スーパーコンピューティング ニュース SUPERCOMPUTING NEWS _{東京大学情報基盤センター}スーパーコンピューティング ニュース

Vol.24, Special Issue 1

2022.8

特集: 2021年度 「若手・女性利用者推薦」 採択課題 成果報告







表紙の画像は、以下の課題より提供を受けております。

- Numerical simulation of pure solutal Marangoni convection in a shallow rectangular cavity (Graduate School of Engineering Science, Osaka University ZHANG JIANGAO)
- ・分子動力学計算で明らかにする金属結合タンパク質のダイナミクス (筑波大学 計算科学研究センター 森田 陸離)
- ・階層型直交格子と再帰的なフィッティングによる回転翼の非定常空力予測 (東京大学大学院 工学系研究科 航空宇宙工学専攻 菅谷 圭祐)

特集:2021 年度「若手・女性利用者推薦」採択課題 成果報告

2021 年度特集号発行にあたって 1 下川辺隆史(東京大学情報基盤センター)
Numerical simulation of pure solutal Marangoni convection 4 in a shallow rectangular cavity 4 ZHANG JIANGAO (Graduate School of Engineering Science, Osaka University) 4
分子動力学計算で明らかにする金属結合タンパク質のダイナミクス10 森田陸離(筑波大学計算科学研究センター)
PaCS-MD と異常検知を援用したタンパク質構造遷移経路の探索 17 原田隆平(筑波大学計算科学研究センター)
分子動力学シミュレーションで解明する維持メチル化酵素(DNMT1)の 活性化メカニズム
降着円盤乱流における Alfvén 的揺動と圧縮的揺動の配分
階層型直交格子と再帰的なフィッティングによる回転翼の非定常空力予測31 菅谷圭祐(東京大学大学院工学系研究科航空宇宙工学専攻)
ダイマー粒子を用いたガラスの Johari-Goldstein beta 緩和の理解
公共データを活用した転写因子結合ダイナミクスの解析45 植野和子(国立国際医療研究センター ゲノム医科学プロジェクト戸山プロジェクト)
ディープラーニングによる高精度マルウェア分析

下川辺 隆史

東京大学情報基盤センター

1. はじめに

本特集号は,東京大学情報基盤センター(以下,センター)で実施した 2021 年度「若手・女 性利用者推薦」制度に採択され,実施した課題の成果報告をまとめたものです。

2. 若手·女性利用者推薦

センターでは,若手研究者(40歳以下,学生を含む)及び女性研究者(年齢は問わない)に よる,スーパーコンピュータ,大規模ネットワーク機器などの大型計算資源を使用した研究を 対象とした公募型プロジェクトを実施しています。

センターの教員による審査の上,年間で数十件程度の優れた研究提案を採択しております。 採択された課題では申請した計算資源を無料で使用することができます。

前期・後期の半年単位で募集を行う一般枠と、学部学生・大学院生を対象とし、主に夏期に おける利用を想定したインターン制度があります。一般枠(前期・後期)では、1 人で行う研 究課題を募集します。一般枠の課題は、1 年または半年単位(後期は半年のみ)の実施となり ます。インターン制度では、1 人で行う研究課題または 2 人以上のメンバーで構成された研究 グループで行う研究課題を募集します。

一般枠で採択された課題のうち、特に優れた課題で「学際大規模情報基盤共同利用・共同研 究拠点(JHPCN)」の萌芽型共同研究課題の条件を満たすものについては、本センターより同拠 点萌芽型共同研究課題として推薦する予定です。同拠点共同研究課題審査委員会で審査の上、 JHPCNの萌芽型共同研究課題としても採択された場合、毎年7月に開催されるJHPCNのシンポ ジウムにて発表の機会が与えられる場合があります。本制度に採択された課題は終了後、得ら れた成果をもとに、「学際大規模情報基盤共同利用・共同研究拠点(JHPCN)」の公募型共同研究 (一般課題、国際課題、企業課題)等へと進展することが大いに期待されます。なお、インタ ーン制度で採択された課題はJHPCN 萌芽型共同研究課題としての推薦は行いません。

本制度の詳細は、以下の Web ページをご覧ください。 https://www.cc.u-tokyo.ac.jp/guide/young/

3. 2021 年度「若手・女性利用者推薦」前期・後期・インターン 採択課題

2021 年度「若手・女性利用者推薦」課題公募では,前期,後期,インターンに,それぞれ表 1 に示した 10 件,表 2 に示した 8 件,表 3 に示した 1 件の課題が採択されました。本特集号で は前期・後期・インターンで実施された 19 件の課題のうち 9 件の課題の成果報告が掲載されて います。なお,採択された課題のうち 7 件の課題については,2022 年 6 月 27 日に開催された 「若手・女性利用者推薦」成果報告会¹で,その成果が報告されました。

¹ https://www.cc.u-tokyo.ac.jp/events/y_report/2021/2021_report.php

課題名	代表者名	所属
シミュレーションを活用した原子力発電所のリス	九伊 坐于的	東京大学大学院 工学系
ク評価*	入床 儿太郎	研究科
Numerical simulation of solutal Marangoni	ZHANG	大阪大学 基礎工学研究
convection in a shallow rectangular cavity	JIANGAO	科
分子動力学計算で明らかにする金属結合タンパク	木田 陸磁	筑波大学 計算科学研究
質のダイナミクス	秋口 座艇	センター
PaCS-MD と異常検知を援用したタンパク質構造遷	百田 陈亚	筑波大学 計算科学研究
移経路の探索	原田 隆平	センター
分子動力学計算で解明する維持メチル化酵素	化田 灯袋	筑波大学 理工情報生命
DNMT1 の活性化メカニズム	木口 扣軋	学術院
降着円盤乱流における Alfven 的揺動と圧縮的揺動	川盂 送亚	東北大学 学際科学フロ
の配分 [†]	川山 仲平	ンティア研究所
異なるガラス系における降伏臨界性・限界安定性の	大山 倫司	東京大学 総合文化研究
普遍性について [‡]		科
ディンプルの敷設および脈動冷却流によるガスタ		
ービン翼後縁部カットバック面上フィルム冷却の	徳武 太郎	東京農工大学 工学府
高性能化*		
物理法則に基づいた深層学習による多孔質弾性変	正	地球環境産業技術研究機
形の解析*	茨 须	構
キャリアドープされた Kitaev 模型候補物質の繰り	万 世 - 如甬	市市十学 田学 조研究科
込み群による研究 [‡]	油井 须另	木爪八子 生子ボ则"九杆

表1 2021年度「若手・女性利用者推薦」前期 採択課題

前期に採択された全ての課題の実施期間は1年です(採択課題の中に実施期間が半年の課題は ありませんでした)。

†「若手・女性利用者推薦」成果報告会で成果報告された課題です。

‡ 論文投稿中などのために、本特集号には成果報告が掲載されていない課題です。これらの成 果報告は、次号以降に掲載される予定です。

課題名	代表者名	所属
Feedforward型のNeuralNetworkを用いた多体電子	乾 去地	東京大学大学院 工学系
系波動関数の基底状態の解法 ^{†‡}	乾 辛地	研究科
ダイマー粒子を用いたガラスの Johari-Goldstein	百五 蕃亚	東京大学大学院 総合文
beta 緩和の理解 [†]	日口 黒干	化研究科
グラフ構造を入力として応力分布を出力とする		
GAN ベースの有限要素法の Surrogate モデルの研究	中井 優	東京大学 工学部
†‡		
公共データを活用した転写因子結合ダイナミクス	枯眠 和乙	国立国際医療研究センタ
の解析†	旭野 和丁	<u> </u>
波形インバージョンによる地球マントル最下部の	十炮 政众	東京大学大学院 理学系
地震波速度構造推定‡	八锅石刀	研究科
地震波形インバージョンによるマントル最下部の		再与十分十分时 由公文
S・P 波速度構造同時推定 -地球深部の熱・化学進	佐藤 嶺	果尔人子人子阮 理子术
化の理解に向けて-*		柳九科
分子動力学計算を用いた DNA 分解酵素の失活メカ	十済 十年	長崎大学 生命医科学域
ニズムの解明†‡	入僶 八倒	(医学系)
ディープラーニングによる高精度マルウェア分析	二括 力应	東京大学大学院 情報理
Ť	二間 刀杯	工学系研究科

表 2 2021 年度「若手・女性利用者推薦」後期 採択課題

後期採択課題の実施期間は半年です。

* 「若手・女性利用者推薦」成果報告会で成果報告された課題です。

‡ 論文投稿中などのために、本特集号には成果報告が掲載されていない課題です。これらの成 果報告は、次号以降に掲載される予定です。

表3 2021 年度「若手・女性利用者推薦」インターン 採択課題

課題名	代表者名	所属
階層型直交格子と再帰的なフィッティングによる	茜公 土井	東京大学大学院 工学系
回転翼の非定常空力予測	百合 王竹	研究科

インターン採択課題の実施期間は半年です。

Numerical simulation of pure solutal Marangoni convection in a shallow rectangular cavity

ZHANG JIANGAO

Graduate School of Engineering Science, Osaka University

1. Introduction

Marangoni convection usually occurs in a liquid layer with a free surface driven by the variation of free surface tension due to temperature gradients or concentration gradient. This convection has been of interest to researchers due to its rich dynamical features and its wide existence in nature and industrial processes, such as oceanography, droplets, thin-film coatings, and crystal growth.

Numerous fruitful research findings on Marangoni convection have been reported ^[1-4]. However, most of the studies only considered the pure thermal Marangoni effect, while the solutal Marangoni convection plays an important role in coating and crystal growth. To the best of our knowledge, there are few works focused on the solutal Marangoni flow. It would be an interesting phenomenon to investigate since it occurs in some processes. In this work, we perform a series of three-dimensional numerical simulations to study pure solutal Marangoni convection in a shallow rectangular cavity to shed further light on flow characteristics and flow pattern transition.

2. Numerical method

We consider a cavity filled with fluid, as shown in Fig. 1. The top boundary is the free surface, while the other boundaries are the cavity walls. The Cartesian coordinate system is adopted. The length, width, and depth of the cavity are L, L, and 0.1L, respectively. High and low concentration values, C_h and C_l , are set at the four boundary corners, and the linear solutal boundary condition is applied, as shown in Fig. 1(b). The Marangoni convections along the free surface is driven by the surface tension gradient due to the imposed concentration difference.

For simplification, we assume that (i) the free surface does not deform under the effect of fluid flow; (ii) the fluid is incompressible and Newtonian, and its physical properties are constant except for surface tension; (iii) the no-slip boundary condition is applied except for the top free surface, while the solutal Marangoni force is taken into consideration on the free surface.

The governing equations of the fluid flow in the cavity are written in dimensionless forms:

$$\nabla \cdot \mathbf{V} = \mathbf{0} \tag{1}$$

$$\frac{\partial \mathbf{V}}{\partial \tau} + \mathbf{V} \cdot \nabla \mathbf{V} = -\nabla P + \nabla^2 \mathbf{V} \tag{2}$$

$$\frac{\partial \Phi}{\partial \tau} + \mathbf{V} \cdot \nabla \Phi = -\frac{1}{Sc} \nabla^2 \Phi \tag{3}$$

where $\mathbf{V} = (V_x, V_y, V_z)$ is the dimensionless velocity vector, the dimensionless concentration is defined as $\Phi = (C - C_1)/(C_h - C_l)$, τ and P are respectively the dimensionless time and pressure. $Sc = \mathbf{v}/D$ is Schmidt number, where \mathbf{v} is the kinematic viscosity and D is the diffusion coefficient. In this study, the calculation has been performed for the fluid with Sc = 10.



Fig.1: The computational domain and the boundary conditions.

The boundary conditions are defined as follows:

$$\Phi = y \quad (x = 0) \tag{4}$$

$$\Phi = 1 - v \quad (x = 1) \tag{5}$$

$$\Phi = x \quad (v = 0) \tag{6}$$

$$\Phi = 1 - x \quad (y = 1) \tag{7}$$

$$\frac{\partial \Phi}{\partial z} = 0 \quad (z = 0) \tag{8}$$

and for the free surface (z = 0.1),

$$\frac{\partial V_x}{\partial z} = -Ma \frac{\partial \Phi}{\partial x}, \quad \frac{\partial V_y}{\partial z} = -Ma \frac{\partial \Phi}{\partial y}, \quad V_z = 0, \quad \frac{\partial \Phi}{\partial z} = 0$$
(9)

The initial conditions are expressed as follows (τ = 0):

$$V_{x} = V_{y} = V_{z} = 0 \tag{10}$$

$$\Phi = x + y \quad (x + y \le 1) \tag{11}$$

$$\Phi = x + y \quad (x + y > 1) \tag{12}$$

The solutal Marangoni number is defined as

$$Ma = -\sigma_C \frac{(C_h - C_l)L}{\mu\nu}$$
(13)

where $\sigma_c = \partial \sigma / \partial C$ (<0) is the surface tension coefficient of the concentration field.

The finite volume method with nonuniform grid is applied to discrete the governing equations and boundary conditions, which are solved by the pressure-implicit splitoperator (PISO) algorithm. The computation is carried out by using the open source software OpenFOAM. Details of the implementation of the numerical procedure and the validation of the numerical method has been conducted in our previous work.

3. Results and discussion

3. 1. Basic flow

Figure 2 shows the concentration and velocity distributions along the AB and CD diagonals at different solutal Marangoni numbers. The positions of AB and CD lines are shown in Fig. 2 (b). Large concentration drops appear near the boundary corners due to the low mass diffusivity so that the concentration gradient in the middle part of the cavity is smaller than that of the corners, as seen in Fig. 2 (a). Also, as the solutal Marangoni number becomes larger, the concentration drops increase, and the solutal Marangoni effect enhances, resulting in a stronger fluid flow. Furthermore, every peak of the curves in Fig. 2 (b) coincides with each vortex-center position.



Fig.2: Distributions of concentration Φ (a) and velocity U (b) along the diagonals on the free surface. The solid and dash profiles represent the diagonals AB and CD, respectively.

3. 2. Oscillatory flow

In order to extract the three-dimensional disturbances, a fluctuation quantity $(\delta \Phi)$ during one period is introduced as follows,

$$\delta \Phi(x, y, z) = \Phi(x, y, z) - \frac{1}{\tau_p} \int_{\tau_0}^{\tau_0 + \tau_p} \Phi(x, y, z) d\tau$$
(14)

Figure 3 presents the computed snapshots of concentration fluctuations on the free surface and the corresponding space-time diagram, which consists of a series of parallel stripes. It is evident from Fig. 3(a-c) that in all three cases the maximum fluctuation of concentration is proportional to *Ma* number.



