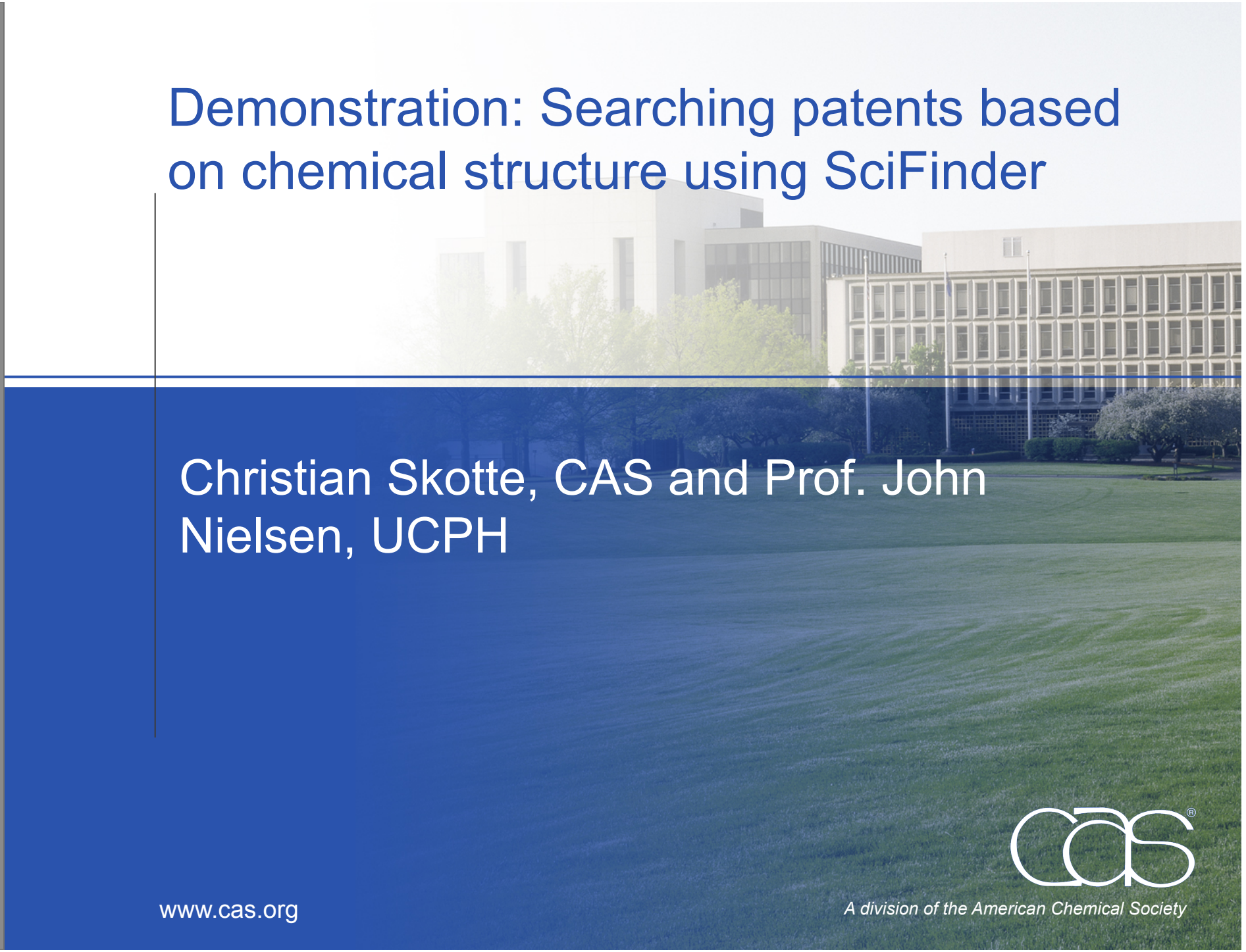
Demonstration: Searching patents based on chemical structure using SciFinder

