CAS STNext®

THE DERWENT FRAGMENTATION GENERATOR!

CAS
A division of the
American Chemical Society

Agenda

- Derwent World Patents Index database (aka DWPI)
- Derwent fragmentation code system
- Derwent fragmentation code generator
- Search example
  - Generate script of codes from structure
  - Run script
  - Display corresponding DWPI records

Derwent World Patents Index (aka DWPI)

- Patent family based database
- Value-add indexing
  - Enhanced titles and abstracts
  - Derwent classes and Manual codes
  - Polymer indexing
  - Specific compound indexing
  - Fragmentation code indexing

Requires a Derwent CPI subscription to search

Derwent World Patents Index (aka DWPI)

- Patent family based database
- Value-add indexing
  - Enhanced titles and abstracts
  - Derwent classes and Manual codes
  - Polymer indexing
  - Specific compound indexing
  - Fragmentation code indexing

DWPI Chemical Fragmentation Codes

- Derwent fragmentation indexing system pre-dates DCR and Markush records – back to 1963
- Only way to search for substances prior to introduction of DCR and Markush indexing
- Fragmentation codes are still in use today, in conjunction with DCR and Markush indexing
- Fragmentation codes include structural and non-structural codes
- Non-structural codes include
  - Processes/Apparatus
  - Pharmaceutical and Agricultural Activities, Properties and Uses

DWPI Chemical Fragmentation Codes

- CPI subscription required to search fragmentation codes
- Specific and Markush substance indexing searchable directly within DWPI records
  a) Unique Derwent substance indexing for more than 500.000 DWPI records
  b) Only comprehensive source of Markush indexing on STNext from 1963-1977Substances are represented as separate paragraphs of linked codes in DWPI records - (P) operator
- Each code represents a part or aspect of the substance
- **STNext can be used to generate and upload the appropriate combination of structural codes!**

Derwent frag codes in DWPI record

```
AN   2022-06309B [2022004]   WPIX Full-text
TI   Synthesizing nicotinicate compounds under non-metallic conditions,
     comprises e.g. performing oxidative dehydrogenation ketene cyclization
     reaction of substituted ethenamine and substituted cyclopropanol in
     organic solvent using tetramethylpiperidine nitrogen oxide as oxidant
DC   B02; B03
IN   SUN Y; ZHAN J; ZHU L
PA   (UAYN-C) UNIV ANYANG NORMAL
CYC  1
PIA  CN 113444037     A  20210928 (2022004)* ZH
ADT  CN 113444037  A CN 2021-10890233 20210804
PRAI CN 2021-10890233      20210804
```

```
M2 *29*   F012 F013 F016 F431 G010 G100 J0 J2 J011 J211 M1 M113 M210 M211
          M212 M240 M272 M281 M320 M413 M510 M521 M531 M540 M710 M720 N204
          N213 N242 N261 N305 N311 N312 N318 N334 N411 N513  M905 M904
          DCN: RBRMX8-N RBRMX8-P
          DCR: 3217831-N 3217831-P
M2 *30*   C000 C017 D000 D012 D019 D020 D022 D029 D030 D
          F014 F019 F020 F029 F111 F199 F211 F299 F431 G
          G011 G012 G013 G015 G019 G020 G021 G022 G029 G
          G040 G050 G111 G112 G113 G221 G299 G553 G563 G
          H583 H584 H600 H601 H602 H607 H608 H609 H641 H642 H643 H683 H684
          H685 H689 J011 J012 J211 L142 L199 L511 L599 L960 M113 M115 M116
          M119 M121 M136 M144 M146 M210 M211 M212 M213 M214 M215 M216 M231
          M232 M240 M250 M272 M280 M281 M282 M283 M311 M320 M321 M322 M323
          M331 M332 M342 M344 M349 M352 M360 M36
          M381 M382 M383 M391 M392 M393 M412 M41
          M520 M521 M522 M523 M530 M531 M532 M53
          M720 N204 N213 N242 N261 N305 N311 N31
          M905 M904
          MCN: A000-8QS01-N A000-8QS01-P
```

This paragraph contains the frag codes associated with DCR compound 3217831.

