



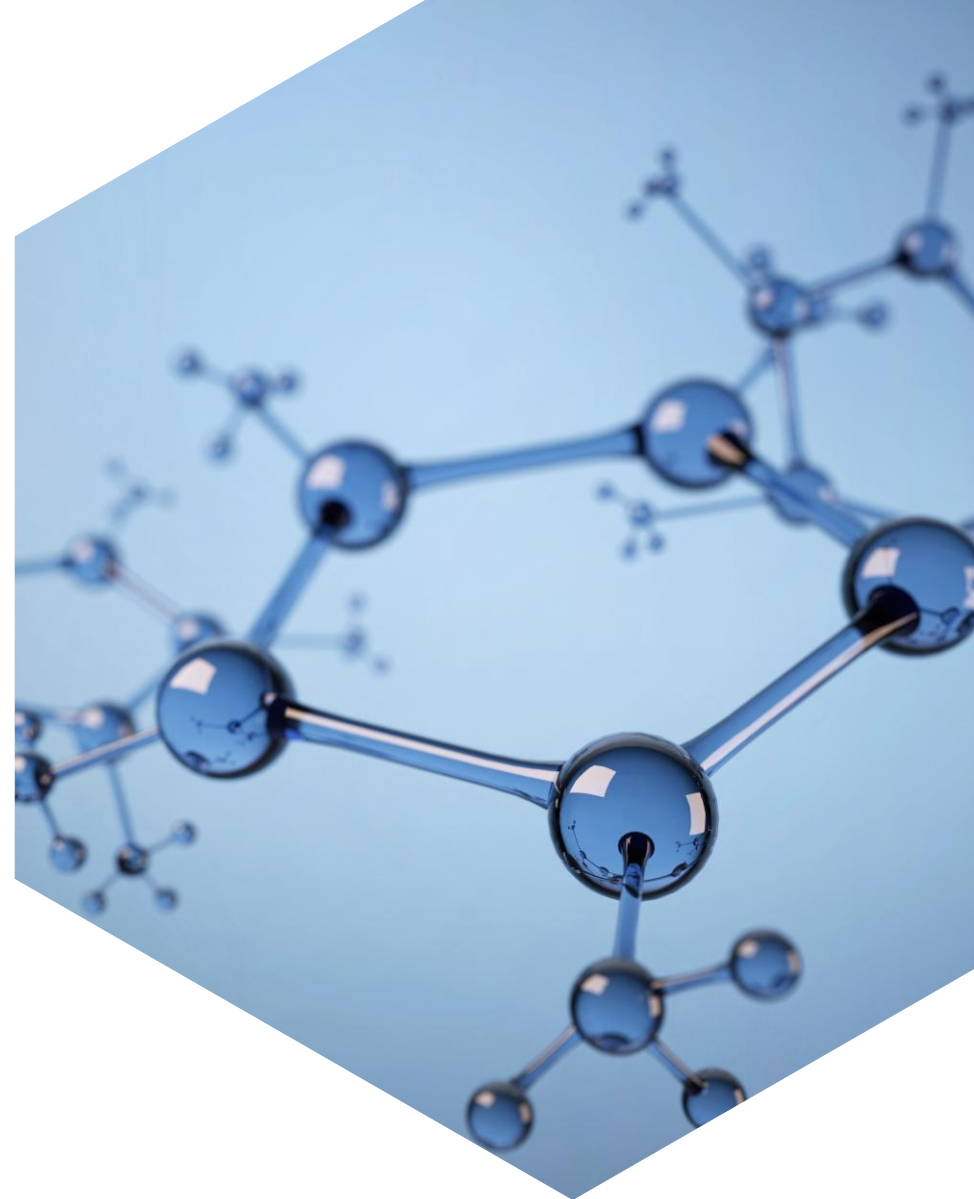
ENHANCING SEARCHING CAPABILITIES WITH INCHI STRINGS AND INCHI KEYS IN CAS STNEXT

Jim Brown – FIZ Karlsruhe

© 2025 American Chemical Society. All rights reserved.

AGENDA

- What is InChI?
- Implementation of InChI in STNext databases
- Examples
- Considerations





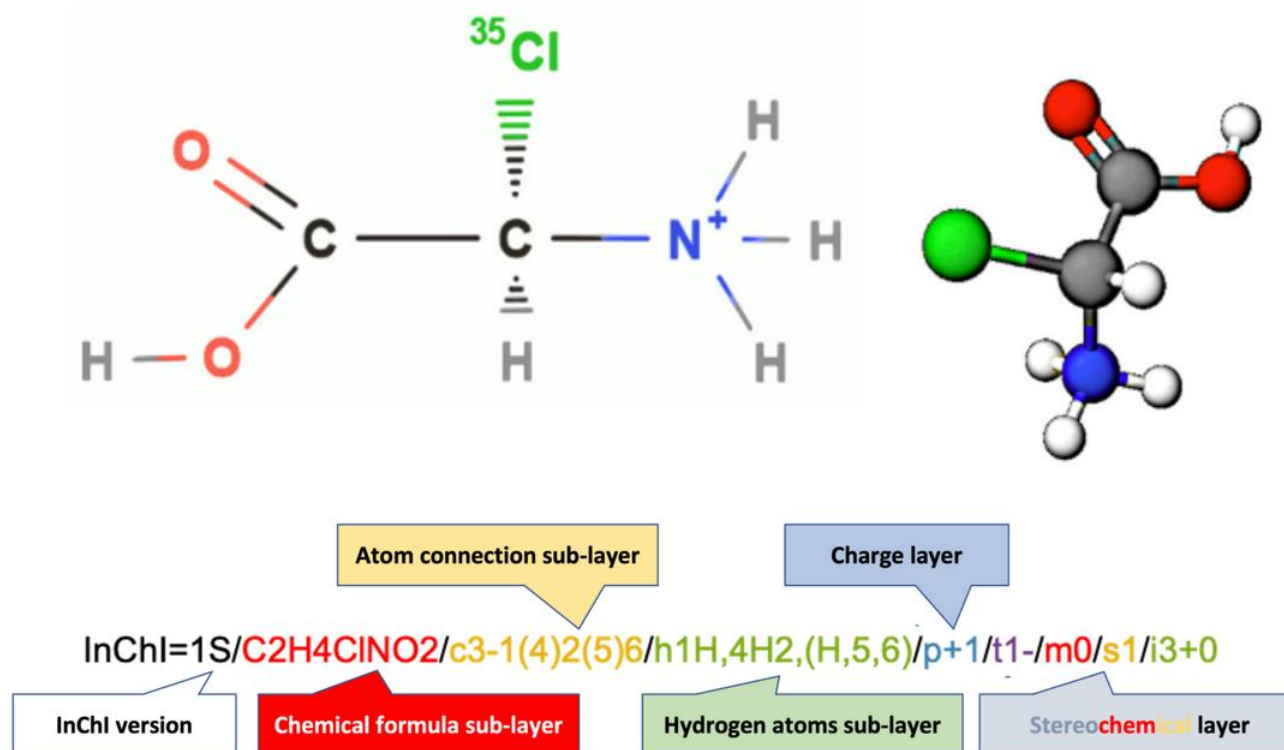
WHAT IS INCHI?

InChI – International Chemical Identifier

- Unique identifier for chemical structures developed to support the communication about specific chemical compounds.
- Developed by IUPAC and InChI Trust; released 2005
- Standardized, non-proprietary digital identifier generated from the structure by the freely available, open source InChI algorithm

InChI – layered structure

Standard InChI defines chemical structures by four levels of information - connectivity, isotopy, stereochemistry and electricity.

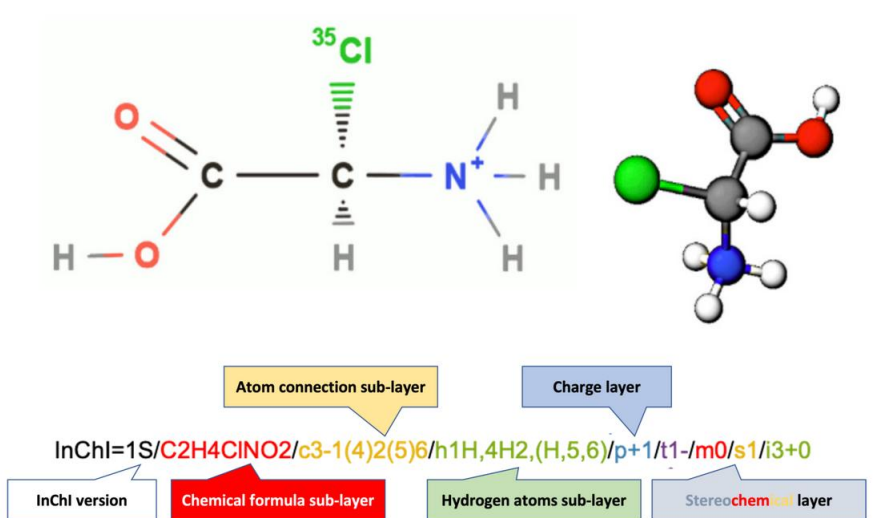


Source <https://www.inchi-trust.org/>

Standard InChIKey

Fixed-length, browser-searchable string derived from standard InChI.

Standard InChI



Standard InChIKey



InChIKey=UWPWWENWLZPQGU-WRFRXMDISA-O

Main layer

Stereochemistry & isotopes

version

protons

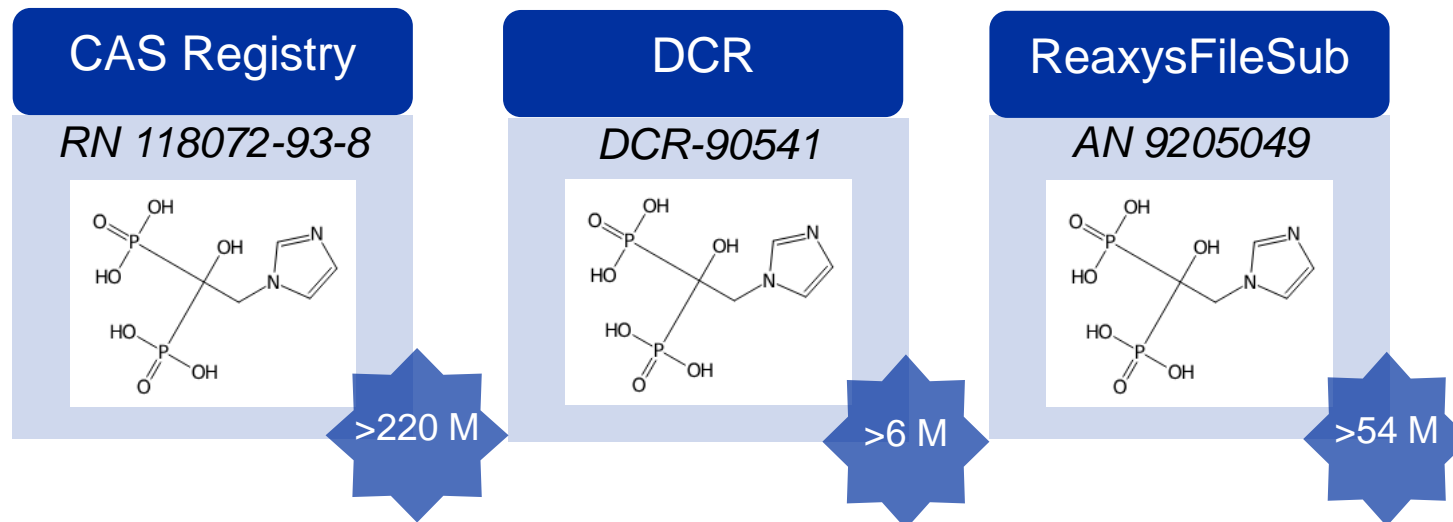
Source <https://www.inchi-trust.org/>



IMPLEMENTATION OF INCHI IN STNEXT DATABASES

STNext - InChI data in STNext databases

Major structure databases provide >280 million structure records with InChIs



INCH InChI=1S/C5H10N2O7P2/c8-5(15(9,10)11,16(12,13)14)3-7-2-1-6-4-7/h1-2,4,8H,3H2,(H2,9,10,11)(H2,12,13,14)

