STRUCTURE SEARCHING DERWENT WORLD PATENTS INDEX® (DWPI\textsuperscript{SM}) USING STN EXPRESS\textsuperscript{®}: Part I

INTELLECTUAL PROPERTY SOLUTIONS

DONALD WALTER
AUGUST 16, 2011
AGENDA

• Part I
  – THE PROBLEM; WHY IS IT DIFFICULT TO SEARCH CHEMICAL STRUCTURES IN PATENTS?
  – SOLUTION; DWPI STRUCTURAL REPRESENTATIONS
  – DWPI PATENT COVERAGE
  – COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?
  – COMPOUND COVERAGE – WHAT INFORMATION ABOUT THE COMPOUNDS IS COVERED?
  – STRUCTURE OF A DCR RECORD

• Part II
  – CHEMICAL STRUCTURE SEARCH EXAMPLE
  – CHEMICAL FRAGMENTATION CODE SEARCH EXAMPLE
  – MULTIFILE STRUCTURE SEARCH EXAMPLE
THE PROBLEM; WHY IS IT DIFFICULT TO SEARCH CHEMICAL STRUCTURES IN PATENTS?

- Although patents are easily word searchable, chemical structures are often not represented by words.
- For example, US 7,956,219 ...

1. A compound of formula I:

![Chemical Structure Diagram]

where

A is a methylene group or is absent;
G is O;
R₁ is selected from the group consisting of hydrogen, hydroxyl, thio, amino, optionally substituted lower alkyl, lower alkylamino, arylamino, aralkylamino, heteroarylamino, heteroaralkylamino, lower alkylxoy, arylxoy, heteroaryloxy, cycloalkyloxy, cycloheteroalkyloxy, aralkyloxy, heteroaralkyloxy, (cycloalkyl) alkylxoy, (cycloheteroalkyl) alkylxoy, lower alkylthio,
AGENDA

• Part I
  – THE PROBLEM; WHY IS IT DIFFICULT TO SEARCH CHEMICAL STRUCTURES IN PATENTS?
  – SOLUTION; DWPI STRUCTURAL REPRESENTATIONS
    – DWPI PATENT COVERAGE
    – COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?
    – COMPOUND COVERAGE – WHAT INFORMATION ABOUT THE COMPOUNDS IS COVERED?
  – STRUCTURE OF A DCR RECORD

• Part II
  – CHEMICAL STRUCTURE SEARCH EXAMPLE
  – CHEMICAL FRAGMENTATION CODE SEARCH EXAMPLE
  – MULTIFILE STRUCTURE SEARCH EXAMPLE
New amino acid analogues are L-arginine transport modulators useful in the treatment of conditions associated with underactivity of nitric oxide synthetic pathway e.g. into cells are new.  

Specific compound numbers - role  

Generic (Markush) compound numbers - role  

Chemical fragmentation codes
AGENDA

• Part I
  – THE PROBLEM; WHY IS IT DIFFICULT TO SEARCH CHEMICAL STRUCTURES IN PATENTS?
  – SOLUTION; DWPI STRUCTURAL REPRESENTATIONS
  – DWPI PATENT COVERAGE
  – COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?
  – COMPOUND COVERAGE – WHAT INFORMATION ABOUT THE COMPOUNDS IS COVERED?
  – STRUCTURE OF A DCR RECORD

• Part II
  – CHEMICAL STRUCTURE SEARCH EXAMPLE
  – CHEMICAL FRAGMENTATION CODE SEARCH EXAMPLE
  – MULTIFILE STRUCTURE SEARCH EXAMPLE
DWPI PATENT COVERAGE

• Pharmaceutical
  – Patents stated to be of pharmaceutical or veterinary interest, as well as those that refer to compounds used as intermediates in the manufacture of pharmaceutical or veterinary products.
  – Patents on compositions used for diagnosis or analysis in the pharmaceutical and/or veterinary fields (e.g. stains for bacterial pathogens).
  – Patents on artificial sweeteners, chemical warfare agents, and plaque-disclosing compositions.
  – Patents dealing with the production of formulations, e.g. tablets, pills, capsules, suppositories, aerosols, etc. Also patents on devices specifically designed for dispensing pharmaceuticals, e.g. syringes, child-proof closures, calendar pill boxes, aerosol devices, etc.
**DWPI PATENT COVERAGE**

