特集: ウイルス感染症に取り組む生物物理学

総説

SARS-CoV-2タンパク質構造モデルによる薬剤候補予測

土方敦司¹, 塩生くらら¹, 中江 摂¹, 塩生真史¹, 太田元規², 金谷重彦³, 白井 剛¹

¹長浜バイオ大学バイオサイエンス学部

²名古屋大学大学院情報学研究科

³奈良先端科学技術大学院大学先端科学技術研究科

The novel coronavirus disease (COVID-19) pandemic has emerged in late 2019 and rapidly spread all over the world. In order to assist structure-based discovery efforts for repurposing drugs against the infectious disease, we constructed homology models of SARS-CoV-2 proteins. We identified several potential drugs by comparing the ligand molecules in the template structures with approved or experimental drugs and compounds of natural drugs, including carfilzomib, sinefungin, tecadenoson, and trabadenoson, that would be further investigated for their potential for treating COVID-19.

homology modeling / protein structure / drug repurposing / infectious disease

1. はじめに

2019 年 12 月に中国・武漢市で見つかった新型コロナウイルス (SARS-CoV-2) 感染症 (COVID-19) は瞬く間に世界中に拡大し, 2020 年 3 月 11 日に WHO は COVID-19 のパンデミックを宣言した. 2020 年 9 月の段階で, 全世界の感染者は 2800 万人, 死者数は 90 万人に達している¹⁾. 我が国においても, 社会活動および大学等における研究・教育活動が制限されるなど, 多大な影響を及ぼしている. この国難とも呼ぶべき事態を収束させるには, 一刻も早い治療薬ならびにワクチンの開発が必要であり, 各国の研究機関や製薬企業において COVID-19 に対する治療薬・ワクチンの研究開発が行われている. 新薬を一から設計・開発することは時間もコストもかかるため, 既存医薬品の転用 (ドラッグリポジショニング) の検討が最も近道であり, 実際にいくつかの既存医薬品について治験が行われている.

抗ウイルス治療薬の開発には大きく分けて二つの戦略が挙げられる. 一つ目は, ウイルスの宿主細胞内での生活環を抑えること, もう一つは, 宿主免疫系の暴走をコントロールすることである. 第 1 の戦略に基づいてウイルスタンパク質の機能を阻害する薬物候補を

探索する上で, タンパク質立体構造情報は極めて有用である. 本総説では, 我々が最近行った SARS-CoV-2 タンパク質の立体構造の知識ベースモデリングに基づく COVID-19 治療薬候補予測について述べる²⁾.

2. SARS-CoV-2 タンパク質の立体構造

SARS-CoV-2 は, 一本鎖プラス鎖 RNA ウイルスに分類され, そのゲノム配列は 2020 年 1 月に解読・公開された³⁾. SARS-CoV-2 のゲノムサイズは約 3 万塩基であり, 11 の遺伝子 (Open Reading Frame, ORF) がコードされている. それぞれの遺伝子は, 1 種類の非構造タンパク質 (orf1ab), 4 種類の構造タンパク質 (スパイク (S) タンパク質, エンベロープ (E), メンブレン (M), スクロオカプシドタンパク質 (N)) および 6 種類の付属タンパク質 (ORF3a, ORF6, ORF7a, ORF7b, ORF8 および ORF10) をコードしている (図 1 左).

orf1ab は翻訳後, 自身がコードするパピリン様プロテアーゼ (nsp3, PL-pro) およびメインプロテアーゼ (M-pro) によって切断され, 16 種類のタンパク質 (nsp1 から nsp16) に分かれる. 図 1 右は, 現時点までの SARS-CoV-2 タンパク質の立体構造決定の状況を

Knowledge-based Modeling of SARS-CoV-2 Proteins and Predicting its Potential Drugs

Atsushi HIJIKATA¹, Clara SHIONYU-MITSUYAMA¹, Setsu NAKAE¹, Masafumi SHIONYU¹, Motonori OTA², Shigehiko KANAYA³ and Tsuyoshi SHIRAI¹

¹Faculty of Bioscience, Nagahama Institute of Bio-Science and Technology

²Graduate School of Information Science, Nagoya University

³Graduate School of Science and Technology, Nara Institute of Science and Technology

図 1 (左) SARS-CoV-2 ウイルスゲノムの遺伝子マップ。コードするタンパク質の種類ごとに色分けしている:非構造タンパク質 (青), 構造タンパク質 (オレンジ), 付属タンパク質 (緑)。(右) 各タンパク質の立体構造カバー領域。横軸はアミノ酸残基数を表す。

