



STN[®]: Stereochemistry in the CAS REGISTRYSM File



Today's presenters are...



Brenda
Poulter



Jim
Blake



Kathy
Schmidt

Stereochemistry is...

- ...A concept as old as organic chemistry itself
- ...A quality that was observed long before the structural foundation was discovered
- ...Present when the rotation of plane polarized light is observed

Stereochemistry is essential to everyday life

- DNA, RNA and ATP (as well as amino acids) are the basis for all life
 - All are enantiomerically pure in living systems
- Stereochemistry is nature's way of assuring an exact fit between a molecule (perhaps a drug) and a biological receptor

Goals for this seminar include...

- Review the basics of stereochemistry
- Review how full and partial stereochemical moieties are handled in the CAS registration process
- Use nomenclature to locate stereoisomeric substances in REGISTRY
- Determine whether to search a stereo moiety or a flat moiety
- Create a stereochemical structure query and execute a search

Review: What are enantiomers?

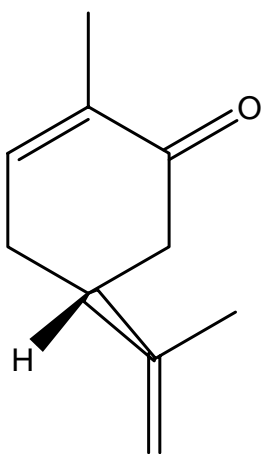
- Enantiomers are non-superimposable mirror image molecules
- Enantiomers have identical chemical and physical properties
- They differ in how they rotate polarized light
 - A pair of enantiomers rotate the plane of polarized light by equal amounts
 - A 1:1 mixture of the two do not rotate light, this is called a racemate or racemic mixture
- They differ in their interaction with other chiral molecules
 - Very important in the field of biology

Other types of stereochemistry exist

- Compounds may exist as diastereomers
 - Stereoisomers that are *not* mirror images
 - Differ by spatial orientation of atoms
 - Have different physical properties
- Drugs frequently have to be optically pure
 - Different physical properties translate to different effects on the body

Stereochemical differences can produce physically detectable effects

- Stereoisomers may have different smells and tastes:

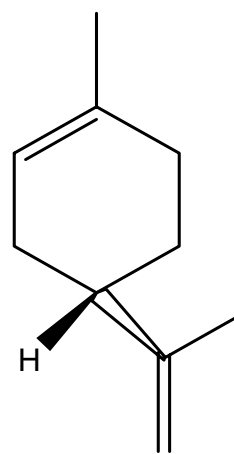


(R)-Carvone

Spearmint
6485-40-1

(S)-Carvone

Caraway
2244-16-8



(S)-Limonene

lemons
5989-54-8

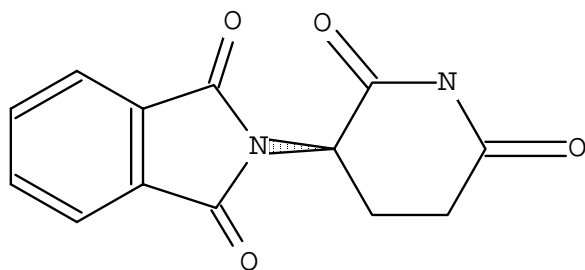
(R)-Limonene

oranges
5989-27-5

- Both the nose and the tongue are excellent stereochemical detectors

Stereochemistry influences biological activity

Significant effects:



(R)-Thalidomide

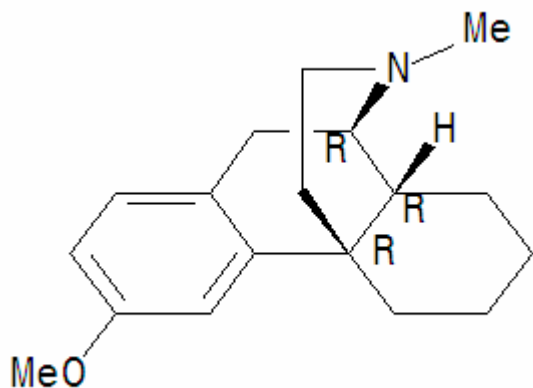
Sedative

2614-06-4

(S)- Thalidomide

Birth Defects

841-67-8



Levomethorphan

Painkiller

125-70-2

Dextromethorphan

Anti cough agent

125-71-3

REGISTRY contains more than 6 million stereo-searchable structures (09/05)

- Not including sequences
- STEREOSEARCH file segment
 - FS STEREOSEARCH
 - Text-searchable field (/FS)
- Many other substances exist with chiral centers or double bonds, but have no stereo specified within REGISTRY
 - Not reported by authors

More than 250,000 substances have the term “stereoisomer” in their CA Index Name

- Over 237,000 have “flat” structures with no stereo shown or searchable
- The term is used when stereochemistry is known but cannot be described using nomenclature rules

Stereochemical indexing policies have changed over the years

- Before 1997 (prior to 14th CI)
 - Partial stereo was registered as “flat” substance with no stereo shown
- Starting in 1997 (beginning with the 14th CI)
 - Partial stereo registered as shown in original document

Many different types of stereochemistry are captured in REGISTRY

- Types of stereochemistry
 - Double bond (E/Z and *cis/trans*)
 - Chiral (R/S, α/β , D/L, *cis/trans*)
 - Relative
 - Absolute
- Racemates (+/-)
 - Indexed with no stereo specified
 - (+-) used with stereoparents

