

Markush searches in PATENTSCOPE

Global Databases Division
June 2023



Agenda

- Definition of a Markush formula
- Challenges of Markush searches
- Markush implementation
- Markush data information and coverage
- Markush search functionality
- Case studies
- Q/A session

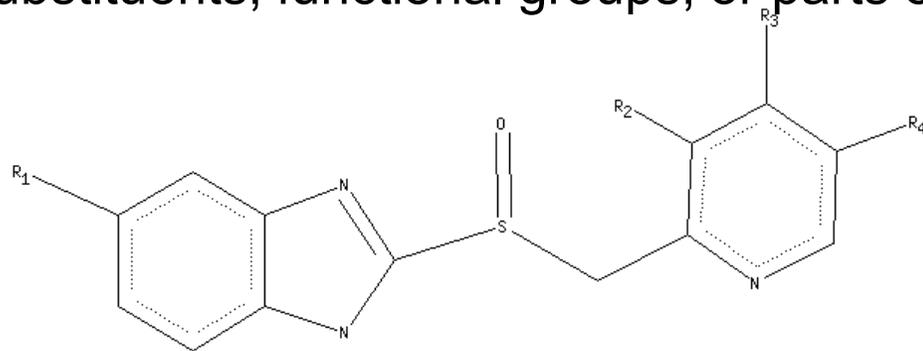
Definition of a Markush formula

- Markush formulae are named after Eugene A. Markush, founder of the Pharma Chemical Corporation in New Jersey.

He was involved in a legal case that set a precedent for generic chemical structure patent filing. The patent filing was US Application 611,637, filed January 9, 1923. Markush was awarded a patent from the US Patent Office for "*Pyrazolone Dye and Process of Making the Same*" on August 26, 1924 (Wikipedia)

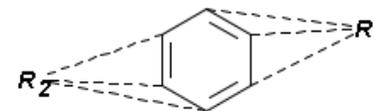
- Representation of families of chemical compounds with «varying parts», usually with the purpose to specify patents claims

A Markush formula uses placeholders, such as R, X, Y, etc., to represent variations of substituents, functional groups, or parts of the structure.

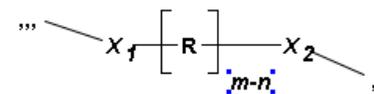


Substitution variation: R₁ = H, CH₃, OH ...

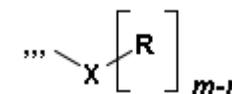
Position variation: Variable points of attachment



Frequency variation: Repeating groups



Multipliers



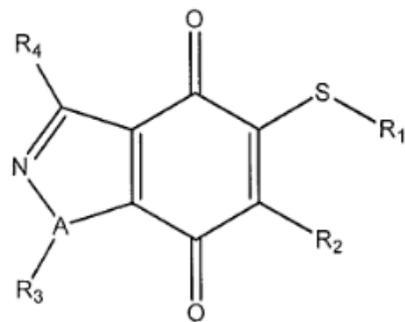
Homology variation: Generic nodes

CHK, CYC, ARY, ...

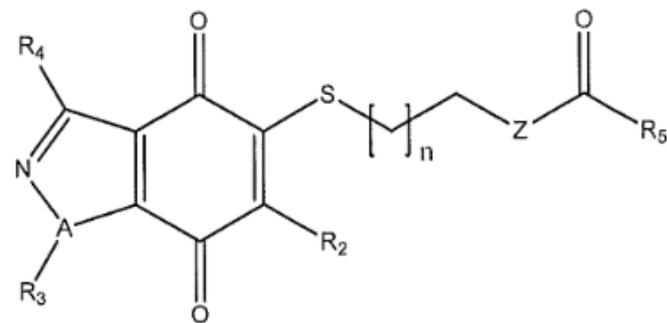
Challenges of Markush searches

- Ambiguity
- Availability
- Complexity
- Size & Scope
- Standard representation

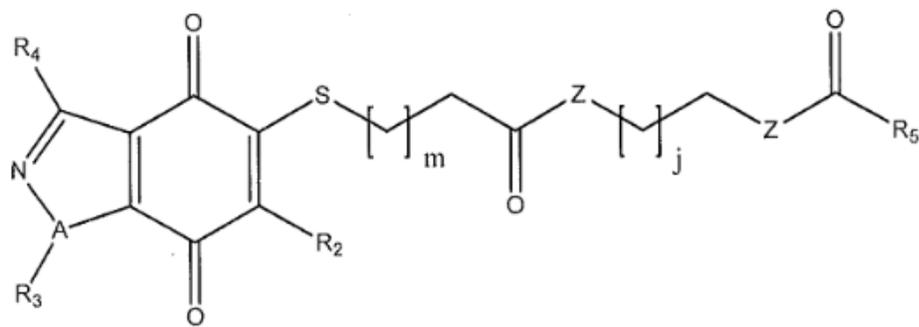
An example: WO2012064632



Ia



Ib



Ic

where:

each Z is independently selected from the group consisting of N, O, S and CH₂;

A is O (then R₃ is null) or N

n=1-5; m=1-5; j=1-4

E is N, CH or CH₂;

R₁, R₂, R₃, R₄, and R₅ individually can represent H, lower straight-chain alkyl (e.g. alkyl groups containing one to ten carbon atoms, such as methyl, ethyl or hexyl) or branched chain lower alkyl, aromatic (e.g. phenyl, naphthyl) substituted aromatic, heteroaromatic (e.g., pyridyl, pyrimidinyl, pyrazole, imidazole, triazole, oxazole, isoxazole, thiazole, isothiazole, oxadiazole, thiadiazole, pyridazine, triazine, indole, indazole, benzoisoxazole, benzoxazole, benzoisothiazole, benzothiazole), substituted heteroaromatic, benzyl, substituted benzyl.

R₁, R₂, R₃, R₄, and R₅ individually can also represent $-(\text{CH}_2)_n\text{-Ar}$, where n is preferably 1 to 3 and Ar is aromatic, substituted aromatic, heteroaromatic or substituted heteroaromatic. A and B individually can also represent $-\text{OR}$, $-\text{SR}$, $-\text{NR}_2$ (where R individually represents H, straight chain or branched chain lower alkoxy, aromatic, substituted aromatic, heteroaromatic, substituted heteroaromatic), halogen, aldehyde, carboxylic acid, $-\text{COOR}$, $-\text{CONHR}$, $-\text{CONR}_2$, $-\text{OCONHR}$, $-\text{OCONR}_2$, $-\text{NCONHR}$, $-\text{NCONR}_2$, $-\text{SCONR}_2$.

R₁, R₂, R₃, R₄, and R₅ can also be individually selected from $-(\text{CH}_2)_n\text{-R}'$, (where R' individually represents H, straight chain or branched chain lower alkoxy, aromatic, substituted aromatic, heteroaromatic, substituted heteroaromatic, halogen, aldehyde, carboxylic acid, $-\text{COOR}$, $-\text{CONHR}$, $-\text{CONR}_2$, $-\text{OCONHR}$, $-\text{OCONR}_2$, $-\text{NCONHR}$, $-\text{NCONR}_2$, $-\text{SCONR}_2$ (where R individually represents H, straight chain or branched chain lower alkoxy, aromatic, substituted aromatic, heteroaromatic, substituted heteroaromatic)).

In some embodiments, R₁, R₂, R₃, R₄, and R₅ are each independently selected from the group consisting of: H, alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkylalkyl, cycloalkylalkenyl, cycloalkylalkynyl, heterocyclo, heterocycloalkyl, heterocycloalkenyl, heterocycloalkynyl, aryl, arylalkyl, arylalkenyl, arylalkynyl, heteroaryl, heteroarylalkyl, heteroarylalkenyl, heteroarylalkynyl, alkoxy, halo, mercapto, azido, cyano, formyl, carboxylic acid, hydroxyl, nitro, acyl, aryloxy, alkylthio, amino, alkylamino, arylalkylamino, disubstituted amino, acylamino, acyloxy, ester, amide, sulfoxyl, sulfonyl, sulfonate, sulfonic acid, sulfonamide, urea, alkoxyacylamino, and aminoacyloxy.

Markush search implementation

- Users' request: free of charge Markush search option
 - ➔ before chemical searches incomplete
- Challenge: implement a, first-level, easy to use and maintain Markush feature
- Free of charge for logged-in users
- Development: Infochem developed it by building further on the PATENTSCOPE Chemsearch components
- License: Clarivate Analytics Markush Data

PATENTSCOPE Markush data and coverage

- Introduction
- Country coverage
- Markush variability types in patents
- Resource Indexing

