

## ЛАБОРАТОРНА РОБОТА № 5-6

ТЕМА: Дослідження параметрів білків за допомогою комп'ютерних програми. Вирівнювання послідовностей і філогенетичний аналіз

**Мета:** ознайомитися з теоретичними основами парного та множинного вирівнювання. Навчитися здійснювати різні типи вирівнювань для встановлення структурних, функціональних і еволюційних відносин між послідовностями біологічних макромолекул.

### Теоретична частина

Вирівнювання послідовностей дає змогу встановити відповідність залишків у ланцюгах, які є основним засобом у біоінформатиці. Методи порівняння послідовностей нуклеотидів нуклеїнових кислот та амінокислот білків допомагають знаходити функціонально подібні або еволюційно споріднені об'єкти.

Вирівнювання послідовностей у біоінформатиці – метод порівняння нуклеотидних (ДНК, РНК) або амінокислотних (білки) послідовностей знаходженням схожих ділянок, що може бути наслідком функціональних, структурних або еволюційних зв'язків між послідовностями.

Порівняння нових нуклеотидних послідовностей ДНК і РНК, а також амінокислотних послідовностей білків відбувається на основі уже відомих, розшифрованих послідовностей та необхідне для виявлення загальних структурних і функціональних особливостей, а також для дослідження еволюційних зв'язків.

Мета вирівнювання послідовностей полягає в тому, щоб визначити ступінь подібності двох послідовностей і якщо вона достатньо висока зробити правомірний висновок про їхню гомологічність.

Залежно від кількості послідовностей, які порівнюються, розрізняють парне та множинне вирівнювання.



Рисунок 1 - Види вирівнювання послідовностей нуклеотидів та амінокислот

**Парне вирівнювання** здійснюється між двома послідовностями.

1. *Глобальне вирівнювання* – форма глобальної оптимізації, яка знаходить відповідності нуклеотидних або амінокислотних залишків по всій довжині досліджуваних послідовностей. Глобальне вирівнювання може бути застосоване, якщо дві послідовності гомологічні по всій довжині.

Результати вирівнювання послідовностей розміщуються в рядках матриці так, що елементи, які збігаються (нуклеотиди або амінокислоти), розміщені один під одним (в одній колонці).

```

Blast:

Query 15 FQQAWANPKHAWAQVNGETRLTQNLIIILERETR 47
        F W PKHA +QVNG T ++Q+ IIL R R
Sbjct 14 FHHNWTRPKHASSQVNGHTEMSQHNIIILRRVPR 46

CLUSTAL:

THE12851.1      MGKKGYKRNEYNNPFQQAWANPKHAWAQVNGETRLTQNLIIILERETRKRS- 50
WP_104057486.1 MSTK-DQLDPQSQAFHHNWTRPKHASSQVNGHTEMSQHNIIILRRVPRSGRR 50
               *..* : : .: *:: *:.***** :*****.*::*: ***.* *.

EMBOSS Needle:

THE12851.1      1 MGKKGYKRNEYNNPFQQAWANPKHAWAQVNGETRLTQNLIIILERETRKRS      50
                 |..|. :.....|...|..|||.|:||||.|.:|:|..|..|...
WP_104057486.  1 MSTKD-QLDPQSQAFHHNWTRPKHASSQVNGHTEMSQHNIIILRRVPRSGR      49

```

Рисунок 2 - Три різних приклади вирівнювання

У цих прикладах послідовності записуються одна під одною, а в рядку між ними різними символами позначені різні відносини між амінокислотами. Пробілом (відсутністю символу) позначають відсутність зв'язку між амінокислотами як по гомології, так і за функцією; символами «\*», "|" або ж літерою (BLAST) - однакові амінокислоти; ":" або "+" - близькі за властивостями; «.» - подібні за властивостями.

Ще приклад вирівнювання:

```

AAB24882      TYHMCQFHCRVYVNNHSGEKLYECNERSKAFSCPSHLQCHKRRQIGKTHEHNQCCKAFPT 60
AAB24881      -----YECNQCCKAFAQHSSLKCHYRTHIGKPYECNQCCKAFSK 40
               ****; ,***; * *:* * :****;.* *****.,

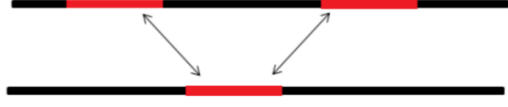
AAB24882      PSHLQYHERTHTGKPYECHQCQAFKCSLLQRHKRTHTGKPYE-CNQCCKAFAQ- 116
AAB24881      HSHLQCHKRTHTGKPYECNQCCKAFSQHGLLQRHKRTHTGKPYMNVINMVKPLHNS 98
               **** *:******;***;* : ,***** : *.: :

```

Рисунок 3 - Глобальне вирівнювання послідовностей амінокислот

У даному типі вирівнювання символом «\*» позначені відповідності (однакові); «.» – подібні за властивостями; «:» – близькі за властивостями; «прогаліни» позначають невідповідності, «-» позначає вставки (інсерції, від англ. insertion) і видалення (делеції, від англ. deletion), які необхідно зробити в обох послідовностях, щоб досягнути максимальної кількості відповідностей.

2. Локальне вирівнювання ідентифікує подібні ділянки у межах довгих послідовностей, які суттєво відрізняються на більшій частині своєї протяжності. Локальне вирівнювання – може бути застосоване для порівняння послідовностей із частковою гомологією.



Можливі також вставки та видалення всередині збігаючої частини.

3. *Вирівнювання, яке перекривається*, призначене для порівняння послідовностей, у яких збігаються тільки кінцеві ділянки



4. *Точкове вирівнювання (dot plot)* застосовується для загального дослідження послідовності з метою виявлення повторів і вибору фрагмента для множинного вирівнювання.

Точкове вирівнювання та вирівнювання, яке перекривається передбачають використання тільки парного порівняння послідовностей.

**Множинне вирівнювання** послідовностей застосовують для одночасного порівняння більш ніж двох послідовностей. Цей вид вирівнювання значно інформативніший, ніж парний.

Зазвичай, після здобуття результатів секвенування (або після завантаження даних із баз Genbank, UniProt) потрібно виконати процедуру вирівнювання нуклеотидних (або амінокислотних) послідовностей. Процедура вирівнювання забезпечує знанням про те, які ділянки ДНК або білка гомологічні, і саме на основі відмінностей у нуклеотидних або амінокислотних замінах гомологів зможемо потім реконструювати філогенетичне дерево.

