

SciFinder<sup>N</sup>

Biosequence Search

BLAST, Nucleotides, Proteins,  
Commonly Determined Regions  
and MOTIFS

# Table of Contents

I.	Biosequence Search using a FASTA File-Nucleotide .....	3
II.	Biosequence Search using a FASTA File -Protein.....	5
III.	Biosequence -Influenza Virus using the following Protein Sequence (H1N1-2009) .....	11
IV.	Commonly Determined Region Search.....	21
V.	MOTIF Search.....	27

# I. Biosequence Search using a FASTA File-Nucleotide

FASTA File (Nucleotide File) for SARS-CoV-2 virus

NC\_000001:Homo sapiens chromosome 1, GRCh38.p13..

Copy the highlighted Genetic Nucleotide String and input it into the BLAST Sequence Upload location

GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTGTTCAACTCACAGTTT

Biosequences Settings Prior to Search in this screen

The screenshot displays the 'Biosequences' search interface. At the top, there are tabs for 'BLAST', 'CDR', and 'Motif', with 'BLAST' selected. Below the tabs is a text input field containing the nucleotide sequence: GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTGTTCAACTCACAGTTT. To the right of the input field are buttons for 'Upload Sequence' and 'Clear Search'. Below the input field is a 'Sequence Type' dropdown menu with 'Nucleotide' selected and 'Protein' as an option. Below that is a 'Search Within' section with radio buttons for 'Nucleotides' (selected) and 'Proteins'. There is also a checked checkbox for 'Include NCBI Sequences'. Below these is a 'Limit Total Sequence Results to:' dropdown menu set to '10'. At the bottom right is a blue button with a magnifying glass icon and the text 'Start Biosequence Search'. Below the main search area is an 'Advanced Biosequence Search' section with various parameters: 'Alignment Identity %' (80), 'Query Coverage %' (90), 'BLAST Algorithm' (BLASTn), 'Match with Gaps?' (No), 'Word Size' (11), 'E-Value' (10), 'Gap Costs' (Existence 5 Extension 2), 'Reward for Match' (2), 'Penalty for Mismatch' (-3), and 'Exclude Low Complexity Regions' (No).

After entering the Nucleotide String, click “Start Biosequence Search” & results are generated.

January 11, 2022

2:28 PM

🔗 Biosequences

Sequence Type: Nucleotide  
Search Within: Nucleotides  
NCBI Included: Yes  
BLAST Algorithm: BLASTn  
Alignment Identity: 80%  
Query Coverage: 90%

GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTGGTTC  
AACTCAGTTTT

View Results

Edit Search

Complete

View All

## View results

**BLAST Search Details**

Sequence Type: Nucleotide  
Search Within: Nucleotides  
BLAST Algorithm: BLASTn  
NCBI Included: Yes  
Alignment Identity: 80%  
Query Coverage: 90%  
E-Value: 10  
Match with Gaps?: No  
Gap Costs: Existence 5  
Extension 2  
Word Size: 11

**Bioscape Analysis**

Visually explore sequence similarity with a new tool.  
Learn more about Bioscape.

Create Bioscape Analysis

Filter by

E-Value: 0 to 10<sup>6</sup>

Query Coverage %: 0 to 100

Subject Coverage %: 0 to 100

Alignment Identity %: 0 to 100

**Biosequences (1)** View: Expanded

References

Query Details: GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTGGTTCAACTCAGTTTT View More

1 Alignment Identity: 100%

Query 1 60

Subject 1 60

Matches: 60  
Mismatches: 0

View Less

Alignment Subject References

Alignment Data  
BLAST Score: 120  
E-Value: 9.334e-23

```
Q 1 GATTTGGGGT TCAAAGCAGT ATCGATCAA TAGTAAATCC ATTTGGTTCAA CTCACAGTTT 60
S 1 GATTTGGGGT TCAAAGCAGT ATCGATCAA TAGTAAATCC ATTTGGTTCAA CTCACAGTTT 60
```

The results returned compares the inputted Genetic Nucleotide String against those found by the Biosequences Search. In the above case, there is 100% matching of the complete Genetic String (60 matches, 0 mismatches). The graphical representation shows identical corroboration of inputted versus “found” strings.

## II. Biosequence Search using a FASTA File -Protein

FASTA File (**Protein Search**) - Calcium-modulated protein

Sequence ID: [XP\\_022517811.1](#)

Copy this File to the BLAST Sequence Upload location. Each letter highlighted letter in this protein string is an amino acid.

```
MADQLTEEQIAEFKEAFSLFDKDGDTITTKELGTVMRSLGQNPTAEALQDMINEVDADGNGTIDFPEFLTMMARKM  
KDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGGQVNYEEFVQMMTAK
```

Settings Prior to Search

### Biosequences

Enter a protein or nucleotide string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST CDR Motif Upload Sequence Clear Search

```
MADQLTEEQIAEFKEAFSLFDKDGDTITTKELGTVMRSLGQNPTAEALQDMINEVDADGNGTIDFPEFLTMMARKM  
KDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGGQVNYEEFVQMMTAK
```

Sequence Type:  
 Nucleotide  Protein

Search Within:  
 Nucleotides  Proteins

Include NCBI Sequences

Limit Total Sequence Results to:  
10

---

Advanced Biosequence Search ^ Adjust Parameters for Short Sequences | Reset All

Alignment Identity %  Match with Gaps?  Yes  No Gap Costs

Query Coverage %  Word Size  Scoring Matrix

BLAST Algorithm  E-Value  Exclude Low Complexity Regions  Yes  No

Press "Start Biosequence Search"

Recent Search History

December 28, 2021

2:18 PM

🔍 Biosequences

Sequence Type: **Protein**  
Search Within: **Proteins**  
NCBI Included: **Yes**  
BLAST Algorithm: **BLASTp**  
Alignment Identity: -  
Query Coverage: **90%**

MADQLTEEQIAEFKREAFSLFDKDGDTITTKELGTVMRSLGQNPTEAELQDMI  
NEVDADGNGTIDFPEFLTMARKMKDTSSEEIIEAFRVFDKDGNGYISAAEL  
RHVMTNLGEKLTDEEVDEMIREADIDGGQVNYEEFVQMMTAK

View Results

Edit Options

Searching...

Results will expire on  
Jan 28, 2022.

