Evolution of client-focused search reports:

Exemplified compound tabular display linked to citing publications

Maddy Marley
Senior Information Scientist
Global Patents
GSK

AVM2019

GSK Information Science colleagues worldwide

Working in Partnership with the attorneys in Global Patents

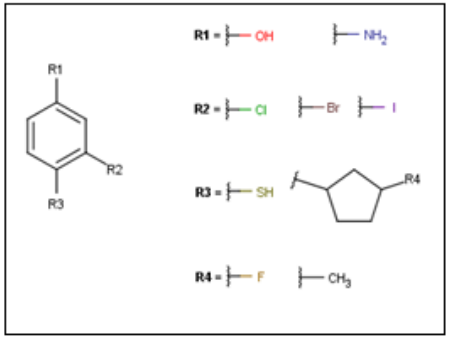
Chemistry structure reports

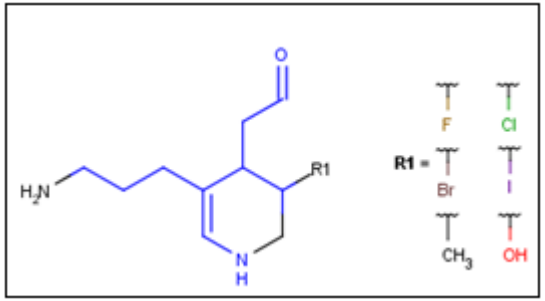
- Produced by the chemistry information scientists for attorneys
- Enable Attorneys to strategically draft their patent applications and discover/overcome any FTO hurdles
- Enable Attorneys to achieve **best possible protection** for GSK's products

  i.e. valid and enforceable patents
- Support the Pharma Consumer and Vaccines businesses
  - **Small molecules, formulations, adjuvants**

Chemistry Reports

- Discovery/pre-candidate/candidate selections
  - Patentability
    - Compounds within Markush
    - Compounds peripheral to Markush + similar use
    - Any source, anywhere. Include proprietary databases
- First time in Human, Commit to Medicine Development, Commit to Phase 3
  - Freedom-to-operate
    - Narrow genus around the candidate structure or process intermediates
    - Patent claims, patents in-force or pending

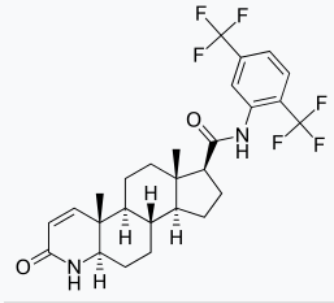


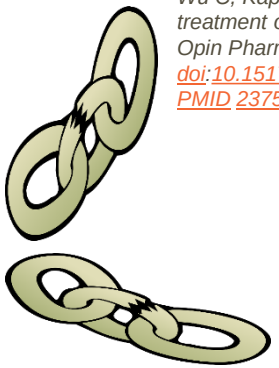


Attorney feedback on reports

- **In order to clearly analyse <u>the compounds</u>** of interest:
  - Present them in a table
  - No duplicates
  - **<u>Link</u>** the compounds to the references
- **In order to view the references**
  - Provide at least one family member, containing the structures (“basic”)
  - Provide the journal source

- **Reports delivered in a WORD document**





*Wu C, Kapoor A (2013). "Dutasteride for the treatment of benign prostatic hyperplasia". Expert Opin Pharmacother. **14** (10): 1399–408. doi:10.1517/14656566.2013.797965. PMID 23750593.*

An oral pharmaceutical composition of dutasteride

EP 2395975

Sources used for Chemistry reports:

- Chemical Abstracts on New STN, STN Classic, STNext
  - Compounds found here are the crown jewels or CROWN OF THORNS
- DWPI
- Reaxys Desktop
  - Covers PubChem, EMolecules and other substance collections
- Virtual libraries, chemical catalogues , Chemical Repositories
  - Zinc, Enamine Real, Pubchem
- Markush Databases

