

**STN INTERNATIONAL**

リフレッシュセミナー

# MARPAT ファイル — 検索テクニック

---

# 目次

## A 概要

CAS FILES における特許中の化学物質索引 .....	1
MARPAT ファイルとは .....	3

## B MARPAT ファイルの検索の仕組み

検索結果のコントロール .....	13
マッチレベル .....	14
参考: 一般式テキスト・ショートカット .....	17
MARPAT ファイルの検索の仕組み .....	18
構造作図ポイント 1 .....	22
構造作図ポイント 2 .....	24
一般式属性 .....	26
構造作図ポイント 3 .....	30
検索できないデータ .....	32
構造作図ポイント 4 .....	33
構造作図ポイント 5 .....	34
構造質問式改善の指針 .....	36

## C 検索テクニック

CASLINK の応用 .....	39
参考: 重複文献除去 .....	50
多成分物質 .....	56
CSS 検索 .....	60
SYSTEM LIMIT を超えた場合 .....	62
STRUCTURE TOO LARGE .....	64
参考: アップロードした構造質問式中のマッチレベルの確認 .....	66

## A 概要

MARPAT ファイルの収録内容と, CAplus/CA ファイルとの関係を説明します



# A 概要

## CAS FILES における特許中の化学物質索引

### ■ CAS FILES における特許中の化学物質索引

MARPAT ファイルで得られる特許は、すべて CAplus/CA ファイルに収録されている。  
 しかし、化学物質の索引方針および検索機能の違いによって、それぞれのファイルで得られる特許に違いが生じる。

- 書誌情報、抄録、索引が作成され、CAplus/CA ファイルに収録される。
- 実施例・特許請求範囲中の重要な特定の化学物質は、CAS 登録番号で収録される。

(18) 日本特許庁 (J P)	(12) 公開特許公報 (A)	(11) 特許出願公開番号 特開2002-20372 (1992-20372A)
(21) 出願番号 C07D 233/07 C07B 41/06	(22) 優先権 3 0 0	(71) 出願人 P I C O T D 233/07 C O T B 41/06 3 0 0 4 C 0 0 8 4 M 0 3 8
(23) 発明の名称 1 0 0 2 8 1 4 1 . 0	(24) 発明日 平成 12 年 6 月 8 日 (2000.6.8)	(72) 発明者 バヤエル・アクチエンゲゼルシャフト BAYER AKTIENGESELLSCHAFT CHIAI 博士 フイロン製薬有限会社 藤原 一孝 氏 ターゼン (徳島県)
(25) 公開日 平成 12 年 6 月 5 日 (2000.6.5)	(26) 特許請求の範囲の総数 全 5 項	

- MARPAT ファイルには、特許請求範囲中のマルクーシュ構造が収録される
- MARPAT ファイルには CA ファイルと同じ特許の情報が含まれている

実施例

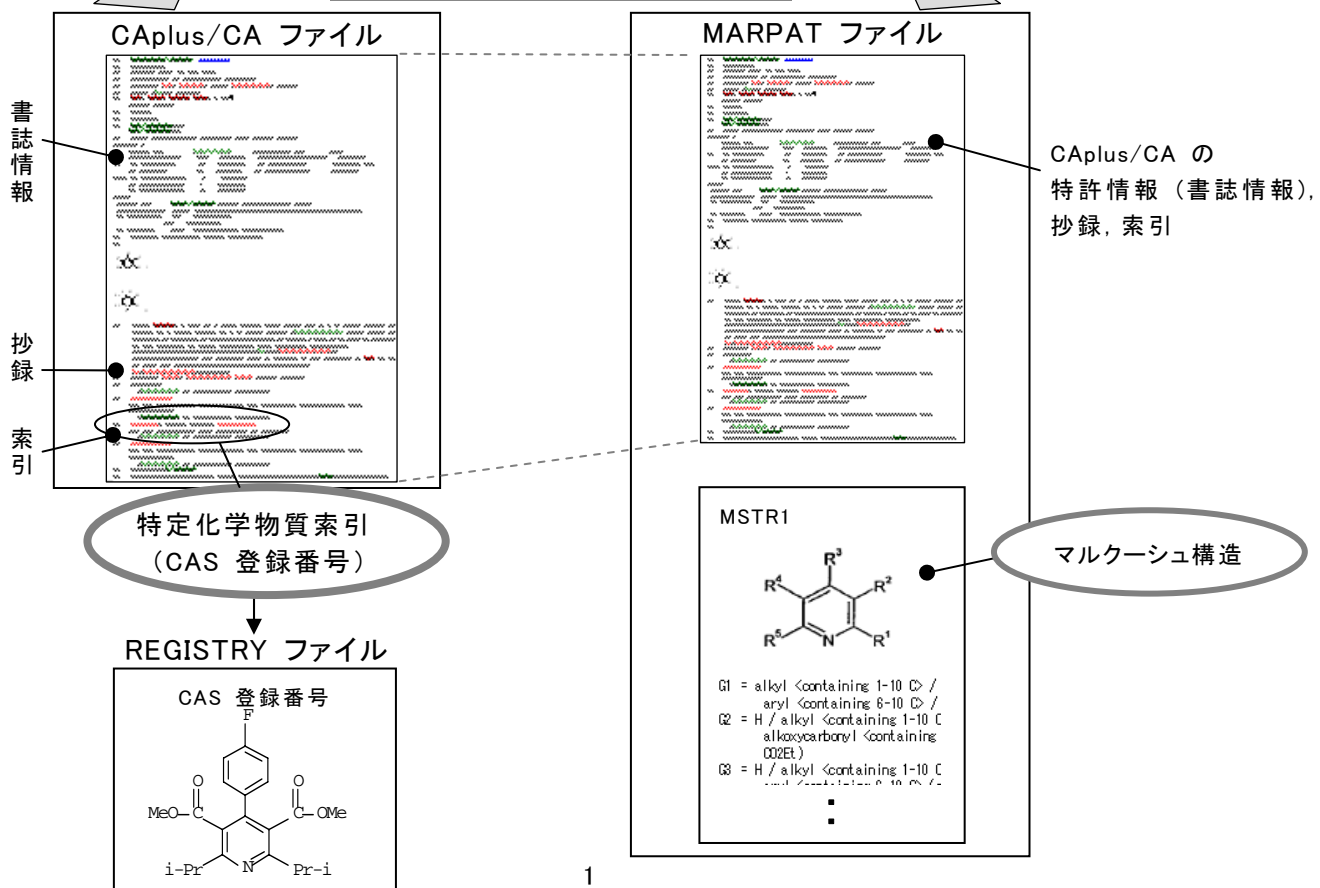
MeO-C(=O)-C<sub>5</sub>H<sub>3</sub>(F)-C(=O)-OMe  
i-Pr

特許請求範囲\*

R<sup>1</sup>, R<sup>5</sup> = C1-10 alkyl, C6-10 aryl; R<sup>2</sup>, R<sup>4</sup> = H, C1-10 alkyl, CO<sub>2</sub>R<sup>6</sup> ...

\* 発明の詳細な説明から索引する場合もある

REGISTRY/CAplus の検索に加え MARPAT を検索することによって化学物質関連特許の網羅的な調査が可能に



## A 概要

### CAS FILES における特許中の化学物質索引

#### ■ MARPAT ファイルと, REGISTRY ファイル, CAplus ファイルの比較

2009 年 8 月現在

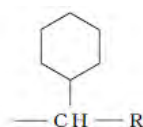
	MARPAT	REGISTRY	CAplus/CA
レコード構成	文献単位 (特許ファミリー単位)	化学物質単位	文献単位 (特許はファミリー単位)
概要	CA ファイルに収録した 特許クレーム中の マルクーシュ構造*を収録	化学構造, 配列, 名称, 分子式, 物性値, 参照文献情報等を収録	特許 (59 特許発行機関) 雑誌論文 学会会議録等を収録
特徴	マルクーシュ構造*が 検索可能	・ 収録物質数が多い ・ 構造検索が可能	・ 速報性 ・ 収録文献数が多い ・ 抄録および物質索引が 充実している
収録件数 (収録年)	約 32 万レコード (すべて特許レコード)  (CA 由来 1988 年～) (INPI 由来 1961 年～)	約 4,786 万レコード (配列以外)(1907 年～)  約 6,244 万レコード (配列)(1957 年～)	CAplus ファイル : 約 3,123 万レコード (特許は約 678 万件) (1808 年～)  CA ファイル : 約 2,825 万レコード (特許は約 542 万件) (1808 年～)
収録物質	有機化合物	○ (構造)	○ (CAS 登録番号)
	ポリマー	×	○ (CAS 登録番号)
	無機化合物	×	○ (CAS 登録番号)
	タンパク質・ 核酸配列	×	○ (CAS 登録番号)
更新頻度	毎日	毎日	CAplus ファイル : 毎日 CA ファイル : 毎週
タイムラグ	特許発行後約 2ヶ月以内 (CA ファイルに収録された 後に MARPAT ファイルに 収録)	/	CAplus ファイル : 主要国特許は発行後 1~2 日以内に収録 CA ファイル: 3~6 週間



#### マルクーシュ構造とは

化学の分野の特許中に見られる複雑な構造式。  
複数の選択肢がある置換基について, 選択肢を可変構造式で示したもの。  
マーカッシュ構造, マーカッシュ形式とも呼ばれる。

(マルクーシュ構造の例)



(Rは, C1~4のアルキル基, アルケニル基, アリール基を示す。)

## A 概要

### MARPAT ファイルとは

■ MARPAT ファイルは, Chemical Abstracts Service (CAS) が作成するマルクーシュ (Markush) 構造を含む特許情報のファイルである (マルクーシュ構造とは化学構造の一般式のこと). クレーム中のマルクーシュ構造から特許調査ができるファイルである.

- レコード単位

- 文献 (特許) 単位 . レコード番号 (AN) は CA ファイルと同じ (CA 抄録番号).

- 収録対象特許

- CAplus/CA ファイル収録対象\* (59 特許発行機関) と同じ

\* <http://www.cas.org/expertise/cascontent/caplus/patcoverage/> 参照

- 収録対象化合物

- 有機化合物, 有機金属化合物, 低重合の物質 (重合度が 10 までの物質)

- 収録対象外の化合物

- 合金, 金属酸化物, 無機塩, 金属間化合物, ポリマー

- 検索できない置換基で表現された物質 (例: "Ph-R" R is any substituent, "halogenated alkenediols")

- 特定の化学物質 (REGISTRY ファイルに収録されるような構造定義が明確な物質)

- マルクーシュ構造の収録源と収録年

- 1988 年 (特許発行年) 以降 - CAplus/CA ファイルのベーシック特許から収録 (CAS 作成)

マルクーシュ構造の存在位置		MARPAT ファイル中の記載
特許請求範囲	マルクーシュ構造が収録される	Patent location: claim 1
発明の詳細な説明	特許請求範囲中にマルクーシュ構造がない場合に収録される	Patent location: disclosure
	発明の詳細な説明のマルクーシュ構造が, 特許請求範囲のマルクーシュ構造を包括する場合に収録する	Patent location: claim 10 Note: also incorporates broader disclosure

- 1961~1987 年 (特許発行年) - INPI より提供されたデータ (ベーシック特許とは限らない)

マルクーシュ構造の存在位置		MARPAT ファイル中の記載
特許請求範囲または発明の詳細な説明	INPI が所有するデータの中で, CA 収録特許であるもの	Patent location: claims Note: record may include structures from disclosure

A 概要

MARPAT ファイルとは

■ MARPAT ファイルレコード例 : ALL 表示形式 (日本の公開番号 2002 年 20372)

**BIB**

CAplus/CA  
ファイルの  
書誌情報  
(IC, CC, CLASS  
は索引情報)

**ABS**

CAplus/CA  
ファイルの  
抄録

**IND**

CAplus/CA  
ファイルの  
索引情報

レコード番号 (AN) は CA ファイルと同じ (CA 抄録番号)

ベーシック特許から MSTR 作成  
(1988 年以降)

注:  
MARPAT ファイルでは CAplus/CA ファイル  
の情報を表示することができるが、検索する  
ことはできない

AN 136:102298 ●MARPAT TI Preparation of substituted IN Norbert, Lui; Panskus, Hans; Schnatterer, Albert PA Bayer A.-G., Germany SO Jpn. Kokai Tokkyo Koho, 5 pp. CODEN: JKXXAF DT Patent LA Japanese IC ICM C07D213-807 ICS C07B061-00 CC 27-16 (Heterocyclic Compounds (One Hetero Atom)) FAN. CNT 2	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">PATENT NO.</th> <th style="text-align: left;">KIND</th> <th style="text-align: left;">DATE</th> <th style="text-align: left;">APPLICATION NO.</th> <th style="text-align: left;">DATE</th> </tr> </thead> <tbody> <tr> <td>PI JP 2002020372 ●</td> <td>A2</td> <td>20020123</td> <td>JP 2001-169465</td> <td>20010605</td> </tr> <tr> <td>DE 10111874</td> <td>A1</td> <td>20011213</td> <td>DE 2001-10111874</td> <td>20010313</td> </tr> <tr> <td>PRAI DE 2000-10028141</td> <td>20000608</td> <td></td> <td></td> <td></td> </tr> <tr> <td>DE 2001-10111874</td> <td>20010313</td> <td></td> <td></td> <td></td> </tr> <tr> <td>DE 2000-10028414</td> <td>20000608</td> <td></td> <td></td> <td></td> </tr> </tbody> </table>	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE	PI JP 2002020372 ●	A2	20020123	JP 2001-169465	20010605	DE 10111874	A1	20011213	DE 2001-10111874	20010313	PRAI DE 2000-10028141	20000608				DE 2001-10111874	20010313				DE 2000-10028414	20000608			
PATENT NO.	KIND	DATE	APPLICATION NO.	DATE																											
PI JP 2002020372 ●	A2	20020123	JP 2001-169465	20010605																											
DE 10111874	A1	20011213	DE 2001-10111874	20010313																											
PRAI DE 2000-10028141	20000608																														
DE 2001-10111874	20010313																														
DE 2000-10028414	20000608																														

CLASS PATENT NO. CLASS PATENT FAMILY CLASSIFI	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">PATENT NO.</th> <th style="text-align: left;">CLASS</th> <th style="text-align: left;">PATENT FAMILY CLASSIFI</th> </tr> </thead> <tbody> <tr> <td>JP 2002020372</td> <td>ICM ICS IPC1</td> <td>C07D0213-807 C07B0061-00 C07D0213-807 [ICM, 7]; C07D0213-00 [ICM, 7, C*]; C07B0061-00 [ICS, 7]</td> </tr> <tr> <td></td> <td>IPCR</td> <td>C07D0213-807 [I, A]; C07B0061-00 [I, C*]; C07B0061-00 [I, A]; C07D0213-00 [I, C*]; C07D0213-80 [I, A]; C07D0213-803 [I, A]</td> </tr> <tr> <td>DE 10111874</td> <td>IPC1</td> <td>C07D0213-133 [ICM, 7]; C07D0213-79 [ICS, 7]; C07D0213-00 [ICS, 7, C*]</td> </tr> <tr> <td></td> <td>IPCR</td> <td>C07D0213-00 [I, C*]; C07D0213-80 [I, A]; C07D0213-803 [I, A]</td> </tr> <tr> <td></td> <td>ECLA</td> <td>C07D213/79B4; C07D213/80C9</td> </tr> </tbody> </table>	PATENT NO.	CLASS	PATENT FAMILY CLASSIFI	JP 2002020372	ICM ICS IPC1	C07D0213-807 C07B0061-00 C07D0213-807 [ICM, 7]; C07D0213-00 [ICM, 7, C*]; C07B0061-00 [ICS, 7]		IPCR	C07D0213-807 [I, A]; C07B0061-00 [I, C*]; C07B0061-00 [I, A]; C07D0213-00 [I, C*]; C07D0213-80 [I, A]; C07D0213-803 [I, A]	DE 10111874	IPC1	C07D0213-133 [ICM, 7]; C07D0213-79 [ICS, 7]; C07D0213-00 [ICS, 7, C*]		IPCR	C07D0213-00 [I, C*]; C07D0213-80 [I, A]; C07D0213-803 [I, A]		ECLA	C07D213/79B4; C07D213/80C9
PATENT NO.	CLASS	PATENT FAMILY CLASSIFI																	
JP 2002020372	ICM ICS IPC1	C07D0213-807 C07B0061-00 C07D0213-807 [ICM, 7]; C07D0213-00 [ICM, 7, C*]; C07B0061-00 [ICS, 7]																	
	IPCR	C07D0213-807 [I, A]; C07B0061-00 [I, C*]; C07B0061-00 [I, A]; C07D0213-00 [I, C*]; C07D0213-80 [I, A]; C07D0213-803 [I, A]																	
DE 10111874	IPC1	C07D0213-133 [ICM, 7]; C07D0213-79 [ICS, 7]; C07D0213-00 [ICS, 7, C*]																	
	IPCR	C07D0213-00 [I, C*]; C07D0213-80 [I, A]; C07D0213-803 [I, A]																	
	ECLA	C07D213/79B4; C07D213/80C9																	

OS CASREACT 136:102298; MARPAT 136:102298 GI	<div style="display: flex; align-items: center;"> <div style="text-align: center; margin-right: 20px;"> <p>I</p> </div> <div style="text-align: center; margin-right: 20px;"> <p>II</p> </div> </div>
---	---

AB Title compds. I (R1, R5 = C1-10 alkyl, C6-10 aryl; R2, R4 = H, C1-10 alkyl, CN, CO2R6; R6 = C1-10 alkyl; R3 = H, C1-10 alkyl, (un)substituted C6-10 aryl) are prepared by reaction of 1,4-dihydropyridine II (R1-R5 = same as I) with Me nitrite in the presence of acids containing <20% oxidizing components. 4-(4-Fluorophenyl)-2,6-diisopropyl-3,5-di(methoxycarbonyl)-1,4-dihydropyridine was oxidized with Me nitrite in 1,4-dihydropyridine was oxidized with Me nitrite in the presence of HCl at 60° to give 98% 4-(4-fluorophenyl)-4-(4-fluorophenyl)-4-(4-fluorophenyl)-2,6-diisopropyl-3,5-di(methoxycarbonyl)pyridine.	pyridine prepn; hydroxyridine oxidn methyl nitrite (preparation of substituted pyridines) 122549-42-2P RL: IMF (Industrial manufacture); SPN (Synthetic preparation); PREP (Preparation) (preparation of substituted pyridines) 624-91-9, Methyl nitrite 132008-67-4 RL: RCT (Reactant); RACT (Reactant or reagent) (preparation of substituted pyridines)
--	---