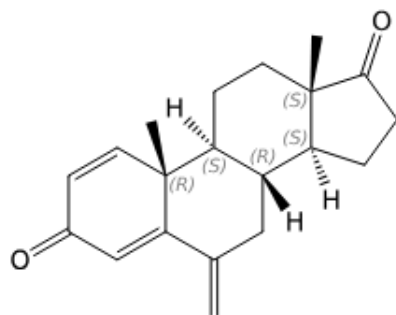
INKY XRASPMIURGNCCH-UHFFFAOYSA-N

INCH – InChI (IUPAC International Chemical Identifier); *INKY* - InChIKey

CAS Registry and DCR with InChIs and InChIKeys

IUPAC International Chemical Identifiers widely used since 20 years

RN **107868-30-4** REGISTRY
ED Entered STN: 02 May 1987
CN Androsta-1,4-diene-3,17-dione, 6-methylene- (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN 6-Methyleneandrosta-1,4-diene-3,17-dione
MF C20 H24 O2
INCH InChI=1S/C20H24O2/c1-12-10-14-15-4-5-18(22)20(15,3)9-7-16(14)19(2)8-6-13(21)11-17(12)19/h6,8,11,14-16H,1,4-5,7,9-10H2,2-3H3/t14-,15-,16-,19+,20-/m0/s1
INKY BFYIZQONLCFLEV-DAELLWK TSA-N



Absolute stereochemistry shown

InChIs are available in all standard display formats

InChI (INCH)

- Unique identifier created from chemical structure
- Depends on structure indexing rules of database producer

InChiKey (INKY)

- Hashed version of the full InChI
- Standardized to a length of 27 characters
- Facilitates database and Web searches

InChIKey retrieval capabilities

Layered design fully implemented on STNext

XUJNEKJLAYXESH-REOHCLBHSA-N

Core structure

Stereochemistry/Isotopes

Protonation state

=> S XUJNEKJLAYXESH-REOHCLBHSA-N/INKY

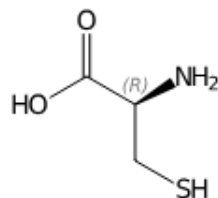
L1 2 XUJNEKJLAYXESH-REOHCLBHSA-N/INKY

=> S XUJNEKJLAYXESH-REOHCLBHSA/INKY

L2 11 XUJNEKJLAYXESH-REOHCLBHSA/INKY

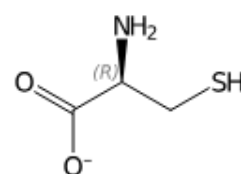
=> S XUJNEKJLAYXESH/INKY

L3 92 XUJNEKJLAYXESH/INKY



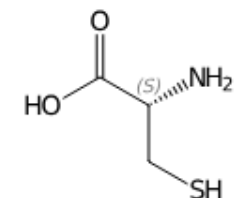
RN 52-90-4 L-cysteine

XUJNEKJLAYXESH-REOHCLBHSA-N



RN 19237-87-7 L-cysteine, ion

XUJNEKJLAYXESH-REOHCLBHSA-M



RN 921-01-7 D-cysteine

XUJNEKJLAYXESH-UWTATZPHSA-N

Value of InChI Keys on STNext

Useful cross database reference keys

- **Analyzing and sorting structure search results by InChIKeys**
 - Especially useful for evaluating substructure search results
- **Selective search**
 - For a specific stereoisomer or a specific isotope-labeled form of a chemical compound
 - Independent from chemical name variations and availability of stereospecific structure search

Collecting substance data from various resources

- From STNext substance databases
- From Internet resources
- **Identifying duplicate structures across databases**
 - Evaluation of multile structure search results and looking at unique chemical compounds

Searching InChI Keys on STNext

- **InChI Keys can be searched at three different levels**
 - Entire InChI Key
 - First two parts
 - First part only
- **InChI codes (or strings) cannot be searched, only displayed**

Searching InChI Keys on STNext

- InChI Keys can be searched at three different levels
 - Entire InChI Key – Core structure + Stereochemistry/Isotopes + Protonation site
 - First two parts – Core Structure + Stereochemistry/Isotopes
 - First part only – Core structure

Search run in
CAS REGISTRY.

```
=> S XUJNEKJLAYXESH-REOHCLBHSA-N/INKY  
  
L1          2 XUJNEKJLAYXESH-REOHCLBHSA-N/INKY  
  
=> S XUJNEKJLAYXESH-REOHCLBHSA/INKY  
  
L2          11 XUJNEKJLAYXESH-REOHCLBHSA/INKY  
  
=> S XUJNEKJLAYXESH/INKY  
  
L3          92 XUJNEKJLAYXESH/INKY
```



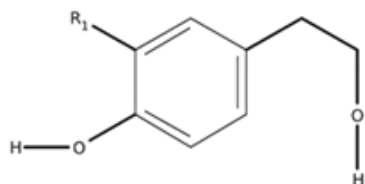
EXAMPLES

Example 1

Analysis of an answer set for core structures

=>

Uploading structure file: ...



R-Group Definitions

R1: O,S,H,F,Cl

Node Attributes

Ring Nodes : 1 2 3 4 5 6

Chain Nodes : 7 8 9 10 11 12

L1 STRUCTURE UPLOADED

=> S L1 CSS FUL

L2 8 SEA CSS FUL L1

R-group defined as 5 different atoms

Closed substructure search for structure with a variable group in DCR.

Example 1 (cont.)

Analysis of an answer set for core structures

```
=> ANA L2 INKY LEN14
```

```
L3          ANALYZE L2 1- INKY LEN 14 :          6 TERMS
```

ANALYZE of answer set for
core structures by using the
14-character INKY string

```
=> D 1-6
```

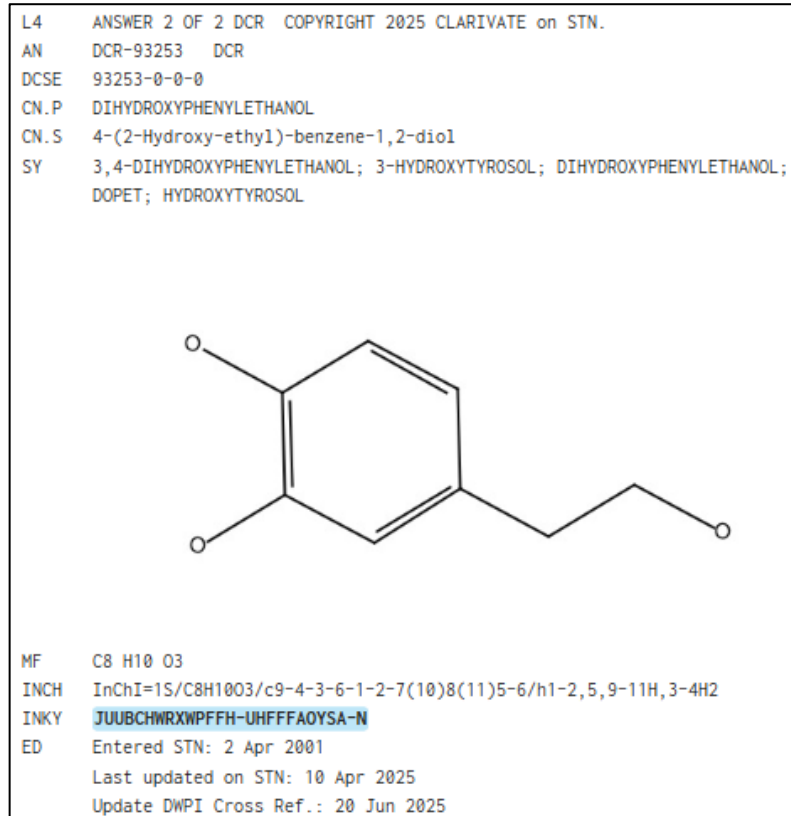
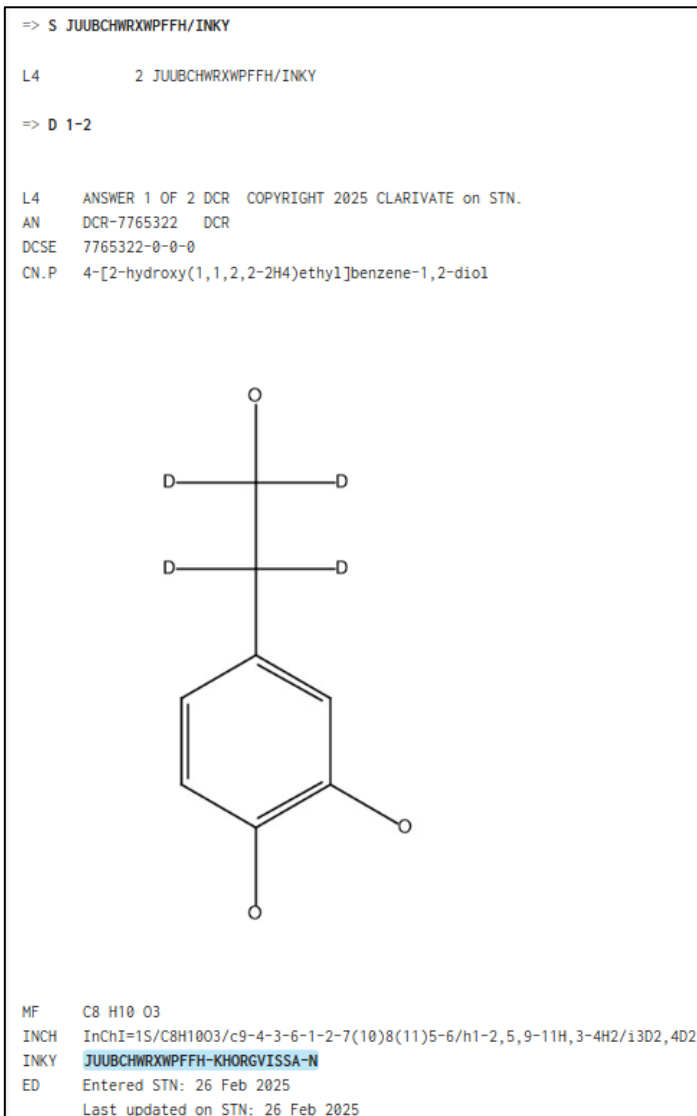
```
L3          ANALYZE L2 1- INKY LEN 14 :          6 TERMS
```

TERM #	# OCC	# DOC	% DOC	INKY
1	2	2	28.57	JUUBCHWRXWPFFH
2	2	2	28.57	YCCILVSKPBXVIP
3	1	1	12.50	PQWLPSXCNUCESA
4	1	1	14.29	TVZJWJYHPVHFGY
5	1	1	14.29	VPWOUSYKQMPMMJ
6	1	1	14.29	WYPBWLOUTCEOJL

***** END OF L3 *****

- InChIKeys (/INKY) are indexed and can also be analyzed in three different lengths which represent different degrees of precision.
- 14 characters: core structure (connectivity and bonds)
- 25 characters: core, stereochemistry, isotopes
- 27 characters: fully defined structure; core, stereochemistry, isotopes and charge.

Example 1 (cont.)



Example 1 (cont.)

=> S YCCILVSKPBXVIP/INKY

L5 2 YCCILVSKPBXVIP/INKY

=> D 1-2

L5 ANSWER 1 OF 2 DCR COPYRIGHT 2025 CLARIVATE on STN.

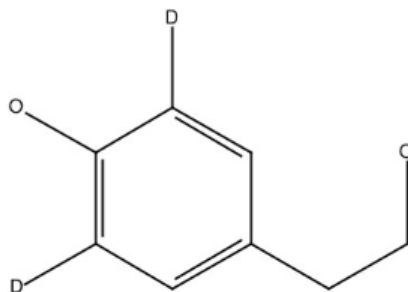
AN DCR-7623334 DCR

DCSE 9344-0-0-1

CN.P TYROSOL (BIS-2H)

CN.S 4-(2-HYDROXYETHYL)(2,6-XFFFDH2)PHENOL

SY TYROSOL (BIS-2H)



CMT isotope labeled, dideuterated

MF C8 H10 O2

INCH InChI=1S/C8H10O2/c9-6-5-7-1-3-8(10)4-2-7/h1-4,9-10H,5-6H2/13D,4D

INKY YCCILVSKPBXVIP-NMQOAUERSA-N

ED Entered STN: 11 Dec 2024

Last updated on STN: 26 Dec 2024

L5 ANSWER 2 OF 2 DCR COPYRIGHT 2025 CLARIVATE on STN.