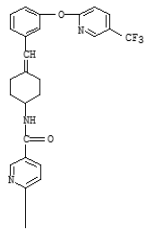
Focus on Chemical Abstracts Registry/CA files

– *Reporting Option 1: "BIB ABS HITSTR"*

- If hit structures belong to >1 reference they are repeat displayed
- Compounds that are mixtures are space-consuming
- Split page–compounds are sometimes undecipherable

3-Pyridinecarboxamide, 6-(1-pyrrolidinyl)-N-[4-[[3-[[5-(trifluoromethyl)-2- pyridinyl]oxy]phenyl]methylene]cyclohexyl]- (CA INDEX NAME)

PAGE 1-A



PAGE 2-A

AN 2017:1847128 HCAPLUS Full-text
TI Crystallization of bisulfite derivatives of enantiomerically enriched verbenone
AU Kovalenko, V. N.; Prat'ko, A. S.; Prokhorevich, K. N.
CS Belarusian State University, Minsk, 220030, Belarus
SO Russian Journal of Organic Chemistry (2017), 53(10), 1598-1600
CODEN: RJOCEQ; ISSN: 1070-4280
DOI 10.1134/S1070428017100190
PB Pleiades Publishing, Ltd.
DT Journal; (online computer file)
LA English
AB After sepn. of cryst. bisulfite derivs. of enantiomerically enriched (1S)- and (1R)-verbenones, steam distn. of the filtrates afforded (1S)- and (1R)-verbenones whose optical purity was higher by 30 and 20%, resp., than that of the initial enantiomers.

CM 1

CRN 7631-90-5
CMF H2 O3 S . Na

• Na

CM 2

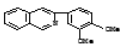
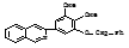
CRN 80-57-9
CMF C10 H14 O

RE.CNT 3 THERE ARE 3 CITED REFERENCES AVAILABLE FOR THIS RECORD
ALL CITATIONS AVAILABLE IN THE RE FORMAT

Focus on Chemical Abstracts Registry/CA files

– *Reporting Option 2* - TABLE TOOL (STN Express)

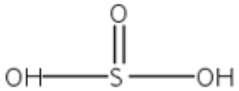
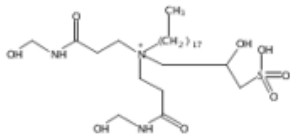
- Impossible to fit-to-size in a report without losing legibility
- Patents and non-patent records mixed together resulting in empty cells
- Hit structures repeat-displayed

Title	Patent Information					Source	Patent Assignee/Corporate Source	Hit Structure
One substrate, two modes of C-H functionalization: a metal-controlled site-selectivity switch in C-H arylation reactions						Organic Letters (2017), 19(1), 262-265 CODEN: ORLEF7; ISSN: 1523-7052	Department of Chemistry, Indian Institute of Science Education and Research Bhopal, Madhya Pradesh, 462066, India	CAS Registry Number 67237-89-2 HCAPLUS Chemical or Trade Name Isoquinoline, 3-(3,4-dimeth INDEX NAME) 
Antimicrobial agents	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE	U.S., No pp. given CODEN: USXXAM	Rutgers, The State University of New Jersey, USA; University of Medicine and Dentistry of New Jersey; Rutgers, The State University of New Jersey	CAS Registry Number 1352036-15-7 HCAPLUS Chemical or Trade Name Isoquinoline, 3-[3,4-dimethoxy-5-(phenylm (CA INDEX NAME) 
	US 8933096	B2	20150113	US 2012-13702936	20121207			CAS Registry Number 1352036-40-8 HCAPLUS

Focus on Chemical Abstracts Registry/CA files

– *Reporting Option 3 New STN, Excel output*

- Lots of editing necessary – no saveable custom display format
- Does not overcome repeat-displaying the structures
- The All Hit Structures link doesn't work in a Word document
  - Unless you "associate" the file containing the structures- this becomes awkward

