

※2024年10月 Demis Hassabis氏とJohn M. Jumper氏のノーベル化学賞受賞を受けて、AlphaFold3の使い方についての講義資料を再編集しました。

バイオインフォマティクス

(2024年6月10日)

第9回

タンパク質の構造予測

近畿大学・理工学部
生命科学科
飯田 慶 (いいだ けい)

タンパク質構造予測の変革

AIの発展により、タンパク質構造予測の状況「も」大きく変化している。



(wired.jp)

By DeepMind Technologies
2010年創業、2014年よりGoogle傘下

タンパク質構造予測の変革

AIの発展により、タンパク質構造予測の状況「も」大きく変化している。



(wired.jp)

By DeepMind Technologies
2010年創業、2014年よりGoogle傘下

AlphaFold3を使ってみよう（1）

① a)~c)のいずれかの方法でアクセス

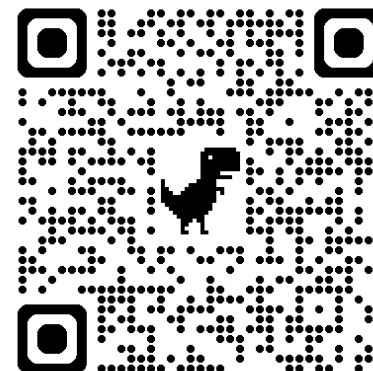
a) alphafold serverで検索



b) URLを打ち込む
(スペルミスに注意！)

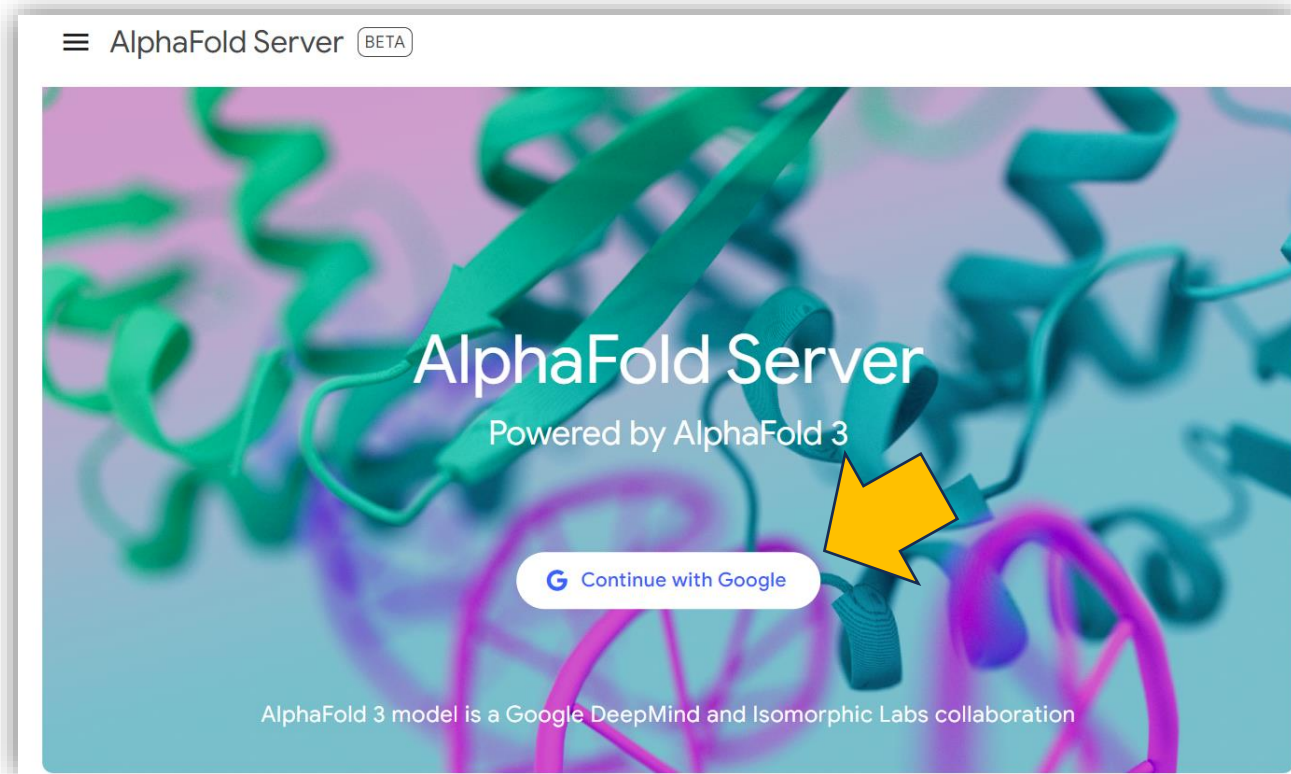
<https://alphafoldserver.com/>

c) バーコードから
アクセス



AlphaFold3を使ってみよう（2）

② “Continue with Google”を選択



AlphaFold3を使ってみよう（3）

- ③ グーグルのアカウントでログインする、chromeでログインしている人は、そのままアカウントが表示されているので、そのまま先に進む。
（近畿大学の方は大学アカウントでOK）

☰ AlphaFold Server BETA

Complete Profile

Google account
慶飯田
kiida@life.kindai.ac.jp

3-1. 確認

☐ Subscribe to receive news and product updates from AlphaFold Server. Your information will be used in accordance with [Google's privacy policy](#). You may opt out at any time.

