
DERWENT
WORLD PATENTS INDEX®

Chemistry Resource
on STN

Copyright © 2007 The Thomson Corporation

Edition 2
ISBN: 978 1 905935 08 6

Copyright © 2007 The Thomson Corporation

Visit the Thomson Scientific web site at scientific.thomson.com

First edition published May 2001

Second edition published June 2007

ISBN: 1 903836 05 2 (Edition 1)

ISBN: 978 1 905935 08 6 (Edition 2)

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means – electronic, mechanical, recording, photocopying or otherwise – without express written permission from the copyright owner.

Contents

1	Introduction	1
2	Customer Service Information	3
	2.1 Customer Technical Support	3
	2.2 Other Sources of Help and Information	5
3	Chemistry Resource	7
	3.1 Introduction	7
	3.2 STN	7
4	Sample Record	9
5	Structure Searching	11
	5.1 Structure Modelling and a Simple Start	11
	5.2 Subset Searching	12
6	Crossing over into the bibliographic segment	15
7	Current Awareness Searching	19
8	DCR Search & Display Fields	23
	8.1 Field Code	23
	8.2 Predefined Display Formats:	24
9	Identifiers	27
	9.1 DCR Number (AN.S)	27
	9.2 Structured DCR Number (DCSE)	29
	9.3 Structure Cross Reference (SCR syn XCR)	30
	9.4 Other structure identifiers (SDCN, SDRN, SRIN)	32
10	Formula Fields	35
	10.1 Molecular Formula (MF)	35
	10.2 Standardized Molecular Formula (SMF)	35
	10.3 Component Molecular Formula	36
	10.4 Element Symbol (ELS) and Element Symbol Count (ELS.CNT)	36
	10.5 Number of Components (NC)	38
	10.6 Number of Fragments (NFRAG)	39
11	Chemical Name Fields	41
	11.1 Chemical Name (CN)	41
	11.2 Chemical Name Preferred (CN.P)	41
	11.3 Systematic Chemical Name (CN.S)	41
	11.4 Chemical Name Segment (CNS)	42
	11.5 Synonym Name (SY)	42

12	Text Data Fields	43
12.1	Classification Codes (CC)	43
12.2	Comment (CMT)	44
12.3	List of available Substance Descriptors	44
13	Structure Display	47
13.1	Structure (STR)	47
14	Update Dates	49
14.1	Entry Date Chemical Repository (EDCR)	49
14.2	Update Date Chemical Repository (UPCR)	49
14.3	Update Date DWPI Cross Reference (UPWX)	49
15	Supplementary Fields	51
15.1	Controlled term (CT)	51
15.2	Controlled Term, Drug Activity (CT.DA)	52
15.3	Controlled Term, Mechanism of Action (CT.MA)	53
15.4	Drug Registry Name (DDRN)	54
15.5	Multipunch Code (MPC)	56
15.6	Molecular Weight (MW)	56
15.7	Substructure Search Terms (SS)	57
16	Appendix	59
16.1	Definitions of Substance Descriptors/ Classification Codes	59
17	Frequently Asked Questions	69
	Content	69
	Index	71

1 Introduction

For over 50 years, Thomson Scientific, the world's leading patent information publisher, has been foremost in providing scientific and technical intelligence to business, industry, government and research institutes throughout the world.

The *Chemistry Resource* (DCR) database gives users the chance to keep abreast of technological advances by easily searching the chemical structures within the *Derwent World Patents Index* (DWPI). Studies have shown that over 70% of the information contained within patents is never published anywhere else, so patents provide a unique source of information on new technology.

STN International® is an online information service operated jointly by the Chemical Abstracts Service (CAS®), Columbus, OH; FIZ Karlsruhe, Germany; and The Japan Science and Technology Corporation (JST). STN® provides access to a wide range of scientific and technical databases, accessed either via the internet or using STN Express® communication software. FIZ Karlsruhe, the European partner of STN International, provides over half of the databases available on STN. Among others, this includes *Derwent World Patents Index*, *Derwent World Patents Index First View*SM, *GENESEQ*TM and *Patents Citation Index*.

This User Guide describes the content, coverage and search capabilities of the *Chemistry Resource* database as it is implemented on the Scientific and Technical Information Network (STN). *Chemistry Resource* is provided on STN International by FIZ Karlsruhe GmbH, Karlsruhe, Germany. Following detailed information on database content, separate chapters detail each of the search and display fields available.

For further information concerning DCR and other information services offered by Thomson Scientific, please contact your local office.

2 Customer Service Information

2.1 Customer Technical Support

Expert advice and support is available via our Customer Technical Support staff, to provide a fast and efficient response to all your enquiries. Our experienced Technical Support staff has an in-depth knowledge of the products and services offered by Thomson Scientific and are familiar with the various command languages.

From general customer queries through to technical questions, the Technical Support department is there to help you.

Email (all regions)

Please use the Feedback Form on the following Web page:
scientific.thomson.com/support

Your message will automatically be directed to your nearest Support Center.

Europe, Middle East and Africa

Thomson Scientific
77 Hatton Garden
London EC1N 8JS
United Kingdom

Phone: +44 207 433 4999

Americas

Thomson Scientific
3501 Market Street
Philadelphia
PA 19104
USA

Phone: +1 800 336 4474 ext. 1591

+1 215 386 0100

Fax: +1 215 386 6362

Japan

Thomson Corporation K.K.
Palaceside Building
5F East
1-1-1 Hitotsubashi 1-Chome
Chiyoda-ku
Tokyo 100-0003
Japan

Phone: 0800 888 8855 (from Japan only)
+81 3 5218 6500
Fax: +81 3 5218 6536

Asia Pacific (Singapore)

Thomson Scientific
16 Collyer Quay
22nd Floor
Hitachi Tower
Singapore 049318

Phone: +65 6879 4118
Fax: +65 6223 2634

China

Thomson Scientific
Room 407, Raycom InfoTech Park Tower A
No. 2 Kexueyuan South Road
Haidian District
Beijing 100080
P.R.China

Phone: +86 10 8286 2099
Fax: +86 10 8286 2088

Korea

Thomson Scientific
13F Hungkuk Life Insurance Bldg. 226
Shinmunro 1-ga, Jongro-gu
Seoul 110-786
Korea

Phone: 0800108100
+82 2 2076 8100
Fax: +82 2 2076 8122

Australia and New Zealand

Thomson Scientific
Level 5
1000 Harris Street
Pyrmont, NSW 2009
Australia

Phone: Australia 1800 007 214
New Zealand 0800 443 162
Fax: +65 6223 2634

Customer Support Center

Optimize your use of Thomson Scientific products and services. In our comprehensive Support Center you are just a click away from the reference materials you need.
scientific.thomson.com/support/

STN International

Fiz Karlsruhe
STN Service Centre Europe
Hermann-von-Helmholtz-Platz 1
76344 Karlsruhe

Tel: +49 7247 808 555
Fax: +49 7247 808 259
E-mail: helpdesk@fiz-karlsruhe.de
<http://www.fiz-karlsruhe.de/>

2.2 Other Sources of Help and Information

User Guides

To help you search *Derwent World Patents Index* more effectively, Thomson Scientific provides a comprehensive range of user guides, all written by specialists in their field. The user guides listed below include topics relevant to searching the database. The majority of these are available as PDF files on our web site at:
scientific.thomson.com/support/userguides

3 Chemistry Resource

3.1 Introduction

The *Chemistry Resource* (DCR) database is a chemical structure database for searching specific compounds indexed in *Derwent World Patents Index*[®] (DWPI) bibliographic records. DCR indexing commenced in DWPI Update 199916. The database is searchable both by chemical structure and by various text fields, allowing simple access to the DWPI database by specialist and non-specialist chemical searchers alike. *Chemistry Resource* runs in parallel to, and to a certain extent replicates, current subscriber Chemical Indexing (Fragmentation Codes) for patents classified in Chemical Patents Index (CPI) Sections B (Pharmaceuticals), C (Agrochemicals) and/or E (General Chemicals).

DCR Numbers, which are unique identifiers for specific chemical compounds, form the link between the DCR chemical structure database and corresponding bibliographic indexing in DWPI. The online functionality that provides this connection reflects the capabilities of the STN systems involved.

Up-to-date news about DCR, including Frequently Asked Questions can be found on the Thomson Scientific website:

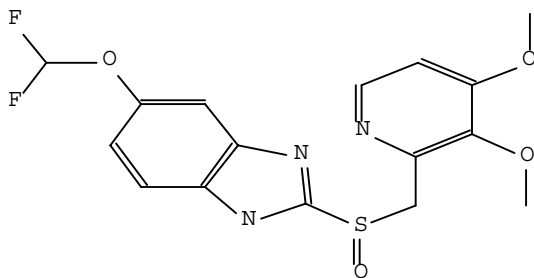
<http://scientific.thomson.com/products/chemistryresource/>

3.2 STN

DCR on STN is available to all searchers, using standard structure searching techniques, via either STN Express or STN on the Web. The database is provided as a seamless part of the DWPI files WPINDEX, WPIDS and WPIX, rather than as a separate file.

4 Sample Record

```
L7 ANSWER 17 OF 17 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN.S DCR-111250
DCSE 111250-0-0-0
CN.P PANTOPRAZOLE
CN.S 5-Difluoromethoxy-2-(3,4-dimethoxy-pyridin-2-ylmethane
sulfinyl)-1H-benzoimidazole
SY CONTROLOC; INIPOMP; PANTOLOC; PANTOPRAZOLE; PANTOZOL;
PEPTAZOL; PROTIUM; PROTONIX; PROTONIX-IV; RIFUN; SKF-96022;
SOMAC
```

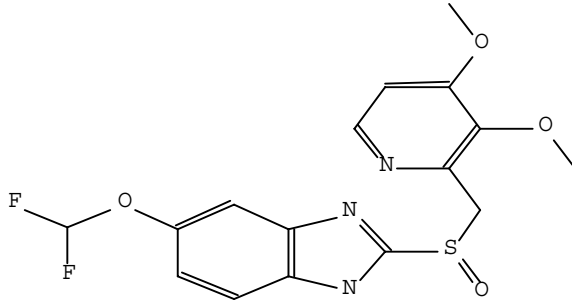


```
MF C16 H15 F2 N3 O4 S
SMF C16 H15 F2 N3 O4 S *1; TOTAL *1; TYPE *1
MW 383.3765
SDCN R22667
```

AN.S contains the *DCR* number primary key which is the unique and unambiguous structure identifier. This is also used for crossing over to the bibliographic (*DWPI*) file segment.

DCSE contains the enhanced *DCR* number which contains information about stereochemistry, isotopes or charges and may serve to aggregate related compounds by masking parts of it.

L7 ANSWER 16 OF 17 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN.S DCR-159347
DCSE 111250-0-1-0
CN.P PANTOPRAZOLE SODIUM
SY PANTOPRAZOLE SODIUM; PROTIUM; SOMAC
CM 1
Na
CM 2



MF C16 H15 F2 N3 O4 S . Na
SMF C16 H15 F2 N3 O4 S *1; TYPE *2; TOTAL *2; Na *1
MW 406.3635

5 Structure Searching

Structure searching is available on STN Express, command line and STN on the Web. The following are features relating to the structure searching of DCR. (additional manuals covering general structure searching techniques are available from STN)..

Structure search modes available in DCR are:

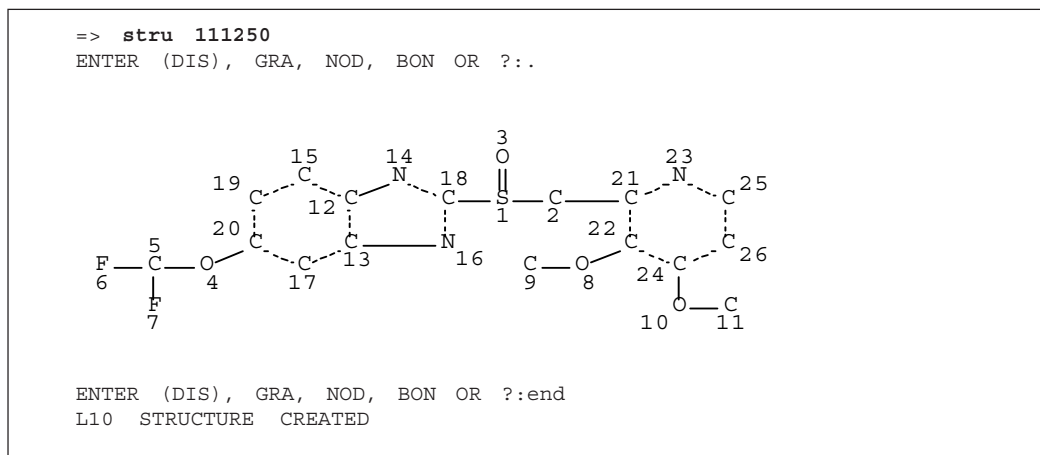
- n Substructure (SSS)
- n Closed Substructure (CSS)
- n Family (FAM)
- n Exact Match (EXA)

Structure search scopes available are:

- n Sample (SAM)
- n Full Substructure Search (FUL)
- n Subset Search (based on answer sets resulting from structure and text searches)
- n Range can be set.

5.1 Structure Modelling and a Simple Start

In DCR either the templates built-in at STN or the structures in the database can be used as templates on the command line. In the following example, an already known chemical compound is called by its DCR number and used as a template:



```

=> s l10 full
FULL SEARCH INITIATED 11:15:14
FULL SCREEN SEARCH COMPLETED -      20 TO ITERATE

100.0% PROCESSED      20 ITERATIONS          19 ANSWERS
SEARCH TIME: 00.00.05

L11      19 SEA SSS FUL L10

```

5.2 Subset Searching

Subset searching is suitable for refining structure searches or when combining text searches with structure searches. If the subset search is based on a previously conducted substructure search the charge incurred is considerably reduced compared to a full substructure search.

