



STN Workshops

MARPAT Basics

January 2001

Copyright © 2001 American Chemical Society
All Rights Reserved
Printed in the USA

QUOTING OR COPYING OF MATERIAL FROM THIS PUBLICATION
FOR EDUCATIONAL PURPOSES IS ENCOURAGED, PROVIDED
ACKNOWLEDGMENT IS MADE OF THE SOURCE OF SUCH MATERIAL.

At the end of this workshop, you will be able to

- Enhance substructure search results of the patent literature by including MARPAT in a structure search strategy
- Use CASLINK to search MARPAT and REGISTRY simultaneously
- Draw structures which can be used in substructure searches of REGISTRY and MARPAT
- Specify Match Level in structure queries to affect recall/precision of MARPAT search results

Before you begin

This workshop is designed for the information professional who needs to enhance their structure search skills of the patent literature. It is designed for the searcher who is skilled in using STN Express software to draw structures, has a basic understanding of the REGISTRY/CAplus databases, and has experience using substructure search tools.

Table of Contents

MARPAT Overview

| | |
|--|----|
| Database Content..... | 3 |
| Database at a Glance | 5 |
| MARPAT Record..... | 6 |
| MARPAT Application | 9 |
| Specific Compounds | 9 |
| Prophetic Substances..... | 10 |
| When No Answers Are Found in REGISTRY..... | 13 |
| Retrieval Differences – REGISTRY and MARPAT..... | 17 |

CASLINK

| | |
|--|----|
| Overview | 21 |
| Formats showing textual information | 31 |
| Formats showing structures..... | 31 |
| Illustration: FQHIT vs. FHIT in MARPAT | 39 |
| Review MARPAT Markush Displays..... | 42 |
| Review: Searching CASLINK | 45 |
| Skills Practice..... | 46 |

Precision Tools

| | |
|---|----|
| Balancing Precision and Recall in MARPAT..... | 51 |
| Match Level in MARPAT..... | 52 |
| Match Level Assumptions in STN Express | 53 |
| Tips for setting CLASS Match Levels | 54 |
| Tips for setting ATOM Match Levels..... | 55 |
| Tips for setting ANY Match Levels..... | 55 |
| Assigning Match Levels in STN Express | 58 |
| Single Nodes | 58 |
| Multiple Nodes..... | 59 |
| Verifying Match Levels Assignments..... | 60 |
| Skills Practice..... | 69 |

Appendix

| | |
|---|----|
| Generic Groups | 73 |
| Fields in MARPAT Structure Displays..... | 73 |
| Codes Found in Markush Structures | 75 |
| Precision Qualifiers - EX and SC..... | 76 |
| Substitution Qualifiers - SO and SR | 77 |
| Occurrence Counts | 77 |
| Generic Text Shortcuts..... | 79 |

MARPAT OVERVIEW

In this section, you will learn

- About the content of the MARPAT database
- How the MARPAT database enhances patent recall

Database Content

The MARPAT database, produced by CAS and available only on STN, contains structural representations of the Markush structures that appear in patent claims. A *Markush structure* condenses a set of implied substances into a single representation. The Markush structures are searchable in MARPAT.

Background Illustration

Markush structure in a patent claim:

| | |
|----------------------------------|---|
| United States Patent [19] | [11] Patent Number: 5,635,449 |
| Langevine et al. | [45] Date of Patent: Jun. 3, 1997 |

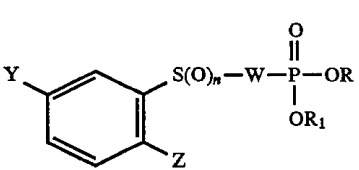
[54] **ARYLTHIOALKYL- AND ARYLTHIOALKENYLPHOSPHONIC ACIDS AND DERIVATIVES THEREOF USEFUL AS HERBICIDAL AGENTS**

S. J. Hays, et al, Journal of Medicinal Chemistry, 33, pp. 2916-2924 (1990).
C. E. Nakamura, et al, Biochemical Pharmacology, 35, pp. 122-126 (1982).

•
•
•

What is claimed is:

1. A compound having the structural formula



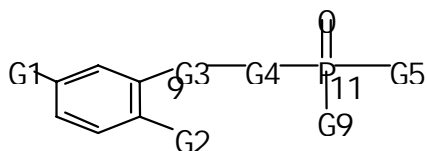
wherein

- Y is hydrogen or halogen;
- Z is NH₂ or OR₂;
- R₂ is hydrogen, C₁-C₄alkylcarbonyl or benzoyl;
- n is an integer of 0, 1 or 2;
- W is $-(CH_2)_4-$, $-CH_2CH=CHCH_2-$ or $-CH_2CH_2CH=CH-$; and
- R and R₁ are each independently hydrogen, C₁-C₄alkyl, C₁-C₄alkylcarbonyloxymethylene or an alkali metal, ammonium or organic ammonium cation, provided that where W is $-(CH_2)_4-$ and Z is NH₂ then at least one of R and R₁ is other than C₁-C₄ alkyl.

Markush structure.

Corresponding entries in MARPAT:

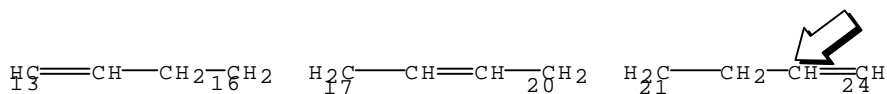
MSTR 1



Base structure - variables represented by G-groups.

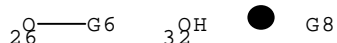
G1 = H / F / Cl / Br / I
 G2 = NH₂ / OH / alkylcarbonyloxy<(1-4)> / OCHO / OCOPh
 G3 = S / S(O) / SO₂
 G4 = CH₂CH₂CH₂CH₂ / 13-9 16-11 / 17-9 20-11 / 21-9 24-11

G-group definitions.

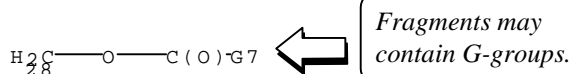


G5 = OH / 26 / 32

Definitions may include structural fragments. Node numbers indicate where the fragment is inserted in the structure.

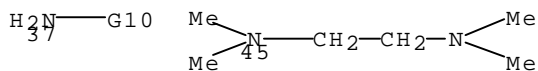


G6 = alkyl<(1-4)> / 28



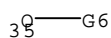
Fragments may contain G-groups.

G7 = alkyl<(1-4)> / (SC Bu-t)
 G8 = alkali metal atom / NH₃ /
 R<TX "organic ammonium cation"> / (SC 37 / 45)



G9 = OH / 35

Indication of where in the patent document the information for the MARPAT structure was obtained.



G10 = Pr-i / c-hexyl
 MPL: claim 1
 NTE: substitution is restricted

Database at a Glance

| Feature | Notes |
|----------------------------------|--|
| Source of information | CAplus patents with Markush structures |
| Time period covered | Patents published since Jan 1, 1988 |
| Countries covered | All countries covered in CAplus |
| Types of compounds covered | All classes of compounds, <i>except alloys, metal oxides, inorganic salts, intermetallics, and polymers</i> |
| Source of the Markush structures | <ul style="list-style-type: none">■ Structure from claim – or disclosure if no Markush in the claims■ Additional details from disclosure |
| Search access points | <ul style="list-style-type: none">■ Substructure searching■ Text terms from MPL, NTE, and other text fields associated with the Markush structure |
| CAplus information | Displayable, not searchable |
| “Work-in-progress” | Records found in MARPATpreviews |

MARPAT Record

All MARPAT records are also in Cplus. They have the same Accession Number (AN). Cplus information can be displayed in MARPAT.

AN 133:290336 MARPAT
TI Coordination compounds with ligands of a nitrogen heterocycle and Organic electroluminescent device using these complexes
IN Kim, Kong-Kyeom; Son, Se-Hwan; Kim, Ok-Hee; Yoon, Seok-Hee; Bae, Jae-Soon; Lee, Youn-Gu; Kim, Hyo-Seok
PA LG Chemical, Ltd., S. Korea
SO PCT Int. Appl., 47 pp.
CODEN: PIXXD2
DT Patent
LA English
IC ICM C07F001-00
ICS C07F003-00; C09K011-06; H05B033-14
CC 78-7 (Inorganic Chemicals and Reactions)
Section cross-reference(s): 28, 73
FAN.CNT 1

| PATENT NO. | KIND | DATE | APPLICATION NO. | DATE |
|---------------|------|----------|-----------------|----------|
| WO 2000058315 | A1 | 20001005 | WO 2000-KR289 | 20000330 |

W: CA, CN, JP
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE
PRAI KR 1999-11160 19990331

AB Disclosed are new coordination compds. having light-emitting and electron-transporting characteristics, such as ZnL2 (HL = 3-(2-benzothiazolyl)-4-hydroxy-2H-1-benzopyran-2-one). Also disclosed are org. EL (electroluminescent) devices using these coordination compds. As electron-transporting materials. The coordination compds. were used to form a light-emitting layer with or without doping of another light-emitting material. The coordination compds. can also be used in an electron-transporting layer of the org. EL device. The org. EL devices incorporating the coordination compds. have high thermal stability. Thus ZnL2 was prepd. by 1st reacting 4-hydroxycoumarin with Ph isothiocyanate, followed by ring closure and reaction with Zn(OAc)2.

ST zinc benzothiazole deriv complex prepn electroluminescent devices
IT Electroluminescent devices
(zinc benzothiazole deriv. complex as electron-transporting material for)

IT 7429-90-5D, Aluminum, complexes with nitrogen-contg. heterocycles
7439-93-2D, Lithium, complexes with nitrogen-contg. heterocycles
7439-95-4D, Magnesium, complexes with nitrogen-contg. heterocycles
7440-41-7D, Beryllium, complexes with nitrogen-contg. heterocycles
7440-55-3D, Gallium, complexes with nitrogen-contg. heterocycles
7440-66-6D, Zinc, complexes with nitrogen-contg. heterocycles
7440-74-6D, Indium, complexes with nitrogen-contg. heterocycles
RL: DEV (Device component use); USES (Uses)
(electron-transporting material for electroluminescent devices)

BIB information, including patent family members when available.

Cplus abstract.

Cplus subject and CAS RN indexing.

(continued on next page)

IT 198-55-0, Perylene 2085-33-8, Alq3
 RL: DEV (Device component use); USES (Uses)
 (light-emitting material for electroluminescent devices using coordination compds. with nitrogen-contg. heterocycles as electron-transporting material)

IT 14994-75-3P 18706-64-4P 26004-25-1P 54289-77-9P 71651-78-0P
 299158-90-0P 299158-91-1P 299158-92-2P 299158-93-3P
 RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation)
 (prepn. and reactant for prepn. of zinc benzothiazole deriv. Complexes as electron-transporting material for electroluminescent devices)

IT 299158-87-5P 299158-88-6P 299158-89-7P
 RL: DEV (Device component use); SPN (Synthetic preparation); PREP (Preparation); USES (Uses)
 (prepn. as electron-transporting material for electroluminescent devices)

IT 103-72-0, Phenyl isothiocyanate 105-53-3, Diethyl malonate 122-39-4, Diphenylamine, reactions 557-34-6, Zinc acetate 1076-38-6, 4-Hydroxycoumarin 1677-46-9
 RL: RCT (Reactant)
 (reactant for prepn. of zinc benzothiazole deriv. complexes as electron-transporting material for electroluminescent devices)

RE.CNT 6

RE

Citations from the patent.

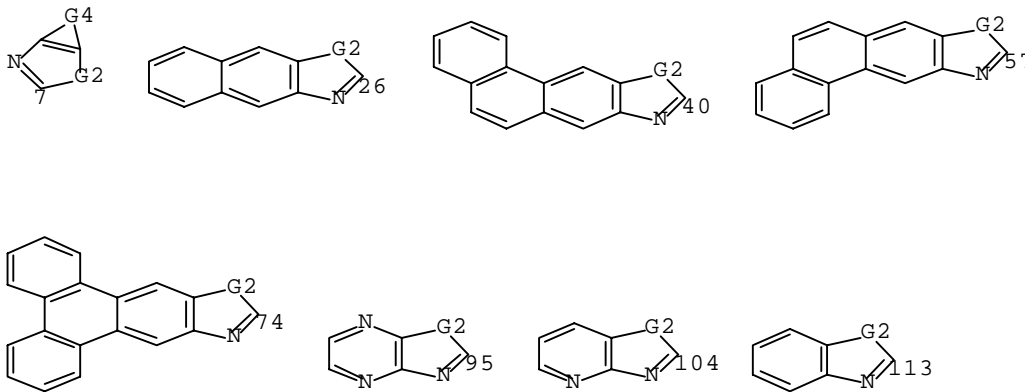
- (1) Chen, C; US 6020078 A 2000 CAPLUS
- (2) Lg Chemical Ltd; WO 9837736 A1 1998 CAPLUS
- (3) Lg Chemical Ltd; WO 9963023 A1 1999 CAPLUS
- (4) Sanyo Electric Co Ltd; EP 0743809 A2 1996 CAPLUS
- (5) Shinko Electric Industries Co Ltd; EP 0801518 A2 1997 CAPLUS
- (6) Xerox Corporation; EP 0862353 A2 1998 CAPLUS

MSTR 1

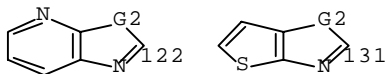
G1—G5 G8

Markush structure from the patent claims.

