

## 日本医療研究開発機構 創薬等ライフサイエンス研究支援基盤事業 事後評価報告書



### I 基本情報

補助事業課題名：（日本語）創薬等ライフサイエンス研究支援基盤事業  
（プログラム名）（英語）Platform Project for Supporting Drug Discovery and Life Science Research

実施期間：平成29年4月1日～令和4年3月31日

補助事業担当者 氏名：（日本語）寺田 透  
（英語）Tohru Terada

補助事業担当者 所属機関・部署・役職：  
（日本語）国立大学法人東京大学・大学院農学生命科学研究科・准教授  
（英語）The University of Tokyo, Graduate School of Agricultural and Life Science, Associate Professor

### II 補助事業の概要

本事業では、比較モデリング法を用いたタンパク質の立体構造予測、ドッキング法を用いた複合体立体構造予測、分子動力学（MD）法を用いた予測構造の精密化、安定性の評価、分子間相互作用解析、結合自由エネルギー計算、ダイナミクス解析、量子化学計算を用いた酵素の反応触媒機構の解析など分子シミュレーション法を用いて、外部研究者のライフサイエンス研究を支援した。23件の支援を行い、すべて完了した。うち、16件で成果を論文にまとめ公表した。

支援課題「X線結晶構造とMDシミュレーションを用いた、薬剤排出膜トランスポーターの多剤認識、輸送メカニズムの解明」では、多剤排出トランスポーターMdfAについて、H<sup>+</sup>と薬剤の対向輸送に関わるタンパク質の立体構造変化のメカニズムを解析した。MdfAは、細胞外側が開いたoutward open構造と、細胞内側が開いたinward open構造をとる。本課題では、薬剤排出に重要な2つの酸性残基、E26、D34がとりうる4種類のプロトン化状態のそれぞれについて、outward open構造、inward open構造のそれぞれから出発したMDシミュレーションを実施した。その結果、D34がプロトン化すると、outward open構造、inward open構造のいずれからでも、細胞外側、細胞内側の両方が閉じたoccluded構造に移行することが示され、この残基のH<sup>+</sup>輸送と共役したタンパク質構造変化における役割を明らかにした（Nagarathinam *et al.*, *Nat. Commun.* **9**, 4005 (2018)）。

支援課題「催涙因子合成酵素LFSの反応機構解析」では、タマネギに存在し、前駆体(E)-1-propensulfenic acid

(1-PSA) を催涙因子 (Z)-propanthial S-oxide (PTSO) に変換する酵素 lachrymatory factor synthase (LFS) の反応触媒機構を解析した。ここでは、LFS の結晶構造を用いて、酵素・基質複合体のモデリングを行い、MD シミュレーションを用いて安定な複合体構造を探索した。さらに、quantum mechanics/molecular mechanics (QM/MM) 法を用いて酵素の反応触媒機構を決定した (Arakawa *et al.*, *ACS Catal.* **10**, 9–19 (2020))。

支援課題「アブラナ科植物の自家不和合性に関する構造機能解析」では、アブラナ科植物における自家不和合性反応のメカニズムを解析した。ここでは、雌ずい側因子の受容体キナーゼ SRK と花粉側因子の SP11 が、同じハプロタイプ由来のとき、特異的に結合すると考えられている。本研究ではまず、ハプロタイプ 8 の SRK/SP11 複合体の結晶構造と、ハプロタイプ 9 の SRK/SP11 複合体の結晶構造をもとに、立体構造未知の 22 種類の SRK、9 種類の SP11 の立体構造予測を行った。合計 24 種類の SRK、11 種類の SP11 を、配列類似性に基づいて 2 つのクラスに分け、それぞれのクラスについて、総当たりで複合体モデルを作成した。続いて、それぞれの複合体モデルについて、molecular mechanics-generalized Born surface area (MM-GBSA) 法を用いて結合自由エネルギーを計算した。この結果、同じハプロタイプが複合体を形成しているとき、結合自由エネルギーは特に大きな負の値をとることが示され、提唱されている自家不和合性反応のメカニズムを支持する結果が得られた (Murase *et al.*, *Nat. Commun.* **11**, 4916 (2020))。