Title	Hit Structure Images
Crystallization of bisulfite derivatives of enantiomerically enriched verbenone	 • Na All Hit Structures
Preparation of secondary amide quaternary ammonium salt-type hydroxypropyl sodium sulfonate as asphalt emulsifier	 • Cl⁻ • Na All Hit Structures

Enter an IT genius

Maz's Macro

Maz Mazumdar • 2nd

Web solutions developer at Monitor

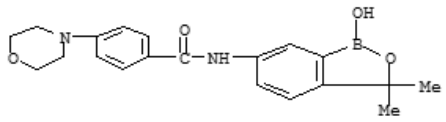
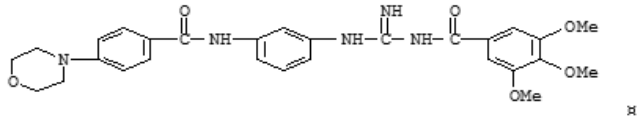
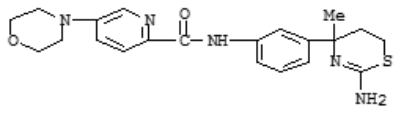
Monitor • Sheffield University

United Kingdom • 106

Connect | Send InMail | More...

Premjit Mazumdar (Maz for short!) is an analyst programmer with over 20 years experience. He specializes in Microsoft technologies, creating and supporting Web, SharePoint, Windows and Microsoft Office applications. He has 6 years experience in the legal sector and over 17 years experience in the pharmaceutical industry, which he first entered in 1989, when he joined Wellcome as an Information Scientist. His interest in programming commenced very shortly afterwards, sparked by a desire to improve the processes with which he and his colleagues conducted their work, allied to a wish to vastly improve the standards of presentation of the analytical reports themselves.

```
    Call IAFunctions.RestoreSettings2(mMyInfo)
    Call IAFunctions.ClearSelectionObject(True)

    MsgBox "Error! Details of the error follow. The macro will then terminate." _
    & vbCrLf & vbCrLf _
    & "Error: " & vbTab & errNumber & vbCrLf _
    & "Where: " & vbTab & errSource & vbCrLf _
    & "What: " & vbTab & errDescription & vbCrLf & vbCrLf _
    & "" _
    , vbCritical, mcsAppName
    On Error Resume Next
    Unload frmExCTF_Progress
    Exit Sub
End Sub


Sub ExCTF2(dummy As Variant)
    'called by the Activate event of frmExCTF_Progress
    'all errors bubbled up to sub ExCmpdTableFormatter from here on

    Const RefAboveTitle As String = "Reference "
    Const ReferencesBookmarkPrefix As String = "Ref"
    Const CompNumberBookmarkPrefix As String = "CompNum"
    Const CompNumPrefix As String = "Compd."
    Const graphicPlaceholder As String = "XX"
    Const STNPromptCharacter As String = "=>"

    On Error GoTo errh

    Dim bMiscDataExists As Boolean
    Dim PlaceHolderRange As Range
    Dim ResultsTable As Range
```

10	Benzamide, N-(1,3-dihydro-1-hydroxy-3,3-dimethyl-2,1-benzoxaborol-6-yl)-4-(4-morpholinyl)-(CA INDEX NAME) 	preparation of benzoxaborole mols. as antiprotozoal agents useful in the treatment of protozoal infection Reference -- 2
11	Benzamide, N-[imino[[3-[[4-(4-morpholinyl)benzoyl]amino]phenyl]amino]methyl]-3,4,5-trimethoxy-(CA INDEX NAME) 	drug candidate; preparation of acylguanidines as inhibitors of Hedgehog protein signalling pathway for treatment of car net dis dia --
12	2-Pyridinecarboxamide, N-[3-(2-amino-5,6-dihydro-4-methyl-4H-1,3-thiazin-4-yl)phenyl]-5-(4-morpholinyl)-(CA INDEX NAME) 	prc tre Al Ref

Fulltext link for patents and non-patents

Reference 2

TI Preparation of benzoxaborole molecules as antiprotozoal agents and their use in the treatment of protozoal infection

AN 2011:200035 HCAPLUS Full-text

DN 154:259801

IN Chen, Daitao; Orr, Matthew; Sligar, Jessica; Jacobs, Robert; Plattner, Jacob J.

