Modeling a Complex Biological Network with Temporal Heterogeneity: Cardiac Myocyte Plasticity as a Case Study

Amin R. Mazloom[1], Kalyan Basu[2], Subhrangsu S. Mandal[3], and Sajal K. Das[2]

[1] Department of Pharmacology & Systems Therapeutics
Mount Sinai School of Medicine, New York, NY 10029, USA
amin.mazloom@mssm.edu
[2] Department of Computer Science & Engineering
[3] Department of Chemistry & Biochemistry
University of Texas at Arlington, Arlington, Texas, TX 76019, USA

Abstract. Complex biological systems often characterize nonlinear dynamics. Employing traditional deterministic or stochastic approaches to quantify these dynamics either fail to capture their existing deviant effects or lead to combinatorial explosion. In this work we devised a novel approach that projects the biological functions within a pathway to a network of stochastic events that are random in time and space. By applying this approach recursively to the object system we build the event network of the entire system. The dynamics of the system evolves through the execution of the event network by a simulation engine which comprised of a time prioritized event queue. As a case study we utilized the current method and conducted an in-silico experiment on the metabolic plasticity of a cardiac myocyete. We aimed to quantify the down stream effects of insulin signaling that predominantly controls the plasticity in myocardium. Intriguingly, our in-silico results on transcription regulatory effect of insulin showed a good agreement with experimental data. Meanwhile we were able to characterize the flux change across major metabolic pathways over 48 hours of the in-silico experiment. Our simulation performed a remarkable efficiency by conducting 48 hours of simulation-time in less that 2 hours of processor time.

1 Introduction

A complex system is a subset of a world comprised of many components whose interactions with the rest of the world or another subset is properly defined. The behavior of a complex system could properly perceived through the aggregate effects of its components. An organism could be viewed as a system or collection of systems at different hierarchical levels. The boundary of a system is defined by its components which for a bio-system could range from organism to organ, tissue, cell, molecule, and atom. The degree of complexity between levels grows exponentially from top to bottom. In the current study a cell draws the system boundary and molecules are the interacting components of this system. The

J. Zhou (Ed.): Complex 2009, Part I, LNICST 4, pp. 467–486, 2009.

respond of a cell to an exogenous signal (antigens, hormones, pressure or temperature change, etc) shaped from endogenous activities within the cellular networks that attempt to maintain the cell homeostasis. To gain insight on the dynamics of this respond at the system level, interaction of the underlying components must be properly characterized in time and space. In this study we propose a novel stochastic discrete event-based methodology [1] to conduct system level in-silico experiments on a typical eukaryotic cell. We elucidate our method by deploying that on the cardiac myocyte along with insulin signal as a case study to validate the significance of our approach. The idea is to learn the system level effects of an exogenous signal through the changes imposed on the dynamics of Signal Transduction Network, (STN), Transcription Regulatory Network (TRN), and Metabolic Network (MTN) of the cell following the perception of the signal. The rest of the paper is organized as follows: we allotted the rest of this section to provide a background on the quantitative models proposed for the heart cell, on section two the modeling steps in the proposed approach along with the simulation algorithm is described, section three devoted to application of the current schema on the insulin signaling pathway in the heart-cell, section four provides the results for the in-silico experiment, section five discusses the key pros and cons of the approach, and we end the paper by drawing conclusion in section six.

1.1 Background

Since 1960 that Denis Noble proposed the first model of the heart, numerous quantitative models have been proposed for the heart from organ to cell level. These models could broadly be classified into two main families: physiological models and pathway models. The former class includes a combination of mechanical and biochemical models that focuss on capturing the physiological and electrophysiological dynamics of heart and its tissues under different physiological and biochemical conditions. Models in [2,3] are typical examples of this class of models. This top-down modeling approach offers a course grain analysis, therefore is not suitable for detail analysis of intra-cellular networks. The latter class has a more microscopic focus which intends to model one or more pathways from the cellular networks and seeks to quantify their dynamics, discover new pathways, complexes, etc. Instances of such models could be found in [4,5]. In order to predict the dynamics of biological functions latter class subscribes to either one of the following approaches: i) deterministic approach where a sets of ordinary differential equations (ODE) or partial differential equations (PDE) is formed based on biochemical reactions and diffusion to address the rate of change in concentrations of molecular parts. These ODEs/PDEs are then numerically solved to determine the dynamics of the underlaying system. ii) The stochastic approach which comprises strains of Gillespie [6] algorithm to approximate Chemical Master Equation (CME) [7], where the system is mapped into sets of chemical kinetic equations which evolves in Monte Carlo steps. Arkin and Samiolov [8] have shown that non-classical behavior of biological networks cause their dynamics to substantially diverge from their average. Therefore, deterministic

approaches based on classical chemical kinetic (CCK) which assumes equilibrium across the entire course of system's evolution would not be an appropriate method to model many of biological systems. This claim remains valid even for the systems with higher molecular abundance. Although Gillespie based family of algorithms are suitable for capturing the behavior of biological functions; however, despite the efficiency enhancements they archived owing to approximation techniques such as *tau-leaping*, they still suffer from high computational complexity. None of these approaches promises a suitable model for a system that comprises network of biological functions (e.g. transcription function, signaling phospo-interaction, metabolic reaction) that manifest several order of magnitude difference in their temporal dynamics. In a system whose components (i.e. network nodes) manifest such temporal heterogeneity sequential evolution of fast processes would exhaust the execution of slower ones. The complex network of cellular processes ($\bigcup\{$STN, TRN, MTN$\}$) in a cardiac myocyte is a typical example of above systems.

To layout a modeling frame-work that contend to such heterogeneity in an in-silico experiment we introduce the concept of myocardial event (*myvent*) that accounts for an individual biological process within a cardiac myocyte. Noting that dynamics of the system is captured through changes in the count of molecular parts in the course of an in-silico experiment. These dynamics evolves through the execution of a network of *myevents*. Each *myevent* is an object from a specific class of *myevents* that has a random execution time with known first and second moments. The ability to bundle one or more *myevents* of a same class grants a *mesoscopic scale* to this modeling approach. This property avoids exhaustive computations for a system that comprises a network of processes that are temporally heterogeneous.

