Chapter 1

Introduction

1.1 Motivation

‘Rice is life’, a quote itself describes the importance of this crop for half of the world’s population [6]. In 2004 United Nations declared the year as the international year of rice. A year devoting to a commodity is precious and shows the importance of rice not only as a staple food but a source of employment for rice producers, processors and traders worldwide [6]. Many households in Asia depends on rice as the main source of nutrients where over 35% to 70% of all calories come from rice. Major source of nutrition like carbohydrates, sugars, fat, protein comes from rice for more than 3 billion people in the world [7]. Food security for billions of people in the world specially in the developing countries, is a critical issue. Providing adequate nutrition for a healthy life is important and a challenge for the poor countries. So instead of providing other nutrition sources/foods for them it is economical to grow more nutritional rice. India in particular, largely dependent on rice and it is one of the major producer of this crop. In 2025 rice production should be doubled as almost 4.5 billion people are going to depend on it as their daily nourishment [7].

Like other cereals, this important crop is also under stress of climate change which can decrease rice grain production and therefore increase the risk of hunger in
several parts of the world [8]. To secure food for billions of people we need to improve yield and quality of rice. Its a decade long subject of interest to improve the yield and quality of rice with the rising demand. From the wild relatives, genes were identified to improve the yield and quality of modern rice [9]. Ideotype approach and crossing between different varieties can improve biomass of hybrid rice [10]. Another notable advancement has been made in producing Vitamin A enriched golden rice [11]. However, the goal is not fully achieved. To meet this challenge of designing an efficient stress tolerant high yielding more nutritious rice cultivar, it is needed to understand the cellular physiology of rice. Metabolism is one of the central player to control the cellular physiology.

Metabolic modelling creates the circumstances in-silico, to observe metabolic responses of an organism. Thus, several conditions could be created using corresponding data and useful predictions can be made. This can improve our biological knowledge about the organism, hence, the predictions can suggest significant direction to design rational experiments.

1.2 Plant cell structure and metabolism

Metabolism is the chemical transformation always occurring within cells of living organisms necessary for responding various perturbations and production of essential molecules and cell components [12]. Plant metabolism achieved remarkable feats for molecular transformation specially capturing sunlight for its energy feeding.¹

Plant cells are eukaryotic cells and has complex matrix structure divided into many organelles, each for a specific task. Figure 1.1 shows a typical plant cell structure. Plants can produce their own food, thus they have some structural differences in comparison to other eukaryotic organisms. Most notable of these is chloroplast. Chloroplast, a type of plastid, contains a green colored pigment called chlorophyll

¹For detail review, interested readers are referred to [13] and hereafter, wherever possible, readers are directed to detail descriptions when any discussion reaches its limit according to the scope of this thesis
which absorbs sunlight to produce necessary energy for producing food by a process called photosynthesis. The stored energy is then used to form organic molecules which is being transported from chloroplast to keep up other necessary metabolic tasks. Mitochondria is a power generating organelle in most eukaryotic cells, responsible for generating most of the cell’s source of chemical energy, ATP. However, mitochondria is also involved in number of other activities from plant development to performance, and has several properties to interact with the specialized features of plant cell metabolism [14]. Golgi apparatus’s membrane bounded structure is responsible for modifying, sorting and packaging of macromolecules for cell secretion. Endoplasmic reticulum (ER) is a network of membranes surrounded with the outer membrane of nuclear envelope and is involved in synthesis of protein (at RER), lipid and carbohydrate metabolism (at SER)\(^2\). Vacuole is a membrane bounded organelle filled with water containing inorganic and organic molecules. Cytoplasm comprises the previously discussed organelles (called compartments) and a gel-like substance called cytosol\(^3\). The cytosol holds a larger part of cell volume and other compartments float in the cytosol. Thus, it provides a support mechanism to them. Although majority of cytosol mass is water, it contains many chemicals that control cell metabolism. Many metabolic pathways are spread over cytosolic fluid mass to the specific compartments it holds. By inter compartmental

\(^2\)RER: rough endoplasmic reticulum and SER: smooth endoplasmic reticulum, the cytosolic face of the RER are situated with ribosomes which are the sites of protein synthesis

