



A control theoretic three timescale model for analyzing energy management in mammalian cancer cells

Abhijit Dasgupta^b, Abhisek Bakshi^d, Nirmalya Chowdhury^c, Rajat K. De^{a,*}

^a Machine Intelligence Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India

^b Department of Data Science, School of Interdisciplinary Studies, University of Kalyani, Kalyani, Nadia 741235, West Bengal, India

^c Department of Computer Science & Engineering, Jadavpur University, Kolkata 700032, India

^d Department of Information Technology, Bengal Institute of Technology, Basanti Highway, Kolkata 700150, India



ARTICLE INFO

Article history:

Received 2 May 2020

Received in revised form 26 November 2020

Accepted 13 December 2020

Available online 29 December 2020

Keywords:

Pathway integration

Timescale

MIMO

Support vector regression

Genetic algorithm

Warburg effect

ABSTRACT

Interaction among different pathways, such as metabolic, signaling and gene regulatory networks, of cellular system is responsible to maintain homeostasis in a mammalian cell. Malfunctioning of this cooperation may lead to many complex diseases, such as cancer and type 2 diabetes. Timescale differences among these pathways make their integration a daunting task. Metabolic, signaling and gene regulatory networks have three different timescales, such as, ultrafast, fast and slow respectively.

The article deals with this problem by developing a support vector regression (SVR) based three timescale model with the application of genetic algorithm based nonlinear controller. The proposed model can successfully capture the nonlinear transient dynamics and regulations of such integrated biochemical pathway under consideration. Besides, the model is quite capable of predicting the effects of certain drug targets for many types of complex diseases. Here, energy and cell proliferation management of mammalian cancer cells have been explored and analyzed with the help of the proposed novel approach. Previous investigations including *in silico/in vivo/in vitro* experiments have validated the results (the regulations of glucose transporter 1 (glut1), hexokinase (HK), and hypoxia-inducible factor-1 α (HIF-1 α) among others, and the switching of pyruvate kinase (M2 isoform) between dimer and tetramer) generated by this model proving its effectiveness. Subsequently, the model predicts the effects of six selected drug targets, such as, the deactivation of transketolase and glucose-6-phosphate isomerase among others, in the case of mammalian malignant cells in terms of growth, proliferation, fermentation, and energy supply in the form of adenosine triphosphate (ATP).

© 2020 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cellular decision making and responses are orchestrated by a set of complex biochemical pathways/networks. Broadly speaking, biochemical pathways/networks can be categorized as metabolic pathways, gene regulatory networks (GRNs), and signaling pathways. A metabolic pathway is a coherent set of biochemical reactions catalyzed by a number of enzymes. It helps a living organism to transform an initial (source) compound into a final (target) compound and energy. On the other hand, fundamental information processing and control mechanisms in a cell are performed by GRNs. Regulatory genes code for proteins that activate or inhibit the expression of other genes. Thus, a complex web of

interactions, called a GRN, in terms of activation and inhibition of genes, is formed. Moreover, signaling pathways contain a series of specific actions in a cell in which a signal is passed from the environment and accordingly the cell responds. These pathways are of diverse nature and are interacting with one another. They also form a hierarchy. For example, an entire organism can be thought of as a huge network of interacting organs, each of which, in turn, is a network of interacting tissues. A tissue is a collection of interacting cells performing similar functions. A cell may be considered as a huge network of interacting components to constitute aforesaid biochemical networks. Moreover, the influence of environment and other factors on enzymes or gene regulation needs to be considered to make a study on how an organism is responsive to environmental changes. Therefore, it is necessary to focus not only on individual processes or pathways but also on their integration. Fig. S1 in Supplementary material depicts the interaction

* Corresponding author at: Machine Intelligence Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata, West Bengal 700108, India.

E-mail address: rajat@isical.ac.in (R.K. De).

among metabolic, signaling and gene regulatory networks related to central carbon metabolism in a mammalian cell. Attempts to elucidate such huge biochemical networks on the structural/ functional basis face the problem of combinatorial explosion. There exist several investigations on modeling each of these pathways individually [1–9]. However, the study on the integration of gene regulation, metabolism, and signaling events (pathways) is not much.

Among several approaches one of the most common is to explore metabolic pathway is Flux Balance Analysis (FBA) [10,11]. FBA can analyze only steady-state response of a system rather transient response. However, Metabolic Control Analysis [12] overcomes the limitation of FBA. Yet, it does not provide any supervisory controller that can regulate the enzyme/metabolite concentration to meet some particular requirements of a cell. Recently, we have developed a control theoretic approach to solve

this problem [13–15]. Some investigations on crosstalk mechanism [16] and identifying properties [17,18] of signaling networks have also been developed. Since the activation of regulatory mechanism depends on corresponding gene expression level thus any GRN can be modeled mathematically by coupled ordinary differential equations (ODEs), Boolean network or Bayesian network [19].

Response time of the interconnected complex subsystems may vary from each other. In other words, they can be treated as complex subsystems of different timescales. Biochemical pathways, such as metabolic, signaling and regulatory networks show this kind of phenomena. Timescale differences among metabolic, signaling and gene regulatory networks make their integration a daunting task. Several researchers have tried to model the dynamic behavior of a system by stoichiometric reconstruction of integrated biochemical pathways [20]. Here, the authors have considered two different timescales. The metabolic and signaling networks have

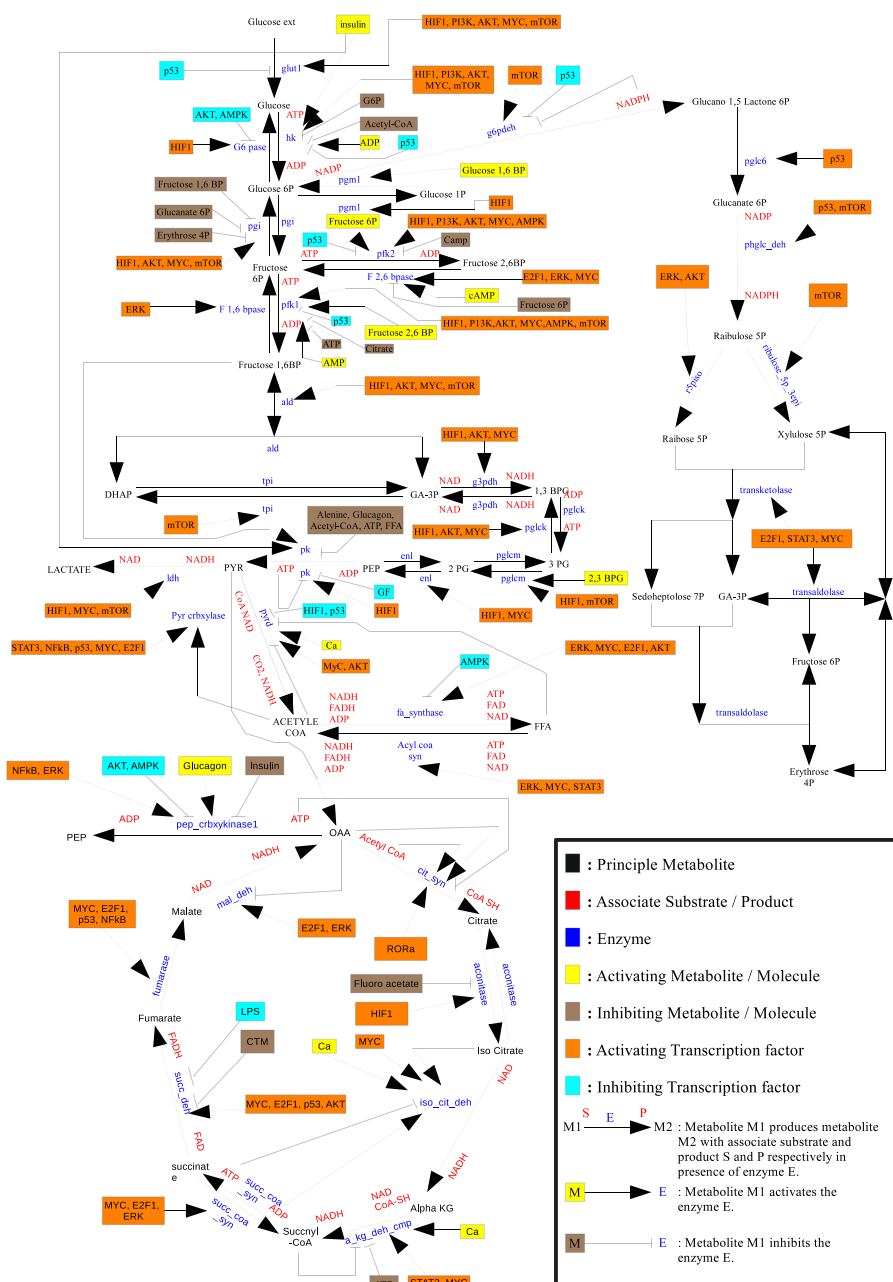


Fig. 1. Integrated metabolic, signaling and gene regulatory networks associated with central carbon metabolism of mammalian cells.

been considered in fast timescale, whereas GRN has been considered in slow timescale. However, the metabolic, signaling and gene regulatory networks participate in three different timescales. Previous investigation [21] show that proteins in *Escherichia coli* take seconds to minutes to express. It will be longer in the case of mammalian systems. On the other hand, metabolite concentrations can respond in seconds to microseconds [22]. Gene regulatory events take minutes to hours [20]. Thus, integration of these biochemical pathways becomes a three timescale problem, where metabolic, signaling and gene regulatory networks are ultrafast, fast and slow respectively.

In this study, we develop a three timescale multiple input and multiple output (MIMO) model based on support vector regression (SVR) and genetic algorithm based controller to simulate the dynamic behavior of integrated signaling, metabolic and gene reg-

ulatory networks specifically responsible for mammalian carbon metabolism in normal and cancer cells. Here, the metabolic, signaling and GRN have been mathematically modeled by sets of ODEs with proper three timescale selection, feedback, allosteric effects and perturbation. In this context, we have collected all pathway information under consideration from KEGG database [23] and literature. Finally, we have developed a genetic algorithm (GA) based controller to drive the change of concentrations/expression levels of certain metabolites/proteins/genes with respect to time in a desired fashion. Thus, the proposed model can also predict the possible effects of certain drug targets of specific diseases through GA controller.

The central carbon metabolic (CCM) network comprises glycolysis, tricarboxylic acid (TCA) cycle, pentose phosphate pathway (PPP) and free fatty acid (FFA) metabolism. Glycolysis consumes

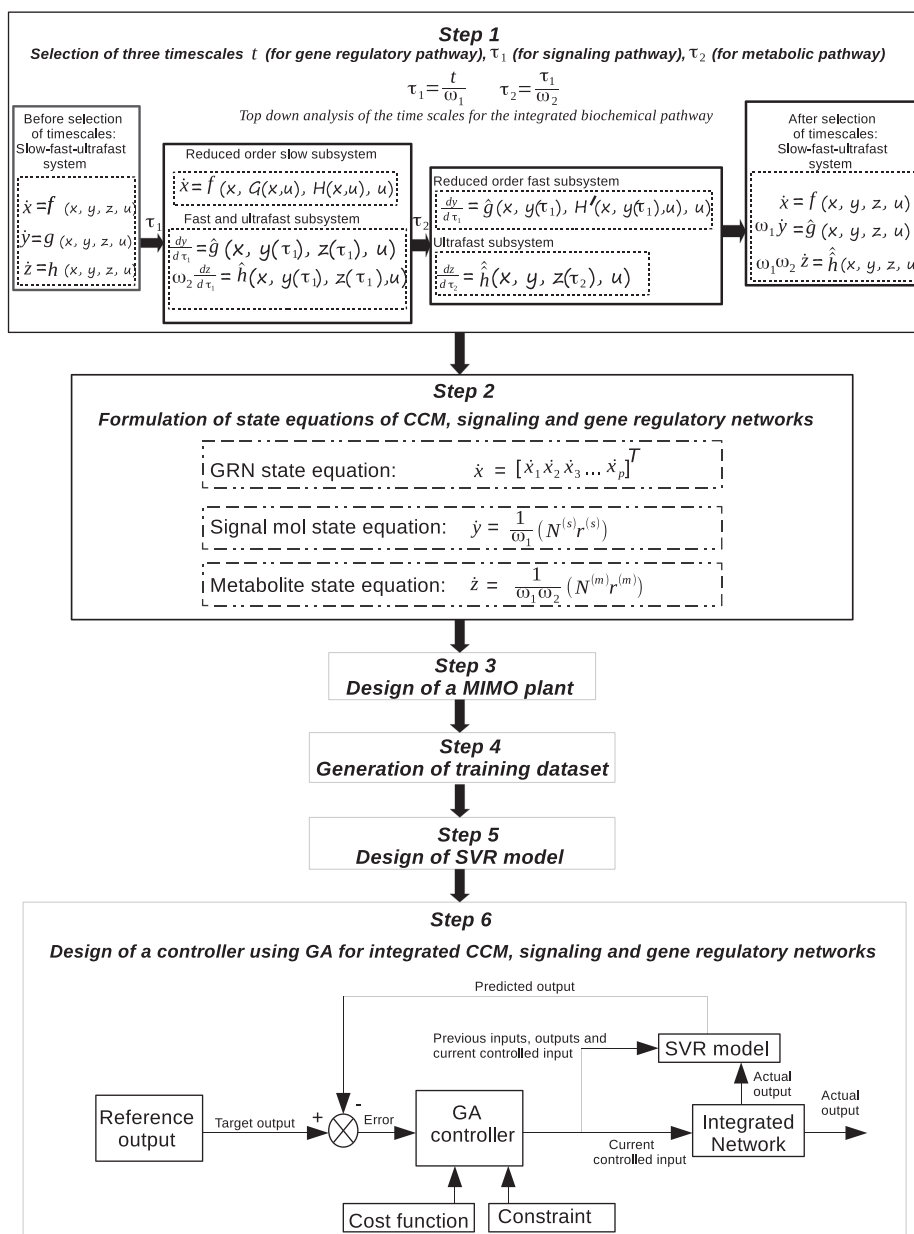


Fig. 2. Flowchart of the entire methodology: Here slow, fast and ultrafast subsystems correspond to gene regulatory, signaling and metabolic pathways respectively.

glucose to produce pyruvate and energy in the form adenosine triphosphate (ATP), although most of the energy is generated from TCA cycle taking pyruvate as input. Here, PPP is involved in macromolecular synthesis including reduced nicotinamide adenine dinucleotide phosphate (NADPH) and ribose 5P. FFA is formed by acetyl CoA of TCA cycle. However, a mammalian tumor cell slows down oxidative phosphorylation by inhibiting TCA cycle activities and produces abnormal amount of lactate. This effect is called “Warburg effect” [24,25] which may lead to cancer [26]. In this context, we have analyzed the energy and cell proliferation management of CCM pathway along with corresponding signaling and gene regulatory networks in mammalian cancer cells. Here, we have considered the interactions among different enzymes/proteins, transcription factors and genes associated with central carbon metabolism to capture the cellular dynamics during normal and malignant conditions in mammalian cells. These detailed interactions can be found in [Supplementary Tables S1–S8](#) (abbreviations of different molecules can be found in [Supplementary Table S10](#)). [Fig. 1](#) depicts the integrated metabolic, signaling and gene regulatory networks associated with central carbon metabolism. We also have analyzed the possible effects of six drug targets, such as deactivation of pyruvate kinase, glucose-6-phosphate dehydrogenase, transketolase, ribose 5P isomerase, glucose-6-phosphate isomerase and finally activation of pyruvate kinase, on mammalian malignant cells.

In accordance with previous investigations including some *in vivo* and *in vitro* experiments [24,27,26,28–45,25,46–65], the simulation results also suggest that abnormal over expressions of certain proteins and genes depicted in [Table 3](#) play significant roles to maintain growth, proliferation and energy supply in mammalian cancer cells to survive. Under expressions of prolyl-hydroxylases (PHD) and tumor suppressing protein p53 are also important in this context. Besides, switching of pyruvate kinase (M2 isoform) to its inactive dimer or active tetrameric form helps to synthesize energy in the form of ATP, and macromolecular precursors for cell growth and proliferation according to the requirements of mammalian malignant cells.

We have also compared the proposed model with the following theoretical models. In this context, FBA based model [66] has depicted a similar set of malfunctioning enzymes as identified by the proposed model in cancer cells for its growth. However, FBA based model does not involve transient analysis of the behavior of molecules in cancer pathways. Previous differential equation and optimization-based models [67–69] have also depicted similar mutation as depicted by the proposed model in cancer cells. However, these models do not consider the three timescale nature of the integrated CCM pathways. Besides, they have not explored the outcome of probable drug targets to control energy metabolism in cancer cells.

Here the proposed model is able to predict that deactivation of glucose-6-phosphate dehydrogenase and ribose 5P isomerase may slow down growth and proliferation of cancer cells. However, glucose-6-phosphate dehydrogenase deactivation may not be able to reduce fermentation and energy supply in malignant cells. On the other hand, though ribose 5P isomerase deactivation may inhibit cell fermentation, it may not get success to stop energy supply in mammalian cancer cells. In this context, deactivation of transketolase and glucose-6-phosphate isomerase may be considered as potential drug targets to resist cell growth, fermentation and proliferation as well as energy supply in human cancer cells. Activation of pyruvate kinase (M2 isoform) may also reduce cancer progression in terms of cell growth, proliferation and fermentation. Besides, it may fail to decrease ATP production in tumor cells. However, deactivation of pyruvate kinase may be a poor choice as a drug target for cancer therapy.

In summary, the proposed novel model can tackle three time-scale nature of any integrated biochemical pathway comprising metabolic, signaling and gene regulatory networks. Subsequently, it can capture the nonlinear transient dynamics of the integrated network of both normal and perturbed human cells. This has been validated appropriately by some other methods/results [70–72,66–69]. Moreover, using proposed GA controller, effects of drug targets on diseased cells have been explored and analyzed. Thus, it can help to find out a novel therapeutic goal for complex diseases, such as cancer and type 2 diabetes.

Table 1
Different mathematical symbols and their description.

| Symbols | Description |
|---|---|
| $\mathbf{x}, \mathbf{y}, \mathbf{z}$ and \mathbf{u} | Vectors representing gene expression levels, protein/signaling molecule expression levels, concentration of metabolites and external inputs (Supplementary Table S11) respectively. |
| p, s, m and c | Dimension of $\mathbf{x}, \mathbf{y}, \mathbf{z}$ and \mathbf{u} respectively. |
| i, j, k and l | Index for gene, signaling molecule, metabolite and external inputs respectively. |
| t, τ_1 and τ_2 | Timescales for gene regulatory, signaling and metabolic networks respectively. |
| $\dot{\mathbf{x}}, \dot{\mathbf{y}}$ and $\dot{\mathbf{z}}$ | Derivatives of \mathbf{x}, \mathbf{y} and \mathbf{z} respectively, with respect to time t . |
| ω_1 and ω_2 | Slow to fast and Fast to ultrafast timescale ratios. |
| $\mathbf{f}(\cdot), \mathbf{g}(\cdot)$ and $\mathbf{h}(\cdot)$ | Activities corresponding to gene regulatory, signaling and metabolic networks respectively. |
| $\hat{\mathbf{g}}(\cdot)$ and $\hat{\mathbf{h}}(\cdot)$ | Fast and ultrafast subsystems with respect to τ_1 . |
| $\hat{\mathbf{h}}(\cdot)$ | Ultrafast subsystem with respect to τ_2 . |
| $\mathbf{G}(\cdot)$ and $\mathbf{H}(\cdot)$ | Quasi-steady state equilibrium of fast and ultrafast subsystems. |
| $\mathbf{H}'(\cdot)$ | Quasi-steady state equilibrium of ultrafast subsystem. |
| e_i, b_i and d_i | Expression, basal production and decay rates respectively for i^{th} gene. |
| $\mathcal{K}_{ij}^{(g)}$ | Binding rate constant for j^{th} transcription factor required to express i^{th} gene. |
| $\mathcal{F}_{ik}^{(g)}, \mathcal{F}_{il}^{(g)}$ and $\mathcal{F}_{ij}^{(g)}$ | Binding constants for i^{th} gene corresponding to k^{th} metabolite, l^{th} external input and j^{th} signaling molecule/transcription factor respectively. |
| $\mathbf{N}^{(s)}, \mathbf{r}^{(s)}$ and q | Interaction matrix, interaction rate vector and number of interactions for signal transduction pathway. |
| $r_{\sigma}^{(s)}$ | σ^{th} element of $\mathbf{r}^{(s)}$. |
| $\mathcal{K}_{ij}^{(s)}$ | Strength of binding constant of j^{th} signaling molecule with j^{th} signaling molecule. |
| $\mathcal{F}_{jk}^{(s)}, \mathcal{F}_{jl}^{(s)}$ and $\mathcal{F}_{ij}^{(s)}$ | Binding constants for j^{th} signaling molecule corresponding to k^{th} metabolite, l^{th} external input and j^{th} signaling molecule respectively. |
| $\mathbf{N}^{(m)}, \mathbf{r}^{(m)}$ and n | Stoichiometric matrix, metabolic flux vector and number of reactions for metabolic pathway. |
| $r_{\rho}^{(m)}$ | ρ^{th} element of $\mathbf{r}^{(m)}$. |
| $\mathcal{K}_{\rho}^{(m)}$ and $\mathcal{K}_{\rho}^{(m)}$ | The kinetic rate constant and Michaelis Menten constant respectively for ρ^{th} metabolic reaction. |
| $\mathcal{F}_{kk}^{(m)}$ and $\mathcal{F}_{kl}^{(m)}$ | Feedback constants for k^{th} metabolite corresponding to k^{th} metabolite and l^{th} external input respectively. |
| $\theta, \tau, \mathbf{v}, d, \mathcal{D}$ and ϕ | Time, number of past inputs, input vector for SVR, dimension of input vector \mathbf{v} , dataset and number of samples respectively. |
| μ and a or b | Index for output and \mathbf{v} respectively. |
| $d', \mathbf{F}, \mathbf{w}_{\mu}, \delta(\cdot)$ and β_{μ} | Dimension of projected feature space, projected feature space, the orientation of the hyperplane in F , mapping from d dimensional input space to d' dimensional transformed feature space F and the bias respectively. |
| $\gamma_{\mu}, \hat{\gamma}_{\mu}, \mathcal{P}_{\mu}$ and \mathcal{Q}_{μ} | μ^{th} element of actual output, μ^{th} element of predicted output, convex optimization problem and quadratic programming problem respectively. |
| $\varepsilon_{\mu}, \varphi_a, \varphi_a^*$ and \mathcal{E} | Tolerable error, slack variables and a trade-off respectively. |
| $\zeta_a, \zeta_a^*, \zeta_b$ and ζ_b^* | Lagrange multipliers. |
| α_b | $(\zeta_b - \zeta_b^*)$. |
| C and G | Number of chromosomes and generations respectively. |

2. Method

Here, we describe the proposed methodology for developing the model that will mimic the behavior of an integrated pathway system. The methodology involves six steps. At first, timescales of metabolic, signaling and gene regulatory networks have been selected, and the basic dynamics of the integrated pathway system has been described. Then, we have formulated the state equations for the integrated pathway. In the third step, an appropriate MIMO plant has been developed by tuning its parameters. Subsequently, training dataset has been generated from the appropriately perturbed MIMO plant/system. Thereafter, SVR model is developed to approximate the MIMO plant. Finally, GA controller is applied on the plant as a model predictive controller to capture nonlinear transient cellular dynamics. Fig. 2 depicts the flowchart of the entire methodology. Here we have considered central carbon metabolism related metabolic (glycolysis, TCA cycle and pentose phosphate pathways as well as FFA synthesis/consumption), signaling and gene regulatory networks throughout the entire methodology. All the mathematical symbols used in this article are defined in Table 1. However, each of these symbols has again been described whenever it appears first.

2.1. Selection of three timescales and the basic dynamics of an integrated pathway system

In an integrated biochemical pathway system, involving metabolic, signaling and gene regulatory networks, the rate of a metabolic reaction is controlled by an enzyme/protein, which is originated from an expressed gene participating in a gene regulatory network. The expression of this gene is in turn regulated by one or more transcription factors (proteins) which are expressed/activated by signaling molecules in a cascade of signal transduction pathway.

Let x_i be the expression level of i^{th} ($i = 1, 2, \dots, p$) gene generating i^{th} enzyme involved in a certain metabolic pathway. Similarly, let y_j be the expression level of j^{th} ($j = 1, 2, \dots, s$) protein/signaling molecule, involved in a signal transduction pathway. Let z_k be the concentration of k^{th} ($k = 1, 2, \dots, m$) metabolite involved in the metabolic pathway under consideration. Besides, there are c external inputs u_l ($l = 1, 2, \dots, c$) applied to the system, depicting perturbation from environment and other pathways to the current integrated pathway system. The terms $\mathbf{x}, \mathbf{y}, \mathbf{z}$ and \mathbf{u} are p, s, m and c dimensional vectors, respectively, for the pathway system.

