

バクテリア細胞質の全原子分子モデルと大規模分子 動力学シミュレーション



杉田 有治

国立研究開発法人理化学研究所
杉田理論分子科学研究室
主任研究員

タンパク質や核酸などの生体分子は異なる細胞環境において固有の機能を発現する。従来の計算機の演算能力では不可能であった細胞内環境をあらわに考慮した分子動力学シミュレーションを行い、様々な計測と比較していくことがこの課題の目標である。このようなアプローチは特に、X線結晶構造解析で得られるタンパク質1分子の原子レベルの立体構造と、光学顕微鏡等で明らかになるマクロな情報の中間に存在する未開拓の領域を研究するために今後重要になるだろう。

我々は、細胞環境の一つと考えられる細胞内分子混雑環境を考慮し、タンパク質、核酸、代謝物、水などの溶媒が全原子レベルの解像度で含まれるモデルを構築し、分子動力学計算を用いたダイナミクス解析を行った。2つのシステムには1170万原子あるいは1億原子が含まれており、全原子分子動力学計算として世界最大級の生体分子系のシミュレーションである。用いたソフトウェアは計算科学研究機構を中心に開発したGENESIS(Generalized-ensemble simulation system)であり、「京」などの超並列計算機を効率良く利用するために必要な様々なアルゴリズムが導入されている。このソフトウェアを用いることで例えば1170万原子系のダイナミクスを「京」を用いて1日16ns程度計算することが可能である。

この解析で注目しているのは、タンパク質の安定性と分子運動、特に、並進や回転拡散運動、タンパク質とATPなどの代謝物の相互作用、タンパク質間の相互作用などである。全原子分子動力学計算としてはこれまで全く例のないシミュレーションであるため、シミュレーションのみならず解析においても様々な困難が生じたが、細胞内分子混雑系に関する新しい知見が複数発見された。例えば、分子混雑系に含まれるタンパク質の一つであるPGKは、実験的にもその立体構造安定性が良く調べられているものである。分子混雑中ではコンパクトな形状をとり、溶液中ではやルーズな構造を維持することがFRETなどで明らかになっている。我々は分子混雑中のみならず水中、塩溶液中など異なる環境でのPGKの構造安定性を調べるための分子動力学計算を実施した。我々の計算結果は、実験結果と良く一致するPGKの構造変化を示していたが、その分子機構は従来行われた粗視化分子モデルを用いたシミュレーションとは異なり、塩効果の重要性を示すものであった。上記のような新しい知見を統合して細胞内でのタンパク質の動的構造をより深く理解し、将来的には創薬応用の革新に接続していきたいと考えている。

上皮成長因子応答経路の1分子粒度シミュレーション

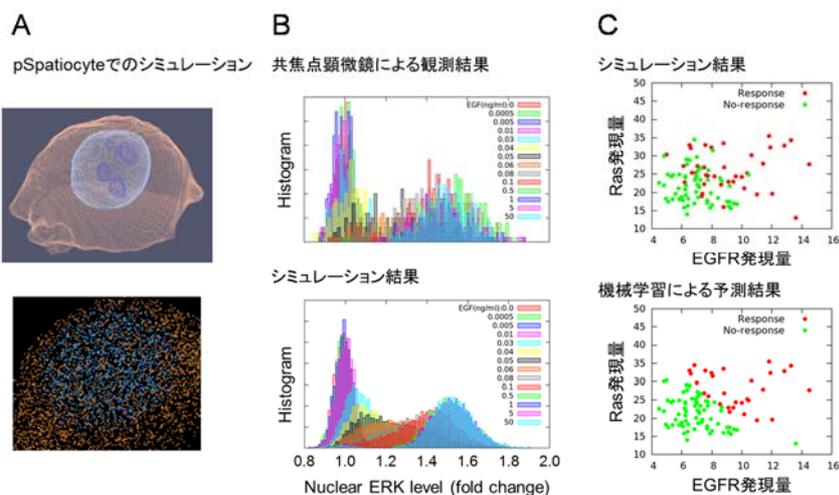


高橋 恒一

理化学研究所生命システム研究センター
生化学シミュレーション研究チーム
チームリーダー

本課題では、細胞内分子混雑や細胞の形態、分子の局在、輸送などの細胞環境を考慮した細胞まるごと規模での1分子粒度反応ネットワークシミュレーションのための高性能計算ソフトウェア **pSpatocyte** を開発し、これを用いて哺乳動物細胞における上皮成長因子応答経路の1分子粒度モデルを開発した。

上皮成長因子応答経路は増殖や分化などの細胞運命を制御するMAPKであるERKの活性化を引き起こす信号伝達経路のうち主要なもの1つである。共焦点顕微鏡を用いてEGFPで蛍光標識されたERK分子の核移行を観察すると、遺伝的に同一な細胞群に全く同一の条件で増殖因子刺激を与えても、因子の濃度によっては細胞間で応答のばらつき（応答不均一性）が生じる場合がある。本研究では「京」を用いて分子混雑などの細胞環境を考慮した1分子粒度シミュレーションを実行し、ERKの応答不均一性を再現する事に成功した。また、その発生机序が特定のタンパク分子の外因的ノイズと関連する事を示唆する結果を得た。さらに、このシミュレーション結果を利用して、機械学習により遺伝子の発現レベルから確率的な細胞応答の事前予測が可能である事を示唆する結果を得た。



