REGISTRY/ZREGISTRY (CAS REGISTRY[SM])

THE CHOICE OF PATENT EXPERTS™

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 REGISTRY 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
- Experimental and predicted property data and tags and spectra data

File Type

Numeric, Structure

Features

Alerts (SDIs)	Biweekly In addition, SMARTracker, an automatic crossfile current-awareness search, (SDI XFILE) using a REGISTRY search profile in CA, HCA, ZCA, CAplus, HCAplus or ZCAplus may be run weekly or biweekly (weekly is the default).				
CAS Registry Number® Identifiers	☑	Keep & Share	☑	SLART	☑
Learning Database	☑	STN Easy®	☑	Structures	☑

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
- Experimental and predicted property data and tags
- Spectra
- List of databases and regulatory listings containing information on the substances
- Super roles and document type information from CAplus[SM]
- Displayable information for 10 most recent references for substances

File Size

More than 140 million organic and inorganic substances; more than 71.1 million sequences (4/18)

Coverage

Early 1800s to the present

Updates

Daily

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	• The CAS Registry System - the computer system that automatically identifies and catalogs all substance information that CAS selects from journal articles, patents, conference proceedings, and other published sources • GenBank® (registered trademark of the U.S. Department of Health and Human Services) - a database repository containing millions of sequences • Special collections of substance data such as substance information listed on various international, national, and state regulatory inventories and lists as well as reputable web collections of substances to which CAS has assigned CAS Registry Numbers
User Aids	• Online Helps (HELP DIRECTORY lists all help messages available) • STNGUIDE
Clusters	• CASLINK • CASRNS • HCASLINK • NUMERIC • STRUCTURE STN Database Clusters information (PDF).
Related Databases	LREGISTRY
Pricing	Enter HELP COST at an arrow prompt.

SEARCH and DISPLAY Field Codes

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

You can also search any REGISTRY search term including structures directly in CAplus using REG1stRY. To search a REGISTRY term in CAplus, enter the SEARCH command and your term followed by the REGISTRY field code, followed by /REG, e.g., SEARCH FENFLURAMINE/CN/REG. The REGISTRY search and crossover to CAplus are performed automatically, and only the final CAplus answer set L-number is shown. Enter HELP FIRST at an arrow prompt in CAplus for more information.

The POLYLINK command is available for more comprehensive searching of condensation polymers. Enter HELP POLYLINK at an arrow prompt for online information.

The SEQLINK command is available for more comprehensive, CAS Registry Number-based searching of protein and nucleic acid sequences. Enter HELP SEQLINK at an arrow prompt for online information.

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
Component Class Identifier(codes or terms as a bound phrase)	/CCI	S MXS/CCI S CCS/CCI	CCI
Class Identifier (codes or terms as a bound phrase)	/CI	S MXS/CI S ALLOY/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 20040101/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
Number of References in the CA File **(2)**	/REF.CA	S L1 AND REF.CA<=10	REF
Number of References in the CA File for Non-Specific Derivatives **(2)**	/REF.CAD	S L3 AND 1/REF.CAD	REF
Number of References in the CAplus File **(2)**	/REF.CAPLUS	S L2 NOT REF.CAPLUS>10	REF
Polymer Class Term (code or text)	/PCT	S POLYAMINE/PCT S PM/PCT	PCT
Polymer Class Term Count **(2)**	/PCT.CNT	S 2-3/PCT.CNT	PCT
Registry Number Locator	/LC	S TSCA/LC S GENBANK/LC S L1 AND CA/LC	LC
Replacing CAS Registry Number	/RR	S 50-01-1/RR	RR
Source of Registration	/SR	S CHEMICAL LIBRARY/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
Editor Note * **(1)**	/ENTE	S ?ENZYM?/ENTE	ENTE
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.

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
Patent Annotation **(4)**	/PNTE (/FEAT)	S RATTUS/PNTE	PNTE
Patent Country **(4)**	/PC	S US/PC	PNTE
Patent Number **(4)**	/PN	S WO2000056771/PN	PNTE
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.
(4) Field contains data only for sequences indexed from patents since October 1999.

Property Search Fields (1)

Search Field Name	Default Units	Search Code	Search Examples	Display Codes
Bioconcentration Factor **(1)**	none	/BCF	S 4000-5000/BCF	BCF
Bioconcentration Factor pH Values **(1)**	none	/BCF.PH	S 4000-5000/PCF(P)7/BCF.PH	BCF
Bioconcentration Factor Temperature **(1)**	deg C	/BCF.T	S 25/BCF.T	BCF
Boiling Point **(1)**	deg C	/BP	S 150-155/BP	BP
Boiling Point Pressure **(1)**	Torr	/BP.P	S 166/BP(P)3/BP.P	BP
Density **(1)**	g/cm**3	/DEN	S DEN>=1.002	DEN
Density Pressure **(1)**	Torr	/DEN.P	S 800/DEN.P	DEN
Density Temperature **(1)**	deg C	/DEN.T	S 1.0021-1.02/DEN (P) 20/DEN.T	DEN
Electric Conductance **(1)**	Siemens	/ECON	S 377 SIEMENS/ECON	ECON
Electric Conductance Temperature **(1)**	deg C	/ECON.T	S 377/ECON (P) 25/ECON.T	ECON
Electric Conductivity **(1)**	S/cm	/ECND	S 1400-1900/ECND	ECND
Electric Conductivity Temperature **(1)**	deg C	/ECND.T	S 1400-1900/ECND (P) 1000/ECND.T	ECND
Electric Resistance **(1)**	ohm	/ERES	S 30-70/ERES	ERES
Electric Resistance Temperature **(1)**	deg C	/ERES.T	S 900-960/ERES (P) 235/ERES.T	ERES
Electric Resistivity **(1)**	ohm*cm	/EREST	S EREST>=6600	EREST
Electric Resistivity Temperature **(1)**	deg C	/EREST.T	S EREST>=6600 (P) 25/EREST.T	EREST
Enthalpy of Vaporization **(1)**	kJ/mol	/HVAP	S 100-110/HVAP	HVAP
Enthalpy of Vaporization Pressure **(1)**	Torr	/HVAP.P	S 760/HVAP.P	HVAP
Experimental Properties (includes content of /ETAG and /FA) **(2)**	none	/EPROPS	S TENSILE STRENGTH/EPROPS	EPROP, ETAG
Experimental Property Tags **(2,3)**	None	/ETAG	S MASS SPECTRA/ETAG	ETAG
Field Availability **(2)**	none	/FA	S HD/FA	FA
Field Not Available **(2)**	none	/FNA	S PYRAZOL? AND HD/FNA	Not displayed
Flash Point **(1)**	deg C	/FP	S FP<250	FP
Freely Rotatable Bonds **(1,4)**	none	/FRB	S 2-5/FRB	FRB
Glass Transition Temperature **(1)**	deg C	/TG	S 7-8/TG	TG