(c) $Ma = 2.5 \times 10^4$

Fig.3: Snapshots of the surface concentration fluctuations (left) and the corresponding space-time diagram (right) along the AB line (see the definition in Fig.2(b)).

At $Ma = 1.3 \times 10^4$, the flow has destabilized and bifurcates to an oscillatory flow. The wave pattern propagates from the low concentration spot (point C, D) to the central region of the free surface, which is synchronously accompanied by the weakening of concentration fluctuation during the propagation process, as shown in the space-time diagram in Fig. 3(a). As solutal Marangoni number increases, i.e., at $Ma = 2 \times 10^4$, concentration fluctuation almost covers the whole free surface, as seen in Fig. 3(b). Also, the wave pattern transits from the spoke pattern to the U-type pattern. In Fig. 3(a) and (b), the wave patterns always keep symmetric along the diagonals, which is called the "SW1" hereafter.

When the solutal Marangoni number is larger enough, i.e., at $Ma = 2.5 \times 10^4$, the rotation wave (RW) appears. This secondary instability should be attributed to the appearance of the pitchfork bifurcation, as pointed previously ^[6].

With the further increase of solutal Marangoni number, as shown in Fig. 4, the fluctuation of concentration close to the high concentration spot (point A, B) further enhances and covers the most region of the free surface. Wave pattern only keeps symmetric along the line AB, called the "SW2". Moreover, as the quasi-oscillatory flow appears, the oscillation period becomes larger, the stripes are more complex, shown in the space-time diagram of Fig. 4.



Fig.4: Snapshots of the surface concentration fluctuations (left) and the space-time diagram (right) along the CD line when $Ma = 3.25 \times 10^4$.

4. Conclusions

A series of three-dimensional simulations have been carried out on the solutal Marangoni convection with a linear solutal boundary condition. The following conclusions are obtained from this work:

(1) At the relatively small solutal Marangoni number, concentration drops appear due to the low mass diffusivity, which significantly influences the fluid-structure.

(2) Once the flow destabilizes, as the Marangoni effect enhances, the wave patterns change from SW1 (Symmetric wave 1) to RW (Rotation wave), and then to SW2 (Symmetric wave 2).

参考文献

 M. K. Smith, "Instability mechanisms in dynamic thermocapillary liquid layers" Phys. Fluids 29, 3182-3186, 1986.

[2] L. Peng, Y. R. Li, W. Y. Shi, N. Imaishi, "Three-dimensional thermocapillary-buoyancy flow of silicone oil in a differentially heated annular pool" Int. J. Heat Mass Transfer 50, 872-880, 2007.

[3] H. Liu, Z. Zeng, L. Yin, Z. Qiu, L. Zhang, "Effect of the Prandtl number on the instabilities of the thermocapillary flow in an annular pool" Phys. Fluids 31, 034103, 2019.
[4] T. Yano, M. Hirotani, K. Nishino, "Effect of interfacial heat transfer on basic flow and instability in a high-Prandtl-number thermocapillary liquid bridge" Int. J. Heat Mass Transfer 125, 1121-1130, 2018.

[5] J. M. Zhan, Z. W. Chen, Y. S. Li, Y. H. Nie, "Three-dimensional double-diffusive Marangoni convection in a cubic cavity with horizontal temperature and concentration gradients" Phys. Rev. E 82, 066305, 2010.

[6] J. M. Zhan, Z. W. Chen, Y. S. Li, Y. H. Nie, "Three-dimensional double-diffusive Marangoni convection in a cubic cavity with horizontal temperature and concentration gradients" Phys. Rev. E 82, 066305, 2010.

分子動力学計算で明らかにする

金属結合タンパク質のダイナミクス

森田陸離

筑波大学計算科学研究センター

1. 背景

生体内には多種の微量金属が存在しており、様々な生命現象に関与している。タンパク質に金 属が配位されることでその活性や機能が制御されるためには、細胞内において局所的に金属イオ ン濃度が調整される必要がある。このような過程においてメタロチオネイン(MT)は金属シグ ナルの伝達・調整を担う^{1,2}。MTの主な機能としては(1)金属の貯蔵、(2)有害金属の解毒、(3) 抗酸化作用である。(1)MTは金属イオンを結合あるいは放出することで、金属イオンの濃度を 局所的に調整することができる³。通常MTは亜鉛イオンに結合している。(2)水銀やカドミウム などの有害な重金属が臓器に蓄積すると中毒症状を起こしてしまう^{4,5}。重金属への暴露は細胞 においてMTの発現を誘導することが知られている。MTは親和性の違いに基づいて亜鉛イオン の代わりに重金属イオンと結合する。この時、放出された亜鉛イオンが転写因子を活性化し、M Tの転写を誘導する。(3)活性酸素種のひとつであるヒドロキシラジカルに対して、MTはラジ カルスカベンジャーとして機能する⁶。有害金属の解毒と同様に、ラジカルと反応したシステイ ンは亜鉛イオンを放出し、MTの転写を誘導する。

MTの構造に着目すると、わずか60アミノ酸程度からなるタンパク質であり、全体の3分の 1を含硫アミノ酸であるシステインが占める⁷。酸化状態のシステインは4つ組になって四面体 を形成し、主に亜鉛を配位する。その構造はフレキシブルであり、X線結晶解析による決定は困 難であったが、NMRによる構造が報告されている^{8,9}。しかし、典型的な構造のみならず、金属イ オンの個数や配位によって多様な構造を取りうることが指摘されている。

そこで、本研究では**分子動力学計算**(MD)によって構造をサンプリングすることを目的とした。MDは、運動方程式に基づいて原子の運動をシミュレーションする。MDはタンパク質の構造変化について時空間的に高解像度の情報を得ることができる。さらに、周囲の環境を任意に変えることが可能であるため、実験では取得困難な条件でのシミュレーションを行える。本研究では、マウス MT-1 のβドメインを使用して比較的長い時間(10 μs)のシミュレーションを行った。

2.方法

マウス MT-1 の β ドメインの立体構造は Protein Data Bank (PDB) より入手した (PDB ID: 1DFT) ⁹。すでに配位している金属や水分子は予め取り除いた。タンパク質側鎖の水素原子は GROMACS 2019¹⁰ に実装されている gmx2pdb を使用して補完された。タンパク質についての力場は Amber ff14SB¹¹ を使用し、亜鉛については改良されたモデル¹² を使用した。MTは 20Å のマージンを持 ったボックスに配置され、TIP3P モデル¹³で溶媒和された。亜鉛イオンを含む系(Holo-MT)につ いては、40 個の亜鉛イオンを追加した。また、系を中和するために水分子をランダムに選択し、 Na+または C1-に置換した。クーロン相互作用の計算には Particle Mesh Ewald 法¹⁴を使用した。 すべてのシミュレーションには GROMACS 2019 を使用した。系を平衡化するために以下の処理を 行った。(1)10,000 ステップのエネルギー最小化、(2)V-rescale サーモスタット¹⁵を使用した NVT 条件下で 100 ps の平衡化、(3)Berendsen カップリング¹⁶を使用した NPT 条件下で 100 ps の 平衡化。最後に NPT 条件下(300 K、1 atm)で 10 μs のプロダクションランを行った。

3. 結果と考察

まず、亜鉛イオンを含まない条件 (Holo-MT) においてMTの構造をサンプリング、分類した(図 1)。ランダムコイル構造がいくつかみられた他は、全体の 92% が球状の構造をとった(図3)。 この代表構造は NMR でみられるものと類似しているが、複数の短いαヘリックスを持つという特 徴がある。





距離地図を用いた階層的クラスタリングによって分類された Holo-MT の構造群とその代表構造を示す。カッコ 書きの数字は 1000 構造中その構造群が占める数を表す。

一方で、亜鉛イオンを含まない条件(Apo-MT)においても構造サンプリングおよび分類を行った(図2)。Holo-MTと同様にランダムコイル構造が現れたが、全体の91%を占めたのは伸展したβシート構造であった(図3)。



Apo-MT のクラスタリング結果を図1と同様に示す。球状の構造は初期構造に由来すると考えられる。

構造の可逆性を調べるために、Apo-MT の構造を 5 つ抽出し、亜鉛イオンを加えて追加のシミ ュレーションを行った。この時、マイクロ秒のオーダーでαヘリックス構造が形成された。この 結果は、Apo-MT と Holo-MT が局所的な亜鉛イオン濃度の変化に迅速に応答するのに十分な構造 可逆性を持つことを示している。実験による二次構造の決定には円二色性スペクトルが使用され てきた¹⁷⁻¹⁹。これまでの実験による研究では金属イオンの添加された条件においてMTの構造変化を抽出することに焦点が当てられていた。そのため、Apo-MT がどのような構造を取りうるかという情報はなかった。この点において、本研究で得られた Apo-MT の構造は興味深いものであり、今後の実験的な検証が期待される。



図3 (左) Holo-MT の代表構造。(右) Apo-MT の代表構造 Holo-MT、Apo-MT の代表構造を拡大して示す。9箇所のシステインはスティック表示で示されている。

Holo-MT の構造に着目すると、平均して 1.55 個の亜鉛イオンが結合していた。NMR 構造では 3 個の亜鉛イオンが観察される。この違いについては 2 つの原因が推測される。1 つ目は、亜鉛イ オンのパラメータの問題である。本研究では亜鉛イオンを方向性を持たない 1 つの粒子として取 り扱った。一般的に亜鉛イオンは 4 つの配位子と方向性を持った配位合をとる。これを改善する ためにカチオン性ダミー原子モデル (CADA) が考案されている²⁰。CADA では質量を持つ粒子の周 りに電荷を分散して持つ 4 つの原子を配置することで、より精度良く配位結合を表現する事がで きる。このようなモデルの使用で得られる結果に違いが生じる可能性は残されている。一方で、 従来の単粒子モデルを使用してもよい結果が得られている結果は複数存在している。2 つ目の可 能性は、時間スケールの問題である。亜鉛が配位するターンオーバーは 10⁻⁸ 秒程度で生じる²¹。 これは、NMR の時間分解能(10⁻⁴ 秒)やタンパク質の構造変化にかかる時間スケール(10⁻⁶ 秒) と比べて非常に高速であり、NMR では平均的な原子の存在のみが記録されている可能性に注意し なければならない。

次に、塩濃度についての懸念を検証した。Holo-MT において、多くの亜鉛イオンおよび中和の ためのカウンターイオンが導入された。その結果として塩濃度が上昇しており、Holo-MT の構造 変化が生じた可能性がある。そこで、亜鉛イオンの代わりにナトリウムイオンを加えて3 μs の シミュレーションを行った(Highsalt-MT)。その結果、Apo-MT と同様にβシート構造が見られた (図4、6)。このことから、単なる塩濃度の上昇ではなく、亜鉛イオンの存在が Holo-MT の構 造形成に必要であることが示唆される。



高塩濃度環境下でのMTの構造クラスタリング結果を図1と同様に示す。全体の構造数は300スナップショットである。

また、Holo-MTでは9つのシステインのうち2番目と3番目のシステイン(C2、C3)に亜鉛イ オンが安定して結合した。この理由として、C2およびC3が亜鉛イオンでキャップされることが 構造形成に関わるのではないかと考えた。そこで、C2およびC3のチオール基(-S⁻)をプロトン でキャップすることで、チオレート基(-SH)として 3 μs のシミュレーションを行った(SemiApo-MT)。その結果、Holo-MT で見られたようなαヘリックスを持つ球状の構造が多くみられた(図 5、6)。このことから Holo-MT 構造はC2およびC3のキャッピングによって誘導される事が示 唆された。しかし、なぜ亜鉛イオンがC2およびC3に安定して結合するのかは不明である。



SemiApo-MT についてのクラスタリング結果を図4と同様に示す。



図 6 (左) Highsalt-MT の代表構造。(右) SemiApo-MT の代表構造 Highsalt-MT および SemiApo-MT の代表構造を示す。

C2の周辺で金属に配位している残基を調べた。その結果、Asp2がC2と共に亜鉛イオンと結合 していることが明らかになった。Asp2はMTのβドメイン唯一の酸性アミノ酸であり、哺乳類の MTにおいて進化的に保存されている。従来はシステインのみが亜鉛結合に関わるとされていた が、Asp2も重要な働きをしていることが明らかになった。EF ハンドなどの一部の金属結合モチ ーフは主にアスパラギン酸を利用している。Asp2 に対する点変異実験によって結合親和性への 寄与を調べることが期待される。



図7 Asp2の周辺構造

Asp2 周囲の亜鉛配位構造を示す。紫の点線は金属配位に関与すると思われる結合を表し、その距離を示した。

4. 結論

本研究では外部環境を変えた MD によって亜鉛イオンの有無による MT の構造多様性を明らか にした。金属のある環境で形成される a ヘリックスは構造を安定させ、MT の金属配位を確実に する可能性がある。また、金属の無い環境で形成される β シートは不安定であった。先行研究に よると、Apo-MT はプロテアーゼによる処理で Holo-MT より早く分解されることがわかっている。 このことから金属イオンが欠乏した環境では不必要となる MT が不安定な β シートを形成して速 やかに分解されることが考えられる。また、MT は金属を保持したり放出したりすることで間接 的に機能を発揮すると考えられている。しかし、MT が構造変化するという事実より、異なるタン パク質とMT が結合することが示唆される。このような相互作用の詳細な解明は金属シグナルの 伝達の分子メカニズムを明らかにするだろう。無数のタンパク質から MT の結合パートナーを発 見するには、更なる実験的あるいは情報科学的研究が必要になる。

謝辞

本研究は 2021 年度東京大学情報基盤センター若手・女性利用(「分子動力学計算で明らかにす る金属結合タンパク質のダイナミクス」)によって支援され、スーパーコンピューター Reedbush を用いて実施された。

参考文献

1

Kagi, J. H. & Valee, B. L. Metallothionein: a cadmium- and zinc-containing protein from equine renal cortex. *J. Biol. Chem.* **235**, 3460-3465, doi:10.1016/s00219258(18)64490-4 (1960).

