

化学结构式的高效检索技巧

更高效易用的检索分析工具

培训讲师：刘冬梅



目录

CONTENTS

01 化学多检索入口使用场景

- 子结构检索与相似结构检索的应用场景
- 组合结构式搜索的应用场景
- 批量搜索的应用场景
- 性质搜索的应用场景
- 化学高亮，化合物二次提取的应用场景

02 化学检索画图工具应用

- 多种画图软件在智慧芽做图工具有兼容性
- 环状结构任意位置的检索
- 同族基团和广义原子简化Markush化学结构的检索
- R基团固定多类别取值检索

03 化学与专利数据库的巧妙联合搜索

- 检索式构建(实现查全，多类技术信息运算，去噪音)
- 快速确定原研药公司专利布局策略
- 风险排查巧查全

01

PART

化学多检索入口使用场景

- 子结构检索与相似结构检索的应用场景
- 组合结构式搜索的应用场景
- 批量搜索的应用场景
- 性质搜索的应用场景
- 化学高亮，化合物二次提取的应用场景

结构搜索

patsnap

输入化合物身份信息或绘制上传化合物结构式进行多类型化合物及专利检索。

The screenshot displays the patsnap chemical search interface. At the top, there is a navigation bar with 'chemical by patsnap', '搜索' (Search), 'LabBench', and '邮件提醒' (Email Reminder). Below this, a search bar is labeled '输入物质身份信息或导入结构图片' (Enter substance identity information or import structure image). To the right of the search bar, a red box highlights '化合物身份信息 (通用名称、结构式名称、分子式、CAS号、SMILES、INCHIkey等)' (Compound identity information (General name, structural name, molecular formula, CAS number, SMILES, INCHIkey, etc.)).

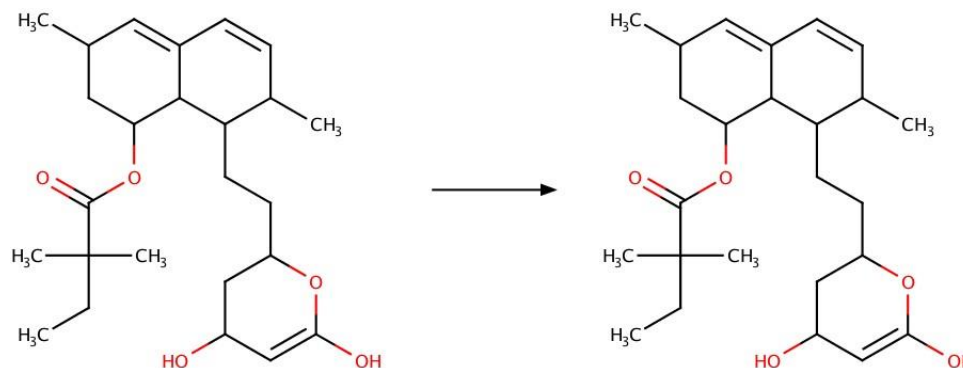
Below the search bar, there are four tabs: '结构式搜索' (Structure search), '组合结构式搜索' (Combination structure search), '性质搜索' (Property search), and '批量搜索' (Batch search). The '结构式搜索' tab is active. To the right of these tabs, a red box highlights '常见问题及功能' (Common questions and functions).

The main workspace is divided into two sections. The left section is a chemical structure editor, labeled '构建编辑化合物结构式' (Construct and edit compound structure). It features a toolbar with various drawing tools and a large canvas displaying the 'Marvin JS by ChemAxon' logo. The right section is the '搜索选项' (Search options) panel. It includes a '添加关键词' (Add keywords) section with a dropdown menu set to '专利名称/摘要/权利要求' (Patent title/abstract/claims). Below this, the '化合物出现位置' (Compound location) section has checkboxes for '专利名称/摘要/权利要求' and '说明书' (Specification), both of which are checked. The '搜索方式' (Search method) section has a dropdown menu set to '结构式' (Structure) and a '马库什' (Markush) option. Below this, the '精确搜索' (Exact search) section has a radio button selected, and the '相似搜索' (Similar search) section has a dropdown menu set to '0.8'. A red box highlights '选择化合物相似度' (Select compound similarity). The '子结构搜索' (Substructure search), '超结构搜索' (Superstructure search), and '配方搜索' (Recipe search) sections are also present. At the bottom of the search options panel, there is a '全部 国家' (All countries) section and a '搜索偏好设置' (Search preference settings) section. A red box highlights '选择想要检索的国家和地区' (Select the countries and regions you want to search). At the bottom right, there is a '在线客服' (Online customer service) button.

At the bottom of the interface, there are two buttons: '搜索专利' (Search patents) and '搜索化合物' (Search compounds).

