

日本医療研究開発機構
次世代治療・診断実現のための創薬基盤技術開発事業
事後報告書

I 基本情報

研究開発課題名：（日本語）天然化合物及びITを活用した革新的医薬品創出技術
研究開発項目①「ITを活用した革新的医薬品創出基盤技術開発」
（英語）Development of core technologies for innovative drug development based upon natural products and IT
Theme 1: Development of core technologies for innovative drug development based upon IT

研究開発実施期間：2013年8月7日～2018年3月31日

研究開発代表者 氏名：（日本語）嶋田 一夫
（英語）ICHIO SHIMADA

研究開発代表者 所属機関・部署・役職：
（日本語）次世代天然物化学技術組合 プロジェクトリーダー
／国立大学法人東京大学大学院薬学系研究科 教授
（英語）Project Leader, Technology Research Association for Next generation natural products chemistry
／Professor, Graduate School of Pharmaceutical Sciences, Faculty of Pharmaceutical Sciences, The University of Tokyo

II 研究開発の概要

研究開発課題名：天然化合物及びITを活用した革新的医薬品創出技術 研究開発項目①「ITを活用した革新的医薬品創出基盤技術開発」

革新的 in silico シミュレーション/スクリーニング技術開発として、タンパク質の動的構造変化を考慮した、高速・高精度のタンパク質/リガンド複合体、およびタンパク質/タンパク質複合体モデリング手法を新規開発し、複数の例で活性化化合物を見出すことに成功した。また、タンパク質の構造解析技術として、膜タンパク質の安定化試料作製法や GraDeR 法、クライオ電顕による単粒子解析法および高分子量タンパク質に対する高感度 NMR 測定法および NMR を用いた医薬品候補化合物の最適化法の確立などを行い、GPCR および クローディン、イネキシンなどの生物学的重要な創薬標的となるタンパク質の立体構造決定に成功した。

1. 分担研究開発課題名：革新的 in silico シミュレーション/スクリーニングソフトウェアの開発

i) 化合物リガンド・データベースの開発

スクリーニング用に入手可能な化合物データベース (DB) の新規構築を目的として、市販化合物等約 4200 万件の化合物情報を収集し、LigandBOX-DB として 720 万件を公開した (Kawabata *et al.*, *Biophysics*, 2013)。欧州 EBI のタンパク質-化合物相互作用データベースである ChEMBL のデータを基にして、ドッキングスコアを記述子とし自動的に QSAR モデルを生成する docking-score QSAR 法を開発した。本手法での誤差は、1 kcal/mol と、通常のドッキングソフトの 10-100 倍の高精度を実現した (Fukunishi *et al.*, *Mol. Informatics*, 2017; *Ibid.* 2018)。

ii) 化合物設計・合成評価用ソフトウェアの新規開発

合成容易性予測では、既存の医薬品など 250 種類以上について、合成科学者の合成難易度評価を再現する機械学習ソフトを開発した (Fukunishi *et al.* *J. Chem. Inf. Model.*, 2014; Fukunishi *et al.*, *Curr. Pharm. Des.* 2016)。分子の合成展開の候補を示すソフトウェアでは、既存の化合物を合理的にフラグメントに分解し、これらのフラグメントを、与えた母核に連結して、合成展開の可能性を提示するソフトウェアを開発・公開した。

iii) タンパク質の動的構造変化を考慮した、高速・高精度のタンパク質/リガンド複合体、およびタンパク質/タンパク質複合体モデリング手法の新規開発

通常の分子動力学 (MD) 計算よりもはるかに高い効率で構造探索を行い、より正確な構造アンサンブルを構築できる V-AUS 法 (Higo *et al.*, *J. Comput. Chem.*, 2015) および VcMD 法 (Higo *et al.*, *J. Phys. Chem.*, 2017; Hayami *et al.*, *J. Comput. Chem.*, 2018) を開発し、静電相互作用を高速に行う ZMM 法と先進的な並列化手法を組み込んだ myPresto/psygene, omegagene に搭載した。この応用によりタンパク質/リガンドおよびタンパク質/タンパク質複合体のモデリングを自由エネルギー地形に基づき行う一方、NMR、X 線結晶解析グループおよび製薬企業との協力により、自由度の高いタンパク質複合体に対する種々のモデルを作成した (Iida *et al.*, *J. Comput. Chem.*, 2016; Kasahara *et al.*, *Nucl. Acids Res.*, 2018)。

