

東京工業大学 TSUBAME 共同利用
令和元年度利用終了課題 利用成果報告書集

東京工業大学 学術国際情報センター
共同利用推進室

<https://www.gsic.titech.ac.jp/tsubame>

本報告書集および個別の報告書の PDF ファイルは、以下の URL にあります。

令和元年度採択実績および利用終了課題報告書

https://www.gsic.titech.ac.jp/kyodou/kadai_h31

■令和元年度産業利用 利用成果報告書 一覧

申請課題名 所属機関／利用課題責任者	頁
ファイラー分散ポリマー複合材料の相分離構造シミュレーション 先端素材高速開発技術研究組合／本田 隆	1
車載レーダにおけるターゲット散乱とレーダ反射断面積の FDTD 法による詳細解析 マツダ株式会社技術研究所／山本 雅史	5
中分子創薬のための天然変性蛋白質と中分子の相互作用の研究 株式会社 PRISM BioLab／高島 一	9
大規模記事コーパスを用いたニューラルネットワークの応用に関する研究 株式会社朝日新聞社／田森 秀明	13

■令和元年度学術利用 利用成果報告書別紙の提出免除課題 一覧 (利用成果を論文/学会等にて発表した要旨等の提出により提出免除)

申請課題名 所属機関／利用課題責任者
論文/学会等における発表済利用成果の情報
ニューラルネットワークに基づく音声合成 国立情報学研究所コンテンツ科学研究系／山岸順一
[1] Y. Yasuda, <i>et al.</i> , “Effect of choice of probability distribution, randomness, and search methods for alignment modeling in sequence-to-sequence text-to-speech synthesis using hard alignment”, Proc. ICASSP, 6724–6728, 2020. arXiv:1910.12383
[2] Y. Zhao, <i>et al.</i> , “Transferring neural speech waveform synthesizers to musical instrument sounds generation”, Proc. ICASSP, 6269–6273, 2020. arXiv:1910.12381
[3] E. Cooper, <i>et al.</i> , “Zero-Shot Multi-Speaker Text-To-Speech with State-of-the-art Neural Speaker Embeddings”, Proc. ICASSP, 6184–6188, 2020. arXiv:1910.10838
[4] S. Kato, <i>et al.</i> , “Rakugo speech synthesis using segment-to-segment neural transduction and style tokens — toward speech synthesis for entertaining audiences”, Proc. SSW10, 111–116, 2019. DOI: 10.21437/SSW.2019-20

- [5] Y. Yasuda, *et al.*, “Initial investigation of encoder–decoder end-to-end TTS using marginalization of monotonic hard alignments”, Proc. SSW10, 211–216, 2019. DOI:[10.21437/SSW.2019-38](https://doi.org/10.21437/SSW.2019-38)
- [6] X. Wang and J. Yamagishi. “Neural Harmonic-plus-Noise Waveform Model with Trainable Maximum Voice Frequency for Text-to-Speech Synthesis”, Proc. SSW10, 1–6, 2019. DOI:[10.21437/SSW.2019-1](https://doi.org/10.21437/SSW.2019-1)
- [7] F. Fang, *et al.*, “Speaker Anonymization Using X-vector and Neural Waveform Models”, Proc. SSW10, 155–160, 2019. DOI:[10.21437/SSW.2019-28](https://doi.org/10.21437/SSW.2019-28)

メタサーフェスの光学応答解析

東京農工大学／岩見健太郎

「Demonstration of focal length tuning by rotational varifocal moiré metalens in an ir-A wavelength」
Optics Express, Vol. 28, Issue 24, pp. 35602–35614 (2020) DOI:[10.1364/OE.411054](https://doi.org/10.1364/OE.411054)

ノードローカルバーストバッファの研究

筑波大学計算科学研究センター／建部修見

「GHOSTZ PW/GF: Distributed Parallel Homology Search System for Large-scale Metagenomic Analysis
Proc. BPOD 2019, pp.3492–3700, 2019 DOI:[10.1109/BigData47090.2019.9006499](https://doi.org/10.1109/BigData47090.2019.9006499)

新興再興感染症研究への分子動力学シミュレーションの応用

国立感染症研究所病原体ゲノム解析研究センター／横山勝

「Allosteric regulation of HIV-1 capsid structure for Gag assembly, virion production, and viral infectivity by a disordered interdomain linker.」 J. Virol., 93, e00381–19, 2019. DOI:[10.1128/JVI.00381-19](https://doi.org/10.1128/JVI.00381-19)

近似ベイズ推定を用いた分散並列深層学習

理化学研究所革新知能統合研究センター／モハマッド エムティヤズ カーン

「Practical Deep Learning with Bayesian Principles」
The 33rd Conference on Neural Information Processing Systems 32 (NeurIPS 2019)

X線多波回折現象の計算機シミュレーション

東京大学工学系研究科／沖津康平

- [1] 「X線多波(n波)動力学理論とその数値解法および放射光実験による検証」
日本放射光学会誌 Vol.33 No.2 pp61–80.
- [2] K. Okitsu *et al.*, “Experimentally obtained and computer-simulated X-ray asymmetric eight-beam pinhole topographs for a silicon crystal”, Acta Cryst. **A75** (2019) 474–482. DOI:[10.1107/S2053273319001499](https://doi.org/10.1107/S2053273319001499)
- [3] K. Okitsu *et al.*, “Experimentally obtained and computer-simulated X-ray non-coplanar 18-beam pinhole topographs for a silicon crystal”, Acta Cryst. **A75** (2019) 483–488. DOI:[10.1107/S2053273319002936](https://doi.org/10.1107/S2053273319002936)

機能物性解析のための計算科学技術の開発

豊橋技術科学大学／後藤仁志

- [1] 「グラフ畳み込みニューラルネットワークによる立体配座—エネルギー相関」
応用物理学会 有機分子・バイオエレクトロニクス分科会会誌
- [2] Yoshiki Kato; Shinji Hamada; Hitoshi Goto, “Validation Study of QSAR/DNN Models Using the Competition Datasets”, Molecular Informatics, 39, 1900154 (2020). DOI:[10.1002/minf.201900154](https://doi.org/10.1002/minf.201900154)

■令和元年度学術利用 利用成果報告書 一覧

申請課題名	所属機関／利用課題責任者	頁
ハイドロゲル保水性発現機構の分子論的解明	大阪大学大学院基礎工学研究科／松林伸幸	19
HPC を利用した自然言語処理技術の研究	情報通信研究機構 ユニバーサルコミュニケーション研究所／鳥澤 健太郎	21
マルチ GPU による心血管系の血流の数値シミュレーション	東北大学材料科学高等研究所／水藤寛	23
第一原理計算による電池・触媒メカニズム解明と新物質探索	物質・材料研究機構／館山 佳尚	25
ボルテックスジェネレータ周りの熱流動直接数値解析	大阪府立大学／桑田 祐丞	27
LRnLA アルゴリズムを用いた物理シミュレーション	法政大学情報科学部／善甫 康成	31
近赤外線を利用する色素の探索	自然科学研究機構アストロバイオロジーセンター / 国立天文台／小松勇	35
高解像度画像を使った広域の家屋及び道路の深層学習による自動判別システムの開発(3)	東京大学空間情報科学研究センター／柴崎 亮介	37
TSUBAME3.0 における TensorFlow アプリケーションの性能評価	電気通信大学／三輪 忍	41
分子動力学シミュレーションを用いた抗関節リウマチ分子標的薬の設計	長岡工業高等専門学校／和久井直樹	43
GPGPU による巨大地震の長周期地震動シミュレーション	国立研究開発法人防災科学技術研究所／藤原広行	47
計算化学による固体分散体中の薬物とキャリアの相互作用の評価	千葉大学大学院薬学研究院製剤工学研究室／東 顕二郎	53
サブナノ粒子の構造及び性質に関する理論的研究	京都大学福井謙一記念研究センター／春田 直毅	57

申請課題名 所属機関／利用課題責任者	頁
ダイヤモンド NV 中心の励起状態に関する理論的研究 京都大学福井謙一記念研究センター／春田 直毅	59
流動摩擦抵抗予測法の構築に向けた粗面乱流直接数値解析 大阪府立大学工学研究科／桑田 祐丞	61
深層学習による高解像度化手法の水文学への応用 熊本大学／石田桂	65
損傷船体の縦曲げ最終強度と波浪中安全性評価に関する研究 神戸大学海洋底探査センター／橋本 博公	67
抗体の安定性解析のための分子動力学シミュレーション 東京大学大学院医学系研究科／石川俊平	71

TSUBAME 共同利用 令和元年度 産業利用 成果報告書

利用課題名 フィラー分散ポリマー複合材料の相分離構造シミュレーション
 英文: Simulation of phase separated structures of polymeric materials with fillers

本田 隆
 Takashi Honda

先端素材高速開発技術研究組合
 Research Association of High-Throughput Design and Development for
 Advanced Functional Materials (ADMAT)
<http://www.admat.or.jp/>

粒子フィラーと高分子混合系における粒子フィラーの凝集状態を高分子の SCF 理論において、粒子フィラーを球形に形状を固定された DIP(Diffuse Interface Particle)として取り扱い、古典的な Flory-Huggins 理論を改良した相図と比較することにより、高分子熔融体中のフィラー粒子が一種の相分離におけるスピノーダル条件となり凝集することが MPI+GPGPU 大規模計算により確認することができた。

The aggregate state of particle fillers in polymer melt was treated as a state of DIPs (Diffuse Interface Particle) whose shape is fixed to a sphere shape in a polymer melts with SCF theory and compared with the phase diagram which improved the classical Flory-Huggins theory. In this method it was confirmed by a large-scale calculation of MPI + GPGPU that particle fillers aggregate under a spinodal condition in a kind of phase separation.

Keywords: SCF theory, polymer composite, particle filler, fiber filler, plate filler, diffuse interface particle

背景と目的

高分子材料は、改質剤として有機物質や、充填材とし無機物質であるフィラー等様々な添加物がされ、その物性が改良されて用いられることが多い。タイヤ材料ではカーボンブラックやシリカが添加され、最近では CNT を添加し電気伝導性・熱伝導性に優れた材料が開発されている。しかし、高分子材料は、高分子量ゆえ、同一重量中の分子数は低分子に比べ少なく、異種物質との混合時にはエントロピー効果による自由エネルギーの低下量が小さくなる。また、高分子間の相互作用や高分子の絡み合い等にも影響を受けるので、通常相分離構造をとると考えなくてはならない。よって、フィラー・高分子複合材料の構造解析・評価は難しく、材料の開発期間は長時間となる場合が多い。

本課題においては、高性能な高分子複合材料の開発期間短縮のため、高分子の SCF 法(密度汎関数法の一つ:以降では簡単のため SCF 法と記載する)[1,2] を利用し、フィラーの凝集構造と高分子マトリックスの相分離構造を同時にシミュレートできるプログラム Soft Blends Analyzer(SOBA)[3]の改良をすることを目的

とし、SOBA を利用した具体的なアプリケーションとして、高分子とフィラーを混合した場合の相分離の有無をシミュレーションにより予測する試みを行った。

概要

1. 理論:フィラーを充填した系の高分子の SCF 法

フィラーの凝集構造をシミュレートするためにフィラーの境界に濃度勾配がある溶液粒子(以降では Diffuse Interface Particle: DIP)を利用した高分子の SCF 計算を用いた。DIP の形状は次のモデル式で記述される。

$$\phi_p(\mathbf{r}) = \frac{\tanh\{\alpha(R-|\mathbf{r}-\mathbf{r}_0|)\}+1}{2} \quad (1)$$

ここで、 $\phi_p(\mathbf{r})$ はフィラーの濃度分布、 R はフィラー粒子の半径、 \mathbf{r}_0 は粒子中止の位置、 α は界面の幅を設定するパラメータである。

フィラーを充填した系の高分子の SCF 法の説明は、先の報告書等[4, 5]に記載したので詳細は割愛する。

2. 理論:フィラーと高分子混合系の相図

古典的な理論として、高分子混合系の相図を予測するには、Flory-Huggins の混合の自由エネルギーを利

用する方法がある。ポリマー A 種と B 種の Flory-Huggins の混合の自由エネルギーの理論式は次の通りである。

$$\frac{\Delta G}{k_B T} = \frac{\phi_A}{N_A} \ln \phi_A + \frac{\phi_B}{N_B} \ln \phi_B + \chi_{AB} \phi_A \phi_B \quad (2)$$

ここで、 N_A はポリマー A の一本鎖が占める体積、 N_B はポリマー B の一本鎖が占める体積、 ϕ_A, ϕ_B はそれぞれ、ポリマー A と B の体積分率、 χ_{AB} はポリマー A と B が隣り合った場合に上昇するエネルギーに関するパラメータであり、Flory-Huggins の χ パラメータと呼ばれる。右辺において、第 1 項と第 2 項はポリマー同士の混合におけるエントロピー効果による自由エネルギーを、第 3 項は異種ポリマーが隣り合うことによるエンタルピー効果による自由エネルギーを表す。この式を高分子とフィラーの混合系に展開すれば、次のような式を考えることができる。

$$\frac{\Delta G}{k_B T} = \frac{\phi_A}{N_A} \ln \phi_A + \frac{\phi_F}{V_F} \ln \phi_F + \chi_{AF} \chi_F \phi_A \phi_F \quad (3)$$

ここで、指数 F はフィラーを意味し、 V_F は半径 R の球としてフィラー 1 個の占める体積であり、 χ_F はフィラーを利用した場合の χ_{AF} に対する補正パラメータである。フィラー中心部は外部と接触することができなくなるので、フィラーが外部と接触できる体積の割合は、近似的に (フィラーの表面部分の微小体積) / (フィラーの体積) である。よって、 χ_F は次のように考えることができる。

$$\chi_F = \left[\int_{R-\delta r}^R 4 \pi r^2 dr \right] / V_F \quad (4)$$

ここで、 δr はフィラーの表面部分の微小体積を計算するための任意の表面厚さである。

このポリマー・フィラーの混合の自由エネルギーの第 3 式を利用して、ポリマー・フィラー混合系の相図を考察し、さらに SOBA による高分子・フィラー相分離シミュレーションを比較することにより、簡易的ではあるがこれまで理論的に解釈することが難しかったポリマー内のフィラーの凝集現象を相分離として評価することが可能となる。

結果および考察

1. モデル系

鎖長 $N=10$ のポリマー・溶融系に $R=2$ 、 $\alpha=3$ の DIP フィラーを添加した場合をベンチマーク用のモデルとして評価することとした。

2. フィラーと高分子混合系の相図

第 3 式を利用した相図を OCTA/POTAG3 を改良して表示できるようにした。計算条件として、第 4 式における δr の値が必要であるが、これは、第 1 式を体積積分することにより DIP の体積を算出して理想的な球の体積との差を取ることで余剰な体積を計算できるので、その余剰体積がフィラーの表面部分の微小体積に相当すると仮定して δr を求めた。その値は $\delta r=0.15$ であった。図 1 に示すのが計算した相図であり、 χ_{AF} を 0.2 刻みで 2.0 まで変化させたところ、 $\chi_{AF}=1.8 \sim$ で相分離の臨界点に達した。POTAGE の出力形式で三角相図となっているが、上の頂点がフィラー、下の左右の点が同一のポリマーを意味する。相図に示す水平方向の赤い線はスピノーダル線であり、この線の内側でポリマーとフィラーは強く相分離するはずであり、上下に斜め引かれた青いラインはタイラインであり、タイラインの末端を繋ぐ水平線がバイノーダル線である。計算の収束性の問題で、中央部のバイノーダル線は表示されていない。図 1 で予想されるのは、 $\chi_{AF}=1.8 \sim 2.0$ でフィラーの体積分率 $\phi_F=0.3 \sim 0.4$ の領域がスピノーダル領域になることである。

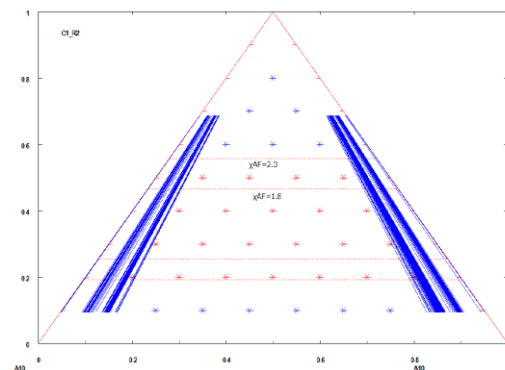


図 1 拡 Flory-Huggins 理論によるポリマー・フィラー混合系の相図

上の頂点がフィラー、下が同一のポリマーである赤い線がスピノーダル線、青い線がタイラインである。

3. ポリマーブレンド中の DIP

SOBA を利用した計算は、システムサイズ 128^3 、メッシュサイズ 256^3 のサイズのセルにポリマーとフィラーを投入した系である。ここでの単位は全て無次元化したものである。フィラーの初期位置はランダムな配置とし、10 ステップの SCF 計算事に DIP に作用する相互作用

用を計算し、仮想的な易動度 $D_e=0.015$ での動力学計算を行った。このような動力学をすると DIP の運動が止まらないので、運動させる度に 0.999 を D_e に掛け、運動が停止するようにした。全 SCF ステップは 100,000 ステップとした。計算には GPGPU 8 枚の MPI 並列計算を用いた。一例として、DIP の体積分率 $\phi_F=0.1$ 、 $\chi_{AF}=2.0$ とした場合の最終構造を図 2 に示す。

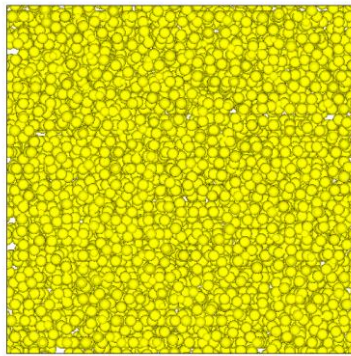


図 2 動力学後の DIP の分散構造の例

$\phi_F=0.1$ 、 $\chi_{AF}=2.0$ 、表示は $\phi_F=0.5$ の値の等値面

図 1 の相図と比較するため、 $\phi_F=0.4$ の最終構造から得られた DIP 間の二体相関関数を図 3 に示す。 χ_{AF} は図 1 の相図で予想された相分離の臨界点近傍 $\chi_{AF}=1.8$ 近辺で変化させた。図 3 で特徴的なのは、 $\chi_{AF}=1.6 \sim 2.0$ で $r=1 \sim 2$ の領域で、わずかにピークが認められることである。 $R=2$ であるので、二体相関関数は R の 2 倍の 4 程度から出現するべきものであるが、ソフトな粒子としてあるので、 χ_{AF} の増大によりスピノーダル領域に突入し、DIP 同士が凝集し重なり出していると考えられることができる。

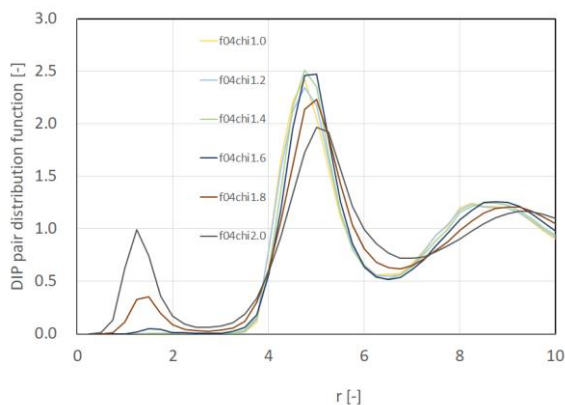


図 3 DIP の 2 体相関関数

$\phi_F=0.4$ の場合

この重なり合いの確認は、系の散乱関数からも確認することができる。図 4 に示すのは、 $\phi_F=0.4$ の場合の散乱関数であり、 $\chi_{AF}=1.6 \sim 2.0$ において、広角側の粒子散乱の特徴的なピークの連続が低減している。これは、粒子同士が重なり、散乱に寄与する粒子形状が粒子の重なりにより消失し出したためと考えられる。

以上の結果から拡張 Flory-Huggins 理論の $\delta r=0.15$ で計算し予測された臨界点 $\chi_{AF}=1.8$ 近傍である $\chi_{AF}=1.6$ 以上 で DIP が凝集を開始しているといえる。スピノーダル条件と DIP の重なりが同等な意味を持つと一概に結論することは難しいが、高分子の SCF 計算を基本としたスキームであり、臨界点を境として大きく相分離構造が変化するはずであるので、スピノーダル条件により DIP の重なりが起きたと考えることは妥当であると考えられる。

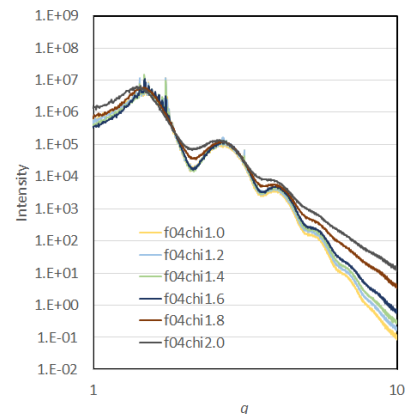


図 4 DIP の散乱関数

$\phi_F=0.4$ の場合

まとめ、今後の課題

理論的な評価が難しい、高分子とフィラー混合系における粒子の凝集の評価を拡張 Flory-Huggins 理論により相図を予想し、さらに SOBA を利用した MPI+GPGPU 並列計算シミュレーションで、予想されたスピノーダル条件で、DIP の凝集による重なりが発生し、フィラーの凝集が起こっていることをシミュレートできたと考える。よって、高分子中のフィラーの凝集条件は、理論的には拡張 Flory-Huggins 理論により非常に簡易・高速に推定することが可能であるという成果が得られた。また、大規模計算を行うことにより高分子とフィラーというスケールの異なる成分の混合系においても、フィラーを模擬した DIP の 2 体相関関数や散乱

関数を大きな誤差を伴うことなく計算することができ、MPI+GPGPU 大規模計算の利点を示せたといえる。

開発したプログラムは、<http://octa.jp/> より SOBA Ver.2. や SUSH11.0 として公開予定である。

今後は、粒子の動力学における CUDA を利用した並列計算の改良や、様々な高分子複合材料のアプリケーションの例を増やしてゆきたい。

謝辞

この成果は、国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の委託業務 (P16010) の結果得られたものです。

参考文献

- [1] 公益社団法人 新化学技術推進協会 (編集)、“増補版 高分子材料シミュレーション: OCTA 活用事例集”、化学工業日報社 (2017).
- [2] T. Honda and T. Kawakatsu, “Applications of SUSHI in OCTA system” in *Nanostructured Soft Matter: Experiments, Theory and Perspectives*, A. V. Zvelindovsky, ed., Springer-Verlag (2007).
- [3] <http://octa.jp/> にて公開中 (アカウント取得と login が必要であるが利用は無料) (2019).
- [4] 本田 隆、“フィラー分散ポリマー複合材料の相分離構造シミュレーション”、TSUBAME 共同利用 平成 30 年度 産業利用 成果報告書
- [5] H. Morita, M. Toda and T. Honda, *Polymer Journal*, **48**, 451 (2016).

TSUBAME 共同利用 令和元年度 産業利用 成果報告書

車載レーダにおけるターゲット散乱とレーダ反射断面積の FDTD 法による詳細解析
Detailed analysis of target scattering and radar reflection cross section in automotive radar by FDTD method

山本 雅史
Masashi Yamamoto

マツダ株式会社
Mazda Motor Corporation
<http://www.mazda.co.jp>

車載レーダは、自動運転や安全運転支援システムにおいて走行環境認識のための重要なデバイスである。レーダ性能向上のためには、車両ターゲットの様々な位置、角度変化に対する受信電力変動を考慮した上で、アンテナ・車載化要件を設定する必要がある。従来から RCS (Radar Cross Section) はターゲットからの反射波強度を把握する上で重要な特性であるが、車両ターゲットからの反射波の受信電力と比較解析した例は少ない。今回、TSUBAME3.0 を活用した FDTD (Finite Difference Time Domain) 法シミュレーションにより比較解析を行った。その結果、5GHz 周波数帯において、FDTD 法でシミュレーションした車両ターゲットからの反射波の受信電力と、RCS とレーダ方程式から算出した受信電力の相対値がほぼ一致していることを確認できた。

An automotive onboard radar is an important sensing device for autonomous driving systems and ADAS (advanced driver-assistance systems) to realize environment recognition functions. RCS (Radar Cross Section) is an important characteristic to estimate intensity of reflected wave from a target, but there are few examples of studies on comparing the RCS of a vehicle target with the received power of an antenna. In this study, comparison analysis was conducted by a FDTD (Finite Difference Time Domain) simulation using TSUBAME3.0. The received power of the reflected wave from a vehicle target was computed by the FDTD simulation and also calculated from the RCS and the radar equation at the 5GHz frequency band. Both results were compared after normalization. It was confirmed that they were almost the same.

Keywords: Automotive radar, RCS, FDTD, Electromagnetic simulation, Radar equation

1. 背景と目的

自動運転システムや安全運転支援システムにおいて車載レーダは走行環境認識のための重要なセンシングデバイスである。レーダ性能向上のためには、車両ターゲットの様々な位置、角度変化に対する受信電力変動を考慮した上で、アンテナ、信号処理、車載化条件を設計する必要がある。従来から RCS (Radar Cross Section) [1] がターゲットからの反射波強度を把握する上で重要な特性となっているが、車両ターゲットにおける受信電力と比較解析した例は少なく、受信電力の増減と RCS との関係性やその関係性の要因となるターゲットからの反射波発生と電波伝搬の詳細は十分に説明されていない。

レイトレース法や物理光学近似 (PO) 法を活用して RCS と受信電力を数値解析で求める手法があるが、

電波の波動性と粒子性 (光線性) のうち粒子性を重視した再現のため、近似を含んでおり再現精度に課題があることと、その結果の要因を解明するためのターゲットの反射、散乱、回折現象とアンテナまで到達する過程の電波伝搬の物理現象を精緻に可視化して解析することは困難である。一方、電波の振る舞いを現すマクスウェル方程式を離散化することで高精度に解く FDTD (Finite Difference Time Domain) 法 [2] がある。FDTD 法は、電波の波動性と粒子性の双方を再現し、解析空間全体を時間領域で解く手法のため、ターゲットへの入射からアンテナまでの伝搬過程を詳細に解析して可視化を可能にする。しかし、FDTD 法は波長の 10 分の 1 以下のセルで解析領域を分割して計算するため、車載レーダのミリ波帯の解析には膨大な計算機メモリと計算時間を必要とする問題がある。

マルチノード、マルチ GPU 計算機である TSUBAME3.0 を用いることで FDTD 法が抱える計算資源(メモリ量と計算時間)の問題を解決できる。本プロジェクトでは、TSUBAME3.0 上で高次精度差分スキームの FDTD(2,4)法と実スケール 3D モデルを用いることで、ターゲットからの反射波を波動性と粒子性の双方で高精度に再現、解析することを可能にした。

今回は、アンテナに対する車両ターゲットの距離と角度を変化させた際の RCS と受信電力の関係性を把握することを目的とした。

2. 計算条件

FDTD 法で車両ターゲットからの反射波を再現し、受信電力を観測する解析モデルを図 1 に、計算条件を表 1 に示す。今回、車両形状による影響を考慮して RCS 特性と受信電力を解析するために、実車両の 3D CAD データを FDTD 法に取り込み計算した。ただし、車両の金属部からの反射が支配的であるとの仮定のもと、タイヤのゴムやウィンドウのガラス等の誘電体材料は取り除いた金属部のみのデータを用いた。車両ターゲットはレーダアンテナ送信軸上に配置し、アンテナに対して車両背面が正対する角度を 0 度として、車両の角度を ±45 度の範囲で変化させた。アンテナとターゲット間の距離は 5.76m、11.1m、25m の 3 条件とした。地面からのアンテナ高さはセダンのレーダ搭載位置を想定して 0.5m とした。受信電力の観測点は、アンテナの直前にアンテナと同じ高さに設定した。周波数はマイクロ波帯の 1GHz と 5GHz の 2 条件でそれぞれ計算し、FDTD 法で求めた受信電力と、RCS とレーダ方程式から算出した受信電力を比較した。レーダ方程式は、直接波と金属床の反射波の干渉を考慮した式(1)を用いた。

$$P_r = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} \left| 1 + \gamma \frac{R_1}{R_2} \exp\left(-j \frac{2\pi}{\lambda} (R_2 - R_1)\right) \right|^4 \quad (1)$$

ここで、 P_r : 受信電力、 P_t : 送信電力、 R : ターゲットまでの距離、 R_1 : 直接波の伝搬距離、 R_2 : 金属床反射波の伝搬距離、 G : アンテナ利得、 λ : 波長、 σ : RCS、 γ : 金属床の反射係数である。

車載レーダで用いられている周波数である、76GHz 帯や 24GHz 帯での大規模解析シミュレーションは、1 条件の計算でおおよそ 50 時間以上要すると見込む。

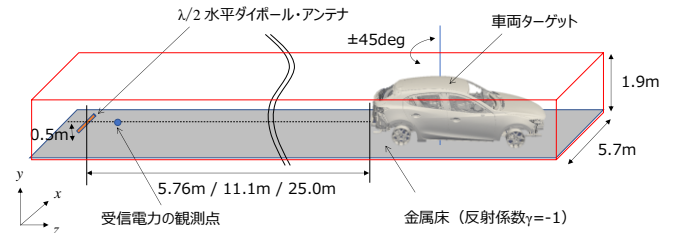


図 1. 車両ターゲットの反射解析モデル

表 1. FDTD 法の計算条件

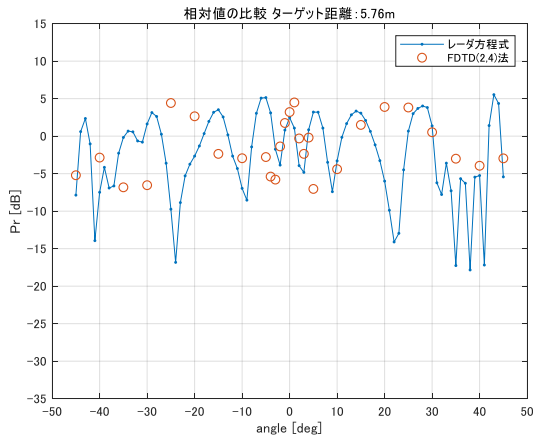
周波数	1.0 [GHz] / 5.0 [GHz]
解析空間	192x64x1024 [cell] / 768x320x5120 [cell]
セルサイズ	2.998x10 ⁻² [m] / 5.996x10 ⁻³ [m] (≒1/10 λ)
計算手法	FDTD (2,4)
クーラン数 CFL	1.3632x10 ⁻¹
時間ステップ	1.36x10 ⁻¹¹ [sec] / 2.73x10 ⁻¹² [sec]
放射源	½波長ダイポールアンテナ+連続正弦波
送信電力	7.614 [dBm] / 7.575 [dBm]
計算ステップ数	80,000 [回] / 200,000 [回]
吸収境界	PML 32 層、R ₀ =1.0 ⁻³² 、M=4
変数の型	float (GPU)