2 Approach

Observations confirm that at the molecular level the cellular behavior arises from the stochastic interaction between molecular parts [8,9]. Such observations is the key motivation in applying stochastic discrete event-based (SDE) method in capturing the dynamics of a cellular function. Hence, identifying molecular functions in a cardiac myocyte and mapping those into sets of *myvents* is fundamental to our approach. Each *myevent* has three attributes: (i) The stochastic physicochemical model that approximates the temporal dynamics for a typical class of myevents (i.e. cytoplasmic reaction, membrane reaction, transcription, etc.), (ii) The molecular resources (input/output) associated with a *myevent*, (iii) The compartment(s) within or across which a *myevent* is executed. In the current study for the first attribute of the a *myevent* we either adopt a physicochemical model from the literature or replace that with a relevant probability distribution that can approximate the experimental data. Also to avoid further complexity we only consider cytoplasm, nucleus, and mitochondria compartments for a cardiac myocyte. In a SDE in-silico experiment, *simulation time* is the representation of the *physical time* of the system being modeled. Each event

is associated a *time-stamp* indicating when that event occurs in the physical system being simulated. The event *time-stamp* is computed from the knowledge of the previous event that has triggered the current event, together with the event execution-time which is a realization of the random number that characterizes the event dynamics. The dynamics of resource utilizations with progression in time unveil the complete internal picture of a complex biological system at the molecular level. Applying this doctrine to study the system level dynamics in a cardiac myocyte, demands the following check-list for characterizing the system parameters: (i) Identify the list of discrete *myevents* that can be included in the model based on the available knowledge of the system; (ii) Identify the resources of interest for the execution of the *myevents* function which are being used by the biological process for each discrete event; (iii) Compute the time taken to complete this biological discrete event. For this purpose, it is important to mathematically relate all event parameters which affect the interaction of the resources in a particular biological function; (iv) Identify the next *myevent* or set of *myevents* initiated on the completion of a *myevent*. If multiple discrete *myevents* are possible after completion of a *myevent*, the next *myevent* is chosen probabilistically, based on the biological pathway of the function being modeled. This probability calculation depends on the *myevent-set* and the properties of the *myevents* within the set.

Once the above check list is satisfied the discrete event simulator scheduler which is a time prioritized event queue pops individual *myevents* from the queue and system proceeds. Upon the execution of each *myevent*, molecular resources of the system is updated, system time is moved forward, and new *myevents* are pushed into the event queue from the *next-myevents* list of current *myevent*. The pseudo code for the algorithm that governs a SDE in-silico experiment is given in Fig. 1.(left) and the simulation engine architecture is depicted in Fig. 1.(right).

```
SDES Algorithm for Cardiac myocyte :-
1.begin
2. Initialize:Random number generators
, global parameters,
myevent parameters,compartments,
simulator clock, trigger event
3.  while the event queue is not empty
4.  begin
5.    pop the myevent from head of queue
6.    update the simulation time
7.    Invoke the model
8.    retrieve the input resources
9.    if all the input resources are available
10.   begin
11.     Utilize the random number generator
to determine event times
for the next sets of events
12.     update molecular resources
13      generate and push next set of
myevents to the event queue
14.   end
15. end
16.end.
```

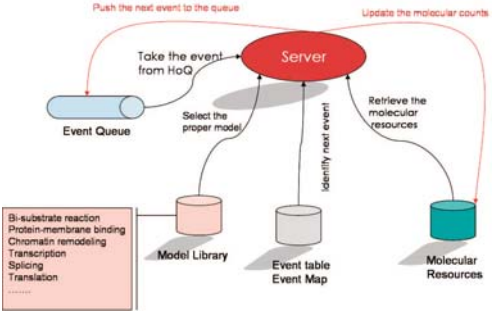

Fig. 1. (Left) The SDES algorithm pseudo code for cardiac myocyte model; (Right) The architecture of simulator engine

In a compartmentalized environment there will be two superclass of *myevents* apart from their types: the *local myevent* ($e_{ij}, \quad i = j$) and *cross compartment myevent* ($e_{ij}, \quad i \neq j$). Subscripts i and j are the source and destination compartments across which the event is executed. Execution of a local *myevent* only effects molecular resources in the compartment local to that *myevent*, where a cross compartment *myevent* potentially changes the resources in both source and destination compartments.

3 Modeling the Cardiac Myocyte Plasticity

Metabolic plasticity is the capacity of a cell to adopt to alternative available metabolic substrates as the source for its energy requirement. In order to model the plasticity of a cardiac myocyte at the system level first we need the to identify the pathways in signal transduction, transcription regulatory, and metabolic networks that pertain to such functionality. Further, identify the *myevents* that comprise each biological function within these pathways. Then associate each *myevent* with the proper stochastic model, input resources, and output resources. Subsequently interconnect those *myevents* in a recursive fashion to form the *myevent* network of that function. The above process that maps a biological functions from its physiological context to an event based context is referred to as *eventology* of that function. By recursively applying the *eventology* to the system we can form the event network of the whole system.

3.1 Eventology of the Signaling Pathways

Cardiac myocytes should have flexibility in their fuel selection in order to be consistent in meeting their energy requirements. Metabolic flux modulation could be regulated at many levels, two of the promising flux modulations in cardiac myocytes are through the control of metabolite uptake and gene expression level [10]. Insulin which is an essential peptide hormone of endocrine system that secretes from β-cells in pancreas is predominantly involved in the fuel selection at both levels. Although the propagation of the insulin signal within the cell influences divers cellular functions such as mitogenic, cell growth, etc.; however, in this work we focus on the signaling information that culminates on the two modulatory effects.

The insulin signal is sensed by binding the insulin to insulin receptors *(INSR)* on the membrane of cardiac myocytes and belong to the family of ligand-activated tyrosine kinase (RTK) receptors [11]. The information of the insulin signal is propagated within the cell through a non-linear signaling network [12]. The inherent robustness is the *de-facto* rule of survival in the evolutionary process of biological systems. Therefore, most of these systems are robust to the large set of stresses and demonstrate the butterfly effect to substantially smaller sets. Setting this fact *vis-a-vis* the complexity of system enables us to reduce the complexity by two strategies: i) by eliminating the components or aggregating their detail to a higher level where it is proven or speculated to have lesser impact on the objective system, ii) exclude a subset of the system from the analysis

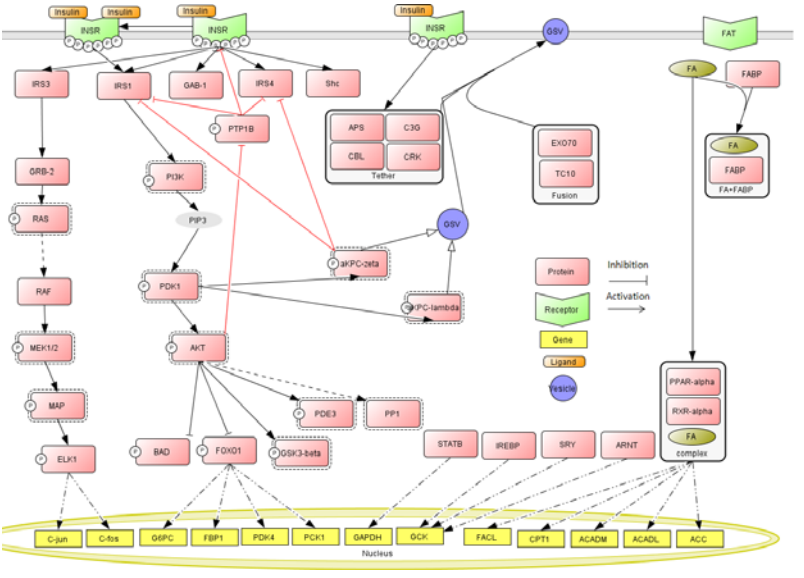

Fig. 2. The insulin (left) and fatty acid (right) signal transduction networks diagram

with the assumption that the rest of the system is in the equilibrium interaction with the current subset. With this strategy we have abstracted the insulin signal transduction hierarchy from excessive details and included those components where a consensus exists on their impact on the cell metabolism [10].

