

LREGISTRY

LREGISTRY ファイルは REGISTRY ファイルの使い方を学習するためにつくられた練習用データベースです。このファイルは、Chemical Abstracts Service (CAS) の化学物質登録システム (Registry System) によって登録された約 125,000 の物質を収録しており、構造検索および辞書検索が可能です。レコードは、LCA ファイルおよび LCASREACT ファイル中にて索引された物質で構成されています。

すべての物質レコードにはユニークな CAS 登録番号[®]、CA 索引名が収録されています。物質レコードには同義名、分子式、合金成分表、ポリマークラス、核酸とタンパク質の配列、構造図などが含まれており、これらのすべてを検索および表示することができます。

後方一致検索は、化学物質名セグメントフィールド (/CNS) と注記フィールド (/NTE) で可能です。

収録内容

LREGISTRY ファイルは、文献に記載されたすべての種類の化学物質を収録しています。合金、配列、配位化合物、鉱物、混合物、ポリマーおよび塩を含むあらゆる種類の無機および有機物質が収録されています。

収録源

雑誌文献、特許、学会会議録に記載された新規物質、および法規制台帳の物質を収録する CAS の化学物質登録システム

ファイル内容

約 123,000 件のレコード

レコードの件数は一定

レコードの内容は REGISTRY ファイルと同時に更新される

検索補助資料

STN 構造検索ユーザーガイド (和訳版)

REGISTRY ファイル辞書検索ユーザーガイド (和訳版)

REGISTRY File: Basic Name Segment Dictionary (和訳版)

REGISTRY File: POLYMER CLASS TERMS (和訳版)

CA Index Guides (Naming & Indexing of Chemical Substances for CA を含む)

Finding and Verifying CAS Registry Numbers on STN

How to Search for CAS Registry Numbers in the CAS Registry File

Nucleic Acid Sequences on STN: A Quick Reference Guide

Polymer Information on STN: A Quick Reference Guide

Protein Sequences on STN: A Quick Reference Guide

REGISTRY File: Biosequence Searching Manual

Screen Dictionary (Adding Screens in Structure Searching を含む)

Searching CASLINK Quick Reference Card

Searching Coordination Compounds

Searching for Polymer Information

ヨーロッパ

STN カールスルーエ

FIZ Karlsruhe
P.O. Box 2465
76012 Karlsruhe
Germany
Phone: +49-7247-808-555
Fax: +49-7247-808-259
E-mail: helpdesk@fiz-karlsruhe.de
Internet: www.stn-international.de

日本

STN 東京

一般社団法人 化学情報協会
〒113-0021 東京都文京区本駒込6-25-4 中居ビル
Phone: 0120-003-462 (Help Desk)
: 0120-151-462 (上記以外)
Fax: 03-5978-4090
E-mail: support@jaici.or.jp (Help Desk)
customer@jaici.or.jp (上記以外)
Internet: www.jaici.or.jp

北アメリカ

STN コロンバス

CAS
P.O. Box 3012
Columbus, Ohio 43210-0012 U.S.A
CAS Customer Care:
Phone: 800-753-4227 (North America)
614-447-3700 (worldwide)
Fax: 614-447-3751
E-mail: help@cas.org
Internet: www.cas.org

検索補助資料

STNote4: Send and View Graphic Images in STNmail
 STNote6: CAS Registry Enhancements
 STNote7: Current awareness on STN
 Structuring Functional Groups
 Using Stereosearch Quick Reference Card
 Using the CAS Registry File on STN Student Manual
 Using the CAS Registry File on STN Structure Searching Student Manual
 オンラインヘルプ (HELP DIRECTORY で利用できるすべてのヘルプメッセージのリストを表示できます)
 STNGUIDE

データベース製作者

Chemical Abstracts Service
 2540 Olentangy River Road
 P. O. Box 3012
 Columbus, OH 43210-0012 USA
 Phone: (+1)614-447-3600
 Fax: (+1)614-447-3798

データベース代理店

(社) 化学情報協会
 〒113-0021 東京都文京区
 本駒込6-25-4 中居ビル
 TEL 03-5978-3601
 FAX 03-5978-3600
 URL: <http://www.jaici.or.jp/>

SEARCHおよびDISPLAYフィールド

後方一致検索が利用できるフィールドはアスタリスク (*) が付与されています。

| フィールド | SEARCH コード | SEARCH 例 | DISPLAY コード |
|---|---------------|--|-------------------------|
| ベーシックインデックス 名称フラグメント 分子式フラグメント ¹⁾ 累積索引コード | なし または/BI | S TOSYL S DIMETHYL ADIPATE S 6CI S 1, 1(W) DICHLORO S C5H10BR202 | AF, CN, IN, MF |
| CAS登録番号 | /RN | S 97-77-8/RN S 97-77-8 | RN, AR, DR, PR CI |
| クラス識別子 (コードまたはフレーズ) | /CI | S MXS/CI S ALLOY/CI | CRN DEF |
| 成分CAS登録番号 | /CRN | S 79-10-7/CRN | 表示されない |
| 物質の定義 | /DEF | S HYDROCARBONS/DEF | 表示されない |
| 入力日 ²⁾ | /ED | S 890810/ED | |
| フィールドの存在 (コードまたはフレーズ) | /FA | S RSD/FA AND L5 S MATERIAL COMPOSITION/FA | |
| ファイルセグメント (頭字語または単語) | /FS | S 3D/FS S PROTEIN/FS S PS/FS S NUCLEIC/FS | FS |
| ポリマー分類用語 (コードまたはテキスト) | /PCT | S POLYAMINE/PCT S PM/PCT | PCT |
| CAS登録番号所在 | /LC | S TSCA/LC | LC |
| 更新日 ²⁾ | /UP | S UP>=890000 | 表示されない |