Thus the proposed nonlinear state space model for the integrated pathway can be defined as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}) \tag{1}$$

$$\dot{\mathbf{y}} = \mathbf{g}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}) \tag{2}$$

$$\dot{\mathbf{z}} = \mathbf{h}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}) \tag{3}$$

Here $\mathbf{f}(\cdot)$, $\mathbf{g}(\cdot)$ and $\mathbf{h}(\cdot)$ represent activities corresponding to gene regulatory, signaling and metabolic networks. The time (t) derivatives of \mathbf{x} , \mathbf{y} and \mathbf{z} are represented by $\dot{\mathbf{x}}$, $\dot{\mathbf{y}}$ and $\dot{\mathbf{z}}$ respectively. State variable \mathbf{x} is slow, whereas \mathbf{y} and \mathbf{z} are fast and ultrafast state variables respectively. The symbols ω_1 and ω_2 correspond to slow to fast and fast to ultrafast timescale ratios. Here, ω_1 and ω_2 are small positive quantities. According to previous investigation [21], proteins in *Escherichia coli* take minutes to express. It will be longer in the case of mammalian systems. However, metabolite concentrations can alter in seconds [22]. On the other hand, gene regulatory events take hours [20]. Based on these assumptions, we have considered $\omega_1 = 1/60$ and $\omega_2 = 1/60$, and thereby the

stretched timescales as $\tau_1 = t/\omega_1$ and $\tau_2 = \tau_1/\omega_2 = t/(\omega_1\omega_2)$. Thus the timescales for gene regulatory, signaling and metabolic networks are t , τ_1 and τ_2 respectively.

We now obtain the fast ($\hat{\mathbf{g}}(\cdot)$) and ultrafast ($\hat{\mathbf{h}}(\cdot)$) subsystems corresponding to signaling and metabolic networks from the top down timescale decomposition [73,74] as

$$\omega_1 \dot{\mathbf{y}} = \frac{d\mathbf{y}}{d\tau_1} = \hat{\mathbf{g}}(\mathbf{x}, \mathbf{y}(\tau_1), \mathbf{z}(\tau_1), \mathbf{u}) \tag{4}$$

and

$$\omega_1 \omega_2 \dot{\mathbf{z}} = \omega_2 \frac{d\mathbf{z}}{d\tau_1} = \hat{\mathbf{h}}(\mathbf{x}, \mathbf{y}(\tau_1), \mathbf{z}(\tau_1), \mathbf{u}) \tag{5}$$

Here, $\mathbf{y}(\tau_1)$ and $\mathbf{z}(\tau_1)$ are instantaneous \mathbf{y} and \mathbf{z} -values, respectively, in τ_1 timescale. Moreover, considering $\omega_1 \rightarrow 0$, if we solve Eqs. (4) and (5) algebraically for \mathbf{y} and \mathbf{z} , we get

$$\mathbf{y} = \mathbf{G}(\mathbf{x}, \mathbf{u}) \tag{6}$$

and

$$\mathbf{z} = \mathbf{H}(\mathbf{x}, \mathbf{u}) \tag{7}$$

Two vector functions $\mathbf{G}(\mathbf{x}, \mathbf{u})$ and $\mathbf{H}(\mathbf{x}, \mathbf{u})$ represent the quasi-steady state equilibrium [75] of the fast ($\hat{\mathbf{g}}(\cdot)$) and ultrafast ($\hat{\mathbf{h}}(\cdot)$) subsystems. Moreover, being a slow state variable in timescale t , \mathbf{x} does not change significantly with respect to fast timescale τ_1 . Thus, the reduced slow gene regulatory subsystem can be defined using top-down timescale decomposition [73,74] as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{G}(\mathbf{x}, \mathbf{u}), \mathbf{H}(\mathbf{x}, \mathbf{u}), \mathbf{u}) \tag{8}$$

Again, the fast and ultrafast subsystems corresponding to signaling and metabolic networks can be treated as a two timescale system. The ultrafast subsystem corresponding to metabolic network ($\hat{\mathbf{h}}(\cdot)$) is given by

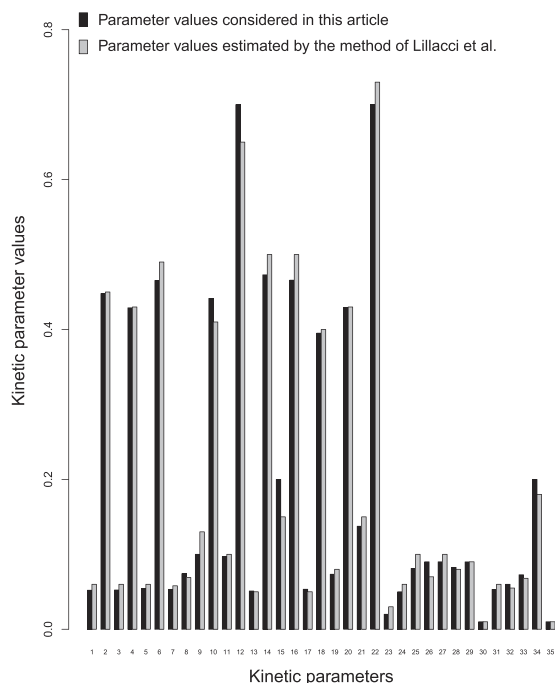


Fig. 3. Comparison of kinetic parameter values considered in the present article with those estimated by the method of Lillacci et al. [70]. Here first 22 kinetic parameters represent glycolytic Michaelis-Menten constants, whereas last 13 kinetic parameters correspond to glycolytic kinetic rate constants.

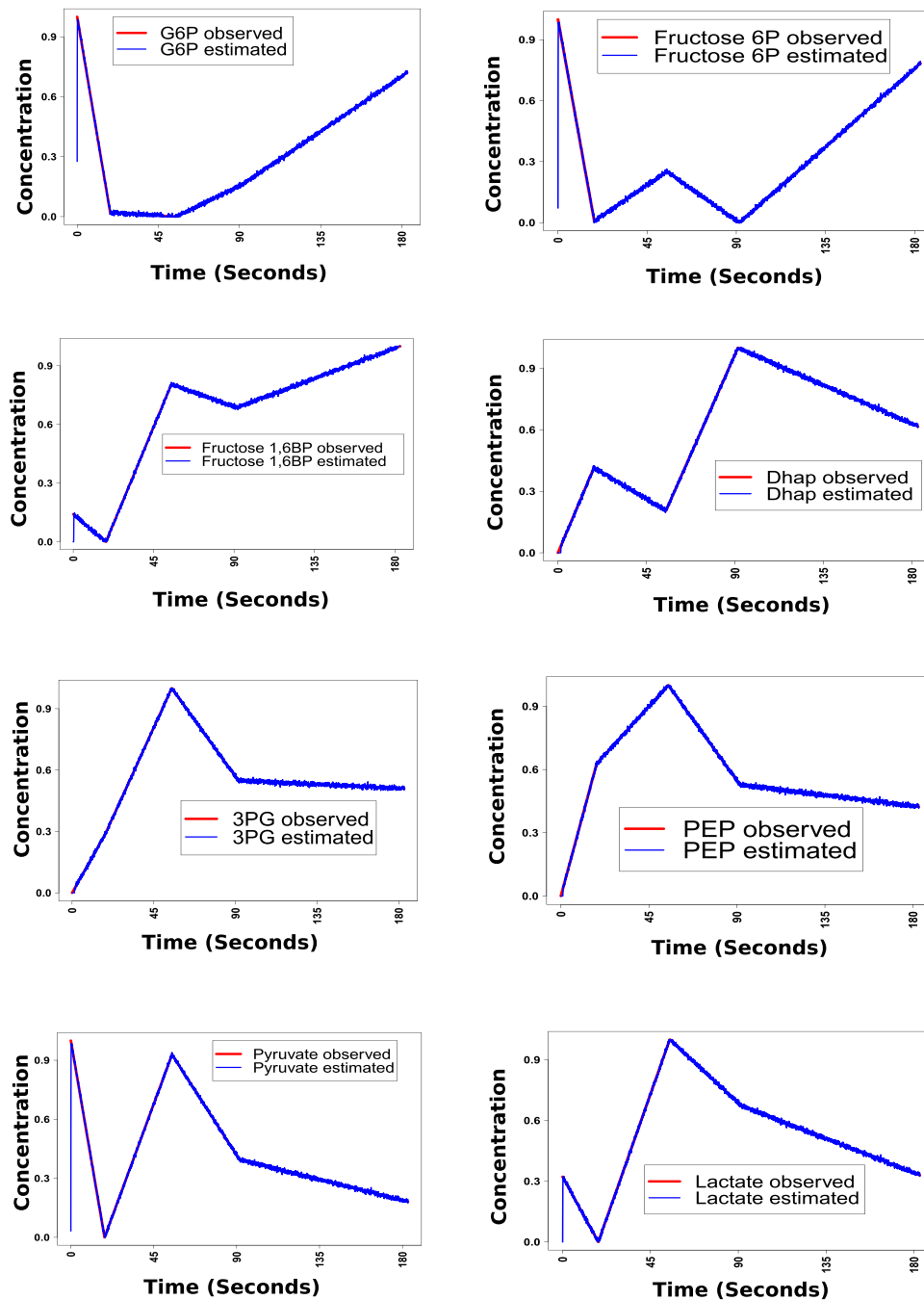


Fig. 4. Estimated states have been overlapped with observed true values of the corresponding states during the estimation of 35 kinetic parameters using hybrid extended Kalman filtering based method [70].

$$\omega_2 \frac{dz}{d\tau_1} = \frac{dz}{d\tau_2} = \hat{\mathbf{h}}(\mathbf{x}, \mathbf{y}, \mathbf{z}(\tau_2), \mathbf{u}) \quad (9)$$

Now, if we solve Eq. (9) algebraically considering $\omega_2 \rightarrow 0$, we can get

$$0 = \hat{\mathbf{h}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}) \Rightarrow \mathbf{z} = \mathbf{H}'(\mathbf{x}, \mathbf{y}(\tau_1), \mathbf{u}) \quad (10)$$

The term $\mathbf{H}'(\mathbf{x}, \mathbf{y}(\tau_1), \mathbf{u})$ represents the quasi steady state equilibrium [75] of the ultrafast subsystem $\hat{\mathbf{h}}(\cdot)$. Here, \mathbf{x} and \mathbf{y} do not change appreciably with respect to τ_2 . Thus, the reduced fast signaling network can be defined as

$$\frac{dy}{d\tau_1} = \hat{\mathbf{g}}(\mathbf{x}, \mathbf{y}(\tau_1), \mathbf{H}'(\mathbf{x}, \mathbf{y}(\tau_1), \mathbf{u}), \mathbf{u}) \quad (11)$$

Thus,

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}) \quad (12)$$

$$\dot{\mathbf{y}} = \frac{1}{\omega_1} \hat{\mathbf{g}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}) \quad (13)$$

$$\dot{\mathbf{z}} = \frac{1}{\omega_1 \omega_2} \hat{\mathbf{h}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}) \quad (14)$$

Table 2
Test accuracy of SVR model with different train to test data ratio.

| Serial Number | Train:test | Test accuracy |
|---------------|------------|---------------|
| 1 | 90:10 | 99.90% |
| 2 | 80:20 | 99.50% |
| 3 | 70:30 | 99.50% |
| 4 | 65:35 | 99.40% |
| 5 | 60:40 | 95.25% |

Table 3
Illustrating significant regulations of different metabolites, transcription factors and genes in mammalian cancer cells compared to that in normal ones.

| Network type | Name of the molecules | Significant observation in cancer cells compared to normal ones | References for validation | |
|-----------------|--|---|---------------------------|---------------|
| Metabolic | ATP | Over production | [46–49] | |
| | Ribose-5P | Over production | [51–53] | |
| | Lactate | Over production | [109,25,105,107] | |
| | GA3P | Over production | [46–49] | |
| | Pyruvate | Over production | [46–49] | |
| | Glucose | Over production | [46,107,105] | |
| | consumption/ G6P production | | | |
| | Glucose production | Less production | [46] | |
| | PEP | Less production | [46–49] | |
| | Fructose 6P | Less production | [96] | |
| | NADH | Less production | [50,103] | |
| | Signaling | HIF-1 α | Over expressed | [28,54] |
| | | P13K | Over expressed | [31–33,55] |
| | | AKT | Over expressed | [31–33,56,57] |
| mTOR | | Over expressed | [31–33,55] | |
| MYC | | Over expressed | [34,35,58,59,65] | |
| ERK | | Over expressed | [38–40] | |
| STAT3 | | Over expressed | [41,42,60–62] | |
| NF- κ B | | Over expressed | [39,43,44,63] | |
| p53 | | Under expressed | [36] | |
| PHD | | Under expressed | [29,30] | |
| Gene regulatory | | HK | Over expressed | [27,45,106] |
| | | Glut 1 | Over expressed | [27,45,106] |
| | | LDH | Over expressed | [109,25,105] |
| | | PFK 1 | Over expressed | [46,104] |
| | PFK 2 | Over expressed | [46] | |
| | Glyceraldehyde-3-phosphate dehydrogenase | Over expressed | [46] | |
| | Pyruvate kinase | Switching alternatively from low to high and vice versa | [47–49,104,64] | |
| | Glucose-6-phosphate dehydrogenase | Over expressed | [51–53] | |
| | Phospho-gluco dehydrogenase | Over expressed | [51–53] | |
| | Ribose 5P isomerase | Over expressed | [51–53] | |
| | Transketolase | Over expressed | [51–53] | |
| | Transaldolase | Over expressed | [51–53] | |

Here we can determine the value of \mathbf{u} , in terms of \mathbf{x} , \mathbf{y} and \mathbf{z} , considering the equilibrium state of the system.

2.2. Formulation of state equations of integrated biochemical pathways

Here we have shown how state equations for gene regulatory, signaling and metabolic networks are formed. Let a gene x_i be

expressed by s' ($1 \leq s' < s$) transcription factors y_j ($j' = 1, \dots, s'$) included in a signaling pathway. Thus, \dot{x}_i can be written as

$$\dot{x}_i = \begin{cases} e_i + b_i - d_i, & \text{if } \prod_{j'=1}^{s'} y_{j'} > 0 \\ b_i - d_i, & \text{if } \prod_{j'=1}^{s'} y_{j'} = 0 \end{cases} \quad (15)$$

The terms e_i and b_i represent the expression and basal production (of mRNA) rates [76] of x_i respectively, whereas d_i denotes the decay rate. The expression rate e_i of a particular gene x_i depends on the binding of many required transcription factors. Absence of any of those transcription factors, the expression of that particular gene is driven by its basal production and decay rate only. Now, e_i can be defined as

$$e_i = \prod_{j'=1}^{s'} \mathcal{K}_{ij'}^{(g)} y_{j'} \quad (16)$$

where $\mathcal{K}_{ij'}^{(g)}$ is the binding rate constant for j'^{th} transcription factor required to express the gene x_i .

Let us consider m' ($1 \leq m' < m$) metabolites z_k and c' ($1 \leq c' < c$) external inputs u_l acting as the cofactors [77] to activate an i^{th} gene and $\prod_{j'=1}^{s'} y_{j'} > 0$. As a result, the expression rate e_i of the gene x_i is enhanced by a multiplication factor. Thus, \dot{x}_i can be rewritten as

$$\dot{x}_i = \prod_{k=1}^{m'} (1 + \mathcal{F}_{ik}^{(g)} z_k) \prod_{l=1}^{c'} (1 + \mathcal{F}_{il}^{(g)} u_l) \prod_{j'=1}^{s'} \mathcal{K}_{ij'}^{(g)} y_{j'} + b_i - d_i \quad (17)$$

Here $\mathcal{F}_{ik}^{(g)}$ and $\mathcal{F}_{il}^{(g)}$ represent the constants corresponding to z_k and u_l to take care of the relative strength of binding with i^{th} gene [13,78,14]. We have considered binding constants to be in $[0, 1)$. Zero binding constant signifies no binding corresponding to a molecule. Whereas, higher value of a binding constant indicates stronger binding corresponding to the molecule. On the other hand, if m' metabolites z_k , s'' ($1 \leq s'' < s$) signaling molecules/transcription factors $y_{j''}$ ($j'' \neq j', j'' = 1, \dots, s''$) and c' external inputs u_l slow down the activation of the i^{th} gene and $\prod_{j'=1}^{s'} y_{j'} > 0$, the expression rate e_i is decreased by a fraction. Thus, Eq. (17) can be modified as

$$\dot{x}_i = \frac{\prod_{j'=1}^{s'} \mathcal{K}_{ij'}^{(g)} y_{j'} + b_i}{\prod_{k=1}^{m'} (1 + \mathcal{F}_{ik}^{(g)} z_k) \prod_{l=1}^{c'} (1 + \mathcal{F}_{il}^{(g)} u_l) \prod_{j''=1}^{s''} (1 + \mathcal{F}_{ij''}^{(g)} y_{j''})} - d_i \quad (18)$$

$(0 < \mathcal{F}_{ik}^{(g)} z_k, \mathcal{F}_{il}^{(g)} u_l, \mathcal{F}_{ij''}^{(g)} y_{j''} < 1)$

The term $\mathcal{F}_{ij''}^{(g)}$ represent the binding constant corresponding to $y_{j''}$ with i^{th} gene.

A signal transduction pathway can be defined by its interaction matrix $\mathbf{N}^{(s)}$ which contains the information about interactions among the signaling molecules including transcription factors. If there are q interactions with initial rate vector $\mathbf{r}^{(s)}$ (in timescale τ_1) of dimension q involving s signaling molecules including transcription factors, the order of $\mathbf{N}^{(s)}$ becomes $s \times q$. It may be mentioned here that $\mathbf{r}^{(s)}$ is analogous to a flux vector in a metabolic pathway. The $(j, \sigma)^{\text{th}}$ element of $\mathbf{N}^{(s)}$ becomes -1 if j^{th} signaling molecule participates in σ^{th} interaction. If j^{th} signaling molecule

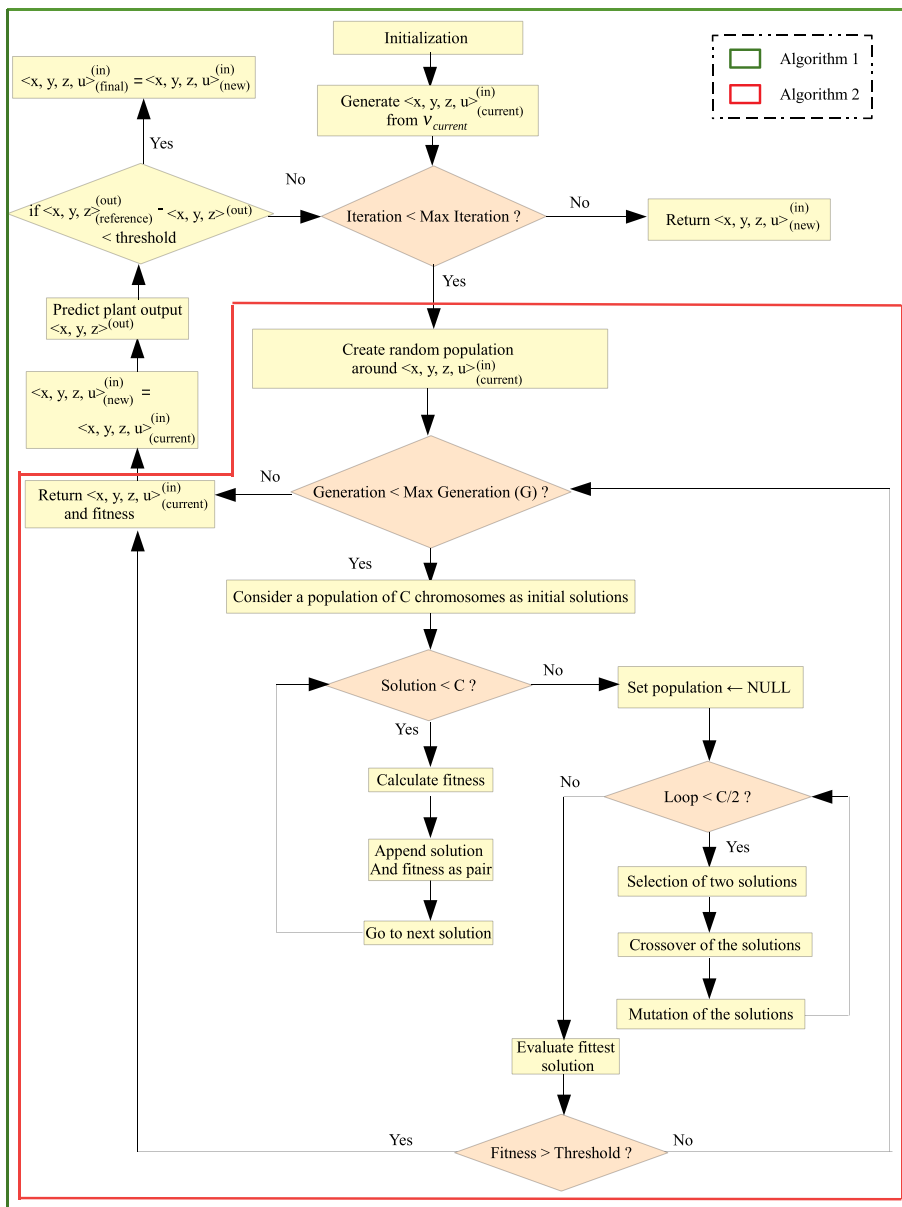


Fig. 5. Illustrating the flowchart showing the functionalities of two Algorithms S1 and S2 for designing GA-based controller.

is produced/activated by σ^{th} interaction, the $(j, \sigma)^{th}$ element of $\mathbf{N}^{(s)}$ becomes +1.

