

# Chemical information in the new world of patents

Chris McCue  
Vice President, Marketing  
ACS Division of Chemistry and the Law  
Spring 2011 ACS Meeting

## What I will discuss today

- **Challenges in finding chemistry in patents**
- **CAS' complete coverage of disclosed research**
- **Expert indexing increases discoverability**
- **Exhaustive coverage of substances from patents**

# CAS supports the mission of the ACS

## ACS Mission

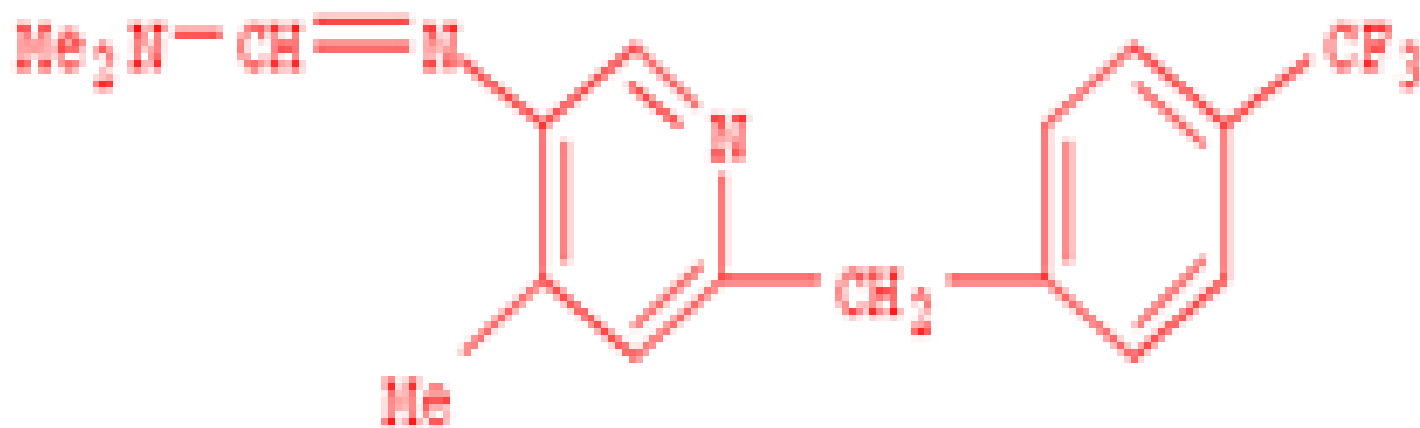
To advance the broader chemistry enterprise and its practitioners for the benefit of Earth and its people.

## CAS Mission

To provide the world's best digital research environment to search, retrieve, analyze, and link chemical information.



## How in the world would you find this substance?



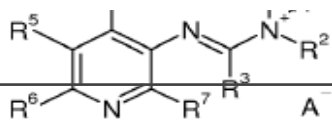
# How in the world would you find this substance?

## 1. Preparation of heteroarylamidines as fungicides useful in controlling microorganism

No Inventor data available

From Eur. Pat. Appl. (2010), EP 2264011 A1 20101222. Language: German, Database: CAPLUS

The invention relates to heteroarylamidines of formula I as fungicides, their prepn. and useful in controlling unwanted microorganisms by application of compds. of formula I to the microorganism habitat. Compds. of formula I wherein R1 and R2 are independently (un)substituted C1-12 alkyl, C2-12 alkenyl, C2-12 alkynyl, C3-8 cycloalkyl, etc; R1R2 taken together to form (un)substituted heterocycle; R3 is H and (un)substituted C1-12 (mono/di)alkylamine; R4, R5, R6 and R7 are independently H, CN, SH and derivs., etc.; are claimed. Example compd. II was prepd. by cross-coupling of 3-amino-6-bromopyridine with 2-phenoxyphenylboronic acid followed by condensation with N,N-ethyl(methyl) formamide di-Me acetal. All the invention compds. were evaluated for their fungicidal activity (some data given).



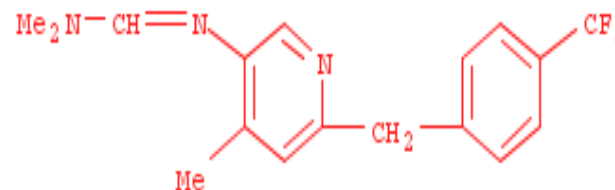
In den Fällen, in denen die Verbindungen (I) als Salze einer Säure HA vorliegen, ist die Säure HA in der Spalte HA von Tabelle 1 aufgelistet.

EP 2264 011 A1

**CAS Registry Number:** 1258551-34-6

C<sub>17</sub> H<sub>18</sub> F<sub>3</sub> N<sub>3</sub>

Methanimidamide, N,N-dimethyl-N'-[4-methyl-6-[[4-(trifluoromethyl)phenyl]methyl]-3-pyridinyl]-



# CAS databases provide the world's most complete and authoritative coverage of chemistry and related science

## Substances

>57M organic and inorganic substances

Updated daily (~12K)

Substances reported back to 1802

2.8 billion experimental and predicted properties

## Reactions

>29M single and multi-step reactions

>13M synthetic preparations

Updated weekly (30K-50K)

Reactions back to 1840

## CAS Databases

## Literature

>33M patent, journal, dissertations, etc.