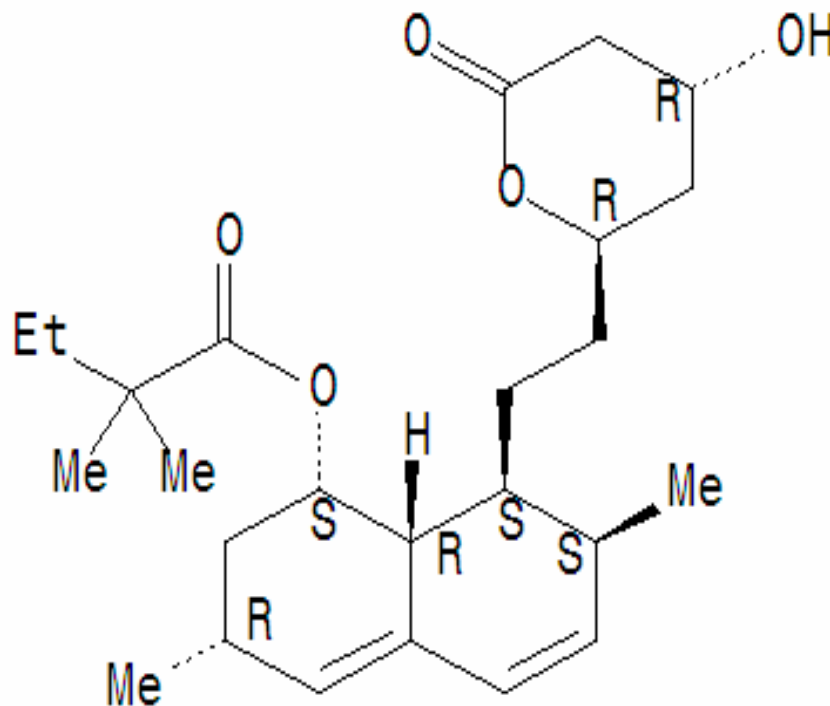
Tips for chemical name searching

- Common search fields
 - Basic Index
 - Names are segmented into chemically significant fragments (truncation unnecessary)
 - Chemical Name (/CN)
 - Complete name with locants and stereo (use right-hand truncation)
 - Chemical Name Segment (/CNS)
 - Names broken into natural segments (use simultaneous left and right truncation (SLART))

Example: Chemical Name field for Zocor

Butanoic acid, 2,2-dimethyl-, (**1S,3R,7S,8S,8aR**)-1,2,3,7,8,8a-hexahydro-3,7-dimethyl-8-[2-[(**2R, 4R**)-tetrahydro-4-hydroxy-6-oxo-2H-pyran-2-yl]ethyl]-1-naphthalenyl ester

Absolute stereochemistry.



/CN CA Index Name shown

Example: Possible nomenclature search terms for Zocor

- Basic Index (/BI) stereo search terms
 - 1S 3R 7S 8S 8AR 2R 4R
- Chemical Name Segment (/CNS) stereo search terms
 - 1S,3R,7S,8S,8AR/CNS
 - 2R,4R/CNS

Tips for chemical name searching

If the chemical name contains	Then your action is
Superscripts Subscripts Italic letters/numbers	Ignore the italicization or the fact that a character is super/subscripted => E DICHLOROMETHANE-D2 /CN
Greek letters	Spell out the name of the Greek letter and place a period before and after the name of the Greek letter => E .ALPHA.-ACETYLNAPHTHALENE /CN
Primes (apostrophes)	Place quotation marks (" ") around the entire name => S "N,N'-DIMETHYL-1,2-ETHANEDIAMINE" /CN
Parentheses	Place quotation marks (" ") around the entire name => S "2-(1-ACETOXYETHYL)FURAN" /CN
Brackets	Replace brackets with parentheses and place quotation marks (" ") around the entire name => S "BENZO(B)THIOPHENE" /CN

Stereospecific nomenclature includes many specific conventions

Stereochemistry is listed in various ways:

- Chiral centers
 - D, L
 - R, S
 - alpha, beta
- Double bonds
 - cis, trans
 - E, Z

With so many designated conventions for stereochemical naming, searching without the stereo portion may be best.

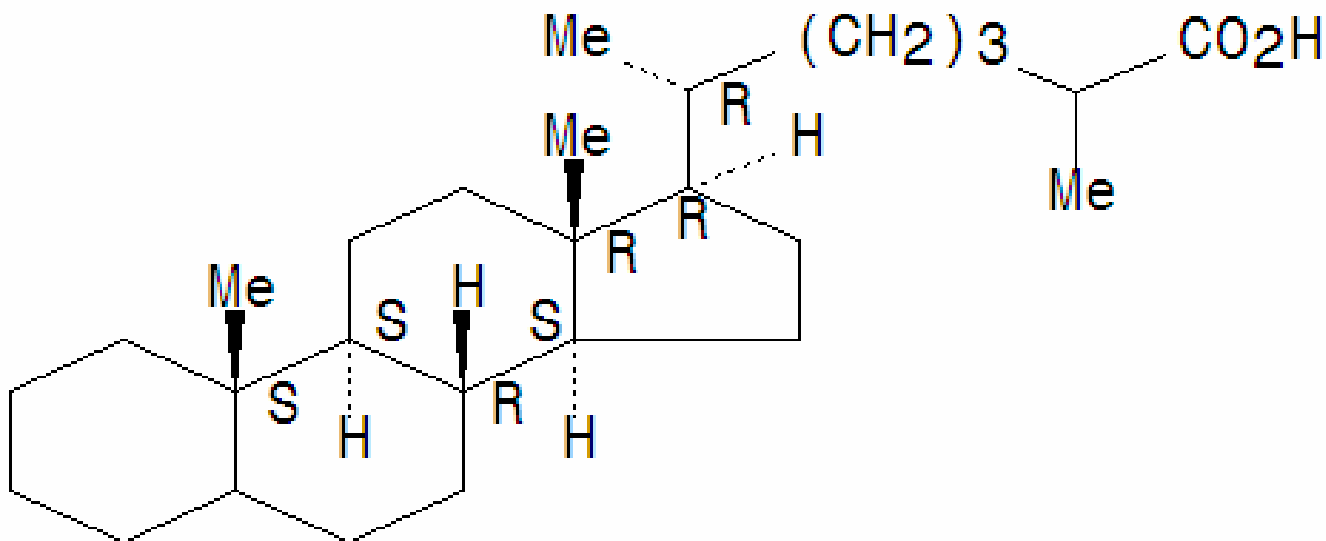
Stereo notation can appear in various places in substance names

- L-Lysinamide, D-alanyl-L-phenylalanyl-
- ...-8-oxo-, [6R-[6.alpha.,7.beta.(R*)]]-
- ...amino]-, (2R,3R)-2,3-dihydroxy...
- ...phenoxy]ethyl]-, [S-(Z)]-

Substance names may also imply stereochemistry

- For example, stereoparents
 - Cholestan-26-oic acid

Absolute stereochemistry.



