TRANSFER Round Two – More Tips and Tricks

Jim Brown

FIZ Karlsruhe

STNext®

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS®
A DIVISION OF THE
AMERICAN CHEMICAL SOCIETY

What is TRANSFER?

- A way to extract terms from a set and search them
- Extract from one field or multiple fields
  - Up to 50,000 terms
- You can search them in the same field or a different field
- Search the terms in the same database or in a different database

How to use TRANSFER

- Enter database in which the search is to be conducted
- Novice Mode – TRANSFER
  - Put in L-number answer set of interest
  - The field code of the info you wish to extract
  - Answer numbers
- Expert Mode – TRA Lx *field code answer number(s)*
- HELP EFIELDS for list of valid extraction fields in that database

TRANSFER – Round 2

- Round 1 - https://www.stn-international.com/en/training-center/recorded-e-seminars/stn-coffee-lecture-transfer-tips-and-tricks
  - TRANSFER in one database
  - TRANSFER across multiple databases
  - TRANFER WITH “ ”
- Round 2
  - Requalifying extracted info to a new field
  - TRA PNK vs. PN

TRANSFER with requalification

- Have to use when the data is in a different field than in the original database
  - CAplus to CAS Formulations
  - DWPI to DCR
- TRA Lx *original data field* 1- */new data field*
  - CAplus to CAS Formulations – **TRA Lx AN 1- /CPAN**
  - DWPI to DCR - **TRA Lx DCR 1- /AN.S**

The CASFORMULATIONS database may also provide additional relevant details on the formulation

```
=> FILE CAPLUS
=> S L21 OR L25
L27 4 L21 OR L25

=> S L27 AND CASFORMULTNS/OS
L28 3 L27 AND CASFORMULTNS/OS

=> FILE CASFORM
=> S L28

COMBINATION OF STRUCTURE AND TEXT TERMS NOT VALID

=> TRA AN L28 /CPAN

L29          TRANSFER L28 1- AN :        6 TERMS
L30          25 L28/CPAN
```

Use the Other Source (OS) field to refine to patent results that are in the CASFORMULATIONS database.

TRANSFER the CAplus Accession Number (AN) and search in the CPAN field to find formulations indexed from the patent records.

DWPI to DCR Search example

- Find patents published in 2020-2021 that are assigned to Eli Lilly. Take a look at specific structures Derwent has indexed for those family records, and possibly refine that structure search.

TRANSFER and requalifying the data points

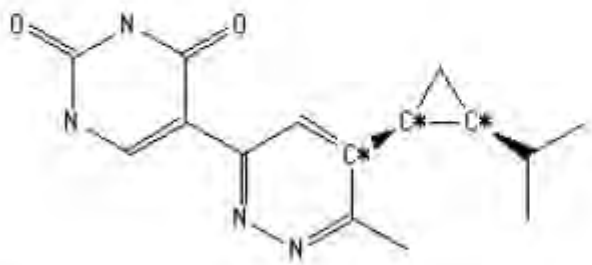
```
'BI,BIEX' IS DEFAULT SEARCH FIELD FOR 'WPINDEX' FILE

=> S ELIL-C/PACO AND 2020-2021/PY.B

          6407 ELIL-C/PACO
       5112039 2020-2021/PY.B
L1         106 ELIL-C/PACO AND 2020-2021/PY.B

=> TRA L1 DCR 1- /AN.S

SELECT IS APPROXIMATELY 64% COMPLETE
L2          TRANSFER L1 1- DCR :    1262 TERMS
L3         470 L2/AN.S
L4             QUE  TERMS FROM L2/AN.S WITH NO HITS:   792 TERMS

=> D L3 ALL 1-3
```

Use the Derwent Patent Assignee Codes to improve comprehensiveness.

This TRANSFER command takes the indexed DCR terms from the Eli Lilly patent family records and transfers them to the corresponding DCR records (L3) by requalifying the DCR terms to the DCR Accession Number field (AN.S.)

DCR Displays

L3 ANSWER 1 OF 470 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN
AN.S **DCR**-5314175
DCSE 5314175-1-0-0
CN.S 5-[5-((1R,2S)-2-Isopropyl-cyclopropyl)-6-methyl-pyridazin-3-yl]-1H-pyrimidine-2,4-dione



MF C15 H18 N4 O2
SMF C15 H18 N4 O2 *1; TYPE *1; TOTAL *1
MW 286.33
SDCN RCZGZY

DCR Displays

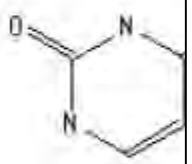
L3 ANSWER 1 OF 470 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S DCR-5