\(^3\)Cytosol is the jel-like structure and cytoplasm is collectively cytosol with the organelles other than the content of the nucleus which is called nucleoplasm
transport mechanism, cytosol collects many essential nutrients from its surrounding organelles, and transforms and transports them in other compartments when required. The nucleus is referred as the control center of the cell. It contains DNA which encodes the genetic information that controls the development and function of living organisms by spreading the instruction through gene expression. Cell membrane separates the interior of the cell from outside environment and controls the movement of substances in and out of the cells through its microscopic channels plasmodesmata. Cell wall situated outside the cell membrane provides physical support and protection to the cell. Cell wall forms part of transport system for water and other solutes to pass through plant tissues. Cellulose is the major chemical component in the cell wall.

1.2.1 Important metabolic functions

In this thesis we analyzed several plant metabolic pathways including the pathways briefly described below. For detailed description of these important metabolic pathways, readers are referred to [15, 16].

1.2.1.1 Photosynthesis

Photosynthesis, a well-known biochemical process occur in plants (also algae and cyanobacteria) which capture the light energy (Sun light) to produce chemical energy such as 3 carbon sugars which are used to form glucose ($C_6H_{12}O_6$, see equation 1.1) that can be later released as a supply of energy for all activities of the organism. Thus in this process plants make their own food by taking inorganic nutrients and photons. In this process oxygen is released as byproduct, so it is photosynthesis that maintains huge atmospheric oxygen levels (and other organic compounds - source of energy) necessary for life on earth.

A trivial representation of the entire process is:

$$6CO_2 + 6H_2O (+\text{Light energy}) \rightarrow C_6H_{12}O_6 + 6O_2 \quad (1.1)$$
Photosynthesis occurs in two parts: cyclic and non-cyclic photophosphorylation (together can be referred to as light reactions) that take place on the thylakoid membranes and Calvin cycle that take place in the stroma of chloroplasts. Figure 1.2 shows the overall process. Light reactions are responsible for capturing light energy and forming *currency metabolites* ATP and NADPH, which are then used in the Calvin cycle. Brief description of both processes are described in the next sections.

### 1.2.1.2 Light reactions

The phosphorylation of ADP to form ATP in photosynthesis using the energy of sunlight is called photophosphorylation. It occurs in two ways - cyclic and non-cyclic. Light reactions occur in thylakoid, where the conversion of light energy to chemical energy is initiated. Thylakoid contains pairs of photosystems called photosystem I (PSI) and photosystem II (PSII). Photosystems contain a network of photon capturing pigment molecules called chlorophyll. Within chlorophyll the absorbed light energy excites electrons to a higher state.

Non-cyclic photophosphorylation starts with PSII where the energized electrons are passed from the reaction center of PSII to an electron transport chain (ETC). The electrons lost by PSII are replaced by electrons which come from oxidation.
of water in a process called photolysis, producing free electrons and oxygen gas. As electron pass through ETC, the energy from the electron are used to pump hydrogen ion from the stroma to the thylakoid (Cytochrome b6f complex), creating a concentration gradient. This gradient powers a protein called ATP synthase which phosphorylates ADP to form ATP. The low energy electron from PSII arrives in PSI. Inside the PSI the low energy electron re-energizes and passes through ETC where they are used to reduce NADP$^+$ to NADPH.

In cyclic photophosphorylation only PSI is used, where electrons from PSI comes to cytochrome b6f and returns to PSI (hence called cyclic) and creates the proton gradient which is used to form ATP by ATP synthase as before. In this process NADPH is not produced and as electron is returned to its origin, the oxidation of water is not required to release electron into PSI. This process is important to produce more ATP and stop the production of NADPH.