ST IT Oxidation	(preparation of substituted pyridines) 122549-42-2P RL: IMF (Industrial manufacture); SPN (Synthetic preparation); PREP (Preparation) (preparation of substituted pyridines) 624-91-9, Methyl nitrite 132008-67-4 RL: RCT (Reactant); RACT (Reactant or reagent) (preparation of substituted pyridines)
--------------------	---

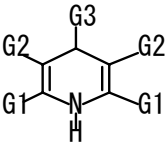
4

## A 概要

### MARPAT ファイルとは

**MSTR**  
 クレーム中の  
 マルクーシュ  
 構造  
 (MARPAT  
 特有の情報)

**MSTR 1**

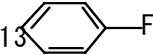


基本骨格 1  
(G グループを含めた構造)

G1 = alkyl <containing 1-10 C> / aryl <containing 6-10 C> / (Example: Pr-i)

G2 = H / alkyl <containing 1-10 C> / CN / alkoxy carbonyl <containing 1-10 C> / (Examples: CO2Me / CO2Et)

G3 = H / alkyl <containing 1-10 C> / aryl <containing 6-10 C> (opt. substd. by 1 or more G4) / (Example: 13)



G4 = halo / NO2 / alkoxy carbonyl <containing 1-10 C> / CN / alkyl <containing 1-10 C>

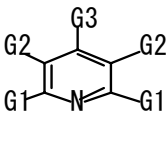
Patent location: claim 1

マルクーシュ構造は基本骨格と置換基で索引される

↓

このマルクーシュ構造を構造検索することで化合物特許を調査するのが MARPAT ファイル

**MSTR 2**

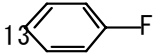


基本骨格 2

G1 = alkyl <containing 1-10 C> / aryl <containing 6-10 C> / (Example: Pr-i)

G2 = H / alkyl <containing 1-10 C> / CN / alkoxy carbonyl <containing 1-10 C> / (Examples: CO2Me / CO2Et)

G3 = H / alkyl <containing 1-10 C> / aryl <containing 6-10 C> (opt. substd. by 1 or more G4) / (Example: 13)



G4 = halo / NO2 / alkoxy carbonyl <containing 1-10 C> / CN / alkyl <containing 1-10 C>

Patent location: claim 1

基本骨格 1 中の G グループ (置換基) の定義

\* 数字は構造フラグメント中の結合点

基本骨格 2 中の G グループ (置換基) の定義

マルクーシュ構造記載位置

マルクーシュ構造記載位置

#### ・ MSTR 中の記載

MSTR 中の記載	内容	例
(Specifically claimed)	別のクレーム中 (従属項など) で、権利請求されている置換基を収録	(Specifically claimed: Me) メチル基が権利請求されている
(Example)	実施例中に記載されている具体的な例示物質を選択的に収録	(Example: Pr-i) イソプロピル基が例示として特許中に記載されている
(subst. by) substituted by	<u>必ず</u> 置換する	(subst. by 1 or more aryl) 1 以上の aryl 基が置換する
(opt. subst. by.) optionally substituted by	<u>任意</u> の置換基の存在を表す	(opt. subst. by 1 or more G4) 1 以上の G4 が置換していてもよい

## A 概要

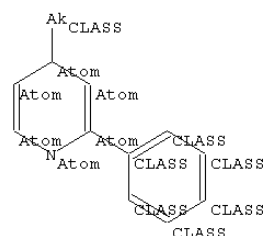
### MARPAT ファイルとは

#### ■ MARPAT ファイルの検索.

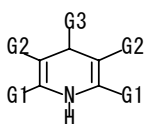
- MARPAT ファイルは、クレーム中のマルクーシュ構造から特許調査をするためのファイルなので、原則として構造検索を利用する。

構造質問式

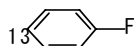
MARPAT ファイルの回答



MSTR 1



- G1 = alkyl <containing 1-10 C> /  
**aryl <containing 6-10 C>** / (Example: Pr-i)
- G2 = H / alkyl <containing 1-10 C> / CN /  
 alkoxycarbonyl <containing 1-10 C> /  
 (Examples: CO2Me / CO2Et)
- G3 = H / **alkyl <containing 1-10 C>** /  
 aryl <containing 6-10 C> (opt. substd. by  
 1 or more G4) / (Example: 13)



- G4 = halo / NO2 / alkoxycarbonyl <containing 1-10 C> /  
 CN / alkyl <containing 1-10 C>  
 Patent location: claim 1

- MARPAT ファイルで実行可能な辞書検索フィールド

検索項目	コード	説明	入力例
レコード番号	/AN	CA 抄録番号	S 109:73345/AN
基本索引	/BI	NTE, DER, MPL	S SALT#
	なし	STE 中の単語	
入力日	/ED	入力日	S 20070321/ED
更新日	/UP	更新日	S 20070321/UP

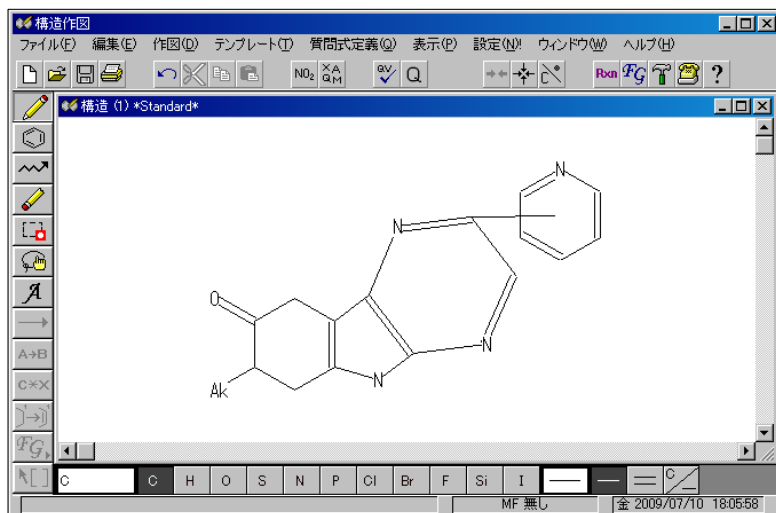
- MARPAT ファイルでは、CAplus/CA ファイルと同じ文献情報を表示することができるが上表以外のフィールドを検索することはできない。
- 書誌情報、抄録、索引情報を検索するときは、CAplus/CA ファイルにクロスオーバーし、検索する。

## A 概要

### MARPAT ファイルとは

- REGISTRY ファイルの検索で回答が得られなかった場合でも、MARPAT ファイルでは回答が得られる場合がある。

- 検索例:



=> FILE REGISTRY

=>

Uploading C:\My\_Document\STN Express 8.4\Queries\sample.str

L1 STRUCTURE UPLOADED

=> S L1

L2 0 SEA SSS SAM L1

=> S L1 FUL

L3 0 SEA SSS FUL L1

REGISTRY ファイルで部分構造検索のフルファイル検索を実行したが回答は得られなかった

=> FILE MARPAT

:

=> S L3

L4 0 SEA SSS SAM L1

=> S L3 FUL

:

L5 3 SEA SSS FUL L1

MARPAT ファイルでは 3 件ヒットした

A 概要

MARPAT ファイルとは

=> D BIB FQHIT 1-3

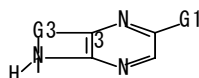
L5 ANSWER 1 OF 3 MARPAT COPYRIGHT 2009 ACS on STN  
 AN 127:293247 MARPAT [Full-text](#)  
 TI Preparation of pyrrolopyrazines as GABA<sub>A</sub> receptor ligands  
 IN Blum, Charles; Hutchison, Alan  
 PA Neurogen Corp., USA  
 SO U.S., 16 pp., Cont.-in-part of U.S. 5,606,059.  
 CODEN: USXXAM

DT Patent  
 LA English

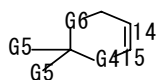
FAN.CNT 3

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 5668283	A	19970916	US 1995-486595	19950607
	US 5286860	A	19940215	US 1992-975409	19921112
	US 5606059	A	19970225	US 1995-436252	19950512
	US 5910590	A	19990608	US 1997-931648	19970916
PRAI	US 1992-975409		19921112		
	US 1995-436252		19950512		
	WO 1993-US10870		19931110		
	US 1995-486595		19950607		

**MSTR 2A**



G1 = pyridyl (opt. substd. by (1-2) G2)  
 G3 = 14-3 15-1



G4 = (0-2) CH<sub>2</sub>  
 G5 = alkyl <containing 1-6 C>  
 G6 = C(0)

Derivative: and pharmaceutically acceptable non-toxic salts  
 Patent location: disclosure

RE.CNT 2 THERE ARE 2 CITED REFERENCES AVAILABLE FOR THIS RECORD  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

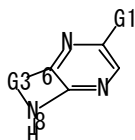
A 概要

MARPAT ファイルとは

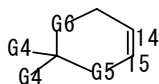
L5 ANSWER 2 OF 3 MARPAT COPYRIGHT 2009 ACS on STN  
 AN 126:199581 MARPAT [Full-text](#)  
 TI aryl substituted pyrrolopyrazines as a new class of GABA brain receptor ligands  
 IN Blum, Charles; Hutchison, Alan  
 PA Neurogen Corporation, USA  
 SO U.S., 16 pp., Cont.-in-part of U.S. 5,286,860.  
 CODEN: USXXAM  
 DT Patent  
 LA English  
 FAN. CNT 3

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 5606059	A	19970225	US 1995-436252	19950512
	US 5286860	A	19940215	US 1992-975409	19921112
	WO 9411374	A1	19940526	WO 1993-US10870	19931110
	W: AT, AU, BB, BG, BR, BY, CA, CH, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, KZ, LK, LU, MG, MN, MW, NL, NO, NZ, PL, RO, RU, SD, SE, SK, UA, US, VN				
	RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG				
	US 5668283	A	19970916	US 1995-486595	19950607
PRAI	US 1992-975409		19921112		
	WO 1993-US10870		19931110		
	US 1995-436252		19950512		

**MSTR 2A**



G1 = pyridyl (opt. substd. by (1-2) G2)  
 G3 = 14-6 15-8



G4 = alkyl <containing 1-6 C>  
 G5 = (0-2) CH2  
 G6 = C(O)

Derivative: and pharmaceutically acceptable non-toxic salts  
 Patent location: disclosure

RE. CNT 8 THERE ARE 8 CITED REFERENCES AVAILABLE FOR THIS RECORD  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

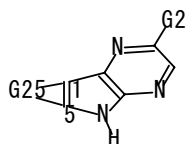
A 概要

MARPAT ファイルとは

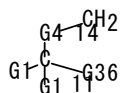
L5 ANSWER 3 OF 3 MARPAT COPYRIGHT 2009 ACS on STN  
 AN 120:245160 MARPAT [Full-text](#)  
 TI Preparation of indolopyrazines and related compounds as brain GABAa agonists, antagonists, or inverse agonists  
 IN Blum, Charles; Hutchison, Alan  
 PA Neurogen Corp., USA  
 SO U.S., 17 pp.  
 CODEN: USXXAM  
 DT Patent  
 LA English  
 FAN.CNT 3

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 5286860	A	19940215	US 1992-975409	19921112
	WO 9411374	A1	19940526	WO 1993-US10870	19931110
	W: AT, AU, BB, BG, BR, BY, CA, CH, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, KZ, LK, LU, MG, MN, MW, NL, NO, NZ, PL, RO, RU, SD, SE, SK, UA, US, VN				
	RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG				
	AU 9455526	A	19940608	AU 1994-55526	19931110
	US 5606059	A	19970225	US 1995-436252	19950512
	US 5668283	A	19970916	US 1995-486595	19950607
PRAI	US 1992-975409		19921112		
	WO 1993-US10870		19931110		
	US 1995-436252		19950512		

**MSTR 1**



G1 = alkyl <containing 1-6 C>  
 G2 = pyridyl (opt. substd. by (1-2) G3)  
 G4 = C(0)  
 G25 = 14-1 11-5



G36 = (1-2) CH2  
 Derivative: or pharmaceutically acceptable non-toxic salts  
 Patent location: claim 1  
 Note: also incorporates claim 3

RE.CNT 6 THERE ARE 6 CITED REFERENCES AVAILABLE FOR THIS RECORD  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

## *B MARPAT* ファイルの検索の仕組み

MARPAT ファイルの検索の仕組みについて紹介します



## B MARPAT ファイルの検索の仕組み

### 検索結果のコントロール

#### ■ 検索結果のコントロール

##### ・ 検索の種類

###### － SSS (部分構造検索)

構造質問式に一致する回答のほか、置換可能な位置に置換基が付いた回答が得られる。

###### － CSS (閉構造部分構造検索)

構造質問式に一致する回答が得られる。結合非水素数 (CONNECT) で特に置換可能なように指定した位置以外は置換を許さない。

##### ・ マッチレベル

###### － クラス (CLASS)

特定原子または一般式グループで収録されている回答が得られる。

###### － 原子 (ATOM)

特定原子で収録されている回答のみが得られる。

###### － 不定 (ANY)

特定原子、一般式グループ、または R グループで収録されている回答が得られる。元素数レベルは「限定しない」設定にする。

##### ・ 元素数レベル

###### － 限定

マッチレベル「クラス」を指定したノードについて、元素数レベル「限定」を指定すると一般式グループがヒットした場合、その元素の種類と数の定義が合致する回答が得られる。

###### － 限定しない

マッチレベル「クラス」を指定したノードについて、元素数レベル「限定しない」を指定すると、元素数レベル「限定」を指定した場合に得られる回答に加え、一般式グループがヒットした場合、その元素の種類と数の定義が合致しない回答や一般式グループの定義がない回答も得られる。

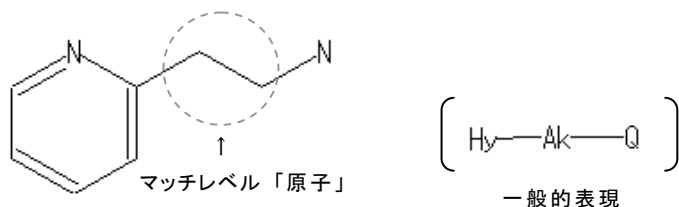
## B MARPAT ファイルの検索の仕組み

### マッチレベル

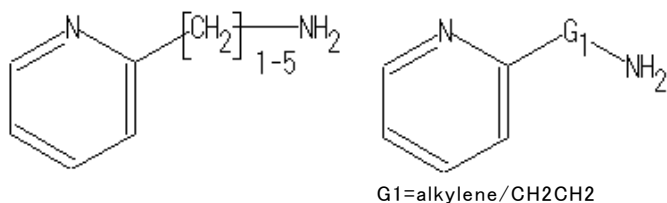
#### ■ マッチレベル - 原子

- ・ 構造質問式中の具体的な原子や一般式グループに対してマッチレベル「原子」を指定した場合は、具体的な原子にのみマッチする。
  - 例えば、構造質問式の Ak はレコード中の CH3 とはマッチするが alkyl とはマッチしない。
  - ヘテロ原子はレコード中の具体的原子とのみマッチする
  - 質問式の Cl または X は、レコード中の Cl とのみマッチ (X の場合は Cl, Br など) し、X とはマッチしない。

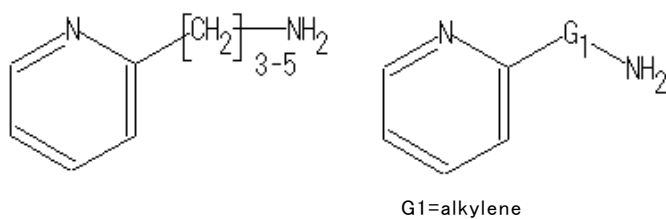
#### [構造質問式]



#### [検索される構造]



#### [検索されない構造]



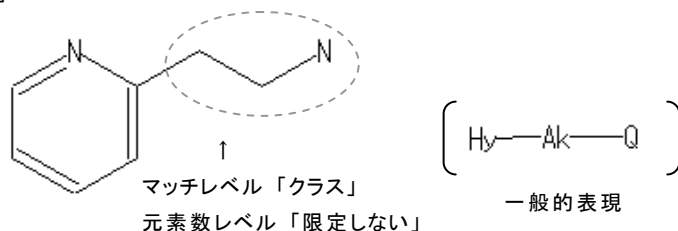
## B MARPAT ファイルの検索の仕組み

### マッチレベル

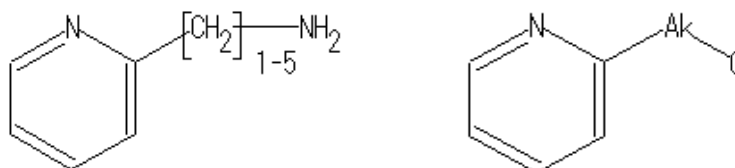
#### ■ マッチレベル - クラス

- ・ 構造質問式中の具体的な原子や一般式グループに対してマッチレベル「クラス」を指定した場合は、具体的な原子と一般式グループ双方にマッチする。
- ・ 構造質問式中から導かれた一般式グループと、レコード中に収録されている構造から導かれた一般式グループがマッチすることはない。
  - 例えば、鎖式炭素原子にクラスに指定して検索すると、レコード中の Ak にマッチし、さらにレコード中で使用されている acyl, alkyl, alkenyl, alkynyl, alkylidene, aralkyl, hydrocarbyl などの一般式テキスト・ショートカット (P.17 参照) ともマッチする。
  - 炭素鎖とマッチするテキスト・ショートカットを含むグループはこの他にも diloweralkylamino, alkynyloxy などがある。また CH<sub>2</sub>-CH<sub>2</sub>-alkyl, C(O)-alkylene-CH<sub>2</sub>-CH<sub>2</sub>-NH<sub>2</sub> のようにテキスト・ショートカットが原子グループと混在している場合もある。
  - ヘテロ原子にクラスを指定して検索すると、Q, M, X とマッチする。

