

*** 目次 ***

核酸・ペプチド・タンパク質の登録例	1
核酸・ペプチド・タンパク質の検索手順	3
化学物質名称, CAS 登録番号 (CAS RN [®]), GenBank 番号	3
構造検索	4
BLAST 配列検索	5
(参考) Advanced Biosequence Search 画面 (BLAST 検索のパラメータ設定)	7
CDR 配列検索	8
Motif 配列検索	9
Bioscape	10
同主鎖の配列を持つ物質の検索	12

核酸・ペプチド・タンパク質の登録例

核酸・ペプチド・タンパク質は一つでも配列が異なれば別物質として収録される。

- 化学修飾された配列，側鎖の置換基の異なる配列，同位体で置換された配列等は，同じ配列でも別物質として収録される。
- GenBank から収録された配列は，1 GenBank 番号につき 1 物質として収録される。

■ 核酸（9 以上の塩基配列を持つ物質）のレコード例

関連情報のリンク

CAS RN®

化学物質名

配列

特徴表

**特許番号と配列の
記載位置，配列番号**

Substance Detail (1 of 1)

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

CAS Registry Number
2128329-99-5

Image Not Available

Unspecified
DNA (synthetic clone WO2017-140839 oligonucleotide) (ACI)

Nucleic Acid Sequence
Sequence Length: 115
34 a, 27 c, 27 g, 27 t
modified

Related Sequences (1)

Other Names and Identifiers
1 Other Name for this Substance
1: PN: WO2017140839 SEQID: 2 claimed DNA

Sequence Details

Sequence: DNA: linear

1	catgttcgat	gaggcaccgat	agatgtacgc	tttgacatac	gcttggacaa
51	tacttgagca	gtcgccagat	ataggatgtt	gcaagctccg	tgatgccac
101	aaaccaaaaa	cctcg	-	-	-

Sequence Modifications

Type	Location	Description
modified base	cytidine-1	5'-phosphate

Patent Annotations

Source: Not Given
Reference: WO2017140839, SEQID 2: claimed

配列長

核酸タイプと数

**同主鎖の配列を持つ物質の
検索ボタン**



大部分の核酸・ペプチド・タンパク質のレコードには，構造図が収録されていないが，水素以外の元素数が 252 以下でさらに原文献から構造が特定できる場合は収録される。構造図が収録されている場合は，構造検索が可能（配列長が 9 以上の核酸については，構造検索できない）。

【構造図が収録されているレコード例】

CAS RN
69718-12-3

CAS Name
Adenosine, 2'-deoxyadenylyl-
(3'-5')-2'-deoxy-5,6-[(1,2-
diphenyl-1,2-ethenediylid)...

Substance Detail

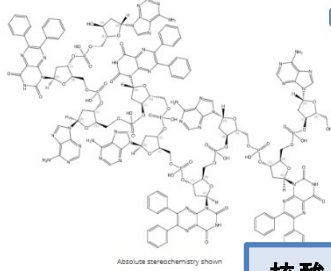
Reactions (2)

Synthesize (2)

Create Retrosynthesis Plan

References (2)

Suppliers (0)



Absolute stereochemistry shown

核酸

CAS RN
1334605-59-2

CAS Name
L-Lysine, N²-[4-bis(2-{4-amino-5,
6-dihydro-7-(methylamino)-2,5-
dioxypyrimido[4...

Substance Detail

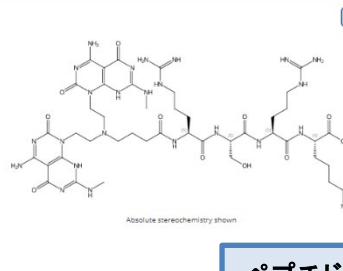
Reactions (15)

Synthesize (15)

Create Retrosynthesis Plan

References (3)

Suppliers (0)



