

# タンパク質生成合成と数理モデル

清水 義宏<sup>1\*</sup>・松浦 友亮<sup>2</sup>

## はじめに

AnfinsenらはリボヌクレアーゼAに変性剤を加えて失活させた後に、変性剤を除去すると酵素活性が回復することを見だし、タンパク質の折り畳みがアミノ酸配列によって一義的に決まるとする、いわゆるAnfinsenのドグマを提唱した<sup>1)</sup>。しかしながら、変性剤の添加およびその除去といった実験は細胞内で起こるタンパク質の折り畳みとは大きく異なり、近年ではドグマから外れた例外的な事象が報告されている。細胞内では開始コドンに対応したメチオニン残基からスタートし、アミノ末端からカルボキシル末端の方向へ、リボソーム上にてアミノ酸の重合が行われる。合成される新生ポリペプチド鎖はリボソーム大サブユニット内にある細いトンネルを抜け、アミノ末端から順にリボソーム外へと送り込まれ、同時に重合中であっても、タンパク質の折り畳みが開始される(co-translational folding)。したがって、細胞内においては、ポリペプチド鎖の重合速度とタンパク質の折り畳みには密接な関係性が存在し、たとえば、概日時計タンパク質をコードする核酸配列を、アミノ酸配列は変えずにコドン使用頻度のみを変更することによってシステムの頑強さが失われるといった報告が近年なされている<sup>2,3)</sup>。

すなわち、細胞内における新生ポリペプチド鎖が折り畳まれ、その機能が発揮されるまでを深く理解するためには、新生ポリペプチド鎖が生成される過程であるタンパク質合成システム自体を深く理解し、その一次配列から重合速度などを予測可能なものにする必要がある。しかしながら、リボソーム上におけるポリペプチド鎖の重合は、tRNAのアミノアシル化にはじまり、開始複合体の形成、アミノアシルtRNAのリボソームへの結合、ペプチド転移反応、tRNAおよびmRNAのリボソーム上での転位反応、ペプチジルtRNAの加水分解にはじまる翻訳終結反応など、多数の生化学反応が織り成す大規模な多段階反応であり、非常に複雑な分子ネットワークを読み解く必要がある。こうした理由から、これまでもタンパク質合成の数理モデル化などが取り組まれてきたが、タンパク質合成全体のモデルとしては主に粗視化されたモデルに留まっている<sup>4,5)</sup>。

そこで、より精細なタンパク質合成のモデルを構築し、タンパク質合成全体のダイナミクスを高い精度で予測可能なシミュレータを開発することを目的に、筆者らは、再構成型の無細胞タンパク質合成システムであるPURE system<sup>6)</sup>をベースに、タンパク質合成反応の全成分シミュレータ(ePURE)を、常微分方程式(ordinary differential equation: ODE)を用いて構築した<sup>7)</sup>。本稿ではePUREの概要と、ePUREを用いて観察されたタンパク質合成反応のダイナミクスから得られる知見などについて紹介する。

## タンパク質合成反応全成分シミュレータの構築

前述したように、タンパク質合成反応はリボソームを基盤としたポリペプチド鎖重合反応以外にもアミノ酸をtRNAに付加させるアミノアシル化や、ATPなどのエネルギー源をADPやAMPなどから再生するエネルギー再生反応などを連携させることが必要である(図1A)。さらに、ポリペプチド鎖重合反応についても、開始・伸長・終結段階を経る複雑な反応であることが知られている。すべてのモデルを視認できる形で一括に記述することは不可能であるため、全体の反応を26のサブモデルに区画化し、個々のサブモデルを記述した後に全体のモデル