Christian Skotte, CAS and Prof. John Nielsen, UCPH

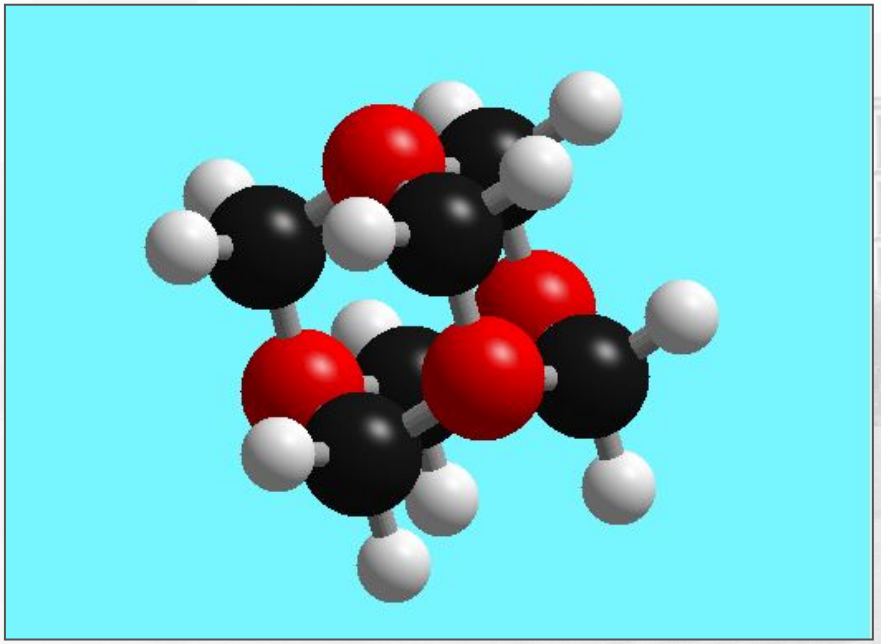


www.cas.org

A division of the American Chemical Society

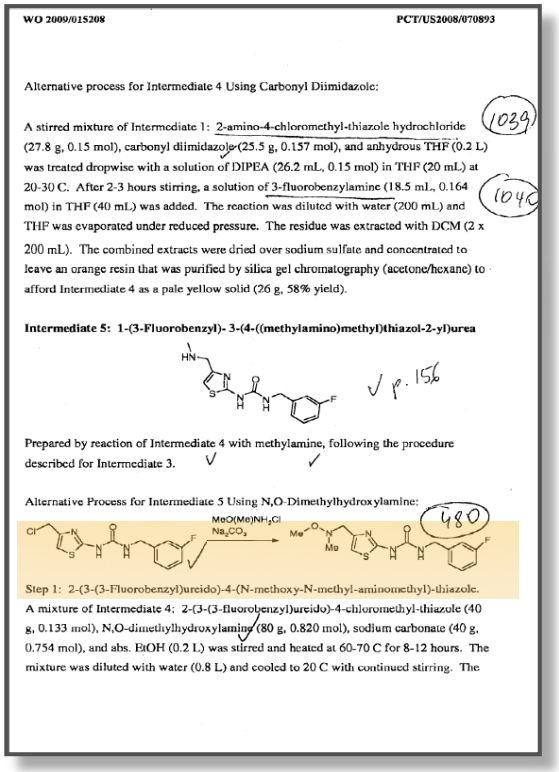
What criteria must a substance meet to be included in the CAS REGISTRY?

- Identified by CAS as coming from a reputable source, including but not limited to patents (63 patent offices), journals (9000 journals) , chemical catalogs, and web-based substance collections
- Described in largely unambiguous terms
- Characterized by physical methods
- Described in a patent document example or claim
- Consistent with the laws of atomic covalent organization
- Generic chemistry of small molecules are separately added to the MARPAT database



For complex chemistry, CAS chemists classify substance information and verify graphical processes and structures

1. Review reaction and structure



WO 2009/015208 PCT/US2008/070893

Alternative process for Intermediate 4 Using Carbonyl Diimidazole:

A stirred mixture of Intermediate 1: 2-amino-4-chloromethyl-thiazole hydrochloride (27.8 g, 0.15 mol), carbonyl diimidazole (25.5 g, 0.157 mol), and anhydrous THF (0.2 L) was treated dropwise with a solution of DIPEA (26.2 mL, 0.15 mol) in THF (20 mL) at 20-30 C. After 2-3 hours stirring, a solution of 3-fluorobenzylamine (18.5 mL, 0.164 mol) in THF (40 mL) was added. The reaction was diluted with water (200 mL) and THF was evaporated under reduced pressure. The residue was extracted with DCM (2 x 200 mL). The combined extracts were dried over sodium sulfate and concentrated to leave an orange resin that was purified by silica gel chromatography (acetone/hexane) to afford Intermediate 4 as a pale yellow solid (26 g, 58% yield).

