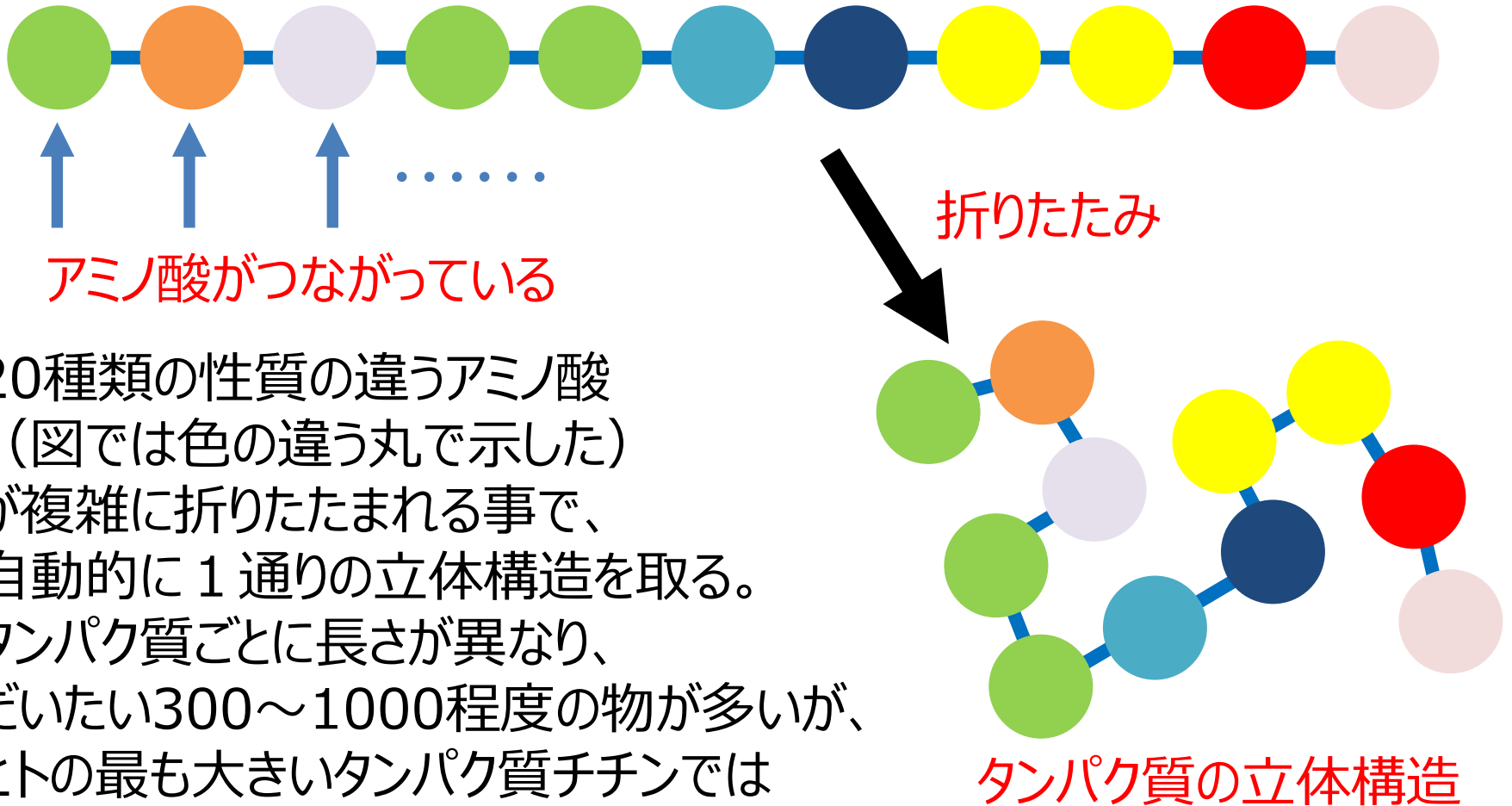


AlphaFold3の使い方

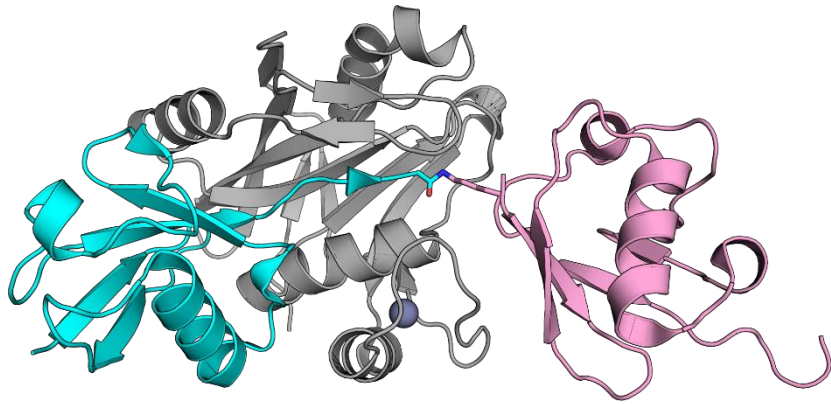
文責：鳥取大学工学部・永野研究室・准教授
佐藤裕介

一本のタンパク質が折りたたまれることで 仕事ができるようになる

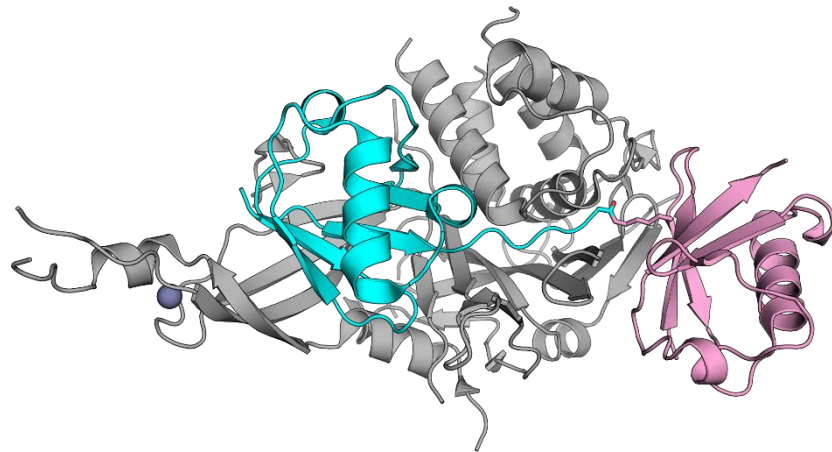


20種類の性質の違うアミノ酸
(図では色の違う丸で示した)
が複雑に折りたたまれる事で、
自動的に1通りの立体構造を取る。
タンパク質ごとに長さが異なり、
だいたい300～1000程度の物が多いが、
ヒトの最も大きいタンパク質チチンでは
3万4350個のアミノ酸が、
これまで発見された最大のタンパク質PKZILLA-1は
4万5212個のアミノ酸が連結してできている。

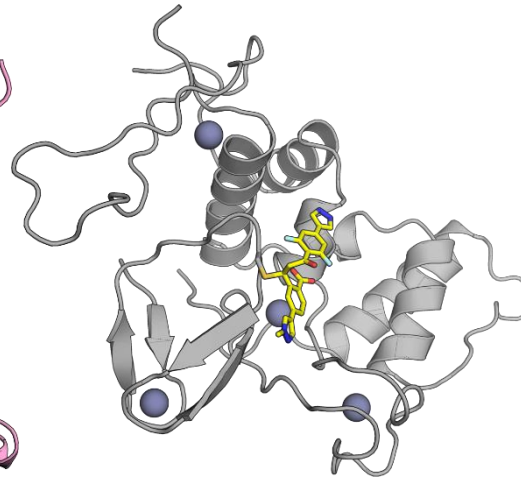
実際に構造が解明された様々なタンパク質



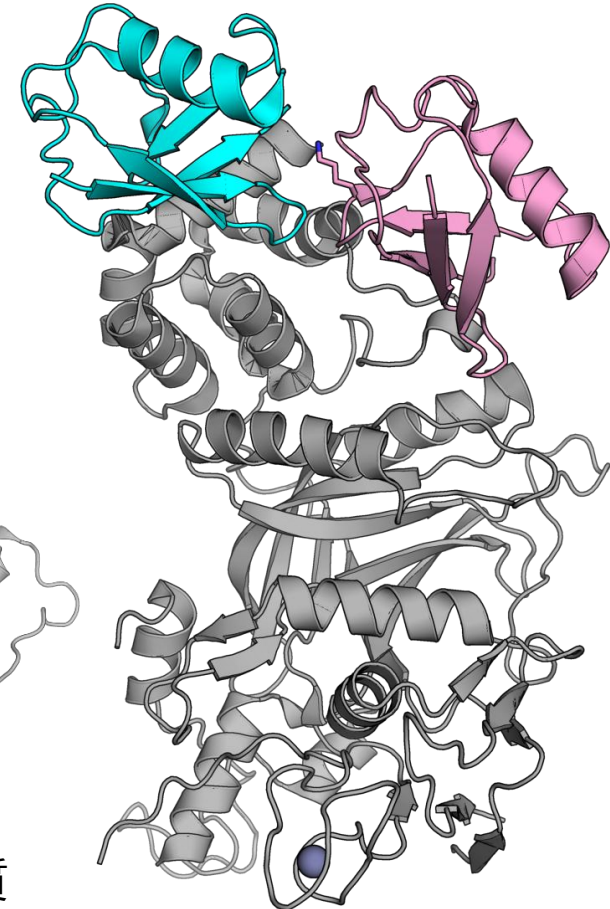
タンパク質の輸送に関与するタンパク質



アルツハイマー病に関与するタンパク質



免疫反応に
関与するタンパク質



不要なタンパク質除去に
関連するタンパク質

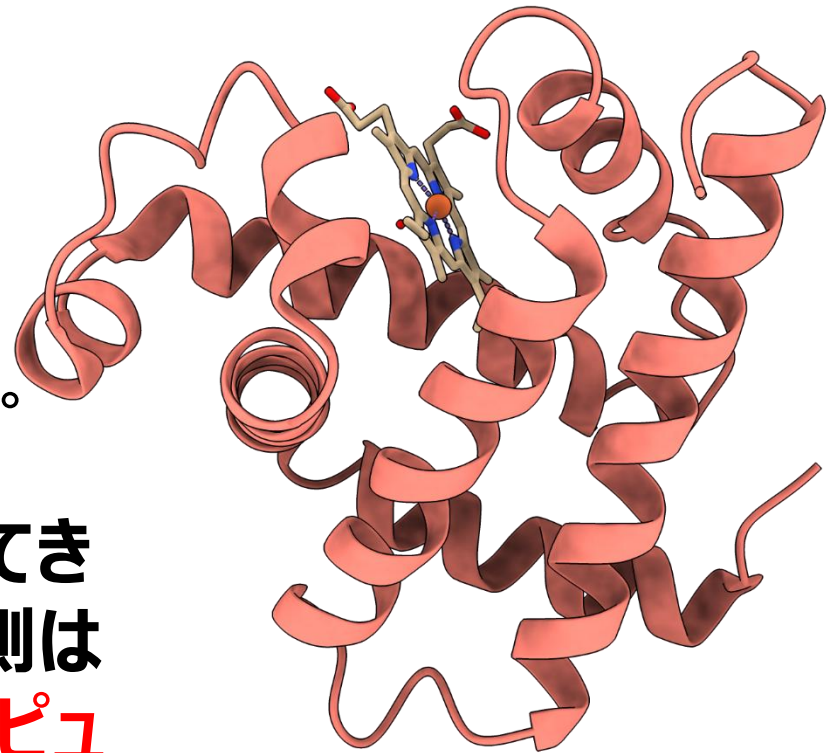
タンパク質の構造を観察することで、
タンパク質の機能を詳しく知ることができる

タンパク質の構造を知るには

1958年にX線結晶構造解析という方法で、153個のアミノ酸がつながってできた、ミオグロビンというタンパク質の構造が世界ではじめて解明されました。

それ以来、アミノ酸がどのような順番でつながると、どのようなタンパク質の構造になるのか、という事がずっと研究されてきました。

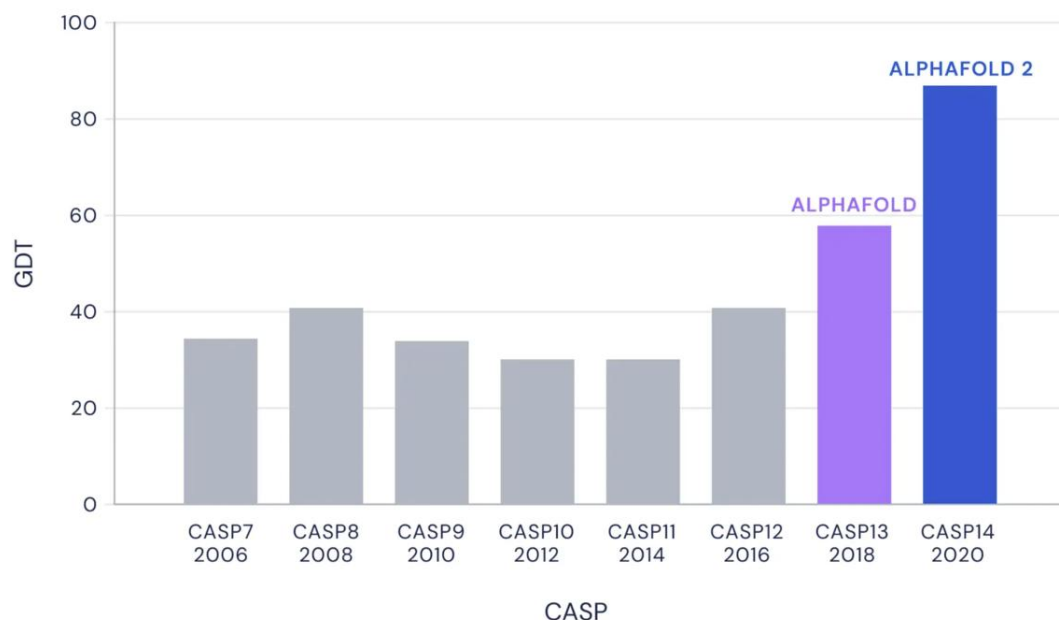
ところが、50年以上研究されてきたにも関わらず、立体構造予測は極めて難しく、**超高性能のコンピューターを使用してもほとんど成功していませんでした。**



ミオグロビンの構造

2020年、DeepMind社によってAlphaFold2が開発された

※DeepMind社は囲碁棋士に初めて勝利した囲碁ソフトAlphaGoでも有名



<https://www.itmedia.co.jp/news/articles/2012/01/news053.html>

しかし、2年に1回開かれるタンパク質構造予測のコンテスト(CASP)で2018年にAlphaFoldが、2020年にその後継のAlphaFold2がぶっちぎりの1位を獲得すると、事態は一変しました。上の図では、棒グラフの高さが構造の予測の正確さを示していますが、縦軸の90を越すと、ほぼ正確な構造と言えます。

AlphaFold2はついにほぼ正確な構造の予測に成功したのです。

2024年、AlphaFoldの開発者デミス・ハサビス氏と
ジョン・ジャンパー氏は、「タンパク質の構造予測」
の業績によりノーベル賞を受賞!!