Absolute stereochemistry shown

ペプチド

■ ペプチド・タンパク質（4 以上のアミノ酸残基を持つ物質）のレコード例

関連情報のリンク

CAS RN®

分子式

化学物質名

配列

特徴表

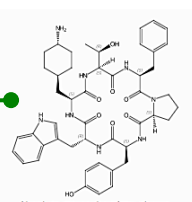
特許番号と配列の
記載位置，配列番号

Substance Detail (1 of 1)

References (7) Reactions (0) Suppliers (0)

CAS Registry Number
216259-64-2

構造図
(水素以外の元素数が 252 以下の場合，構造検索可能)



Absolute stereochemistry shown

分子式 C₄₇H₅₆N₈O₈

化学物質名 Cyclo[3-(trans-4-aminocyclohexyl)-L-alanyl-L-threonyl-L-phenylalanyl-L-prolyl-L-tyrosyl-D-tryptophyl] (9Cl, ACI)

配列長 Protein/Peptide Sequence
Sequence Length: 6 modified (modifications unspecified)

同主鎖の配列を持つ物質の検索ボタン Related Sequences (2)

Other Names and Identifiers

Canonical SMILES
O=C1NC(C(=O)N(C(=O)N(C(=O)N2CCCC2C(=O)NC1CC3=CC=C(O)C=C3)CC=4C=CC(=CC4)C(O)C)CC5CC(N)(CS)CC6=7C=CC=CC76

Isomeric SMILES
C(C=1C=2C(=NC1)=CC=CC2)[C@@H]3C(=O)N(C@@H)C[C@@H]4C[C@@H](N)CC4[C](=O)N(C@@H)[C@@H](C)O(C(=O)N(C@@H)CC5=CC=CC=C5)C(=O)N6[C@@H](C(=O)N(C@@H)C7=CC=C(O)C=C7)C(=O)N3)CCC6[H][H]

2 Other Names for this Substance
70: PN: US20020042374 PAGE: 10 claimed protein
74: PN: US6268342 SEQID: 80 claimed protein