The insulin signal transduction network (STN) that has the above property could be found in *KEGG pathway database* [13]. We imported their STN and modified the original version based on data published elsewhere to include some of missing components that were necessary for our work as well as excluded the excessive details. Fig. 2 shows the signal transduction networks for insulin and fatty acid that we used in our in-silico experiment. The *myevent* diagram for the insulin signaling pathway of Fig. 2 is depicted in Fig. 3. The color code is used to represent the *myevents* with similar physicochemical (e.g phosphorylation, activation, transport) class. The physicochemical class of each *myevent* was explored from literature. A *myevent* whose physicochemical class was unidentified was assigned to a biochemical reaction class. Noting that, since signal transduction and transcription regulatory networks are interrelated we included a subset TRN that is affected by the insulin in the event diagram. Noting that Fig. 3 does not include the events that pertain to the fatty acid signal which is partially depicted in right corner of Fig. 2. In the *myevent* legend the three capital letters following the name of each *myevent* specifies the class of that *myevent* (e.g. TRN: transcription, ACT: activation, PPR: phosphorylation, INA: insulin receptor activation). A self feedback in the event digram induces the signal prorogation by one fold from the feedback point. These loops are added to the map empirically by comparing the in-silico results and experimental data.

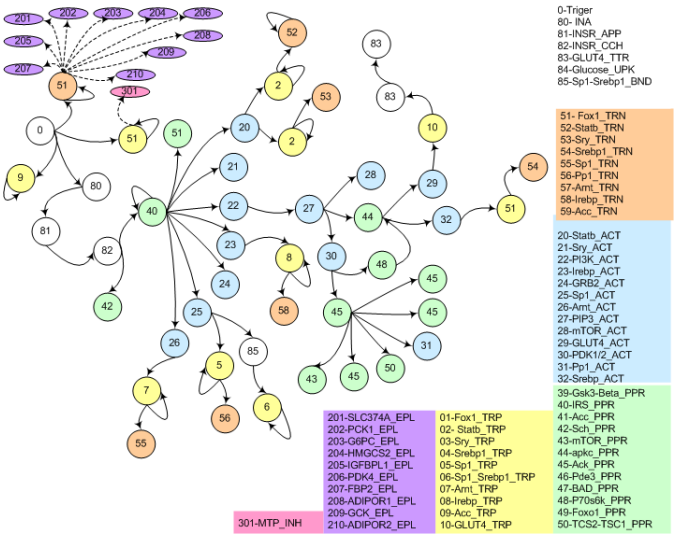

Fig. 3. The *myevent* diagram above depicts selected events for insulin STN in Fig. 2. Events with purple color belong to TRN.

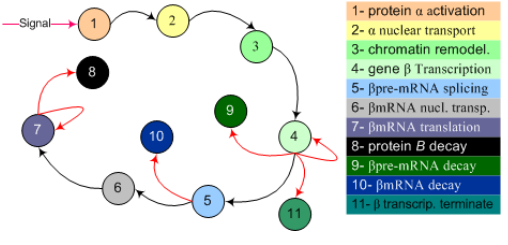

Fig. 4. Event diagram of protein synthesis in eukaryotes

3.2 Models for Glucose and Fatty Acid Uptake

Glucose and Fatty Acids comprise > %90 of the energy resources of cardiac tissues [14]. Hence, in this work we decided to focus on modeling uptake pathways of these two substrates. In adult heart Glucose is taken into the cell mainly by glucose transporter 4 (Glut4) [15]. Insulin promotes the Glut4 membrane transport through two parallel pathways. These two pathways which both originate from the insulin receptor protein (*INSR*) activation complement each others role in mediating the glucose uptake. Phosphorylated $aPKC\lambda/\xi$ is a down stream product of a phosphorylation signaling cascade from the first pathway which enables the Glut4 vesicle transporters (GSV) to move to the vicinity of the membrane [16]. Upon activation *INSR* phospho-activates the *APS* protein that initiates the second pathway. Activation of *APS* initiates a sequence of activations and interactions that involve *m*ore than seven proteins [17]. Through a sequence of complex interactions the Glut4 which at the time is present in the vicinity of

plasma membrane is first docked then tethered and ultimately fussed into the membrane [16]. In our in-silico experiment we abstract the outlined process into following three *myevents*: GSV activation *myevent* (associated with reaction model in [18] discussed later in the paper), GSV transport *myevent* (associated with diffusion model in [19], and Glut4 tether *myevents* (associated with in model [20]). Fig. 5.(a) shows the discrete even representation of the glucose uptake. We made a subtle modification to the model in [20] which mainly includes adding a capacity to each membrane receptor to handle the group transport activity for GSV. For the Fatty Acid (FA) uptake we focused to model the mechanism for which a strong consensus exists and tried to implement that for a long chain fatty acid (LCFA). Note that choosing a fatty acid with different chain length (e.g. short or medium) mainly affects the oxidation reactions and not the uptake mechanism. The plasma isoforms of FATBP and FAT/36 can participate in passive diffusion by increasing the dissociation rate of albumin, and in facilitated transport by interacting with FATP and importing the FA into cytoplasm [21]. This system adjusts the rate of FA uptake with mitochondrial demand to avoid accumulation of FA in cytoplasm which could be hazardous for the cell. Once the FA entered the cytoplasm, it then binds to cytoplasmic isoform of FATBP and transported to the vicinity of the mitochondrial outer membrane. *Acyl-CoA* Synthase (ACS) converts the Long Chain Fatty Acid (LCFA) to *LC_acyl-CoA*. To participate in the β-Fatty Acid oxidation *LC_acyl-CoA* should be transported into the mitochondria. To cross the impermeable mitochondria membrane the fatty acid transport pathway utilizes the *Carnitine palymitoyltransferase (CPT)* system. CPT composed of $L-$carnitine, *acylcarnitnie translocase* (ACT) and two transfer proteins i.e. *CPT1* and *CPT2* [22]. *Carnitine palymitoyltransferase 1* is a transmembrane protein located on the outer membrane of mitochondria and delivers the *LC_acyl-CoA* to carnitine to form *LC_acylcarnitine*. ACT hands the *LC_acyl-CoA* over to CPT2 through the intermembrane space. The second transfer protein replaces the carnitine group of *LC_acylcarnitine* with *CoA* and releases the *LC_acyl-CoA* in the mitochondria to participate in the β-fatty acid oxidation pathway [22]. *CPT1* is sensitive to *Malonyl_CoA* which is the product of *Acetly_CoA* carboxylation in cytoplasm this reaction is catalyzed by *Acetyl_CoA Carboxylase (ACC)*. Hence *Malonyl_CoA* is a negative regulator of β-fatty acid oxidation. The event-based model of the FA uptake depicted in Fig. 5.(b). The associated model for the designated *myevents* are as follows: FA uptake associated with model in [20], FA transport and FA mitochondrial transport both associated with fast reaction model (described in supplementary materials[1]). In the FA mitochondrial transport we have modeled the process by breaking the transport between the metabolic and signaling networks. More specifically the binding FA to CPT1 is handled by the signaling network as one bimolecular reaction [18]. Shuttling *LC_acylcarnitine* to the CPT2 is handled by a metabolic reaction which is catalyzed by *CPT2*. The reason for breaking the event between metabolic and signaling network originates from the set of

[1] Supplementary section not included due to the space limitations and is available upon request from the corresponding author.