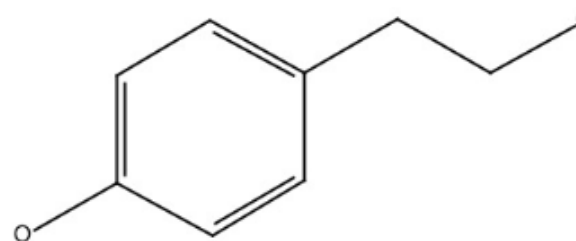
AN DCR-9344 DCR

DCSE 9344-0-0-0

CN.P TYROSOL

CN.S 4-(2-Hydroxy-ethyl)-phenol

SY 4-HYDROXYPHENETHYL ALCOHOL; HYDROXYETHYLPHENOL-2-4;
HYDROXYETHYLPHENOL-2-PARA; HYDROXYPHENETHYL-ALCOHOL-4;
HYDROXYPHENETHYL-ALCOHOL-PARA; HYDROXYPHENYLETHANOL-2-PARA;
HYDROXYPHENYLETHANOL-4-2; HYDROXYPHENYLETHANOL-4-BETA;
PARA-HYDROXYPHENYLETHANOL-BETA; TYROSOL; TYROSOL-PARA



MF C8 H10 O2

INCH InChI=1S/C8H10O2/c9-6-5-7-1-3-8(10)4-2-7/h1-4,9-10H,5-6H2

INKY YCCILVSKPBXVIP-UHFFFAOYSA-N

ED Entered STN: 13 Aug 1999

Last updated on STN: 10 Apr 2025

Update DWPI Cross Ref.: 18 Jun 2025

Example 2

Targeted search for a specific stereoisomer and its preparation

```
=> S D-PENICILLAMINE/CN
```

```
FILE 'DCR'
```

```
L6          0 D-PENICILLAMINE/CN
```

```
FILE 'REGISTRY'
```

```
L7          1 D-PENICILLAMINE/CN
```

```
FILE 'REAXYSFILESU'
```

```
L8          7 D-PENICILLAMINE/CN
```

Search for the chemical name of the stereoisomer D-penicillamine finds

- 0 hits in DCR
- 1 hit in REGISTRY
- 7 hits in REAXYSFILESUB

Example 2 (cont.)

ReaxysFileSub records

```
L8      ANSWER 1 OF 7 REAXYSFILESU COPYRIGHT 2025 ELSEVIER INC. on STN.  
AN      11346972      REAXYSFILESU  
CN      d-penicillamine; disodium EDTA; OTTENS Mixed Berry; phosphoric  
        acid; propylparaben; sodium methylparaben; sodium saccharine;  
SD      mixture (composition partially given)  
...  
L8      ANSWER 2 OF 7 REAXYSFILESU COPYRIGHT 2025 ELSEVIER INC. on STN.  
AN      11346973      REAXYSFILESU  
CN      d-penicillamine; disodium EDTA; phosphoric acid; propylparaben;  
        sodium methylparaben; sodium saccharine; water; purified;  
SD      mixture (composition partially given)  
...
```

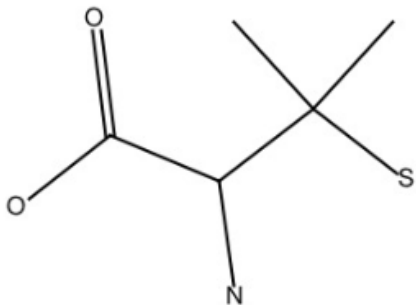
Mixtures of D-penicillamine

Multicomponent substances with D-penicillamine include the name of the component D-penicillamine in the chemical name.

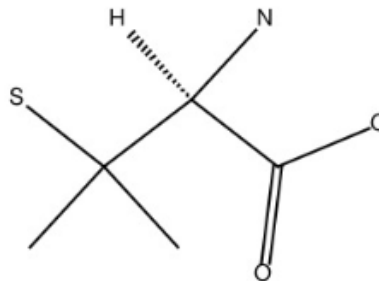
Example 2 (cont.)

3 ReaxysFileSub records of interest

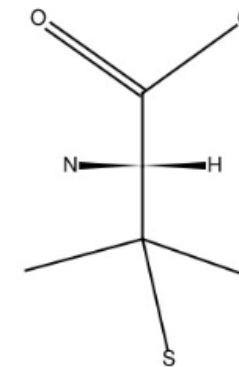
L13 ANSWER 5 OF 7 REAXYSFILESU COPYRIGHT 2025 ELSEVIER INC. on STN.
AN 4291203 REAXYSFILESU
CN **D-penicillamine**; 2-Amino-3-mercapto-3-methyl-butyric acid
SD acyclic
MF C5 H11 N O2 S
CMF C5 H11 N O2 S
LSF C5H11NO2S
INCHI VVNCNSJFMMFHPL-UHFFFAOYSA-N
AINCHI VVNCNSJFMMFHPL-QDQILVOLCP
MW 149.214
MARKREF.CNT 0
REC 13
ED Entered STN: 13 Jul 2020
Last updated on STN: 19 Jan 2024



L13 ANSWER 6 OF 7 REAXYSFILESU COPYRIGHT 2025 ELSEVIER INC. on STN.
AN 1722375 REAXYSFILESU
RN **52-67-5**
CN 3,3-dimethyl-D-cysteine; **D-penicillamine**; penicillamine; 3B2; -thiovaline; DPA
SD acyclic
MF C5 H11 N O2 S
CMF C5 H11 N O2 S
LSF (CH3)2(HS)CCH(NH2)COOH
INCHI VVNCNSJFMMFHPL-VKHYHEASA-N
AINCHI VVNCNSJFMMFHPL-GFBCBKCJDA
MW 149.214
MARKREF.CNT 2
REC 4713
ED Entered STN: 14 Jul 2020
Last updated on STN: 20 Jun 2025



L13 ANSWER 7 OF 7 REAXYSFILESU COPYRIGHT 2025 ELSEVIER INC. on STN.
AN 1722374 REAXYSFILESU
RN **1113-41-3**
CN L-penicillamine; **D-penicillamine**; penicillamine
SD acyclic
MF C5 H11 N O2 S
CMF C5 H11 N O2 S
LSF HSC(CH3)2CH(NH2)COOH
INCHI VVNCNSJFMMFHPL-GSVOUGTGS-A-N
AINCHI VVNCNSJFMMFHPL-ZKQKRMKDJ
MW 149.214
MARKREF.CNT 0
REC 893
ED Entered STN: 14 Jul 2020
Last updated on STN: 20 Jun 2025



Example 2 (cont.)

DCR records for InChI Keys from ReaxysFileSub

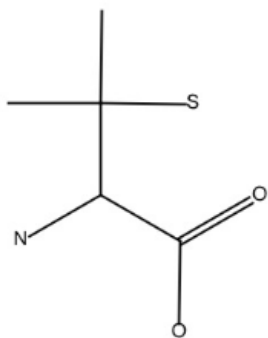
In this record, info on stereochemistry is given in the comment field.

=> S (VVNCNSJFMMFHPL-UHFFFAOYSA-N OR VVNCNSJFMMFHPL-VKHYHEASA-N OR VVNCNSJFMMFHPL-GSVOUGTGA-N)/INKY

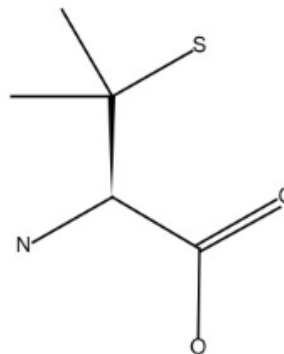
1 VVNCNSJFMMFHPL-UHFFFAOYSA-N/INKY
1 VVNCNSJFMMFHPL-VKHYHEASA-N/INKY
1 VVNCNSJFMMFHPL-GSVOUGTGA-N/INKY
L14 3 (VVNCNSJFMMFHPL-UHFFFAOYSA-N OR VVNCNSJFMMFHPL-VKHYHEASA-N OR VVNCNSJFMMFHPL-GSVOUGTGA-N)/INKY

=> D 1-3

L14 ANSWER 1 OF 3 DCR COPYRIGHT 2025 CLARIVATE on STN.
AN DCR-1150233 DCR
DCSE 8198-0-0-0
CN.P PENICILLAMINE
CN.S 2-Amino-3-mercapto-3-methyl-butyric acid
SY PENICILLAMINE

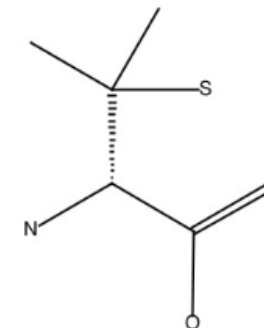


L14 ANSWER 2 OF 3 DCR COPYRIGHT 2025 CLARIVATE on STN.
AN DCR-961512 DCR
DCSE 8198-2-0-0
CN.P L-PENICILLAMINE
CN.S (R)-2-Amino-3-mercapto-3-methyl-butyric acid
SY L-PENICILLAMINE; PENICILLAMINE-L



MF C5 H11 N O2 S
INCH InChI=1S/C5H11NO2S/c1-5(2,9)3(6)4(7)8/h3,9H,6H2,1-2H3,(H,7,8)/t3-/m1/s1
INKY VVNCNSJFMMFHPL-GSVOUGTGA-N
ED Entered STN: 11 Oct 2004
Last updated on STN: 20 Jun 2025
Update DWPI Cross Ref.: 18 Jun 2025

L14 ANSWER 3 OF 3 DCR COPYRIGHT 2025 CLARIVATE on STN.
AN DCR-8198 DCR
DCSE 8198-1-0-0
CN.P PENICILLAMINE
SY PENICILLAMINE



CMT D-isomer

MF C5 H11 N O2 S
INCH InChI=1S/C5H11NO2S/c1-5(2,9)3(6)4(7)8/h3,9H,6H2,1-2H3,(H,7,8)/t3-/m0/s1
INKY VVNCNSJFMMFHPL-VKHYHEASA-N
ED Entered STN: 3 Jun 1999
Last updated on STN: 13 Apr 2022
Update DWPI Cross Ref.: 20 Jun 2025

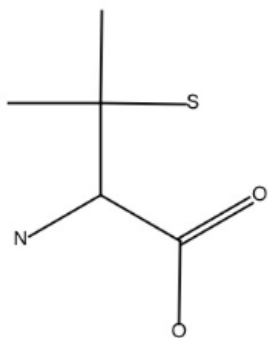
Example 2 (cont.)

DCR records for InChI Keys from ReaxysFileSub

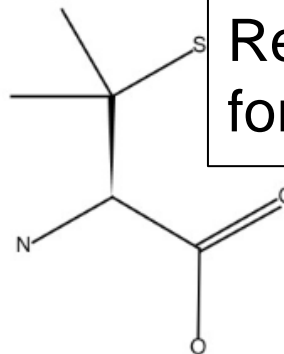
```
=> S (VVNCNSJFMMFHPL-UHFFFAOYSA-N OR VVNCNSJFMMFHPL-VKHYHEASA-N OR VVNCNSJFMMFHPL-GSVOUGTGA-N)/INKY  
  
1 VVNCNSJFMMFHPL-UHFFFAOYSA-N/INKY  
1 VVNCNSJFMMFHPL-VKHYHEASA-N/INKY  
1 VVNCNSJFMMFHPL-GSVOUGTGA-N/INKY  
L14 3 (VVNCNSJFMMFHPL-UHFFFAOYSA-N OR VVNCNSJFMMFHPL-VKHYHEASA-N OR  
VVNCNSJFMMFHPL-GSVOUGTGA-N)/INKY
```

=> D 1-3

```
L14 ANSWER 1 OF 3 DCR COPYRIGHT 2025 CLARIVATE on STN.  
AN DCR-1150233 DCR  
DCSE 8198-0-0-0  
CN.P PENICILLAMINE  
CN.S 2-Amino-3-mercapto-3-methyl-butyric acid  
SY PENICILLAMINE
```



```
L14 ANSWER 2 OF 3 DCR COPYRIGHT 2025 CLARIVATE on STN.  
AN DCR-961512 DCR  
DCSE 8198-2-0-0  
CN.P L-PENICILLAMINE  
CN.S (R)-2-Amino-3-mercapto-3-methyl-butyric acid  
SY L-PENICILLAMINE; PENICILLAMINE-L
```



```
MF C5 H11 N O2 S  
INCH InChI=1S/C5H11NO2S/c1-5(2,9)3(6)4(7)8/h3,9H,6H2,1-2H3,(H,7,8)/t3-/m1/s1  
INKY VVNCNSJFMMFHPL-GSVOUGTGA-N  
ED Entered STN: 11 Oct 2004  
Last updated on STN: 20 Jun 2025  
Update DWPI Cross Ref.: 18 Jun 2025
```

The corresponding
ReaxysFileSub record is
for D-Penicillamine.

```
L14 ANSWER 3 OF 3 DCR COPYRIGHT 2025 CLARIVATE on STN.  
AN DCR-8198 DCR  
DCSE 8198-1-0-0  
CN.P PENICILLAMINE  
SY PENICILLAMINE
```

```
CMT D-isomer  
MF C5 H11 N O2 S  
INCH InChI=1S/C5H11NO2S/c1-5(2,9)3(6)4(7)8/h3,9H,6H2,1-2H3,(H,7,8)/t3-/m0/s1  
INKY VVNCNSJFMMFHPL-VKHYHEASA-N  
ED Entered STN: 3 Jun 1999  
Last updated on STN: 13 Apr 2022  
Update DWPI Cross Ref.: 20 Jun 2025
```

Example 2 (cont.)