## Biosequence Search is Completed

Recent Search History

December 28, 2021

2:18 PM

🔍 Biosequences

Sequence Type: **Protein**  
Search Within: **Proteins**  
NCBI Included: **Yes**  
BLAST Algorithm: **BLASTp**  
Alignment Identity: -  
Query Coverage: **90%**

MADQLTEEQIAEFKREAFSLFDKDGDTITTKELGTVMRSLGQNPTEAELQDMI  
NEVDADGNGTIDFPEFLTMARKMKDTSSEEIIEAFRVFDKDGNGYISAAEL  
RHVMTNLGEKLTDEEVDEMIREADIDGGQVNYEEFVQMMTAK

View Results

Edit Search

Complete

Results will expire on  
Jan 28, 2022.

Choose "View Results". The following 3 screens are all for search results returned.

References



Query Details MADQLTEEQIAEFKEAFSLFDKDGDTITT KELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPPEFLTMMARKMKDT... View Less

```
MADQLTEEQIAEFKEAFSLFDKDGDTITT KELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYISAAE
LRHVMTNLGEKLTDEEVDEMIREADIDGGQVNYEEFVQMMTAK
```

1

Alignment Identity: 100%



Matches: 149  
Mismatches: 0

View Less

Alignment

Subject

References

References

Alignment Data

BLAST Score: 759

E-Value: 1.12043e-93

```
Q      1  MADQLTEEQI AEFKEAFSLF DKDGDGTITT KELGTVMRSL GQNPTEAELQ DMINEVDADG NGTIDFPPEFL 70
      |||
S     39  MADQLTEEQI AEFKEAFSLF DKDGDGTITT KELGTVMRSL GQNPTEAELQ DMINEVDADG NGTIDFPPEFL 108
Q     71  TMMARKMKDT DSEEEIREAF RVFDKDGNGY ISAAELRHVM TNLGEKLTDE EVDEMIREAD IDGGQVNYE 140
      |||
S    109  TMMARKMKDT DSEEEIREAF RVFDKDGNGY ISAAELRHVM TNLGEKLTDE EVDEMIREAD IDGGQVNYE 178
Q    141  EFVQMMTAK 149
      |||
S    179  EFVQMMTAK 187
```

Click on "Subject". In this screen and the one above, 149 amino acids of the query 973 length Genetic String match the returned results. In this case, the Protein string (subject) found is a small portion of the queried string.

References



Query Details **MADQLTEEQIAEFKEAFSLFDKDGDTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDPFPEFLTMMARKMKDT...** [View More](#)

1 Alignment Identity: 100%



Matches: 149  
Mismatches: 0

[View Less](#) ▾

Alignment **Subject** References

References

CAS Registry Number: -  
Length: 973 aa

Sequence

```

1  MHHHHHSSG LVPRGSGMKE TAAAKFERQH MDSPDLGTMA DQLTEEQIAE FKEAFSLFDK DGDGTITTKELGTVMRSLGQ
81  NPTEAELQDM INEVDADGNG TIDPFPEFLTM MARKMKDTS EEEIREAFRV FDKDNGYIS AELRHVMTN LGEKLTDEEV
161  DEMIREADID GDGQVNYEEF VQMMTAKGRP MGGATSAQIE MNKKREAEFQ KMRDLEEAT LQHEATAATL RKKHADSVAE
241  LGEQIDNLQR VKQKLEKEKS EMKMEIDDLA SNMEVSKSK GNLEKMCRTL EDQVSELKTK EEEQQRLINE LTAQRGRLQT
321  ESGEYSRQLD EKDSLVSQLS RKGQAFQTI EELKRQLEEE IKAKSALAHA LQSSRHDCDL LREQYEEEQE AKAELQRAMS
401  KANSEVAQWR TKYETDAIQR TEELEEAKK LAQRLQDAE HVEAVNAKCA SLEKTKQRLQ NEVEDLMIDV ERTNAACAAL
481  DKKQRNFDKI LAEWKQKYEE THAELEASQK ESRSLSTELF KIRNAYEESL DHLETLKREN KNLQQEISDL TEQIAEGGKR
561  IHELEKIKKQ IEQEKSELQA ALEEAASLE HEEGKILRIQ LELNQVKSEI DRKIAEKDEE IDQLKRNHIR VVESMQSTLD
641  AEIRSRNDAI RLKKRMEGDL NEMEIQLNHS NRMAAEALRN YRNTQGILKD TQLHLDDALR GQEDLKEQLA MVERRANLLQ
721  AEIEELRATL EQTERSRIA EQELLDASER VQLLHTQNTS LINTKKKLET DISQIQGEME DIVQEARNAE EKAKKAITDA
    
```

Click on "References"



References



Query Details MADQLTEEQIAEFKEAFSLFDKDGDTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPPEFLTMMARKMKDT... [View More](#)

1 Alignment Identity: 100%



Matches: 149  
Mismatches: 0

[View Less ▾](#)

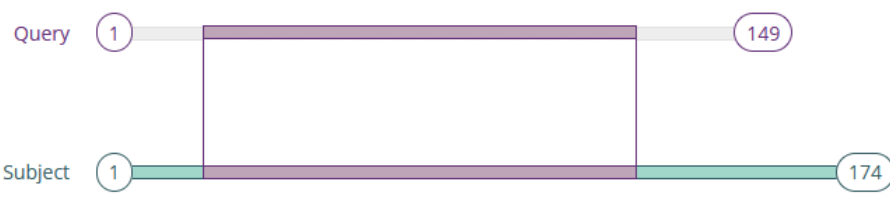
- Alignment
- Subject
- References
- References

モータータンパク質、人工筋肉、マイクロアクチュエータおよびこれらの製造方法  
 Assignee: 国立大学法人北陸先端科学技術大学院大学  
 JP2020050606 A | Seq ID No: 12

Scroll to the next Biosequence found in the original search.