配列 Sequence Details

Sequence: cyclic				
1	ATFPYV	-	-	-

特徴表 Sequence Modifications

Type	Location	Description
stereo	tryptophan-6	D

特許番号と配列の記載位置，配列番号 Patent Annotations

Source: Not Given
Reference: US6268342, SEQID 80: claimed

Source: Not Given
Reference: US20020042374, PAGE 10: claimed

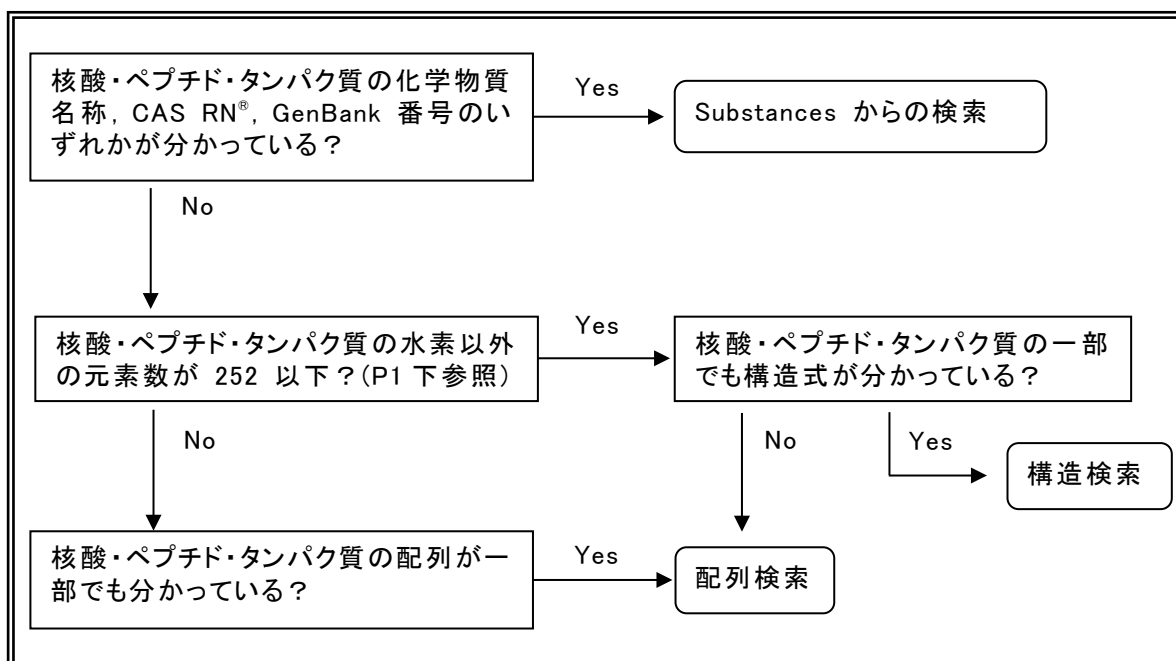


登録ルールの例外

- ・ 2002 年以降の特許・雑誌由来の配列は，修飾基も含めて同一の配列であっても同一レコードにならず特許・雑誌ごとに別レコードとして収録される場合がある。
- ・ GenBank 由来の配列と文献・特許由来の配列が同じ配列であっても同一レコードにならず，別レコードとして収録される場合がある。

核酸・ペプチド・タンパク質の検索手順

■ 核酸・ペプチド・タンパク質の検索方針



■ 化学物質名称, CAS RN®, GenBank 番号

Search

Search by Substance Name, CAS RN, Patent Number, etc.

化学物質名称

GABAA receptor GABRA1 (human brain)

Draw

Use Advanced Search for Molecular Formula, Substance Property, or Experimental Spectra

Search

Search by Substance Name, CAS RN, Patent Number, etc.

CAS 登録番号 (CAS RN)

1448554-62-8

Draw

Use Advanced Search for Molecular Formula, Substance Property, or Experimental Spectra

Search

Search by Substance Name, CAS RN, Patent Number, etc.

GenBank 番号

AAH30696

GenBank AAH30696 ← この番号を選択

GenBank AAH30696 (Translated from: GenBank BC030696)

