

編集後記

構造活性相関部会・ニュースレター <1 April, 2021>

SAR News No.40

「目次」

///// Perspective/Retrospective /////					
SARS-CoV-2 RNA 依存性 RNA ポリメラーゼ	の分子基盤				
	理化学研究所	関根	俊一	•••	1
///// Cutting Edge /////					
機械学習とバーチャルスクリーニングを通し	ごた				
SARS-CoV-2 RNA 依存性 RNA ポリメラーゼ	の阻害剤予測				
株式会社 Elix 二	Nazim Medzhidov、	結城	伸哉	•••	11
理論創薬及び関連分野における SARS-CoV-2	2研究動向 2020~2	2021/3			
	理化学研究所	高谷	大輔	•••	23
///// SAR Presentation Award /////					
受賞コメントおよび発表要旨				•••	30
///// Activities /////					
<報告>					
第 48 回構造活性相関シンポジウム 開催報告	E -				
	理化学研究所	本間	光貴	•••	42
<会告>					
構造活性フォーラム 2021 会告					
	横浜市立大学	池口	満徳	•••	45

••• 46

SARS-CoV-2 RNA 依存性 RNA ポリメラーゼの分子基盤

理化学研究所 関根 俊一

1. はじめに

2020年は予期せぬ異例の年になった。2019年末に中国の武漢で報告された新型コロナ ウイルスによる感染症(COVID-19)が、2020年3月頃には世界各地で広がりを見せ、瞬く間 にほぼすべての国に広がってパンデミックとなった。日本も例外ではなく、我々もこれまで に経験したことのない生活を余儀なくされている。本稿執筆時点(2021年2月)で、国内 で40万人以上が感染し、死者は7,000人を超える。感染を抑えるためには、ワクチンの接 種による集団免疫の獲得が不可欠な状況であるが、並行して感染者を治療する治療薬の開 発も急務である。

COVID-19 の原因である重症急性呼吸器症候群コロナウイルス 2 (SARS-CoV-2) は、 ニドウイルス目コロナウイルス亜科に属するβコロナウイルスの一種である。(+)1 本鎖 RNA をゲノムに持つ RNA ウイルスであり、そのゲノムサイズは約 30kb と RNA ウイルスの中で は最大級である。コロナウイルスのゲノムの前半部分 2/3 は 2 つの長いオープンリーディン グフレーム (ORF1a, ORF1b) を含んでおり、これらは非構造タンパク質 (nsp1~16) をコー ドしている(*1, 2*)。ORF1a から nsp1~11 を含む pp1a というポリタンパク質が翻訳される。ま た、リボソームのフレームシフトにより、ORF1a に続けて ORF1b が翻訳されたときにのみ 生じる pp1ab というポリタンパク質は nsp1~10 に加えて nsp12~16 を含む。pp1a と pp1ab は 3C 様プロテアーゼ (3CLpro, nsp5) およびパパイン様プロテアーゼ (PLP, nsp3) でプロセス され、個々の非構造タンパク質(nsp)となる(3)。ゲノムの後半部分 1/3 は 4 種類の構造タン パク質(spike (S), nucleocapsid (N), membrane (M), envelope (E))に加え、いくつかのアクセサリ ータンパク質をコードしている。この領域からは、それぞれの遺伝子をコードした分節化し た mRNA が合成される。

コロナウイルスのゲノム RNA の複製には nsp7~16 タンパク質が深く関わっている。 nsp12 は RNA 合成を担う RNA 依存性 RNA ポリメラーゼ (RdRp)の本体であり、他の nsp タンパク質がサブユニットとして加わることで、ゲノム複製に必須の機能を付与している と考えられている。感染症をコントロールするためには、ウイルスの複製機構の理解とワク チンや特効薬の開発が鍵となるが、ウイルス RNA の複製を担う RdRp は抗ウイルス薬の開 発の主要なターゲットのひとつである。抗 SARS-CoV-2 薬として広く認知されているレム デシビルは RdRp を標的としたヌクレオチドアナログである。世界的な流行が発生してか ら、SARS-CoV-2 の複製や阻害の分子メカニズムに関する研究報告が急激に増加した。本稿 では、最近の知見を構造生物学的視点で俯瞰したい。

2. SARS-CoV-2 RNA 依存性 RNA ポリメラーゼ (RdRp)の構造

コロナウイルスの RNA の複製・転写を司る RNA 依存性 RNA ポリメラーゼ (RdRp) は、ウイルスがコードする 3 種類の nsp タンパク質 (nsp7, nsp8, nsp12) からなる複合体酵素 である。nsp12 タンパク質は、RNA 合成を触媒する RdRp の本体であり、コロナウイルスの 複製サイクルにおけるキーコンポーネントである (6)。nsp12 はそれ単独ではほとんど活性 がなく、その機能には 2 種類のアクセサリーサブユニット(nsp7, nsp8)が必要で、それらが会 合して活性のある RdRp コア複合体を形成する(4, 5)。



図 1. SARS-CoV-2 RdRp の構造

A. SARS-CoV-2 RdRp (nsp12-nsp82-nsp7) のクライオ電子顕微鏡構造 (PDB: 6M71 を基に作図)。B-D. RNA が結合した SARS-CoV-2 RdRp の伸長複合体のクライオ電子顕微鏡構造 (PDB: 6YYT を基に作図)。

2020 年 4 月に入り、海外の複数の研究機関から SARS-CoV-2 RdRp のクライオ電子顕微 鏡構造が相次いで報告された(6-9)。SARS-CoV-2 の RdRp の構造は、2002 年に発生した SARS-CoV や中東呼吸器症候群関連コロナウイルス MERS の RdRp の構造とよく似ている(8)。 nsp12 は、ウイルスの RNA ポリメラーゼに典型的な"右手"型の RdRp ドメイン、ニドウイ ルスの RdRp に固有の RdRp-associated nucleotidyltransferase (NiRAN) ドメイン、それらをつ なぐ Interface ドメインからなっている (図 1A)。 RdRp コア複合体には、1 分子の nsp7 サ ブユニットと2 分子の nsp8 サブユニットが含まれている。nsp7 と nsp8 は nsp12 の"右手"ド メインの"親指(thumb)"部分に結合し、もう 1 分子の nsp8 は"指(fingers)"に結合する。RNA 合成の活性部位は、"手のひら(palm)"上に位置し、5 つの保存されたモチーフ (モチーフA ~E) で構成されている。モチーフ C には、RNA 合成に必須のアスパラギン酸残基(D760, D761)が存在し、ここに活性に必要な金属イオン(Mg²⁺, Mn²⁺)が結合し、それを介して基質で あるヌクレオシド 3 リン酸(NTPs)と新生 RNA 鎖の 3′末端が結合する(9)。

続けて報告された RNA を結合した状態の RdRp 複合体の構造によれば、コロナウイル スの RdRp は他のウイルスの RdRp には見られない特徴的な RNA 結合様式を持っている (図 1B-D) (6, 7)。2 分子の nsp8 サブユニットは、C 末端側の球状ドメインを介して RdRp の本体に結合する一方で、N 末端側の長いヘリックス("extension")を介して伸長中の2重 鎖 RNA に結合している。このヘリックスは伸長中の RNA に沿うように伸びており、RNA の2 重らせんのおよそ2 巻き分をカバーしている。長いヘリックスの片面には正電荷をも ったアミノ酸残基が規則的に配置されており、RNA のリン酸バックボーンと相互作用して いる。全体として、RdRp の活性中心に近い約 10 塩基対の部分を nsp12 の"右手"がホールド し、それより下流の約 20 塩基対にわたる部分を2 つの nsp8 extension がガイドするかたち になっており、RNA 合成途中の RNA の脱落を防ぐのに役立っていると考えられる。

3. RdRp 阻害剤 -レムデシビル-

新型コロナウイルス感染症の発生以来、レムデシビルやアビガンといった薬剤の名前 をニュース等で頻繁に目にするようになった。これらは、ウイルスの RNA 合成の材料とな るヌクレオチド類似物質の前駆体であり、RdRp の働きを阻害する化合物である。特に、レ ムデシビル (remdesivir) は、元々はエボラ出血熱の治療薬として開発されていたものである が、現在では COVID-19 治療薬としての認知度が高い。米国では、FDA(食品医薬品局)が 2020 年 5 月に重症入院患者を対象に緊急時の人道的使用の許可を与え、のちに正式承認さ れた。日本でも、FDA の使用許可を受けて特例承認されている。



4残基目まで伸長した段階で、 C1'-シアノ基と立体障害を生じる

図 2. SARS-CoV-2 RdRp によるレムデシビルの取り込みと阻害機構

A. RNA の 3'末端にレムデシビルを取り込んだ SARS-CoV-2 RdRp の構造 (PDB: 7BV2 を基 に作図)。B. レムデシビルを取り込んだ後、さらに 3 塩基伸長した状態の構造 (PDB: 7B3C を基に作図)。

レムデシビルはコロナウイルスを含む一本鎖 RNA ウイルスに対して抗ウイルス活性 を示し、その標的は RdRp である。プロドラッグであり、体内(細胞内)に取り込まれて保 護基が外れ、三リン酸が付加されてレムデシビル三リン酸(RTP)の状態となってはじめて 阻害活性を示す(10)。RTP は ATP アナログであり、アデニンに類似した環とリボースの 1' 位から突き出た 1'-シアノ基を持っているのが特徴である。2020 年には、レムデシビルの作 用に関する生化学・構造生物学研究が相次いで報告された。それによれば、レムデシビルの作 用に関する生化学・構造生物学研究が相次いで報告された。それによれば、レムデシビルは アデノシン(A)と同様に RNA 鎖の 3'末端に Uを鋳型として取り込まれる(9)(図 2A)。しか し、この段階では転写は止まらずに、RdRp は最大で 3 残基分のヌクレオチドをさらに追加 することができる(11)。したがって、レムデシビルはいわゆる chain terminator ではないが、 4 残基を超えて RNA が伸長するのを阻害する。レムデシビルを取り込んだ RdRp のクライ オ電子顕微鏡構造によれば、その原因は、レムデシビルの 1'-シアノ基が立体障害となり、 RNA が 4 残基伸長した時点で立体障害が生じ、それ以上のトランスロケーションが妨げら れるためである(12)(図 2B)。構造と配列の比較によれば、レムデシビルの阻害様式は他の RNA ウイルスの RdRp にも適用可能で、ヌクレオチドアナログをベースとした幅広いスペ クトルをもった抗ウイルス薬を設計できる可能性が示唆されている。

4. エキソヌクレアーゼ

コロナウイルスは RNA ウイルスの中で最大サイズのゲノムを持つ(13)。大きいゲノム サイズを維持するためには RNA 合成の精度が要求されるため、RNA 合成中のエラーの校 正機構を持っているのがコロナウイルスの大きな特徴である。nspl4 はエキソヌクレアーゼ (ExoN)活性を持つサブユニットで、RdRp コア複合体に RNA 合成に影響を与えずに結合す ることができ、転写エラーの校正を司ると考えられている。ExoN は、ゲノムサイズ 20 kb 以上のニドウイルスにおいて保存されており(14,15)、その変異は RNA 複製の精度に影響を 与える(16-18)。一般的な RNA ウイルスの複製精度 (10⁻³~10⁻⁵) と比較して、SARS-CoV の低 い突然変異率 (10⁻⁶~10⁻⁷) は ExoN 活性で説明される。また、nspl4 の ExoN 活性を欠いたコ ロナウイルス変異体は、レムデシビルに対してより高い感受性を示し(19)、ExoN 活性が薬 剤の有効性を鈍らせている可能性が示唆されている(20)。



図 3. nsp14-nsp10 複合体とエキソヌクレアーゼ

A. nsp14-nsp10 複合体の構造。B. nsp14 の ExoN 活性部位。(A, B ともに SARS-CoV の結晶構造 PDB: 5C8S を基に作図)

コロナウイルスがコードする nsp14 タンパク質は 2 つのドメインからなり、N 末端ド メインが ExoN 活性を、C 末端ドメインが N7-MTase 活性(後述)をそれぞれ司る(*15*)(図 3A)。nsp14 は ExoN ドメインを介して nsp10 タンパク質と安定な複合体を形成する(*21*)。 ExoN ドメインは DEDD スーパーファミリーのエキソヌクレアーゼに類似した構造を持っ ている(図 3B)。また、コロナウイルスに特有の 2 つの Zn フィンガーモチーフをもってお り、これらは ExoN 活性に必須である。また、nsp10 も ExoN 活性に必須である。nsp10 と nsp14 との間には広範な相互作用が見られ、nsp10 は nsp14 の ExoN ドメインの構造を安定 化させ、活性を促進すると考えられる(*22*)。nsp14 がどのようにして RNA の二重らせんに取 り込まれたミスマッチ塩基やレムデシビル等を除去するのかは不明であるが、nsp13 タンパ ク質(後述)が 5' \rightarrow 3'ヘリカーゼ活性で RNA を巻き戻し、nspl4 の ExoN の活性部位に RNA の 3'末端を導くのではないかというモデルが提唱された(23)。

5. RNA ヘリカーゼ

ウイルスの複製転写複合体は、多くの場合 RNA 合成に必須の RNA ヘリカーゼを含ん でいる。コロナウイルスの nsp13 タンパク質は、スーパーファミリー1B (SF1B)に属する NTP 依存性ヘリカーゼであり、RNA ないし DNA を 5' → 3'の極性をもってほどく活性を示す (24-26)。ヘリカーゼ活性の他に、RNA 5'-triphosphatase 活性を有しており、mRNA のキャッ ピングにも関与している可能性が示唆されている(27, 28)。最近、nsp13 を結合した RdRp 複 合体のクライオ電顕構造が報告され、RdRp コア複合体 (nsp12-nsp82-nsp7) に対して 2 分子 の nsp13 が結合することが明らかになった(23, 29) (図 4)。 nsp13 の RNA ヘリカーゼの極性 (5' → 3') は RdRp の転写の方向性(3' → 5')とは反対であるため、nsp13 は単純に RNA 合成 の processivity に寄与するだけでなく、上述の nsp14 (ExoN)による転写エラーの校正および ニドウイルスに特有の分節化 mRNA を合成するためのテンプレートスイッチングに役立っ ている可能性が示唆された。



図 4. RNA ヘリカーゼを結合した SARS-CoV-2 RdRp の構造

A, B. nsp13 サブユニット (RNA ヘリカーゼ)を結合した SARS-CoV-2 RdRp の構造 (PDB: 6XEZ を基に作図)。

6. RNA の 5 [´] 末端のキャッピング

コロナウイルス RNA は、その 5 末端にキャップ構造を有する(30, 31)。5'キャップ構造 は真核生物の mRNA の特徴であるが、ウイルスはウイルス RNA の 5'末端にキャップ構造 をもつことで mRNA を模倣していると言える。これは RNA の安定性、宿主の免疫の回避、 およびウイルスタンパク質の翻訳に重要である。ピコルナウイルスや C 型肝炎ウイルスの ように、ウイルス RNA の翻訳に Internal ribosome entry site (IRES)を使用することでキャッ プの問題を回避しているウイルスとは異なり、コロナウイルスは他の多くのウイルスと同 様に RNA をキャップする機構を獲得し、RNA は真核生物の mRNA と同様にキャップ依存 的に翻訳される。



