Computational modelling of the glycinergic synapse

Elzbieta Al-Saidi

A thesis submitted in partial fulfilment of the requirements of
Oxford Brookes University
for the award of the degree Doctor of Philosophy

The work has been fulfilled according to a Studentship Agreement (19/09/2016)
between Oxford Brookes University and the Institute of Cytochemistry and
Molecular Pharmacology, Moscow, Russia.
September 2020

Acknowledgements

The last four years spent pursuing my PhD in the Cell Systems Modelling Group will always be in my memory. This time has been challenging and fun, and has not only expanded my scientific knowledge, but also strenghtened my character. Without the support of people who have offered their unlimited help, for whom I am deeply grateful, it would not have been possible.

My biggest thanks goes to my supervisors, Dr Mark Poolman, Professor David Fell, and Professor Isabel Bermudez-Diaz who have been incredibly helpful and have taught me a great lot. Their encouragement, belief, and the time they gave me to pass on their knowledge and experience will always be appreciated. Apart from their educational input, I am deeply grateful for their enthusiasm, non-scientific conversations, and guidance. I would also like to express my thanks to Dr. Yaroslav R. Nartsissov who initiated this project and gave a principal scientific aim to the work.

I would like to thank the colleagues within the Cell Systems Modelling Group, who have made this time enjoyable and did not hesitate to answer my questions and offer help and discussions, even in their own busy periods. Thank you, Dr Noah Mesfin, Dr Kailash Adhikari, Dr Dipali Singh, Dr Hassan Hartman, and Yanica Said, as well as the passing members, Dr Maite Diaz-Calvo, Dr Nicole Pearcy, and Dr Rupert Norman. To the Cell Systems Modelling Group as a whole, thank you for the fun travels, and teaching me how to appreciate a good curry- you certainly made this journey less bland.

A special thanks also goes to the other members of the Oxford Brookes University who have been supportive. I would like to thank Professor David Carter, without whom I may not have began this journey, as well as Professor Caroline Griffiths who has supervised me in my undergraduate years and shared her enthusiasm for science with me. I would also like to thank Professor David Meredith for helpful discussions regarding the experimental work, and Teresa Minguez-Vinas for helping me in the laboratory. Finally, a big thank you goes to the staff at the Faculty of Health and Life Science, in particular Catherine Joyejob and Jill Organ for their administrative help.

A huge thank you goes to the people in my life who have made me smile and helped me throughout these years- to all the colleagues from various departments, for the time spent enjoying lunches in the sun. I would also like to thank my friends from outside of the University, with a special thank you to Hernan. Thank you for every moment you supported me, our conversations, and the time we spent in misty Oxford pubs- without you, my time in this city would not have been the same.

My deepest gratitude goes to family- my beloved grandparents Janina and Stanislaw Stasiak for always offering warm words of advice and support, and my brothers Jakub and Basym.

Finally, I would like to deeply thank my mum, Beata Stasiak. Thank you for your positivity, kindness, and strength. You have always encouraged me, and passed your curious nature and a desire for learning on to me which resulted in this thesis, for which I will always be grateful.

Abstract

Glycine is the simplest and smallest amino acid present in human proteins. It has many important roles in human as an intermediate of biosynthetic one carbon metabolism pathways, component of structural proteins, and in the brain acting as the second major inhibitory neurotransmitter via the activation of the glycine receptor (GlyR). Glycinergic neurons are present in various brain regions including the forebrain and the retina, although highest densities can be found in the lower brain regions of the central nervous system (CNS) such as the brain stem, where glycine is important in the regulation of muscle contraction, sensory inputs, and others.

Given its diverse role, dysregulation in glycine biosynthesis and neurotranmission is implicated in numerous conditions detrimental to human health such as some subtypes of glioma, as well as genetic diseases caused by mutations to the components of the glycinergic synapse, for instance non-ketotic hyperglycinemia (NHK) which leads to severe neurological problems. Moreover, the control of glycine concentration time course via the inhibition of glycine reuptake has been suggested as a potential novel treatment for several conditions, including schizophrenia, alcoholism, and others.

In order to advance the understanding of glycine regulation in the human CNS, computational modelling of the various aspects of the glycinergic synapse were deployed. A metabolic model of the CNS was constructed with the use of publicly available databases as a derivation of a generic model of human metabolism, Recon2.2. The experimentally determined value of the human oxygen to glucose index (OGI), indicating glucose brain consumption was used to validate the model via the estimation of the background ATP demand which was similar to that reported in the scientific literature.

In order to further contextualise the model to represent CNS metabolism, a method for the integration of single cell RNA-sequencing dataset containing the expression values of different types of CNS cells was developed.

Flux Balance Analysis (FBA) methods were applied to the model in order to establish the most likely pathways of glycine and ATP production across different CNS cell types, suggesting the compartmentalization of the serine/glycine metabolism in CNS cells. Furthermore, FBA was used to assess whether serine or glycine can be used as main substrates in increasing cellular ATP and biosynthetic demands, with the results suggesting that serine can be used for ATP formation by the mitochondria via its breakdown to glycine, producing NADH for use in the Electron Transport Chain (ETC).

Interactions between proteins composing the glycinergic synapse, including the glycine transporter proteins 1 and 2 (GlyT1, GlyT2) which are responsible for the termination of signal via glycine reuptake determine the time course of glycine released upon stimulation, and therefore the inhibitory post-synaptic currents (IPSC) Cl^- currents via GlyR. For the purpose of increasing the understanding of the events shaping the glycine time course, a kinetic model of the synapse was developed integrating previous experimental observations. The model was analysed by looking at the effects of changing its parameters on the resulting amplitudes and decay time constants of glycine and the resulting IPSC.

In addition, electrophysiological two electrode voltage clamp (TEVC) techniques were used in order to characterize the effects of a newly identified NKH-causing mutation at position Ser 407 on GlyT1 expressed in *Xenopus laevis* oocytes. The mutation lead to the inhibition of transport via the GlyT1 as compared to wild type transporter, and the result was integrated into the kinetic synapse model suggesting that the loss of GlyT1 function leads to a higher maximum concentration of glycine which persists in the synapse for longer, leading to an increased time course of IPSC at the postsynaptic neuron.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Glycine is the simplest amino acid with numerous function in mammals, where it is a component of structural proteins, an intermediate metabolite of cellular biosynthetic pathways as well as in the brain functioning as an inhibitory neurotransmitter. Although glycine or its precursor, serine, can be supplied from the diet, due to their neuroactive properties uptake of these amino acids is restricted by the blood brain barrier (BBB) and their *de novo* brain biosynthesis occurs inside the organ.

Glycinergic inhibitory synapses are located mostly to the lower brain regions such as the brain stem and the spinal cord, where glycine regulates the movement of skeletal muscles, breathing rhythms, and others. Disruption of glycinergic neurotransmission resulting from mutations in proteins controlling its concentration lead to rare conditions with severe phenotypes, for instance the inborn metabolic condition, nonketotic hyperglycinemia (NKH) which results in death in over 50% of affected individuals in the first neonatal week [3]. Furthermore, modulation of glycine time course across the glycinergic synapse in different brain regions via the inhibition of its reuptake has been proposed as a treatment for numerous conditions, including schizophrenia, alcoholism, and others [4].

With respect to cellular metabolism overall, conversion of serine to glycine is the main source of one carbon (1C) units used in the biosynthetic pathways of one carbon metabolism and thus pathways of serine/glycine metabolism are implicated in a subset of cancers. In the brain, increased glycine concentrations and correlated with the presence and stage of human gliomas [5, 6], and upregulation of genes encoding enzymes related to serine/glycine has been found in a subset of cancers from different tissues [7–10].

Much of what is known about glycine biosynthetic pathways specific to the brain is based on radiolabelled carbon incorporation studies and immunohistochemical staining of enzymes performed between 1970s - 19s90s, with scant research performed on the topic in recent years. Given the importance of glycine metabolism to human health, increasing the understanding of its initial biosynthesis within the brain, as well as the interaction between proteins composing the glycinergic synapse which regulate its concentration time course can aid in the design of novel chemotherapeutic agents and neuromodulator compounds.

1.2 Glycine and the CNS

1.2.1 Brain composition and neuronal communication

The brain is composed of various cell types with by far the most important being the neuron, first discovered by Cajal in 1886 and responsible for conducting information across the central nervous system. Neurons are specialized cells, with extensions from the cell body (or the soma) called the axons, which relay information away from the cell body and dendrites, which integrate incoming signals towards the soma. The axons vary in length, with the longest being the sciatic nerve which extends from the base of the spinal cord to the big toe of the foot and measures at around one meter. Information is carried across the axon as an electrical impulse known as an action potential, thus in order to prevent the dissipation of ionic flow and ensure speed of transmission, many axons are coated in myelin, a fatty substance which insulates the cord along its length. Communication in the form of electrical signals is enabled due to a difference in charge contained within the fluid external to the cell and that within the cell cytosol. The resting voltage of a neuron is around -70mV although varies according to the cell and is set up by the active transport of ions across the compartments, as well as the presence of negatively charged proteins within the cell.

The synapse comes from the Greek word for ‘conjunction’ and is the site of communication between two neurons. It is composed of the axon belonging to the presynaptic cell, a tiny space called the synaptic cleft, and the dendrite of the opposing postsynaptic neuron. Transmission of signal across the synapse is relayed via electrical or chemical means, with each having a specific advantage. The electrical synapse directly transmits ionic fluxes from one cell to the other via the gap junction made up of specialized proteins and is a faster method of talk, allowing cells to fire in unison.

The chemical synapse on the other hand is slower, and involves the release of a neurotransmitter from the flattened structure of the axonal bouton, which binds to a neurotransmitter receptor protein on the opposing dendrite. Neurotransmitters available in the cytosol of the presynaptic cell are packaged into synaptic vesicles which, triggered by an increase in Ca^{2+} concentrations in the soma, fuse with the membrane and release the messenger across the synaptic cleft. There are two main types of receptor proteins, the ligand gated ion channels (LGICs) which directly carry an ionic flux, and G-coupled receptor proteins which trigger the opening of a neighbouring ion channel protein. Upon neurotransmitter binding, the LGIC protein changes its conformation and allows the flow of ions into the postsynaptic neuron. Amino acids composing the pore of the protein determine the electrochemical environment which favour the flow of differently charged ions. Excitatory chemical synapses which form the majority within the CNS contain neurotransmitter receptors which favours the flow of positively charged ions, leading to an increase in the cell voltage and thus the probability of cell depolarization. Inhibitory chemical synapses contain receptors which carry negatively charged ions, leading to a decrease in the probability of neuronal firing. A single neuron contains dendrites which forms connections with numerous other neurons, both excitatory and inhibitory thus response is generated within the cell if there is a large net disbalance between inhibition and excitation within a region of the soma.

Although it was originally thought that one neuron synthesizes and releases only one type of neurotransmitter, it is now known that many neurons co-release different neurotransmitters which may be concentrated in separate, or the same synaptic vesicles. The list of neurotransmitters accepted so far is still growing, with more than 20 classified neurotransmitters and more potential candidates under review. Major categories include amino acid neurotransmitters or their derivatives (eg. glutamate and glycine), peptides (eg. endorphins) and monoamines (eg. dopamine, serotonin).

Neurotransmitter receptors bind their respective agonist, but may also contain binding sites for partial agonists such as other neurotransmitters which lead to a lower receptor response relative to its response to the main agonist. They may also contain binding sites for other molecules which modulate receptor function which may be native or non-native such as alcohols, anaesthetics, and psychoactive substances like cannabis.

Apart from neurons, the brain is composed of different supporting cell types collectively termed the glial cells which outnumber neurons, and constitute between 30-60% of the total mammalian brain mass [11, 12]. Glial cells are further subdivided into four major classes, microglial cells, astrocytes, oligodendrocytes and their precursor cells, NG2-glia [13]. These cells have many important functions which are yet not fully understood. Astrocytes are by far the most studied glial cell type and are important in providing metabolic support to neurons, as well as in the maintenance of ionic gradients. Oligodendendrocytes secrete the myelin which ensheaths the neuronal axons of some neurons but their function is thought to extend beyond that, also to include providing metabolic support to neurons and others [14].

Brain tissue is divided into the grey and white matter, with the former composed of neuronal cell bodies, dendrites, as well as supporting glial cells. White matter refers to the mass containing myelinated axons as well as glial cells, including oligodendrocytes and others.

1.2.2 Brain anatomy overview

The human brain weights around 3 pounds in the adult and its weight to body weight ratio is one of the highest in mammals. Lower brain regions including the cerebellum and the brain stem control the more primal functions shared across the animal kingdom whereas the outer layer, the cerebral cortex, is the origin of more sophisticated activities, such as the understanding of consequence. The cerebral cortex composes around 80% of the weight of the brain and is thought to have expanded more recently in the course of human brain evolution in comparison to the lower brain region. The highly developed cerebral cortex, and high brain to body mass ratio are an indicator of higher intelligence in primates overall.

Communication between the brain and the rest of the body is mediated by neuronal connections and has been classified into two distinct system, the peripheral and the central nervous systems (PNS and CNS). The former includes the nerves outside of the brain as well as the spinal cord, with the main function of relaying information between organs and limbs to the central nervous system which is composed of the spinal cord and the brain.

The brain is enclosed in bone termed the cranium, which together with the bones protecting the face form the skull. Between the cranium and the brain itself is the meninges, which is a three layered tissue mass functioning as protective tissue as well as housing blood vessels and cavities which allows the flow of cerebrospinal fluid (CSF). The latter is a clear fluid that is secreted by the choroid plexus within the brain's ventricular system, composed of four hollow spaces located to the center of the brain and associated tubes which allow the flow of CSF around the brain. Its main functions are to provide mechanical protection, increase the brains buoyancy thus protecting the lower brain regions from the the effects of its mass, as well as being important in the clearance of metabolic waste products and maintenance of cell homeostasis.

Anatomy of the brain can be generalised to three major sections, the cerebral cortex (or cerebrum), composing the outer layer of the brain, the brain stem which resides in the deep mid-section of the brain, and the cerebellum which lies in the back of the organ.

The cerebrum is composed of the right and left hemispheres joined by a tract of nerves named the corpus callosum which communicates signals between the two hemispheres. Different rates of expansion of the grey and white matter during embryonic development leads to an increase in the surface of the organ in a compact form, leading to its wrinkled appearance. The tissue depth can be divided into six cortical layers which contain different types and structural architectures of neurons and supporting cells, and by region which distinguishes four major divisions, the frontal, paretial, occipital and temporal lobes (Figure 1.1). Much of what is known about the functions attributed to each of these areas come from observation of animal and human with injuries to a specific area, as well as functional MRI (fMRI) studies which observe changes in regional blood flow in response to stimuli or given task.

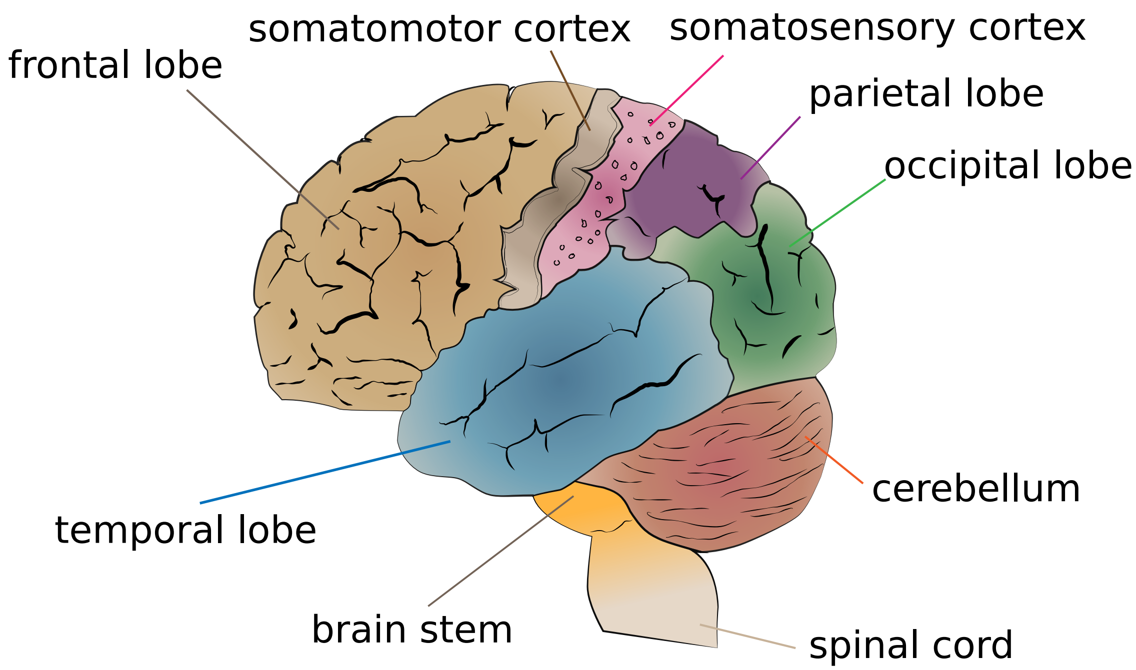


Figure 1.1: **Main regions of the human brain** The brain can be divided into three major areas, the cerebral cortex which forms the outer layer which is further subdivided into four lobes, the brain stem, and the cerebellum. Although some brain regions such as the somatomotor cortex are known to have very specific functions, much of the higher order phenomena such as the understanding of sound is a result of the neuronal networks connecting different brain regions which work in tandem. Adapted from Wikimedia Commons: https://commons.wikimedia.org

The frontal lobe is located at the front of the brain and is the largest of all lobes. It is implicated in a variety of human traits and abilities, including personality, decision making, emotional responses, speech and language - for instance, patients with damage to the frontal lobe may perform normally when testing for intellectual ability but have problems with emotional regulation.

Further towards the back of the organ lies the somatomotor cortex, which receives signals from the lower brain regions and is responsible for conscious movement. It contains populations of neurons which directly map onto different parts of the body- this specific mapping feature is also seen in the neighbouring somatosensory complex which is involved in the processing of sensory signals from the peripheral nervous system. The somatosensory cortex lies next to the paretial lobe, which is involved in higher order functions such as sensorimotor planning, learning, spatial recognition, and the ability to differentiate objects shapes and relationships in space.

The occipital lobe is the smallest out of all, and lies in the most posterior region of the brain. It is the visual center of the brain, receiving and processing information from the retina and deals with aspects of colour and object recognition, depth perception, facial recognition and others.

Lastly, the temporal lobe is located in the lower area on each side of the brain and is generally implicated in processing of auditory and visual signals, important in language recognition, production of speech and its perception, memory formation and others.

Deep in the core of the brain, above the brain stem lie the thalamus and the hypothalamus, which are important for movement, perception, and vital bodily functions. The thalamus receives information from the lower brain regions that make up the senses of touch, taste, hearing and sight, then modulates and relays the information into the higher regions of the cerebral cortex. The hypothalamus sits below the thalamus and is important in maintaining body homeostasis, such as temperature control, sleep cycles, appetite and emotions. It controls the interactions between the body's nervous system and the endocrine system, receiving information from different areas of the body and in turn stimulating organs and glands to release hormones.

The brain stem is composed of three sections, the pons, the midbrain and the medulla oblongata. It is continuous with the spinal cord and is crucial in the regulation of autonomic body functions, for example heart rate, breathing rhythms, as well the diameter of blood vessels. It is also important in the integration of signals which control voluntary skeletal muscle movements.

Another major area controlling movement is the cerebellum, which is located to the back of the brain and relays information from the motor area of the cerebral cortex to the spinal cord, whilst also receiving impulses from the periphery. In result, it is crucial in the coordination of voluntary muscle actions which work in maintaining posture, balance, as well as in the production of speech.

Extending from the brain stem is the spinal cord, mostly responsible for motor control of extremities as well as for relaying sensory signals from the rest of the body to the brain. It is a cylindrical bundle of neurons and supporting cells running within the vertebral canal, surrounded by spinal meninges filled with CSF which functions to protect the internal tissue. The interior of the spinal cord is composed of the grey matter, surrounded by white matter (Figure 1.2). Cross-sectional anatomy of the grey matter divides the spinal cord into dorsal (posterior) and ventral (anterior) horns, with a central canal in the middle that carries CSF from the ventricular system and transports nutrients to the surrounding tissue. Neurons within the dorsal horn receive sensory inputs such as pain from the afferent dorsal axon nerve cells of the peripheral nervous system and relays that information to the central nervous system. The ventral horn contains neurons which extend their neuronal axons via the ventral root and further synapse to their respective body segment, thus carrying information from the central to the peripheral nervous system.

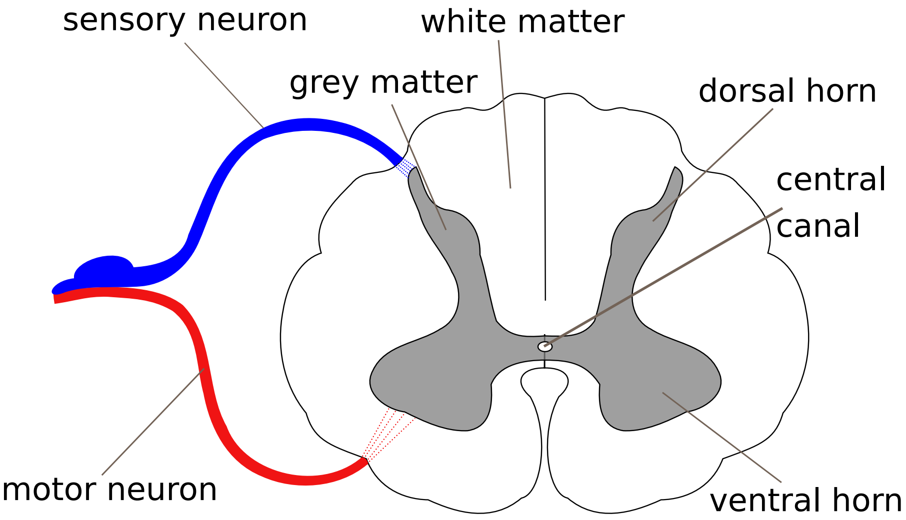


Figure 1.2: **Transverse section of the spinal cord** Grey matter which forms the center of the tissue is divided into dorsal and ventral horns. Spinal nerves couples emerge from the space between two vertebrae on each side of the spine. The dorsal horn contains neuronal cells incoming from the dorsal root ganglion and carries sensory inputs relayed from the periphery to the central nervous system. The ventral horn neurons carry information from the brain to the periphery, synapsing onto skeletal muscle cells. Adapted from Wikimedia Commons: https://commons.wikimedia.org

1.2.3 Brain energy metabolism

The average human adult brain comprises around 2% of body weight, yet accounts for 20% of the oxygen consumption [15, 16], a proxy for the total amount of energy utilization by this relatively small organ. The biochemical pathways of the ATP production in brain are the same as for other human cells, involving the oxidation of glucose via glycolysis, and subsequently mitochondrial oxidative phosphorylation.

Early studies of *in vivo* cerebral metabolism based on measurements of the arterio-venous differences in oxygen and glucose levels showed that glucose is consumed in nearly stoichiometric amounts with the oxygen required for the complete oxidation of glucose [15]. The molar ratio consumption of these substrates is the oxygen-to-glucose index (OGI), and is a widely used measure to describe the fate of glucose in the brain, with a ratio of six suggesting that all glucose is utilized via the mitochondrial oxidative pathways. A recent meta-analysis by Blazey et al. [17] showed that in the human, experimentally recorded values of OGI range from 4.5 to 7.5, with a median of 5.5. The discrepancy between this value and the theoretical value of six suggests glucose processing via non-oxidative pathways, incomplete oxidation leading to a net release of lactate, or others. Value of OGI above six suggest utilization of alternative energy substrates by the brain. Since the OGI is close to the value of six, it is assumed that glucose is the near exclusive energy substrate for cerebral ATP production overall.

Development of the 2-[^{14}C]deoxyglucose (2-DG) tracer technique [18] allowed for the determination of glucose utilization rates across all brain regions simultaneously in animal, and revealed that the rate of glucose consumption is unequal across the brain [19], with the grey matter requiring more ATP than white matter overall [19]. The majority of the metabolic energy consumption in the grey matter is required to sustain processes occurring at chemical synapses which underlie neuronal signalling. This includes restoration of ionic gradients necessary for the propagation of action potentials, reuptake of neurotransmitter, and others. Since excitatory glutamatergic synapses are the predominant type of chemical synapse, outnumbering inhibitory synapses by 9 to 1, glutamatergic neurotransmission is the major consumer of ATP in the brain. This has been supported by studies showing a co-localization of mitochondria with the glutamatergic synapse components [20], as well as the observations that glucose use correlates with the rates of conversion of glutamate to glutamine [21]. Regional glucose consumption is linked to neuronal activation, and blocking neuronal signalling via deep anaesthesia in the cortex leads to an 80-90% reduction in energy consumption [21] showing that in the grey matter, neuronal activation is the major consumer of energy, as opposed to cellular maintenance processes. Given the prevalence of glutamate in the brain, the majority of research regarding brain energy utilization has focused on neurons and the supporting glial cells which compose the glutamatergic synapse.

The Astrocyte-Neuronal Lactate Shuttle

Early studies using isolated cell cultures showed that CO_2 production in neurons is greater than that in the astrocyte [22, 23], suggesting a higher rate of mitochondrial oxidative phosphorylation in the former. Pellerin and Magistretti [24] observed that upon stimulation with glutamate, astrocytes increase their rate of 2-DG phosphorylation and lactate release, which led these authors to propose a mechanistic model for the metabolic cooperativity between these two cell types, named the Astrocyte-Neuronal Lactate Shuttle (ANLS). This theory proposes that there is dissimilar preference for the use of glucose and lactate to fuel energy production in astrocytes and neurons, which work synergistically to support glutamatergic neurotransmission. Stimulated by glutamate uptake, astrocytes oxidise glucose to lactate which is released into the extracellular environment. Lactate is then taken up by the surrounding neurons which oxidise it for energy production. The authors state that astrocytes have a higher activity of glycolysis in comparison to neurons, which produce ATP via oxidative phosphorylation. According to the ANLS theory, neuronal glucose is primarily consumed via the pentose phosphate pathway (PPP), which produces NADPH required for the reduction of glutathione (GSH), necessary for the neutralization of reactive oxygen species (ROS) produced in the oxidative mitochondrial pathways. Evidence supporting the theory was summarised in a review by Magistretti and Allaman [25] and includes, amongst experimental observations, a summary of the differential expression of genes encoding the enzymes of central metabolism which led the authors to conclude that the enzymatic signature contained in these cells supports the ANLS.

However, no technique has yet been developed which may directly prove the theory *in vivo*, and there exists multiple lines of evidence against the ANLS. Astrocytes contain high densities of mitochondria, and express enzymes of the Tricarboxylic acid Cycle (TCA) and well as those encoding components of the Electron Transport Chain (ETC) [26] suggesting the activity of oxidative metabolism. As opposed to the findings reported by Pellerin and Magistretti [24], other groups showed that glutamate does not have an effect on 2-DG phosphorylation, and later studies utilizing *in vivo* NMR spectroscopy showed that the fraction of oxygen utilization by astrocytes is similar to the average volume fraction occupied by these cells in the brain [27, 28]. Another objection to the ANLS theory states that it does not take into account the capability of astrocytes to oxidise glucose-derived neurotransmitters, most notably glutamate [29–31] which is present in large concentrations intracellularly and may significantly contribute to the total ATP production in these cells.

Transport of glucose and lactate is mediated by the monocarboxylate transporters (MCT) which are expressed across various CNS cell types [32], thus utilization of lactate in one cell which originates from another may take place, but it remains underdetermined whether each cell type has a preference for either substrate for the production of energy. As summed up in a recent review by Dienel [33], it is likely that the metabolic pathways of energy production in the brain require a more complex description, which depends on factors such as substrate availability, physiological state of the cell, and others. These factors are difficult to control in the organism *in vivo*, and thus computational modelling can greatly aid in the probing of questions pertinent to changes in the metabolic pathways utilized by different CNS cell types.

Blood Brain Barrier

Metabolic pathways in the brain are dependent on the passage of oxygen, as well as carbon and nitrogen sources from the circulating blood to the brain. The restrictive properties of brain microvasculature, composed of endothelial cells held together by tight junctions, blood vessels, immune system cells, astrocytes and others is collectively termed the blood brain barrier (BBB). Movement of molecules across the BBB is a regulated process which allows for maintenance of CNS homeostasis and protection of the brain against pathogens, toxins, and disease. Hydrophilic substrates such as glucose must transverse the BBB via carrier-mediated systems of transporters which are asymmetrically expressed in the luminal (blood side) and abluminal (brain side) membranes of the endothelial cells [34]. Glucose is transported across the BBB via the Glucose transport protein (GLUT), but during starvation the brain can consume acetoacetate and hydroxybutyrate to support its metabolic needs by uptake via monocarboxylate transporters, the expression of which increases in response to prolonged fasting [35]. Uptake of amino acids is a tightly regulated process, mediated by a system of facilitative amino acid transporters expressed on each side of the endothelial cell. In general, movement of essential, neutral amino acids is greater than that of nonessential amino acids and since the concentrations of all amino acids (except for glutamate) is higher in the circulating blood than in the CSF, amino acids most likely leave the brain against a concentration gradient [36]. The uptake of neurotransmitter amino acids such as glutamate and glycine is minimal, in order to ensure no uncontrolled activation of their respective receptors, therefore these amino acids are mostly synthesized within the brain rather than imported from blood circulation.

1.2.4 Glycine metabolism in the CNS

Pioneering experiments studying the role of glycine in the CNS by Aprison et al. [37] showed that it is widely distributed across the brain, with highest concentrations found in the brain stem and the spinal cord in the *post mortem* brains of vertebrates, which was later confirmed in autopsied human brains by Perry et al. [38]. Measurements of the rate of labelled C^{14} atom incorporation from glucose and serine into glycine by Shank et al. [39] showed that the ratio of carbon label from glucose was similar to that obtained from serine, indicating that the latter is the primary source of glycine in the CNS via the serine hydroxymethyltransferase (SHMT) catalyzed reaction. Another potential source of glycine suspected at the time was from the transamination of glyoxylate by the serine-pyruvate transaminase enzyme, which was shown in rat spinal cord extracts *in vitro* [40]. However, ratios of labelled serine and glycine derived from glucose as shown by Shank et al. [39] suggested that this pathway does not significantly contribute to glycine biosynthesis in the brain *in vivo*. In addition, expression of genes encoding the SHMT is high in comparison to those encoding the transaminase enzyme in the brain, which has mostly been found in the human liver so far [41] thus it is unlikely that glyoxylate is available in the brain.

Serine hydroxymethyltransferase catalyzes the reversible conversion of serine and tetrahydrofolate into glycine and methylene-THF . Apart from its role in glycine metabolism, it is an important enzyme in the reactions of one-carbon metabolism of folates in which it significantly contributes to the availability of 1C units for the synthesis of purines, deoxythymidine monophosphate (dTMP), as well as methionine which is adenylated to S-adenosyl methionine (SAM), a methylation cofactor [42]. In humans, two genes encoding SHMT1 and SHMT2 are present encoding the cytoplasmic and mitochondrial enzyme isoforms, respectively. This allows for the compartmentalization of folate metabolism, thus metabolic pools involved in folate cycling can be regulated independently in the cell. Experiments aiming to explain the function of SHMT compartmentalization showed that each isoenzyme regulates a different cellular process - SHMT2 is mainly responsible for the biosynthesis of glycine and 1C groups, whereas SHMT1 utilizes the mitochondrially-derived 1C groups for the production of purines, methionine, and dTMP [43]. Mice deficient in SHMT2 experience respiratory problems and die within the neonatal period [44], but those deficient in SHMT1 are viable, thus only the former has been shown to be essential in rodent models [43].

Degradation of glycine in the CNS is mediated by the glycine cleavage system (GCS) which is composed of four proteins loosely associated in the mitochondrial membrane which catalyze a rather complex reaction resulting in the release of methylene-THF , CO_2 and NH_3 [45]. This reaction was shown to be readily reversible in the rat liver mitochondria *in vitro* [45], although it is uncertain if this may occur *in vivo* [46]. Although the expression of SHMT is not cell-specific, immunohistochemical staining of the various regions of a rat brain by Sato et al. [47] showed that the P-protein component of the GCS was confined to astrocytes, suggesting that glycine breakdown is mediated exclusively by these cells.

Folate-cycling metabolic pathways are essential in cellular biosynthetic processes, thus the disruption of these pathways has long been identified as potential treatment against cancer. Antifolate chemotherapeutic agents were the first class of antimetabolite drugs to be administered clinically [48], and include currently widely used medications such as methotrexate and pemetrexed which affect tetrahydrofolate biosynthesis [49]. It follows that serine/glycine metabolism has been identified as a potential target for novel chemotherapy drug development, as has been previously reviewed in several papers [50–52]. Relevant to brain malignancy, increased glycine concentrations are positively associated with the presence and stage of human gliomas [5, 6]. In addition, overexpression of genes related to serine/glycine biosynthetic pathways, phosphoglycerate dehydrogenase (PHGDH), SHMT2 and GCS has been found in a subset of human cancers [7–10], including gliomas [6, 53]. The ability of serine to support proliferation was shown in human cell cultures in *in vitro* [54], suggesting that this amino acid may be capable of supporting not only biosynthetic, but also energy production pathways in a cell although exactly how serine/glycine metabolic pathways support cell proliferation is not known.

1.2.5 Glycinergic neurotransmission

Following the studies by Aprison et al. [37] indicating high glycine concentrations in the ventral horn of the spinal cord, subsequent electrophysiological experiments showed that glycine application causes a hyperpolarization of motorneurons located in this region, which further implied that glycine is an inhibitory neurotransmitter [55]. Glycine was found to increase Cl^- ion conductance of the neuronal membrane, and was inhibited by the alkaloid strychnine, the binding of which was Cl^- and glycine-dependent in the spinal cord of a rat [56]. Subsequently, the subunits composing the rat glycine receptor (GlyR) protein were purified by affinity chromatography by Pfeiffer et al. [57], establishing conclusive evidence for glycine's role as an inhibitory neurotransmitter in the CNS.

High sequence homology with the nicotinic acetylcholine receptor (nAchR) led to the classification of GlyR as a member of the Cys-loop family of ligand-gated ion channels (LGICs), which share structural features [58]. LGICs can be expressed as homomeric or heteromeric proteins, composed of five subunits arranged symmetrically in a ring with a central ion-conducting pore. They are transmembrane proteins, containing an extracellular ligand-binding domain, a transmembrane domain, and an intracellular domain which takes part in receptor modulation [59]. The heteromeric GlyR is composed of the α subunit, four isoforms of which have been identified [60] and the β subunit, which is mainly involved in the anchoring of the receptor to the membrane via the protein gephyrin [61] . Although α subunits can assemble to form functional homomeric GlyRs, expression of which has been found in embryonic neurons [62], the distribution of the heteromeric receptor is more widespread in the adult brain [63], thus it is likely to mediate majority of the glycine-induced inhibition.

The glycinergic inhibitory synapse is composed of GlyR expressed in the membrane of the postsynaptic neuron, as well as proteins conveying the events of glycine release and reuptake (Figure 1.3). Cytoplasmic glycine localized in the presynaptic neuron is concentrated in the synaptic vesicle via the vesicular inhibitory amino acid transporter (VIAAT), which uptakes both glycine and the other major inhibitory neurotransmitter, γ-aminobutyric acid (GABA). Inhibitory glycinergic neurons may therefore utilize glycine only, or mixed neurotransmitter GABA and glycine which is thought to be determined by the relative concentrations of each in the cytoplasm, affecting the final concentrations in the synaptic vesicle [64]. Fusion of the synaptic vesicle is initiated by the binding of Ca^{2+} ions, the concentrations of which increases following an action potential arrival in the presynaptic cell. Upon release, glycine diffuses across the synaptic cleft, binding to the GlyR present on the opposing neuron which increases its Cl^- conductance.

Termination of this process is mediated partially by passive diffusion of glycine away from the synapse, as well as the active reuptake by two transporter proteins, the glycine transporter 1 and 2 (GlyT1, GlyT2). Up to date, five variants of the GlyT1 (a-e) and three variants of GlyT2 (a-c) have been identified which occur as a result of alternative gene promoter usage and/or splicing. The glycine transporters are members of the SLC6 Na^+ /Cl^- -dependent neurotransmitter transporters family which also include transporters for other amino acids such as taurine, as well as neurotransmitters dopamine, serotonin, and others [65]. The SLC6 Na^+ /Cl^- family proteins share structural, and thus functional features which determine the mechanism of transport. Translocation of ligand is coupled with Na^+ and Cl^- ions, utilizing the energy provided by their electrochemical gradients to convey neurotransmitter transport. Although the types of ions co-transporter is the same for all transporters across the family, the numbers of ions transported on each cycle differ. Electrophysiological experiments in combination with radiolabelled glycine uptake studies by Roux and Supplisson [66] showed that GlyT1 binds Na^+ , Cl^- and glycine in a 2:1:1, whereas GlyT2 in a 3:1:1 ratio. The difference in the stoichiometry has important implications, since the binding of the extra Na^+ by GlyT2 provides a bigger driving force for glycine reuptake by this protein, and affects its ability for reverse glycine transport [66]. Indeed, the cellular distribution of each transporter isoform, in combination with studies using pharmacological agents which inhibit reuptake suggest that they have different roles in the CNS. GlyT1 is expressed across most regions of the brain, and is located mostly to the glial cells surrounding the synapse. Across the inhibitory glycinergic synapses, GlyT1 is thought to regulate the time course of the neurotransmitter transient. GlyT1-deficient mice have respiratory defects and die within the first postnatal day. Measurements of the time course of the inhibitory postsynaptic currents (IPSC) in the brain stem of these mice shows that the IPSC decay time constant is increased, thus the phenotype is likely caused by over-inhibition of the respiratory centers, due to increase glycine concentrations which persist for longer in the synapse [67]. Apart from its role across the inhibitory synapses, in other regions of the brain such as the forebrain GlyT1 is important in the regulation of N-methyl-D-aspartate (NMDA) receptor which binds glycine as a co-substrate with glutamate [68]. On the other hand, GlyT2 expression closely follows that of the GlyR [69], where it is thought to be mostly implicated in the refilling of the presynaptic vesicle with glycine, necessary for effective uptake by the VIAAT [70], which has a low affinity for glycine [71]. The amplitude of IPSC recorder from the spinal cords of mice deficient in GlyT2 show a marked reduction, suggesting insufficient glycine concentrations in the presynaptic cell.

1.2.6 Glycine and human health

Glycine receptor immunoreactivity was found to be particularly high in the lower brain regions of the brain stem and the spinal cord, the retina [72] but also in the higher brain regions such as the hippocampus and the forebrain structures of the cerebral cortex [73]. In the spinal cord, glycinergic interneurons regulate motor control, such as via the reciprocal inhibition of excitatory neurons coordinating the antagonistic relaxation of skeletal muscle [74].

Mutations in glycinergic synapse components, most notably GlyR, as well as GlyT2 results in hyperkplexia, which is characterised by uncontrolled startle reflex, as well as neonatal hypertonia in

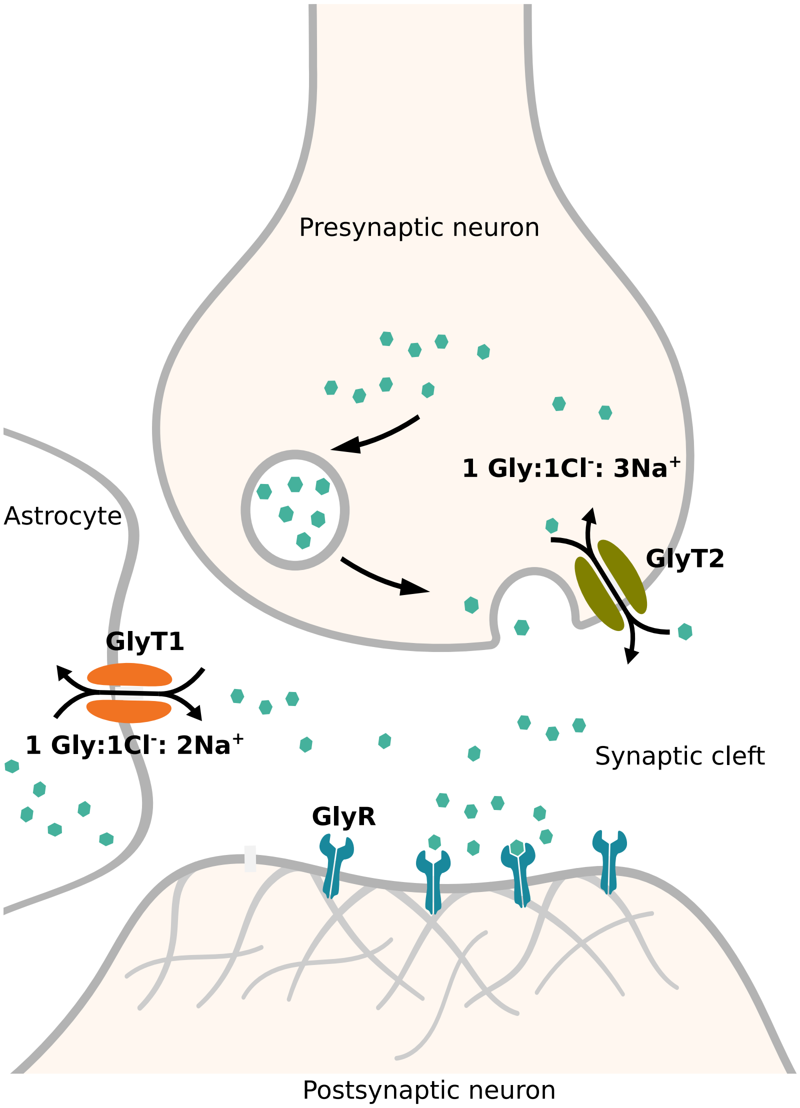


Figure 1.3: **Glycinergic inhibitory synapse** Glycine available in the cytosol is packaged into synaptic vesicles which fuse with the membrane releasing the neurotransmitter, triggered by an increase in the concentration of Ca^{2+} ions. Glycine diffuses across the synaptic cleft, binding to the GlyR proteins present on the opposing neuron, which allows an influx of Cl^- ions leading to cell hyperpolarization. Glycine concentration is reduced partially by passive diffusion away from the synaptic cleft, as well as active reuptake by two transporter proteins, GlyT1 and GlyT2. Adapted from Wikimedia Commons: https://commons.wikimedia.org

response to tactile, auditory and other stimuli [75]. Although the onset of the condition may occur at any age, it most often presents in early life, and in the neonatal period can cause severe apnea and even death. As the disease progresses, life-threatening events subside but are replaced with startle responses which present as muscle stiffenings in the affected individuals. The condition is caused by reduced glycinergic inhibition, and current treatments include the use of benzodiazepines which enhance GABA-mediated inhibitory action, compensating for the loss of glycinergic inhibition.

In the spinal cord, glycine is also involved in the regulation of the somatosensory system, and reduction in the inhibitory neurotransmission is thought to contribute to the increased sensitivity in neuropathic pain. Therefore, increasing glycine concentrations in these regions was identified as a possible treatment option which may lead to analgesic effects in affected individuals. Since the BBB restricts the uptake of glycine, research has focused on increasing its concentration in the brain *in situ*, leading to the development of GlyT inhibitors, which act by decreasing glycine reuptake during neurotransmission. Although various GlyT inhibitors have been shown to reduce nociception in animal pain models, acute inhibition was also shown to cause side effects including respiratory distress and at certain concentrations, an increase in nociception [76–78]. In addition, it has been proposed that GlyT inhibitors could be used in order to control seizures in epileptic patients unresponsive to currently available GABA therapy, as well as in the facilitation of controlling alcoholism, obsessive compulsory disorders (OCD), and others via the dampening of the dopaminergic reward systems in the nucleus accumbens of the midbrain.

Glycine is an essential co-agonist of the NMDA receptors which are critical in the cellular mechanisms responsible for learning and memory, the disruption of which is thought to be implicated in the pathophysiology of schizophrenia [4]. It was therefore hypothesized that regulation of glycine concentrations across synapses containing NMDA receptor activity can be used to alleviate symptoms in patients, and the most advanced clinical trials of GlyT inhibitors were performed in the pharmaceutical development of novel schizophrenia treatments. In the first instance, GlyT1 inhibitors were developed with the goal of increasing NMDA receptor activity in the forebrain, but it has been since recognised that effects may also be relayed by inhibitory glycinergic circuits regulating dopamine and glutamate across various areas of the brain, alleviating negative symptoms of schizophrenia in patients, such as poor working memory [4]. Although numerous compounds were shown to be promising in controlling schizophrenia symptoms in animals models, bitopertin, a GlyT1 inhibitor went the furthest in clinical trials, however phase III results show that it was ineffective across a group size of roughly 600 schizophrenia patients as compared to placebo [79].

Dysregulation in glycinergic inhibition can also be caused by defects in enzymes involved in glycine biosynthesis. Nonketotic hyperglycinemia (NKH) is an inborn condition which is characterised by lethargy, hypotonia, and apnea in newborns with mortality rates at 50% within the first neonatal week [3], with surviving infants experiencing severe motor and cognitive impairments. It is caused by mutations in the GCS, responsible for glycine breakdown and thus the condition is recognised by the elevated glycine concentrations in the cerebrospinal fluid, which most likely leads to overactivation of inhibitory glycinergic and excitatory NMDA synapses. To date, no effective treatments for NKH have been developed [3].

Majority of patients presenting with NKH are born with mutations in the GCS components, although recently a novel NKH-causing mutation was identified in a child which resides in the Ser 407 position of the GlyT1 [80]. This is the first time a mutation in the GlyT1 was linked to NKH, and the functional effect of amino acid replacement at Ser 407 position of GlyT1 has not yet been established.

1.3 Computational modelling

1.3.1 Introduction

The application of mathematics to describe and study interactions within biological systems emerged independently across various disciplines aiming to rigorously describe diverse phenomena including laws governing biochemical reactions, electrical properties of neuronal cells, or the growth of populations of a species. These systems differ widely, varying in their complexity with respect to the number of interacting components, relative time scales in which events occur, and physical constrains they are under, leading to an equally wide range of theoretical descriptions and methods of analysis. The overall goal of all, however is to develop a mathematical abstraction of a system based on previous observation for the purpose of understanding the system and furthermore, predict its future behaviours. Mathematical modelling also allows one to set up experiments in systems that would prove difficult or otherwise impossible to perform in a laboratory setting, which is limited to the techniques available.

Modern mathematical models describing protein interactions are rooted in principles derived from the experimental observations in the late 19th and early 20th centuries. In 1864 Guldberg and Waage described the law of mass action, which states that the rate of a chemical reaction in equilibrium is proportional to the concentration of its product. Considering a reaction of the form:

$$aA + bB \underset{k_r}{\overset{k_f}{\rightleftharpoons}} cD + dD \tag{1.1}$$

where A and B are the substrates, C and D are products, a, b, c and d are the stoichiometric coefficients describing the molar proportions of the species, and k_f and k_r are the reaction rate constants. The *order* of reaction is related to how many molecules are changed in a reactions, thus the above example reaction is *second order*. Applying the law of mass action yields two relationships:

$$V_f = k_f A^a B^b \tag{1.2}$$

$$V_r = k_r C^c D^d \tag{1.3}$$

At equilibrium, the rates of the overall forward and reverse reactions are equal, thus there is no change in the composition of the medium. This relationship can be expressed as:

$$k_f A^a B^b = k_r C^c D^d \tag{1.4}$$

which allowed for the description of a new constant, the *equilibrium constant* , K_{eq} :

$$K_{eq} = \frac{k_f}{k_r} = \frac{C^c D^d}{A^a B^b} \tag{1.5}$$

The equilibrium constant expresses the relationship between the concentrations of substrates and products and remains the same regardless of the concentrations or absolute rates, at a given temperature and pressure. Applying the law of mass balance, differential equations describing the rate of reactant concentration changes over time can be written, for instance the rate of change of C can be expressed as:

$$\frac{dC}{dt} = V_f - V_r = k_f A^a B^b - k_r C^c D^d \tag{1.6}$$

This description of chemical reactions provide the foundation for deriving differential equations that quantify the relationships between concentrations and reactions rates which can be measured via experimental manipulation of variables in the equation.

Observations of enzyme-catalyzed reaction rates indicated features which did not follow the law of mass action. It was noticed that the thermal stability of the enzyme is higher in the presence of the subtrate than in its absence, leading to the idea of the enzyme-substrate complex. The enzyme was not destroyed, even after catalysing the hydrolysis of a much higher weight of its substrate than itself. Furthermore, there existed an optimal temperature at which the reaction proceeded the fastest, unlike second-order kinetic reaction rates [81]. Amongst others, these observations led to the description of biological catalysts and the kinetic reaction rate formalisms which followed. As described by Cornish-Bowden [82], the equation describing the rate of progress of enzyme-catalyzed reactions was first described by Henri, and the advantage of the use of the equation as well as further experimental contributions achieved by Michaelis and Menten. The Michaelis Menten equation is of the following form:

$$V = \frac{V_{max} S}{K_m + S} \tag{1.7}$$

and describes the kinetics of saturable reactions involving a catalyst which converts product into substrate, without itself taking part in the reaction. The parameters of the equation are V_{max} , the maximum forward rate of the reaction, S, the concentration of substrate and K_m, the Michealis constant. Enzyme catalysts function by increasing the speed of a reaction via lowering the activation energy required for the reaction to proceed under given conditions, without affecting the equilibrium of the reaction. They bind their substrates at the catalytic site at which the reaction occurs, and may undergo some structural changes during the reaction progress but overall are not changed in the reaction, thus are capable of catalysing multiple rounds of the reaction.

Although the Michaelis Menten equation could successfully fit reaction rate curves of many single-substrate enzymes, it could not account for the sigmoidal shapes of rate curves observed for other types of reactions. This type of curve implied a different method of regulation, with a higher sensitivity of substrate to its protein. As previously, novel observations of how the rate of a reaction changes according to the substrate concentrations lead to a need for a novel explanation. The first model of cooperativity was proposed by Hill who used it to fit the rate of oxygen binding to haemoglobin, and is of the form:

$$V = \frac{V_{max} S^h}{K_{eq}^h + S^h} \tag{1.8}$$

It can be seen that this equation is very similar to the Michaelis Menten equation, with the exception of the introduction of a new parameter, h, which is the Hill's coefficient of cooperativity. The Hill equation is a *phenomenological* model, meaning that it fitted the experimental observations but did not provide a mechanistic interpretation for what happens between the substrate and protein. This was offered by the development of *mechanistic* models of cooperativity, which were derived based on the assumed sequence of substrate binding and the associated conformational changes that the protein undertakes and include the Adair equation, the Koshland-Nemethy-Filmer (KNF) model, as well as the Monod-Wyman-Changeux (MWC) model.

Cooperativity manifests as an apparent higher affinity of the enzyme to the substrate following the binding of a preceeding substrate. It is an instance of *allostery*, which is the phenomena observed when an effector binding at one site of the enzyme induces a change in the conformation of the enzyme far away in space from its binding site. The MWC model [83] assumes that the enzyme is composed of several subunits which possess at least one axis of symmetry that can exist in different global conformational states even in the absence of any bound substrate. The affinity of the substrate differs towards each global conformational state, and its binding shifts the equilibrium towards one or the other conformational state, inducing a global concerted movement in all subunits. This is in contrast to the KNF model, which assumes that the binding of ligand induces independent conformational changes in protein subunits.

Since their conception, these well-known equations have been applied to describe the progress of various types of proteins interacting with a substrate. The Michealis Menten equation, as well as the Hill equation have the advantage of simplicity, and contain terms which can be easily interpreted and compared between different protein-substrate interactions. Thus, outside of enzymology they have found application in the field of quantitative pharmacology to describe the rates of transport of substrates across biological membranes by transporter proteins, as well as ligand-receptor binding reactions. Computational fitting of these equations to the reaction rates curves allows for the estimation of the maximal reaction rate, as well as concentrations of substrate which produces half-maximal effects, parameters that can be easily understood. In addition, the extension of these equations by terms describing the action of inhibitors is used to directly compare the effects of different compounds on the same protein, and has been applied in the field of drug discovery.

The MWC model has also seen wide use in biology. Experimental X-ray crystallography studies have shown that numerous enzymes, as well as signalling and transport proteins are oligomeric with one or more axis of symmetry [84], which is in agreement with the original MWC postulates. Since its conception, the model has been applied to describe the activation of G-protein coupled receptors [85], chaperones [86], and others [87]. Perhaps one of the earliest protein families the MWC was applied to were the ligand-gated ion channel receptors (LGICs) (Section 1.2.5), including the nicotinic Acetylcholine receptor and later on, the glycine receptor. Since the original MWC did not provide a good fit to some of the experimental data obtained from the single-channel study of LGICs, multiple models have been proposed which built upon the original assumptions, nevertheless retaining the overall features predicted by the MWC [88].

1.3.2 Kinetic models

Due to the normally low levels of expression in cells, study of proteins often involves overexpression of the protein in question in model organisms, which is far from the *in vivo* environment the protein normally operates in. This is invaluable in elucidation of the single protein structure and function but lacks the context within which the protein operates. Kinetic modelling of systems combines the knowledge of numerous single-protein reactions and can aid in the understanding of component interactions within the system as a whole. Kinetic models are a type of mathematical model and include sets of equations which describe the time evolution of events within a system composed of multiple reactions. Often, multiple kinetic mechanistic models may be proposed for a set of experimental observations and that which provides the best fit to the data offers an insight to the sequence of events that take place.

The importance of studying component interactions is central to the field of systems biology, which has strong philosophical underpinnings characterized by a holistic approach to understanding life. This is in opposition to reductionistic approaches, which is the idea that complex phenomena can be understood via the study of its principal components. For instance, although individual properties of enzymes can be quantified in isolated experiments *in vitro*, this will not provide an understanding of their action within a cell cytoplasm which contains numerous enzymes that form metabolic pathways which operate simultaneously.

Kinetic models in biology are often composed of a set of mass balance first order differential equations (ODEs). This requires the definition of the species involved in each reaction, as well as the appropriate kinetic rate equations which describes the speed of progress of each reaction. Which appropriate rate equation to use is decided based upon previous experimental observations and research by the modeller. Naturally, the model aims to represent a simplified nature of the biology in question, thus experimental data is normally collected prior to model construction.

Consider the set of biochemical reactions in Figure 1.4, which can be described by a system of ODEs as follows:

$$\begin{aligned}
\tfrac{dA}{dt} &= v_{R_1} + v_{R_3} - v_{R_2} \\
\tfrac{dB}{dt} &= v_{R_2} - v_{R_3} - v_{R_4} - v_{R_5} \\
\tfrac{dC}{dt} &= v_{R_4}
\end{aligned}$$

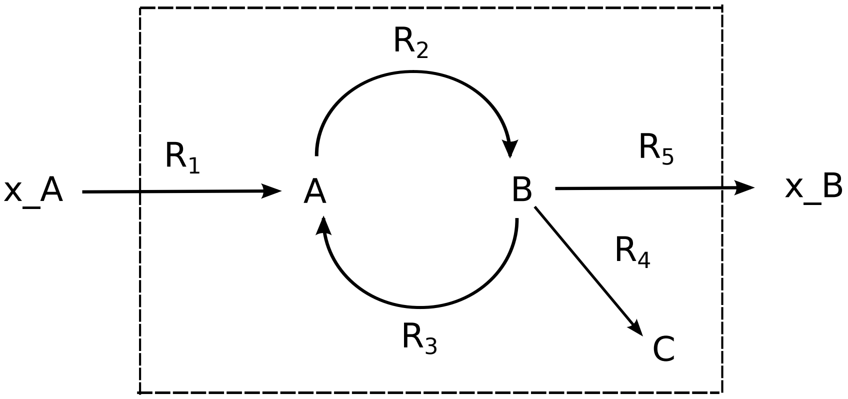
For convenience, it is assumed that the concentrations of external metabolites x_A, x_B are constant thus are not changing in time. Biologically, these metabolites are those whose concentration is high inside the cell thus changes in their concentration are considered negligible, or those which are maintained at a constant level outside of the system. The internal metabolite concentrations $(A - C\ )$ are non-constant and are a function of initial concentrations, time, t, and the kinetic parameters defined in the rate equations describing the progress of reactions $v_{R_1} - v_{R_5}$. The kinetic rate equations used may be in the form of mass action statements, as in Equation 1.6, the Michaelis Menten equation, or other. The initial conditions, at $t = 0$, including metabolite concentrations and the kinetic parameters of rate equations are set by the modeller and provide a starting point for the model simulation. The system of ODEs is integrated by a computer, where the evolution of the dependent parameters, which are the metabolite concentrations and reaction rates are calculated over the independent parameters, t, as well as the kinetic parameters.

Although ODE models are most common, kinetic models vary greatly in their mathematical composition. For instance, biophysical models of neurotransmitter diffusion away from the site of release are often composed of partial differential equations (PDEs) capable of calculating species concentration evolutions in time as well as space [89]. A big distinction within kinetic modelling is their division into deterministic and stochastic models, with the ODEs systems falling in the former. Stochastic models are more complex, and take into account random variability in the behaviour of biological systems which is needed to be taken into account in the modelling of small scale systems.

1.3.3 Structural models

Structural modelling of metabolism represents metabolic interconversions within a system under study, but is concerned only with the analysis of the network based on its topology rather than reaction time evolutions. The model is composed of a set of enzyme-catalysed reactions including their stoichiometry, but excluding the definition of rate equations and kinetic parameters that describe their progress. The set of reactions in the system follows the fundamental laws of mass conservation as all models describing material transfers and this, in combination with a high level of metabolite connectivity in the network imposes a powerful set of constraints from which insight can be gained. For instance, structural models can be used to discover those reactions in the system which are essential for a given pathway, or those sets of reactions which must proceed at proportionate velocities in the network. It can be argued that the analysis of the network structure is of most importance in modelling overall, and may also provide insight into the possible rates at which reactions proceed, even if they are not known. Furthermore, the appropriate choice of rate equations and collection of kinetic parameters for kinetic modelling often proves difficult, and structural modelling circumvents these issues altogether.

The set of ODEs describing the network in Figure 1.4 can be represented in matrix format:



$$\begin{array}{c|ccccc|} & R_1 & R_2 & R_3 & R_4 & R_5 \\ A & 1 & -1 & 1 & 0 & 0 \\ B & 0 & 1 & -1 & -1 & -1 \\ C & 0 & 0 & 0 & 1 & 0 \end{array}$$

Figure 1.4: **Representation of a reaction network in a matrix format** An example of a small biochemical reaction network can be seen in the diagram. Species with the 'x_' prefix are considered external to the system, thus are not included in the stoichiometry matrix. System of the internal reactions can be represented mathematically in a stoichiometry matrix, in which the rows describes metabolites and columns describe reactions. Elements of the matrix are the stoichiometric coefficients, where the negative signs mean the metabolite is a substrate and positive sign mean the metabolite is a product. If a metabolite does not take part in a reaction, the stoichiometric coefficient is zero.

$$\begin{bmatrix} \frac{dA}{dt} \\ \frac{dB}{dt} \\ \frac{dC}{dt} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} v_{R_1} \\ v_{R_2} \\ v_{R_3} \\ v_{R_4} \\ v_{R_5} \end{bmatrix}$$

where the changes in the concentrations of metabolites, collected in the most left vector, can be calculated from the *stoichiometry* matrix, multiplied by the vector of reaction rates. In short notation, this can be written as:

$$\frac{d\mathbf{s}}{dt} = \mathbf{N}\mathbf{v} \tag{1.9}$$

where $\mathbf{s}$ is the vector of internal metabolite concentrations, $\mathbf{N}$ is the stoichiometry matrix, and $\mathbf{v}$ is the vector of reaction rates. The internal stoichiometry matrix includes the set of reaction stoichiometries which involve the transformations of metabolites considered internal to the system. External metabolites are excluded, since their concentration is assumed to be constant. The stoichiometry matrix rows correspond to metabolites, whereas columns correspond to reactions, and the entries are filled with the stoichiometric coefficient of a metabolite taking part in a specific reaction (Figure 1.4). For example, since one mole of A is produced in the reactions R_1 and R_3, the entry in the stoichiometry matrix is positive one. One mole of A is consumed in the reaction R_2, thus the entry is minus one, and since A does not take part in reactions R_4 and R_5 the entries are zero.

The stoichiometry matrix allows for the inclusion of all reactions present in the system in a single mathematical notation. Matrix transformations, common to the field of mathematics allow for the discovery of interesting properties belonging to the system.

Structural modelling assumes that the system is in *steady state*, meaning that the changes in metabolite concentrations, $\mathbf{s}$, which are a product of the concentrations of metabolites $\mathbf{N}$ and reaction rates, $\mathbf{v}$ is equal to zero, ie.

$$\frac{d\mathbf{s}}{dt} = \mathbf{N}\mathbf{v} = 0 \tag{1.10}$$

The assumption of steady state means that the system maintains the internal concentrations constant over time, such that the net consumption of external metabolites that are consumed equals the net formation of excreted products. Fluctuations in enzyme concentrations which catalyse internal reactions, such as degradation or synthesis happen on a much longer time scale than enzymatic transfomations, thus are considered external to the system [90]. The steady state assumption is very powerful, as the relationships defined by the directionality of the reactions, stoichiometric coefficients, as well connectivities between the metabolites only can provide lots of information regarding the interactions between these variables. The addition of reactions with known rates further allows for the calculation of a physically meaningful rates vector $\mathbf{v}$ [91], as described in the following sections.

Analysis of the right nullspace of the stoichiometry matrix

The relationships between reaction rates contained within vector **v** as imposed by the steady state constraint have useful biological interpretations [92]. Linear algebra can be applied to the matrix in a set of transformation, from which insight into the boundaries within which the biological system operates can be gained. The complete set of possible solutions to the Equation 1.11 forms a subspace of the stoichiometry matrix called the null space, or the *kernel* of the matrix, **K**, thus Equation 1.11 becomes:

$$\mathbf{NK} = 0 \tag{1.11}$$

An example of a kernel of the stoichiometry matrix shown in Figure 1.4 is:

$$K = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}$$

The row order of the kernel is the same as the column order of **N**, and each column in the kernel represents a flux distribution vector at steady state. Any steady state flux vector can be further written as a scaled combination of the kernel column vectors, for example:

$$K = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} v_{R_1} \\ v_{R_2} \\ v_{R_3} \\ v_{R_4} \\ v_{R_5} \end{bmatrix} = \begin{bmatrix} 1w_1 & +0w_2 \\ 1w_1 & +1w_2 \\ 0w_1 & +1w_2 \\ 0w_1 & +0w_2 \\ 1w_1 & +0w_2 \end{bmatrix}$$

From these examples, a number of important properties of the network can be understood. In the first instance, one can identify reactions in the network which cannot carry flux under any condition as given by those rows all elements of which are equal to zero, in this example, reaction 4. These reactions are considered *dead*, and are often a result of incomplete network annotation. Secondly, the kernel identifies sets of reactions which carry proportional or identical flux at steady state, which can be seen as proportional row vectors of **K**, in the above example reactions 1 and 5. These reaction sets are named *enzyme subsets*, and they identify reactions that are active in unison, therefore which reactions must *all* be active in a solution for *any* to be active; removal of any reaction from a subset renders the others dead [93]. Graphically, enzyme subset row vectors are parallel to each other in the solution space. An extension to the idea of enzyme subsets was to assess the correlation between all row vectors which was described by Poolman et al. [94], who formally defined the correlation between kernel vectors as the cosine angle between each vector in the orthogonal kernel form. The orthogonal kernel is defined as that which satisfies the relationship $KK^T = I$. The resulting value was termed the *reaction correlation coefficient* and is given by:

$$\phi_{ij} = cos(\theta_{ij}^{\mathbf{K}}) = \frac{\mathbf{K}_i \mathbf{K}_j^T}{\sqrt{\mathbf{K}_i \mathbf{K}_i^T}\sqrt{\mathbf{K}_j \mathbf{K}_j^T}} \tag{1.12}$$

where $\mathbf{K}$ is the orthogonal kernel matrix with m rows, $i, j \in 1...n$ and $\theta_{ij}^{\mathbf{K}}$ is the angle between rows i and j in $\mathbf{K}$. Mathematically, the RCC is identical to the Pearson's correlation coefficient, and can take any value between +1 and -1 where values of +1 identify reactions which carry proportional flux (identical to enzyme subsets) and 0 indicate those which are completely disconnected. Collection of all the ϕ_{ij} in a matrix notation forms the *dissimilarity* matrix, to which clustering algorithms can be applied such as the Weighted Pair Group Method using Arithmetic Averaging (WPGMA) algorithm in order to generate metabolic correlation trees, that can be used to visualise the relationships between entities [94].