Property Search Fields (1) (cont'd)

Search Field Name	Default Units	Search Code	Search Examples	Display Codes
Hydrogen Acceptors **(1,4)**	none	/HAC	S 1-3/HAC	HAC
Hydrogen Donors **(1,4)**	none	/HD	S HD<=5	HD
Hydrogen Donor/Acceptor Sum **(1)**	none	/HDAS	S 12/HDAS	HDAS
Koc (Organic Carbon Adsorption Constant) **(1)**	none	/KOC	S 100-200/KOC	KOC
Koc pH Values **(1)**	none	/KOC.PH	S 100-200/KOC(P)7/KOC.PH	KOC
Koc Temperature **(1)**	deg C	/KOC.T	S 25/KOC.T	KOC
LogD **(1)**	none	/LOGD	S 2.21/LOGD	LOGD
LogD pH Values **(1)**	none	/LOGD.PH	S 2.21/LOGD (P) 10/LOGD.PH	LOGD
LogD Temperature **(1)**	deg C	/LOGD.T	S 25/LOGD.T	LOGD
LogP **(1,4)**	none	/LOGP	S LOGP<=3	LOGP
LogP Temperature **(1)**	deg C	/LOGP.T	S 25/LOGP.T	LOGP
Magnetic Moment **(1)**	uB	/MM	S MM<=0.98	MM
Magnetic Moment Temperature **(1)**	K	/MM.T	S 0.021/MM (P) 10/MM.T	MM
Mass Intrinsic Solubility **(1)**	g/L	/ISLB.MASS	S 1.3/ISLB.MASS	ISLB.MASS
Mass Solubility **(1)**	g/L	/SLB.MASS	S 1.4/SLB.MASS	SLB.MASS
Mass Solubility pH **(1)**	none	/SLB.PH	S 0.17/SLB.PH	SLB.MASS
Mass Solubility Temperature **(1)**	deg C	/SLB.T	S 0.17/SLB.PH (P) 25/SLB.T	SLB.MASS
Median Lethal Dose **(1)**	mg/kg	/LD50	S 741-745/LD50	LD50
Median Lethal Dose Organism	none	/LD50.ORGN	S 741-745/LD50 (P) MOUSE/LD50.ORGN	LD50
Median Lethal Dose Route of Administration	none	/LD50.RTE	S 450-520/LD50 (P) ORAL/LD50.RTE	LD50
Melting Point **(1,3)**	deg C	/MP	S MP<=30	MP
Melting Point Pressure **(1)**	Torr	/MP.P	S 70/MP(P)2/MP.P	MP
Melting Point Solvent	none	/MP.SOL	S ACETIC ACID/MP.SOL	MP
Molar Intrinsic Solubility **(1)**	mol/L	/ISLB.MOL	S 1.6E-14/ISLB.MOL	ISLB.MOL
Molar Solubility **(1)**	mol/L	/SLB.MOL	S SLB.MOL>=6	SLB.MOL
Molar Solubility pH Values **(1)**	none	/SLB.PH	S SLB.MOL>=6 (P) 7-10/SLB.PH	SLB.MOL
Molar Solubility Temperature **(1)**	deg C	/SLB.T	S SLB.MOL>=6 (P) 25/SLB.T	SLB.MOL
Molar Volume **(1)**	cm**3/mol	/MVOL	S 31.1/MVOL	MVOL
Molar Volume Temperature **(1)**	deg C	/MVOL.T	S 25/MVOL.T	MVOL
Molar Volume Pressure **(1)**	Torr	/MVOL.P	S 760/MVOL.P	MVOL
Molecular Weight **(1,4)**	none	/MW	S MW<200	MW
Optical Rotatory Power **(1)**	deg	/ORP	S 70-80/ORP	ORP
Optical Rotatory Power Concentration **(1)**	g/100mL	/ORP.C	S 0.12/ORP.C	ORP
Optical Rotatory Power Pathlength **(1)**	dm	/ORP.LEN	S 43/ORP.LEN	ORP
Optical Rotatory Power Solvent	none	/ORP.SOL	S METHANOL/ORP.SOL	ORP
Optical Rotatory Power Temperature **(1)**	deg C	/ORP.T	S 70-80/ORP(P)20/ORP.T	ORP
Optical Rotatory Power Wavelength **(1)**	nm	/ORP.W	S 546/ORP.W	ORP
pKa **(1)**	none	/PKA	S PKA<=-0.62	PKA
pKa Temperature **(1)**	deg C	/PKA.T	S 25/PKA.T	PKA
pKa Type	none	/PKA.TYP	S PKA<=0.52 (P) MOST ACIDIC/PKA.TYP	PKA

Property Search Fields (1) (cont'd)