支援課題「グラム陰性細菌の鉄結合タンパク質の鉄配位様式と結合強度の分子動力学による解析」では、グラム陰性細菌 *Thermus thermophilus* の鉄取込みに関わる鉄結合タンパク質 TtFbpA がとる、pH によって異なる 2 つの鉄配位様式それぞれについて量子化学計算を行い、リガンドの分子種とプロトン化状態を決定した。また、鉄結合部位周辺のタンパク質構造の違いを見出し、配位形式の違いに加えてタンパク質構造の違いが鉄結合強度の違いに寄与している可能性があることを明らかにした (Lu *et al.*, *Metallomics*, **11**, 2078–2088 (2019))。

支援課題「ブラシノステロイド情報伝達に特異的な DNA 認識機構の解析」では、植物ホルモン・ブラシノステロイドの情報伝達系で機能する転写因子 BIL1 と、同じ DNA 配列を認識する別の転写因子 MYC2 の間で、DNA 配列の認識機構の違いを解析した。BIL1、MYC2 と DNA との複合体の結晶構造に対して MD シミュレーションを実行し、分子間相互作用の違いを明らかにした。さらに、DNA の塩基置換に伴う結合自由エネルギーの変化を計算し、実験と一致する結果を得た (Nosaki *et al.*, *Sci. Rep.* **11**, 3879 (2021))。

支援課題「アルツハイマー病、家族性認知症及び 2 型糖尿病に対する新規予防・治療薬の創出」では、タンパク質分解に関わる Cullin-RING ligases (CRLs) における分子認識メカニズムを解析した。CRLs では、Cul2/ElonginBC または Cul4A/DDB1 が分解のターゲットである基質認識タンパク質と結合するが、ElonginBC に結合するタンパク質には BC-box、DDB1 に結合するタンパク質には H-box と呼ばれるコンセンサス配列が存在することが知られている。依頼者らは、この認識は排他的ではなく、BC-box、H-box の両方の性質を併せ持つ配列が存在することを見出した。本課題では、28 の配列について ElonginBC および DDB1 との複合体のモデリングを行い、いずれも無理なく ElonginBC と DDB1 の両方に結合できることを示した (Yasukawa *et al.*, *Cell Rep.* **30**, 3478–3491 (2020))。

支援課題「LRRTM2 変異体の分子動力学シミュレーション」では、シナプス後終末のオーガナイザー (分化を誘導する細胞接着因子) である LRRTM2 と、シナプス前終末のオーガナイザーである Neurexin 1 $\beta$  (Nrxn1 $\beta$ ) の細胞外領域の複合体の結晶構造に対して MD シミュレーションを実行し、分子間相互作用を解析した。特に、分子界面から離れた変異により活性が低下する他グループの実験結果について検証を行った。この変異は、分子間相互作用には影響を与えないことが示され、変異により発現量が低下したためであることが確認された (Yamagata *et al.*, *Nat. Commun.* **9**, 3964 (2018))。

支援課題「innexin-6 のヘミチャネルの分子動力学シミュレーション」では、隣接する細胞の細胞質同士を直接つなぐギャップ結合チャネルタンパク質 innexin-6 (INX-6) の電子顕微鏡構造に対して、MD シミュレーションを用いてチャネル内外の脂質二重層のモデリングを行った。この結果、チャネル内部では脂質二重層が外部の脂質二重層より細胞内側にずれていること、一部の脂質分子はサブユニット界面に侵入していることが明らかとなった (Burendei *et al.*, *Sci. Adv.* **6**, eaax3157 (2020))。