リン酸化酵素反応機構の理解に向けた自由エネルギー解析



林 重彦

京都大学大学院理学研究科
教授

細胞内信号伝達は、複数のリン酸化酵素のリン酸化反応により行われる。それらのリン酸化酵素は、その酵素自身が別のリン酸化酵素によりリン酸化を受けることにより、リン酸化活性が変化し、信号伝達が制御される。従って、リン酸化酵素の反応活性は、細胞内信号伝達の鍵となる化学的過程であり、主な創薬ターゲットの一つとなっている。そこで、本研究では、細胞環境内におけるリン酸化酵素の反応活性の分子機構を明らかにし、細胞内信号伝達の理解と制御に向けた分子論的基盤を構築することを目指す。

このような、タンパク質複合体の酵素反応を解析する際には、複合体の柔らかい構造で起こりうるタンパク質の大きな構造変化を捉えることが必要となる。最近、タンパク質分子に特徴的な遅く大きな構造熱揺らぎが、酵素の高い触媒活性重要であることが実験的に示唆されているが、直接的な証拠がなく大きな論争となっている。そのような酵素活性の分子機構を分子シミュレーションにより探るために、近年、広く用いられ成功を収めている手法が、quantum mechanical/molecular mechanical (QM/MM) 法である。この方法は、分子軌道法や密度汎関数法などの量子化学的 (QM) 手法と、生体分子の分子シミュレーションで通常用いられている分子力場に基づく分子力学的 (MM) 手法をハイブリッドするものであり、化学反応に関わる領域を QM 法で取り扱う際に、そのまわりのタンパク質環境の影響を MM 法に基づき記述することにより、酵素反応活性に関わる触媒場を非常に効率良く考慮することを可能にする。

しかしながら、そのような QM/MM 法の成功の一方で、限界も明らかになってきた。最も大きな困難は、タンパク質分子に特徴的な遅く大きな構造熱揺らぎを考慮するための十分な統計サンプルが取れないことである。この問題を解決するために、我々は、新規な QM/MM 自由エネルギー法 (QM/MM RWFE-SCF 法) を開発した。この手法では、分子動力学シミュレーションで得られたタンパク質環境の統計サンプルにより構成される自由エネルギー曲面上で QM 分子の反応解析を行うことにより、酵素反応に伴う大きな構造熱揺らぎの効果を考慮することが可能になる。QM/MM RWFE-SCF 法を用いて、Ras-GAP 信号伝達 G タンパク質複合体の酵素反応の解析を行ったところ、反応遷移状態生成の伴い、タンパク質の大きな構造変化が観測され、それにより、反応活性化エネルギーが減少し、酵素触媒活性が高まっていることが示された。

MD-SAXS 法を用いた核内 DNA 結合タンパク質の機能構造シミュレーション



池口 満徳

横浜市立大学
大学院生命医科学研究科
教授

分子動力学シミュレーションは、タンパク質や核酸等について、機能に重要な動的な性質を明らかにするための欠くことのできない方法になっているが、先端計測実験と連携することにより、その価値を高めることができる。放射光等を用いた X 線小角散乱法は、タンパク質の溶液構造を計測するのに効果的な実験手法であるが、得られるデータが低解像度にとどまるという問題がある。そこで、X 線小角散乱法を、スーパーコンピュータによる分子動力学シミュレーションと組み合わせることにより、実験的にも整合性のある溶液構造の詳細なモデルを構築することができる。

本研究では、粗視化分子動力学(MD)シミュレーションと X 線小角散乱(SAXS)実験を組み合わせる MD-SAXS 法を開発した。粗視化 MD シミュレーションは、複数の原子をまとめて取り扱うことで、ヌクレオソームのような巨大複合体でも十分にシミュレーションが可能となる。複数のタンパク質や核酸の立体構造について、本方法を適用して粗視化モデルによる SAXS プロファイル計算を行い、SAXS 実験結果と比較したところ、精度よく一致することがわかった。

さらに、本方法を、ヒストンバリエントを含むヌクレオソームに適用した。研究対象は、染色体に広く存在し一般的なカノニカルヌクレオソーム、染色体のセントロメア領域に存在しこの領域の形成に重要な役割を持つ CENP-A ヒストンバリエントを含むヌクレオソーム、転写が活発に行われている染色体領域に存在する H2A.B ヒストンバリエントを含むヌクレオソームの 3 種のヌクレオソームである。スーパーコンピュータ「京」を用いて、様々なパラメータのもとで3万回近いシミュレーションを繰り返したところ、SAXS プロファイルによく合う構造分布を導出することに成功した。その構造分布を解析することにより、ヒストンバリエントの違いによって、DNA とヒストンの相互作用は大きく異なり、DNA の形状に影響を与えていることが明らかになってきた。

粗視化分子モデルを用いた信号伝達経路上の リン酸化酵素複合体とクロマチンの動的モデリング



高田 彰二

京都大学理学研究科
教授

(1) 信号伝達経路上のリン酸化酵素複合体の動的モデリング

細胞が周囲の環境に適応するためには、外部からの刺激を核内まで伝達させて遺伝子発現を変化させる機構が必要となる。シグナル伝達経路の代表的モデルである MAPK カスケードでは、外からの刺激を契機に、Ras, Raf(MAPKKK)が活性化し、さらに MEK1(MAPKK)、ERK2(MAPK)へとリン酸化カスケードが起こる。本研究の目的は哺乳類の MAPK カスケードにおける信号伝達の分子機構に迫る為に、実験的に未解明な MEK1(MAPKK)と ERK2(MAPK)の複合体の構造モデリングをすることである。我々は、(i)粗視化シミュレーションにより 2 つの有力構造を、バイオインフォマティクス的手法により有力構造を 1 つ、構築した。これらの 3 つの複合体では、MEK の触媒部位と ERK の活性化ループ(リン酸化サイト)がある程度近くに存在する事が分かった。特に、前 2 者は、独立な方法[Wang 等(杉田研)]により作成された代表的な複合体構造とよく似ている事が分かった。