Another property that can be calculated from the right null space is the set of sub-pathways that form small 'routes' within the network which can operate by themselves, called *elementary modes.* An elementary mode is a flux distribution such that it agrees with the steady state assumption, reversibility constraints imposed by the network, and cannot be decomposed into smaller modes (pathways). The last requirement means that the deletion of any reaction within an elementary mode renders all of the other reactions in that elementary mode unable to carry flux under steady state, thus identifying small stand-alone networks within the whole system (Figure 1.5). The concept was introduced by Leiser and Blum [95], later developed by Fell, D.A. [96], and the first algorithm for the calculation of elementary modes was described by Schuster and Hilgetag [97].

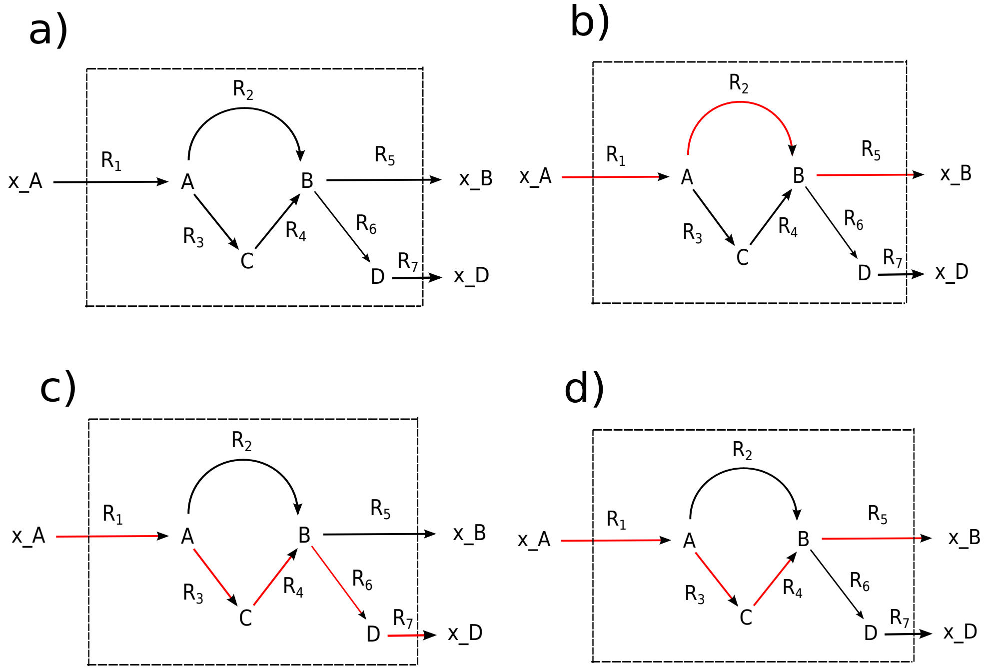


Figure 1.5: **Elementary modes within a small network** Reactions which form elementary modes within the hypothetical network *a* are indicated in red arrows. Elementary modes b and c are irreversible and have identical net stoichiometries, converting x_A to x_B.

Analysis of the left nullspace of the stoichiometry matrix

Just as the calculation of the right null space of the stoichiometry matrix is informative of the relationships between reaction fluxes, calculation of linear dependencies between row vectors can be used to derive relationships between metabolites. This information is found within the left nullspace and is defined as the complete set of vectors by which the transpose of the stoichiometry matrix, N^T , is multiplied such that the solution equals the null vector, ie.:

$$\mathbf{K^L N} = 0 \tag{1.13}$$

and can be calculated from:

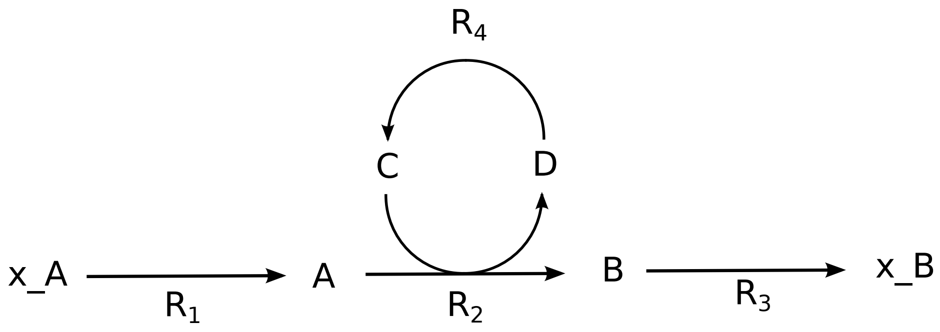
$$\mathbf{K^L} = \mathbf{K}(\mathbf{N^T})^T \tag{1.14}$$

where $\mathbf{K^L}$ is the nullspace, and $\mathbf{N^T}$ is the transpose of the stoichiometry matrix, $\mathbf{N}$. The left null space contains information about the conservation relationships between the concentrations of all metabolites within the model which is useful both in the area of structural and kinetic modelling. Linear dependencies allow for the reduction of the system's dimension since the dependent species can be calculated from the concentrations of the independent species [90]. With respect to metabolic modelling, the left nullspace can be used to identify moieties whose total concentration is constant, named *conserved moieties*, an example of which can be seen in Figure 1.6. Biologically, conserved moieties are atom groups which remain unchanged and often are cofactors such as NAD which take part in many reactions. Although their concentrations must vary according to initial biosynthesis and final degradation, it is assumed the interconversions between moieties is much faster compared to their overall turnovers [98].

Flux Balance Analysis

Methods of structural analysis can describe the relationships between model metabolites in the network, as well as reaction rates, without the knowledge of their absolute values. For example, analysis of the right null space can be used to address questions such as- what is the effect of doubling rate of reaction 1 on reaction 3? However, without the knowledge of some of the reaction fluxes these effects cannot be quantified. This is achieved by including experimentally determined flux values via the transport reactions defining the inputs and outputs. An example of such experimental data are the rates of production of bacterial biomass components which can be determined from the culture growth curves.

A commonly used technique for analysing metabolic models which makes use of the experimentally determined flux values is Flux Balance Analysis (FBA), which applies Linear Programming (LP) methods of optimization to calculate flux distributions under a set of constraints. Linear programming is a technique applied to a system of linear equations which finds the most 'optimal' solution



$$\begin{array}{c|cccc} & R_1 & R_2 & R_3 & R_4 \\ \hline A & 1 & -1 & 0 & 0 \\ B & 0 & 1 & -1 & 0 \\ C & 0 & 1 & 0 & -1 \\ D & 0 & -1 & 0 & 1 \end{array}$$

$$\frac{dA}{dt} = R_1 - R_2$$

$$\frac{dB}{dt} = R_2 - R_3$$

$$\frac{dC}{dt} = R_2 - R_4$$

$$\frac{dD}{dt} = R_4 - R_2$$

Figure 1.6: **Moiety conservation relationships** Set of the differential equations describing the rate of metabolite production can be derived from the left null space of the stoichiometry matrix. It can be seen from this example that metabolites **C** and **D** are produced and consumed by the same reactions in opposite directions, thus are conserved across the system. This means that the concentration of one of these metabolites only is needed to determine the other, reducing the dimensionality of the system. Although these relationships can be easily derived by hand for the network above, this does not hold true for larger networks including hundreds of reactions, thus providing an obvious advantage of their derivation by the transformation of the stoichiometry matrix.

for an *objective function* under a set of defined constraints. It is a widely used technique in mathematics, economics, and engineering. The application of LP to metabolic modelling was first realised by Fell and Small [99] and Watson [100], and developed by Varma and Palsson [101] which resulted in the FBA methods now widely used in structural modelling [102, 103]. The problem definition is set up by the modeller thus rendering the technique applicable to a plethora of models, and thus problems. The objective function is a value defined by the modeller which can be minimized or maximized, with respect to a set of constraints. An example of an FBA problem can be given as below:

$$\begin{aligned} &\text{minimise} \quad : \quad \textstyle\sum_{i=1}^{m} \cdot |v_j| \\ &\text{subject to} \quad \begin{cases} N_{n,m} \cdot v = 0 \\ v_{\mathrm{x_B}} = 1 \end{cases} \end{aligned} \tag{1.15}$$

where the objective function is minimization of the total flux of the reaction vector, v_j, subject to the steady state assumption which states that the reaction vector v, multiplied by the stoichiometry matrix, N containing n metabolites taking part in m reactions must equal zero, with an additional constraint on the value of export of a metabolite, $v_{\mathrm{x_B}}$ measured experimentally equal to one. In FBA, the objective function may be that of flux minimization or maximization, under any set of constraints on reactions internal or external to the system.

1.3.4 Construction and analysis of genome scale metabolic models

Thanks to the developments in the genome-sequencing techniques which allow for rapid identification of a set of genes, and therefore enzyme present, experimental data is readily available for model reconstruction. This is reflected in the high number of published models describing the metabolism of more than 6000 organisms, and applied to a variety of problems in the fields of biotechnology, human health, and others [104].

Although structural models vary in their network coverage, genome scale metabolic models (GSM) are most commonly developed, which are comprised of all the known metabolic reactions within an organisms derived from a complete genome sequence. GSMs have successfully been used to predict genomic knockout targets for the disruption of pathogenic bacterial growth [105], increase in production yields of biotechnologically useful compounds in yeast [106], and others.

Advances in the field have enabled reconstruction of metabolic models has expanded from the description of single-celled prokaryotes, to eukaryotic cells containing compartmentalised metabolism,

and finally models of higher eukaryotic organisms including models of human tissues. Reconstruction of models describing human metabolism differs to those of single-celled organisms because of the high heterogeneity in protein expression across different organs and cell types. This has led to the development of cell specific metabolic models such as the liver hepatocyte model HepatoNet1 [107] or the model of astrocyte and neuron interactions by Çakr et al. [108]. Context-specific models can be achieved by extensive literature review or a more automated method of -omics data integration into a more generic model. In 2007, two genome-scale models (GSM) comprising collective human cells metabolism were released - the Edinburgh Human metabolic network [109] and Recon1 [110]. Since then, the Recon model has undergone several revisions, with the most current version, Recon3D [111], available interactively online at https://www.vmh.life/.. The availability of these models as templates for all human metabolism has allowed for the development of numerous tailored, cell- and context- specific models by the method of integration of -omics data [112–114].
As an overview, the reconstruction of metabolic models requires the collection of a set of reactions to include, transformation of this data into a computational object and finally, development of a computational toolset for the analysis of the model.

Databases

Model reconstruction relies heavily on the use of well annotated biochemical databases which provide information about the association of genes with proteins, and the stoichiometries and directionality of the reactions they catalyze, which is lacking from gene expression data only. Two commonly used biochemical databases are BioCyc [115] and KEGG [116] with the former used in the current study. BioCyc contains extensive biochemical annotations for hundreds of reactions further subdivided into organisms-specific datasets which provides a great convenience during model annotation. Organism-specific datasets can be accessed online or downloaded as raw files from biocyc.org/download.shtlm in the *.dat* format, which can be explored in any user-specific environment. Models can be derived from the set of organism-specific reactions in the database, or else provide a good resource in the process of manual gap filling during model curation. Another resource relevant to model curation is the BRENDA [117] database, which encapsulates information about the properties of metabolic enzymes such as their substrates and products as well as related kinetic parameters, and others.
The generic human-tissue model reconstructions such as Recon [111] can be of use as a template upon which more specialised tissue and cell-specific models can be built. This requires the collection of gene and protein expression datasets which are used to constrain the network to contain those reactions that are likely present in the system under study. Since the first mapping of the human genome was completed in 2001 [118], an interest in defining normal and pathological states in human has led to an explosion of datasets including genomes, transcriptomes and proteomes. As the availability of data increased, so did the need for their organisation into single resources, making them more accessible. The Human Protein Atlas https://www.proteinatlas.org/ is a Web-based resource

which aims to map all human proteins present at the tissue and cellular level in non- and pathological states in the body, integrating data from numerous experiments. More specialized tissue resources have also been developed, including the Allen Brain Map https://portal.brain-map.org/ which contains transcriptomic datasets for mouse and human brain across different anatomical sections of the brain and at pathological states, including glioma. These resources are available freely online and are crucial in the generation of tissue-specific models, as levels of -omics expressions are informative of the likely presence of enzyme and thus, active reaction to include in the model.

Integration of transcriptomics data with metabolic models

Reconstruction of context-specific models requires a template model and an expression dataset, for instance the transcriptome or the proteome extracted from the tissue. In the first instance, the dataset content is mapped to reactions present in the model by relating genes encoding the enzymes which catalyze each reaction to the genes present in the dataset. In the next step, the expression values within the dataset are applied as constraints in an FBA problem, with the overall objective of translating a higher expression value to a higher probability of the enzyme-catalyzed reaction appearing in the solution. The statistical analysis of the data, and the method by which it is applied to the model are dependent on the type of data and the choice of the modeller, thus several algorithms have been developed so far. Some of these algorithms lead to the complete disqualification of reactions catalysed by enzymes with low expression, such as FASTCORE [119] or INIT [120] whilst others lead to minimization of flux via reactions without their complete removal, such as GIMME [121]. The choice of algorithm potentially results in great variation between the models extracted from the same dataset, which has been recently evaluated by Opdam et al. [122]. Overall, the authors found that the biggest contributor to the content of a model lies in the algorithm threshold selection which determines whether a reaction is included or not.

Transcriptomic datasets are often used for integration with GSMs since the experimental methods of RNA sequencing from tissues are widely available, and provide a wide coverage of the genes expressed in cells. Conventional cell culture sequencing methods pool the expression of multiple cells grown *in vitro* into an averaged dataset. Alternatively, the tissue can be dissected and the expression of different types of cells averaged. However, organs are composed of widely diverse populations of cells as a result of differential gene expression, which cooperate to perform the organ's function. Gene expression is heterogenous, even amongst the same or similar cell types [123]. A better understanding of the function of an individual cell can be achieved by the analysis of individual cell transcriptomes. This has been allowed by the development of single-cell RNA-sequencing (scRNA-seq) methods, pioneered by Eberwine et al. [124] who sequenced the genomes of individual neurons from the rat hippocampus. Single cell RNA-seq requires the isolation of individual cells, followed by parallel sequencing of the reverse-transcribed cDNA libraries, after which cells are sorted by some specific property. One of the methods of cell isolation is fluorescence-activated cell sorting (FACS), which is a commonly used technique that distinguishes cells by a known extracellular protein marker against which a fluorescently labelled antibody is raised. Tagged cells are later separated by flow cytometry which sorts them by the fluorescent tag attached. Once separated, mRNA is reverse transcribed at which step it may be barcoded with a unique identifier. The tagging of the genome allows for the counting of individual mRNA molecules present and relating them back to a single cell [125].

The application of scRNA-seq to metabolic models can have obvious uses in the building of cell-specific metabolic models, which could be further compounded to generate models describing the interactions of different cells. This has been so far explored in a single study by Zhang et al. [126] who used scRNA-seq to explore NAD$^+$ biosynthesis in mouse tissues.

Model curation

The initial reconstruction of a model from a genome annotation in combination with a publicly available database defines the first model draft. However, the databases often contain numerous errors which can be a result of unspecific or incomplete annotation [127]. An example of such errors is the inconsistent naming of metabolites which take part of numerous reaction, or the erroneous inclusion of reactions in the database. Since these sorts of errors often result in biologically impossible reaction sets, they can be tracked down by the modeller based on known *a priori* physical constraints. Once traced, they require manual correction from the modeller on the basis of thorough literature research, which is the model curation step. As it is relatively simple to generate a first model draft, manual curation is often the most lengthy step of the model building process but necessary for the construction of a high-quality model which does not violate known laws of nature.

Mass conservation

In order to respect the principle of mass conservation, the total atom transfer from a set of substrates to products must remain equal. These issues can be identified by the comparison of the sums of atoms present on each side of the equation, given that the chemical formulas for the metabolites taking part in the reaction are present in the database. Another way of discovering mass conservation problems developed by Gevorgyan et al. [128] utilizes mixed integer linear programming (MILP) applied to the left nullspace of the stoichiometry matrix which contains information about the conservation relationships amongst metabolites (Section 1.3.3). It allows for the identification of metabolites which cannot be assigned a positive molecular mass and thus rendering the network *stoichiometrically inconsistent.* For instance, consider the two reactions:

$$\begin{aligned} &A \rightarrow B \\ &A \rightarrow B + C \end{aligned}$$

It can be deduced from the first equation that the chemical formulas of metabolites A and B are the same, suggesting that it describes some example of isomerization. The second equation can only hold true if metabolite C does not have mass, which clearly cannot be the case biologically, meaning that one out of the two statements, or both, must be false.

Presence of such inconsistencies is often a result of incorrect database annotations. Reactions may be reported with the incorrect or no chemical formula, which is commonly a result of generically named metabolites that are not assigned a chemical formula. For instance, BioCyc contains a data entry for the compound '3-HYDROXY-ISOBUTYRATE' which represents a compound class with no chemical formula attached- the specific instance of this metabolite is found with the identifier 'CPD-12174' which needs to be replaced in the model. Another frequent cause of inconsistencies is the definition of polymers in the database which do not have a defined length, such as the generically named metabolite 'Omega-3-Fatty-Acids' which does not have a chemical formula in BioCyc.

Atomic inconsistencies are resolved manually via the replacement of metabolite for its correct counterpart, or its specific instance.

Energy and Redox conservation

The principle of the conservation of energy mean that it remains constant over time, which also applies for conservation of redox, therefore ATP or the reducing agents NAD(P) should not be formed in the model in the absence of any inputs. Since ATP and reducing cofactors take part in numerous reactions, incorrect reaction directionality often results in internal reaction cycles that spontaneously generate these metabolites. In addition to directionality issues, a major cause of these types of errors is the presence of two reactions with a similar stoichiometry but utilizing a different type of cofactor.

Linear programming can be used to identity such issues by setting a demand on the production of a metabolite in question in the absence of any inputs. It is apparent that in this case no output should be possible thus if the LP problem is feasible, it will contain problematic reactions. Offending reactions can also be spotted by the calculation of the enzyme subsets from the right kernel of the stoichiometry matrix (Section 1.3.3). Given that two reactions in a subset are carrying flux in opposite direction, no LP solution exists in which they may both carry flux, leading to an *inconsistent enzyme subset* [127]

Reaction identified by these methods are manually compared against the Brenda database and their directionality corrected. In the instance where two identical reactions utilizing different cofactors are found, that which utilizes the cofactor with a lower affinity to the enzyme in question is removed.

Computational tools

Application of mathematical analysis methods described in Section 1.3.3 is only feasible if the computation is done, unsuprisingly, by a computer. Although the first kinetic ODE models such as the Michaelis Menten model were integrated by hand, even models composed of a few reactions only are generally algebraically insoluble. Technological advances have provided a solution to this problem, and allowed for the definition of more complex models and the development of the theories described previously. Many modern computational tools extend beyond their use only as a vehicle for model construction by containing functions which greatly aid in their development, curation and simulation processes.

Given the extensive use of computational modelling, various tools have been developed for different purposes. Early computational software tools were developed to aid the simulation of biochemical kinetic models such as BIOSSIM [129] and SCAMP [130]. With the maturation of the structural modelling field came the development of software for structural modelling, including METATOOL written in 1999 [93] after which others followed. More recent software packages are written as Graphical User Interface (GUI) programs which enables users with a less developed computational background to build models. The various software bundles available differ in their characteristics such as their computational representation of the network, the database they rely on, and tools available for the automation of reconstruction which has been summarized in recent reviews [131, 132].

ScrumPy software which was used in the current study was derived from the SCAMP [130] and is a collection of Python packages which can be used for the building and analysis of both kinetic, and structural models [133]. The packages can be accessed via IDLE, Integrated Development Environment for Python and therefore requires the user to have a knowledge of the Python programming language. Although GUIs are available for the plotting of data generated from models, model analysis is mostly done in the Python language and the model itself is treated as a computational object. This was a deliberate design of the software, as it allows the user to build upon the analysis packages already available, and make changes to the packages according to the need of the model, both of which add to the flexibility of the software. As SCAMP [130] was written in the C programming language, the choice of Python for the software development was also deliberate owing to the advantages Python offers described below.

1. Python is a high level programming language, meaning that its syntax in significantly abstracted from the binary machine code fed to the computer. Some operations, for example memory assignment to data are performed without the need of user input, making it easier to code in.
2. In comparison to other programming languages such as C, its' syntax is very simple which

makes it easy to learn for a novice.

3. It is an open source language which makes it freely available to everyone.
4. There are a number of scientific packages already available in Python, such as the SciPy (Scientific Programming) package which contains algorithms with ubiquitous use in all computational modelling. Due to the widespread use of Python in the scientific community, the development and support for these packages is likely to be better than that for a new package developed by a single person with the same use.

ScrumPy models are, at the lowest level, stoichiometry matrix objects with an array of sophisticated methods attached which aid in the analysis of the matrix. Methods of analysis described in Section 1.3.3 such as the calculation of reaction control coefficients, generation of metabolic trees or calculation of enzyme subsets are readily available in ScrumPy. In addition to model analysis, the software offers a collection of packages which aid in the reconstruction of models from BioCyc files as well as the following steps of model curation. It has been successfully used for model building in several previous PhD projects as well as publications [105, 134, 135]. Both kinetic and structural model construction and analysis in the current study were done in ScrumPy. The software was developed by Mark Poolman and is downloadable from http://mudshark.brookes.ac.uk/ScrumPy.

1.3.5 Kinetic modelling of neuronal transmission

Application of mathematics to study the properties of the brain is collectively termed *computational neuroscience*, a name coined in 1988 by Sejnowski et al. [136], and emerged independently from the *systems biology* field which includes the biochemical reaction theories described in the previous sections [137]. It encompasses theoretical modelling of the brain at different scales, beginning with the microscopic scale of neurotransmitter channel activation, electrical properties of whole neurons, events happening at the scale of synapses and finally the interactions of whole networks of neurons [138].

The description of the brain as composed of individual neurons was first given by Cajal in 1886 who studied the anatomy of the central nervous system (CNS) and lay the foundations of modern neuroscience. The idea of a synapse was introduced in 1897 by Sherrington, and the description of neuronal excitability in terms of Resistor-Capacitor (RC) electrical circuits was proposed by Hermann, based on the much earlier experiments of Galvani on the effects of electricity on the contraction of frog muscles in the late 18th century [139], the birth of electrophysiology. This lead to the development of models describing the electrical excitability of neuronal membranes. The first of these models, the leaky integrate and fire (LIF) model was developed in 1907 based on experiments by Lapique, and introduced by Stein in 1965, which models a neuronal spike with equations taken from RC circuits theory in physics.

The action potential ODE model derived later by Hodkin and Huxley provided a more detailed description of the change of a neuron's membrane voltage, in terms of changes in the concentrations of Na^+ and K^+ ions. These pioneering experiments accurately described the time courses of macroscopic currents, ie. those resulting from the summed responses of multiple channels in a neuronal cell.

Neher and Sakmann developed techniques to record microscopic currents through single channels, thus proving that the macroscopic currents are a result of summed responses of many individual protein channels in operation. The translation of the single-channel currents into mechanistic models describing the activation of a single channel by its ligand was theoretically described by Colquhoun and Hawkes [140]. Based on these theories, the information contained within the single-channel data is used to predict the series and speed of conformational changes that a protein undergoes upon ligand binding, in a similar way that chemical reactions are described. A system of differential equations is proposed, to which the single-channel current data is fitted- that which fits the data best is chosen as the most likely mode of channel activation.

The detailed descriptions of single channels have been integrated into larger kinetic models which may describe the interaction of different channels, changes in the concentration of ions, and others [141].

Modelling of electrogenic proteins

Elucidation of mechanisms behind electrogenic protein activation by a ligand usually involves the collection of currents via the protein under different conditions, such as varying ligand and ionic concentration, membrane voltages, and others. It is now known that ligand activation induces conformational changes in the protein which was confirmed at the molecular level by direct imaging via X-ray crystallography studies that capture single molecules in different states. The first of these structures was published by Doyle et al. [142] and described the structure of bacterial Na^+ channel. Electrophysiological experiments aiming to predict these structural changes involve the collection of data, and subsequently the fitting of a multiple state kinetic model which is capable to describe it well. Multiple schemes are proposed, and optimization algorithms are used to calculate which scheme can describe the data best, and is the most mechanistically probable based on previous information.

As previously mentioned, electrophysiological techniques based themselves in the description of a neuron and the proteins within via RC- circuit equations formalism. The aqueous solutions inside and outside of the neuron are filled with charged particles, thus making them good conductors in comparison to the bilipid membrane, which does not allow the movement of ions and thus can be described as a capacitor, C. Electrical potential, V, across the membrane is a result the difference in the concentration of ions, thus the charge, Q, on either side of the membrane, which is set up by active transport, and the equation for membrane capacitance can be written:

$$C = Q/V \tag{1.16}$$

If the voltage across the capacitor changes, a current will flow which is given by:

$$I = C\frac{dV_m}{dt} \tag{1.17}$$

The current is a result of the movement of ions via receptor pores, due to the difference in the electrochemical potential. The Nernst equation describes the reversal potential for the ion, which is an individual property of each ion, and is given by the equation:

$$E_{rev} = \frac{RT}{zF}ln\frac{c_{out}}{c_{in}} \tag{1.18}$$

where R is the is the universal gas constant, T is the temperature in Kelvin, z is the valence of the ion, F is the Faraday constant, and c are the concentrations of ions outside and inside the cell. Voltage drop across the membrane, V_m, is a result of the voltage drop across the resistor, R_i as well as the voltage drop across the 'battery', given by the Nernst potential for the ion, E_{rev}, which can be written as:

$$V_m = I_i R_i + E_{rev} \tag{1.19}$$

and since the conductance, g, is the reciprocal of resistance, R_i, the equation can be rearranged to give:

$$I_i = g_i(V_m - E_{rev}) \tag{1.20}$$

Since the current is the rate of charge flow, and the voltage is proportional to the current multiplied by the capacity the equation becomes:

$$C\frac{dV}{dt} = g_i(V_m - E_{rev}) \tag{1.21}$$

which is the ordinary differential equation for the membrane voltage change across a circuit. In electrophysiological recordings, parameters such as the membrane voltage, concentration of ions defining the strength of the 'battery', as well ligand concentrations can be manipulated by the experimenter in order to calculate the currents via channels, and others.

Electrophysiology techniques and TEVC

Recording techniques in electrophysiology all revolve around the manipulation of the membrane voltage of the system in question, called a voltage clamp. This is done both in the collection of macroscopic currents from a whole cell, meaning those generated as a summed response of individual currents via a number of channels, or microscopic currents via single-channels in the patch clamp techniques developed by Neher and Sakmann. Two-electrode voltage clamp (TEVC) is a technique developed by Hodgkin and Huxley for the collection of macroscopic currents from whole cells. It can be used in neurons dissected from the organism which allows one to obtain data on the currents from native channels expressed in the neuron. Since neurons typically express more than one receptor protein, pharmacological blocks are often used in order to single out currents from the receptor in question. In these experiments the level of expression of the ion channels cannot be manipulated by the experimenter thus the currents are often low in amplitude. In addition, the small size of cells makes the experimental set up difficult. Genetic manipulation of mammalian oocytes excised from the African clawed frog *Xenopus laevis* to express heterologous proteins of interest provided a solution to these issues and since has become a widely used technique in electrophysiological recordings. Foreign cDNA or mRNA prepared *in vitro* is injected into the oocyte which translates the genetic material into proteins which express in the cell membrane. The large size of the oocyte - around 1mm diameter - makes oocyte handling more easy, including molecular material injections as well as collection of electrophysiological data. In addition, the oocyte does not depend on substrates taken up from the environment, expression of endogenous transporter proteins is low, meaning recordings can be taken in low background noise [143].

In TEVC, the oocyte expressing the protein of interest is placed in an aqueous solution of known salt concentrations inside a recording chamber. The intracellular concentration of ions and charged proteins *versus* the concentration of ions in the bathing solution generates a voltage across the lipid bilayer of the oocyte. Two electrodes are used to impale the cell, one of which injects current and the other measures the membrane voltage, V_m. This allows the experimenter to aritificially control the membrane potential via setting of the V_m0.1cm to a chosen command potential, V_c, both of which are connected to a clamping amplifier. The ionic flux that is carried via the protein of interest upon simulation - for instance application of a ligand - displaces the V_m0.1cm which is compared to the V_c applied by the experimenter. In order to counteract the deviation, the output voltage from the amplifier forces the current electrode to pass a current directly proportional and opposite in polarity to the ionic flux via the protein. It is this current which is measured, which changes as a function of the macroscopic currents via the recombinant protein. A schematic diagram of the TEVC setup can be seen in Figure 1.7.

Given the previously mentioned advatantage of TEVC recordings in *X.* oocytes, this system has been widely applied in the study of the properties of electrogenic receptor and transporter proteins.

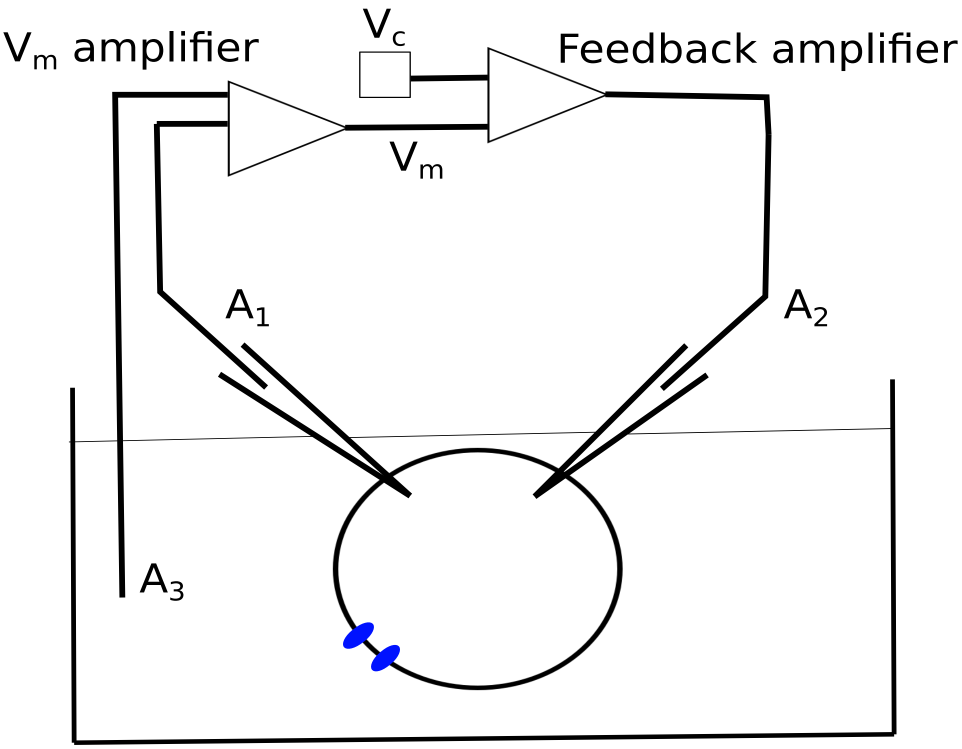


Figure 1.7: **TEVC recording arrangement** Oocyte is placed in a chamber which contains an ionic solution. The voltage-measuring amplifier receives input from the measuring electrode (A_1) and the bath electrode (A_3) and determines the voltage which approximates V_m. The membrane voltage is sent to the feedback amplifier which compares it to the command voltage (V_c) as set manually in the TEVC instrument. The deviation of V_m from V_c is corrected by the feedback amplifier connected to the current-injecting electrode (A_2). The feedback current is the measure equal and opposite in polarity to the ionic currents through protein upon agonist application.

1.4 Aims and thesis structure

The various aspects of glycine metabolism and function in the CNS are in need of further clarification, and are adressed in the current thesis. Much of what is known about the initial pathways of glycine biosynthesis *specific to the brain* comes from studies performed in rats in the 1970s, and has not been revisited in the light of new experimental findings. For instance, it is now widely accepted that brain metabolism is compartmentalized across different cell types, as is the case in the cycling of the major neurotransmitter, glutamate. Expression of genes encoding the proteins involved in glycine metabolism in monocultures of astrocytes and neurons suggests the compartmentalization of glycine biosynthesis, but has not been conclusively proven. Furthermore, glycine metabolism *in vivo* takes place in the background of large demands for ATP required for neurotransmission which should be taken into account in a holistic approach to the study of metabolism. Therefore, one of the major aims of this project is to discover the metabolic pathways of glycine biosynthesis in the context of the background CNS metabolism via the use of computational modelling techniques, and integration of scRNA-seq dataset which allows for the study of glycine across different cell types.

As a metabolic intermediate of cellular biosynthetic pathways, evidence points to the role of serine/glycine in the support of cancer cell proliferation, with a potential of developing novel treatments that disrupt these pathways. However, questions such as whether these two amino acids are equally capable of supporting proliferation, and the possible metabolic pathways utilized are underdetermined which is also addresses in the current work via the analysis of a metabolic model representing the CNS.

With respect to glycine's role as a neurotransmitter, it has been proposed that regulation of its concentration via the manipulation of its reuptake via GlyT1 in different brain regions can be used as a treatment for numerous conditions. However, the most advanced clinical trials of GlyT1 inhibitor bitopertin did not result in a positive outcome, and studies of alternative inhibitor compounds show contrasting result. Increasing the understanding of the interaction between the glycinergic synapse components could therefore be of use in the future development of neuromodulators. The direct effect of GlyT1 inhibition on the glycinergic currents is difficult to measure in the organism *in vivo*, and computational methods can aid in the understanding of the mechanism at the molecular level. Therefore, the second major aim of this work was the development and analysis of a computational model which integrates previous kinetic studies of single proteins which compose the glycinergic synapse for the purpose of better understanding how glycine is modulated and its effect on the glycinergic Cl^- currents via the GlyR.

Lastly, a point mutation at position Ser 407 in the GlyT1 was identified by Alfadhel,M. et al. [80] as a novel cause for NKH, but the effect of the mutation on currents via the GlyT1 has not yet been characterised. Another major aim of the work presented in this thesis was to perform electrophysiological experiments on the Ser 407 GlyT1 mutant in order to functionally characterise the effect of

this mutation, and observe its effects on glycinergic currents by integrating the experimental results into the kinetic model.
In summary, the main objectives of this thesis are:

1. To use metabolic modelling in order to better understand pathways of glycine biosynthesis in different cell types of the CNS, in the context of brain energy demands.
2. To use metabolic modelling to discover the involvement of serine/glycine metabolic pathways in increasing cellular ATP and biosynthetic needs.
3. To use kinetic modelling in order to better understand the interaction of proteins involved in glycine release and reuptake across the glycinergic synapse, and how the time course of glycine concentration increases affect the Cl^- currents via the GlyR.
4. To use electrophysiological TEVC techniques in order to characterise the effect of the Ser 407 GlyT1 mutation on the resulting currents via the transporter.
5. To integrate the results from the experimental studies with the kinetic model of the glycinergic synapse for the purpose of relating the effect of the mutation on the resulting Cl^- via the GlyR.

The structure of the thesis is as follows:

Chapter 1 contains an introduction to the concepts relevant to the biology of glycine metabolism in the CNS, as well as an overview of the computational and experimental methods used in the current thesis.

Chapter 2 describes the construction of a metabolic model of the CNS, and the analysis of the model for the estimation of energy and biosynthetic needs which are used in later sections for the discovery of glycine biosynthetic pathways.

Chapter 3 describes the integration of a scRNA-seq dataset with the CNS model, and analysis of the pathways of energy as well as glycine production used by the different CNS cell types. The ability of serine and glycine to support biomass and energy production is also described in this chapter.

Chapter 4 describes the construction and analysis of a kinetic model of a glycinergic synapse. A detailed scheme of GlyR activation by glycine taken from Burzomato [144] is analysed, and built upon by the addition of reactions describing glycine release and reuptake by GlyT1 and GlyT2. The full model is used to look at the effects of changing components of the glycine synapse on the resulting Cl- currents, including the result of the GlyT1 Ser 407 mutation which abolished currents via the GlyT1, as described in Chapter 5.

Chapter 5 describes the experimental results from electrophysiological study of the effect of Ser 407 mutation in the GlyT1 on the rate of glycine transport.

Chapter 6 contains an overview of the main results presented in the thesis, as well as an overall discussion of the results.

Chapter 2

Metabolism of CNS glycine in the context of energy and biomass requirements

2.1 Introduction

This chapter describes the developement and curation of a CNS metabolic model in the ScrumPy modelling software. The model was derived from a pre-existing GSM model of human metabolism, Recon2.2 which encompasses reactions present in all human cells and tissues [145]. Previous models pertinent to CNS metabolism have been published, including the astrocyte and neuron model by Çakr et al. [108], as well as a model of the human astrocyte network by Martín-Jiménez et al. [146]. Both of these models are cell-specific, thus the development of the CNS model was motivated by a need for a more generic reconstruction in the first instance. Further contextualisation of the generic reconstruction is applied in Chapter 3, where metabolism in different cell types is investigated via the integration of transcriptomic data.

The purpose of the model reconstruction was to probe questions relating to glycine metabolism in the CNS, subject to the constraints imposed by substrate availability, as well as the metabolic context within which biosynthesis occurs. The first is determined by the rate of metabolite passage across the blood brain barrier (BBB), thus second to model curation, constraints on model import were added according to literature values. In the next step, uptake constraints were used to estimate energy and biomass turnover rates within the brain. The oxygen to glucose consumption index (OGI) is a previously well defined parameter and is between 5.5- 6.0 in human brain (Introduction, Section 1.2.3). This value range was used to estimate the metabolic state of the cell, and variations in the energy and biosynthetic demands were assessed via their effect on the OGI. Estimation of these values, as well as the uptake rates of metabolites were used later to provide a meaningful cellular context in which to evaluate glycine biosynthesis. This was achieved by the comparison of pathways utilized in the absence of the biosynthetic and energy demands, and later on in their presence. Pathways utilized within each scenario were compared to elucidate the most likely routes of glycine production in the background of cellular metabolism.

2.2 Methodology

All steps of model construction and analysis were completed ScrumPy (Chapter 2, Section 1.3.4). The model was derived from the Recon2.2 GSM [145] under the assumption that it has undergone more extensive curation in comparison to the MetaCyc database of human metabolism. The Recon2.2 model was downloaded from http://identifiers.org/biomodels.db in the systems biology markup language (SBML) format and converted into a ScrumPy model via identification of equivalent MetaCyc species.

Model metabolite names were translated in the first instance, followed by reaction names. Metabolites were matched to MetaCyc counterparts by their unique identifiers (ChEBI, InChi, SMILES) found in the annotation files or via the application of the Python Levenshtein string matching algorithm. Around 43% of the Recon2.2 reactions represented unique, metabolic reactions which were subject to translation. In the first instance an annotation of the Recon2.2 was downloaded from http://metanetx.org which contained MetaCyc equivalents for some of the reactions in the model. Additional reactions were matched via their enzyme EC numbers and gene associations, and the remaining reactions were matched by their metabolite stoichiometry. In result, the model Recon2.2 derivation covered 41% of metabolites and 63.6% of reactions present in the original model.

2.2.1 Model curation

Derivation of a reaction set allowed for the generation of a first model draft which required further revision in order to ensure it becomes a faithful representation of CNS metabolism. More specifically, that it contains a set of reactions *relevant* to the system in question and thus reactions known to be inactive in the brain were manually removed. Enzyme ‘inactivity’ was determined as zero mean expression of enzyme RNA across various regions of the human cortex, as identified from the Allen Brain Map. Secondly, the model must obey the laws of nature known *a priori* (Introduction, Section 1.3.4) thus inconsistencies were corrected semi-manually and are described in the following passages.

Mass conservation

Reactions which violated mass conservation were identified via comparison of the sums of molecular masses on each side of the equation, as well via the analysis of the left null space of the stoichiometry matrix (Introduction, Section 1.3.3). Atomic inconsistencies resulting from C, N, S, O, P imbalances were resolved manually via the replacement of metabolite for its correct counterpart, or its specific instance. Since H imbalances are mostly caused by the incorrect protonation state of the metabolite which may be missing from the database, they were resolved automatically via the addition of a H atom on either side of the equation.

Energy and Redox conservation

Flux Balance Analysis was used in order to identify reactions which lead to spontaneous energy formation. In the LP formulation, a positive value was set as the flux constraint on a hypothetical ATPase reaction which catalyses the breakdown of ATP into ADP and P_i, in the presence of zero constraint values on all the import reactions. A similar approach was used to ensure redox conservation, where flux via a NAD(P) oxidation reaction was chosen as the objective function. If the problem was feasible, the flux distributions returned contained offending reactions which required curation. In a different approach, offending reactions were identified by the analysis of the right null space of the stoichiometry matrix (Chapter 2, Section 1.3.4)

Reaction identified by these methods were manually compared against the Brenda database and directionality corrected. A second major cause of imbalance was the presence of reactions with very similar stoichiometries, differing only by the cofactors involved. This was solved via removal of the reaction which utilized the cofactor with a lower affinity to the enzyme in question, as defined in Brenda.

2.2.2 General model properties

The final version of the curated model is composed of a top level file which includes 11 subfiles (Figure 2.1) that represent :

- six internal compartments of a cell, including:
    - Cytoplasm (907 reactions)
    - Mitochondrium (395 reactions)
    - Peroxisome (156 reactions)
    - Endoplasmic Reticulum (125 reactions)
    - Nucleus (66 reactions)
    - Lysosome (22 reactions)

- three metabolite transport reaction files split into extracellular transporters (44 reactions), intercompartmental transporters (402 reactions) and biomass exporters (36 reactions)
- a representation of the electron transport chain (ETC)
- internal ‘extra reactions’ file added during model curation

Out of the total of 2193 model reactions, 674 are ‘dead’ reactions which cannot carry flux at steady state.

Internal compartments

Assignment of reactions to compartments was kept consistent with Recon2.2. However, three compartments were removed in the curation process. The golgi apparatus was removed due to a low number of identified reactions (7) which were moved to the cytoplasm. The ‘boundary’ compartment, which in the original model represents extracellular fluids did not contain any metabolic reactions thus was also removed. Lastly, the mitochondrial intermembrane space compartment contained five reactions describing the accumulation of H^+ in the intermembrane space, ATP synthase reaction, an aspartate-glutamate shuttle, and the heat-producing uncoupling protein catalyzed reaction. This compartment was removed and the ETC added as a separate module.

Transport reactions

The movement of metabolites in the system was split into three logical components, more specifically;

- **Extracellular uptake and release**

  Import and export of metabolites in the model was added manually based on studies of metabolite transport across BBB membranes. Substrates are composed of glucose, amino acids, ketone bodies and fatty acids known to be essential for the compositions of membrane lipids. Pantothenate used for the formation of vitamin B, and choline used in the biosynthesis of phosphatidylcholine are obtained from diet [147] thus also comprise cellular import. Lastly, transport of inorganic phosphate ($\mathrm{P_i}$, $\mathrm{P}P_\mathrm{i}$) and ions was added. A list of all transport reactions can be seen in Table 2.1, including both uptaken and released metabolites.

- **Biomass exporters**

  Most structural models of metabolism utilize cellular growth as a metabolic objective. As this condition applies to model describing bacterial cultures or plant metabolism, it does not apply to mature, differentiated neuronal cells which do not divide [148]. Nevertheless, under steady state these cells require the replacement of their contents including metabolic enzymes, phosholipids and neurotransmitter receptors. This demand on cell metabolism was added in the form of a set of export reactions for micromolecules making up cellular components. The latest revision of the human GSM, Recon3D [111] contains a biomass turnover equation which was used in the CNS model. The complete list of biomass components can be seen in Table 2.1. Manual gap filling was performed to ensure that the model is capable of exporting each compound from uptake metabolites, and a separate model file was designated that describes their export. Apart from formation, turnover processes also include degradation of cell components however this was not taken into account in the CNS model.

- **Inter-compartmental exchange**

  This module contains reactions which transport metabolites between the internal compartments of the cell, as originally defined in the Recon2.2 model.

Electron Transport Chain

Production of ATP via oxidative phosphorylation in the mitochondria is driven by the electrochemical gradients established by H^+ accumulation in the mitochondrial intermembrane space. H^+ accumulation is achieved by mitochondrial complexes complexes I ,III and IV which take part in redox reactions that harness the energy produced by the movement of electrons for H^+ translocation. The ETC model file is a representation of this process. It includes reactions catalyzed by complexes I - IV, with the net transfer of 10 mitochondrial H^+ to a H^+_{int} species, representing H^+ located in the mitochondrial intermembrane space. The last reaction is a representation of the ATP synthase reaction, fixed in the direction of ATP formation where one ATP molecule requires the transport of four H^+_{int} back into the matrix, of the form;

$$1.0\ \mathrm{ADP} + 4.0\ \ \mathrm{H}^+_{int} + 1.0\ \mathrm{P_i} \rightarrow 1.0\ ATP + 3.0\ \mathrm{H}^+ + 1.0\ \mathrm{H_2O}$$

The resulting model P/O ratio is 2.5, assuming two electrons used in the process enter via complex I and pass via all three complexes I,II, III and IV, which is in line with previously reported

Table 2.1: Complete set of model transport reactions

Class	Import	Export	Biomass
General	Glucose	Lactate	
	Oxygen	CO_2	
	Water	NH_3	
	Proton		
	Coenzyme A		
	Pantothenate		
	Choline		
	Na^+ , Cl^-		
	P_i, PP_i		
Amino acids	Valine	Glycine	Valine
	Leucine	Serine	Leucine
	Isoleucine	Proline	Isoleucine
	Methionine	Threonine	Methionine
	Phenylalanine	Asparagine	Phenylalanine
	Tryptophan	Glutamate	Tryptophan
	Tyrosine	Glutamine	Tyrosine
	Lysine	Aspartate	Lysine
	Histidine		Histidine
	Arginine		Arginine
	Ornithine		Asparagine
	Cysteine		Proline
			Glutamine
			Glutamate
			Alanine
			Threonine
			Aspartate
			Serine
			Cysteine
			Glycine
Fatty acids	Palmitate		
	Linoleate		
	Linolenic acid		
	Docosahexanoate		
	β-Hydroxybutyrate		
	Ketobutyrate		
Phospholipids, Lipids			Phosphatidylserine
			Phosphatidyl-myoinositol
			Phosphatidylethanolamine
			Phosphatidylglycerol
			Phosphatidylcholine
			Phosphatidylserine
			Sphingomyelin
			Cardiolipin
			Cholesterol
Nucleotides			DGTP
			DCTP
			DATP
			GTP
			TTP
			UTP
			CTP

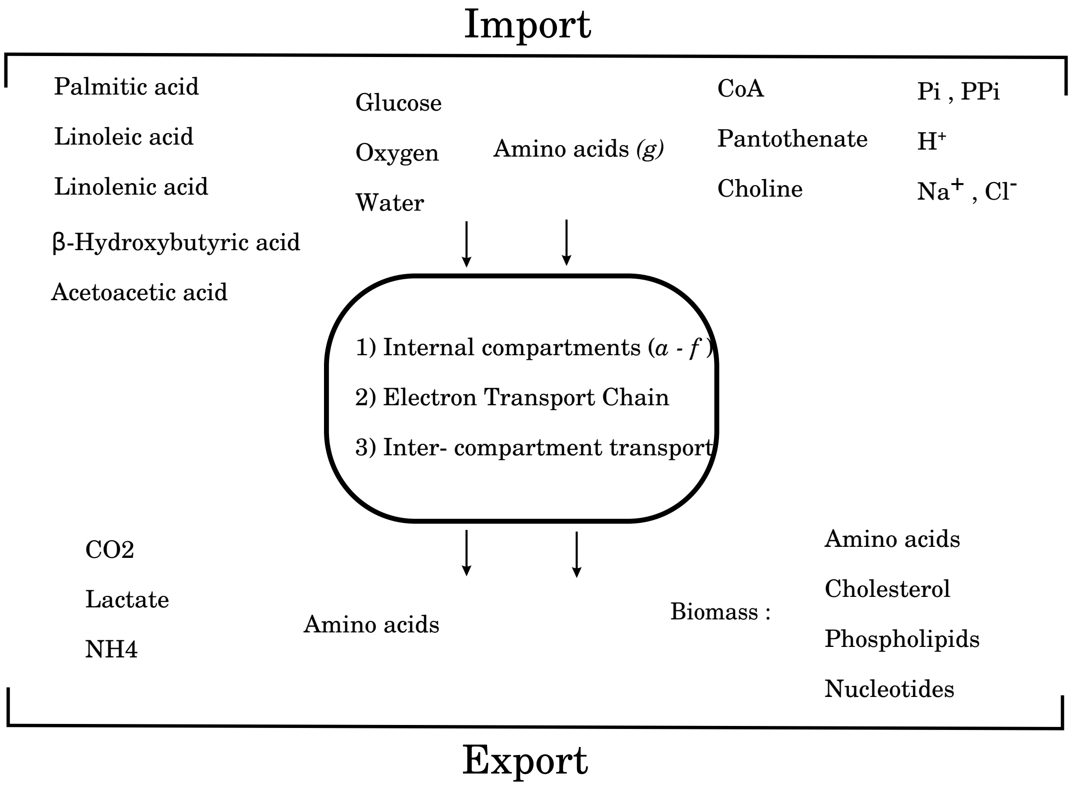


Figure 2.1: **Composition of the final model.** Internal compartments include (a) cytoplasm, (b) mitochondria, (c) peroxisome, (d) endoplasmic reticulum ,(e) nucleus, (f) lysosome, as well as the ETC and the inter- compartmental transport reactions. Cell substrates can be divided into several subclasses, C and N sources (glucose, amino acids), long and short chain fatty acids, metabolites obtained from diet and small molecules and ions. Export of metabolites includes the general by-products lactate and CO_2, amino acids and biomass components. Biomass export can be further subdivided into amino acids, phospholipids, nucleotides and cholesterol. All transport reactions are summarised in Table 2.1

values in human [149].

ATP production via the ETC relies on the availability of reducing agents in the mitochondria, mostly in the form of NADH produced in glycolysis via the oxidation of glyceraldehyde-3-phosphate. Since the mitochondrial membrane is impermeable to NAD it its oxidised and reduced form, electrons from NADH are carried across the membrane mostly via two mechanisms, the malate-aspartate shuttle and the glycerol-phosphate shuttle with the former constituting the primary mechanisms for NADH generation in the brain [150]. Apart from their role in providing electrons for the ETC, these shuttles are important for the regeneration of NAD required for the operation of glycolysis. Increasing evidence also points to the importance of lactate-pyruvate cycling in the regulation of cytosolic NAD levels via the enzyme lactate dehydrogenase (LDH) [151] which interconverts these monocarboxylates in the following reaction:

$$1.0\ \text{Pyruvate} + 1.0\ \ \text{NADH} + 1.0\ \ \text{H}^{+} \leftrightarrow 1.0\ \text{Lactate} + 1.0\ \text{NAD}$$

Lactate dehydrogenase is present both in the cytosol as well as in the mitochondrial intermembrane space (mLDH) and the direction of the reaction is dependant on the specific isoform of the enzyme as well as the concentrations of each compound [151]. Cycling of these compounds is also dependant on the presence of monocarboxylate transporters (MCT) in the intermembrane space which may co-transport each metabolite with a proton, or act as exchangers where the uptake of one of the metabolites is accompanied by the release of the other on the other side of the membrane [152].

The above named processes, including the malate-aspartate shuttle, the glycerol phosphate shuttle, as well as the LDH, mLDH and MCT-catalysed reactions were included in the model. Since the model is a simplification of the biological entity, reactions taking place in the intermembrane space such as the mLDH were included in mitochondrial compartment.

The last file, 'Extra Reactions' contains internal metabolic reactions that were manually added during model curation such as those used to gap-fill pathways of biomass components production.

2.3 Results

Formulation of LP problems

The following results sections describe the output of FBA-type problems which utilizes LP to impose a set of constraints on the flux distributions of model reactions. Across all solutions, the objective function was minimization of the sum of all reactions fluxes, subject to the steady state condition. Additional constraints were added according to the problem, including a specific cellular objective subject to constraints on the upper limits of fluxes in the transport reactions and/or a fixed flux value of biomass export reactions. The LP can be expressed as follows:

$$\text{minimise} \quad : \quad \sum_{i=1}^{m} |v| \cdot j_i$$

$$\text{subject to} \quad \begin{cases} N_{n,m} \cdot v = 0 \\ v_{bj} = b_j \\ 0 \leq v_{tj} \leq t_j \\ v_{\text{ATPase}} = 1 \end{cases} \tag{2.1}$$

where $|v|$ represents the vector of all reactions fluxes weighed by an objective coefficient vector, j_i , values of which are all set to 1 in the current chapter. Steady state condition is imposed where the stoichiomentry matrix, N of n metabolites and m reactions multiplied by the reaction flux vector $|v|$ must equal zero. The vector of biomass transporter fluxes, v_{bj} is constrained to single values for each component in a flux vector, b_j. Vector of transport reaction fluxes, v_{tj}, is constrained only by the upper limit of flux for each reaction defined in vector t_j. In this LP formulation, the specific objective function is breakdown of ATP via the ATPase reaction which is applied by setting the flux via the ATPase, v_{ATPase}, to a value of one. Apart from the steady state condition, constraints are varied according to the problem in order to probe the specific metabolic objective in question.

Metabolite uptake constraints

The rate of uptake of import metabolites was constrained to physiologically relevant values, as summarised in Table 2.2. Since the concentrations of compounds in the blood vary according to the dietary intake and physiological state, the constraints were formulated such that the upper limit is constrained only, as in Equation 3.1. This should result in physiologically relevant solutions, without over-constriction of the model to produce the desired result. Amino acids include all of the import amino acids, as specified in Table 2.1.

Table 2.2: Upper limits on the rate of metabolite uptake in the model

Metabolite	Uptake limit(μmolg^{-1}min^{-1})	Reference
Glucose	0.6	[154–158]
Amino acids	0.01	[159–163]
Palmitate	0.0015	[164]
Linolenic acid	0.0001	[165]
Linoleic acid	0.0001	[165]
Decosahexanoate	0.0001	[164]

Although intracellular levels of brain glutamate is high, passage of this amino acid across the BBB is highly restricted thus majority of its synthesis takes place inside the organ from anaplerosis of glutamine as well as the TCA intermediate, alpha-keto-glutarate [153]. There is a net uptake of amino acids into the brain which serve as amino donors for these reaction in the first instance. Since in this chapter metabolism of the brain as an organ is considered, substrates available are constrained to those amino acids which pass the BBB, excluding glutamine and glutamate. In Chapter 3, where metabolism of different CNS cell types is assessed, large intracellular concentrations of glutamine and glutamate are taken into account and these amino acids are allowed as uptake metabolites during model analysis.

Biomass export constraints

The sum of fluxes for biomass exporter reactions in Recon3D representing cell maintenance without the production of replication precursors is 113.7 μmolg^{-1}min^{-1} . It can be seen that there is a large discrepancy between the magnitude of the total biomass output and the biological upper limits in the uptake of glucose and nitrogen sources as in Table 2.2. The origin of the absolute flux values provided in Recon3D is not well documented in the original publication thus the reason for such a discrepancy is unknown. Since the total biomass output was infeasible according to the uptake rates gathered from literature, it was adjusted in the current model. In addition, the ratio of compounds in the original equation were changed to better represent CNS cell metabolism.

The flux value for each individual biomass transporter as in Recon3D represents a percentage of the total biomass output, $\sum b_j$. Similarly, components of the biomass transport can be grouped into classes such as ‘amino acids’, which can also be represented as a percentage of $\sum b_j$. Since neuronal cells have a high lipid composition, the percentage of ‘lipids’ class, including phospholipids and lipids (Table 2.1) was adjusted to 24% [166] of $\sum b_j$ with the remaining percent demand on the other compound classes adjusted accordingly. Within each class, the percent of a specific compound out of the class total was kept the same as in the original publication. For example, if cholesterol composes 7% of the total lipids class then although the flux demand on its export increases by adjustment of lipids to 24% of biomass sum, its percent composition of lipids class remains the same at 7%. In each LP problem, the ratio of each biomass component was used to calculate the b_j from the specified $\sum b_j$.

2.3.1 Pathways of ATP production in the model

The brain has large requirements for ATP synthesis with the majority of energy used to re-establish ionic concentration gradients necessary for signal transmission (Introduction, Section 1.2.3). This requirement will be referred to as the ATP demand (ATP_d) in the following sections and is abstracted in the model by setting a constraint on the rate of ATP hydrolysis by the ATPase-catalyzed reaction, v_{ATPase}, of the form:

$$1.0\ \text{ATP} + 1.0\ \ \text{H}_2\text{O} \rightarrow 1.0\ ADP + 1.0\ \text{H}^+ + 1.0\ \text{P}_\text{i}$$

In order to establish the pathways of energy production from glucose, as well as to estimate the upper limit on ATP synthesis in the presence of physiological uptake constraints (Table 2.2) and constraints on the uptake of ketone bodies 3-hydroxybutyrate (HB) and 3-ketobutyrate (KB), an increasing demand on the ATPase reaction was formulated in an LP :

$$\begin{array}{lll} \text{minimise} & : & \sum_{i=1}^{m} |v| \cdot j_i \\ \text{subject to} & & \begin{cases} N_{n,m} \cdot v = 0 \\ 0 \leq v_{tj} \leq t_j \\ v_{\text{ATPase}} = a_j \\ v_{\text{KB}_{\text{tx}}} = 0 \\ v_{\text{HB}_{\text{tx}}} = 0 \end{cases} \end{array}$$

(2.2)

In each FBA problem the changing variable was the a_j, which was incremented from 0.2 to 21 μmolg^{-1}min^{-1} . The maximal feasible rate of ATP production in the model was 16.7 μmolg^{-1}min^{-1} , shortly after glucose uptake saturation at 15.6 μmolg^{-1}min^{-1} . Within the range of ATP_d = 0.2-16.7 μmolg^{-1}min^{-1} the solutions remained the same. There is a consumption of glucose which is processed via the pathways of glycolysis and subsequently complete oxidation via the TCA cycle and oxidative phosphorylation in the mitochondria (Figure 2.2). The OGI index remains at six across all solutions up to glucose saturation. Pyruvate formed in the cytosol is converted to lactate via lactate dehydrogenase (LDH), oxidizing NADH to NAD. Subsequently, lactate is co-transported with a H^+ via the mitochondrial membrane where it is oxidized back to pyruvate by the reverse reaction of mitochondrial LDH (mLDH), releasing a H^+ and reducing NAD to NADH. No lactate export is present, and the uptake of oxygen is unlimited thus lactate acts an intermediate of oxidative phosphorylation under normal oxygen tension.

ATP demand increase in the background of biomass turnover

The OGI index provides a good overview of the degree of coupling between glucose and O_2 consumption and therefore the fate of glucose. In addition, the values for the human brain OGI have been previously well established to range from 5-6 under 'normal' physiological conditions [17]. Therefore in the next set of analysis the effect of ranging demands for biomass turnover and energy production on the resulting OGI index are measured. It is anticipated that an increase in ATP_d on its own will increase the coupling between O_2 and glucose consumption and therefore OGI, as an increasingly higher proportion of ATP is produced via mitochondrial oxidative phosphorylation. On the other hand, an increase in biomass demand is expected to lower OGI because glucose is being used in biosynthetic reactions, such as the NADPH-generating PPP and reactions diverting glycolytic intermediates to amino acids and lipid production.

In order to confirm these predictions, OGI was determined as a function of changing ATP_d at a bm_d of 0.09 $\mu molg^{-1}min^{-1}$. The demand for increased ATP_d was formulated in a LP as in Equation 3.1 with the added constraint on the sum of biomass export. Upon each sequential LP formulation the value of the v_{ATPase} was incremented from 0.2- 21 $\mu molg^{-1}min^{-1}$.

As the ATP_d increases between 0.2-16.6 $\mu molg^{-1}min^{-1}$, there is an hyperbolic rise in the OGI from 0.8 to 5.4 which is coupled to increased glucose and oxygen uptake. Conversion of cytoplasmic pyruvate to lactate and its subsequent transport across the mitochondrial membrane and conversion back to pyruvate remains as in the previous scan with no additional bm_d (Section 2.3.1). Reactions of lactate to pyruvate conversions, as well as those composing the TCA cycle and ETC components increase linearly across the scan until $ATP_d = 16.6$ $\mu molg^{-1}min^{-1}$ at which they hit a plateau. Although there is a stable linear increase in the TCA cycle reactions, there is a change in the diversion of glucose which at lower biomass demands is shunted to the PPP pathways. From 0.2- 6.7 $\mu molg^{-1}min^{-1}$ all of the glucose imported is phosphorylated to glucose-6-phosphate, which subsequently enters the oxidative PPP via conversion to 6-phospho D-glucono-1,5-lactone by glucose-6-phosphate dehydrogenase. Until $bm_d = 6.7$ $\mu molg^{-1}min^{-1}$, the second step of glycolysis is active in the direction of fructose-6-phosphate to glucose-6-phosphate conversion. Reactions of the non-oxidative PPP catalysed by transketolase and transaldolase enzymes are supplying fructose-

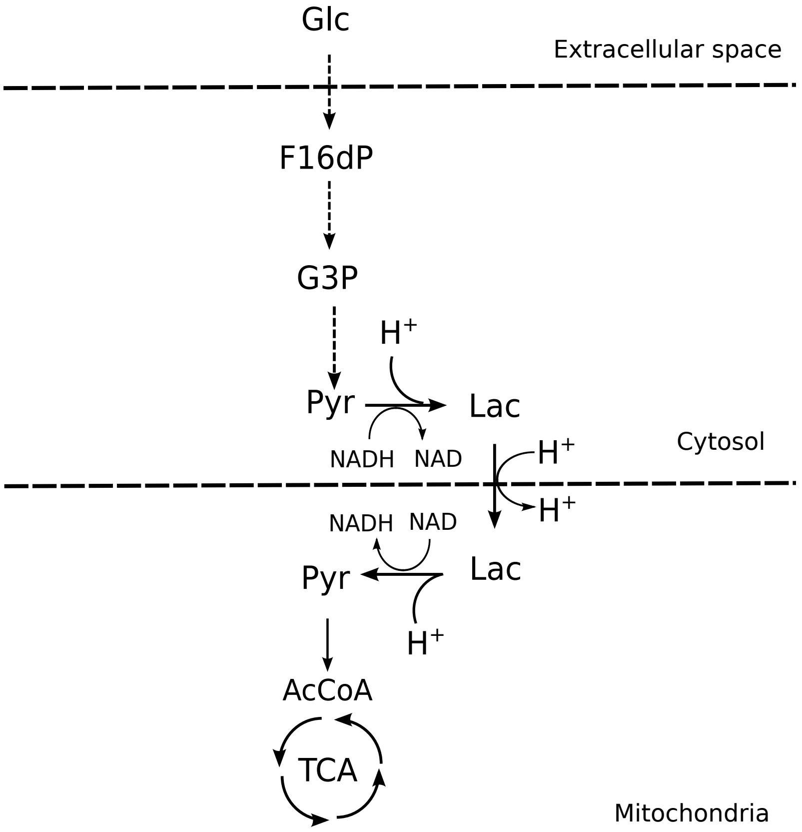


Figure 2.2: **Pathway of ATP production in the model** Glycolytic product pyruvate is used to make lactate under normal oxygen tension. Next, lactate is co-transported with a H^+ and re-converted back to pyruvate via mLDH. Dashed arrows represent a condensed pathway, where intermediate reaction are not shown.