Let a signaling molecule y_j be activated by certain signaling molecules $y_{j'}$ ($j' \neq j$) in a signaling interaction with initial rate $r_{\sigma}^{(s)}$, σ^{th} ($\sigma = 1, \dots, q$) element of signaling interaction rate vector $\mathbf{r}^{(s)}$. Consequently, the initial interaction rate $r_{\sigma}^{(s)}$ depends on the binding of such signaling molecules activating y_j . If some metabolites z_k , signaling molecules $y_{j''}$ ($j'' \neq j'$) and external inputs u_l slow down the activation of y_j , a certain fraction can be multiplied with the initial interaction rate. Thus, we can write using modified mass action law [78] as

$$r_{\sigma}^{(s)} = \frac{\prod_{j'=1}^{s'} \mathcal{K}_{jj'}^{(s)} y_{j' \neq j}}{\prod_{k=1}^{m'} (1 + \mathcal{F}_{jk}^{(s)} z_k) \prod_{l=1}^{c'} (1 + \mathcal{F}_{jl}^{(s)} u_l) \prod_{j''=1}^{s''} (1 + \mathcal{F}_{jj''}^{(s)} y_{j''})}$$

where $\sigma = \sigma(j)$ is the index for the reaction through which y_j is activated. The term $\mathcal{K}_{jj'}^{(s)}$ represents the strength of binding of $y_{j'}$ with y_j , whereas $\mathcal{F}_{jk}^{(s)}$, $\mathcal{F}_{jl}^{(s)}$ and $\mathcal{F}_{jj''}^{(s)}$ represent binding constants corresponding to z_k , u_l and $y_{j''}$ for the interaction activating y_j . Again, if some metabolites z_k and external inputs u_l accelerate the activation of y_j , the initial interaction rate is increased by a multiplication factor. Thus, Eq. (19) becomes

$$r_{\sigma}^{(s)} = \prod_{k=1}^{m'} (1 + \mathcal{F}_{jk}^{(s)} z_k) \prod_{l=1}^{c'} (1 + \mathcal{F}_{jl}^{(s)} u_l) \prod_{j'=1}^{s'} \mathcal{K}_{jj'}^{(s)} y_{j' \neq j} \quad (20)$$

Thus, based on Eq. (13), the differential equation representing the dynamics of the signaling network can be rewritten as

$$\dot{\mathbf{y}} = \frac{1}{\omega_1} (\mathbf{N}^{(s)} \mathbf{r}^{(s)}) \quad (21)$$

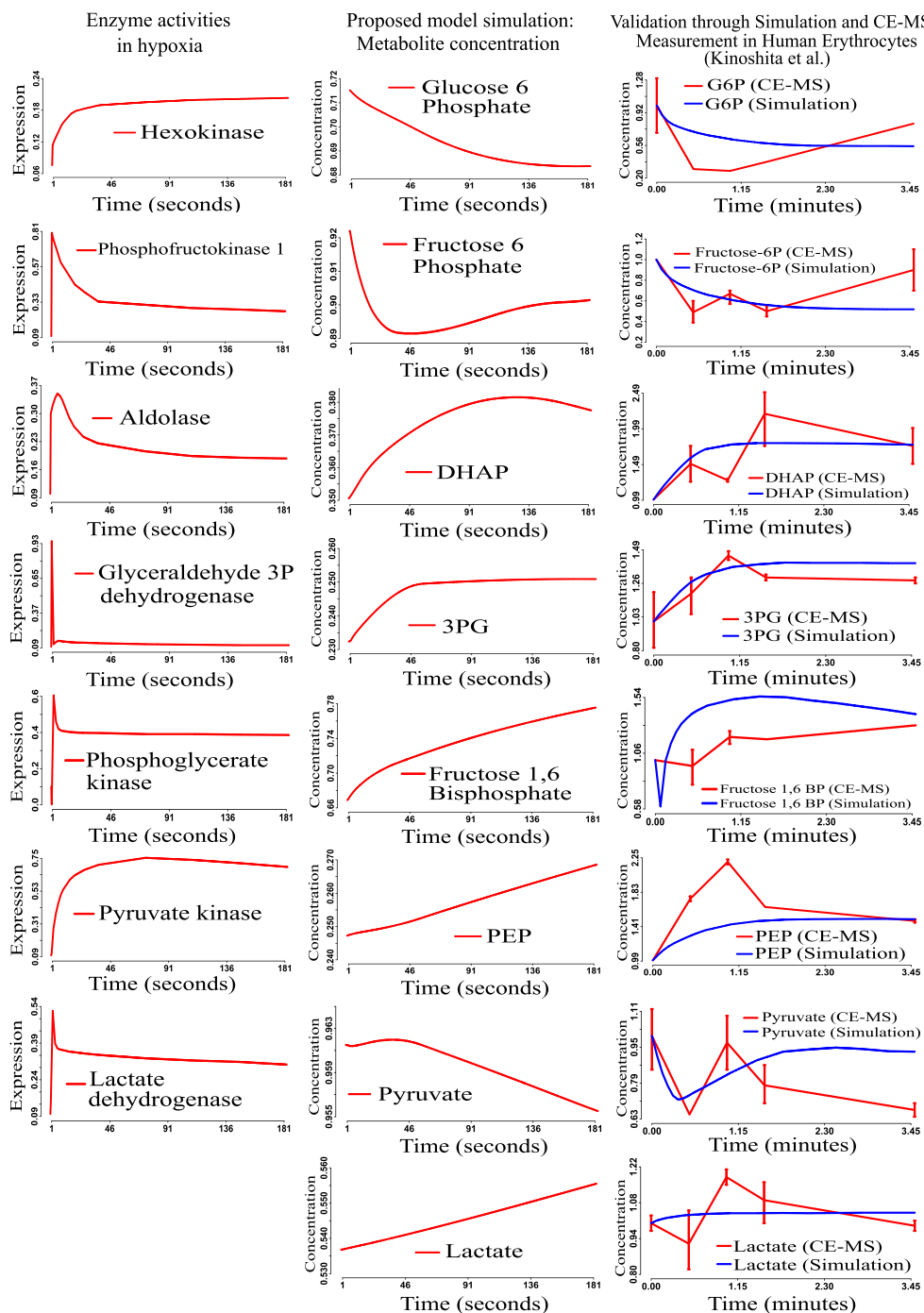


Fig. 6. Validation of the proposed model during hypoxia through CE-MS measurement and simulation in human erythrocytes (Kinoshita et al. [71]).

Similar to the signal transduction pathway, a specific stoichiometry is also associated with a metabolic pathway. Thus, we can consider a stoichiometric matrix $\mathbf{N}^{(m)}$ of order $m \times n$ for the metabolic network under consideration, where m is number of metabolites participating in n metabolic reactions in the network with metabolic flux vector $\mathbf{r}^{(m)}$ (in timescale τ_2) of dimension n . The $(k, \rho)^{th}$ element of $\mathbf{N}^{(m)}$ becomes +1 if k^{th} metabolite is produced through ρ^{th} reaction. On the other hand, if k^{th} metabolite is consumed in ρ^{th} reaction, the $(k, \rho)^{th}$ element of $\mathbf{N}^{(m)}$ becomes -1.

Let us consider a metabolite z_k be produced by m' ($1 \leq m' < m$) substrates $z_{k'}$ ($k' \neq k, k' = 1, \dots, m'$) in a metabolic reaction with initial rate $r_{\rho}^{(m)}$, ρ^{th} ($\rho = 1, \dots, n$) element of $\mathbf{r}^{(m)}$. This reaction is catalyzed by an enzyme/protein \mathcal{E}_{ρ} (representing the expression level) produced from an expressed gene. Furthermore, let us assume that m'' ($1 \leq m'' < m$) metabolites $z_{k''}$ ($k'' \neq k', k'' = 1, \dots, m''$) and c' external inputs u_i accelerate (noncompetitively or allosterically) the production of z_k . Thus, the initial reaction rate $r_{\rho}^{(m)}$ is increased by a multiplication factor. Then, we can write using modified Michaelis Menten kinetic equation [13,78,14] as

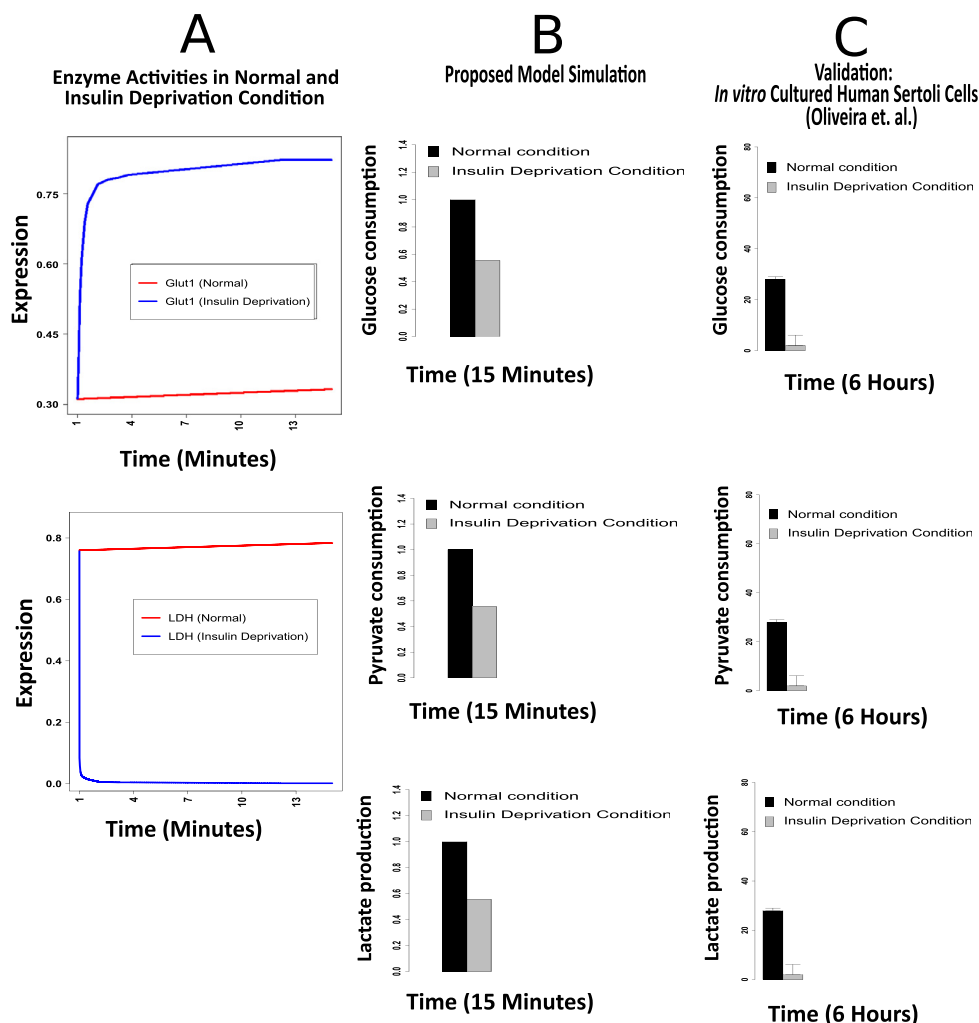


Fig. 7. Validation of the proposed model during insulin deprivation through *in vitro* observations from cultured human sertoli cells (Oliveira et al. [72]).

$$r_{\rho}^{(m)} = \frac{\prod_{k'=1}^{m'} (1 + \mathcal{F}_{kk'}^{(m)} Z_{k'}) \prod_{l=1}^c (1 + \mathcal{F}_{kl}^{(m)} u_l) \mathcal{K}_{\rho}^{(m)} \mathcal{E}_{\rho} \prod_{k'=1}^{m'} Z_{k'}}{\mathcal{K}_{\rho}^{(m)} \mathcal{E}_{\rho} + \left(\prod_{k'=1}^{m'} Z_{k'} \right)} \quad (22)$$

where $\rho = \rho(k)$ (could be one-to-many mapping) is the index for the reaction through which Z_k is produced. The terms $\mathcal{K}_{\rho}^{(m)}$ and $\mathcal{K}_{\rho}^{(m)}$ represent the kinetic rate constant and Michaelis Menten constant respectively. Besides, the terms $\mathcal{F}_{kk'}^{(m)}$ and $\mathcal{F}_{kl}^{(m)}$, for a k^{th} metabolite, represent feedback constants corresponding to $Z_{k'}$ and u_l . Again, if m' metabolites $Z_{k'}$ and c' external inputs u_l slow down (noncompetitively or allosterically) the production of Z_k , the initial reaction rate $r_{\rho}^{(m)}$ is decreased by a fraction. Thus, we can modify Eq. (22) as

$$r_{\rho}^{(m)} = \frac{\mathcal{K}_{\rho}^{(m)} \mathcal{E}_{\rho} \prod_{k'=1}^{m'} Z_{k'}}{\left(\mathcal{K}_{\rho}^{(m)} \mathcal{E}_{\rho} + \left(\prod_{k'=1}^{m'} Z_{k'} \right) \right) \prod_{k'=1}^{m'} (1 + \mathcal{F}_{kk'}^{(m)} Z_{k'}) \prod_{l=1}^c (1 + \mathcal{F}_{kl}^{(m)} u_l)} \quad (23)$$

Now, based on Eq. 14, the differential equation representing the dynamics of the metabolic network can be rewritten as

$$\dot{\mathbf{z}} = \frac{1}{\omega_1 \omega_2} (\mathbf{N}^{(m)} \mathbf{r}^{(m)}) \quad (24)$$

2.3. Design of a MIMO plant

The initial concentrations/expressions of metabolites, signaling molecules and genes have been considered in (0, 1) randomly (Supplementary Table S9). Moreover, the kinetic parameter values have also been initialized randomly within the same interval (Supplementary Tables S12–S15). Based on these initial values, we have simulated the proposed three timescale state space model of an integrated biochemical pathway by solving ODEs in Eqs. 17 (or 18), 21 and 24 with proper timescale selection, and solved for \mathbf{u} based on \mathbf{x} , \mathbf{y} and \mathbf{z} values under equilibrium condition. If the model fails to mimic the known behavioral pattern of the integrated biochemical pathway under consideration, we have altered slightly the initial values of kinetic parameters in (0, 1) with the help of previous knowledgebase by trial and error until satisfactory known behavior is captured. Once the proposed model has mimicked the known nonlinear dynamics of the integrated biochemical pathway (as described in SubSection 3.1), we have considered it as normal MIMO plant with the selected values of initial concentrations/expressions and kinetic parameters. This MIMO plant mimics the normal behavior of the concerned integrated pathway.

Results of the three timescale model

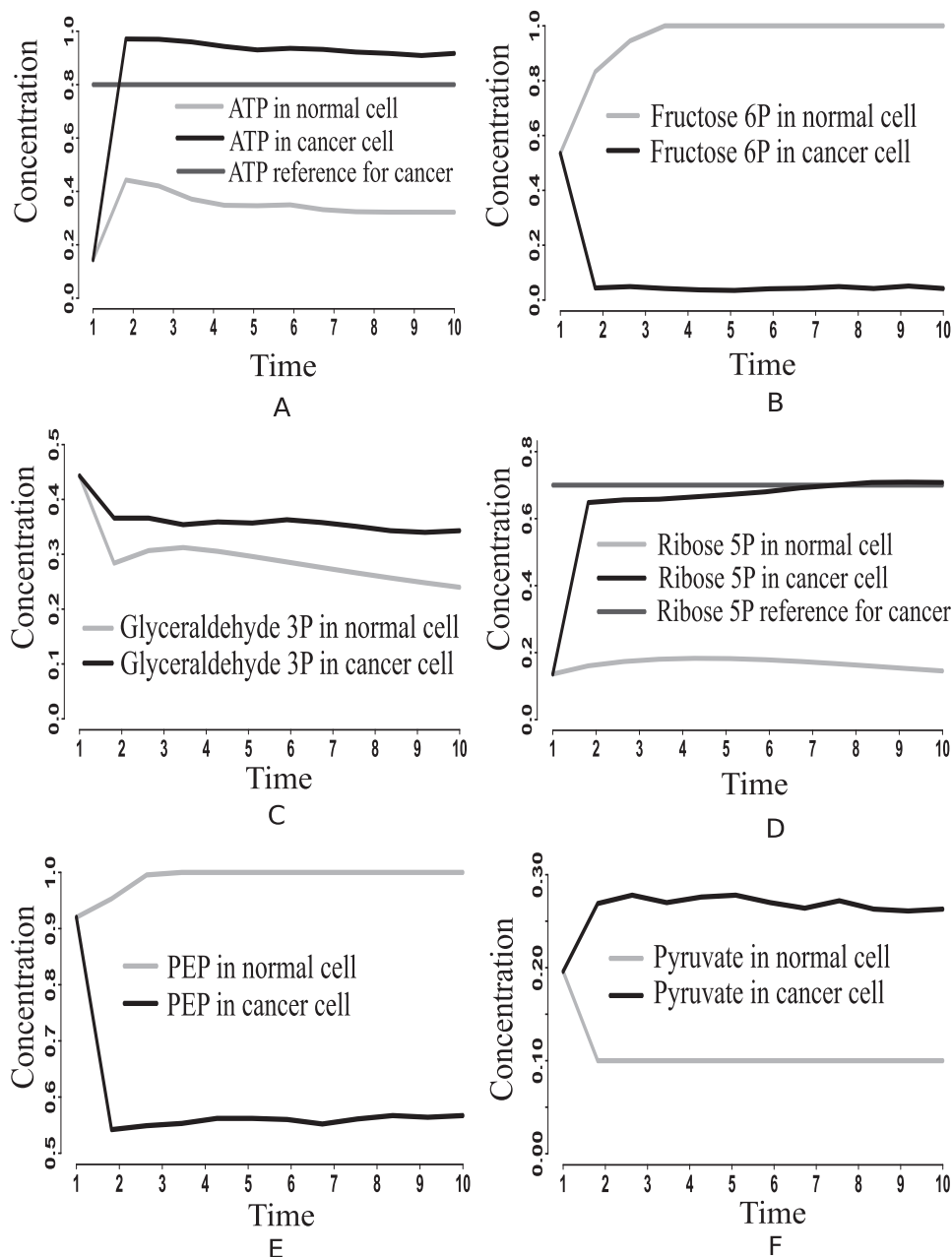


Fig. 8. Alterations of (A) ATP, (B) fructose 6P, (C) glyceraldehyde-3P (GA3P), (D) ribose 5P, (E) PEP, and (F) Pyruvate in the integrated biochemical pathway related to carbon metabolism in mammalian cancer cells compared to that in normal ones.

In this context, we have checked whether trial and error based parameter estimation resembles other previous method. For this purpose, we have compared the values of 35 kinetic parameters considered here with those estimated by a method based on hybrid extended Kalman filtering [70]. Fig. 3 depicts that the parameter values in (0, 1) considered in this article are quite similar with those estimated by hybrid extended Kalman filtering based method [70]. To estimate these 35 kinetic parameter values in (0, 1) using hybrid extended Kalman filtering, time dependent true values in (0, 1) of some observed states need to be provided. Consequently, we have used capillary electrophoresis mass spectrom-

etry (CE-MS) measured concentration values (normalized in (0, 1)) of eight metabolites in hypoxia-induced metabolic alterations in human erythrocytes [71] as observed states. Here we have externally applied appropriate signal in (0, 1) for corresponding enzyme activities related to hypoxia in human erythrocytes (as described in SubSection 3.1). During the estimation of these 35 parameter values using hybrid extended Kalman filtering, we have noticed that estimated states have been overlapped with observed states as depicted in Fig. 4. However, we could not verify the other parameter values using this way due to lack of required observed states.

Results of the three timescale model

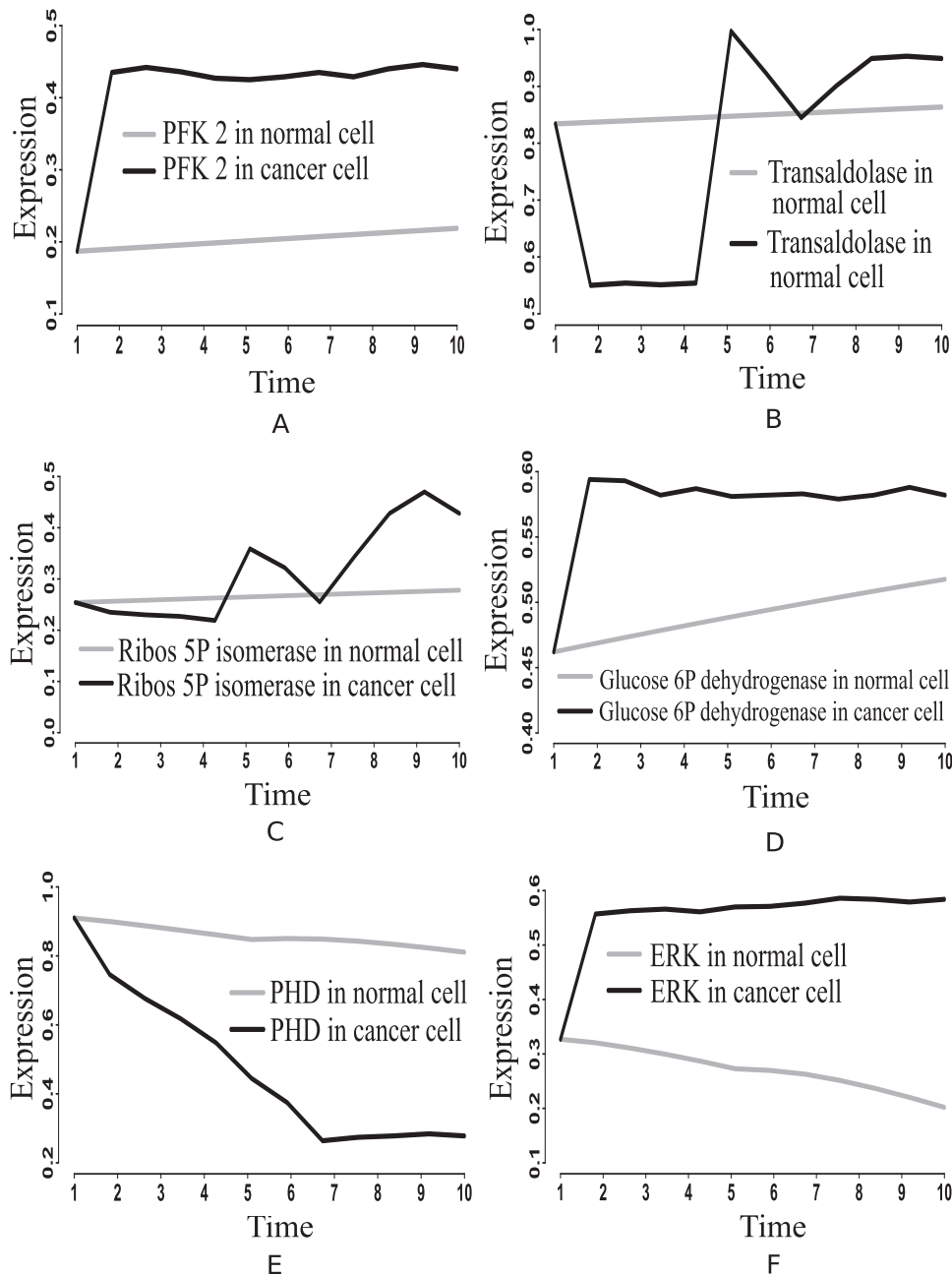


Fig. 9. Alterations of (A) PFK2, (B) transaldolase, (C) ribose 5P isomerase, (D) glucose-6-phosphate dehydrogenase, (E) PHD, and (F) ERK in the integrated biochemical pathway related to carbon metabolism in mammalian cancer cells compared to that in normal ones.

Now, we perturb the normal MIMO plant by knocking out genes (in this study, genes producing the enzymes pyruvate dehydrogenase and pyruvate carboxylase in carbon metabolism) to capture the nonlinear dynamics associated with a particular biological phenomenon (in this study, “Warburg effect” [24,25]). As a result, the normal plant becomes the desired perturbed MIMO plant under consideration.

2.4. Generation of training dataset

So far we have developed the desired perturbed MIMO plant with consideration of appropriate three timescales. It does not

reflect the mutated regulations of a mammalian cancer cell. However, it contains the perturbations associated with “Warburg effect”. In order to capture the altered dynamics of a cancer cell, we need a model predictive controller to be applied on an approximate model corresponding to the MIMO plant under consideration [79]. We are going to develop such an approximate model using support vector regression. Thus, the past input and output data simulated from the MIMO plant under investigation are collected to serve as a training dataset for the SVR-based approximation. Here we have found that consideration of τ (=19) number of past inputs and corresponding outputs along with current input (at time θ), from the simulated data as an input sample of SVR, and current out-

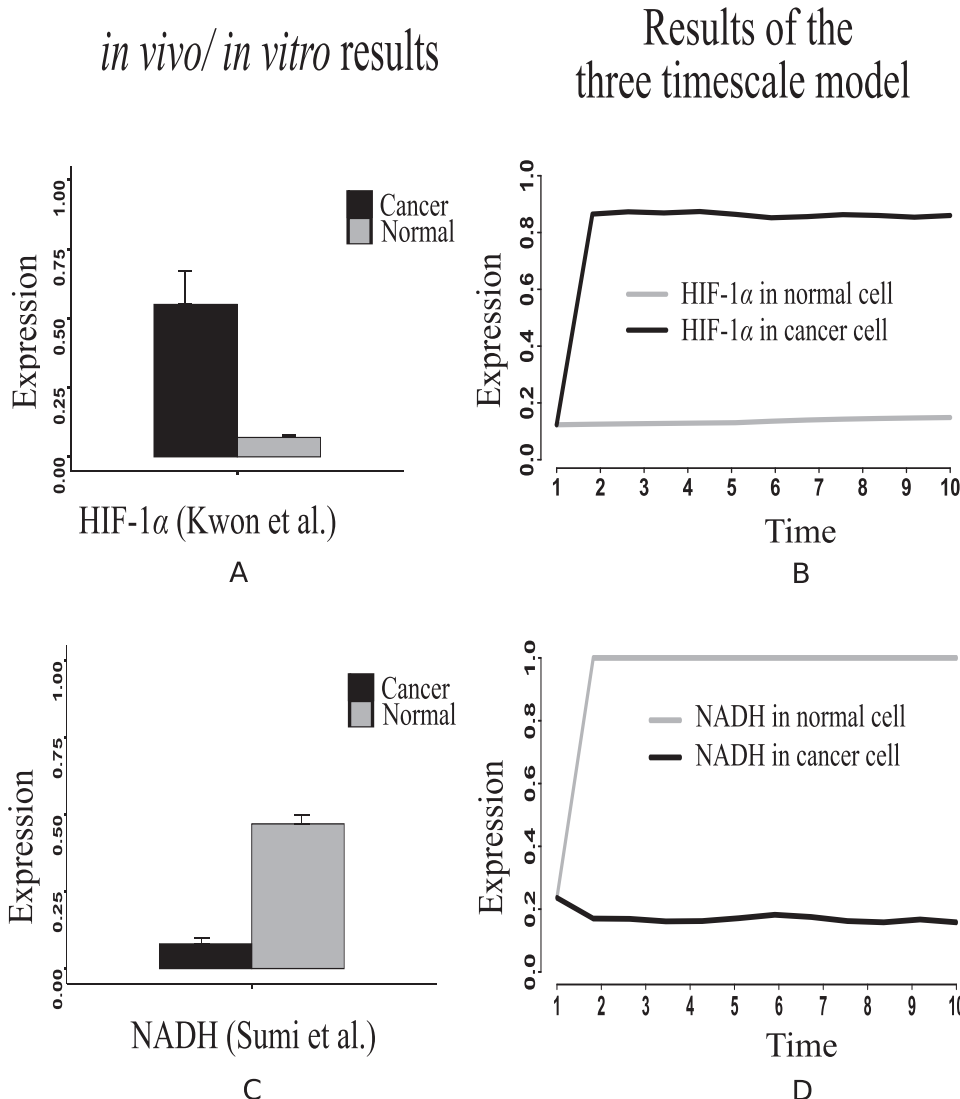


Fig. 10. It depicts: (A) Higher expression level of HIF-1 α measured in human hepatocellular carcinoma (HCC) specimens compared to the non-cancerous tissue specimens [54], (B) Similar behavior of HIF-1 α expression shown by our simulation results, (C) Lower concentration of NADH in cancer cells compared to normal ones as per the *in vivo/ in vitro* experiment performed by Sumi et al. [103], (D) Similar behavior of NADH depicted by our computational results.

put (at time θ) as corresponding supervised signal of the regression model, has generated maximum test accuracy (99.4%) with mean squared error (MSE) threshold of 0.05. Now, we are going to discuss how such training data have been generated.