Introduction

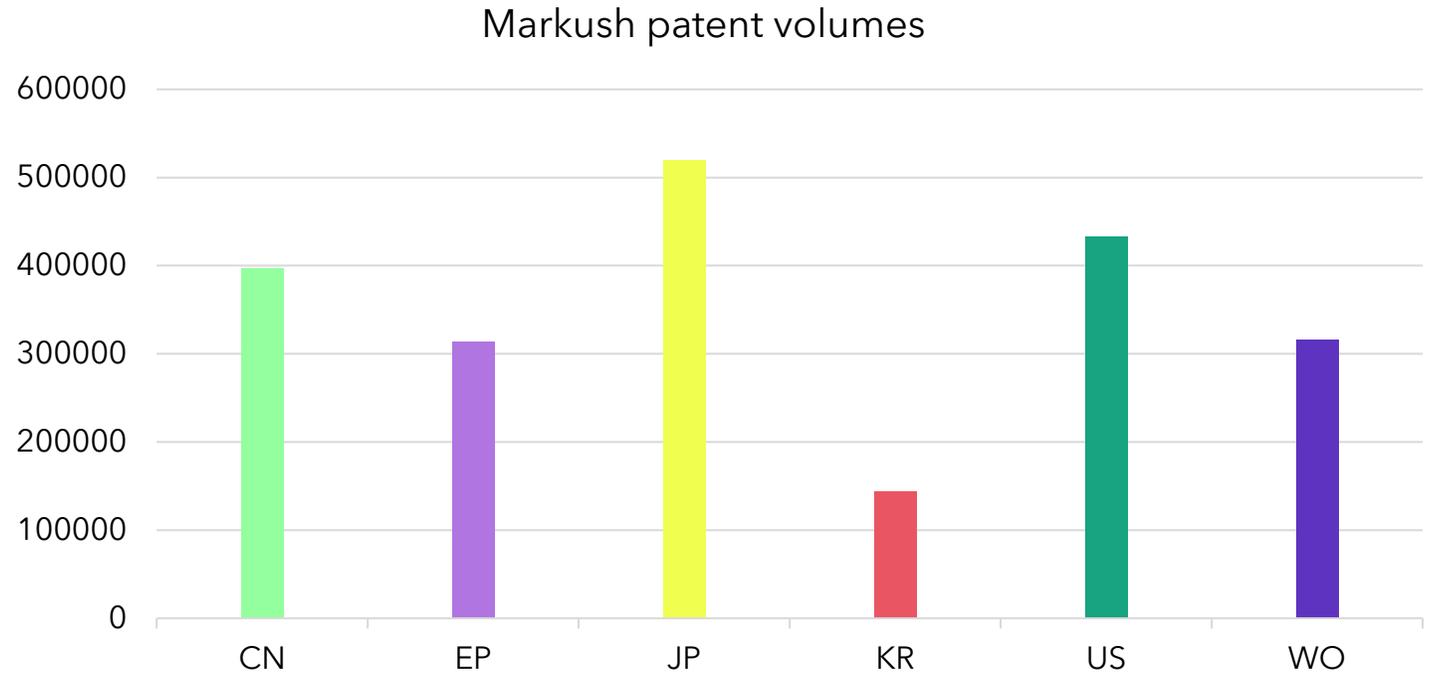
Derwent Markush Resource

Database of 2.4 million editorially curated Markush structures from two sources:

- Clarivate's Derwent World Patents Index (DWPI) editorial process
 - covering the period 1987 to date
 - 2.2 million structures
- Plus Markush structures indexed by INPI (Institute National de la Propriété Industrielle), FR
 - covering the period 1961 to 1998
 - 0.2 million structures
- Each indexed structure is assigned a unique Markush Compound Number, which provides a link from the structure record to the corresponding patent.
- The database is also available via the CAS STNext platform (file DWPIM)

Markush country coverage on PATENTSCOPE

Authority	Start date
China	2008
European Patents	1978
Japan	1987
Korea	2008
PCT/WO	1978
United States	1978



Derwent Markush Resource structure coverage

Indexing policy: Markush structures are indexed principally from the claims, but structures from the description may also be covered where they differ from the claimed structures.

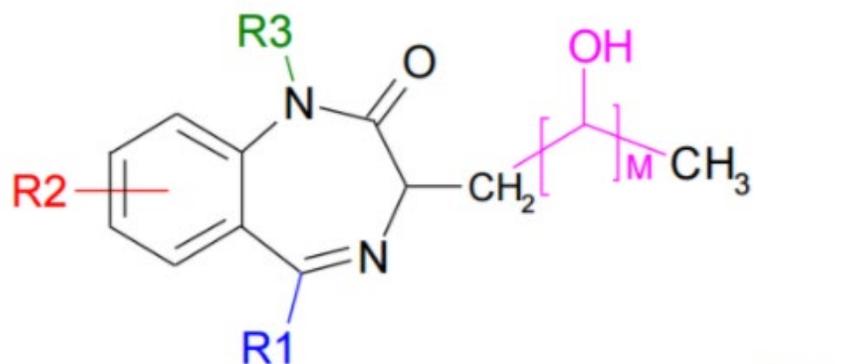
Full Coverage:

- Organic compounds
- Organometallics
- Inorganic salts and metal oxides

Partial coverage:

- Alloys
- Intermetallics
- Polymers

Markush variability types in patents



$R_1 = \text{propyl} / \text{CH}_3\text{CH}(\text{Br})\text{CH}_3 / \text{OMe}^+ /$ ← 

$R_2 = \text{F}$

$M = 2-5$

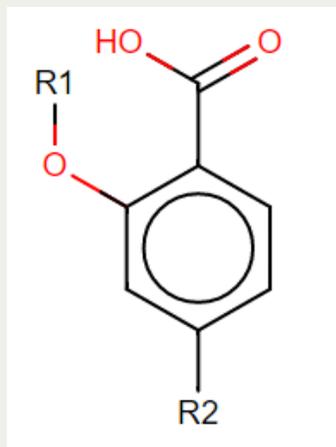
$R_3 = \text{alkyl C1-6}$

- Substituent variation
- Position variation
 - e.g. variable point of attachment to a ring
- Frequency variation (repeating groups)
- Homology group variation
 - e.g. alkyl, alkenyl, aryl etc
- All of the above features are captured in Derwent Markush Resource

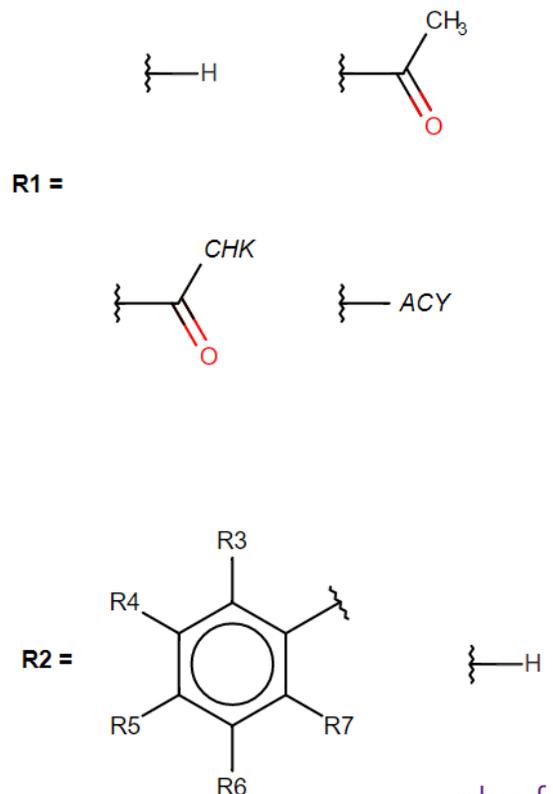
Derwent Markush Resource Indexing

The DWPI editorial team re-draws Markush from the patent according to our indexing rules to capture the Markush core plus all possible variations, e.g.

Markush core



Variable group definitions



plus further definitions for R3 - R7

Generic groups - Superatoms

As well as ordinary atom types (C, N, O, S, Cl etc), Derwent Markush includes generic 'superatoms' that are unique to our database

Acyclic / Cyclic

CHK (Alkyl, Alkylene)

CHE (Alkenyl, Alkenylene)

CHY (Alkynyl, Alkynylene)

ARY (Aryl)

CYC (Cycloaliphatic)

HEA (Monocyclic heteroaryl)

HET (Monocyclic nonaromatic)

HEF (Fused heterocyclic)

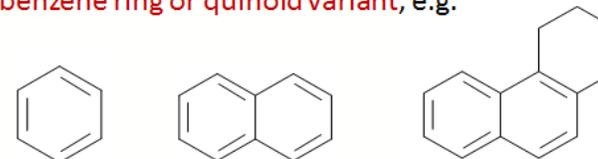
CHK saturated carbon chain (no C=C or C≡C)

CHE carbon chain containing C=C (no C≡C)

CHY carbon chain containing C≡C

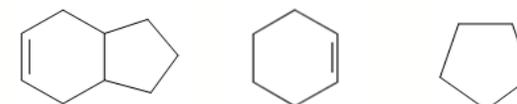
ARY

carbocyclic systems which have **at least one benzene ring or quinoid variant**, e.g.



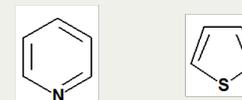
CYC

Cycloaliphatic, monocyclic or fused, non-aromatic carbocyclic systems, **which do not fit with the definition of the ARY superatom**, e.g.



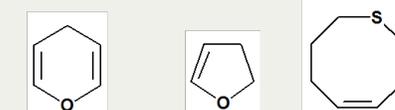
HEA

monocyclic, aromatic heterocycle (heteroaryl),
- either fully unsaturated 6-rings
- or 5-rings with 2 double bonds



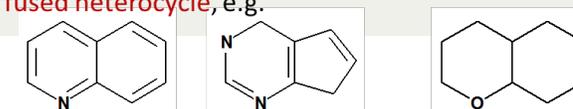
HET

Nonaromatic, monocyclic heterocycle, which **do not fit with the definition of the HEA superatom**, e.g.



HEF

fused heterocycle, e.g.

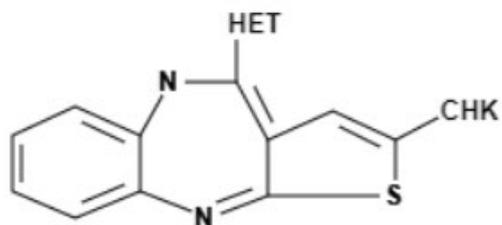


PATENSCOPE search query using generic atoms

Convert structure

Upload structure

Structure editor



Edit Atom Properties

Atom properties

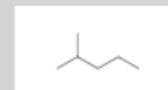
Query atom

Generic atom

Acyclic Hydrocarbons (linear or branched, no rings):

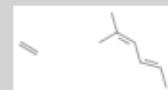
CHK

saturated C-chain



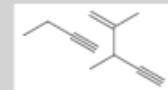
CHE

unsaturated C-chain, no triple bond



CHY

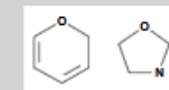
unsaturated C-chain, with triple bond



Heterocyclic Systems (at least one hetero atom):

HET

monocyclic, non-aromatic



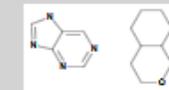
HEA

monocyclic, aromatic



HEF

polycyclic, aromatic and/or non-aromatic



Carbocyclic Systems (mono- or polycyclic rings, no hetero atoms):