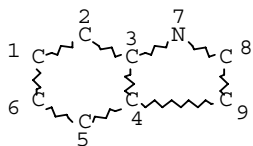
5.2.1 Subset Searching a Text Search Answer Set

Please note that the text search needs to be pointing to the DCR file segment text data otherwise the subset search will not yield any results.

```

=> stru none
ENTER (DIS), GRA, NOD, BON OR ?:gra r65
ENTER (DIS), GRA, NOD, BON OR ?:nod 7 n
ENTER (DIS), GRA, NOD, BON OR ?:.

```



```

ENTER (DIS), GRA, NOD, BON OR ?:end
L1 STRUCTURE CREATED

```

```

=> e a/cc
E#   FILE      FREQUENCY  TERM
--   ----      -
**** START OF FIELD ****
E3   DWPIX     0  ->      A/CC
E4   DWPIX    2863  ALKALOIDS/CC
E5   DWPIX    121   ALLOYS/CC
E6   DWPIX    393   ANTHRACYCLINES/CC
E7   DWPIX    105   ANTIBODIES/CC
E8   DWPIX    1020  BARBITURATES/CC
E9   DWPIX    4802  BENZODIAZEPINES/CC
E10  DWPIX    1358  BETA LACTAMS/CC
E11  DWPIX    24    BORANES/CC
E12  DWPIX    6     CARBOHYDRATE/CC

```

```

=> s e4
L2      2863 ALKALOIDS/CC

```

```

=> s 11 sss full sub=12
FULL SUBSET SEARCH INITIATED 13:35:34
FULL SUBSET SCREEN SEARCH COMPLETED - 1966 TO ITERATE

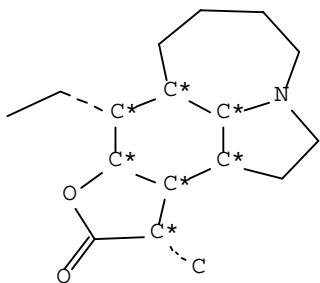
100.0% PROCESSED 1966 ITERATIONS 551 ANSWERS
SEARCH TIME: 00.00.05

L3 551 SEA SUB=L2 SSS FUL L1

=> d 17 all

L3 ANSWER 17 OF 551 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN.S DCR-1195457
DCSE 107809-5-0-0
CN.P NEOSTENINE
SY NEOSTENINE

```



```

MF C17 H27 N O2
SMF C17 H27 N O2 *1; TOTAL *1; TYPE *1
MW 277.4102
SRIN 13070
SDCN RAKC9Q
CC ALKALOIDS

```

5.2.2 Subset Searching a Structure Search Answer Set

```

=> s 14 sss full sub=13
FULL SUBSET SEARCH INITIATED 16:03:00
FULL SUBSET SCREEN SEARCH COMPLETED - 113 TO ITERATE

100.0% PROCESSED 113 ITERATIONS 29 ANSWERS
SEARCH TIME: 00.00.02

L5 29 SEA SUB=L3 SSS FUL L4

=> s 14 sss full sub=13
FULL SUBSET SEARCH INITIATED 16:04:49
FULL SUBSET SCREEN SEARCH COMPLETED - 113 TO ITERATE

100.0% PROCESSED 113 ITERATIONS 29 ANSWERS
SEARCH TIME: 00.00.01

```

L6 29 SEA SUB=L3 SSS FUL L4

=> d 29

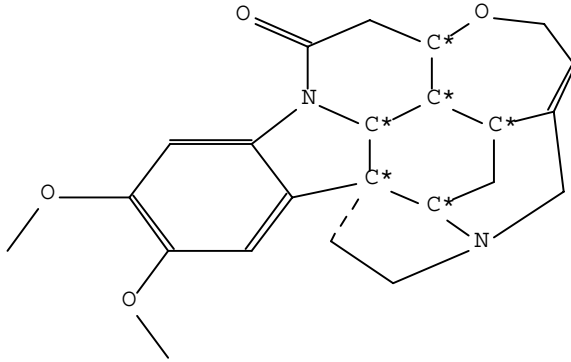
L6 ANSWER 29 OF 29 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN

AN.S DCR-3457

DCSE 3457-1-0-0

CN.P BRUCINE

SY 2,3-DIMETHOXYSTRICHNIDIN-10-ONE; BRUCINE; BRUZIN; CANIRAMIN;
DIMETHOXYSTRYCHNINE; VOMICINUM



MF C23 H26 N2 O4

6 Crossing over into the bibliographic segment

Starting from the results of the Pantoprazol search the corresponding bibliographic documents can be retrieved by re-qualifying with /DCR. The DCR references in the chemical and enhanced polymer coding fields are then searched for the structure identifiers laid down in the answer set of the structure search.

```
=> s l11/dcr
L12      332 L11/DCR

=> d max hitstr

L12 ANSWER 12 OF 332 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN      2006-767310 [78] DWPIX Full-text<<LOGINID::20070302>>
ED      20061204
DNC     C2006-237782 [78]
TI      Use of a proton pump inhibitor e.g. omeprazole, lansoprazole
        in the treatment of sleeping disturbance due to silent
        gastroesophageal reflux
DC      B02
IN      FERNSTROEM P; HASSELGREN G
PA      (ASTR-C) ASTRAZENECA AB
CYC     111
PI      WO 2006118534 A1 20061109 (200678)* EN 22[0]
ADT     WO 2006118534 A1 WO 2006-SE535 20060503
PRAI    US 2005-680932P 20050512
        SE 2005-1041 20050504
IPCI    A61K0031-4164 [I,C]; A61K0031-4184 [I,A]; A61K0031-4427 [I,C];
        A61K0031-4439 [I,A]; A61P0001-00 [I,C]; A61P0001-04 [I,A];
        C07D0235-00 [I,C]; C07D0235-28 [I,A]; C07D0401-00 [I,C];
        C07D0401-12 [I,A]; C07D0401-14 [I,A]; C07D0471-00 [I,C];
        C07D0471-04 [I,A]
AB      WO 2006118534 A1 UPAB: 20061204
        NOVELTY - In the treatment of sleeping disturbance due to
        silent gastroesophageal reflux, a proton pump inhibitor (PPI)
        is administered. ACTIVITY - Hypnotic; Antiinflammatory;
        Gastrointestinal-Gen.. Patients suffering from sleeping
        disturbance due to silent gastroesophageal reflux were
        evaluated. A total of 53 reflux events, which were associated
        with 41 awakenings and 128 arousals were observed. All reflux
        events were associated with either an arousal or awakening or
        both. Subjects with reflux were analyzed pre- and post-
        treatment with omeprazole. After treatment with omeprazole the
        number of awakenings preceded by reflux events decreased from
        3.7 plus minus 0.9 - 1.3 plus minus 0.5. The number of
        arousals preceded by reflux events decreased from 11.6 plus
        minus 3.8 - 1.5 plus minus 0.8 and the total time (pH less
        than 4) decreased from 38.7 plus minus 13.7 - 5.3 plus minus
        1.6 minutes.
        MECHANISM OF ACTION - Proton pump inhibitor; H+ ATPase
        inhibitor; K+ ATPase inhibitor.
        USE - For treating sleeping disturbance due to silent
```

gastroesophageal reflux (claimed).
 ADVANTAGE - The use of proton pump inhibitor improves sleep; reduces risk of developing esophagitis; prevents development of Barrett's esophagus/adenocarcinoma and reduces the use of hypnotics in this group of patients. It also limits the amount of fluid excreted by the stomach, reduces intervariability between patients and shows more effective acid secretion inhibition than therapeutic amounts of other drugs with this effect.

TECH ORGANIC CHEMISTRY - Preferred Compound: The proton pump inhibitor is a substituted benzoimidazole compound of formula (Ia).

ABEX SPECIFIC COMPOUNDS - Use of omeprazole, lansoprazole, pantoprazole, rabeprazole, esomeprazole, tenatoprazole, ilaprazole, leminoprazole their salts and/or enantiomer as the PPI, are specifically claimed.

EXAMPLE - No suitable example is given.

IT UPIT 20061204
 76120-CL 76120-USE; 99135-CL 99135-USE; 111250-CL 111250-USE;
 269446-CL 269446-USE; 109574-CL 109574-USE; 730862-CL 730862-
 USE; 99239-CL 99239-USE; 93863-CL 93863-USE; 1393483-CL
 1393483-USE; 1393484-CL 1393484-USE; 1393485-CL 1393485-USE;
 1393486-CL 1393486-USE

FS CPI
 MC CPI: B05-A01B; B06-D05; B14-D03; B14-E10A; B14-J01B1; B14-L12
 CMC UPB 20061204

M2 *01* C216 D012 D022 D711 F012 F013 F014 F015 F431 H5
 H521 H541 H8 K0 K4 K442 L922 M210 M211 M240 M272
 M282 M311 M321 M342 M373 M391 M412 M431 M511 M521
 M530 M540 M781 M782 P420 P445 P616 P617 P714 M905
 M904
 DCN: R04401-K R04401-M R04401-U
 DCR: 76120-K 76120-M 76120-U

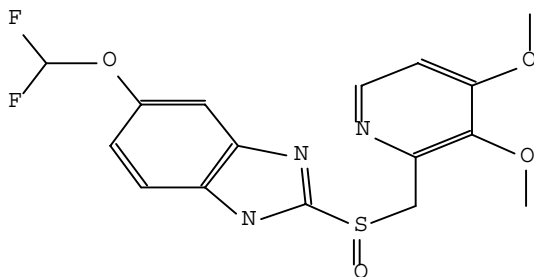
M2 *02* C216 D013 D711 F012 F013 F014 F431 H5 H521 H6 H685
 H8 K0 K4 K442 L922 M210 M211 M240 M281 M311 M312
 M321 M332 M342 M344 M362 M373 M391 M412 M431 M511
 M521 M530 M540 M781 M782 P420 P445 P616 P617 P714
 M905 M904
 DCN: R22683-K R22683-M R22683-U
 DCR: 99135-K 99135-M 99135-U

M2 *03* C216 D012 D022 D711 F012 F013 F014 F431 H5 H522
 H541 H6 H601 H608 H684 H8 K0 K4 K442 L922 M210 M211
 M272 M282 M311 M322 M342 M343 M362 M373 M391 M412
 M431 M511 M521 M530 M540 M781 M782 P420 P445 P616
 P617 P714 M905 M904
 DCN: R22667-K R22667-M R22667-U
 DCR: 111250-K 111250-M 111250-U

M2 *04* C216 D012 D022 D711 F012 F013 F014 F015 F431 H5
 H521 H541 H8 K0 K4 K442 L922 M210 M211 M240 M272
 M282 M311 M321 M342 M373 M391 M412 M431 M511 M521
 M530 M540 M781 M782 P420 P445 P616 P617 P714 M905
 M904
 DCN: RA1IY2-K RA1IY2-M RA1IY2-U
 DCR: 269446-K 269446-M 269446-U

...

AN.S DCR-111250
CN.P PANTOPRAZOLE
CN.S 5-Difluoromethoxy-2-(3,4-dimethoxy-pyridin-2-ylmethane
sulfinyl)-1H-benzimidazole
SDCN R22667



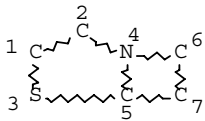
In order to allow for effective cross-over from the chemical repository to the bibliography segment in *DWPI*, Thomson Scientific has equipped DCN and DRN numbers with the corresponding DCR numbers for chemical and enhanced polymer coding. This was part of the 2006 *DWPI* reload, where the backfile for these numbers was populated and run against a DCN and DRN correspondence list.

7 Current Awareness Searching

Structure searching for current awareness purposes can be conducted in the DCR segment of *DWPI* either by setting up an SDI or running your own scripts. Structure SDIs can be set up to deliver the results in hardcopy or softcopy form or as an online answer set delivered to your online account. The latter is recommended if subsequent crossover into the bibliography segment is required. Below the procedure to accomplish this is illustrated. Please note that after crossing over into the bibliography segment you'll probably need to confine the results to those *DWPI* documents that have been updated and which contain a reference to the chemical compound from your structure search result set.

