

構造バイオインフォマティクス基礎 第4回 2016.05.17

タンパク質の構造予測 (補足資料)

清水謙多郎
shimizu@bi.a.u-tokyo.ac.jp

構造モデリングのためのアラインメントの作成(1)

ターゲットの配列をクエリとしてBLAST検索した結果

「Multiple alignment」を選択

アラインメントの作成

構造モデリングのためのアラインメントの作成(2)

いったんチェックを外す

アラインメントの作成

構造モデリングのためのアラインメントの作成(3)

最初の行のクエリと
2HT5の2つをチェック

アラインメントの作成

構造モデリングのためのアラインメントの作成(4)

再度アラインメントをとる

Accession	Description	Links
gi Query_46169	p1	
3C7E.A	Chain A, Crystal Structure Of The Native 1918 H1N1 Neuraminidase From A Crystal With L...	
3B5Q.A	Chain A, Neuraminidase Of A6RE1VC M85SC11198 H1N1 ST1049	
3H5D.A	Chain A, The Crystal Structure Of Neuraminidase From American Green-winged Tealwate...	
2HT5.A	Chain A, N1 Neuraminidase	
3C32.A	Chain A, N1 Neuraminidase H294s + Clofazim	
3C5Z.A	Chain A, N1 Neuraminidase H274y + Zanamiv	
4B7Q.A	Chain A, H1N1 2009 Pandemic Influenza Virus: Resistance Of The Q229 Neuraminidase Mutat...	
4B7J.A	Chain A, H1N1 2009 Pandemic Influenza Virus: Resistance Of The Q229 Neuraminidase Mutat...	
4B7B.A	Chain A, H1N1 2009 Pandemic Influenza Virus: Resistance Of The Q229 Neuraminidase Mutat...	
3H5S.A	Chain A, The 2009 Pandemic H1N1 Neuraminidase H1 Lacks The 150-cavity In Its Active Site...	PubChem BioAssay info linked to 3H5S.A
4QHP.A	Chain A, Crystal Structure Of The 2009 Pandemic H1N1 Influenza Virus Neuraminidase With A...	
2HT7.A	Chain A, N1 Neuraminidase	
4B3E.A	Chain A, Influenza Neuraminidase In Complex With A Stereomulated Analogue Of Clofazim	
4E51.A	Chain A, Influenza Neuraminidase In Complex With Arterial Compound (3s,4r,5r)-4-(acetamyl...	
4M31.A	Chain A, Influenza Neuraminidase In Complex With A Novel Arterial Compound	
4M33.A	Chain A, The Crystal Structure Of Neuraminidase From A H3N1 Influenza Virus Isolated From...	
4M47.A	Chain A, Influenza Neuraminidase In Complex With A Novel Arterial Compound	
3C9H.A	Chain A, Influenza Na In Complex With Compound 5	
4G81.A	Chain A, Synthesis And Evaluation Of Novel 3-C-Alkylated-Neuracem Derivatives As Probes...	
5H8N.A	Chain A, The Crystal Structure Of Neuraminidase From Agyptal on Washington 11988.6/2014	
2HT5.A	Chain A, N1 Neuraminidase	
33AL.A	Chain A, Crystal Structure Of Influenza A Virus Neuraminidase N5	PubChem BioAssay info linked to 33AL.A

アラインメントの作成

5

構造モデリングのためのアラインメントの作成(5)

アラインメントの作成

アラインメントの作成

6

構造モデリングのためのアラインメントの作成(6)

FASTA plus gaps

アラインメントの作成

7

構造モデリングのためのアラインメントの作成(7)

```
>Query_46169 p1
VILTGNSSLCPISGNWAIYSKDNIGIRIGSKGDVVFVIREPFFISCSHLECRFTFFLTQGALLND
KHSNGTVKDRSPYRTLMSCPVGEAPSPYNSRFSVAVWSASACHDGMGLWTIGISGPDNGA
VAVLKYNGIITDTIKSWRNRIILRTQSEACVNGSCFTIMTDGSPNGQASYKILKIEKKG
VTKSIELNAPNYHYEECSYCPDTGKVMCVCRDNWHGNSNRPWVSPDQNLDYQIYICSGVF
GDNRPNDG--TGSCGPVSSNGANGIKGFSFRYDNGVWIGRTKSTSSRSGFEMIDWPNGW
TETDSSFVSVRQDVAITDWSGYSGSFVQHPELTGLDCMRPCFWVELIRGQPKENTIWTSG
SSISFGVNSDVTGWVSWPDGAEIPFSI---
>gi|114794506|pdb|2HT5|A Chain A, N8 Neuraminidase
TYMNNTEAICDAKGFAPFSKDNIGIRIGSRGHIFVIREPFFVSCSPIECRTFFLTQGSLLND
KHSNGTVKDRSPYRTLMSVEVQSPNVYQARFEAVWASATACHDGKKNMTVGVTPGDSKA
VAVIHYGGVPTDVVNSWAGDILRTOESSCTCIQDCYVWMTDGPANRQAQYRIYKANQGR
IIGQTDISFNGGHIIECSYCPNDGKVEVCVCRDGTGTRNRPVLVISPDLRYRVLGACGIP
SDTPRGEDTQFTGSCSPMGNGQYGVKGFGRQGTDVWVGRTISRTSRSGFELIRIKNGW
TQTSKEQIRKQVVDNLNWSGYSGLPPLVVELSGKDCLVPCFWVEMIRGKPEKTIWTSS
SSIVMCGVDVEVADNWSHWDGAILPFDIDKM
```

- FASTA形式のファイルが得られる
- Chimeraに入力するには、「Aligned FASTA」形式を選択。そのため、ファイル修飾子を「.fasta」などとする
- Chimeraは、指定されたテンプレート構造をダウンロードするので、アラインメントファイルの「2HT5」を「2HTQ」に変えておく
- ChimeraでModellerを起動したとき、「Include non-water HETATM ...」にチェック

アラインメントの作成

8

構造モデリングのためのアラインメントの作成(8)

なぜ2つの配列でアラインメントを取り直すのか？

- モデリングで利用するテンプレートとターゲットのアラインメントファイルが直接的に得られるから。実は、スライド(1)で、「Download」を実行し、全体のマルチプルアラインメントのファイルを得た上で、2つの配列を取り出してもよい。その方が、類縁タンパク質の性質を加味したアラインメントが行える可能性がある。しかし、全体の配列のアラインメントなので、余計なギャップが入ることがある。ベストなのは、適当な配列を選択してマルチプルアラインメントをとり、そこから2つの配列のアラインメントを取り出す方法だろう。

BLASTによって得られたアラインメントを使っていないのでは？

- ここで得られるアラインメントは、NCBIのマルチプルアラインメント手法COBALTによるものである。BLASTで得られる、2つの配列の間のアラインメントは、高スコアで一致した部分を、動的プログラミングでつなぐアラインメントである。COBALTは、保存されたドメインと局所的な配列類似性をもとにアラインメントをとる手法で、両者は異なる。COBALTの手法の方が、テンプレートの性質を生かした効果的なアラインメントが得られる可能性があると考えられる。

そのほか、モデリングの精度向上のためのさまざまな手法については、講義資料を参考にしてください。