6-phosphate and glyceraldehyde-3-phosphate to the subsequent pathways of glycolysis thus allowing for linear increase in the downstream glycolytic pathways feeding the TCA cycle.

At all biomass demands there is an uptake of ketobutyrate which is converted to acetoacetyl-CoA in the cytoplasm, further converted to acetyl-CoA. Acetyl-CoA is next used to feed the TCA cycle via conversion to citrate in the mitochondria, or for the formation of malonyl-CoA used in lipid biosynthesis. The contribution of ketobutyrate to the total acetyl-CoA pool decreases in comparison to its derivation from pyruvate, reflecting the relatively stable uptake (range 0.14-0.15 μmolg^{-1}min^{-1}) of ketobutyrate and large increase in the rate of glucose uptake.

At the point of the metabolic switch, ATP$_d$ = 16.6 μmolg^{-1}min^{-1} , glucose uptake is saturated and 3-hydroxybutyrate uptake switches on. Reactions present in the subsequent solutions but absent from the previous are the pathways of hydroxybutyrate breakdown into acetoacetyl-CoA, and subsequently acetyl-coA in the mitochondria. Thus, the metabolic switch occurs when the cell has run out of glucose and begins to use ketone bodies for acetyl-CoA production, explaining the rapid rise in oxygen consumption. This is reflected in the rate of OGI increase which switched from exponential to linear rise. Across the dataset, there is a stable uptake of nitrogen sources including Threonine, Tyrosine, Isoleucine, Leucine , Histidine, Phenylalanine, Lysine, Methionine, Tryptophan, Arginine and Cysteine. Finally, fatty acids palmitate and linoleate uptake is saturated at 0.0015 μmolg^{-1}min^{-1} and 0.0001 μmolg^{-1}min^{-1} , respectively across the whole dataset. Coenzyme A is attached to palmitate to form palmityl-CoA, which takes part in subsequent reactions of lipid production. Linoleate is used to form steraoyl- CoA which also takes part in lipid biosynthetic pathways.

Estimation of physiologically relevant energy and biomass demands

In order to estimate a range in which the combination of ATP$_d$ and bm$_d$ produce an OGI in the 5-6 range, in the next set of analyses the effect of increasing the bm$_d$ on OGI at various ATP$_d$ was measured. Each subsequent LP formulation was defined as in Equation 3.1), where the bm$_d$ was increased from 0 to 0.9 μmolg^{-1}min^{-1} at a constant value of ATP$_d$. This was repeated over a series of ATP$_d$ ranging from 1- 20.8 μmolg^{-1}min^{-1} .

At the lower range of ATP demands, OGI decreases rapidly as a function of biomass demand increase (Figure 2.5) quickly approaching non-physiological brain OGI index values below 5. At a higher range of ATP$_d$ = $\geq$ 18.6 μmolg^{-1}min^{-1} the OGI $\geq$ 6 even at the lowest biomass demands and increases linearly (Figure 2.5). As seen in Section 2.3.1, the increase in OGI is a reflection of glucose saturation and uptake of ketone bodies for energy production. Since the model represents the brain under 'normal', fully-fed conditions when the main substrate for energy production is glucose (Section 1.2.3), the higher ATP$_d$ values can be excluded. Within the middle region of the graph (ATP$_d$ = 7.6 -16.4 μmolg^{-1}min^{-1}) there is a stable decrease in the OGI as the bm$_d$ increases, as anticipated. Therefore, within this range of values of ATP$_d$ the model produces a

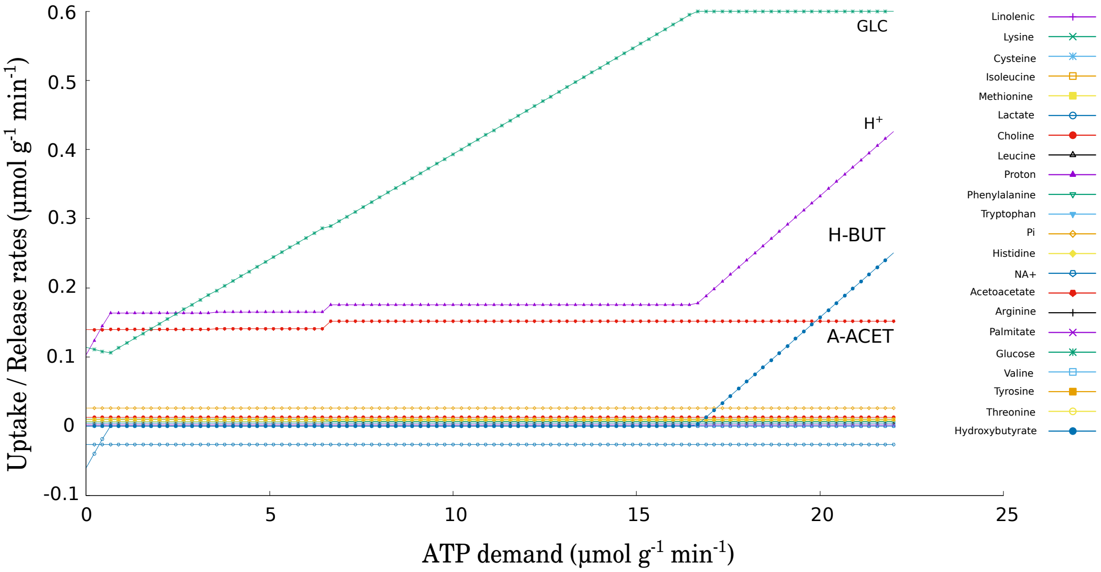


Figure 2.3: **Effect of increasing the ATP demand in the presence of a biomass demand** Glucose is able to support an ATP demand increase only up to ≈16 μmolg^{-1}min^{-1} , after which the uptake of ketone bodies begins. The complete list of precursors required for biomass production can be seen in the legend.

physiologically relevant OGI.

However, this result does not constrain the ATP$_d$ and bm$_d$ to any specific value because any combination of ATP$_d$ 7.6 -16.4 μmolg^{-1}min^{-1} and bm$_d$ 0- 0.09 μmolg^{-1}min^{-1} will produce an OGI that falls between 5-6. Since glucose uptake of 0.6 μmolg^{-1}min^{-1} defines the most extreme limit of glucose consumption which at resting state is closer to 0.2-0.4 μmolg^{-1}min^{-1} , an additional constraint on these values can be applied by specifying a limit on glucose consumption. Under the assumption that neurons have a relatively low biomass turnover but maximise their ATP production, specific values for ATP$_d$ and bm$_d$ were obtained by running the same analysis over a larger number of data points, ranging ATP$_d$ from 0.001- 17 μmolg^{-1}min^{-1} and bm$_d$ from 0.11- 0.10 μmolg^{-1}min^{-1} . Maximum feasible ATP$_d$ and bm$_d$ were searched for such that they satisfy the following criteria;

- OGI $\geq$ 5.4
- glucose uptake rate $\leq$ 0.35 μmolg^{-1}min^{-1}

The final estimations which satisfy the above are ATP$_d$ = 9.6 μmolg^{-1}min^{-1} and bm$_d$ = 0.027 μmolg^{-1}min^{-1} which are the constraints used in the following model analysis sections. The ATP demand is in line with the value of 9.5 μmolg^{-1}min^{-1} for the forward rate of the ATPase reaction in the grey matter of the human brain [167]. Solving an LP formulation using the demands

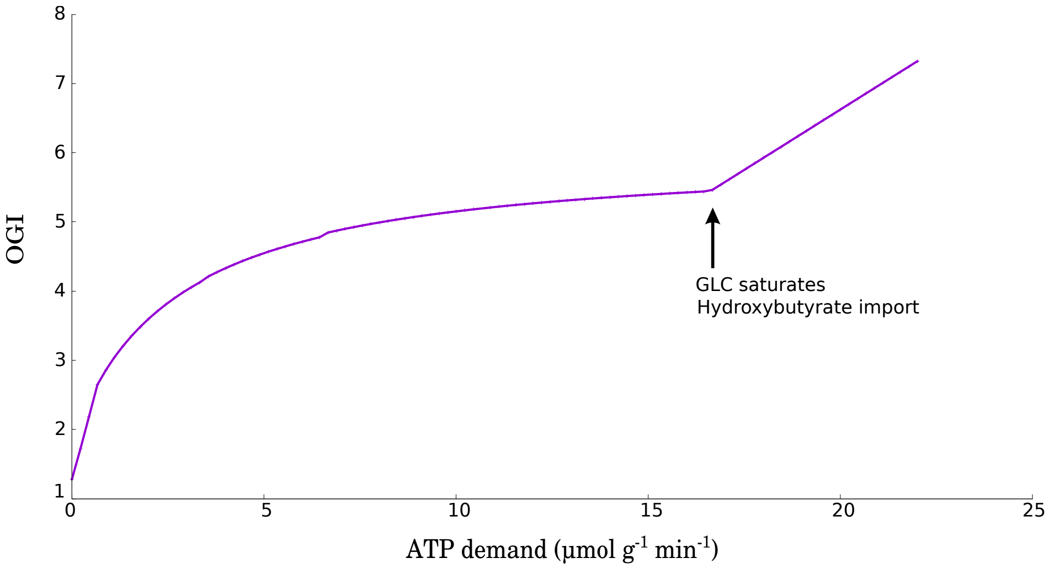


Figure 2.4: **OGI increase across the ATP demand scan** The trend of OGI increase is hyperbolic when glucose is the only used substrate as more glucose is processed via oxidative phosphorylation. The linear trend begins upon glucose transport saturation, when ketone body uptake switches on.

set to the calculated values results in the rate of malate dehydrogenase reaction which catalyses malate to oxaloacetate conversion equal to 0.63 μmolg^{-1}min^{-1} . Using this reaction as a proxy to the rate of the overall TCA cycle, this value is similar to the rate of the TCA cycle estimates from kinetic modelling of NMR data in human brain, ranging from 0.6 μmolg^{-1}min^{-1} to 0.77 μmolg^{-1}min^{-1} [168–170].

It is worth to note that the these values are only an approximation for use in the subsequent analysis, and the true values will be a reflection of the substrate availability and physiological state of the brain in a given moment.

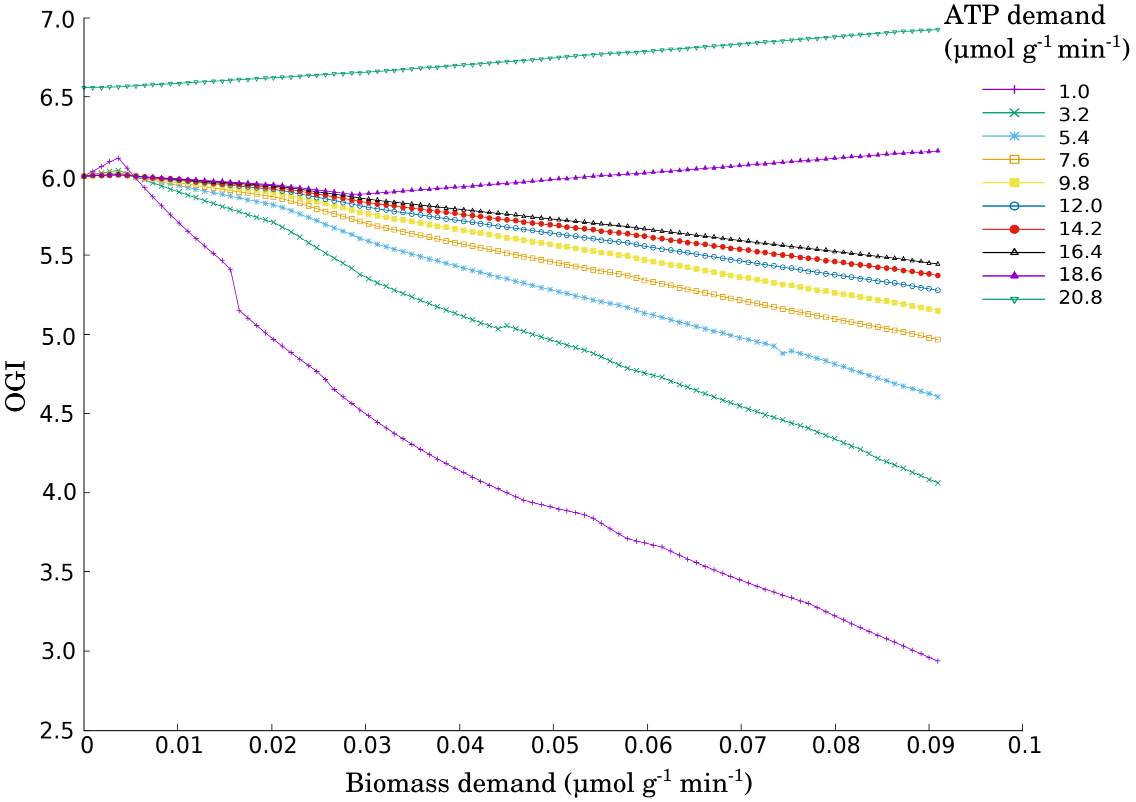


Figure 2.5: **Effect of changing the ATP and biomass demands on the OGI values** Biomass was increased at various ATP$_d$ ranging from 1.0 - 20.8 μmolg^{-1}min^{-1} . At the lower ATP$_d$ (1.0 - 5.4 μmolg^{-1}min^{-1}) there is an unstable and rapid decrease in the OGI which is not as dramatic in the regions of ATP$_d$ 7.6 -16.4 μmolg^{-1}min^{-1} . Above the values of ATP$_d$ 18.6 μmolg^{-1}min^{-1} there is an increase in the OGI, likely reflecting fast glucose uptake saturation and use of ketone bodies for energy production.

2.3.2 Pathways of glycine production in the model

It has previously been established that CNS glycine is primarly synthesized from glucose-derived serine by the enzyme serine hydroxymethyltransferase 2 (SHMT2) in the mitochondria [39]. Given the slow uptake of serine by the BBB transporters [161] and alternative fates of glucose, questions pertinent to serine metabolism in the first instance remain. In this section the model is used to identify potential pathways of serine/glycine biosynthesis. The relationships between glycine biosynthesis and overall cell metabolic function of energy and biomass production using the rates established in Section 2.3.1 are also explored.

In the first instance, an estimate for the glycine synthesis rate was required. Although the most accurate rates of neurotransmitter formation are determined via MRS, to my knowledge, the rate of glycine synthesis via these methods is not available in the literature. Therefore it was taken from pioneering studies on glycine metabolism by Shank and Aprison [171] performed in rats in which synthesis rates were calculated for white (WM) and grey matter(GM) of different brain areas. In the grey matter of the spinal cord, the rate of glycine synthesis from serine equalled 0.035 $\mu\text{molg}^{-1}\text{min}^{-1}$. Although glycine *synthesis* rate is not equal to its *export*, due to lack of alternative data, this rate was chosen as the demand for glycine export in the subsequent analysis sections.

Glycine biosynthesis in the presence of uptake constraints

Firstly the pathways of glycine biosynthesis without the need for energy production were assessed. The FBA problem was formulated in an LP as follows:

$$
\begin{aligned}
&\text{minimise} \quad &&: \ \textstyle\sum_{i=1}^{m} |v| \cdot j_i \\
&\text{subject to} &&\begin{cases} N_{n,m} \cdot v = 0 \\ 0 \leq v_{tj} \leq t_j \\ v_{\text{gly}_{\text{tx}}} = a_j \end{cases}
\end{aligned}
\tag{2.3}
$$

given constraints and a demand for glycine export, $a_j = 0.035$ $\mu\text{molg}^{-1}\text{min}^{-1}$. Solving the LP produces the solution as in Figure 2.6. Histidine is the most highly consumed metabolite the uptake of which is saturated at 0.1 $\mu\text{molg}^{-1}\text{min}^{-1}$. There is a small uptake of threonine at 0.005 $\mu\text{molg}^{-1}\text{min}^{-1}$, and a net efflux of glycine, lactate and CO_2 (both at 0.0025 $\mu\text{molg}^{-1}\text{min}^{-1}$). Glycine biosynthesis takes place in the mitochondria via the forward reaction of SHMT2, as well as the reverse action of the glycine cleavage system (GCS) which balances all of the methylene-THF produced in the former. Thus, the carbon released during serine breakdown is not ‘wasted’, rather used for the formation of the desired product which is further exported into the cytoplasm. Twenty percent of glycine formed is used to synthesize serine via the reverse reaction of the cytoplasmic SHMT1, and the remainder released extracellularly.

The majority of serine is formed from glutamate released in the histidine degradation pathways and 3-phosphopyruvate that is generated from the reverse reactions of glycolysis, via the glycolytic intermediate 3-phospho-glycerate. The pathway of threonine degradation leads to the formation of 2-oxobutanoate via threonine deaminase which is transported into the mitochondria and degraded into succinate, with the production of NADH available to the TCA cycle. In the next step, succinate oxidation to fumarate takes place leading to reduction of ubiquinone to ubiquinol, the transport of 2 H^+_{int} via complex III and subsequently ATP formation via the ATP synthase. Fumarate is used to produce phospho-enol-pyruvate which is transported into the mitochondria for the formation of 3-hydroxypyruvate required for serine synthesis.

Histidine is the major source of N in the solution, contributing 80% of the total as well as 100% of the total glutamate production. The net release of lactate results from the conversion of malate to pyruvate in the cytoplasm via the malic NADP enzyme, with a reduction of NADP to NADPH and the release of CO_2. This is the only reaction providing reduced NADPH, which is necessary for the conversions of 5,10-methenyl-THF to methylene-THF in the cytoplasm and mitochondria.

Glycine biosynthesis in the presence of an ATP demand

In this section the effect of adding an ATP demand on glycine biosynthesis is described. The LP formulation is the same as in Equation 2.3 with an additional constraint on flux via the ATPase reaction set at 9.6 μmolg^{-1}min^{-1} .

Addition of glycine synthesis in the background of ATP production increases the OGI from 6 (Section 2.3.1) to 6.1 . This is expected, as additional O_2 is utilised for the oxidation of amino acids used for glycine production. Overall, TCA cycle reactions comprise 18.15% of the total solution flux, glycolytic reactions carry 5.6% of total flux and PPP is not active.

The majority (85%) of ATP synthesis occurs via oxidative phosphorylation in the mitochondria. As seen in Section 2.3.1, LDH-catalyzed pyruvate to lactate cycling takes place.

Apart from glucose and oxygen, the amino acids isoleucine, histidine and threonine are used as substrates (Figure 2.7). Comparison of results for the demand of an equal amount of ATP (Section 2.3.1) shows there is a slight decrease in the rate of glucose uptake as well as the rate of conversion of pyruvate to lactate. On the other hand, the rate of mitochondrial ATP synthase is slightly increased. Thus, the addition of a demand for glycine synthesis lowers the dependancy of mitochondrial ATP generation on glucose uptake which is fuelled by the C skeletons cleaved from threonine and isoleucine breakdown, which lead to the formation of TCA cycle intermediates succinyl-CoA and acetyl-CoA, respectively. In addition, pathways of isoleucine and threonine degradation produce NADH for use by the ETC. Nitrogen used for the formation of serine phosphate is derived

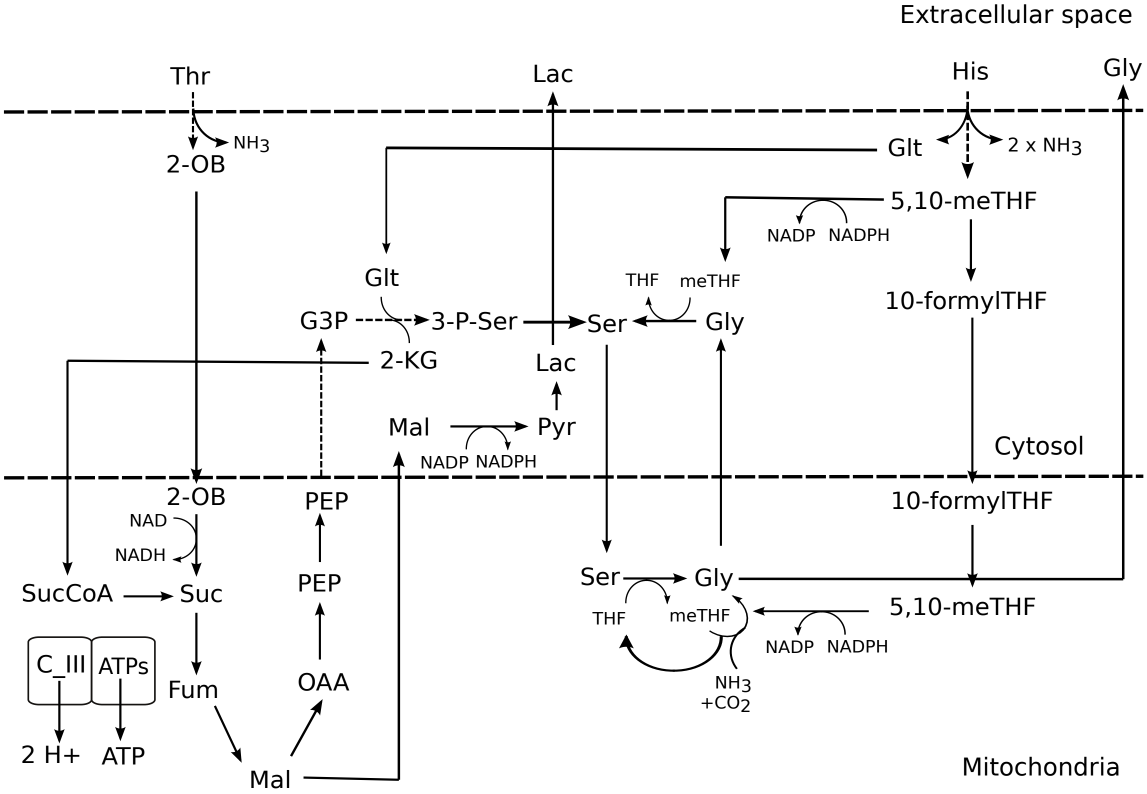


Figure 2.6: **Pathway of glycine biosynthesis in the model** In the absence of other biological objectives, glycine is made mostly from histidine and threonine. This pathway is linked to glycolysis, since hydroxypyruvate is required for serine biosynthesis in the first instance. In addition, the C skeletons of amino acids are sunk in the mitochondria via production of ATP by the ATP synthase.

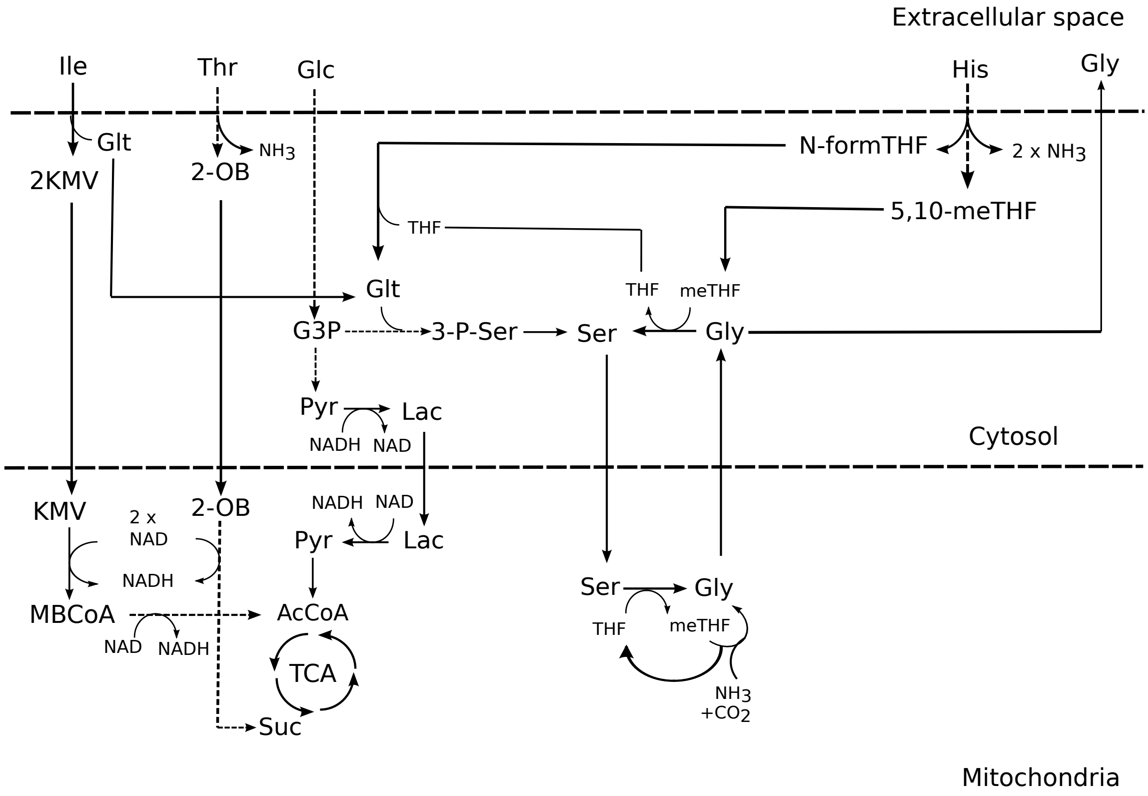


Figure 2.7: **Glycine biosynthesis in the presence of an ATP demand** Addition of ATP production to the metabolic objective induces the uptake of isoleucine, which provides glutamate and 2-keto-methyl-valerate, and the latter is degraded to acetyl-CoA. Folate groups are no longer cycled across the two compartments, as in the previous results.

from glutamate, which in turn comes from pathways of isoleucine (54%) and histidine (34%) degradation. In comparison to the previous result, histidine remains the primary source of NH_3 in the solution (58%) however isoleucine is the major source of glutamate used for the phosphorylation of glycolytic intermediate 3-hydroxypyruvate to form serine. Cytosolic serine is transported into the mitochondria, where 100% is used for glycine biosynthesis via SHMT2. As seen previously in Section 2.3.2, the contribution of SHMT2 to total glycine synthesis is approximately 50% with the other half produced via the reverse action of the GCS. Eighty-five percent of glycine is released extracellularly and the remaining 15% used to fuel the production of cytoplasmic serine via reverse SHMT1.

Response of serine/glycine biosynthetic pathways to an increase in biomass turnover

Pathways of serine/glycine production are coupled to biomass increase in human cells due to the involvement of these metabolites in one carbon metabolism as well as lipid biosynthesis (Introduction, Section 1.2.4). In order to explore the possible changes in the serine/glycine biosynthetic pathways the sum of the biomass flux was increased in the background of an ATP demand, with no forced glycine export. This was formulated in an LP as in Equation 3.1, where the sum of biomass flux was increased from 0 to 0.1 μmolg^{-1}min^{-1} at a constant value of ATP$_d$ = 9.617 μmolg^{-1}min^{-1} . There is an overall increase in glucose consumption but it does not reach transport saturation. Analysis of individual reaction fluxes of glycolysis shows that the pathways leading to the formation of glyceraldehyde 3-phosphate from glucose decrease, however there is an increased flux in the downstream reactions leading to pyruvate synthesis. This is because glucose-6-phosphate conversion to 6-phospho D-glucono-1,5-lactone increases linearly, as well as the subsequent reactions of the PPP which contribute to over 90% of NADPH production from bm$_d$ = 0.011 μmolg^{-1}min^{-1} onwards. Consequently, there is a linear drop in the OGI value which equals 6.0 when there is no biomass demand and reaches 4.1 at maximal demand.

Across all solutions, there is a glycine and serine turnover since, apart from their role in phospholipids (serine) and purine (glycine) biosynthesis, they are themselves components of biomass precursors. In the range of biomass demands 0- 0.09 μmolg^{-1}min^{-1} total serine is synthesized from glucose and glutamate in the cytoplasm (Figure 2.8,a). Serine that is not used in biosynthetic pathways is transported into the mitochondria for glycine biosynthesis via SHMT2 and reverse GCS. Up to 0.05 μmolg^{-1}min^{-1} , around 15% of glycine is synthesized from serine in the cytoplasm via SHMT1 which is the main source of methylene-THF used in inosine monophosphate biosynthesis. This reaction becomes inactive from 0.07 μmolg^{-1}min^{-1} onwards which correlates with a large increase in the uptake of histidine. At this point, the rate of glycolytic reactions downstream of glyceraldehyde-3-phosphate, including the conversion of pyruvate to lactate, increases and the rate of serine transport into the mitochondria decreases. Consequently, there is a decrease in the percent glycine production via SHMT2 and an increase in its biosynthesis from the reverse GCS. At the higher biomass demands ($\geq$0.09 μmolg^{-1}min^{-1}) reverse reaction of SHMT1 switches on and cytoplasmic glycine begins to be used for serine formation as seen in Section 2.3.2.

The reactions of histidine degradation (Figure 2.8,c) form an enzyme subset which leads to the formation of 5,10-methenyl-THF. Therefore, if the cell acquires its C and N atoms via this pathway there is an obligatory production of methylated THF. Due to the steady state condition, this is balanced by a net decrease in methylene-THF formation from serine as well as reverse SHMT1-catalysed reaction which acts both to replenish the cells tetrahydrofolate and as a sink for surplus methylene-THF connecting glycine and histidine metabolic pathways via tetrahydrofolate cycling reactions.

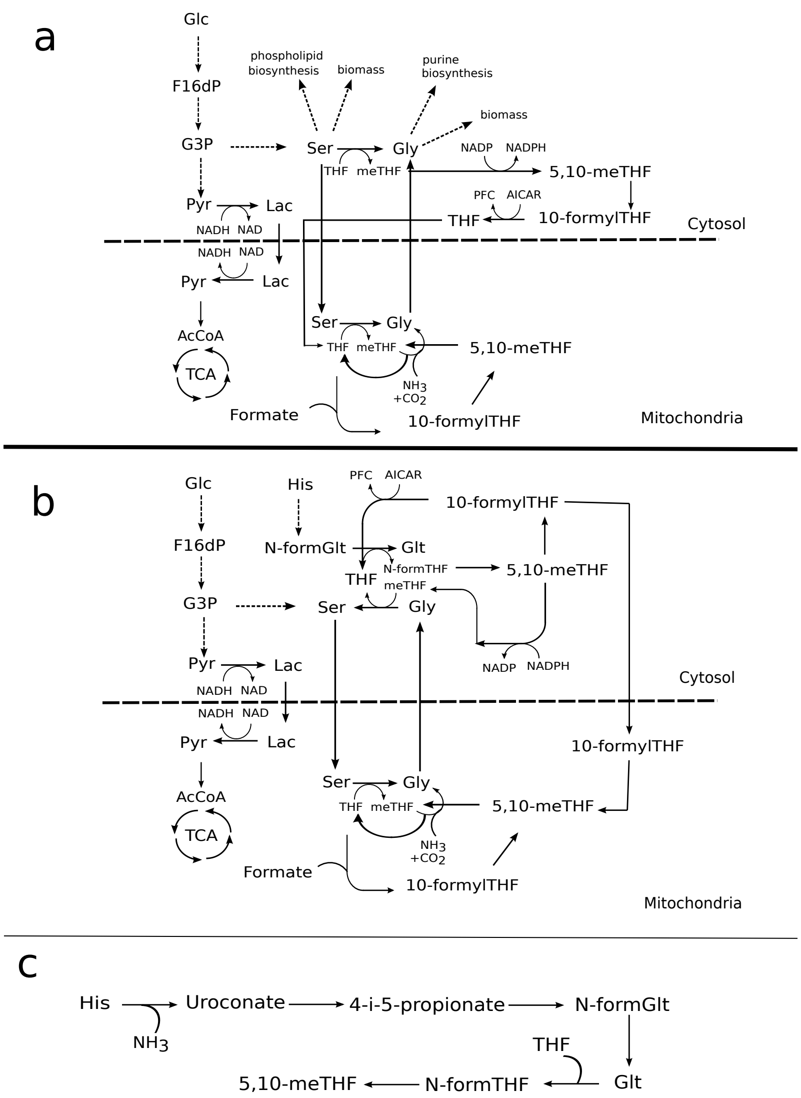


Figure 2.8: **Pathways of serine/glycine metabolism at different biomass demands** At lower biomass demands ranging up to 0.07 μmolg^{-1}min^{-1} (a) cytoplasmic SHMT1 operates in the glycine- forming direction, providing methylene-THF groups required for pathways of IMP synthesis. Carbon atoms enter the folate pool from serine, as well as formate in the mitochondria. Serine is used for glycine formation, as well as in the pathways of phospholipids biosynthesis (inluding phosphatidylcholine, phosphatidylethanolamine, and phosphatidylserine). Glycine is used in the formation of purine nucleotides. Tetrahydrofolate crosses the compartments and is regenerated in the cytoplasm from AICAR and 10-formylTHF in the intermediate pathways of IMP biosynthesis. At higher biomass demands, histidine uptake and degradation pathways switch on. In the range of bm$_d$ 0.07-0.09 μmolg^{-1}min^{-1} , cytoplasmic SHMT1 becomes inactive. It begins to carry flux again at bm$_d$ $\geq$0.09 μmolg^{-1}min^{-1} , where SHMT1 uses methylene-THF and glycine for serine formation, regenerating tetrahydrofolate in the process (b). Since pathways of histidine degradation form an enzyme subset (c) by necessity they must carry equal flux, leading to consumption of tetrahydrofolate and production of 5,10-methyleneTHF, adding an additional entry of carbons into the folate pool. Conversion of glycine to serine may therefore be active to consume the additional C atoms.

2.4 Discussion

ATP production in the model

The model was used to establish the likely pathways of ATP production in an ATP scan. At all energy demands, glucose is processed via glycolysis for pyruvate formation in the cytosol. Next, pyruvate is converted to lactate via LDH leading to the regeneration of cytosolic NAD. Lactate is co-transported with a H^+ into the mitochondria where it is re-converted to pyruvate via the mLDH, leading to the formation of NADH for use in the TCA cycle. Given there was no constraint applied on the uptake rate of oxygen, these results are not representative of the typical description of lactate production under low oxygen tension occurring due to a mismatch between the rate of glycolysis and the TCA cycle [172].

Contrary to the classical hypothesis, it has previously been proposed that in all tissues lactate is constantly being produced under aerobic conditions and is the natural end-product of glycolysis [173]. The ability of the brain to use lactate as an energy substrate has been established in neuronal cells *in vitro* [174] as well as in human brains *in vivo*. Whole-brain glucose uptake rates in human are decreased in the presence of circulating plasma lactate [175], showing that it can be used as an energy substrate. The ability of mitochondria to oxidise lactate relies on the presence of transporters across the mitochondrial membrane. Both lactate and pyruvate are transported across biological membranes by a family of the monocarboxylic acid transporters (MCTs) the expression of which is widespread across the brain and has been shown in the mitochondrial membranes of rat neurons [176]. Following import, the expression of mLDH is required and was demonstrated in mitochondrial extracts from isolated human astrocytes, capable of generating ATP from lactate [177].

Oxidative phosphorylation in the mitochondria is dependant on the availability of reducing equivalents, mainly NADH. In a similar manner, the rate of glycolysis will be determined by the regeneration of NAD in the cytosol. In the model, both of these requirements are met by the inter-compartmental conversions between pyruvate and lactate where the conversion of mitochondrial lactate to pyruvate provides NADH needed for oxidative phosphorylation without the activity of the malate-aspartate shuttle. Since the model is a simplification of the system, it does not contain the intermembrane space compartment, which is the location of mLDH. Biologically, the system would still require the operation of the malate-aspartate shuttle, which generates NADH in the mitochondrial matrix, as required for ETC activity- thus the absence of the malate-aspartate shuttle reactions is likely an artefact of the model definition. Nevertheless, it has been proposed that the inter-compartmental conversions of pyruvate and lactate is important in the homeostasis of NAD/NADH levels in the cytosol. Secondly, proton release which accompanies lactate oxidation in the intermembrane space contributes to the proton gradient forming the proton motive force for use by the ETC, or may also be used indirectly by the malate-aspartate shuttle for transport into the matrix [151].

The model predicts the maximal rate of ATP production from glucose at 16.7 μmolg^{-1}min^{-1} which is lower in comparison to the estimated values of 30-45 μmolg^{-1}min^{-1} for the rat brain grey matter, or 20 μmolg^{-1}min^{-1} for rat brain as a whole [19]. This could be due to a difference in the rates of glucose consumption in a rat versus human. In addition, import of lactate or other possible energy substrates was not allowed in the simulation, thus the rate is representative of ATP production from glucose only. Brain energy production also relies on the catabolism of amino acids, most notably glutamate which is widely used across the brain. It has been previously shown that astrocytes catabolize some of the uptaken glutamate to CO_2 *in vitro* [178]. Thus, the slightly lower estimation of 16.7 μmolg^{-1}min^{-1} for the rate of ATP biosynthesis is possibly a result of incomplete representation of substrate availability in the brain. In addition, the model does not depict *all* of the processes active within the brain at any moment, such as the energy-consuming pathways of vesicle recycling across chemical synapses, and others.

In the next step, the effect of adding a biomass demand on energy production was assessed in the presence of ketone body uptake. In comparison to the previous results where the OGI ratio is six across all solutions, showing complete glucose oxidation, the OGI increases as the ATP demand goes up. This is because at lower biomass demands majority of glucose is shunted to the oxPPP pathways, generating NADPH used mostly in lipid biosynthesis. Increasing the demand for ATP causes a large increase in the uptake of oxygen used in oxidative phosphorylation, explaining the higher coupling between the consumption of the two. Upon glucose exhaustion, there is an increase in the uptake of hydroxybutyrate. The trend of OGI increase is hyperbolic as glucose is used as the primary energy source and switches to linear upon the increase of hydroxybutyrate uptake thus, the trend of OGI increase is informative of the type of substrate used.

It has been shown that ketone bodies can protect the viability of rat hippocampal neurons from glucose deprivation *in vitro* [179]. However, this scenario is applicable under glucose deprivation, causing a high concentrations of circulating ketone bodies produced by the liver and is not relevant in fully-fed conditions. Lastly, the physiological range for brain OGI subject to glucose uptake rates were used in order to estimate the ATP and biomass turnover demands. The resulting forward rate of the ATPase reaction was estimated at 9.6 μmolg^{-1}min^{-1} , which is in line with the previously published estimations at 9.5 μmolg^{-1}min^{-1} in the grey matter of human brain [167]. At this value of ATP demand and a biomass turnover demand of 0.03 μmolg^{-1}min^{-1} , the rate of malate dehydrogenase reaction of the TCA cycle is 0.6 μmolg^{-1}min^{-1} . If this reaction is used as a proxy for the overall rate of the TCA, the value of 0.6 μmolg^{-1}min^{-1} is in line with previously published results for the rate of the TCA cycle ranging from 0.6- 0.77 μmolg^{-1}min^{-1} [168–170]. Thus, the OGI ratio can be meaningfully used for the predictions of the rate at which oxidative phosphorylation in the mitochondria occurs. These values provided an estimate for use in the following model analysis sections.

Glycine biosynthesis in the model

In the first instance, pathways of glycine synthesis in the model were assessed in the absence of an energy demand. Under equal constraints on the rate of uptake of imported amino acids, the preferred substrates were histidine and threonine. The involvement of histidine breakdown pathways in the supply of methylated THF is known although its' role in serine/glycine metabolism has not been described. Reactions of histidine breakdown produce two molecules of NH_3 as well as glutamate, thus providing a rich source of nitrogen for glycine metabolism, meanwhile taking part in the reactions of tetrahydrofolate -mediated methyl transfers the metabolism of which is tightly coupled to serine/glycine biosynthetic pathways. Given that regeneration of tetrahydrofolate is required by the mitochondrial glycine production, histidine degradation provides a pathway that can balance the tetrahydrofolate transfers in glycine metabolism without the involvement of purine, methionine or choline metabolism. Although the majority of glycine is formed via the mitochondrial pathways, a proportion of glycine is used to regenerate serine in the cytoplasm.

Previous studies imply that the roles of cytoplasmic and mitochondrial SHMT differ. Mitochondrial SHMT2 has been found to be overexpressed in different cancer cells [180] where SHMT1 is not. Growth of Chinese Hamster ovary (CHO) cell lines mutant for the mitochondrial SHMT2 are dependant on glycine addition to the media [181], and later studies tracking the incorporation of isotopically labelled serine show that majority of methyl groups incorporated into downstream methionine in the cytoplasm are derived from the mitochondrial SHMT2 reaction [182]. On the other hand, studies on cytoplasmic SHMT1-knockout mice show that although the deficiency affects one carbon metabolism in the organism, the mice are viable thus the gene is not essential [43]. In line with these findings, mitochondrial SHMT2 is the main glycine producer in the model, and the backflux of glycine into serine seems only to function as a sink for 25% of methylene-THF groups formed in the pathways of histidine breakdown. The remainining 75% of methylene-THF is sunk into glycine production via reverse GCS. In the model, glycine is produced in nearly equal amounts by the forward conversion of serine to glycine via SHMT2 and the reverse reaction of GCS, which balances the methylene-THF produced in the former. Yet, the primary role attributed to GCS is glycine degradation, as opposed to its synthesis (Introduction, Section 1.2.4). In the pathway of serine production of glucose, carbons at positions 3 and 4 constitute the carboxyl α carbon of serine, and carbon at position 1 constitutes the β carbon of serine. Via the SHMT-catalyzed reaction, the α carbon of serine is incorporated into glycine and the latter released as methylene-THF [39]. Although the reversibility of GCS, producing glycine from methylene-THF was doubted to occur *in vivo* due to thermodynamic constraints [46], formation of two molecules of glycine from serine has been shown by Motokawa and KIKUCHI [45], and later the direct incorporation of the β carbon of serine into glycine via methylene-THF confirmed by the same group in liver mitochondrial extracts. In the spinal cord of rats, Shank et al. [39] showed that although glycine formation from radiolabelled [3,4-^{14}C]glucose is higher than that from [1-^{14}C]glucose, the latter produces glycine suggesting its biosynthesis from the β carbon of serine and methylene-THF . The pathway of glycine production seen in the model is based on the assumption that the reaction can proceed in the reverse direction under physiological conditions, as suggested by previous experimental results.

The breakdown of threonine produces N atoms required for reverse GCS, and leads to the production of oxobutanoate which is next transported into the mitochondria, where it is converted into succinate. The next step of TCA cycle, oxidation of succinate to fumarate, also leads to the reduction of ubiquinone to ubiquinol which results in the transport of two H^+ across the mitochondrial intermembrane space. In addition, the conversion of phosphopyruvate to 3-phosphoserine leads to the formation of 2-oxoglutarate which additionally feeds the TCA cycle pathways. In sum, the biosynthetic pathways of glycine/serine metabolism induce ATP formation by oxidative phosphorylation in the mitochondria due to the requirement for sinking the C skeletons of amino acids used.

In the last step, the increase of biomass production was addressed in a biomass turnover scan, in the background of an ATP demand. As expected, as the demand for lipid production goes up there is a decrease in the OGI index. This is a result of disproportionate oxygen consumption to glucose, as more glucose is converted to D-glucono-1,5-lactone by D-glucopyranose-6-phosphate enzyme and subsequent pathways of NADPH generation by the PPP.

Serine to glycine conversion is the major producer of methylated THF groups [183] which is the case in the model simulations. At lower biomass demands, there is no uptake of histidine and both mitochondrial and cytoplasmic SHMT operate in the glycine-forming direction. Given the low percent turnover of glycine via SHMT1, the primary role of this reaction is likely the production of methylene-THF in the cytoplasm. As the uptake of histidine switches on, cytoplasmic SHMT1 is no longer active until it switches on again at very high biomass demands of 0.09 μmolg^{-1}min^{-1} , where SHMT1 begins to operate in the opposite direction. It has previously been proposed that serine formation from glycine can be used as a sink for additional methyl groups produced in histidine degradation, meanwhile regenerating tetrahydrofolate [184]. As the SHMT2-catalyzed reaction in the mitochondria proceeds forward at all biomass demands, the change in the directionality of SHMT1 across the solutions suggests further that the function of this pathways is to balance the folate derivatives which take part in biosynthetic reactions. The compartmentalization of serine/glycine thus provides metabolic flexibility to requirements imposed by the steady-state on one carbon metabolism.

It is surmised that in the presence of a low demand of biomass turnover, serine/glycine metabolism is unlikely to occur from histidine. Increase in serine and glycine is stoichiometrically linked to the increase in methylene-THF . The folate pool is supplied with C atoms mainly from serine to glycine conversions, as well as formate in the mitochondria. Utilization of histidine introduces a novel pool of methylene-THF which requires to be balanced by the system. This has the effect of slowing down the rate of serine biosynthesis from glucose, and glycine biosynthesis in the mitochondria via the forward reaction of SHMT2 which is the most likely mode of glycine biosynthesis [39]. Production of serine from alternative amino acids, likely threonine and isoleucine as well as glucose does not impose the additional strain on the system. One interpretation of the co-dependencies between the metabolism of histidine, serine and glycine and reactions reliant on the folate intermediates is that overall, this would provide a disadvantage to the cell. It can be imagined that maintaining cell homeostasis is far more simple in a system with a higher number of control points, as opposed to a highly intertwined set of entities. More dependencies would lead to a decrease in both the external and internal entries of regulation. Following the logic, if a cell acquires its N atoms from amino acids other than histidine, there is no additional demand on the system to regulate levels of folate cofactors.

2.4.1 Conclusion

In summary, the major insights of this chapter can be split into those pertinent to energy and biomass generating pathways, and those specific to glycine production under normal physiological conditions. Simulations of energy production constrained by literature-defined upper limits of substrate uptake across the BBB showed an emergence of the lactate-pyruvate shuttling under normal physiological conditions. Although several studies suggest this phenomenon may be a common occurance in the brain (Section 2.4), the model is an incomplete representation of the cell exemplified by the absence of the intermembrane space compartment which highlights the limitations resulting from the simplification of metabolic processes in computational modelling overall.

Nevertheless, the maximal rates of ATP production predicted by the model were not too far from experimental findings (Section 2.4). It should be noted that the estimations were taken from rat brain ATP estimations as available in the literature. The differences in the values are likely due to an incomplete representation of all substrates available in the brain at a given moment, such as recycled amino acids. Furthermore, not all ATP-consuming pathways that fall outside of strictly metabolic processes are represented in the model, such as neurotransmitter vesicle recycling.

Another major insight is the establishment of a method to visualise and quantify changes in the metabolic pathway utilizaton in parallel to variations in the OGI values, as energy and biomass needs of the system are changed. Measurement of the OGI changes is also reflected in the type of substrate used to support these processes, such as a switch to ketone body catabolism as seen in Section 2.3.1.

Related to glycine metabolism, the contribution of amino acid metabolism to ATP generation in the mitochondria and following, the rates of glucose uptake were highlighted (Section 2.3.2). In addition, the possible importance of the reverse action of the GCS in glycine production from serine was shown which is supportedby some experimental findings, although previously debated in the literature (Section 2.4).

Furthermore, contribution of different amino acids to glycine metabolism was assessed in relationship to tetrahydrofolate cycling in the cell. A major limitation of these findings is that glutamate, which is a major amino acid that relays N balancing in the brain, was not present in the simulations. This is taken into account in Chapter 3, where metabolism of glycine across different CNS cell types is assessed, as opposed to the whole brain metabolism.

Chapter 3

Integration of single-cell RNA-seq data

3.1 Introduction

In the previous Chapter, the CNS model was contextualized via manual removal of reactions likely inactive in the brain, application of constraints on substrate uptake rates across the BBB and fine-tuning energy and biosynthetic demands to the brain OGI. Nevertheless, this represents metabolism of the whole organ which is made up of heterogenous cell types, that at the highest glance split into neuronal and glial cells (Introduction, Section1.2.1). Classification of brain cell types has proven complex, partially due to different criteria for specifying subgroups [185]. For instance, some neurons have previously been named according to the brain region they populate rather than function, as is the case for the GABA-ergenic Purkinje neurons present in the cerebellar cortex.

Naturally, these various cell types differ in their function and thus they show a varied signature of structural and metabolic protein content. Efforts towards the improvement of brain cell type classification have come from the advances in the single-cell RNA sequencing (scRNA-seq) technologies which classify and group single cells based on their mRNA expression (Introduction, Section 1.3.4). Computational methods are used to cluster cells based on their transcriptome which are identified by their known molecular markers and in reverse, can help to identify novel molecular markers unique to a cell type.

Apart from classification, these studies can aid in our understanding of the function of these cells- the individual transcriptomes can provide clues on the metabolic pathways the cell uses for energy production and peripheral functions. One of the theories pertinent to the differences in brain cell bioenergetics implies the differential glucose utilization and cycling between excitatory glutameric cells and astrocytes, as described by Pellerin and Magistretti [24] in the astrocyte-neuron lactate-shuttle (ANLS) hypothesis (Introduction, Section 1.2.3). Although evidence supporting the ANLS theory exists in the form of experimental observations of substrate uptake in cultured cells *in vitro* and differential gene expression, direct evidence is still lacking.

In the following chapter, the CNS model is contextualized further via the integration of a scRNA-seq dataset from Zeisel et al. [186]. As described in the Introduction, Section 1.3.4, transcriptomics datasets used for integration with metabolic models are usually the expression patterns of a bulk of tissue. Indeed, majority of the conclusions reached regading the natural world are based on the mean representation of a system which allows one to describe the patterns of behaviour for greater predictability. However, can looking at the pattern of gene expression of a single cell reveal the *possible* states of a cell and its function? Moreover, randomness of gene expression can provide benefits to a cell. The advantage of stochasticity in cellular transcription and translation seen in bacteria is the emergence of sub-populations resistant to its external environments, increasing survival. In eukaryotes, the random gene-protein expression may lead to the differentiation of cellular sub-types from a homogenous population with a unique, specialised function [187]. The possible gains of studying cell heterogeneity in human is apparent in tumour research, where the knowledge of the various mutations within a population can aid in the design of combinatorial treatments [188]. It may be also possible that the study of single-cell expressions in non-cancerous tissue may lead to the discovery of novel, albeit rare cell phenotypes, which was the reason for the choice of the scRNA-seq dataset used in the current study.

In Chapter 2, Section 2.3.1, the brain ATP demand was estimated by fixing the value to the overall brain OGI index. The assumption that follows from this is that glucose is used as the main energy substrate, and energy production occurs mostly via oxidative phosphorylation in the mitochondria. Further to the ANLS hypothesis, this may not hold true for different cell types of the CNS. In order to test the possible differences in energy production across brain cell types, scRNA-seq was applied to the CNS model and the pathways of ATP production across different cell types analysed. In the original publication by Zeisel et al. [186] cells were clustered into cell classes by their genetic expression. In order to test whether the heterogeneity based on the genetic signature overlays with clustering based on the metabolic pathways utilized by cells, pathways of ATP production were used to cluster cells within the dataset.

In the next section, the scRNA-seq dataset was used to analyse possible differences in glycine production in each cell type. Results from Chapter 2, Section 2.3.2 show that serine/glycine metabolic pathways depend on the overall balance of folate derivatives. This may explain the reason for serine hydroxymethyltransferase (SHMT) compartmentalization in the cell. The purpose for the compartmentalization, as well as the implication of an irreversibility constrain on the GCS were assessed in the model weighed by the median enzyme expression of the interneuron cell cluster.

In the last section, the ability of different amino acids to support cellular growth was analysed. As described in the Introduction, Section 1.2.4 expression of genes related to serine/glycine metabolic pathways has been found in a subset of human cancer cells, including gliomas [6–10, 53]. In addition, serine has previously been found to support human cancer cell line proliferation *in vitro* [54]. Further to the interest in the development of novel molecular inhibitors of enzymes involved in serine/glycine biosynthesic pathways for the treatment of cancer [189–191], the model was used to assess the potential of these substrates to support cellular growth.

3.2 Methodology

The dataset was downloaded from http://linnarssonlab.org/cortex/ in a .txt format and organised into a matrix, columns of which were populated with cell IDs and rows with gene names. The entries contained the absolute mRNA counts for 19,972 genes across 3005 individuals cells separated from the rat somatosensory (SS) cortex and CAI region of the hippocampus [186]. It is important to note that expression of mRNA does not always correlate with protein levels thus the use of a quantitative cell proteome would provide more insightful, were it available. Upon inspection of several proteomic datasets [192] it became clear that they usually do not identify more than $\approx$ 300 proteins, which is likely due to difficulties arising from the experimental methods used to obtain proteomic data. Moreover, single cell proteomic datasets are not yet in existence [193]. The scRNA-seq dataset used in this study was chosen as it contains a high coverage or genes in cells separated from two distinct brain regions.

Statistical analysis of the dataset

The majority of data counts (81.2%) equalled zero, which could be an artefact of considerable RNA loss during processing which particularly affects transcripts present in low count numbers [125]. Distribution of the remaining 19% of the data is highly skewed showing a large variability in the numbers of counts with a *Std.dev.* = 19.6 and normality test p-value = 0.0 (Figure 3.1). Overall, out of the positive mRNA counts, 90% are between 1 and 6 with a median value of 2 .

Clustering of the data in the original publication split the cells into six different “level 1” classes and a further 47 sub-classes. Level 1 classification includes glial cells (astrocytes, oligodendrocytes and microglia), neuronal cells (pyramidal CA1, pyramidal SS, interneurons) and endothelial cells. There is an unequal representation of each cell class, with a much larger population of pyramidal CA1 and oligodendrocytes as compared to the other cell types (Table 3.1). As stated in the original publication, there is a difference in the mRNA content between glial and neuronal cells with the former containing fewer total transcripts overall.

In order to assess gene up- or down-regulation across the different level 1 cell classes, the proportion of a specific cell class positive for a gene in question out of all cells positive for that gene was compared to the overall proportion of the cell class in the whole dataset by the χ^2 test. This did not take into account the overall cell numbers, nor the absolute number of mRNA counts, which may affect the results.

Application of the dataset to the CNS model

Genes that do not encode metabolic enzymes which catalyse reactions present in the model were reduntant for the next steps, and thus removed from the dataset. Out of 1747 metabolic reac-

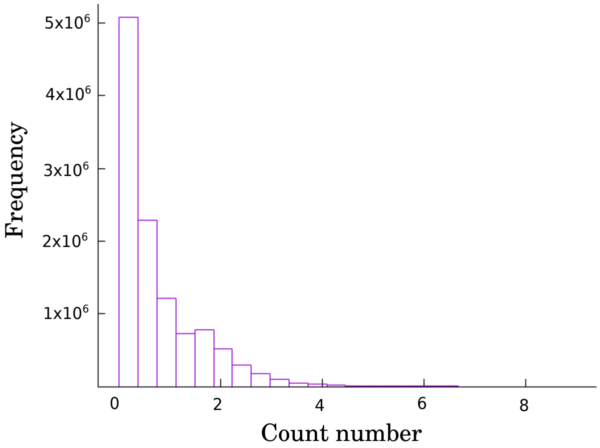


Figure 3.1: **Distribution of gene counts in the scRNA-seq dataset** Data is highly skewed towards the right tail, thus a vast majority of counts are in the lower count numbers, with exceptions of single transcripts present in extremely large counts.

tions in the model, 75% were assigned to one or more gene associations across 860 unique genes. Gene-reaction relationships were identified from the original Recon2.2 model annotations or the HumanCyc database. Out of the 860 genes in the model 86% (740) were present in the transcriptome, which was reduced to include the entries for only these genes across all 3005 cells.

Distribution of the 'trimmed' dataset shows that the percentage of zero counts was slightly lower, at 75.3%. Similarly to the original dataset, after the removal of zero counts the distribution is highly skewed, with few genes present in very high copy numbers.

Given the high proportion of zero counts in the dataset disqualification of reactions based on an expression threshold was not a viable option. Instead, for the purpose of correction for the false-negative bias a zero entry for any gene was replaced by its median value across the whole dataset. Corrections were not cell-specific because of the differences in the numbers of zero counts for each level 1 class. The overall mRNA content is higher for neuronal cell types in comparison to glial cell types (Table 3.1). In effect, the higher number of overall positive counts brings the median value down in comparison to cells with less overall positive hits.

Since the original model annotations did not contain information about the genes associated with the model transport reactions, these reactions were weighed by a coefficient equal to the median of the dataset.

A number of reactions had more than one gene association defined by AND/OR relationships. In order to maximise the feasibility of the objective function, these relationships were not taken into account and the maximum mRNA count for any of the genes associated with the reaction was

Table 3.1: Number of cells in the dataset *versus* zero counts present for each cell type

Level 1 class	Cell number	Zero counts (/740 genes)	Zero counts (/19972 genes)
Pyramidal CA1	939	556	15726
Pyramidal SS	399	580	16412
Interneurons	290	545	15967
Oligodendrocytes	820	630	18365
Astrocytes-ependymal	224	664	19179
Microglia	98	710	19273
Endothelial-mural	236	714	19315

The total number of genes in the original dataset from Zeisel et al. [186] was 19972, out of which 740 genes were associated with one or more reactions in the model.

set as the single expression value. These values were applied in the LP formulation as objective coefficient weights for individual reaction fluxes. Since all the FBA calculations used minimization of the sum of flux as the primary objective function, the weight was defined as the reciprocal of the mRNA count. During minimization, the higher weighting factor makes it less likely for the reaction to appear in a solution. Zero count values were constrained by an arbitrarily large objective coefficient weight of 10, meaning that reactions were not completely disqualified from the solution. The LP formulation can be expressed as follows:

$$\begin{aligned} &\text{minimise} && : \ \textstyle\sum_{i=1}^{m} |v| \cdot j_i \\ &\text{subject to} && \begin{cases} N_{n,m} \cdot v = 0 \\ v_{bj} = b_j \\ 0 \leq v_{tj} \leq t_j \\ v_{\text{ATPase}} = 1 \end{cases} \end{aligned} \tag{3.1}$$

Where the variables have an equal definition to those described in Chapter 2, Equation 2.3. The specific objective function varied according to the problem. The values for the objective coefficient weights from scRNA-seq dataset were collected in the vector j_i and applied in the LP formulation. In all results sections apart from Section 3.3.2, the weights dictionary was calculated from the median expression values of a given cell class. Responses of microglia were not analysed in the following sections because of the small representation of these cells in the dataset, at 98 individual cells only. Endothelial cells were also not taken into account because they are not considered glial or neuronal cells.

Clustering of cells based on FBA solutions

In Section 3.3.2, the scRNA-seq was clustered based on the solutions returned from FBA problems weighted by the gene expression values for each individual cell. The solutions were organised in a matrix as follows:

$$\mathbf{C}_{i,j} = \begin{matrix} v_{11} & v_{11} & \dots & v_{1j} \\ v_{21} & v_{22} & \dots & v_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ v_{i1} & v_{i2} & \dots & v_{ij} \end{matrix} \tag{3.2}$$

where matrix rows, i, represent reaction names and columns, j, represent individual cell IDs. Matrix element v_{ij} contains the flux carried by i_{th} reaction in the solution of j_{th} cell, thus each column vector contains an individual solution. The transpose of this matrix was used to calculate a dissimilarity matrix (Introduction, Section 1.3.3) containing Pearson's correlation coefficients between the solution vectors. The resulting matrix was visualized as a hierarchical tree via the application of a Weighted Pair Group Method using Arithmetic Averaging (WPGMA) clustering algorithm.

3.3 Results

3.3.1 Pathways of energy metabolism in different CNS cell types

In order to establish the possible differences in the ATP production pathways across different cell types, expression of genes related to ATP production was analysed directly by comparison of the expression levels, and later via FBA simulation of the model weighted by the scRNA-seq values with the objective function of ATP production in the presence of glucose and lactate.

Expression pattern of genes related to central carbon metabolism

The trends in the expression pattern of genes related to central metabolism are summarised in Table 3.2 as follows:

1. Pyruvate dehydrogenase kinase (PDK) catalyzes the phosphorylation of pyruvate dehydrogenase (PDH) thus lowering its activity and flux of pyruvate to acetyl-CoA in the mitochondria. The expression of the pyruvate dehydrogenase kinase isozyme 4 (PDK4) is more frequent in astrocytes in comparison to the other cell types, slightly lower in pyramidal CA1 cells and considered ‘normal’ for all other cell types. However, PDK can be transcribed into four isoenzymes and PDK4 is the least expressed type, with a significantly higher number of cells positive for the remaining three (PDK1, PDK2, PDK3). As opposed to PDK4 expression, PDK1-3 are more frequently expressed in pyramidal CA1 cells and below the expected numbers in oligodendrocytes and astrocytes.

2. The overall expression of the PKM gene, which encodes pyruvate kinase that catalyses the dephosphorylation of phosphoenolpyruvate to pyruvate is widespread across all the cell types (2427 cells positive). It is expressed in higher proportion in pyramidal CA1 cells, pyramidal SS and interneurons. On the contrary, less than expected numbers of oligodendrocytes express the gene and it is present at normal levels in astrocytes. Thus, neuronal cell types have a higher PKM expression overall.

3. GLO1 encodes the components of the glyoxylate system, lactoylglutathione lyase and hydroxyacylglutathione hydrolase which are important in detoxification of methylglyoxal produced in glycolysis. Expression of the GLO1 gene is widespread and normally distributed across all cell types.

4. Genes encoding the subunits of lactate dehydrogenase (LDHA-D) which interconverts pyruvate to lactate are slightly higher in all neuronal cell types and normal in astrocytes and oligodendrocytes. This could be suggestive of high rates of pyruvate to lactate conversion, although the opposite also applies given that the reaction is reversible.

5. Differences in the distribution of the expression of genes encoding glycolytic enzymes can be seen in Figure 3.2. A cell was counted as positive for expression if any of the genes encoding the subunits of enzyme in question were present. The expression of nearly all glycolytic enzymes is higher in neuronal cell types and slightly low in astrocytes and oligodendrocytes.

6. In comparison, expression of genes encoding the PPP enzymes shows a split between the oxidative and non-oxidative PPP (Figure 3.3). In general, a smaller number of cells express enzymes of the oxPPP (maximum 712 cells) which are slightly upregulated in neuronal cells and downregulated in oligodendrocytes and astrocytes. Genes relevant to the non-oxidative PPP are more prevalent, with normal expression distribution in both types of pyramidal cells and astrocytes, and slighly higher levels in oligodendrocytes and interneurons. Although the genes of oxPPP are overexpressed in neuronal cell types, glycolytic genes are overall more widespread and significantly overexpressed in pyramidal cells.

Cell type specific ATP production in the presence of glucose only

In this step a systems-level approach was applied and substrate preference for ATP production in each cell type was investigated in a set of FBA problems. In the first instance, the demand for ATP synthesis, v_{ATPase} was increased from 1 to 20 μmolg^{-1}min^{-1} in the background of uptake restrictions (Section 2.3, Table 2.2) which can be expressed in the LP formulation below:

$$\begin{aligned} &\text{minimise} && : \textstyle\sum_{i=1}^{m} |v| \cdot j_i \\ &\text{subject to} && \begin{cases} N_{n,m} \cdot v = 0 \\ 0 \leq v_{tj} \leq t_j \\ v_{\mathrm{ATPase}} = a_j \end{cases} \end{aligned} \tag{3.3}$$

Overall, the metabolic pathways used by interneurons, pyramidal CA1 and pyramidal SS cells were the same, thus they were grouped together into neuronal cell responses. Similarly, the same solutions appeared for oligodendrocytes and astrocytes, which were grouped into glial cell responses. Both cells utilised pathways of glucose uptake and processing via glycolysis and subsequently oxidative phosphorylation in the mitochondria. As seen in the previously described results (Chapter 2, Section 2.3.1), cytoplasmic pyruvate is converted to lactate which is further transported across the mitochondrial membrane and converted back to pyruvate, generating NADH. The single difference between the solutions for each cell type is that glial cells consume less glucose and oxidise it

Table 3.2: Distribution of cell types across cells positive for genes related to central carbon metabolism.