1.2.1.3 Calvin cycle

The Calvin cycle is a part of photosynthesis where inorganic atmospheric carbon dioxide is converted into organic molecules. The set of reactions involved are also called carbon fixation because it fix the carbon into organic form in the plant leaves. The cycle was discovered by Melvin Calvin, James Bassham and Andrew Benson at the University of California, Berkeley. Energy molecules such as ATP, generated by the light reactions are used as the energy source here. The key enzyme Rubisco plays an important role in this pathway by catalyzing ribulose-1,5-bisphosphate (RuBP) and carbon dioxide to a three carbon sugar 3-phosphoglycerate (PGA). In this process RuBP regenerated (figure 1.3), hence, completes the cycle. Surplus glyceraldehyde 3-phosphate (GAP) produced in this cycle is used to produce glucose and other carbohydrates.
Figure 1.3: A representation of partial central carbon metabolism in plants. Calvin cycle is shown in green, Pentose Phosphate Pathway is shown in red, Glycolysis is shown in blue and TCA cycle is shown in pink (without ETC).
1.2.1.4 Pentose phosphate pathway

Pentose phosphate pathway (PPP) is an alternative route for the breakdown of glucose-6-phosphate, aiming to generate NADPH and precursors for various biosynthetic pathways. There are two phases of this pathway: oxidative and non-oxidative. In the first stage NADPH is generated and in the second stage it is used for the production of 5-carbon sugars. In plants most steps occur in plastids. Ribose-5-phosphate (R5P) and erythrose-4-phosphate (E4P) are generated in the second phase and is used in the synthesis of nucleotides and aromatic amino acids, respectively. Figure 1.3 shows the PPP occurring in plants.

1.2.1.5 Glycolysis

Glycolysis converts glucose into pyruvate. Plant glycolysis occurs both in cytosol and plastids. This multi-step metabolic pathway (several enzymes operating within the process) can be separated into two phases: investment phase of ATP and pay off phase in which ATP is produced. In the first and third step where glucose is converted to glucose-6-phosphate and fructose-6-phosphate is converted to fructose-1,6-bisphosphate, respectively; ATP is required as an energy source. In the later stage, each reaction occur twice per glucose molecule (since glucose leads to two triose sugars in the investment phase), so there is a production of two NADH and four ATP molecules and has a net yield of two molecules of ATP and NADH in the pay-off phase. See figure 1.3 for the pictorial view of this description.

1.2.1.6 TCA cycle

Tricarboxylic acid cycle (TCA cycle), or the Krebs cycle is an energy production pathway (cellular respiration) in the form of ATP in mitochondria of eukaryotic cells and in the cytosol of prokaryotes. This pathway also provides precursors of certain amino acids and NADH which is used in other reactions and in electron transport chain (ETC). In the first step two carbon acetyl group from acetyl-CoA
is transferred to four carbon oxaloacetate (OAA) to form six carbon compound citrate (Cit). One of the main source of acetyl-CoA is the decarboxylation of pyruvate that comes from glycolysis. The citrate goes through series of chemical transformation and at the end four carbon oxaloacetate is regenerated and completes the cycle. Here, oxygen is used and carbon dioxide is released. Many TCA cycle intermediates (such as 2-KG) are also used as precursors for the biosynthesis of other molecules. NADH fed into the oxidative phosphorylation (electron transport) which uses their structure (inner membrane) and energy released by the oxidation of nutrients to form ATP. Overall the TCA cycle works together with ETC to generate ATP efficiently and that’s why mitochondria is often called the power house of the cell. The full cyclic TCA cycle is just a one way of its representation, many other non-cyclic modes might occur depending on the cellular need [17]. Full cyclic TCA cycle is shown in figure 1.3.

1.3 Metabolic model

Complete genome sequences of number of organisms are available and now it is a matter of bioinformatics and/or computational approach to extract biologically significant information from them. In this process a model (replica representation of a system) is reconstructed by collecting the reactions present in the system (e.g., a cell) which represents its metabolic capabilities, hence, called metabolic model. The first step of this reconstruction begins with collecting the genome-scale reaction information available in the biochemical reaction databases.