Setting the SDI up:

```
=> filDWPIx
...
=> stru penicl
ENTER (DIS), GRA, NOD, BON OR ?:.



ENTER (DIS), GRA, NOD, BON OR ?:.end
L1 STRUCTURE CREATED

=> sdi
ENTER QUERY L# FOR SDI REQUEST OR (END):11
ENTER SDI REQUEST NAME, (AA001/S), OR END:PENICL/S
ENTER COST CENTER (NONE) OR NONE:.
ENTER TYPE OF SEARCH (SSS), CSS, FAMILY, OR EXACT:.
ENTER TITLE (NONE):Carbapenam SSS
ENTER METHOD OF DELIVERY (OFFLINE), ONLINE, OR EMAIL:online
ELIMINATE PREVIOUSLY SEEN ANSWERS WITH EACH SDI RUN? Y/(N):.
HIGHLIGHT HIT TERMS? (Y)/N:y
ENTER SDI RUN FREQUENCY - WEEKLY, (EVERYUPDATE), MONTHLY, OR ?:.
ENTER SDI EXPIRATION DATE 'YYYYMMDD' OR (NONE):.
QUERY L1 HAS BEEN SAVED AS SDI REQUEST 'PENICL/S'
```

Collecting the results:

=> d sav/a

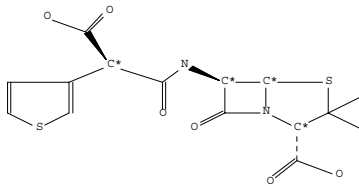
NAME	CREATED	NOTES/TITLE
-----	-----	-----
PENICL20/A	24 MAR 2007	8 ANSWERS IN FILEDWPIX

=> act penicl20/a

TITLE: CARBAPENAM SSS
 L1 STR
 L2 8 SEA FILE=WPIX SSS SDI L1

=> d

L2 ANSWER 1 OF 8 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
 AN.S DCR-108920
 DCSE 108920-1-0-0
 CN.P TICARCILLIN
 CN.S 6-(2-Carboxy-2-thiophen-3-yl-acetylamino)-3,3-dimethyl-7-oxo-4-thia-1-aza-bicyclo[3.2.0]heptane-2-carboxylic acid
 SY AERUGIPEN; TICARCILLIN; TICARPEN; TRIACILLIN



MF C15 H16 N2 O6 S2

Crossing over into the bibliography segment:

=> s 12/dcr

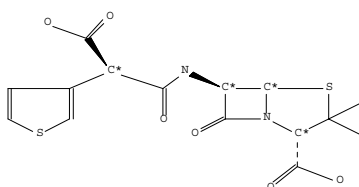
L3 1607 L2/DCR

=> d 1 full hitstr

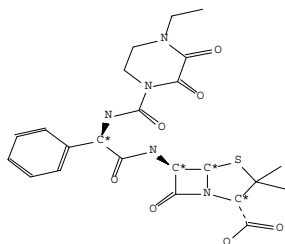
L3 ANSWER 1 OF 1607 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
 AN 2007-200390 [20] DWPIX Full-text<<LOGINID::20070326>>
 DNC C2007-073120 [20]
 TI Biocompatible release system useful for e.g. drug delivery system comprises inorganic component dispersed inside polymer matrix and has lamellar structure with neutralized charge to intercalate within the structure of active principle
 DC A96; B05; B07; C03; C07; D22
 IN BOLOGNESE A; CALIFANO L; CALIGNANO A; COSTANTINO U; MARENZI G; SAMMARTINO G; VITTORIA V
 PA (BOLO-I) BOLOGNESE A; (CALI-I) CALIFANO L; (CALI-I) CALIGNANO A; (COST-I) COSTANTINO U; (MARE-I) MARENZI G; (SAMM-I) SAMMARTINO G; (VITT-I) VITTORIA V
 CYC 113
 PI WO--2007010584 A2 20070125 (200720)* EN 31[1]
 ADT WO--2007010584 A2 2006WO-IT0000556 20060721
 PRAI 2005IT-RM0000393 20050722
 IPCI A61K [,S]

AB WO 2007010584 A2 UPAB: 20070322
 NOVELTY - Biocompatible release system comprises a polymer matrix; an inorganic component dispersed inside the matrix and having a lamellar structure with a neutralized net positive or negative charge able to intercalate within the lamellar structure of a pharmacologically active principle, establishing an ionic type bond with the inorganic component. The combination of the inorganic component and of the active principle constitutes the intercalation compound.
 DETAILED DESCRIPTION - An INDEPENDENT CLAIM is included for preparation of the release system involving treating the inorganic component in such a way as to confer a net positive or negative charge, then combining it with the active principle also in an ionic form or transformed in such a way as to confer an ionic nature, thereby obtaining an intercalation compound which is then mixed with the polymer matrix.

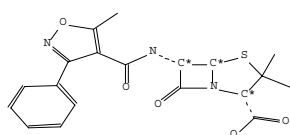
...
 CN.P TICARCILLIN
 CN.S 6-(2-Carboxy-2-thiophen-3-yl-acetylamino)-3,3-dimethyl-7-oxo-4-thia-1-aza-bicyclo[3.2.0]heptane-2-carboxylic acid
 MF C15 H16 N2 O6 S2



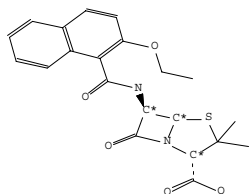
CN.P PIPERACILLIN
 CN.S 6-{2-[4-Ethyl-2,3-dioxo-piperazine-1-carbonyl]-amino}-2-phenyl-acetylamino}-3,3-dimethyl-7-oxo-4-thia-1-aza-bicyclo[3.2.0]heptane-2-carboxylic acid
 MF C23 H27 N5 O7 S



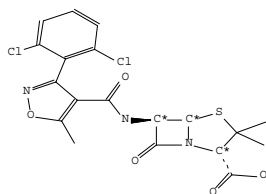
CN.P OXACILLIN
 CN.S 3,3-Dimethyl-6-[(5-methyl-3-phenyl-isoxazole-4-carbonyl)-amino]-7-oxo-4-thia-1-aza-bicyclo[3.2.0]heptane-2-carboxylic acid
 MF C19 H19 N3 O5 S



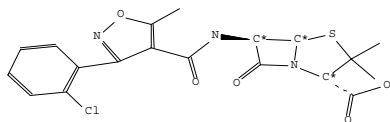
CN.P NAFCILLIN
 CN.S 6-[(2-Ethoxy-naphthalene-1-carbonyl)-amino]-3,3-dimethyl-7-oxo-4-thia-1-aza-bicyclo[3.2.0]heptane-2-carboxylic acid
 MF C21 H22 N2 O5 S



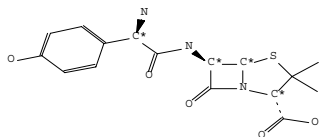
CN.P DICLOXACILLIN
 CN.S 6-{[3-(2,6-Dichloro-phenyl)-5-methyl-isoxazole-4-carbonyl]-amino}-3,3-dimethyl-7-oxo-4-thia-1-aza-bicyclo[3.2.0]heptane-2-carboxylic acid
 MF C19 H17 Cl2 N3 O5 S



CN.P CLOXACILLIN
 CN.S 6-{[3-(2-Chloro-phenyl)-5-methyl-isoxazole-4-carbonyl]-amino}-3,3-dimethyl-7-oxo-4-thia-1-aza-bicyclo[3.2.0]heptane-2-carboxylic acid
 MF C19 H18 Cl N3 O5 S



CN.P AMOXICILLIN
 CN.S 6-[2-Amino-2-(4-hydroxy-phenyl)-acetyl-amino]-3,3-dimethyl-7-oxo-4-thia-1-aza-bicyclo[3.2.0]heptane-2-carboxylic acid
 MF C16 H19 N3 O5 S



8 DCR Search & Display Fields

The set of fields for search, select, sort and display available for DCR is listed below. There aren't any 'mixed' fields where search terms for DCR and the bibliography segment of *DWPI* are available side by side. Hence there aren't any items from DCR indexed in the basic index (/BI) of *DWPI*.

8.1 Field Code

Code		Name
AN.S	SEA/DIS	<i>Chemistry Resource</i> Number, DCR Segment
CC	SEA/DIS	Classification Code
CMF	SEA	Component Molecular Formula
CMF.CNT	SEA	Component Molecular Formula Count
CMT	SEA/DIS	Comment
CN	SEA/DIS	Chemical Name
CN.P	SEA/DIS	Chemical Name Preferred
CN.S	SEA/DIS	Systematic Chemical Name
CNS	SEA	Chemical Name Segment
CT*	SEA/DIS	Controlled Term
CT.DA*	SEA/DIS	Controlled Term Drug Activity
CT.MA*	SEA/DIS	Controlled Term Mechanism
DDRN*	SEA/DIS	Drug Registry Name
DCSE	SEA/DIS	<i>Chemistry Resource</i> Number, DCR Segment
EDCR	SEA/DIS	Entry Date Chemical Repository
ELS	SEA	Element Symbol
ELS.CNT	SEA	Element Symbol Count
FRAGMF	SEA	Fragment Molecular Formula
FRAGMF.CNT	SEA	Fragment Molecular Formula Count
MF	SEA/DIS	Molecular Formula
MPC*	SEA/DIS	Multi Punch Code
MW	SEA/DIS	Molecular Weight
NC	SEA	Number of Components
NFRAG	SEA	Number of Fragments

Code		Name
SCR	SEA/DIS	Structure Cross Reference
SDCN	SEA/DIS	Structure Segment Compound Number
SDRN	SEA/DIS	Structure Segment Registry Number
SMF	SEA/DIS	Standardized Molecular Formula
SRIN	SAE/DIS	Structure Segment Ring Index Number
SS*	SEA/DIS	Substructure Term
STR	DIS	Chemical Structure Display
SY	SEA/DIS	Synonym Name
UPCR	SEA/DIS	Update Date Chemical Repository
UPWX	SEA/DIS	Update Date <i>DWPI</i> Cross Reference

* Sparsely occupied fields

8.2 Predefined Display Formats:

FORMAT

Default format: STD

TRIAL - CN, CN.S, MF, STR

SCAN - CN, CN.S, MF, STR

STD - AN.S, DCSE, CN, CN.S, STR, SCR, CMT, MF

Syn

IDE

ISTD - AN.S, DCSE, CN, CN.S, STR, SCR, CMT, MF

ALL - AN.S, DCSE, CN, CN.S, STR, SCR, CMT, SMF, MW

syn SDCN, SDRN

FULL

IALL - AN.S, DCSE, CN, CN.S, STR, SCR, CMT, SMF, MW, SRIN,

syn SDCN, SDRN

IFULL

MAX - AN.S, DCSE, CN, CN.S, STR, SCR, CMT, MF, SMF, MW, SRIN, SDCN, SDRN, DDRN*, CC, CT, SS, MPC*

IMAX - AN.S, DCSE, CN, CN.S, STR, SCR, CMT, MF, SMF, MW, SRIN, SDCN, SDRN, DDRN*, CC, CT*,
SS, MPC*

8.2.1 Related *DWPI* format

HITSTR - The DCR hit record which led to the retrieval of the bibliographic record.

9 Identifiers

9.1 DCR Number (AN.S)

DCR Numbers are unique compound identifiers that have been assigned by Thomson Scientific on a regular basis from 1999, (reaching back to 1987 for some compounds). DCR

Numbers are found in both segments of WPIDS, WPIX and WPINDEX: in the AN.S field (DCR segment) and in the Indexing Terms (IT), Chemical Coding (M0-6) and Polymer Indexing (PLE) fields (bibliographic segment).

The DCR number can be up to ten-digits (there is scope for longer numbers and new additions). In the primary key index field for the DCR segment (/AN.S) numbers are indexed with a 'DCR-' prefix :

```
=> e 0/an.s
E#      FILE      FREQUENCY  TERM
--      - - - -  - - - - - - - -
****  START OF FIELD  ****
E3      WPIX      0      -> 0/AN.S
E4      WPIX      1      DCR-1/AN.S
E5      WPIX      1      DCR-10/AN.S
E6      WPIX      1      DCR-100/AN.S
E7      WPIX      1      DCR-1000/AN.S
E8      WPIX      1      DCR-10000/AN.S
E9      WPIX      1      DCR-100000/AN.S
E10     WPIX      1      DCR-1000001/AN.S
E11     WPIX      1      DCR-1000002/AN.S
E12     WPIX      1      DCR-1000003/AN.S
```

In the bibliographic segment, DCR Numbers are also indexed along with their appropriate Role Qualifiers if available. Role Qualifiers can be used to refine the search further, if this is required. There are two distinct sets of roles available depending on the field. The (T) proximity operator should be used to link the DCR Numbers to the chosen Role Qualifier in the Indexing Terms (IT) or Chemical Coding (M0-6) fields. Entries in the Polymer Indexing don't carry a role.

All references to DCR reference entries in the bibliography segment of *DWPI* are indexed in a universal search field (/DCR) which can be used to cross over from the structure to the bibliography segment of the *DWPI* file regardless whether the references stem from Indexing Terms, Chemical Coding or Polymer Indexing. This has been improved since the *DWPI* reload in 2006, where backfile indexing of DCR Numbers for DCN and DRN entries. The entries in /DCR have their roles indexed if available, with different role types depending on the origin of the DCR number. Those numbers genuinely indexed in the Indexing Terms field have their DCR type roles, those generated from DCN and DRN numbers take on these roles.

9.1.2 DCR Number Roles

The DCR numbers occur in the index terms (IT) section, in the chemical coding, and in the polymer coding section. The DCR numbers are indexed in /DCR, /IT, /M0-M6 and /PLE index fields.

DCR Roles can be searched on their own, or linked with DCR Numbers in the IT/KW or the /DCR fields. The (T) proximity operator is used to link Roles to DCR Numbers, e.g.

=> S (87874 (T) PRD) /IT; S L2/DCR (T) NEW/IT.

See HELP DCR for further background information.

There are two different versions of roles potentially attached to the DCR numbers: The shortened version in the chemical coding, and the extended version in the index term section.

Both are indexed in /DCR accordingly.

The following DCR roles are available from 1999 onwards in the IT section (/IT).

Role	Definition	Scope Notes
CL	CLAIM	Applied to compounds present in the patent claims (1999-date).
EX	EXAMPLE	Applied to compounds present in the examples, but not in the claims (from update 200253).
DISC	DISCLOSURE	Applied to compounds present in the disclosure, but not in the claims nor in the examples (from update 200253)
NEW	NEW	Substance, process, or apparatus claimed or described as new. (Before 1999 rarely applied.)
PRD	PRODUCED	Production or manufacture of substance or apparatus is claimed or described.
USE	USE	Use of substance or apparatus is claimed or described.
DET	DETECTED	Applied to the keyword for a condition or substance which has been detected as a result of testing.
RCT	REACTANT	Applied to starting materials or products defined in terms of starting materials (1987-date)
RGT	REAGENT	Applied to reaction components apart from starting materials e.g. catalysts, purifying agents (1987-date)
CMP	COMPONENT	Applied to components of a mixture (1987-date)
PUR	PURIFIED	
REM	REMOVED	
TES	TESTED	
ST	SALT	Applied to alkali or alkaline earth metal salts of organic acids; also to certain salts of organic bases e.g. hydro halides, acetates.

A set of single-letter roles is also available for DCR numbers, identical to those available for DWPI Compound Numbers (DCN). (see below).

Here is an example for linking a role with the structure identifier (CL= claimed/):

```
=> s l1/dcr(t)cl/it
      31 L1/DCR
      66622 CL/it
L3    31 L1/DCR(T)CL/it    <- Corresponding bibliographic records
```

9.2 Structured DCR Number (DCSE)

The structured DCR Number (/DCSE) has a logical format, with isomers and salts sharing a common 1-8 digit numerical stem - which can be searched without needing to use truncation.