1) 分子式フラグメントをベーシックインデックスで検索するときにはスペースなしで入力します。

2) 数値演算子または範囲指定による検索が可能な数値検索フィールドです。

化学物質名称フィールド

| フィールド | SEARCH コード | SEARCH 例 | DISPLAY コード |
|------------------|---------------|--|----------------|
| 化学物質完全名称 | /CN | S 1-CHLORO-1, 3-BUTADIENE/CN S INTERFERON. ALPHA. 1?/CN | CN, IN |
| 化学物質名セグメント* 1) | /CNS | S IMINO/CNS S ?QUAT?/CNS NOT AQUA | CN, IN |
| 見出し語母核 | /HP | S BENZOIC ACID/HP | CN, IN |
| 索引名セグメントー見出し語母核 | /INS. HP | S METHYLETHYL/INS. HP | CN, IN |
| 索引名セグメントー非見出し語母核 | /INS. NHP | S ACRYLO/INS. NHP | CN, IN |
| 他の名称セグメント | /ONS | S ANILINE/ONS | CN |

1) 後方一致を利用する時には入力語は最低4文字必要です。

分子式フィールド

| フィールド | SEARCH コード | SEARCH 例 | DISPLAY コード |
|---|---------------|---|----------------|
| 原子数 ¹⁾ | /ATC | S 5/ATC | 表示されない |
| 元素数 ¹⁾ | /ELC | S 7-9/ELC | 表示されない |
| 物質に対する元素数 ¹⁾ | /ELC. SUB | S ELC. SUB>=8 | 表示されない |
| 元素式 ²⁾ | /ELF | S AL CO LA O/ELF | AF, MF |
| 元素比 xx ¹⁾ (xx=CH, CN, CO, HC, HN, HO, NC, NH, NO, OC, OH, またはON) | /ELR. xx | S 3. 1666667/ELR. CH S 1-2/ELR. CN S ELR. CO<=1 | 表示されない |
| 元素種 | /ELS | S B/ELS AND H/ELS | 表示されない |
| 多成分物質に対する元素種 | /ELS. MCF | S (N (XA) P)/ELS. MCF | 表示されない |
| 分子式量 ¹⁾ | /FW | S 420-460/FW | 表示されない |
| 物質組成 ³⁾ | /MAC | S 1-5 ND/MAC | STR |
| 分子式 ⁴⁾ | /MF | S C7H3BR2F02/MF S C4H4O4. 2NA/MF S C24 H37 OS P3/MF | AF, MF |
| 成分数 ¹⁾ | /NC | S F/ELS NOT NC>=2 | 表示されない |
| 周期律グループ | /PG | S B6/PG S LNTH/PG | 表示されない |
| 相対組成 | /RC | S FE. CR. NI/RC | 表示されない |
| 特定元素数 ¹⁾ | /元素記号 | S 7/SI | 表示されない |

1) 数値演算子または範囲指定検索のできる数値検索フィールドです。

2) 元素式では、各元素間にスペースを入力しなければなりません。

3) 数値およびテキストフィールドの組み合わせ。組成は数値であり、数値演算子または範囲指定による検索が利用できます。成分はテキスト用語です。

4) 分子式は、スペースを入れても、入れなくても入力できます。

環データ

| フィールド | SEARCH コード | SEARCH 例 | DISPLAY コード |
|--|---------------|--------------------------------|----------------|
| 環系の元素式 ¹⁾ (および成分構造内のEAの存在数) | /EA | S C4N-C5N/EA S 2 C3N0-C6/EA | RSD |
| 最小環の元素式 ¹⁾ (および環系内のEASの存在数) | /EAS | S C5N04/EAS S >9 C6/EAS | 表示されない |
| 環系の元素配列 ¹⁾ (および成分内のESの存在数) | /ES | S NCOC2-C6/ES S 1-3 O2C4/ES | RSD, SRSD |
| 最小環の元素配列 ¹⁾ (および環系内のESSの存在数) | /ESS | S FE3/ESS S >=2 SC2SC2/ESS | 表示されない |

(続く)

環データ

| フィールド | SEARCH コード | SEARCH 例 | DISPLAY コード |
|---|---------------|---|----------------|
| 環系数 ²⁾ | /NRS | S 7/NRS | 表示されない |
| 成分内の環系数 ²⁾ | /CNRS | S 4-5/CNRS | 表示されない |
| 最小環の数(物質全体) ²⁾ | /NR | S 10/NR | 表示されない |
| 成分内の最小環の数 ²⁾ | /CNR | S CNR>=12 | 表示されない |
| 環系内の最小環の数 ²⁾ | /NRRS | S 5-6/NRRS | 表示されない |
| 環系の原子数 ²⁾ | /RATC | S 4/RATC | 表示されない |
| 環系の構成元素種 ¹⁾ (および環系内のRELの存在数) | /REL | S SE/REL S 5 P/REL | 表示されない |
| 環系内の元素種の数 ²⁾ | /RELC | S 6/RELC | 表示されない |
| 環系の構成元素式 ^{1), 3)} (および成分内のRELFの存在数) | /RELF | S C N O P/RELF S >3 C N O/RELF | 表示されない |
| 環系識別子 ¹⁾ (および成分内のRIDの存在数) | /RID | S 31779. 1. 2/RID S 1938/RID S >=2 1949. 52/RID | RSD, SRSD |
| 最小環の環の大きさ ^{1), 2)} (および環系内のSZSの存在数) | /SZS | S 8/SZS S 5 4/SZS | 表示されない |
| 環系式 ¹⁾ (および成分内のRFの存在数) | /RF | S C20AGN4/RF S 5 C10/RF | RSD |
| 環系の環の大きさ ¹⁾ (および成分内のSZの存在数) | /SZ | S 3-4-5/SZ S 3 5-5-6/SZ | RSD |