- 2 Margoshes, M. & Vallee, B. L. A Cadmium Protein from Equine Kidney Cortex. J. Am. Chem. Soc. 79, 4813-4814, doi:DOI 10.1021/ja01574a064 (1957).
- 3 Suhy, D. A., Simon, K. D., Linzer, D. I. & O'Halloran, T. V. Metallothionein is part of a zinc-scavenging mechanism for cell survival under conditions of extreme zinc deprivation. J. Biol. Chem. 274, 9183-9192, doi:10.1074/jbc.274.14.9183 (1999).
- 4 Durnam, D. M. & Palmiter, R. D. Transcriptional regulation of the mouse metallothionein-I gene by heavy metals. *J. Biol. Chem.* 256, 5712-5716, doi:10.1016/s0021-9258(19)69264-1 (1981).
- 5 Heuchel, R. *et al.* The transcription factor MTF-1 is essential for basal and heavy metal-induced metallothionein gene expression. *The EMBO Journal* 13, 2870-2875, doi:10.1002/j.1460-2075.1994.tb06581.x (1994).
- 6 Dalton, T., Palmiter, R. D. & Andrews, G. K. Transcriptional induction of the mouse metallothionein-I gene in hydrogen peroxide-treated Hepa cells involves a composite major late transcription factor/antioxidant response element and metal response promoter elements. *Nucleic Acids Res.* 22, 5016-5023, doi:10.1093/nar/22.23.5016 (1994).
- 7 Winge, D. R. & Miklossy, K. A. Domain nature of metallothionein. J. Biol. Chem. 257, 3471-3476, doi:10.1016/s0021-9258(18)34802-6 (1982).
- 8 Peterson, C. W., Narula, S. S. & Armitage, I. M. 3D solution structure of copper and silver-substituted yeast metallothioneins. *FEBS Lett.* 379, 85-93, doi:10.1016/0014-5793(95)01492-6 (1996).
- 9 Zangger, K., Öz, G., Armitage, I. M. & Otvos, J. D. Three-dimensional solution structure of mouse [Cd7]-metallothionein-1 by homonuclear and heteronuclear NMR spectroscopy. *Protein Sci.* 8, 2630-2638, doi:10.1110/ps.8.12.2630 (2008).
- 10 Abraham, M. J., van der Spoel, D., Lindahl, E., Hess, B. & the GROMACS development team. GROMACS User Manual version 2019. (2019).
- Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696-3713, doi:10.1021/acs.jctc.5b00255 (2015).
- Macchiagodena, M., Pagliai, M., Andreini, C., Rosato, A. & Procacci, P. Upgrading and Validation of the AMBER Force Field for Histidine and Cysteine Zinc(II)-Binding Residues in Sites with Four Protein Ligands. *J. Chem. Inf. Model.* 59, 3803-3816, doi:10.1021/acs.jcim.9b00407 (2019).
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 79, 926-935, doi:Doi 10.1063/1.445869 (1983).
- 14 Ewald, P. P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. Ann Phys 369, 253-287, doi:10.1002/andp.19213690304 (1921).
- 15 Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling.

J. Chem. Phys. 126, 014101, doi:10.1063/1.2408420 (2007).

- Berendsen, H. J. C., Postma, J. P. M., Vangunsteren, W. F., Dinola, A. & Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* 81, 3684-3690, doi:Doi 10.1063/1.448118 (1984).
- 17 Cai, W. & Stillman, M. J. (Mercury)18-metallothionein. J. Am. Chem. Soc. 110, 7872-7873, doi:10.1021/ja00231a045 (2002).
- 18 Rupp, H. & Weser, U. Circular dichroism of metallothioneins. A structural approach. Biochim. Biophys. Acta 533, 209-226, doi:10.1016/0005-2795(78)90565-2 (1978).
- 19 Scheller, J. S., Irvine, G. W. & Stillman, M. J. Unravelling the mechanistic details of metal binding to mammalian metallothioneins from stoichiometric, kinetic, and binding affinity data. *Dalton Trans* **47**, 3613-3637, doi:10.1039/c7dt03319b (2018).
- 20 Pang, Y. P. Novel zinc protein molecular dynamics simulations: Steps toward antiangiogenesis for cancer treatment. *J. Mol. Model.* 5, 196-202, doi:DOI 10.1007/s008940050119 (1999).
- 21 Gray, H. B. & Langford, C. H. Ligand Substitution Dynamics. *Chemical & Engineering News Archive* 46, 68-76, doi:10.1021/cen-v046n015.p068 (1968).

PaCS-MD と異常検知を援用したタンパク質構造遷移経路の探索

原田隆平

筑波大学計算科学研究センター

1. はじめに

2021 年度(前期)スーパーコンピューター若手・女性利用者推薦制度の採択課題「PaCS-MDと 異常検知を援用したタンパク質構造遷移経路の探索」の研究成果を報告する.報告内容は,採択 期間において開発した生体機能に関係するタンパク質の構造変化を効率的に探索する計算手法 (構造サンプリング法)の開発である.

2. イントロダクション

生体内ではタンパク質をはじめとする生体分子が機能を担い、生命を維持している.タンパ ク質は構造揺らぎを巧みに利用することで機能を発揮している.例えば、ドメイン間の大規模 な構造変化により基質(リガンド)を認識したり、他のタンパク質と会合して複合体構造を形成 したり、様々な構造変化が生体機能に関係している.これらの構造変化を追跡することは、生 体機能の解析において非常に重要である.構造変化を直接追跡する計算科学的手法として分子 動力学シミュレーション(MD)は強力であり、理論研究者のみならず実験研究者にも利用され ている.MD は、運動方程式を逐次的に解くことで時々刻々変化するタンパク質ダイナミクスを 追跡できる.しかしながら、生体機能に関係するタンパク質の大規模構造変化は通常のMD が追 跡可能な時間スケールよりも遥かに長時間で観測される「生物学的レアイベント」であるため、 抽出が困難である.MD 専用計算機[1,2]などを利用すれば、タンパク質のサイズは小さいもの のミリ秒に及ぶ長時間 MD が可能になってきてはいるものの、これらの極めて高価な計算機を利 用できる研究者は限られており、全ての研究者に解放されているわけではない、故に、現実的 かつ合理的な計算コストによりタンパク質の長時間ダイナミクスを抽出する「レアイベントサン プリング手法」の開発が望まれている.

本研究では、タンパク質の構造遷移探索法として開発したカスケード選択型分子動力学計算 (PaCS-MD: Parallel Cascade Selection MD) [3] と機械学習の1つである異常検知を援用し、生 体機能に重要なタンパク質の長時間ダイナミクス(レアイベント)を効率的に抽出する「異常検 知型 PaCS-MD (ad-PaCS-MD)」を開発した.具体的には、「構造変化のし易さ」と「構造の異常さ」 を異常検知により対応付け異常度が高い構造を特定し、PaCS-MDの初期構造に採用する.PaCS-MD の各サイクルにおいて「異常度が高い構造 = 遷移確率が高い構造」から短時間 MD を繰り返すこ とで構造遷移を誘起し、構造変化(遷移経路)を自動的に予測する ad-PaCS-MD を開発する.こ こで ad-PaCS-MD の開発に先立ち、PaCS-MD の構造サンプリング戦略の概要を記載する.

PaCS-MD はタンパク質の構造変化(構造遷 移)を効率的に探索可能なレアイベントサン プリング手法である、図1 に示すように PaCS-MDの具体的な計算手順は、(1) 複数の短 時間MDから遷移確率の高い分子構造を選択し. (2) 選択構造を初期構造として再び短時間 MD を実行し繰り返すことで効率的な構造探索を 可能にする、非常に単純な計算スキームであ るが、(1)と(2)からなるサイクルを繰り返すだ けで効率的に生体分子の構造遷移を誘起させ ることができる、常に遷移確率の高い分子構 造を選択して選択した分子構造から短時間 MD を繰り返すサイクルにより、生体機能に関係 する重要な構造変化を効率的に誘起する. PaCS-MD は様々なターゲットに適用され、生体 機能に重要な構造遷移の抽出に成功してきた. [4-6]



図 1: PaCS-MD の概念図

PaCS-MD の詳細に関してはレビュー論文

[7,8] を参照にされたい. PaCS-MD は、バイアスをかけることなく構造間の遷移経路を探索する. 従来法では、構造遷移を促進するため系の温度を高温にして MD を実行するか,系の特定の部分 に外力や外部ポテンシャルといった「バイアス」を付加することがほとんどである「9,10].従 来法に対して本手法は、反応座標を設定してしまえば、遷移確率が高い初期構造を選択し、短 時間 MD により構造リサンプリングを繰り返していくだけなので、バイアスは使用しない.更に、 パラメーターを最適化するための準備計算も必要としない。本手法により効率的に構造探索が 可能であるのは、構造リサンプリングを通して「稀にしか発生しない分子構造」の出現確率を選 択的に上昇させているからである。初期構造の選択にあたり様々な反応座標が考えられる。あ らかじめ遷移先の終構造が分かっている場合に関しては、終構造に対する平均自乗距離 (root-mean-square deviation: RMSD) が考えられる. 例えば、構造リサンプリングにより得ら れるトラジェクトリを RMSD の値でランキングし, 値の小さい順にスナップショットを複数個選 択し, 短時間 MD の初期構造として構造探索していくことにより, 徐々に終構造へ近づいていき, レアイベントを抽出することができる.ここで、反応座標は任意に指定可能であり、RMSD に限 らず指定できる.実際の適用では、抽出したいレアイベントを的確に記述する合理的な反応座 標を指定することが重要になる.また、遷移先の構造が未知である場合、PaCS-MDの拡張も必要 になる. PaCS-MD の拡張にあたり、遷移先の終構造が未知である場合、かつ、反応座標の指定が できるだけシステムに依存しないようする、という2点が重要になってくる.



図 2: ad-PaCS-MD の設計思想

本研究では PaCS-MD の拡張として、タンパク質の構造遷移を効率的に抽出する PaCS-MD と機 械学習(異常検知)を援用することで生体機能に重要なタンパク質の長時間ダイナミクス(レア イベント)を効率的に抽出する「異常検知型 PaCS-MD (ad-PaCS-MD)」を開発した(図2).PaCS-MD は,構造遷移を誘起する確率が高い初期構造を特定し、短時間の分子動力学計算(MD)を繰り返す ことで効率的にレアイベントを抽出する.本研究では PaCS-MD を拡張し、より効率的にレアイベ ントを抽出可能な方法へ発展させた.具体的には、「構造変化のしやすさ」と「構造の異常さ」を 異常検知により対応付けて異常度が高い構造を特定し、PaCS-MDの初期構造に採用する.対応付け は、タンパク質の残基間距離を要素とする距離行列を画像として蓄積して機械学習し、異常度が 高い距離行列を他の準安定状態に構造遷移しやすいタンパク質構造とみなした.ad-PaCS-MDの各 サイクルにおいて「異常度が高い構造 = 遷移確率が高い構造」から短時間 MD を繰り返すことで 構造遷移を誘起し、構造変化(遷移経路)を自動的に予測可能な ad-PaCS-MD を開発した.以上の 開発思想に基づき Reedbush-H 上で ad-PaCS-MD を実装した.

ad-PaCS-MD のサンプリング効率を調べるため,加水分解酵素である T4 リゾチーム(T4L) に 適用して Open-Closed 構造遷移をレアイベントとして抽出できるか検証した.具体的には T4L の Open 構造から MD を実行し,得られたトラジェクトリを機械学習したのち,異常検知システムを構 築した.この異常検知システムを用いてトラジェクトリに含まれる異常度の高い(出現確率の低 い)タンパク質構造を抽出し,短時間 MD を繰り返した.結果として,T4L の Open 構造からスター トして Closed 構造へ至る遷移経路を抽出することに成功した.ad-PaCS-MD から得られたトラジ ェクトリをもとにマルコフ状態モデルを構築し,自由エネルギーを計算した(図 3 左).自由エネ



図 3: ad-PaCS-MD(左)と 通常の MD(右)の探索効率の比較

ルギー地形が示すように、Open 構造からスタートして Closed に至る 2 構造間の遷移経路を抽出 することができている.通常の MD ではこのようなレアイベントを抽出することが難しいた め、Open 構造からスタートしても Closed 構造に至る遷移経路を抽出することはできない(図 3 右).以上より、生体分子に ad-PaCS-MD を適用することで、現実的な計算コストで効率的にレアイ ベントを抽出でき、生体機能の解析に貢献することが期待される.現在、ad-PaCS-MD を複合体の解 離プロセス抽出に適用した応用研究を実施している.複合体の場合は、タンパク質間相互作用を 距離行列で的確に記述できるため、詳細な解離課程を抽出できる.

4. まとめと今後の展望

PaCS-MDをad-PaCS-MDとして拡張することにより,遷移先の構造が未知である場合にも適用 可能になり,出発構造の周辺に存在する順安定状態への遷移経路が自動的に抽出できるように なる.また,タンパク質の残期間距離に基づく距離行列を初期構造に採用しているため、シス テム依存で反応座標を定義する必要がなく,汎用性が高い.今後はad-PaCS-MDを適用し、タン パク質とタンパク質,あるいは、タンパク質-基質といった複数の分子や化合物が結合して複合 体を形成する秩序形成過程や解離過程といった複雑な構造変化を抽出していきたい.拡張にあ たり、距離行列は複合体に関しても容易に定義できる.

5. 謝辞

本研究は、東京大学情報基盤センター「若手・女性利用者推薦」2021 年度(前期)の採択 課題として、Reedbush-Hを利用して実施された.

参考文献

- D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R.O. Dror, M.P. Eastwood, J.A. Bank, J.M. Jumper, J.K. Salmon, Y.B. Shan, W. Wriggers, *Science* 330 (2010) 341-346.
- [2] K. Lindorff-Larsen, S. Piana, R.O. Dror, D.E. Shaw, *Science* 334 (2011) 517-520.
- [3] R. Harada, A. Kitao, *J. Chem. Phys.* 139 (2013) 035103.
- [4] T. Baba, R. Harada, M. Nakano, Y. Shigeta, J. Comput. Chem. 35 (2014) 1240-1247.
- [5] R. Sato, R. Harada, Y. Shigeta, Biophysics and Physicobiology 13 (2016) 311-319.
- [6] J. Fujita, R. Harada, Y. Maeda, Y. Saito, E. Mizohata, T. Inoue, Y. Shigeta, H. Matsumura, J. Struct. Biol. 198 (2017) 65-73.
- [7] R. Harada, Y. Takano, T. Baba, Y. Shigeta, Phys. Chem. Chem. Phys. 17 (2015) 6155-6173.
- [8] R. Harada, Bull. Chem. Soc. Jpn. 91 (2018) 1436-1450.
- [9] G.M. Torrie, J.P. Valleau, J. Comput. Phys. 23 (1977) 187-199.
- [10] G.M. Torrie, J.P. Valleau, J. Chem. Phys. 66 (1977) 1402-1408.