GenBank AAH30196

Draw

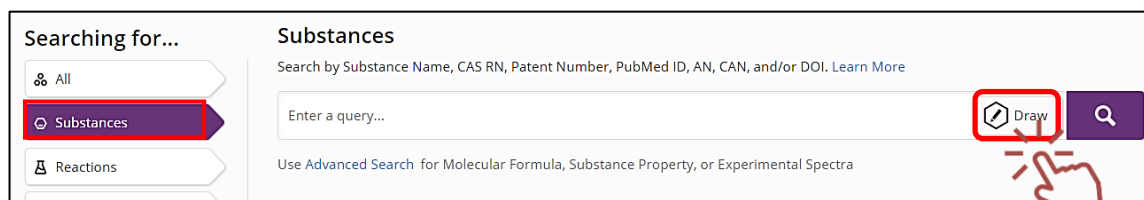


GenBank 番号の入力ルール

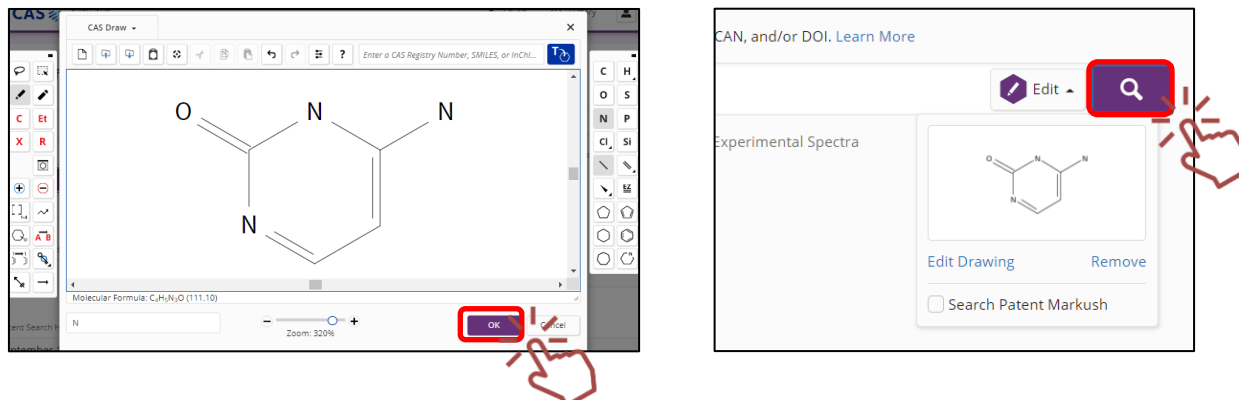
GenBank 番号から検索する場合は番号の前に「GenBank」を入力して検索します。検索ボックスに GenBank 番号を入力すると、オートサジェスト機能で「GenBank」の文字を含む候補が表示されるので、検索したい番号を選択してください。

■ 構造検索

- ① Substances から Draw をクリックし、構造を作図する。



- ② 作図が終了したら OK をクリックして、検索を実行する。



- ③ 回答が表示される。得られた回答が多かった場合は、Filter by (選択した項目に限定) から Substance Class (物質の種類) のフィルターを用いて核酸やペプチド・タンパク質に限定すると探しやすい。

■ BLAST 配列検索

Biosequences から配列質問式を使って、BLAST ホモロジー検索を実行することができます。

- ① Biosequences をクリックします。
- ② 配列質問式を入力するか Upload (.txt または .fasta) で呼び出します。
- ③ 配列質問式を選択します。
- ④ 回答の配列を選択します。
- ⑤ 回答の上限数を選択します。
- ⑥ (任意) パラメータを設定します。
- ⑦ 検索を実行します。

The screenshot shows the 'Biosequences' search page. On the left is a sidebar with navigation links: All, Substances, Reactions, References, Suppliers, and Biosequences (highlighted). The main area is titled 'Searching for...' and 'Biosequences'. It contains a search bar with the text 'gcggtttgctcttcttcttgcg' (annotated with a red box and '2'). Above the search bar are tabs for 'BLAST' (annotated with '1'), 'CDR', and 'Motif'. To the right of the search bar are buttons for 'Upload Sequence' and 'Clear Search'. Below the search bar is a section for 'Advanced Biosequence Search' (annotated with a red dashed box and '6'), which includes various parameters like 'Sequence Identity %' (80), 'Query Coverage %' (90), 'Word Size' (11), 'BLAST Algorithm' (BLASTn), 'E-Value' (10), 'Match with Gaps?' (No), 'Gap Costs' (Existence 5 Extension 2), 'Reward for Match, Penalty for Mismatch' (2, -3), and 'Exclude Low Complexity Regions' (No). To the right of the search bar are three more controls: 'Sequence Type' (Nucleotide selected, annotated with '3'), 'Search Within' (Nucleotides selected, annotated with '4'), and 'Limit Total Sequence Results to' (100, annotated with '5'). At the bottom right is a 'Start Biosequence Search' button (annotated with '7') with a hand icon pointing to it.

- ⑧ 配列検索には時間がかかることがあります。Search History で「Searching」と表示されている場合は検索実行中です。検索が終わると Searching の表示が Complete に変わります。回答を表示するには、View Results ボタンを押します。

The screenshot shows the 'Search History' page. The top bar includes the CAS SciFinder logo, a search bar, and icons for Draw, Search, Favorites, and User. The main area is titled 'Search History (375)'. It shows a list of search results. The first result is dated 'September 2, 2021' and has a time of '3:01 PM'. It is labeled 'Biosequences' and shows details: 'Sequence Type: Nucleotide', 'Search Within: Nucleotides', 'BLAST Algorithm: BLASTn', 'Alignment Identity: 80%', and 'Query Coverage: 90%'. To the right of these details are three buttons: 'View Results' (highlighted with a red box and a hand icon), 'Edit Search', and 'Complete' (highlighted with a green box).