This paragraph contains the frag codes associated with Markush structure record A000-8QS01.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

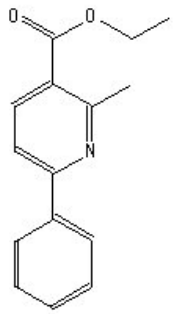
CAS
A division of the American Chemical Society

Derwent frag codes in DWPI record

```
M2 *29*   F012 F013 F016 F431 G010 G100 J0 J2 J011 J211 M1 M113 M210 M211
          M212 M240 M272 M281 M320 M413 M510 M521 M531 M540 M710 M720 N204
          N213 N242 N261 N305 N311 N312 N318 N334 N411 N513  M905 M904
          DCN: RBRMX8-N RBRMX8-P
          DCR: 3217831-N 3217831-P
```

```
L1   ANSWER 1 OF 1  WPIX COPYRIGHT 2022   CLARIVATE ANALYTICS on STN
AN.S DCR-3217831
DCSE 3217831-0-0-0
CN.S 2-Methyl-6-phenyl-nicotinic acid ethyl ester


STR
```




```
MF   C15 H15 N O2
```

F012 – Position 2 of the pyridine is substituted
F013 – Position 3 of the pyridine is substituted
F016 – Position 6 of the pyridine is substituted
F431 – Pyridine ring
G010 – 1 substitution on benzene ring
G100 – Unfused benzene – no other carbocycle present
J0 – COOH/Derivative
J2 – Ester
J011 – 1-COOH or derivative
J211 – 1- Ester attached to a heterocyclic ring
. . .
N204 – Bonds broken – Carbocyclic ring opened
N213 – C-C broken – C=C
N242 – C-O broken – C-O other
N261 – H-N bond broken
. . .

Structural codes Non-Structural codes

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

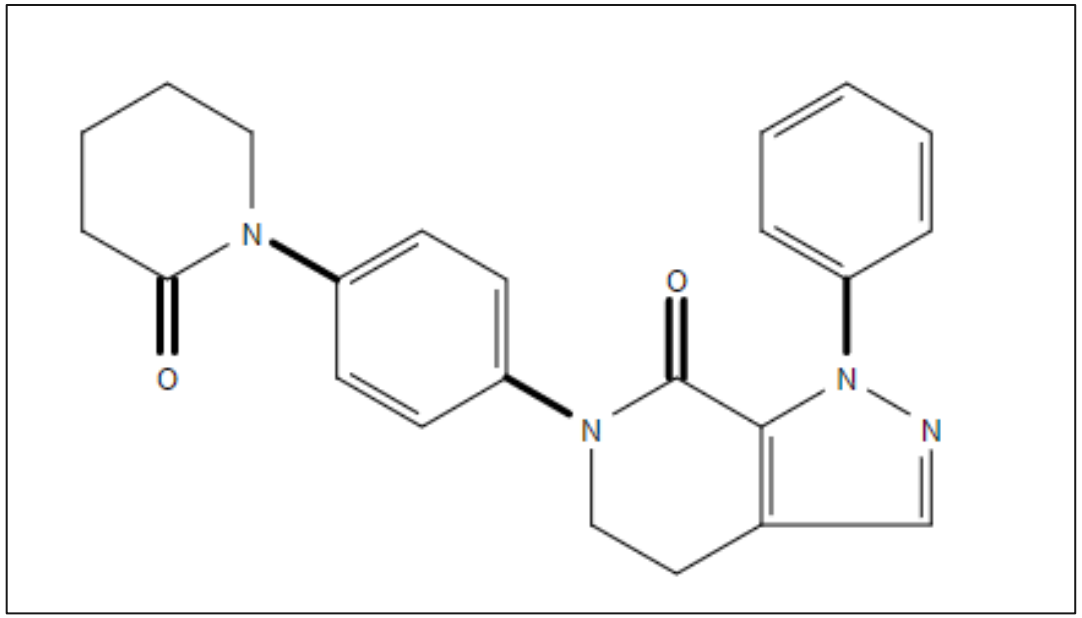
CAS
A division of the American Chemical Society