図 5. RNA キャッピング

A. nsp12のNiRAN ドメインのGTase 活性部位(PDB: 7CYQ を基に作図)。B. nsp14のN7-MTase 活性部位(PDB: 5C8S を基に作図)。C. SARS-CoV-2 nsp16-nsp10 複合体の構造(PDB: 7JYY を基に作図)。D. nsp16の2'-O-MTase 活性部位の構造。

このキャップ構造は、4 段階の連続した反応で生成されると考えられている(*32*)。①まず、 RNA 5'-triphosphatase (RTPase)が、新生 RNA 鎖の 5'末端の三リン酸基 (5'-pppN) からリン 酸基を1つ取り除いて 5'-ppN を生成する。②次に、guanylyltransferase (GTase)が GTP を加 水分解して GMP を 5'-ppN に転移し、キャップコア構造を生成する (GpppN)。③次いで、 N7-MTase がグアニンの N7 位をメチル化し、cap-0 構造 (m⁷GpppN) を生成する。④最後に、 2'-O-MTase が一番目と二番目のヌクレオチドのリボースの 2'-OH をメチル化し、それぞれ cap-1, cap-2 構造を生成する。

①の RTPase 反応は nsp13 が担うと考えられているが、詳細は不明である。②の GTase については、最近 nsp12 の NiRAN ドメインに GTP 結合部位が同定され、GTase 活性部位が あることが示唆された(33)(図 5A)。③については、nsp14 の N7-MTase ドメインがその役割 を担う(図 5B)。N7-MTase ドメインは非典型的な MTase fold をもったドメインであるが、メチル基のドナーとなる SAM (S-adenosyl-L-methionine)と受容体の GpppN が結合するポケ ットを有しており、SAM のメチル基をグアニンの N7 近傍に固定して、インライン機構で メチル転移を実現していると考えられる(22)。④の 2'-O-MTase の活性部位は nsp16 にあり、m⁷GpppA-RNA に特異的にキャップ付加を行って cap-1 構造を生成する(図 5C,D)。nsp16 タ

ンパク質は nsp10 と安定な複合体を形成する。nsp10 は、nsp16 の 2'-O-MTase 活性に必須で あり、nsp16 の SAM 結合ポケットを安定化し、受容体となる RNA 鎖の結合溝を拡張するこ とで、基質結合に寄与し、nsp16 を活性化していると推定されている(*34-36*)。

7. 終わりに

現在我々は、それ以前はまったく予想もしなかったウイルス感染症のパンデミックの 最中にあり、いわゆる疫病が前時代的なものでも対岸の火事でもないことを痛感している。 緊急承認されたレムデシビルは幅広いスペクトルを持ったヌクレオチド型阻害剤である。 一方、WHOは大規模な治験の結果、レムデシビルが COVID-19 に効果がなかったとしてい る。上述したように、コロナウイルスの複製・増殖には多くのウイルスタンパク質や宿主因 子の働きが不可欠であり、それらはウイルスを制御するためのターゲットとなりうる。 COVID-19 のパンデミックは、皮肉にも SARS-CoV-2 に関する研究を急速に促進させている が、その中から第二第三の特効薬が現れることを期待したい。また一方で、長期的な視野に 立ち、コロナウイルスに限らず病気の原因となりうるウイルス等の感染・増殖の仕組みを理 解し、それらをコントロールする方法をあらかじめ準備し、今後に備えていくことが肝要で あろう。

参考文献

- J. Cui, F. Li, Z. Shi, Origin and evolution of pathogenic coronaviruses. *Nature reviews*. *Microbiology* 17, 181–192 (2019).
- J. Ziebuhr, E. Snijder, A. Gorbalenya, Virus-encoded proteinases and proteolytic processing in the Nidovirales. *The Journal of general virology* 81, 853-879 (2000).
- Y. Báez-Santos, S. St John, A. Mesecar, The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antiviral research* 115, 21-38 (2015).
- 4. L. Subissi *et al.*, One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 3900-3909 (2014).
- R. Kirchdoerfer, A. Ward, Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nature communications* 10, (2019).
- Q. Wang *et al.*, Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. *Cell* 182, 417-428 (2020).
- 7. H. Hillen *et al.*, Structure of replicating SARS-CoV-2 polymerase. *Nature* 584, 154–156 (2020).
- 8. Y. Gao *et al.*, Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* (*New York, N.Y.*) **368**, 779-782 (2020).
- W. Yin *et al.*, Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science (New York, N.Y.)* 368, 1499-1504 (2020).

- R. Eastman *et al.*, Remdesivir: A Review of Its Discovery and Development Leading to Emergency Use Authorization for Treatment of COVID-19. *ACS central science* 6, 672–683 (2020).
- C. Gordon, E. Tchesnokov, J. Feng, D. Porter, M. Götte, The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *The Journal of biological chemistry* 295, 4773-4779 (2020).
- 12. G. Kokic et al., Mechanism of SARS-CoV-2 polymerase inhibition by remdesivir. bioRxiv, (2020).
- E. Snijder *et al.*, Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *Journal of molecular biology* 331, 991-1004 (2003).
- A. Gorbalenya, L. Enjuanes, J. Ziebuhr, E. Snijder, Nidovirales: evolving the largest RNA virus genome. *Virus research* 117, 17-37 (2006).
- E. Minskaia *et al.*, Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 103, 5108-5113 (2006).
- 16. L. Eckerle, X. Lu, S. Sperry, L. Choi, M. Denison, High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *Journal of virology* **81**, (2007).
- 17. L. Eckerle *et al.*, Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. *PLoS pathogens* **6**, (2010).
- E. Smith, H. Blanc, M. Surdel, M. Vignuzzi, M. Denison, Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *PLoS pathogens* 9, (2013).
- 19. M. Agostini *et al.*, Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. *mBio* **9**, (2018).
- 20. A. Shannon *et al.*, Remdesivir and SARS-CoV-2: Structural requirements at both nsp12 RdRp and nsp14 Exonuclease active-sites. *Antiviral research* **178**, (2020).
- Y. Ma *et al.*, Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. *Proceedings of the National Academy of Sciences of the United States of America* 112, 9436-9441 (2015).
- 22. M. Bouvet *et al.*, RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 9372-9377 (2012).
- J. Chen *et al.*, Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. *Cell* 182, 1560-1573 (2020).
- 24. N. Lee *et al.*, Cooperative translocation enhances the unwinding of duplex DNA by SARS coronavirus helicase nsP13. *Nucleic acids research* **38**, 7626-7636 (2010).

- 25. J. Tanner *et al.*, The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5' to 3' viral helicases. *The Journal of biological chemistry* **278**, 39578-39582 (2003).
- 26. A. Seybert, A. Hegyi, S. Siddell, J. Ziebuhr, The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity. *RNA (New York, N.Y.)* **6**, 1056-1068 (2000).
- K. Ivanov, J. Ziebuhr, Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. *Journal of virology* 78, 7833–7838 (2004).
- K. Ivanov *et al.*, Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. *Journal of virology* 78, 5619-5632 (2004).
- 29. L. Yan *et al.*, Architecture of a SARS-CoV-2 mini replication and transcription complex. *Nature communications* **11**, (2020).
- S. Nallagatla, R. Toroney, P. Bevilacqua, A brilliant disguise for self RNA: 5'-end and internal modifications of primary transcripts suppress elements of innate immunity. *RNA biology* 5, 140-144 (2008).
- E. Decroly, F. Ferron, J. Lescar, B. Canard, Conventional and unconventional mechanisms for capping viral mRNA. *Nature reviews. Microbiology* 10, 51–65 (2011).
- 32. Y. Furuichi, A. Shatkin, Viral and cellular mRNA capping: past and prospects. *Advances in virus research* **55**, 135-184 (2000).
- L. Yan *et al.*, Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription Complex Reveals an Intermediate State in Cap Synthesis. *Cell*, 184-193 (2020).
- E. Decroly *et al.*, Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. *PLoS pathogens* 7, (2011).
- 35. Y. Chen *et al.*, Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *PLoS pathogens* 7, (2011).
- 36. M. Rosas-Lemus *et al.*, High-resolution structures of the SARS-CoV-2 2'- O-methyltransferase reveal strategies for structure-based inhibitor design. *Science signaling* **13**, (2020).

///// Cutting Edge /////

機械学習とバーチャルスクリーニングを通じた SARS-CoV-2 RNA 依存性 RNA ポリメラーゼの阻害剤予測

株式会社 Elix Nazim Medzhidov、結城 伸哉

1. はじめに

重症急性呼吸器症候群コロナウイルス2(SARS-CoV-2)の世界的な拡散により、COVID-19パンデミックが発生した。SARS-CoV-2は、一本鎖プラス鎖RNAゲノム[(+)ssRNA]を有 し、ニドウイルス目コロナウイルス科ベータコロナウイルス属に属する[1]。COVID-19は、 非定型肺炎を特徴とする致死性の呼吸器疾患である[2]。COVID-19治療薬は現時点で存在せ ず、計算機を使用した解析をはじめ、様々な手法でドラッグリポジショニングの取り組みが 行われてきた[3-5]。本研究では、複数の機械学習モデルを使用し、それらを分子ドッキング アプローチと組み合わせ、ウイルス複製機構の重要な構成要素であるSARS-CoV-2 RNA依存 性RNAポリメラーゼ(RdRp)に対し阻害活性を有しうる抗ウイルス薬および抗炎症薬のス クリーニングを行った。

2. コロナウイルスと SARS-Cov-2

2.1 概要

コロナウイルスは、様々な動物に感染する多様なウイルス群であり、ヒトでは軽度から 重度の呼吸器感染症を引き起こす可能性がある。人獣共通感染源の高病原性コロナウイル スである重症急性呼吸器症候群コロナウイルス(SARS-CoV)と中東呼吸器症候群コロナウ イルス(MERS-CoV)は、2002年と2012年にそれぞれヒトに感染し、致死的な呼吸器疾患 を引き起こした。こうしてコロナウイルスは、21世紀の公衆衛生上の懸念事項として認識 されるようになった[6]。2019年末、高い感染力を有する SARS-CoV-2が中国の武漢市で報 道され、その後世界でもウイルス性肺炎の感染が広がりを見せた。COVID-19と呼ばれるこ の病気は、感染者数と流行地域数の両方で過去に発生した SARS や MERS を上回っており、 世界の公衆衛生に並々ならぬ脅威をもたらしている。

2.2 SARS-CoV-2 の起源と拡散

2019年12月下旬、中国・武漢の複数の保健施設から原因不明の肺炎患者のクラスター が報告された[7]。患者の多くは、発熱、咳、胸部不快感などのウイルス性肺炎の症状を呈 し、重症の場合呼吸困難や両側肺浸潤など、SARSやMERSの患者と同様の症状が見られた [7,8]。患者の多くは、武漢市の中心部にある華南海鮮卸売市場と何かしらの関連を持って いたとされる[9,10]。12月31日、武漢市衛生委員会は原因不明の肺炎発生を世界保健機関 (WHO)に通知した[11]。

2.3 Genomic Characterization of SARS-CoV-2 / SARS-CoV-2 の遺伝的特徴

SARS-CoV-2 は SARS-CoV と比して 79%、MERS-CoV と比して 50%のゲノム配列同一 性を有する[12]。SARS-CoV-2 は他のベータコロナウイルスと共通のゲノム構造を持つ。 SARS-CoV-2 のゲノム構造は他のベータコロナウイルスと共通しており、レプリカーゼ (ORF1a/ORF1b)、スパイク(S)、エンベロープ(E)、メンブレン(M)、ヌクレオカプシド (N) の 6 つのオープンリーディングフレーム (ORF1a/ORF1b) が 5′から 3′の順に配列 されている。また、遺伝子には、付属タンパク質をコードする 7 つの ORF が散在している [1]。SARS-CoV-2 がコードする多くのタンパク質の長さは、SARS-CoV の対応するタンパク 質と近い。4 つの構造遺伝子(上記 S、E、M、N)のうち、SARS-CoV-2 は S 遺伝子を除い て SARS-CoV と 90%以上のアミノ酸同一性を共有している[12]。レプリカーゼ遺伝子は 5′ ゲノムの 3 分の 2 をカバーし、大きなポリタンパク質(pp1ab)をコードしており、ウイル ス複製に関与する 16 の非構造タンパク質をコードする。SARS-CoV-2 非構造タンパク質の ほとんどは、SARS-CoV の構成アミノ酸と 85%以上同じアミノ酸で構成されている[1]。

2.4 SARS-CoV-2 RdRp を対象とした理由

RdRps は、RNA テンプレートから RNA 合成を触媒するマルチドメインタンパク質で、 ウイルスのゲノム複製と転写プロセスを担う[13]。コロナウイルス属が属するニドウイルス 目は、ウイルスの複製と転写を促進すべく ORF1a および ORF1b ウイルスポリプロテインの 開裂産物として産生される非構造タンパク質 (NSP) によって作動する RNA 合成専用の複 雑な機構を持つという特徴を有する[14]。

ウイルスの宿主細胞への侵入は、感染開始の重要な過程である。SARS-CoV-2 は、ウイル ススパイク(S)タンパク質の受容体結合ドメインを利用して宿主細胞上のアンジオテンシ ン変換酵素 2 (ACE2)に結合し、エンドソームを介して細胞内に侵入する[15]。細胞内侵入 の後、ウイルスゲノム RNA が放出され、ウイルスタンパク質の翻訳やウイルスゲノム複 製のための鋳型として機能する。RNA 依存性 RNA ポリメラーゼ (RdRp)は、非構造タン パク質 12 (nsp12)としても知られており、新しいウイルスのゲノム RNA の産生を担うウ イルスレプリカーゼ複合体の重要な構成要素である[16]。

ほとんどすべての RNA ウイルスは、そのゲノムにコードされた RdRp を有する。配列の 類似性が比較的低いにもかかわらず、さまざまな RNA ウイルス由来の RdRp は、「右手 モデル」として説明される 3 つの構成ドメイン(「親指ドメイン」、「手のひらドメイン」、

「指ドメイン」)と同様の構造的類似性を有しており、これは逆転写酵素とも類似している。種々のウイルスから得られた RdRp の配列解析から、活性部位と手のひらドメインに重要な残基が保存されていることが明らかになっている[13,17-19]。また、SARS-CoV-2 RdRpの構造解析の結果、RdRpの一般的な構造、保存された A-G モチーフの存在、主要な触媒アミノ酸の保存、C型肝炎ウイルス(HCV)やポリオウイルス RdRp との構造的な類似性が確認された[20]。RdRp はウイルスの複製サイクルにおいて重要な役割を果たしつつ人体に対応するものが存在しないため、治療中に望ましくない副作用が生じるリスクを軽減できる