(2) クロマチンの動的モデリング

ヌクレオソーム 3 量体の構造動態とセントロメア特異的構造：セントロメアは、動原体が結合するサイトを作るクロマチンの特別な領域であり、そのことによって染色体の分離を可能にしている。動原体の形成機構を詳細にするために、セントロメア特異的なクロマチンの構造情報が不可欠であるが、これまでにその構造はほとんど明らかになっていなかった。我々は、実験的に合成されたヌクレオソーム 3 量体の X 線溶液散乱データと粗視化分子シミュレーションの協力によって、通常のものとはセントロメア特異的な二つのヌクレオソーム 3 量体の構造を明らかにした。

アセチル化の効果：ヒストンアセチル化は、代表的なエピジェネティック制御であり、概ねクロマチン構造をほぐくことによって遺伝子発現を活性化することが知られている。しかし、その詳細な機構は不明である。本研究では、ヌクレオソーム 3 量体からなる系を例にとり、様々なヒストンテールアセチル化が、その構造をどのように変えるのかを研究した。アセチル化されるヒストンテールによって、顕著に異なる構造変化がみられることを明らかにした。

クロマチン構造と転写因子動態の連関：核内混み合い環境を実現する為に、20 ヌクレオソームからなるクロマチン様構造を構築し、その中における 3 つの核内蛋白質、大型の転写因子 p53、小型の転写因子 HMGB1、非転写因子 ERK の拡散運動等を研究した。ヌクレオソーム濃度 0.5mM の混雑環境では、p53 の拡散は著しく抑制されているが、ERK はやや遅くなるものの十分に拡散が可能である。HMGB1 はその中間的ふるまいを示す。その分子機構について、構造をもとに解析を行った。

リガンドによるタンパク質の構造変化と結合自由エネルギー



藤谷 秀章

東京大学
先端科学技術研究センター
特任教授

疾患標的タンパク質の立体構造を X 線結晶構造解析や NMR 法などで明らかにして、このタンパク質によく結合する新しい低分子化合物を設計する Structure Based Drug Design (SBDD)を成功させるには何が必要であろうか。100 万個の化合物ライブラリーからタンパク質のポケットに結合する化合物をドッキング・シミュレーションで選別しても、実際に結合する化合物を見出す確率は極めて低い。実際のタンパク質は周りを水分子に囲まれて化合物が入る前のアポ構造で揺らいでおり、化合物は水分子やアミノ酸側鎖を押し退けて結合位置に入り込む。この侵入過程で化合物特性に応じたタンパク質の構造変化が起きるので、結合する化合物によってタンパク質立体構造は異なる。タンパク質と低分子有機化合物を統一的に記述する FUJI 力場を用いて、タンパク質と化合物の結合自由エネルギーを求めるには、MP-CAFEE 計算の初期構造が熱平衡状態にある事が前提になる。結合しない化合物をタンパク質の結合ポケットに押し込めても、熱平衡状態を求める途中で化合物はポケットから外に飛び出す筈である。MP-CAFEE で正しい結合自由エネルギーが得られないのは、計算の初期構造に使用する熱平衡状態を求める処で失敗しているのが普通である。特に創薬標的として関心が高いキナーゼでは複合体結晶構造から熱平衡状態を得るのが難しい場合があり慎重な計算が必要である。

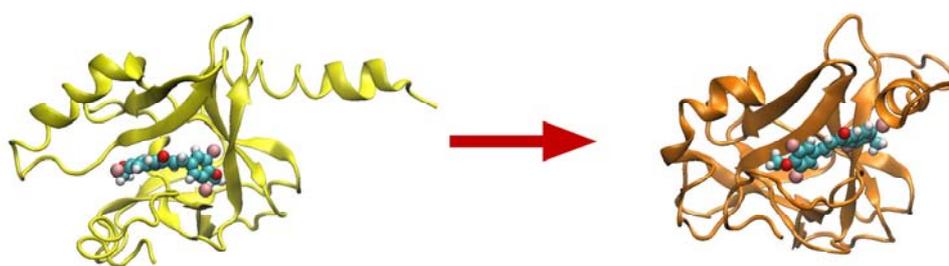


Figure 1. X 線結晶構造解析で得られたタンパク質のリガンド結合ポケットに新たな低分子化合物を入れて分子動力学計算を実行するとタンパク質形状が大きく変化する事例.