そこで今回は、計算負荷が小さい 1GHz と 5GHz の 2 つの周波数帯で距離と角度の多様条件で RCS と受信電力の関係の傾向を事前に把握することで、今後行う 76GHz 帯や 24GHz 帯での受信特性と伝搬特性の仮説を立て、計算条件の絞り込みを行いたいと考えた。

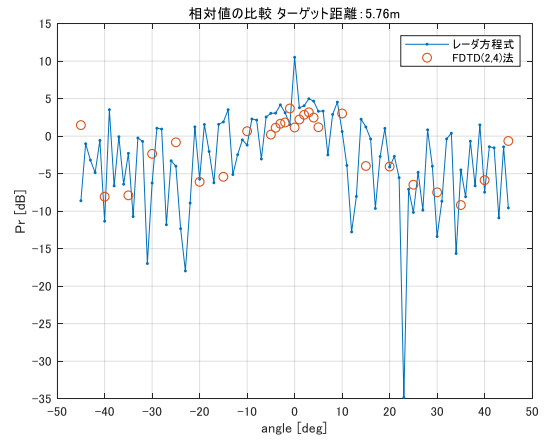
3. 結果および考察

FDTD 法シミュレーションにより求めた受信電力と、RCS とレーダ方程式から求めた受信電力をターゲットまでの距離別に比較した。比較に際して、角度が ±45 度の範囲で平均値を求め、平均値が 0dB となるように正規化した相対値を用いた。この理由は、FDTD 法とレーダ方程式で、アンテナ特性である送信電力とアンテナゲインの相違の影響を無くすためである。

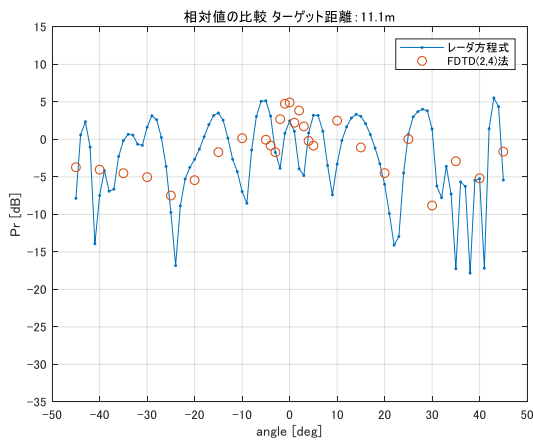
1GHz でターゲットまでの距離別の結果を図 2(a) ~ (c) に、5GHz の結果を図 3(a) ~ (c) に示す。この結果より、1GHz ではターゲットまでの距離に関わらず FDTD 法とレーダ方程式の結果は、極端な信号レベルの相違ではないが、ピークおよび落ち込みの現れる角度や、ターゲット角度に対する信号レベル変化に一致



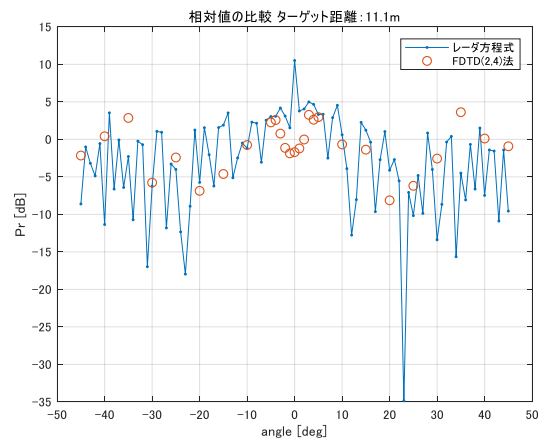
(a) ターゲットまでの距離: 5.76m



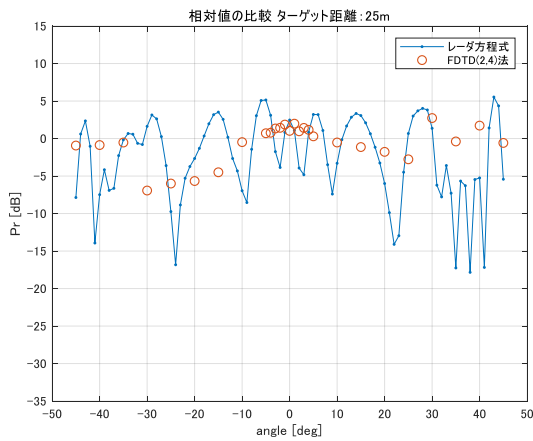
(a) ターゲットまでの距離: 5.76m



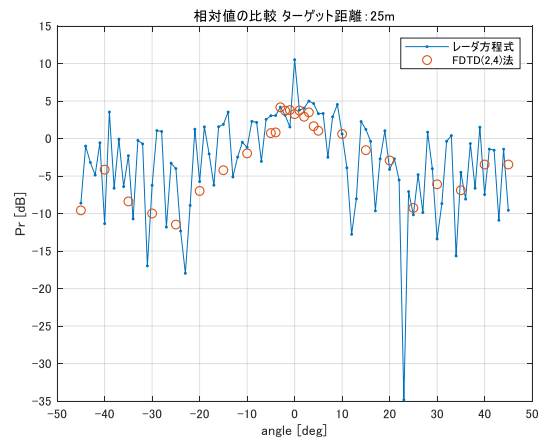
(b) ターゲットまでの距離: 11.1m



(b) ターゲットまでの距離: 11.1m



(c) ターゲットまでの距離: 25.0m



(c) ターゲットまでの距離: 25.0m

図 2. 1GHz の受信電力の相対値の比較

図 3. 5GHz の受信電力の相対値の比較

は見られなかった。一方、5GHz ではいずれのターゲットまでの距離でも、ターゲット角度変化に対する信号レベルおよび増減変化の特徴がほぼ一致していることを確認した。

FDTD 法で求めた受信電力と、RCS とレーダ方程

式から算出した受信電力の間の相関を評価するため、相関係数を条件ごとに求めた結果を表 2 に示す。5GHz では 5.76m と 25m の条件でかなり強い相関があり、11.1m で相関があった。1GHz では弱い相関が相関がないことを確認した。

表 2. FDTD 法とレーダ方程式の結果の相関係数

ターゲットまでの距離	周波数: 1GHz	周波数: 5GHz
5.76m	0.006	0.663
11.1m	0.256	0.381
25.0m	0.107	0.772

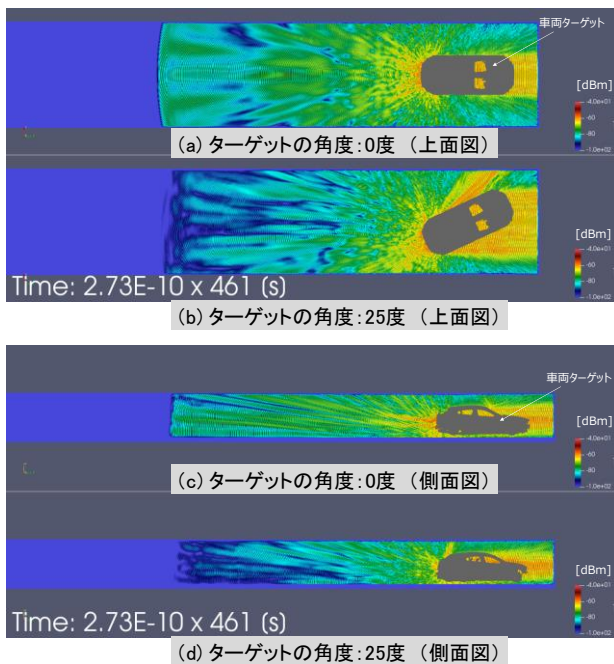


図 4. 5GHz、距離 25m の散乱界による電力空間分布

5GHz 連続正弦波の車両ターゲットからの散乱界の過渡状態における電界強度を受信電力に換算して可視化した結果を図 4(a)～(d)に示す。受信電力の空間分布はコンター表示しており、赤色になるほど電力が高く、青色になるほど電力が低いことを表している。ターゲットまでの距離は 25m の条件である。図 4(a)と(b)は、車両ターゲットが 0 度と 25 度の条件で同一時刻の電力分布の上面図であり、図 4(c)と(d)は側面図である。

車両ターゲットが 0 度の場合では、車両のリアパネルとバンパー付近で強い反射が起こり、その強い反射波がアンテナ方向に伝搬していることを確認できる。また、25 度の場合は車両の側面で強い反射が起こり車両ボディに沿って強く伝搬しているのに対して、アンテナ方向への反射は全体的に弱く、特にアンテナ送信軸上は反射波が干渉して信号レベルが落ち込むヌルとな

っていることを確認できる。この結果は、図 3 で示した、車両ターゲット角度が 0 度で信号レベルが高く、25 度で大きく落ち込み、その差が 15dB～20dB であった要因を示唆している。つまり、車両ターゲット角度が 25 度で受信電力が大きく低下したのは、車両ターゲットでのアンテナ方向への反射が弱いことに加えて、反射波が干渉したためと推察する。

4. まとめ、今後の課題

車両ターゲットに 1GHz と 5GHz の連続正弦波を照射する条件で、ターゲットからの反射波の受信電力を FDTD 法によりシミュレーションした結果と、RCS とレーダ方程式から算出した結果を相対値で比較した。その結果、5GHz 周波数帯では、ターゲット距離と角度変化によらず 2 者がほぼ一致していることを確認できた。また、5GHz ではアンテナに対する車両ターゲットの角度変化により、受信電力で 15dB～20dB 程度の差異が生じており、車両ターゲットからアンテナ方向への反射の強さの違いに加えて反射波の干渉が影響していることが分かった。

今回の 1 条件の計算には、最大で 2 ノードの 8GPU を使用し、メモリ量は 115GB、計算時間は 2.6 時間であった。

今後、車両ターゲットの形状と大きさが異なる場合と、周波数が実際の車載レーダの 76GHz や 24GHz の場合において、RCS と受信電力の関係性の解明に取り組む。

参考文献

- [1] Eugene F. Knott, 他, Radar Cross Section (2nd Edition), scitech publishing, (2004)
- [2] 宇野 亨, FDTD 法による電磁界およびアンテナ解析, コロナ社, p.22(2009)
- [3] 園田 潤, 大規模電波伝搬解析に適した波動方程式に基づく FDTD(2,M)法の数値分散と並列計算特性, 信学技報, AP2009-12, p.7-12(2009)
- [4] 園田 潤, 高次 FDTD 法とクラスタを用いた並列計算による大規模電波伝搬解析に関する研究, 東北大学博士学位論文, p.78-83(2005)

TSUBAME 共同利用 令和元年度 産業利用 成果報告書

利用課題名 中分子創薬のための天然変性蛋白質と中分子の相互作用の研究
英文: Study on intrinsically denatured protein-middle molecule interactions for drug discovery

高島 一

Hajime Takashima

株式会社 PRISM BioLab
PRISM BioLab Co., Ltd.
<http://www.prismbiolab.com/>

蛋白質間相互作用(PPI)を指向した創薬が注目を集めているが、どのように化合物が標的蛋白質を認識し結合しているかについてはまだ十分な理解が得られていない。今回我々は、構造の柔軟な蛋白質複合体として mSin3-NRSF/REST 複合体を例にとり、その結合を阻害する α ヘリックス模倣化合物 mS-11 と mSin3 の相互作用を、GA-based mD-VcMD で計算した。その結果、mS-11 はその模倣配列部分である"LIML"配列領域に最も安定に結合することを示し、またその結合過程において分子配向の秩序化が起きることが分かった。これらの結果は、PPI を指向した我々の分子創出戦略が有効である事を示しており今後の PPI 創薬やドラッグデザインへの応用が期待できる。

For drug discovery of protein-protein interaction (PPI), it is important to understand how PPI inhibitors can recognize and bind to a target protein. In this article, we calculated interaction and binding of mS-11, alpha-helical mimetic compound of NRSF/REST protein, and its binding protein mSin3 using GA-based mD-VcMD method. The calculation results showed that mS-11 binds to "LIML" sequence region, which is an alpha-helical part of NRSF/REST and corresponds to the mimetic region of mS-11. We also found ordering of molecular orientation during its binding process. These results indicate that our drug design strategy using alpha-helical mimetic compound is effective for PPI drug discovery.

Keywords: PPI (protein-protein interaction) inhibitor, alpha-helical mimetics, mD-VcMD, enhanced sampling, free energy

背景と目的

近年は蛋白質間相互作用を阻害ないし促進できる新しい分野の創薬が注目を集めており、多くの国内外企業も PPI 創薬への取り組みを始めている。従来の酵素や受容体を標的とした低分子創薬では、それらの天然の基質分子の多くが単離・同定されていることから、それらの類似構造を持つ化合物として低分子医薬品の探索が可能であった。また、酵素・基質は比較的、蛋白質立体構造が堅いものが多く、構造解析と薬物ドッキングシミュレーション計算からの医薬品の探索も可能な場合が多かった。

しかし、PPI を指向した創薬では、類似化合物の探索が困難であり、標的蛋白質の構造が柔軟であるためにドッキングシミュレーション計算での探索も困難、そして、それらの結合様式も予測が困難で分子の改良も困難である。このような状況の中、これまで株式会社

PRISM BioLab では主に α ヘリックス構造を模倣した独自の中分子化合物ライブラリーを開発してきた。また、これらの新規中分子を DISC(産学協働スクリーニングコンソーシアム)にも提供してきた。このような α ヘリックス構造を模倣した分子が標的蛋白質へどのように結合できるのかを計算シミュレーションで支援することで、効率的かつ低コストなドラッグデザインが実現できれば、より広く社会で PPI 創薬が普及することが期待できる。

転写制御因子 NRSF/REST は、転写コリプレッサーである mSin3 と結合することで多くの神経特異的な遺伝子の発現を制御している。我々はこれまでに、mSin3 に結合して NRSF/REST との蛋白質-蛋白質間結合を特異的に阻害する新規化合物 mS-11 を見出した(Bioorganic & Medicinal Chemistry Letters 27 (2017) 4705-4709)。mS-11 は、NRSF/REST 蛋白質の α ヘリックスを形成している"LIML"配列領域

(アミノ酸1文字表記法)を模倣するよう設計された化合物である。当該研究では、mS-11とmSin3の結合と解離を司る自由エネルギー地形を、分子動力学シミュレーションを用いて算出し、結合解離過程を再現してそのメカニズムを解明し理解する事を目指した。

通常の分子動力学シミュレーションでは、分子間の結合・解離の自由エネルギー地形を算出するのは一般に困難である。そこで、シミュレーションが局所安定構造にとらわれず、構造空間を広く探索するために、共同研究者(肥後順一)の開発した「genetic-algorithm に支援された、多次元仮想系と相互作用する共役した分子動力学法(GA-based mD-VcMD)」を適用することで解決を図った。シミュレーションの実行ではTSUBAMEのGPUを用い多数の独立なrunsを流して2分子の結合・解離過程の広範な立体構造を探索し、得られた立体構造に熱力学的重みを割り付け、それを構造空間にマッピングする(すなわち、これが自由エネルギー地形の算出である)。それにより、mS-11との結合様式(最も自由エネルギーが低い複合体構造)のみならず、準安定な複数の複合体構造とその自由エネルギー値を算出することが原理上可能である。

今回、実施に256runからなるGA-based mD-VcMD計算を実行し、mS-11のmSin3蛋白質への結合様式のみならずその結合メカニズムについても考察することができた。mS-11は、その模倣配列部分に相当するNRSF/RESTの“LIML”配列結合領域に最も安定に結合することを示した。また、mS-11がmSin3に接近するにつれ、配向の秩序化が起きることを示した。

概要

製薬・医療健康分野において、中分子創薬は、エピゲノム調整や遺伝子の調整に必要な蛋白質間相互作用を阻害・ないし促進できる新しい分野の薬(新規モダリティー)として注目を集めており、多くの国内企業も中分子への取り組みを始めている。しかし、中分子が実際に生体において、どうやって標的蛋白質を認識し、結合し、薬効を発揮できるのか、といったことには、十分な

理解が得られていない。今回我々は、構造の柔軟な蛋白質複合体としてmSin3-NRSF/REST複合体を例にとり、その結合を阻害する α ヘリックス模倣化合物mS-11とmSin3の相互作用をGA-based mD-VcMDで計算した結果、mS-11とmSin3の最安定結合様式のみならず準安定な複数の複合体構造、またその結合過程を解明することができた。

結果および考察

Fig. 1aは、mSin3(緑色の分子)周囲のmS-11の存在確率を等密度マップで示したものである。H1およびH2は、mSin3に含まれる4本のヘリックスのうちの一つである。残りの二つ(H3とH4)は紙面の後ろに隠れている。マップの色が赤→青→マゼンタ→深緑となるにつれ存在密度が下がる(密度の数値を説明するには、専門的な長い説明が必要なので、ここでは述べない)。高密度の場所、つまりmS-11がそこにくと熱力学的に安定になる場所は、赤い等密度マップの部分である。ここには、二つのクラスター(A and B)が存在しており、クラスターAの方がBよりも安定である(Aの方がBよりも大きい)。mS-11はmSin3の周囲に広く分布できるが(深緑色の等密度マップ)、高密度の安定結合領域は、クラスターAとBに限られている。

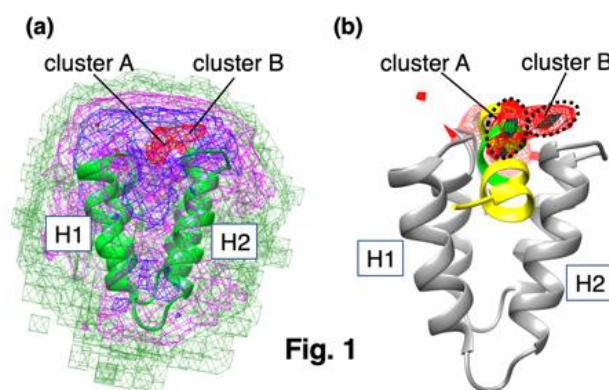


Fig. 1bは、NRSF/REST(黄色と緑色の部分)がmSin3に結合した時の位置を示す。緑色の部分がNRSF/RESTの“LIML”配列であり、mS-11が模倣した部位である。図は、クラスターAがLIML配列の位置と一致することを示しており、我々の予想通りの結合が生じたことを示している。

さて、mSin3 には疎水性のクレフトがあり、NRSF/REST はそこに結合する。一方、mS-11 には長い疎水性側鎖を持っている。クラスターA に位置する mS-11 は、その長い疎水性側鎖を mSin3 のクレフトにはめ込むことで構造安定化が得られている (Fig. 2a)。図中、点線の円がその疎水性クレフトの位置を示している。なお、図では、クラスターA の中でサンプルさ

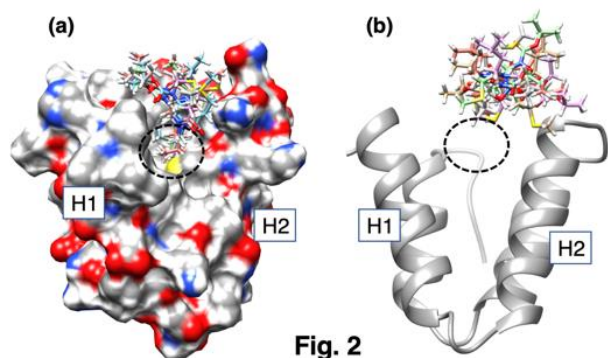


Fig. 2

れた mS-11 の構造をいくつか重ねて描いている。

一方、クラスターBに位置する mS-11 は、疎水のクレフトに mS-11 の分子のどの部分も入りこんでいない (Fig. 2b)。クラスターA が B よりも安定なのは、疎水性クレフトに mS-11 の疎水性側鎖が入り込むことが重要なのだと推測できる。これらの結果から、mS-11 は設計したコンセプトの通り NRSF/REST の”LIML”領域に結合して mSin3 に結合するのを阻害することを、計算科学的に示すことができた。

次に、mS-11 が mSin3 に接近するためのメカニズムを調べた。mS-11 は二つのリングが結合することで近似的に平面構造をとっているが、二つのリングの間に微妙な角度のずれがあり、細かく見るとアーチ型になっている (Fig. 3b)。そこで、アーチ型の凸の方向に並行

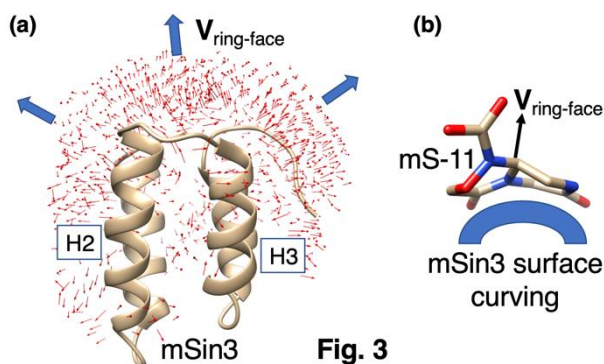


Fig. 3

なベクトル ($V_{\text{ring-face}}$) を定義した。そのベクトルを mSin3 の周囲で調べると Fig. 3a のようになる。小さ

な赤い矢印の集団が、そのベクトルの場を示す。mS-11 は、自身の微妙な凹み部分を mSin3 の分子表面に当てることで、mSin3 と結合することがわかる。

まとめ、今後の課題

上記のように、本研究は、mS-11 が mSin3 のクレフトに結合する最安定なサイトでの知見 (阻害効果) だけでなく、結合していく過程に関しても知見 (分子配向のオーダーリング) を与えた。ここで紹介したことは代表的な結果であり、より深い解析から様々な知見が得られており、本研究は現在論文にまとめているところである。これらの結果は PPI 創薬を指向した我々の α ヘリックス模倣分子創出戦略が有効である事を示しており、今後の PPI 創薬やドラッグデザインへの応用が期待できる。

一方、本研究から得られなかったこともある。ここでは詳細は述べなかったが、本研究では周期的境界条件を使って系を設定した。溶媒で満たされた周期 box の中心に mSin3 を置き、その周囲を mSin3 が運動する。この box の大きさは、mS-11 が mSin3 に対して会合・乖離を実現するには十分な大きさであった。しかし、mSin3 からの影響が完全に消えるほどには box が大きくなかった。つまり、本研究からは、結合自由エネルギー (結合状態と完全な解離状態の自由エネルギーの差) が算出できなかった。box をさらに大きく設定し、mS-11 がバルクな溶液中に浮かんでいるとみなせるほど mS-11 が mSin3 から離れた領域まで計算できれば、結合自由エネルギーが本研究から直接的に計算できる期待できる。それが今後の課題である。

共同研究者: 肥後順一 (兵庫県立大学)

小規模リソースにおける生成型要約のためのスタイル転移

Style transfer for abstractive summarization under low resource condition

利用課題責任者 田森 秀明
Hideaki Tamori

株式会社朝日新聞社メディアラボ
Media Lab, The Asahi Shimbun Company

<https://cl.asahi.com>

概要

本研究では記事本文と見出しの組の大規模な訓練データと、小規模な短文要約のデータを組み合わせて、見出し生成と要約生成を同時に学習させることで高い性能を持つ短文要約生成器を得る。短文要約のデータだけで学習した場合と比較すると、見出し生成と要約生成を同時に学習することで、15.4 ポイントの ROUGE-1 スコアの改善が得られた。

In this research project, we tried to obtain a high-performance summary generator by training headline and summary generation simultaneously. We did this by mixing large article bodies and headlines paired with small summary data. We improved the ROUGE-1 score by 15.4 points by training headline and summary generation at the same time, in comparison to using only summary data.

1 はじめに

Rushら [1] の研究以降、新聞記事コーパスを用いた生成型要約の研究が盛んに行われている。現在の生成型要約の成功は、Transformer[2] などのエンコーダ・デコーダモデル (EncDec) の発展、および何百万事例という規模の大量の訓練データに支えられている。

ところが、大量の訓練データを利用できるドメインは限られている。新聞記事コーパスに関しては、記事本文と紙面見出しの組のデータは豊富に公開されているが、記事本文と要約の組のデータは入手しづらい。例えば、朝日新聞社が公開している Japanese News Corpus (JNC) は記事の先頭 3 文と見出しの組を 1,828,231 件収録している。一方、朝日新聞社の要約サービスの一つである、ニュース速報配信サービス

(ANDES)¹では 2019 年 6 月時点までに約 3 万件の記事と要約の組みしか蓄積されていない²。高々数万件の訓練データを用い要約モデルを学習するだけでは、高品質な要約が得られない。

そこで、本研究では記事本文と見出しの組の大規模な訓練データと、小規模な記事要約のデータを組み合わせて、高品質な要約器を構築する手法を探索する。具体的には、記事からの見出し生成と要約生成を同時に学習することを、生成スタイルの転移と見なし、転移学習により高いパフォーマンスの要約生成器を獲得する。要約生成をスタイル転移として捉えることで、異なる長さの要約を異なる部署・担当者が執筆している場合など、要約の「スタイル」の差も考慮できる。既に公開している JAPANESE MULTI-LENGTH SUMMARIZATION CORPUS (JAMUL) の拡張版として、朝日新聞社が配信した新幹線向け短文要約を 30,656 件収録した JAMUL 2020 を構築し、見出し生成から要約生成へのスタイル転移の実験を行う。

結果から、短文要約のデータだけでモデルを学習した場合と比較すると、スタイル転移でモデルを学習することで、15.4 ポイントの ROUGE-1 スコアの改善が得られた。また、最善の 1 文をシステム要約としてみならず抽出型のオラクル要約に関して、ROUGE-1 で 4.8 ポイントの改善が見られた。

2 JAMUL 2020

2.1 コーパスについて

今回、人見ら [3] が提案した異なる長さの見出しが 1 記事につき 4 つ付与されたコーパス JAMUL について

¹<http://www.asahi.com/information/haishin/service>

²配信する記事を厳選していることと、人的コストの観点から要約を付与できる数に限りがある。

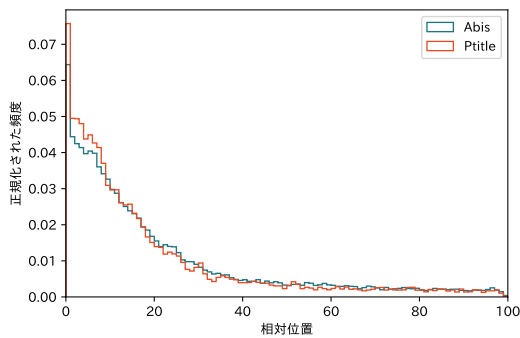


図 1: 正解要約に含まれる bi-gram に関する記事中の相対位置のヒストグラム

拡張した。具体的には、朝日新聞社が展開するニュース速報配信サービスから見出しと要約のデータを抽出し、JAMUL 2020 を構築・公開³する。

JAMUL 2020 は 2014 年 5 月から 2019 年 6 月までに朝日新聞社が展開する要約サービス ANDES で配信された 30,656 件のデータから構成される。それぞれの記事に対して最大で 5 種類の見出し・要約が付与されている。内訳は、新聞紙面向けの見出し (Print)、朝日新聞デジタル向け見出し (Large)、携帯端末向け見出し (Middle)、電光掲示板向け見出し (Short)、新幹線の電光掲示板などに配信される短文要約 (Sum) である。Print 以外に関しては、表示するデバイスやレイアウトの都合からそれぞれ異なる文字数の上限がある。上限はそれぞれ、Large が 26 文字、Middle が 13 文字、Short が 10 文字、Sum が 50 文字となっている。本研究では、Print・Large・Sum のみを用いる。

2.2 見出しと要約の比較

一般的に、見出しと要約は共に記事の重要な部分の抜粋である。このことから着目する単語や文がある程度共通していることが想定される。そこで JAMUL 2020 における Print と Sum についていくつかの分析を行った。

まず、後述する検証データ (4.1 節) における Print と Sum でどれだけ単語が重複しているのかを確認していく。具体的には、Print をシステム出力、Sum を参照要約とした際の ROUGE の F1 スコアを求めた。結果は ROUGE-1 が 42.8、ROUGE-2 が 20.8、ROUGE-L が 35.1 となり、ある程度単語が重複していることが伺える。次に、重複している単語の品詞について MeCab[4] を用いて調べていく。頻度順に名詞が 68 %、助詞が 19 %、記号が 6 % となり、それ以外については 3 % 未満であった。重複する 7 割近くの単語が名詞であり、キーワード選択の手がかりとして効果的であることが

伺える。

最後に、Print と Sum の位置バイアスについて確認する。具体的には、見出しと要約における各 bi-gram が記事中のどの位置に存在するかの頻度について、図 1 にヒストグラムで可視化した。図からは位置バイアスがほぼ一致していることが分かる。つまり、見出しと要約で着目する部分が殆ど同じであり、転移学習に効果的であることが期待できる。

3 提案手法

3.1 タスク定義

入力文 X を One-hot ベクトルからなる長さ I の系列とする。ここで、 $x_i \in 0, 1^{|V_s|}$ は、入力 X の i 番目のトークン (単語) を表す。また、 V_s は語彙であり、 $|V_s|$ は語彙 V_s に含まれる単語数を表す。以降、 $X = (x_1, \dots, x_I)$ の略記法として $x_{1:I}$ を用いる。同様に出力側の要約について、 $y_j \in 0, 1^{|V_t|}$ は長さが J の出力 Y における、 j 番目の 1-hot ベクトルと定義する。 $|V_t|$ は出力側語彙 V_t の語彙数とする。ここで、 Y は常に専用のトークンを 2 つ含むと仮定する。具体的には、 y_0 が $\langle bos \rangle$ であり、 y_{J+1} が $\langle eos \rangle$ である。また、Rush ら [1] の提案したヘッドライン生成タスクに倣い、 $I > J$ とする。また、短文要約と見出しのデータを区別して学習するためにコントロールコード c を導入する。具体的には、 x_0 に c が与えられるものとする。EncDec は以下の条件付き確率をモデル化する。

$$p(Y|X, c) = \prod_{j=1}^{J+1} p(y_j | y_{0:j-1}, X, c) \quad (1)$$

\mathcal{D} を訓練データの集合、 θ を EncDec で訓練するパラメータの集合だとする。最適なパラメータ θ' を、次の損失関数 $G(\theta)$ を \mathcal{D} 上で最小化することによって求める。

$$G(\theta) = \frac{1}{|\mathcal{D}|} \sum_{(X, Y) \in \mathcal{D}} \ell_{trg}(Y, X, c, \theta), \quad (2)$$

$$\ell_{trg}(Y, X, c, \theta) = -\log(p(Y|X, c, \theta))$$

推論時には、訓練したパラメータのもと、式 1 で定義した条件付き確率の積を最大化する系列をビーム探索する。

3.2 モデル

EncDec として生成型要約で高いパフォーマンスが報告されている Transformer[2] と BERTSUMABS[5] を用いる。その上で、見出し生成と要約生成のスタイルを分離して学習するために、2 つの観点からコントロールコードを導入する。一つ目はタスクについてのコントロールコード [6]、二つ目は長さについてのコント

³https://cl.asahi.com/api_data/jnc-jamul.html

記事: サッカーのスペイン1部リーグで13日、バルセロナのFWリオネル・メッシ(31)=アルゼンチン=が本拠地のエイバル戦で得意の左足でゴールを決め、リーグ史上初の通算400ゴールの偉業を成し遂げた...