Програм, які допомагають здійснити вирівнювання секвенованих послідовностей, безліч. Найпоширеніші такі (більшість із них безкоштовна) як:

- ClustalW
- MUSCLE
- T-Coffee
- BLAST
- Mafft
- KAlign
- MEGA

На сьогодні множинне вирівнювання послідовностей реалізується кількома програмами, доступними в режимі онлайн.

Одна з найуживаніших програм прогресивного множинного вирівнювання ClustalW. Ця програма дає змогу вирівнювати як нуклеотидні, так і амінокислотні послідовності. Основне призначення програми ClustalW – виконання множинного вирівнювання, обчислення еволюційних дистанцій між послідовностями, визначення характеру і типу амінокислотних замін.

Отримані вирівнювання можна відобразити в чорно-білій або кольоровій гамі, залежно від властивостей амінокислот (у разі вирівнювання амінокислот). Консервативність і напівконсервативність амінокислотних замін визначається відповідно до таблиці.

*Таблиця. Один із можливих способів забарвлення амінокислотних залишків при візуалізації множинного вирівнювання амінокислотних послідовностей*

Колір	Тип залишку	Амінокислоти
жовтий	неполярні залишки	Gly, Ala, Ser, Thr
зелений	гідрофобні	Cys, Val, He, Leu, Pro, Phe, Tyr, Net, Trp
фіолетовий	полярні	Asn, Gin, His
червоний	негативно заряджені	Asp, Glu
синій	позитивно заряджені	Lys, Arg

Велика частина програмного забезпечення працює безпосередньо з форматом FASTA. У біоінформатиці, формат FASTA – це текстовий формат файлу для збереження нуклеотидних або

амінокислотних послідовностей, у якому нуклеотиди або амінокислоти позначаються літерами. Цей формат дає змогу передавати опис цих послідовностей і короткий коментар до них. Назва формату походить від програмного пакета FASTA, але він уже став незалежним стандартом у біоінформатиці. Послідовність у цьому форматі починається з назви, перед якою ставлять символ «>». Перше слово після цього символу зазвичай – ідентифікатор послідовності – такий як номер послідовності в базі даних UniProt або GenBank.

```
>sp|Q3TDX8|NB5R4_MOUSE Cytochrome b5 reductase 4 OS=Mus
musculus GN=Cyb5r4 PE=2 SV=3
MLNVPSQAFFPAGSQQRVSSQGRSKVPLKQGRSLMDWIRLTKSGKDLTGLKGGLIEVTEE
ELKKNKEDCWICIRGFVYNVSPYMEYHPGGEDELMRAAGADGTDLFNEVHRWVNYESM
LKECLVGRMAVKPAVPKDCHEGKRVNLGMLPKSQMSDTLPRDVTDTLPREDLSSPSYDWF
QTESSVTIVVYTKQKNISLDSVIVDLQDDSLRAEAVIKDHSYLVHVGLSHEVQENFSVRV
IENVGKIEIVLQKKEVSWQCLGDHLEKHDSFIPKKTGLYYRRCQLISKEDVTHDTRLF
CMLPSTHQLQVPVQGHVYLKLSVTGAEIVKPYTPVSDSLLSDFKEPVLSPNKYICFLIK
IYPAGLFTPELDRQLIGDFISVSGPEGDFKVKLQEVEDLFLAAGTGFTPMVTVLNYAL
SHMSSLRKVKLMFFNKTEDDIWRCQLEKLALREKRFDFVFLSAPSPWNGKQGHISRA
LLSEFLQRSSENSRAFLCICGPTPFTDEGIRLLHDLNFSDDDEIHGFTA
```

Рисунок 3 - Приклад формату FASTA амінокислотної послідовності

У FASTA-форматі використовують однобуквені коди для нуклеотидів і амінокислот, задані Міжнародним об'єднанням біохімії і Міжнародним об'єднанням чистої і прикладної хімії.

Філогенетичний аналіз у систематиці описує взаємозв'язки між таксонами і допомагає зрозуміти історію еволюційних відносин між живими організмами. Еволюційну історію, відновлену в результаті філогенетичного аналізу, зображають у вигляді розгалужених, деревоподібних діаграм, які подають передбачуваний родовід спадкових взаємозалежностей між молекулами, організмами або і тим, і тим.

Філогенетичне дерево (еволюційне дерево, дерево життя) – умовна схема, що відображає еволюційні взаємозв'язки між різними видами, іншими таксонами, генами або іншими об'єктами, що мають загального предка. Вершини філогенетичного дерева діляться на три класи: листя, вузли і (максимум один) корінь. Листя – це кінцеві вершини, тобто ті, в які входять рівно по одному ребру; кожен лист відображає деякий вид живих організмів (або інший об'єкт, схильний до еволюції, наприклад білковий домен). Кожен вузол представляє еволюційну подію: розділення предкового виду на два або більше, які надалі еволюціонували незалежно. Корінь представляє загального предка всіх даних об'єктів. Ребра дерева філогенезу прийнято називати «гілками».

## Практична частина

### Завдання 1

Виконайте вирівнювання амінокислотної послідовності супероксиддисмутази людини проти амінокислотної послідовності відповідного білка щура, дрізофіли, кишкової палички. Побудуйте філогенетичне дерево. Вкажіть найспорідненіші види.

#### Хід дослідження:

1. Увійти на веб-ресурс UniProt (<http://www.uniprot.org>). У полі пошуку ввести запит для супероксиддисмутази людини (**human superoxide dismutase**) і натиснути «Search».

UniProtKB 1,287 results

BLAST Align Map IDs Download Add View: Cards Table Share

- O14618 · CCS\_HUMAN  
Copper chaperone for superoxide dismutase · Gene: CCS · Homo sapiens (Human) · 274 amino acids · Evidence at protein level · Annotation score: 5.5  
#Chaperone  
1 domain · 1 PTM · 1 reviewed variant · 3 interactions · 7 3D structures · 14 reviewed publications
- P00441 · SODC\_HUMAN  
Superoxide dismutase [Cu-Zn] · Gene: SOD1 · Homo sapiens (Human) · EC:1.15.1.1 · 154 amino acids · Evidence at protein level · Annotation score: 5.5  
#Antioxidant #Oxidoreductase #Amyotrophic lateral sclerosis #Disease variant #Neurodegeneration  
12 PTMs · 81 reviewed variants · 79 interactions · 2 diseases · 131 3D structures · 92 reviewed publications
- P04179 · SODM\_HUMAN  
Superoxide dismutase [Mn], mitochondrial · Gene: SOD2 · Homo sapiens (Human) · EC:1.15.1.1 · 222 amino acids · Evidence at protein level · Annotation score: 5.5  
#Oxidoreductase

2. Из результатів пошуку вибираємо той, який відповідає супероксиддисмутазі людини (наприклад з ідентифікатором P00441).