DCSE 531417

CN.S 5-[5-(

pyrimi



MF C15 H1

SMF C15 H1

MW 286.33

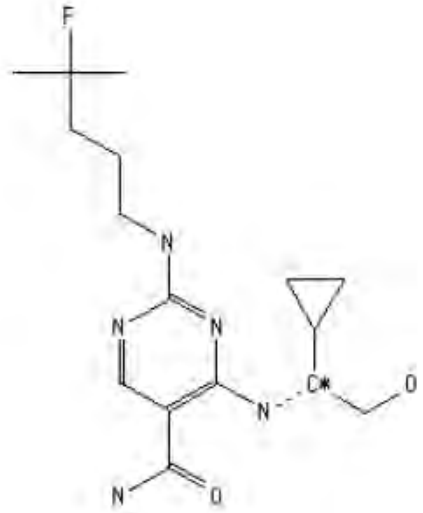
SDCN RCZGZY

L3 ANSWER 2 OF 470 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S DCR-5284678

DCSE 5284678-1-0-0

CN.S 4-((R)-1-Cyclopropyl-2-hydroxy-ethylamino)-2-(4-fluoro-4-methyl-pentylamino)-pyrimidine-5-carboxylic acid amide



MF C16 H26 F N5 O2

SMF C16 H26 F N5 O2 *1; TYPE *1; TOTAL *1

MW 339.41

SDCN RCYVHU

DCR Displays

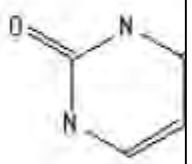
L3 ANSWER 1 OF 470 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S **DCR**-5

DCSE 531417

CN.S 5-[5-(

pyrimi



MF C15 H1

SMF C15 H1

MW 286.33

SDCN RCZGZY

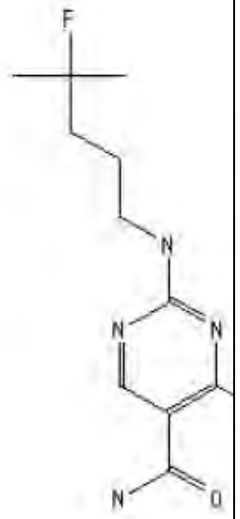
L3 ANSWER 2 OF 470 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S **DCR**-528467

DCSE 5284678-1-

CN.S 4-((R)-1-C

pentylamin



MF C16 H26 F

SMF C16 H26 F

MW 339.41

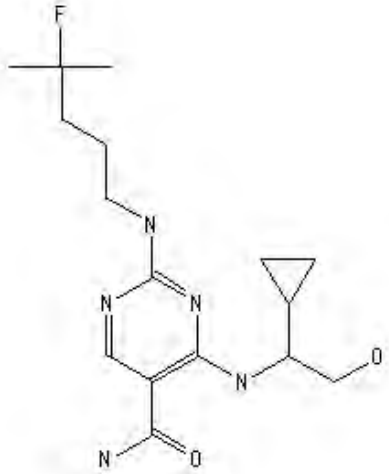
SDCN RCYVHU

L3 ANSWER 3 OF 470 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S **DCR**-5284677

DCSE 5284677-0-0-0

CN.S 4-(1-Cyclopropyl-2-hydroxy-ethylamino)-2-(4-fluoro-4-methyl-pentylamino)-pyrimidine-5-carboxylic acid amide



