

## Rho-GTP アーゼによるキナーゼコンピューティング: 解析とシミュレーション

劉 健勤 下原勝憲

(株)国際電気通信基礎技術研究所 先端情報科学研究部  
京都府相楽郡精華町光台2-2-2、619-0288

**あらまし:** 新しい分子計算パラダイムとしてキナーゼコンピューティングを提案し、その理論的な解析とシミュレーションを行った。さらに生体分子による実装を目指し、計算論的モレウェアについて考案した。

**キーワード** 分子計算、キナーゼコンピューティング、LSシステム。

## Kinase Computing in Rho-GTPases: Analysis and Simulation

Jian-Qin LIU and Katsunori SHIMOHARA

Information Sciences Division, ATR International  
Hikaridai 2-2-2, Seika-cho, Soraku-gun, Kyoto, 619-0288, JAPAN

**Abstract:** This paper presents our recent research results on kinase computing – a new class of molecular computing – by theoretical analysis and simulated synthesis experiments. The corresponding molecular biological issues on potential computational moleware *in situ* are also discussed.

**Keywords:** molecular computing, kinase computing, LS-systems.

### 1. はじめに

次世代コンピュータの創出を目指す分子計算の分野において、分子エレクトロニクスを用いたシリコンの限界を超える速度の達成など、従来にない計算原理が注目されている。例えば、DNA、RNA、たんぱく質など生体分子を利用した生体分子計算の基本機構としてセルフアセンブリー（自己組織化）のメカニズムがある[1,2]。しかし単一分子レベルでの操作コストが高く、しかも、例えば  $10^{15}$  個 DNA 分子など、巨大な分子数の合成は非常に難しいのが現状である [3]。

従って、細胞内のみならず細胞間の信号処理を扱う能力は未だ生体分子計算システムでは実装されていない。

本論文では、細胞内と細胞間通信の信号伝達経路を用いた新しい計算パラダイムとして“キナーゼコンピューティング”を提案する。そして、チューリングマシンと等価な計算モデルを用いて形式的な解析を行う。また、動的未知環境下でのロバストネス（頑強性）の確保に必要な機能を創り出す“パスウェイ書き換え系”という新し

い書き換え系に基づく“LSシステム”を提案し、動作検証のシミュレーションを行う。また、データ構造やオブジェクト指向プログラミング技法など中間的なモデリング・アルゴリズムを検討する。

### 2. LSシステム

最初にチューリングマシンと等価な計算モデルを考える。LSシステムを

$\langle V, T, D, U, E, Y, Z, PTs, Q \rangle$

の九項目により定義する。ここで、

V: アルファベット集合;

T: 終端記号の集合(Vの部分集合);

D:  $\{1,0\}$ の集合;

U: 頂点集合(Vの部分集合);

E: 辺の集合;

Y:  $\langle HE, U \rangle$ の形をもつハイパーグラフの集合; そのうち HE は超辺の集合;

HR と VR という操作を行う;

Z: 局所濃度(Vの部分集合);

PTs: 経路いわゆるパスウェイの集合。

[経路] $::=[経路] \cup$  [二つ頂点を持つ単一超辺]  
s.t. Q の操作

ここで、  
Q : HE、V、E によるルール:

$$Q = \{Q_i\}$$

$i$  は有限の自然数。

条件付きの  $Q$  を用いて異構造の複数経路集合から成る異構造  $LS$  システムができる。

“キナーゼコンピューティング”のコンセプトは、 $LS$  システムに基づく細胞とその中核としてのキナーゼを自己調整できる信号経路を使うことである。生体分子計算の新しいパラダイムとして、以下の三つの重要な利点が考えられる:

- (1) 複数の計算対象クラス間のダイナミックな相互作用を自動的に調整するメカニズムを創発できる;
- (2) カテゴリーに基づく厳密な計算理論を備えるシステムにより万能計算能力を保證できる;
- (3)  $NP$  問題解決のために要する時間はシステム内部の複雑さに依存するが、それを動的操作によって線形に抑えることができる。

なお、一つ一つの  $DNA$  分子の自己組織化原理により分子システムを実装すると高いコストを要する。このコストを削減するために、本論文では細胞中の信号経路を支える分子集団の生化学特性を利用した方法を提案する。この方法により実装できる可能性が高まることが期待できる。また、生体分子固有の機能やメカニズムに基づく  $OS$  などのシェルシステムや自動的なツールを用いて、汎用システムとしての生体分子コンピュータの開発も可能となるであろう。

### 3. $LS$ システムの計算能力の検討

このモデルでは、外部から擾乱がある場合、グラフ集合からグラフ集合までのハイパーグラフの写像であり、かつ、論理矛盾のない文法の間で変換できる最小グラフとして一つ経路が存在する。そして、グラフ全体の表現とチューリングマシンによっ

