Multiscale modeling of cellular epigenetic states and tissue patterning: stochasticity in molecular network, chromatin folding in cell nucleus, and cell-cell interactions in tissue patterning

Jie Liang,[1,*] Youfang Cao,[2] Gamze Gürsoy,[1] Hammad Naveed,[3] Anna Terebus,[1] & Jieling Zhao[1]

[1]Program in Bioinformatics, Department of Bioengineering, University of Illinois at Chicago, IL, 60305, USA; [2]Theoretical Biology and Biophysics (T-6) and Center for Nonlinear Studies (CNLS), Los Alamos National Laboratory, Los Alamos, NM, 87545, USA; [3]Toyota Technological Institute at Chicago, 6045 S. Kenwood Ave. Chicago, Illinois 60637, USA;

*Address all correspondence to: Jie Liang, Program in Bioinformatics, Department of Bioengineering, University of Illinois at Chicago, IL, 60305, USA

Program in Bioinformatics, Department of Bioengineering, University of Illinois at Chicago, IL, 60305, USA; Tel: 312-355-1789; jliang@uic.edu

ABSTRACT: Genome sequences provide the overall genetic blueprint of cells, cells possessing the same genome can exhibit diverse phenotypes. There is a multitude of mechanisms controlling cellular epigenetic states and that dictate the behavior of cells. Among these, networks of interacting molecules, often under stochastic control, depending on the specific wirings of molecular components and the physiological conditions, can have a different landscape of cellular states. In addition, chromosome folding in three-dimensional space provides another important control mechanism for selective activation and repression of gene expression. Fully differentiated cells with different properties grow, divide, and interact through mechanical forces and communicate through signal transduction, resulting in the formation of complex tissue patterns. Developing quantitative models to study these multi-scale phenomena and to identify opportunities for improving human health requires development of theoretical models, algorithms, and computational tools. Here we review recent progress made in these important directions.

KEY WORDS: cell modeling, stochastic network, epigenetic state, chromatin folding, vertex cell model, epithelial topology,

I. INTRODUCTION

With the completion of the sequencing of human genomes and the advent of individual genomes[1–3], we are gaining full access to the genetic blueprints of human cells. However, much of the behavior of cells depends on the control of expression of genes. Cells from different tissues posses the same genome but exhibit different properties and perform different functions. Understanding how genetic information is selectively expressed and how epigenetic control establishes the specific expression patterns of proteins are fundamental problems of biology and problems with important implications in the treatment of disease.

As large scale epigenetic profiling studies have shown, covalent modification such as DNA methylation and post-translational modifications of proteins play important roles in controlling the expression of genes[4,5]. The epigenetic states of cells can also arise from the network of multiple molecular components

0278–940X/14/$35.00 © 2016 by Begell House, Inc. www.begellhouse.com

itself[6–9]. Often under stochastic control, different biological phenotypes can result from the same molecular network, as there may exist multiple peaks in the stationary probabilistic landscape of the reaction network[6,9]. With specific wirings, cells can maintain heritable states depending on the physiological conditions without covalent bond modifications[6–9].

Furthermore, recent studies based on fluorescence *in-situ* hybridization (FISH), chromosome conformation capture and related techniques have demonstrated the importance of the spatial organization of genome in determining cellular phenotypes[10–14]. The uncovering of massive amount of long range interactions in chromatins in different cells[11–14] suggests the existence of a complex three-dimensional folding landscape of chromosomes. The discovery of hierarchical structural units such as topologically associated domains (TAD) points to the existence of complex folding machineries that provide cooperative control of gene expression.

Quantitative understanding of both stochastic networks and genome folding requires the development of fundamental theory, models, and algorithms so that effective computational analysis can be efficiently carried out. Once differentiation of cells is complete and different cell types are formed, another important task is to understand and model how populations of different cells interact and form different patterns of tissue, and how insight into complex processes such as wound healing can be gained through computational studies.

In this review article, we examine recent progress in the development of theoretical model, algorithms, and computational methods for computing the probability landscape of stochastic network, for predicting three-dimensional structures of folded chromosomes, and for understanding tissue pattern formation.

II. STOCHASTIC NETWORK AND DISCRETE CHEMICAL MASTER EQUATION

Networks of interacting molecules are the basis of the regulatory machineries of cells. When the copy numbers of molecules involved are small (μM to nM range, or $1-10^2$ copies of molecules in a $10\mu M^3$ cell)[15,16], stochasticity can play significant roles in cellular processes such as gene regulation, protein synthesis, and signal transduction[15,17–21]. Well-known examples of cellular epigenetic states controlled by a stochastic network include the decision network of switching between the lysogenic and the lytic phases of phage lambda[6,15,22,23], and the competence transition network of *Bacillus subtilis*[24].

A stochastic network is characterized by its time-dependent probability landscape, namely, the probability of the network to be at each possible microstate, namely, the vector of copy numbers of molecular species in the network. Knowledge of the full probability landscape of a network will identify all epigenetic states, which correspond to the peaks in the landscape, as well as transitions and transition probabilities between them. The basic theoretical framework for studying a stochastic network is given by the discrete chemical master equation (dCME)[25,26]

$$\frac{dp(\boldsymbol{x},t)}{dt}=\sum_{\boldsymbol{x}'}\Big[A(\boldsymbol{x}',\boldsymbol{x})p(\boldsymbol{x}',t)-A(\boldsymbol{x},\boldsymbol{x}')p(\boldsymbol{x},t)\Big], \tag{1}$$

where $p(\boldsymbol{x},t)$ is the probability of the network in the microstate $\boldsymbol{x}$, $A(\boldsymbol{x}',\boldsymbol{x})$ is the matrix of transition rates between microstates $\boldsymbol{x}$, and $\boldsymbol{x}'$.

A. Computing probabilistic landscape

However, analytically solving the dCME is not possible except for a few simple models with one or two nodes. The stochastic simulation algorithm (SSA) and related techniques have been widely used for studying stochastic networks[25,27]. By generating Monte Carlo samples of trajectories of reactions of a network using SSA, one can infer properties of the stochastic network through analysis of millions of sampled trajectories[25]. However, SSA is inefficient in sampling rare events, as most computing time is spent on following high-probability paths. In addition, the assessment of convergence becomes difficult when the network is complex[28–31]. Overall, it is

generally difficult to know if an accurate solution to the dCME of a network has been found. One does not know if all major probabilistic peaks have been identified and different epigenetic states located, or if important ones with significant probability mass in the usually high dimensional space are undetected. It is also difficult to know if the locations of identified probabilistic peaks are correctly mapped. One also does not know if a computed probabilistic landscapes is overall erroneous and how such errors can be quantified. Furthermore, the best possible accuracy one can achieve with a given finite computing resources is generally unknown. One also does not know what computing resource is required so solutions with accuracy within a predefined tolerance can be obtained. These issues can be resolved using recently developed ACME method as described in a later section.

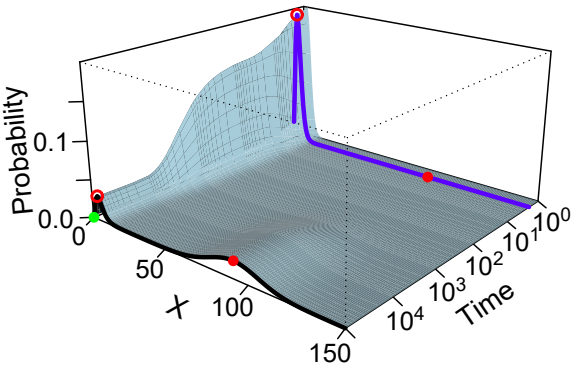

FIG. 1: The time-evolving probability and transition probability of rare events of the bistable Schlögl model. The blue and black curves show the landscape at $t = 2$ and at the steady state, respectively. The two high probability regions at the steady state (black curve) are located at $x = 4$ (red circle on black curve) and $x = 92$ (red dot on black curve), respectively. They are separated by a high barrier of low probability. The initial state $x = 0$ (green dot) is near the first peak, and the target state (red dot) is at the center of the second peak. The probability landscape at time $t = 2$ (blue curve) shows a much sharper peak centered at $x = 3$ (red circle on blue curve). The transition from $x = 0$ to $x = 92$ within $t = 2$ is a rare event and the transition paths have a steep barrier to cross. The probability of this rare event can be sampled effectively using the ABSIS method (adapted from reference[31]).

1. Biased Sampling for Barrier Crossing

To improve the efficiency of sampling rare events, Kuwahara and Mura developed the weighted SSA (wSSA) algorithm by biasing each reaction rate using a pre-determined constant, with the overall summation of reaction rates unchanged[28]. Further development introduced biases to reaction selection and significantly improved sampling efficiency for rare events[29,30]. These methods, however, do not address the issue of crossing barriers, which arises in multistable networks with a complex probability landscape. In reference[31], a general theoretical framework for obtaining optimized biases in sampling individual reactions for estimating probabilities of rare events was developed. In addition, a practical algorithm called *adaptively biased sequential importance sampling* (ABSIS) method for efficient probability estimation was also given[31]. By using a look-ahead strategy and by enumerating short paths from the current microstate, the reaction-specific and state-specific forward and backward moving probabilities of the system were estimated, which are then used to bias reaction selections. The ABSIS algorithm can automatically detect barrier-crossing regions, and can adjust bias adaptively, with bias determined by the outcome of exhaustively generated short paths[31]. Test results on the biochemical networks (see Fig 1 for the example of the Schögle bistable model) showed that the ABSIS method can accurately and efficiently estimate rare event probabilities, often with smaller variance than other importance sampling algorithms[31].