1) 存在数は検索フィールドの最初に入力しなければなりません。これは数値であり、数値演算子または範囲指定検索ができます。

2) 数値演算子または範囲指定による検索が可能な数値検索フィールドです。

3) 元素式は、各元素間にスペースを入力しなければなりません。

配列検索

| フィールド | SEARCH コード | SEARCH 例 | DISPLAY コード |
|-----------------------|---------------|--|----------------|
| 注記* ¹⁾ | /NTE | S CYCLIC/NTE S ?CHLORO?/NTE S OAA-17/NTE | NTE |
| 核酸数 ^{2), 3)} | /NA. CNT | S 12-42/NA. CNT | NA |
| 核酸タイプ ³⁾ | /NA | S 12-42 A/NA S G/NA | NA |
| 配列長 ²⁾ | /SQL | S 4-20/SQL S SQL<=500 | SQL |

1) 後方一致では、入力語は少なくとも4文字含まれることが必要です

2) 数値演算子または範囲指定による検索が可能な数値検索フィールドです。

3) 核酸配列データのみ含まれるフィールドです。

制限検索コード

| フィールド | SEARCH コード | SEARCH 例 | DISPLAY コード |
|--------------------|---------------------------|------------------------|----------------|
| イタレーションが完全に行われた回答 | /COMPLETE ¹⁾ | S L4/COM ²⁾ | 表示されない |
| イタレーションが不完全に終わった回答 | /INCOMPLETE ¹⁾ | S L4/INC ²⁾ | 表示されない |

1) 最初の3文字まで省略可能です。

2) REGISTRYファイルで作成された回答セットのL番号にのみ制限が可能です。

構造検索の質問式

| 質 問 式 ^{1), 2)} | 検 索 例 |
|---|--|
| STRUCTUREコマンドを用いて作成された、またはSTN Express/ STN on the Webからアップロードされた構造のL番号 (L番号の間でブール演算子が利用できる) | SEARCH L1 FAM SAM SEA L1 AND L2 SSS FUL |
| SCREENコマンドを用いて作成されたスクリーンセットのL番号 (L番号の間でブール演算子が利用できる) | S L3 OR L4 SSS SAM |
| STRUCTUREコマンドを用いて作成された、またはSTN Express/ STN on the Webからアップロードされた構造のL番号とSCREENコマンド を用いて作成されたスクリーンセットのL番号の組み合わせ (L番号の間でブール演算子が利用できる) | S L1 AND L2 NOT L3 |

1) 構造検索の回答セットのL番号は、辞書検索の用語と組み合わせることができます (例えば、S L3 AND TSCA/LC)。

2) 配列検索の質問式については、配列検索の質問式の項を参照下さい。

構造検索の種類

| 種 類 ¹⁾ | 内 容 | SEARCH コード | SEARCH 例 |
|-----------------------------------|--|---------------|--|
| Substructure 部分構造検索 (デフォルト) | 構造質問式を満足する物質を検索。 すべての開かれた位置に追加の置換基が 結合してもよい。他の成分が含まれても よい。 | SSS | SEARCH L1 SSS FUL S L2 OR L3 SSS SAM S L7 SSS RAN |
| Closed Substructure 閉構造部分構造検索 | 構造質問式に完全に一致する物質を検索。 CONNECTによって許された位置に置換基が 結合してもよい。他の成分が含まれても よい。 | CSS | SEARCH L1 CSS FUL S L2 NOT L3 CSS S L4 OR L5 CSS RAN |
| Family ファミリー検索 | 構造質問式に完全に一致する物質を検索。 他の成分が含まれてもよい。 | FAM | S L6 FAM SAM |
| Exact 完全一致検索 | 構造質問式に完全に一致する物質を検索。 | EXA | SEA L5 EXA FUL |

1) 配列検索の種類については、配列検索の種類項を参照下さい。

構造検索の範囲

| 範 囲 | 内 容 | SEARCH コード | SEARCH 例 |
|-------------------|---|---------------|--|
| Sample (デフォルト) | ファイルの固定された5%を検索 | SAM | SEARCH L3 EXA SAM |
| Full Range | ファイル全体 (100%) を検索 ユーザーが指定した範囲内で検索 | FUL RAN | S L6 NOT L7 SSS SAM S L5 OR L8 SSS FUL S L4 RAN=(110507-58-9,) S L3 FAM RAN=(109784-14-7, 109904-92-9) |
| Subset Sample | LREGISTRYファイルの検索で得られた 回答セットの中をサンプル検索 | SUB SAM | S L7 CSS SUB=L5 SAM |
| Subset Range | LREGISTRYファイルの検索で得られた 回答セットの中をユーザーが指定した 範囲で検索 | SUB RAN | S L3 SUB=L2 RAN=(, 50-11-3) |
| Subset Full | LREGISTRYファイルの検索で得られた 回答セット全体 (100%) を検索 | SUB FUL | S L8 SUB=L6 FAM FUL |

DISPLAYおよびPRINT形式

カスタム表示形式のどのような組合せでも、また定型表示形式のどのような組合せでも利用できます。しかし、カスタム表示形式と定型表示形式を組み合わせることはできません。複数の表示形式は、カンマまたはスペースで区切ってください。フィールドは、入力した順に表示されます。