Print(19文字): メッシがスペインリーグ400点の金字塔
Sum(48文字): サッカー・スペイン1部リーグで、バルセロナのメッシが通算400ゴール。リーグ史上初の歴代最高記録

c	ドメイン	入力
タスク	Print	@Print サッカーのスペイン...
	Sum	@Sum サッカーのスペイン...
長さ	Print	@Len19 サッカーのスペイン...
	Sum	@Len48 サッカーのスペイン...

表1: コントロールコード(c)を導入した際の入力例。

ロールコード [3] である。コントロールコードの具体例については表1にまとめた。最終的に、EncDecとコントロールコードの組み合わせとして以下の4つを用いる。

Conditional Transformer(Task) Keskarら [6] が提案したコントロールコードをTransformerに適用したモデル。コントロールコードは@Printと@Sumの2つを用いる。

Conditional Transformer(Length) 人見ら [3] が提案した手法で、コントロールコードは7~50文字を表す計44個を用いる。

Conditional BertSumAbs(Task) Liuら [5] が提案した要約モデルにコントロールコードを導入したモデル。コントロールコードは@Printと@Sumの2つを用いる。

Conditional BertSumAbs(Length) Liuら [5] が提案した要約モデルに、コントロールコードは7~50文字を表す計44個を用いる。

4 実験

4.1 データセット

JAMUL 2020 訓練データとして、2017年以前のものを用い、比較実験の都合から同一記事にSum・Largeの両方が付与されたものを抽出した。検証・評価データについては、JNCとの重複を避けるため2018年以降の組みから構成した。その後、比較実験と第2.2節の分析の都合から同一記事にSum・Large・Printの3つの要約が付与されている組みのみを抽出した。その後、ランダムに半分を抽出し検証データとし、残りを評価データとした。最終的に、訓練データは23,049件、検証データは2,321件、評価データは2,322件となった。

JNC 人見ら [3] が提案した見出しと記事の対からなる大規模なコーパス。JNCは朝日新聞の記事とPrintの対を2007年から2016年にかけて収集したコーパス

で、1,932,398件からなる。人見ら [3] の用いた前処理スクリプト⁴によって得られた1,523,468件を転移学習を目的とした訓練データとして用いる。

4.2 ベースライン

ベースラインとして、以下の6手法と提案手法を比較する。

Extractive Oracle 入力記事の各文からROUGE-1とROUGE-2の合計が最大になる1文⁵をシステム出力とみなす。

LEAD-1 入力記事の第1文をシステム出力とする。

Transformer(Sum) Sumのみで訓練されたTransformer。

Transformer(Sum+Large) Paulusら [7] が導入した設定で、入力として見出し・特殊トークン・記事を結合したものを使用し、訓練されたTransformer。今回は見出しとしてLargeを用いる。

Transformer+SP-token 人見ら [3] の手法で、入力の先頭に長さの情報を持ったスペシャルトークを追加する。なお、JNCので訓練したケース。

BertSumAbs Liuら [5] が提案した生成型要約モデル。SumのみでFine-tuningする。

4.3 実装の詳細

Transformerの実装としてFairseq⁶ [8] を用いた。各種パラメータについては学習率を除き、Vaswaniら [2] のWMT 2014のEnglish-Germanにおけるbase modelのものを用いた。学習率については0.0007⁷に変更した。また、単語埋め込みベクトルの学習に用いる重み行列については入力側・出力側で別々のものを用いた。BERTの訓練済みモデルとしては日本語のWikipediaで事前学習を行っているJapaneseBERT⁸を用いた。BERTSUMの実装は著者実装⁹を用い、コントロールコードの追加には👉 Transformers¹⁰を用いた¹¹。また、BERTSUMABS以外の実験における語彙については、SentencePiece¹²で構築した。merge operationを8000と定め、入力文、出力文を結合したコーパスから語彙の構築を行った。長さについてのコントロールコードを用いたモデルについては、評価時に48文字を表現したコントロールコードを用いる。推論時のビーム幅は5とした。また、生成結果・正解要約共

⁴<https://github.com/asahi-research/Gingo>

⁵<https://github.com/nlpyang/BertSum> を参考にした。

⁶<https://github.com/pytorch/fairseq>

⁷WMT 2014のEnglish-GermanにおけるVaswaniらのスコアをFairseq上で再現した時の学習率。

⁸<https://github.com/cl-tohoku/bert-japanese>

⁹<https://github.com/nlpyang/PreSumm>

¹⁰<https://github.com/huggingface/transformers>

¹¹追加されたコントロールコードの重みベクトル・バイアスベクトルの初期化についてはDevlinら [9] に従った。

¹²<https://github.com/google/sentencepiece>

	事前学習データ	訓練データ	R-1	R-2	R-L	ALD
Conditional Transformer(Task)	-	JNC, Sum	57.5	32.9	46.3	2.6
Conditional Transformer(Length)	-	JNC, Sum	57.5	32.8	46.4	1.4
Conditional BERTSUMABS(Task)	日本語 Wikipedia	JNC, Sum	52.5	28.7	37.3	19.2
Conditional BERTSUMABS(Length)	日本語 Wikipedia	JNC, Sum	53.4	28.6	38.3	8.4
Extractive Oracle	-	-	52.7	28.7	40.4	24.0
LEAD-1	-	-	47.5	23.8	36.3	28.9
Transformer(Sum)	-	Sum	32.1	7.7	23.9	5.7
Transformer(Sum+Large)	-	Sum, Large	32.9	7.7	23.8	3.8
Transformer+SP-token	-	JNC	42.5	18.0	30.7	2.1
BERTSUMABS	日本語 Wikipedia	Sum	45.0	21.8	32.8	24.0

表 2: 実験結果. 上から 4 段目までが提案手法.

に MeCab [4] により分かち書きをして ROUGE を求めた.

4.4 実験結果

単語の被覆率による評価 F 値ベースの ROUGE-1, ROUGE-2, ROUGE-L を用いた. それぞれの略記法として, R-1, R-2, R-L を用いる. 表 2 が実験結果である. ROUGE のスコアにおいて, 提案手法である Conditional Transformer が, 最も高いパフォーマンスとなった.

そこで Conditional Transformer とその他手法について比較して行く. まず, 抽出型のアプローチである Extractive Oracle・LEAD-1 の 2 手法と比べていく. 提案手法が Extractive Oracle に対して R-1 で 4.8 ポイント, LEAD-1 に対して R-1 で 10 ポイントの改善が確認された. 次に, Sum だけで訓練されたモデルと比較すると, 提案手法が R-1 で 25.4 ポイントの改善が見られた. ヘッドラインを追加の入力として与えるケースについてもほぼ同様の結果となった. Transformer+SP-token と比較していくと, R-1 で 15 ポイントの改善が確認された. このことから, 単純に見出しのみを学習データに用いるだけでは不十分であることが分かる. 続いて, BERTSUMABS と比較すると, 提案手法は R-1 で 12.5 ポイントの改善が見られた.

長さの制約についての評価 生成された要約が長さの制約をどれくらい正確に守れているかを評価するために以下の式から *Average Length Difference(ALD)* を求めた.

$$ALD = \frac{1}{n} \sum_{i=1}^n |l_i - len| \quad (3)$$

ここで, n は生成された要約の数を, l_i は生成された要約の文字数を, len は正解要約の文字数を表す. スコアの解釈としては, 正解要約から平均何文字ずれているかを表現している.

表 2 を ALD について見てみると, 抽出型要約である Extractive Oracle・LEAD-1 については平均 24~29 文字ずれているのに比べ, Conditional Transformer は平均 1~3 文字程度のずれであった. 厳密な長さの制約がある場合, 制約に合わせて生成ができる生成型

要約の方が本タスクにおいて有効であるといえる.

次に, コントロールコードの種類による影響を見ていく. 具体的に, タスクについて制御するケースと長さについて制御するケースについて比較する. 結果から, 長さについて制御した方がより厳密に制約を守れていることが分かった. これは, タスクについてのコントロールコードが長さについて幅を持って学習しているのに対し, 長さについてのコントロールコードは 1 文字ずつ学習できるためである. 一方, タスクについて制御を行うと, 見出しにしか出現しないスペースなどの表現が生成された要約に全く現れなくなるなど, ライティングスタイルについての改善が確認された.

また, Conditional BERTSUMABS は Conditional Transformer に比べ正確に出力長をコントロールできないことが分かった. これは事前学習時に長さの情報について学習できていないことが原因であると考えられる.

5 関連研究

要約において, 長さの制約を考慮したコーパスは JAMUL[3] と DUC 2004¹³がある. しかし, JAMUL は 1,489 件, DUC 2004 は 500 件のデータで構成されており, EncDec の枠組みで訓練・検証・評価の 3 つに分割することが困難であった. また生成型要約の小規模データに対する取り組みとして, Tilk ら [10] は見出し生成タスクに取り組んだ. 約 300,000 件のデータサイズを小規模データと定義し, 事前学習によるパフォーマンスの改善を提案した. Parida ら [11] は単一文書要約において, 90,000 件のデータサイズを小規模データと定義し, 擬似データによるデータ拡張でパフォーマンスの改善に取り組んだ. ただし, 追加の要約データが豊富にあることが前提であるため, 今回の実験設定とは異なる. これらの先行研究に比べると, 我々が扱う小規模データは約 23,000 件であるという点でより小規模なデータに取り組んでいる.

データサイズについて, 我々より小規模で実験を行なった Zhang ら [12] の研究では, 英語の単一文書要約のコーパス CNN/Dailymail[13] において 1,000 件

¹³<https://duc.nist.gov/duc2004/>

程度で高いパフォーマンスを出せることを示している。しかし、事前学習に 1.5B のデータが必要であり、我々は 1.5M を Style transfer に用いているため、追加で用いるデータが大幅に小規模である。

6 おわりに

本研究では、小規模リソースを用いた要約生成におけるパフォーマンスの改善に取り組んだ。提案手法は、コントロールコードによるスタイル分離することで、見出し生成と要約生成を同時に学習し、スタイル転移によるパフォーマンスの改善を行った。今回構築したコーパス、JAMUL 2020 の評価データにおいては抽出型のオラクル要約に対して R-1 で 4.8 ポイント、その他既存手法に対しても大幅な改善が確認された。

謝辞

全ての計算は、東京工業大学のスパコン TSUB-AME3.0 を用いて行った。

参考文献

- [1] Alexander M. Rush, Sumit Chopra, and Jason Weston. A neural attention model for abstractive sentence summarization. In *In EMNLP*, pp. 379–389, 2015.
- [2] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *NIPS*, pp. 5998–6008, 2017.
- [3] Yuta Hitomi, Yuya Taguchi, Hideaki Tamori, Ko Kikuta, Jiro Nishitoba, Naoaki Okazaki, Kentaro Inui, and Manabu Okumura. A large-scale multi-length headline corpus for analyzing length-constrained headline generation model evaluation. In *INLG*, 2019.
- [4] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. Applying conditional random fields to japanese morphological analysis. In *EMNLP*, pp. 230–237, 2004.
- [5] Yang Liu and Mirella Lapata. Text summarization with pretrained encoders. In *EMNLP-IJCNLP*, pp. 3728–3738, 2019.
- [6] Nitish Keskar, Bryan McCann, Lav Varshney, Caiming Xiong, and Richard Socher. Ctrl: A conditional transformer language model for controllable generation. In *arXiv preprint*, 2019.
- [7] Romain Paulus, Caiming Xiong, and Richard Socher. A deep reinforced model for abstractive summarization. In *ICLR*, 2018.
- [8] Myle Ott, Sergey Edunov, Alexei Baevski, Angela Fan, Sam Gross, Nathan Ng, David Grangier, and Michael Auli. fairseq: A fast, extensible toolkit for sequence modeling. In *NAACL-HLT*, pp. 48–53, 2019.
- [9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *NAACL*, pp. 4171–4186, 2019.
- [10] Ottokar Tilk and Tanel Alumäe. Low-resource neural headline generation. In *NEWSUM*, pp. 20–26, 2017.
- [11] Shantipriya Parida and Petr Motlicek. Abstract text summarization: A low resource challenge. In *EMNLP-IJCNLP*, pp. 5993–5997, 2019.
- [12] Jingqing Zhang, Yao Zhao, Mohammad Saleh, and Peter J. Liu. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. *ArXiv*, 2019.
- [13] Karl Moritz Hermann, Tomáš Kočiský, Edward Grefenstette, Lasse Espeholt, Will Kay, Mustafa Suleyman, and Phil Blunsom.

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 ハイドロゲル保水性発現機構の解明

英文: Molecular analysis of mechanism of water retention in hydrogel

利用課題責任者 松林伸幸

Nobuyuki Matubayasi

所属 大阪大学 基礎工学研究科

Affiliation Division of Chemical Engineering, Graduate School of Engineering Science, Osaka University
URL

邦文抄録(300字程度)

三次元網目構造の内部に水を蓄えたハイドロゲルは、加圧に対して水を離さない優れた保水性を示すことから高吸水性製品に幅広く応用されている。このような保水性を示す分子論的機構には未解明な部分が多い。そこで、本研究課題では分子動力学シミュレーションを用いた保水性発現機構の解明を目指す。自由エネルギー計算などの熱力学的安定性や水素結合ネットワーク構造などの静的性質の解析を行う。

英文抄録(100 words 程度)

In order to investigate the molecular mechanism by which hydrogels retain water, this research project aims to elucidate the mechanism of water retention using all-atom molecular dynamics simulation. We analyze thermodynamic stability such as free energy calculation and static properties such as hydrogen bond network structure.

背景と目的

ハイドロゲルは、構成する高分子の三次元網目構造によって多量の水を保持することができる材料であり、その特性である吸水性を活かして、オムツなどに利用されている。ゲル中の網目構造は、水素結合などの非共有結合に起因する物理架橋、または共有結合による化学架橋によって秩序化されている。物理・化学架橋双方でゲル構造を形成することが知られている poly vinyl alcohol (PVA) ハイドロゲルは、高含水率・高強度という性質に加え生体適合性の高さから人工血管などとして生物医学分野で広く使用されている。

ハイドロゲルの構造や物性は、含水率や PVA 中の親水基の水素結合のような分子レベルの相互作用に大きく依存する。近年では、網目高分子と相互作用している含有水の運動性を測定する実験も可能になっているが、分子レベルの構造を詳細に観察することは難しい。先行研究では、分子動力学(MD)シミュレーションを用いて PVA ハイドロゲル内の水分子の運動性や PVA ゲルの弾性率などからゲルの性質に迫る解析が行われてきた。しかし、分子間構造や含有水に関わるエネルギー的な議論はされていない。そこで本プロジェクトでは、分子レベルの相互作用を露わに取り入れた MD

シミュレーションとエネルギー表示理論を組み合わせることで、PVA ハイドロゲル中の水分子のエネルギー解析を行う。また、構造情報の解析からハイドロゲル中の分子レベル情報を明らかにした。

概要

重合度 200 の PVA を 50 本と PVA のモル分率が 20 wt%となるように水分子をランダムに配置することで初期構造を作成し、NPT アンサンブルで 300 K と 200 K への降温と昇温を繰り返すことでゲルをモデリングした。これは、実験における凍結融解法に準じた手法である。次いで、ゲルを取り出し、生じた空隙に水分子を挿入し、NPT アンサンブルで計算を行った。PVA の力場には GAFF を採用し、MD 計算は Gromacs 2016.5 を用いた。水分子の動径分布関数(RDF)、そして PVA と水の混合溶媒にさらに水分子を加えたときの自由エネルギー変化をエネルギー表示法による溶媒和自由エネルギー計算から求めた。

結果および考察

Fig.1.(a)に PVA ハイドロゲルの膨潤度を、(b)に真空から PVA ハイドロゲル中への水分子の移行自由エネルギー(溶媒和自由エネルギー)を示す。膨潤し始める

領域で、PVA の寄与が減少し、水の寄与が増大する。系全体の溶媒和自由エネルギーは純水中の値に近く、変化は小さい。つまり膨潤領域では純水中並みの安定度で水が溶解する。

次に、水の周りの配位数を Fig.2 に示す。ここで配位数とは、動径分布関数を第 1 ピーク範囲で積分することで得られる。ある水分子から見た水分子の個数(●)、PVA のヒドロキシ基の個数(▲)を示す。含水量の増加によってヒドロキシ基の配位数は減少するのに対し、同時に水の配位数は増加する。つまり、含水量が増加するほど水分子同士が凝集することが分かった。

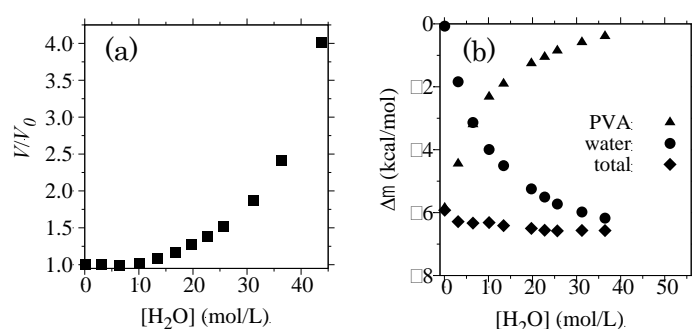


Fig. 1. (a) Degree of swelling in PVA hydrogel. (b) The solvation free energy $\Delta\mu$ of water molecule.

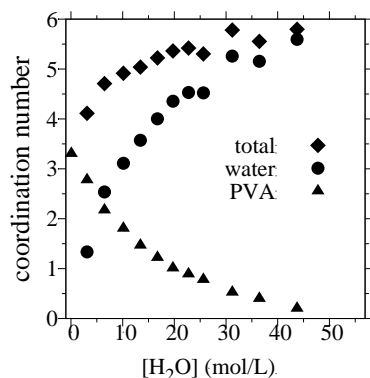


Fig. 2. Coordination number of water and PVA around

まとめ、今後の課題

全原子分子動力学シミュレーションとエネルギー表示理論を組み合わせることで PVA ハイドロゲル中の原子レベル情報を取り入れたエネルギー情報と構造情報の解析を行った。膨潤領域では純水中並みの安定度で水が溶解すること、含水量の増加に伴い水分子同士の凝集が見られることがわかった。今後の課題として、緩和弾性率や水分子の拡散係数といったゲルのダイナミクスと関連した量の解析を行う予定である。

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 HPC を利用した自然言語処理技術の研究

英文: High Performance Computing for Natural Language Processing Technology Research

利用課題責任者

鳥澤 健太郎

所属

国立研究開発法人情報通信研究機構 ユニバーサルコミュニケーション研究所

データ駆動知能システム研究センター

<http://www2.nict.go.jp/direct/>

邦文抄録(300 字程度)

情報通信研究機構では、対話システム WEKDA を初めとする複数の自然言語処理アプリケーションで、深層学習による様々な分析を行っている。近年発表されている大規模なニューラルネットワークは、一枚の GPU に学習パラメータを収容することが困難であることから、研究代表者らはニューラルネットワークを分割し複数の GPU に格納して分散計算を行う、モデルパラレルと呼ばれる分散学習を行うフレームワーク RaNNC を開発してきた。本課題では RaNNC を拡張し、より複雑なネットワークの学習を可能とすると共に、データパラレルとのハイブリッド学習を実現した。BERT-Large を 5 倍以上の規模に拡大したネットワークを 256 枚の GPU を用いて事前学習し、BERT-Large より優れた学習性能が得られることを確認した。

英文抄録(100 words 程度)

NICT has been developing language processing applications including WEKDA, which intensively use deep learning to analyze a large-scale Web archive. Since some modern networks have billion-scale parameters and do not fit to single GPU, we have been developing RaNNC, which is a deep learning framework for model parallelism. RaNNC divides a neural network into smaller parts and distributes them onto multiple GPUs. In this work, we extended RaNNC to learn more complicated networks with the hybrid data/model parallelism. We trained a BERT network that has five times more parameters than the original BERT using RaNNC. Our pre-training with 256 GPUs showed that the scaled-up BERT results in a better training loss than BERT-Large.

Keywords: 自然言語処理, 大規模情報分析, テキスト分析, モデルパラレル

背景と目的

近年、深層学習で用いられるニューラルネットワークの大規模化が進んでいる。深層学習では GPU 等のアクセラレータを用いるのが一般的であるが、メモリサイズの制約から、巨大なネットワークの学習は容易ではない。そこで計算の並列化と、GPU1 枚あたりの必要メモリ削減の目的のため、データパラレルと呼ばれる、ミニバッチをさらに分割して学習する方式が広く用いられている。しかし、データパラレルではニューラルネットワーク自体が各 GPU に複製されるため、極めて多数のパラメータを持つネットワークは、GPU のメモリに収めることができず、データパラレルでは学習できない。そこで研究代表者らは、ネットワークを分割するモデルパラレルを自動的に行うフレームワーク RaNNC (Rapid Neural Network Connector) を開発してきた。

本課題では、RaNNC を拡張し、より複雑なネットワークへの対応やスケーラビリティの強化に取り組み、BERT-Large[Devlin 2018]の 5 倍以上のパラメータ数を持つネットワークを、256 枚の GPU で学習することに成功した。

概要

本課題では、モデルパラレルのための深層学習フレームワーク RaNNC の機能強化を行い、TSUBAME 上で動作検証や性能測定を行った。RaNNC は、既存の深層学習フレームワーク(現在は PyTorch¹に対応)が出力する計算グラフを複数の部分グラフに分割し、それぞれの部分グラフを異なる GPU に配置して計算する。各 GPU 上では、既存の深層学習フレームワーク

¹ <https://pytorch.org/>

の計算エンジンを用いて部分グラフを計算するが、入出力を MPI によって通信することにより、部分グラフを結合する。課題開始時点で、AlexNet などの基本的かつ小規模なニューラルネットワークを、数枚程度の GPU 上での学習できる段階まで実現済みであった。課題開始後の進捗は、以下の通りである。

(1) 複雑なモデルへの対応

各 GPU に配置された部分グラフは、PyTorch の計算エンジンを用いて計算を行うが、本来は部分グラフの計算が想定されておらず、グラフの構造によっては、様々な問題が生じる。一例として、計算グラフへの入力値が、複数の計算オペレータから共有される場合の勾配計算が挙げられる。逆伝播において、共有される入力値にはそれぞれの計算オペレータ由来の勾配が加算されるが、PyTorch の計算エンジンはこのようなケースを想定しないため、加算の操作が競合し、正しく計算できない。そのため RaNNC では独自に排他処理を実装している。その他、通信タイミングの制御や複雑なデータ型の通信などを強化し、BERT において、PyTorch のみを用いる場合と全く等価な結果を得られることを確認した。

(2) データパラレルとのハイブリッド化

本課題では、極めて大規模なネットワークの学習を目的としているため、モデルパラレルのみでは、学習に要する時間が極めて長くなり、実用性に乏しい。そのため、モデルパラレル・データパラレルのハイブリッドで、GPU 数百枚規模の学習を可能とするよう拡張した。

本課題では、こうした機能の検証及び速度測定に TSUBAME を使用した。

結果および考察

以上に述べた成果を適用し、BERT の学習を試みた。異なる作成者による PyTorch を用いた BERT モデル²³を対象に、ニューラルネットワークを定義する部分の実装を改変することなく、モデルパラレル・データパラレルのハイブリッドでの学習が実現できた。モデルパラレルを可能とする既存のフレームワークとして、Mesh-TensorFlow[Shazeer 2018]、Megatron-LM [Shoeybi 2019]などがあるが、これらを適用するには、既存のモデルの大幅な改修が必要となる。無改変でのモ

デルパラレルを実現した研究やフレームワークは、研究代表者らの知る限り他にない。

また、BERT-Large の隠れ層サイズを 1920(原論文では 1024)、総数を 54 層(同 24 層)に大規模化したネットワーク(パラメータ総数にして BERT-Large の約 5 倍以上)を、GPU256 枚を用いて事前学習を行った(本実験は課題代表者らの計算設備で実行した)。計算資源確保の問題から、学習途中で中断しているが、図 1 に示すように、BERT-Large と比べて顕著に学習誤差が減少していることが確認できた。

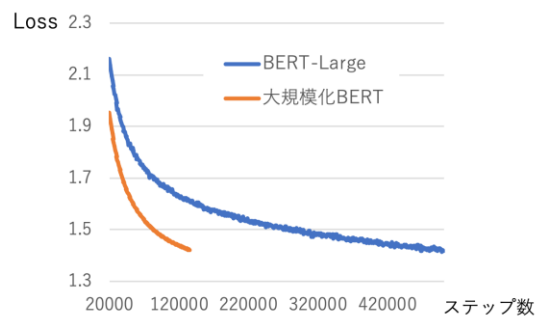


図 1: 大規模化 BERT の事前学習

まとめ、今後の課題

独自に開発したモデルパラレル深層学習フレームワーク RaNNC の強化を行い、BERT-Large の 5 倍以上の規模のネットワークを学習可能とした。今後、GPU 利用効率を向上するなどの更なる改善を進めつつ、T5[Raffel 2019]を初めとする、より巨大なネットワークの学習を進める。

参考文献

- [Devlin 2018] Jacob Devlin et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2019)*, pp. 4171–4186, 2018.
- [Shazeer 2018] N. Shazeer, Y. Cheng, N. Parmar, D. Tran, A. Vaswani, P. Koanantakool, P. Hawkins, H. Lee, M. Hong, C. Young, R. Sepassi, and B. Hechtman. Mesh-TensorFlow: Deep learning for supercomputers. *In Neural Information Processing Systems*, 2018.
- [Shoeybi 2019] Mohammad Shoeybi and Mostofa Patwary and Raul Puri and Patrick LeGresley and Jared Casper and Bryan Catanzaro. Megatron-LM: Training Multi-Billion Parameter Language Models Using Model Parallelism. arXiv: 1909.08053, 2019.
- [Raffel 2019] Colin Raffel et al., Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer, arXiv: 1910.10683, 2019.

² <https://github.com/NVIDIA/DeepLearningExamples/>

³ <https://github.com/huggingface/transformers>

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 マルチ GPU による心血管系の血流の数値シミュレーション
 英文: GPU-based numerical simulation of cardiovascular blood flow

利用課題責任者 水藤寛
 Hiroshi Suito

所属 東北大学 材料科学高等研究所
 Affiliation Advanced Institute for Materials Research, Tohoku University
 URL <https://www.wpi-aimr.tohoku.ac.jp/jp/index.html>

邦文抄録

安定化有限要素法は、複雑な境界形状を持つ領域での 3 次元非圧縮性 Navier-Stokes 方程式の数値計算に広く用いられているが、その GPU による高速化には、実装上いまだ困難な点が多い。そこで、本研究課題では、マルチ GPU による大規模計算が行える TSUBAME の計算環境を利用し、安定化有限要素法による 3 次元非圧縮性 Navier-Stokes 方程式のマルチ GPU 並列解法を構築し、血流などの体内流れについての評価を行った。

英文抄録

A numerical method for simulating blood flow is presented. The Navier-Stokes equation is used as a governing equation and discretized using the streamline upwind Petrov-Galerkin(SUPG)/pressure stabilizing Petrov-Galerkin(PSPG) stabilized finite element method(FEM). The system of linear equation was solved iteratively at each time step via the generalized product bi-conjugate gradient (GPBi-CG) algorithm. We developed a multi-GPU implementation of the GPBi-CG.

