LREGISTRY[SM] (CAS REGISTRY[SM] Learning Database)

Subject Coverage

- All types of inorganic and organic substances, including alloys, coordination compounds, minerals, mixtures, polymers, salts, high throughput screening (HTS) compounds as well as nucleic acid and protein sequences
- Substances included in LREGISTRY meet the following criteria:
  - Identified by CAS as coming from a reputable source, including but not limited to patents, journals, chemical catalogs, and selected substance collections on the web
  - Described in largely unambiguous terms
  - Characterized by physical methods or described in a patent document example or claim
  - Consistent with the laws of atomic covalent organization

File Type

Numeric, Structure

Features

Alerts (SDIs)	Not available				
CAS Registry Number® Identifiers	☑	Keep & Share	☑	STN® AnaVist™	☐
Learning Database	☑	SLART	☑	STN Easy®	☐

Record Content

- CAS Registry Numbers
- CA index names and commonly used chemical names and trade names
- Molecular formulas
- Structure diagrams
- Sequence data
- Alloy composition tables
- Classes for polymers
- Ring analysis data
- List of databases and regulatory listings containing information on the substances

File Size

More than 124,578 records

Coverage

Selected from LCA and LCASREACT

Updates

None since it is a closed file.

Language

English

Database Producer

Chemical Abstracts Service
2540 Olentangy River Road
P.O. Box 3012
Columbus, Ohio 43210-0012 USA
Phone: 800-753-4227 (North America)
Phone: 614-447-3700 (worldwide)
Fax: 614-447-3751
Email: help@cas.org

Sources	CASRNs from LCA and LCASREACT
User Aids	• Online Helps (HELP DIRECTORY lists all help messages available) • STNGUIDE
Clusters	• LEARNING STN Database Clusters information (PDF).
Related Databases	REGISTRY
Pricing	Enter HELP COST at an arrow prompt.

SEARCH and DISPLAY Field Codes

Fields that allow left truncation (/CNS and /NTE) are marked with an asterisk (*).

Substance Data Fields

Search Field Name	Search Code	Search Examples	Display Codes
Basic Index (contains name fragments, molecular formula fragments, and Collective Index codes) **(1)**	None (or /BI)	S TOSYL S DIMETHYL ADIPATE S 6CI S 1,1(W)DICHLORO S C5H10BR2O2	AF, CN, IN, MF
CAS Registry Number	/RN	S 97-77-8/RN S 97-77-8	RN, AR, DR, PR
Class Identifier (codes or terms as a bound phrase)	/CI	S MXS/CI S ALLOY/CI	CI
Component Registry Number	/CRN	S 79-10-7/CRN	CRN
Definition	/DEF	S HYDROCARBONS/DEF	DEF
Deleted CAS Registry Number	/DR	S 50-83-9/DR	DR
Entry Date **(2)**	/ED	S 1985/ED	ED
Field Availability (codes or terms as a bound phrase)	/FA	S RSD/FA AND L5 S MATERIAL COMPOSITION/ FA	Not displayed
File Segment (acronyms or single words)	/FS	S PROTEIN/FS S PS/FS S NUCLEIC/FS	FS
Polymer Class Term (code or text)	/PCT	S POLYAMINE/PCT S PM/PCT	PCT
Registry Number Locator	/LC	S TSCA/LC S GENBANK/LC S L1 AND CA/LC	LC
Replacing CAS Registry Number	/RR	S 50-21-5/RR	RR
Source of Registration	/SR	S GENBANK/SR	SR
Update Date **(2)**	/UP	S UP>=20040101	Not displayed

(1) Formula fragments searched in the Basic Index must be entered without spaces.
(2) Numeric search field that may be searched using numeric operators or ranges.

Nomenclature Fields

Search Field Name	Search Code	Search Examples	Display Codes
Chemical Name	/CN	S 1-CHLORO-1, 3-BUTADIENE/CN S INTERFERON .ALPHA.1?/CN S GENBANK M12334/CN	CN, IN
Chemical Name Segment * **(1)**	/CNS	S IMINO/CNS S ?QUAT?/CNS NOT AQUA	CN, IN
Heading Parent	/HP	S BENZOIC ACID/HP	CN, IN
Index Name Segment Heading Parent	/INS.HP	S METHYLETHYL/INS.HP	CN, IN
Index Name Segment NonHeading Parent	/INS.NHP	S ACRYLO/INS.NHP	CN, IN
Other Name Segment	/ONS	S ANILINE/ONS	CN

(1) With left truncation, the input term must contain at least 4 characters.