⑨ View Results ボタンをクリックすると、配列検索の結果が表示されます。

検索条件

ヒットした回答配列の記載特許をまとめて検索

回答の並び替え

エクスポート

アライメントの概略図
(質問配列の 1~21 番目のコードが
回答配列の一部と一致している)

アライメントの類似性*

Subject タブ: 回答配列の詳細情報

References タブ: 出典文献情報

Alignment タブ: アライメント*・スコア値*・期待値*

絞り込み (フィルター)

E-Value (期待値) *
Query Coverage %*
Subject Coverage %*
Alignment Identity %*

記載文献 (特許もしくは
非特許) の検索

記載文献

* 用語説明

用語	内容
E-Value (期待値)	データベース中の配列に対してマッチする際の統計的有意性の閾値
Query Coverage %	類似領域 ÷ 配列質問式 × 100
Subject Coverage %	類似領域 ÷ 回答配列 × 100
Alignment Identity % (アライメントの類似性)	一致したコード ÷ 類似領域 × 100
Alignment (アライメント)	配列質問式と回答配列のどこが類似した領域か特定できるように並べたもの
BLAST Score (スコア値)	配列質問式と回答配列の対応するコードの一致度をスコアとして計算したもの

(参考) Advanced Biosequence Search 画面 (BLAST 検索のパラメータ設定)

The screenshot shows the 'Advanced Biosequence Search' interface. A red box labeled '1' highlights the 'Adjust Parameters for Short Sequences' link. Other numbered callouts point to specific input fields: '2' points to 'Sequence Identity %' (80), '3' to 'Query Coverage %' (90), '4' to 'BLAST Algorithm' (BLASTn), '5' to 'Match with Gaps?' (No), '6' to 'Word Size' (11), '7' to 'E-Value' (10), '8' to 'Gap Costs' (Existence 5 Extension 2), '9' to 'Reward for Match, Penalty for Mismatch' (2, -3), and '10' to 'Exclude Low Complexity Regions' (No).

- 1 配列長が 30 以下の短い配列質問式の際に用いるパラメータ
- 2 配列質問式に対して回答配列の類似領域中で一致する割合 (一致したコード ÷ 類似領域 × 100)
- 3 配列質問式の配列長に対して、ヒットした回答の配列と重複している長さの割合 (類似領域 ÷ 配列質問式 × 100)
- 4 検索タイプ

検索タイプ	検索機能	質問式	回答
BLASTn	塩基配列の質問式に類似した塩基配列を検索	核酸	
Mega BLASTn	塩基配列の質問式により類似した塩基配列を高速で検索	核酸	
BLASTn-short	短い塩基配列の質問式に類似した塩基配列を検索	核酸	
tBLASTn-fast	データベース中の塩基配列をアミノ酸配列に翻訳した配列の中から、アミノ酸配列の質問式に類似した配列を高速で検索	タンパク質	核酸
BLASTp	アミノ酸配列の質問式に類似したアミノ酸配列を検索	タンパク質	
BLASTp-fast	アミノ酸配列の質問式に類似したアミノ酸配列を高速で検索	タンパク質	
BLASTp-short	短いアミノ酸配列の質問式に類似したアミノ酸配列を検索	タンパク質	
BLASTx-fast	塩基配列の質問式をアミノ酸配列に翻訳して、データベース中の類似したアミノ酸配列を高速で検索	核酸	タンパク質

- 5 ギャップを考慮してマッチさせる/させない
- 6 検索用の文字列の長さ
- 7 期待値 (デフォルトは 10)
- 8 Open Gap Cost と Extend Gap Cost の組み合わせ
- 9 マッチやミスマッチの場合のペナルティスコア
- 10 低複雑度領域のマスクフィルタリングの使用

■ CDR 配列検索

Biosequences で CDR タブを選択すると、抗体や T 配列細胞受容体の CDR を指定した配列検索を実行することができます。

- ① Biosequences 画面で CDR タブをクリックします。
- ② 配列質問式を入力するか Upload (.txt または .fasta) で呼び出します。
- ③ 回答の上限数を選択します。
- ④ 検索を実行します。

Searching for...