UniProtKB 1,287 results

BLAST Align Map IDs Download Add View: Cards Table Customize columns Share

Entry	Entry Name	Protein Names	Gene Names	Organism	Length
<input type="checkbox"/> O14618	CCS_HUMAN	Copper chaperone for superoxide dismutase[...]	CCS	Homo sapiens (Human)	274 AA
<input checked="" type="checkbox"/> P00441	SODC_HUMAN	Superoxide dismutase [Cu-Zn][...]	SOD1	Homo sapiens (Human)	154 AA
<input type="checkbox"/> P04179	SODM_HUMAN	Superoxide dismutase [Mn], mitochondrial[...]	SOD2	Homo sapiens (Human)	222 AA
<input type="checkbox"/> A0A384NL29	A0A384NL29_HUMAN	Superoxide dismutase[...]		Homo sapiens (Human)	222 AA
<input type="checkbox"/> P08294	SODE_HUMAN	Extracellular superoxide dismutase [Cu-Zn][...]	SOD3	Homo sapiens (Human)	240 AA
<input type="checkbox"/> V9HWC9	V9HWC9_HUMAN	Superoxide dismutase [Cu-Zn][...]	HEL-S-44	Homo sapiens (Human)	154 AA
<input type="checkbox"/> Q9WU84	CCS_MOUSE	Copper chaperone for superoxide dismutase[...]	Ccs, Ccsd	Mus musculus (Mouse)	274 AA
<input type="checkbox"/> P00442	SODC_BOVIN	Superoxide dismutase [Cu-Zn][...]	SOD1	Bos taurus (Bovine)	152 AA
<input type="checkbox"/> P08228	SODC_MOUSE	Superoxide dismutase [Cu-Zn][...]	Sod1	Mus musculus (Mouse)	154 AA

3. Заходимо на цей результат пошуку і знаходимо амінокислотну послідовність. Потім натискаємо **Sequence** і завантажуюмо:

UniProt BLAST Align Peptide search ID mapping SPARQL UniProtKB Advanced | List Search

Function  
Names & Taxonomy  
Subcellular Location  
Disease & Variants  
PTM/Processing  
Expression  
Interaction  
Structure  
Family & Domains  
Sequence  
Similar Proteins

Entry Feature viewer Publications External links History

### Sequence<sup>1</sup>

Tools **Download** Add Highlight Copy sequence

Length 154  
Mass (Da) 15,936

Last updated 2007-01-23 v2  
Checksum<sup>1</sup> 25CA38DA8D564483

MATKAVCVLK<sup>10</sup> GDGVPVQGIIN<sup>20</sup> FEQKESNGPV<sup>30</sup> KVWGSIKGL<sup>40</sup> EGLHGFVHE<sup>50</sup> FGDNTAGCTS<sup>60</sup> AGPHFNPLSR<sup>70</sup> KHGGPKDEER<sup>80</sup> HVGDLGNVTA<sup>90</sup> DKDGVADVSI<sup>100</sup> EDSVISLSGD<sup>110</sup>  
HCCIIGRTL<sup>120</sup> VV HEKADDLGK<sup>130</sup> GNEESTKTGN<sup>140</sup> AGSRLACGVI<sup>150</sup> GIAQ

### Computationally mapped potential isoform sequences<sup>1</sup>

There is 1 potential isoform mapped to this entry

BLAST Align Add View all

Entry	Entry name	Gene name	Length
<input type="checkbox"/> H7BYH4	H7BYH4_HUMAN	SOD1	135

Features  
Showing features for sequence conflict<sup>1</sup>.

4. Отримуємо амінокислотну послідовність у форматі FASTA, яку зберігаємо у заздалегідь створеному файлі.

```
>sp|P00441|SODC_HUMAN Superoxide dismutase [Cu-Zn] OS=Homo sapiens OX=9606 GN=SOD1 PE=1 SV=2
MATKAVCVLKGDDGVPVQGIINFEQKESNGPVKVVWGSIKGLTEGLHGFVHVEFGDNTAGCTS
AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVDVSIEDSVISLSGDHCCIIGRTL VV
HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ
```

5. Кроки 1 – 4 виконуємо для супероксиддисмутази щура (rat superoxide dismutase - ідентифікатор **P07632**), дрозофіли (*Drosophila melanogaster* superoxide dismutase - ідентифікатор **P61851**), кишкової палички (*Escherichia coli* superoxide dismutase - ідентифікатор **P0AGD1**).

6. Здійснюємо вирівнювання послідовностей із використанням програми BLAST - <https://blast.ncbi.nlm.nih.gov>

U.S. National Library of Medicine NCBI: National Center for Biotechnology Information

## BLAST<sup>®</sup>

Home Recent Results Saved Strategies

### Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

**igBLAST 1.8.0 released**  
A new version of igBLAST is now available  
Wed, 15 Nov 2017 16:00:00 EST [More BLAST news...](#)

### Web BLAST

**Nucleotide BLAST**  
nucleotide → nucleotide


**blastx**  
translated nucleotide → protein

**tblastn**  
protein → translated nucleotide

**Protein BLAST**  
protein → protein

7. Вибираємо опцію «Protein BLAST», оскільки ми виконуємо вирівнювання амінокислотних послідовностей.

8. Під час множинного вирівнювання амінокислотних послідовностей вибираємо функцію «Align two or more sequences».


**National Library of Medicine**  
 National Center for Biotechnology Information

[Log in](#)

---

**BLAST® » blastp suite**

[Home](#)
[Recent Results](#)
[Saved Strategies](#)
[Help](#)

**Align Sequences Protein BLAST**

blastn **blastp** blastx tblastn tblastx

BLASTP programs search protein subjects using a protein query. more... [Reset page](#) [Bookmark](#)

**Enter Query Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) Query subrange [?](#)

From   
To

Or, upload file   [?](#)

Job Title

**Align two or more sequences** [?](#) Enter a descriptive title for your BLAST search [?](#)


**Enter Subject Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) Subject subrange [?](#)

From   
To

Or, upload file   [?](#)

9. Відкривається два поля.