予測医療に向けた階層統合シミュレーション

「抗血小板薬効予測のためのマルチスケール血栓症シミュレーションと筋繊維の集合体としての骨格筋シミュレーション」



高木 周

東京大学・大学院工学系研究科
教授

本講演では、課題3全体として取り組んでいる「階層統合シミュレーション」の中から、血栓症の初期過程である血小板粘着のシミュレーションと抗血小板薬のモデリング、および、脳神経系-筋骨格系統合シミュレータを用いたパーキンソン病の病態再現に向けた連成計算手法の開発について説明する。血小板粘着のシミュレーションについては、これまでは、後藤ら（東海大）が行ってきたフローチャンバー実験に関して、同程度のせん断率を持つ単純せん断流れのシミュレーションを行い、血小板粘着に対する赤血球の影響、すなわちヘマトクリット値依存性について議論を行ってきた。これに対し、昨年度から本年度にかけて、実際のフローチャンバーの流動条件と同じ条件でのシミュレーションが達成された。その結果について報告する。また、東海大で行っている抗血小板薬のモデリングのためのフローチャンバー実験の結果より、血小板表面および内部の生化学反応をモデリングし、多数の粘着血小板が存在する系でそのモデルを取り入れた大規模シミュレーションを実施した。その結果について説明する。パーキンソン病については、銅谷ら(OIST)が構築した大脳基底核-視床-運動野の連成モデルと中村ら（東大）が開発している全身モデルを結びつけるものとして、個別の骨格筋について、脳から来るスパイクシグナルが脊髄・運動ニューロンを経由して、筋繊維を収縮させ、さらに筋繊維の集合体としての骨格筋全体が収縮するマルチスケール骨格筋シミュレータの開発を行ってきた。この部分について最近の進展を説明する。

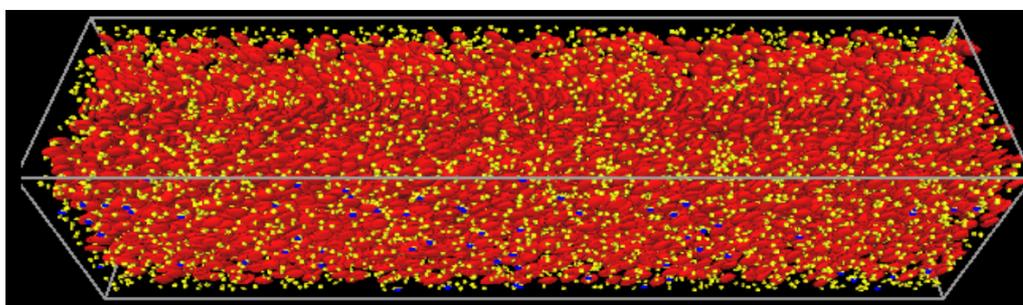


Fig. 1 Actual scale simulation on platelet adhesions in the flow chamber experiment
(Blue spots correspond to the platelet adhesion locations.)

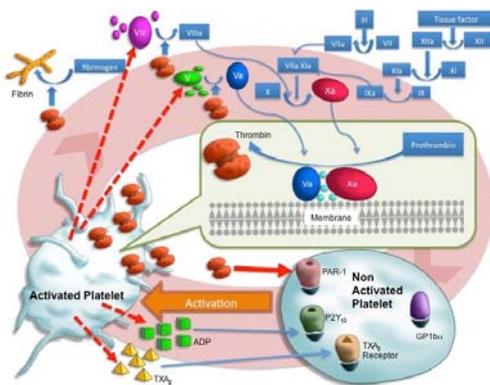
抗血小板薬薬効予測を可能とする階層統合シミュレータの開発に向けたモデリング



後藤 信哉

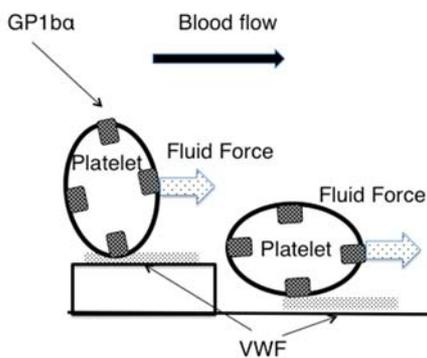
東海大学 医学部 内科学系循環器内科学
教授

心筋梗塞、脳梗塞などの発症に至る冠動脈、脳動脈閉塞血栓には血小板細胞が含まれる。血小板細胞活性化シグナルおよび受容体阻害薬であるアスピリン、クロピドグレルなどの抗血小板薬による心筋梗塞、脳梗塞発症予防効果が臨床的に確認されていることから、これらの疾病と「血小板細胞」の間には論理的関連があると推論される。

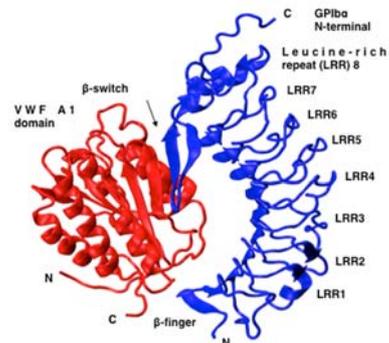


心筋梗塞、脳梗塞を惹起する冠動脈、脳血管の閉塞血栓のサイズはミリメートルオーダーである。Fig. 1 に示す「血小板細胞」の直径は2-5マイクロメートルなので、マイクロメータスケールの血小板細胞の集積、活性化がミリメートルスケールの血管閉塞に至るモデルが必要となる。活性化血小板上のリン脂質を介する凝固カスケードの活性化反応を取り込んだモデルを公表した (左図: Goto S, et al. *Circulation*, 2015)。

Fig. 1 に示す血小板細胞は、動脈血流に抵抗して血管壁損傷部位に集積する。生物学的実験により血小板の接着は血小板細胞上の GPIb α と血管壁損傷部位の VWF の相互作用による。



両者の相互作用モデルを Fig. 1 に示す。分子動力学計算の手法を用いて、結晶構造から GPIIb α と VWF の可溶性結合構造を予測計算した (Fig. 2)。Fig. 2 から分子間力予測計算を行い、Fig. 1 の Fluid Force を予測する。さらに、複数血小板の流体した相互作用のモデルを作成する。分子スケール (ナノメートル)、細胞スケール (マイクロメートル)、臓器全身スケール (ミリメートル) の連成計算により、薬剤介入可能な分子スケールの介入結果の細胞スケール (機能評価可能)、臓器全身スケール (疾病発症) の論理の一貫化を目指したモデリングを行った。