3-2. クリック

Continue to Terms of Service

AlphaFold3を使ってみよう（４）

注意事項は目を通しておく。

AlphaFold Server Terms of Service

[AlphaFold Serverの使用には、Googleの利用規約](#)および[AlphaFold Serverの追加利用規約](#)が適用されます。

知っておくべき重要なこと

1. AlphaFold Server は、個人および非営利組織 (大学、非営利団体、研究機関、教育機関、政府機関) による非営利目的、またはジャーナリズム目的でのみご利用いただけます。
2. AlphaFold Server またはその出力を使用しないでください。
 - a. 営利組織に代わって行う調査を含むあらゆる商業活動に関連して;
 - b. GlideやAutoDockなど、タンパク質とリガンドまたはペプチドとの結合または相互作用を予測する自動化システムにおいて、または
 - c. AlphaFold Server と同様に、生体分子構造予測のための機械学習モデルまたは関連技術をトレーニングします。
3. お客様は、継続的な使用が[AlphaFold Server 出力の利用規約](#)およびお客様が行った変更の対象となることを明確に通知する要件を含む、当社の規約に従って AlphaFold Server 出力を公開、共有、および適応することができます。

AlphaFold3を使ってみよう（5）

If applicable, please confirm your:

- **Organization, university, or other affiliation(s):**

Organization
KINDAI University, JAPAN



5-1. 大学名を入れる。近大生の場合、例の通りでOK

As a reminder, AlphaFold Server is not available for: (i) commercial work, even if on behalf of a non-commercial organization; (ii) commercial organizations. The only exception is journalism which is permitted.

☒ I accept the [Google Terms of Service](#) and [AlphaFold Server Additional Terms of Service](#)



5-2. チェックを入れる

5-3. クリック

Accept and continue



AlphaFold3を使ってみよう（6）

プライバシーポリシーについて

（近大生の場合、大学のアカウントがGoogle発行のもの。ここでのプライバシーポリシーもそれに準じている）

Using AlphaFold Server:

Last Modified: May 8, 2024 | [Download PDF](#)

By using AlphaFold Server, you acknowledge you have read the [Google Privacy Policy](#) and this AlphaFold Server Privacy Notice:

1. In accordance with the “Retaining Your Information” section of the [Google Privacy Policy](#), we may retain time-stamped records of the sequences you input into AlphaFold Server and the output it generates (“**User Record**”) for an extended period of time to monitor compliance with, enforce and investigate potential violations of the Terms.
2. Notwithstanding the “Exporting, removing & deleting your information” section of the [Google Privacy Policy](#), **deleting your Google account or amending or deleting your AlphaFold Server history will not delete your User Record**. We take steps to protect your privacy, including