Results for Search result #2 - Choose "Alignment"

2 Alignment Identity: 100%



Matches: 149  
Mismatches: 0

[View Less ▾](#)

- Alignment
- Subject
- References
- References

Alignment Data  
 BLAST Score: 759  
 E-Value: 1.12043e-93

```

Q      1  MADQLTEEQI AEFKEAFSLF DKDGDGTITT KELGTVMRSL GQNPTEAELQ DMINEVDADG NGTIDFPPEFL 70
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
S      15  MADQLTEEQI AEFKEAFSLF DKDGDGTITT KELGTVMRSL GQNPTEAELQ DMINEVDADG NGTIDFPPEFL 84
Q      71  TMMARKMKDT DSEEEIREAF RVFDKDGNGY ISAAELRHVM TNLGEKLTDE EVDEMIREAD IDGDGQVNYE 140
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
S      85  TMMARKMKDT DSEEEIREAF RVFDKDGNGY ISAAELRHVM TNLGEKLTDE EVDEMIREAD IDGDGQVNYE 154
Q     141  EFVQMMTAK 149
      | | | | | | | |
S     155  EFVQMMTAK 163
    
```

## Choose "Subject"

2 Alignment Identity: 100%

Query 1 149

Subject 1 174

Matches: 149  
Mismatches: 0

View Less ▾

Alignment Subject References References

CAS Registry Number: [1145807-38-0](#)  
Length: 174 aa

Sequence

```
1  MGSSHHHHH HSSGMADQLT EEQIAEFKEA FSLFDKGDG TITTKELGTV MRSLGQNPT E AELQDMINEV DADGNGTIDF
81  PEFLTMMARK MKDITDSEEEI REAFRVFDK GNGYISAAEL RHVMTNLGEK LTDEEVDEMI READIDGDGQ VNYEEFVQMM
161 TAKSSGLVPR GSHM
```

## Choose "References"

2 Alignment Identity: 100%

Query 1 149

Subject 1 174

Matches: 149  
Mismatches: 0

View Less ▾

Alignment Subject References References

**A fusion tag comprising an affinity tag and an EF-hand motif containing polypeptide and methods of use thereof**  
Assignees: UNIVERSITY HEALTH NETWORK; MCCLUSKEY, ANDREW, J.; POON, GREGORY, M. K.; GARIÉPY, JEAN  
WO2009046520 A1 | Seq ID No: -

### III. Biosequence -Influenza Virus using the following Protein Sequence (H1N1-2009)

<https://www.ncbi.nlm.nih.gov/nucore/CY065935>

(Influenza A virus (A/Guam/NHRC0001/2009(H1N1)) segment 1, complete sequence

Note – (Copy the sequence below completely...the Sequence Search will disregard blank spaces)

MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPA

LRMKWMMAMRYPITADKRIMDMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWRN

GPTTSTVHYPKVYKTYFEKVERLKHGTFGPVHFRNQVKIRRRVDTNPGHADLSAKEAQ

DVIMEVFPNEVGARILTSESQLAITKEKKEELQDCKIAPLMVAYMLERELVRKTRFL

PVAGGTGSVYIEVLHLTQGTCWEQMYTPGGEVRNDDVDQSLIAARNIVRRAAVSADP

LASLLEMCHSTQIGGVRMVDILRQNPTEEQAVDICKAAIGLRISSEFSFGGFTFKRTS

GSSVKKEEEVLTGNLQTLKIRVHEGYEFTMVGRRATAILRKATRRLIQLIVSGRDEQ

SIAEAIIVAMVFSQEDCMIKAVRGDLNFVNRRANQRLNPMHQLLRHFQKDAKVLVFNWG

IESIDNVMGMIGILPDMTPSTEMSLRGIRISKMGVDEYSSTERVVVSIDRFLRVRDQR

GNVLLSPEEVSETQGTEKLTITYSSMMWEINGPESVLVNTYQWIIRNWEIVKIQWSQ

DPTMLYNKMEFEPFQSLVPKATRSRYSGFVRTLFQQMRDVLGTFDTVQIIKLLPFAAA

PPEQSRMQFSSLTVNVRGSGLRILVRGNSPVFNYNKATKRLTVLGKDAGALTEPDEG

TSGVESAVLRGFLILGKEDKRYGPALSINELSNLAKGEKANVLIGQGDVVLVMKRKRDR

SSILTDSQTATKRIRMAIN

Select "Protein" as the Sequence Type and "Proteins" as the Search Type.

## Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

## Biosequences

Enter a protein or nucleotide string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST CDR Motif

```
MERIKELRDLSMSQSRTRREILTKTTVDHMAIIKKYTSGRQEKPNALRMKMMAMRYPITADKRIMDMI PERNEQGQTL  
WSKINDAGSDRVMVSPPLAVTWNNRNGPTTSTVHYPKVYKTYFEKVERLKHGTFGPFVFRNQVKIRRRVDTNPGHADL  
SAKEAQDVIMEVVPNEVGARILTSESQLAITKEKKEELQDCKIAPLMVAYMLERELVRRTRFLPVAGGTGSVYIEV  
LHLTQGTQWEQMYTFGGEVRNDDVDQSLIIAARNIVRRAAVSADPLASLLEMCHSTQIGGVRMVDILRQNPTEEQAV  
DICKAAIGLRISSSFSFGGTFKRTSGSSVKKKEEVLGTGNLQTLKIRVHEGYEEFTMVGRRATAILRKRRLIQLI  
VSGRDEQSI AEAIIVAMVFSQEDDCMIKAVRGDLPVNRANQRLNPMHQLLRHFQKDAKLVLFQNWGIESIDNVGMIG  
ILPDMTSTEMSLRGIRISRMGVDEYSSTERVVVSIDRFLRVRDQQRGNVLLSPPEEVSSTQGTTEKLITITYSSMMWEI  
NGPESVLVNTYQWIRNWEIVKIQSDPTMLYKMEPEFPQSLVPKATRSRYSGFVRTLFQQRDVLGTFDTVYQIIK  
LLPFAAAPPEQSRMQPSSLTNNVRSGLRILVRGNSPVFNYNKATRLTLVLGK DAGALTEDEPDEGTSGVESAVLRGP  
LILGKEDKRYGPALSINELSNLAKGEKANVLIGQGDVVLVMKRRRDSIILTDSQTATKRIRMAIN
```

[Advanced Biosequence Search](#)

Sequence Type:  
 Nucleotide  **Protein**

Search Within:  
 Nucleotides  **Proteins**

Include NCBI Sequences

Limit Total Sequence Results to:  
1000

## Start the Biosequence search

January 10, 2022

2:51 PM

**Biosequences**

Sequence Type: **Protein**  
Search Within: **Proteins**  
NCBI Included: **Yes**  
BLAST Algorithm: **BLASTp**  
Alignment Identity: -  
Query Coverage: **90%**

[View All](#)

```
MERIKELRDLSMSQSRTRREILTKTTVDHMAIIKKYTSGRQEKPNALRMKMMAM  
RYPITADKRIMDMI PERNEQGQTLWSKINDAGSDRVMVSPPLAVTWNNRNGPTT  
STVHYPKVYKTYFEKVERLKHGTFGPFVFRNQVKIRRRVDTNPGHADLSAKEA  
QDVIMEVVPNEVGARILTSESQLAITKEKKEELQDCKIAPLMVAYMLERELV  
RKTRFLPVAGGTGSVYIEVLHLLTQGTQWEQMYTFGGEVRNDDVDQSLIIAAR...
```

+

The Biosequence Search is completed. Click on the "View results" to view the results.