- **Agricultural and veterinary**
  - Pest control agents such as insecticides, miticides, rodenticides, molluscicides, slugicides, vermicides (nematocides, anthelmintics, etc.), pest repellents and attractants, and soil fumigants. Also biological control using microorganisms, predators, or natural products.
  - Chemical warfare agents
  - Plant growth control agents such as herbicides, weedicides, defoliants, desiccants, fruit drop and set controllers, rooting compounds, sprouting inhibitors, growth stimulants and retardants, moss and lichen controllers. Also plant genetics.
  - Plant disease control agents such as fungicides, viricides, timber preservatives, and bactericides.
  - Soil improvement agents such as fertilisers, trace metal additives, bacterial action control stimulants, and soil consolidation agents (if used for agricultural purposes).
  - Veterinary products such as disease control agents, nutritional agents, and veterinary vaccines.
• Other chemical compounds (non-polymeric), as well as the apparatus and novel catalysts for producing them. For example, compounds used in cosmetics, fuels, dyes and more).

• Includes their production, purification, use, detection, removal, and phase changes

• Exceptions
  – Monomers taking part in a polymerisation reaction unless the patent is concerned with the production or purification of the monomer
  – Starting materials for a chemical reaction unless the patent is concerned with the production or purification of the starting material.
  – Polymerisation catalysts are not normally classified in Section E unless the novelty of the invention is the catalyst.
  – Mixtures of compounds described as a cut (i.e. hydrocarbon feedstock) in a petrochemical process (usually)
  – Highly complex, non-stoichiometric compounds, e.g. those used as fluorescent materials, might not be but simpler compounds normally are.
  – Solvents and very common reagents, e.g. water
AGENDA

• Part I
  – THE PROBLEM; WHY IS IT DIFFICULT TO SEARCH CHEMICAL STRUCTURES IN PATENTS?
  – SOLUTION; DWPI STRUCTURAL REPRESENTATIONS
  – DWPI PATENT COVERAGE
  – COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?
  – COMPOUND COVERAGE – WHAT INFORMATION ABOUT THE COMPOUNDS IS COVERED?
  – STRUCTURE OF A DCR RECORD

• Part II
  – CHEMICAL STRUCTURE SEARCH EXAMPLE
  – CHEMICAL FRAGMENTATION CODE SEARCH EXAMPLE
  – MULTIFILE STRUCTURE SEARCH EXAMPLE
COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?

The following specific and Markush structures are indexed with BCE Chemical Fragmentation Codes and the Merged Markush Service;

- All compounds and reaction intermediates stated to be novel
- Products of new processes
- New uses of known materials
- Materials detected and detecting agents
- Detection media
- Materials recovered or purified in new ways
- Materials removed and removing agents (only since 1977, unless they were the only chemicals in the invention that could be indexed)
- Components of compositions that are essential to the invention
- Novel catalysts (1970+)
- Activities, properties, and uses
- Chemical formulations and apparatus
COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?

The following **specific** structures are also indexed with Derwent Registry numbers

- It is on the list of ~2100 compounds in the DRN
- Significant compounds or significant non-metallic elements mentioned in the claims or examples
  - An example of an insignificant compound is a solvent mentioned in a process in which any solvent may be used.

More details are in “CPI Chemical Indexing Guidelines Indexing of Chemical and Pharmaceutical Patents”,
http://science.thomsonreuters.com/scientific/m/pdfs/mgr/chemical_index_guidelines.pdf
COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?

The following specific structures are also indexed with the Derwent Chemistry Resource

• All claimed compounds up to a maximum of ~99. This number is reduced if a Markush is also needed. (Max no. of DCR + Markush records =99.).

• At least 1 example, which should be the best example illustrating the invention (usually the one in the abstract). If the abstract (best) example is also claimed, then another should be selected.