>10K scientific journals covered

Patents from 61 patent offices

Updated daily (~3K)

Nearly 300M citations back to 1997

Coverage back to early 1800s

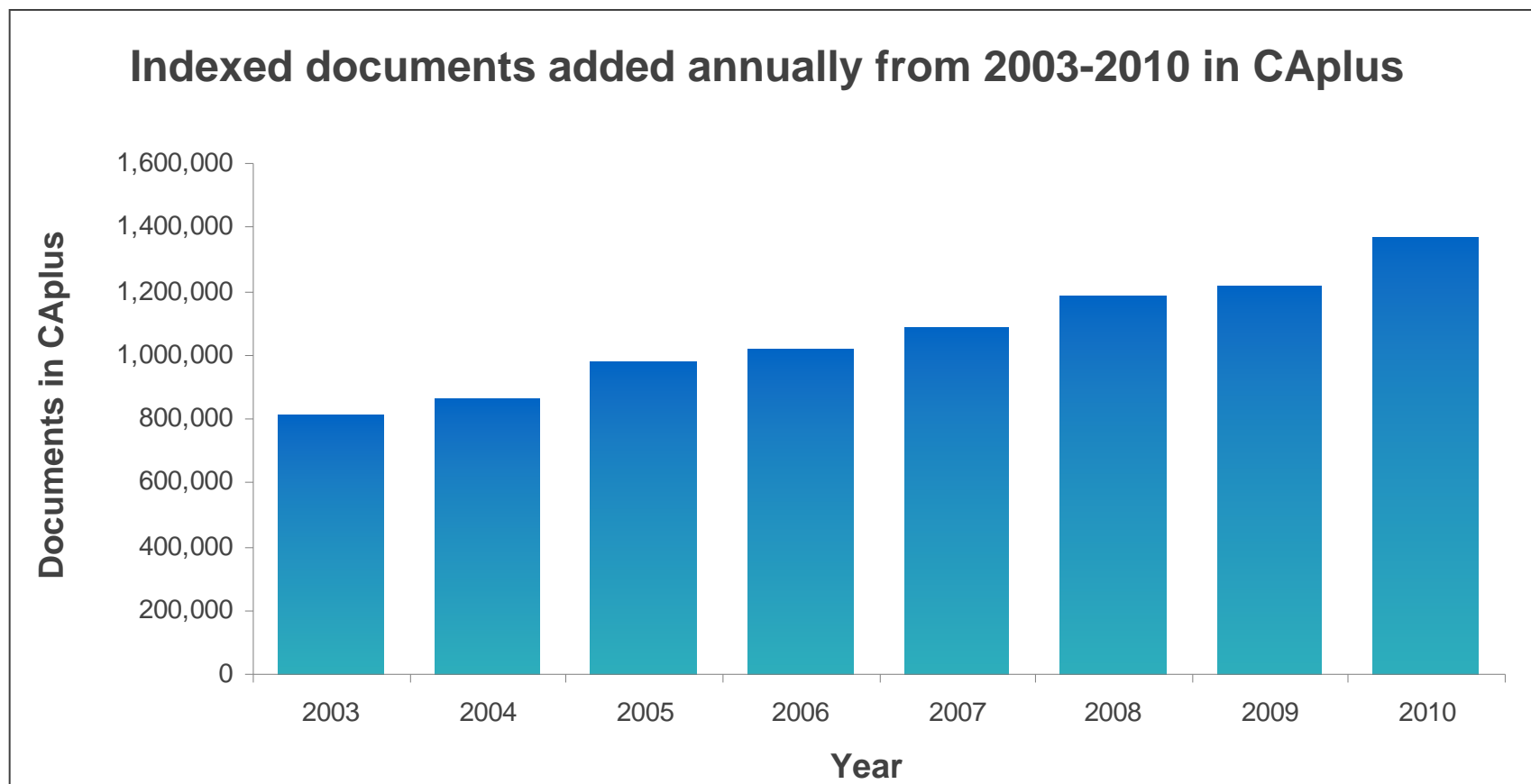
## Markush

Markush structures representing organic and organometallic substances in patents