CYC

aliphatic



ARY

at least one aromatic ring



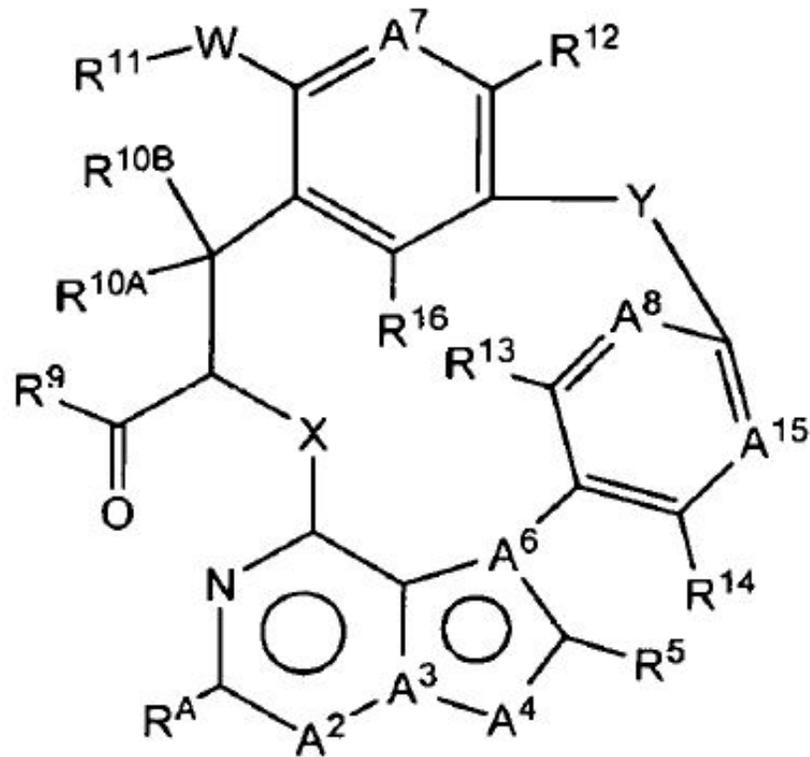
OK

Cancel

Indexing the Markush structure - what features do we look out for?

- Indexing conventions
 - Keto-enol tautomerism (keto form is the preferred one)
 - Amidine normalisation (amidine/guanidine groups have normalised bonds not single and double bonds)
- Use of DWPI Markush terminology and shortcuts
 - Use of superatoms terms (CHK, ARY etc.) & shortcuts (CO₂, SO₃ etc.)
- Include any exceptions and provisos mentioned in the patent
 - e.g. where at least one of R1 & R2 is not H
 - e.g. if R1 = value x, then R2 = a subset of its possible listed values)
- Allow for system limits
 - Check for large numbers of possible values for each variable group, especially when these include further variable sub-groups ('nesting')
 - Each indexed structure can contain up to 1023 atoms and 50 R-groups, so larger Markush structures are split into 2 or more records

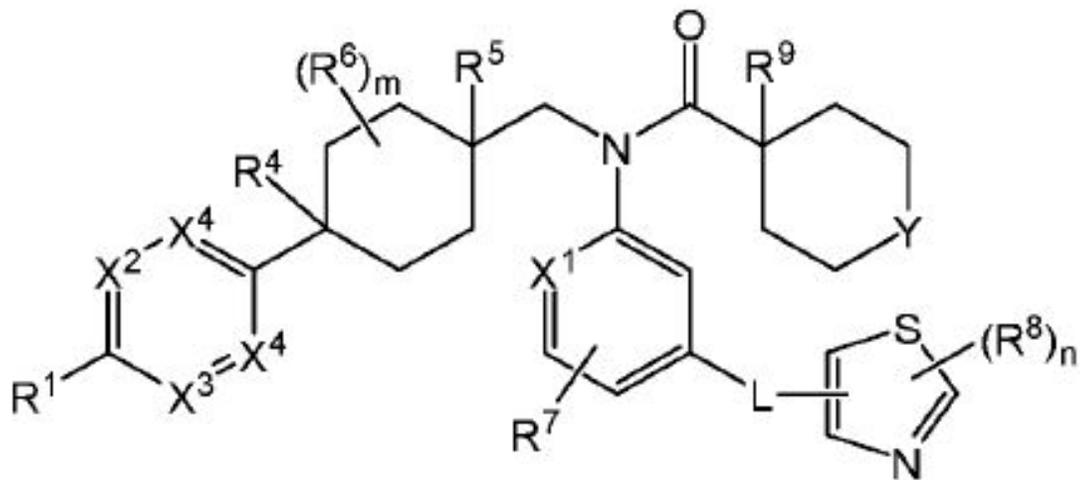
Markush preanalysis 1



(I)

- Variables X and Y within the macrocycle - are they always present?
- Substituents R10A and R10B on one C may fuse
- Substituents R5 and RA may generate tautomers or normalisation if adjacent to double bonds

Markush preanalysis 2



Formula (I)

- R⁶ and R⁸ may be present or absent depending on values of m and n
- Is linker L always present?
- Any of R⁴-R⁶ together may form a ring bridge: R⁴+R⁵ most likely
- Can pairs of R⁶ form a ring or oxo?
- Will any or all of R⁸ generate a tautomer or localised normalisation?
- Might any of X²-X⁴ form a fused ring?

Markush searches in PATENTSCOPE

- Enumeration
- Indexation
 - exact search (XSS)
 - substructure search (SSS)

Enumeration

- Up to the 500 simplest compounds per Markush formulae (frequency variation for Repeating Groups and Multipliers up to $n=4$)
- Index the Inchikeys enumerated compounds in a new PATENTSCOPE search field called 'ENUM'
- Search compounds expressed in Markush formulae as easily as compounds explicitly cited in patents.

Enumerated

PATENTSCOPE Chemical compounds search ▼

[Convert structure](#) [Upload structure](#) [Structure editor](#) [Found compounds](#) [Found Markush Formulas](#)

Search type
Compound name ▼ Type an accepted name, commercial name, CAS name, IUPAC name

Search for scaffold

Include enumerated Markush structures omeprazole

Offices
All ▼

[Reset](#) [Show in editor](#) [Exact Structure Search](#)

CHEM:(SUBDBMMJDZJVOS-UHFFFAOYSA-N) OR ENUM:(SUBDBMMJDZJVOS-UHFFFAOYSA-N)



32,162 results Offices all Languages en Stemming true Single Family Member false Include NPL false



Sort: Relevance ▾ Per page: 100 ▾ View: All+Image ▾

< 1/322 >

Download ▾ Machine translation ▾

1. **0993830** NEW STABILIZED GALENIC FORMULATIONS COMPRISING AN ACID LABILE BENZIMIDAZOLE COMPOUND AND ITS PREPARATION

EP - 19.04.2000

Int.Class [A61K 9/26](#) Appl.No 99116334 Applicant ESTEVE LABOR DR Inventor BALLESTER RODES MONTSERRAT

New stabilized galenic formulations comprising an acid labile benzimidazole compound and its preparation. The new stable oral pharmaceutical formulations are prepared covering an inert nucleus with a first layer containing an acid labile benzimidazole compound of formula I and a water soluble polymer, a second isolation layer consisting in a water soluble polymer, talc and a pigment, and a final enteric coating containing a polymer, a plasticizer and talc.



2. **0773025** NEW STABLE GALENIC FORMULATIONS CONTAINING AN ACID-LABILE BENZIMIDAZOL COMPOUND, AND PRODUCTION PROCESS

EP - 14.05.1997

Int.Class [A61K 9/26](#) Appl.No 96901349 Applicant ESTEVE LABOR DR Inventor BALLESTER RODES MONTSERRAT

New stable galenic formulations containing an acid-labile benzimidazol compound, and production process. Said formulations comprise a neutral nucleus on which is applied a layer containing the active ingredient and comprised of the benzimidazol compound having the general formula [I], a water-soluble polymer and non-alkaline reaction vehicles, and on which is applied a second isolating layer which comprises a water-soluble polymer, a pigment and talcum, and a last enteric layer which contains a polymer, a plastifier and talcum.



3. **1997511257** 酸不安定性ベンズイミダゾールを含有する新規安定型ガレニック製剤、及びその製造方法

JP - 11.11.1997

Int.Class [A61K 31/4439](#) Appl.No 1996523278 Applicant Inventor パレスター・ローデス、 モントセラット



Markush searches by indexation (XSS and SSS)

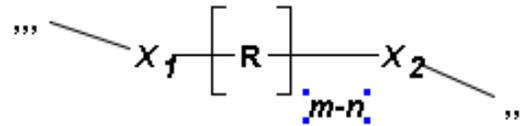
- Each Markush structure in the database is indexed and stored as a superstructure (graph) that encodes all variations (substitution, position, frequency* and homology)
- Each User's query is converted on-the-fly also to a superstructure
- A query is executed by checking if the searched superstructure matches one by one the superstructure of each stored Markush structure in the database (matching of graphs: nodes, links and variations)

Limitation of indexing of the superstructures

Frequency variation:

- For Multipliers: $\dots \diagdown X \begin{matrix} \text{[R]} \\ \text{m-n} \end{matrix}$

Up to n=9 groups can be attached to the superstructure, m is ignored.