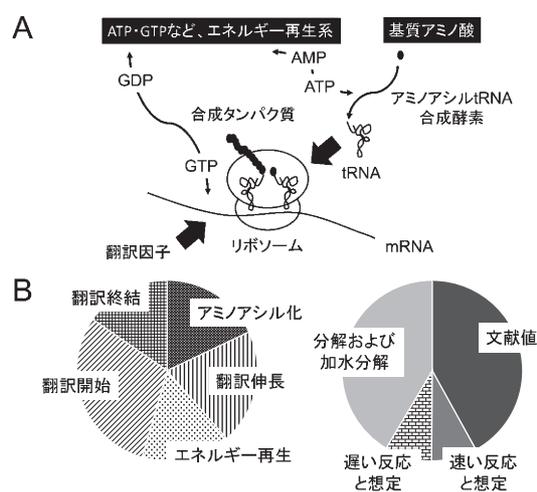


図1. タンパク質合成反応全成分シミュレータ(ePURE)の構築。(A) さまざまな酵素反応と連携したタンパク質合成システム。(B) ePUREに用いた構成要素の分類(左)および割り振られた反応速度定数の分類(右)。

へと融合させる方針を取ることとした。個々のサブモデルに含まれる構成要素や反応については重複可能かつ抽象的な形で記述可能な一定のルールを設け、CellDesignerというソフトウェア用<sup>8)</sup>、視認できる形で記述を行った。個々のサブモデルはSystems Biology Markup Language (SBML) 言語<sup>9)</sup>によって保存され、SBMLに準拠したアプリケーションプログラミングインターフェース (API) ライブラリ<sup>10)</sup>によってソフトウェア的にアクセスすることが可能である。すべてのサブモデルを記述した後、このAPIライブラリなどを用い、ソフトウェア的にこれらを一つの統合化されたモデルへと変換を行っている。

一般的に酵素触媒反応などでは、単純な反応であっても、induced-fitなどによって生じるタンパク質構造変化を含む多段階な反応形態を取るものがあることが知られており、タンパク質合成においてもそうしたダイナミクスが観察されている<sup>11)</sup>。しかしながら、こうした複雑な反応形態までも記述することは困難であるため、モデル生成においては、一次および二次反応のみに着目することとした。またすべての反応は可逆反応として扱い、不可逆反応の場合は逆反応の反応速度定数を0とすることで不可逆反応として記述することとした(構成要素の分解反応のみ不可逆反応とした)。合成するポリペプチドについては、ePURE構築時はMet-Gly-Gly (MGGペプチド)のトリペプチド合成を行うこととしたが、個々のモデルは抽象化されており、現時点ではソフトウェア的にGFPなどのタンパク質を含む任意のポリペプチドを合成可能な形に改良されている(未発表)。

最終的に、MGGペプチドを合成するモデルは、27の開始要素を含む、241の構成要素および968の反応で構成され、それぞれの反応に対してODEが割り振られた。241の構成要素に対しては機能的分類を行い5つのクラスにそれぞれ分類された(図1B左)。また、968の反応に対しては、過去の文献を参考に、反応速度定数が割り振られた。これらの反応速度定数についてもそれらがどのように割り振られたかにしたがって、それぞれ分類された(図1B右)。反応速度定数はすべてが既知ではないため、多数の反応速度定数が推定的なものであったが、以下に記述するように、MGG合成について妥当なシミュレーション結果が得られている。

### ペプチド合成シミュレーション

最適化されたPURE systemの構成要素濃度<sup>12)</sup>を参考に開始要素濃度を決定し、0秒から1000秒までのシミュレーションをMatlab (Mathworks) 上で行ったところ、

PURE systemで観察された合成速度とほぼ同様のペプチド合成が観察された。一般的な無細胞タンパク質合成で観察される1分弱のラグタイムの後、時間依存的な直線性のペプチド合成が観察されており(図2A)、構築したモデルおよび割り振られた反応速度定数が、推定的なものを含むにもかかわらず、妥当なものであることが示唆された。また、モデル内のさまざまな構成要素の濃度を調節することによって、翻訳開始過程や伸長過程などの個々の反応過程を観察した結果、これらも実験において観察されたものと一致する妥当な結果が得られている。さらに、個々の開始要素濃度を変化させてペプチド合成を検証した結果についても概ね実験結果と一致する妥当な結果が観察されており、これらのことから構築したモデルが実際の反応の特徴を反映したものであることが示された。