Gene	Pyramidal SS	Pyramidal CA1	Interneurons	Oligodendrocytes	Astrocytes	Microglia
PDK1 (750)	++ (p= 1.9×10^{-12})	++ (p= 0.0)	N	- - (p= 0)	- (p= 1.7×10^{-05})	N
PDK2 (498)	N	++ (p= 3.7×10^{-11})	N	- (p= 1.0×10^{-05})	N	N
PDK3 (795)	N	++ (p= 0.0)	N	N	- (p= 2.6×10^{-08})	- (p= 1.2×10^{-05})
PDK4 (96)	N	- (p= 7.4×10^{-05})	N	N	+ + (p= 0.0)	N
3.3×10^{-14})	- (p= 3.3×10^{-03})	- (p= 2.6×10^{-04})				
PKM (2427)	+ (p= 2.2×10^{-05})	++ (p= 1.9×10^{-14})	+ (p= 2.1×10^{-04})	- - (p= 1.9×10^{-14})	N	- (p= 2.5×10^{-04})
ATP5B (2804)	N	N	N	N	N	N
GLO1 (2391)	N	N	N	+ (p= 2.2×10^{-03})	N	- (p= 4.2×10^{-07})
LDHA (2317)	+ (p= 1.7×10^{-07})	++ (p=0.0)	+ (p= 2.8×10^{-03})	- - (p= 0.0)	N	- (p= 2.8×10^{-03})
LDHB (2559)	+ (p= 8.4×10^{-04})	+ (p= 8.1×10^{-06})	+ (p= 4.8×10^{-03})	N	N	- (p= 3.8×10^{-08})
LDHD (246)	+ (p= 1.1×10^{-03})	+ (p= 4.1×10^{-03})	+ (p= 1.5×10^{-06})	N	N	N

The plus and minus signs represent up- or down- regulation of genes, where + +/- - p= $\leq 1\times10^{-09}$, +/- p= $\geq 1\times10^{-09}$, and N shows a 'normal' distribution, ie. p= ≥ 0.005.

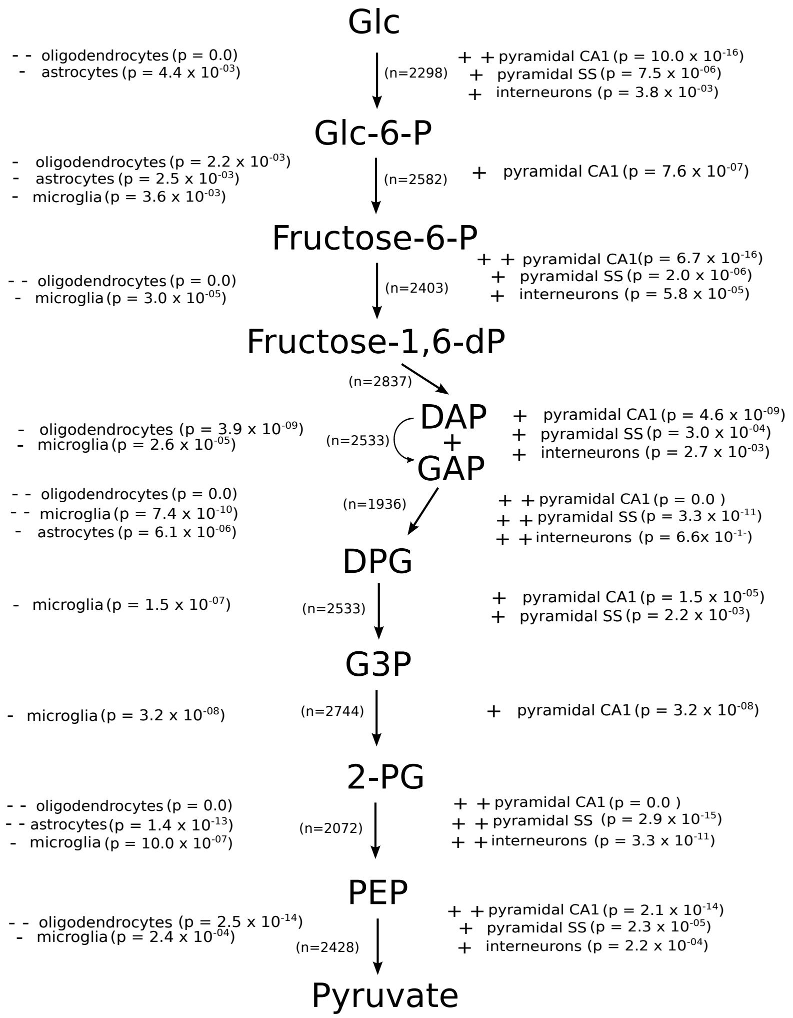


Figure 3.2: **Distribution of cell types amongst cells positive for genes encoding glycolytic enzymes** Overall there is a high expression of glycolytic gene across the whole dataset. Neuronal cell types compose a higher proportion of these cells, with a slightly lower representation of oligodendrocytes and astrocytes. The plus and minus signs represent up- or down- regulation of genes, where + +/- - p= $\leq 1 \times 10^{-09}$, +/- p= $\geq 1 \times 10^{-09}$, and N shows a 'normal' distribution, ie. p= ≥ 0.005.

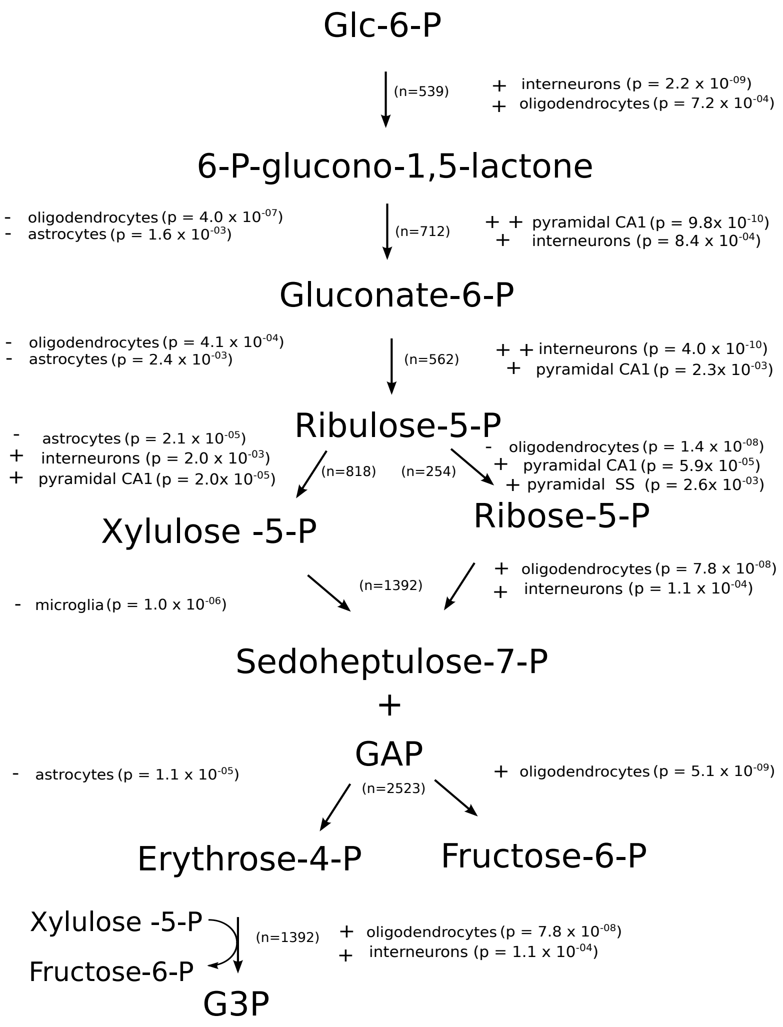


Figure 3.3: **Distribution of cell types amongst cells positive for genes encoding the Pentose Phosphate Pathway** There is an unequal expression of genes of the PPP, with genes encoding enzymes of the oxidative PPP present in fewer cells than the pathways of non-oxidative PPP. Distribution of cell types is more equal than the distribution of genes encoding glycolysis and less homogenous overall, where only the non-oxidative PPP pathways are consistently overexpressed in oligodendrocytes. The plus and minus signs represent up- or down- regulation of genes, where + +/- - p= $\leq 1 \times 10^{-09}$, +/- p= $\geq 1 \times 10^{-09}$, and N shows a 'normal' distribution, ie. p= ≥ 0.005.

completely, whereas neuronal cells saturate the uptake of glucose even at the lower energy demands and release the surplus as lactate (Figure 3.4). It follows that the OGI index for glial cells is a perfect ratio of six across all solution up to 18 μmolg^{-1}min^{-1} , whereas there is a linear increase in the OGI for neuronal cell types (Figure 3.4, b), where the increasing ATP$_d$ causes the cell to use more glucose directly for energy production and not 'wasting'.

The values of ATP$_d$ at which the OGI falls into the previously recorded range for the whole brain (5.4-6.0) are between ≈17-19 μmolg^{-1}min^{-1} for neuronal cells. At an ATP$_d$ of 17.5 μmolg^{-1}min^{-1} , neuronal OGI = 5.44 and lactate efflux is 18% of glucose uptake, in line with the range previously reported in isotopic label incorporation studies [108]. Thus, this value for ATP$_d$ will be used to constrain the background ATP$_d$ in neuronal cells in the subsequent problems. Since the glial cell OGI index remained consistent across the range at six, to allow for direct comparison of solutions the same value for ATP$_d$ will be used to constrain glial energy demands. This assumption may not hold true, as it is thought that neuronal cells have a higher energy demand than glial cells (Introduction, Section 1.2.3) meaning that energy consumption of a glial cell should be adjusted to a lower value. However, reactions utilized by glial cells remain constant across all energy demands thus the absolute value of energy production is irrelevant.

Cell type ATP production in the presence of glucose and lactate

In this section, the substrate preference of each cell type was assessed. The ATP$_d$ was increased from 1-30 μmolg^{-1}min^{-1} in the background of an equal uptake constraint on glucose as in Section 3.3, with the addition of a bi-directional lactate transporter reaction which was left unconstrained. Neuronal cell types utilized the same metabolic pathways across all ATP$_d$ as in the presence of glucose only. Up to ATP$_d$ $\approx$ 20 μmolg^{-1}min^{-1} , glucose consumption is saturated and there is a net release of lactate, which begins to be imported at higher demands. On the contrary, the response of the two glial cell types differs- oligodendrocytes show a preference for glucose uptake with no net lactate release. Similarly to neurons, lactate begins to be consumed upon glucose saturation at ATP$_d$ $\approx$ 20 μmolg^{-1}min^{-1} . On the other hand, astrocytes show a preferential lactate consumption at all ATP$_d$. It is imported into the mitochondria for pyruvate production and full oxidation in the TCA cycle, and glycolysis is inactive altogether (Figure 3.5).

To sum up, in the absence of lactate uptake, neuronal cell types do not optimize their glucose consumption at low energy demands and release lactate as a byproduct whereas glial cell types oxidise glucose fully at all energy demands. When lactate is available, ATP production pathways remain the same for neuronal cell types and oligodendrocytes, which begin to use lactate as a substrate only upon glucose saturation. However, astrocytes show a preference for lactate at all energy demands which is the sole substrate imported and is directly shunted into the mitochondria for complete oxidation.

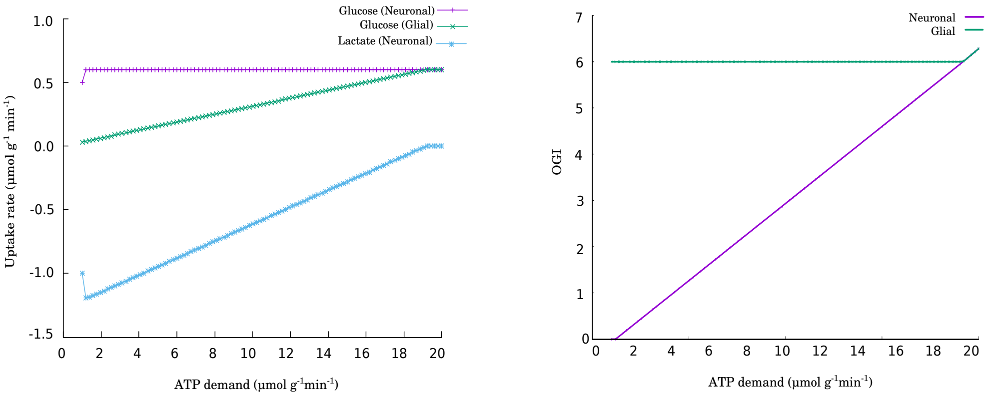


Figure 3.4: **Uptake of glucose *versus* lactate export across a range of ATP demands in different cell types** Neuronal cells saturate their glucose uptake at all ATP_d and release lactate up to ≈ 18 $\mu\text{molg}^{-1}\text{min}^{-1}$, when lactate release stops. In effect, OGI increases linearly as more O_2 is consumed by the mitochondria to oxidise glucose. On the contrary, glial cells do not release lactate at any ATP_d and seem to 'optimise' their glucose uptake as all that is imported is used for energy production.

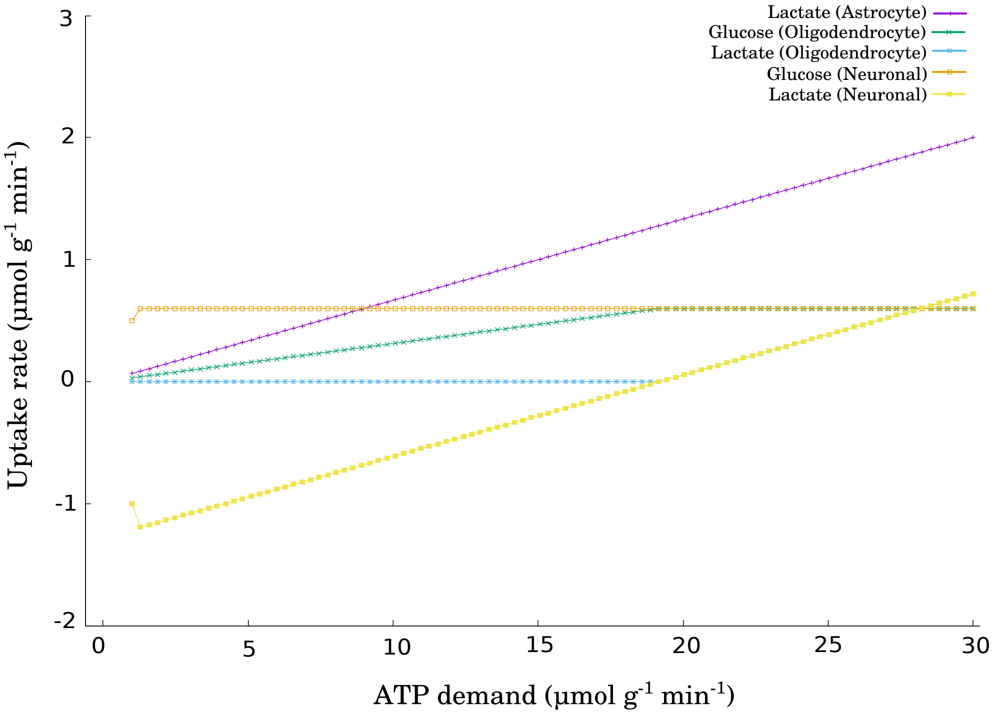


Figure 3.5: **Uptake of glucose and lactate across an increasing ATP demand in different cell types** Neuronal cells and oligodendrocytes prefer to use glucose and only begin to import lactate only at $\text{ATP}_d \approx 20\ \mu\text{molg}^{-1}\text{min}^{-1}$ and onwards. Astrocytes do not import glucose and have a preference for lactate at all ATP_d .

3.3.2 Clustering of cells based on their metabolic response to an ATP demand

Previous solutions represent the generalised responses of each cell class, without taking into account the stochasticity of gene expression in individual cells. In order to test whether clustering based on a metabolic objective function will produce similar result as in the original publication, as well as to identify whether it can lead to the discovery of novel phenotypes, cells were clustered by their response to an ATP demand. In each LP formulation the objective coefficient weights vector was weighted by the expression values of an individual cell, in the background of an ATP_d at 17.5 μmolg^{-1}min^{-1} under constraints as in Equation 3.3.

The metabolic tree generated from the solutions can be seen in Figure 3.6 and is discussed below. Due to a large number of individual tree leaves the nodes were condensed, with each node representing a group of cells that produced the same solution. Overall, 95% of all cells utilize one of the two pathways of ATP synthesis described in the previous Section 3.3.1.

Node L13 contains cells that completely oxidise glucose for energy production, with no lactate release and a perfect OGI of six. As expected, the majority of the cells belonging to this group are glial cells: oligodendrocytes (45%) and astrocytes (13%), although there is a significant representation of neuronal cells in this node, with 15% pyramidal CA1, 7% pyramidal SS and 3% interneurons.

The other major group, node L14, contains cells that release lactate and thus do not use all the glucose for ATP production, with an OGI of 5.4. It follows that an overwhelming majority of cells in this group are neuronal, with 55% pyramidal CA1, 24% pyramidal SS and 19% interneurons. There is a small representation of glial cell types, with 2% oligodendrocytes and 0.16% astrocytes.

Out of the remaining nodes, only L2 and L7 contained more than 60 cells. Node L2 is a cluster of 65 cells made up largely of microglial cells (88%) which, apart from glucose, uptake ketobutyrate which is transported into the mitochondria and convert it to acetoacetyl-CoA and subsequently acetyl-CoA. Node L7 is a cluster of 65 cells made up of a large population of oligodendrocytes at 75% which completely oxidise glucose to ATP. The single difference between the solution of L13 and L7 is the succinate-CoA synthetase-catalysed reaction, which cleaves Coenzyme A from succinyl-CoA in the formation of succinate and a nucleoside triphosphate. In the solution of L7 cells, this reaction phosphorylates GDP to form GTP, instead of ATP.

Clustering based on the metabolic objective of ATP production does not result in the differentiation of cells to the same extent as genetic expression does- cells split roughly into two groups of neuronal and glial types without further subdivisions. The small population of microglial cells (98) were the only cell class well differentiated from others as 65/98 belonged to a distant node L2 which, apart from glucose, imported ketone bodies for energy production.

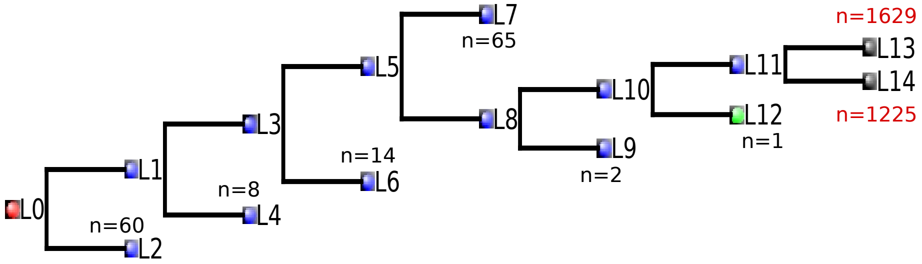


Figure 3.6: **Clustering of cells based on the pathways used for ATP production** Majority of cells produce ATP via one of the two pathways described in Section 3.3.1, belonging to nodes L13 and L14. Out of the remaining nodes, only L2 and L7 having a significant representation of 60 cells or more. Branch length does not represent the distance between nodes.

3.3.3 Pathways of glycine metabolism across different CNS cell types

For the purpose of comparing pathways of glycine production in neuronal *versus* glial cells, in the following section the median mRNA expression for each cell type was applied to the model. In the second step, the model was contextualized further and the expression data for a median interneuron was applied to assess glycine biosynthesis under different scenarios.

Constraints applied on glucose and amino acid uptake

The availability of substrate *within* the brain is not constrained by the BBB, and it is accepted that metabolites are exchanged between different cell types as is the case in the glutamate-glutamine cycle [194]. Quantification of amino acid levels in post-mortem human brain and spinal cord reveals a relatively high concentration of these amino acids in their free form across all regions of the grey and white matter. Similarly, glycine and serine are concentrated in the white matter of the spinal cord, correlating with their function in this region [195]. Consequently, the uptake of glutamine, glutamate and serine were left unbounded in the following LP formulations. The upper bounds on the remaining metabolites introduced in Chapter 2, Section 2.3, Table 2.2 remained constant.
Results from Section 3.3.1 support the notion that in the presence of glucose and lactate, neuronal cells have a preference for glucose and a high rate of oxidative phosphorylation. Based on this assumption, the ATP demand of the cell was fixed to 17.5 $\mu \mathrm{mol g}^{-1}\mathrm{min}^{-1}$ in the subsequent FBA problems.

Expression of genes related to glycine metabolism and transport

Gene expression of enzymes related to the biosynthesis and transport of serine and glycine are summarised in Table 3.3.

1. Genes encoding the enzymes catalysing serine biosynthesis from glucose, PHGDH (phosphoglycerate dehydrogenase), PSAT1 (phosphoserine aminotransferase), PSPH (phosphoserine phosphatase), are widely expressed in the dataset and are significantly overexpressed in oligodendrocytes. The gene PHGDH, encoding phosphoglycerate dehydrogenase which catalyzes the first step of the pathway converting 3-phospho-glycerate to 3-hydroxypyruvate is also significantly overexpressed in astrocytes. The expression of PHGDH and PSAT1 which encodes phosphoserine aminotransferase is significantly lower in all three types of neuronal cells.

2. The distribution of SHMT genes encoding the two isoenzymes of serine hydroxymethyl transferase is normal across all cell types. The expression of mitochondrial enzyme, SHMT2 is higher overall than that of cytoplasmic SHMT1.

3. The expression of GCSH, which encodes the GCS component protein H is high across the dataset with no significant up- or down-regulation in any cell type.
4. Genes PEPD and CNDP2 encode dipeptidase enzymes, prolidase and carnosinase which catalyze the hydrolysis of amino acid pairs. The expression of these genes is widespread across the dataset, and significantly higher in pyramidal CA I cells. The expression of CNDP2 is significantly high in interneurons.
5. The expression of serine and glycine transporters shows a distict pattern. SLC6A9 encodes the glycine transporter 1 protein responsible for the uptake of glycine [196] and is significantly overexpressed in oligodendrocytes and underexpressed in pyramidal CA I cells. The same pattern is present in the expression profile for SLC1A4, encoding the Asct-1 which is involved in the transport of serine, alanine and threonine [197]. The expression of SLC7A10, which encodes Asc1 involved in the Na^+ -independent transport of glycine, alanine, serine and cysteine [198] is significantly high in astrocytes.

Pathways of glycine biosynthesis in different cell types

Glycine export was fixed at 0.035 μmolg^{-1}min^{-1} in the background of an ATP demand at 17.7 μmolg^{-1}min^{-1} and constraints on substrate uptake. The LP formulation can be described as follows:

$$\text{minimise} \quad : \quad \sum_{i=1}^{m} |v| \cdot j_i$$

$$\text{subject to} \begin{cases} N_{n,m} \cdot v = 0 \\ 0 \leq v_{tj} \leq t_j \\ v_{\text{ATPase}} = 17.5 \\ v_{\text{gly}_{\text{tx}}} = a_j \end{cases} \tag{3.4}$$

As can be seen in Figure 3.7, both cell types import serine which is transported into the mitochondria for glycine formation via the mitochondrial SHMT2. Carbon released in this reaction is recycled back into glycine via the reverse action of GCS, with each reaction contributing 50% to total glycine production. The single difference between solutions is the source of N atoms used by the reverse GCS-catalysed reaction and serine production from hydroxypyruvate. Neuronal cells

Table 3.3: Distribution of each cell type across cells positive for genes related to serine and glycine biosynthesis and transport

Gene	Pyramidal SS	Pyramidal CA1	Interneurons	Oligodendrocytes	Astrocytes	Microglia
PHGDH (1150)	- - (p= 0.0)	- - (p= 0.0)	- - (p= 0.0)	+ + (p= 0)	+ + (p= 0.0)	- (p= 5.0×10^{-6})
PSAT1 (1604)	- (p= 1.2×10^{-0})	- - (p= 6.7×10^{-16})	- (p=2.7×10^{-08})	+ + (p= 0.0)	+ (p= 1.0×10^{-03})	- (p= 3.3×10^{-08})
PSPH (1149)	N	N	N	+ + (p= 0)	N	- (p= 2.3×10^{-05})
SHMT1 (84)	N	N	N	N	N	N
SHMT2 (413)	+ (p= 7.7×10^{-04})	N	N	N	N	N
PEPD (1315)	N	+ +(p= 3.5×10^{-13})	N	+ (p= 2.2×10^{-03})	- - (p= 7.5×10^{-11})	N
CNDP2 (1032)	+ (p= 4.4×10^{-03})	+ (p= 4.2×10^{-08})	+ + (p= 9.7×10^{-15})	- (p= 1.3×10^{-07})	- (p= 1.0×10^{-07})	N
GCSH (1277)	N	+ (p= 2.4×10^{-06})	+ (p= 1.6×10^{-04})	N	N	- (p= 3.0×10^{-06})
SLC6A9 (786)	- (p= 1.4×10^{-05})	- - (p= 4.4×10^{-16})	N	+ + (p=0.0)	N	- (p= 4.0×10^{-04})
SLC7A10 (323)	N	- (p= 7.0×10^{-08})	- (p= 6.7×10^{-05})	N	+ + (p= 0.0)	N
SLC1A4 (779)	- (p= 1.4×10^{-05})	- - (p= 4.4×10^{-16})	N	+ + (p=0.0)	N	- (p= 4.0×10^{-04})
+						

The plus and minus signs represent up- or down- regulation of genes, where + +/- - p= $\leq 1\times10^{-09}$, +/- p= $\geq 1\times10^{-09}$, and N shows a 'normal' distribution, ie. p= ≥ 0.005.

take up glutamine, which undergoes deamination in the mitochondria to form glutamate by the glutaminase enzyme. Glutamate is sunk into the TCA cycle intermediate, 2-oxoglutarate via the glutamate dehydrogenase-catalysed reaction which releases NH_4 and reduces NAD to NADH. In the glial cells, glutamate is used as the substrate and all NH_4 is provided by the glutamate dehydrogenase -catalysed reaction (Figure 3.7). The preferential uptake of each amino acids matches well with the glutamate-glutamine cycle theory.

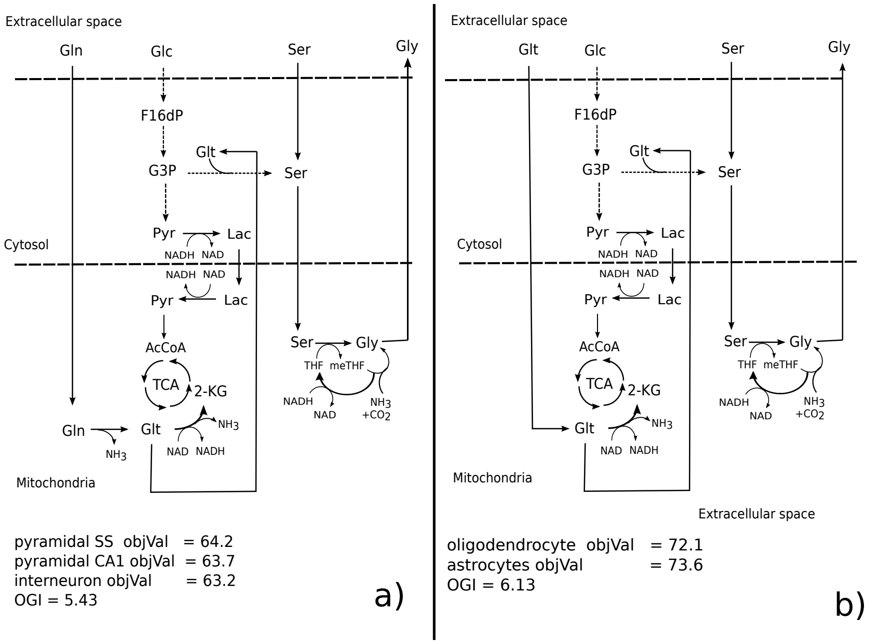


Figure 3.7: **Pathways of glycine synthesis in different cell types** Pathways of glycine production remain the same, with the exception of the N source. Glial cells show a preference for glutamate uptake (b), whereas neuronal cells consume glutamine which is converted to glutamate in the mitochondria (a).

Alternative pathways of glycine synthesis in an interneuron

In the previous solutions seen in Chapter 2, Section 2.3.2, glycine is synthesized equally via the reverse GCS and the forward reaction of SHMT2, where reverse GCS consumes methylene-THF generated in the mitochondria by the forward reaction of SHMT2. The implication of setting the GCS as irreversible on the preferred pathway utilization are assessed in the next results section. The purpose of SHMT cell compartmentalization was also assessed by analysing the effect of the knockout of each reaction on the resulting pathways of glycine production. In each problem, the constraints were set as in equation 3.4, and the flux of the reaction in question was set to zero. reactions described were sequentially restricted to zero in each next problem described as follows.

1. **Knockout of mitochondrial SHMT2** Glycine is synthesized via the cytoplasmic SHMT1 and reverse GCS. Serine is not directly transported across the mitochondrial membrane, rather C atoms cleaved from serine are involved in several folate transformation reactions with the overall effect of generating mitochondrial methylene-THF required for the GCS-catalysed reaction. The net flux of serine and glutamate import remains the same as in Section 3.3.3, where serine import flux is 50% of glycine export. Folates are known to be transporter across mitochondrial membranes by mitochondrial folate transporters which are required for glycine biosynthesis [199].

2. **Setting GCS to irreversible in the glycine-degradation direction** Glycine production via SHMT in both compartments is inactive. Instead, serine is transported into the peroxisome where it is used as a substrate in the serine-pyruvate transaminase-catalysed reaction:

$$1.0 \text{ Serine} + 1.0 \text{ Pyruvate} + 1.0 \text{ H}^+ \rightarrow 1.0 \text{ Alanine} + 1.0 \text{ Hydroxypyruvate}$$

Alanine and glyoxylate produce glycine and pyruvate via the enzyme alanine-glyoxylate aminotransferase. Glyoxylate used in this reaction is formed from hydroxypyruvate in the peroxisome, via a pathway which converts the latter to glycoaldehyde, and subsequently glycolate in the cytoplasm. This pathway leads to the formation of H_2O_2 which is neutralised by the catalase-mediated reaction to O_2 and H_2O. The serine-pyruvate transaminase reaction occurs both in the peroxisome and in the mitochondria, thus upon the knockout of the former the latter switches on. Alanine used in the mitochondrial alanine-glyoxylate aminotransferase-catalysed reaction is synthesized from glutamate, and glyoxylate is generated in the same peroxisomal pathway as described. Knockout of both pathways utilizing alanine causes the cell to change the amino donor. Glyoxylate is produced via the equal peroxisomal pathway however, glycine production occurs in the cytoplasm from glyxoylate and glutamate in the following reaction:

$$1.0\ \text{Glutamate} + 1.0\ \text{Glyoxylate} + 1.0\ \text{H}^+ \rightarrow 1.0\ \text{Glycine} + 1.0\ 2-\text{Oxoglutarate}$$

Across all solutions the import of serine is equal to the export of glycine, at 0.035 μmolg^{-1}min^{-1} . There is no net turnover of NADP in the solution. The expression of the AGXT gene encoding the serine-pyruvate transaminase enzyme is zero across the scRNA-seq, therefore this pathway of glycine production is unlikely to occur in brain cells.

Knockout of the glyoxylate biosynthetic pathway If none of the above pathways can be used and GCS is set as irreversible, glycine is synthesized in the cytoplasm via SHMT1. The methylene-THF produced which can no longer be sunk by the GCS is cycled back to THF with a release of CO_2. Upon inspection of this solution, futile cycles catalysing NAD(P)H conversions were spotted, such as the reaction cycle:

$$\begin{gathered}1.0\ \text{Hydroxyacetone} + 1.0\ \text{NADPH} + 1.0\ \text{O}_2 + 1.0 \\ \text{H}^+ \rightarrow 1.0\ \text{Methylglyoxal} + 1.0\ \text{NADP} + 1.0\ \text{H}_2\text{O}\end{gathered}$$

$$1.0\ \text{Hydroxyacetone} + 1.0\ \text{NADPH} \leftarrow 1.0\ \text{Methylglyoxal} + 1.0\ \text{NADP}$$

Overall, there are no net metabolic transformations that take place in this cycle. Four equivalent cycles were identified involving the conversions of the toxic metabolite methylglyoxal, as well as fatty acid interconversions. Upon the removal of these cycles, glycine production occurs both via the cytoplasmic and mitochondrial SHMT enzymes which contribute equally to the total glycine production (Figure 3.8, a). Methylene-THF produced in the cytoplasm which can no longer be sunk into mitochondrial glycine is used to form methionine. Further pathways of methionine degradation switch on, utilizing S-adenosyl methionine in the pathway of choline formation from ethanolamine. In this pathway, three S-adenosyl methionine molecules are used for the methylation reaction producing three molecules of S-adenosyl-homocysteine (SAM) which is further cleaved to adenosine and homocysteine. The methylene-THF produced in the mitochondria is recycled to cytoplasmic THF via folate transformations and a release of CO_2. Uptake of serine is 0.041 μmolg^{-1}min^{-1} , exceeding glycine export.

Knockout of cytoplasmic SHMT1 Glycine biosynthesis occurs entirely from SHMT2 in the mitochondria. As in the previous result, methylene-THF produced in the mitochondria is recycled back into tetrahydrofolate with a release of the C atom as CO_2. This reaction reduces NADP to NADPH, which is balanced by the oxidation of reduced glutathione (GSH) to glutathione (GSSG) in cysteine transfers reactions (Figure 3.8, b). There is a net formation of H_2O_2 which is neutralised by GSH. Uptake of serine is equal to the export of glycine, at 0.035 μmolg^{-1}min^{-1} .

Knockout of mitochondrial SHMT2 As in the case of equal glycine production from both SHMT enzymes, there is a net production of choline via the methionine degradation pathways. However, there is no transport of glycine into the mitochondria and production occurs entirely by SHMT1, which can be seen in Figure 3.8, c. Apart from methylene-THF sink into 5-methyl-THF, methylene-THF is converted into 5,10-methenyl-THF which is subsequently recycled into tetrahydrofolate required for the SHMT1- catalyzed reaction. Net uptake of serine exceeds glycine export at 0.043 μmolg^{-1}min^{-1} .

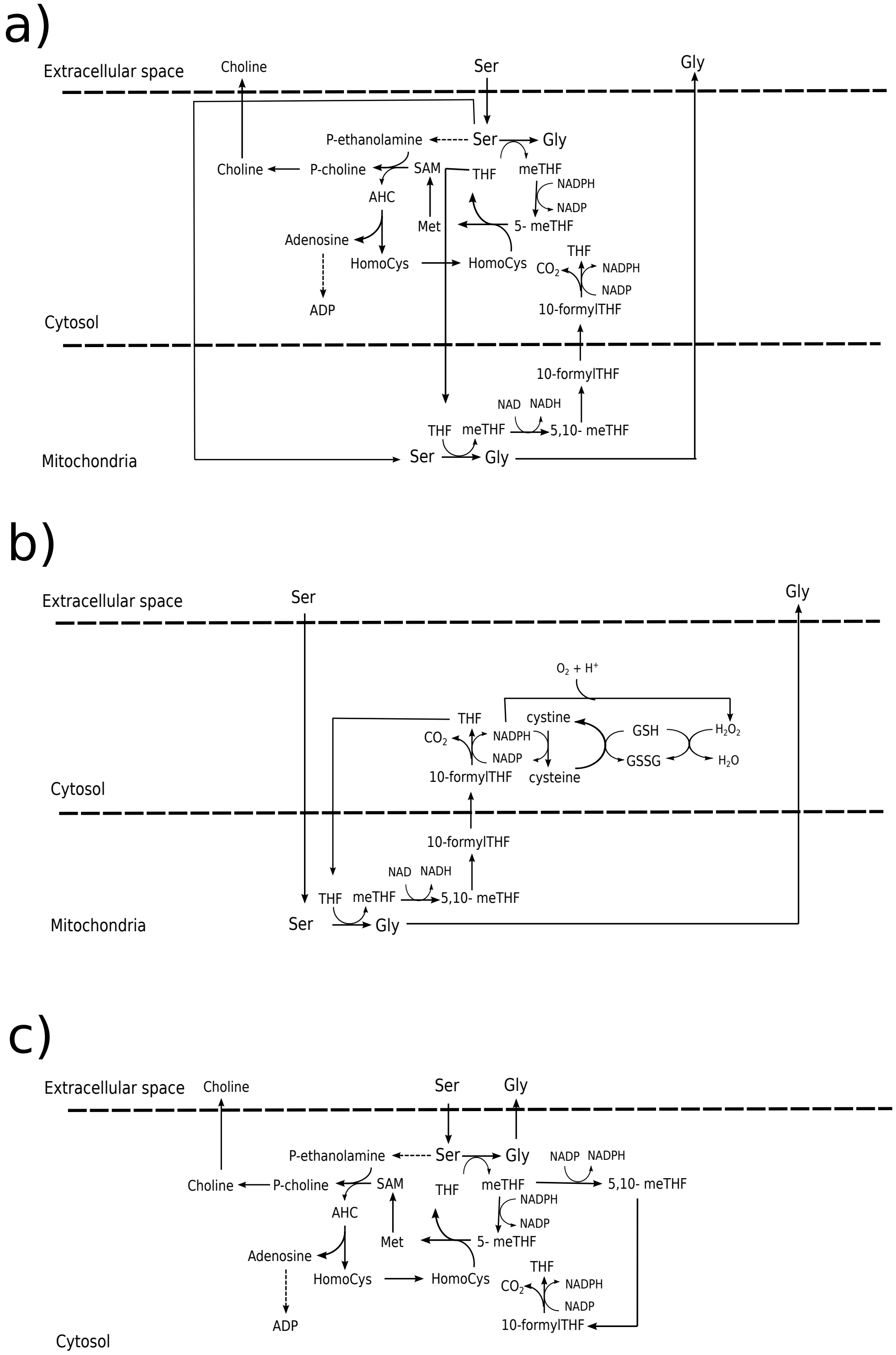


Figure 3.8: **Effect of SHMT knockouts** When the GCS-catalysed reaction is set as irreversible in the glycine degrading direction and NAD(P)H cycling reactions are removed, glycine is formed from both SHMT enzymes equally (a). Glycine formation entirely from SHMT2 results in a stop in methionine formation, and all of the C atoms from methylene-THF are sunk into CO_2 in the cytoplasm (b). Glycine production from SHMT1 only supports methionine formation, as in solution (a).

3.3.4 Amino acid uptake at increasing energy and biomass demands

In this section the preferences for amino acids uptake during cell growth were assessed. In each FBA problem the ATP_d was kept constant at 17.5 μmolg^{-1}min^{-1} and the total biomass ranged from 0.025 to 0.055 μmolg^{-1}min^{-1} . Restrictions on glucose uptake remained the same, at the upper limit of 0.6 μmolg^{-1}min^{-1} , and all fatty acid uptake was restricted to 0.015 μmolg^{-1}min^{-1} .

Preference of substrate in the presence of free amino acid uptake

In the first instance, the uptake of all imported amino acids including glutamate, glutamine, serine and glycine was left unconstrained. Across all solutions the cells saturated glucose uptake and released lactate, with a linear increase in the uptake of glutamate and export of aspartate (data not shown). There is a small increase in the rates of glycine and serine uptake which are used in the biosynthesis of purines and lipids via the reaction pathways seen in Chapter 2, Section 2.3.2.

In the next problem formulation, glutamate uptake was restricted to 0.01 μmolg^{-1}min^{-1} which resulted in increased uptake of serine and isoleucine, and a release of glycine (Figure 3.9, b). As seen in Section 3.3.1 (increases in the demands of ATP_d only), glucose uptake is saturated at all biomass demands. It follows that as more oxygen is consumed in the pathways of amino acid oxidation there is a linear increase in the OGI across solutions (5.58-5.82). Glucose utilization is shunted towards the oxidative PPP thus reducing glycolytic flux, without an effect on the flux via the TCA. Isoleucine is the biggest contributor to the total imported C moles across all solutions (Figure 3.9 a) and undergoes transamination with 2-oxoglutarate to produce glutamate and 2-keto-3-methyl-valerate. Both of these compounds are transported into the mitochondria, where glutamate is used to directly fuel the TCA cycle via its conversion to 2-oxoglutarate by the glutamate dehydrogenase enzyme which releases significant NH_4 .

As biomass demand increases, serine becomes a major source of C atoms via its conversion to the glycolytic intermediate, 2-phosphoglycerate. Firstly, it exchanges an amino group with pyruvate to form hydroxypyruvate and alanine via the serine-pyruvate transaminase reaction. Hydroxypyruvate is next converted to glycerate in a reaction which oxidises NADH to NAD via hydroxypyruvate reductase. Finally, glycerate is converted to the glycolytic intermediate 2-phosphoglycerate via glycerate 2-kinase which uses ATP as the phosphate donor.

Although the flux carried by this pathway as well as reactions utilizing serine and glycine for biomass formation increase, the rates of increase do not correspond with the large increase in the rate of serine uptake. Large increases in the rate of serine import correspond directly to its transport into the mitochondria and cleavage into glycine (Figure 3.9,c). As seen in previous solutions, the total methylene-THF produced in this reaction is sunk into glycine formation via reverse GCS, with both reactions carrying equal flux. Glycine is next transported into the cytoplasm and released. Since the two transport reactions do not carry equal flux and there is no import of glycine, that which is used for biosynthetic reactions is formed completely from serine.

If the large increase of serine uptake is directly correlated with glycine export, and all the methylene-THF formed in the mitochondria is sunk into glycine, what is the purpose of this pathway? Analysis of the solution shows that the reverse GCS acts as primary sink for NH_4 formed as a byproduct in the glutamate to 2-oxoglutarate conversion reaction. At higher biomass demands, more than 90% of N is neutralised by this pathway (Figure 3.9, d), thus the primary purpose of the serine/glycine transformation is to balance the excess N formed during isoleucine degradation.

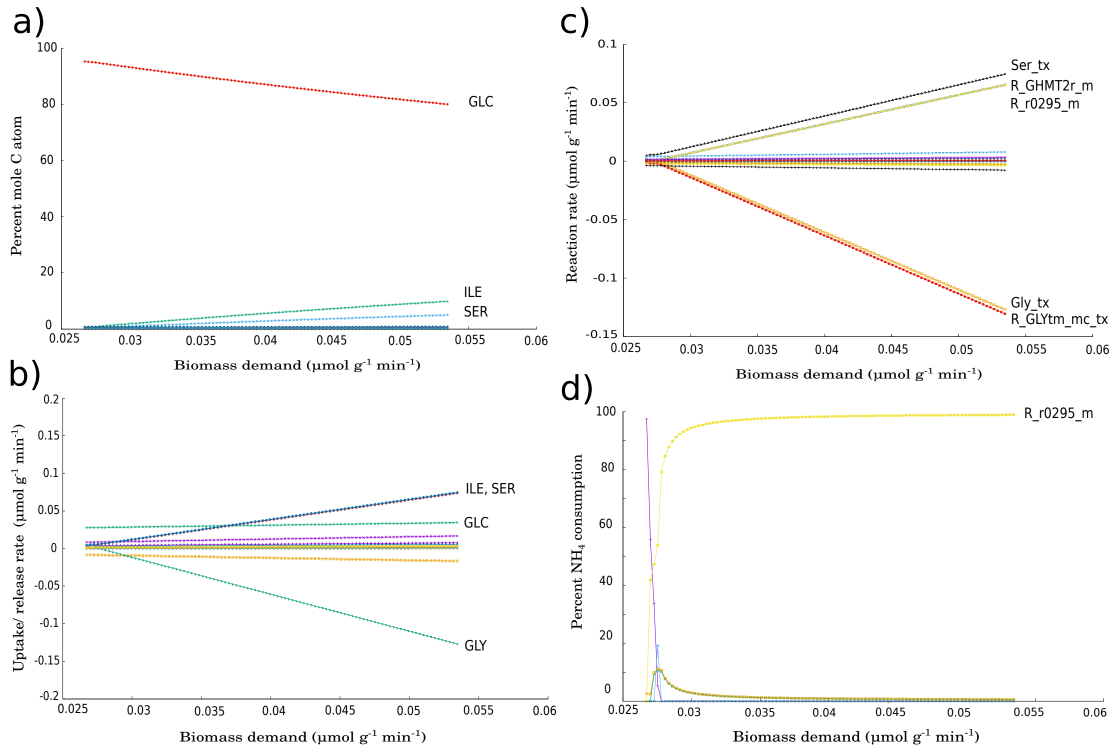


Figure 3.9: **Cell uptake of isoleucine and serine for biomass production** In the absence of additional glucose or glutamate, the cell begins to use isoleucine and serine (b). Isoleucine, along with serine are the major C donor across all solutions (a). The rate of serine uptake corresponds to equivalent increases in the rate of its transport to the mitochondria, conversion to glycine (reacID: R_GHMT2r_m), the reverse reaction of GCS (reacID: R_r0295_m), and subsequent glycine release (reacID: Gly_tx), (c). This reaction cycle may be acting as a primary sink for NH_4 produced in glutamate degradation pathways, as reverse GCS is the reaction with highest N consumption across all solutions (d).

In the next scan, both glutamate and isoleucine were constrained to 0.1 μmolg^{-1}min^{-1} . In this solution, the cell begins to uptake glutamine and continues to consume large amounts of serine and export glycine (Figure 3.10, b). The conversions of serine to hydroxypyruvate increase proportionally to serine uptake (Figure 3.10, c), indicating that this pathway becomes a major contributor to energy production as can be seen in Figure 3.10, a. Glutamine also contributes to biomass production via its deamination to glutamate in the mitochondria which – as previously – is sunk into 2-oxoglutarate with a net production of NH$_4$. Reverse GCS remains the major N sink (Figure 3.10, d) produced in this pathway.

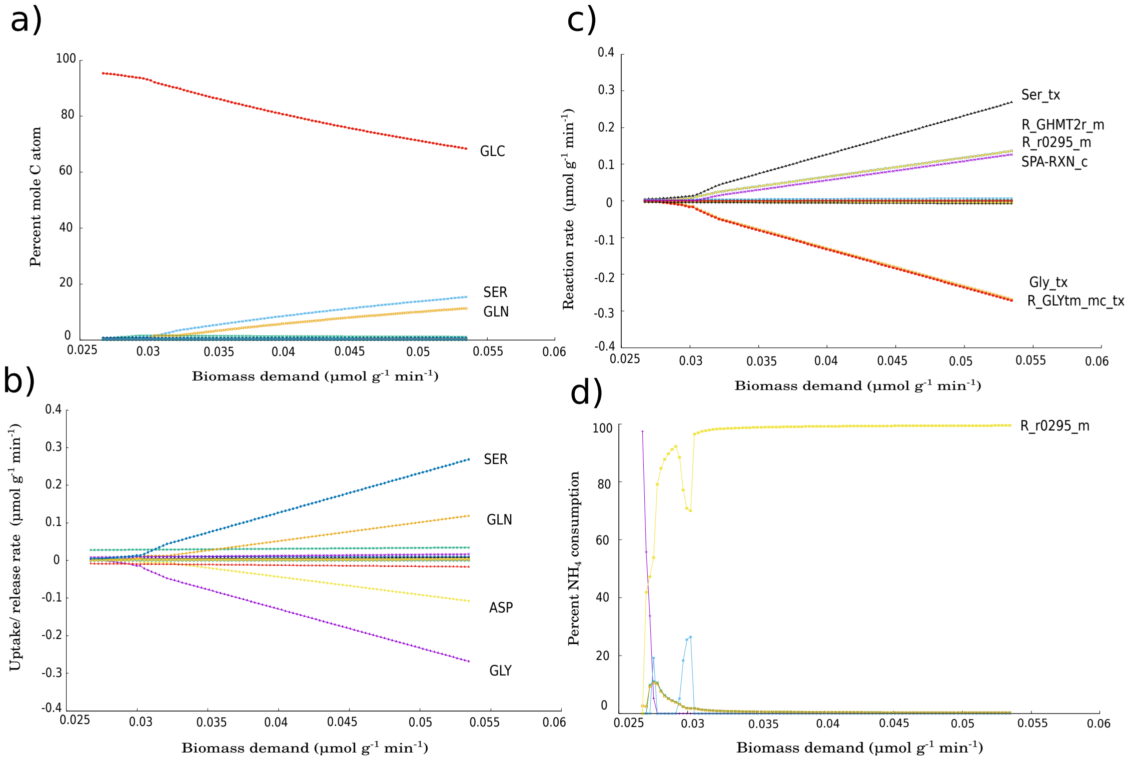


Figure 3.10: **Cell uptake of serine and glutamine for biomass production** Upon restriction of isoleucine import, the cell begins to uptake large amounts of serine as well as glutamine (b). The former becomes a major C donor (a) after glucose via the serine-pyruvate aminotransferase reaction in the cytoplasm (reacID: SPA-RXN_c) which leads to formation of 2-phosphoglycerate (c). As in the previous solution, reverse GCS remains the major consumer of NH$_4$ produced by glutamate degradation (d).

Production of ATP from serine and glycine

Biomass turnover requires an increase in cellular ATP consumption. The ability of each amino acid to support energy production only was tested in the model. In each problem formulation, ATP demand was incremented from 1-20 μmolg^{-1}min^{-1} in the presence of a zero flux constraint on all transport reactions except for O_2 , CO_2 , H_2O , H^+ and either serine or glycine.
Out of the two amino acids only serine was able to support an ATP demand increase. Serine was able to support ATP production in the mitochondria via a number of pathways as follows:

1. Serine conversion to cysteine via the enzyme cystathionine-β-synthase, with the latter converted to pyruvate in the cytoplasm via cystathionine-γ-lyase.

2. Serine conversion to cystathione via the enzyme cystathionine-β-synthetase, which is next sunk into 2-oxoglutarate in the cytoplasm.

3. Serine conversion to aminocrylate via the enzyme serine-dehydratase which is next converted to pyruvate in the cytoplasm.

4. Serine conversion to 3-phosphoglycerate via the reversal of its pathway of biosynthesis from glucose.

5. Cleavage of serine into glycine in the mitochondria, leading to the production of NADH for use in the ETC (Figure 3.11)

.

Pathway 5, ATP production from serine via cleavage into glycine leading to the generation of NADH in the mitochondria can be seen in Figure 3.11. Serine uptake is directly correlated with glycine release (Figure 3.12, a). On the other hand, when the only substrate for energy production was glycine the cell could not make ATP at any demand. This shows that these two amino acids are not freely interconvertible.

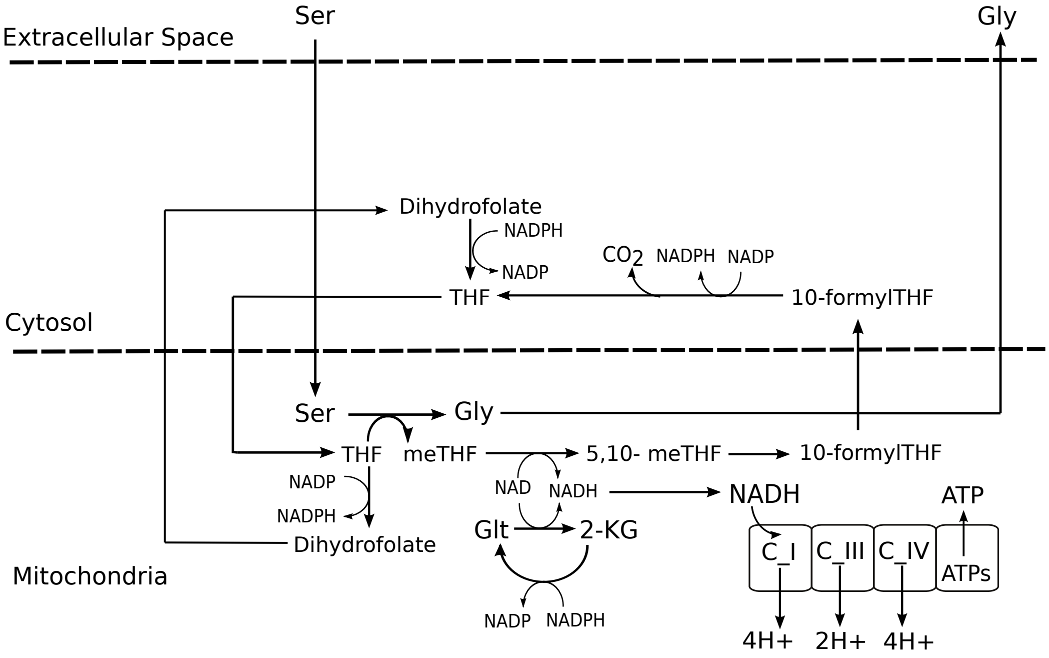


Figure 3.11: **Serine to glycine conversion supports ATP formation via the ETC** Serine can support ATP formation by inducing NAD reduction via the reaction converting methylene-THF to 5,10-methenyl-THF . The methylene-THF is generated by conversion of serine to glycine via mitochondrial SHMT2. This pathway requires the cycling of folate groups across the mitochondrial membrane.

Interconversion of serine and glycine

Reversibility of the SHMT enzymes would suggest that cells should have a capability to freely convert serine and glycine. This is not the case given previous results which show that only serine is capable of supporting energy increases. In this set of analyses, the export rate of each amino acid was varied from 0.001- 0.35 μmolg^{-1}min^{-1} in the background of constraints on amino acids uptake and an unconstrained uptake of the other amino acid.

In the absence of glucose uptake, the model is capable of producing unlimited amounts of glycine from serine only (Figure 3.12, c). In reverse, production of serine from glycine is only feasible in the presence of glucose (Figure 3.12, b) or formate (Figure 3.12, d). In the former, glycine is used to produce glutamate in the cytoplasm which acts as the amino donor in the pathways of serine

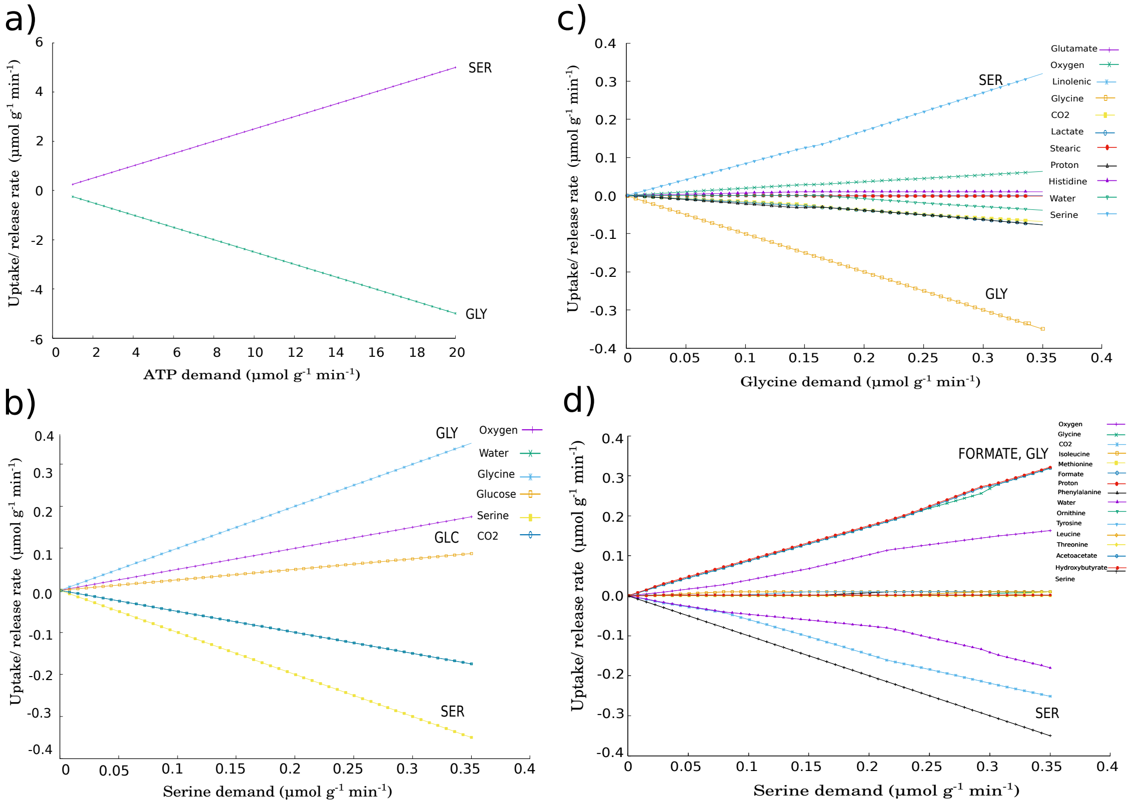


Figure 3.12: **Glycine alone cannot support serine synthesis** Serine can support ATP and glycine production without the use of any other substrate (a, c). On the other hand, production of serine from glycine is not possible unless the uptake of glucose (b) or formate (d) is allowed in the model.

production from glucose. When the model is allowed formate uptake in the absence of glucose, serine production occurs via both the glycolytic pathway, as well as the reverse flux of both SHMT enzymes. These contribute to serine biosynthesis in different ratios, with mitochondrial SHMT2 producing ≈ 65%, cytoplasmic SHMT1 ≈ 25% and the glycolytic pathway ≈ 8% of the total serine biosynthesis. These ratios remain the same across all solutions.

Ability of serine and glycine to support increasing biomass demands

It is apparent that serine is one of the preferred amino acids used during biomass increase along with glutamate, isoleucine and glutamine. Secondly, production of glycine in the mitochondria may act as a N sink for a cell utilizing glutamate to feed the TCA cycle. However, it remains unclear whether serine may act as a sole substrate for biomass production.

In the next set of FBA problems the uptake of all amino acids was constrained to 0.01 μmolg^{-1}min^{-1} . Serine and glycine uptake was constrained to 1 μmolg^{-1}min^{-1} , in the presence of constraints on glucose and fatty acids uptake as described previously. Biomass was increased from 0.03 to 0.14 μmolg^{-1}min^{-1} . The uptake of amino acids was allowed to a limit because the complete knockout of amino acids import renders the solution infeasible.

Serine alone can sustain biomass increases up to $\approx$ 0.12 μmolg^{-1}min^{-1} (Figure 3.13 a) by donating its C atoms to 2-phosphoglycerate in the cytoplasm as described in the previous section. Knockout of the hydroxypyruvate pathway in the presence of the same restrictions leads to the formation of the intermediates of glycolysis and ETC as described in the presence of an ATP demand only (Section 3.3.4). Across all solutions, glycine is exported proportionally to serine uptake and thus remains the main sink for N atoms generated in these reactions. Setting the GCS to unidirectional in the glycine degradation direction stops glycine export, and the cell begins to release asparagine which becomes the N sink.

In the presence of glycine and a restriction on serine uptake, biomass increase can only be supported up to $\approx$ 0.056 μmolg^{-1}min^{-1} . Glycine begins to be imported only at very high demands which causes an increase in H^{+} import and arginine export. Across all solutions, C atoms are acquired by the oxidation of amino acids other than glycine up to their saturation of uptake (Figure 3.13).

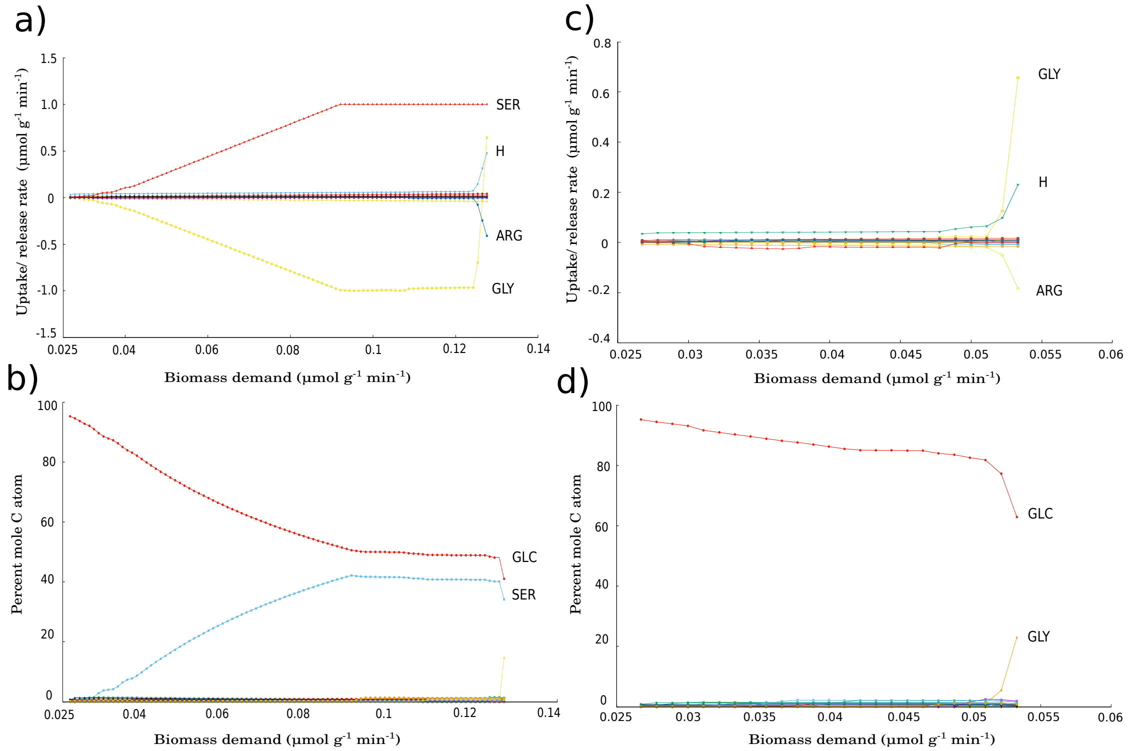


Figure 3.13: **Ability of serine and glycine to sustain biomass increase** Serine can support growth up to $\approx 0.13\ \mu\text{molg}^{-1}\text{min}^{-1}$ (a) and is the second C donor after glucose (b). In the presence of glycine, solution is feasible only up to a total of $\approx 0.057\ \mu\text{molg}^{-1}\text{min}^{-1}$ total biomass flux. Glycine begins to be imported only at the very high biomass demands, which results in a large increase of H^{+}import (c).

3.4 Discussion

Energy production pathways in different cell types

Expression of all glycolytic genes is widespread across the dataset and upregulated in neuronal cell types, as can be seen in Table 3.2. This includes the gene encoding pyruvate kinase (PKM), which catalyzes the dephoshorylation of phosphoenolpyruvate to pyruvate. Differential splicing of PKM in human results in two isoforms, PKM1 and PKM2 where the latter has been associated with higher expression in human cancer cells [200]. Since PKM1 is constitutively active, but the activation of PKM2 relies on the binding of allosteric activators [201] PKM2 limits glucose oxidation and favours biosynthetic-diverging pathways of glycolysis. However, the enhancement of PKM1 has also been shown to favour tumour growth in tumour mouse models, thus evidence towards tumour-enhancing overexpression of both isozymes exist. Analysis of transcriptomic data from different types of purified brain cells by Zhang et al. [202] showed that PKM1 is mostly expressed in neurons, and PKM2 in astrocytes which have a better ability to modulate their glycolytic flux, in comparison to neurons. Although the products of the differential splicing cannot be distinguished in the scRNA-seq, the expression of PKM is widespread across all cell types showing no specific pattern of expression.

