

No.	質問	回答
	古井講師への事前質問FAQ2/n[(参照)講義資料:スライド48] (「Boltz-1で複数のポリペプチド鎖に対する構造予測の方法 (MSAの準備方法)が分かりませんでしたので、ご紹介ください。 localcolabfoldで複数ポリペプチド鎖に使うa3mを流用しましたが、動作しませんでした。」)のご回答について 追加でお伺いしたいことがあります。	json形式のmsaというよりは、jsonの中身に入っているa3mのファイルを書き出し、それを指定する方法となっています。
1	「解決策2:localcolabfoldで単量体として予測し、a3mファイルをそれぞれMSAとして指定する(unpaired非対応)」も試してみたのですが、1つの蛋白質しか立体構造が発生しない形でした。「解決策3:alphafold3_toolsで個々のMSA情報を抽出し、ファイルに保存する」を試したいと考えますが、Boltz-1はjson形式のmsaも利用可能という理解でよろしいでしょうか？その場合のyamlファイルでの記載例もご紹介ください。 また、解決策3ではBoltz-1の推論コードも修正する必要があるということでしょうか。	以下の最後のセクションでこの処理に関するハンズオンを追記しました。 https://github.com/YumizSui/AJACS-PSP-2025/blob/main/notebooks/AJACS_Boltz_1.ipynb
2	最近ではLLMを基盤としてヒトの発想を超えるような科学を行おうとする動き(AI for Science など)があると思います。タンパク質立体構造予測の界限でも、人の発想にはないような予測結果から新しい知見が得られたような事例はありますか？	この辺は悩ましいところです。構造生物学的には、立体構造予測で、例えば分子置換をテンプレートで使っ てうまくとけたということもありますので、うまく使っている人は非常に多いです。しかし、人の発想にな いような予測結果となると、それがどのように活用されているかは、それぞれの分野次第です。LLMがす ごくキャッチーだったということもありますが、一つの道具なので、うまく使うということを、それぞれの業 界でやっていると思います。
3	AlphaFold2と3では、タンパク質構造の予測精度はどちらが良いのでしょうか？	明確にAlphaFold3が良いと思います。ただし、天然変性領域が多く含まれているタンパク質を予測する ときは、AlphaFold3は注意が必要です。
4	拡散モデルの利用により変わった点は具体的に何でしょうか？	AlphaFold2はアミノ酸のかたまりをベースに座標を決めていましたが、AlphaFold3では基本的には 原子単位で座標を決める点が大きな違いです。技術的には、Structure Module (AlphaFold2) と呼 ばれるニューラルネットワークが、AlphaFold3では拡散モデルベースに置き換わっています。
5	分子量 約 10 K(約 90 アミノ酸残基)のタンパク質が研究対象で す。相同性のあるタンパク質の構造についての報告はありません。 本タンパク質の配列を用いて、AlphaFold2 (正確には ColabFold) を半年ほど置いて2回試したところ、2回で全く異な る構造が予測されました。設定はデフォルトのままで、変更等はして いません。このような事例はよくあるのでしょうか？また、どうした ら正確性の高いと考えられる結果を得ることができるのでしょうか？	おそらくモデルとデータベースが異なっていることに起因していると思われます。各種データベースのアップ デートや、AlphaFold自体のモデルのアップデートが行われており、そこで違いが起きたのではないかと 思います。基本的には新しい実行結果のほうが信じられると思いますが、信頼度という点では微妙な差 なので、どちらがもっともらしいかを人間の目で検討する方が良いと思います。
6	既知のタンパク質と全く配列類似性が無いタンパク質が新たに見つ かった場合は、MSAが構築できないのでAlphaFold2-3では立 体構造が予測困難ということでしょうか？	MSAが一切なくてもある程度予測は可能です。de novoデザインタンパク質でも構造予測がうまくいっ た事例もあります。MSAが豊富に取れるタンパク質に比べると精度は劣ると思われます。