重要な治療標的と見なされている。

3. ドラッグリポジショニングにおける人工知能の利用

3.1 分子ドッキングと機械学習によるアプローチの比較

SARS-CoV-2 に有効な治療薬の開発が急務となっており、これまで多くの研究が行われ てきた。最近では、SARS-CoV-2 の異なるタンパク質を標的とした様々なドラッグリパーパ シングの候補を提案する in silico 創薬報告が数多く発表されている[5,21-24]。これらの研究 のほとんどは従来の分子ドッキング解析を利用しており、標的やリガンドの立体構造に関 する情報を必要とする。この情報が得られれば、標的タンパク質上の指定された領域に対し てドッキングシミュレーションを行い、タンパク質とリガンドの結合エネルギーを予測す ることができる。しかし、標的タンパク質の立体構造取得は困難で、費用も時間もかかる。 さらに、結合エネルギーが良好であても、リガンドによる阻害活性が高いとは限らない。

3.2 SARS-CoV-2 ドラッグリパーパシングにおける人工知能の利用

創薬における機械学習活用の技術開発は、従来の創薬プロセスを短縮し、新薬創出にかかるコストを削減できる可能性がある[25,26]。最近では、立体構造に関する情報が得られない場合でも比較的短期間で潜在的な薬剤候補を見つけることができる代替手段として、機械学習を用いたアプローチがいくつか検討されている。

英国に拠点を置く BenevolentAI 社は、種々の科学文献から得られる情報を統合した AI 由来のナレッジグラフを多く活用している[4]。同グループは、宿主タンパク質 AAK1 の阻 害を標的とし、関節リウマチの治療薬として承認されているバリシチニブを同定した。バリ シチニブは、クラスリンを介したエンドサイトーシスを阻害することで、細胞のウイルス感 染を抑制すると考えられている。 バリシチニブが有する抗炎症性は、COVID-19 患者に多く 見られるサイトカインレベルの上昇に有効である可能性が高い[27]。同様に、Beckらは、 Molecule Transformer-Drug Target Interaction (MT-DTI)と呼ばれる深層学習の技術を活用した 薬物-標的相互作用モデルを応用し、SARS-COV-2 関連のプロテアーゼとヘリカーゼを標的 とする可能性を有する市販の抗ウイルス薬を予測した論文を発表している[3]。このモデル は分子を 1 次元の文字列表す SMILES(Simplified molecular-input line-entry system)とアミノ 酸(AA)配列を使用しているため、実験で3次元結晶構造が確認されていない標的タンパ ク質を扱うことが可能である。また、米国に拠点を置く Atomwise 社は、新たな抗ウイルス 剤の開発を目指し、複数のコロナウイルスに高度に共通して保存されている SARS-CoV-2 タ ンパク質結合部位をターゲットにした研究に注力している。具体的には、AtomNetの深層畳 み込みニューラルネットワーク技術[28]を使い、数百万の仮想化合物をスクリーニングして いるほか、学術研究者と15のパートナーシップを結んで、モデルで予測された化合物を in vitro アッセイで確認している[29]。

複数のウイルスの RdRp において構造上の類似性および活性部位での主要アミノ酸の保

存が認められ、また、広範な抗 RdRp 薬であるレムデシビルが同定されたことで、有効な RdRp 阻害薬の化学構造に類似パターンが存在する可能性が示された。このことから、潜在 的な RdRp 阻害剤を同定するため、教師あり機械学習アルゴリズムを実装することとした。

4. Discovery of SARS-CoV-2 RdRp Inhibitors / SARS-CoV-2 RdRp 阻害剤の探索 4.1 データセット収集及びモデルの学習

まずは、HCV、デングウイルス、ポリオウイルス、インフルエンザウイルスの RdRp に 対し、過去の実験で既に活性値が確認されている低分子を含むデータセットを構築した。デ ータセットは PubChem [30]と ChEMBL [31]のバイオアッセイから取得した。分類モデルの 学習を目的とし、既に実験で活性値(IC50/EC50)が知られているエントリを選択し、活性値 と閾値から二値の活性ラベル(活性あり・なし)を割り当てた。学習に使用する活性のカッ トオフ閾値は5µMとした。最終的に構築されたデータセットは、活性ラベル付きの化合物 1,356(656 種類が不活性、700 種類が活性あり)が含まれている。バリデーションデータセ ットは、データセット全体から 20%の化合物(活性化合物と不活性化合物)を無作為に選 択したものを使用した。学習データセットは、残りの 80%の化合物を使用した。学習デー タセット中の分子は HCV、ポリオウイルス、デングウイルス、インフルエンザウイルスの RdRp に対して阻害活性を示した。これらの阻害剤を機械学習モデルに学習させることで、 機械学習のモデルに有効な RdRp 阻害剤の化学的特徴の学習が可能になると考えられる。次 に、当該モデルが既知の前臨床および臨床 RdRp 阻害剤を同定する能力を評価した。最後に、 RdRp に対する阻害活性を潜在的に有する物質の候補を特定するため、FDA 承認の抗ウイル ス薬および抗炎症薬を、当該モデルでスクリーニングした。また、SARS-CoV-2の RdRp タ ンパク質に対する抗ウイルス薬や抗炎症薬の分子ドッキング解析も併せて実施した。今回 の機械学習モデルでは、RdKit [32]を用いて化合物を分子フィンガープリントに変換したも のを入力特徴量として使用した。フィンガープリントは、Morgan フィンガープリント[33] とトポロジカルフィンガープリントを使用しており、トポロジカルフィンガープリント に関しては、結合数が最小1から最大7までの化合物のすべてのサブグラフを抽出して計 算した。モデルの実装は、主に scikit-learn ライブラリ[34]を使用した。

受信者操作時特性曲線下の面積(以下、ROC-AUC、又は AUROC とする) 0.8 を超えた モデルは複数ある。具体的には、グラフ畳み込みネットワーク、メッセージパッシングネッ トワーク、ランダムフォレスト分類器(円形フィンガープリント、トポロジカルフィンガー プリント)、リッジ分類器(円形フィンガープリント)、ラッソ分類器(トポロジカルフィン ガープリント)、3 層マルチレイヤーパーセプトロン(円形フィンガープリント)、XGBoost 分類器(トポロジカルフィンガープリント)が、0.8 を超えていた。そのうちの1つである 円形フィンガープリントのランダムフォレスト分類器は ROC-AUC 値が 0.9 を超えていた。 精度(ACC)は、ランダムフォレスト分類器が 84%と最も良い数値を示した。

4.2 前臨床 RdRp 阻害剤のテストセットを用いたモデル評価

結果検証には、ROC-AUCに基づく上位3つのモデルを使用し、既知の前臨床 RdRp 阻害 剤のテストセットを対象とした予測を行った。ROC-AUCと精度に加え、真陽性(TP)、真 陰性(TN)、偽陽性(FP)、偽陰性(FN)の割合も表1.1に示す。

モデル	AUROC	ACC	Confidence Interval (alpha=0.05)	ТР	TN	FP	FN
GraphConv	0.700	0.700	[0.558, 0.842]	0.65	0.75	0.25	0.35
RandomForest (C)	0.725	0.725	[0.587, 0.863]	0.50	0.95	0.05	0.50
3-layer MLP (C)	0.625	0.625	[0.475, 0.775]	0.50	0.75	0.25	0.50

表 1.1 テストデータに対するモデルのパフォーマンス

ランダムフォレスト分類器は陰性例の検出に非常に優れていたが(真陰性率95%)、陽性 例について見てみると、検出可能な活性分子は半数に留まった(真陽性率50%)。グラフ畳 み込みモデル等の他のモデルは、より多くの活性分子を検出できるが(真陽性率は65%)、 真陰性率は75%に低下し、より多くの偽陽性を検出してしまう。更に、異なるモデルの出 力間の相関はあまり高くない。これらの事情を勘案し、アンサンブルモデルを使用すること とした。

今回の実験では、RBF カーネルを用いた単純な Support Vector Machine が最も良く機能した。このモデルは、最も良い結果を示した上位 10 個のモデルの出力値を入力特徴量として使用している。最初に、元の学習セットを用いて個々のモデルの学習を行った。次に、バリ デーションセットを 2 つの同じサイズのサブセットに分割した。そのうちの 1 つをアンサ ンブルモデルの学習に使用し、もう 1 つは検証セットに充てた。アンサンブルモデルでは、 テストセットを用いた他のすべてのモデルの性能をわずかに上回った(表 1.2)。

Dataset	AUROC	ACC	Confidence Interval (alpha=0.05)	ТР	TN	FP	FN
Validation	0.875	0.875	[0.819, 0.931]	0.871	0.879	0.129	0.121
Test	0.750	0.750	[0.616, 0.884]	0.600	0.900	0.100	0.400

表 1.2 アンサンブルモデルの結果

モデルの評価は、バリデーションセットと既知の前臨床 RdRp 阻害剤のテストセットのパフォーマンスに基づき実施した。最も良い性能を示した3つのモデルと(表1.1参照)、アンサンブルモデルの予測結果を分析した。モデルの学習はリガンドの情報のみを使用し、標的である SARS-CoV-2 RdRp の立体構造は使用していない。そのため、標的タンパク質とリガンドの立体構造情報に基づく従来の分子ドッキングアプローチと比較することが望ましいと考えられる。そこで、SARS-CoV-2 RdRp(PDB ID: 6m71)の活性部位に対する抗ウイルス・抗炎症データセットのバーチャルスクリーニングを AutoDock Vina を用いて行った[35]。

5. 機械学習とドッキングシミュレーションの結果

5.1 抗ウイルスデータセットの結果

解析の結果、抗ウイルス薬と抗炎症薬の両方のデータセットから標的候補を複数同定した。本章では具体的に紹介する。

我々のモデルは抗ウイルスデータセットからレムデシビルを同定した。レムデシビルは SARS-CoV-2 RdRp を標的とすることが確認されたヌクレオシドアナログであり、COVID-19 患者の治療薬として米国 FDA によって承認されている。レムデシビルは陽性対照としてテ ストセットに含まれていた。なお、上述のモデルで同定されたバロキサビル・マルボキシル、 TMC-310911 (ASC09)、およびユミフェノビル (アルビドール) は、COVID-19 を対象とし た臨床試験でも研究が進んでいる。臨床試験登録番号と識別番号は、バロキサビル・マルボ キシルが ChiCTR2000029544、ASC09 が NCT04261907、ユミフェノビルが NCT04350684 で ある。バロキサビル・マルボキシルはインフルエンザウイルスの RdRp に作用し[36]、TMC-310911 は HIV-1 に対して開発されたプロテアーゼ阻害剤であり[37]、ユミフェノビルはヘマ グルチニン (HA) 糖タンパク質を標的にしてウイルスの細胞内への侵入を抑制する抗イン フルエンザ薬[38]である。

これらの薬剤候補に加え、最も良い性能を示した 4 つのモデル (表 1.1 の 3 つのモデル 及びアンサンブルモデル、以下「最良モデル」とする)が、RdRp 阻害剤の可能性として、 抗ウイルスデータセットからベクラブビルとアスナプレビルを同定した。ベクラブビルは、 HCV RdRp (NS5B)の非ヌクレオシド系阻害剤である[39]。SARS-CoV-2 と同様、HCV も一 本鎖エンベロープ型ポジティブセンス RNA ウイルスである。SARS-CoV-2 と HCV RdRp の 活性部位は構造的類似性を示しており、両者とも触媒部位の保存アミノ酸を保有している [20]。ベクラブビルの SARS-CoV-2 RdRp に対する結合エネルギーは-9.2 kcal/mol となり、本 候補の阻害性を示唆している。別の抗 HCV 薬であるアスナプレビルは、HCV のプロテアー ゼを標的とすることが知られている。全ての最良モデルが潜在的な抗 RdRp 候補としてアス ナプレビルを同定し、AutoDock Vina を用いた結合エネルギーは、-7.5 kcal/mol であった。 他の候補としては、パリタプレビル、ファルダプレビル、シメプレビル、ベドロプレビル

(HCV プロテアーゼ阻害剤)、レディパスビル、オダラスビル、およびベルパタスビル (HCV NS5A 阻害剤) が挙げられる (表 1.3 参照、4 つの最良モデルのうち少なくともいずれか 2 つで予測され、SARS-CoV-2 RdRp に対する結合エネルギーも比較的低いため)。我々のモデルは RdRp 阻害剤のみを対象としていたが、HCV のプロテアーゼまたは NS5A タンパク質を標的とするいくつかの抗 HCV 薬も RdRp 阻害剤として分類された。なお、これらの候補分子は、分子ドッキング解析に基づく SARS-CoV-2 の RdRp に対する結合エネルギーが良好であると予測された。これらの分子が実際に RdRp に作用するかどうか確認するためには、実験を通じて更なる検証が必要である。

物質名	対象物質を予測した 予測モデルの数	SARS-CoV-2 RdRp との 結合エネルギー (kcal/mol)
Beclabuvir	4	-9.2
Asunaprevir	4	-7.5
Paritaprevir	3	-10.5
Faldaprevir	3	-9.6
Odalasvir	3	-8.8
Simeprevir	3	-8.7
Vedroprevir	3	-8.6
Velpatasvir	3	-8.6
Telaprevir	3	-8.3
Dolutegravir	3	-8.0
Sofosbuvir	3	-6.9
Uprifosbuvir	3	-6.8
Entecavir	3	-6.6
Lobucavir	3	-6.6
Trifluridine	3	-6.3
Nevirapine	3	-6.1
Ledipasvir	2	-9.2
Ruzasvir	2	-8.1
Baloxavir marboxil	2	-8.0
TMC-310911(ASC09)	2	-7.9
Adafosbuvir	2	-7.8
Remdesivir	2	-7.5
Saquinavir	2	-7.2
Abacavir	2	-7.1
Maribavir	2	-7.1
Elvitegravir	2	-6.6
Vidarabine	2	-6.5
Efavirenz	2	-6.3
Valganciclovir	2	-6.2
Valomaciclovir	2	-6.2
Sorivudine	2	-6.1
Ibacitabine	2	-6.1
Idoxuridine	2	-5.9
Fialuridine	2	-5.9
Didanosine	2	-5.8
Umifenovir	2	-5.8

表 1.3 RdRp に作用する可能性のある抗ウイルス薬一覧 及び結合エネルギー(計算は AutoDock Vina を使用)