### 擬定常状態(Quasi-Stationary State : QSS)を利用した反応初期過程の観察

構築したePUREはコンピュータシミュレータであるため、すべての構成要素の濃度変化を辿ることが可能である(図2B)。すなわち、実際の反応液を用いた実験系では情報の取得が難しい反応中間体の濃度遷移や、定常状態に至るまでのごく早い反応初期過程の反応ダイナミクスを推定することが可能である。しかしながら、ほぼすべての構成要素の濃度変化が一定となる定常状態と比較して、反応初期過程では多数の構成要素がさまざまなパターンの濃度遷移を辿っており、こうした複雑な反応ダイナミクスを理解可能な形で把握するためには、抽象化やパターン化などの情報処理過程を経る必要がある。そこで、複雑系を理解するための概念の一つであるカオス履歴<sup>13)</sup>から着想を得て、個々の構成要素の擬定常状態(Quasi-Stationary State : QSS)に着目した解析を行うことにより、反応初期過程を評価しようと試みた。

構成要素の濃度遷移に着目すると(図2B)、反応が定常状態に至る前段階において、一定の間、濃度変化が不変である時間帯がそれぞれの構成要素に観察されており、こうした濃度変化が不変である時間帯においては、その構成要素がQSSにあると定義し、個々の構成要素がQSSにあるかそれともQSSにない濃度変化を伴った状態であるかに二分することによって、それらの状態がどのように遷移するかを図示した(図3A)。その結果、QSSにある構成要素はしばらくの間徐々に増加していくが(Phase I)、特定の時間(約0.5秒)にピークを迎え、その後はQSSにある構成要素が増加と減少を繰り返す時間帯を経て(Phase II)、再び増加し(Phase III)、最終的な

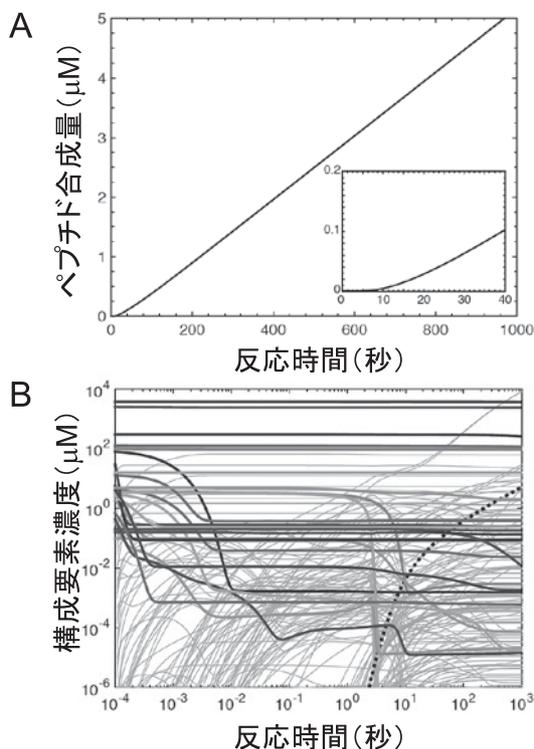


図2. ePUREによるペプチド合成シミュレーション。(A) 時間依存的なMGGペプチド合成。枠内は0秒から40秒までを拡大したもの。(B) 時間依存的な全構成要素の濃度遷移(両対数グラフ)。点線は合成ペプチド濃度遷移, 太い実線は開始要素の濃度遷移, 中間体などのその他の構成要素の濃度遷移は細い実線で示している。図は文献7より引用し, 改変を行った。

定常状態へと至ることが明らかとなった。したがって, QSSに着目した解析によると, この反応初期過程は3つの段階を経て定常状態へと至ることが明らかとなり, 複雑に見えた状態遷移が理解できる形に落とし込めることが明らかとなった。

さらに, QSSにある構成要素同士が個々の反応式において同時に出現する場合は, 「それらの構成要素が関連する」と定義することによってお互いに関連しあうQSSにある構成要素をクラスタリングし, 反応ネットワーク構造がどのように変化していくかを視覚化したところ(図3B), Phase Iでは, 個々のクラスターが融合しながら少しずつ大型化し, その数を減少させていく成長段階にあったが, Phase IIへと変化する特定の時間において, クラスターの小型化と数の増加を伴うクラスターの崩壊が起こり, Phase IIでは, クラスターの成長と崩壊が繰り返し起こっていた。Phase IIIに入ると, 個々のクラスターが融合していき, 最終的に大きな一つのクラスターに至ることによって最終的な定常状態へと変化している。