For example: 3-Methyl-cyclotetradec-5-enone isomers (structured DCR number stem 270633)

```
=> s 270633/dcse
L1    6    270633/DCSE

=> e 270633/dcse
E1    1    270630/DCSE
E2    1    270630-0-0-0/DCSE
E3    6    -> 270633/DCSE
E4    1    270633-1-0-0/DCSE
E5    1    270633-2-0-0/DCSE
E6    1    270633-3-0-0/DCSE
E7    1    270633-4-0-0/DCSE
E8    1    270633-5-0-0/DCSE
E9    1    270633-6-0-0/DCSE
E10   2    270638/DCSE
E11   1    270638-1-0-0/DCSE
E12   1    270638-2-0-0/DCSE

=> s e3
L2    6    270633/DCSE
```

9.2.1 DCR Number Format

The stem of a structured DCR Number is the same for many related compounds (see above), but with suffixes to indicate, e.g. stereochemistry, salts, isotopes and physical forms. The format is as follows:

00000000-00-00-00

00000000	1 to 8-figure sequential number (allows up to 100 million compounds)
00	First suffix for stereo isomers (number from 1-99)
00	Second suffix for salts (number from 1-99)
00	Third suffix to deal with other cases such as physical forms, isotopes, tautomers, etc. (number from 1-99)

Numbers are not filled out with leading "0"s so relevant numbers appear online with a minimum of 4 digits, i.e. 1-0-0-0, etc.

The characters after the first hyphen (the first suffix) are for stereochemistry: 0 the default, used for compounds with no stereo centres or where the stereochemistry is not defined. For any compounds with stereo centres, the next available number is used, i.e. 1 for the first stereoisomer encountered, 2 for the next one, etc up to 99.

The characters after the second hyphen (the second suffix) are for salts: 0 The default, used for the free acid or free base. The next available number is then used for the next salt encountered. The use is restricted to Group I and II or "simple" metal or amine salts of acids, simple (inorganic) salts of bases e.g. halogens etc. N.B. Inorganics and organometallic complexes will have unique identifiers, as will most organic salts consisting of an organic acid and organic base.

The last characters (the third suffix) are for other cases where related compounds might be associated: These include isotopes, tautomers, different physical forms and other cases not covered above. This section is also used if there is a need for a special version of a structure for a particular file or service. Sequential numbers will be assigned whenever different forms appear. 0 is the default, used for the parent compound.

9.3 Structure Cross Reference (SCR syn XCR)

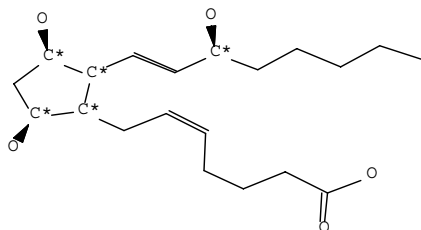
When chemical structures have related structures in DCR the related compounds can be cross referenced. The format is DCR number followed by colon and a descriptor of the relationship, e.g. SEE ALSO or ISOMER. There can be multiple cross references in one DCR record.

```
=> e
E13   WPIX           1           101946 : SEE ALSO/SCR
E14   WPIX           1           10240  : SEE ALSO/SCR
E15   WPIX           1           102739 : SALT PARENT/SCR
E16   WPIX           1           102861 : SEE ALSO/SCR
E17   WPIX           1           103181 : DERIVATIVE OR PARTIAL/SCR
E18   WPIX           1           103245 : SEE ALSO/SCR
E19   WPIX           1           103524 : SEE ALSO/SCR
E20   WPIX           1           103537 : SEE ALSO/SCR
E21   WPIX           1           103743 : ISOMER/SCR
E22   WPIX           1           103781 : SEE ALSO/SCR
E23   WPIX           1           103843 : DERIVATIVE OR PARTIAL/SCR
E24   WPIX           1           103918 : PRECURSOR/SCR

=> s e21
L3     1           "103743 : ISOMER"/SCR

=> d all scr

L3 ANSWER 1 OF 1 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN.S   DCR-111370
DCSE   93389-5-0-0
CN.P   PGF2-ALPHA-EPI-8
CN.S   7-[3,5-Dihydroxy-2-(3-hydroxy-oct-1-enyl)-cyclopentyl]-hept-5-
       enoic acid
SY     PGF2-ALPHA-EPI-8
```

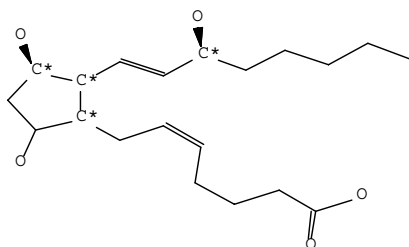


MF C20 H34 O5
 SMF C20 H34 O5 *1; TOTAL *1; TYPE *1
 MW 354.4836
 SDCN RA03RZ
 CC PROSTAGLANDINS
 SCR 103743 : ISOMER

=> s DCR-103743/an.s
 L4 1 DCR-103743/AN.S

=> d all

L4 ANSWER 1 OF 1 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
 AN.S DCR-103743
 DCSE 93389-2-0-0
 CN.P PGF2
 CN.S 7-[3,5-Dihydroxy-2-(3-hydroxy-oct-1-enyl)-cyclopentyl]-hept-5-enoic acid
 SY GLANDIN-N; HORSAFERTIL; PANACELAN-F; PGF2; PROSTAGLAN;
 PROSTAGLANDIN-F2-ALPHA; PROSTAMODIN-F; U-14583



MF C20 H34 O5
 SMF C20 H34 O5 *1; TOTAL *1; TYPE *1
 MW 354.4836
 SDCN RA0CZ6
 CC PROSTAGLANDINS

9.4 Other structure identifiers (SDCN, SDRN, SRIN)

DCR records can also include, where applicable, other (older) compound numbering systems which exist in WPIDS, WPIX and WPINDEX. Three DCR fields are used for these systems: SDCN (Compound Number, DCR segment), SDRN (Registry Number, DCR segment), and SRIN (Ring Index Number, DCR segment). The corresponding fields in the bibliographic segment are DCN (Compound Number), DRN (Registry Number) and RIN (Ring Index Number).

In order to extract DCNs, DRNs and RINs from the DCR segment and then search them in bibliographic segment of WPIDS, WPIX and WPINDEX either the SELECT or TRANSFER commands can be used.

9.4.1 Structure Segment Compound Number (SDCN)

Compound Numbers (DCN) are Merged Markush Service (MMS) Compound Numbers, for specific compound entries in the MMS database on Questel.Orbit. MMS compound number indexing is available in *DWPI* on all hosts from 1987 onwards for patents classified in Sections B (Pharmaceuticals), C (Agrochemicals) and/or E (General Chemicals).

9.4.2 DWPI Compound Number (DCN) Roles

DCN roles are searchable appended to individual *DWPI* Compound Numbers in the DCN field of the bibliographic segment of the *DWPI* file, e.g. S R00708-P/DCN. They can be searched on their own, or linked to an L-numbered answer set with a proximity operator. The following DCN roles are available from 1987 onwards, except as indicated.

Role	Definition/Notes
A	Substance Analysed/Detected
C	Catalyst
D	Detecting Agent
E	Excipient
K	Known Compound
M	Component of a Mixture
N	New Compound
P	Known Compound Produced
Q	Product Defined in Terms of Starting Materials
R	Removing/Purifying Agent
S	Starting Material
T	Therapeutically Active
U	Use of a Single Compound
V	Reagent
X	Substance Removed
Z	Miscellaneous

9.4.3 Structure Segment Registry Number (SDRN)

Over 2,000 commonly-occurring chemicals encountered in the claims and examples of patent specifications in *DWPI* sections B, C, and E have been indexed with unique Registry Numbers since 1981 (*DWPI* Update 198127).

From 1984 (*DWPI* Update 198401) the use of Registry Numbers was extended to cover *DWPI* sections A, D and H; and from update 198407, to the remaining chemical sections F, G, and J-M.

Section A (Plasdoc) has a separate list of Registry Numbers for about 750 compounds (or groups of compounds). Of these, approximately 350 are identical to those used in the other Chemical Patents Index (CPI) sections and have the same numbers. The 400 additional Section A compounds have been allocated numbers in the 5,000 series. These numbers in the 5,000 series were discontinued from *DWPI* Update 199501 on the introduction of the Enhanced Polymer Indexing system.

Registry Numbers are searchable with or without the role letter. Since Registry Numbers are only applied to specific compounds in claims and examples, a search by Registry Number alone does not retrieve unspecified compounds contained within a Markush structure. Registry Numbers do, however, give retrieval of high relevance.

DCR numbers which have been auto generated from the corresponding Registry Numbers are available in the Chemical Coding field.

DWPI Registry Number (DRN) Roles

The Registry Numbers are indexed in the bibliographic segment of *DWPI* with and without the following roles:

Role	Definition/Notes
S	Intermediate or starting material
P	Compound produced
U	Use of a compound (single use or as a mixture)

The roles are indexed on their own as well.

Searching the Registry Numbers field (/DRN) in the bibliographic part of *DWPI* requires the appropriate level of subscription. Please contact Thomson Scientific or STN to determine your subscription level (See Chapter 2 for full contact details).

9.4.4 Structure Segment Ring Index Number (SRIN)

Ring Index Numbers (RIN) are codes assigned to chemical ring systems that are not precisely defined by appropriate *DWPI* Chemical Fragmentation Codes (M0-M6). They are searchable in *DWPI* from 1972 onwards, for patents classified in Sections B (Pharmaceuticals), C (Agrochemicals) and/or E (General Chemicals).

Patents sometimes mention general types of rings rather than specifying the exact ring system involved in an invention e.g. “aryl” or “aromatic heterocyclic ring system”. To enable more specific searches on ring systems, began assigning ring numbers from *The Ring Index* (Patterson, Capell and Walker, 2nd edition, American Chemical Society, and its supplements) to patent indexing records in 1972. These Ring Index Numbers are five digit numbers that appear in the (S)RIN fields of the *Derwent World Patents Index* database. Although the “Patterson Ring Index” is used as a guide, not all of the Ring Index Numbers are used, since Thomson Scientific does not distinguish between levels of unsaturation or different tautomers.

Ring systems encountered in patent documents but not found in the “Patterson Ring Index” are assigned to RINs by Thomson Scientific numbering from 40,000 onwards.

In the same field “Rarer Fragment Numbers” are included. They were used during the period 1972-1975 to describe less common chemical fragments and were given numbers from 70,000 onwards.

Thomson Scientific has now stopped assigning new RINs (update 199901) but continues to apply existing RINs. Ring index numbers are searchable in the bibliographic part of *DWPI* by eligible subscribers only.

10 Formula Fields

10.1 Molecular Formula (MF)

This formula has been calculated from the topological structure data. Molecular formula fragments are separated by dots in this type of Molecular Formula. Individual atoms plus stoichiometric factors are separated by spaces. This molecular formula can also contain words like 'complex'.

```
AN.S   DCR-151227
DCSE   49376-1-1-0
CN.P   PHENOXYMETHYLPENICILLIN POTASSIUM
...
MF     C16 H18 N2 O5 S . K
SMF    C16 H18 N2 O5 S *1; K *1; TOTAL *2; TYPE *2
```

10.2 Standardized Molecular Formula (SMF)

This type of molecular formula was introduced in order to improve the indexing for compounds not adequately searchable by structure searching. Hence it was mainly designed for retrieval of co-ordination compounds and salts, but a structured molecular formula is available for all chemical compounds for consistency.

It is a searchable text field that contains terms corresponding to chemical fragments. Each formula fragment represents the molecular formula of the ion or ligand, arranged according to the Hill standard. Individual fragments are separated by semicolons. Stoichiometry factors are linked to each formula fragment by an asterisk. If the stoichiometry is unknown, the factor is left out. The total number of fragments in the compound is shown as 'TOTAL* #', the total number of different types of fragments is indicated by 'TYPE* #'. Elements within a fragment are separated by spaces. There are no rules for the ordering of formula fragments within the SMF.

```
AN.S   DCR-186734
DCSE   186734-0-0-0
CN.P   COBALT TRIS-ETHYLENEDIAMINETRICHLORIDE
...
MF     C2 H8 N2 . 3 Cl . Co
SMF    C2 H8 N2 *1; Cl *3; Co *1; TOTAL *5; TYPE *3
MW     154.4852
SDCN   R07658
```

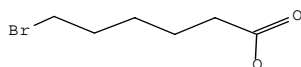
10.3 Component Molecular Formula

Multi-component compounds have the molecular formulae of their contributing fragments or components from the structured molecular formula (SMF) indexed in a separate field.

```
=> s e6
L6      5871  "NA  *1"/CMF
```

```
=> d max
```

```
L6  ANSWER 1 OF 5871 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN.S  DCR-1438098
DCSE  8975-0-1-0
CM    1
      Na
CM    2
```



```
MF      C6 H11 Br O2 . Na
SMF     C6 H11 Br O2 *1; Na *1; TOTAL *2; TYPE *2
MW      218.0472
SDCN    RAPIDK
```

/FRAGMF is a synonym for /CMF and can be used in lieu.