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

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	RASOGIRNYLA	X
CDR2	AASTLQS	X
CDR3	QRYNRAPYT	X

Limit Total Sequence Results to: 20000

Start Biosequence Search

- ⑤ 検索が終わると検索履歴中の *Searching* の表示が *Complete* に変わります。View Results ボタンをクリックすると、配列検索の結果が表示されます。

ヒットした回答配列の記載特許をまとめて検索

CDR 配列検索では図中の特定のエリアの回答に限定できる

回答の並び替え

アライメントの類似性

アライメントの概略図

Subject タブ: 回答配列の詳細情報

References タブ: 出典文献情報

Alignment タブ: アライメント・スコア値・期待値

記載文献(特許もしくは非特許)の検索

Biosequences (20,000)

Sort: Alignment Identity View: Collapsed

Query Details [View More](#)

151

Alignment Identity: 100%

Matches: 9 Mismatches: 0

Subject 1 107

View Less

Alignment Subject References

Alignment Data

BLAST Score: 67

E-Value: 0.0362696

CDR1 2 ASOGIRNYL 10

|||||||

S 25 ASOGIRNYL 33

152

Alignment Identity: 100%

Matches: 27 Mismatches: 0

Subject 1 107

■ Motif 配列検索

Biosequences で Motif タブを選択すると、短いパターン配列を検索することができます。

- ① Biosequences 画面で Motif タブをクリックします。
- ② 配列質問式を入力します。角括弧を用いることにより、代替残基を含めた配列検索を行うことができます。
- ③ (任意) パラメータを設定します。
- ④ 配列質問式を選択選択します
- ⑤ 回答の上限数を選択します。
- ⑥ 検索を実行します。

Searching for...

All Substances Reactions References Suppliers Biosequences

Biosequences

Enter a protein or nucleotide string. Learn more about Biosequence Search.

BLAST CDR **Motif** Clear Search

[L][ALKV]G[FL][VI]D[AG]DG

Sequence Type: Nucleotide **Protein**

Limit Total Sequence Results to: 20000

Start Biosequence Search

Advanced Biosequence Search ^

Query Coverage % 90 E-Value 10 Reset All

- ⑦ 検索が終わると検索履歴中の *Searching* の表示が *Complete* に変わります。View Results ボタンをクリックすると、配列検索の結果が表示されます。

Biosequences (4,737)

Sort: Alignment Identity View: Collapsed

References

Query Details View More

> Seq 1: 1 LAGFVDADG 9

Alignment Identity: 100%

Matches: 9 Mismatches: 0

Query 1 9

Subject 1 354

View Less

Alignment Subject References

Alignment Data

BLAST Score: 63

E-Value: 0.429846

Q 1 LAGFVDADG 9

S 13 LAGFVDADG 21

References タブ: 出典文献情報

Alignment タブ: アライメント・スコア値・期待値

Subject タブ: 回答配列の詳細情報

代替残基を含めた検索結果の場合は、Query Details の配列をクリックして表示されるリストから別の配列を選択することができます

アライメントの概略図

アライメントの類似性

回答の並び替え

ヒットした回答配列の記載特許をまとめて検索

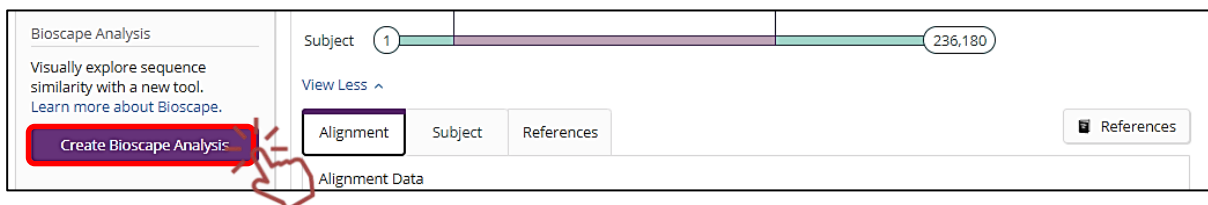
記載文献(特許もしくは非特許)の検索

■ Bioscape (バイオスケープ)

