

タンパク質立体構造解析



デイヴィッド・ベイカー
(ワシントン大)
© Nobel Prize Outreach.
Photo: Clément Morin



デミス・ハサビス
(Google DeepMind)
© Nobel Prize Outreach.
Photo: Clément Morin



ジョン・ジャンパー
(Google DeepMind)
© Nobel Prize Outreach.
Photo: Clément Morin

2024年、ノーベル化学賞と物理学賞が Artificial Intelligence (AI) 技術に——。10月9日、AIを用いたタンパク質構造予測と設計に革新をもたらした3人の科学者、米・ワシントン大のデイヴィッド・ベイカー (David Baker)、Google DeepMind¹ のデミス・ハサビス (Demis Hassabis)、同社のジョン・ジャンパー (John M. Jumper) にそれぞれノーベル化学賞が授与された [1]。

「生命の維持に不可欠であるタンパク質の立体構造を正確に予測し、新規のタンパク質を設計する能力が確立すれば、人類は計り知れない利益を享受することになる」との選考理由である。

折しもその前日に発表されたノーベル物理学賞受賞者は、「人工ニューラルネットワークによる機械学習を可能にした基礎的発見と発明」で今日の深層学習 (DeepLearning; DL) 技術の基礎を築いた米・プリンストン大のジョン・ホップフィールド (John Hopfield)、カナダ・トロント大のジェフリー・ヒントン (Geoffrey Hinton) であった。

2024年のノーベル賞はさながらAIの年となった。

¹ 英・アルファベット社 (Alphabet) の子会社。2010年に創業したディープマインドテクノロジー社 (DeepMind Technologies) が2014年にグーグル社 (アルファベット社が親会社) に買収され、2023年4月に現在の名称となった。

1. タンパク質の構造予測を革新するAI技術

物質や材料の「機能」は「構造」と密接に関連している。タンパク質の機能もまたその立体構造に依存しており、その構造を正確に解明することは、**生物学的な機能と疾患メカニズム**を解明するために不可欠である。今回ノーベル化学賞を受賞した3人は、タンパク質構造の設計や予測を革新する以下のAI技術を創製した。

(1) Rosetta

1990年代末、ベイカーらは、タンパク質を構成するアミノ酸の配列からエネルギー的に安定な立体構造を予測し、目的の機能を持つ新規タンパク質を設計するソフトウェア「**ロゼッタ (Rosetta)**」の開発を始めた。2004年には、自然界に存在しないタンパク質「Top7」の設計に成功したことを報告 [2]。この技術によって、特定の分子と結合する人工抗体や環境中の有害物質の分解酵素など、様々な用途に特化したタンパク質を設計できるようになった。COVID-19 ワクチン開発においても、ウイルスの表面タンパク質を模倣した人工タンパク質の設計につながった。

(2) AlphaFold

ハサビスとジャンパーらはタンパク質構造予測モデル **AlphaFold (AF)** を開発した。DLを基盤とするこのモデルの登場は、『**タンパク質の折りたたみ問題**』 (『タンパク質の折りたたみ (フォールディング)』次頁) と呼ばれる50年来の課題をほぼ解決した』として注目を集め、バイオテクノロジーや創薬分野への応用が急速に進んでいる。

2万報を超えてなお引用されるAlphaFold 2論文

表1に“AlphaFold”をタイトルに含む主な論文を示す。ハサビスらの論文 (灰色帯) が上位に複数報ある。中でもAFの改良版であるAlphaFold 2 (AF2) の詳細技術と成果をまとめた「AFによる高精度なタンパク質構造予測 (Nature, 2021)」 [3] は、2025年1月時点で**約2万報の論文から引用され** (図1)、それらの論文の研究分野は「生化学・遺伝学・分子生物学」が約3割を占めるほか、「化学」「医学」「コンピュータ科学」など多岐にわたっている (図2)。