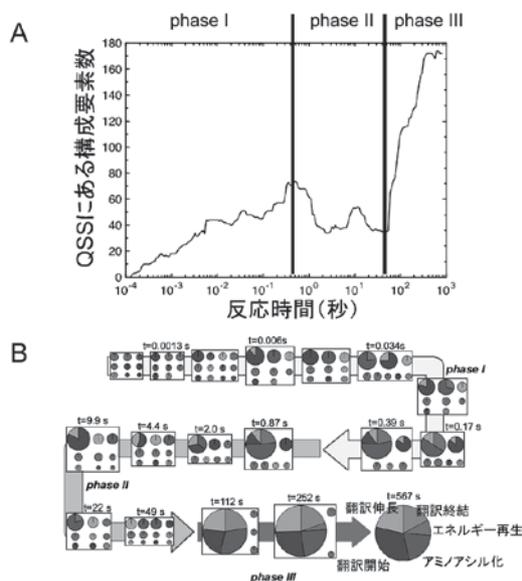


図3. QSSを用いた反応初期過程の観察。(A) 時間依存的なQSSにある構成要素数の推移(時間軸が対数となる片対数グラフ)。(B) お互いに関連しあうQSSにある構成要素を元にクラスタリングしたクラスターの時間的推移。個々の円グラフはクラスターにある構成要素数に比例して大型化するよう表示した。円グラフ内にある要素は図1B左に分類された構成要素の区分を表す。図は文献7より引用し, 改変を行った。

一般的な反応ネットワークの成長について, 個々のクラスターが融合のみを繰り返して最終的に大きなクラスターへと成長していく場合は直感的に理解可能である。しかしながら, QSSに着目した解析においては, 個々のクラスターは融合と崩壊を繰り返して成長している。こうした崩壊はどのような原因で起こるのであろうか。個々の構成要素に着目してさらなる解析を行った結果, Phase IからPhase IIへと至る0.5秒の段階はアミノアシル化されていないtRNAがほぼすべてアミノアシル化された結果, 枯渇してしまうタイミングであり, tRNAのアミノアシル化に相当するネットワークが大きな状態変化を起こすタイミングであった。また, Phase IIでは, 全体のネットワークにおいてボトルネックとなりうる翻訳開始過程の遅い反応速度に起因して, クラスターの崩壊と成長が繰り返されており, これらの遅い反応過程がすべて起動した後, Phase IIIへと移行していることが明らかとなった。したがって, QSSに着目した解析によって, 基質の枯渇による状態変化やボトルネックとなりうる反応過程を明らかにすることが可能であり, 全体の反応過程に影響を与える大きな状態変化を割り出し, 全体の反応のさまざまな局面を明らかにすることに役立つことが明らかとなった。

### 反応速度定数に対する全体の反応の頑強性

前述したように、多くの反応速度定数が推定的なものであるにもかかわらず、実際の反応液を用いて観察された実験結果と一致する、妥当なシミュレーション結果が得られたことから、反応速度定数に対する全体の反応の頑強性が示唆される。そこで、ゼロでない数値を持つ483の反応速度定数に対して個々に変化を与え、MGGペプチドの合成速度（合成速度）および反応が定常状態へと至るまでの時間（ラグタイム）の二点を指標に、それらの網羅的な変化が与える影響の観察を行った<sup>14)</sup>。その結果、反応速度定数を100倍または0.01倍にした際に合成速度およびラグタイムの二点に1.5倍以上の変化を与える感受性の高い定数は全体の6%であり、残りの94%は変化を与えても全体の反応にはほぼ影響を与えない感受性の低い定数であることが明らかとなった（図4）。また、たとえば、定数を100倍にして合成速度にもっとも変化を与える定数であってもその変化は2.2倍程度であることから、反応速度定数に対する全体の反応の頑強性が強く示唆された。

感受性の高い反応速度定数に相当する反応はどのようなものであろうか。相当する反応を割り出したところ、翻訳開始反応および翻訳終結反応の一部に相当し、これらの反応が全体の反応のダイナミクスを決定するボトルネックとなりうる反応であることが明らかとなった。終結反応の一部の正反応の反応速度定数を増加させると合成速度が向上し、逆反応の定数を減少させると合成速度が低下していた、開始反応の一部についても同様の傾向

