



Chemical Fragmentation Code の応用

(化学構造検索 On DWPI)

Feb., 2007

Thomson Corporation
Thomson Scientific

Online Business Manager

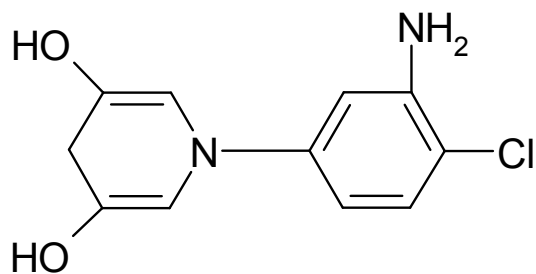
ツー ツオン
褚 冲

THOMSON
★

1. 化学構造式にコードを付与してみる
2. Chemical Fragmentation Codeを修正する
3. DWPIで構造検索 (Chemical Fragmentation Codes と DCR)
4. コード修正のまとめ
5. Multi Files, Multi Toolsを利用する検索参考例

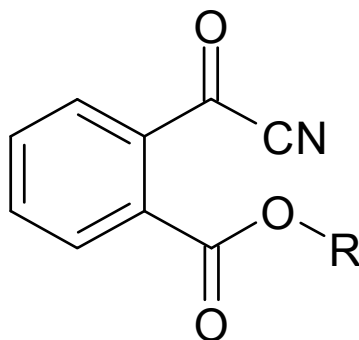
THOMSON
★

Example I

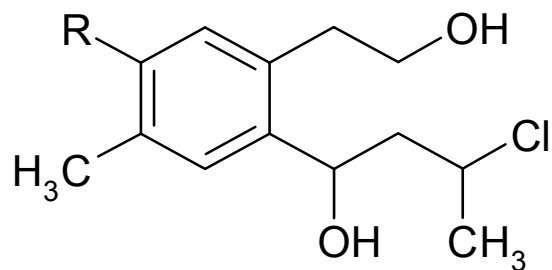


Example II

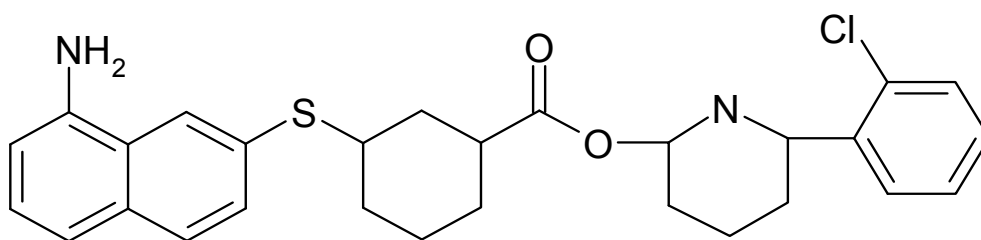
R=H or Alkyl



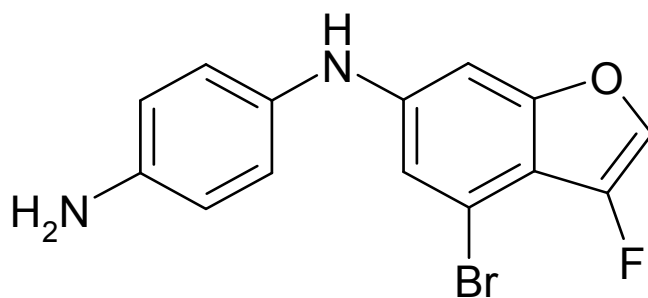
Example III

R=H, Alkyl(1-6C)

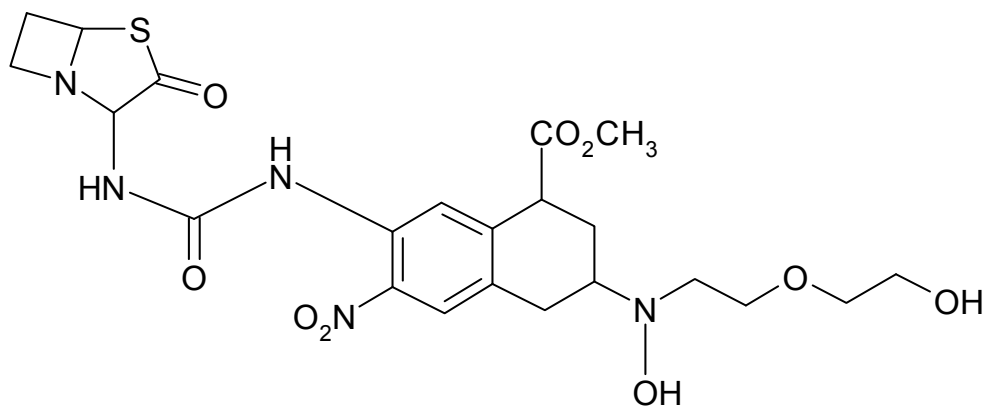
Example IV



Example V



Example VI

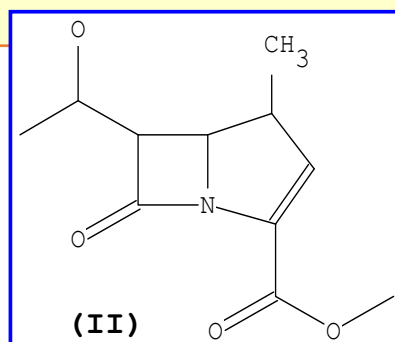


1. 化学構造式にコードを付与してみる
2. Chemical Fragmentation Codeを修正する
3. DWPIで構造検索 (Chemical Fragmentation Codes と DCR)
4. コード修正のまとめ
5. Multi Files, Multi Toolsを利用する検索参考例

Fragmentation code 修正の例

課題1: DWPI fragmentation codeを用いて包括的な式で carbapenem の誘導体を検索する

[構造 (II)]



Substructure検索用のコードに修正に使われる一般的な手法

1. Negation codes, 例えば H1等を削除。
2. 基本構造(Basic Group): M4グループコードをより高い順位のコードを含める。例えば、M411。
3. D01/D02 や G01/G02/G03関連コードを加えて、環上置換位置の許容数を広げる
4. 一般官能基が存在する数を考慮する。例えば、J111, J112。
5. Hydroxyl はether/esterになる可能性はあるか? Amineがamideになるか? 可能であれば、ORするかこれらのコードを削除するか
6. カーボンチェンコード(e.g. M280, M320)を編集・削除。
7. 環系の数関連のコードを考慮する。例えば、M511, M512。
8. 環間コードのオプションM1 codeを考慮する。例えば、M131 (>C=O), M132 (other carbon), など
9. 必要に応じ、ジェネリックコードを加える。例えば、H600, F020。
10. その他: 色違いのコードをORする場合の注意点

Fragmentation code command file の発生と編集...

標準の fragmentation code はMarkush Topfrag
や STN Expressを で発生させる

```

=>S (D690 (P) H401 (P) H481 (P) J211 (P) J521 (P) M240 (P) M331 (P) M412) /M0, M2, M3, M4 ¥>_line1
=>S _line1 (P) (M511 (P) M520 (P) M530 (P) M540) /M2, M3, M4 ¥>_line2
=>S _line2 (P) (M210 (P) M281 (P) M312 (P) M321 (P) M340 (P) M342 (P) M391) /M2, M3, M4 ¥>_line3
=>S _line3 (P) ((M270 OR M272) (P) (M370 OR M373)) (M2, M3, M4 ¥>_line4)
=>S _line4 (P) 41252/RIN ¥>_line5
=>S _line5 (P) (D013 (P) D019 (P) J011 (P) "L941" (P) M211) /M2, M3, M4 ¥>_line6
=>S (_line1 (P) M900/M0) OR (_line2 (P) M901/M2, M3, M4) OR (_line5 (P) M902/M2, M3, M4) ¥>_line7
=>S _line7 OR _line6 ¥>_line8
=>S _line8 (NOTP) (H1 OR H2 OR H3 OR H5 OR H6 OR H7 OR H9 OR J1 OR J3 OR J4) /M2, M3, M4 ¥>_line9
=>S _line9 (NOTP) (J6 OR J9 OR K1 OR K2 OR K3 OR K4 OR K5 OR K6 OR K7 OR K8) /M2, M3, M4 ¥>_line10
=>S _line10 (NOTP) (K9 OR "L1" OR "L2" OR "L3" OR "L4" OR "L5" OR "L6") /M2, M3, M4 ¥>_line11
=>S _line11 (NOTP) ("L7" OR "L8" OR M1) /M2, M3, M4 ¥>_line12
  
```

この例はすべての化学関連を検索対象に、M0 – M4.

Note: Line1 = Black; Line2 = Red; Line3-5 = Blue; Line6 = Green

Step 1: negation codesを削除

```

=>S (D690 (P)H401 (P)H481 (P)J211 (P)J521 (P)M240 (P)M331 (P)M412)/M0,M2,M3,M4 ¥>_line1
=>S _line1 (P) (M511 (P)M520 (P)M530 (P)M540)/M2,M3,M4 ¥>_line2
=>S _line2 (P) (M210 (P)M281 (P)M312 (P)M321 (P)M340 (P)M342 (P)M391)/M2,M3,M4 ¥>_line3
=>S _line3 (P) ((M270 OR M272) (P) (M370 OR M373))/M2,M3,M4 ¥>_line4
=>S _line4 (P) 41252/RIN ¥>_line5
=>S _line5 (P) (D013 (P)D019 (P)J011 (P) "L941" (P)M211)/M2,M3,M4 ¥>_line6
=>S (_line1 (P)M900/M0) OR (_line2 (P)M901/M2,M3,M4) OR (_line5 (P)M902/M2,M3,M4) ¥>_line7
=>S _line7 OR _line6 ¥>_line8

=>S _line8 (NOTP) (H1 OR H2 OR H3 OR H5 OR H6 OR H7 OR H9 OR J1 OR J3 OR J4)/M2,M3,M4
  ¥>_line9
=>S _line9 (NOTP) (J6 OR J9 OR K1 OR K2 OR K3 OR K4 OR K5 OR K6 OR K7 OR K8)/M2,M3,M4
  ¥>_line10
=>S _line10 (NOTP) (K9 OR "L1" OR "L2" OR "L3" OR "L4" OR "L5" OR "L6")/M2,M3,M4 ¥>_line11
=>S _line11 (NOTP) ("L7" OR "L8" OR M1)/M2,M3,M4 ¥>_line12

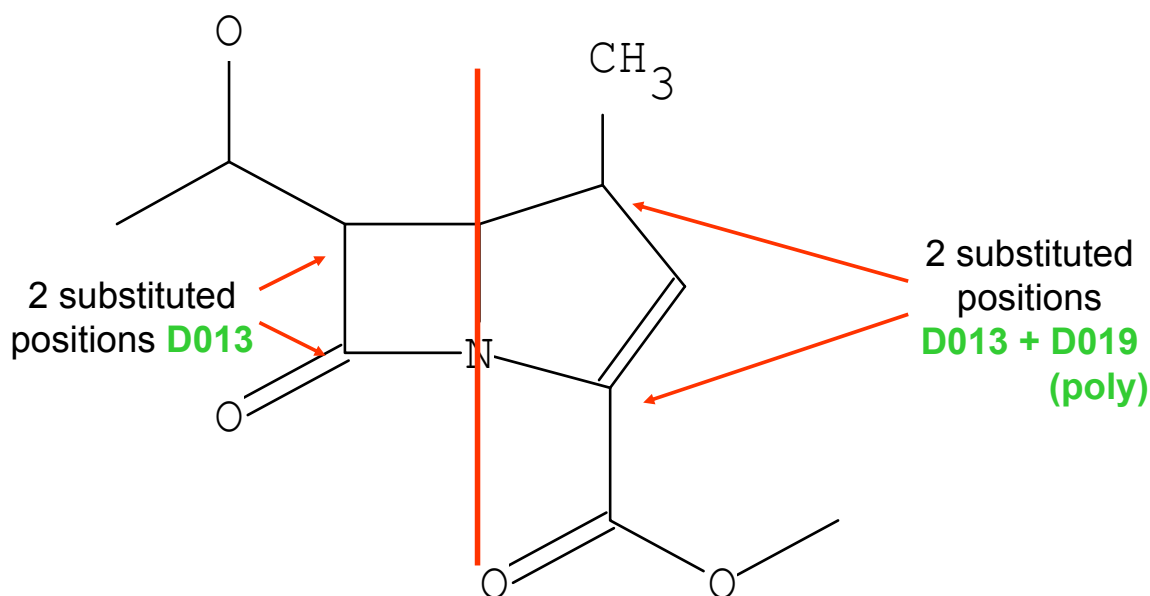
```