5.2 抗炎症性データセットの結果

抗ウイルス薬は、患者の負荷軽減には役立つが、ウイルス誘発性肺炎そのものに直接対 処するものではない。新型コロナウイルスによる肺炎は、SARS-CoV-2 が肺で引き起こす炎 症の結果生じるものである[40]。したがって、肺炎を発症した COVID-19 の患者には、肺の 炎症を抑えるための追加治療が必要となる可能性がある。上述のモデルは数多の抗ウイル ス薬から RdRp 阻害剤を同定する能力を有するだけでなく、抗炎症薬に対しても同様の効果 を発揮しうることから、SARS-CoV-2 RdRp に対する RdRp 阻害剤の特性と結合エネルギー の予測に焦点を当て、解析を行った。抗炎症性データセットの分析の結果、全ての最良モデ ルが、天然物質であるベツリン酸とルピオールが抗 RdRp 活性を有していると予測すること が明らかになった。他にも、リフィテグラスト、アントラフェニン、ウルソール酸、酢酸デ キサメタゾン、リン酸プレドニゾロンは、最良モデルのうち少なくとも 2 つのモデルによっ て予測されており、また、-7.5 から-9.5 kcal/mol の範囲の結合エネルギーで SARS-CoV-2 RdRp の活性部位に結合することが予測された(表 1.4)。なお、ベツリン酸とウルソール酸は、 HIV に対する抗ウイルス活性が報告されている五環式トリテルペノイドである[41]。

物質名	対象物質を予測した 予測モデルの数	SARS-CoV-2 RdRp との 結合エネルギー (kcal/mol)		
Betulinic Acid	4	-7.4		
Lupeol	4	-7.2		
Lifitegrast	3	-9.5		
Antrafenine	3	-8.7		
Ursolic acid	3	-8.0		
Floctafenine	3	-7.1		
Cimicoxib	3	-7.0		
Acemetacin	3	-6.8		
Morniflumate	3	-6.8		
Loteprednol	3	-6.8		
Polmacoxib	3	-6.8		
Andrographolide	3	-6.7		
Dexamethasone acetate	2	-7.6		
Prednisolone phosphate	2	-7.5		
Cortisone acetate	2	-7.3		
Mometasone furoate	2	-7.3		
Prednicarbate	2	-7.1		
Deflazacort	2	-7.1		
Clobetasone	2	-6.8		
Rimexolone	2	-6.8		
Robenacoxib	2	-6.8		
Hydrocortisone probutate	2	-6.8		
Mometasone	2	-6.6		
Diflunisal	2	-6.5		
Lumiracoxib	2	-6.5		
Etoricoxib	2	-6.5		
Clobetasol	2	-6.5		
Apremilast	2	-6.5		
Bisindolylmaleimide I	2	-6.5		
Talniflumate	2	-6.3		
NS-398	2	-6.2		
Firocoxib	2	-5.6		
Dimethyl sulfone	2	-3.0		

表 1.4 RdRp に作用する可能性のある抗炎症性薬一覧 及び結合エネルギー (計算は AutoDock Vina を使用)

6. 結論

COVID-19の世界的パンデミックの社会的・経済的影響は、世界中の数多くの人々の生活 に深刻な影を落としている。多様な生態系において、人命を奪う可能性のあるウイルスを保 有する動物が存在し、また、グローバルに人々が移動する現代社会においては、病原体の世 界的な伝播が容易になっていることから、将来、新たなパンデミックの出現によってより脆 弱な環境にさらされ得る。有効な治療薬の発見のために多大な努力がなされているが、これ までのところ、COVID-19の患者の治療に役立つ薬剤特定は未だ発展途上であり、多くの死 亡者が報告され続けている。これは、創薬のプロセス全体を迅速化できるよう、現在の創薬 の方法を慎重に、しかし新しい方法で見直していく必要があることを示唆している。本論文 では、機械学習アルゴリズムに基づくスクリーニングを通じ、SARS-CoV-2 RdRp 阻害剤の 候補を予測した。今後は、当該分子が SARS-CoV-2 RdRp に対し活性を有するかどうか、モ デルだけでなく実際の実験を通じて検証することが望ましいと考える。

7. 謝辞

本論文の作成に当たり、多くの人々のご協力を頂きました。特に、丁寧な原稿の修正と翻 訳に協力して頂いた加藤嵩侑氏、伊東美絵氏(ともに株式会社 Elix)に感謝の意を表します。 今回のパンデミックは編集現時点でも、多くの人々に多大なる影響を及ぼしています。人々 の安全で健康な生活のために、命がけで COVID-19 の患者のケアをして下さっている医療 従事者の皆様にも感謝の意を表し、謝辞とさせて頂きます。

参考文献

- Chan, J. F.-W. et al.: Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging Microbes & Infections*, 9, 221–236 (2020). http://doi:10.1080/22221751.2020.1719902.
- [2] Zhou, P. et al.: A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579, 270–273 (2020). http://doi:10.1038/s41586-020-2012-7.
- [3] Beck, B. R. et al.: Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. Computational and Structural Biotechnology Journal, 18, 784–790 (2020). http://doi:10.1016/j.csbj.2020.03.025.
- [4] Richardson, P. et al.: Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. The Lancet, 395, e30–e31 (2020). http://doi:10.1016/s0140-6736(20)30304-4.
- [5] Kandeel, M.; Al-Nazawi, M.: Virtual screening and repurposing of FDA approved drugs against COVID-19 main protease. Life Sciences, 251, 117627 (2020). http://doi:10.1016/j.lfs.2020.117627.

- [6] Cui, J. et al.: Origin and evolution of pathogenic coronaviruses. *Nature Reviews Microbiology*, 17, 181–192 (2019). http://doi:10.1038/s41579-018-0118-9.
- Zhu, N. et al.: A novel coronavirus from patients with pneumonia in China, 2019. New England Journal of Medicine, 382, 727–733 (2020). http://doi:10.1056/nejmoa2001017.
- [8] Gralinski, L. E.; Menachery, V. D.: Return of the coronavirus: 2019-nCoV. *Viruses*, 12, 135 (2020). http://doi:10.3390/v12020135.
- [9] Deng, S.-Q.; Peng, H.-J.: Characteristics of and public health responses to the coronavirus disease 2019 outbreak in China. *Journal of Clinical Medicine*, 9, 575 (2020). http://doi:10.3390/jcm9020575.
- [10] Jiang, S. et al.: An emerging coronavirus causing pneumonia outbreak in Wuhan, China: calling for developing therapeutic and prophylactic strategies. *Emerging Microbes & Infections*, 9, 275–277 (2020). http://doi:10.1080/22221751.2020.1723441.
- [11] Wu, Z.; McGoogan, J. M.: Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China. *JAMA*, **323**, 1239–1242 (2020). http://doi:10.1001/jama.2020.2648.
- [12] Lu, R. et al.: Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, **395**, 565–574 (2020). http://doi:10.1016/s0140-6736(20)30251-8.
- [13] Venkataraman, S. et al.: RNA dependent RNA polymerases: Insights from structure, function and evolution. *Viruses*, **10**, 76 (2018). http://doi:10.3390/v10020076.
- [14] Snijder, E. J. et al.: The nonstructural proteins directing coronavirus RNA synthesis and processing. Advances in Virus Research, 96, 59–126 (2016). http://doi:10.1016/bs.aivir.2016.08.008.
- [15] Lan, J. et al.: Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, 581, 215–220 (2020). http://doi:10.1038/s41586-020-2180-5.
- [16] Masters, P. S.: The Molecular biology of coronaviruses. *Advances in Virus Research*, 66, 193–292 (2006). http://doi:10.1016/s0065-3527(06)66005-3.
- [17] Bruenn, J. A.: A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. *Nucleic Acids Research*, **31**, 1821–1829 (2003). http://doi:10.1093/nar/gkg277.
- [18] Xu, X. et al.: Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. *Nucleic Acids Research*, **31**, 7117–7130 (2003). http://doi:10.1093/nar/gkg916.
- [19] Černý, J. et al.: Evolution of tertiary structure of viral RNA dependent polymerases. *PLoS ONE*, 9, e96070 (2014). http://doi:10.1371/journal.pone.0096070.
- [20] Gao, Y. et al.: Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*, 368, 779–782 (2020). http://doi:10.1126/science.abb7498.

- [21] Hage-Melim, L. I. da S. et al.: Virtual screening, ADME/Tox predictions and the drug repurposing concept for future use of old drugs against the COVID-19. Life Sciences, 256, 117963 (2020). http://doi:10.1016/j.lfs.2020.117963.
- [22] Quimque, M. T. J. et al.: Virtual screening-driven drug discovery of SARS-CoV2 enzyme inhibitors targeting viral attachment, replication, post-translational modification and host immunity evasion infection mechanisms. Journal of Biomolecular Structure and Dynamics, , 1–23 (2020). http://doi:10.1080/07391102.2020.1776639.
- [23] Naik, B. et al.: High throughput virtual screening reveals SARS-CoV-2 multi-target binding natural compounds to lead instant therapy for COVID-19 treatment. International Journal of Biological Macromolecules, 160, 1–17 (2020). http://doi:10.1016/j.ijbiomac.2020.05.184.
- [24] Jiménez-Alberto, A. et al.: Virtual screening of approved drugs as potential SARS-CoV-2 main protease inhibitors. Computational Biology and Chemistry, 88, 107325 (2020). http://doi:10.1016/j.compbiolchem.2020.107325.
- [25] Zhavoronkov, A. et al.: Deep learning enables rapid identification of potent DDR1 kinase inhibitors.
 Nature Biotechnology, **37**, 1038–1040 (2019). http://doi:10.1038/s41587-019-0224-x.
- [26] Fleming, N.: How artificial intelligence is changing drug discovery. *Nature*, 557, 855–857 (2018). http://doi:10.1038/d41586-018-05267-x.
- [27] Stebbing, J. et al.: COVID-19: combining antiviral and anti-inflammatory treatments. *The Lancet Infectious Diseases*, 20, 400–402 (2020). http://doi:10.1016/s1473-3099(20)30132-8.
- [28] Wallach, I. et al.: AtomNet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery. *arXiv*, (2015).
- [29] https://www.businesswire.com/news/home/20200521005238/en/Atomwise-Partners-with-Global-Research-Teams-to-Pursue-Broad-Spectrum-Treatments-Against-COVID-19-and-Future-Coronavirus-Outbreaks
- [30] Kim, S. et al.: PubChem 2019 update: improved access to chemical data. *Nucleic Acids Research*, 47, D1102–D1109 (2018). http://doi:10.1093/nar/gky1033.
- [31] Gaulton, A. et al.: ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, 40, D1100–D1107 (2012). http://doi:10.1093/nar/gkr777.
- [32] Landrum, G.: RDKit: Open-Source Cheminformatics Software. null, (2016).
- [33] Rogers, D.; Hahn, M.: Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, 50, 742–754 (2010). http://doi:10.1021/ci100050t.
- [34] Pedregosa, F. et al.: Scikit-Learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830 (2011).
- [35] Trott, O.; Olson, A. J.: AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31, 455–461 (2010). http://doi:10.1002/jcc.21334.

- [36] Giacchello, I. et al.: Insights into RNA-dependent RNA polymerase inhibitors as anti-influenza virus agents. *Current Medicinal Chemistry*, 27 (2020). http://doi:10.2174/0929867327666200114115632.
- [37] Dierynck, I. et al.: TMC310911, A novel human immunodeficiency virus type 1 protease inhibitor, shows in vitro an improved resistance profile and higher genetic barrier to resistance compared with current protease inhibitors. *Antimicrobial Agents and Chemotherapy*, 55, 5723–5731 (2011). http://doi:10.1128/aac.00748-11.
- [38] Gasparini, R. et al.: Compounds with anti-influenza activity: present and future of strategies for the optimal treatment and management of influenza. Part I: Influenza life-cycle and currently available drugs. *Journal of preventive medicine and hygiene*, 55, 69–85 (2014).
- [39] Gentles, R. G.: Discovery of Beclabuvir: A potent allosteric inhibitor of the hepatitis C virus polymerase. *Topics in Medicinal Chemistry*, 31, 193–228 (2019). http://doi:10.1007/7355_2018_38.
- [40] Tay, M. Z. et al.: The trinity of COVID-19: immunity, inflammation and intervention. *Nature Reviews Immunology*, 20, 363–374 (2020). http://doi:10.1038/s41577-020-0311-8.
- [41] Xiao, S. et al.: Recent progress in the antiviral activity and mechanism study of pentacyclic triterpenoids and their derivatives. *Medicinal Research Reviews*, 38, 951–976 (2018). http://doi:10.1002/med.21484.

///// Cutting Edge /////

理論創薬及び関連分野における SARS-CoV-2 研究動向 2020~2021/3

国立研究開発法人理化学研究所 生命機能科学研究センター 高谷大輔 https://researchmap.jp/ykgkhakase

1. はじめに

理論創薬研究では基礎研究として、新規手法開発、データ蓄積、モデルケースに対する性能評価等が行われる。またここ最近では多くの分野で人工知能 (AI) の開発が盛んになり、創薬分野も例外ではなく、AI の創薬等への応用[1]や次世代創薬 AI の構築についての研究も始まっている。[2] これらの研究は主に「平常時」に行われることを前提としており、いったん創薬ターゲットが決定したならば基礎研究で培った、いわば "総合力"が問われるわけである。2020 年初頭から猛威を振るった新型コロナウイルスによるパンデミックは一年以上経った 2021 年 3 月時点において、首都圏における緊急事態宣言が 2021 年 3 月 21 日をもって解除されたが、リバウンドを警戒するなど、いまだに収束の気配は見えていない。

この期間において新型コロナウイルスを対象とする創薬研究開発は盛んに行われ、従来の医薬品の適用拡大を目指すドラッグリポジショニング及び新規阻害剤探索からワクチン開発まで と、その対象は低分子から高分子まで多様となった。Wang らの報告では 2020 年 10 月 17 日時 点で 177 品のワクチン候補が開発され、そのうち 54 品のヒトに対する臨床試験が開始されてい た。[3] それから半年が経過した本稿執筆時点 (2021 年 3 月中旬ごろ)の同じデータソースを参 照すると[4]、合計 204 品のワクチン中 80 品の臨床試験が実施されており、世界中で開発が急ピ ッチで進んでいることが読み取れる。また治療薬についても 498 品が開発され、411 もの臨床試 験が開始されている。特に製薬各社での開発が進み、日本においても 2021 年 2 月 17 日からワ クチン接種が開始されるなど[5]、概ね一年程度で未知のウイルスに対するワクチンが世に現れ た事になり、まさしく人類による「叡智と実践」の結果と言えるだろう。