**National Library of Medicine**  
 National Center for Biotechnology Information

[Log in](#)

---

**BLAST® » blastp suite**

[Home](#)
[Recent Results](#)
[Saved Strategies](#)
[Help](#)

**Align Sequences Protein BLAST**

blastn **blastp** blastx tblastn tblastx

BLASTP programs search protein subjects using a protein query. more... [Reset page](#) [Bookmark](#)

**Enter Query Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) Query subrange [?](#)

From   
To

Or, upload file   [?](#)

Job Title

**Align two or more sequences** [?](#) Enter a descriptive title for your BLAST search [?](#)

**Enter Subject Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) Subject subrange [?](#)

From   
To

Or, upload file   [?](#)

**Program Selection**

Algorithm  blastp (protein-protein BLAST)  
Choose a BLAST algorithm [?](#)

Search protein sequence using Blastp (protein-protein BLAST)  
 Show results in a new window

В перше поміщаємо амінокислотну послідовність супероксиддисмутази людини, у друге – послідовності супероксиддисмутази щура, дрозофіли, кишкової палички. Натискаємо «**BLAST**»

BLAST **blastp** align blastx tblastn tblastx

Align Sequences Protein BLAST

BLASTP programs search protein subjects using a protein query, more... [Reset page](#) [Bookmark](#)

**Enter Query Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) **Query subrange**

>sp|P00441|SODC\_HUMAN Superoxide dismutase [Cu-Zn] OS=Homo sapiens OX=9606 GN=SOD1 PE=1 SV=2  
 MATKAVCVLKGDPVQGVVHFEQKESNGPVKWGSIKGLTEGLHGFHVHEF  
 GDNTAGCTS

From   
 To

Or, upload file

Job Title   
 Enter a descriptive title for your BLAST search

Align two or more sequences

**Enter Subject Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) **Subject subrange**

>sp|P07632|SODC\_RAT Superoxide dismutase [Cu-Zn] OS=Rattus norvegicus OX=10116 GN=Sod1 PE=1 SV=2  
 MAMKAVCVLKGDPVQGVVHFEQKASGEPVWVSGQITGLTEGEHGFHVHGY  
 GDNTQGCTT

From   
 To

Or, upload file

**Program Selection**

Algorithm  blastp (protein-protein BLAST)  
 Choose a BLAST algorithm

**BLAST**  Show results in a new window

Так виглядають внесені послідовності порівнюваних організмів у другому полі:

blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins&PROGRAM=blastp&BLAST\_PROGRAMS=blastp&PAGE\_TYPE=BlastSearch&... [Reset page](#) [Bookmark](#)

**Enter Query Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) **Query subrange**

GDNTAGCTS  
 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDH  
 GIGRITLVV  
 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ

From   
 To

Or, upload file

Job Title   
 Enter a descriptive title for your BLAST search

Align two or more sequences

**Enter Subject Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) **Subject subrange**

>sp|P07632|SODC\_RAT Superoxide dismutase [Cu-Zn] OS=Rattus norvegicus OX=10116 GN=Sod1 PE=1 SV=2  
 MAMKAVCVLKGDPVQGVVHFEQKASGEPVWVSGQITGLTEGEHGFHVHGY  
 GDNTQGCTT  
 AGPHFNPHSKKHGGPADEERHVGDLGNVAAGKDVANVSIEDRIVLSGEH  
 SIIGRTMNV  
 HEKQDDLKGGNEESTKTGNAGSRLACGVIGIAQ  
 >sp|P61851|SODC\_DROME Superoxide dismutase [Cu-Zn] OS=Drosophila melanogaster OX=7227 GN=Sod1 PE=1 SV=2  
 MIVKAVCVINGDAKGTVFFEQESSGTPVKVSGEVCGLAKGLHGFHVHEFGD  
 NTNGQIMSSG  
 PHFNPFYGKEHGAPVDENRHLGDLGNIEATGDCPTKVNITDSKITLFGADSIIG  
 RTLVVHA  
 DADDLGGGHELKSKSTGNAGARIGCGVIGIAKV  
 >sp|P0AGD1|SODC\_ECOLI Superoxide dismutase [Cu-Zn] OS=Escherichia coli (strain K12) OX=83333 GN=sodC PE=1 SV=1  
 MKRFSLAIALVATGAQAASEKVEMLNLYTSQGVGQSIGVITITETDKGLEFS  
 PDLKALP  
 PGEHGFHIAKGSQPATKDGKASAAESAGGHLDPQNTGKHEGPEGAGHL  
 GDLPALVNN

From   
 To

Or, upload file

10. Отримуємо результат множинного вирівнювання послідовностей у людини та обраних організмів:



[← Edit Search](#) Save Search Search Summary ▾

How to read this report? BLAST Help Videos Back to Traditional Results Page

Job Title	sp P00441 SODC_HUMAN Superoxide dismutase...
RID	ZC7YNNY5114 Search expires on 03-18 06:40 am Download All ▾
Program	Blast 2 sequences Citation ▾
Query ID	lcl Query_4488789 (amino acid)
Query Descr	sp P00441 SODC_HUMAN Superoxide dismutase [Cu-Zn] ( ...
Query Length	154
Subject ID	lcl Query_4488791 and 2 more subject(s) (amino acid)
Subject Descr	See details ▾
Subject Length	480

Filter Results

Percent Identity  to  E value  to  Query Coverage  to

[Filter](#) [Reset](#)

**Descriptions** Graphic Summary Alignments

Sequences producing significant alignments

Download ▾ Select columns ▾ Show 100 ▾ ?

select all 3

[Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	sp P07632 SODC_RAT Superoxide dismutase [Cu-Zn]. OS=Rattus norvegicus OX=10116 GN=Sod...		258	258	100%	1e-94	83.12%	154	Query_4488791
<input checked="" type="checkbox"/>	sp P61851 SODC_DROME Superoxide dismutase [Cu-Zn]. OS=Drosophila melanogaster OX=7227...		186	186	100%	1e-65	61.04%	153	Query_4488792
<input checked="" type="checkbox"/>	sp P0AGD1 SODC_ECOLI Superoxide dismutase [Cu-Zn]. OS=Escherichia coli (strain K12) OX=83...		49.3	49.3	87%	2e-12	31.03%	173	Query_4488793

Натискаємо на Alignments (вирівнювання)

