction RECON 2.04 downloaded from the VMH Lab (https://vmh.uni.lu) [102]. The R package 'g2f' [37] was used to identify and fill the gaps using all no gene-associated reactions included in RECON 2.04, as well as to identify and remove all blocked reactions from the reconstruction.

All reactions involved in the conversion of extracellular glutamate, glycine, cysteine and glucose to extracellular glutamine, glycine, serine-D, reduced glutathione, lactate, and ATP respectively were added. Exchange reactions were limited to components of the Dulbecco's Modified Eagle Medium (DMEM) as input and gliotransmitters (glutamine, D-serine, ATP, glutamate), reduced glutathione, lactate, glucose, nitric oxide, prostaglandins and leukotrienes as output. Finally, syntax, mass-charge validation and creation of SBML files were carried out through the 'minval' R Package [103]. Reaction limits (upper and lower bounds) were constrained proportionally to the mean gene expression reported for genes included in Gene-Protein-Reaction (GPR) [5] associated to each reaction in samples of 47 to 63 years old using the 'exp2flux' R package [104]. All Flux Balance Analysis (FBA) were performed using the 'sybil' [105] R Package running under R 3.3.1 [106].

Flux Balance Analysis

FBA is a linear optimization method for simulating metabolism that allows identifying the set of reactions involved in the production of a biological response within a metabolic model [107]. The metabolic reactions are represented internally as a stoichiometric matrix (S), of size $m \times n$, where m represents the compounds and n the reactions; the entries in the matrix are the stoichiometric coefficients of the metabolites participating in a reaction [108]. The flux through all of the reactions in a network is represented by the vector v, which has a length of n. The concentrations of all metabolites are represented by the vector x, with length m. The systems of mass balance equations at steady state, $\frac{d_x}{d_t} = 0$ or $S \times v = 0$. FBA seeks to maximize or minimize an objective function which can be any linear combination fluxes, to obtain a flux for each reaction, indicating how much each reaction contributes to the objective function [107].

Table 4-1.: Main metabolic capabilities associated to astrocytes represented as the set of objective functions used to evaluate neuroprotective effects of Tibolone under inflamed scenarios

ID	FORMULA REACTION	DESCRIPTION
Glu2Gln	1 glu_L[e] ⇒ 1 gln_L[e]	Glutamate - Glutamine Cycle
Gly2SerD	1 gly[e] ⇒ 1 ser_D[e]	Glycine to D-serine conversion
Glc2Lac	1 glc_D[e] ⇒ 2 lac_L[e]	Lactate production from Glucose
Glc2ATP	1 glc_D[e] ⇒ 36 atp[e]	ATP production from Glucose
Cys2GTHRD	1 cys_L[e] + 1 glu_L[c] + 1 gly[c] ⇒ 1 gthrd[e]	Catch of Cysteine to produce reduced Glutathione

FBA for healthy, inflammated and medicated scenarios was resolved using GLPK 4.60, setting the generic human biomass reaction included in RECON 2.04 and each one of reactions described in table **4-1** as objective functions. Models were analyzed by comparing fluxes between scenarios, metabolites production rate and a sensitivity analysis.

Metabolic Scenarios

To test neuroprotective effects of tibolone during astrocytic metabolic inflammation we define three different metabolic scenarios. A 'healthy' scenario, where palmitate uptake rate was freely set by optimizer; an 'inflamed' scenario, where uptake rate of palmitate was forced to be stable in the mean of the half maximal inhibitory concentration (IC50) value for all objective functions included in table **4-1**. IC50 values were calculated through a robustness analysis performed using uptake of palmitate ('EX_hdca(e)' in RECON 2.04) as control

reaction and 1000 points in the range from 0 to 1 $mMgDW^{-1}h^{-1}$ for each objective function. Uptake value where each objective function reached IC50 was selected and subsequently averaged. Finally, a medicated scenario, defined as an inflamed scenario that includes 279 reactions associated with estradiol-derivated compounds and ten specific reactions associated with Tibolone action mechanism not included in RECON 2.04 described in table **4-2**.

Table 4-2.: Set of reactions added to recreate the medicated scenario model over the astrocyte tissue-specific model. Reactions are the representation of tibolone metabolism in the brain following the reported by Kloosterboer (2004).

ID	FORMULA REACTION	DESCRIPTION
T1	tibolone[e] ⇔	Tibolone exchange reaction
T2	tibolone[e] ⇔ a3OHtibolone[e]	3αhidroxytibolone interconvertion
T3	tibolone[e] ⇔ b3OHtibolone[e]	3βhidroxytibolone interconvertion
T4	tibolone[e] ⇒ d4tibolone[e]	Δ4tibolone isomer formation
T5	b3OHtibolone[e] ⇒ d4tibolone[e]	Δ4tibolone isomer formation from 3β-hidroxytibolone
T6	a3OHtibolone[e] ⇒ estradiol[c]	Estradiol receptor agonist action mechanism of 3α-hidroxytibolone
T7	b3OHtibolone[e] ⇒ estradiol[c]	Estradiol receptor agonist action mechanism of 3β-hidroxytibolone
T8	d4tibolone[e] ⇒ prgstrn[c] + tststerone[c]	Progesterone and androgen receptor activation by tibolone Δ^4 isomer
T9	a3OHtibolone[e] ⇔ a3SOtibolone[e]	3αhidroxytibolone interconvertion to sulfated inactive compounds
T10	a3SOtibolone[e] ⇒	Tibolone inactive form in blood

Metabolic Changes

Metabolic changes across metabolic scenarios were measured through two different approximations. Flux differences for each reaction between optimized scenarios were measured using the fold change calculated as described in equation 4-1.

$$foldChange = \frac{valueModel2 - valueModel1}{|valueModel1|} \tag{4-1}$$

Pro-inflammatory, Anti-inflammatory and Tibolone Action Mechanism Associated Enzymes

Enzymes involved in pro-inflammatory and anti-inflammatory responses as well as in the tibolone action mechanism were identified through sensitivity analysis as follows: Pro-inflammatory enzymes, are those that catalyze reactions that being knocked out allows an increase of objective function value. Anti-inflammatory enzymes are those associated with reactions that have a fold-change greatest equal to 2, and at being knocked out reduce even more the objective function value. Tibolone action mechanism associated enzymes are those that catalyze reactions that being knocked out inhibit entirely the metabolic effect of tibolone.

4.3. Results

All data, code, software and output files used in the developing of this work, are available to be downloaded from GitHub URL: https://github.com/dosorio/masterThesis as a free repository.

Tissue Specific Metabolic Model

Generated astrocyte tissue-specific model describe the metabolism of 1956 compounds in a total of 2747 biochemical reactions associated to 1262 unique genes. Biochemical reactions include 60 exchange and 1080 transport reactions (79 % gene associated, facilitated or active transport) as is shown in figure **4-1A**. To describe the astrocyte tissue-specific metabolic

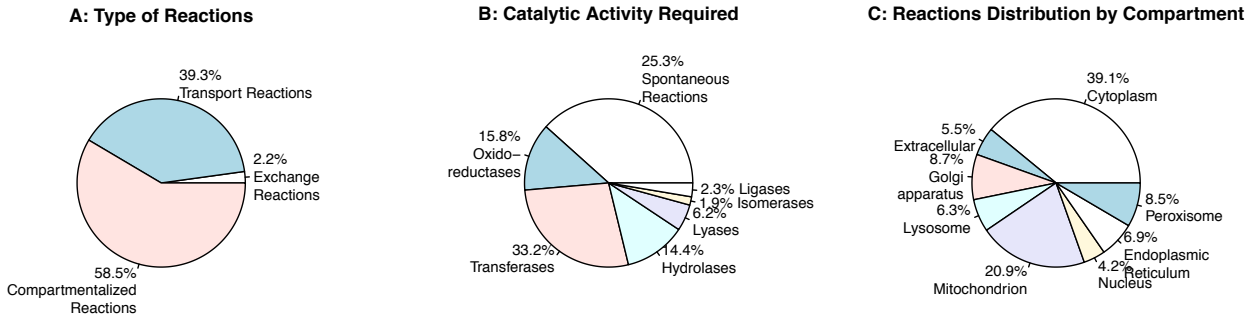

Figure 4-1.: Distribution of biochemical reactions included in the astrocyte tissue-specific metabolic model, classification based in **A:** Type of reaction, **B:** Catalytic activity required and **C:** Associated compartment.

model, reactions were classified on the basis of required enzymatic activity to be catalyzed according to their Enzyme Commission (E.C) numbers (Fig. **4-1B**), sub-cellular locations

according to metabolites compartment (Fig. **4-1C**), and metabolic pathways assigned in the KEGG database (Fig. **4-2**). Based on the associated enzyme to each biochemical reaction 33.2 % of them are catalyzed by a transferase enzyme, 15.8 % by an oxidoreductase, 14.4 % by a hydrolase, 6.2 % by a lyase, 2.3 % by a ligase, 1.9 % by an isomerase enzyme and 25.3 % of them are spontaneous reactions without enzyme or gene associated. In the classification shown in figure **4-1C**, the cytosolic and mitochondrial reactions contributed to 60 % of the total reactions in the model. The other 40 % of reactions are distributed in six other compartments as follows: 8.7 % occurs in Golgi apparatus, 8.5 % in the peroxisome, 6.9 % in the endoplasmic reticulum, 6.3 % in the lysosome, 4.2 % in nucleus, finally 5.5 % of them occurs outside the cell, in the extracellular space. Reactions included in astrocyte model are associated with 113 metabolic pathways reported in the KEGG database [6]. Almost 50 % reactions are associated to 10 main metabolic pathways, highly related to astrocytes metabolism and neuron support metabolic functions [62, 68, 109–112]. Entirely distribution of reactions in metabolic pathways is shown in figure **4-2**.

Healthy Scenario

As previously reported by the Das *et al.* wet lab, healthy human astrocytes grow up in DMEM culture medium [113]. Our metabolic simulation allows to predict an astrocytes slow grow rate (0.37 $\mathrm{mMgWD^{-1}h^{-1}}$) under DMEM medium.

In our healthy scenario (Fig. **4-3**), astrocytes activate the 52 % of model reactions and prefer a glucose-based metabolism, equal than found by Çakir *et al.* and Bhowmick *et al.* in resting conditions [68, 114]. Glucose is catabolized and constitutively released by astrocytes as lactate without any stimuli [115]. This observation is highly expected due astrocytes release large amounts of lactate in the extracellular space which can be used by neurons to supply their energy needs [71]. In our simulations, other gliotransmitters are synthetized and released by astrocytes only under specific stimulus (objective functions) and their release rate was used as a reference to the comparison between scenarios (Fig. **4-6**).