ノーベル賞の受賞対象となる化学的発見は、
その重要性が十分に検証されてから
授与されるので通常数十年程度かかります。

しかし、AlphaFoldはAlphaFold2の発表から
わずか4年でノーベル賞を受賞しており、
精度の高いタンパク質構造予測が
いかに重要であることを示しています。

そして、2024年には、AlphaFold2を超えた性能の
AlphaFold3が無料で誰でも使えるようになりました!!
つい最近に開発され、ノーベル賞を受賞したプログラムを
誰でも使えるというのは、本当にすごい時代になりました

AlphaFoldの使い方

AlphaFoldの予測結果を見るためには2つの方法があります。

1. AlphaFold Protein Structure Database (AlphaFold PDB)に掲載されたデータを見る

良いところ: DeepMind社がすでに予測した構造情報が掲載されているので、自分で予測しなくてもすぐに予測結果を観察できる。これまでゲノム解析されたほとんどの生物のタンパク質構造をすぐに見ることができる。

悪いところ: タンパク質は複数が組み合わさって働くが、単体の構造しか登録されていないため、どのように組み合わさるのが予測できない。

2. 自分でAlphaFold3を走らせる

良いところ: アミノ酸配列さえわかれば、タンパク質の構造予測が可能。タンパク質の組み合わさった形も予測が可能。

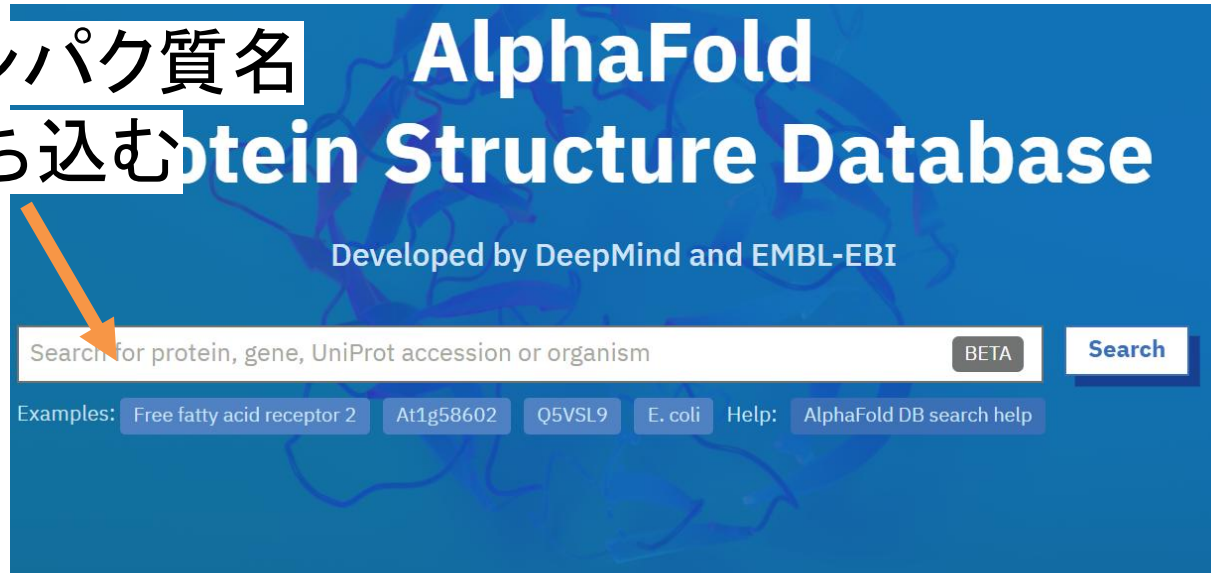
悪いところ: 10-30分くらいかかる。

1. AlphaFold PDBの使い方

<https://alphafold.ebi.ac.uk/>

↑のURLへ移動する

ここにタンパク質名
を打ち込む



AlphaFold Protein Structure Database

Developed by DeepMind and EMBL-EBI

Search for protein, gene, UniProt accession or organism

BETA Search

Examples: Free fatty acid receptor 2 At1g58602 Q5VSL9 E. coli Help: AlphaFold DB search help

AlphaFold DB provides open access to protein structure predictions for the human proteome and 20 other key organisms to accelerate scientific research.

ヒトの5万種のタンパク質ほぼすべて+ほとんどの生物種の構造を片っ端から予測し掲載してある。

『Npl4 human』で検索した場合の例

Showing all search results for Npl4 human

1 - 20 of 20313 results

クリック

Filter by:

Organism

- ☐ Homo sapiens (20294)
- ☒ **Danio** rerio (4)
- ☐ Mus musculus (3)
- ☐ Rattus norvegicus (3)
- ☐ Arabidopsis thaliana (1)
- ☐ Caenorhabditis elegans (1)
- ☐ Candida albicans (strain SC5314 / ATCC MYA-2876) (1)
- ☐ Dictyostelium discoideum (1)

Nuclear protein localization protein 4 homolog

Q8TAT6 (NPL4_HUMAN)

Protein Nuclear protein localization protein 4 homolog

Gene NPLOC4

Source Organism Homo sapiens [search this organism](#)

UniProt Q8TAT6 [go to UniProt](#)

NPL4 domain-containing protein

A4I1H3 (A4I1H3_LEIIN)

Protein NPL4 domain-containing protein

Gene LINJ_25_1320

様々な生物種が登録されているので、
タンパク質名 + 生物種名 で検索するのが良い。
ここでは、ヒトのNpl4というタンパク質の構造を検索している。
自分でも興味のあるタンパク質の名前を検索してみましょう。

構造の予測結果はすぐに見ることができる

3D viewer ②

Model Confidence:

- Very high (pLDDT > 90)
- Confident (90 > pLDDT > 70)
- Low (70 > pLDDT > 50)
- Very low (pLDDT < 50)

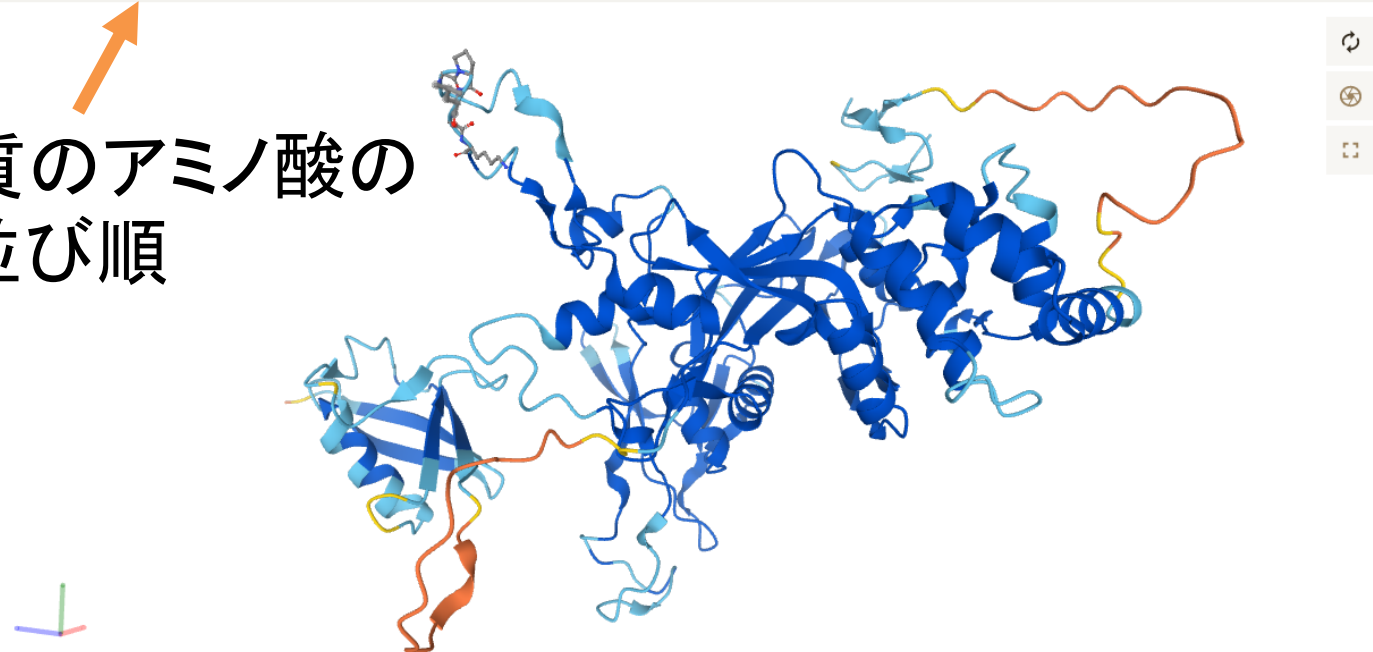
AlphaFold produces a per-residue confidence score (pLDDT) between 0 and 100. Some regions below 50 pLDDT may be unstructured in isolation.

構造の
信頼性
の凡例

Sequence of AF-Q8TAT6-F1 1: Nuclear pro... A

251 261 271 281 291 301 311 321 331 341 351 361
GNQHFG YLYGRYTEHKDIPLGIRAEVAAIYEPPQIGTQNSLELLED PKAEVVDEIAAKLGLRKVGW IFTDLVSEDT RKGTVRYSRNKDTYFLSSEE CITAGDFONKHPNMCRLSPD GHFGSK
371 381 391 401 411 421 431 441 451 461 471 481
FVTA VATGGFDNQVHFEGYQVSNQCMALVRDECLPCKDAPELG YAKESSEQYVPDVFYKDVDF KFGNEITQLARPLPVEYLII DITTTFPKDPVYTFISQNP FPIENRDVLGETQDFHSL
491 501 511 521 531 541 551 561 571 581 591 601
AT YLSQNTSSVPLDTISDFHLL LFLVTNEVMPLQDSISLLE AVRTRNEELA QTWKRSEQWATIEQLCSTVG GQLPGLHEYGA VGGSTHTAT AAMWACQHCT FMNQPGTGHC EMCSLPRT

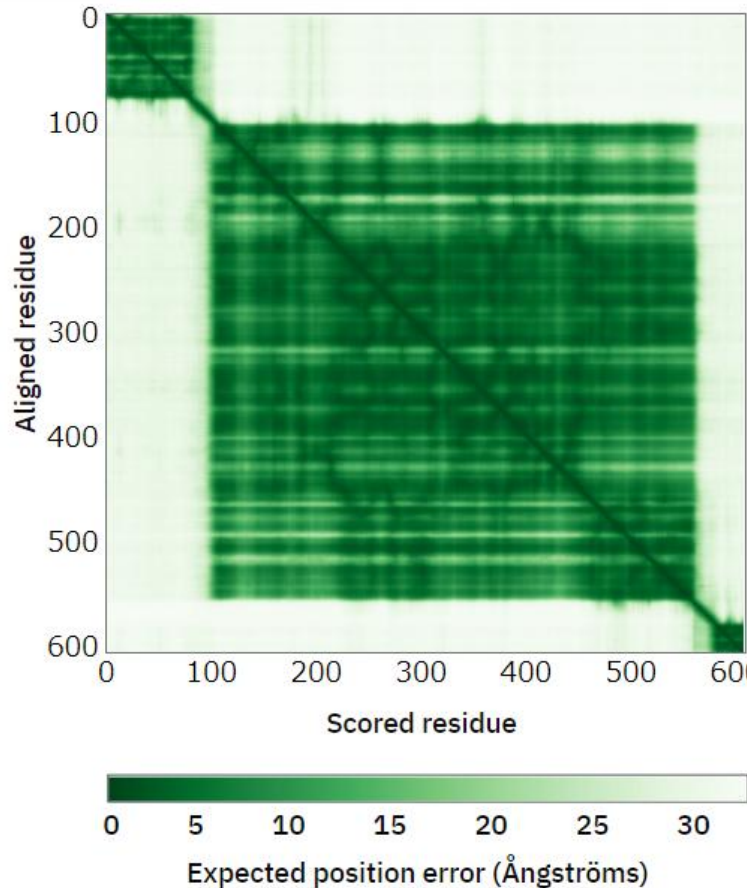
タンパク質のアミノ酸の
並び順



構造を左クリックで回転、右クリックで移動、
ホイールで拡大・縮小ができる。

青いところは構造予測結果の信頼性が高いと考えられるところ、
オレンジのところは信頼性が低いと考えられるところ。

PAEの見方 (難しければとばしてOK)