Intermediate 5: 1-(3-Fluorobenzyl)- 3-(4-((methylamino)methyl)thiazol-2-yl)urea

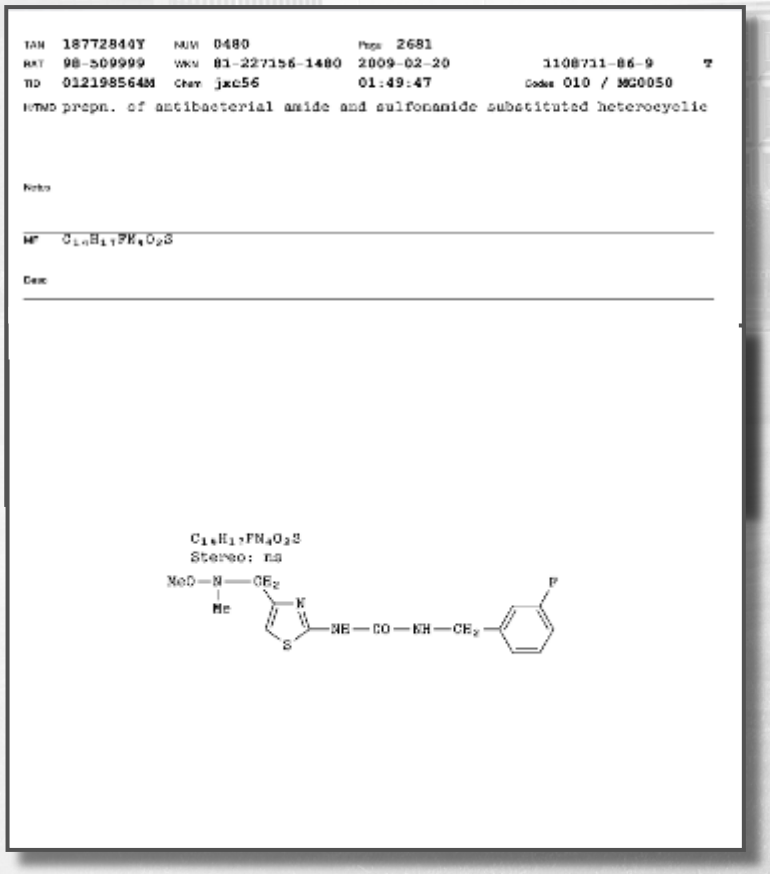
Prepared by reaction of Intermediate 4 with methylamine, following the procedure described for Intermediate 3.

Alternative Process for Intermediate 5 Using N,O-Dimethylhydroxylamine:

Step 1: 2-(3-(3-Fluorobenzyl)ureido)-4-(N-methoxy-N-methyl-aminomethyl)-thiazole.

A mixture of Intermediate 4: 2-(3-(3-fluorobenzyl)ureido)-4-chloromethyl-thiazole (40 g, 0.133 mol), N,O-dimethylhydroxylamine (80 g, 0.820 mol), sodium carbonate (40 g, 0.754 mol), and abs. EtOH (0.2 L) was stirred and heated at 60-70 C for 8-12 hours. The mixture was diluted with water (0.8 L) and cooled to 20 C with continued stirring. The

2. Create registration record



TAN 18772844Y NUM 0480 Page 2681
RXT 98-509999 WKN 81-227156-1480 2009-02-20 1108711-86-9 T
TD 0121985648M Chem jxc56 01:49:47 Codes 010 / MG0050
RTND prepn. of antibacterial amide and sulfonamide substituted heterocyclic