Cancel

Continue

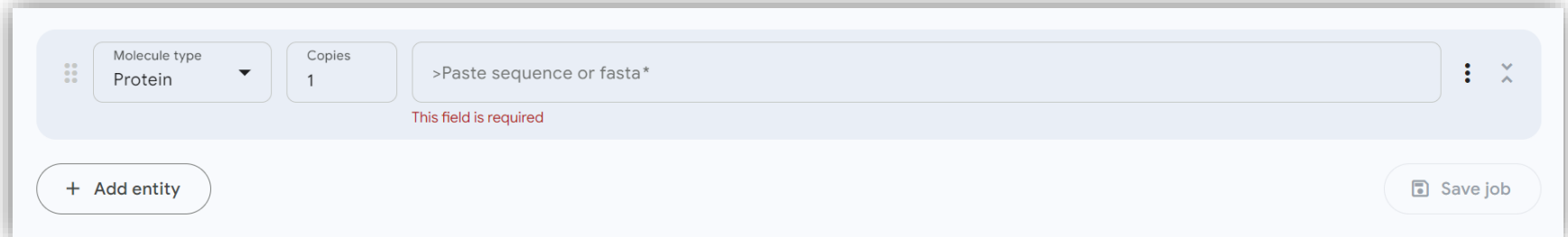
目を通した後クリック



AlphaFold3を使ってみよう（7）

※アカウント登録後はここから作業開始

タンパク質のアミノ酸配列を入れるウィンドウがあることを確認



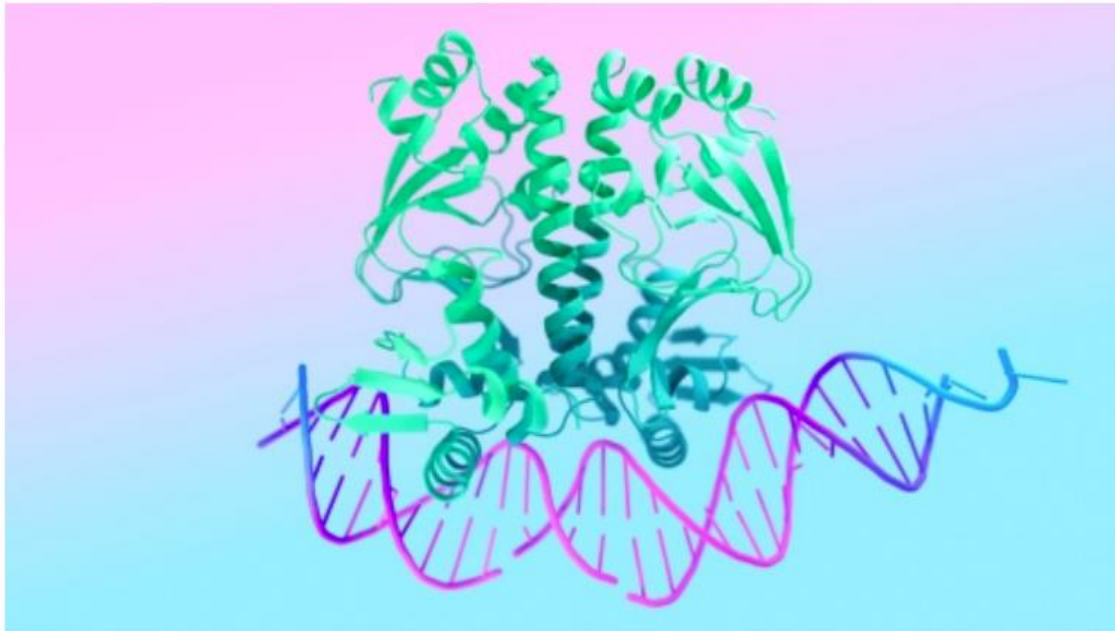
The image shows the input interface of the AlphaFold3 web application. It features a light blue header bar with a grid icon on the left. The main input area contains three elements: a 'Molecule type' dropdown menu set to 'Protein', a 'Copies' input field set to '1', and a large text input field labeled '>Paste sequence or fasta*'. Below the text input field, a red error message reads 'This field is required'. To the right of the text input field are three vertical dots and a close icon. At the bottom left is a button labeled '+ Add entity', and at the bottom right is a button labeled 'Save job' with a floppy disk icon.

題材

DNA塩基配列の”編集”を行うために設計されたタンパク質と、DNA二重鎖、金属イオンの複合体を予測してみる。

従来、タンパク質の構造予測だけでも困難だった。

AlphaFold3はこれを高い精度で行い、さらに分子間相互作用まで予測する！



グーグル・ディープマインドが開発したAlphaFold3はタンパク質と他の分子の相互作用を予測できる

(出所：グーグル・ディープマインド)

[画像のクリックで拡大表示]

(日経クロステックHPより)

題材

ホーミングエンドヌクレアーゼ (メガヌクレアーゼとも呼ばれる); 微生物の可動性介在配列内にコードされているDNA切断酵素。
塩基配列を認識してDNA切断をするため、DNA編集を目的としたタンパク質デザインの標的となる。

(アミノ酸配列を変えた時に、どのような塩基配列を認識するか予測が重要！)

Nucleic Acids Research

核酸研究 2017年8月21日; 45(14):8621–8634.

PMCID: PMC5737575

2017年6月20日にオンラインで公開されました。doi: [10.1093/nar/gkx544](https://doi.org/10.1093/nar/gkx544)

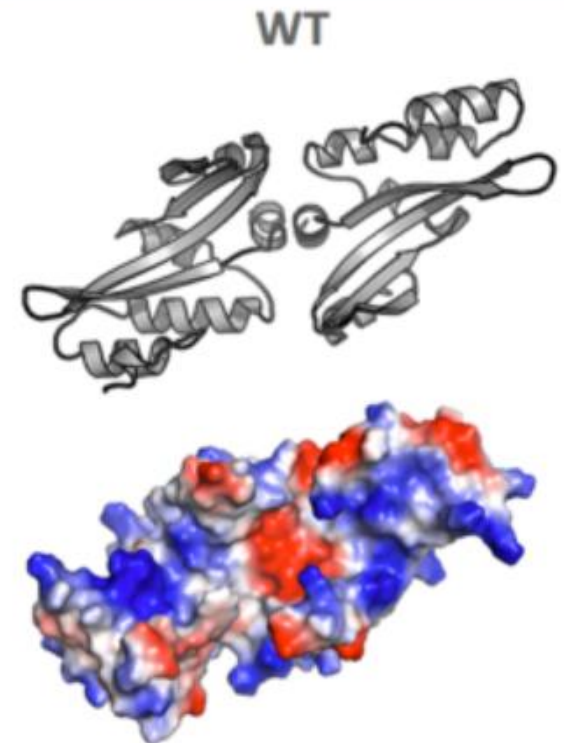
ID: [28637173](https://pubmed.ncbi.nlm.nih.gov/28637173/)

結晶構造解析により、メガヌクレアーゼの標的特異性の顕著な可塑性と効率的な再コーディングが明らかになった。

レイチェル・ワーサー、¹ ジャズミン・P・ハリナン、¹ アビゲイル・R・ランバート、¹ カイル・ハイブズ、² マーク・ボグソン、² ジョーダン・ジャージュール、² ロベルト・ガリジ、³ ニコライ・ウィンドビラー、³ アンドレア・クリサンティ、³ トニー・ノーラン、³ バリー・L・ストッダード¹

▶ 著者情報 ▶ 記事注記 ▶ 著作権とライセンス情報 [PMC免責事項](#)