HITおよびKWIC表示形式を利用するためには、検索時にハイライトはONであることが必要です。

CM (成分番号) フィールドは多成分物質のレコードで表示されますが、カスタム表示形式ではなく、表示で指定することはできません。

辞書フィールドコード

| 形式 | 英語名 | 内容 | 入力例 |
|--------------------|---|------------------------------|------------------|
| AF | Alternate Molecular Formula | 非優先分子式 | D L4 1-4 AF |
| AR | Alternate Registry Number | 非優先CAS登録番号 | D L1 3 AR |
| CCI | Component Class Identifier | 成分クラス識別子 | D CCI 1, 3-5 |
| CCN ¹⁾ | Condensed Chemical Name | 圧縮型化学物質名称 | D 20 CCN |
| CDES | Component Descriptor | 成分ディスクリプタ | D CDES 5-10 |
| CI | Substance Class Identifier | クラス識別子 | D 1-3, 7, 8 CI |
| CIL | Component Isotope at Unknown Location | 位置不明の成分同位体 | D CIL |
| CMF | Component Molecular Formula | 成分分子式 | D L1 CMF 3 |
| CN | Chemical Name | 化学物質名称 | D CN |
| COMP ²⁾ | Composition | 組成 | D L7 |
| CRN | Component Registry Number | 成分CAS登録番号 | D 1, 3, 6 CRN L5 |
| DEF | Definition | 定義 | D DEF |
| DES | Descriptor | ディスクリプタ | D DES 2 |
| DR | Deleted Registry Number | 削除されたCAS登録番号 | D L8 DR 1-3 |
| FCN ¹⁾ | Full Chemical Name | すべての化学物質名称 | D FCN L3 7 |
| FS | File Segment | ファイルセグメント | D 1, 4 FS |
| IL | Isotope at Unknown Location | 位置不明の同位体 | D IL |
| IN | CA Index Name | CA索引名 | D IN L1 4 |
| LC | Registry Number Locator | CAS登録番号所在 | D LC 3, 4 |
| MF | Molecular Formula | 分子式 | D MF |
| PCT | Polymer Class Term | ポリマー分類用語 | D L3 PCT |
| PR | Preferred Registry Number | 優先 CAS 登録番号 | D 5, 3 PR |
| RN | CAS Registry Number | CAS登録番号 | D L4 RN 3 |
| RR | Replaced Registry Number | 置換されたCAS登録番号 | D L3 2 RR |
| RSD ³⁾ | Ring System Data | 環系データ | D RSD |
| SCN ⁴⁾ | Short Chemical Name | 簡略型化学物質名称 | D 5-9 SCN |
| SR | Source of Registration | 登録情報源 | D SR 1, 3 L12 |
| SRSD ⁵⁾ | Short Ring System Data | 簡略型環データ | D SRSD |
| STF | Flat Structure (no stereo indicated) | 平面構造図 (立体情報なし) | D L9 1 3 |
| STR ⁶⁾ | Structure Diagram (includes stereo bonds and R/S/E/Z labels when available) | 構造図 (あれば立体結合、R/S/E/Z ラベルを表示) | D STR |
| STS ⁶⁾ | Stereo Structure (includes stereo bonds when available) | 立体構造図 (あれば立体結合を表示) | D CN STS |

1) 名称は、表示コードCNで表示されます。

2) 合金および表形式無機化学物質に対する成分情報、成分CAS登録番号をまとめた表示形式の表示です。

3) EA, ES, SZ, RF, RID, およびRIDの存在数をまとめた表形式の表示です。

4) CA索引名およびその他のすべての名称は、表示コードCNで表示されます。

5) EA, RID, およびRIDの存在数をまとめた表形式の表示です。

6) 立体構造はグラフィック機能付き通信ソフト、またはSTN on the Webを利用した場合およびオフラインプリントで表示可能です。

配列フィールドコード

| 形式 | 英語名 | 内容 | 入力例 |
|------|---------------------------|--------------|-------------|
| NA | Nucleic Acid | 核酸 | D 6 9 11 NA |
| NTE | Note | 注釈 | D NTE |
| SEQ | Sequence (1-letter codes) | 配列 (1-文字コード) | D SEQ |
| SEQ3 | Sequence (3-letter codes) | 配列 (3-文字コード) | D SEQ3 1-10 |
| SQL | Sequence Length | 配列長 | D L3 SQL |

定型表示形式

| 形式 | 内容 | 入力例 |
|--------------------|--|------------------|
| SQD | RN, AR, PR, DR, RR, FS, SQL, NTE, SEQ | D 5 SQD |
| SQD3 | RN, AR, PR, DR, RR, FS, SQL, NTE, SEQ3 | D 2-4 SQD3 |
| SQIDE | RN, CN, DEF, AR, PR, DR, RR, FS, SQL, NA, NTE, SEQ, MF, AF, CI, PCT, SR, LC, IL, DES, STR | D L4 SQIDE |
| SQIDE3 | SQIDEと同様、タンパク質配列は3文字コードで表記 | D L4 SQIDE3 |
| SNQ | RN, CN, AR, PR, FS, SQL, DR, RR | D SNQ L5 6-9 |
| ALL | すべての存在するフィールドおよび名称、配列データ | DISPLAY L1 1 ALL |
| FIDE | 配列データを除く、すべての存在するフィールドおよび名称 (RN, CN, DEF, AR, PR, FS, DR, RR, MF, AF, CI, PCT, SR, LC, IL, DES, RSD, CRN, CMF, CCI, CDES, CIL, STR, COMP) | D FIDE 3 7 L6 |
| IDE | FIDEと同様、名称は50項目まで表示されるが、RSDは表示されない (デフォルトはIDE) | D IDE L10 |
| REG | CAS登録番号 (RN, DR, AR, PR, RR) | D REG |
| SAM | IN, SQL, MF, CI, STR, COMP | D L3 1-18 SAM |
| SCAN ¹⁾ | IN, SQL, MF, CI, STR, COMP (回答番号なしのランダム表示) | D SCAN |
| HIT ²⁾ | ヒットタームを含むすべてのフィールド | D HIT 5-10 |
| KWIC ²⁾ | ヒットタームとその前後20語 | D KWIC 5-10 |