Inflammated Scenario

In the model, the healthy scenario was perturbed to generate the inflamed scenario. The palmitate-induced IC50 value calculated for a set of metabolic functions (Table **4-1**) was 0.208 ± 0.024 $\mathrm{mMgDW^{-1}h^{-1}}$. Calculated IC50 value is the same (0.2 mM) used by Liu *et al.* in wet lab to induce an astrogliosis reaction [116]. Inflammed scenario increases the demand for L-asparagine, L-aspartate, iron, D-glucose, L-glutamate, histidine, L-serine and the release of L-glutamine and lactate (Fig. **4-3**). This response is typical of astrocytes in astrogliosis where neuroinflammation lead homeostatic disturbances [117], such as iron accumulation, within CNS cells [70]. Iron accumulation has been demonstrated in several neurodegenerative diseases as AD, and PD, where it has been postulated to promote disease

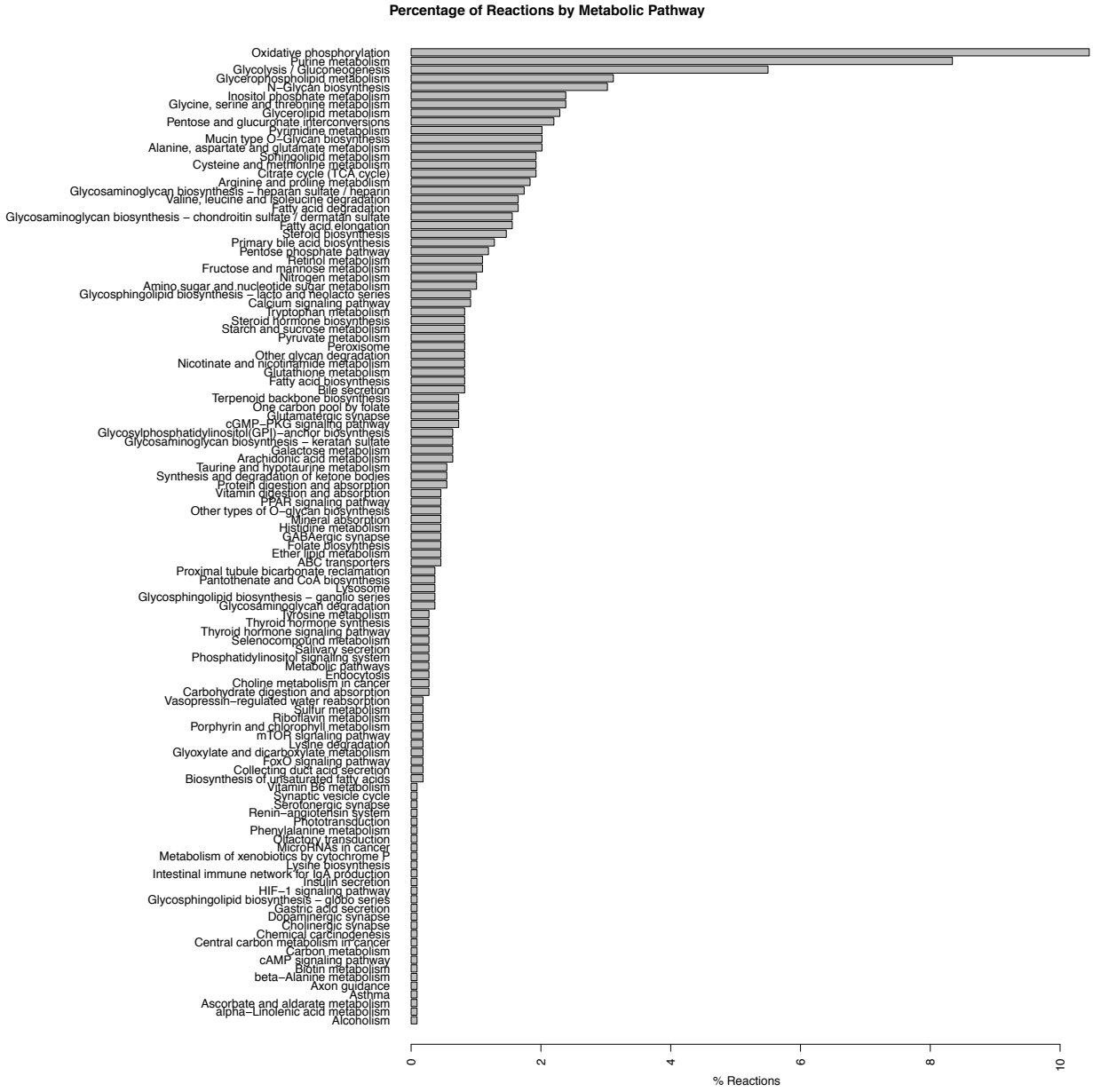

Figure 4-2.: Pathways associated with biochemical reactions included in the astrocyte tissue-specific metabolic model. Pathway association was assigned based in the categorization of the KEGG database.

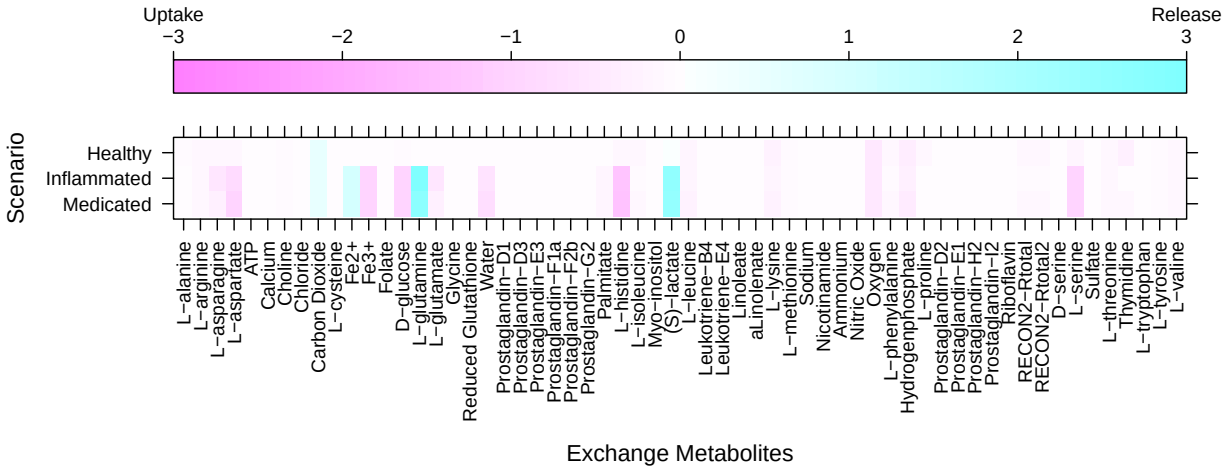

Figure 4-3.: The exchange rate of metabolites between metabolic scenarios using the generic biomass reaction included in RECON 2.04 as the objective function

by augmenting microglial pro-inflammatory activity, altering mitochondrial function, and inducing ROS production [118].

Inflammation although induce the neuronal release of glutamate that may result in the recruitment of neurons in the neuroinflammatory process [119]. Glutamate uptake into astrocytes disinhibits glycolytic enzymes that result in glucose uptake; this glucose is generally processed glycolytically, leading to the synthesis and release of lactate [70]. As neurons cannot generate glutamine from glutamate owing to the lack of the glutamine synthetase enzyme, uptake glutamate is returned to neurons via synaptic clefts in the form of glutamine [120].

Histidine uptake increase was previously reported and suggested as a biomarker of metabolic inflammation [121]; it acts as a free-radicals scavenger and could reduce the levels of IL-6, TNF-α, CRP and inhibit the $H_2O_2^-$ and TNF-α induced by IL-8 secretion [122, 123]. Aspartate, present in the brain as N-Acetyl-L-aspartate (NAA) is synthesized and stored in the neurons but is hydrolyzed in glial cells [124]. NAA act as an anti-proliferation, antiangiogenic, and anti-inflammatory molecule through the decrease of the amount of prostaglandin E2 (PGE2) in astroglial cells [125]. L-Asparagine, in turn, acts as a regulator of ammonia toxicity through the increase of Na^+ intracellular concentration when is co-transported inside astrocytes [126]; asparagine induce a Ca^{2+} response comparable to GABA-induced Ca^{2+} transients in a dose-dependent manner [127].

L-serine and L-asparagine uptake increase may be related to the cell survival process that

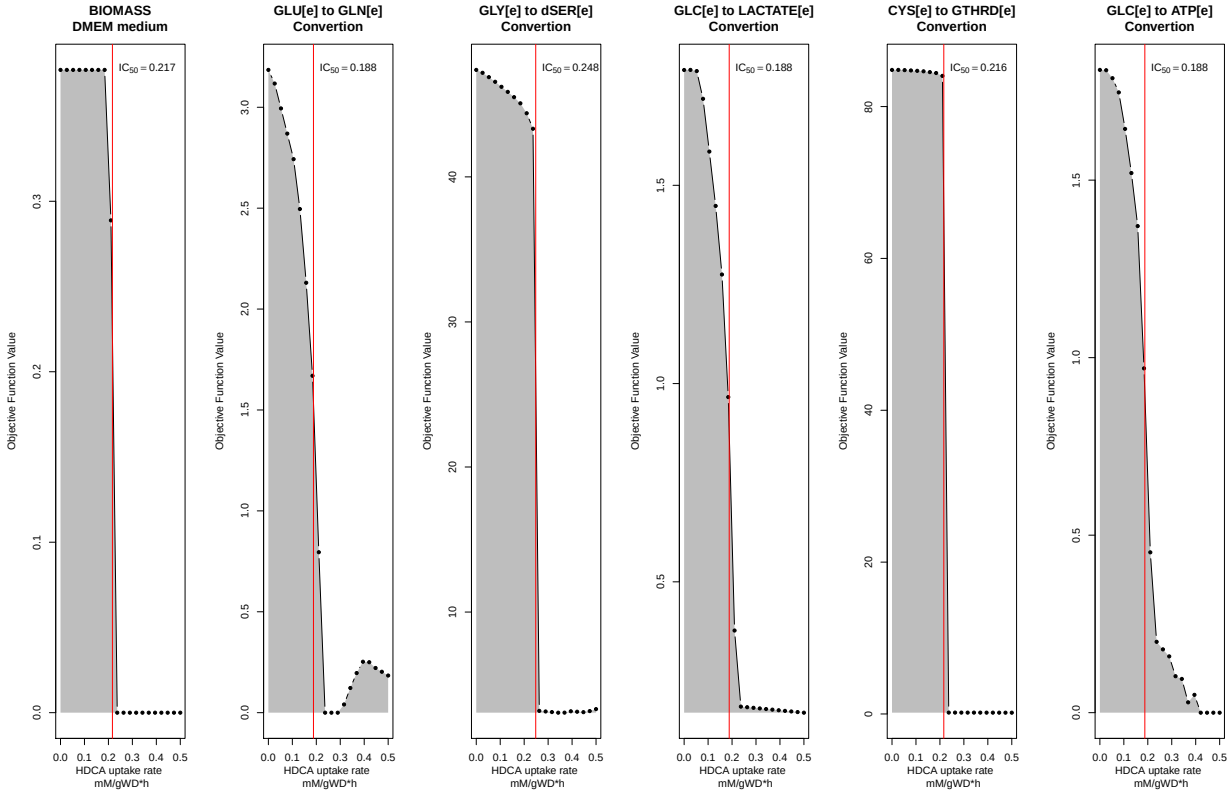

Figure 4-4.: Robustness analysis to calculate palmitate-induced IC50 value for each objective function described in table **4-1**. The red line represents the calculated IC50 value.

switches cellular metabolism to be highly dependent of nonessential amino acids available in extracellular space such as glutamine, serine, glycine, arginine, and asparagine [128]. Moreover, under the inflamed scenario, our astrocyte model release a very small amount of prostaglandin D2. The release of this prostaglandin was previously associated to induce the depolarization and potentiate the actions of simultaneously applied transmitters such as GABA, taurine, glutamate, and aspartate in astrocytes [129].