残基間の位置関係は
Predicted aligned error (PAE)で
その信頼性が予測されている。

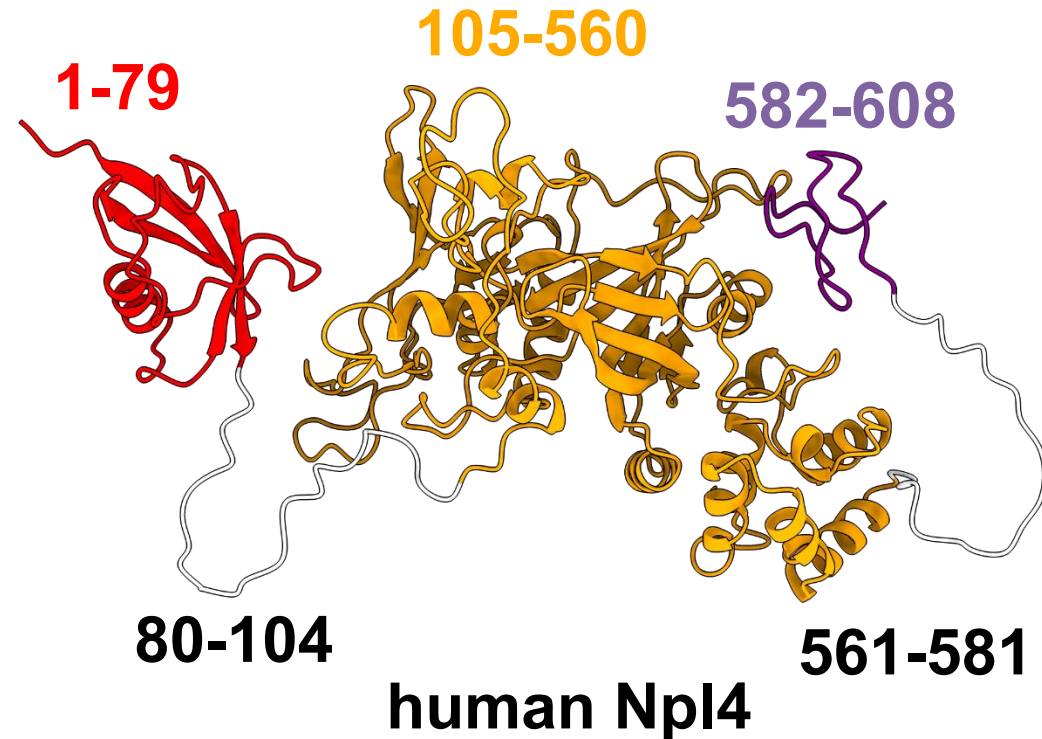
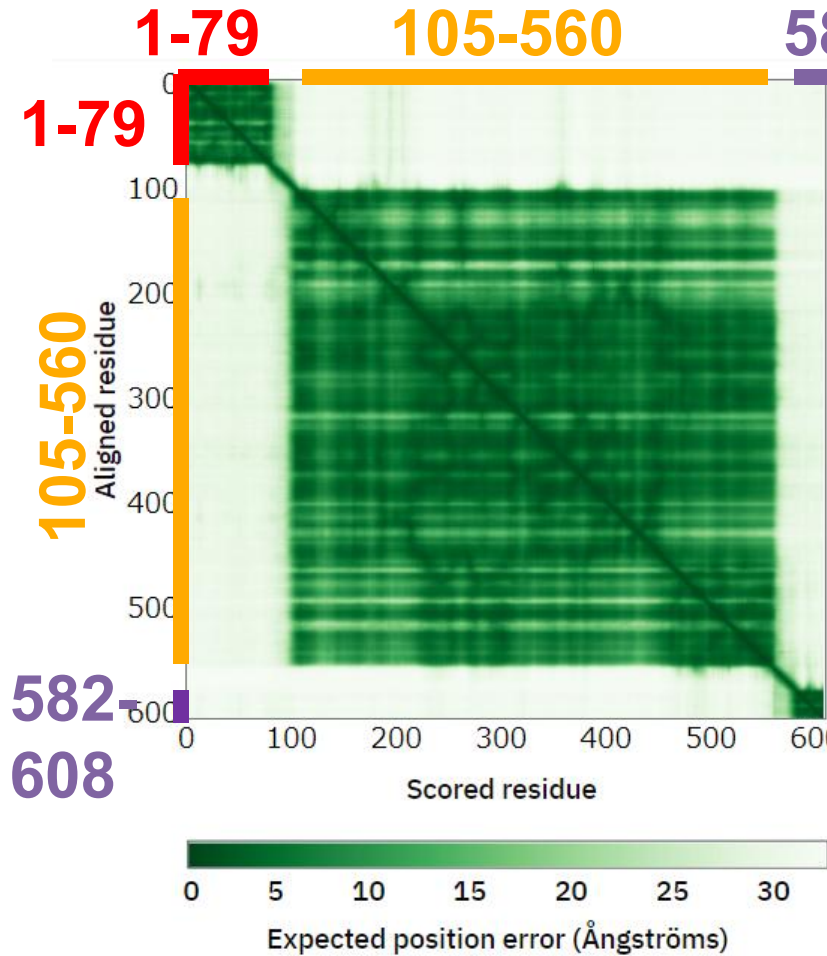
縦/横軸はともにタンパク質の
アミノ酸番号を意味する。

**x番目の残基とy番目の残基の
位置関係のずれを予測**している。

$x = y$ の場合、位置関係は同じ場所になる
ので、ずれるはずがない。
つまり、常に予測からのずれは
 0 Å となっている

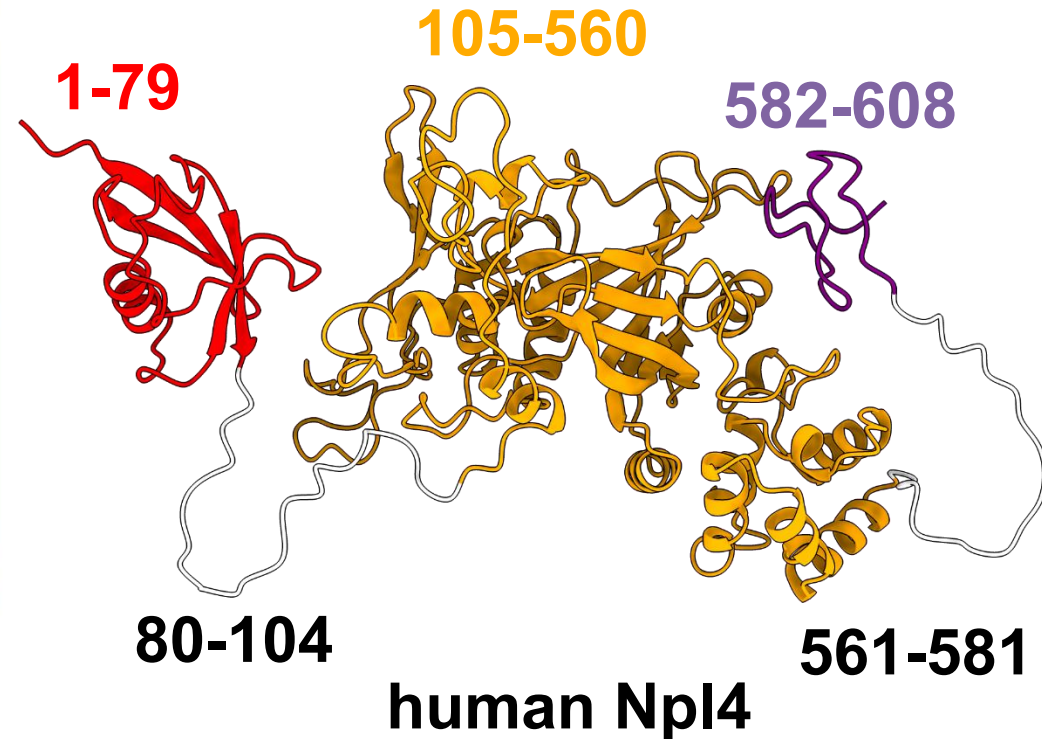
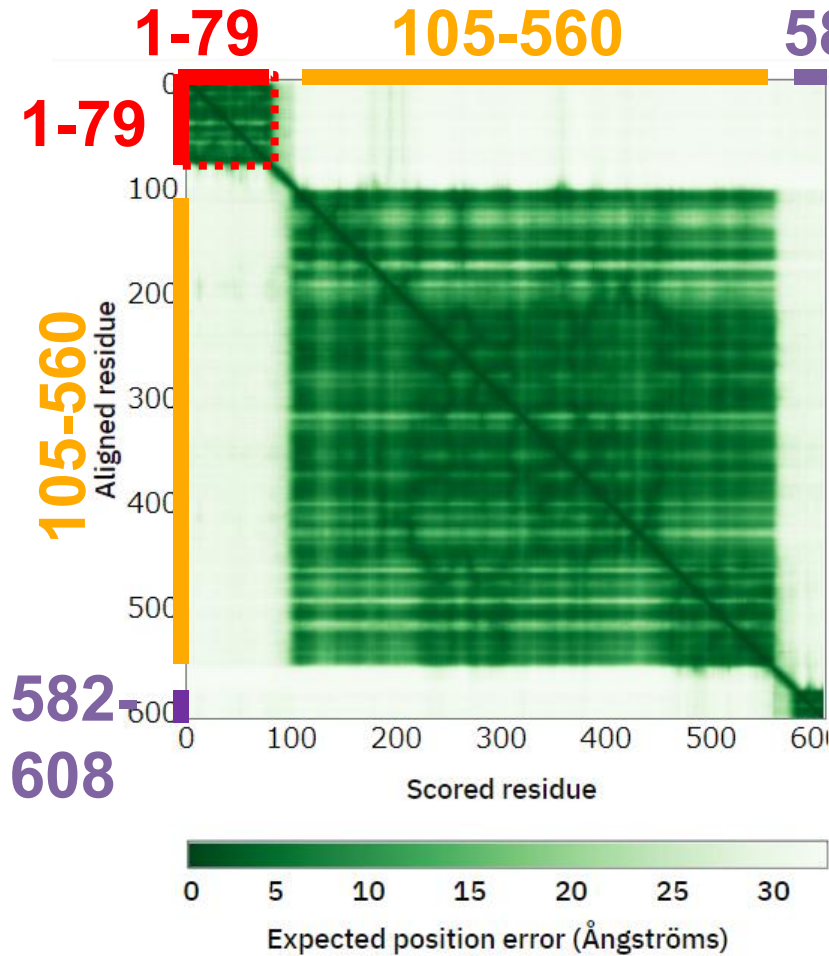
大事なのは、 $x = y$ ではないところの色

PAEの見方 (難しくればとばしてOK)



AlphaFold構造を見慣れない人がこの構造を見ると、**105-560の領域**と**582-608の領域**が結合しているように感じてしまう。しかし、今から説明するPAEを理解すれば、両者の位置関係はまったく信用できない、つまり、結合していない事がわかる。

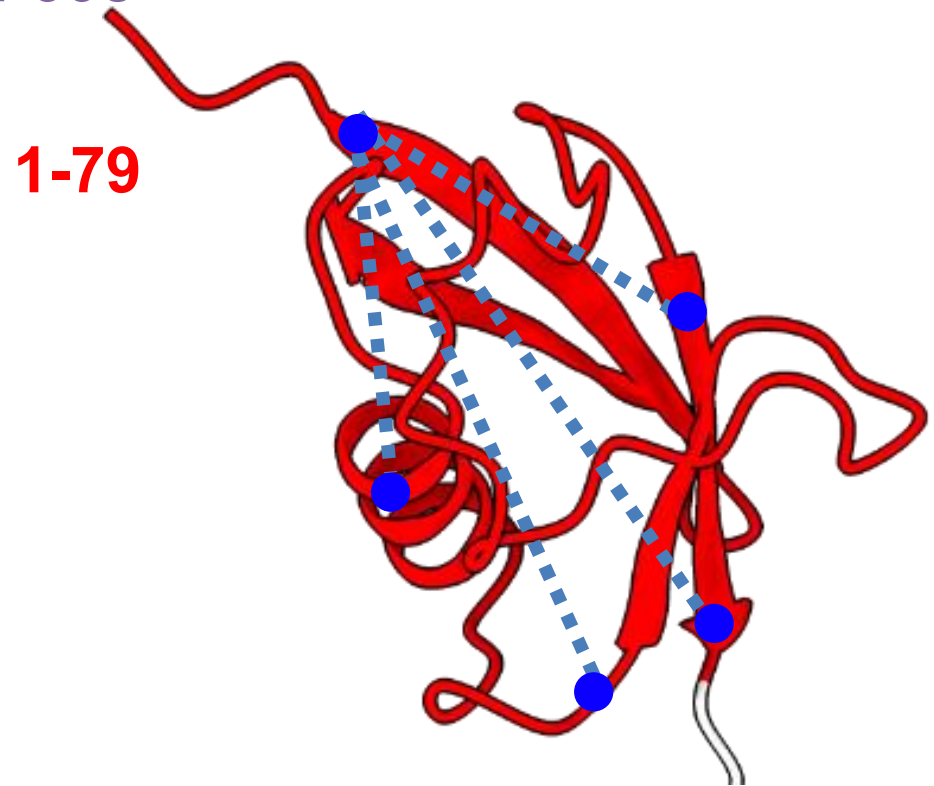
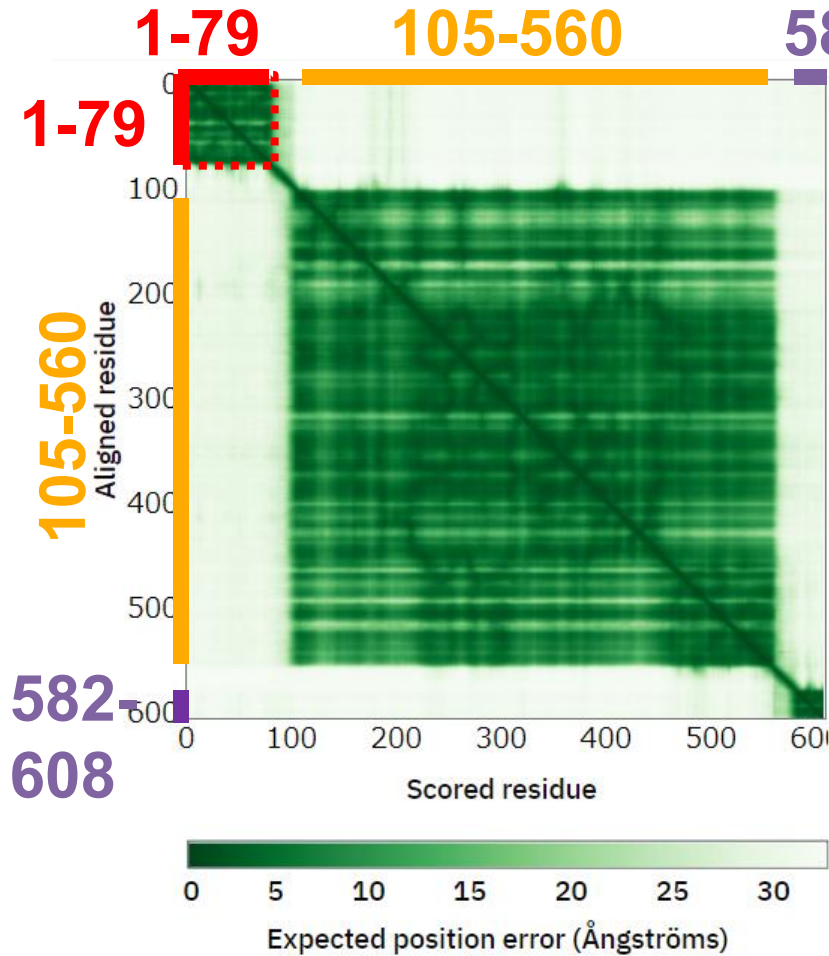
PAEの見方 (難しくればとばしてOK)