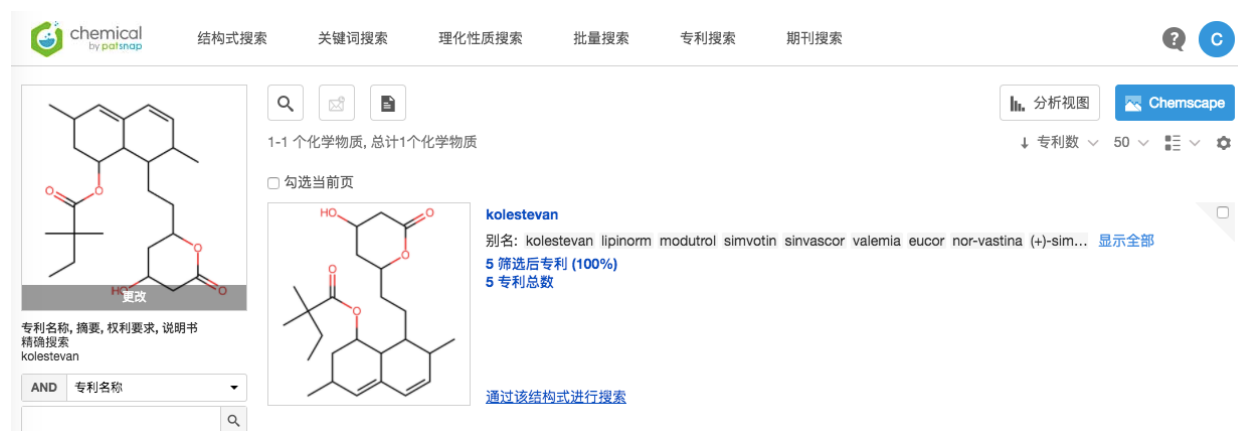
检索类型	目标结构是否包含提问结构	提问结构是否包含目标结构
Exact structure search 精确结构检索	是	是
Substructure search 子结构检索	是	
Superstructure search 超结构检索		是
Similarity structure search 相似结构检索	通过二进制指纹和Tanimoto参数来计算化学结构的相似度 (会从化学结构的空间伸展方向、电荷分布等各方面来对化学结构进行综合性比较)	

- 查询跟提问结构式完全一样的化合物



提问结构式

命中化学结构式



chemical by patsnap

结构式搜索 关键词搜索 理化性质搜索 批量搜索 专利搜索 期刊搜索

1-1 个化学物质, 总计1个化学物质

☐ 勾选当前页

专利名称, 摘要, 权利要求, 说明书
精确搜索
kolestevan

AND 专利名称

通过该结构式进行搜索

kolestevan
别名: kolestevan lipinorm modutrol simvotin sinvascor valemia eucor nor-vastina (+)-sim... 显示全部
5 筛选后专利 (100%)
5 专利总数

精确结构检索： 匹配条件的调整

patsnap

应用场景：

1. 同时搜索：

- 电荷不同的化合物
- 同位素不同的化合物
- 互变异构体和立体异构体

2. 在绘制结构时键的位置绘制不对，可以用精确结构找到相应的结构

搜索偏好设置

搜索设置

- ☒ 允许立体异构体
- ☒ 允许替代的原子质量 (同位素)
- ☒ 允许替代的原子电荷
- ☒ 允许互变 (异构) 体

结构类型

- ☐ 包括溶剂
- ☐ 包括无机化合物



全部 国家



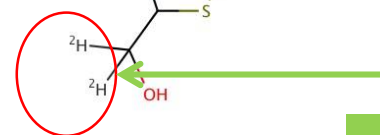
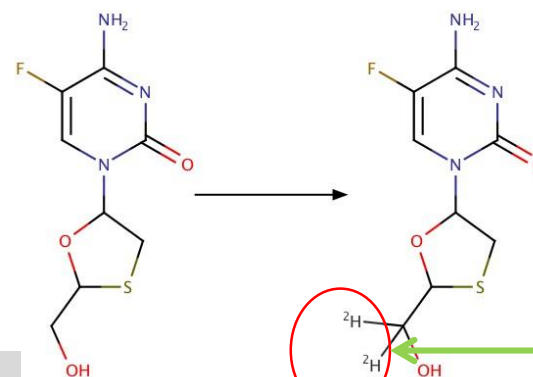
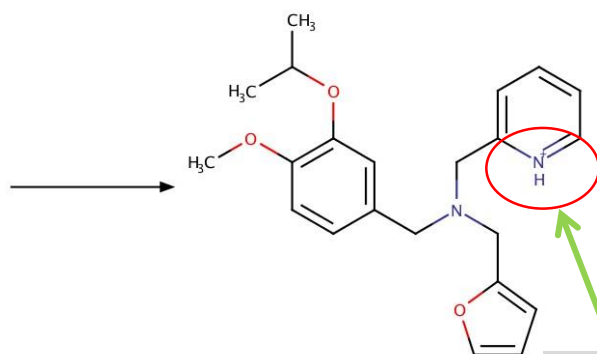
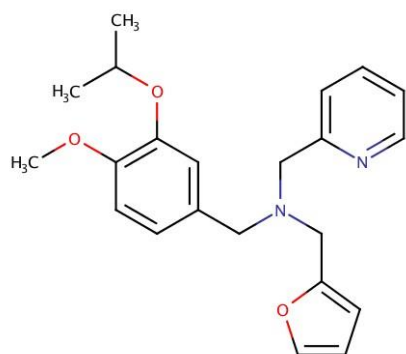
搜索偏好设置