Entry of pyruvate into the TCA cycle is regulated by pyruvate dehydrogenase, which catalyzes the formation of acetyl-CoA from pyruvate in the mitochondria. This enzyme is regulated via phosphorylation by the pyruvate dehydrogenase kinase enzyme which exists in four isoenzymes, encoded by the genes PDK1-4. Activities of each isoenzyme have been determined *in vitro* showing that the isoenzymes vary in their phosphorylation rates at three serine sites [203]. Overall, PDK2 has the highest rate of phosphorylation of serine site 1, whereas PDK3 has the highest rate of phosphorylating serine site 2 and serine site 3 is only phosphorylated by PDK1 [203]. In the same study by Zhang et al. [202] it was found that the expression of PDK4 is higher in astrocytes than in neurons, which may act to prevent pyruvate entry into the TCA cycle and promote lactate formation by lactate dehydrogenase. In the current study it was found that PDK4 expression is indeed higher in astrocytes, but the expression of PDK1-3 genes is higher overall and more prevalent in pyramidal CA I thus neuronal, rather than glial cell types.

Lactate dehydrogenase, which interconverts pyruvate and lactate is encoded in humans by four genes, LDHC. LDHA and LDHB encode the subunits which compose the heteromeric enzyme defined by random combinations of the two subunits, however the affinity of LDHA to pyruvate is higher than lactate and *vice versa*, which, along with the concentrations of lactate and pyruvate present, determine in which direction the reaction will proceed [204]. The expression of LDHA and LDHB in the scRNA-seq is widespread across the whole dataset with both genes expressed in slightly higher numbers in neuronal cell types, as seen in Table 3.2. Overall, these expression patters suggest that glycolysis is present in all cell types, and perhaps higher in neuronal cells in comparison to glial.

Methylglyoxal is a product of spontaneous breakdown of dihydroxyacetone phosphate (DHAP), an intermediate of glycolysis and is associated with numerous health-related conditions [205]. Magistretti and Allaman [25] reasoned that higher expression of GLO genes in astrocytes in comparison to neurons better equips these cells to a high glycolytic activity which leads to methylglyoxal accumulation. In this study, the GLO1 gene, which encodes the components of the glyoxylate system, lactoylglutathione lyase and hydroxyacylglutathione hydrolase is one of the most widely expressed genes analysed, suggesting the ubiquitous need of methyglyoxal detoxification in brain cells which is not cell-specific.

As seen in Figure 3.3, the expression of genes encoding the enzymes of the oxidative PPP is lower than glycolytic genes overall. There is no significant correlation between the expression patterns of these genes. Genes of the non-oxidative PPP are more widely expressed, but there is not a significant difference in the expression levels across cells. This further supports the notion that glucose in all cells is mostly processed via glycolysis, rather than diverging biosynthetic pathways such as the PPP.

Further to the expression patterns, ATP production in different cell types was analysed in an FBA problem ATP scan (Section 3.3.1). The pathways utilized by each cell type, neuronal and glial were the same with a difference of lactate efflux. Glial cells consumed only the amount of glucose required for the production of ATP, at all demands. Although the conversion of pyruvate to lactate occurred in the cytosol, all of the lactate was oxidised in the mitochondria. Suprisingly, neuronal cells saturated glucose uptake and released lactate at all energy demands, with the proportion of lactate release decreasing at increasing energy demands. This is in direct contrast with the ANLS hypothesis, supported by previous studies such as that by Walz and Mukerji [206] in which higher lactate release rates were seen from cultured astrocytes in comparison to neurons. In their study, the rate of neuronal release reached astrocytic release rates when uncoupling of oxidative phosphorylation in the mitochondria was achieved by the addition of dinitrophenol, suggesting that the difference accounts for the varying rates of glycolysis and oxidative phosphorylation in the two cell types. Higher rates of lactate release by astrocytes were also reported by Itoh et al. [207] *in vitro*. However, stimulation of PDH via the administration of dichloroacetate to conscious rats, reducing the rate of astroglial lactate release, did not affect normal brain function, suggesting that astroglial lactate release and cycling is not essential for neuronal function. In addition, it has been shown by Díaz-García, C. et al. [208] that during stimulation neuronal cells uptake glucose and process it via glycolysis, increasing the NADH ratio without the uptake of astrocytic lactate. Thus, although these cells may have different metabolic capabilities, neuronal cells are not dependent on astrocytic lactate release.

So what could account for neuronal lactate release? Sobieski et al. [209] found that rat neuron-astrocyte co-cultures can survive long-term glucose deprivation and sustain neurotransmission. Cell starvation induced large lactate releases which remained in the media, supporting cell survival. The origin of lactate was likely astrocytic, although in the same study glucose-deprived neurons maintained neurotransmission whether grown in the presence of astrocytes or not. The authors concluded that neuronal cells may use lactate release as a way to store energy. Given that neuronal cells have a high ATP demand, but brain glycogen reservoir is restricted to glial cells [210], lactate efflux could provide neurons with an energy source in the surrounding fluid. Whether this happens *in vivo*, when the cell is not subjected to stress conditions which may cause it to rewire its metabolism is not certain. It can be imagined that the experimental treatment of the tissue prior to the collection of mRNA expression values in the scRNA-seq dataset used in the current study could lead to stress-induced changes in genetic expression, leading to the results seen in this study.

In the next model analyses (Section 3.3.1), both glucose and lactate were allowed to be uptaken and the ATP demand ranged from 1-30 μmolg^{-1}min^{-1} . All cells, apart from astrocytes preferred to consume glucose up to its saturation at ATP$_d \approx 20$ μmolg^{-1}min^{-1} . This is in direct contrast with substrate preference of primary cultures of astrocytes and neurons reported by Bouzier-Sore et al. [211], although, it has been previously shown that both cell types can take up glucose and lactate as substrates [207]. Overall, the high rate of lactate release and preferential use of lactate by astrocytes is in direct contrast to the ANLS hypothesis

In Section 3.3.2, ATP production was used as the objective function to cluster cells based on the solutions they produce. Overall, the heterogeneity emerging from the data was less apparent than that resulting from clustering based on genetic expression. Two main groups of cell clusters appeared, neuronal and glial cell types with the former releasing lactate. Out of the seven main classes identified in the original publication, only microglia were separated in a distant node. Microglial cells utilized both ketobutyrate and glucose for the production of acetyl-CoA in the mitochondria, which was not seen in any other solution set. Lastly, a separate cluster of 65 cells composed of mostly oligodendrocytes emerged, with a preference for GTP production via the TCA cycle reaction catalyzed by succinate CoA synthase.

It seems obvious that the main pathways of energy production in a cell will remain similar across cell types, given that cellular need of energy production is fundamental and constant across the lifespan of the cells. In his book ‘What is life?’, Schrödinger [212] wrote that stochastic behaviour within a system is a disadvantage for its existence in an ordered- at least on the macro scale- universe. For example, the Brownian motion of a particle would make it difficult for it to plan its journey which is trivial for bigger organisms like humans, thus defining a scale to living organisms. As energy production is a fundamental need of cells to sustain life, stochasticity would make it less robust to external factors, such as gene expression. In fact, the stability of the ETC machinery as a means of energy production is proved by its emergence in bacteria about 2×10^9 years ago and its use in mammals still today [213]. Rather, cellular heterogeneity is a result of the more specialised cellular functions contained in the total genetic expression of the cell. It is therefore surmised that clustering based on more specific objective functions will produce a greater cell heterogeneity. Nevertheless, the emergence of two small populations of cells in this study suggest that scRNA-seq could be used further in metabolic modelling for the identification of rare phenotypes. In addition, it may prove useful in the elucidation of metabolic interactions between cell types, for example energy substrate or neurotransmitter cycling.

Glycine biosynthesis pathways in different cell types

Expression of genes related to glycine and serine metabolism and transport can be seen in Table 3.3. Amongst metabolic enzyme genes, there was a significantly uneven distribution of genes encoding enzymes of serine biosynthesis from the glycolytic intermediate 3-phosphoglycerate. Phosphoglycerate dehydrogenase (PHGDH) catalyses the conversion of 3-phosphoglycerate to phosphohydroxypyruvate which is the first step of glycolysis diverging pathway of serine biosynthesis from glycine. Expression of PHGDH is significantly low in all neuronal cell types, and significantly high in oligodendrocytes and astrocytes. Overexpression of PHGDH in astrocytes has previously been found by Shimizu et al. [214] who found its low expression in cultured neurons and high expression in astrocytes. In addition, it was later shown by the same group that astrocyte-derived serine enhances the survival and growth of cultured hippocampal neurons [215]. This lead the authors to conclude that serine is produced in the astrocytes and shuttled across to neurons. The expression patterns in the scRNA-seq are in line with these findings.

Phosphoserine aminotransferase (PSAT1) and phosphoserine phosphatase (PSPH) catalyze the next two steps of this pathways with the final production of serine, and the genetic expression of these genes is significantly high in oligodendrocytes. These cells are classically attributed the role of providing axonal insulation, and given serine is a phospholipids precursor high demands of serine would be expected.

Expression of cytoplasmic SHMT1 is much lower than its mitochondrial counterpart, SHMT2 suggesting a bigger contribution of the latter to the total glycine pool. Another possible source of glycine is from glyoxylate, which require the peroxisomal or mitochondrial enzyme serine-pyruvate aminotransferase, encoded by AGXT and AGXT2, respectively. These genes were not present in the scRNA-seq dataset used, implying this pathway is inactive in the brain. According to the Allen Brain Map, mean transcriptome expression across all cell types is zero. This is in line with findings by Shank et al. [39], who suggested that glycine is unlikely to be derived from glyoxylate in the brain.

Genes encoding dipeptidase enzymes prolidase (PEPD) and carnosinase (CNDP2) are highly expressed in the dataset, with numbers of cells positive in the same range as for the PHGDH gene. Prolidase was significantly overexpressed in pyramidal CA I cells. It preferentially hydrolyzes dipeptides containing proline or hydroxyproline, and is important in the recycling of collagen in cells [216] of which glycine is a major constituent. Collagen is known to be the component of structural tissues in the body, but a recent paper by Seppänen, A. et al. [217] describing immunohistochemical staining of the human brain showed high levels of collagen expression in neuronal, but not glial cells. The function of collagen in the brain is not known, however its presence requring production and consumption could represent a substantial source and sink for glycine in neurons.

The dipeptidase encoded by CNDP2 is significantly overexpressed in interneurons. It has been previously been reported that this protein has a preference for hydrolyzation of cysteine-glycine bonds *in vitro*, and is involved in the degradation of glutathione [218]. High rates of oxidative phosphorylation in neuronal cells generates oxidative stress, rendering these cells in need of a high requirement for ROS scavenging mechanisms, thus GSH turnover which will have an effect on the cellular glycine pools.

The expression of GCSH, encoding the glycine cleavage system H protein, is widespread in the dataset and nearly triple that of SHMT2. Expression is consistent across the dataset, with no significant over or under expression in any specific cell type which is in contrast with findings by Sato et al. [47]. The number of GCSH-positive cells in comparison to SHMT2-positive cells could imply that glycine degradation is more widespread to its biosynthesis. This would support the notion that glycine biosynthesis occurs in one cell type and is shuttled to others. On the other hand, if GCS is reversible *in vivo*, it could provide a source of glycine although subject to the availability of methyl groups produced in cellular biosynthetic pathways.

Lastly, the expression of serine/glycine transporters is non-normal across the dataset. SLC6A9 endodes the sodium-dependant glycine transporter 1 (GlyT1) which modulates glycine concentrations in glycinergic as well as glutamatergic NMDA synapses. It readily uptakes glycine but is also capable of glycine release which is a function of the membrane voltage [219]. SLC1A4 encodes the Asct-1 transporter which has been shown to efficiently uptake serine *in vitro* [220]. The expression of these transporters was significantly high in oligodendrocytes, low in pyramidal CA I cells and normal across all other cell types. The high expression of GlyT1 in oligodendrocytes is suprising, since these cells are usually attributed the role of producing myelin which wraps around the neuronal axon, providing insulation. However, novel evidence shows that the roles of these cells in the CNS may extend further, for example in providing trophic support to neuronal axons [221]. It could be possible that oligodendrocytes, along with astrocytes are the main locations of serine and glycine biosynthesis which are shuttled into the neuronal cells, where it is required for neurotransmission. In addition, given the close proximity of the oligodendrocyte to the neuronal axon, these cells may have a function in the reuptake of glycine during neurotransmission, particularly in the uptake of extra-synaptic neurotransmitter spill-over.

It is important to add that, since the number of gene copies at any given moment is related to the stability of the protein, it is difficult to draw conclusions regarding the actual amounts of functional protein within cells based on RNA expression only, and without the knowledge of the half-lives of these proteins. Further evidence would ideally be supplemented by proteomics data which quantifies protein expression.

The difference in glycine production pathways across cell types was assessed via FBA in Section 3.3.3. The presence of transporters for glutamate, glutamine and serine, as well as the differential expression of PHGDH in different cell types imply that serine and glutamate required for glycine biosynthesis are likely recycled across different cell types as is the case in the glutamate-glutamine shuttle. Under these conditions, glycine production in the background of an ATP demand occurs from serine and glutamate in glial cell types, and serine and glutamine in neuronal cell types which is in line with the preferred amino acid uptake by each cell type proposed in the glutamine-glutamate shuttle hypothesis [222]. As in the pathways of glycine biosynthesis in Chapter 2, glycine is synthesized entirely from serine via the mitochondrial SHMT2 and the reverse GCS, which utilizes the methyl group released in that reaction.

In Section 3.3.3, the effect of knocking out the various components of serine/glycine biosynthetic pathways was investigated. Loss of the mitochondrial SHMT2 lead to the utilization of its cytoplasmic counterpart, along with the reverse GCS formation. The metabolic flexibility has previously been seen in *in vitro* studies in human cells, where SHMT2-knockout cell lines were capable of supporting methyl group production in the cytoplasm, the source of which is mostly mitochondrial before the knockout [223]. When the GCS was set as unidirectional in the glycine-catabolising direction, both SHMT enzymes became inactive altogether and glycine was synthesized from serine via the glyoxylate pathway. This pathway requires the enzyme serine-pyruvate aminotransferase and since genes encoding the enzyme were absent from the scRNA-seq this pathway is unlikely. Implicit in the LP formulation is the assumption of a metabolic Occam's razor, which is implemented by describing the problem as one of flux minimization. Via the glyoxylate pathway, serine utilization is higher than that produced in solutions allowing GCS reversibility. However, this pathway does not require NADP and its reduced equivalent, and does not produce methylated groups that need to be sunk under steady state condition. One explanation for the appearance of this pathway in the solution is due to its disconnection from these requirements.

Glyoxylate production causes the generation of H_2O_2, has been shown to inhibit glucose oxidation in rat brain slices *in vitro* [224] and proved toxic when applied to rats *in vivo* [225]. Furthermore, significant glycine production from glyoxylate has been shown only in the human liver so far [41], thus this pathway is unlikely to occur in the brain.

Knockout of the glyoxylate pathway in the absence of NADPH cycling reactions and a unidirectional GCS causes the cell to utilize both SHMT1 and SHMT2 for glycine production. Serine uptake exceeds glycine export, and methionine biosynthetic pathways switch on, leading to the release of choline and production of ADP. Mitochondrial SHMT2 knockout produces a similar solution in which glycine biosynthesis causes a switch in biosynthetic pathways. According to Meléndez-Hevia and de Paz-Lugo [46], the stoichiometry of SHMT-catalysed reactions enforces a metabolic inflexibility in the system, where glycine formation causes an obligatory methylene-THF production posing an inability of the system to regulate each independently, assuming that GCS is irreversible. This is what is seen in the model when glycine production occurs from cytoplasmic SHMT1, in the absence of NADPH cycling reactions. It is possible that this is a result not only of the GCS irreversibility constraint, but also of the requirement for NADPH balance. On the other hand, knockout of the cytoplasmic SHMT1 results in a solution in which glycine biosynthesis occurs entirely by the mitochondrial SHMT2, without the presence of other biosynthetic pathways. NADPH generated by cycling is instead utilized for the reduction of GSSG to GSH, necessary for oxidative stress protection. The total serine import is equal to that of glycine export, and lower than that in the presence of cytoplasmic SHMT1. The methylene-THF produced is transported into the cytoplasm in the form of 10-formyl-tetrahydrofolate, which is broken down to tetrahydrofolate and CO_2. Since this pathway requires crossing of folates across the mitochondrial membrane, it would be dependent on the rates of folate transporters in a cell- according to Wagner [226], the transport of folates across mitochondrial membranes is slower than that of serine and glycine which may impose a constraint on the rate of glycine biosynthesis via this pathway.

As surmised in the Discussion section of Chapter 2, modularization of metabolism gives the system the advantage of flexibility. Glycine formation via both reversal of the GCS and forward SHMT2 proves as the most ‘economical’ pathway, since serine consumption equals half of the glycine production flux and both glycine C atoms are derived wholly from serine. It is also advantageous as it does not require the additional need for a methyl sink, which is also the case for glycine biosynthesis entirely via mitochondrial SHMT2. The latter is less preferred, because it does not allow for the incorporation of two C atoms from one serine molecule into glycine. The additional advantage is provided by the fact that utilization of mitochondrial SHMT2 leads to the reduction of NAD to NADH, linking glycine biosynthesis to energy production in the mitochondria. A link between SHMT2 and energy production was seen in cultured human cells knockout of SHMT2, which caused a decrease in the TCA cycle intermediates which was not seen on SHMT1 knockout [223]. On the other hand, utilization of cytoplasmic SHMT1 exerts a pressure on the system which leads to methionine formation, and higher serine consumption. If glycine, and so methylene-THF production is less efficiently achieved by the cytoplasmic SHMT1, it is possible that depletion of TCA intermediates in their study was caused by an increase in the channeling of glucose towards serine biosynthesis.

Serine/glycine metabolic pathways and biomass turnover

The importance of the metabolic pathways involving serine and glycine for one-carbon metabolism is well established. A previous metabolomics study performed in a panel of cancer cell lines showed that an increase in glycine uptake and expression of biosynthetic genes correlated positively with cell proliferation, and generation of methylene-THF from glycine [227]. Furthermore, increased glycine concentrations are positively associated with the presence and stage of human gliomas [5, 6]. In

Section 3.3.4, the preference for amino acids uptake, and the abilities of serine or glycine to support increased biomass and ATP demands were investigated. Out of all model uptake amino acids, glutamate, glutamine, isoleucine and serine were preferred as substrates when glucose uptake is limited. Biomass increase requires ATP, and in a set of experiments on cultured neurons Divakaruni et al. [228] showed that inhibition of mitochondrial MCTs, thus blocking pyruvate entry does not decrease cell viability. Glutamate abundance was decreased, without a decrease in the levels of TCA cycle intermediates showing that glutamate is readily oxidised to support mitochondrial energy production. In addition, the ability of glutamine to sustain bioenergetic cellular needs and cancer cell proliferation has been previously well established [229, 230] supporting the use of glutamate and glutamine in the model.

The upregulation of genes related to serine and glycine metabolic pathways in a subset of cancer cell lines including gliomas [7–10] is suggestive of the utilization of these amino acids to support cellular growth demands. Uptake of serine and corresponding glycine efflux has been shown in cultured colon cancer cell lines HTC116 and RKO by Labuschagne et al. [54], and later in neurosphere-forming glioma cell lines by Kim et al. [231], in line with the model results. In the model, the rate of serine uptake increase was proportional to its conversion to glycine in the mitochondria, generating methylene-THF required for biosynthetic reactions. Serine was also able to support a demand in ATP and biomass increases via a number of pathways, the feasibility of which are described as follows:

1. Serine conversion to the glycolytic intermediate 2-phosphoglycerate requires the enzyme serine pyruvate aminotransferase, encoded by AGXT1 enzyme which, as previously discussed, is unlikely to be present in the brain.

2. Production of pyruvate from cysteine derived from serine requires the enzyme cystathionine-β-synthase which is a component of the transsulfuration pathways involved in the control of cellular methylation (via S-adenosyl methionine) and antioxidant status (via GSH production) [232]. Cystathionine-β-synthase in encoded by the CBS gene, which was found in 217 cells in the scRNA-seq dataset, showing a low expression of this enzyme in the brain. Knock-out of the PKM genes in a pancreatic cancer cell line has previously been shown that cysteine can maintain cell viability by its catabolism to pyruvate in the cytosol [233], although it is unclear whether the upregulation of this pathway could be responsible for energy production in the brain.

3. Another alternative pathway of serine contribution to energy production was its conversion to pyruvate via the enzyme serine dehydrytase, encoded by the enzyme SDS which was found in only 3 cells in the scRNA-seq dataset. Indeed, serine dehydrytase has been mostly found in the human liver where it exhibits low activity [234], thus this pathway is unlikely to be able to sustain energy production in the brain.

4. Production of the glycolytic intermediate 3-phosphoglycerate from serine, via the reversal of its phosphorylated pathway of synthesis is unlikely to occur since the reversal of the serine phosphatase, catalysing the dephosphorylation of 3-phosphoserine to serine is highly energetically unfavourable, as reported by Lund et al. [235].

5. Finally, ATP production via the catabolism of serine to glycine which leads to a transfer of electrons from methylene-THF to NAD, producing 5,10-methenyl-THF and NADH as seen in the model has also been shown experimentally by Yang et al. [236] via deuterium labelling of the H atom of serine in normal and cancerous cell lines. In their study, the fractional contribution of serine to NADH was increased in hypoxia-induced states, thus serine may act as a source of ATP production in a context-dependent manner.

Conversion of serine to glycine is a main producer of methylene-THF necessary for biosynthetic pathways. The model results also suggest that this reaction may play a role as a nitrogen sink. Silencing of the GCS in cell lines expressing high levels of SHMT2 was shown by Kim et al. [231] to be detrimental to cell proliferation. In addition, the authors found a correlation between the level of SHMT2 expression and sensitivity to GCS knockouts, suggesting that in those cell lines the overexpression of both enzymes is required for the high proliferation. If the GCS is a reversible process, the model results suggest that apart from methylene-THF generation, GCS may play a role as a nitrogen sink in a cell oxidising amino acids for its biosynthetic needs. Ammonia must be exported from cells and excreted in the urea cycle, or otherwise deposited into a reservoir of amino acids, for example cytosolic glutamine which is a major sink of free NH_4 in the human brain [237]. As pointed out by Spinelli et al. [238], the question of how tumour cells which increase their rate of amino acids utilization manage the surplus ammonia locally remains underdetermined. Since the vasculature of tumours is less organised than that of normal tissue, the bulk usually contains hypoxic areas which is less well cleared from the excess nitrogen which may disrupt the 'normal' pathways of nitrogen balancing in the cell. Kim et al. [231] described that that SHMT2 suppression had a pronounced effect in glioma cell lines under hypoxic conditions, which was not the case under normal oxygen tension. The authors suggest that the serine/glycine dependence of these tumour may therefore be context-dependant, which is in line with the increase in mitochodrial NADH production from serine reported by Yang et al. [236].

Although serine could sustain energy and biomass increases in the model, the same was not true for glycine. The two amino acids were not freely convertible, as serine formation from glycine required the addition of glucose for its production via the glycolytic pathway. Alternatively, the addition of methylene-THF source was required as shown by the addition of formate. This is in line with the findings by Labuschagne et al. [54], who showed that glycine media supplementation cannot sustain cell proliferation in the same manner that serine can. Furthermore, the authors showed that application of glycine alone lead to the inhibition of glycine flux into purines, and increase in its rate of conversion to serine, depleting the methylene-THF pool. It is unclear why glycine could not sustain proliferation given it can provide methylene-THF via GCS.

It is interesting to note that in both of the described studies, the cells showed high levels of SHMT, GCS and glycine efflux at the same time. If the GCS is unidirectional it would suggest a high rate of glycine breakdown into CO_2 and NH_3 , and thus little glycine release, in contrast with what is seen experimentally as well as in the model results.

It has been previously realised that alterations in the these biosynthetic pathways of serine/glycine metabolism provide an attractive target for the development of cancer treatments, for example PHGHD inhibitors [190, 239] which shows contrasting results in their effectiveness to inhibit proliferation [240], which is likely context-dependent. The model results support previous experimental observations in which serine was shown capable of supporting cellular needs of energy and biomass demands, which does not hold true for glycine. Knockout of the pathways involved in the metabolism of these amino acids highlights the flexibility of metabolism which may explain the reason for which inhibition of individual reactions is not effective. The model could be further used in order to aid the elucidation of the set of reactions involved in serine/glycine metabolism which may be targeted in the treatment of cancers displaying upregulation of enzymes involved in these pathways.

3.4.1 Conclusion

In summary, the major insights of this chapter relate to the establishment of a novel method for integration of scRNA-seq datasets with the metabolic model which was used to shed light on the differences in ATP production pathways in different cell types, as well pathways related to metabolism of glycine. Energy production in glial ad neuronal cells was assessed in the context of the ANLS hypothesis. Both statistical analysis of gene expression as well as model simulations resulted in findings contrary to the assumptions made by the ANLS. Expression of genes encoding enzymes composing the glycolytic pathway suggests this pathway is active both in glial and neuronal cells. Preferential uptake of glucose was shown for neuronal cells during model simulations. Perhaps more suprising was the finding that neuronal cells saturated glucose uptake and released lactate, and that astrocytes preferred lactate as an energy substrate at all ATP demands. As multiple lines of experimental evidence exist supporting and disproving the ANLS theory (Section 3.4) it will remain a worthy topic of further exploration and debate.

Analysis of the expression of genes related to serine and glycine metabolic pathways showed alignment with previous findings, supporting the notion that serine is likely produced in astrocytes and shuttled across to neurons, following the high expression of PHGDH in astrocytes. Th breakdown of glycine was found to be likely active in all cell types, following the ubiqutous expression of the GCSH gene across all cell types, which contrasts with the localization of glycine catabolism to astrocytes only [47]. This finding also further supports the theory for an inter-cellular serine and glycing cycling mechanism within the brain, further supporting by the high expression of genes encoding serine and glycine transporters, GlyT-1 and Asct-1. All brain cells types were lacking the expression of genes encoding the enzymes of the glyoxylate to glycine conversion pathway, resonating with previous findings (Section 3.4) thus glycine is likely not synthesized from glyoxylate in the brain. On the other hand, the high expression of dipeptidases which decompose GSH as well as collagen suggest that a significant pool of glycine is recycled, rather than synthesized *de novo* from serine. The presence and significance of collagen in the brain is an area requiring further exploration.

A suprising outcome of the analysis of gene expression is the finding that serine and glycine metabolic genes are highly present in oligodendrocytes. This can be related to the most well-known function of these cells which is the production of myelin however it may also be relevant to novel roles being attributed to these cells. High expression of genes encoding serine and glycine transporters could hint to potentially two areas. In the first instance, it is possible that oligodendrocytes play a function in regulating neurotransmission during synaptic glycine release by the uptake of extra-synaptic glycine, supported by high GlyT-1 gene expression. In the second case, oligodendrocytes could also be important in providing metabolic support to neurons supported by the high expression of the genes encoding the serine and glycine metabolic enzymes. Although outside of the scope of this study, it is an interesting topic for further exploration.

The knockout of each SHMT subtype during glycine demand simulations lead to insights regarding their role in the context of background metabolism. Knockout of SHMT2 led to the 'metabolic inflexibility' in the system as described by Meléndez-Hevia and de Paz-Lugo [46] during which glycine and methylene-THF cannot be regulated independently. Production of glycine via SHMT2 only does not lead to an engagement of other biosynthetic pathways, although it relies on the cycling of folate groups across the mitochondrial membrane. Since metabolic models do not incorporate the time component, kinetics of transport proteins are not captured which is a limitation of this type of modelling, although it is not the focus of such models.

Increase in biomass demands in the presence of different amino acids showed that serine is one of the main preferred substrates for the support of ATP and biomass generating pathways. Analysis of the different routes from serine to energy formation suggests that the most likely pathway to ATP generation is via the supply of NADH to the ETC following serine to glycine breakdown. The current study also highlighted a possible importance of the reverse GCS reaction in sinking the nitrogen surplus in hypoxic areas of tumour, such that the nitrogen released by utilization of serine is deposited in glycine released from the cell.

Although serine was found to support increases in biomass and energy demands, the reverse was not true for glycine, consolidating that these two amino acids are not interchangeable and glycine likely cannot support proliferation.

Major limitations to the above findings are related to the nature of the scRNA-seq datasets which in big proportion are populated with zero gene counts. This is likely due to the difficulties arising in the experimental procedures that generate these datasets, which may lead to many false negative hits. The complication arising from that is the statistical analysis of the dataset and pertinent to metabolic modelling, definition of a method for the integration of the dataset with the metabolic model such that it provides the most biologically relevant solution. Bulk transcriptomic data has been incorporated into metabolic models for a considerate amount of time by different groups, thus various methods exist which differ in the key parameters such as gene expression thresholds specifying 'active' enzymes to be allowed in the solutons. The effects of applying various methods has been assessed in relationship to the resulting solutions. It is likely that as this field matures, a similar comparison can be done for the method of integration of scRNA-seq with metabolic models. Nevertheless, application of scRNA-seq can provide insights of cell specific metababolic phenotypes. Clustering of pathways utilized by the model weighed by the genetic expression of individual cells showed that this method can be used to assess rare and novel celullar phenotypes (Section 3.3.2). Utilization of transcriptomic data- whether coming from single cells or bulk tissue- is limiting in a number of ways. The number of gene copies is related to the stability of the protein in the cell, thus copy numbers cannot be directly translated to protein expression. Furthermore, post-translational modifications will be determinant of the actual activity of the enzymes- and although the enzyme may be present in a given cell, the direction with which the reaction will proceed is a function of the substrate and product concentrations, which is not taken in the account in the current study.

Chapter 4

Kinetic model of the glycinergic synapse

4.1 Introduction

Glycinergic neurotransmission in the human CNS is important in the regulation of excitatory signalling pathways involving the movement of skeletal muscles, control of pain perception, and others (Introduction, Section 1.2.5). Mutations in genes encoding components of the glycinergic synapse lead to serious, albeit rare conditions such as hyperkplexia, or the metabolic disorder nonketotic hyperglycinemia (NKH). Moreover, there is an interest in the development of glycine transporter 1 and 2 (GlyT1 and GlyT2) inhibitors for the treatment of neuropathic pain and others, the mechanism of action of which is to regulate the time course of glycine in the synaptic cleft by blocking its reuptake.

NKH is normally caused by mutations in the glycine cleavage system (GCS) leading to impaired glycine breakdown, thus increased glycine levels in the CNS. A novel caused for NKH phenotype in a child was reported by Alfadhel,M. et al. [80], who identified a Ser 407 → Gly point mutation in the GlyT1 protein, responsible for reuptake of glycine following its synaptic release. Chapter 5 describes the results from functional characterization of the GlyT1 mutant via electrophysiological recordings in *X.* oocyte and it was found that mutation at the position Ser 407 cause a knockout in GlyT1 function.

The amplitude and the length of the inhibitory post-synaptic currents (IPSC) which are disrupted in disease states are a result of the interaction between the proteins composing the glycinergic synapse. Thus, the focus of this chapter was the development and analysis of a computational kinetic model reproducing the processes of glycine release and reuptake across the synaptic cleft.

Previous experimental studies have established the kinetic parameters for the GlyR, as well as GlyT1 and GlyT2 in isolation- the kinetic model described in the current Chapter was developed with the aim of probing the interaction between these proteins as a whole system.

In the first instance, a detailed scheme of GlyR activation was adapted from Burzomato [144] based on the single-channel recording of $\alpha 1\beta$ GlyR expressed in HEK293 cells in the ScrumPy modelling software. This provided a starting point for the glycinergic synapse model, which was expanded by the addition of reactions describing glycine release from the pre-synaptic cell as well as processes which consume glycine, more specifically reactions describing its reuptake by GlyT1 and GlyT2, as well as loss resulting from diffusion.

Firstly, the GlyR scheme and the resulting IPSC were analysed in isolation. Steady state and dynamic currents were recorded from GlyR stimulated with varying glycine and compared to those reported in previous studies. Secondly, the GlyR model structure and associated kinetic parameters were assessed via the application of local sensitivity analysis, as well as the mathematical reduction of the model to a Monod-Wyman and Changeaux (MWC) type equation (Introduction, Section 1.3.1).

One of the advantages of computational modelling is that models can be applied to probe questions relating to events which - such as neurotransmission - happen on a millisecond scale and may otherwise be difficult to control in an experimental setting. Following, the full glycine synapse model incorporating the detailed GlyR scheme was used to assess the effects of changing the time course and the amplitudes of neurotransmitter released during synaptic simulation. Since these parameters are largely affected by the rate of glycine reuptake, the model was used to estimate the parameters describing GlyT1 and GlyT2 transport incorporating previous experimental observations. Finally, the glycinergic synapse model was used to look at the effects of modulating GlyT1 and GlyT2 reuptake on the glycinergic currents, such as that resulting from the Ser 407 mutation which causes a functional knockout of GlyT1.

4.2 Methodology

All steps of model construction and analysis were completed in ScrumPy (Introduction, Section 1.3.4). Unless otherwise stated, rate equations following mass-action type kinetics were automatically generated by the software in the following form:

$$V = k_f \times (S - P/K_{eq}) \tag{4.1}$$

Where the k_f is the forward reaction rate, S and P represent substrate and product concentrations, and K_{eq} is the equilibrium constant calculated as the ratio of the forward and reverse reactions,

i.e. $K_{eq} = \frac{k_f}{k_r}$. All model concentrations and rates are in Ms^{-1}or s^{-1}.
Model curve fitting was performed in Python using the scipy.optimize.curve_fit package which uses the Levenberg-Marquardt algorithm. The concentration-response curves generated from the model were normalized to the values of the maximum response and fitted to the one-component Hill equation of the form:

$$A = \frac{1.0}{1.0 + (EC_{50}/S)^h} \tag{4.2}$$

Where the EC$_{50}$ is the concentration of substrate eliciting half-maximal response, h is the Hill coefficient of cooperativity (H_{coeff}), and S is the substrate concentration. In order to compare the the speed of signal decrease, 10-90% of the decay phase of glycine and IPSC responses were fitted with a single-exponential equation of the form:

$$A_{max} \times e^{\frac{A}{k}} \tag{4.3}$$

where A is the amplitude, A_{max} is the maximum amplitude, and k is the microscopic decay constant.

4.2.1 Model description

Glycine receptor model

Glycine receptor model (Figure 4.1) and the associated rate constants (Appendix B,6.2) were taken from Burzomato [144] and are based on single-channel recordings from heteromeric rat $\alpha 1 \beta$ GlyR. The receptor can reside in three global conformational states, resting (R), flipped (F), and open (O) and binds glycine which is the agonist (A). Glycine binding affinities to each global conformational state R and F are equal, thus the increase in substrate affinity to protein is a result of a concerted change in the protein conformation stabilised by a higher number of ligands bound, in line with the MWC model of ligand binding cooperativity. The model is composed of 11 reactions between 10 model species. Reactions 1-3, 7 and 8 represent glycine binding to the receptor in different global conformations resting (R) and flipped (F). Reactions 4-6 represent receptor isomerization from the R to F conformation which can happen in the presence any number of ligands bound. Receptor openings (states OA - OA$_3$) can proceed only from the flipped conformations (FA - FA$_3$) and are represented by reactions 9-11.
The V$_{\text{max}}$ and K_{eq} parameters were calculated from the microscopic rate constants taking into account the number of binding sites available in each conformation. For example, the unbound state R contains three available binding sites, but no glycine molecules bound, therefore the K_{eq} is calculated as $3 \times k_f/k_r$ and the V$_{\text{max}}$ $= 3 \times k_f$. The K_{eq} associated with reactions r4 and r6 (Figure 4.1) were calculated from the K_{eq} describing neighbouring reactions in order to respect the constraints imposed by microscopic reversibility on the internal model cycles.

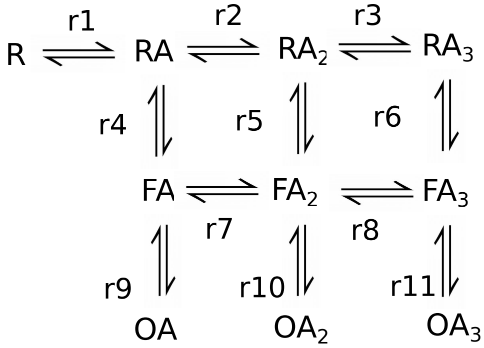


Figure 4.1: **Glycine receptor scheme** The scheme is described by a total of 11 reactions describing the rate of up to three molecules of glycine (A) binding to two global conformational states resting (R) and flipped (F) (reactions 1-3, 7-8) as well as the global conformational changes from the R to F conformations (reactions 4-6) and from the F to the open (O) conformations (reactions 9-11).

Inhibitory post-synaptic currents (IPSC) carried by n receptors are a product of channel conductance and the reversal potential (E_{rev}) of Cl$^-$ ions. The equation used to calculate the current is derived from Ohm's law and assumes a linear current-voltage relationship of the form:

$$IPSC = (g_{max} \times P_{open} * R_n) \times (\Delta\Psi - E_{rev}) \tag{4.4}$$

where $IPSC$ is the current through n receptors, $g_{max} = 45.9 \times 10^{-12}$S is the maximum channel conductance [241], P_{open} is the channel open probability, $R_n = 100$ is an estimation of the number of receptors, $\Delta\Psi$ = -60mV is the membrane voltage and $\mathrm{E_{rev}}$ = -70mV is the reversal potential for Cl$^-$ ions. The assumption of linear current-voltage characteristics is justified based on studies on the electrophysiological signature of GlyR expressed in HEK293 cells showing only small amount of current rectification due to voltage change for heteromeric GlyR [242]. Upon model simulation, the time course of P_{open} which drives the IPSC is calculated from the sum of the open channel conformations ie. $P_{open} = OA + OA2 + OA3$, driven by the binding of ligand.

Glycine synapse model

The glycinergic synapse model can be seen in Figure 4.2 and is composed of the GlyR model described in the previous section as well as reactions associated with release and reuptake of glycine

Table 4.1: Concentrations of Na$^+$ and glycine in the model

Parameter	Model value (mM)	Reference
Gly_i, astrocyte	2	[243]
Gly_i, neuron	10	[66, 244]
Na_i^+, astrocyte	17	[245]
Na_i^+, neuron	30	[246]
Na_e^+, synaptic cleft	150	[66]

across the synapse. Reaction 12 represents glycine release from the synaptic vesicle upon its fusion with the presynaptic membrane and its rate was calculated from a simple mass-action equation of the form:

$$V = k_f \times S \tag{4.5}$$

This equation was also used to describe simplified neurotransmitter diffusion away from the synaptic cleft (r13), as well as the slow processes which act as a sink for the glial (r16) and neuronal (r17) glycine. Reactions describing glycine release and diffusion were adjusted to the slower uptake processes. Reactions 14 and 15 represent bi-directional transport by GlyT1 and GlyT2, respectively. Rate of transport was calculated from a two-substrate Michaelis Menten equation of the form:

$$V = \frac{(V_f \times Gly_e \times Na_e^{+h})}{(K_{Gly_e} + Gly_e) \times (K_{Na_e^+} + Na_e^{+h})} - \frac{(V_r \times Gly_i \times Na_i^{+h})}{(K_{Gly_i} + Gly_i) \times (K_{Na_i^+} + Na_i^{+h})} \tag{4.6}$$

Where the V_f is the maximal reaction rate, Gly_e, Gly_i , Na_e^+ and Na_i^+ are the extracellular and intracellular glycine and Na$^+$ concentrations, associated K parameters are the equilibrium constants for each species, and h is the H_{coeff} for Na$^+$. Maximal forward and reverse reaction rates, V_f or V_r, are calculated in the model from the microscopic turnover rate of the transport cycle, k multiplied by the transporter number, T and a constant, c :

$$V = k \times T \times c \tag{4.7}$$

where $T = 1000$, and $c = 1\times10^{-4}$ which were estimated in the current study. Values of the microscopic turnover rates of each transporter can be seen in the Appendix B, 6.2. The equilibrium constant of glycine binding has been shown to vary according to membrane voltage, and large rises in the affinity of glycine to both transporters can be seen at positive voltage potentials [66, 219]. The values of parameters in the model are based on a membrane voltage of -60mV. Intracellular concentrations of Na$^+$ and glycine describe the astrocytic and neuronal concentrations and can be seen in Table 4.1. Na$^+$ concentrations were treated as model parameters thus were not allowed to vary during model simulation. ScrumPy model file can be found in Appendix A, 6.2.

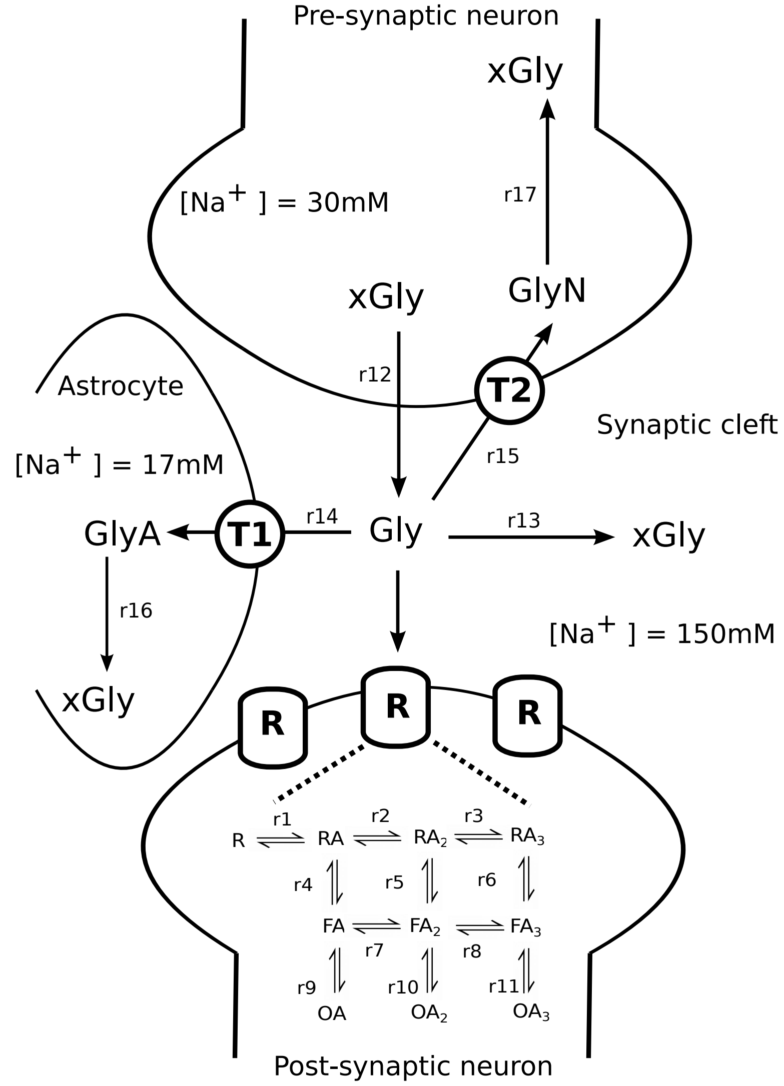


Figure 4.2: **Glycinergic synapse model** Glycine is released from the presynaptic neuron (r12) and diffuses across the synaptic cleft, binding to the glycine receptors (R) on the opposing neuron. Termination of neurotransmission is regulated by the removal of glycine via passive diffusion (r13) as well as uptake into the astrocyte by the GlyT1 (r14) and back into the pre-synaptic neuron by GlyT2 (r15). Concentrations of Na^+ which drive transporter uptake are based on physiological values and can be seen in Table 4.1

.

4.3 Results

4.3.1 Interactions of glycine with the GlyR

Characterisation of the steady state and dynamic GlyR responses

Glycine receptor model responses were analysed over a range of glycine concentrations in order to confirm that the model implementation can realistically reproduce experimentally-recorded IPSC. Concentration response curve of the P_{open} at varying glycine concentrations from 10 to 1000μmol can be seen in Figure 4.3. The curve was fitted to a Hill equation with the resulting parameters $\mathrm{EC}_{50} = 60.9\mu$mol and $H_{\mathrm{coeff}} = 1.54$ which are in agreement with the values obtained in the original study [144]. The receptor is saturated at a glycine concentration of $\approx 300\mu$mol with a maximum P_{open}= 0.94 which correctly reflects the high P_{open} for $\alpha 1\beta$ GlyR previously reported at 0.7-0.9 by Legendre [247].

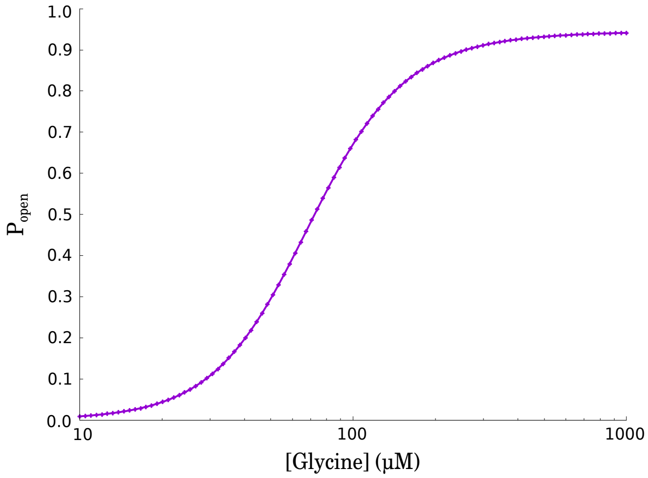


Figure 4.3: **Concentration response curve for glycine binding to GlyR** Steady state glycine open probability was calculated over a range of glycine concentration 10-1000μmol. Fit of the Hill equation resulted in parameters $\mathrm{EC}_{50} = 60.9$ and $H_{\mathrm{coeff}} = 1.54$.

Next, time courses of dynamic model occupancies calculated from the model stimulated with 1mM glycine for 10ms were compared to GlyR model stimulation from a study by Colquhoun and Lape [88] which utilised a different method for model states calculations. Results from both models are in close agreement, confirming that the ScrumPy GlyR implementation can accurately reproduce the kinetic behaviour of GlyR described in the original study (data not shown).

Sensitivity analysis of the GlyR scheme

In order to establish which receptor model parameters mostly affect the resulting glycinergic currents, local sensitivity analysis was performed on the model kinetic constants. Since the IPSC is directly proportional to the sum of the open conformations, the sensitivities of steady state $[O]$ to a small change in the K_{eq} describing reactions 1-11 were calculated over a range of glycine concentrations from 10-300μmol. The quantitative value resulting from the analysis is the response coefficient , $R^{[O]}_{K_{eq}}$, formally defined as:

$$R^{[O]}_{K_{eq}} = \frac{\partial ln[O]}{\partial ln K_{eq}} \tag{4.8}$$

where $[O]$ is the sum of the open conformations and K_{eq} is the equilibrium constant of a reaction. Values of $R^{[O]}_{K_{eq}}$ may vary from 0 to 1, where higher values indicate a bigger impact of changing the independent variable, K_{eq} , on the dependent variable, $[O]$. The $R^{[O]}_{K_{eq}}$ calculated for each GlyR reaction is plotted against glycine concentrations in Figure 4.4, a). It is apparent that the overall sensitivity of $[O]$ drops as the concentration of glycine increases, which is expected given that the receptor experiences half saturation at $\approx$ 60μmol. As glycine concentration increases, the receptor resides in the tri-liganded, open state thus the remaining states will not be occupied frequently. The $[O]$ is most sensitive to variations in the rates of the initial glycine binding to the R conformation (RBind1), the binding of the second glycine to the F conformation (FBind3) as well as the opening in the tri-liganded state (Open3). High $R^{[O]}_{K_{eq}}$ over the range can also be seen for K_{eq} describing the rate of receptor binding the second glycine molecule (RBind2), as well as a change in the global conformation from the double-liganded R to F states (Flip2). At low concentrations of glycine, $\leq$ $\approx$20μmol, the rate of the second glycine binding to the F conformation (FBind2) as well as receptor flipping in the single-liganded state (Flip1) also seem to have a large effect on the $[O]$, which becomes lower as the glycine input increases. On the other hand, $R^{[O]}_{K_{eq}}$ of the K_{eq} describing receptor opening from a single-bound conformation (Open1), the binding of a third glycine molecule to the R conformation (RBind3) as well as flipping of the triple-bound R to F (Flip3) are consistently low across all concentration ranges.

These results raise a question regarding the occupancy of these states which are products of the reactions with the lowest response coefficients. In order to determine if the results from the sensitivity analysis identify which states are hardly occupied at all glycine concentrations, the sum of flux via each reaction was recorded following a 10ms model simulation with ranging glycine concentrations from 10-300μmol (Figure 4.4, b). The sum of fluxes correlate with the $R^{[O]}_{K_{eq}}$ as given by the sensitivity analysis, thus suggesting that even at low glycine concentrations receptor states OA and RA$_3$ are not visited frequently. According to the kinetic parameters from the original publication, the receptor will bind two glycine molecules before switching to the flipped conformational state, followed by the binding of the third glycine and subsequent opening. These results were confirmed to match similar findings to those reported by Colquhoun and Lape [88] who identified similar ‘routes’ of GlyR transitions.

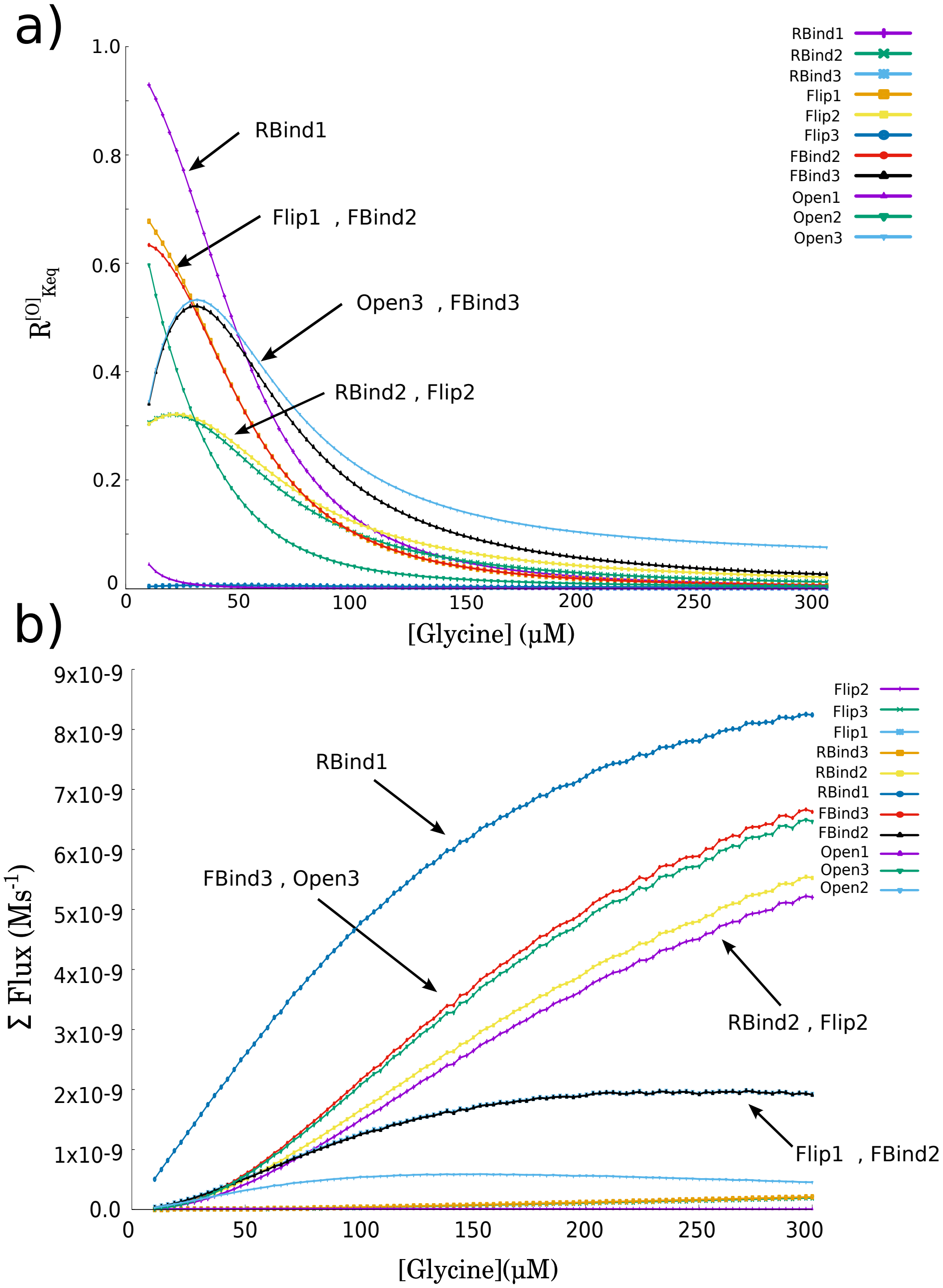


Figure 4.4: **Sensitivity of $\mathbf{P_{open}}$ to the $\mathbf{K_{eq}}$ of GlyR transitions** Top graph (a) shows the $R^{[O]}_{K_{eq}}$ of the all model reactions K_{eq} at varying glycine concentration. At concentrations $\geq$ $\approx$20μmol, the $R^{[O]}_{K_{eq}}$ correlate with the sum of flux carried by each reaction in a dynamic model simulation with 1mM glycine for 10msec (b).

Reduction of the GlyR scheme to a MWC-type equation

The GlyR model by Burzomato [144] is based on the Monod Wyman and Changeaux (MWC) model which explains ligand binding cooperativity under several assumptions. However, the model is not a complete MWC scheme thus it does not describe the rates of glycine binding to the open conformations, nor the ratios of receptors in the different conformational states, R, F and O in the absence of ligand. This is because the experimental data which the scheme was fit to does not contain information about these states- openings in the absence of ligand are very rare, and any conformational changes that are experimentally recorded are evoked in the *presence* of agonist; currents are not generated through closed states in the agonist *absence.*

Hess's law states that the enthalpy of a chemical process from reactant to product must remain the same regardless of the path taken. Since the GlyR scheme is composed of internal cycles, Hess's law can be applied to the scheme in order to calculate missing parameters which could not be obtained experimentally.

Treatment of the scheme as for the original MWC model allowed for the derivation of two equations which describe the steady state fractions of receptor in each conformational state (Equation 4.14) as well as the fractional occupancy of the receptor (Equation 4.13) which is the fraction of ligand bound to the receptor.

The overall goal of the derivation was to express each species in terms of the substrate concentration, [S] and substrate dissociation constants for [O],[F] and [R] ; K_O, K_F and K_R respectively. Firstly, all of the species were expressed in terms of the concentration of the open, unliganded state, [O], following which it was set to 1 to allow for the elimination of concentrations from the equations altogether.

The following constants were defined: $\alpha = \frac{[S]}{K_O}$, $c = \frac{K_O}{K_F}$, $d = \frac{K_F}{K_R}$.

Therefore: $\alpha c = \frac{[S]}{K_F}$, $\alpha cd = \frac{[S]}{K_R}$, $M = \frac{F}{O}$ and $L = \frac{R}{F}$.

Secondly, each species was re-written in terms of [O] and the concentration of substrate, [S], as well as the dissociation constants for each global conformations which were written in terms of individual sites available. For example; $[OA_1] = 3 \times [O] \times \frac{[S]}{K_O}$ because there are 3 vacant substrate sites available in the [O] conformation but only one site occupied in the $[OA_1]$ conformation therefore the probability of the first event is three, however the second only one. This resulted in a set of equations:

$[\mathrm{O}] = 1$

$[\mathrm{OA}_1] = 3[O]\frac{[S]}{K_O} = 3\alpha$

$[\mathrm{OA}_2] = [OA_1]\frac{[S]}{K_O} = (3 \times 1)[O](\frac{[S]}{K_O})^2 = 3\alpha^2$

$[\mathrm{OA}_3] = [OA_2]\frac{[S]}{K_O} = (3 \times 1 \times \frac{1}{3})[O](\frac{[S]}{K_O})^3 = 3\alpha^3$

$[\mathrm{F}] = \mathrm{M}\ [\mathrm{O}] = \mathrm{M}$

$[\mathrm{FA}_1] = 3[F]\frac{[S]}{K_F} = 3M[O]\frac{[S]}{K_F} = 3M\alpha c$

$[\mathrm{FA}_2] = 3[F_1]\frac{[S]}{K_F} = (3 \times 1)M[O](\frac{[S]}{K_F})^2 = 3M\alpha^2c^2$

$[\mathrm{FA}_3] = 3[F_2]\frac{[S]}{K_F} = (3 \times 1 \times \frac{1}{3})M[O](\frac{[S]}{K_F})^3 = 3M\alpha^3c^3$

$[\mathrm{R}] = \mathrm{L}\ [\mathrm{F}] = \mathrm{LM}[\mathrm{O}] = \mathrm{LM}$

$[\mathrm{RA}_1] = 3[R]\frac{[S]}{K_R} = 3LM[O]\frac{[S]}{K_R} = 3LM\alpha cd$

$[\mathrm{RA}_2] = 3[R_1]\frac{[S]}{K_F} = (3 \times 1)LM[O](\frac{[S]}{K_R})^2 = 3LM\alpha^2c^2d^2$

$[\mathrm{RA}_3] = 3[R_2]\frac{[S]}{K_R} = (3 \times 1 \times \frac{1}{3})LM[O](\frac{[S]}{K_F})^3 = 3LM\alpha^3c^3d^3$

The fractional saturation equation describes the amount of agonist bound to receptor. The number of sites occupied by the agonist forms the numerator, where each species is multiplied by the number of agonist it has bound. The denominator is the sum of all the agonist sites available, regardless if the agonist is bound or not thus each species is multiplied by three.

$$y = \frac{[OA_1]+2[OA_2]+3[OA_3]+[FA_1]+2[FA_2]+3[FA_3]+[RA_1]+2[RA_2]+3[RA_3]}{3\times([O]+[OA_1]+[OA_2]+[OA_3]+[F]+[FA_1]+[FA_2]+[FA_3]+[R]+[RA_1]+[RA_2]+[RA_3])} \tag{4.9}$$

The fraction of receptor in the open conformation, $\sum[O]$ is the sum of all the open conformations divided by the sum of the total receptor, ie.:

$$\sum[O] = \frac{[O]+[OA_1]+[OA_2]+[OA_3]}{[O]+[OA_1]+[OA_2]+[OA_3]+[F]+[FA_1]+[FA_2]+[FA_3]+[R]+[RA_1]+[RA_2]+[RA_3]} \tag{4.10}$$

Substituting in the concentrations of species in terms of agonist and dissociation constants described previously yields:

$$y = \frac{3\alpha+6\alpha^2+3\alpha^3+3M\alpha c+6M\alpha^2c^2+3M\alpha^3c^3+3LM\alpha cd+6LM\alpha^2c^2d^2+3LM\alpha^3c^3d^3}{3\times(1+3\alpha+3\alpha^2+\alpha^3+M+3M\alpha c+3M\alpha^2c^2+M\alpha^3c^3+LM+3LM\alpha cd+3LM\alpha^2c^2d^2+LM\alpha^3c^3d^3])} \tag{4.11}$$

$\therefore$

$$y = \frac{3\alpha(1+2\alpha+\alpha^2)+3M\alpha c(1+2\alpha c+\alpha^2c^2)+3LM\alpha cd(1+2\alpha cd+\alpha^2c^2d^2)}{3\times((1+3\alpha+3\alpha^2+\alpha^3)+M(1+3\alpha c+3\alpha^2c^2+\alpha^3c^3)+LM(1+3\alpha cd+3\alpha^2c^2d^2+\alpha^3c^3d^3))} \tag{4.12}$$

which reduces to the equation describing the fractional saturation:

$$y = \frac{\alpha(1+\alpha)^2+M\alpha c(1+\alpha c)^2+LM\alpha cd(1+\alpha cd)^2}{(1+\alpha)^3+M(1+\alpha c)^3+LM(1+\alpha cd)^3} \tag{4.13}$$

Thus the equation for receptor fraction in the open conformation becomes:

$$\sum[O] = \frac{(1+\alpha)^3}{(1+\alpha)^3+M(1+\alpha c)^3+LM(1+\alpha cd)^3} \tag{4.14}$$

Which is equivalent to the following equation describing the sum of the open conformations in terms of the substrate concentration and dissociation constants:

$$\sum[O] = \frac{(1+\frac{[S]}{K_O})^3}{(1+\frac{[S]}{K_O})^3+M(1+\frac{[S]}{K_F})^3+LM(1+\frac{[S]}{K_R})^3} \tag{4.15}$$

The values for the parameters in the above equations are calculated as follows:

$$d = \frac{K_F}{K_R} = \frac{8^{-6}}{520^{-6}} = 0.0154$$

$$Ld = \frac{[RA_1]}{[FA_1]}$$

but

$$L = \frac{[F]}{[R]}$$

$\therefore$

$$L = \frac{1}{Ld} = \frac{1}{0.006 \times 0.0154} = 10822$$

During the original model parameter fitting the model did not include reactions for agonist binding to the open conformations thus the value for K_O was not constrained by microscopic reversibility. Calculating the K_O from the internal cycles returned two different values, disagreeing with model assumptions that the speed of agonist binding to each conformation is the same, regardless of how many agonists are bound.

Parameters c and M were thus calculated as follows:

$$\frac{\alpha_3}{\beta_3} \times \frac{\beta_1}{\alpha_1} = c^2 = 0.067$$

$\therefore$

$$c = \sqrt{0.067} = 0.26$$

$$\frac{\alpha_3}{\beta_3} = Mc^3$$

$\therefore$

$$M = \frac{\alpha_3}{\beta_3} \times \frac{1}{c^3} = 3.13$$

and

$$\frac{\alpha_1}{\beta_1} = Md$$

$\therefore$

$$M = \frac{\alpha_1}{\beta_1} \times \frac{1}{c} = 3.13$$

Which give similar results for M (rounded to two decimal places).

Thus the resulting set of parameters for the equations are:

- d= 0.015
- c= 0.26
- L $\approx 11\times10^{3}$
- M = 3.13

Values for the dissociation constants are $K_R = 520\times10^{-6}$ and $K_F = 8\times10^{-6}$.

As for K_O, if:

$$c = \frac{K_O}{K_F}$$

then:

$$K_O = c \times K_F = 0.26 \times 8\times10^{-6} = 2.07\times10^{-6}$$

Full scheme in terms of the MWC parameters described in the derivation can be seen in Figure 4.5. The dissociation constant for glycine open conformation, K_O, equalled $2.07 \times 10^{-}6$ Ms^{-1}. As expected, glycine binding to an already open receptor is extremely rapid. The high value for the L parameter reveals that the fraction of receptors in the flipped conformation is extremely low in the absence of ligand. It follows that the ratio between flipped and open conformations in ligand absence is small, implying that there is a high energetic cost in the initial change in receptor conformation from resting to flipped state, followed by receptor opening which is less energetically demanding. In order to confirm these equations are good approximations for the model output at steady state, open receptor fractions were calculated at glycine inputs ranging from 10 to 1000μmol from the full model and Equation 4.14. The results can be seen in Figure 4.6. Fitting of a Hill equation to both curves resulted in values of the EC$_{50}$ and H_{coeff} which are near identical (Full model EC$_{50}$ = 71.1μmol, $H_{\text{coeff}} = 2.27$, Equation 1 EC$_{50}$ = 69.3μmol, $H_{\text{coeff}} = 2.46$). These results show that the full scheme can be successfully reduced to Equation 4.14 capable of reproducing steady-state GlyR fractions, and therefore IPSC.