まず、**1-79番目**の残基を見てみると、**1-79番目**の残基同士ではその位置関係のずれの予測は小さい（濃い緑色）

つまり、**1-79**では各残基の位置関係が正しいと予想されている。

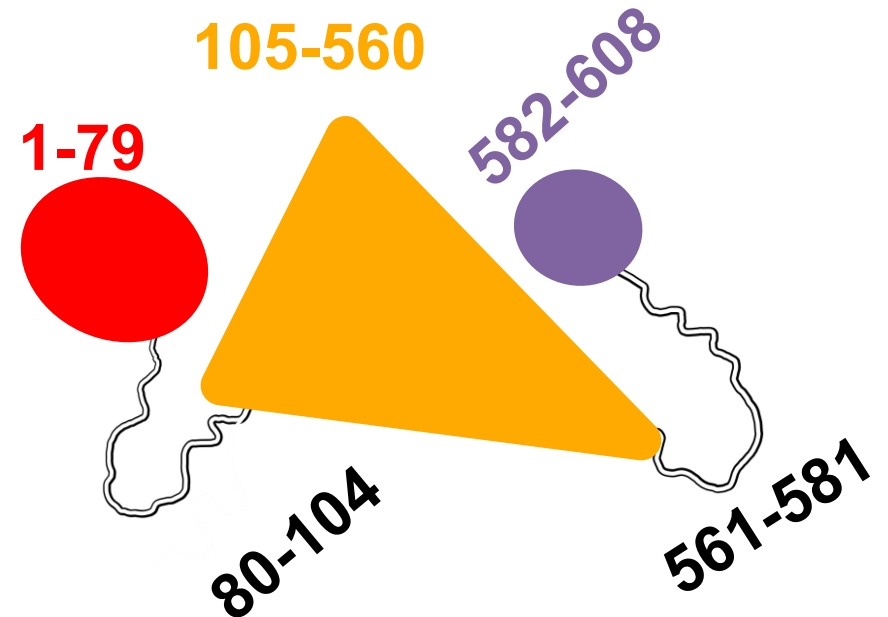
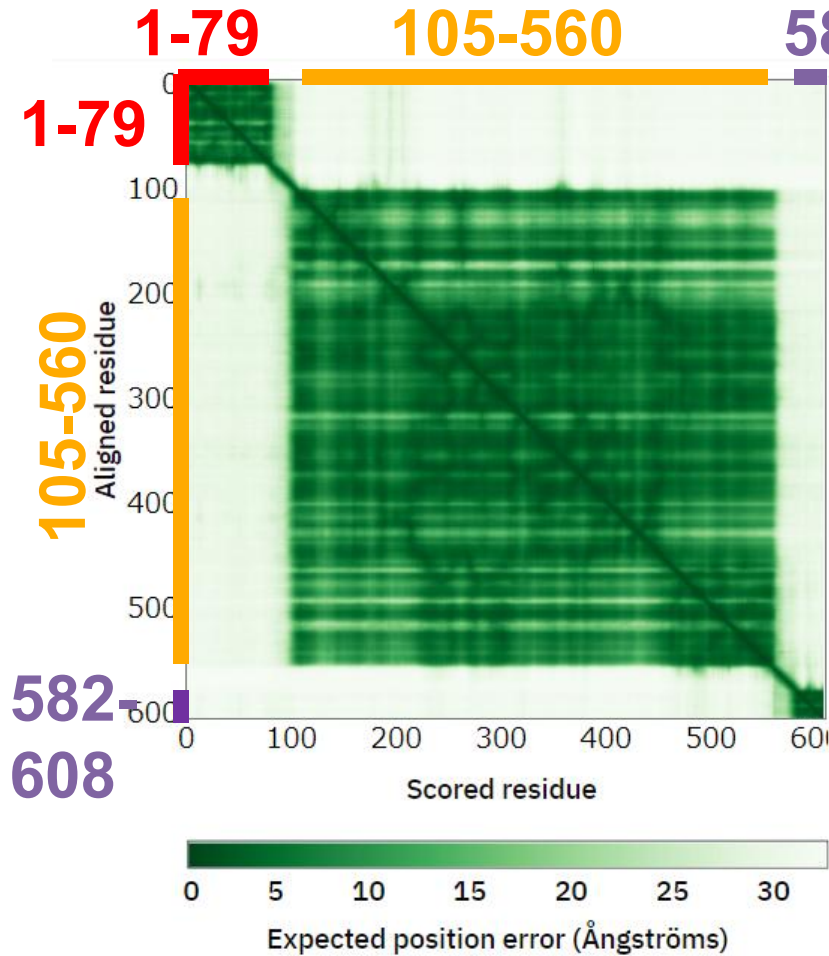
PAEの見方 (難しくればとばしてOK)



ちょっとわかりづらいかもなので補足説明。

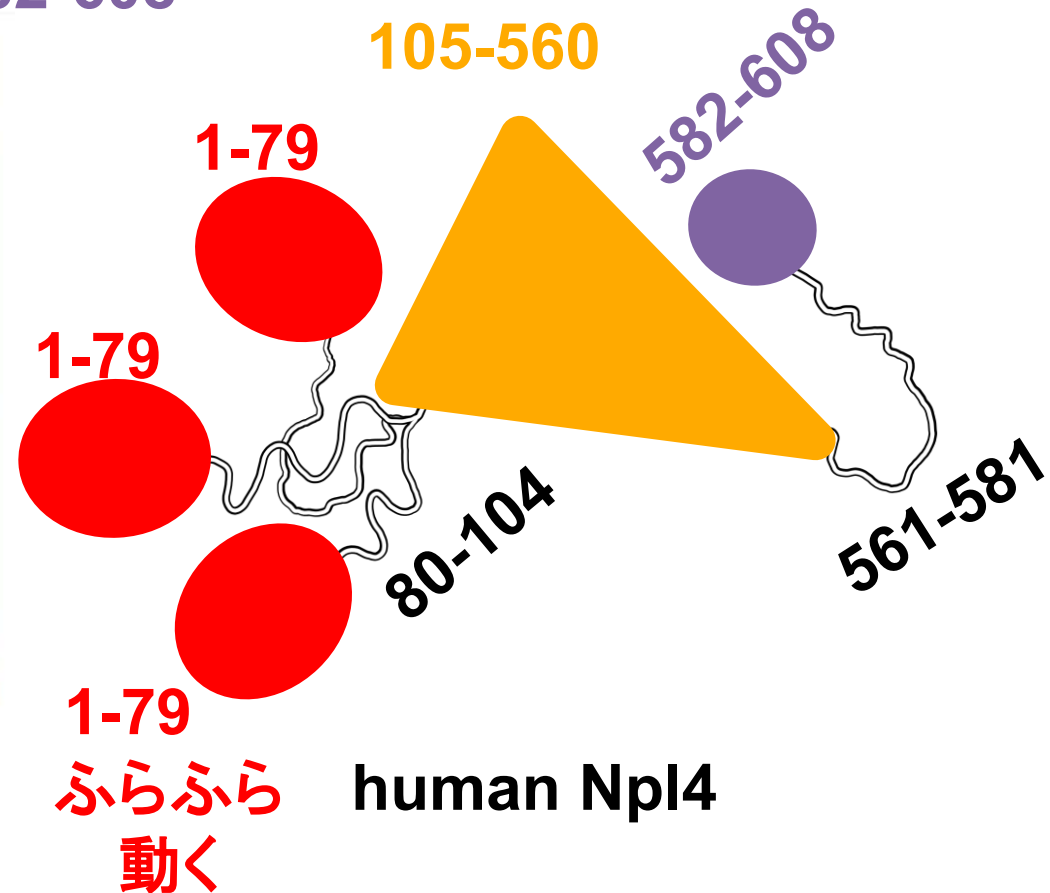
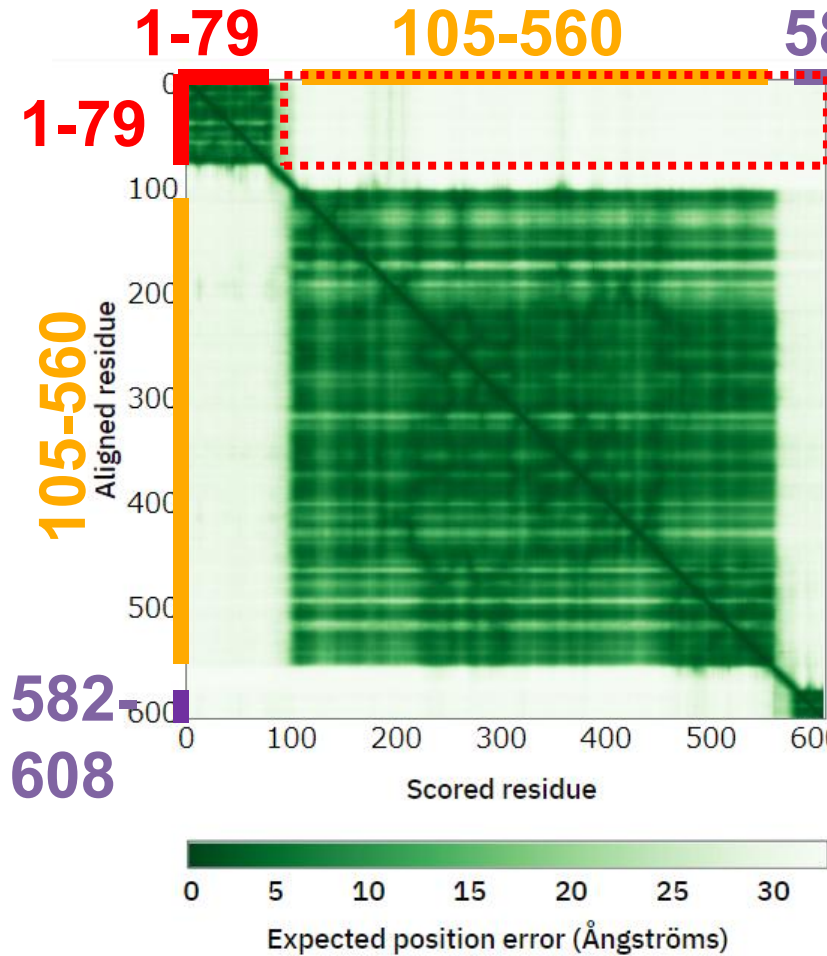
PAEに**正方形**が現れた場合、1-79の任意の2つの地点の位置関係は常に予測の信頼性が高いという意味。位置関係がずれたりしないという事は、**きっちりと折りたたまれた部分(=ドメイン)と見なすことができる**という意味。

PAEの見方 (難しければとばしてOK)



ここから、説明をわかりやすくするため、
タンパク質の予測構造を模式図で示していく

PAEの見方 (難しければとばしてOK)

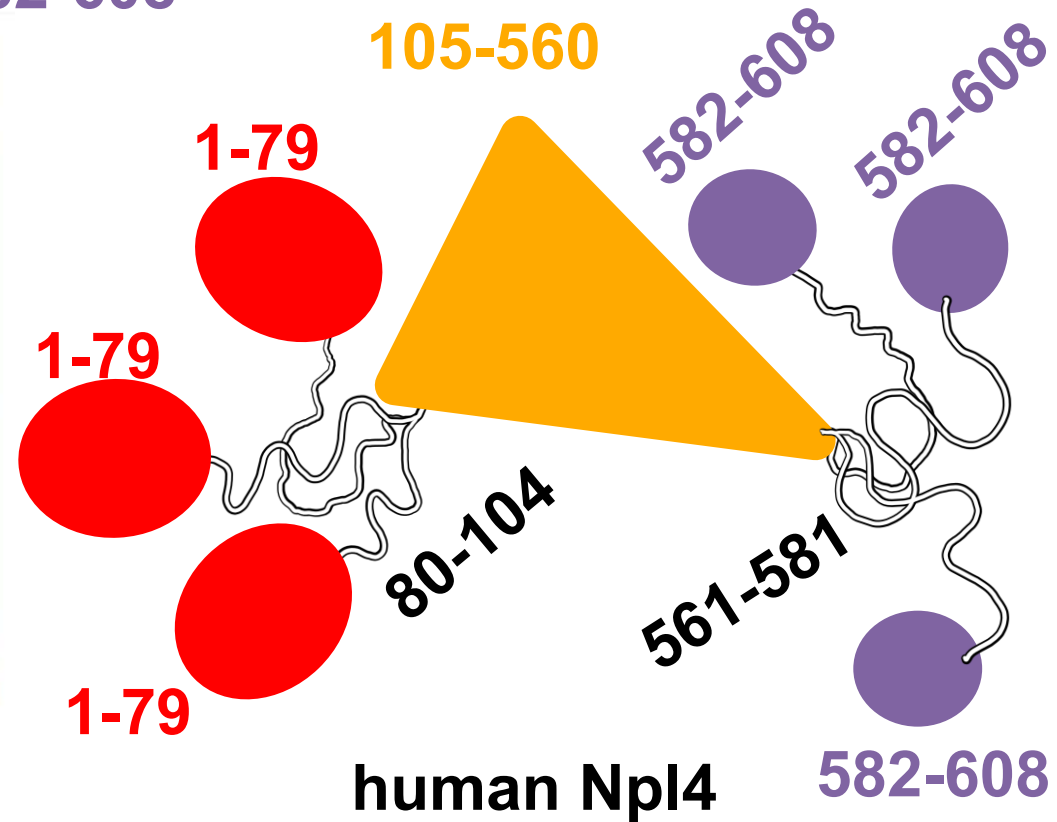
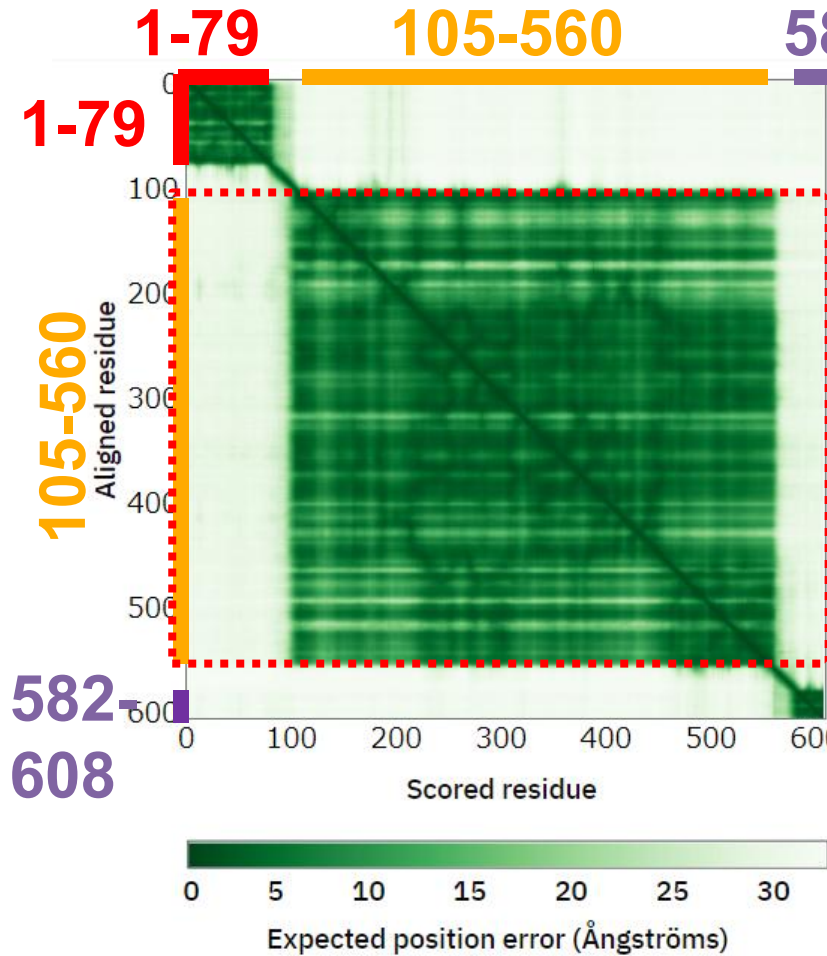


もう一度**1-79番目**の残基を見てみると、**80-600**の残基に対し
その位置関係のずれの予測は大きい (PAEが白色で表示されている)