最後にある2桁のNegation codes 例. H1 Amine

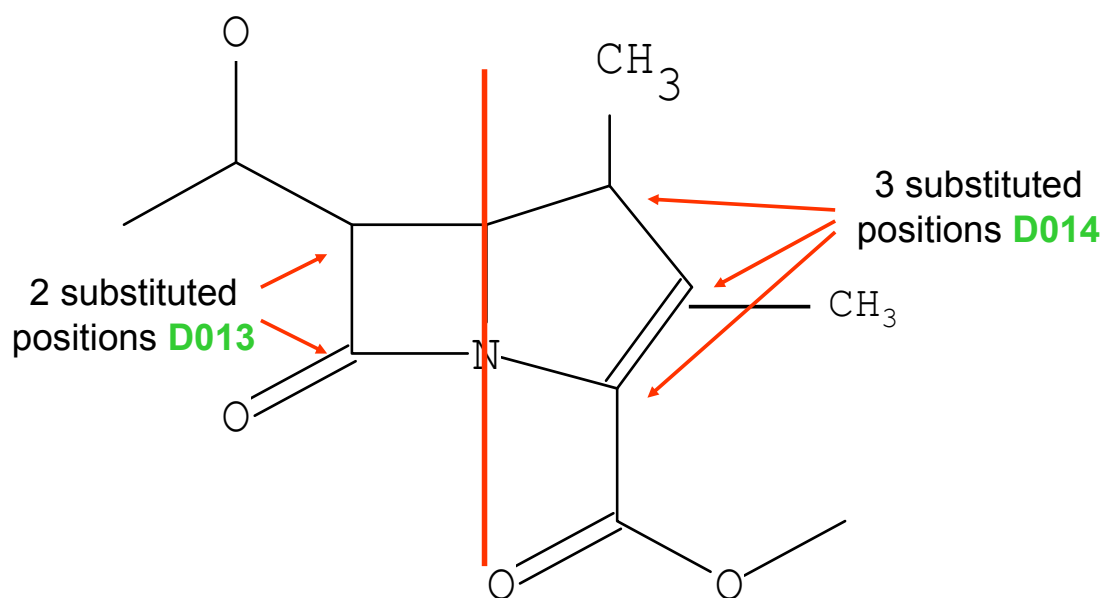
Steps 2-10: コードを修正する例

2. M412 に、 M411を ORする (例えば、塩を含む)
3. 2-置換 D013に、 3-置換D014 をORする。
4. 1-ester J211に、 >1の J212 をORする。
1-oxo J521に、 >1の J522 をORする。
1-carboxy derivative J011 を削除。
5. Hydroxyl H401/H481 を削除。
6. Carbon chain (M2/M3) の選択肢をORする
7. 1-ring M511に、 >1のM512, M513 ORする。
No-ring コードM520, M530, M540 を削除。
8. この式にはM1 codes はない。(環間コードはない)
9. このケースではジェネリックコード(eg. H600)を加えることを考慮しない。
10. このケースでは色違いのコードをORすることを考慮しない。

Fused-heterocyclic 環上の置換位置コード



Fused-heterocyclic 環上の置換位置コード

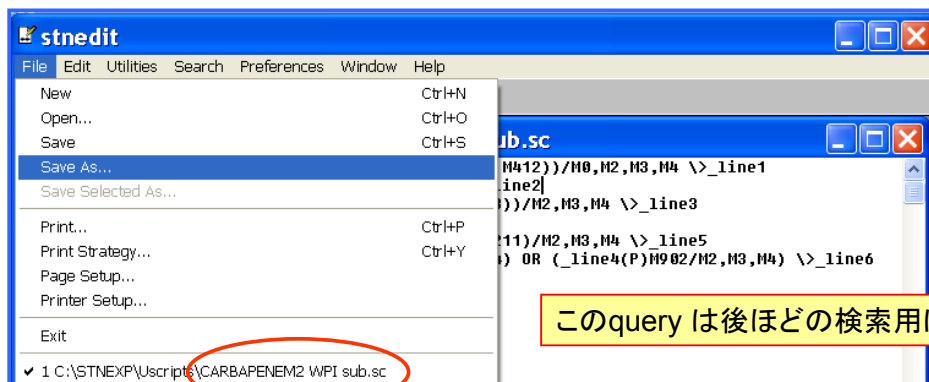


修正後の式を保存

編集後の STN command file

```

=>S (D690(P)J21!(P)J52!(P)M240(P)M331(P)(M411 OR M412))/M0,M2,M3,M4 ¥>_line1
=>S _line1(P)(M511 OR M512 OR M513)/M2,M3,M4 ¥>_line2
=>S _line2(P)(M210(P)(M281 OR M282 OR M283)(P)(M321 OR M322 OR M323))/M2,M3,M4¥>_line3
=>S _line3(P)41252/RIN ¥>_line4
=>S _line4(P)(D013(P)(D019 OR D014)(P)"L941"(P)M211)/M2,M3,M4 ¥>_line5
=>S (_line1(P)M900/M0) OR (_line2(P)M901/M2,M3,M4) OR (_line4(P)M902/M2,M3,M4) ¥>_line6
=>S _line6 OR _line5 ¥>_line8
  
```



コードのORする時 (可変基を含む化合物の検索式の修正)

検索対象の化合物に可変基を含む場合、さまざまな選択枝のコードをORするとき、十分注意する必要があります。特に選択枝を異なる色のコードで検索するとき、細心の注意が必要です。コードをORするとき、コードの行を下記の表で確認してください。

コード	黒	赤	青	緑	緑# (先行コード黒)	緑# (先行コード青)
黒	1	※	※	※	1	※
赤	※	2	※	※	※	※
青	※	※	3	※	※	3
緑	※	※	※	4	※	※
緑# (先行コード黒)	1	※	※	※	1	※
緑# (先行コード青)	※	※	3	※	※	3

1=検索式の第1行(黒コード)
 2=検索式の第2行(赤コード)
 3=検索式の第3行(青コード)
 4=検索式の第4行(緑コード)
 ※については下記の(iii)をご参照ください。

i) 選択枝が全て同じ色のコードで索引されている場合
 この場合には、コードを同一行でORすることができます。

ii) 選択枝が異なる色のコードで索引され、一部のコードに#が付いている場合
 この場合には、コードの色に注目し、表を参照してコードを追加する行を決定します。

iii) 選択枝が異なる色のコードで索引され、#が付いているコードがない場合
 これは最も難解なケースですが、3つの解決策が考えられます。
 (iii-a) 可変基のコードを削除し、他のコードで検索して、所望のレコードをヒットします。
 (iii-b) 選択枝が記載されている最下段のコードをORします。この検索によると、所望のレコードをヒットしますが、ノイズも検索されます。
 (iii-c) 選択枝によって異なる構造式に対して、完全に別々の検索式を作成し、最終の回答をORします。

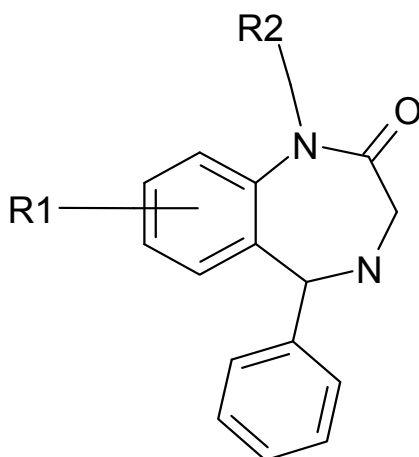
ジェネリックフラグメンテーションコード

検索に含むことが可能なジェネリックフラグメンテーションコードがあります。ジェネリックフラグメンテーションコードは、特許中の一般情報を表わすために索引されています。網羅的な検索をする場合、検索式にジェネリックフラグメンテーションコードを含みます。

注) TOPFRAGはジェネリックフラグメンテーションコードを発生しないことが多いので、必要な場合にはマニュアルでジェネリックフラグメンテーションコードを追加しなければなりません。

コード	定義
A100	Alkali metal, general (Black)
A200	Alkaline earth metal, general (Green)
A300	Group IIIA-VA, general (Green)
A400	First transition metal series, general (Black)
A500	Second transition series, general (Black)
A600	Third transition series, general (Black)
A700	Lanthanides, general (Black)
A800	Actinides, general (Red)
B000	Noble gases, general (Black)
C000	Halogen, general (Green)
D010	Non-specific substitution on the heterocyclic ring in a fused system (Green)
D020	Non-specific substitution on the carbocyclic ring in a fused system (Green)
D040	Fused ring heterocycle, general (Black)
F010	Non-specific substitution on a mononuclear heterocycle (Green)
F020	Aromatic mononuclear heterocycle, general (Black)
F021	Non-aromatic mononuclear heterocycle, general (Black)
G001	Non-specific substitution on benzene (Green)
G002	Non-specific substitution on fused aromatic ring (Green)
G003	Non-specific substitution on fused or unfused alicyclic ring (Green)
G040	Aromatic carbocycle, general (Black)
G050	Unfused alicyclic ring, general (Black)
G051	Fused alicyclic ring system, general (Black)
H600	Halogen, general (Black)

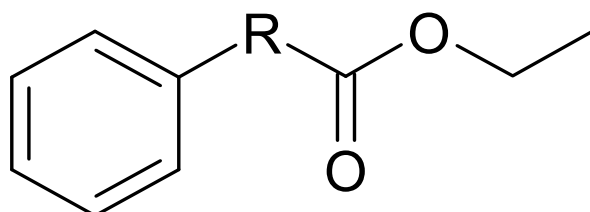
Example A



R1=Halogen
R2=1-6C Alky

Example A**STN****Black:** D780,G100,H211,J521,M412,M531, (H600 or H601 or H602 or H603 or H604),**Red:** M113, M511, M520, M540 (not by Topfrag)**Blue:** M21!, M281, M320, (M270 or M273#)**Ring:** (RIN 01829)**Green:** D014, (D021 or D022), G010, H641,L941

- Using MTF [HAL,VPA; CHK(1C-6C)] (STN)
- S (D780(P)G100(P)H211(P)J521(P)M412(P)M531)/M0,M2,M3,M4
- S L1(P)(H601 OR H602 OR H603 OR H604)/M0,M2,M3,M4
- S L2(P)(M113(P)M511(P)M520)/M2,M3,M4
- S L3(P)(M281(P)M320(P)M210(P)(M270 OR M273))/M2,M3,M4
- S L4(P)01829/RIN
- S L5(P)(D014(P)G010(P)H641(P)"L941"(P)(D021 OR D022))/M2,M3,M4
- S (L2(P)M900/M0) OR (L3(P)M901/M2,M3,M4) OR (L5(P)M902/M2,M3,M4)
- S L7 OR L6
- S L8(NOTP)(H1 OR H3 OR H4 OR H5 OR H7 OR H8 OR H9 OR J0 OR J1 OR J2)/M2,M3,M4
- S L9(NOTP)(J3 OR J4 OR J6 OR J9 OR K1 OR K2 OR K3 OR K4 OR K5 OR K6)/M2,M3,M4
- S L10(NOTP)(K7 OR K8 OR K9 OR "L1" OR "L2" OR "L3" OR "L4" OR "L5")/M2,M3,M4
- S L11(NOTP)("L6" OR "L7" OR "L8")/M2,M3,M4

Example C

R=O or NH

Example C

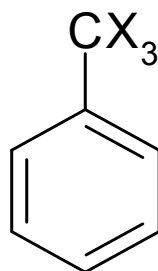
- Black: L472, (L460, L462#) ←
- Red:
- Blue
- Green:

R=Oの場合
カーボネートは黒コード (L472) で検索されます。

R=NHの場合
ウレタンは緑コード (L462) で検索されます。
コードL462には#が付いているので、先行コードL460
(黒コード) でも検索することができます。
黒コードの行でORします。
(L472 OR L462 OR L460)

- 参考:
- Using MTF [R=G1=O or N] (Dialog)
- S M0,M2,M3=(G100(S)M414(S)M531(S)(L460+L462+L472))
- S S1(S)M2,M3=(M212(S)M281(S)M320(S)(M270+M272))
- S S2(S)M2,M3=G010
- S (S1(S)M0=M900)+(S1(S)M2,M3=M901)+(S2(S)M2,M3=M902)+S3
- S S4(NOT S)M2,M3=(H1+H2+H3+H4+H5+H6+H7+H8+H9+J0+J1+J2+J3+J4+J5+J6)
- S S5(NOT S)M2,M3=(J9+K1+K2+K3+K4+K5+K6+K7+K8+K9+L1+L2+L3+L5+L6+L7)
- S S6(NOT S)M2,M3=(L8+L9+M1)

Example D



X=F, Cl, Br, I

Example D

黒コード: G100、(H600、H602、H603、H604)、H685、M414、M531 (CF3は、H601では索引されていません。)

赤コード: 適応コード無し

青コード: M280、M311、M321、M344、M391 (炭素数1の炭素鎖は、M332では索引されていません。)

緑コード: H686、M353#

置換基-CX3は、黒コードH685(-CF3)、緑コードH686(-CX3、ただしXは塩素、臭素、ヨウ素)を使用して検索することができます。

(iii-a) 可変基のコードを削除します。

この例の場合、可変基のコードは唯一の官能基を定義するものなので、可変基のコードを削除する方法はお勧めできません。

(iii-b) 可変基のコードを緑の行でORします。

検索は適切に行われますが、検索結果にはノイズが含まれます。

(iii-c) Xがフッ素の場合と、塩素、臭素、ヨウ素の場合とで別々の検索式を作成します。

最も正確な結果を得ることができます。

完全な回答セットを作成するために下記の検索式をORします。(ネグーションコードは省略しています)

Xがフッ素の場合

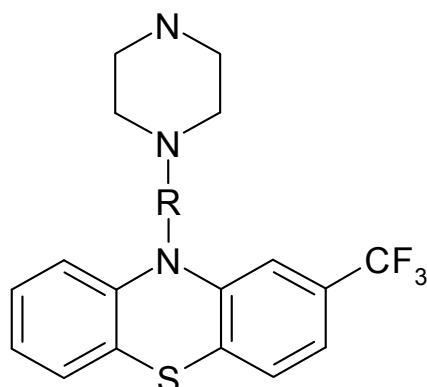
S (G100(P)H685(P)M414(P)M531)/M0,M2,M3,M4
 S L1
 S L2(P)(M280(P)M311(P)M321(P)M344(P)M391(P)(M353 OR M350))/M2,M3,M4
 S L3(P)G010/M2,M3,M4
 S (L1(P)M900/M0) OR (L2(P)M901/M2,M3,M4) OR (L3(P)M902/M2,M3,M4) OR L4

Xが塩素、臭素、ヨウ素の場合

S (G100(P)(H600 OR H602 OR H603 OR H604)(P)M414(P)M531)/M0,M2,M3,M4
 S L6
 S L7(P)(M280(P)M311(P)M321(P)M344(P)M391(P)(M353 OR M350))/M2,M3,M4
 S L8(P)(G010(P)H686)/M2,M3,M4
 S (L6(P)M900/M0) OR (L7(P)M901/M2,M3,M4) OR (L8(P)M902/M2,M3,M4) OR L9

S L5 OR L10 (Xがフッ素の場合 と Xが塩素、臭素、ヨウ素の場合を合わせる)