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5737575/>



解析に使う、タンパク質と核酸の配列

>Meganuclease_I-OnuI_3rd

MASSRRESINPWILTGFADAEGSFGLSILNRNRGTARYHTRLSFTIMLHNKDKS
ILENIQSTWKVGSILNNGDHYVSLVVYRFEDLKVIIDHFEKYPLITQKLGDYKLF
KQAFSVMENKEHLKENGIELVRIKAKMNWGLNDELKKAFFPENISKERPLINK
NIPNFKWLAGFTSGDGSFFVRLRKSNNARVRVQLVFEISQHIRDKNLMNSLI
TYLGCGHIYEGNKSERSWLQFRVEKFSDINDKIIPVFQENTLIGVKLEDFEDWC
KVAKLIEEKKHLTESGLDEIKKIKLNMNKGR

>DNA1

GGGGGCATGCAGATCCACAGGCGCG

>DNA2

GGCCGCGCCTGTGGGATCTGCATGCC

AlphaFold3を使ってみよう（8）

Molecule type
Protein

Copies
1

>Paste sequence or fasta*

⋮ ×

This field is required

+ Add entity

Save job



Meganucleaseのアミノ酸配列を張り付ける(300残基であることを確認)

Molecule type
Protein

Copies
1

MASSRRRESIN	PWILTGfADA	EGSFGLSILN	RNRGTARYHT
RLSFTIMLHN	KDKSILENIQ	STWKVGSILN	NGDHYVSLVV
YRFEDLKVII	DHFEKYPLIT	QKLG DYKLFK	QAFSVMENKE
HLKENGIKEL	VRlKAKMNWG	LNDELKKAfP	ENISKERPLI
NKNIPNFKWL	AGFTSGDGSF	FVRLRKSNVN	ARVRVQLVFE
ISQHIRDKNL	MNSLITYLGC	GHIYEGNKSE	RSWLQFRVEK
FSDINDKIIP	VFQENTLIGV	KLEDFEDWCK	VAKLIEEEKKH
LTESGLDEIK	KIKLNMNKGR		

⋮ ×

AlphaFold3を使ってみよう (9)

The interface shows a single entry for a protein sequence. The 'Molecule type' is set to 'Protein' and 'Copies' is 1. The sequence is displayed in a grid with residue numbers 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300. The sequence is: MASSRRRESIN PWILTGFADEGSFGLSILNRNRGTARYHT RLSFTIMLHN KDKSILENIQSTWKVGSILNNGDHYVSLVV YRFEDLKVII DHFEKYPLITQKLG DYKLFKQAFSVMENKE ENISKERPLI HLKENGikel VRIKAKMNWGLNDELKKAFP ARVRVQLVFE NKNIPNFKWL AGFTSGDGSF FVRLRKSNNV ARVRVQLVFE ISQHIRDKNL MNSLITYLGC GHIYEGNKSE RSWLQFRVEK FSDINDKIIP VFQENTLIGV KLEDFEDWCK VAKLIEEKKH LTESGLDEIK KIKLNMNKGK.

Buttons: + Add entity, Save job

+ Add entryをクリックして、
入力枠を増やす

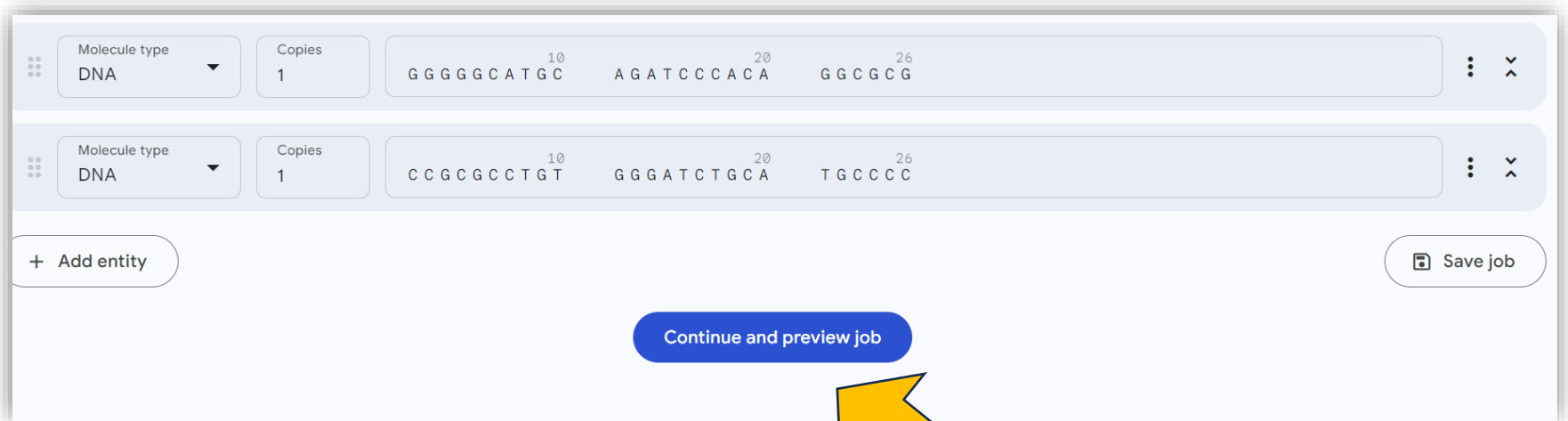


The interface now shows multiple entries. The top entry is a protein sequence (Molecule type: Protein, Copies: 1). Below it is a new entry for a DNA sequence (Molecule type: DNA, Copies: 1) with a text input field labeled '>Paste sequence or fasta*'. The protein sequence is now split into two columns, with residue numbers 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300. The DNA sequence is: >Paste sequence or fasta*.