つまり、**1-79**は他の領域との位置関係が信頼できない。

これは、**1-79**は他の領域に対してふらふらと動く事を意味する。

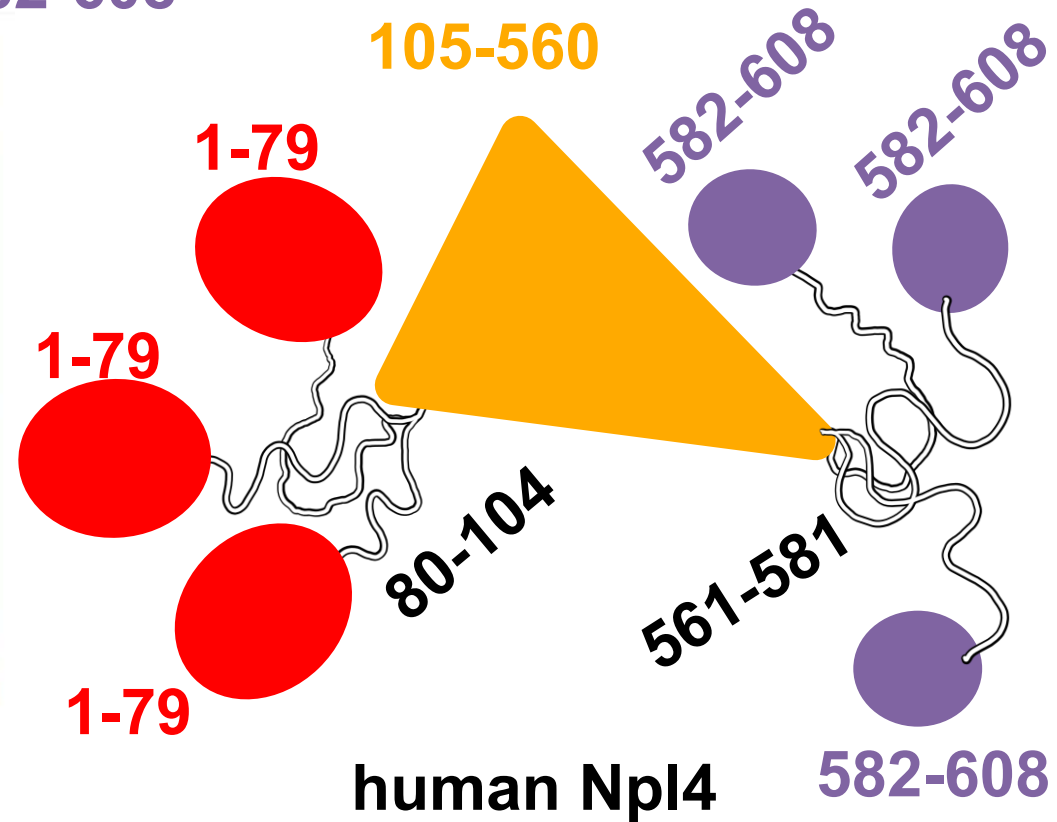
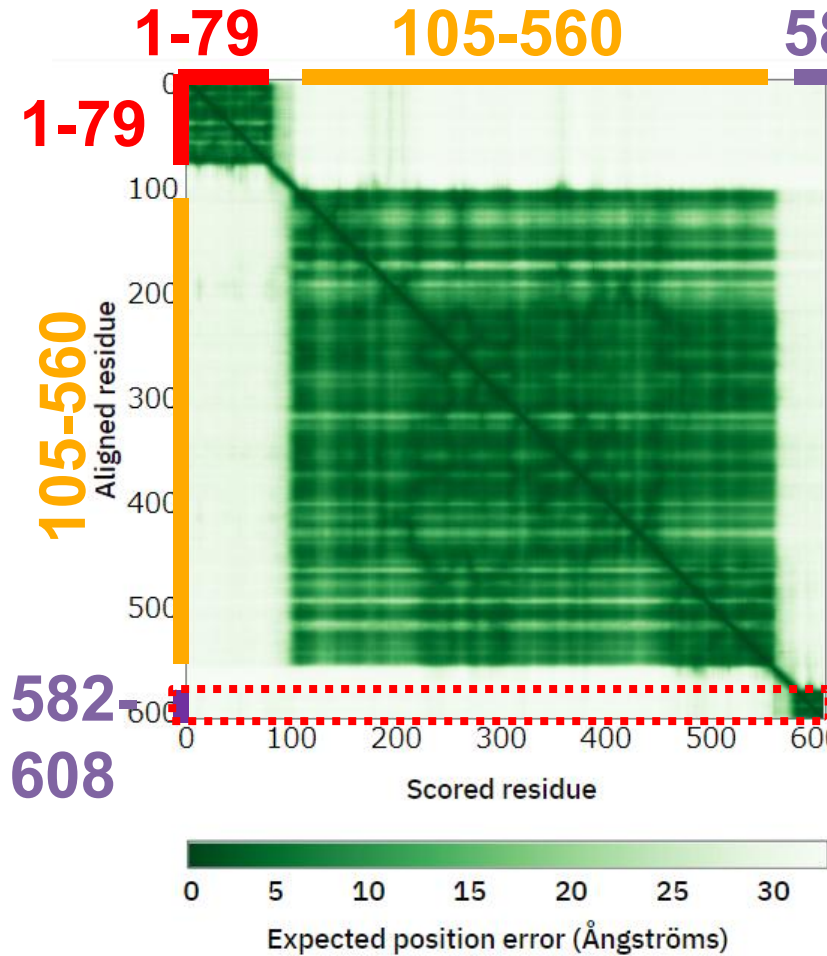
PAEの見方 (難しければとばしてOK)



同様に**105-560**の領域を見てみると、**105-560**の領域同士でのみ、その位置関係のずれの予測は小さい（濃い緑色）

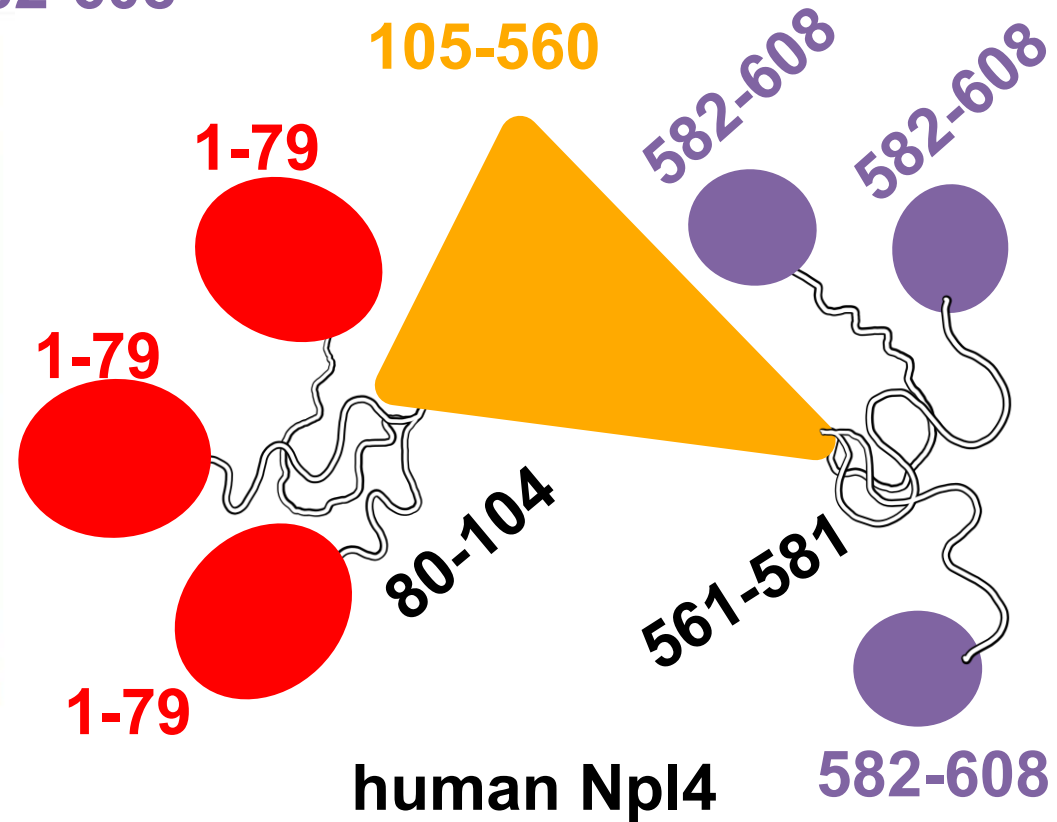
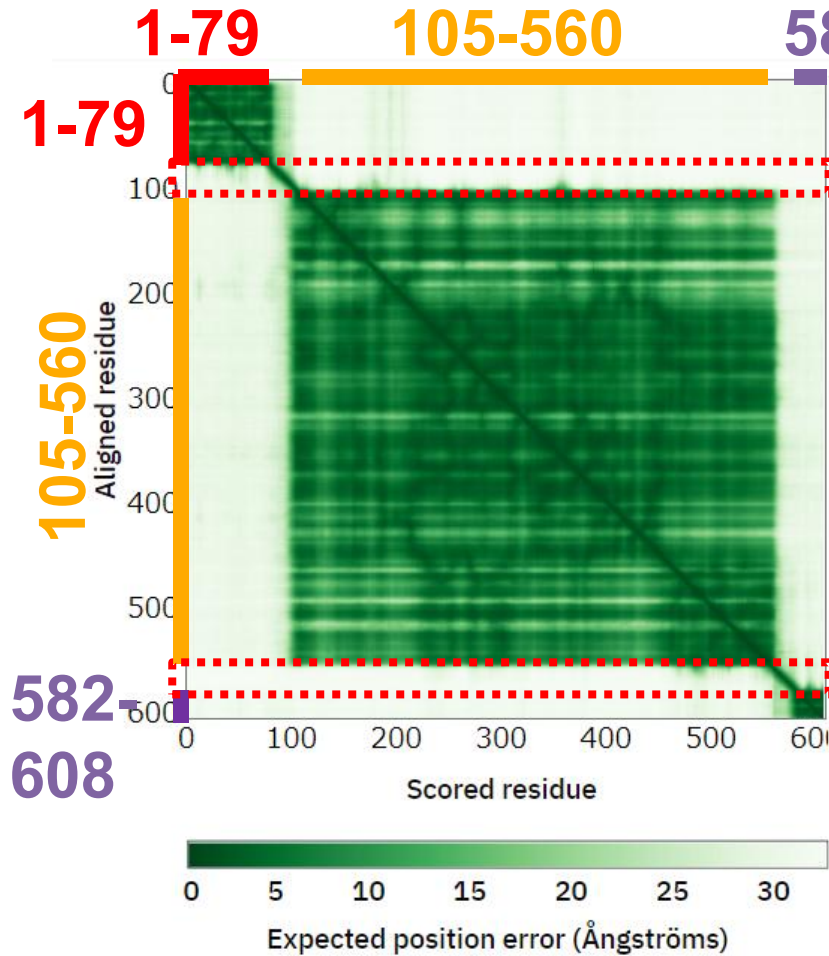
つまり、**105-560**はひとかたまりのドメインとして予測されているが、他の領域との位置関係は信頼性にかけ、ふらふらと動くと予測されている。

PAEの見方 (難しくればとばしてOK)



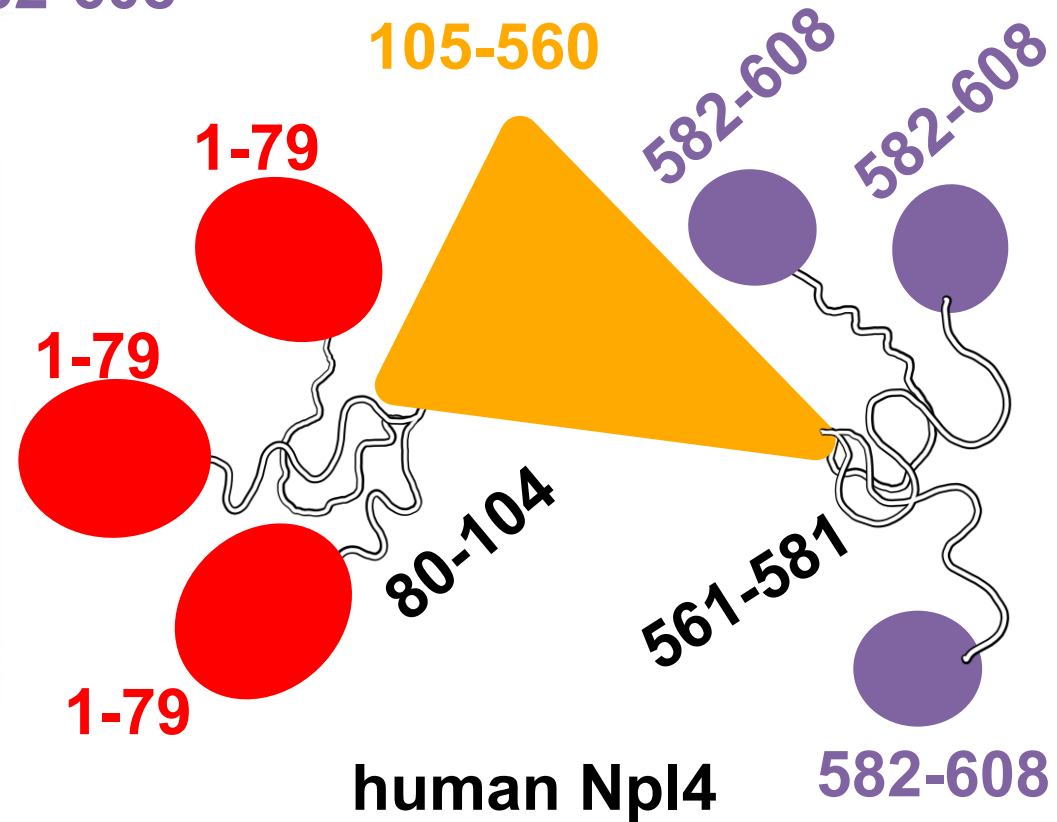
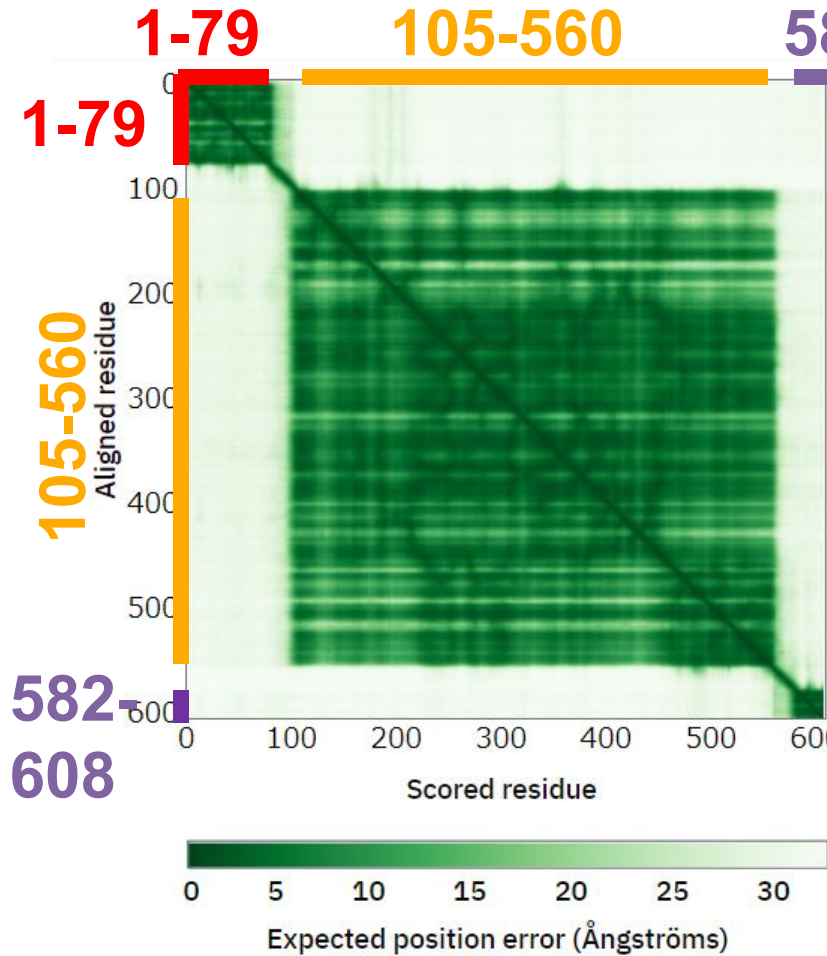
同様に**582-608**の領域を見てみると、**582-600**の領域同士でのみ、その位置関係のずれの予測は小さい（濃い緑色）
つまり、**582-608**はひとかたまりのドメインだが、他の領域との位置関係は信頼性に関わらず、ふらふらと動くと予測されている。