Molecular Formula Fields

Search Field Name	Search Code	Search Examples	Display Codes
Atom Count **(1)**	/ATC	S 5/ATC	Not displayed
Element Count **(1)**	/ELC	S 7-9/ELC	Not displayed
Element Count for Substance **(1)**	/ELC.SUB	S ELC.SUB>=8	Not displayed
Element Formula **(2)**	/ELF	S AL CO LA O/ELF	AF, MF
Element Ratio, xx (where xx = CH, CN, CO, HC, HN, HO, NC, NH, NO, OC, OH, or ON) **(1)**	/ELR.xx	S 3.1666667/ELR.CH S 1-2/ELR.CN S ELR.CO<=1	Not displayed
Element Symbol	/ELS	S B/ELS AND H/ELS	Not displayed
Element Symbol for Multicomponent Formula	/ELS.MCF	S (N (XA) P)/ELS.MCF	Not displayed
Formula Weight **(1)**	/FW	S 420-460/FW	Not displayed
Material Composition **(3)**	/MAC	S 1-5 ND/MAC	STR
Molecular Formula **(4)**	/MF	S C7H3BR2FO2/MF S C4H4O4.2NA/MF S C24 H37 OS P3/MF	AF, MF
Number of Components **(1)**	/NC	S F/ELS NOT NC>=2	Not displayed
Periodic Group	/PG	S B6/PG S LNTH/PG	Not displayed
Relative Composition	/RC	S FE.CR.NI/RC	Not displayed
Specific Element Count **(1)**	/Element Symbol	S 7/SI	Not displayed

(1) Numeric search field that may be searched using numeric operators or ranges.
(2) Formulas must be entered with spaces between the elements.
(3) Combined numeric and text field. Composition terms are numeric and may be searched using numeric operators or ranges. Component terms are text terms.
(4) Formulas may be entered with or without spaces.

Limiting Search Codes

Search Field Name	Search Code	Search Examples	Display Code
Answers completely iterated	/COMPLETE **(1)**	S L4/COM **(2)**	Not displayed
Answers incompletely iterated	/INCOMPLETE **(1)**	S L4/INC **(2)**	Not displayed

(1) The code may be abbreviated to the first three letters.
(2) Only an L-number for an answer set created in LREGISTRY may be limited.

Ring Analysis Data Fields

Search Field Name	Search Code	Search Examples	Display Codes
Elemental Analysis for Ring System (and number of occurrences of EA in a component structure) **(1)**	/EA	S C4N-C5N/EA S 2 C3NO-C6/EA	RSD
Elemental Analysis for Smallest Ring (and number of occurrences of EAS in a ring system) **(1)**	/EAS	S C5NO4/EAS S >9 C6/EAS	Not displayed
Elemental Sequence for Ring System (and number of occurrences of ES in a component structure) **(1)**	/ES	S NCOC2-C6/ES S 1-3 O2C4/ES	RSD, SRSD
Elemental Sequence for Smallest Ring (and number of occurrences of ESS in a ring system) **(1)**	/ESS	S FE3/ESS S >=2 SC2SC2/ESS	Not displayed
Number of Ring Systems **(2)**	/NRS	S 7/NRS	Not displayed
Number of Ring Systems in a Component **(2)**	/CNRS	S 4-5/CNRS	Not displayed
Number of Rings (number of smallest rings) **(2)**	/NR	S 10/NR	Not displayed
Number of Rings in a Component (number of smallest rings) **(2)**	/CNR	S CNR>=12	Not displayed
Number of Rings in Ring System **(2)**	/NRRS	S 5-6/NRRS	Not displayed
Ring Atom Count **(2)**	/RATC	S 4/RATC	Not displayed
Ring Element (and number of occurrences of REL in a ring system) **(1)**	/REL	S SE/REL S 5 P/REL	Not displayed
Ring Element Count **(2)**	/RELC	S 6/RELC	Not displayed
Ring Elemental Formula (and number of occurrences of RELF in a component structure) **(1,3)**	/RELF	S C N O P/RELF S >3 C N O/RELF	Not displayed
Ring Identifier (and number of occurrences of RID in a component structure) **(1)**	/RID	S 31779.1.2/RID S 1938/RID S >=2 1949.52/RID	RSD, SRSD
Ring Size of Smallest Ring (and number of occurrences of SZS in a ring system) **(1,2)**	/SZS	S 8/SZS S 5 4/SZS	Not displayed
Ring System Formula (and number of occurrences of RF in a component structure) **(1)**	/RF	S C20AGN4/RF S 5 C10/RF	RSD
Size for the Ring System (and number of occurrences of SZ in a component structure) **(1)**	/SZ	S 3-4-5/SZ S 3 5-5-6/SZ	RSD

(1) The number of occurrences must be entered first in the search field. It is a numeric term and may be searched using numeric operators or ranges.
(2) Numeric search field that may be searched using numeric operators or ranges.
(3) Formulas must be entered with spaces between the elements.