[← Edit Search](#) Save Search Search Summary ▾

How to read this report? BLAST Help Videos Back to Traditional Results Page

Job Title	sp P00441 SODC_HUMAN Superoxide dismutase...
RID	ZC7YNNY5114 Search expires on 03-18 06:40 am Download All ▾
Program	Blast 2 sequences Citation ▾
Query ID	lcl Query_4488789 (amino acid)
Query Descr	sp P00441 SODC_HUMAN Superoxide dismutase [Cu-Zn] ( ...
Query Length	154
Subject ID	lcl Query_4488791 and 2 more subject(s) (amino acid)
Subject Descr	See details ▾
Subject Length	480

Filter Results

Percent Identity  to  E value  to  Query Coverage  to

[Filter](#) [Reset](#)

**Descriptions** Graphic Summary **Alignments**

Sequences producing significant alignments

Download ▾ Select columns ▾ Show 100 ▾ ?

select all 3

[Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	sp P07632 SODC_RAT Superoxide dismutase [Cu-Zn]. OS=Rattus norvegicus OX=10116 GN=Sod...		258	258	100%	1e-94	83.12%	154	Query_4488791
<input checked="" type="checkbox"/>	sp P61851 SODC_DROME Superoxide dismutase [Cu-Zn]. OS=Drosophila melanogaster OX=7227...		186	186	100%	1e-65	61.04%	153	Query_4488792
<input checked="" type="checkbox"/>	sp P0AGD1 SODC_ECOLI Superoxide dismutase [Cu-Zn]. OS=Escherichia coli (strain K12) OX=83...		49.3	49.3	87%	2e-12	31.03%	173	Query_4488793

і отримуємо вирівнювання амінокислотних послідовностей для кожного організму у порівнянні з людиною:

Щур:

Descriptions Graphic Summary **Alignments**

Alignment view Pairwise [Restore defaults](#) [Download](#)

2 [?](#)

[Download](#) [Graphics](#) [Next](#) [Previous](#) [Descriptions](#)

**sp|P07632|SODC\_RAT Superoxide dismutase [Cu-Zn] OS=Rattus norvegicus OX=10116 GN=Sod1 PE=1 SV=2**  
 Sequence ID: **Query\_28057** Length: 154 Number of Matches: 1

Range 1: 1 to 154 [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
258 bits(660)	2e-94	Compositional matrix adjust.	128/154(83%)	137/154(88%)	0/154(0%)
Query 1	MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFGDNTAGCTS	60			
Sbjct 1	MA KAVCVLKGDPVQGI+FEQK S PV V G I GLTEG HGFHVH+GDNT GCT+MAMKAVCVLKGDPVQGIHFEQKASGEPVWVSGQITGLTEGEHGFHVHGYDNTQGCTT	60			
Query 61	AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLWV	120			
Sbjct 61	AGPHFNP S+KHGGP DEERHVGDLGNV A KDGVA+VSIED VISLSG+H IIGRT+VVAGPHFNPHSKKHGGPDAEERHVGDLGNVAAGKDVANVSIEDRVISLSGEHSIIGRTMVV	120			
Query 121	HEKADDLGGKGGNEESTKTGNAGSRLACGVIGIAQ	154			
Sbjct 121	HEK DDLGGKGGNEESTKTGNAGSRLACGVIGIAQ	154			

[Download](#) [Graphics](#) [Next](#) [Previous](#) [Descriptions](#)

**sp|P61851|SODC\_DROME Superoxide dismutase [Cu-Zn] OS=Drosophila melanogaster OX=7227 GN=Sod1 PE=1 SV=2**  
 Sequence ID: **Query\_28058** Length: 153 Number of Matches: 1

Дрозофіла:

[Download](#) [Graphics](#) [Next](#) [Previous](#) [Descriptions](#)

**sp|P61851|SODC\_DROME Superoxide dismutase [Cu-Zn] OS=Drosophila melanogaster OX=7227 GN=Sod1 PE=1 SV=2**  
 Sequence ID: **Query\_28058** Length: 153 Number of Matches: 1

Range 1: 1 to 152 [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
186 bits(471)	1e-65	Compositional matrix adjust.	94/154(61%)	113/154(73%)	2/154(1%)
Query 1	MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFGDNTAGCTS	60			
Sbjct 1	M KAVCV+ GD +G + FEQ+ S PVKV G + GL +GLHGFHVHFGDNT GC S MWWKAVCVINGD--AKGTVFFEQESSGTPVKVSGEVCGLAKGLHGFHVHFGDNTNGCMS	58			
Query 61	AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGADVSIEDSVISLSDGHCIIIGRTLWV	120			
Sbjct 59	+GPHFNP ++HG P DE RH+GDLGN+ A D V+I DS I+L G IIGRT+VVSGPHFNPNYKHEGAPVDENRHLGDLGNIEATGDCPTKVNITDSKITLFGADSIIGRTVV	118			
Query 121	HEKADDLGGKGGNEESTKTGNAGSRLACGVIGIAQ	154			
Sbjct 119	H ADDLG+GG+E S TGNAG+R+ CGVIGIA+HADADDLGGGGHLSKSTGNAGARIGCGVIGIAK	152			

E. Coli^

**sp|P0AGD1|SODC\_ECOLI Superoxide dismutase [Cu-Zn] OS=Escherichia coli (strain K12) OX=83333 GN=sodC PE=1 SV=1**

Sequence ID: **Query\_4488793** Length: **173** Number of Matches: **1**

Range 1: 36 to 172 [Graphics](#)

[Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
49.3 bits(116)	2e-12	Compositional matrix adjust.	45/145(31%)	67/145(46%)	18/145(12%)
Query 16	QGIINFEQKESNGPVKVGWSIKGLTEGLHGFHVHEFGD-----NTAGCTSAGPHFNP				67
	Q I + E++ ++ +K L G HGFH+H G + SAG H +P				
Sbjct 36	QSIGSVTITETDKGLEFSPDLKALPPGEGHGFHIAKAGSCQPATKDGKASAAESAGGHLDP				95
Query 68	LSR-KHGGPKDEERHVGDLGNVTADKDGVA-DVSIEDSVISLSGDHCIIGRTLVVEKAD				125
	+ KH GP + H+GDL + + DG A D I + SL I + L+VH D				
Sbjct 96	QNTGKHEGP-EGAGHLGDLPALVWVNDGKATDAVIAPRLKSLDE---IKDKALMVHVGDD				151
Query 126	DLGKGGNEESTKTGNAGSRLACGVI 150				
	++ +++ G G R ACGVI				
Sbjct 152	NM----SDQPKPLGGGGERYACGVI 172				

Натискаючи **Graphics**, можна подивитися де присутні зміни в послідовностях амінокислот в організмах порівняно з людиною.