January 10, 2022

2:42 PM

**Biosequences**

Sequence Type: **Protein**  
Search Within: **Proteins**  
NCBI Included: **Yes**  
BLAST Algorithm: **BLASTp**  
Alignment Identity: -  
Query Coverage: **90%**

[View All](#)

```
MERIKELRDLSMSQSRTRREILTKTTVDHMAIIKKYTSGRQEKPNALRMKMMAM  
RYPITADKRIMDMI PERNEQGQTLWSKINDAGSDRVMVSPPLAVTWNNRNGPTT  
STVHYPKVYKTYFEKVERLKHGTFGPFVFRNQVKIRRRVDTNPGHADLSAKEA  
QDVIMEVVPNEVGARILTSESQLAITKEKKEELQDCKIAPLMVAYMLERELV  
RKTRFLPVAGGTGSVYIEVLHLLTQGTQWEQMYTFGGEVRNDDVDQSLIIAAR...
```

Results returned from Search – zero Mismatches

BLAST Search Details

Sequence Type: Protein  
 Search Within: Proteins  
 BLAST Algorithm: BLASTp  
 NCBI Included: Yes  
 Alignment Identity: -  
 Query Coverage: 90%  
 E-Value: 10  
 Match with Gaps?: No  
 Gap Costs: Existence 11  
 Extension 1  
 Word Size: 3

Bioscape Analysis

Visually explore sequence similarity with a new tool.  
 Learn more about Bioscape.

Create Bioscape Analysis

Filter by

E-Value

0 to 10<sup>6</sup>

Query Coverage %

0 to 100

## Biosequences (1,000)

Sort: Alignment Identity - View: Expanded -

References

Query Details MERIKELRDLMSQSRTRREILTKITVDHMAIHKYTSGRQKQKFPALRMGMGMRYFITADKRMIMIPERNEQGGTLNSKINDAGS... View More

Alignment Identity: 100%

Query 1 759

Subject 1 759

Matches: 759  
 Mismatches: 0

View Less

Alignment Subject References

Alignment Data

BLAST Score: 3857  
 E-Value: 0

Q	1	MERIKELRDLMSQSRTRREILTKITVDHMAIHKYTSGRQKQKFPALRMGMGMRYFITADKRMIMIPERNEQGGTLNSKINDAGS...	70
S	1	MEIKELRDLMSQSRTRREILTKITVDHMAIHKYTSGRQKQKFPALRMGMGMRYFITADKRMIMIPERNEQGGTLNSKINDAGS...	70
Q	71	NEQQQLNSKINDAGSDRUMVSPPLAVTWNKNGPTTSTVHYPKVTYTFEKVRLKSGTGFVHFRANQVK...	140
S	71	NEQQQLNSKINDAGSDRUMVSPPLAVTWNKNGPTTSTVHYPKVTYTFEKVRLKSGTGFVHFRANQVK...	140
Q	141	IRRVVDNFGHADLSAKEAQDVIMEVFFNEVGARILTSEQLAITREKKEELQDCKIAPIMVAYMLERE...	210
S	141	IRRVVDNFGHADLSAKEAQDVIMEVFFNEVGARILTSEQLAITREKKEELQDCKIAPIMVAYMLERE...	210

View the 2<sup>nd</sup> Search returned with 1 Mismatch

There is one amino acid (red line in the graphical representation) different between the sequence search and the #2 Query Result returned. There is one mismatch.

The graphic shows the location where the mismatch occurs.

The Alignment Data shows the exact point in the sequence where the mismatch occurs between the two amino acids.

Alignment Identity: 99.87%

Matches: 758  
 Mismatches: 1

Query 1 759

Subject 1 759

View Less

Alignment Subject References




Alignment Data

BLAST Score: 3856  
 E-Value: 0

Q	351	TLKIRVHEGYEFTMVGRRA TAILRKATER LIQLIVSGRD EQSIAEAIIVAMVPSQEDCM IKAVRGDLNF	420
S	351	TLKIRVHEGYEFTMVGRRA TAILRKATER LIQLIVSGRD EQSIAEAIIVAMVPSQEDCM IKAVRGDLNF	420
Q	421	VNRAHQRLNPMHQLLRHFQKDAKVLQNWGIESIDNVGMIGILPDMTPSTEMSAGIRISPMVDEYSS	490
S	421	VNRAHQRLNPMHQLLRHFQKDAKVLQNWGIESIDNVGMIGILPDMTPSTEMSAGIRISPMVDEYSS	490
Q	491	TERVVVSIDRFLRVRDQRGNVLLSPEEVSE TQSTEKLTIIT YSSSMMEIN GPESVLVNTYQNIIRNWEIV	560
S	491	TERVVVSIDRFLRVRDQRGNVLLSPEEVSE TQSTEKLTIIT YSSSMMEIN GPESVLVNTYQNIIRNWEIV	560

The “Subject” Tab shows cross references to other similar proteins

### Substance Detail

Reference (1) Reactions (0) Suppliers (0)   


CAS Registry Number  
**1639314-78-5**

Image Not Available

#### Unspecified

RNA-dependent RNA polymerase (Influenza A virus strain A/California/07/09 gene PB2) (ACI)

Protein/Peptide Sequence  
Sequence Length: 759

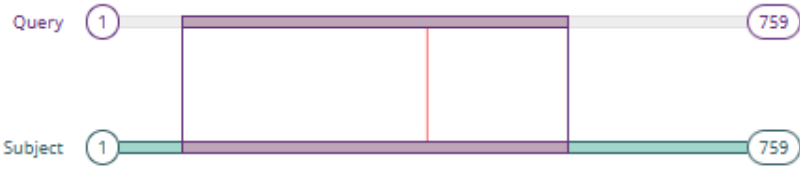
 Related Sequences (20)

[Expand All](#) | [Collapse All](#)

- Other Names and Identifiers
- Sequence Details
- Additional Details


Choosing the “Reference” Tab brings up the documentation where the sequence is found

2 Alignment Identity: 99.87%



Matches: 758  
Mismatches: 1

View Less ▾

Alignment Subject **References** 

**流感病毒重配**  
Assignees: 诺华股份有限公司; 合成基因组疫苗股份有限公司  
CN105722976 A | Seq ID No: -

**Influenza virus reassortment**  
Assignees: SEQIRUS UK LIMITED; SYNTHETIC GENOMICS, INC.  
US20170326227 A1 | Seq ID No: 17