• Further examples input at analysts discretion, but more should be selected if there are examples which are structurally dissimilar from those claimed, but still representative of the Markush.

• Selected examples should be "real" not prophetic; i.e. should have supporting data such as preparative data, activity data etc.

• Compounds from the disclosure can be indexed at the Analysts discretion. Usually these would be if there are no (or few) claimed or exemplified compounds, or if there are novel disclosed compounds that are not claimed.
**COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?**

<table>
<thead>
<tr>
<th>Indexing year</th>
<th>Patent says</th>
</tr>
</thead>
<tbody>
<tr>
<td>1963 BCE</td>
<td>BCE</td>
</tr>
<tr>
<td>1981 BCE, MMS</td>
<td>DRN¹</td>
</tr>
<tr>
<td>1987 BCE, MMS</td>
<td>DRN¹, DCR</td>
</tr>
<tr>
<td>1999 BCE, MMS</td>
<td>DRN¹, DCR</td>
</tr>
</tbody>
</table>

Note 1; DRNs are applied if there is a number representing the compound

---

*Patented Aug. 26, 1924.
1,506,316

**UNITED STATES PATENT OFFICE.**

EUGENE A. MARKUSH, OF JERSEY CITY, NEW JERSEY, ASSIGNOR TO PHARMA-CHEMICAL CORPORATION, A CORPORATION OF NEW YORK.

PYRAZOLONE DYER AND PROCESS OF MAKING THE SAME.

Application filed January 9, 1923. Serial No. 611,087.

No Drawing.

To all whom it may concern:

I, Eugene A. Markush, a citizen of the United States, residing at Jersey City, in the county of Hudson and State of New Jersey, have invented certain new and useful Improvements in Pyrazolone Dyes and Processes of Making the Same, of which the following is a full, clear, and exact description.

I have discovered that the diazo compound acetate is utilized in the reacting mixture. Tests, such as with congo paper, should be taken at frequent intervals during the progress of the reaction to be sure that the solution remains alkaline or free from inorganic acids.

A satisfactory test for determining the presence of an excess of pyrazolone in the reacting mixture may be made as follows: A drop of the coupled solution is cooled out.**
AGENDA

• Part I
  – THE PROBLEM; WHY IS IT DIFFICULT TO SEARCH CHEMICAL STRUCTURES IN PATENTS?
  – SOLUTION; DWPI STRUCTURAL REPRESENTATIONS
  – DWPI PATENT COVERAGE
  – COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?
  – COMPOUND COVERAGE – WHAT INFORMATION ABOUT THE COMPOUNDS IS COVERED?
  – STRUCTURE OF A DCR RECORD

• Part II
  – CHEMICAL STRUCTURE SEARCH EXAMPLE
  – CHEMICAL FRAGMENTATION CODE SEARCH EXAMPLE
  – MULTIFILE STRUCTURE SEARCH EXAMPLE
COMPOUND COVERAGE – WHAT INFORMATION ABOUT THE COMPOUNDS IS COVERED*?

* Besides the structure itself

- USES of novel and known materials
- Synthesis
- Chemical formulations and apparatus
- Materials detected and detecting agents
- Detection media
- Materials recovered or purified in new ways
- Materials removed and removing agents (only since 1977, unless they were the only chemicals in the invention that could be indexed)
- Components that are essential to the inventive formulation
- Novel catalysts (1970+)
- Activities, properties, and uses
AGENDA

• Part I
  – THE PROBLEM; WHY IS IT DIFFICULT TO SEARCH CHEMICAL STRUCTURES IN PATENTS?
  – SOLUTION; DWPI STRUCTURAL REPRESENTATIONS
  – DWPI PATENT COVERAGE
  – COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?
  – COMPOUND COVERAGE – WHAT INFORMATION ABOUT THE COMPOUNDS IS COVERED?
  – STRUCTURE OF A DCR RECORD

• Part II
  – CHEMICAL STRUCTURE SEARCH EXAMPLE
  – CHEMICAL FRAGMENTATION CODE SEARCH EXAMPLE
  – MULTIFILE STRUCTURE SEARCH EXAMPLE
## STRUCTURE OF A DCR RECORD