Buttons: + Add entity, Save job

AlphaFold3を使ってみよう (10)

10-1. 配布したfastaファイルから
DNA1, DNA2の塩基配列を張り付ける。



The screenshot shows the AlphaFold3 web interface. It features two input rows for DNA sequences. Each row has a 'Molecule type' dropdown set to 'DNA', a 'Copies' field set to '1', and a text area for the sequence. The first row contains the sequence 'GGGGGCATGC' (positions 1-10), 'AGATCCCACA' (positions 20-29), and 'GGCGCG' (positions 26-31). The second row contains 'CCGCGCCTGT' (positions 1-10), 'GGGATCTGCA' (positions 20-29), and 'TGCCCC' (positions 26-31). Below the input rows are buttons for '+ Add entity', 'Save job', and 'Continue and preview job'. A large yellow arrow points to the 'Continue and preview job' button.

Molecule type	Copies	Sequence
DNA	1	GGGGGCATGC ¹⁰ AGATCCCACA ²⁰ GGCGCG ²⁶
DNA	1	CCGCGCCTGT ¹⁰ GGGATCTGCA ²⁰ TGCCCC ²⁶

+ Add entity

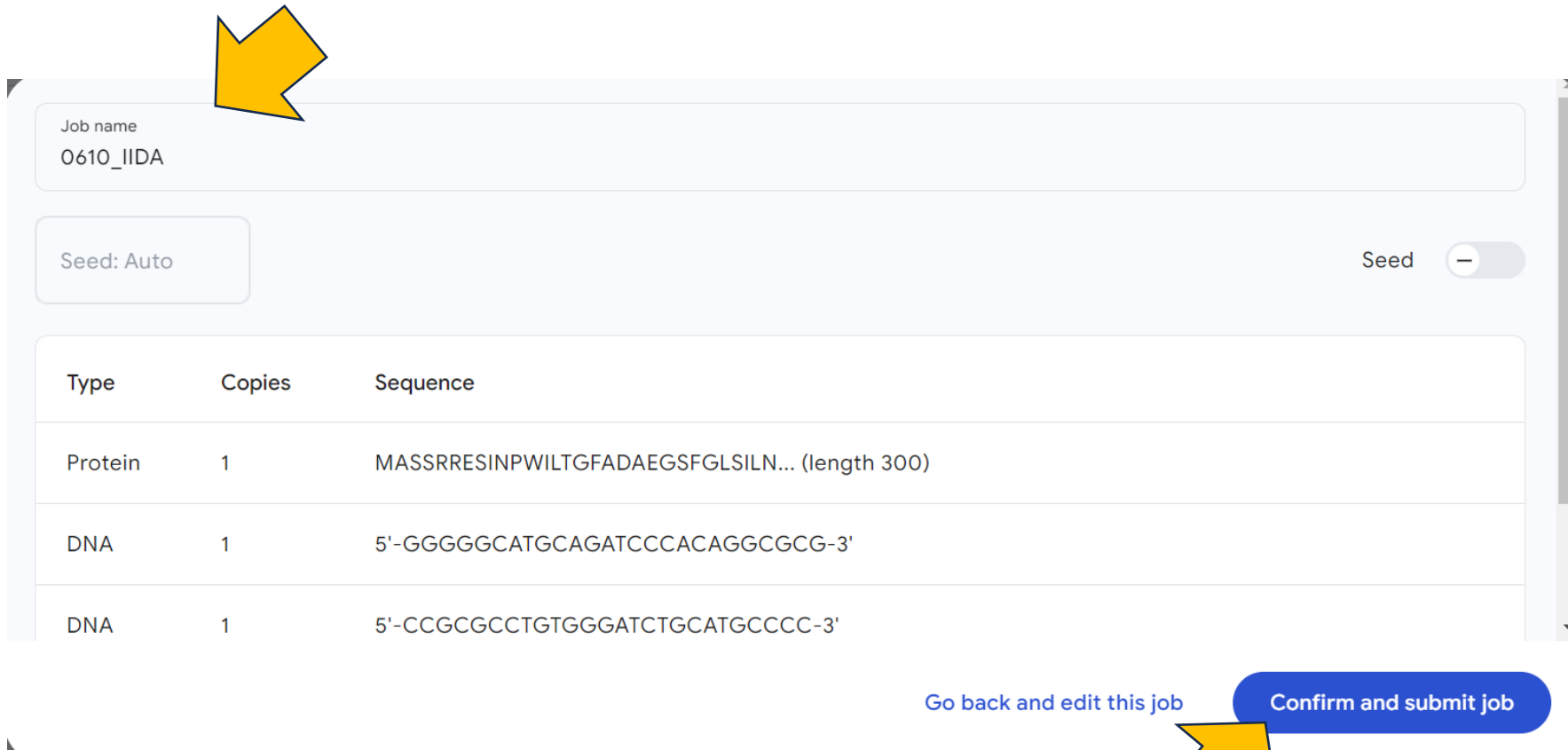
Save job

Continue and preview job

10-2 すべての配列を張り付いたら
クリック

AlphaFold3を使ってみよう (11)

11-1. Jobタイトルを入力する (任意)