2. Direct solution of dCME models

Several methods have been developed towards the goal of directly computing the full probability landscape of a stochastic network. These include the finite state projection (FSP), the sliding window method, the finite buffer dCME method, as well as several other techniques[23,32–36]. The FSP method is based on a truncated projection of the state space and uses numerical techniques to compute the time-

evolving probability landscapes[32,37,38]. However, the use of an absorbing boundary leads to the accumulation of errors as time proceeds, therefore making it unsuitable to study long-time and steady state behavior of a network. The sliding window method is also based on truncation of the state space. To ensure small truncation error, a large number of states must be included, as the size of the state space takes the form of a n-dimensional hypercube, with n the number of molecular species. This makes it difficult to achieve the desired level of accuracy.

A bottleneck problem for solving the dCME directly is to have an efficient and adequate account of the discrete state space. As the copy number of each of the n molecular species takes an integer value, conventional methods of state enumeration incorporate all vertices in a n-dimensional hypercube nonnegative integer lattice, which has an overall size of $O(\prod_{i=1}^{n} b_i)$, where b_i is the maximally allowed copy number of molecular species i. State enumeration rapidly becomes intractable, both in storage and in computing time. To address this issue, we developed the finite buffer discrete CME (fb-dCME) method for efficient enumeration of the state space[33]. It uses a buffer queue with a fixed number of molecular tokens to keep track of the remaining number of states that can be enumerated. Instead of including every state in a hypercube, it examines only states that can be reached from a given initial state. This approach has been extended in the Accurate Chemical Master Equation (ACME) method[39,40], in which the stochastic network is decomposed into independent components, each equipped with its own finite state space controlled by a separate buffer queue. Molecular species in each component can be transformed into each other through mass-balance reactions.

The ACME method is provably optimal in space and in time required for state enumeration, and exhibits very effective use of the overall finite state space. The size of the overall truncated state space is $O(\prod_j \binom{B+n_j}{n_j})$, a product of the volumes of n_j-simplices, with n_j the number of molecular species in the j-th component of the network. While maintaining the same maximum copy number of molecules, the truncated space is dramatically smaller than the n-dimensional hypercube, which has an overall size of $O(\prod_j B^{n_j})$. The reduction of state space is roughly $\prod_j (n_j!)$, therefore the size of the state space is much smaller than that generated using the conventional hypercube method.

Furthermore, a broad mathematical framework for analytical estimation of steady state probability solutions has been developed[39]. Importantly, an overall upper-bound estimation of the truncation errors has been obtained from theoretical analysis using the techniques of aggregation of microstates by buffer usage, factorization of a quotient matrix, and stochastic ordering[39,40]. This enabled estimation of truncation error *a priori* without actually carrying out costly trial solutions of the dCME for the purpose of error estimation. It also allows determination if the truncation error for a given computing environment is within a pre-defined tolerance threshold. In addition, it informs on the minimal size of state space required to achieve a given level of accuracy.

The ACME method works for arbitrary networks, including those with strong couplings between species. With this optimal method for state enumeration, numerical methods for solving large linear systems can be applied to directly solve the dCME equation.

B. Multistability of common motifs

Network motifs consisting of 1–4 nodes are small subunits that occur frequently and often have defined functions[41]. They are the small building blocks of more complex regulatory networks. It is important to understand their roles in maintaining the multistability of different epigenetic states.

Studies based on deterministic models showed that multistability can arise only when feedback and cooperativity are present[42]. However, multistability can arise from stochastic network motifs which would go undetected if deterministic network models are used. The smallest network motif exhibiting stochastic multistability is that of the self-regulating gene. It was shown that a probabilistic landscape

with multiple peaks can appear when the rates of gene expression in bound and unbound forms of DNA are well separated and switching between the bound and unbound states is slow[43]In this case, the multistability, however, cannot be captured using deterministic ordinary differential equation (ODE) models or simple Fokker-Plank approximation of the dCME[43]. Overall, stochastic multistability in simple network motifs remain poorly characterized.

The ACME method can be applied to study the stochastic behavior of motifs of gene regulatory network. As an illustration, we discuss the motif of a single input module[44]. There are three genes in this network. $GeneA$, $GeneB$, and $GeneC$ express protein A, B, and C, respectively (Fig 2A). Protein A inhibits both $GeneB$ and $GeneC$, both in turn activate the expression of protein A. Multistability exist in the probabilistic landscape of this simple motif computed using the ACME method, where bistabilities for expression of proteins A, B, and C can be clearly seen (Fig 2B and 2C).

With the ACME method, we can examine systematically the stochastic behavior of all major network motifs with 3–4 nodes, including negative feed-back loops[45], feed-forward loops[46], and four-nodes bifan networks[47]. For example, the results of computation of the probability landscape using ACME for the incoherent feed-forward loop (Fig 2D) showed that the stationary probabilistic landscape of the motif changes from that of one peak (E) to that of six peaks (F) when network parameters are altered. With the ACME method, we can characterize the probability landscapes of these motifs and identify key factors defining their stochastic multistabilities. We can now explore the full parameters space and construct phase diagrams of multistability of many network motifs. These results can be useful for constructing synthetic gene networks for desired stochastic properties.

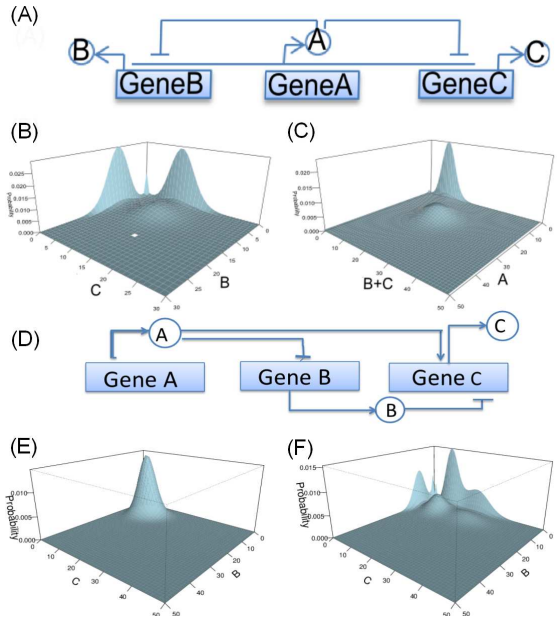

FIG. 2: Single input network module. (A) Its network architecture. Steady state probability landscapes for (B) proteins B and C, and (C) proteins A and $B + C$. (D) The network architecture of the incoherent feed-forward loop. Steady state probability landscapes (E) with one peak and with (F) six peaks.

C. Epigenetic decision network of cellular fate

Bacteriophage lambda is a virus that infects *E. coli* cells (Fig 3). A molecular circuitry controls phage lambda to choose between two productive modes of development, namely, the lysogenic phase and the lytic phase (Fig 3A). In the lysogenic phase, phage lambda represses its developmental function, integrates its DNA into the chromosome of the host *E. coli* bacterium, and is replicated in cell cycles for potentially many generations. When threatening DNA damage occurs, for example, when UV irradiation increases, phage lambda switches from the epigenetic state of lysogeny to the lytic phase and undergoes massive replications in a single cell cycle, releasing 50–100 progeny phages upon lysis of the *E. coli* cell[22].

To study how lysogeny is maintained and how it transitions to the lytic state, we used a simplified

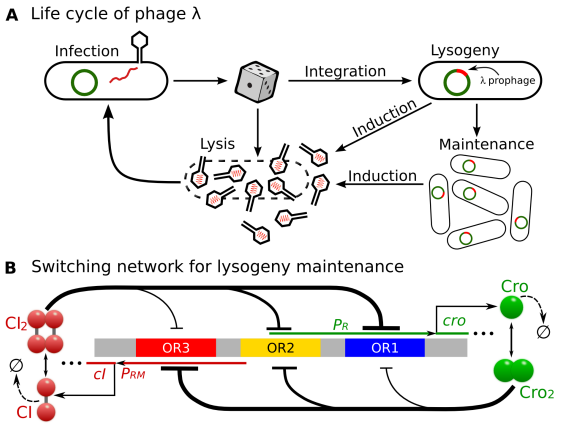

FIG. 3: Different selection of cell fate of *E. coli* infected by phage lambda and a model of the epigenetic circuit for lysogeny maintenance. (a) The lysogenic and lytic phases of phage lambda. (b) A simplified model of the epigenetic switch for lysogeny maintenance. (Modified from reference[23]).