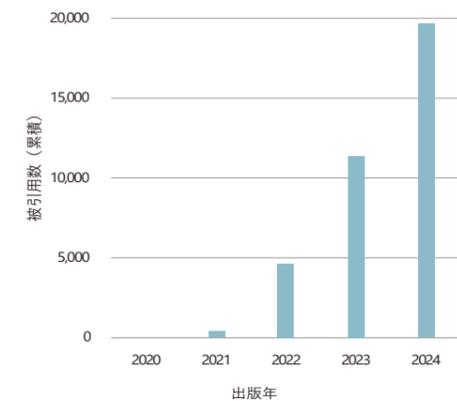


図1 AF2論文 (Nature, 2021) [3] の被引用数 (累積) Scopus 2025.2.4

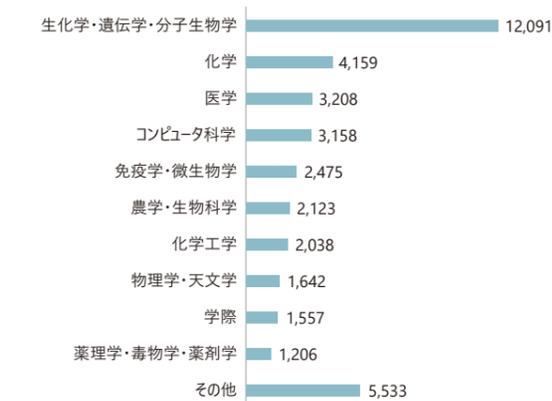


図2 AF2論文 (Nature, 2021) [3] を引用する論文の分野 Scopus 2025.2.4

表1 “AlphaFold” をタイトルに含む論文 (被引用数上位20報および日本著者上位6報)

Scopus 2025.1.3

被引用数	内容	掲載誌	掲載年	著者 (責任)	著者 (責任)
20,316	AFによる高精度なタンパク質構造予測	Nature	2021	J. Jumper D. Hassabis	DeepMind
4,089	AFタンパク構造DB: タンパク質配列空間の構造網羅率を大幅拡大	Nucleic Acids Res.	2022	S. Velankar D. Hassabis	欧・バイオインフォマティクス研究所 DeepMind
1,158	AF3による生体分子相互作用の正確な構造予測	Nature	2024	D. Hassabis J.M. Jumper M. Jaderberg	Google DeepMind Google DeepMind 英・アイソモーフックラボ
308	AFと天然変性タンパク質への影響	J. Mol. Biol.	2021	R.V. Pappu	セントルイス・ワシントン大
244	CASP14でのAFの適用と改善	Proteins Struct. Funct. Bioinformatics	2021	J. Jumper D. Hassabis	DeepMind
228	CASP13におけるAF	Bioinformatics	2019	M. Alquraishi	ハーバード大医学大学院
227	2024年時点のAFタンパク構造DB: 2.14億超のタンパク質配列構造をカバー	Nucleic Acids Res.	2024	R.P. Joosten; A. Perrakis	蘭・がん研究所
215	AlphaFill: リガンドとコファクターでAFモデルをリッチにする	Nat. Methods	2023	J. Jumper D. Hassabis	Google DeepMind 欧・バイオインフォマティクス研究所
161	多様なヘテロ二量体タンパク質複合体をベンチマークとしてAFの精度を試験	Protein Sci.	2022	A. David	インペリアルカレッジロンドン
159	AF Database of Protein Structures: 生物学者向けガイド	J. Mol. Biol.	2022	B.G. Pierce	メリーランド大
140	AF2: なぜ機能するのか、タンパク質の配列・構造・機能の関係理解への影響	J. Chem. Inf. Model.	2021	J. Skolnick	ジョージア工科大
140	AFによる原子精度でのタンパク質構造予測	Nat. Methods	2022	J. Jumper D. Hassabis	DeepMind
129	生物学におけるAI革命: AFがもたらす喜びと危うさ	EMBO Rep.	2021	T.K. Sixma	蘭・がん研究所
121	AFがAI創薬を加速: 新規CDK20低分子阻害剤の効率的な発見	Chem. Sci.	2023	A. Aspuru-Guzik A. Zhavoronkov	トロント大 インシリコメディシン上海
110	AFが告げる生物学と医学におけるデータ駆動型の革命	Nat. Commun.	2022	J.M. Thornton	欧・分子生物研究所
105	AFとモンテカルロ木探索を用いた大規模タンパク質複合体の構造予測	Mol. Syst. Biol.	2022	P. Bryant	スウェーデン・生命のための科学研究所
105	AFによる抗菌薬探索のための分子ドッキング予測ベンチマーキング	Nat. Med.	2021	J.J. Collins	MIT
102	統合クライオ電子顕微鏡とAFによる核膜孔複合体の細胞質リング構造	Science	2022	H. Wu	ハーバード大医学大学院
99	AFを用いて単一の突然変異がタンパク質の安定性と機能に与える影響を予測する	PLoS ONE	2023	D.N. Ivankov	露・スコルボ科技大
94	AFとAIタンパク質フォールディング革命の次なる展開	Nature	2022	E.Callway	記載なし
:	:	:	:	:	:
23	AFとAbAdaptを用いた抗体特異的エピトープ予測の改善	ChemBioChem	2022	ダロン・スタンダー	阪大微研
12	溶解性を考慮したタンパク質結合ペプチドをAFで設計	Biomedicines	2022	小杉孝嗣 大上雅史	東工大
9	タンパク質進化の統計的傾向: AFデータベースからの教訓	Mol. Biol. Evol.	2022	唐乾元; 金子邦彦	理研 東大; コペンハーゲン大
8	AF2によるVIPR2とKS-133の間の結合メカニズムとVIPR2機能に重要なS-S結合 (Cys25 - Cys192) 形成の解明	Biochem. Biophys. Res. Commun.	2022	坂元孝太郎 広川貴次	一丸ファルコス 筑波大
8	タンパク質間相互作用を標的とする環状ペプチドをAFで設計	Int. J. Mol. Sci.	2023	大上雅史	東工大
4	SpatialPPI: AFd Multimerによる三次元空間タンパク質間相互作用予測	Comput. Struct. Biotechnol. J.	2024	大上雅史	東工大