臓器全身スケール (ミリメートル) の連成計算により、薬剤介入可能な分子スケールの介入結果の細胞スケール (機能評価可能)、臓器全身スケール (疾病発症) の論理の一貫化を目指したモデリングを行った。

心疾患の合理的治療のためのマルチスケール マルチフィジックス心臓シミュレーション



久田 俊明

東京大学名誉教授

(株) UT-Heart 研究所 代表取締役会長

心臓は ATP 加水分解による生化学反応をエネルギー源として、化学・電気・力学の諸現象に広く派生するマルチフィジックス問題を構成する。また空間尺度としては、タンパク分子から細胞、組織、臓器を経て血液拍出に至るマルチスケール問題を構成している。臨床で日常的に用いられる心電図や血圧などのマクロ現象については古くから多くの医学・生理学的研究がなされてきたが、一方で近年著しい発展を続ける分子生物学によるミクロレベルの知見との因果関係は、その間に大きなスケール差と複雑な相互作用を介したブラックボックスが介在しているため、もはや専門家にとっても明らかでなく、両者を合理的に説明し予測することは困難な状況に至っている。この観点から、私達は 2001 年頃よりミクロとマクロを架橋する心臓シミュレータ UT-Heart の研究開発に取り組んできたが、「京」の出現によりシミュレーションが現実のものとなった[1][2][3]。本戦略プログラムでは特にサルコメアの確率的動力学モデルから心臓の拍動へと繋ぐマルチスケールシミュレーションによる基礎医学の研究、その成果を用いた外科領域を中心とする臨床研究、更にはマルチスケール冠循環シミュレーションなどを行って来た。講演ではそれらの概要を実用化への展望を含めて述べる。なお本研究は幾つかの病院の協力のもとに倫理委員会の承認を得て、UT-Heart 研究チーム（筆者、杉浦清了、鷺尾 巧、岡田純一、假屋太郎）と富士通（株）未来医療開発センターの共同によりなされたものである。

[1] S. Sugiura, T. Washio, A. Hatano, J. Okada, H. Watanabe, and T. Hisada, Multi-scale simulations of cardiac electrophysiology and mechanics using the University of Tokyo heart simulator, *Prog Biophys. Mol. Biol.*, 110 (2012), pp. 380-389.

[2] T. Washio, J. Okada, S. Sugiura, and T. Hisada, Approximation for cooperative interactions of a spatially-detailed cardiac sarcomere model, *Cell. Mol. Bioeng.*, 5 (2012), pp. 113-125.

[3] T. Washio, J. Okada, A., Takahashi, K., Yoneda, Y., Kadooka, S. Sugiura, and T. Hisada, Multiscale Heart Simulation with Cooperative Stochastic Cross-Bridge Dynamics and Cellular Structures, *SIAM Multiscale Modeling & Simulation*, 11:4 (2013), pp. 965-999.

ヒトの運動器、運動神経、感覚神経からなるシステムの統合シミュレーション



中村 仁彦

東京大学

大学院情報理工学系研究科 知能機械情報学専攻
教授

私たちはロボティクスの研究で開発した計算アルゴリズムをヒトの骨格系に適用することを2000年頃から始めました。ヒトの全身の骨格系を155自由度のリンク系として扱い、骨格系に筋の付着位置に従って張られた約1000本の糸が独立に張力を変えると、操り人形のような筋骨格モデルを作りました。

現在、スーパーコンピュータ「京」を用いて、詳細な数学モデル K-Body を開発しています。各筋は質量と連続体とし有限要素法で解き、骨格の剛体系との連成運動を計算します。生理学の文献から筋の運動ニューロンと筋紡錘の感覚ニューロンの数を調べ、すべての筋に外挿し、脊髄の内外のニューロンプールに空間的に配置しました。有限要素計算エンジンには理研が開発した V-Biomech を採用し、ニューロンは Leaky Integral and Fire Model による Spiking Neuron としプログラミング言語 NEST を用いて配置しました。図1に有限要素法の全身の筋、図2に脊髄のニューロンプールを示します。

課題3「予測医療に向けた階層統合シミュレーション」(代表：高木周)では沖縄科学技術大学院大学のチームが開発する脳の大脳基底核と皮質のモデルとつないで、大脳基底核で生じるドーパミン受容体の変性に関わるパーキンソン症候群のシミュレーションによる再現を目指しています。