Derwent Fragmentation Code Search

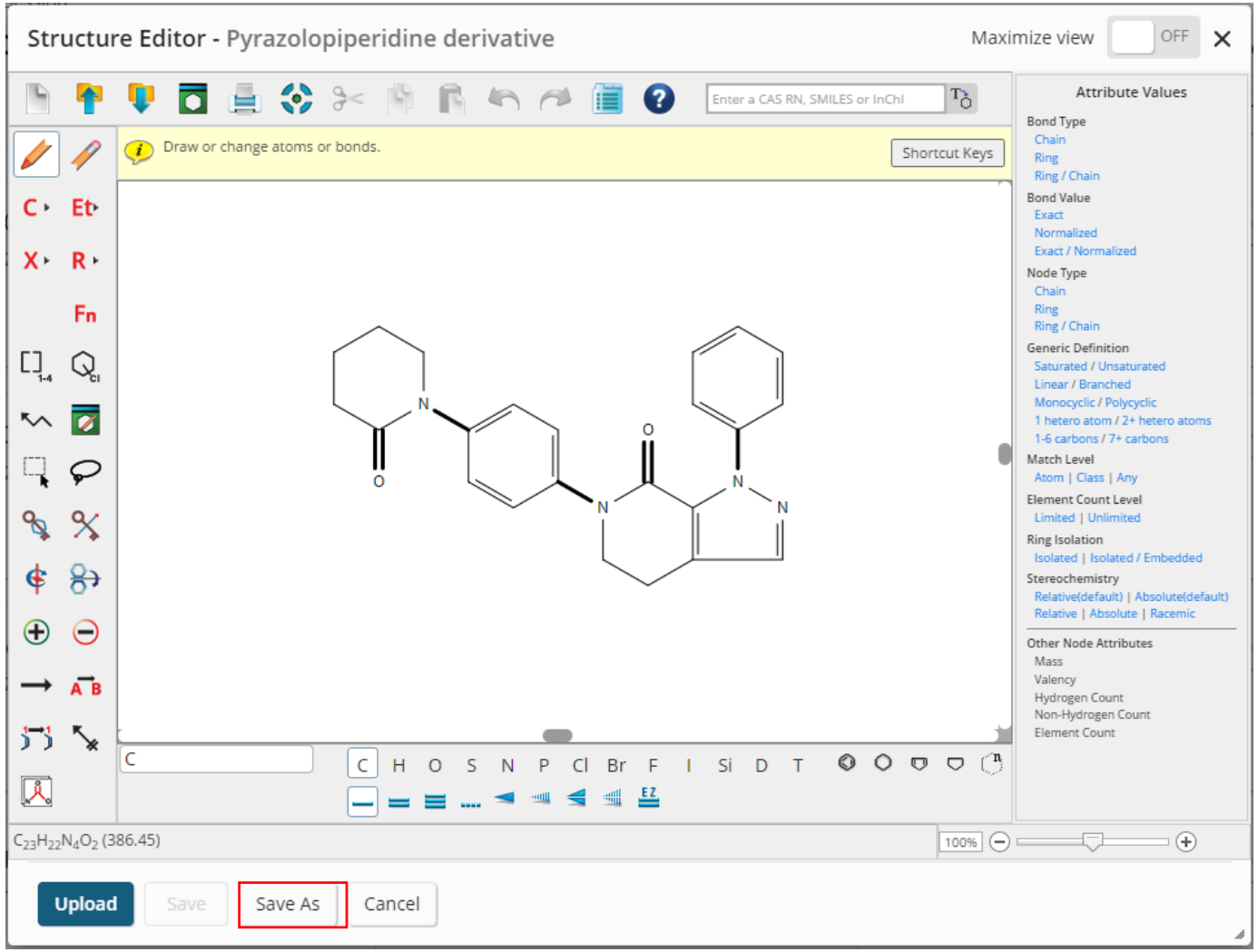
1. Draw structure using STNext Structure Editor
2. Click on Save As, name structure, Click on Save
3. Close STNext Structure Editor
4. Click on ellipsis for structure
5. Choose Generate FragCode Script
6. (Re)name, Click on Generate Script
7. Click on Run, in WPIX or WPIDS
8. Display records in preferred format