Corresponding DWPI search

=> FILE WPINDEX

=> S L14(T)(PRD OR P)/RL

882 L14

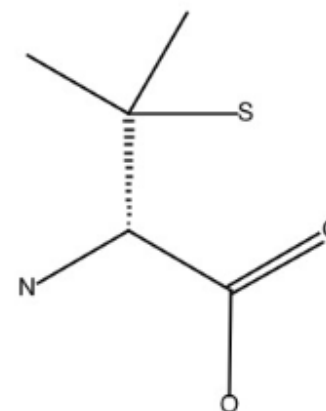
1161220 PRD/RL

1374864 P/RL

L15 45 L14(T)(PRD OR P)/RL

=> D BIB HITSTR 1-10

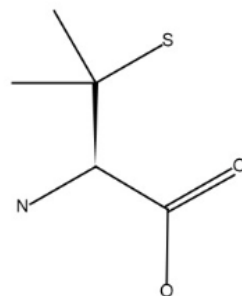
L15 ANSWER 2 OF 45 WPINDEX COPYRIGHT 2025 CLARIVATE on STN
AN 2024-B87529 [2024094] WPINDEX Full-text
TI Synthesizing D-penicillamine, involves taking penicillin G potassium salt as raw material, performing condensation and ring opening with mono-substituted hydrazine in mixed solvent under action of catalyst, adding acetic acid, acidifying, adding n-heptane, crystallizing, obtaining D-penicillamine
DC B02; J04
IN CUI C; HU Y; YANG M; YU S
PA (CASA-N) CAS ALDONE DALIAN PHARM TECHNOLOGY CO
CYC 1
PI CN 118851961 A 20241029 (2024094)* ZH
CN 118851961 B 20250124 (2025014) ZH
ADT CN 118851961 A CN 2024-10889777 20240703; CN 118851961 B CN 2024-10889777 20240703
FDT CN 118851961 B Previous Publ CN 118851961 A
PRAI CN 2024-10889777 20240703
AN.S DCR-8198
CN.P PENICILLAMINE
MF C5 H11 N O2 S
STR



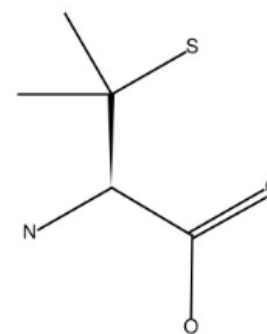
Example 2 (cont.)

Corresponding DWPI search

L15 ANSWER 3 OF 45 WPINDEX COPYRIGHT 2025 CLARIVATE on STN
AN 2024-B4750R [2024095] WPINDEX Full-text
TI Preparing L-penicillamine, comprises e.g. nitrating 3,3-dimethylacrylate, subjecting to Michael addition with thio reagent in presence of alkaline catalyst, reducing nitro group to amino group to obtain racemic penicillamine, resolving racemic by ester hydrolase, and performing alkaline hydrolysis
DC B05; D16
IN CHEN H; DONG C; HUANG J; MAN Y; WEI S; WEI Y; XIAO Y; XIONG Z; YUAN L; ZHANG G; ZHANG H
PA (ASYM-C) TIANJIN ASYMCHEM PHARM CO LTD
CYC 1
PI [CN 118834148](#) A 20241025 (2024095)* ZH
[CN 118834148](#) B 20250121 (2025010) ZH
ADT [CN 118834148](#) A [CN 2024-11316816](#) 20240920; [CN 118834148](#) B [CN 2024-11316816](#) 20240920
FDT [CN 118834148](#) B Previous Publ [CN 118834148](#) A
PRAI [CN 2024-11316816](#) 20240920
AN.S DCR-961512
CN.P L-PENICILLAMINE
CN.S (R)-2-Amino-3-mercapto-3-methyl-butyric acid
MF C5 H11 N O2 S
STR



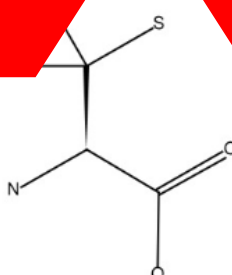
L15 ANSWER 6 OF 45 WPINDEX COPYRIGHT 2025 CLARIVATE on STN
AN 2023-991165 [2023081] WPINDEX Full-text
TI Producing L-penicillamine useful e.g. as synthetic raw material for pharmaceuticals involves cyclizing D-penicillamine, acting D-amino acid oxidase on cyclic D-penicillamine compound, reducing thiazoline compound and ring-opening compound
DC B05; D16
IN IWAMOTO Y; IWASAKI A
PA (KANF-C) KANEKA CORP
CYC 138
PI [WO 2023176487](#) A1 20230921 (2023081)* JA 68[0]
ADT [WO 2023176487](#) A1 [WO 2023-JP7888](#) 20230302
PRAI [JP 2022-184513](#) 20221118
[JP 2022-42906](#) 20220317
AN.S DCR-961512
CN.P L-PENICILLAMINE
CN.S (R)-2-Amino-3-mercapto-3-methyl-butyric acid
MF C5 H11 N O2 S
STR



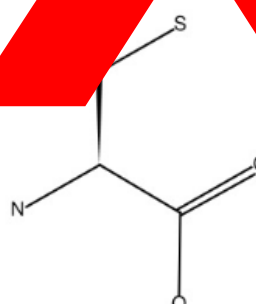
Example 2 (cont.)

Corresponding DWPI search

L15 ANSWER 3 OF 45 WPINDEX COPYRIGHT 2025 CLARIVATE on STN
AN 2024-B4750R [2024095] WPINDEX Full-text
TI Preparing L-penicillamine, comprises e.g. nitrating 3,3-dimethylacrylate, subjecting to Michael addition with thio reagent in presence of alkaline catalyst, reducing nitro group to amino group to obtain racemic penicillamine, resolving racemic by ester hydrolase, and performing alkaline hydrolysis
DC B05; D16
IN CHEN H; DONG C; HUANG J; MAN Y; WEI S; WEI Y; XIAO Y; XIONG Z; YUAN L; ZHANG G; ZHANG H
PA (AS) SYNGEN PHARM CO LTD
CYC 1
PI CN 118834148 20241025 (2024095) B CN 2024-11316816
ADT CN 118834148 20240920 CN 118834148 B CN 2024-11316816
FDT CN 118834148 B CN 2024-11316816
PRAI CN 2024-11316816
AN.S DCR-961512
CN.P L-PENICILLAMINE
CN.S (R)-2-Amino-3-mercaptopropionic acid
MF C5 H11 N O2 S
STR

C[C@H](CS)C(=O)O

L15 ANSWER 6 OF 45 WPINDEX COPYRIGHT 2025 CLARIVATE on STN
AN 2023-991165 [2023081] WPINDEX Full-text
TI Producing L-penicillamine useful e.g. as synthetic raw material for pharmaceuticals involves cyclizing D-penicillamine, acting D-amino acid oxidase on cyclic D-penicillamine compound, reducing thiazoline compound and ring-opening compound
DC B05; D16
IN IWAMOTO Y; IWASAKI A
PA (KANF-C) KANFKA CORP
CYC 138
PI WO 2023176487 2023081 (2023081) B CN 2024-11316816
ADT WO 2023176487 2023081 (2023081) B CN 2024-11316816
PRAI JP 2022-184513 2022081 (2022081) B CN 2024-11316816
JP 2022-42906 2022081 (2022081) B CN 2024-11316816
AN.S DCR-961512
CN.P L-PENICILLAMINE
CN.S (R)-2-Amino-3-mercaptopropionic acid
MF C5 H11 N O2 S
STR

C[C@H](CS)C(=O)O

These records were captured because the DCR record for L-penicillamine was included in the strategy.

Example 2 (cont.)

CAS REGISTRY records for InChI Keys from ReaxysFileSub

=> FILE REGISTRY

=> S (VVNCNSJFMMFHPL-UHFFFAOYSA-N OR VVNCNSJFMMFHPL-VKHYMYEASA-N OR VVNCNSJFMMFHPL-GSVOUGTGSA-N)/INKY

L16
1 VVNCNSJFMMFHPL-UHFFFAOYSA-N/INKY
2 VVNCNSJFMMFHPL-VKHYMYEASA-N/INKY
1 VVNCNSJFMMFHPL-GSVOUGTGSA-N/INKY
4 (VVNCNSJFMMFHPL-UHFFFAOYSA-N OR VVNCNSJFMMFHPL-VKHYMYEASA-N OR
VVNCNSJFMMFHPL-GSVOUGTGSA-N)/INKY

=> D 1-4

L16 ANSWER 1 OF 4 REGISTRY COPYRIGHT 2025 ACS on STN
RN 2920635-83-0 REGISTRY
ED Entered STN: 18 Apr 2023
CN INDEX NAME NOT YET ASSIGNED
FS STEREOSEARCH
MF (C5 H11 N O2 S)x
INCH InChI=1S/C5H11NO2S/c1-5(2,9)3(6)4(7)8/h3,9H,6H2,1-2H3,(H,7,8)/t3-/m0/s1
INKY VVNCNSJFMMFHPL-VKHYMYEASA-N
CI PMS
PCT Polyamide, Polyamide formed
SR CA
LC STN Files: CA, CAPLUS

L16 ANSWER 2 OF 4 REGISTRY COPYRIGHT 2025 ACS on STN
RN 1113-41-3 REGISTRY
ED Entered STN: 16 Nov 1984
CN L-Valine, 3-mercapto- (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN 3-Mercapto-L-valine
CN Valine, 3-mercapto-, L- (8CI)
OTHER NAMES:
CN (+)-Penicillamine
CN (2R)-2-Amino-3-methyl-3-sulfanylbutanoic acid
CN (R)-2-Amino-3-mercapto-3-methylbutanoic acid
CN (R)-Penicillamine
CN L-Penicillamine
CN NSC 241261
FS STEREOSEARCH
DR 16414-53-2, 73706-37-3
MF C5 H11 N O2 S
INCH InChI=1S/C5H11NO2S/c1-5(2,9)3(6)4(7)8/h3,9H,6H2,1-2H3,(H,7,8)/t3-/m1/s1
INKY VVNCNSJFMMFHPL-GSVOUGTGSA-N

Example 2 (cont.)

CAS REGISTRY records for InChI Keys from ReaxysFileSub

```
L16 ANSWER 3 OF 4 REGISTRY COPYRIGHT 2025 ACS on STN
RN 52-67-5 REGISTRY
ED Entered STN: 16 Nov 1984
CN D-Valine, 3-mercapto- (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN 3-Mercapto-D-valine
CN Valine, 3-mercapto-, D- (8CI)
OTHER NAMES:
CN (-)-Penicillamine
CN (2S)-2-Amino-3-methyl-3-sulfanylbutanoic acid
CN (2S)-2-Amino-3-methyl-3-sulfanylbutanoic acid
CN (2S)-2-Azaniumyl-3-methyl-3-sulfanylbutanoate
CN (S)-3,3-Dimethylcysteine
CN (S)-Penicillamine
CN  $\beta$ -Thiovaline
CN 2-Amino-3-mercapto-3-methylbutanoic acid
CN Artamin
CN Cuprenil
CN Cuprimine
CN Cupripen
CN D-(-)-Penicillamine
CN D-3-Mercaptovaline
CN D-Mercaptovaline
CN D-Penamine
CN D-Penicillamine
```

```
INCH InChI=1S/C5H11NO2S/c1-5(2,9)3(6)4(7)8/h3,9H,6H2,1-2H3,(H,7,8)/t3-/m0/s1
INKY VVNCNSJFMMFHPL-VKHYHEASA-N
```

```
L16 ANSWER 4 OF 4 REGISTRY COPYRIGHT 2025 ACS on STN
RN 52-66-4 REGISTRY
ED Entered STN: 16 Nov 1984
CN Valine, 3-mercapto- (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN 3-Mercaptovaline
CN DL-Valine, 3-mercapto-
CN Valine, 3-mercapto-, DL- (8CI)
OTHER NAMES:
CN ( $\pm$ )-Penicillamine
CN  $\beta,\beta$ -Dimethylcysteine
CN  $\beta$ -Mercaptovaline
CN  $\beta$ -Thiovaline
CN 2-Amino-3-mercapto-3-methylbutanoic acid
CN 2-Amino-3-mercapto-3-methylbutyric acid
CN 2-Amino-3-methyl-3-sulfanylbutanoic acid
CN 2-Azaniumyl-3-methyl-3-sulfanylbutanoate
CN Cysteine, 3,3-dimethyl-
CN DL- $\beta$ -Mercaptovaline
CN DL-3-Mercaptovaline
CN d1-DMC
CN DL-Penicillamine
CN DMC
CN NSC 22880
CN NSC 44656
DR 5139-02-6
MF C5 H11 N O2 S
INCH InChI=1S/C5H11NO2S/c1-5(2,9)3(6)4(7)8/h3,9H,6H2,1-2H3,(H,7,8)
INKY VVNCNSJFMMFHPL-UHFFFAOYSA-N
```

Example 2 (cont.)