Keywords: Navier-Stokes equation, SUPG/PSPG, FEM, multi-GPU

背景と目的

循環器系における血流評価においては、臓器内や毛細血管床など多孔性媒質中の流れの評価も重要となる。そこで本研究ではそのような状況のモデルとして微視的な流路を設定し、そこでの流れの解析を GPU を用いて高速に行う手法を構築することとした。なお本研究は主に、東北大学材料科学高等研究所の Viet Q.H. Huynh、及び研究協力者である埼玉県環境科学国際センターの鈴木和将によって実施された。

近年、安定化有限要素法は、複雑な境界形状を持つ領域での 3 次元圧縮性 Navier-Stokes 方程式の数値計算に広く用いられているが、その GPU による高速化には、実装上いまだ困難な点が多い。本研究では、マルチ GPU による大規模計算が行える TSUBAME の計算環境を利用し、安定化有限要素法による 3 次元非圧縮性 Navier-Stokes 方程式のマルチ GPU 並列解法を構築し、生体内流れの数値シミュレーションを行うことを目的とした。特に微細な組織内での幾何形状の特徴とマクロな流れの構造と

の関係性に着目して研究を行った。

概要

単純化した系として、三次元空間にランダムに球を配置した計算モデル(多孔性媒質、図 1)を用いた。球径を数段階に変化させたモデルを作成した。

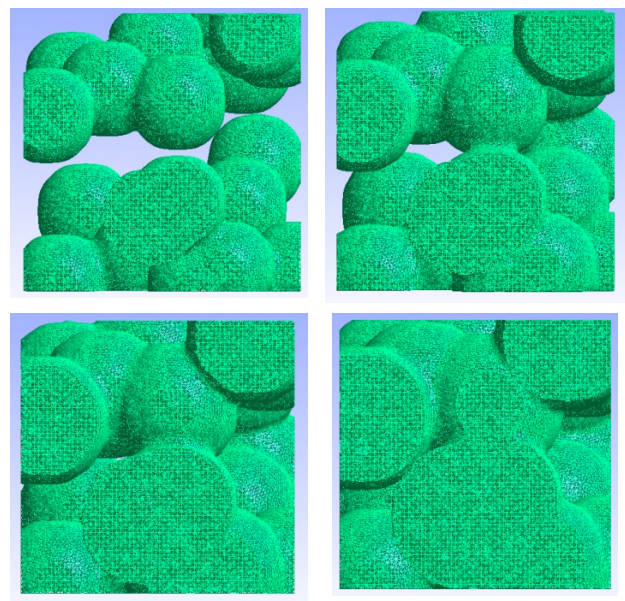


図 1 多孔性媒質モデル (R=11, 13, 15, 17)

支配方程式には、非圧縮性 Navier-Stokes 方程式、連続の式を用いた。

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{1}{Re} \frac{\partial}{\partial x_j} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \text{ in } \Omega, (1)$$

$$\frac{\partial u_i}{\partial x_i} = 0 \text{ in } \Omega. (2)$$

ここで、 u は流速、 p は圧力、 Re は Reynolds 数、 Ω は計算領域である。

上記(1),(2)の支配方程式に対して、安定化有限要素法(SUPG/PSPG 法)を適用すると以下のように弱形式が導かれる。

$$\begin{aligned} & \int_{\Omega} w_i \left(\frac{\partial u_i}{\partial t} + \bar{u}_j \frac{\partial u_i}{\partial x_j} \right) d\Omega - \int_{\Omega} \frac{\partial w_i}{\partial x_i} \\ & + \int_{\Omega} \frac{1}{Re} \frac{\partial w_i}{\partial x_j} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \\ & + \sum_{e=1}^{n_{el}} \int_{\Omega} \left(\beta \bar{u}_k \frac{\partial w_i}{\partial x_k} \right) \left(\frac{\partial u_i}{\partial t} + \bar{u}_j \frac{\partial u_i}{\partial x_j} + \frac{\partial p}{\partial x_i} \right) d\Omega = 0, (3) \end{aligned}$$

$$\begin{aligned} & \int_{\Omega} q \frac{\partial u_i}{\partial x_i} d\Omega \\ & + \sum_{e=1}^{n_{el}} \int_{\Omega} \left(\beta \frac{\partial q}{\partial x_i} \right) \left(\frac{\partial u_i}{\partial t} + \bar{u}_j \frac{\partial u_i}{\partial x_j} + \frac{\partial p}{\partial x_i} \right) d\Omega = 0, (4) \end{aligned}$$

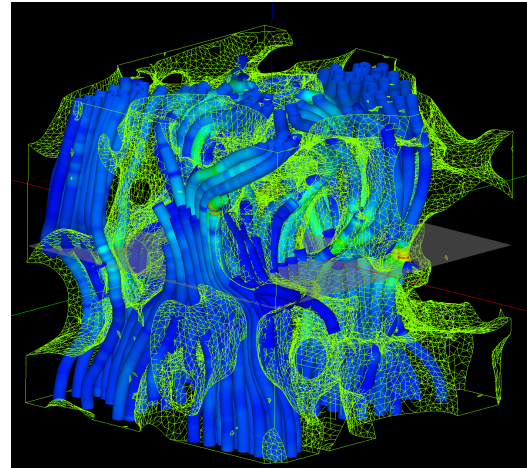
ここで、 w_i, q は、それぞれ式(1),(2)に対する Galerkin 項の重み関数である。また、 \bar{u}_i は移流速度を表し、 β は安定化パラメータである。

これら(3),(4)式に対して、空間の離散化要素には、四面体 1 次要素を用い、時間方向の離散化には、Crank-Nicolson 法を適用した。なお、最終的に得られる連立一次方程式の解法には GPBi-CG 法を適用した。本研究では、計算の高速化を達成するために、NVIDIA 社の統合開発環境である CUDA と cuBLAS, cuSPRASE 等のライブラリを用いて GPBi-CG の並列化を行い、GPU スパコンを利用して計算を実施した。

結果および考察

多孔性媒質モデルに対して、数値シミュレーションにより流れ場計算を実施した。流れの状況を流線として図化した結果を図 2 に示す。図の上部が流入面、下部が流出面であり、流入面は、一様流を与えた Dirichlet 条件である。図に示されるように、モデル領域の球が結合する箇所や分岐部で大きな

流れの変化がみられた。また、球径を変化させて数値シミュレーションを行った結果、径が大きくなるに



たが、流路が狭くなり、流速が速くなった。

図 2 多孔性媒質内流れの様子(流線)

そこで、球径の違いによる流れの影響を定量的に評価するため、流入境界と流出境界の間の圧力を変化させ、数値解析を行った。図 3 に圧力差と流量の関係を示す。球径が大きくなるほど、圧力損失が大きくなった。また、調べた範囲では、圧力差と流量の間に線形関係が成り立っていることが示された。

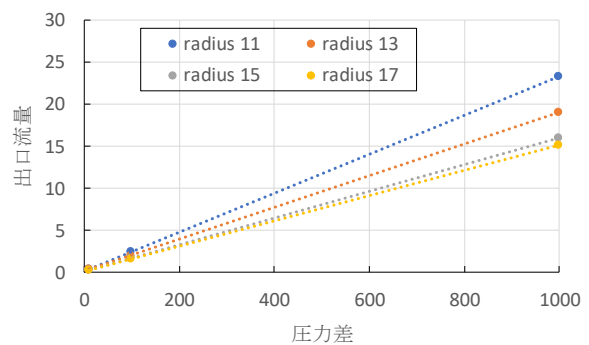


図 3 圧力差と出口流量との関係

まとめ、今後の課題

本研究の発展として形状の幾何学的な取り扱いをさらに進め、流れと流路形状の関係を明らかにする数理モデルの構築を目指すことが必要であると考えられる。

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

第一原理計算による電池・触媒メカニズム解明と新物質探索
Materials science of batteries and catalysts via first principles calculation

館山 佳尚
Yoshitaka Tateyama

物質・材料研究機構
National Institute for Materials Science
<https://www.nims.go.jp/group/cs/>

邦文抄録

蓄電池・触媒の多くは電極および電解質(液)から構成され、電極—電解質界面における化学反応および酸化還元反応が最も重要な素過程となっている。しかしながら埋もれた界面のその場観察はいまだに難しく、界面過程の微視的機構は不明な点が多い。我々はこのような界面過程について電子・原子状態を高精度に扱うことのできる第一原理計算を用いた理論計算研究を行ってきた。本利用課題では TSUBAME を利用して、全固体電池の酸化物電極—硫化物電解質界面状態に対する第一原理計算研究を遂行し、その微視的構造の解明を行った。その結果、硫化物電解質では界面相互拡散が十分起きることが示された。

英文抄録(100 words 程度)

We have been working on reactions around buried interfaces between electrolyte and electrolyte in battery and catalyst by means of density functional theory based first-principles calculations with certain accuracy. In this project, we set up the first-principles calculation scheme on TSUBAME, and investigated microscopic structures of oxide cathode – sulfide electrolyte interfaces in all-solid-state batteries. It then turned out that the mutual diffusion across the interfaces is highly possible.

Keywords: 第一原理計算、蓄電池、触媒、表面・界面

背景と目的

蓄電池・触媒の多くは電極—電解液界面における化学反応および酸化還元反応が最も重要な素過程となっている。しかしこれらの界面過程の原子スケール描写は実験的にもいまだ難しく、予言性の高い理論計算のニーズが非常に高い。実際に当研究室では、第一原理計算解析をもとに電池・触媒反応における界面の微視的過程について予測・提案を行ってきており、それが実験研究の動機・駆動力となるケースを何度か経験してきた。また当研究室では「京」向けに開発した第一原理サンプリングコード stat-CPMD を始め、第一原理計算とより広範囲な構造・反応サンプリングを可能にするモジュール開発も進めている。

このような背景の下、本プロジェクトでは TSUBAME において、これらの計算プログラムの高効率実行を行い、電池・触媒系の界面過程に関する理論計算メカニズム解析をさらに発展的に進めることを目的とする。

概要

全固体電池などの蓄電固体デバイス内に存在する異なる材料間のヘテロ固固界面の電子・イオン状態を高精度・高効率に解析可能な計算手法の開発に最近成功した。この第一原理計算手法を用いて、全固体電池の固固界面の安定な微視的構造の探索を実行し、界面相互拡散の可能性について明らかにした。

結果および考察

本研究では、異なる材料間のヘテロ固固界面における原子・イオンの格子不整合、集団緩和、局所緩和に関するあらゆる最適化と、高効率構造予測計算手法 CALYPSO 法の界面構造探索への適用、電子状態を高精度に取り扱える密度汎関数理論(DFT)計算を全て組み合わせることで開発したヘテロ固固界面の高精度シミュレーション手法を用いて、硫化物系全固体電池の酸化物正極—硫化物電解質界面の微視的構造探索を実施した。その結果、硫化物電解質で

は界面相互拡散が十分起きうることが示された。

まとめ、今後の課題

当研究室で使用している第一原理計算プログラムを TSUBAME 上で実行可能な状態にし、全固体電池界面に関する計算を実行した。

今後も第一原理計算プログラムの TSUBAME 上での運用を拡充し、電池・触媒メカニズムに関するさらなる第一原理計算研究を実行していく予定である。

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 ボルテックスジェネレータ周りの熱流動直接数値解析

英文: Direct numerical simulation of turbulent heat transfer over a surface with vortex generators

利用課題責任者: 桑田 祐丞

First name Surname: Yusuke Kuwata

所属: 大阪府立大学

Affiliation: Osaka Prefecture University

URL: <http://www2.me.osakafu-u.ac.jp/htlab/>

邦文抄録(300 字程度)

ボルテックスジェネレータを代表とする人口粗さによる運動量・熱輸送の増大効果を調べるために、半球粗面乱流熱伝達の直接数値解析を行った。解析の結果、粗さによって運動量・熱輸送ともに増大するが、運動量輸送の増大効果は熱輸送の増大効果に比べて大きくなることが明らかになった。この運動量・熱輸送の非相似性は粗さによる運動量輸送の増大効果が強く表れるほど顕著にみられた。半球個数が一定の場合、スパン方向の半球間隔が短い粗面において、非相似性が強く表れるいっぽうで、主流方向の間隔の短い粗面では非相似性は弱く、少ない摩擦抵抗増で効率的に熱輸送効果を高められることが明らかになった。

英文抄録(100 words 程度)

Direct numerical simulations of turbulent heat transfer over surfaces with hemispheres are conducted to discuss the augmentation of the momentum and heat transfer due to the wall-roughness. The results show that the momentum and heat transfer is enhanced due to the wall-roughness; however, the enhancement of the momentum transfer is larger than that of the heat transfer. This dissimilarity is found to be more substantial as the augmentation of the momentum transfer increases. When the roughness density is fixed, the dissimilarity is substantial for the surface with narrower spanwise spacing between the neighboring spheres while it is less substantial for the surface with narrower streamwise spacing.

Keywords: vortex generator, lattice Boltzmann, heat transfer, turbulence, direct numerical simulation

背景と目的

ボルテックスジェネレータを始めとする人口粗さは運動量や熱輸送を活発化させることが知られており、近年ではボルテックスジェネレータを用いた積極的な熱流動制御が盛んに行われている。しかしボルテックスジェネレータの形状・配置には多くの組み合わせが存在し、運動量や熱輸送の増大効果に多大な影響を与えるが、実験的な形状・配置の最適化には多大なコストを要するうえに、ボルテックスジェネレータをすべて格子解像してシミュレーションを行うのも工学的には現実的とは言えない。そこで本研究では、複数 GPU 並列された格子ボルツマン法コードを用いて、比較的少ない解析コストでボルテックスジェネレータを模した半球粗面の直接的な熱流動解析を行う。半球の個数や配置を系統的に変えた解析を行うことで、ボルテックスジェネレータの効率的な設置方

法を模索するとともに、熱輸送や運動量輸送の増大メカニズムを明らかにする。本解析によって、半球個数が一定の場合、スパン方向の半球間隔が短い粗面では運動量輸送が熱輸送に比べて大きく増大することが分かった。いっぽうで、主流方向の間隔の短い粗面では、少ない摩擦抵抗増で効率的に熱輸送効果を高められることが明らかになった。

概要

近年の数値解析・実験技術の進化により、高度な流体制御が可能になりつつある。その1つであるボルテックスジェネレータは、航空機・自動車などの空力性能の向上のために使用されるほか、内燃機関のシリンダー内や熱交換器内部の熱流動を制御する方法としての応用が期待されている。工学的には、

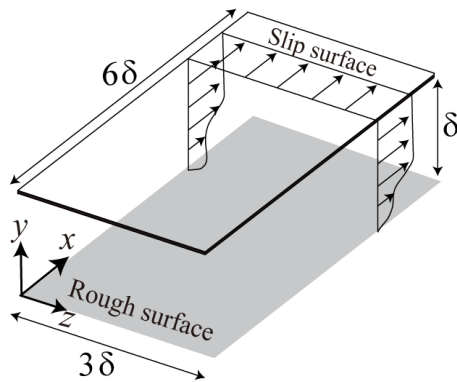


図1 粗面オープンチャンネル流れの解析系

乱流モデルを用いて、複数のボルテックスジェネレータの並ぶ壁面での熱流動を低コストに予測することが求められるが、ボルテックスジェネレータの配置や個数が熱流動場に与える影響などに関しては十分に理解されているとは言えない。本研究では、格子ボルツマン法を用いて、ボルテックスジェネレータを模した半球粗面乱流の熱流動解析を行う。解析系は図1に示すように、2次元平行平板チャンネルの半分領域となっており、下壁には半球を千鳥配置した

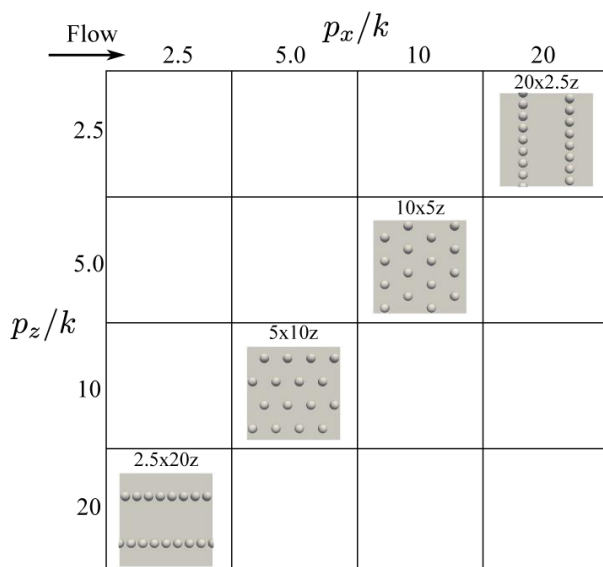


図2 ボルテックスジェネレータを模した半球を千鳥配置した粗面。半球個数は一定で、主流・スパン方向の半球間隔を変えている。

粗面、上面には滑り壁条件とした。下壁に敷き詰めた半球粗面は滑りなし条件・等熱流束加熱条件を与えている。主流(x)・スパン方向(z)には周期境界条件としており、主流方向に圧力勾配を付与することで流体を駆動した。半チャンネル幅と摩擦速度を基にしたレイノルズ数で660として解析行った。また、流体には空気を仮定しプラントル数0.71として、計算格子数は解析ケースによって異なるが3~6億点である。対象とした半球粗面は図2に示すように、半球個数を一定として半球を千鳥配置した粗面で、スパン方向と主流方向の半球ピッチを変更させた4つのケースで解析を行った。速度場の解析には3次元27方向速度多緩和時間格子ボルツマン法[1]を用い、温度場の解析には3次元19方向速度正規化格子ボルツマン法[2]を用いた。

結果および考察

瞬時の温度を可視化した結果を図3に示す。下壁近傍では、壁乱流特有の主流方向に伸びたストリーク構造と半球によって生じる乱れが混在した様子が分かる。下壁近傍には、非常に微細な乱流変動も計算格子によってとらえられており、非物理的な数値振動なども見られないことが確認できる。

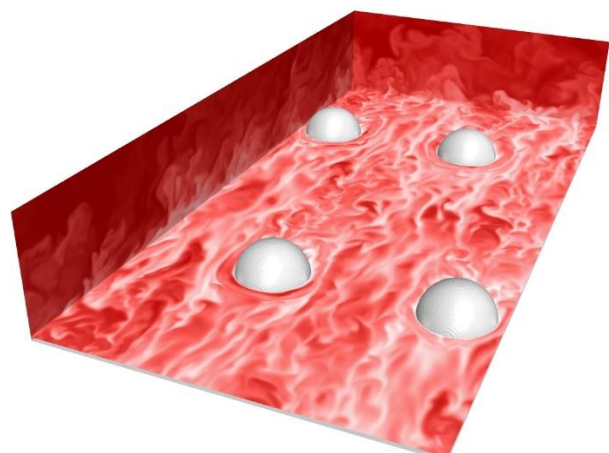


図3 半球粗さ粗面の瞬時温度の可視化結果。粗面は一樣等熱流束で加熱されており、白は高温領域、赤は低温領域を示す。

次に粗面による運動量と熱輸送の増大効果を調査するために、ボルテックスジェネレータを設置することによる壁面摩擦係数の増大率を図4、スタントン数の増大率を図5に示す。これらの結果から、ボルテックスジェネレータを設置することで、壁面摩擦係数・スタントン数ともに増大しており、運動量・熱輸送が増大したことが分かる。しかし、どのケースにおいても、壁面摩擦係数の増大効果はスタントン数の増大効果に比べて大きく、運動量の増大効果は熱輸送の増大効果に比べて大きいことがわかる。この運動量と熱輸送の非相似性は、スパン方向の半球間隔が狭いほど顕著に表れている。主流方向の半球間隔の短い粗面では、摩擦係数の増加率をスタントン数の増加率で除した値が1.4倍であるのに対して、スパン方向の半球間隔の短い粗面では2.4倍になっている。つまり、主流方向の間隔の短い粗面では、少ない摩擦抵抗増で効率的に熱輸送効果を高められることを意味している。

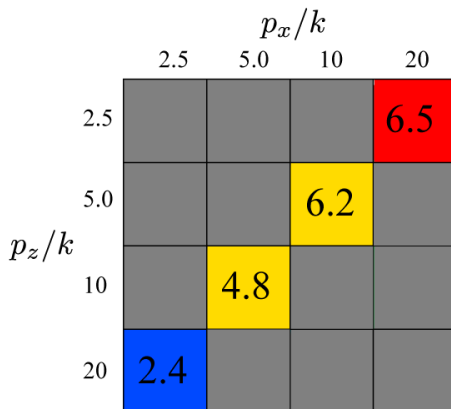


図4 滑面と比較した壁面摩擦係数の増大率.

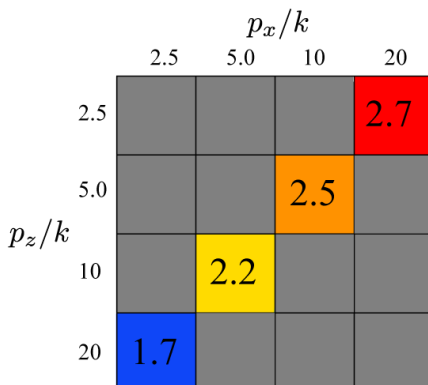


図5 滑面と比較したスタントン数の増大率.

まとめ、今後の課題

格子ボルツマン法を用いてボルテックスジェネレータを模した半球粗面乱流熱流動の直接数値解析を行った。半球個数を固定した場合においても、半球の配置によって運動量と熱輸送の増大効果は大きく異なることが分かった。とくに、主流方向の間隔の短い粗面では、少ない摩擦抵抗増で効率的に熱輸送効果を高められることが明らかになった。

参考文献

- [1] Suga, K., et al. *Computers & Mathematics with Applications* 69.6 (2015): 518-529.
- [2] Suga, K. et al., *International Journal of Heat and Fluid Flow* 68 (2017): 225-236.

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 LRnLA アルゴリズムを用いた物理シミュレーション
 英文: Simulation of Physical Processes with LRnLA Algorithms

善甫 康成
 Yasunari Zempo

法政大学 情報科学部
 Computer and Information Sciences
<http://cis.k.hosei.ac.jp/>

地震解析のための DiamondSeism コードは、毎秒最大 14 億セル更新のパフォーマンスで、層状媒体の地震弾性波をシミュレートできる性能を持っている。この性能は LRnLA 並列アルゴリズムを使用することで達成されたものである。昨年報告したシミュレーションを継続し、更に多くの地震情報を取得している。表面波、屈折波、二次波など総てのタイプの弾性波を考慮に入れた 3D の有限差分モデルにより、地震応答の正確な予測ができるようになった。3D クロスステンシルを持つ新しいモデルを用いて LRnLA コードの最適化のテストを実行した。この新しいアルゴリズムは、今後様々な物理シミュレーションコードの開発で使用できると期待できる。

The code DiamondSeism simulates seismic elastic waves in a layered media with a performance up to 1.4 billion cell updates per second. This performance is achieved with the use of parallel LRnLA algorithms. We continue the simulation batch that has been reported in the previous year, and obtain new seismograms in the extended domain. Taking account of all types of elastic waves: surface wave, refracted waves, secondary waves, finite difference modelling in 3D media gives accurate predictions of the seismic response. We have performed tests for optimization of LRnLA codes on a model with a 3D cross stencil. The new algorithm will be used in further physical simulation code development.

Keywords: LRnLA Algorithm, Hooke's law, synthetic seismogram, finite difference, cross stencil

背景と目的

永久凍土(permafrost)がある地域での地震波トモグラフィーを用いる際に注意しなければならないことの一つに不凍土層(Talik)の存在の有無がある。不凍土層があると地震計の記録に干渉によるノイズが発生することが知られている。3次元全波数値シミュレーションを使用すると、推測した位置にある不凍土層を含む地震層モデルから理論的な地震波が得られる。その記録を解析することにより不凍土層に起因した干渉の検出が可能になる。更に地震波トモグラフィーの解析について補正を行い、地震波を正確にとらえることが可能である。

計算コストの観点から、不凍土層は低速度層であるので、解析の解像度は空間ステップによって決まり、それがモデルの全計算量を決める。

4次近似の数値スキームを使用すれば、空間ステップは一番短い波長の8分の1程度で十分であり、必要な精度が得られるモデルリングが可能である。不凍土層のP波の伝播速度は $V_p \sim 2400$ m/s, S波の伝播速度は $V_s \sim 1550$ m/s であるので、長さが ~ 10 km程度のシミュレーションを行うと、データ量は50GB以上になる。また幾つかの震源位置での地震解

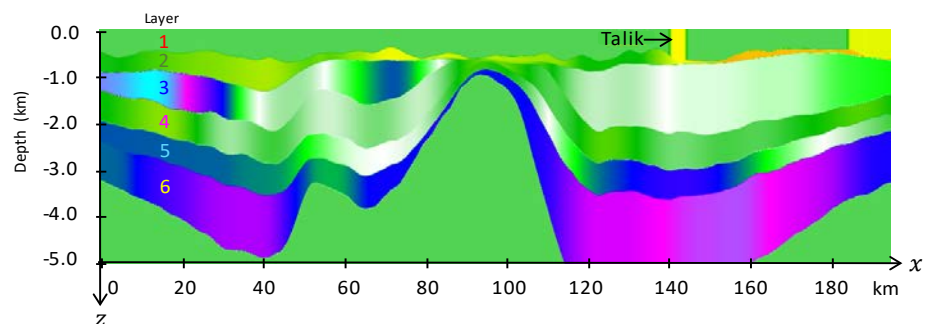


図 1. 地震層モデル。Layer1 の 141200m-144925m が Talrik 層である。地震源は 134925m から 151200m の空間にあるとした。

析の結果を得るには、必要な計算を 1000 ケース以上行う必要がある。

このためには効率的なアルゴリズムを持つ並列コードと高性能計算機の利用が不可欠となる。本プロジェクトでは、GPU を使用して地震波の伝播現象をシミュレーションする高性能コード DiamondSeism を many-GPU の環境である TSUBAME3.0 に移植し不凍土層を含む地震層モデル計算を実施した[1]。

加えて並列効率とメモリ効率を更に向上させるため、新しい LRnLA アルゴリズムと実装法を開発し、その試験を実施した。

内容

地震波の伝播の数値モデルはフックの法則とニュートン運動の法則からなる[2]。モデルは、応力テンソルと変位速度ベクトルの時間発展をあらわす。使用されている Finite-Difference 数値スキームは空間で4次近似、時間で二次近似である[3]。境界条件として PML を使用し、地震波源のため TF/SF 法を使用している。なお本モデルは以前、DTmaxwell コード[4]を基本に TSUBAME2.5 と TSUBAME3.0 上で開発してきた解析を実施してきた実績のあるものである。

図 1 に地震層モデルを示す。地表から地球内部に向かって主要な構造が 6 層ある。これらの層を用い図 2 に示す層のパラメータ依存性を考慮しシミュレーションを行った。密度の空間依存性、P 波および S 波の伝播速度との空間依存性を、それぞれ図 2、図 3 に示す。本プロジェクトで用いたモデルは 3 次元データを用いた 3 次元解析であるが、1 断面をとり 2 次元の表示としている。伝播速度が低い Talik 層は Layer1(赤線)の中の 141200 m – 144925 m に示してある。

本コードの特徴は、LRnLA アルゴリズムの使用していることである。この数値計算スキームでは局所依存性を考慮し、メモリバンド幅への負担をできるだけ最小化・緩和する。また LRnLA アルゴリズムは、全メモリ階層と全並列法を考慮し temporal blocking[5] の考え方を更に進めた形になっている。GPU の場合、計算のデータを比較的に大きいレジスタファイル内に閉込め CUDA-スレッドの並列性を使う[4]。

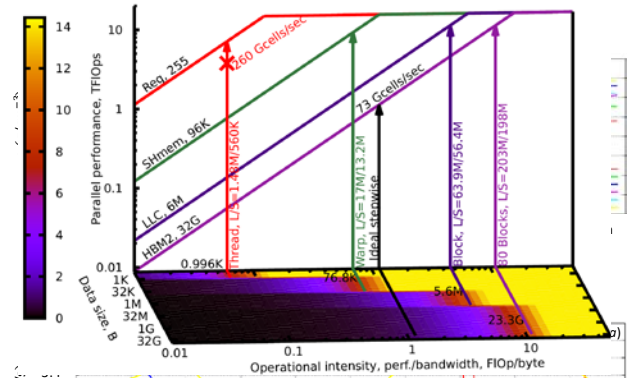


図 4. Roofline 解析による性能評価

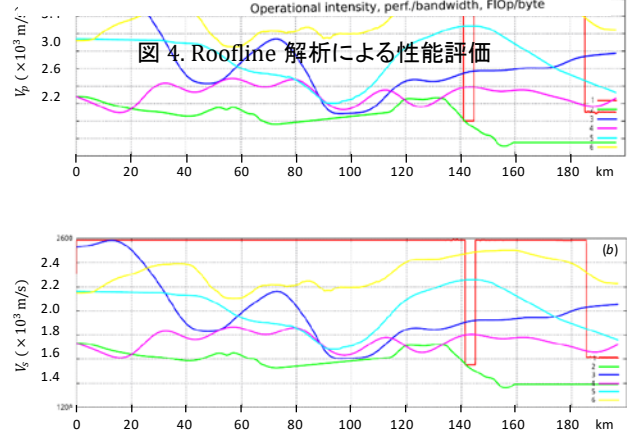


図 3 P 波(a)および S 波(a)の伝播速度との空間依存性

本プロジェクトでは、スレッドの並列に加え、新しい試みとして CUDA-warp 並列と shared メモリアクセス数を減じて実装効率増加させるという実装法を導入した。

図 5 に RoofLine[6] 解析による性能評価の結果を示す。3 次元クロスステンスルを用いたコードのパフォーマンスの見積と NVidia P100GPU での実施結果である。黒い矢印は LRnLA アルゴリズムを用いない従来型コードの最大パフォーマンスの見積である。これは全データ(23.3GB)をデバイスメモリ(HBM2)に保存し、1 回全データのアップデートのため、数値毎にロードして、再度 1 回セーブすることが必要であるということから算出した理想的な場合の見積結果である。

LRnLA アルゴリズムを用いた場合の上限を赤い矢印で示す。従来型のコードとは様子が大きく異なる。その理由は、デバイスメモリからのロードが一番時間が掛かるが、一旦ロードすると、既にロードされているデータを用い多数の計算が行われるため、operational intensity が非常に高く保たれるからである。単純に LRnLA アルゴリズムを用いた結果を図 4 中に桃色の矢印で示す(一番右)。3 次元のデータを DiamondTorre アルゴリズムにて計算した時のデータ量から算出した推測値である。また DiamondTorre 間のデータ交換のため、LLC cache を使用可能と仮定した結果を、二番目の 'Block' という紫色の矢印で示している。また CUDA-block 中の計算では CUDA-スレッドの並列化を使用するので、更に DiamondTorre 間のデータ交換のた

め、LLC cache を使用可能と仮定した結果が、右から二番目の'Block'という紫色の線で示した結果である。かなり処理能力が向上していることがわかる。

また CUDA-block 中の計算では CUDA-スレッドの並列化を使用する。Warp-間のデータ交換には shared メモリを使用する。これを利用した場合の上限値を見積ったのが緑の矢印で示した値である、もちろん CUDA-warp 中のデータ効果は shuffle ベクトル演算を用いるため、更に性能が向上することが期待できる(赤の矢印)。その結果、パフォーマンスの上限値は 490 cell update per second であると思える。実際のテストでは、オーバーヘッドのためと思われるが、260 cell update per second であった。

結果および考察

TSUBAME3.0 のノード内の 4 つの NVIDIA Tesla P100 を用い計算を行った。

震源の振動数は 30 Hz である。モデルの 134925 m から 151200 m までの空間内で 50 m ごとに 325 個の地震源があるとした。地震波センサーは地表に 25 m 毎に 650 個設置し、シミュレーション領域は地震計を設置した線に沿って長さ 10 km であり、深さは 5 km である。空間のメッシュ数は $1500 \times 600 \times 256$ 個である。地震計の線の垂線の方向で 128 メッシュ内には PML 層も含んでいるが、球形の波面を持つ地震波の伝播距離にからすると減衰には十分であると考えられる。

計算データは 13 GB となるので、TSUBAME のノード内では、4 つの独立なシミュレーションを実行することが可能である。時間ステップは ~ 0.667 秒であるため、5 秒の伝播を図るためには 15000 ステップ必要である。現在のパフォーマンスは 3.1×10^9 cell update per second であり一つの計算は 12 分程度で済む。

計算結果の例として、図 4 では理論的に求めた地震計の記録が表している。これは、 $t-x$ のグラフで変位の速度の V_z である。 z -軸は下に向いている。不凍土層による反射が良く見てとれる。

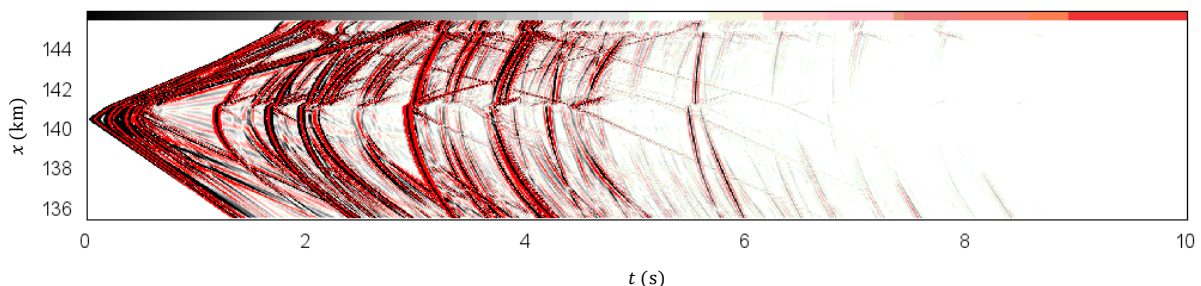


図 6. 理論的に求めた地震波。震源付近での変位の速度の大きさを表している。

まとめと今後の課題

我々が、これまでに開発した LRnLA アルゴリズムを使ったコードを地震波の伝播計算に展開した。また LRnLS に関して GPU のハードに合わせた新しいアルゴリズムの開発も行った。その特徴は、3次元データにおける計算の局所化とレジスター内でのデータ交換である。CUDA を用い実施法の単純化がさらに必要でありさらに改良を続ける予定である。完成すれば、電磁場、地震波のシミュレーションで使用可能になり、現在使用している地震波解析コード DiamondSeism より高いパフォーマンスが期待できる。

参考文献

- [1] T. Levchenko, V. Rok, V. Levchenko, A. Perepelkina, Y. Zempo, "Computer modelling specifics of the geological structure with contrasting inhomogeneities under the permafrost conditions" GEO Eurasia-2019, Materials of the International geologic and geophysics conference and exhibition (Feb. 4-7, 2019, Moscow), pp. 814-817 <https://www.gece.moscow/> https://drive.google.com/file/d/1Z7o5H8ZQ6EYqgp_atcbBEGlaKiEMFiarh/view?usp=sharing
- [2] Aki, Keiiti, and Paul G. Richards, "Quantitative seismology", 2002.
- [3] Levander Alan R. Fourth-order finite-difference P-W seismograms // Geophysics. 53(11), p1425-1436., 1988
- [4] A. Zakirov, V. Levchenko, A. Perepelkina, Y. Zempo, "High performance FDTD algorithm for GPGPU supercomputers", J. Phys.: Conf. Ser. 759 012100, 2016
- [5] Endo, Toshio. "Applying Recursive Temporal Blocking for Stencil Computations to Deeper Memory Hierarchy." 2018 IEEE 7th Non-Volatile Memory Systems and Applications Symposium (NVMSA). IEEE, 2018.
- [6] Williams, Samuel, Andrew Waterman, and David Patterson. "Roofline: an insightful visual performance model for multicore architectures." Communications of the ACM 52.4 (2009): 65-76.