Search example

Create a Derwent fragmentation code strategy for this structure



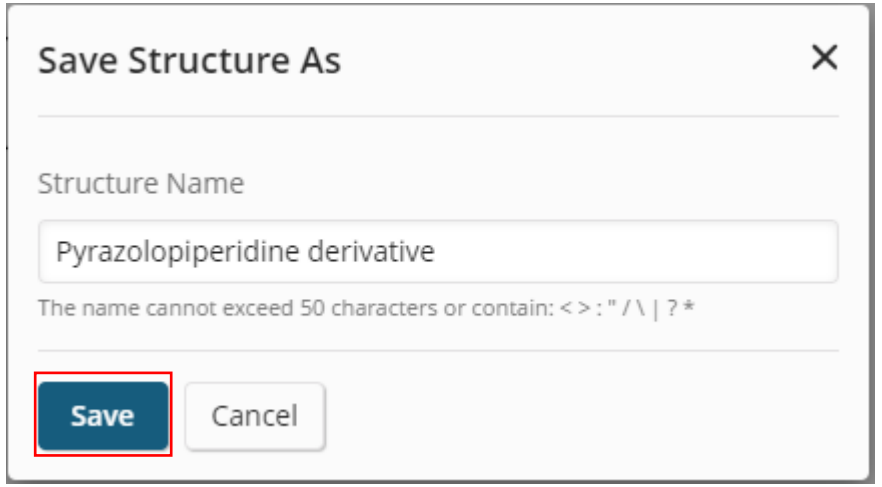
Draw Structure in STNext Structure Editor

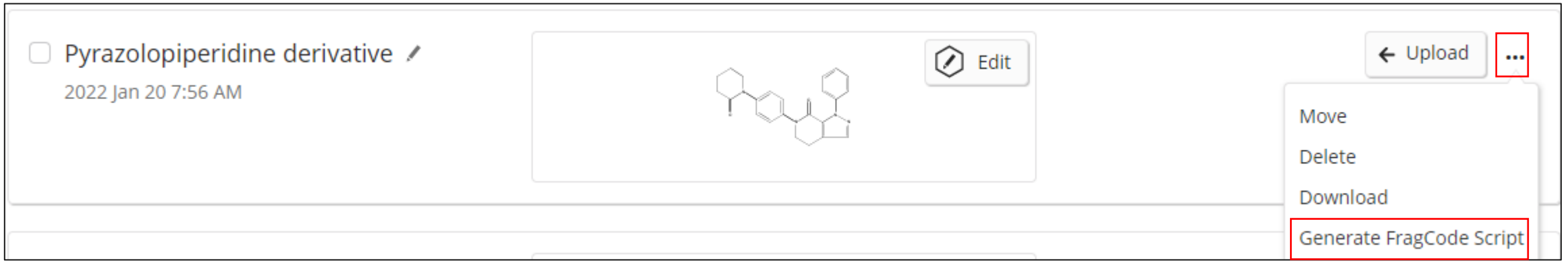


FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Save Structure As, Generate FragCode Script