Examples of named stereoparents

- Alkaloids
- Amino Acids
- Peptides
- Proteins
- Carbohydrates
- Cyclitols
- Nucleosides & Nucleotides
- Steroids
- Terpenes

Tips for working with stereoparents

- Stereochemistry is assumed unless otherwise specified
 - Stereo for other chiral centers is stated
- To see stereo for the stereoparent
 - EXPAND & SEARCH on name in /CN
 - DISPLAY name and structure

Use nomenclature to conduct a stereosearch

Search Question:

Modified amino acids are used for many therapeutic and industrial areas. The stereoisomer

β -(3,4-Dihydroxyphenyl)- α -L-alanine

is used for what particular therapeutic areas?

Search Strategy

To use nomenclature to conduct a stereosearch...

- Step 1. Modify the nomenclature for the search
 - Add any necessary quotation marks, Greek letters, etc.
- Step 2. Search the stereospecific compound name in REGISTRY
- Step 3. Display the compound of interest
- Step 4. Find and display associated CAplusSM documents

Modify the nomenclature for the search

The given name of :

β -(3,4-Dihydroxyphenyl)- α -L-alanine

Becomes:

“.beta.-(3,4-Dihydroxyphenyl)-.alpha.-L-alanine”

Search the stereospecific compound name in REGISTRY

=> **FILE REGISTRY**

=> **S ".beta.-(3,4-Dihydroxyphenyl)-.alpha.-L-alanine"/CN**

The RTF transcript will automatically convert search notation back into Greek symbols:

=> **S " β -(3,4-Dihydroxyphenyl)- α -L-alanine"/CN**

L3 1 ".beta.-(3,4-DIHYDROXYPHENYL)-.alpha.-L-ALANINE"/CN

Display the compound of interest

=> D L3 IDE

L3 ANSWER 1 OF 1 REGISTRY COPYRIGHT 2005 ACS on STN
RN 59-92-7 REGISTRY

ED Entered STN: 16 Nov 1984

CN L-Tyrosine, 3-hydroxy- (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Alanine, 3-(3,4-dihydroxyphenyl)-, L- (8CI)

OTHER NAMES:

CN (-)-3,4-Dihydroxyphenylalanine

CN (-)-Dopa

CN β -(3,4-Dihydroxyphenyl)- α -L-alanine

CN β -(3,4-Dihydroxyphenyl)-L-alanine

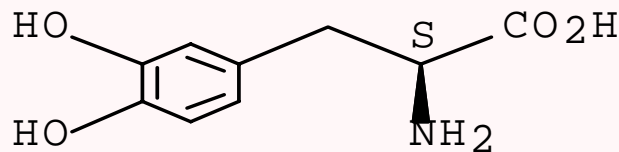
CN β -(3,4-Dihydroxyphenyl)alanine



FS STEREOSEARCH

DR 25525-15-9, 23734-74-9, 72572-99-7, 72573-00-3,
88250-23-1, 34241-25-3

Absolute stereochemistry.



Find and display associated CAPLUS documents

=> FILE CAPLUS; S L3

=> D L4 3003

L4 ANSWER 3003 OF 11149 CAPLUS COPYRIGHT 2005 ACS on STN
AN 2002:244634 CAPLUS [Full-text](#)
DN 136:257277
TI L-Dopa derivatives for prevention and treatment of
Parkinson disease
IN Sudo, Junichi; Nagai, Koji; Higashiyama, Kimio;
Takayama, Kozo; Iwase, Hiroaki; Kakuno, Katsuhiko
PA Japan
SO Jpn. Kokai Tokkyo Koho, 19 pp.
CODEN: JKXXAF
DT Patent

Stereospecific nomenclature may also appear in the abstract of a CAplus record, in addition to the indexing

AB L-Dopa derivs. (I; R1 = H, C1-16 alkyl; R2, R3 =H, C1-16) and their pharmacol. acceptable salts are claimed for prevention and treatment of Parkinson disease. I (0.5-15% weight) can be formulated into transdermal preps., including sheets, containing skin absorption promoters hydrocarbons, high alcs., high fatty acids and their esters, glycols, etc., oils and water-soluble excipients carboxyvinyl polymer, hydroxypropyl cellulose, alcs., 1-5% weight lemonene and 1-menthol, and 15-45% weight ethanol.

A Word of Advice



- Search by trade name or by systematic name in the /CN index
- Search by name and stereo fragments in the Basic Index

Structure searching may also be used to locate stereochemical compounds

- Flat structure or stereospecific?
- Tips for compounds containing double bonds
- Chiral stereochemistry

Search the flat structure query first as a FULL search

- Search the stereo query using a subsequent SUBSET search if necessary
 - Isolate the “flat” answers with no stereo specified
- Why take this approach?
 - “Flat” search yields all stereoisomers as well as structures with no stereo shown
 - “Flat” answers may be valid retrievals
 - Stereospecific search misses any answers that do not fit the query stereochemistry

Tips for drawing double bond queries

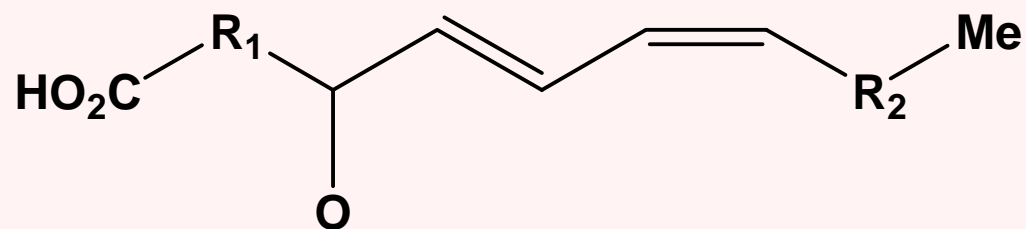
- Singly-bonded attachments must give clear and unambiguous representation of double bond geometry
- Two attachments to the same node must NOT be on the same side of the axis of the double bond
- Only one or the singly attached nodes may be on the same axis as the double bond

Tips for drawing double bond queries

- If there is only one singly-bonded attachment, it cannot be on the double bond axis
- Double bonded node may *not* be:
 - Shortcut
 - G-group
 - Generic group symbol
 - Hydrogen isotope
 - Chiral center

Double bond search example

Search Question: Find compounds with the structure shown below.



R1 = R2 = saturated, linear, unsubstituted
carbon chain

How many substances have the exact double
bond geometry?