新型コロナウイルスは SARS-CoV-2 と呼ばれるように、10 年以上前に流行した SARS-CoV と の遺伝子配列相同性が 79%ほどあり、その派生型であると報告されている。[6] 創薬のターゲッ トとしては、ウイルスの増殖に関与するメインプロテアーゼ (Mpro)、パパイン様プロテアーゼ (PLPro) および RNA 依存性 RNA ポリメラーゼ (RdRp)、ウイルスの構造タンパク質であるスパ イクタンパク質 (Spike protein) 等が挙げられる。また宿主となるヒト細胞膜に存在するセリン プロテアーゼの TMPRSS2 も注目されている。[6] 筆者は現在 LBDD 及び SBDD を主体としたイ ンシリコスクリーニング担当として複数の創薬研究に関わっているが、実は北里大学の生物分 子設計学教室に在籍していた当時、梅山秀明先生 (現・名誉教授) 及び竹田一志鷹真由子先生 (現・教授)の指導の元、SARS-CoV のターゲットタンパク質の立体構造モデリングに参加した ことがあり、何かと因縁のあるターゲットである。[7] 本稿では、2020 年初頭から 2021 年 3 月 末の期間で、主に創薬に関連する計算科学者らを中心に行われたいくつかの新型コロナウイル スに対する創薬および関連研究を紹介しつつ、著者らの所属するグループにより行われた新型 コロナウイルス関連タンパク質に対する FMO 計算と FMODB への登録を主として報告する。

2. SARS-CoV-2 関連タンパク質構造の PDB への登録数の推移

SARS-CoV-2 を構成する構造タンパク質および非構造タンパク質 (Non-structural protein; NSP) の実験構造が多く報告された。さらに SBDD によるアプローチも行われ、阻害剤との複合体構造も Protein Data Bank (PDB) に数多く登録された。SARS-CoV-2 のターゲットタンパク質で注目されている 4 種のタンパク質 (Mpro、PLpro、RdRp、Spike protein) の登録数について、下記にその推移を示した。(図 1)



図 1: PDB に登録された Mpro、PLpro、RdRp、Spike Protein の実験構造数の月毎の推移。(2021 年 3 月 19 日 までのデータを使用;合計 579 エントリー) それぞれ 6LU7.A、6YVA.A、7BV2.A、7BZ5.A (S1 サブユニッ ト RBD 部分) の配列と 99%以上の配列相同性をもつエントリーを PDB REST サービス[8]を用いて JSON 形式で入手、それぞれ PDBID 数を一ヶ月ごとに集計した。登録日データは'initial_release_date'を使用、描 画には Matplotlib を用いた。[9]

上記のカテゴリの合計は 579 エントリーであった。2020 年 1 月 30 日に WHO による PHEIC (国際的に懸念される公衆衛生上の緊急事態)の宣言がなされてから、[10][11] 僅かの期間で多くの構造ベースの研究が行われてきた事がわかる。2020 年 2 月に公開された SARS-CoV-2 の Mproの結晶構造 (PDBID: 6LU7)から始まり、毎月途切れることなく構造が登録されている。特に Mpro については 2020 年 3 月の登録数が 82 エントリーと本期間中で最も多かった。また 12 月ごろに Mpro の登録数について 2 つ目のピークが現れる。これは大規模 X 線結晶構造スクリーニングの結果、多数の低分子との複合体構造が登録されたためである。前者は 1,250 種以上のフラグメント化合物に対してスクリーニングを行い、23 の非共有結合性化合物を含む 74 の化合物を検出、合計 96 の複合体構造を PDB に登録している。[12] 今後フラグメント分子の相互作用を基にした阻害剤、また例えば FBDD 的なアプローチによる相互に連結した強力な Mpro 阻害剤が得られると期待される。一方、後者は医薬品や臨床試験入り化合物を中心にスクリーニング対象としていることから、フラグメント分子に比べて大きめの化合物の複合体構造が得られており、30 以上のエントリーが登録されている。[13] 後者はよりドラッグリポジショニングを重視したと言えるだろう。

スパイクタンパク質については、6月ごろから登録数が増え始めた。スパイクタンパク質は巨 大なタンパク質であり、抗体医薬のターゲットタンパク質としての側面があり、実際に PDB 登 録情報のタイトル等に抗体関連のキーワード ("antibody"、"nanobody"等) を含むエントリーは少 なく見積もっても100を超えていた。スパイクタンパク質を標的とした抗体医薬では、ヒトACE2 を認識するレセプター結合ドメイン (receptor binding domain; RBD) と結合する抗体や、それぞ れ異なる部位に 2 種の抗体が結合し、いわゆるカクテル療法を目指すケースもあり (PDBID: 6XDG) [14]、その種類は多様である。また、スパイクタンパク質は変異残基が観察されやすいタ ンパク質であり、ウイルスの感染力に影響を与える可能性が示唆されており、[6][15] 構造生物 学的に関心が高かった事がうかがえる。ウイルスの複製に関与する RdRp の登録数が全体的に少 ないのは、nsp12 で 900 残基ある巨大なタンパク質であることなどが関係するのかもしれない。 中でも Yin らはクライオ電子顕微鏡法を用いて RdRp (NSP12-NSP7-NSP8) とレムデシビルとの 複合体構造を報告し、SARS-CoV-2に対する本薬剤の相互作用を示した。[16] また PLpro は SARS-CoV-2 に存在する主要なプロテアーゼの一種であるが、Mpro に比べれば登録数は少ないが定期 的に登録されている。PLpro はウイルスの増殖以外にも宿主の免疫系への関与が指摘されている こともあり[17]、今後の研究が期待される。全般的に本稿執筆時点の 2021 年 3 月中旬において も PDB への登録は続いており、この傾向は今後もしばらく続くと筆者は見込んでいる。

3. FMO 法による生体高分子に対する計算データを蓄積したデータベース~FMODB

FMO 法は北浦らによって提唱された量子化学計算法であり、アミノ酸等残基単位、または阻害剤などの低分子単位を参考にフラグメントに分割し、原子数の非常に多いタンパク質等の生体高分子に対しても現実的な時間内で量子化学計算が可能である。[18] また FMO 法では、フラグメント間 (ダイマー)の相互作用エネルギーを IFIE/PIEDA により精密に評価できる。タンパク質及び核酸分子などの生体高分子とリガンド分子、タンパク質―タンパク質間相互作用解析だけでなく、機械学習モデルなどの相互作用記述子の一つとして近年注目が集まっている。この様に IFIE/PIEDA の蓄積は重要であることから、著者らのグループでは FMODB として理研やFMODD コンソーシアムで計算された計算データを定期的に公開している。[19]

4. FMODB からの SARS-CoV-2 ターゲットタンパク質の計算データ公開

毎週ごとに追加される PDB データに対して FMO 計算を実施することとした。本計算は 2020 年 3 月ごろから開始された。アップデート自体は本稿執筆時点でも二週間を目安に定期的に行っている。(https://drugdesign.riken.jp/FMODB/covid-19.php)



図 2: FMODB の Web サイトにおける最新の COVID-19 特集ページ



図 3: FMODB で公開済みの計算データ登録数の推移。横軸は FMODBID が FMODB へ登録された日とした。(2021 年 3 月 24 日時点のデータを使用)

2020年4月の公開当初は65の実験構造についてFMO計算を実施した。[20](図2)ホームページの各タンパク質のリボン図に貼られたリンクから計算データを参照でき、阻害剤などのリガンド分子が存在する場合はタンパク質との相互作用図を通常のFMODBと同じ手順で参照できる。一年近く経過した2021年3月時点では合計510の計算データ登録に至っている。(図3)FMODB に登録する SARS-CoV-2のPDBデータの選定は現時点では自動化されておらず、一つのPDBID について計算条件の異なる複数の計算結果がある場合もある。今後も継続的にアップデートを行う予定である。以下ではMpro、RdRp、Spike protein について行われた計算科学からのアプローチや、FMO計算を用いた解析例について紹介する。

5. Mpro を対象とした研究

SARS-CoV-2 の Mpro は翻訳されたポリタンパク質を切断する機能を有し、原稿執筆時点でも PDB に新規の実験構造が報告されるなど、図 1 で示したとおり、構造データが豊富に使用でき ることから、SARS-CoV-2 の阻害剤探索において最も創薬研究が盛んなターゲットタンパク質と 言えるだろう。2020 年 2 月に SARS-CoV 及び MERS-CoV 等の複数のコロナウイルスの Mpro を 阻害する阻害剤 N3 との結晶構造 (PDBID:6LU7) が 2020 年 2 月に PDB で公開された。[21] 筆 者は本構造公開をきっかけに創薬の競争が始まったなという印象をもったが、国内でも HIV プ ロテアーゼ阻害剤のネルフィナビルが Mpro に結合することを計算モデルにより示した報告[22] は記憶に新しい。FMO 計算の観点からは、本複合体結晶構造を対象とした FMO 計算がなされ ている。[23] 本研究では阻害剤 N3 を 5 つのフラグメントにマニュアルで分割し、各フラグメン トの相互作用を定量的に分析している。例えば N3 のフラグメント1 に注目した場合、フラグメ ント 1 中のペプチド結合部分の窒素原子と Thr190 との水素結合が非常に強いことが示されてい る。計算データは FMODB からも参照することができる。(図 4、FMODBID: R1GK8)



図 4: (左) 阻害剤 N3 の結合サイトの周辺残基 (PDBID: 6LU7) 各点線は IFIE を示す。(右) フラグメント1 (残基名:02J、右図中のオキサゾール環とアミド結合からなるフラグメント) からの 4.5Å 以内にあるフラグ メントの PIEDA を表示した。(FMODBID: R1GK8) グラフの残基番号 191 は Thr190 の主鎖のカルボニル基 が水素結合に関与している影響が大きい。

6. RdRp を対象とした研究

SARS-CoV-2 の RdRp (NSP12) はウイルスの RNA の複製に必須のタンパク質であり、NSP7 および NSP8 と複合体を形成することでその機能を強くする。レムデシビルはエボラウイルスに対する阻害剤として開発され、プロドラッグであり、体内でリン酸化されたのち合成中の RNA 鎖に取り込まれる事で伸長を阻害する。エボラウイルス以外にもコロナウイルスに対しても効果がある事が知られていた。[24] 2020 年 4 月 22 日にクライオ電子顕微鏡法を用いて得られたレムデシビルとの複合体構造 (PDBID: 7BV2) が PDB から公開された。[16] 加藤らは本構造に対して FMO 計算を行い、レムデシビル周辺の残基及び核酸残基との相互作用を解析した。[25] 本解析では RNA ポリマーに取り込まれたレムデシビルと周辺残基の IFIE/PIEDA 解析を実施し、実験構造中で確認されていた 5 つの代表的な相互作用残基の中で Thr687、Asn691、Asp760 がより重要であること相互作用エネルギーの観点から示している。FMO 計算データは FMODB から参照可能である。(図 5、FMODBID: 1JL3Z)



図 5: (左) レムデシビルと周辺残基 (PDBID: 7BV2) (右) レムデシビルフラグメント (残基名:F86) から 4.5Å にあるフラグメントの PIEDA を表示 (FMODBID: 1JL3Z)

7. スパイクタンパク質を対象とした研究

スパイクタンパク質は SARS-CoV-2 の構造タンパク質の一つであり、ヒトの細胞に侵入する 際に宿主のアンジオテンシン変換酵素 2 (ACE2) を認識し、ACE2 と結合するドメインを特に receptor binding domain (RBD) と呼ぶ。2020年3月11日には RBD と ACE2-B0AT1 複合体構造が PDB から公開された。[26] スパイクタンパク質は他の SARS-CoV-2 のタンパク質と比べて変異



図 6: N501Y の変異部位 (RBD; シア ン) と抗体 (COVOX-269; マゼンタと オレンジ) の複合体構造 (PDBID: 7NEG) が入りやすい部位として知られる。例えばイギリスや南ア フリカで確認されている変異株 (VUI-202012/01 その後 VOC-202012/01 に変更、B 1.1.7 系統[27]) においては、ス パイクタンパク質に複数の変異が報告されている。[28] 特に RBD の Asn501 の Tyr501 への変異 (N501Y)(図 6) で は、ACE2 への結合親和性が増加する可能性が報告されて いる。[15] スパイクタンパク質を対象に複数の計算科学 のアプローチがなされている。 阪大の渡邉らは抗体 (もし くは関連するペプチド)とスパイクタンパク質の複合体 構造 12 エントリーに対して FMO 計算を実施し、PIEDA 相互作用の観点から相互作用な重要な 9 つの残基を示し ている。[29] また、理研の渡邉らはスパイクタンパク質と 抗体複合体構造に対して IFIE/PIEDA 解析を実施、XH/π相 互作用を正確に評価することが相互作用の理解に重要で あることを示している。さらにN501Yの変異によりACE2 と相互作用エネルギーが増加することを示している。[30] (FMODBID: 4NZVN、7J11K など) また N501Y を対象とし た分子動力学計算による解析も bioRxiv に報告されており、 本変異により相互作用エネルギーが向上した事が示され ている。[31] また Supasa らは N501 または Y501 をもつス パイクタンパク質に対する各抗体の Kd 値を報告し、変化 がない抗体が多いものの、特に変化の大きい抗体種につい て報告している。[32]

8. 終わりに

2020 年初頭から SARS-CoV-2 関連の創薬ターゲットとなる構造や、阻害剤との複合体の実験 構造は PDB からすぐに公開された。またインターネットの普及により研究は世界中を巻き込ん で行われ、新規の阻害剤情報等もすぐに報告され、理論創薬分野ではインシリコスクリーニング から相互作用研究のような基礎研究まで幅広く行われた。治療薬開発では特に海外のワクチン 開発等が先行しているが、今後いわゆるチーム・ジャパンまたは日本発創薬の成果も続々と公開 されていくと期待される。

SARS-CoV、SARS-CoV-2 と本ウイルスの創薬研究に関わってきた著者だが、どうも縁がある ようである。「2 度あることは 3 度ある」ではないが、コロナウイルスに限らず、十数年先に新 たなパンデミックが起きる可能性もゼロではないだろう。私も含めアカデミア創薬に関わる研 究者は己の知識及び技術のアップデートを絶やすことなく自己の研鑽を続け、得られた知識を 惜しむことなくコミュニティで共有し、その時に備えていただきたい。

謝辞

本稿を執筆するにあたり原稿チェック、グラフや図表の作成を手伝っていただいた理化学研 究所 生命科学機能研究センター 制御分子設計研究チームのみなさまに深く感謝いたします。