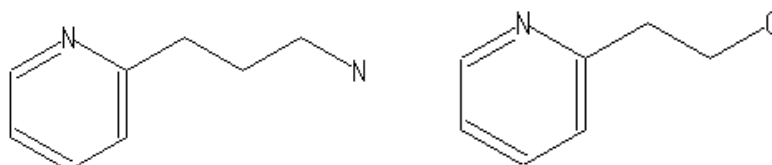
#### [構造質問式]



#### [検索される構造]



#### [検索されない構造]



マッチレベル「クラス」の検索は非常に強力であるが、構造質問式の全体ないし一部でもファイル中の一般式グループにマッチさせたくない場合は、マッチレベル「原子」の指定によって構造質問式中の原子や一般式グループと、レコード中の構造とのマッチングを制限する

## B MARPAT ファイルの検索の仕組み

### マッチレベル

#### ■ マッチレベル - 不定

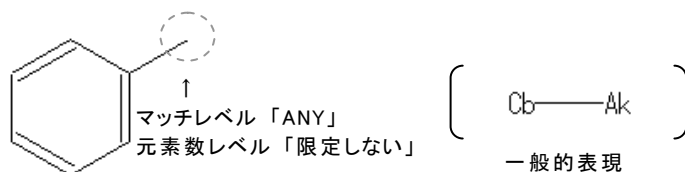
- ・ 構造質問式中の具体的な原子や一般式グループに対してマッチレベル「不定」を指定した場合は、具体的な原子、一般式グループおよび R のすべてにマッチする。

#### R ノード

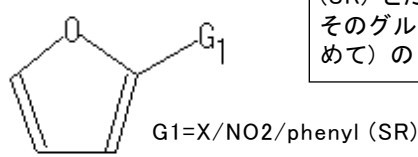
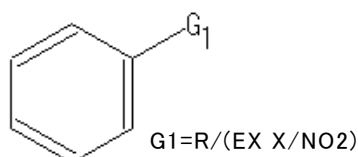
- 特許の中で「有機原子団」「電子吸引基」などのように記述されている、構造式で表現できないグループを表すのにレコード中の構造式では R を用いている。
- その特許で R グループを記述するのに用いられている語句は R の後に続く TX フィールドに記載されている。このテキストは特許に書かれていた記述に対応している。検索はできない。

- ・ クラスまたは原子を指定したノードは R とはマッチしない。
- ・ 特許に置換が許される (SO) あるいは置換が必要である (SR) ことが述べられていても、置換基が具体的に明示されていない場合には R が用いられる。

#### [構造質問式]



#### [検索される構造]



G グループの後に (SO) あるいは (SR) とだけ記されている場合は、そのグループに任意の数 (ゼロを含めて) の R が置換しえることを示す

#### ■ マッチレベル「不定」を用いる場合

- ・ 特許中にある付加的な構造情報が R の具体例としてあげられていれば「クラス」「原子」の指定でもヒットできるが、「不定」が必要なのは以下の場合である。
  - 質問構造のノードはそのまま残しておいて検索範囲を広げたい場合
  - 広範囲の回答を得たい場合
  - (SO) または (SR) をマッチさせたい場合
  - acyl をマッチさせたいとき

## B MARPAT ファイルの検索の仕組み

参考：一般式テキスト・ショートカット

### ■ 参考：一般式テキストの定義

#### 1. AK (直鎖, 分岐, 飽和, 不飽和を問わず, 1 以上の炭素鎖) に関するもの

- Alkanoyl Alkyl-C(O)-, H-C(O)-
- Alkenyl 1 以上の二重結合を持ち, 三重結合を持たず, 炭素数 2 以上の Ak  
Ak<EC (2-)C, BD (1-)D (0)T> -  
1 価 (たとえば, CH<sub>2</sub>-CH=CH-CH<sub>2</sub>- )
- Alkenylene 1 以上の二重結合を持ち, 三重結合を持たず, 炭素数 2 以上の Ak  
- Ak<EC (2-)C, BD (1-)D (0)T> -  
2 価 (たとえば, -CH=CH- )
- Alkenylenedioxy -O-alkenylene-O-
- Alkoxy alkyl-O- (Alkyloxy または alkoxy)
- Alkyl 単結合のみを持つ飽和の AK (線上 / 枝分かれ)  
1 価 (たとえば, Me-, Et-, t-Bu- )
- Alkylene -alkyl-  
2 価 (たとえば, -CH<sub>2</sub>- )
- Alkylidene alkyl=  
1 価 (たとえば, =CH-CH-CH<sub>3</sub> )
- Alkynyl 二重結合を持たず, 以上の三重結合を持ち, 炭素数 2 以上の AK  
Ak<EC (2-)C, BD (1-)T (0)D> -  
1 価 (たとえば, CH=C-, HC=C-CH<sub>2</sub>- )
- Lower 炭素数 1-6
- Loweralkyl 炭素数 1-6 の単結合のみを持つ AK
- Perhaloalkyl 水素全てがハロゲンで置き換えられた Alkyl

#### 2. CB (単環, 多環, 飽和, 不飽和を問わず炭素環), HY (単環, 多環, 飽和, 不飽和を問わず 1 以上のヘテロ原子を持つ環) に関するもの

- Aryl 1 以上の芳香環, 6 以上のノーマライズド結合, 1 以上の 6 員環をもつ CB
- Arylene -aryl-  
2 価 (たとえば, phenylene)
- Cycloalkenyl 1 以上の二重結合を持ち, 三重結合を持たない CB  
1 価 (たとえば, cyclopentadienyl, cyclohexenyl)
- Cycloalkyl 単結合のみを持つ CB  
1 価 (たとえば, cyclopropyl, decahydronaphthyl)
- Heteroaryl 1 以上の芳香環, 6 以上のノーマライズド結合, および 1 以上の 6 員環を持つ HY, または, 1 以上の芳香環, 2 以上の二重結合, および 1 以上の 5 員環を持つ HY (たとえば, pyridyl, benzopyranyl)

#### 3. 混合

- Acyl AK (置換基はあってもよい) と結合しているカルボニル基で, カルボニルの炭素が結合点である  
R と結合しているカルボニル基で, カルボニルの炭素が結合点である  
CHO
- Aralkyl 1-3 の aryl 基と結合している alkyl 基
- Hydrocarbyl AK  
CB  
CB に結合している AK で, AK が結合点である  
AK に結合している CB で, CB が結合点である  
CB に結合している CB で, 片方の CB が結合点である

## B MARPAT ファイルの検索の仕組み

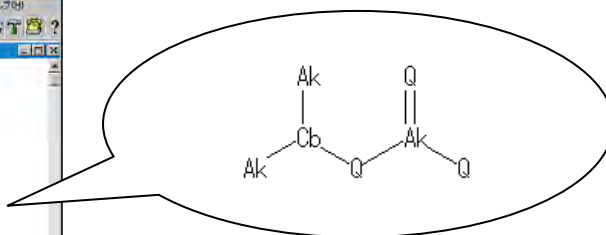
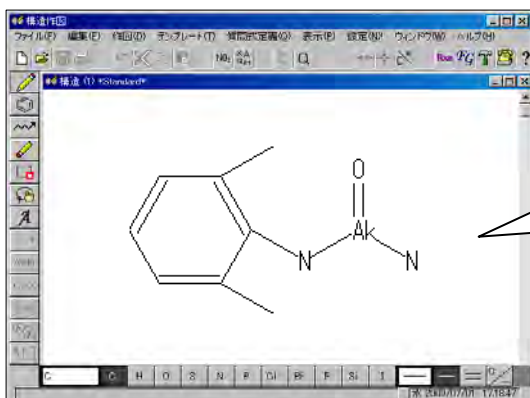
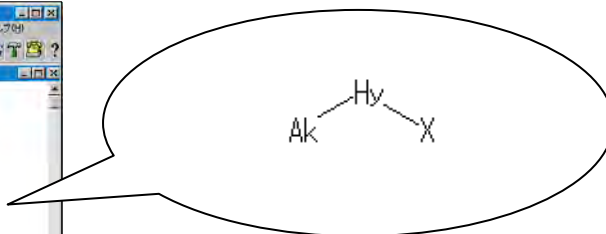
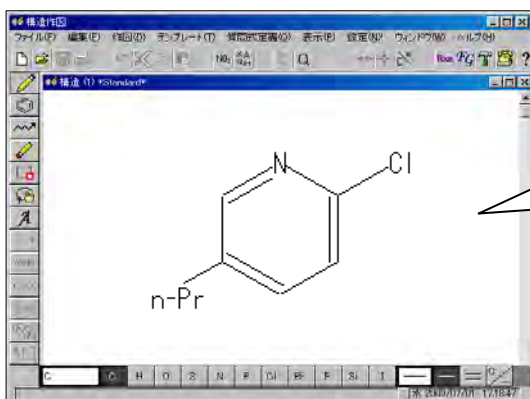
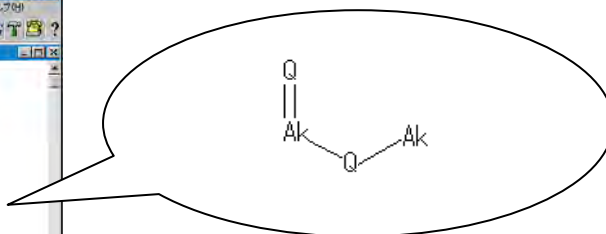
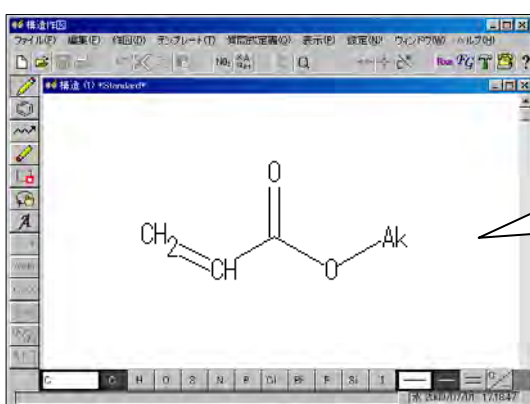
### MARPAT ファイルの検索の仕組み

■ MARPAT ファイルでは、構造質問式中のマッチレベルの指定によって得られる回答が異なる。

- ・ これは、システムが構造質問式中の特定の原子や属性の情報と共に、「構造質問式から導かれる一般式グループ」を識別に用いているためである。
- ・ MARPAT ファイルの検索システムについての理解を深めるため、構造質問式を作図する際に、それが表現する可能性のある一般式を同時に考えてみると役立つ。

[実際に作図した構造質問式]

[構造質問式から導かれる一般的表現]



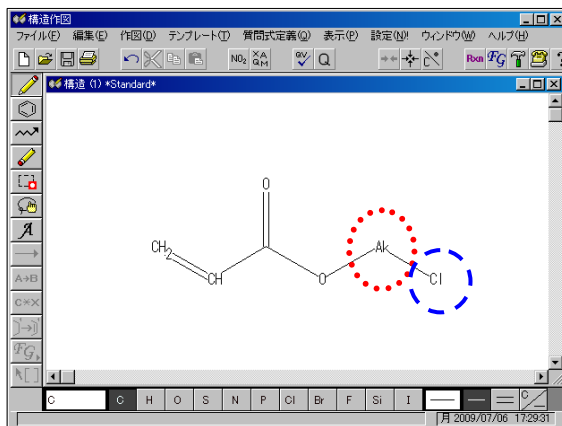
## B MARPAT ファイルの検索の仕組み

### MARPAT ファイルの検索の仕組み

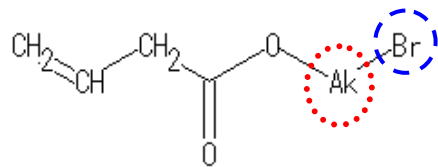
#### ■ MARPAT ファイルの検索の仕組み

- ・ システムは一般式グループのうちどれが原子から導かれたものであるか、どれが本来質問式に含まれていたものを識別する。
- ・ 構造質問式から導かれる一般式グループと同様に、「MARPAT ファイルに収録されている構造から導かれる一般式グループ」も存在し、これも識別に用いられる。

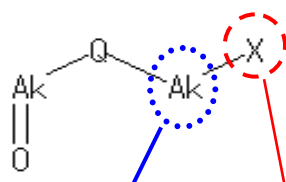
[実際に作図した構造質問式]



[レコード中の構造]



[構造質問式から導かれる一般的表現]

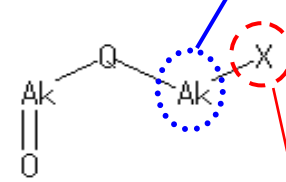


構造質問式中に本来含まれていた一般式グループ

構造質問式中の原子から導かれた一般式グループ

[レコード中の構造から導かれる一般的表現]

レコード中に本来含まれていた一般式グループ



レコード中の構造から導かれた一般式グループ

#### ポイント

構造質問式から導かれた一般式グループと、レコード中の構造から導かれた一般式グループがマッチすることはない。

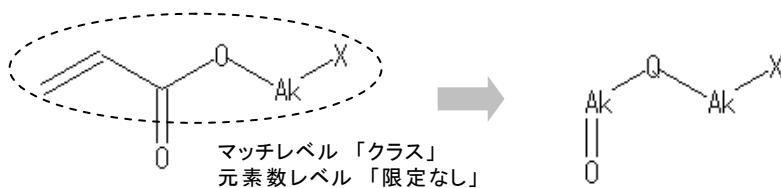
## B MARPAT ファイルの検索の仕組み

### MARPAT ファイルの検索の仕組み

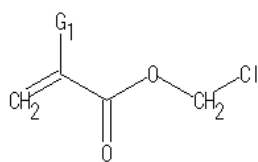
- ・ マッチレベル「クラス」を指定することによって、構造質問式から導かれた一般式グループと、レコード中の構造に本来存在していた一般式グループをマッチさせることができる。
- ・ マッチレベル「クラス」を指定しても、構造質問式から導かれた一般式グループが、レコード中の構造から導かれた一般式グループとマッチすることはない。

[構造質問式]

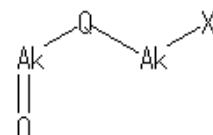
[一般的表現]



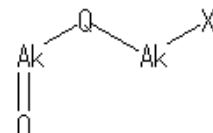
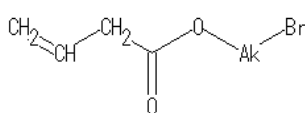
[レコード中の構造]



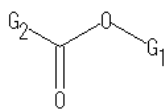
G1=H/Me



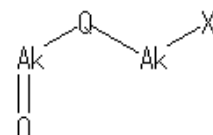
ヒットする



ヒットしない



G1=alkyl (SO X)/Ph  
G2=alkenyl<2-6>/loweralkyl



ヒットする

MARPAT ファイルでは以下のすべての構造を比較して、検索に使用することによってノイズの少ない検索が実行される仕組みになっている。

- ・ 構造質問式中に元から存在する具体的な原子・一般式グループ
- ・ 構造質問式から導かれた一般式グループ
- ・ レコード中の構造中に元から存在する具体的な原子・一般式グループ
- ・ レコード中の構造から導かれた一般式グループ

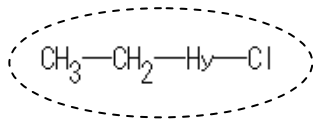
## B MARPAT ファイルの検索の仕組み

### MARPAT ファイルの検索の仕組み

#### ■ 演習

- ・ 構造質問式およびレコード中の構造から導かれる一般的表現の構造を書き、「ヒットする」「ヒットしない」にチェックを付ける。

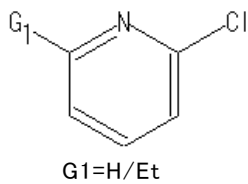
[構造質問式]



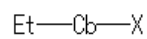
マッチレベル 「クラス」  
元素数レベル 「限定なし」

[一般的表現]

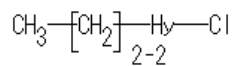
[レコード中の構造]



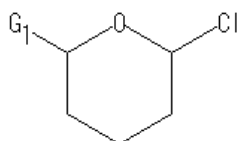
ヒットする/ヒットしない



ヒットする/ヒットしない



ヒットする/ヒットしない



G<sub>1</sub>=alkyl<1-6>/alkenyl<2-8>

ヒットする/ヒットしない

## B MARPAT ファイルの検索の仕組み

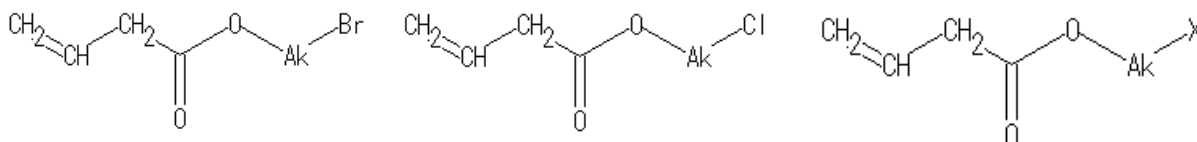
### 構造作図ポイント 1

#### 構造作図ポイント 1

マッチレベルでコントロールする前に、「特定の原子を用いた作図」をするのか、または「一般式グループを用いた作図」をするのかを検討する。

特に回答が少ない場合は、検討してみるとよい。

- 目的物質として以下のような回答をすべて含めたい場合



- 構造質問式

構造質問式の具体的な原子 (Cl) を X (ハロゲン) に変更する

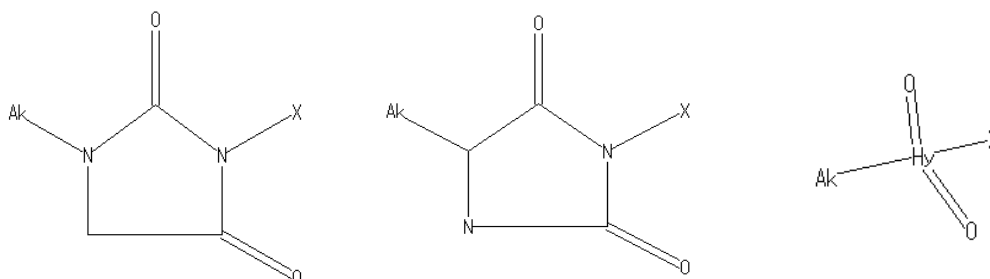
=> S L1 FUL  
L3      3299 SEA SSS FUL L1