1) この表示形式のオンライン・ディスプレイ料金は無料です。

SCANはコマンドに続けて入力します。例: D SCAN または DISPLAY SCAN

2) HITおよびKWICは、MAC、RC、およびCRNを除くすべての辞書項目の表示形式および配列の表示形式で利用できます。KWICはDEFとLCを除くすべてのフィールドに対してHITと同様です。ヒットした個々のタームがハイライトされるDEFとLCを除くすべてのフィールドでは、ヒットタームを含むフィールド全体がハイライトされます。RSDの表は、ハイライトなしで表示されます。NTEフィールドでは、ヒットタームを含む表の行がハイライトなしで表示されます。SEQとSEQ3は、ヒットしたアミノ酸コードが、下線および配列のヒットしたアミノ酸コードの位置の表示によってハイライトされます。

配列検索の質問式

| 質問式 | SEARCH 例 |
|--|--------------------------------|
| 一般アミノ酸の1文字コード ¹⁾ | S LAGLL/SQSP |
| 一般アミノ酸および特殊アミノ酸の3文字コード ^{1), 2)} | S 'LEU-ALA-GLY-LEU-LEU' /SQSFP |
| コードまたはコードの列を一重引用符で囲む。 | S F'HCY-STA' LF/SQSP |
| 各コード間にハイフンを入力する。 | S 'GLP' AGYSK/SQEP |
| 核酸の1文字コード ³⁾ | S 'CYS-ASN-THR-ALA' /SQEP |
| | S ATTTTTTTTTT/SQEN |
| | S AAGGTTACTA/SQSN |

1) 矢印プロンプトで、HELP AACと入力することによって一般アミノ酸の1および3文字コード表が表示されます。

2) 矢印プロンプトで、HELP AAUと入力することによって特殊アミノ酸の3文字コード表が表示されます。

3) 矢印プロンプトで、HELP NUCと入力することによって核酸コード表が表示されます。

配列検索の種類

核酸およびタンパク質配列データは、SEQフィールドに1文字コードが、タンパク質のみSEQ3で3文字コードが表示されます。

| 種 類 | 定 義 | SEARCH コード | SEARCH 例 |
|-----------------|--|---------------|---|
| タンパク質、完全配列 | 質問式に一致する配列を検索。 質問式は完全に定義されていなければならない。 | /SQEP | S YADAIF/SQEP S ' CYS-ASN-THR-ALA' /SQEP |
| タンパク質、完全配列ファミリー | 質問式に一致する配列と質問式の アミノ酸の等価ファミリーに 相当する配列を検索。 ¹⁾ | /SQEFP | S YGGFL/SQEFP S ' TYR-GLY-GLY-PHE-LEU' /SQEFP |
| タンパク質、部分配列 | 質問式に完全に一致する配列と 質問式の配列を含む配列を検索。 多様な記号が利用できる。 | /SQSP | S LAGLL/SQSP S F' HCY-STA' LF/SQSP |
| タンパク質、部分配列ファミリー | 質問式に完全に一致する配列と 質問式の配列を含む配列と 質問式のアミノ酸の等価ファミリー に相当する配列を含む配列を検索。 多様な記号が利用できる。 ¹⁾ | /SQSFP | S ATCXAW/SQSFP S ' LEU-ALA-GLY-LEU-LEU' /SQSFP |
| 核酸、完全配列 | 質問式に完全に一致する配列を検索。 核酸の曖昧コードが利用できる。 | /SQEN | S ATTTTTTTTT/SQEN |
| 核酸、部分配列 | 質問式に完全に一致する配列と、 質問式の配列を含む配列を検索。 核酸の曖昧コードおよび多様な 記号が利用できる。 | /SQSN | S AAGTTACTA/SQSN |

1) タンパク質の等価ファミリーは以下のとおりです。

| | |
|------------------|------------|
| P, A, G, S, T | (弱疎水性、中性) |
| Q, N, E, D, B, Z | (親水性、酸アミド) |
| H, K, R | (親水性、塩基性) |
| L, I, V, M | (疎水性) |
| F, Y, W | (疎水性、芳香族) |
| C | (架橋) |

部分配列の多様な記号 (/SQSP, /SQSFP, および/SQSN)¹⁾

| 記 号 | 機 能 | SEARCH 例 |
|------------------|---|--|
| [] | 代替残基を特定 | S LGP[VL]/SQSP S LGP[' VAL' ' LEU']/SQSP |
| [-] | 特定または代替残基を除く | S LGP[-H]/SQSP S LGP[-' HIS']/SQSP S LGP[-HL]/SQSP |
| {m} | 直前の配列または配列質問式 (L#, E#, または保存された質問式) をm回繰り返す | S (FL) {2}/SQSP S L4 {2}/SQSP S NAME/Q {3}/SQSP S (CTG) {2}/SQSN S TAA (TAAA) {2}/SQSN |
| {m, u} または {m-u} | 直前の配列または配列質問式 (L#, E#, または保存された質問式) をmからu回 繰り返す | S GG (FL) {1, 2}/SQSP S L3 {1, 3}/SQSP S NAME/Q {1, 4}/SQSP S (CTG) {1, 3}/SQSN |