Sequence Fields

Search Field Name	Search Code	Search Examples	Display Codes
Notes * **(1)**	/NTE	S CYCLIC/NTE S ?CHLORO?/NTE	NTE
Nucleic Acid Count **(2,3)**	/NA.CNT	S 12-42/NA.CNT	NA
Nucleic Acid Type **(3)**	/NA	S 12-42 A/NA	NA
Sequence Length	/SQL	S G/NA S SQL<=500	SQL

(1) With left truncation, the input term must contain at least 4 characters.
(2) Numeric search field that may be searched using numeric operators or ranges.
(3) Field contains data only for nucleic acid sequences.

Structure Search Terms

Terms(1)	Search Examples
L-numbers of structures built using the STRUCTURE command or uploaded from STN Express® (Boolean logic allowed between the L-numbers)	SEARCH L1 FAM SAM SEA L1 AND L2 SSS FUL
L-numbers of screen sets created using the SCREEN command (Boolean logic allowed between the L-numbers)	S L3 OR L4 SSS SAM
L-numbers of structures built using the STRUCTURE command or uploaded from STN Express combined with L-numbers of screen sets created using the SCREEN command (Boolean logic allowed between L-numbers)	S L1 AND L2 NOT L3

(1) The L-number answer set from a structure search may be combined with dictionary terms, e.g., S L3 AND TSCA/LC.

Types of Structure Searching

Type	Definition	Search Code	Search Examples
Substructure (default)	Search for substances that match the query. Substitution is allowed at all open positions. Additional components may be retrieved.	SSS	SEARCH L1 SSS FUL S L2 OR L3 SSS SAM S L7 SSS
Closed Substructure	Search for substances that match the query exactly. Substitution is allowed at positions opened by CONNECT. Additional components may be retrieved.	CSS	SEARCH L1 CSS FUL S L2 NOT L3 CSS S L4 OR L5 CSS RANGE
Family	Search for substances that match the query exactly. Additional components may be retrieved.	FAM	S L6 FAM SAM
Exact	Search for substances that match the query exactly.	EXA	SEA L5 EXA FUL

Scopes of Structure Searches

To create an L-number answer set containing candidate structures that have passed the screening step of your structure search, enter EXTEND on the search command line or enter SET EXTEND ON or SET EXTEND ON PERM at an arrow prompt (=>). For details, enter HELP SET EXTEND at an arrow prompt.

Type	Definition	Search Code	Search Examples
Sample **(1)** (default)	Search a fixed 5% of the file.	SAM	SEARCH L3 EXA SAM S L6 NOT L7 SSS SAM
Full	Search 100% of the file.	FUL	S L5 OR L8 SSS FUL
Range	Search a user-specified portion of the file.	RAN	S L4 RAN= (110507-58-9,) S L3 FAM RAN= (109784-14-7, 109904-92-9)
Subset Sample	Search a fixed sample of an answer set created by a search in LREGISTRY.	SUB SAM	S L7 CSS SUB=L5 SAM
Subset Range	Search a user-specified portion of answer set created by a search in LREGISTRY.	SUB RAN	S L3 SUB=L2 RAN=(,50-11-3)
Subset Full	Search 100% of an answer set created by a search in LREGISTRY.	SUB FUL	S L8 SUB=L6 FAM FUL

(1) EXTEND is not valid with SAMPLE.

Sequence Search Terms

Terms	Search Example
One-letter codes for common amino acids **(1)** Three-letter codes for common and uncommon amino acids **(1) (2)** Enclose codes or strings of codes in single quotes. Use dashes to separate codes in strings. Single letter codes for nucleic acids **(3)**	S LAGLL/SQSP S 'LEU-ALA-GLY-LEU-LEU'/SQSFP S F'HCY-STA'LF/SQSP S 'GLP'AGYSK/SQEP S 'CYS-ASN-THR-ALA'/SQEP S ATTTTTTTTTT/SQEN S AAGGTTACTA/SQSN

(1) Enter HELP AAC at an arrow prompt to display a table of the 1- and 3-letter codes for common amino acids.
(2) Enter HELP AAU at an arrow prompt to display a table of the 3-letter codes for uncommon amino acids.
(3) Enter HELP NUC at an arrow prompt to display a table of the codes for nucleic acids.