Let us consider the MIMO plant under consideration takes an instance $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{(\theta-\tau)}^{(in)}$ as input at time $(\theta - \tau)$ and generate corresponding output instance $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{(\theta-\tau)}^{(out)}$. Using the output instance $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{(\theta-\tau)}^{(out)}$ at time $(\theta - \tau)$, a new instance of \mathbf{u} has been determined, followed by formation of the new input instance $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{(\theta-(\tau-1))}^{(in)}$ for the MIMO plant at time $(\theta - (\tau - 1))$. Likewise, we have generated 19 (*i.e.*, $\tau = 19$) past inputs and outputs to form a current input instance $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{\theta}^{(in)}$ and corresponding output instance $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{\theta}^{(out)}$ at time θ . In general, the total number of elements in $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle^{(in)}$ is $(p + s + m + c)$, while that in $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle^{(out)}$ is $(p + s + m)$. Here we have considered $p = 37$, $s = 29$, $m = 41$ and $c = 27$ for carbon metabolic pathway obtained from KEGG.

The output of the MIMO plant depends not only on the current input but also on its past inputs and outputs. We have, therefore, considered an input vector \mathbf{v} to be applied to SVR model as.

$$(\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{(\theta-19)}^{(in)}, \langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{(\theta-19)}^{(out)}, \langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{(\theta-18)}^{(in)}, \langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{(\theta-18)}^{(out)}, \dots, \langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{\theta}^{(in)}).$$

Thus, the number of inputs to SVR model becomes $(19 \times (2 \times (p + s + m) + c) + (p + s + m + c)) = 4713$, which is the dimension (d) of \mathbf{v} . Now, the current output instance $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{\theta}^{(out)}$ has been attached to \mathbf{v} , as its supervised signal, to form one sample to be included in a dataset \mathcal{D} . If there are ϕ such samples in \mathcal{D} , let us call it as \mathcal{D}_{ϕ} . Here we have considered $\phi = 60000$. We have selected 65% data from \mathcal{D}_{ϕ} as training data and remaining 35% data as test data to design the SVR model to be discussed in the following section. During training of the SVR model, we aim to utilize a minimum number of a training pattern in order to improve generalization performance. To assess the best possible ratio between train and test data, we have observed the performance with different ratios. The accuracy of such cases is given in Table 2. Here, it

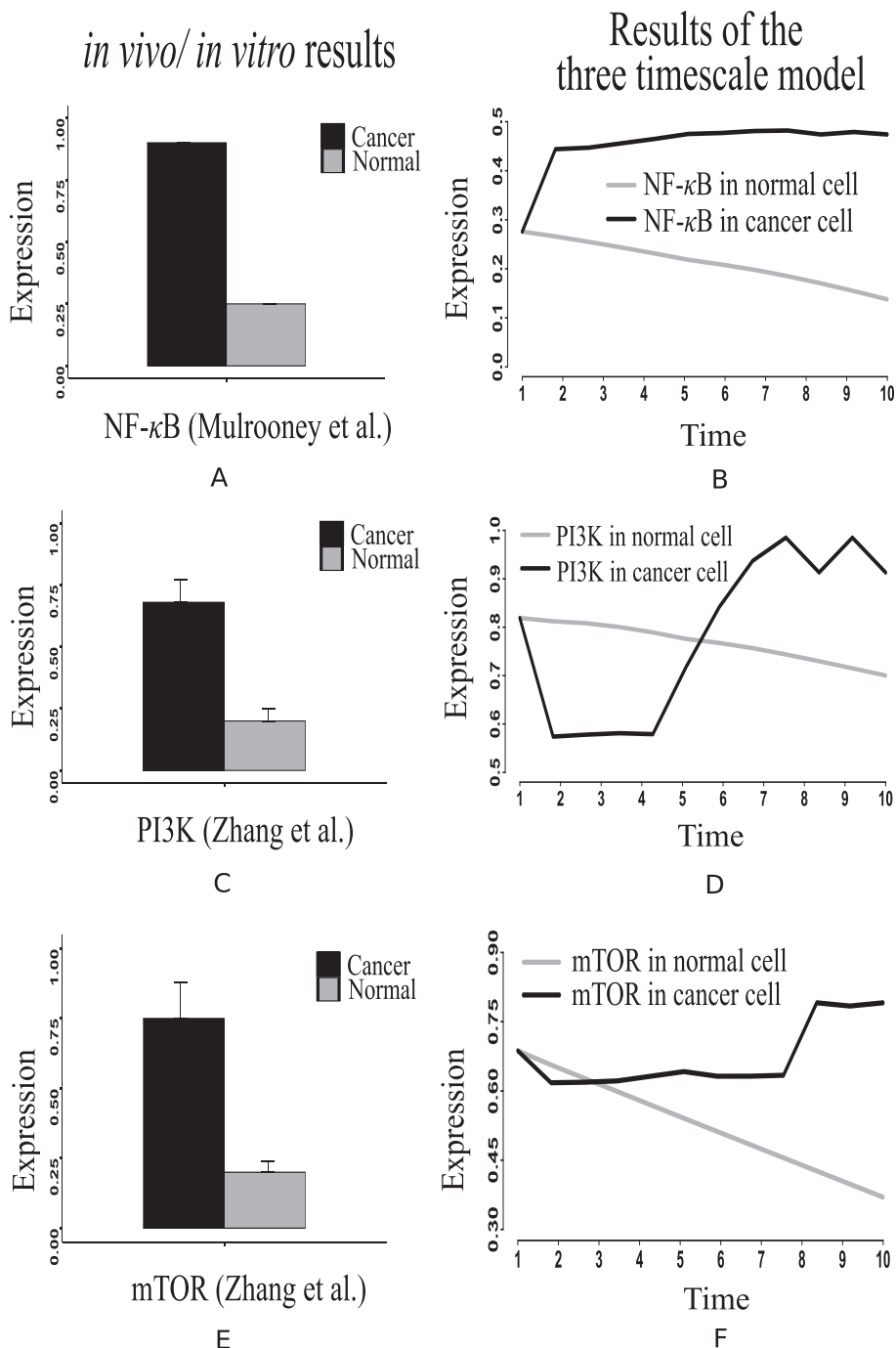


Fig. 11. (A) Mulrooney et al. [63] has depicted that relative expression of NF- κ B significantly increases in cancer (CT-2A astrocytoma) compared to normal ones. (B) Our computational results show similar behavior of NF- κ B in cancer cells compared to that in normal ones. (C) Higher expression of PI3K has been captured in the *in vivo/in vitro* experiment performed by Zhang et al. [55] in the case of colon cancer as compared to the normal tissue samples, (D) Similar difference in the expression levels of PI3K in cancer and normal cells have been captured by the proposed computational model, (E) mTOR expression is higher in colon cancer compared to normal one as per the *in vivo/in vitro* experiment performed by Zhang et al. [55], and (F) Similar altered behavior of mTOR can be found in our computational results.

can be concluded that the train to test ratio as 65:35 may be the best trade-off between the number of patterns used for training and the test accuracy. Thus we have considered train to test data ratio accordingly.

2.5. Design of SVR model

From control theoretic point of view, it is difficult to apply online optimization on actual nonlinear MIMO plant with small sampling periods [80,81]. Moreover, due to large number of vari-

ables involved in an optimization problem, and requirement of high sampling rate, it may not be possible to manage online optimization. However, it will be easier if we can provide a model that suitably approximates the nonlinear dynamics of actual MIMO plant. We also think that such implementation may be relevant for future study to determine drug dosages in real time. Even this kind of modelling will be effective in future when actual input-output mapping of the MIMO plant is unknown, whereas the inputs and corresponding outputs with time are known. The training dataset \mathcal{D}_ϕ helps in developing the approximate MIMO SVR

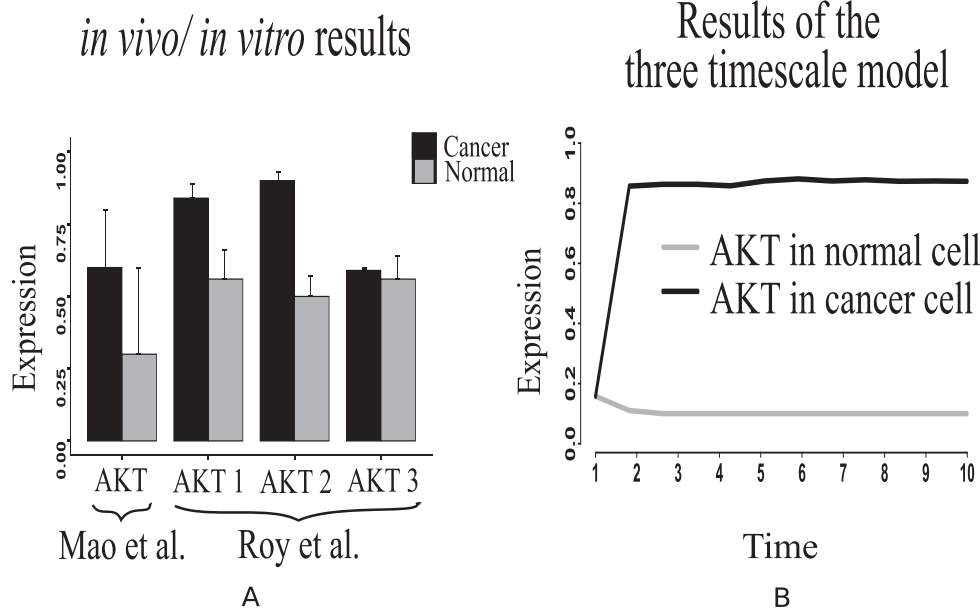


Fig. 12. (A) Immunohistochemical scores measured by Mao et al. [56] has depicted that AKT expression is significantly higher in pancreatic cancer (91 cases) than that in normal pancreas (51 cases). Besides, Roy et al. [57] has shown the differentially expressed AKT isoforms in normal and malignant oral tissues through immunohistochemical analysis of the human samples, (B) Similar altered behavior of AKT has been shown by the computational results generated from the proposed model.

model to mimic the nonlinear dynamics of the MIMO plant/system under consideration. In this context, it should be mentioned that according to some recent studies [82–84], SVR has better performance accuracy than an artificial neural network (ANN) (particularly, multi-layer perceptron (MLP)) and genetic programming (GP). In general, the computational complexity can be evaluated depending mainly on training time. SVR depends on solving a quadratic problem, where a vector and bias need to be calculated. Even during training, only support vectors are selected regarding the kernel. Besides, SVR considers that the number of support vectors is governed by very limited number of training samples. In addition, most of the kernels compute simple dot product. Even SVR may act better than ANN as per the consistency with physical behavior [85]. Moreover, it has been claimed that the SVR technique not only achieves high consistency but also shows a greater degree of generalization to unseen test data compared to some other methods [86]. Besides, a less number of parameters involved in the training phase results in less computation time than that of an ANN [87].

Now, we are going to discuss how a MIMO SVR model can be designed using support vector regression [88,89].

An approximate MIMO SVR model has been developed by combining multiple input single output (MISO) SVR models for each μ^{th} ($1 \leq \mu \leq (p + s + m)$) element of actual output $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{\theta}^{(out)}$. Let a^{th} input be \mathbf{v}_a ($1 \leq a \leq \phi$) of dimension d . For an a^{th} input, the MISO SVR model for μ^{th} element of $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{\theta}^{(out)}$ in a projected feature space F of dimension d' ($d' > d$) can be written as

$$\hat{\gamma}_{\mu} = \mathbf{w}_{\mu}^T \delta(\mathbf{v}_a) + \beta_{\mu} \quad (25)$$

Here $\hat{\gamma}_{\mu}$ represents μ^{th} element of predicted output $\langle \hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}} \rangle_{\theta}^{(out)}$. The term \mathbf{w}_{μ} represents a vector perpendicular to the hyperplane to be approximated for the regression problem under consideration in the projected feature space F corresponding to μ^{th} element of $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{\theta}^{(out)}$. In other words, \mathbf{w}_{μ} determines the orientation of the

hyperplane in F . On the other hand, $\delta(\cdot)$ is a mapping from d dimensional input space to d' dimensional feature space F . The bias term is represented by β_{μ} determining the position of the hyperplane in F corresponding to μ^{th} element of $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{\theta}^{(out)}$.

Based on Vapnik's ε -insensitive loss function, the convex optimization problem corresponding to Eq. (25) can be formulated as [90]

$$\min_{\mathbf{w}_{\mu}, \beta_{\mu}, \varphi_a, \varphi_a^*} \mathcal{P}_{\mu} = \frac{1}{2} \|\mathbf{w}_{\mu}\|^2 + \mathcal{C} \sum_{a=1}^{\phi} (\varphi_a + \varphi_a^*) \quad (26)$$

subject to:

$$\begin{aligned} \gamma_{\mu} - \mathbf{w}_{\mu}^T \delta(\mathbf{v}_a) - \beta_{\mu} &\leq \varepsilon_{\mu} + \varphi_a, \text{ for } a = 1, \dots, \phi \\ \mathbf{w}_{\mu}^T \delta(\mathbf{v}_a) + \beta_{\mu} - \gamma_{\mu} &\leq \varepsilon_{\mu} + \varphi_a^*, \text{ for } a = 1, \dots, \phi \\ \varphi_a, \varphi_a^* &\geq 0, \text{ for } a = 1, \dots, \phi \end{aligned}$$

Here, the upper limit of tolerable error is a very small positive real quantity ε_{μ} for μ^{th} element of $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{\theta}^{(out)}$, whereas φ_a and φ_a^* are slack variables. The term γ_{μ} represents the actual value of μ^{th} element of $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{\theta}^{(out)}$ generated by the MIMO plant taking $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{\theta a}^{(in)}$ (extracted from \mathbf{v}_a) as an a^{th} input. The trade-off between flatness of $\hat{\gamma}_{\mu}$ and tolerance level of deviations ($> \varepsilon_{\mu}$) is represented by the constant \mathcal{C} (> 0).

Based on the notion of Lagrange multipliers and Karush Kuhn Tucker (KKT) condition [91], the dual corresponding to Eq. (26), in the form of a quadratic programming problem with incorporation of kernel function, can be written as

$$\begin{aligned} \min_{\zeta_a, \zeta_a^*} \mathcal{Q}_{\mu} &= \frac{1}{2} \sum_{a=1}^{\phi} \sum_{b=1}^{\phi} (\zeta_a - \zeta_a^*) (\zeta_b - \zeta_b^*) \mathcal{K}(\mathbf{v}_a, \mathbf{v}_b) + \varepsilon_{\mu} \sum_{a=1}^{\phi} (\zeta_a + \zeta_a^*) \\ &\quad - \sum_{a=1}^{\phi} \gamma_{\mu} (\zeta_a - \zeta_a^*) \end{aligned} \quad (27)$$

subject to:

in vivo/ in vitro results

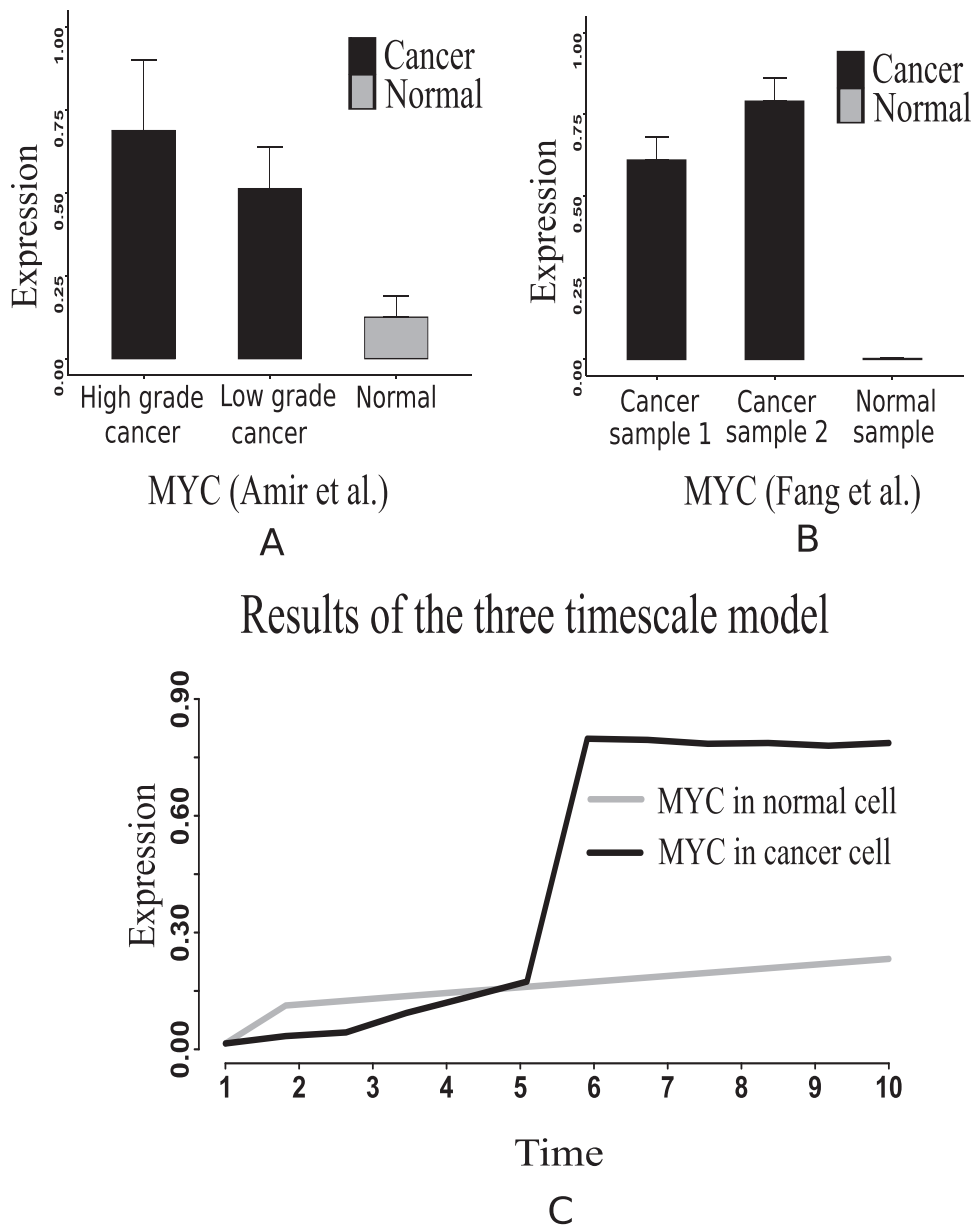


Fig. 13. (A) & (B) depict that the relative m-RNA level protein expression [58,59] of MYC is significantly higher in cancer cell than normal ones. (C) depicts a similar behavior of MYC generated by the proposed computational model.

$$0 \leq \zeta_a, \zeta_a^* \leq \mathcal{C},$$

$$\sum_{a=1}^{\phi} (\zeta_a - \zeta_a^*) = 0, \text{ for } a = 1, \dots, \phi$$

Here the term $\kappa(\mathbf{v}_a, \mathbf{v}_b)$ denotes the kernel function. Besides, $\zeta_a, \zeta_a^*, \zeta_b$ and ζ_b^* stand for Lagrange multipliers. The solution of Eq. (27) provides the optimum values of $\zeta_a, \zeta_a^*, \zeta_b$ and ζ_b^* . Moreover, $\mathbf{w}_\mu - \sum_{b=1}^{\phi} (\zeta_b - \zeta_b^*) \delta(\mathbf{v}_b) = 0$ must hold for optimality. Thus, for an a^{th} input, the support vector kernel expansion of Eq. (25) can be obtained from

$$\hat{\gamma}_\mu = \sum_b \alpha_b \kappa(\mathbf{v}_a, \mathbf{v}_b) + \beta_\mu \tag{28}$$

The summation has been carried out over all the support vectors \mathbf{v}_b . Here $\alpha_b = (\zeta_b - \zeta_b^*)$. An input vector \mathbf{v}_b corresponding to a non-zero α_b is called a support vector. The value of β_μ is determined when $|\gamma_\mu - \hat{\gamma}_\mu| = \epsilon_\mu$ and $0 \leq \zeta_b - \zeta_b^* \leq \mathcal{C}$ hold for each support vector \mathbf{v}_b . Thus, Eq. (28) provides the required μ^{th} element of predicted output $\langle \hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}} \rangle_0^{(out)}$, taking \mathbf{v}_a as input, corresponding to μ^{th} element of actual output $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_0^{(out)}$.

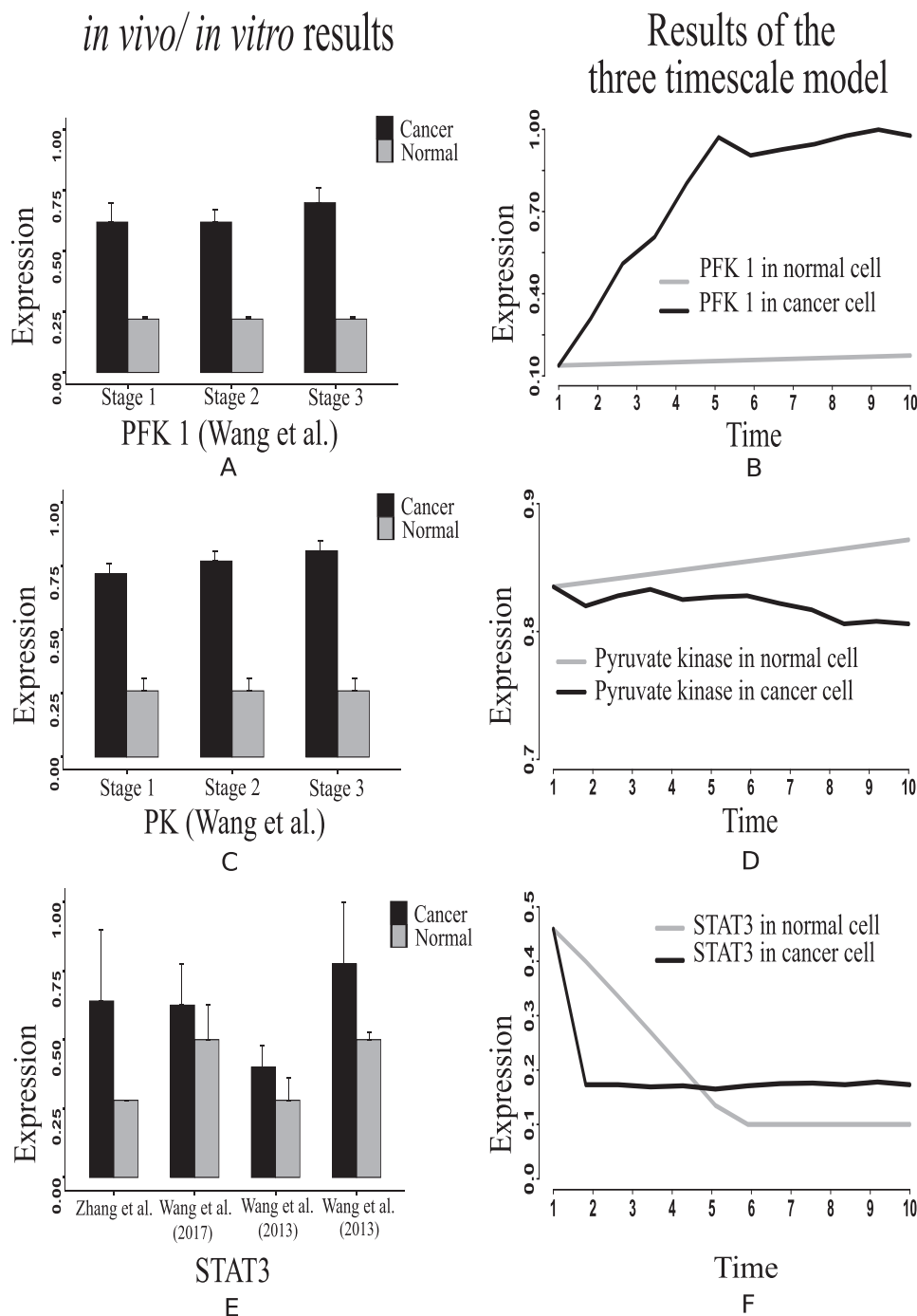


Fig. 14. Activities of (A) PFK1 [104] and (C) pyruvate kinase in breast cancer and paracancer tissues, expressed as units per gram of protein (U/gprot) [104]. Such altered behaviors of (B) PFK1 and (D) pyruvate kinase have been captured by the proposed computational model. The previous *in vivo/ in vitro* studies have illustrated that (E) STAT3 is over expressed in cancer cells, such as breast cancer [60], gastric cancer [61] and colon cancer [62] cells compared to normal ones. Similar observation regarding (F) STAT3 can be found in our computational results.