(続く)

部分配列の多様な記号 (/SQSP, /SQSFP, および/SQSN) ¹⁾

| 記号 | 機能 | SEARCH 例 |
|--------------------------|--|--|
| ? または {0,1} または {0-1} | 直前の配列または配列質問式 (L#, E#, または保存された質問式) をゼロまたは1回繰り返す | S FLRRI (RP)?K/SQSP S FLRRI (RP) {0,1}K/SQSP S L1 {0-1}NN/SQSP S NAME/Q {0,1}NN/SQSP S CAT (CGA) {0,1}GGAC/SQSN S KLK (WD) {0,}N/SQSP S KLK (WD)*N/SQSP S L1 {0-}NN/SQSP S NAME/Q {0,}NN/SQSP S CAT (CTG) {0,}TATT/SQSN |
| * または {0,} または {0-} | 直前の配列または配列質問式 (L#, E#, または保存された質問式) をゼロまたはそれ以上繰り返す | S KLK (DLE) {1,}/SQSP S KLK (DLE)+/SQSP S L2 {1-}/SQSP S NAME/Q {1,}/SQSP S CAT (CTG) {1,}TATT/SQSN |
| + または {1,} または {1-} | 直前の配列または配列質問式 (L#, E#, または保存された質問式) を1またはそれ以上繰り返す | S L1&L3/SQSFP S L2&L5 {1,3}/SQSP S NAME1/Q {2}&NAME2/Q/SQSP S E1&E3/SQSP |
| & | 配列表記または質問式 (L#, E#, または保存された質問式) を結合する | |

上記の他に、脱字符号 (^)、垂直バー (|)も利用できます。

^ は、配列フィールドの最初または最後の配列フラグメントを指定することができます。

| は、代替の配列質問式を分けるために使用する記号です。

1) 部分配列質問式の様々な指定法に関する詳細は、REGISTRYファイルで矢印プロンプトの次にHELP SQQと入力することによって表示されます。

部分配列検索のギャップ記号 (/SQSP, /SQSFP, および/SQSN)

| 記号 | 機能 | SEARCH 例 |
|-------------------------------------|-------------------|--|
| . | 1残基のギャップ | S SY. RPG/SQSP S SY. . RPG/SQSFS S AAG. . . TGC/SQSN |
| {m} または [m.] | m残基のギャップ | S SY. {2}RPG/SQSP S SY[2.]RPG/SQSP |
| {m, u} または .{m-u} | mからu残基のギャップ | S GFF. {2,10}LSS/SQSP S GFF. {2-10}LSS/SQSP S AAG. {2,5}TGC/SQSN |
| : または .? または . {0,1} または . {0-1} | ゼロまたは1残基のギャップ | S AGA:SRI/SQSFPS S AGA.?SRI/SQSFP S AGA. {0,1}SRI/SQSFP S AGA. {0-1}SRI/SQSFP |
| * または . {0,} または . {0-} | ゼロまたはそれ以上の残基のギャップ | S HLC.*TYG/SQSP S HLC. {0,}TYG/SQSP S HLC. {0-}TYG/SQSP S AAGGCAGATG.*GCAA/SQSN |
| + または . {1,} または . {1-} | 1またはそれ以上の残基のギャップ | S SY.+TH/SQSFP S SY. {1,}TH/SQSFP S SY. {1-}TH/SQSFP S TCCTG.+GTGG/SQSN |

SELECTおよびSORTフィールド

SELECT コマンドは、回答セットの指定したフィールドから抽出した語句にE番号またはL番号を付与します。

SORT コマンドは、検索結果を指定したフィールドのアルファベット順または数値順に並べ替えます。

(該当項目はY、該当しないものはNで表示されています。)

| フィールド | フィールドコード | SELECT ¹⁾ | SORT |
|----------------------------|----------|-------------------------|------|
| 非優先分子式 | AF | Y ²⁾ | N |
| 非優先CAS登録番号 | AR | Y ³⁾ | N |
| CA索引名 | IN | Y ⁴⁾ | Y |
| CAS登録番号 | RN | Y | Y |
| 化学物質完全名称 | CN | Y ⁵⁾ | N |
| クラス識別子 | CI | Y | N |
| 成分クラス識別子 | CCI | Y ⁶⁾ | N |
| 成分分子式 | CMF | Y ⁷⁾ | N |
| 成分CAS登録番号 | CRN | Y | N |
| 定義 | DEF | Y | N |
| 削除されたCAS登録番号 | DR | Y ³⁾ | N |
| 環系の元素式 | EA | Y | N |
| 環系の元素配列 | ES | Y | N |
| ファイルセグメント | FS | Y | Y |
| すべての化学物質名称 | FCN | Y | N |
| 分子式 | MF | Y | N |
| 化学物質名称 | NAME | Y ⁸⁾ | N |
| 核酸配列 (完全一致検索形式) | SQEN | Y | N |
| 核酸配列 (部分配列検索形式) | SQSN | Y | N |
| ポリマー分類用語 | PCT | Y | N |
| 優先CAS登録番号 | PR | Y ³⁾ | N |
| タンパク質配列 (完全一致ファミリー検索形式) | SQEFP | Y | N |
| タンパク質配列 (完全一致検索形式) | SQEP | Y | N |
| タンパク質配列 (部分配列ファミリー検索形式) | SQSFP | Y | N |
| タンパク質配列 (部分配列検索形式) | SQSP | Y | N |
| CAS登録番号所在 | LC | Y | N |
| CAS登録番号および化学物質名称 | CHEM | Y ⁹⁾ (デフォルト) | N |
| 置換されたCAS登録番号 | RR | Y ⁵⁾ | N |
| 環系識別子 | RID | Y | N |
| 環系式 | RF | Y | N |
| 配列 (1文字コード) | SEQ | Y | N |
| 配列 (3文字コード) | SEQ3 | Y | N |
| 配列長 | SQL | N | Y |
| 簡略型化学物質名称 | SCN | Y ⁴⁾ | N |
| 環系の環の大きさ | SZ | Y | N |
| 登録情報源 | SR | Y | N |