=> S L4 FUL  
L6      3971 SEA SSS FUL L4

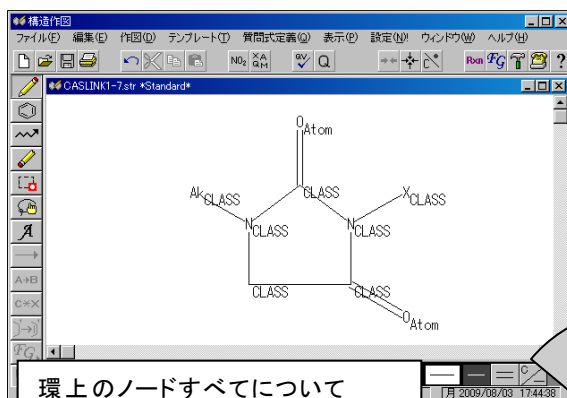
## B MARPAT ファイルの検索の仕組み

### 構造作図ポイント 1

- 目的物質として以下の物質を回答に含めたい場合



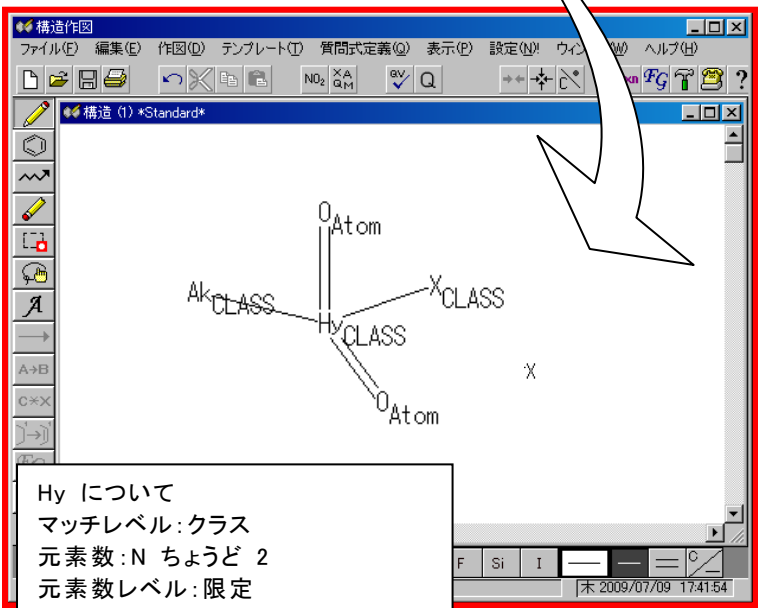
- 構造質問式



環上のノードすべてについて  
 マッチレベル: クラス  
 元素数レベル: 限定  
 孤立環

=> S L1 FUL  
 L3            12164 SEA SSS FUL L1

具体的な原子で作図した環を  
 Hy (一般式グループ) に変更する



Hy について  
 マッチレベル: クラス  
 元素数: N ちょうど 2  
 元素数レベル: 限定  
 一般式属性: 単環式

=> S L4 FUL  
 L6            14410 SEA SSS FUL L4

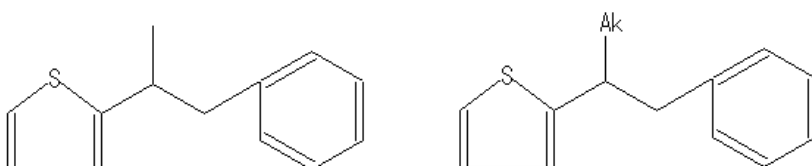
## B MARPAT ファイルの検索の仕組み

### 構造作図ポイント 2

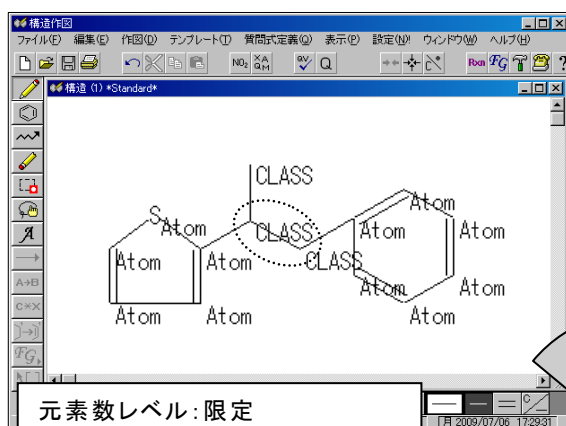
#### 構造作図ポイント 2

分岐・側鎖が置換する位置やその原子の種類・数が重要な場合は、マッチレベル「原子」を指定する。

- 目的物質として以下の物質を回答に含めたい場合



- 構造質問式

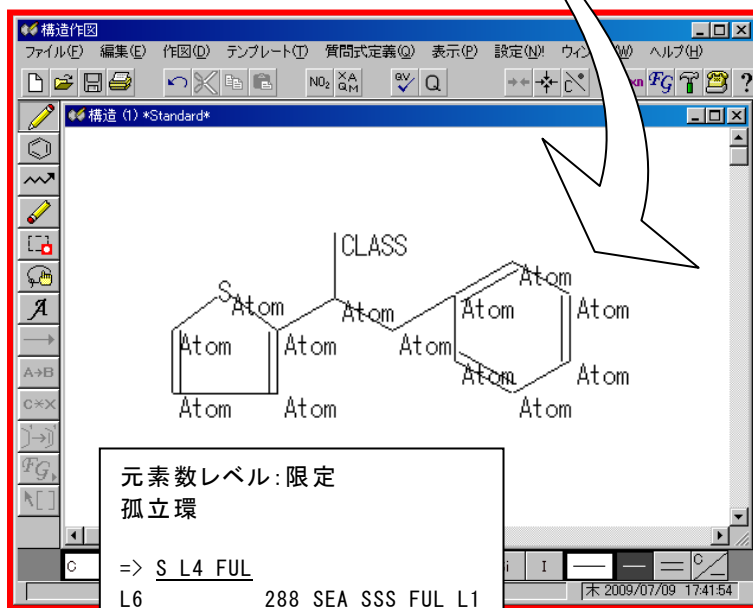


元素数レベル: 限定  
孤立環

=> S L1 FUL

L3 1720 SEA SSS FUL L4

点線で囲った部分の炭素原子に  
マッチレベル「原子」を指示する



元素数レベル: 限定  
孤立環

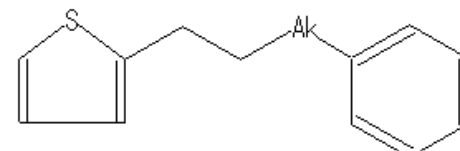
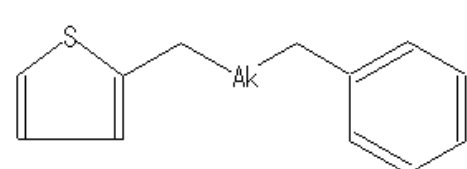
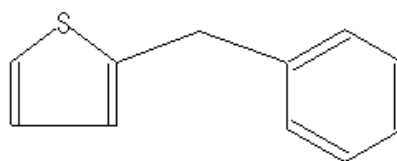
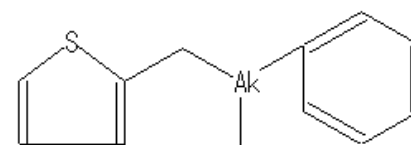
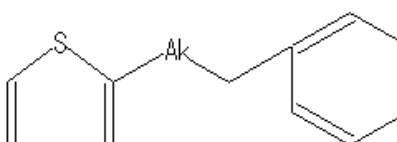
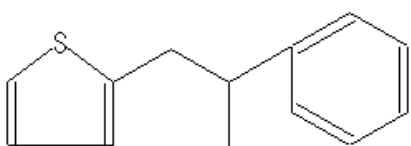
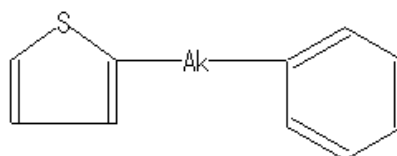
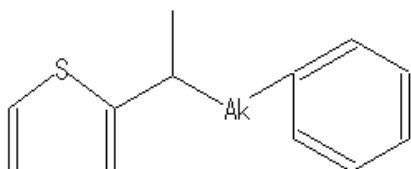
=> S L4 FUL

L6 288 SEA SSS FUL L1

## B MARPAT ファイルの検索の仕組み

### 構造作図ポイント 2

#### ■ 左ページの検索でヒットしない回答



このようにマッチレベルの指定で回答が変化するので、回答として要求されるものとマッチレベルの指定によって回答中に含まれるもの、含まれなくなるものの両方を考慮する必要がある

## B MARPAT ファイルの検索の仕組み

### 一般式属性

■ 構造質問式とレコード中の構造を照合する際には、一般式属性も用いられる。

- ・ 構造質問式中の具体的原子からは、一般式グループが導かれると同時に、一般式属性 (GGC) も生成される。
  - HIC : 炭素原子 7 個以上
  - LOC : 炭素原子 6 個以下
  - HIQ : ヘテロ原子 2 個以上
  - LOQ : ヘテロ原子 1 個
  - SAT : 飽和 (すべて単結合)
  - UNS : 不飽和 (単結合以外の結合が 1 個以上)
  - BRA : 分岐 (グループ中の原子 1 個以上がグループ中のほかの原子 3 個以上と結合)
  - LIN : 直鎖 (グループ中の原子はすべてグループ中のほかの原子 2 個以下と結合)
  - PCY : 多環式
  - MCY : 単環式
- ・ 構造質問式中の具体的原子から導かれる一般式グループに対して生成される一般式グループカテゴリの組み合わせは、検索のタイプ (SSS または CSS) によって異なる。
  - SSS 検索では、構造質問式から導かれた Ak に対して HIC と LOC の両方が生成される。
  - 一般式属性を生成する際に水素数は考慮されない。

構造質問式	環	検索タイプ	一般的表現	自動的に付与される一般式属性
		SSS	Ak	HIC/LOC UNS BRA
		CSS	Ak	LOC UNS BRA
	孤立	SSS	Cb	LOC SAT LIN MCY
		CSS	Cb	LOC SAT LIN MCY
	非孤立	SSS	Cy	HIC/LOC SAT/UNS LIN/BRA MCY/PCY
		CSS	Cb	LOC SAT LIN MCY
	孤立	SSS	Hy	LOC SAT LIN MCY LOQ
		CSS	Hy	LOC SAT LIN MCY LOQ
	非孤立	SSS	Hy	LOC/HIC SAT/UNS LIN/BRA MCY/PCY LOQ/HIQ
		CSS	Hy	LOC SAT LIN MCY LOQ

- ・ 構造質問式中に元から存在する一般式グループに特に指定していない場合は、それぞれ次のような一般式属性の組み合わせが生成される。
  - Ak : HIC/LOC SAT/UNS LIN/BRA
  - Cb : HIC/LOC SAT/UNS LIN/BRA MCY/PCY
  - Hy : HIC/LOC SAT/UNS LIN/BRA MCY/PCY HIQ/LOQ
  - Cy : HIC/LOC SAT/UNS LIN/BRA MCY/PCY HIQ/LOQ

## B MARPAT ファイルの検索の仕組み

### 一般式属性

#### ■ レコード中の構造に含まれている一般式グループに対する属性

- レコード中の一般式グループにそれを限定する値が付随している場合は、自動的に一般式属性が生成される。

```
G4      = 8 / heterocycle <containing 1 heteroatom, 1 N, 8 C, aromatic, 6 normalized bonds, 2 C fusion atoms, bicyclic, (1) 5-membered ring, (1) 6-membered ring> (opt. substd.) / heterocycle <containing 1 heteroatom, 1 N, 12 C, aromatic, 12 normalized bonds, 4 C fusion atoms, tricyclic, (1) 5-membered, (2) 6-membered rings> (opt. substd.)
```

元素数から発生する一般式属性	: HIC/LOC HIQ/LOQ
結合から発生する一般式属性	: SAT/UNS
環の数から発生する一般式属性	: MCY/PCY
結合数から発生する一般式属性	: LIN/BRA
縮合する元素の数から発生する一般式属性	: MCY/PCY

- 一般式グループのテキスト・ショートカットもそれぞれ適当な一般式属性を持つ。

```
- 例: alkyl      Ak HIC/LOC SAT LIN/BRA
      alkenyl   Ak HIC/LOC UNS LIN/BRA
      alkynyl   Ak HIC/LOC UNS LIN/BRA
      loweralkyl Ak LOC SAT LIN/BRA
      cycloalkenyl Cb HIC/LOC UNS LIN/BRA MCY/PCY
      aryl      Cb HIC/LOC UNS LIN/BRA MCY/PCY
      heteroaryl Hy HIC/LOC UNS LIN/BRA MCY/PCY
```

- テキストでの値が付随していない場合は、デフォルトの一般式属性が生成される。

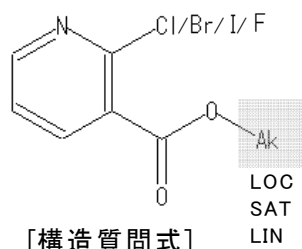
## B MARPAT ファイルの検索の仕組み

### 一般式属性

■ 構造質問式とレコード中の構造を照合する際には、一般式属性も照合に用いられる。

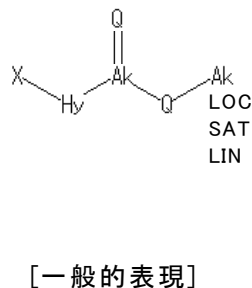
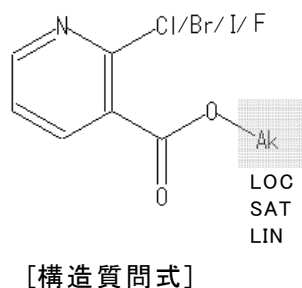
・ 下図の Ak にマッチレベル「クラス」(元素数レベル「限定」)を指定すると ...

- 原子数 1-6 で単結合のみを持つ炭素直鎖とマッチする。
- LOC SAT LIN である任意の Ak とマッチする。



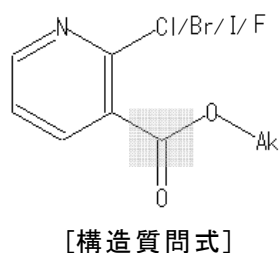
・ 下図の Ak (LOC, SAT, LIN) にマッチレベル「原子」を指定すると ...

- 原子数 1-6 で単結合のみを持つ炭素直鎖とマッチする。
- CH<sub>3</sub> や -(CH<sub>2</sub>)<sub>5</sub>-CH<sub>3</sub> とはマッチするが t-Bu とはマッチしない。



・ 下図のカルボキシル基の C にマッチレベル「クラス」(元素数レベル「限定しない」)を指定すると、部分構造の場合、Ak に対するすべての属性の対が生成される。 その結果...

- カルボキシル基の C とマッチする。
- Ak とマッチする。

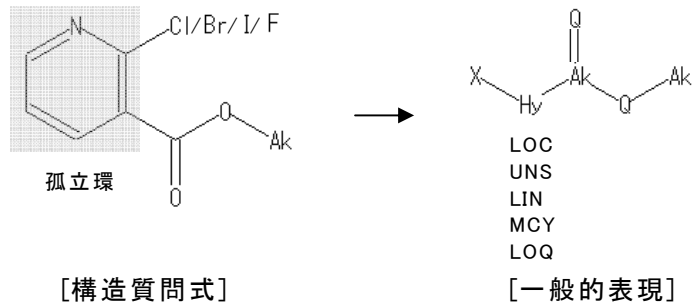


## B MARPAT ファイルの検索の仕組み

### 一般式属性

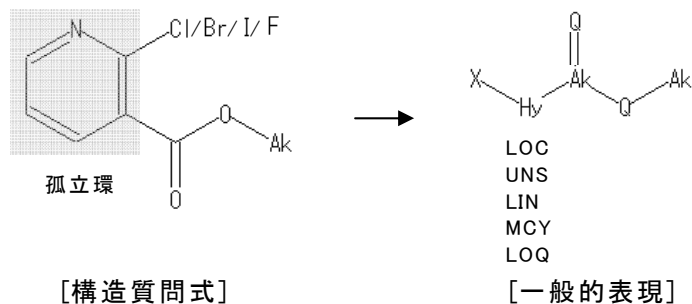
- 下図のピリジン環にマッチレベル「クラス」(元素数レベル 限定) を指定すると ...

- ピリジンとマッチする.
- LOC UNS LIN MCY LOQ である Hy とマッチする.
- HIC SAT BRA PCY HIQ である Hy とはマッチしない.



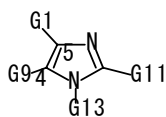
- 下図のピリジン環にマッチレベル「クラス」(元素数レベル 限定しない) を指定すると ...

- ピリジンとマッチする.
- Hy とマッチする.



- 回答例 (ピリジン環の部分にマッチレベル「クラス」(元素数レベル 限定しない) に指定したため heterocycle でヒットした回答)

#### MSTR 1



G13 = 74

074—G14

G14 = heterocycle <containing zero or more O, zero or more S, 1 or more double bonds>(opt. substd. by (up to 3) G22)

G22 = F / alkoxy carbonyl <containing 1-4 C>

Derivative: and acid addition salts and metal complexes

Patent location: claim 1

このレコード中の heterocycle から自動的に発生する一般式属性は HIQ/LOQ HIC/LOC LIN/BRA SAT/UNS MCY/PCY であり、これが構造質問式から導かれる一般式属性と一致するためにこのような回答が得られる

## B MARPAT ファイルの検索の仕組み

### 構造作図ポイント 3

#### 構造作図ポイント 3

具体的な構造を作図した環に対してマッチレベル「クラス」に指定する場合は、元素数レベル「限定」を指定するとノイズを減らすことができる。

#### ■ 構造質問式

元素数レベル「限定しない」、孤立環  
↓  
Hy (LOC SAT LIN MCY HIQ)  
[一般的表現]

=> S L1 FUL  
L3 60508 SEA SSS FUL L4

クラスを指定した環の元素数レベルを「限定」する

元素数レベル「限定」、孤立環  
↓  
Hy (LOC SAT LIN MCY HIQ)  
[一般的表現]

=> S L1 FUL  
L3 32572 SEA SSS FUL L4

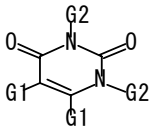
## B MARPAT ファイルの検索の仕組み

### 構造作図ポイント 3

#### ■ 回答例

##### ・ 元素数レベル「限定」

**MSTR 1**



レコード中の一般式グループを限定する値から自動的に導かれる一般式属性は LOC SAT/UNS LIN MCY LOQ/HIQ

G1 = heterocycle <containing 5-6 atoms, 1-3 heteroatoms, zero or more N, zero or more O, zero or more S (no other heteroatoms), 5- to 6-membered monocyclic ring> (opt. substd. by 1 or more G5)

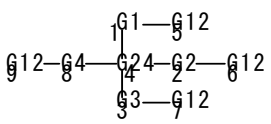
Patent location: claim 1

Note: or tautomers, solvates, hydrates, prodrugs or pharmaceutically acceptable salts

Stereochemistry: or stereoisomers

##### ・ 元素数レベル「限定しない」

**MSTR 1**



レコード中の一般式グループを限定する値から自動的に導かれる一般式属性は LOC/HIC SAT/UNS LIN/BRA MCY/PCY LOQ/HIQ

G12 = heterocycle <containing 6 or more atoms, 1 or more heteroatoms, 1 or more N, attached through 1 or more N, 0 or more double bonds, (+1) charge, mono- or polycyclic, 1 or more 6-membered rings> (opt. substd. by 1 or more G13)

G13 = heterocycle <containing 1 or more heteroatoms, 3-24 C> (opt. substd.)