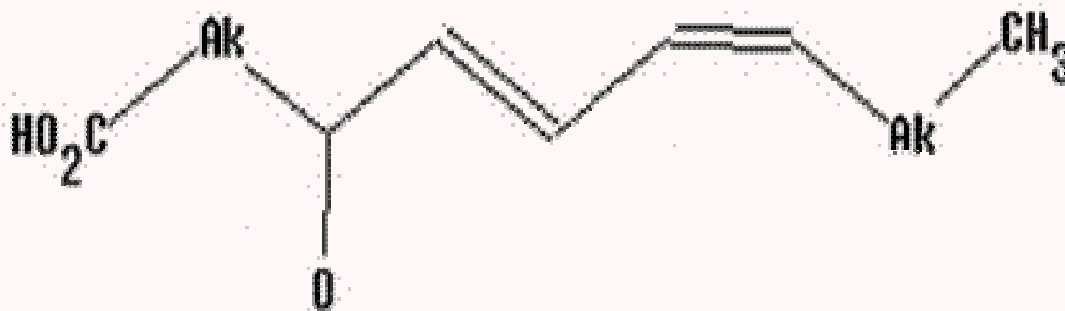
Search Strategy

To find stereo compounds containing double bonds...

- Step 1. Draw the double bond in the desired orientation
- Step 2. Search the structure without stereo in REGISTRY and review answers
- Step 3. Isolate substances with no stereo shown
- Step 4. Search for substances with desired stereo by changing query settings
- Step 5. Isolate other stereoisomers

Draw the double bond in the desired orientation

Search query:



Ak = saturated, linear carbon chain

NON-H attachments = Exactly 2

Search the structure without stereo in REGISTRY

=> S L1 FULL

FULL SEARCH INITIATED 09:35:30

FULL SCREEN SEARCH COMPLETED - 19059 TO ITERATE
100.0% PROCESSED 19059 ITERATIONS

57 ANSWERS

SEARCH TIME: 00.00.01

L3 57 SEA SSS FUL L1

L1 is a query with *no stereo specified*. The uploaded query structure was displayed and a SAMPLE search was run.

Review the answers

=> D SCAN

L3 57 ANSWERS REGISTRY COPYRIGHT 2005 ACS on STN
IN 10,12-Pentadecadienoic acid, 9-hydroxy- (9CI)
MF C15 H26 O3



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

HOW MANY MORE ANSWERS DO YOU WISH TO SCAN? (1):0

The initial search retrieved 57 answers

- Isolate those with stereo shown

=> S L3 AND STEREOSEARCH/FS

L4 43 L3 AND STEREOSEARCH/FS

- Isolate substances with no stereo shown

=> S L3 NOT L4

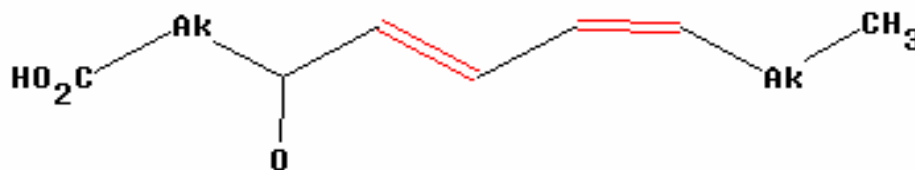
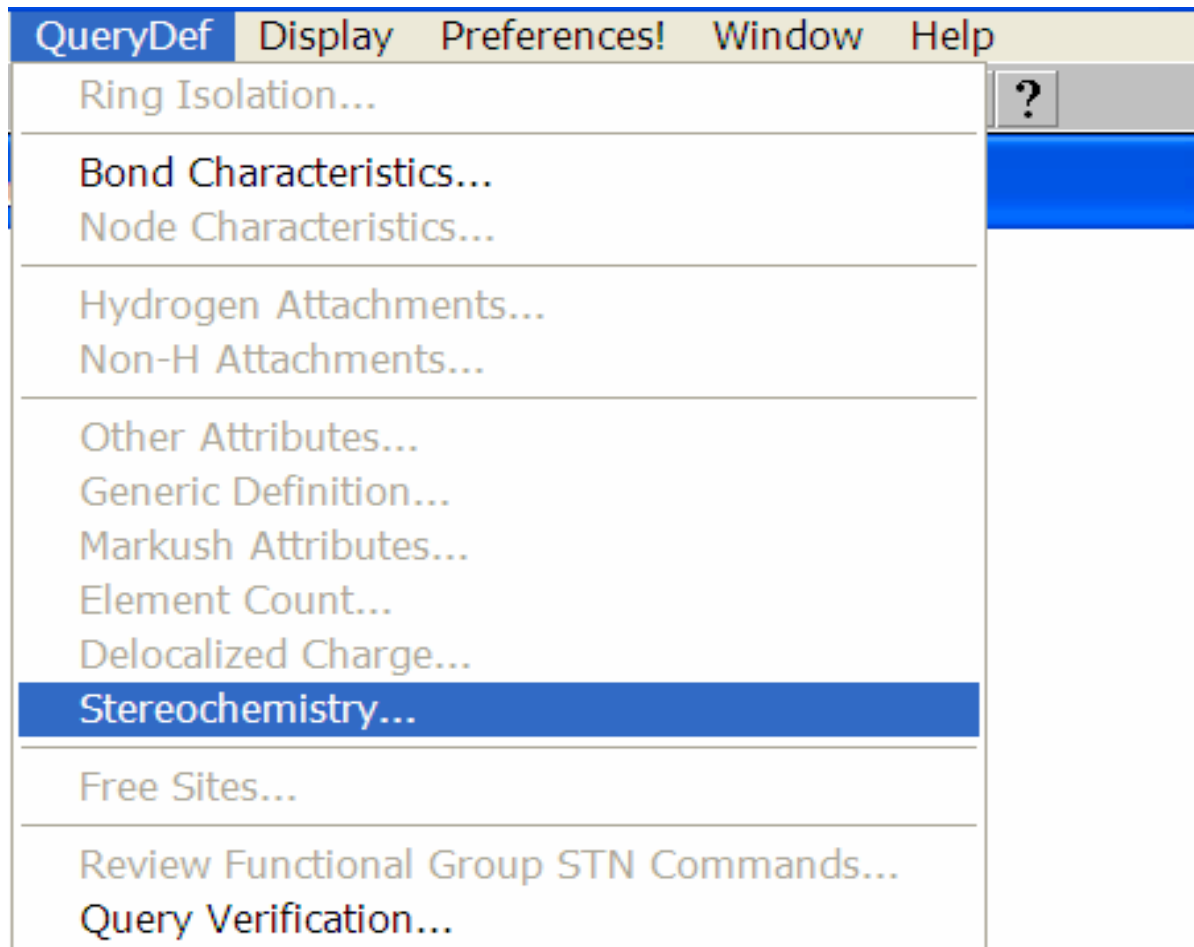
L5 14 L3 NOT L4