7	リン酸化についてAlphaFoldで解析できますか？	AlphaFold3のインプットに反映させることができます。しかし、その結果を本当に信じられるかは怪しい です。(講師は翻訳後修飾を専門としていないため、)翻訳後修飾を専門に扱っている方が使ってみたとき の感想が、より重要だと思います。
8	リン酸化されたタンパク質の構造予測はできますか？	同上
9	この領域で、計算量に関する課題はありますか？ 計算時間がかか りすぎて解けないような問題はどのような問題でしょうか？	あります。AlphaFold3は比較的速くなりましたが、AlphaFold2のときは、1個のタンパク質に数時間か かっていたので、計算を削減するということは重要な課題だと思います。AlphaFold3は、結果的に拡散 モデルベースのニューラルネットワークにしたことで、全体が割と小さいニューラルネットワークになり、計 算時間が速くなりました。速く計算できるということは重要です。特に、大量のメタゲノム配列に使いたい、 いろいろなタンパク質の相互作用の予測に使いたいといったときには、大量に予測することになるので、 計算時間には課題があります。
10	「1アミノ酸違いに弱い」とのことですが、逆に1アミノ酸変異で大き く構造が変わった場合、その信頼性についてはどう考えたら良いで しょうか。「それは意味がある！」という感じでしょうか。	予測結果が変わってもその結果に意味があるかはわかりません。ただ、タンパク質の残基コンタクトペアの 片方のアミノ酸が変わったときに、アラインメント上では保存されているはずなのに異なる残基になったと いうことに意味を見出して情報抽出されると思うので、AlphaFoldの中ではコンタクトが崩れた・相互作 用が減弱したとして構造が開くようになる、というような予測結果を返す可能性はあります。すなわち、変 異箇所が構造の相互作用によく寄与していた部分だったときには、そのような意味がある(かもしれない) 議論ができると思います。
11	コラーゲン研究者なのですが、歴史的には非常に古くから構造が知 られている三重鎖が似ても似つかない構造に予想されていて驚い たことがあります。細胞外タンパク質は苦手なのでしょうか？	天然変性タンパク質の構造予測は難しいです。コラーゲンもそうですが、例えばアルツハイマーに関係する ようなAβのようなタンパク質も、うまく予測できません。実際に実験的に取れている構造がPDBにある にも関わらず、予測結果がうまくいかない例もあります。それぞれケースバイケースです。信頼度となる pLDDTなどのスコアとともに判断することが大事です。細胞外タンパク質が必ずしも苦手ということでは ありません。
12	Blastでhitしてきたアミノ酸配列群の立体構造をAlphaFoldで予 測した際、pLDDTとPAEが高いものほどリコンビナントタンパク質 を作りやすいという傾向はありますか？	文献を確認したわけではないですが、そのような議論をしている論文があると聞いたことがあります。 pLDDTやPAEが高いほど「自然なタンパク質」らしいタンパク質なはずですので、リコンビナントも作りや すいという相関がある(?)ようです。
13	いろいろなアプリケーションで予測をしてどのアプリケーションでも スコアが高めな類似構造、という場合と、それぞれのアプリケーショ ンのトップスコアで構造が似ていない構造、どちらの構造をより信 頼すべきですか？	「どのアプリケーションでもスコアが高めな類似構造」の方が安心できそうです。機械学習系の予測ツール でも、多数決が精度を担保することが経験的によく使われます。(全員が間違えている可能性もあります が)

14	素人質問ですが、AlphaFold を使って低分子とのドッキング・シミュレーションは可能なのでしょうか？あるいは、今後、できる様になるのでしょうか？	AlphaFold3で低分子とタンパク質の「複合体構造」の予測ができるようになりました。ただし、アフィニティが予測できるわけではないことに注意が必要です。いわゆるドッキングスコアや結合自由エネルギーの値がほしいときは、別のツールを使うことになります。
15	立体構造を検索するFoldSeekについて、Web版のFoldSeekとGitHubのFoldSeekは同じモデルを用いているのでしょうか(同じ結果が得られるでしょうか)？	Foldseekは基本的な検索方法が固まっている(3次元の部分構造をアルファベットと見立てて配列アラインメントすること、web/githubで同じコードを使っているはずなので、違いは無いと思います。(バージョン更新によって、例えばアルファベットの定義の仕方を改良したりした、とかが起きると、結果が変わる可能性はあります)