G1 = 7 / (SC 26 / 40 / 57 / 74 / 95 / 104 / 113 / 122 / 131)



(continued on next page)



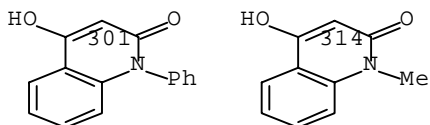
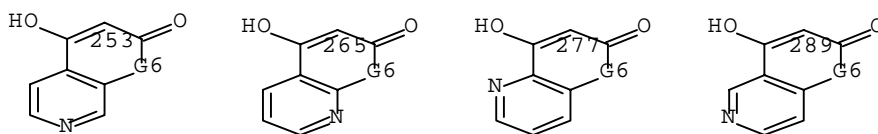
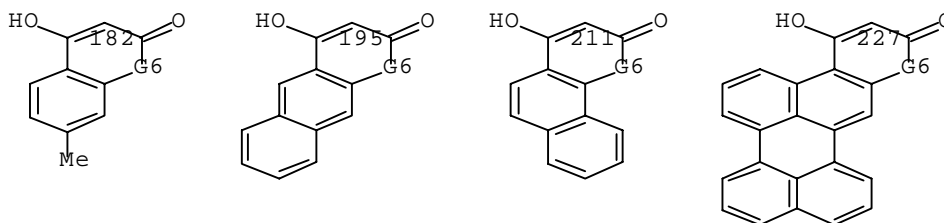
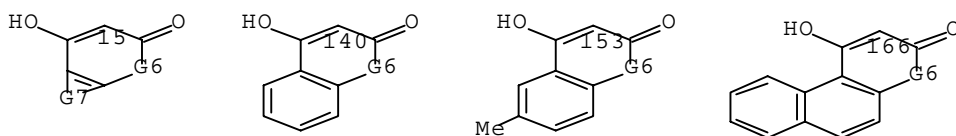
G2 = CH2 (SO) / O / S / Se / 8



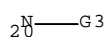
G3 = alkyl (SO) / aryl (SO)

G4 = R<TX "moiety to form aromatic or heterocyclic ring">

G5 = 15 / (SC 140 / 153 / 166 / 182 / 195 / 211 / 227 / 253 / 265 / 277 / 289 / 301 / 314)



G6 = O / S / 20



G7 = R<TX "moiety to form aromatic or heterocyclic ring">

G8 = R<TX "metal"> / (SC Li / Be / Zn / Mg / Ga / In / Al)

MPL: claim 1

NTE: as complex with G8

MARPAT Application

Searching both MARPAT and REGISTRY enhances substructure search recall of the patent literature by retrieving both

- Specific compound matches, from REGISTRY
- Prophetic compound matches, from MARPAT

Techniques for conducting a substructure search in MARPAT are similar to those used in REGISTRY.

Specific Compounds

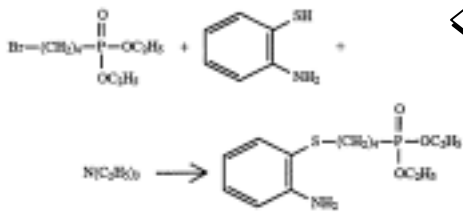
REGISTRY contains the CAS RNs for substances indexed for the document records in CAplus. For patent documents, the following substances are indexed:

- Substances explicitly defined in patent claims (1980-present)
- Substances appearing in patent examples with supporting information

Example:

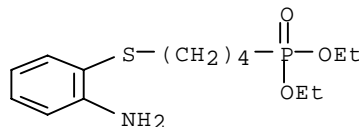
Compound indexed from patent Example 1:

EXAMPLE 1
Preparation of Diethyl {4-[(o-aminophenyl)thio]butyl}phosphonate



Triethylamine (0.1 mol) is slowly added to a mixture of diethyl (4-bromobutyl)phosphonate (0.1 mol) and o-aminothiophenol (0.1 mol) in tetrahydrofuran at 0° C. The reaction mixture is stirred at 0° C. for 4 hours, at room temperature for 16 hours and poured into water. The aqueous mixture is extracted with ether. The organic extracts are combined, dried over anhydrous magnesium sulfate and concentrated in vacuo to obtain a residue. Column chromatography of the residue using silica gel, a 1:1 ethyl acetate/hexane solution and a 1:9 methanol/ethyl acetate solution gives the title product as a clear yellow liquid which is

CAS RN 191411-58-2



AN 127:62046
TI Preparative and derivative

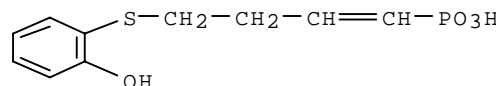
alkenylphosphonic acids

IT 191411-58-2P 191411-61-7P 191411-63-9P 191411-65-1P 191411-67-3P
191411-69-5P 191411-71-9P 191411-73-1P 191411-74-2P 191411-75-3P
191411-76-4P 191411-77-5P 191411-78-6P 191411-79-7P 191411-80-0P
191411-81-1P 191411-82-2P 191411-83-3P 191411-84-4P 191411-85-5P
191411-86-6P 191411-87-7P 191411-88-8P 191411-89-9P 191411-90-2P
191411-91-1P 191411-94-6P 191411-95-7P
191411-99-1P 191412-00-7P

Compound indexed from claim 4:

4. The compound according to claim 3 [4-[(o-hydroxyphenyl)thio]-1-butenyl]phosphonic acid.
5. The compound according to claim 3 diethyl {4-[(o-aminophenyl)thio]butyl}phosphonate.
6. The compound according to claim 3 dipivalate ester of bis(hydroxymethyl) {4-[(o-hydroxyphenyl)thio]butyl}phosphonate.

CAS RN 191411-86-6



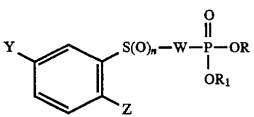
Prophetic Substances

Prophetic substances, e.g., those represented by Markush structures in patent claims, are not generally indexed in CAPLUS.

The Markush representations in MARPAT provide structure-searchable access to the substances covered by patent claims and disclosures, such as the prophetic substances.

Example:

What is claimed is:
 1. A compound having the structural formula



wherein
 Y is hydrogen or halogen;
 Z is NH₂ or OR₂;
 R₂ is hydrogen, C₁-C₄alkylcarbonyl or benzoyl;
 n is an integer of 0, 1 or 2;
 W is $-(CH_2)_4-$, $-CH_2CH=CHCH_2-$ or $-CH_2CH_2CH=CH-$; and
 R and R₁ are each independently hydrogen, C₁-C₄alkyl, C₁-C₄alkylcarbonyloxymethylene or an alkali metal, ammonium or organic ammonium cation, provided that where W is $-(CH_2)_4-$ and Z is NH₂, then at least one of R and R₁ is other than C₁-C₄ alkyl.

2. The compound according to claim 1 wherein
 Y is hydrogen, F or Br;
 n is an integer of 0 or 1;
 W is $-(CH_2)_4-$ or $-CH_2CH_2CH=CH-$; and
 R and R₁ are each independently hydrogen, C₁-C₄alkyl, C₁-C₄alkylcarbonyloxymethylene or an alkali metal or organic ammonium cation.

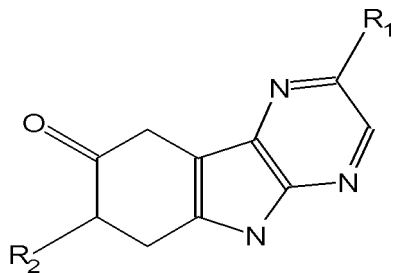
3. A compound selected from the group consisting of
 {4-[(o-hydroxyphenyl)thio]-1-butenyl}phosphonic acid;
 {4-[(o-hydroxyphenyl)thio]-1-butenyl}phosphonic acid;
 dilithium {4-[(o-aminophenyl)thio]butyl}phosphonate;
 dipivalate ester of bis(hydroxymethyl) {4-[(o-hydroxyphenyl)thio]butyl}phosphonate;
 {4-[(o-hydroxyphenyl)thio]butyl}phosphonic acid, compound with N N N',N'-tetramethylethylenediamine;

30 {4-[(o-hydroxyphenyl)sulfinyl]butyl}phosphonic acid;
 {4-[(o-hydroxyphenyl)thio]butyl}phosphonic acid, aryl-
 butyrate ester; and mixtures thereof.
 4. The compound according to claim 3 {4-[(o-
 hydroxyphenyl)thio]-1-butenyl}phosphonic acid.
 5. The compound according to claim 3 dilithium {4-[(o-
 aminophenyl)thio]butyl}phosphonate.
 6. The compound according to claim 3 dipivalate ester of
 bis(hydroxymethyl) {4-[(o-hydroxyphenyl)thio]
 butyl}phosphonate.
 40 7. The compound according to claim 2 {4-[(o-
 hydroxyphenyl)thio]butyl}phosphonic acid, compound with
 N,N,N',N'-tetramethylethylenediamine.
 8. *Each specific compound generated based
 on the Markush variables is not indexed
 for CAPLUS.*
 hydroxy ester.
 10. A mixture comprising a compound selected from the
 group consisting of
 {4-[(o-aminophenyl)thio]butyl}phosphonic acid;
 {4-[(o-hydroxyphenyl)sulfinyl]butyl}phosphonic acid;
 and
 {4-[(o-hydroxyphenyl)thio]-1-butenyl}phosphonic acid,
 in combination with an aliphatic or cycloaliphatic
 amine.
 11. The mixture of claim 10 wherein the aliphatic amine
 is cyclohexylamine.
 12. The mixture of claim 10 wherein the cycloaliphatic
 amine is isopropylamine.
 60 13. The mixture of claim 11 wherein cyclohexylamine is in
 an amount greater than the compound.
 14. The mixture of claim 12 wherein isopropylamine is in
 an amount greater than the compound.
 65 15. The mixture according to claim 13 comprising
 {4-[(o-aminophenyl)thio]butyl}phosphonic acid and
 cyclohexylamine in a ratio of 1:2.

Consider searching REGISTRY and MARPAT when

- It is important to comprehensively search all structural possibilities covered in patent claims
- REGISTRY searches turn up no hits
- It is important to cover "incompletely defined" compounds that might not match your query specifications in REGISTRY

Search Question: What has been reported on compounds with the following structural characteristics?



R1 = pyridyl ring (isolated/embedded)
R2 = carbon chain of any length
No additional fusion on the polycyclic ring
Additional substitution allowed at all open sites

Searching REGISTRY for specific compound matches:

=> FILE REGISTRY

=>

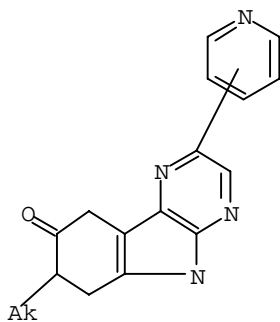
Uploading mar1.str

L1 STRUCTURE UPLOADED

=> D L1

L1 HAS NO ANSWERS

L1 STR



(continued on next page)

```
=> S L1 SSS SAM
```

```
•  
•  
•
```

```
L2          0 SEA SSS SAM L1
```

```
=> S L1 SSS FULL
```

```
•  
•  
•
```

```
L3          0 SEA SSS FUL L1
```

The REGISTRY search turns up no substances.

When No Answers Are Found in REGISTRY

"No hits" in REGISTRY means that no substances represented by the structure query were indexed in CAPlus from 1967-present (or in CAOLD from the 1957-1966 *Molecular Formula Indexes*).

The substance of interest could be encompassed in the prophetic substances covered by the Markush structure in a patent claim.

Searching MARPAT for prophetic substance matches:

```
=> FILE MARPAT
```

```
=> S L3 SSS SAM
```

```
SAMPLE SEARCH INITIATED 13:55:20 FILE 'MARPAT'  
SAMPLE SCREEN SEARCH COMPLETED - 7 TO ITERATE  
100.0% PROCESSED 7 ITERATIONS 0 ANSWERS  
SEARCH TIME: 00.00.03
```

```
FULL FILE PROJECTIONS: ONLINE **COMPLETE**  
                        BATCH  **COMPLETE**  
PROJECTED ITERATIONS: 7 TO 299  
PROJECTED ANSWERS: 0 TO 0
```

```
L5          0 SEA SSS SAM L1
```

Use a SAMPLE search in MARPAT to verify the search will run to completion.

(continued on next page)

=> S L3 SSS FULL

FULL SEARCH INITIATED 13:54:52 FILE 'MARPAT'
FULL SCREEN SEARCH COMPLETED - 89 TO ITERATE
100.0% PROCESSED 89 ITERATIONS 3 ANSWERS
SEARCH TIME: 00.00.09

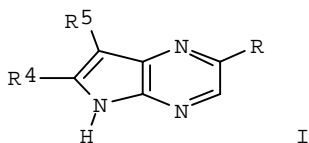
L4 3 SEA SSS FUL L1

The MARPAT search retrieves 3 patents which discuss these compounds as GABA_A receptor ligands.

=> D 1-3 BIB ABS

L6 ANSWER 1 OF 3 MARPAT COPYRIGHT 2000 ACS
AN 127:293247 MARPAT
TI Preparation of pyrrolopyrazines as GABA_A receptor ligands
IN Blum, Charles; Hutchison, Alan
PA Neurogen Corp., USA
SO U.S., 16 pp. Cont.-in-part of U.S. 5,606,059.
CODEN: USXXAM
DT Patent
LA English
FAN.CNT 3

| | PATENT NO. | KIND | DATE | APPLICATION NO. | DATE |
|------|---------------|------|----------|-----------------|----------|
| PI | US 5668283 | A | 19970916 | US 95-486595 | 19950607 |
| | US 5286860 | A | 19940215 | US 92-975409 | 19921112 |
| | US 5606059 | A | 19970225 | US 95-436252 | 19950512 |
| PRAI | US 92-975409 | | 19921112 | | |
| | US 95-436252 | | 19950512 | | |
| | WO 93-US10870 | | 19931110 | | |
| GI | | | | | |



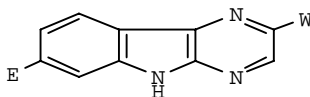
AB Title compds. [I; R = (un)substituted Ph, -thienyl, -pyridyl; R4R5 = (CH₂)_nCR₁R₂ZCH₂ or CH:CR₆Z₁:Z₂; R₁,R₂,R₆ = H or alkyl; Z = NR₃, CO, CH(OH), alkoxy carbonylmethine, etc.; R₃ = H, (phenyl)alkyl, Ph, alkoxy, etc.; Z₁ = NCR₁₅; R₁₅ = H, halo, (phenyl)alkyl, etc.; Z₂ = CH or N; n = 0-2] were prepd. Thus, 2,5-F(MeO)C₆H₃COCHO was cyclocondensed with H₂NCH₂CONH₂ and the product converted in 4 steps to I [R = C₆H₄3(OMe)F-5,2, R₄R₅ = (CH₂)₄]. Data for biol. Activity of I given.