<table>
<thead>
<tr>
<th>DCR number (used to find the bibliographic record covering this cpd)</th>
<th>AN.S  DCR-90406</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferred chemical name</td>
<td>CN.P  CEPHALOSPORIN-C</td>
</tr>
<tr>
<td>Systematic chemical name</td>
<td>CN.S  3-Acetoxyethyl-7-(5-amino-5-carboxypentanoylamino)-8-oxo-5-thia-1-aza-8-oxo-5-thia-1-aza-bicyclo[4.2.0]oct-2-ene-2-carboxylic acid</td>
</tr>
<tr>
<td>Synonyms</td>
<td>SY  CEPHALOSPORIN C; CEPHALOSPORIN-C; VIFAZOLIN</td>
</tr>
<tr>
<td>Molecular formula</td>
<td>MF  C16 H21 N3 O8 S</td>
</tr>
<tr>
<td>Segment molecular formulae</td>
<td>SMF  C16 H21 N3 O8 S *1; TOTAL *1; TYPE *1</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>MW  415.4251</td>
</tr>
<tr>
<td>Structure Derwent Compound Number</td>
<td>SDCN  R00220</td>
</tr>
<tr>
<td>Structure Derwent Registry Number</td>
<td>SDRN  0220</td>
</tr>
<tr>
<td>Class of compound</td>
<td>CC  BETA LACTAMS</td>
</tr>
</tbody>
</table>
AGENDA

• Part I
  – THE PROBLEM; WHY IS IT DIFFICULT TO SEARCH CHEMICAL STRUCTURES IN PATENTS?
  – SOLUTION; DWPI STRUCTURAL REPRESENTATIONS
  – DWPI PATENT COVERAGE
  – COMPOUND COVERAGE – WHICH COMPOUNDS ARE COVERED?
  – COMPOUND COVERAGE – WHAT INFORMATION ABOUT THE COMPOUNDS IS COVERED?
  – STRUCTURE OF A DCR RECORD

• Part II
  – CHEMICAL STRUCTURE SEARCH EXAMPLE
  – CHEMICAL FRAGMENTATION CODE SEARCH EXAMPLE
  – MULTIFILE STRUCTURE SEARCH EXAMPLE
FOR SUPPORT and TRAINING

• Customer Service
  For subscriptions, pricing and renewals
  http://science.thomsonreuters.com/custserv/

• Technical Support
  For access, content, searching, troubleshooting
  and technical issues.
  http://science.thomsonreuters.com/techsupport/

• Training
  For product or IP-related training options.
  http://ip.thomsonreuters.com/training/ti/

---

Offices around the World

US, Canada & Latin America
Phone: +1 800 336 4474
Fax: +1 215 386 2911
ts.info.us@thomsonreuters.com

Europe, Middle East and Africa
Tel: +44 (0)20 7433 4000
Fax: +44 (0)20 7433 4001
ts.info.emea@thomsonreuters.com

Japan
Phone: +81 3 5218 6500
Fax: +81 3 5218 7840
ts.info.jp@thomson.com

Asia Pacific (Singapore office)
Phone: +65 6411 6888
Fax: +65 6223 2634
ts.support.asia@thomson.com
Structure Searching Derwent World Patents Index® (DWPI\textsuperscript{SM}) using STN Express\textsuperscript{®}: Part II

Robert Austin – FIZ Karlsruhe
Agenda

• Chemical structure search example
  – The DWPI Chemistry Resource (DCR)
• Chemical fragmentation code search example
  – Generating and searching queries using STN Express
• Multifile structure search example
  – Searching in combination with REGISTRY/CAplus™
Search Question:
Search for DWPI patent references to specific carbapenem derivatives of substructure (I)

Learn more about the basics of structure searching: [http://www.cas.org/support/stngen/stndoc/structure.html](http://www.cas.org/support/stngen/stndoc/structure.html).
Chemical structure search options