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 近赤外線を利用する色素の探索

英文: Theoretical search for pigments utilizing near infrared radiation

利用課題責任者

Yu Komatsu

所属

NINS-ABC / NAOJ

<http://abc-nins.jp/>

邦文抄録(300 字程度)

近赤外線を利用する色素分子は生体バイオイメージングや色素増感電池などの分野で待望されている。光合成の長波長限界に関しても近年更新されており、その実現のされ方、既存の分子をどのように代替できるかは興味深い問題である。ここでは、より赤い光を利用し、光合成色素として機能し得る分子を探索するために、量子化学計算によって広い条件で網羅的に物性を調べた。バクテリオクロロフィル、フタロシアニンなどの分子を鋳型に、中心金属(Mg, Ca, Ni, Zn, Sr, Pd, Cd, Ba, Pt, Hg, Pb, H₂)、4種の溶媒などの条件を変えて計算した。その結果、中心金属をかえるだけで最低励起状態の波長に80nm以上の大きな幅が見られること、原子半径が大きいもののPbであるとより長波長を吸収して、既存の分子に近い物性を示すことなどが推定された。

英文抄録(100 words 程度)

Organic dye molecules utilizing NIR radiation are anticipated in the field of bioimaging, dye-sensitized solar cells and so on. In this study, in order to explore molecules as photosynthetic pigments which absorbs redder radiation in varieties of conditions by quantum chemistry calculations: molecules like bacteriochlorophylls and phthalocyanines as templates, central metal (Mg, Ca, Ni, Zn, Sr, Pd, Cd, Ba, Pt, Hg, Pb, H₂) and four solvent conditions. As a result, we found that the estimated wavelengths at the lowest excited states range over more than 80 nm with different central metals, and pigments with Pb, whose radius is relatively large, absorb longer wavelength radiation and possess physical chemical properties close to existing pigments.

Keywords: quantum chemistry, photosynthesis, DFT, astrobiology

背景と目的

酸素を発生するタイプの光合成細菌であるシアノバクテリアを近赤外光下で生育すると、より長波長を吸収するクロロフィル *f* が光化学反応に利用されるようになったことが近年発見された (Nürnberg *et al.*, 2018, *Science* 360:1210–1213)。このように、光合成の長波長限界はなお更新されているところである。一方、天文学では、太陽より低温の恒星を持つ惑星が観測ターゲットであり、そこでの光合成の痕跡を同定することに興味を持たれている (Kiang *et al.*, 2007b, *Astrobiology*, 7:252–274)。また、生体バイオイメージングや色素増感電池などの医療・環境・エネルギーなどの分野でも近赤外線を利用する色素分子の提案が望まれている。

上記のような様々な研究領域で、近赤外線を有効利

用して、機能性を持った分子を効果的に探索することが期待されている。光合成の長波長限界に関しては、いくつらな推定があるもののその実現のされ方はわからないことが多い。これは、既存の色素分子を他のもので代替可能であるかという問題に置き換えることができる。昨今のハイスループットスクリーニングの方法論の重要性も鑑みつつ、理論的な土台を均一の基準で押さえておくことは有効である。

本プロジェクトでは、近赤外線を利用し、光合成色素として機能し得る分子を探索するために、色素、中心金属、溶媒などの広い条件で網羅的に、量子化学計算による物性の依存性を調べた。その結果、依存性のより小さなバクテリオクロロフィル *b* であっても中心金属をかえるだけで最低励起状態の波長に~80nmの差が出ること、原子半径が大きいもののPbがより赤くてさらに興味深い性質を示す可能性があること、などがわかった。

概要

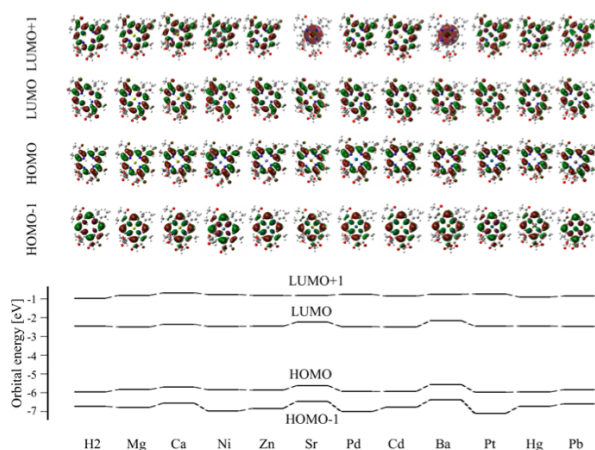


図 1 バクテリオクロロフィル *b* の中心金属を替えた分子に関して、推定された HOMO/LUMO 周辺の電子状態と軌道エネルギー(真空中)。

近赤外線を光吸収する分子をいくつか用意し(バクテリオクロロフィル *b*、フタロシアニン、ジベンゾポルフィセンなど)、中心金属として Mg(天然のクロロフィルが有する)、H2(クロロフィルに対してフェオフィチン系の分子に相当)、Zn(クロロフィルにおいて発見されている)、Ca, Ni, Sr, Pd, Cd, Ba, Pt, Hg, Pb、溶媒として、真空、水、メタノール、ベンゼン(PCM)の広い条件で、物性を調べた(図1)。現実的な計算コストで評価するために、CAM-B3LYP/Def2tzvp//B3LYP-D3/Def2tzvp の精度で、光学的特性、酸化還元的性質、中心金属の脱離の程度を評価した。

結果および考察

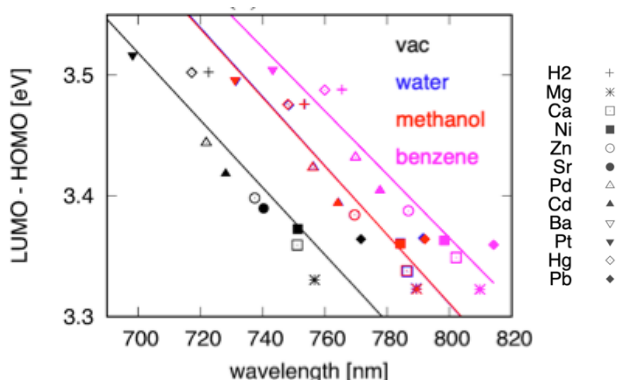


図 2 バクテリオクロロフィル *b* の中心金属を替えたもの(溶媒4種)について、推定された最低励起状態の波長に対して、HOMO-LUMO ギャップを示した。溶媒毎に fit した直線をプロットした。

天然の光合成色素で最も長波長に吸収極大を持つバクテリオクロロフィル *b* の、TDDFT で見積もった光学的特性に関して、中心金属、溶媒の条件を変えたときの依存性が図2に示されている。同じ色素分子をベースにしつつ、110 nm 以上の幅があることが推定された。中心金属の variation としては、80nm 以上差がある。より赤い条件を見ると、例えばベンゼン中では真空中と比較して~40nm レッドシフトしている。中心金属に関しては、天然に見られる Mg はより長波長に位置し、Pb はそれより若干赤く真空中では~17nm 長波長にシフトしていた。逆に最も青い側に位置していたのは Pt であった。

さらに酸化還元的な性質、イオン化ポテンシャル、電子親和力を計算したところ、大きな原子半径を持つ Ca, Sr, Ba などの条件では、他の中心金属と比較していずれも小さな値を取ることがわかった。さらに、金属と色素の間の結合エネルギーを評価したところ、先の Ca, Sr, Ba は様々な溶媒で脱離しやすいことがわかった。よってこれらは、色素としての役割を担うには適していないという結論を得た。一方、Pb に関しては、原子半径が比較的大きいにも関わらず、Mg などの値からは大きく外れておらず、既存の光合成色素の中心金属を代替する候補となる可能性がある。これらの傾向は他のクロロフィルや人工色素のフタロシアニンでも同様に見られた。

まとめ、今後の課題

より赤くて、光合成色素として機能しうる分子条件を、広く均一に調べた。その結果、1) 中心金属を変えるだけで 80nm 以上の幅が見られること(溶媒を含むと 110nm 以上)、2) Pb であると Mg よりわずかに赤く、既存の分子と近い物性を示すこと、3) Ca, Sr, Ba などの半径の大きなものは適していないこと、などがわかった。

ここまではかなり広い条件で網羅的に依存性をみたが、これを踏まえてさらに精密な評価をする必要がある。例えば実験で得られる酸化還元的電位を理論的に精度良く予言するのは難しいが重要である。これまでの単量体の推定から、天然に見られる2量体(スペシャルペア)の評価にシフトし、妥当に酸化還元的電位と結びつける評価基準を作ることがさらなる課題である。

なお、関連研究に関して、物性研と国立天文台のスパコンも利用したことを追記する。

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 高解像度画像を使った広域の家屋及び道路の深層学習による自動判別システムの開発(3)
 英文: System development of automatic building and road network detection using high resolution
 imagery by deep-learning 3

柴崎 亮介
 Ryosuke Shibasaki

東京大学 空間情報科学研究センター
 Center for Spatial Information Science, University of Tokyo
<http://csis.u-tokyo.ac.jp>

本プロジェクトでは衛星画像データ内に見られる建物・家屋を深層学習によって自動認識する処理を、地図データとして利用するためのマッピング処理を自動化することで、広域に展開するためのシステム開発を実施した。TSUBAME の GPU クラスタと連携したシステムを開発した。昨年度までの対象地域(ラオス、スリランカ、ケニア、モザンビーク)に加えて、ザンビアを対象として学習データを追加し、ResNet をバックボーンとした UNet モデルを構築した。その結果、Accuracy は 0.9916、IoU (Intersect of Union) は 0.6057 (検出結果と正解データが互いに約 75% 重なり合う) となり、昨年に比べて短い処理時間で高い精度を達成することができた。今後は、成果データの配信方法も視野に入れて、品質管理・改良を効率的に進める手順・方法を研究開発し、当システムへの実装を進める。

This project developed a system to apply building mapping to broad areas by deep learning to satellite images. The system was developed with TSUBAME's GPU clusters. In addition to the regions covered by the previous year (Laos, Sri Lanka, Kenya, and Mozambique), We added Zambia to the target areas for training data set and applied a UNet model with ResNet as the backbone. As a result, it achieved an accuracy of 0.9916 and an IoU of 0.6057 in a shorter computing time than last year's. For future works, we are going to develop processes and methods for efficient quality control and improvement as well as dissemination of the mapping outputs and implement those to the system.

Keywords: deep learning, satellite images, building mapping, large-scale processing

背景と目的

持続可能な開発課題 (Sustainable Development Goals; SDGs) に Goal 11 “Make cities and human settlements” として掲げられているように、都市の成長を適切に管理するためのデータ基盤の整備には早急な対応が求められている。地球観測データは、このようなニーズに対応するデータ資源として、特に広域性と高頻度の特長が、都市成長の観測において高い有用性が期待されている。近年、Google Earth をはじめとした高分解能衛星画像データの普及により、個別の建物を衛星画像から検出し、衛星画像データの位置情報とあわせることで正確にマッピングする可能性が議論され、機械学習や深層学習の適用によって、自動的に衛星画像内の建物を検出する研究がなされてきた。建物や道路のマッピングの自動化は、広域にわたる地図作

成を低コストで実現し、インフラ開発や公衆衛生の課題において、都市部だけでなく郊外の社会経済の状況を概観するのに有用である。広域にわたる自動マッピングへのニーズは高まる一方、これまでの研究の成果は、単一シーンの衛星画像データといった限られた領域でのみ処理されており、全国土といった広域を対象とした処理は未だ試みられていない。

本プロジェクトでは、昨年度の成果等をもとに、深層学習による自動マッピングアルゴリズムを、ザンビアに拡張したほか、高度なモデルをである ResNet32 の適用を試みた。

概要

社会基盤情報として人口分布、交通ネットワークの整備は必須であるが、途上国、僻地の情報基盤の整備は乏しい状況にある。オープンな利用が広まりつつ

ある高解像度衛星画像から機械学習、ディープラーニングなどの手法で家屋や道路ネットワークを自動検出し、地図データと利用することによる広域・社会基盤情報の整備に寄与する。

本プロジェクトでは、東工大 TSUBAME をデータ処理の基幹とし、データ収集・管理の機能を連携させることで、任意の領域について、家屋マッピングを自動化することができた。前年度のトレーニングデータにザンビアを加えて、ResNet32 を適用することで Accuracy は 0.9916、IoU は 0.6057 を達成した。

結果および考察

ディープラーニングに入力する学習データを効率的に作成するためのシステムをオープンソース GIS である QGIS を使って構築し、インターネット地図サービスの高分解能衛星画像に見られる建物を、約 300mx300m の解像度 0.3m のタイル画像を単位として GIS のポリゴンデータとして記録する手順を整備した(図 1)。ザンビア全土を対象として 9,720 タイルをトレーニングサンプルとして抽出し、53,336 の建物をポリゴンデータとして作成した。

建物検出の精度を向上させるにあたり、ディープラーニングの処理において様々な画像認識に共通する低レベル特徴量を転用することで計算時間を短縮する転移学習を適用した。昨年度に使用した U-net モデルに、多数の層の組み込みを可能にする ResNet32 による重み付けを適用することで転移学習を実装した。実装にあたっては Segmentation Models¹を用いた。95,321 の画像パッチを学習データとして入力した結果、Accuracy は 0.9916、IoU (Intersect of Union) は 0.6057 (検出結果と正解データが互いに約 75% 重なり合う) となり、昨年比べて短い処理時間で高い精度を達成することができた。

都市部・郊外部とも良好な結果が得られた(図 2、図 3)。また、前年度の課題であった郊外部の散在する建物の検出結果が良好であった。(図 4、図 5、図 6)。しかし、大きな建造物の一部が検出結果のから欠落したり、樹林等で一部が隠れる建物が検出されない例が確認された(図 7、図 8)。これらのケースに対し

ては、学習データを追加することで改善する必要がある。

作成した建物検出モデルをもとに、ユーザーがウェブブラウザから任意の緯度経度を入力し、建物の検出結果を GIS データとして出力・ダウンロードするシステムを開発した。このシステムを使って、2019 年 7 月 16 日に公開ワークショップ『衛星画像を用いた「自動家屋の登録システム」ハンズオンセミナー』を開催した。

まとめ、今後の課題

ResNet32 の適用により、昨年度より高い精度で高分解能衛星画像内の建物を自動で検出するモデルを構築し、作成したモデルを使ったエンドユーザーインターフェースをウェブ上に作成し、それを使ったワークショップを開催した。また、ディープラーニングによるモデル構築のほか、効率的に学習データを追加するための仕組みを整備した。今後は、モデルの見落としや誤検出について、利用者からのエラー報告を学習データとして使うための仕組みを整備し、さらにモデルの精度を向上させ、広域に展開させる。

¹ https://github.com/qubvel/segmentation_models



図 1 QGIS を使った学習データ作成



図 5 郊外部の検出結果例 2



図 6 郊外部の検出結果例 3



図 2 ザンビア・Kabwe 周辺マッピング結果。赤色は検出された建物を表す。



図 7 郊外部の不検出例

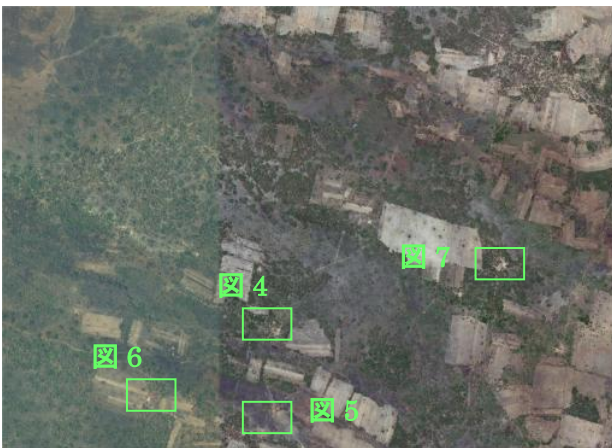


図 3 ザンビアの郊外部マッピング結果。赤色は検出結果を表す。



図 8 都市部の建造物検出の一部欠落の例



図 4 郊外部の検出結果例 1

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 TSUBAME3.0 における TensorFlow アプリケーションの性能評価
 英文: Performance Evaluation of TensorFlow Applications on TSUBAME3.0

利用課題責任者

三輪 忍

所属

電気通信大学 大学院情報理工学研究科

<http://www.uec.ac.jp>

邦文抄録(300 字程度) 深層学習は膨大な計算資源を必要とすることから、高性能計算環境上で分散処理されることが多い。分散学習に使用するデータセットの量は今後さらに増大すると予想されることから、高性能計算機システムの I/O 性能はますます重要になると考えられる。ところが、分散学習における I/O 性能を評価した研究は少なく、評価が十分とは言えない現状がある。そこで本研究では、いくつかの代表的な畳み込みニューラルネットワークを用いて、訓練データセットサイズ等のパラメータを変更した場合の TensorFlow による分散学習性能を評価した。その結果、メインメモリに収まらないサイズの訓練データセットを使用する場合でも、ファイルシステムから計算ノードへの訓練データの転送が性能上の問題とならないことを確認した。

英文抄録(100 words 程度) Deep learning requires many computing resources so that distributed processing with HPC systems is widely used. Since the amount of datasets used for distributed deep learning is expected to increase in the future, I/O performance will become much more important in HPC systems. However, the I/O performance of distributed deep learning has not been well investigated. In this study, we test the performance of distributed deep learning with TensorFlow and some representative CNNs on TSUBAME3.0 while varying some parameters such as training dataset sizes. Our experimental results do not show that data transfer from the file system to a compute node is a bottleneck in distributed deep learning even in case of using large datasets.

Keywords: TensorFlow, Distributed Deep Learning, CNN, Performance Evaluation, I/O

背景と目的

自動運転を始めとする幅広い分野で活用されている深層学習は多くの計算資源を必要とすることから、クラウド等の高性能計算環境において複数の計算機を使用して実行(分散学習)されることが多い。一方、深層学習に使用する訓練データセットは、多いものでは既に 3.5TB に達しており[1]、将来的にはさらに増大することが見込まれる。したがって、次世代の AI 向け高性能計算環境においては I/O 性能が重要になる可能性が高いことから、本プロジェクトでは将来の分散学習の要求 I/O 性能を明らかにすることを目的とする。

高性能計算環境における分散学習性能評価はいくつか行われているが[2,3]、I/O 性能に着目して分散学習性能を評価した先行事例は少ないのが現状である。これは、従来の性能評価実験は、次世代の分散学習アプリケーションを想定しておらず、評価対象を現存するシステムアーキテクチャとアプリケーションの組み合わせ

せとした結果、I/O が分散学習性能を律速するユースケースがほとんどなかったためと考えられる。

本プロジェクトでは、TSUBAME3.0 と TensorFlow を用いて、将来のユースケース(具体的には訓練データセットサイズが計算ノードのメインメモリサイズを超える場合)を想定した上で分散学習の性能評価と分析を行い、次世代の高性能計算環境に必要とされる I/O 性能を明らかにすることを試みた。評価の結果、当初の予想に反し、上述のユースケースにおいては、ファイルシステムから計算ノードへの訓練データの転送が性能上の問題とならないことがわかった。

概要

本実験では、TSUBAME3.0 の計算ノード 1 台を評価に使用する。TSUBAME3.0 の各計算ノードは 4 台の GPU を有しており、これら 4 台の GPU を用いて分散学習を行う。訓練データセットは、高速・大容量の外

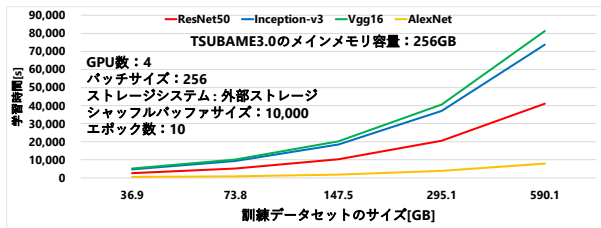


図 1. 訓練データセットサイズと学習時間の関係

部ストレージである Lustre ファイルシステム上に配置する。

評価には 4 つの畳み込みニューラルネットワーク (ResNet50, Inception-v3, Vgg16, AlexNet) を使用する。評価に使用した TensorFlow のバージョンは 1.12 である。評価した分散学習方式は同期型のデータ並列である。訓練用データセットとして、ILSVRC2012 データセットの学習用データ(データサイズ: 147.5GB、画像枚数: 約 120 万枚)をスケールして使用した。

結果および考察

異なるサイズの訓練データセットを用いて分散学習を行った場合の学習時間を図 1 に示す。グラフの横軸は訓練データセットサイズ、縦軸は学習時間である。バッチサイズは 256 とし、10 エポック分学習を行った。

グラフより、いずれのニューラルネットワークも学習時間は訓練データセットサイズに比例して増加する。特に、訓練データセットサイズが 147.5GB 以下(メインメモリサイズ以下)の領域と 295.1GB 以上(メインメモリサイズ以上)の領域を比較すると、学習時間の増分はほとんど同じであることがわかる。当初の予想では、訓練データセットサイズがメインメモリサイズを超える場合に I/O へのアクセスが多発し、学習時間の大幅な増加が起こると考えていたが、今回の実験ではそのような傾向は見られなかった。

図 2 に、AlexNet を学習中の GPU の使用状況とデータ転送状況のプロファイリング結果を示す。プロファイリングには NVIDIA Visual Profiler を使用した。図の横方向は時刻を表しており、各行はイベントの種類を表している。また、青色と黄土色の四角は、それぞれ計算/データ転送イベントを表している。なお、実験に使用した訓練データセットのサイズは 295.1GB である。

図より、訓練データを GPU へ転送する時間は性能上のボトルネックとなっていないことがわかる。メインメ

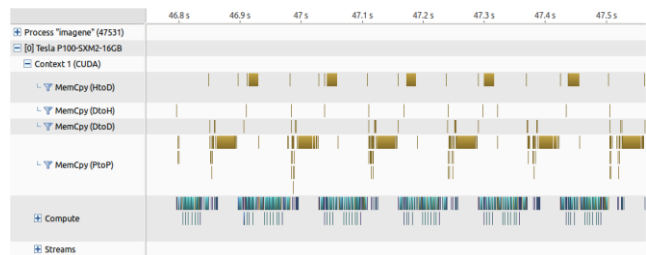


図 2. 学習中の GPU の使用状況とデータ転送状況

メモリから GPU メモリへのデータ転送イベントの時間は、GPU による計算イベントの時間によって完全に隠ぺいされており、前者のイベントは学習時間に影響を及ぼしていない。また、図には示されていないが、メインメモリから GPU メモリへのデータ転送の前に発生するはずの、外部ストレージからメインメモリへのデータ転送の時間も GPU の計算時間によって隠ぺいされている。なお、GPU による計算は GPU 間のデータ転送によって中断されているが、これは複数 GPU 間で同期処理を行い、学習結果を共有しているためである。

まとめ、今後の課題

本実験では、メインメモリサイズを超える規模の訓練データセットを用いてノード内/データ並列方式による学習を行うケースを想定し、TSUBAME3.0 上で TensorFlow による分散学習の性能評価を行った。本実験により分散学習と TensorFlow に関する数々の知見が得られたものの、今回のユースケースでは I/O が分散深層学習の性能を律速する状況は残念ながら確認できなかった。今後は、I/O に対する要求性能がより厳しいユースケースにおける性能評価を行う必要があると考えている。

参考文献

- [1] T. Kurth, et al., “Exascale Deep Learning for Climate Analytics,” SC18, pp. 51:1-51:12 (2018).
- [2] X. Li, G. Zhang, et al., “Performance Analysis of GPU-Based Convolutional Neural Networks,” ICPP, pp. 67-76 (2016).
- [3] F. Chowdhury, et al., “I/O characterization and performance evaluation of BeeGFS for deep learning,” ICPP, pp. 1-10 (2019).

TSUBAME 共同利用 令和元年度 共同利用 成果報告書

利用課題名 分子動力学シミュレーションを用いた抗関節リウマチ分子標的薬の設計
 英文: Molecular design of anti-rheumatic drugs with molecular dynamics simulation

利用課題責任者 和久井直樹
 First name Surname Naoki Wakui

所属 独立行政法人国立高等専門学校機構 長岡工業高等専門学校
 Affiliation National Institute of Technology (KOSEN), Nagaoka College
 URL <http://www.nagaoka-ct.ac.jp/>

邦文抄録(300 字程度)

関節リウマチは全身性の自己免疫疾患であり、滑膜関節炎と破壊性関節病変により身体機能障害をきたす疾患である。関節リウマチの原因タンパク質として Protein Arginine Deiminase4(以下、PAD4)が同定されている。本研究では PAD4 選択的阻害剤として知られる GSK199 と PAD4 の相互作用ならびに GSK199 と PAD4 のアイソザイムである PAD1、PAD2、PAD3 との相互作用を分子動力学シミュレーションを用いて解析した。相互作用解析の結果、PAD4 と GSK199 の間にだけ π - π スタッキングが形成されることが分かった。 π - π スタッキングを形成する Phe634 に対応する PAD1~PAD3 のアミノ酸の側鎖は芳香環を持たないことから、Phe634 が GSK199 の標的選択性に関与していることが示唆された。

英文抄録(100 words 程度)

Rheumatoid arthritis is a systemic autoimmune disease that causes physical dysfunction due to synovial arthritis and destructive joint lesions. Protein Arginine Deiminase4 (PAD4) has been identified as a causative protein in rheumatoid arthritis. In this study, we analyzed the interaction between GSK199, which are known as selective PAD4 inhibitor, and PAD4 and the interaction between GSK199 and PAD1, PAD2, and PAD3, which are PAD4 isozymes, with molecular dynamics simulations. These interaction analyses showed that π - π stacking was formed only between PAD4 and GSK199. the amino acids corresponding to Phe634, which forms the π - π stacking, from PAD1 to PAD3 do not have an aromatic ring, suggesting that Phe634 is involved in the target selectivity of GSK199.