Corresponding HCAplus search

=> S L16/PREP

6863 L16
9261531 PREP/RL
L17 302 L16/PREP
(L16 (L) PREP/RL)

=> D BIB HITSTR 1-10

L17 ANSWER 1 OF 302 HCAPLUS COPYRIGHT 2025 ACS on STN

[PatentPak PDF](#) | [PatentPak PDF+](#) | [PatentPak Interactive](#)

AN 2025:1078465 HCAPLUS Full-text

DN 191:9138

TI Green preparation D-penicillamine using acidic cationic resin

IN Chen, Xiaodong; Zhou, Liang; Chu, Changhu; Wu, Teng; Liu, Xuejun

PA Shanghai Neutan Pharma Research and Development Co., Ltd., Peop. Rep. China

SO Faming Zhuanli Shenqing, 7pp.

CODEN: CNXXEV

DT Patent

LA Chinese

FAN.CNT 1

PPPI

PATENT NO.	KIND	DATE	LANGUAGE	PatentPak
CN 119977889	A	20250513	Chinese	PDF PDF+ Interactive

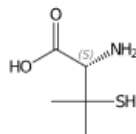
PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
CN 119977889	A	20250513	CN 2023-11495766	20231110

PRAI [CN 2023-11495766](#)

PSPI

PATENT NO.	KIND	STATUS	STATUS DATE
CN 119977889	A	Alive	20250522

IT [52-67-5 P](#)
RL: IMF (Industrial manufacture); SCLM (Substance claimed); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); [PREP](#) ([Preparation](#)); USES (Uses)
(green prepn. D-penicillamine using acidic cationic resin)
RN [52-67-5](#) HCAPLUS
CN D-Valine, 3-mercapto- (CA INDEX NAME)



Absolute stereochemistry shown

L17 ANSWER 2 OF 302 HCAPLUS COPYRIGHT 2025 ACS on STN

AN 2024:2594088 HCAPLUS Full-text

TI Self-assembly of differently charged trimesic based lithocholic amphiphiles and their assessment on antimicrobial and biostimulant properties

AU Ezhumalai, Nishanthi; Panchalingam, Santhiya; Govindaraju, Kasivelu; Kannan, Malaichamy; Kasthuri, Jayapalan; Rajendiran, Nagappan

CS Department of Polymer Science, Guindy Campus, University of Madras, Tamil Nadu, 600025, India

SO Colloids and Surfaces, B: Biointerfaces (2025), 246, 114391

CODEN: CSBBEQ; ISSN: 0927-7765

DOI 10.1016/j.colsurfb.2024.114391

PB Elsevier B.V.

DT Journal; (online computer file)

LA English

IT INDEXING IN PROGRESS

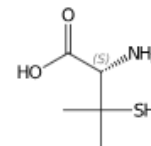
IT [52-67-5 P](#)

RL: AGR (Agricultural use); BPN (Biosynthetic preparation); PRP (Properties); RCT (Reactant); RGT (Reagent); BIOL (Biological study); [PREP](#) ([Preparation](#)); RACT (Reactant or reagent); USES (Uses)

(trimesic based lithocholic ester functionalized zwitterionic; self-assembly of differently charged trimesic based lithocholic amphiphiles and their assessment on antimicrobial and biostimulant properties)

RN [52-67-5](#) HCAPLUS

CN D-Valine, 3-mercapto- (CA INDEX NAME)



Absolute stereochemistry shown

OSC.G 2 THERE ARE 2 CAPLUS RECORDS THAT CITE THIS RECORD (2 CITINGS)
RE.CNT 50 THERE ARE 50 CITED REFERENCES AVAILABLE FOR THIS RECORD
ALL CITATIONS AVAILABLE IN THE RE FORMAT

Example 2 (cont.)

Corresponding HCAplus search

L17 ANSWER 5 OF 302 HCAPLUS COPYRIGHT 2025 ACS on STN
PatentPak PDF | PatentPak PDF+ | PatentPak Interactive
AN 2024:2295125 HCAPLUS Full-text
DN 188:355886
TI Preparation of L-penicillamine
IN Xiong, Zhengchang; Xiao, Yi; Zhang, Haiming; Yuan, Long; Dong, Changming;
Wei, Yiran; Chen, Hui; Wei, Shanshan; Zhang, Guangwei; Man, Yi; Huang, Jie
PA Tianjin Asymchem Pharmaceutical Co., Ltd., Peop. Rep. China
UO ASYMCHEM INC.
UOS Asymchem
SO Faming Zhuanli Shenqing, 26pp.
CODEN: CNXXEV
DT Patent
LA Chinese
FAN.CNT 1
PPPI

PATENT NO.	KIND	DATE	LANGUAGE	PatentPak
CN 118834148	A	20241025	Chinese	PDF PDF+ Interactive
CN 118834148	B	20250121	Chinese	PDF

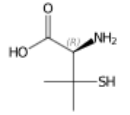
PI

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
CN 118834148	A	20241025	CN 2024-11316816	20240920
CN 118834148	B	20250121		
PRAI CN 2024-11316816		20240920		

PSPI

PATENT NO.	KIND	STATUS	STATUS DATE
CN 118834148	A	Alive	20241031
CN 118834148	B	Alive	20241031

IT 1113-41-3 P
RL: SCLM (Substance claimed); SPN (Synthetic preparation); PREP (Preparation)
(prepn. of L-penicillamine)
RN 1113-41-3 HCAPLUS
CN L-Valine, 3-mercapto- (CA INDEX NAME)



Absolute stereochemistry shown
Rotation (+)

L17 ANSWER 7 OF 302 HCAPLUS COPYRIGHT 2025 ACS on STN
PatentPak PDF | PatentPak PDF+ | PatentPak Interactive
AN 2024:1464016 HCAPLUS Full-text
DN 187:205681
TI Cellulose ether and preparation method thereof
IN Xing, Jianlei; Jiang, Shuzheng; Yu, Xiaoliang; Jiao, Ningning; Xie, Haoran
PA Wuqiao Qiyuan Cellulose Co., Ltd., Peop. Rep. China
SO Faming Zhuanli Shenqing, 9pp.
CODEN: CNXXEV
DT Patent
LA Chinese
FAN.CNT 1
PPPI

PATENT NO.	KIND	DATE	LANGUAGE	PatentPak
CN 118290656	A	20240705	Chinese	PDF PDF+ Interactive
CN 118290656	B	20241101	Chinese	PDF

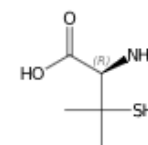
PI

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
CN 118290656	A	20240705	CN 2024-10492788	20240423
CN 118290656	B	20241101		
PRAI CN 2024-10492788		20240423		

PSPI

PATENT NO.	KIND	STATUS	STATUS DATE
CN 118290656	A	Alive	20240718
CN 118290656	B	Alive	20240718

IT 1113-41-3 DP, reaction products with dialdehyde CM-cellulose
RL: IMF (Industrial manufacture); SCLM (Substance claimed); TEM (Technical or engineered material use); PREP (Preparation); USES (Uses)
(prepn. method of cellulose ether)
RN 1113-41-3 HCAPLUS
CN L-Valine, 3-mercapto- (CA INDEX NAME)



Absolute stereochemistry shown
Rotation (+)

Example 2 (cont.)

Corresponding HCAplus search

L17 ANSWER 5 OF 302 HCAPLUS COPYRIGHT 2025 ACS on STN
PatentPak PDF | PatentPak PDF+ | PatentPak Interactive
AN 2024:2295125 HCAPLUS Full-text
DN 188:355886
TI Preparation of L-penicillamine
IN Xiong, Zhengchang; Xiao, Yi; Zhang, Haiming; Yuan, Long; Dong, Changming;
Wei, Yiran; Chen, Hui; Wei, Shanshan; Zhang, Guangwei; Man, Yi; Huang, Jie
PA Tianjin Asymchem Pharmaceutical Co., Ltd., Peop. Rep. China
UO ASYMC
UOS Asymche
SO Faming Zh
CODEN: CNXX
DT Patent
LA Chinese
FAN.CNT 1
PPPI

PATENT NO.	DATE	LANGUAGE	PatentPak
CN 118834148	20241021	Chinese	PDF PDF+ Interactive
CN 118834148	20241021	Chinese	PDF

PI

PATENT NO.	SECTION NO.	DATE
CN 118834148	1	20241021
CN 118834148	2	20250121
CN 2024-113	1	20240920

PRAI CN 2024-113

PSPI

PATENT NO.	IND	STATUS	DATE
CN 118834148	A	Alive	20241021
CN 118834148	B	Alive	20241031

IT 1113-41-3 P

RL: SCLM (Substance claimed); SPN (Synthetic preparation); **PREP** (Preparation)

(prepn. of L-penicillamine)

RN 1113-41-3 HCAPLUS

CN L-Valine, 3-mercapto- (CA INDEX NAME)

NC(C(C(=O)O)S)C

Absolute stereochemistry shown
Rotation (+)

L17 ANSWER 7 OF 302 HCAPLUS COPYRIGHT 2025 ACS on STN
PatentPak PDF | PatentPak PDF+ | PatentPak Interactive
AN 2024:1464016 HCAPLUS Full-text
DN 187:205681
TI Cellulose ether and preparation method thereof
IN Xing, Jianlei; Jiang, Shuzheng; Yu, Xiaoliang; Jiao, Ningning; Xie, Haoran
PA Wuqiao Qiyuan Cellulose Co., Ltd., Peop. Rep. China
SO Faming Zhuanli Shenqing, 9pp.
CODEN: CNXXEV
DT Patent
LA Chinese
FAN.CNT 1
PPPI

PATENT NO.	DATE	LANGUAGE	PatentPak
CN 118290656	20240705	Chinese	PDF+ Interactive
CN 118290656	20241101	Chinese	

PI

PATENT NO.	SECTION NO.	DATE
CN 118290656	A	2024-10492788
CN 118290656	B	20240423
CN 2024-10492788		

PRAI CN 2024-10492788

PSPI

PATENT NO.	DATE	STATUS	DATE
CN 118290656	20240705	Alive	20240705
CN 118290656	20241101	Alive	20241101

IT 1113-41-3 DP, products with dia... cellulose

RL: IMF (Industrial manufacture); SCLM (Substance claimed); TEM (Technical or engineering use); **PREP** (Preparation)

(prepn. method of cellulose ether)

RN 1113-41-3 HCAPLUS

CN L-Valine, 3-mercapto- (CA INDEX NAME)

NC(C(C(=O)O)S)C

Absolute stereochemistry shown
Rotation (+)

Example 2 (cont.)

Rerun DCR/WPINDEX search using just one InChI Key

```
=> FILE DCR
```

```
=> S VVNCNSJFMMFHPL-VKHYHEASA-N/INKY
```

```
L18          1 VVNCNSJFMMFHPL-VKHYHEASA-N/INKY
```

```
=> FILE WPINDEX
```

```
=> S L18(T)(PRD OR P)/RL
```

```
538 L18
```

```
1161220 PRD/RL
```

```
1374864 P/RL
```

```
L19          30 L18(T)(PRD OR P)/RL
```

Original search
retrieved 45 records.

Example 2 (cont.)