In our inflamed scenario, astrocytes activate the 46.6 % of model reactions (5.6 % less than healthy scenario) and affect the biomass flux rate of 586 reactions in comparison with the healthy scenario. Main metabolic changes occur in the activation of oxidative phosphorylation, histidine metabolism, and fatty acid degradation pathways; as well as an inactivation of TCA and glycolysis pathways (Fig. **4-5**). Inflammation affects in a negative way all metabolic objective functions evaluated except the release of D-serine. In comparison to the healthy scenario, growth rate over DMEM medium decrease in a 15.6 %, the catch of cysteine to produce reduced glutathione in a 59.3 %, conversion of glucose to ATP in a 72 %, and to

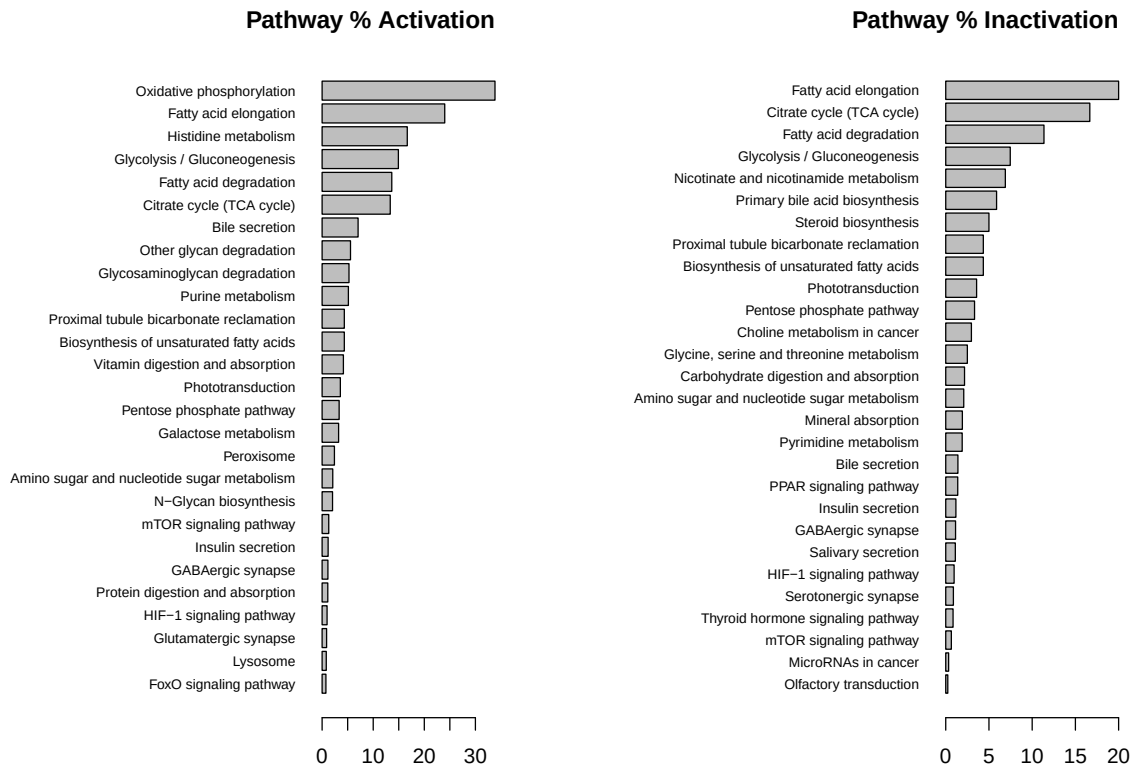

Figure 4-5.: Metabolic pathways affected by metabolic inflammation. Percentage of activation and inactivation was calculated in comparison with genes associated with each pathway in the KEGG database

lactate in a 74.4 %; finally, conversion of extracellular glutamate in glutamine reduces in a 67.7 % (Fig. **4-6**).

Based on performed sensibility analysis, we identify two pro-inflammatory reactions candidate to be knocked out (Table **4-3**), that when being blocked increases the value of the objective function above the maximum value set (in an 11.45 % and 5.14 % respectively) for the inflammated scenario. Reactions are associated to the formimidoyl-transferase cyclodeaminase (FTCD) enzyme and the Aquaporin-8, a water transport protein.

FTCD enzyme, previously reported as over-expressed in high-fat diets [130], contribute with one-carbon units from histidine degradation to the folate pool [131]. In turn, the Aquaporin 8, generally associated to ammonia and water transport [132], has been proposed as a biomarker for inflammation processes where was in contrary way to our observations, highly correlated to cellular defence against severe oxidative stress [133].

Table 4-3.: Set of reactions with pro-inflammatory potential identified through a sensibility analysis over inflamed scenario.

ID	REACTION DESCRIPTION	H. FLUX	I. FLUX	FOLD CHANGE
FTCD	Formimidoyltransferase cyclodeaminase	0.39	1.28	2.28
H2Otm	H2O transport mitochondrial	-0.26	2.44	10.44

Table 4-4.: Set of reactions with anti-inflammatory potential identified through a sensibility analysis over inflamed scenario.

ID	REACTION DESCRIPTION	H. FLUX	I. FLUX	FOLD CHANGE
AKGMALtm	α-ketoglutarate/malate transporter	-0.17	-1.3	-6.85
NADH2_u10m	NADH dehydrogenase mitochondrial	0.12	0.37	2.17
r0639	Lauroyl-CoA: acetyl-CoA C-acyltransferase.	0.02	0.09	4.04
r0653	cMyristoyl-CoA: acetyl-CoA C-myristoyl transferase	0.02	0.09	4.04
r0714	(S)-3-Hydroxyhexadecanoyl-CoA: NAD^+ oxidoreductase	0.02	0.09	4.04
r0716	(S)-3-Hydroxyhexadecanoyl-CoA hydrolyase	0.02	0.09	4.04
r0718	(S)-3-Hydroxytetradecanoyl-CoA: NAD+ oxidoreductase	0.02	0.09	4.04
r0720	(S)-3-Hydroxytetradecanoyl-CoA hydrolyase	0.02	0.09	4.04

As well as pro-inflammatory reactions, eight anti-inflammatory reactions were identified. Identified reactions (Table **4-4**) have a change between scenarios greatest equal to 2-fold and when being blocked decreased, even more, the value of the objective function in comparison with the healthy scenario. The majority of identified reactions (r0639, r0653, r0714, r0716, r0718 and r0720) are involved in the fatty acid elongation in mitochondria through acyl-COA association [134]. The elongation system, is responsible for the addition of two carbon units to the carboxyl end of a fatty acid chain, and play an important role in the maintenance of membrane lipid composition as well as in the generation of precursors for cell signaling molecules (such as eicosanoids and sphingosine-1 phosphate), energy production, and other unknown pathways involving with cancer growth. [135].

Medicated Scenario

The inflamed scenario with the 279 reactions associated with tibolone and estradiol-derivated compounds metabolism, defined as our medicated scenario displays several neuroprotective effects. Medicated scenario increase the demand of L-aspartate and in turn decreases the demand for L-asparagine, L-glutamate and the release of L-glutamine in comparison to the inflamed scenario (Fig. **4-3**).

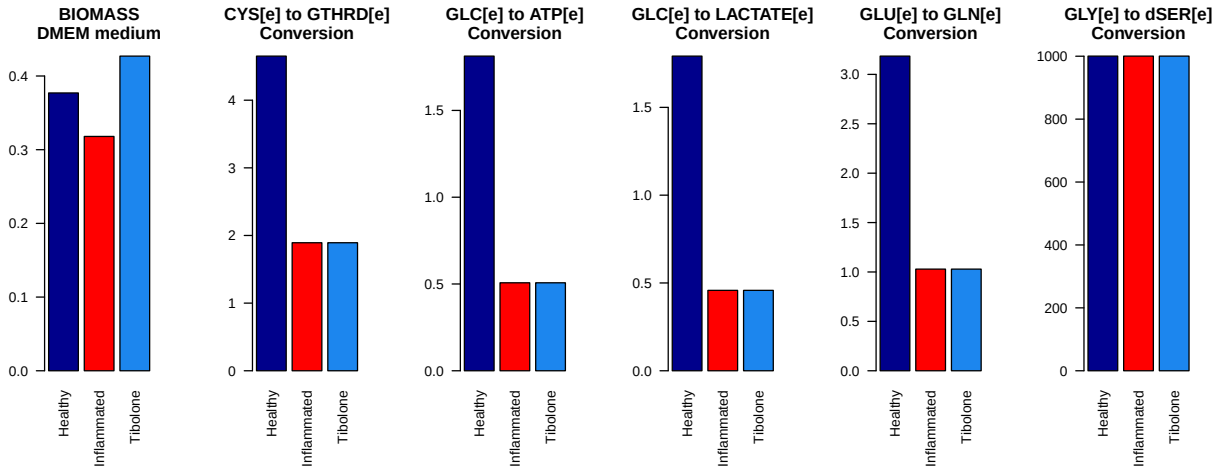

Figure 4-6.: The response of the main astrocytes metabolic capabilities to different modeled scenarios.

The reduction of L-glutamate and L-glutamine uptake/release rate mediated by Tibolone effects could be associated with a neuroprotective effect through a reduction of neurotoxicity mediated by L-glutamate in astrocytes [66]. L-glutamate is a contributing factor in neuronal damage induced by inflammation, traumatic brain injury, stroke, and in most of the chronic neurodegenerative diseases, such as PD and AD [136].