16	AlphaFold2で出力された構造をinputにしてMD Simulationを実行し、構造変化を追跡するという方法は妥当でしょうか？	比較的小さいタンパク質であれば妥当だと思います。大きなものになると、初期構造依存性とサンプリングが大変になることから、AlphaFold2の初期構造が変だったときにMDの結果が変になる可能性は考えられます。
17	出力された構造の物理的な妥当性(静電的な相互作用、疎水性物質が細胞質方向に大きく露出してしまうなど)を評価するうえで有用なツールなどがありますか？Kihara LabのDAQはどうですか？	DAQ法、DAQ-scoreはクライオ電験でとった構造の構造決定上のミスを検出するものなので、直接的には使えません。(これを予測構造の評価に使う、という発想はあると思います。)
18	2つのタンパク質のドッキング予測する際に、2つのタンパク質の全体の配列を使うのと、ドッキングに関係しそうな部分に削ってドッキングするのは、どちらが良いですか？	末端の領域をカットしたり、disorder領域が多すぎるような場合に、構造ドメインのみを抽出してからドッキングさせた方が、うまくいくことは多いようです。
19	環状や分岐など1本鎖でないペプチドをリガンドとして、タンパク質との複合体構造は予測することはできますでしょうか？	head-to-tailのペプチド結合型の環状ペプチドは、Boltz-1などで予測できるようになっています。6の字型や8の字型、チオエーテル環状化などは、Boltz-1などにSMILESとして入力することで(一応)予測させることができます。
20	酵素活性を向上させようという取り組みがしたい場合、活性部位に変異を与えるのが良いかと思っておりましたが、点変異の予測が難しいとなると、別の手法の方が良いのでしょうか。	活性部位でどのような(他の分子との)相互作用がありそうかというのを、AlphaFoldなどで推測するという使い方はできると思いますが、基本的にはもう少し精密な方法(MDシミュレーションなど)を検討するのが良いかと思います。
21	糖鎖との複合体の構造予測は出来ますか？	AlphaFold3やBoltz-1などで予測を実行することはできます。が、糖鎖との複合体については、ちゃんと良い精度が出るのかはわかりません(大上は糖鎖の知見がないのでなんとも言えません)。
22	タンパク質周囲の脂質などの影響を解析できるモデルはあるのでしょうか？	基本的にはAIモデル系のツールはあまり無いのではと思います。脂質の影響などはMDシミュレーション・生物物理界限で広く研究はされていると思います。
23	大上先生は、pLDDTとPAEでは、まずどちらに注目して予測の正確性の判断をなさることが多いですか？	大きなタンパク質でなければ、基本的にはpLDDTを見ます。ドメインが複数になってくるとPAEも見ます。
24	AlphaFold2(localcolabfold)を用いて、天然変性タンパク質を予測する研究をしているのですが、天然変性タンパク質の構造予測にはAlphaFoldは不向きですか。また、これから天然変性タンパク質の予測ツールは発展していくでしょうか	PDBにある(世の中にある解けた)構造を信じているという性質上、天然変性を正しく扱えるようになるのはまだまだ先だと思います。「天然変性領域かどうか」の予測の精度は高いですが、「protein Bの結合によってprotein Aの天然変性領域が〇〇という構造を取る」みたいなことをことを正確に予測できるようにはなかなかならなそうです。もちろん、動きが大きいタンパク質など、がんばって全部のコンフォメーションを予測させようという流れはあり、徐々に解決するかもしれません。
25	10〜20アミノ酸くらいの短いペプチドとタンパク質のAlphaFold2による複合体予測は精度はどの程度でしょうか？	それまでにあったタンパク-ペプチドドッキングツール(galaxypepdock, glide peptide mode, cabsdockなど)と比べて、AlphaFold2のタンパク-ペプチド複合体予測の精度は格段に高いと評価されています。それでも、体感で成功率3〜5割くらいだと思います(成功の定義にもよりますが)。