- For Repeating Groups:

Only one group (m=n=1) is expanded within the superstructure, m and n are ignored

Search by indexation (graph search)

- **XSS** (**EX**act **S**ubstructure **S**earch)
 - Closed SSS (all atoms “as drawn”)
 - Components will be also hits
 - Standard atom query features are allowed
(atom types A, Q and atom query properties)
 - Standard bond query features are allowed (any, S/D, S/A)
 - Stereo configurations will be ignored
 - Any hit may be a set of structures
- **SSS** (**S**ub**S**tructure **S**earch)
 - Normal SSS
 - Any hit may be a set of structures

Node match Levels

- **Match Level ATOM**

Query nodes match only specific atom nodes in target (default for rings)
Components will be also hits

- **Match Level CLASS**

Normal SSS: Any hit may be a set of structures

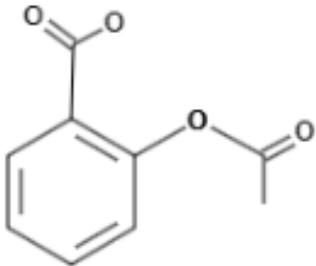
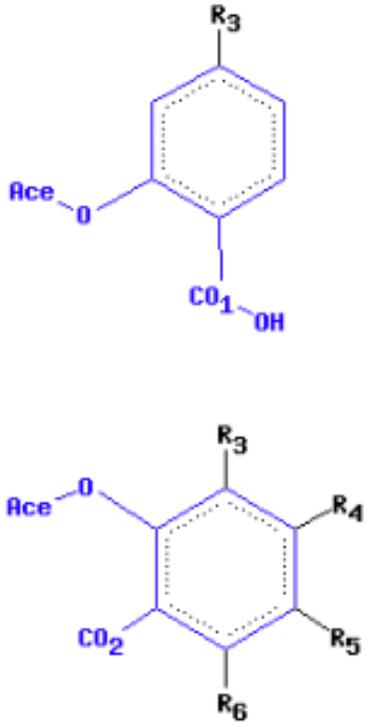
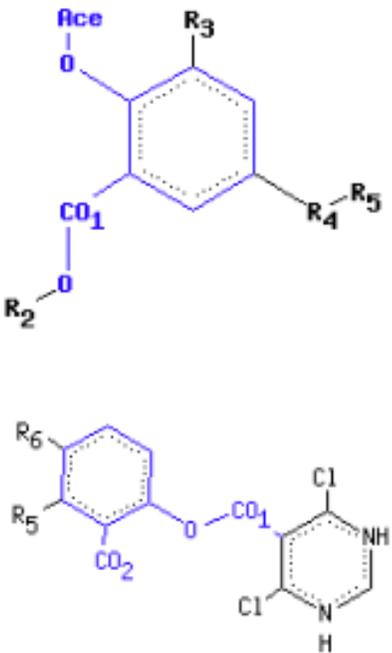
- **Match Level ANY**

Query nodes match additionally XX (XX may be any undefined structural fragment in target)

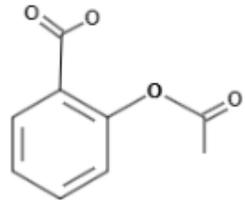
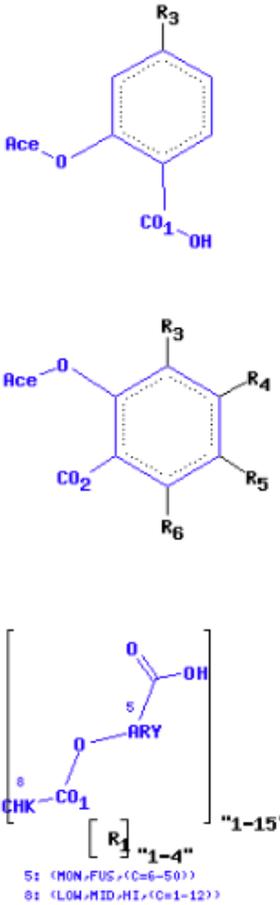
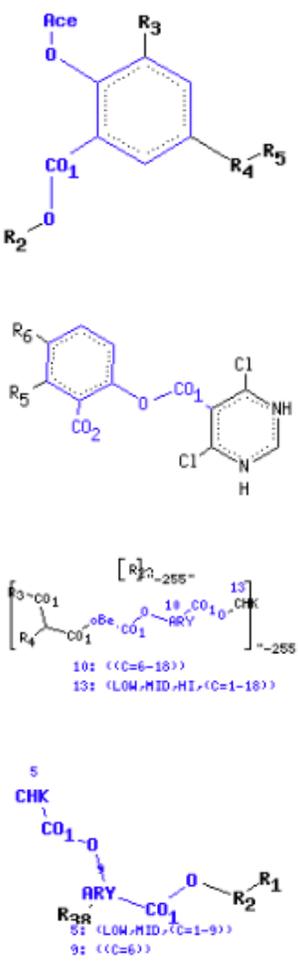
=> Implementation in PATENTSCOPE

All nodes of the query have assigned Match Level CLASS automatically, to retrieve as many relevant hits as possible. Matches on XX are less relevant.

Examples for Match Level ATOM

Query	Hits XSS	Hits SSS
<p data-bbox="591 349 715 382">Aspirine</p>  <p>The query structure shows a benzene ring with a carboxylic acid group (-COOH) at the 1-position and an acetoxy group (-O-CO-CH3) at the 2-position.</p>	 <p>The first XSS hit is a benzene ring with an acetoxy group (-O-Ace) at the 1-position, a hydroxyl group (-CO1-OH) at the 2-position, and a substituent R3 at the 3-position. The second XSS hit is a benzene ring with an acetoxy group (-O-Ace) at the 1-position, a carboxylate group (-CO2) at the 2-position, and substituents R3, R4, and R5 at the 3, 4, and 5 positions, respectively, and R6 at the 6-position.</p>	 <p>The first SSS hit is a benzene ring with an acetoxy group (-O-Ace) at the 1-position, a carboxylate group (-CO1) at the 2-position, a substituent R3 at the 3-position, and substituents R4 and R5 at the 4 and 5 positions, respectively, and R2 at the 6-position. The second SSS hit is a benzene ring with a carboxylate group (-CO2) at the 1-position, an acetoxy group (-O-CO1) at the 2-position, and substituents R5 and R6 at the 3 and 4 positions, respectively. It is linked via an ester bond to a pyridine ring with chlorine atoms at the 2 and 6 positions and a hydrogen atom at the 4-position.</p>

Examples for Match Level CLASS

Query	Hits XSS	Hits SSS
<p data-bbox="715 272 817 301">Aspirine</p> 	 <p data-bbox="1029 1143 1233 1182">5: (MON,FUS,(C=5-9)) 8: (LOW,MID,HI,(C=1-12))</p>	 <p data-bbox="1396 1196 1564 1239">5: (LOW,MID,(C=1-9)) 9: ((C=6))</p>

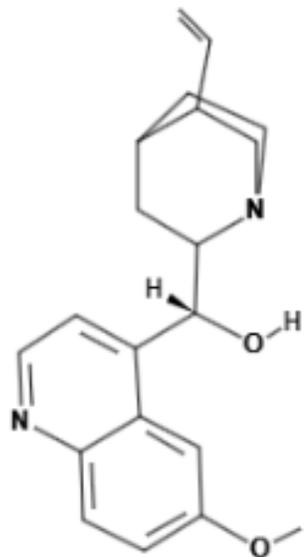
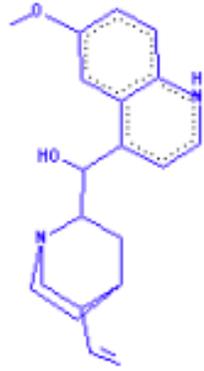
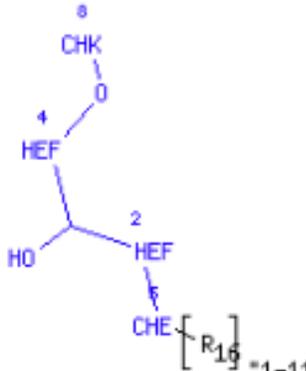
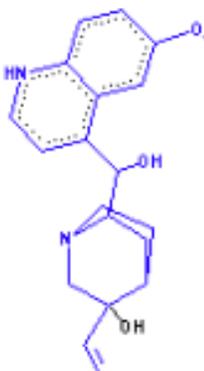
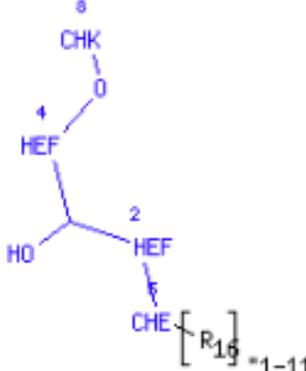
Element Count Levels

- **Element Count Level LIMITED**
Any explicit or implicit defined element count of the query nodes must match on the element count of the target nodes (default)
- **Element Count Level UNLIMITED**
All element counts of the query nodes will be ignored when matching on target nodes.

=> Implementation in PATENTSCOPE

all nodes have assigned the Element Count Level LIMITED.

Examples for Element Count LIMITED

Query	Hits XSS	Hits SSS
<p>Chinine</p> <p>chiral</p> 	  <p>2: ((C)=2,(N)=0,(O)=0,(S)=0,(X)=0)</p>	  <p>2: ((C)=2,(N)=0,(O)=0,(S)=0,(X)=0)</p>

Generic Nodes for extended search: Chains

CHK: linear or branched C-chain with 1- n ($n \geq 1$) C-Atoms

CHE: linear or branched C-chain with 2- n ($n > 1$) C-Atoms and at least 1 double bond, but no triple bond.

CHY: linear or branched C-chain with 2- n ($n > 1$) C-Atoms and at least 1 triple bond

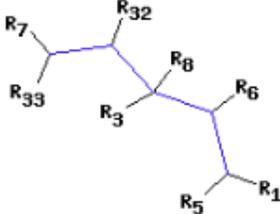
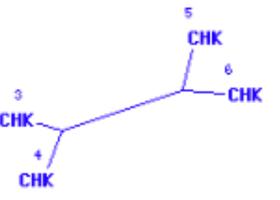
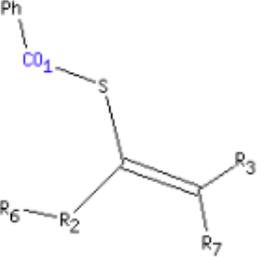
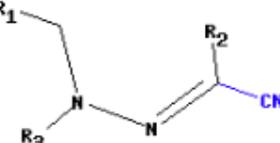
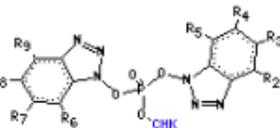
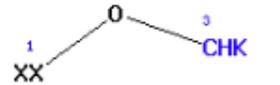
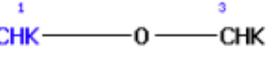
Generic Nodes for extended search: Rings

- CYC: carbocyclic, non-aromatic, mono- or polycyclic ring, without hetero atoms and arbitrary number of ring atoms (size \geq 3).
- ARY: carbocyclic, mono- or polycyclic ring, without hetero atoms, where at least one of the rings is a six-membered, aromatic ring. The size must be six or larger (size \geq 6).
- HET: monocyclic, non-aromatic ring with at least one hetero atom and arbitrary size (size \geq 3).
- HEA: monocyclic, aromatic ring with at least one hetero atom and size 5 or 6.
- HEF: polycyclic, aromatic or non-aromatic ring with at least one hetero atom and size \geq 4.