Rerun CAS REGISTRY/HCAplus search using just one InChI Key

```
=> FILE REGISTRY
```

```
=> S VVNCNSJFMMFHPL-VKHYHEASA-N/INKY
```

```
L20          2 VVNCNSJFMMFHPL-VKHYHEASA-N/INKY
```

```
=> FILE HCAPLUS
```

```
=> S L20/PREP
```

```
5952 L20
```

```
9261531 PREP/RL
```

```
L21          206 L20/PREP
```

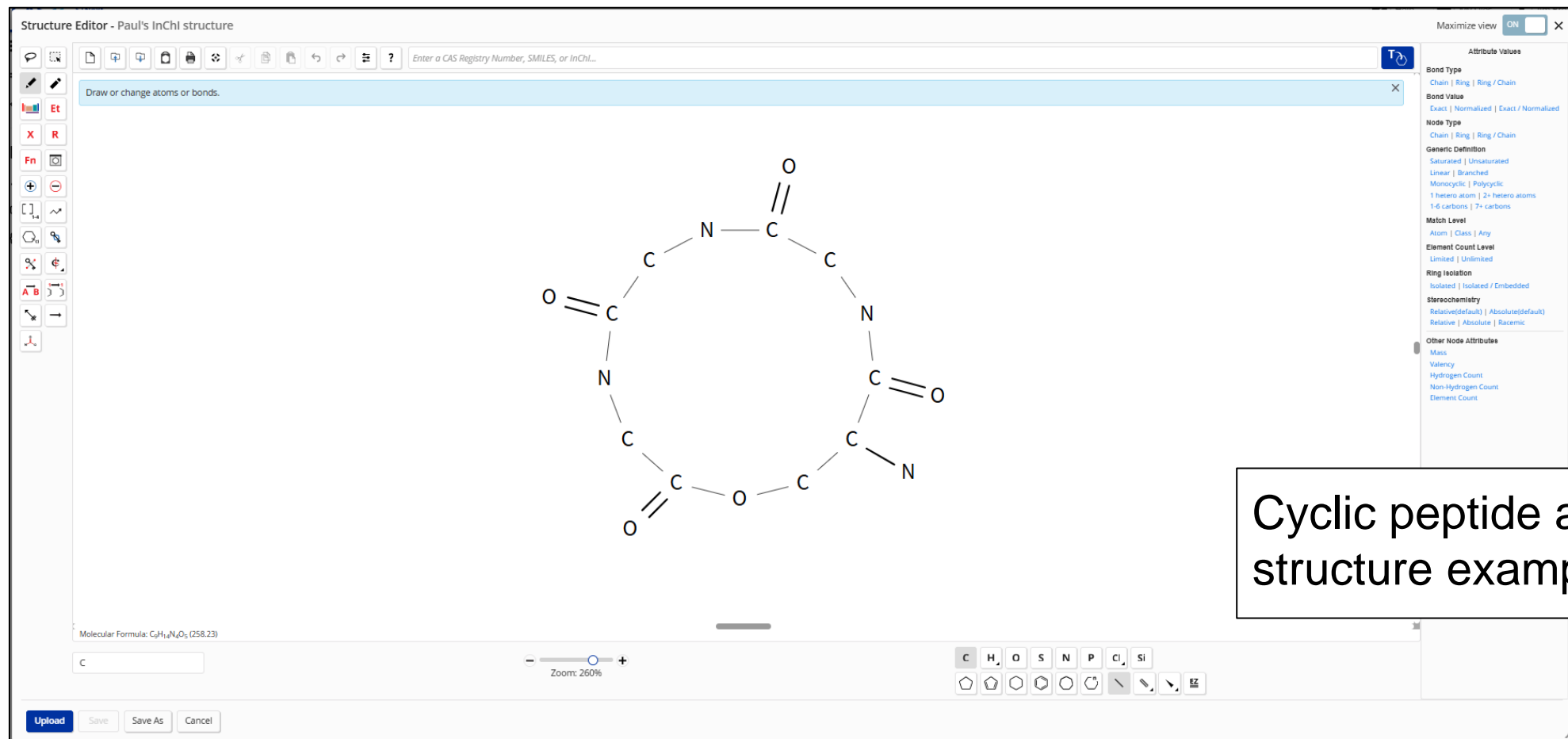
```
(L20 (L) PREP/RL)
```

Monomer and polymer have the same InChI Key due to the monomer-based indexing.

Original search retrieved 302 records.

Example 3

Find unique substance records across REGISTRY, DCR & REAXYSFILESUB



Example 3 (cont.)

Find unique substance records across REGISTRY, DCR & REAXYSFILESUB

```
=> S L1 SSS FULL

FULL SEARCH INITIATED 14:11:43
FULL SCREEN SEARCH COMPLETED - 12917 TO ITERATE

100.0% PROCESSED      12917 ITERATIONS      1029 ANSWERS
SEARCH TIME: 00.00.01

L2      1029 SEA SSS FUL L1

=> ANA L2 INKY 1-

L3      ANALYZE L2 1- INKY : 1022 TERMS

=> D

L3      ANALYZE L2 1- INKY : 1022 TERMS

TERM #   # OCC   # DOC   % DOC INKY
-----
1         2       2    0.19 OXKLHOSMVDGTBS-UHFFFAOYSA-N
2         2       2    0.19 QSPKBORZTSNKNP-MMMVMMOASA-N
3         2       2    0.19 RBHIOSHGUQQZQF-MMMVMMOASA-N
4         2       2    0.19 SGGJJTTZBRPIKP-UHFFFAOYSA-N
5         1       1    0.10 ABCJMLIOAYXJJP-HFWKHEOBSA-N
6         1       1    0.10 ACYGACFVUBXFPG-JFAYZGHXSA-N
7         1       1    0.10 ADNGAVBKDGODMA-UGFNBYPTSA-N
8         1       1    0.10 ADVMDYNGUKYGDK-FAFIKCBISA-N
9         1       1    0.10 AHTNGRTYKCUSSD-TWMHPYOZSA-N
10        1       1    0.10 AHUYNAIWJQEFGX-FFBZIVEISA-N
1012 MORE TERMS WITH AN OCCURRENCE COUNT OF 1
```

Search is run in CAS
REGISTRY.

Example 3 (cont.)

Find unique substance records across REGISTRY, DCR & REAXYSFILESUB

```
=> S L1 SSS FULL

FULL SEARCH INITIATED 14:18:11
FULL SCREEN SEARCH COMPLETED -      0 TO ITERATE

    0.0% PROCESSED    6263191 ITERATIONS                    141 ANSWERS
SEARCH TIME: 00.00.01

L4          141 SEA SSS FUL L1

=> ANA L4 INKY 1-

L5          ANALYZE L4 1- INKY :    139 TERMS

=> D

L5          ANALYZE L4 1- INKY :    139 TERMS

TERM #      # OCC  # DOC  % DOC INKY
-----
    1         2     2   1.42 HJEBQUNJIQPONN-ANVYOYFISA-N
    2         2     2   1.42 HJEBQUNJIQPONN-KTURZQIXSA-N
    3         1     1   0.71 AENANIALMPCGPR-OTRWZRNCSA-N
    4         1     1   0.71 AERTVQZWEKLOH-KUPAKDCRSA-N
    5         1     1   0.71 AJIQHWJOMQKGQ-FMCGYTSFSA-N
    6         1     1   0.71 AVVXSZZSJQUPHA-QWZOZRLTSA-N
    7         1     1   0.71 AXMVMQQOBZPNAY-FFRZOQKGSAN-N
    8         1     1   0.71 BEJRLMHCSIOJDI-IDPDROOQSA-N
    9         1     1   0.71 BFEKVXPJQSHMSY-YJZKUPLCSA-N
   10         1     1   0.71 BJWIDOUFUOSKJW-ABBHOYPZSA-N
129 MORE TERMS WITH AN OCCURRENCE COUNT OF 1
```

Search is run in DCR.

Example 3 (cont.)

Find unique substance records across REGISTRY, DCR & REAXYSFILESUB

```
=> S L1 SSS FULL

FULL SEARCH INITIATED 14:23:04
FULL SCREEN SEARCH COMPLETED -      0 TO ITERATE

    0.0% PROCESSED   54922368 ITERATIONS                1201 ANSWERS
SEARCH TIME: 00.00.02

L6          1201 SEA SSS FUL L1

=> ANA L6 INKY 1-

L7          ANALYZE L6 1- INKY :    1167 TERMS

=> D

L7          ANALYZE L6 1- INKY :    1167 TERMS

TERM #    # OCC  # DOC  % DOC INKY
-----
   1         3     3   0.25 DXMMUZNSJCZZLR-LOYRBTAXSA-N
   2         3     3   0.25 ZRZMGQJQHXS-AFRDPKMSA-N
   3         2     2   0.17 HCIYRJZEGWHZMT-PUDJQCGISA-N
   4         2     2   0.17 LMBFAGIMSUYTBN-MPZNNTNKS-A-N
   5         2     2   0.17 MEAXCSPGSRHSOY-CNZUMQHHS-A-N
   6         2     2   0.17 NUNXIPWJBFWITR-UPQQPLLJSA-N
   7         2     2   0.17 PZQLEZBFSMMAHY-LNHLIAMYSA-N
   8         2     2   0.17 QFEWDDBTSSPCY-BZKBKYEGB-A-N
   9         2     2   0.17 QHJMQPOXERDDJU-YYMGAGDFSA-N
  10         2     2   0.17 QSPKBORZTSNKNP-MMMVMMOASA-N
  11         2     2   0.17 RDSXDSVHMBEZXH-ZTHUPLLRSA-N
  12         2     2   0.17 RGMTULDDXQGTSL-PUDJQCGISA-N
  13         2     2   0.17 UARCSGORGPIBZ-ZTHUPLLRSA-N
  14         2     2   0.17 YHMOPLGVQZVMEH-APIHUWKFS-A-N
  15         2     2   0.17 YYIFRWLTUAIOT-CNICKXRFSA-N
```

Search is run in ReaxysFileSub.

Example 3 (cont.)

Find unique substance records across REGISTRY, DCR & REAXYSFILESUB

```
=> S L6 NOT (L3 OR L5)
```

```
L9      QUE  TERMS FROM L3 WITH NO HITS:   469 TERMS
```

```
L11     QUE  TERMS FROM L5 WITH NO HITS:   130 TERMS
```

```
L12     645 L6 NOT (L8 OR L10)
```

```
=> D
```

```
L12     ANSWER 1 OF 645 REAXYSFILESU COPYRIGHT 2025 ELSEVIER INC. on STN.
```

```
AN      63390503  REAXYSFILESU
```

```
CN      Ile(sub)2(/sub)(sup)S(/sup)-Leu(sub)10(/sub)-teixobactin
```

```
MF      C58 H96 N12 O14 S
```

```
CMF     C58 H96 N12 O14 S
```

```
LSF     C58H96N12O14S
```

```
INCHI   OHQJALSJOZTACI-ATPZDXELSA-N
```

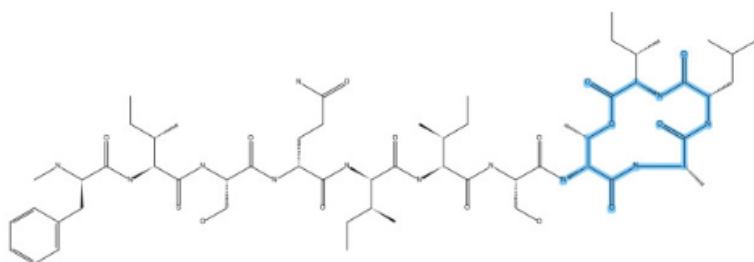
```
MW      1217.54
```

```
MARKREF.CNT 0
```

```
REC      1
```

```
ED      Entered STN: 13 Jun 2025
```

```
        Last updated on STN: 13 Jun 2025
```



L3 = REG set ANALYZEd for InChI Keys.
L5 = DCR set ANALYZEd for InChI Keys.

Unique record from ReaxysFileSub.

Example 3 (cont.)

Find unique substance records across REGISTRY, DCR & REAXYSFILESUB

=> S L4 NOT (L3 OR L7)

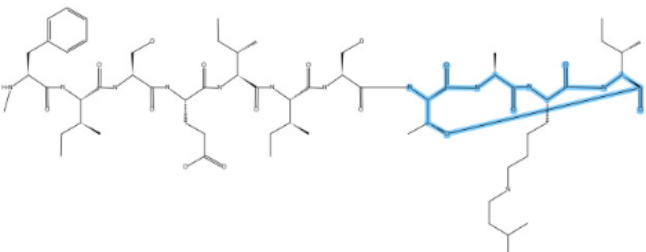
L14 QUE TERMS FROM L3 WITH NO HITS: 1011 TERMS

L16 QUE TERMS FROM L7 WITH NO HITS: 1158 TERMS

L17 128 L4 NOT (L13 OR L15)

=> D

L17 ANSWER 1 OF 128 DCR COPYRIGHT 2025 CLARIVATE on STN.
AN DCR-7664927 DCR
DCSE 7664927-1-0-0



MF C63 H106 N12 O16

INCH InChI=1S/C63H106N12O16/c1-14-35(7)48(72-56(83)44(64-13)31-41-23-19-18-20-24-41)59(86)69-45(32-76)57(84)68-43(26-27-47(78)79)55(82)71-50(37(9)16-3)61(88)73-49(36(8)15-2)60(87)70-46(33-77)58(85)75-52-40(12)91-63(90)51(38(10)17-4)74-54(81)42(67-53(80)39(11)66-62(52)89)25-21-22-29-65-30-28-34(5)6/h18-20,23-24,34-40,42-46,48-52,64-65,76-77H,14-17,21-22,25-33H2,1-13H3,(H,66,89)(H,67,80)(H,68,84)(H,69,86)(H,70,87)(H,71,82)(H,72,83)(H,73,88)(H,74,81)(H,75,85)(H,78,79)/t35-,36-,37-,38-,39-,40+,42-,43-,44-,45-,46-,48-,49-,50-,51-,52-/m0/s1

INKY NGFYWFKKUYEPFF-RDZWRUAUSA-N

ED Entered STN: 10 Jan 2025
Last updated on STN: 10 Jan 2025
Update DWPI Cross Ref.: 10 Jan 2025

L3 = REG set ANALYZEd for InChI Keys.
L7 = ReaxysFileSub set ANALYZEd for InChI Keys.