26	複合体としてもAlphaFold2よりもAlphaFold3のほうが予測精度が高いのでしょうか？	AlphaFold3の論文内でテストされた結果によると、そのようです。特に抗体抗原複合体の予測精度が上がっています。
27	アミノ酸の長さによって計算時間はどれくらい変わりますでしょうか？	AlphaFold2, AlphaFold3, その他、また利用する計算機環境(GPU)でも結構変わりますが、AlphaFold2/ColabFoldの時は200aa程度で30分、1000aaで10時間ぐらいの計算時間でした。
28	AlphaFold3でZn fingerが3つある構造を予想させたところ、1つしかZnが入った構造しか出てこないのですが、改善できますか？	なかなか改善させることは難しいかもしれません。Zn fingerモチーフ自体が正しく予測できていれば、あとからZnを手で置くなどが必要だと思います。
29	二量体や四量体などを形成するようなタンパク質の複合体予測において、片方を野生型もう片方を変異体にするようなことは可能でしょうか。	配列をWT/MTそれぞれ入力することで可能です。WTPROTEINSEQ:MTPROTEINSEQみたいなイメージです。
30	動物以外のタンパク質(例えば植物)におけるAlphamissenseのようなツールはあるのでしょうか？	もしかしたらあるかもしれません。機能が失活しそう／構造が壊れそうな変異を推定したいという観点で見れば、予測できないことは無いと思います。
31	Diffusionモデルを使用したAlphaFold3では非天然アミノ酸を含む10〜15程度の環状ペプチドの構造予測は可能でしょうか？	SMILESに変換して入力することで可能です
32	翻訳後修飾の精度があまりよくないとの話でしたが、糖鎖(O型なしはN型)も含めて構造予測は可能でしょうか？	入力できる糖鎖修飾は、以下に説明があります。(入力できるだけで、その影響を正しく反映して予測できているかは怪しいです。) https://alphafoldserver.com/faq#can-i-model-glycosylated-proteins
33	CASP16の結果やAIの歴史などのご紹介もありましたが、AlphaFold系以外だと、なかなか太刀打ちできていない状況なのでしょうか。(そもそも、それ以外の方向での研究者もAlphaFold系に移っていることなどもあるのでしょうか)	もちろんCASP17で話が変わるかもしれませんが、現状はAlphaFold系の手法がトップです。David Bakerらが作ったRoseTTAFoldもAlphaFold2のアーキテクチャに着想を得て作られています。

34	多量体形成するタンパク質Aと、酵素Bからなる融合タンパク質ABにおいて、予測ツールを用いた融合タンパク質の多量体形成予測は可能でしょうか。現状では、おそらく酵素Bの分子サイズに引っ張られ、タンパク質Aのドメインでの多量体形成がうまく行きません。	うまくできる場合もあると思いますが、一般にはタンパク質のサイズが大きくなってくると怪しくなってきます。
35	立体構造予測においてフォールディングされる順番を加味する研究などはあるのでしょうか？翻訳の順番や、シャペロンのようなものがくっついているかどうかで、同じ配列でも取りやすい構造は変わりそうな気がしました。ホモロジーの情報に内包されていそうな気がします。	構造予測だけで議論することは難しいかもしれませんが(PDBの登場頻度＝人間の興味が強い構造に引きずられる)が、構造予測＋フォールディングシミュレーションのアプローチはいろいろ検討されていると思います。最後に紹介した東大の新井先生の仕事も近いと思います。
36	pLLDなどの予測信頼性の数値の計算式は公開されているのでしょうか？	ニューラルネットワークによる予測が使われています。詳細はAlphaFold2論文のSupporting InformationのAlgorithm 29に記載があります。
37	AF3でドッキングや機械学習のモデルに使用できないのは、なぜでしょうか。	AlphaFold serverの利用規約でGoogleが禁じているからです。AlphaFold3のソースコード版ではそのような制限はありません。