The overall metabolism of a cell and its responses at varying conditions can be studied using a metabolic model of the cell. The reactions, their stoichiometries, the kinetic parameter and the directionality of a reaction are in general used to describe a metabolic model. Beginning of this process is the collection of the set of reactions of the system under investigation. Several experimentally validated and manually curated reaction or pathway databases are present such as BioCyc [18], KEGG [19], RiceCyc (http://www.gramene.org/pathway/ricecyc.html),
BRENDA (www.brenda-enzymes.info), etc. The steps of the reconstruction follows: collection of the reactions, curation of them and presenting it in a software (metabolic modelling tool) readable format. Initial draft of a model may have missing information of metabolites, improper reaction stoichiometry, duplication of reactions [20] and sometime missing reaction(s) in a pathway. Thus, an initial version of a metabolic model may not be ready to observe responses therefore, lacks prediction ability. Therefore, model curation is very important, despite being the most tedious part of any metabolic model reconstruction. Sometime it needs rigorous literature survey to make a model work (like taking nutrients and producing all biomass precursors). So reconstruction and result analysis is a simultaneous process and after each refinement model prediction is improved. Final version of a model should be capable of predicting responses when different constraints are imposed on them. These constraints are analogous of several environmental and stress conditions (described later) and then it becomes capable of generating hypothesis about actual cellular behavior.

As reactions are used to build the model and represents the metabolism of an organism, it becomes a significant platform to gain detail knowledge about the system and biologically meaningful predictions. The activity from gene to enzyme level can be supplied into a metabolic model using additional constraints to make a ‘complete picture’. When the reactions present in the cell’s entire genome are used [3], then it is referred as a genome-scale model (GSM). In specific, a genome-scale structural metabolic model is the list of reactions (extracted from databases) aiming to represent the entire metabolic network of a species.

1.3.1 Databases used for the construction of metabolic models

Small network models representing small section of the cellular metabolism (simply a single pathway) can be constructed by manual reaction addition, however, metabolic networks containing more than hundreds of reactions, usually starts
with collecting the reaction information from reaction databases. Usually reactions from these databases are rewritten in a particular format (automatically using a computer code) that can be read by metabolic modelling softwares like ScrumPy [21]. These type of softwares/tools are designed to read models, prepare it for method application and gives model responses and analysis scope (would be discussed later in detail). Several reaction databases are available and they contain metabolic information of many organisms which are used to build genome-scale models. BioCyc contains pathway/genome databases of completely sequenced genomes of prokaryotic and eukaryotic species. MetaCyc is a database of experimentally elucidated metabolic pathways from all domains of life and curated from scientific literatures [22]. RiceCyc database has more than thousand reactions of annotated rice (Oryza sativa japonica ‘Nipponbare’) genome sequence under Gramene, (www.gramene.org) [23] a comprehensive genome mapping database for cereals. Gramene includes annotated genomic information of other important crops like maize, sorghum, millet, wheat, etc. KEGG is also a pathway database of several species which has reaction and metabolite information necessary for model building. These databases have the information of gene → protein → enzyme → reaction → metabolites that reflects the extended version of central dogma of molecular biology thus, representing itself as a metabolism of an organism. Hence, the models build by these databases, in principle, are capable of predicting actual responses of an organism.

1.3.1.1 Inconsistencies present in databases

If error present in the reaction(s) is induced in the model then it can influence the metabolic responses. Then the prediction by the model might become wrong as it might provide non-realistic cellular feature. Most of the time there is a high chance of having contradiction with biologically significant features in the first version of any metabolic network model. These contradictions arise mainly due to inconsistencies present in its originating databases. Problem existing in the databases gives rise to some errors in reconstructed network, such as the presence of ‘orphan’ and ‘dead-end’ metabolites, disconnected subnetwork within an organism.
and violation of mass conservation of reactions [5]. Poolman et al. (2006) identified some features of databases that give rise to the discrepancies in the metabolic network. These are non-unique metabolite identifiers, same metabolites reported as both substrate and product in a reaction, reactions having incorrect stoichiometry and incorrect empirical formula of metabolites, etc. Apart from these, some error in databases could not be found computationally and it needs expert biological knowledge to reconstruct more accurate models. Table 1.1 taken from Poolman et al. (2006) summarizes the errors or the inconsistencies present in KEGG and BioCyc databases in their earlier versions.