**リアソータントインフルエンザAウイルス**  
Assignees: ノバルティス アーゲー; シンセティック ジェノミクス ヴァクシンス. インコーポレーテッド  
JP2016501020 A | Seq ID No: 17

Repeat the same search for the Nucleotides you have previously, but this time remove part of the sequence

Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

### Biosequences

Enter a protein or nucleotide string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST CDR Motif Upload Sequence Clear Search

```
GATTGGGGTTCAAAGCAGTATCGATCAAATAGTAATCCATTGTTCAACTCACAGTTT
```

Sequence Type:  
 Nucleotide  Protein

Search Within:  
 Nucleotides  Proteins  
 Include NCBI Sequences

Limit Total Sequence Results to:  
1000

**Start Biosequence Search**

Advanced Biosequence Search ^ Adjust Parameters for Short Sequences | Reset All

Alignment Identity %  Match with Gaps?  Gap Costs

80  Yes  No Existence 5 Extension 2

Now Rerun the search

Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

### Biosequences

Enter a protein or nucleotide string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST CDR Motif Upload Sequence Clear Search

```
GATTGGGGTTCAAAGCAGTATCGATCAAATA|
```

Sequence Type:  
 Nucleotide  Protein

Search Within:  
 Nucleotides  Proteins  
 Include NCBI Sequences

Limit Total Sequence Results to:  
1000

**Start Biosequence Search**

Advanced Biosequence Search ^ Adjust Parameters for Short Sequences | Reset All

Alignment Identity %  Match with Gaps?  Gap Costs

80  Yes  No Existence 5 Extension 2

The Updated search is found. Now the number of mismatches is much more pronounced and is shown in the highlighted red vertical lines shown in the graphics for the 2<sup>nd</sup> search returned.

**BLAST Search Details**

Sequence Type: Nucleotide  
 Search Within: Nucleotides  
 BLAST Algorithm: BLASTn  
 NCBI Included: Yes  
 Alignment Identity: 80%  
 Query Coverage: 90%  
 E-Value: 10  
 Match with Gaps?: No  
 Gap Costs: Existence 5  
 Extension 2  
 Word Size: 11

**Bioscape Analysis**

Visually explore sequence similarity with a new tool. Learn more about Bioscape.

[Create Bioscape Analysis](#)

Filter by

^ E-Value

0 to 10<sup>6</sup>

^ Query Coverage %

**Biosequences** (11) Sort: Alignment Identity View: Expanded

References

Query Details: GATTTGGGGTTCAAAGCAGTATCGATCAAATA [View More](#)

---

**1** Alignment Identity: 100%

Query 1 32

Subject 1 60

Matches: 32  
Mismatches: 0

[View Less](#)

Alignment Subject References

Alignment Data  
 BLAST Score: 64  
 E-Value: 5.68136e-8

```

Q      1  GATTTGGGGT TCAAAGCAGT ATCGATCAA TA 32
      | | | | | | | | | | | | | | | | | |
S      1  GATTTGGGGT TCAAAGCAGT ATCGATCAA TA 32
      
```

---

**2** Alignment Identity: 87.1%

Query 1 32

Subject 1 1,368

Matches: 27  
Mismatches: 4

[View Less](#)

Alignment Subject References

Alignment Data  
 BLAST Score: 42  
 E-Value: 0.0384269

```

Q      2  ATTTGGGGTT CAAAGCAGTA TCGATCAAAT A 32
      | | | | | | | | | | | | | | | |
S     422  ATTTGGGGTC TGAAGCAGTG TCGATCAAAT A 452
      
```

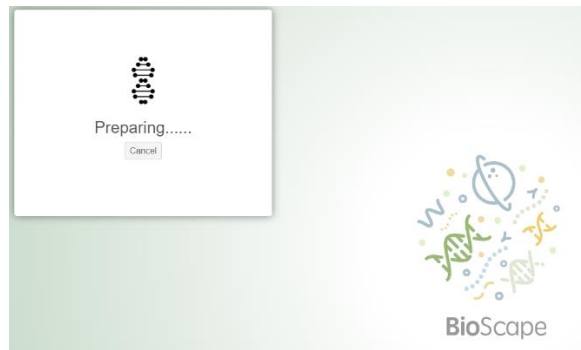
The #3 search returned show more mismatched (red vertical lines) nucleotides than does search #2.





Now Plot the Analysis of the Sequences by choosing “Create Bioscape Analysis”

The database searches for content



Results returned and plotted

Scroll with your mouse to Zoom in or Out of the Plotted variable screen



Click on one of the Silos to see the results returned – 3 Patents are returned

The screenshot shows a search interface with a 'Sequence Similarity' slider ranging from 80% to 100%. Below the slider is a small bar chart with colored bars. To the right, a window titled '321092-37-9' displays search results for 'Relevant Patents' (3) and 'Sequence Length' (937). The 'Domain Sequence' is shown as a long string of nucleotide bases (A, T, C, G).

Zoomed in View of the Patents returned

This is a zoomed-in view of the patent search results for patent number 321092-37-9. It shows the 'Relevant Patents' (3) and 'Sequence Length' (937). The 'Domain Sequence' is displayed as a long string of nucleotide bases (A, T, C, G).

Click on the Patent Number(s) found (in this case #3)

## References (1)

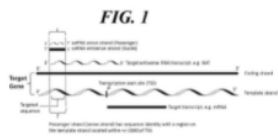
View: Partial Abstract ▾

[Substances](#) [Reactions](#) [Citing](#) [Download](#) [Email](#) [Save](#)

1

### Use of short activating RNA for modulating target gene expression for diagnostic, prognostic and therapeutic applications

By: Saetrom, Paal; Stovner, Endre Bakken  
World Intellectual Property Organization, WO2016170348 A2 2016-10-27 | Language: English, Database: CAplus



**FIG. 1**  
The invention relates to oligonucleotides, e.g., saRNAs useful in upregulating the expression of a target gene and therapeutic compositions comprising such oligonucleotides. Methods of using the oligonucleotides and the therapeutic compositions are also provided.

[PatentPak](#) [Full Text](#) [Substances \(0\)](#) [Reactions \(0\)](#) [Citing \(3\)](#) [Citation Map](#)

Choose PatentPak to View the Patents.