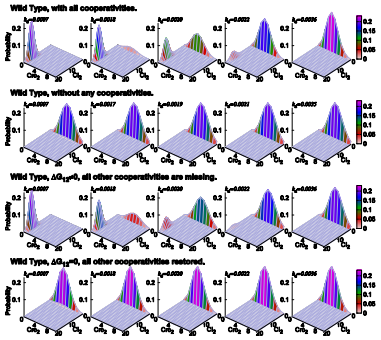

FIG. 4: The probability landscape of the epigenetic circuits of lysogeny maintenance in phage lambda. (Row 1) For wild type phage lambda, at the CI degradation rate of $k_d = 7.0 \times 10^{-4}$/s, probability landscape centers at locations with high copy numbers of Cl_2 and close to 0 copy of Cro_2. This corresponds to the lysogenic phase of phage lambda. When k_d increases from $k_d = 1.8 \times 10^{-3}$/s to 2.2×10^{-3}/s, the peak located at lysogenic phase gradually diminishes, whereas the peak located at lytic phase gradually increases. At about $k_d = 2.0 \times 10^{-3}$/s, phage lambda has about equal probability to be in either lysogenic or lytic phase. When CI is degraded at a faster rate of $k_d = 3.6 \times 10^{-3}$/s, the probability landscape centers at locations where there are higher copy numbers of Cro dimer and close to 0 copy of CI. This corresponds to the lytic phase of phage lambda. (Row 2) When all cooperativities are removed from the model, lysogeny cannot be achieved. (Row 3) When only the cooperativity of ΔG_{12} is restored, wild-type behavior is largely restored. (Row 4) When all other cooperativities except ΔG_{12} are restored, lysogeny still cannot be achieved (Modified from reference[23]).

stochastic model for the molecular regulatory network that controls the epigenetic switch in phage lambda (Fig 3b)[23]. With a total of 54 reactions involving 13 molecular species and around 1.7 million microstates, the effects of UV irradiation can be modeled by increasing the CI degradation rates k_d due to the response of the SOS system. The steady state probability associated with each of these microstates are computed from dCME after state enumeration[23]. Fig 4 (row 1) shows the probability landscape of phage lambda projected to the subspace of CI and Cro dimers at five different UV irradiation conditions, each modeled with a different CI degradation rate k_d. With a high copy number of the Cl_2 repressor, the lysogenic phase of the phage lambda is maintained, whereas a high copy number of Cro_2 protein signifies the lytic phase[48]. A clear picture of the landscape in lysogeny, at the start of transition, during mid-transition, at the end of transition, and in lysis can be seen.

Calculation results also showed that wild-type phage lambda can maintain a constant level of repressor over a wide range of repressor degradation rates and is stable against UV irradiation, ensuring heritability of the lysogenic state. Furthermore, it can switch efficiently to the lytic state once repressor degradation increases past a high threshold by a small amount. Novel findings including the role of cooperativity, effects of mutations and hair-triggers, and the origin of network robustness, are discussed in reference[23].

The example of the decision network of phage lambda demonstrated that realistic systems can now be directly studied using the ACME method. A detailed network of 16-dimensional MAPK network that can be studied using the ACME method is shown in Figure 5, with its probability landscape

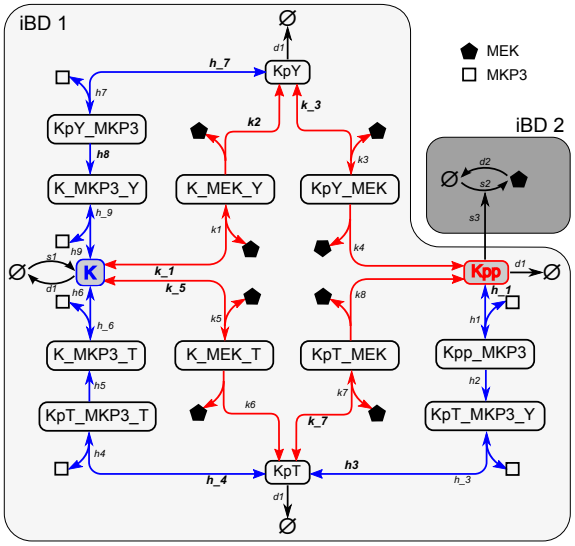

FIG. 5: A detailed network model of the MAPK cascade. The ERK(K) phosphorylation is catalyzed by the kinase MEK, whereas MEK synthesis is up-regulated by dual phosphorylated ERK(Kpp). Detailed reactions during the dual phosphorylation process of the ERK(K), the synthesis and degradation of MEK are explicitly modeled. Red and blue arrows represent phosphorylation and dephosphorylation reactions, respectively. Bidirectional arrows represent reversible reactions. The network can be partitioned into two components, as shown in two shaded areas of different color. There are a total of 16 molecular species and 35 individual reactions in the network. (Modified from reference [40]).

computed exactly[40]. This is a significant improvement over current practice in which one introduces simplifications to reduce this network into that of 3-4 nodes, with consequences unknown, so SSA simulation can be carried out.

III. CHROMATIN FOLDING AND CELLULAR STATES

Recent development of fluorescence *in-situ* hybridization (FISH), chromosome conformation capture (3C) and related techniques (4C, 5C, Hi-C) enabled large-scale discovery of long-range chromatin looping interactions among distant chromosomal elements[10–14], as well as the identification of Topologically Associated Domains(TAD)[49–51]. Understanding the spatial organization of the genome in a cell nucleus provides an important route towards gaining insight into the mechanisms of gene activities and the maintenance of cellular epigenetic states[52].

A. Global scaling properties of chromosome folding

FISH and chromosome conformation capture studies revealed important global scaling properties of the human genome, including the relationship between contact probability $P_c(s)$ and loop length s[13], the spatial distance $R^2(s)$ between elements and the loop length s, as well as the leveling-off effects of spatial distances between genomic elements as their genomic distances increases[10,11].

The fractal globule (FG) model[13] was the first model developed to describe the global folding properties of the human genome, as it can explain the scaling relationship between $P_c(s)$ and s. However, it does not account for the leveling-off effects observed in FISH experiments[10,11]. Subsequently, the Strings and Binders Switch (SBS) model was developed, which pointed to a more heterogeneous structural ensemble, in which the scaling properties of the individual structures depend on the concentration of binder molecules such as architectural proteins[53]. However, scaling in the SBS model strongly depend on the choice of model parameters, and all observed scaling properties cannot be accounted for using a fixed set of parameters.

Chromosomes reside within the severely confined space of the cell nucleus. However, the direct effects of nuclear confinement on chromatin folding and compaction are not known. A detailed computational model, named Constrained Self-Avoiding Chromatin (C-SAC), was developed for studying the folding properties of chromosomes (Fig 6A)[54]. In C-SAC, nuclear confinement is explicitly modeled and ensemble of chromatin chains are generated inside the severely confined cell nucleus. This was achieved by overcoming the challenging problem of attrition in sampling using the geometrical sequential important sampling technique[54–61]

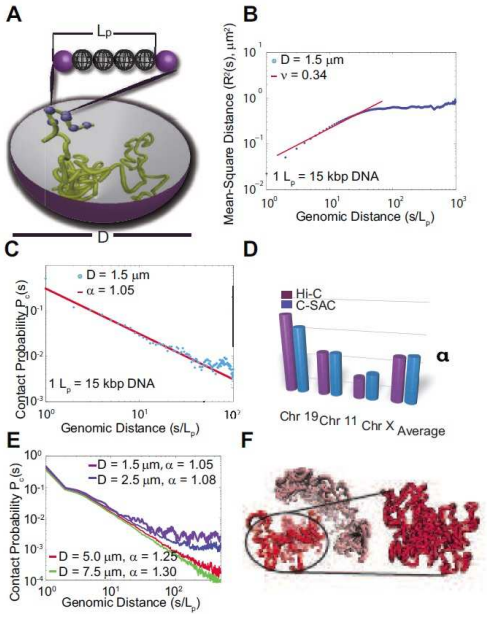

FIG. 6: The physical model of C-SAC chains and its scaling properties. **(A)** Schematic representation of the C-SAC model. A chromatin fiber is represented by a self-avoiding polymer chain with a persistence length L_p. Polymers are grown as chains inside a spherical confined space of a diameter D. **(B)** The scaling of $R^2(s)$ from 10,000 chains of length 1,000L_p in $\log_{10}$ scale. **(C)** The scaling of contact probability $P_c(s)$. **(D)** Comparison of exponent α of $P_c(s)$ between C-SAC and Hi-C data[13]. Values of α for different chromosomes[13,53] were compared to those calculated separately for different clusters of C-SAC chains. **(E)** $P_c(s)$ *vs.* chain length s for different confinement sizes D. **(F)** A random C-SAC chromatin chain with independent substructures. A singular domain-like conformation is shown in detail (adapted from reference[54]).

Results using C-SAC showed that spatial confinement of the nucleus is responsible for much of the experimentally observed scaling behavior of chromosome folding (Fig 6B-E)[10,11,13]. The C-SAC model also predicts the formation of highly interactive sub-structures that may give rise to topological domains[62] (Fig 6F). A recent study further suggested that the segregated organization of chromosome is due to the onset of glassy dynamics when long chromosome chains are confined into small nuclei[63]. These findings highlight the importance of nuclear confinement and how its size change may regulate epigenetic programming of cells, as in the case of transitioning from stem cells to differentiated cells[53,54].

B. Genome-wide structures of multi-chromosomes in yeast

Understanding different cellular states in mammalian cell differentiation requires knowledge of genome-wide interactions of chromatins. Budding yeast is an excellent model system for constructing genome-wide models of multi-chromosomes, as its transcription machineries have been well studied and the roles of genome organization well understood[64].