Patent location: claim 1

Note: substitution is restricted

一般属性は原子数として HIQ または LOQ の 2 種類しか存在しないため、元素数レベルを「限定」しないとノイズが多数得られる場合がある。そのような場合は、元素数レベルを「限定」する。

元素数指定なしの Ak や Hy で作図し、マッチレベル「クラス」の場合は、どのような一般式属性を付加しても、元素数レベルの違い（限定/限定しない）は回答に影響しない

## B MARPAT ファイルの検索の仕組み

### 検索できないデータ

#### ■ MARPAT ファイルで検索できないデータ

##### (1) 作図時 : 利用できない属性

- 結合属性「環/鎖」の指定
- Ak について元素数 0 (ゼロ) の指定

##### (2) 検索時 : 検索タイプ・機能

- 完全一致検索, ファミリー検索
- 複数の構造質問式の L 番号をブール演算子で掛け合わせる検索
- スクリーン検索

##### (3) 検索は実行できるが, 検索に考慮されない (検索できない) レコード中のデータ

- 一般式グループテキストに対する置換基・置換数
- 一般式グループテキストの属性 (\* 元素数など一部検索可能)
- G グループの選択枝の個数

#### 参考 : 条件が存在する複雑なマルクーシュ構造の収録

- ・ クレームに「ただし R=Me ならば R1≠OH であり...」といった複雑な条件が記載されているマルクーシュ構造があった場合, MARPAT ファイルには以下の優先順位で構造が収録される.
  - (1) 可能なかぎり一つの MSTR で収録する. → 構造検索可能
  - (2) MSTR を複数に分割して収録する. → 構造検索可能
  - (3) substitution is restricted のように表記されるのみ. → 基本索引検索 (構造検索不可)
- ・ 上記 (3) のように NTE フィールドにテキスト情報として記載され, すべてのバリエーションが MSTR に入らない場合がある.

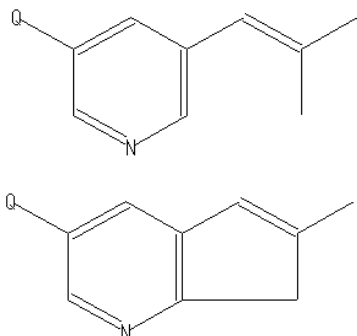
## B MARPAT ファイルの検索の仕組み

### 構造作図ポイント 4

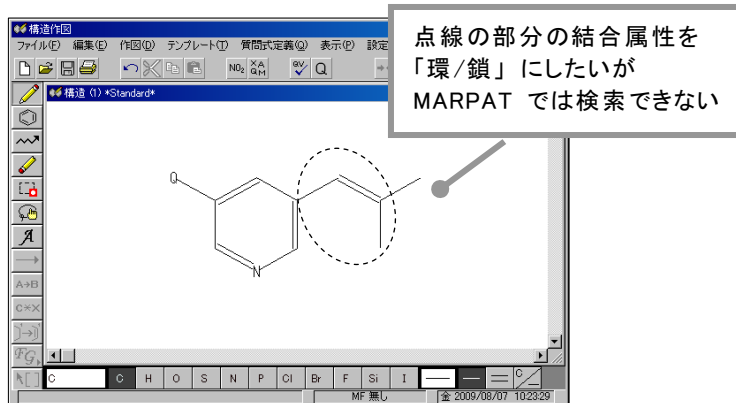
#### 構造作図ポイント 4

MARPAT ファイルでは、結合属性「環/鎖」に作図して検索することができない。そのような場合は、G グループなどその他の作図を利用する。

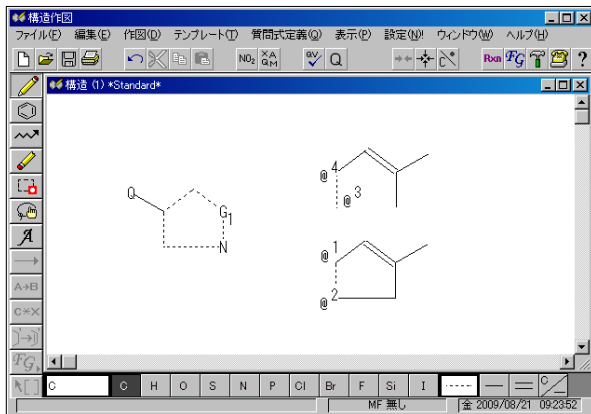
#### 【目的の物質】



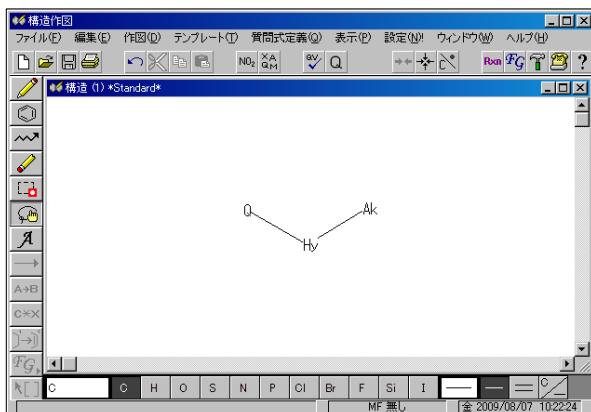
#### 【作図したい構造図案】



- ・ 作図の代替例 1 : G グループの利用



- ・ 作図の代替例 2 : 一般式グループ記号の利用



## B MARPAT ファイルの検索の仕組み

### 構造作図ポイント 5

#### 構造作図ポイント 5

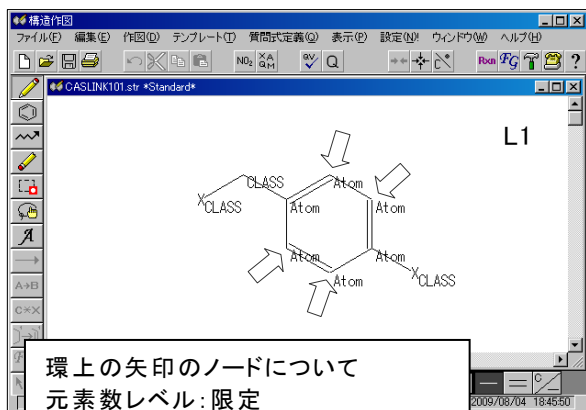
MARPAT ファイルに収録されている構造中の

- ・ 一般式グループテキストに対する置換基・置換基数
- ・ 一般式グループテキストの属性 (\* 元素数など一部検索可能)
- ・ G グループの選択枝の個数

は検索することができないため、構造質問式中で水素を作図したり、置換基制御のために結合水素数、または結合非水素数を指定しても、回答中にノイズが含まれることがある。

ただし、すべてこれらの指定が無視されるわけではなく、上記以外の部分については有効であるため、必要部位については作図する際に指定する。

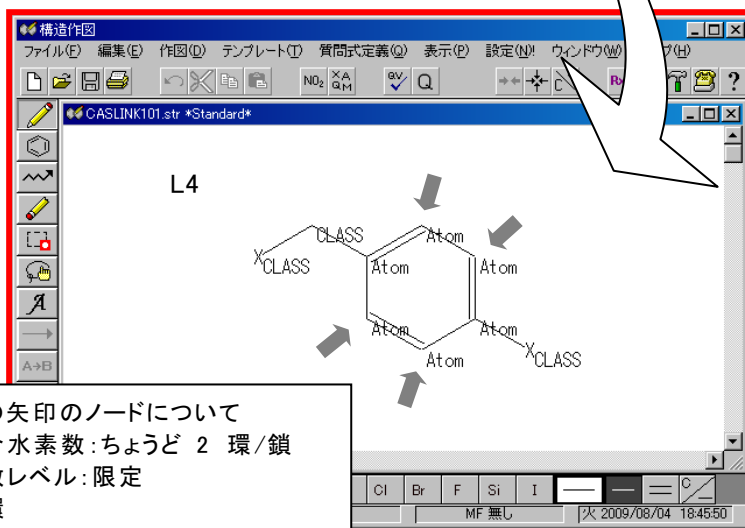
#### 【作図する構造図】



環上の矢印のノードについて  
元素数レベル: 限定  
孤立環

=> S L1 FUL  
L3 40015 SEA SSS FUL L1

矢印のノードに置換基が不要な場合は  
結合非水素数などを指定する



環上の矢印のノードについて  
非結合水素数: ちょうど 2 環/鎖  
元素数レベル: 限定  
孤立環

=> S L1 FUL  
L6 9008 SEA SSS FUL L4

## B MARPAT ファイルの検索の仕組み

### 構造作図ポイント 5

=> FILE MARPAT

=>

Uploading C:\My\_Document\STN Express 8.4\Queries\CASLINK102.str  
L1 STRUCTURE UPLOADED

=> S L1

:

=> S L1 FUL

:

L3 40015 SEA SSS FUL L1

=>

Uploading C:\My\_Document\STN Express 8.4\Queries\CASLINK101.str  
L4 STRUCTURE UPLOADED

=> S L4

:

=> S L4 FUL

:

L6 9008 SEA SSS FUL L4

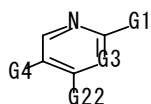
=> S L3 NOT L6

L7 31007 L3 NOT L6

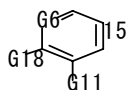
=> D FQHIT

L7 ANSWER 1 OF 31007 MARPAT COPYRIGHT 2009 ACS on STN

#### MSTR 1



G4 = 15



G11 = 26

~~26~~ 12-G13

G13 = 28

~~28~~ 25-G26

G26 = Ph (opt. substd. by 1 or more G27)

G27 = CF<sub>3</sub> / F

Patent location:

claim 1

Note:

or pharmaceutically acceptable salts, or solvates

## B MARPAT ファイルの検索の仕組み

### 構造質問式改善の指針

#### ■ 構造質問式の特定性

- ・ MARPAT ファイルでも REGISTRY ファイルなど、他の構造検索ファイルと同様、構造質問式の定義が厳密なほど、自動的に生成される一般式属性の精度も高く、より適合率の高い回答を得ることができる。

#### ■ 一般式グループ記号

- ・ 回答数が少ないときには、構造質問式中の具体的な原子を一般式グループ記号に変更する。
- ・ 一般式グループの内容を正確に指定するには一般式属性や元素数を指定する。

#### ■ 一般式属性

- ・ マッチレベル「クラス」を指定した構造質問式中の一般式グループに一般属性や元素数を指定しないと、それに対応するすべての一般式グループと、該当する具体的な原子の構造すべてがマッチする。
- ・ 構造質問式を限定してもよい場合は、一般式属性や元素数を指定し、元素数レベルを「限定」する。

#### ■ 環の孤立

- ・ 環系の孤立を指定しないと、生成する一般式属性が非常に広汎なものになる。

- ピリジン環を非孤立/孤立させたときに発生する、一般式グループは

孤立させないと	Hy* HIC/LOC UNS LIN/BRA MCY/PCY HIQ/LOQ
孤立させると	Hy* LOC UNS LIN MCY LOQ

- ベンゼン環を非孤立/孤立させたときに発生する、一般式グループは

孤立させないと	Cy* HIC/LOC UNS LIN/BRA MCY/PCY HIQ/LOQ
孤立させると	Cb* LOC UNS LIN MCY

#### ■ 結合

- ・ 結合を可変にするとノイズを生じる場合がある。できれば、不定結合やエグザクト/ノーマライズド結合よりも、ノーマライズドやエグザクト結合を指定する。

## C 検索テクニック



## C 検索テクニック

### CASLINK の応用

#### ■ CASLINK の応用

- ・ CASLINK は特別な検索の流れが組み込まれている STN のファイルクラスターのひとつである。REGISTRTY ファイルから CAplus ファイルへのクロスオーバー検索と CAplus ファイルと MARPAT ファイルの重複除去が自動的に行われる。
- ・ CASLINK で得られた回答は、CAplus ファイルの検索フィールドで絞り込むことができる。

#### 応用 1 : 絞り込み検索

効率よく回答をチェックするには、特許出願日や技術用語を組み合わせて絞り込み検索を実行し、より適合率の高い回答から優先的に確認する。

#### 応用 2 : サブセット検索

構造検索を実行する前に、キーワードや書誌情報による検索を実行し、その回答を母集合にしたサブセット検索を実行することもできる。

- ・ CASLINK では MARPAT ファイルを優先する CAplus ファイルと MARPAT ファイルの重複除去が自動的に行われる。

#### 応用 3 : 重複文献除去の優先順位

重複文献除去の優先順位を変更したい場合は、CASLINK の途中で作成される L 番号を用いて DUP REM コマンドを実行する。

#### 応用 4 : INCOMPLETE の回答が得られた場合

CASLINK の実行途中で構造検索結果に INCOMPLETE の回答が含まれた場合は、該当ファイルに入って、/COM の検索を行った後に、DUP REM コマンドを実行する。

- ・ CASLINK の中で、重複除去した回答を保存することができるが、呼び出した回答に対して、さらに別の検索式を演算することはできない。

#### 応用 5 : 保存

個別のファイルで保存すれば、後日呼び出した回答でも複雑な条件を組み合わせた検索が実行できるので、さらに活用度が増す。

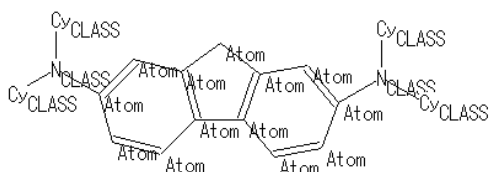
## C 検索テクニック

### CASLINK の応用

#### 応用 1 : 絞込み検索

効率よく回答をチェックするには、特許出願日や技術用語を組み合わせて絞り込み検索を実行し、より適合率の高い回答から優先的に確認する。

#### ■ 構造質問式



環は孤立

#### ■ 検索例

=> FILE CASLINK  
:

=> S L1  
:

=> S L1 FUL

S L1 SSS FUL FILE=REGISTRY

:  
100.0% PROCESSED 47973 ITERATIONS  
SEARCH TIME: 00.00.03

958 ANSWERS

L4 958 SEA SSS FUL L1

S L4 SSS FUL FILE=MARPAT

:  
100.0% PROCESSED 10810 ITERATIONS  
SEARCH TIME: 00.00.03

170 ANSWERS

L5 170 SEA SSS FUL L1

S L4 FILE=CAPLUS

L6 314 FILE CAPLUS

SET DUPORDER FILE  
SET COMMAND COMPLETED

DUP REM L5 L6

PROCESSING COMPLETED FOR L5

PROCESSING COMPLETED FOR L6

L7 414 DUP REM L5 L6 (70 DUPLICATES REMOVED)

ANSWERS '1-170' FROM FILE MARPAT

ANSWERS '171-414' FROM FILE CAPLUS

## C 検索テクニック

### CASLINK の応用

```
=> S (ELECTROPHOT? OR ELECTROLUMI?) AND 19900827>=AD  
S (ELECTROPHOT? OR ELECTROLUMI?) AND 19900827>=AD FILE=CAPLUS  
L8 32689 FILE CAPLUS
```

```
=> S L7 AND L8
```

```
S L6 AND L8 FILE=CAPLUS  
L9 7 FILE CAPLUS
```

```
S L5 AND L8 FILE=CAPLUS  
L10 3 FILE CAPLUS
```

```
S L10 AND L5 FILE=MARPAT  
L11 3 FILE MARPAT
```

```
DUP REM L11 L9  
PROCESSING COMPLETED FOR L11  
PROCESSING COMPLETED FOR L9  
L12 9 DUP REM L11 L9 (1 DUPLICATE REMOVED)  
ANSWERS '1-3' FROM FILE MARPAT  
ANSWERS '4-9' FROM FILE CAPLUS
```

```
=> S L12 AND P/DT
```

```
S L9 AND P/DT FILE=CAPLUS  
L13 7 FILE CAPLUS
```

```
S L11 AND P/DT FILE=CAPLUS  
L14 3 FILE CAPLUS
```

```
S L14 AND L11 FILE=MARPAT  
L15 3 FILE MARPAT
```

```
DUP REM L15 L13  
PROCESSING COMPLETED FOR L15  
PROCESSING COMPLETED FOR L13  
L16 9 DUP REM L15 L13 (1 DUPLICATE REMOVED)  
ANSWERS '1-3' FROM FILE MARPAT  
ANSWERS '4-9' FROM FILE CAPLUS
```

```
=> D 1-9 TI
```

```
L16 ANSWER 1 OF 9 MARPAT COPYRIGHT 2009 ACS on STN DUPLICATE 1  
TI Electrophotographic photoreceptors using fluorene derivative  
charge transporting agent
```

```
L16 ANSWER 2 OF 9 MARPAT COPYRIGHT 2009 ACS on STN  
TI Overcoated electrophotographic photoreceptor containing metal  
acetylacetonate in polymer layer
```

```
L16 ANSWER 3 OF 9 MARPAT COPYRIGHT 2009 ACS on STN  
TI Hydroxyarylamine compounds for electrophotographic photoreceptors with  
improved solvent resistance
```

```
L16 ANSWER 4 OF 9 CAPLUS COPYRIGHT 2009 ACS on STN  
TI Photoconductive imaging members with fluorene polyester hole transporting  
layers
```