Example F



R = direct bond, -O- or 2-6C alkylene

Example F

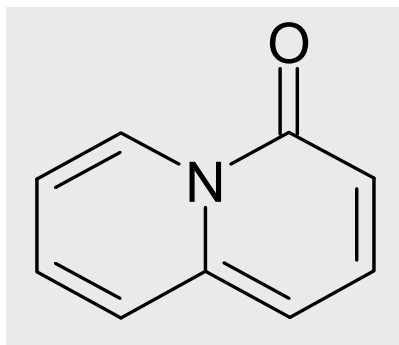
- Rが結合の場合、-N-N-結合は下記のコードで検索されます。
- H212 (黒コード)
- (K640 OR K600) (先行コードが黒の緑コード)
-
- Rが-Oの場合、-N-O-N-結合は下記のコードで検索されます。
- H212 (黒コード)
- (K810 OR K800) (先行コードが黒の緑コード)
- Rが炭素数2～6個のアルキレンの場合、-N-アルキレン-N-は下記のコードで検索されます。
- H182, H202 (黒コード)
- (M312 OR M313 OR M314 OR M315), M342 (青コード)
- 表を参照すると、Rが結合または-O-の場合は第1行(黒コードの行)でORできますが、Rが炭素数2～6個のアルキレンで青コードである場合、別々の検索式を作成して検索するのが最もよい方法です。
-
- Rが結合または-O-の場合(STN):
- =>S ("E800"(P)F553(P)H212(P)H685(P)M412(P)((K600 OR K640) OR (K800 OR K810)))/M0,M2,M3,M4 ¥>_line1
- =>S _line1(P)(M511(P)M521(P)M530(P)M540)/M2,M3,M4 ¥>_line2
- =>S _line2(P)(M280(P)M311(P)M321(P)M344(P)M391(P)(M350 OR M353))/M2,M3,M4 ¥>_line3
- =>S _line3(P)(D011(P)D022(P)F011)/M2,M3,M4 ¥>_line4
- =>S (_line1(P)M900/M0) OR (_line2(P)M901/M2,M3,M4) OR (_line3(P)M902/M2,M3,M4) ¥>_line5
- =>S _line5 OR _line4 ¥>_line6
-
- Rが炭素数2～6個のアルキレンの場合(STN):
- =>S ("E800"(P)F553(P)H182(P)H202(P)H685(P)M412)/M0,M2,M3,M4 ¥>_line1
- =>S _line1(P)(M511(P)M521(P)M530(P)M540)/M2,M3,M4 ¥>_line2
- =>S _line2(P)(M280(P)M311(P)(M312 OR M313 OR M314 OR M315) (P)M342(P)M344(P) M391)/M2,M3,M4 ¥>_line3
- =>S _line3(P)((M350 OR M353)(P)(M380 OR M383))/M2,M3,M4 ¥>_line4
- =>S _line4(P)(D011(P)D022(P)F011)/M2,M3,M4 ¥>_line5
- =>S (_line1(P)M900/M0) OR (_line2(P)M901/M2,M3,M4) OR (_line4(P)M902/M2,M3,M4) ¥>_line6
- =>S _line6 OR _line5 ¥>_line7

1. 化学構造式にコードを付与してみる
2. Chemical Fragmentation Codeを修正する
3. DWPIで構造検索 (Chemical Fragmentation Codes と DCR)
4. コード修正のまとめ
5. Multi Files, Multi Toolsを利用する検索参考例

課題2:

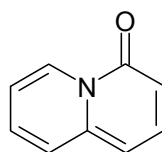
下記化学構造を含む発明を調査

(STN上で行う例)



2006.08.20

I : STN Expressで構造式を書いて、コードを発生する。



```

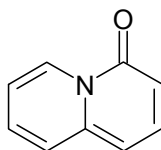
=>S (D660 (P) J521 (P) M412) /M0,M2,M3 ¥>_line1
=>S _line1 (P) (M511 (P) M520 (P) M530 (P) M540) /M2,M3 ¥>_line2
=>S _line2 (P) (M280 (P) M320) /M2,M3 ¥>_line3
=>S _line3 (P) 01686/RIN ¥>_line4
=>S _line4 (P) (D011 (P) "L941") /M2,M3 ¥>_line5
=>S (_line1 (P) M900/M0) OR (_line2 (P) M901/M2,M3) OR (_line4 (P) M902/M2,M3) ¥>_line6
=>S _line6 OR _line5 ¥>_line7
=>S _line7 (NOTP) (H1 OR H2 OR H3 OR H4 OR H5 OR H6 OR H7 OR H8 OR H9 OR J0) /M2,M3 ¥>_line8
=>S _line8 (NOTP) (J1 OR J2 OR J3 OR J4 OR J6 OR J9 OR K1 OR K2 OR K3 OR K4) /M2,M3 ¥>_line9
=>S _line9 (NOTP) (K5 OR K6 OR K7 OR K8 OR K9 OR "L1" OR "L2" OR "L3" OR "L4") /M2,M3 ¥>_line10
=>S _line10 (NOTP) ("L5" OR "L6" OR "L7" OR "L8" OR M1) /M2,M3 ¥>_line11

```

L14 49 Hits

L21 7 Hits (1999-2006/PY.B)

II : 発生されたコードを編集



```

• =>S (D660(P)M412)/M0,M2,M3 ¥>_line1
• =>S _line1 ¥>_line2
• =>S _line2 ¥>_line3
• =>S _line3(P)01686/RIN ¥>_line4
• =>S _line4(P) "L941"/M2,M3 ¥>_line5
• =>S (_line1(P)M900/M0) OR (_line2(P)M901/M2,M3) OR (_line4(P)M902/M2,M3) ¥>_line6
• =>S _line6 OR _line5 ¥>_line7

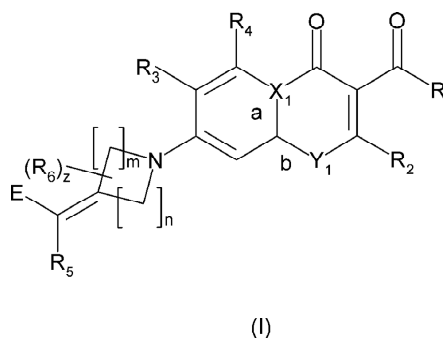
```

L21 263 Hits

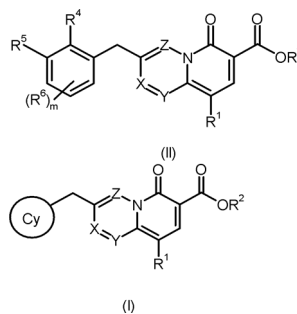
L32 84 Hits (1999-2006/PY.B)

編集前, 得られなかったヒット (214 Hits) –Examples I ...

- AN 2006-510201 [52] WPIX [Full-text](#)
- TI New 7-amino alkylidenyl-heterocyclic quinolone and naphthyridone compounds used to treat or prevent condition caused by or contributed to by bacterial infection.
- PA (GRAN-I) GRANT E B; (MACHI-I) MACIELAG M J; (PAGE-I) PAGET S D; (WEID-I)
- PI US 2006052359 A1 20060309 (200652)* 75



- AN 2006-263360 [27] WPIX [Full-text](#)
- TI New quinolizinone derivatives are HIV integrase inhibitors, useful for preventing AIDS and treating HIV infection.
- PA (ARAM-I) ARAMAKI H; (INOI-I) INOUE M; (KAWA-I) KAWAKAMI H; (MATS-I) MATSUZAKI Y; (NAKA-I) NAKAMURA H; (SATO-I) SATOH M; (SHIN-I) SHINKAI H; (YAMA-I) YAMATAKA K; (NISB) JAPAN TOBACCO INC
- PI WO 2006033422 A1 20060330 (200627)* JA 211



編集前, 得られなかったヒット(214 Hits) –Examples II ...

- AN 2004-238715 [22] WPIX [Full-text](#)
- TI New quinolone derivatives used for treating microbial infections.
- PA (PROC) PROCTER & GAMBLE CO
- PI WO 2004014893 A2 20040219 (200422)* EN 147 C07D401-04

(I)

- AN 2003-783897 [74] WPIX [Full-text](#)
- TI Pharmaceutical containing (new) quinolizone compound useful as antimicrobial for treating infections with Gram-positive and Gram-negative bacteria, mycoplasma, chlamydia and resistant microbes.
- PA (SANY) SANKYO CO LTD; (UBEI) UBE IND LTD
- PI JP 2003261566 A 20030919 (200374)*

(I) (a) (b)

編集前, 得られなかったヒット(214 Hits) –Examples III ...

- AN 2003-486265 [46] WPIX [Full-text](#)
- TI Preparation of heterocyclic ring-containing 4-oxoquinolizone compounds, useful against Gram positive or negative and anaerobic microorganisms, comprises reaction of 4-oxoquinolizones and heterocyclic compounds.
- PA (SATO) SATO SEIYAKU KK
- PI JP 2003012670 A 20030115 (200346)* 15

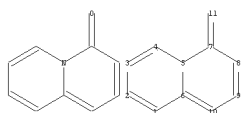
(I) (II) (III)

- AN 1999-619697 [53] WPIX [Full-text](#)
- TI New heterocyclic derivatives are broad spectrum antibacterial agents, useful for treating bacterial and fungal infections.
- PA (ABBO) ABBOTT LAB
- PI US 5977133 A 19991102 (199953)* 11

(I)

III : DCR on STNで検索

=> Uploading C:\YSTNEXP\Queries\4-oxoquinolizine_DCR.str



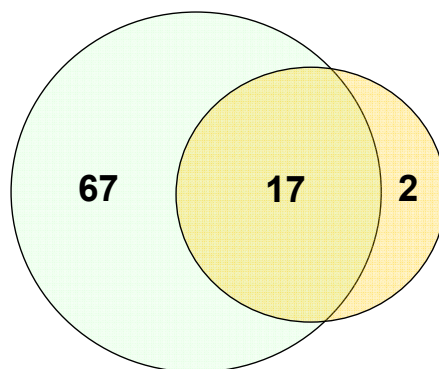
```

chain nodes :
11
ring nodes :
1 2 3 4 5 6 7 8 9 10
chain bonds :
7-11
ring bonds :
1-2 1-6 2-3 3-4 4-5 5-6 5-7 6-10 7-8 8-9 9-10
exact/norm bonds :
1-2 1-6 2-3 3-4 4-5 5-6 5-7 6-10 7-8 7-11 8-9 9-10
Match level :
1:Atom 2:Atom 3:Atom 4:Atom 5:Atom 6:Atom 7:Atom 8:Atom 9:Atom 10:Atom
11:CLASS
L1 STRUCTURE UPLOADED
=> s L1 full
FULL SEARCH INITIATED 11:24:05 FILE 'WPIX'
FULL SCREEN SEARCH COMPLETED - 10758 ITERATIONS
100.0% PROCESSED 10758 ITERATIONS 210 ANSWERS
SEARCH TIME: 00.00.06
L2 210 SEA sss FUL L1
=> s L2/dcr
L3 19 L2/DCR

```

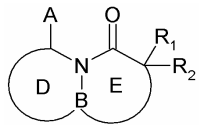
L2 210 Hits (DCR)**L3 19 Hits (DWPI)**

IV : DCR と Chemical Fragmentation Codesの結果

**84 Hits (1999-2006/PY.B)****19 Hits (DCR ->DWPI)**

Codesでしかヒットしなかった (67 Hits) –Examples ...

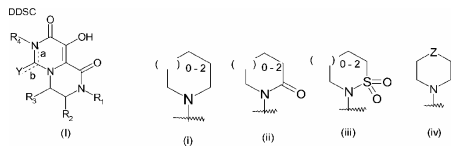
- AN 2003-618069 [58] WPIX [Full-text](#)
- TI New bicyclic lactam derivatives are matrix metalloproteinase inhibitors useful for the treatment of e.g. allergy, anorexia, asthma, atherosclerosis, Behcet's disease.
- PI WO 2003055856 A2 20030710 (200358)* EN



(I)

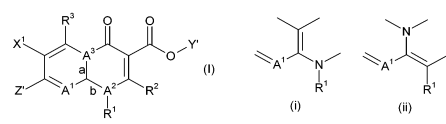
- AN 2005-746986 [76] WPIX [Full-text](#)
- TI New bicyclic uracil derivatives as HIV integrase inhibitor useful for treating and preventing HIV infection and AIDS.
- PI WO 2005092099 A1 20051006 (200576)* EN 135
- PA (MERI) MERCK & CO INC

DDSC



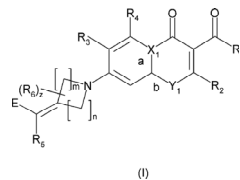
(i) (ii) (iii) (iv)

- AN 2005-435681 [45] WPIX [Full-text](#)
- TI Anti acid-fast bacterial agent useful for treating infection caused by Gram-positive and Gram-negative bacteria e.g. Mycobacterium tuberculosis comprises pyridone carboxylic acid derivatives.
- PI AU 2004240167 A1 20050113 (200545)*
- PA (DAUC) DAIICHI PHARM CO LTD



(i) (ii)

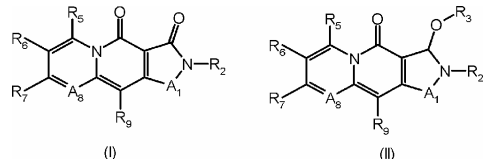
- AN 2006-510201 [52] WPIX [Full-text](#)
- TI New 7-amino alkylidene-heterocyclic quinolone and naphthyridone compounds used to treat or prevent condition caused by or contributed to by bacterial infection.
- PA (GRAN-I) GRANT E B; (MACI-I) MACIELAG M J; (PAGE-I) PAGET S D; (WEID-I)
- PI US 2006052359 A1 20060309 (200652)*



(I)

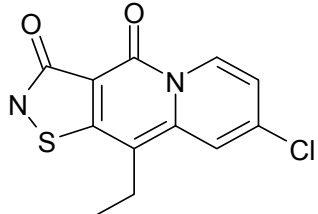
DCRLしかヒットしなかった (2 Hits) –Example 1 ...