Nevertheless, there is some controversy on the extent that the genome organization of yeast is dictated by the architectural properties of the cell nucleus. Recent computational studies showed that yeast chromosomes behave as randomly folded flexible self-avoiding polymer chains subject to the constraints derived from nuclear landmarks and nuclear confinement[65,66]. However, correlations of modeled inter-chromosomal interactions with experimentally captured interactions are modest at best[65,66]. The question whether the genome organization of budding yeast is fully dictated by physical tethering of the landmarks and the excluded-volume of the self-avoiding chains is unresolved. Another study showed that once observed interactions are corrected using a null model, the genome organization of the budding yeast no longer exhibit the properties of randomly folded polymer chains with constraints[67].

To shed light into the nature of the organization of yeast chromosome, we developed a three-dimensional chromosome model incorporating constraints derived from electron/light microscopy experiments to mimic the nuclear environment and its effects on the folding of the yeast genome (Fig 7A). Large ensembles of model genomes (∼200,000) were generated using different constraints. Our results showed that the organization of the yeast genome, including interchromosomal interactions is indeed dictated to a large extent by the confinement of the cell nucleus and the physical tethering of the

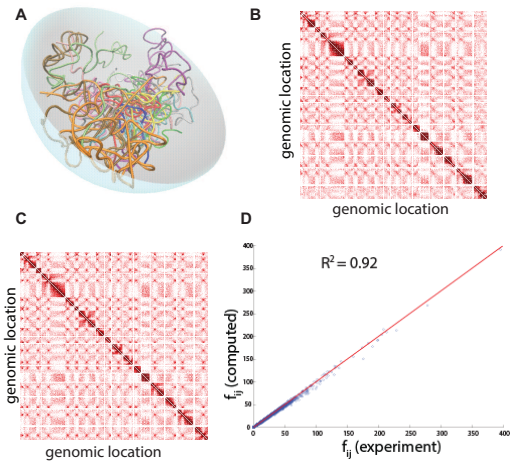

FIG. 7: Multi-chromosome model of budding yeast. **(A)** Schematic representation of yeast genome in the confined nuclear space. Different chromosomes are represented in different colors. **(B)** Heatmap of interactions that are taken from biochemical experiments[14]. Each block represents a chromosome and darker color indicates higher frequency interaction. **(C)** Heatmap of interactions that are averaged from predicted model genomes built without using any biochemical data. The heatmap generated using predicted models is very similar to the measured heatmap in (C). **(D)**. Experimentally measured interaction frequencies correlate well with predicted interaction frequencies from our model (see reference[61] for more details.

centromeres[61].

Comparison of computatioanlly generated ensembles of folded chromosomes g with those from 3C-based studies showed that the majority of measured interactions regulating important cellular functions are captured by our model (at an accuracy of 92%, Fig. 4B-D)[61]. Furthermore, preliminary study suggests that these computationally captured interactions can be used to extract biologically specific interactions from experimental data, as they arise from the nuclear architecture of the cell nucleus and can be subtracted. This would allow predictions of novel transcription regulatory mechanism in budding yeast that are experimentally testable.

C. Spatial structures of functional loci of chromatin in different cell states

Experimental data from chromosome conformation capture studies can capture detailed promoter-enhancer interactions at the locus level. However, the captured information is intrinsically two-dimensional pairwise contact frequencies. Such captures may also be limited by the distribution of restriction enzyme sites and the sequence mappability. Constructing 3D structures of a gene locus can help to obtain detailed structural understanding of promoter-enhancer interactions and how they may affect transcriptional machineries and regulate cellular epigenetic states. Several polymer models of chromatin in genome or locus level have been developed[13,14,53,63,65–74], and many can generate structures of chromatins from 3C/5C measurements, capturing many known interactions.

There are many difficulties in constructing accurate 3D models of a gene locus. Random interactions from non-specific formaldehyde fixation[75] introduce complexity, as it is challenging to distinguish biologically relevant interactions from non-specific random spatial interactions. As measurements of 3C/5C are from a mixture of cell populations, full ensembles of 3D chromatin chains need to be generated to collectively best describe the bulk measurements and to reveal specific interactions from experiments[76].

We have developed a method that can generate well-sampled ensembles of high-resolution chromatin chains of a specific locus. Our method can remove non-specific physical interactions from 3C measurements and can incorporate data of specific interactions. We have applied our method to study the chromatin structures of the α-globin locus, which is differentially expressed in normal (GM12878) and cancer (K562) cell lines[71] (Fig 8A-B). Comparison of differences in their spatial structures will help to understand how spatial organization of the genome affects the expression level of this important gene and will allow identification and interrogation of patterns of spatial interactions. Preliminary results showed that the differential expression of α-globin is

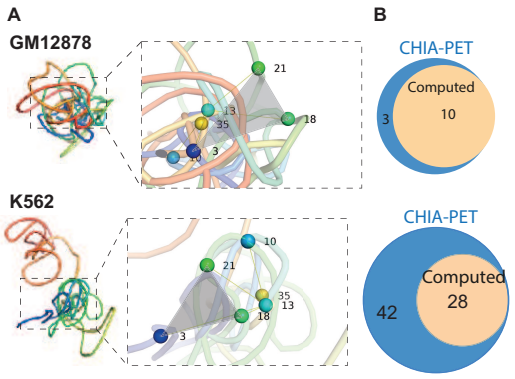

FIG. 8: Structural models of the α-globin gene domain. (A) Representative three-dimensional chromatin chains of the α-globin gene domain in the GM12878 and the K562 cell lines. The α-globin gene domain has only one domain and is more compact in the GM12878 cell. In the K562 cell, the α-globin gene domain has two distinct domains that form a more extended structure. Close-up views of selected sites for comparison between cell lines are also shown, with specific nodes noted. **(B)** Comparison between the number of CTCF- (top) and RNAPol2-mediated (bottom) interactions based on the predicted model (beige) and based on measurements from an independent experimental study using the technique of CHIA-PET (blue, reference[77]).

strongly influenced by the folding landscape of chromatin, such as the formation of different chromatin globules, in agreement with previous studies[71]. In addition, our model predicts specific novel interactions that are not captured by 3C experiments because of primer design. A subset of such predicted interactions are found to be biological important as verified by an independent studies (Fig 8C)[77].

IV. CELL MODEL FOR TISSUE PATTERN FORMATION

Fully differentiated cells under epigenetic control grow, divide, and migrate to form tissue. Regulation of cell growth and cell-cell interactions plays fundamental roles in tissue formation, organ development, and cancer progression[78–81]. The coordinated efforts of a large number of cells to form an organ is a complex process that is not yet fully understood. Tissue formation occurs with precision and persistence, extending beyond individuals and even generations. We do not yet have the full picture of how changes in properties of individual cells such as cell size, shape, geometry, lineage, division, growth rate, and death affect tissue formation and the whole organism. Neither do we have sufficient information on how and when cell-cell interactions become important. It is challenging to design and conduct experimental studies to identify and differentiate specific effects of individual cell attributes and cell-cell interactions on tissue and organ development. Computational studies can generate useful hypothesis that can complement experimental studies in providing important insights.

A. Mechanical models of cell

Several computational methods have been developed to model tissue formation. These include the cellular Potts model, center-based model, finite element models and vertex models. Below we briefly summarize these methods (see Sëma et al.[82] for a detailed comparison).

The cellular Potts model is used for studying cell behavior, where each cell is modeled as a collection of about 25–50 lattice sites[83]. Each lattice site can be modeled as a square, triangle, or a hexagon. Size of the cells is constant and neighboring cells interact with specific binding energy, which models the underlying biology, *e.g.*, cadherin interactions[84]. Cell shape and topology are not modeled directly in the cellular Potts model and realistic modeling of cell shapes often requires extensive post-processing. Metropolis moves are often used to indirectly model cell proliferation. The state of a randomly picked lattice site is first proposed to change into that of its neighbor. This proposal is then evaluated for acceptance or rejection according to certain probability determined by energy calculation. Cell motions are also implemented as a result of energy minimization after stochastic fluctuations of flips of lattice sites introduced by Metropolis moves[84].

The center-based model approximates each cell by an isotropic, elastic, and adhesive sphere[85]. Cell

growth, division, and migration can be modeled by this model. It is specifically designed to study details of pair-wise cell forces where cellular interactions can be treated as interactions between homogeneous elastic sticky spheres[85]. However, detailed descriptions of cell shapes are not included, and any shape deviation from sphere is ignored[85]. As a result, center based models are not well-suited to study details of the biological problems that involve dynamic changes in cell shape and topology.

Finite element methods model cell shapes realistically but have inflexible boundary conditions and cannot model dynamic changes in cell shape such as cell growth, cell migration, cell birth, and cell apoptosis[86–88].

Vertex models minimizes the energy under forces acting on cell junctions (represented as vertices) to model cell shapes. These models are specifically designed to study packing and remodeling of epithelial cells[89,90]. They can be used to study cell birth, growth, migration, and apoptosis at varying degrees as they can realistically model cell size, shape, and elasticity[91–95]. However, there are some limitations in the current vertex models. For example, cells are always polygonal and do not have boundaries with curvature. Cell growth is also not modeled in detail. In addition, initial conditions require a plural number of cells (*e.g.*, $\geq$16 cells), often with periodic boundary conditions. Cell death can only be modeled for the special case associated with a specific type of topological change[90].

Chimeric methods such as the viscoelastic cell model by Jamali et al.[96] and the immersed boundary framework by Rajniak[97] can model realistic cell shape, cell growth, cell division, cell motion, and cell-cell interactions. However, they are not suitable for simulating large tissues due to the model choice of representing the shape of a cell by a network of linear Voigt elements or by a collection of boundary points connected through linear springs, which leads to substantial computational overhead.