iv) 最新の GPU 及びメニーコア PC クラスタを用いたスクリーニングソフトの高速化・高精度化

MD 計算プログラム myPresto/psygene を GPU 専用ソフト (myPresto/psygene-G) に加えて、比較的小さな系に対して単一の GPU コアで高速に計算を実施する myPresto/omegagene を開発・公開した (Kasahara *et al.*, *Biophys. Physicobiol.*, 2016)。一方、タンパク質-薬物ドッキングソフト myPresto/sievgene の並列化を行い、オリジナルな myPresto/sievgene の精度を保ったまま、メニーコア PC 上で 20-70 倍の高速化を達成した。

v) 極めて高い精度の結合力を推定できる力場パラメーター計算と分子シミュレーション技術の開発

ドッキング計算でのスクリーニングによる結合力推定のため、半経験的量子化学計算ソフト MOPAC を改良して各化学結合の種類に応じて原子部分電荷を補正し、高精度の原子電荷を与える MOPAC AM1BCC 電荷を計算できるようにした。この手法は項目 i) における LigandBOX-DB でのデータ構築に用いられ、また一般にも公開した。一方、マルチカノニカル MD 計算による構造探索と SRPG 法を組み合わせ、リガンドの binding/unbinding のパスに沿った熱力学積分を行い、リガンドの結合自由エネルギーを高精度に予測する一般的な手法を開発した (Nguyen *et al.*, *J. Chem. Inf. Model.*, 2015; Bekker *et al.*, *J. Chem. Theory Comput.*, 2017)。

vi) ユーザ・インターフェースの開発

Windows、Mac、Linux など代表的な OS で動作可能なグラフィックインターフェイス (GUI) である “myPresto portal” を開発し、新 myPresto (myPresto version 5) として公開した (<https://www.mypresto5.jp/>)。myPresto の種々のソフトウェア類の入出力を整理して GUI から容易に実行できるようにし、化合物の分子編集、タンパク質系の準備、分子シミュレーション、化合物ドッキングなどがマウス操作で行える。また、自前の PC クラスタ等のハードを必要とせずクラウド上の PC クラスタを利用する実証実験も行った。GUI 製品の市販化が組合企業 2 社から始まり、関連ソフト開発やサービスも複数の企業で開始された。

vii) 新 myPresto による探索的実証研究

水溶性 PI5P4K キナーゼの阻害化合物を、200 万化合物から myPresto により選択した 1,167 種類を 200 種に絞り込

み、NMR法による活性検証を実施してIC₅₀が10 μM以下の阻害化合物を12個、1 μM以下のものを2個見出した。その複合体のX線結晶構造解析も実施した。転写制御因子mSin3の構造からmyPrestoの種々のソフトウェアとDBにより阻害化合物を選択し、NMR法により50種以上の活性化化合物を見出した。企業との共同研究により農薬開発を進めアゴニスト活性を持つものを見出した。

2. 分担研究開発課題名：核磁気共鳴法（NMR）によるタンパク質の生理的条件下における動的立体構造取得技術の開発

i) 創薬標的タンパク質の高感度NMRを可能とする試料調製法及び測定法の開発

① 昆虫細胞発現系を用いた重水素化タンパク質調製法の開発および実証研究

多くの創薬標的タンパク質の発現において必要な昆虫細胞発現系において、重水素標識によるNMRシグナルの高感度化（5倍以上）を世界で初めて達成した。開発した手法は投稿論文に発表して（Kofuku *et al.*, *Angew. Chem. Int. Ed.*, 2014）、プロトコール化した。さらに、従来用いていたメチオニンに加えてアラニン・イソロイシン・ロイシン・スレオニン・バリンをプローブとするための標識法の開発にも成功した。

上記6種類のアミノ酸は、創薬標的タンパク質として重要な膜タンパク質の膜貫通領域に多く存在することから、膜タンパク質の高感度NMR解析に特に有用であり、膜タンパク質を標的とした創薬開発に寄与できる。実際に創薬標的タンパク質であるβ₂ARおよびP2X受容体（Minato *et al.*, *Proc. Natl. Acad. Sci.* 2016）について、脂質二重膜中での動的構造を初めて解析することに成功した。したがって、生理的環境である脂質二重膜中での動的構造解析が可能となり、薬物候補化合物が結合した状態での薬効発現をより正確に予測できるようになった。