支援課題「tRNA を標的とするリボヌクレアーゼの基質認識機構解明に向けた構造モデルの構築」では、抗菌活性を持つバクテリオシンの 1 つ Colicin D と、その基質である tRNA の複合体のモデリングを行った。ここでは、構築したモデルのそれぞれについて MD シミュレーションを行い、複合体が安定に保持され、反応に必要な原子配置も維持されるモデルを得た (Ogawa *et al.*, *RNA Biol.* **18**, 1193–1205 (2021))。

支援課題「新規配列を持つ糖加水分解酵素分子の揺らぎ」では、真菌由来の  $\beta$ -1,2-グルカンを加水分解する酵素 TfSGL と基質との複合体の結晶構造に対して MD シミュレーションを実行し、活性中心を構成する残基のダイナミクスと、分子間相互作用を解析した。この結果、結晶構造はパッキングの影響を受けておらず、水溶液中でも基質との相互作用が維持されることが示された (Tanaka *et al.*, *J. Biol. Chem.* **294**, 7942–7965 (2019))。

支援課題「リプログラミング因子変異体の DNA 相互作用における分子動力学シミュレーション」では、リプログラミング効率を高める KLF4 の変異体について、この変異が KLF4 と DNA の間の親和性を高めるメカニズムを、MD シミュレーションにより解析した。この結果、変異により、DNA に対する KLF4 の結合様式がわずかに変化し、両者の間の相互作用が増加することで、親和性が高くなっていることが示唆された (Borisova *et al.*, *iScience*, **25**, 103525 (2022))。

支援課題「SIRT1 とリガンドの分子間相互作用の分子動力学シミュレーション」では、加齢による病気 (2 型糖尿病、神経変性、ガン、アルツハイマー病など) を防ぐ働きがあることが知られている Sirtuin 1 (SIRT1) の N 端ドメイン (NTD) と、これを活性化する resveratrol および KPMF-8 との相互作用を、MD シミュレーションを用いて解析した。この結果、いずれの化合物も NTD への結合構造には複数の安定構造があり、その間を速いタイムスケール (< 1  $\mu$ s) で交換していることが明らかとなった (Zhang *et al.*, *Commun. Biol.* **4**, 209 (2021))。

支援課題「放線菌由来メチル基転移酵素の機能解析」では、放線菌の benzastatin の生合成にかかわる酵素 BezA の結晶構造をもとに、BezA とメチル基供与体 S-adenosylmethionine (SAM)、基質 geranyl diphosphate (GPP) の 3 者複合体のモデルを構築した。続いて、MD シミュレーションを用いてモデルを精密化し、QM/MM 計算により触媒反応機構を明らかにした (Tsutsumi *et al.*, *Angew. Chem. Int.* **61**, e202111217 (2022))。

支援課題「光合成細菌の RC-LH1 の MD シミュレーション」では、光合成細菌の光合成反応中心 (RC) とアンテナ色素タンパク質複合体 (LH1) が結合した超分子複合体 RC-LH1 の電子顕微鏡構造について、MD シミュレーションを行った。この結果、複合体の外部と ubiquinone-10 (U10) の結合部位を隔てる 2 本の膜貫通ヘリックスの間が広がり、そこを通過して U10 が出入りしていることが示唆された (Bracun *et al.*, *Sci. Adv.* **7**, eabf8864 (2021))。

支援課題「Large pore channel の分子動力学シミュレーション」では、細胞膜を介した物質の輸送に関わるヒトの pannexin-1 (PANX1) の電子顕微鏡構造に対して MD シミュレーションを実行し、サブユニット界面に存在する脂質分子の運動を解析した (Kuzuya *et al.*, *Sci. Signal.* **15**, eabg6941 (2022))。

支援課題「新規フルクトース含有糖合成酵素の分子動力学シミュレーション」では、口腔内のビフィズス菌から発見された新規酵素 difructose dianhydride I synthase/hydrolase の結晶構造に対して MD シミュレーションを実行し、酵素と、基質・生成物との相互作用、および基質ポケット周辺の水和構造を解析した。また、酵素内部の空洞をリガンド分子が移動する様子を明らかにした (Kashima *et al.*, *J. Biol. Chem.* **297**, 101324 (2021))。