(continued on next page)

L6 ANSWER 2 OF 3 MARPAT COPYRIGHT 2000 ACS
 AN 126:199581 MARPAT
 TI aryl substituted pyrrolopyrazines as a new class of GABA brain
 Receptor ligands
 IN Blum, Charles; Hutchison, Alan
 PA Neurogen Corporation, USA
 SO U.S., 16 pp. Cont.-in-part of U.S. 5,286,860.
 CODEN: USXXAM
 DT Patent
 LA English
 FAN.CNT 3

| | PATENT NO. | KIND | DATE | APPLICATION NO. | DATE |
|------|---|------|----------|-----------------|----------|
| | ----- | --- | ----- | ----- | ----- |
| PI | US 5606059 | A | 19970225 | US 95-436252 | 19950512 |
| | US 5286860 | A | 19940215 | US 92-975409 | 19921112 |
| | WO 9411374 | A1 | 19940526 | WO 93-US10870 | 19931110 |
| | W: AT, AU, BB, BG, BR, BY, CA, CH, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, KZ, LK, LU, MG, MN, MW, NL, NO, NZ, PL, RO, RU, SD, SE, SK, UA, US, VN | | | | |
| | RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG | | | | |
| | US 5668283 | A | 19970916 | US 95-486595 | 19950607 |
| PRAI | US 92-975409 | | 19921112 | | |
| | WO 93-US10870 | | 19931110 | | |
| | US 95-436252 | | 19950512 | | |

GI



AB Indolopyrazines I (W = 2-, 3-thienyl, 2-, 3-, 4-pyridyl, Ph, optionally mono- or disubstituted with halogen, OH, lower alkyl, alkoxy, NH₂, mono- or dialkylamino; E = H, lower alkyl), highly selective agonists, antagonists or inverse agonists for GABA_A brain receptors, are claimed.

L6 ANSWER 3 OF 3 MARPAT COPYRIGHT 2000 ACS
 AN 120:245160 MARPAT
 TI Preparation of indolopyrazines and related compounds as brain GABA_A
 agonists, antagonists, or inverse agonists
 IN Blum, Charles; Hutchison, Alan
 PA Neurogen Corp., USA
 SO U.S., 17 pp.
 CODEN: USXXAM
 DT Patent
 LA English

(continued on next page)

FAN.CNT 3

| | PATENT NO. | KIND | DATE | APPLICATION NO. | DATE |
|------|---|------|----------|-----------------|----------|
| PI | US 5286860 | A | 19940215 | US 92-975409 | 19921112 |
| | WO 9411374 | A1 | 19940526 | WO 93-US10870 | 19931110 |
| | W: AT, AU, BB, BG, BR, BY, CA, CH, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, KZ, LK, LU, MG, MN, MW, NL, NO, NZ, PL, RO, RU, SD, SE, SK, UA, US, VN | | | | |
| | RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG | | | | |
| | AU 9455526 | A1 | 19940608 | AU 94-55526 | 19931110 |
| | US 5606059 | A | 19970225 | US 95-436252 | 19950512 |
| | US 5668283 | A | 19970916 | US 95-486595 | 19950607 |
| PRAI | US 92-975409 | | 19921112 | | |
| | WO 93-US10870 | | 19931110 | | |
| | US 95-436252 | | 19950512 | | |
| | • | | | | |
| | • | | | | |
| | • | | | | |

Helpful HINT

Search the L-number resulting from the REGISTRY substructure search to take advantage of a lower, “extended” search fee in MARPAT.

For additional information on file-specific charges see HELP COST in the file.

Retrieval Differences – REGISTRY and MARPAT

The way substructure matches are made in each database is different.

In **REGISTRY**, only specific substituents are present in the database compounds. Substructures match against specific substituents.

In **MARPAT**, substructures may match against

- Specific substituents: H, CF₃, CN, etc.
- Generic substituents: X, HY, alkoxy-carbonyl, etc.

Answers in each database are different:

- Answers in REGISTRY are compounds. The bibliographic references and abstracts associated with each compound (from patents and journals) are available in CAPlus.
- Answers in MARPAT are references to patents.

CASLINK

In this section, you will learn how to

- Use the CASLINK feature to substructure search in MARPAT and REGISTRY simultaneously
- Review MARPAT Markush displays

Overview

The CASLINK tool provides one-step searching of both the

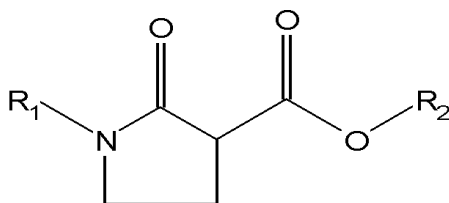
- Specific substances in REGISTRY
- Prophetic substances in MARPAT and MARPATpreviews

CASLINK also retrieves the corresponding references to the specific substances from CAplus and eliminates any duplicate hits between CAplus and MARPAT/MARPATpreviews.

Basic Search Steps To use CASLINK

- Step 1 Building and save the structure.
- Step 2 Upload the structure.
- Step 3 Test the query.
- Step 4 *(Option)* Revise the query.
- Step 5 Run a FULL substructure search.
- Step 6 Display results.

Search Question: *Locate references on the following substances:*



R1 = Any non-hydrogen substituent

R2 = any ring or chain carbon atom

No additional fusion on the ring

Additional substitution allowed at all open sites

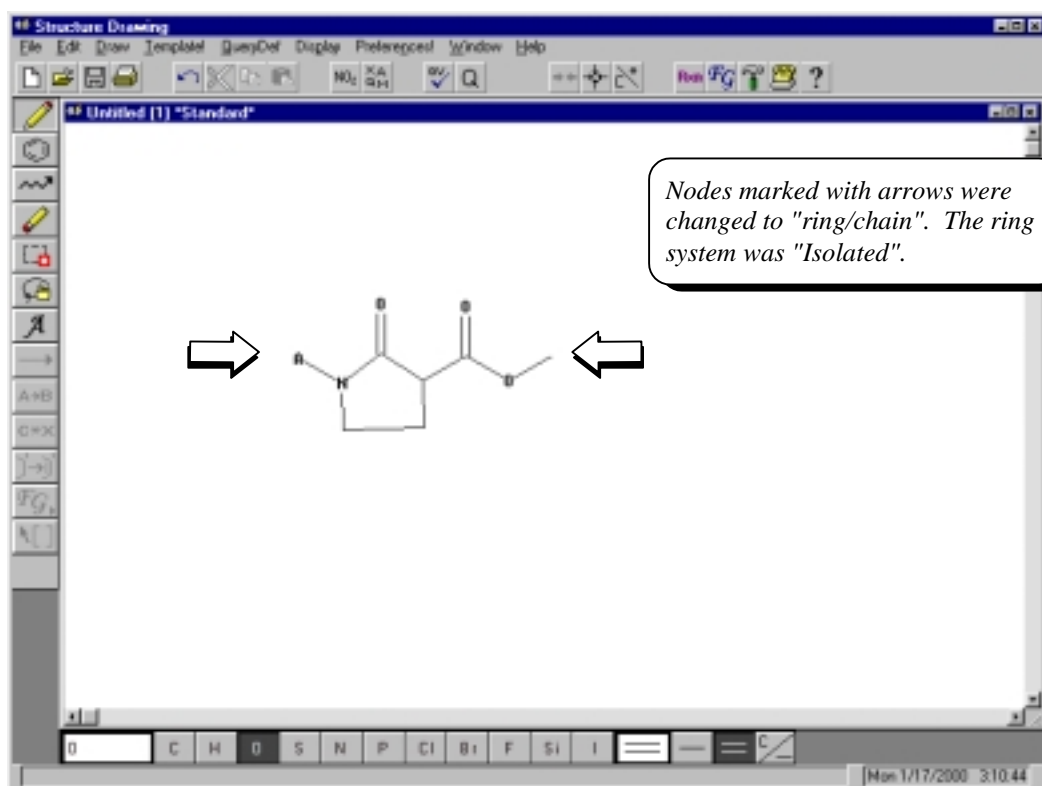
Worksheet

| Requirement | Your Action |
|--|--------------------|
| R1 = any non-hydrogen substituent | |
| R2 = any ring/chain carbon atom | |
| No ring fusion | |
| Substitution allowed at all open sites | |

Step 1: Build and Save the Structure

The same types of structure queries can be searched in REGISTRY and MARPAT. MARPAT queries may contain

- Specific atoms and shortcuts
- Variable groups
- G-groups
- Specific bonds
- Unspecified bonds
- Isolated rings
- Ring/chain nodes



note

Ring/chain bond characteristics are ignored in MARPAT

Step 2: Upload the Structure

Uploading a structure in CASLINK is accomplished in 3 steps:

1. Enter the CASLINK cluster of files.
2. Perform the upload.
3. Verify the structure online.

Entering CASLINK:

```
=> FILE CASLINK
```

```
FILE 'REGISTRY' ENTERED
```

```
FILE 'MARPAT' ENTERED
```

```
FILE 'MARPATPREV' ENTERED
```

```
FILE 'CAPLUS' ENTERED
```

```
CLUSTER 'CASLINK' ENTERED
```

```
Predefined command sequences will be executed in  
REGISTRY, MARPAT, MARPATPREV, and CAPLUS.
```

*Use the FILE command to enter the
CASLINK multifile cluster of files.*

Performing the upload:

```
=>
```

```
Uploading marl.str
```

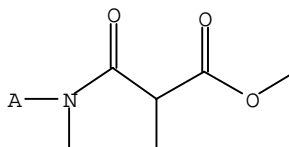
```
L1 STRUCTURE UPLOADED
```

Verifying the structure:

```
=> D L1
```

```
L1 HAS NO ANSWERS
```

```
L1 STR
```



Step 3: Test the Query

Testing the substructure query is a 2-step process:

1. Run a SAMPLE search.
2. Evaluate answers.

Running a SAMPLE search:

```
=> S L1 SSS SAM
```

```
S L1 SSS SAM FILE=REGISTRY
```

```
SAMPLE SEARCH INITIATED 15:34:31 FILE 'REGISTRY'  
SAMPLE SCREEN SEARCH COMPLETED - 284 TO ITERATE  
100.0% PROCESSED 284 ITERATIONS 15 ANSWERS  
SEARCH TIME: 00.00.01
```

```
FULL FILE PROJECTIONS: ONLINE **COMPLETE**  
                        BATCH **COMPLETE**  
PROJECTED ITERATIONS: 4669 TO 6691  
PROJECTED ANSWERS:   68 TO 532
```

```
L2 15 SEA SSS SAM L1  
1 FILES SEARCHED...
```

```
S L2 SSS SAM FILE=MARPAT
```

```
SAMPLE SEARCH INITIATED 15:34:34 FILE 'MARPAT'  
SAMPLE SCREEN SEARCH COMPLETED - 526 TO ITERATE  
100.0% PROCESSED 526 ITERATIONS 3 ANSWERS  
SEARCH TIME: 00.00.10
```

```
FULL FILE PROJECTIONS: ONLINE **COMPLETE**  
                        BATCH **COMPLETE**  
PROJECTED ITERATIONS: 9195 TO 11845  
PROJECTED ANSWERS:   3 TO 164
```

```
L3 3 SEA SSS SAM L1  
1 FILES SEARCHED...
```

SAMPLE structure searches in CASLINK are automatically run in REGISTRY and MARPAT.



Check the full-file projections in both databases.



Evaluating answers:

Each answer from REGISTRY is a specific compound that matches the structure query.

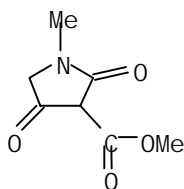
Each answer from MARPAT is a document record (patent) in which the Markush structure contained fragments that matched the structure query.

D SCAN is used to determine if the REGISTRY and MARPAT answers are on target. D SCAN content is file-specific:

=> D SCAN L2

Structure matches from REGISTRY.

L2 15 ANSWERS REGISTRY COPYRIGHT 2000 ACS
IN 3-Pyrrolidinecarboxylic acid, 1-methyl-2,4-dioxo-, methyl ester
(9CI)
MF C7 H9 N O4

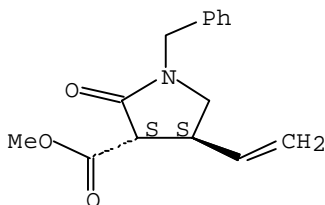


Note: Substance identification information from REGISTRY.

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

L2 15 ANSWERS REGISTRY COPYRIGHT 2000 ACS
IN 3-Pyrrolidinecarboxylic acid, 4-ethenyl-2-oxo-1-(phenylmethyl)-, Methyl ester, (3R,4R)-rel- (9CI)
MF C15 H17 N O3

Relative stereochemistry.



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

(continued on next page)

=> D SCAN L3

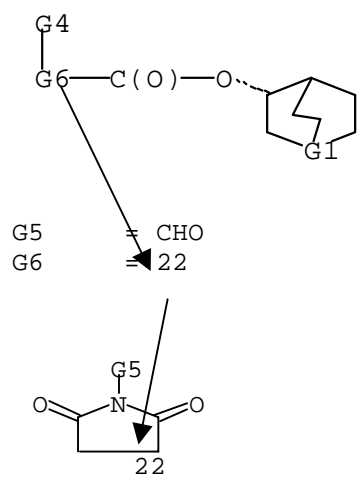
Structure matches from
MARPAT.

L3 3 ANSWERS MARPAT COPYRIGHT 2000 ACS
IC ICM C07D453-02
ICS A61K031-435
CC 27-16 (Heterocyclic Compounds (One Hetero Atom))
Section cross-reference(s): 1
TI Preparation of R-3-quinuclidinol esters with heterocyclylcarboxylic
acids
as muscarinic M1 antagonists
ST heterocyclylcarbonyloxyazabicyclooctane prepn muscarinic M1
antagonist; quinuclidinol piperidinylocarbonyl muscarinic antagonist;
piperidinylocarbonyloxyquinuclidine muscarinic antagonist;
pyrrolidinylocarbonyloxyquinuclidine muscarinic antagonist
IT Ulcer inhibitors
(quinuclidinol esters)
IT Emphysema
Heart, disease or disorder
(treatment of, with quinuclidinol esters)
IT Bronchodilators
(antiasthmatics, quinuclidinol esters)
●
●
●
IT 133904-91-3P 133904-92-4P 133904-93-5P 133904-94-6P 133904-95-7P
133904-96-8P 133904-97-9P 133904-98-0P 133904-99-1P 133905-00-7P
133905-01-8P 133905-02-9P 133905-03-0P 133905-04-1P 133905-05-2P
133905-35-8P 133905-36-9P 133905-37-0P 133905-38-1P 133905-39-2P
133905-40-5P 133905-41-6P 133905-42-7P 133905-43-8P 133905-44-9P
133905-45-0P 133905-46-1P 133905-47-2P 133905-48-3P 133905-49-4P
133905-50-7P 133905-51-8P 133905-52-9P 133905-53-0P 133905-54-1P
133905-55-2P
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as muscarinic M1 antagonist)
IT 133905-19-8P
RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation)
(prepn., hydrolysis, and cyclization of)
IT 4553-07-5 133905-23-4
RL: RCT (Reactant)
(reaction of, in prepn. of muscarinic antagonist)
IT 133905-33-6
RL: RCT (Reactant)
(transesterification reaction of, with quinuclidinol)

Note: Title and indexing
terms for each MARPAT
answer.