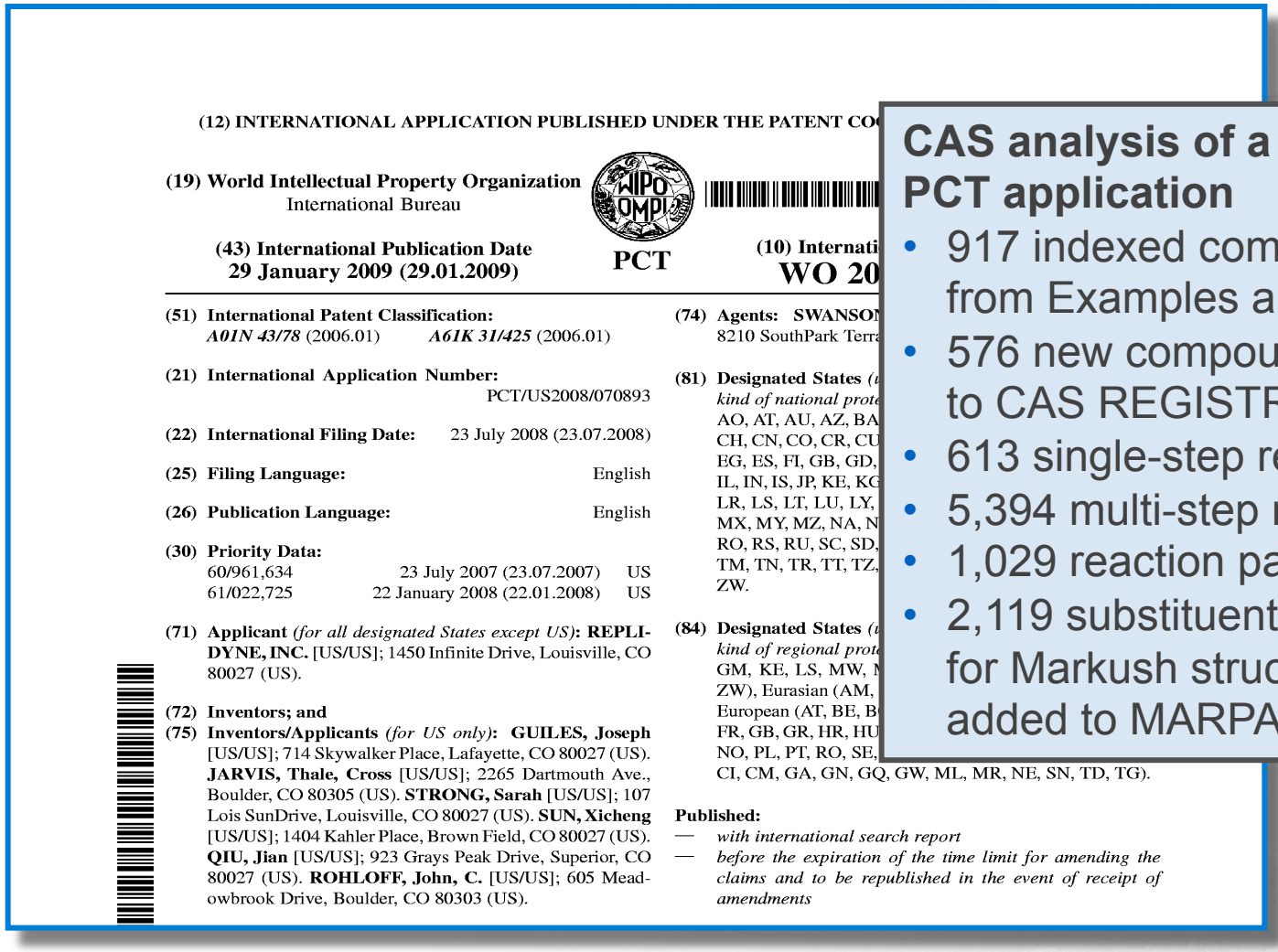
Notes

MF $C_{14}H_{17}FN_4O_2S$

Desc

$C_{14}H_{17}FN_4O_2S$
Stereo: ns

Since 1997, patents have provided more new small molecules than have journals



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT CO

(19) World Intellectual Property Organization
International Bureau

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

PCT

(10) Internati
WO 20

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

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

(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)*: **REPLIDYNE, INC.** [US/US]; 1450 Infinite Drive, Louisville, CO 80027 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **GUILES, 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).

(74) Agents: SWANSO
8210 SouthPark Terra

(81) Designated States (
kind of national prote
AO, AT, AU, AZ, BA
CH, CN, CO, CR, CU
EG, ES, FI, GB, GD,
IL, IN, IS, JP, KE, KG
LR, LS, LT, LU, LY,
MX, MY, MZ, NA, N
RO, RS, RU, SC, SD,
TM, TN, TR, TT, TZ,
ZW.

(84) Designated States (
kind of regional prote
GM, KE, LS, MW, M
ZW), Eurasian (AM,
European (AT, BE, B
FR, GB, GR, HR, HU
NO, PL, PT, RO, SE,
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 analysis of a typical PCT application

- 917 indexed compounds from Examples and Claims
- 576 new compounds added to CAS REGISTRY
- 613 single-step reactions
- 5,394 multi-step reactions
- 1,029 reaction participants
- 2,119 substituent definitions for Markush structures added to MARPAT®

CAS indexing policies from patents

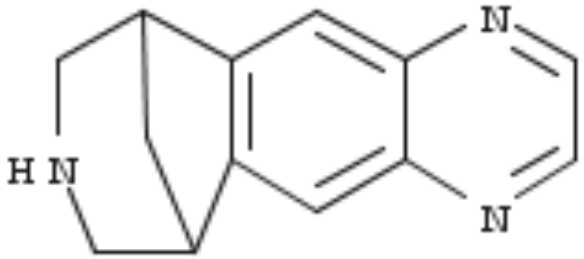
- **Primary indexing is from claims and examples**
- **Rarely, indexing is from other parts of patents**
- **All indexing must represent the novelty of the patent. Novelty is decided on basis of claims, author emphasis, and analyst judgment.**
- **All exemplified compounds with supporting data from 63 patent authorities.**
  - Indexing is based on novelty in addition to data justification. Data may be physical characteristics (melting point), spectral data, yield ranges, or bioassay data. The author of the patent may have a comprehensive bioassay statement to justify all examples. This statement may be anywhere in the patent.
- **Exemplified compounds (prophetic) without data from 9 major patent offices 1998-**
- **All claimed compounds**
- **Generic claims with more than a few combinations are indexed in the Marpat database and can be searched in SciFinder too**