PAEの見方 (難しくればとばしてOK)



最後に、**80-104**や**561-581**の領域を見てみると、 $x = y$ 以外の領域に対して、全てずれの予測が大きい。このような場合、他の領域に対しての位置関係はすべて信頼性にかかけ、構造をとらない**紐状の領域**と予測される。

PAEの見方 (難しくればとばしてOK)



以上をまとめると、

Npl4は3つのドメイン構造 (**1-79**, **105-560**, **582-608**) が
柔らかな紐状の部分 (**80-104**, **561-581**) で連結されている。
紐状の部分が柔らかいので、ドメイン構造同士はフラフラと動くと考えられる

2. 自分でAlphaFold3を走らせる

**Googleアカウント(無料)があれば、
誰でも無料でインストールなしで
AlphaFold3を走らせる事ができます。**

**※GoogleアカウントはG-mailを使っている人は
持っているはずです。これも簡単に登録できます。**

**なお、今回紹介する方法では実際の計算は
ネットで接続したGoogle DeepMind社の
AlphaFold Serverで行うので、
使用するパソコンのスペックは
低くても問題ありません**

<https://alphafoldserver.com>

↑のURLで、AlphaFold serverを使えます

AlphaFold Serverの初回起動時には登録が必要

<https://alphafoldserver.com/>

に移動したら、
ここをクリック。

Googleアカウントを選択するように
言われるので、アカウントを選択。
アカウント毎に結果が記録されるので、
どのアカウントで予測を行ったのかは
覚えておくこと。



色々と英語で説明されるので、チェックボックスにチェックを入れつつ、
Continueというアイコンをクリックしていく

適当にポチポチしていけば、勝手に登録されるはず

最低限これだけやれば、高精度の構造予測が可能！

AlphaFold Server BETA

[Server](#) [About](#) [FAQ & Guides](#) ▾

Remaining jobs: 8

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

[Learn more](#) ^

- 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

Ok, got it

複合体中に
含まれる数

ここに配列を
入力

ここで、protein, DNA
ligandなどを指定

ここで配列の
種類を増やす

Entity type
Protein ▾

Copies
1

>Paste sequence or fasta
Input

This field is required

Entity type
Protein ▾

Copies
1

>Paste sequence or fasta
Input

This field is required

+ Add entity

入力できたらここをクリック

Continue and preview job

Save job

入力したあとの状態は、3 ページ後に示しました

なお、アミノ酸配列はUniprotから持ってくると良い

<https://www.uniprot.org/>

Find your protein

UniProtKB • human NPL4 Advanced List Search

Examples: Insulin, APP, Human, P05067, organism_id:9606

①調べたいタンパク質の名前や、
どの生物のタンパク質なのか、
といった情報を入力して、Search
今回は、human NPL4を検索

Function **Q8TAT6 · NPL4_HUMAN**

Names & Taxonomy Protein¹ Nuclear protein localization protein 4 homolog Amir

Subcellular Location Gene¹ NPLOC4 Protein exi

Disease & Variants Status¹ UniProtKB reviewed (Swiss-Prot) Annotator

PTM/Processing Organism¹ Homo sapiens (Human)

Expression

Interaction

Structure

Family & Domains **Sequence & Isoform** Similar Proteins

Function¹
The ternary complex containing UFD1, VCP and NPLOC4 binds ubiquitinated proteins and is necessary for the export of misfolded protein proteasome. The NPLOC4-UFD1-VCP complex regulates spindle disassembly at the end of mitosis and is necessary for the formation of a

③個別の結果のページに移動したら、
左のタブのSequence & isoform
をクリック

UniProtKB 351 results

Entry	Accession	Protein Name	Gene Name	Organism	Length
Q8TAT6	NPL4_HUMAN	Nuclear protein localization protein 4 homolog	NPLOC4	Homo sapiens (Human)	608 AA
Q97733	NPL4_HUMAN	Nucleosome assembly protein 1-like 4	NAP1L4	Homo sapiens (Human)	271 AA
P0C949	NPL4_HUMAN	Nuclear protein localization protein 4	NPL4	Cryptosporidium parvum var. meleagridis (Cryptosporidium parvum)	608 AA
A48138	NPL4_HUMAN	Nuclear protein localization protein 4	NPL4	Alphaproteobacteria (strain N162) (Rhodospirillum rubrum)	602 AA
P0C951	NPL4_HUMAN	Nuclear protein localization protein 4	NPL4	Cryptosporidium parvum var. meleagridis (Cryptosporidium parvum)	608 AA
Q10714	NPL4_HUMAN	Nuclear protein localization protein 4	NPL4	Candida albicans (strain ATCC 90026) (Candida albicans)	607 AA
Q10715	NPL4_HUMAN	Ubiquitin receptor factor 1 (UBR1-associated degradation protein 1)	UBR1	Homo sapiens (Human)	207 AA
Q9H4C2	NPL4_HUMAN	Small GTP-binding protein	SRP	Homo sapiens (Human)	77 AA
Q9H4C3	NPL4_HUMAN	FAH-associated factor 1	FAH1	Homo sapiens (Human)	441 AA
Q9H4C4	NPL4_HUMAN	UBR1 domain-containing protein 7	UBR1	Homo sapiens (Human)	491 AA
P50072	NPL4_HUMAN	Translocated endoplasmic reticulum chaperone 1	CHP1	Homo sapiens (Human)	604 AA
Q9H4C5	NPL4_HUMAN	FAH-associated factor 1	FAH1	Homo sapiens (Human)	441 AA
Q43323	NPL4_HUMAN	UBR1 domain-containing protein 1	UBR1	Homo sapiens (Human)	297 AA
Q9H4C6	NPL4_HUMAN	E3 ubiquitin-protein ligase RNF125	RNF125	Homo sapiens (Human)	232 AA
Q9H4C7	NPL4_HUMAN	ATPase family gene 2 protein homolog 1	ATP2B	Homo sapiens (Human)	720 AA
Q9H4C8	NPL4_HUMAN	BIRC6 type 1 and C1orf100 type 2 protein homolog 1	BIRC6	Homo sapiens (Human)	510 AA
P42119	NPL4_HUMAN	ERAD-associated E3 ubiquitin-protein ligase (ERAD3)	ERAD3	Homo sapiens (Human)	1,074 AA
Q9H4C9	NPL4_HUMAN	Translocated endoplasmic reticulum chaperone 1	CHP1	Homo sapiens (Human)	604 AA
Q9H4C10	NPL4_HUMAN	Translocated endoplasmic reticulum chaperone 1	CHP1	Homo sapiens (Human)	604 AA
P44812	NPL4_HUMAN	Translocated endoplasmic reticulum chaperone 1	CHP1	Homo sapiens (Human)	604 AA

②でてきた結果のタンパク質名や
生物名に注意して、対応する
Entry列のコードをクリック

Q8TAT6-1
This isoform has been chosen as the canonical sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.

Tools Download Add Highlight **Copy sequence**

Length 608
Mass (Da) 68,120

See also sequence in UniParc or sequence clusters in UniRef

Last updated 2007-01-23 v3
MD5 Checksum¹ EC580A48328CE69180CC5F33AC78930

MAESIIRV SPQGVKRTA TKRETAATFL KKVAKKEFGQ NNGFVSYIHR NKTGEITASS NKSLLNLKIK HGDLFLFLPS SLAIGPSSHE TSVPVPGKVF GAPRVVEDEI DQVLSKQDGK IYRSRDPQCC RHGFLGKCVH

CVLPFFDED YLNLHLEPVK HMFHAYIRK LTGGADKGF VALENISCKI KSGCGEHLFW PWIGICTKCP SATLNLQKRY RHVQNZIFEN HTVADRFLDF WRKTGNQHFQ YLYGRYTEHK DIPLGIRAEV AAIYEPPQZG

TQNSLELLED PKAEVDEIA AKLGLRKGW IFTDLVSEDIT RKGTVRYSRN KDTYFLSSEE CTAGDFQNK HPBPCRLSPD GHFGSKFVTA VATGGPQKQV HFEGYQVSMQ CHALVRDECL LPCKDAPELG YAKESSESEQY

VPDVFYKQVD KFGNEITQLA RPLPVEYLII DIITTFPKDP VYTFISIQFP FPIENRDLVG ETQDFHSLAT YLSQNTSVVF LOTISDFHL LFLVTNEVMP LQDSISLILLE AVRTNRNELA QTKWRSEQMA TIEQLCSTVG

GLPLGHELVG AVGGSTHTAT AAMACQKCT FHRQPGTGK ENCLSPRT

④Copy sequenceをクリックすると
配列がコピーできるので、これを
AlphaFoldに入力する

参考（NPL4とUFD1の配列）

自分の興味のあるタンパク質で予測を試みるのが良いですが、
とりあえず練習したい人は下の配列をコピーして予測してみてください。

NPL4 MAESIIIRVQSPDGVKRITATKRETAATFLKKVAKEFGFQNNGFSVYINRNKTGEITASSNK
SLNLLKIKHGDLLFLFPSSLAGPSSEMETSVPFGFKVFGAPNVVEDEIDQYLSKQDGKIYR
SRDPQLCRHGPLGKCVHCVPLEPDEDYLNHLEPPVKHMSFHAYIRKLTGGADKKGKFVAL
ENISCKIKSGCEGHLWPNGICTKCQPSAITLNRQKYRHVDNIMFENHTVADRFLDFWR
KTGNQHFQYLYGRYTEHKDIPLGIRAEVAAIYEPPIGTQNSLELLEDPKAEVVDEIAAKL
GLRKVGWIFTDLVSEDTRKGTVRYSRNKDTYFLSSEECITAGDFQNKHPNMCRLSPDGH
FGSKFVTAVATGGPDNQVHFEGYQVSNQCMALVRDECLLPCKDAPELGYAKESSESEQYV
PDVFYKDVDKFGNEITQLARPLPVEYLIIDITTTFPKDPVYTFSSISQNPFPPIENRDVLGETQ
DFHSLATYLSQNTSSVFLDTISDFHLLLFLVTNEVMPLQDSISLLEAVRTRNEELAQTWK
RSEQWATIEQLCSTVGGQLPGLHEYGAVGGSTHTATAAMWACQHCTFMNQPGTGHCE
MCSLPRT

UFD1 MFSFNMFDHPIPRVFQNRSTQYRCFSVSMLAGPNDRSDVEKGGKIIMPPSALDQLSRLN
ITYPMLFKLTNKNSDRMTHCGVLEFVADEGICYLPHWMMQNLLLEEGGLVQVESVNLQV
ATYSKFQPPQSPDFLDITNPKAVLENALRNFACLTTGDVIAINYNEKIYELRVMETKPKAV
SIIECDMNVDFDAPLGYKEPERQVQHEESTEGEADHSGYAGELGFRAFSGSGNRLDGKK
KGVEPSPSPIKPGDIKRGIPNYEFKLGKITFIRNSRPLVKKVEEDEAGGRFVAFSGEGQSLR
KKGRKP

実際はこんな感じに入力する（例：NPL4-UFD1）

NPL4とUFD1は1:1結合なので、どちらも1と入力しておく

Entity type Protein Copies 1

NPL4

MAESIIIRVQ	SPDGVKRITA	TKRETAATFL	KKVAKEFGFQ	NNGFSVYINR	NKTGEITAS
NKSLNLLKIK	HGDLLFLFPS	SLAGPSSEME	TSVPPGFKVF	GAPNVVEDEI	DQYLSKQDGK
IYRSRDPQLC	RHGPLGKCVH	CVPLEPFDED	YLNHLEPPVK	HMSFHAYIRK	LTGGADKGF
VALENISCKI	KSGCEGHLPW	PNGICTKCQP	SAITLNRQKY	RHVDNIMFEN	HTVADRFLDF
WRKTGNQHFG	YLYGRYTEHK	DIPLGIRAEV	AAIYEPPQIG	TQNSLELLED	PKAEVVDIEA
AKLGLRKVGW	IFTDLVSEDT	RKGTVRYSRN	KDTYFLSSEE	CITAGDFQNK	HPNMCRLSPD
GHFGSKFVTA	VATGGPDNQV	HFEGYQVSNQ	CMALVRDECL	LPCKDAPELG	YAKESSESQY
VPDVFYKDVD	KFGNEITQLA	RPLPVEYLII	DITTTFPKDP	VYTFSSISQNP	FPIENRDVLG
ETQDFHSLAT	YLSQNTSSVF	LDTISDFHLL	LFLVTNEVMP	LQDSISLLE	AVRTRNEELA
QTWKRSEQWA	TIEQLCSTVG	GQLPGLHEYG	AVGGSTHTAT	AAMWACQHCT	FMNQPGTGHC
EMCSLPRT					

Entity type Protein Copies 1

UFD1

MFSFNMFDP	IPRVFQNRFS	TQYRCFSVSM	LAGPNDRSDV	EKGGKIIMPP	SALDQLSRLN
ITYPMLFKLT	NKNSDRMTHC	GVLEFVADEG	ICYLPHWMMQ	NLLLEEGLLV	QVESVNLQVA
TYSKFQPGSP	DFLDITNPKA	VLENALRNFA	CLTTGDVIAI	NYNEKIYELR	VMETKPDKAV
SIIECDMNVD	FDAPLGYKEP	ERQVQHEEST	EGEADHSGYA	GELGF	
GVEPSPSPIK	PGDIKRGIPN	YEFKLGKITF	IRNSRPLVKK	VEEDE	
RKKGRKP					

Uniprotから取ってきた配列はここに入力

+ Add entity

Continue and preview job

Save job

入力できたらここをクリック

今回は、NPL4とUFD1というタンパク質の相互作用を調べてみます²⁶

Continue preview jobを押すとこの画面が表示される

The screenshot shows a web interface for managing a job. At the top, there is a 'Job name*' field containing '2025-09-09_10:48'. Below it, a message says 'We've generated a job name for you, please edit'. There is a 'Seed: Auto' button and a 'Seed' toggle switch. A table lists job details with columns 'Type', 'Copies', and 'Sequence'. The table has two rows, both for 'Protein' with 'Copies' of 1. The first row's sequence is 'MAESIIIRVQS (length 608)' and the second is 'MFSFNMFDHPIPRVFQNRFTSTQYRCFSVSM ... (length 307)'. Below the table, it says 'Remaining jobs: 30'. At the bottom, there are two buttons: 'Go back and edit this job' and 'Confirm and submit job'. Orange arrows point from text boxes to the job name, the seed toggle, and the submit button.