Search Field Name	Default Units	Search Code	Search Examples	Display Codes
Polar Surface Area	A**2 (Angstrom**2)	/PSA	S 3.24/PSA	PSA
Property Note **(4)**	none	/PNT	S MP<=30 (P)DECOMP/PNT S MP=110(P)EXACT/PNT	EPROP
Property Source Note **(2)**	none	/PSO	S L1 NOT ACD/PSO S L1 AND CAS/PSO	PROP
Property Type **(2)**	none	/PTYP	S L1 AND PREDICTED/PTYP S EXPERIMENTAL/PTYP	PROP
Reference Accession Number in CA **(2)**	none	/RAN.CA	S 90:102811/RAN.CA	EPROP
Refractive Index **(1)**	none	/RI	S 1.427/RI	RI
Refractive Index Temperature **(1)**	deg C	/RI.T	S 1.427/RI(P)25/RI.T	RI
Refractive Index Wavelength	nm	/RI.W	S 500-589.3/RI.W	RI
Spectra	none	/SPEC	S IR/SPEC S NMR SPECTRA/SPEC S MASS SPECTRA/SPEC S PROTON SPECTRA/SPEC	SPEC
Tensile Strength	MPa	/TS	S 42/TS	TS
Tensile Strength Temperature	deg C	/TS.T	S 200-315/TS (P) 190/TS.T	TS
Uncertainty Range	none	/UR	S BP=200(P)UR<=10	EPROP
Vapor Pressure **(1)**	Torr	/VP	S .0001-.0002/VP	VP
Vapor Pressure Temperature **(1)**	deg C	/VP.T	S .0001-.0002/VP(P)25/VP.T	VP

(1) Numeric search field that may be searched using numeric operators or ranges.
(2) Field containing text terms which are not searchable with numeric operators or ranges.
(3) For a list of search terms, refer to REGISTRY: Tagged Experimental Properties at: http://www.cas.org
(4) The Property Note (/PNT) field contains property-related note terms such as COMP, SUBLM, and POLYMORPH when this information is reported. The /PNT field also contains the terms EXACT, CLOSED RANGE, OPEN RANGE. To limit your retrievals to EXACT, CLOSED RANGE, or OPEN RANGE values, combine your numeric query with one of these terms using the (P) operator.

CAplus Super Roles and Document Type Search Fields (1)

Search Field Name	Search Code	Search Examples	Display Codes
Document type	/DT.CA	S JOURNAL/DT.CA	DT.CA
Super roles for specific substances	/RL	S RACT/RL	RL
Super roles for non-specific derivatives	/RLD	S RACT/RLD	RLD
Super roles for specific substances and non-specific derivatives	/RLS	S USES/RLS	RLS
Super roles for specific substances from patents	/RL.P	S PREP/RL.P	RL.P
Super roles for non-specific derivatives from patents	/RLD.P	S PREP/RLD.P	RLD.P
Super roles for specific substances and non-specific derivatives from patents	/RLS.P	S PREP/RLS.P	RLS
Super roles for specific substances from non-patent documents	/RL.NP	S PREP/RL.NP	RL.NP
Super roles for non-specific derivatives from non-patent documents	/RLD.NP	S PROPERTIES/RLD.NP	RLD.NP
Super roles for specific substances and non-specific derivatives from non-patent documents	/RLS.NP	S COMBINATORIAL?/RLS.NP	RLS

(1) Enter HELP ROLES at an arrow prompt in the file for a list of CAplus super roles that are searchable in REGISTRY.

Limiting Search Codes

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

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

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 REGISTRY.	SUB SAM	S L7 CSS SUB=L5 SAM
Subset Range	Search a user-specified portion of answer set created by a search in REGISTRY.	SUB RAN	S L3 SUB=L2 RAN=(,50-11-3)
Subset Full	Search 100% of an answer set created by a search in REGISTRY.	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

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

Symbol	Function	Search Examples
* 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 REGISTRY 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 7.0 or higher for Windows®.

DISPLAY and PRINT Formats

You may use any of the substance information field codes, property field codes, CAplus super roles, and document type field codes to DISPLAY and PRINT answers. You may also use any of the CA document reference field codes or predefined formats, but these must always be combined with one of the substance information or property fields or 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 (1)	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
ENTE	Editor Note	D ENTE
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
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

DISPLAY and PRINT Formats (cont'd)