PA Anacor Pharmaceuticals, Inc., USA

SO PCT Int. Appl., 202pp. CODEN: PIXXD2

DT Patent

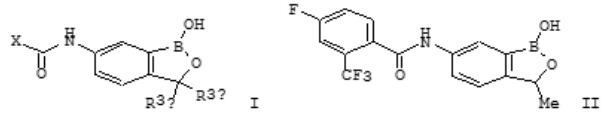
LA English

PI WO 2011019618 A1 20110217 WO 2010-US44787 20100806

PRAI US 2009-61234213 P 20090814 US 2010-61315774 P 20100319 WO 2010-US44787 W 20100806

ASSIGNMENT HISTORY FOR US PATENT AVAILABLE IN LSUS DISPLAY FORMAT

OS CASREACT 154:259801; MARPAT 154:259801GI



I II

IT 1266120-43-7P [Compd.10] RL: ADV (Adverse effect, including toxicity); PAC (Pharmacological activity); PKT (Pharmacokinetics); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses) (preparation of benzoxaborole mols. as antiprotozoal agents useful in the treatment of protozoal infection)

Maz's macro – enables clients to clearly analyse the exemplified compounds –GOAL!

- Runs in Word
- Produces a Table of Compounds from the BIB ABS HITSTR display (STN Express, STNext)
  - No duplicates
- Each compound contains hyperlink(s) to references (s)
- References are in BIB ABS HITIND format, in underneath the table
- Each CASRN in the reference is hyperlinked back to the table compound

- **BUT**
  - Macro written in 2001 for Word 97 – still in use today! (Word 2016)
  - It is very tricky to install, and often falls over
  - Still a problem with large structures (pg1a/pg2a) and references look "old-fashioned"
  - Does not work at all for New STN output
  - GSK needed to seek long-term solution to move forward with post-processing

Enter BizInt!

- Post-search processing market leader
- Highly respected and influential to patent information professionals
- Long-standing relationship with GSK of >10yrs, small organisation, trustworthy, good at listening
- Patent/Non-Patent Smart Charts extensively used in the GP information team, robust product
  - Changes we have requested in the past (example, with Orbit output) have been implemented quickly
- Experienced with STN, and with working together with database providers
- Strong track record for providing post processing solution from many applications/platforms.
- BizInt is an export option in many databases including those key to our work:
  - STN Express/STNext
  - New STN
  - Orbit
  - Cortellis
- BizInt Reference Rows allows merging of the “best” features of databases into one row
  - Orbit patent family legal status and claims with Chemical Abstracts structure displays

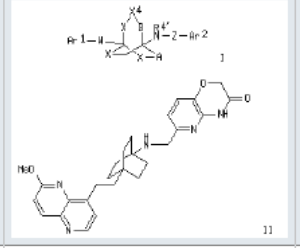
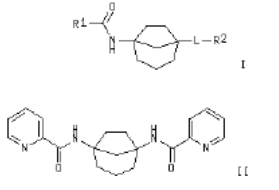
BizInt – Initial Release March 2018 New STN

Software for Business Intelligence
BizInt Smart Charts

gsk

L4 (2158317-14-5 OR 2159095-42-6 OR 2160562-95-6 OR 2160606-03-9 OR 2160606-04-0 OR 2161401-82-5)/rn
CAplus, REGISTRY