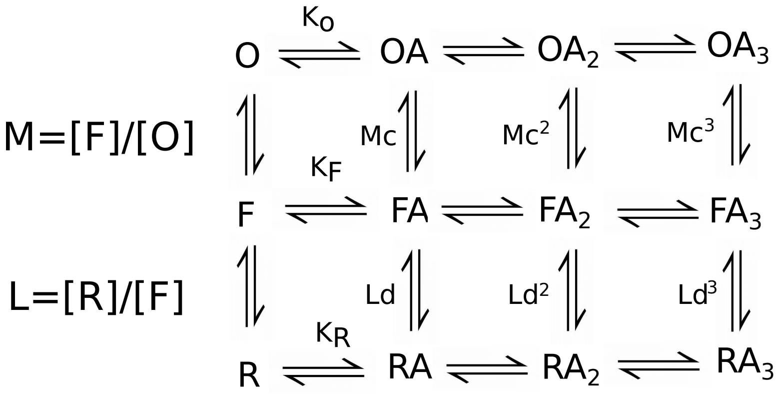


Figure 4.5: **Full MWC scheme describing the glycine receptor** GlyR scheme rewritten in its full form, with the addition of unbound global conformations F and O, the occupancy of which is extremely rare. Rates of the global conformation changes are written in terms of receptor concentrations.

H

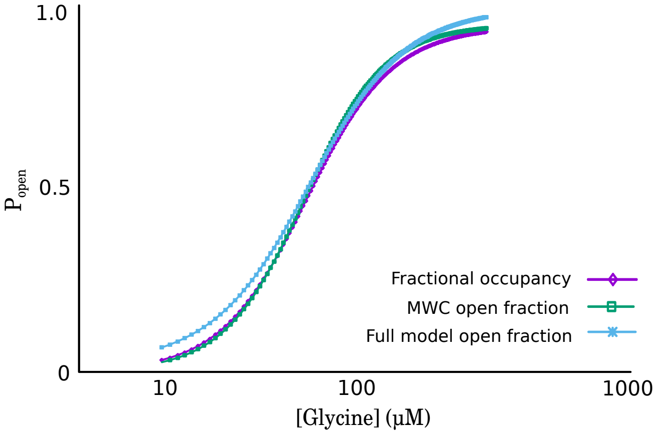


Figure 4.6: **Concentration response curves calculated from the MWC *versus* full model** P_{open} at steady state was calculated from the full model (blue) and Equation 4.14 (green) over a range of glycine concentrations (10-1000μmol) resulting in near identical curves.

4.3.2 Glycine synapse model

Effect of the neurotransmitter profile on the glycinergic currents

Ca^+-triggered glycine release from the presynaptic cleft is a fast event happening at the order of milliseconds, followed by rapid neurotransmitter reuptake. Thus the activation of GlyR at the post-synaptic density happens under non-equilibrium conditions, shaped by the duration of neurotransmitter transients in the synaptic cleft [248]. In the next section, the effects of changing the kinetic parameters of glycine transporters on glycinergic transients in the synaptic cleft were assessed under physiological conditions. The maximal concentration of glycine and the time constant of its decay, τ_{decay}, have been previously estimated in rat lumbar motorneurons by Beato [249] and equalled 2.2-3.5mM and 0.6-0.8ms, respectively, providing a starting point for model simulations. Figure 4.7 shows the results of model simulation with 4mM glycine for 1ms in the absence of transport reactions. Synaptic glycine concentration reached a peak of 3.5mM and was removed by diffusion with a τ_{decay}= 0.5ms, resulting in an IPSC current maximum amplitude, I_{max}, of 413pA with a τ_{decay}= 8.78 ms. It can be seen in Figure 4.7 b), c) that the time course of current decay closely follows the time course of P_{open} with the latter reaching a maximum value of 0.9, close to the values achieved under saturating glycine transients seen in Section 4.3.1. The resulting IPSC decay time is within the range of previously published values recorder from glycinergic neurons of the spinal cord, ranging from 5 to 15ms [249–251].

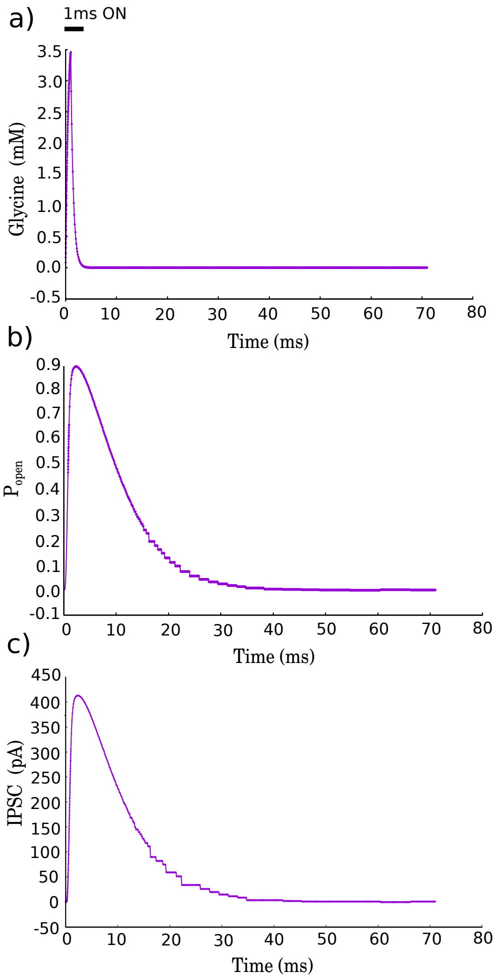


Figure 4.7: **IPCS evoked by fast synaptic glycine transients** The model was simulated with a 1ms pulse of 4mM glycine in the absense of the transport reactions. Top graph (a) shows the time course of glycine pulse, leading to a high maximum GlyR P_{open} of 0.9 which resulted in an IPCS with a $\tau_{decay} = 8.78$ ms. The staircase effect seen in the decay phases of P_{open} and IPSC amplitudes is an artefact of the solver software used.

Effect of varying the concentration and duration of released glycine transients on IPSC

Apart from the processes which remove glycine during neurotransmission, the neurotransmitter transients may also be affected by the rate of vesicular release, the concentration within the vesicle, as well as the length of the release event. The model sensitivity of the IPSC to released glycine concentration and duration of release were assessed by varying these two parameters over a range of glycine concentrations from 0.5 to 5.0mM and durations of release between 0.2 and 3ms. The results can be seen in Figure 4.8, where the top graphs show a variation in the resulting IPSC τ_{decay}(a) and maximum amplitude (b). Effects of changing the independent variables on glycine maximum amplitude (c) can be seen in the bottom graph. Overall, there is a higher variability in the resulting amplitudes of both glycine and IPSC which increase as the concentration and the duration of glycine pulse increase, as expected. Maximum concentrations of glycine show a uniform increase across the two independent variables. On the other hand, the increase in the maximal IPSC reached is non-linear, reflecting receptor saturation. The decay time constant of the glycine peaks does not significantly change across all of the concentrations and durations tested and remains stable at 0.5ms, showing that the shape of the neurotransmitter decay is shaped mostly by the rate of its consumption by diffusion, rather than the length of release. However, there is an increase in the decay time constant of the IPSC from ≈ 6.5 to 9ms across the range. Significant differences can be seen at the extreme values of independent variables, with the majority of the IPSC decay time constant values falling in between 8 and 9ms. Given that the $\tau_{decay}$of glycine does not significantly change, the variations in IPCS time courses and amplitudes are mostly driven by the top glycine concentrations reached. In addition, long, low concentration glycine transients elicit currents with a similar time course as fast, high concentration release events.

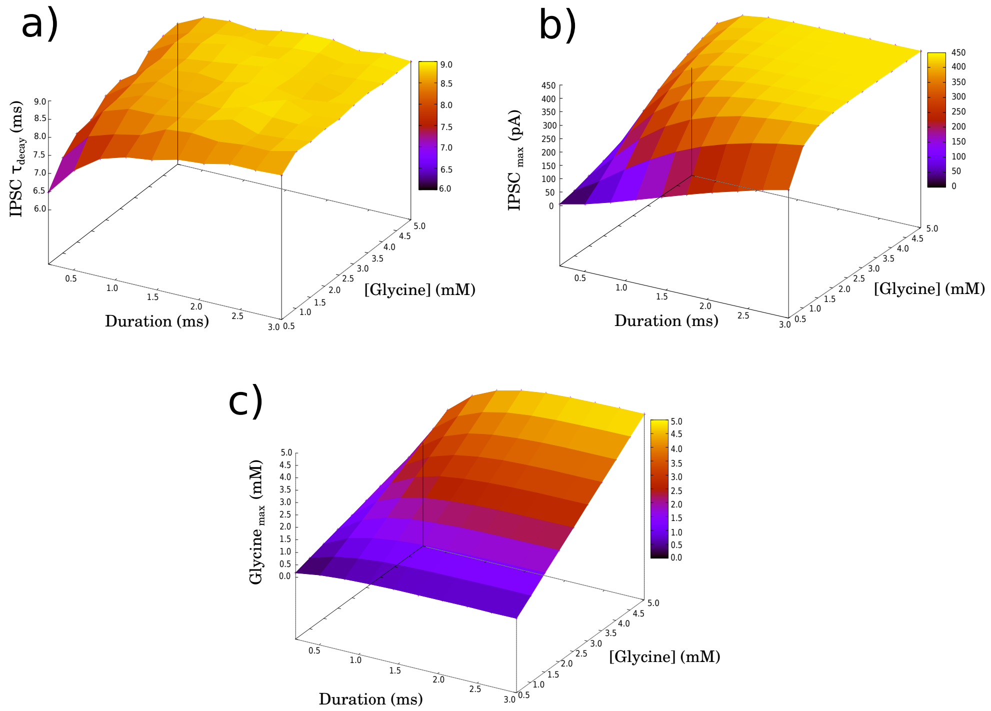


Figure 4.8: **Effects of varying glycine release concentration and duration on the resulting IPSC** Current amplitudes varied between 5.6- 434.1pA with a decay time constant variation from 6.48- 8.98ms, showing a much higher differences in the top amplitudes reached. Glycine amplitudes varied from 0.17- 4.99mM with a decay time constant remaining equal at approximately 0.5ms.

4.3.3 Estimation of the GlyT1 and GlyT2 kinetic parameters

Glycine and sodium binding affinities in the absence of product

In the first instance, kinetics of glycine uptake by GlyT1 and GlyT2 in the absence of products were established. Figure 4.9 shows steady state rates of Na^+ (a) and glycine (b) transport by GlyT1 and GlyT2 in the presence of extracellular glycine only. Na^+ concentration response curves were calculated in the presence of 1mM extracellular glycine over a range of Na^+ concentrations ranging from 0.1 to 150mM. Fits of the Hill equation to the Na^+ binding curve for GlyT1 resulted in transporter EC_{50} = 8.4mM and a H_{coeff} = 1.2. This is in line with the previously reported values of $EC_{50} \approx$ 7mM and H_{coeff} mean value of 1.4 in *X.* oocytes expressing GlyT1 voltage held at -70mV reported by Roux and Supplisson [66]. Glycine receptor 2 EC_{50} equalled 44.2mM with a H_{coeff} of 1.9, which also agrees with the $EC_{50} \approx$ 30mM and H_{coeff} mean value of 2.2 reported in

the same study. GlyT1 binds extracellular Na^+ with a higher affinity than GlyT2, however with a lower Na^+ cooperativity which can be seen by the right shift of the curve.
Rate of glycine uptake (Figure 4.9, b) for each transporter was plotted in the presence of extracellular $Na^+ = 150$mM, over a range of glycine concentrations 1-1000μmol. Fits of the Hill equation resulted in an $EC_{50} = 21.2\mu$mol and a $H_{\text{coeff}} = 1.0$ for GlyT1 and an $EC_{50} = 22.0\mu$mol and a $H_{\text{coeff}} = 1.0$ for GlyT2, which also agree with the results reported by Roux and Supplisson [66].

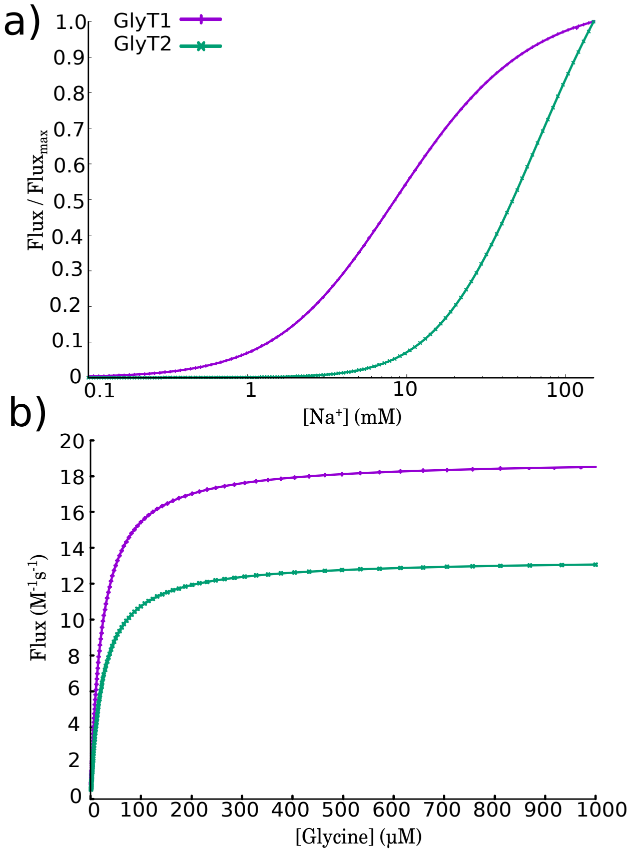


Figure 4.9: **Concentration response curves of Na^+ and glycine binding to GlyT1 and GlyT2** GlyT1 has a higher affinity for Na^+ than GlyT2 (a), although the latter displays higher cooperativity of binding which is a result of the transport stoichiometry, with GlyT2 transporting $3Na^+$ in comparison to $2Na^+$ import per transport cycle for GlyT1. Bottom graph shows maximal rates of steady state glycine transport in the presence of 150mM Na^+ .

Reverse transport of glycine by GlyT1 and GlyT2

Reversal of transporter flux is a function of the equilibrium potential of each transporter, which depends on the electrochemical gradients of the substrate and co-substrate ions, as given by the Nernst equation:

$$E_t = \frac{RT}{(n_{Na} - n_{Cl})} ln \left(\frac{{Na^+}_e^{nNa} {Cl^-}_e^{nCl} Gly_e}{{Na^+}_i^{nNa} {Cl^-}_i^{nCl} Gly_i} \right) \tag{4.16}$$

Where R is the gas constant, T= 298K, F is the Faraday constant, and n_{Na} is the number of ions coupled to transport. Previous studies have shown that the ability of reverse transport by each transporters differ, with GlyT1 showing a significant rate of glycine transport in the reverse direction as the internal Na$^+$ is increased [66, 252, 253] in comparison to GlyT2 [66], the function of which is to ensure a high synaptic glycine concentration required for vesicular refilling (Introduction, Section 1.2.5). Aubrey et al. [252] have reported the glycine binding affinity to the internal facing GlyT1 conformation at ≈ 4.3mM, which was the starting point for the estimation of the kinetic parameters determining GlyTs reversal that agree with experimental observations. Figure 4.10 shows the difference in maximal reverse transport rates by each transporter in the presence of 2mM and 10mM glycine, the estimated astrocytic and neuronal concentrations, respectively. The rate of transport was calculated in the first instance assuming that the internal glycine K_{eq} is equal for each transporter at 4.3mM, and the affinity of the intracellular Na$^+$ to the transporter is the same as for the extracellular Na$^+$ at 6mM and 30mM for GlyT1 and GlyT2, respectively. Extracellular and intracellular Na$^+$ were set to 150mM and 50mM in all calculations. It can be seen that at both concentrations of internal glycine the maximal reverse rate is higher for GlyT1, which has a higher affinity for Na$^+$ in both intracellular and extracellular-facing conformations. Assuming an astrocytic glycine concentration of 2mM, reversal of transport to the extracellular glycine-consuming direction occurs when extracellular glycine approaches 6.5μmol , in comparison to 3μmol for GlyT2. In the presence of higher intracellular glycine concentrations of 10mM, both transporters show a significant reverse transport until extracellular glycine reaches concentrations of ≈ 7μmol and 20μmol for GlyT2 and GlyT1, respectively. Across all observations, the direction of transport at concentrations of extracellular glycine $\geq$ 20μmol is in the extracellular to intracellular direction, even if the internal glycine concentration is high, which shows that the transport cycle is mostly driven by the high Na$^+$ gradients across the compartments.

Assuming that neuronal glycine concentration is ≈ 10mM, given the previously described parameters GlyT2 would show significant glycine leaks from the presynaptic neuron in the absence of neurotransmission. This does not agree with results by Roux and Supplisson [66], who demonstrated that GlyT2 reversal transport does not occur at a significant rate from *X.* oocytes. The authors concluded that GlyT2 affinities to intracellular glycine and Na$^+$, or the rate of reverse transport must be lower than those in the extracellular facing state. In order to estimate the binding affinities of intracellular Na$^+$ and glycine to GlyT (Equation 4.6), these parameters were ranged from 30-100mM and 4-20mM, respectively. The intracellular glycine concentration was set to 10mM, in the presence of 50mM intracellular and 150mM extracellular Na^+ . The results can be seen in Figure 4.11 and show that any combination of Na^+ K_{eq} from 80-100mM, and glycine K_{eq} 16-20mM would produce a rate of reverse transport at less than $0.5Ms^{-1}$. For the next model simulations values of Na^+ $K_{eq} = 80$mM and glycine $K_{eq} = 15$mM were chosen in order to ensure minimal glycine reverse transport via GlyT2.

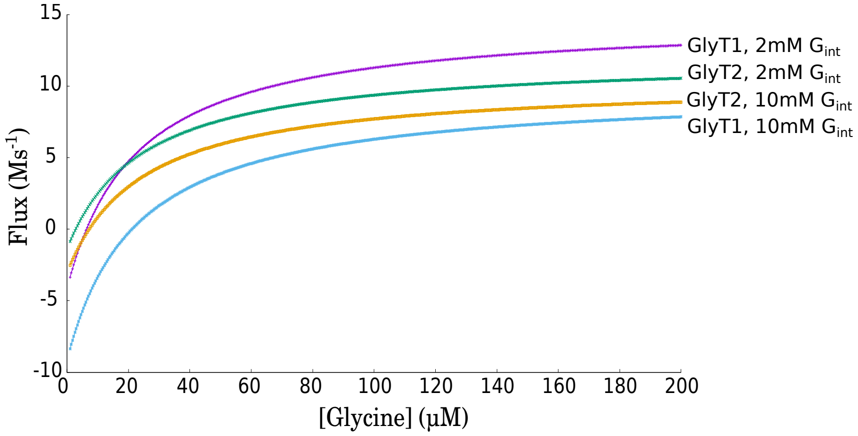


Figure 4.10: **Maximal rates of reverse transport by GlyT1 and GlyT2** Steady state reaction rates were calculated for each transporter at internal glycine concentrations of 2mM and 10mM, under the assumption that the affinity of internal glycine binding to each transporter is equal. Na^+ affinities to each transporter were assumed to be the same towards extracellular and intracEllular-facing conformations. GlyT1 has a higher potential for reverse transport at both concentrations tested.

Reverse glycine transport by GlyT1 and GlyT2 under physiological conditions

The effect of increasing intracellular glycine concentrations on the extracellular glycine levels and subsequent GlyR activation was measured. Figure 4.12 shows the results from a 10ms model simulation in the presence of extracellular Na^+ at 150mM, astrocytic Na^+ at 17mM, and neuronal Na^+ at 30mM. Top graph, a) , shows the maximum extracellular glycine released by each transporter at varying internal glycine concentrations, in the absence of other uptake processes. The resulting maximum P_{open} can be seen in b). The bottom graphs show results from equivalent 10ms simulations at varying internal glycine concentrations but in the presence of all model uptake processes. GlyT2 does not release glycine from the neuronal cell at concentrations high enough for

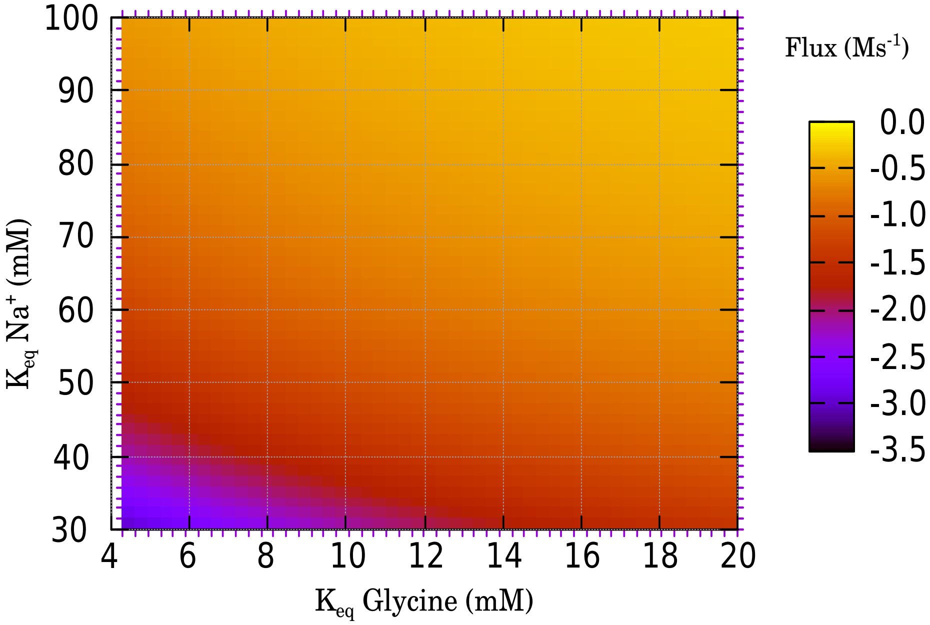


Figure 4.11: **Effects of changing the GlyT2 Na$^+$ and glycine affinities on the rate of reverse transport** Affinities of intracellular Na$^+$ and glycine were ranged in the presence of 10mM intracellular glycine to assess in which range the parameter combinations produce minimal glycine extrusion in the absence of neurotransmission.

GlyR activation even at very high intracellular glycine concentrations $\geq$ 15mM. In comparison, the model shows that as the astrocytic glycine concentration increases, there is a considerable glycine efflux into the extracellular space, leading to low levels of GlyR openings. However, as the uptake processes of GlyT2 and diffusion are allowed, the small concentrations of astrocytic glycine released are effectively cleared, keeping the concentration in the sub-micromolar range not high enough for GlyR activation.

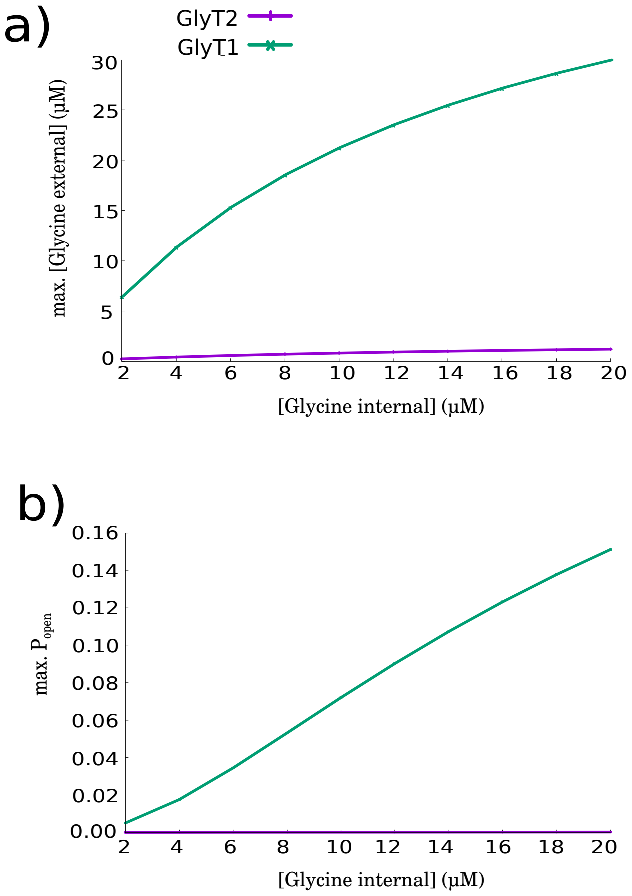


Figure 4.12: **Glycine release at varying intracellular concentrations under physiological conditions** In each simulation the starting intracellular astrocytic (GlyT1) or neuronal (GlyT2) glycine concentration was increased from 2-20mM. Increases of astrocytic glycine leads to its release via GlyT1 (a), which in the absence of GlyT2 and diffusion leads to GlyR activation (b). In the presence of GlyT2 and diffusion processes, astrocytic glycine release is negligible due to its fast consumption in the synapse, and does not lead to any significant GlyR activation. GlyT2 does not release glycine in concentrations high enough for GlyR activation even at very high neuronal concentrations.

4.3.4 Effects of GlyT1 and GlyT2 uptake on the time course of glycinergic currents

Effect of changing transporter numbers on the IPSC

In order to identify the effects of changing GlyT1 and GlyT2 densities on the resulting currents, transporter numbers were ranged from 100-2000 and the resulting maximum IPSC amplitude and decay time constant were measured in a dynamic model simulation with 4mM glycine for 1ms given the initial GlyT1 and GlyT2 parameters:

Table 4.2: Initial set of GlyT1 and GlyT2 kinetic parameters

Paremeter	GlyT1 model value (Ms^{-1}, s^{-1})	GlyT2 model value (Ms^{-1}, s^{-1})
k_f	2.00×10^{1}	2.00×10^{1}
k_r	2.00×10^{1}	2.00×10^{1}
K_{eq} Gly$_e$	22.60×10^{-6}	24.60×10^{-6}
K_{eq} Gly$_i$	4.30×10^{-3}	17.00×10^{-3}
K_{eq} Na_e^+	6.00×10^{-3}	30.00×10^{-3}
K_{eq} Na_i^+	6.00×10^{-3}	8.00×10^{-2}
Hill Na^+	1.4	2.2

Kinetic parameters used in Equation 4.6 to describe GlyT1 and GlyT2 transport, where k_f and k_r are used for the calculation of the V_f and V_r values, respectively. K_{eq} refer to the equilibrium constants of Na^+ and glycine in the forward (extracellular, e) and reverse (intracellular, i) reaction directions.

As can be seen in Figure 4.13, increasing the number of GlyT2 leads to a faster decaying IPSC, decreasing its decay time constant from 12ms to 9.5ms. Increasing the number of GlyT1 had the opposite effect, and caused a longer decay of the IPSC from 9.5ms to 10.5ms. Reduction in IPSC decay time constant and amplitude seen upon GlyT2 increase is expected, as the higher number of transporters leads to faster removal of the released glycine, leading to a shorter time course of glycine across the synaptic cleft. This observation is in line with experimental results in which the pharmacological inhibition of both GlyT1 and GlyT2 leads to a slow down in the IPSC decay, increasing its decay time constant by ≈ 1-5 ms [249, 251, 254]. On the contrary, the increases in the τ_{decay} seen as GlyT1 numbers are increased do not fit with the experimental observations described. This effect can be explained by the difference in the reverse transport abilities of each transporter, given the parameters in Table 4.2. Figure 4.13, c), d) shows the maximal concentration of astrocytic and neuronal glycine reached during each simulation, from a starting concentration of 2mM and 10mM, respectively. As expected, increasing the number of GlyT1 leads to an increase in the total astrocytic glycine uptake, whereas increase in the GlyT2 leads to its decrease as the competition between the two transporters leads to a higher accumulation of glycine in the neuron. Similarly, increasing the number of GlyT2 increases neuronal glycine, however this effect is also seen when GlyT1 is increased. As the number of GlyT1 increases, there is an increase in the rate of reverse reaction via GlyT1 leading to a higher concentration of astrocytic release followed by GlyT2 reuptake. Figure 4.14 shows the results of a 4mM glycine, 1ms simulation in the presence of GlyT2= 1000 and GlyT1=1000 (a) and GlyT1= 2000 (b). At the end of the 1ms simulation, there is a large overshoot of the flux where GlyT1 is operating in the glycine-releasing direction, which is proportional to GlyT2 uptake flux leading to the accumulation of astrocytically- released glycine in the presynaptic cell. This effect is not seen in the absence of GlyT1, when the flux via GlyT2 returns to zero immediately after the end of the glycine release. As the number of GlyT1 is doubled, there is an increase in the rate of glycine uptake but also in the rate of astrocytic release post-simulation, leading to higher rates of GlyT2 reuptake. According to these model parameters, glycine released during synaptic activation is uptaken in higher concentrations into the astrocyte, followed by slower release of the astrocytic glycine which is returned back to the pre-synaptic cell by GlyT2.

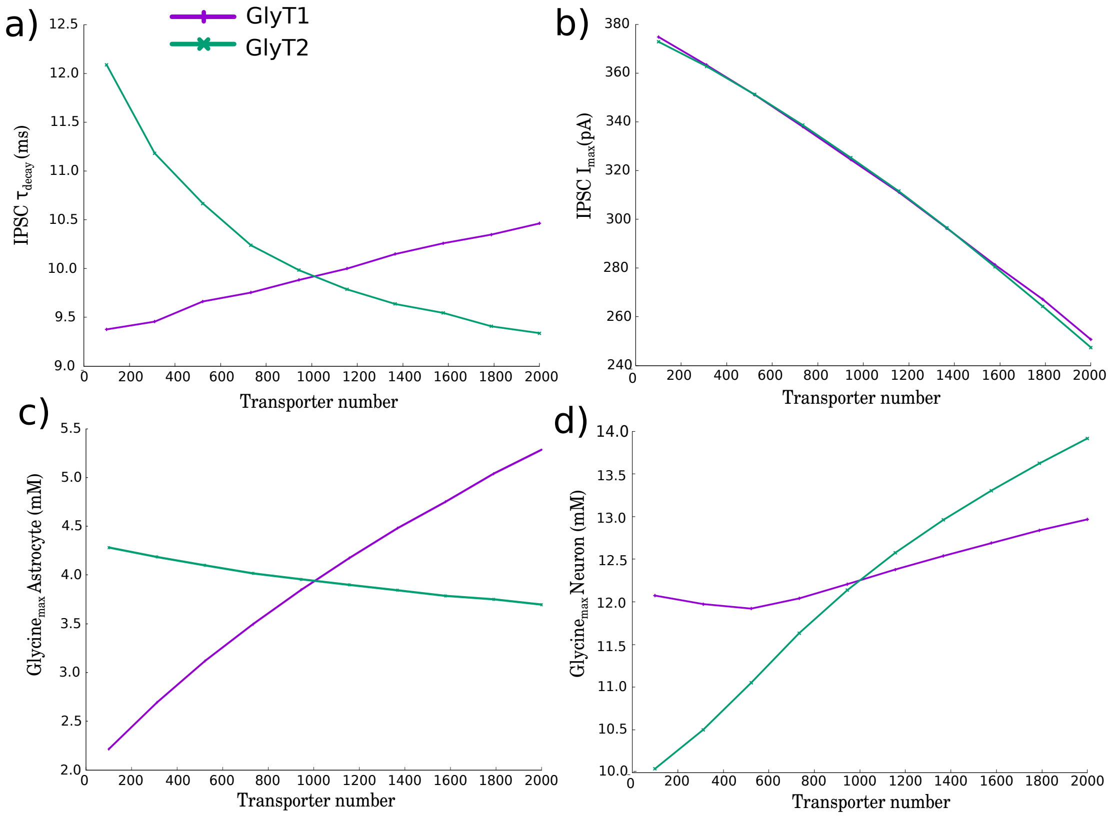


Figure 4.13: **Effect of varying GlyT numbers on the decay time constant and maximum amplitude of IPSC** The model was simulated with 4mM glycine for 1ms at varying GlyT1 and GlyT2 numbers. The resulting IPSC τ_{decay} and maximum amplitudes reached can be seen in a), b). Maximal concentrations of astrocytic and neuronal glycine reached can be seen in c), d)

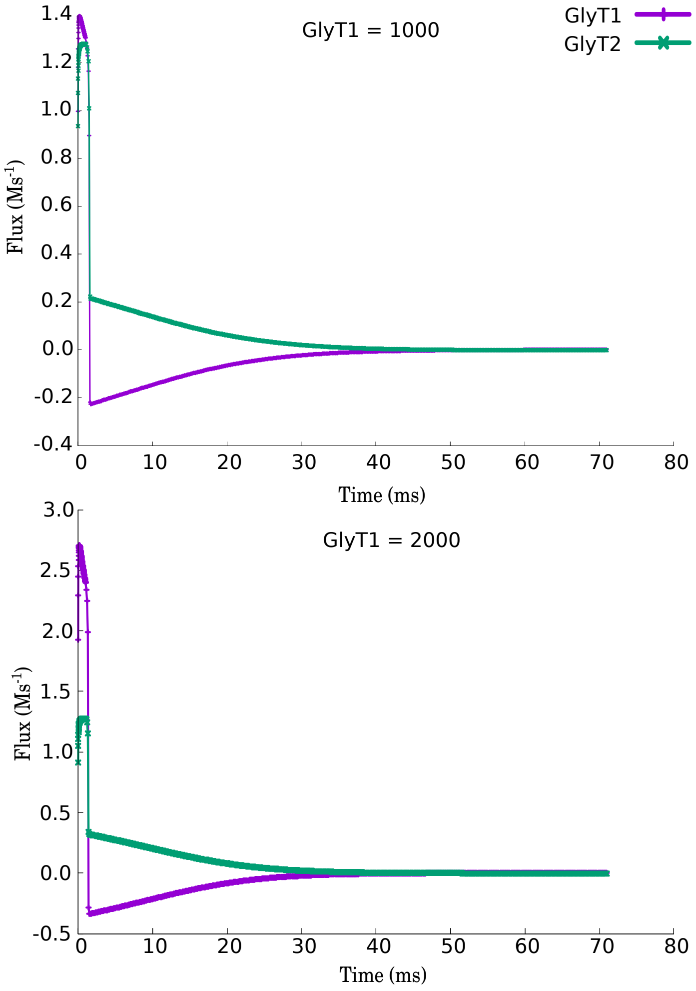


Figure 4.14: **Maximal rates of GlyT1 reversal at increased GlyT1 numbers** The model was simulated with 4mM glycine for 1ms in the presence of GlyT2=1000 and GlyT1= 1000 (a), and GlyT2=2000 (b). It can be seen that the maximal rate of GlyT1 overshoot increases at the GlyT1 number is doubled, and the rate of release is directly proportional to GlyT2 reuptake.

Glycinergic IPSCs in the presence of both transporters set to n=1000 leads to an IPSC with a τ_{decay}= 9.93ms and I_{max}= 321pA. Complete block of GlyT2, equivalent to experimental inhibition leads to an increase in the τ_{decay} to 12.47ms with an I_{max}= 377pA. In comparison, complete block of GlyT1 leads to a decrease of τ_{decay} to 9.22ms and an I_{max}= 379pA. This does not agree with experimental observations of pharmacological GlyT1 block in rat spinal cord neurons [249, 251, 254], suggesting that model rates describing the reverse reactions of GlyT1 are too fast. Thus, new model parameters were searched for such that they reflect the experimental observations. Setting the rate of GlyT1 reverse reaction, kr_{14} at values more than 5s^{-1} did not lead to IPSC τ_{decay} values lower than those in the absence of GlyT1 than in its presence, even at very high glycine and Na^{+} equilibrium constant values. Figure 4.15 shows the values of τ_{decay} IPSC calculated across a range of glycine and Na^{+} equilibrium constants at kr_{14} = 5s^{-1}, following a 4mM, 1ms glycine model simulation. Lowest value of the decay time constant was measured at glycine K_{eq} = 5.7mM, Na^{+} K_{eq} = 29mM at which the τ_{decay}= 8.7ms. Thus, these rates of GlyT1 were set in the next model simulations. In line with previous observations that the rate of GlyT2 reverse reaction is slower than that of GlyT1, the rate of the reverse GlyT2 reaction, kr_{15}, was also set to 5s^{-1}.

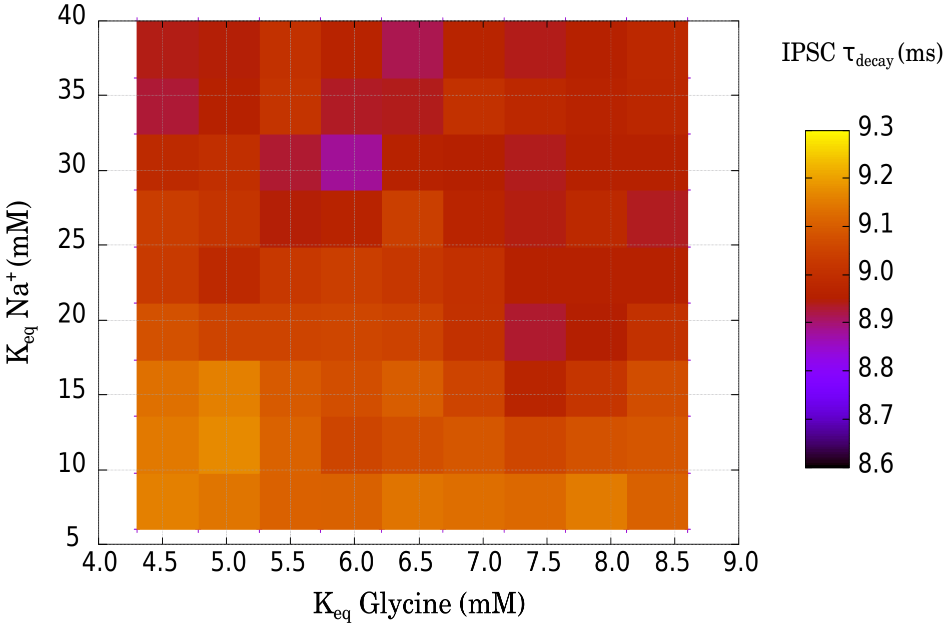


Figure 4.15: **Effects of changing GlyT1 intracellular glycine and Na$^+$ affinities on the time course of IPSC** The model was simulated with 4mM glycine for 1ms at varying intracellular glycine K_{eq} ranging from 4.3-8.6mM and Na$^+$ K_{eq} ranging from 6-40mM at the microscopic rate constant of 4s^{-1}.

The effect of changing the number of transporters on the resulting glycine peaks and IPSC amplitude and kinetics were re-calculated with the new model parameters:

Figure 4.16 shows results from model simulation with 4mM glycine for 1ms over a range of GlyT numbers 0-3000. Figure 4.16, a) shows that increasing the number of GlyT1 causes an overall fall in the τ_{decay} (range 8.5 - 7.9ms) similarly to GlyT2 which causes a fall in the τ_{decay} in the same range. There is steady fall in the maximum amplitudes of IPSC reached for both transporters, which is more pronounced in response to change in GlyT1 (range 379- 67pA) in comparison to GlyT2 (range 359- 125pA) showing that GlyT1 reuptake has a stronger effect on the IPSC (Figure 4.16, b). The highest P_{open} reached across all simulations was 0.82 when GlyT2 n= 1000 and there was a complete block of GlyT1 in comparison to the maximum P_{open} of 0.94 seen during GlyR simulation with saturating glycine in the absence of transport in Section 4.3.1. This suggests that transport reactions have a significant effect on GlyR activation during fast synaptic transmission.

Table 4.3: Final set of GlyT1 and GlyT2 kinetic parameters

Paremeter	GlyT1 model value (Ms^{-1}, s^{-1})	GlyT2 model value (Ms^{-1}, s^{-1})
k_f	20.00×10^{0}	20.00×10^{0}
k_r	5.00×10^{0}	5.00×10^{0}
K_{eq} Gly$_e$	22.60×10^{-6}	24.60×10^{-6}
K_{eq} Gly$_i$	5.70×10^{-3}	17.00×10^{-3}
K_{eq} Na$_e^+$	6.00×10^{-3}	30.00×10^{-3}
K_{eq} Na$_i^+$	2.90×10^{-2}	8.00×10^{-2}
Hill Na$^+$	1.4	2.2

Kinetic parameters used in Equation 4.6 to describe GlyT1 and GlyT2 transport, where k_f and k_r are used for the calculation of the V_f and V_r values, respectively. K_{eq} refer to the equilibrium constants of Na$^+$ and glycine in the forward (extracellular,e) and reverse (intracellular, i) reaction directions.

The change in the shape of IPSC results from significant changes in the τ_{decay} and maximum amplitude of glycine transients, which range from 0.02- 0.34ms and 0.6- 2.9mM (Figure 4.16, c), d). The relatively large effects on glycine peaks contrast with the small effects on the IPSC τ_{decay} suggesting that the dynamics of GlyR activation play a major role in shaping the resulting IPSC. The unstable fall in the τ_{decay} of IPSC, in comparison to the trend seen in time constant of glycine decay is likely due to the non-linear trends in GlyR activation due to varying glycine transients. As expected, maximal concentrations of astrocytic and neuronal glycine reached increase as the number of GlyT1 and GlyT2 increase (Figure 4.16, e, f). Neuronal concentrations reach up to 14.2mM at GlyT1= 1000, GlyT2= 3000, suggesting that at high GlyT2 densities nearly all of the released neurotransmitter will be taken back up into the presynaptic cell. Overall, the concentrations of glycine taken up into the astrocyte and neuron is highly dependant on the ratios of GlyT1 and GlyT2 expressed across the synapse which compete for glycine reuptake.

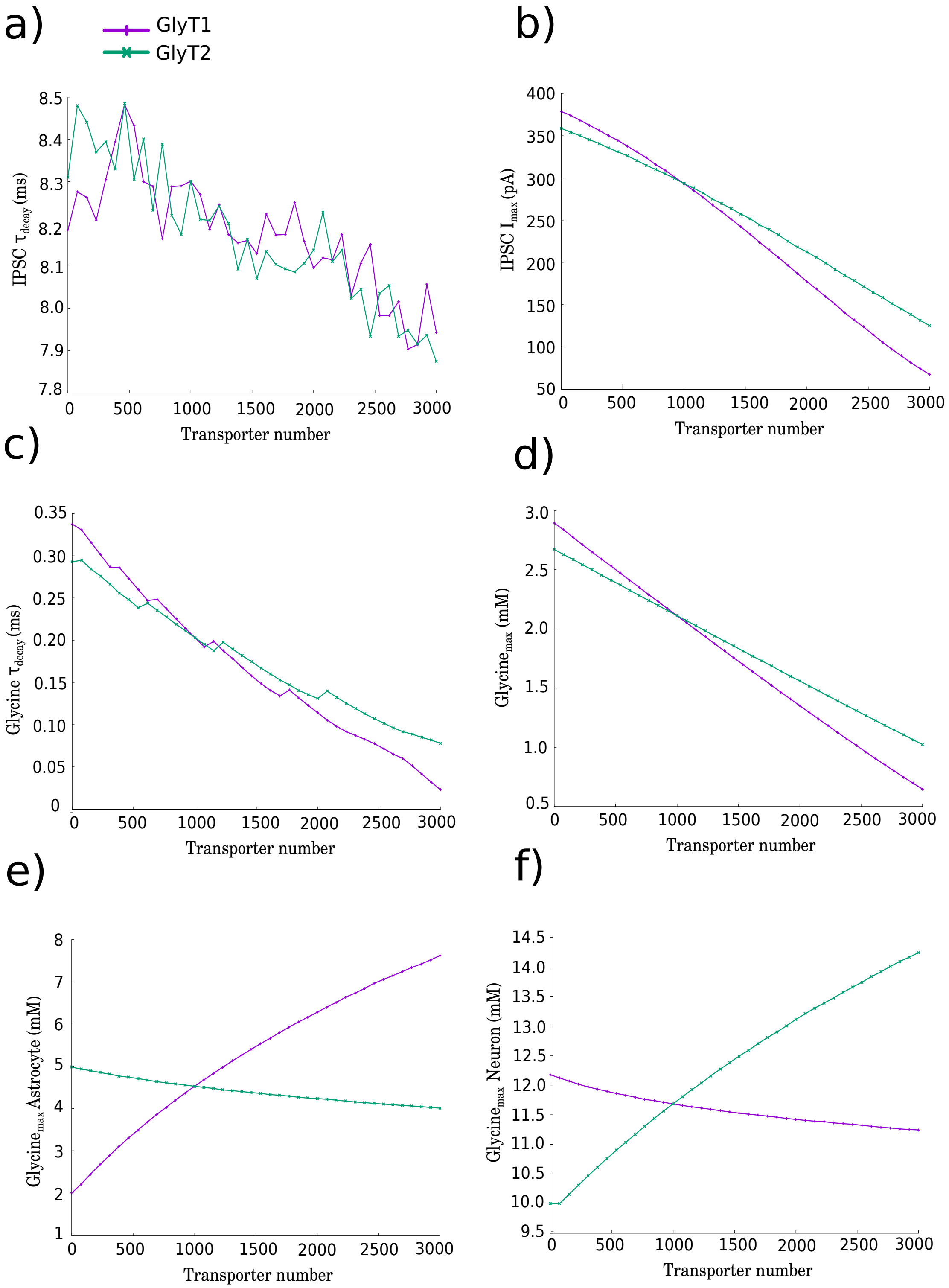


Figure 4.16: **Change in the decay time constant and maximum amplitude of IPSC at varying GlyT numbers** The effect of increasing the GlyT1 and GlyT2 was measured given model parameters as in Table 4.3. The model was simulated with 4mM glycine for 1ms at varying transporter numbers. Graphs a), b) show the effects seen in IPSC τ_{decay} and maximal amplitudes. Glycine transient $\tau_{decay}$and maximal amplitudes can be seen in c), d). The maximal astrocytic and neuronal glycine concentrations reached in each simulation can be seen in d), e).

Regulation of glycine reuptake by GlyT1 and GlyT2

Since GlyT1 operates close to its reversal potential, it has previously been hypothesized that the depolarization of the astrocytic cell may lead to non-vesicular glycine releases across the synaptic cleft [66]. This was explored in a model simulation with the intracellular Na^+ concentration raised by 25mM, an estimated concentration rise during cell depolarization [255]. The model simulations were performed under an assumption that the astrocytic and neuronal glycine removal processes are significantly slower than the kinetics of GlyT1 and GlyT2, ie. the system is closed. Figure 4.17 shows results of model simulation with 4mM glycine for 1ms in the presence of 17mM (a, c) and 42mM (a, c) astrocytic Na^+ , when the sink processes for astrocytic (r16) and neuronal (r17) glycine are set to zero, allowing for the accumulation of glycine inside the compartments. Simulation of the model in the presence of 17mM astrocytic glycine resulted in a glycine transient reaching a maximum concentration of 2.1mM with a τ_{decay}= 0.2ms that lead to a maximum P_{open}= 0.64 and IPSC with a maximum amplitude of 293pA with a τ_{decay}= 9ms. These results are very similar in the presence of 43mM astrocytic glycine, where glycine amplitude equalled 2.13mM with a τ_{decay}= 2.1ms, leading to a P_{open}=0.65 and an IPSC amplitude of 296pA with a τ_{decay}= 8.9ms. Thus, raising the astrocytic Na^+ does not have an effect on the GlyR activation which happens at a fast time scale. However, it can be seen from Figure 4.17, a), b), that there is a significant difference in the size of the GlyT1 reverse flux which happens after the fast glycine accumulation in the astrocyte following neurotransmission. Rate of GlyT1 reversal is proportional to glycine reuptake via GlyT2, thus increasing astrocytic Na^+ affects the final concentrations of neuronal glycine under the assumption that the processes which consume glycine in the astrocyte are slower than the kinetics of GlyT1 reverse transport. The maximal astrocytic and neuronal glycine concentrations reached at Na^+ = 17mM are 4.6mM and 12.7mM, respectively suggesting that even in the presence of baseline Na^+ concentrations GlyT1-released glycine is taken back up into the pre-synaptic cell. Raising the astrocytic Na^+ to 43mM leads to maximal astrocytic and neuronal glycine concentrations of 4.5mM and 13.8mM, thus although the maximal GlyT1 uptake following neurotransmission is unaffected, the concentration of neuronal glycine is increased.

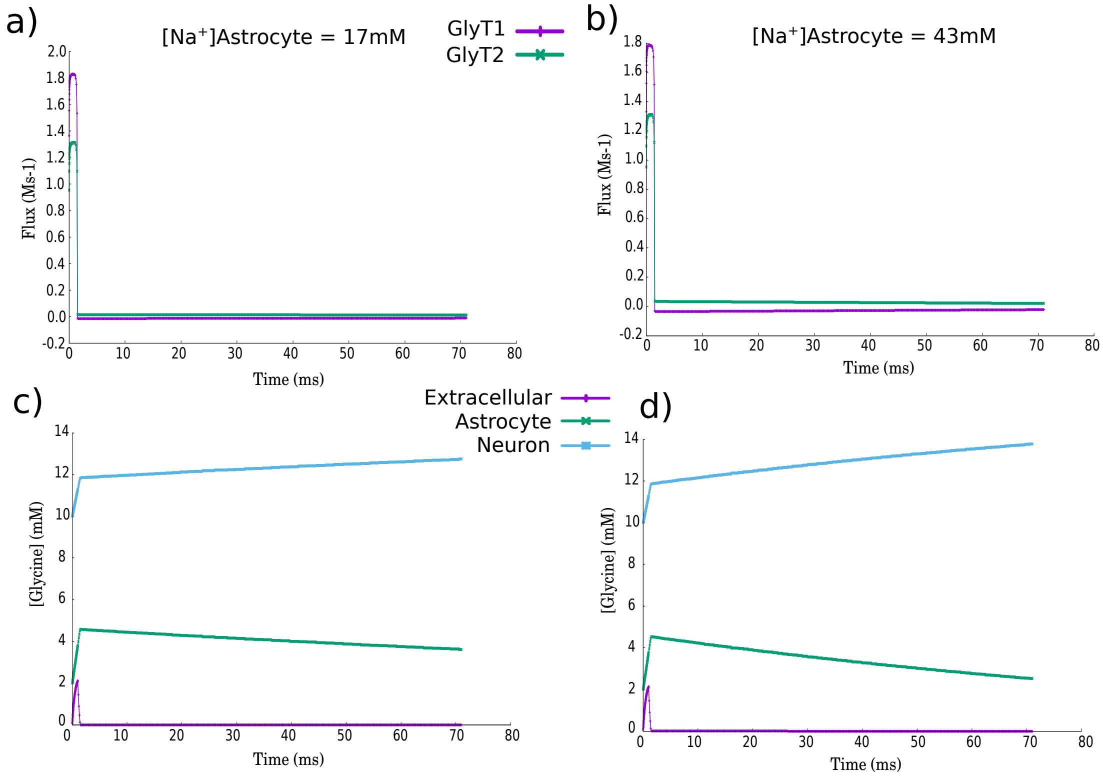


Figure 4.17: **Rate of GlyT1 reverse transport across different astrocytic Na$^+$ concentrations** Graphs a), c) show flux carried by GlyT1 and GlyT2 during a 4mM glycine simulation for 1ms at astrocytic Na$^+$ = 17mM, and the resulting intracellular accumulations of astrocytic and neuronal glycine. Graphs b), d) show results from the same model simulation at astrocytic Na$^+$ = 43mM.

4.4 Discussion

Glycine receptor model

Understanding of events which take place at the molecular level upon ligand binding to a receptor is important in elucidating how protein structure relates to function. Development of experimental techniques which allowed for the collection of current recordings via single channel receptors, and subsequently theory for the analysis of the data to model receptor activation has greatly advanced our understanding of receptor gating.

In Section 4.3.1, the overall features of the GlyR scheme activation arising from its topology and associated kinetic parameters were analysed. Although sensitivity analysis is concerned with model results at steady state, the trends seen for the the response coefficients over a concentration range followed those of the total sum of flux via each reaction in a dynamic simulation. This is not suprising, given that the equilibrium constant is a ratio of the microscopic rate constants for the forward and reverse reaction which themselves can be thought of as the number of events per second, or the probability of an event taking place. Regarded as probabilities, it is intuitive to think that the concentration of *any* receptor species will be most sensitive to those reactions (events) which have a higher probability of taking place.

These methods led to the identification of the most likely events taking place during receptor activation. Although the explanation for the protein activation provided by the MWC theory imposes the existence of all global conformations, R F and O in all non-, single-, double- and triple-liganded conformations, the likelihood of the receptor residing in each state differs.

According to the model results, openings of the receptor from a single-liganded state which would be expected to take place in low, non-saturating glycine concentrations is very rare. The opening of the receptor will happen mostly from triple- liganded states, as seen from the high total flux carried by reactions FBind3 and Open3, whereas openings from a double-liganded state mostly take place at below half-saturating glycine concentrations of 50μmol. At around 30μmol glycine, there is a crossover of the response coefficient trends for two reaction pairs, Flip1 and FBind2, and FBind3 and Open3. This would suggest that the activation of the GlyR below 30μmol will proceed with the binding of one molecule to state R, followed by an isomerization of the protein into the F state, to which another one or two molecules will bind leading to receptor opening. Above 30μmol, the likely route of activation is the binding of two glycine molecules to the R state, followed by its isomerization to the F state and subsequent opening. At all glycine concentrations, reactions RBind3 and Flip3 are descibed by very low response coefficients and carry little flux. In contrast, reaction FBind3 carries large flux across all concentrations. This suggests that the binding of the third glycine molecule will nearly always happen in the global F conformation. As the equilibrium constant is related to the standard Gibbs free energy of the reaction, it is the energy provided by glycine binding that leads to the stabilisation of the protein in each state, thus the binding of two

glycine molecules is predicted to provide the energy needed for the isomerization of the protein.
In the following section, the model was reduced to a three state MWC equation which was used to reproduce steady state full model results. Reduction of the scheme led to the estimation of the isomerization equilibrium constants between the global conformations states R and F and equalled $\approx 11\times10^3$ as well as states F and O which equalled 3.13. In addition, it allowed for the estimation of glycine affinity to the open conformation which equalled 2.07×10^{-6}M. According to these values, in the absence of ligand the majority of the receptor will be found in the R conformation. Once receptor isomerizes, there is a 65-fold increase in glycine affinity to the F conformation but only a 4-fold change in affinity from the flipped to the open conformation. This suggests that the majority of energy provided by glycine binding is used for the first step of receptor conformational change into the F conformation, from which the second isomerization steps leading GlyR opening is relatively less energetically costly. Given the previous results from sensitivity analysis, the binding of two glycine molecules will provide the energy needed for the conformational 'flip'.
Before the aplication of the MWC to GlyR by Burzomato [144], the model was applied to the nicotinic receptors (AChRs) by Auerbach [256] who used mutant-receptor constructs to show the applicability of the model to explain the differential effects of mutations on the ligand affinity *versus* global protein isomerization, as well as to estimate the rates of AChRs openings in the absence of ligand. As stated by the authors, the knowledge of the different MWC parameters may aim in the engineering of receptor function by combining mutations the effects of which can be quantitatively measured. Apart from the advantage of model reduction, a desirable feature in the area of mathematical modelling, the derivation of the MWC equation for the GlyR may prove useful in the identification of the effects of mutations, as well as design of experiments. The equation provides a clear separation between parameters which affect ligand binding (c, d) and global isomerizations of the protein (M, L). The effects of mutations are usually predicted from their AA position based on the crystal structures of receptors. However, in the absence of this information the equation could potentially be used during curve fitting to predict the effect of a ligand, inhibitor, or a mutation has on the receptor. This would require the collection of enough data point for the concentration-response curves such that the fitting of the numerous equation parameters is feasible.

Glycine synapse model

The time course of the neurotransmitter transient in the synapse is a result of many factors, including the maximal concentration and length of release from the vesicle, the geometry of the synaptic cleft and density of the solution affecting the rate of diffusion, the expression and locations of the transporter proteins and others. In Section 4.3.2, the model was simulated with peaks of glycine varying in concentration and duration of release in the absence of transporters in order to see how presynaptic release affects the IPSC time course. Previous estimates of neurotransmitter peak and length of persistence in the synapse from glutamatergic synapses resulted in values of 1-3mM with a decay time constant of up to 1.2ms [257, 258]. A more recent estimation of glycine concentration across inhibitory synapses by Beato [249] provided a range of values for concentrations between 2.2-3.5mM glycine with time decay constant of 0.6-0.9ms which was the starting point for model simulations. Varying the vesicular glycine concentrations and duration of release in the model had no effect on the decay time constant of glycine, only the maximal amplitudes reached. Since IPSC are a result of the occupancy of the open states calculated from the kinetic model of GlyR there is a hyperbolic increase in the maximal amplitudes and decay time constants of the IPSC upon GlyR saturation. The model shows that vesicular glycine concentrations of 2-5mM released for longer than 1ms would lead to IPSC which change only by $\approx 10\%$ in amplitude and even less in the speed of decay. On the other hand, in the region of durations of neurotransmitter transient less than those determined experimentally (1ms) the resulting IPSC significantly vary depending on the length and the initial concentration, agreeing with the observations that receptor activation happens under non-equilibrium [248]. The condition of non-equilibrium activation allows for a greater flexibility in the regulation of GlyR currents such as dynamic changes in transporter expression, or modulation via modification such as phosphorylation which affect the maximum rate of transport [259].

It has been previously proposed that the process of diffusion is faster than active transporter reuptake and accounts entirely for the clearance of glycine [260]. However, numerous studies have since shown that pharmacological inhibition of transporters in spinal cord tissues reduces IPSC, suggesting that active reuptake significantly affects the glycine transients experienced by the GlyR. In order to determine the effects of GlyTs on IPSC, realistic estimates for the kinetic parameters describing their reuptake were searched for.

Two models varying in their complexity have been previously published for GlyT1 [252, 261] and a more recent study describing the kinetics of both transporters under the same experimental conditions was published by Erdem et al. [219]. They provide invaluable information regarding transporter activation due to glycine, and the current study aimed to bring together kinetic parameters estimated in these studies as well as those from the electrophysiological analysis of currents via GlyTs [66], highlighting further systemic constraints on these parameters that emerge from the interactions between these proteins and the effect it has on the IPSC.

Affinity of glycine to GlyT2 in the intracellular-facing conformation must be low enough such that glycine release is not seen even at high neuronal concentrations $\geq$ 10mM, as predicted to be required for successful competition of glycine with GABA for the import into the vesicle by the low-affinity vesicular transporters [66, 244, 262]. The model results show that values of Na^+ and glycine equilibrium constants which satisfy this condition must be more than two orders of magnitude bigger than the affinities of the substrates to the extracellular-facing conformations.

In a similar manner, the higher limit on the rate of reverse transport by GlyT1, which is the summed rate of all of the protein conformational transitions including substrate binding as well as isomerization could be estimated. According to model results, this value must be at least 4-fold slower than that of the forward transport otherwise the reversal of flux stimulated by large glycine astrocytic concentration increased during neurotransmission will lead to a longer glycine transient in the synapse, and so IPSC which is not seen experimentally. The lower intracellular Na^+ binding affinities to GlyT1 estimated are in line with the model by Cherubino et al. [261] but disagree with the predictions made by Aubrey et al. [252]. In contast, lower intracellular glycine binding affinities to GlyT1 align with the model by Aubrey et al. [252] but disagree with the other model. This is because the kinetic schemes described in these studies must follow microscopic reversibility, which is not imposed by the phenomenological Michaelis Menten equation used in the current study. Detailed analysis of GlyT kinetics by Erdem et al. [219] confirm that the affinities of both substrates are lower to the intracellular-facing transporter states who explained this feature as a result of substrate binding cooperativity.

Estimation of the kinetic parameters for GlyTs allowed to look at the competition between the transporters for glycine reuptake. The model results show that GlyT1 uptakes a higher proportion of released glycine than GlyT2. If the processes which remove glycine from the astrocyte such as its metabolic breakdown by the GCS are much slower than the kinetics of GlyT1 reversal, fast accumulation of astrocytic glycine leads to its release from the astrocyte back into the synaptic cleft. The astrocytically-released glycine is taken back up into the presynaptic neuron by GlyT2 where it is repackaged into the synaptic vesicles. Whether this could happen *in vivo*, and the importance of GlyT1-mediated refilling of the neuronal glycine concentrations depends on the expression of each transporter, as well as proximity of these proteins.

Apart from slow, long term glycine releases following glycine neurotransmission, glial GlyT1 glycine release may occur due to cell depolarization, as shown by the model results. The possibility of non-vesicular release of neurotransmitter was initially described by Attwell et al. [255], and the relevance of GlyT1 reversal in the activation of NMDAR receptors in glutamatergic synapses suggested by Roux and Supplisson [66]. Later on it was shown that glycine can be released from Bergmann glial cells upon membrane depolarization [263] as well from astrocytes stimulated with dopamine [253]. As astrocytic release of glycine via GlyT1 could have an advantage in the modulation of NMDAR across glutamatergic synapses, the relevance of astrocyte depolarization in the inhibitory neurons of the lower brain is less known. Expression patterns of glycine biosynthetic enzymes seen in Chapter 4 show that astrocytes and oligodendrocytes are the main locations of initial glycine production. Since glycine is used not only as a neurotransmitter, but is important in the formation of glutathione needed for redox balance in neuronal cells, one carbon metabolism regulation and others, it is likely that *de novo* biosynthesis is required in which case GlyT1 reversal could play a central role in supplying neuronal glycine. Since the expression of SLC6A9 encoding GlyT1 was high not only in astrocytes, but also in oligodendrocytes , the depolarization of both of these types of cells may be important in glycine homeostasis. These results could shed a light on the implication of long-term pharmacological block of GlyT1 across inhibitory synapses- if GlyT1 releases are necessary for GlyT2- mediated accumulation of glycine in the pre-synaptic neuron, long-term effects of persistent inhibition could overall decrease glycinergic inhibition by emptying the required glycine pools. This should be taken into account during the design and dosage of GlyT1 inhibitors which are developed for the treatment of numerous conditions, as related to the location of GlyT1 expression [264].

Point mutation at the position Ser 407 in GlyT1 was identified as novel cause for in a child. In both GlyT1, as well as 46% sequence identical dopamine transporter serine at position 407 is conserved, and in the latter was found to be located at the sodium and chloride binding pocket [80]. It was surmised in the original publication that replacement of amino acid at this site changes the chemical properties of the protein, resulting in the disrupted functioning of the transporter. Indeed, replacement of this amino acid in wild-type human GlyT1 expressed in *X.* oocyte lead to inhibiton of currents elicited by GlyT1, as described in Chapter 5. Setting the number of transporters in the glycine synapse model to zero, equivalent to the functional knockout of GlyT1 was seen to significantly increase the maximum concentration of glycine reached, as well as lenghten its persistance in the synapse, leading to a longer GlyR activation as seen by the increase in the IPSC decay time constant. GlyT1 knockout mouse models have previously shown that severe disruption of glycine clearance leads to sustained inhibition in the respiratory centers located in the brainstem [265], causing a disruption in the development of respiratory patterns as seen in the GlyT1 mutant patient.

Application of the synapse model allowed for the confirmation of the effects of GlyT1 mutations on the resulting IPSC at the synapse scale. The model could be extended further to quantify the effects of mutations which may reduce the affinities of GlyT to glycine without causing complete inhibition. In addition, transporter equations could be expanded by terms describing inhibitor binding and kinetic constants measured experimentally can be further integrated into the model. Thus, the model forms a starting point which may be expanded upon to predict how the effect of different changes in one component affect glycinergic neurotranmission overall.

4.4.1 Conclusion

In summary, the major insights of this chapter relate to single protein GlyR activation upon glycine binding, and those pertinent to the regulation of glycine concentration across the whole glycinergic synapse. Sensitivity analysis of the GlyR scheme gave insight to the sequence of events upon GlyR activation, predicting the different activation pathways at non- and saturating glycine concentrations. In addition, this led to the discovery of GlyR receptor occupancies, showing that some receptor states such as the triple bound resting state will be hardly visited at any concentrations.