Format (1)	Content	Examples
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
BCF	Tabular display of Bioconcentration Factor	D IDE BCF 1-5
BP	Tabular display of Boiling Point	D BP 1-2
DEN	Tabular Display of Density	D DEN 1-2
ECND **(7)**	Tabular display of Electric Conductivity	D ECND
ECON **(7)**	Tabular display of Electric Conductance	D ECON
ERES **(7)**	Tabular display of Electric Resistance	D ERES
EREST **(7)**	Tabular display of Electric Resistivity	D EREST
FP	Tabular display of Flash Point	D IDE FP 1-5
FRB	Tabular display of Freely Rotatable Bonds	D IDE FRB 1-2
HAC	Tabular display of H Acceptors	D HAC 1
HDAS	Tabular display of H Donor/Acceptor Sum	D HDAS
HD	Tabular display of H Donors	D HD HAC
HVAP	Tabular display of Enthalpy of Vaporization	D IDE HVAP 1-5
ISLB.MASS	Tabular display of Mass Intrinsic Solubility	D ISLB.MASS
ISLB.MOL	Tabular display of Molar Intrinsic Solubility	D ISLB.MOL
KOC	Tabular display of Organic Carbon Adsorption	D IDE KOC 1-5
LD50 **(7)**	Tabular display of Median Lethal Dose	D LD50
LOGD	Tabular display of LogD	D LOGD 2-5
LOGP	Tabular display of LogP	D LOGP
MM **(7)**	Tabular display of Magnetic Moment	D MM
MP	Tabular display of Melting Point	D MP 1-2
MVOL	Tabular display of Molar Volume	D MVOL
MW	Tabular display of Molecular Weight	D MW
ORP	Tabular display of Optical Rotatory Power	D ORP
PKA	Tabular display of pKa	D PKA
PRFA (FA) **(9)**	Property fields available	D PRFA
PSA	Tabular display of Polar Surface Area	D PSA
RI	Tabular display of Refractive Index	D RI 1-2
SLB.MASS	Tabular display of Mass Solubility	D SLB.MASS
SLB.MOL	Tabular display of Molar Solubility	D SLB.MOL
SPEC **(10)**	Spectra	D SPEC
SPEC.B11NMR **(10)**	Boron-11 NMR Spectra	D SPEC.B11NMR
SPEC.C13NMR **(10)**	Carbon-13 NMR Spectra	D SPEC.C13NMR
SPEC.F19NMR **(10)**	Fluorine-19 NMR Spectra	D SPEC.F19NMR
SPEC.H1NMR **(10)**	Proton NMR Spectra	D SPEC.H1NMR
SPEC.IR **(10)**	IR Absorption Spectra	D SPEC.IR
SPEC.MASS **(10)**	Mass Spectra	D SPEC.MASS
SPEC.N15NMR **(10)**	Nitrogen-15 NMR Spectra	D SPEC.N15NMR
SPEC.O17NMR **(10)**	Oxygen-17 NMR Spectra	D SPEC.O17NMR
SPEC.P31NMR **(10)**	Phosphorus-31 NMR Spectra	D SPEC.P31NMR
SPEC.PROTONNMR **(10)**	Proton NMR Spectra	D SPEC.PROTONNMR
SPEC.RAMAN **(10)**	Raman Spectra	D SPEC.RAMAN
SPEC.SI29NMR **(10)**	Silicon-29 NMR Spectra	D SPEC.SI29NMR
TG **(7)**	Tabular display of Glass Transition Temperature	D TG
TS **(7)**	Tabular display of Tensile Strength	D TS
VP	Tabular display of Vapor Pressure	D IDE VP 1-5
DT.CA	CAplus document type	D DT.CA
RL	CAplus super roles	D RL
RL.NP	CAplus super roles from non-patents	D RL.NP
RL.P	CAplus super roles from patents	D RL.P
RLD (RL.D)	CAplus super roles for non-specific derivatives	D RLD

DISPLAY and PRINT Formats (cont'd)

Format (1)	Content	Examples
RLD. NP	CAplus super roles for non-specific derivatives from non-patents	D RLD.NP
RLD.P	CAplus super roles for non-specific derivatives from patents	D RLD.P
RLS (RLS.NP, RLS.P)	CAplus super roles for the specific substance and its non-specific derivatives	D RLS
ALL **(10)**	All available substance information, CAplus super roles and document types, property tables (EPROP, ETAG, PPROP), and BIB ABS IND for the 10 most recent CA references	DISPLAY L1 1 ALL
EPROP	Tabular display of experimental property data	D EPROP 2-3
EPROPS	EPROP, ETAG	D EPROPS
ETAG	Tabular display of experimental property tags (one line for each unique property)	D ETAG
ETAGFULL	Tabular display of all experimental property tags	D ETAGFULL
FIDE	All substance information including all names and RSD, CAplus super roles and document types, and all property tables (EPROP, ETAG, PPROP)	D FIDE
IALL **(10)**	ALL, indented with text labels for CAplus data	D IALL
IDE	Up to 50 names and other substance data except for RSD (IDE is the default)	D IDE
IDERL	IDE plus CAplus super roles and document types	D IDERL L10
MAX **(10)**	All available substance information, CAplus super roles and document types, property tables (EPROP, SPEC, ETAGFULL, PPROP), and BIB ABS IND for the 10 most recent CA references	D MAX
OPROP	Tabular display of experimental property data with original document units	D OPROP
PPROP (CALC)	Tabular display of predicted (calculated) property data	D PPROP
PROP (APROPS) **(10)**	Tabular display of property data and tags (EPROP, ETAG, PPROP)	D PROP
QRD	IDE and the rows of the table containing the hit property values or tags	D QRD
REG	CAS Registry Numbers (RN, DR, AR, PR, RR)	D REG
SAM	IN, SQL, MF, CI, STR, COMP	D L3 1-18 SAM
SCAN **(9,11)**	IN, SQL, MF, CI, STR, COMP (random display without answer numbers)	D SCAN
HIT **(12)**	Fields containing hit terms	D HIT 5-10
KWIC **(12)**	Hit terms plus 20 words on either side (KeyWord in Context)	D KWIC 5-10

(1) In addition to these substance field codes and formats, bibliographic information, except for FAN.CNT, FAN, FAM, and FBIB, for the 10 most recent documents that cite the substance in CA can be displayed if combined with at least one substance field, e.g., D RN TI AU. The substance code or format must be given first. The bibliographic formats are found on the CA Database Summary Sheet.
(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.
(9) No online display charge for this option.
(10) Spectral images are displayed with the SPEC format only in STN on the Web and via STN Express. They are available in prints sent only as EMAIL.
(11) SCAN must be specified on the command line, i.e., D SCAN or DISPLAY SCAN.
(12) 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.