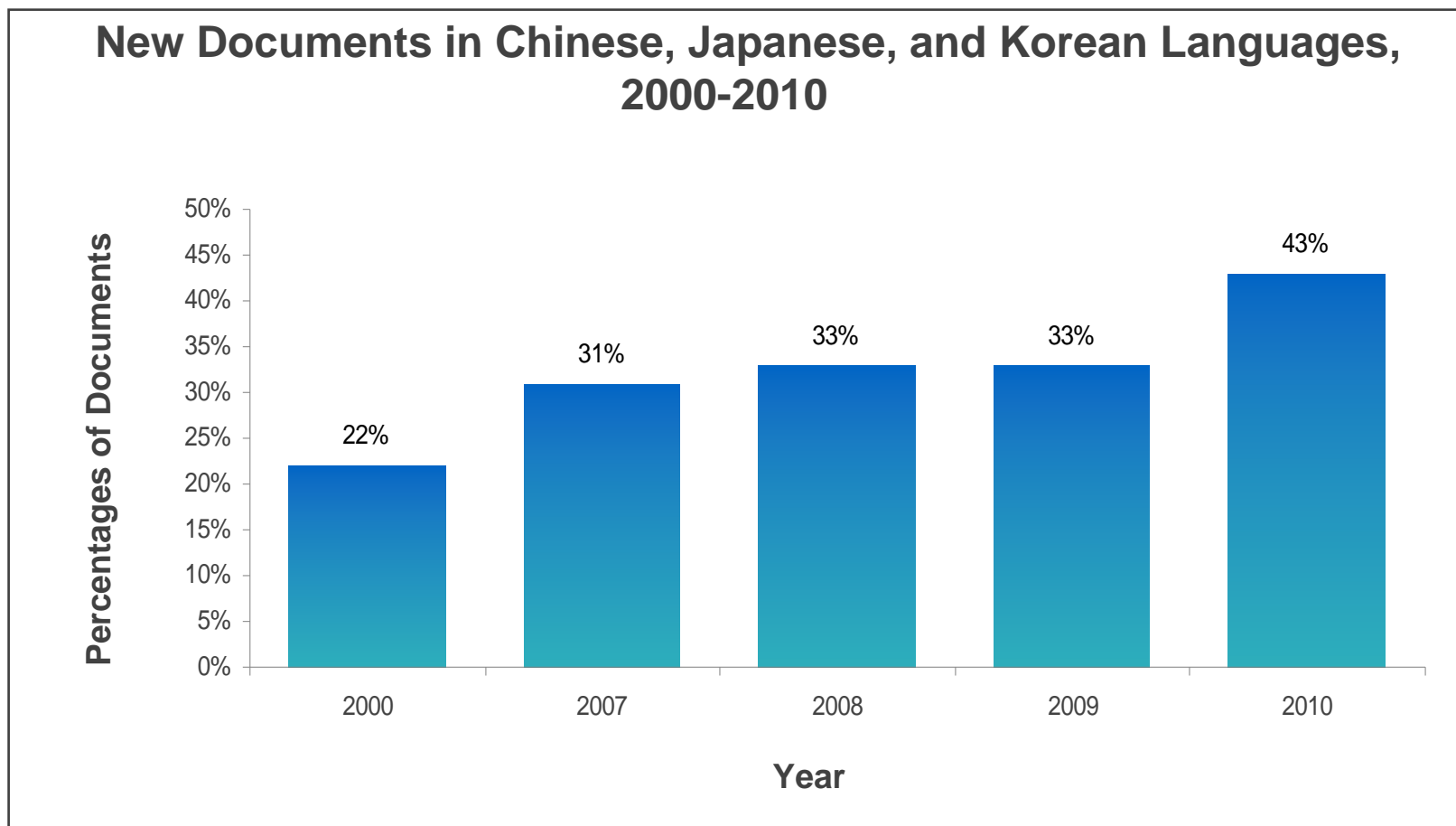
Coverage back to 1961

Updated daily (~150-200 Markush structures)

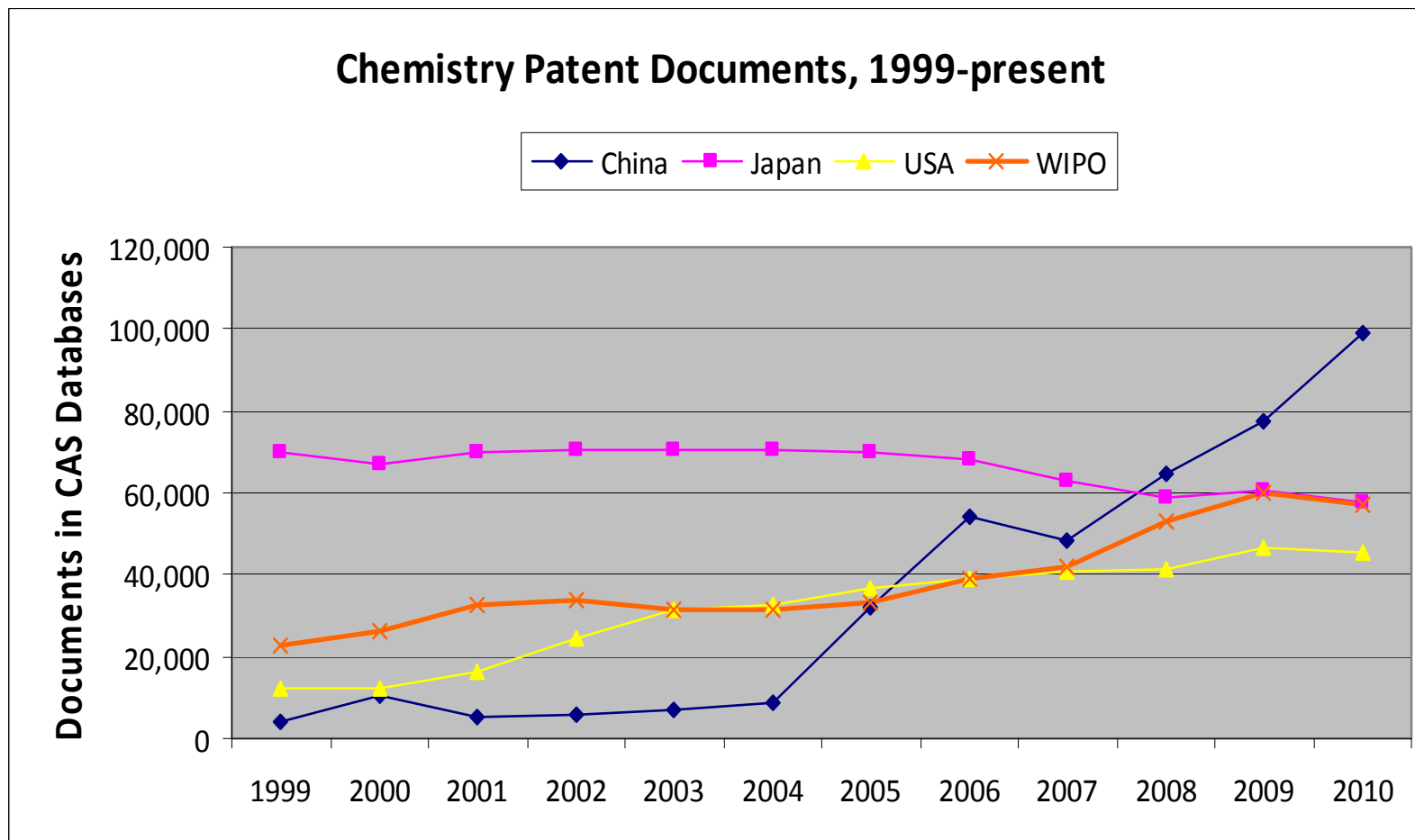
## Growth in published chemical literature has been strong in the last decade