また、以下の 6 件の高度化課題に取り組んだ。

高度化課題「リガンド結合シミュレーション解析法の高度化」では、粗視化モデルを用いたリガンド結合シミュレーションのトラジェクトリから、マルコフ状態モデルを構築し、リガンドの長時間のふるまいを明らかにする方法を開発した。また、粗視化モデルのポテンシャルエネルギー関数に、リガンド非結合状態と結合状態の 2 つの参照構造をエネルギー極小状態とする double-basin ポテンシャルを組み合わせることで、リガンド結合に伴うタンパク質の立体構造変化を再現する手法を開発した。さらにパラメータを調整することで、conformational selection と induced fit の 2 つの経路を再現できることを示した (Negami *et al.*, *Chem. Phys. Lett.* **742**, 137144 (2020))。

高度化課題「分子シミュレーション法の高度化」では、天然変性領域を含むタンパク質について、SAXS データを再現する構造アンサンブルを構築した。ここでは、粗視化 MD シミュレーションを用いて生成した構造アン

サンプルに、最大エントロピー法に基づく加重法を適用した。支援課題「アブラナ科植物の自家不和合性に関する構造機能解析」では、SP11の立体構造予測に、ソフトウェア Rosetta を用いた *ab initio* 立体構造予測と accelerated MD による安定構造探索を適用した。支援課題「innexin-6 のヘミチャネルの分子動力学シミュレーション」では、巨大な環状複合体の内部（チャネル）に脂質二重層を構築する手法を開発した。

高度化課題「結合リガンド・リガンド結合部位予測法の高度化」では、タンパク質と様々なリガンドの組み合わせについて、結合リガンドとリガンド結合部位を予測する手法を開発した。金属イオンの結合部位予測では、アミノ酸配列から、Ca<sup>2+</sup>、Co<sup>2+</sup>、Cu<sup>2+</sup>、Cu<sup>+</sup>、Fe<sup>3+</sup>、Fe<sup>2+</sup>、Hg<sup>2+</sup>、Mg<sup>2+</sup>、Mn<sup>2+</sup>、Ni<sup>2+</sup>、Zn<sup>2+</sup>の結合部位を機械学習と配列類似性に基づいて予測する方法を開発した。特に、Zn<sup>2+</sup>、Mn<sup>2+</sup>、Fe<sup>3+</sup>については、先行研究に比べて高い accuracy が得られた (Tian *et al.*, *Adv. Bioinform. Chem.* **1**, 025–036 (2019))。

高度化課題「機械学習を用いたクライオ電子顕微鏡の撮影効率化」は、産学連携高度化研究課題として実施した。クライオ電子顕微鏡の試料は、多数の穴が規則的に並んだカーボン被膜で覆われた金属製のグリッドに、タンパク質溶液を塗布し、急速に冷却することで作成される。タンパク質の立体構造を高い分解能で決定するためには、タンパク質が適切な厚さの氷の膜に包埋されていることが重要である。本課題では、撮影効率の向上を目指して、機械学習を用いて適切な厚さの氷が張られた膜穴を判別する手法を開発した。膜穴の検出に YOLOv3、膜穴の判別に Xception を用いて判別器を構築し、膜穴の低倍率画像を学習することで、93%以上の accuracy で判別できることを示した (Yokoyama *et al.*, *Biophys. Rev.* **12**, 349–354 (2020))。また、電子顕微鏡制御ソフトウェア SerialEM と連携し、膜穴を検出・選別する PyHoleFinder を作成し公開した (<https://github.com/tterada-utokyo/PyHoleFinder>)。