Here, we have experimentally chosen radial basis kernel function with kernel coefficient as 10, regularization parameter as 100 and $\epsilon_{\mu}=0.0025$ to achieve maximum test accuracy.

2.6. Design of a controller using a genetic algorithm (GA)

In this study, we have applied a GA [89] on the SVR model as a model predictive controller to control some outputs of the MIMO plant at specific values. There are many reasons behind choosing GA as an optimization technique. Firstly, a genetic algorithm (GA) [92,93] is a stochastic process. Secondly, it is a vigorous search

technique, which does not require any information about the structure of the function to be optimized. Such a situation is very common to address a biological system, particularly during mutation. Thirdly, GA is very efficient in handling highly complicated non-linear problems, such as an integrated biochemical pathway system. Besides, due to its inherent parallelism, GA can easily be implemented in a distributed environment. Such a characteristic is very much helpful in handling a large number of parameters/variables in a biological system. Fourthly, GA may be able to avoid being trapped in a locally optimal solution, unlike traditional methods. Besides, GA can efficiently be applied to different large

Results of the three timescale model

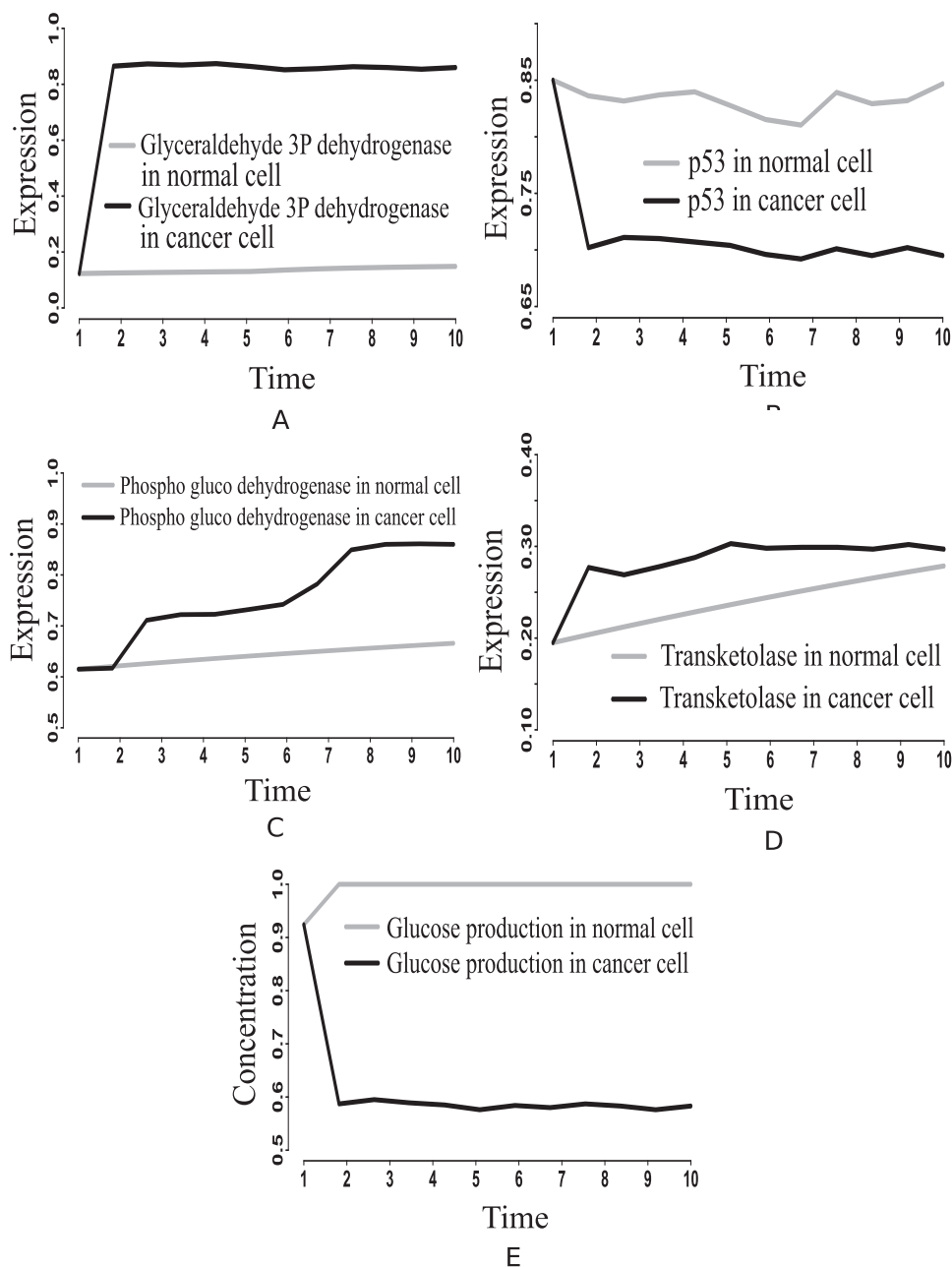


Fig. 15. Alterations of (A) glyceraldehyde-3-phosphate dehydrogenase, (B) p53, (C) phospho-gluco dehydrogenase, (D) transketolase, and (E) glucose production in the integrated biochemical pathway related to carbon metabolism in mammalian cancer cells compared to that in normal ones.

scale real-world problems with the requirement of multi-objective optimization, such as the particular problem addressed in this article.

Such a GA based control mechanism has been depicted in Fig. 2 (Step 6). Here the error between SVR generated predicted output and (user-specified) reference output drives the GA based controller to produce a controlled input based on some cost functions and constraints. The integrated biochemical network (*i.e.*, the MIMO plant) accepts the controlled input to compute actual output. The GA controller has been developed using two algorithms (Supplementary Algorithms S1 and S2). Supplementary Algorithm

S1 deals with the working principles of a GA controller. On the other hand, the Supplementary Algorithm S2 drives some outputs of the MIMO plant to target values using Supplementary Algorithm S1 iteratively. Now, we are going to discuss Supplementary Algorithms S1 and S2 in brief. Here, Fig. 5 depicts the flowchart indicating the role and sequence of two Algorithms S1 and S2 for designing GA-based controller.

- Supplementary Algorithm S1 shows how proposed GA controller works. Here the GA controller accepts $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{current}^{(in)}$ and $\mathbf{v}_{current}$ as input. At first, it generates C chromosomes (set

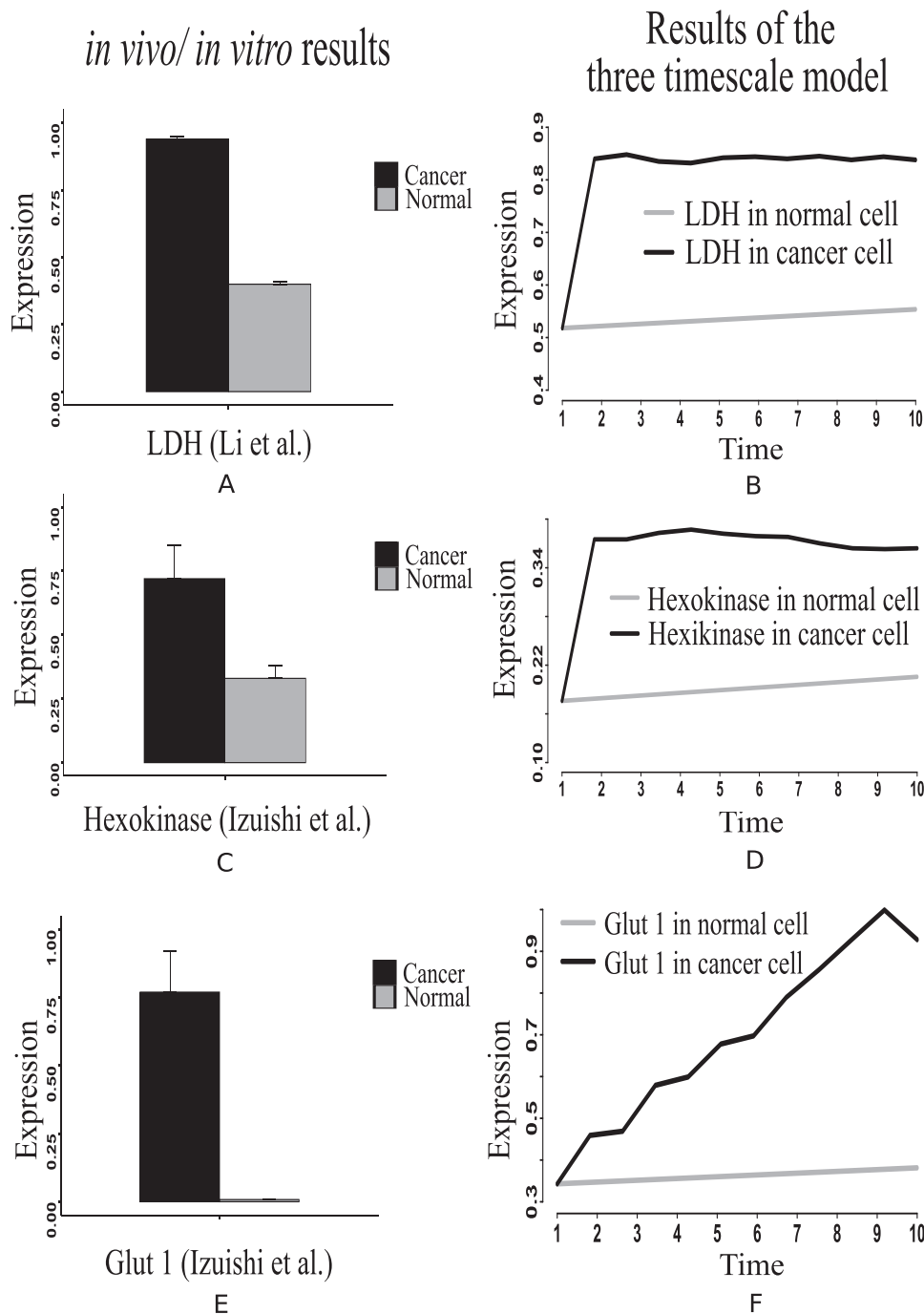


Fig. 16. Li et al. [105] has shown that (A) LDH has significantly been up regulated in Dppa4 overexpressed cancer cell, (C) Expression of hexokinase (HK) mRNA levels have been found higher in cancer [106] compared to normal liver tissue, (E) Higher glut1 mRNA levels has been found as 92-fold higher in Meta specimens [106] compared to normal liver tissue. Similar altered behaviors of (B) LDH, (D) HK and (F) glut1 have been found in our computational results.

of possible solutions) randomly, forming a population. However, it creates chromosomes around $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u})_{current}^{(in)}$ after getting an optimal solution. For each generation, the fitness value of each chromosome has been computed. Thereafter, the chromosomes are modified by selection, crossover and mutation. In this context, a set of chromosomes is selected according to fitness values greater than a weight value. Subsequently, the crossover is performed by swapping a part of chromosome with corresponding part of another chromosome according to a crossover index. In this context, it should be mentioned that as the size of the output vector is $107 (=p + s + m)$ as men-

tioned earlier), we have considered crossover index as 53. This index is nearly half of the size of the output vector. It is followed by mutation to modify chromosomes with mutation probability 0.7. Here, we have used uniform crossover [94,95], i.e., multi-point crossover, to reduce the mutation bias. These steps are carried out for G generations until the fittest chromosome, whose fitness value is greater than a predefined threshold value, has been evolved as an optimal solution $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u})_{new}^{(in)}$.

- **Supplementary Algorithm S2** depicts how some outputs of the MIMO plant are controlled to get desired responses with repetitive use of GA controller (**Supplementary Algorithm S1**). The

in vivo/ in vitro results

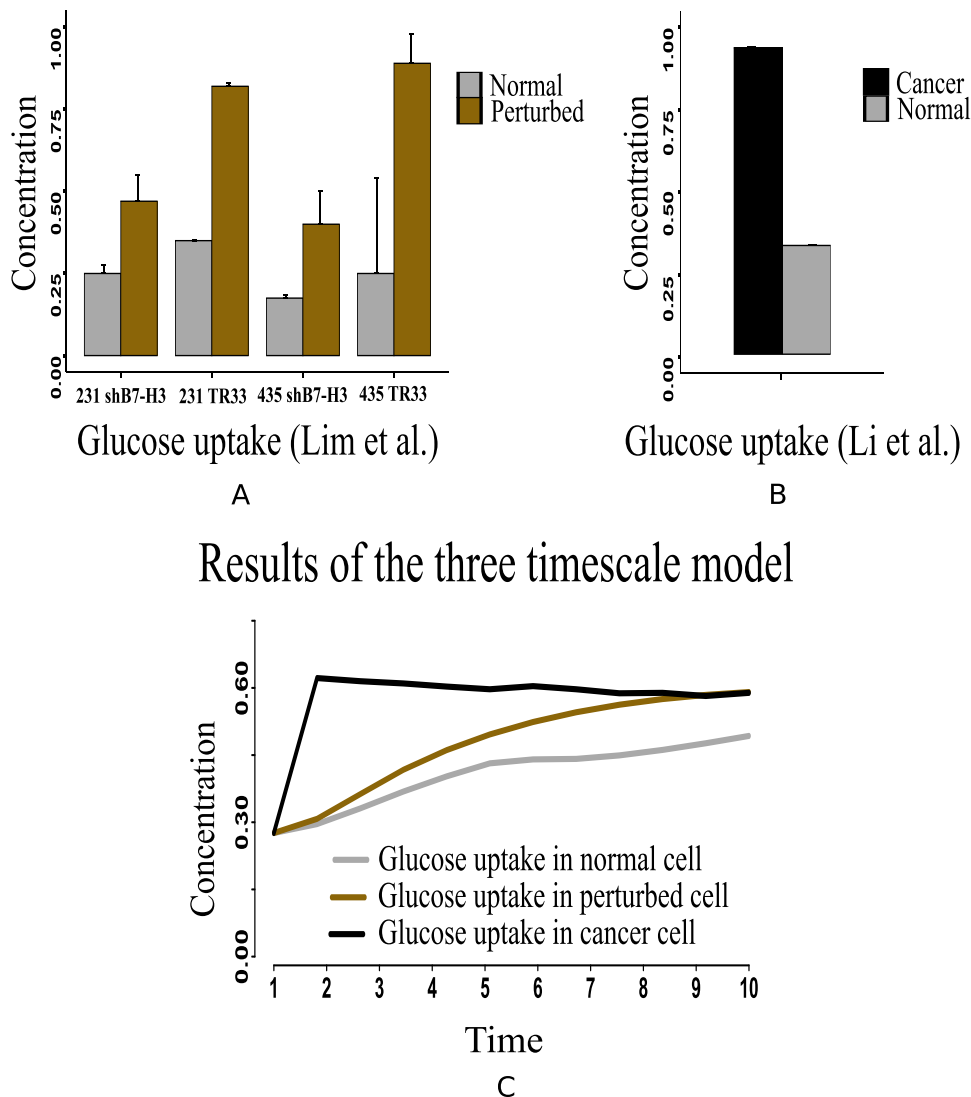


Fig. 17. (A) Higher glucose uptake has been measured by Lim et al. [107] in B7-H3 knockdown cells grown in normoxic or perturbed conditions (hypoxia) for 24 h, (B) Li et al. [105] has shown that glucose uptake increases significantly in Dppa4 overexpressed cancer cell, and (C) Our computational results depict similar behavior of glucose uptake in cancer cells compared to that in normal and perturbed ones.

execution of [Supplementary Algorithm S2](#) needs the trained SVR model, actual MIMO plant, initial SVR input $\mathbf{v}_{current}$ and reference output $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{reference}^{(out)}$. As per aforesaid discussion, the GA controller receives $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{current}^{(in)}$ (extracted from $\mathbf{v}_{current}$) and $\mathbf{v}_{current}$ itself to compute an optimal solution $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{new}^{(in)}$. However, during iteration > 1 of [Supplementary Algorithm S2](#), if the fitness value of an optimal solution (chromosome) becomes less than that of previous iteration, the GA controller tries again to find a better optimal solution using the same $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{current}^{(in)}$. After obtaining an optimal solution $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{new}^{(in)}$, the initial SVR input has been updated to \mathbf{v}_{new} . Subsequently, we have computed predicted output $\langle \hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}} \rangle^{(out)}$ and actual output $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle^{(out)}$ from the SVR model and MIMO plant respectively, using \mathbf{v}_{new} and $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{new}^{(in)}$. Here we have considered the reference output $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{reference}^{(out)}$ containing target values for some

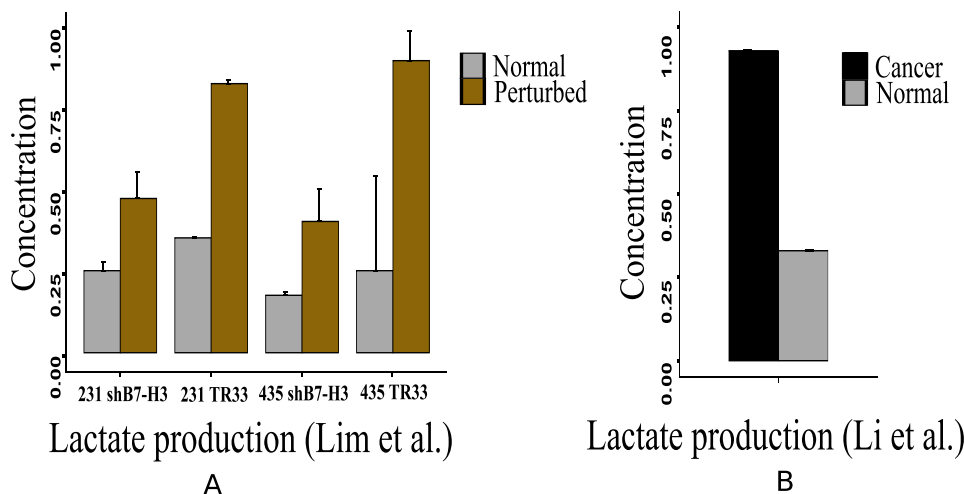
specific elements (for example, 0.95 for ATP and 0.7 for ribose 5P), whereas others remain at the same values as computed in $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle^{(out)}$. We have finally obtained the final solution $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u} \rangle_{final}^{(in)}$ when the L_1 -norm of the difference between $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle_{reference}^{(out)}$ and $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle^{(out)}$ becomes less than a predefined threshold value after several iterations.

Finally, it should be mentioned that the average value of G is 10 for convergence of Algorithm S1, whereas 30 iterations on average are needed for Algorithm S2 to converge.

3. Results and discussion

This section has been divided into three subsections. Firstly, we are going to discuss about MIMO plant validation through analysis

in vivo/ in vitro results



Results of the three timescale model

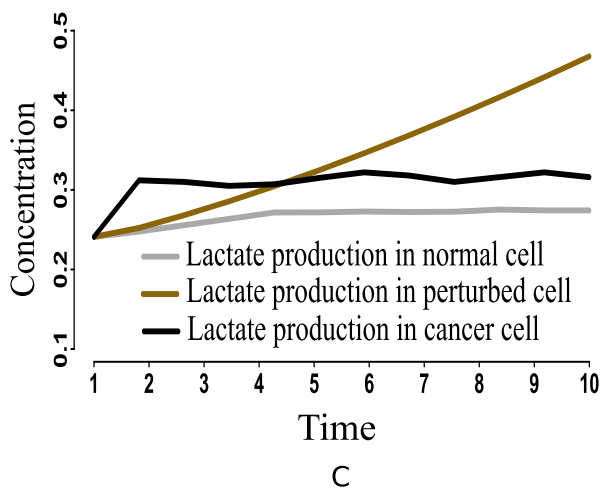


Fig. 18. (A) Higher Lactate production has been captured by the measurement of Lim et al. [107] in B7-H3 knockdown cells grown in normoxic or perturbed conditions (hypoxia) for 24 h. (B) Higher lactate production has been shown by Li et al. [105] in Dppa4 overexpressed cancer cell, and (C) Similar altered behavior of lactate production has been depicted by our computational results in cancer cells compared to that in normal and perturbed ones.

of both normal and perturbed behavior of the integrated biochemical pathways under consideration. Secondly, the altered behavior of the integrated biochemical pathways in mammalian cancer cells compared to normal ones will be discussed. Finally, the third subsection deals with the description of the effects of six possible drug targets in mammalian cancer cells.

3.1. MIMO plant validation and comparison with other mathematical prediction

Here, we have monitored the concentrations of key molecules during simulation of the proposed three timescale state space model for normal as well as perturbed integrated biochemical pathways related to carbon metabolism. In normal scenario (Fig. S2 in Supplementary material), both glucose consumption and production increase. Glucose is broken down into glucose-6P

(G6P). Subsequently, the energy in the form of ATP is consumed. Thus, increasing glucose consumption signifies enhancement of G6P production. The production of ribose 5P and NADPH increases by utilizing higher amount of G6P through PPP. On the other hand, slowing down glyceraldehyde-3P production indicates higher flux through later phase of glycolysis by consuming higher amount of glyceraldehyde-3P. Higher concentration of phosphoenolpyruvate (PEP) supports our claim. Similarly, slowing down the production of pyruvate and acetyl CoA indicates higher flux through TCA cycle and fatty acid synthesis. Enhanced concentrations of oxaloacetate (OAA), citrate and FFA are in conformity with our observation. ATP production increases with higher flux through TCA cycle and later phase of glycolysis. However, after a while ATP decreases due to its higher consumption at early phase of glycolysis. Previous investigations [11,96,97] validate the aforesaid behavior of the proposed model under normal scenario.

Table 4

Illustrating the significance of certain rational drug targets in terms of management of energy and cell proliferation in mammalian cancer cells.

| Drug target | Energy (ATP) production | Cell proliferation indicated by the production of ribose 5P and NADPH | Glucose utilization | p53 expression | Glut1 expression | Remarks | References for validation |
|---|-------------------------|---|-------------------------|-------------------------|-------------------------|---|---------------------------|
| Deactivation of pyruvate kinase | Moderately decreases | Higher production of ribose 5P and NADPH continue | Significantly increases | Significantly decreases | No significant change | Probably not a good choice | - |
| Deactivation of glucose-6-phosphate dehydrogenase | Minutely increases | Both ribose 5P and NADPH significantly decrease | Increases | Significantly increases | Decreases | May be efficient to reduce cell proliferation | [113,111] |
| Deactivation of transketolase | Moderately decreases | Both ribose 5P and NADPH significantly decrease | No significant change | Moderately increases | Significantly decreases | Probable significant | [113–116] |
| Deactivation of ribose 5P isomerase | No significant change | Both ribose 5P and NADPH decrease | Moderately decreases | Significantly increases | Significantly decreases | May be a good choice | [117] |
| Deactivation of glucose-6-phosphate isomerase | Significantly decreases | Both ribose 5P and NADPH significantly decrease | Decreases | Increases | Decreases | Probable effective choice | [118] |
| Activation of pyruvate kinase | Moderately increases | Both ribose 5P and NADPH significantly decrease | Moderately decreases | Significantly increases | Significantly decreases | May be efficient to reduce cancer progression | [47] |

We have further compared the model with a previous simulation result [71] for an environment in cancer cells, particularly, hypoxia condition [98]. Here, we have perturbed the present model to incorporate the mutation exhibiting hypoxia condition. Moreover, the simulation results have been compared with CE-MS measurements during hypoxia in human erythrocytes [71]. As hypoxia in human erythrocytes enhances the expression of certain glycolytic enzymes including hexokinase, aldolase and pyruvate kinase [71], we have applied these enzyme expression values in (0, 1) to the model following their activities as described in the previous investigation [71]. Fig. 6 depicts that our simulation results almost follow not only the CE-MS measurements but also the simulation results obtained by the study [71] in response to the enzyme activities due to hypoxia in human erythrocytes. We have calculated mean squared error (MSE) between our simulated concentrations of eight metabolites (G6P, fructose-6P, fructose 1, 6 bisphosphate, Dhap, 3PG, PEP, pyruvate, and lactate) with time, and the corresponding CE-MS measurements of the same due to aforementioned enzyme activities during hypoxia in human erythrocytes. The MSE values, corresponding to these eight metabolites, are (0.07337, 0.06031, 0.074389, 0.12327, 0.13879, 0.14829, 0.34182, and 0.05727). In addition, we have calculated the MSE between the simulation results of Kinoshita et al. [71] and the same CE-MS measurements as before. Here, we have found MSE values as (0.05610, 0.027992, 0.10074, 0.12518, 0.12673, 0.15056, 0.15697, and 0.17443). It indicates that MSE values, corresponding to our simulation results, are slightly higher for G6P, Fructose-6P, 3PG, and Pyruvate than the same corresponding to Kinoshita et al. On the other hand, our simulation results are closer to CE-MS measurements than that of Kinoshita et al. for Fructose 1, 6 bisphosphate, Dhap, PEP, and Lactate in terms of the MSE values.