38	ゲノムにおける点変異とかRNA同士の融合遺伝子の解析などに、AlphaFoldシリーズのモデルは利用できるのでしょうか？例えば、DNAシークエンスから、タンパク質のシークエンスを推定して、そこから構造を解析。その構造の変化によって、がんの病的バリエーションになりうるかどうかを推定できればと思っています。	ゲノム→タンパク質への翻訳後に、構造予測させることはできます。DNAのバリエーションが構造にどう影響し、そうかを推定する(まさにAlphaMissenseと似た発想ですが)という点にも使えると思いますが、ゲノム言語モデルといった専用のツール(Evoなど)もあるので、そちらを使ったほうがいいかもしれません。
39	AlphaMissenseにおいて、変異がPathogenicと判断される根拠となる考えはなんなのでしょうか？あるタンパクの変異体を作成したい場合(例えば熱安定性を上げたいなどの目的)、Pathogenicではない残基を狙うことは有効でしょうか？	構造のコアの部分の変異は構造に強い影響がありそう(壊れたりしそう)、だからpathogenicだ、という発想です。Pathogenicではない残基を狙うこと自体は有効かもしれませんが、逆にPathogenicだと判断されるような重要な残基箇所の相互作用を強化するような変異をあえて入れる(構造がより安定になるかも)という戦略もあるかもしれません。
40	AlphaFold2よりAlphaFold3の方が蛋白質の立体構造予測精度(蛋白-蛋白複合体予測精度も?)も高くなっているというお話ですが、AlphaFold2とChai-1/Boltz-1の蛋白質立体構造予測精度に優劣はありますでしょうか？	基本的には、AlphaFold2と、AlphaFold3/AlphaFold3-like toolとで、後者の方が明確によくなっています。AlphaFold3-like toolはどれもあまり変わらないので、いずれを使っても良いと思います。
41	AlphaFold3において予測したタンパク質構造において、リガンドが結合する座標はどのように予測できるのでしょうか？	リガンドを指定すると、このあたりに予測するだろうというところを拡散モデルで予測しています。
42	AlphaFold3などの共有結合ドッキングについて、大上先生の発表では予測構造にリガンドをドッキングすると変な位置に配位してしまうこともあるとおっしゃっていたと思います。一方で共有結合する箇所が分かっている場合、AlphaFold3などでの共有結合ドッキングを行った場合、信頼性が高いのでしょうか。	共有結合の場所が合っていれば、おそらく大丈夫だと思いますが、精度についてはあまり評価した結果がありません。学習の度合いによっても異なると思います。
43	AlphaFold, Chau-1, Bolz-1につきまして、リガンド側の構造や安定配座などは考慮されるのでしょうか？また、複合体を予測する際、リガンドが結合した上での構造変化も含めて予測されるのでしょうか？	リガンドの構造・配座はタンパク質の構造に合わせて予測します。必ずしも安定配座になっているかはわかりません(なっていないと思います)。また、立体化学や結合長がおかしかったりすることがあります(Boltz-1のような最新版だと改善していると思います)。
44	ある1つのタンパクと他のタンパクの相互作用を網羅的に(A vs B, A vs C,,,,A vs Zのように)調べることは可能でしょうか？	A:B, A:C, A:D, を順に実行して比べる、という単純な方法はもちろんできます。AlphaPulldown: https://github.com/KosinskiLab/AlphaPulldown や、我々が開発しているSpatialPPI: https://github.com/ohuelab/SpatialPPIv2 などでも可能です。
45	がんなどの疾患では正常ではないバリエーションタンパクが出来ますが、正常なものと疾患での異常なものとで、親和性が変わる推定結合タンパクのスクリーニングができるモデルはあるのでしょうか？	AlphaPulldown: https://github.com/KosinskiLab/AlphaPulldown や、我々が開発しているSpatialPPI: https://github.com/ohuelab/SpatialPPIv2 などでも予測してみることが可能ですが、異常なバリエーションタンパク質を正しくモデリングできてそうかという点が重要で、ケースバイケースになってきます(うまくいくものもあるかもしれませんが、という程度)。