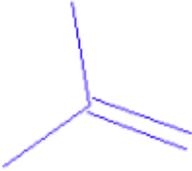
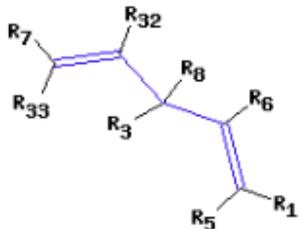
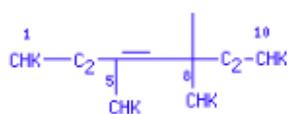
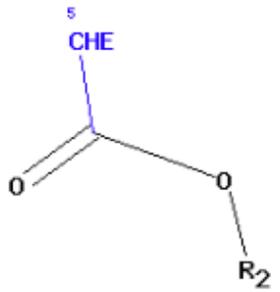
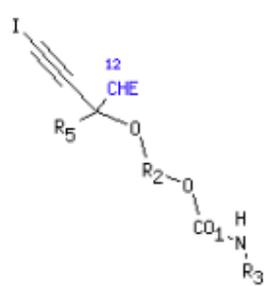
Matches for free sites, chain and ring extensions

Query-Node	Generic XSS Target Node (no free sites in query)	Generic SSS Target Node (free sites in query)
C, CHK	CHK	CHK CHE, CHY CYC, ARY, HET, HEA, HEF
C=C, CHE	CHE	CHE CHY CYC, ARY, HET, HEA, HEF
C#C, CHY	CHY	CHY CYC, ARY, HET, HEA, HEF
CYC, Carbocycle	CYC	CYC ARY, HEF
ARY, aromatic ring	ARY	ARY HEF
HET, hetero cycle	HET	HET HEF
HEA, aromatic hetero cycle	HEA	HEA HEF
HEF, fused hetero cycle	HEF	HEF

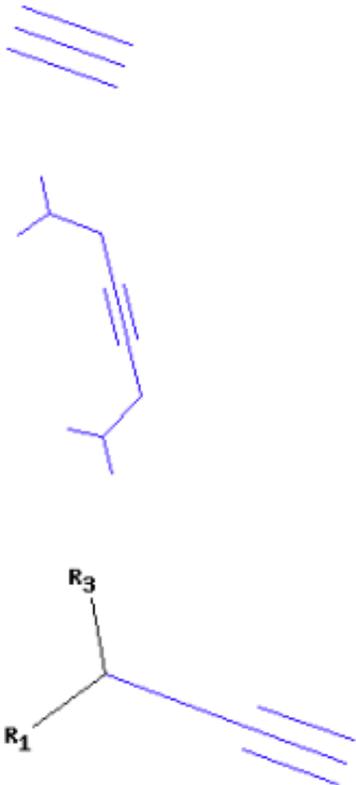
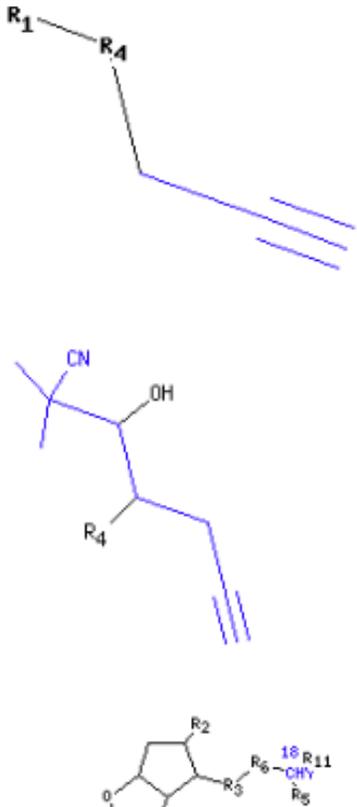
Examples for Match of CHK (C-chain, saturated)

Query	Hits XSS	Hits SSS
<p><i>CHK</i></p>	<p><i>C₁₂</i></p> <p><i>iBu</i> — <i>iPr</i></p>  <p><i>CHK</i></p> <p><i>CHK</i> — <i>CHK</i> — <i>CHK</i></p>  <p> <i>3: <STR,BRA,LOW,MID,HI,(C=...></i> <i>4: <STR,BRA,LOW,MID,HI,(C=...></i> <i>5: <STR,BRA,LOW,MID,HI,(C=...></i> <i>6: <STR,BRA,LOW,MID,HI,(C=...></i> </p>	    <p> <i>1: <<POLYSACCHARIDE RESIDUE,>></i> <i>3: <MID,HI,(C=8-28)></i> </p>  <p> <i>1: <LOW,(C=2-3)></i> <i>3: <MID,HI,(C=8-16)></i> </p>

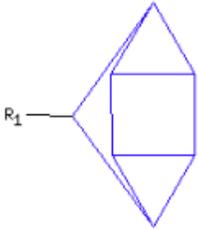
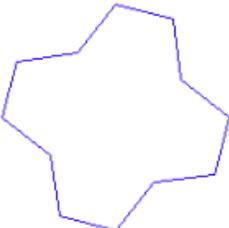
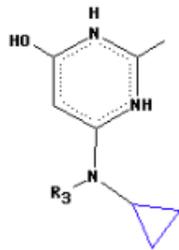
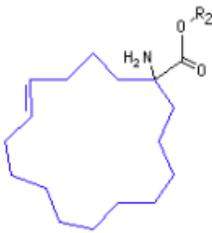
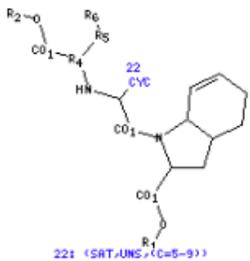
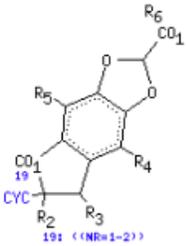
Examples for Match of CHE (C-Chain, unsaturated with double bonds)

Query	Hits XSS	Hits SSS
<p><i>CHE</i></p>	   <p>1: (LOW,MID,HI,(C=4-12)) 5: (LOW,MID,HI,(C=4-12)) 6: (LOW,MID,HI,(C=4-12)) 10: (LOW,MID,HI,(C=4-12))</p>	 <p>5: (MID,HT)</p>  <p>12: (STR,BRA,LOW,(C=2-6))</p>
	<p><i>CHE</i></p>  <p>3: (STR,BRA,LOW,MID,(C=2-7)) 4: (STR,BRA,LOW,MID,(C=1-7))</p>	

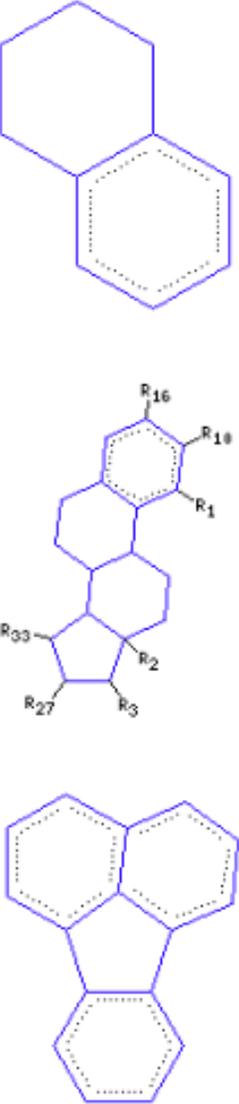
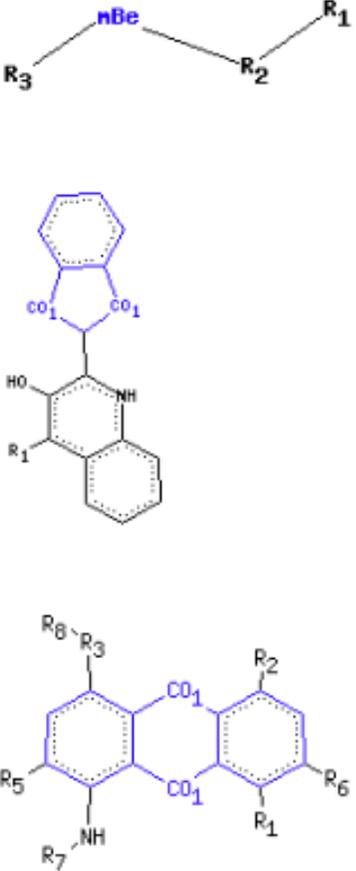
Examples for Match of CHY (C-Chain, unsaturated with triple bonds)

Query	Hits XSS	Hits SSS
<p>CHY</p>	 <p>1 CHY 1: (STR,BRR,LO W,MID,HI,C...</p>	 <p>18: (STR,BRR,LOW,MID,(C=2-10))</p>

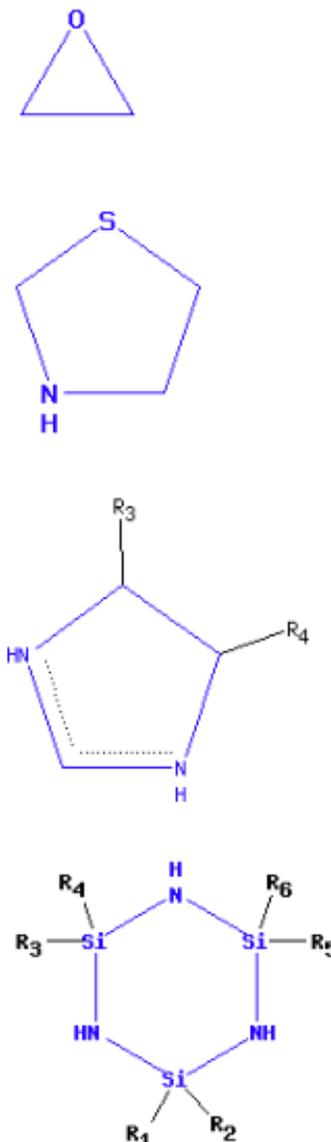
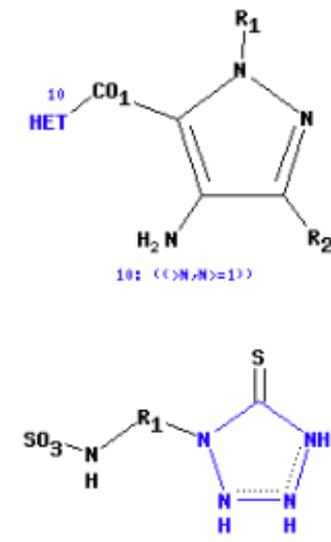
Examples for Match of CYC (Carbocycle)