# Chinese, Japanese, and Korean language documents account for 43% of new CPlus database records in 2010



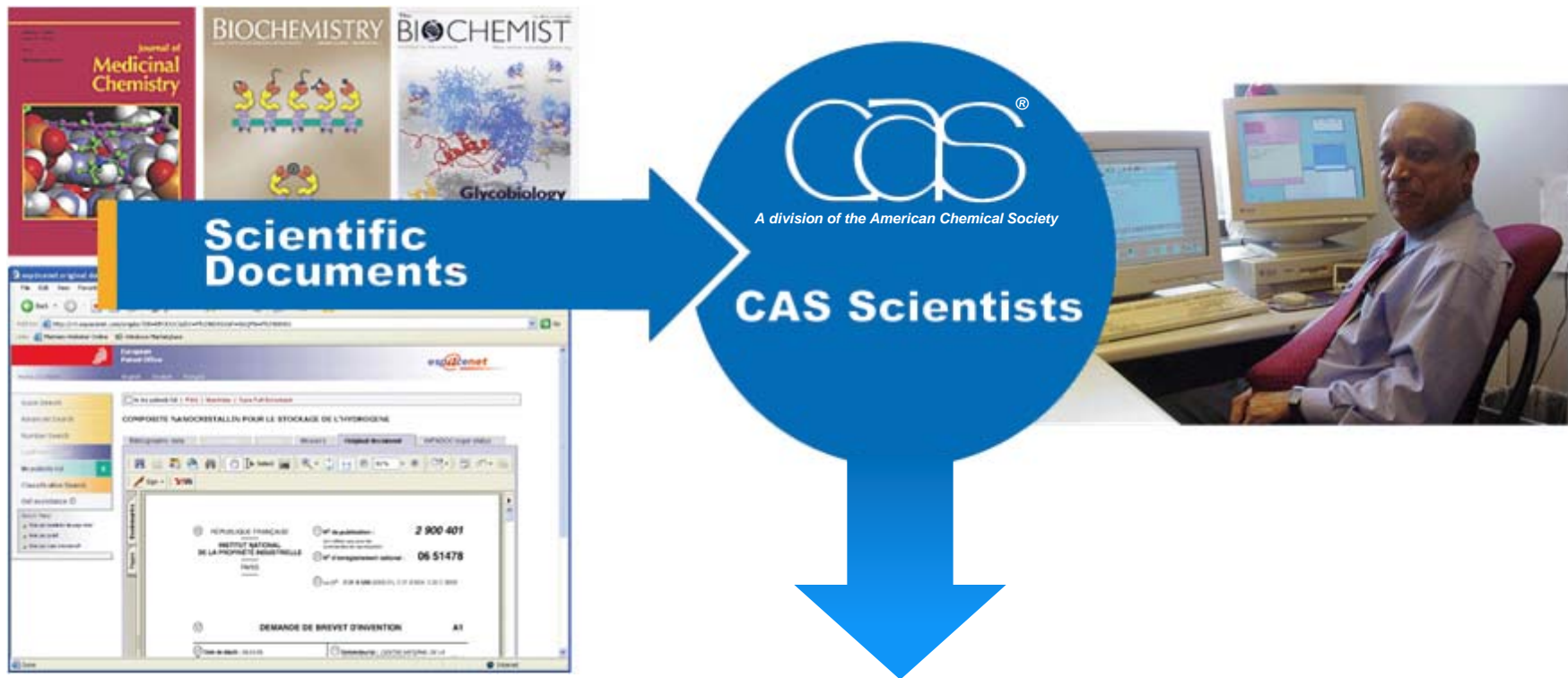
# China has established a clear lead in the volume of published patent applications on an annual basis



## CAS provides the most timely coverage of important patent authorities

- **Bibliographic information and abstracts from original documents are added quickly**
  - Within 2 days for 9 major patent authorities:  
US, WO, EP, CA, DE, FR, GB, JP, RU
  - No more than 14 days for Chinese and Korean patent offices (typically 5 days)
  - Within 7 days for 1,500 “core” journals
  - Machine translations added for some countries
- **Daily updates provide additional information**
  - Intellectually enhanced titles and abstracts
  - Subject and substance indexing added within 27 days for major authorities

# CAS chemists and global partners perform the indexing that adds value to CAS database records



## Highest-quality indexing

## CAS indexing is maintained in a hierarchical structure that enhances retrieval

CA Index Terms	Chemical Compound Classes	Taxonomic Vocabulary
<ul style="list-style-type: none"><li>• &gt; 16,200 subject headings (index terms)</li><li>• &gt; 78,900 subject headings with linked terms</li><li>• Related concepts</li><li>• Date ranges of use for each term</li></ul>	<ul style="list-style-type: none"><li>• &gt; 11,280 frequently indexed substances</li><li>• Frequently cited substances from the past year</li><li>• Broader terms for each substance</li></ul>	<ul style="list-style-type: none"><li>• &gt; 244,400 controlled taxonomic terms</li><li>• Relationships between taxonomy and broader or non-taxonomic vocabulary</li></ul>

# CAS indexing provides historical and synonymous terms

```
=> E FIREPROOF MATERIALS+ALL/CT
E1      76      --> Fireproof materials/CT
                HNTE Valid heading during volumes 1-20 (1907-1926) only.
E2      31958   NEW  Fire-resistant materials/CT
***** END *****
```

```
=> E E2+ALL
E1      28139   BT1  Materials/CT
E2      31958   --> Fire-resistant materials/CT
                HNTE Valid heading during volume 21 (1927) to present.
E3      1249    OLD  Fire-resistant materials, Flame-retardant
                materials/CT
E4      76      OLD  Fireproof materials/CT
E5      UF      UF   Fire-resistant material/CT
E6      UF      UF   Fire-resisting materials/CT
E7      UF      UF   Fire-retardant materials/CT
E8      UF      UF   Flame retardants/CT
E9      UF      UF   Flame-resistant materials/CT
E10     UF      UF   Flame-retardant material/CT
E11     UF      UF   Flame-retardant materials/CT
E12     10725   RT   Fireproofing/CT
E13     36617   RT   Fireproofing agents/CT
E14     40460   RT   Heat-resistant materials/CT
E15     37868   RT   Refractories/CT
E16     RTCS   RTCS  Decabromodiphenyl ether/CT
E17     RTCS   RTCS  Phenyl dichlorophosphate/CT
E18     RTCS   RTCS  Zinc borate/CT
***** END *****
```

All Used For (UF) terms are accommodated by using the CA index term *Fire-resistant materials*