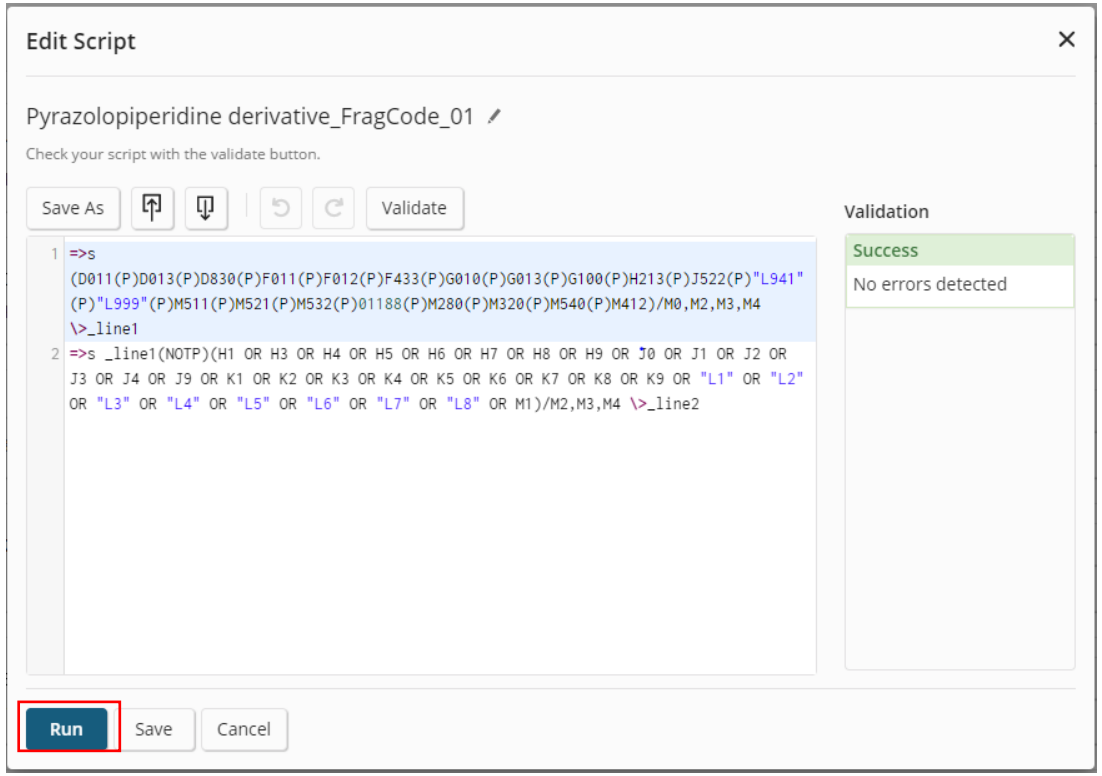




FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Generate and Edit Script



FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Fragmentation code search in WPIX

```
=> s (D011(P)D013(P)D830(P)F011(P)F012(P)F433(P)G010(P)G013(P)G100(P)H213(P)J522(P)"L941"(P)"L999"
(P)M511(P)M521(P)M532(P)01188(P)M280(P)M320(P)M540(P)M412)/M0,M2,M3,M4

         12811 D011/M0
        282239 D011/M2
         84855 D011/M3
         46137 D011/M4
         12811 D013/M0
        272302 D013/M2
         83694 D013/M3
         45409 D013/M4
```

. . .

```
        245722 "L8"/M2
         78309 "L8"/M3
           976 "L8"/M4
        439858 M1/M2
        203155 M1/M3
         67743 M1/M4
L3           9 L2(NOTP)(H1 OR H3 OR H4 OR H5 OR H6 OR H7 OR H8 OR H9 OR J0 OR
               J1 OR J2 OR J3 OR J4 OR J9 OR K1 OR K2 OR K3 OR K4 OR K5 OR K6
               OR K7 OR K8 OR K9 OR "L1" OR "L2" OR "L3" OR "L4" OR "L5" OR
               "L6" OR "L7" OR "L8" OR M1)/M2,M3,M4
```