Exercise: Searching patents based on **chemical structure** using **SciFinder**

Christian Skotte, CAS and Prof. John Nielsen, UCPH

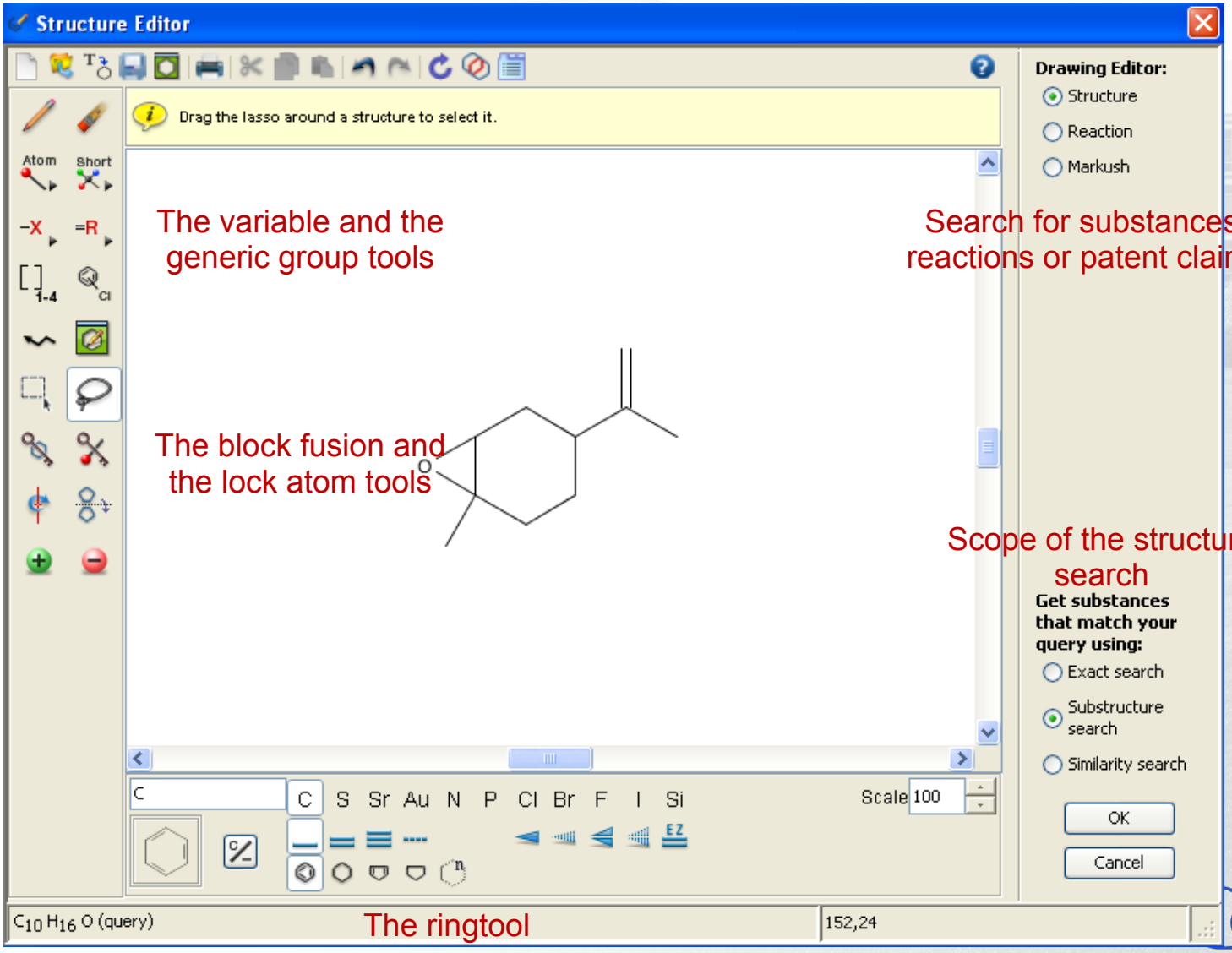
A division of the American Chemical Society

Search example I - Varenicline

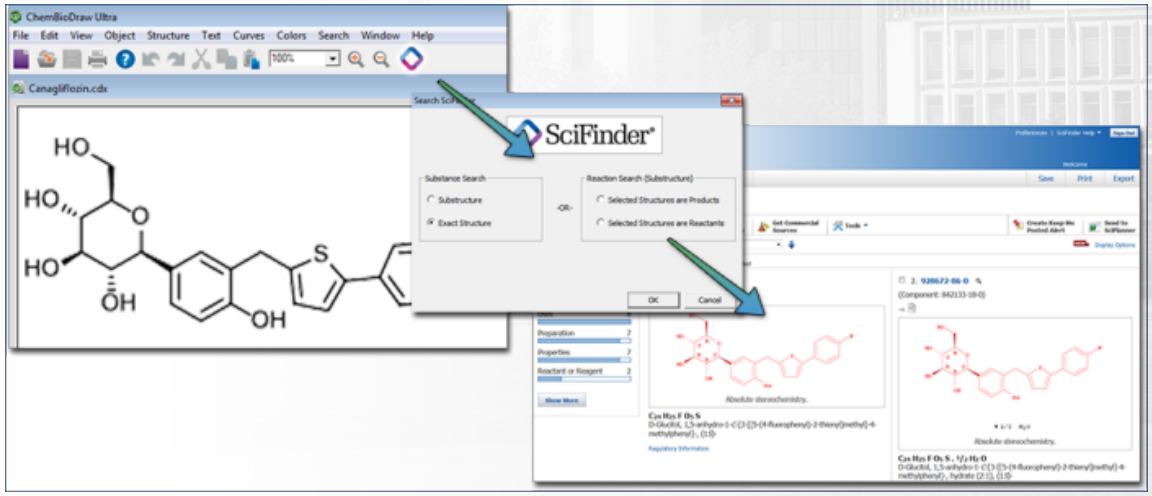


- **Search as a chemical name**
- **Search Varenicline as structure, structure exact**
- **What are the differences in answers in Registry and the corresponding references in Caplus?**