**sp|P07632|SODC\_RAT Superoxide dismutase [Cu-Zn] OS=Rattus norvegicus OX=10116 GN=Sod1 PE=1 SV=2**  
 Icl|Query\_28057

Sequence Viewer 3.47.0

Query\_28057: 1..154 (154 aa)

Contact | Copyright | Disclaimer | Privacy | Accessibility  
 National Center for Biotechnology Information, US National Library of Medicine  
 8600 Rockville Pike, Rockville, MD USA 20894

**sp|P61851|SODC\_DROME Superoxide dismutase [Cu-Zn] OS=Drosophila melanogaster OX=7227 GN=Sod1 PE=1 SV=2**  
 Icl|Query\_28058

Sequence Viewer 3.47.0

Query\_28058: 1..153 (153 aa)

Contact | Copyright | Disclaimer | Privacy | Accessibility  
 National Center for Biotechnology Information, US National Library of Medicine  
 8600 Rockville Pike, Rockville, MD USA 20894

**National Library of Medicine**  
 National Center for Biotechnology Information

[Log in](#)

NCBI Sequence Viewer 3.49.0    [Help](#)    [Release Notes](#)

**sp|P0AGD1|SODC\_ECOLI Superoxide dismutase [Cu-Zn] OS=Escherichia coli (strain K12) OX=83333 GN=sodC PE=1 SV=1**  
 Icl|Query\_4488793

Sequence Viewer 3.49.0

Query\_4488793: 25..173 (149 aa)

Contact | Copyright | Disclaimer | Privacy | Accessibility  
 National Center for Biotechnology Information, US National Library of Medicine  
 8600 Rockville Pike, Rockville, MD USA 20894

11. Для побудови філогенетичного дерева натиснути **Descriptions**

[< Edit Search](#) [Save Search](#) [Search Summary](#) ▾[? How to read this report?](#) [▶ BLAST Help Videos](#) [↶ Back to Traditional Results Page](#)

Job Title sp|P00441|SODC\_HUMAN Superoxide dismutase...

RID [0C9472FR114](#) Search expires on 03-08 00:36 am [Download All](#) ▾

Program Blast 2 sequences [Citation](#) ▾

Query ID lc|Query\_28055 (amino acid)

Query Descr sp|P00441|SODC\_HUMAN Superoxide dismutase [Cu-Zn ...

Query Length 154

Subject ID lc|Query\_28057 and 2 more subject(s) (amino acid)

Subject Descr [See details](#) ▾

Subject 513

Length

## Filter Results

Percent Identity  to

E value  to

Query Coverage  to

[Filter](#) [Reset](#)

[Descriptions](#) [Graphic Summary](#) [Alignments](#)

## Sequences producing significant alignments

[Download](#) ▾ [Select columns](#) ▾ Show  [?](#) select all 2 sequences selected[Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> sp P07632 SODC_RAT Superoxide dismutase [Cu-Zn] OS=Rattus norvegicus OX=10116 GN=Sod1 PE=1 SV=2		258	258	100%	2e-94	83.12%	154	Query_28057
<input checked="" type="checkbox"/> sp P61851 SODC_DROME Супероксиддисмутаза [Cu-Zn] OS=Drosophila melanogaster OX=7227 GN=Sod...		186	186	100%	1e-65	61.04%	153	Query_28058

## Потім Distance tree of results

Job Title sp|P00441|SODC\_HUMAN Superoxide dismutase...

RID [0C9472FR114](#) Search expires on 03-08 00:36 am [Download All](#) ▾

Program Blast 2 sequences [Citation](#) ▾

Query ID lc|Query\_28055 (amino acid)

Query Descr sp|P00441|SODC\_HUMAN Superoxide dismutase [Cu-Zn ...

Query Length 154

Subject ID lc|Query\_28057 and 2 more subject(s) (amino acid)

Subject Descr [See details](#) ▾

Subject 513

Length

## Filter Results

Percent Identity  to

E value  to

Query Coverage  to

[Filter](#) [Reset](#)

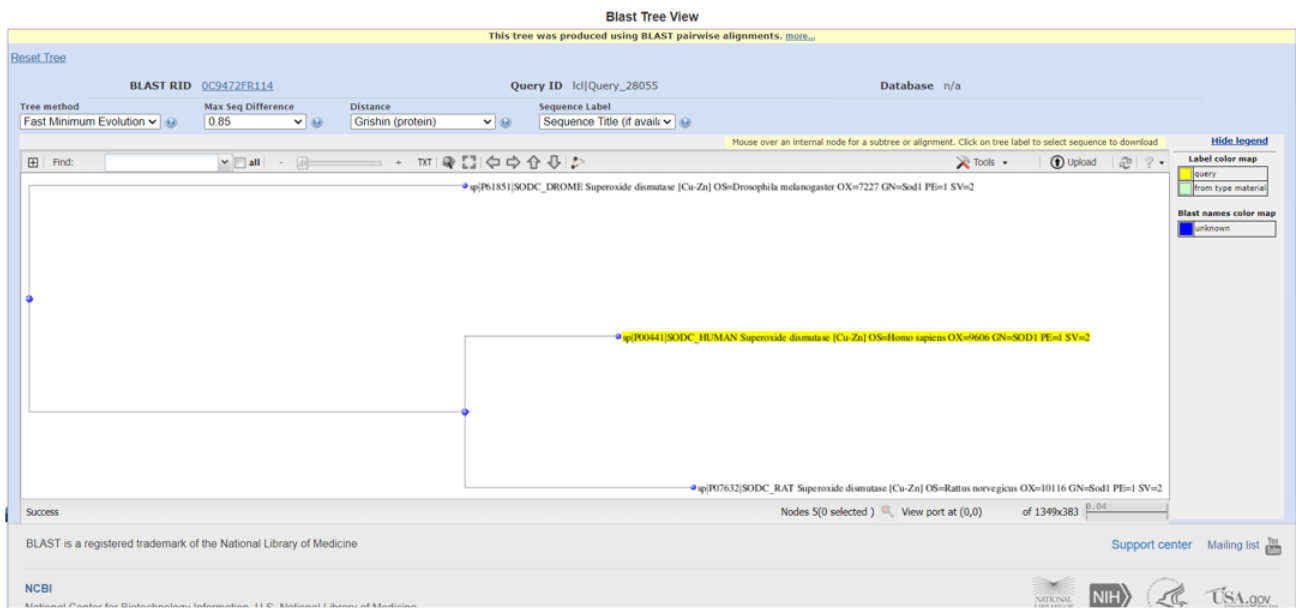
[Descriptions](#) [Graphic Summary](#) [Alignments](#)