② 高分子量タンパク質の新規運動性解析法の開発

創薬標的タンパク質の動的構造情報を抽出するうえで問題となっていた分子量限界を克服し、リガンドの結合解離や酵素活性に重要なマイクロ秒からミリ秒の時定数で生じる動的構造情報を抽出するための多量子緩和解析法を開発した。開発した手法はプロトコール化が行われ、コントロールタンパク質を用いた一連の検討から有効性を実証したうえで投稿論文として発表し、細胞内シグナル伝達を担うGαi3タンパク質とその発がん性変異体のがん化機構の解明（Toyama *et al.*, *Nat. Commun.* 2017）、および分子量20万を超える膜タンパク質KirBacチャネルの機能解明に成功した（Toyama *et al.*, *J Am Chem Soc.* 2016）。

ii) 創薬標的タンパク質に対するリガンド結合部位の精密同定法の開発

動的構造の考慮により化合物の特異性と親和性を両立させる新たな技術の確立

タンパク質と複合体を形成した時の局所の運動性と空間相補性を指標に、結合状態で運動性が高く、空間相補性が低い個所を特定することで、より親和性が強く、特異性が向上したリガンドを得る方法を確立した

（Mizukoshi *et al.*, *Angew Chem Int Ed* (2016)。加えて、本手法を発展させ、ライブラリー化合物に多く含まれるフッ素核への適用を可能にした（Tokunaga Y *et al.*, *Molecules.* (2017)）。

iii) 創薬標的タンパク質の動的構造情報抽出法の開発及び得られた情報の計算科学的手法への導入

① GPCRのシグナル選択性を制御する動的構造の情報を抽出する方法の開発・応用

μオピオイド受容体、P2X創薬標的膜タンパク質における、活性を決定する動的構造平衡を解明することに成功した（Okude *et al.*, *Angew. Chem. Int. Ed.*, 2015, Minato *et al.*, *Proc. Natl. Acad. Sci.*, 2016）。さらに、β₂アドレナリン受容体、アデノシンA_{2A}受容体、CXCR4、μオピオイド受容体に加えて、δオピオイド受容体においても構造平衡が薬効度およびシグナル選択性に関与していることが明らかとなった。

また、アレスチンシグナルを活性化する状態である、C末端領域がリン酸化されたβ₂ARにおいて、リン酸化された部位と膜貫通領域が相互作用することにより、β₂ARがアレスチンシグナルの活性化に適した構造となることが明らかになった。このように、シグナル選択性を制御する動的構造の情報を抽出する方法を確立した（Shiraishi *et al.*, *Nat. Commun.* 2017）。

② 新 myPresto による探索的実証研究

新 myPresto により複数の GPCR のリガンド候補化合物を設計して、実際に合成した。活性を評価したところ、候補化合物の一つが部分アゴニストとして機能することが明らかとなった。さらに、NMR により解析したところ、本候補化合物存在下において、GPCR が活性型と不活性型が同程度存在する構造平衡状態にあることを示した。したがって、一連の新 myPresto の手法が、実際の新規化合物の設計に有用であることが実証された。

3. 分担研究開発課題名：X線及び電子線によるタンパク質及びその化合物複合体の精緻立体構造取得技術の開発

i) 構造不安定かつ高分子量タンパク質の構造解析向けに構造を安定化する技術の開発

構造が不安定で構造解析が不可能であった高分子量タンパク質の構造解析を可能にするために、変異を導入することで構造を安定化する技術開発を、GPCR であるエンドセリン受容体 B 型 (ET_BR) で行い、この技術開発に成功して、アゴニストとの複合体、リガンドフリー受容体、さらに 2 種類のアンタゴニストとの複合体の構造解析に成功して、*Nature* 誌などに発表した (*J. Mol. Biol.*, **428**, 2265-2274 (2016), *Nature*, **537**, 363-368 (2016), *Nature Structure & Molecular Biology*, **24**, 758-764 (2017))。この方法で作製した変異体の機能が WT と変わらないことも確認して、本技術のプロトコールを作成した。

ii) 生理的条件下における精緻な立体構造情報を取得する技術の開発

①電子線結晶学を用いることで生理的条件下における電位感受性 Na⁺チャンネルの構造解析に成功し、電位感受性 Na⁺チャンネルが速い gating をする機構の一端が理解できるようになった。この例の様に、膜内で膜タンパク質の構造を解析する技術開発を行い、膜内での構造解析の重要性と有用性を示した (*J. Mol. Biol.*, **425**, 4074-4088 (2013), *eLIFE*, **06119**, 10.7554 (2015))。