Types of Sequence Searches

Sequence data for protein and nucleic acid sequences are displayed in the SEQ field with 1-letter codes and the SEQ3 field with 3-letter codes for proteins only.

Type	Definition	Code	Examples
Sequence Exact, Protein	Search for sequences that match the query. The query must be completely defined.	/SQEP	S YADAIF/SQEP S 'CYS-ASN-THR-ALA'/SQEP
Sequence Exact Family, Protein	Search for sequences that match the query and those in which family-equivalent substitution of the query amino acids occur **(1)**.	/SQEFP	S YGGFL/SQEFP S 'TYR-GLY-GLY- PHE-LEU'/SQEFP
Subsequence, Protein	Search for exact answers plus sequences in which the query sequence is embedded. Variability symbols are allowed.	/SQSP	S LAGLL/SQSP S F'HCY-STA'LF/SQSP
Subsequence Family, Protein	Search for exact subsequences, and answers in which family-equivalent substitution of the query amino acids occurs **(1)**.	/SQSFP	S ATCXAWV/SQSFP S 'LEU-ALA-GLY-LEU-LEU'/SQSFP
Sequence Exact, Nucleic Acid	Search for sequences that match the query. Ambiguity codes for nucleic acids are allowed.	/SQEN	S ATTTTTTTTTT/SQEN
Subsequence, Nucleic Acid	Search for exact answers, plus sequences in which the query sequence is embedded. Ambiguity codes for nucleic acids and variability symbols are allowed.	/SQSN	S AAGGTTACTA/SQSN

(1) The families of amino acid equivalents retrieved in protein family searches are:

P, A, G, S, T	(neutral, weakly hydrophobic)
Q, N, E, D, B, Z	(hydrophilic, acid amine)
H, K, R	(hydrophilic, basic)
L, I, V, M	(hydrophobic)
F, Y, W	(hydrophobic, aromatic)
C	(cross-link forming)

Variability Symbols for Subsequence Searches (/SQSP, /SQSFP, and /SQSN)(1,2)

Symbol	Function	Search Examples
[ ]	To specify alternate residues	S LGP[VL]/SQSP S LGP['VAL''LEU']/SQSP
[-]	To exclude a specific residue or alternate residues	S LGP[-H]/SQSP S LGP[-'HIS']/SQSPSP S LGP[-HL]/SQSP
{m}	To repeat the preceding sequence or sequence query (L#, E#, or saved query) m times	S (FL){2}/SQSP S L4{2}/SQSP S NAME/Q{3}/SQSP S (CTG){2}/SQSN S TAA(TAAA){2}/SQSN
{m,u} or {m-u}	To repeat the preceding sequence or sequence query (L#, E#, or saved query) m to u times	S GG(FL){1,2}/SQSP S L3{1,3}/SQSP S NAME/Q{1,4}/SQSP S (CTG){1,3}/SQSN
? or {0,1} or {0-1}	To repeat the preceding sequence or sequence query (L#, E#, or saved query) zero or one time	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
* or {0,} or {0-}	To repeat the preceding sequence or sequence query (L#, E#, or saved query) zero or more times	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
+ or {1,} or {1-}	To repeat the preceding sequence or sequence query (L#, E#, or saved query) one or more times	S KLK(DLE){1,}/SQSP S KLK(DLE)+/SQSP S L2{1-}/SQSP S NAME/Q{1,}/SQSP S CAT(CTG){1,}TATT/SQSN
&	To join together sequence expressions or queries (L#s, E#s, or saved queries)	S L1&L3/SQSFP S L2&L5{1,3}/SQSP S NAME1/Q{2}&NAME2/Q/SQSP S E1&E3/SQSP

(1) For more information on specifying variability in subsequence queries, enter HELP SQQ at an arrow prompt.

(2) In addition, the caret and the vertical bar may be used. The caret is used at the beginning or end of a sequence to search for that sequence at the beginning or end of a sequence field. The vertical bar is the symbol for alternation, i.e., it is used to separate alternate sequence queries.