(continued on next page)

MSTR 1



Hit portion of the first hit Markush structure in the answer.

MPL: claim 1
STE: diastereoisomers and mixtures

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

Step 5: Run a FULL Search

When a FULL structure search is requested in CASLINK, STN does the following:

1. Runs FULL structure searches in REGISTRY, MARPAT, and MARPATpreviews.
2. Searches the REGISTRY answers in CAPlus.
3. Removes duplicate answers between CAPlus and MARPAT/ MARPATpreviews and creates a single answer set.

```
=> S L1 SSS FULL

S L1 SSS FUL FILE=REGISTRY
FULL SEARCH INITIATED 15:37:01 FILE 'REGISTRY'
FULL SCREEN SEARCH COMPLETED - 5669 TO ITERATE
100.0% PROCESSED 5669 ITERATIONS 411 ANSWERS
SEARCH TIME: 00.00.01

L4 411 SEA SSS FUL L1
1 FILES SEARCHED...

S L4 SSS FUL FILE=MARPAT
FULL SEARCH INITIATED 15:37:03 FILE 'MARPAT'
FULL SCREEN SEARCH COMPLETED - 10664 TO ITERATE
11.8% PROCESSED 1261 ITERATIONS 9 ANSWERS
.
.
.
99.4% PROCESSED 10605 ITERATIONS 53 ANSWERS
100.0% PROCESSED 10664 ITERATIONS 54 ANSWERS
SEARCH TIME: 00.02.44

L5 54 SEA SSS FUL L1
1 FILES SEARCHED...

S L5 SSS FUL FILE=MARPATPREV
FULL SEARCH INITIATED 15:39:49 FILE 'MARPATPREV'
FULL SCREEN SEARCH COMPLETED - 40 TO ITERATE
100.0% PROCESSED 40 ITERATIONS 0 ANSWERS
SEARCH TIME: 00.00.04

L6 0 SEA SSS FUL L1
1 FILES SEARCHED...
```

(continued on next page)

```
S L4 FILE=CAPLUS
L7      161 FILE CAPLUS
  1 FILES SEARCHED...
```

Retrieval of references from CAplus with the REGISTRY answer set.

```
DUP REM L6 L5 L7
L6 HAS NO ANSWERS
PROCESSING COMPLETED FOR L6
PROCESSING COMPLETED FOR L5
PROCESSING COMPLETED FOR L7
```

Elimination of duplicates between CAplus, MARPAT, and MARPATpreviews.

```
L8      209 DUP REM L6 L5 L7 (6 DUPLICATES REMOVED)
        ANSWERS '1-54' FROM FILE MARPAT
        ANSWERS '55-209' FROM FILE CAPLUS
```

Unique set of document records from the SSS

CASLINK Search Results

| Database | Number of Answers |
|----------------|----------------------------------|
| REGISTRY | 411 compounds |
| CAplus | 161 references (155 unique hits) |
| MARPAT | 54 References (48 unique hits) |
| MARPATpreviews | 0 |

Step 6: Display Results

Formats showing textual information:

| Use this format | If you want to display |
|-----------------|---|
| BIB/IBIB | Bibliographic information, including closely-related family members |
| ABS | Abstract |
| IND | CAplus indexing |

Formats showing structures:

| Use this format | If you want to display |
|-----------------|---|
| HITSTR | In CAplus: The structure(s) of the specific compounds(s) that caused the answer to be retrieved |
| FQHIT | In MARPAT: The hit portion only for the first hit Markush structure in the answer |
| FHIT | In MARPAT: The hit Markush structure in the answer with the hit portions highlighted |
| MSTR | In MARPAT: All the Markush structures associated with an answer <i>Note:</i> This display can be very long for complex Markush structures |

note

Additional display formats are available. Consult the database summary sheet or type *HELP FORMAT FILE=MARPAT* in CASLINK.

Displaying MARPAT hits:

=> D 5, 35 BIB ABS FQHIT

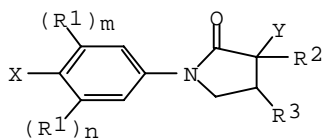
L8 ANSWER 5 OF 209 MARPAT COPYRIGHT 2000 ACS
 AN 114:81576 MARPAT
 TI Preparation of 1-phenylpyrrolidones as herbicide
 IN Woolard, Frank X.
 PA ICI Americas, Inc., USA
 SO U.S., 11 pp.
 CODEN: USXXAM
 DT Patent
 LA English
 IC ICM C07D207-26
 ICS C07D403-06; A01N043-36
 NCL 071095000
 CC 27-10 (Heterocyclic Compounds (One Hetero Atom))
 Section cross-reference(s): 5

DUPLICATE 6

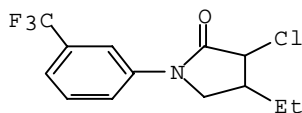
Answer retrieved from the
 REGISTRY/CaPlus and
 MARPAT searches.

FAN.CNT 1

| PATENT NO. | KIND | DATE | APPLICATION NO. | DATE |
|------------|------|----------|-----------------|----------|
| US 4956006 | A | 19900911 | US 1988-291081 | 19881227 |



I

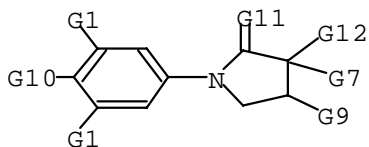


II

AB Title compds. I [R1 = halo, CF3, cyano, Me, CF2CHF2, OCF2CHF2, OCHF2, OCF3, SMe, SMe, SO2Me, (substituted) oxyiminoalkyl, (substituted) pyridyloxy; R2 = halo, cyano, CO2H, (substituted) carbalkoxy; R3 = C1-4 alkyl, C2-4 alkenyl; X, Y = H, halo; Z = 0, S; m, n = 0, 1] useful as herbicides, were prepd. m-(N-n-Butyl)aminobenzotrifluoride was subjected to amidation with ClCOCH2CO2Et, diazotization, cyclization, chlorination, hydrolysis, and decarboxylation to afford I (R1 = CF3, R2 = Cl, R3 = Et, X = Y = H, Z = 0, m = 1, n = 0) (II). The trans isomer of II at 4 lb/acre showed .gtoreq.90% control as a pre-emergence herbicide against yellow nutsedge and a variety of grasses and broadleaf weeds.

(continued on next page)

MSTR 1



G7 = alkoxy carbonyl<(1-8)> (SO (1-) G8)

G11 = O

MPL: claim 1

Unique MARPAT hit.

L8 ANSWER 35 OF 209 MARPAT COPYRIGHT 2000 ACS

AN 120:245064 MARPAT

TI Preparation of N-(difluorobenzodioxolyl)-2-pyrrolidones, as herbicides

IN Rempfler, Hermann

PA Ciba-Geigy A.-G., Switz.

SO PCT Int. Appl., 49 pp.

CODEN: PIXXD2

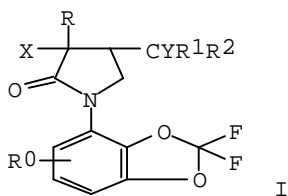
DT Patent

LA English

FAN.CNT 1

| | PATENT NO. | KIND | DATE | APPLICATION NO. | DATE |
|------|--|------|----------|-----------------|----------|
| PI | WO 9324483 | A1 | 19931209 | WO 1993-EP1097 | 19930505 |
| | W: AU, BB, BG, BR, CA, CZ, FI, HU, JP, KP, KR, LK, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SK, UA, VN | | | | |
| | RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG | | | | |
| | CA 2109514 | AA | 19931123 | CA 1993-2109514 | 19930505 |
| | AU 9340647 | A1 | 19931230 | AU 1993-40647 | 19930505 |
| | AU 662226 | B2 | 19950824 | | |
| | EP 604602 | A1 | 19940706 | EP 1993-909898 | 19930505 |
| | R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LI, LU, MC, NL, PT, SE | | | | |
| | JP 06509363 | T2 | 19941020 | JP 1993-500115 | 19930505 |
| | BR 9305532 | A | 19941220 | BR 1993-5532 | 19930505 |
| | US 5312800 | A | 19940517 | US 1993-64243 | 19930520 |
| | ZA 9303548 | A | 19931122 | ZA 1993-3548 | 19930521 |
| PRAI | CH 1992-1655 | | 19920522 | | |
| | WO 1993-EP1097 | | 19930505 | | |

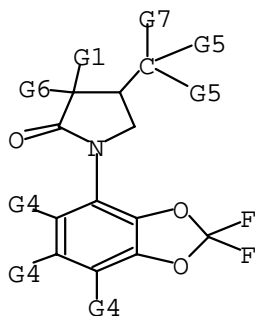
GI



(continued on next page)

AB Pyrrolidinones I (R = H, (substituted) C1-C4 alkyl, (substituted) Ph; R0 = H, halo; R1, R2 = H, C1-C4alkyl; X = H, Br, Cl, F, NC, R1O2C, R2R1NCO, H2N, R3CONH, wherein R3 = C1-4 alkyl; Y = H, Br, Cl) and the diastereoisomeric forms thereof, are prepd. N-allyl-N-(dichloroacetyl)-2,3-(difluoromethylenedioxy)aniline (prepn. given) was added to Cu(I)Cl and 2,2'-bipyridyl in diethylene glycol di-Me ether to give I (R = R0 = R1 = R2 = H, X = Y = Cl) which ar 4 kg/ha completely controlled Avena, Setaria, Sinapis, and Stellaria.

MSTR 1



G6 = alkoxy carbonyl<(1-4)>
MPL: claim 1
STE: or diastereoisomeric forms

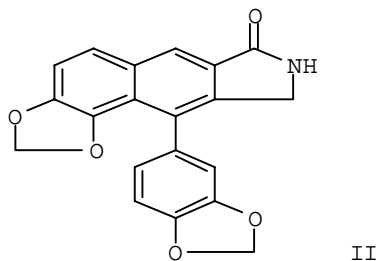
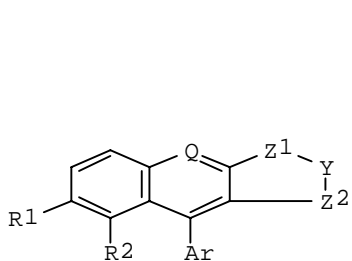
Displaying REGISTRY/Caplus hits:

=> D 62 69 BIB ABS HITSTR

L8 ANSWER 62 OF 209 CAPLUS COPYRIGHT 2000 ACS
 AN 1998:147315 CAPLUS
 DN 128:204735
 TI Preparation of naphtholactams and lactones for use as bone
 Morphogenetic protein active agents
 IN Marui, Shogo; Hazama, Masatoshi; Notoya, Kohei; Ogino, Masaki
 PA Takeda Chemical Industries, Ltd., Japan; Marui, Shogo; Hazama,
 Masatoshi; Notoya, Kohei; Ogino, Masaki
 SO PCT Int. Appl., 300 pp.
 CODEN: PIXXD2
 DT Patent
 LA English
 FAN.CNT 1

*Patent reference
from Caplus.*


| | PATENT NO. | KIND | DATE | APPLICATION NO. | DATE |
|------|---|------|----------|-----------------|----------|
| PI | WO 9807705 | A1 | 19980226 | WO 1997-JP2858 | 19970819 |
| | W: AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CU, CZ, EE, GE, HU, IL, IS, KG, KR, KZ, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, SL, TJ, TM, TR, TT, UA, US, UZ, VN, YU, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM | | | | |
| | RW: GH, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG | | | | |
| | AU 9738660 | A1 | 19980306 | AU 1997-38660 | 19970819 |
| | EP 920416 | A1 | 19990609 | EP 1997-935809 | 19970819 |
| | R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI | | | | |
| | JP 11005779 | A2 | 19990112 | JP 1997-224015 | 19970820 |
| PRAI | JP 1996-218353 | | 19960820 | | |
| | JP 1997-107617 | | 19970424 | | |
| | WO 1997-JP2858 | | 19970819 | | |
| OS | MARPAT 128:204735 | | | | |
| GI | | | | | |



(continued on next page)

AB Naphtholactams and lactones I [R1 = R2 = H, OH, alkyl, alkoxy, halogen; R1R2 = fused ring such as OCH2O, OCH2CH2O, CH2CH2O, etc.; Q = substituted or unsubstituted carbon, N, N(O); Z1 = C1-3-alkylene, oxo or thioxo contg. alkylene; Z2 = C1-3-alkylene, oxo or thioxo contg. alkylene; Y =methylene, S, S(O), NH, substituted N; Ar = aryl, heteroaryl], helioxanthanin analogs, were prepd. for use as bone morphogenetic protein active agents. Thus, lactam II was prepd. starting from helioxanthanin and was tested for induction of alk. phosphatase prodn. in cultured murine osteoblasts.

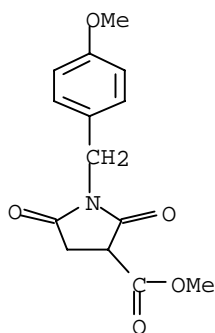
IT **203937-45-5**

RL: RCT (Reactant) 

(prepn. of naphtholactams and lactones for use as bone morphogenetic protein active agents)

RN 203937-45-5 CAPLUS

CN 3-Pyrrolidinecarboxylic acid, 1-[(4-methoxyphenyl)methyl]-2, 5-dioxo-, methyl ester (9CI) (CA INDEX NAME)



Note: The hit structure is a reactant used in the preparation of substances in this patent.