## Detailed indexing has also been added to more than 50,000 World Traditional Medicine patents

IT Analgesics  
 Ascites  
 Aspongopus  
 Astragalus  
 Atractylodes macrocephala  
 Aucklandia  
 Bupleurum  
 Cirrhosis  
 Edema  
 Epimedium  
 Fallopia japonica  
 Hepatitis  
 Ipomoea  
 Isatis  
**Ligusticum chuanxiong**  
 Liver disease  
 Manis  
 Natural products, pharmaceutical

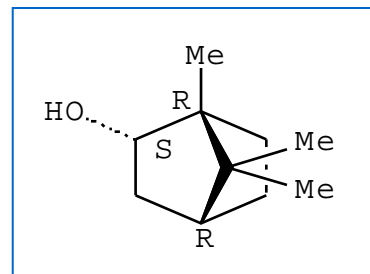
•••  
 IT Abdominal disease  
 (abdominal distention)  
 IT Disease, animal  
 (asthenia)  
 IT Medical goods  
 (bags)  
 IT Viral hepatitis  
 (chronic)

•••  
 IT **507-70-0**  
 RL: BSU (Biological study  
 unclassified); BIOL (Biological study)

CA Indexing includes:

- Taxonomic terms
- Disease terms
- Substance class terms
- Specific substances

*Ligusticum chuanxiong* is one of the 50 fundamental herbs in traditional Chinese medicine.



*Borneol* (CAS Registry Number 507-70-0) aids in Chuanxiong absorption.

# Patent applications are particularly rich sources of substance information

Dr. Mark Willi from CAS completes substance indexing 15 days after WIPO publishes the patent



## CAS chemist analysis of a single PCT application

- WO 2009015208
- 917 indexed compounds from Examples and Claims
- 576 new compounds added to CAS REGISTRY
- 613 single-step reactions
- 5,394 reactions with multiple steps
- 1,029 reaction participants
- Markush structure with 2,119 substituent definitions

# Indexing for major patent office documents is added quickly to the CAS databases

  
 US 20110054170A1

(19) **United States**  
 (12) **Patent Application Publication** (10) Pub. No.: US 2011/0054170 A1  
 McQuade et al. (43) Pub. Date: **Mar. 3, 2011**

(54) SIX-MEMBERED N-HETEROCYCLIC CARBENE-BASED CATALYSTS FOR ASYMMETRIC REACTIONS

(75) Inventors: D. Tyler McQuade (US); Jin Kyoong Park (US); Matthew D. Lackey, Hickory, NC (US); Tallahassee, FL (US)

(73) Assignee: FLORIDA STATE UNIVERSITY RESEARCH FOUNDATION, TALLAHASSEE, FL (US)

(21) Appl. No.: 12/870,901

(22) Filed: Aug. 30, 2010

Related U.S. Application Data

(60) Provisional application No. 61/238,531, 2009.

ED Entered STN: 04 Mar 2011  
 AN 2011:271713 CAPLUS Full-text  
 DN 154:259821  
 TI Six-membered n-heterocyclic carbene-based catalysts for asymmetric reactions  
 IN McQuade, D. Tyler; Park, Jin Kyoong; Rexford, Matthew D.; Lackey, Hershel H.  
 PA Florida State University Research Foundation, USA  
 SO U.S. Pat. Appl. Publ., 22pp.  
 CODEN: USXXCO

DT Patent  
 LA English

FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 20110054170	A1	20110303	US 2010-870901	20100830
PRAI	US 2009-23836				

ASSIGNMENT HISTORY F

● ● ●

IT 3943-97-3P

1185248-86-5P

1252913-49-2P

**1266612-72-9P**

RL: SPN (Synthesis)

(preparation of)

catalysts for

● ● ●

RN 1266612-72-9 REGISTRY

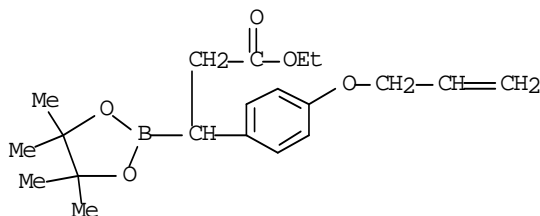
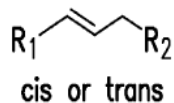
ED Entered STN: 07 Mar 2011

CN 1,3,2-Dioxaborolane-2-propanoic acid, 4,4,5,5-tetramethyl-  
 .beta.-[4-(2-propen-1-  
 yloxy)phenyl]-, ethyl ester (CA INDEX NAME)