Job name
0610_IIDA

Seed: Auto

Seed ☐

Type	Copies	Sequence
Protein	1	MASSRRESINPWILTFADAEGSFGLSILN... (length 300)
DNA	1	5'-GGGGGCATGCAGATCCCACAGGCGCG-3'
DNA	1	5'-CCGCGCCTGTGGGATCTGCATGCCCC-3'

[Go back and edit this job](#) [Confirm and submit job](#)

1日の実行数に上限があるので
間違えないように！

11-2. クリック
(計算が始まるので、待つ)

AlphaFold3を使ってみよう (12)

Name	Modified
<input checked="" type="checkbox"/> 7rce_modified_IIDA	2024-06-08 11:10

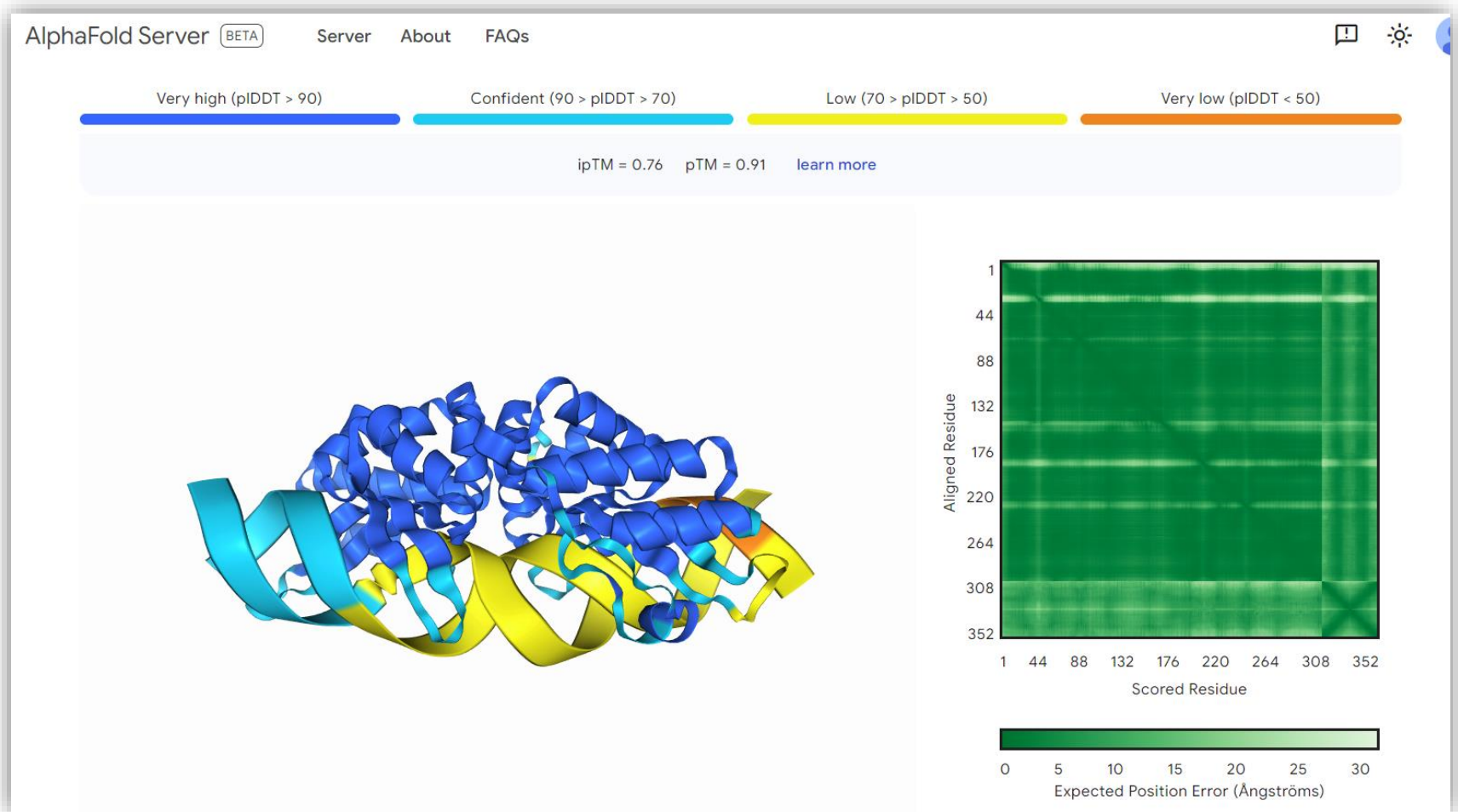
(AlphaFold severに戻り)
12-1. ここがチェックマーク
に変わっていたら計算終了

12-2.
ここクリックして、“open
results”を選択

	✓ Open results
	↓ Download
	📄 Clone and reuse
	✎ Rename
	🗑 Delete
Modified	⋮
2024-06-08 11:10	

AlphaFold3を使ってみよう (13)

結果の確認



- 予測の確からしさは
カラーバーで表示される

- 原子間の相対位置関係
の確からしさ

正解構造を確認してみる

① PDBjで ID; 7RCEを検索

<https://pd bj.org/mine/summary/7rce?lang=ja>

7RCE

Third stage reengineered variant of I-OnuI with specificity enhancing substitutions