- Search for substances with desired stereo
 - Add stereo to query

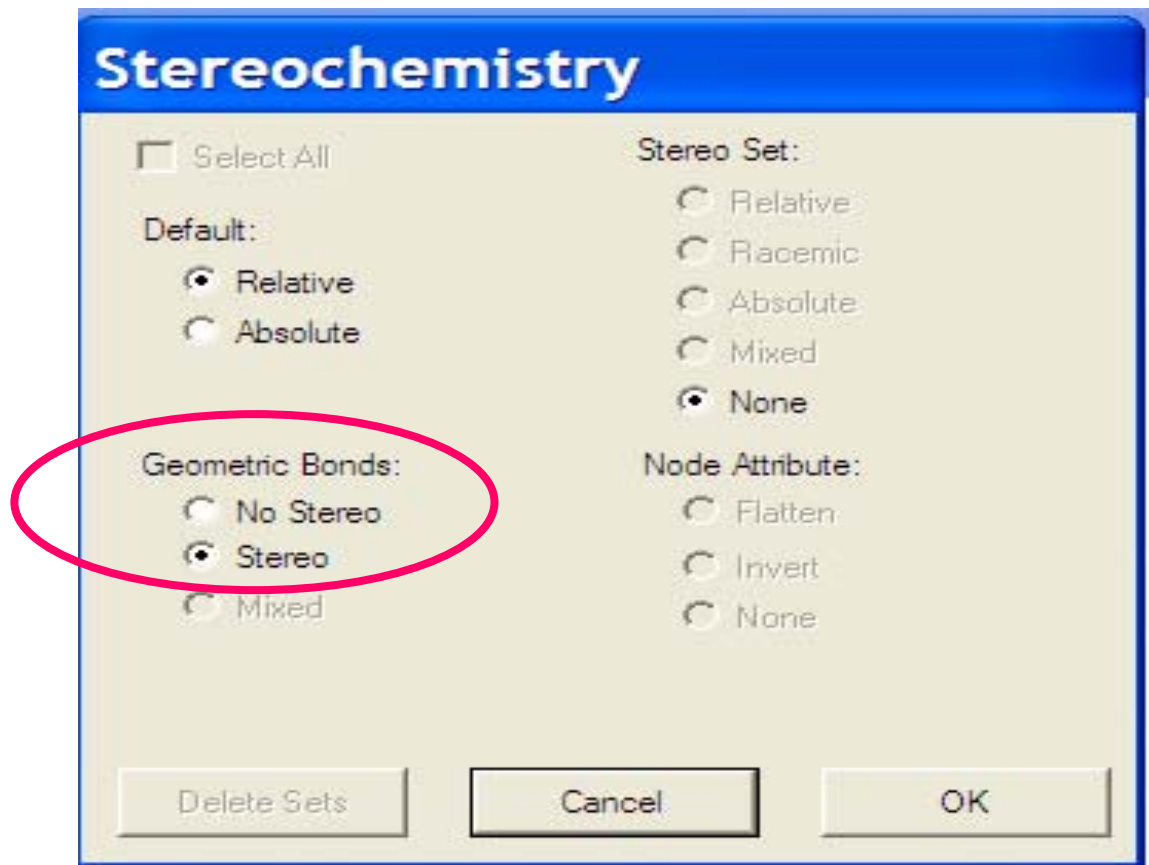
Add stereochemical information to the query

- Draw the double bond in the desired orientation (already done)
- Highlight the double bond(s)
- Open the **QueryDef** menu and select **Stereochemistry**
- Click the **Stereo** radio button under **Geometric Bonds**
- Click **OK**

Open the QueryDef menu and select **Stereochemistry**



Click the ***Stereo*** radio button under **Geometric Bonds**



Upload the stereo query and do a subset search

L6 STRUCTURE UPLOADED

=> S L6 SUB=L4 FUL

FULL SUBSET SEARCH INITIATED 09:38:00

FULL SUBSET SCREEN SEARCH COMPLETED - 43 TO ITERATE

100.0% PROCESSED 43 ITERATIONS 22 ANSWERS

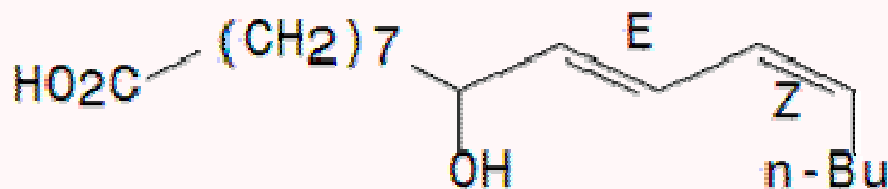
SEARCH TIME: 00.00.01

L7 22 SEA SUB=L4 SSS FUL L5

Review the resulting answers

=> D SCAN

L7 22 ANSWERS REGISTRY COPYRIGHT 2005 ACS on STN
IN 10,12-Heptadecadienoic acid, 9-hydroxy-,
(E,Z)- (9CI)
MF C17 H30 O3



Isolate other stereoisomers

- Z,E; Z,Z; E,E

=> S L4 NOT L7

L8

21 L4 NOT L7

Double bond search summary

- 57 total answers
 - 22 with desired E,Z stereo
 - 21 with other stereo
 - 14 with no stereo specified

Tips for drawing stereo queries

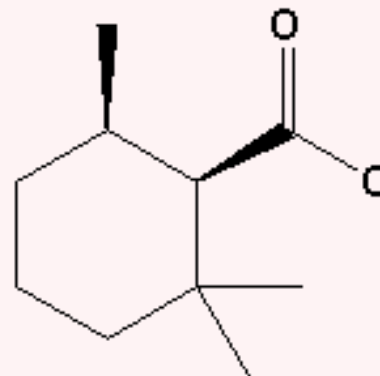
- Stereo and flat bonds must give a clear, unambiguous representation of a chiral center
- Two or three flat bonds must define the reference plane
- Only one or two stereo bonds may be shown from the chiral center

Tips for drawing stereo queries

- A chiral center may not be a shortcut, G-group, generic group symbol, or hydrogen isotope
- No two nodes attached to a chiral center may be the same, but their differences may be implied
 - e.g., two carbon atoms in an open substructure search

Example: Find a fragrance composition

Search Question: Stereoisomers of the structure below are used in fragrance compositions. How many substances share the structure? How many substances have the same relative stereochemistry? How many have the same absolute stereochemistry? The ring is isolated.