<table>
<thead>
<tr>
<th></th>
<th>KEGG</th>
<th>BioCyc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Reactions</td>
<td>6576</td>
<td>5071</td>
</tr>
<tr>
<td>Total metabolites</td>
<td>5538</td>
<td>4846</td>
</tr>
<tr>
<td>Unbalanced Reactions</td>
<td>6.9</td>
<td>1.3</td>
</tr>
<tr>
<td>Same metabolite</td>
<td>1.5</td>
<td>2.1</td>
</tr>
<tr>
<td>Orphan Metabolites</td>
<td>41.2</td>
<td>51.9</td>
</tr>
<tr>
<td>Dead-end metabolites</td>
<td>11.6</td>
<td>7.6</td>
</tr>
</tbody>
</table>

Table 1.1: Discrepancies present in KEGG and BioCyc. Adapted from [5]. Results expressed as the percentages of totals.

It appears that automatically reconstructed large-scale metabolic models from these databases are not fully consistent. Thus, the challenging task for the metabolic modellers is to build a biologically significant model that can predict close to the realistic behavior of an organism. An iterative manual and computational model curation process can solve these problems as shown in figure 1.4.

### 1.3.2 Types of metabolic model

Majority of metabolic models are of two types: kinetic and structural models. In kinetic modelling approach, the cellular process under investigation is defined by reactions described by nonlinear differential equations called reaction rate equations, values of kinetic parameters and initial metabolite concentrations, etc [5, 24].
To capture the time evolution of a process inside a reacting system, various types of data are needed for representing the kinetic parameters that control the process. Complex biological processes depend on several parameters at a time to reach from an initial to final state and the calculation of such processes are not easy. Thus, kinetic models are used to predict relatively small sections of a cell’s dynamic behavior, depending on many parameters acting on the section. Due to non-availability of different enzyme-kinetics data and complex computational approaches; the kinetic modelling technique is not used for large metabolic networks. On the other hand, structural modelling approach is relatively less complex and normally independent of the requirement of kinetic data. It is able to predict results well in agreement with experimental validations [3, 25, 26]. In structural modelling, reactions with proper stoichiometry and directionality are required to represent the metabolism of an organism. After the reconstruction, gene to reaction relationship for a given condition could be imposed. The system is considered under steady state, i.e., the production and consumption of any metabolite is balanced and no long term change in their concentration would occur. Different approaches have been developed to analyze structural models (a network of reactions) which are discussed in Chapter 2. To date many structural models are available, simulations of them generate possible cellular metabolic responses and new hypothesis that can be experimentally verified. Figure 2.1 shows a simple (‘toy’) structural model.

1.4 Metabolic modelling

1.4.1 Systems biology

Present day understanding is not confined to small number of variables but it is now possible to study biological systems with a global analytical approach [27]. Understanding the biological phenomena in genome level necessitates observation of several organelle synergy within a cell. The purpose of systems biology is to study the biological system not as a small section of the whole but as an integrated
system of genes, proteins, enzymes, reactions, metabolites and pathways studied altogether.

### 1.4.2 Representation and target

After a metabolic model is reconstructed using databases, it compiles into a mathematical model so that different methods can be applied. An iterative process of model validation gives a comprehensive analysis scope of the model responses and allows identification of key features of metabolism such as stress response, growth yield, etc. This knowledge can then be applied to metabolic engineering for novel outcome.

Metabolic modelling aims to gain comprehensive insight into the molecular mechanism and major advantage of this is that it allows to observe responses of an organism in the perspective of the entire network (hence, genome) level. So with *in-silico* investigation of plant specially by metabolic modelling, it becomes possible to zoom into the living systems deeper; the flow of this process is shown in figure 1.4.
1.5 Large scale models of plant metabolism