- AN 2006-522238 [53] WPIX [Full-text](#)
- TI New 1-thia-2,4a-diaza-cyclopenta(b)naphthalene-3,4-dione useful for treating or preventing bacterial or protozoal infections e.g. urinary tract infection.
- PI WO 2006074317 A1 20060713 (200653)* EN 91 C07D513-00



(I) (II)

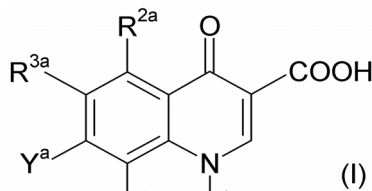
- CMC UPB 20060817
- M2 *27* D011 D012 D013 E850 H6 H602 H621 J5 J522 L9 L941 L999 M210 M212 M240 M281 M320 M412 M511 M520 M530 M540 M710 M904 M905 P001 P200 P210 P220 P241 P310 P420 P714 P723 P735 P820 P941 P943 R032
- DCN: **RANIDG-N**
- DCR Record:
- DCN: **RANIDG**



DCRLしかヒットしなかった(2 Hits) –Example 2 ...

- AN 2004-286224 [27] WPIX [Full-text](#)
- TI O/W type emulsion used in production of microcapsules used for treating periodontal disease comprises organic solvent dispersion liquid in which oil phase contains pyridone carboxylic acid compound, in vivo degradable polymer and zinc oxide.

PI JP 2003300882 A 20031021 (200427)*



- CMC UPB 20040611
- M1 *35* C017 C100 C800 C801 C803 C804 C805 C806 C807 D014 D019 **D660** F011 F013 F433 G030 G530 H1 H100 H122 H2 H201 H6 H601 H621 J0 J011 J1 J111 J5 J521 L9 **L941** M1 M116 M210 M211 M240 M281 M320 **M411** M423 M431 M511 M521 M530 M541 M640 M782 M904 M905 R022 R033

RIN: 01686

DCN: RAD41A-K; RAD41A-M

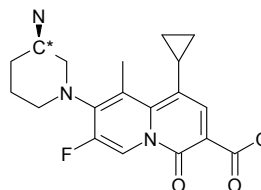
DCR Record:

AN.S DCR-158998

DCSE 158998-1-1-0

CN.P A-99058-1

CN.S 8-(3-Amino-piperidin-1-yl)-1-cyclopropyl-7-fluoro-9-methyl-4-oxo-4H-quinolizine-3-carboxylic acid; **hydrochloride**



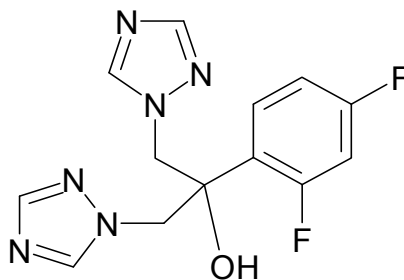
No M412 but M411

課題3:

抗真菌剤 Fluconazoleの誘導体に関する特許を包括的に検索

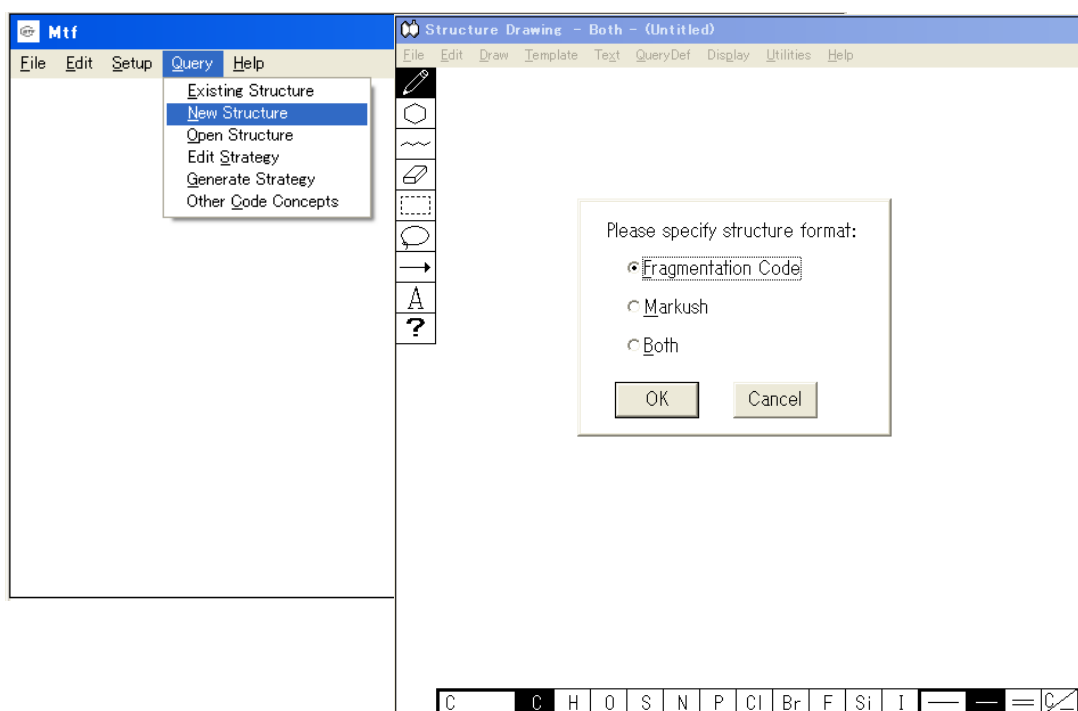
(Dialog上で行う例)

構造式

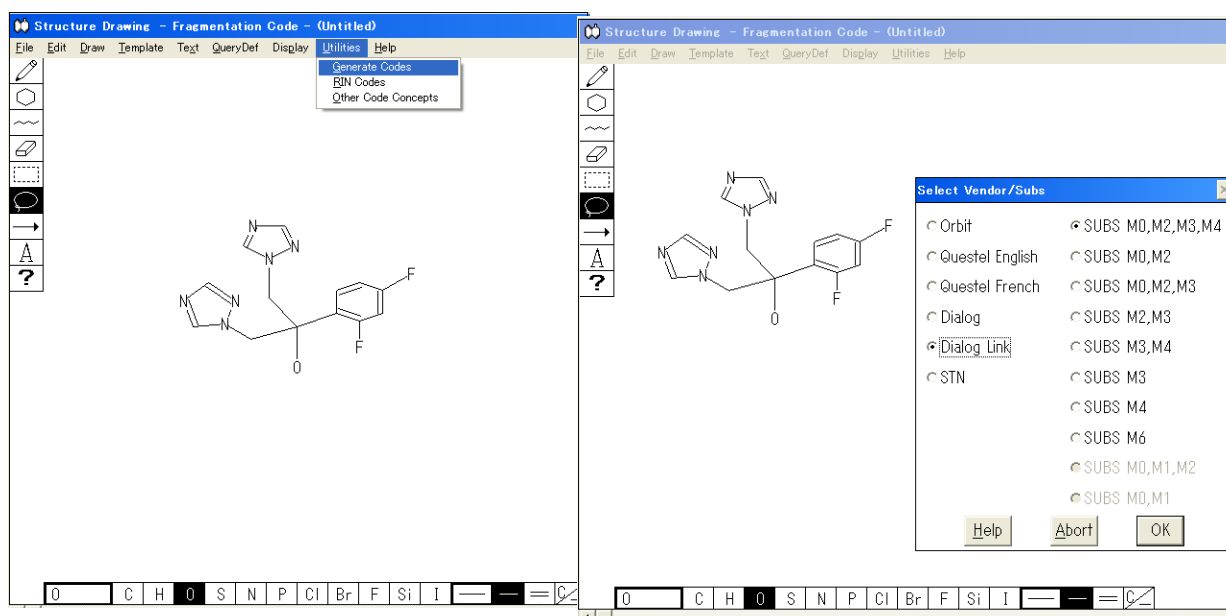


基本特許: US Pat. 4,404,216 (Pfizer)

Markush Topfragで構造式を作図して検索式を発生---1



Markush Topfragで構造式を作図して検索式を発生---2



Fluconazole (MTF:Fragcode 未修正)

S M0,M2,M3,M4=(F570(S)F599(S)G100(S)H212(S)H401(S)H481(S)H601(S)H608(S)M413)

S S1(S)M0,M2,M3,M4=M531

S S2(S)M2,M3,M4=M522

S S3(S)M2,M3,M4=(M280(S)M313(S)M321(S)M332(S)M344(S)M391(S)(M370+M373))

S S4 AND RR=00096

S S5(S)M2,M3,M4=(F011(S)F019(S)G015(S)H642)

S (S2(S)M0=M900)+(S3(S)M2,M3,M4=M901)+(S5(S)M2,M3,M4=M902)+S6

S S7(NOT S)M2,M3,M4=(H1+H3+H5+H7+H9+J0+J1+J2+J3+J4+J5+J6+J9+K0+M1)

Fluconazole (MTF:Fragcode 修正へ)

S M0,M2,M3,M4=(F570(S)F599(S)G100(S)H212(S)H401(S)H481(S)H601(S)H608(S)M413)

S S1(S)M0,M2,M3,M4=M531

S S2(S)M2,M3,M4=M522

S S3(S)M2,M3,M4=(M280(S)M313(S)M321(S)M332(S)M344(S)M391(S)(M370+M373))

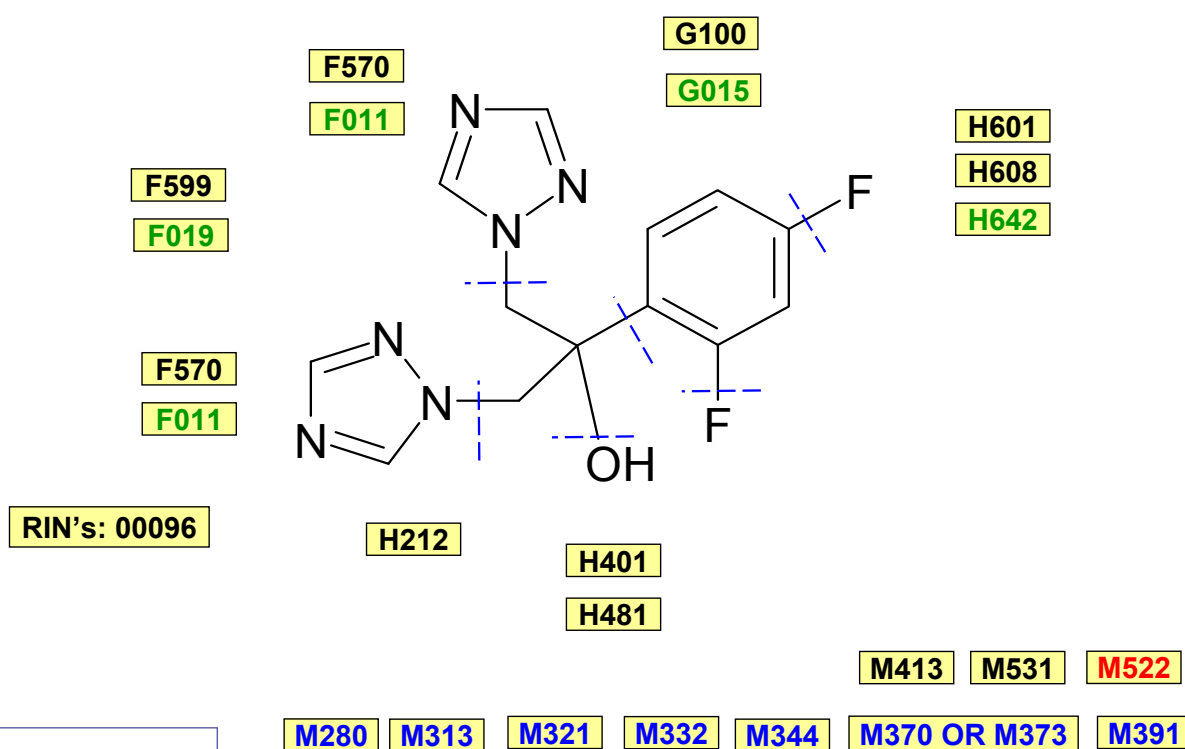
S S4 AND RR=00096

S S5(S)M2,M3,M4=(F011(S)F019(S)G015(S)H642)

S (S2(S)M0=M900)+(S3(S)M2,M3,M4=M901)+(S5(S)M2,M3,M4=M902)+S6

S S7(NOT S)M2,M3,M4=(H1+H3+H5+H7+H9+J0+J1+J2+J3+J4+J5+J6+J9+K0+M1)

Fluconazole



Fluconazole (Fragcode 修正後、例)

S M0,M2,M3,M4=(F570(S)F599(S)H212(S)H601(S)H608(S)(M412 OR M413))