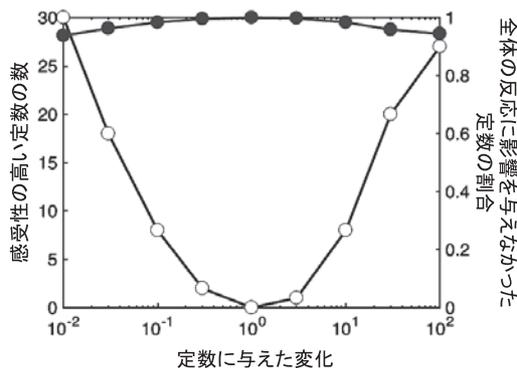


図4. 反応速度定数に変化を与えた際に全体のペプチド合成に影響を与えた定数。白丸は影響を与えた感受性の高い定数の数を表し、黒丸は全体の定数のうち、影響を与えなかった定数の割合を表す。図は文献14より引用し、改変を行った。

が観察されたが、合成速度の他に、ラグタイムへの影響も観察された。通常、多段階の反応において全体の反応速度を決定するいわゆる律速段階が一義的に決定されるとされているが、これらの結果は、大規模な反応ネットワークにおいては、律速段階以外にもボトルネックとなりうる反応が複数存在する可能性があり、今回のような解析を行うことによりそのような全体の反応のダイナミクスを決定する感受性の高い反応速度定数を明らかにすることができることを示している。

### おわりに

タンパク質合成反応について詳細なモデルを、単純な常微分方程式を用いて構築したePUREはwebを通じて誰でも扱えるようになっており<sup>15)</sup>、個々の研究者がさまざまなシミュレーションを行うことが可能である。また、MGGペプチド以外の、GFPなどのタンパク質を含む任意のポリペプチドを合成可能な形への改良を進めており、新生ポリペプチド鎖の合成とタンパク質の折り畳みの連携などにも将来的には応用可能と考えており、さらなる発展が期待される。

また、今回行ったQSSを用いた解析や反応速度定数の網羅的な変化によるボトルネックとなりうる反応の割り出しなどのアプローチはタンパク質合成以外の他の大規模な反応ネットワークの特性を明らかにする際にも大いに参考になると考えており、今回行ったアプローチを一般化することによって、さまざまな生物反応プロセスを明らかにする試みに応用していければと考えている。

### 文 献

- 1) Anfinsen, C. B.: *Science*, **181**, 223 (1973).
- 2) Zhou, M. *et al.*: *Nature*, **495**, 111 (2013).
- 3) Xu, Y. *et al.*: *Nature*, **495**, 116 (2013).
- 4) Karzbrun, E. *et al.*: *Phys. Rev. Lett.*, **106**, 048104 (2011).
- 5) Stógbauer, T. *et al.*: *Integr. Biol. (Camb)*, **4**, 494 (2012).
- 6) Shimizu, Y. *et al.*: *Nat. Biotechnol.*, **19**, 751 (2001).
- 7) Matsuura, T. *et al.*: *Proc. Natl. Acad. Sci. USA*, **114**, E1336 (2017).
- 8) Funahashi, A. *et al.*: *BIOCILICO*, **1**, 159 (2003).
- 9) Hucka, M. *et al.*: *Bioinformatics*, **19**, 524 (2003).
- 10) Bornstein, B. J. *et al.*: *Bioinformatics*, **24**, 880 (2008).
- 11) Milon, P. *et al.*: *Nat. Struct. Mol. Biol.*, **19**, 609 (2012).
- 12) Kazuta, Y. *et al.*: *J. Biosci. Bioeng.*, **118**, 554 (2014).
- 13) Kaneko, K. and Tsuda, I.: *Chaos*, **13**, 926 (2003).
- 14) Matsuura, T. *et al.*: *ACS Synth. Biol.*, **7**, 1964 (2018).
- 15) PURE system simulator: <https://sites.google.com/view/puresimulator> (2019/5/6).