- Exact search (**EXA**) – Retrieves specific compounds and isotopes
- Family search (**FAM**) – Retrieves specific compounds, isotopes, salts and mixtures
- Closed Substructure Search (**CSS**) – Allows for substitution at defined positions
- Substructure Search (**SSS**) – Allows for substitution at any position
- Sample search (**SAM**) – Free-of-charge pre-search
- Subset search (**SUBSET**) – Structure search on a database subset
- Batch search (**BATCH**) – For broad structure queries which may reach system limits
How to run a structure search in DWPI

1. Draw & save the structure query in standard format with STN Express (or STN on the Web)

2. Upload the structure query to DWPI on STN

3. Run the structure search, e.g.:
   a. Sample search => S L1 SSS SAM
   b. Full search => S L1 SSS FULL

4. Retrieve DWPI patent records
   E.g.: => S L2/DCR

5. Display DWPI patent records with DCR hit structures (HITSTR) in-context
   E.g.: => D L3 1- FULLG HITSTR
Draw & save the structure query in standard format with STN Express.
Upload the structure query to DWPI on STN

Upload the query with the ‘Q’ button.
Run a sample structure search

=>
Uploading C: . . . .\My Documents\STN Express 8.4\Queries\CARBAPENEM.str

L1 STRUCTURE UPLOADED

=> D
L1 HAS NO ANSWERS
L1 STR

Structure attributes must be viewed using STN Express query preparation.

=> S L1 SSS SAM
SAMPLE SEARCH INITIATED 15:58:51 FILE 'WPIX'
SAMPLE SCREEN SEARCH COMPLETED - 1
100.0% PROCESSED 157 ITERATIONS 50 ANSWER

L2 50 SEA SSS SAM L1

50 compounds are retrieved (L2).

Option: display the query (L1), to verify that the Upload was successful.

The uploaded structure query (L1).

Run a substructure (SSS) sample (SAM) search using the query (L1).
Review some answers using D SCAN

The effectiveness of the query (L1) may be assessed by reviewing some records (L2), e.g. using the free-of-charge SCAN format.

HOW MANY MORE ANSWERS DO YOU WISH TO SCAN? (1): 3
Run a full structure search

Run a substructure (SSS) full-file (FUL) search using the query (L1).

1210 compounds are retrieved (L3).

678 DWPI patent records are retrieved (L4).
Producing biosensor for identifying anti-infective substances by selecting clones with anti-infective-responsive reporter expression and identifying clones exhibiting different anti-infective-responsive reporter expression as biosensors.
Searches may be refined using roles

- “Roles” in DWPI describe the function of a compound in the patent, e.g.
  - Compound is prepared, purified or part of a mixture
  - Compound is claimed or from the examples
- DCR compounds can have two types of roles assigned in DWPI
  - 2-3-letter codes in the IT-display (DCR Roles)
  - Single-letter codes in the CMC display (DCN Roles)
- Roles can be linked to DCR numbers in the DCR search field using the (T)-proximity operator
Example: DCR role searching for preparations

=> S L3/DCR (T) (PRD OR P OR NEW OR N)/DCR
L5 290 L3/DCR (T) (PRD OR P OR NEW OR N)/DCR

=> D TRIAL HIT 2

L1 ANSWER 2 OF 290 WPIX COPYRIGHT 2011
AN 2011-G80386 [201143] WPIX
TI Preparing carbapenem derivatives, useful as antibacterial agent, comprises contacting or incubating a carbapenem substrate with ThnQ enzyme
DC B02; D16
IPCI A61K0031-407 [I,A]; A61P0031-00 [I,A]; C07D0477-20 [I,A]; . . . .
MC CPI: B02-P; B06-D04; B11-A02A3; B14-A01; D05-A02A; D05-C
IT UPIT 20110708
108756-CL 108756-PRD; 2512871-CL 2512871-PRD; 2512872-CL 2512872-PRD; 2512870-CL 2512870-PRD; 1109-67201-CL . . . .
CMC UPB 20110708
RIN: 41252
DCN: R04390-K R04390-P
DCR: 108756-K 108756-P

L3 = carbapenem structure search.