1) ヒットタームだけを抽出させるには、HITを使います。例: SEL HIT CN

2) /MFが付与されます。

3) /RNが付与されます。

4) /CNが付与されます。

5) CA索引名、アルファベット順の最初の50名称、およびすべてのヒットした名称が抽出されます。

6) /CIが付与されます。

7) /BIが付与されます。

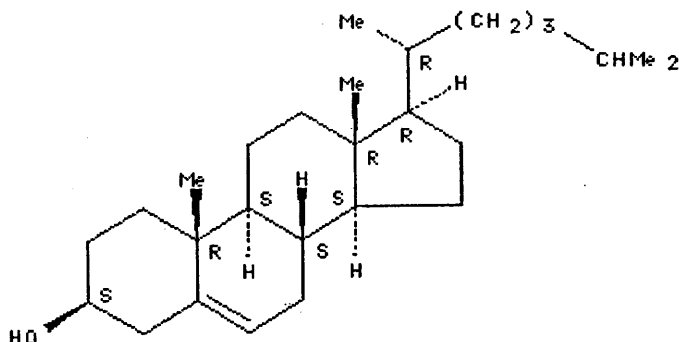
8) 倒置された名称を除くすべての名称が抽出され、/BIが付与されます。

9) AR, DR, PR, RN, RRおよび倒置された名称を除くすべての名称が抽出され、/BIが付与されます。

サンプルレコード

IDE (デフォルト) 形式での表示

RN 57-88-5 LREGISTRY
 CN Cholest-5-en-3-ol (3.beta.)- (9CI) (CA INDEX NAME)
 OTHER CA INDEX NAMES:
 CN Cholesterol (8CI)
 OTHER NAMES:
 CN (-)-Cholesterol
 CN .DELTA.5-Cholesten-3.beta.-ol
 CN 3.beta.-Hydroxycholest-5-ene
 CN 5:6-Cholesten-3.beta.-ol
 CN Cholest-5-en-3.beta.-ol
 CN Cholesterin
 CN Cholesteryl alcohol
 CN Dythol
 CN Lidinit
 CN Lidinite
 CN Provitamin D
 FS STEREOSEARCH
 MF C27 H46 O
 CI COM
 LC STN Files: ANABSTR, BEILSTEIN*, BIOBUSINESS, BIOSIS, CA, CAOLD,
 CAPREVIEWS, CASREACT, CEN, CHEMINFORMRX, CHEMLIST, CBNB, CIN,
 CJACS, CSCHM, CSNB, DDR, DETHERM*, DRUGR, DRUGU, EMBASE, GMELIN*,
 HODOC*, IFICDB, IFIPAT, IFIUBB, IPA, MEDLINE, MRCK*, MSDS-OHS,
 MSDS-SUM, NAPRALERT, PDLCOM*, PIRA, PNI, PROMT, RTECS*, SPECINFO,
 TOXLINE, TOXLIT, USAN, VTB
 (*File contains numerically searchable property data)
 Other Sources: DSL**, EINECS**, TSCA**
 (**Enter CHEMLIST File for up-to-date regulatory information)
 DES 4:3B. CHOLEST



IDE (デフォルト) 形式での表示

RN 91386-77-5 LREGISTRY
 CN Interferon .alpha.1 (human leukocyte protein moiety reduced),
 1-L-serine- (9CI) (CA INDEX NAME)
 FS PROTEIN SEQUENCE
 MF Unspecified
 CI MAN
 LC STN Files: CA

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
 *** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***

SQIDE形式での表示 (タンパク質配列レコード)

RN 91386-77-5 LREGISTRY
 CN Interferon .alpha.1 (human leukocyte protein moiety reduced),
 1-L-serine- (9CI) (CA INDEX NAME)
 FS PROTEIN SEQUENCE
 SQL 166

SEQ 1 SDLPETHSLD NRRTLMLLAQ MSRISPSSCL MDRHDFGFPO EEFDGNQFQK
 51 APAISVLHEL IQQIFNLFTT KDSSAAWDED LLDKFCTELY QQLNDLEACV

SQIDE形式での表示 (タンパク質配列レコード) (続き)

101 MQEERVGETP LMNADSLAV KKYFRITLY LTEKKYSPCA WEVVRAEIMR
 151 SLSLSTNLQE RLRRKE
 MF Unspecified
 CI MAN
 LC STN Files: CA

SQIDE形式での表示 (GenBankからの核酸配列レコード)

RN 91449-61-5 LREGISTRY
 CN Deoxyribonucleic acid (Tikaut provirus 5'-long terminal repeat)
 (9CI) (CA INDEX NAME)
 FS NUCLEIC ACID SEQUENCE
 SQL 641
 NA 186 a 170 c 160 g 125 t
 NTE doublestranded