Bioscape は配列検索で得られた回答集合から、配列の類似性により解析したマップを作成する機能です。マップ内には関連特許の件数が 3D で示されるため配列関連特許を視覚的に分かりやすく解析することができます。システム制限（一度に解析できる配列の件数）は 1,000 件です。

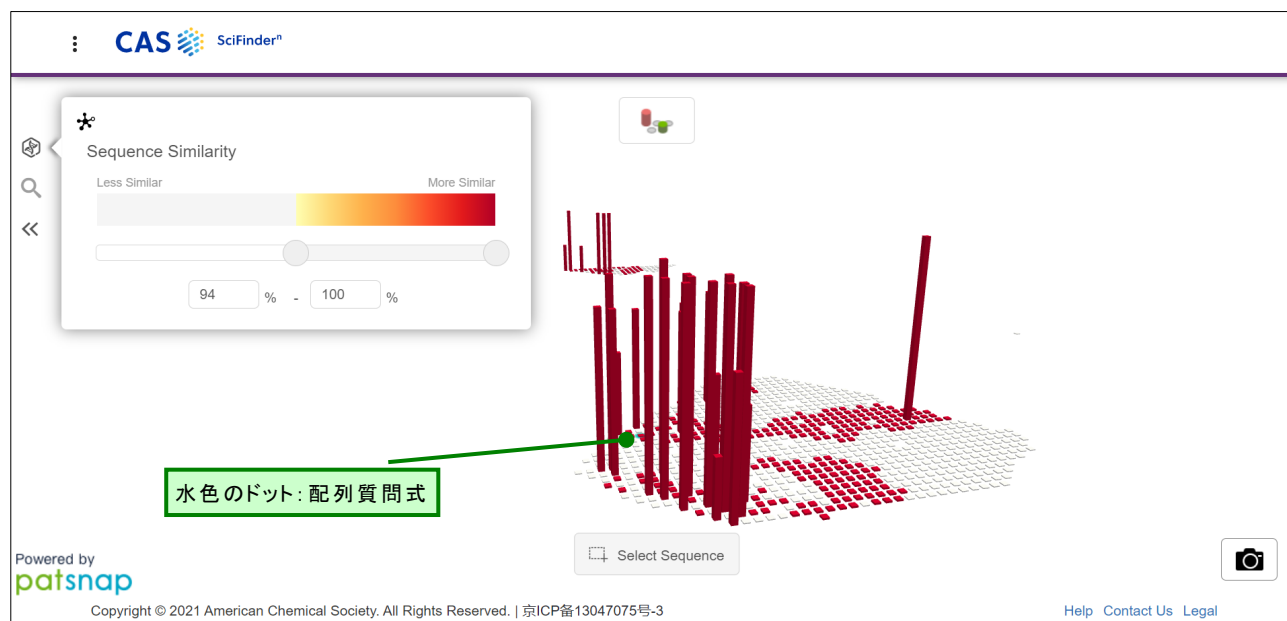
Bioscape 解析マップ作成手順

配列検索結果の画面左側に表示される「Create Bioscape Analysis」ボタンをクリックします。



別のタブに Bioscape の解析結果が表示されます。実行直後は、Sequence Similarity で解析した配列の類似性によるマップが表示されます。検索に用いた配列質問式が水色のドットです。その周囲に類似性が高い配列が赤色で、類似性が低い物質は黄色で示されます。

3D マップのバーの高さは、特許の件数を表しています。バーをクリックすると、該当する配列のモーダルウィンドウが表示され、その中の数字のリンクをクリックすると特許を表示することができます。