7RCE の概要

エントリーDOI	10.2210/pdb7rce/pdb
分子名称	I-OnuI_e-hPD1-d, DNA (26-MER), CALCIUM ION, ... (6 entities in total)
機能のキーワード	meganuclease, gene editing, protein engineering, dna binding protein-dna complex, dna binding protein/dna
由来する生物種	Synthetic construct

ダウンロード

- Sequence (fasta)
- PDBx/mmCIF
- PDBML (ヘッダのみ (no-atom))
- PDB形式 (全ての情報)
- 検証レポート (PDF)
- EDMap 2fo-fc (MTZ)
- [More...](#)

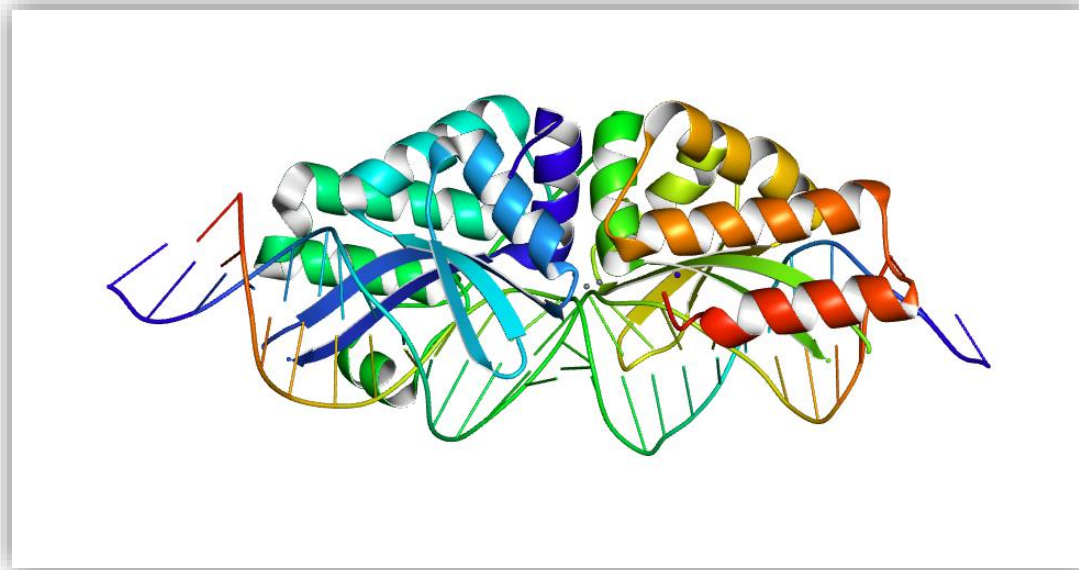
構造

[非対称単位を表示](#)

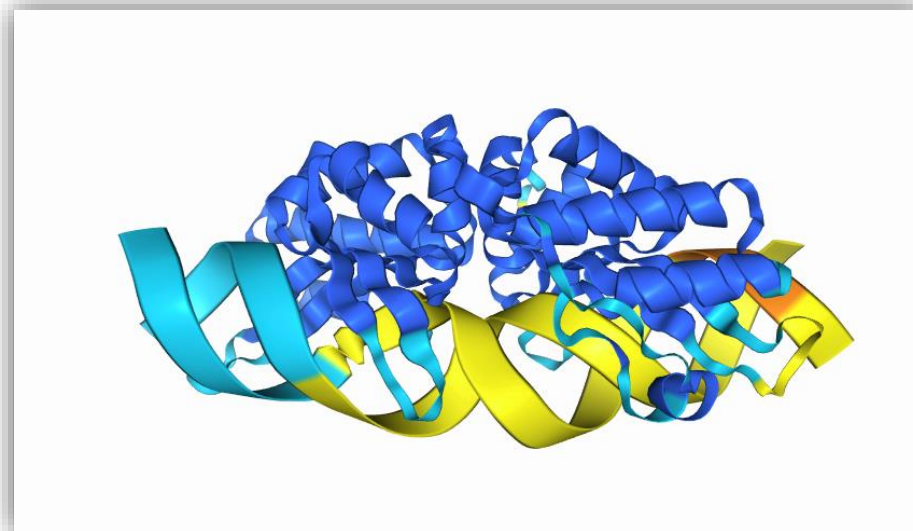
②ここをクリックすると実験的に決められたタンパク質立体構造を見ることができる

正解構造を確認してみる

正解構造



予測構造



PDB情報の確認

PDBjで ID; 7RCEを検索したあとの画面から



7RCE

分子構成を確認する

エンティティ									
Entity ID	鎖名	説明	種類	鎖長	化学式量	分子数	データベース名(アクセス番号)	由来する生物種	エンティティの一般名
1	A	I-OnuI_e-hPD1-d	polymer	300	34983.3	1		Synthetic construct	
2	B	DNA (26-MER)	polymer	26	8064.2	1		Synthetic construct	
3	C	DNA (26-MER)	polymer	26	7917.1	1		Synthetic construct	
4	A	CALCIUM ION	non-polymer		40.1	3	Chemie (CA)		
5	A	SODIUM ION	non-polymer		23.0	1	Chemie (NA)		
6		water	water		18.0	8	Chemie (HOH)		

予測には使わなかったカルシウムイオン、ナトリウムイオンが入っていることがわかる

金属イオンを追加する

AlphaFold Server BETA

[Server](#)

[About](#)

[FAQs](#)

Remainin

AlphaFold Server allows you to model a structure consisting of many biological molecules

- Remaining jobs refresh each day
- Jobs can be up to 5,000 tokens - see more details on token calculation, accepted formats, seed selection and other features in our [FAQ](#)
- ⋮ Use the entity bar to chemically modify proteins and nucleic acids
- 🗨️ Get in touch with the AlphaFold team if you have any questions

Explore these examples of structures to see it in action – try them out without using your quota until you begin editing!