Query	Hits XSS	Hits SSS
<p data-bbox="868 239 927 268">:CYC:</p>	<div data-bbox="983 229 1182 458">  </div> <div data-bbox="983 505 1212 733">  </div> <div data-bbox="983 791 1090 872"> <p data-bbox="1014 791 1029 805">1</p> <p data-bbox="1021 811 1072 833">CYC</p> <p data-bbox="991 836 1090 872">1: <UNS> <C=3-30></p> </div> <div data-bbox="983 933 1090 1015"> <p data-bbox="1014 933 1029 948">1</p> <p data-bbox="1021 953 1072 976">CYC</p> <p data-bbox="991 979 1090 1015">1: <UNS,FUS> <DBO=2,NR=2></p> </div>	<div data-bbox="1268 229 1447 479">  </div> <div data-bbox="1268 525 1480 758">  </div> <div data-bbox="1268 811 1518 1072"> <p data-bbox="1386 872 1411 886">22</p> <p data-bbox="1393 892 1444 915">CYC</p> <p data-bbox="1319 1053 1472 1072">22: <SRT,UNS> <C=5-9></p>  </div> <div data-bbox="1268 1133 1454 1379"> <p data-bbox="1276 1300 1302 1315">19</p> <p data-bbox="1284 1320 1335 1343">CYC</p> <p data-bbox="1304 1358 1411 1376">19: <NR=1-2></p>  </div>

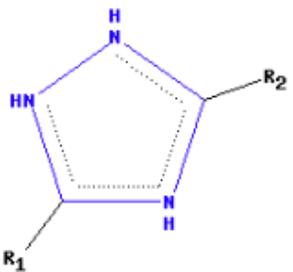
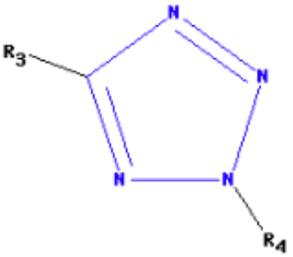
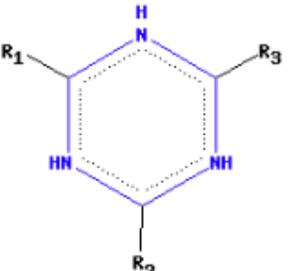
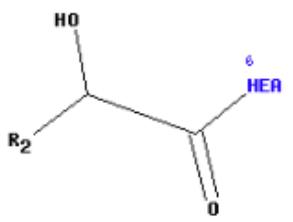
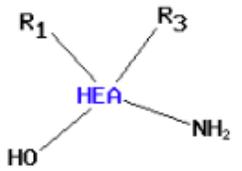
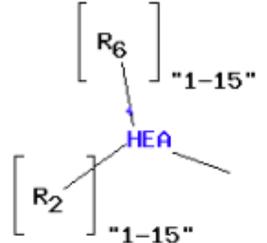
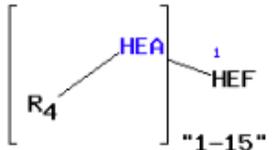
Examples for Match of ARY (Aromatic Carbocycle)

Query	Hits XSS	Hits SSS
<p>ARY</p>		

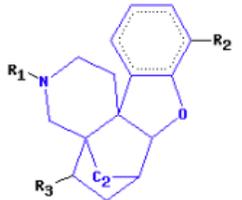
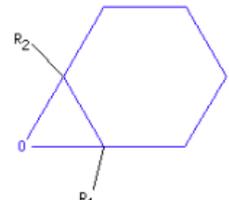
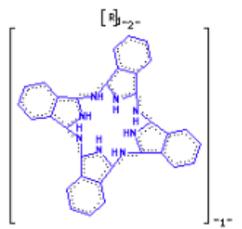
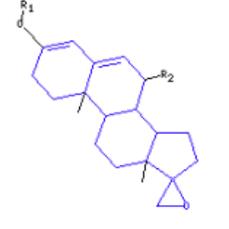
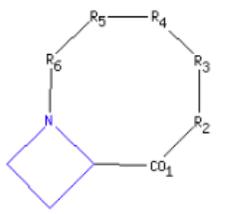
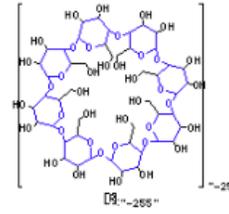
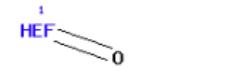
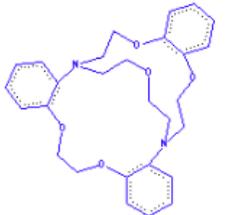
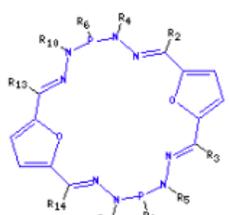
Examples for Match of HET (Heterocycle)

Query	Hits XSS	Hits SSS
<i>HET</i>		

Examples for Match of HEA (Hetero Aromatic Cycle)

Query	Hits XSS	Hits SSS
HEA	 <p>1 HEA 1: << N>=1>></p>  	 <p>6: <<(N,N)=1>></p>   <p>4: <<(N,N)=1>></p>  <p>1: <<(NR)=5,(N)=6,(O)=0,(S)=0,>></p>

Examples for Match of HEF (Hetero Fused Cycle)

Query	Hits XSS	Hits SSS
HEF		
		
		
	<p>1 HEF 11 ((RA)=4,DBC >=1,NR=2,O=1))</p>	<p>1 HEF 11 ((RA)=6,NR=2,O=1))</p> 
		

Generic Nodes Properties

Enhanced Generic Chain Properties	Keywords
branched / linear (default: any)	bra / lin
number of carbon atoms (default: >1)	LOW(1-4), MID(5-7), HI(>7), size=m-n

Enhanced Generic Ring Properties	Keywords
saturated (default) / unsaturated	sat / unsat
number of rings (range allowed)	NR=m -n
number of ring atoms (range allowed)	RA=m-n
number of heteroatoms (range allowed)	NHet=m-n
monocyclic (default) / polycyclic	mon / pol
number of carbon atoms (range allowed)	NC=m-n
Element count (range allowed)	ES m-n ES: Element symbol

PATENTSCOPE Chemical compounds search ∨

Convert structure

Upload structure

Structure editor

Found compounds

Found Markush Formulas

Search type

Compound name

Type an accepted name, commercial name, CAS name, IUPAC name

aspartame

Search for scaffold

Include enumerated Markush structures

Offices

All

Reset

Show in editor

Exact Structure Search

Convert structure

Upload structure

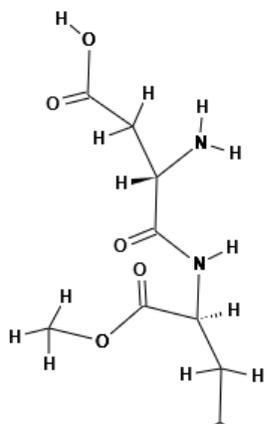
Structure editor

Found compounds

Found Markush Formulas



chiral



InChI: InChI=1S/C14H18N2O5/c1-21-14[20]11[7-9-5-3-2-4-6-9]16-13[19]10[15]8-12[17]18/h2-6,10-11H,7-8,15H2,1H3,[H,16,19][H,17,18]/t10-,11-/m0/s1

InChiKey: IAQZJIPTCAWIRG-QWRGUYRKSA-N

Molecular Formula: C14H18N2O5

Molecular Weight: 294.3072 g/mol



Search for scaffold

Include enumerated Markush structures

Offices

All

Substructure Search

Exact Search

Reset

▼ Markush Search

Substructure Search

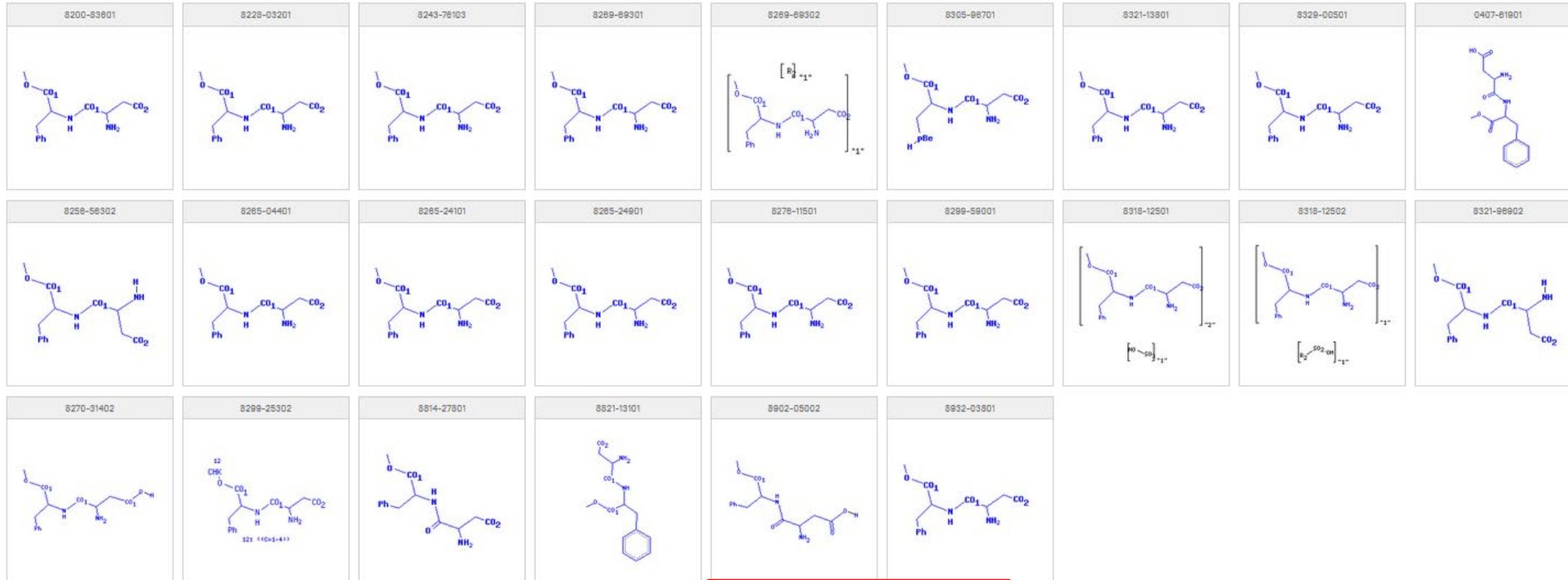
Exact Structure Search

Evaluate

search results (41 hits found, 7.9% searched)