46	今回ご紹介いただいたツールを用いて、酵素の活性が向上するような変異の探索は可能でしょうか？事例等がございましたらご教示いただきたいです。	AlphaFold3は点変異の予測精度が上がっています。論文もありますが、AlphaFold2よりは良いです。他の専用ツール(Flex ddGやFoldX)などのほうが精度が良いと思いますが、使ってみて比べたり、そのようなことをやっている論文を調べてみるのが良いと思います。
47	ローカルで実施すると仮定した場合、どれぐらいのCPU及びGPUが必要になりますか。	基本的にGPUがあれば、Google Colaboratoryに載るくらいのは大体できます。AlphaFold3についてはSSDが必要で、Multiple Sequence Alignmentを計算するときには計算コストがかかります。ColabFoldはウェブサーバーでMSAを計算するのですが、配列をネット環境に流したくない場合にはこれを完全にローカル化する必要があり、メモリが760GBくらい、SSDが1～2TBなど、結構な計算リソースが必要になります。
48	これらツールを用いて、あるタンパク質と特定のリガンドが結合するか否かを判断することは可能でしょうか。	一応予測させることはできますが、そのスコアがどの程度正しいかはわかりません。
49	AlphaFold3で構造予測させた時、いくつかの構造が予測されたのですが、どれが可能性が高いか判定する方法がありますか？	AlphaFold3で調べる限りは、ipTMやpTM、それらを統合して得られるランキングスコアが出てくるので、それをもとに判定できます。
50	AlphaFold2で予測された3D構造から、N末、C末がどちらかわかる方法はありますか？	分子ビューワー(pymolなど)で配列と構造の位置を確認するのが手っ取り早いです。

51	Boltz-1で入力するMSAファイルはどのように作ったら良いでしょうか？ファイルはどのような形式なのでしょう？	use msa serverというオプション(引数)を指定すれば、MSAを指定しなくてもCSVファイルが作られます。その作られたファイルの一部のリガンドを他のリガンドに変えて、アミノ酸配列は同じまま予測するということに、その作られたファイルを参照すればMSAを使い回すことができます。基本的には、Boltz-1でできるCSVファイルか、ColabFoldで出てくるa3mとよばれるファイル形式のものを使うとよいと思います。GitHubのリポジトリの中に、具体的にどのようなファイルがあるのかを見られます。
52	MMseqs2サーバーを使用せずに(配列を外部に出さずに)MSAを取得する方法には、どのような方法がありますでしょうか？	講義資料をご参照ください。
53	boltz1の予測に使用するMSAを自分で指定する場合とありましたが、MSAのデータベースのようなものがあるのでしょうか？サーバーからMSAを取得しない場合、どのようにMSAを準備すべきでしょうか？	MSAのデータベースがあり、ウェブ上から取得しています。そうでない場合は、qiitaの記事(講義スライド7)をご参照ください。*発展的な内容なので難しいかもしれません。
54	Chai-1サーバー版の利用は、入力したクエリーや実行結果が、運営者から利用される可能性もあると考えてよろしいでしょうか。	わかりません。基本的には、ウェブに投げる情報はすべて、そういう可能性が無くはないと思って使用してください。
55	ccdやmsaなどでオンラインが必要のように感じましたが、ローカルでの操作を行う際、スタンドアローンでBoltz-1は実行できますでしょうか？	CCDの参照はローカルで完結しています。また、MSAについては、MSAを入力として取らなくするか、Qiitaの記事(講義スライド7)のようにMMSeqs2をローカルで動かすかの2つの方法があります。また、AlphaFold3はMMSeqs2でなくjackhmmerを使っているので、ローカルで実行できます。
56	ColabFoldで無料で予測できるアミノ酸配列長は1000残基程度らしいのですが、それより長い配列を予測する方法を教えてください。	これは利用するGPUのメモリに依存します。Colab無料版はT4という廉価なGPUが割当られます。解決方法は、(1)Colab有料版でA100 GPUなどを使う、(2)手元にGPU環境を用意してlocalColabFoldを使う、(3)AlphaFoldserver(AlphaFold3)を使う(上限5000aa)、などがあります。