In our medicated scenario, astrocytes activate 46.6 % of model reactions, equal than in inflamed scenario. Nevertheless, tibolone effects affect the biomass flux rate of 948 reactions in comparison with the inflamed scenario; main metabolic changes occurs through the activation of several metabolic pathways with neuroprotective actions associated such as taurine metabolism, which has been shown to be tissue-protective in many models of oxidant-induced injury [137], gluconeogenesis which is accelerated and facilitate the conversion of fatty acids into ketone bodies under steroid-mediated effects [138], calcium and PPAR signaling pathways (Fig. **4-7**).

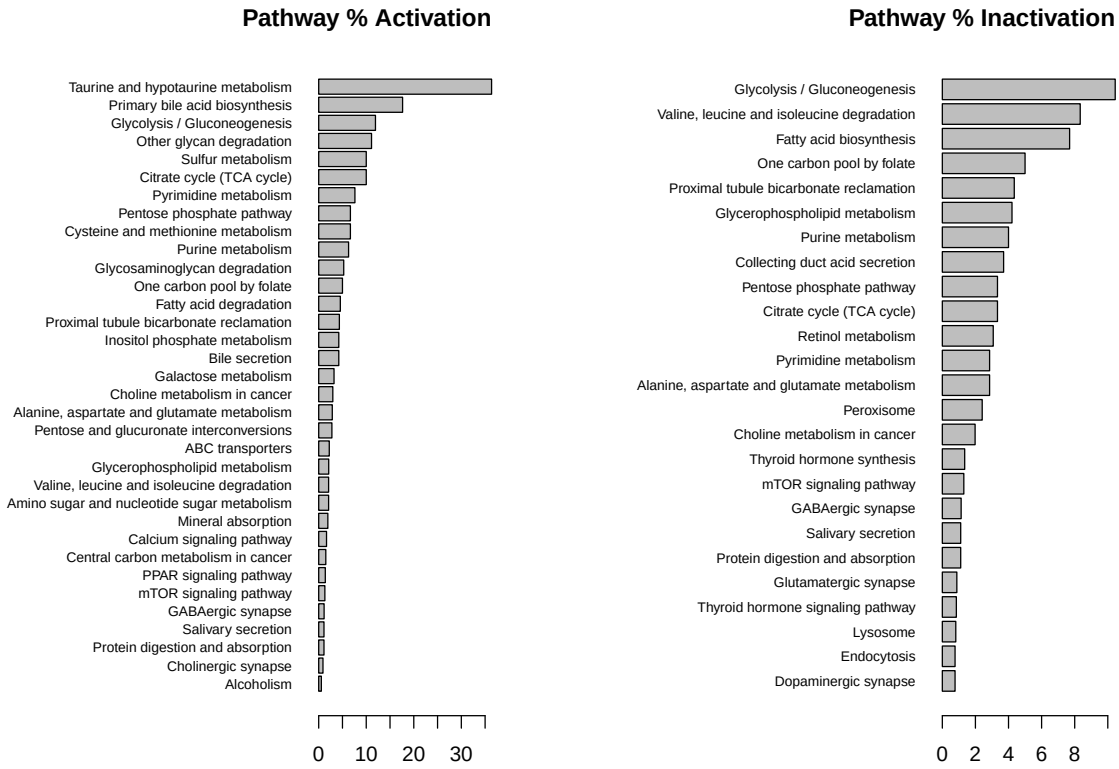

Figure 4-7.: Metabolic pathways affected by tibolone effects over inflamed scenario. Activation and inactivation percentage was measured in comparison with genes associated to each pathway in the KEGG database

Tibolone does not show directly effect over the main associated neuron-supportive capabilities (Table **4-1**) by affected by inflammation, however, it shows a high related activity as a potentiator of growth rate (Fig. **4-6**). An increase of growth rate (13.26 % higher than in healthy scenario) could be associated with an increase in cell viability or as an increase in the proliferative potential in cells [139]. Due that proliferative potential was not evidenced in the inflamed scenario, this proliferative potential must not be associated with astrogliosis mediated proliferation in astrocytes [57]. However, it could be associated with the proliferative side effect of steroid compounds as was previously reported in other cell types where tibolone was tested in wet lab [140, 141].

Based on performed sensibility analysis over the 289 reactions associated with tibolone and estradiol-derivated compounds metabolism, we identify a set of four reactions that when being individually knocked out, block entirely the tibolone effects (Table **4-2**). Identified reactions are catalyzed by an alcohol dehydrogenase (E.C. 1.1.1.1) and a Cytochrome P450 associated with the PPAR signaling pathway. Both enzymes were previously reported with

Table 4-5.: Set of reactions associated with tibolone required to execute its neuroprotective effects. Reactions were identified through a sensibility analysis over the medicated scenario.

ID	REACTION DESCRIPTION	GENES IN ASTROCYTE DATA
r0739	Alcohol Dehydrogenase	ADH4, ADH5, ADH7
r2518	ATP-binding Cassette (ABC)	ABCD3
RE1804M	Cholestanetriol 26-monooxygenase	CYP27A1
RE1807M	Cholestanetriol 26-monooxygenase	CYP27A1

ROS reduction through redox reactions mediated by alcohol dehydrogenase (ADH) and posterior release associated to a cytochrome P450 [142, 143].

4.4. Conclusion

In this work, a tissue-specific metabolic network for mature astrocyte has been developed, and three different scenarios were modeled. Modeled scenarios allowed identify the metabolic changes between a healthy and an inflamed scenario as well as from an inflamed to a tibolone medicated scenario. The model was capable of yielding results which were in correspondence to the experimentally proved metabolic processes [113, 116]. From our study, adverse effects associated with the increase of palmitate uptake were described based on exchange, metabolite production, and metabolic pathways perturbed under inflammatory response. Sensibility analysis performed through constrained-based modeling approach and FBA methods permitted recognize two possible reactions and their associated enzymes susceptible to be knocked out to reduce inflammatory processes.

Based on literature reports we modeled a tibolone medicated scenario used to identify and describe the neuroprotective effects of this synthetic neurosteroid under an inflamed scenario in astrocytes; our main results suggest that tibolone execute their neuroprotective effects through a reduction of neurotoxicity mediated by L-glutamate in astrocytes. L-glutamate [66]. We also found a tibolone associated increase in growth rate probably in concordance to previously reported side effects of steroids in other human cell types [140, 141]. Identified enzyme associated reactions with tibolone effects and their action mechanisms are highly consistent with reported previously by our associated wet lab [74, 144].

A. MINVAL algorithms

MINVAL algorithm 1: isValidSyntax

input : A set of stoichiometric reaction with the following format: `H2O[c] + Urea-1-carboxylate[c] <=> 2 CO2[c] + 2 NH3[c]`. Where arrows and plus signs are surrounded by a "space character". It is also expected that stoichiometry coefficients are surrounded by spaces. It also expects arrows to be in the form => or <=>. Meaning that arrows like ==>, <==>, -> or -> will not be parsed and will lead to errors.

output: A boolean value `TRUE` or `FALSE` for each stoichiometric reaction

```
foreach stoichiometric reaction do
    create an empty boolean vector V;
    if each metabolite has only just one coefficient then
        add TRUE to V
    else
        add FALSE to V
    if metabolites coefficients are not surrounded by parentheses then
        add TRUE to V
    else
        add FALSE to V
    if arrow symbol is between blank spaces then
        add TRUE to V
    else
        add FALSE to V
    if arrow symbol is <=> or => then
        add TRUE to V
    else
        add FALSE to V
    if metabolites names are separated by a plus symbol (+) between blank spaces then
        add TRUE to V
    else
        add FALSE to V
    if substituents position are joined by an hyphen to the metabolite name then
        add TRUE to V
    else
        add FALSE to V
    if all elements of V are TRUE then
        return TRUE
    else
        return FALSE
```

MINVAL algorithm 2: isBalanced

input :

reactionList: A set of stoichiometric reaction with the following format: `H2O[c] + Urea-1-carboxylate[c] <=> 2 CO2[c] + 2 NH3[c]`. Where arrows and plus signs are surrounded by a "space character". It is also expected that stoichiometry coefficients are surrounded by spaces. It also expects arrows to be in the form => or <=>. Meaning that arrows like ==>, <==>, -> or -> will not be parsed and will lead to errors.

referenceData: A chemical table containing data to evaluate the balance

ids: A mandatory id of metabolites id column in the referenceData

mFormula: An optional id of molecular formula column in the referenceData

mWeight: An optional id of molecular weight column in the referenceData

mCharge: An optional id of net charge column in the referenceData

output: A boolean value `TRUE` or `FALSE` for each stoichiometric reaction

```
1  foreach stoichiometric reaction do
2      extract reactants ;                                   /* Applying regular expressions */
3      extract reactants coefficients ;                      /* Applying regular expressions */
4      multiply reactants by coefficients;
5      map reactants in referenceData using the given id;
6      if all reactants were mapped then
7          if mFormula is given then
8              split molecular formulas and sum all atoms ;  /* Applying regular expressions */
9          else
10             if mWeight is given then
11                 sum all molecular weights
12             else
13                 if mCharge is given then
14                     sum all molecular charges
15                 else
16                     return NA
17     else
18         return NA
19     extract products ;                                    /* Applying regular expressions */
20     extract products coefficients ;                       /* Applying regular expressions */
21     multiply products by coefficients;
22     map products in referenceData using the given id;
23     if all products were mapped then
24         if mFormula is given then
25             split molecular formulas and sum all atoms ;  /* Applying regular expressions */
26         else
27             if mWeight is given then
28                 sum all molecular weights
29             else
30                 if mCharge is given then
31                     sum all molecular charges
32                 else
33                     return NA
34     else
35         return NA
36 if sum of reactants is equal to sum of products then
37     return TRUE
38 else
39     return FALSE
```

MINVAL algorithm 3: orphanReactants

input : A set of stoichiometric reaction with the following format: `H2O[c] + Urea-1-carboxylate[c] <=> 2 CO2[c] + 2 NH3[c]`. Where arrows and plus signs are surrounded by a "space character". It is also expected that stoichiometry coefficients are surrounded by spaces. It also expects arrows to be in the form => or <=>. Meaning that arrows like ==>, <==>, -> or -> will not be parsed and will lead to errors.

output: A set of metabolites that are not produced in any other reaction or just are involved in just one reaction.

```
create three empty vectors m, r and p;
foreach stoichiometric reaction do
    split stoichiometric reaction by arrow (<=>) and plus symbol (+);
    remove stoichiometric coefficients;
    remove all blank spaces;
    add metabolites to m;
compute the absolute frequency for each metabolite;
remove from m vector all metabolites with frequencies greater or equal than 2;
for all irreversible reactions do
    r ← extract reactants ;                          /* Applying regular expressions */
    p ← extract products ;                           /* Applying regular expressions */
return elements of ((r ∉ p) ∪ m)
```