②浮腫や視神経脊髄炎治療の鍵となるアクアポリン - 4 (AQP4) について、生理的条件下で構造解析することで初めてチャンネル内の水分子を分離して観察できることを示していた。また、その阻害剤候補として見出していたアセタゾーラミド (AZA) はヒトの AQP4 を十分に阻害できないことを見出した。それゆえ、ラット AQP4 と AZA との複合体の構造解析を解析し (*Microscopy*, **65**, 177-184 (2016))、その分解能を 2.8 Å まで向上させることで、ヒトの阻害剤を開発する構造情報を得た。IT 創薬として阻害剤設計を可能にする精緻な立体構造情報を得る技術を開発した。

iii) 既知の立体構造情報に基づいて新規の標的タンパク質の立体構造を効率よく解析する技術の開発

①脊椎動物のギャップ結合チャンネルである、コネキシンの構造を結晶学で解析していたので、無脊椎動物のギャップ結合チャンネル、イネキシンの構造を同じく結晶学で解析することを目指した。低い分解能 (10 Å) でも既知の構造から興味深い解析が出来た (*J. Mol. Biol.*, **428**, 1227-1236 (2016))。さらに効率の良い解析を可能にするために、単粒子解析法を用いた膜タンパク質の高分解能の構造解析を行う上で必要な界面活性剤を除く方法を開発した (*Structure*, **23**, 1769-1775 (2015))。この GraDeR 法を開発することによって、イネキシンが形成するギャップ結合の構造を単粒子解析法を用いて、短期間に 3.3 Å 分解能で構造解析することに成功した (*Nature Commun*, **7**, 13681 (2016))。この例の様に、既知の情報を活用して効率よく構造解析する方法の開発に成功した。特に、単粒子解析が強力なことが確認できた。

②既知構造のホモログがある可溶性タンパク質の構造解析は試料の精製が出来れば、単粒子解析法を用いることで、数週間での構造モデル作製が可能な技術開発を行った。

iv) タンパク質/化合物複合体の立体構造を解析する技術の開発

①クローディン関連の構造解析で 2 報を *Science* 誌などに発表した (*Science*, **344**, 304-307 (2014), *Science*, **347**, 775-778 (2015))。

② i) において記述した ET_BR とアゴニスト及びアンタゴニストとの複合体の構造を解析して *Nature* 誌などに発表した (*J. Mol. Biol.*, **428**, 2265-2274 (2016), *Nature*, **537**, 363-368 (2016), *Nature Structure & Molecular Biology*, **24**, 758-764 (2017))。

- ③ ii) において記述したアクアポリン 4 と阻害剤の複合体の構造を 10 Å、3 Å、2.8 Å という 3 種類の分解能で解析して、3 Å より高い分解能の解析が創薬において重要であることを確認し、Drug Rescuing を実証できるようにした。
- ④ iii) において記述したギャップ結合チャンネルの高効率の単粒子解析は特筆できる。多く創薬標的タンパク質の構造が速く解析出来るようになった。

本プロジェクトの成果の 1 つとして、株式会社 CeSPIA を 2017 年 4 月 17 日に創設した。株式会社 CeSPIA では、製薬企業からの依頼に答えて構造解析を実施している。

The database of chemical compounds has been developed for docking screening procedures based on the electronic catalogues of the industries, and new methods applying machine-learning approach have been developed to estimate the values of activities and LogS/LogD of chemical compounds from public databases as the “Big Data”. Several software programs have been created for designing and synthesizing new chemical compounds and for modeling the complex structures of pairs of ligand-protein by considering their dynamic natures. New ensemble docking procedures with several new algorithms have been also developed to compute an accurate binding free energy by the thermodynamic integration along an appropriate binding/unbinding path that is found from the free-energy landscape. Based on the above methods and software including a GUI portal with the cloud usage, a new program suite “myPresto version 5” has been designed, developed and released from our Web site, <https://mypresto5.jp/>.

By applying the above new software and systems, putative hit compounds have been found for several target proteins, in particular, a transcription factor and a kinase.

Dynamic properties of membrane proteins in lipid bilayers are difficult to investigate owing to the low sensitivities of their NMR signals. We successfully increased the NMR sensitivities by developing novel stable isotope labeling methods in baculovirus-insect cell expression systems and revealed that the conformational equilibria of β_2 -adrenergic receptor (β_2 AR) and P2X₄ channel in lipid bilayers are important for their functions. We also developed the NMR methods to characterize chemical exchange processes of proteins on the order of microseconds to milliseconds, which could be applied to KirBac1.1, with an apparent molecular mass of 200 kDa. Using these methods, we showed that the oncogenic mutation of G α accelerates GDP dissociation by shifting the equilibrium towards the excited conformation.