分子動力学シミュレーションで解明する

維持メチル化酵素(DNMT1)の活性化メカニズム

保田拓範

筑波大学大学院 理工情報生命学術院 生物学学位プログラム

1. はじめに

2021 年度(前期)スーパーコンピューター若手・女性利用者推薦制度の採択課題「分子動力学 シミュレーションで解明する維持メチル化酵素(DNMT1)の活性化メカニズム」の研究成果を報告 する.なお,本稿では概略のみを記載するため,詳細については Journal of Molecular Biology(JMB)誌に掲載されている論文(<u>https://doi.org/10.1016/j.jmb.2021.167371</u>)を参照し て頂きたい.

2. イントロダクション

すべての細胞は基本的に同一の DNA 配列を有しているものの、固有の性質を持つ、これは、エピ ゲノムと呼ばれる DNA 修飾が異なるためである. DNA メチル化は代表的なエピゲノムの一種で、組 織や細胞で特異的な遺伝子発現を誘導することが知られている.そのため,細胞分化を始めとす る様々な生体内のプロセスにおいて重要な役割を担っている. 例えば、ゲノムインプリンティン グや X 染色体の不活性化は,DNA メチル化によって引き起こされるイベントの代表例である[1]. この DNA メチル化は主に2種類存在することが知られている.1 つ目は **de novo メチル化**で,メチ ル化パターンの確立を担っている.より詳細には、初期胚形成の間に、両鎖ともメチル化されてい ない DNA 鎖をメチル化する役割を持つ、この過程により、メチル化パターンが確立されると、細胞 分裂の間,そのメチル化パターンを維持することが必要になり,これが2つ目のメチル化機構に あたる維持メチル化である.具体的には,DNA 複製の際,新生鎖に DNA メチル化パターンをコピー し,エピゲノム情報を親細胞から娘細胞に引き継ぐという役割を担っている[2,3].一般に,この ような DNA メチル化は DNA メチル化酵素ファミリー (Dnmt1, Dnmt2, Dnmt3A, Dnmt3B, Dnmt3L) によ り制御されている[4]. DNA メチル化酵素(Dnmt)ファミリーをその機能により大分する と、Dnmt3A、3B、3L は主に de novo メチル化に関与し、Dnmt1 は主に維持メチル化に関与している [5]. 本報告書のテーマでもある維持メチル化酵素1 (Dnmt1) に着目すると,機能的にはヘミメチ ル化 DNA にメチル基を付加する酵素であるといえる.また、DNA メチル化パターンは遺伝子発現に 影響するため, Dnmt1は細胞の維持に不可欠な役割を担っており, 細胞癌を始めとした様々な疾患 と密接に関連していることが知られている[6].

Dnmt1 の活性は DNA 複製時にのみ要求されるため高度に制御される必要があり,結晶構造解析 及び分子生物学的な実験により様々な活性制御メカニズムが報告されている.結晶構造解析では, 複数の Dnmt1 の立体構造が決定され,約 1600 アミノ酸残基からなり,特徴的なドメインを持つこ とが明らかになってきた[7,8].着目すべき点として,Replication Foci Targeting Sequence domain ドメイン(RFTS) が Dnmt1 の触媒ドメインに入り込んでいることが明らかになった.つま り, DNA 複製時以外は Dnmt1 の DNA 結合部位が露出しておらず,活性がない. そのため,活性化には Dnmt1 の大規模な構造変化が必要であると考えられてきた.一方,分子生物学的な実験によ り,RFTS が Dnmt1 をヘミメチル化 DNA へと部位特異的にリクルートする役割を担っていることが 明らかにされている[9,10].より詳細には,Ubiquitin like PHD and RING finger domains 1 (Uhrf1)がターゲット DNA の Histone H3 tail (H3T) をユビキチン化する[11].更に,この Uhrf1 依存的なユビキチン化を,RFTS が認識することで DNMT1 がヘミメチル化部位にリクルートされる ことになる[12,13].このように,ヒストンのユビキチン化は Dnmt1 による部位特異的な DNA メチ ル化のマーカーとみなされている.さらに,近年の実験的・計算科学的な研究により,2箇所でモ ノユビキチン化された H3T が RFTS に結合し,RFTS は触媒部位から乖離,Dnmt1 が活性を持つこと が明らかになった[14].この際,Dnmt1 は複数のユビキチン化パターン (K14/K18, K14/K23, K18/23)を有する H3T と結合し,活性化されることが明らかになっている.興味深いことに,上述 のユビキチン化パターンの違いは,Dnmt1 との結合親和性を変化させることも明らかになってい る.具体的には,K14 および K18 の2箇所でモノユビキチン化 H3T (K14/K18) は Dnmt1 に対して最 も高い結合活性を示し,K14 および K23 の2箇所でモノユビキチン化された H3T (K14/K23) は最 も低い結合活性を示すが,H3T を結合していない2つのユビキチン分子(Free Ubiquitin)は Dnmt1 にほとんど結合しない[14].

このような先行研究により, DNMT1 が2箇所でユビキチン化された H3T により活性化され, ヘミ メチル化部位で特異的に活性を持つというメカニズムが明らかになってきた.しかし, 活性化に おける H3T の役割を始めとした詳細な分子メカニズムやユビキチン化パターンの差により結合 活性の差が生じるメカニズムについては不明であった.そこで, 本研究ではユビキチン化ヒスト ンによる Dnmt1 活性化メカニズムの解明を目的として,(1) K14 / K18 でモノユビキチン化され た H3T (2) K14 / K23 でモノユビキチン化された H3T (3) H3T と結合していない 2 つのユビキチ ン, の3システムに関して全原子分子動力学計算(MD)シミュレーションを実施した(図-1). 最 終的に, H3T の構造的な柔軟性が 2 つのユビキチンの構造状態を制御しており, 2 つのモノユビキ チン間の相互作用を阻害することで Dnmt1 との結合を促進する役割を持つことがわかった.



第1図:計算で使用した構造およびヒストンH3 tailのアミノ酸配列

(a) K18/K23の2箇所でモノユビキチン化されたH3T(青)と2つのユビキチン分子(Ub1,Ub2,赤)の複合体結晶構造(PDBid:5WV0)(b) Dnmt1(緑)と2箇所でモノユビキチン化されたH3Tの複合体結晶構造(PDBid:5WV0).(c)H3Tのアミノ酸配列.3つのユビキチン化部位(K14,K18,K23)はハイライト(赤色)で表示.

3. 結果

Dnmt1 の活性化メカニズムを分子レベルで解明するため,(1)K14/K18,(2)K14/K23 でモノユビ キチン化された H3T,(3)H3T に結合していない2つのユビキチン分子(Free ubiquitin)の3つ のシステムについて全原子 MD シミュレーションを実施した.解析の結果,H3T が存在する場合は, ユビキチン化パターンにより,2つのユビキチンにおける,分子間距離及びコンタクトが異なる ことが明らかになった.より詳細には,K14/K18システムではK14/K23システムよりユビキチン間 の分子間距離が長く,分子間コンタクトが多くなっていた.更に,Free ubiquitinの場合,K14/K18 および K14/K23 で H3T と結合したユビキチン分子に比べ,分子間距離が短く,コンタクトが多く なることが明らかになった.(図-2)これらの結果から,2つのユビキチン間における分子間相互作 用は,ユビキチン化のパターンによって調節されていることが示唆された.



第2図: 異なるユビキチン化パターンにおけるユビキチン間の相互作用 (a-c)ユビキチン間の分子間距離における自由エネルギー地形(PMF).(d-f)ユビキチン分子間のコンタクト数に ついての頻度.(a, d) K14/K18のモノユビキチン化H3T.(b, e) K14/K23のモノユビキチン化H3T.(c, f) H3Tと結 合していない2つのユビキチン分子.(a-c)の各 PMFの値は & Tでスケーリング.(d-f)では,コンタクト数がゼロ の場合を除いて,最も高いピークをハイライト(黒色)で表示.

続いて、ユビキチン化部位間のH3Tの距離に着目すると、K14-K18の方がK14-K23と比べて長くなっていることが明らかになった.K14-K18間のH3Tは4残基であるのに対し、K14-K23間は10残基あることを考慮すると非常に興味深い結果であり、ユビキチン化部位間におけるH3Tの構造的な特徴がその差異を生んでいるのではないかと考察した.そのため、ユビキチン化部位間(K14-K23)のH3TペプチドについてのみMDシミュレーションを実施した.その結果、K14からK18までの領域が構造的に硬く、K18からK23までの領域は柔らかいことがわかった.つまり、K14/K23の

H3T のユビキチン化部位間の距離は,K14/18 のユビキチン化部位間の距離よりも構造上の性質から短くなり,これらのユビキチン化部位間の構造がユビキチン分子間相互作用を調節していることが示唆された(図-3).



第3図:H3Tが2つのユビキチンに与える影響

H3T ペプチドの構造的な特徴により,2つのユビキチン分子(Ub1,Ub2)の分子間距離にどのような影響を与えるかを示す模式図.

MD シミュレーションを実施した3つの系において,ユビキチン間の分子間相互作用は先行研究 で測定された Dnmt1 に対する結合活性とよく一致した.具体的には,実験的に得られた Dmnt1 に 対する結合力は,計算で推定されたユビキチン分子間相互作用が強くなるにつれて弱くなること がわかった.以上の解析から,H3Tの構造的な柔軟性が2つのユビキチンの構造状態を制御してお り,H3T はユビキチン間の分子間相互作用を阻害するスペーサーとして働くことで Dnmt1 への結 合を促進していると結論づけた.(図-4)



第4図: ユビキチン間相互作用と DNMT1 との結合活性の関係

ヒストン H3T が存在しない場合,ユビキチン間の相互作用が大きくなり, DNMT1 との結合活性が小さくなる.一方でH3T が存在する場合,相互作用が小さくなることで DNMT1 との結合が促進される.

4. 謝辞

本研究は、東京大学情報基盤センター「若手・女性利用者推薦」2021 年度(前期)の採択 課題として、Reedbush-H を利用して実施された.

参考文献

[1] Li E, Beard C, Jaenisch R. Role for DNA Methylation in Genomic Imprinting. Nature. 1993;366:362-5.

[2] Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet. 2010;11:204-20.

[3] Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. Annu Rev Biochem. 2005;74:481-514.

[4] Okano M, Xie SP, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet. 1998;19:219-20.

[5] Gaudet F, Hodgson JG, Eden A, Jackson-Grusby L, Dausman J, Gray JW, et al. Induction of tumors in mice by genomic hypomethylation. Science. 2003;300:489-92.

[6] Zhang ZM, Liu S, Lin K, Luo YF, Perry JJ, Wang YS, et al. Crystal Structure of Human DNA Methyltransferase 1. J Mol Biol. 2015;427:2520-31.

[7] Song JK, Rechkoblit O, Bestor TH, Patel DJ. Structure of DNMT1-DNA Complex Reveals a Role for Autoinhibition in Maintenance DNA Methylation. Science. 2011;331:1036-40.

[8] Takeshita K, Suetake I, Yamashita E, Suga M, Narita H, Nakagawa A, et al. Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1). P Natl Acad Sci USA. 2011;108:9055-9.

[9] Easwaran HP, Schermelleh L, Leonhardt H, Cardoso MC. Replication-independent chromatin loading of Dnmt1 during G2 and M phases. Embo Rep. 2004;5:1181-6.

[10] Leonhardt H, Page AW, Weier HU, Bestor TH. A Targeting Sequence Directs DNA Methyltransferase to Sites of DNA-Replication in Mammalian Nuclei. Cell. 1992;71:865-73.

[11] Citterio E, Papait R, Nicassio F, Vecchi M, Gomiero P, Mantovani R, et al. Np95 is a histone-binding protein endowed with ubiquitin ligase activity. Mol Cell Biol. 2004;24:2526-35.

[12] Nishiyama A, Yamaguchi L, Sharif J, Johmura Y, Kawamura T, Nakanishi K, et al. Uhrf1dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. Nature. 2013;502:249-+.

[13] Qin WH, Wolf P, Liu N, Link S, Smets M, La Mastra F, et al. DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. Cell Res. 2015;25:911-29.

[14] Ishiyama S, Nishiyama A, Saeki Y, Moritsugu K, Morimoto D, Yamaguchi L, et al. Structure of the Dnmt1 Reader Module Complexed with a Unique Two-Mono-Ubiquitin Mark on Histone H3 Reveals the Basis for DNA Methylation Maintenance. Mol Cell. 2017;68:350-60 e7.

降着円盤乱流における Alfvén 的揺動と圧縮的揺動の配分

川面 洋平

東北大学 学際科学フロンティア研究所

1. はじめに

ブラックホールは物理学者、天文学者、そして一般市民をも惹きつける謎に満ちた天体である。 しかし、ブラックホールは光を発しないため観測することが困難である。ではブラックホールを "見る"ためにはどうすればよいのだろうか。手がかりはブラックホールを取り巻くプラズマの 「降着流」にある。降着流中でプラズマは乱流状態になっており、このゆらぎのエネルギーが散 逸することでプラズマは高温に加熱される。高温になったプラズマは電磁波を放射するため、こ の放射を観測することで間接的にブラックホールの情報を得ることができる(図 1)。Event Horizon Telescope (EHT)は、銀河中心にある巨大ブラックホールの降着流からの放射を捉える 国際プロジェクトであり、ドーナツ状のブラックホールの影を捉えることに成功した[1,2]。し かし、観測結果を物理的に解釈するための理論モデルの構築がまだ不完全であるため、実際にブ ラックホール周辺でどのようなプラズマが存在しているのか確定していない。そこで本研究では、 かつてない超高解像度の電磁流体力学シミュレーションを行い、降着流中のプラズマ乱流の物理 的性質を明らかにする。これにより、EHT が観測したブラックホールの影をより高精度に解釈す ることが可能になると期待される。

降着流におけるプラズマ乱流の性質の中でも特に重要となるのが、「乱流中のAlfvén的揺動と 圧縮的揺動の比」である。Kawazura et al., 2020ではジャイロ運動論と呼ばれるモデルの直接 数値シミュレーションを行い、微小スケールにおけるイオンと電子の加熱比が慣性領域における Alfvén的揺動と圧縮的揺動の比の増加関数になることを明らかにした[3]。降着円盤からの電磁 放射は電子温度によって決定されるので、イオンと電子の加熱配分を担っている Alfvén 的揺動 と圧縮的揺動の比を求めることは非常に重要である。



図1. 本課題の対象の概念図。降着流中の乱流によってプラズマが高温に加熱され電磁波を放射 する。こうして生成された電磁波によってブラックホールの間接的な観測が可能となる。本課題 では、加熱の源となる乱流の特性を明らかにする。