高度化課題「抗原・抗体ドッキングモデル安定性評価・精密化法の開発」と高度化課題「タンパク質と低分子化合物の複合体結晶構造に特有な相互作用を学習した AI の開発」は「AI を活用したインシリコ創薬支援」の一環として、2020年10月より取り組んだ。前者では、低解像度の抗原・抗体複合体モデルに基づいて、分子界面付近のループ構造を変化させた多数のモデルを生成し、MD シミュレーションを実行して安定なモデルを選択する方法を開発した。後者では、分子内・分子間相互作用を模したネットワーク構造を持つグラフニューラルネットワークを構築し、タンパク質と低分子化合物の相互作用と結合自由エネルギー予測する AI を開発した。これをベンチマークセットの CASF-2016 に適用したところ、予測値と実験値の相関係数は 0.751 となり、ドッキングシミュレーションで良く用いられる GoldScore (0.416) や AutoDock Vina (0.604) を上回った。また、同様にグラフニューラルネットワークを用いて、結晶構造と結晶構造に近い (RMSD が 2~4 Å) ドッキング構造を判別する AI を開発した。

We provided scientific support for 26 life-science research projects conducted by outside researchers, using comparative modeling, docking simulation, molecular dynamics (MD) simulation, and quantum chemical calculation. We completed all the projects and published peer-reviewed papers in journals in 16 projects. We studied a multidrug transporter termed MdfA using MD simulations. We identified an acidic residue that plays an important role in the structural change coupled with the H<sup>+</sup> transport [Nagarathinam *et al.*, *Nat. Commun.* **9**, 4005 (2018)]. A structural model of the complex between lachrymatory factor synthase (LFS) and its substrate was constructed by using docking simulation and MD simulation. The catalytic mechanism was determined by quantum chemical calculation [Arakawa *et al.*, *ACS Catal.* **10**, 9–19 (2020)]. We studied a pair of proteins termed SRK and SP11 involved in *Brassica* self-incompatibility. The amino-acid sequences of these proteins are different between different haplotypes. It is hypothesized that only the proteins from the same haplotype can form a stable complex. Structural models of SRK and SP11 from various haplotypes were constructed and binding free energies between the SRK and the SP11 models were calculated. We found that the pairs from the same haplotype showed the lowest free energies in accordance with the hypothesis [Murase *et al.*, *Nat. Commun.* **11**, 4916 (2020)]. We studied the iron coordination structure of a protein termed TtFbpA involved in iron uptake in *Thermus thermophilus*. It adopts two different