🔗 Protein-RNA-Ion: PDB 8AW3

🔗 Protein-Glycan-Ion: PDB 7BBV

🔗 Protein-DNA-Ion: PDB 7RCE



金属イオンを追加する

Molecule type: Protein, Copies: 1

MASSRRRESIN PWILTGFADE EGSFGLSILN RNRGTARYHT RLSFTIMLHN KDKSILENIO
STWKVGSILN NGDHYVSLVV YRFEDLKVII DHFEKYPLIT QKLGDYKLFK QAFSVMENKE
HLKENGIKEL VRIKAKMNWG LNDELKKAFF ENISKERPLI NKNIPNFKWL AGFTSGDGSF
FVRLRKSNVN ARVRVQLVFE ISQHIRDKNL MNSLITYLGC GHIYEGNKSE RSWLQFRVEK
FSDINDKIIP VFQENTLIGV KLEDFEDWCK VAKLIEEKHH LTESGLDEIK KIKLNMNKGR

Molecule type: DNA, Copies: 1

GGGGGCATGC AGATCCCACA GGCGCG

Molecule type: DNA, Copies: 1

CCGCGCCTGT GGGATCTGCA TGCCCC

Molecule type: Ion, Copies: 3

Ca²⁺

Molecule type: Ion, Copies: 1

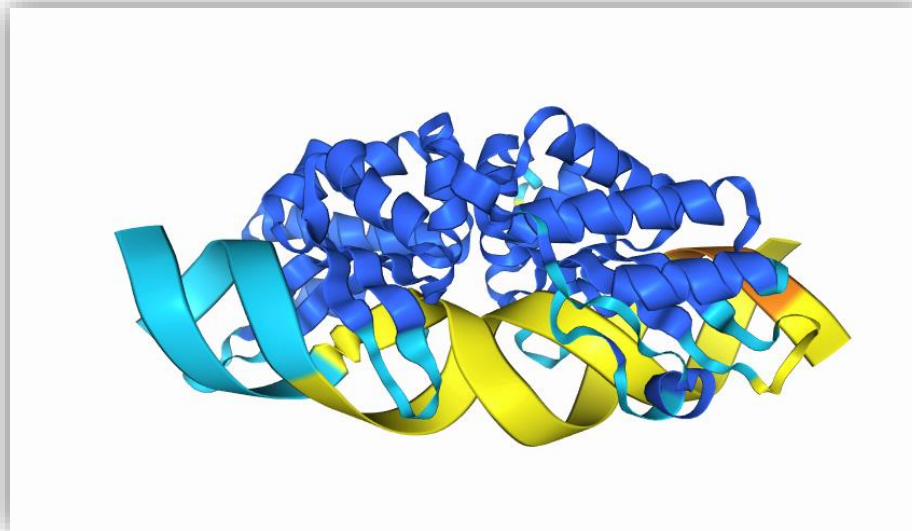
Na⁺

1回目と同様に予測を実行する。

AlphaFoldがデモ用に提示しているデータなので、計算はすぐ終わる

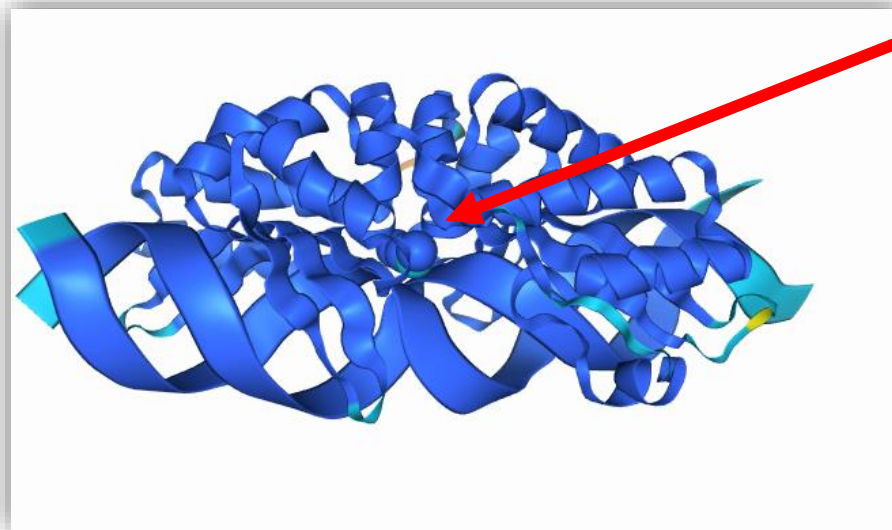
金属イオンを追加する

1回目
金属イオンなし



2回目
金属イオンあり

構造はほぼ同様だが、
予測精度が上がっている



この辺りに金属イオンが入っている。
これによりDNAの構造が”AIにとって”
“違和感のないもの”
になった。

つまり金属イオンの
必要性を予測している！