Keywords: Molecular dynamics, Drug discovery, Protein Arginine Deiminase

背景と目的

関節リウマチは全身性の自己免疫疾患であり、滑膜関節炎と破壊性関節病変により身体機能障害をきたす疾患である。患者数も年々増加傾向にあり、特に 50 代から 70 代の女性での発症が多いのが特徴である。既存の関節リウマチに対する治療は免疫抑制剤または生物学的製剤の使用による寛解が限界である。リウマチ専門医の数が限られ、生物学的製剤を使用する場合は月 1 回の点滴投与または 2 週間に 1 回の皮下注射となっており、特に患者数の多い 50 代から 70 代の患者に対しての身体的負担が大きい。これまでの基礎研究結果から、関節リウマチ患者はシトルリン化したタンパク質に対する自己抗体を多く生産していることが明らかになっており、タンパク質をシトルリン化する酵素として Protein Arginine Deiminase(以下、PAD)が同定されている。PADには PAD1、PAD2、PAD3、PAD4、

PAD6 の 5 つのアイソザイムが存在しており、PAD1 から PAD4 はシトルリン化活性を有しているが PAD6 は多数の変異があるためシトルリン化活性がない。PAD1 から PAD4 のうち、PAD4 のシトルリン化活性が高くなることが関節リウマチの発症と進行に寄与していることから、標的タンパク質として注目されている¹⁾。しかし、アイソザイム間でのアミノ酸配列一致度が 50%を超えているため、PAD4 阻害剤として開発した場合であっても PAD1、PAD2、PAD3 に結合し阻害する場合がある。標的タンパク質以外を阻害してしまう場合、副作用などの重篤な症状が出る場合があり、高い標的選択性を持った PAD4 阻害剤の開発が求められている。本研究では PAD4 のシトルリン化活性阻害を作用機序とする PAD4 選択的阻害剤の設計を分子動力学シミュレーションにより行うことを目的とした。TSUBAME3.0 でサポートされている Schrödinger Small Molecule Drug

Discovery Suite を用いて、アイソザイム (PAD1~PAD3) と候補化合物の複合体構造のモデリングならびに分子動力学シミュレーションを実施し、PAD4 への選択性に寄与していると考えられる相互作用を同定した。

概要

関節リウマチの原因タンパク質として PAD4 が同定されている。PAD4 選択的阻害剤の開発には、PAD4 との amino 酸配列一致度が 50% を超えるアイソザイムとの基質認識機構の違いを明らかにする必要がある。本研究では、PAD4-阻害剤複合体とアイソザイム-阻害剤複合体に対して分子動力学シミュレーションを行い、基質認識機構に影響を与える amino 酸残基の同定を行った。

結果および考察

複合体構造のモデリング

Lewis らの先行研究より、GSK199 と呼ばれる化合物が PAD4 を選択的に阻害する可能性があることが示唆された²。また、Lewis らは PAD4 と GSK199 との複合体構造の決定に成功している。PAD4 を選択的に阻害する化合物を設計するためには、PAD4 のアイソザイムとして知られている PAD1、PAD2 および PAD3 と GSK199 との相互作用と PAD4 と GSK199 の相互作用の差異を見つける必要がある。しかし、これらアイソザイムと GSK199 の複合体構造は未だ報告されていないため、PAD1~PAD3 と GSK199 の複合体構造をモデリングする必要がある。PAD4/GSK199 複合体構造に基づき、ホモロジーモデリングと呼ばれる手法を用いて PAD1/GSK199 複合体構造、PAD2/GSK199 複合体構造、PAD3/GSK199 複合体構造のモデリングを行った。

GSK199 とタンパク質間の相互作用解析

TSUBAME3.0 および D. E. Shaw Research がアカデミア向けに無償提供している Desmond を用いて各複合体構造に対する 200 ns の分子動力学シミュレーションを実施し、分子動力学シミュレーションの結果得られたトラジェクトリに対して相互作用解析を行った。PAD1~PAD4 の相互作用を比較したところ、

PAD4/GSK199 複合体にのみ、Phe634 との π - π スタッキングによる相互作用が全シミュレーション時間のうち約 80% の割合で形成されることが確認された (図 1)。PAD1~PAD4 の amino 酸配列を比較してみると、PAD4 の Phe634 に対応する PAD1、PAD2 および PAD3 ではそれぞれ Leu634、Ser636 および Thr635 であった。これらの amino 酸はいずれも側鎖に芳香環がないため、PAD4/GSK199 複合体で見られたような π - π スタッキングが形成できないと考えられる。

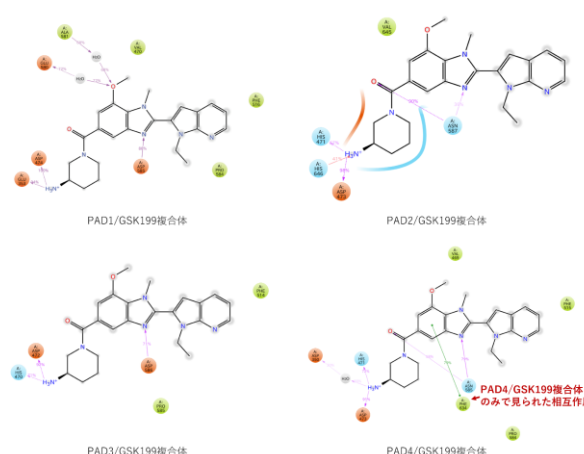


図 1 各複合体構造における相互作用

GSK199 とタンパク質間の結合自由エネルギーの推定

Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) 法を用いて、GSK199 と各タンパク質との結合自由エネルギーを推定した。Run1、Run3 および Run4 において PAD4/GSK199 複合体がエネルギー的に有利であることが示唆された。

表 1 結合自由エネルギー推定結果

複合体	ΔG_{bind} (kcal mol ⁻¹)			
	Run1	Run2	Run3	Run4
PAD1/GSK199	-72.6±3.3	-63.4±4.3	-67.0±4.6	-70.4±5.7
PAD2/GSK199	-72.5±7.9	-81.9±5.2	-73.7±7.1	-75.4±3.5
PAD3/GSK199	-65.3±2.3	-63.0±6.0	-62.1±5.1	-63.0±4.5
PAD4/GSK199	-74.7±4.8	-78.6±5.6	-77.4±6.7	-77.4±5.7

まとめ、今後の課題

PAD4 選択的阻害剤として報告されている化合物

GSK199 の標的選択性を明らかにするために、複合体構造のモデリング、分子動力学シミュレーション、相互作用解析、結合自由エネルギーの推定を行った。相互作用解析の結果、GSK199 の標的選択性には Phe634 との π - π スタッキングが影響していることが示唆された。PAD4/GSK199 複合体にのみ π - π スタッキングがあるため、結合自由エネルギーを調べた際に、PAD4/GSK199 複合体が最も安定であると考えられる。今後は GSK199 よりも PAD4 選択性が高くなる化合物の設計を行う必要がある。

参考文献

1. Schellekens, Gerard A., *et al.* "Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies." *The Journal of clinical investigation* 101.1 (1998): 273-281.
2. Lewis, Huw D., *et al.* "Inhibition of PAD4 activity is sufficient to disrupt mouse and human NET formation." *Nature chemical biology* 11.3 (2015): 189.

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 GPGPU による巨大地震の長周期地震動シミュレーション
 英文: Long-period ground motion simulation for mega-earthquakes using GPGPU

藤原広行
 Hiroyuki Fujiwara

国立研究開発法人 防災科学技術研究所
 National Research Institute for Earth Science and Disaster Resilience
<http://www.bosai.go.jp>

発生が想定されている巨大地震であってもその詳細な震源像を事前に予測しモデル化することは極めて困難であるため、震源像の不確実性を考慮した多数のシナリオに基づく地震動シミュレーションを行い、予測結果のばらつきを定量的に評価することが必要である。本研究では、主要な海溝型および活断層型地震である南海トラフ地震、糸魚川静岡構造線、および上町断層帯を対象として、申請者らが TSUBAME の GPU 環境に対応させた地震動シミュレータ GMS により地震波伝播シミュレーションを実施し、周期 1-2 秒以上の長周期地震動を計算した。47 ケースの震源破壊シナリオについてシミュレーションを実施し、それらの計算結果の統計処理によって長周期地震動ハザード評価を行った。

Long-period ground motion for the anticipated mega-earthquakes at the Nankai Trough and major active faults were simulated by the 3-D finite-difference computation of seismic wave propagation, using a software called the Ground Motion Simulator (GMS) adopted to GPU on TSUBAME. Simulations for a total of 47 cases were conducted, considering different rupture scenarios for each earthquake. Based on the statistical analysis of the simulation results, we evaluated the probabilistic seismic hazard for each earthquake.

Keywords: 長周期地震動, 三次元差分法, 南海トラフ地震, 活断層型地震

背景と目的

巨大地震による長周期地震動(長周期とは、ここでは周期 1 秒~20 秒程度以上とする)には、大きく分けて 2 種類の現象が知られている。一つは、地震波の長周期成分が減衰せずに遠方まで伝わり、100km~数 100km 程度離れた平野部で、振幅の大きい長周期の揺れが長時間続く現象である。2011 年東北地方太平洋沖地震発生時には、東京や、大阪においても観測され、超高層ビルでの揺れやエレベータ停止などが報告された。来るべき南海トラフ巨大地震の発生時には震源が陸地に近い分、大都市圏において長周期長周期地震動がさらに顕著になることが危惧される。

もう一つは、巨大地震の断層ごく近傍における断層破壊による長周期パルスである。1995 年兵庫県南部地震や 2016 年熊本地震で見られたように、内陸の活断層型地震発生時に、断層近傍の都市部において構造物に直接大きな被害を及ぼし、数 m の顕著な地表変位(地面のずれ)を伴うこともある。

これら長周期地震動のハザード評価(危険度評価)のためには震源破壊シナリオを模した震源モデルと 3 次元地下構造モデルを用いた波動伝播シミュレーションによる数値計算が有効である。巨大地震の震源モデルを事前に予測することは極めて困難であることから、破壊シナリオの不確実性を考慮して多数の震源モデルを設定し、個々の震源モデルに対する予測に加え、予測結果がどの程度のばらつきを持つかを定量的に評価することが重要となる。そのためには計算コストの高い大規模シミュレーションを多数行う必要がある。

本課題では、主要な海溝型地震と活断層地震である南海トラフ地震、糸魚川静岡構造線断層帯(静岡、長野、新潟県)および上町断層帯(大阪府)の地震について、TSUBAME の GPU 環境を利用することで多数回の大規模地震波伝播シミュレーションを実施し、それに基づいた地震ハザード評価を行った。南海トラフ地震と糸魚川静岡構造線断層帯の地震はマグニチュード 8 クラスの巨大地震である。上町断層帯の地震はマグニチ

ュード 7 程度であるが、大都市の直下に存在する活断層であり、注視すべき極めて重要な地震である。

概要

南海トラフ地震、糸魚川静岡構造線、上町断層帯の地震を対象として、震源パラメータの不確実性や破壊不均質性を考慮して複数ケースの震源モデルを設定した。図 1, 2, 3 にそれぞれの震源モデルの例を示す。

まず、南海トラフ地震については地震調査研究推進本部（地震本部）の長期評価（https://www.jishin.go.jp/main/chousa/kaikou_pdf/nankai_2.pdf）で例示された 15 の震源域のうち、マグニチュード 8 程度に相当する二つの震源域についてアスペリティの配置や破壊開始点を変えた計 7 通りを設定した。南海トラフ地震について同様の検討を試行してきたが、本年度は周期 1 秒以上の地震動をより適切に表現できるよう、破壊不均質性を付与した震源モデルに更新した。

糸魚川静岡構造線断層帯、上町断層帯については、地震本部の全国地震動予測地図（https://www.jishin.go.jp/evaluation/seismic_hazard_map/shm_report/shm_report_2018/）の断層モデルを基礎として、アスペリティは一、破壊開始点、断層上端深さ等を変えた 20 通りをそれぞれの地震に対して設定した。

地下構造モデルは防災科学技術研究所で作成している全国 3 次元地下構造モデルに地震本部によるプレート構造を付与した。それぞれの地震の計算領域の大きさ、差分法の格子サイズ、格子数を表 1 に示す。一般に、地表に近づくほど地震波速度が遅く、剛性率が小さくなるため、地表付近まで断層が達する活断層型の地震については計算精度が保たれるよう留意する必要がある。特に上町断層帯は大阪堆積盆地内に位置し、地震波速度が遅いため、安定して計算を行うために計算格子サイズや時間刻み幅（サンプリング）を調整した。

地震波伝播シミュレーションでは、地下構造モデルを水平方向に 8×6 に分割し、それぞれの部分領域に 1 つの GPU を割り付けて計算処理を行った。計算された地震動速度の東西、南北、上下の 3 成分をそれぞれ出力した。出力ファイルには、陸域に約 2km 間隔で設定した観測点（南海トラフの場合、約 5 万点）での毎タイムステップの速度値が記録されており、南海トラフの場合

1ファイルで約 40GB である。

TSUBAME3.0 の 48GPU (24 ノード) を使用した実行時間は平均 3 時間程度であった。

結果および考察

出力された速度値に対して、バンドパスフィルタ、ダウンサンプリングのポスト処理を行い、各ケースについて速度応答スペクトル (SV) や最大速度値 (PGV) の空間分布を求めた。図 4, 5, 6 に PGV 分布図の例を示す。南海トラフ地震や糸魚川静岡構造線断層帯の地震では、破壊シナリオ（例えば破壊開始点位置）によって地震動強さの空間分布が大きく異なることが分かる。一方、上町断層帯の地震は、堆積盆地構造をなす大阪平野内での地震動の増幅が大きい、詳細に見るとやはり破壊シナリオにより地震動強さの分布は異なる。

それぞれの地震について、代表地点における SV、PGV の条件付き超過確率を求めた。地震動の超過確率とは、ある地震（例えば上町断層帯の地震）が発生する条件下で、ある地点（例えば大阪市役所）における予測地震動はケースごとにばらつきがあるので、それぞれのケースの確からしさを考慮した場合の、ある一定の地震動強さ（例えば PGV で 50 cm/s）以上の揺れになる確率である。ここでは、それぞれのケースは同程度に確からしいと仮定した。横軸に地震動強さ、縦軸に超過確率をプロットするとハザード曲線になる。例として、図 7 に大阪市、神戸市、京都市における上町断層帯についての PGV のハザード曲線を示す。

まとめ、今後の課題

TSUBAME3.0 の GPU 環境を用いて、南海トラフ、糸魚川静岡構造線断層帯、上町断層帯の地震を対象とした多数の震源モデルに基づく長周期地震動シミュレーションを行い、震源モデルの違いによるばらつきを含んだ長周期地震動のハザード評価を行った。

今回は海溝型地震については震源モデルのアップデートを行い、また活断層型地震についても複数ケースの計算を始めることができた。今後、より計算効率を高め、ハザード評価に十分なばらつきを有する母集団となるケースの計算を進めたい。

表1 計算領域と格子数

	計算領域：南北 × 東西 × 深さ [km]	最小格子サイズ：南北 × 東西 × 深さ方向 [m]	格子数 [億個]	サンプリング [Hz]	タイムステップ [回]
南海トラフ	763 × 868 × 100	105 × 105 × 100	68	160	80600
糸魚川静岡構造線	411 × 422 × 100	105 × 105 × 100	14	160	49600
上町断層帯	322 × 261 × 100	45.5 × 45.5 × 100	46	260	49600

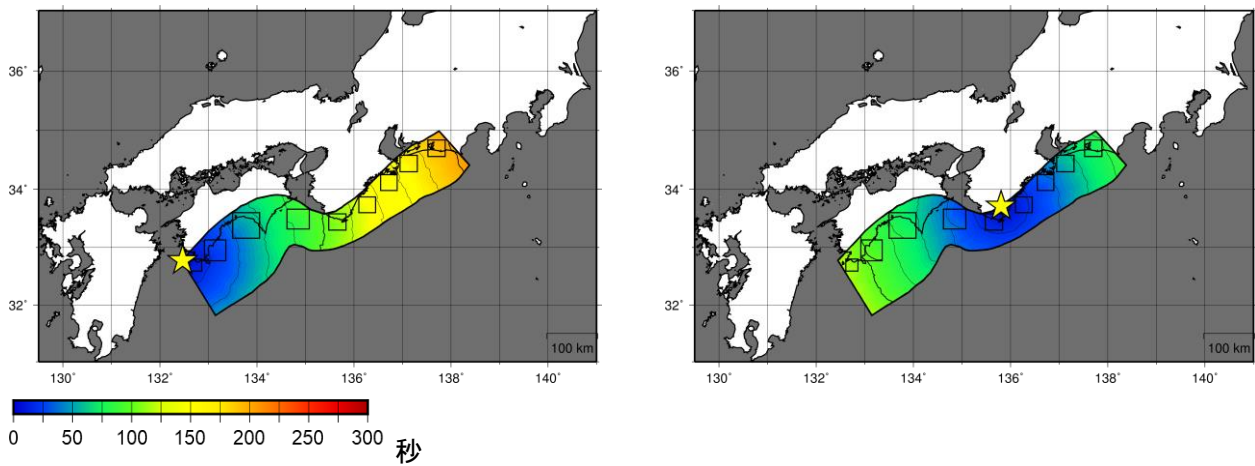


図1 南海トラフ地震の震源断層モデルの例. 色は破壊時刻分布を表す.

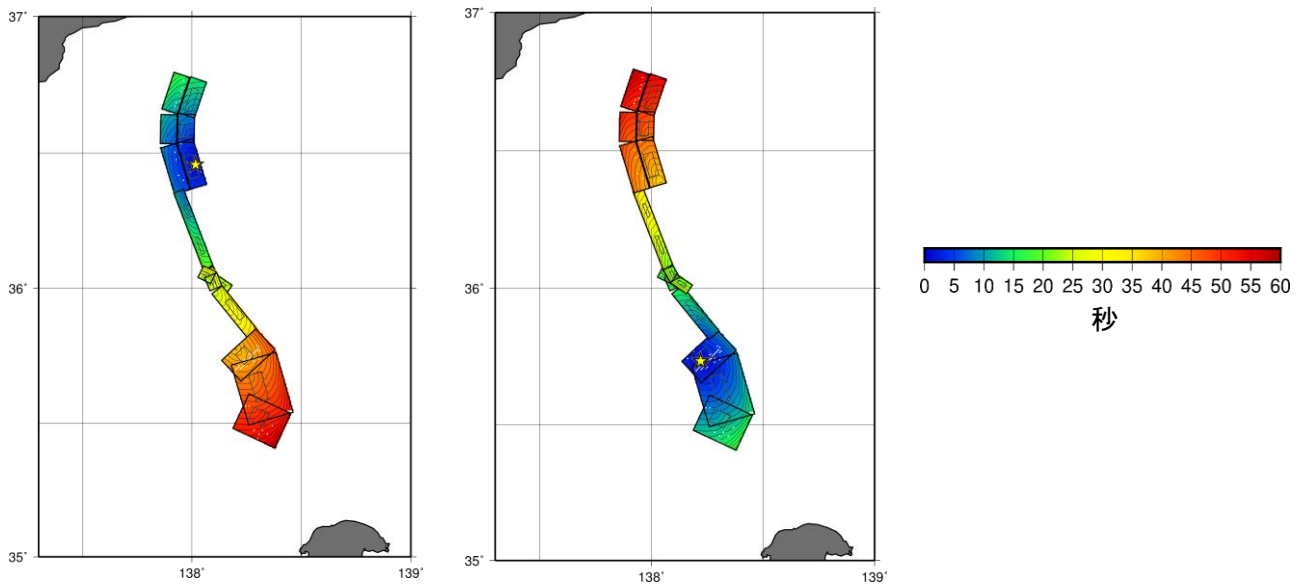


図2 糸魚川静岡構造線断層帯の震源断層モデルの例. 色は破壊時刻分布を表す.

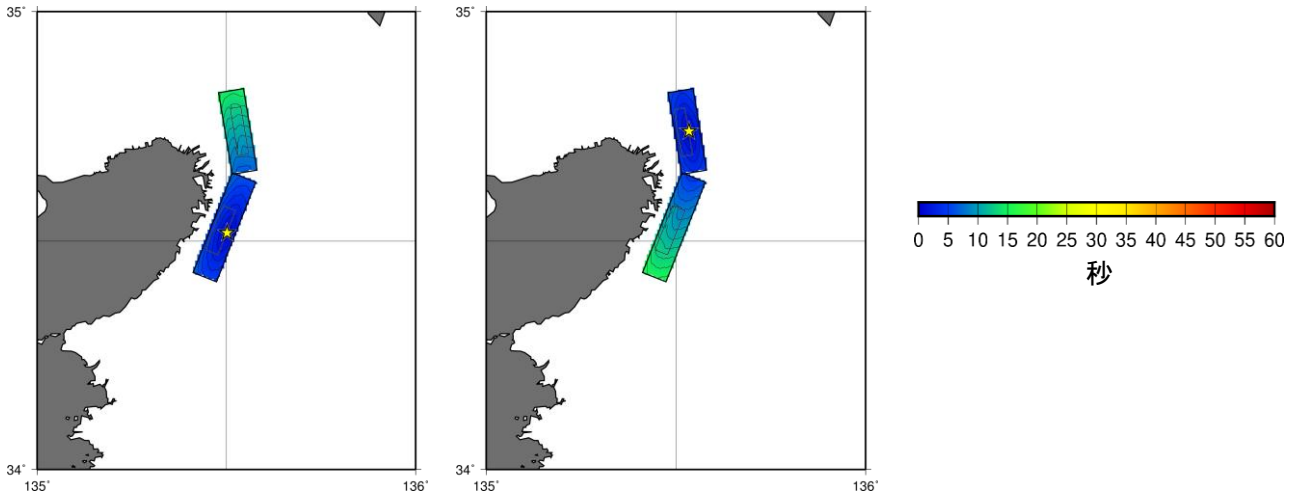


図3 上町断層帯の震源断層モデルの例. 色は破壊時刻分布を表す.

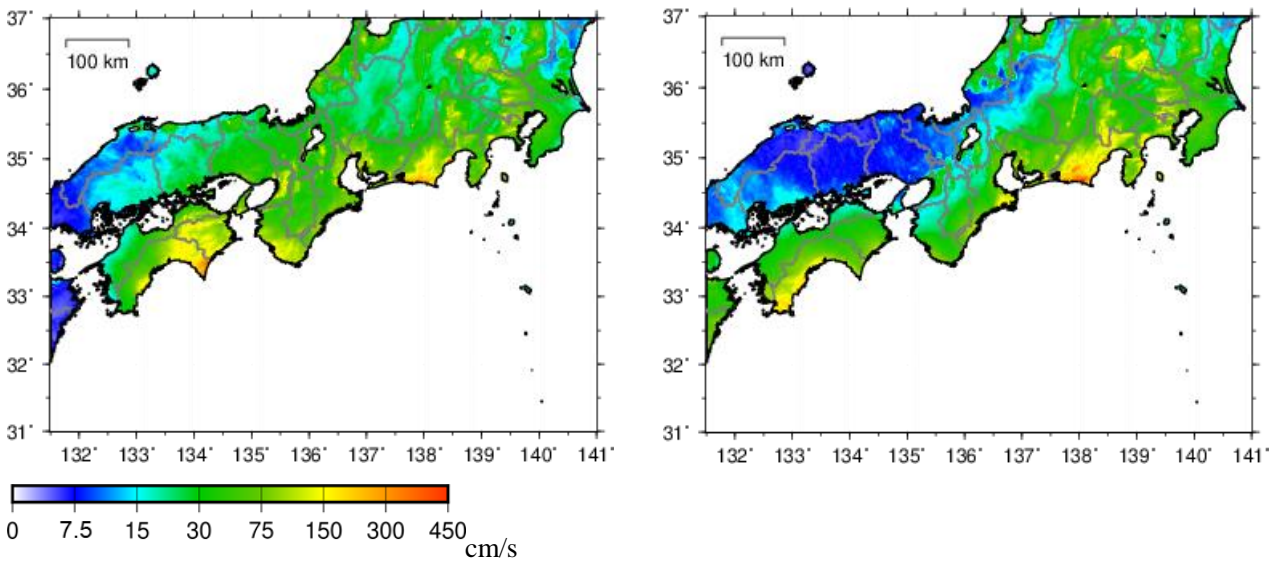


図4 南海トラフ地震の予測地震動(PGV 分布)の例.

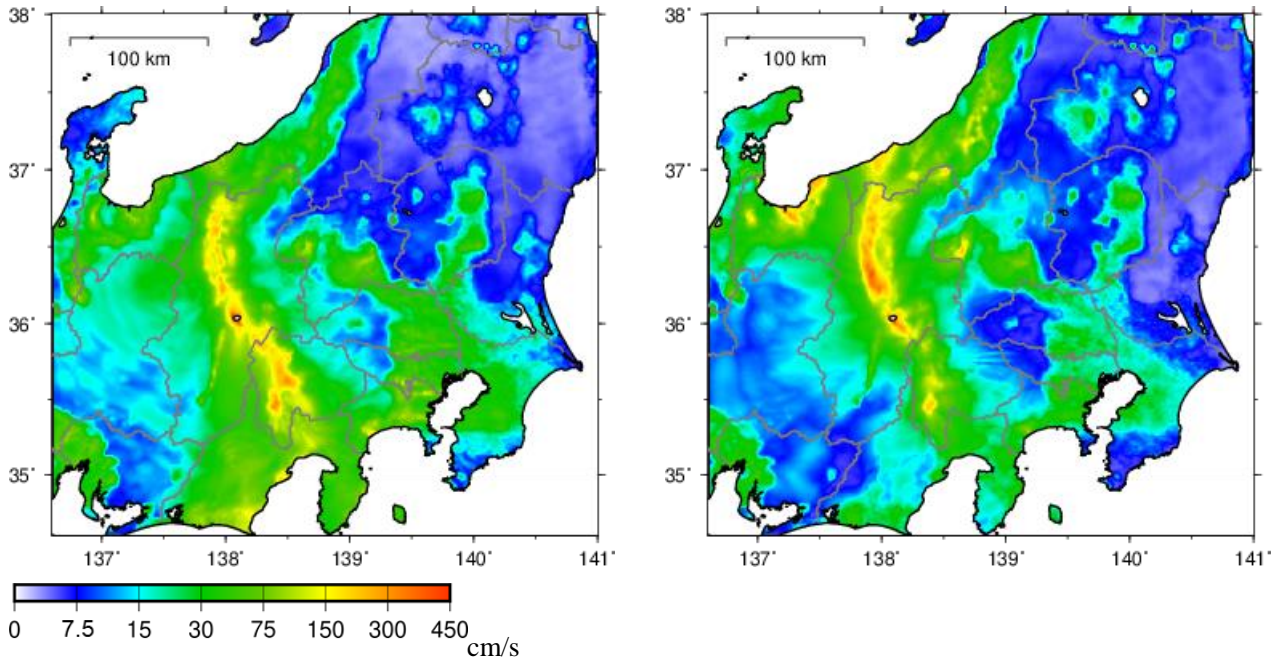


図5 糸魚川静岡構造線の予測地震動(PGV 分布)の例.

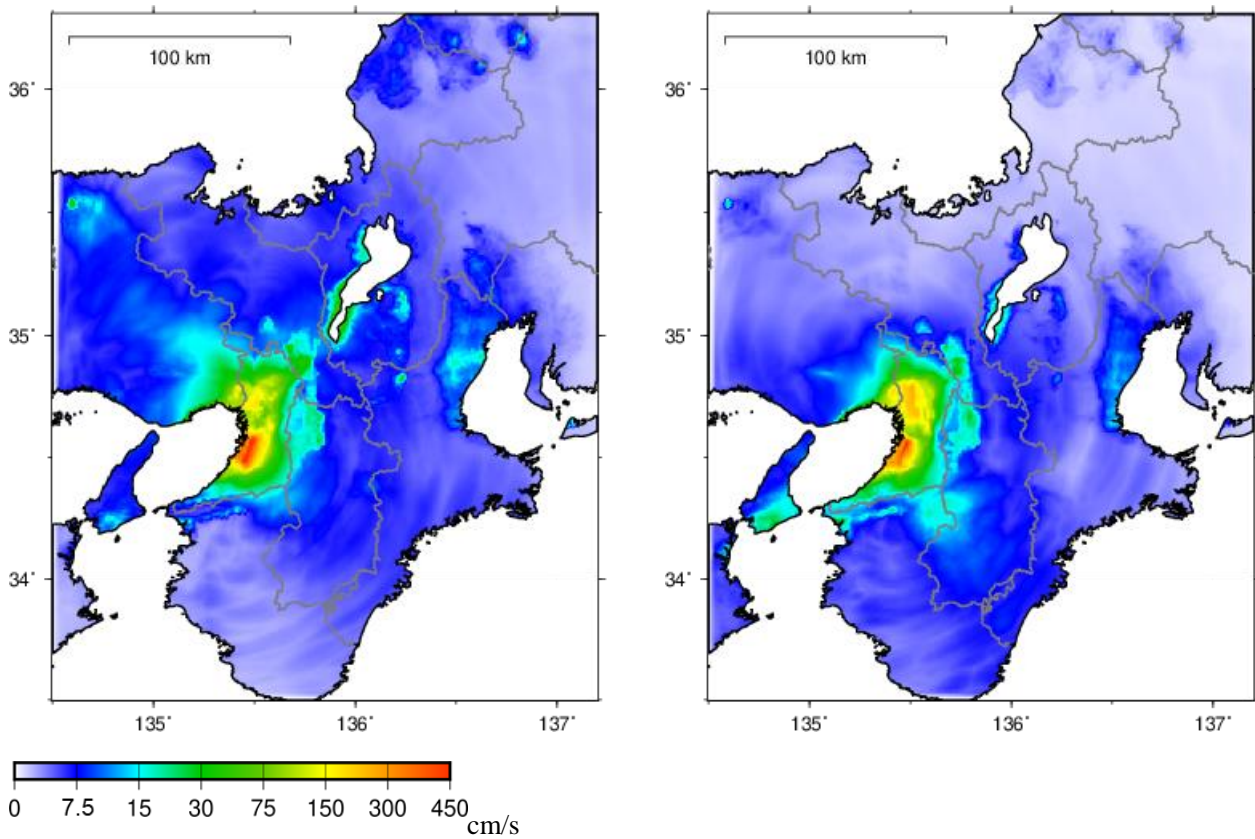


図6 上町断層帯の予測地震動(PGV 分布)の例.

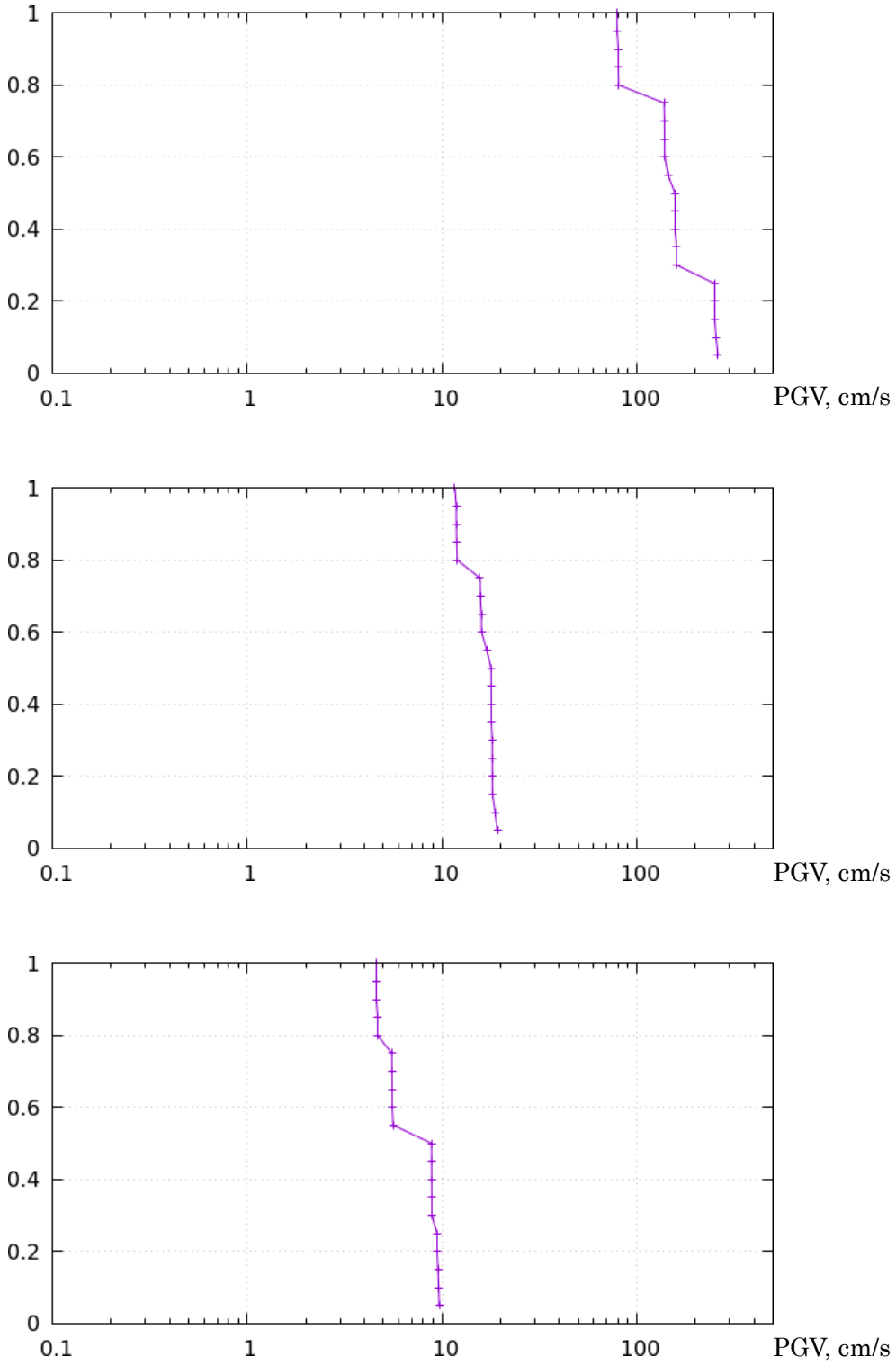


図7 上町断層帯の地震に対する主要地点(上から大阪市, 神戸市, 京都市)におけるPGVのハザード曲線.