10.4 Element Symbol (ELS) and Element Symbol Count (ELS.CNT)

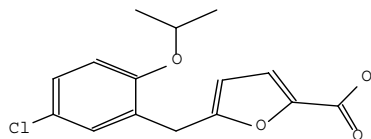
Compounds can be retrieved in *DCR* by searching the element components of the molecular formula. For example (ELS= Element Symbol; Na= Sodium; Cl= Chlorine):

```
=> s (na and cl)/els
      9868      NA/ELS
      184521    CL/ELS
L20   1284      (NA AND CL)/ELS
```

```
=> d scan
```

```
L20 1284 ANSWERS DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
```

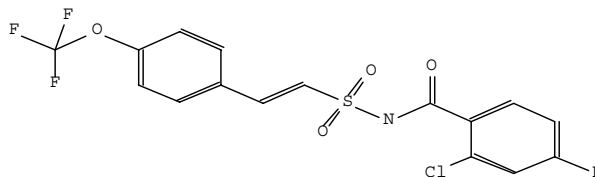
```
MF      C15 H15 Cl O4 . Na
CM      1
      Na
CM      2
```



```
HOW MANY MORE ANSWERS DO YOU WISH TO SCAN? (1):2
```

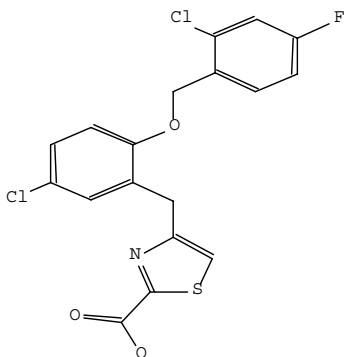
```
L20 1284 ANSWERS DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
```

```
MF      C16 H10 Cl F4 N O4 S . Na
CM      1
      Na
CM      2
```



L20 1284 ANSWERS DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN

MF C18 H12 Cl2 F N O3 S . Na
 CM 1
 Na
 CM 2



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

Elements can also be searched by numerical count within the formula using the Element Symbol Count Field (ELS.CNT). This makes use of the (T) proximity operator (default), and is also range searchable.

For example (ELS.CNT= Element Symbol Count; O= Oxygen)

```
=> s o 2-3/els.cnt
      870298 O/ELS
      637224 2-3/ELS
L21 345323 O 2-3/ELS.CNT
      (O/ELS (T) 2-3/ELS)
```

=> d scan

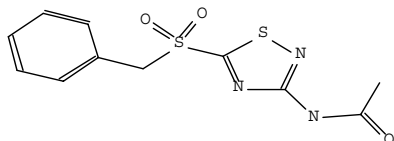
L21 345323 ANSWERS DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN

CN.S 1-(4-CHLORO-PHENYL)-7-[1-(4-FLUORO-PHENYL)-METHYLIDENE]-
 1,4,6,7-TETRAHYDRO-PYRANO[4,3-C]PYRAZOLE-3-CARBOXYLIC ACID
 METHYL-PHENYL-AMIDE
 MF C27 H21 Cl F N3 O2

HOW MANY MORE ANSWERS DO YOU WISH TO SCAN? (1):2

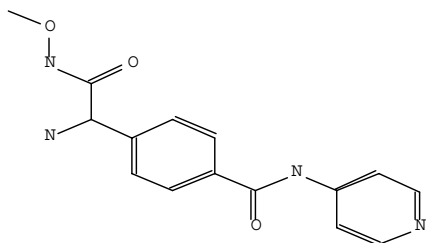
L21 345323 ANSWERS DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN

CN.S N-(5-Phenylmethanesulfonyl-1,2,4-thiadiazol-3-yl)-acetamideN-
(5-Phenylmethanesulfonyl-[1,2,4]thiadiazol-3-yl)-acetamide
MF C11 H11 N3 O3 S2



L21 345323 ANSWERS DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN

CN.S 4-(Amino-methoxycarbonyl-methyl)-N-pyridin-4-yl-benzamide
MF C15 H16 N4 O3



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

10.5 Number of Components (NC)

The number of components in a structured molecular formula can be numerically searched.

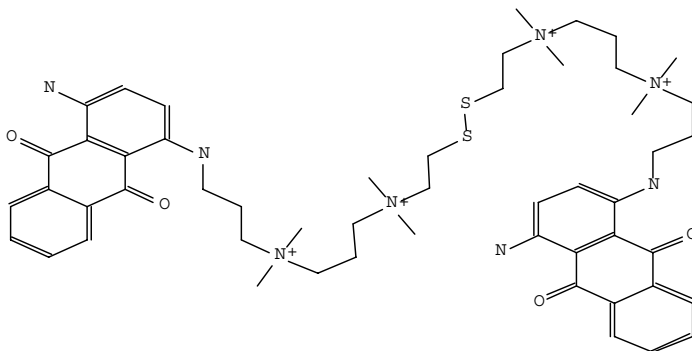
The value is visible in SMF with the heading 'TOTAL'. It will be highlighted there if the value had been searched for.

```
=> e 0/nc
E#      FILE      FREQUENCY      TERM
--      -
**** START OF FIELD ****
E3      DWPIX      0      --> 0/NC
E4      DWPIX      873481     1/NC
E5      DWPIX      52833     2/NC
E6      DWPIX      17121     3/NC
E7      DWPIX      9057      4/NC
E8      DWPIX      6615      5/NC
E9      DWPIX      1616      6/NC
E10     DWPIX      1472      7/NC
E11     DWPIX      696       8/NC
E12     DWPIX      515       9/NC

=> s e8
L22     6615 5/NC
```

=> d max

L22 ANSWER 1 OF 6615 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
 AN.S DCR-1435821
 DCSE 1435821-0-1-0
 CM 1
 Br
 CM 2



CMT 1:4 ratio
 MF 4 Br . C52 H74 N8 O4 S2
 SMF Br *4; C52 H74 N8 O4 S2 *1; TOTAL *5; TYPE *2
 MW 1019.2528
 SRIN 03618
 SDCN RAPH6

10.6 Number of Fragments (NFRAG)

The number of unique fragments in a structured molecular formula can be numerically searched.

The value is visible in SMF with the heading 'TYPE'. It will be highlighted there if the value had been searched for.

```

=> e 0/nfrag
E#   FILE      FREQUENCY   TERM
--   -
**** START OF FIELD ****
E3   DWPIX      0           -> 0/NFRAG
E4   DWPIX     873517      1/NFRAG
E5   DWPIX     72390       2/NFRAG
E6   DWPIX     15036       3/NFRAG
E7   DWPIX     3264        4/NFRAG
E8   DWPIX     701         5/NFRAG
E9   DWPIX     136         6/NFRAG
E10  DWPIX     39          7/NFRAG
E11  DWPIX     12          8/NFRAG
E12  DWPIX     4           9/NFRAG

```

=> e 0/nfrag

E#	FILE	FREQUENCY	TERM
---	---	---	---
****	START OF	FIELD	****
E3	DWPIX	0	-> 0/NFRAG
E4	DWPIX	873517	1/NFRAG
E5	DWPIX	72390	2/NFRAG
E6	DWPIX	15036	3/NFRAG
E7	DWPIX	3264	4/NFRAG
E8	DWPIX	701	5/NFRAG
E9	DWPIX	136	6/NFRAG
E10	DWPIX	39	7/NFRAG
E11	DWPIX	12	8/NFRAG
E12	DWPIX	4	9/NFRAG

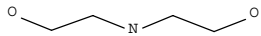
=> s e8

L23 701 5/NFRAG

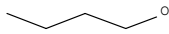
=> d max

L23 ANSWER 1 OF 701 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN

AN.S DCR-1436372
 DCSE 1436372-0-0-0
 CM 1
 Al
 CM 2
 Ti
 CM 3



CM 4



CM 5



CMT 2:1:2:1:1 ratio

MF 2 C4 H11 N O2 . Al . C3 H8 O . 2 C4 H10 O . Ti

SMF Al *1; C3 H8 O *1; C4 H10 O *2; C4 H11 N O2 *2; TOTAL *7;
 TYPE *5; Ti *1

MW 314.2396

SDCN RAPHRB

11 Chemical Name Fields

11.1 Chemical Name (CN)

The CN field provides one step searching for names appearing in both the CN.P and SY fields (see below). Any multiple segment names appearing in this field are searchable and expandable as a single bound phrase. If you wish to search or expand the individual fragments of chemical names the Chemical Name Segment (CNS) field should be used instead.

11.2 Chemical Name Preferred (CN.P)

This is often, but not always, the first name encountered for the compound by Thomson Scientific analysts. It has real no search significance over and above those names which appear in the SY field, so for complete retrieval CN.P should be searched in combination with the SY field, using the CN search field (see above). Names can originate from any product dealing with chemical substances, e.g. the Thomson Scientific *Drug File (DDF)* database (file DRUGU/DDFU).

Consequently names which appear here do not just originate from patent references in *DWPI*. Any multiple segment names appearing in this field are searchable and expandable as a single bound phrase. If you wish to search or expand the individual fragments of chemical names the Chemical Name Segment (CNS) field should be used instead.

11.3 Systematic Chemical Name (CN.S)

This field is populated by many multiple segment systematic names, separated by hyphens and spaces. The names are generated automatically from the structure drawing using Beilstein AUTONOM[®] software. Each systematic name is searchable and expandable as a single bound phrase. If you wish to search or expand the individual fragments of chemical names the Chemical Name Segment (CNS) field should be used instead.

11.4 Chemical Name Segment (CNS)

The CNS field provides one step searching for name segments appearing in the CN.P, SY or CN.S fields. Multiple segment names are searchable and expandable in this field, as the separate name segments. Chemical names are fragmented for this purpose, at all non-alphanumeric characters, e.g. a space or a hyphen. Simultaneous left and right truncation (SLART) can be used to search for name fragments. This can be particularly useful in conjunction with the term operator (T). If you prefer to search or expand chemical names as a bound phrase the Chemical Name (CN) and/or Systematic Chemical Name (CN.S) fields should be used instead.

11.5 Synonym Name (SY)

Synonym names to the preferred name (CN.P), as encountered by editorial staff, are recorded here. Names can originate from any product dealing with chemical substances, e.g. *DDF* (file DRUGU/DDFU). Consequently names which appear here do not just originate from patent references in *DWPI*. Any multiple segment names appearing in this field are searchable and expandable as a single bound phrase. If you wish to search or expand the individual fragments of chemical names the Chemical Name Segment (CNS) field should be used instead.

12 Text Data Fields

12.1 Classification Codes (CC)

Substance Descriptors are keywords that relate to classes of compounds, designed for retrieving groups of substances which are difficult, or impossible, to retrieve using a structure query, e.g. general Alkaloids. They are searched in the Classification Code (/CC) field in WPIDS, WPIX or WPINDEX. An alphabetical list of available Substance Descriptors is given below. These are also available by online by entering HELP SDC at the STN command prompt (=>).

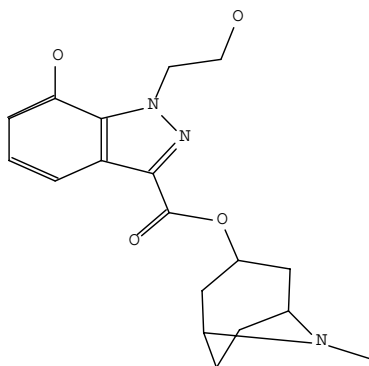
For example searching 'Alkaloids' in the CC field will retrieve all DCR references to alkaloid compounds. Crossover from DCR to DWPI will therefore retrieve all patents which have references to alkaloid compounds.

```
=> s alkaloids/cc
L1      493      ALKALOIDS/CC

=> d scan

L1      493      DWPIDS      COPYRIGHT      2001      INFORMATION      LTD

AN.S      DCR-334564
CN.S      7-Hydroxy-1-(2-hydroxy-ethyl)-1H-indazole-3-carboxylic acid
           8-methyl-8-aza-bicyclo[3.2.1]oct-3-yl ester
MF        C18 H23 N3 O4
```



12.2 Comment (CMT)

This is a free text field containing structure descriptions. This is usually provided if there is no structure available to display.

```
=> e peptidase/cmt
E#      FILE      FREQUENCY    TERM
---      -
E1      DWPIX       8           PEPPER/CMT
E2      DWPIX       1           PEPSIN/CMT
E3      DWPIX       8           -> PEPTIDASE/CMT
E4      DWPIX      232         PEPTIDE/CMT
E5      DWPIX      16          PEPTIDES/CMT
E6      DWPIX       1           PEPTIDO/CMT
E7      DWPIX       2           PEPTIDOLYCAN/CMT
E8      DWPIX       1           PEPTIDOLYCAN/CMT
E9      DWPIX       1           PEPTIDOMIMETIC/CMT
E10     DWPIX       6           PEPTIDYL/CMT
E11     DWPIX       1           PEPTIDYLGLYCINE/CMT
E12     DWPIX       2           PEPTIDYLPROLINE/CMT
```

```
=> s e3
L1      8      PEPTIDASE/CMT
```

```
=> d
```

```
L1 ANSWER 1 OF 8 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN.S   DCR-1231534
DCSE   1231534-1-0-0
CN.P   FURIN
SY     DIBASIC PROCESSING ENZYME; DIBASIC-PROCESSING-ENZYME; FURIN;
      PAIRED BASIC AMINO ACID RESIDUE CLEAVING ENZYME; PAIRED-BASIC-
      AMINO-ACID-RESIDUE-CLEAVING-ENZYME; PROHORMONE CONVERTASE NO
      STRUCTURE DIAGRAM AVAILABLE FOR THIS ACCESSION NUMBER
CMT    Belongs to peptidase family, cleaves paired basic amino acid
      residues.
MF     Unknown
```

12.3 List of available Substance Descriptors

This is the complete list of all controlled term keywords which are searchable in the /CC field.

```
ALKALOIDS
ALLOYS
ANTHRACYCLINES
ANTIBODIES
BARBITURATES
BENZODIAZEPINES
BETALACTAMS
BORANES
CARBOHYDRATES
  glycoproteins
  polysaccharides
  cyclodextrins
```