て選択された頂点群の関係が論理的に証明できる  $rLS$  システムを以下に示す。

**定理 1**  $rLS$  モデル  $\Phi$  はチューリングマシンと同様な文字系列 (ストリング) でのテープ読取能力を持つ。

**定理 2**  $rLS$  モデル  $\Phi$  は動的擾乱がある環境においてもチューリングマシンの計算ができる。

**証明(スケッチ):**

外部からの擾乱因子を  $\delta(t)$  と仮定する。モデル  $\Phi$  の素子群集団を相互作用させることにより、次のような経路終端記号集合を得る。

$$PT(t) = \{PT_p(t) \mid \delta(t) \text{ に対する } \Delta\delta(t)\}$$

これを補償する経路が存在する。擾乱がないときにはチューリングマシンの計算を

$$\beta_i \rightarrow \beta_{i+1}$$

とする、 $i$  は時間  $t$  の離散表示である、 $\beta_i$  と  $\beta_{i+1}$  はストリングとしての終端記号である。 $\beta_i$  と  $PT_p(t)$  が等しくなる前提のもと、

$$\beta_{i+1} \neq PT_p(t+1)$$

の場合、 $\beta_{i+1}$  を生み出す経路は、 $HR$  と  $VR$  のグラフ書き換えルールを有限回 ( $n+m$  回) 適用することにより

$$PT_p(t+1+m+n)' = \beta_{i+1}$$

と表現できる。

$(t+1+m+n)$  に対するストリングを  $i+1$  で再定義することが必要である。その操作を経路集合全体に有限回繰り返すことで、終端記号全体を書き換えるプロセスをチューリングマシンが認識できることを確認した。そしてチューリングマシンの計算論的定義において、モデル  $\Phi$  の書き換えプロセスは時間  $t$  に対して

$$\gamma_i \vdash \gamma_{i+1}$$

である。ここで  $\gamma_i, \gamma_{i+1}$  は終端記号  $V$  に属し、かつ、システムの状態を示す。従って関連酵素数は有限であり経路の複雑性の限界は  $2^n$  になる。経路の写像能力は  $N \rightarrow N$  の全空間をカバーする。

Q.E.D.

**定理 3**  $rLS$  モデル  $\Phi$  は終端記号に対する読取プロセスを停止できる。

このモデルから導出したアルゴリズムは、細胞中に生体分子集合を対象とするプログラムを作る。局所的な濃度分布は異なるが、複数の TM 操作のテープと等価な生体分子の濃度がダイナミックに変化する。よって、マルチ TM の機能を備える経路集合間の相互作用が実現され、複雑性の向上が達成された。

#### 4. キナーゼコンピューティングのシミュレーション

中間視的なプロテオーム計算論的化学の観点では、データ構造は特定のハイパーグラフで表現される経路となる。C 言語でオブジェクト指向型のシミュレーションシステムを試作した。ハイパーグラフで表現された経路の生成、生体細胞信号伝達経路の経路更新と頑健な経路書き換えのオペレータ集合をモジュールとして構築した。また、モジュール間の通信の信号パターンが異なる構造間でも同期するようにした。

入力から出力までの写像空間中に一定のトポロジー構造ができ、不変量を経路に関する機能測度で定義する。言わば一定入力から出力まで素子に対する経路変化を調整するのである。経路をランダムに区分すると頂点数を増やし、これに応じて辺の集合を最小化する。なお、大域的に最適化するアルゴリズムから並列性を得る。

提案モデルのシミュレーションにおいて“はしご”現象を発見した。図 1 は 50 変数 30 節の 3-SAT 問題に本手法を適用した結果である。系列 1 は本方法の適用結果であり、DNA 計算 (系列 3) の指数 ( $2^{50}$ ) と系列 2 の線形 (50) に比較して、最も小さい酵素数 (14 個) となり、系列 2 (50) の半分以下のコスト削減を達成できた。つまり提案モデルでは酵素に関する計算複雑性は

$O(n/2)$  である (X: 3-SAT 問題の節数、Y: 3-SAT 問題解決の分子数、1: 酵素数、2: DNA 数、3: DNA 数)。ここで分子数は  $[0,5][5,10],[10,15], \dots$  であり、それぞれの階段を使って指数関数的な増長が避けられ、一つ一つ分子を作らなくて良いという利点がある。

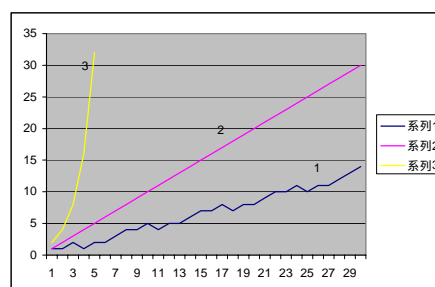


図 1 3-SAT 問題解決の分子数と節数の関係

#### 5. Rho-GTPase による実装への考え

ポストゲノム時代の分子生物学の急速な進展により、様々な新しい現象が発見されつつある。例えば、細胞内の複雑性に関する知見が生化学の理論から次々と明らかにされつつある。

貝淵、黒田、天野の Rho 族 GTPアーゼに関する (本文では Rho-GTP アーゼと表記する) 細胞信号伝達経路の研究成果 [4] は生命科学のフロンティア分野に大きく貢献すると考える。計算機科学的な観点では、GTPアーゼによる細胞生体分子信号伝達経路集合は、その相互作用のために、分子計算機の実装材料としての強い潜在力を有する。