示す。

SARS-CoV-2 は、2002 年から 2003 年にかけて発生した重症急性呼吸器症候群 (SARS: severe acute respiratory syndrome) の病原体である SARS コロナウイルス (SARS-CoV) とゲノム配列一致度が約 80% と非常によく類似しており、コードするタンパク質の多くが高度に保存されている (配列一致度:28%–99%)。また、SARS 発生当時から現在まで、数多くの SARS-CoV タンパク質の立体構造が実験的に決定されている。2020 年 1 月末時点の Protein Data Bank Japan (PDBj) (<https://pdj.org>) のデータを用いて、SARS-CoV-2 タンパク質に対するアミノ酸配列の相同性検索を行ったところ、26 タンパク質中 17 個については、有意な配列類似性を示す立体構造既知タンパク質が存在し、そのうち、nsp4 を除く 16 個は SARS-CoV のタンパク質の立体構造であった。さらに、2020 年 1 月から 8 月末までの期間に PDBj に登録された SARS-CoV-2 タンパク質を時系列順で調べてみると、2 月頃から徐々に登録が増え始め、わずか 8 ヶ月余りの間に、部分構造も含めて 16 種類の SARS-CoV-2 タンパク質の立体構造が決定されている。これはこのウイルスがコードする全タンパク質のアミノ酸配列長のうち実に 54% について立体構造の情報が得られた計算になる (図 2)。

同時期における、ウイルスタンパク質をターゲットとした治験の状況を調べてみると、RNA 依存性 RNA ポリメラーゼ (RdRp) および M-pro について、立体構造が登録された時期とほぼ同時期に、それらをターゲットとした治験が行われている (図 2 に示した治験は明確に SARS-CoV-2 タンパク質を標的とした医薬品のうち、2020 年 8 月末までになんらかの結果が公表されたものに限られる)。いずれの医薬品も既存抗ウイルス薬の転用を狙って行われたものである。しかしながら、これまでのところ COVID-19 に対して有意な有効性が示されたのはレムデシビルのみである⁴⁾。

3. タンパク質-リガンド立体構造複合体情報に基づいた薬剤探索アプローチ

本総説で紹介する研究では、2020 年 1 月末時点で構造未知であった SARS-CoV-2 タンパク質のうち 17 個について、知識ベースモデリング (タンパク質構造をホモロジーモデリングで予測し、低分子を含む複合体構造の構築を既知複合体構造情報との重ね合わせに基づいて行う構造モデリング) によって立体構造モデルを構築した (図 3)。これは SARS-CoV-2 の全タンパ

図 2

(上) 2020 年 1 月から 8 月までの期間の SARS-CoV-2 タンパク質の立体構造情報の時系列変化。本総説のモデリングが行われた時点を赤矢印で示した。(下) 同時期における、ウイルスタンパク質をターゲットとした医薬品の治験の開始と終了時期。

図 3

SARS-CoV-2 タンパク質の立体構造。ORF3a, ORF8 以外は知識ベースモデリングによって我々が構築したモデル。赤丸で囲った部分 (M-pro, 2oMT, ACE2) はモデル構造に含まれるリガンド結合部位を示す。ORF3a および ORF8 は、最近になって実験的に決定された。

ク質のアミノ酸配列長の 65% に相当する。

多くのタンパク質は、ホモまたはヘテロ多量体の形で立体構造モデルが得られた。例として、M-pro はホモ二量体、nsp10 はエクソヌクレアーゼ (ExoN) および 2-O'-リボースメチル基転移酵素 (2oMT) とそれぞれヘテロ二量体で阻害剤複合体のモデルが得られた。

また、S タンパク質はホモ三量体を形成し、さらにレセプター結合ドメイン (RBD) とヒトのアンギオテンシン変換酵素 2 (ACE2) との複合体モデルが得られた。

これらの立体構造モデルに基づいて医薬品の候補 (ウイルスタンパク質の特定の部位に結合する化合物) を探し出すインシリコ解析手法として、バーチャルス

クリーニングが挙げられる。この手法は典型的には、標的となるタンパク質の立体構造上で医薬分子を結合させたい部位を定義し、化合物ライブラリーの分子を網羅的に計算機上でドッキングし、エネルギー関数などの評価関数で結合安定性を評価して候補（ヒット）化合物を抽出する。近年では並列計算・機械学習を利用することで著しい高速化が実現しているものの、標的タンパク質や化合物ライブラリーが絞り込めていない状況での適用は容易でない。

そこでこの研究では、知識ベースモデリングによりこの過程の単純なショートカットを試みた。既知複合体構造で標的タンパク質ホモログに結合しているリガンドはファーマコフォア（リガンド結合部位などの化合物を結合させたい箇所が特異的に認識する構造的特徴）を内包している可能性が高く、そのリガンドと類似の構造を持ち、構造モデルに無理なくドッキング可能な既承認薬があれば、その分子は治療薬候補になりうると予想される。