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
CAplus Document Type	DT.CA	Y	N
CAplus super roles	RL	Y	N
CAplus super roles from non-patents	RL.NP	Y	N
CAplus super roles from patents	RL.P	Y	N
CAplus super roles for non-specific derivatives	RLD	Y	N
CAplus super roles for non-specific derivatives from non-patents	RLD.NP	Y	N
CAplus super roles for non-specific derivatives from patents	RLD.P	Y	N
CAplus super roles for the substance and its non-specific derivatives	RLS	Y	N
CAplus super roles for the substance and its non-specific derivatives from non-patents	RLS.NP	Y	N
CAplus super roles for the substance and its non-specific derivatives from patents	RLS.P	Y	N
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
Density	DEN	N	Y
Deleted CAS Registry Number	DR	Y **(3)**	N
Editor Note	ENTE	Y	N
Elemental Analysis for Ring System	EA	Y	N
Elemental Sequence for Ring System	ES	Y	N
Entry Date	ED	Y	Y
Experimental Property Tags	ETAG (TAGS)	Y	N
Experimental Properties	EPROP	Y **(7)**	N
File Segment	FS	Y	Y
Freely Rotatable Bonds	FRB	N	Y
Full Chemical Name	FCN	Y **(4)**	N
H Acceptors	HAC	N	Y
H Donors	HD	N	Y
Melting Point	MP	N	Y
Molecular Formula	MF	Y	N
Molecular Weight	MW	N	Y
Names	NAME	Y **(8)**	N
Optical Rotatory Power	ORP	N	Y
Nucleic Acid Sequence (exact search form)	SQEN	Y	N
Nucleic Acid Sequence (subsequence search form)	SQSN	Y	N
Patent Number	PN	Y **(9)**	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

SELECT, ANALYZE, and SORT Fields (cont'd)

FIELD NAME	FIELD CODE	ANALYZE SELECT (1)	SORT
References	REF	N	Y
Reference Accession Number in CA	RAN.CA	Y	N
Refractive Index	RI	N	Y
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 **(12)**	N
Sequence (3-letter codes)	SEQ3	Y **(12)**	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

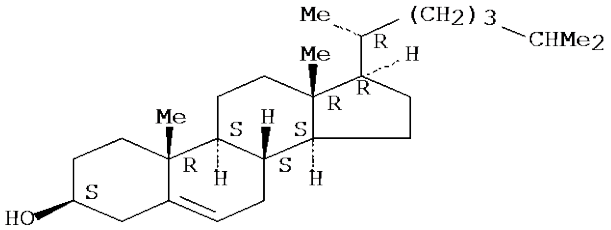
(1) HIT may be used to restrict terms extracted to terms that match the search expression used to create the answer set, e.g., SEL HIT CN.
(2) /MF is appended to the terms created by SELECT.
(3) /BI is appended to the terms created by SELECT.
(4) /CN is appended to the terms created by SELECT.
(5) CA 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) Extracts patent number from the Patent Annotation (PNTE) field.
(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.
(12) /SQSP is appended to the terms created by SELECT.

Sample Records

DISPLAY IDE

RN **57-88-5** REGISTRY
ED Entered STN: 16 Nov 1984
CN Cholest-5-en-3-ol (3β)- (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN (3β)-Cholest-5-en-3-ol
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 80356-14-5, 80356-33-8, 209124-38-9, 218965-24-3, 262418-13-3, 378185-03-6, 676322-57-9, 732297-95-9, 793670-51-6, 849593-11-9, 856708-55-9
MF C27 H46 O
CI COM
SR CA
LC STN Files: ADISNEWS, ANABSTR, BIOSIS, BIOTECHNO, CA, CABA, CAPLUS, CASREACT, CBNB, CHEMCATS, CHEMLIST, CIN, DDFU, DRUGU, IPA, MEDLINE, MSDS-OHS, NAPRALERT, PIRA, REAXYSFILE*, RTECS*, TOXCENTER, USPAT2, USPATFULL, USPATOLD, VETU
(*File contains numerically searchable property data)
Other Sources: DSL**, EINECS**, TSCA**
(**Enter CHEMLIST File for up-to-date regulatory information)

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

238846 REFERENCES IN FILE CA (1907 TO DATE)
12209 REFERENCES TO NON-SPECIFIC DERIVATIVES IN FILE CA
243605 REFERENCES IN FILE CAPLUS (1907 TO DATE)

DISPLAY SQIDE (Protein Sequence Record)

```
RN   91386-77-5  REGISTRY
CN   Interferon α1 (human leukocyte protein moiety reduced), 1-L-serine-
     (9CI)  (CA INDEX NAME)
FS   PROTEIN SEQUENCE
SQL  166

SEQ      1 SDLPETHSLD NRRTLMLLAQ MSRISPSSCL MDRHDFGFPQ EEFDGNQFQK
        51 APAISVLHEL IQQIFNLFTT KDSSAAWDED LLDKFCTELY QQLNDLEACV
       101 MQEERVGETP LMNADSILAV KKYFRRITLY LTEKKYSPCA WEVVRAEIMR
       151 SLSLSTNLQE RLRRKE
MF   Unspecified
CI   MAN
LC   STN Files:   CA, CAPLUS
DT.CA  CAplus document type:  Conference
RL.NP  Roles from non-patents:  PREP (Preparation)
               1 REFERENCES IN FILE CA (1907 TO DATE)
               1 REFERENCES IN FILE CAPLUS (1907 TO DATE)
```

DISPLAY SQIDE (Nucleic Acid Sequence Indexed by CAS)

```
RN   91449-61-5  REGISTRY
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
DT.CA  CAplus document type:  Journal
RL.NP  Roles from non-patents:  BIOL (Biological study); PRP (Properties)
               1 REFERENCES IN FILE CA (1907 TO DATE)
               1 REFERENCES IN FILE CAPLUS (1907 TO DATE)
```

DISPLAY SQIDE (Nucleic Acid Sequence Registered from GenBank®)