57	イオンチャネルのopen-close-inactevateのように特定の条件下(電位、熱、圧力など)で構造が変化するようなタンパク質について、stateを指定したような形でAlphaFol2などで予測できるのでしょうか？それとも、最も安定なstate(イオンチャネルのclose型)で構造が予測されるのでしょうか？	stateを指定して構造予測することは今のところできないです(close型のキーになっているcontact残基ペアに対してMSAを破壊することで無理やりそのstateを回避して予測させるような方法は提案されています https://www.nature.com/articles/s42003-025-07791-9)。基本的には安定なstateが予測される事が多いですが、PDBでよく見かける構造が出力されていると取ることもできます。
58	AlphaFold3やBoltz-1のGithub版を使用したい場合、どれくらいのスペックの計算機を用意する必要がありますか。GPUは要りますか。	基本的にGPUは必須です。AF3の場合は配列データベースが必要であるためBoltz-1よりもより多くのスペックが必要です。詳細は https://github.com/google-deepmind/alphafold3/blob/main/docs/performance.md をご覧ください。
59	構造予測した結果がウェット実験の結果と一致しなかった場合でも、その配列が予測構造をとる「潜在的能力」があるという解釈でウェット実験の結果に新たな解釈を加える、ということは考えられますでしょうか？	「PDB(既知構造)の中に類似配列でその構造を取っているものがある」という意味で可能性を吟味することはできそうです。
60	1アミノ酸変異をしたタンパク質に対するBoltz-1で予測した構造に対する、ドッキングシミュレーションの妥当性を教えてください。また、結合位置はどのツールを使って予測するのがおすすめでしょうか？	タンパク質ドッキングの場合は一定程度関連すると思います(https://www.biorxiv.org/content/10.1101/2024.05.25.595871v1)。リガンドドッキングについて変異の影響を調査した研究は知る限りまだないと思いますが、基本的には従来のドッキング手法と併用するのが良いと思います。結合位置(活性部位)の予測についてはAlphaFold3-like手法の性能は高いと思いますが、必ず信頼できる性能ではない点に留意が必要です。
61	PROTACのような3者複合体構造予測は難しいとのことですが、将来的に予測の精度があがる見込みはありますでしょうか？やはり実験的な構造情報が蓄積しないと難しいですか？	デグレーダー、molecular glue含め、とてもホットな分野だと思います。実験的な構造が蓄積されて解決するのか、ということも怪しいですが、学習に足る構造が今は全然足りない状況だとは思います。E3 ligaseとの複合体が必ずしも安定的であるのかどうか、みたいな、根本的な議論もあると思います。
62	AlphaFoldやBoltz-1に限らず、予測された立体構造を人の目で見て判断する場合、最低限どの程度の知識レベルが求められますでしょうか？入門的な教科書や論文はありますでしょうか？	まずは、「実験医学別冊 最強のステップUPシリーズ AlphaFold時代の構造バイオインフォマティクス実践ガイド」(羊土社)(https://www.yodosha.co.jp/yodobook/book/9784758122764/) がおすすめです。どちらかというと生物学的なバックグラウンドが要求されることが多く、ケースバイケースになりがちですので、そのタンパク質自体の背景知識を調べることも重要です。
63	Chai-1, Boltz-1もAlphaFoldとの比較もあってか、立体構造学習のカットオフが 2021年に設定されているかと思います。一方で、その後もたくさんのタンパク質構造が解かれています。Boltz-1は学習が可能なようですが、新たなタンパク質構造を学習させることについて、予測の向上の観点から、どのようにお考えになりますか？	精度向上の可能性は高いです(特に低分子やほかの分子とタンパク質の複合体)。モデルアップデートは今後、AlphaFold本家や、それぞれの手法でも進むと思います。