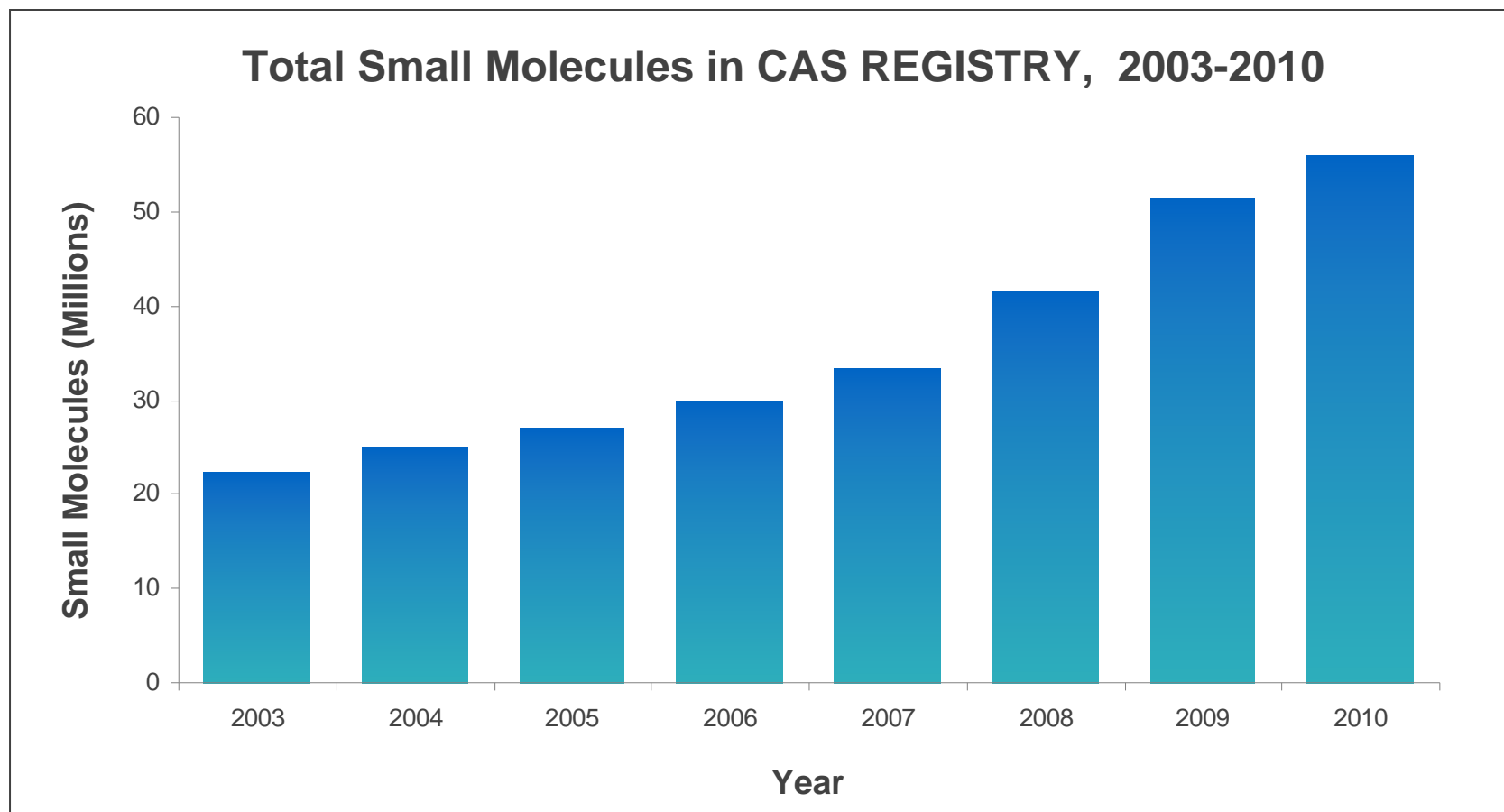
MF C20 H29 B O5

SR CA

LC STN Files: CA, CAPLUS, USPATFULL



# CAS REGISTRY is the complete source for chemical substance information



# In 2010, patents contributed 46% of all new small molecule registrations

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
29 January 2009 (29.01.2009)

PCT

(10) International Publication Number  
**WO 2009/015208 A1**

(51) International Patent Classification:  
A01N 43/78 (2006.01) A61K 31/425 (2006.01)

(74) Agents: SWANSON & BRATSCHEUN, L.L.C. et al.;  
8210 SouthPark Terrace, Littleton, CO 80120 (US).

(21) International Application Number:  
PCT/US2008/070893

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(22) International Filing Date: 23 July 2008 (23.07.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/961,634 23 July 2007 (23.07.2007) US  
61/022,725 22 January 2008 (22.01.2008) US

(71) Applicant (for all designated States except US): REPLI-DYNE, INC. [US/US]; 1450 Infinite Drive, Louisville, CO 80027 (US).

(72) Inventors; and

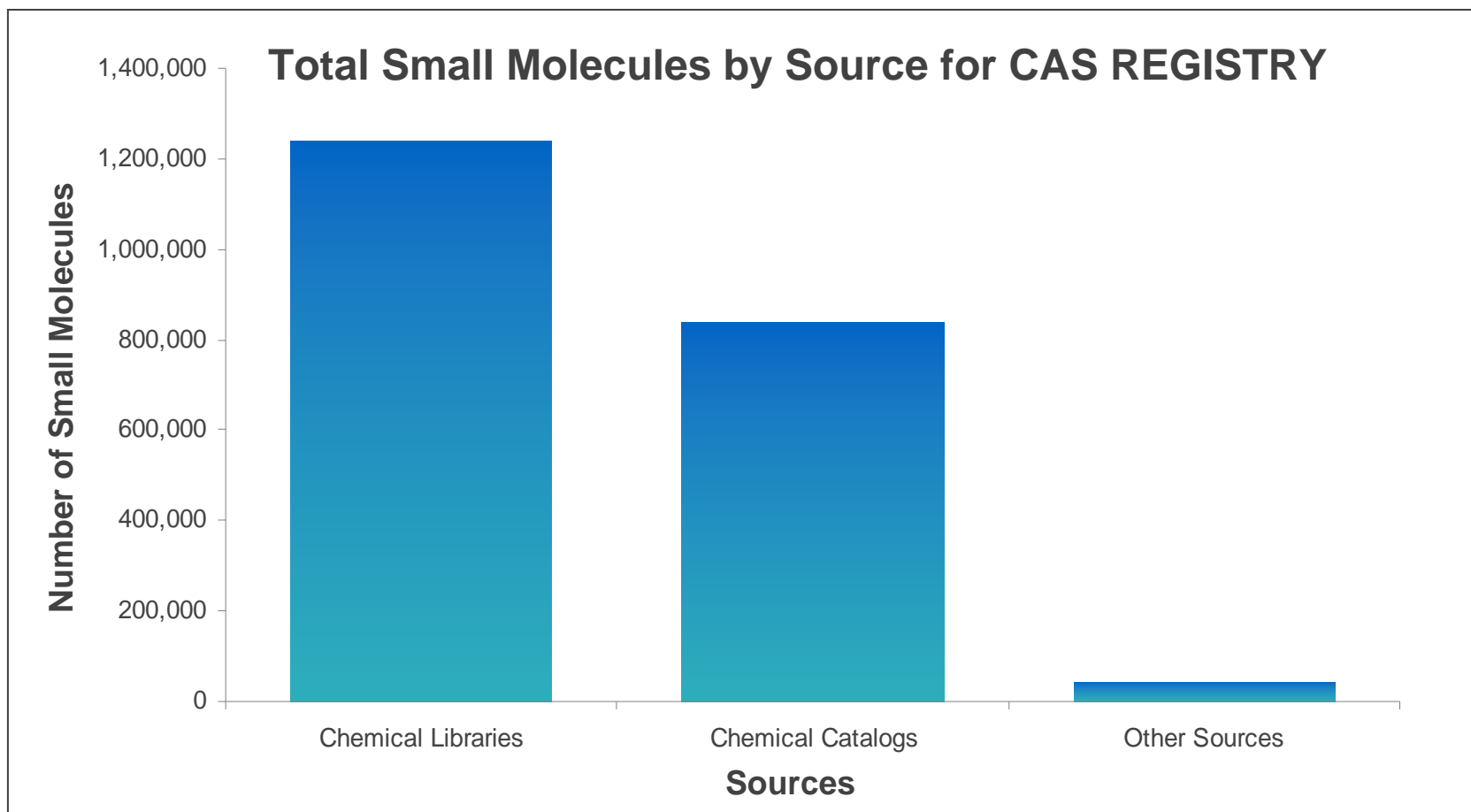
(75) Inventors/Applicants (for US only): GUILLES, Joseph [US/US]; 714 Skywalker Place, Lafayette, CO 80027 (US). JARVIS, Thale, Cross [US/US]; 2265 Dartmouth Ave., Boulder, CO 80305 (US). STRONG, Sarah [US/US]; 107 Lois SunDrive, Louisville, CO 80027 (US). SUN, Xicheng [US/US]; 1404 Kahler Place, Brown Field, CO 80027 (US). QIU, Jian [US/US]; 923 Grays Peak Drive, Superior, CO 80027 (US). ROHLOFF, John, C. [US/US]; 605 Meadowbrook Drive, Boulder, CO 80303 (US).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