## References (1)

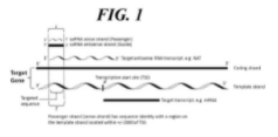
View: Partial Abstract ▾

[Substances](#) [Reactions](#) [Citing](#) [Download](#) [Email](#) [Save](#)

1

### Use of short activating RNA for modulating target gene expression for diagnostic, prognostic and therapeutic applications

By: Saetrom, Paal; Stovner, Endre Bakken  
World Intellectual Property Organization, WO2016170348 A2 2016-10-27 | Language: English, Database: CAplus



**FIG. 1**  
The invention relates to oligonucleotides, e.g., saRNAs useful in upregulating the expression of a target gene and therapeutic compositions comprising such oligonucleotides. Methods of using the oligonucleotides and the therapeutic compositions are also provided.

[PatentPak](#) [Full Text](#) [Substances \(0\)](#) [Reactions \(0\)](#) [Citing \(3\)](#) [Citation Map](#)

Patent	Language	Kind Code	PatentPak Options
<a href="#">WO2016170348</a>	English	A2	<a href="#">PDF</a>
WO2016170348	English	A3	<a href="#">PDF</a>
JP2018512876	Japanese	T	<a href="#">PDF</a>
US20180305689	English	A1	<a href="#">PDF</a>
JP2021035360	Japanese	A	<a href="#">PDF</a>
US20210363525	English	A1	<a href="#">PDF</a>

Open any PatentPak PDF and Save it to your PC. Open the Saved PDF. Now Search with your original query.

GATTTGGGGTTCAAAGCAGTATCGATCAAATA

Choose View Full Text for the PDF document via EspaceNet



☆ **WO2016170348A2** SARNA COMPOSITIONS AND METHODS OF USE

[Bibliographic data](#) [Description](#) [Claims](#) [Drawings](#) [Original document](#) [Citations](#) [Legal events](#) [Patent family](#)

Register ↗ **Global Dossier** ↗

**Applicants** MINA THERAPEUTICS LTD [GB] +  
**Inventors** SÆTROM PÅL [NO]; STOVNER ENDRE BAKKEN [NO] +

**Classifications**

**CPC** A61P1/00 (EP); A61P11/00 (EP); A61P13/00 (EP); A61P17/00 (EP); A61P25/00 (EP); A61P25/02 (EP); A61P27/02 (EP); A61P3/00 (EP); A61P35/00 (EP); A61P37/02 (EP); A61P37/06 (EP); A61P43/00 (EP); A61P5/00 (EP); A61P9/00 (EP); C12N15/113 (EP,US); C12N2310/11 (US); C12N2310/14 (EP,US); C12N2310/20 (EP,US); C12N2310/34 (US);

**Priorities** US201562150892P 2015-04-22; US201562150893P 2015-04-22; US201562150895P 2015-04-22; US201562150897P 2015-04-22; US201562150900P 2015-04-22; US201562150904P 2015-04-22; US201562150908P 2015-04-22

**Application** GB2016051116W 2016-04-21

**Publication** [WO2016170348A2](#) 2016-10-27

**Published as** EP3286318A2; JP2018512876A; JP2021035360A; US2018305689A1; US2021363525A1; [WO2016170348A2](#); WO2016170348A3; WO2016170348A8

EN  FR

SARNA COMPOSITIONS AND METHODS OF USE

Abstract

## IV. Commonly Determined Region Search

Commonly Determined Regions (CDR's) are used to find the complimentary regions found in antibodies (also monoclonal antibodies – aka mAb's). CDR's, Antibodies and mAB's bind to antibodies antigens. An antigens is a substance introduced within the host body such as a chemical, pollen, virus or bacteria. Monoclonal antibodies are made synthetically by pharmaceutical and biotech manufacturers and antibodies are found and formed naturally within organism's bodies.

Monoclonal antibodies help fight off infection the body cannot compensate and resolve itself such as during Chemotherapy and Radiation treatments.

Any of the following example CDR's can be inputted into the SciFinderN tool within the CDR field and will find antibodies based on the antigen search needed.

Below are the examples with can be inputted into the CDR tabs.

WASTRES  
LVSNLES

DASNRAT  
SYTFH

PNPLKAM  
GYMH

For this Examples, input LVSNLES into the CDR input. It does not matter if the input is capitalized.

Searching for...  
All  
Substances  
Reactions  
References  
Suppliers  
Biosequences  
Retrosynthesis

Biosequences  
Enter a protein string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST CDR Motif Upload Sequence Clear Search

CDR1 LVSNLES  
CDR2 Enter CDR2...  
CDR3 Enter CDR3...

Include NCBI Sequences  
Limit Total Sequence Results to: 1000  
Start Biosequence Search

The system starts its search.

March 21, 2022

Biosequences  
9:29 AM

NCBI Included: No  
Results will expire on  
Apr 20, 2022.

CDR1: LVSNLES  
CDR2: -  
CDR3: -

View Results  
Edit Options  
Searching

View the results – The CDR found a perfect match against the database of possibilities. Over 1000 CDR matches were found.

CDR Segments

Select a segment below to view individual or intersecting CDR results:

CDR1 1,000

CDR2 0

CDR3 0

Apply

Reset Segments

Bioscape Analysis

Visually explore sequence similarity with a new tool. [Learn more about Bioscape.](#)

Create Bioscape Analysis

Filter by

E-Value

0 to 10<sup>6</sup>

Query Coverage %

Biosequences (1,000)

Sort: Alignment Identity View: Expanded

References

Query Details View Less

CDR1 LVSNLES

CDR2

CDR3

1 Alignment Identity: 100%

Subject 1 112

CDR1

Matches: 7  
Mismatches: 0

View Less

Alignment Subject References

Alignment Data

BLAST Score: 49  
E-Value: 71.5995

CDR1	1	LVSNLES	7
S	54	LVSNLES	60

Let's rerun the search with 2 CDR's inputted. Leave LVSNLES in the search and add WASTRES as the 2<sup>nd</sup> CDR search. Start the Biosequence.

Searching for...

All

Substances

Reactions

References

Suppliers

Biosequences

Retrosynthesis

Biosequences

Enter a protein string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST CDR Motif

Upload Sequence Clear Search

CDR1 LVSNLES

CDR2 WASTRES

CDR3 Enter CDR3...

Include NCBI Sequences

Limit Total Sequence Results to: 1000

Start Biosequence Search

System Search Begins

March 21, 2022

📌 Biosequences
NCBI Included: No
CDR1: LVSNLES
CDR2: WASTRES
CDR3: -
View Results

9:41 AM
Results will expire on Apr 20, 2022.
Edit Options

Searching.

**View the results. Now there is a cross reference with multiple CDR regions within the database. Only the WASTRES CDR is found in any of the searches run. Both CDR's are not found within the same sequences returned.**

CDR Segments

Select a segment below to view individual or intersecting CDR results.