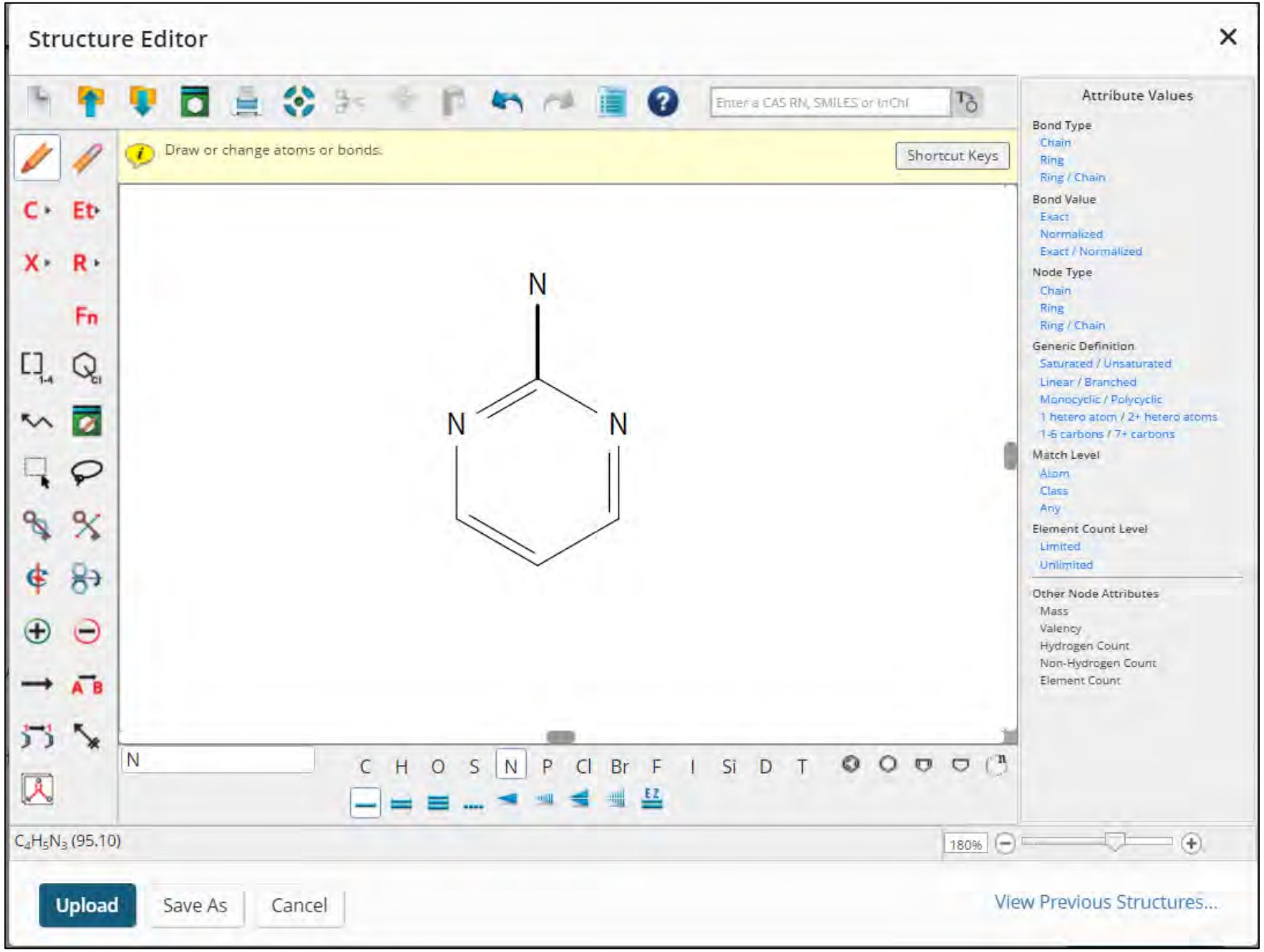
MF C16 H26 F N5 O2

SMF C16 H26 F N5 O2 *1; TYPE *1; TOTAL *1

MW 339.41

SDCN RCYVHT

Refine set by structure search



Substructure Search

```
L5      STRUCTURE UPLOADED

=> S L5 SSS FULL

FULL SEARCH INITIATED 18:44:31
FULL SCREEN SEARCH COMPLETED -   189826 TO ITERATE

100.0% PROCESSED     189826 ITERATIONS                    134346 ANSWERS
SEARCH TIME: 00.00.14

L6       134346 SEA SSS FUL L5

=> S L3 AND L6

L7           12 L3 AND L6

=> D 1-3
```

12 of the 470 DCR records from the original set have structure of interest.

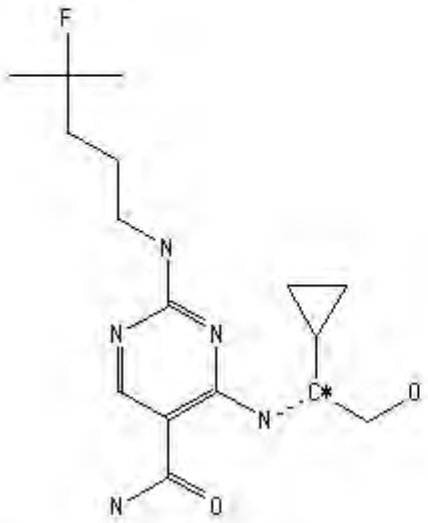
Displays from refined set

L7 ANSWER 1 OF 12 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S **DCR**-5284678

DCSE 5284678-1-0-0

CN.S 4-((R)-1-Cyclopropyl-2-hydroxy-ethylamino)-2-(4-fluoro-4-methyl-pentylamino)-pyrimidine-5-carboxylic acid amide



MF C16 H26 F N5 O2

Displays from refined set

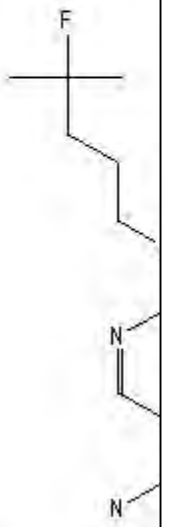
L7 ANSWER 1 OF 12 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S DCR-5