2. タンパク質構造解析手法

すべての生物の細胞中に存在するタンパク質。「酵素としての触媒作用」「細胞構造の維持」「シグナル伝達」「分子運搬」など、生体内のほぼすべての生命活動に関与している。このタンパク質は、20種類の「標準アミノ酸」がペプチド結合により縮合重合したポリペプチド鎖である (『タンパク質の立体構造』)。遺伝情報に基づいて合成されたアミノ酸配列 (一次構造) は、特定の立体構造へと「折りたたまれ

ること」で機能を発揮する。

無数の原子間相互作用や周囲の化学的環境の影響を受けながら、タンパク質は緻密に、ときに曖昧に自己組織化している。本来はエネルギー的に最適な構造へと折りたたまれるが、細胞内の環境 (温度, PH, イオン濃度) などによっては、不安定な領域も随所に散在し得る [4]。

タンパク質の「折りたたみ」構造を解析すべく様々な実験手法が試みられており、現在、**X線結晶構造解析法**に加え、**核磁気共鳴分光**

No. 25 AI × 医療シリーズ

タンパク質立体構造解析

法（Nuclear Magnetic Resonance; NMR）、クライオ電子顕微鏡（Cryo-EM）解析法が3主流である（表2）。

(1) X線結晶構造解析法

1950年代、結晶化したタンパク質に照射したX線の回折光を測定して電子密度マップを構築することで、立体構造を導出するX線結晶構造解析法が確立された。タンパク質立体構造データベース（Protein Data Bank; PDB、[PDB](#)）「構造生物学者が構築するタンパク質構造データベース」に蓄積されている膨大なタンパク質構造の多くは、このX線結晶構造解析法により決定されたものである。

(2) 核磁気共鳴分光法（NMR）

NMRは、強磁場下におかれた試料に電磁波を照射し、特定の原子核スピンの応答して放出するNMR信号を検出することで、分子構造に関する情報を得る。この手法の強みは、結晶化を必要とせず、溶液中におけるタンパク質の構造、さらには分子の動的挙動や相互作用に関する情報をも提供できる点にある。

(3) クライオ電子顕微鏡（Cryo-EM）解析法

Cryo-EM解析法は、特に巨大なタンパク質複合体や結晶化が困難な試料の解析に適している。試料を結晶化せずそのまま急速凍結して観察できることは、結晶化が必要なX線結晶構造解析法に対する利点である。さらに、近年は、技術革新により大幅に解像度が向上（2.0Å以下[8, 9]）したこともあって単なる補助手法から、構造解析の主役のひとつへと躍進を遂げつつある。