2. 二次元分割擬スペクトル法コード CALLIOPE の開発

上記の通り乱流中の Alfvén 的揺動と圧縮的揺動の比が重要であるが、これまでの計算でこの 値が求められたことはなかった。Alfvén 的揺動と圧縮的揺動の比を求めるためには乱流の慣性 領域に到達する必要があるが、降着円盤の乱流を駆動している磁気回転不安定性(MRI)は注入 領域がフーリエ空間でブロードバンドであるため、従来の数値解像度では慣性領域に到達するこ とが出来ていなかった[4]。そこで本研究では、擬スペクトル法を用いて精度を上げることで実 行的な解像度の改善を行った。擬スペクトル法では各ステップでFFTと逆FFTを行う必要がある ため、並列分割の仕方が鍵となる。これまで降着円盤のシミュレーションで広く用いられてきた 擬スペクトル法コード SNOOPY[5]では1次元分割が用いられてきた。そのため、N³のグリッドに 対してNプロセスしか使うことが出来ず、計算規模に制約がかかっていた。そこで本研究では二 次元分割のライブラリ P3DFFT[6]を用いることで N³のグリッドに対して N²フロセスまで使える コード CALLIOPE[7]を開発した(図 2)。



図 2. CALLIOPE で用いている二次元分割。左は実空間グリッド、右はフーリエ空間グリッドに対応。

3. 結果

まず、CALLIOPE の並列性能に関して示す[7]。図3はOakforest-PACS及び富岳において測定したストロングスケーリングである。グリッド数は2048³に固定した。Oakforest-PACSではスレッド数4のときが最も性能が高かった。この図からCalliopeの高い並列性能が見て取れる。

次に、CALLIOPEを用いて行った MRI 乱流の結果を示す[8]。図4は時間発展を示している。特に 下パネルから、 Alfvén 的揺動と圧縮的揺動の比は1対2であることが分かる。この比率はプラ ズマのパラメータやシミュレーションのグリッド数を変えても不変であることを確認した。次に 図5は乱流のエネルギースペクトルを示している。最も興味深いのは圧縮的揺動が波数の-3/2 乗のべき乗則になっていることである。これまで多くの MRI シミュレーションで-3/2 乗に近い スペクトルが見られていたが、ここまできれいにスペクトルの傾きが見えた例はない。さらにこ のスペクトルは圧縮的揺動が支配的であることも見て取れる。この性質はプラズマのパラメータ βによらないことも分かる。



図 3. Oakforest-PACS 及び富岳における CALLIOPE のストロングスケーリング。



図 4. MRI 乱流シミュレーションの時間発展。(上) Alfvén 的揺動(u_1 、 δB_1)と圧縮的揺動($u_{||}$ 、 $\delta B_{||}$)のエネルギー、(中央) Alfvén 的揺動と圧縮的揺動のエネルギー注入率(I)と散逸率(D)、 (下) 注入率と散逸率の Alfvén 的揺動と圧縮的揺動の比。



図 5. エネルギースペクトル。左から右に β=0.1,1,10 のシミュレーションに対応する。

4. まとめ

本研究では、二次元分割擬スペクトル法コード CALLIOPE を開発し、高い並列性能を発揮できる ことを確認した。次に CALLIOPE を用いて MRI 乱流のシミュレーションを行い、Alfvén 的揺動と圧 縮的揺動の比が 2 対1 になること、パワースペクトルが波数の-3/2 乗になることを明らかにし た。以上の結果は論文[7]及び[8]で出版済みである。現在、今後はさらなる高解像度シミュレー ションを行うとともに、様々な乱流解析を行う予定である。

参考文献

- EHT Collaboration 2019 First M87 Event Horizon Telescope results. I. The shadow of the supermassive black hole. Astrophys. J. Lett. 875, L1.
- [2] EHT Collaboration 2022 First Sagittarius A* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way. Astrophys. J. Lett. 930, L12.
- [3] Kawazura, Y., Schekochihin, A. A., Barnes, M., TenBarge, J. M., Tong, Y., Klein, K. G. & Dorland, W. 2020 Ion versus electron heating in compressively driven astrophysical gyrokinetic turbulence. Phys. Rev. X 10, 041050.
- [4] Lesur, G. & Longaretti, P. Y. 2011 Non-linear energy transfers in accretion discs MRI turbulence. I. Net vertical field case. Astron. Astrophys. 528, A17.
- [5] https://ipag.osug.fr/~lesurg/snoopy.html
- [6] https://p3dfft.readthedocs.io
- [7] Kawazura, Y. 2022 Calliope: Pseudospectral shearing magnetohydrodynamics code with a pencil decomposition. Astrophy. J. 928 113.
- [8] Kawazura, Y., Schekochihin, A. A., Barnes, M., Dorland, W., and Balbus, S. A. 2022 Energy partition between Alfvénic and compressive fluctuations in magnetorotational turbulence with near-azimuthal mean magnetic field. J. Plasma Phys. 88, 905880311.

階層型直交格子と再帰的なフィッティングによる回転翼の

非定常空力予測

菅谷圭祐

東京大学大学院 工学系研究科 航空宇宙工学専攻

1. 序論

近年,複数の電動ファンを有する電動・ハイブリッド航空機や,回転翼を有し垂直離着陸が可 能な Urban Air Mobility といった,革新的な航空機が注目されている.従来の航空機と比べ,革 新的な航空機周りの空気の流れは,電動ファンなどの回転する構成部品の影響を受けるため,複 雑である.そのため数値流体力学 (Computational Fluid Dynamics, CFD)を用い,回転する構成部 品周りの非定常な流れを精度よく予測することは,革新的な航空機の空力設計において重要であ る.一方,CFDを空力設計により積極的に利用するためには,機体や部品周りの空間を離散化し た計算格子を,自動で生成できることが望ましい.そこで,複雑な形状に対し自動・ロバストに 計算格子を生成可能な直交格子法が注目されている.筆者らの研究グループでは,直交格子法に 基づく格子生成手法として,Recursive Fitting法 (RF法)を提案している^{1,2}.このRF法には, 壁面近傍の格子の情報を再帰的に修正することで,物体形状を近似的に再現した格子を自動で生 成できるという特徴がある.このRF法を,物体とともに計算格子を移動・回転させる移動格子 法³と組み合わせることで,回転する物体周りの非定常な流れを予測できると考えられる.

本研究の目的は、回転翼の非定常乱流解析に向けた、自動格子生成が可能な新しい手法を開発 することである.本研究で利用するプログラムは、東京大学 李家今村研究室で開発中の階層型 直交格子を用いる圧縮性流体ソルバー UTCart をベースに開発する⁴.計算格子の生成には RF 法 を用い、移動格子法と組み合わせることで物体の移動を再現する.加えて壁面近傍の乱流境界層 を計算するため、壁関数を用いて壁面摩擦応力を予測する.計算対象は、Caradonna and Tung に より実験がなされた回転翼である⁵. コレクティブピッチ角(回転翼のブレードと気流のなす角 度)が一定の条件で、回転翼の角速度を変更した二通の解析を行い、Caradonna and Tung の実験 と比較することで、開発する手法の予測精度を評価する.

2. 計算手法

2.1.階層型直交格子の生成

図1は物体の周りに生成された階段状の直交格子の例である.格子のセルで流体領域に存在す るものを流体セル(fluid cell)と呼び,物体壁面(body surface)と交差するセルを壁面セル(wall cell)と呼ぶ.また流体セルと壁面セルの間のセル境界を壁面境界(wall face)と呼び,壁面境界 を有する流体セルを壁面隣接セルと呼ぶ.

2. 2. Recursive Fitting 法(RF法)

図1のように、単純に物体の周りに生成された直交格子の壁面境界は、物体壁面に沿わない.

そこで RF 法では、物体に向けて壁面隣接セルを拡大し、物体形状を近似的に再現した格子を自動で生成する.図2に RF 法で生成される格子の例を示す.格子の拡大には、変形前の直交格子の壁面境界を構成する頂点に対する、物体壁面の再近傍点を利用する.加えて、RF 法では壁面境界を構成する追加の点を計算し、セルの体積やセル境界の面積ベクトルといった、流体計算に必要な情報を修正する.追加点を再帰的に導入することで、例えば航空機の翼胴結合部といった、特徴線と呼ばれる形状が不連続に変化する部分でも、物体壁面を近似的に再現した格子を自動で生成できる.RF 法の詳細は文献^{1,2}も参照されたい.

2. 3. 支配方程式

移動格子法を用いる場合の非定常 Navier-Stokes 方程式は次式である.

$$\frac{\partial}{\partial t} \iiint_{V} \boldsymbol{Q}_{c} dV + \iint_{S} \{ (\boldsymbol{F}(\boldsymbol{Q}_{c}, \boldsymbol{n}) - \boldsymbol{Q}_{c} W_{n}) - \boldsymbol{F}_{v} \} dS = 0,$$
(1)

式(1)の Q_c は保存変数である. V,S はそれぞれ検査体積と検査体積の境界である. またベクト ル n は検査体積に対し外向きの単位ベクトルである. F,F_v はそれぞれ非粘性流束と粘性流束 である. W_n は格子の速度ベクトルと単位法線ベクトル n の内積である. 式(1)には,検査体積 の移動に伴う流束 Q_cW_n が含まれている.本研究では文献⁶を参考に,非粘性流束と格子の移動 に伴う流束を次式で計算する.

$$F(\boldsymbol{Q}_{c},\boldsymbol{n}) - \boldsymbol{Q}_{c}W_{n} = \mathcal{T}_{c}^{-1}\boldsymbol{F}(\mathcal{T}_{c}\boldsymbol{Q}_{c},\boldsymbol{n}), \qquad (2)$$

$$\mathcal{T}_{c} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -W_{x} & 1 & 0 & 0 & 0 \\ -W_{y} & 0 & 1 & 0 & 0 \\ -W_{z} & 0 & 0 & 1 & 0 \\ \frac{|W|^{2}}{2} & -W_{x} & -W_{y} & -W_{z} & 1 \end{bmatrix}.$$
(3)

式(2), (3)では,格子に対する気流の相対速度から計算される保存変数である $T_c Q_c$ を計算した後, 逆行列 T_c^{-1} と流束の積を計算することで,保存変数 Q_c に対する流束を評価している.

本研究では、式(1)-(3)をセル中心有限体積法で離散化する.非粘性流束と粘性流束はそれぞれ SLAU⁷と中心差分で評価する. 翼端マッハ数が遷音速の場合は、Hishida のリミタ⁸を利用する. 時間積分には MFGS 陰解法⁷と Dual Time Stepping による二次精度後退差分を利用する. 乱流モ デルは、Spalart–Allmaras (SA) モデル⁹をベースにした DDES-p⁴ である. さらに、壁面での摩擦 応力は壁関数を用いて評価する.

2. 4. 壁関数の実装

本研究では,壁面境界条件は非滑り・断熱壁である. 直交格子を用いて物体周りの乱流を計算 するために,壁関数を用いて壁面摩擦速度 u_r と壁面摩擦応力 $\tau_{wall} = \rho_{wall} u_r^2$ を予測する. Newton–Raphson 法を用い,以下の式を用い反復計算で摩擦速度を決定する.

$$u_{\tau}^{m+1} = u_{\tau}^{m} - \frac{g(u_{\tau}^{m})}{g'(u_{\tau}^{m})}, \qquad (4)$$

$$g(u_{\tau}) = u_{\tau} f_{wall}(y^{+}) - (u_{t,1} - W_t), g'(u_{\tau}) = f_{wall}(y^{+}) + u_{\tau} \left\{ \frac{\partial f_{wall}}{\partial u_{\tau}}(y^{+}) \right\},$$
(5)

$$\frac{\partial f_{wall}}{\partial u_{\tau}} = \frac{\partial y^{+}}{\partial u_{\tau}} \frac{\partial f_{wall}}{\partial y^{+}} = \frac{y^{+}}{u_{\tau}} \frac{\partial f_{wall}}{\partial y^{+}} , \qquad (6)$$

式(5)の $u_{t,1}$ は,壁面隣接セルのセル中心値における,壁面に対し接線方向の速度である. f_{wall} は SA 壁関数⁹である. また $y^+ = y_1 u_\tau / v_1$ は無次元壁面距離である. W_t は壁面の速度ベクトル のうち,壁面の接線方向の成分であり,式(5)で利用される. さらに, Crocco–Busemannの関係式 ¹⁰から壁面での温度を計算する.

図3は,壁関数の予測に利用される点の比較である.直交格子と壁関数を組み合わせた解析の 先行研究では、Image Point (IP) が導入されている(図3a)^{4,9}. IP は壁面から垂直な方向に *d*_{1P} 離 れた位置に配置される.また *d*_{1P} はユーザー指定のパラメータである. IP を導入することで摩 擦速度の予測に利用される点の壁からの距離を一定にすることができるが、計算結果が *d*_{1P} に 依存する場合がある.計算結果の *d*_{1P} への依存性を無くすため、本研究では IP を導入せず、そ のかわり壁面隣接セルのセル中心の値を利用する(図3b).先行研究の RF 法と壁関数を組み合わ せた静止する物体周りの解析では、壁面隣接セルを用いて壁面摩擦を予想した場合でも乱流境界 が再現されることが報告されている¹.





図 2: RF 法で生成される格子の例



図3:壁関数による摩擦速度の予測に利用される点の比較

3. 検証計算

3. 1. 問題設定

本研究では、二枚のブレードから構成される Caradonna-Tung 回転翼⁵周りの流れを計算する. ブレードの形状は矩形翼で、翼型は NACA0012 である. またブレードに捩りとテーパーは無い. コレクティブピッチ角と角速度を変更した実験がなされており、本研究ではコレクティブピッチ 角が 8 [deg]、翼端でのマッハ数 M_{tip} が 0.439 と 0.890 の条件を計算する. 翼弦長 C = 7.5 [in.]と 音速基準のレイノルズ数は 4.44 × 10⁶ である.

3.2.計算格子の設定

図4に計算格子の設定を示す.ブレードに対する気流の速度は,翼根から翼端に向けて増加す る.そのため,翼端側で,気流の相対速度から計算されるレイノルズ数が増加する.そこでブレ ード中央での格子幅を基準に,翼端側では格子を細分化し,翼根側ではより大きな格子幅とする. 翼端側の格子幅は,翼弦長の1/500である.また非定常空力を精度良く予測するためには,回転 翼後流の非定常な流れを精度良く計算することが重要である.そこで,後流を局所的に等間隔に 細分化する.格子のセル数は76,141,294である.