Type	Copies	Sequence
Protein	1	MAESIIIRVQS (length 608)
Protein	1	MFSFNMFDHPIPRVFQNRFTSTQYRCFSVSM ... (length 307)

予測ファイルの名前。
日時は自動入力される。
タンパク質名を書いておくと便利。

初期乱数を自分で決められる
基本的にOFFで良い

ここをクリックすれば予測開始
このくらいなら10分程度で
予測完了する

配列入力画面の少し下に予測結果リストが表示される

✓ Completed ✓ Saved draft ✓ In progress ✓ Examples ✓ Failed				
<input type="checkbox"/>	Name	Modified		
<input type="checkbox"/>	2025-09-09_NPL4-UFD1	2025-09-09 10:54	⋮	
<input type="checkbox"/>	2025-05-10_KPNB1-	2025-05-10 17:25	⋮	
<input type="checkbox"/>	2025-05-09_UBE2Q2	2025-05-09 13:56	⋮	
<input type="checkbox"/>	2025-05-09_UBE2R2	2025-05-09 13:58	⋮	
<input type="checkbox"/>	2025-05-09_UBE2M		⋮	
<input type="checkbox"/>	2025-05-09_UBE2I-		⋮	
<input type="checkbox"/>	2025-05-09_UBE2F-		⋮	
<input type="checkbox"/>	2025-05-09_UBE2D3-		⋮	
<input type="checkbox"/>	2025-05-09_UBE2D2-	2025-05-09 13:56	⋮	
<input type="checkbox"/>	2025-05-08_UBE2E2	2025-05-08 19:14	⋮	

予測ファイルの名前。

予測中はくるくる回る

予測が終わると✓マークになる

ここをクリックすると
予測ファイル名の変更や
予測結果の確認、ダウンロード、
削除、など様々な事が可能

予測が終わったもののファイル名をクリックすれば
予測結果のページに飛べる

参考 (ついでに知っておくと良い情報)

AlphaFold Server BETA **Server** About FAQ & Guides ▾

残り回数。1日30回。
毎朝9時に上限まで回復。

Remaining jobs: 8

別のページに
飛んでしまったら
このServerをクリック。
このページに戻って
これです。

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

[Learn more](#) ^

- 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

Ok, got it

ここをクリックすると
exampleが自動入力
されます。入力方法の
勉強にどうぞ

入力情報の全削除

Clear

各配列情報の削除、
翻訳後修飾の入力、

Entity type
Protein

Copies
1

>Paste sequence or fasta
Input

This field is required

Entity type
Protein

Copies
1

>Paste sequence or fasta
Input

This field is required

+ Add entity

Save job

Continue and preview job

予測結果のくるくる回るマークが✓マークになると、予測終了
予測は5回されるが、そのうち最も良好なデータが表示される

2025-09-09_NPL4-UFD1

← Back ⬇ Download 📄 Clone and reuse 🗉 Feedback on structure

5回のデータを全部みたい場合は
ここをクリックしてダウンロード

Very high (pLDDT > 90)

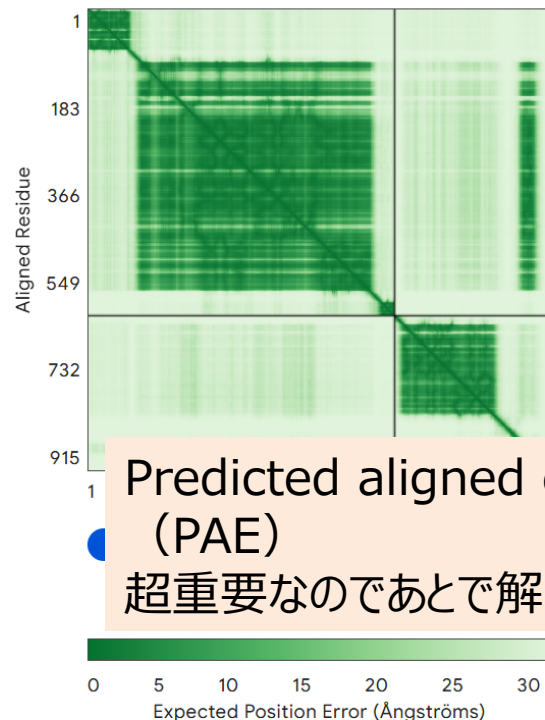
High (70 > pLDDT > 70)

50)

pLDDTの凡例

ipTM = 0.66 pTM = 0.57 [learn more](#)

ipTMは複合体の
予測信頼度が
高いかどうか、
pTMは単体の
予測信頼度が
高いかどうかの尺度。
0.8を越すと信頼性が
高いが、それ以下でも
正解のことは多く、
あまり気にしなくて良い。



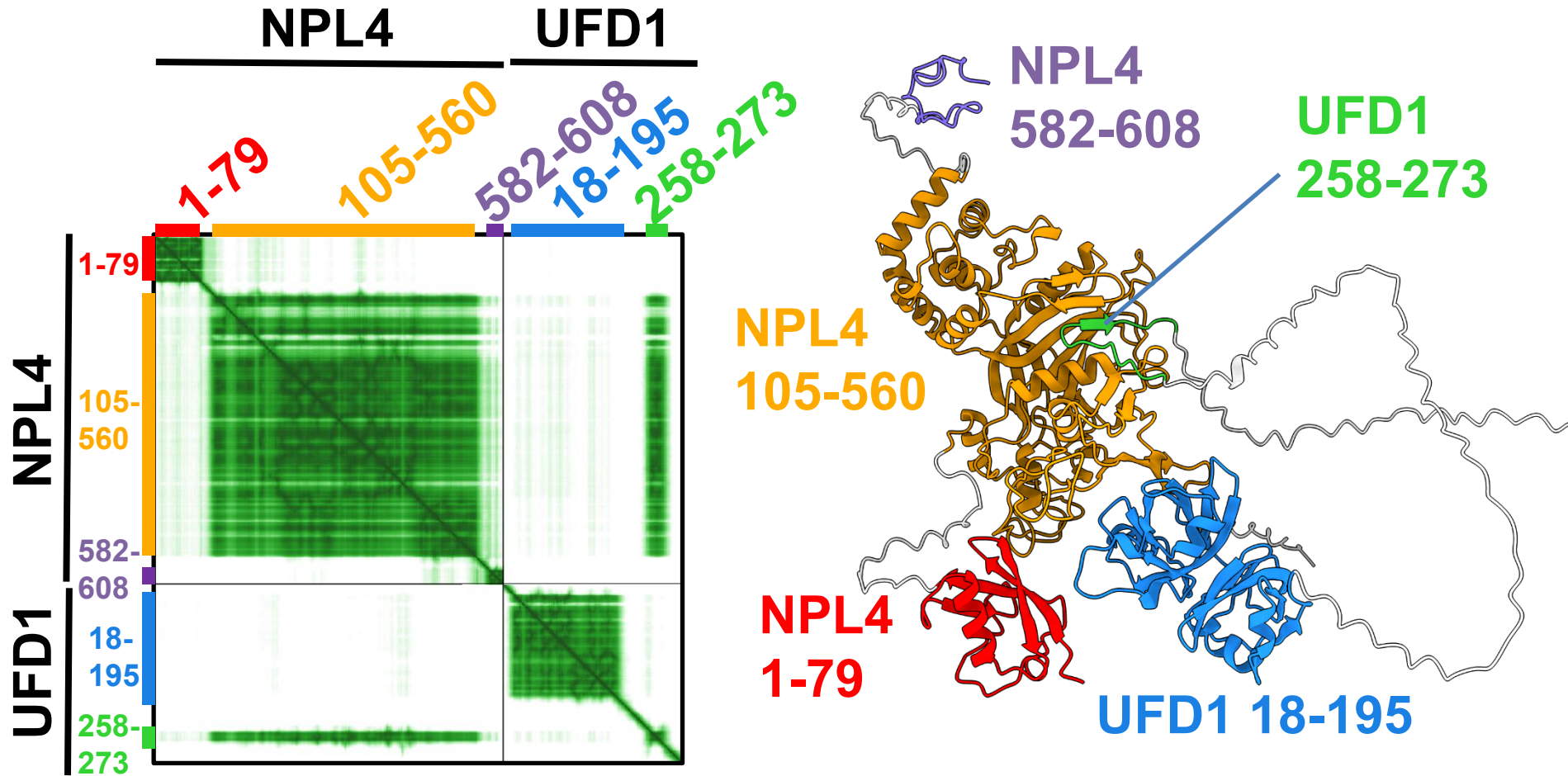
スパナマークをクリックすると、見た目を変えることができる

できることが多すぎて
かえって混乱してしまうが、
とりあえず、左図の
①Polymerの右端の…
②Set Coloring
③Chain Property
④Chain Id
の順番でクリックすると、
タンパク質毎に色分けされる