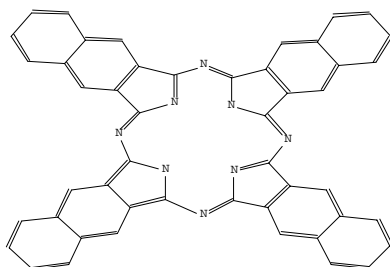
CARBORANES
CROWNETHERS
CYCLIC PEPTIDES *see* PEPTIDES
CYCLODEXTRINS *see* CARBOHYDRATES
DENDRIMERS
ENZYME *see* PROTEINS
FATTY ACID *see also* UNSATURATED FATTY ACIDS
FLAVONOIDS
FULLERENES
GLYCOPROTEINS *see* CARBOHYDRATES and PROTEINS
HALOCARBONS
HETEROFULLERENES
HETEROPOLY ACIDS
LIPOPROTEINS
METALLOCENES
NOBLE GASES
NUCLEOSIDES
NUCLEOTIDES
 oligonucleotides
OLIGONUCLEOTIDE *see* NUCLEOTIDES
OTHER NATURAL PRODUCTS
PEPTIDES
 cyclic peptides
PHOSPHOLIPIDS
POLYMERS
POLYSACCHARIDES *see* CARBOHYDRATES
PROSTAGLANDINS
PROTEINS
 enzymes
 glycoproteins
RETINOIDS
SAPONINS
SILICONES
STEROIDS *see* SAPONINS
TAXANES
TERPENES
TETRACYCLINES
UNSATURATED FATTY ACIDS *see also* FATTY ACIDS
ZEOLITES

13 Structure Display

13.1 Structure (STR)

Topological structures in *DCR* are displayed using the structure display software employed across STN (standard STN structure conventions apply). However, it is important to note that the structures are drawn at Thomson Scientific employing a different set of software and that the coordinates from the connection tables (also provided by Thomson Scientific) are used as the basis for the displays rather than algorithmically calculated coordinates, as with other files on STN. There may be differences therefore in the displays for chemical compounds between *DCR* and other STN structure databases like Beilstein or CAS Registry.

```
AN.S   DCR-134037
DCSE   134037-0-0-0
CN.P   2,3-NAPHTHOPHTHALOCYANINE
SY     2,3-NAPHTHOPHTHALOCYANINE;  NAPHTHALOCYANINE;
      NAPHTHOPHTHALOCYANINE,2,3-;  NAPHTHOPHTHALOCYANINE,2,3-;
      PHTHALOCYANINE,2,3-NAPHTHO-
```



There are limits of sizes fitting on the screen, and if the chemical structure cannot be represented as a topological structure, an error message will be displayed.

```
AN.S   DCR-7659
DCSE   7659-0-0-0
CN.P   BUCKMINSTERFULLERENE
SY     BUCKMINSTERFULLERENE;  BUCKMINSTERFULLERENE C60;  FULLERENE C60;
      FULLERENE-C60
```

NO STRUCTURE DIAGRAM AVAILABLE FOR THIS ACCESSION NUMBER

```
CMT    A carbon sixty fullerene
MF     C60
```


14 Update Dates

14.1 Entry Date Chemical Repository (EDCR)

When a new structure record enters the database it receives a 'time stamp', in this case the entry date.

14.2 Update Date Chemical Repository (UPCR)

Whenever a structure record enters the database or a structure record is amended, an update date is created.

14.3 Update Date *DWPI* Cross Reference (UPWX)

Whenever a *DCR* structure record is referenced in the bibliographic part of *DWPI* the structure record receives a 'time stamp', the *DWPI* cross reference update date.

All three update dates can be different for any given *DCR* structure record, for example:

```
=> d an.s upcr edcr upwx
```

```
L1 ANSWER 1 OF 71 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN  
AN.S DCR-1401361  
UPCR 20070115  
EDCR 20070102  
UPWX 20070119
```

For structure SDIs UPWX should be employed.

15 Supplementary Fields

There are some supplementary fields available which don't fit one of the former categories. Some of them are sparsely populated and hence of limited value, but may be useful on occasion.

15.1 Controlled term (CT)

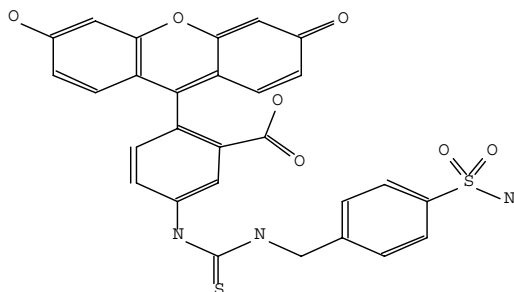
The controlled terms field is the sum of its constituents CR.DA and CT.MA (see below), created for Thomson Scientific's *Drug File (DDF)* database, also available on STN.

```
=> e carbonic/ct
E#      FILE      FREQUENCY    TERM
--      - - - - -
E1      DWPIX      4           CARBOHYDRATE-METABOLISM-STIMULANT/CT
E2      DWPIX      1           CARBOHYDRATE-METABOLISM-STIMULANT./CT
E3      DWPIX      0           -> CARBONIC/CT
E4      DWPIX      1           CARBONIC-ANHYDRASE-I-INHIBITOR/CT
E5      DWPIX      2           CARBONIC-ANHYDRASE-I-INHIBITORS/CT
E6      DWPIX      2           CARBONIC-ANHYDRASE-II-INHIBITOR/CT
E7      DWPIX      2           CARBONIC-ANHYDRASE-II-INHIBITORS/CT
E8      DWPIX      1           CARBONIC-ANHYDRASE-III-INHIBITOR/CT
E9      DWPIX      40          CARBONIC-ANHYDRASE-INHIBITOR/CT
E10     DWPIX      1           CARBONIC-ANHYDRASE-INHIBITOR./CT
E11     DWPIX      39          CARBONIC-ANHYDRASE-INHIBITORS/CT
E12     DWPIX      1           CARBONIC-ANHYDRASE-IX-INHIBITOR/CT
```

```
=> s e9
L4      40      CARBONIC-ANHYDRASE-INHIBITOR/CT
```

```
=> d max ct
```

```
L4      ANSWER 1 OF 40 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN.S    DCR-1122772
DCSE    1122772-0-0-0
CN.S    2-(6-Hydroxy-3-oxo-3H-xanthen-9-yl)-5-[3-(4-sulfamoyl-benzyl)-
        thioureido]-benzoic acid
```



```

MF      C28 H21 N3 O7 S2
SMF    C28 H21 N3 O7 S2 *1; TOTAL *1; TYPE *1
MW      575.6235
SDCN   RAM2LB
CT     CARBONIC-ANHYDRASE-II-INHIBITORS; CARBONIC-ANHYDRASE-
      INHIBITORS; CARBONIC-ANHYDRASE-IX-INHIBITORS
      CARBONIC-ANHYDRASE-II-INHIBITOR; CARBONIC-ANHYDRASE-IX-
      INHIBITOR; CARBONIC-ANHYDRASE-INHIBITOR

```

15.2 Controlled Term, Drug Activity (CT.DA)

This field contains controlled drug activity terms from *DDF*. Since this requires compounds appearing both in *DWPI* and *DDF*, the number of compounds in *DCR* having this field occupied is limited.

```

=> e a/ct.da
E#      FILE          FREQUENCY    TERM
--      -
E1      DWPIX          2          5-HT-7-ANTAGONISTS/CT.DA
E2      DWPIX          8          5-HT-UPTAKE-INHIBITORS/CT.DA
E3      DWPIX          0  ->      A/CT.DA
E4      DWPIX          12         ABL-TYROSINE-KINASE-INHIBITORS/CT.DA
E5      DWPIX          8          ABORTIFACIENTS/CT.DA
E6      DWPIX          2          ABRASIVES/CT.DA
E7      DWPIX          73         ACARICIDES/CT.DA
E8      DWPIX          10         ACAT-INHIBITORS/CT.DA
E9      DWPIX          108        ACE-INHIBITORS/CT.DA
E10     DWPIX          10         ACETYL-COA-CARBOXYLASE-INHIBITORS/CT.DA
E11     DWPIX          5          ACIDIFIERS/CT.DA
E12     DWPIX          16         ACTH-AGONISTS/CT.DA

```

```

=> s e10
L2      10      ACETYL-COA-CARBOXYLASE-INHIBITORS/CT.DA

```

```

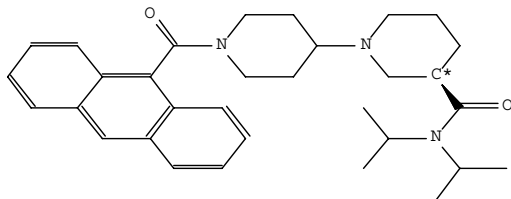
=> d max ct

```

```

L2      ANSWER 1 OF 10 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN.S    DCR-785697
DCSE    785697-1-0-0
CN.P    CP-640188
CN.S    1'-(Anthracene-9-carbonyl)-[1,4']bipiperidinyl-3-carboxylic
      acid diisopropylamide
SY      CP-640188

```



```

MF      C32 H41 N3 O2
SMF    C32 H41 N3 O2 *1; TOTAL *1; TYPE *1
MW      499.7025

```

SDCN RABVON
 CT ACETYL-COA-CARBOXYLASE-INHIBITORS; ANTIARTERIOSCLEROTICS;
 TRIAL-PREP.ACETYL-COA-CARBOXYLASE-INHIBITOR

15.3 Controlled Term, Mechanism of Action (CT.MA)

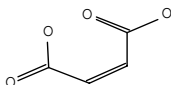
This field contains controlled mechanism of action terms from *DDF*. Since this requires compounds appearing both in *DWPI* and *DDF*, the number of compounds in *DCR* having this field occupied is limited.

```
=> e serotonin/ct.ma
E#      FILE      FREQUENCY  TERM
--      -
E1      DWPIX      1          SEROTININERGIC-1/CT.MA
E2      DWPIX      1          SEROTININERGIC-1D/CT.MA
E3      DWPIX      0          --> SEROTONIN/CT.MA
E4      DWPIX      1          SEROTONIN-1A SEROTONINERGIC
          ACTIVITY./CT.MA
E5      DWPIX      1          SEROTONIN-1A-RECEPTOR-LIGAND/CT.MA
E6      DWPIX      1          SEROTONIN-2-LIGAND/CT.MA
E7      DWPIX      1          SEROTONIN-2B-LIGAND/CT.MA
E8      DWPIX      10         SEROTONIN-ANTAGONIST/CT.MA
E9      DWPIX      1          SEROTONIN-ANTAGONIST./CT.MA
E10     DWPIX      2          SEROTONIN-DEPLETOR/CT.MA
E11     DWPIX      2          SEROTONIN-RECEPTOR PARTIAL-AGONIST./CT.MA
E12     DWPIX      6          SEROTONIN-RECEPTOR-LIGAND/CT.MA
```

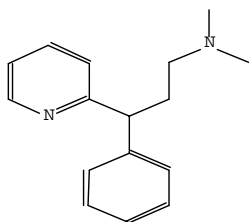
```
=> s e8
L3      10      SEROTONIN-ANTAGONIST/CT.MA
```

```
=> d max ct
```

```
L3      ANSWER 1 OF 10 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN.S    DCR-151191
DCSE    103826-0-1-0
CN.P    PHENIRAMINE MALEATE
CN.S    Dimethyl-(3-phenyl-3-pyridin-2-yl-propyl)-amine; compound with
          but-2-enedioic acid
SY      ALLER-G; ALTERGIAN; ANTIHISTONE; AVIL; AVIL-RETARD; DANERAL;
          DANERAL-SA; FENAMINE; FENAMINE-SLOW; HEMARIL; INHISTON; LARIL;
          MALEATE-PHENIRAMINE; MEDOPHEN; METRON; PHENIL; PHENIRAMIN;
          PHENIRAMINE MALEATE; PHENIRAMINE-MALEATE; PHYLLAXENE; PIRIEX;
          PM-241; QUIL; S-108; TRIMETON; TRIPOTON
CM      1
```



CM 2



MF C4 H4 O4 . C16 H20 N2
 SMF C16 H20 N2 *1; TOTAL *2; TYPE *2; C4 H4 O4 *1
 MW 356.4253
 SDCN R17806
 CT ANTIHISTAMINES-H1
 Antihistamine-H1; serotonin-antagonist; enhances effects of
 adrenaline.

15.4 Drug Registry Name (DDRN)

If a cross-reference is available between *DDF* and *DWPI*, it will be located in the DDRN (Drug Registry Name) field. This can be used to cross-over between both files.

```
=> e dr0121037/ddrn
E#   FILE      FREQUENCY   TERM
--   ----      -
E1   DWPIX      1           DR0120945/DDRN
E2   DWPIX      1           DR0121029/DDRN
E3   DWPIX      1           --> DR0121037/DDRN
E4   DWPIX      1           DR0121039/DDRN
E5   DWPIX      1           DR0121051/DDRN
E6   DWPIX      1           DR0121052/DDRN
E7   DWPIX      1           DR0121054/DDRN
E8   DWPIX      1           DR0121056/DDRN
E9   DWPIX      1           DR0121059/DDRN
E10  DWPIX      1           DR0121062/DDRN
E11  DWPIX      1           DR0121063/DDRN
E12  DWPIX      1           DR0121064/DDRN
```

```
=> d ddrn
DDRN DR0121037
```

A corresponding DRUGU record:

```
AN 2004-10239 DRUGU B P
TI Isozyme-nonspecific N-substituted bipiperidylcarboxamide
acetyl-CoA carboxylase inhibitors reduce tissue malonyl-CoA
concentrations, inhibit fatty acid synthesis, and increase
fatty acid oxidation in cultured cells and in experimental
animals.
AU Harwood H J Jr; Petras S F; Shelly L D; Zaccaro L M;
Perry D A; Makowski M R; Hargrove D M; Martin K A; Tracey W R;
Chapman J G
CS Pfizer
LO Groton, Conn., USA
SO J.Biol.Chem. (278, No. 39, 37099-111, 2003) 6 Fig. 3 Tab. 56
```

Ref. CODEN: JBCHA3 ISSN: 0021-9258

AV Dept. of Cardiovascular + Metabolic Diseases, Pfizer Global
Research + Development, Groton Labs., Pfizer, Inc., Groton, CT
06340, U.S.A. (16 authors; e-mail:
h_james_harwood@groton.pfizer.com).