L8 ANSWER 69 OF 209 CAPLUS COPYRIGHT 2000 ACS

AN 1998:48125 CAPLUS

DN 128:88742

TI A New Palladium-Catalyzed Intramolecular Allylation to Pyrrolidin-2-ones

AU Giambastiani, Giuliano; Pacini, Barbara; Porcelloni, Marina; Poli, Giovanni

CS Dipartimento di Chimica Organica Ugo Schiff, Florence, I-50121, Italy

SO J. Org. Chem. (1998), 63(3), 804-807

CODEN: JOCEAH; ISSN: 0022-3263

PB American Chemical Society

DT Journal

LA English

OS CASREACT 128:88742

AB A novel palladium(0)-catalyzed cyclization of (Z)-RCH2CON(CH2Ph)CH2CH:CHCH2OAc [I, R = CO2Me, Ac, CN, SO2Ph, SPh, P(O)(OEt)2, Cl] to 3,4-disubstituted pyrrolidin-2-ones has been developed. The new approach relies upon the concomitant generation

Journal reference from CAplus.

(continued on next page)

of stabilized acetamide enolate anions and of a .pi.-allyl-palladium appendage, properly tethered by a nitrogen atom. Reaction conditions have been optimized for I [R =CO2Me] and then applied to the other substrates. A broad range of acetamide anion stabilizers [R = CO2Me, Ac, Cn, P(O)(OEt)2, SO2Ph] allowed the desired intramol. C-C bond formation. The cyclizations gave exclusively 5-exo-trig ring closure, thereby affording the 3,4-disubstituted pyrrolidin-2-ones with total diastereoselection. Complete trans preference was unequivocally demonstrated for the model reaction via a NOESY expt.

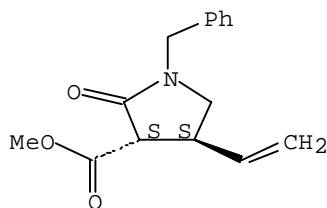
IT **201045-68-3P**

RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation)
(palladium-catalyzed intramol. allylation of
acetoxymethylacetamides to pyrrolidinones)

RN 201045-68-3 CAPLUS

CN 3-Pyrrolidinecarboxylic acid, 4-ethenyl-2-oxo-1-(phenylmethyl)-,
Methyl ester, (3R,4R)-rel- (9CI) (CA INDEX NAME)

Relative stereochemistry.



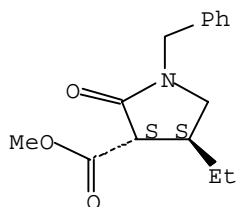
IT **142613-06-7P**

RL: SPN (Synthetic preparation); PREP (Preparation)
(palladium-catalyzed intramol. allylation of
acetoxymethylacetamides to pyrrolidinones)

RN 142613-06-7 CAPLUS

CN 3-Pyrrolidinecarboxylic acid, 4-ethyl-2-oxo-1-(phenylmethyl)-,
Methyl ester, trans- (9CI) (CA INDEX NAME)

Relative stereochemistry.



*This answer had two hit
structures in the
REGISTRY search.*

Helpful HINT

The final CASLINK answer set can be further refined with CPlus search terms. CASLINK does all the processing needed for this type of search to work.

```
=> S L8 AND PATENT/DT
```

```
•  
•  
•
```

```
DUP REM L13 L12 L9
```

```
L13 HAS NO ANSWERS
```

```
PROCESSING COMPLETED FOR L13
```

```
PROCESSING COMPLETED FOR L12
```

```
PROCESSING COMPLETED FOR L9
```

```
L14          87 DUP REM L13 L12 L9 (6 DUPLICATES REMOVED)  
              ANSWERS '1-54' FROM FILE MARPAT
```

Illustration: FQHIT vs. FHIT in MARPAT

With *FQHIT display format*, only the portions of the Markush structure that caused it to be a "hit" are displayed. The format is useful to

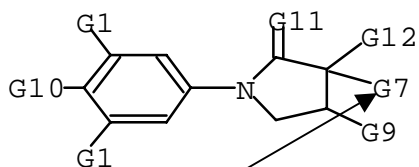
- Get a general idea of the context in which a structure hit
- Determine if a structure query is too broad
- Display large numbers of MARPAT hits

=> D 5 FQHIT

L8 ANSWER 5 OF 209 MARPAT COPYRIGHT 2000 ACS

DUPLICATE 6

MSTR 1



G7 = alkoxy carbonyl<(1-8)> (SO (1-) G8)

G11 = O

MPL: claim 1

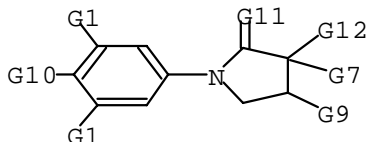
With *FHIT display format*, all G-group definitions are shown. It is useful when you need to see the entire context in which a structure hit.

=> D 5 FHIT

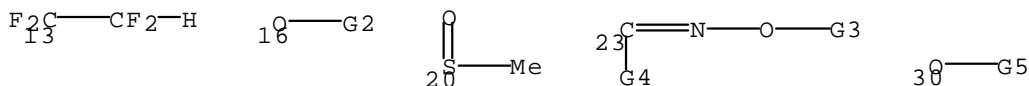
L8 ANSWER 5 OF 209 MARPAT COPYRIGHT 2000 ACS

DUPLICATE 6

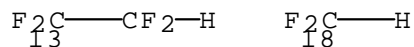
MSTR 1



G1 = H / X / CF3 / CN / Me / 13 / 16 / SMe / 20 / SO2Me / 23 / 30 / (SC Cl)



G2 = 13 / 18 / CF3



G3 = Me / CPh

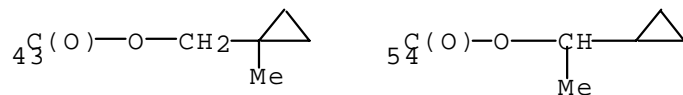
G4 = H / Me

G5 = pyridyl (SO (1-) G6)

G6 = X / CF3

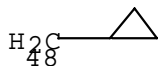
G7 = X / CN / CO2H / **alkoxycarbonyl<(1-8)>** (SO (1-) G8) / cycloalkyloxycarbonyl<(3-8)> (SO (1-) G13) / (SC Cl / CO2Et) / (EX 54 / 43)

The parts that caused the answer to be a "hit" are highlighted.



(continued on next page)

G8 = X / CF3 / Ph / cycloalkyl<(3-7)>
G9 = alkyl<(1-4)> (SO (1-) cyclopropyl) /
cycloalkyl<(3-4)> (SO (1-) Me) / alkenyl<(2-4)> / (SC Et /
CH=CH2) / (EX 48)



G10 = H / X / (SC F)
G11 = O / S
G12 = H / X / (SC Cl)
G13 = X / CF3 / Ph / alkyl<(1-5)>
MPL: claim 1

Helpful HINT

Markush displays can be very complicated.

Option: Consider displaying the associated CAplus abstract. Many contain structural summaries of the claimed substances.

Reviewing MARPAT Markush Displays

Markush displays in MARPAT consist of

- Base structure
- G-groups defining the variability in the structure

| G-groups in MARPAT displays may contain | Example |
|--|---|
| Real atoms | G1 = O / S / NH / CH2 |
| Variable groups | G2 = H / X / CF3 / CN / Me |
| Variable groups with deeper levels of definition | G3 = OH // Hy <EC (3-7) A (1-2) Q (1-2) N (0) OTHERQ, RC (1), RS (1) X7> (SO) / NMe2 |
| Structural fragments identified by node numbers | G4 = 13 / 18 / CF3 <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> $\begin{array}{c} \text{F}_2\text{C} \\ \\ \text{I}_3 \end{array} \text{---CF}_2\text{---H}$ </div> <div style="text-align: center;"> $\begin{array}{c} \text{F}_2\text{C} \\ \\ \text{I}_8 \end{array} \text{---H}$ </div> </div> |
| Generic text shortcuts | G5 = OH / alkoxy / X / NH2 / alkoxycarbonyl / CO2H / |
| Textual information describing a non-structural entity | G6 = R < TX " protecting group "> / CH2Ph / CH2CH=CH2 |

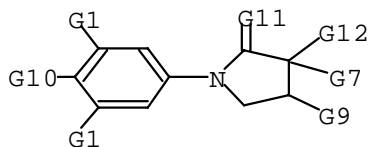
Example:

=> d 5 fhit

L8 ANSWER 5 OF 209 MARPAT COPYRIGHT 2000 ACS

DUPLICATE 6

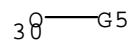
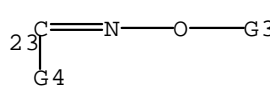
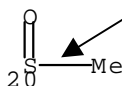
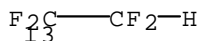
MSTR 1



Base structure with G-groups

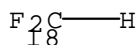
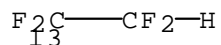
G1 = H / X / CF3 / CN / Me / 13 / 16 / SMe / 20 / SO2Me
23 / 30 / (SC Cl)

G-group options are separated by slashes.



G2 = 13 / 18 / CF3

G-groups in structural fragments are also defined.



G3 = Me / CPh

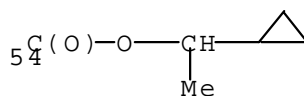
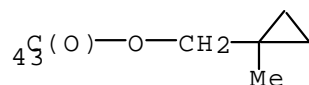
G4 = H / Me

G5 = pyridyl (SO (1-) G6)

G6 = X / CF3

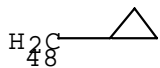
G7 = X / CN / CO2H / **alkoxycarbonyl<(1-8)>** (SO (1-) G8) / cycloalkyloxycarbonyl<(3-8)> (SO (1-) G13) / (SC Cl / CO2Et) / (EX 54 / 43)

Textual information may further define generic fragments.



(continued on next page)

G8 = X / CF3 / Ph / cycloalkyl<(3-7)>
G9 = alkyl<(1-4)> (SO (1-) cyclopropyl) /
cycloalkyl<(3-4)> (SO (1-) Me) / alkenyl<(2-4)> / (SC Et /
CH=CH2) / (EX 48)



G10 = H / X / (SC F)
G11 = O / S
G12 = H / X / (SC Cl)
G13 = X / CF3 / Ph / alkyl<(1-5)>
MPL: claim 1



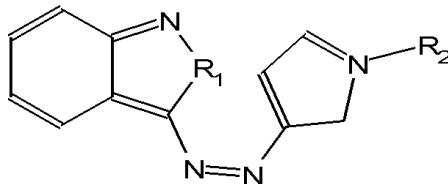
The source of the information in the Markush structure. This information is searchable in the Basic Index.

Review: Searching CASLINK

| | | |
|--------|--------------------------------|--------------------------|
| Step 1 | Build and save the structure | |
| Step 2 | Upload the Structure | |
| | A. Enter CASLINK | => FILE CASLINK |
| | B. Perform the upload | |
| | C. Verify the structure online | => D L# |
| Step 3 | Test the query | |
| | A. Run a SAMPLE search | => S L# SSS/CSS SAM |
| | B. Evaluate answers | => D SCAN L# |
| Step 4 | (Option) Revise the query | |
| Step 5 | Run a FULL substructure search | => S L# SSS/CSS FULL |
| Step 6 | Display results | |

Skills Practice

1. Use CASLINK to search REGISTRY/CAplus and MARPAT to locate references discussing substances with the following structure



R₁ = O or S

R₂ = nitrogen atom in a ring or chain

Any type of substitution is allowed at all open sites

Rings may be isolated or embedded in larger ring systems

- a. Will the query run to completion in REGISTRY and MARPAT?
- b. Do the SAMPLE search structures meet the query requirements?
- c. Record the number of hits retrieved in the FULL search:

| | |
|----------------|-----|
| REGISTRY | ___ |
| MARPAT | ___ |
| MARPATpreviews | ___ |
| CAplus | ___ |

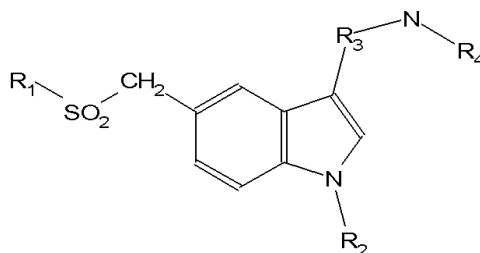
- d. Display the final answer set using the following formats:

For MARPAT answers use BIB ABS FQHIT

For CAplus answers use BIB ABS HITSTR

- e. Display the full Markush structure for one of the MARPAT answers.

2. Use CASLINK to search REGISTRY/CAplus and MARPAT to locate referenc discussing substances with the following structure



R1 = nitrogen in a ring or chain

R2 = anything, including hydrogen

R3,R4 = alkyl chain with any type of substitution

Nitrogen-containing ring may be isolated or embedded in a larger ring system

Any substitution is allowed at all open sites

- Do the sample structures meet the query requirements?
- Record the number os hits retrieved in the FULL search:

REGISTRY _____

MARPAT _____

MARPATpreviews _____

CAplus _____

Unique hits after duplicates were removed _____

- Display the final answer set using the following formats:

For MARPAT answers use BIB ABS FQHIT

For CAplus answers use BIB ABS HITSTR

- Display the full MARKUSH structure for one of the MARPAT answers.