## Sequences producing significant alignments

[Download](#) ▾ [Select columns](#) ▾ Show  [?](#) select all 2 sequences selected[Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> sp P07632 SODC_RAT Superoxide dismutase [Cu-Zn] OS=Rattus norvegicus OX=10116 GN=Sod1 PE=1 SV=2		258	258	100%	2e-94	83.12%	154	Query_28057
<input checked="" type="checkbox"/> sp P61851 SODC_DROME Супероксиддисмутаза [Cu-Zn] OS=Drosophila melanogaster OX=7227 GN=Sod...		186	186	100%	1e-65	61.04%	153	Query_28058

Отримуємо філогенетичне дерево.



Зробити висновок про те, які організми є найспорідненішими. Чому кишкова паличка відсутня у філогенетичному дереві.

## Завдання 2.

На основі амінокислотної послідовності аспартатамінотрансферази людини за допомогою інтернет-ресурсів визначте молекулярну масу та ізоелектричну точку даного ферменту.

### Хід виконання

1. Знаходимо амінокислотну послідовність у FASTA-форматі для аспартатамінотрансферази людини (human aspartate aminotransferase) через веб-ресурс UniProt (див. завдання 1).

```
>sp|P17174|AATC_HUMAN Aspartate aminotransferase, cytoplasmic OS=Homo sapiens OX=9606
GN=GOT1 PE=1 SV=3
MAPPSVFAEVPQAQPVLVFKLTADFREDPDPRKVN LGVGAYRTDDCHPWVLPVVKKVEQK
IANDNSLNHEYLPILGLAEFRSCASRLALGDDSPALKEKRVGGVQSLGGTGALRIGADFL
ARWYNGTNNKNTPVYVSSPTWENHNAVFS AAGFKDIRSYRYWDAEKRLDLQGF LNDLEN
APEFSIVVLHACAHNPTGIDPTPEQWKQIASVMKHRFLFPFFDSAYQGFASGNLERDAWA
IRYFVSEGFEEFFCAQSFSKNFGLYNERVGNLTVVGKEPESILQVLSQMEKIVRITWSNPP
AQGARIVASTLSNPELFEWGTGNVKT MADRILLTMRSELRARLEALKTPGTWNHITDQIGM
FSFTGLNPKQVEYLVNEKHIYLLPSGRINVSGLT TKNLDYVATS IHEAVTKIQ
```

2. Увійти на веб-ресурс ExPASy (<https://www.expasy.org>). У меню головного пошукового екрана зліва вибираємо опцію «**Proteins & Proteomes**»

Posted 6 March 2023 - **Studying the diversity of microorganisms in biological samples?** The latest version of [mOTUs](#) allows the profiling of metagenomes for >33,000 species-level operational taxonomic units.











# Expasy

Swiss Bioinformatics Resource Portal

e.g. [BLAST](#), [UniProt](#), [MSH6](#), [Albumin](#)...

- Genes & Genomes**
  - Genomics
  - Metagenomics
  - Transcriptomics
- Proteins & Proteomes**
- Evolution & Phylogeny**
  - Evolution biology
  - Population genetics

## SIB Resources

 <p><b>Nextstrain</b> Impact of pathogen genome data on science and public health</p>	 <p><b>SwissOrthology</b> One-stop shop for orthologs</p>	 <p><b>STRING</b> Protein-protein interaction networks and enrichment analysis</p>	 <p><b>ASAP</b> Web-based, cooperative portal for single-cell data analyses</p>
 <p><b>Cellosaurus</b></p>	 <p><b>SwissRegulon Portal</b></p>	 <p><b>Bgee</b></p>	 <p><b>SwissLipids</b></p>

Posted 6 March 2023 - **Studying the diversity of microorganisms in biological samples?** The latest version of [mOTUs](#) allows the profiling of metagenomes for >33,000 species-level operational taxonomic units.









# Expasy

Swiss Bioinformatics Resource Portal

















e.g. [BLAST](#), [UniProt](#), [MSH6](#), [Albumin](#)...


- Genes & Genomes**
  - Genomics
  - Metagenomics
  - Transcriptomics
- Proteins & Proteomes**
- Evolution & Phylogeny**
  - Evolution biology
  - Population genetics

## SIB Resources

 <p><b>STRING</b> Protein-protein interaction networks and enrichment analysis</p>	 <p><b>SwissLipids</b> Knowledge resource for lipids</p>	 <p><b>UniProtKB/Swiss-Prot</b> Protein knowledgebase</p>	 <p><b>SWISS-MODEL</b> Protein structure homology-modelling</p>
 <p><b>Glyco@Expasy</b></p>	 <p><b>Rhea</b></p>		

3. Вибрати ресурс «Compute pI/Mw».

 <p><b>UniLectin</b> Structural and functional classification of lectins</p>	 <p><b>Decrease redundancy</b> Redundancy reduction in a set of sequences</p>	 <p><b>MALDI PepQuant</b> Quantify MALDI peptides from Phenix output</p>	 <p><b>PROSITE</b> Protein family and domain database</p>
 <p><b>Compute pI/MW</b> Isoelectric point and molecular weight from protein sequence</p>	 <p><b>SWISS-MODEL Repository</b> Protein structure homology models</p>	 <p><b>SwissTargetPrediction</b> Target prediction for bioactive small molecules</p>	 <p><b>RandSeq</b> Random protein sequence generator</p>
 <p><b>OrthoLogger</b> Accurate and scalable inference of groups of orthologs</p>	 <p><b>GlycoSiteAlign</b> Multiple alignment of sequences around glycosylation sites</p>	 <p><b>Swiss Mass Abacus</b> Intuitive calculator of peptide and glycopeptide masses</p>	 <p><b>HAMAP SPARQL endpoint</b> Use HAMAP + SPARQL to generate portable annotation pipelines</p>
 <p><b>Dotlet</b> Regions of close similarity in 2 sequences</p>	 <p><b>Rhea SPARQL endpoint</b> SPARQL access to the Rhea DB</p>	 <p><b>SWISS-MODEL Workspace</b> Fully automated protein structure homology-modeling server</p>	 <p><b>HAMAP-Scan</b> Scanning of protein sequences against HAMAP families</p>



e.g. BLAST, UniProt, MSH6, Albumin...