MINVAL algorithm 4: orphanProducts

input : A set of stoichiometric reaction with the following format: `H2O[c] + Urea-1-carboxylate[c] <=> 2 CO2[c] + 2 NH3[c]`. Where arrows and plus signs are surrounded by a "space character". It is also expected that stoichiometry coefficients are surrounded by spaces. It also expects arrows to be in the form => or <=>. Meaning that arrows like ==>, <==>, -> or -> will not be parsed and will lead to errors.

output: A set of metabolites that are not produced in any other reaction or just are involved in just one reaction.

```
create three empty vectors m, r and p;
foreach stoichiometric reaction do
    split stoichiometric reaction by arrow (<=>) and plus symbol (+);
    remove stoichiometric coefficients;
    remove all blank spaces;
    add metabolites to m;
compute the absolute frequency for each metabolite;
remove from m vector all metabolites with frequencies greater or equal than 2;
for all irreversible reactions do
    r ← extract reactants ;                          /* Applying regular expressions */
    p ← extract products ;                           /* Applying regular expressions */
return elements of ((p ∉ r) ∪ m)
```

B. G2F algorithms

G2F algorithm 1: additionCost

input :

reaction: A stoichiometric reaction with the following format: `H2O[c] + Urea-1-carboxylate[c] <=> 2 CO2[c] + 2 NH3[c].` Where arrows and plus signs are surrounded by a "space character". It is also expected that stoichiometry coefficients are surrounded by spaces. It also expects arrows to be in the form => or <=>. Meaning that arrows like ==>, <==>, -> or -> will not be parsed and will lead to errors.

reference: A set of stoichiometric reactions with the same format of reaction

output: The addition cost of a stoichiometric reaction based in the metabolites that compound a reference

1 refMet ← extract all metabolites from reference ; `/* Applying regular expressions */`
2 rxnMet ← extract all metabolites from reaction ; `/* Applying regular expressions */`
3 **return** $\frac{(\mid \text{rxnMet} \mid - \mid \text{rxnMet} \in \text{refMet} \mid)}{\mid \text{rxnMet} \mid}$

G2F algorithm 2: blockedReactions

input : A valid model for the '`sybil`' R package. An object of class `modelorg`.
output: A set of ID's associated to reactions without flux under all scenarios.

1 create a empty vector F;
2 **foreach** *reaction in the model* **do**
3 set all objective coefficients as 0;
4 assign selected reaction as objective function;
5 optimize the model;
6 identify reactions with flux different to 0;
7 add IDs of identified reactions to F;
8 **return** model reaction IDs $\notin$ F

G2F algorithm 3: gapFill

input :

reactionList: A set of stoichiometric reactions with the following format: `H2O[c] + Urea-1-carboxylate[c] <=> 2 CO2[c] + 2 NH3[c]`. Where arrows and plus signs are surrounded by a "space character". It is also expected that stoichiometry coefficients are surrounded by spaces. It also expects arrows to be in the form => or <=>. Meaning that arrows like ==>, <==>, -> or -> will not be parsed and will lead to errors.

reference: A set of stoichiometric reaction with the same format of `reactionList`

limit: An addition cost value to be used as a limit to select reactions to be added. Is calculated as $\frac{NumberNewMetabolites}{NumerOfMetabolites}$ for each reaction.

woCompartment: A boolean value `TRUE` to define if compartment labels should be removed of the `reactionList` stoichiometric reactions, `FALSE` is used as default.

consensus: A boolean value `TRUE` to define if reactionList and newReactions should be reported as a unique vector or `FALSE` if just newReactions should be reported.

output: A set of stoichiometric reactions that fill the model gaps.

if *woCompartment is TRUE* **then**
  remove compartments of `reactionList` and `reference` metabolites ; `/* Applying regular expressions */`
`orphanOriginal` ← `orphanReactants`(*reactionList*) ∪ `orphanProducts`(*reactionList*);
to.add ← `reactionList`;
do
  Compute `additionCost`(*reference*);
  Select stoichiometric reactions with `additionCost`(*reference*) ≤ `limit`;
  Extract `orphanReactants`(*reactionList*);
  Compute |`orphanReactants`(*reactionList*) ∈ `orphanOriginal`|;
  Add stoichiometric reaction that include the `orphanReactants`(*reactionList*) within selected stoichiometric reactions with `additionCost`(*reference*) ≤ `limit` to `to.add` array;
while |*orphanOriginal* ∈ `orphanReactants` *(reactionList)*| ≥ |*orphanOriginal* ∈ `orphanReactants`(*reactionList* ∪ *to.add*)|;
`newReactions` ← `unique`(*to.add*);
do
  Compute `additionCost`(*reference*);
  Select stoichiometric reactions with `additionCost`(*reference*) ≤ `limit`;
  Extract `orphanProducts`(*reactionList*);
  Compute |`orphanProducts`(*reactionList*) ∈ `orphanOriginal`|;
  Add stoichiometric reaction that include the `orphanProducts`(*reactionList*) within selected stoichiometric reactions with `additionCost`(*reference*) ≤ `limit` to `to.add` array;
while |*orphanOriginal* ∈ `orphanProducts` *(reactionList)*| ≥ |*orphanOriginal* ∈ `orphanProducts(unique`(*reactionList* ∪ *to.add*))|;
`newReactions` ← `unique`(*to.add* ∪ *newReactions*) ∉ `reactionList`;
if *consensus is TRUE* **then**
  return `reactionList` ∪ `newReactions`
else
  return `newReactions`

C. EXP2FLUX algorithms

EXP2FLUX algorithm 1: fluxDifferences

input :

model1: A valid model for the 'sybil' package.

model2: A valid model for the 'sybil' package. Must have the same reactions (reaction number and reaction identifiers) as "model1" with different restrictions.

foldReport: A threshold value to be reported. All reactions with a greater or equal fold change than the given threshold are reported.

output: The calculated fold change for the fluxes of two given metabolic models.

```
1 fluxModel1 ← extract fluxDistribution(optimize(model1));
2 fluxModel2 ← extract fluxDistribution(optimize(model2));
3 foreach reaction ∈ (model1 ∩ model2) do
4     if fluxModel1_reaction is equal to 0 then
5         return fluxModel2_reaction
6     else
7         if (1-(fluxModel2_reaction/fluxModel1_reaction ≤ foldReport) then
8             return 1-(fluxModel2_reaction/fluxModel1_reaction)
```

EXP2FLUX algorithm 2: exp2flux

input :

model: A valid model for the 'sybil' package.

expression: A valid ExpressionSet object (one by treatment).

organism: A valid organism identifier for the KEGG database. List of valid organism identifiers are available in: http://rest.kegg.jp/list/organism.

typeID: A string to define the type of ID used in GPR's. One of entrez or kegg must be given

missing: A character string specifying the value to be used in missing cases; must be one of min, 1q, mean, median, 3q, or max.

scale: A boolean value to specify if data must be scaled to assign a value of 1000 as max.

output:

```
1  if organism and typeID is given then
2      Download from the KEGG database the metabolic pathways associated to each gene
3  foreach non exchange reaction ∈ model do
4      if GPR is given then
5          split associated GPR by 'or' conectors to extract complex;
6          foreach complex do
7              split complex by the 'and' conector to extract associated genes;
8              foreach gene ∈ complex do
9                  if gene ∈ ExpressionSet then
10                     extract associated expression value from the ExpressionSet;
11                     return expression value for selected gene
12                 else
13                     if organism and typeID is given then
14                         identify associated pathways to gene;
15                         select main pathway (more genes present in the model);
16                         extract associated expression values to main pathway;
17                         compute the selected metric for missing to all expression data associated to the selected
                             pathway
18                     else
19                         return 0
20             find the minimal expression value of all genes associated in the complex;
21             return identified min value
22         compute sum the minimal expression values associated to all genes included in the complex that conform the
             GPR;
23         return computed sum
24     else
25         compute the selected metric for missing to all expression data included in the ExpressionSet;
26         return computed value
27     replace lower.bound with computed value multiplied by -1;
28     replace upper.bound with computed value
29 return the model with lower and upper bound modified
```

Bibliography

[1] William Lingran Chen, David Z Chen, and Keith T Taylor. Automatic reaction mapping and reaction center detection. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(6):560–593, 2013.

[2] James B Hendrickson. Comprehensive System for Classification and Nomenclature of Organic Reactions. *Journal of Chemical Information and Computer Sciences*, 37(97):850–852, 1997.

[3] A. Lambert, J. Dubois, and R. Bourqui. Pathway preserving representation of metabolic networks. *Computer Graphics Forum*, 30(3):1021–1030, 2011.

[4] Jong Myoung Park, Tae Yong Kim, and Sang Yup Lee. Constraints-based genome-scale metabolic simulation for systems metabolic engineering. *Biotechnology Advances*, 27(6):979–988, 2009.

[5] Ines Thiele and Bernhard Ø Palsson. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, 5(1):93–121, 2010.

[6] M Kanehisa. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28(1):27–30, 2000.

[7] Ron Caspi, Tomer Altman, Richard Billington, Kate Dreher, Hartmut Foerster, Carol A. Fulcher, Timothy A. Holland, Ingrid M. Keseler, Anamika Kothari, Aya Kubo, Markus Krummenacker, Mario Latendresse, Lukas A. Mueller, Quang Ong, Suzanne Paley, Pallavi Subhraveti, Daniel S. Weaver, Deepika Weerasinghe, Peifen Zhang, and Peter D. Karp. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Research*, 42(D1):D459–D471, 2014.

[8] David Croft, Antonio Fabregat Mundo, Robin Haw, Marija Milacic, Joel Weiser, Guanming Wu, Michael Caudy, Phani Garapati, Marc Gillespie, Maulik R. Kamdar, Bijay Jassal, Steven Jupe, Lisa Matthews, Bruce May, Stanislav Palatnik, Karen Rothfels, Veronica Shamovsky, Heeyeon Song, Mark Williams, Ewan Birney, Henning Hermjakob, Lincoln Stein, and Peter D'Eustachio. The Reactome pathway knowledgebase. *Nucleic Acids Research*, 42(D1):D472–D477, 2014.

[9] A. Chang, I. Schomburg, S. Placzek, L. Jeske, M. Ulbrich, M. Xiao, C. W. Sensen, and D. Schomburg. BRENDA in 2015: exciting developments in its 25th year of existence. *Nucleic Acids Research*, 43(D1):D439–D446, 2015.