DCSE 52846

CN.S 4-((R

penty



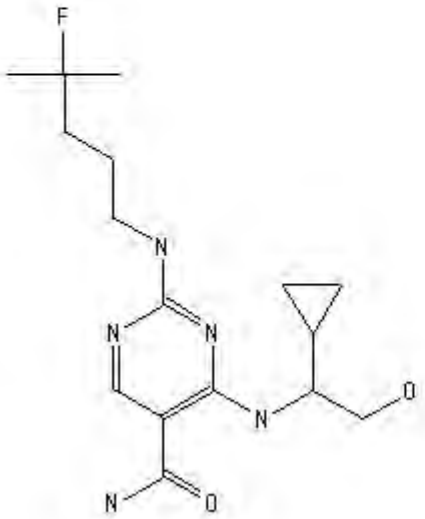
MF C16 H

L7 ANSWER 2 OF 12 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S DCR-5284677

DCSE 5284677-0-0-0

CN.S 4-(1-Cyclopropyl-2-hydroxy-ethylamino)-2-(4-fluoro-4-methyl-pentylamino)-pyrimidine-5-carboxylic acid amide



MF C16 H26 F N5 O2

Displays from refined set

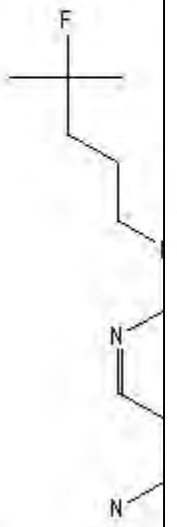
L7 ANSWER 1 OF 12 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S **DCR**-5

DCSE 528467

CN.S 4-((R)

pentyl



MF C16 H2

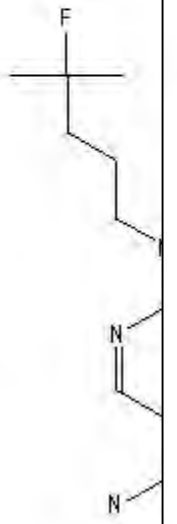
L7 ANSWER 2 OF 12 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S **DCR**-5284677

DCSE 528467

CN.S 4-(1-(

pyrimi



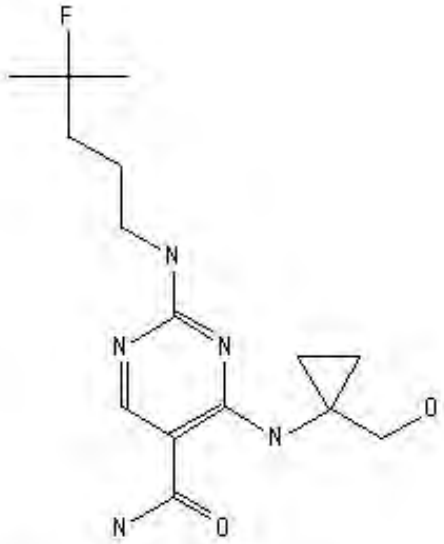
MF C16 H2

L7 ANSWER 3 OF 12 WPINDEX COPYRIGHT 2021 CLARIVATE ANALYTICS on STN

AN.S **DCR**-5284676

DCSE 5284676-0-0-0

CN.S 2-(4-Fluoro-4-methyl-pentylamino)-4-(1-hydroxymethyl-cyclopropylamino)-pyrimidine-5-carboxylic acid amide



MF C15 H24 F N5 O2

TRA PN vs. PNK

- PN – extracts the publication (patent) number from a set
- PNK – extracts the publication (patent) number and their corresponding kind codes from a set
  - Can result in a cleaner set
  - Only available for four databases
  - CAplus, DWPI, INPA, ReaxysFileBib
- Quick test – 5 chemical & 5 pharma companies, 2019-2021, across DWPI and CAplus

TRANSFER - PNK vs. PN

- Create a set in Derwent World Patents Index
  - Derwent Patent Assignee codes
  - Date range for basic publication date
- TRA PNK in CAplus
- TRA PN CAplus
- Compare results

Conclusions

- In all ten searches, the TRANSFER count was identical using PNK vs. PN
- Anecdotal evidence but too small a sample size to be conclusive
- Both PN and PNK work
  - With PN search, you *may* get a few false drops that you wouldn't get with PNK

Summary

- TRANSFER is a way to extract terms from a set and search them
  - Search the terms in the same database or in a different database
- The searcher can limit the terms extracted with TRANSFER by using WITH “”
- Requalify when needed to speak across different data fields
- PN vs. PNK

Contact Us

CAS **help@cas.org**
www.cas.org

FIZ Karlsruhe **helpdesk@fiz-karlsruhe.de**
www.stn-international.de

STNext®

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS®
A DIVISION OF THE
AMERICAN CHEMICAL SOCIETY