Specifying Gaps in Subsequence Searches (/SQSP, /SQSFP, and /SQSN)

Symbol	Function	Search Examples
.	A gap of one residue	S SY.RPG/SQSP S SY..RPG/SQSPS S AAG...TGC/SQSN
.{m} or [m.]	A gap of m residues	S SY.{2}RPG/SQSP S SY[2.]RPG/SQSP
.{m,u} or .{m-u}	Gap of m to u residues	S GFF.{2,10}LSS/SQSP S GFF.{2-10}LSS/SQSP S AAG.{2,5}TGC/SQSN
: or .? or .{0,1} or .{0-1}	Gap of zero or one residues	S AGA:SRI/SQSFPS S AGA.?SRI/SQSFP S AGA.{0,1}SRI/SQSFP S AGA.{0-1}SRI/SQSFP
.* or .{0,} or .{0-}	Gap of zero or more residue	S HLC.*TYG/SQSP S HLC.{0,}TYG/SQSP S HLC.{0-}TYG/SQSP S AAGGCAGATG.*GCAA/SQSN
.+ or .{1,} or .{1-}	A gap of one or more residues	S SY.+TH/SQSFP S SY.{1,}TH/SQSFP S SY.{1-}TH/SQSFP S TCCTG.+GTGG/SQSN

CAS Registry BLAST® Similarity Searching

Similarity searching of peptides and nucleotides in LREGISTRY using the BLAST® (Basic Local Alignment Search Tool) algorithm is also available only for commercial accounts via STN® on the Web℠ or using STN Express 8.3 or higher for Windows®.

DISPLAY and PRINT Formats

Individual substance information fields may not be combined with substance predefined formats, e.g., D IDE RSD is not a valid request.

Multiple codes must be separated by commas or spaces. The fields are displayed or printed in the order requested.

Highlighting must be ON during SEARCH in order to use the HIT and KWIC formats.

The CM (Component Number) field appears in records for multicomponent substances, but it is not a custom display field and cannot be used in display or print requests.

Format	Content	Examples
AF	Alternate Molecular Formula	D L4 1-4 AF
AR	Alternate Registry Number	D L1 3 AR
CCI	Component Class Identifier	D CCI 1,3-5
CCN **(2)**	Condensed Chemical Name	D 20 CCN
CI	Substance Class Identifier	D 1-3,7,8 CI
CIL	Component Isotope at Unknown Location	D CIL
CMF	Component Molecular Formula	D L1 CMF 3
CN	Chemical Name	D CN
COMP**(3)**	Composition	D L7
CRN	Component Registry Number	D 1,3,6 CRN L5
DEF	Definition	D DEF
DR	Deleted CAS Registry Number	D L8 DR 1-3
ED	Entry Date	D ED
FCN **(2)**	Full Chemical Name	D FCN L3 7
FS	File Segment	D 1,4 FS
IL	Isotope at Unknown Location	D IL
IN	CA Index Name	D IN L1 4
LC	Registry Number Locator	D LC 3,4
MF	Molecular Formula	D MF
PCT	Polymer Class Term	D L3 PCT
PR	Preferred Registry Number	D 5,3 PR
REF	Number of references in CA, CAplus	D REF
RN	CAS Registry Number	D L4 RN 3
RR	Replacing CAS Registry Number	D L3 2 RR
RSD **(4)**	Ring System Data	D RSD
SCN **(5)**	Short Chemical Name	D 5-9 SCN
SR	Source of Registration	D SR 1,3 L12
SRSD **(6)**	Short Ring System Data	D SRSD
STF **(7)**	Flat Structure (no stereo indicated)	D L9 1 3
STR **(8)**	Structure Diagram (includes stereo bonds and R/S/E/Z labels when available)	D L4 STR
STS **(7,8)**	Stereo Structure (includes stereo bonds when available)	D STS
NA	Nucleic Acid	D 6 9 11 NA
NTE	Note	D NTE
PNTE	Patent Annotation	D PNTE
SEQ	Sequence (1-letter codes)	D SEQ

DISPLAY and PRINT Formats (cont'd)