参考文献

- [1] https://linc-ai.jp/ (2021 年 3 月アクセス)
- [2] https://www.amed.go.jp/program/list/11/02/001_02-04.html (2021 年 3 月アクセス)
- [3] Wang, M.-Y., Zhao, R., Gao, L.-J., et al. SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. *Front. Cell. Infect. Microbiol.* 10:587269, 2020 doi: 10.3389/fcimb.2020.587269
- [4] https://biorender.com/covid-vaccine-tracker (2021 年 3 月 16 日にアクセス)
- [5] https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000121431_00218.html (2021 年 3 月 26 日アク セス)
- [6] Hu, B., Guo, H., Zhou, P. et al. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 19, 141–154 (2021).
- [7] Takeda-Shitaka, M., Takaya, D., Chiba, C. et al. Protein structure prediction in structure based drug design *Curr. Med. Chem.* **11** (5), 551-558 (2004)
- [8] Rose, Y., Duarte, J. M., Lowe, R. et al. RCSB Protein Data Bank: Architectural Advances Towards Integrated Searching and Efficient Access to Macromolecular Structure Data from the PDB Archive, *J. Mol. Biol.* 166704, 2020 DOI: 10.1016/j.jmb.2020.11.003
- [9] Hunter, J. D., "Matplotlib: A 2D Graphics Environment," Comput Sci Eng, 9 (3) 90-95, 2007.
- [10] https://www.who.int/publications/m/item/covid-19-public-health-emergency-of-internationalconcern-(pheic)-global-research-and-innovation-forum (2021年3月26日アクセス)
- [11] https://www.forth.go.jp/topics/202001311700.html (2021 年 3 月 26 日アクセス)
- [12] Douangamath, A., Fearon, D., Gehrtz, P. et al. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *Nat Commun* 11, 5047 2020 https://doi.org/10.1038/s41467-020-18709-w
- [13] Günther, S., Reinke, P. Y. A., Fernández-García, Y. et al. Inhibition of SARS-CoV-2 main protease by allosteric drug-binding bioRxiv 2020.11.12.378422; doi: https://doi.org/10.1101/2020.11.12.378422
- [14] Hansen, J., Baum, A., Pascal, K.E. et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail *Science* 369, 1010–1014, 2020
- [15] Gu, H., Chen, Q., Yang, G. et al. Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy. Science 369 (6511), 1603-1607, 2020 doi: 10.1126/science.abc4730.
- [16] Yin, W., Mao, C., Luan, X. et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science*. 368 (6498), 1499-1504,2020
- [17] Shin, D., Mukherjee, R., Grewe, D. et al. Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* 587, 657–662 (2020).
- [18] Kitaura, K., Ikeo, E., Asada, T. et al. Fragment molecular orbital method: an approximate computational method for large molecules, *Chem. Phys. Lett.*, **313**, 701-706 (1999).
- [19] Takaya, D., Watanabe, C. Nagase, S. et al. FMODB: The World's First Database of Quantum Mechanical Calculations for Biomacromolecules Based on the Fragment Molecular Orbital Method J. Chem. Inf. Model. 61 (2), 777-794, (2021)
- [20] https://www.amed.go.jp/news/release_20200417-02.html (2021 年 3 月 26 日アクセス)

- [21] Jin, Z., Du, X., Xu, Y. et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* 582, 289–293 (2020). https://doi.org/10.1038/s41586-020-2223-y
- [22] Ohashi, H., Watashi, K., Saso, W. et al. Multidrug treatment with nelfinavir and cepharanthine against COVID-19 bioRxiv 2020.04.14.039925; doi: https://doi.org/10.1101/2020.04.14.039925
- [23] Hatada, R., Okuwaki, K., Mochizuki, Y., et al. Fragment Molecular Orbital Based Interaction Analyses on COVID-19 Main Protease – Inhibitor N3 Complex (PDB ID: 6LU7) J. Chem. Inf. Model. 60 (7), 3593-3602, 2020
- [24] Warren, T., Jordan, R., Lo, M. et al. Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. *Nature* 531, 381–385 (2016). https://doi.org/10.1038/nature17180
- [25]Kato, K., Honma, T., Fukuzawa, K. Intermolecular interaction among Remdesivir, RNA and RNAdependent RNA polymerase of SARS-CoV-2 analyzed by fragment molecular orbital calculation J. Mol. Graph., 100, 107695, 2020
- [26] Yan, R., Zhang, Y., Li, Y. et al. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2 Science 367 (6485), 1444-1448, 2020
- [27] https://www.mhlw.go.jp/content/10900000/000741774.pdf (2021,3月25日アクセス)
- [28] Rahimi, F., Talebi Bezmin Abadi, A. Implications of the Emergence of a New Variant of SARS-CoV-2, VUI-202012/01. Arch Med Res. S0188-4409 (21) 2021 doi: 10.1016/j.arcmed.2021.01.001. Epub ahead of print.
- [29] Watanabe, K., Watanabe, C., Honma, T. et al. Intermolecular Interaction Analyses on SARS-CoV-2 Receptor Binding Domain and Human Angiotensin-Converting Enzyme 2 Receptor-Blocking Antibody/peptide Using Fragment Molecular Orbital Calculation. *ChemRxiv*. Preprint. https://doi.org/10.26434/chemrxiv.14135138.v1
- [30] Watanabe, C., Okiyama, Y., Tanaka, S. et al. Molecular Recognition of SARS-CoV-2 Spike Glycoprotein: Quantum Chemical Hot Spot and Epitope Analyses *Chem. Sci.*, 2021, DOI: 10.1039/D0SC06528E.
- [31] Luan, B., Wang, H., Huynh, T. Molecular Mechanism of the N501Y Mutation for Enhanced Binding between SARS-CoV-2's Spike Protein and Human ACE2 Receptor *bioRxiv* 2021.01.04.425316; doi: https://doi.org/10.1101/2021.01.04.425316
- [32] Supasa, P., Zhou, D., Dejnirattisai, W. et al. Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera *Cell*, in press, 2021 https://doi.org/10.1016/j.cell.2021.02.033

///// SAR Presentation Award /////

SAR Presentation Award について

「SAR Presentation Award」は、構造活性相関シンポジウムにおける若手研究者の発表を奨励 し、構造活性相関研究の発展を促進するため、2010年度に創設された。当初は応募制とし て審査対象講演の募集を行った。2012年度からは、正式名称を「構造活性相関シンポジウ ム優秀発表賞」(英語表記 SAR Presentation Award)と定めた。

2020 年度 SAR Presentation Award について

2020年度は、第48回構造活性相関シンポジウム(オンライン開催)における45歳以下 の発表者(日本薬学会会員または受賞後に日本薬学会に入会いただける方)による一般講 演(ロ頭発表・ポスター発表)を選考対象とすることとした。

2020 年度 SAR Presentation Award 受賞者(五十音順)

口頭発表	:	笠原	慶亮	(東京大学大学院)
口頭発表	:	千葉	峻太朗	(理化学研究所)
ポスター発表	:	伊藤	朱里	(横浜市立大学大学院)
ポスター発表	:	髙橋	真帆	(横浜市立大学大学院)
ポスター発表	:	原田	祥季	(北里大学大学院)

受賞者の選考について

2020年12月12・13日にオンラインにて各審査員から提出頂いた審査票を集計し、ロ頭 発表2名、ポスター発表3名を受賞者として選出した。ロ頭発表の審査は点数方式、ポス ター発表の審査は3演題選出するという方式で行った。授賞式は、第48回構造活性相関シ ンポジウムの閉会式において行った。後日、受賞者には、賞状と副賞を贈呈した。なお、 審査にあたっての審査項目は下記の通りである。

口頭発表審査項目

- a) 講演要旨: 講演要旨は発表内容を反映して適切に作成されているか。
- b) 講演資料: スライドは専門領域の異なる参加者にもわかりやすく、見やすく、かつ発 表時間に見合って適切に作成されているか。
- c) プレゼンテーション:発表時に参加者にわかりやすく説明しているか。
- d) 研究の目的:研究の背景と目的、先行研究との関係、研究の新規性あるいは有用性が 明確になっているか。
- e)研究成果:価値のある成果が得られているか。
- f) 質疑応答: 質問等に対し、的確な応答・議論がなされたか。活発な討論がなされたか。
- g) 将来性:研究内容について、将来の発展が期待できるか。

審査員

第48回構造活性相関シンポジウムに参加した2020年度常任幹事および幹事

<受賞者コメント>

KO-01

氏名 笠原 慶亮(かさはら けいすけ) 所属 東京大学大学院 工学系研究科 バイオエンジニアリング専攻 演題 計算デザイン Supercharging 抗体の物性機能解析

この度は、第48回構造活性相関シンポジウム優秀発表賞(ロ頭)を賜りまして、大変光栄に 存じます。ご評価いただきました先生方、ならびに日本薬学会構造活性相関部会の先生方に厚く 御礼申し上げます。

私共の研究は、抗体表面の電荷を変化させることによる物性機能への影響や制御の可能性に ついて詳細に議論することを目的としております。本発表では、Rosetta を用いて抗体表面への 荷電アミノ酸変異導入デザインを行い、発現させた Supercharging 抗体について安定性解析・相 互作用解析を行いました。また、実験結果と分子動力学シミュレーションを組み合わせることに より、表面電荷の大幅な変化が抗体の物性に及ぼす影響について考察しました。さらに私共は、 荷電抗体の溶媒環境依存的な物性変化や抗原結合における速度論パラメータについて、実験と 計算科学を組み合わせてより詳細な議論を展開したいと考えております。今回の受賞を励みに、 医薬品・材料応用を指向した荷電抗体デザインの可能性を広げていく所存です。

最後に、本研究を進めるにあたりご指導を賜りました、津本浩平教授、長門石曉特任准教授、 黒田大祐講師、田部亜季特任助教、河出来時さんに深く感謝申し上げます。

本発表につきましては要旨の掲載を控えさせていただきます。(編集委員会)

<受賞者コメント>

KO-06

氏名 千葉 峻太朗(ちば しゅんたろう) 所属 理化学研究所 医薬プロセス最適化プラットフォーム推進グループ 演題 幾何学的相互作用解析と機械学習を組合わせた抗体の側鎖構造予測

このたび、第48回構造活性シンポジウム優秀発表賞(ロ頭)を賜り、大変光栄に存じます。 長年にわたり構造活性相関部会およびシンポジウムを運営されてきた歴代の先生方に、本発表 の機会をいただきましたことを深くお礼申し上げます。同時に、ウェブ開催を決断し、しかも円 滑な運営を実現された実行委員会の皆様に深くお礼申し上げます。

抗原・抗体の構造モデリングは抗体の物性や抗原への親和性を調整するうえで重要です。本研 究ではモデル構造(側鎖周辺のローカルな構造)が、多数の実験構造からの集約された情報から みて、妥当であるか判定する手法開発を目指しました。このために、水素結合のような典型的な 相互作用に加え、いわゆる弱い相互作用(CH-O, CH-πなど)も加味した幾何学的相互作用解析 を、機械学習に組み合わせる手法を考案し、従来法と同等以上の正確度を有する手法開発に成功 しました。いまはハイパーパラメータ探索や、学習データ数を大幅に増やすなどの課題に対応す ることで、さらなる改良を目指しています。今後は、開発した手法を利用して抗体の物性や抗原 への親和性調整へも取り組みたいと考えています。

最後に、共同発表者の小甲裕一先生、野沢優翼先生、柳田駿介先生、佐藤美和先生、池口満徳 先生、大田雅照先生に感謝申し上げます。

KO-06

幾何学的相互作用解析と機械学習を組合わせた 抗体の側鎖構造予測

○千葉峻太朗¹、小甲裕一²、野沢優翼²、柳田駿介²、 佐藤美和²、池口満徳^{1,3}、大田雅照¹

(¹理研·MIH、²三井情報、³ 横浜市大·生命医科研) E-mail: shuntaro.chiba@riken.jp

1. 背景と目的

抗体を治療薬、診断薬とするためには、抗体の 安定性などの物性や抗原に対する親和性を改変 する必要がある。このため、抗体のアミノ酸変異 が行われる。アミノ酸変異は通常1残基ずつ実施 し、親和性や安定性が向上したものを組み合わせ て、さらなる改変が実施される。我々は、通常の 1 残基変異ではない新しい変異法として、2 残基 を同時に変異し、変異残基同士の協同的効果によ る親和性改善を試み、これに成功した[1]。興味深 いことに、変異2残基を個別に1残基ごとに変異 させると親和性低下あるいは発現せず、2 残基を 同時に協同的に変異させることが重要であるこ とも判明した。しかしながら、多残基同時変異は、 組合わせが膨大であるため、実験の前に、膨大な 多残基変異体モデルから、効果の見込める変異体 を計算機によって絞り込むことが求められる。

多数の多残基同時変異体から有望なものを絞 り込むためには、抗原-多残基同時変異体複合体モ デル構造から、その親和性変化を予測・判断する ことが求められる。この場合は計算機上でモデリ ングした構造が妥当であることが必要である。し かしながら、モデル構造が妥当かどうかを判定す ることは、アミノ酸側鎖ロータマーの組合せや水 分子の配置の大きな多様性のため、容易ではない。

本研究では、抗原・抗体のモデル構造(側鎖周 辺のローカルな構造)が妥当であるか判定する手 法開発を目指した。このため、水素結合のような 典型的な相互作用に加え、いわゆる弱い相互作用

(CH-O, CH-πなど)も加味した幾何学的相互作用 解析を、機械学習に組み合わせる手法を考案し、 その正確度を検証した。

2. 方法

妥当構造の判定の機械学習モデル構築のため、 妥当構造/非妥当構造のデータセットを作成した。 まず、PDBからダウンロードした約80個の抗原-抗体複合体(解像度 < 2 Å,抗原はタンパク質ま たはペプチド)の各構造について、抗体アミノ酸 のうち、抗原に接しているアミノ酸を側鎖モデリ ングの対象とした。これら各アミノ酸を SP (single point) とよぶ。各 SP にさらに接しているアミノ 酸を選択し、その SP と組み合わせた 2 アミノ酸 も、側鎖モデリングの対象とした。各組合せを DP (double point) とよぶ。SP または DP および その周辺構造を水も含めて FoldX[2]を用いて再モ デリングした構造のうち、PDB 構造に近い構造 (symmetry-corrected RMSD < 1 Å) [3]を妥当構造、 遠い構造 (同 RMSD \geq 1 Å) を非妥当構造と定 義した。約 70 種類の PDB 由来構造を学習セット、 残りを検証セットとした。最終的に、学習セット と検証セットの(妥当, 非妥当) 構造数は、(7200, 3000) と (900, 400) になった。

各モデル構造に対して、幾何学的相互作用解析 [1]を適用し、記述子を作成した。具体的には、 SP/DP アミノ酸ー周囲のアミノ酸間、SP/DP アミ ノ酸ー水分子ー周囲のアミノ酸間、(側鎖モデリン グ対象アミノ酸間(DP の場合のみ)について、 水素結合と VDW 相互作用に加え、弱い相互作用 を含め、各相互作用の判定条件に合致する数をカ ウントした。加えて、SP/DP アミノ酸ロータマー 頻度および SP/DP アミノ酸と立体衝突している周 辺の原子数も記述子に含めた。以上の記述子の種 類は 530 個程度である。記述子化した学習セット を用いて、Random forest による分類モデルを構築 し、検証セットに適用した。