coordination structures in a pH-dependent manner. We determined the types of the ligand molecules coordinating to the iron and their protonation states based on quantum chemical calculation [Lu *et al.*, *Metallomics*, **11**, 2078–2088 (2019)]. The difference in the DNA recognition mechanism was investigated between two transcription factors termed BIL1 and MYC2, which recognize the same DNA sequence. MD simulations were performed for the two proteins in complex with the same DNAs and binding free-energy changes upon replacement of the base pairs in the recognition sequence were calculated [Nosaki *et al.*, *Sci. Rep.* **11**, 3879 (2021)]. The molecular recognition mechanism was investigated for Cullin-RING ligases (CRLs). In CRLs, a pair of Cul2/ElonginBC or Clu4A/DDB1 recognizes substrate binding proteins. The substrate binding proteins that bind to ElonginBC have the BC-box sequence and those that bind to DDB1 have the H-box sequence. The collaborators found sequences that have characteristics of both the BC-box and H-box. We constructed structural models complexed with ElonginBC or DDB1 for each of 28 candidate sequences. We found that all the sequences reasonably formed complexes both with ElonginBC and DDB1 [Yasukawa *et al.*, *Cell Rep.* **30**, 3478–3491 (2020)]. The intermolecular interactions between a postsynaptic organizer LRRTM2 and a presynaptic organizer Nrnx1 $\beta$  were investigated by using MD simulation [Yamagata *et al.*, *Nat. Commun.* **9**, 3964 (2018)]. Structural models of the hemichannel of a gap-junction channel innexin-6 (INX-6) embedded in a lipid bilayer were constructed [Burendei *et al.*, *Sci. Adv.* **6**, eaax3157 (2020)]. A structural model of the complex between a bacteriocin, Colicin D and its substrate tRNA was constructed by using MD simulation [Ogawa *et al.*, *RNA Biol.* **18**, 1193–1205 (2021)]. The dynamics and the interactions with the substrate were investigated for an enzyme termed TfSGL by using MD simulation [Tanaka *et al.*, *J. Biol. Chem.* **294**, 7942–7965 (2019)]. We investigated the mechanism of how a mutation of a reprogramming factor, KLF4 increases the affinity to DNA using MD simulation [Borisova *et al.*, *iScience*, **25**, 103525 (2022)]. The ligand-binding and -unbinding dynamics were investigated for a system of a protein termed SIRT1, which is involved in prevention of age-related diseases, and the compounds that activate SIRT1 by using MD simulation [Zhang *et al.*, *Commun. Biol.* **4**, 209 (2021)]. A structural model of the ternary complex of a methyltransferase, BezA with *S*-adenosylmethionine and the substrate was constructed. The catalytic mechanism was determined by quantum chemical calculation [Tsumumi *et al.*, *Angew. Chem. Int.* **61**, e202111217 (2022)]. The dynamics of a redox mediator, ubiquinone-10 in the reaction center (RC)-light-harvesting complex 1 (LH1) supercomplex of a photosynthetic bacterium was investigated by using MD simulation [Bracun *et al.*, *Sci. Adv.* **7**, eabf8864 (2021)]. The dynamics of the lipid molecules between the subunits of a large-pore membrane channel, pannexin-1 was investigated by using MD simulation [Kuzuya *et al.*, *Sci. Signal.* **15**, eabg6941 (2022)]. The intermolecular interactions and the hydration structure in the catalytic pocket were investigated for the complex between an enzyme, difructose dianhydride I synthase/hydrolase and its substrates [Kashima *et al.*, *J. Biol. Chem.* **297**, 101324 (2021)].

In addition, we conducted the following projects to sophisticate the methodology. We developed a method to construct a Markov state model describing the ligand-binding dynamics from trajectories of coarse-grained MD simulations. A method to reproduce the conformation change upon ligand binding in a coarse-grained MD simulation was also developed [Negami *et al.*, *Chem. Phys. Lett.* **742**, 137144 (2020)]. Coarse-grained MD simulation was applied to construct a structural ensemble that reproduce the SAXS profile of a protein that contains intrinsically disordered regions. A Bayesian maximum entropy method was used to determine the weight of each structure in the ensemble. In the structure prediction of SP11, an *ab initio* structure prediction software, Rosetta was used to generate the initial models and the accelerated MD method was used to refine them. A method to generate a structural model of a lipid bilayer within the INX-6 channel using MD simulation was developed. We developed a method to predict metal-binding sites of a protein from its amino-acid sequence based on machine learning and sequence similarity [Tian *et al.*, *Adv. Bioinform. Chem.* **1**, 025–036 (2019)]. We also developed a deep learning-based method to identify “good” regions of a cryo-electron microscopy grid [Yokoyama *et al.*, *Biophys. Rev.* **12**, 349–354 (2020)]. Moreover, we developed a program termed PyHoleFinder that finds holes in holey carbon films on a cryo-electron microscope grid and can work with SerialEM. This program is available at <https://github.com/tterada->

utokyo/PyHoleFinder. To facilitate AI-driven drug development, we developed a method to refine structural models of antibody-antigen complex. We also developed graph neural network-based methods to predict the binding free energy of a small molecule to a protein and to select a docking structure close to the crystal structure ( $\text{RMSD} < 2 \text{ \AA}$ ) from those with RMSDs between  $2 \text{ \AA}$  and  $4 \text{ \AA}$ .