[10] Timothy Jewison, Yilu Su, Fatemeh Miri Disfany, Yongjie Liang, Craig Knox, Adam Maciejewski, Jenna Poelzer, Jessica Huynh, You Zhou, David Arndt, Yannick Djoumbou, Yifeng Liu, Lu Deng, An Chi Guo, Beomsoo Han, Allison Pon, Michael Wilson, Shahrzad Rafatnia, Philip Liu, and David S. Wishart. SMPDB 2.0: Big Improvements to the Small Molecule Pathway Database. *Nucleic Acids Research*, 42(D1):D478–D484, 2014.

[11] A. Gevorgyan, M. G. Poolman, and D. a. Fell. Detection of stoichiometric inconsistencies in biomolecular models. *Bioinformatics*, 24(19):2245–2251, 2008.

[12] Meiyappan Lakshmanan, Geoffrey Koh, Bevan K S Chung, and D.-Y. Lee. Software applications for flux balance analysis. *Briefings in Bioinformatics*, 15(1):108–122, 2014.

[13] Hyun Uk Kim, Tae Yong Kim, and Sang Yup Lee. Metabolic flux analysis and metabolic engineering of microorganisms. *Mol. BioSyst.*, 4(2):113–120, 2008.

[14] Ed Reznik, Pankaj Mehta, and Daniel Segrè. Flux Imbalance Analysis and the Sensitivity of Cellular Growth to Changes in Metabolite Pools. *PLoS Computational Biology*, 9(8):e1003195, 2013.

[15] Nelson E. Vega-Vela, Cynthia Jimenez, George E Barreto, and Janneth Gonzalez. Metabolic Reconstruction of Glutamate-Glutamine Cycling: A Flux Balance Approach. In *Frontiers in Cellular Neuroscience*, volume 9, 2015.

[16] Scott A Becker, Adam M Feist, Monica L Mo, Gregory Hannum, Bernhard Ø Palsson, and Markus J Herrgard. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nature protocols*, 2(3):727–38, 2007.

[17] Rasmus Agren, Liming Liu, Saeed Shoaie, Wanwipa Vongsangnak, Intawat Nookaew, and Jens Nielsen. The RAVEN Toolbox and Its Use for Generating a Genome-scale Metabolic Model for Penicillium chrysogenum. *PLoS Computational Biology*, 9(3), 2013.

[18] Kirill Degtyarenko, P. de Matos, Marcus Ennis, Janna Hastings, Martin Zbinden, A. McNaught, R. Alcantara, Michael Darsow, Mickaël Guedj, and Michael Ashburner. ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Research*, 36(Database):D344–D350, 2007.

[19] T. Bernard, A. Bridge, A. Morgat, S. Moretti, I. Xenarios, and M. Pagni. Reconciliation of metabolites and biochemical reactions for metabolic networks. *Briefings in Bioinformatics*, 15(1):123–135, 2014.

[20] A. Ravikrishnan and K. Raman. Critical assessment of genome-scale metabolic networks: the need for a unified standard. *Briefings in Bioinformatics*, 16(6):1057–1068, 2015.

[21] Balázs Szappanos, Károly Kovács, Béla Szamecz, Frantisek Honti, Michael Costanzo, Anastasia Baryshnikova, Gabriel Gelius-Dietrich, Martin J Lercher, Márk Jelasity, Chad L Myers, and Others. An integrated approach to characterize genetic interaction networks in yeast metabolism. *Nature genetics*, 43(7):656–662, 2011.

[22] Ning Chen, Ioscani Jimenez del Val, Sarantos Kyriakopoulos, Karen M Polizzi, and Cleo Kontoravdi. Metabolic network reconstruction: advances in in silico interpretation of analytical information. *Current opinion in biotechnology*, 23(1):77–82, 2012.

[23] Rasmus Agren, Liming Liu, Saeed Shoaie, Wanwipa Vongsangnak, Intawat Nookaew, and Jens Nielsen. The RAVEN toolbox and its use for generating a genome-scale metabolic model for Penicillium chrysogenum. *PLoS Comput Biol*, 9(3):e1002980, 2013.

[24] Hal Alper, Yong-Su Jin, J F Moxley, and G Stephanopoulos. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in Escherichia coli. *Metabolic engineering*, 7(3):155–164, 2005.

[25] Stephen S Fong, Anthony P Burgard, Christopher D Herring, Eric M Knight, Frederick R Blattner, Costas D Maranas, and Bernhard O Palsson. In silico design and adaptive evolution of Escherichia coli for production of lactic acid. *Biotechnology and bioengineering*, 91(5):643–648, 2005.

[26] I. Thiele, N. Vlassis, and R. M. T. Fleming. FASTGAPFILL: efficient gap filling in metabolic networks. *Bioinformatics*, 30(17):2529–2531, 2014.

[27] Bernhard Palsson. Metabolic systems biology. *FEBS Letters*, 583(24):3900–3904, 2009.

[28] Vinay Satish Kumar, Madhukar S Dasika, and Costas D Maranas. Optimization based automated curation of metabolic reconstructions. *BMC bioinformatics*, 8(1):1, 2007.

[29] Doug Howe, Maria Costanzo, Petra Fey, Takashi Gojobori, Linda Hannick, Winston Hide, David P Hill, Renate Kania, Mary Schaeffer, Susan St Pierre, and Others. Big data: The future of biocuration. *Nature*, 455(7209):47–50, 2008.

[30] Alex Bateman. Curators of the world unite: the International Society of Biocuration. *Bioinformatics*, 26(8):991, 2010.

[31] Benjamin D Heavner and Nathan D Price. Transparency in metabolic network reconstruction enables scalable biological discovery. *Current opinion in biotechnology*, 34:105–109, 2015.

[32] Meiyappan Lakshmanan, Geoffrey Koh, Bevan K S Chung, and Dong-Yup Lee. Software applications for flux balance analysis. *Briefings in bioinformatics*, page bbs069, 2012.

[33] Jennifer L Reed, Trina R Patel, Keri H Chen, Andrew R Joyce, Margaret K Applebee, Christopher D Herring, Olivia T Bui, Eric M Knight, Stephen S Fong, and Bernhard O Palsson. Systems approach to refining genome annotation. *Proceedings of the National Academy of Sciences*, 103(46):17480–17484, nov 2006.

[34] Vinay Satish Kumar, Madhukar S Dasika, and Costas D Maranas. Optimization based automated curation of metabolic reconstructions. *BMC Bioinformatics*, 8(1):212, 2007.

[35] Vinay Satish Kumar and Costas D Maranas. GrowMatch: An Automated Method for Reconciling In Silico/In Vivo Growth Predictions. *PLoS Computational Biology*, 5(3):e1000308, mar 2009.

[36] Ines Thiele, Nikos Vlassis, and Ronan M T Fleming. FASTGAPFILL: efficient gap filling in metabolic networks. *Bioinformatics*, 30(17):2529–2531, sep 2014.

[37] Kelly Botero, Daniel Osorio, Janneth Gonzalez, and Andres Pinzon. *g2f: Find and Fill Gaps in Metabolic Networks*, 2016.

[38] Minoru Kanehisa, Susumu Goto, Masahiro Hattori, Kiyoko F Aoki-Kinoshita, Masumi Itoh, Shuichi Kawashima, Toshiaki Katayama, Michihiro Araki, and Mika Hirakawa. From genomics to chemical genomics: new developments in KEGG. *Nucleic acids research*, 34(suppl 1):D354—-D357, 2006.

[39] Nathan D Price, Jennifer L Reed, and Bernhard Ø Palsson. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nature Reviews Microbiology*, 2(11):886–897, 2004.

[40] Liming Liu, Rasmus Agren, Sergio Bordel, and Jens Nielsen. Use of genome-scale metabolic models for understanding microbial physiology. *FEBS letters*, 584(12):2556–2564, 2010.

[41] Jan Schellenberger, Richard Que, Ronan M T Fleming, Ines Thiele, Jeffrey D Orth, Adam M Feist, Daniel C Zielinski, Aarash Bordbar, Nathan E Lewis, Sorena Rahmanian, and Others. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2. 0. *Nature protocols*, 6(9):1290–1307, 2011.

[42] Samuel M D Seaver, Christopher S Henry, and Andrew D Hanson. Frontiers in metabolic reconstruction and modeling of plant genomes. *Journal of experimental botany*, 63(6):2247–2258, 2012.

[43] Amlt Varma and Bemhard O Palsson. Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use. *Bio/technology*, 12, 1994.

[44] Bernhard Palsson. In silico biology through "omics". *Nature biotechnology*, 20(7):649–650, 2002.

[45] Nathan E Lewis, Harish Nagarajan, and Bernhard O Palsson. Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. *Nature Reviews Microbiology*, 10(4):291–305, 2012.

[46] Nadine Töpfer, Camila Caldana, Sergio Grimbs, Lothar Willmitzer, Alisdair R Fernie, and Zoran Nikoloski. Integration of genome-scale modeling and transcript profiling reveals metabolic pathways underlying light and temperature acclimation in Arabidopsis. *The Plant Cell*, 25(4):1197–1211, 2013.

[47] Matthew A Oberhardt, Bernhard Ø Palsson, and Jason A Papin. Applications of genome-scale metabolic reconstructions. *Molecular systems biology*, 5:320, 2009.

[48] Daniel Machado and Markus Herrgård. Systematic Evaluation of Methods for Integration of Transcriptomic Data into Constraint-Based Models of Metabolism. *PLoS Computational Biology*, 10(4):e1003580, 2014.

[49] Markus W Covert and Bernhard Ø Palsson. Transcriptional regulation in constraints-based metabolic models of Escherichia coli. *Journal of Biological Chemistry*, 277(31):28058–28064, 2002.

[50] Mats Åkesson, Jochen Förster, and Jens Nielsen. Integration of gene expression data into genome-scale metabolic models. *Metabolic engineering*, 6(4):285–293, 2004.

[51] Markus W Covert, Eric M Knight, Jennifer L Reed, Markus J Herrgard, and Bernhard O Palsson. Integrating high-throughput and computational data elucidates bacterial networks. *Nature*, 429(6987):92–96, 2004.

[52] Jennifer L Reed. Shrinking the metabolic solution space using experimental datasets. *PLoS Comput Biol*, 8(8):e1002662, 2012.

[53] Anna S Blazier and Jason A Papin. Integration of expression data in genome-scale metabolic network reconstructions. *Frontiers in physiology*, 3:299, 2012.

[54] Daniel Machado and Markus Herrgård. Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput Biol*, 10(4):e1003580, 2014.

[55] Caroline Colijn, Aaron Brandes, Jeremy Zucker, Desmond S Lun, Brian Weiner, Maha R Farhat, Tan-Yun Cheng, D Branch Moody, Megan Murray, and James E Galagan. Interpreting expression data with metabolic flux models: predicting Mycobacterium tuberculosis mycolic acid production. *PLoS Comput Biol*, 5(8):e1000489, 2009.