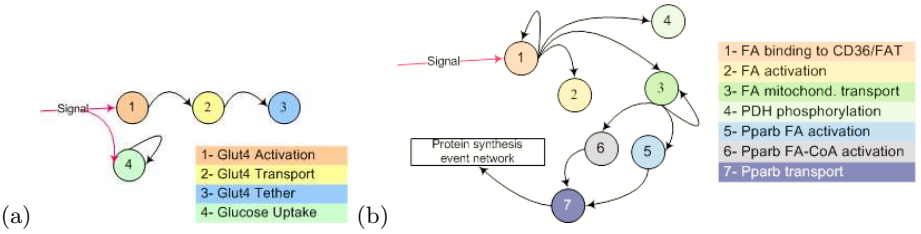

Fig. 5. (a) The *myevent* diagram for glucose uptake process. (b) Fatty acid uptake *myevent* diagram.

metabolic reactions that we used to model the metabolic reaction network in cardiac myocytes.

3.3 Eventology of Protein Synthesis

Down stream effect of a signal might effect the gene expression and consequently the protein synthesis. Protein synthesis is the core process of life which involves a very complex and not completely known regulatory mechanism. Although, we are still far behind from a comprehensive and detailed quantitative model of protein synthesis; however, our knowledge of *central dogma* is just enough to propose an event based abstraction that meets the requirements to fit into the in-silico experiment paradigm.

Protein synthesis in eukaryotes compromises an orchestrated sequence of events including: chromatin remodeling, gene transcription, pre-mRNA splicing, mRNA nuclear transport and mRNA translation. These events involve sophisticated evolution and regulatory mechanisms whose detail discussion is beyond the scope this paper. Hence, we briefly browse through the major concepts that will contribute to our modeling effort. Transcription and translation in eukaryotes is very complex and much of their details yet not properly understood. General mechanism of transcription and translation discussed in [23] and elsewhere. In [24] general concepts involved in mammalian gene transcription is described and a qualitative model for their assembly is proposed. Transcription and RNA II-TFIIB are structurally analyzed in [25] and mechanism of RNA II elongation is discussed in [26]. Binding of TATA Box Proteins (TBP) is essential for gene expression, in [27] regulation of gene expression by TBP is elucidated. Kinetic analysis of gene transcription is provided in [28]. Following the gene expression the pre-mRNA will be spliced to generate messenger mRNA. Each mRNA should be transported to cytoplasm and translated by the ribosomal proteins (tRNA) to give birth to the protein it encodes. The process of transporting the mRNA to cytoplasm is referred to as nuclear transport which it self has divers and complex mechanisms [29]. Also the kinetics of mRNA nuclear transport is studied in [30]. Following the export of mRNA to the cytoplasm ribosomal protein (tRNA) translates the codons in mRNA to the proper amino acids. The mechanism of translation initiation is given in [31] while the molecular mechanism of

translation is described in [32]. Also it has been shown that protein synthesis is non-linear and has a bursty dynamic [33].

Nevertheless, the details of gene expression is far more complicated than described above, we have abstracted the protein synthesis process as a network of *myevents*. These *myevents* could be categorized into two classes of *explicit* and *implicit* events based on the mechanism of their initiation. Former, includes those *myevents* whose trigger is explicitly indicated in the qualitative models such as transcription event, splicing event, etc. The latter class includes those *myevents* that will be executed although they are not explicitly included in the qualitative models, examples of those include: protein decay, mRNA decay, transcription termination, etc.

The protein synthesis is the product of collaborative effort between the transcription regulatory and signaling networks. Therefore, suppose an external signal in its downstream activates the transcription factor α, upon activation α is transported into the nucleus. Further assume that as a result gene β is affected. The effect of α on gene expression is interpreted by the gene regulatory network, our abstracted mechanism of gene regulation will be discussed shortly. In the case of positive regulation, $\beta - mRNA$ is produced, transported to cytoplasm, and translated to protein B. Fig. 4 shows the event diagram of this model where the red arrows point to the *implicit events* and black arrows to the *explicit events*. As observed in the diagram a *myevent* could belong to both categories and the only difference is the mechanism for triggering an event with respect to the qualitative model.

For chromatin remodeling we assumed to have SWI/SNF remodeling complex, since it is the most preserved remodeling complex across eukaryotes and has no sequence specificity [34]. Researchers in [35] and elsewhere have reported the data on different aspects of chromatin remodeling. We were able to fit their reported data (result not shown) for the rate of nucleosome *in-cis* translocation (base-pair/sec) into a gamma distribution ($\alpha = 2.50 \pm 0.17, \beta = 4.67 \pm 0.35$). Therefore, temporal dynamics of the remodeling *myevent* is modeled with gamma distribution. Also we assumed a nucleosome occupancy of 0.3 for all promoter regions. For the transcription event we used the model proposed in [36]. This model uses a birth and death Markov chain to determine the rate of the transcript production. They have modeled the process based on number of RNA PII that binds to the gene and the elongation rate of RNA PII. We have adapted and calibrated the model to become consistent for eukaryotic based on the parameters given in [37,38] (e.g. basal RNA PII elongation rate (40 bases/s), etc.). For the splicing *myevent* we assume to have a *constitutive* splicing [39] where each a pre-mRNA spliced at a rate of 0.25 per minute [30] which is negative exponentially distributed around the mean. For the pre-mRNA and mRNA decay *myevents* we applied the exponential decay processes with a rates according to to half life of these species reported in [40,33]. We used a simple stochastic diffusion model proposed in [19] to estimate temporal dynamics of mRNA nuclear transport based on kinetics reported in [41]. For estimation of translation *myevents* time, we applied the markov model proposed in [36] for translation in prokaryotes

and calibrate the parameters based on experimental data in [38,30]. The protein decay *myevent* has an exponential decay process with rate reported in [42]. The transcription termination event has a constant time which we obtained empirically while calibrating the simulator.

3.4 Transcription Regulatory Network and In-Silico Regulatory Model

More than 150 genes have been identified that are positively or negatively regulated by the insulin [43]. Amongst genes affected by insulin, < 50 genes reported as myocardial genes [44]. We sought to collect as many genes that has been reported and is regulated by either insulin or fatty acid signaling pathways in the heart muscle cell [12,43]. To abstract the expression and inhibition of the target gene *'X'*, we attribute each gene with a status flag and a time stamp. The status flag can hold one of the following three states: being expressed(BE), already inhibited (AI), or no activity (NA). The time stamp indicates the time for the last change in the status flag of the gene. Transition of the gene status from NA to either BE or AI is triggered by the transcription *myevents*. To handle the transition form BE or AI to NA a specific Gene Status Check (GSC) event is predicted that is executed periodically and compare the target gene time stamp with current time. If the difference between the two times is greater that a *GENE_HOLD_STATUS* constant then it shifts the gene status to NA. Based on current model there is no direct shift between BE and AI states.