The thermodynamic properties of a ligand in the bound state affect its binding specificity. We experimentally evaluated the local dynamics and the surface complementarity of weak-affinity ligands in the receptor-bound state by forbidden coherence transfer analysis in free-bound exchange systems (Ex-FCT). The Ex-FCT analyses successfully provided information for the rational design of a ligand with higher affinity and preferable thermodynamic properties for p38 α .

G-protein-coupled receptor (GPCR) ligands impart differing degrees of signaling in the G-protein and arrestin pathways, in phenomena called "biased signaling". We found that the intracellular cavity of μ -opioid receptor exists in an equilibrium between closed and multiple open conformations and that the population of each open conformation determines the G-protein- and arrestin-mediated signaling levels. We also revealed that the phosphorylated C-terminal region of β_2 AR, which is important for arrestin-mediated signaling, adheres to either the intracellular side of the transmembrane region or lipids, and that the phosphorylation of the C-terminal region allosterically alters the transmembrane conformation of β_2 AR. These findings provide insight into the biased signaling of GPCRs and will be helpful for development of ligands with reduced side effects.

By electron crystallography of two-dimensional crystals, we analyzed structure of a gap junction channel of innexin (invertebrate GJ channel), and observed a plug density which was similar density observed in gap junction channel of connexin (vertebrate GJ channel). We revealed that innexin forms the gap junction channel by 16 subunits (hemi-channel by 8 subunits), while connexin channel is formed by 12 subunits (hemi-channel by 6 ones). By single particle method, structure of the gap junction channel of innexin was very effectively analyzed at 3.3 Å resolution (**Nature Commun**, **7**, 13681 (2016)). Single particle analysis enabled us to observe detailed structure of innexin channel including short helix (possibly forming plug helix) at N-terminal side as well as cytoplasmic domain, which was not observed by crystallographic analyses of connexin and might be importantly related with gating mechanisms of the gap junction channel.

By electrophysiological experiments, we confirmed that two important residues in the prokaryotic voltage-gated Na⁺ channels for binding of Local Anesthetics including quaternary ammoniums (**FEBS J.**, **283**, 2881-2895 (2016)). For a new strategy of drug development named Drug Rescuing, structure of water channel AQP4 and Acetazolamide complex was analyzed by electron crystallography (**Microscopy**, **65**, 177-184 (2016))

A GPCR; the endothelin type B receptor (ET_BR) which is involved in several important physiological processes, such as, maintenance of vascular tone, humoral homeostasis, neural crest cell development and neurotransmission. We identified thermostabilizing mutations on 11 residues of ET_BR. Furthermore, the combination of the five residues among these 11 mutations gave 17°C higher T_m value than that of wild type. This mutant was confirmed to exhibit an affinity for the endogenous agonist endothelin-1 and activate G_q similar to the wild type (**JMB**, **428**, 2265-2274 (2016)). Utilizing this thermostable mutant, we analyzed structures of ET_BR in complex with ET-1 as well as in the ligand-free form. We observed the agonist-induced conformational changes and identified 36 residues in the receptor which bind ET-1 in a virtually irreversible manner. The ET-1 structure binding in the receptor was completely different from any analyzed structures of ET-1 alone without binding in the receptor (**Nature**, **537**, 363-368 (2016)). The results strongly suggest importance of structure analysis of a ligand binding in the receptor. The complex structures of ET_BR with antagonists were also analyzed by X-ray crystallography (**Nature Structure & Molecular Biology**, **24**, 758-764 (2017)).

For future attempt to develop Drug Delivery system, we analyzed structure of claudin-15 and made clear the transmembrane segments form a tight left-handed four-helix bundle and two extra cellular segments form a characteristic β-structure which we named a Palm (**Science**, **344**, 304-307 (2014)). Furthermore, we analyzed structure of claudin-19 and C-terminal domain of *Clostridium Perfringens* Enterotoxin (C-CPE) complex, and found key interactions for disrupting the tight junction strands (**Science**, **347**, 775-778 (2015)).

Based on advanced techniques of structure analysis developed in this project, a business venture, CeSPIA inc., was founded for structure determination service of drug target proteins and accommodates requests from pharmaceutical companies.