Use (T)-operator to link DCR numbers to their roles (/DCR).

PRD, P – preparation roles.
NEW, N – new compound roles.

DCR-108756: Thienamycin.
DCR searches may also be refined using chemical fragmentation codes (CMC)

Example: Q263 = sunscreen agent. Q262 = skin protection (except Q263). P94+ = burn, wound and skin treatments.

DCR-97728: Imipenem.
Search Question:
Search for DWPI patent references to all carbapenem derivatives of substructure (I)
How to run a fragmentation code search

1. Draw & save the structure query in *WPI format*
2. Select Query and *Generate WPI strategy* from the main STN Express window
3. Select the fields to be searched, e.g. /M0,M2, and the strategy is generated and saved
4. Logon to STN, access WPIDS or WPIX, then select Query and Run Command File
5. Select the command file – this runs line-by-line automatically – and review the results
Open, Draw and save the query structure in “WPI” format

Open a new WPI format drawing window.

Note: By convention, unlike an STN structure search, all unsubstituted positions are assumed to be H.

Draw and save the WPI structure.
Alternatively a standard format structure can be converted to a *WPI format* structure.

A standard STN structure query can usually be converted to *WPI format*, for generating a fragmentation code query.
Use STN Express to generate the fragmentation code script.
Select the search fields (subheadings) and generate the fragmentation codes

STN Express search field options:

/M0 Pre-1970 Pharma/agrochem
/M2 1970-date Pharma/agrochem
/M3 1970-date Other chemicals (excl. M4)
/M4 1970-date Dyes & pigments

Other fragmentation code fields

/M1 1970-date Natural products
/M5 1963-1999 Steroids
/M6 1976-date Galenicals/formulations
DWPI fragmentation code strategy is generated in *Command File* format

**Tip:** Check that any manual edits you make are valid codes, with *Utilities, Check Command File.*
Use *Run Command File* to execute the fragmentation code search online.
The command file code query runs automatically line-by-line.

STN Express links and groups codes together into correctly formatted queries.

Note: A relatively small answer set (L16) is retrieved because this is not a substructure search.
Fragmentation code searches may also be refined with role codes

=> S L16 (P) (M720 OR M710)/M2
L17 78 L16 (P) (M710 OR M720)/M2

=> D AN TI HITCMC

L17 ANSWER 1 OF 78 WPIX COPYRIGHT 2011 AN 2009-F45171 [200966] WPX
TI New 4-alkyl-7-oxo-1-azabicyclo-(3,2,0)-hept-2-ene-2-carboxylic acid derivative for use in preparing medicine for treating and/or preventing infectious diseases CMC UPB 20091015


RIN: 00904 41252

MCN: 1058-36001-N

This hit is a Markush compound.

L16 = carbapenem code search.
M720 – preparation role code.
M710 – new compound role code.

The HITCMC format is often helpful for displaying hit fragmentation paragraphs.
Option: Edit the fragmentation code query, to run a full substructure search....
Run the edited fragmentation code query

=> S (D690(P)J11! (P)J52! (P) (M412 OR M411))/M0,M2
L18 3236 (D690(P)J11! (P)J52! (P) (M412 OR M411))/M0,M2

=> S L18(P) (M511 OR M512 OR M513)/M2
L19 3218 L18(P) (M511 OR M512 OR M513)/M2

=> S L19(P) (M321 OR M322 OR M323)/M2
L20 3100 L19(P) (M321 OR M322 OR M323)/M2

=> S L20(P)41252/RIN
L21 1275 L20(P)41252/RIN

=> S L21(P) (D013(P) (D011 OR D019 OR D014) (P) "L941")/M2
L22 1267 L21(P) (D013(P) (D011 OR D019 OR D014) (P) "L941")/M2

=> S (L18(P)M900/M0) OR (L19(P)M901/M2) OR (L21(P)M902/M2) OR L22
L23 1285 (L18(P)M900/M0) OR (L19(P)M901/M2) OR (L21(P)M902/M2) OR L22

=> S L23(NOTP) "L1"/M2
L24 1283 L23(NOTP) "L1"/M2

Note: This is a substructure search (L24).