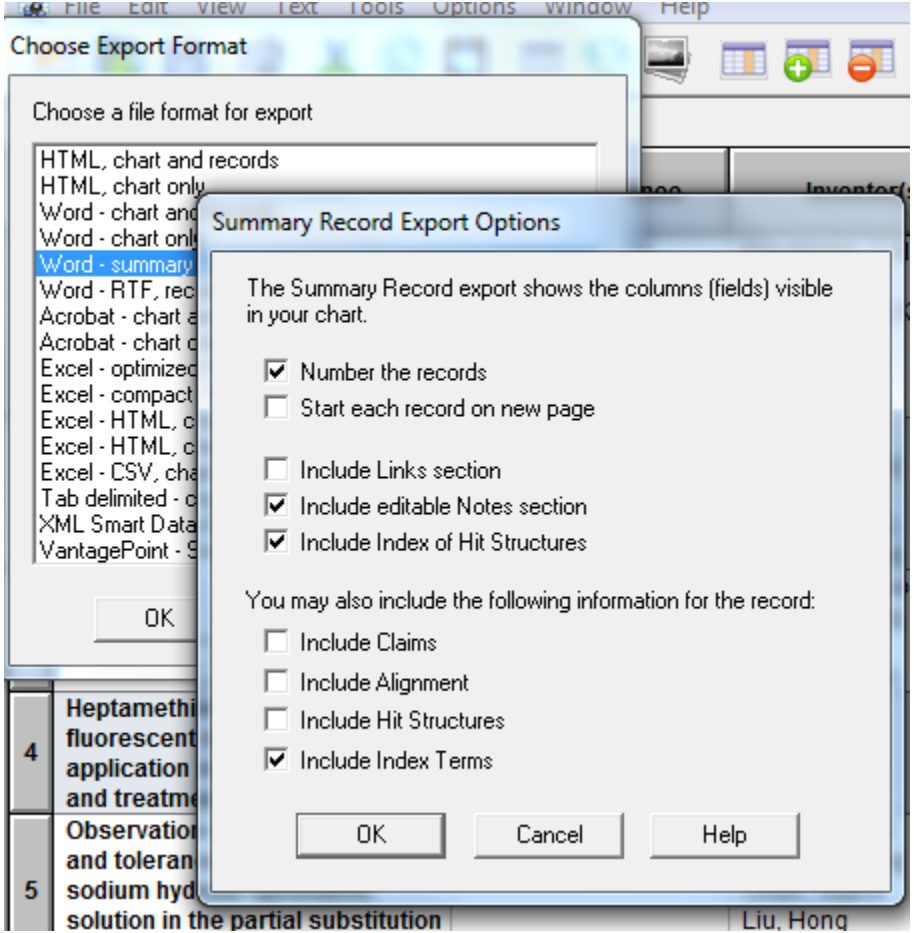
CAplus	5
REGISTRY	6

	Title	Patent Assignee	Abstract	Patent Family			Priority Date	Graphic Information
				Patent	Kind	Date		
1	**Substituted cyclohexylamine compounds**	Epizyme, Inc., USA	The present disclosure provides substituted cyclohexylamine compds. to treat a disorder responsive to the blockade of SMYD proteins such as SMYD3 or SMYD2. Compds. of the present disclosure are esp. useful for treating cancer. MSTR 1 Assembled	WO 2016040502	A1	20160317	2014-09-10	
				CA 2960271	A1	20160317		
				AU 2015315171	A1	20170316		
				EP 3190891	A1	20170719		
				JP 2017527576	T	20170921		
				US 20170362191	A1	20171221		
				US 10106510	B2	20181023		
2	**Bridged bicyclic compounds as antibacterial agents and their preparation and use for the treatment of bacterial infections**	Kyorin Pharmaceutical Co., Ltd., Japan Merck Sharp & Dohme Corp.	Bridged bicyclic compds. of formula I are disclosed herein, along with their pharmaceutically acceptable salts, hydrates and prodrugs. Also disclosed are compns. comprising such compds., methods of prepg. such compds. and methods of using such compds. as antibacterial	WO 2013003383	A1	20130103	2011-06-27	
				CA 2840060	A1	20130103		
				AU 2012275499	A1	20131219		
				EP 2723737	A1	20140430		
				JP 2014518267	T	20140728		
				US 20140243302	A1	20140828		
	Preparation of bicyclo[3.2.1]octyl amide derivatives useful in treating central nervous system diseases or disorders	H. Lundbeck A/S, Den.	The title amide derivs. I [L = NHCO or CONH; R1, R2 = alkyl, cycloalkyl, aryl, etc.] which act as allosteric modulators of the metabotropic glutamate receptor 5, were prepd. Thus, reacting bicyclo[3.2.1]octane-1,5-diamine dihydrochloride (prepn. given) with picolinic acid afforded 16% II. Exemplified compds. I were tested their allosteric mGluR5 activity (data given for representative compds. I)	WO 2012088365	A1	20120628	2010-12-22	
				AR 84457	A1	20130515		
				TW I538905	B	20160621		
				CA 2821937	A1	20120628		
				US 20120190686	A1	20120726		
				US 8921370	B2	20141230		
				AU 2011348188	A1	20130704		
				AU 2011348188	B2	20160616		
				IL 226981	A	20130731		
				CN 103369961	A	20131023		