Furthermore, it was shown that the model can be reduced to fewer equations which accurately predict time-dependant GlyR activation, as well as one MWC-type equation which can be used to calculate steady state occupancies, and therefore IPCS. The parameters defining the MWC-type equation are also clearly split between those which define substrate binding, and those which relate to global protein isomerization as a result of GlyR composition. Furthermore, parameters which are very difficult to obtain experimentally such as the binding of glycine to the open state were determined which may provide insightful in probing the effects of mutations as well as activator and inhibitor molecules on GlyR activation (Section 4.4).

A holistic model which combines the results of experimental kinetic studies on single proteins was also established. The systemic constraints resulting from the model give an estimation on the glycine transporter kinetic parameters. It was shown that in order to keep glycine concentrations high enough for successful vesicle refilling in the neuron, the affinity of glycine to GlyT2 in the intracellular facing state has to be two-fold lower than that in the extracellular facing state. A similar finding is reported for the speed at which glycine binds the GlyT1 in different states. The rate at which GlyT1 binds glycine to GlyT1 must be four-fold slower to the intracellular facing state than that to the extracellular facing state, otherwise the glycine rapidly accumulating in the astrocyte would compete with the depleting extracellular glycine, causing GlyT1 to release astrocytic glycine back into the synapse and thus lenghtening the time course of IPCS (Section 4.4).

Under the assumption that metabolic processes sinking astrocytic glycine are slower than GlyT1 kinetics, it is possible that GlyT1-mediated astrocytic release of low concentrations of glycine may play a role in the refilling of vesciles in the synapse, which could be a mechanism for ensuring high neuronal concentrations required. This would take place at a longer time scale than the fast events of vesicular release upon activation, and the concentrations should be low enough such that they don't cause GlyR activation.

It was also shown that depolarization of the astrocyte, in the presence of assumed 2mM concentrations of intracellular glycine would cause GlyT1 to rapidly release glycine across the synapse. Whether this mechanism is important across the inhibitory synapses for the role of neuronal refilling, metabolism, or other functions is a topic for further study.

Lastly, the model results show that the inhibition of GlyT1, whether resulting from mutations such as that causing the novel NKH phenotype (Chapter 5) or via a pharmacological block would lead to a higher concentration of glycine and its longer persistance in the synaptic cleft, leading to longer GlyR activation. These findings can provide insights to the mechanism of GlyT1 inhibition and its resulting phenotype, such as the persistant over-inhibition of respiratory centers in patients.

The synapse model relies on experimentally determined parameters of different proteins which were collected under varying conditions, which is a major limitation of this study. The model also does not show include other transporter proteins which may be regulating glycine concentrations, such as the Asc-1 transporter which has recently been reported to regulate glycinergic neurotransmission

[266]. Lastly, the equations do not take into account the space dimension, meaning effects such as synapse geometry, locations of the proteins, and the diffusion processes are not represented. In the future state, the model can be expanded upon via additional reactions and the space dimension.

Chapter 5

Effects of Ser 407 mutations on the GlyT1 transporter

5.1 Introduction

Glycine transporter 1 (GlyT1) regulates inhibitory neurotransmission by active reuptake of glycine into glial cells wrapping the glycinergic synapses. This ensures rapid clearance of glycine from the synaptic cleft (Introduction, Section 1.2.5). The importance of GlyT1 for glycinergic signalling is supported by the finding that loss of astrocytic GlyT1 results in a lethal phenotype in newborn mice due to an over-inhibition of respiratory centers in the brain stem [265]. A similar phenotype has been reported in a child presented with nonketotic hyperglycinema (NKH) by Alfadhel,M. et al. [80], who identified a missense Ser 407 → Gly mutation in the GlyT1. Since nearly all cases of NKH are a result of variants in the genes encoding the catabolic enzymes involved in glycine degradation,the glycine cleavage system (GCS) [267], the GlyT1 mutation discovered by Alfadhel,M. et al. [80] is a novel cause for NKH.

GlyT1 is a member of the SLC6 Na^+ /Cl^- -dependent neurotransmitter transporters family which also includes transporters for other amino acids such as taurine, as well as neurotransmitters dopamine, serotonin and others [65]. All members of the SLC6 family utilize the high electrochemical gradients across biological membranes to couple the transport of Na^+ and Cl^- to ligand translocations. Much of what is known about the structure of these transporters as well as the locations of ligand and co-substrate binding sites is based on the resolution of the crystal structure of the bacterial leucine transporter (LeuT) which shows a high degree of sequence similarity with mammalian SLC family members [1] in selected regions. Later, the crystal structure of the *Drosophila melanogaster* dopamine transporter (DAT) was obtained which shares 46% sequence identity with the GlyT1, providing additional information regarding the topology of the mammalian transporters. Overall, the SLC6 are composed of 12 transmembrane helices (TM) with intracellular N- and C-termini and a large extracellular loop 2 (EL2) containing a pair of cysteine residues which form a disulfide bridge and residues containing glycosylation sites [268]. The 12 TM regions buried within the lipid bilayer contain the site of ligand binding, and form two inverted bundles of five TM regions which move against one another during transport [1, 269].

Serine at the position 407 in GlyT1 is located in the transmembrane helix (TM7) region. It is a conserved residue across human GlyT1, GABA, dopamine as well as the serotonin transporters [1], and is located in a cavity binding Na^+ and Cl^- ions. It is surmised that Ser407 mutation leads to a disruption of transporter function by affecting its co-substrate binding affinities.

In this chapter, an attempt was made to characterize the effects of Ser407 on the transporter function. In the first instance, currents via the wild type GlyT1 (wtGlyT1) expressed in *Xenopus laevis* oocytes were recorder to establish a positive control. Introduction of Ser 407 → Gly, and Ser 407 → Tyr mutations led to complete inhibition of currents through GlyT1. Lastly, in order to characterise the effect of mutation a FLAG epitope tag was introduced into the extracellular loop 2 (EL2) of wtGlyT which also caused complete abolishment of currents as compared to the wild type.

Result of GlyT1 functional disruption caused by the Ser407 mutation was incorporated into a kinetic model of the glycinergic synapse as described in Chapter 4, confirming that the mutation leads to an increase in the maximum glycine peak reached across the synapse, which persists in the synapse for a longer time period, increasing GlyR activation.

5.2 Materials and methods

5.2.1 Molecular Biology

Materials

All restriction endonucleases (RE), enzymes and buffers were purchased from New England Biolabs (USA). The E.N.Z.A. Plasmid DNA Kit used for plasmid elution from competent *E.*coli was purchased from Omega Bio-Tek (USA). Plasmid elution from ligation/RE reactions was done using the Monarch PCR & DNA Cleanup Kit purchased from New England Biolabs (USA). Transcription of cDNA was performed using the mMessage mMachine kit purchased from Thermo Fisher Scientific (USA). All oligonucleotide primers were purchased from Eurofins Genomics (UK).

Cloning of the wild type GlyT1 sequence into the pSUNOT vector

Human glycine transporter 1 cDNA (GenBank ID:BC156979.1) was obtained from Source Bioscience (UK). The sequence was amplified in a PCR cycle with a forward primer containing the

XbaI recognition site and a reverse primer containing the EcoRV recognition site. The subcloning primer sequences can be seen below:

Primer	5'-3' oligonucleotide sequence
Forward	TAA TCC TAG AGG CAC CAT GAG CGG CGG AGA CAC GCG GGC TGC
Reverse	GGC GCG ATA TCT TAT ATC CGG GAG TCC TGG AGG CGG CTG GA

The reaction mixture was set up as follows:

Component	Volume(μL)
GlyT1 template (50ng/μL)	1
Forward primer (125ng/μL)	1
Reverse primer (125ng/μL)	1
dNTPS(2mM)	5
Pfu buffer(10X)	5
Dimethyl Sulfoxide (DMSO)	3
High Fidelity *Pfu* enzyme	1
Nuclease free water	33
Total	50

Next, a PCR cycle was ran with the following conditions:

Step	Number of cycles	Temperature (°C)	Time
1	1	95	2 min
2	15	95	30 s
2	↓	60	1 min
2		72	3 min
3	1	72	10 min

The resulting product was digested with restriction enzymes XbaI and EcoRV and subcloned into the pSUNOT vector using a standard protocol for T4 DNA ligase (Promega), using a 3:1 pSUNOT vector to GlyT1 insert ratio. pSUNOT is an in-house modified PCI (Promega) expression vector. In order to confirm the subcloning was successful the resulting material was sequenced (Eurofins). The resulting pSUNOT plasmid containing the wild type GlyT1 insert can be seen in Figure 5.1.

Plasmid amplification

Plasmids were amplified in competent *E.coli* cells (X-Gol, Stratagene). Briefly, competent cells were thawn on ice, then plasmid cDNA material was added and cells were returned to ice for 30min. Next, cells were heat-shocked at 42°C for 42s, and 0.5mL of Luria Broth (LB) was added

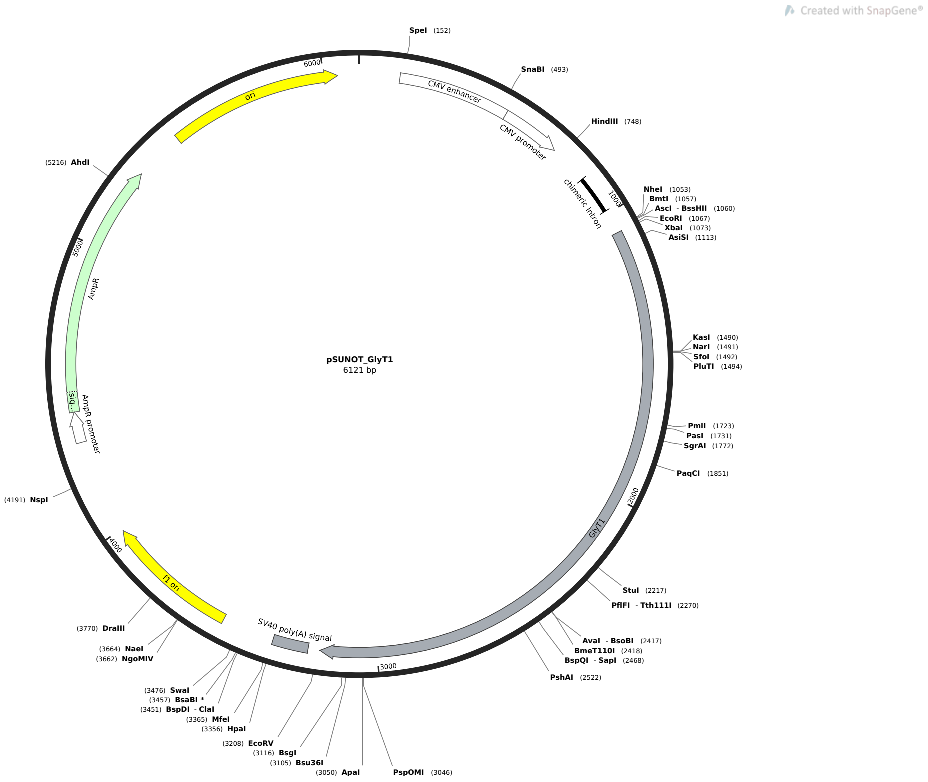


Figure 5.1: **pSUNOT-GlyT1 plasmid map** The wtGlyT1 was subcloned into the pSUNOT vector downstream the CMV promoter site enhancing the transcription of the target gene. The total plasmid size was 6121 base pairs including the 2122 base pairs wtGlyT1 located between the XbaI and EcoRV restriction enzyme sites.

to the eppendorf which was incubated at 37°C for 45mins. The cells were plated on LB-agar plates containing ampicillin at 100μg/mL and left to grow overnight for 16h in a 37°C incubator. The next day colonies were selected from the plate using a sterile pipette tip and inoculated into 10mL of Terrific Broth (TB) media (Thermo Fisher Scientific, USA) containing ampicillin at 100μg/mL. Cells were grown overnight for 16h in a shaking incubator at 37°C . The next day tubes were spun at 3000rpm for 8min, supernatant removed, and plasmid was extracted from the pellet using the E.N.Z.A. Plasmid DNA Kit (USA) according to the kit instructions. Plasmid was collected on the filter of a QIAprep column and eluted with 50μL of nuclease-free water (NFW). The resulting concentration measured in a NanoDrop One spectrophotometer (Thermo Fisher Scientific, USA). Plasmid was kept at -20°C until needed.

Synthesis of RNA for *X.*oocyte injection

DNA was linearised downstream of the insert site using a unique SmiI restriction enzyme site in a reaction incubated at 37°C for 2.5h. Next, the linear plasmid was eluted with the Monarch PCR DNA Cleanup Kit and was used to generate capped RNA with the mMessage mMachine kit (Thermo Fisher Scientific,USA), according to the manufacturers instructions. Reaction mixture was left to incubate at 37°C for 2.5h after which 1μL of DNAse was added for 30mins to remove the template cDNA material. Next, 25μL of lithium chloride and 30μL of NFW was added to the resulting RNA and the reaction mix was left overnight at -20°C. The next day, sample was spun at -4°C for 30mins and the supernatant carefully removed. One millilitre of 70% EtOH was added to the sample to wash the RNA pellet, after which the sample was re-spun at 4°C for 20 minutes. EtOH was carefully removed and the RNA pellet resuspended in 15μL of NFW. The concentration of RNA in the resulting sample was measured in a NanoDrop One spectrophotometer (Thermo Fisher Scientific, USA) after which the sample was aliquoted and stored in a -80°C freezer until the day of injection.

Introduction of GlyT1 mutations

Nucleotide substitutions at position Ser407 were introduced into the wild type GlyT1 according to the QuickChange Site-Directed mutagenesis kit (Stratagene, The Netherlands). The reaction mixture was set up as follows:

Component	Volume(μL)
DNA template (50ng/μL)	1
Forward primer (125ng/μL)	1
Reverse primer (125ng/μL)	1
dNTPS(2mM)	5
Pfu buffer(10X)	5
Dimethyl Sulfoxide (DMSO)	3
High Fidelity *Pfu* enzyme	1
Nuclease free water	33
Total	50

Next, a PCR cycle was ran with the following conditions:

Step	Number of cycles	Temperature (°C)	Time
1	1	95	1 min
2	15	95	30 s
2	↓	55	1 min
2		68	6.5 min
3	1	68	6.5 min

In order to remove methylated parent DNA, 1μL of Dpn I was added at the end of the cycle and the reaction was left to incubate at 37°C for 1 hour. Two mutations were introduced overall, Ser 407 → Gly and Ser 407 → Tyr. The primers used to introduce each mutation can be seen below:

Mutation	5'-3' oligonucleotide sequence (forward primer)
Ser 407 → Gly	ACC AAC TGT GCC ACC GGC GTC TAT GCT GGC TTC
Ser 407 → Tyr	ATC ACC AAC TGT GCC ACC ACT GTC TAT GCT GGC

In order to confirm the mutations were successful the resulting material was sequenced (Eurofins).

Insertion of the FLAG epitope into the EL2 of GlyT1

Insertion of the FLAG- tag was performed via the PCR-driven overlap extension method following the protocol described in Heckman and Pease [270]. Overlap PCR was ran in two steps using the primers below:

Mutation	5'-3' oligonucleotide sequence
Primer A	ATT ATC TAG AGC CAC CAT GAG CGG CGG AGA C
Primer B	TTT ATC GTC ATC GTC TTT GTA ATC CAG TGA CAG CTT CAG CAC GT
Primer C	GAT TAC AAA GAC GAT GAC GAT AAA GAC ATT GGG AAC TTT GG
Primer D	ATA ATG ATA TCT TAT CCG GGA GTC CTG GAG GCG G

In the first step, two reaction mixtures containing primers A and B, and primers C and D were set up as follows:

Component	Volume(μL)
DNA template (50ng/μL)	1
Primer 1 (A or C) (125ng/μL)	1
Primer 2 (B or D) (125ng/μL)	1
dNTPS(2mM)	5
Pfu buffer(10X)	5
Dimethyl Sulfoxide (DMSO)	3
High Fidelity *Pfu* enzyme	1
Nuclease free water	33
Total	50

The PCR reaction was set up with the following conditions:

Step	Number of cycles	Temperature (°C)	Time
1	1	95	1 min
2	15	95	30 s
2	↓	55	1 min
2		68	6.5 min
3	1	68	6.5 min

Products from the resulting PCR reactions containing overlapping fragments were purified and used as the template in the next PCR step. The reaction was set up as follows:

Component	Volume(μL)
AB template (78ng/μL)	0.6 (50ng)
CD template (64.8ng/μL)	0.77 (50ng)
Forward primer A (125ng/μL)	1
Reverse primer D (125ng/μL)	1
dNTPS(2mM)	5
Pfu buffer(10X)	5
Dimethyl Sulfoxide (DMSO)	3
High Fidelity *Pfu* enzyme	1
Nuclease free water	32.6
Total	50

The PCR was set up with the following conditions:

Step	Number of cycles	Temperature (°C)	Time
1	1	95	1 min
2	31	95	30 s
2	↓	55	1 min
2		72	1.5 min
3	1	72	10 min

Product from the second PCR reaction was digested with XbaI and EcoRV and subcloned into the pSUNOT vector using a standard protocol for T4 DNA ligase (Promega). In order to confirm the insertion was successful the resulting material was sequenced (Eurofins). The resulting plasmid was used used to make RNA for injection in the *X.*oocytes, as described previously.

5.2.2 Electrophysiology

Materials

ORG Mg^{+} solution used in oocyte preparation contained 82mM NaCl, 2mM KCl, 2mM MgCl and 5mM HEPES at pH 7.6. ORG Ca^{2+} solution in which oocytes were stored contained 82mM NaCl, 2mM KCl, 2mM $CaCl_2$ and 5mM at a pH of 7.6 supplemented with Antibiotic/Antimycotic Solution x100 (10,000 units penicillin, 10 mg streptomycin and 25g amphotericin B per mL) and amikacin purchased from Merck (Sigma-Aldrich,UK). Collagenase Type I was purchased from Merck (Sigma-Aldrich,UK). SOS buffer used to fill the bath chamber during the electrophysiology recordings contained 100mM NaCl, 2mM NaCl, 1.8mM $CaCl_2$, 1mM $MgCl_2$ and 5mM HEPES. Glycine was obtained from Sigma-Aldrich (UK).

Preparation and microinjection of *Xenopus* oocytes

Xenopus oocytes were purchased from Xenopus one (Chicago, USA) or Xenopus Express (France). *Xenopus* care and experimental procedures were in accordance with the UK Home Office regulations and were approved by the Animal Use Committee of Oxford Brookes University. Briefly, *X.laevis* were housed in the animal house of Oxford University in black tanks filled with dechlorinated water ($\geq$ 15 L per toad) that was kept in a temperature-controlled room (18°C). The animals were kept under a fixed 12 h light/dark cycle. Frogs were fed twice a week with amphibian food pellets. Ovarian lobes were harvested from aneasthesised *Xenopus* (MS222 solution -0.5% tricane, pH 7.4) via surgical laparotomy and prepared as described below.

The ovarian lobes were washed in ORG Mg^{+} solution in a shaker at 150rpm for 10 minutes. Next they were placed in a Petri dish containing ORG Mg^{+} and the oocytes were separated into smaller groups of 5-6, after which they were placed back on a shaker for a second wash. Oocytes were then defolliculated via the addition of ORG Mg^{+} solution containing collagenase at 2mg/mL to the Petri dish, which was left on the shaker at 150rpm for approximately 1 hour. After the removal of the extracellular membrane, healthy stage IV- V oocytes were selected based on their appearance and placed into ORG Ca^{2+} solution containing Antibiotic/Antimycotic Solution x100 and amikacin, then incubated at 17°C.

The next day, unhealthy oocytes were removed from the plate and healthy oocytes selected for microinjection with RNA material. The needles used for microinjection were made from Drummond glass capillaries (Sartorius,UK) and pulled in a Narishige PC-10 micropipette puller (Narishige, Japan). The needle was filled with mineral oil and the RNA material, and afterwards loaded onto a Nanoject II microinjector (Drummond, USA). Oocytes were injected into the center of the vegetal pole. The volume of RNA injected varied from 50.5nL- 119.6nL in the attempt to increase transporter expression, however this factor did not affect amplitude of currents recorded. Next, the oocytes were placed separately into 96 well-plates in ORG Ca^{2+} supplemented with kanamycin at 2 g/ml, 1x Antibiotic Antimycotic Solution and 5% horse serum, pH 7.6. and incubated at 17°C. Solution was changed every day until the day of experiment. Oocytes were used for recording 5-10 days after transfection on the day of the highest transporter expression.

Electrophysiological recordings of GlyT1 currents

Whole cell currents through individual oocytes expressing GlyT1 were measured via TEVC (Introduction, Section 1.3.5). The oocyte was placed in a 30μL recording chamber (Digitimer Ltd, UK), bathed in SOS buffer (100mM NaCl, 2mM NaCl, 1.8mM $CaCl_2$, 1mM $MgCl_2$ and 5mM HEPES) and voltage clamped at -80mV. Electrodes used for the recording were made from borosicate capillary glass (Harvard Apparatus, GC 150 TF) and pulled in a vertical electrode puller (Narishige PP-83). Next, the electrodes were filled with 3M KCl and connected to the Oocyte Clamp OC-725C

(Warner Instruments, USA) instrument upon which the impedance was checked. Only electrodes with an impedance between 0.2- 2M Ω were used in the subsequent recording. Once the V_m of the oocyte was stable, SOS buffer containing glycine at various concentrations was used to perfuse the chamber via a gravity driven perfusion system. In order to add or remove solutions, a valve allowing the flow of compound into the chamber was manually manipulated.

Data analysis of the concentration response curves (CRC)

All data plots and analysis were performed in the GraphPad Prism software. In order to determine glycine affinity the maximal response was measured over a range of glycine concentrations. The amplitude of response was normalised to the response generated by the maximum effective concentration (I_{max}) of 1000μM. Resulting values were input into GraphPad Prism software which fitted the values to the one-component Hill equation of the form:

$$I = \frac{I_{max}}{1 + (EC_{50}/[A])^h} \tag{5.1}$$

Where EC_{50} is the concentration of agonist which induces 50% of the I_{max}, h is the Hill coefficient estimation of cooperativity, and [A] is the agonist concentration.

Concentration response curve was plotted based on cumulative results from a number (n) of independent experiments chosen on the basis of the EC_{50} 95% confidence interval (CI.) range values.

5.3 Results

Response of the wild type GlyT1 to glycine

Five glycine concentrations ranging from 1-1000μM were applied via bath perfusion of oocytes expressing the wtGlyT1. Representative responses can be seen in Figure 5.2. In the first instance I_{max} was determined via application of 1000μM glycine. Maximal responses in individual eggs were varied depending on the day of injection, as well as amongst oocytes injected on the same day which is a reflection of the high variation in protein expression at the cell membrane. Oocytes were perfused with different solutions every 3 to 6 minutes in order to ensure the removal of residual compound between recordings thus allowing for the recovery of the transporter. Figure 5.3 shows glycine concentration response curve from n=6 individual oocytes. Fit of the Hill equation to the curve resulted in glycine EC$_{50}$ = 25.7μmol (18.9-34.0 CI.) and H_{coeff} = 0.95 (0.64-1.25 CI.) which is in line with the previously reported value of EC$_{50}$ $\bar{2}$2.60μmol and H_{coeff} = 0.78 [66].

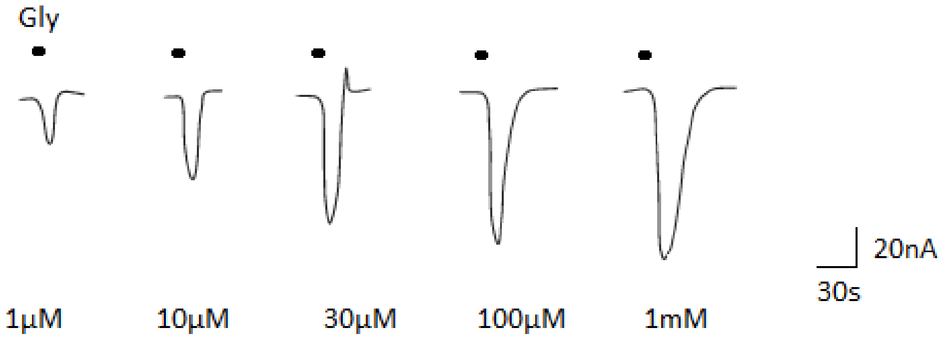


Figure 5.2: **Amplitudes of the wtGlyT1 responses to glycine at varying concentrations** Representative Cl^{-} currents via the wtGlyT1 at concentration of glycine applied ranging from 1μmolto 1mM. Maximal responses upon 1mM glycine application were approximately 60nA, but ranged from one experiment to another as determined by membrane transporter expression.

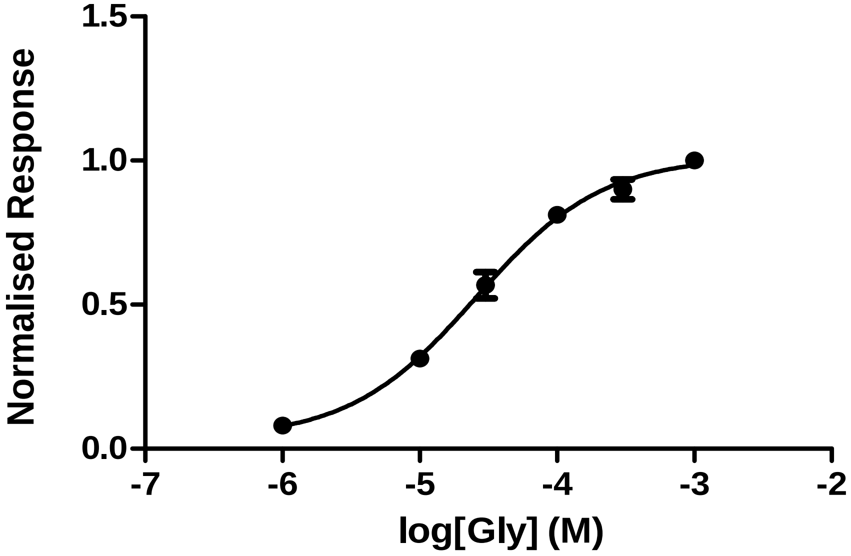


Figure 5.3: **Concentration response curve of glycine binding to the wtGlyT1** Fits of the Hill equation to the curve resulted in an $EC_{50} = 25.7\mu$mol (18.9-34.0 CI.) and $H_{\text{coeff}} = 0.95$ (0.64-1.25 CI.).

Effects of S407 mutations on GlyT1

After the establishment of the wild type GlyT1 transporter EC_{50} an attempt was made to determine the effect of point mutations at position Ser407 on the GlyT1 currents. In the first instance, mutation Ser 407 → Gly was introduced. No currents could be detected in oocytes expressing this construct. Thus a second, less dramatic mutation Ser 407 → Tyr was introduced which produced ambiguous results. Currents collected through the Ser 407 → Tyr mutant in oocytes injected on different days could not be resolved due to the small amplitude of the I_{max}, ranging from 5-20nA. Since the baseline noise in the recordings varied up to 10nA, it was not possible to establish a stable baseline current under these conditions.

Insertion of FLAG epitope tag

The absence of Cl^- currents through oocytes injected with the mutant transporter does not univocally resolve the effect of the mutation. It remains underdetermined whether the mutation causes a functional knockout of GlyT1 inserted in the membrane, or whether it has an effect on the correct protein folding and delivery to the membrane in the first instance.

Electrophysiological recordings cannot resolve this issue which requires the visualization of the protein in the membrane. Standard techniques utilizing synthetic or animal-derived antibodies binding to the protein of interest include immunocytochemistry, wester blotting and others. However, most of these techniques require the homogenization of the cell thus the detection of intra- and extra-cellular proteins from ruptured cells. Visualisation of membrane proteins can be achieved via antibody binding to segments of proteins that protrude extracellularly from the membrane. A commonly used technique is the insertion of small peptide sequence in the protein for which strongly-binding antibodies are available, such as the FLAG epitope which is is an eight amino acids epitope sequence motif (DYKDDDDK) for which strongly binding antibodies are commonly available.

For the purpose of visualization of GlyT1 mutants in the membrane, the FLAG tag was inserted into the extracellular loop of the protein. The extracellular loop 2 (EL2) is the largest across the mammalian SLC6 family, thus to minimizes the probability of functional disruption of FLAG insertion, EL2 was chosen as the site of insert. In order to predict the location of EL2, bioinformatics tools capable of predicting the secondary structure and topology of transporters were used, as can be seen in Table 5.1.

Once the amino acid sequence composing EL2 was estimated, the location of FLAG insertion was chosen such that it minimizes the probability that the FLAG insert will have an effect on protein function. Previously recorder amino acids conserved across the family members were avoided, as well as positions close to the potential glycosylation sites (Figure 5.4). Location of insert was chosen as Asp278, since the amino acid in itself forms the beginning of the the FLAG sequence.

Insertion of the FLAG tag was performed via the PCR-driven overlap extension method following the protocol described in Heckman and Pease [270] (Materials and Methods, Section 5.2). Briefly, insertion of the sequence was performed in two PCR-reaction steps. Primers were designed such that in the first cycle, two fragments of the desired product are generated containing overlapping sections that include the insert. In the second PCR cycle, these fragments are denatured and used as a template in the following reaction. The single-stranded products hybridize at the overlapping regions, leading to the amplification of the product driven by the addition of primers flanking the 5'-3' beginning of sequence and reverse complement of the 3'-5' end of sequence. Thus, the second PCR reaction generates the desired product containing the amino acid sequence insert which can be inserted into a vector of interest.

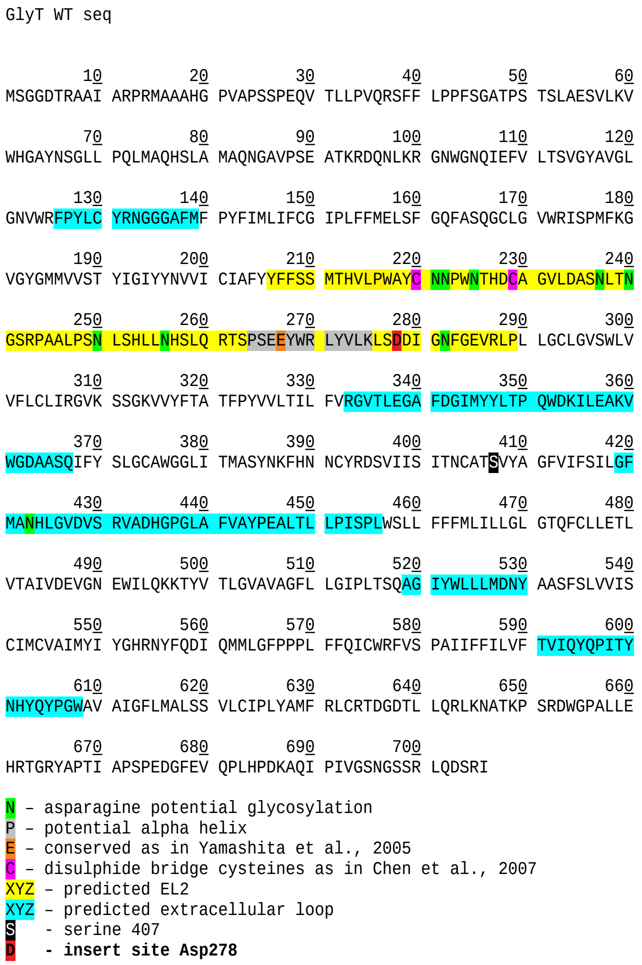
X. oocytes were injected with the GlyT1-FLAG RNA on the same day as the wtGlyT1 RNA which

Table 5.1: Bioinformatic tools used to predict the sequence of amino acids composing EL2.

Bioinformatics tool	EL2 amino acids position
TMHMM	210-285
MEMSAT	210-289
TOPCONS	205-285
Phobius	209-287
Philius	209-285
TMpred	209-289
SOSUI	210-281
Amino acid length	71-80

Sequence of amino acids composing EL2 as predicted by each tool can be seen in the right hand column.

was used as a positive control. The oocytes were testes 5-10 days post-injection however no currents could be detected via the wtGlyT-FLAG construct. Given that currents could be measured via the wtGlyT1 it is expected that the FLAG insertion caused a knockout of transporter function.



```
GlyT WT seq

        10         20         30         40         50         60
MSGGDTRAAI ARPRMAAAHG PVAPSSPEQV TLLPVQRSFF LPPFSGATPS TSLAESVLKV

        70         80         90        100        110        120
WHGAYNSGLL PQLMAQHSLA MAQNGAVPSE ATKRDQNLKR GNWGNQIEFV LTSVGYAVGL

       130        140        150        160        170        180
GNVWRFPYLC YRNGGGAFMF PYFIMLIFCG IPLFFMELSF GQFASQGCLG VWRISPMFKG

       190        200        210        220        230        240
VGYGMMVVST YIGIYYNVVI CIAFYYFFSS MTHVLPWAYC NNPWNTHDCA GVLDASNLTN

       250        260        270        280        290        300
GSRPAALPSN LSHLLNHSLQ RTSPSEEYWR LYVLKLSDDI GNFGEVRLPL LGCLGVSWLV

       310        320        330        340        350        360
VFLCLIRGVK SSGKVVYFTA TFPYVVLTIL FVRGVTLEGA FDGIMYYLTP QWDKILEAKV

       370        380        390        400        410        420
WGDAASQIFY SLGCAWGGLI TMASYNKFHN NCYRDSVIIS ITNCATSVYA GFVIFSILGF

       430        440        450        460        470        480
MANHLGVDVS RVADHGPGLA FVAYPEALTL LPISPLWSLL FFFMLILLGL GTQFCLLETL

       490        500        510        520        530        540
VTAIVDEVGN EWILQKKTYV TLGVAVAGFL LGIPLTSQAG IYWLLLMDNY AASFSLVVIS

       550        560        570        580        590        600
CIMCVAIMYI YGHRNYFQDI QMMLGFPPPL FFQICWRFVS PAIIFFILVF TVIQYQPITY

       610        620        630        640        650        660
NHYQYPGWAV AIGFLMALSS VLCIPLYAMF RLCRTDGDTL LQRLKNATKP SRDWGPALLE

       670        680        690        700
HRTGRYAPTI APSPEDGFEV QPLHPDKAQI PIVGSNGSSR LQDSRI
```

N – asparagine potential glycosylation
P – potential alpha helix
E – conserved as in Yamashita et al., 2005
C – disulphide bridge cysteines as in Chen et al., 2007
XYZ – predicted EL2
XYZ – predicted extracellular loop
S - serine 407
D - insert site Asp278

Figure 5.4: **Amino acids spanning extracellular loops of wtGlyT1 as predicted by bioinformatics tools** Prediction of the amino acids sequence composing EL2 is highlighted in yellow. FLAG insertion was chosen at position Asp278 in order to avoid possible interaction with residues that form secondary structures, as predicted by the bioinformatics tools used (Table 5.1). Glutamate at position 267 is a conserved residue across members of the SLC6 family [1]. The two cysteine residues are a potential site of a disulfide bridge formation as has been found in the equivalent EL2 cysteines in *Drosophila melanogaster* dopamine transporter, where mutation of either or both amino acids produced non-functional transporters [2].

5.4 Discussion

In this chapter the effect of the Ser407 mutation in the GlyT1 was characterised. Introduction of a mutation in this amino acid position completely abolished currents via the transporter expressed in *X.* oocyte. In the first instance, the polar serine was replaced with glycine which is a non-conservative mutation. In order to establish whether a less dramatic mutation would lead to reduced function, serine was replaced with tyrosine which has similar biochemical properties. However, in both cases no currents could be recorder via GlyT1 mutants. Given that the maximal amplitudes of currents were very small (maximal values of $\approx$ 200nA) even via the wtGlyT, it is possible that small currents, indistinguishable from the background noise were carried via the tyrosine mutant, thus further investigation is required for complete characterisation of the Ser407 mutation.

Transmembrane helix 7 contains a number of amino acids which are conserved across the family. In the bacterial leucine transporter, Na^+ binding occurs in the unwound regions of TM1 and TM6 halfway across the lipid bilayer and is stabilised by residues in TM7. Crystal structure of the mammalian DAT sharing higher sequence similarity to GlyT1 further revealed that the Cl^- ion binding site is located close to the Na1 binding site and is coordinated by residues in TM6, TM2, and Ser356 of TM7 which is the conserved serine residue aligning with GlyT1 Ser407. A detailed study including comprehensive sequence alignment between the prokaryotic LeuT and numerous other members of the mammalian SLC6 transporters including GlyT1 showed that the residues comprising the agonist binding sites are highly conserved across all members, supporting the applicability of LeuT structure in predicting Ser407 site function [271]. Importance of amino acids comprising TM7 is further supported by previous directed mutagenesis studies of related transporters. Replacement of the Ser356 residue to glycine or alanine in the DAT expressed in COS cells significantly reduced the uptake of dopamine [272]. Similarly, uptake assays in the rat serotonin transporter expressed in HeLa cells show that mutation in the equivalent TM7 serine causes a 40% or more reduction in serotonin transport [273]. Altogether, these findings support the results in the current study suggesting a crucial role of Ser407 in transporter function. Replacement of the polar serine residue with less polar glycine in the hydrophobic TM7 region likely disrupts GlyT1 function directly via changing the chemical environment in the co-substrate binding pocket.

In the second section, a small peptide FLAG epitope was inserted into the large extracellular loop (EL2) of the transporter in order to further examine the effect of the mutation. In ordert to establish a positive control against which currents via the FLAG-tagged mutants could be established, the wtGlyT1-FLAG construct was expressed in *X.* oocyte and compared against the wtGlyT1. It was found that FLAG insertion at the position Asp278, located at the end of EL2 sequence near TM4 in the lipid bilayer caused a complete inhibition of currents via the transporter. What could be the cause of functional disruption?

As compared to the highly conserved TM7, amino acids of EL2 show a lower sequence conservation between members of the SLC6 family [271]. Since TM7 comprises the substrate binding site, and all the SLC6 members transport Na^+ and Cl^- alike, the high conservation is expected in comparison to EL2, which is involved in post-translational modifications of the transporter. The EL2 of LeuT is relatively small in comparison to mammalian SLC6 transporters and does not share high sequence similarity. In the mammalian SLC6, EL2 contains residues important for proper transporter folding, trafficking and surface expression.

The SLC6 family members are subject to various post-translational modifications, with the EL2 identified as an important site for N-linked glycosylation, forming bonds with the asparagine residues [268]. In the GlyT1 transporter EL2 there are eight potential glycosylation sites (Figure 5.4, green) out of which Asn282 resides closest to the Asp278, location of the FLAG insert. However, only four of these sites were postulated by Smith et al. [274] to reside in the EL2 region feasible of undergoing glycosylation (Arn237, Arn240, Arn250 and Arn256). This was later confirmed by Olivares, L. et al. [275] who showed that progressive mutagenesis of these residues lead to reduced transport activity via GlyT1. In their study, enzymatic deglycosylation resulted in reduced activity of the transporter-thus N-linked glycosylation is not indispensible to GlyT1 function. As determined by GlyT1 biotinylation, mutagenesis revealed a lower expression of transporter at the surface membrane, showing that it is important in protein trafficking and stability. Reduced protein expression at the cell membrane was also found for unglycosylated serotonin [276] and the GABA transporters [277]. As Asn282 is located closer to the TM4 region in the hydrophobic membrane, it is less likely that it would be involved in N-linked glycosylation, but may nevertheless affect the conformation of EL2, affecting glycosylation of the GlyT1.

It is also possible that the location of insert was too close to the TM4 region- predictions from the bioinformatics tools as in Table 5.1 show that the last amino acid comprising EL2 is likely glutamine at position 285, however the SOSUI prediction estimates it at the glycine residue at position 281. Furthermore, the Asp278 is located close to a conserved glycine (Gly284) residue across human glycine, dopamine, norepinephrine, serotonin and GABA transporters [271]. Since glycine does not have a side chain, it has higher confromational flexibility and is therefore often found in loops, and regions of rigid helix formations breaks [278]. Conservation of Gly284 and its equivalents in the other human SLC6 members may imply an important role of this amino acid in the formation of the tertiary structure of EL2.

Altogether, these findings suggest a crucial role of the serine residue at position 407 in the human GlyT1, which likely takes part in the binding of Na^+ and Cl^- ions in the ligand binding cavity. Future investigation of the effect of this mutation could involve visualisation of the protein to establish correct folding and trafficking to the membrane. This could be achieved via insertion of the FLAG epitope in a position which does not disrupt transporter function, such as an alternative extracellular loop known to be less important in protein trafficking and stability. Alternative methods of protein identification which may cause a smaller interference, such as surface biotinylation to assess relative expressions of the mutant and wtGlyT1 may also be used.

5.4.1 Conclusion

In summary, the major insights of this chapter relate to the effects of a novel NKH-causing Ser 407 → Gly mutation. It was found that replacement of the serine with amino acids glycine and tyrosine leads to the inhibition of currents at all concentrations tested. Furthermore, insertion of the FLAG epitope, which was attempted in order to establish the mechanism via which the mutation disrupts transport, showed that FLAG insertion at position Asp 278 in the EL2 also abolishes currents. Further investigation is required to elucidate the consequence of these genetic manipulations on the GlyT1 structure and expression in the cell membrane.

The sensitivity of the TEVC set up used for the recording of GlyT1-mediated currents determines the resolution with which signal is captured. Oocytes injected with the wild type GlyT1 mRNA showed maximal transporter expression after at least five days post injection, thus maintaning the viability of the cell in the media until the day of experiment was a difficulty. Many oocytes were discarded, reducing the pool of material to test experimentally. In addition, since the transporter tranfers a net one charge upon one transport cycle, the current amplitudes recorder were low, at roughly 200nA maximal responses collected. The majority of recordings measured at maximal glycine concentration (1mM) applications were approximately 60nA. At these low currents, good sensitivity of the instrument is required, such that the background noise can be clearly distinguished from responses at sub-saturating glycine concentrations. This needs to be taken into account in the future experiments utilising *X.*oocytes as an expression system for glycine transporters.

Chapter 6

Conclusions

The dual function of glycine in the CNS as a central metabolite, as well as an inhibitory neurotransmitter across different brain regions establish its importance in human health. Metabolic pathways of glycine synthesis in this organ are subject to the specificity of brain metabolism in comparison to other tissue, as determined by the restricted substrate availability as well as the brain's high energetic needs. Since glutamate and GABA dominate neurotransmission in the brain overall, relatively less research has been dedicated to the study of glycine in the CNS. Recent interest in the understanding of pathways of its metabolism has been sparked by numerous studies showing increased expression of the enzymes involved in serine/glycine interconversions in a subset of cancers, creating a novel target for the development of chemotherapeutic agents.

F Pertinent to glycine's action as an inhibitory neurotransmitter, regulation of its concentrations in the brain *in situ* has been identified as a potential new method for the treatment of numerous conditions, as proven by the efforts towards the development of several GlyT1 inhibitors so far. However, the effect of these compounds on glycinergic transmission is difficult to measure in the live animal, and although *in vitro* experimental studies of individual proteins regulating glycine concentrations may provide detailed descriptions of their activation mechanism, they need to be combined in order to look at how the interactions between them determine GlyR-mediated IPSC across a synapse.

Work presented in this thesis employed computational modelling techniques with the overall aim of increasing the understanding of glycine metabolism, as well as glycinergic neurotransmission relayed across a synapse. Electrophysiological methods were employed in order to elucidate the effects of NKH-causing Ser 407 $\rightarrow$ Gly mutation in the GlyT1, which was integrated into the kinetic model of the glycinergic synapse to assess its effects on glycinergic currents.

6.1 Overview of the result

The main aim of Chapter 2 was the construction and curation of a metabolic model representing CNS metabolism, built on the template of a more generic human tissue reconstruction, Recon2.2 [145]. In order to contextualise the model to represent brain metabolism, experimentally recorder transport rates of substrates across the BBB, in combination with the well-established brain OGI parameter were used to estimate the energy and biosynthetic demands of the brain. The response of the model to changing ATP and biomass demands were assessed by looking at the effects on the OGI, which establishes a model system capable of directly relating variations in the OGI to substrate consumption and metabolic pathway utilization. Estimations of the background metabolic requirements allowed for the elucidation of glycine biosynthetic pathways. Increase in the demand for glycine was seen to slightly decrease the rate of glucose consumption, as well as the conversion of pyruvate to lactate showing that the mitochondrial SHMT2-catalyzed reaction contributes to ATP production in the mitochondria.

Heterogeneity in the gene expression across different CNS cell types suggests a variability in the metabolic pathways utilized by these cells. In Chapter 3, a method for the integration of a single-cell RNA-seq dataset was developed for the purpose of assessing the possible differences in the metabolic pathways of energy and glycine production. The motivation for the choice of the dataset was two-fold- in the first instance, the dataset contains the genetic signature of both excitatory and inhibitory neurons, with the latter most likely to be glycinergic. In addition, it contains data from astrocytes and oligodendrocytes- as previously mentioned, the majority of studies of brain metabolism have focused on glutamatergic, excitatory neurons and the associated astrocytes, and relatively scant research has been conducted looking at the more rare cell types contained within the dataset. Lastly, previous studies assessing the expression of genes involved in glycine metabolism suggested compartmentalization of these pathways across different cell types, which was addresses in the current study.

Further to Chapter 2 results, pathways of ATP production in each cell type were evaluated. In contrast to the ANLS hypothesis, genetic expression as well as FBA applied to the model weighted by the median expression of each cell type suggest high glycolytic activity in the neuron, and a preference for glucose uptake accompanied by lactate release. Suprisingly, when both substrates were present, astrocytes showed a preference for lactate uptake, whereas oligodendrocytes utilized glucose. Although these results require experimental validation, they show that application of scRNA-seq data to metabolic modelling may provide useful in the investigation of the metabolism of less studied cell types, such as oligodendrocytes.

In the background of an ATP demand, glycine biosynthesis occured via the mitochondrial SHMT2 from serine and glutamine in the neuronal cell types, and glutamate in glial cell types. Expression of the SHMT2 enzyme was more widespread than SHMT1 across the dataset, confirming its primary role in glycine biosynthesis. Both SHMT2 as well as the GCSH gene expressions were normally distributed across different cell types, which is in contrast to previous findings that located the GCS to astrocytes only [47]. Genes encoding enzymes catalyzing the phosphorylated pathway of serine biosynthesis from glucose were mostly expressed in astrocytes and oligodendrocytes, strongly suggesting that *de novo* glycine biosynthesis is indeed compartmentalized across different CNS cell types.

Genetic expression of enzymes related to the glyoxylate pathway of glycine biosynthesis was minimal in the brain, consolidating results from previous studies which suggested that this pathway does not contribute to glycine biosynthesis in this organ [39].

Knockout of the enzymes involved in glycine metabolism in the model showed the metabolic flexibility of the cell, which could provide useful in the design of a strategy for knockout targets aiming to disrupt aberrant glycine biosynthesis, such as in cancer. In support of previous experimental studies in cell cultures [236], model results suggest that serine catabolism to glycine may support mitochondrial ATP production by providing the ETC with NADH, which could partially explain the upregulation of serine consumption and large glycine releases in a subset of cancer types, especially in the context of hypoxia. This could establish the role of this pathway as not only to provide proliferating cells with the biosynthetic precursors, but also with an associated high demand for energy which dominates cellular metabolism. However, further experimental validation is required in order to confirm whether serine can support high cellular ATP demands.

High rates of serine uptake and glycine export were seen in the model which utilized isoleucine and glutamine as biomass was increased, with the reverse GCS acting as a sink for the nitrogen atoms generated in the pathways of amino acids breakdown. This could suggest that glycine production may occur as a method of nitrogen assimilation in cancers utilizing amino acids for energy production, especially in hypoxic areas of tumours in which the clearance of metabolic by-products is disrupted.

Chapter 4 and 5 results are concerned with glycine's role as an inhibitory neurotransmitter. In Chapter 4, detailed analysis of the GlyR scheme from Burzomato [144] led to the discovery of the likely routes of receptor activation at sub-saturating and saturating glycine concentrations, increasing the understanding of receptor dynamics. Derivation of the MWC-type equation from the GlyR scheme allowed for the reduction of the full model to only one equation, clearly separating the kinetic parameters related to receptor gating and affinity for substrates, which may aid in the elucidation of the effects of mutations on receptor function, and provide a quantitative measure for the differences in the energy provided by the binding of partial agonists [256].

Furthermore, a kinetic model integrating the detailed GlyR scheme, as well as reactions describing the speed of reuptake by GlyT1 and GlyT2 was developed, allowing for the estimation of the upper limits on the rates of reverse transport by each glycine transporter. The model establishes a system in which theoretical predictions of the effects of protein mutations, inhibitor binding, and ionic concentration changes across compartments can be studied. For instance, the depolarization of the astrocyte expressing GlyT1 was shown to induce glycine release across the synapse, which was rapidly uptaken by GlyT2 into the neuronal cell.

Suprisingly, the expression of the SLC6A9 gene encoding the GlyT1 as seen in Chapter 3 was significantly high in oligondendrocytes, suggesting that this cell type may play a role in glycine metabolism and/or neurotransmission, although whether this is the case requires further investigation.

The model was also used to assess the effects of varying GlyT numbers on glycine transient concentration and time course, as well as the resulting IPSC, showing that much larger changes can be seen in the former in comparison to the latter. The decay time constant of resulting IPSC was decreased upon increasing GlyT numbers, but changes were less dramatic than that on the decay time constant of glycine, confirming that IPSC persistence is mostly dependent on the deactivation kinetics of the GlyR. This implies that the regulation of the glycinergic neurotransmission via pharmacological inhibition of GlyTs is limited to a range within which variation in GlyT-mediated reuptake has an effect on GlyR deactivation.

Experimental results presented in Chapter 5 confirmed that the novel NKH-causing mutation Ser 407 → Gly in the human GlyT1, as reported by Alfadhel,M. et al. [80] causes a knockout of the transporter function, as no currents could be detected in *X.* oocytes expressing the mutant. This result was integrated into the glycinergic synapse model to show that, as anticipated, disruption of GlyT1 function causes an increase in the maximum amplitude and decay time of the resulting IPSC via the GlyRs.

6.2 Outlook

Results presented in the current work are non-exhaustive, and both require experimental validation and provide opportunity for further research.

With respect to the metabolic model of the CNS, improvements to the method of scRNA-seq integration method can be added. Due to the physical limitations of experimental techniques owing to the low amount of genetic material present in the first instance, these datasets inherently suffer from a higher number of possible false-negative values and higher noise resulting from stochasticity of gene expression, in comparison to bulk RNA sequencing data. Up to date, integration of scRNA-seq data into metabolic models presented in this thesis has been previously attempted by one group only [279], thus the method is lacking rigorous assessment of the choices made during analysis. In the current thesis, the outcome of some of these choices on the feasibility of the objective function were assessed. For example, it was found that the difference in the results generated by the model was minimal whether the median, or the maximal value was chosen as the single expression value used in the objective function weights for a reaction catalyzed by an enzyme with multiple gene relationships. However, owing to the novelty of this technique much theoretical consideration is still needed, which is a task certainly worth pursuing. Application of the scRNA-seq data can increase our future understanding of the differences in the metabolic pathway utilization on a scale of single cells, and furthermore, models can be combined in order to look at cooperation between cell populations. Since metabolic cooperation between cells is largely dependent on the transport systems available, more attention could be given during the model curation process to transport reactions, which could also be weighted by their expression values which wasn't taken into account in the current study.

Although gene expression values analysed in the current study suggest serine/glycine metabolism compartmentalization, this hypothesis can be further tested theoretically by analysing these pathways in a multi-cellular model, weighted by the expression of individual cell types, as described previously. These results would require experimental validation, for example by looking at radiolabelled glucose incorporation rates into downstream metabolites involved in glycine metabolic pathways in cellular co-cultures. In return, such experimental data could be integrated into the metabolic model to consolidate the validity of its predictions.

Results from Chapter 3 which suggest that serine degradation to glycine via SHMT2 in the mitochondria may play an important role in supporting ATP production could be expanded upon by experimental validation in cancer cell lines, increasing the understanding of the role of this amino acids in context-specific malignancy formation. Inhibition of phosphoglycerate dehydrogenase has previously been identified as a novel cancer treatment, and was previously found to affect the activity of central carbon metabolic pathways such as the ETC and glycolysis [280], implying that this pathway exhibits regulation over cellular metabolism beyond serine/glycine biosynthesis and proliferation. The model could be used further to explain the observations from experimental studies such as the aforementioned, and supply novel hypothesis regarding the effects of serine/glycine metabolism on the overall cell metabolism. This may help in the design of future treatment strategies applicable to cancers of different human tissue cell types.

With respect to the kinetic model of the glycinergic synapse, in the first instance the model would greatly benefit from the integration of experimental measurements of the reverse transport rates by the GlyT1 and GlyT2. The values estimated in the current thesis provide a range within which these parameters fall, nevertheless are approximations deduced from the estimations of glycine concentrations within the astrocyte and neuronal compartments, which have not been directly measures *in vivo*. As ample data is available on the forward transport rates of each protein, only one study by Aubrey et al. [252] has measured the affinity of glycine to GlyT1 in the intracellular-facing protein conformation, which is more difficult to set up experimentally. Expression of each transporter in a model system such as the *X.* oocytes, and manipulation of the internal glycine and ionic concentrations would allow for the calculation of dose-response curves for these proteins to ligand in intracellular-facing states. Integration of these kinetic parameters would improve the validity of the current model, which could be further applied to the study of the effects of GlyT inhibitors in development, such as bitopertin, on the resulting glycinergic currents.

Lastly, the functional knockout resulting from the Ser 407 mutation in the human GlyT1 could be further explained via experimental validation to establish whether the mutant protein is expressed in the membrane in the first instance. This was attempted by the insertion of the FLAG epitope, however could not be confirmed since the insertion also led to functional disruption. These experiments could by continued by the insertion of the FLAG epitope in a different region of the EL2, or a different extracellular loop altogether. Nevertheless, disruption of transporter function via the FLAG insertion in itself is of interest, which if further studied could increase the knowledge of the significance of EL2 in transporter function.

Appendix A

Glycinergic synapse model ScrumPy file.

```
x_Gly= 0
Gly   = 0
GlyG = 0.002
GlyN = 0.01

## Sodium concentrations, M

x_Na   = 0.15          ## extracellular Na
x_NaG = 0.017          ## glial Na
x_NaN = 0.03           ## neuronal Na



R      = 1e-12
R_No = 100             ##receptor number

T = 1e-4
GlyT1 =   1000         ##transporter number
GlyT2 =   1000




Gly1R= 0
Gly2Fopen= 0
Gly3Fopen= 0
Gly3F= 0
Gly2F= 0
```

```
Gly1F= 0
Gly3R= 0
Gly2R= 0
Gly1Fopen= 0

Gly3DesF= 0
Gly3DesS= 0


#
##
### Gly release and point diffusion
##
#

V_GlyDiff=2.0e3
V_GlyRel= 2.0e3


GlyRel:
    x_Gly -> Gly
    (V_GlyRel* x_Gly)




GlyDiff:
    Gly -> x_Gly
    (V_GlyDiff*Gly)


#
##
### Gly binding to non-flipped receptor
##
#
```

```
V_GlyR_kf= 0.59e6
V_GlyR_kr= 300

V_RBind1= V_GlyR_kf* 3                        ##forward rate
    constants , M-1 s-1
K_RBind1 = V_RBind1/V_GlyR_kr                 ##equilibrium constant
   M taking into account the statistical factors

V_RBind2= V_GlyR_kf* 2
K_RBind2= V_RBind1/(V_GlyR_kr*2)



V_RBind3=  V_GlyR_kf
K_RBind3 = V_RBind3/(V_GlyR_kr*3)



RBind1:
     Gly + R <> Gly1R
     ~(V_RBind1, K_RBind1)

RBind2:
     Gly + Gly1R <> Gly2R
     ~(V_RBind2, K_RBind2)

RBind3:
     Gly + Gly2R <> Gly3R
     ~(V_RBind3, K_RBind3)



#
##
### Flipping of the bound receptor
##
#

V_Flip1= 1.8e2
#K_Flip1= 0.006
```

```
V_Flip2 = 6.8e3
K_Flip2= 0.378

V_Flip3 = 20.9e3
#K_Flip3 = 23.22

Flip1:
    Gly1R <> Gly1F
    ~(V_Flip1,K_Flip1)

Flip2:
    Gly2R <> Gly2F
    ~(V_Flip2, K_Flip2)

Flip3:
    Gly3R <> Gly3F
    ~(V_Flip3 ,K_Flip3)



#
##
### Gly Binding to the flipped receptor
##
#

V_FBind2= 3.0e8
K_FBind2= 1.25e5

V_FBind3 = 1.50e8
K_FBind3= 4.17e4

FBind2:
    Gly + Gly1F <> Gly2F
    ~(V_FBind2 ,K_FBind2)

FBind3:
```

```
    Gly + Gly2F   <> Gly3F
    ~(V_FBind3,K_FBind3)



#
##
### Receptor opening
##
#

V_Open1 = 4.2e3
K_Open1 = 1.235

V_Open2= 2.8e4
K_Open2 = 13.3

V_Open3= 1.29e5
K_Open3= 18.43

Open1:
    Gly1F <> Gly1Fopen
    ~(V_Open1, K_Open1)

Open2:
    Gly2F <> Gly2Fopen
    ~(V_Open2, K_Open2)

Open3:
    Gly3F <> Gly3Fopen
    ~(V_Open3, K_Open3)


K_Flip1 = K_Flip2*K_RBind2/ K_FBind2
K_Flip3 = K_Flip2*K_FBind3/K_RBind3



#
```

```
##
### Transporter uptake
##
#
## parameters from Roux and Supplisson: Glycine transporters have
    different stoichiometries



## GlyT1 parameters

kcat_GlyT1   = 20
kcatr_GlyT1 = 5

keq_GlyT1    = 22.6e-6
keq_GlyGT1   = 5.7e-3

h_NaT1   = 1.4
h_NaGT1 = 1.4

keq_NaT1   = 6e-3
keq_NaGT1 = 2.9e-2


## GlyT2 parameters

kcat_GlyT2 = 20
kcatr_GlyT2= 5

keq_GlyT2    = 24.6e-6
keq_GlyNT2   = 17.0e-3

h_NaT2   = 2.2
h_NaNT2 = 2.2

keq_NaT2   = 30e-3
keq_NaNT2 = 8e-2
```

```
GlyT1_tx:
    Gly <> GlyG
    (   ((kcat_GlyT1 * GlyT1 * T)* Gly * (x_Na**h_NaT1)) /((keq_GlyT1
       + Gly) * (keq_NaT1  + x_Na)**h_NaT1))- (((kcatr_GlyT1 * GlyT1 *
       T) * GlyG * (x_NaG**h_NaGT1)) / ((keq_GlyGT1 + GlyG) *
       (keq_NaGT1  + x_NaG)**h_NaGT1))



GlyT2_tx:
    Gly <> GlyN
    (   ((kcat_GlyT2 * GlyT2 * T)* Gly * (x_Na**h_NaT2)) /((keq_GlyT2
       + Gly) * (keq_NaT2  + x_Na)**h_NaT2))- (((kcatr_GlyT2 * GlyT2 *
       T) * GlyN * (x_NaN**h_NaNT2)) / ((keq_GlyNT2 + GlyN) *
       (keq_NaNT2  + x_NaN)**h_NaNT2))



## Current via the transporters

e= 1.6021766208e-19                 ## electron charge, C
I_GlyT1 := GlyT1_tx * e
I_GlyT2 := GlyT2_tx * 2 * e


#
##
### Removal of glial and neuronal glycine
##
#

Gly_rel= 1e1              ## adjusted slow processes
```

```
GlyGRelease:                ## sink for glial glycine

    GlyG -> x_Gly
    (Gly_rel* GlyG)



GlyNRelease:                ## sink for neuronal glycine

    GlyN -> x_Gly
    (Gly_rel* GlyN)


Closed := (R + Gly1R + Gly2R + Gly3R + Gly1F + Gly2F + Gly3F) * 1e12

Open := (Gly1Fopen + Gly2Fopen + Gly3Fopen) * 1e12

Flipped := (Gly1F + Gly2F + Gly3F) *1e12




#
##
### Electrophys calculations
##
#

gmax= 45.9e-12       ## single channel conductance, A , from Scott et
    al, 2015

V= -60e-3            ## membrane voltage, V
E_rev= -70e-3        ## reversal potential for chloride ions, V

g := gmax * (Open)

I := gmax * (Open) * (V- E_rev)              ## Current , A
```

```
PicoI := I * 10e12                              ## Current, pA

I_patch := (gmax * R_No) * (Open) * (V- E_rev)* 10e12         ##
    Current through n*GlyR, pA

I_p := Open * R_No

Sum := Open + Closed


#
##
### MWC sums
##
#

Ko = 2.07e-6
Kf = 8e-6
Kr = 520e-6
M = 3.13
L = 10823.0


MWCOpen := (1+(Gly/Ko)**3) / ( (1+(Gly/Ko)**3) + (M*(1+(Gly/Kf)**3)) +
    (L*M*(1+(Gly/Kr)**3)) )


Saturation := ((Gly/Ko)*(1+(Gly/Ko))^2 + M*(Gly/Kf)*(1+(Gly/Kf))^2 +
    M*L*(Gly/Kr)*(1+(Gly/Kr))^2) / (1+(Gly/Ko))^3 + M*(1+(Gly/Kf))^3 +
    M*L*(1+(Gly/Kr))^3)

MWCFlipped := (M*(1+(Gly/Kf)**3)) / ( (1+(Gly/Ko)**3) +
    (M*(1+(Gly/Kf)**3)) + (L*M*(1+(Gly/Kr)**3)) )

MWCResting := (L*M*(1+(Gly/Kr)**3)) / ( (1+(Gly/Ko)**3) +
    (M*(1+(Gly/Kf)**3)) + (L*M*(1+(Gly/Kr)**3)) )
```

Appendix B

List of kinetic parameters used in the glycinergic synapse model.