3. 結果とまとめ

3-1. 検証セットを用いた判別性能

学習セットでの判別の正確度は 1.0 であった。 一方、検証セットにおける正確度は平均的に 0.79 であった (図 1 の All)。PDB ごとの正確度にはば らつきが大きいが、検証セットの抗体の CDR に 類似する CDR が学習セットに含まれているほど、 高い傾向にあった。しかしながら、1yqv や 4i77 のように、CDR 一致度が 40-50%程度であっても 正確度が高いものも存在した。以上のことから、 過学習傾向がみられるため、ハイパーパラメータ



■Accuracy ●Max CDR identity 図 1. 検証セットの PDB ごとの正確度 (Accuracy)および検証セットの各 PDB の抗 体の相補性決定領域(CDR)と学習セットの 各 PDB の CDR の Identity の最大値(Max CDR identity)。All はすべての検証セットの PDB の 結果を合わせたもの。

の最適化、記述子選択に加えて、学習データをさ らに増やすことでさらに汎化性能の向上が期待 される。

3-2. 重要な記述子

図2に示したように、記述子のなかで重要度が 高いものを調べると、ロータマー頻度(rot_freq) と立体衝突数(crash)が1位、2位となり、直感 的にも妥当な結果が得られた。興味深いのは、 CH-π、CH-O、NH-πなどの弱い相互作用が、妥当 構造の判別性能に大きく寄与していたことであ



図 2. 記述子の重要度(Feature importance)の 上位 10 個。重要度は Permutation importance に よって計算した。すなわち、全学習セットで作 成したモデルの Out-of-bag のデータの正確度 から、Out-of-bag の各記述子に対応するデータ をシャッフルした場合に計算した正確度を引 いたものとして定義した。AgAb#は側鎖モデリ ング対象残基一抗原または抗体間、M#は対象 残基一対象残基間、W#は水を介する相互作用 を表す。

る。これらは、側鎖モデリングの際に強い相互作 用のみならず、いわゆる弱い相互作用を考慮する ことが重要であることを示唆している。なお、今 回は、FoldXによって作成した妥当構造/非妥当構 造の判別を行ったため、別の側鎖モデリング手法 で作成したデータセットでは異なる順位になる ことも考えられるため注意を要する。

3-3. FoldX との比較

FoldX による側鎖モデリングの正確度と、本研 究の手法の正確度を比較するため、SP または DP およびその周辺構造を水も含めて FoldX を用いて 再モデリングした構造のうち、エネルギー最低の ものを FoldX の予測する妥当構造と定義した。こ の構造を PDB 構造と比較して RMSD <1 Å の場合 を真陽性、そうでない場合を偽陽性として正確度 (この場合は精度と一致)を計算した。検証には、 検証セットの PDB(図1)を利用した。一方、同 じ検証セットに対して、本研究の学習済みモデル を適用し正確度を調べた。

全検証セットまたは CDR 一致度が低い PDB の みで検証する場合では、本研究の予測正確度は FoldX を上回るか匹敵するものだった(表)。FoldX は抗体側鎖の変異導入に伴う抗原抗体間親和性 予測において、実験データ再現性が最も良いプロ グラムの一つであり[4]、側鎖構造予測についても 同様に高い性能を有することが予想され、本研究 の側鎖妥当性判定手法も、高い性能を有すること が期待される。

本研究では、既存プログラムに匹敵するか上回 る性能を有する側鎖構造妥当性判定モデルを開 発できた。今後、記述子選択、ハイパーパラメー タ最適化に加え、学習セットを増やすことで、さ らなる性能の向上を目指していきたい。

表.本研究(Ours)とFoldXの比較

Test set	All PDB ^a		Low CDR identity ^b		
Method	Ours	FoldX	Ours	FoldX	
Accuracy	0.84	0.78	0.80	0.80	

(a) 検証セットのすべての PDB を利用。(b)
 CDR identity < 50%の PDB のみを考慮。

4. 謝辞

計算資源をご提供いただいた Hokusai BigWaterfall システムに感謝します。

5. 参考文献

1. Chiba et al., *Sci Rep* **10**, 17590 (2020). 2. Guerois et al., *J Mol Biol* **320**, 369 (2002). 3. Bell et al., *J Cheminform* **11**, 40 (2019). 4. Sirin et al., *Protein Sci* **25**, 393 (2016).

<受賞者コメント>

KP-13

氏名 髙橋 真帆(たかはし まほ)

所属 横浜市立大学大学院 生命医科学研究科

演題 ニューロトロフィン受容体 TrkAd5 と結合ペプチド間の相互作用様式の解明

この度は、第48回構造活性相関シンポジウム SAR Presentation Award を頂き、 誠にありがと うございます。ご評価下さいました審査員の先生方、並びに日本薬学会構造活性相関部会の先生 方に、厚く御礼申し上げます。当日は、アカデミアの先生方や製薬企業の方々と議論させて頂き ました。自身の研究内容の改善点も明らかになり非常に有意義な時間となりました。

本研究では、TP1の結合ポーズを解明するために、横浜市立大学のスーパーコンピュータを用いて、拡張アンサンブルシミュレーションを行いました。その結果、NMRの実験結果と一致する、TP1の結合ポーズを明らかにすることができ、結合に大事な相互作用部位の特定もできました。今後は、この知見を活かして、NMRとの共同研究により、結合能を上げるデザインに取り組んでいきたいと思います。

最後に、研究を進めるにあたりご指導いただきました、池口満徳教授、浴本亨助教、理化学研 究所の山根努上級研究員、共同研究先として多くを学ばせていただきました、高橋栄夫教授、鈴 木里佳研究員に心より感謝申し上げます。

KP-13

ニューロトロフィン受容体 TrkAd5 と 結合ペプチド間の相互作用様式の解明

○髙橋真帆¹、浴本亨¹、鈴木里佳¹、山根努²、高橋栄夫
 ¹、池口満徳^{1,2}

(¹横浜市大・生命医、²理化学研究所・MIH)

E-mail: w205425d@yokohama-cu.ac.jp

1. 背景と目的

細胞表面受容体チロシンキナーゼAドメイン 5(TrkAd5)は、Trk 受容体ファミリーの1つで、神経成長因子(NGF)と特異的に結合する。NGFは、 疼痛状態において発現量が増加するため、 TrkAd5-NGF間の結合阻害が、神経変性疾患の創 薬標的になると期待されている。これまでに、フ ァージディスプレイ法により、TrkAd5と結合する TrkA-binding peptide1(TP1)が発見され、NMRを用 いた相互作用サイトの特定が行われてきたが、 TrkAd5-TP1 複合体構造の構造は解かれておらず、 具体的な相互作用様式を構造的に解明すること が求められていた。

そこで、我々は、NMR 実験データと拡張アン サンブル手法による構造サンプリング分子動力 学シミュレーションを相補的に組み合わせた NMR-REST 法を用いて、TrkAd5-TP1 の結合様式 を全原子レベルで求め、相互作用解析を実施した。 TP1 は 10 残基のペプチドであるが、構造が柔軟で あり、かつ、NMR データから示唆された結合サ イトは広い面であった。通常の分子動力学シミュ レーション (MD) では限られた結合ポーズのみ しかサンプリングできなかったが、拡張アンサン ブル手法では、複数の結合ポーズがサンプリング でき、その結合ポーズが NMR データと一致する ことを示す。

2. 方法

TrkAd5-NGF 結晶構造を参照に、TrkAd5-TP1 構 造をモデリングし、通常の MD と拡張アンサンブ ル手法である gREST 法[1]による構造サンプリン グを実施した。

3. 結果とまとめ

拡張アンサンブル法による構造サンプリング により、NMR 実験データと一致した、結合ポー ズのアンサンブルモデルを得ることができた(図 -1)。通常 MD の範囲では1つの結合ポーズのみ がサンプリングされたが、gREST 法では、2つの 結合ポーズが出現した。REST から得た構造アン サンブルモデルを用いて、TrkAd5-TP1 間のコンタ クトを解析したところ、コンタクトの頻度が高い 領域と、TP1 結合に伴う化学シフト変化量が大き い領域がよく一致していた。2つの典型的な結合 ポーズの相互作用様式を解析したところ、TP1 の 少数の残基が TrkAd5 の特定の領域と強く相互作 用していることもわかった。



4. 参考文献

1. Kamiya M, Sugita Y, J. Chem. Phys. 149, 072304-1-11 (2018),

<受賞者コメント>

KP-21

氏名 伊藤 朱里(いとう あかり)

所属 横浜市立大学 生命医科学研究科

演題 中分子シクロスポリンAとシクロスポリンEの分子ダイナミクスの比較

この度は名誉ある賞を頂戴し、大変光栄に思います。会期中は、多くの専門家の方々に質問やア ドバイスをいただきました。様々なバックグラウンドをもつ方々の考え方を知ることができ、私 の研究生活において非常に意義深いものとなりました。

本研究は、経口投与可能な中分子医薬品シクロスポリン A と、その代謝物であるシクロスポリ ンEの膜透過性の差が生まれる要因に分子動力学シミュレーション (MD) から迫るものであり ます。膜中でシクロスポリンが安定に存在する位置や、構造変化する過程の中間的な構造がわか っただけでなく、中間構造の特徴が CsA と CsE で異なることもわかりました。しかし膜透過過 程そのものはシミュレーションできていないので、今後も研究を進めて、膜透過のメカニズムを 明らかにしていきたいと思います。

これらの研究を進めるにあたりご指導を賜りました池口満徳教授をはじめ、多くのアドバイス をくださった浴本亨助教、理化学研究所の山根努上級研究員、生命情報科学研究室の皆様に心よ り感謝申し上げます。今回の経験を活かして、残りの研究生活をより一層精進していきたいと思 います。

KP-21中分子シクロスポリンAとシクロスポリンEの
分子ダイナミクスの比較

 ○伊藤朱里¹、浴本亭¹、山根努²、池口満徳^{1,2} (¹横浜市大・生命医、²理研・MIH)
 E-mail: w205406f@yokohama-cu.ac.jp

1. 背景と目的

シクロスポリン A(CsA)は、免疫抑制剤として 用いられる、経口投与可能な中分子医薬品である。 11 残基から成る環状ペプチドで、高い膜透過性を 有している。シクロスポリン E(CsE)は CsA の代 謝物である。CsA と CsE のアミノ酸の配列におけ る違いは、MVA か VAL かという 1 残基のみであ るにも関わらず、CsA は CsE の約 10 倍膜透過性 が高い[1]。

分子量が大きいにも関わらず、CsA が高い膜透 過性を持つ要因は、溶媒環境に応じて動的構造変 化を起こすためであると示唆されている。これま でに、シクロフィリンーCsA 複合体の結晶構造 (PDBID: 1CWA)と、クロロホルム溶媒中の CsA

結晶構造(CSD: DEKSAN)が解かれている。タ ンパク質に結合した CsA は環が拡がった、親水性 の原子を溶媒中に露出させた open 構造をとって いたが、クロロホルム中の CsA は環がタイトで、 親水性の原子が分子内で水素結合を形成した closed 構造をとっていた。その open/closed 構造が 親水/疎水環境での構造であると想定し、溶媒から 膜中、膜中から溶媒環境へ open/closed 構造を変化 させながら膜透過をするモデルが提案されてい る。また、結晶構造では、open/closed 構造間で、 隣り合う N メチルロイシンの主鎖二面角が trans/cis に変化していた事から、二面角の変化と 構造変化の関係も示唆されている。しかし、実際 の膜水系での透過過程は明らかにされておらず、 メカニズムの解明が求められている。クロロホル ム中の CsE の結晶構造 (CSD:SUONUN) も解か れており、分子内水素結合の様式が CsA とは異な る closed 構造であった。しかし、主鎖レベルでは 本質的に CsA と CsE は一致しており、膜透過性 の違いは、解明されていない。

そこで、本研究では、分子動力学シミュレーション(MD)によって、水中や膜中のシクロスポリンの構造ダイナミクスを求め、CsAとCsEの膜透過性の違いが生まれる要因を明らかにすることを目的とした。MDには、我々が開発を進めて

いる、CsAの7残基に関与する、Nメチルアミノ酸の主鎖に対応した新規CHARMM力場を用いた。

2. 方法

CsA の open/closed 構造、及び、CsE の closed 構造を水中に配置した溶液系と、膜の中心に配置 した膜水系を用意し、それぞれ、 $1-3 \mu s \cap MD を$ 実施した。(合計 12 本の MD、計 13.5 μs)

3. 結果とまとめ

膜中における、CsA と CsE の重心位置を解析し たところ、膜の中心に配置したシクロスポリンは シミュレーション全体を通して膜から出ること はなく、自発的な膜透過イベントはなかったが、 膜の中心ではなく、膜分子のアシル鎖付近の位置 にいる傾向が大きいという、CsA と CsE に共通の、 膜中で安定な位置があることがわかった。

また、主鎖構造の特徴を PCA で解析したとこ ろ、CsA と CsE では膜中/水中でとりやすい構造 が異なり、構造変化における特徴も異なることが わかった。結晶構造で得られている closed/open 構 造まわりの溶液構造だけでなく、その中間構造が あることがわかった。その構造間の構造変化を解 析したところ、CsA では closed 構造のような細長 い構造が、open 構造に近い丸い構造へと変化して いくことがわかった。対照的に、CsE では、closed 構造が捻じれるように構造変化していくことが わかった。本研究で実施したµsの MD の範囲で は、2つのメチルロイシン間二面角の cis/trans 変 化は見られなかった。しかし、MD 中では、CsA の closed/open 構造とは異なる中間構造が得られ たため、必ずしも、cis/transの状態が、closed/open 構造に対応しないことがわかった。

4. 参考文献

1. Ahlbach L C, Lexa W K, Bockus T A, Chen V, Crews P, Jacobson P M, Lokey S R, *Future Med. Chem.* **7**, 2121-30 (2015).