Search Strategy

To find stereo compounds...

Step 1. Search the “flat” structure

- Draw and save flat structure
- Upload and search flat structure
- Don't forget to run a SAMPLE search first

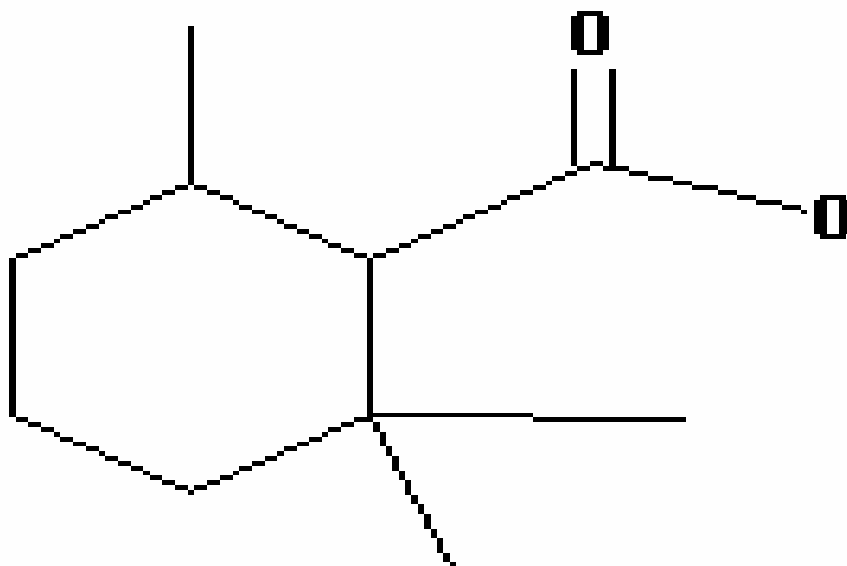
Step 2. Choose the desired stereo option

- Relative stereosearch
- Absolute stereosearch

Option 1: Perform a relative stereosearch

- Relative stereochemistry
 - Default for stereo bonds
 - Use the “flat” structure
 - Right-click in the **Bond** window or open the **Draw** menu and select **Bond...**
 - Select the desired stereo bond
 - Place in query
 - Save as new query

Assign relative stereo: Right-click in the **Bond** box

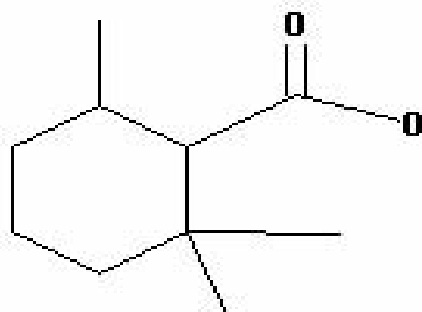


Bond box.

A screenshot of a software interface showing a toolbar. The toolbar is a horizontal bar with a blue top edge and a light beige background. It contains several icons: a box with the letter 'F', a box with 'Si', a box with 'I', a box with a single horizontal line, a box with a single horizontal line on a dark background, a box with two parallel horizontal lines, and a box with a 'C' and a diagonal line. A purple arrow points from the 'Bond box.' text to the single horizontal line icon.

Select the desired stereochemistry

Stereo from plane.



Bond Selection

- Single
- Double
- Triple
- Unspecified

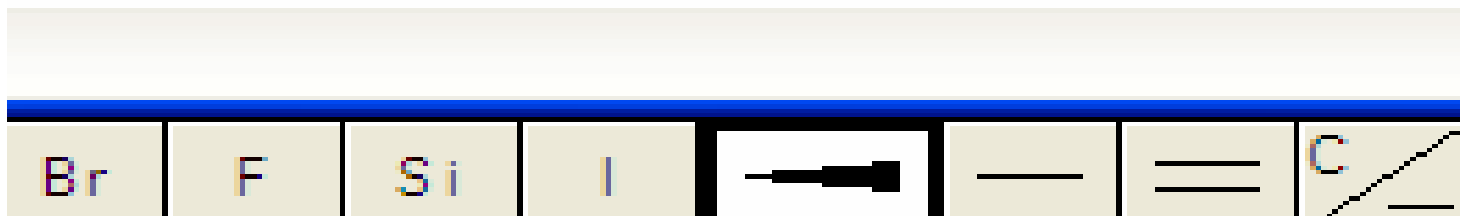
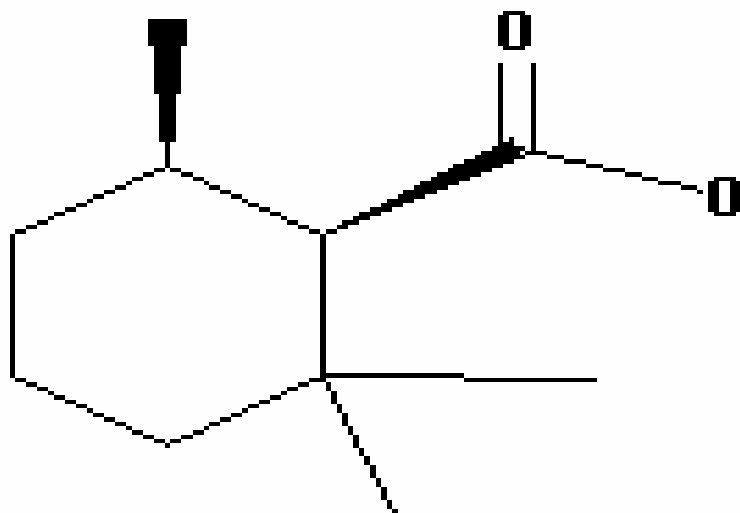
- Stereo thick bond
- Stereo from plane
- Stereo into plane
- Stereo double from plane
- Stereo double into plane
- Stereo unspecific

Cancel

Multiple Use

Single Use

Apply the stereo bond as desired
within the query



A horizontal toolbar for chemical drawing. From left to right, it contains buttons for: Br, F, Si, I, a wedge bond symbol, a dash bond symbol, a double bond symbol, and a carbon atom symbol (C) with a diagonal slash.