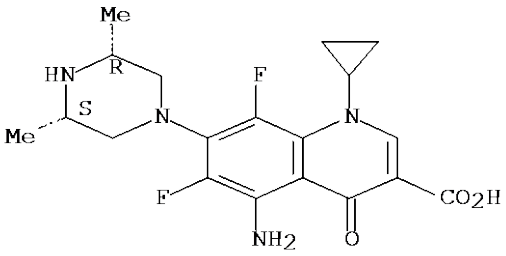
```
RN   139065-61-5  REGISTRY
CN   GenBank M12334 (9CI)  (CA INDEX NAME)
FS   NUCLEIC ACID SEQUENCE
SQL  35
NA   6 a   8 c   14 g   7 t

SEQ      1 ggaggctcat ttgcagttga ggccagcagg tcggc
MF   C342 H428 N138 O209 P34
CI   MAN
SR   GenBank
LC   STN Files:   GENBANK
```

RN **110871-86-8** REGISTRY
ED Entered STN: 24 Oct 1987
CN 3-Quinolinecarboxylic acid, 5-amino-1-cyclopropyl-7-[(3R,5S)-3,5-dimethyl-1-piperazinyl]-6,8-difluoro-1,4-dihydro-4-oxo-, rel- (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN 3-Quinolinecarboxylic acid, 5-amino-1-cyclopropyl-7-(3,5-dimethyl-1-piperazinyl)-6,8-difluoro-1,4-dihydro-4-oxo-, cis-
CN rel-5-Amino-1-cyclopropyl-7-[(3R,5S)-3,5-dimethyl-1-piperazinyl]-6,8-difluoro-1,4-dihydro-4-oxo-3-quinolinecarboxylic acid
OTHER NAMES:
CN AT 4140
CN CI 978
CN CP 103826
CN Parox
CN PD 1315-1
CN PD 131501
CN Salocin 120
CN Spara
CN Sparcin
CN Sparfloxacin
CN Zagam
FS STEREOSEARCH
MF C19 H22 F2 N4 O3
CI COM
SR CA
LC STN Files: ADISINSIGHT, ADISNEWS, BIOSIS, CA, CABA, CAPLUS, CASREACT, CBNB, CHEMCATS, CHEMLIST, CIN, DDFU, DRUGU, EMBASE, IMSRESEARCH, IPA, MEDLINE, MSDS-OHS, PATDPASPC, PS, REAXYSFILE*, RTECS*, TOXCENTER, USPAT2, USPATFULL
(*File contains numerically searchable property data)
DT.CA CAplus document type: Book; Conference; Journal; Patent
RL.P Roles from patents: ANST (Analytical study); BIOL (Biological study); NANO (Nanomaterial); OCCU (Occurrence); PREP (Preparation); PROC (Process); PRP (Properties); PRPH (Prophetic); RACT (Reactant or reagent); USES (Uses)
RLD.P Roles for non-specific derivatives from patents: ANST (Analytical study); BIOL (Biological study); NANO (Nanomaterial); PREP (Preparation); PROC (Process); USES (Uses)
RL.NP Roles from non-patents: ANST (Analytical study); BIOL (Biological study); FORM (Formation, nonpreparative); MSC (Miscellaneous); NANO (Nanomaterial); OCCU (Occurrence); PREP (Preparation); PROC (Process); PRP (Properties); RACT (Reactant or reagent); USES (Uses)
RLD.NP Roles for non-specific derivatives from non-patents: ANST (Analytical study); BIOL (Biological study); FORM (Formation, nonpreparative); PREP (Preparation); PRP (Properties); USES (Uses)

Ring System Data

Elemental Analysis EA	Elemental Sequence ES	Size of the Rings SZ	Ring System Formula RF	Ring Identifier RID	RID Occurrence Count
C3	C3	3	C3	1.13.1	1
C4N2	NC2NC2	6	C4N2	46.383.1	1
C5N-C6	NC5-C6	6-6	C9N	591.79.40	1

Relative stereochemistry.



Experimental Properties (EPROP)

PROPERTY (CODE)	VALUE	CONDITION	NOTE
Melting Point (MP)	266-269 deg C (decomp)	Solv: chloroform (67-66-3), ethanol (64-17-5)	(1) IC
Melting Point (MP)	266-269 deg C (decomp)		(2) APC
Melting Point (MP)	258-260 deg C		(3) IC
Melting Point (MP)	250-253 deg C (decomp)		(4) IC

(1) Miyamoto, Teruyuki; Journal of Medicinal Chemistry 1990 V33(6) P1645-56 CAPLUS
(2) "Drugs - Synonyms and Properties" data were obtained from Ashgate Publishing Co. (US) CAPLUS
(3) Matsumoto, Junichi; EP 221463 A2 1987 CAPLUS
(4) Petersen, Uwe; DE 3711193 A1 1988 CAPLUS

Experimental Property Tags (ETAG)

PROPERTY	NOTE
Acid/Base Dissociation Constant (Ka/Kb)	(1) CAS
ADME (Absorption, Distribution, Metabolism, Excretion)	(2) CAS
9 more tags shown in the MAX or ETAGFULL formats	
Emission/Luminescence Spectra	(3) CAS
Half-Life (Biological)	(4) CAS
6 more tags shown in the MAX or ETAGFULL formats	
IR Absorption Spectra	(5) CAS
2 more tags shown in the MAX or ETAGFULL formats	
Minimum Inhibitory Concentration	(5) CAS
34 more tags shown in the MAX or ETAGFULL formats	
Potential of Electrode Reaction	(6) CAS
Proton NMR Spectra	(7) IC
1 more tag shown in the MAX or ETAGFULL formats	
UV and Visible Absorption Spectra	(8) CAS
3 more tags shown in the MAX or ETAGFULL formats	
UV and Visible Emission/Luminescence Spectra	(9) CAS
1 more tag shown in the MAX or ETAGFULL formats	

(1) Llinas, Antonio; Journal of Chemical Information and Modeling 2008 V48(7) P1289-1303 CAPLUS
(2) Kihira, Tetsunari; Journal of Infection and Chemotherapy 2004 V10(2) P97-100 CAPLUS
(3) Wang, Jian; Science in China, Series B: Chemistry 2008 V51(1) P31-40 CAPLUS
(4) Kudo, Masakiyo; Iryo Yakugaku 2004 V30(1) P8-12 CAPLUS
(5) Efthimiadou, Eleni K.; Bioorganic & Medicinal Chemistry Letters 2008 V18(14) P4033-4037 CAPLUS