S S1

S S2

S S3

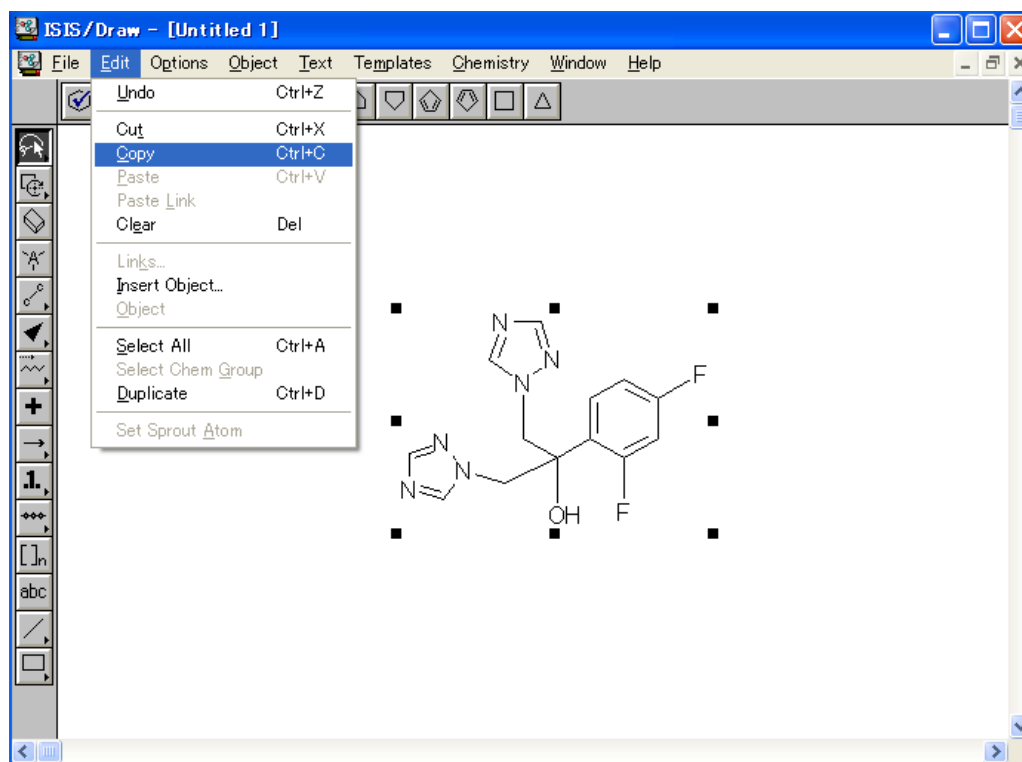
S S4 AND RR=00096

S S5(S)M2,M3,M4=(F011(S)F019(S)G015)

S (S2(S)M0=M900)+(S3(S)M2,M3,M4=M901)+(S5(S)M2,M3,M4=M902)+S6

保存して、その後使う

ISISで作図、構造式をコピー



DialogLink 5にペーストして, DCR (File 355)で検索する

DialogLink 5.0 - [Dialog Session]

File Edit View Session Configuration Window Help

Help and Infor
DialogLink Mats
Set Links

Command Previous Reports Chemical Structure

Launch Drawing Program

Define Search

Choose Search Type:
Substructure

BEGIN CSS Collection(s):

- 390: Beilstein Facts
- 389: (ONTAP Beilstein Facts)
- 355: Derwent Chemistry Resource
- 445: IMS R&D Focus
- 447: IMS Patent Focus
- 947: IMS Patent Focus (Subscribers)
- 955: IMS R&D Focus (Subscribers)
- 302: Index Chemicus
- 128: Pharmaprojects
- 928: Pharmaprojects (Subscribers)
- 452: Prous Drug Data Report
- 453: Prous Drugs of the Future

Disconnected

DCRで構造検索

Begin 355

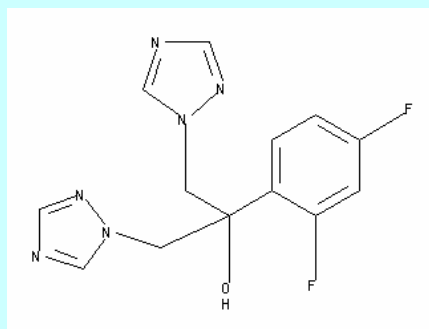
[File 355] **Derwent Chemistry Resource** UD=200674

(c) 2006 The Thomson Corporation. All rights reserved.

CSS Structure 2006.11.27 11.57.29.125.mol As SSS

S1 30 CSS STRUCTURE 2006.11.27 11.57.29.125.MOL AS SSS

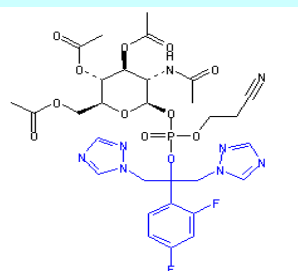
? Query Molecule:



? t s1/19/1,10,20,30

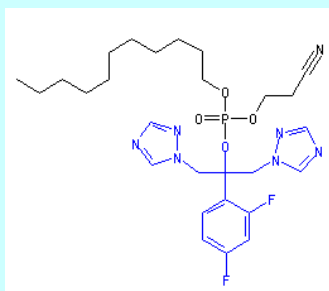
0001025045

DCR Number: 1025045
 External DCR Number: 1025045-1-0-0
 Specific Compound Number: RAGSPO
 Systematic Chemical Name: Acetic acid
 3,4-diacetoxy-5-acetylamino-6-((2-cyano-ethoxy)-[1-(2,4-difluoro-phenyl)-2-[1,2,4]triazol-1-yl-1-[1,2,4]triazol-1-ylmethyl-ethoxy]-phosphoryloxy)-tetrahydro-pyran-2-ylmethyl ester
 Molecular Weight: 768.6305
 Molecular Formula (Orig): C30 H35 F2 N8 O12 P
 Molecular Formula (Comp): C30H35F2N8O12P
 Structured Molecular Formula
 Fragment and Multiplier: C30H35F2N8O12P 1
 Number of Fragments: 1
 Total Fragments: 1
 Ring Index Numbers: 00096
 Substance Descriptor: CARBOHYDRATES



001025036

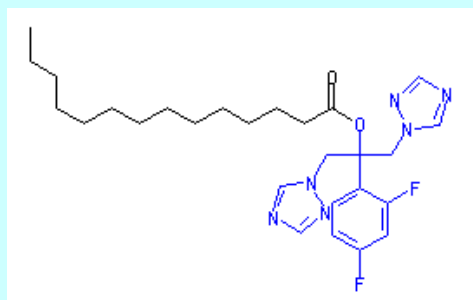
DCR Number: 1025036
 External DCR Number: 1025036-0-0-0
 Specific Compound Number: RAGSPF
 Systematic Chemical Name: Phosphoric acid 2-cyano-ethyl ester
 1-(2,4-difluoro-phenyl)-2-[1,2,4]triazol-1-yl-1-[1,2,4]triazol-1-ylmethyl-ethyl ester undecyl ester
 Molecular Weight: 593.619
 Molecular Formula (Orig): C27 H38 F2 N7 O4 P
 Molecular Formula (Comp): C27H38F2N7O4P
 Structured Molecular Formula
 Fragment and Multiplier: C27H38F2N7O4P 1
 Number of Fragments: 1
 Total Fragments: 1
 Ring Index Numbers: 00096



? t s1/19/1,10,20,30

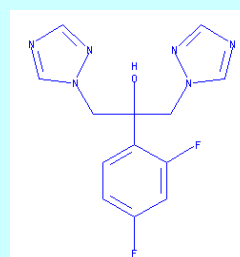
0001025023

DCR Number: 1025023
 External DCR Number: 1025023-0-0-0
 Specific Compound Number: RAGSP2
 Systematic Chemical Name: Tetradecanoic acid
 1-(2,4-difluoro-phenyl)-2-[1,2,4]triazol-1-yl-1-[1,2,4]triazol-1-ylmethyl-ethyl ester
 Molecular Weight: 516.6397
 Molecular Formula (Orig): C27 H38 F2 N6 O2
 Molecular Formula (Comp): C27H38F2N6O2
 Structured Molecular Formula
 Fragment and Multiplier: C27H38F2N6O2 1
 Number of Fragments: 1
 Total Fragments: 1
 Ring Index Numbers: 00096
 Substance Descriptor: FATTY ACIDS



0000095294

Preferred Name: FLUCONAZOLE
 DCR Number: 95294
 External DCR Number: 95294-0-0-0
 Specific Compound Number: R20553
 DDF Structure ID: FLUCONAZO
 Systematic Chemical Name: 2-(2,4-Difluoro-phenyl)-1,3-bis-[1,2,4]triazol-1-yl-propan-2-ol
 Synonyms: ARNAZOLE; BIOZOLENE; DIFLUCAN; DIFLUCAN-GINECOLOGICO; ELAZOR; FLUCONAZOLE; FUNGATA; MYCOSYST; QUAMATEL; TRIFLUCAN; UK-49858
 Molecular Weight: 306.2748
 Molecular Formula (Orig): C13 H12 F2 N6 O
 Molecular Formula (Comp): C13H12F2N6O
 Structured Molecular Formula
 Fragment and Multiplier: C13H12F2N6O 1
 Number of Fragments: 1
 Total Fragments: 1
 Ring Index Numbers: 00096
 DDF Descriptors
 Activity: FUNGICIDES
 Substructure: AMINOALCOHOL; ARALKYLAMINE; TRIAZOLE; ARYLFLUORIDE
 Mechanism: CYTOCHROME-P450-INHIBITOR; active against: Cryptococcus neoformans.; highly selective inhibitor of sterol C-14 alpha-demethylation



DCR (File 355) から DWPI (File 350)へ

? MAP DN T S1

Temp SearchSave "TC003" stored
3 Select Statements, 30 Search Term(s)
SearchSave TC003

DCR番号を抽出して一時的保存"TC003"

1 SearchSave(s), 30 Search Term(s)

? B 350

[File 350] Derwent WPIX 1963-2006/UD=200675 (c) 2006 The Thomson Corporation. All rights reserved.

S1 3198 S M0,M2,M3,M4=(F570(S)F599(S)H212(S)H601(S)H608(S)(M412 OR M413))

S2 3198 S S1

S3 3198 S S2

S4 3198 S S3

S5 2149 S S4 AND RR=00096

S6 1816 S S5(S)M2,M3,M4=(F011(S)F019(S)G015)

S7 1922 S (S2(S)M0=M900)+(S3(S)M2,M3,M4=M901)+(S5(S)M2,M3,M4=M902)+S6

修正後Frag code で得られたヒット

DCR (File 355) から DWPI (File 350)へ

? EXS TC003

EXS: S DN=1025007 + DN=1025022 + DN=1025023 + DN=1025024 + DN=1025025 +
DN=1025026 + DN=1025027 + DN=1025028 + DN=1025029 + DN=1025030 + DN=1025031 +
DN=1025032 + DN=1025036 + DN=1025037 + DN=1025038 + DN=1025039 + DN=1025040 +
DN=1025041

1 DN=1025007

(省略)

1 DN=1025041

S8 1 S DN=1025007 + DN=1025022 + DN=1025023 + DN=1025024 + DN=1025025 +
DN=1025026 + DN=1025027 + DN=1025028 + DN=1025029 + DN=1025030 + DN=1025031 +
DN=1025032 + DN=1025036 + DN=1025037 + DN=1025038 + DN=1025039 + DN=1025040 +
DN=1025041

EXS: S DN=1025042 + DN=1025043 + DN=1025044 + DN=1025045 + DN=369997 +
DN=370000 + DN=426859 + DN=426860 + DN=426872 + DN=786792 + DN=838611 +
DN=95294

1 DN=1025042

(省略)

1 DN=838611

366 DN=95294

S9 370 S DN=1025042 + DN=1025043 + DN=1025044 + DN=1025045 + DN=369997 +
DN=370000 + DN=426859 + DN=426860 + DN=426872 + DN=786792 + DN=838611 +
DN=95294

EXS: S S8:S9

DCR をDWPIへトランスファーした結果

S10 370 S S8:S9

DCR (File 355) から DWPI (File 350)へ

S11	2278	S S7 OR S10
S12	1908	S S7 NOT S10
S13	356	S S10 NOT S7
S14	14	S S7 AND S10
S15	1	S PN=US 4404216
S16	1	S S7 AND S15
S17	0	S S10 AND S15
S18	1	S S11 AND S15
S19	1	S S12 AND S15
S20	0	S S13 AND S15
S21	0	S S14 AND S15
S23	10988	S MC=B14-A04?
S24	9096	S MC=B12-A02?
S25	20083	S S23 OR S24
S26	328	S S11 AND S25
S27	1	S S26 AND S15

DCR + Fragcode結果

DCRと Fragcodeの結果の関係

Fluconazole 基本特許

Manual Code

Manual Codeで絞る

基本特許が集合に入っている

B14-A04 Antifungal general and other
1994

B14-A04A . Aspergillus
1994

Previous code(s): B12-A02C

B14-A04B . Candida
This organism commonly causes thrush.
1994

Previous code(s): B12-A02C

B14-A04C . Trichophyton, Microsporium
This code covers treatment of e.g. ringworm, tinea, Athlete's foot.
1994

Previous code(s): B12-A02C

B12-A02 Antifungal, antialgal, antilichen general
1963-1993

Now coded as: B14-A04+,

B12-A02A . Antialgal
1986-1993

B12-A02B . Antilichen
1986-1993

Now coded as: B14-B08

B12-A02C . Antifungal
1986-1993

Now coded as: B14-A04+

B12-A03 Antileptotic
1963-1993

Fluconazole 特許のDWPI レコード

- WPI Acc no: 1982-07498J/198250
2-Difluorophenyl-1,3-bis-triazolyl-propan-2-ol - having antifungal activity
Patent Assignee: PFIZER LTD (PFIZ)
Inventor: RICHARDSON K

- Patent Family (23 patents, 24 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
GB 2099818	A	19821215	GB 198117379	A	19810606	198250	B
US 4404216	A	19830913	US 1982383866	A	19820601	198339	E
.....							
JP 1990033691	B	19900730	JP 198297494	A	19820607	199034	E

- Alerting Abstract** GB A

2-(2,4-Difluorophenyl)-1,3-bis (1H-1,2,3-triazol-1-yl)-propan-2-ol of formula (I) and its salts are new.

(I) has antifungal activity and may be used to treat fungal infections in humans and animals. Unlike the corresponding 2,4-dichloro cpd., (I) is not teratogenic. (I) is useful in treating topical fungal infections in man such as those caused by species of Candida, Trichophyton, Microsporium or Epidermophyton, mucosal infections caused by Candida albicans, and systemic infections caused by C.albicans, Cryptococcus neoformans, Aspergillus fumigatus, Coccidioides, Paracoccidioides, Histoplasma or Blastomyces. In tests on mice, infected with C.albicans (I) gives at least 50% protection at dosages below 0.5 mg/kg (p.o. or i.v.). Daily dosages of (I) are generally in the range 0.1-5 mg/kg a divided doses (oral or parenteral). (I) may also be applied topically or in the form of a suppository or pessary.