In comparison with some other mathematical models in analyzing energy management in mammalian cancer cells, we have found that a constraint-based flux balance model [66] of CCM pathways in cancer cells has found a set of enzymes, such as lactate dehydrogenase, playing important roles in cancer growth. Similar results have been found using our model. We have elaborately discussed the results in the following sections. However, such a flux balance analysis based modeling does not depict the transient behavior of the molecules involved in the cancer pathway. Another differential equation based model [67] has explored a quantitative relationship between the hypoxia intensity and the intracellular lactate levels in cancer cells. Besides, it has predicted some important regulators of the glycolysis pathway only. Similar type of model [68] has been developed for targeting energy metabolism in pancreatic cancer.

Another model [69] based on optimization technique has shown the lactate-driven coupling for fulfilling energy requirements in cancer cells. However, unlike the proposed model, they do not consider three kinds of pathways with different response time. Moreover, they do not explore the effects of activating/deactivating the probable drug targets to control energy metabolism in cancer cells.

For more validation of the proposed model, we have compared our simulation results with *in vitro* observations from cultured human sertoli cells during insulin deprivation [72]. Such a condition in human sertoli cells reduces the expression level of lactate dehydrogenase (LDH). Besides, the expression level of glut1 has increased [72]. Incorporation of these two enzyme expressions in (0, 1) into our model, we can notice similar behavior of initial glucose consumption, pyruvate consumption and lactate production with the corresponding behavior in cultured human sertoli cells during insulin deprivation (Fig. 7).

In the context of present investigation, we have finally simulated the situation of knocking out certain enzymes, viz., pyruvate dehydrogenase, pyruvate carboxylase, acyl-CoA synthetase, fatty acid synthase, phosphoenolpyruvate carboxykinase 1 and succinyl CoA synthetase, to slow down oxidative phosphorylation through mitochondria leading to “Warburg effect” [24,25]. In order to knock out these enzymes, we have considered the kinetic rate constants as zero for the reactions catalyzed by the above enzymes. Considering the aforesaid perturbed condition, we have monitored the behavioral alteration of different molecules compared to that in normal situation as depicted in Figures S3 and S4 in [Supplementary material](#). Enhancement of pyruvate and lactate production as well as slowing down the production of FFA, citrate and succinate compared to that under normal condition indicates low oxidative phosphorylation. Subsequently, consumption of glucose increases to compensate less ATP production, which leads to reduction of glucose production compared to that under normal condition. In spite of this situation, ATP production decreases significantly. As most of the glucose is utilized through glycolytic flux, NADPH production decreases compared to that in normal mammalian cells. However, production of fructose 6P, acetyl CoA, ribose 5P, fructose 2,6 bisphosphate, PEP and protein kinase B (AKT) does not show any significant difference compared to that under normal condition. Interestingly, expression level of p53 increases to inhibit glucose transporter 1 (glut1) [99], which is responsible for transportation of glucose across the plasma membranes of mammalian cells [100]. Thus, higher expression of p53 leads to suppressing the production of glucose in glycolysis pathway. As a

Case 1 (Pyruvate kinase deactivation)

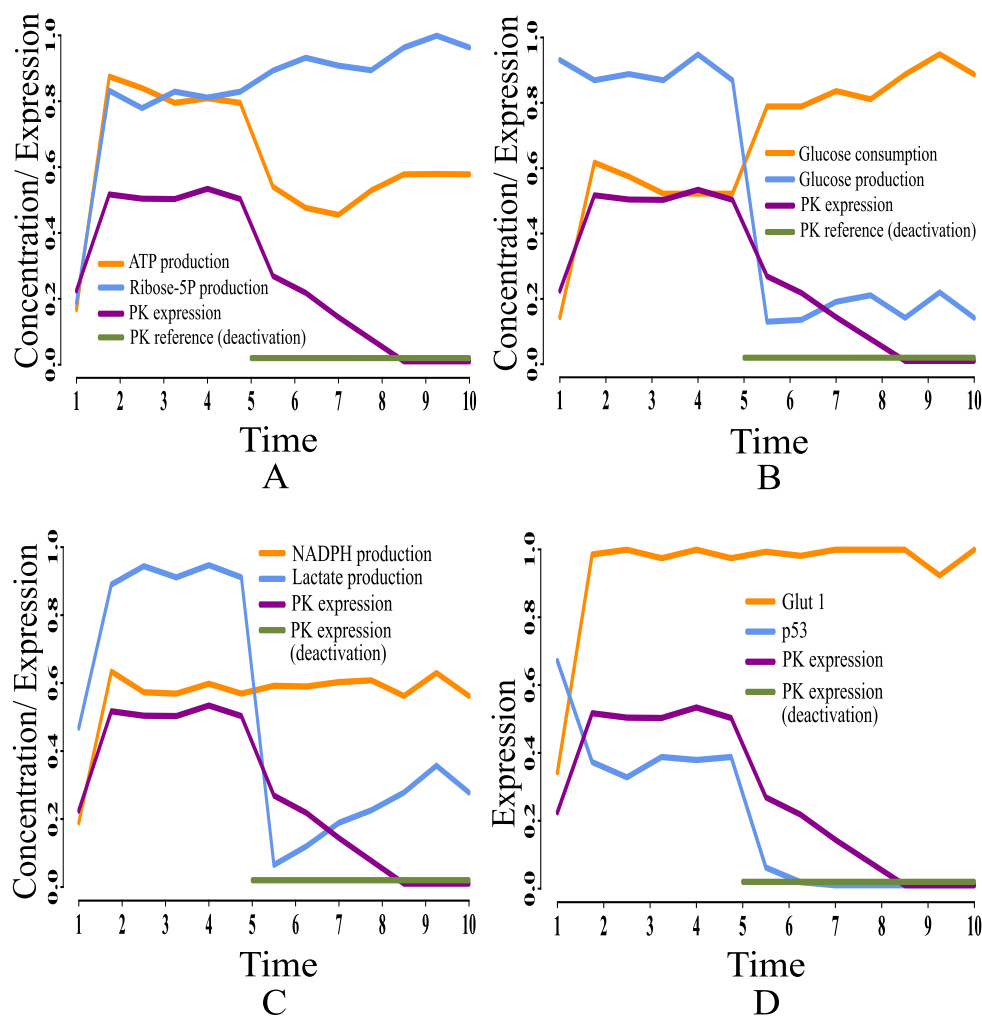


Fig. 19. Case 1: Deactivation of pyruvate kinase as a drug target. After time labeled 5.0 along X-axis, effects of pyruvate kinase deactivation can be noticed.

result, glucose concentration decreases rapidly (Fig. S3 in Supplementary material) leading to cell apoptosis [101].

Nevertheless, mammalian cancer cells somehow manage both energy, in the form of ATP, and cell proliferation to survive [102]. In this study, we aim at investigating the probable mutated regulation that drives mammalian cancer cells to survive despite “Warburg effect”. In this context, we have considered the aforesaid perturbed state space model of integrated biochemical pathways related to carbon metabolism as a MIMO plant of our interest. Thereafter, we have applied the GA controller to achieve high concentrations of ATP (0.95) and ribose 5P (0.70). Here, high ATP will supply constant energy in mammalian cancer cells to survive, while ribose 5P plays an important role in nucleotide synthesis promoting cell growth and proliferation [27]. Thus, the proposed MIMO plant along with GA controller mimics the altered nonlinear dynamics of mammalian cancer cells. We are now going to describe the probable mutated regulations of cancer cells in the following subsection.

3.2. Analysis of regulations in mammalian cancer cells

Here, we have considered the integrated CCM pathway with incorporation of possible mutation responsible for energy man-

agement in cancer cells. The necessary pathway and information have been obtained from KEGG database [23] and literature. As a result, oncogenic somatic and germline mutation have automatically been taken care. The simulation results have found that ATP and ribose 5P production in mammalian cancer cells becomes sufficiently higher than in normal ones following the reference outputs to GA controller (Figs. 8A and 8D). In this situation, we have monitored the altered behavior of other molecules that assist cancer cells to survive and grow. We have found that the expression level of PHD decreases significantly (Fig. 9E) compared to that in normal cells. As PHD accelerates the degradation of hypoxia-inducible factor-1 α (HIF-1 α) [108,78], the expression level of HIF-1 α is quite higher (Fig. 10B) due to less expressed PHD. Subsequently, the expression levels of phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K) (Fig. 11D), AKT (Fig. 12B), mammalian target of rapamycin (mTOR) (Fig. 11F), MYC (Fig. 13C), and extracellular signal-regulated kinases (ERK) (Fig. 9F) increase significantly in cancer cells compared to that in normal ones. Besides, signal transducer and activator of transcription 3 (STAT3) and nuclear factor kappa-light-chain-enhancer of activated B cells (NF- κ B) are highly expressed (Fig. 14F and 11B). On the other hand, p53 (Fig. 15B) shows under expression compared to normal mammalian cells.

Case 2 (Glucose-6-phosphate dehydrogenase deactivation)

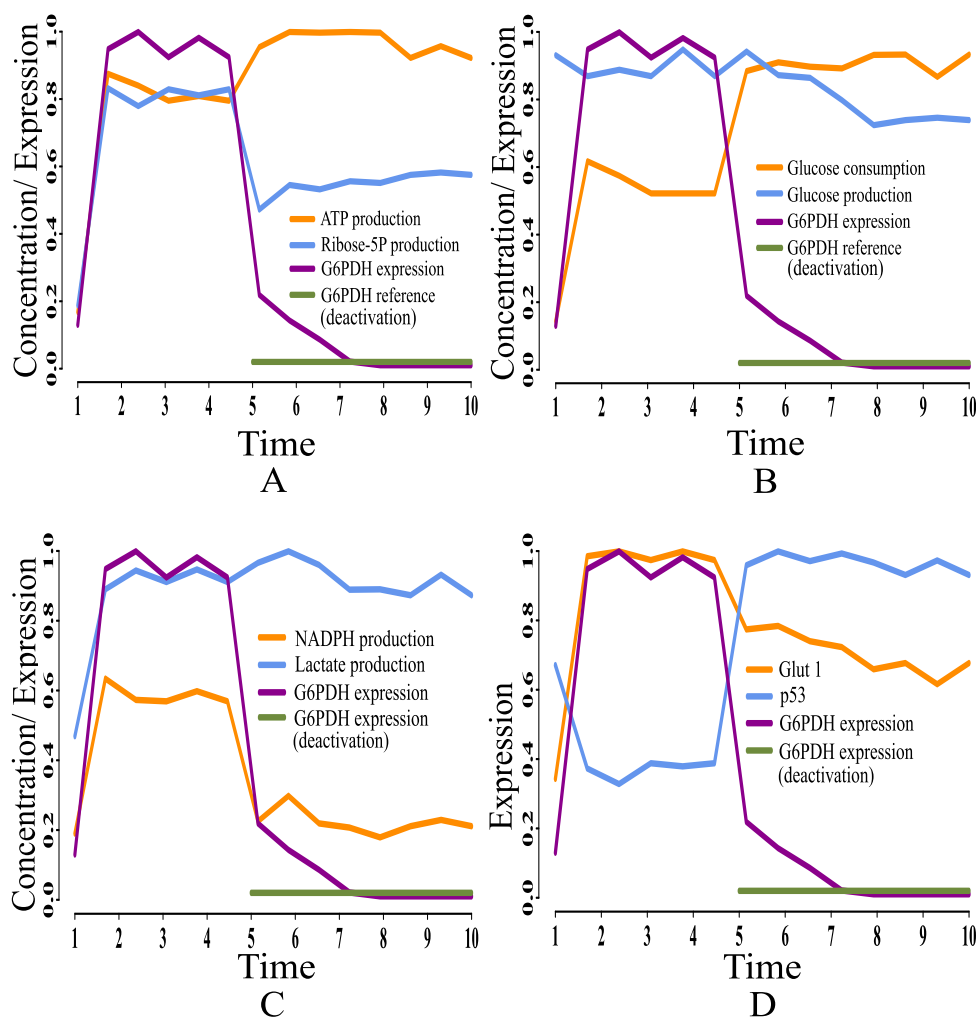


Fig. 20. Case 2: Deactivation of glucose-6-phosphate dehydrogenase as a drug target. After time labeled 5.0 along X-axis, effects of glucose-6-phosphate dehydrogenase deactivation can be noticed.

In support of our simulation results, we have found some previous investigations, including *in vivo* and *in vitro* experiments [28–44,64,65], which show similar kinds of behavior. Besides, we have collected the recent *in vivo/in vitro* data [54,56,57,55,58–63] of human cancer cells to compare different mutations verified in experimental laboratories with our simulation results. Here, we have normalized the experimental data in [0, 1]. According to a clinical study [28], the data, collected from patients with breast, cervical and endometrial cancers at early stage, show high mortality rate of the patients having tumors with over expressed HIF-1 α transcription factor. PHD inhibition promotes more stabilization of HIF-1 α in cancer cells than in normal ones [29,30]. In this context, Kwon et al. [54] has found higher expression level of HIF-1 α measured in human hepatocellular carcinoma (HCC) specimens compared to the non-cancerous tissue specimens (Fig. 10A). Previous studies [31–33] have shown that PI3K, AKT and mTOR express significantly higher in tumor tissues of patients suffering from ovarian, gastric and prostate cancer compared to normal individuals. Experiment [55] with human colon cancer sample has depicted higher expression of PI3K compared to normal ones (Fig. 11C). Immunohistochemical scores measured by Mao et al. [56] has depicted

that AKT expression is significantly higher in pancreatic cancer (91 cases) than that in normal pancreas (51 cases) (Fig. 12A). Besides, Roy et al. [57] has shown the differentially expressed AKT isoforms in normal and malignant oral tissues through immunohistochemical analysis of the human samples (Fig. 12A). On the other hand, higher mTOR expression has been found in colon cancer compared to normal one as per the *in vivo/in vitro* experiment performed by Zhang et al. [55] (Fig. 11E).

Evidences [34,35,65] suggest that over expression of MYC triggers certain genes to promote growth and proliferation of cancer cells. Similar behavior of MYC in cancer cells has been found in two recent *in vivo/in vitro* experiments [58,59] (Fig. 13A and 13B). It is reported that tumor suppressor p53 is under expressed [36] in cancer cells due to lysine methylation [37]. Moreover, evidences [38,39,40] have shown that highly expressed ERK in cancer cells promotes over expression of N-cadher protein involved in metastasis. Subsequently, previous investigations [41,42] have demonstrated that enhanced expression of STAT3 signaling protein either inhibits apoptosis or accelerates cell proliferation, angiogenesis, and finally metastasis. Consequently, it leads to cancer initiation and progression. Besides, over expression of NF- κ B plays an

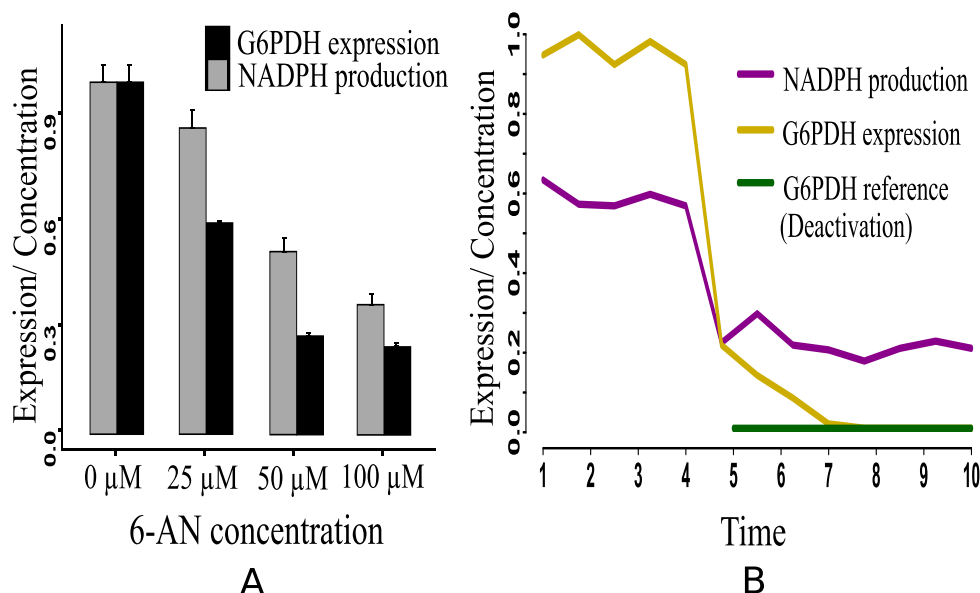


Fig. 21. Effect of glucose-6-phosphate dehydrogenase deactivation (case 2): (A) The relative enzymatic activity (expression level) of glucose-6-phosphate dehydrogenase controls the NADPH production [111]. Lower expression of glucose-6-phosphate dehydrogenase leads to reduce the concentration of NADPH. Here, expression of glucose-6-phosphate dehydrogenase and NADPH concentration have been assessed in A549/DDP cells according to various concentrations of 6-Aminonicotinamide (6-AN) (unit in μM), (B) The proposed computational model depicts similar behavior of NADPH due to the deactivation of glucose-6-phosphate dehydrogenase.

important role in this context [39,43,44]. The previous *in vivo/in vitro* studies have supported the aforementioned claim by illustrating over expressed STAT3 in certain cancer cells, such as breast cancer [60], gastric cancer [61] and colon cancer [62] cells compared to normal ones (Fig. 14E). Besides, the evidence [63] has also depicted that relative expression of NF- κ B significantly increases in cancer cells (CT-2A astrocytoma) compared to normal ones (Fig. 11A).

The present simulation results demonstrate higher expression of hexokinase (HK) and glut1 (Figs. 16D and F) in cancer cells. Previous *in vivo/in vitro* result [106] depicting higher expression of HK mRNA levels (Fig. 16C) in cancer compared to normal liver tissue has validated the computational result. This experiment [106] has also shown higher glut1 mRNA levels (Fig. 16E) as 92-fold higher in Meta specimens compared to normal liver tissue. Here, enhanced expressions of HIF-1 α , PI3K, AKT, MYC and mTOR are responsible for activating HK and glut1 more in cancer cells. As a result, glucose consumption increases (Fig. 17C). Consequently, glucose production decreases as depicted in Fig. 15E. Previous investigations [27,45] support our claim. They have claimed that enhanced glut1 increases the utilization of glucose by anabolic pathways. Besides, highly expressed MYC activates lactate dehydrogenase (LDH) more in cancer cells. It leads to fermentation of glucose [109,25] through enhanced glycolysis [46]. Consequently, lactate production increases. The present results have followed the above claims by showing higher expression of LDH (Fig. 16B) as well as enhanced lactate production (Fig. 18C). In this context, Lim et al. [107] has measured higher glucose uptake (Fig. 17A) in B7-H3 knockdown cells grown in normoxic or perturbed conditions (hypoxia) for 24 h. Besides, Li et al. [105] has shown that glucose uptake increases significantly in Dppa4 overexpressed cancer cells (Fig. 17B). This experiment [105] has also shown that LDH has significantly been up regulated (Fig. 16A) in Dppa4 overexpressed cancer cells. As a result, higher lactate production (Fig. 18B) has been found here [105]. Lim et al. [107] has also found similar higher lactate production (Fig. 18A) in B7-H3 knockdown cells as mentioned before. These experimental results provide a strong support to the present computational results.

Higher expression levels of phosphofructokinase 1 (PFK1) (Fig. 14B), phosphofructokinase 2 (PFK2) (Fig. 9A), and glyceraldehyde-3-phosphate dehydrogenase (Fig. 15A) also confirm enhanced glycolysis in mammalian cancer cells. In support of the computational results, a previous *in vivo/in vitro* result [104] has depicted higher PFK1 expression (Fig. 14A) in breast cancer and paracancer tissues, expressed as units per gram of protein (U/gprot). Consequently, higher amount of fructose 6P is utilized to produce more fructose 2,6 biphosphate and fructose 1,6 biphosphate than in normal cells. Besides, higher amount of fructose 2,6 biphosphate accelerates the break down of fructose 6P into fructose 1,6 biphosphate [96]. That is why simulation result has shown reduction of fructose 6P production (Fig. 8B). Similarly, the concentrations of PEP drops (Fig. 8E) due to its over consumption to produce higher amount of pyruvate and ATP (Figs. 8F and 8A) compared to normal cells. Here, expression of pyruvate kinase switches alternatively from low to high and vice versa as depicted in Fig. 14D. Evidences [47–49,64] have shown that pyruvate kinase (M2 isoform) switches to its inactive dimer form or active tetrameric form according to the requirements of mammalian cancer cells. When pyruvate kinase (M2 isoform) is in tetrameric form, the flux through glycolysis enhances with sufficient amount of ATP production. As a result, production of intermediate glycolytic metabolites, such as glyceraldehyde 3P (Fig. 8C) and fructose 1,6 biphosphate, increases. Conversely, dimer form of pyruvate kinase (M2 isoform) promotes macromolecular synthesis from glycolytic intermediate metabolites to continue cell growth and proliferation. An evidence [104] has also reported about higher pyruvate kinase expression (Fig. 14C) in breast cancer and paracancer tissues, expressed as units per gram of protein (U/gprot). Thus, cancer cells manage energy in the form of ATP to survive in spite of slow oxidative phosphorylation under consideration. Here, reduction of reduced nicotinamide adenine dinucleotide (NADH) production (Fig. 10D) indicates successful incorporation of slow oxidative phosphorylation [50] into the proposed model. An *in vivo/in vitro* experiment, performed by Sumi et al. [103] depicting lower concentration of NADH (Fig. 10C) in cancer cells compared to normal ones, supports the computational result.

Case 3 (Transketolase deactivation)

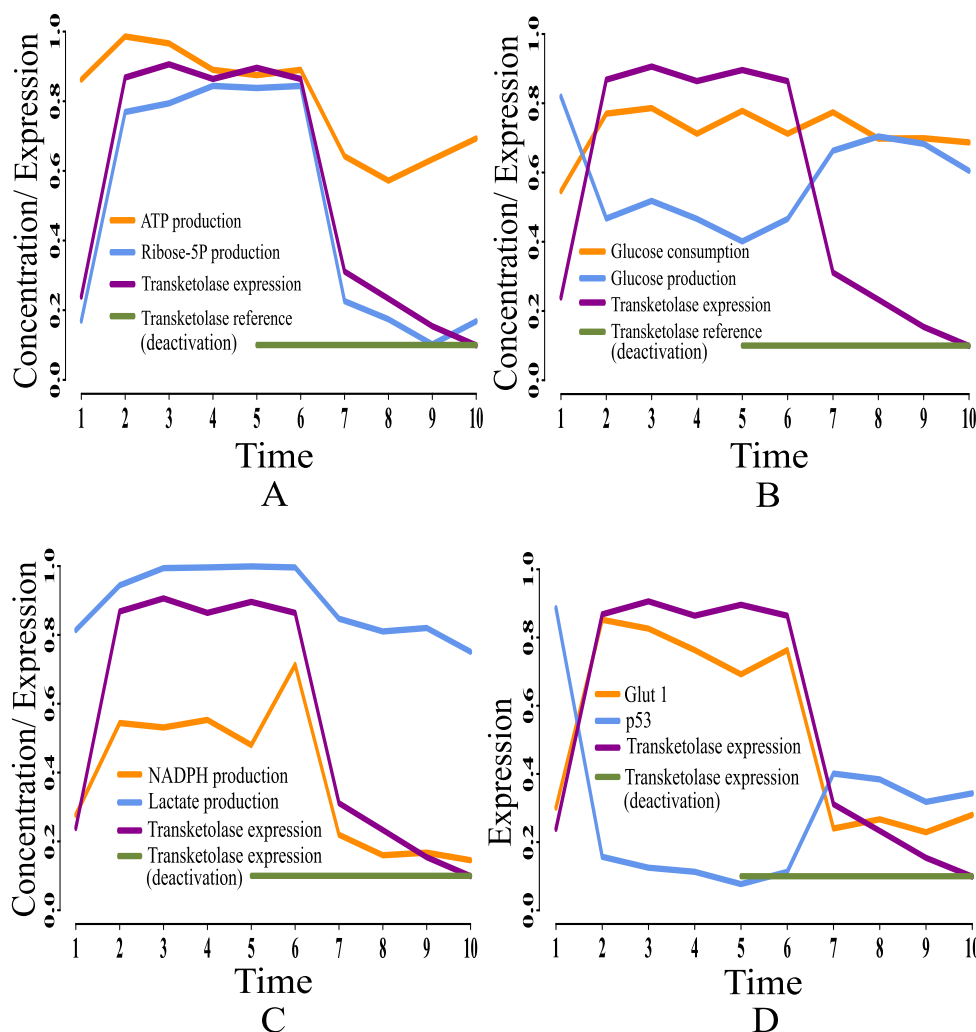


Fig. 22. Case 3: Deactivation of transketolase as a drug target. After time labeled 5.0 along X-axis, effects of transketolase deactivation can be noticed.