PRECISION TOOLS

In this section, you will learn to

- Specify Match Level to achieve desired recall from MARPAT

Balancing Precision and Recall in MARPAT

The structure search process matches elements of the structure query with database structures. Because of the differences in the structures in REGISTRY and MARPAT, unique retrieval options are possible in MARPAT.

| Database | Retrieval possibilities |
|----------|---|
| REGISTRY | Specific atoms |
| MARPAT | <ul style="list-style-type: none">■ Specific atoms■ Generic groups that match the query definition |

Each atom in a query structure is assigned a *default Match Level*. Match Level controls how query atoms match specific atoms and generic groups in the MARPAT database. Three Match Level options are possible:

- Atom – retrieves only real atoms
- Class – retrieves real atoms and generic groups
- Any – retrieves real atoms, generic groups, and R-groups containing unstructurable substituents

note

Match Level is ignored in REGISTRY

Match Level in MARPAT

| Match | Retrieval | Example | |
|--------------|---|---------------|---|
| | | Query atom | Database hit |
| Atom | Real atoms | Br | BR |
| | | X | Br, Cl, F, etc |
| | | Pyridine ring | Pyridine ring |
| | | Hy | Pyridine, thiophene, benzofuran, etc |
| Class | <ul style="list-style-type: none"> ■ Real atoms ■ Generic groups (Q X M Ak Hy Cb Cy) | Br | Br, X |
| | | X | Br, Cl F, etc., X |
| | | Pyridine ring | Pyridine ring, Hy |
| | | Hy | Pyridine, thiophene, benzofuran, etc., Hy |
| Any | <ul style="list-style-type: none"> ■ Real atoms ■ Generic groups ■ Any R^b | Br | Br, X, r |
| | | Pyridine ring | Pyridine ring, Hy, R |

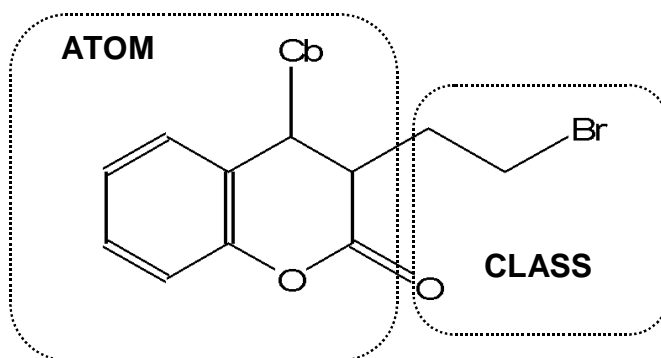
^a Class assigned to rings retrieves the R group defined as "group to form a ring". Generally, class does not retrieve R groups. ^b"R" groups represent unstructurable portions of the database structure, e.g., "organic group", "electron withdrawing group". ANY matches any R group, regardless of its definition.

Match Level Assumptions in STN Express

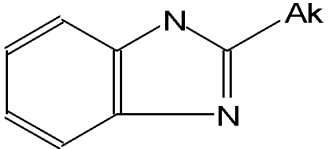
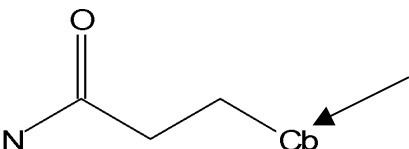
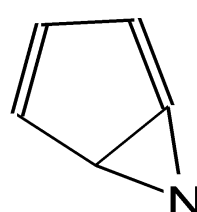
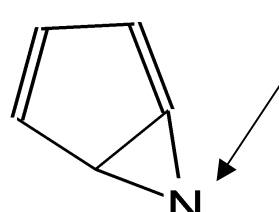
STN Express assumes a Match Level for each specific and generic atom in a query structure.

| By default, Match Level is set as | For the following parts of a structure |
|-----------------------------------|---|
| Atom | <ul style="list-style-type: none">■ All atoms in a ring system■ Hy, Cb, Cy |
| Class | All chain atoms in a structure |

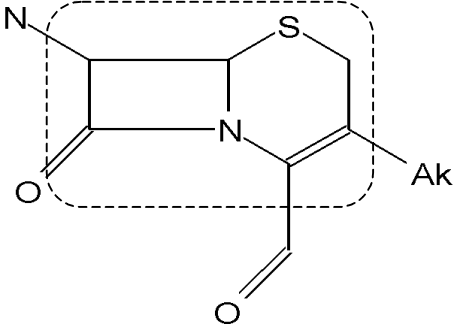
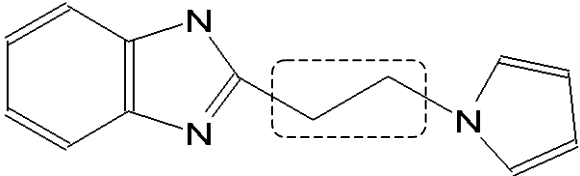
Illustration



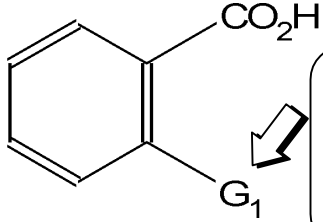
Tips for setting CLASS Match Levels

| If in the query there is | Consider this Match Level option |
|---|---|
| A specific <i>ring system</i> | <p>When you want to retrieve the real atom ring system, plus generic groups that encompass the ring definition, change Match Level for all atoms in the ring to CLASS</p> <div style="display: flex; align-items: center; justify-content: center;">  <div style="border: 1px solid black; border-radius: 10px; padding: 5px; margin-left: 20px;"> <p><i>Change Match Level for all ring atoms to CLASS.</i></p> </div> </div> |
| <i>Hy, Cb, Cy</i> | <p>When you want to retrieve real atom rings matching the ring generic group, plus the generic group, change Match Level on the generic system to CLASS</p> <div style="display: flex; align-items: center; justify-content: center;">  <div style="border: 1px solid black; border-radius: 10px; padding: 5px; margin-left: 20px;"> <p><i>Change Match Level to CLASS.</i></p> </div> </div> |
| M, Q, X, or A as <i>part of a ring system</i> | <p>When you want to match specific elements, as well as the corresponding generic groups, change the Match Level on those atoms to CLASS.</p> <div style="display: flex; align-items: center; justify-content: center;">  </div> |
| M, Q, X, or A as <i>part of a ring system</i> | <p>When you want to match specific elements, as well as the corresponding generic groups, change the Match Level on those atoms to CLASS.</p> <div style="display: flex; align-items: center; justify-content: center;">  <div style="border: 1px solid black; border-radius: 10px; padding: 5px; margin-left: 20px;"> <p><i>Change Match Level to CLASS. Rest of ring atoms are ATOM.</i></p> </div> </div> |

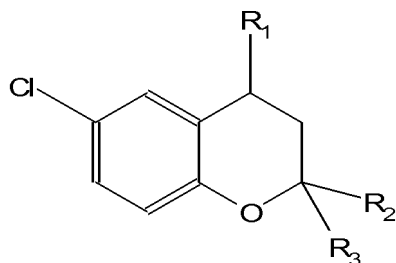
Tips for setting ATOM Match Levels

| If in the query there is | Consider this Match Level option |
|---|---|
| <p>An important ring system that must be explicitly present in each retrieval</p> | <p>Leave the default for the ring system at ATOM</p>  |
| <p>A chain spacer of a certain length that must be present</p> | <p>Change the Match Level for the chain to ATOM</p>  |

Tips for setting ANY Match Levels

| If in the query there is | Consider this Match level option: |
|--|--|
| <p>A substituent that is often described generically in patent claims, e.g., an electron withdrawing group</p> | <p>Assigning Match Level ANY to those types of substituents.</p>  <p><i>Change Match Level for the G1 substituents to ANY to also match R which might be defined as an "electron withdrawing group."</i></p> <p>G1=NO₂, CN, X</p> |

Search Question: *Locate references discussing compounds with the following structure:*



R1 = heterocyclic ring with at least one =O attached

R2, R3 = any type of carbon chain (substituted or unsubstituted)

The oxygen-containing ring may be isolated or embedded in a larger ring system

Any substitution at all open sites

All of the atoms in the structure, except for the benzopyran ring, may match real atoms or generic groups in Markush structures.

The benzopyran ring may match only real atoms.

Worksheet

| Requirement | Your action |
|--|--------------------|
| <p>R1 = heterocyclic ring system</p> <p>R1 must have a least one =O attached to it</p> <p>R2, R3 = carbon chain which optionally may be substituted</p> <p>No ring fusion</p> <p>Substitution allowed at all other open sites</p> <p>Cl may match Cl or X</p> <p>R1, R2 may match a real atoms or generic groups</p> <p>The benzopyran ring cannot match generic ring descriptions</p> | |

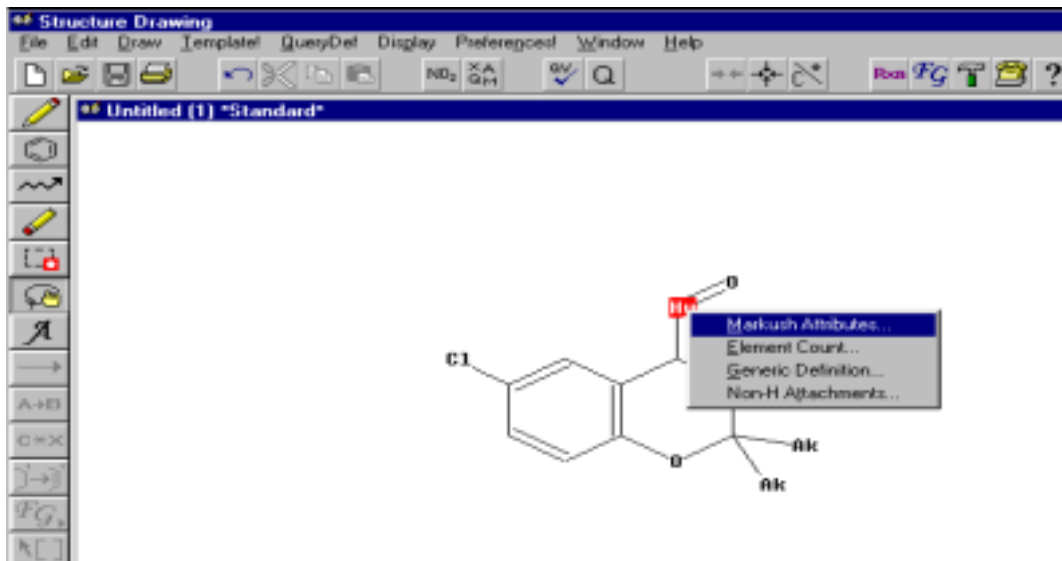
Assigning Match Levels in STN Express

Match Levels may be changed on a single node or a group of nodes.

Single nodes

To change the Match Level on a single node, do the following:

1. Right click on the node. A pop-up menu appears. Select *Markush Attributes*.



2. The **Markush Attributes** dialog box appears. Do the following:
 - a. Click the radio button associated with the Match Level of interest.
 - b. Click **OK**.



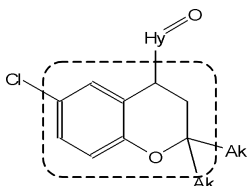
note

The "Limited" check box specifies how element counts will match candidate generic groups with Match Level CLASS.

When checked, only generic groups with assigned element counts that match those in the query structure are retrieved. For best precision, "Limited" should be checked.

Example:

"Limited" on the circled part of the structure



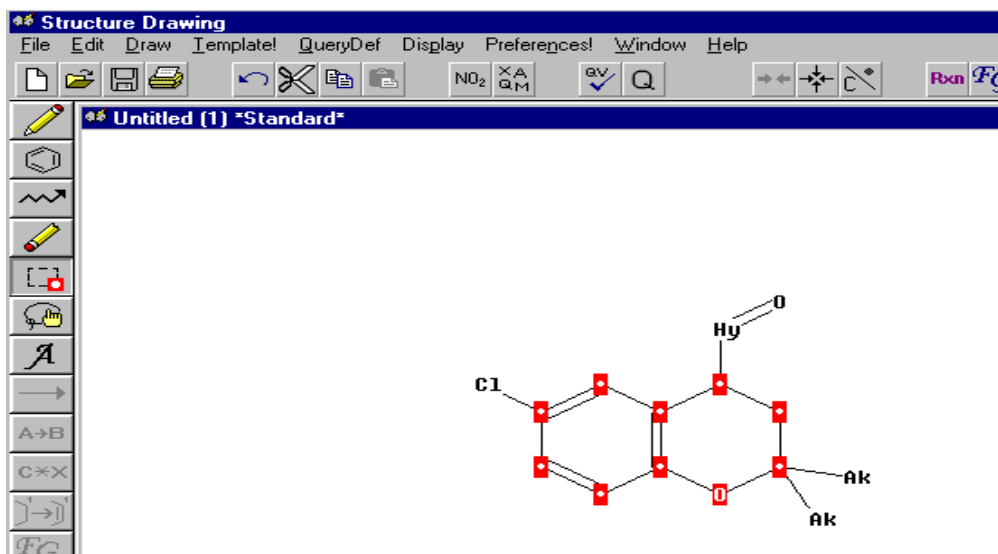
will retrieve:

- Benzopyran real atom structure
- Hy<EC (1-2)Q (1-) O>
- Hy<EC (1-2) Q (0-1) O (8-10) C, RC (1-2)>

Multiple nodes

To change the Match Level on multiple nodes all at once, do the following:

1. Highlight the nodes using the highlighting tool.



- From the **Query Def** pull-down menu, select *Markush Attributes*.

An alert dialog box appears. Click **OK**.

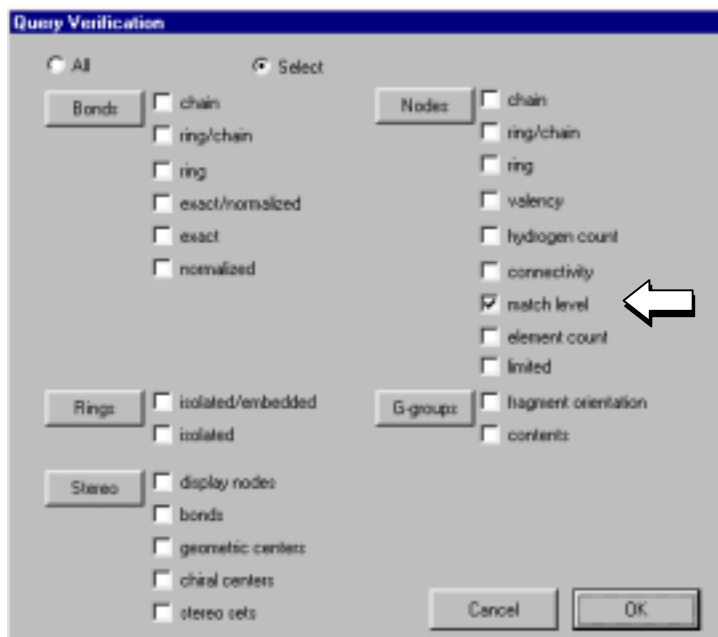
The *Markush Attributes* dialog box appears. Do the following:

- Click the radio button associated with the Match Level of interest.
- Click **OK**.

Verifying Match Levels Assignments

To verify Match Level assignments for nodes in a structure, do the following:

- From the **Query Def** pull-down menu, select *Query Verification*.
- The **Query Verification** dialog box appears. Do the following:
 - Click the *Select* radio button.
 - Click in the *match level* box.
 - Click **OK**.



- A **Query Verification** pop-up dialog appears and Match Levels for all atoms display. Click **OK**.