(6) Abdel Ghani, N. T.; Analytical Sciences 2007 V23(9) P1053-1058 CAPLUS
(7) Miyamoto, Teruyuki; Journal of Medicinal Chemistry 1990 V33(6) P1645-56 CAPLUS
(8) Salem, Maissa Yacoub; Chemical & Pharmaceutical Bulletin 2006 V54(12) P1625-1632 CAPLUS
(9) Sun, Chun-Yan; Spectrochimica Acta, Part A: Molecular and Biomolecular Spectroscopy 2011 V82(1) P375-382 CAPLUS

Predicted Properties (PPROP)

PROPERTY (CODE)	VALUE	CONDITION	NOTE
Bioconc. Factor (BCF)	1.0	pH 1 25 deg C	(1)
Bioconc. Factor (BCF)	1.0	pH 2 25 deg C	(1)
Bioconc. Factor (BCF)	1.0	pH 3 25 deg C	(1)
Bioconc. Factor (BCF)	1.0	pH 4 25 deg C	(1)
Bioconc. Factor (BCF)	1.0	pH 5 25 deg C	(1)
Bioconc. Factor (BCF)	1.0	pH 6 25 deg C	(1)
Bioconc. Factor (BCF)	1.0	pH 7 25 deg C	(1)
Bioconc. Factor (BCF)	1.0	pH 8 25 deg C	(1)
Bioconc. Factor (BCF)	1.0	pH 9 25 deg C	(1)
Bioconc. Factor (BCF)	1.0	pH 10 25 deg C	(1)
Boiling Point (BP)	640.4+/-55.0 deg C	760 Torr	(1)
Density (DEN)	1.436+/-0.06 g/cm**3	20 deg C 760 Torr	(1)
Enthalpy of Vap. (HVAP)	99.36+/-3.0 kJ/mol	760 Torr	(1)
Flash Point (FP)	341.1+/-31.5 deg C		(1)
Freely Rotatable Bonds (FRB)	4		(1)
H acceptors (HAC)	7		(1)
H donors (HD)	4		(1)
Hydrogen Donors/Acceptors Sum (HDAS)	11		(1)
Koc (KOC)	1.0	pH 1 25 deg C	(1)
Koc (KOC)	1.0	pH 2 25 deg C	(1)
Koc (KOC)	1.0	pH 3 25 deg C	(1)
Koc (KOC)	1.0	pH 4 25 deg C	(1)
Koc (KOC)	1.0	pH 5 25 deg C	(1)
Koc (KOC)	3.32	pH 6 25 deg C	(1)
Koc (KOC)	9.54	pH 7 25 deg C	(1)
Koc (KOC)	9.26	pH 8 25 deg C	(1)
Koc (KOC)	2.82	pH 9 25 deg C	(1)
Koc (KOC)	1.0	pH 10 25 deg C	(1)
LOGD (LOGD)	-0.51	pH 1 25 deg C	(1)
LOGD (LOGD)	-0.50	pH 2 25 deg C	(1)
LOGD (LOGD)	-0.50	pH 3 25 deg C	(1)
LOGD (LOGD)	-0.47	pH 4 25 deg C	(1)
LOGD (LOGD)	-0.27	pH 5 25 deg C	(1)
LOGD (LOGD)	0.33	pH 6 25 deg C	(1)
LOGD (LOGD)	0.79	pH 7 25 deg C	(1)
LOGD (LOGD)	0.77	pH 8 25 deg C	(1)
LOGD (LOGD)	0.26	pH 9 25 deg C	(1)
LOGD (LOGD)	-0.54	pH 10 25 deg C	(1)
LOGP (LOGP)	2.598+/-1.429	25 deg C	(1)
Mass Intrinsic Solubility (ISLB.MASS)	0.0029 g/L	25 deg C	(1)
Mass Solubility (SLB.MASS)	3.8 g/L	pH 1 25 deg C	(1)
Mass Solubility (SLB.MASS)	3.7 g/L	pH 2 25 deg C	(1)
Mass Solubility (SLB.MASS)	3.6 g/L	pH 3 25 deg C	(1)
Mass Solubility (SLB.MASS)	2.8 g/L	pH 4 25 deg C	(1)
Mass Solubility (SLB.MASS)	0.86 g/L	pH 5 25 deg C	(1)
Mass Solubility (SLB.MASS)	0.14 g/L	pH 6 25 deg C	(1)
Mass Solubility (SLB.MASS)	0.043 g/L	pH 7 25 deg C	(1)
Mass Solubility (SLB.MASS)	0.043 g/L	pH 8 25 deg C	(1)
Mass Solubility (SLB.MASS)	0.15 g/L	pH 9 25 deg C	(1)
Mass Solubility (SLB.MASS)	1.1 g/L	pH 10 25 deg C	(1)
Mass Solubility (SLB.MASS)	0.039 g/L	Unbuffered Water pH 7.47	(1)