B. Dynamic vertex model of cells

We have recently developed a physical model of cell and a simulation algorithm that incorporates cell size, shape, lineage, growth rate, death rate, and different cell-cell interactions (Fig 9[82]). Our method can model monolayered tissue formation by following the growth process of either a single cell, or a group of cells with arbitrarily pre-arranged spatial relationship, unlike previous studies that must start with a specific pre-existing cellular pattern[90,92,95]. Our method can model cell proliferation and programmed cell death without the constraints from unrealistic boundary conditions and can incorporate different growth rates due to the effect of specific growth factors. Moreover, our model represents the geometry of cells more accurately, with inner cells treated as polygons and outer cells as disc segments, as seen in *in vivo* studies[98].

In our method, we assume force equilibrium and adopt small step sizes in cell volume changes. Time-specific, cell-type specific, and location specific parameter values for cell growth or shrinkage and tissue mechanical properties can be introduced. Our implementation of the model using the *Half-Edge* data structure is very efficient and robust, and provides additional benefit of no overhead with regards to time and storage in maintaining the list of neighboring cells for each cell. We were able to simulate tissues with a large number of cells ($\sim$20,000) in a very short amount of time (*e.g.*, $<$1 hour)[82,99].

C. Pattern formation of tissue

We have used our dynamic vertex model[82,99] to examine the mechanisms regulating topological changes in tissue formation. Specifically, we studied the effects of orientation of division plane, differential proliferation, and mechanical forces on animal epithelial cells[100,101]. By incorporating cell rearrangements, we succeeded in reproducing the commonly observed topological distributions of cells in natural proliferating animal epithelial, regardless of the orientation of division plane. Different schemes

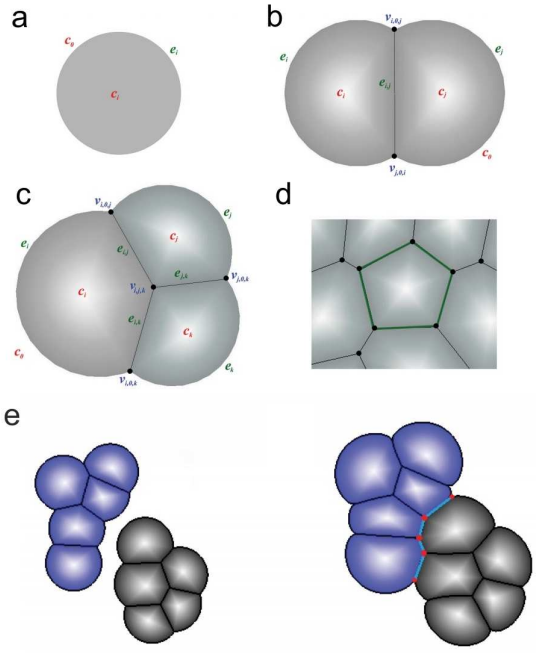

FIG. 9: Dynamic vertex model: (a) An isolated cell is modeled as a disk. (b) A cell is modeled as a disk segment when contacting other cell(s). An outer edge e_i is an arc or a circle, representing the boundary between cell c_i and the outside medium (denoted as c_0). An inner edge $e_{i,j}$ occurs when a cell c_i is in contact with another cell c_j. Their shared boundary is modeled as a straight line segment. When two cells c_i and c_j make contact, their outer edges (arcs) e_i and e_j intersect at two vertices $v_{i,0,j}$ and $v_{j,0,i}$ which are also the two end-points of the inner edge $e_{i,j}$ (c) When three cells c_i, c_j and c_k intersect, they form a vertex $v_{i,j,k}$. (d) A cell completely surrounded by other cells is represented as a polygon. (e) Fusion of two growing tissues. New edges and vertices formed are highlighted.

of division plane generated different frequency of hexagonal cells, consistent with experimental observations. Moreover, in our simulations of proliferating cells interfacing quiescent cells, we found that differential proliferation and increased tension on the boundary together reproduced the observed topological changes. Our simulation showed that both division plane orientation and mechanical forces play important roles in cell topology in animal proliferating epithelial[100,101].

We have also used our dynamic vertex model to study the mechanisms of regulating tissue elongation in drosophila wing[102]. Our simulation on the effects of directional cues and reduced cell size on tissue elongation indicated that oriented cell divisions and oriented mechanical forces act as directional cues during tissue elongation. In addition, we found that reduced cell size may significantly promote tissue elongation in conjunction with the directional cues. Our findings suggest that cell divisions without cell growth play essential roles in tissue elongation[102].

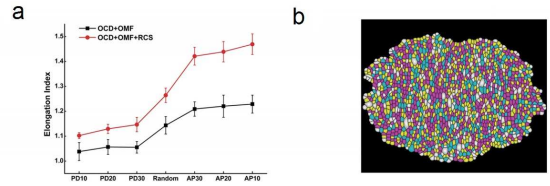

FIG. 10: Simulation of tissue elongation. (a) Oriented cell divisions drive tissue elongation, but only to a limited extent (black). Reduced cell size, when combined with oriented cell divisions, enhances tissue elongation (red). (b) Morphology of the elongated tissue (see reference[102]) for details.

We have also used our dynamic vertex model to study the effects of mechanical properties on tumor invasion[103]. Our simulations show increased adhesion to extra-cellular matrix (ECM) and decreased adhesion among tumor cells result in invasive tumor behavior. Moreover, our simulations show increased stiffness and increased degree of degradation of ECM promote tumor invasion, resulting in invasive tumor morphologies[103]. We have further improved this vertex based model into a stochastic spatial dynamic model by incorporating the inhibition growth rate, proliferation and differentiation probabilities of individual cells through feedback loops controlled by secreted factors from neighboring cells[104]. Our simulations show that with proper strengths of inhibition to growth and stem cell divisions, the tissue is capable of achieving a homeostatic size control[104].

D. Dynamic finite element model and wound healing

Built upon the dynamic vertex model, we have further developed a novel cell model approach based on dynamic finite element method (dFEMC) that can model explicitly cells with realistic cell elasticity, cell-cell interactions, and can model complex cell shapes and their dynamic changes during cell proliferation, cell migration and cell apoptosis. In addition, it can be used to study a large population of cells in the tissue [105].

In dFEMC, we assume cells have linear elasticity [106]. Both external and internal stresses can deform the cell. Cell deformation in turn can generate internal stresses. Following standard elastic models [107], we use strain tensor to measure the local deformation of the cell. Denote the displacement of a point $\boldsymbol{x}$ in cell as $\boldsymbol{u}(\boldsymbol{x}) = [u_1(\boldsymbol{x}), u_2(\boldsymbol{x})]^T$, the strain tensor ϵ of $\boldsymbol{x}$ can be defined as $\epsilon = [\epsilon_{x_1}, \epsilon_{x_2}, \epsilon_{x_2x_1}]^T = B\boldsymbol{u}$. Taking into account the properties of cell material, the stress tensor σ of internal stress in a cell in response to cell deformation can be written as $\sigma = D\epsilon$. With the strain and stress tensors defined, the elastic strain energy of a cell, along with work done by external forces $(\boldsymbol{x})$ can be written as $E(\boldsymbol{u}) = \frac{1}{2}\int_\Omega \boldsymbol{u}^T B^T DB\boldsymbol{u}d\boldsymbol{x} - \int_\Omega f^T \boldsymbol{u}dV$.

In our model, the boundary $\partial\Omega$ of a cell Ω with specific shape is represented by a set of discrete points $\partial\Omega = \{\boldsymbol{x} = [x_1, x_2]^T\}$. Triangular mesh tilling up the body of the cell Ω is then constructed using the farthest point sampling method based on Delaunay triangulations [108], which generates internal discrete point set $\{\text{Int}(\Omega) : \boldsymbol{x} = [x_1, x_2]^T\}$ and partitions the cell domain into a set of triangular elements $\{T_\Omega : T_e = [\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3]^T : \boldsymbol{x}_i \in \partial\Omega \cup \text{Int}(\Omega))\}$ (Figure 11a and 11b).

Each triangular element T_e is defined by three nodes either from $\partial\Omega$ or $\text{Int}(\Omega)$. The strain energy of each T_e under external force and/or internal forces is at its minimum when $\partial E(\boldsymbol{u})/\partial\boldsymbol{u} = 0$. This can be used to obtain the equilibrium solution of displacement describing cell deformation, which can be written as $K^e u^e = f^e$, where K^e is the element stiffness matrix of T_e, u^e the displacement vector of nodes in T_e, f^e the discredited force vector exerted on nodes in T^e. Upon assemblage of all elementary stiffness matrix into a global stiffness matrix, we obtain a large sparse linear system $Ku = f$. The behavior of the whole cell system at a specific time then can be simulated by solving this linear equation.

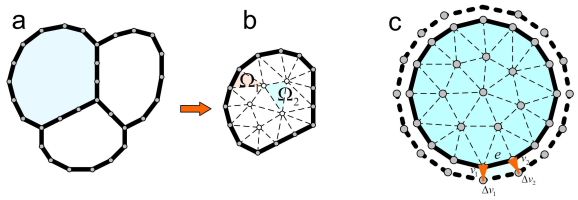

FIG. 11: Geometric democratization of cells: (a) An example of a toy model of a tissue consisting of 3 cells. The boundary of each cell is represented by a set of vertices. (b) Each cell is tiled by a triangular mesh generated using the farthest point sampling method based on Delaunay triangulation [108]. (c) Cell grows incrementally, with volume change attributed to individual boundary element that sum to ΔV (adapted from references [82,105]).