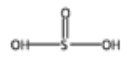
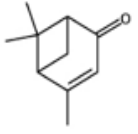
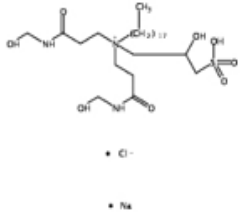
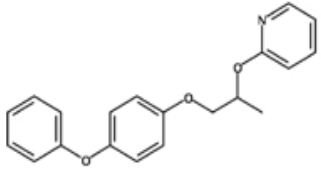
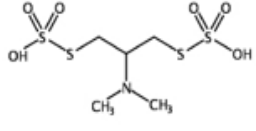
BizInt – Initial Release March 2018 New STN

Software for Business Intelligence
BizInt Smart Charts

gsk



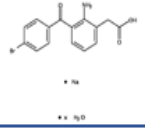
Subsquent releases - STNext

	Substance	Structure	Reference
1	2161401-82-5 Sulfurous acid, compounds, sodium salt, compd. with 4,6,6-trimethylbicyclo[3.1.1]hept-3-en-2-one (1:1:1)	CM1 CRN 7631-90-5  • Na CM2 CRN 80-57-9 	crystn. of bisulfite derivs. of enantiomerically enriched verbenone *Reference 1*
2	2160562-95-6 1-Octadecanaminium, N,N-bis[3-[(hydroxymethyl)amino]-3-oxopropyl]-N-(2-hydroxy-3-sulfopropyl)-, chloride, sodium salt (1:1:1)	 • Cl⁻ • Na	
3	2158317-14-5 Index name not yet assigned	CM1 CRN 95737-68-1  CM2 CRN 29547-00-0 	

Software for Business Intelligence

BizInt Smart Charts

gsk

6	2159035-42-6 Benzeneacetic acid, 2-amino-3-(4-bromobenzoyl)-, sodium salt, hydrate (1:1:?)	 • Na • x H₂O	observation on the availability and tolerance of 0.1% bromfenac sodium hydrate ophthalmic soln. in the partial substitution of glucocorticoid after LASEK *Reference 5*

References

1.

Title:	**Crystallization of bisulfite derivatives of enantiomerically enriched verbenone**
Patent Assignee:	Belarusian State University
Patent Family:	Patent Kind Date
Abstract:	After sepn. of cryst. bisulfite derivs. of enantiomerically enriched (1S)- and (1R)-verbenones, steam distn. of the filtrates afforded (1S)- and (1R)-verbenones whose optical purity was higher by 30 and 20%, resp., than that of the initial enantiomers.
Source:	Russian Journal of Organic Chemistry (20171031) Vol. 53, No. 10, CODEN: RJOCEQ, ISSN: 1070-4280, pp. 1598-1600
Priority Date:	

Notes

Comment: Not seen in the priority filing search

Index Terms:

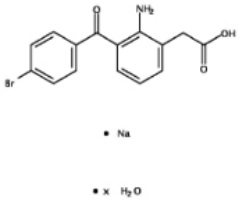
2161401-82-5P (Cmpd. 1) Preparation (PREP); Reactant (RCT); Reactant or Reagent (RACT); Synthetic Preparation (SPN) (crystn. of bisulfite derivs. of enantiomerically enriched verbenone)

2.