The structure drawing editor

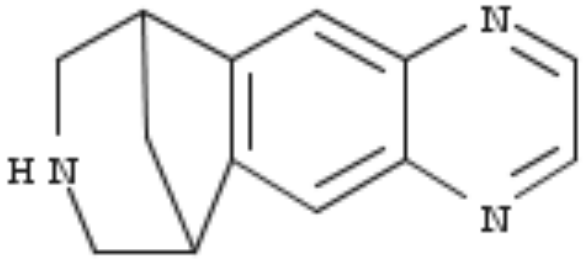


An alternative to using the structure drawing editor



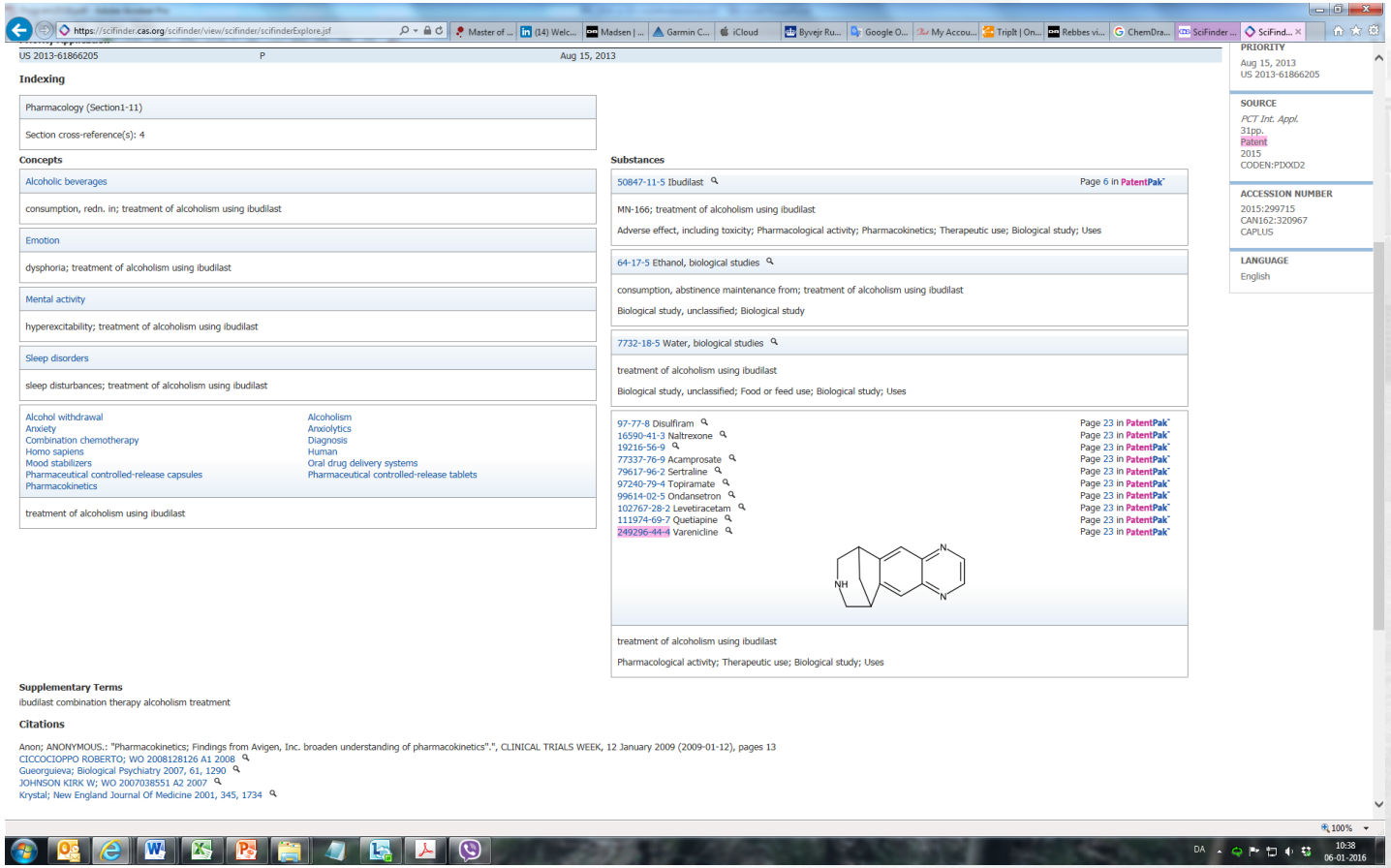
How many do have ChemDraw?