<受賞者コメント>

KP-25

氏名 原田 祥季(はらだ よしき)

所属 北里大学大学院薬学研究科

演題 タンパク質複合体立体構造予測における Quality 評価法の予測精度の比較検証

この度は、第48回構造活性相関シンポジウム優秀発表賞(ポスター)を頂き、大変光栄に存 じます。評価をしてくださった先生方、並びに日本薬学会構造活性相関部会の先生方に厚く御礼 申し上げます。

本研究では、タンパク質複合体立体構造予測に適用可能な4種類のQuality評価法の予測精度の比較検証を行いました。その結果、タンパク質複合体ターゲットの構造予測の難易度による得意不得意といった、それぞれのQuality評価法の特徴を知ることができました。また、本研究の知見を基に2020年度に開催されたCASP14に参加し、Assembly部門において39チーム中3位の結果を残すことができました。今後は、本研究結果に基づき、新たなQuality評価法の開発を進めていきたいと考えています。

これらの研究を進めるにあたりご指導を賜りました竹田-志鷹真由子教授、並びに、北里大学 薬学部生物分子設計学教室の皆様に心より感謝申し上げます。今回このような栄誉ある賞を頂 き、ありがたく思うと同時に、発表時に皆様から頂いた貴重な意見を吸収し、今後の研究に取り 組んで参りたいと存じます。

KP-25

タンパク質複合体立体構造予測における Quality 評価法の予測精度の比較検証

 ○原田祥季、清田泰臣、竹田一志鷹真由子 (北里大・薬)
 E-mail: shitakam@pharm.kitasato-u.ac.jp

1. 背景と目的

創薬研究において、タンパク質間相互作用を立 体構造の上で理解することは重要である。しかし ながら、実験構造が解明されていないタンパク質 複合体は数多く存在しており、コンピュータを用 いたタンパク質複合体の立体構造予測(複合体モ デリング)の重要性が増している。複合体モデリ ングは、単体分子の構造予測、相互作用部位の予 測、分子の相対配置の予測、相互作用界面の構造 予測など複雑なステップからなっており、多数の 複合体モデルの候補構造が発生する。したがって、 モデル構造候補群の中から質の良いモデル構造 を選び出す Quality 評価というステップが重要な 役割を果たす。そこで本研究では、複合体モデリ ングにおける Quality 評価に焦点を当て、Quality 評価法の予測精度の比較検証を行った。

2. 方法

本研究では「4 つの Quality 評価法」 ① Rosetta score [1], 2 VoroMQA [2], 3 SOAP-PP [3], 4 SOAP-Protein-OD [3] の予測精度を比較検証した。 検証には、タンパク質立体構造予測の国際コンテ ストである CASP13 (2018 年開催)の Assembly 部門(複合体モデリング部門)[4] で出題された 42 個のターゲットを用いた。それぞれのターゲッ トにおいて、CASP13 のホームページで公開され ているモデル構造候補群に対して Quality 評価法 を適用し、最も質が良いと予測されるモデル構造 を選択した。それぞれの Quailty 評価法が選択し たモデル構造を実験構造と比較し、より実験構造 に類似したモデル構造を選択できた Quality 評価 法を優れた Quality 評価法とした。モデル構造と 実験構造との比較においては、(i) Jaccard score [5] (相互作用部位の一致度)、(ii) F1 score [5] (界面 のアミノ酸残基間コンタクトの一致度)、(iii) IDDT [6] (局所構造の類似度)、(iv) GDT-TS (全体 構造の類似度)の「4 つの評価基準」で類似度を 判定した。

3. 結果とまとめ

4 つの評価基準それぞれにおいて、各ターゲッ トでのベストなモデル構造(モデル構造候補群の 中で最も実験構造に類似したモデル構造)を真と したときの ROC 曲線の AUC を計算し、4 つの Quality 評価法の評価基準ごとの予測精度を比較 検証した。全 42 ターゲットの評価基準ごとの AUC の平均値を求めたところ、(i) Jaccard score と (ii) F1 score では ③ SOAP-PP が、(iii) IDDT と (iv) GDT-TS では ② VoroMQA が優れていること が分かった。

また、42 個のターゲットを構造予測の難易度 (Easy、Medium、Hard)で分類し、評価基準ごと のAUCの平均値を難易度別に求めた。その結果、 全 42 ターゲットでの比較検証とは異なる傾向が みられた。したがって、ターゲットの難易度など の条件によって、これらの Quality 評価法を使い 分けることで、それぞれを単独で用いるよりも予 測精度が高くなることが示唆された。

本研究での検証結果は、複合体モデリングの 予測精度の向上につながるものと考えられる。参 考までに、CASP13 Assembly 部門に参加した研究 グループのモデリング結果と 4 つの Quality 評価 法の仮想的な比較も行った。

4. 参考文献

1. Rebecca F. Alford, et al. J. Chem. Theory Comput 13, 3031–3048 (2017).

2. Olechnovic K, et al. *Proteins* **85**, 1131–1145 (2017).

3. Dong G.Q, et al. *Bioinformatics* **29**, 3158-3166 (2013).

4. Dmytro Guzenko, et al. Proteins 87, 1190-1199 (2019).

5. Aleix Lafita, et al. Proteins. 86, 247-256 (2018).

6. Mariani V, et al. *Bioinformatics* **29**, 2722–2728 (2013).

///// Activities /////

第48回構造活性相関シンポジウム

日時:令和2年12月10日(木)・11日(金) 会場:オンライン開催 主催:日本薬学会構造活性相関部会 協賛:情報計算化学生物学会・日本化学会・日本農薬学会 後援:日本農芸化学会

12月10日及び11日の2日間にわたって、第48回構造活性相関シンポジウムが開催された。 今回は、新型コロナウイルスの感染拡大を防ぐために、初めてオンラインによる開催となった。 メインの招待講演・ロ頭発表のセッションを Zoom によって、ポスターセッション及び懇親会は Remoを用いて実施した。

初めてのオンライン開催でトラブルの可能性が相当あること、及び企業の協賛・展示によって 開催費用の相当部分を賄える見込みであることから、こちらも初めての試みであるが、参加費は 無料とした。無料であることで参加のハードルが下がり、参加登録者は292人、最多同時接続数 は10日で180人程度、11日で170人程度と、最近数年の中では最も多い参加者に集まっていた だくことができた。

本シンポジウムは、低分子創薬の停滞を打破するためのシミュレーション、AI (artificial intelligence) による定量的予測をテーマとして開催した。このテーマに沿って、招待講演は、AI の創薬応用の専門家である横浜市大の寺山先生、住友化学株式会社において量子化学計算による 分子設計を行っている原田先生、京大で京、富岳を用いたシミュレーションに基づく予測を行っている荒木先生にお願いした。若手でこの分野をリードする先生方からの講演は大変盛り上がり、活発に質疑応答が行われた。

オンライン開催の様子について記す。Zoomはすでに多くの学会・セミナーでも使われており、 今回も大きなトラブルなく招待講演・ロ頭発表セッションを実施することができた。他の学会で は発表終了後に座長以外との質疑応答や拍手がなく寂しいという状況が見られたことに対する 改善策として、座長、座長補佐はパネリスト資格として自分の担当セッション以外の際にも質問 を積極的に行うことと、拍手をすることを奨励した。



図 1. Zoom のメインセッションにおける実行委員及び座長による拍手の様子

Remoは、テーブル毎に発表やホワイトボードでの議論ができるため、ポスターセッションや 懇親会に利用した。他の学会における利用実績が少ないソフトウェアであり、トラブルを心配し たが、一部のネットワーク接続状態の悪い参加者以外では概ねトラブルは無い様子であった。 参加者からアンケートを実施したところ、44人から回答があった。5段階で評価してもらった ところ、ロ頭発表・招待講演が平均4.56、ポスターが4.32、懇親会が4.03と、概ね「非常に良 い」と「良い」の間の評価を得ることでき、好評であった。



図 2. ポスターセッションと懇親会に利用した Remo の画面

SAR Award の審査結果についてもオンラインで集計し、2日目の昼前には集計を締め切り、受 賞者を決定した。以下に受賞者を紹介する。今年は、学生による優秀なロ頭発表、ポスターが多 く、受賞者5名のうち、4名を学生が占めた。

- KO-01 口頭笠原慶亮(かさはらけいすけ)東京大学大学院修士1年計算デザイン Supercharging 抗体の物性機能解析
- KO-06 ロ頭 千葉峻太朗(ちばしゅんたろう) 理研 MIH研究員 幾何学的相互作用解析と機械学習を組合せた抗体の側鎖構造予測
- KP-13 ポスター 髙橋真帆(たかはしまほ) 横浜市立大学大学院 修士1年 ニューロトロフィン受容体 TrkAd5 と結合ペプチド間の相互作用様式の解明
- **KP-21** ポスター 伊藤朱里(いとうあかり) 横浜市立大学大学院 修士1年 中分子シクロスポリンAとシクロスポリンEの分子ダイナミクスの比較
- KP-25 ポスター 原田祥季(はらだよしき) 北里大学大学院 修士1年 タンパク質複合体立体構造予測における Quality 評価法の予測精度の比較検証



図 3. SAR Award 受賞式における受賞者の様子

一般参加者の音声による質疑や、1対1のコミュニケーションなどまだ改善点はあるが、初め てのオンライン開催にしては、成功と言ってよいと考えている。事務局の幸研究員、高谷研究員、 実行委員会の皆さん、構造活性相関部会の幹事、常任幹事の皆さん、協賛・展示いただいた企業 の皆さん、招待講演を快く引き受けてくださった先生方、発表者・一般参加者の皆さんにこの場 を借りて深く感謝したい。

第48回構造活性相関シンポジウム

実行委員長 本間 光貴

///// Activities /////

<会告>

構造活性フォーラム 2021

「次期スーパーコンピュータ「富岳」時代の計算創薬」

主催: 日本薬学会構造活性相関部会

会期: 2021年6月4日(金)

会場: Zoom によるオンライン開催

フォーラムホームページ: http://www.qsarj.org/forum2021

開催趣旨:2021年3月に、日本のスーパーコンピュータのフラッグシップ「富岳」の運用が開始になった。「富岳」は「京」に比して、ソフトウエア性能で100倍以上の性能向上が謳われている。「富岳」のような膨大な計算能力を手にしたとき、計算創薬技術はどのような変革を迎えるだろうか。さらに、「富岳」では、これまでのシミュレーション研究のみならず、人工知能(AI)技術も取り込んだ研究進展が期待されている。深層学習などのAI技術や大規模データ解析が、どのように計算創薬を変革させるだろうか。そこで、本フォーラムでは、今後、実現が期待される、スーパーコンピューティングを用いた計算創薬および、計算創薬におけるAI展開など、次世代の計算創薬について議論したい。

- プログラム:
 - 講演1.「分子動力学シミュレーションによる蛋白質の動的構造と基質結合:分子混雑環境の影響とSARS-CoV-2スパイク蛋白質に関する計算分子混雑環境における蛋白質と 阻害剤の結合と相互作用」

杉田 有治 (理化学研究所)

講演 2. 「インシリコシミュレーションで探るタンパク質複合体の結合親和性とキネティックス」

北尾 彰朗(東京工業大学)

講演 3. 「分子動力学シミュレーションを用いた抗体設計の可能性」

山下 雄史 (東京大学)

講演 4. 「武田薬品における創薬化学研究への分子動力学計算の活用」

高木 輝文 (武田薬品工業株式会社)

講演 5.「創薬研究における化合物生成モデルの現状と課題」

小倉 圭司(塩野義製薬株式会社)

講演 6. 「溶液理論で得たタンパク質水和の包括的解析」

吉留 崇 (東北大学)

参加登録および申込締切日: 5月21日(金)までに、フォーラムホームページから事前参加登録をお願いいたします。参加人数が上限に達しましたら、参加登録を打ち切らせていただきますのでご了承ください。

- 参加費: (薬学会会員・非会員ともに)無料
- 問合先: 構造活性フォーラム 2021 実行委員会 池口 満徳(実行委員長)
 〒230-0045 横浜市鶴見区末広町 1-7-29 横浜市立大学
 Tel: 045-508-7232 Fax: 045-508-7367 E-mail: ike@yokohama-cu.ac.jp

部会役員人事 2020年度 常任世話人 2021/3/31 現在 大田雅照 (理化学研究所) 部会長 副部会長 服部一成(塩野義製薬(株)) 副部会長 本間光貴 (理化学研究所) 会計幹事 川下理日人 (近畿大学) 竹田-志鷹 真由子(北里大学) 庶務幹事 広報幹事 加藤博明 (広島商船高等専門学校) SAR News 編集長 幸 瞳 (理化学研究所) ホームページ委員長 高木達也(大阪大学) 2021年度部会役員人事は後日ホームページにてお知らせします。

構造活性相関部会の沿革と趣旨

1970年代の前半、医農薬を含む生理活性物質の活性発現の分子機構、立体構造・電子構造の計算や活性 データ処理に対するコンピュータの活用など、関連分野のめざましい発展にともなって、構造活性相関と 分子設計に対する新しい方法論が世界的に台頭してきた。このような情勢に呼応するとともに、研究者の 交流と情報交換、研究発表と方法論の普及の場を提供することを目的に設立されたのが本部会の前身の構 造活性相関懇話会である。1975年5月京都において第1回の「懇話会」(シンポジウム)が旗揚げされ、1980 年からは年1回の「構造活性相関シンポジウム」が関係諸学会の共催の下で定期的に開催されるようになっ た。

1993年より同シンポジウムは日本薬学会医薬化学部会の主催の下、関係学会の共催を得て行なわれるこ ととなった。構造活性相関懇話会は1995年にその名称を同研究会に改め、シンポジウム開催の実務担当グ ループとしての役割を果すこととなった。2002年4月からは、日本薬学会の傘下組織の構造活性相関部会と して再出発し、関連諸学会と密接な連携を保ちつつ、生理活性物質の構造活性相関に関する学術・研究の 振興と推進に向けて活動している。現在それぞれ年1回のシンポジウムとフォーラムを開催するとともに、 部会誌のSAR Newsを年2回発行し、関係領域の最新の情勢に関する啓蒙と広報活動を行っている。本部会 の沿革と趣旨および最新の動向などの詳細に関してはホームページを参照頂きたい。 (https://sar.pharm.or.jp/)

編集後記

今号では公開時点においても未だに猛威を振るう SARS-CoV-2 をターゲットとした創薬について、4 人 の先生にご寄稿をお願いしました。Perspective/Retrospective では、理化学研究所の関根俊一先生から SARS-CoV-2 の創薬ターゲットの中でも特に複雑な機構をもつ RdRp について丁寧にご紹介いただきまし た。Cutting Edge では、株式会社 Elix の結城伸哉先生、Nazim Medzhidov 先生からいち早く発表されていま したドラッグリポジショニング研究について、理化学研究所の高谷大輔先生から PDB に登録されている SARS-CoV-2 ターゲットや FMO 計算を活用した解析結果について、それぞれご紹介いただきました。 COVID-19 関連の研究をとおして、研究で成果を出すスピードはもちろん大事ですが、大手医学誌で COVID-19 に関する論文が撤回されたこともありますし、研究対象となる事象、生物種やタンパク質、自 分たちが用いる手法についての理解も決して疎かにしてはいけないと強く感じます。ご多忙の中、快くご 執筆していただいた先生方に深く感謝申し上げます。昨年 12 月に開催した構造活性相関シンポジウムのご 報告および6月の構造活性フォーラムの会告も掲載しておりますので、お目通しいただければ幸いです。 (編集委員会)

SAR News No.40 2021 年 4 月 1 日

発行:日本薬学会 構造活性相関部会長 大田雅照

SAR News 編集委員会

(委員長)幸瞳、河合健太郎、清田泰臣、合田浩明、田上宇乃、仲西功

*本誌の全ての記事、図表等の無断複写・転載を禁じます。