Uploading the query

=> FILE CASLINK

=>

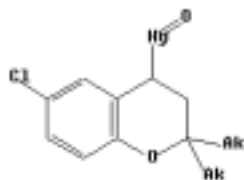
Uploading mar2.str

L1 STRUCTURE UPLOADED

=> D L1

L1 HAS NO ANSWERS

L1 STR



CLASS match level on:
Cl, Ak, O, Hy

ATOM match level on:
All atoms in the benzopyran ring.

Testing the query

=> S L1 SSS SAM

```
S L1 SSS SAM FILE=REGISTRY
SAMPLE SEARCH INITIATED 16:02:42 FILE 'REGISTRY'
SAMPLE SCREEN SEARCH COMPLETED - 967 TO ITERATE
100.0% PROCESSED 967 ITERATIONS 0 ANSWERS
SEARCH TIME: 00.00.01
```

```
FULL FILE PROJECTIONS: ONLINE **COMPLETE**
                        BATCH **COMPLETE**
PROJECTED ITERATIONS: 17476 TO 21204
PROJECTED ANSWERS: 0 TO 0
```

```
L2 0 SEA SSS SAM L1
  1 FILES SEARCHED...
```

```
S L2 SSS SAM FILE=MARPAT
SAMPLE SEARCH INITIATED 16:02:43 FILE 'MARPAT'
SAMPLE SCREEN SEARCH COMPLETED - 90 TO ITERATE
100.0% PROCESSED 90 ITERATIONS 4 ANSWERS
SEARCH TIME: 00.00.05
```

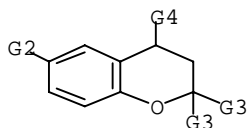
```
FULL FILE PROJECTIONS: ONLINE **COMPLETE**
                        BATCH **COMPLETE**
PROJECTED ITERATIONS: 1233 TO 2367
PROJECTED ANSWERS: 4 TO 200
```

```
L3 4 SEA SSS SAM L1
  1 FILES SEARCHED...
```

=> D SCAN L3 FQHIT

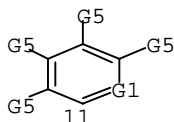
```
L3 4 ANSWERS MARPAT COPYRIGHT 2000 ACS
```

MSTR 1



(continued on next page)

G1 = N
G2 = Cl
G3 = alkyl<(1-8)>
G4 = 11

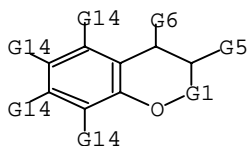


G5 = OH
MPL: claim 1

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

L3 4 ANSWERS MARPAT COPYRIGHT 2000 ACS

MSTR 1A



G1 = 11



G2 = alkyl<(1-6)>
G3 = alkyl<(1-6)>
G7 = Hy<EC (4-5) C (1-2) N (0) OTHERQ, AN (1-) N (1-) C,
BD (1) D (0) T, RC (1), RS (1) E6 (0) OTHER> (SO)
G14 = X
G28 = O
DER: and salts
MPL: claim 1

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

Note that FQHIT shows only the fragments that caused the structure to hit. For example, the complete definition of G6 is not shown - only the part that caused the answer to be a hit (G7) is shown.

Running the FULL search

```
=> S L1 SSS FULL

S L1 SSS FUL FILE=REGISTRY
FULL SEARCH INITIATED 16:04:47 FILE 'REGISTRY'
FULL SCREEN SEARCH COMPLETED - 19531 TO ITERATE
100.0% PROCESSED 19531 ITERATIONS 6 ANSWERS
SEARCH TIME: 00.00.03

L4          6 SEA SSS FUL L1
  1 FILES SEARCHED...

S L4 SSS FUL FILE=MARPAT
FULL SEARCH INITIATED 16:04:51 FILE 'MARPAT'
FULL SCREEN SEARCH COMPLETED - 1935 TO ITERATE
  85.7% PROCESSED 1658 ITERATIONS 32 ANSWERS
100.0% PROCESSED 1935 ITERATIONS 42 ANSWERS
SEARCH TIME: 00.00.24

L5          42 SEA SSS FUL L1
  1 FILES SEARCHED...

S L5 SSS FUL FILE=MARPATPREV
FULL SEARCH INITIATED 16:05:17 FILE 'MARPATPREV'
FULL SCREEN SEARCH COMPLETED - 12 TO ITERATE
100.0% PROCESSED 12 ITERATIONS 0 ANSWERS
SEARCH TIME: 00.00.02

L6          0 SEA SSS FUL L1
  1 FILES SEARCHED...

S L4 FILE=CAPLUS
L7          14 FILE CAPLUS
  1 FILES SEARCHED...

DUP REM L6 L5 L7
L6 HAS NO ANSWERS
PROCESSING COMPLETED FOR L6
PROCESSING COMPLETED FOR L5
PROCESSING COMPLETED FOR L7
L8          51 DUP REM L6 L5 L7 (5 DUPLICATES REMOVED)
              ANSWERS '1-42' FROM FILE MARPAT
              ANSWERS '43-51' FROM FILE CAPLUS
```

Displaying results

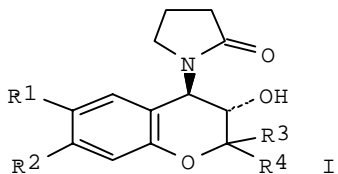
Answers from MARPAT

=> D 1 BIB ABS FQHIT

L8 ANSWER 1 OF 51 MARPAT COPYRIGHT 2000 ACS DUPLICATE 1
 AN 121:255814 MARPAT
 TI 3-hydroxy-4-(2-oxopyrrolidonyl)benzopyran potassium channel
 activators
 IN Eszenyi, Tibor; Sebok, Peter; Frank, Laszlo; Papp, Gyula; Timar,
 Tibor; Bartik, Tamas
 PA Alkaloida Vegyeszeti Gyar Rt., Hung.
 SO PCT Int. Appl., 46 pp.
 CODEN: PIXXD2
 DT Patent
 LA English
 FAN.CNT 1

| | PATENT NO. | KIND | DATE | APPLICATION NO. | DATE |
|------|---|------|----------|-----------------|----------|
| PI | WO 9414799 | A1 | 19940707 | WO 1993-HU79 | 19931220 |
| | W: CZ, JP, KZ, LV, PL, RU, SK, UA, US | | | | |
| | RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE | | | | |
| | HU 72741 | A2 | 19960528 | HU 1992-4049 | 19921219 |
| | EP 674633 | A1 | 19951004 | EP 1994-902996 | 19931220 |
| | R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LI, LU, MC, NL, PT, SE | | | | |
| | LV 10955 | B | 19960620 | LV 1995-178 | 19950616 |
| | US 5703113 | A | 19971230 | US 1995-481440 | 19951108 |
| PRAI | HU 1992-4049 | | 19921219 | | |
| | WO 1993-HU79 | | 19931220 | | |

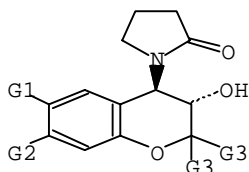
GI



AB The title compds [I; R1 = C1-6 alkyloxy, alkenyloxy, alkynyloxy, cycloalkyloxy, aryloxy, aralkyloxy, hydroxy, nitro, cyano, formyl, C1-6 acyl, C1-6 alkylcyano group; R2 = C1-6 alkyloxy, alkenyloxy, alkynyloxy, cycloalkyloxy, aryloxy, aralkyloxy, hydroxyl; R3, R4 = C1-6 alkyl group; R1R2 = alkylendioxy], useful as K channel activators for treatment of hypertension, are prepd. by reacting a (un)substituted 4-chromanone with metal hydrides, dehydrating the resulting 4-chromanol, converting the resultant 2H-chromene into a bromohydrin, and reacting the bromohydrin with 2-pyrrolidone. Thus, (+/-)-6,7-dimethoxy-3,4-dihydro-2,2-dimethyl-trans-4-(2-oxo-1 pyrrolidonyl)-2H-benzo[b]pyran-3-ol was prepd. And demonstrated pA2 of 5.78.+-.0.24 on isolated rabbit arteria mesenteric anterior (after Auton, 1987).

(continued on next page)

MSTR 1

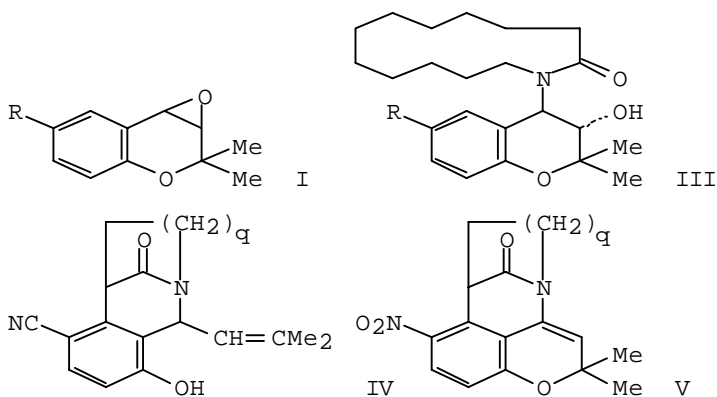


G1 = X
 G3 = alkyl<(1-6)>
 MPL: claim 1

=> D 45 51 BIB ABS HITSTR

Answers from CAPLUS/REGISTRY.

L8 ANSWER 45 OF 51 CAPLUS COPYRIGHT 2000 ACS
 AN 1988:112204 CAPLUS
 DN 108:112204
 TI Unexpected products from the reactions of 6-cyano-, and
 6-nitro-3,4-epoxy-3,4-dihydro-2,2-dimethyl-2H-1-benzopyran with
 azacyclotridecanone
 AU Cassidy, Frederick; Evans, John M.; Smith, Duncan M.; Stemp,
 Geoffrey
 CS Med. Res. Cent., Beecham Pharm. Res. Div., Pinnacles/Harlow/Essex,
 CM19 5AD, UK
 SO Tetrahedron Lett. (1987), 28(21), 2403-6
 CODEN: TELEAY; ISSN: 0040-4039
 DT Journal
 LA English
 OS CASREACT 108:112204
 GI



AB Addn. of epoxybenzopyrans I (R = H, Cl) with azacyclotridecanone
 (II) gave 10% and 34% benzopyrans III (R = H, Cl), resp., whereas
 the same reaction of I (R = cyano, NO₂) with II gave 26% pyran ring
 cleavage product IV and 3% tetracyclic chromene V, resp.

(continued on next page)

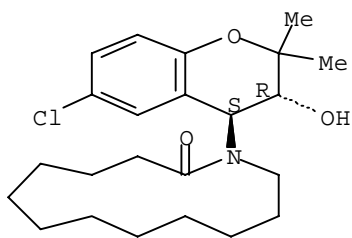
IT 113123-16-3P

RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, by addn. of epoxybenzopyran with azacyclotridecanone)

RN 113123-16-3 CAPLUS

CN Azacyclotridecan-2-one, 1-(6-chloro-3,4-dihydro-3-hydroxy-2,2-dimethyl-2H-1-benzopyran-4-yl)-, trans- (9CI) (CA INDEX NAME)

Relative stereochemistry.



L8 ANSWER 51 OF 51 CAPLUS COPYRIGHT 2000 ACS

AN 1983:488053 CAPLUS

DN 99:88053

TI Pharmaceutically active benzopyran compounds

IN Evans, John Morris; Buckingham, Robin Edward; Willcocks, Kenneth

PA Beecham Group PLC, UK

SO Eur. Pat. Appl., 65 pp.

CODEN: EPXXDW

DT Patent

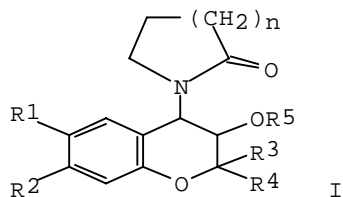
LA English

FAN.CNT 2

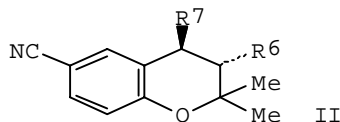
| | PATENT NO. | KIND | DATE | APPLICATION NO. | DATE |
|----|---|------|----------|-----------------|----------|
| PI | EP 76075 | A1 | 19830406 | EP 1982-304946 | 19820921 |
| | EP 76075 | B1 | 19861120 | | |
| | R: AT, BE, CH, DE, FR, GB, IT, LI, LU, NL, SE | | | | |
| | AT 23718 | E | 19861215 | AT 1982-304946 | 19820921 |
| | FI 8203272 | A | 19830326 | FI 1982-3272 | 19820923 |
| | FI 73675 | B | 19870731 | | |
| | FI 73675 | C | 19871109 | | |

•
•
•

GI



I



II

(continued on next page)

AB Benzopyrans I [one of R1 and R2 = H, the other alkyl- or alkoxy-carbonyl, allyl-carbonyloxy, alkylhydroxymethyl, NO₂, cyano, Cl, CF₃, alkylsulfinyl, -sulfonyl, alkoxy-sulfinyl, -sulfonyl, carboxamido, alkoxy-carboxamido, alkyl (un)substituted aminosulfinyl, -sulfonyl, -carbonyl, alkylsulfinyl-, -sulfonylamino, alkoxy-sulfinyl-, -sulfonylamino, CH:CH₂ terminally substituted by alkyl-carbonyl, NO₂, cyano, C(C1-6 alkyl):NOH, C(C1-6 alkyl):NNH₂; one of R3 and R4 = H, C1-4 alkyl, the other C1-4 alkyl, CR₃R₄ = C3-6 spiroalkyl; R5 = H, C1-3 alkyl, C1-8 acyl; n = 1, 2; lactam group trans to OR₅], useful as antihypertensives, were prepd. Etherifying 4-NCC₆H₄OH with Me₂CClC.tplbond.CH in H₂O-CH₂Cl₂ contg. NaOH and PhCH₂N+Me₃ OH- 5.5 days at room temp. gave 4-NCC₆H₄OCMe₂C.tplbond.CH which was cyclized to chromene II (R₆R₇ = bond). Bromination in Me₂SO contg. A little H₂O gave bromohydrin II (R₆ = Br, R₇ = OH) which was cyclized with NaOH in H₂O-dioxane to give epoxide II (R₆R₇ = O). Aminolysis of this with H₂N(CH₂)₃CO₂H and NaHCO₃ in refluxing H₂O-EtOH, sepn. of the most polar product by chromatog., and lactamizing in refluxing PhMe 2 h gave trans-I (R₁ = cyano, R₂ = R₅ = H, R₃ = R₄ = Me, n = 1) (III). At 1 mg/kg orally (rats) III gave -47.+- .1 and -35 .+- .4% change in systolic blood pressure after 1 and 2 h, resp.