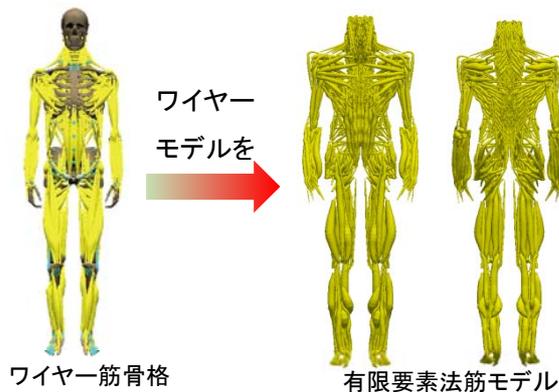


Figure 1. FEM Musculoskeletal Model

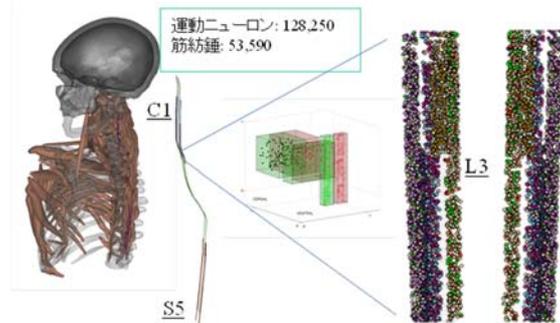


Figure 2. Motor and Muscle-Spindle Neuron Pool

コンプライアントな生体筋・腱系に駆動されるヒト骨格系の運動制御に対して大脳基底核が果たす役割の数理モデル化



野村 泰伸

大阪大学 大学院基礎工学研究科
機能創成専攻 生体工学領域
教授

本研究では、静止立位姿勢および定常二足歩行の神経制御の数理モデル化を実施した。特に、これらの運動を生成する神経制御が実現すべき運動の特徴として、直立姿勢あるいは周期的歩行（歩行サイクル）の平衡状態からの偏差の揺らぎとして表出する関節の柔軟性、および、運動揺らぎの時間パターンが示すフラクタル性あるいはべき乗則に従う長期相関に注目した。さらに、直立姿勢の安定性に関して、体性感覚情報の神経伝達時間に起因する時間遅れ誘因性不安定化を回避する姿勢制御メカニズム、および、対応する制御器のパラメータ変化に対する姿勢安定化メカニズムのロバスト性を解析した。運動計測データに表出する特徴は、姿勢の安定化メカニズムが機能した結果として生成されると考えられる。我々は、安定性と柔軟性という一見相反する運動特性を実現する制御戦略の有力候補の一つである間欠制御仮説を提案し、その数理モデル化を通じて、ヒト身体運動制御系の大規模マルチスケールモデル構築に貢献する。

A Possible Neural Basis of Intermittent Control during Human Standing and Walking

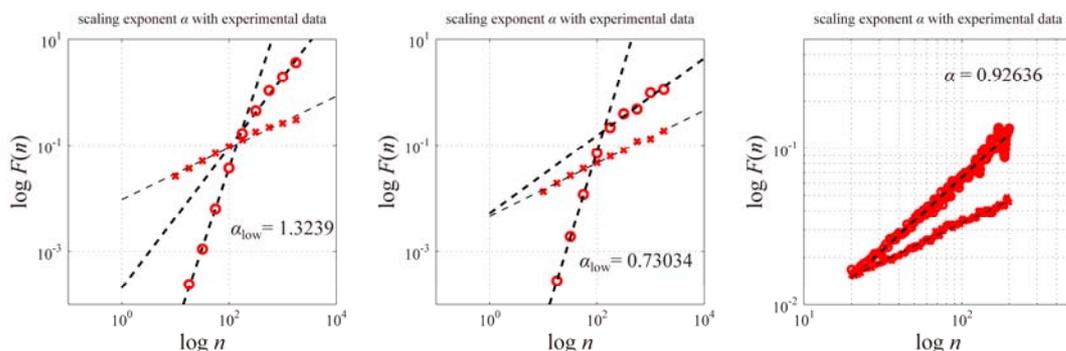
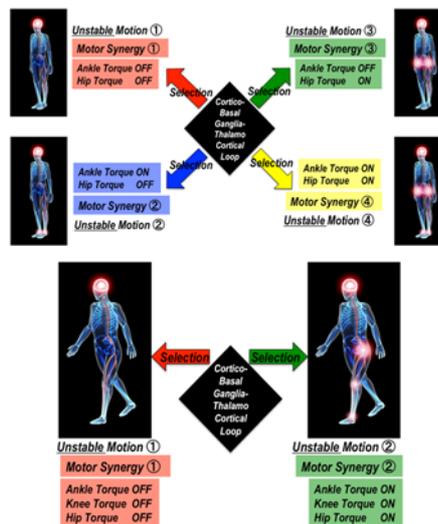


Figure 1. Detrended Fluctuation Analysis for postural sway in a healthy subject (left) and a patient with Parkinson's disease (middle) and for gait cycle variability in a healthy subject (right).