In many physiological process, cells migrate under chemical and mechanical cues. For example, during the re-epithelialization process in wound healing, cells migrate to the wound bed and repair the tissue. A rapidly moving cell often adapts its shape dynamically and may completely rebuild its cytoskeleton and adhesive structures during the migration process [110]. To gain understanding of the influence of intercellular adhesions and transmission of mechanical signal on the collective migration of cells, it is important to take into account dynamic changes in cell shape, along with changes in cell behavior associated with mechanical signal transmission.

We have applied our dFEMC method to study cell proliferation, apoptosis and tissue fusion [105]. We were able to reproduce suppression of tumor tissue growth under the treatment of curcumin and DMSO which influences the growth rate of cancer cells, as observed in clinical experiment [105,111,112]. Differential suppression effects of curcumin and DMSO treatments from our simulation were consistent with experimental observations [109], where curcumin exhibited stronger suppression effect on cancer cells

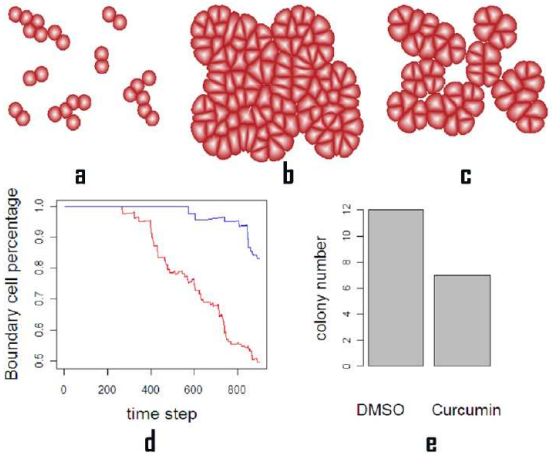

FIG. 12: Cancer cell growth under suppression induced by antitumor agents. (a). The initial state of cell growth. (b). Tumor tissue after 800 time steps using growth parameter modeling the effects of suppression from DMSO. (c). Tumor tissue after 800 time steps using growth parameter modeling the effects of suppression from curcumin. (d). The percentage of boundary cells, namely, the number of cell on boundary of the tissue over the total cell number, at each time step for DMSO treatment (red) and for curcumin treatment (blue). The higher boundary cells percentage corresponds to a more scattered cell population. (e). The colony number for DMSO and curcumin treatments. We count every 9 cells as one colony. The colony number ratio between the two treatments, 12 against 7, is consistent with the colony number ratio between the two treatments found in experiment studies[109] (see reference[105] for details).

than DMSO agent (Figure 12).

The dFEMC method can be further applied to study skin wound healing, a complex process involving many different cell types, signal transduction networks, as well as mechanical cell-cell interactions (Fig 13). It can help to delineate the differential roles of chemical signal and mechanical force in regulating processes such as re-epithelialization and granulation during wound healing, and can identify the specific effect of cytokines such as TGF-β in reconstruction of granulation tissue during wound healing.

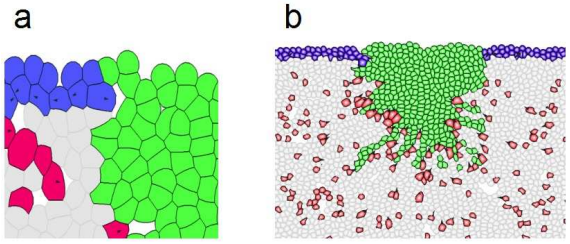

FIG. 13: Application of the dFEMC model to study wound healing. (a) Detailed view of the wound tissue including multiple cell types and clot elements. Green: wound element; Blue: keratinocyte cells; Red: fibroblast cells; Gray: elements of extra-cellular matrix. (b) A global view of a snapshot of a simulation of the process of wound healing.

V. CONCLUSIONS

Multiscale modeling of biological systems can aid in formulating theoretical models and can reveal fundamental insight into biological systems. It also has the promise to uncover new opportunities for disease treatment and health interventions. Previous computational studies at the molecular levels of proteins and RNAs have already generated important insight (not reviewed here, but see[113,114]). Computational techniques such as master equation treatment of protein folding[115] and sequential importance sampling of polymers can serve as the basis for developing effective computational techniques for studying stochastic network and for understanding chromosome folding[55–60]. The intricate geometry and topology of macromolecule with tens of thousands of atoms bears strong resemblance to problems encountered in studying cells and tissue[116,117]

The development of computational algorithms such as the ACME method for computing the probability landscape of a stochastic network discussed here is significant. With orders of magnitude reduction in the size of the state space, it enables accurate solution of the dCME for a large class of problems, whose solutions were previously unobtainable. For example, the state space of the MAPK network (Fig 5) can be reduced by 6–9 orders (*e.g.*, from 1.0 $\times 10^{16}$ to 6.2 $\times 10^{6}$), allowing a stochastic problem otherwise unsolvable to be computed on a desk-

top computer[39,40]. With explicit formulas for error bounds, the accuracy of the solution of a stochastic network can now be reliably estimated, and we gain the certainty of knowing if there are any probability peaks missing from our computational solution, hence no longer suffer from possible unknown unknowns. We can also assess rapidly the best possible accuracy we can achieve for a given computing environment. Such advancement has important implications, as we will be able to access hidden peaks and valleys of complex probabilistic landscapes, allowing quantitative models of epigenetic states corresponding to disease and healthy states, and the rare stochastic transitions between them to be developed and computed[118,119].

There is a large amount of data on chromatin interactions available from studies using chromosome conformation capture and related techniques. We now have access to detailed knowledge of pairwise interactions among genomic elements, many of which are important for nuclear activities. Computational construction of three-dimensional structures of chromatin can overcome significant limitations in these experimental techniques and can covert 2D mappings of genomic interactions into 3D spatial models of gene loci as well as the full genome. With genome-wide epigenetic studies pointing to potential enhancers along the 1D genome, 3D models of chromatins can facilitate interpretation and integration of epigenetic information, enabling mechanistic understanding of the temporal and spatial pattern of genome organization associated with cell differentiation. We expect future development of computational predictions of structural models of gene locus and global folding of genome will aid in the discovery of many novel higher order gene-enhancer interactions. Such knowledge will help to understand the control mechanism of the epigenetic states of stems cells and differentiated cells, with important implications in regenerative medicine.

Fully differentiated cells of different types grow, divide, interact, and migrate to form tissues with complex pattern. A fundamental question is how mechanical properties and chemical signaling specifically affect cell behavior and tissue formation. As mechanical stress exerted at cell-cell junctions and its transmission between cells cannot be easily measured experimentally, computational cell models such as those described here can play important roles in identifying the roles of mechanical cues and chemical cues in regulating the growth, division, and collective migration of different cell types. The dynamic finite element model (dFEMC) of cells discussed here can realistically model cell growth, death, migration, along with accurate description of cell shapes. It can also integrate effects of chemical network and mechanical forces, and can be used to study cellular pattern formation in complex processes such as wound healing, with the promise of generating new treatment strategies.

ACKNOWLEDGMENT

This work is supported by NIH grant GM079804, NSF grant MCB-1415589, and the Chicago Biomedical Consortium with support from the Searle Funds at The Chicago Community Trust.

REFERENCES

1. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001 Feb;409(6822):860–921.
2. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science. 2001 Feb;291(5507):1304–1351.
3. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012 Nov;491(7422):56–65.
4. Histone acetylation and an epigenetic code. Bioessays. 2000 Sep;22(9):836–845.
5. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. Science. 2001 Aug;293(5532):1068–1070.
6. Zhu XM, Yin L, Hood L, Ao P. Calculating biological behaviors of epigenetic states in the phage λ life cycle. Functional & Integrative Genomics. 2004;4(3):188–195.
7. Dodd IB, Micheelsen MA, Sneppen K, Thon G. Theoretical analysis of epigenetic cell memory by nucleosome modification. Cell. 2007;129(4):813–822.