To distinguish relative stereo, run a subset search on the “flat” answer set

```
=> S L4 SUB=L3 FUL
```

```
FULL SUBSET SEARCH INITIATED 09:04:16
```

```
FULL SUBSET SCREEN SEARCH COMPLETED 228 TO ITERATE
```

```
100.0% PROCESSED          228 ITERATIONS
```

```
69 ANSWERS
```

```
SEARCH TIME: 00.00.01
```

```
L5          69 SEA SUB=L3 SSS FUL L4
```

Review relative stereo answers using D SCAN

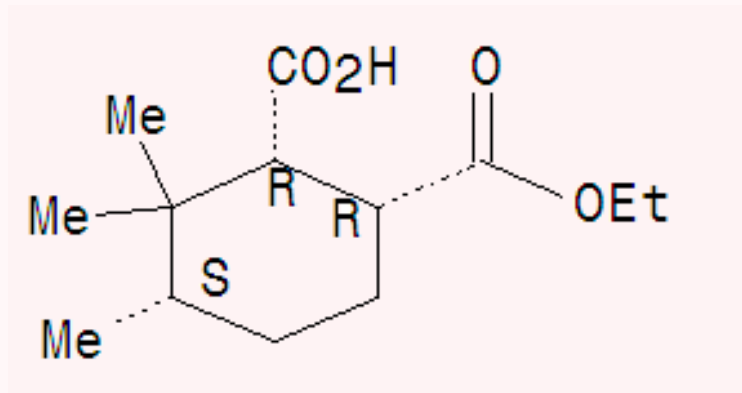
=> **D SCAN**

L5 69 ANSWERS REGISTRY COPYRIGHT 2005 ACS on
STN

IN 1,2-Cyclohexanedicarboxylic acid, 3,3,4-
trimethyl-, 1-ethyl ester, (1R,2R,4S)-rel-
(9CI)

MF C13 H22 O4

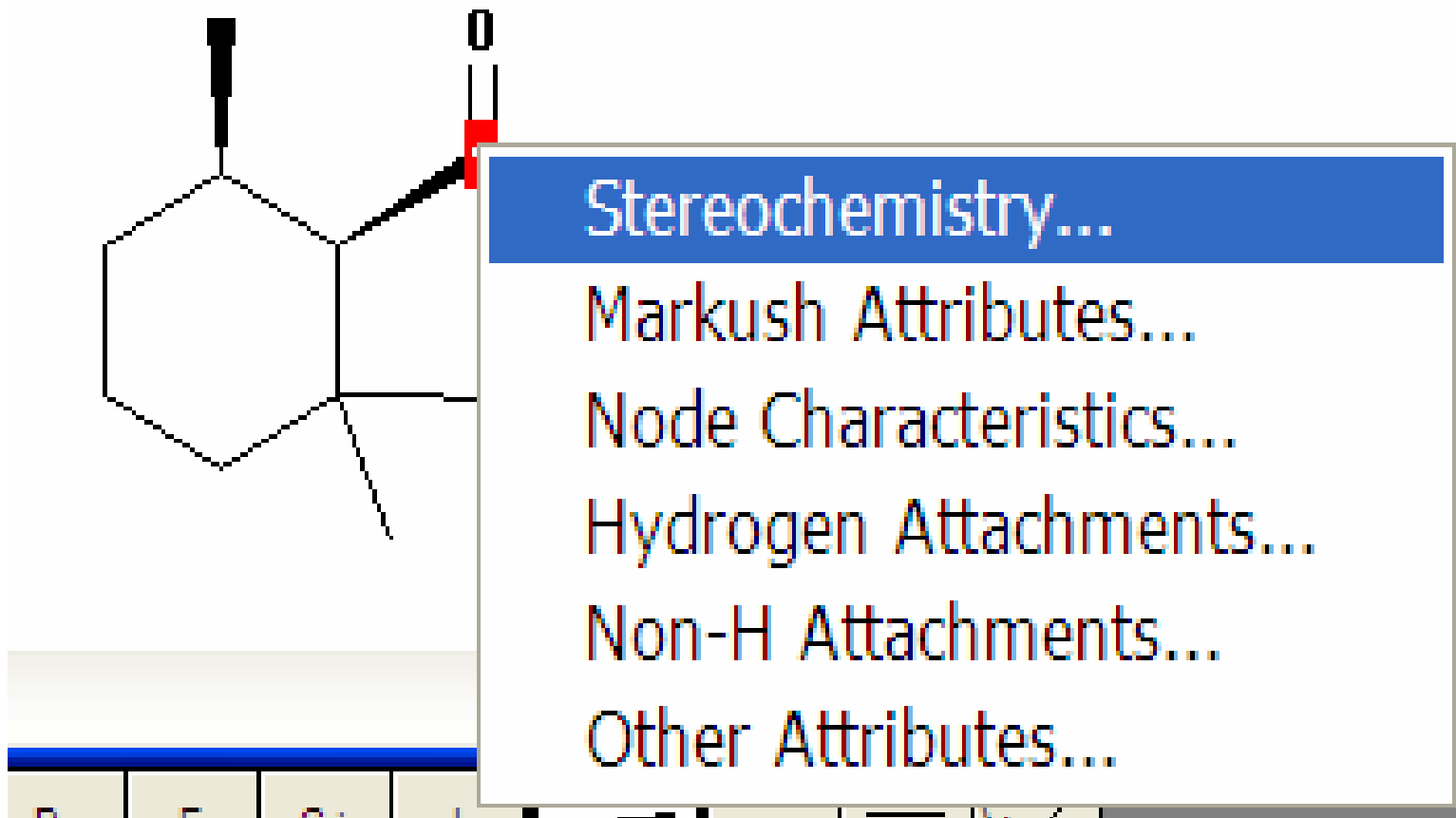
Relative stereochemistry.