技術用語や出願日による限定

MARPAT ファイル単独ではこのような検索は実行できないが、CASLINK の場合は利用可能

CAplus ファイルの回答には非特許文献も含まれるため  
資料種類 (P/DT) による検索で特許に限定

MARPAT ファイルの書誌情報中はヒットタームハイライトなし

C 検索テクニック

CASLINK の応用

L16 ANSWER 5 OF 9 CAPLUS COPYRIGHT 2009 ACS on STN  
 TI Photoconductive imaging members with polyurethane hole transporting layers

L16 ANSWER 6 OF 9 CAPLUS COPYRIGHT 2009 ACS on STN  
 TI **Electrophotographic** photoreceptors using fluorenylidene derivative as charge-transporting agent

L16 ANSWER 7 OF 9 CAPLUS COPYRIGHT 2009 ACS on STN  
 TI Organic charge transport layer in **electrophotographic** photoreceptor

L16 ANSWER 8 OF 9 CAPLUS COPYRIGHT 2009 ACS on STN  
 TI **Electrophotographic** photoconductor

L16 ANSWER 9 OF 9 CAPLUS COPYRIGHT 2009 ACS on STN  
 TI **Electrophotographic** charge-generating tetrakisazo pigments

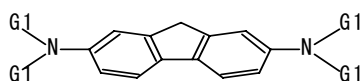
=> D 1 BIB FQHIT

MARPAT ファイルの回答は BIB FQHIT 表示形式などの  
 マルクーシュ構造の表示される表示形式で表示する

L16 ANSWER 1 OF 9 MARPAT COPYRIGHT 2009 ACS on STN DUPLICATE 1  
 AN 117:58899 MARPAT [Full-text](#)  
 TI Electrophotographic photoreceptors using fluorene derivative  
 charge-transporting agent  
 IN Kikuchi, Norihiro; Kanamaru, Tetsuo; Senoo, Akihiro; Tanaka, Takakazu  
 PA Canon K. K., Japan  
 SO Jpn. Kokai Tokkyo Koho, 9 pp.  
 CODEN: JKXXAF  
 DT Patent  
 LA Japanese  
 FAN. CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	JP 04078859	A	19920312	JP 1990-192674	19900720
	JP 2728967	B2	19980318		
PRAI	JP 1990-192674		19900720		

**MSTR 1**



G1 = Ph (opt. substd. by 1 or more G2)  
 Patent location: claims

C 検索テクニック

CASLINK の応用

=> D 6 BIB HITSTR

CAPLUS ファイルの回答は BIB HITSTR 表示形式で表示すると  
ヒットした化合物の構造を経済的に確認できる

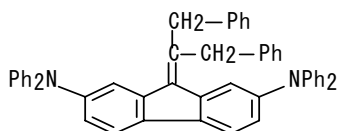
L16 ANSWER 6 OF 9 CAPLUS  
 AN 1991:153904 CAPLUS [Full-text](#)  
 DN 114:153904  
 OREF 114:25857a, 25860a  
 TI **Electrophotographic** photoreceptors using fluorenylidene derivative as charge-transporting agent  
 IN Suzuki, Koichi; Kikuchi, Norihiro; Kanamaru, Tetsuo  
 PA Canon K. K., Japan  
 SO Jpn. Kokai Tokkyo Koho, 12 pp.  
 CODEN: JKXXAF  
 DT **Patent**  
 LA Japanese  
 FAN. CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	JP 02167552	A	19900627	JP 1988-320708	19881221 <--
PRAI	JP 1988-320708		19881221		
IT	<b>132871-15-9</b>	<b>132871-22-8</b>	<b>132871-23-9</b>	<b>132871-24-0</b>	
	<b>132871-29-5</b>	<b>132871-30-8</b>	<b>132871-31-9</b>		

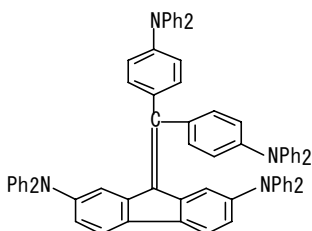
RL: USES (Uses)

(charge-transporting agent, **electrophotog.** photoreceptor using)

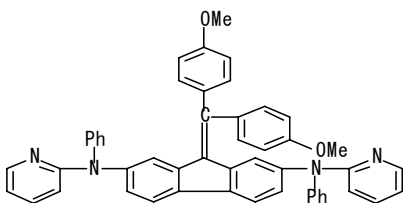
RN 132871-15-9 CAPLUS  
 CN 9H-Fluorene-2, 7-diamine, N2, N2, N7, N7-tetraphenyl-9-[2-phenyl-1-(phenylmethyl)ethylidene]- (CA INDEX NAME)



RN 132871-22-8 CAPLUS  
 CN 9H-Fluorene-2, 7-diamine, 9-[bis[4-(diphenylamino)phenyl]methylene]-N2, N2, N7, N7-tetraphenyl- (CA INDEX NAME)



RN 132871-23-9 CAPLUS  
 CN 9H-Fluorene-2, 7-diamine, 9-[bis(4-methoxyphenyl)methylene]-N2, N7-diphenyl-N2, N7-di-2-pyridinyl- (CA INDEX NAME)



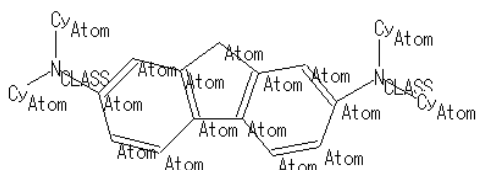
## C 検索テクニック

### CASLINK の応用

#### 応用 2 : サブセット検索

構造検索を実行する前に、キーワードや書誌情報による検索を実行し、その回答を母集合にしたサブセット検索を実行することができる

#### ■ 構造質問式



環は孤立

#### ■ 検索例

=> FILE CASLINK  
:

技術用語や出願日による検索

=> S (ELECTROPHOT? OR ELECTROLUMI?) AND 19900827>=AD AND P/DT

S (ELECTROPHOT? OR ELECTROLUMI?) AND 19900827>=AD AND P/DT FILE=CAPLUS  
L17 32689 FILE CAPLUS

=> S L1 SUB=L17 SAM  
:

あらかじめ作成したキーワードや書誌情報の検索で得られた回答セットを母集合にしたサブセット検索（サンプル検索）

S L1 SSS SAM FILE=REGISTRY  
FILE 'REGISTRY'  
SAMPLE SEARCH INITIATED 17:11:2  
SAMPLE SCREEN SEARCH COMPLETED -

\* この手法は特にイタレーションが INCOMPLETE になるような構造質問式を用いた検索時に特に有効

80.6% PROCESSED 2000 ITERATIONS  
INCOMPLETE SEARCH (SYSTEM LIMIT EXCEEDED)  
SEARCH TIME: 00.00.01

48 ANSWERS

FULL FILE PROJECTIONS: ONLINE \*\*COMPLETE\*\*  
BATCH \*\*COMPLETE\*\*  
PROJECTED ITERATIONS: 46652 TO 52628  
PROJECTED ANSWERS: 728 TO 1654

L21 48 SEA SSS SAM L1

S L21 SUBSET=L20 SSS SAM FILE=MARPAT  
FILE 'MARPAT'  
SAMPLE SUBSET SEARCH INITIATED 17:11:29  
SAMPLE SUBSET SCREEN SEARCH COMPLETED - 27 TO ITERATE

100.0% PROCESSED 27 ITERATIONS

0 ANSWERS

L22 0 SEA SUB=L20 SSS SAM L1

## C 検索テクニック

### CASLINK の応用

=> S L1 SUB=L17 FUL

S L17 AND MARPAT/OS FILE=CAPLUS  
L23 1843 FILE CAPLUS

S L17 AND NONINDEXED/FS FILE=CAPLUS  
L24 181 FILE CAPLUS

S L23 FILE=MARPAT  
L25 1843 FILE MARPAT

S L1 SSS FUL FILE=REGISTRY  
FILE 'REGISTRY'  
FULL SEARCH INITIATED 17:11:36  
FULL SCREEN SEARCH COMPLETED - 47973 TO ITERATE

100.0% PROCESSED 47973 ITERATIONS 958 ANSWERS  
SEARCH TIME: 00.00.03

L26 958 SEA SSS FUL L1

S L26 SUBSET=L25 SSS FUL FILE=MARPAT  
FILE 'MARPAT'  
FULL SUBSET SEARCH INITIATED 17:11:39  
FULL SUBSET SCREEN SEARCH COMPLETED - 633 TO ITERATE

100.0% PROCESSED 633 ITERATIONS 3 ANSWERS  
SEARCH TIME: 00.00.01

L27 3 SEA SUB=L25 SSS FUL L1

S L26 AND L17 FILE=CAPLUS  
L28 7 FILE CAPLUS

DUP REM L27 L28  
PROCESSING COMPLETED FOR L27  
PROCESSING COMPLETED FOR L28  
L29 9 DUP REM L27 L28 (1 DUPLICATE REMOVED)  
ANSWERS '1-3' FROM FILE MARPAT  
ANSWERS '4-9' FROM FILE CAPLUS

フルファイル検索の結果得られる回答数は P.41 と同じく 9 件

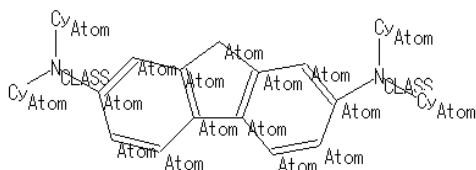
## C 検索テクニック

### CASLINK の応用

#### 応用 3 : 重複文献除去の優先順位

重複文献除去の優先順位を変更したい場合は, CASLINK の途中で作成される L 番号を用いて DUP REM コマンドを実行する

#### ■ 構造質問式



環は孤立

#### ■ 検索例

=> FILE CASLINK

:

=> S L1 FUL

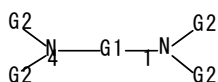
:

L7 381 DUP REM L5 L6 (57 DUPLICATES REMOVED)  
ANSWERS '1-124' FROM FILE MARPAT  
ANSWERS '125-381' FROM FILE CAPLUS

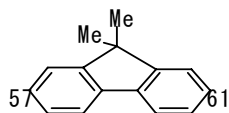
=> D L7 TI FQHIT

L7 ANSWER 1 OF 381 MARPAT COPYRIGHT 2009 ACS on STN **DUPLICATE 1**  
TI Aromatic amine derivatives and organic electroluminescence devices using the same

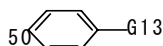
#### MSTR 1



G1 = 57-4 61-1



G2 = 50



Patent location:

claim 1

Note:

substitution is restricted

RE.CNT 10 THERE ARE 10 CITED REFERENCES AVAILABLE FOR THIS RECORD  
ALL CITATIONS AVAILABLE IN THE RE FORMAT

CASLINK では自動的に MARPAT ファイルを優先した重複文献除去が実行される

MARPAT ファイルの構造は複雑で確認に手間取ることがある

C 検索テクニック

CASLINK の応用

=> SET DUP FILE  
SET COMMAND COMPLETED

=> DUP REM L6 L5  
PROCESSING COMPLETED FOR L6  
PROCESSING COMPLETED FOR L5  
L8 381 DUP REM L6 L5 (57  
ANSWERS '1-314' F  
ANSWERS '315-381' FROM FILE MARPAT

重複除去時の優先順位を変更するには、  
(1) あらかじめ SET DUP FILE を入力  
(2) DUP REM コマンドを用いて重複除去を実行

L5 : MARPAT ファイルのフルファイル検索結果  
L6 : REGISTRY → CAplus ファイルの検索結果  
(P.40 参照)

=> D L8 BIB HITSTR

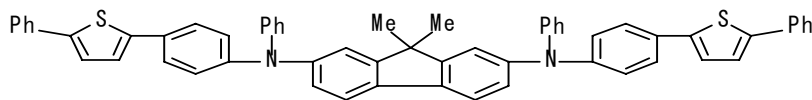
L8 ANSWER 1 OF 381 CAPLUS COPYRIGHT 2009 ACS on STN DUPLICATE 1  
AN 2008:255678 CAPLUS [Full-text](#)  
DN 148:275087  
TI Aromatic amine derivatives and organic electroluminescence devices using  
the same  
TIJP 芳香族アミン誘導体とそれを用いる有機エレクトロ・ルミネッセンス装置  
[機械翻訳]  
IN Yabunouchi, Nobuhiro; Kawamura, Masahiro; Kawamura, Hisayuki; Hosokawa,  
Chishio; Matsunami, Shigeyuki; Yoshinaga, Tadahiko; Kijima, Yasunori  
PA Idemitsu Kosan Co., Ltd., Japan; Sony Corporation  
SO PCT Int. Appl., 77pp.  
CODEN: PIXXD2  
DT Patent  
LA Japanese  
FAN. CNT 3

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 2008023759	A1	20080228	WO 2007-JP66356	20070823
	:				
	US 20080108839	A1	20080508	US 2007-834157	20070806
	US 20080106190	A1	20080508	US 2007-834203	20070806
	US 20080108832	A1	20080508		
	EP 2055700	A1	20080508		
	:				
	IN 2009CN00955	A	20090313		
	KR 2009036139	A	20090313		
PRAI	JP 2006-226121	A	20061221		
	WO 2007-JP66356	W	20070823		
OS	MARPAT 148:275087				
IT	<b>1007387-67-8P</b>				

前ページと同じレコード  
重複除去時の優先順位を変更したため、CAplus ファイルが  
優先されるようになった

CAplus ファイルでは、化学物質の構造を HITSTR 表示  
形式で表示できるので構造の確認に便利

RL: IMF (Industrial manufacture); TEM (Technical or engineered material  
use); PREP (Preparation); USES (Uses)  
(formation of aromatic amine derivs. for organic electroluminescence devices)  
RN 1007387-67-8 CAPLUS  
CN 9H-Fluorene-2,7-diamine, 9,9-dimethyl-N2,N7-diphenyl-N2,N7-bis[4-(5-phenyl-  
2-thienyl)phenyl]- (CA INDEX NAME)



RE. CNT 10 THERE ARE 10 CITED REFERENCES AVAILABLE FOR THIS RECORD  
ALL CITATIONS AVAILABLE IN THE RE FORMAT

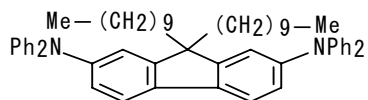
## C 検索テクニック

### CASLINK の応用

=> D 200 BIB HITSTR

L8 ANSWER 200 OF 381 CAPLUS COPYRIGHT 2009 ACS on STN  
AN 2004:670129 CAPLUS [Full-text](#)  
DN 141:303847  
TI Two-photons and beyond: 2, 3 and 4 photon absorption in conjugated fluorenes  
TIJP : 2を超えた共役なフルオレンにおける二光子, 3, および4光子吸収 [機械翻訳]  
AU Belfield, Kevin D.; Hernandez, Florencio E.; Cohanoschi, Ion; Bondar, Mykhailo V.;  
Van Stryland, Eric W.  
CS Department of Chemistry and School of Optics: CREOL and FPCE, University of Central  
Florida, Orlando, FL, 32816-2366, USA  
SO PMSE Preprints (2004), 91, 346-347  
CODEN: PPMRA9; ISSN: 1550-6703  
PB American Chemical Society  
DT Journal; (computer optical disk) ●  
LA English  
IT **289892-10-0**, 9,9-Didecyl-2,7-bis(N,N-diphenylamino) fluorene  
RL: PRP (Properties)  
(multi-photon absorption in conjugated fluorenes)  
RN 289892-10-0 CAPLUS  
CN 9H-Fluorene-2,7-diamine, 9,9-didecyl-N2,N2,N7,N7-tetraphenyl- (CA INDEX NAME)

MARPAT ファイルではヒットしなかった回答  
(CAplus ファイルの非特許文献レコード)



OSC.G 1 THERE ARE 1 CAPLUS RECORDS THAT CITE THIS RECORD (1 CITINGS)  
RE.CNT 4 THERE ARE 4 CITED REFERENCES AVAILABLE FOR THIS RECORD  
ALL CITATIONS AVAILABLE IN THE RE FORMAT

C 検索テクニック

CASLINK の応用

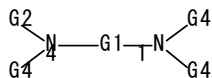
=> D 315 BIB FQHIT

315 番目以降は MARPAT ファイルのみでヒットした回答  
(CAplus ファイルではヒットしなかった回答)

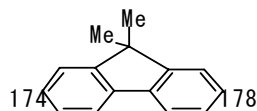
L8 ANSWER 315 OF 381 MARPAT COPYRIGHT 2009 ACS on STN  
 AN 148:297671 MARPAT [Full-text](#)  
 TI Organic electroluminescent device  
 IN Jinde, Yukitoshi; Yabunouchi, Nobuhiro; Kuma, Hitoshi; Hosokawa, Chishio  
 PA Idemitsu Kosan Co., Ltd., Japan  
 SO PCT Int. Appl., 85pp.  
 CODEN: PIXXD2  
 DT Patent  
 LA Japanese  
 FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 2008023623	A1	20080228	WO 2007-JP65916	20070815
	W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW RW: AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG, BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM US 20080049413 A1 20080228 US 2006-521317 20060915 EP 2061103 A1 20090520 EP 2007-792550 20070815 R: AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LI, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR, AL, BA, HR, MK, RS KR 2009042272 A 20090429 KR 2009-703517 20090220 PRAI JP 2006-224958 20060822 US 2006-521317 20060915 WO 2007-JP65916 20070815				

**MSTR 1**



G1 = 174-4 178-1



G2 = biphenyl  
 G4 = Ph

Patent location: claim 1  
 Note: substitution is restricted

RE.CNT 14 THERE ARE 14 CITED REFERENCES AVAILABLE FOR THIS RECORD  
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

## C 検索テクニック

参考: 重複文献除去

- DUPLICATE コマンドを用いれば、重複文献レコードに同じ DUPLICATE 番号を付与して 3 種類の処理 (除去・識別・検出) を実行することができる。

=> DUP\*<sup>1</sup> [オプション] [回答セットの L 番号\*<sup>2</sup>]