The Command File runs automatically line-by-line.
Review fragmentation code answers

The fragmentation code hit structure (FRAGHITSTR) format displays the DCR structure for any specific compounds retrieved.
Option: Use the (NOTP) operator to exclude hits already found via the DCR structure search

=> S L24 (NOTP) L4
L25  819 L24 (NOTP) L4

=> D AN TI HITCMC

L25  ANSWER 1 OF 819  WPIX COPYRIGHT 2011  THOMSON REUTERS on STN
AN  2011-H38000 [201146]  WPIX
TI  New carbapenem derivatives useful for treatment of infection by bacteria e.g. methicillin-resistant staphylococcus aureus
CMC  UPB  20110720
M2  *17*  C316  D013  D014  D690  F011  F012  F014  F423  G010  G011  .  .  .
     H682  H683  H684  H685  H686  H689  H8  H9  J0  J012  J013  J014  J111
     K399  K442  K499  L143  L145  L199  L640  L650  L660  L699  L9  L941  M1
     M126  M142  M210  M211  M212  M213  M214  M215  M216  M231  M232  M233
     M240  M262  M271  M272  M273  M281  M282  M283  M311  M312  M313  M314
     M372  M373  M381  M382  M383  M391  M392  M393  M412  M511  M521  M531
     M540  M630  M640  M650  M710  P210  P220  P241  P420  P714  M905  M904
RIN:  41252
MCN:  1111-00001-N

This hit is a Markush compound.
If we explore this answer further...

Use the IFULLG format to see patent family, abstract, and selected drawing image.
### Display formats for reviewing chemical fragmentation codes in DWPI

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>D CODE</td>
<td>All codes, indexing and classes</td>
</tr>
<tr>
<td>D CMC</td>
<td>All chemical codes, M0-M6</td>
</tr>
<tr>
<td>D M2, etc</td>
<td>Pharma/Agrochem only, etc</td>
</tr>
<tr>
<td>D RIN</td>
<td>Ring Index Numbers only</td>
</tr>
<tr>
<td>D HITCMC</td>
<td>The hit fragmentation code paragraph(s) found in a fragmentation code search</td>
</tr>
<tr>
<td>D FRAGHITSTR</td>
<td>The DCR structure for any specific compounds retrieved in a filtration code search</td>
</tr>
</tbody>
</table>
Search Question:
Search for all patent references of cetirizine, including salts and mixtures
Multifile structure searching using DCR/DWPI and REGISTRY/CAplus

1. Prepare a suitable standard structure query for REGISTRY/CAplus and DCR/DWPI
2. Run the structure search in REGISTRY/CAplus
3. Display CAplus records of interest
4. Run the structure search in DCR/DWPI, and remove duplicates between CAplus and DWPI
5. Display any additional DWPI records