TSUBAME 共同利用 令和元年度 TSUBAME 利用 成果報告書

利用課題名 計算化学による固体分散体中の薬物とキャリアの相互作用の評価

英文: Evaluation of Interaction between Drug and Carrier in Solid Dispersion
by Computational Chemistry

利用課題責任者 東 顕二郎

First name Surname Kenjirou Higashi

所属 千葉大学大学院薬学研究院製剤工学研究室

Affiliation Laboratory of Pharmaceutical Technology, Graduate School of Pharmaceutical Sciences,
Chiba UniversityURL <http://www.p.chiba-u.jp/lab/seizai/index.html>

邦文抄録(300 字程度)

近年、難水溶性薬物の溶解性改善手法として固体分散体が広く研究されており、新規添加剤である α -glucosyl rutin (Rutin-G)が注目されている。分子動力学シミュレーションにより、難水溶性薬物のモデルとして carbamazepine (CBZ)を用い、実験データを基にしたシミュレーション条件で CBZ/Rutin-G 固体分散体を調製した。その結果、シミュレーションで得られたガラス転移温度と実験から求めた値がほぼ一致したことから、固体分散体のシミュレーションに成功したと考えられた。フラグメント分子軌道計算から、固体分散体において、CBZとRutin-Gでは水素結合のみならず疎水的な相互作用が働いており、これが固体分散体の安定性に寄与していることが示唆された。

英文抄録(100 words 程度)

Nowadays, solid dispersions have been widely investigated to improve the solubility of poorly water-soluble drugs, and α -glucosyl rutin (Rutin-G), a new solid dispersion carrier, has been attracting an attention. Carbamazepine (CBZ) was used as a model of poorly water-soluble drugs, and CBZ/Rutin-G solid dispersion was calculated by molecular dynamics simulation under the conditions based on experimental data. The glass transition temperature obtained by the simulation almost agreed with the experimental one, declaring the success of solid dispersion simulation. Besides, fragment molecular orbital calculation indicated that not only hydrogen bonds but also hydrophobic interactions between CBZ and Rutin-G contributed to the stabilization of solid dispersion.

Keywords: 非晶質, 固体分散体, MD シミュレーション, FMO 計算

背景と目的

約 8 割の新規医薬品候補薬物は水に溶けにくく、これが経口製剤化において大きな障害となっている。難水溶性薬物の溶解性改善手法として薬物非晶質化が挙げられる。薬物の非晶質状態は結晶状態と比較しエネルギー的に高い状態であるため、高い溶解度及び溶解速度を示す。しかし、薬物の非晶質状態は熱的に不安定であるため、保存過程に再結晶化することで、溶解性が下がることが問題となる。そこで、薬物を非晶質状態で添加剤に分散させることで、薬物の非晶質状態を安定化する固体分散体を利用されている。

固体分散体の添加剤として、高分子担体が汎用されている。しかしながら、高分子担体の分子量は大きいいため、最終的な製剤の体積が大きくなってしまふ。我々の研究グループでは新規機能性食品添加剤である α -

glucosyl rutin (Rutin-G)を固体分散体に用いる添加剤として研究している。Rutin-G (Fig.1)は Rutin(ビタミン P)を酵素処理により glucosyl 化した化合物であり、疎水基である flavonol 部分と親水基である糖鎖を有する。これまでに、低細胞毒性と高溶解度を有する Rutin-Gを用い、種々の難水溶性薬物を非晶質化することで、薬物の溶解度及びバイオアベイラビリティを改善できることが報告されている。また、固体 NMR 測定の結果より、モデル薬物として carbamazepine (CBZ)を用いて調製した CBZ/Rutin-G 固体分散体において、CBZ のアミド基と Rutin-G の flavonol 部分間の相互作用形成が示唆された。しかしながら、固体 NMR のみでは、具体的な相互作用の評価が困難であった。

本研究では、分子動力学 (MD)シミュレーションにより、初めて CBZ と Rutin-G を含む固体分散体の構造

を分子レベルで明らかとした。さらに、フラグメント分子軌道(FMO)計算から、CBZとRutin-G間の相互作用エネルギーを評価し、固体分散体の安定化メカニズムを解明した。

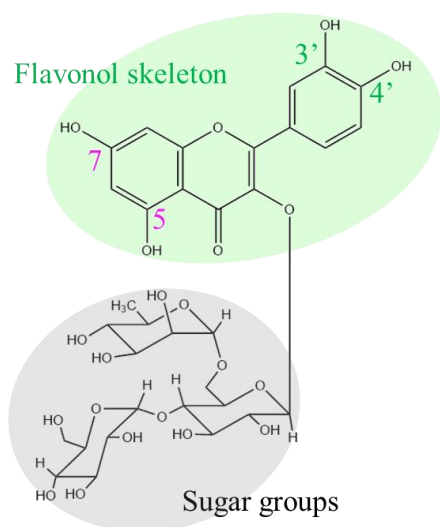


Fig. 1. The chemical structure of Rutin-G.

概要

1. MDシミュレーションによる固体分散体の評価

一辺 50 Å の立方体のセルに 50 分子の CBZ と 80 分子の Rutin-G をランダムに配置させ、周期境界条件を設定した。そして、セル全体をエネルギー極小化させ、初期構造 (Fig. 2) とした。Melt-quench 法を想定し、複数の異なる冷却条件 (Fig. 3) で MD シミュレーションを行った。MD 計算は NAMD プログラムを用い、力場を MMFF94x、カットオフ距離を 8-10 Å、時間刻みを 1 fs、NTP アンサンブルとした。CBZ/Rutin-G 固体分散体中の全ての水素をエネルギー極小化させ、CBZ と Rutin-G 間の水素結合を解析した。

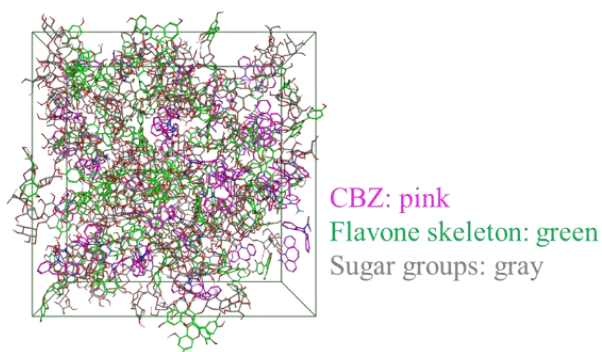


Fig. 2. The initial structure of CBZ/Rutin-G mixture.

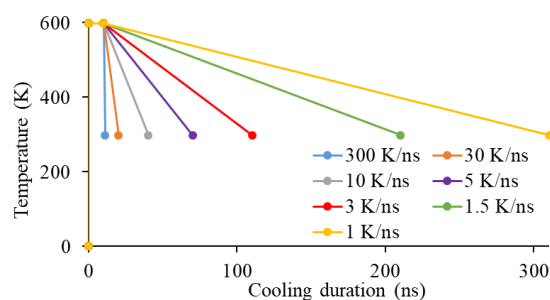


Fig. 3. Different protocols using in MD simulation.

2. FMO 計算による固体分散体の評価

MD シミュレーションで得られた固体分散体から、各 CBZ 分子及びこの CBZ 分子の 10 Å 以内の全て分子を含む構造をそれぞれ取り出し、FMO 計算を実施した。FMO 計算には、ABINIT-MP プログラムを用い、計算手法を MP2/6-31G 法とした。PIEDA (相互作用エネルギー成分分割解析) 計算により、得られた相互作用エネルギーをさらに静電、交換反発、電荷移動、分散相互作用エネルギーに分割した。

結果および考察

1. MDシミュレーションによる固体分散体の評価

Table 1 に示したように、冷却速度は 3 K/ns 以下に設定すると、固体分散体は同様な体積を示すことが明らかとなった。計算コストを考慮し、以降の検討では冷却速度を 3 K/ns と設定し、固体分散体の MD シミュレーションを行うこととした。

Table 1. The average system volume depending cooling rate during equilibration process (n=3, mean±S.D.)

Cooling rate (K/ns)	Volume (nm ³)
300	104.8 ± 0.3
30	103.9 ± 0.8
10	102.7 ± 0.6
5	103.2 ± 0.2
3	102.3 ± 0.4
1.5	102.1 ± 0.2
1	102.4 ± 0.2

MD シミュレーションの冷却過程において、CBZ/Rutin-G 固体分散体のガラス転移が認められた。このガラス転移温度 (T_g) は 169°C であり、示差走査熱量測定により実験的に求めた $T_g = 149°C$ と近い

値を示した。

Fig. 4 に固体分散体中の CBZ と Rutin-G 間の水素結合の統計結果を示した。CBZ は主に Rutin-G の糖鎖と水素結合を形成することが示唆された。

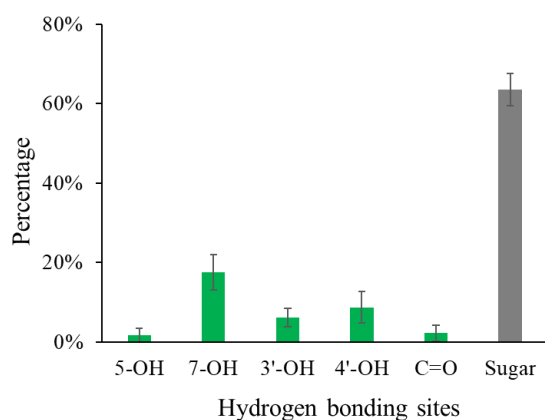


Fig. 4. Percentage for different hydrogen bonding sites of Rutin-G with CBZ in solid dispersion (n = 3, mean mean±S.D.).

2. FMO 計算による固体分散体の評価

FMO 計算により、固体分散体中の 50 分子の CBZ と周囲の分子間の相互作用エネルギーを計算した。CBZ は Rutin-G の糖鎖との相互作用エネルギーが最も大きいことが明らかになった。また、PIEDA 法により、分割した相互作用エネルギーを Fig. 5 に示した。一般的に、水素結合は静電力や電荷移動力が主要成分であるのに対して、CH/ π や π - π などの疎水的な相互作用は分散力が主な成分である。固体分散体においては CBZ と Rutin-G の間で水素結合のみならず、疎水的な相互作用も形成されており、これらが固体分散体の安定化に寄与していると推察した。

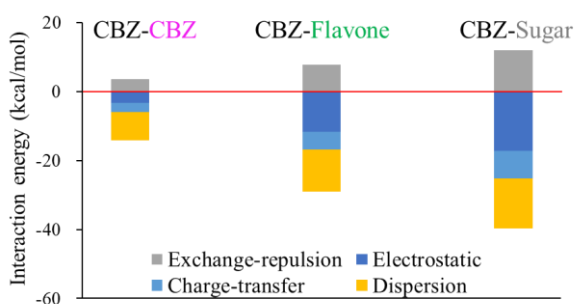


Fig. 5. The average interaction energy between CBZ and different fragments (n = 50).

まとめ、今後の課題

本研究では、MD シミュレーション及び FMO 計算に

より、固体分散体の調製及び解析を行った。CBZ/Rutin-G 固体分散体においては Rutin-G の flavonol 部分および糖鎖との相互作用により CBZ の非晶質が安定化されていることが示唆された。今後、冷却過程における CBZ と Rutin-G 間の相互作用形成様式を解析し、本固体分散体の安定化機構をより詳細に検討する予定である。今後は、本手法を様々な薬物/添加剤固体分散体に適用し、本手法の有効性・妥当性を検討する必要がある。将来的には、本手法を固体分散体の薬物と添加剤の最適組成のスクリーニング法として確立することが期待される。

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 サブナノ粒子の構造及び性質に関する理論的研究
 英文: Theoretical Study on Structures and Properties of Subnano Particles

利用課題責任者 春田 直毅
 First name Surname Naoki Haruta

所属 京都大学 福井謙一記念研究センター
 Affiliation Fukui Institute for Fundamental Chemistry, Kyoto University
 URL <https://sites.google.com/site/nharuta1987/>

邦文抄録(300字程度) サブナノ粒子は粒径 1 nm 前後の極小粒子であり、第 3 の材料として注目を浴びている。本研究課題では、様々なサイズの酸化スズサブナノ粒子をモデリングし、密度汎関数理論を用いることで、これまで未解明であったサブナノ粒子特有の振動分光スペクトルの帰属に成功した。ピーク位置のサイズ依存性は、表面の配位不飽和性、構造歪み、OH 基表面修飾が粒子サイズの減少に伴って増加し、Sn-O 結合の強度が弱まることが原因であることが分かった。この Sn-O 結合の強度変化は、酸素原子供給能の変化を表しており、実験で報告されている CO 酸化に対する酸化スズサブナノ粒子の高い触媒活性を説明する。

英文抄録(100 words 程度) Subnano particles (ca. 1 nm) possess completely different physical and chemical properties as compared to bulk and nanoscale materials. In this project, we achieved the successful assignment of experimental vibrational spectra for subnano tin oxide by means of density functional theory calculations. According to our computational result, particle size dependence on the peak position is related to the weakening of Sn-O bonds on the particle surface. This is consistent to the experimentally reported high catalytic activity for CO oxidation.

Keywords: subnano, nano, tin oxide, Raman spectroscopy, density functional theory

背景と目的

サブナノ粒子は粒径 1 nm 前後の極小粒子であり、バルクやナノ粒子と全く異なる光学・磁気・触媒特性を有することから、第 3 の材料として注目を浴びている。サブナノ粒子を「使える材料」として作り出す技術は長い間存在しなかったが、近年になって精密合成が可能となり、それ以来、大規模な機能探索が行なわれている。しかし、トライアルアンドエラーや経験則に基づく研究開発が続けられており、ab initio 計算による理解と予測の方法論の確立が課題となっていた。

本プロジェクトでは、これまで未解明であったサブナノ粒子特有の振動分光スペクトルの帰属問題を合理的なモデリングと密度汎関数理論を用いた理論計算によって解決し、サブナノ粒子が示す特異な触媒活性との関連性を解明するという成果を得た。

概要

最近、Kuzume らのグループによって、酸化スズサ

ブナノ粒子の超高感度ラマン計測が行われた(図1)。それによると、酸化スズサブナノ粒子はバルクとは全く異なるスペクトル形状を示すとともに、粒子サイズが小さくなるにつれて、ピーク位置が低波数シフトすることが分かった。また一方で、Inomata らにより、酸化スズをサブナノ化すると、CO 酸化に対する触媒活性が著しく向上することも報告されている。

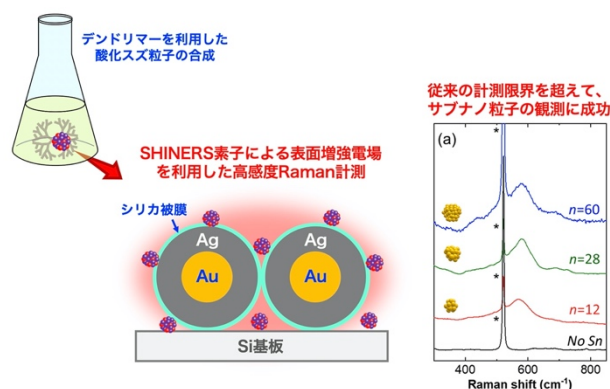


図1 酸化スズサブナノ粒子のラマン分光測定

本研究課題では、XPS 測定で報告されている酸化数に基づき、様々なサイズの酸化スズサブナノ粒子 Sn_{12}O_x , Sn_{28}O_x , Sn_{60}O_x を合理的にモデリングし、密度汎関数理論を用いて、これらのラマンスペクトルの理論シミュレーションを行った。

結果および考察

我々の計算によれば、粒子表面が水和したモデルを採用することで、実験スペクトルが完全に再現されることが分かった(図2)。一方で、水和していないモデルは焼成したサブナノ粒子の実験スペクトルを再現した。

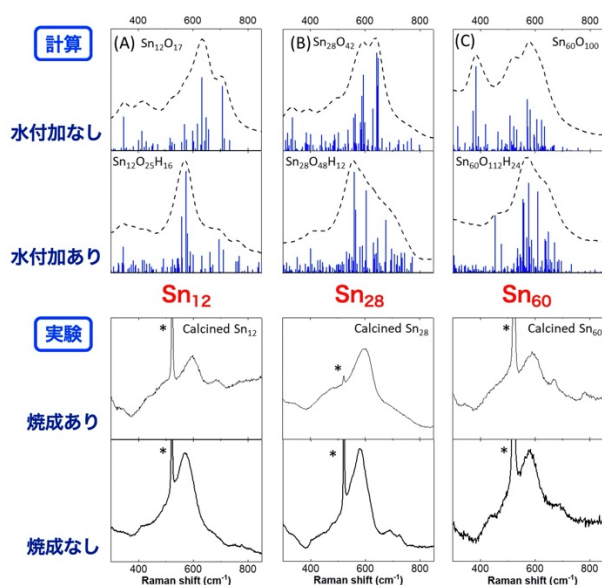


図2 酸化スズサブナノ粒子のラマンスペクトル

ラマン測定の実験報告によれば、粒子サイズが小さくなるほど、スペクトルのピーク位置が低波数シフトする(図3)。これは、表面の配位不飽和性、構造歪み、OH基表面修飾が粒子サイズの減少に伴って増加し、Sn-O結合の強度が弱まることの原因であることが分かった。このSn-O結合の強度変化は、酸素原子供給能の変化を表しており、CO酸化に対する高い触媒活性の原因であると考えられる。

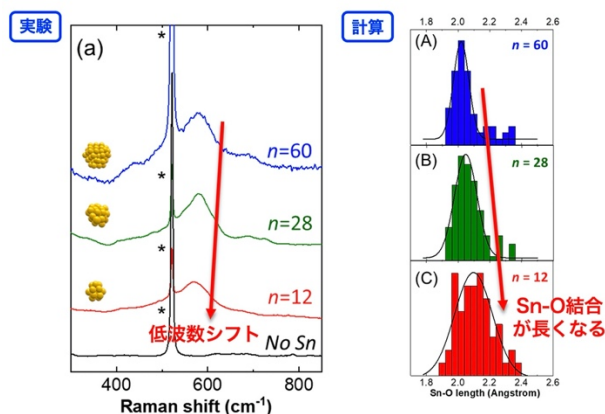


図3 ラマンピークシフトとSn-O結合距離

まとめ、今後の課題

今回、これまで未解明であったサブナノ粒子特有の振動分光スペクトルについて、合理的なモデリングと密度汎関数理論を用いた理論計算によって再現し、帰属することに成功した。同時に、サブナノ粒子が示す特異な触媒活性との関連性を解明することができた。

今後は、さらに広汎なサブナノ粒子について振動分光スペクトルと触媒機能の解析を行い、ab initio 計算の立場からサブナノサイエンスの確立を目指す。

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 ダイヤモンド NV 中心の励起状態に関する理論的研究
 英文: Theoretical Study on Excited States of NV Centers in Diamond

利用課題責任者 春田 直毅
 First name Surname Naoki Haruta

所属 京都大学 福井謙一記念研究センター
 Affiliation Fukui Institute for Fundamental Chemistry, Kyoto University
 URL <https://sites.google.com/site/nharuta1987/>

邦文抄録(300 字程度) ダイヤモンド窒素空孔中心(NV 中心)は、ダイヤモンド中に生じる格子欠陥の一種であり、光応答性のある安定なスピンを持つことから、様々な応用が期待されている。本プロジェクトでは、NV 中心の特異な発光特性の起源を明らかにするため、時間依存密度汎関数理論を用いた励起状態計算と振電相互作用解析を行った。NV 中心の縮退した SOMO は主軸と直交する方向にのみ大きく非局在化し、これが異方性のある強い電子遷移の原因であることが分かった。また、この異方的な軌道が原因となって、電子状態と表面水素の振動が相互作用し、発光量子収率を低下させていることも明らかとなった。この結果は、主軸と直交する方向における異種元素ドーピングや表面修飾によって、NV 中心の光学特性が大きく改善される可能性を示唆する。

英文抄録(100 words 程度) Diamond nitrogen-vacancy center (NV center) is a defect in a diamond. It has optically sensitive stable spins, and thus its various applications are expected. In this project, we performed excited state calculations and vibronic coupling analyses based on time-dependent density functional theory calculations to clarify the origin of its unique photoluminescent property. According to our computational result, its degenerate SOMOs are delocalized only in the orthogonal direction to the principal axis. This is the origin of its strong anisotropic electronic transition. Further, such SOMOs cause vibronic coupling with surface hydrogen vibrations, lowering its photoluminescence yield. This could be improved by doping or surface modification.

Keywords: diamond NV center, excited state, density functional theory, vibronic coupling

背景と目的

ダイヤモンド窒素空孔中心(NV 中心)は、ダイヤモンド中に生じる格子欠陥の一種であり、光応答性のある安定なスピンを持つことがその特徴である(図1)。磁場、電場、温度の高感度計測や量子情報技術への応用が期待されており、近年、精力的な研究開発が行なわれている。しかし、ab initio 計算に基づく理論評価の報告例は極めて少なく、基礎物性の向上や新たな応用開発を目的として、現在もトライアルアンドエラーによる探索が行われていることが問題とされてきた。

本プロジェクトでは、未解明であった NV 中心の特異な発光特性の起源を時間依存密度汎関数理論を用いた励起状態計算と振電相互作用解析によって解決し、NV 中心の光学特性向上のための新たな設計指針を得るという成果を得た。

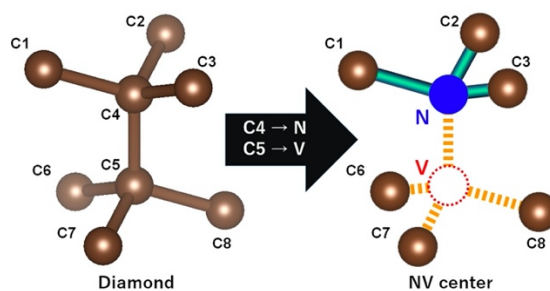


図1 ダイヤモンドと NV センター

概要

本研究課題ではまず、NV 中心の site symmetry を保持する C_{3v} 対称のモデルクラスターを系統的に作成する方法論を確立した。これらのモデルクラスターに対して、時間依存密度汎関数理論を用いた励起状態計算を行い(図2)、さらに振電相互作用解析を行った。

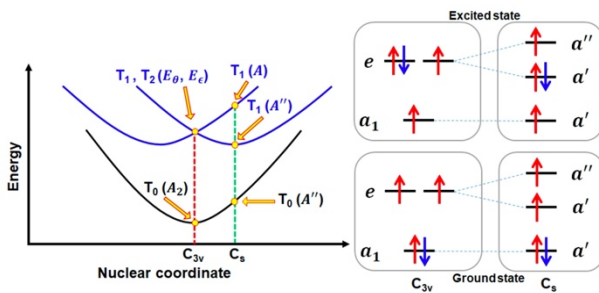


図2 NV センターの電子構造

結果および考察

NV 中心を含む最小単位炭素クラスターのまわりに、1~4 層の炭素層を加え、ダングリングボンドを水素原子で終端することで、 $C_{42}H_{42}N$, $C_{140}H_{98}N$, $C_{316}H_{174}N$, $C_{598}H_{270}N$ を作成し、拡張 Hückel 計算を行った。そうしたところ、SOMO-1 の非縮退軌道はクラスター全体に非局在化するのに対し(図3)、SOMO の縮退軌道は主軸と直交する方向にのみ大きく非局在化することが分かった(図4)。この結果は、密度汎関数理論を用いた計算でも同様であった。

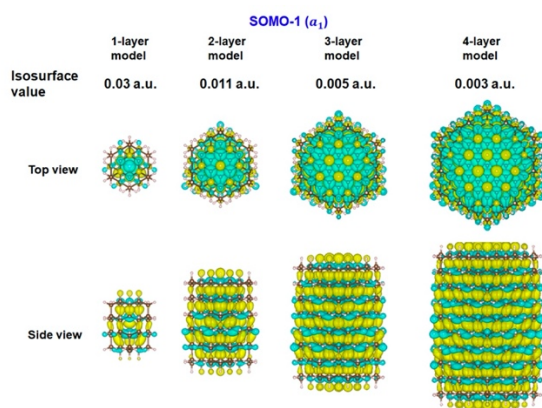


図3 NV 中心の SOMO-1 軌道

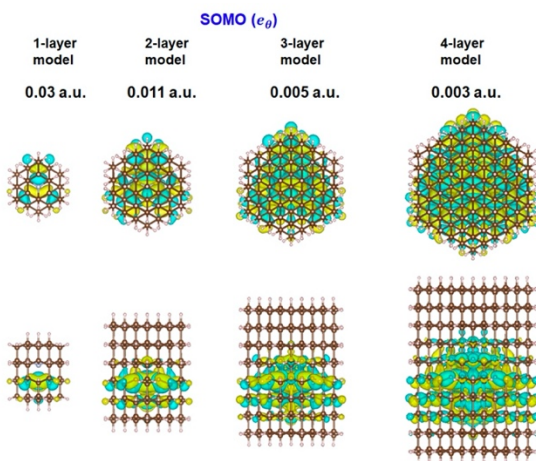
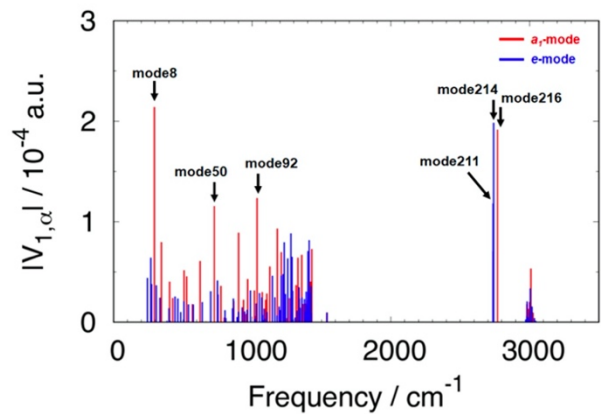


図4 NV 中心の縮退した SOMO の 1 つ

次に、時間依存密度汎関数理論を用いた励起状態計算を行ったところ、異方的に非局在化した SOMO が、異方性のある強い電子遷移の原因であることが分かった。同時に、この異方的な軌道が原因となって、電子状態と表面水素の振動が相互作用し、発光量子収率を低下させていることも明らかとなった(図5)。この結果は、主軸と直交する方向における異種元素ドーピングや表面修飾によって、NV 中心の光学特性が大きく改善される可能性を示唆するものである。

図5 NV 中心の最低励起状態 T_1 における対角振電子相互作用定数

まとめ、今後の課題

今回、時間依存密度汎関数理論を用いた励起状態計算と振電相互作用解析によって、NV 中心の異方性のある強い電子遷移の原因、発光量子収率の低下要因を明らかにすることに成功した。さらに、異種元素ドーピングや表面修飾によって、NV 中心の光学特性が大きく改善される可能性も明らかになった。

今後は、作成したモデルクラスターに対して、実際に異種元素ドーピングや表面修飾を行うことで、光学特性がどのように変化するかを理論計算から明らかにし、NV 中心の物性向上やデバイス開発のための合理的な予測と提案を行うことを目指す。

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 流動摩擦抵抗予測法の構築に向けた粗面乱流直接数値解析

英文: Direct Numerical Simulation of rough wall turbulence toward the prediction of rough wall skin friction

利用課題責任者: 桑田 祐丞

First name Surname: Yusuke Kuwata

所属: 大阪府立大学

Affiliation: Osaka Prefecture University

URL: <http://www2.me.osakafu-u.ac.jp/htlab/>

邦文抄録(300 字程度)

粗さの特徴量であるひずみ度や有効勾配の複合的な影響を調査するために、粗面の高さを固定して粗さの特徴量を系統的に変化させた粗面を用いて、粗面乱流の直接数値解析を行った。解析は3次元27方向速度の多緩和時間格子ボルツマン法を用いて行い、局所細密格子法によって計算コストの低減を図った。ひずみ度が正で有効勾配が大きな粗面、つまり急峻な起伏を持つ山が支配的な粗面において、壁面摩擦係数が大きくなることが分かった。また、ひずみ度が壁面摩擦抵抗に与える影響は有効勾配の大きさに関わらず現れるが、急峻な起伏を有する有効勾配が大きな粗面において、ひずみ度の影響が強く表れることが明らかになった。

英文抄録(100 words 程度)

Combined effect of the skewness and effective slope of the rough surface is studied by means of the direct numerical simulations of turbulence over systematically varied rough surfaces. The simulations are performed by the D3Q27 multiple-relaxation-time lattice Boltzmann method, and zonal grid refinement technique is used to circumvent the high cost of simulating rough wall turbulence. The skin friction coefficient of the rough surface is found to be larger for the surface with the positive skewness and large effective slope value, which represents the steep peak-dominated surface. The influence of the skewness on rough wall turbulence emerges irrespective of the effective slope value; however, the influence of the skewness is more substantial for the surface with the large effective slope.