3.3.計算の手順

本研究では,非定常空力を予測する.このため回転翼後流の流れを精度良く予測することが重要である.一方計算の後処理の観点から,異なる翼端マッハ数の計算でも同じ計算ステップ数で解析することが望ましい.そこで,無次元時間刻み幅を,*M_{tip}* = 0.439 では 0.179, *M_{tip}* = 0.890 では 0.0882 とする. 翼端マッハ数と音速から計算されるクーラン数は,後流でおおよそ 1 である.さらに,どちらの翼端マッハ数の設定でも 3,600 計算ステップで回転翼が 1 回転する. 翼端マッハ数が 0.890 の場合は翼端で衝撃波が発生する.そこで,非粘性流束評価にリミタを利用する.

図 5 は推力係数 (C_T)と Figure of Merit (*FM*)の履歴である.計算は以下の手順で行う.はじめ に内部反復を2回とし、回転翼を10回転させる(図 5a (1)).この計算の目的は、計算の初期段 階に生じる擾乱をブレード下流に移流させることである.次に、内部反復を10回に変更し、回 転翼を12回転させる(図 5a (2), (3)).10回の内部反復により、密度の内部残差は一桁減少する. 時間平均された空力係数と表面圧力係数分布は、最終の4回転を平均して計算する(図 5a (3)).

3. 4. 計算資源

本研究には東京大学情報基盤センターが運営する Oakbridge-CX スーパーコンピュータシステムを用いる.計算格子の分割には METIS ライブラリを用いる. Message Passing Interface (MPI) により,896 並列で計算する.計算に要する時間 (wall clock time) は,翼端マッハ数が 0.439 の場合は88 時間,翼端マッハ数が 0.890 の場合は96 時間である. 翼端マッハ数が 0.890 の場合は 非粘性流束評価にリミタを利用するため,計算時間が増加する.

3.5.計算結果

図5は推力係数とFigure of Meritの履歴である.履歴には振動が観察されるが、最終4回転は 準定常状態である.表1は C_T とFMの時間平均値である.翼端マッハ数が増加することで、 C_T は 増加し、FMは減少する. C_T が増加する傾向は実験でも観察されている.またブレード端で生じる衝撃波が原因で、FMは減少していると考えられる.図 6,7 はブレードの 50%、80%、96% 位置での、表面圧力分布の時間平均値である.どちらの翼端マッハ数の結果も、実験値を定性的・定量的に再現できている.図8はQ値の等値面である.ブレード端から下流に向けて、らせん状の渦が生じている.以上より、本研究の計算は実験を良好に再現できている.

4. 結論

本研究の目的は、回転翼の非定常乱流解析に向けた、自動格子生成が可能な新しい手法を開発 することであった.計算格子の生成には、物体形状を近似的に再現した格子を自動で生成できる RF 法を用いた. RF 法を移動格子法と組み合わせることで、回転翼の回転といった物体の移動を 再現した.また壁面近傍の乱流境界層を計算するため、壁関数を用いて壁面摩擦応力を予測した. 検証計算として、Caradonna and Tung により実験がなされた回転翼周りの流れを解析した.計算 には東京大学情報基盤センターが運営する Oakbridge-CX スーパーコンピュータシステムを用い た.1ケース当たりの計算時間(wall clock time)はおおよそ 90~100時間であった.推力係数の 翼端マッハ数への依存性は、Caradonna and Tung の実験と定性的に一致した.また表面圧力係数 分布は、実験と定量的に一致した.

謝辞

本研究は東京大学 今村太郎 准教授の指導の下行われた.また本研究は東京大学情報基盤セン ター「若手・女性利用者推薦」2021 年度インターン課題として行われた.計算結果の可視化に は, Intelligent Light University Partner Program (UPP) の提供を受け, FieldView を用いた.ここに 感謝の意を表す.

表1:空力係数の時間平均値

翼端マッハ数	推力係数	Figure of Merit
0.439	0.00516	0.543
0.890	0.00577	0.457



図4:格子の設定



図 7:表面圧力係数分布 $(M_{tip} = 0.439)$



(a) $M_{tip} = 0.439$



(b) M_{tip} = 0.890図 8:Q 値の等値面(軸方向速度で色付け)

 Sugaya, K., and Imamura, T., "Aerodynamic Analysis of Common Research Model at Low Speed Conditions Using Recursive Fitting Method and Wall Function," AIAA Scitech 2022 Forum, AIAA Paper 2022-0448, January 2022.

https://doi.org/10.2514/6.2022-0448

- Sugaya, K., and Imamura, T., "A Recursive Fitting Approach for Automatic Grid Generation Around Aircraft Configuration," Journal of the Japan Society for Aeronautical and Space Sciences, Vol. 68, No. 3, 2020, pp. 115–122 (in Japanese). https://doi.org/10.2322/jisass.68.115
- Yu, M. and Wang, L., "A high-order flux reconstruction/correction procedure via reconstruction formulation for unsteady incompressible flow on unstructured moving grids," Computers & Fluids, Vol. 139, 2016, pp. 161–173.

https://doi.org/10.1016/j.compfluid.2016.05.028

- Imamura, T., and Tamaki, Y., "Unsteady Flow Simulation Around Two-Wheel Main Landing Gear Based on Compressible Navier-Stokes Solver with Immersed Boundary Method," AIAA AVIATION 2020 FORUM, AIAA Paper 2020-2579, June 2020. https://doi.org/10.2514/6.2020-2579
- Caradonna, F. X. and Tung, C., "Experimental and analytical studies of a model helicopter rotor in hover," Conference Paper NASA-TM-81232, NASA, September 1981. https://ntrs.nasa.gov/search.jsp?R=19820004169
- Nishikawa, H., "Roe Flux in Deforming Mesh," Web-published note. http://doi.org/10.13140/RG.2.2.20266.57288
- Shima, E. and Kitamura, K., "Parameter-Free Simple Low-Dissipation AUSM-Family Scheme for All Speeds," AIAA Journal, Vol. 49, No. 8, 2011, pp. 1693–1709. http://doi.org/10.2514/1.J050905
- Hishida, M., Hashimoto, A., Murakami, K., and Aoyama, T., "A new slope limiter for fast unstructured CFD solver FaSTAR," JAXA Special Publication: Proceedings of 42nd Fluid Dynamics Conference / Aerospace Numerical Simulation Symposium 2010, JAXA, February 2011, pp. 85–90, JAXA-SP-10-012 (in Japanese).

http://id.nii.ac.jp/1696/00004716/

- Berger, M. and Aftosmis, M. J., "Progress Towards a Cartesian Cut-Cell Method for Viscous Compressible Flow," 50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, Aerospace Sciences Meetings, 2011, AIAA Paper 2012-1303, January 2013. https://doi.org/10.2514/6.2012-1301
- 10. White, F. M., Viscous Fluid Flow, McGraw Hill, New York, 1974, pp. 578-581, 627-629.

ダイマー粒子を用いたガラスの Johari-Goldstein beta 緩和の 理解

白石薫平

東京大学 大学院総合文化研究科 広域科学専攻

1. はじめに

本稿では、東京大学情報基盤センター「若手・女性利用」2021 年度(後期)の採択課題「ダイ マー粒子を用いたガラスの Johari-Goldstein beta 緩和の理解」の研究成果を報告する。

2. イントロダクション

液体を冷やすと凍ることは、我々が日常的に目にしている現象である。微視的には、液体状態 では乱雑な配置を取っていた構成分子が、融点以下になると規則正しく整列して結晶になること に相当する。しかし、液体を急激に冷却すると、構成分子が乱雑な配置のまま固まりガラスが生 成される。これを**ガラス転移**という[1, 2, 3]。熱力学的な相転移とは大きく異なり、ガラス転 移近傍では液体の微視的構造はほとんど変化しないのに、ダイナミクスの急激なスローダウンが 生じる。このように、物理学におけるガラス転移は数多くの未解決問題を残す現象であり、ガラ ス転移は熱力学的な相転移なのかという基本的な論点でさえ明確な答えは得られていない。

ガラス転移にまつわる論点は多岐に亘るが、ここでは過冷却液体中の緩和に着目する。液体を 冷却してガラスを生成する際、液体は融点での結晶化を避けて過冷却液体という状態に入る。こ の過冷却液体状態では、微視的な構成分子は熱揺らぎによって運動しており、十分時間が経過す ると、微視的な配置同士の相関が失われる。これが緩和である。低温では緩和時間が発散するこ とが知られ、このような過冷却液体中における緩和時間の増大の起源を理解することがガラス転 移研究の大問題となっている。実験研究では、分子性液体の過冷却液体状態において、系の緩和 が幅広い時間スケールで進むことが実証されている。その緩和中には主要な2つの過程が含まれ ており、短時間で進む過程を Johari-Goldstein beta 緩和、長時間で起こる過程を alpha 緩和と 呼ぶ[4,5]。これまでのガラス転移のシミュレーション研究は主に後者の解明に力点が置かれて きた。それらの研究では、球や円盤などの等方粒子の分子動力学シミュレーションによって各粒 子の運動を観察する[6]。その結果、alpha 緩和の前には、周囲の粒子が形作るケージ構造にトラ ップされることで自由な運動が阻害され、ランダムな熱運動によってケージを抜け出す運動が alpha 緩和に相当するという理解が確立された(第1図)。微視的な運動の情報が全て手に入る 分子動力学シミュレーションによって得られたケージ構造という理解は、コロイド分散系の実験 でも観測されている[7]。しかしながら不満は残る。前述の通り分子性液体の実験で普遍的に観 測される Johari-Goldstein beta 緩和が、等方粒子系のシミュレーションには欠落しているので ある。



第1図:ケージトラップとケージ脱出の模式図。

左図:低温で粒子は、周囲の粒子が形成するケージにトラップされ、ケージ内で熱揺らぎによる振動運動をしている。右図:長時間が経過すると熱揺らぎによってケージを脱出する。

Johari-Goldstein beta 緩和が多くの分子性液体で普遍的に観測されている事実は、過冷却液体の緩和において「ケージ構造へのトラップ・脱出」という過程以外の実空間的な運動様態が、 普遍的に存在していることを意味している。この普遍的な緩和過程の実空間描像を確立すること が本研究の目的である。

3. シミュレーション

本研究では、Fragiadakis と Roland によって提案された非対称ダイマーモデルの分子動力学 シミュレーションを行った[8]。1 つのダイマー分子は 2 つの原子から構成されており、両者の 距離は 0.5 に固定されている。原子同士は Lennard-Jones ポテンシャルで相互作用する:

$$V(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right].$$

パラメータ ϵ_{ij} 、 σ_{ij} は同一分子に所属する原子同士で異なっており、とくにパラメータ σ_{ij} で非対称 性をコントロールする。このような非対称ダイマー分子を80:20の割合で混合した2成分系を考 える。

まず、能勢フーバー熱浴による粒子数 N、体積 V、温度 T が一定の分子動力学シミュレーショ ンを行い、平衡液体配置を生成した[9]。低温では緩和時間が増大し通常の分子動力学シミュレ ーションによる平衡化と位相空間中の適切なサンプリングが不可能になるので、レプリカ交換法 を実装して平衡配置のサンプリングを行った[10, 11]。そして得られた平衡配置から粒子数 N、 体積 V、エネルギーE が一定の分子動力学シミュレーションを改めて行い、得られた軌道の微視 的な動力学を解析した。時間発展の際に原子同士の距離を固定するため RATTLE 法を用いた[9]。

4. ダイマー系の緩和過程

まず、ダイマー重心の緩和を明らかにするため、平均二乗変位 (MSD)

$$|\Delta r(t)|^{2} = \frac{1}{N} \sum_{i} |r_{i}(t) - r_{i}(0)|^{2}$$

を計算する。但し、 $r_i(t)$ は時刻tにおけるダイマーiの重心の位置座標を表す。高温T = 2.25から低温T = 0.403までの各温度の軌道に対して MSD を計算した結果を第2図に示した。等方粒子系同様[6]、短時間では弾道過程 $|\Delta r(t)|^2 \propto t^2$ が、長時間では拡散過程 $|\Delta r(t)|^2 \propto t$ が見られる。低温では中間部にプラトーが出現する。等方粒子系の MSD に現れるプラトーが1段階なのとは一転して、ダイマー系のプラトーは2段階のプラトーを示す。プラトー間に現れる MSD の増加が、Johari-Goldstein beta 緩和に相当する。

さらに、ダイマーの回転運動に対する特徴付けとして、下記の量を導入する:

$$\theta(t) = \frac{1}{N} \sum_{i} acos[\hat{d}_{i}(t) \cdot \hat{d}_{i}(0)].$$

但し、*â*_{*i*}(*t*)は時刻*t*におけるダイマー*i*の単位方向ベクトルを表す。これは、ダイマーの方向ベクトルが初期配置からどれだけ回転したかを定量化する、回転運動の MSD に相当する量である。第2回に結果を示した。並進の MSD と同じく、中間部に2段階のプラトーが現れることが分かる。ここで、Johari-Goldstein beta 緩和の後のプラトーの高さに着目する(第2図中に点線で表

した)。並進の MSD が表す重心の変位は、Johari-Goldstein beta 緩和後に 0.1 に満たない程度 であるのに対して、回転の MSD が表す方向ベクトルの回転はπ/4以上と大きいことが分かる。す なわち、Johari-Goldstein beta 緩和においては、並進位置の変化に比べて強い回転運動が生じ ている。よって、Johari-Goldstein beta 緩和は、周囲の分子が形成するケージ内の回転運動で あると結論できる。





左図の平均二乗変位では、高温では短時間の弾道過程から長時間の拡散過程へすぐに移行している。低温では 中間部に2段階のプラトーが現れ、プラトー間の緩和が Johari-Goldstein beta 緩和である。右図の方向ベクト ルの回転でも同様の3段階緩和が見られる。Johari-Goldstein beta 緩和後のプラトーの高さは、並進の平均二 乗変位が粒子直径よりも短い0.1程度であるのに対して、方向ベクトルの回転は $\pi/4$ 程度で大きいという顕著な 違いが存在する。両者とも、T = 2.25(赤)からT = 0.403(紫)までの結果を示した。

5. Johar i-Goldstein beta 緩和の実空間描像

ここでは、前節までで明らかになったダイマーの緩和の実空間的な理解を、1分子レベルの微 視的運動として理解することを目指す。そこで、分子同士の接続に分類を導入する。まず、ダイ マーという二原子分子間の接続を考えるため、原子間接続の本数での分類を考える。ただし、原 子間接続は2つの原子間の距離が閾値より小さいかで判定する。扱っているモデルから明らかな ように、1本の分子間接続は原子間接続が4本から1本までの4通りに分けることができる。さ らに、このダイマー分子は大原子と小原子から成る非対称性を持つため、1本の原子間接続が大 原子・大原子、大原子・小原子、小原子・小原子の3通りが存在する。以上を勘案すると、1本 の分子間接続には下表の通り12通りの状態が存在することが分かる。