LA English
DT Journal
AB The effects of acetyl-CoA carboxylase (ACC)1 and ACC2
inhibition by CP-640186, CP-640188 and CP-610431 were studied.
The pharmacokinetics of CP-640-188 was determined in rats. The
ACC inhibitors inhibited fatty acid synthesis and increased
fatty acid oxidation in liver, adipose, heart and muscle
tissue in-vitro and in rats and mice in-vivo after i.p. and
p.o. administration. The results suggest that isozyme-non-
selective inhibition may reduce risk factors associated with
metabolic syndrome.

SH B Biochemistry
P Pharmacology
CC 8 Pharmacokinetics
22 Endogenous Compounds
58 Vasoactive
72 New Drugs
73 Trial Preparations
CT OBESITY *OC; BODY-WEIGHT *OC; CL-316243 *RC; RAT *FT;
MOUSE *FT; IN-VIVO *FT; HEP-G2-CELL *FT; LIVER *FT; HEART *FT;
MUSCLE *FT; ADIPOSE-TISSUE *FT; IN-VITRO *FT; LIPID-METAB.
*FT; DRUG-COMPARISON *FT; I.P. *FT; P.O. *FT;
ANTIARTERIOSCLEROTIC *FT; ACETYL-COA-CARBOXYLASE-INHIBITOR
*FT; TRIAL-PREP. *FT; ACETYL-COA-CARBOXYLASE-INHIBITORS *FT;
ANTIARTERIOSCLEROTICS *FT; NEW *FT; LAB.ANIMAL *FT;
HEPATOBLASTOMA *FT; TUMOR-CELL *FT; TISSUE-CULTURE *FT;
INJECTION *FT
[01] CP-640186 *PH; CP-640186 *DM; DR0121039 *RN; I.V. *FT;
PHARMACOKINETICS *FT; INJECTION *FT; PH *FT; DM *FT
[02] CP-610431 *PH; DR0121035 *RN; PH *FT
[03] CP-640188 *PH; DR0121037 *RN; PH *FT
FA AB; LA; CT
FS Literature

15.5 Multipunch Code (MPC)

This field contains multipunch codes from *DDF*. Since this requires compounds appearing both in *DWPI* and *DDF*, the number of compounds in *DCR* having this field occupied is limited. In addition, multipunch codes have been discontinued in *DDF*.

```
=> s e235
L9      1      "PGG2 *PI"/MPC

=> d hit

L9 ANSWER 1 OF 1 DRUGU COPYRIGHT 2007 THE THOMSON CORP on STN
MPC      [02] PGG2 *PI; PGH2 *PI

=> d mpc

L9 ANSWER 1 OF 1 DRUGU COPYRIGHT 2007 THE THOMSON CORP on STN
MPC      [02] PGG2 *PI; PGH2 *PI
```

15.6 Molecular Weight (MW)

A fully range searchable molecular weight field is available, which may prove useful, e.g. in refining large answer sets retrieved using the Element Symbol (ELS) or Element Symbol Count (ELS.CNT) fields.

For example (MW<100= Molecular Weight of less than 100; ELS= Element Symbol; K= Potassium):

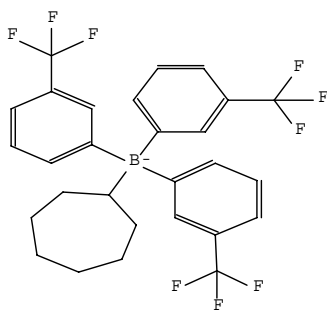
```
=> s mw<100 and k/els
      1604 MW<100
      308 K/ELS
L3    36 MW<100 AND K/ELS
```

```
=> d tri mw 1-3
L3 ANSWER 5 OF 36 DWPIDS COPYRIGHT 2001 INFORMATION LTD
AN.S DCR-208582
MF C28 H25 B F9 . C4 H12 N
CM 1
```



```
CM 2
MW 56.1049
```

```
L3 ANSWER 6 OF 36 DWPIDS COPYRIGHT 2001 INFORMATION LTD
AN.S DCR-133148
CN.P POTASSIUM PEROXIDE
MF K . O2
CM 1
```



```
CM 2
MW 71.096
```

```
L3 ANSWER 7 OF 36 DWPIDS COPYRIGHT 2001 INFORMATION LTD
AN.S DCR-131846
CN.P POTASSIUM SULFIDE update if required.
MF H K . H2 S
CM 1
```



```
CM 2
MW 71.164
```

15.7 Substructure Search Terms (SS)

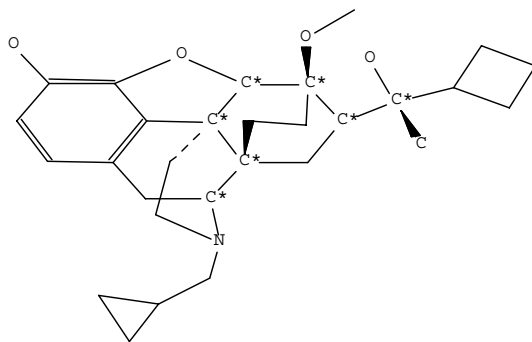
This field contains substructure search terms from *DDF*. Since this requires compounds appearing both in *DWPI* and *DDF*, the number of compounds in *DCR* having this field occupied is limited.

```
=> e alkaloid/ss
E#      FILE      FREQUENCY    TERM
--      -
E1      DWPIX      239          ALDEHYDE/SS
E2      DWPIX      137          ALDIMINE/SS
E3      DWPIX      952          -> ALKALOID/SS
E4      DWPIX      28           ALKANE/SS
E5      DWPIX      121          ALKYL BROMIDE/SS
E6      DWPIX      452          ALKYL CHLORIDE/SS
E7      DWPIX      1066         ALKYL FLUORIDE/SS
E8      DWPIX      36           ALKYL IODIDE/SS
E9      DWPIX      2            ALUMINUM/SS
E10     DWPIX      12           ALUMINUM-COMPLEX/SS
E11     DWPIX      22           ALUMINUM-SALT/SS
E12     DWPIX      1            AMERICIUM/SS
```

```
=> s e3
L2      952      ALKALOID/SS
```

```
=> d max ss
```

```
L2      ANSWER 1 OF 952 DWPIX COPYRIGHT 2007 THE THOMSON CORP on STN
AN.S    DCR-1271357
DCSE    1271357-1-0-0
```

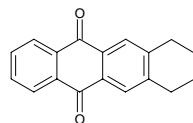


```
MF      C29 H39 N O4
SMF     C29 H39 N O4 *1; TOTAL *1; TYPE *1
MW      465.6385
SRIN    06766
SDCN    RALXWI
CC      ALKALOIDS
SS      ALKALOID; BRIDGE-STRUCT.; COND.RING; CYCLOHEXANE; MORPHINAN;
        PHENOL; ETHER; ISOBENZOFURAN; CYCLOPROPANE; CYCLOBUTANE;
        BENZOFURAN; AMINOALCOHOL
```

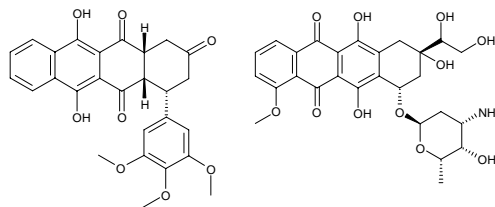

16 Appendix

16.1 Definitions of Substance Descriptors/ Classification Codes

SUBSTANCE DESCRIPTOR	DESCRIPTION
ALKALOIDS	Organic nitrogen-containing bases, mainly of plant origin. This descriptor is only used when identified as such in the source document. Examples are morphine, caffeine, atropine, and strychnine.
ALLOYS	A metal that consists of an intimate mixture of two or more metallic elements.
ANTHRACYCLINES	A class of compounds containing the following ring system, the degree of saturation and substitution can vary.



e.g.

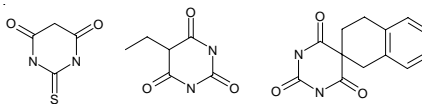


ANTIBODIES

A blood serum protein of the globulin fraction which is formed in response to the introduction of an antigen. Only used when identified as such in the source document.

BARBITURATES

Used for ALL derivatives of barbituric acid, including thio analogues. e.g.

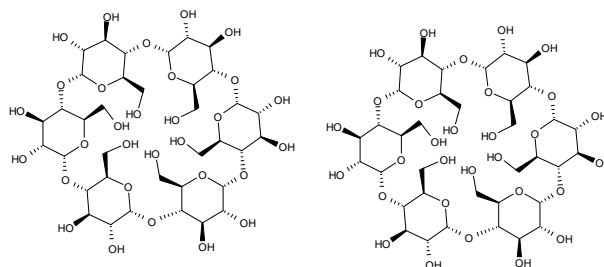


SUBSTANCE DESCRIPTOR	DESCRIPTION
BENZODIAZEPINES	<p>Used when a benzene ring is condensed to a 7 membered ring containing 2 nitrogen atoms (in any position), other atoms in the ring being carbon. Regardless of the degree of saturation or substitution. e.g.</p> <div data-bbox="712 451 1208 638" style="text-align: center;"> </div>
BETA LACTAMS	<p>Keyword applied to compounds containing the beta lactam group condensed to thiazine or thiazole ring i.e. cephalosporins, penicillins, regardless of the degree of saturation or substitution. Basic ring structure shown.</p> <div data-bbox="712 820 816 886" style="text-align: center;"> </div> <p>e.g.</p> <div data-bbox="712 975 1097 1120" style="text-align: center;"> </div>
BORANES	<p>Group of compounds that contain boron and hydrogen only. The simplest example is diborane B₂H₆. The larger borane molecules have open or closed polyhedra of boron atoms</p>
CARBOHYDRATES	<p>Polyhydroxyaldehydes (or polyhydroxyketones) or substances that yield these on hydrolysis. The general molecular formula of carbohydrates is C_x(H₂O)_y.</p> <p>Any compound containing a sugar moiety is assigned the keyword carbohydrate, the definition for a sugar sets the lower limit of size, such that compounds must contain at least 2 stereocentres. Therefore glycoaldehyde (HOCH₂CHO) and glyceraldehyde (HOCH₂CHOHCHO) are both excluded because neither contains 2 stereocentres.</p>
polysaccharides	<p>A polysaccharide is a compound which contains at least 5 adjacent sugar residues (or their derivatives) linked via ether or thioether linkages. The term carbohydrate is also used.</p>

SUBSTANCE DESCRIPTOR **DESCRIPTION****cyclodextrins**

Cyclic oligomers of glucose in which the individual glucose units are connected by 1,4-bonds. The terms carbohydrate and polysaccharide are also used.

e.g.

**glycoproteins**

Any protein with carbohydrate group attached. The terms protein and carbohydrate are also used.

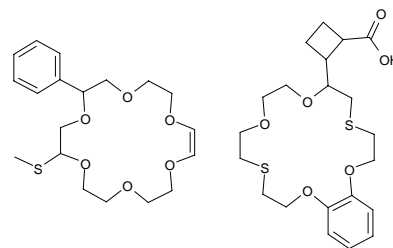
CARBORANES

Boron cluster compounds with one or more of the polyhedral vertices replaced by carbon, e.g. C₂H₁₂B₁₀

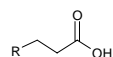
CROWN ETHERS

Macrocyclic compounds with O or S hetero atoms as the donor atoms in their ring structure and having the property of incorporating cations into their cavities. The best known crown ethers are the macrocyclic polyethers containing the repeating unit $(-OCR_2CR_2)_n$, where R is most commonly H and are named in the form: x-crown-y, where x is the total number of atoms in the ring and y is the number of oxygens.

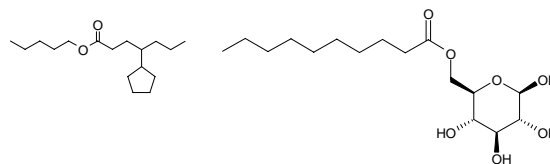
e.g.



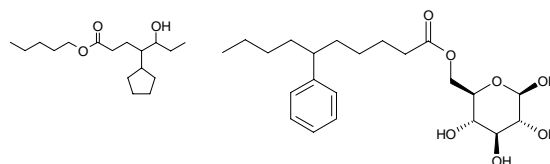
SUBSTANCE DESCRIPTOR	DESCRIPTION
DENDRIMERS	<p>Globular structures in which well-defined branches radiate from a central core, becoming more branched and crowded as they extend out to the periphery. Some dendrimers have a diameter of more than 10 nm and a molecular weight exceeding 1 million Daltons.</p> <p>The second type of dendritic structure is the hyperbranched polymer. This type of polymer also has a fractal pattern of chemical bonds, but its branches don't emanate from a central core. Hyperbranched polymers can have either random or fairly regular architectures.</p> <p>The term also applies to organometallics with dendrimer ligands.</p>
FATTY ACIDS	<p>Any straight- or branched-chain, unsubstituted, saturated monocarboxylic acid with a total of 3 or more C atoms, includes derivatives such as esters and amides, and includes analogues with the cycloalkyl substituents in the chain. For unsaturated use UNSATURATED FATTY ACID</p>



e.g.

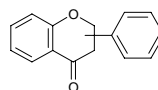


The keyword is NOT applied to the following due to the substitution on the alkyl chain (only cycloalkyl substituents are allowed) :-

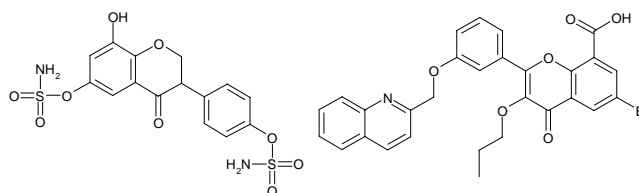


SUBSTANCE DESCRIPTOR**DESCRIPTION****FLAVONOIDS**

Compounds containing a benzopyran ring substituted at C-2 or C-3 by an aryl group, the degree of saturation and substitution can vary, basic structure shown.



e.g.

**FULLERENES**

Giant closed-cage molecules that are formed entirely of carbon in the sp² hybridised state and are arranged to form adjoining pentagonal and hexagonal rings.

