

# آموزش مقایسه همردیفه تواله‌هاک پروتئین در قالب فرمت PDB

## تهیه و تدوین:

حسن رسوله (مرکز تحقیقات بیولوژیک پزشکی دانشگاه علوم پزشکی کرمانشاه)

مقایسه همردیفه یا alignment روشی برای مقایسه دو توالی با یکدیگر است که از ابزاری به نام BLAST برای انجام آن استفاده می‌شود. دقت داشته باشید که blast ابزار انجام مقایسه همردیفه است و بین این دو مقوله یعنی alignment و blast تفاوت‌های اساسی وجود دارد. با توجه به اینکه این مبحث برای علاقه‌مندی در تدوین شده است که تازه شروع به یادگیری بیوانفورماتیک کرده‌اند لذا سعی شده است مطالب مذکور با بیانی ساده و روان برای این عزیزان تشریح گردد تا دوستان بتوانند به راحتی اصول پایه و اولیه بیوانفورماتیک را یاد گیرند و مجبور نباشند برای شرکت در ورکشاپ‌هایی که با قیمت‌های فضایی در حال حاضر در کشور برگزار می‌شود شرکت نمایند. همانطوری که گفته شد blast ابزار انجام مقایسه همردیفه است. به طور کلی چند نوع مقایسه همردیفه یا همترازی وجود دارد. نوع اول مقایسه همردیفه مقایسه یک توالی با توالی دیگر است که به آن مقایسه همردیفه دوگانه یا pairwise alignment می‌گویند. نوع دوم مقایسه همردیفه مقایسه یک توالی با چندین توالی مختلف (گاهی حتی با میلیون‌ها توالی که در یک پایگاه داده وجود دارد) صورت می‌گیرد. این نوع از مقایسه همردیفه معمولاً برای شناسایی توالی‌هایی که دارای حداقل یا حداکثر تشابه به توالی query (یعنی توالی که شما آن را با توالی‌های دیگر مقایسه می‌کنید) می‌باشد. در پایگاه NCBI ابزارهای مختلفی از BLAST وجود دارد. در حال حاضر ابزار دیگری که در پایگاه PDB در آدرس <http://www.rcsb.org/pdb> موجود است معرفی می‌شود. این ابزار به شما کمک می‌کند که CHAIN‌های مختلف یک پروتئین را با یکدیگر مقایسه کنید و نقاط تشابه موجود در این فایل‌ها را شناسایی کنید. برای استفاده از این ابزار مراحل زیر را دنبال کنید.

۱- ورود به سایت PDB در آدرس:

<http://www.rcsb.org/pdb/home/home.do#Category-analyze>

۲- پس از ورود به سایت PDB بایستی از منوی سمت چپ گزینه ANALYZE را انتخاب کنید. پس از انتخاب این گزینه صفحه جدیدی مطابق با شکل زیر برای شما به نمایش در خواهد آمد:

The screenshot displays the RCSB PDB website's 'Sequence & Structure Alignment' tool. The interface includes a sidebar with navigation options and a main content area with a descriptive text, a 3D protein structure image, and input fields for PDB IDs (1CIV and 1SMK), chain IDs (A), and alignment methods (blast2seq). The 'Align' button is highlighted in blue.

۳- اکنون در بخش **Sequence and structure alignment** شماره دسترسی توالی‌های پروتئین خود را وارد نمایید. در اینجا ما از دو شماره دسترسی **1CIV** و **1SMK** استفاده کرده‌ایم. شما می‌توانید شماره دسترسی توالی‌های دلخواه‌تون را در این قسمت وارد کنید. permutations.

This close-up view shows the input fields for the alignment tool. The PDB IDs '1CIV' and '1SMK' are circled in red. Below them are dropdown menus for 'Select Associated Chain ID' (both set to 'A') and a dropdown for 'blast2seq'. The 'Align' button is also visible.

۴- بعد از درج کد دسترسی توالی‌ها اکنون در قسمت **select associated chain ID** شما بایستی **chain**‌های دلخواه‌تون را با هم مقایسه کنید. در اینجا ما دو **chain A** پروتئین‌های مذکور را برای مقایسه با یکدیگر انتخاب کرده‌ایم:

both rigid-body and flexible alignments and detection of circular permutations.



1CIV ↔ 1SMK

Select Associated Chain ID

A (Seq: SVTSSDQIQAPLPAKQKPECF) ↔ A (Seq: RAKGGAPGFKVAILGAAGGIC)

A (Seq: SVTSSDQIQAPLPAKQKPECFGVFCLTYDL...) ↔

blast2seq ↔ Align More Options

۵- بعد از این مرحله بایستی در گزینه **select comparison method** نوع روش مقایسه خود را انتخاب نمایید. در اینجا ما از گزینه **blast2seq** (یعنی مقایسه دو توالی با یکدیگر) را انتخاب کرده‌ایم. گزینه‌های دیگری نیز در این صفحه وجود دارند که هر کدام از آنها مربوط به نوع خاصی از مقایسه هم‌ردیفی هستند و با توجه به نیاز کاربر انتخاب می‌شوند:

Needleman-Wunsch, and Smith-Waterman) and structure alignmer (FATCAT, CE, Mammoth, TM-Align, TopMatch).