Sort by natural

[1 of 2] 1 2 24



[1 of 2] 1 2 24

Markush search results (41 hits found, 7.9% searched)

Offices
All

Reset

Clear all

Select all

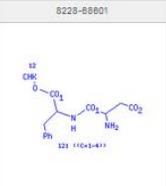
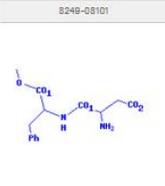
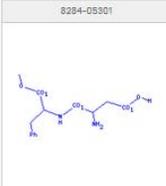
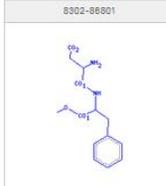
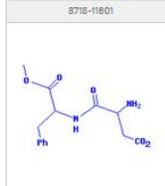
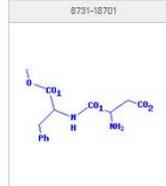
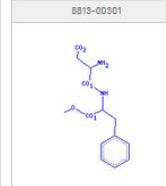
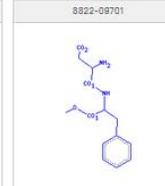
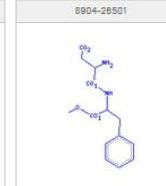
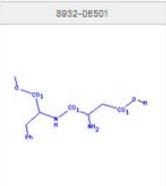
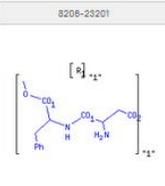
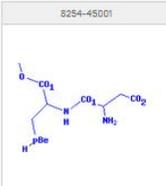
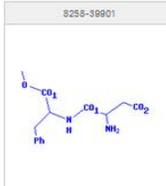
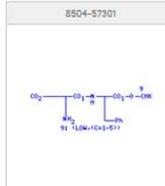
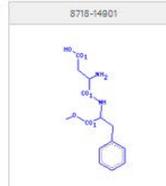
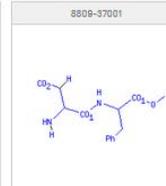
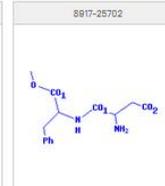
Search

Batch

Convert structure Upload structure Structure editor Found compounds Found Markush Formulas

search results [41 hits found, 7.9% searched] Sort by natural

[2 of 2] 1 2 24

8228-88601 	8249-08101 	8284-05301 	8302-86801 	8716-11601 	8731-18701 	8815-00301 	8822-09701 	8904-26501 
8932-08501 	8206-23201 	8254-45001 	8258-39901 	8504-57301 	8719-14901 	8809-37001 	8917-25702 	Show more...

[2 of 2] 1 2 24

Markush search results [41 hits found, 7.9% searched]

Offices
All

Reset Clear all Select all Search Batch

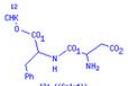
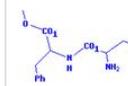
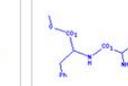
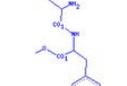
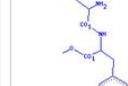
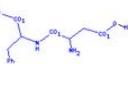
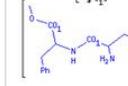
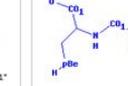
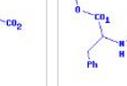
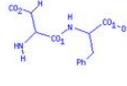
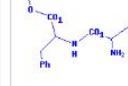
PATENTSCOPE Chemical compounds search

Convert structure Upload structure Structure editor Found compounds Found Markush Formulas

search results [41 hits found, 7.9% searched]

Sort by natural

[2 of 2] 1 2 24

8228-69601 	8246-08101 	8284-05301 	8302-86801 	8718-11801	8731-18701	8813-00301 	8822-08701 	8904-26501 
8992-06501 	8208-23201 	8254-45601 	8258-98901 	<div data-bbox="1039 578 1439 806"><p>Markush batch</p><p>There is currently one batch in the queue. Do you want to proceed?</p><p>Name aspartame</p><p>Cancel Launch</p></div>		8808-37001 	8917-29702 	Show more...

[2 of 2] 1 2 24

Markush search results [41 hits found, 7.9% searched]

Offices
All

Reset Clear all Select all Search Batch

Markush batches

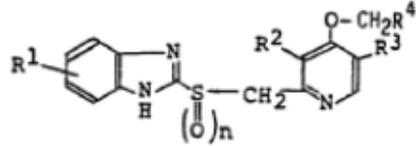
These are your Markush searches executed in batch mode in PATENTSCOPE.

Date	Name	Type	Status	#	
14.10.2021 17:16	MJIHNNLFOKEZEW-UHFFFAOYSA-N	Exact	COMPLETED	0	 
08.06.2023 09:35	omeprazole_june2023	Exact	COMPLETED	148	 
12.06.2023 15:02	aspartame	Exact	QUEUED	0	 

Markush information display: example EP0174726

[0006] The present invention relates to

[1] pyridine derivatives of the formula (I)



I

wherein R¹ is hydrogen, methoxy or trifluoromethyl, R² and R³ are independently hydrogen or methyl, R⁴ is a C₁₋₄ fluorinated alkyl, and n denotes 0 or 1, or their pharmacologically acceptable salts and

2. EP0174726 - PYRIDINE DERIVATIVES AND THEIR PRODUCTION



National Biblio. Data Description Claims Compounds **Markush** Documents

PermaLink Machine translation ▼

Office

European Patent Office

Application Number

85305458

Application Date

31.07.1985

Publication Number

0174726

Publication Date

19.03.1986

Publication Kind

B1

IPC

A61K 31/44

A61K 31/4409

A61K 31/4418

A61K 31/4427

A61P 1/04

C07D 213/68

[View more classifications](#)

CPC

A61P 1/04

C07D 213/68

C07D 213/89

C07D 401/12

Applicants

TAKEDA CHEMICAL INDUSTRIES, LTD.

Inventors

NOHARA, AKIRA

MAKI, YOSHITAKA

Title

[DE] Pyridin-Derivate und deren Herstellung.

[EN] PYRIDINE DERIVATIVES AND THEIR PRODUCTION

[FR] Dérivés de pyridine et leur préparation.

Abstract

Other related publications

[BG60415](#)

[DK356485](#)

[ES546152](#)

[ES8607288](#)

2. EP0174726 - PYRIDINE DERIVATIVES AND THEIR PRODUCTION



National Biblio. Data Description Claims Compounds **Markush** Documents

PermaLink

Markush Nr.

8265-43501

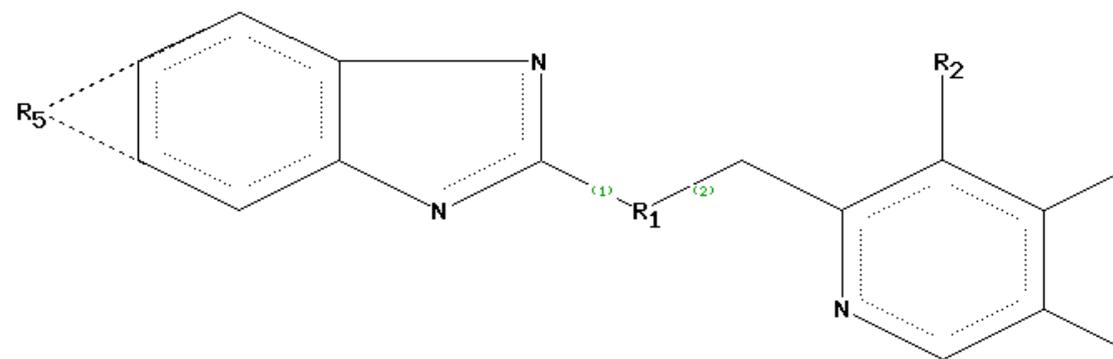
▶ Markush formula

▶ Enumerated compounds

Markush Nr.

8265-43501

▼ Markush formula



R1 =

R2 =

r

Markush Nr.

8265-43501

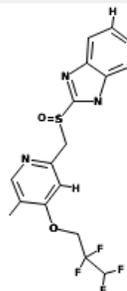
▶ Markush formula

▼ Enumerated compounds

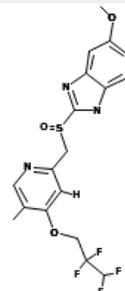
Note: These structures have been created automatically. Please use the original Markush definition in the PDF version for legal matters

1 2 3 4 5 6 7 8 9 10

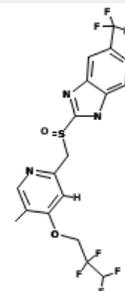
FCNYJRDBGFIPEL-UHFFFAOYSA-N



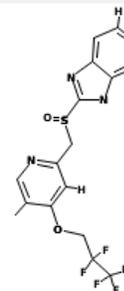
YJSGNOIHFPMTL-UHFFFAOYSA-N



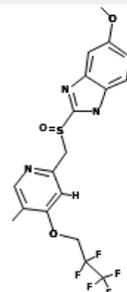
RGVCLMGEEKDNBP-UHFFFAOYSA-N



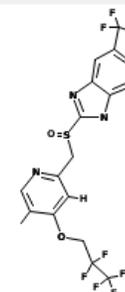
LLUMYBCVFOSPTD-UHFFFAOYSA-N



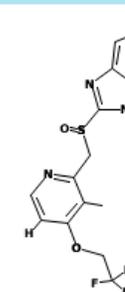
KTUAGXLCJVEKGX-UHFFFAOYSA-N



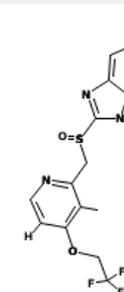
OYZRPLGVRJWVOP-UHFFFAOYSA-N



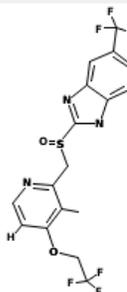
MJIHNNLFOKEZEW-UHFFFAOYSA-N



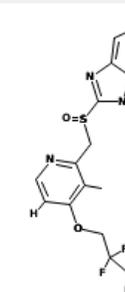
SOVFHOUJKUQOBG-UHFFFAOYSA-N



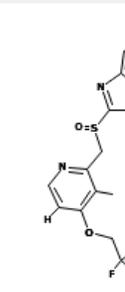
FTMFDXWHZNPGR-UHFFFAOYSA-N



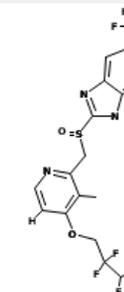
UPLPMFJOVHZZKT-UHFFFAOYSA-N



XPBLNAHUILTEPL-UHFFFAOYSA-N



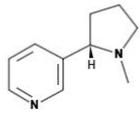
FLZRETSUHOXVNH-UHFFFAOYSA-N



Case study 1: Nicotine in Markush formulae

PATENTSCOPE Chemical compounds search ▾

Convert structure Upload structure **Structure editor** Found compounds Found Markush Formulas