Number of C atoms = 2(10 + m) with 12 pentagonal rings and m hexagonal rings

Number of rings = 12 + (n - 20)/2 where n = number of C atoms

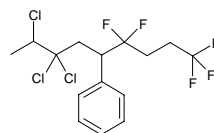
The nanotubes are very large tubular fullerenes and are not considered as a separate class of molecule. The tubular shape is the result of the large number carbon atoms which form hexagonal rings. The tube is sealed at each end due to the presence of pentagonal rings.

See also HETEROFULLERENE

HALOCARBONS

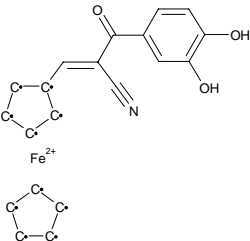
A compound containing a carbon skeleton which is poly-substituted with, one or more halogens, no other heteroatoms are present.

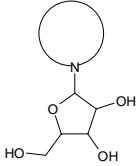
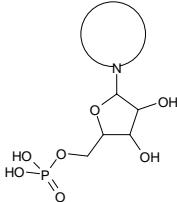
e.g.

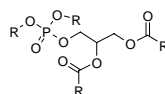
**HETEROFULLERENES**

Fullerenes where one or more carbon atoms have been replaced by another atom.

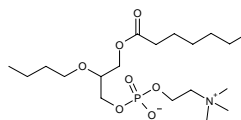
See also FULLERENE

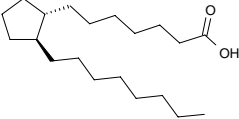
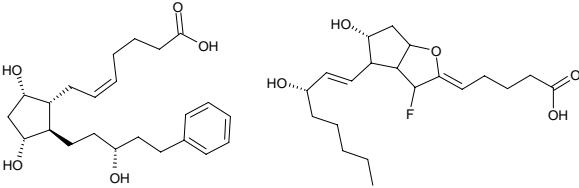
SUBSTANCE DESCRIPTOR	DESCRIPTION
HETEROPOLY ACIDS	<p>Definition: These are compounds that satisfy the following formula:</p> $\text{H}_x\text{A}_y\text{M}_z\text{O}_w$ <p>where:</p> <p>A = phosphorous, silicon, boron or arsenic M = transition metal (normally molybdenum, vanadium or tungsten)</p> <p>x = > 0 y = > 0 z = > 0</p> <p>This substance descriptor also includes salts of the acids in which some or all of the hydrogen atoms are replaced by cations, most commonly ammonium or alkali metal cations. It is also possible to have structures in which some of atoms M are replaced by a second transition metal (niobium being the most common one) so that the heteropolyacid contains two metals plus the metalloid A.</p> <p>Examples of heteropolyacids are: H₃PW₁₂O₄₀, H₄PWW₁₁O₄₀, Na₉PW₁₅Nb₃O₆₂</p>
LIPOPROTEINS	<p>Any compound containing a protein and a lipid moiety. This descriptor is only used when identified as such in the source document.</p>
METALLOCENES	<p>An organometallic compound that contains at least one cyclopentadienyl group, or its derivative, bonded to the central metal atom. Derivatives of the cyclopentadiene ligand which are also included within this definition are those with rings fused onto the cyclopentadienyl ring e.g. indene and fluorene.</p> <p>e.g.</p> 

SUBSTANCE DESCRIPTOR	DESCRIPTION
NUCLEOSIDES	<p>A nucleoside is a compound which contains a sugar residue attached via N to a cyclic base group. The base group is usually derived from purine or pyrimidine groups, or their ring modified derivatives including the thia derivatives. The more usual base groups are adenine, cytosine, thymine, uracil, and guanine residues.</p> <p>Below is the basic structure for a nucleoside system, the sugar moiety can be substituted and the keyword is still applied for deoxy/dideoxy analogues.</p>
	
NUCLEOTIDES	<p>A nucleoside with a phosphate group attached to the sugar moiety.</p>
	
oligonucleotides	<p>Compounds containing 3 or more nucleotide residues which are linked via the phosphate groups. Usually denoted with single letter codes representing the nucleoside bases e.g. TTUUGGCATU</p>
OTHER NAURAL PRODUCTS	<p>Natural Products not covered by any other substance descriptor, e.g plant extracts"</p>
PEPTIDES	<p>A compound formed by the linking of two or more amino acids by CO-NH groups. For peptides containing fifty or more residues use the term PROTEIN instead.</p>
cyclic peptides	<p>As for peptide, but part or all of the peptide chain forms a ring</p>
PHOSPHOLIPIDS	<p>Esters of fatty acids formed with alcohol components containing a phosphate group.</p>

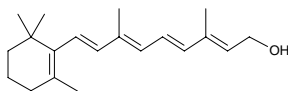


e.g.

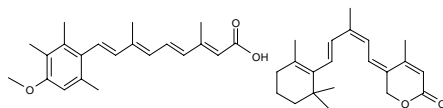


SUBSTANCE DESCRIPTOR	DESCRIPTION
POLYMERS	A macromolecule with five or more structural repeat units.
PROSTAGLANDINS	Compounds that are derived from 20-carbon unsaturated carboxylic acids with a cyclopentane ring i.e. analogues of prostanoic acid. The degree of saturation and substitution can vary.
	<p data-bbox="719 499 812 513">Prostanoic acid</p> 
	e.g.
	
PROTEINS	Peptides with a specific sequence of 50 or more residues. The term PEPTIDE is not also applied.
enzymes	Any of a large class of protein substances produced by living cells, which act as biocatalysts in biochemical reactions. Enzymes are typically composed of a protein part (the apoenzyme) and a non-protein part (the coenzyme) necessary for activity. The term protein is also used.
	This descriptor is only used when identified as such in the source document or if the enzyme name is obviously identifiable i.e. name ending in -ase.
glycoproteins	Any protein with carbohydrate group attached. The terms protein and carbohydrate are also used.
RETINOIDS	Synthetic analogues of vitamin A. The keyword is applied regardless of the degree of saturation and substitution.

Vitamin-A



e.g.

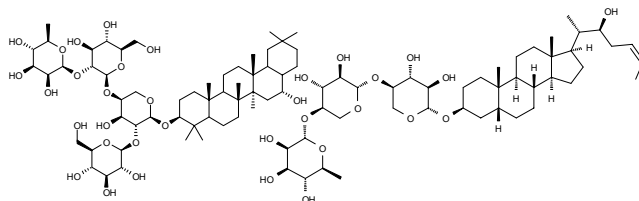


SUBSTANCE DESCRIPTOR

DESCRIPTION

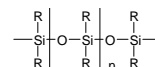
SAPONINS

Saponins are glycosides of steroids and triterpenes and are widely distributed in plants and some marine organisms
 e.g.



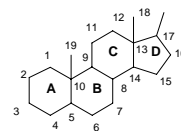
SILICONES

Any of a large group of siloxane polymers based on a structure consisting of alternate silicon and oxygen atoms with various organic radicals attached to the silicon.

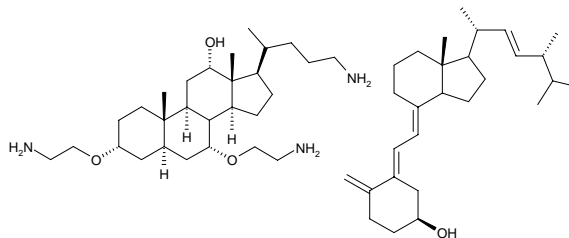


STERIODS

A compound is considered a steroid if it contains a cyclopenta(a)phenanthrene ring, term is also applied for ring broken analogues.

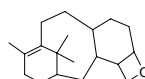


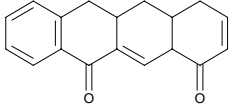
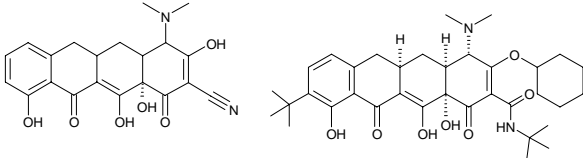
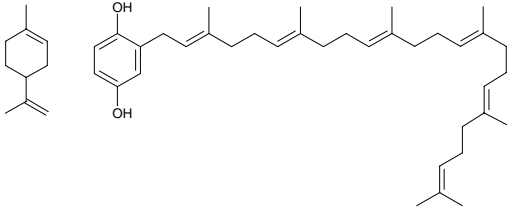
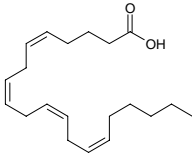
e.g.



TAXANES

Any compound containing the below basic ring system, the degree of saturation and substitution can vary.



SUBSTANCE DESCRIPTOR	DESCRIPTION
TETRACYCLINES	<p>A class of compounds containing the below ring system, the degree of saturation and substitution can vary. Typically you will find the configuration of the double bonds as show below:-</p>
	
	e.g.
	
TERPENES	<p>A group of unsaturated hydrocarbons present in plants, consisting of isoprene units $\text{CH}_2:\text{C}(\text{CH}_3)\text{CH}:\text{CH}_2$. Monoterpenes have 2 units, sesquiterpenes have 3 units, diterpenes have 4 units etc.</p>
	e.g.
	
UNSATURATED FATTY ACIDS	<p>As for FATTY ACID except one or more double and/or triple bonds are to be found in the chain.</p>
	e.g.
	
ZEOLITES	<p>An (alumino)silicate or analogous structure of the porous tectosilicate (infinite 3-dimensional network) type that facilitates the exchange of anions and/or cations and the sorption/desorption of water molecules. Must also satisfy the formula of an extended structure.</p>

17 Frequently Asked Questions

Content

1. How many compounds will be indexed from a given patent document? Is there really a limit of 25 compounds?

Initially, Thomson Scientific's aim was to build its experience in applying this type of structural indexing. The initial guidelines provided to Thomson Scientific analysts contained the following regarding number limits:

The analyst, where possible, should index up to 25 Compounds for new compounds, and 50 Compounds which have already been registered and are being cited. As Thomson Scientific's experience has grown and customer feedback has been received, a set of guidelines has been devised to meet both customer and internal requirements. As part of these revised guidelines, these number limits have been phased out.

2. What are the detailed selection procedures for selecting these compounds?

Thomson Scientific believes the selection criteria for the input of structural information into the *Chemistry Resource* are of great importance. The selection guidelines that analysts follow are summarised below.

- 1 Cover all claimed compounds.
- 2 Cover at least the main (best) example. If there are few claimed compounds, more examples will be selected.
- 3 For compounds outside the claims, only real (i.e. those with good supporting physical or biological data) will be selected.
- 4 Where compounds are selected from a range of examples, those which best illustrate the structural diversity of patent coverage will be selected.

Supporting these guidelines are new effective quality checks which will enable Thomson Scientific to continue providing high quality indexing.

3. How is the Systematic Name field generated? Why is the systematic name field sometimes empty?

The CN.S field is populated by running *Chemistry Resource* records through the Beilstein AUTONOM software. At the time Thomson Scientific was creating *Chemistry Resource*, several naming programs were evaluated and AUTONOM gave the best results. AUTONOM is not 100% successful in generation of names, especially for organometallic, inorganics and some other complex structures. Consequently, some entries especially from the back file have

gone online without systematic names but this number will be minimised over time and, as these records are cited in the future, fill in the gaps over time.

4. What are Substance Descriptors and how are they applied and used?

The Substance Descriptors are keywords that relate to classes of compounds. They are applied intellectually by Thomson Scientific analysts and are designed for retrieving groups of substances which are difficult, or impossible, to retrieve using a single structure query, e.g. general Alkaloids. The Substance Descriptors are stored in the CC field of the *Chemistry Resource* segment of *DWPI*.

For example searching 'Beta Lactams' in the CC field will retrieve all *Chemistry Resource* references to beta lactam compounds. Crossover from *Chemistry Resource* to *Derwent World Patents Index* will therefore retrieve all patents which have references to beta lactam compounds, including those for which no specific beta lactam was specified. (N.B. - when no specific examples or Markush formula are given or when the structure is not fully defined, the concept is indexed using only the Substance Descriptor. Searching for BETA LACTAMS/CC can retrieve additional references to beta lactams, that a straightforward structure search would not).

Index

B

- Bibliographic segment 15
- Bibliographic Segments 20

C

- Chemical Name Fields
 - Chemical Name Segment 42
 - Synonym Name 42
 - Systematic Chemical Name 41
- Classification Codes 43, 59
- Comment 44
- Component Molecular Formula 36
- Controlled term 51
 - Drug Activity 52
 - Mechanism of Action 53
- Crossing over to bibliography segment 20
- Current Awareness Searching 19
- Customer Technical Support 3

D

- DCR Number 27
- DCR Number Format 29
- DCR Number Roles 28
- DCSE 9
- Definitions 59
 - Substance Descriptors 59
- Display Fields 23
- Drug Registry Name 54
- DWPI Compound Number Roles
 - DCN 32
- DWPI Registry Number Roles
 - DRN 33

E

- Element Symbol
 - ELS 36
- Element Symbol Count
 - ELS CNT 36

F

- Formula Fields 35
- Frequently Asked Questions 69

M

- Molecular Formula
 - MF 35
- Molecular Weight 56
- Multipunch Code 56

N

- Number Format 29
- Number of Components 38
- Number of Fragments
 - NFRAG 39

P

- Predefined Display Formats 24

R

- Ring Index Numbers 34
- Role Qualifiers 27

S

- Sample Record 9
- SDI
 - Setting up 19
- Search Fields 23
- Searching
 - Current Awareness 19

Setting up the SDI 19
Standardized Molecular Formula
SMF 35
Structure Cross Reference
SCR 30
Structure Display 47
Structure identifiers 32
Structure Modelling 11
Structure searching 11
Structure Segment Compound
Number 32
SDCN 32
Structure Segment Registry Number
SDRN 33
Structure Segment Ring Index Number
SRIN 34
Structured DCR Number
DCSE 29
Subset searching 12
Substance Descriptors 44, 59
Substructure Search Terms 57
Supplementary Fields 51

T

Text Data Fields 43

U

Update Dates 49

W

WPIDS 43