Apply  
Reset Segments

Bioscape Analysis

Visually explore sequence similarity with a new tool. [Learn more about Bioscape.](#)

Create Bioscape Analysis

Filter by

E-Value: 0 to 10<sup>6</sup>

Query Coverage %: 0 to 100

Subject Coverage %: 0 to 100

Alignment Identity %: 0 to 100

**Biosequences** (1,000) Sort: Alignment Identity View: Expanded

References

Query Details [View Less](#)

- > CDR1 LVSNLES
- > CDR2 WASTRES
- > CDR3

1 Alignment Identity: 50%

Matches: 7  
Mismatches: 7

View Less

Alignment Subject References

Alignment Data  
BLAST Score: 27  
E-Value: 3.387

CDR2	1	WASTRES	7
S	56	WASTRES	62

2 Alignment Identity: 50%

Matches: 7  
Mismatches: 7

View Less

**Try a different CDR to search with LVSNLES. In this case use PNPLKAM.**

## Recent Search History

[View All Search History](#)

March 21, 2022

Biosequences 10:07 AM	NCBI Included: No Results will expire on Apr 20, 2022.	CDR1: LVSNLES CDR2: PNPLKAM CDR3: -	<a href="#">View Results</a> <a href="#">Edit Options</a> <i>Searching.</i>
--------------------------	--	---	---

**There were no searches return with both CDR's in the returned information**

**Rerun the search but use NCBI sequences**

Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

### Biosequences

Enter a protein string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST | **CDR** | Motif | Upload Sequence | Clear Search

Include NCBI Sequences

Limit Total Sequence Results to: 1000

[Start Biosequence Search](#)

CDR1	LVSNLES	<input type="text"/>
CDR2	PNPLKAM	<input type="text"/>
CDR3	Enter CDR3...	<input type="text"/>

**Both CDR's were found, but no sequences where both CDR's are found in the same sequence.**



CDR Segments

Select a segment below to view individual or intersecting CDR results.

Apply  
Reset Segments

Bioscape Analysis

Visually explore sequence similarity with a new tool.  
[Learn more about Bioscape.](#)

Create Bioscape Analysis

Filter by

E-Value  
0 to 10<sup>6</sup>

Query Coverage %  
0 to 100

Subject Coverage %  
0 to 100

Alignment Identity %  
0 to 100

## Biosequences (1,000)

Sort: Alignment Identity View: Expanded

References

Query Details View Less

> CDR1  
LVSNLES

> CDR2  
PNPLKAM

> CDR3

1 Alignment Identity: 50%

Subject 1 118

CDR2 1 7

Matches: 7  
Mismatches: 7

View Less

Alignment Subject

Alignment Data  
BLAST Score: 27.5  
E-Value: 3.14668

CDR2	1	PNPLKAM	7
S	99	PNPLKAM	105

2 Alignment Identity: 50%

Subject 1 112

CDR1 1 7

Matches: 7  
Mismatches: 7

Try another CDR to compare against WATRES. Use GYYMH as the 2<sup>nd</sup> CDR input. Check the “Include NCBI Sequences” box.

Searching for...

- All
- Substances
- Reactions
- References
- Suppliers
- Biosequences**
- Retrosynthesis

## Biosequences

Enter a protein string, or upload a .txt or .fasta file. [Learn more about Biosequence Search.](#)

BLAST CDR Motif Upload Sequence Clear Search

CDR1 WASTRES

CDR2 GYYMH

CDR3 Enter CDR3...

Include NCBI Sequences

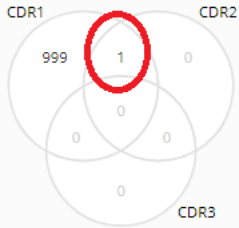
Limit Total Sequence Results to:  
1000

Start Biosequence Search

Both CDR's are found within the search and data displayed as found.

CDR Segments

Select a segment below to view individual or intersecting CDR results.



Apply  
Reset Segments

Bioscape Analysis

Visually explore sequence similarity with a new tool.  
Learn more about Bioscape.

Create Bioscape Analysis

Filter by

E-Value  
 0 to 10<sup>6</sup>

Query Coverage %  
 0 to 100

Subject Coverage %  
 0 to 100

References



Query Details View Less

- > CDR1  
WASTRES
- > CDR2  
GYMH
- > CDR3

1 Alignment Identity: 100%



Matches: 12  
Mismatches: 0

View Less

Alignment	Subject	References
Alignment Data BLAST Score: 50 E-Value: 48.8762		
CDR1	1	WASTRES 7
S	56	WASTRES 62
CDR2	1	GYMH 5
S	415	GYMH 419

## V. MOTIF Search

MOTIF is a protein genetic pattern that has chemical or biological activity/function. MOTIFS are typically 10 to 30 characters long but can be much larger. Zinc finger proteins are very common MOTIFS and can 1000 to 2000 characters in length. This MOTIF is a protein and is 1133 amino acids long.

Here is an example of a Homo Sapien Ring Finger to try.

```
MPVQAPQWTFDLSPICTQTFDETIKPIISLGC GHTVCKMCLNK
LHRKACPFDQTTINTDIELLPVNSALLQLVGAQVPEQQPITLCSGVEDTKHYEEAKKC
VEELALYLKPLSSARGVGLNSTTQSVLSRPMQRKLVTLVHCQLVEEEGRIRAMRAARS
LGERTVTELILQHQPQLSSNLWAAVRARGCQFLGPAMQEEALKLVLLALEDGSALS
RKVLVLFVVQRLEPRFPQASKTSIGHVVQLLYRASC FKVTKRDESSLMQLKEEFRTY
EALRREHDSQIVQIAMEAGLRIAPDQWSSLLYGDQSHKSHMQSIIDKLQTPASFAQSV
QELTIALQRTGDPANLNRLRPHLELLANIDPSPDAPPPTWEQLENGLVAVRTVVHGLV
DYIQNHSSKKGADQQPPQHSKYKTYMCRDMKQRGGCPRGASCTFAHSQEELEKFRKMN
KRLVPRRPLSASLGQLNEVGLPSAAILPDEGAVDLP SRKPPALPNGIVSTGNTVTQLI
PRGTDPSYDSSLKPGKIDHLSSAPGSPDLESVPKSI SALPVNPHSIPPRGPADLP
PMPVTKPLQMVPRGSQLYPAQQT D V Y Y Q D P R G A A P P F E P A P Y Q Q G M Y T P P P Q C V S R F
VRPPSAPEPAPPYLDHYPPYLQERVVNSQYGTQPQQY PPIYPSHYDGRRVYPAPSYT
REEIFRESPIEIPPAAVPSYVPESRERYQQIESYYPVAPHPTQIRPSYLREPPYSR
LPPPPQPHPSLDELHRRRKEIMAQLEERKVISPPPFAPSPTLPPTFHPEEFLDEDLKV
AGKYKNDYSQYSPWSCDTIGSYIGTKDAKPKDVVAAGSVEMMNVESKGMRDQRDLQ
RRAAETSDDDLIPFGDRPTVSRFGAISRTSKTIYQGAGPMQAMAPQGAPTKSINISDY
SPYGTGGWGASPYSPHQNIPSQGHFSE RERISMSEVASHGKPLPSAEREQLRLELQQ
LNHQISQQTQLRGL EAVSNRVLVQREANTLAGQS QPPPPPPKWPGMISSEQLSLELH
QVEREIGKRTRELSMENQCSDLMKSKLNTSKQAENGQPEPQNKVPAEDLTLTFS DVPNGSALTQENISLLSNKTS LNL
EDPEGGDNDNSQRSGVTPSSAP
```