Search example I - Varenicline



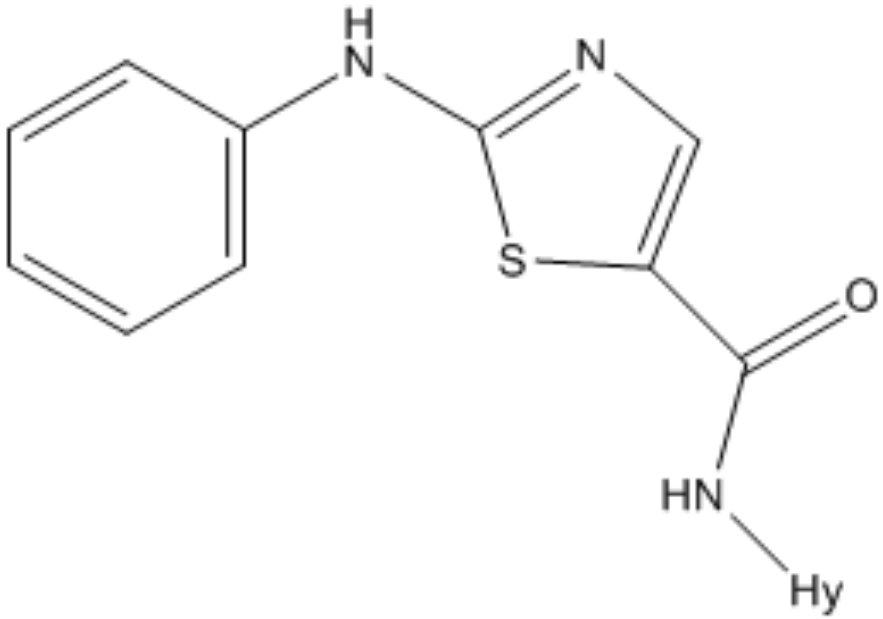
- Search as a chemical name
- Search Varenicline as structure, structure exact
- What are the differences in answers in Registry and the corresponding references in Caplus?

Search example I – Varenicline - PatentPak



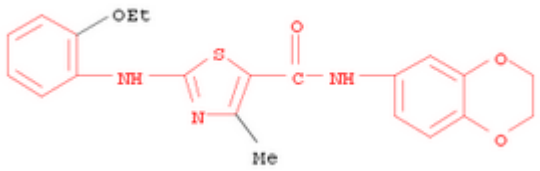
- **PatentPak, a workflow solution. Allows you to track down the actual chemistry in the patent. Provides you with a searchable pdf and a mark-up function!**

Search example II

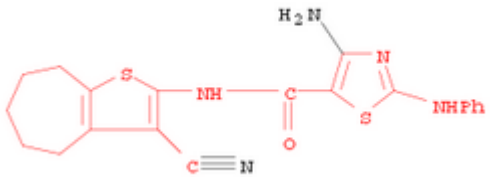


- Hy is the variable that retrieve all heterocycles
- Work with the lock atom tool. Toggle on and off on the Hy. What happens?
- How do you easily obtain answers where the generic heterocycle is substituted only once?

Answers from example II

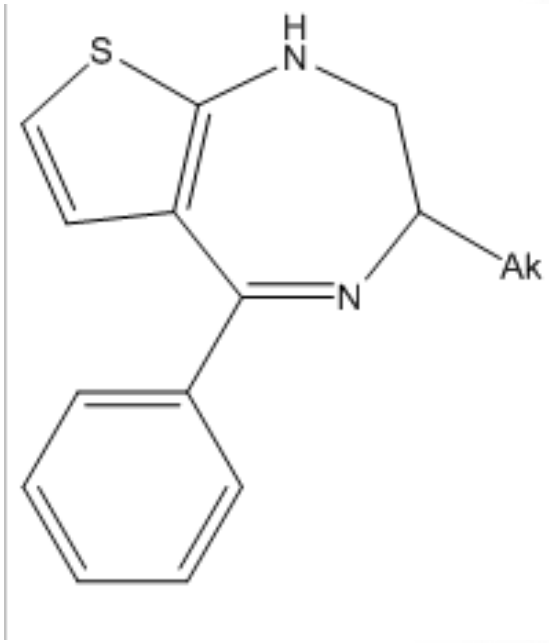


- Example of answers with closed Hy node



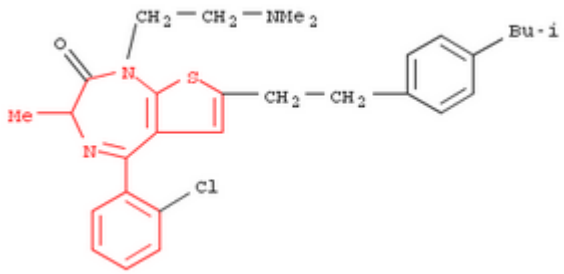
- Example of answers with exactly one substitution on the Hy