Title Terms /Index Terms/Additional Words: DI; FLUOROPHENYL; TRIAZOLYL; PROPANE; OL; ANTIFUNGAL; ACTIVE; PROPANOL

- File Segment: GPI
DWPI Class: B03; C02
Manual Codes (CPI/A-N): B07-D13; **B12-A02**; B12-B05; C07-D13; C12-A02; C12-B05

Chemical Indexing

Derwent Registry Numbers: 0443-S

- Chemical Fragment Codes (M2):
01 M903 F011 F019 F570 F599 G015 G100 H2 H212 H4 H401 H481 H6 H601 H608
H642 H8 M280 M313 M321 M332 M344 M373 M391 M413 M510 M522 M531 M540
M640 M650 M710 P001 **P241**

1. 化学構造式にコードを付与してみる
2. Chemical Fragmentation Codeを修正する
3. DWPIで構造検索 (Chemical Fragmentation Codes と DCR)
4. コード修正のまとめ
5. Multi Files, Multi Toolsを利用する検索参考例

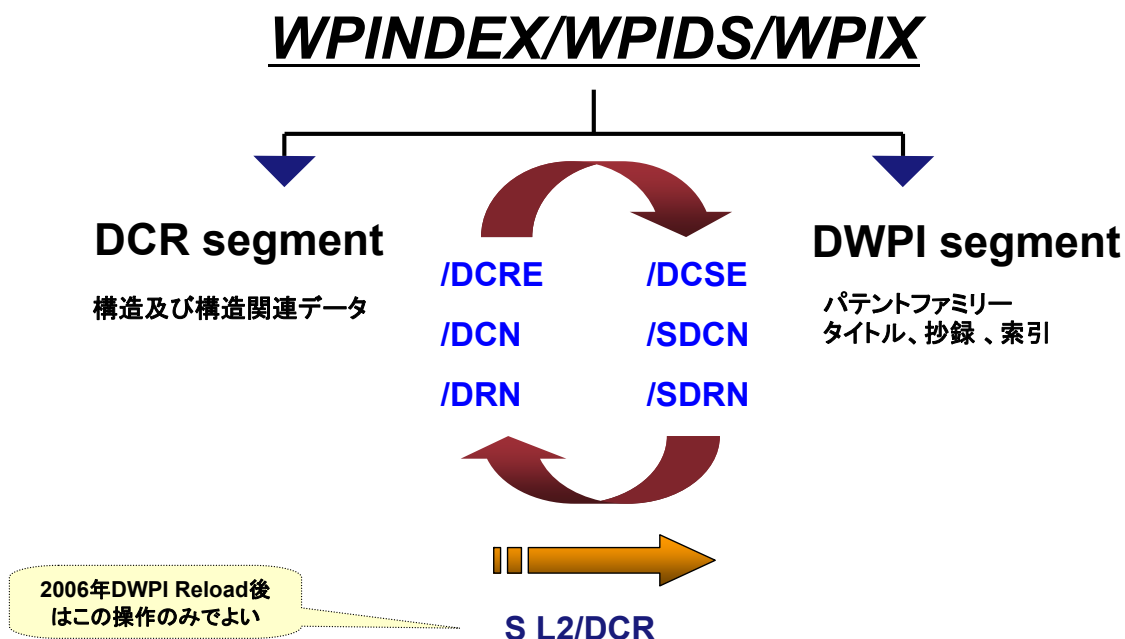
STN DCR セグメント → DWPI セグメント

- **STEP1:** 化学構造検索 (完全一致・部分一致・類似構造 ;
その他、物質名称 (/CN.P)、分子式 (/MF)、分子量 (/MW))



- **STEP2:** ヒットしたDCRレコードの集合(eg.L2)をDWPIセグメントにクロス:
方法:s L2/DCR
該当化合物が関連する特許レコードを特定

STN DCR セグメント → DWPI セグメント



Dialog DCR (File:355) → DWPI (File: 352, 350)

- **STEP1:** 化学構造検索 (完全一致・部分一致・類似構造; DialogLink 5のみ可
その他、物質名称 (NA=)、分子式 (MF=))
- **STEP2:** ヒットしたDCRレコード中の**"DCR Number"**を確認
方法1: 出力 (TYPE) して確認する
方法2: MAP DN T Sn (Snは任意の集合番号) DCR番号 (DN)を抽出
- **STEP3:** DWPI (ファイル352, 350)に接続し、
DCR番号検索を実行 (例. EXS TC003)
該当化合物が関連する特許レコードを特定

化学構造検索は以下のHPをご参照ください:

<http://database.g-search.or.jp/service/dialoglink5/dialoglink5.html>

Substructure検索用のコードに修正に使われる一般的な手法

1. Negation codes, 例えば H1等を削除。
2. 基本構造(Basic Group): M4グループコードをより高い順位のコードを含める。例えば、M411。
3. D01/D02 や G01/G02/G03関連コードを加えて、環上置換位置の許容数を広げる
4. 一般官能基が存在する数を考慮する。例えば、J111, J112。
5. Hydroxyl はether/esterになる可能性はあるか? Amineがamideになるか? 可能であれば、ORするかこれらのコードを削除するか
6. カーボンチェンコード(e.g. M280, M320)を編集・削除。
7. 環系の数関連のコードを考慮する。例えば、M511, M512。
8. 環間コードのオプションM1 codeを考慮する。例えば、M131 (>C=O), M132 (other carbon), など
9. 必要に応じ、ジェネリックコードを加える。例えば、H600, F020。
10. 色違いのコードをORする場合は注意する必要がある。

Chemical Fragmentation Codesの特徴(メリット)

特許からMarkush化学構造が索引されている最も古いシステムのひとつ(1963年)

パテントからMarkushと特定物質両方索引されている

書き難い概念をFragmentation Codeで分かりやすく表現できる場合が多い

他のテキストタームと組み合わせて検索するのに便利

(例: キーワード、IPCやMCなど)

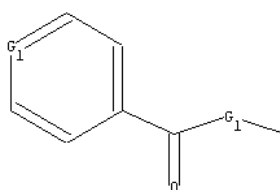
構造検索料がかからない

STN Expressから発生したコードを使う場合(注意すべき)参考例

G1-G1 vs. G1-G2

(ソフトを使ってG-groupを含む構造からコードを発生させる際...)

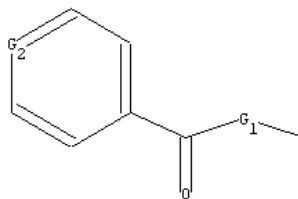
G1-G1 vs. G1-G2 -(1)



G1-G1 を使った場合, (ここでは G1= C or N)

- =>S (M281(P)M320)/M2 ¥>_line1
- =>S M900/M0,M2 OR M901 OR (_line1(P)M902) OR _line1 ¥>_line2
- =>S _line2(P)(F431(P)J311(P)M413)/M0,M2 ¥>_line3
- =>S _line3(P)(M521(P)M530(P)M540)/M2 ¥>_line4
- =>S _line4(P)(M211(P)(M270 OR M273))/M2 ¥>_line5
- =>S _line5(P)(F014(P)J011)/M2 ¥>_line6
- =>S _line2(P)(G100(P)J581(P)M414(P)M531)/M0,M2 ¥>_line7
- =>S _line7(P)(M212(P)(M260 OR M262))/M2 ¥>_line8
- =>S _line8(P)G010/M2 ¥>_line9
- =>S (M900/M0(P)(_line3 OR _line7)) OR (M901/M2(P)(_line4 OR _line7)) ¥>_line10
- =>S _line10 OR (M902/M2(P)(_line5 OR _line8)) OR _line6 OR _line9 ¥>_line11
- =>S _line11(NOTP)(H1 OR H2 OR H3 OR H4 OR H5 OR H6 OR H7 OR H8 OR H9 OR J1)/M2 ¥>_line12
- =>S _line12(NOTP)(J2 OR J4 OR J9 OR K0 OR M1)/M2 ¥>_line13

G1-G1 vs. G1-G2 -(2)



G1, G2 を使った場合(copied form
G1 をコピー又は新作成), ここ
では G1=G2= C or N:

```

=>S (M281(P)M320)/M2 >_line1
=>S M900/M0,M2 OR M901 OR (_line1(P)M902) OR _line1 >_line2
=>S _line2(P)(F431(P)M413)/M0,M2 >_line3
=>S _line3(P)(M521(P)M530(P)M540)/M2 >_line4
=>S _line4(P)F014/M2 >_line5
=>S _line2(P)J311/M0,M2 >_line6
=>S _line6(P)(M211(P)(M270 OR M273))/M2 >_line7
=>S _line7(P)J011/M2 >_line8
=>S _line2(P)J581/M0,M2 >_line9
=>S _line9(P)(M212(P)(M260 OR M262))/M2 >_line10
=>S _line2(P)(G100(P)M414(P)M531)/M0,M2 >_line11
=>S _line11(P)G010/M2 >_line12
=>S _line2(P)J331/M0,M2 >_line13
=>S M900/M0 >_line14
=>S (_line11(P)(_line9 OR _line13)) OR (_line3(P)(_line6 OR _line9)) >_line15
=>S _line14(P)_line15 >_line16
=>S M901/M2 >_line17
=>S (_line11(P)(_line9 OR _line13)) OR (_line4(P)(_line6 OR _line9)) >_line18
=>S _line17(P)_line18 >_line19
=>S _line16 OR _line19 >_line20
=>S M902/M2 >_line21
=>S (_line11(P)(_line7 OR _line10)) OR (_line4(P)(_line7 OR _line10)) >_line22
=>S _line21(P)_line22 >_line23
=>S _line20 OR _line23 OR (_line12(P)(_line8 OR _line10)) >_line24
=>S _line24 OR (_line5(P)(_line8 OR _line10)) >_line25
=>S _line25(NOTP)(H1 OR H2 OR H3 OR H4 OR H5 OR H6 OR H7 OR H8 OR H9 OR J1)/M2 >_line26
=>S _line26(NOTP)(J2 OR J4 OR J9 OR K0 OR M1)/M2 >_line27
*****
=>S (_line11(P)(_line9 OR _line13)) OR (_line3(P)(_line6 OR _line9)) >_line15
Simplified: (G100 (P) (J581 OR J331)) OR (F431 (P) (J311 OR J581)) =>_line15
*****

```

1. 化学構造式にコードを付与してみる
2. Chemical Fragmentation Codeを修正する
3. DWPIで構造検索 (Chemical Fragmentation Codes と DCR)
4. コード修正のまとめ
5. Multi Files, Multi Toolsを利用する検索参考例

Multi-Files, Multi-Toolsによる検索参考例

同じ化学構造を CAS registry とWPIで検索 (STN Expressを使う例)

本参考例の目的は
各ツールの多様性を検証するものではなく、Multi-File, Multi-Toolsを用
いる必要のある場合、如何に検索結果をまとめるかという参考例です。

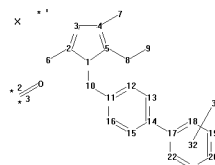
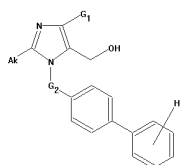
同じ化学構造を CAS registry とWPIで検索－1 (STN Expressを使う)

=> FILE REG

=>

Uploading C:\Presentn\BCE\trngxmS.str

- chain nodes :
- 6 7 8 9 10 23 25 26 31
- ring nodes :
- 1 2 3 4 5 11 12 13 14 15 16 17 18 19 20 21 22
- chain bonds :
- 1-10 2-6 4-7 5-8 8-9 10-11 14-17 25-26
- ring bonds :
- 1-2 1-5 2-3 3-4 4-5 11-12 11-16 12-13 13-14 14-15 15-16 17-18 17-22 18-19 19-20 20-21 21-22
- exact/norm bonds :
- 1-2 1-5 1-10 2-3 2-6 3-4 4-5 4-7 8-9 10-11 25-26
- exact bonds :
- 5-8 14-17
- normalized bonds :
- 11-12 11-16 12-13 13-14 14-15 15-16 17-18 17-22 18-19 19-20 20-21 21-22
- G1:CO2H[*1]
- G2:CH2.O.S[*2-*3]
- Match level :
- 1:Atom 2:Atom 3:Atom 4:Atom 5:Atom 6:CLASS 7:CLASS 8:CLASS 9:CLASS 10:CLASS 11:Atom 12:Atom 13:Atom 14:Atom 15:Atom 16:Atom 17:Atom 18:Atom 19:Atom 20:Atom 21:Atom 22:Atom 23:CLASS 25:CLASS 26:CLASS 31:CLASS 32:CLASS
- Generic attributes :
- 6:
- Saturation : Saturated
- Number of Carbon Atoms : less than 7
- 31:
- Type of Ring System : Monocyclic
- L1 STRUCTURE UPLOADED



Registryで構造検索

同じ化学構造を CAS registry とWPIで検索－2 (STN Expressを使う)

=> S L1

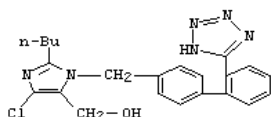
- 100.0% PROCESSED 35 ITERATIONS 2 ANSWERS
- SEARCH TIME: 00.00.01
- FULL FILE PROJECTIONS: ONLINE **COMPLETE**
- BATCH **COMPLETE**
- PROJECTED ITERATIONS: 346 TO 1054
- PROJECTED ANSWERS: 2 TO 124
- L2 2 SEA CSS SAM L1

サンプルサーチ

=> D L2 2

- L2 ANSWER 2 OF 2 REGISTRY COPYRIGHT 2002 ACS
- RN 148206-92-2 REGISTRY
- CN 1H-Imidazole-5-methanol, 2-butyl-4-chloro-1-[[2'-(1H-tetrazol-5-yl)][1,1'-biphenyl]-4-yl]methyl]-, compd. with pyridine (1:1) (9CI) (CA INDEX NAME)
- MF C22 H23 Cl N6 O . C5 H5 N
- SR CA
- LC STN Files: CA, CAPLUS
- CM 1
- CRN 114798-26-4
- CMF C22 H23 Cl N6 O