The input to the a transcription *myevent* is a transcription factor *'T'*. Execution of a transcription event indicates that resource for *'T'* is available. The non-empty set g includes all the genes that are up/down regulated by transcription factor *'T'*, upon execution of a transcription event one of these genes is selected for the status change with probability $p = \frac{1}{|g|}$. Based on wether the selected gene *'X'* belongs to up-regulated or down-regulated subset of g, its flag is changed accordingly. The set of transcription factors (TFs) that we included in our in-silico along with their target genes is available in supplementary material.

3.5 Metabolic Reaction Model

Glycolysis I, TCA cycle, pyruvate metabolism, and β-fatty acid oxidation pathways are the major pathways dedicated to precursor substrates metabolism in cardiac myocytes. In our experiment we considered the set of metabolic reactions that comprises the above metabolic. This set composed of 109 reactions consistent with human metabolic reactions reported in BiGG database [45]. Each reaction is identified by a unique reaction ID that we borrowed from the original record in the BiGG (list of these reactions is given in supplementary materials).

For a metabolic reaction *myevent* we consider a *lumped metabolic event* whose effects on metabolites is based on the Flux Balance Analysis (FBA) approach [46]. Implementing such strategy requires a metabolic *myevent* to be local. However, keeping metabolic *myevents* local will cause metabolite explosion in some compartment (e.g. mitochondria) and metabolite starvation in the others

(e.g. cytosol). To circumvent this issue we define a new cross compartment event called *metabolite squad myevent (MetabSquad)* that executed regularly every τ *squad* unit of time and redistributes the metabolite across pairs of *neighbor compartments*. $\Omega_k(i,j) \leq 1, \quad \forall i,j : i \neq j$ is the portion of metabolite k molecules in compartment i to be transported to neighbor compartment j. This ratio is estimated in an iterative fashion. A pair of cellular compartments that can have direct molecular transport between themselves are called *neighbor compartments*.

We employed FBA approach to determine the flux across the metabolic reactions. From a reaction flux we can determine the change in molecular-count of a specific metabolite in the entire set of metabolic reactions in an arbitrary epoch during experiment, given the steady state condition. The essence of the FBA for a metabolic reaction founded on the assumption that the cell tends to maximize the biomass yield in the steady state condition. The emerging problem is then mapped into a linear optimization problem where the solution to this problem are optimum fluxes across sets of metabolic reactions given: the reactants, products, and enzymes concentrations. The method to manipulate the flux for a reaction across each metabolic *myevent* inter-arrival time briefly includes following steps: (i) determine active reactions from the availability of their participant molecular parts, (ii) determine the reaction direction by comparing the equilibrium constant of the reaction to the ratio of $\sum[procuts]$ to $\sum[reactants]$, where brackets in brackets indicate the concentration, (iii) determine the weight of a reaction with respect to all set of reactions (weight of a reaction is inversely proportional to the number of reactions in which its reactants participate), (iv) determine the reaction flux during time t_{mtb} with respect to the enzyme turnover number and metabolite constrains. t_{mtb} is the inter-arrival time between two metabolic *myevetns* which could be set to an arbitrary constant value. The dilemma for setting t_{mtb} value is choosing between efficiency and the precision of the simulation (i.e. large versus short periods).

Noting that any *myevent* that appeared in the event networks and was not discussed individually was modeled using stochastic reaction model proposed in [18]. The model equation for the reaction between reactants A and B is replicated here:

$$p(t_{A\cdot B}) = e^{(-\lambda t_{A\cdot B})}, \quad \lambda^{-1} = \frac{n_A (r_A + r_B)^2}{V} \sqrt{\frac{8\pi k_B T (m_A + m_B)}{m_A \cdot m_B}} \exp\left(\frac{-E_{Act}}{k_B T}\right) \quad (1)$$

In the above equation λ is reaction rate, n_A is molecular count of reactant A, r_x and m_x are the average molecular radius and molecular mass of reactant x, respectively, k_B is the Boltzmann constant, T is the absolute temperature and E_{Act} is the activation energy of the reaction. Noting that, this model subjects to further approximations given in supplementary materials.

4 In-Silico Results

The traces of plasticity is also observable in the expression profile of those genes contributing to a specific substrate metabolism. On the other hand, metabolism

of an abundant metabolite subjects to the promising availability of the transport proteins and metabolic enzymes specific to that metabolite. Hence, a higher gene expression profile is expected for the underlaying genes. Van Bilsen and his colleagues [47] conducted an experiment for the rat heart and identified the expression patterns for some of the genes contributing to the glucose and fatty acid oxidations in the rat heart.

In such experiment the cardiac myocytes were forced to follow a certain pattern in substrate (glucose and fatty acid) metabolism. The pattern imposed by feeding the model animals with glucose rich food for 8 hours (*feeding period*) and then letting them starve for the next 40 hours(*fasting period*). The starvation forces the body to release the fat stored in adipocytes into the blood. This, would let the other cells (e.g. cardiac myocyte) to uptake and oxidize the fatty acids for their functions, which obligates activating the fatty acids-dependent uptake and oxidation pathways. To validate our approach we utilized the proposed methodology to conduct the above experiment in-silico at the molecular level. To date of this paper no in-silico simulation tool or quantitative model has been reported to have the capacity of capturing the system-level dynamics of a cellular network for such a pro-long duration (i.e 48 hours).

To design the experiment we supplied 1.4 nM of each signaling proteins, 1.4 nM of each metabolic enzymes, and a basal level of transcript for each of the genes listed in supplementary materials. Also 11 mM of the Octadecaontate_(n-18:0) which is a saturated stearic fatty acid was supplied as the exogenous fatty acid resource. To mimic the short feeding period followed by a longer fasting period we supplied the initial concentration of the glucose such that it would last for ~ 8 hours, where fatty acid concentration would last for entire course of experiment. Hence, for the 40 hours following the initial 8 hours of experiment only fatty acid would be available as the metabolic substrate. Noting that we suspended the insulin signal once the glucose supply reached %5 of its initial concentration. The choice of the stearic or palmitic fatty acid would not skew the results since the stearic acid is converted to a palmitic acid by metabolic reaction *R_FAOXC180* which is an oxidation-reduction reaction in β-fatty acid oxidation pathway.

The fold change in concentration of the transcripts for those genes whose data could be validated with published data is depicted in Fig. 6. Comparison is shown at two time points for the feeding scenario discussed earlier between the in-silico and empirical results. As observed the *CPT1* which is a member of CPTS increased during the fasting and *ACADL* which is Long-chain specific acyl-CoA dehydrogenase was also induced during that period. The in-silico results shows that HK2 (hexokinase 2) which is a glycolysis pathway metabolic enzyme remained constant during fasting where the empirical data suggested reduction by half fold for the same period. This may suggests a potential inhibitory regulation which is not included in our simulation. Although both results agree on the increase for Fatty acid-binding protein (*FABP*) during the fasting period; however, in-silico results show a significantly higher fold which demands for further regulatory mechanism not implement by our gene regulatory model.