Format	Content	Examples
SEQ3	Sequence (3-letter codes)	D SEQ3 1-10
SQD	RN, AR, PR, DR, RR, FS, SQL, NA, NTE, PNTE, SEQ	D 5 SQD
SQD3	RN, AR, PR, DR, RR, FS, SQL, NA, NTE, PNTE, SEQ3	D 2-4 SQD3
SQIDE	RN, CN, DEF, AR, PR, DR, RR, FS, SQL, NA, NTE, PNTE, SEQ, MF, AF, CI, PCT, SR, LC, IL, STR, REF	D L4 SQIDE
SQIDE3	Same as SQIDE except that 3-letter codes are used for protein sequences	D L4 SQIDE3
SQL	Sequence Length	D L3 SQL
SQN	RN, CN, AR, PR, FS, SQL, DR, RR, REF	D SQN L5 6-9
ALL	All available substance information	DISPLAY L1 1 ALL
FIDE	All substance information including all names and RSD	D FIDE
IDE	Up to 50 names and other substance data except for RSD (IDE is the default)	D IDE
REG	CAS Registry Numbers (RN, DR, AR, PR, RR)	D REG
SAM	IN, SQL, MF, CI, STR, COMP	D L3 1-18 SAM
SCAN **(9,10)**	IN, SQL, MF, CI, STR, COMP (random display without answer numbers)	D SCAN
HIT **(1)**	Fields containing hit terms	D HIT 5-10
KWIC **(1)**	Hit terms plus 20 words on either side (KeyWord in Context)	D KWIC 5-10

(1) HIT and KWIC are available for all dictionary fields except MAC, RC, and CRN, and in all biosequence fields. KWIC is the same as HIT for all fields except DEF and LC. The entire field containing hit terms is highlighted except for DEF and LC in which the individual terms are highlighted. The entire RSD table is displayed without highlighting. For NTE, row(s) of the table containing the hit terms is displayed without highlighting. For SEQ and SEQ3, the amino acid codes causing the hit to be highlighted by underlining and also by a statement of their position in the sequence.
(2) Names are displayed with CN code. This is a custom display only.
(3) This is a tabular display that lists composition information and Component Registry Numbers for alloys and tabular inorganic substances.
(4) This is a tabular display that lists EA, ES, SZ, RF, RID, and RID Occurrence Count.
(5) The CA Index Name and all OTHER NAMES are displayed with CN code. This is a custom display only.
(6) This is a tabular display that lists EA, RID, and RID Occurrence Count.
(7) Custom display format.
(8) Stereo structure diagrams are available only on graphics terminals and offline prints.
(9) No online display charge for this option.
(10) SCAN must be specified on the command line, i.e., D SCAN or DISPLAY SCAN.

SELECT, ANALYZE, and SORT Fields

The SELECT command is used to create E-numbers or an L-number containing terms taken from the specified field in an answer set.

The ANALYZE command is used to create an L-number containing terms taken from the specified field in an answer set.

The SORT command is used to rearrange the search results in either alphabetic or numeric order of the specified field(s).

FIELD NAME	FIELD CODE	ANALYZE SELECT (1)	SORT
Alternate Molecular Formula	AF	Y **(2)**	N
Alternate Registry Number	AR	Y **(3)**	N
CA Index Name	IN	Y **(4)**	Y
CAS Registry Number	RN	Y **(3)**	Y
Chemical Name	CN	Y **(5)**	N
Class Identifier	CI	Y	N
Component Class Identifier	CCI	Y **(6)**	N
Component Molecular Formula	CMF	Y **(3)**	N
Component Registry Number	CRN	Y	N
Definition	DEF	Y	N
Deleted CAS Registry Number	DR	Y **(3)**	N
Elemental Analysis for Ring System	EA	Y	N
Elemental Sequence for Ring System	ES	Y	N
Entry Date	ED	Y	Y
File Segment	FS	Y	Y
Full Chemical Name	FCN	Y **(4)**	N
Molecular Formula	MF	Y	N
Names	NAME	Y **(8)**	N
Nucleic Acid Sequence (exact search form)	SQEN	Y	N
Nucleic Acid Sequence (subsequence search form)	SQSN	Y	N
Polymer Class Term	PCT	Y	N
Preferred Registry Number	PR	Y **(3)**	N
Protein Sequence (exact family search form)	SQEFP	Y	N
Protein Sequence (exact search form)	SQEP	Y	N
Protein Sequence (subsequence family search form)	SQSFP	Y	N
Protein Sequence (subsequence search form)	SQSP	Y	N
Registry Number Locator	LC	Y **(10)**	N
Registry Numbers and Names	CHEM	Y **(11)** (default)	N
Replacing CAS Registry Number	RR	Y **(5)**	N
Ring Identifier	RID	Y	N
Ring System Formula	RF	Y	N
Sequence (1-letter codes)	SEQ	Y **(9)**	N
Sequence (3-letter codes)	SEQ3	Y **(9)**	N
Sequence Length	SQL	N	Y
Short Chemical Names	SCN	Y **(4)**	N
Size for the Ring System	SZ	Y	N
Source of Registration	SR	Y	N