大規模生命データ解析：「京」で初めて見えた世界



宮野 悟

東京大学医科学研究所ヒトゲノム解析センター
教授

2013年より、大規模生命データ解析は、秋山泰（東工大）、松田秀雄（阪大）、そして宮野悟（東大）の3チームからなるグループに集中的に編成され、がんのシステム異常の網羅的解析、寒冷刺激で転換する脂肪細胞のメカニズム、そして細菌叢のメタゲノム解析という3つの方向に研究を展開してきた。その結果、宮野等は「京」による世界最大規模の網羅的がんの薬剤感受性・耐性遺伝子ネットワーク解析を行い、薬剤耐性・感受性を予測する方法を構築し、薬剤に関して、がんの個別化医療及び創薬基盤のひとつを構築した。また、30年前に発見されたがん遺伝子 MYC は数千の遺伝子を制御していることが知られ、18,000以上の論文が発表されている。肺がんの分子メカニズムの研究者である高橋隆教授（名大）との連携により、約8000サンプルの肺腺がん患者の遺伝子発現データを「京」で解析し、MYCの転写活性を制御する新たなロング・ノンコーディング RNA (MYMLR) を発見した。これは MYC の制御にロング・ノンコーディング RNA が関与していることを世界で初めて示したものである。肥満にとっては悪玉ともいわれる白色脂肪細胞組織が寒冷刺激によりベージュ（褐）色に変化する。このベージュ細胞組織は骨格筋の100倍の熱産生能力をもつアンチメタボ細胞である。松田等は、河田輝雄教授（京大農）と連携し、その全遺伝子ネットワークを「京」で解析した結果、熱産生の鍵遺伝子を惹起する新たなメカニズムを明らかにした。さらに、培養細胞では寒冷刺激でベージュ化するが、個体内では寒冷刺激をうけてもベージュ化しない細胞がある。マイクロ RNA を含む全遺伝子ネットワークを「京」で解析した結果、ベージュ化を抑制する機構を解明しつつある。秋山等は、従来法(BLASTX)の約100倍の高速化を実現し、「京」で超並列化・高速計算する GHOST-MP を開発（世界最深度で遠いホモロジーを探索するツール）し、細菌叢のメタゲノム解析に応用した。粘膜免疫を専門とする植松智教授（東大）と連携した研究では、GHOST-MP を用いて、海外渡航歴およびコレラワクチンの履歴がない IgA+ のヒト糞便検体とコントロール・ヒト糞便検体(IgA-)のメタゲノムを比較解析した。その結果、門レベルでは区別がつかなかったが、属レベルで IgA+ 群において有意に増加している大腸菌を発見した。この大腸菌は コレラトキシンと構造・機能的に類似し、Cholera Toxin B (CTB) -IgA と交差反応する heat-labile enterotoxin (LT) を産生するものである。このことは、LT 産生をする病原性の大腸菌を常在化させる菌群がいると考えられる。コレラのワクチンはまだ確立されたものがない状況であり、また IgA 抗体は持続期間が短い。今後これを同定することで、新たなワクチン開発につながる可能性が見えた。

大規模生体分子ネットワーク解析による脂肪細胞のエネルギー消費への転換機構の解明



松田 秀雄

大阪大学 大学院情報科学研究科
教授

ヒトを含む哺乳類には、大別して2種類の脂肪細胞があり、それぞれ白色脂肪細胞、褐色脂肪細胞と呼ばれている。白色脂肪細胞は中性脂肪の形でエネルギーを貯蔵するのに対して、褐色脂肪細胞は脂肪を分解して体温の恒常性維持のために熱として放出する。興味深いことに、継続的な寒冷刺激を受けると、一部の部位の白色脂肪細胞は褐色脂肪細胞に似た色となり熱産生を行うようになる。この転換は「褐色化」と呼ばれ、褐色化した白色脂肪細胞はベージュ脂肪細胞と呼ばれる。脂肪細胞の褐色化は、蓄積された脂肪の分解とエネルギー消費を伴うことから、新しい視点からの肥満是正の方策として注目を集めているが、一部の白色脂肪細胞のみしか褐色化しない原因は不明であった。

我々のグループは、京都大学農学研究科の河田照雄教授のグループの協力を得て4°Cで飼育されたマウス個体のベージュ脂肪細胞からマイクロアレイにより時系列遺伝子発現データを取得し、発現変動が見られた約1万個の遺伝子に対するネットワーク解析を行った。その結果、褐色化機構の一部が免疫細胞の一種であるマクロファージの分泌するサイトカインによって調節されていることを明らかにした（下図の点線で囲まれた部分）。さらに、褐色化機構へのマイクロRNAの関与に注目して、前述のベージュ脂肪細胞から約300個のマイクロRNAの時系列発現データを取得し、マイクロRNAを含めた生体分子ネットワークの解析を行った。これらの大規模生体分子ネットワーク解析は「京」を使うことで初めて実現できたものであり、本講演では解析により得られた結果について説明する。

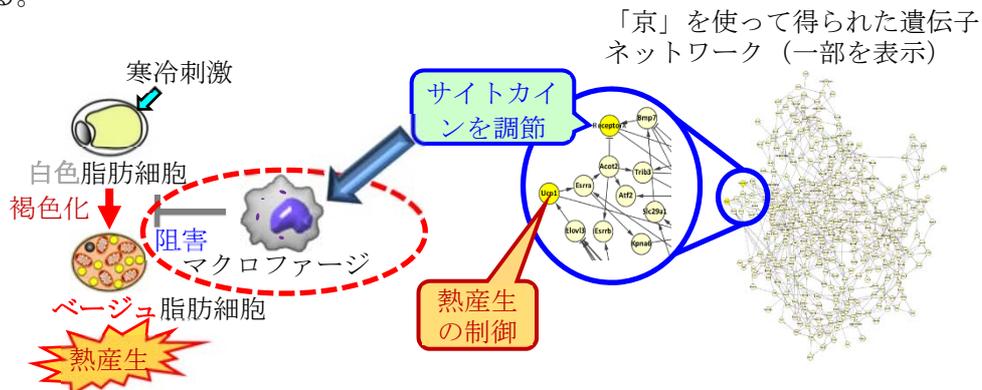


Figure 1. 解析で明らかになった脂肪細胞のエネルギー消費に向けての転換機構

次世代シーケンサデータ解析のための 情報処理システムの開発



秋山 泰
東京工業大学
大学院情報理工学研究科
計算工学専攻
教授

ゲノムを基軸とした大規模生命データ解析から、生命プログラムとその多様性を理解するために、次世代シーケンサから産出される大量のゲノム配列情報を高速に解析する情報処理システムの開発が危急の課題となっている。本研究ではその要求に答えるため、メタゲノム解析パイプラインの開発を行った。メタゲノム解析パイプラインは、環境中の細菌叢から直接抽出したDNAから次世代シーケンサによって得られる配列情報に対して解析を行い、環境中に存在する細菌の種類や遺伝子、その相対存在度を明らかにする。