8. Ptashne M. On the use of the word 'epigenetic'. Curr Biol. 2007 Apr;17(7):R233–6.
9. Liang J, Qian H. Computational Cellular Dynamics Based on the Chemical Master Equation: A Challenge for Understanding Complexity. J Computer Sci Tech. 2010;25(1):154–168.
10. Mateos-Langerak J, Bohn M, de Leeuw W, Giromus O, Manders EMM, Verschure PJ, et al. Proceedings of the National Academy of Sciences of the United States of America;.
11. Jhunjhunwala S, Van Zelm MC, Peak MM, Cutchin S, Riblet R, van Dongen JJ, et al. The 3D Structure of the Immunoglobulin Heavy-Chain Locus: Implications for Long-Range Genomic Interactions. Cell. 2008;133(2):265–279.
12. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science. 2002;295(5558):1306–1311.
13. Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326(5950):289–293.
14. Duan Z, Andronescu M, Schutz K, McIlwain S, Kim YJ, Lee C, et al. A three-dimensional model of the yeast genome. Nature. 2010;465(7296):363–367.
15. Arkin A, Ross J, McAdams HH. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. Genetics. 1998;149(4):1633–1648.
16. Morishita Y, Aihara K. Noise-reduction through interaction in gene expression and biochemical reaction processes. Journal of Theoretical Biology. 2004;228(3):315–325.
17. Kuwahara H, Gao X. Stochastic effects as a force to increase the complexity of signaling networks. Sci Rep. 2013;3:2297.
18. McAdams HH, Arkin A. Stochastic mechanisms in gene expression. Proceedings of the National Academy of Sciences of the United States of America. 1997;94(3):814–819.
19. Hasty J, Pradines J, Dolnik M, Collins JJ. Noise-based switches and amplifiers for gene expression. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(5):2075–2080.
20. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. Regulation of noise in the expression of a single gene. Nat Genet. 2002;31(1):69–73.
21. Levin MD. Noise in gene expression as the source of nongenetic individuality in the chemotactic response of Escherichia coli. FEBS Letters. 2003;550(1-3):135–138.
22. Ptashne M. A Genetic Switch: Phage Lambda Revisited. Cold Spring Harbor Laboratory Press; 3 edition; 2004.
23. Cao Y, Lu HM, Liang J. Probability landscape of heritable and robust epigenetic state of lysogeny in phage lambda. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(43):18445–18450.
24. Schultz D, Jacob EB, Onuchic JN, Wolynes PG. Molecular level stochastic model for competence cycles in Bacillus subtilis. Proceedings of the National Academy of Sciences. 2007;104(45):17582–17587.
25. Gillespie DT. Exact stochastic simulation of coupled chemical reactions. Journal of Physical Chemistry. 1977;81:2340–2361.
26. Van Kampen NG. Stochastic processes in physics and chemistry. Amsterdam: North Holland; 1992.
27. Cao Y, Gillespie DT, Petzold LR. The slow-scale stochastic simulation algorithm. J Chem Phys. 2005;122(1):14116.
28. Kuwahara H, Mura I. An efficient and exact stochastic simulation method to analyze rare events in biochemical systems. The Journal of Chemical Physics. 2008;129:165101.
29. Daigle BJ, Roh MK, Gillespie DT, Petzold LR. Automated estimation of rare event probabilities in biochemical systems. The Journal of Chemical Physics. 2011;134:044110.
30. Roh MK, Daigle BJ, Gillespie DT, Petzold LR. State-dependent doubly weighted stochastic simulation algorithm for automatic characterization of stochastic biochemical rare events. Journal of Chemical Physics. 2011;135(23):234108.
31. Cao Y, Liang J. Adaptively Biased Sequential Importance Sampling for Rare Events in Reaction Networks with Comparison with Exact Solutions from Finite Buffer dCME Method. J Chem Phys. 2013;In press:x.
32. Munsky B, Khammash M. The finite state projection algorithm for the solution of the chemical master equation. The Journal of Chemical Physics. 2006;124(4):044104.
33. Cao Y, Liang J. Optimal enumeration of state space of finitely buffered stochastic molecular networks and exact computation of steady state landscape probability. BMC Systems Biology. 2008;2(1):30.
34. MacNamara S, Bersani AM, Burrage K, Sidje RB. Stochastic chemical kinetics and the total quasi-steady-state assumption: application to the stochastic simulation algorithm and chemical master equation. The Journal of chemical physics. 2008;129(9):095105.
35. Wolf V, Goel R, Mateescu M, Henzinger T. Solving the chemical master equation using sliding windows. BMC Systems Biology. 2010;4(1):42.
36. Jahnke T. On reduced models for the chemical master equation. Multiscale Modeling & Simulation. 2011;9(4):1646–1676.

37. Sidje RB. Expokit: a software package for computing matrix exponentials. ACM Transactions on Mathematical Software (TOMS). 1998;24(1):130–156.
38. Munsky B, Khammash M. A multiple time interval finite state projection algorithm for the solution to the chemical master equation. Journal of Computational Physics. 2007;226(1):818 – 835.
39. Cao Y, Terebus A, Liang J. State space truncation with quantified effects for accurate solution of discrete Chemical Master Equation. Bulletin Math Biology. 2016;.
40. Cao Y, Terebus A, Liang J. Accurate Chemical Master Equation Solution Method with Multi-Finite Buffers for Time-Evolving and Steady State Probability Landscapes and First Passage Times. SIAM: Multiscale Modeling and Simulation. 2016;.
41. Alon U. An introduction to systems biology: design principles of biological circuits. Chapman and Hall/CRC press, Mathematical and Computational Biology (Book 10); 2006.
42. Siegal-Gaskins D, Grotewold E, Smith GD. The capacity for multistability in small gene regulatory networks. BMC Systems Biology. 2009;3(1):1.
43. Duncan A, Liao S, Vejchodskỳ T, Erban R, Grima R. Noise-induced multistability in chemical systems: Discrete versus continuum modeling. Physical Review E. 2015;91(4):042111.
44. Terebus A, Cao Y, Liang J. Exact computation of probability landscape of stochastic networks of Single Input and Coupled Toggle Switch Modules. In: Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE. IEEE; 2014. p. 5228–5231.
45. Tyson JJ, Novák B. Functional motifs in biochemical reaction networks. Annual review of physical chemistry. 2010;61:219.
46. Mangan S, Itzkovitz S, Zaslaver A, Alon U. The incoherent feed-forward loop accelerates the response-time of the gal system of Escherichia coli. Journal of molecular biology. 2006;356(5):1073–1081.
47. Kashtan N, Itzkovitz S, Milo R, Alon U. Topological generalizations of network motifs. Physical Review E. 2004;70(3):031909.
48. Johnson AD, Poteete AR, Lauer G, Sauer RT, Ackers GK, Ptashne M. λ Repressor and cro–components of an efficient molecular switch. Nature. 1981;294(5383):217–223.
49. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012;485(7398):381–385.
50. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485(7398):376–380.
51. Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell. 2013;153(6):1281–1295–385.
52. Fraser P, Bickmore W. Nuclear organization of the genome and the potential for gene regulation. Nature. 2007;447(7143):413–417.
53. Barbieri M, Chotalia M, Fraser J, Lavitas LM, Dostie J, Pombo A, et al. Complexity of chromatin folding is captured by the strings and binders switch model. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(40):16173–16178.
54. Gürsoy G, Xu Y, L KA, Liang J. Spatial Confinement is a major Determinant of Folding Landscape of Human Chromosomes and Genetic Programming of Cell. Nucleic Acids Research. 2014;42(13):8223–8230.
55. Zhang J, Chen R, Tang C, Liang J. Origin of scaling behavior of protein packing density: A sequential Monte Carlo study of compact long chain polymers. J Chem Phys. 2003;118:6102–6109.
56. Zhang J, Chen Y, Chen R, Liang J. Importance of chirality and reduced flexibility of protein side chains: A study with square and tetrahedral lattice models. J Chem Phys. 2004;p. 592–603.
57. Zhang J, Lin M, Chen R, Liang J. Discrete State Model and Accurate Estimation of Loop Entropy of RNA Secondary Structures. J Chem Phys. 2008;128(125107):1–10.
58. Lin M, Zhang J, Lu HM, Chen R, Liang J. Constrained proper sampling of conformations of transition state ensemble of protein folding. J Chem Phys. 2011 Feb;134(7):075103.
59. Tang K, Zhang J, Liang J. Fast protein loop sampling and structure prediction using distance-guided sequential chain-growth Monte Carlo method. PLoS Comput Biol. 2014 Apr;10(4):e1003539.
60. Tang K, Wong SW, Liu JS, Zhang J, Liang J. Conformational sampling and structure prediction of multiple interacting loops in soluble and -barrel membrane proteins using multi-loop distance-guided chain-growth Monte Carlo method. Bioinformatics. 2015 Aug;31(16):2646–2652.
61. Gürsoy G, Xu Y, Liang J. computational predictions of structures of multichromosome of budding yeast. Conf Proc IEEE Eng Med Biol Soc. 2014;Accepted.
62. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485(7398):376–380.
63. Kang H, Yoon Y, Thirumalai D, C H. Confinement-Induced Glassy Dynamics in a Model for Chromosome Organization. PRL. 2015;115(19):198102–198107.
64. Taddei A, Gasser S. Structure and function in the budding