表2 代表的なタンパク質の構造解析法（実験手法）

	X線結晶構造解析法	核磁気共鳴分光法(NMR)	クライオ電子顕微鏡解析法
原理	結晶化した試料にX線を照射し、得られた回折パターンから電子密度分布を計算することで、原子レベルの3D構造モデルを構築	強磁場中で電磁波を照射された原子核が吸収・放出するNMR信号を解析し、距離情報や立体配座を導出することで、3D構造を推定	試料を急速凍結し、電子顕微鏡により高分解能構造情報を取得。2D元画像をコンピュータで解析・統合することにより、3D構造を再構成
結晶化	必要 ^{*1}	不要（溶液状態で解析）	不要（急速凍結）
適用範囲	小型～中型タンパク質、結晶化可能な複合体（解像度1.0Å以下）	小型～中型タンパク質、動的構造の解析	中型～大型タンパク質、大型タンパク質複合体、膜タンパク質、ウイルス（解像度2.0Å以下 ^{*3} ）
その他	現在PDBに蓄積されている構造の多くはこの手法で決定された	動的挙動や相互作用を観察可能 非破壊分析	生理的条件に近い状態で解析 コンフォメーション、動的構造にも適用可能
課題	結晶化困難なタンパク質には不向き 静的構造のみに対応 ^{*2}	高濃度試料の調製が困難な場合あり 同位体標識が必要 スペクトル解析が複雑	試料の放射線損傷リスク 小さな分子の解析 装置や解析技術が高価

^{*1} 結晶スポンジ法など結晶化が不要な技術も開発されている[10]。

^{*2} X線自由電子レーザー(X-ray free-electron laser; XFEL)によるタンパク質酵素の動的構造機能解析なども行われている[11]。

^{*3} 近年大幅に高解像度化が進んでいる。Electron Microscopy Data Bank(EMDB)にCryo-EMの解像度統計が提供されている[12]。

タンパク質の折りたたみ（フォールディング）

アミノ酸が直鎖状に並んだポリペプチド鎖（一次構造）は自発的に折りたたまれ、熱力学的に最も安定した三次構造（あるいは複数鎖からなる四次構造）を形成する。この過程は自己組織化によって進行し、タンパク質は固有の立体構造を獲得することで初めて本来の機能を発揮できるようになる。折りたたみに失敗した場合、アルツハイマー病や鎌状赤血球症などの難治性疾患の一因となることもある。したがって、「アミノ酸配列から最終的な立体構造がどのように決定されるかを予測すること」こそが「タンパク質の折りたたみ問題」の核心的な課題として長年にわたり追究されてきた。

1961年、NIHのクリスチャン・アンフィンセン（Christian Anfinsen）は「タンパク質の天然構造は自由エネルギー最小の状態であること」を示す実験結果を発表した。リボヌクレアーゼ酵素の再構成実験を通じて、タンパク質のアミノ酸配列がその立体構造と生物学的活性を決定するという原理を明確に示したのである。この理論は「アンフィンセンのドグマ」として現在も知られている[5, 6]。

1969年、コロンビア大のサイラス・レヴィンソール（Cyrus Levinthal）は、タンパク質が膨大な構造選択肢の中から、どのようにして極めて短時間で正しい構造に到達するのかという点に注目し、「レヴィンソールのパラドックス」として知られる問題を提起した。彼の計算によれば、もしすべての構造を無作為に試すのであれば、タンパク質が正しい構造を見つけるには宇宙の年齢を超える時間が必要になるはずである。が、実際には短時間に折りたたみが行われる。この矛盾から、「タンパク質はランダムな探索ではなく、何らかの効率的かつ決定的なメカニズムに従って折りたたまれているに違いない」との仮説を提示した[7]。この提言は折りたたみ問題の本質を一段と浮き彫りにし、研究者たちの関心を「構造の安定性」から「折りたたみ経路の力学」へと導く転換点となった。

タンパク質の立体構造

タンパク質は、生物の細胞や組織を構成し、酵素やホルモンなど生命活動のあらゆるプロセスを支える基本的な高分子化合物である。その機能的特性は、アミノ酸配列（一次構造）だけでなく、分子全体の「立体構造」に強く依存する。

1951年、ライナス・カール・ポーリングらは「一次構造」からαヘリックスやβシートなどの「二次構造」が形成され、これらが折りたたまれることで「三次構造」が形成されると予測した。現在、タンパク質の構造は次の4つの階層的レベルに分類される[13]。