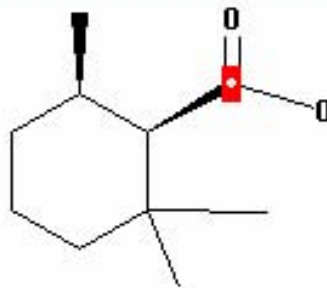
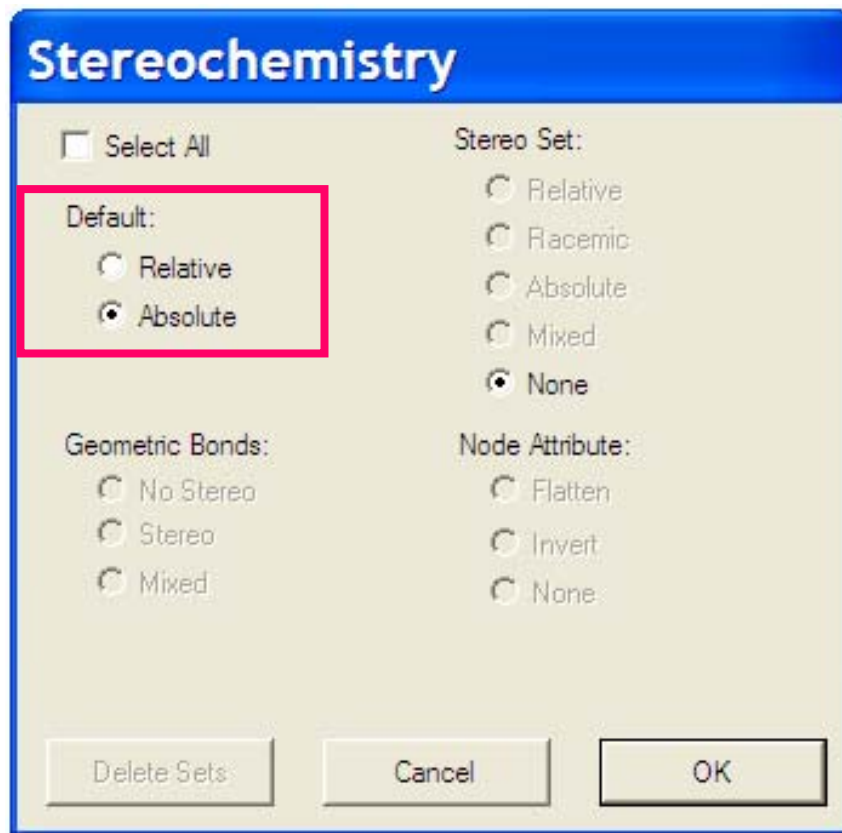
Option 2: Run an absolute stereo search

- Draw the structure in the same manner as for a relative stereo search
- One extra step:
 - Assign “Absolute”

Right-click on the node or bond of interest and select ***Stereochemistry***



Click the ***Absolute*** radio button under **Default**



To distinguish absolute stereo,
run a subset search on the
“flat” answer set

```
=> S L8 SUB=L3 FUL
```

```
FULL SUBSET SEARCH INITIATED 09:04:16
```

```
FULL SUBSET SCREEN SEARCH COMPLETED 228 TO ITERATE
```

```
100.0% PROCESSED      228 ITERATIONS      69 ANSWERS
```

```
SEARCH TIME: 00.00.01
```

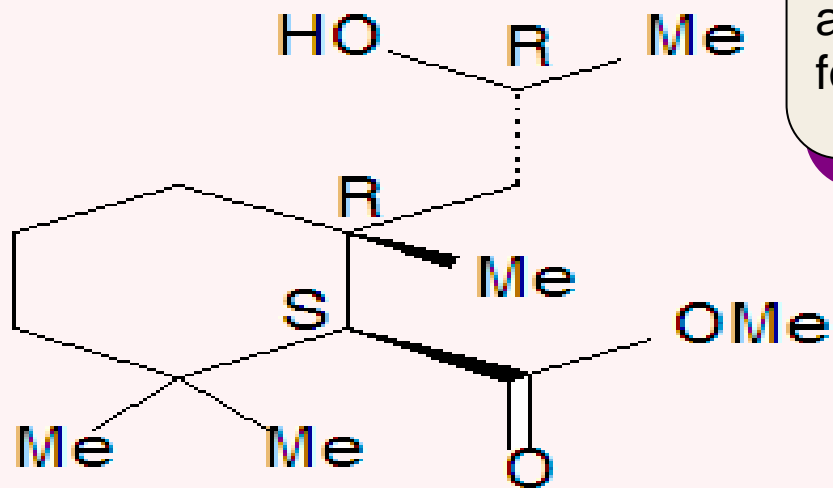
```
L9      11 SEA SUB=L3 SSS FUL L8
```

Review absolute stereo answers using D SCAN

=> **D SCAN STR**

L9 11 ANSWERS REGISTRY COPYRIGHT 2005 ACS on
STN

Absolute stereochemistry.



You may specifically display fields within SCAN that are already a part of the SCAN format, e.g., STR.

A Word of Advice



- Searching stereochemistry can be best achieved by using a flat structure, in most cases
- In nomenclature searching
 - use the exact name with stereo in the /CN field
 - Use name and stereo fragments in the Basic Index

Additional resources

STN Express web page

www.cas.org/ONLINE/STN/discover.html

STN Express learning resources

www.cas.org/ONLINE/STN/expresources.html

“Using the CAS REGISTRY File on STN”

www.cas.org/ACAD2/casregis.pdf

“Structure Searching in the CAS REGISTRY File on STN”

www.cas.org/ACAD/strcsrch.pdf

Reference e-seminars

For more tips on performing the types of searches and techniques used in this seminar please see:

- STN: Structure and Substructure Searching Tips
- STN: Advanced Structure Search Techniques: Ring Information

Questions and answers...



SCIENTISTS SERVING SCIENCE



Products Customer Care About CAS What's New CAS Learning Center

Meeting Center Training Center Support Center **Event Center** More Services ▼ My WebEx

Date & Time	Event
October 2005	
October 25, 2005 1:00 pm- 2:00 pm EDT	STN: Finding Clinical Trial and Drug Pipeline Information
November 2005	
November 29, 2005 1:00 pm- 2:00 pm EDT	STN: Multifile Patent Searching
December 2005	
December 13, 2005 1:00 pm- 2:00 pm EDT	STN: Using Boolean Operators in Structure Searching
January 2006	
January 12, 2006 8:30 am- 9:30 am EDT	STN: Using Boolean Operators in Structure Searching

<http://casevents.webex.com>