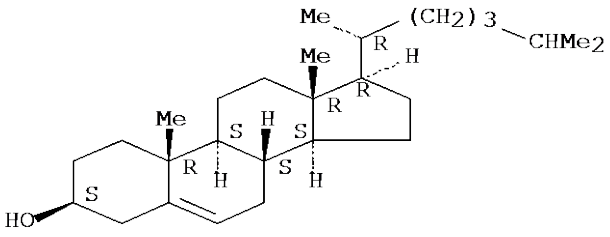
(1) HIT may be used to restrict terms extracted to terms that match the search expression used to create the answer set, e.g., SEL HIT CN.
(2) /MF is appended to the terms created by SELECT.
(3) /BI is appended to the terms created by SELECT.
(4) /CN is appended to the terms created by SELECT.
(5) CA Index Name, first 50 names in alphabetical order, and any additional hit names are extracted.
(6) /CI is appended to the terms created by SELECT.
(7) /FA is appended to the terms created by SELECT.
(8) All names except inverted names are extracted and /BI is appended to the terms created by SELECT. For nucleic acids from the GenBank database, NAME extracts GenBank Locus ID and GenBank numbers. GenBank numbers may be used as search terms in the GenBank database or other STN databases such as MEDLINE®.
(9) /SQSP is appended to the terms created by SELECT.
(10) E-numbers containing the files listed in this field may be used in the FILE and INDEX commands in place of the file names.
(11) AR, DR, PR, RN, RR, and all names except inverted names are extracted and /BI is appended to the terms created by SELECT.

Sample Records

DISPLAY IDE

```
RN   57-88-5  LREGISTRY
CN   Cholest-5-en-3-ol (3β)-  (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN   Cholesterol (8CI)
OTHER NAMES:
CN   (-)-Cholesterol
CN   Δ5-Cholesten-3β-ol
CN   3β-Hydroxycholest-5-ene
CN   5:6-Cholesten-3β-ol
CN   Cholest-5-en-3β-ol
CN   Cholesterin
CN   Cholesteryl alcohol
CN   Dythol
CN   Lidinit
CN   Lidinite
CN   NSC 8798
CN   Provitamin D
CN   SyntheChol
FS   STEREOSEARCH
DR   849593-11-9, 856708-55-9, 732297-95-9, 793670-51-6, 80356-14-5,
     80356-33-8, 209124-38-9, 218965-24-3, 262418-13-3, 378185-03-6,
     676322-57-9
MF   C27 H46 O
CI   COM
LC   STN Files:   ADISNEWS, AGRICOLA, ANABSTR, BIOSIS, BIOTECHNO, CA, CABA,
       CAPLUS, CASREACT, CBNB, CHEMCATS, CHEMINFORMRX, CHEMLIST, CIN, CSNB,
       DDFU, DETHERM*, DRUGU, EMBASE, IFICDB, IFIPAT, IFIUDB, IPA, MEDLINE,
       MRCK*, MSDS-OHS, NAPRALERT, PIRA, REAXYSFILE*, RTECS*, SPECINFO,
       TOXCENTER, ULIDAT, USAN, USPAT2, USPATFULL, VETU
         (*File contains numerically searchable property data)
     Other Sources:   DSL**, EINECS**, TSCA**
         (**Enter CHEMLIST File for up-to-date regulatory information)

Absolute stereochemistry.
```



LREGISTRY

DISPLAY CCN

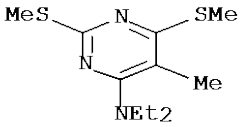
```
CN   Methanamine, conjugate acid (9CI)  (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN   Methylamine, conjugate acid (8CI);  Methylamine, conjugate acid of (7CI)
OTHER NAMES:
CN   Methanaminium;  Methylammonium;  Methylammonium cation;
     Methylammonium ion;  Monomethylammonium;  Monoprotonated methylamine;
     Protonated methylamine
```