\*<sup>1</sup> マルチファイル検索の回答セットのほか、単一ファイルの回答セットも指定できる。  
また、回答セットも複数個指定することができる

\*<sup>2</sup> DUPLICATE が実行できるファイルは => HELP DUFIL で確認できる。

- ・ 処理内容はオプションで指定する。

オプション	機能	内容
<u>REMOVE</u>	重複文献除去	優先ファイルのレコードを残して重複文献除去した回答セットを作成する
<u>IDENTIFY</u>	重複文献識別	重複文献を識別し、重複文献レコードも含めた回答セットを作成する。重複文献除去はしない。
<u>ONLY</u>	重複文献検出	重複文献レコードのみの回答セットを作成する

- 優先的に回答を残すファイルを指定して重複文献除去する方法

- ・ マルチファイル検索の回答セットを重複文献除去する場合、除去されずに残る回答レコードの優先順位は FILE コマンドでマルチファイル環境に入るファイルの順序に相当する。優先的に回答を残したいファイルがある場合は、その順序でマルチファイル環境に入る。

=> FIL [優先的に回答を残したいファイルの順にファイル名を指定\*]

:

=> DUP REM [マルチファイル検索の回答セットの L 番号]

- ・ 単一ファイルで別々に検索した複数の回答セットに対して重複文献除去することもできる。  
その場合は、優先的に残したいファイル順に回答セットの L 番号を入力する。

=> DUP REM [優先的に回答を残したいファイルの順に回答セットの L 番号を指定]

- SET DUPORDER FILE によって、重複文献除去 (DUPLICATE REMOVE) または重複文献識別 (DUPLICATE IDENTIFY) 後の回答レコードをファイルごとにまとめることができる。

- ・ この設定変更をしない場合 (SET DUPORDER DEFAULT)、回答レコードはファイルの区別なく新しい順に並ぶ。

## C 検索テクニック

参考: 重複文献除去

### ■ STN の重複文献識別のしくみ

- STN では重複文献レコードを以下の検出キーで識別している。

レコードのタイプ	検出キー
非特許文献	発行年, 記事・資料種類, 開始ページ, 巻, 号, ISSN, CODEN, 著者, 標題* <sup>1</sup> など
特許	特許発行年, 特許発行国, 特許番号* <sup>2</sup>

\*<sup>1</sup> 補助的に利用

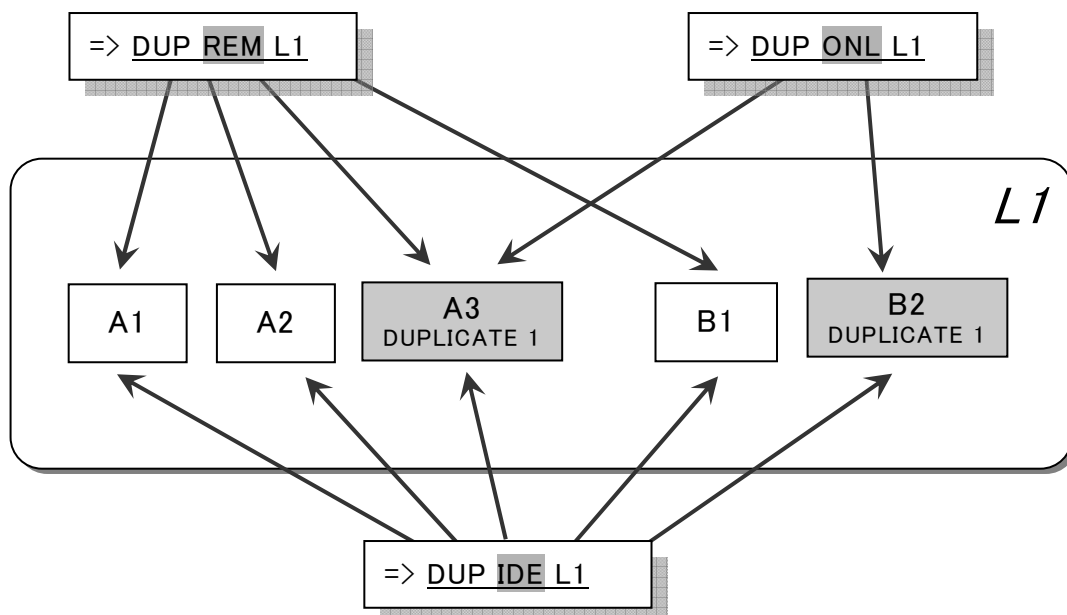
\*<sup>2</sup> CPlus/CA ファイルや WPI ではベーシック特許の特許番号のみ使用

- 雑誌などの非特許文献では、標題は検出キーとして補助的に利用されるのみである。このため翻訳などによって標題の異なる文献でも、重複文献除去することができる。
- 特許の検出キーには標題は含まれていない。このため、標題の翻訳、情報付加、書き換えなどは一切影響しない。
- CPlus/CA ファイルや WPI などの特許レコードでは、ベーシック特許の特許番号を検出キーに用いている。対応特許の特許番号は照合しないため、同一発明の特許レコードが除去されない場合がある。

### ■ DUPLICATE コマンドのしくみ

L1 : マルチファイル検索の回答セット  
 A1, A2, A3 : A ファイルのレコード  
 B1, B2 : B ファイルのレコード

A3 と B2 は重複文献  
 → 新たに作成される回答セットに含めるレコード



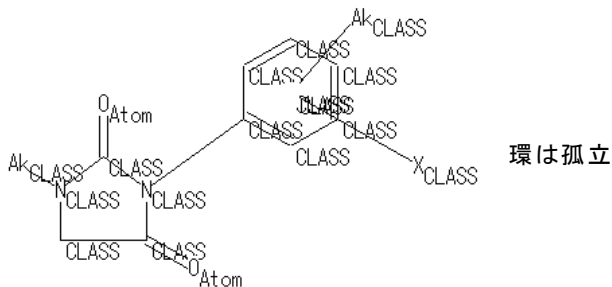
## C 検索テクニック

### CASLINK の応用

#### 応用 4 : INCOMPLETE の回答が得られた場合

CASLINK の実行途中で構造検索結果に INCOMPLETE の回答が含まれた場合は、該当ファイルに入って、/COM の検索を行った後に、DUP REM コマンドを実行する

#### ■ 構造質問式



#### ■ 検索例

=> FILE CASLINK

:

=> S L1

:

=> S L1 FUL

S L1 SSS FUL FILE=REGISTRY

FILE 'REGISTRY'

FULL SEARCH INITIATED 18:03:56

FULL SCREEN SEARCH COMPLETED - 102338 TO ITERATE

100.0% PROCESSED 102338 ITERATIONS ( 2 INCOMPLETE) 280 ANSWERS

SEARCH TIME: 00.00.03

L4 280 SEA SSS FUL L1

S L4 SSS FUL FILE=MARPAT

FILE 'MARPAT'

FULL SEARCH INITIATED 18:04:00

FULL SCREEN SEARCH COMPLETED - 66132 TO ITERATE

100.0% PROCESSED 66132 ITERATIONS

SEARCH TIME: 00.00.37

L5 2122 SEA SSS FUL L1

S L4 FILE=CAPLUS

L6 30 FILE CAPLUS

DUP REM L5 L6

L7 2140 DUP REM L5 L6 (12 DUPLICATES REMOVED)

ANSWERS '1-2122' FROM FILE MARPAT

ANSWERS '2123-2140' FROM FILE CAPLUS

REGISTRY ファイルの回答中に 2 件 INCOMPLETE の回答が含まれていたが CASLINK ではイタレーションが COMPLETE した回答に限定せずに自動的に重複文献除去まで実行される

## C 検索テクニック

### CASLINK の応用

=> D HIS

```
FILE 'CAPLUS, MARPAT, REGISTRY' ENTERED AT 14:33:09 ON 29 JUL 2009
L1          STRUCTURE UPLOADED
L2          4 FILE REGISTRY
L3          50 FILE MARPAT
L4          280 FILE REGISTRY
L5          2122 FILE MARPAT
L6          30 FILE CAPLUS
L7          2140 DUP REM L5 L6 (12 DUPLICATES REMOVED)
```

=> FILE REGISTRY

=> S L4/COM  
L8 278 L4/COM

イタレーションが INCOMPLETE になったファイルに入り /COM を用いてイタレーションが完了した回答に限定する

=> FILE CAPLUS

=> S L8  
L9 29 L8

イタレーションが完了した回答に限定した REGISTRY ファイルの回答セットを CAPLUS ファイルにクロスオーバーする

=> SET DUP FILE  
SET COMMAND COMPLETED

重複文献除去を実行する

=> DUP REM L9 L5

```
FILE 'CAPLUS' ENTERED AT 18:08:50 ON 04 AUG 2009
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2009 AMERICAN CHEMICAL SOCIETY (ACS)
```

```
FILE 'MARPAT' ENTERED AT 18:08:50 ON 04 AUG 2009
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2009 American Chemical Society (ACS)
PROCESSING COMPLETED FOR L9
PROCESSING COMPLETED FOR L5
L10         2139 DUP REM L9 L5 (12 DUPLICATES REMOVED)
           ANSWERS '1-29' FROM FILE CAPLUS
           ANSWERS '30-2139' FROM FILE MARPAT
```

## C 検索テクニック

### CASLINK の応用

#### 応用 5 : 保存

個別のファイルで保存すれば、後日呼び出した回答でも複雑な条件を組み合わせた検索が実行できるので、さらに活用度が増す

■ CASLINK で得られた回答を保存する場合は、

- ・ CASLINK 環境で保存する
- ・ CAplus, MARPAT ファイルの各回答を、それぞれのファイルで保存する

の 2 通りの手法がある。

- ・ CASLINK 環境で保存する場合

- CASLINK で得られた重複除去後の回答セットを保存した場合は、後日、呼び出して使用する際に、絞りこみ検索などに使用できない。
- 回答表示のみ後日実行したい場合などに使用する。

=> FILE CASLINK

=>

```
Uploading C:\My_Document\STN Express 8.4\Queries\CASLINK1.str
L1      STRUCTURE UPLOADED
```

:

=> S L1 FUL

:

```
L7      381 DUP REM L5 L6 (57 DUPLICATES REMOVED)
        ANSWERS '1-124' FROM FILE MARPAT
        ANSWERS '125-381' FROM FILE CAPLUS
```

=> SAVE L7 TEMP CASLINK/A

```
ANSWER SET L7 HAS BEEN SAVED AS 'CASLINK/A'
```

---

=> FILE CASLINK

=> ACT CASLINK/A

```
L1      STR
L2 (    958)SEA FILE=REGISTRY SSS FUL L1
L3 (    124)SEA FILE=MARPAT SSS FUL L1
L4 (    314)SEA FILE=CAPLUS SPE=ON  ABB=ON  PLU=ON  L2
L5      381 DUP REM L3 L4 (57 DUPLICATES REMOVED)
```

=> S L5 AND ?LUMINE? ●

S L4 AND ?LUMINE? FILE=CAPLUS

COMBINATION OF STRUCTURE AND TEXT TERM

COMMAND STACK INTERRUPTED. ENTER "DISPLAY HISTORY"

TO SEE WHICH COMMANDS WERE EXECUTED.

:

CASLINK で得られた重複除去後の回答セットを保存した場合は、後日、呼び出して使用する際に、絞りこみ検索などに使用できない

## C 検索テクニック

### CASLINK の応用

- CAplus, MARPAT ファイルの各回答を、それぞれのファイルで保存する場合

=> FIL CASLINK

=>

Uploading C:\My\_Document\STN Express 8.4\Queries\CASLINK1.str

L1 STRUCTURE UPLOADED

:

=> S L1 FUL

:

L7 381 DUP REM L5 L6 (57 DUPLICATES REMOVED)

ANSWERS '1-124' FROM FILE MARPAT

ANSWERS '125-381' FROM FILE CAPLUS

=> FILE MARPAT

:

=> SAV TEMP L5 MARPAT/A

ANSWER SET L5 HAS BEEN SAVED AS 'MARPAT/A'

=> FILE CAPLUS

:

=> SAV TEMP L6 REGCA/A

ANSWER SET L5 HAS BEEN SAVED AS 'REGCA/A'

MARPAT ファイルのフルファイル検索  
結果 (L5) を保存する

CAplus ファイルのフルファイル検索  
結果 (L6) を保存する

---

=> FILE MARPAT

:

=> ACT MARPAT/A

L1 STR

L2 139 SEA FILE=MARPAT SSS FUL L1

=> FILE CAPLUS

:

=> ACT REGCA/A

L3 STR

L4 ( 5)SEA FILE=REGISTRY SSS FUL L3

L5 4 SEA FILE=CAPLUS ABB=ON PLU=ON L4

=> S L2 OR L5

139 L2

L6 142 L2 OR L5

=> S L6 AND ?LUMINE?

373835 ?LUMINE?

L7 1 L6 AND ?LUMINE?

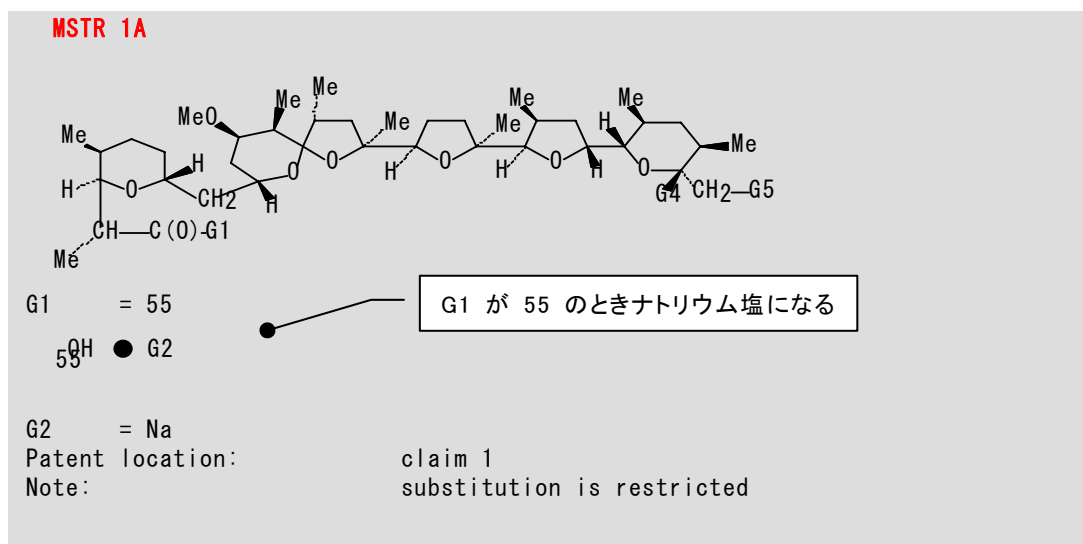
● 後日、呼び出してから、複雑な検索も実行できる

## C 検索テクニック

### 多成分物質

- MARPAT ファイルにも塩や対イオンのような多成分物質が収録されている。

- ・ レコード例



- ・ 塩を検索する場合は以下の手法を考慮する。

#### 方法 1 : 特定の塩の検索

同一構造作図画面中に複数の構造を作図して検索する  
(MARPAT ファイルでは複数の構造質問式をブール演算で結合することはできない)

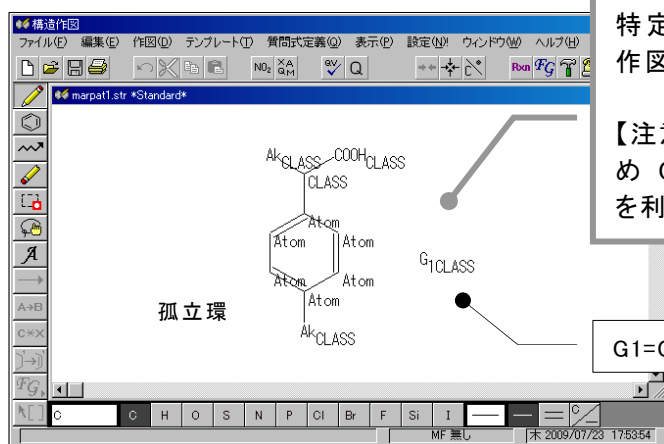
#### 方法 2 : 一般的な塩の検索 (注記フィールド検索)

レコードによっては、塩については注記フィールドに記載されるだけで、構造は遊離の酸の構造のみしか収録されていない場合があるため、注記フィールド中に salt と記載されているレコードを検索する

## C 検索テクニック

### 多成分物質

#### ■ 検索例 1 : 特定の塩の検索



特定の多成分物質を検索したい場合は同一作図画面中に必要な成分を作図して検索する

【注意】REGISTRY での塩の検索とは異なるため CASLINK や REGISTRY からの連続検索を利用する場合はこの作図は利用しない

G1=Ca/K/Na

=> FILE MARPAT

=> S L1 FUL

L3 459 SEA SSS FUL L1

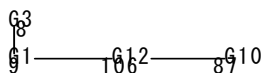
=> S L3/COM

L4 455 L3/COM

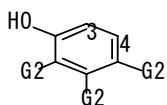
=> D FQHIT 3

L4 ANSWER 3 OF 455 MARPAT COPYRIGHT 2009 ACS on STN

#### MSTR 1A



G1 = 3-8 4-106



G2 = alkyl <containing 1-22 C>

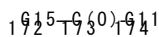
G4 = alkali metal atom

G10 = 104



G11 = 0

G12 = 172-9 174-87



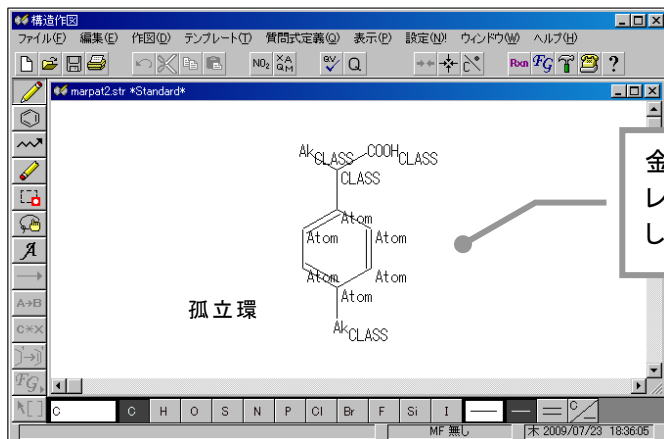
G15 = alkylene <containing 1-4 C>

Patent location: claim 1

## C 検索テクニック

### 多成分物質

#### ■ 検索例 2 : 一般的な塩の検索 (注記フィールド検索)



金属が特定されていない一般的な塩のレコードを検索する場合は酸のみ作図して検索する

=> FILE MARPAT

Uploading C:\My\_Document\STN Express 8.4\Queries\marpat2.str  
L5 STRUCTURE UPLOADED

=> S L5 FUL

L7 6340 SEA SSS FUL L5

=> S L7/COM

L8 6337 L7/COM

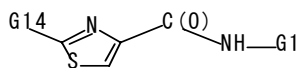
=> S L8 AND SALT?