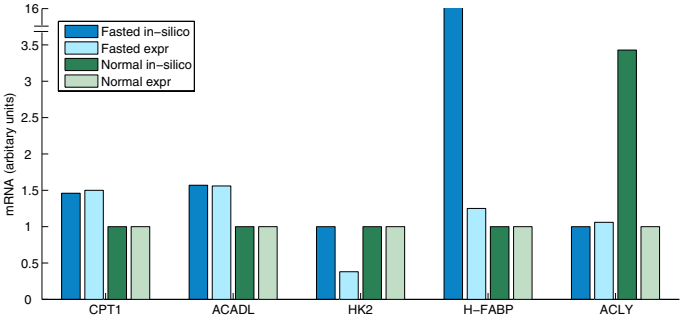

Fig. 6. Change in the expression profile of selected myocardial genes for *normal* feeding period (after 8 hours) and *fasting* period after 48 hours. In-silico results and empirical data are shown in blues and greens, respectively.

This proposition stays valid for ATP-citrate synthase (*ACLY*) too, but this time during the feeding period.

In Fig. 7.(a) we have shown the transcription regulatory effect of current feeding scenario at four time points for the entire set of genes. The genes that induced by *FOXO1* show exponential increase in their expression profile after 5 hours of simulation. This happened because the insulin signal which would negatively regulate those genes gradually diminished. Although the increase in the expression level of these genes was expected; however, the reported quantities in their expression profile subject to further validation with empirical data. Many of the genes involved in fatty acid transport and oxidation pathway show a one to two folds increase which is in agreement with the experimental data reported elsewhere. Since very limited data was available on negative regulatory effects of current transcription factors we were not successful to capture their negative regulatory effects on the gene expression profiles.

To further observe the metabolic plasticity of a myocardial cell we also looked into the metabolic fluxes, ATP synthesis and some substrates concentration profile during the course of experiment. Fig. 7.(b) shows that during the early hours experiment both exogenous substrates were highly utilized in energy production of the cell, as a result *ATP* concentration increased exponentially. The concentration of *D-Glucose-6P* follows an exponential decay which indicates a very high utilization of glucose in cell. After initial raise in the concentration of intermediate metabolites, for the hours between 6 to 20 we observe a decline in the slope of *Stearoyl-CoA(18:0CoA)* decay. The smoother slope is the consequence of negative regulation of *CPT1* by *Malonyl-CoA* as well as marginal inhibitory effect of insulin signal on fatty acid transport system [48], which we incorporated in the event network as a slow reaction *myevent* on *FAT/CD36*. Reduction of *Malonyl-CoA* concentration was followed by increased the activity of CPT1 which further increased the rate of fatty acid oxidation after the first day (24 hours).

Fig. 8 shows the fluxes across all active metabolic reaction in Glycolysis I, TCA cycle, pyruvate metabolism, and β-fatty acid oxidation pathways during the course of in-silico experiment. The radius of circles show *log(flux)* value of

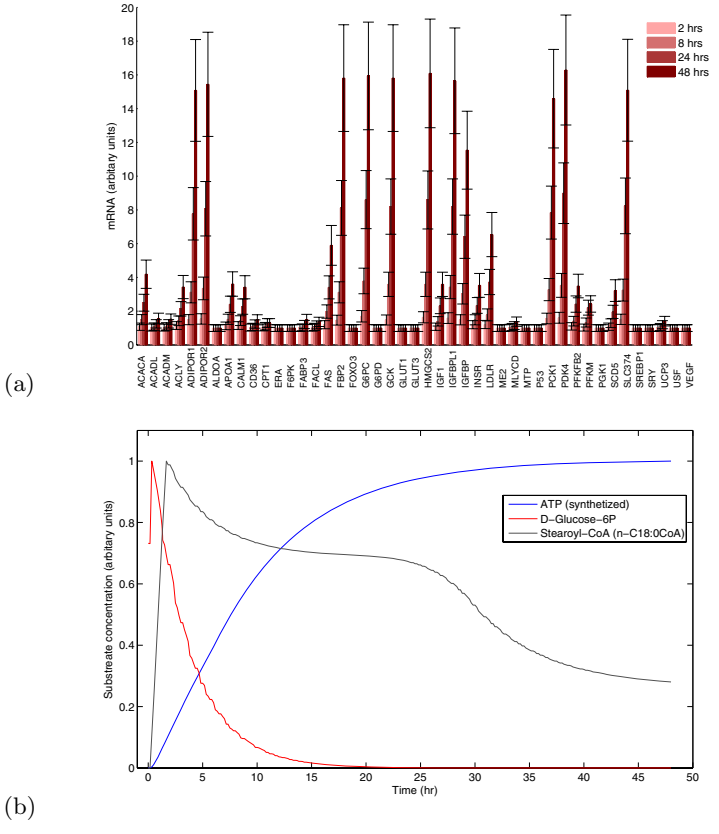

Fig. 7. Effect of 8 hours period of normal feeding followed by 40 hours of fasting on: (a) gene expression profile for all the genes in transcription regulatory network underlaying the current in-silico experiment. (b) concentration of *D-Glucose-6P*(red), *Stearoyl-CoA(18:0CoA)* (gray), and ATP in blue.

the reactions which were measured in 45 minute intervals, x and y axis are the time and reaction index, respectively. From here we reconfirm that roughly there was no flux across glycolysis pathway after 8 hours, where the flux across fatty acid reactions fluxes varied but sustained during the entire course of experiment.

Following is a list of selected parameters along with their values that we used in the simulation: Average myocardial cell volume $= 40 \times 10^{-15}\ m^3$ reported in [49], nucleus volume is $\sim$ %10 of the cell volume [23], in myocardial cell mitochondria occupies $\sim$ %30 of the cell volume [50], from data in literature we estimated there are $\sim$ 4660 cardiac myocytes per 1mg wet cells (varies among the samples). To convert any molecular counts in heart muscle to nano-Molar concentration we divided the counts by 240.88×10^2. The weight per amino acid was considered 0.11 KDa and an average weight of a eukaryotic cell $\sim 10^{-9}$ grams. The activation energy $9 < E_a < 21\ \ k_bT$ was used for the reaction model governed by Eqn. 1 and temperature was set to $T = 300\ \ K$. The complete list

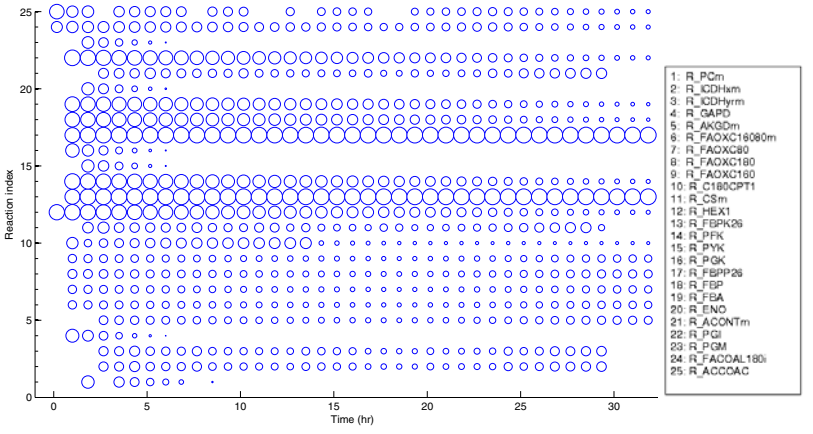

Fig. 8. Reaction fluxes for 25 active reactions: radius of each circle represents the log value of the flux per reaction, numbers on the y axis correspond to the reaction indices on the list to right of the chart

of parameters and their values is available upon the request from corresponding author (amin.mazloom@mssm.edu).