First large scale model of plant was published in 2009 on barely seed metabolism [28]. The model includes 257 biochemical and transport reactions. The predictions of this model in oxygen depletion and enzyme deletion conditions were in agreement with experimental observations. In addition, predicted growth rates in different conditions were in accordance with published experimental results. After few months, genome-scale metabolic model of heterotrophic Arabidopsis cell (Arabidopsis thaliana) was constructed using AraCyc database [25]. This model consists of 1253 metabolites and 1406 reactions. The model is simulated at varying ATP demands and the observations are in agreement with estimates in prokaryotes and yeast. In 2010, C4 genome-scale model was constructed for mesophyll and bundle sheath cells to observe C4 photosynthesis [29]. It has 1588 reactions, 1755 metabolites, 83 inter-organelle transporters and 29 external transporters. This C4 photosynthetic model predicts the classical C4 biosynthesis pathway and important metabolic interactions in both the cells. Then compartmentalized tissue specific genome-scale model of Arabidopsis thaliana was constructed [30]. First photosynthetic genome-scale model of crop plant rice (Oryza sativa) was published in 2013 representing a mesophyll cell of an expanding leaf [3]. The model behaviors were investigated at varying light intensities (i.e., varying the incident photon), and some important interactions between chloroplast and mitochondria in different light levels along with role of photorespiration in high light were identified. Soon afterwards the effect of flood and drought stresses were investigated in a rice model representing two tissue types: germinating seeds and photorespiring leaves [31]. Multiscale metabolic modelling (MMM) on barley plant is implemented to integrate organ specific models in a whole plant dynamic model to find its metabolic behavior [32].
1.6 Model predictions

To date many evidences are available to show that the prediction from models are consistent with experiments. In the first GSM of photosynthetic plant, i.e., rice, we have showed that the model could predict different metabolic states of a rice leaf which are supported by known experimental evidences [3]: plants sometimes get exposed to superoptimal light and the excess energy received by the photon should get expend somewhere so as to prevent any damage and photorespiration is among them [33], quantum demand becomes high when nitrate used as the nitrogen source [34]. Model of Arabidopsis [25] predicts that with additional ATP demand (such as required for growth and maintenance) full TCA cycle operates with higher flux through glycolysis and TCA cycle and this flux response is matched with isotope labeling experiments [35]. Here, activity of rubisco in non-photosynthetic model is interesting but the role of rubisco for lipid synthesis and flux distribution from G6P and PGA was identical for certain oilseeds [36] in low energy solutions. The estimation of total ATP requirement shows prediction ability of this model. When a trade-off between ATP and NADPH production used for cell maintenance along with metabolite transport cost in the Arabidopsis flux balance model, improve accuracy of fluxes are predicted that is confirmed by $^{13}$C-MFA [4]. Further, introduction of day and night condition in Arabidopsis GSM representing C3 leaf metabolism predicts citrate would be synthesized during night by TCA cycle and stored in the vacuole and in day it is exported and metabolized to 2-oxoglutarate and this prediction is consistent with isotopic labeling experiments [37]. This study also predicts that there is no overall energetic advantage to Crassulacean Acid Metabolism (CAM). Central carbon metabolism is responsible for converting sugars into metabolic precursors and prediction of fluxes of this is important for in-silico studies. Genome-scale model of Arabidopsis accurately predicts fluxes occurring in vivo, measured by $^{13}$C metabolic flux analysis (MFA) in the glycolysis and TCA cycle; in increased temperature and hyperosmotic stress the model was able to predict the corresponding flux changes [38]. This study shows usability of metabolic modelling for prediction of fluxes or
cellular strategies in stress conditions. Conventional TCA cycle\(^4\) is the one way within number of its reaction organization possibilities [17]. Genome-scale models of Arabidopsis and rice both predict non-cyclic TCA cycle modes depending on the cellular need [3, 25] and it is well established that in plants alternative TCA flux modes are feasible and could be advantageous for metabolic flexibility [17, 39]. The reconstruction of Arabidopsis GSM helped de Oliveira Dal’Molin et al. (2010) to predict that to compensate for photorespiration, 30%-50% increase in photosynthesis was needed which matches experimental measurements. It’s a worldwide concern to increase biofuel from plants but its a challenge to improve biofuel from cellulose and decrease lignin content for efficient and cost effective production [41]; in this thesis, applying the gene expression data as an additional constraint we observe that the rice GSM can predict high cellulose and low lignin contents similar to the experimental work (see section 2.7.1). Large scale network study identified ‘superessential’ reactions that could be common for all metabolic networks (some of them experimentally confirmed for different organisms) and can serves as a drug target [42]. Using a genome-scale model of \textit{E. Coli} it is observed that incoming and outgoing flux of proton from the medium is important for maximizing cellular growth [43]. Using FBA, many researchers analyze bacterial models in the context of maximization of biomass precursors (optimizing growth) that matches experimental evidences [26, 44–46]. \textit{In-silico} metabolic model of \textit{E. Coli} quantitatively predicts the growth potential in a gene deletion study of central metabolic pathway [47]. Feist et al. (2007), using updated metabolic network of \textit{E. Coli} shows that predicted growth and acetate secretion rate are matched with experiment using glucose and oxygen uptake rates as modelling constraints. In these results, authors also found complete agreement in the modelling and experimental flux direction of central metabolism. Key aspects of network functionality, robustness and gene regulation are possible to predict simultaneously using theoretical analysis of structural models [49].