L9 71016 SALT?  
3576 L8 AND SALT?

=> D FQHIT 2 3 7

L9 ANSWER 2 OF 3576 MARPAT COPYRIGHT

#### MSTR 1



G14 = Ph (opt. substd. by (1-2) G15)

G15 = alkenyl <containing 2-6 C>  
(opt. substd. by 1 or more G22)

G22 = CO2H

Patent location:

Note:

Note:

Note:

Stereochemistry:

claim 1

substitution is restricted

additional oxo formation also disclosed

and salts

and enantiomers and diastereomers

塩に関する情報は構造で表現できない場合もある

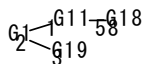
例えば "pharmaceutically acceptable salt" のような表現が特許中に出現した場合は、注記フィールド中にそのテキスト情報のみ記述される

C 検索テクニック

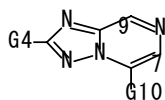
多成分物質

L9 ANSWER 3 OF 3576 MARPAT COPYRIGHT 2009 ACS on STN

**MSTR 1**



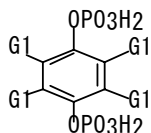
G1 = 9-1 7-3



G4 = alkyl <containing 1-4 C>  
 (opt. substd. by 1 or more G28)  
 G28 = CO<sub>2</sub>H / Ph (opt. substd. by (1-3) G29)  
 G29 = alkyl <containing 1-4 C>  
 Patent location: claim 1  
 Note: or salts, solvates, or solvates of salts  
 Note: additional ring formation also claimed

L9 ANSWER 7 OF 3576 MARPAT COPYRIGHT 2009 ACS on STN

**MSTR 1**



G1 = alkyl <containing 1-6 C> (opt. substd. by G2)  
 G2 = CO<sub>2</sub>H  
 Patent location: claim 1  
 Note: or cosmetically acceptable salts, solvates, or hydrates  
 Stereochemistry: and optical or geometric isomers

## C 検索テクニック

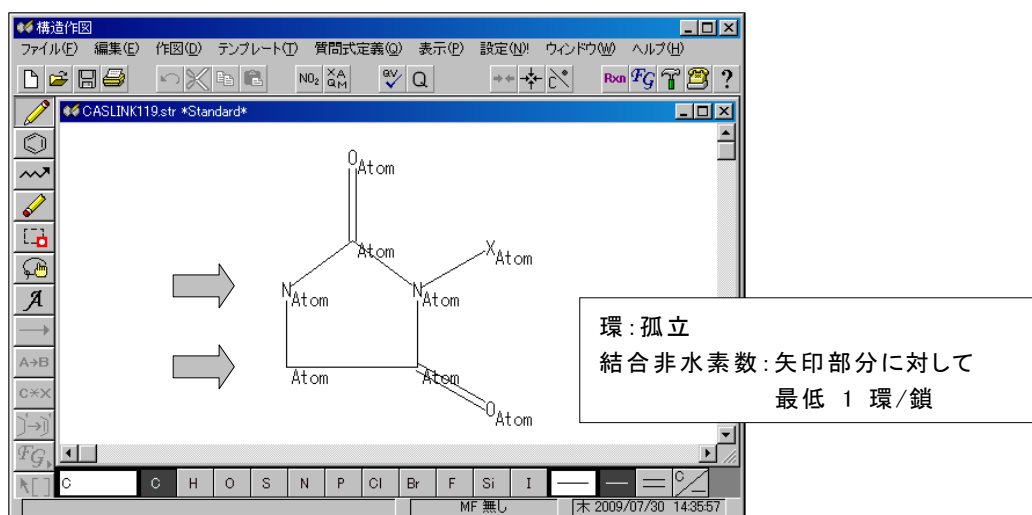
### CSS 検索

#### ■ MARPAT ファイルにおける CSS 検索

- ・ 置換は許容しない
  - 構造質問式中のノードに結合非水素数を指定しない限り、置換は許容しない。
- ・ 置換を許容したい部位には指定が必要
  - 置換を許容したいノードには、必ず結合非水素数で適切な指定をする。
- ・ 環の孤立化
  - 構造質問式中の環は自動的にすべて「孤立」の指定になる。
- ・ Ak は炭化水素になる
  - 置換を許容したい場合は、結合非水素数で適切な指定をする。

#### ■ MARPAT ファイルにおける CSS 検索利用上の注意点

- ・ MARPAT ファイルでは「環の孤立」よりも、結合非水素数の指定が優先されるため孤立環の指定をした環上のノードに結合非水素数で結合の存在を指定して CSS 検索を実行すると、環の孤立は無視される。
  - 例えば、下図のように作図すると . . .



#### SSS, CSS 検索と属性（環の孤立および結合非水素数）の関係

- ・ SSS 検索の場合 : 環の孤立 > 結合非水素数
- ・ CSS 検索の場合 : 環の孤立 < 結合非水素数

## C 検索テクニック

### CSS 検索

#### ■ 検索例

=> FILE MARPAT  
:

=> S L1  
L2 3 SEA SSS SAM L1

=> S L1 FUL  
L3 91 SEA SSS FUL L1

=> S L1 CSS  
L4 3 SEA CSS SAM L1

=> S L4 CSS FUL  
L5 94 SEA CSS FUL L1

=> S L5 NOT L3  
L6 3 L5 NOT L3

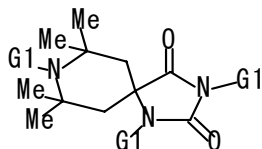
=> D FQHIT 2-3

L6 ANSWER 2 OF 3 MARPAT COPYRIGHT 2009 ACS on STN

SSS 検索を実行すると 91 件の回答セットが得られる

MARPAT ファイルでは「環の孤立」よりも、結合非水素数の指定が優先されるため CSS 検索を実行すると 94 件の回答セットが得られる

#### MSTR 1

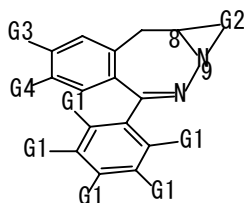


G1 = Cl  
Patent location: claim 1

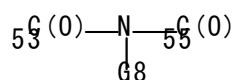
CSS 検索では、環の孤立が無視され、結合非水素数の指定が有効になるため、縮合環、スピロ環の回答が得られる

L6 ANSWER 3 OF 3 MARPAT COPYRIGHT 2009 ACS on STN

#### MSTR 1B



G2 = 53-8 55-9



G8 = F  
Derivative: and isomers and physiologically useful salts  
Patent location: claim 1

## C 検索テクニック

### SYSTEM LIMIT (30 分) を超えた場合

- MARPAT ファイルではイタレーション検索のシステム制限以外に、検索時間の制限があるため、30 分以内に検索が終了しないと、SYSTEM LIMIT EXCEEDED のメッセージが表示され、検索は中断される。

```
=> FILE MARPAT
:

=> S L1 FUL
FULL SEARCH INITIATED 16:33:27
FULL SCREEN SEARCH COMPLETED -
:

INCOMPLETE SEARCH (SYSTEM LIMIT EXCEEDED)
SEARCH TIME: 00.30.05

FULL FILE PROJECTIONS:  ONLINE  **COMPLETE**
                        BATCH   **COMPLETE**
PROJECTED ITERATIONS:   197862 TO 205338
PROJECTED ANSWERS:      18 TO    132

L3          18 SEA SSS FUL L1
```

SYSTEM LIMIT EXCEEDED のメッセージが表示され、検索が中断される

- ・ フルファイル検索が中断した場合の対処

- システムは新しいレコードから古いレコードの順に検索を実行するため、(中断して)得られた回答セット中の一番古い回答のレコード番号を確認し、それ以前のレコードを対象にした RANGE 検索を実行する

- ・ RANGE 検索 (範囲指定検索)

- MARPAT ファイルでは、レコード番号、CA の巻・号、西暦 (AN の年) を指定した RANGE 検索が実行できる。
- RANGE 検索はフルファイル検索のみ。サンプル検索は指定できない。したがって、検索実行と同時にフルファイル検索料金がかかるので、よく確認した上で実行する。

- 入力例

```
=> S L1 RAN=(, 111:12345) ← レコード番号 111:12345 以前のレコードを検索する
=> S L1 RAN=(111:12345,) ← 111:12345 以後のレコードを検索する
=> S L1 RAN=(, V111) ← CA の 111 巻以前のレコードを検索する
=> S L1 RAN=(, 1992) ← 1992 年以前のレコードを検索する
=> S L1 RAN=(1992, 1998) ← 1992-1998 年のレコードを検索する
```

## C 検索テクニック

SYSTEM LIMIT (30 分) を超えた場合

### ・ RANGE 検索例

```

=> FILE MARPAT
:

=> S L1 FUL
FULL SEARCH INITIATED 16:33:27
FULL SCREEN SEARCH COMPLETED -
:

```

SYSTEM LIMIT EXCEEDED のメッセージが表示され、  
検索が中断される

```

INCOMPLETE SEARCH (SYSTEM LIMIT EXCEEDED)
SEARCH TIME: 00.30.05

FULL FILE PROJECTIONS:  ONLINE **COMPLETE**
                        BATCH  **COMPLETE**
PROJECTED ITERATIONS:   197862 TO 205338
PROJECTED ANSWERS:     18 TO 132

L3          18 SEA SSS FUL L1

=> D AN 18

```

● システムは新しいレコードから順に検索するため  
一番古い回答のレコード番号を確認

```

L3 ANSWER 18 OF 18 MARPAT CO
AN 114:163736 MARPAT

=> S L1 RAN=(,114:163736)
RANGE SEARCH INITIATED 16:42:16
RANGE SCREEN SEARCH COMPLETED -

```

● 確認した古いレコード番号以前 (古い) のレコードを  
対象に RANGE 検索する

```

L4          1 SEA RAN=(,114:163736) SSS L1

=> S L3 OR L4

```

● 回答をまとめる

参考 : 上記のように検索すると 2 回分の構造検索料がかかる。

参考 : MARPAT のシステム制限値

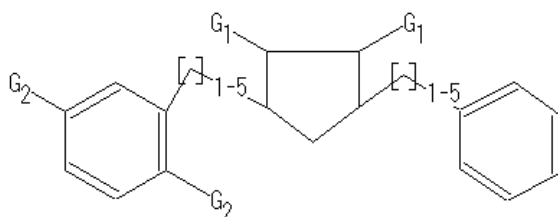
Search Scope	Iterations	Answers	Minutes
Online SAMPLE	2,000	50	5
Subset SAMPLE	2,000	50	5
Online FULL	200,000	200,000	30
Subset FULL	200,000	200,000	30
BATCH FULL	200,000	200,000	360
Online RANGE	200,000	200,000	30
Subset RANGE	200,000	200,000	30
BATCH RANGE	200,000	200,000	360

## C 検索テクニック

### STRUCTURE TOO LARGE

- 構造検索を実行しようとするとき「STRUCTURE TOO LARGE」の警告が表示され、検索できない場合がある。

構造質問式：



G1=C/S/N/P/Cl/Br/F/I  
G2=OH/SH/MeO/EtO/n-PrO/l-PrO/n-BuO/t-BuO/s-BuO

=> FILE MARPAT

=>

Uploading C:\My\_Document\STN Express 8.4\Queries\sample.str

L1 STRUCTURE UPLOADED

=> S L1

STRUCTURE TOO LARGE - SEARCH ENDED ●

「個々の G グループの中味が多すぎる」  
「G グループの数が多すぎる」などの理由で  
構造検索を実行することができない場合  
がある

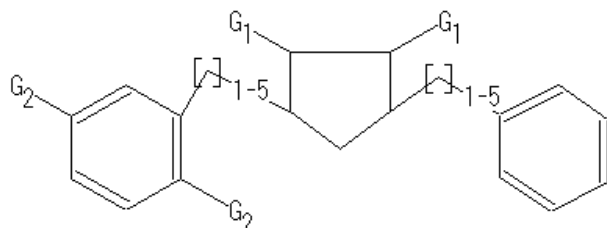
- ・ STRUCTURE TOO LARGE になった場合の対処
  - G グループの代わりに、その定義を含むより広義の一般式グループ記号を用いる。
  - 定義の広い一般式グループ記号の代わりに、可変原子を用いる。
  - G グループ中の特定の環を含めている場合、それらを Hy, Cb, Cy (属性, 元素数も指定) に置き換える。
  - 炭素 (C) の繰り返しグループの指定を Ak に置き換える。
  - 大きな繰り返しグループは繰り返しの指定をやめ、他の構造と結合させずにフラグメントのみ作図する。

## C 検索テクニック

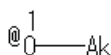
### STRUCTURE TOO LARGE

#### 再検索

構造質問式：



G1=C/X/Q  
G2=OH/SH/@1



G1=C/S/N/P/Cl/Br/F/I を指定していたが ...

Cl/Br/F/I → X (ハロゲン)  
S/N/P → Q (C, H 以外のノード)

に変更

G2=OH/SH/MeO/EtO/n-PrO/l-PrO/n-BuO/t-BuO/s-BuO を  
指定していたが ...

MeO/EtO/n-PrO/l-PrO/n-BuO/t-BuO/s-BuO → O-Ak

に変更

=> FILE MARPAT

=>

Uploading C:\My\_Document\STN Express 8.4

L1 STRUCTURE UPLOADED

=> S L1

SAMPLE SEARCH INITIATED 15:59:48 FILE 'MARPAT'

SAMPLE SCREEN SEARCH COMPLETED - 9556 TO ITERATE

20.9% PROCESSED 2000 ITERATIONS  
INCOMPLETE SEARCH (SYSTEM LIMIT EXCEEDED)  
SEARCH TIME: 00.00.18

0 ANSWERS

FULL FILE PROJECTIONS: ONLINE \*\*COMPLETE\*\*  
BATCH \*\*COMPLETE\*\*  
PROJECTED ITERATIONS: 187331 TO 194909  
PROJECTED ANSWERS: 0 TO 0

L2 0 SEA SSS SAM L1

=> S L1 FUL

FULL SEARCH INITIATED 16:00:29 FILE 'MARPAT'

FULL SCREEN SEARCH COMPLETED - 192814 TO ITERATE

L3 16 SEA SSS FUL L1

参考：一般的なシステム制限値回避のための対処

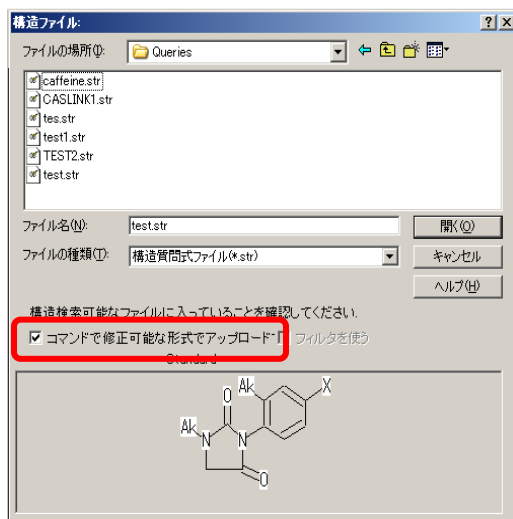
- 結合をより限定的な定義に変更する (例:「エグザクト/ノーマライズド」を「エグザクト」に変更する)
- マッチレベルをより限定的な定義に変更する (例:「クラス」を「原子」に変更する)
- 環を孤立化する
- 一般式属性や元素数を指定する

## C 検索テクニック

参考 : アップロードした構造質問式中のマッチレベルの確認

### ■ 参考 : アップロードした構造質問式中のマッチレベルの確認

- ① STN Express の場合は、アップロードする構造質問式のファイルを選択する画面(下図)で「コマンドで修正可能な形式でアップロード」にチェックを付けて「開く」をクリックする。



\* STN on the Web の場合は、特に指定しなくても「コマンドで修正可能な形式でアップロード」にチェックを付けたのと同じ設定になっている。

- ② アップロードした構造質問式を => D QUE コマンドで表示すると、属性が表示され、マッチレベルや元素数レベルも確認することができる。

NODE ATTRIBUTES:			
NSPEC	IS R	AT	1
NSPEC	IS R	AT	2
NSPEC	IS R	AT	3
NSPEC	IS R	AT	4
NSPEC	IS R	AT	5
NSPEC	IS C	AT	6
NSPEC	IS C	AT	7
NSPEC	IS R	AT	8
NSPEC	IS R	AT	9
NSPEC	IS R	AT	10
NSPEC	IS R	AT	11
NSPEC	IS R	AT	12
NSPEC	IS R	AT	13
NSPEC	IS C	AT	14
NSPEC	IS C	AT	16

DEFAULT MLEVEL IS ATOM

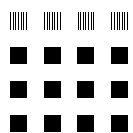
→ デフォルトのマッチレベルは「原子」  
(1 2 3 4 5 7 14 の各ノード)

MLEVEL IS CLASS AT 6 8 9 10 11 12 13 15 16

→ 6 8 9 10 11 12 13 15 16 の各ノードのマッチレベルは「クラス」

DEFAULT ECLEVEL IS LIMITED

→ デフォルトの元素数レベルは「限定」



# **JAICI** 社団法人 化学情報協会

## 情報事業部

〒113-0021 東京都文京区本駒込6-25-4 中居ビル

サービス全般 TEL: 0120-151-462

E-mail: [customer@jaici.or.jp](mailto:customer@jaici.or.jp)

ヘルプデスク TEL: 0120-003-462

E-mail: [support@jaici.or.jp](mailto:support@jaici.or.jp)

FAX: 03-5978-3600 URL: [www.jaici.or.jp](http://www.jaici.or.jp)