The screenshot displays the Biosequence Search interface. On the left, a sidebar titled "Searching for..." lists various search categories: All, Substances, Reactions, References, Suppliers, Biosequences (highlighted), and Retrosynthesis. The main area is titled "Biosequences" and contains a search input field with the text "Enter a protein or nucleotide string. [Learn more about Biosequence Search.](#)". Below the input field are three tabs: "BLAST", "CDR", and "Motif" (selected). A "Clear Search" button is located to the right of the tabs. The search results are displayed in a text area, showing a long sequence of amino acids. On the right side of the interface, there are several options: "Sequence Type:" with "Nucleotide" and "Protein" buttons (Protein is selected); "Include NCBI Sequences" with a checked checkbox; "Limit Total Sequence Results to:" with a dropdown menu set to "1000"; and a "Start Biosequence Search" button.

72 Searches are returned with different degrees of Matching.

4 Searches are returned with 100% Matching including #1. The search also returns the Organism each MOTIF is found in. This shows similarities between different species for this protein.

**Motif Search Details**  
 Sequence Type: Protein  
 NCBI Included: Yes  
 Query Coverage: 90%  
 E-Value: 10

**Bioscape Analysis**  
 Visually explore sequence similarity with a new tool.  
[Learn more about Bioscape.](#)  
 Create Bioscape Analysis

**Filter by**  
 E-Value: 0 to 10<sup>6</sup>  
 Query Coverage %: 0 to 100  
 Subject Coverage %: 0 to 100  
 Alignment Identity %: 0 to 100

**Organisms**  
 Homo sapiens (3)  
 Macaca fascicularis (3)  
 Papio anubis (3)

**Biosequences (72)** Sort: Alignment Identity View: Expanded

**Query Details** > Seq 1: 1 MPVQAPQNTDFLSCPICQTQFDETIRKPISLGCGHVCKMCLNKLHRKACPFQDQTINTDIELLPVNSALLQL... [View More](#)

**1** Alignment Identity: 100%  
 Query 1 (1,133)  
 Subject 1 (1,133)  
 Matches: 1,133  
 Mismatches: 0

**Alignment Data**  
 BLAST Score: 8577  
 E-Value: 0

```

Q 1 MPVQAPQNTDFLSCPICQTQFDETIRKPISLGCGHVCKMCLNKLHRKACPFQDQTINTDIELLPVNSALLQL 70
S 1 MPVQAPQNTDFLSCPICQTQFDETIRKPISLGCGHVCKMCLNKLHRKACPFQDQTINTDIELLPVNSALLQL 70
Q 71 LQLVGAQVPEQQPITLCSGVEDTKHYEEAKKCVVEELALYLKPLSSARGVGLNSTTQSVLSRPMQRKLVTL 140
S 71 LQLVGAQVPEQQPITLCSGVEDTKHYEEAKKCVVEELALYLKPLSSARGVGLNSTTQSVLSRPMQRKLVTL 140
Q 141 VHCQLVEEEGRIRAMRAARSLGERTVTELI LQHQNPPQLSSNLWAAVRARGCQFLGPAMQEEALKLVLLA 210
S 141 VHCQLVEEEGRIRAMRAARSLGERTVTELI LQHQNPPQLSSNLWAAVRARGCQFLGPAMQEEALKLVLLA 210
  
```

**2** Alignment Identity: 100%  
 Query 1 (1,133)  
 Subject 1 (1,109)  
 Matches: 1,109  
 Mismatches: 0

System Search included the NCBI DB.

March 21, 2022

**Biosequences**  
 11:08 AM

Sequence Type: Protein  
 NCBI Included: Yes  
 Query Coverage: 90%  
 E-value: 10

Results will expire on  
 Apr 20, 2022.

MPVQAPQNTDFLSCPICQTQFDETIRKPISLGCGHVCKMCLNKLHRKACPFQDQTINTDIELLPVNSALLQLVGAQVPEQQPITLCSGVEDTKHYEEAKKCVVEELALYLKPLSSARGVGLNSTTQSVLSRPMQRKLVTLVHCQLVEEEGRIRAMRAARSLGERTVTELI LQHQNPPQLSSNLWAAVRARGCQFLGPAMQEEALKLVLLALEEDGSALSRKLVLFVQRLERPRFPQASKTISGHVVQLLYRASCFCVTKRDEDSLMLQKKEEFRTYEALRREHDSQIVQIAMEAGLRIAPDQWSSLLYGDQSHKSHMQSIIDKLTQSPASFAQSVQELTIALQRTGDPANLNLRPHLELLANIDPSPDAPPPTEQLENGLVAVRTVVHGLVDYIQNHSSKKGADQQQPQHSKYKTYMCRDMKQRGGCPRGASCTFAHSQEELEKFRKMNKRLVPRRPLSASLGQLNEVGLPSAAILPDEGAVDLPKRPPALPNGIVSTGNTVTQLIPRGTDPSYDSSLKPGKIDHLSAPGSPDLESVPSISALPVNPHSIPRGPADLPMPVTKPLQWVPRGSQLYPAQQTVDVYQDPRGAAPPFEPAPYQQGMYTTPPQCYSRFRPPSAPEPAPPYLDHYPPYLQERVVNSQYGTQP...

[View Results](#)  
[Edit Options](#)  
 Searching...

Further in the research results the matches decrease and mismatched increase. Here is result #72.