原子間接続の本数と原子間接続の種類に着目することで、1本の分子間接続を12通りに場合分けできる。Lは 大原子、Sは小原子を意味する。

状態	内訳			
4	LL, LS, LS, SS			
3A	LL, LS, SS			
3B	LL, LS, LS			
3C	LS, LS, SS			
2A	LL, LS			
2B	LS, SS			
2C	LS, LS			
2D	LL, SS			
1A	LL			
1B	LS			
1C	SS			
0	_			

ここで、分子動力学シミュレーションで生成した軌道中で、それぞれ分子対がどのような状態遷移を経て解離していくかを調べた。分子間接続の状態遷移図で表現すると、分子間接続の実に99%が下図に示す状態遷移に従うことが分かった。この遷移を図示すると、第3図になる。



第3図:ダイマーの分子間接続状態の遷移図。

ダイマーの分子対の99%以上が、図中に赤矢印で示した状態遷移に従って解離していく。

第1表:ダイマーの分子間接続の分類と構成の内訳。

ほぼ全ての分子間接続が従うこの状態遷移の各状態を第4図に図示した。この図中の「4」から「1A」までの各状態を、Johari-Goldstein beta緩和のタイムスケールでは行きつ戻りつしながら、全体の傾向としては1Aに向かう方向に遷移していく。この結果より、Johari-Goldstein beta緩和は分子の回転に伴う原子ボンドの組み換え運動に相当し、並進位置の変化をあまり含まないことが分かる。他方、alpha緩和は状態1Aから状態0への変化を起こしていることが分かった。この結果より、alpha緩和は分子ボンド切断を伴う並進位置変化に相当することが明らかになった。



第4図:ダイマーの解離過程の実空間描像。

Johari-Goldstein beta 緩和の時間スケールで分子対は、図中の4と示した状態から1Aと示した状態までを 行きつ戻りつしながら、1Aの方向へ遷移していく。この原子間接続の繋ぎ替えを推進するのは、第2図で議論し た分子の回転運動である。alpha 緩和の時間スケールでは、最後に残った大原子間の接続(第1表でのLLに相 当)が並進運動により切断され、分子間接続が解離する。各図左上に示した状態は、第1表で導入した状態分類 に対応している。

6. まとめ

本研究では、分子性過冷却液体に現れる Johari-Goldstein beta 緩和という緩和過程を、二原 子分子モデルである非対称ダイマーの分子動力学シミュレーションによって調べた。このモデル には Johari-Goldstein beta 緩和が現れ、その運動は分子の並進位置をあまり変えないケージ内 回転運動であった。一分子レベルの運動に着目した詳細な解析により、Johari-Goldstein beta 緩和は原子間接続の繋ぎ替えに、alpha 緩和は分子間接続の切断に相当することが明らかになっ た。

参考文献

[1] L. Berthier and G. Biroli, Theoretical perspective on the glass transition and amorphous materials, Reviews of Modern Physics 83, 587--645 (2011)

[2] P. G. Debenedetti and F. H. Stillinger, Supercooled liquids and the glass transition, Nature 410, 259--267 (2001)

[3] F. H. Stillinger, A Topographic View of Supercooled Liquids and Glass Formation, Science 267, 1935--1939 (1995) [4] G. P. Johari and M. Goldstein, Viscous Liquids and the Glass Transition. II. Secondary Relaxations in Glasses of Rigid Molecules, The Journal of Chemical Physics 53, 2372--2388 (1970)

[5] K. L. Ngai and M. Paluch, Classification of secondary relaxation in glass-formers based on dynamic properties, The Journal of Chemical Physics 120, 857-873 (2004)

[6] W. Kob and H. Andersen, Testing mode-coupling theory for a supercooled binary Lennard-Jones mixture I: The van Hove correlation function, Physical Review E 51, 4626 (1995)

[7] E. R. Weeks and D. A. Weitz, Properties of Cage Rearrangements Observed near the Colloidal Glass Transition, Physical Review Letters 89, 095704 (2002)

[8] D. Fragiadakis and C. M. Roland, Molecular dynamics simulation of the Johari-Goldstein relaxation in a molecular liquid, Physical Review E 86, 020501 (2012)

[9] M. P. Allen and D. J. Tildesley, Computer Simulation of Liquids, Oxford University Press (2017)

[10] K. Hukushima and K. Nemoto, Exchange Monte Carlo Method and Application to Spin Glass Simulations, Journal of the Physical Society of Japan 65, 1604-1608 (1996)
[11] R. Yamamoto and W. Kob, Replica-exchange molecular dynamics simulation for supercooled liquids, Physical Review E 61, 5473 (2000)

公共データを活用した転写因子結合ダイナミクスの解析

植野和子

国立国際医療研究センター ゲノム医科学プロジェクト 戸山プロジェクト

1. はじめに

DNA に保存された生命の遺伝情報は、複製により子孫に受け継がれるのと同時に RNA への「転 写」・タンパク質への「翻訳」の過程を経て表現型として表出する。一連の遺伝子発現メカニズ ムの制御において最も重要なステップであると考えられるのが、ゲノム DNA から RNA ポリメラー ゼ II (Po12) によりメッセンジャーRNA (mRNA) が写し取られる「転写」である。転写因子は DNA の特異的な塩基配列を認識し結合するタンパク質の一群であり、遺伝情報を RNA に転写する過程 を促進または抑制することで、種々の細胞種毎の遺伝子発現状態をオーガナイズしている。ヒト ゲノム上には、およそ 1,800 前後の転写因子がコードされており、それらの働きを網羅的に解明 することが、究極的には生命の理解につながると考えられている。

一方近年では、転写因子の人為的操作によって細胞運命を工学的にリプログラミングすること が可能になりつつある。終末分化した体細胞を初期化し、あらゆる細胞への分化能を有する誘導 多能性幹細胞(iPS 細胞)を誘導する技術がもっとも有名だが¹、最近では様々な転写因子の組 み合わせを導入することで、多種多様な細胞を直接誘導するダイレクトリプログラミング技術も 盛んに開発が進められており²、いずれも再生医療に向けた応用が期待されている。今後は、そ れらのリプログラミング技術をより安全・安定的なものに改良し、臨床応用可能なレベルにまで 発展させるために、導入転写因子セットが、ゲノム DNA の「どこ」に結合し「何の」遺伝子を「ど のように」制御するのかを明らかにすることが必要とされている。

2. 転写因子 Foxa3 は「動く」

我々の研究グループでは、マウスの皮膚や胎仔由来の線維芽細胞にHnf4αおよびFoxa(ファミ リータンパク質であるFoxa1、Foxa2、Foxa3のいずれか一つ)を強制発現させることにより作製 される誘導肝細胞様細胞(iHep 細胞)³を研究対象として、転写因子の挙動と機能、およびリプ ログラミングにおける役割を解析してきた。特に**次世代シークエンサー(NGS)**を用いたゲノム ワイドな転写因子の結合解析を通して、2つの転写因子が互いに影響しながら細胞の転写状態を 生体の肝細胞に近い状態に誘導するメカニズムの一端を明らかにした⁴。

その解析データの中から我々は、リプログラミング誘導転写因子 Foxa ファミリーのうち Foxa3 のみが、標的遺伝子のコード領域(genebody)上を Po12 と共に移動し、なおかつその様な挙動 をとること自体が iHep 細胞へのダイレクトリプログラミングの進行にとって必須であることを 見出した(図1)。従来、転写因子は DNA上の認識配列に強固に結合し、その部位に転写の活性 化や抑制に関わる補因子をリクルートすることで転写を制御するという静的モデルが一般的に 信じられてきたが、この発見は転写因子の既存イメージを大きく覆すものである。また、このよ うな Foxa3 のダイナミックに「動く」分子挙動は、リプログラミング細胞だけでなく生体由来の 肝細胞においても確認することができることから、生体の細胞内において何らかの生理的な役割 を果たしているものと考えられる。



第1図: Foxa3 および Pol2の genebody 上での結合分布。

左図に Foxa3の、右図に活性型 Pol2の遺伝子領域上での結合分布を示す metagene plot を記載した。それぞ れの色別プロットは全遺伝子をそれらの転写量で10分割したカテゴリーを示しており、Foxa3は Pol2と同様 に転写量が減少すると genebody 上の結合が減少することを示している。このことは Foxa3が Pol2と結合を共に している可能性を強く示唆している。また、転写開始点 (TSS) に強いピーク、転写終結点 (TES) 下流に弱いピ ークが出現し、その間の genebody に裾野を引くような分布パターンとなることも、Foxa3が genebody 上を移動 するという「動く」分子モデルをサポートする結果である。

3. 転写因子の大規模公共データベースとその再解析

Foxa3と同様の分子挙動を示す転写因子の例はこれまでに報告が無く、上述したような「動く」 分子モデルが一般的なものなのか、それとも Foxa3 特有のものなのかは全くわからない。転写因 子のゲノムワイドな結合データは、論文投稿時にパプリックドメインにデポジットされることが 求められることもあり、RAW データおよびその一次加工データが公共データベースに大量に格納 されている(例: Gene Expression Omnibus⁵)。その登録数は日々増加しており、それらのデー タを再解析することで、Foxa3 と似た「動く」分子挙動を示す転写因子をスクリーニングできる のではないかと考えた。

しかし、転写因子の結合データのほとんどは、クロマチン免疫沈降シークエンシング法 (ChIPseq) という手法で解析されたものであるが、ChIP-seq のプロトコールには様々なバリエーショ ンがあり、かつ品質において玉石混合であることからも、無作為に収集したデータを機械的に解 析するだけでは意味のある結果を得るのは難しいと考え、データの品質や実験手法が統一されて いる大規模プロジェクトのデータベースから情報を収集することにした。

その意味で、現時点で最も信頼性が高く大量のデータを格納しているのが、アメリカ国立衛生 研究所の主導で進められている ENCODE (ENCyclopaedias Of DNA Elements) 計画のデータベー スである⁶。このデータベースには 4,756 のヒトおよびマウスの細胞・組織から得られた ChIPseq データが収められており (2022 年 6 月現在)、実験条件やサンプルなどの metadeta も整備さ れている。よって本プロジェクトでは ENCODE データベースから大量の ChIP-seq データを取得 し、転写因子の結合パターンのスクリーニングを行うこととした。

4. 転写因子結合の公共データの再解析

本プロジェクトでは、パイロットスタディとして 40 種類の転写因子データセットを取得し解 析を行った。転写因子の選定基準は、①実験的に「動く」転写因子であることが示されている FOXA3 (Foxa3 のヒトホモログ)、②実験により「動かない」転写因子であることが示されている FOXA1 および FOXA2 (Foxa1 および Foxa2 のヒトホモログ)、③必ず遺伝子上を動く分子である POLRA2 (Po12 のサブユニット)、④TSS 近傍への結合が示唆されている MYC/MAX ファミリー転写因子、

⑤他の FOX ファミリー分子、⑥AP-1 や CTCF など代表的な転写因子、についてそれぞれ種々の細胞をサンプルとして実験された ChIP-seq データを集めた。今回収集したデータは全てヒト由来の細胞を用いて実験したものである。

一般的に、ChIP-seq データなどの NGS データは short reads と呼ばれる 25~150 bp ほどの大 量の配列断片が格納された fastq と呼ばれるファイルを解析の起点として、マッピング、カバレ ッジの定量、ピークコールなどを続けて行う(図2)。ENCODE データベースでは、fastq に加え て一次解析後のファイル (BAM および bigwig) も配布しているため、それらをデータベースより 取得して解析を行った。



第2図:転写因子結合解析におけるNGSデータの解析フロー。

ChIP-seq 実験により得られたサンプルから配列データ (fastq)の取得、さらにそこからマッピングデータ (BAM/SAM)、カバレッジデータ (bigwig/wig)、ピークデータ (bed)を取得する流れを示した。マッピング解析 は非常に大きな計算資源とストレージを要求するため、今回は一次解析データであり、比較的ファイルサイズの 小さい bed 及び bigwig を ENCODE データベースより取得して解析した。

まず取得したピーク情報 (bed) を用いて、TSS 周辺のシグナルを確認した。TSS と重複してい るピークが少ないデータは、十分な s/n 比が得られていない (または遺伝子近傍には結合しない 転写因子) と判断し解析対象から除外した (2/40)。続いてカバレッジデータ (bigwig) を用い て TSS 周囲の領域でシグナルを 500 bp ウインドウで平均値を算出した。その分布が genebody に おいて「裾野」を持つパターンになるか、それとも TSS のピーク以外はシグナルが低い平坦なパ ターンになるかどうかを、転写因子の「動く」「動かない」の判定基準とした (図3)。具体的に は、500bp 幅のウインドウに分割した TSS 周辺の領域について各遺伝子座の ChIP-seq シグナル の平均値を算出し、1番目と2番目の分布に有意な差があるものを「動く」転写因子、差がなく フラットな低い値の分布を示すものを「動かない」転写因子と定義した。



第3図:転写因子の挙動を解析する計算方法。

転写因子の遺伝子近傍における挙動を解析した数値をウインドウ毎の boxplot で示している。「動かない」こと が既知の因子 Foxal (左図) と、動くことが明らかな Pol2 (中図) のデータを示す。いずれも TSS に強いピーク があるが、Foxal がその周囲の領域ではシグナルが低くフラットとなるパターンを示すのに対し、Pol2 ではシグ ナルが段階的に下降する裾野のようなパターンを示す。MYC は Pol2 と同じような結合パターンを示した(右図)。

解析の結果、Foxa1 および Foxa2 はこれまでの実験データの通り「動かない」転写因子である と判定された。一方 Foxa3 は、肝臓由来の細胞株による ChIP-seq データでは自分達の実験結果 と同様に Po12 と同じ「動く」パターンを示したが、血球由来の細胞株から得られたデータでは 「動かない」結合パターンとなった。このことは、同じ転写因子でも細胞種の違いなどのコンテ クストにより挙動を変化させる可能性があることを示唆している。他の転写因子の中で、MYC/MAX ファミリーの ChIP-seq データ全てが「動く」転写因子のパターンを示した。MYC/MAX は転写が 活性化している遺伝子のプロモーターに積極的に結合する「invasion」と呼ばれる現象が知られ ており⁷、これまでの解析からは明らかにされてこなかったが、Foxa3 と同様に Po12 と共に「動 く」ことで機能を発揮する転写因子である可能性が示唆された。

5. まとめ

ENCODE データベースから取得した 40 個の転写因子結合データを再解析することで、Pol2 と類 似した、遺伝子上を「動く」パターンを示す転写因子として、MYC/MAX ファミリーを同定した。 今回の結果は、我々が構築した公共データの再解析法が転写因子の遺伝子上の動きのパターンを 判別するうえで有効であることを示唆しており、今後は ENCODE データベースに格納される全 ChIP-seq データ等を用いて、より大規模な解析を実施したい。

参考文献

- [1] K. Takahashi and S. Yamanaka, Cell, 126, pp. 663-676, (2006).
- [2] K. Horisawa and A. Suzuki, Jpn Acad Ser B Phys Biol Sci., 96, pp. 131-158, (2020).
- [3] S. Sekiya and A. Suzuki, Nature, 475, pp. 390-393, (2011).
- [4] K. Horisawa, M. Udono, K. Ueno, Y. Ohkawa, M. Nagasaki, S. Sekiya, A. Suzuki, Mol Cell, 79, pp. 660-676, (2020).
- [5] https://www.ncbi.nlm.nih.gov/geo/
- [6] https://www.encodeproject.org/
- [7] A. Sabò, T.R. Kress, M. Pelizzola, S. Pretis, M.M. Gorski, A. Tesi, M.J. Morelli, P. Bora, M. Doni, A. Verrecchia, C. Tonelli, G. Fagà, V. Bianchi, A. Ronchi, D. Low, H. Müller, E. Guccione, S. Campaner, B. Amati, Nature, 511, pp. 488-492, (2014).