【コントロールパネル】画面左にはコントロールパネルが表示され、下記の操作を実行できます。

	Chart	マップに反映する類似度の設定
	Search	解析結果内の該当物質の検索

【範囲の選択】

中央下部にある **Select Sequence** をクリックすると、ポインタが範囲指定用のツールに変わり、マウスをドラッグすることでその範囲に含まれる配列をまとめて選択することができます。選択された配列と関連特許の件数が左のボックス内に表示されます。

また、画面右下には View Patent のリンクが表示されます。このリンクから、選択した配列の特許をまとめて表示することができます。

範囲選択を終了するには Exit をクリックします。

[illegible]

同主鎖の配列を持つ物質の検索

同じ配列であっても異なる CAS RN® が付与されたレコードが存在することがあります。

- ・ 配列の主鎖は同じであっても、化学修飾、側鎖の置換基の異なるもの、同位体で置換された配列は別配列として別レコードに収録されます。化学修飾などの情報は特徴表に記載されます。
- ・ GenBank 由来の核酸配列では、更新前と更新後の配列は、同じ配列であっても別レコードとして収録されています。
- ・ 登録ルールの例外によって、同じ配列であっても別レコードとして収録されている場合があります。

■ Related Sequences

同主鎖の配列を持つ配列を調べるには、Related Sequences 機能を使います。

核酸やタンパク質の物質の詳細情報を表示し、Related Sequences ボタンをクリックすると同主鎖の配列を持つ物質を検索できます。

同主鎖の配列レコードが 12 件存在する

Related Sequences (12)

Sequence: Linear
1 MRKSPGLSDP LWARILLST LTGRSYGGPS LQELKQNTT VFTRILDRLL
51 DGYNRLRPG LGERVTEVKT DIFVTSFGPV SDHMEYITD VFFRGSWYDE
101 RLKFKGPMIV LRLNLWASK INTPTDFHNN QKKSVAHMT MPNKLRLITE
151 DGTLLYTNRL TYRAECPMHL EDPMDAHAC PLKFGSYAYT RAEVYVETR
201 EPARGVVAE DGSRLNGYDL LGQTVDSGIV QGSTGEYVVM TTHFLKRLKI
251 GYFVIGTYLP CINTVILSQV SFNLNRESVP ARTVFGVTTV LTTWTLSSISA
301 RNSLPKXVAY TANDVFJAVC YAFVFSALIE FATVWYFTR GYAWDGKSYV
351 PEKPKVWDP LIKKNNTYAP TATSYTPNLA RGPGLATIA KSAITIEPEV
401 KPETHPEPK KTFNSVSKID RLSRIAPLL FGFINLVYVA TYLNREPQLK
451 APTPHQ - - -

同主鎖の配列を持つ物質

CAS No.	Substance Name	Sequence Length
479879-70-4	Unspecified Gamma-aminobutyric acid (GABA) A receptor alpha 1 (human clone MGC26564 IMAGE4820972) (9CI)	456
176229-25-7	Unspecified GABAA receptor GABRA1 (human brain)	456
2098448-41-8	Unspecified GABAA receptor GABRA1 (human)	456
1700675-57-5	Unspecified GABAA receptor GABRA1 (human)	456
1700675-53-1	Unspecified GABAA receptor GABRA1 (human precursor)	456
1448554-62-8	Unspecified Protein/Peptide Sequence	456
1448104-12-8	Unspecified GABAA receptor GABRA1 (human clone PR296331-SEQID-23)	456
1190908-30-5	Unspecified GABAA receptor GABRA1 (human)	456
1172644-36-8	Unspecified GABA (gamma-aminobutyric acid) A1 receptor alpha 1 (human)	456
1087358-21-1	Unspecified GABAA receptor GABRA1 (human)	456
628759-25-1	Unspecified Protein (human clone US030219744-SE QID-331)	456
528642-91-3	Unspecified Protein (human clone US569662-SEQID-331)	456