Note: A comprehensive STN search would also include CAS MARPAT and DWPI Chemical Fragmentation Codes.
Run a sample structure search

```
=> FILE REGISTRY

=>
Uploading C:\. . . .\My Documents\STN Express 8.4\Queries\cetirizine.str

L1   STRUCTURE UPLOADED

=> D
L1 HAS NO ANSWERS
L1   STR

Structure attributes must be viewed using STN Express query preparation.

=> S L1 FAM SAM
SAMPLE SEARCH INITIATED 18:58:00 FILE 'REGISTRY'
SAMPLE SCREEN SEARCH COMPLETED - 14
.
.
.
L2   3 SEA FAM SAM L1

The uploaded structure query (L1).

Option: display the query (L1), to verify that theUpload was successful.

Run a family (FAM) sample (SAM) search using the query (L1).

3 compounds are retrieved (L2).```
The effectiveness of the query (L1) may be assessed by reviewing some records (L2), e.g. using the free-of-charge SCAN format.
Run a full structure search

Run a family (FAM) full-file (FUL) search using the query (L1).

The full file structure search retrieves 85 REGISTRY records (L3).

The full file structure search (L3) retrieves 761 patent records (P/DT) in CAplus (L4).
Example: CAplus answer retrieved

Display CAplus patent records with in-context hit structures (HITSTR).
Repeat the structure search in DCR

=> FILE WPIX
Access the Derwent World Patents Index (DWPI) (files WPINDEX, WPIDS or WPIX).

=> S L1 FAM FUL
FULL SEARCH INITIATED 18:59:24 FILE 'WPIX'
FULL SCREEN SEARCH COMPLETED - 49 ITERATIONS
100.0% PROCESSED 49 ITERATIONS
44 ANSWERS
SEARCH TIME: 00.00.02

L5 44 SEA FAM FUL L1

Repeat the structure search using the same query (L1).

The full file family search retrieves 44 DCR records (L5).

=> S L5/DCR

The 44 DCR records (L5) retrieve 611 patent records (L6).
Retrieve the unique DWPI records

Use TRANSFER PNK to crossover the CAplus results (L4) to DWPI (L8).

The DCR/DWPI search (L6) adds 117 unique inventions (L9) to the REGISTRY/CAplus results (L4).

PNK = The patent number/kind code field. This is the most precise field to use when crossing over results between STN patent databases.
Review the unique DWPI records

=> D AN TI 1-50

L9 ANSWER 1 OF 117 WPIX COPYRIGHT 2011 THOMSON REUTERS on STN
AN 2011-J36917 [201149] WPIX
TI Interleukin-containing composition, comprises interleukin-1 and a cyclooxygenase inhibitor, which is diclofenac

L9 ANSWER 8 OF 117 WPIX COPYRIGHT 2011 THOMSON REUTERS on STN
AN 2010-K89255 [201064] WPIX
TI Oil in water emulsion aerosol foam, useful for treating a skin disease e.g. acne and psoriasis, composition comprises active agent, water, oil, oil miscible organic solvent, surfactant component and propellant

L9 ANSWER 13 OF 117 WPIX COPYRIGHT 2011 THOMSON REUTERS on STN
AN 2010-F66310 [201037] WPIX
TI Making levocetirizine involves resolving pair of diastereomer of piperazine compound into single diastereomer, and converting single diastereomer of piperazine compound with 1-((4-chloro-phenyl)-phenyl-methyl)-piperazine into levocetirizine

L9 ANSWER 38 OF 117 WPIX COPYRIGHT 2011 THOMSON REUTERS on STN
AN 2009-J67366 [200936] WPIX
TI Orally disintegrating tablet useful for delivering main ingredient e.g. cetirizine hydrochloride for treating allergy, comprises main ingredient having specified water solubility and does not contain binder
Example: DWPI answer retrieved

| L1   | ANSWER 13 OF 117 WPIX COPYRIGHT 2011 THOMSON REUTERS on STN |
| AN   | 2010-F66310 [201037] WPIX |
| TI   | Making levocetirizine involves resolving pair of diastereomer of piperazine compound into single diastereomer, and converting single diastereomer of piperazine compound with 1-((4-chloro-phenyl)-phenyl-methyl)-piperazine into levocetirizine |
| DC   | B03; B05 |
| IN   | FIRET J J; ZHU J |
| PA   | (SYNT-N) SYNTON BV |
| CYC  | 122 |

| PIA  | WO 2010057515 A1 20100527 (201037)* EN 35[0] |
|      | NL 1037485 C 20100727 (201140) NL |
| PRAI | WO 2008-EP9977 20081121 |
| AN.S | DCR-174423 |
| CN.P | LEVOCETIRIZINE |
| CN.S | 2-{4-[(4-Chloro-phenyl)-phenyl-methyl]-piperazine | acid |
| SDCN | RA0WVA |
Summary

• Chemical structure search example
  – The DWPI Chemistry Resource (DCR)
• Chemical fragmentation code search example
  – Generating and searching queries using STN Express
• Multifile structure search example
  – Searching in combination with REGISTRY/CAplus
For more information ...

CAS
E-mail: help@cas.org
Support and Training:
www.cas.org

FIZ Karlsruhe
helpdesk@fiz-karlsruhe.de
Support and Training:
www.stn-international.de