to

- Select Comparison Method -

**Sequence Alignment**

- blast2seq
- Needleman-Wunsch
- Smith-Waterman

**Structure Alignment**

- jFATCAT - rigid
- jFATCAT - flexible
- jCE algorithm
- jCE Circular Permutation
- External: FatCat
- External: Mammoth
- External: TM-Align
- External: TopMatch

- Select Comparison Method -

the PDB archive and  
ial features include st  
etection of circular

1SMI

Select



A (S

Align M

۶- پس از انتخاب روش مقایسه هم‌ردیفی اکنون بایستی شما گزینه **Align** را انتخاب نمایید. این گزینه به شما کمک خواهد کرد تا دو فایل **pdb** مذکور را با همدیگر مقایسه کنید. پس انتخاب این گزینه صفحه خروجی شما به صورت زیر می‌باشد:

Calculate pairwise sequence or structure alignments.

blast2seq Sequence Alignment Results

<b>Alignment Details:</b> <b>Query:</b> <i>NADP-MALATE DEHYDROGENASE</i>  <b>Query:</b> PDB ID: 1CIV Chain ID: A EC number: 1.1.1.82	<b>Subject:</b> <i>Malate dehydrogenase, glyoxysomal</i>  <b>Subject:</b> PDB ID: 1SMK Chain ID: A EC number: 1.1.1.37
---	--

Comparison Method

Select these two chains for other comparison: --- Select Comparison Method ---  
 Click [here](#) to align other protein chains. Back to the all vs. all search results for 1CIV.A or 1SMK.A

Text Representation of the blast2seq Sequence Alignment

Query : 1CIV.A  
 Subject: 1SMK.A  
 Query=  
 (385 letters)  
 >  
 Length = 326

Acknowledgements

Click [Here](#) to show the acknowledgements related to our structure comparison feature.

Comparison Tool

Calculate pairwise sequence or structure alignments.



```
>
Length = 326

Score = 34.3 bits (77), Expect = 5e-06
Identities = 65/253 (25%), Positives = 93/253 (36%), Gaps = 28/253 (11%)

Query: 45 VAVSGAAGMISNHLFLKASGEVFGPDQPISLKLGSERSFAALEGVAMELEDSLYPLLR 104
      VA+ GAAG I L + + +S+ L          GV ++ +
Sbjct: 11 VAILGAAGGIGQPLAMLKMNPL-----VSVLHL---YDVVNAPGVTADISHMDTGAVV 61

Query: 105 QVSIQIDPYEIFQDAEWALLIGA-KPRGPGMERADLLDINGQIFAEQKALNAVASPNVK 163
      + +G E      +++ A PR PGM R DL IN I  + + A P
Sbjct: 62 RGLFGQQQLEAALTGMDLIIVPAGVPRKPGMTRDDLDFKINAGIVKTLCEGI-AKCCPRAI 120

Query: 164 VMVVGNPCNTNALIC---LKNAPNIPPKNFHALTRLDENRAKQQLALKAGVFYDKVSNVT 220
      V ++ NP N+ I   K A   PK   +T LD RA  +A G+ V
Sbjct: 121 VNLISNPNVNSTVPIAAEVFKAGTYDPKRLLGVTMLDVVRANTFVAEVLGLDPRDVPV 180

Query: 221 IWGNHSTIQVPDFLNAKIHGIPVTEVIRDRKWEDEFTHMVQIRGG--VLIKKWGRSSA- 277
      + G+ T +P   K   P +   + +L D N   GG V+ K G SA
Sbjct: 181 VGGHAGVTILPLLSQVK---PPSSFTQEISYLTDRIQN-----GGTEVVEAKAGAGSAT 232

Query: 278 ---ASTAVSIVDA 287
      A AV DA
Sbjct: 233 LSMAYAAVKFADA 245

Score = 15.8 bits (29), Expect = 1.9
Identities = 8/28 (28%), Positives = 12/28 (42%)
```

```

Sbjct: 121 VNLIDNPFVNDIVFIAAEVFKKAGITDFKRLLGVIIMLDVVRANIFVAELVGLDFRDVDFV 180

Query: 221 IWGNHSTIQVPDFLNAKIHGIPVTEVIRDRKWEDEFNMQIRGG--VLIKKWGRSSA- 277
      +G+ T+P K P+ + +L D N GG V+ K G SA
Sbjct: 181 VGGHAGVTILPLLSQVK---PPSSFTQEEISYLIDRIQN-----GGTEVWEAKAGAGSAT 232

Query: 278 ---ASTAVSIVDA 287
      A AV DA
Sbjct: 233 LSMAYAAVKFADA 245

Score = 15.8 bits (29), Expect = 1.9
Identities = 8/28 (28%), Positives = 12/28 (42%)

Query: 346 KKIKKSEDELLAEKKCVAHLTGEGIAVC 373
      +K + D+L + EGIA C
Sbjct: 88 RKPGMTRDDLFKINAGIVKILCEGIAKC 115

Lambda K H
      0.317 0.135 0.403

Gapped
Lambda K H
      0.267 0.0410 0.140

Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 261

```

اکنون تمامی مشخصات دو **chain** شما در دسترس است و می‌توانید نقاط تشابه آنها را مشاهده نمایید. چنین مقایساتی هنگام ایجاد مدل‌های سه بعدی از پروتئین‌ها بر اساس الگوی **pdb** از اهمیت بالایی برخوردار است. شما اکنون می‌توانید آنالیزهای موردنظر را ذخیره کنید و در رابطه با آن تفسیرهای مناسبی را ارائه دهید.