Unique record from DCR.

Example 3 (cont.)

Find unique substance records across REGISTRY, DCR & REAXYSFILESUB

=> S L2 NOT (L5 OR L7)

L19 QUE TERMS FROM L5 WITH NO HITS: 128 TERMS

L21 QUE TERMS FROM L7 WITH NO HITS: 614 TERMS

L22 471 L2 NOT (L18 OR L20)

=> D

L22 ANSWER 1 OF 471 REGISTRY COPYRIGHT 2025 ACS on STN

RN **3082422-60-1** REGISTRY

ED Entered STN: 09 May 2025

CN INDEX NAME NOT YET ASSIGNED

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C74 H105 Br N18 O16

INCH InChI=1S/C74H105BrN18O16/c1-6-40(3)22-27-60(96)85-57(34-59(79)95)70(104)89-54(31-43-23-25-46(75)26-24-43)68(102)91-55(32-44-35-80-49-17-10-8-15-47(44)49)69(103)88-53(21-14-30-78)67(101)87-51(19-12-28-76)65(99)82-37-61(97)84-52(20-13-29-77)66(100)90-56(33-45-36-81-50-18-11-9-16-48(45)50)71(105)93-64-42(5)109-74(108)58(39-94)86-62(98)38-83-72(106)63(41(4)7-2)92-73(64)107/h8-11,15-18,23-26,35-36,40-42,51-58,63-64,80-81,94H,6-7,12-14,19-22,27-34,37,5H3,(H2,79,95)(H,82,99)(H,83,106)(H,88,103)(H,89,104)(H,90,100)(H,91,102)(H,92,93,94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115,116,117,118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139,140,141,142,143,144,145,146,147,148,149,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,172,173,174,175,176,177,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194,195,196,197,198,199,200,201,202,203,204,205,206,207,208,209,210,211,212,213,214,215,216,217,218,219,220,221,222,223,224,225,226,227,228,229,230,231,232,233,234,235,236,237,238,239,240,241,242,243,244,245,246,247,248,249,250,251,252,253,254,255,256,257,258,259,260,261,262,263,264,265,266,267,268,269,270,271,272,273,274,275,276,277,278,279,280,281,282,283,284,285,286,287,288,289,290,291,292,293,294,295,296,297,298,299,300,301,302,303,304,305,306,307,308,309,310,311,312,313,314,315,316,317,318,319,320,321,322,323,324,325,326,327,328,329,330,331,332,333,334,335,336,337,338,339,340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357,358,359,360,361,362,363,364,365,366,367,368,369,370,371,372,373,374,375,376,377,378,379,380,381,382,383,384,385,386,387,388,389,390,391,392,393,394,395,396,397,398,399,400,401,402,403,404,405,406,407,408,409,410,411,412,413,414,415,416,417,418,419,420,421,422,423,424,425,426,427,428,429,430,431,432,433,434,435,436,437,438,439,440,441,442,443,444,445,446,447,448,449,450,451,452,453,454,455,456,457,458,459,460,461,462,463,464,465,466,467,468,469,470,471,472,473,474,475,476,477,478,479,480,481,482,483,484,485,486,487,488,489,490,491,492,493,494,495,496,497,498,499,500,501,502,503,504,505,506,507,508,509,510,511,512,513,514,515,516,517,518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,535,536,537,538,539,540,541,542,543,544,545,546,547,548,549,550,551,552,553,554,555,556,557,558,559,560,561,562,563,564,565,566,567,568,569,570,571,572,573,574,575,576,577,578,579,580,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,597,598,599,600,601,602,603,604,605,606,607,608,609,610,611,612,613,614,615,616,617,618,619,620,621,622,623,624,625,626,627,628,629,630,631,632,633,634,635,636,637,638,639,640,641,642,643,644,645,646,647,648,649,650,651,652,653,654,655,656,657,658,659,660,661,662,663,664,665,666,667,668,669,670,671,672,673,674,675,676,677,678,679,680,681,682,683,684,685,686,687,688,689,690,691,692,693,694,695,696,697,698,699,700,701,702,703,704,705,706,707,708,709,710,711,712,713,714,715,716,717,718,719,720,721,722,723,724,725,726,727,728,729,730,731,732,733,734,735,736,737,738,739,740,741,742,743,744,745,746,747,748,749,750,751,752,753,754,755,756,757,758,759,760,761,762,763,764,765,766,767,768,769,770,771,772,773,774,775,776,777,778,779,780,781,782,783,784,785,786,787,788,789,790,791,792,793,794,795,796,797,798,799,800,801,802,803,804,805,806,807,808,809,810,811,812,813,814,815,816,817,818,819,820,821,822,823,824,825,826,827,828,829,830,831,832,833,834,835,836,837,838,839,840,841,842,843,844,845,846,847,848,849,850,851,852,853,854,855,856,857,858,859,860,861,862,863,864,865,866,867,868,869,870,871,872,873,874,875,876,877,878,879,880,881,882,883,884,885,886,887,888,889,890,891,892,893,894,895,896,897,898,899,900,901,902,903,904,905,906,907,908,909,910,911,912,913,914,915,916,917,918,919,920,921,922,923,924,925,926,927,928,929,930,931,932,933,934,935,936,937,938,939,940,941,942,943,944,945,946,947,948,949,950,951,952,953,954,955,956,957,958,959,960,961,962,963,964,965,966,967,968,969,970,971,972,973,974,975,976,977,978,979,980,981,982,983,984,985,986,987,988,989,990,991,992,993,994,995,996,997,998,999,1000,1001,1002,1003,1004,1005,1006,1007,1008,1009,1010,1011,1012,1013,1014,1015,1016,1017,1018,1019,1020,1021,1022,1023,1024,1025,1026,1027,1028,1029,1030,1031,1032,1033,1034,1035,1036,1037,1038,1039,1040,1041,1042,1043,1044,1045,1046,1047,1048,1049,1050,1051,1052,1053,1054,1055,1056,1057,1058,1059,1060,1061,1062,1063,1064,1065,1066,1067,1068,1069,1070,1071,1072,1073,1074,1075,1076,1077,1078,1079,1080,1081,1082,1083,1084,1085,1086,1087,1088,1089,1090,1091,1092,1093,1094,1095,1096,1097,1098,1099,1100,1101,1102,1103,1104,1105,1106,1107,1108,1109,1110,1111,1112,1113,1114,1115,1116,1117,1118,1119,1120,1121,1122,1123,1124,1125,1126,1127,1128,1129,1130,1131,1132,1133,1134,1135,1136,1137,1138,1139,1140,1141,1142,1143,1144,1145,1146,1147,1148,1149,1150,1151,1152,1153,1154,1155,1156,1157,1158,1159,1160,1161,1162,1163,1164,1165,1166,1167,1168,1169,1170,1171,1172,1173,1174,1175,1176,1177,1178,1179,1180,1181,1182,1183,1184,1185,1186,1187,1188,1189,1190,1191,1192,1193,1194,1195,1196,1197,1198,1199,1200,1201,1202,1203,1204,1205,1206,1207,1208,1209,1210,1211,1212,1213,1214,1215,1216,1217,1218,1219,1220,1221,1222,1223,1224,1225,1226,1227,1228,1229,1230,1231,1232,1233,1234,1235,1236,1237,1238,1239,1240,1241,1242,1243,1244,1245,1246,1247,1248,1249,1250,1251,1252,1253,1254,1255,1256,1257,1258,1259,1260,1261,1262,1263,1264,1265,1266,1267,1268,1269,1270,1271,1272,1273,1274,1275,1276,1277,1278,1279,1280,1281,1282,1283,1284,1285,1286,1287,1288,1289,1290,1291,1292,1293,1294,1295,1296,1297,1298,1299,1300,1301,1302,1303,1304,1305,1306,1307,1308,1309,1310,1311,1312,1313,1314,1315,1316,1317,1318,1319,1320,1321,1322,1323,1324,1325,1326,1327,1328,1329,1330,1331,1332,1333,1334,1335,1336,1337,1338,1339,1340,1341,1342,1343,1344,1345,1346,1347,1348,1349,1350,1351,1352,1353,1354,1355,1356,1357,1358,1359,1360,1361,1362,1363,1364,1365,1366,1367,1368,1369,1370,1371,1372,1373,1374,1375,1376,1377,1378,1379,1380,1381,1382,1383,1384,1385,1386,1387,1388,1389,1390,1391,1392,1393,1394,1395,1396,1397,1398,1399,1400,1401,1402,1403,1404,1405,1406,1407,1408,1409,1410,1411,1412,1413,1414,1415,1416,1417,1418,1419,1420,1421,1422,1423,1424,1425,1426,1427,1428,1429,1430,1431,1432,1433,1434,1435,1436,1437,1438,1439,1440,1441,1442,1443,1444,1445,1446,1447,1448,1449,1450,1451,1452,1453,1454,1455,1456,1457,1458,1459,1460,1461,1462,1463,1464,1465,1466,1467,1468,1469,1470,1471,1472,1473,1474,1475,1476,1477,1478,1479,1480,1481,1482,1483,1484,1485,1486,1487,1488,1489,1490,1491,1492,1493,1494,1495,1496,1497,1498,1499,1500,1501,1502,1503,1504,1505,1506,1507,1508,1509,1510,1511,1512,1513,1514,1515,1516,1517,1518,1519,1520,1521,1522,1523,1524,1525,1526,1527,1528,1529,1530,1531,1532,1533,1534,1535,1536,1537,1538,1539,1540,1541,1542,1543,1544,1545,1546,1547,1548,1549,1550,1551,1552,1553,1554,1555,1556,1557,1558,1559,1560,1561,1562,1563,1564,1565,1566,1567,1568,1569,1570,1571,1572,1573,1574,1575,1576,1577,1578,1579,1580,1581,1582,1583,1584,1585,1586,1587,1588,1589,1590,1591,1592,1593,1594,1595,1596,1597,1598,1599,1600,1601,1602,1603,1604,1605,1606,1607,1608,1609,1610,1611,1612,1613,1614,1615,1616,1617,1618,1619,1620,1621,1622,1623,1624,1625,1626,1627,1628,1629,1630,1631,1632,1633,1634,1635,1636,1637,1638,1639,1640,1641,1642,1643,1644,1645,1646,1647,1648,1649,1650,1651,1652,1653,1654,1655,1656,1657,1658,1659,1660,1661,1662,1663,1664,1665,1666,1667,1668,1669,1670,1671,1672,1673,1674,1675,1676,1677,1678,1679,1680,1681,1682,1683,1684,1685,1686,1687,1688,1689,1690,1691,1692,1693,1694,1695,1696,1697,1698,1699,1700,1701,1702,1703,1704,1705,1706,1707,1708,1709,1710,1711,1712,1713,1714,1715,1716,1717,1718,1719,1720,1721,1722,1723,1724,1725,1726,1727,1728,1729,1730,1731,1732,1733,1734,1735,1736,1737,1738,1739,1740,1741,1742,1743,1744,1745,1746,1747,1748,1749,1750,1751,1752,1753,1754,1755,1756,1757,1758,1759,1760,1761,1762,1763,1764,1765,1766,1767,1768,1769,1770,1771,1772,1773,1774,1775,1776,1777,1778,1779,1780,1781,1782,1783,1784,1785,1786,1787,1788,1789,1790,1791,1792,1793,1794,1795,1796,1797,1798,1799,1800,1801,1802,1803,1804,1805,1806,1807,1808,1809,1810,1811,1812,1813,1814,1815,1816,1817,1818,1819,1820,1821,1822,1823,1824,1825,1826,1827,1828,1829,1830,1831,1832,1833,1834,1835,1836,1837,1838,1839,1840,1841,1842,1843,1844,1845,1846,1847,1848,1849,1850,1851,1852,1853,1854,1855,1856,1857,1858,1859,1860,1861,1862,1863,1864,1865,1866,1867,1868,1869,1870,1871,1872,1873,1874,1875,1876,1877,1878,1879,1880,1881,1882,1883,1884,1885,1886,1887,1888,1889,1890,1891,1892,1893,1894,1895,1896,1897,1898,1899,1900,1901,1902,1903,1904,1905,1906,1907,1908,1909,1910,1911,1912,1913,1914,1915,1916,1917,1918,1919,1920,1921,1922,1923,1924,1925,1926,1927,1928,1929,1930,1931,1932,1933,1934,1935,1936,1937,1938,1939,1940,1941,1942,1943,1944,1945,1946,1947,1948,1949,1950,1951,1952,1953,1954,1955,1956,1957,1958,1959,1960,1961,1962,1963,1964,1965,1966,1967,1968,1969,1970,1971,1972,1973,1974,1975,1976,1977,1978,1979,1980,1981,1982,1983,1984,1985,1986,1987,1988,1989,1990,1991,1992,1993,1994,1995,1996,1997,1998,1999,2000,2001,2002,2003,2004,2005,2006,2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020,2021,2022,2023,2024,2025,2026,2027,2028,2029,2030,2031,2032,2033,2034,2035,2036,2037,2038,2039,2040,2041,2042,2043,2044,2045,2046,2047,2048,2049,2050,2051,2052,2053,2054,2055,2056,2057,2058,2059,2060,2061,2062,2063,2064,2065,2066,2067,2068,2069,2070,2071,2072,2073,2074,2075,2076,2077,2078,2079,2080,2081,2082,2083,2084,2085,2086,2087,2088,2089,2090,2091,2092,2093,2094,2095,2096,2097,2098,2099,2100,2101,2102,2103,2104,2105,2106,2107,2108,2109,2110,2111,2112,2113,2114,2115,2116,2117,2118,2119,2120,2121,2122,2123,2124,2125,2126,2127,2128,2129,2130,2131,2132,2133,2134,2135,2136,2137,2138,2139,2140,2141,2142,2143,2144,2145,2146,2147,2148,2149,2150,2151,2152,2153,2154,2155,2156,2157,2158,2159,2160,2161,2162,2163,2164,2165,2166,2167,2168,2169,2170,2171,2172,2173,2174,2175,2176,2177,2178,2179,2180,2181,2182,2183,2184,2185,2186,2187,2188,2189,2190,2191,2192,2193,2194,2195,2196,2197,2198,2199,2200,2201,2202,2203,2204,2205,2206,2207,2208,2209,2210,2211,2212,2213,2214,2215,2216,2217,2218,2219,2220,2221,2222,2223,2224,2225,2226,2227,2228,2229,2230,2231,2232,2233,2234,2235,2236,2237,2238,2239,2240,2241,2242,2243,2244,2245,2246,2247,2248,2249,2250,2251,2252,2253,2254,2255,2256,2257,2258,2259,2260,2261,2262,2263,2264,2265,2266,2267,2268,2269,2270,2271,2272,2273,2274,2275,2276,2277,2278,2279,2280,2281,2282,2283,2284,2285,2286,2287,2288,2289,2290,2291,2292,2293,2294,2295,2296,2297,2298,2299,2300,2301,2302,2303,2304,2305,2306,2307,2308,2309,2310,2311,2312,2313,2314,2315,2316,2317,2318,2319,2320,2321,2322,2323,2324,2325,2326,2327,2328,2329,2330,2331,2332,2333,2334,2335,2336,2337,2338,2339,2340,2341,2342,2343,2344,2345,2346,2347,2348,2349,2350,2351,2352,2353,2354,2355,2356,2357,2358,2359,2360,2361,2362,2363,2364,2365,2366,2367,2368,2369,2370,2371,2372,2373,2374,2375,2376,2377,2378,2379,2380,2381,2382,2383,2384,2385,2386,2387,2388,2389,2390,2391,2392,2393,2394,2395,2396,2397,2398,2399,2400,2401,2402,2403,2404,2405,2406,2407,2408,2409,2410,2411,2412,2413,2414,2415,2416,2417,2418,2419,2420,2421,2422,2423,2424,2425,2426,2427,2428,2429,2430,2431,2432,2433,2434,2435,2436,2437,2438,2439,2440,2441,2442,2443,2444,2445,2446,2447,2448,2449,2450,2451,2452,2453,2454,2455,2456,2457,2458,2459,2460,2461,2462,2463,2464,2465,2466,2467,2468,2469,2470,2471,2472,2473,2474,2475,2476,2477,2478,2479,2480,2481,2482,2483,2484,2485,2486,2487,2488,2489,2490,2491,2492,2493,2494,2495,2496,2497,2498,2499,2500,2501,2502,2503,2504,2505,2506,2507,2508,2509,2510,2511,2512,2513,2514,2515,2516,2517,2518,2519,2520,2521,2522,25