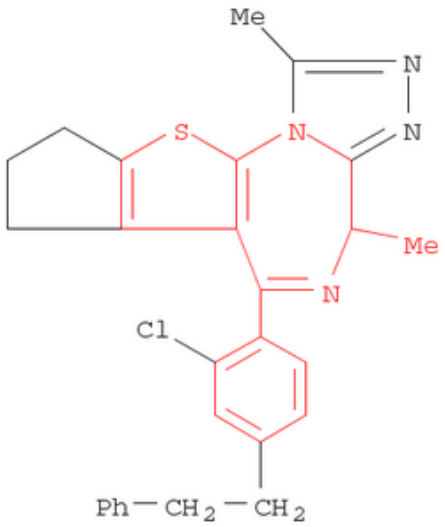
Search example III



- Use the ring tool to create the 7 membered ring
- The ”Ak” is a variable. These are found in the ”X”
- “Ak” will retrieve any chain. Branched, linear, saturated or unsaturated
- Use the ”lock for fusion tool” to retrieve only ring systems that looks like this one
- ….
- An answer set can be sorted by various parameters. ”Number of references”, ”Molecular weight” and others
- Molecules are stored in the Registry file. Patent references are stored in the CAplus file. Find corresponding patent references.

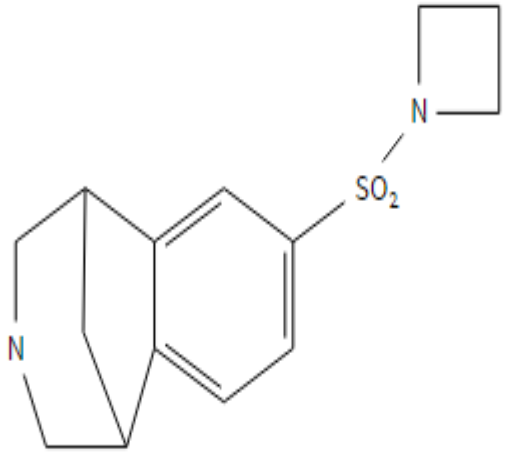
Answers for the example above





- **Substructure search with rings closed for fusion - 68**
- **Substructure search with rings opened for fusion - 398**

Search example IV:
Can you operate in the space of no actual hits?



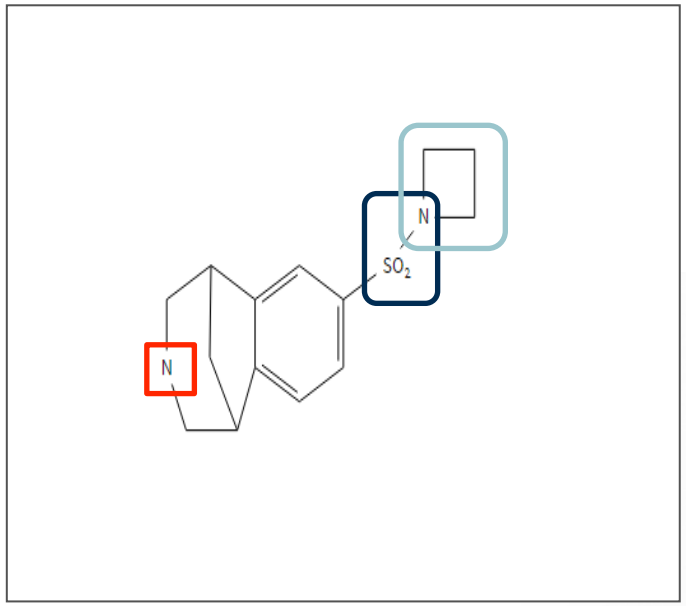
Markush matching and reassembling

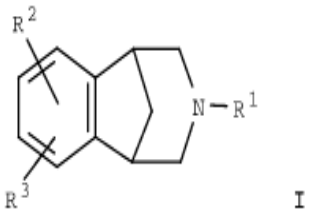
9. Preparation of aryl-fused azapolycyclic compounds as nicotine binding inhibitors

By: Brooks, Paige Roanne Palmer; Coe, Jotham Wadsworth

Assignee: Pfizer Products Inc., USA

The invention discloses the prepn. of aryl-fused azapolycyclic compds., such as I [R^1 = H, alkyl, unconjugated alkenyl, benzyl, $X(CO)R^{13}$, CH_2CH_2O-alkyl; R^2, R^3 = H, alkenyl, alkynyl, hydroxy, nitro, amino, halo, cyano, SO_qalkyl, (q = 0 - 2), alkylamino, CO_2R^4, $CONR^5R^6$, $SO_2NR^7R^8$, COR^{13}, $X(CO)R^{13}$; R^2 and R^3, together with the carbons to which they are attached form a 4-7 membered monocyclic ring or a 10-14 membered bicyclic ring; R^4-R^8, R^{13} = H, alkyl or R^5 and R^6, or R^7 and R^8 together with nitrogen to which they are attached, form a pyrrolidine, piperidine, morpholine, azetidine, piperazine, thiomorpholine; X = alkylene], and their pharmaceutically acceptable salts, as nicotine binding inhibitors (IC50 < 10 μM) in the treatment of neurol. and psychol. disorders. Thus, aryl-fused azapolycyclic compd. I (R^1-R^3 = H) was prepd. via a multistep synthetic sequence starting from 2-fluorobromobenzene via a cycloaddn. with cyclopentadiene and an amination with triethylbenzylammonium chloride.





I