Large scale network models are used to predict functional effect of a drug in human [50]. A number of human metabolic models have been constructed [51] to visualize

\(^{4}\)Full cyclic TCA cycle shows a cycle of Acetyl-CoA converted to different intermediates and reproducing it again
the effect of perturbations caused by the drug in the organism in quick and less expensive manner. Using metabolic model of human kidney, specific metabolic enzymes that causes the side effects of a drug target are identified, the method used is also useful for the implication of personalized medicine [50]. Prior to this analysis, the human metabolic network was used to predict alternative drug targets observing correlated reaction sets [52]. Genome-scale model analysis also suggests that partial inhibition of small number of drug targets is more efficient than complete inhibition of a single target [53]. *E. Coli* metabolic network is used to design algorithm that can stop production of target compound with minimum side effects [54].

### 1.7 Importance of metabolic modelling

This is discussed at the end of this chapter as now I can concisely explain the importance of metabolic modelling. As of now we have come across vast application of modelling and its prediction capabilities in plants (quality, quantity, stress tolerance improvement, etc.) and in bacterial and human metabolic models (e.g., drug target identification). It is well established now that with this holistic approach living systems can be better understood. Different biotic and abiotic conditions can be created and responses in effect of those of a living object can be obtained in less time and analysis could result important biological discoveries (that could be really challenging only with experiments). Using computational techniques we also can check beneficial metabolic changes in a species [55]. Collectively, outcome of the model building and its analysis can enrich the knowledge of the biochemistry and also can help in metabolic engineering. So experimental and theoretical researchers may come in a win-win situation collaboratively so as to reach their common goal quickly.
1.8 Aim and organization of the thesis

With the advancement of many modelling and high throughout techniques, there is a need to utilize their features for novel outcome. This thesis presents the utilization of a genome-scale metabolic model of rice to identify some complex biological features encoded into rice (Oryza sativa) metabolism.

Rice is an important crop for almost half of the world’s human population for their major caloric intake. Understanding its system level responses should help metabolic engineers to design high stress tolerant and more yielding cultivars. This improvement is required to deal with the future challenge of food security for billions of people worldwide.

Here, I proceed to describe the methodology used to extract meaningful information from metabolic models in Chapter 2. Then the curation process of model reactions is described in Chapter 3. The important reactions which are necessary to produce major biomass components and effects of their absence in genome level are identified in Chapter 4. To pave the way for further understanding of the less known phenomena of inter-organelle metabolite transport, the influence of inter-compartmental transporters in genome level is analyzed in Chapter 5. The procedure constructed for this is verified here using experimental data, taken from rice leaf cell. Plants must face different light intensities throughout the daytime and over the seasons. This routine change in incident photon and in effect of this the alteration in metabolic level is a parallel process. The metabolic readjustments during change in photon absorption are analyzed in Chapter 6, to enhance our understanding in this change and it’s inherent adaptation techniques. The effect of variation of ATP/NADPH ratio has also been presented here.

The work presented here, intent to predict the important metabolic features imprinted in rice genome. Our insight into the complex metabolic interactions and method development to observe the metabolic responses of rice should help biotechnologists to design rational experiment for improvement/understanding this plant. Rice, can be a template to understand other crop plants [56], hence,
this work is also useful to improve knowledge about other crops. Thus, the work in this thesis and it’s future developments will gain much insight into rice metabolism for human benefits.