サンプルをチェック



同じ化学構造を CAS registry とWPIで検索－3 (STN Expressを使う)

=> S L1 Full

- 100.0% PROCESSED 671 ITERATIONS 22 ANSWERS
- SEARCH TIME: 00.00.03
- L3 22 SEA CSS FUL L1

フルサーチ

=> FILE CA; S L3

.....(省略)

- L4 1424 L3

=> S L4 AND P/DT

- 3649188 P/DT

- L5 149 L4 AND P/DT

=> FILE WPIX;

.....(省略)

=> TRANSFER L5 1-

- ENTER DISPLAY FIELDS (TI) OR ? : APPS
- SELECT IS APPROXIMATELY 87% COMPLETE
- L6 TRANSFER L5 1- APPS : 1529 TERMS
- SEARCH OF L6 IS APPROXIMATELY 66% COMPLETE
- L7 177 L6

結果をCAに

結果を特許に限定

DWPIに接続

CAの結果をDWPIにトランスファー

=> S (F521(P)G100(P)H401(P)H481(P)M240(P)M413(P)M532)/M0,M2,M3

- 1423 F521/M0
- 21177 F521/M2
-(省略)
- 61925 "L9"/M2
- 37525 "L9"/M3
- L54 733 L53(NOTP)("L4" OR "L5" OR "L6" OR "L7" OR "L8" OR "L9")/M2,M3

Chemical Frag Codesを実行

同じ化学構造を CAS registry とWPIで検索－４ (STN Expressを使う)

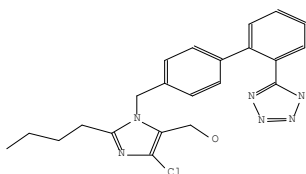
```

=> S L1
.....(省略)
• 100.0% PROCESSED 0 ITERATIONS 0 ANSWERS
• SEARCH TIME: 00.00.01
• FULL FILE PROJECTIONS: ONLINE **COMPLETE**
• BATCH **COMPLETE**
.....(省略)
• L55 0 SEA CSS SAM L1
=> S L1 Full
• ENTER TYPE OF SEARCH (SSS), CSS, FAMILY, OR EXACT:CSS
.....(省略)
• L56 3 SEA CSS FUL L1

=> D L56 IALLG 3

• L56 ANSWER 3 OF 3 WPIX COPYRIGHT 2002 DERWENT INFORMATION LTD
• ACCESSION NUMBER: DCR-99616
• DERWENT CHEM.RES.NO.: 147167-0-1-0
• PREF. CHEMICAL NAME: LOSARTAN POTASSIUM
• SYNONYM: COSARR; COZAAR; COZAAREX; DUP-753; HYZAAR; LOSARTAN;
LOSARTAN POTASSIUM; LOSARTAN-POTASSIUM; MK-954; POTASSIUM
LOSARTAN; POTASSIUM-LOSARTAN
.....(省略)
• CM 1
•
• K
•
• CM 2

```



THOMSON

69

Copyright 2007 Thomson Corporation

CA用の構造をDCR検索に

DCRサンプルサーチ及びフルサーチを実行

DCRの結果をチェック

同じ化学構造を CAS registry とWPIで検索－５ (STN Expressを使う)

```

=> S L56/DCR
L57 47 S L56/DCR
.....(省略)

=> S L54 OR L57
• L58 738 L54 OR L57

=> S L7 OR L58
• L59 843 L7 OR L58

=> S L58 NOT L7; S L58 AND L7; S L7 NOT L58
• L60 666 L58 NOT L7
• L61 72 L58 AND L7
• L62 105 L7 NOT L58

=> S L57 NOT L54; S L57 NOT L7
• L63 5 L57 NOT L54
• L64 7 L57 NOT L7

• => DUPLICATE IDENTIFY L59,L5
.....(省略)
• PROCESSING COMPLETED FOR L59
• PROCESSING COMPLETED FOR L5
• L65 887 DUPLICATE IDENTIFY L59 L5 (INCLUDES 68 SETS OF DUPLICATES)

```

DCRの結果をDWPIIにクロス

BCE codesの結果とDCR由来の結果をOR

CAからDWPIIにトランスファーした結果をBCE codesの結果、DCR由来の結果と合わせると843件得られる

DWPIIからユニークなヒット666件

両方ヒットしたのは72件

CAからユニークなヒット105件

BCE codesに対してDCRからユニークなヒットとして5件
CAに対してDCRからユニークなヒット7件

最後に得られた集合を重複除去 (Duplicate IDE) やファミリー (Fsort) をまとめて仕上げる

THOMSON

70

Copyright 2007 Thomson Corporation

同じ化学構造を CAS registry とWPIで検索－6

(STN Expressを使う)

```

=> FSORT L65
• SET SMARTSELECT ON
• SET COMMAND COMPLETED
• SET HIGHLIGHTING OFF
• SET COMMAND COMPLETED
• SEL L65 1- PN,APPS
• SELECT IS APPROXIMATELY 51% COMPLETE
• L66 SEL L65 1- PN APPS : 8369 TERMS
• 'L66' DELETED
• L66 887 FSO L65
• 117 Multi-record Families Answers 1-255
• Family 1 Answers 1-2
• Family 2 Answers 3-4
• .....(省略)
• Family 117 Answers 254-255
• 632 Individual Records Answers 256-887
• 0 Non-patent Records
• SET SMARTSELECT OFF
• SET COMMAND COMPLETED
• SET HIGHLIGHTING DEF
• SET COMMAND COMPLETED
  
```

まとめ:

STN Express Standard
Format で作図

STN Express WPI
Format で作図

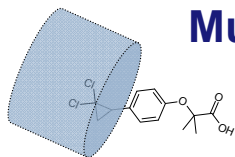


より包括的なサーチ
が必要とされる場:

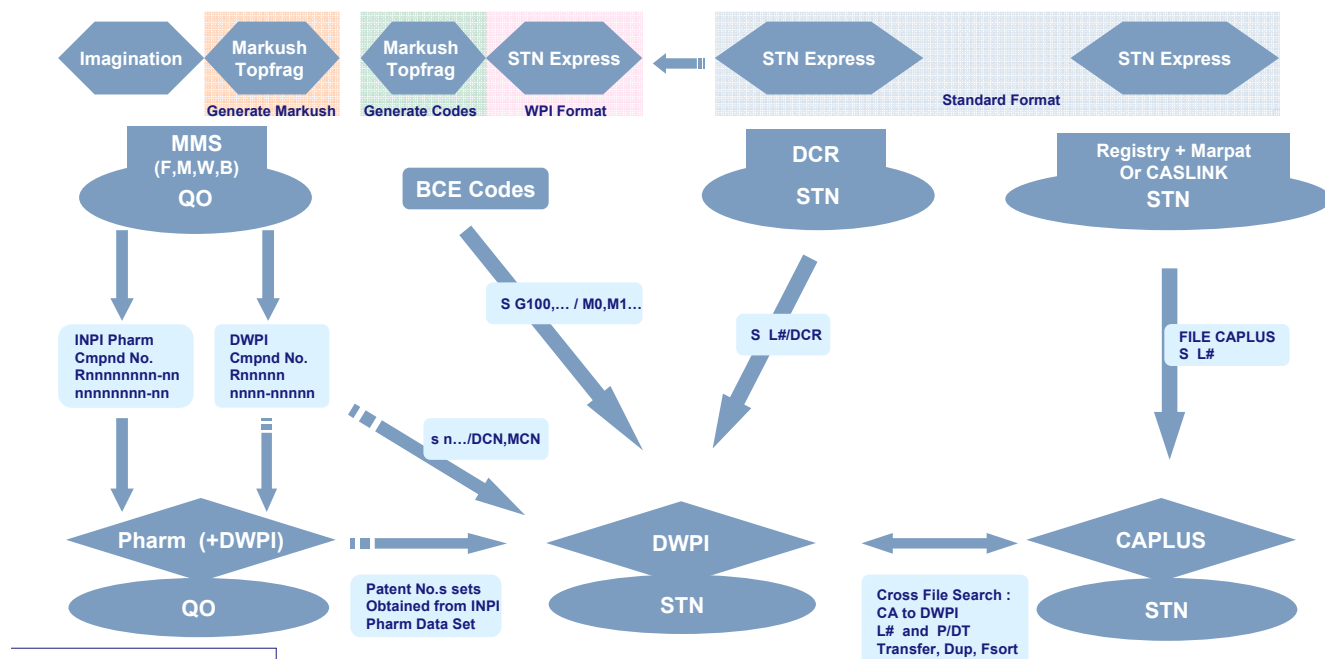
MARPAT やCASLINKを加える

MMS (Merged Markush Service)
の利用を加える

Chemical Frag (BCE) codesの
式に修正を加える

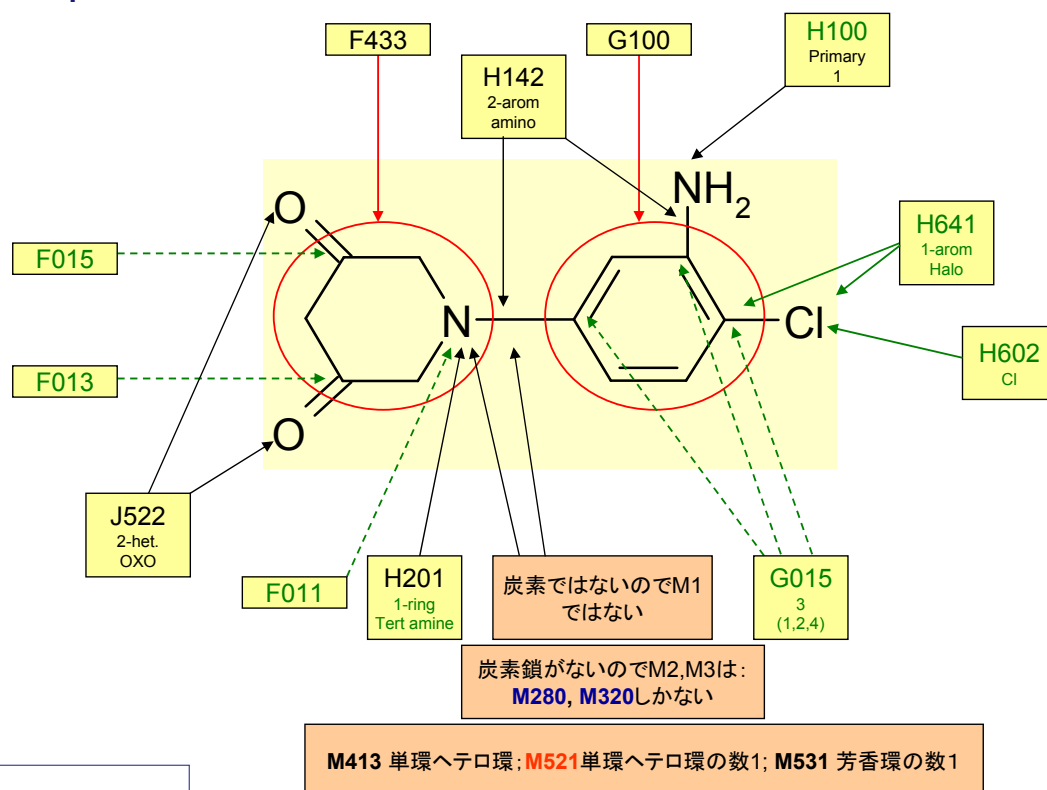


Multi-Files, Multi-Hostsにおける化学特許の検索 (STN を例として)



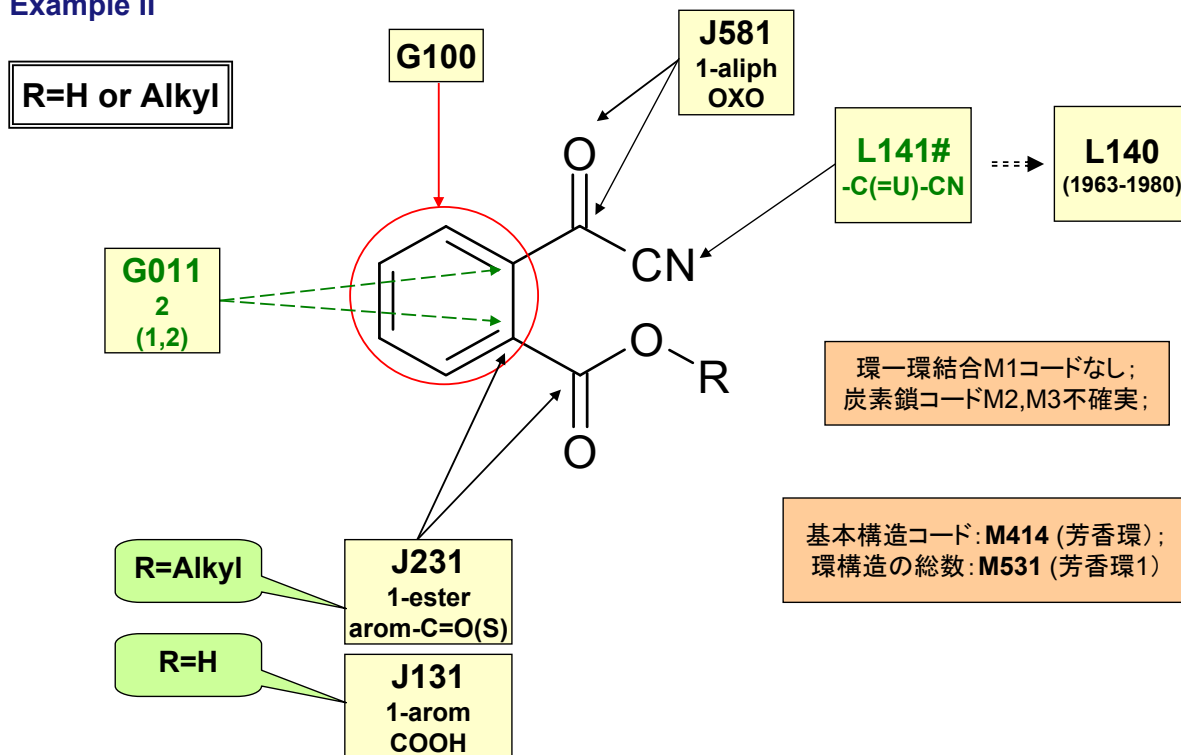
化学構造式にコードを付与してみる
---付与されるコードは---

Example I

Example I
検索式

- S M0,M2,M3=(F433(S)G100(S)H142(S)H201(S)H602(S)J522(S)M413(S)M531)
- S S1(S)M2,M3=M521
- S S2(S)M2,M3=(M280(S)M320)
- S S3(S)M2,M3=(F011(S)F013(S)F015(S)G015(S)H100(S)H641)
- S (S1(S)M0=M900)+(S2(S)M2,M3=M901)+(S3(S)M2,M3=M902)+S4
- S S5(NOT S)M2,M3=(H3+H4+H5+H7+H8+H9+J0+J1+J2+J3+J4+J6+J9+K0+M1)