64	ColabFoldの予測データをAutoDock Vinaのドッキングシミュレーションする際ですが、古井講師の講義資料(スライド10)に記載のようにipTM0.8が指標でしょうか？ipTMの値については、どの程度の妥当性があるかは相関(結晶構造との相関比較)ありますでしょうか？	ipTM>0.8が成功かどうかの指標になっています。AlphaFold-Multimer論文(https://www.biorxiv.org/content/10.1101/2021.10.04.463034v2)にDockQとipTMの相関に関する図が載っています。
65	AlphaFold Databaseでは、アイソフォームによる違いはどのように扱われていますか。	一番有名なアイソフォーム(UniProtでtopに出てくる配列)が予測されています。異なるアイソフォームは自分で予測する必要があります。
66	dockingとco-foldの一番大きな違いはなんですか	dockingは受容体の構造が与えられており、co-foldは受容体とリガンドの配列・SMILES情報から複合体構造を直接予測します。(dockingという言葉が比較的幅広いため、co-foldもdockingと呼ぶことがあります)

67	Boltz-1で、トータルの配列数が1000残基を超えるような大規模なタンパク質複合体を予測する場合、ran out of memoryで途中でfailします。メモリ容量が大きいGPUで再計算する以外で、回避策はご存じでしょうか？	現状は難しいかもしれませんが、将来的なアップデートでbfloat16を用いたメモリ効率化が導入される可能性があります。 https://github.com/jwohlwend/boltz/pull/264
68	disorder領域とのドッキングを予測する際には、どのプログラムを利用するのが現状は信頼が高いでしょうか？Alphafold3でも結合しているようには見えますが、どの妥当性評価が悩ましいと感じていました。	現状では信頼に足るプログラムはなく、なかなか手を出すのが難しいと思います。
69	AlphaMissenseは公開されていますか。何かしらの方法で使うことはできるでしょうか。	AlphaMissense自体は公開されていません(倫理的な理由が挙げられています)が、TogoVarやClinVarとの紐づけはされています。
70	低分子化合物との結合を解析できるWebツールはありますか	AlphaFold3-likeなリガンド複合体構造予測をWebツールで実行する場合は、Chai laboratoryを使うか、Boltz-1をColaboratory上で使うかのいずれかになります。 また、リガンドドッキングシミュレーションの意味であれば、我々はあまり試したことはありませんが、SwissDock (https://www.swissdock.ch/) でAutoDock Vinaを使ったりリガンドドッキングがWeb上で実行できるようです。
71	ドメイン間がリンカー様の配列で結合されているようなマルチドメインのタンパク質構造の予測において、構造予測を改善するためにできる工夫はあるでしょうか。ドメイン間のインタラクションが既知の箇所に拘束をかけるといった方法でしょうか。	ドメイン間の相互作用が既知の箇所に拘束をかけるのは有力な選択肢です。また、拘束なしの場合の予測結果と比較することも有効だと思います。
72	営利目的で活性部位の基質のドッキングの様子を調べたい場合、どのツールが適しているでしょうか？	Boltz-1(もしくはChai-1)のGithub版を利用するのが適していると思います。
73	予測した立体構造をもとに、立体構造のデータベース等から相互作用する可能性のあるタンパク質を探索するようなツールはあるでしょうか？	相互作用予測ツールは(inputが配列のものや構造のものなど、様々ですが)数多く提案されています。 AlphaFoldの結果に基づいて予測を行うAlphaPulldown (https://academic.oup.com/bioinformatics/article/39/1/btac749/6839971) が有名です。また、大上研でもSpatilaPPI (https://github.com/ohuelab/SpatialPPIv2) というツールを開発しています。
74	AlphaFoldで予測したタンパク質構造同士の類似性を比較する際に、RMSDを算出すると思いますが、この際は、炭素原子のみを比較するのがよいでしょうか？	C α のみを使ってRMSDを算出することが多いですが、重原子でのRMSDを用いることもあります。