ディープラーニングによる高精度マルウェア分析

三橋力麻

東京大学大学院情報理工学系研究科 システム情報学専攻

1. はじめに

サイバーセキュリティの現場では、セキュリティアナリストとマルウェアの戦いが続いている。 マルウェアの構造や動作は従来よりも複雑さを増しているため、セキュリティアナリストは、新 しいマルウェアを入手すると、その一つ一つを丁寧に解析せざるを得ない状況にあり、その作業 負担は膨大になっている。その一方で、マルウェアの攻撃はますます高度になっており、日々発 見される新しい脆弱性や、検知動作の検知を回避する仕組みを取り込みながら、次々に新しいマ ルウェアが量産されている。ただし、マルウェアは一度開発されると、検知を回避するためプロ グラムの構造の一部を変更した「亜種」が数多く生産される特徴がある。そこで、未知のデータ セットの中から、マルウェア亜種を自動的に既知の仲間に分類できれば、セキュリティアナリス トによるマルウェア解析の負担を大きく減らすことができる。本研究では、ディープラーニング を活用した高精度なマルウェア分析を実現し、セキュリティアナリストによるマルウェア解析作 業の補助や新たなマルウェア亜種の速やかな検出などマルウェアによる被害を早期に防止する ためのシステムの実現を目指す。

2. 背景

2.1 マルウェア画像

マルウェアプログラムのバイナリコードを8ビットずつに区切ることで、グレースケール画像 に変換する手法が知られている[1]。図1はこの手法を用いて作成されたマルウェア画像のサン プルである。マルウェア画像は、専門家の知識を必要とせずに生成することが可能であり、既存 の画像分類手法と親和性が高い特徴がある。



Adialer.C

Agent. FYI

図1: マルウェア画像のサンプル

2. 2. Convolutional Neural Network (CNN)

画像を分類するためのディープラーニングモデルとして、Convolutional Neural Network (CNN) が広く使われている。EfficientNet[5]は Google が 2019 年に発表した最新モデルであり、 EfficientNetB0 から EfficientNetB7 まで 8 モデルが用意されているが、このうち EfficientNetB6 と EfficientNetB7 は、他のモデルに比べて高い分類精度を持つとされる[6]。 その一方で、これらのモデルは、入力画像のサイズが大きく¹、また多くの学習パラメータを有す るため、学習時間が長くかかる課題がある。Wisteria-A (Aquarius)に搭載された複数の高速な GPUを使用することによって、この課題の解決を図る。

2. 3. 転移学習とファインチューニング

転移学習は、ある領域で得られた知識を別の領域に適用する手法である。転移学習では、高品 質なデータが大量に存在する領域から知識を移すことで、データ量が少ない領域でも分類精度の 高いモデルを作成することができる。TensorFlow[7]が提供する EfficientNet モデルは、 ImageNet[8]のデータベースに含まれる自然物の画像(植物、動物、風景、乗り物、スポーツ用 品、家具等)を事前に学習している。

ファインチューニングは転移学習の一種であり、学習済みパラメータの一部を残しつつ、残り の層を再学習することによって、分類対象のデータセットの特性に合ったモデルを構築する手法 である。しかし、マルウェア画像の分類にディープラーニングを適用した先行研究[2]-[4]では、 モデルやファインチューニングのバリエーションが限られており、自然物画像の知識がマルウェ アの分類にどのくらい役立つのかは明らかにされていない。

3. 設計

3. 1. ディープラーニングモデルとハイパーパラメータ

ディープラーニングモデルは EfficientNetB6 と EfficientNetB7 を選択する。各モデルにお ける出力層 (output layer) の活性化関数は softmax を使用する。全結合層 (fully-connected layer) の活性化関数は Relu を使用し、出力数は 128 に設定する。ハイパーパラメータについて は、損失関数 (loss function) は categorical cross entropy、最適化アルゴリズム (Optimizer) は SGD、学習率 (learning rate) は 1e-4、モメンタム (momentum) は 0.9、エポック数 (epochs) は 30、バッチサイズ (batch size) は 64 とする。

3. 2. ファインチューニング

高い分類精度を得るためには、高性能なモデルを使うことに加えて、データセットに合わせた ファインチューニングが重要である。しかし、実際に現れるマルウェアの数は攻撃の傾向に左右 されるため、最適なパラメータを事前に把握することは難しい。そこで、5段階にファインチュ ーニングしたモデルを用意して本システムに組み込むことにする。凍結した事前学習パラメータ の比率は(1) 1 (Frozen all)、(2) 3/4 (Frozen 3/4)、(3) 1/2 (Frozen 1/2)、(4) 1/4 (Frozen 1/4)、(5) 0 (Frozen none)とする。

¹入力画像のサイズは、EfficientNetB6 が 528×528、EfficientNetB7 が 600×600 である。

ファインチューニングおいて、事前学習した自然物の知識とマルウェア画像の知識を使う比率 は層の単位で指定する。表1に事前学習した自然物の知識を利用する層の数(すなわち、事前学 習パラメータを凍結する層の数)を示す。このとき、各層の持つパラメータ数は異なるため、凍 結した事前学習パラメータの比率は厳密に 3/4、1/2、1/4 にはなっていない。また、この表に示 した層の数と名前は、TensorFlow に含まれる Keras の実装に基づいている。

モデル	事前学習した自然物の知識を利用する層の数 (カッコ内は Keras の実装における層の名称)				
	Frozen all	Frozen 3/4	Frozen 1/2	Frozen 1/4	Frozen none
EfficientNetB6	666	633	560	485	1
	(top_activation)	(block7a_projec	(block6g_add)	(block6b_add)	(input_1)
		t_bn)			
EfficientNetB7	813	780	692	587	1
	(top_activation)	(block7b_add)	(block6i_add)	(block6b_add)	(input_1)

表1:ファインチューニングによって事前学習パラメータを凍結する層の数

3.3.評価指標

以下の式で定義される指標を用いて、分類結果を評価する。ここで、 tp_i は真陽性、 fp_i は偽陽性、 fn_i は偽陰性、 tn_i は真陰性、lはクラス数を表す。

$$\begin{aligned} Accuracy &= \frac{1}{l} \cdot \sum_{i=1}^{l} \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i} \\ Precision &= \frac{1}{l} \cdot \sum_{i=1}^{l} \frac{tp_i}{tp_i + fp_i} \\ Recall &= \frac{1}{l} \cdot \sum_{i=1}^{l} \frac{tp_i}{tp_i + fn_i} \\ F\text{-score} &= \frac{1}{l} \cdot \sum_{i=1}^{l} \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \end{aligned}$$

3. 4. 評価方法

評価方法には、クロスバリデーションの一種である層化 10 分割クロスバリデーションを用い る。この方法は、データセットのクラス毎に、サンプルデータの 90%をトレーニングデータ,残 りの 10%をテストデータとする分割を 10 回おこなう。このとき、すべてのサンプルデータを一 度はテストデータとして使用する。最終的な評価値は、10 回のテストで得られた結果を平均し て算出する。

4. 評価

4.1.システム環境

ハードウェアはWisteria-A (Aquarius)の8ノードを使用した。1ノードに8つのGPU (NVIDIA A100)を搭載しているため、同時に64GPUを使うことが可能である。ソフトウェアはWisteria-A (Aquarius)が提供する TensorFlow2.4.1を選択した。また、TensorFlowの機能のうち、API は Keras、分散処理は tf. distribute. Strategy を使用した。

4. 2. ジョブスケジュール

Wisteria-A (Aquarius)におけるジョブの最大同時実行本数は8本、ジョブの最大投入数は16 本である。図2にジョブスケジュールの全体像を示した。クロスバリデーションでおこなう10 回のテストのうち、1回目のテストは各モデルを個別のジョブとして設定した。2回目以降のテ ストは各モデルをまとめて1つのジョブとして設定した。この方法では、1回目から7回目のテ ストを投入したところでジョブの最大投入数に達するため、8回目から10回目のテストはジョ ブに空きが出るのを待って投入した。



図 2: ジョブスケジュールの全体像

4.3.データセット

マルウェア画像のデータセットとして、Maling を選択した。Maling は、Natarajら[1]によっ て作成されたデータセットであり、PE32 形式の Windows マルウェアプログラムを PNG 形式のマ ルウェア画像に変換したものである。データセットに含まれるマルウェア画像は 9,339 枚あり、 25 のクラスに分かれている。各画像にはマルウェアファミリーの名前がラベルとして付けられ ているため、このラベルに基づいて画像分類の正誤を判定した。

4.4 学習の実行結果

学習に要した計算時間は10時間35分であった。仮にこの計算を1GPUで行った場合は、約20日を要することが見込まれる。Wisteria-A (Aquarius)を使うことによって、約半日の実用的な計算時間でマルウェア分析を完了できることを確認した。

4.5.マルウェア画像の分類結果

図3は、EfficientNetB6およびB7のファインチューニングモデルによって得られた分類精度 (Accuracy)を示したものである。



図 3: EfficentNet6 および B7 のファインチューニングモデルによって得られた分類精度(Accuracy)

表 2 に上位 5 モデルの分類精度を示した。最も分類精度が良かったのは、EfficientNetB6 Frozen 1/4 モデルであった。Frozen 1/4のファインチューニングモデルは、自然物画像の基本 的な特徴を使いつつ、大部分は再学習したマルウェア画像の特徴を用いている。このモデルが高 い分類精度を示した背景には、自然物画像はマルウェア画像とはあまり似ていないことがあると 考えられる。ただし、自然物画像の特徴を全く使用していない EfficientNetB6 Frozen none モ デルよりも分類精度は高いことから、自然物画像の特徴は分類精度の向上にある程度の効果があ ったと考えられる。

表 2 において、EfficientNetB6 Frozen 1/4 モデルの Accuracy と F-score を比べると、Accuracy が 92.07%に対して F-score が 78.39%であり、13.68%の低下がみられた。この原因を調査するため混同行列を検証した。

Model	Frozen	Accuracy	Precision	Recall	F-score
EfficinetNetB6	1/4	0.9207	0.8104	0.7906	0.7839
EfficinetNetB6	1/2	0.8990	0.7808	0.7494	0.7419
EfficinetNetB7	1/4	0.8375	0.6704	0.6463	0.6383
EfficinetNetB7	none	0.6827	0.5425	0.5388	0.5264
EfficinetNetB6	none	0.6450	0.5030	0.5103	0.4973

表 2: 上位 5 モデルの分類精度

図4はEfficientNetB6 Frozen 1/4 モデルの混同行列である。クロスバリデーションのテストを10回行った中で、最も良い分類精度が得られた回の結果を示した。対角線上に青で示した箇所は、正しく分類されたサンプルを示している。大部分が正しく分類されているが、Autorun. K、Malex.gen!J、Swizzor.gen!Iなどの比較的サンプル数が少ないクラス(赤枠で示した箇所)に誤分類がみられた。サンプル数が少ないクラスにおける分類精度の向上が今後の課題に挙げられる。



図 4: EfficientNetB6 Frozen 1/4 モデルの混同行列

5. おわりに

本研究では、ファインチューニングした最新のディープラーニングモデルを使って、マルウェ ア分析システムを実装した。Wisteria-A (Aquarius)の高速な GPU を用いて、約半日の実用的な 時間でマルウェア分析を完了した。また、マルウェア画像のデータセットを使って分類精度を評 価した結果、EfficientNetB6 Frozen 1/4 モデルで 92.07%の分類精度を得た。サンプル数が少な いクラスにおける分類精度の向上が今後の課題である。

参考文献

- [1] L. Nataraj, S. Karthikeyan, G. Jacob, and B. S. Manjunath, "Malware Images: Visualization and Automatic Classification," in Proceedings of the 8th International Symposium on Visualization for Cyber Security, 2011.
- [2] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, and P. de Geus, "Malicious Software Classification Using Transfer Learning of ResNet-50 Deep Neural Network," in Proceedings of the 16th IEEE International Conference on Machine Learning and Applications (ICMLA), 2017, pp. 1011–1014.
- [3] Y. Mourtaji, M. Bouhorma, and D. Alghazzawi, "Intelligent Framework for Malware Detection with Convolutional Neural Network," in Proceedings of the 2nd International Conference on Networking, Information Systems & Security (NISS19), 2019.
- [4] M. Kalash, M. Rochan, N. Mohammed, N. D. B. Bruce, Y. Wang, and F. Iqbal, "Malware Classification with Deep Convolutional Neural Networks," in Proceedings of the 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS), 2018, pp. 1–5.
- [5] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proceedings of the 36th International Conference on Machine Learning, ICML 2019, pp. 6105-6114.
- [6] EfficientNet:モデルスケーリングとAutoML で最高精度を達成した Google の画像認識技術. https://developers-jp.googleblog.com/2019/07/efficientnet-automl-google.html.
- [7] TensorFlow. https://www.tensorflow.org/.
- [8] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A largescale hierarchical image database," in Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248-255.

東京大学情報基盤センター・スーパーコンピューティングニュース Vol. 24 Special Issue 1 (2022. 8)

【編集】

【発行】

東京大学情報基盤センター 〒277-0882 千葉県柏市柏の葉6-2-3 (電話) 04-7133-4663(ダイヤルイン)