## Compute pI/MW

Proteins & Proteomes Software tool

Compute the theoretical pI (isoelectric point) and Mw (molecular weight) for a list of UniProt Knowledgebase (Swiss-Prot or TrEMBL) entries or for user entered sequences

[Browse the resource website](#)

Hosted by the [Swiss-Prot group](#).

What you can do with this resource

[Protein sequence analysis](#),  
[Protein molecular weight calculation](#),  
[Protein isoelectric point calculation](#)

Browse these keywords in Expasy

[Protein properties](#),  
[Molecular property](#),  
[UniProt accession](#), [Protein sequence](#)

You might also be interested in

поле ресурсу вставляємо скопійовану амінокислотну послідовність і натискаємо «Click here to Compute pI/Mw».

4. У

MAPPSVFAEVPQAQPVLVFKLTADFREDPDPKRVNLGVGAYRTDDCHPWVLPVVKKVEQK  
 IANDNSLNHEYLPILGLAEFRSCASRLALGDDSPALKEKRVGGVQSLGGTGALRIGADFL  
 ARWYNGTNNKNTVPVYVSSPTWENHNAVSAAGFKDIRSYRYWDAEKRGLDLQGFLNDLEN  
 APEFSIVVLHACAHNPTGIDPTPEQWKQIASVMKHRFLFPFFDSAYQGFASGNLERDAWA  
 IRYFVSEGFEFFCAQSFSKNFLYNERVGNLTVVGKEPESILQVLSQMEKIVRITWSNPP  
 AQGARIVASTLSNPELFEWTGNVKTMAADRILTMRSRLRARLEALKTPGTWNHITDQIGM  
 FSFTGLNPKQVEYLVNEKHIIYLLPSGRINVSGLTTKNLDYVATSIHEAVTKIQ

## Compute pI/Mw tool

**Compute pI/Mw** is a tool which allows the computation of the theoretical pI (isoelectric point) and Mw (molecular weight) for a list of UniProt Knowledgebase (Swiss-Prot or TrEMBL) entries or for user entered sequences [reference].

Documentation is available.

## Compute pI/Mw for Swiss-Prot/TrEMBL entries or a user-entered sequence

Please enter one or more UniProtKB/Swiss-Prot protein identifiers (ID) (e.g. *ALBU\_HUMAN*) or UniProt Knowledgebase accession numbers (AC) (e.g. *P04406*), separated by spaces, tabs or newlines. Alternatively, enter a protein sequence in single letter code. The theoretical pI and Mw (molecular weight) will then be computed.

```
ARWYNGTINKNTPVYVSSPTWENHNAVFSAAAGKDIRSYR
YNDKEKRGLDQGLNDLEN
APEFSIVVLHACAHNPTGIDPTPEQWKQIASVMKHRFLFP
FFDSAYQGFASGNLERDAWA
IRYFVSEGEFFCAQSFKNFGLYNERVGNLTVVGKEPES
ILQVLSQMEKIVRITWSNPP
AQGARIIVASTLSNPELFEETGKVTMADRILTMRSELRA
RLEALKTPGTWNHITDQIGM
FSFTGLNPKQVEYLVNEKHIIYLLPSGRINVSGLTTKNLDY
VATSIHEAVTKIQ
```

Or upload a file from your computer, containing one Swiss-Prot/TrEMBL ID/AC or one sequence per line:  Файл не выбран

Resolution:  Average or  Monoisotopic

## 5. Отримуємо результат.

**Expasy**
Compute pI/Mw

---

**Compute pI/Mw**

**Theoretical pI/Mw (average) for the user-entered sequence:**

10	20	30	40	50	60
MAPPSPVFAEV	PQAQPVLVFK	LTADFREDPD	PRKVNLVGVA	YRTDDCHPWV	LPVVKKVEQK
70	80	90	100	110	120
IANDNSLNHE	YLPILGLAEF	RSCASRLALG	DDSPALKEKR	VGGVQSLGGT	GALRIGADFL
130	140	150	160	170	180
ARWYNGTINKN	NTPVYVSSPT	WENHNAVFS	AGFKDIRSYR	YWDAEKRGD	LQGFNDLEN
190	200	210	220	230	240
APEFSIVVLH	ACAHNPTGID	PTPEQWKQIA	SVMKHRFLFP	FFDSAYQGFA	SGNLERDAWA
250	260	270	280	290	300
IRYFVSEGEF	FFCAQSFKN	FGLYNERVGN	LTVVGKEPES	ILQVLSQMEK	IVRITWSNPP
310	320	330	340	350	360
AQGARIIVAST	LSNPELFEET	TGNVKTADR	ILTMRSELRA	RLEALKTPGT	WNHITDQIGM
370	380	390	400	410	
FSFTGLNPKQ	VEYLVNEKHI	YLLPSGRINV	SGLTTKNLDY	VATSIHEAVT	KIQ

Theoretical pI/Mw: 6.53 / 46247.51

Встановлено, що ізоелектрична точка та молекулярна маса аспартатамінотрансферази людини становлять 6,53 та 46247,51 відповідно.

**Завдання для самостійного виконання:**

1. Виконайте вирівнювання амінокислотної послідовності білка людини (*Pancreatic triacylglycerol lipase, Pancreatic alpha-amylase, Gastricsin, Myeloperoxidase, Catalase, Trypsin, Glucose-6-phosphate 1-dehydrogenase, L-lactate dehydrogenase та інших*) проти амінокислотної послідовності відповідного білка ссавця, птаха, риби, мікроорганізму тощо. **Звернути увагу, що об'єкт вибору для кожного студента вказаний на сайті <https://moodle.znu.edu.ua/>. Опишіть яку функцію виконує Ваш білок.**

**Результати роботи надати у вигляді скріншотів результату вирівнювання: 1) амінокислотні послідовності білка людини та інших тварин; 2) графіки порівняння; 3) філогенетичне дерево**

2. На основі амінокислотної послідовності обраного вище ферменту (human *фермент*) людини за допомогою інтернет-ресурсів визначте молекулярну масу та ізоелектричну точку даного ферменту. **Результати роботи надати у вигляді скріншоту за п. 5.**