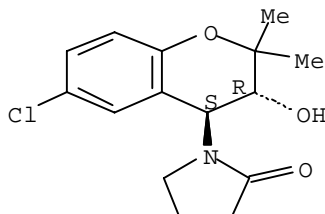
IT 86776-76-3P

RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of)

RN 86776-76-3 CAPLUS

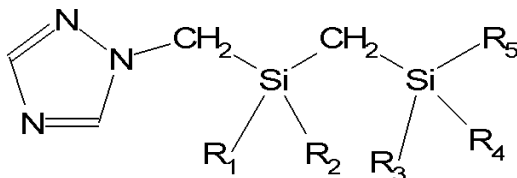
CN 2-Pyrrolidinone, 1-(6-chloro-3,4-dihydro-3-hydroxy-2,2-dimethyl-2H-1-benzopyran-4-yl)-, trans- (9CI) (CA INDEX NAME)

Relative stereochemistry.



Skills Practice

1. Use CASLINK to locate patents on compounds with the following structure:



R₁, R₃, R₄, R₅ = an alkyl chain of any length with any type of substitution

R₂ = any ring system with any type of substitution

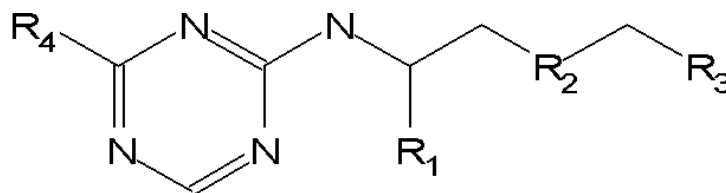
Nitrogen-containing ring may be isolated or embedded in a larger ring system

All the ring systems may match real atom rings or generic groups

The -CH₂-Si-CH₂-Si- chain may match only real atoms

The alkyl groups may match real atoms or generic groups

2. Use CASLINK to locate references discussing compounds with the following structure:



R₁ = carbon chain (substituted or unsubstituted) or a carbocyclic ring

R₂ = O, S, or N

R₃ = Any type of ring system

R₄ = Nitrogen in a chain

The nitrogen-containing ring may be isolated or embedded in a larger ring system

All of the atoms in the structure must match real atoms in Markush structures, except for R₁ and R₃ which may also match generic groups.

APPENDIX

In this section, you will learn to

- Generic groups
- MARPAT display fields
- Codes used in Markush structures, including generic group attributes, precision qualifiers, substitution qualifiers, and occurrence counts
- Generic text shortcuts

Generic Groups

- Ak** Any carbon chain (only first atom need be carbon); any bond value allowed
- Cy** Any cyclic group
- Hy** Any cyclic group with one (1) or more non-carbon atoms
- Cb** Any cyclic group with all carbon atoms

Fields in MARPAT Structure Displays

MSTR

Display label for the Markush Structure

Example: MSTR 4

Translation: This is the 4th Markush structure in the document

VAR G#

Defines the alternatives for a G#

Example: VAR G1 = O / S / 16 / NULL

Translation: G1 is O or S or node 16 or G1 is a direct bond
Node sixteen (16) appears in the structure diagram portion of the display

REP G# =

Defines a REpeating group and the number of times it repeats

Example: REP G3 = (0-7) CH2

Translation: CH2 repeats (0-7) times

CVA

The Conditional Variable statement, currently ignored at search time

Example: VAR G1 = H / OH / X
 VAR G2 = NO2 / Me / Et
 CVA = If G1 = OH THEN G2 = Me

Translation: If G1 is the OH alternative, then G2 must be Me and not NO2 or Et
Searched as G1 = H, OH, X and G2 = NO2, Me, Et (STN Express)
as VAR G1 = H / OH / X and VAR G2 = NO2 / Me / Et

DER

DERivative information that cannot be structured
Searched as single words in the Basic Index (default)

Example: DER: or salts or metal complexes

NTE

General NoTEs
Searched as single words in the Basic Index (default)

Example: NTE: substitution restricted

MPL

Location of the Markush structure in the patent
Searched as single words in the Basic Index (default)

Example: MPL: claim 1

STE

STEReochemistry
Searched as single words in the Basic Index (default)

Example: STE: 41,42-cis

Codes Found in Markush Structures

Generic Group Attributes - AN, AR, BD, CH, DC, EC, FA, RC, RS, TX

- Found in <> following the generic group they modify
- The marked attributes are searched as Generic Definitions (STN Express queries) or Generic Group Categories (STR command queries)
- All others are ignored while searching

| Abbreviation | Example | Translation |
|-----------------------------|------------------------------|---|
| AN Attachment Nodes | AN (3) A AN (2) N | Attached through 3 nodes of any kind (A) Attached through 2 N nodes |
| AR Aryl | AR (1-) AR (0) | 1 or more aromatic rings is not aromatic |
| BD* Bonds | BD (0) T BD (1-) D (2) SE | no triple (T) bonds 1 or more double (D) bonds and two single exact (SE) bonds |
| CH Charge | CH (2) + CH (1) +- | total of two (2) positive charges one (1) positive or negative charge |
| DC** Degree of Connectivity | DC (0) M3 DC (1-) M3 | no branching branched |
| EC*** Element Count | EC (1-8) C EC (2-3) N | 1-8 carbon (C) atoms 2-3 nitrogen (N) atoms |
| FA**** Fusion Atoms | FA (2-4) C FA (2-) C | 2-4 carbon (C) atoms fuse rings 2 or more carbon C atoms fuse rings |
| RC***** Ring Count | RC (2-) RC (1) | polycyclic with 2 or more rings monocyclic |

(continued on next page)

| | | | |
|----|-----------------|---|---|
| RS | Ring Size | RS (2-3) E6 RS (1) M5 (1) X6 and RC (1) | 2-3 6-membered rings 5- to 6-membered ring |
| TX | Text Qualifiers | R <TX “protecting group”> R <TX “residue”> | the patent said “protecting group” the patent said “residue” |

- * Searchable as Generic Definition *Saturated* or *Unsaturated* (STN Express) or *GGC SAT* or *GGC UNS* (STR command)
- ** Searchable as Generic Definition *Linear* or *Branched* (STN Express) or *GGC LIN* or *GGC BRA* (STR command)
- *** Searchable as Generic Definition *7 or more C* or *Less than 7 C* and *2 or more hetero atoms* or *Less than 2 hetero atoms* (STN Express) or *GGC HIC* or *GGC LOC* and *GGC HIQ* or *GGC LOQ* (STR command)
- **** Searchable as Generic Definition *Monocyclic* or *Polycyclic* (STN Express) or *GGC MCY* or *GGC PCY* (STR command)

Precision Qualifiers - EX and SC

- Found in parentheses at the end of the variable definition
- Additional alternatives for the G-group

EX

Alternatives are found in the EXamples (in the disclosure)

Example: VAR G1 = R<TX “leaving group”>/(EX Cl/ Br / I)

Translation: Cl, Br and I are alternatives for G1 and were found in the disclosure
VAR denotes variability G1 is the tag on the variable group

SC

Alternatives are Specifically Claimed

Example: VAR G3 = H / alkyl / alkoxy / (SC Me / OMe)

Translation: Methyl and methoxy are alternatives for G3 and are specifically claimed

Substitution Qualifiers - SO and SR

- Found in parentheses immediately after the group they modify
- In the current implementation, SO and SR are both searched as SO; i.e., as if the substituent were present on the group any number of times, including zero

SO

Alternative is Substituted Optionally by this group

Example: VAR G1 = Ph (SO X) / alkoxy carbonyl (SO CO2H)

Translation: Phenyl is substituted with zero (0) to five (5) halogens
Alkoxy carbonyl is substituted with zero (0) or more carboxy groups

SR

Alternative has Substitution Required by one (1) or more of this group

Currently searched as SO

Example: VAR G1 = Ph (SR X) / alkoxy carbonyl (SR CO2H)

Translation: Phenyl is substituted with one (1) to five (5) halogens
Alkoxy carbonyl is substituted with one (1) or more carboxy groups
Currently searched as zero (0) to five (5) halogens and zero (0) or more carboxy groups

Occurrence Counts

- Occurrence Counts are found in front of the alternative they modify
- Used to limit the substitution by specifying how many times an alternative is present
- The occurrence counts are not currently searchable

(n)

Alternative occurs exactly “n” times

Example: VAR G1 = (1) X / Me

Translation: Exactly one (1) of the G1’s is X
Searched as G1 = X,Me (STN Express); VAR G1=X/Me (STR command)

(n-)

Alternative occurs “n” or more times

Example: VAR G3 = (2-) H / loweralkyl / X

Translation: Two (2) or more of the G3’s are H
Searched as G3 = H, Ak , X (STN Express); VAR G3= H / Ak, X Me (STR command)

(-n)

Alternative occurs zero (0) to “n” times

Example: VAR G6 = (-1) H / alkyl / CH2Ph

Translation: Zero (0) or one (1) of the G6’s are H
Searched as G6 = H, Ak , CH2Ph (STN Express); VAR G6 = H / Ak, CH2Ph (STR command)

(m-n)

Alternative occurs “m” to “n” times

Example: VAR G2 = X / CF3 / Ph (SO (1-2) G3)

Translation: The phenyl (Ph) is optionally substituted by one (1) or two (2) of the variable G3’s
Searched as the phenyl (Ph) substituted by zero (0) to five (5) G3’s

Generic Text Shortcuts

Ak-based

Ak is a chain of 1 or more carbons, linear or branched, saturated or unsaturated

| | |
|------------------------|---|
| Alkanoyl | Alkyl -C(O)-, H - C(O) - |
| Alkenyl | Ak with one or more double bonds, no triple bonds, and two or more carbons, Ak<EC(2-) C,BD (1-) D (0) T> - An unsaturated monovalent radical chain of two or more carbons, branched or linear, containing one or more carbon-to-carbon double bonds, but no triple bonds Formed by the removal of one hydrogen from the corresponding alkene, e.g., CH ₂ - CH = CH - CH ₂ - |
| Alkenylene | - Ak<EC (2-) C,BD (1-) D (0)T> - An unsaturated divalent hydrocarbon chain radical of at least two carbon atoms containing one or more double bonds, but no triple bonds Formed by the removal of two hydrogens from the parent branched or linear alkene, e.g., - CH = CH - |
| Alkenylenedioxy | - O - alkenylene - O - |
| Alkoxy | alkyl - O - also called alkyloxy and alkoxy |
| Alkyl | Ak with all bonds single exact A totally saturated monovalent radical chain, branched or linear Formed from an alkane by removal of one hydrogen, e.g., Me-, Et-, t-Bu- |

| | |
|---------------------|--|
| Alkylene | - alkyl - A divalent saturated hydrocarbon radical Formed by the removal of two hydrogens from the branched or linear parent alkane, e.g., - CH ₂ - |
| Alkylidene | alkyl = A divalent alkyl radical that is attached to the parent by a double bond or two (2) single bonds from the same carbon, e.g., = CH - CH ₂ - CH ₃ |
| Alkynyl | Ak with no double bonds, one (1) or more triple bonds, and two (2) or more carbons, Ak < EC (2-) C, BD (1-) T (0) D> An unsaturated monovalent radical chain of two (2) or more carbons, branched or linear, containing one or more carbon-to-carbon triple bonds, but no double bonds Formed by the removal of one hydrogen from the corresponding alkyne, e.g., $\text{HC}\equiv\text{C}-\text{---} \quad , \quad \text{HC}\equiv\text{C}-\text{CH}_2-\text{---}$ |
| Lower | Any of the “alk” terms may be preceded by the term “lower” which implies a total carbon count one (1) to six (6) carbons for alkanes; 2-6 carbons for alkenes and alkynes. This term is used only when no carbon count is given in the patent |
| Loweralkyl | Ak with all bonds single exact and one (1) to six (6) carbons |
| Perhaloalkyl | Alkyl with all hydrogens replaced by halogen atoms |

Cb / Hy -based

Cb is any monocyclic or polycyclic group containing all carbon atoms with any bond values between atoms.

Hy is any monocyclic or polycyclic group containing one or more non-carbon atoms with any bond values between the atoms.

| | |
|---------------------|--|
| Aryl | <p>Cb with one or more aromatic rings, six (6) or more normalized bonds, and one (1) or more six (6)-membered rings</p> |
| Arylene | <p>- aryl -</p> <p>A divalent aromatic radical</p> <p>Formed by the removal of two (2) hydrogens from two different carbon atoms on the aromatic molecule, e.g., phenylene</p> |
| Cycloalkenyl | <p>Cb with one (1) or more double bonds and no triple bonds</p> <p>An unsaturated monovalent monocyclic or polycyclic radical containing one (1) or more carbon-to-carbon double bonds</p> <p>Formed by the removal of one (1) hydrogen from the corresponding cycloalkene, e.g., cyclopentadienyl, cyclohexenyl</p> |
| Cycloalkyl | <p>Cb with all bonds single exact</p> <p>A saturated monovalent alicyclic radical</p> <p>Formed by the removal of one (1) hydrogen from the corresponding cycloalkane, e.g., cyclopropyl, decahydronaphthyl</p> |
| Heteroaryl | <p>Hy with one (1) or more aromatic rings with six (6) or more normalized bonds and one (1) or more six (6)-membered rings</p> <p>Hy with one (1) or more aromatic rings and two (2) or more double bonds and one (1) or more five (5)-membered rings</p> <p>A monovalent radical derived from an aromatic molecule that contains at least one (1) heteroatom.</p> <p>Formed by removal of one (1) hydrogen from the pyridyl, benzopyranyl</p> |

Hybrids

Acyl Carbonyl bonded to Ak (which may have R's on it), bonded at the carbonyl carbon

Carbonyl bonded to R, bonded at the carbonyl carbon

Formyl

Aralkyl Alkyl bonded to one-to-three (1-3) aryls, bonded at the alkyl

Hydrocarbyl Ak

Cb

Ak bonded to Cb, bonded at Ak

Cb bonded to Cb, bonded to one (1) Cb

Cb bonded to Ak