The simulation engine was entirely written in JAVA where JVM was running in Windows XP environment. The in-silico experiment was conducted on a stand alone Dell XPS-3000 machine, which had dual core 2.1 GHz P4 Intel processors and 4 GB of DDR2 RAM. Forty eight hours of simulation time took approximately two hours of CPU time.

5 Discussion

The proposed approach significantly reduced the computational cost of the experiment. Computation complexity is a major factor that challenges most system-level simulation efforts in biological networks. Appropriate design of the insulin and fatty acid event diagrams which forms the road map for evolution of the systems dynamics is essential in the of success of current approach. Although there is no general role to follow for designing the details of an event diagram, for instance where to incorporate a loop or when to aggregate a group of events (e.g GLUT4 tethering and fusion events); nevertheless both experience and practice as well as relevant data from biochemical and biological experiments are particularly important. Including a loop could increase the speed of the signal propagation in the subnetwork originated from that node by one fold. Furthermore on the border of the signal transduction and metabolic networks consistency in selecting of input/output resources is crucial for the evolution of the system. Also most of the laboratory experiments on myocardium are at the tissue and organ-level. Hence, we often have to apply certain approximations or assumptions to re-scale experimental data to be beneficial for our in-silico

simulation. Such mappings are not always trivial due to the missing vital data components. The accuracy in capturing the dynamics of a *myevents* is directly proportional to the precision of parametric physicochemical model that is associated with the model. In the current study in several cases we applied a single parameter probability distribution, further study is required to replace those distributions with more realist and accurate formalisms. The proposed regulatory mechanism projects the regulatory effect of a TF to a stochastic binary parameter with constant life time. Undoubtedly, real transcription regulatory mechanisms are far more complex, yet we observed that the current model could be a starting point in designing more complex and yet efficient transcription regulatory models that fit the system level simulation paradigm of complex biological systems. Although the outlined approach demonstrated a high capacity for system level simulation of a complex biological system such cardiac myocytes; however, big knowledge gaps in the structure of the target system could significantly diminish these capacity. Also designing a stochastic phytochemical model that can properly capture the temporal behavior a biological process is a particularly challenging task.

6 Conclusion

We established a novel framework in simulating the dynamics of biological networks with temporal heterogeneity across multiple cellular compartments. Inherent stochasticity that exists in the cell environment is conserved in the current in-silico framework. Furthermore, the proposed approach is scalable, very efficient and fairly accurate compared to available methods for system level modeling of biological systems. The promising capacities of the current approach was demonstrated by utilizing that in conducting an in-silico experiment on the metabolic plasticity of cardiac myocytes. We believe that current method could be very constructive in hypothesis testings experiments, drug target analysis, and cellular level study of diseases . Also the application of this approach is not limited to the heart cell and could potentially be applied in any cell lineage. Although the proposed method is sound in efficiency, yet demands substantial work especially in establishing more promising physicochemical models as well as the gene regulatory model to grant further accuracy to the results.

References

1. Zeigler, B.P., Kim, T.G., Praehofer, H.: Theory of Modeling and Simulation, 2nd edn. (2000)
2. Smith, N.P., Crampin, E.J., Niederer, S.A., Bassingthwaighte, J.B., Beard, D.A.: Computational biology of cardiac myocytes: proposed standards for the physiome. J. Exp. Biol. 210(9), 1576–1583 (2007)
3. Kuijpers, N.H.L., ten Eikelder, H.M.M., Bovendeerd, P.H.M., Verheule, S., Arts, T., Hilbers, P.A.J.: Mechanoelectric Feedback as a Trigger Mechanism for Cardiac Electrical Remodeling: A Model Study. Analys. of Biomedical Eng. 36(11), 1816–1835

4. Sedaghat, A.R., Sherman, A., Quon, M.J.: A mathematical model of metabolic insulin signaling pathways. Physiol. Endocrinal Metab 283, E1084–E1101 (2002)
5. Vinnakota, K., Kemp, M.L., Kushmerick, M.J.: Dynamics of Muscle Glycogenolysis Modeled with pH Time Course Computation and pH-Dependent Reaction Equilibria and Enzyme Kinetics. Biophy. J. 91, 1264–1287 (2006)
6. Gillespie, D.T.: A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. 22, 403–434 (1976)
7. van Kampen, N.G.: Stochastic Processes in Physics and Chemistry, 2nd edn. North-Holland, Amsterdam (1992)
8. Samoilov, M.S., Arkin, A.P.: Deviant effects in mulecular reaction. Nat. Compu. Biol. 24, 1235–1240 (2006)
9. Gadgil, C., Lee, C.H., Othmer, H.G.: A stochastic analysis of first-order reaction networks. Bull. Math. Biol. 67, 901–946 (2005)
10. Huss, J.M., Kelly, D.P.: Nuclear receptor signaling and cardiac energetics. Circulation Research 95, 568–578 (2004)
11. Nystrom, F.H., Quon, M.J.: Insulin signaling: metabolic pathways and mechanisms for specificity. Cell Signal 11(8), 563–574 (1999)
12. Saltiel, A.R., Kahn, R.: Insulin signaling and the regulation of glucose and lipid metabolism. Nature 414, 799–806 (2001)
13. Kanehisa, M.: KEGG Data base. Novartis found Sym. 247, 91–101 (2001)
14. Kodde, I.F., van der Stok, J., Smolenski, R.T., de Jong, J.W.: Metabolic and genetic regulation of cardiac energy substrate preference. Comparative Biochem. and Physiol. 146, 26–39 (2006)
15. Gould, G.W., Holman, G.D.: The glucose transporter family: structure, function and tissue specific expression. Biochem. 295(2), 329–341 (1993)
16. Watson, R., Pessin, J.E.: Glut4 Translocation: The last 200 nanomters. Cellular Signalling 19, 2209–2217 (2007)
17. Chang, L., Chiang, S., Saltiel, A.R.: Insulin signaling the regulation of glucose transport. Mlecular Medicine 10, 7–12 (2004)
18. Ghosh, P., Ghosh, S., Basu, K., Das, S.K.: Holding time estimation for reactions in stochastic event-based simulation of complex biological systems. Model. Pract. Theory (2007) doi:10.1016/j.simpat.2007.09.002
19. Ghosh, P., Ghosh, S., Basu, K., Das, S.K.: A diffusion model to estimate the inter-arrival time of charged molecules in stochastic event based modeling of complex biological networks. In: Proc. IEEE Comp. Systems Biol. Conf. (2005)
20. Mazloom, A.R., Basu, K., Das, S.K.: A Random Walk Modelling Approach for Passive Metabolic Pathways in Gram-Negative Bacteria. In: Proce. IEEE CIBCB Conf., pp. 1–8 (2006)
21. Koonen, D.P., Glatz, J.F., Bonen, A., Luiken, J.J.: Long chain fatty acid uptake and FAT/CD36 trasnlocation in heart and skeletal muscle. Biochem. Biophys. Acta. 1736(3), 163–180 (2005)
22. Bonnefont, J.P., Demaugre, F., Prip-Buus, C., Saudubray, J.M., Brevit, M., Abadi, N., Thuillier, L.: Carnitine palmitoyltransferase deficiencies. Mol. Genet. Metab. 68, 424–440 (1999)
23. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J.D.: Molecular Biology of the Cell, 4th rev. edn. Garland Publishing Inc. (2002)
24. Flores, O., Lu, H., Reinberg, D.: Factors Involved in Specific Transcription by Mammalian RNA Polymerase II. Biol. Chemistry 267(4), 2786–2793 (1992)
25. Bushnell, D.A., Westover, K.D., Davis, R.E., Kornberg, R.D.: Structural basis of transcription: an RNA Polymerase II-TFIIB Cocrystal at 4.5 Angstroms. Science 303, 983–988 (2004)