Example II



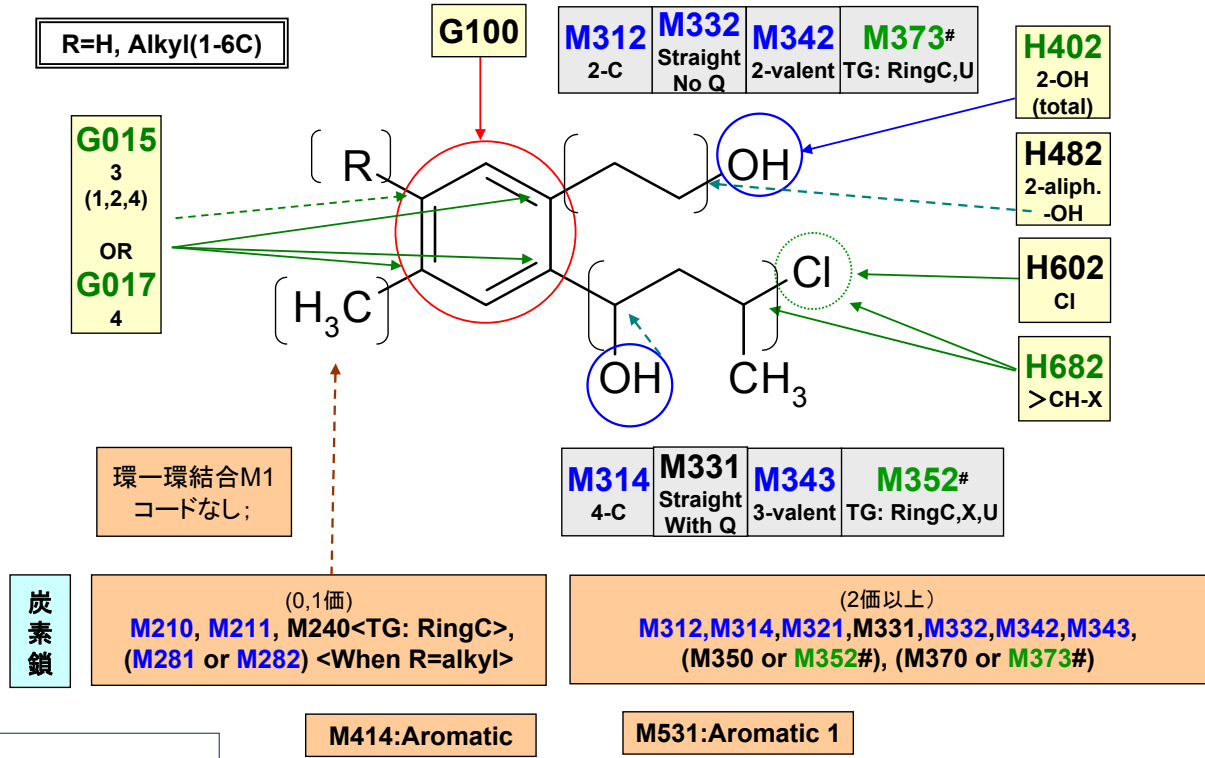
検索式 II

- S M0,M2,M3=(**G100**(S)(**J131** or **J231**)(S) **J581**(S)**M414**(S)**M531**(S) ("L140" or "L141"))
- S S1
- S S2
- S S3(S)M2,M3=(**G011**(S)**J011**)
- S (S1(S)M0=**M900**)+(S2(S)M2,M3=**M901**)+(S3(S)M2,M3=**M902**)+S4
- S S5(NOT S)M2,M3=(H3+H4+H5+H7+H8+H9+J0+J1+J2+J3+J4+J6+J9+K0+M1)

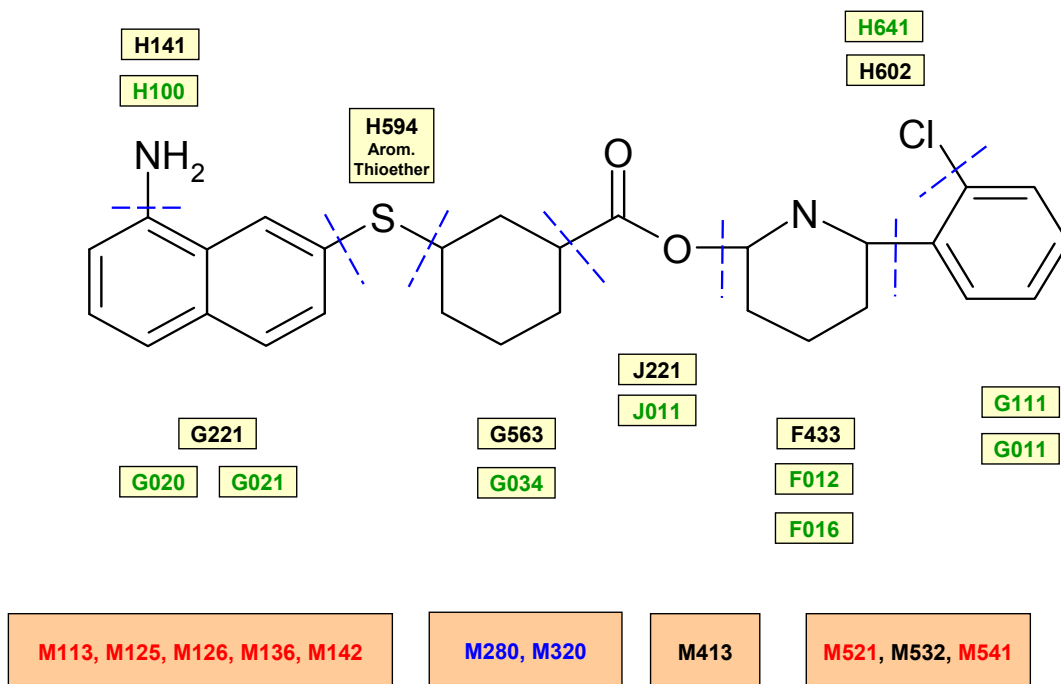
R=CH3 の場合

- S M0,M2,M3=(**G100**(S)**J231**(S) **J581**(S)**M414**(S)**M531**(S) ("L140" or "L141"))
- S S1
- S S2(S)M2,M3=(**M211**(S)**M281**(S)**M320**(S)(**M270** or **M272**))
- S S3(S)M2,M3=(**G011**(S)**J011**)
- S (S1(S)M0=**M900**)+(S2(S)M2,M3=**M901**)+(S3(S)M2,M3=**M902**)+S4
- S S5(NOT S)M2,M3=(H1+H2+H3+H4+H5+H6+H7+H8+H9+J1+J3+J4+J6+J9+K1+K2)
- S S6(NOT S)M2,M3=(K3+K4+K5+K6+K7+K8+K9+L2+L3+L4+L5+L6+L7+L8+L9+M1)

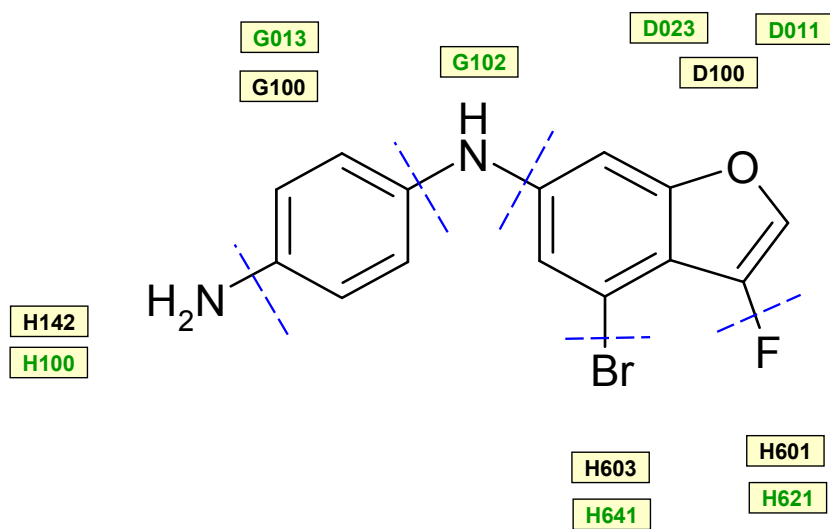
Example III



Example IV



Example V



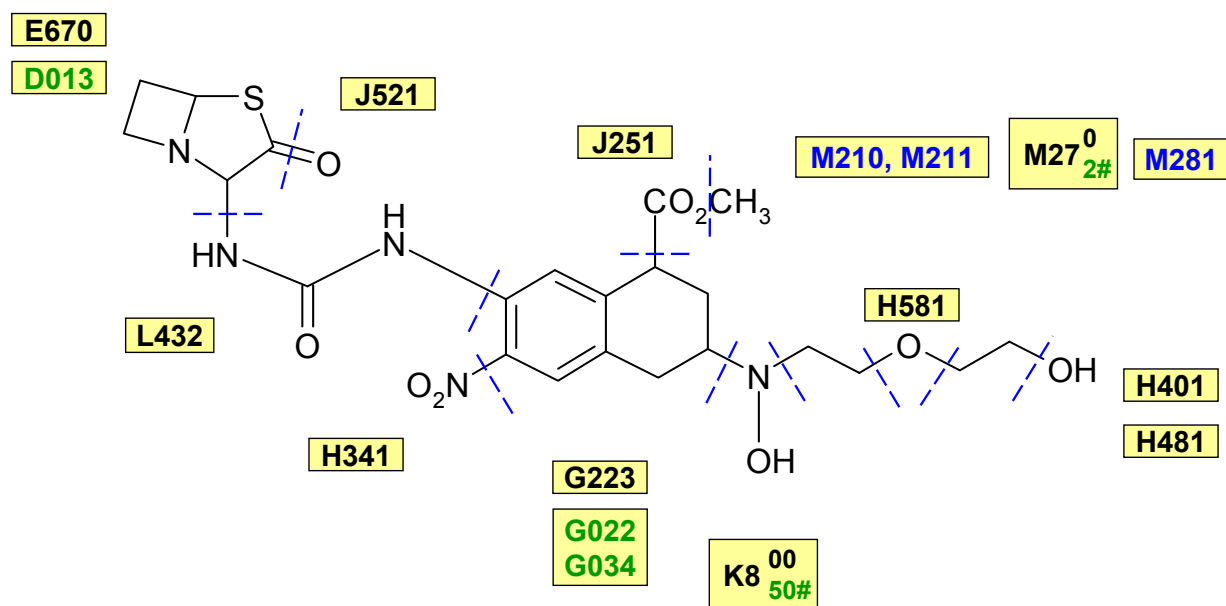
M122, M143

M280, M320

M412

M511, M520, M531

Example VI



M412

M511, M520, M531, M540

Example VI

Use STN Express

```

stnedit
File Edit Utilities Search Preferences Window Help
Check Command File
Next Error
C:\STNEXP\Uscripts\fragcode_ad_070220.sc
=>S ("E670"(P)G223(P)H341(P)H401(P)H481(P)H581(P)J521(P)J581(P)"L432")/M0,M2,M3,M4 ¥>_line1
=>S _line1(P)(M412(P)M531(P)(K800 OR K850))/M0,M2,M3,M4 ¥>_line2
=>S _line2(P)(M125(P)M137(P)M511(P)M520(P)M540)/M2,M3,M4 ¥>_line3
=>S _line3(P)(M210(P)M281(P)M312(P)M322(P)M332(P)M342(P)M392)/M2,M3,M4 ¥>_line4
=>S _line4(P)((M260 OR M262)(P)(M380 OR M383))/M2,M3,M4 ¥>_line5
=>S _line5(P)(D013(P)G022(P)G034(P)"L943"(P)M211)/M2,M3,M4 ¥>_line6
=>S (_line2(P)M900/M0) OR (_line3(P)M901/M2,M3,M4) OR (_line5(P)M902/M2,M3,M4) ¥>_line7
=>S _line7 OR _line6 ¥>_line8
=>S _line8(NOTP)(H1 OR H2 OR H6 OR H7 OR H8 OR J0 OR J1 OR J2 OR J3 OR J4)/M2,M3,M4 ¥>_line9
=>S _line9(NOTP)(J6 OR J9 OR K1 OR K2 OR K6 OR K4 OR K5 OR K6 OR K7 OR K8)/M2,M3,M4 ¥>_line10
=>S _line10(NOTP)("L" OR "L2" OR "L3" OR "L5" OR "L6" OR "L7" OR "L8")/M2,M3,M4 ¥>_line11
¥* M322 - ONE OF CODES M311 TO M316 USED TWICE ONLY (multiplier for

```

どうもありがとうございました

Thomson Scientific / テクニカル・サポート

Phone: 0800-888-8855

Fax: 03-5218-6536

ts.support.jp@thomson.com

Thomson Scientificの日本語ホームページをご覧ください。

<http://www.thomsonscientific.jp>

各種製品情報、関連ニュース、テキストのダウンロードなどができます。

<http://www.thomsonscientific.jp/support/>