一次構造（Primary Structure）

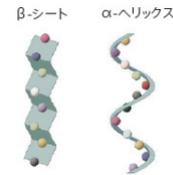
タンパク質を構成する20種類の基本アミノ酸が、特定の順序で直鎖状につながったポリペプチド鎖。この配列順序が、タンパク質の「設計図」となってその特性や機能が決まる。ここで、一つ一つのアミノ酸を「残基（residue）」という。これは、縮合反応により1分子の水を失った状態を指す。



二次構造（Secondary Structure）

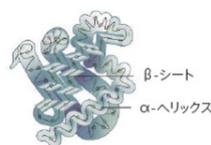
ポリペプチド鎖は、ペプチド結合部分でカルボニル基 (>C=O) とア

ミノ基 (H-N<) の間に形成される水素結合により安定化する。その結果、αヘリックス（らせん状）、あるいはβシート（波状）などの構造で「折りたたまれる」。強度や柔軟性といった物理的特性は、この二次構造によって生まれる。



三次構造（Tertiary Structure）

側鎖同士の疎水性相互作用、イオン結合、ファンデルワールス力、さらにはシステインの側鎖間のジスフィド結合などによって、ポリペプチド鎖はさらに複雑に折りたたまれ、特有の立体構造を形成する。この三次構造こそが、本質的にタンパク質の「機能」にかかわる。



四次構造（Quaternary Structure）

複数のポリペプチド鎖（サブユニット）が非共有結合を介して集合することで形成される機能的複合体。例えば、ヘモグロビンは4つのサブユニットからなる複合体で、酸素を効率的に運搬する。他にも、RNAポリメラーゼやイオンチャンネルなど、多くの重要な生体分子がこのレベルの高次構造を持つ。



構造生物学者が構築するタンパク質構造データベース

構造生物学者が実験的に決定した原子座標と実験情報、およびそれに紐づくメタデータを中核とするProtein Data Bank (PDB)は、1971年に米国でスタートした老舗のデータベースである。米・ブルックヘブン国立研究所と英・ケンブリッジ結晶学データセンターにより共同で創設された。電子化された生体高分子立体構造データのアーカイブを“PDB”と称したのが始まりである[14]。日本では、1979年から阪大蛋白質研究所で継続的にデータ配布が行われてきた。

米日欧三極で運営

1998年までは、主にブルックヘブン国立研究所が管理しており、米・エネルギー省（Department of Energy; DOE）などの支援を受けて、世界中から集まったデータの登録と編集を行っていた。1990年代後半にPDBへの登録件数が飛躍的に増加したため（[「22万を超えたタンパク質立体構造データベースへの蓄積」](#)次頁）、阪大にPDBj、欧州バイオインフォマティクス研究所（European Bioinformatics Institute; EBI）にPDBeがそれぞれ設置された。

1999年、ブルックヘブン国立研究所からラトガー大を中心とするバイオインフォマティクス共同機構（Research Collaboratory for Structural Bioinformatics; RCSB）へと移管され、RCSB PDB、PDBj、PDBeの三極体制が確立した。また、PDBjは2000年にRCSB PDBと連携し、

構造データの受付を試験的に開始、2001年から正式に登録拠点として運用が始まった。2003年、PDBアーカイブの国際的かつ持続的な運営を保障し、世界中の誰もが無料で利用できる体制を維持するため、国際組織worldwide Protein Data Bank（wwPDB）が設立された。

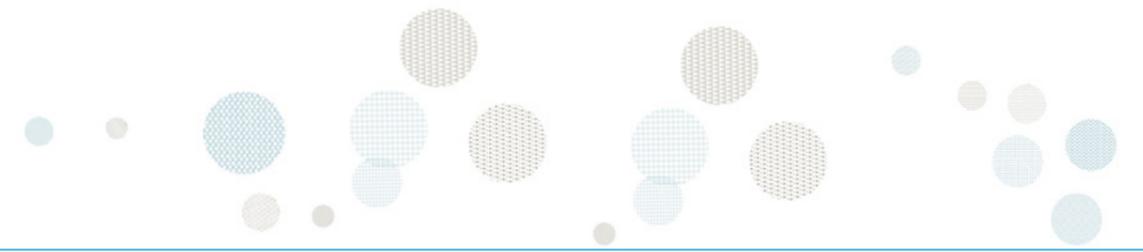
wwPDBの設立から20年となる2023年、PDBjに登録されたデータ件数は全世界の約4分の1（5万件）を突破し、日本の拠点としての存在感を確かなものとしている。

現在では、NMR、クライオ電子顕微鏡で決定された構造データにも対応した共通の登録・編集システムが整備されており、日米欧それぞれの拠点で同一の仕組みに基づいてPDBアーカイブが維持・更新されている。

“It’s humbling every time we train [AlphaFold] on years of effort. Each data point is years of effort from someone.”