[56] Pablo Carbonell. Hepatotoxicity prediction by systems biology modeling of disturbed metabolic pathways using gene expression data. *ALTEX*, pages 1 – 23, 2016.

[57] Kazuhiro Takuma, Akemichi Baba, and Toshio Matsuda. Astrocyte apoptosis: Implications for neuroprotection. *Progress in Neurobiology*, 72:111–127, 2004.

[58] Sofie C. Lange, Lasse K. Bak, Helle S. Waagepetersen, Arne Schousboe, and Michael D. Norenberg. Primary cultures of astrocytes: Their value in understanding astrocytes in health and disease. *Neurochemical Research*, 37:2569–2588, 2012.

[59] Harold K Kimelberg. Functions of Mature Mammalian Astrocytes: A Current View. *The Neuroscientist*, 16(1):79–106, 2010.

[60] Nicola J Allen and Ben Barres. Neuroscience: Glia - more than just brain glue. *Nature*, 457(7230):675–677, 2009.

[61] Michael M Halassa and Philip G Haydon. Integrated brain circuits: astrocytic networks modulate neuronal activity and behavior. *Annual review of physiology*, 72(2):335–355, 2010.

[62] Christian Giaume, Annette Koulakoff, Lisa Roux, David Holcman, and Nathalie Rouach. Astroglial networks: a step further in neuroglial and gliovascular interactions. *Nature reviews. Neuroscience*, 11(2):87–99, 2010.

[63] Cora H. Nijboer, Cobi J. Heijnen, Vincent Degos, Hanneke L M Willemen, Pierre Gressens, and Annemieke Kavelaars. Astrocyte GRK2 as a novel regulator of glutamate transport and brain damage. *Neurobiology of Disease*, 54:206–215, 2013.

[64] Ben Barres. The mystery and magic of glia: a perspective on their roles in health and disease. *Neuron*, 60(3):430–40, 2008.

[65] Jun Shen. Modeling the glutamate-glutamine neurotransmitter cycle. *Frontiers in Neuroenergetics*, 5(1):1–13, 2013.

[66] Francesco Petrelli and Paola Bezzi. Novel insights into gliotransmitters. *Current Opinion in Pharmacology*, 26(Table 1):138–145, 2016.

[67] Andrea R Durrant and Uriel Heresco-Levy. D-Serine in Neuropsychiatric Disorders: New Advances. *Advances in Psychiatry*, 2014:1–16, 2014.

[68] Tunahan Cakir, Selma Alsan, Hale Saybaşili, Ata Akin, and Kutlu O Ulgen. Reconstruction and flux analysis of coupling between metabolic pathways of astrocytes and neurons: application to cerebral hypoxia. *Theoretical biology & medical modelling*, 4(1):48, 2007.

[69] Stephen P. Raps, James C K Lai, Leif Hertz, and Arthur J L Cooper. Glutathione is present in high concentrations in cultured astrocytes but not in cultured neurons. *Brain Research*, 493(2):398–401, 1989.

[70] Mithilesh Kumar Jha, Dong Ho Park, Hyun Kook, In-Kyu Lee, Won-Ha Lee, and Kyoungho Suk. Metabolic Control of Glia-Mediated Neuroinflammation. *Current Alzheimer research*, 13(4):387–402, 2016.

[71] Igor Allaman, Mireille Bélanger, and Pierre J. Magistretti. Astrocyte–neuron metabolic relationships: for better and for worse. *Trends in Neurosciences*, 34(2):76–87, 2011.

[72] Xue Feng Wang and Max S. Cynader. Astrocytes provide cysteine to neurons by releasing glutathione. *Journal of Neurochemistry*, 74(4):1434–1442, 2000.

[73] Nicholas J Maragakis and Jeffrey D Rothstein. Mechanisms of Disease: astrocytes in neurodegenerative disease. *Nature clinical practice. Neurology*, 2(12):679–689, 2006.

[74] Marco Avila-Rodriguez, Luis Miguel Garcia-Segura, Ricardo Cabezas, Daniel Torrente, Francisco Capani, Janneth Gonzalez, and George E. Barreto. Tibolone protects T98G cells from glucose deprivation. *The Journal of Steroid Biochemistry and Molecular Biology*, 144(8):294–303, 2014.

[75] Ghulam Hussain, Florent Schmitt, Jean-Philippe Loeffler, and Jose-Luis Gonzalez de Aguilar. Fatting the brain: a brief of recent research. *Frontiers in cellular neuroscience*, 7(9):144, 2013.

[76] Brent E Masel and Douglas S DeWitt. Traumatic brain injury: a disease process, not an event. *Journal of neurotrauma*, 27(8):1529–1540, 2010.

[77] Qing Yan. *Systems Biology in Drug Discovery and Development*, volume 662 of *Methods in Molecular Biology*. Humana Press, Totowa, NJ, 2010.

[78] Michael T. Fitch and Jerry Silver. CNS injury, glial scars, and inflammation: Inhibitory extracellular matrices and regeneration failure. *Experimental Neurology*, 209(2):294–301, 2008.

[79] James a. Dowell, Jeffrey a. Johnson, and Lingjun Li. Identification of astrocyte secreted proteins with a combination of shotgun proteomics and bioinformatics. *Journal of Proteome Research*, 8(8):4135–4143, 2009.

[80] Anna V. Molofsk, Robert Krenick, Erik Ullian, Hui Hsin Tsai, Benjamin Deneen, William D. Richardson, Ben. Barres, and David H. Rowitch. Astrocytes and disease: A neurodevelopmental perspective. *Genes and Development*, 26:891–907, 2012.

[81] Marta Sidoryk-Wegrzynowicz and Michael Aschner. Role of astrocytes in manganese mediated neurotoxicity. *BMC Pharmacology and Toxicology*, 14(1):23, dec 2013.

[82] Sunita Gupta, Alecia G. Knight, Shruti Gupta, Jeffrey N. Keller, and Annadora J. Bruce-Keller. Saturated long-chain fatty acids activate inflammatory signaling in astrocytes. *Journal of Neurochemistry*, 120(6):1060–1071, 2012.

[83] Sudarshana Purkayastha and Dongsheng Cai. Neuroinflammatory basis of metabolic syndrome. *Molecular Metabolism*, 2(4):356–363, 2015.

[84] Kyoungho Suk. Proteomics-based discovery of biomarkers and therapeutic targets in neurodegenerative diseases: perspective of microglia and neuroinflammation. *Expert Opinion on Therapeutic Patents*, 16(3):237–247, 2006.

[85] Yoram Vodovotz. Translational systems biology of inflammation and healing. *Wound Repair and Regeneration*, 18(1):3–7, 2010.

[86] Yoram Vodovotz, Yoram Vodovotz, Marie Csete, Marie Csete, John Bartels, John Bartels, Steven Chang, Steven Chang, Gary An, and Gary An. Translational Systems Biology of Inflammation. *PLoS Comput Biol*, 4(4), 2008.

[87] Gary An, John Bartels, and Yoram Vodovotz. In silico augmentation of the drug development pipeline: Examples from the study of acute inflammation. *Drug Development Research*, 72(2):187–200, 2011.

[88] Qi Mi, Nicole Yee-Key Li, Cordelia Ziraldo, Ali Ghuma, Maxim Mikheev, Robert Squires, David O Okonkwo, Katherine Verdolini-Abbott, Gregory Constantine, Gary An, and Yoram Vodovotz. Translational systems biology of inflammation: potential applications to personalized medicine. *Personalized Medicine*, 7(5):549–559, 2010.

[89] Durgaprasad Laveti, Manoj Kumar, R Hemalatha, Ramakrishna Sistla, V G M Naidu, Venu Talla, Vinod Verma, Navrinder Kaur, and Ravinder Nagpal. Anti-inflammatory treatments for chronic diseases: a review. *Inflammation & allergy drug targets*, 12(5):349–61, 2013.

[90] Paola Albertazzi, Raffaele Di Micco, and Ettore Zanardi. Tibolone: a review. *Maturitas*, 30(3):295–305, nov 1998.

[91] Katarzyna Wojtal, Michał K. Trojnar, and Stanisław J. Czuczwar. Endogenous neuroprotective factors: neurosteroids. *Pharmacological reports : PR*, 58(3):335–40, 2006.

[92] H.J Kloosterboer. Tibolone: a steroid with a tissue-specific mode of action. *The Journal of Steroid Biochemistry and Molecular Biology*, 76(1-5):231–238, mar 2001.

[93] Michael J Reed and Helenius J Kloosterboer. Tibolone: a selective tissue estrogenic activity regulator (STEAR). *Maturitas*, 48 Suppl 1(4):S4–6, aug 2004.

[94] Cees J. Timmer, H. A M Verheul, and D. P. Doorstam. Pharmacokinetics of tibolone in early and late postmenopausal women. *British Journal of Clinical Pharmacology*, 54(2):101–106, 2002.

[95] M A Altinoz, S B Albayrak, A Karasu, P A Sabanci, M Imer, and A Bilir. The effects of tibolone on the human primary glioblastoma multiforme cell culture and the rat C6 glioma model. *Neurol Res*, 31(9):923–927, 2009.

[96] Helenius J. Kloosterboer. Tissue-selectivity: the mechanism of action of tibolone. *Maturitas*, 48(SUPPL. 1):30–40, 2004.

[97] Ron Edgar, Michael Domrachev, and Alex E Lash. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*, 30(1):207–210, 2002.

[98] Ye Zhang, Steven A. Sloan, Laura E. Clarke, Christine Caneda, Colton A. Plaza, Paul D. Blumenthal, Hannes Vogel, Gary K. Steinberg, Michael S B Edwards, Gordon Li, John A. Duncan, Samuel H. Cheshier, Lawrence M. Shuer, Edward F. Chang, Gerald A. Grant, Melanie G Hayden Gephart, and Ben A. Barres. Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron*, 89(1):37–53, 2016.

[99] Michael Rebhan, Vered Chalifa-Caspi, Jaime Prilusky, and Doron Lancet. GeneCards: integrating information about genes, proteins and diseases. *Trends in Genetics*, 13(4):163, 1997.

[100] Donna Maglott, Jim Ostell, Kim D Pruitt, and Tatiana Tatusova. Entrez Gene: gene-centered information at NCBI. *Nucleic acids research*, 33(suppl 1):D54–D58, 2005.

[101] Marc Carlson. *UniProt.ws: R Interface to UniProt Web Services*, 2016.

[102] Ines Thiele, Neil Swainston, Ronan M T Fleming, Andreas Hoppe, Swagatika Sahoo, Maike K Aurich, Hulda Haraldsdottir, Monica L Mo, Ottar Rolfsson, Miranda D Stobbe, and Others. A community-driven global reconstruction of human metabolism. *Nature biotechnology*, 31(5):419–425, 2013.