Reaction	Parameter	Model value (M , Ms^{-1})	Reference
Glycine receptor activation			
r1 (RBind1)	$\mathrm{V}_{\max 1}$	1.77×10^{6}	[144]
	$K_{eq\ 1}$	5.9×10^{3}	[144]
r2 (RBind2)	$\mathrm{V}_{\max 2}$	1.18×10^{6}	[144]
	$K_{eq\ 2}$	2.95×10^{3}	[144]
r3 (RBind3)	$\mathrm{V}_{\max 3}$	5.9×10^{5}	[144]
	$K_{eq\ 3}$	6.55×10^{2}	[144]
r4 (Flip1)	$\mathrm{V}_{\max 4}$	1.80×10^{2}	[144]
	$K_{eq\ 4}$	$\mathrm{K}_{eq\ 5}\times\mathrm{K}_{eq\ 2}/\mathrm{K}_{eq\ 7}$	[144]
r5 (Flip2)	$\mathrm{V}_{\max 5}$	6.80×10^{3}	[144]
	$K_{eq\ 5}$	0.0378	[144]
r6 (Flip3)	$\mathrm{V}_{\max 6}$	20.90×10^{3}	[144]
	$K_{eq\ 6}$	$\mathrm{K}_{eq\ 5}\times\mathrm{K}_{eq\ 3}/\mathrm{K}_{eq\ 8}$	[144]
r7 (FBind2)	$\mathrm{V}_{\max 7}$	3.00×10^{8}	[144]
	$K_{eq\ 7}$	1.25×10^{5}	[144]
r8 (FBind3)	$\mathrm{V}_{\max 8}$	1.50×10^{8}	[144]
	$K_{eq\ 8}$	4.17×10^{4}	[144]
r9 (Open1)	$\mathrm{V}_{\max 9}$	4.20×10^{3}	[144]
	$K_{eq\ 9}$	1.235	[144]
r10 (Open2)	$\mathrm{V}_{\max 10}$	2.80×10^{4}	[144]
	$K_{eq\ 10}$	1.33×10^{1}	[144]
r11 (Open3)	$\mathrm{V}_{\max 11}$	1.29×10^{5}	[144]
	$K_{eq\ 11}$	1.84×10^{1}	[144]
Glycine release			
r12	kf_{12}	2.00×10^{3}	[281]
Glycine diffusion			
r13	kf_{13}	2.00×10^{3}	-
Glycine Transporter 1			
r14	kf_{14}	20.00	[219, 282]
	kr_{14}	5.00	-
	K_{eq} Gly_{e14}	22.60×10^{-6}	[66]
	K_{eq} Gly_{i14}	5.7×10^{-3}	[252]
	K_{eq} $\mathrm{Na}_{e}^{+}{}_{14}$	6.00×10^{-3}	[66, 219]
	K_{eq} $\mathrm{Na}_{i}^{+}{}_{14}$	2.9×10^{-2}	
	Hill Na_{14}^{+}	1.4	[66]
Glycine Transporter 2			
r15	$\mathrm{kf}_{\_15}$	20.00	[219, 282]
	$\mathrm{kr}_{\_15}$	5.00	-
	K_{eq} Gly_{e15}	24.60×10^{-6}	[66]
	K_{eq} Gly_{i15}	17.00×10^{-3}	-
	K_{eq} $\mathrm{Na}_{e}^{+}{}_{15}$	30.00×10^{-3}	[66, 219]
	K_{eq} $\mathrm{Na}_{i}^{+}{}_{15}$	8.00×10^{-2}	-
	Hill Na_{15}^{+}	2.2	[66]
Glial and neuronal glycine removal processes			
r16	$\mathrm{kf}_{\_16}$	1.00×10^{1}	-
r17	$\mathrm{kf}_{\_17}$	1.00×10^{1}	-

Bibliography

[1] A. Yamashita, S. Singh, T. Kawate, Y. Jin, and E. Gouaux. Crystal structure of a bacterial homologue of Na+/Cl–dependent neurotransmitter transporters. *Nature*, 437(7056), 2005.

[2] R. Chen, H. Wei, E. Hill, L. Chen, L. Jiang, D. Han, and H. Gu. Direct evidence that two cysteines in the dopamine transporter form a disulfide bond. *Molecular and cellular biochemistry*, 298(1-2), 2007.

[3] Y. Chang, W. Lin, Z. Chin, C. Wang, I. Chou, H. Kuo, and F. Tsai. Nonketotic hyperglycinemia: a case report and brief review. *BioMedicine*, 2(2), 2012.

[4] R. J. Harvey and B. K. Yee. Glycine transporters as novel therapeutic targets in schizophrenia, alcohol dependence and pain. *Nature Reviews Drug Discovery*, 12(11):866–885, 2013.

[5] A. Maudsley, R. Gupta, R. Stoyanova, N. Parra, B. Roy, S. Sheriff, N. Hussain, and S. Behari. Mapping of glycine distributions in gliomas. *American Journal of Neuroradiology*, 35(6), 2014.

[6] B. Wang, W. Wang, Z. Zhu, X. Zhang, F. Tang, D. Wang, X. Liu, X. Yan, and H. Zhuang. Mitochondrial serine hydroxymethyltransferase 2 is a potential diagnostic and prognostic biomarker for human glioma. . *Clinical neurology and neurosurgery*, 154, 2017.

[7] K. Birsoy, L. Garraway, and M. Mino-Kenudson. Functional genomics reveals serine synthesis is essential in PHGDH-amplified breast cancer. *Nature*, 476(7360), 2012.

[8] H. Yoshino, N. Nohata, K. Miyamoto, M. Yonemori, T. Sakaguchi, S. Sugita, T. Itesako, S. Kofuji, M. Nakagawa, R. Dahiya, and H. Enokida. PHGDH as a key enzyme for serine biosynthesis in HIF2-targeting therapy for renal cell carcinoma. *Cancer research*, 77(22), 2017.

[9] G. DeNicola, P. Chen, E. Mullarky, J. Sudderth, Z. Hu, D. Wu, H. Tang, Y. Xie, J. Asara, K. Huffman, and I. Wistuba. NRF2 regulates serine biosynthesis in non–small cell lung cancer. . *Nature genetics*, 47(12), 2015.

[10] A. Antonov, M. Agostini, M. Morello, M. Minieri, G. Melino, and I. Amelio. Bioinformatics analysis of the serine and glycine pathway in cancer cells. *Oncotarget*, 5(22), 2014.

[11] F.A. Azevedo, L.R. Carvalho, L.T. Grinberg, J.M. Farfel, R.E. Ferretti, R.E. Leite, W.J. Filho, R. Lent, and S. Herculano-Houzel. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. . *Journal of Comparative Neurology*, 513(5), 2009.

[12] Herculano-Houzel, S. The glia/neuron ratio: how it varies uniformly across brain structures and species and what that means for brain physiology and evolution. *Glia*, 62(9), 2014.

[13] S. Jäkel and L. Dimou. Glial cells and their function in the adult brain: a journey through the history of their ablation. . *Frontiers in cellular neuroscience,*, 11(4), 2017.

[14] T. Philips and J. Rothstein. Oligodendroglia: metabolic supporters of neurons. *The Journal of clinical investigation*, 127(9), 2017.

[15] Kety, S. *Metabolism of the nervous system.* 1957.

[16] Rolfe, D. and Brown, G. Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiological reviews*, 77(3), 1997.

[17] T. Blazey, A. Snyder, M. Goyal, A. Vlassenko, and M. Raichle. A systematic meta-analysis of oxygen-to-glucose and oxygen-to-carbohydrate ratios in the resting human brain. *PLoS One*, 13(9), 2018.

[18] L. Sokoloff, M. Reivich, C. Kennedy, M. Rosiers, C. Patlak, K. Pettigrew, O. Sakurada, and M. Shinohara. The [14C] deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat 1. *Journal of neurochemistry*, 28(5), 1977.

[19] R. Shulman and D. Rothman. *Brain energetics and neuronal activity: applications to fMRI and medicine.* . John Wiley & Sons., 2005.

[20] M. Wong-Riley. Cytochrome oxidase: an endogenous metabolic marker for neuronal activity. *Trends in neurosciences*, 12(3), 1989.

[21] N. Sibson, A. Dhankhar, G. Mason, D. Rothman, K. Behar, and R. Shulman. Stoichiometric coupling of brain glucose metabolism and glutamatergic neuronal activity. . *Proceedings of the National Academy of Sciences*, 95(1), 1998.

[22] Hydén, H. and Lange, P. A kinetic study of the neuron-glia relationship. . *The Journal of Cell Biology*, 13(2), 1962.

[23] Hamberger, A. and Hydén, H. Inverse enzymatic changes in neurons and glia during increased function and hypoxia. *The Journal of Cell Biology*, 16(3), 1963.

[24] L. Pellerin and P. Magistretti. Glutamate uptake into astrocytes stimulates aerobic glycolysis: a mechanism coupling neuronal activity to glucose utilization. *Proceedings of the National Academy of Sciences*, 91(22), 1994.

[25] P. Magistretti and I. Allaman. A cellular perspective on brain energy metabolism and functional imaging. *Neuron*, 86(4), 2015.

[26] D. Lovatt, U. Sonnewald, H.S. Waagepetersen, A. Schousboe, W. He, J. Lin, X. Han, T. Takano, S. Wang, F.J. Sim, and S.A. Goldman. The transcriptome and metabolic gene signature of protoplasmic astrocytes in the adult murine cortex. *Journal of Neuroscience*, 27 (45), 2007.

[27] R. Gruetter, E.R. Seaquist, and K. Ugurbil. A mathematical model of compartmentalized neurotransmitter metabolism in the human brain. *American Journal of Physiology-Endocrinology And Metabolism*, 281(1), 2001.

[28] Blüml, S., A. Moreno-Torres, F. Shic, C.H. Nguy, and B.D. Ross. Tricarboxylic acid cycle of glia in the in vivo human brain. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo*, 15(1), 2002.

[29] Sonnewald, U., Westergaard, N., Petersen, S.B., Unsgård, G., and Schousboe, A. Metabolism of [U-13C] glutamate in astrocytes studied by 13C NMR spectroscopy: incorporation of more label into lactate than into glutamine demonstrates the importance of the tricarboxylic acid cycle. *Journal of neurochemistry*, 61(3), 1993.

[30] L. Hertz and E. Hertz. Cataplerotic TCA cycle flux determined as glutamate-sustained oxygen consumption in primary cultures of astrocytes. *Neurochemistry international*, 43(4-5), 2003.

[31] M. McKenna. Substrate competition studies demonstrate oxidative metabolism of glucose, glutamate, glutamine, lactate and 3-hydroxybutyrate in cortical astrocytes from rat brain. *Neurochemical research*, 37(11), 2012.

[32] E.M. Koehler-Stec, I.A. Simpson, S.J. Vannucci, K.T. Landschulz, and W. Landschulz. Monocarboxylate transporter expression in mouse brain. *American Journal of Physiology-Endocrinology And Metabolism*, 275(3), 1998.

[33] G. Dienel. Lack of appropriate stoichiometry: Strong evidence against an energetically important astrocyte–neuron lactate shuttle in brain. *Journal of Neuroscience Research*, 95(11), 2017.

[34] M. McAllister, L. Krizanac-Bengez, F. Macchia, R. Naftalin, K. Pedley, M. Mayberg, M. Marroni, S. Leaman, K. Stanness, and D. Janigro. Mechanisms of glucose transport at the blood–brain barrier: an in vitro study. *Brain research*, 904(1), 2001.

[35] H. White and B Venkatesh. Clinical review: ketones and brain injury. *Critical Care*, 15(2), 2011.

[36] R.A. Hawkins, R.L. O'Kane, I.A. Simpson, and J.R. Vina. Structure of the Blood–Brain Barrier and Its Role in the Transport of Amino Acids. *The Journal of Nutrition*, 136(1), 2006.

[37] M. Aprison, R. Shank, and R. Davidoff. A comparison of the concentration of glycine, a transmitter suspect, in different areas of the brain and spinal cord in seven different vertebrates. . *Comparative Biochemistry and Physiology*, 28(3), 1969.

[38] T. Perry, K. Berry, S. Hansen, S. Diamond, and C. Mok. Regional distribution of amino acids in human brain obtained at autopsy. . *Journal of neurochemistry*, 18(3), 1971.

[39] R.P. Shank, M.H. Aprison, and C.F. Baxter. Precursors of glycine in the nervous system: Comparison of specific activities in glycine and other amino acids after administration of [U-14C]glucose, [3,4-14C]glucose, [1-14C]glucose, [U-14C]serine or [1,5-14C]citrate to the rat. *Brain Research*, 52, 1973.

[40] G. Johnston and M. Vitali. Glycine-producing transaminase activity in extracts of spinal cord. *Brain research*, 12(2), 1969.

[41] J. Thompson and K. Richardson. Isolation and characterization of an L-alanine: glyoxylate aminotransferase from human liver. . *Journal of Biological Chemistry*, 242(16), 1967.

[42] G. Ducker and J. Rabinowitz. One-carbon metabolism in health and disease. *Cell metabolism*, 25(1), 2017.

[43] A. MacFarlane, X. Liu, C. Perry, P. Flodby, R. Allen, S. Stabler, and P. Stover. Cytoplasmic serine hydroxymethyltransferase regulates the metabolic partitioning of methylenetetrahydrofolate but is not essential in mice. . *Journal of Biological Chemistry*, 283(38), 2008.

[44] H. Tani, S. Ohnishi, H. Shitara, T. Mito, M. Yamaguchi, H. Yonekawa, O. Hashizume, K. Ishikawa, K. Nakada, and J. Hayashi. Mice deficient in the Shmt2 gene have mitochondrial respiration defects and are embryonic lethal. . *Scientific reports*, 8(1), 2018.

[45] Y. Motokawa and G. KIKUCHI. Glycine Metabolism by Rat Liver Mitochondria: II. Methylene Tetrahydrofolate as the Direct One Carbon Donor in the Reaction of Glycine Synthesis. *The Journal of Biochemistry*, 65(1), 1969.

[46] E. Meléndez-Hevia and P. de Paz-Lugo. Branch-point stoichiometry can generate weak links in metabolism: the case of glycine biosynthesis. . *Journal of biosciences*, 33(5), 2008.

[47] K. Sato, S. Yoshida, K. Fujiwara, K. Tada, and M. Tohyama. Glycine cleavage system in astrocytes. *Brain Research*, 567(1), 1991.

[48] M. Visentin, R. Zhao, and I. Goldman. The antifolates. *Hematology/Oncology Clinics*, 26(3), 2012.

[49] S.S. Abolmaali, A.M. Tamaddon, and R. Dinarvand. A review of therapeutic challenges and achievements of methotrexate delivery systems for treatment of cancer and rheumatoid arthritis. . *Cancer chemotherapy and pharmacology*, 71(5), 2013.

[50] M. Yang and K. Vousden. Serine and one-carbon metabolism in cancer. . *Nature Reviews Cancer*, 16(10), 2016.

[51] K. Mattaini, M. Sullivan, and M. Vander Heiden. The importance of serine metabolism in cancer. . *Journal of Cell Biology*, 214(3), 2016.

[52] Amelio, I., Cutruzzolá, F., Antonov, A., Agostini, M., and Melino, G. Serine and glycine metabolism in cancer. . *Trends in biochemical sciences*, 39(4), 2014.

[53] J. Liu, S. Guo, Q. Li, L. Yang, Z. Xia, L. Zhang, Z. Huang, and N. Zhang. Phosphoglycerate dehydrogenase induces glioma cells proliferation and invasion by stabilizing forkhead box M1. *Journal of neuro-oncology*, 111(3), 2013.

[54] C. Labuschagne, N. Van Den Broek, G. Mackay, K. Vousden, and O. Maddocks. Serine, but not glycine, supports one-carbon metabolism and proliferation of cancer cells. . *Cell reports*, 7(4), 2014.

[55] R. Werman, R. Davidoff, and M. Aprison. Inhibitory of glycine on spinal neurons in the cat. *Journal of neurophysiology*, 31(1), 1986.

[56] A. Young and S. Snyder. The glycine synaptic receptor: evidence that strychnine binding is associated with the ionic conductance mechanism. *Proceedings of the National Academy of Sciences*, 71(10), 1974.

[57] F. Pfeiffer, D. Graham, and H. Betz. Purification by affinity chromatography of the glycine receptor of rat spinal cord. . *Journal of Biological Chemistry*, 257(16), 1982.

[58] T. Lynagh and J. Lynch. Molecular mechanisms of Cys-loop ion channel receptor modulation by ivermectin. . *Frontiers in molecular neuroscience*, 5, 2012.

[59] A. Thompson, H. Lester, and S. Lummis. The structural basis of function in Cys-loop receptors. . *Quarterly reviews of biophysics*, 43(4), 2010.

[60] J. Lynch. Molecular structure and function of the glycine receptor chloride channel. . *Physiological reviews*, 84(4), 2004.

[61] J. Fritschy, R. Harvey, and G. Schwarz. Gephyrin: where do we stand, where do we go? *Trends in neurosciences*, 31(5), 2008.

[62] A. Flint, X. Liu, and A. Kriegstein. Nonsynaptic glycine receptor activation during early neocortical development. *Neuron*, 20(1), 1998.

[63] C. Deleuze, M. Runquist, H. Orcel, Rabié, A., G. Dayanithi, G. Alonso, and N. Hussy. Structural difference between heteromeric somatic and homomeric axonal glycine receptors in the hypothalamo-neurohypophysial system. *Neuroscience*, 135(2), 2005.

[64] H. Ishibashi, J. Yamaguchi, Y. Nakahata, and J. Nabekura. Dynamic regulation of glycine-GABA co-transmission at spinal inhibitory synapses by neuronal glutamate transporter. *Jounal of Physiology*, 591(16):3821–32, 2013.

[65] Bröer, S. and U. Gether. The solute carrier 6 family of transporters. . *British journal of pharmacology*, 167(2), 2012.

[66] M.J. Roux and S. Supplisson. Neuronal and glial glycine transporters have different stoichiometries. *Neuron*, 25(2):373–83, 2000.

[67] J. Gomeza, S. Hülsmann, K. Ohno, V. Eulenburg, K. Szöke, D. Richter, and H Betz. Inactivation of the glycine transporter 1 gene discloses vital role of glial glycine uptake in glycinergic inhibition. . *Neuron*, 40(4), 2003.

[68] F. Zafra and Giménez, C. Glycine transporters and synaptic function. *IUBMB life*, 60(12), 2008.

[69] I. Poyatos, J. Ponce, Aragón, C., Giménez, C., and F. Zafra. The glycine transporter GLYT2 is a reliable marker for glycine-immunoreactive neurons. . *Molecular brain research*, 49(1-2), 1997.

[70] K. Aubrey, F. Rossi, R. Ruivo, S. Alboni, G. Bellenchi, A. Le Goff, B. Gasnier, and S Supplisson. The transporters GlyT2 and VIAAT cooperate to determine the vesicular glycinergic phenotype. . *Journal of Neuroscience*, 27(23), 2007.

[71] S.L. McIntire, R.J. Reimer, K. Schuske, R.H. Edwards, and E.M. Jorgensen. Identification and characterization of the vesicular GABA transporter. *Nature Reviews Drug Discovery*, 389(6653):870–6, 1997.

[72] A.N. van den Pol and T. Gorcs. Glycine and glycine receptor immunoreactivity in brain and spinal cord. . *Journal of Neuroscience*, 8(1), 1988.

[73] E. Naas, K. Zilles, H. Gnahn, H. Betz, and C. Becker. Glycine receptor immunoreactivity in rat and human cerebral cortex. *Brain research*, 561(1), 1991.

[74] M. S. Hernandes and L.R. Troncone. Glycine as a neurotransmitter in the forebrain: a short review. *Journal of Neural Transmission*, 116(12):1551–1560, 2009.

[75] J. Davies, S. Chung, R. Thomas, A. Robinson, C. Hammond, J. Mullins, E. Carta, B. Pearce, K. Harvey, and R. Harvey. The glycinergic system in human startle disease: a genetic screening approach. . *Frontiers in molecular neuroscience*, 3, 2012.

[76] H. Hermanns, U. Muth-Selbach, R. Williams, S. Krug, P. Lipfert, R. Werdehausen, S. Braun, and I. Bauer. Differential effects of spinally applied glycine transporter inhibitors on nociception in a rat model of neuropathic pain. *Neuroscience letters*, 445(3), 2008.

[77] K. Morita, N. Motoyama, T. Kitayama, N. Morioka, K. Kifune, and T. Dohi. Spinal antiallodynia action of glycine transporter inhibitors in neuropathic pain models in mice. . *Journal of Pharmacology and Experimental Therapeutics,*, 362(2), 2008.

[78] K. Perry, J. Falcone, M. Fell, J. Ryder, H. Yu, P. Love, J. Katner, K. Gordon, M. Wade, T. Man, G. Nomikos, L. Phebus, A. Cauvin, K. Johnson, C. Jones, B. Hoffmann, G. Sandusky, M. Walter, W. Porter, L. Yang, K. Merchant, H. Shannon, and K. Svensson. Neurochemical and behavioral profiling of the selective GlyT1 inhibitors ALX5407 and LY2365109 indicate a preferential action in caudal vs. cortical brain areas. *Neuropharmacology*, 55, 2008.

[79] D. Bugarski-Kirola, T. Blaettler, C. Arango, W. Fleischhacker, G. Garibaldi, A. Wang, M. Dixon, R. Bressan, H. Nasrallah, S. Lawrie, and J. Napieralski. Bitopertin in negative symptoms of schizophrenia—results from the phase III FlashLyte and DayLyte studies. *Biological psychiatry*, 82(1), 2017.

[80] Alfadhel,M., Nashabat, M., Al Qahtani, H., Alfares,A., Al Mutairi, F., and Al Shaalan, H. Mutation in SLC6A9 encoding a glycine transporter causes a novel form of non-ketotic hyperglycinemia in humans. *Human Genetics*, 135(11), 2016.

[81] Cornish-Bowden, A. John Wiley & Sons., 2013.

[82] A. Cornish-Bowden. One hundred years of Michaelis–Menten kinetics. . *Perspectives in Science*, 4, 2015.

[83] Monod, J., Wyman, J., and Changeux, J. On the nature of allosteric transitions: A plausible model. *J Mol Biol*, 12(1), 1965.

[84] J.P. Changeux. Allostery and the Monod-Wyman-Changeux model after 50 years. *Annual review of biophysics*, 41, 2012.

[85] M. Canals, J.R. Lane, A. Wen, P.J. Scammells, P.M. Sexton, and A. Christopoulos. A Monod-Wyman-Changeux mechanism can explain G protein-coupled receptor (GPCR) allosteric modulation. *Journal of Biological Chemistry*, 287(1), 2012.

[86] O. Yifrach and A. Horovitz. Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL. . *Biochemistry*, 34(16), 1995.

[87] J. Changeux. Allostery and the Monod-Wyman-Changeux model after 50 years. *Annual review of biophysics*, 41, 2012.

[88] D. Colquhoun and R. Lape. Perspectives on: conformational coupling in ion channels: allosteric coupling in ligand-gated ion channels. *Journal of General Physiology*, 140(6):599–612, 2012.

[89] J.L. Smart and J.A. McCammon. Analysis of synaptic transmission in the neuromuscular junction using a continuum finite element model. . *Biophysical journal*, 75(4), 1998.

[90] S. Schuster and D. Fell. *Modeling and Simulating Metabolic Networks*. 2007.

[91] J. Hofmeyr. Steady-state modelling of metabolic pathways: A guide for the prospective simulator. . *Bioinformatics,*, 2(1), 1986.

[92] S. Schuster and C. Hilgetag. On elementary flux modes in biochemical reaction systems at steady state. . *Biological Systems*, 2(02), 1994.

[93] T. Pfeiffer, I. Sanchez-Valdenebro, J. Nuvo, F. Montero, and S. Schuster. METATOOL: for studying metabolic networks. *Bioinformatics (Oxford, England)*, 15(3), 1999.

[94] M. Poolman, C. Sebu, M. Pidcock, and D. Fell. Molecular decomposition of metabolic systems via null-space analysis. *Journal of Theoretical Biology*, 249(4), 2007.

[95] J. Leiser and J.J. Blum. On the analysis of substrate cycles in large metabolic systems. . *Cell biophysics*, 11(1), 1987.

[96] Fell, D.A. *The analysis of flux in substrate cycles*. Springer, 1993.

[97] S. Schuster and C. Hilgetag. On elementary flux modes in biochemical reaction systems at steady state. . *Journal of Biological Systems*, 2(02), 1994.

[98] H. Sauro. Moiety-conserved cycles and metabolic control analysis: problems in sequestration and metabolic channelling. . *BioSystems*, 33(1), 1994.

[99] D. Fell and J. Small. Fat synthesis in adipose tissue. An examination of stoichiometric constraints. . *Biochemical Journal*, 238(3), 1986.

[100] M. Watson. A discrete model of bacterial metabolism. *Bioinformatics*, 2(1), 1986.

[101] A. Varma and B.O. Palsson. Metabolic capabilities of Escherichia coli II. Optimal growth patterns. *ournal of Theoretical Biology*, 165(4), 1993.

[102] A. Raghunathan, J. Reed, S. Shin, B. Palsson, and S. Daefler. Constraint-based analysis of metabolic capacity of Salmonella typhimurium during host-pathogen interaction. . *BMC systems biology,*, 3(1), 2009.

[103] M.G. Poolman, L. Miguet, L.J. Sweetlove, and D.A. Fell. A genome-scale metabolic model of Arabidopsis and some of its properties. *Plant physiology*, 151(3), 2009.

[104] C. Gu, G. Kim, W. Kim, H. Kim, and S. Lee. Current status and applications of genome-scale metabolic models. . *Genome Biology*, 20(1), 2019.

[105] H. Hartman, D. Fell, S. Rossell, P. Jensen, M. Woodward, L. Thorndahl, L. Jelsbak, J. Olsen, A. Raghunathan, S. Daefler, and M. Poolman. Identification of potential drug targets in Salmonella enterica sv. Typhimurium using metabolic modelling and experimental validation. . *Microbiology*, 160(6), 2014.

[106] M. Asadollahi, J. Maury, K. Patil, M. Schalk, A. Clark, and J. Nielsen. Enhancing sesquiterpene production in Saccharomyces cerevisiae through in silico driven metabolic engineering. *Metabolic engineering*, 11(6), 2009.

[107] C. Gille, C. Bölling, A. Hoppe, S. Bulik, S. Hoffmann, K. Hübner, A. Karlstädt, R. Ganeshan, M. König, K. Rother, and M. Weidlich. HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. . *Molecular systems biology*, 6(1), 2010.

[108] T. Çakr, S. Alsan, H. Saybaşl, A. Akn, and K. Ülgen. Reconstruction and flux analysis of coupling between metabolic pathways of astrocytes and neurons: application to cerebral hypoxia. . *Theoretical Biology and Medical Modelling*, 4(1), 2007.

[109] H. Ma, A. Sorokin, A. Mazein, A. Selkov, E. Selkov, O. Demin, and I. Goryanin. The Edinburgh human metabolic network reconstruction and its functional analysis. *Molecular Systems Biology*, 3(1), 2007.

[110] N. Duarte, S. Becker, N. Jamshidi, I. Thiele, M. Mo, T. Vo, R. Srivas, and B. Palsson. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *PNAS*, 104(6), 2007.

[111] E. Brunk et al. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nature Biotechnology*, 36(1), 2018.

[112] Y. Wang, J. Eddy, and N. Price. Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE. *BMC systems biology*, 6(1), 2012.

[113] R. Agren, S. Bordel, A. Mardinoglu, N. Pornputtapong, I. Nookaew, and J. Nielsen. Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. . *PLoS Comput Biol*, 8(5), 2012.

[114] A. Mardinoglu, R. Agren, C. Kampf, A. Asplund, M. Uhlen, and J. Nielsen. Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. *Nature communications*, 5(1), 2014.

[115] P. Karp, C. Ouzounis, C. Moore-Kochlacs, L. Goldovsky, P. Kaipa, D. Ahrén, S. Tsoka, N. Darzentas, V. Kunin, and N. López-Bigas. Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. . *Nucleic acids research*, 33(19), 2005.

[116] M. Kanehisa, S. Goto, M. Hattori, K. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki, and M. Hirakawa. From genomics to chemical genomics: new developments in KEGG. . *Nucleic acids research*, 34(suppl_1), 2006.

[117] I. Schomburg, A. Chang, C. Ebeling, M. Gremse, C. Heldt, G. Huhn, and D. Schomburg. BRENDA, the enzyme database: updates and major new developments. *Nucleic acids research*, 32(suppl_1), 2004.

[118] J. Venter, M. Adams, E. Myers, P. Li, R. Mural, G. Sutton, H. Smith, M. Yandell, C. Evans, R. Holt, J. Gocayne, et al. The sequence of the human genome. *science*, 291(5507), 2001.

[119] N. Vlassis, M. P. Pacheco, and T. Sauter. Fast reconstruction of compact context-specific metabolic network models. . *PLoS Comput Biol*, 10(1), 2014.

[120] R. Agren, L. Liu, S. Shoaie, W. Vongsangnak, I. Nookaew, and J. Nielsen. The RAVEN toolbox and its use for generating a genome-scale metabolic model for Penicillium chrysogenum. . *PLoS Comput Biol*, 9(3), 2013.

[121] S. Becker and B. Palsson. Context-specific metabolic networks are consistent with experiments. *PLoS Comput Biol*, 4(5), 2008.

[122] S. Opdam, A. Richelle, B. Kellman, D. Zielinski, .S Li, and N. Lewis. A systematic evaluation of methods for tailoring genome-scale metabolic models. *Cell systems*, 4(3), 2017.

[123] B. Hwang, J.H. Lee, and D. Bang. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Experimental & molecular medicine*, 50(8), 2018.

[124] J. Eberwine, H. Yeh, K. Miyashiro, Y. Cao, S. Nair, R. Finnell, M. Zettel, and P. Coleman. Analysis of gene expression in single live neurons. . *Proceedings of the National Academy of Sciences*, 89(7), 1992.

[125] A. Saliba, A. Westermann, S. Gorski, and J. Vogel. Single-cell RNA-seq: advances and future challenges. *Nucleic acids research*, 42(14), 2014.

[126] Y. Zhang, M. S. Kim, E. Nguyen, and D. Taylor. Modeling metabolic variation with single-cell expression data. . *bioRxiv*, 2020.

[127] D. Fell, M. Poolman, and A. Gevorgyan. Building and analysing genome-scale metabolic models. *Biochemical Society Transactions*, 38(5), 2010.

[128] A. Gevorgyan, M. Poolman, D. Fell, et al. Detection of stoichiometric inconsistencies in biomolecular models. *Bioinformatics*, 24(19), 2008.

[129] D. Garfinkel. A machine-independent language for the simulation of complex chemical and biochemical systems. *Computers and Biomedical Research*, 2(1), 1968.

[130] H. Sauro. SCAMP: a general-purpose metabolic simulator and metabolic control analysis program. *Comp. Appl. Biosci.*, 9(4), 1993.

[131] L. Wang, S. Dash, C. Ng, and C. Maranas. A review of computational tools for design and reconstruction of metabolic pathways. . *Synthetic and systems biotechnology*, 2(4), 2017.

[132] S. Mendoza, B. Olivier, D. Molenaar, and B. Teusink. A systematic assessment of current genome-scale metabolic reconstruction tools. *Genome biology*, 20(1), 2019.

[133] M. Poolman. ScrumPy:metabolic modelling with Python. *IEE Proceedings-Systems Biology*, 153(5), 2006.

[134] H. Yuan, C. Cheung, M. Poolman, P. Hilbers, and N. van Riel. A genome-scale metabolic network reconstruction of tomato (Solanum lycopersicum L.) and its application to photorespiratory metabolism. *The Plant Journal*, 85(2), 2016.

[135] R. Norman, T. Millat, S. Schatschneider, A. Henstra, R. Breitkopf, B. Pander, F. Annan, P. Piatek, H. Hartman, M. Poolman, and D. Fell. Genome-scale model of C. autoethanogenum reveals optimal bioprocess conditions for high-value chemical production from carbon monoxide. . *Engineering Biology*, 3(2), 2019.

[136] T. Sejnowski, C. Koch, and P. Churchland. Computational neuroscience. *Science*, 241(4871), 1988.

[137] E. De Schutter. Why are computational neuroscience and systems biology so separate? *PLoS Comput Biol*, 4(5), 2008.

[138] R. Kass, S. Amari, K. Arai, E. Brown, C. Diekman, M. Diesmann, B. Doiron, U. Eden, A. Fairhall, G. Fiddyment, and T. Fukai. Computational neuroscience: Mathematical and statistical perspectives. . *Annual review of statistics and its application*, 5, 2018.

[139] M. Piccolino. Animal electricity and the birth of electrophysiology: the legacy of Luigi Galvani. *Brain research bulletin*, 46(5), 1998.

[140] D. Colquhoun and A. Hawkes. On the stochastic properties of single ion channels. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 211(1183), 1981.

[141] A. Destexhe, Z. Mainen, and T. Sejnowski. Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. *Journal of computational neuroscience*, 1(3), 1994.

[142] D.A. Doyle, J. Cabral, R. Pfuetzner, A. Kuo, J. Gulbis, S. Cohen, B. Chait, and R. MacKinnon. The structure of the potassium channel: molecular basis of K+ conduction and selectivity. *science*, 280(5360), 1998.

[143] C. Wagner, B. Friedrich, I. Setiawan, F. Lang, and S. Bröer. The use of Xenopus laevis oocytes for the functional characterization of heterologously expressed membrane proteins. *Cellular Physiology and Biochemistry*, 10(1-2), 2000.

[144] V. Burzomato. Single-Channel Behavior of Heteromeric 1 Glycine Receptors: An Attempt to Detect a Conformational Change before the Channel Opens. *Journal of Neuroscience*, 24 (48):10924–10940, 2004.

[145] N. Swainston, K. Smallbone, H. Hefzi, P.D. Dobson, J. Brewer, M. Hanscho, D.C. Zielinski, K.S. Ang, N.J. Gardiner, J.M. Gutierrez, S. Kyriakopoulos, M. Lakshmanan, S. Li, J.K. Li, V.S. Martinez, C.A. Orellana, L. Quek, A. Thomas, J. Zanghellini, N. Borth, D. Lee, L.K. Nielsen, D.B. Kell, N.E. Lewis, and P. Mendes. Recon 2.2: from reconstruction to model of human metabolism. *Metabolomics*, 12(109), 2016.

[146] C Martín-Jiménez, D. Salazar-Barreto, G. Barreto, and J. González. Genome-scale reconstruction of the human astrocyte metabolic network. . *Frontiers in aging neuroscience*, 9, 2017.

[147] S. Zeisel and K. Costa. Choline: An Essential Nutrient for Public Health. *Nutrition Reviews*, 67(11), 2009.

[148] D. Purves, G. Augustine, D. Fitzpatrick, L. Katz, A. LaMantia, J. McNamara, and S. Williams. *Neuroscience 2nd edition. sunderland (ma) sinauer associates.* Sinauer Associates, 2001.

[149] P. Hinkle. "P/O ratios of mitochondrial oxidative phosphorylation.". *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1706(1-2), 2005.

[150] M. McKenna, H.S. Waagepetersen, A. Schousboe, and U. Sonnewald. Neuronal and astrocytic shuttle mechanisms for cytosolic-mitochondrial transfer of reducing equivalents: current evidence and pharmacological tools. . *Biochemical pharmacology*, 71(4), 2006.

[151] D. Kane. Lactate oxidation at the mitochondria: a lactate-malate-aspartate shuttle at work. . *Frontiers in neuroscience*, 8, 2014.

[152] J. Pérez-Escuredo, V. Van Hée, M. Sboarina, J. Falces, V. Payen, L. Pellerin, and P. Sonveaux. Monocarboxylate transporters in the brain and in cancer. . *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1863(10), 2016.

[153] A. Cooper and T. Jeitner. Central role of glutamate metabolism in the maintenance of nitrogen homeostasis in normal and hyperammonemic brain. *Biomolecules*, 6(2), 2016.

[154] L. Baxter, J. Mazziotta, M. Phelps, C. Selin, B. Guze, and L. Fairbanks. Cerebral Glucose Metabolic Rates in Normal Human Females Versus Normal Males. *Psychiary Research*, 3(1), 1987.

[155] E. Gibbs, W. Lennox, L. Nims, and F. Gibbs. Arterial and cerebral venous blood: arterial-venous differences in man. *Journal of Biological Chemistry*, 144(2), 1942.

[156] H. Benveniste, G. Dienel, Z. Jacob, H. Lee, R. Makaryus, A. Gjedde, F. Hyder, and D. Rothman. Trajectories of Brain Lactate and Re-visited Oxygen-Glucose Index Calculations Do Not Support Elevated Non-oxidative Metabolism of Glucose Across Childhood. *Frontiers in Neuroscience*, 12(631), 2018.

[157] P. Cohen, S. Alexander, T. Smith, M. Reivich, and H. Wollman. Effects of hypoxia and normocarbia on cerebral blood flow and metabolism in conscious man. *Journal of Applied Physiology*, 23(2), 1967.

[158] P. Rasmussen, M. Wyss, and C. Lundby. Cerebral glucose and lactate consumption during cerebral activation by physical activity in humans. *The FASEB Journal*, 25(9), 2011.

[159] W. Olendorf. Brain uptake of radiolabelled amino acids, amines, and hexoses after arterial injection. *The American Journal of Physiology*, 221(6), 1971.

[160] L. Battistin, A. Grynbaum, and A. Lajtha. The uptake of various amino acids by the mouse brain in vivo. *Brain Research*, 29(1), 1971.

[161] W. Olendorf and J. Szabo. Amino acid assignment to one of three blood-brain barrier amino acid carriers. *American Journal of Physology*, 230(1), 1976.

[162] H. Sershen and A. Lajtha. Capillary transport of amino acids in the developing brain. *Experimental Neurology*, 53(2), 1976.

[163] G Knudsen, K. Pettigrew, C. Patlak, and M. Hertz. Asymmetrical Transport of Amino Acids Across the Blood-Brain Barrier in Humans. *Journal of Cerebral Blood Flow & Metabolism*, 10(5), 1990.

[164] A. Karmi, P. Iozzo, A. Viljanen, B. Fielding, K. Virtanen, V. Oikonen, J. Kemppainen, T. Viljanen, L. Guiducci, M. Haaparanta-Solin, K. Någren, O. Solin, and P. Nuutila. Increased brain fatty acid uptake in metabolic syndrome. *Diabetes*, 59(9), 2010.

[165] J. DeMar, H. Lee, K. Ma, L. Chang, S. Rapoport, and R. Bazinet. Brain elongation of linoleic acid is a negligible source of the arachidonate in brain phospholipids of adult rats. *Biochimica et Biophysica Acta*, 9(1), 2006.

[166] W. Norton and S. Poduslo. Neuronal perikarya and astroglia of rat brain: chemical composition during myelination. *Journal of Lipid Research*, 12(1), 1971.

[167] X. Zhu, H. Qiao, F. Du, Q. Xiong, X. Liu, X. Zhang, K. Ugurbll, and W. Chen. Quantitative Imaging of Energy Expenditure in Human Brain. *Neuroimage*, 60(4), 2012.

[168] G. Mason, R. Gruetter, D. Rothman, K. Behar, G. Shulman, and E. Novotny. Simultaneous Determination of the Rates of the TCA Cycle, Glucose Utilization, (X-Ketoglutarate/Glutamate Exchange, and Glutamine Synthesis in Human Brain by NMR. *Journal of Cerebral Blood Flow and Metabolism*, 15(1), 1995.

[169] J. Shen, K. Shen, K. Behar, P. Brown, T. Nixon, G. Mason, O. Petroff, G. Shulman, and D. Rothman. Determination of the rate of the glutamate/glutamine cycle in the human brain by in vivo 13C NMR. *Neurobiology*, 96(14), 1999.

[170] M. Garcia-Espinosa, T. Rodrigues, A. Sierra, M. Benito, C. Fonseca, H. Gray, B. Bartnik, M. Garcia-Martın, P. Ballesteros, and S. Cerdán. Cerebral glucose metabolism and the glutamine cycle as detectedby in vivo and in vitro 13C NMR spectroscopy. *Neurochemistry International*, 45(2), 2004.

[171] R. Shank and M. Aprison. The metabolism in vivo of glycine and serine in eight areas of the rat central nervous system. *Journal of Neurochemistry*, 17(10), 1970.

[172] B. Phypers and J. Pierce. Lactate physiology in health and disease. *Continuing education in Anaesthesia, critical care & pain*, 6(3), 2006.

[173] M. Rogatzki, B. Ferguson, M. Goodwin, and L. Gladden. Lactate is always the end product of glycolysis. *Frontiers in neuroscience*, 9, 2015.

[174] M. Larrabee. Lactate metabolism and its effects on glucose metabolism in an excised neural tissue. *Journal of neurochemistry,*, 64(4), 1995.

[175] D. Smith, A. Pernet, W. Hallett, E. Bingham, P. Marsden, and S. Amiel. Lactate: a preferred fuel for human brain metabolism in vivo. *Journal of Cerebral Blood Flow & Metabolism*, 23 (6), 2003.

[176] T. Hashimoto, R. Hussien, H. Cho, D. Kaufer, and G. Brooks. Evidence for the mitochondrial lactate oxidation complex in rat neurons: demonstration of an essential component of brain lactate shuttles. *PloS one*, 3(8), 2008.

[177] J. Lemire, R. Mailloux, and V. Appanna. Mitochondrial lactate dehydrogenase is involved in oxidative-energy metabolism in human astrocytoma cells (CCF-STTG1). *PloS one*, 3(2), 2008.

[178] A. Yu and L. Hertz. Uptake of glutamate, GABA, and glutamine into a predominatly GABA-ergic and a predominantly glutamatergic nerve cell population in culture. . *Journal of neuroscience research*, 7(1), 1982.

[179] G. Maurer, D. Brucker, O. Bähr, P. Harter, E. Hattingen, S. Walenta, W. Mueller-Klieser, J. Steinbach, and J. Rieger. Differential utilization of ketone bodies by neurons and glioma cell lines: a rationale for ketogenic diet as experimental glioma therapy. . *BMC cancer,*, 11 (1), 2011.

[180] K. Mattaini, M. Sullivan, and M. Vander Heiden. The importance of serine metabolism in cancer. *Journal of Cell Biology*, 241(3), 2016.

[181] L. Chasin, A. Feldman, M. Konstam, and G. Urlaub. Reversion of a Chinese hamster cell auxotrophic mutant. . *Proceedings of the National Academy of Sciences*, 71(3), 1974.

[182] K. Herbig, E. Chiang, L. Lee, J. Hills, B. Shane, and P. Stover. Cytoplasmic serine hydroxymethyltransferase mediates competition between folate-dependent deoxyribonucleotide andS-adenosylmethionine biosyntheses. *Journal of Biological Chemistry*, 277(41), 2002.

[183] W. Wang, Z Wu, Z. Dai, Y. Yang, J. Wang, and G. Wu. Glycine metabolism in animals and humans: implications for nutrition and health. *Amino Acids*, 45(3):463–477, 2013.

[184] D. Fell and R. Steele. Enhancement of histidine and one-carbon metabolism in rats fed high levels of retinol. *The Journal of nutrition*, 112(3), 1982.

[185] H. Zeng and J. Sanes. Neuronal cell-type classification: challenges, opportunities and the path forward. . *Nature Reviews Neuroscience*, 18(9), 2017.

[186] A. Zeisel et al. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*, 347(6226), 2015.

[187] M. Kaern, T. Elston, W. Blake, and J. Collins. Stochasticity in gene expression: from theories to phenotypes. . *Nature Reviews Genetics,*, 6(6), 2005.

[188] Z. Xiao, Z. Dai, and J. Locasale. Metabolic landscape of the tumor microenvironment at single cell resolution. . *Nature communications*, 10(1), 2019.

[189] G. Ducker, J. Ghergurovich, N. Mainolfi, V. Suri, S. Jeong, S. Li, A. Friedman, M. Manfredi, Z. Gitai, H. Kim, and J. Rabinowitz. Human SHMT inhibitors reveal defective glycine import as a targetable metabolic vulnerability of diffuse large B-cell lymphoma. . *Proceedings of the National Academy of Sciences*, 114(43), 2017.

[190] M Pacold, K. Brimacombe, S. Chan, J. Rohde, C. Lewis, L. Swier, R. Possemato, W. Chen, L. Sullivan, B. Fiske, and S. Cho. A PHGDH inhibitor reveals coordination of serine synthesis and one-carbon unit fate. *Nature chemical biology*, 12(6), 2016.

[191] *Discovery of novel SHMT small molecule inhibitors for cancer treatment.*, 2018.

[192] G. Craft, A. Chen, and A. Nairn. Recent advances in quantitative neuroproteomics. *Methods*, 61(3), 2013.

[193] V. Marx. A dream of single-cell proteomics. *Nature methods*, 16(9), 2019.

[194] Y. Daikhin and M. Yudkoff. Compartmentation of brain glutamate metabolism in neurons and glia. *The Journal of nutrition*, 130(4), 2000.

[195] N. Robinson and C. Williams. Amino acids in human brain. . *Clinica Chimica Acta*, 12(3), 1965.

[196] J. Harsing, Z. Juranyi, I. Gacsalyi, P. Tapolcsanyi, A. Czompa, and P. Matyus. Glycine transporter type-1 and its inhibitors. *Current medicinal chemistry*, 13(9), 2006.

[197] R. Hyde, P. Taylor, and H. Hundal. Amino acid transporters: roles in amino acid sensing and signalling in animal cells. . *Biochemical Journal*, 373(1), 2003.

[198] J. Ehmsen, Y. Liu, Y. Wang, N. Paladugu, A. Johnson, J. Rothstein, S. Du Lac, M. Mattson, and A. Höke. The astrocytic transporter SLC7A10 (Asc-1) mediates glycinergic inhibition of spinal cord motor neurons. *Scientific reports*, 6, 2016.

[199] S. Lawrence, J. Hackett, and R. Moran. Tetrahydrofolate recognition by the mitochondrial folate transporter. . *Journal of Biological Chemistry*, 286(36), 2011.

[200] J. Mukherjee, J. Phillips, S. Zheng, J. Wiencke, S. Ronen, and R. Pieper. Pyruvate kinase M2 expression, but not pyruvate kinase activity, is up-regulated in a grade-specific manner in human glioma. *PloS one*, 8(2), 2013.

[201] B. Chaneton, P. Hillmann, L. Zheng, A. Martin, O. Maddocks, A. Chokkathukalam, J. Coyle, A. Jankevics, F. Holding, K. Vousden, and C. Frezza. Serine is a natural ligand and allosteric activator of pyruvate kinase M2. *Nature*, 491(7424), 2012.

[202] Y. Zhang, K. Chen, S. Sloan, M. Bennett, A. Scholze, S. O'Keeffe, H. Phatnani, P. Guarnieri, C. Caneda, N. Ruderisch, and S. Deng. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *Journal of Neuroscience*, 34(36), 2014.

[203] L. Korotchkina and M. Patel. Site specificity of four pyruvate dehydrogenase kinase isoenzymes toward the three phosphorylation sites of human pyruvate dehydrogenase. . *Journal of Biological Chemistry,*, 276(40), 2001.

[204] K. Urbańska and A. Orzechowski. Unappreciated Role of LDHA and LDHB to Control Apoptosis and Autophagy in Tumor Cells. . *International journal of molecular sciences*, 20 (9), 2019.

[205] C. Schalkwijk and C. Stehouwer. ethylglyoxal, a Highly Reactive Dicarbonyl Compound, in Diabetes, Its Vascular Complications, and Other Age-Related Diseases. . *Physiological reviews*, 100(1), 2020.

[206] W. Walz and S. Mukerji. Lactate release from cultured astrocytes and neurons: a comparison. . *Glia*, 1(6), 1988.

[207] Y. Itoh, T. Esaki, K. Shimoji, M. Cook, M. Law, E. Kaufman, and L. Sokoloff. Dichloroacetate effects on glucose and lactate oxidation by neurons and astroglia in vitro and on glucose utilization by brain in vivo. . *Proceedings of the National Academy of Sciences*, 100(8), 2003.

[208] Díaz-García, C., R. Mongeon, C. Lahmann, D. Koveal, H. Zucker, and G. Yellen. Neuronal stimulation triggers neuronal glycolysis and not lactate uptake. . *Cell metabolism*, 26(2), 2017.

[209] C. Sobieski, N. Warikoo, H. Shu, and S. Mennerick. Ambient but not local lactate underlies neuronal tolerance to prolonged glucose deprivation. . *PloS one*, 13(4), 2018.

[210] J. Duran, A. Gruart, O. Varea, I. López-Soldado, J. Delgado-García, and J. Guinovart. Lack of neuronal glycogen impairs memory formation and learning-dependent synaptic plasticity in mice. . *Frontiers in cellular neuroscience*, 13(374), 2019.

[211] A. Bouzier-Sore, P. Voisin, V. Bouchaud, E. Bezancon, J. Franconi, and L. Pellerin. ompetition between glucose and lactate as oxidative energy substrates in both neurons and astrocytes: a comparative NMR study. . *European Journal of Neuroscience,*, 24(6), 2006.

[212] E. Schrödinger. *What Is Life?* Cambridge University Press, 1944.

[213] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular Biology of the Cell.4th edition*. Garland Science, 2002.

[214] M. Shimizu, S. Furuya, Y. Shinoda, J. Mitoma, T. Okamura, I. Miyoshi, N. Kasai, Y. Hirabayashi, and Y. Suzuki. Functional analysis of mouse 3-phosphoglycerate dehydrogenase (Phgdh) gene promoter in developing brain. . *Journal of neuroscience research*, 76(5), 2004.

[215] J. Mitoma, S. Furuya, and Y. Hirabayashi. A novel metabolic communication between neurons and astrocytes: non-essential amino acid L-serine released from astrocytes is essential for developing hippocampal neurons. . *Neuroscience research*, 30(2), 1998.

[216] P. Ledoux, C. Scriver, and P. Hechtman. Expression and molecular analysis of mutations in prolidase deficiency. . *American journal of human genetics*, 59(5), 1996.

[217] Seppänen, A., T. Suuronen, S. Hofmann, K. Majamaa, and I. Alafuzoff. Distribution of collagen XVII in the human brain. *Brain research*, 1158, 2007.

[218] H. Kaur, C. Kumar, C. Junot, M. Toledano, and A. Bachhawat. Dug1p is a Cys-Gly peptidase of the -glutamyl cycle of Saccharomyces cerevisiae and represents a novel family of Cys-Gly peptidases. . *Journal of Biological Chemistry*, 284(21), 2009.

[219] F. Erdem, M. Ilic, P. Koppensteiner, Gołacki, J., G. Lubec, M. Freissmuth, and W. Sandtner. A comparison of the transport kinetics of glycine transporter 1 and glycine transporter 2. *Journal of General Physiology*, 151(8), 2019.

[220] A. Foster, J. Farnsworth, G. Lind, Y. Li, J. Yang, V. Dang, M. Penjwini, V. Viswanath, U. Staubli, and M. Kavanaugh. D-serine is a substrate for neutral amino acid transporters ASCT1/SLC1A4 and ASCT2/SLC1A5, and is transported by both subtypes in rat hippocampal astrocyte cultures. *PloS one*, 11(6), 2016.

[221] K. Nave. Myelination and the trophic support of long axons. *Nature Reviews Neuroscience*, 11(4), 2010.

[222] L. Hertz. The glutamate–glutamine (GABA) cycle: importance of late postnatal development and potential reciprocal interactions between biosynthesis and degradation. . *Frontiers in endocrinology*, 4, 2013.

[223] G. Ducker, L. Chen, R. Morscher, J. Ghergurovich, M. Esposito, X. Teng, Y. Kang, and J. Rabinowitz. Reversal of cytosolic one-carbon flux compensates for loss of the mitochondrial folate pathway. . *Cell metabolism*, 23(6), 2016.

[224] M. Dennis and J. Clarke. Effects of glycine and glyoxylate on cerebral glucose oxidation in vitro. . *Journal of neurochemistry*, 33(1), 1979.

[225] H. Laborit, A. Baron, and J. Olympie. Activité nerveuse centrale et pharmacologie générale comparée du glyoxylate, du glycolate et du glycolaldéhyde. *Agressologie*, 12(3), 1971.

[226] C. Wagner. Biochemical role of folate in cellular metabolism. *Clinical Research and Regulatory Affairs*, 18(3), 2001.

[227] M. Jain, R. Nilsson, S. Sharma, N. Madhusudhan, T. Kitami, A. Souza, R. Kafri, M. Kirschner, C. Clish, and V. Mootha. Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science*, 336(6084), 2012.

[228] A. Divakaruni, M. Wallace, C. Buren, K. Martyniuk, A. Andreyev, E. Li, J. Fields, T. Cordes, I. Reynolds, B. Bloodgood, and L. Raymond. Inhibition of the mitochondrial pyruvate carrier protects from excitotoxic neuronal death. . *Journal of Cell Biology,*, 216(4), 2017.

[229] R. DeBerardinis and T. Cheng. Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. . *Oncogene*, 29(3), 2010.

[230] C. Hensley, A. Wasti, and R. DeBerardinis. Glutamine and cancer: cell biology, physiology, and clinical opportunities. *The Journal of clinical investigation*, 123(9), 2013.

[231] D. Kim, B. Fiske, K. Birsoy, E. Freinkman, K. Kami, R. Possemato, Y. Chudnovsky, M. Pacold, W. Chen, J. Cantor, and L. Shelton. SHMT2 drives glioma cell survival in ischaemia but imposes a dependence on glycine clearance. . *Nature,*, 520(7547), 2015.

[232] A.D. Belalcázar, J.G. Ball, L.M. Frost, M.A. Valentovic, and J Wilkinson. Transsulfuration is a significant source of sulfur for glutathione production in human mammary epithelial cells. *ISRN biochemistry*, 2013.

[233] L. Yu, S. Teoh, E. Ensink, M. Ogrodzinski, C. Yang, A. Vazquez, and S. Lunt. Cysteine catabolism and the serine biosynthesis pathway support pyruvate production during pyruvate kinase knockdown in pancreatic cancer cells. *Cancer & Metabolism*, 7(1), 2019.

[234] L. Sun, M. Bartlam, Y. Liu, H. Pang, and Z. Rao. Crystal structure of the pyridoxal-5-phosphate-dependent serine dehydratase from human liver. *Protein science*, 14(3), 2005.

[235] K. Lund, D. Merrill, and R. Guynn. The reactions of the phosphorylated pathway of L-serine biosynthesis: thermodynamic relationships in rabbit liver in vivo. . *Archives of biochemistry and biophysics*, 237(1), 1985.

[236] L. Yang, J. Canaveras, Z. Chen, L. Wang, L. Liang, C. Jang, J. Mayr, Z. Zhang, J. Ghergurovich, L. Zhan, and S. Joshi. Serine catabolism feeds NADH when respiration is impaired. *Cell Metabolism*, 2020.

[237] M. Adeva, G. Souto, N. Blanco, and C. Donapetry. Ammonium metabolism in humans. . *Metabolism*, 61(11), 2012.

[238] J. Spinelli, H. Yoon, A. Ringel, S. Jeanfavre, C. Clish, and M. Haigis. Metabolic recycling of ammonia via glutamate dehydrogenase supports breast cancer biomass. . *Science*, 358(6365), 2017.

[239] E. Mullarky, J. Xu, A Robin, D. Huggins, A. Jennings, N. Noguchi, A. Olland, D. Lakshminarasimhan, M. Miller, D. Tomita, and M. Michino. Inhibition of 3-phosphoglycerate dehydrogenase (PHGDH) by indole amides abrogates de novo serine synthesis in cancer cells. . *Bioorganic & medicinal chemistry letters,*, 29(17), 2019.

[240] J. Chen, F. Chung, G. Yang, M. Pu, H. Gao, W. Jiang, H. Yin, V. Capka, S. Kasibhatla, Laffitte B., and S. Jaeger. Phosphoglycerate dehydrogenase is dispensable for breast tumor maintenance and growth. . *Oncotarget*, 4(12), 2013.

[241] S. Scott, J. W. Lynch, and A. Keramidas. Correlating structural and energetic changes in glycine receptor activation. *Journal of Biological Chemistry*, 290(9), 2015.

[242] C. Raltschev, F. Hetsch, A. Winkelmann, J. C. Meier, and M. Semtner. Electrophysiological signature of homomeric and heteromeric glycine receptor channels. . *Journal of Biological Chemistry*, 291(34), 2016.

[243] S. Berger, J. Carter, and O. Lowry. The distribution of glycine, GABA, glutamate and aspartate in rabbit spinal cord, cerebellum and hippocampus. . *Journal of neurochemistry*, 28(1), 1977.

[244] P.F. Apostolides and L.O. Trussell. Rapid, activity-independent turnover of vesicular transmitter content at a mixed glycine/GABA synapse. *Journal of Neuroscience*, 33(11), 2013.

[245] J. Chatton, P. Magistretti, and L. Barros. Sodium signaling and astrocyte energy metabolism. *Glia*, 64(10), 2016.

[246] C. Lindsly, C. Gonzalez-Islas, and P. Wenner. Elevated intracellular Na+ concentrations in developing spinal neurons. *Journal of neurochemistry*, 140(5), 2017.

[247] P. Legendre. A reluctant gating mode of glycine receptor channels determines the time course of inhibitory miniature synaptic events in zebrafish hindbrain neurons. *ournal of Neuroscience*, 18(8), 1998.

[248] A. Scimemi and M. Beato. Determining the neurotransmitter concentration profile at active synapses. . *Molecular neurobiology*, 40(3), 2009.

[249] M. Beato. The Time Course of Transmitter at Glycinergic Synapses onto Motoneurons. *Journal of Neuroscience*, 28(29):7412–7425, 2008.

[250] M. Medelin, V. Rancic, G. Cellot, J. Laishram, P. Veeraraghavan, C. Rossi, L. Muzio, L. Sivilotti, and L. Ballerini. Altered development in GABA co-release shapes glycinergic synaptic currents in cultured spinal slices of the SOD1(G93A) mouse model of amyotrophic lateral sclerosis. *Journal of Physiology*, 594(13), 2016.

[251] M. Oyama, S. Kuraoka, S. Watanabe, T. Iwai, and M. Tanabe. Electrophysiological evidence of increased glycine receptor-mediated phasic and tonic inhibition by blockade of glycine transporters in spinal superficial dorsal horn neurons of adult mice. . *Journal of pharmacological sciences*, 133(3), 2017.

[252] K.R. Aubrey, R.J. Vandenberg, and J.D. Clements. Dynamics of Forward and Reverse Transport by the Glial Glycine Transporter, Glyt1b. *Biophysics Journal*, 89(3):1657–1668, 2005.

[253] K. Shibasaki, N. Hosoi, R. Kaneko, M. Tominaga, and K. Yamada. Glycine release from astrocytes via functional reversal of GlyT1. *Journal of neurochemistry*, 140(3), 2017.

[254] A. Bradaïa, R. Schlichter, and J. Trouslard. Role of glial and neuronal glycine transporters in the control of glycinergic and glutamatergic synaptic transmission in lamina X of the rat spinal cord. *The Journal of physiology*, 559(1), 2004.

[255] D. Attwell, B. Barbour, and M. Szatkowski. Nonvesicular release of neurotransmitter. *Neuron*, 11(3), 1993.

[256] A. Auerbach. Thinking in cycles: MWC is a good model for acetylcholine receptor-channels. . *The Journal of physiology*, 590(1), 2012.

[257] J. Clements, R. Lester, G. Tong, C. Jahr, and G. Westbrook. The time course of glutamate in the synaptic cleft. . *Science*, 258(5087), 1992.

[258] J. Trommershäuser, J. Marienhagen, and A. Zippelius. Stochastic model of central synapses: slow diffusion of transmitter interacting with spatially distributed receptors and transporters. *Journal of theoretical biology*, 198(1), 1999.

[259] J. Vargas-Medrano, V. Castrejon-Tellez, F. Plenge, I. Ramirez, and M. Miranda. PKC-dependent phosphorylation of the glycine transporter 1. . *Neurochemistry international*, 59 (8), 2011.

[260] M.J. Titmus, H.E.N.R.I. Korn, and D.S. Faber. Diffusion, not uptake, limits glycine concentration in the synaptic cleft. *Journal of neurophysiology*, 75(4), 1996.

[261] F. Cherubino, E. Bossi, A. Miszner, C. Ghezzi, and A. Peres. Transient Currents in the Glycine Cotransporter GlyT1 Reveal Different Steps in Transport Mechanism. *Journal of Molecular Neuroscience*, 41(2):243–251, 2010.

[262] F. Rousseau, K.R. Aubrey, and S. Supplisson. The glycine transporter GlyT2 controls the dynamics of synaptic vesicle refilling in inhibitory spinal cord neurons. *Journal of Neuroscience*, 28(39):9755–68, 2008.

[263] H. Huang, L. Barakat, D. Wang, and A. Bordey. Bergmann glial GlyT1 mediates glycine uptake and release in mouse cerebellar slices. . *The journal of Physiology*, 560(3), 2004.

[264] A. Armbruster, E. Neumann, V. Kotter, H. Hermanns, R. Werdehausen, and V. Eulenburg. The GlyT1 inhibitor bitopertin ameliorates allodynia and hyperalgesia in animal models of neuropathic and inflammatory pain. . *Frontiers in Molecular Neuroscience*, 10(438), 2018.

[265] V. Eulenburg, M. Retiounskaia, T. Papadopoulos, J. Gomeza, and H. Betz. Glial glycine transporter 1 function is essential for early postnatal survival but dispensable in adult mice. . *Glia*, 58(9), 2010.

[266] H. Safory, S. Neame, Y. Shulman, S. Zubedat, I. Radzishevsky, D. Rosenberg, H. Sason, S. Engelender, A. Avital, S. Hülsmann, J. Schiller, and H. Wolosker. The alanine-serine-cysteine-1 (Asc-1) transporter controls glycine levels in the brain and is required for glycinergic inhibitory transmission . *EMBO Rep.*, 16(5), 2015.

[267] S. Kure, K. Kato, A. Dinopoulos, C. Gail, T.J. DeGrauw, J. Christodoulou, V. Bzduch, R. Kalmanchey, G. Fekete, A. Trojovsky, and B. Plecko. Comprehensive mutation analysis of GLDC, AMT, and GCSH in nonketotic hyperglycinemia. . *Human mutation*, 27(4), 2006.

[268] A.S. Kristensen, J. Andersen, T.N. Jorgensen, L. Sorensen, J. Eriksen, C.J. Loland, K. Stromgaard, and U. Gether. SLC6 neurotransmitter transporters: structure, function, and regulation. *Pharmacological Reviews*, 63(3), 2011.

[269] A. Penmatsa, K.H. Wang, and E. Gouaux. X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature*, 503(7474), 2013.

[270] K. Heckman and L. Pease. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nature protocols*, 2(4), 2007.

[271] T. Beuming, L. Shi, J. Javitch, and H. Weinstein. A comprehensive structure-based alignment of prokaryotic and eukaryotic neurotransmitter/Na+ symporters (NSS) aids in the use of the LeuT structure to probe NSS structure and function. . *Molecular pharmacology*, 70(5), 2006.

[272] S. Kitayama, S. Shimada, H. Xu, L. Markham, D. Donovan, and G. Uhl. Dopamine transporter site-directed mutations differentially alter substrate transport and cocaine binding. *Proceedings of the National Academy of Sciences*, 89(16), 1992.

[273] K. Penado, G. Rudnick, and M. Stephan. Critical amino acid residues in transmembrane span 7 of the serotonin transporter identified by random mutagenesis. *Journal of Biological Chemistry*, 273(43), 1998.

[274] K. Smith, L. Borden, P. Hartig, T. Branchek, and R. Weinshank. Cloning and expression of a glycine transporter reveal colocalization with NMDA receptors. . *Neuron*, 8(5), 1992.

[275] Olivares, L., Aragón, C., Giménez, C., and Zafra, F. The role of N-glycosylation in the targeting and activity of the GLYT1 glycine transporter. *Journal of Biological Chemistry*, 270(16), 1995.

[276] C. Tate and R. Blakely. The effect of N-linked glycosylation on activity of the Na (+)-and Cl (-)-dependent serotonin transporter expressed using recombinant baculovirus in insect cells. . *Journal of Biological Chemistry,*, 269(42), 1994.

[277] G. Cai, P. Salonikidis, J. Fei, W. Schwarz, R. Schülein, W. Reutter, and H. Fan. The role of N-glycosylation in the stability, trafficking and GABA-uptake of GABA-transporter 1: Terminal N-glycans facilitate efficient GABA-uptake activity of the GABA transporter. . *The FEBS journal*, 272(7), 2005.

[278] K. Imai and S. Mitaku. Mechanisms of secondary structure breakers in soluble proteins. *Biophysics*, 1, 2005.

[279] Y. Zhang, M. Kim, E. Nguyen, and D. Taylor. Modeling metabolic variation with single-cell expression data. . *bioRxiv*, 2020.

[280] M.A. Reid, A.E. Allen, S. Liu, M.V. Liberti, P. Liu, X. Liu, Z. Dai, X. Gao, Q. Wang, Y. Liu, and L. Lai. Serine synthesis through PHGDH coordinates nucleotide levels by maintaining central carbon metabolism. *Nature communications*, 9(1), 2018.

[281] X. Lou, V. Scheuss, and R. Schneggenburger. Allosteric modulation of the presynaptic Ca 2+ sensor for vesicle fusion. . *Nature*, 435(7041), 2005.

[282] S. Supplisson and M.J. Roux. Why glycine transporters have different stoichiometries. *FEBS Letters*, 529(1):93–101, 2002.