AFもまたPDBから大きな恩恵を受けている。「AFにタンパク質の立体構造データを学習させる度に、いつも我々は謙虚な気持ちになる。どのデータも実験研究者たちが蓄積してきた成果だからだ」。ジャンパーはPDBをこう讃えた[15]。

「AFによってタンパク質構造決定の実験が効率化し加速する」、「PDB登録データがますます充実する」、そして「AFの予測精度が高まっていく」。このような「好循環」が実現するように、実験と予測が両輪となって駆動する形を希求していきたい。



22 万を超えたタンパク質立体構造データベースへの蓄積

実験的に決定された立体構造のうち、信頼性が高く、かつ公開が許可されたものは PDB に登録されている。その数は 2025 年 1 月時点で 22 万件を超えた (図 3)。

1990 年代後半以降、登録件数が急増している。その原動力のひとつは、日米欧で放射光施設を利用した構造ゲノムプロジェクトが立ち上がったことである。また、ネイチャー誌など主要な国際学術雑誌が、タンパク質構造を扱う論文の掲載に際して、wwPDB への構造登録を投稿条件の一つとして義務づけたことも、登録数の増加に拍車をかけている。

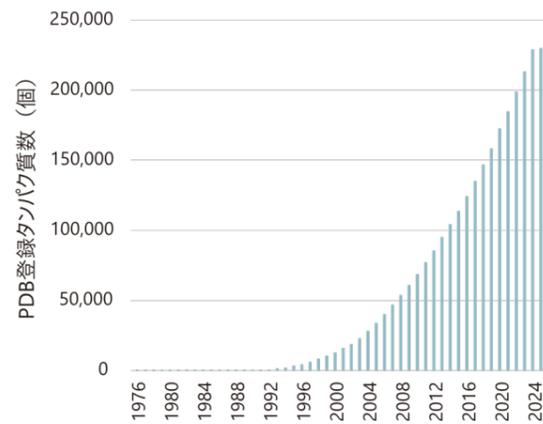


図 3 PDB の登録タンパク質数の推移

<https://pdj.org/info/statistics?lang=ja>

●●● References

- [1] The Nobel Prize in Chemistry 2024, <https://www.nobelprize.org/prizes/chemistry/2024/popular-information/>
- [2] Kim, D.E. et al., Nucleic Acids Res, 32, Web Server issue, pp. W526-531, 2004
- [3] Jumper, J. et al., Nature, 596, 7873, pp. 583-589, 2021
- [4] Callaway, E., Nature, 588, 7837, pp. 203-204, 2020
- [5] Anfinsen, C.B. et al., Proc Natl Acad Sci U S A, 47, 9, pp. 1309-1314, 1961
- [6] Haber, E. and Anfinsen, C.B., J Biol Chem, 236, pp. 422-424, 1961
- [7] Levinthal, C., Journal de chimie physique, 65, pp. 44-45, 1968
- [8] Renaud, J.P. et al., Nat Rev Drug Discov, 17, 7, pp. 471-492, 2018
- [9] Nakane, T. et al., Nature, 587, 7832, pp. 152-156, 2020
- [10] 東京大学工学系研究科・工学部, 結晶スポンジ法 結晶化不要の X 線結晶解析, 2013, https://www.u-tokyo.ac.jp/focus/ja/features/f_00051.html
- [11] 大阪大学 講師 溝端 栄一 (工学研究科 応用化学専攻), X 線自由電子レーザー (XFEL) を用いたタンパク質・酵素の動的構造機能相関の解明, <https://sdgs.osaka-u.ac.jp/research/5821.html>
- [12] EMBL-EBI, EMDb entry resolution in shells per year, https://www.embl.ac.uk/emdb/statistics/emdb_resolution_year
- [13] 竹内敬人ら, 化学, 東京書籍, 2016,
- [14] 中村春木, 栗栖源嗣, 生物物理, 58, 2, 71, 2018,
- [15] Callaway, E., Nature, 634, 8034, pp. 525-526, 2024