今回モデリングした、SARS-CoV-2 タンパク質モデルのうち、M-pro ホモ二量体、S タンパク質-ACE2 複合体、および 2oMT-nsp10 ヘテロ二量体の3つのモデルでは、鋳型構造に結合したリガンド分子が存在したので、KEGG DRUG⁵⁾ならびに KNApSACk⁶⁾ データベースを用いて、それらのリガンドと構造が類似している承認薬ならびに生薬成分を検索した。

M-pro はウイルスタンパク質生産に必須の酵素であり、2002 年の SARS アウトブレイク当時からコロナウイルス疾患の有望な創薬標的とみなされ、多くのペプチドミメティック阻害剤との複合体構造解析がなされている。しかし予想に反して、これらの既知リガンドに高い類似性を示す既存薬分子は見いだされなかった。これは SARS-CoV-2 の M-pro がシステインプロテアーゼであるのに対して、HIV プロテアーゼなど既存の抗ウイルスプロテアーゼ阻害剤の標的の多くがアスパラギン酸／亜鉛プロテアーゼであり、これらの酵素の構造的特徴が大きく異なることが影響したと考えられる。

実際、HIV プロテアーゼ阻害剤であるロピナビル／リトナビル（一般名：カレトラ）は COVID-19 の有望な治療薬と期待されていたが、治験において有効性は示されなかった⁷⁾。また同じく HIV プロテアーゼ阻害剤であるダルナビルも有効性を示せなかったと報告されている（図 2）。これは、ドラッグリポジショニングの観点からは、M-pro が優れた標的ではない可能性を示唆している。一方、本研究で既知リガンドに最も高い類似性を示したのは、カルフィルゾミブであっ

た。カルフィルゾミブは、不可逆的プロテアソーム阻害剤であり、多発性骨髄腫またはワルデンシュトレームマクログロブリン血症の治療薬として承認されている。カルフィルゾミブの標的はスレオニンプロテアーゼ（求核攻撃基 Thr）であるが、M-pro の求核攻撃基 Cys とも反応することが示されている。

ウイルス表面の S タンパク質は、宿主細胞に感染する際にヒト ACE2 を受容体として利用するため、S タンパク質-ACE2 の結合部位は重要なターゲットの一つである。ACE2 は高血圧治療薬の主要な標的である ACE のホモログ（アミノ酸配列一致度 44%）であり、いくつかの ACE-阻害剤複合体構造が既知である。これらの阻害剤と類似した承認薬として、いずれも血圧降下薬であるリシノプリル、エナラプリラート、カプトプリルが見つかった。しかし、これらの分子は S タンパク質-ACE2 の相互作用部位とは異なる位置に結合していたため、S タンパク質との相互作用を直接的に阻害できないと予想された。

現在 S タンパク質を標的とする医薬品として、RBD を抗原とする抗体医薬品の治験が実施されているが、この部位を標的とした低分子医薬品は提案されていなかった。しかし最近筆者らも参加した研究から、タマサキツラフジ由来成分のセファランチン（円形脱毛症・白血球減少症治療薬として承認されている）が、S タンパク質-ACE2 相互作用を阻害し、SARS-CoV-2 感染を抑制することが示された⁸⁾。これは、セファランチンからの展開による COVID-19 感染予防薬開発の可能性を示唆する知見である。

2oMT-nsp10 複合体は、ウイルス RNA の 5' 末端のキャップ構造にメチル基を修飾する酵素の一つである。キャップ構造は、ウイルス RNA を宿主による RNA 分解から防ぐとともに、宿主の翻訳機構を利用して自身のタンパク質を合成する上でも必要不可欠である。モデルの鋳型となった SARS-CoV の 2oMT のリガンド複合体構造からいくつかの既存薬が発見されたが、その一つシネフンギンはストレプトマイセス属の細菌が産生する抗生物質であり、試験管内の実験で抗ウイルス活性が報告されている⁹⁾。さらにアデノシン A1 受容体作動薬（ただし現時点で未承認）であるテカデノソン、セロデノソンおよびトラボデノソンも 2oMT を標的にできることが予想された。このうちテカデノソンとトラボデノソンはフェーズ 1 の治験を通過していることから、安全性は確認されている^{10),11)}。

また、ウイルス RNA 分子の 5' 末端にキャップ構造を付加する酵素（グアニル酸転移酵素）が存在するはずであるが、実際にどのタンパク質がその役割を果た

すのかは今のところ不明であり、今後同定されれば創薬標的になりうる。これらの結果は、現時点ではあまり着目されていない2oMTを含むキャップ合成系酵素群が有効な創薬標的であることを示唆している。