		25 deg C	
Molar Intrinsic Solubility (ISLB.MOL)	0.0000075 mol/L	25 deg C	(1)
Molar Solubility (SLB.MOL)	0.0097 mol/L	pH 1 25 deg C	(1)
Molar Solubility (SLB.MOL)	0.0094 mol/L	pH 2 25 deg C	(1)
Molar Solubility (SLB.MOL)	0.0091 mol/L	pH 3 25 deg C	(1)
Molar Solubility (SLB.MOL)	0.0071 mol/L	pH 4 25 deg C	(1)
Molar Solubility (SLB.MOL)	0.0022 mol/L	pH 5 25 deg C	(1)
Molar Solubility (SLB.MOL)	0.00035 mol/L	pH 6 25 deg C	(1)
Molar Solubility (SLB.MOL)	0.00011 mol/L	pH 7 25 deg C	(1)
Molar Solubility (SLB.MOL)	0.00011 mol/L	pH 8 25 deg C	(1)
Molar Solubility (SLB.MOL)	0.00038 mol/L	pH 9 25 deg C	(1)
Molar Solubility (SLB.MOL)	0.0029 mol/L	pH 10 25 deg C	(1)
Molar Solubility (SLB.MOL)	0.000099 mol/L	Unbuffered Water, pH 7.47, 25 deg C	(1)
Molar Volume (MVOL)	273.2+/-3.0 cm**3/mol	20 deg C, 760 Torr	(1)
Molecular Weight (MW)	392.40		(1)
PKA (PKA)	6.42+/-0.50	Most Acidic, 25 deg C	(1)
PKA (PKA)	8.59+/-0.60	Most Basic, 25 deg C	(1)
Polar Surface Area (PSA)	98.90 A**2		(1)
Vapor Pressure (VP)	2.83E-17 Torr	25 deg C	(1)

This substance may exist in multiple tautomeric forms. The predicted property values in this table are calculated based upon the displayed form and may therefore differ from experimental values based on the actual tautomeric ratio at equilibrium.

(1) Calculated using Advanced Chemistry Development (ACD/Labs) Software V11.02 ((C) 1994-2016 ACD/Labs)

See HELP PROPERTIES for information about property data sources in REGISTRY.

3019 REFERENCES IN FILE CA (1907 TO DATE)
59 REFERENCES TO NON-SPECIFIC DERIVATIVES IN FILE CA
3050 REFERENCES IN FILE CAPLUS (1907 TO DATE)

DISPLAY CCN

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;

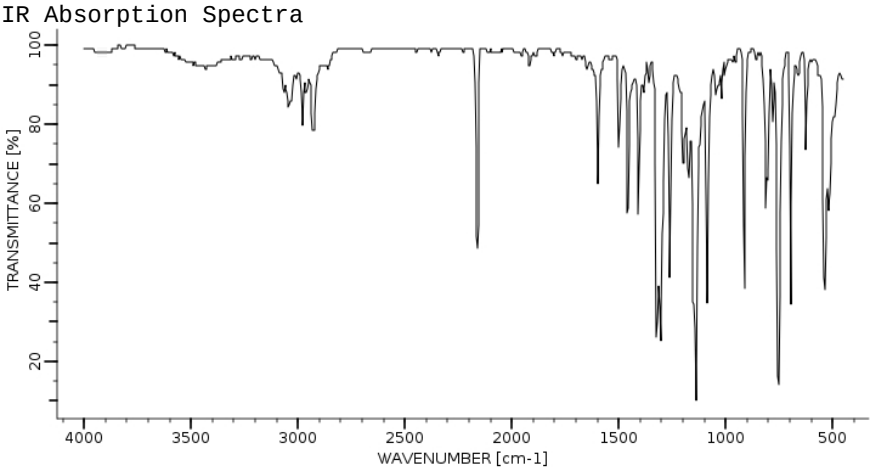
REGISTRY/ZREGISTRY

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

DISPLAY SPEC

L6 ANSWER 1 OF 93225 REGISTRY COPYRIGHT 2016 ACS on STN

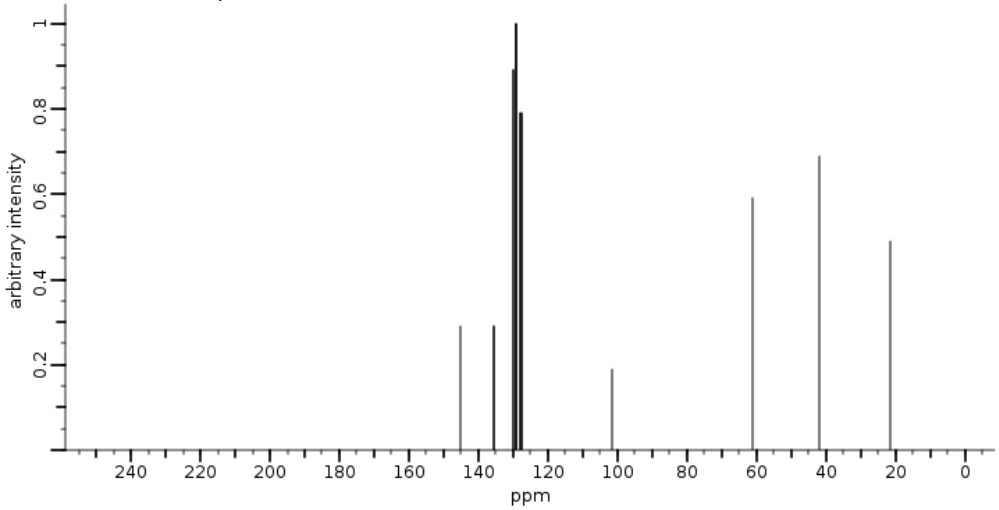
IR Absorption Spectra



Spectrum ID: BR097176
Spectrometer: BIO-RAD DIGILAB FT-IR OR EQUIVALENT
Source: Infrared spectral data from the Bio-Rad/Sadtler IR Data Collection was obtained from Bio-Rad Laboratories, Philadelphia, PA (US). Copyright (C) Bio-Rad Laboratories. All Rights Reserved.

COPYRIGHT 2016 ACS on STN

Carbon-13 NMR Spectra



Spectrum ID: NC_14175
Solvent: chloroform-d (865-49-6)
Standard: tetramethylsilane
Spectrometer: VARIAN CFT-20
Source: Copyright Bio-Rad Laboratories. All Rights Reserved.

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