CONSIDERATIONS

Considerations

Identification of same substances

- Works well for the majority of substances
- Restrictions are based on
 - limitations of the InChI system
 - different indexing by database producers

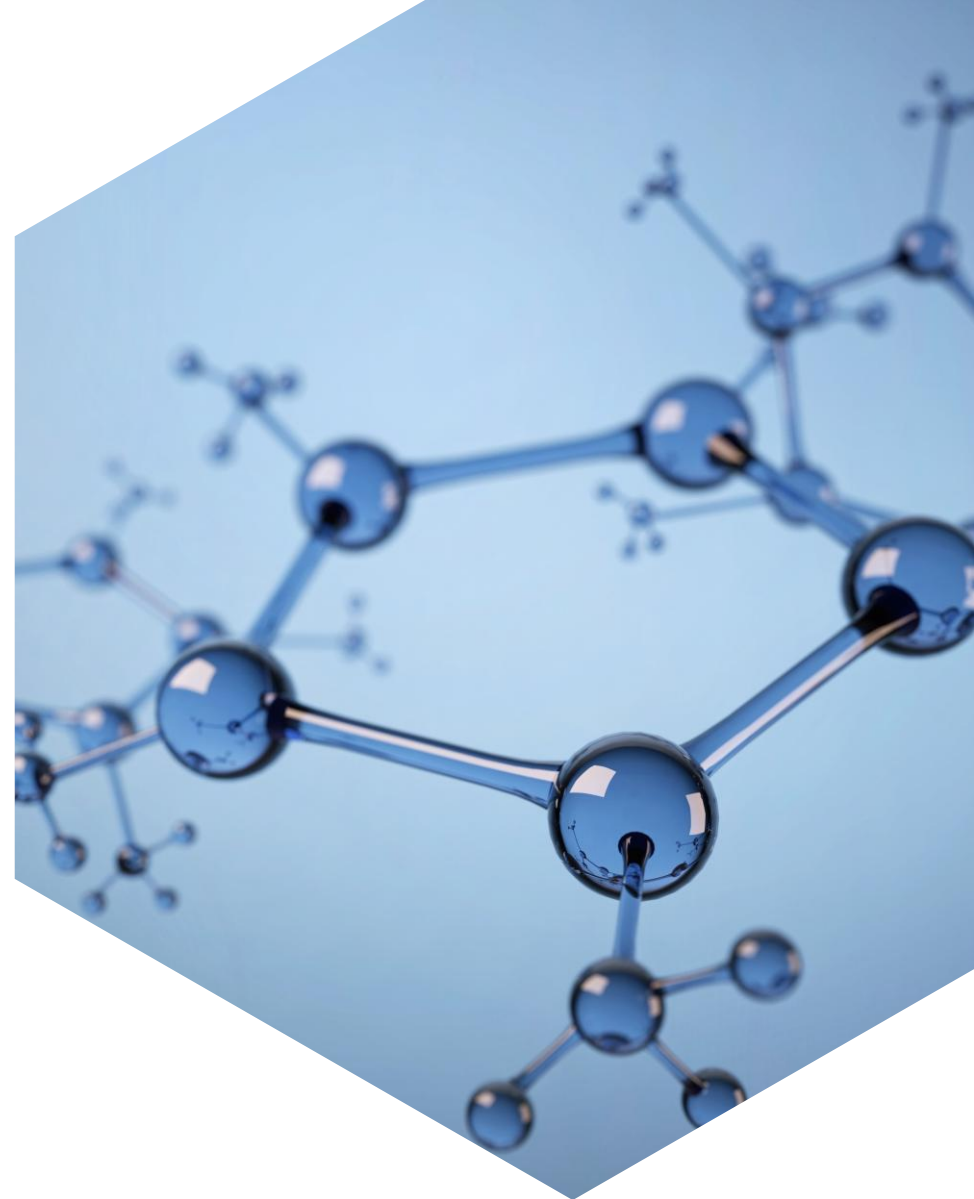
Considerations

Limitation of the InChI system

- Technical limitation of 1023 atoms; not suitable for very large compounds
- InChI needs enhancements for less well defined chemical entities; e.g.,
 - Variable structures
 - Polymers
 - Complex organometallics
 - Biologicals
 - Nanomaterials
 - Mixtures

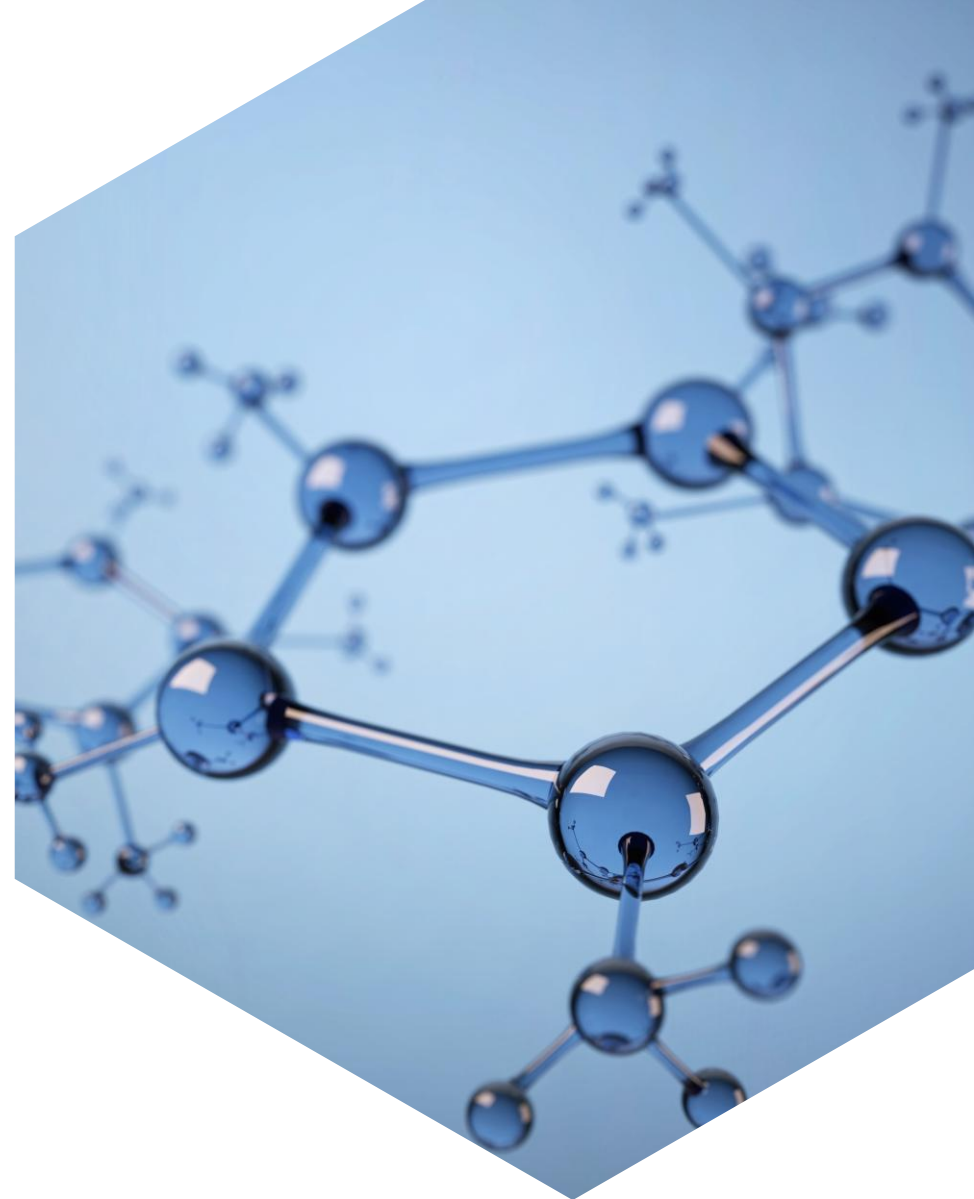
Summary

- InChI Keys are unique substance identifiers that can be used across multiple databases
- Implementation of InChI in STNext databases
 - CAS REGISTRY, DCR, ReaxysFileSub
- Consider source of information, difference in indexing systems between database producers



Special thanks

- Claudia Dauvermann-Gotsche
- Paul Peters



Thank you

Jim Brown

Senior Product Specialist IP
jim.brown@fiz-k.com

CONTACT

CAS

help@cas.org
cas.org

EMEA Help

EMEAhelp@cas.org