Keywords: rough wall turbulence, lattice Boltzmann, grid refinement, surface topology, direct numerical simulation

背景と目的

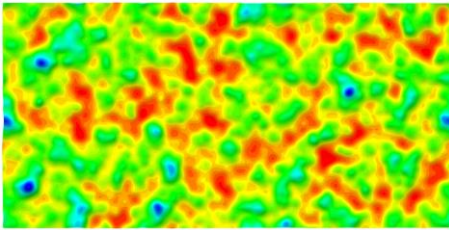
一般的に壁面は流体力学的に滑らかとみなすことができず、たとえ微小な粗さであっても粗さによって流動抵抗が増えることが知られている。表面粗さは表面仕上げ等の加工によって生じるものに加えて、表面の腐食、塗料の劣化などの経年劣化によって引き起こされ、これらは流体機械の性能を低下させることが知られている。表面粗さは、根二乗平均粗さなどの高さに関連する量に加えて、粗さ要素の確立密度分布に関連するひずみ度や波長に関連する有効勾配などによって特徴づけられ、粗さに起因する流動抵抗の増加もこれらの特徴量の影響を強く受けることが報告されている[1, 2]。しかし、既往の研究の多くはこれらの特徴量の個別の影響を調べたもの

が多く、特徴量の複合的な影響を議論された例は少ない。本研究では、これらの特徴量を系統的に変えた粗面を用いて、粗面乱流の直接数値解析を実施する。得られた結果より、これらの特徴量が壁面摩擦係数増大に与える複合的な影響について議論を行う。

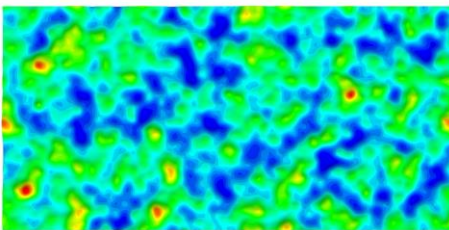
概要

粗面乱流に関する研究の多くは、粗面の幾何的な特徴量の1つに着目し、摩擦抵抗等のスケーリングを行うものがほとんどである。しかし実際には、様々な粗面の特徴量が複雑に影響しあうことで、摩擦抵抗が変化するため、粗面乱流理解するためには

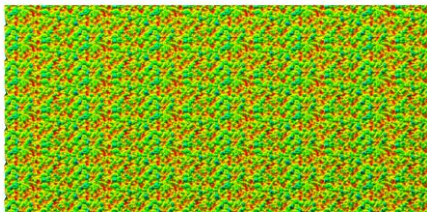
(a) ひずみ度 +0.5, 有効勾配0.1



(b) ひずみ度 -0.5, 有効勾配0.1



(c) ひずみ度 +0.5, 有効勾配0.6



(d) ひずみ度 -0.5, 有効勾配0.6

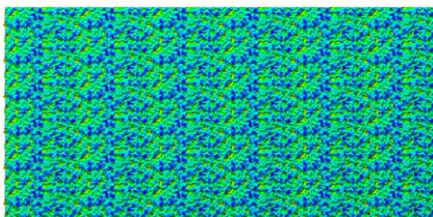


図1 粗面高さを固定してひずみ度と有効勾配を変化させた粗面

個々の粗面特徴量の影響を調査するのみでは不十分である。本研究では、粗面乱流などの複雑な構造の壁面を高効率・高精度に扱うことが可能であるD3Q27多緩和時間格子ボルツマン法プログラム[3]を用いて、粗面乱流構造を直接計算格子で解像する

直接数値解析を実施する。解析対象の粗面は図1に示すようなランダム粗面である。この粗面は様々なサイズの双曲関数型の回転体から構成されており、数値的に生成したものである。オリジナル粗面(図1 a)の粗面高さを反転することで、ひずみ度の符号が反転した粗面(図1 b)を作成した。また、粗面幅を縮小した粗面を敷き詰めることで有効勾配を大きくした粗面(図1 c,d)を作成した。それぞれのひずみ度は0.5, -0.5となり、有効勾配は0.1と0.6とした。なお、有効勾配0.1は波状面として分類され、0.6では粗面として分類される[2]。

解析系は図2に示すように2次元平行平板流の半分の領域を模擬したオープンチャンネルを対象とし、下壁には粗面、上壁には滑り条件を適用しており、主流(x)・スパン方向(z)には周期境界条件としており、主流方向に圧力勾配を付与することで流体を駆動した。半チャンネル幅と摩擦速度を基にしたレイノルズ数で600として解析を行った。局所細密格子法を用いて粗面近傍には解像度の高い格子を用いており、総計算格子点数は約3億点となった。なお、解析コードはMPI並列を行うことで複数GPU解析が可能であり、1ケースあたり28並列で解析をおこなった。

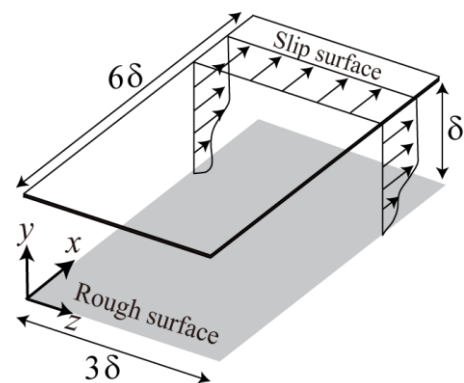


図2 粗面オープンチャンネル流れの解析系

結果および考察

粗面近傍における瞬時の乱流渦を可視化した結果を図3に示す。粗面近傍より活発な乱流渦が生成されており、微細な渦構造まで格子解像されていることが確認できる。粗面乱流では、主流方向に伸びたストリーク構造の発達には粗さによって妨げられるために、一般的な滑面乱流で見られるような乱流渦構造と比べて、秩序性が失われた乱流渦構造となっている。

つぎに、図4に粗面の摩擦速度で無次元化した平均速度分布を示す。比較の為、滑面の直接解析結果[4]を示している。どの粗面ケースにおいても、粗面壁近傍で主流速度が減速しており、平均速度分布は滑面の結果と比べて下方にシフトしていることが分かる。これは、粗面に生じる摩擦抵抗の増大を意味しており、ひずみ度が正のケース(山が支配的な粗面)と有効勾配が大きなケース(急峻な粗面)で特に顕著にみられることが分かる。

最後に、粗面の壁面摩擦係数を滑面の値で除した結果を表1に示す。壁面摩擦係数はどのケースにおいても増加しており、ひずみ度が正の有効勾配が大きなケースでは摩擦係数は滑面と比べて2.8倍程度になっていることが分かる。有効勾配に関わらず、ひずみ度が正のケースでは負のケースに比べて壁面摩擦係数が大きい。有効勾配が小さなケースではひずみ度の正負で10%ほどしか摩擦係数が変化しないのに対して、有効勾配が大きなケースではその値は30%ほどになる。つまり、有効勾配が大きい

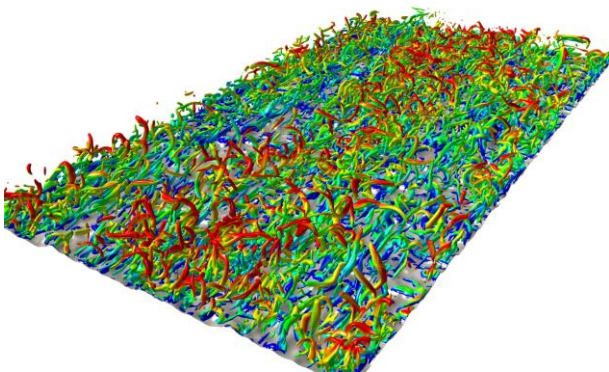


図3 粗面近傍における瞬時乱流渦構造の可視化結果。コンターは瞬時主流速度を表しており、赤は高速領域、青は低速領域を示している。

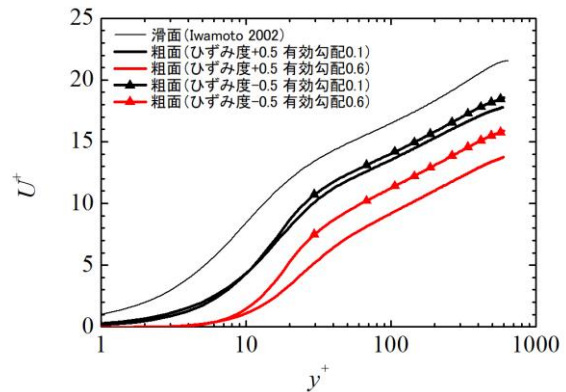


図4 摩擦速度で無次元化した主流平均速度分布。比較の為、滑面の直接数値解析結果[4]を示している。

表1 滑面との壁面摩擦係数比

	有効勾配 0.1	有効勾配 0.6
ひずみ度 +0.5	1.49	2.82
ひずみ度 -0.5	1.47	2.15

なケースほどひずみ度の影響がより顕著に表れることが明らかになった。

まとめ、今後の課題

粗面の特徴量であるひずみ度や有効勾配を系統的に変えた粗面を用いて、粗面乱流の直接数値解析を実施、特徴量の複合的な影響について議論を行った。ひずみ度が正で有効勾配が大きな粗面、つまり急峻な起伏を持つ山が支配的な粗面において、粗面壁の摩擦抵抗が大きくなることが分かった。また、ひずみ度の影響は有効勾配に関わらず現れるが、急峻な起伏を有する有効勾配が大きな粗面において、ひずみ度の影響が強く表れることが明らかになった。今後は、これら粗面特徴量の複合的な影響を取り入れることで、粗面効果を予測する相関式の構築を行う。

参考文献

[1] Flack, Karen A., and Michael P. Schultz. Journal

of Fluids Engineering 132.4 (2010).

[2] Napoli, E., Vincenzo Armenio, and M. De Marchis.

Journal of Fluid Mechanics 613 (2008): 385-394.

[3] Suga, Kazuhiko, et al. Computers & Mathematics with Applications 69.6 (2015): 518-529.

[4] Iwamoto, K., Suzuki, Y. & Kasagi, N. 2002

Database of fully developed channel

flow-thtlab internal report no. ILR-0201, THTLAB,

Dept. of Mech. Engng., The Univ. of Tokyo.

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 深層学習による高解像度化手法の水文学への応用
 英文: Application of Deep Learning Super Resolution Methods to Hydrologic Engineering

利用課題責任者
 Kei Ishida

所属
 Kumamoto University
<https://www.fast.kumamoto-u.ac.jp/home/>

邦文抄録

近年、豪雨・洪水や渇水が世界各地で甚大化・頻発化しており水文学の重要性が増している。しかしながら、既存の手法では現在様々な要因により解析精度・予測精度に限界がある。そこで、近年発展が著しい深層学習に注目する。本課題においては、まず豪雨に多大な影響をもたらす海水面温度を対象とし、深層学習の従来及び新しい高精度化手法を試し比較を行った。低解像度の海水面温度データは ERA20C から得た。また、高解像度の海水面温度データとして OISST を用いた。深層学習手法として、従来の高解像度化手法である SRCNN、新しい CNN 系及び GAN 系の手法である RRDBNet 及び ESRGAN を用いた。精度の評価手法として、RMSE 及び誤差率を用いたところ RRDBNet が最も良い結果を示した。

英文抄録

Hydrology is getting important now because precipitation and flood are getting severe. However, the accuracy of existing approaches in hydrology are not always sufficient for analysis or prediction. The accuracy of analysis and predication should be improved. Deep learning methods are nowadays getting popular in various research fields. This study used deep learning methods to improve the resolution of sea surface temperature data. As the deep learning methods, SRCNN, RRDBNet, and ESRGAN were selected. According to RMSE and error rates, RRDBNet showed the best results among the three methods.

Keywords:

Convolutional Neural Network, Generative Adversarial Network, Sea Surface Temperature, Super Resolution

背景と目的

近年、豪雨・洪水や渇水が世界各地で甚大化・頻発化しており水文学の重要性が増している。

しかしながら、既存の手法では現在様々な要因により解析精度・予測精度に限界がある。

そこで、近年発展が著しい深層学習に注目し、本課題においては、まず豪雨に多大な影響をもたらす海水面温度を対象に、深層学習の従来及び新しい高精度化手法を試し比較を行った。その結果、新しい CNN 系の手法である RRDBNet が最も良い結果を示した。

重要性が増している。しかしながら、既存の手法では現在様々な要因により解析精度・予測精度に限界がある。一方、近年深層学習は様々な分野において注目を集めている。特に情報学習分野における深層学習手法の発展は著しく、新しい手法が目まぐるしい勢いで開発されている。しかしながら、水文学の分野においても深層学習による高解像度化などが徐々に取り入れ始められているが、最新の手法はほとんど取り入れられていない。そこで、本課題において、まず豪雨に多大な影響をもたらす海水面温度に注目し、深層学習の従来及び新しい高精度化手法を試し比較を行った。

概要

近年、豪雨・洪水や渇水が世界各地で甚大化・頻発化している。また、地球温暖化の影響により今後甚大化・頻発化がさらに進む可能性が様々な研究により示唆されている。そのため、それらの事象を扱う水文学の

結果および考察

高解像度化の入力に用いる低解像度の海水面温度データは再解析データである ERA20C から得た。また、学習に用いる高解像度の海水面温度データとして Optimum Interpolation Sea Surface Temperature

(OISST)を選択した。本研究課題においては北米周辺の海域を選択した。また、深層学習手法として、従来の高解像度化手法である SRCNN (Super-Resolution Convolutional Neural Network)、新しい CNN (Convolutional Neural Network) 系及び GAN (Generative Adversarial Network) 系の手法である RRDBNet 及び ESRGAN (Enhanced Super-Resolution Generative Adversarial Network)を用いた。精度の評価手法として、RMSE 及び誤差率を用いた。

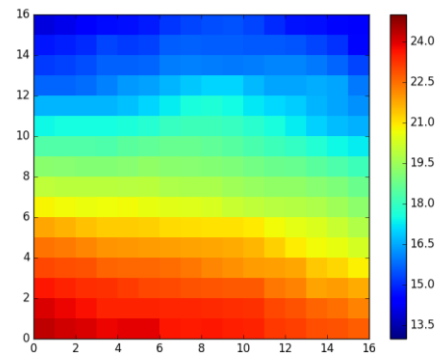
北米大西洋側沿岸において SRCNN, RRDBNet, 及び ESRGAN を用いて高解像度化を行なったとき、RMSE はそれぞれ 1.566 °C, 0.907 °C, 1.096 °C であった。また、誤差率は 8.191 %, 5.794 %, 6.929 % であった。どちらの評価手法においても RRDBNet が最も良い値を示した。また、高解像度化した海水面温度の結果の一例としてカリフォルニア沿岸における 1984 年 1 月 1 日の海水面温度を図1に示す。RRDBNet は高解像度にはなっているが、高解像度データ(OISST)と比較すると改善の必要性があると考えられる。

まとめ、今後の課題

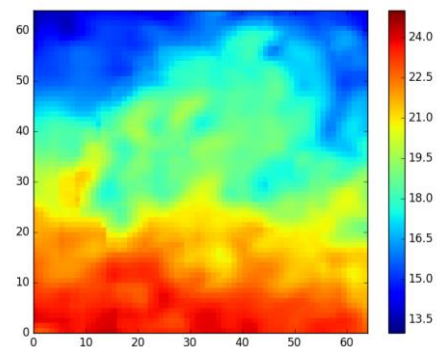
本課題においては海水面温度を対象とし、深層学習の従来及び新しい高精度化手法を試し比較を行った。従来の高解像度化手法である SRCNN, 新しい CNN 系及び GAN 系の手法である RRDBNet 及び ESRGAN を用いた。RMSE と誤差率で比較したところ、RRDBNet が最も良い結果を示した。しかしながら、高解像度データの比較ではまだ改善の必要性があると考えられる。

現在のところ、解析は一部でしか行なっておらず、高解像度化も一部の地域のみである。今後はさらなる解析を進め、また他の地域にも適用することにより、本手法の精度や適用性を検証していく予定である。

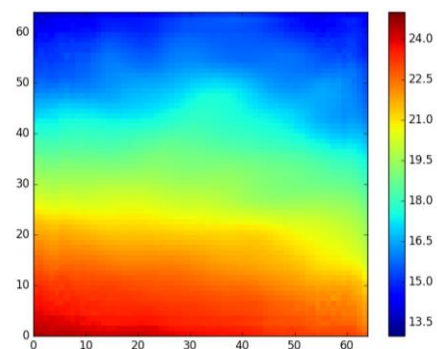
a) 低解像度データ(ERA20C)



b) 高解像度データ (OISST)



c) RRDBNet



d) ESRGAN

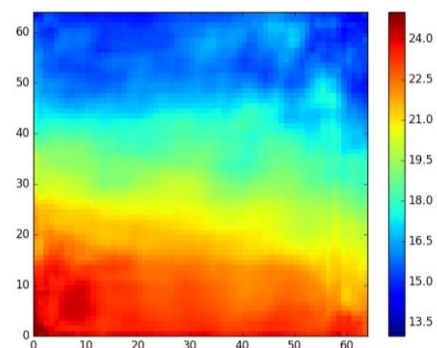


図1 1984/01/01 の海水面温度

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 損傷船体の縦曲げ最終強度と波浪中安全性評価に関する研究

英文: Study on the ultimate bending strength and the safety assessment of damaged ships

橋本 博公

Hirotada Hashimoto

神戸大学 海洋底探査センター

Kobe University

<http://www.k-obec.kobe-u.ac.jp/>

船舶が衝突・座礁した場合、折損・沈没などの大規模損傷に至るかどうかが、至る場合はどの程度の時間が残されているかを、海象条件を考慮して評価する必要がある。本研究は、漂流あるいは被曳航状態の損傷船体について、浸水流動が全船応答および縦曲げ強度に及ぼす影響を調査することを目的とする。TSUBAME3.0 上でのハイブリッド並列計算により、2 億超の粒子を用いた大規模な陽的 MPS 法を用いて、内部浸水を含む非線形流体挙動をシミュレートし、全船応答の付加質量力、復原力等の巨視的流体力に及ぼす影響を明確化することを試みた。外乱条件としては、縦曲げ強度が最も求められる規則向波とし、瘦型船を対象に損傷破孔の有無が波強制力に及ぼす影響について調査を行った。

In this study, the influence of floodwater on a ship response and longitudinal bending strength of a damaged hull was investigated. By using hybrid parallel computation on TSUBAME3.0, nonlinear fluid behavior including internal floodwater was simulated by a large-scale explicit MPS method with 200 million particles. Then the influence on an added mass force and a restoring force of a whole ship was presented. The external disturbance was set as regular head waves in which the longitudinal bending strength is most required, and effects of presence or absence of damage opening on a wave-exciting force was investigated for a slender ship.

Keywords: Damaged ship, Floodwater, Longitudinal bending strength, Explicit MPS, GPGPU

背景と目的

船舶が衝突・座礁した場合、人命・財産・環境への影響を最小化するためには、漂流あるいは被曳航状態の損傷船が、折損や沈没に至るかどうかが、あるいは至る場合にはそれまでの時間を、海象条件を考慮して、即時に精度良く判断する必要がある。そのため、まず上述の状態にある損傷船体を対象として、詳細な数値解析(波浪中弾性応答、縦曲げ崩壊挙動、浸水部流体挙動)、並びに水槽試験を行い、波浪中における損傷船体の縦曲げ崩壊挙動を浸水・傾斜影響を含めて明らかにすることを目的とする。

本研究では、損傷船体の縦曲げ最終強度評価法を確立するにあたり、損傷破孔を介して出入りする浸水滞留水が波浪中の船体応答及び縦曲げ荷重に及ぼす影響を明らかにすべく、スーパーコンピュータ TSUBAME3.0 を使用し、十分な大きさの数値水槽を用いた大規模粒子法によって、損傷破孔の有無が船体に作用する波浪強制力に及ぼす影響について明ら

かにした。

概要

浸水流動が全船応答に及ぼす影響を縦強度の観点から調べた研究はこれまでなく、その第一歩となる検討を行う。数値計算には、研究代表者らが開発してきた複数 GPU を用いての並列計算が可能な陽的 MPS コードを用いる。粒子法の計算コストは、GPU 単体でも大幅に低減させることが可能だが、MPI 通信を用いた複数デバイスによるマルチ GPU 並列プロセス計算を行うことで、2 億超の粒子数を用いた計算を可能としている。MPI によるプロセス並列は、複数の CPU プロセッサで複数のタスクを同時に実行することを可能としており、OpenMPI を並列処理に使用することで計算領域を分割し、それぞれの分割領域における演算を CPU プロセッサに割り当てて処理することで、計算負荷の分散を図っている。また、プログラムにおけるメインストリーム、領域分割の各種処理、分割領域同士の

情報通信(コミュニケーター)はホスト側で逐次処理を行い、分割領域内での粒子法による流体解析では、最も計算コストの高い時間発展部分の全てをデバイス側の演算によって処理する。

数値水槽は Fig.1 に示す通りであり、水槽壁面からの反射波の影響を避けるため、船体に対して十分に大きな水槽サイズを確保し、倒立振り型の造波機の反対側には傾斜角 15 度の消波ビーチを設けた。初期粒子配置に対して、2 次元スライスグリッドによる領域分割の様子を Fig.2 に示す。本研究では、8 ノード、32 GPU を用いて 40 秒間のシミュレーションを行った。なお、計算領域分割数は 64 である。

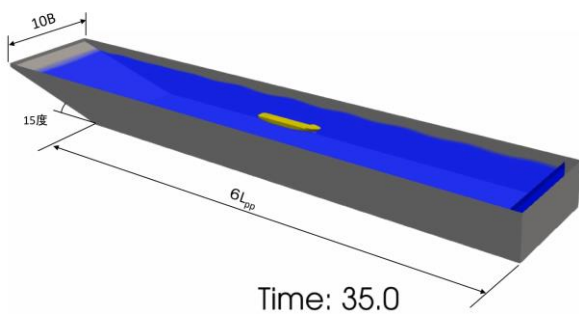


Fig.1 Numerical wave tank

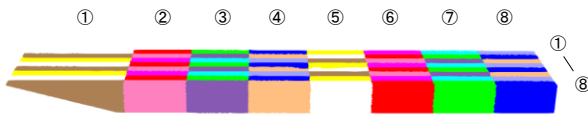


Fig.2 Initial decomposition of the domain

数値水槽に設置した船体は運動を完全固定とし、いわゆるディフракション問題を解くことで、波浪強制力を求めた。非損傷状態と右舷中央部に損傷破孔と損傷区画を有する損傷状態の 2 状態について、シミュレーションを実施した。波条件は、縦曲げモーメントが厳しくなる波長船長比 1.0、波高波長比 0.027 の規則波を真正面から受ける向波と設定した。

結果および考察

非損傷状態におけるシミュレーションのスナップショットを Fig.3 に、ある瞬間における船体表面の圧力分布を損傷/非損傷の両状態について Fig.4 に示す。Fig.4 より、損傷状態では船体中央部の区画に浸水が生じ、内部の滞留水に作用する重力が評価されていることが分かる。



Fig.3 A snapshot of simulation



Fig.4 An example of pressure contour on the hull surface

Fig.5 は、非損傷状態における、上下、左右方向の波強制力、ならびに船体重心周りの横揺れモーメントのシミュレーション結果である。Fig.6 は同じ波条件における損傷状態の結果である。滞留水の影響により、船体に作用する上下方向の浮力が低下していること、幅方向に対する非対称浸水を原因とする非定常横揺れモーメントが生じており、計算結果の妥当性が見て取れる。また、損傷区画の存在により、左右方向にも定常な流体力が生じていることは興味深い。

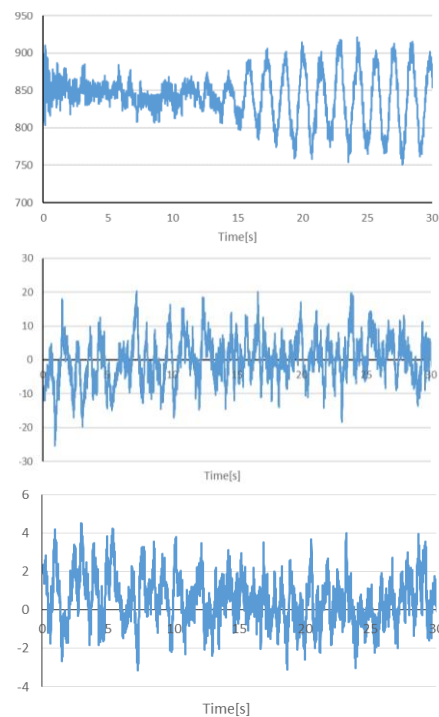


Fig.5 Wave-exciting forces and moment on the intact hull

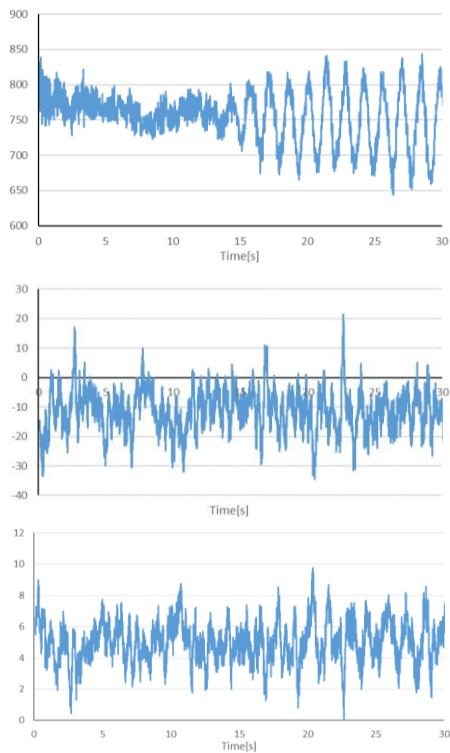


Fig.6 Wave-exciting forces and moment on the damaged hull

まとめ、今後の課題

TSUBAME3.0を用いた大規模な陽的MPSにより、瘦型船の非損傷/損傷状態での波浪強制力の推定を試みた。本計算における空間解像度は、喫水に対して約10個、全幅に対して約30個程度の粒子に留まっているが、波強制力については、損傷区画の存在の影響を定性的に推定できているようである。今後は水槽実験を実施し、その結果と比較することで、縦曲げ強度の解析に求められる精度が得られているかを検証する必要がある。

TSUBAME 共同利用 令和元年度 学術利用 成果報告書

利用課題名 抗体の安定性解析のための分子動力学シミュレーション
 英文: Molecular dynamics simulation for stability analysis of antibodies

利用課題責任者 石川俊平
 Shumpei Ishikawa

所属 東京大学大学院医学系研究科
 Graduate School of Medicine, The University of Tokyo
<https://plaza.umin.ac.jp/prm/>

邦文抄録(300 字程度) リンパ球の抗体(免疫グロブリン)は多様な抗原に対応できるよう、同一個体内の各クローンで異なる DNA 配列を持ち多様な構造のものが作られる。抗体の構造多様性の程度を把握するために、立体構造をシミュレーションによって推定することは重要なアプローチのひとつであるが、多くの抗体の配列から構造を推定するには比較的長時間でのシミュレーションが重要であり、そのために多くの計算リソースが必要である。本課題では抗原受容体の結晶構造情報をもとに多くの種類の抗体の構造を分子動力学シミュレーションによって解析し、抗体の構造多様性に関する結果を得た。

英文抄録(100 words 程度) Antibodies produced by each lymphocyte clone within the same individual have different DNA sequences and protein structures, which enables recognizing a variety of antigens. Although simulation-based inference of 3D structures is one of important approaches to understand the structural diversity of antibodies, relatively long simulation is essential for that purpose, which requires high computational resources to infer structures of many antibodies. We analyzed various kinds of antibody structures by MD simulation based on their crystal structure information and obtained the results as to the structural diversity of antibodies.

Keywords: 抗体、立体構造、分子動力学シミュレーション、GROMACS

背景と目的

本課題では抗原受容体の結晶構造情報をもとに多くの種類の抗体の構造を分子動力学シミュレーションによって解析し、抗体の構造多様性を把握することが目的である。シミュレーションの結果を用いることで、抗体の構造だけでなく、各残基間の相互作用を計算することも可能である。これらは各抗体を特徴付ける特徴量であり、将来的に抗体生成モデルの構築に必要な抗体探索空間の次元削減を行うための基盤データとなる。

概要

リンパ球の抗体(免疫グロブリン)は多様な抗原に対応できるよう、同一個体内の各クローンで異なる DNA 配列を持ち多様な構造のものが作られる。抗体の構造多様性の程度を把握するために、立体構造をシミュレーションによって推定することは重要なアプローチのひとつであるが、構造を推定するには比較的長時間でのシミュレーションが重要であり、そのために多くの抗体

の配列の計算には多くのリソースが必要である。本課題では抗原受容体の結晶構造情報をもとに多くの種類の抗体の構造を分子動力学シミュレーションによって解析し、抗体の構造多様性を把握することが目的である。

結果および考察

本課題では実験的に得られた 336 種類の抗体に対して、水中および血中の 2 つの条件下における分子動力学シミュレーションを行った。その後、トラジェクトリファイルから各残基間のクーロン力およびファンデルワールス力の継時的変化を計算した。

ソフトウェアとしてはTSUBAME3.0にデフォルトでインストールされているGROMACS 2019を使用した。1つの抗体をシミュレーションするのに 1/4 ノード(7CPU、1GPU)を用い、1 ステップあたり 2 フェムト秒、計 50,000,000 ステップ 100 ナノ秒のシミュレーションを行った。シミュレーションを行った原子の数は水分子を含

め 50,000~60,000 であり、1 つの抗体を 100 ナノ秒シミュレーションするのに約 14 時間かかった。

まとめ、今後の課題

本課題では種々の抗体に対して分子動力学シミュレーションを行い、さらに各残基間の相互作用を計算した。今後は各抗体における相互作用の共通点と相違点について検討を行う。また、これらのデータを用いて抗体探索空間の次元削減、および抗体生成モデルの構築を行う。

東京工業大学 TSUBAME 共同利用 令和元年度利用終了課題 利用成果報告書集

発行 : 令和3年 5月

国立大学法人 東京工業大学 学術国際情報センター 共同利用推進室

住所 : 〒152-8550 東京都目黒区大岡山 2-12-1 E2-6

E-mail : kyoyo@gsic.titech.ac.jp

URL : <https://www.gsic.titech.ac.jp/tsubame>

本書に記載の記事・写真等の二次利用を禁じます。これらの情報は著作権法上認められた「私的利用」または「引用」の条件をみたした場合を除いて、著作権者に無断で転載、複製、放送、公衆送信、翻訳、販売、貸与等の利用を禁じます。