[103] Daniel Osorio, Janneth Gonzalez, and Andres Pinzon. minval: MINimal VALidation for Stoichiometric Reactions. Technical report, 2016.

[104] Daniel Osorio, Kelly Botero, Janneth Gonzalez, and Andres Pinzon. *exp2flux: Convert Gene EXPression Data to FBA FLUXes*, 2016.

[105] Gabriel Gelius-dietrich, Abdelmoneim Amer Desouki, Claus Jonathan Fritzemeier, and Martin J Lercher. sybil – Efficient constraint-based modelling in R. 2013.

[106] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.

[107] Jeffrey D Orth, Ines Thiele, and B O Palsson. What is flux balance analysis? *Nature Biotechnology*, 28(3):245–248, 2010.

[108] Karthik Raman and Nagasuma Chandra. Flux balance analysis of biological systems: Applications and challenges. *Briefings in Bioinformatics*, 10(4):435–449, 2009.

[109] M T Fitch and J Silver. Activated macrophages and the blood-brain barrier: inflammation after CNS injury leads to increases in putative inhibitory molecules. *Experimental Neurology*, 148(2):587–603, 1997.

[110] R. Ciccarelli, P. Ballerini, G. Sabatino, M. P. Rathbone, M. D'Onofrio, F. Caciagli, and P. Di Iorio. Involvement of astrocytes in purine-mediated reparative processes in the brain. *International Journal of Developmental Neuroscience*, 19(4):395–414, 2001.

[111] Mustafa Sertbaş, Kutlu Ülgen, and Tunahan Çakir. Systematic analysis of transcription-level effects of neurodegenerative diseases on human brain metabolism by a newly reconstructed brain-specific metabolic network. *FEBS Open Bio*, 4:542–553, 2014.

[112] João V. Sá, Susanne Kleiderman, Catarina Brito, Ursula Sonnewald, Marcel Leist, Ana P. Teixeira, and Paula M. Alves. Quantification of Metabolic Rearrangements During Neural Stem Cells Differentiation into Astrocytes by Metabolic Flux Analysis. *Neurochemical Research*, apr 2016.

[113] Arabinda Das, Naren L. Banik, and Swapan K. Ray. Flavonoids activated caspases for apoptosis in human glioblastoma T98G and U87MG cells but not in human normal astrocytes. *Cancer*, 116(1):164–176, 2010.

[114] Rupa Bhowmick, Abhishek Subramanian, and Ram Rup Sarkar. Exploring the differences in metabolic behavior of astrocyte and glioblastoma: a flux balance analysis approach. *Systems and Synthetic Biology*, 9(4):159–177, dec 2015.

[115] Christelle Le Foll and Barry E Levin. Fatty acid-induced astrocyte ketone production and the control of food intake. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 310(11):R1186–R1192, jun 2016.

[116] Li Liu, Rebecca Martin, and Christina Chan. Palmitate-activated astrocytes via serine palmitoyltransferase increase BACE1 in primary neurons by sphingomyelinases. *Neurobiology of Aging*, 34(2):540–550, 2013.

[117] Rafael Rangel-Aldao. The unfolded protein response, inflammation, oscillators, and disease: a systems biology approach. *Endoplasmic Reticulum Stress in Diseases*, 2(1):30–52, 2015.

[118] Rachel Williams, Cassandra L. Buchheit, Nancy E J Berman, and Steven M. Levine. Pathogenic implications of iron accumulation in multiple sclerosis. *Journal of Neurochemistry*, 120(1):7–25, 2012.

[119] V Parpura and P G Haydon. Physiological astrocytic calcium levels stimulate glutamate release to modulate adjacent neurons. *Proc.Natl.Acad.Sci.U.S.A*, 97(0027-8424):8629–8634, 2000.

[120] Leif Hertz, Ralf Dringen, Arne Schousboe, and Stephen R. Robinson. Astrocytes: Glutamate producers for neurons. *Journal of Neuroscience Research*, 57(4):417–428, 1999.

[121] Yu-Cun Niu, Ren-Nan Feng, Yan Hou, Kang Li, Zhen Kang, Jian Wang, Chang-Hao Sun, and Ying Li. Histidine and arginine are associated with inflammation and oxidative stress in obese women. *British Journal of Nutrition*, 108(01):57–61, 2012.

[122] Yuan Ti Lee, Cheng Chin Hsu, Meng Hsiao Lin, Keh Sen Liu, and Mei Chin Yin. Histidine and carnosine delay diabetic deterioration in mice and protect human low density lipoprotein against oxidation and glycation. *European Journal of Pharmacology*, 513(1-2):145–150, 2005.

[123] Dong Ok Son, Hideo Satsu, and Makoto Shimizu. Histidine inhibits oxidative stress- and TNF-α-induced interleukin-8 secretion in intestinal epithelial cells. *FEBS Letters*, 579(21):4671–4677, aug 2005.

[124] Morris H Baslow. N-acetylaspartate in the vertebrate brain: metabolism and function. *Neurochemical research*, 28(6):941–53, jun 2003.

[125] Leonard T. Rael, Gregory W. Thomas, Raphael Bar-Or, Michael L. Craun, and David Bar-Or. An anti-inflammatory role for N-acetyl aspartate in stimulated human astroglial cells. *Biochemical and Biophysical Research Communications*, 319(3):847–853, 2004.

[126] F a Chaudhry, R J Reimer, D Krizaj, D Barber, J Storm-Mathisen, D R Copenhagen, and R H Edwards. Molecular analysis of system N suggests novel physiological roles in nitrogen metabolism and synaptic transmission. *Cell*, 99(7):769–80, 1999.

[127] Michael Doengi, Daniela Hirnet, Philippe Coulon, Hans-Christian Pape, Joachim W Deitmer, and Christian Lohr. GABA uptake-dependent Ca(2+) signaling in developing olfactory bulb astrocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 106:17570–17575, 2009.

[128] D. R. Green, L. Galluzzi, and G. Kroemer. Metabolic control of cell death. *Science*, 345(6203):1250256–1250256, 2014.

[129] Sean Murphy, Brian Pearce, James Jeremy, and Paresh Dandona. Astrocytes as eicosanoid-producing cells. *Glia*, 1(4):241–245, 1988.

[130] Harshica Fernando, John E. Wiktorowicz, Kizhake V. Soman, Bhupendra S. Kaphalia, M. Firoze Khan, and G. A. Shakeel Ansari. Liver proteomics in progressive alcoholic steatosis. *Toxicology and Applied Pharmacology*, 266(3):470–480, 2013.

[131] Leif Väremo, Camilla Scheele, Christa Broholm, Adil Mardinoglu, Caroline Kampf, Anna Asplund, Intawat Nookaew, Mathias Uhlén, BenteKlarlund Pedersen, and Jens Nielsen. Proteome- and Transcriptome-Driven Reconstruction of the Human Myocyte Metabolic Network and Its Use for Identification of Markers for Diabetes. *Cell Reports*, 11(6):921–933, 2015.

[132] Sapar M. Saparov, Kun Liu, Peter Agre, and Peter Pohl. Fast and selective ammonia transport by aquaporin-8. *Journal of Biological Chemistry*, 282(8):5296–5301, 2007.

[133] Anje a Te Velde, Inge Pronk, Floor de Kort, and Pieter C F Stokkers. Glutathione peroxidase 2 and aquaporin 8 as new markers for colonic inflammation in experimental colitis and inflammatory bowel diseases: an important role for H2O2? *European journal of gastroenterology & hepatology*, 20:555–560, 2008.

[134] Clemente Landriscina, Gabriele V Gnoni, and Ernesto Quaqliariello. Fatty -Acid Bios y nt hesis Present in Microsomes and Mitochondria of Rat Liver. 196:188–196, 1972.

[135] Kenji Tamura, Asami Makino, Françoise Hullin-Matsuda, Toshihide Kobayashi, Mutsuo Furihata, Suyoun Chung, Shingo Ashida, Tsuneharu Miki, Tomoaki Fujioka, Taro Shuin, Yusuke Nakamura, and Hidewaki Nakagawa. Novel lipogenic enzyme ELOVL7 is involved in prostate cancer growth through saturated long-chain fatty acid metabolism. *Cancer Research*, 69(20):8133–8140, 2009.

[136] Barbara Ahlemeyer, Stefan Kölker, Yuan Zhu, Georg F. Hoffmann, and Josef Krieglstein. Increase in glutamate-induced neurotoxicity by activated astrocytes involves stimulation of protein kinase C. *Journal of Neurochemistry*, 82(3):504–515, jul 2002.

[137] Georgia B. Schuller-Levis and Eunkyue Park. Taurine: new implications for an old amino acid. *FEMS Microbiology Letters*, 226(2):195–202, sep 2003.

[138] Nūn Amen-Ra. Humans are evolutionarily adapted to caloric restriction resulting from ecologically dictated dietary deprivation imposed during the Plio–Pleistocene period. *Medical Hypotheses*, 66(5):978–984, jan 2006.

[139] Adam M. Feist and Bernhard O. Palsson. The biomass objective function. *Current Opinion in Microbiology*, 13(3):344–349, 2010.

[140] Graham A. Colditz, Kathleen M. Egan, and Meir J. Stampfer. Hormone replacement therapy and risk of breast cancer: Results from epidemiologic studies. *American Journal of Obstetrics and Gynecology*, 168(5):1473–1480, 1993.

[141] Graham A. Colditz, Susan E. Hankinson, David J. Hunter, Walter C. Willett, JoAnn E. Manson, Meir J. Stampfer, Charles Hennekens, Bernard Rosner, and Frank E. Speizer. The Use of Estrogens and Progestins and the Risk of Breast Cancer in Postmenopausal Women. *New England Journal of Medicine*, 332(24):1589–1593, 1995.

[142] Dominique Pessayre, Alain Berson, Bernard Fromenty, and Abdellah Mansouri. Mitochondria in Steatohepatitis. *Seminars in Liver Disease*, 21(01):057–070, 2001.

[143] F. Sun, M. L. Xie, L. J. Zhu, J. Xue, and Z. L. Gu. Inhibitory effect of osthole on alcohol-induced fatty liver in mice. *Digestive and Liver Disease*, 41(2):127–133, 2009.

[144] Marco Avila-Rodriguez, Luis Miguel Garcia-Segura, Oscar Hidalgo-lanussa, Eliana Baez, Janneth Gonzalez, and George E. Barreto. Tibolone protects astrocytic cells from glucose deprivation through a mechanism involving estrogen receptor beta and the upregulation of neuroglobin expression. *Molecular and Cellular Endocrinology*, 433:35–46, 2016.