予測結果は
左クリックで回転させたり
右クリックで移動させたり
ホイールで拡大したり
できる

⑤の撮影マークで
画像を保存することも
できる

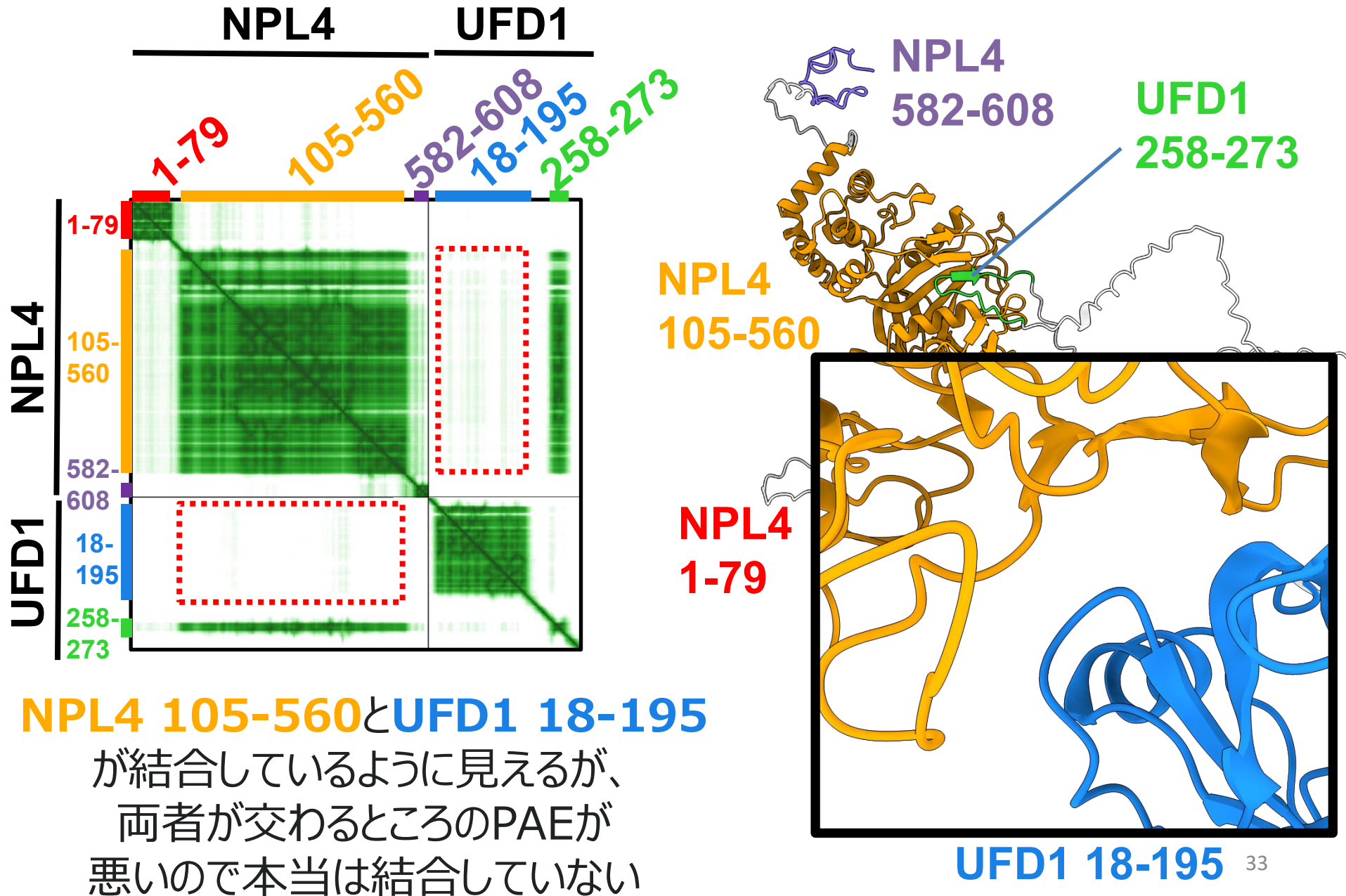
配列を2つ以上入力した場合は、PAEで相互作用を確認



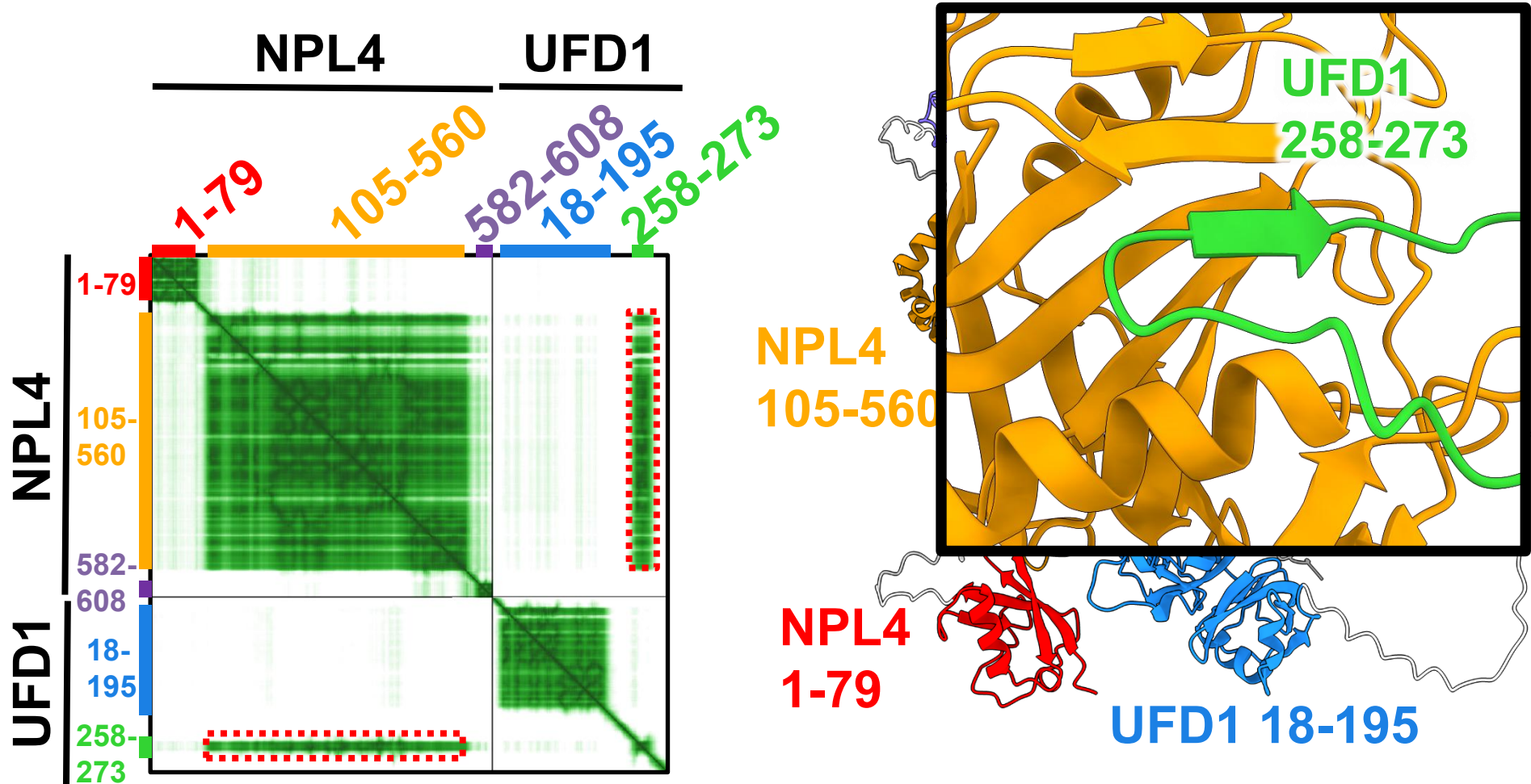
複合体を形成したのかどうかは、構造よりもまずPAEを重視すること。
複合体の時は、入力したタンパク質がすべて連結したPAEが表示される。

**結合した2つの領域は、お互いの位置関係がきっちりするので、
PAEが良くなる、というのが最も大事な点**

配列を2つ以上入力した場合は、PAEで相互作用を確認



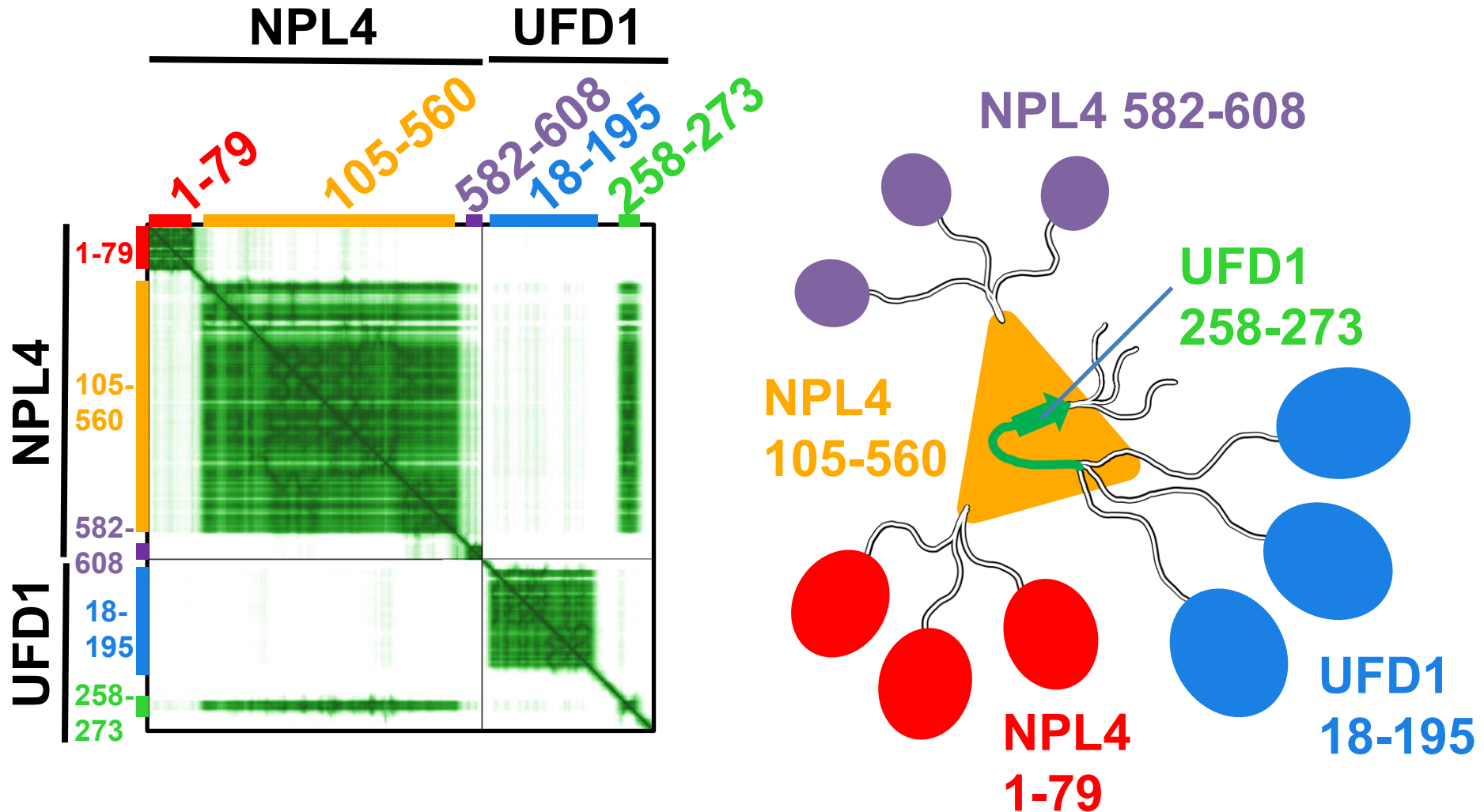
配列を2つ以上入力した場合は、PAEで相互作用を確認



NPL4 105-560と**UFD1 258-273**

が結合しているように見えるし、
両者が交わるところのPAEが
良いので**本当に結合する可能性が高い**

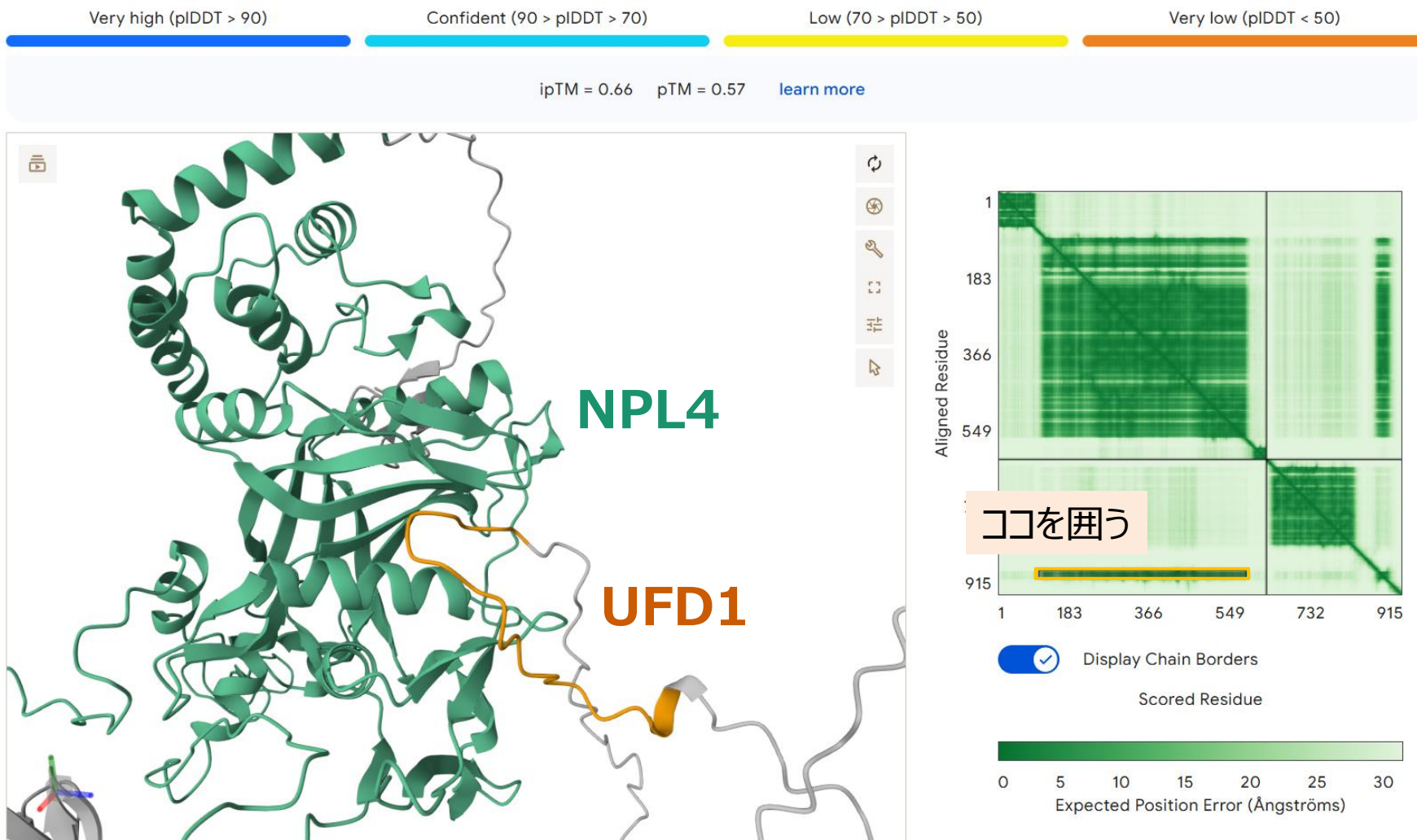
配列を2つ以上入力した場合は、PAEで相互作用を確認



なお、いつも模式図で書くとこんな感じ。

PAEを考慮すると、**NPL4 105-560**と**UFD1 258-273**が相互作用するほかは、フラフラした領域を持つ複合体であることがわかる

なお、AlphaFold ServerのPAEの気になる部分を
左クリックで囲うと、その位置に対応する部分だけ色分け
されるので、相互作用を確認しやすい



もっと詳しく見たい場合、ダウンロードすると以下のファイルが得られる。
.cifファイルから構造が見られるが、そのためにはChimeraX(無料)などのソフトが必要になるので、
今回は省略する

名前	更新日時	種類	サイズ
fold_2025_02_26_nucleosome_1aoi_full_data_0.json	2025/02/26 7:35	JSON ファイル	13,662 KB
fold_2025_02_26_nucleosome_1aoi_full_data_1.json			13,660 KB
fold_2025_02_26_nucleosome_1aoi_full_data_2.json			13,664 KB
fold_2025_02_26_nucleosome_1aoi_full_data_3.json			13,648 KB
fold_2025_02_26_nucleosome_1aoi_full_data_4.json	2025/02/26 7:35	JSON ファイル	13,662 KB
fold_2025_02_26_nucleosome_1aoi_job_request.json	2025/02/26 7:35	JSON ファイル	13,662 KB
fold_2025_02_26_nucleosome_1aoi_model_0.cif	2025/02/26 7:35	CIF ファイル	1,129 KB
fold_2025_02_26_nucleosome_1aoi_model_1.cif			1,129 KB
fold_2025_02_26_nucleosome_1aoi_model_2.cif			1,143 KB
fold_2025_02_26_nucleosome_1aoi_model_3.cif			1,129 KB
fold_2025_02_26_nucleosome_1aoi_model_4.cif	2025/02/26 7:35	CIF ファイル	1,143 KB
fold_2025_02_26_nucleosome_1aoi_summary_confidences_0.json	2025/02/26 7:35	JSON ファイル	3 KB
fold_2025_02_26_nucleosome_1aoi_summary_confidences_1.json			
fold_2025_02_26_nucleosome_1aoi_summary_confidences_2.json			
fold_2025_02_26_nucleosome_1aoi_summary_confidences_3.json			
fold_2025_02_26_nucleosome_1aoi_summary_confidences_4.json	2025/02/26 7:35	JSON ファイル	3 KB
terms_of_use.md			

ファイルサイズの大きな .jsonファイルはChimeraXでPAEを読み込める

メモ帳で開くと入力情報が確認できる

.cifファイルからChimeraX、PyMolで構造を読み込める

ファイルサイズの小さな .jsonファイルからはipTMなど確認できる

メモ帳で開くと規約が確認できる

AlphaFoldの使い方まとめ

1. AlphaFold Protein Structure Database (AlphaFold PDB)に掲載されたデータを見る

<https://alphafold.ebi.ac.uk/>

2. 自分でAlphaFold3を走らせる

データベースにないタンパク質の場合は
自分でプログラムを走らせる

<https://alphafoldserver.com>

タンパク質のアミノ酸配列は下記の
NCBIデータベースから持ってくる

<https://www.uniprot.org/>

おわりに

今回は、AlphaFoldを利用した構造予測について解説しました。

ここまで簡単に構造予測ができるようになると、今後構造研究は必要なのか？と考える人もいると思います。しかし、実は予測結果はタンパク質が起こす化学反応について詳細な議論をするためには正確性が足りません。特に、タンパク質と化合物との予測精度は、まだそれほど高くないようです。

また、タンパク質は仕事をする際に大きく動くことも多いのですが、AlphaFoldではまだ、タンパク質の動きに対応することはできていません。

したがって、今後の構造研究では、化合物の認識や、化学反応、タンパク質の動きを中心に行っていく必要があります。みなさんも興味があれば、大学で構造生物学を学んでみてください。