GTPアーゼによる計算モデルの構造は、DNA 計算とは異なることが明らかになった。生体計算機の実装においては、このモデルでは、細胞信号伝達経路に関する Rho-GTP アーゼを用いる場合の細胞数、分子数、酵素およびキナーゼ数が有限となる。これに関するシミュレーションにより、その実現の可能性は明らかであることを確認した。即ち、Rho 族 GTP アーゼによる計算モデルは生体分子計算機実装の基礎となる有力なモデルである。

## 6 .Rho-GTP アーゼによる計算モデルのシミュレーションの例

図 2 にシミュレーション結果を示す。本シミュレーションにおいて、細胞数 3 個、入力分子数 30 個、出力 30 個、細胞中の分子数 (機能的なたんぱく質、酵素、キナーゼなど) 35 個、関連経路数 52500 個である。

図 2 (1)、(2)、(3)の系列 1 と 2 は通常の経路 A の入力と出力である。図 2 (1)、(2)、(3)の系列 4 と 5 は擾乱の経路 B の入力と出力である。図 2 (1)、(2)、(3)の系列 3 と、図 2 (3)の系列 6 はハイパーグラフ書き換えの経路(C1,C2)の入力と出力である。図 2 (1)では C1 の活性があり経路 A の特徴パターンは 0.000599 (600-) になる。図 2 (2)では C1 の活性が半分になるとき、経路 A の特徴パターンが 0.000507 (500+) になる。図 2 (3)に図 2 (2)C1 の活性に対する C 2 を導入するときの経路 A の特徴パターンは 0.000605 (600+) である。図 2 (1)と図 2 (2)の状態から図 2 (3)の状態まで VR と HR を操作した。これでは酵素数が制御できることを示す。

## 7 . 結論

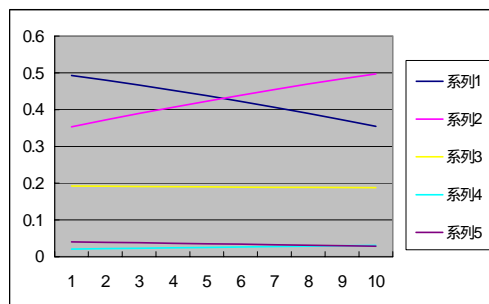
本論文ではLSシステムを提案し、優れた形式計算モデルであることを示した。アルゴリズム性能に関するシミュレーションでは50変数30節の3-SAT問題を解決した。ちなみにDNAコンピューティングによって解決された3-SAT問題の規模は、Qinghua Liuらの4変数4節[2]、坂本らの6変数10節[5]、Braichらの6変数11節[6]である。今後、本研究では生体分子による実装が重要な検討課題となる。

### 参考文献

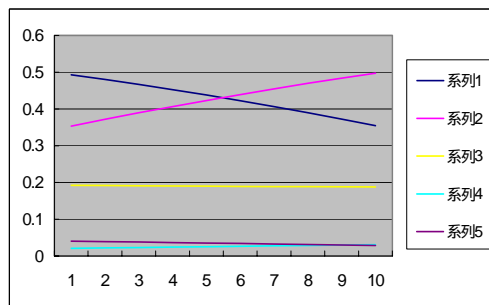
- [1] L.M.Adleman, Molecular computation of solutions to combinatorial problems, Science, 266, 1994; 1021-1024.
- [2] Qinghua Liu, Liman Wang, Anthony G. Frutos, Anne E. Condon, Robert M. Corn and Lloyd M. Smith, DNA computing on surfaces, Nature, 403, 2000; 175-179.
- [3] Mitsunori Ogihara and Animesh Ray, DNA computing on a chip, Nature, 403, 2000; 143-144.
- [4] Kozo Kaibuchi, Shinya Kuroda and

Mutsuki Amano, Regulation of the cytoskeleton and cell adhesion by the Rho family GTPases in mammalian cells, Annu. Rev. Biochem. 68, 1999; 459-486.

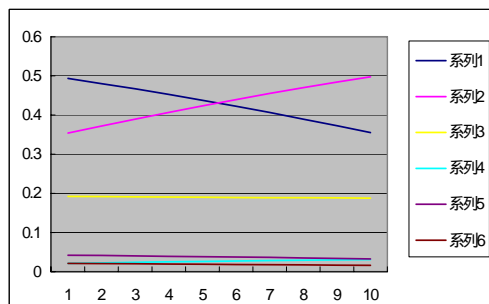
- [5] Kensuku Sakamoto, Hidetaka Gouzu, Ken Komiya, Daisuke Kiga, Shigeyuki Yokoyama, Takashi Yokomori and Masami Hagiya, Molecular computation by DNA hairpin formation, Science, 288, 2000; 1223-1226.
- [6] Ravinderjit S. Braich, Cliff Johnson, Paul W.K. Rothmund, Darryl Hwang, Nickolas Chelyapov, and Leonard M. Adleman, Solution of a satisfiability problem on a gel-based DNA computer, In A.Condon & G.Rozenberg, editors, Pre-proc. of DNA 6, 2000; 31-42.



(1)



(2)



(3)

図 2 経路の信号系列