4. おわりに

近年、特に標的創薬・構造創薬の発展により、創薬過程は著しく合理化されたとされるが、今回のCOVID-19パンデミックはその実力を証明することを突然要求された形である。SARS-CoV-2のゲノムやタンパク質立体構造の情報（図2）がハイペースで蓄積されたことは、基礎生物学の進歩の明確な成果であり、このパンデミックの発生が10年早ければこのような即応は不可能であった。一方で本総説でも概観したように、我々の生活を元に戻せるほど有効な治療薬はまだ発見されていない。SARS-CoV-2タンパク質を標的とした新しい治療薬の開発はまだ道半ばであり、ここで解説した研究もその一助になることを切に願っている。

本総説で紹介した研究はAMED創薬等先端技術支援基盤プラットフォーム（BINDS）の支援により実施されたもので、構造モデルはBINDS（<https://www.binds.jp/SARS-CoV-2/>）またはBSM-Arc（<https://bsma.pdbj.org/featured/covid-19>）から利用可能である。

文 献

- 1) WHO Coronavirus Disease (COVID-19) Dashboard. <https://covid19.who.int/>.
- 2) Hijikata, A. *et al.* (2020) FEBS Lett. **594**, 1960-1973. DOI: 10.1002/1873-3468.13806.
- 3) Wu, F. *et al.* (2020) Nature **579**, 265-269. DOI: 10.1038/s41586-020-2008-3.
- 4) Beigel, J. H. *et al.* (2020) N. Engl. J. Med. **383**, 1813-1826. DOI: 10.1056/NEJMoa2007764.
- 5) Kanehisa, M. *et al.* (2012) Nucleic Acids Res. **40**, D109-D114. DOI: 10.1093/nar/gkr988.
- 6) Afendi, F. M. *et al.* (2012) Plant Cell Physiol. **53**, e1. DOI: 10.1093/pcp/pcr165.
- 7) Cao, B. *et al.* (2020) N. Engl. J. Med. **382**, 1787-1799. DOI: 10.1056/NEJMoa2001282.
- 8) Ohashi, H. *et al.* (2020) bioRxiv. DOI: 10.1101/2020.04.14.039925.
- 9) Kuroda, Y. *et al.* (2019) J. Antibiot. (Tokyo). **72**, 981-985. DOI: 10.1038/s41429-019-0234-4.
- 10) Peterman, C., Sanoski, C. A. (2005) Cardiol. Rev. **13**, 315-321. DOI: 10.1097/01.crd.0000181621.84565.9d.
- 11) Li, G. *et al.* (2018) Invest. Ophthalmol. Vis. Sci. **59**, 383-392. DOI: 10.1167/iovs.17-23212.

土方敦司

土方敦司（ひじかた あつし）

長浜バイオ大学プロジェクト特任講師
日本学術振興会特別研究員，理化学研究所リサーチアソシエイトを経て2012年11月より現職。
研究内容：構造バイオインフォマティクス
連絡先：〒526-0829 滋賀県長浜市田村町 1266
E-mail: a_hijikata@nagahama-i-bio.ac.jp

塩生くらら

塩生くらら（しおにゅう くらら）

長浜バイオ大学プロジェクト研究員
2003年名古屋大学大学院工学研究科修了，2017年より現職。
研究内容：構造情報生物学
連絡先：同上
E-mail: claras@nagahama-i-bio.ac.jp

中江 摂

中江 摂（なかえ せつ）

長浜バイオ大学助手
大阪府立大を経て現職。
研究内容：構造生物学
連絡先：同上
E-mail: s_nakae@nagahama-i-bio.ac.jp

塩生真史

塩生真史（しおにゅう まさふみ）

長浜バイオ大学准教授
2002年名古屋大学大学院理学研究科修了，日本学術振興会特別研究員などを経て，2005年長浜バイオ大学バイオサイエンス学部講師，2014年より現職。
研究内容：構造情報生物学
連絡先：同上
E-mail: m_shionyu@nagahama-i-bio.ac.jp

太田元規

太田元規（おおた もとのり）

名古屋大学大学院情報学研究科教授
1996年早稲田大学修了，遺伝研，東工大を経て2008年より現職。
研究内容：生命情報学
連絡先：〒464-8601 愛知県名古屋千種区不老町
E-mail: mota@i.nagoya-u.ac.jp

金谷重彦

金谷重彦（かなや しげひこ）

奈良先端科学技術大学院大学教授
山形大学工学部を経て2004年より現職。
研究内容：バイオインフォマティクス
連絡先：〒630-0101 奈良県生駒市高山町 8916-5
E-mail: skanaya@gtc.naist.jp

白井 剛

白井 剛（しらい つよし）

長浜バイオ大学教授
名古屋大学，BERIを経て現職。
研究内容：構造情報生物学
連絡先：〒526-0829 滋賀県長浜市田村町 1266
E-mail: t_shirai@nagahama-i-bio.ac.jp