Title:	**Preparation of secondary amide quaternary ammonium salt-type hydroxypropyl sodium sulfonate as asphalt emulsifier**
Patent Assignee:	Shandong University, China (CN)
Patent Family:	Patent Kind Date CN107382794 A 20171124
Abstract:	The title asphalt emulsifier is as shown, and is prepd. from octadecyl amine, alc. solvent, N-

BizInt – STNext

– Structures can be referenced to one or multiple patents or non patents, along with the context in the document (CAS indexing)

6	2159095-42-6 Benzeneacetic acid, 2-amino-3-(4-bromobenzoyl)-, sodium salt, hydrate (1:1:2)		heptamethine cyanine fluorescent dye and application in precise diagnosis and treatment of tumor *Reference 4* observation on the availability and tolerance of 0.1% bromfenac sodium hydrate ophthalmic soln. in the partial substitution of glucocorticoid after LASEK *Reference 5*

-And the CAS RNs in the indexing of the document hyperlink back to the table

optical characteristics better than that of ICG, so that it can be used in photodynamic therapy and tumor imaging with good stability. Heptamethine cyanine near IR fluorescent dye of the invention can be used for the accurate diagnosis and treatment of tumor.

Source: Faming Zhuanli Shenqing CNXXEV, pp. 15

Priority Date: 2017-04-21

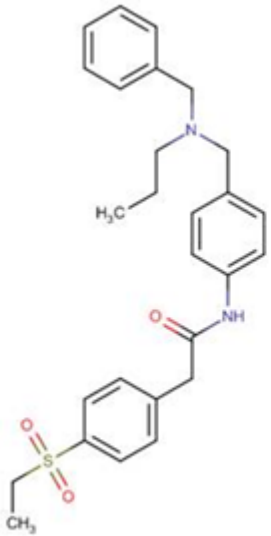
Index Terms:

2160606-03-9P (Cmpd. 4) 2160606-04-0P (Cmpd. 5) Industrial Manufacture (IMF); Preparation (PREP); Reactant (RCT); Reactant or Reagent (RACT) (heptamethine cyanine fluorescent dye and application in precise diagnosis and treatment of tumor)

Future

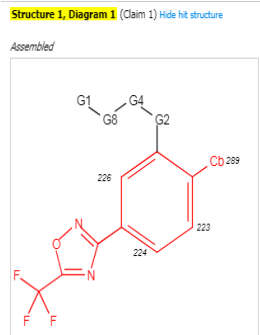
HELP US DISPLAY OPEN-SOURCE DATA

- Example PUBCHEM
- Reaxys offers “deduping” of PubChem records against Reaxys retrieval
  - BUT some PUBCHEM Records offer unique references to substances

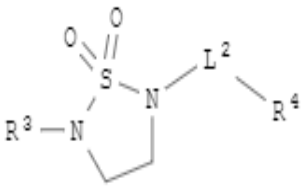
	WO2015131035 (and associated patent family members); see pg. 98 where the compound in question is described as a "certain embodiment" of the invention and a previously published journal article is cited. The patent reference was indexed in **PubChem**.