According to the simulation results, the enzymes (proteins) catalyzing PPP, such as glucose-6-phosphate dehydrogenase (Fig. 9D), phospho-gluco dehydrogenase (Fig. 15C), ribose 5P isomerase (Fig. 9C), transketolase and transaldolase (Fig. 15D and 9B), have shown over expression in cancer cells for macromolecular precursors required for cell growth and proliferation. Under expressed p53 plays an important role in this context. High concentration of ribose 5P (Fig. 8D) confirms enhanced production of cell building material, including DNA, RNA, nucleic acids and histidine, in mammalian cancer cells. Evidences [51–53] have shown that relatively higher expressions of glucose-6-phosphate dehydrogenase, phospho-gluco dehydrogenase, ribose 5P isomerase, transketolase and transaldolase help cancer cells in generating high amount of NADPH and ribose 5P, which are responsible for reactive oxygen species (ROS) reduction as well as the production of high levels of nucleotides for DNA synthesis and repair. It leads to resistance against certain cancer therapies resulting in enhancement of oxidative stress or DNA damage. In this context, it should be mentioned that higher amount of ROS can frequently be observed in cancer cells helping in activation of oncogenes and metastasis. However, further enhancement of ROS beyond a certain threshold induces cell death [110]. Finally, we have summarized the altered

regulations of different metabolites, transcription factors and genes in cancer cells compared to that in normal ones in Table 3.

3.3. Analysis of certain rational drug targets in terms of management of energy and cell proliferation in mammalian cancer cells

Here, we are going to discuss the effects of certain rational drug targets on cancer cells in terms of management of energy and cell proliferation from simulation point of view. These results may have significant implications during *in vivo* and/or *in vitro* experiments. When the proposed model computationally meets the reference target concentrations of ATP and ribose 5P for GA controller, we have observed that the model has mimicked the altered regulation of cancer cells as previously discussed. At that moment, we have set reference target expression levels (high or low) of certain proteins/enzymes as drug targets using GA controller. Although the model is quite capable of analyzing the effect of any drug target, we have considered only six drug targets among others for our study to restrict the size of present article. Here, we have monitored the effects of deactivating pyruvate kinase (reference expression level 0.02), glucose-6-phosphate dehydrogenase (reference expression level 0.02), transketolase (reference expres-

Case 4 (Ribose 5P isomerase deactivation)

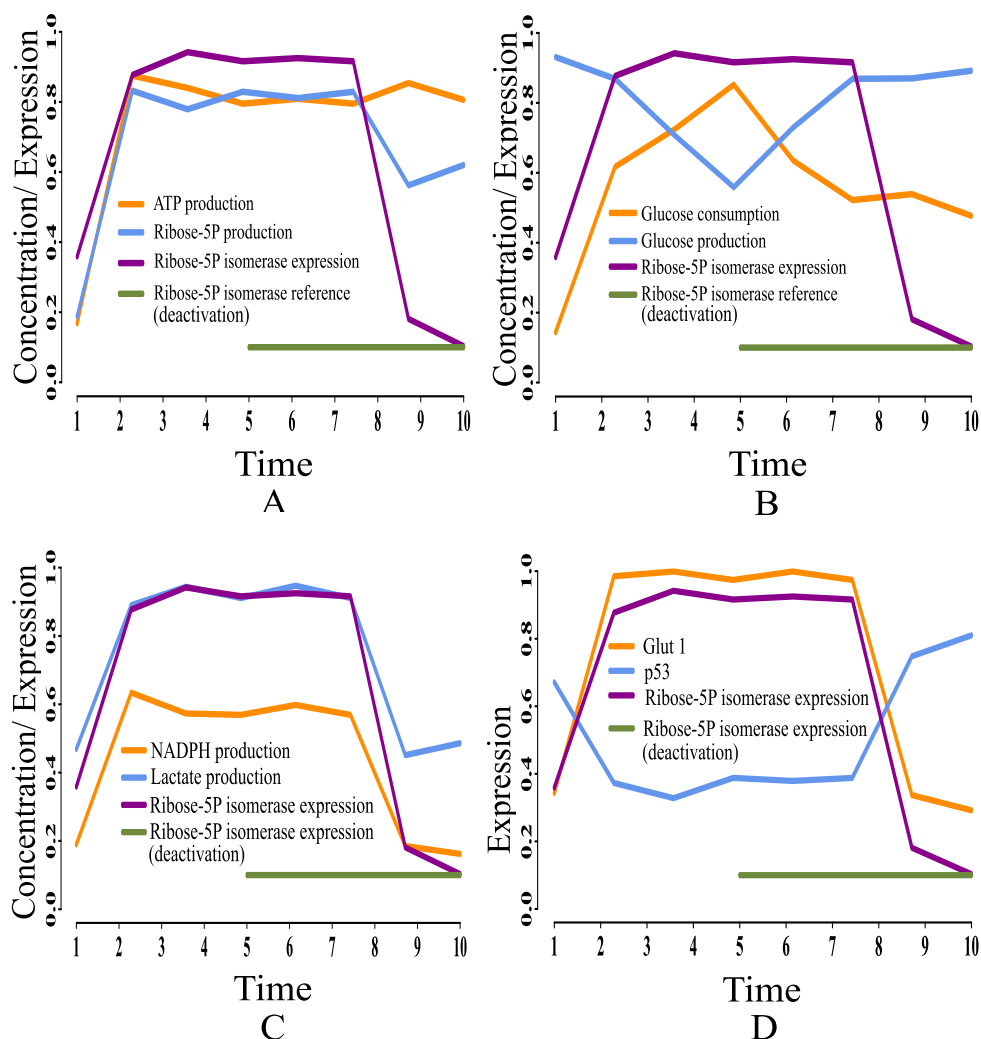


Fig. 23. Case 4: Deactivation of ribose 5P isomerase as a drug target. After time labeled 5.0 along X-axis, effects of ribose 5P isomerase deactivation can be noticed.

sion level 0.09), ribose 5P isomerase (reference expression level 0.09) and glucose-6-phosphate isomerase (reference expression level 0.03). Finally, the effect of activating pyruvate kinase (reference expression level 0.98) has also been observed and analyzed. Table 4 summarizes the significance of the drug targets under consideration in energy supply and cell proliferation.

- Case 1 (Pyruvate kinase deactivation):** According to the simulation results as depicted in Fig. 19, we have found that deactivation of pyruvate kinase cannot reduce the concentration of ribose 5P (Fig. 19A) and NADPH (Fig. 19C). In other words, high proliferation of mammalian cancer cells continues. Besides, enhanced NADPH still prevents from ROS production helping cancer cells to survive in spite of oxidative stress. Although the concentrations of ATP (Fig. 19A) and lactate (Fig. 19C) reduce at the beginning, after a while they increase again significantly. Even pyruvate kinase deactivation cannot reduce high glucose utilization (*i.e.*, high consumption and less production) (Fig. 19B) in mammalian cancer cells. Moreover, high expression of glut1 and under expression of tumor suppression protein p53 (Fig. 19D) cannot be reversed in this case. Thus, pyruvate kinase deactivation may not be a good choice as a drug target for mammalian cancer therapy.

- Case 2 (Glucose-6-phosphate dehydrogenase deactivation):** Glucose-6-phosphate dehydrogenase is the key enzyme (protein) to utilize G6P through PPP. It leads to producing high amount of NADPH and ribose 5P so that nucleotides and fatty acid can sufficiently be synthesized to maintain intracellular redox homeostasis [112]. The present results (Fig. 20) have shown that the concentrations of ribose 5P (Fig. 20A) and NADPH (Fig. 20C and 21B) decrease due to deactivation of glucose-6-phosphate dehydrogenase. In support of this result, a recent *in vivo/in vitro* study has shown how the relative activity (expression level) of glucose-6-phosphate dehydrogenase enzyme controls the NADPH production [111]. Lower expression of glucose-6-phosphate dehydrogenase leads to reduction of the concentration of NADPH (Fig. 21A). Here, expression of glucose-6-phosphate dehydrogenase and NADPH concentration have been assessed in A549/DDP cells according to various concentrations of 6-Aminonicotinamide (6-AN). Besides, expression level of p53 increases (Fig. 20D), whereas glut1 expression decreases (Fig. 20D). These results indicate reduction of cell growth and proliferation due to glucose-6-phosphate dehydrogenase deactivation. However, high amount of glucose (Fig. 20B) is utilized through glycolysis to generate sufficient ATP (Fig. 20A) for cell survival. Subsequently, cell fer-

Case 5 (Glucose 6P isomerase deactivation)

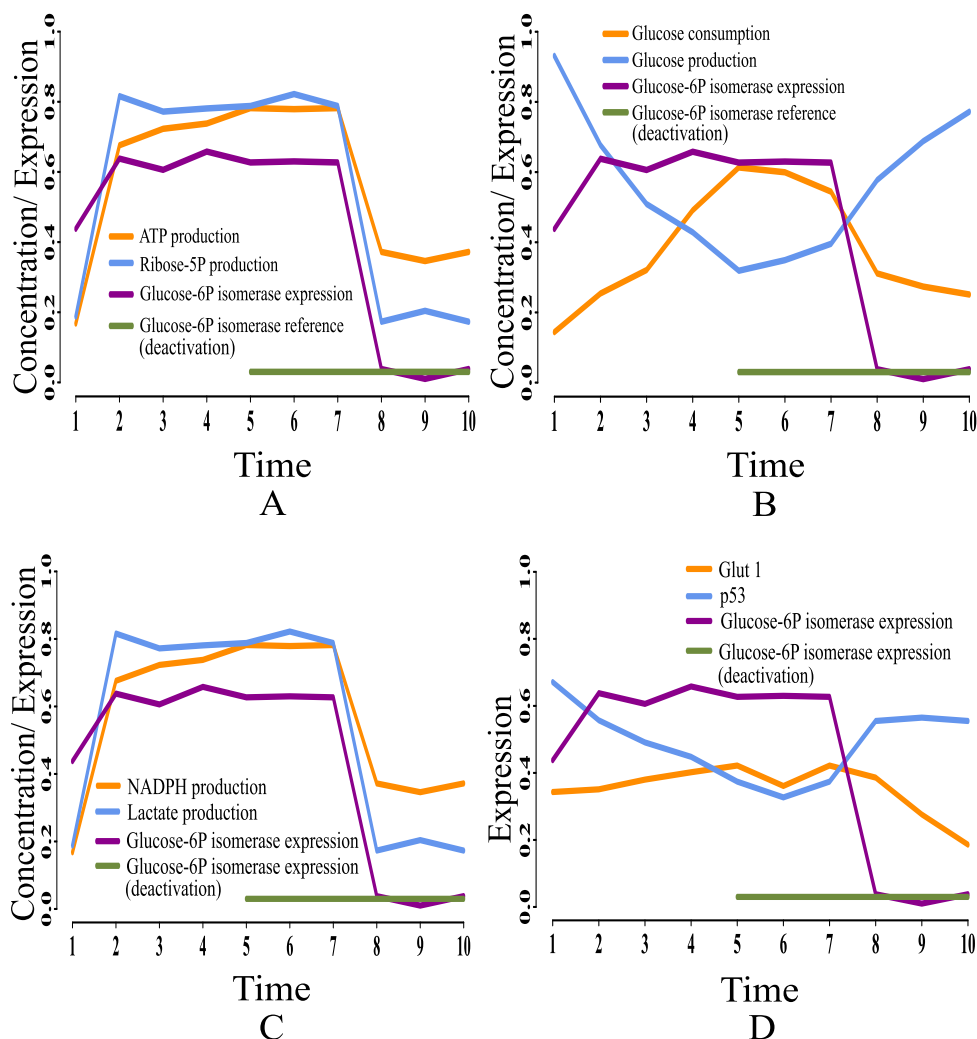


Fig. 24. Case 5: Deactivation of glucose-6-phosphate isomerase as a drug target. After time labeled 5.0 along X-axis, effects of glucose-6-phosphate isomerase deactivation can be noticed.

mentation continues through high amount of lactate production (Fig. 20C). A previous study [113] has shown reduction of cell proliferation by silencing glucose-6-phosphate dehydrogenase in the human breast cancer cell line MCF7. Thus, deactivation of glucose-6-phosphate dehydrogenase as a possible drug target may be efficient to reduce cell growth and proliferation but not be effective in reduction of energy supply and fermentation.

- Case 3 (Transketolase deactivation):** Transketolase has similar importance as glucose 6-phosphate dehydrogenase to maintain cell growth and proliferation leading to metastasis [114]. Simulation results (Fig. 22) demonstrate that the concentration of ribose 5P (Fig. 22A) decreases significantly in accordance with the deactivation of transketolase. Subsequently, enhanced expression of p53 (Fig. 22D) conveys that cell growth and proliferation are inhibited in this case. Reduction of NADPH concentration (Fig. 22C) signifies enhancement of ROS preventing from survival of cancer cells by oxidative stress. Moreover, ATP (Fig. 22A) and lactate (Fig. 22C) production also decrease through reduced glucose utilization (Fig. 22B). Besides, expression level of glut1 decreases as depicted in Fig. 22D. Previous

investigations [113,115,116] have shown similar effects due to transketolase silencing. Thus, it is clear that targeting transketolase may be a potential choice for future cancer therapy.

- Case 4 (Ribose 5P isomerase deactivation):** We have already discussed that up regulation of ribose 5P isomerase plays an important role in cell growth and proliferation of a cancer patient. A previous study [117] has claimed that expression level of ribose 5P isomerase is enhanced in colorectal cancer. Our simulation results (Fig. 23) have shown the effect of silencing ribose 5P isomerase in cancer cells. We have observed that enhanced p53 expression and decreased glut1 expression (Fig. 23D) indicate slowing down of proliferation and growth of cancer cells. In this context, reduction in NADPH (Fig. 23C) and ribose 5P (Fig. 23A) concentration confirm our claim. Subsequently, glucose utilization (Fig. 23B) and cell fermentation (i.e., lactate production as depicted in Fig. 23C) decrease. However, energy supply in the form of ATP (Fig. 23A) in cancer cells is somehow managed in spite of the deactivation of ribose 5P isomerase. Consequently, ribose 5P isomerase might be considered as a biomarker for targeted cancer therapy and prediction.

- Case 5 (Glucose-6-phosphate isomerase deactivation):** Glucose-6-phosphate isomerase or phosphoglucose isomerase (PGI) is a glycolytic enzyme that directs G6P flux into the glycolysis branch. As a result, G6P breaks down into fructose 6P. Deactivation of Glucose-6-phosphate isomerase results in slowing down of glucose utilization (Fig. 24B) as well as lactate (Fig. 24C) and ATP (Fig. 24A) production. Thus “Warburg effect” is somehow reversed. Besides, reduced glut1 expression level along with enhanced p53 expression (Fig. 24D) may slow down growth and proliferation rate of cancer cells. In this context, reduction of NADPH (Fig. 24C) and ribose 5P (Fig. 24A) concentration supports our claim. Here, a previous investigation [118] has shown that down regulation of Glucose-6-phosphate isomerase may suppress “Warburg effect”. Besides, oxidative phosphorylation is activated. However, its impact on tumor growth is minimal except in the case of hypoxia. Thus, glucose-6-phosphate isomerase can be a potential drug target for future cancer therapy.
- Case 6 (Pyruvate kinase activation):** Activation of pyruvate kinase (Fig. 25) may slow down cell growth and proliferation as well as cell fermentation. Decreased concentrations of NADPH, lactate (Fig. 25C) and ribose 5P (Fig. 25A) confirm our

claim. Besides, decreased glucose utilization (Fig. 25B) and glut1 expression (Fig. 25D) support the fact. Here, expression level of p53 (Fig. 25D) also increases to slow down cell growth and proliferation. However, ATP is produced sufficiently in this case as depicted in Fig. 25A. Evidence [47] has claimed that cancer cell growth and proliferation are inhibited by activation of pyruvate kinase (M2 isoform). Thus, pyruvate kinase (M2 isoform) activators may be considered as possible significant drugs for oncogenic treatment.

4. Conclusion

In this study, we have successfully integrated three types of biochemical pathways, viz., metabolic, signaling and gene regulatory networks, keeping their three timescale nature intact. Here, we have developed the integrated state equations considering appropriate timescales as well as all possible perturbations present in the contemplated integrated biochemical pathway. Besides, depending on the training dataset generated by solving the pathway ODEs, SVR based MIMO model has been developed. The MIMO model can mimic the transient nonlinear dynamic behavior of the integrated biochemical pathway under consideration. Moreover,

Case 6 (Pyruvate kinase activation)

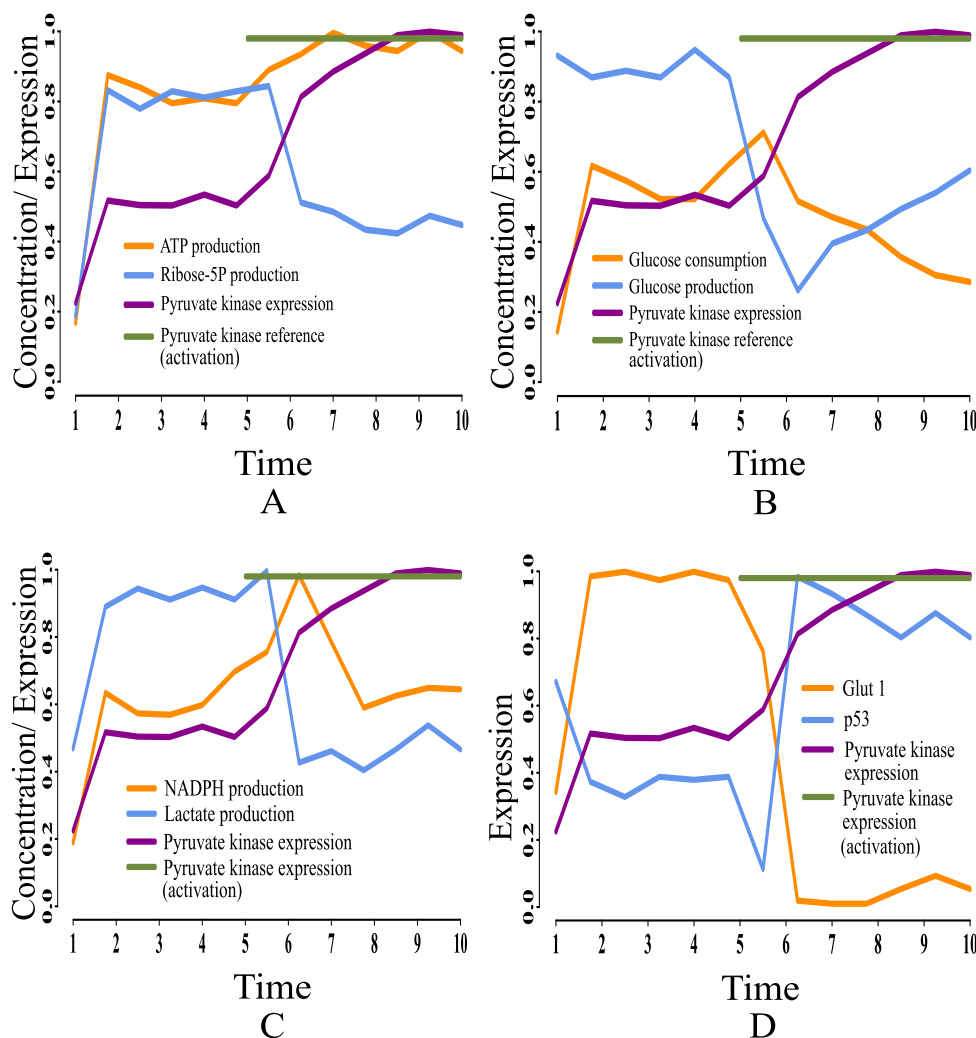


Fig. 25. Case 6: Activation of pyruvate kinase as a drug target. After time labeled 5.0 along X-axis, effects of pyruvate kinase activation can be noticed.

with the help of the GA controller, the model can predict the effect of drug targets applied to complex diseased cells. In order to investigate the effectiveness of the model, we have used our model to explore how mammalian cancer cells are able to manage their growth, proliferation and energy supply to survive. In this context, “Warburg effect” [24,25] has been taken into account. The simulation results have depicted that the model has not only captured the key regulations, but also has been able to predict certain possible drug effects in terms of energy and cell proliferation management in mammalian cancer cells.

According to the results, the proteins or genes HIF-1 α , HK, glut1, AKT, glyceraldehyde-3-phosphate dehydrogenase, phospho-gluco dehydrogenase, ERK, ribose 5P isomerase, mTOR, glucose-6-phosphate dehydrogenase, STAT3, NF- κ B, PI3K, MYC, LDH, PFK1, PFK2, transketolase and transaldolase are up regulated in cancer cells. Besides, PHD and p53 are down regulated. Switching of pyruvate kinase (M2 isoform) between its two oligomeric form, viz., inactive dimer and active tetramer, plays an important role in managing proliferation, growth and energy in mammalian cancer cells. These results have been validated through previous investigations involving *in vivo* and *in vitro* experiments [24,27,26,28–45,25,46–54,56,57,55,58–65]. Besides, other mathematical models [66–69] support the results derived by the proposed model. Among six drug targets under consideration, deactivation of transketolase and glucose-6-phosphate isomerase may be the most potential to slow down cancer progression by reducing cell proliferation, growth, fermentation and energy supply. On the other hand, pyruvate kinase (M2 isoform) activation and ribose 5P isomerase deactivation may reduce cell growth, proliferation and fermentation during cancer. However, they may not be able to stop energy supply in mammalian cancer cells. Although deactivation of glucose-6-phosphate dehydrogenase may slow down cell growth and proliferation, it may fail to stop fermentation and energy supply in the malignant cells. In this context, pyruvate kinase deactivation may be an awful choice as a rational drug target for future cancer therapy.

Finally, it may be mentioned that the experimental values of different kinetic parameters are still unknown or poorly documented. In this study, we have estimated these values by trial and error based on previous knowledgebase. Some of the parameter values considered here have been seen to be close enough with those estimated by the method of Lillacci et al. [70]. The other parameter values could not be checked due to unavailability of required experimental observations. However, trial and error based technique is time consuming and difficult to perform because of large number of such kinetic parameters. This is a drawback of the present methodology. Thus more *in vivo/in vitro* parameter values are needed, and/or appropriate parameter estimation methods based on the theory of machine learning and control theory can be developed for designing more accurate MIMO plant for an integrated biochemical pathway.

Authors' Contributions

AD and RKD initiated control theoretic modeling of metabolic pathways. AD conceptualized the basic idea of three timescale modeling. AD and AB formulated the methodology and implemented it. RKD gave crucial theoretical input. AD and AB wrote the first draft of the manuscript. RKD corrected it. NC read the article and gave fruitful suggestions to edit the manuscript.

Funding

No funding agency has funded this work.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Abhijit Dasgupta acknowledges Digital India Corporation (formerly Media Lab Asia), Ministry of Electronics and Information Technology, Government of India, for providing him a Senior Research Fellowship under the Visvesvaraya Ph.D. scheme for Electronics and IT. Rajat K. De acknowledges SyMeC Project grant [BT/Med-II/NIBMG/SyMeC/2014/Vol. II] given to the Indian Statistical Institute by the Department of Biotechnology (DBT), Government of India.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.csbj.2020.12.019>.