InChI: InChI=1S/C10H14N2/c1-12-7-3-5-10[12]9-4-2-6-11-8-9/h2,4,6,8,10H,3,5,7H2,1H3/t10-/m0/s1
InChIKey: SNICXCGAKADSCV-JTQLQIEISA-N
Molecular Formula: C10H14N2
Molecular Weight: 162.2346 g/mol

Search for scaffold

Include enumerated Markush structures

Offices
All

Reset Markush Search Substructure Search Exact Structure Search Evaluate

CHEM:(SNICXCGAKADSCV-JTQLQIEISA-N) OR ENUM:(SNICXCGAKADSCV-JTQLQIEISA-N)



85,842 results Offices all Languages en Stemming true Single Family Member false Include NPL true



Sort: Pub Date Asc ▾ Per page: 10 ▾ View: All ▾

< 1 / 8,585 ▾ >

Download ▾ Machine translation ▾

1. [4132791](#) USE OF ETOPERIDONE IN PARKINSONISM AND IN OTHER EXTRAPYRAMIDAL SYNDROMES CHARACTERIZED BY TREMORS

US - 02.01.1979

Int.Class [A61K31/495](#) ⓘ Appl.No 05793336 Applicant Aziende Chimiche Riunite Angelini Francesco A.S.R.A.F. S.p.A. Inventor Silvestrini Bruno

Treatment of tremors in Parkinsonism and in other extrapyramidal syndromes by administering an amount of trazodone of the formula ##STR1## AND OF ETOPERIDONE OF THE FORMULA: ##STR2## IN THE RANGE OF ABOUT 25 MG. TO 100 MG. THREE TIMES A DAY.

2. [4141902](#) 1-HALOMETHYL-6-PHENYL-4H-S-[:4,3-A];[:1,4];BENZODIAZEPINES

US - 27.02.1979

Int.Class [C07D243/00](#) ⓘ Appl.No 05138288 Applicant The Upjohn Company Inventor Hester Jackson B.

1-Substituted-6-phenyl-4H-s-triazolo[4,3-a];[:1,4];benzodiazepines of the formula II: ##STR1## wherein R.sub.0 is selected from the group consisting of hydrogen, methyl, ethyl, fluorine, chlorine, and bromine, and X is selected from fluorine, chlorine, and bromine; wherein R.sub.1 is selected from the group consisting of hydrogen, alkyl of 1 to 3 carbon atoms inclusive; wherein R.sub.2, R.sub.3, R.sub.4, and R.sub.5 are selected from the group consisting of hydrogen, alkyl, as defined above, halogen, nitro, cyano, trifluoromethyl, and alkoxy, alkylthio, alkylsulfinyl, and alkylsulfonyl, in which the carbon chain moieties are of 1 to 3 carbon atoms, inclusive, are prepared by reacting a compound of the formula: ##STR2## wherein R.sub.1, R.sub.2, R.sub.3, R.sub.4, and R.sub.5 are defined as above, with an haloacyl chloride or haloacyl bromide of the formula ##STR3## wherein R.sub.0 and X are defined as above and X' is chlorine or bromine. The new products of formula II including their pharmacologically acceptable acid addition salts are useful as sedatives, tranquilizers and muscle relaxants in mammals and birds, and if X.dbd.X".dbd.Cl or Br they are useful as intermediates for cyano, alkylthio and dialkylamino products of high sedative and tranquilizing activity.

3. [RE029934](#) DERIVATIVES OF 5-CARBAMOYLOXYMETHYL-3-SUBSTITUTED-2-OXAZOLIDINONES, PROCESS OF PREPARATION THEREOF AND THERAPEUTIC APPLICATION

US - 13.03.1979

Int.Class [C07D295/20](#) ⓘ Appl.No 05663563 Applicant Delalande S. A. Inventor Fauran Claude P.

A compound of the formula ##STR1## .[:.in which R is H, halogen, alkyl having 1 to 4 carbon atoms or trifluoromethyl, and R.sub.1 and R.sub.2 each is H or dialkylaminoalkyl or together with N comprise methylpiperazino.];. .ladd.in which [1] .ladd. ##STR2## .laddend. is -NH.sub.2 and R is H, m-CH.sub.3, o-CF.sub.3, m-Cl, p-Cl, m-Br, m-F, p-F, o-F, p-CH.sub.3 or m-CH.sub.3, or .ladd. ##STR3## .laddend. and R is H, m-CF.sub.3 or m-F, or .ladd. ##STR4## .laddend. and R is H, m-CF.sub.3 or m-F. PAL The compound is prepared by treating the corresponding 5-hydroxymethyl compound with ammonia or an amine in the presence of phosgene. PAL The compounds possess anti-depressive, myorelaxing, tranquilizing, sedative, analgesic, anti-convulsive, anti-pyretic, anti-inflammatory and uricosuric activities.

4. [4145518](#) FIBERS, FILTER RODS AND OTHER NONWOVEN ARTICLES MADE FROM POLY[1,2-PROPYLENE TEREPHTHALATE]COPOLYESTERS OF TEREPHTHALIC ACID, 1,2-PROPYLENE GLYCOL AND ETHYLENE GLYCOL

US - 20.03.1979

ENUM:(SNICXCGAKADSCV-JTQLQIEISA-N)



0 results Offices all Languages en Stemming true Single Family Member false Include NPL true



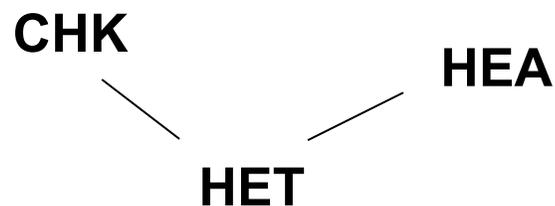
No results were found to match your search
Try modifying your search criteria

After a few seconds, the search by enumeration does not bring back any result...

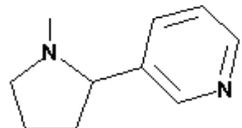
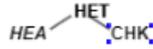
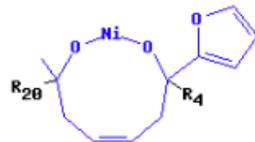
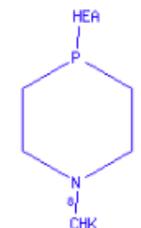
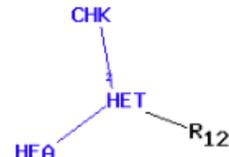
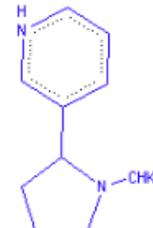
- The XSS search brings 70 Markush hits after 30 minutes (batch mode), which in turn are referred by 132 distinct patent documents

- => If you don't find what you are looking for, use the IC Editor to make the search more generic:

Nicotine can be made more generic with:

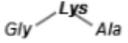
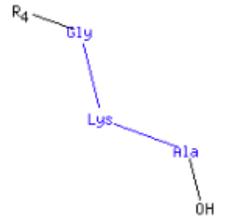
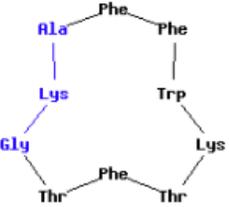
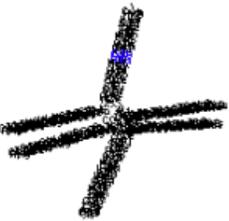


2138 Markush structures => too large to investigate and probably too noisy

Query	Hits XSS
Nicotine 	70 hits
Generic Nicotine, unspecific 	2138 hits   <small>8: ((LOW*(C=1-4))</small>  <small>2: ((R0=4-10,N)=1))</small> 

Case study 2: peptides

Searching peptides in Markush structures?
yes, you can!

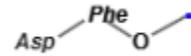
Query	XSS-Hits	SSS- Hits
	0 hits	1723 hits    

query

XSS hits

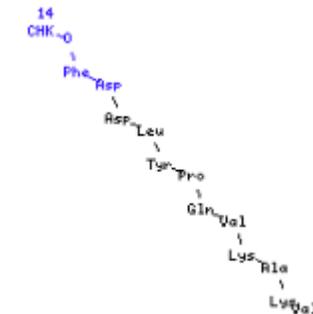
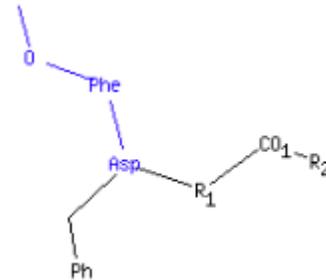
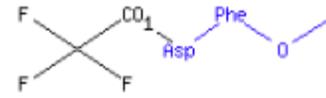
SSS hits

Aspartame
Asp-Phe-Ome



490 hits

590 hits



14: (STR,BRA,LOW)

Participant list

The participant list will be shared with Clarivate & DeepMatter unless you clearly indicate by email patentscope@wipo.int so **you do not agree** to share your contact information.

Q&A





patentscope@wipo.int

Public Request For Information

■ Published February, 2019

Findings:

=> Creation of structured Markush information from scratch for the world patents backfile and front file “out of reach” for WIPO

=> Only one possibility: license the data from a provider and implement search functionality in PATENTSCOPE