26. Sims III, R.J., Belotserkovskaya, R., Reinberg, D.: Elongation by RNA polymerase II: the short and long of it. Genes & Dev. 18, 2437–2468 (2004)
27. Lee, T.I., Young, R.A.: Regulation of gene expression by TBP-associated proteins. Genes & Dev. 12, 1398–1408 (1998)
28. Perez-Ortin, J.E., Alepuz, P.M., Moreno, J.: Genomics and gene transcription kinetics in yeast. TRENDS in Genetics 23(5), 250–257 (2007)
29. Zhang, C., Zobeck, K.L., Burton, Z.F.: Human RNA Polymerase II elongation in slow motion: role of the TFIIF RAP74α1 Helix in nucleoside triphosphate-driven translocation. Molec. and Cell. Biol. 25(9), 3583–3595 (2005)
30. Audibert, A., Weil, D., Dautry, F.: In Vivo Kinetics of mRNA Splicing and Transport in Mammalian Cells. Molc. and Cell. Biol. 22(19), 6706–6718 (2002)
31. Jackson, R.J.: Alternative mechanisms of initiating translation of mammalian mRNAs. Biochem. Society Lectures, pp.1231–1241 (2005)
32. Pestova, T.V., Kolupaeva, V.G., Lomakin, I.B., Pilipenko, E.V., Shatsky, I.N., Agol, V.I., Hellen, C.U.T.: Molecular mechanisms of translation initiation in eukaryotes. Proc. Natl. Acad. Sci. 98(13), 7029–7036 (2001)
33. Lorsch, J.R., Herschlag, D.: Kinetic dissection of fundamental processes of eukaryotic translation initiation in vitro. The EMBO Journal 18(23), 6705–6717 (1999)
34. Saha, A., Wittmeyer, J., Cairns, B.: Chomatin remodeling the industrial revolution of DNA arround histones. Nat. Rew. Mol. Cell Biol. 7, 437–447 (2006)
35. Bustamanate, C., Peterson, C.L., Cairns, B.R., Smith, S.B., Mihardja, S., Grill, S.Q., Saha, A., Smith, C.L., Zhang, Y.: DNA transclocation and loop formation mechanism of charomatin remodeling by SWI/SNF and RSC. Mol. Cell 24, 559–568 (2006)
36. Ghosh, S., Ghosh, P., Basu, K., Das, S.K.: Modeling the stochastic dynamics of gene expression in single cells: a birth and death markov chain analysis. In: IEEE (BIBM) Conf., pp. 308–316 (2007)
37. Singh, S., Yang, H.O., Chena, M., Yu, S.: A kinetic-dynamic model for regulatory RNA processing. J. of Biotechnology 127, 488–495 (2006)
38. Kim, H., Yin, J.: Effects of RNA splicing and post-transcriptional regulation on HIV-1 growth: a quantitative and integrated perspective. IEE Proc. Syst. Biol. 152(3), 138–152 (2005)
39. Sanford, J.R., Caceres, J.F.: Pre-mRNA splicing: life at the centre of the central dogma. J. Cell Science 117, 6261–6263 (2004)
40. Cao, D., Parker, R.: Kinetic dissection of fundamental processes of eukaryotic translation initiation in vitro. RNA Journal 7, 1192–1212 (2002)
41. Kubitscheck, U., Grünwald, D., Hoekstra, A., Rohleder, D., Kues, T., Siebrasse, J.P., Peters, R.: Nuclear transport of single molecules: dwell times at the nuclear pore complex. Cell Biology 168(2), 233–243 (2005)
42. Belle, A., Tanay, A., Bitincka, L., Shamir, R., O'Shea, E.K.: Quantification of protein half-lives in the budding yeast proteome. Proc. Natl. Acad. Sci. 103(35), 13004–13009 (2006)
43. Mounier, C., Poser, B.I.: Transcriptional regulation by insulin: from receptor to the gene. Physiol. Pharmacol 84, 713–721 (2006)
44. Brownsey, R.W., Boone, A.N., Allard, M.F.: Action of insulin on mammalian heart: metabolism, pathology and biochemical mechanism. Cardiovascular Research 34, 3–24 (1997)
45. The BiGG database (2007), `http://bigg.ucsd.edu`

46. Covert, M.W., Schilling, C.H., Palsson, B.: Regulation of gene expression in Flux balance models of metabolism. Theoritical Biology 213, 73–88 (2001)
47. Van der Lee, K.A.J.M., Willemsen, P.H.M., Samec, S., Seydoux, J., Sulloo, A.G., Pelsers, M.M.A.L., Glatz, J.F.C., Van der Vusse, G.J., Van Bilsen, M.: Fasting-induced changes in the expression of genes controling metabolism in rat heart. J. of Lipid Research 42, 1752–1758 (2001)
48. Saddik, M., Gamble, J., Witter, L.A., Lopaschuk, G.D.: Acetyl-CoA carboxylase resulation of fatty acid oxidation in the heart. J. of Biological Chemistry 268(34), 25836–25845 (1993)
49. Maisch, B.: Enrichment of vital adult cardiac muscle cells by continuous silica sol gradient centrifugation. Basic Res. Cardiol. 76, 622–629 (1981)
50. Drake-Holland, A.J., Noble, M.I.M.: Cardiac metabolism. John Wiley and Sons, New York (1983)