Future

- **Export of Markush structures as a table with related references**
  - Marpat/DWPIM
  - Export the hit Markush structure and reference
    - Minimize post-processing and maximize readability and information content in the Markush
    - Use colours to show structure overlap
    - New STN has made a good start, but the exportable “hit” Markush not always reflective of the true claim

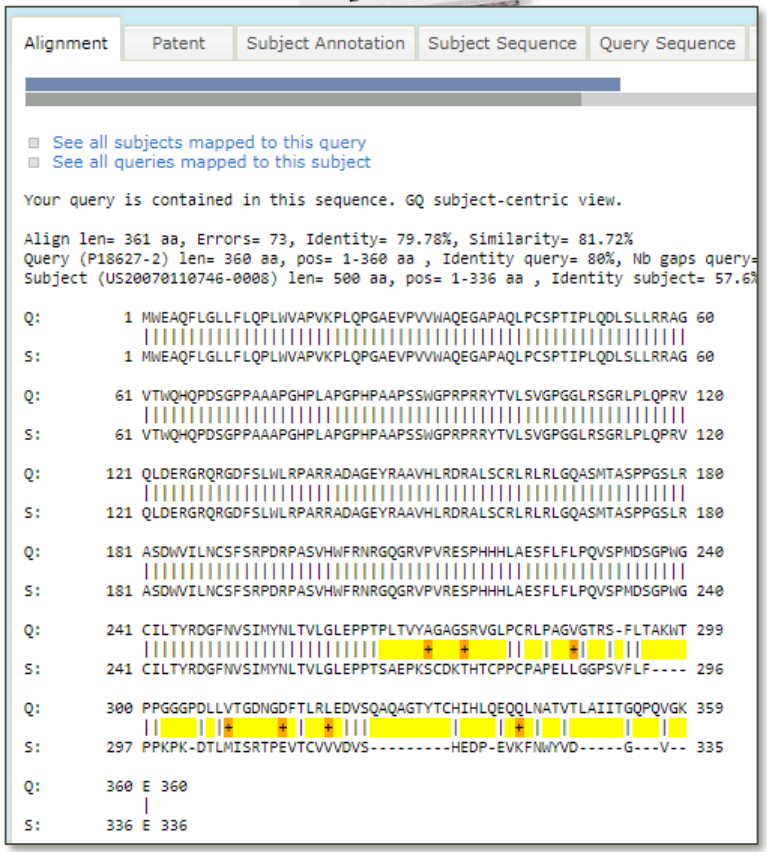


=hit structure from this claim



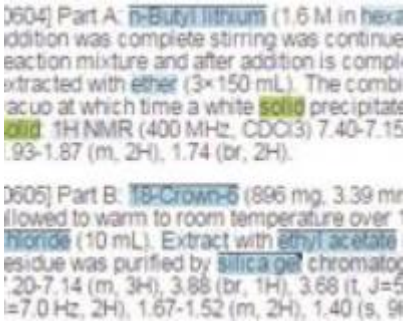
Future

- **Sequence post-processing**
  - Registry, DGene, USgene, PCTGen
  - GenomeQuest



Future

- **Indexing of structures within patents is changing**
  - **Shift from human-indexed to machine-indexed structures**
    - Subscription databases: Semi-IP.com (Chemanalyser), PatBase and Orbit chemical modules
    - Free databases PatentScope and Espacenet
    - Machine-translations improving therefore structures from non-English patents indexed faster



- **Challenges to the Patent Information Professional**
  - Markush searching (for now)
  - Quality And Time and Trust and Recall
  - How do we report multiple source results, without duplicates?

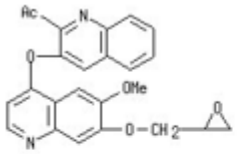
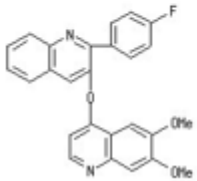
Future

- **One search One Report Workflow**

- Structure-search value-add and full text databases sequentially or simultaneously and offer a similarity-ranked structure output

- 1 Structure Drawing module
- Accessing many structure-searchable databases
- One report:

MOCK UP

	WO 2005080377 , CAPlus WO 2004018430, PatentScope Full-text link Journal of Chromatography (1977), 139(1), 111-20 Reaxys
	WO99123456 , PatBase Chemical Explorer