yeast nucleus. Genetics. 2012;192(1):107–129.
65. Tjong H, Gong K, Chen L, Alber F. Physical tethering and volume exclusion determine higher-order genome organization in budding yeast. Genome research. 2012 Jul;22(7):1295–1305.
66. Wong H, Marie-Nelly H, Herbert S, Carrivain P, Blanc H, Koszul R, et al. A predictive computational model of the dynamic 3D interphase yeast nucleus. Current Biol. 2012;22(20):1881–1990.
67. Ay F, Bunnik EM, Varoquaux N, Bol SM, Prudhomme J, Vert JP, et al. Three-dimensional modeling of the P. falciparum genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. Genome Research. 2014;24:974–988.
68. Zhang B, G WP. Topology, structures, and energy landscapes of human chromosomes. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(19):6062–6067.
69. Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, Dekker J, et al. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. Cell. 2014;157(4):950–963.
70. Rousseau M, Fraser J, Ferraiuolo MA, Dostie J, Blanchette M. Three-dimensional modeling of chromatin structure from interaction frequency data using Markov chain Monte Carlo sampling. BMC Bioinformatics. 2011;12(414).
71. Baù D, Sanyal A, Lajoie BR, Capriotti E, Byron M, Lawrence JB, et al. The three-dimensional folding of the α-globin gene domain reveals formation of chromatin globules. Nature structural & molecular biology. 2010;18(1):107–114.
72. Tokuda N SM Terada T P. Dynamical modeling of three-dimensional genome organization in interphase budding yeast. Biophysical Journal. 2012;102(2):296–304.
73. Wang XJ S, J Z. Inferential modeling of 3D chromatin structure. Nucleic Acids Res. 2015;43(8):e54.
74. Trieu T, Cheng J. Large-scale reconstruction of 3D structures of human chromosomes from chromosomal contact data. Nucleic Acids Res. 2014;42(7):e52.
75. S BA. Large-scale chromatin organization: the good, the surprising, and the still perplexing. Curr Opin Cell Biol. 2014;26:69–78.
76. Ay F, Noble WS. Analysis methods for studying the 3D architecture of the genome. Genome Biol. 2015;16:183.
77. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell. 2012;148(1-2):84–98.
78. Shraiman BI. Mechanical feedback as a possible regulator of tissue growth. Proc Natl Acad Sci U S A. 2005 Mar;102(9):3318–3323.
79. Collinet C, Rauzi M, Lenne PF, Lecuit T. Local and tissue-scale forces drive oriented junction growth during tissue extension. Nat Cell Biol. 2015 Oct;17(10):1247–1258.
80. Hariharan IK. Organ Size Control: Lessons from Drosophila. Dev Cell. 2015 Aug;34(3):255–265.
81. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011 Mar;144(5):646–674.
82. Kachalo S, Naveed H, Cao Y, Zhao J, Liang J. Mechanical model of geometric cell and topological algorithm for cell dynamics from single-cell to formation of monolayered tissues with pattern. PLoS One. 2015;10(5):e0126484.
83. Marée AFM, Grieneisen VA, Hogeweg P. II.2 The cellular Potts model and biophysical properties of cells, tissues and morphogenesis. In: Single-cell-based models in biology and medicine. Basel/Switzerland: Birkhäuser Verlag; 2007. p. 107–136.
84. Merks RM, Perryn ED, Shirinifard A, Glazier JA. Contact-inhibited chemotaxis in de novo and sprouting blood-vessel growth. PLoS Comput Biol. 2008;4(9):e1000163.
85. Hoehme S, Drasdo D. A cell-based simulation software for multi-cellular systems. Bioinformatics. 2010;26(20):2641–2642.
86. Honda H, Tanemura M, Yoshida A. Differentiation of wing epidermal scale cells in a butterfly under the lateral inhibition model-appearance of large cells in a polygonal pattern. Acta biotheoretica. 2000;48(2):121–136.
87. Honda H, Tanemura M, Yoshida A. Estimation of neuroblast numbers in insect neurogenesis using the lateral inhibition hypothesis of cell differentiation. Development. 1990;110(4):1349–1352.
88. Viens D, Brodland GW. A three-dimensional finite element model for the mechanics of cell-cell interactions. J Biomech Eng. 2007 Oct;129(5):651–657.
89. Staple D, Farhadifar R, Röper JC, Aigouy B, Eaton S, Jülicher F. Mechanics and remodelling of cell packings in epithelia. The European Physical Journal E: Soft Matter and Biological Physics. 2010;33(2):117–127.
90. Farhadifar R, Röper JC, Aigouy B, Eaton S, Jülicher F. The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing. Current Biology. 2007;17(24):2095–2104.
91. Nagai T, Kawasaki K, Nakamura K. Vertex dynamics of two-dimensional cellular patterns. J Phys Soc Jpn. 1988;57:2221–24.
92. Nagai T, Honda H. Computer simulation of wound closure in epithelial tissues: Cell˘basal-lamina adhesion. Phys Rev E. 2009 Dec;80:061903. Available from: http://link.aps.org/doi/10.1103/PhysRevE.80.061903.
93. Landsberg KP, Farhadifar R, Ranft J, Umetsu D, Widmann TJ, Bittig T, et al. Increased cell bond tension governs cell sorting at the Drosophila anteroposterior compartment

boundary. Curr Biol. 2009 Dec;19(22):1950–1955.

94. Aegerter-Wilmsen T, Smith AC, Christen AJ, Aegerter CM, Hafen E, Basler K. Exploring the effects of mechanical feedback on epithelial topology. Development. 2010 Feb;137(3):499–506.
95. Aigouy B, Farhadifar R, Staple DB, Sagner A, Roper JC, Julicher F, et al. Cell flow reorients the axis of planar polarity in the wing epithelium of Drosophila. Cell. 2010 Sep;142(5):773–786.
96. Jamali Y, Azimi M, Mofrad MR. A sub-cellular viscoelastic model for cell population mechanics. PLoS One. 2010;5(8).
97. Rejniak KA. An immersed boundary framework for modelling the growth of individual cells: an application to the early tumour development. J Theor Biol. 2007 Jul;247(1):186–204.
98. Hayashi T, Carthew RW. Surface mechanics mediate pattern formation in the developing retina. Nature. 2004 Oct;431(7009):647–652.
99. Naveed H, Li Y, Kachalo S, Liang J. Geometric order in proliferating epithelia: impact of rearrangements and cleavage plane orientation. In: Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE. IEEE; 2010. p. 3808–3811.
100. Li Y, Naveed H, Kachalo S, Xu LX, Liang J. Mechanical forces mediate localized topological change in epithelia. In: Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE. IEEE; 2011. p. 178–181.
101. Li Y, Naveed H, Kachalo S, Xu LX, Liang J. Mechanisms of Regulating Cell Topology in Proliferating Epithelia: Impact of Division Plane, Mechanical Forces, and Cell Memory. PLoS ONE. 2012 08;7(8):e43108.
102. Li Y, Naveed H, Kachalo S, Xu LX, Liang J. Mechanisms of Regulating Tissue Elongation in Drosophila Wing: Impact of Oriented Cell Divisions, Oriented Mechanical Forces, and Reduced Cell Size. PLoS ONE. 2014 02;9(2):e86725.
103. Li Y, Naveed H, Liang J, Xu LX. Effects of mechanical properties on tumor invasion: insights from a cellular model. Conf Proc IEEE Eng Med Biol Soc. 2014;2014:6818–6821.
104. Cao Y, Liang C, Naveed H, Li Y, Chen M, Nie Q. Modeling spatial population dynamics of stem cell lineage in tissue growth. Conf Proc IEEE Eng Med Biol Soc. 2012;2012:5502–5505.
105. Zhao J, Naveed H, Kachalo S, Cao Y, Tian W, Liang J. Dynamic mechanical finite element model of biological cells for studying cellular pattern formation. Conf Proc IEEE Eng Med Biol Soc. 2013;2013:4517–4520.
106. Bischofs I, Safran S, Schwarz U. Elastic interactions of active cells with soft materials. Physical Review E. 2004;69(2):021911.
107. Fung Yc. Foundations of solid mechanics. Prentice Hall; 1965.
108. Moenning C, Dodgson NA. Fast marching farthest point sampling for implicit surfaces and point clouds. Computer Laboratory Technical Report. 2003;565.
109. Tu SP, Jin H, Shi JD, Zhu LM, Suo Y, Lu G, et al. Curcumin induces the differentiation of myeloid-derived suppressor cells and inhibits their interaction with cancer cells and related tumor growth. Cancer prevention research. 2012;5(2):205–215.
110. Keren K, Pincus Z, Allen GM, Barnhart EL, Marriott G, Mogilner A, et al. Mechanism of shape determination in motile cells. Nature. 2008;453(7194):475–480.
111. Jagetia GC, Aggarwal BB. Spicing up of the immune system by curcumin. Journal of clinical immunology. 2007;27(1):19–35.
112. Bertelli G, Gozza A, Forno G, Vidili M, Silvestro S, Venturini M, et al. Topical dimethylsulfoxide for the prevention of soft tissue injury after extravasation of vesicant cytotoxic drugs: a prospective clinical study. Journal of Clinical Oncology. 1995;13(11):2851–2855.
113. Xu Y, Xu D, Liang J, editors. Computational Methods for Protein Structure Prediction and Modeling: Basic Characterization. vol. 1 of Editor-in-Chief: Elias Greenbaum. 233 Spring Street, New York, NY 10013, USA: Springer Science; 2006.
114. Liang J. Experimental and computational studies of the determinants of membrane protein folding. Curr Opin Chem Biol. 2002;6:878–884.
115. Kachalo S, Lu HM, Liang J. Protein Folding Dynamics via Quantification of Kinematic Energy Landscape. Phys Rev Lett. 2006;96(5):058106.
116. Li X, Hu C, Liang J. Simplicial edge representation of protein structures and alpha contact potential with confidence measure. Proteins. 2003;53:792–805.
117. Li X, Liang J. Geometric cooperativity and anticooperativity of three-body interactions in native proteins. Proteins: Structure, Function, and Bioinformatics. 2005;60(1):46–65.
118. Ao P, Kown C, Qian H. On the existence of potential landscape in the evolution of complex systems. Complexity. 2007;12:19–27.
119. Ao P. Stochastic force defined evolution in dynamical systems. Arxiv preprint physics/0302081. 2003;.