## CAS finds novel compounds in many reputable sources

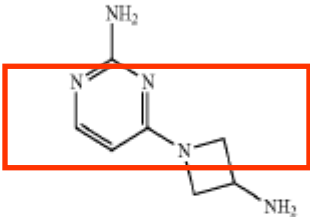
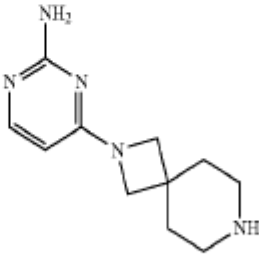
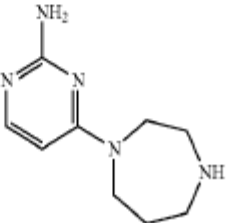




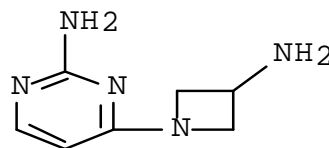
# CA indexing also identifies prophetic substances from patent examples

TABLE 2-continued

Examples 26-51.

Starting diamines Example (HNR <sup>1</sup> R <sup>2</sup> ) Number and CAS#	Product Structures	May be Prepared by the Method of (substituting the indicated starting diamine)
31 tert-butyl azetidin-3-ylcarbamate CAS # 91188-13-5		Example 7A followed by Example 7B
32 tert-butyl 2,7-diazaspiro[3.5]nonane-7-carboxylate CAS # 236406-55-6		Example 7A followed by Example 7B
33 tert-butyl 1,4-diazepane-1-carboxylate CAS # 112275-50-0		Example 7A followed by Example 7B

L1 ANSWER 21 OF 75 REGISTRY COPYRIGHT 2011 ACS on STN  
 RN 1259941-53-1 REGISTRY  
 ED Entered STN: 19 Jan 2011  
 CN 2-Pyrimidinamine, 4-(3-amino-1-azetidiny)-  
 (CA INDEX NAME)  
 MF C7 H11 N5  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER



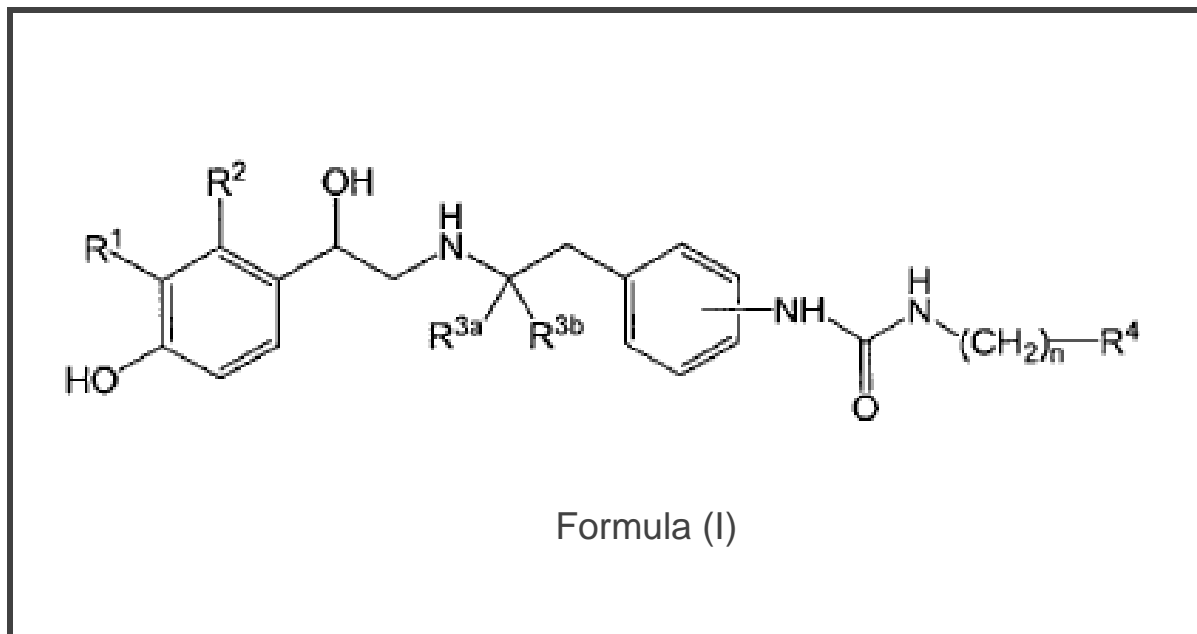
\*\*PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT\*\*