解析パイプラインにおける処理の中核は配列相同性検索となるが、既存の検索プログラムでは十分な速度が得られないため、高速な検索アルゴリズム (GHOSTX アルゴリズム) の考案とその並列化を行った。GHOSTX アルゴリズムは、逐次処理においても BLASTX に対し、高い検索精度を保ちながら 131 – 165 倍の高速化を達成し、MPI/OpenMP を用いたハイブリッド並列版 (GHOST-MP) は、「京」上で 24,576 CPU コアまで良いスケールリングを示した (Fig. 1)。開発した解析パイプラインをヒト口腔内 8 部位 381 サンプルの細菌叢のメタゲノムデータに適用し、オーソロググループの相対存在度を各サンプル間で比較し、サンプル間の関係や特定の口腔部位で特異的な相対存在度を有する遺伝子グループを明らかにした (Fig. 2)。現在はさらに、この解析パイプラインを利用して、腸管免疫の専門家である植松智特任教授 (東京大学医科研) および口腔微生物の専門家である石原和幸教授 (東京歯科大学) と実データ解析の共同研究を実施している。

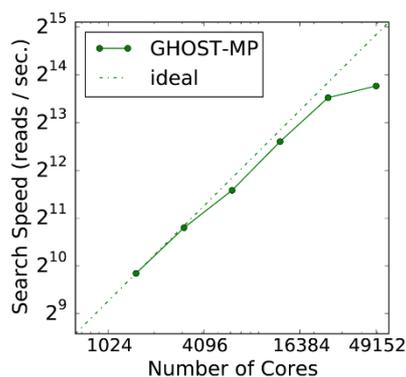


Figure 1. Strong scaling of homology search program (GHOST-MP) on K computer.

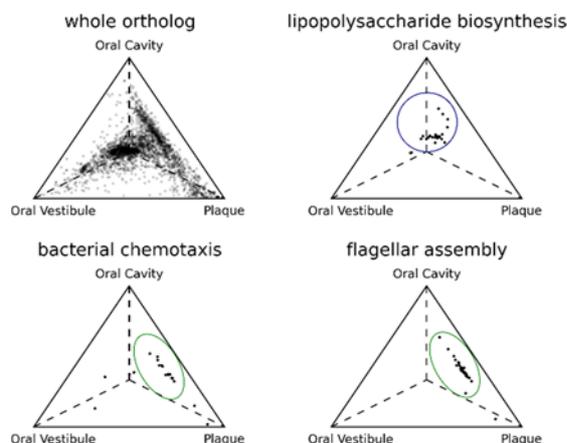


Figure 2. Ternary plot of the relative abundances of the ortholog groups related to a specific biological pathway. Blue and green circles indicate abundance in oral cavity and less abundance in oral vestibule, respectively.

計算生命科学の教育とアウトリーチ ー現在と未来ー



江口 至洋

理化学研究所 *HPCI* 計算生命科学推進プログラム
副プログラムディレクター

計算生命科学は学際的な研究領域であり、古くから計算科学の伝統がある物質・材料や気象等の研究領域に比べ、21世紀に入ってからの生命科学の急速な進展をも踏まえるといまだ萌芽的段階にあると言える。そのため、「教育とアウトリーチ」活動は重要であり、本プログラムでは積極的に進めてきた。

10年、20年先の計算生命科学研究者を育てる意味からも、高校生、大学生、大学院生を対象にした出張講義を行ってきた。高校生や大学生には、個々人の体内にあるタンパク質の動きのシミュレーション動画を見て「えっ。気持ち悪い」という印象と共に、学問領域として生物学が物理、化学、数学と繋がり、本来的に学際的研究領域であるとの実感を得て頂いた。大学院生を対象にした遠隔授業「計算生命科学の基礎」を神戸大学主催で実施して頂いたが、予想外にも民間企業からの参加が多く、計算生命科学への期待が大学のみでなく民間企業においても強く存在することが明らかになった。この遠隔講義は今後も継続して実施していく予定で、将来的には計算生命科学の「標準カリキュラム」に繋がっていくことを期待している。なお、大学院生や民間企業の実験者を対象とした教育では、大阪大学および産業技術総合研究所のご協力を得ている。

広報活動では *BioSupercomputing Newsletter* を年2度発行し、*HPCI* 計算生命科学推進プログラムの研究内容を全国の大学、民間企業の実験者に紹介してきた。また、先進的な研究内容を臨場感を持って理解して頂くために、シミュレーション動画の公開を行い、「Scientific American」の Blog Site などで紹介して頂き、世界的な広報活動が展開できている。これら広報活動は「*HPCI* 戦略プログラム 分野1」のホームページを通しても行われている (<http://www.scls.riken.jp/>)。

研究成果の普及では、「京」互換機である *SCLS* 計算機システムを全国の生命科学研究者に開放し、活発に利用して頂いている。また、大学等研究機関と製薬企業とのコンソーシアムを形成し、「京」の産業利用枠を活用した研究活動を行って頂いている。現在、2つのコンソーシアムが活動しているが、そこには20社以上の製薬企業が参画しており、将来計算生命科学が健康・医療サービスの質的な向上に寄与しうることが期待される。

本プログラムは「京」とともに歩んできたが、2020年稼働予定のポスト「京」においてもさらに計算生命科学が、Tax Payer でもある国民と共にさらに発展していくことを願っている。