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Display - BIB portion

```
=> D BIB KWIC


L3   ANSWER 1 OF 9  WPIX COPYRIGHT 2022   CLARIVATE ANALYTICS on STN
AN   2021-A6614S [2021087]   WPIX Full-text
TI   New bifunctional compounds useful in medicament to treating disease
     associated with endogenous protein recruitment to E3 ubiquitin ligase or
     androgen receptor, preferably e.g. acne, hirsutism, enlarged sebaceous
     glands, alopecia and asthma
DC   B05; D16; D21
IN   TONG Y
PA   (SUZH-N) SUZHOU KINTOR PHARM INC
CYC  135
PIA  WO 2021180160    A1 20210916 (2021087)* ZH 446[17]
     CN 113387930     A  20210914 (2021087)  ZH
ADT  WO 2021180160 A1 WO 2021-CN80167 20210311; CN 113387930 A CN 2021-10246427
     20210305
PRAI CN 2021-10246427      20210305
     CN 2020-11593015      20201229
     CN 2020-10165789      20200311
```

Display – KWIC portion

```
M2 *97*   C000 C100 D010 D011 D012 D013 D014 D019 D020 D021 D022 D023
          D024 D025 D029 D040 D049 D611 D700 D720 D830 D840 D860 D900
          D910 D920 D970 D980 F010 F011 F012 F013 F014 F015 F016 F017
          F019 F020 F021 F029 F423 F433 F710 G001 G002 G003 G010 G011
          G012 G013 G014 G015 G016 G019 G020 G021 G022 G029 G030 G033
          G034 G035 G039 G040 G050 G100 G111 G112 G113 G221 G299 G553
          G563 G599 H100 H102 H103 H121 H122 H123 H141 H142 H2 H211 H212
          H213 H401 H402 H421 H422 H521 H522 H541 H600 H608 H609 H621
          H622 H623 H641 H681 H682 H683 J011 J012 J013 J311 J312 J371 J5
          J521 J522 J523 J592 J599 K130 K140 K620 K640 K830 K850 K899
          K930 L110 L130 L142 L143 L199 L9 L930 L941 L943 L999 M111
          M112 M113 M114 M115 M116 M119 M122 M123 M124 M125 M126 M129 M136
          M141 M143 M210 M211 M212 M213 M214 M215 M216 M231 M232 M233 M240
          M272 M273 M280 M281 M282 M283 M311 M312 M313 M314 M315 M320
          M321 M322 M331 M332 M333 M340 M342 M349 M353 M373 M381 M391
          M412 M431 M511 M512 M513 M521 M522 M523 M530 M531 M532
          M533 M540 M541 M542 M543 M630 M640 M650 M710 M782 M800 P431
          P446 P510 P517 P518 P526 P530 P554 P611 P612 P623.  .  .  P731
          P739 P812 P814 P818 P822 P824 P844 P922 P930 P943 P950  M905
          M904
          RIN: 01160 01168 01172 01174 01175 01186 01188 01190 01262
          01264 01275 01279 07978 11882 46969 47119 47120 48548 50527
          50531 62081
          MCN: 2259-00002-M 2259-00002-N
```

Summary

- Derwent fragmentation code searches improve comprehensiveness of structure searches
- Automatically generated strategy may need to be modified
  - Consider negation codes
  - Consider codes with counts
  - Consider derivatives – hydroxy to ether or ester
- Limitations of the FragCode Generator system
- In-house training available

Derwent fragmentation documentation

- https://clarivate.com/intellectual-property/training-support/derwent/dwpi-reference-center/indexing-user-guides/chemistry-indexing/Derwent CPI Chemical indexing Guide
  - Contains detailed descriptions of how documents are indexed
  - Definitions of all codes
- Derwent CPI fragmentation Codes Chart

For more information…

CONTACT

CAS
help@cas.org
cas.org

FIZ Karlsruhe
EMEAhelp@cas.org

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the
American Chemical Society