1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

## CAS covers Markush structures in patents

- **Markush structures are generic representations of substances in claims examples, disclosures, and descriptions in patents**
- **Coverage includes**
  - Organic, organometallic, and oligomeric substances
  - All languages
  - ~350,000 patents
  - ~800,000 Markush structures
- **~2.5 Markush structures per patent**

Markush structures are typically identified by “Selected from the group consisting of...”

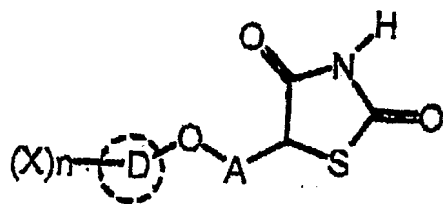


# CAS chemists interpret rich “language” expressions found in Markush structures in patents

## Patent

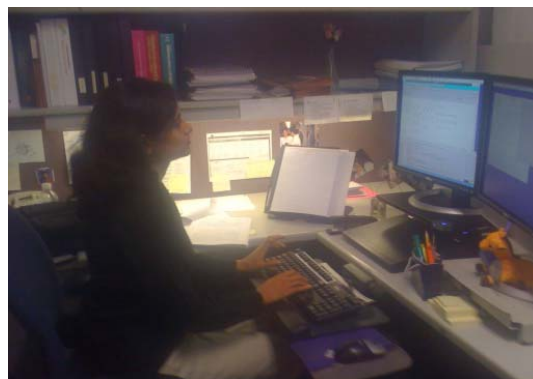
Claims:

1.

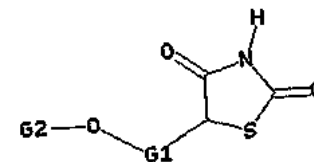


In which:

A represents an alkylene group that may be substituted by up to three halogen atoms or is alkenylene substituted by 1-2 hydroxy groups; D represents optionally substituted aryl or heteroaryl group such as phenyl; X represents halo, alkyl, or alkoxy, and n = 0-5; or salts thereof



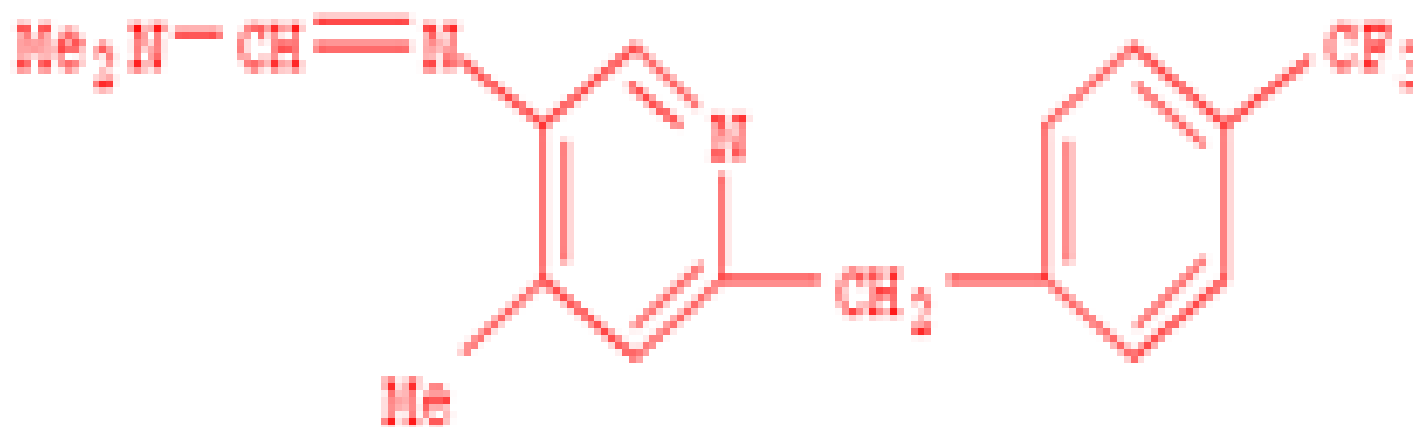
## MARPAT indexing



G1 = alkylene (opt. substd. by up to 3 halo)/alkenylene (substd. by 1-2 OH)

G2 = aryl (opt. substd. by up to 5 G3)/heteroaryl (opt. substd. by up to 5 G3)/Ph (opt. substd. by up to 5 G3)

Now you know how you can find this substance:



## Summary

- **CAS scientists find obscured substances in patents**
- **Patents are an increasingly bountiful source of novel chemistry**
- **Rapid processing of patent information is a CAS priority**
- **CAS databases reflect the changing face of chemistry**
- **CAS keeps current on chemical information trends**

## Interested in learning more about CAS at this conference?

Please consider attending these CAS sessions on Thursday:

- **“Key sources for pharmaceutical and chemical literature and patent searching”**
  - Presenter: Elaine Cheeseman, Ph.D., Science IP
  - Thursday, March 31, 1:30 p.m., Room 213 D, Anaheim Convention Center
- **“What patent searchers need to know from patent attorneys”**
  - Presenter: Mike Mullican, Ph.D., Science IP
  - Thursday, March 31, 2:00 p.m., Room 213 D, Anaheim Convention Center

***Thank You.***