SEQ 1 tgaagacc caccataagg cttagcaagc tagctgcagt aagccattt
 51 tgcaaggcat gaaaaagtac cagagctgag ttctcaaagt caacaacgaa
 101 gtttagttaa agaataaggc tgaacaaaac tgggacagg gccaacagg
 151 atatctgtgg tcgagcagct agggccccgg ctcagggccca agaacagatg
 201 gtactcagat aaagcgaagg gctgaacaaa acgggacagg ggccaacag
 251 gatgggggcc aaacaggata tctgtggtcg agcacctggg ccccggtca
 301 gggccaagaa cagatggtac tcagataaag cgaactaac aacagtttct
 351 ggaaagtccc acctcagttt caagtcccc aaaagaccgg gaaaaacccc
 401 aagccttatt taaactaacc aatcagctcg cttctcgtt ctgtaacccg
 451 cgctttttgc tcccagccct ataaaaaggg taaaaacccc acactcggcg
 501 cccagtcct cggatagact gagtcgccc ggtaccgtg tatccaataa
 551 agccttttgc tgttgcattc gaatcgtgtt ctcgctgatc cttgggaggg
 601 tctcctcaga gtgattgact gccagcctg ggggtctttc a

MF Unspecified
 CI MAN
 LC STN Files: CA, TOXLIT

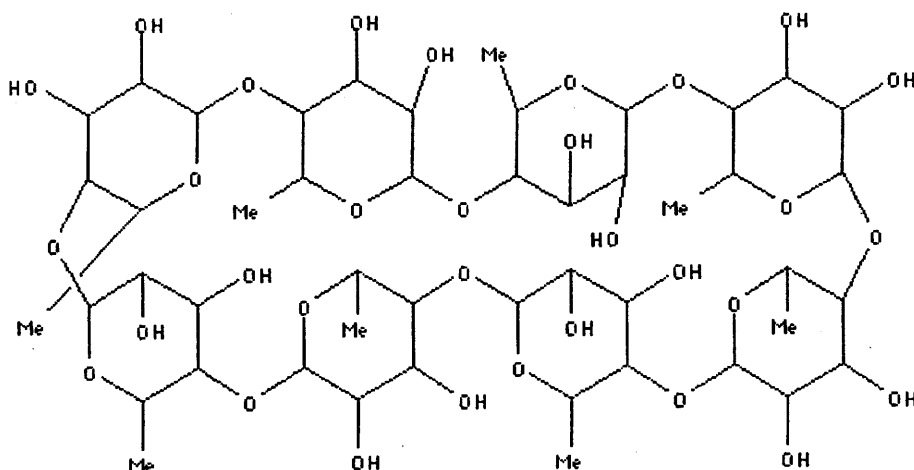
FIDE形式での表示

RN 53784-90-0 LREGISTRY
 CN .gamma.-Cyclodextrin, 6A, 6B, 6C, 6D, 6E, 6F, 6G, 6H-octadeoxy- (9CI) (CA
 INDEX NAME)
 OTHER CA INDEX NAMES:
 CN 2, 4, 7, 9, 12, 14, 17, 19, 22, 24, 27, 29, 32, 34, 37, 39-
 Hexadecaioxanonacyclo[36.2.2.2.3.6.28.11.213.16.218.21.223.26.228.31.2
 33,36]hexapentacontane, .gamma.-cyclodextrin deriv. (9CI)
 MF C48 H80 O32
 LC STN Files: BEILSTEIN*, CA
 (*File contains numerically searchable property data)
 DES 6:GAMMA-CYCLODEXTRIN

Ring System Data

| Elemental Analysis EA | Elemental Sequence ES | Size of the Rings SZ | Ring System Formula RF | Ring Identifier RID | RID Occurrence Count |
|--|---|----------------------------|------------------------------|---------------------------|----------------------------|
| C50-C50-C50- C50-C50-C50- C50-C50- C24016 | OC5-OC5-OC5- OC5-OC5-OC5- OC5-OC5- OCOC2OCOC2OCOC C2OCOC2OCOC2O COC2OCOC2OCOC 2 | 6-6-6-6-6-6- 6-6-40 | C40016 | 14246.1.1 | 1 |

FIDE形式での表示 (続き)



CCN形式での表示

CN Methanaminium, N-[4-[[4-(dimethylamino)phenyl]phenylmethylene]-2,5-cyclohexadien-1-ylidene]-N-methyl-, chloride (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN C. I. Basic Green 4 (8CI); Victoria Green WB (6CI)

OTHER NAMES:

CN Acryl Brilliant Green B; ADC Malachite Green Crystals; Aizen Malachite Green; Aizen Malachite Green Crystals; Aniline green; Astra Malachite Green; Astra Malachite Green B; Astra Malachite Green BXX; Atlantic Malachite Green; Basic Green 4; Basonyl Green 830; Benzal Green; Benzaldehyde green; Bronze Green Toner A 8002; Burma Green B; C. I. 42000; Calcozine Green V; China Green; Diabasic Malachite Green; Diamond Green B extra; Diamond Green BX; Diamond Green P Extra; Green MX; Grenoble Green; Hidaco Malachite Green Base; Hidaco Malachite Green LC; Hidaco Malachite Green SC; Light Green N; Lincoln Green Toner B 15-2900; Malachite green; Malachite Green A; Malachite Green AN; Malachite Green B; Malachite green chloride; Malachite Green CP; Malachite Green Crystals; Malachite Green Crystals BPC; Malachite Green J 3E; Malachite Green Powder; Malachite Green WS; Malachite Lake Green A; Mitsui Malachite Green; New Victoria Green Extra I; New Victoria Green Extra II; New Victoria Green Extra O; Oji Malachite Green; Solid Green Crystals O; Solid Green O; Super Ick Cure; Tertrophene Green M; Tokyo Aniline Malachite Green; Verona Basic Green M; Victoria Green; Victoria Green (basic dye); Victoria Green B; Victoria Green S; Victoria Green WPB