References

- [1] Steuer R, Gross T, Selbig J, Blasius B. Structural kinetic modeling of metabolic networks. *Proc Natl Acad Sci* 2006;103(32):11868–73.
- [2] Hornberg JJ, Bruggeman FJ, Westerhoff HV, Lankelma J. Cancer: a systems biology disease. *Biosystems* 2006;83(2):81–90.
- [3] Kholodenko BN. Cell-signalling dynamics in time and space. *Nat Rev Mol Cell Biol* 2006;7(3):165–76.
- [4] Legewie S, Blüthgen N, Herzog H. Mathematical modeling identifies inhibitors of apoptosis as mediators of positive feedback and bistability. *PLoS Comput Biol* 2006;2(9). e120.
- [5] Klipp E, Liebermeister W. Mathematical modeling of intracellular signaling pathways. *BMC Neurosci* 2006;7(1):S10.
- [6] Aldridge BB, Burke JM, Lauffenburger DA, Sorger PK. Physicochemical modelling of cell signalling pathways. *Nat Cell Biol* 2006;8(11):1195–203.
- [7] Schlitt T, Brazma A. Current approaches to gene regulatory network modelling. *BMC Bioinf* 2007;8(S6):S9.
- [8] Karlebach G, Shamir R. Modelling and analysis of gene regulatory networks. *Nat Rev Mol Cell Biol* 2008;9(10):770–80.
- [9] Hecker M, Lambeck S, Toepfer S, Van Someren E, Guthke R. Gene regulatory network inference: data integration in dynamic models? a review. *Biosystems* 2009;96(1):86–103.
- [10] De Rajat K, Das M, Mukhopadhyay S. Incorporation of enzyme concentrations into fba and identification of optimal metabolic pathways. *BMC Syst Biol* 2008;2(1):65.
- [11] De RK, Tomar N. Modeling the optimal central carbon metabolic pathways under feedback inhibition using flux balance analysis. *J Bioinf Comput Biol* 2012;10(06):1250019.
- [12] Cascante M, Boros LG, Comin-Anduix B, de Atauari P, Centelles JJ, Lee PWN. Metabolic control analysis in drug discovery and disease. *Nat Biotechnol* 2002;20(3):243–9.
- [13] Paul D, Dasgupta A, De RK. Exploring the altered dynamics of mammalian central carbon metabolic pathway in cancer cells: a classical control theoretic approach. *PLoS One* 2015;10(9). e0137728.
- [14] Dasgupta A, Paul D, De RK. A fuzzy logic controller based approach to model the switching mechanism of the mammalian central carbon metabolic pathway in normal and cancer cells. *Mol BioSyst* 2016;12(8):2490–505.
- [15] Dasgupta A, Bandyopadhyay GK, Ray I, Bandyopadhyay K, Chowdhury N, De RK, Mahata SK. Catestatin improves insulin sensitivity by attenuating endoplasmic reticulum stress: *in vivo* and *in silico* validation. *Comput Struct Biotechnol J* 2020;18:464.
- [16] Natarajan M, Lin KM, Hsueh RC, Sternweis PC, Ranganathan R. A global analysis of cross-talk in a mammalian cellular signalling network. *Nat Cell Biol* 2006;8(6):571–80.
- [17] Papin JA, Hunter T, Palsson BO, Subramaniam S. Reconstruction of cellular signalling networks and analysis of their properties. *Nat Rev Mol Cell Biol* 2005;6(2):99–111.
- [18] Pe'er D. Bayesian network analysis of signaling networks: a primer. *Sci STKE* 2005;281. I4.
- [19] Morris MK, Saez-Rodriguez J, Sorger PK, Lauffenburger DA. Logic-based models for the analysis of cell signaling networks. *Biochemistry* 2010;49(15):3216–24.
- [20] Lee JM, Gianchandani EP, Eddy JA, Papin JA. Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Comput Biol* 2008;4(5). e1000086.

- [21] Rosano GL, Ceccarelli EA. Recombinant protein expression in escherichia coli: advances and challenges. *Front Microbiol* 2014;5:172.
- [22] Theobald U, Mailinger W, Baltes M, Rizzi M, Reuss M. In vivo analysis of metabolic dynamics in saccharomyces cerevisiae: I. Experimental observations. *Biotechnol Bioeng* 1997;55(2):305–16.
- [23] Kanehisa M, Goto S. Kegg: kyoto encyclopedia of genes and genomes. *Nucl Acids Res* 2000;28(1):27–30.
- [24] Vander Heiden MG, Cantley LC, Thompson CB. Understanding the warburg effect: the metabolic requirements of cell proliferation. *Science* 2009;324(5930):1029–33.
- [25] Koppenol WH, Bounds PL, Dang CV. Otto warburg's contributions to current concepts of cancer metabolism. *Nat Rev Cancer* 2011;11(5):325.
- [26] Kim JW, Dang CV. Cancer's molecular sweet tooth and the warburg effect. *Cancer Res* 2006;66(18):8927–8930.
- [27] Boroughs LK, DeBerardinis RJ. Metabolic pathways promoting cancer cell survival and growth. *Nat Cell Biol* 2015;17(4):351.
- [28] Semenza GL. Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics. *Oncogene* 2010;29(5):625.
- [29] Nguyen TL, Durán RV. Prolyl hydroxylase domain enzymes and their role in cell signaling and cancer metabolism. *Int J Biochem Cell Biol* 2016;80:71–80.
- [30] Villar VH, Merhi F, Djavaheri-Mergny M, Durán RV. Glutaminolysis and autophagy in cancer. *Autophagy* 2015;11(8):1198–208.
- [31] Li H, Zeng J, Shen K. Pi3k/akt/mtor signaling pathway as a therapeutic target for ovarian cancer. *Arch Gynecol Obstet* 2014;290(6):1067–78.
- [32] Tapia O, Riquelme I, Leal P, Sandoval A, Aedo S, Weber H, Letelier P, Belloio E, Villaseca M, Garcia P, et al. The pi3k/akt/mtor pathway is activated in gastric cancer with potential prognostic and predictive significance. *Virchows Archiv* 2014;465(1):25–33.
- [33] Hamidi A, Song J, Thakur N, Itoh S, Marcusson A, Bergh A, Heldin CH, Landström M. Tgf- β promotes pi3k-akt signaling and prostate cancer cell migration through the traf6-mediated ubiquitylation of p85 α . *Science Signal* 2017;10(486):eaal4186.
- [34] Stine ZE, Walton ZE, Altman BJ, Hsieh AL, Dang CV. Myc, metabolism, and cancer. *Cancer Discovery* 2015;5(10):1024–39.
- [35] Di Giacomo S, Sollazzo M, de Biase D, Ragazzi M, Bellosta P, Pession A, Grifoni D. Human cancer cells signal their competitive fitness through myc activity. *Scientific Rep* 2017;7(1):12568.
- [36] Muller PA, Vousden KH. Mutant p53 in cancer: new functions and therapeutic opportunities. *Cancer Cell* 2014;25(3):304–17.
- [37] Zhu J, Dou Z, Sammons MA, Levine AJ, Berger SL. Lysine methylation represses p53 activity in teratocarcinoma cancer cells. *Proc Natl Acad Sci* 2016;113(35):9822–7.
- [38] Cho Y.Y. Roles of erk1-rsk2 signaling in human cancers. *Proceedings of the American Association for Cancer Research Annual Meeting*, Washington, DC Philadelphia (PA) 2017. <https://doi.org/10.1158/1538-7445.AM2017-3127>.
- [39] Lee KJ, Yoo JW, Kim YK, Choi JH, Ha TY, Gil M. Advanced glycation end products promote triple negative breast cancer cells via erk and nf- κ b pathway. *Biochem Biophys Res Commun* 2018;495(3):2195–201.
- [40] Liu S, Zha J, Lei M. Inhibiting erk/mnk/eif4e broadly sensitizes ovarian cancer response to chemotherapy. *Clin Transl Oncol* 2018;20(3):374–81.
- [41] Siveen KS, Sikka S, Surana R, Dai X, Zhang J, Kumar AP, Tan BK, Sethi G, Bishayee A. Targeting the stat3 signaling pathway in cancer: role of synthetic and natural inhibitors. *Biochim. Biophys. Acta* 2014;1845(2):136–54.
- [42] Khan MW, Saadalla A, Ewida AH, Al-Katranji K, Al-Saoudi G, Giaccone ZT, Gounari F, Zhang M, Frank DA, Khazaie K. The stat3 inhibitor pyrimethamine displays anti-cancer and immune stimulatory effects in murine models of breast cancer. *Cancer Immunol. Immunother.* 2018;67(1):13–23.
- [43] Mohammed S, Harikumar KB. Role of resveratrol in chemosensitization of cancer. In *Role of Nutraceuticals in Cancer Chemosensitization*. Elsevier; 2018:61–76.
- [44] Park MH, Hong JT. Roles of nf- κ b in cancer and inflammatory diseases and their therapeutic approaches. *Cells* 2016;5(2):15.
- [45] Yoon SO, Jeon TJ, Park JS, Ryu YH, Lee JH, Yoo JS, Kim JK, Yoon DS, Oh EJ. Analysis of the roles of glucose transporter 1 and hexokinase 2 in the metabolism of glucose by extrahepatic bile duct cancer cells. *Clin Nucl Med* 2015;40(3):e178–82.
- [46] Bao Y, Mukai K, Hishiki T, Kubo A, Ohmura M, Sugiura Y, Matsuura T, Nagahata Y, Hayakawa N, Yamamoto T, et al. Energy management by enhanced glycolysis in g1-phase in human colon cancer cells in vitro and in vivo. *Mol Cancer Res* 2013;11(9):973–85.
- [47] Adem S, Comakli V, Uzun N. Pyruvate kinase activators as a therapy target: a patent review 2011–2017. *Expert Opin Ther Patents* 2018;28(1):61–8.
- [48] Iqbal MA, Gupta V, Gopinath P, Mazurek S, Bamezai RN. Pyruvate kinase m2 and cancer: an updated assessment. *FEBS Lett* 2014;588(16):2685–92.
- [49] Luo W, Semenza GL. Emerging roles of pkm2 in cell metabolism and cancer progression. *Trends Endocrinol Metab* 2012;23(11):560–6.
- [50] Wang L, Zhang J, Cao Z, Wang Y, Gao Q, Zhang J, Wang D. Inhibition of oxidative phosphorylation for enhancing citric acid production by aspergillus niger. *Microb Cell Factories* 2015;14(1):7.
- [51] Patra KC, Hay N. The pentose phosphate pathway and cancer. *Trends Biochem Sci* 2014;39(8):347–54.
- [52] Kowalik MA, Columbano A, Perra A. Emerging role of the pentose phosphate pathway in hepatocellular carcinoma. *Front Oncol* 2017;7:87.
- [53] Cho ES, Cha YH, Kim HS, Kim NH, Yook JI. The pentose phosphate pathway as a potential target for cancer therapy. *Biomol Therap* 2018;26(1):29.
- [54] Kwon JH, Lee J, Kim J, Jo YH, Kirchner VA, Kim N, Kwak BJ, Hwang S, Song GW, Lee SG, et al. Hif-1 α regulates a2b adenosine receptor expression in liver cancer cells. *Exp Therap Med* 2019;18(6):4231–40.
- [55] Zhang X, Shi H, Tang H, Fang Z, Wang J, Cui S. mir-218 inhibits the invasion and migration of colon cancer cells by targeting the pi3k/akt/mtor signaling pathway. *Int J Mol Med* 2015;35(5):1301–8.
- [56] Mao Y, Xi L, Li Q, Cai Z, Lai Y, Zhang X, Yu C. Regulation of cell apoptosis and proliferation in pancreatic cancer through pi3k/akt pathway via polo-like kinase 1. *Oncol Rep* 2016;36(1):49–56.
- [57] Roy NK, Monisha J, Padmavathi G, Lalruaitluanga H, Kumar NS, Singh AK, Bordoloi D, Baruah MN, Ahmed GN, Longkumar I, et al. Isoform-specific role of akt in oral squamous cell carcinoma. *Biomedicines* 2019;9(7):253.
- [58] Amir H, Khan MA, Feroz S, Bibi N, Nawaz M, Mehmood A, Yousuf A, Khawaja MA, Khadim MT, Tariq A. Carlo-7—a plausible biomarker for bladder cancer. *Int J Exp Pathol* 2019;100(1):25–31.
- [59] Fang H, Liu HM, Wu WH, Liu H, Pan Y, Li WJ. Upregulation of long noncoding rna ccat1-l promotes epithelial–mesenchymal transition in gastric adenocarcinoma. *Oncotargets Ther* 2018;11:5647.
- [60] Zhang Y, Liao S, Fan W, Wei W, Wang C, Sun S. Tunicamycin-induced er stress regulates chemokine ccl5 expression and secretion via stat3 followed by decreased transmigration of mcf-7 breast cancer cells. *Oncol Rep* 2014;32(6):2769–76.
- [61] Wang Z, Si X, Xu A, Meng X, Gao S, Qi Y, Zhu L, Li T, Li W, Dong L. Activation of stat3 in human gastric cancer cells via interleukin (il)-6-type cytokine signaling correlates with clinical implications. *PLoS One* 2013;8(10):e75788.
- [62] Wang Y, Lu Z, Wang N, Zhang M, Zeng X, Zhao W. MicroRNA-1299 is a negative regulator of stat3 in colon cancer. *Oncol Rep* 2017;37(6):3227–34.
- [63] Mulrooney TJ, Marsh J, Urits I, Seyfried TN, Mukherjee P. Influence of caloric restriction on constitutive expression of nf- κ b in an experimental mouse astrocytoma. *PLoS One* 2011;6(3):e18085.
- [64] Pinweha P, Rattanapornsompong K, Charoensawan V, Jitrapakdee S. MicroRNAs and oncogenic transcriptional regulatory networks controlling metabolic reprogramming in cancers. *Comput Struct Biotechnol J* 2016;14:223–33.
- [65] Luo T, Li Y, Nie R, Liang C, Liu Z, Xue Z, Chen G, Jiang K, Liu ZX, Lin H, et al. Development and validation of metabolism-related gene signature in prognostic prediction of gastric cancer. *Comput Struct Biotechnol J* 2020.
- [66] Resendis-Antonio O, Checa A, Encarnación S. Modeling core metabolism in cancer cells: surveying the topology underlying the warburg effect. *PLoS One* 2010;5(8):e12383.
- [67] Hashemzadeh S, Shahmorad S, Rafii-Tabar H, Omid Y. Computational modeling to determine key regulators of hypoxia effects on the lactate production in the glycolysis pathway. *Scientific Rep* 2020;10(1):1–8.
- [68] Roy M, Finley SD. Computational model predicts the effects of targeting cellular metabolism in pancreatic cancer. *Front Physiol* 2017;8:217.
- [69] Capuani F, De Martino D, Marinarì E, De Martino A. Quantitative constraint-based computational model of tumor-to-stroma coupling via lactate shuttle. *Scientific Rep* 2015;5:11880.
- [70] Lillacci G, Khammash M. Parameter estimation and model selection in computational biology. *PLoS Comput Biol* 2010;6(3):e1000696.
- [71] Kinoshita A, Tsukada K, Soga T, Hishiki T, Ueno Y, Nakayama Y, Tomita M, Suematsu M. Roles of hemoglobin allostery in hypoxia-induced metabolic alterations in erythrocytes: simulation and its verification by metabolome analysis. *J Biol Chem* 2007.
- [72] Oliveira P, Alves M, Rato L, Laurentino S, Silva J, Sa R, Barros A, Sousa M, Carvalho R, Cavaco J, et al. Effect of insulin deprivation on metabolism and metabolism-associated gene transcript levels of in vitro cultured human sertoli cells. *Biochim Biophys Acta* 2012;1820(2):84–9.
- [73] Roncero SE. Three-time-scale nonlinear control of an autonomous helicopter on a platform [Ph.D. thesis]. Automation, Robotics and Telematic Engineering, Universidad de Sevilla; 2011.
- [74] Esteban S, Gordillo F, Aracil J. Three-time scale singular perturbation control and stability analysis for an autonomous helicopter on a platform. *Int J Robust Nonlinear Control* 2013;23(12):1360–92.
- [75] Goussis DA. Quasi steady state and partial equilibrium approximations: their relation and their validity. *Combust Theory Model* 2012;16(5):869–926.
- [76] Richard G, Chang H, Cizelj I, Belta C, Julius AA, Amar S. Integration of large-scale metabolic, signaling, and gene regulatory networks with application to infection responses. In: *Decision and Control and European Control Conference (CDC-ECC), 50th Conference. IEEE; 2011. p. 2227–32.*
- [77] van der Knaap JA, Verrijzer CP. Undercover: gene control by metabolites and metabolic enzymes. *Genes Devel.* 2016;30(21):2345–69.
- [78] Ray I, Dasgupta A, De RK. Succinate aggravates nafld progression to liver cancer on the onset of obesity: an in-silico model. *J Bioinf Comput Biol* 2018;16(4):1850008–1–15. doi:10.1142/S0219720018500087.
- [79] Drgoña J, Picard D, Kvasnica M, Helsen L. Approximate model predictive building control via machine learning. *Appl Energy* 2018;218:199–216.
- [80] Canale M, Fagiano L, Milanese M. Fast nonlinear model predictive control using set membership approximation? *IFAC Proc Vol* 2008;41(2):12165–70.
- [81] Hertneck M, Köhler J, Trimpe S, Allgöwer F. Learning an approximate model predictive controller with guarantees. *IEEE Control Syst Lett* 2018;2(3):543–8.
- [82] Hesami M, Naderi R, Tohidfar M, Yousefzadeh-Najafabadi M. Development of support vector machine-based model and comparative analysis with artificial neural network for modeling the plant tissue culture procedures: effect of

- plant growth regulators on somatic embryogenesis of chrysanthemum, as a case study. *Plant Methods* 2020;16(1):1–15.
- [83] Sari PA, Suhatri M, Osman N, Mu'azu M, Dehghani H, Sedghi Y, Safa M, Hasanipannah M, Wakil K, Khorami M, et al. An intelligent based-model role to simulate the factor of safe slope by support vector regression. *Eng Comput* 2019;35(4):1521–31.
- [84] Shen X, Gong X, Cai Y, Guo Y, Tu J, Li H, Zhang T, Wang J, Xue F, Zhu ZJ. Normalization and integration of large-scale metabolomics data using support vector regression. *Metabolomics* 2016;12(5):89.
- [85] Shirzad A, Tabesh M, Farmani R. A comparison between performance of support vector regression and artificial neural network in prediction of pipe burst rate in water distribution networks. *KSCE J Civil Eng* 2014;18(4):941–8.
- [86] Jain RK, Smith KM, Culligan PJ, Taylor JE. Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. *Appl Energy* 2014;123:168–78.
- [87] Dhiman HS, Deb D, Guerrero JM. Hybrid machine intelligent svr variants for wind forecasting and ramp events. *Renew Sustain Energy Rev* 2019;108:369–79.
- [88] Iplikci S. A support vector machine based control application to the experimental three-tank system. *ISA Trans* 2010;49(3):376–86.
- [89] Feng K, Lu J, Chen J. Nonlinear model predictive control based on support vector machine and genetic algorithm. *Chin J Chem Eng* 2015;23(12):2048–52.
- [90] Basak D, Pal S, Patranabis DC. Support vector regression. *Neural Inf Process-Lett Rev* 2007;11(10):203–24.
- [91] Gordon G, Tibshirani R. Karush-kuhn-tucker conditions. *Optimization* 2012;10(725/36):725.
- [92] Sivanandam S, Deepa S. Genetic algorithms. In: *Introduction to genetic algorithms*. Springer; 2008. p. 15–37.
- [93] Kumar S, Jain S, Sharma H. Genetic algorithms. In *advances in swarm intelligence for optimizing problems in computer science 2018*; pp. 27–52..
- [94] Spears WM, De Jong KD. On the virtues of parameterized uniform crossover. *Tech. Rep.*; Naval Research Lab Washington DC; 1995..
- [95] Mirjalili S. Genetic algorithm. In: *Evolutionary algorithms and neural networks*. Springer; 2019. p. 43–55.
- [96] Nelson DL, Lehninger AL, Cox MM. *Lehninger principles of biochemistry*. Macmillan; 2008.
- [97] Shlomi T, Cabili MN, Herrgård MJ, Palsson BØ, Rupp E. Network-based prediction of human tissue-specific metabolism. *Nat Biotechnol* 2008;26(9):1003.
- [98] Al Tameemi W, Dale TP, Al-Jumaily RMK, Forsyth NR. Hypoxia-modified cancer cell metabolism. *Front Cell Develop Biol* 2019;7:4..
- [99] Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. *Kegg: Kyoto encyclopedia of genes and genomes*. *Nucl Acids Res* 1999;27(1):29–34.
- [100] Olson AL, Pessin JE. Structure, function, and regulation of the mammalian facilitative glucose transporter gene family. *Annu Rev Nutr* 1996;16(1):235–56.
- [101] Pradelli L, Villa E, Zunino B, Marchetti S, Ricci J. Glucose metabolism is inhibited by caspases upon the induction of apoptosis. *Cell Death Dis* 2014;5(9). e1406.
- [102] Fadaka A, Ajiboye B, Ojo O, Adewale O, Olayide I, Emuowhochere R. Biology of glucose metabolism in cancer cells. *J Oncol Sci* 2017;3(2):45–51.
- [103] Sumi S, Umemura N, Adachi M, Ohta T, Naganawa K, Kawaki H, Takayama E, Kondoh N, Sumitomo S. The luminance ratio of autofluorescence in a xenograft mouse model is stable through tumor growth stages. *Clin Exp Dental Res* 2018;4(5):174–81.
- [104] Wang G, Xu Z, Wang C, Yao F, Li J, Chen C, Sun S. Differential phosphofructokinase-1 isoenzyme patterns associated with glycolytic efficiency in human breast cancer and paracancer tissues. *Oncol Lett* 2013;6(6):1701–6.
- [105] Li L, Wang Y, Wang Q, Qu J, Wei X, Xu J, Wang Y, Suo F, Zhang Y. High developmental pluripotency-associated 4 expression promotes cell proliferation and glycolysis, and predicts poor prognosis in non-small-cell lung cancer. *Mol Med Rep* 2019;20(1):445–54.
- [106] Izuishi K, Yamamoto Y, Mori H, Kameyama R, Fujihara S, Masaki T, Suzuki Y. Molecular mechanisms of [18f] fluorodeoxyglucose accumulation in liver cancer. *Oncol Rep* 2014;31(2):701–6.
- [107] Lim S, Liu H, Da Silva LM, Arora R, Liu Z, Phillips JB, Schmitt DC, Vu T, McClellan S, Lin Y, et al. Immunoregulatory protein b7-h3 reprograms glucose metabolism in cancer cells by ros-mediated stabilization of hif1 α . *Cancer Res* 2016;76(8):2231–42.
- [108] Dehne N, Brüne B. Hif-1 in the inflammatory microenvironment. *Exp Cell Res* 2009;315(11):1791–7.
- [109] Liberti MV, Locasale JW. The warburg effect: how does it benefit cancer cells?. *Trends Biochem Sci* 2016;41(3):211–8.
- [110] Wang J, Yi J. Cancer cell killing via ros: to increase or decrease, that is the question. *Cancer Biol Ther* 2008;7(12):1875–84.
- [111] Hong W, Cai P, Xu C, Cao D, Yu W, Zhao Z, Huang M, Jin J. Inhibition of glucose-6-phosphate dehydrogenase reverses cisplatin resistance in lung cancer cells via the redox system. *Front Pharmacol* 2018;9:43.
- [112] Zhang C, Zhang Z, Zhu Y, Qin S. Glucose-6-phosphate dehydrogenase: a biomarker and potential therapeutic target for cancer. *Anti-Cancer Agents Med Chem* 2014;14(2):280–9.
- [113] Benito A, Polat IH, Noé V, Ciudad CJ, Marin S, Cascante M. Glucose-6-phosphate dehydrogenase and transketolase modulate breast cancer cell metabolic reprogramming and correlate with poor patient outcome. *Oncotarget* 2017;8(63). 106693.
- [114] Tseng CW, Kuo WH, Chan SH, Chan HL, Chang KJ, Wang LH. Transketolase regulates the metabolic switch to control breast cancer cell metastasis via the α -ketoglutarate signaling pathway. *Cancer Res* 2018;78(11):2799–812.
- [115] Xu LM, Lai RKH, Lin SH, Tse APW, Chiu DKC, Koh HY, Law CT, Wong CM, Cai Z, Wong CCL, et al. Transketolase counteracts oxidative stress to drive cancer development. *Proc Natl Acad Sci* 2016;113(6):E725–34.
- [116] Diaz-Moralli S, Aguilar E, Marin S, Coy JF, Dowerchin M, Antoniewicz MR, Meca-Cortés O, Notebaert L, Ghesquière B, Eelen G, et al. A key role for transketolase-like 1 in tumor metabolic reprogramming. *Oncotarget* 2016;7(32):51875.
- [117] Chou YT, Jiang JK, Yang MH, Lu JW, Lin HK, Wang HD, Yuh CH. Identification of a noncanonical function for ribose-5-phosphate isomerase a promotes colorectal cancer formation by stabilizing and activating β -catenin via a novel c-terminal domain. *PLoS Biol* 2018;16(1). e2003714.
- [118] De Padua MC, Delodi G, Vučetić M, Durivault J, Vial V, Bayer P, Noleto GR, Mazure NM, Ždravlević M, Pouységur J. Disrupting glucose-6-phosphate isomerase fully suppresses the “warburg effect” and activates oxphos with minimal impact on tumor growth except in hypoxia. *Oncotarget* 2017;8(50):87623.