DISPLAY ALL – RING SYSTEM DATA

```
RN   117663-66-8  LREGISTRY
CN   4-Pyrimidinamine, N,N-diethyl-5-methyl-2,6-bis(methylthio)-  (CA INDEX
     NAME)
MF   C11 H19 N3 S2
SR   CA
LC   STN Files:   CA, CAPLUS, CASREACT, REAXYSFILE*
         (*File contains numerically searchable property data)

Ring System Data

Elemental|Elemental| Size of |Ring System|   Ring   |   RID
Analysis |Sequence |the Rings|  Formula  |Identifier|Occurrence
   EA    |   ES    |   SZ    |    RF     |   RID    |  Count
=========+=========+=========+===========+==========+==========
C4N2     |NCNC3    |6        |C4N2       |46.195.39 |1
```



DISPLAY SQIDE (Nucleic Acid Sequence Indexed by CAS)

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

SEQ      1 tgaaagaccc caccataagg cttagcaagc tagctgcagt aacgccattt
        51 tgcaaggcat gaaaaagtac cagagctgag ttctcaaagt caacaacgaa
       101 gtttagttaa agaataaggc tgaacaaaac tgggacaggg gccaaacagg
       151 atatctgtgg tcgagcagct agggccccgg ctcagggcca agaacagatg
       201 gtactcagat aaagcgaagg gctgaacaaa acgggacagg ggccaaacag
       251 gatgggggcc aaacaggata tctgtggtcg agcacctggg ccccggctca
       301 gggccaagaa cagatggtac tcagataaag cgaaactaac aacagtttct
       351 ggaaagtccc acctcagttt caagttcccc aaaagaccgg gaaaaacccc
       401 aagccttatt taaactaacc aatcagctcg cttctcgctt ctgtaacccg
       451 cgctttttgc tcccagccct ataaaaaggg taaaaacccc acactcggcg
       501 ccccagtcct ccgatagact gagtcgcccg ggtacccgtg tatccaataa
       551 agccttttgc tgttgcatcc gaatcgtggt ctcgctgatc cttgggaggg
       601 tctcctcaga gtgattgact gcccagcctg ggggtctttc a
MF   Unspecified
CI   MAN
LC   STN Files:   CA, CAPLUS
```

DISPLAY SQIDE (Nucleic Acid Sequence Registered from GenBank®)

```
RN   91386-53-7  LREGISTRY
CN   DNA (Friend mink cell focus-inducing virus clone Bp-1 3'-long terminal
     repeat) (9CI)  (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN   Deoxyribonucleic acid (Friend mink cell focus-inducing provirus clone Bp-1
     3'-long terminal repeat)
OTHER NAMES:
CN   GenBank K01385
CN   GenBank M16688 (Secondary GenBank Accession Number)
FS   NUCLEIC ACID SEQUENCE
SQL  550
NA   132 a   164 c   134 g   120 t

SEQ      1 tgaaagaccc caccaagttg cttagcctga tagccgcagt aacgccattt
        51 tgcaaggcat ggaaaaatac caaaccaaga atagagaagt tcagatcaag
       101 ggcgggtaca cgaaaacagc taacgttggg ccaaacagga tatctgcggt
       151 aagcagtttc ggcccccggcc cggggaagaa cagatggtca ccgcagttcg
       201 gccccggccc ggggaagaac agatggtccc cagatatggc ccaaccctca
       251 gcagtttctt aagacccatc agatgtttct aggctccccc aaggacctga
       301 aatgaccctg tgccttattt gaattaacca atcagcctgc ttctcgcttc
       351 tgttcgcgcg cttctgcttc ccgagctcta taaaagagct cacaacccct
       401 cactcggcgc gccagtcctc cgacagactg agtcgcccgg gtacccgagt
       451 atccaataaa tcctcttgct gttgcatccg actcgtggtc tcgctgttcc
       501 ttgggagggt ctcctcagag tgattgacta cccgtctcgg gggtctttca
MF   Unspecified
CI   MAN
LC   STN Files:   CA, CAPLUS
```

In North America
CAS
STN North America
P.O. Box 3012
Columbus, Ohio 43210-0012 U.S.A.

CAS Customer Center:
Phone: 800-753-4227 (North America)
614-447-3700 (worldwide)
Fax: 614-447-3751
Email: help@cas.org
Internet: www.cas.org

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

In Japan
JAICI (Japan Association for
International Chemical Information)
STN Japan
Nakai Building
6-25-4 Honkomagome, Bunkyo-ku
Tokyo 113-0021, Japan
Phone: +81-3-5978-3601 (Technical Service)
+81-3-5978-3621 (Customer Service)
Fax: +81-3-5978-3600
Email: support@jaici.or.jp (Technical Service)
customer@jaici.or.jp (Customer Service)
Internet: www.jaici.or.jp