互变异构

立体异构

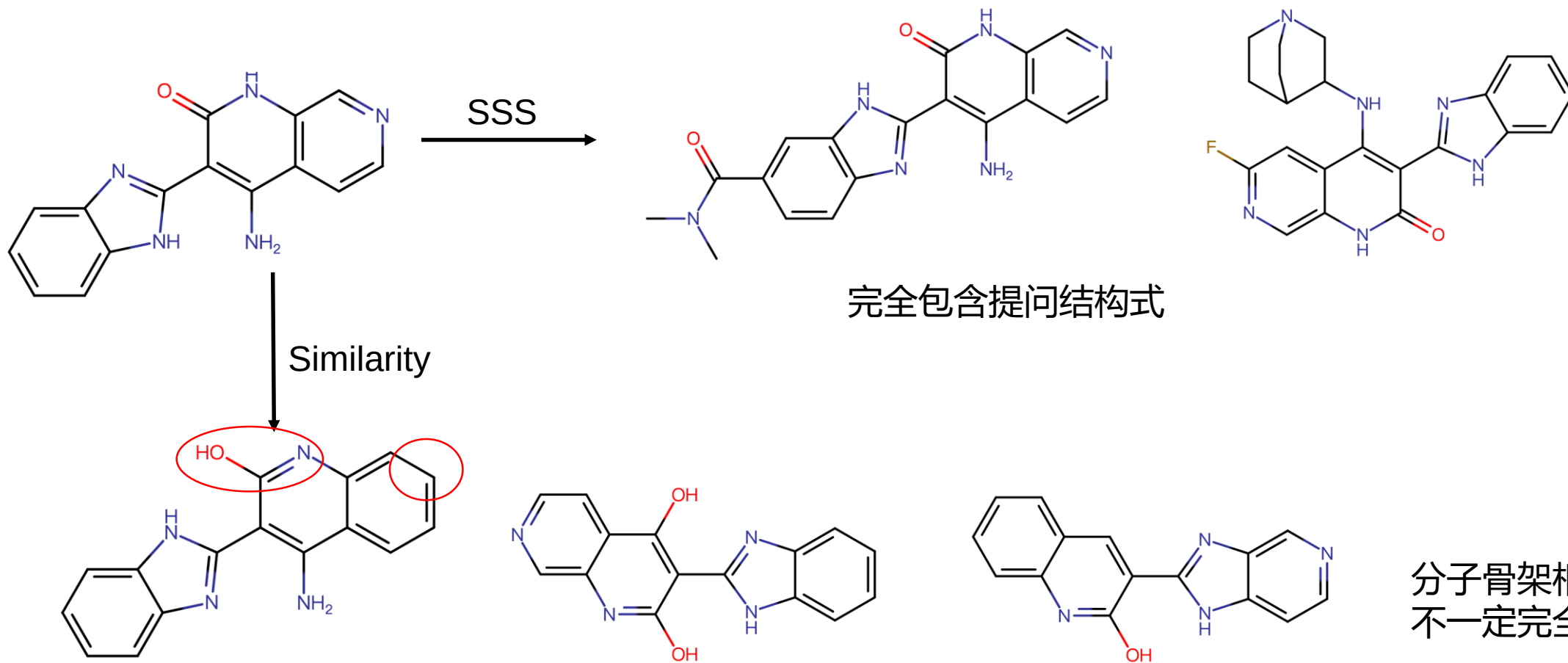
同位素变化

电荷变化



子结构检索(SSS) & 相似性检索(Similarity)

patsnap



案例：盐酸埃克替尼通过相似结构突破原研专利

分子结构	厄洛替尼 原研	埃克替尼 仿制
治疗癌症	非小细胞肺癌	非小细胞肺癌
制药企业	罗氏制药	浙江贝达药业
价格	4300元人民币-5000元人民币/每周	2700元人民币-3500元人民币/每周

以著名的肺癌分子靶向药盐酸埃克替尼为例：

跨国药企在药物研发时，出于专利保护的考虑，会尽可能地覆盖将一个结构的改造空间。埃克替尼找到了一个巧妙的侧链成环结构，整个分子结构相似，风险较低，但差异合理，突破了专利。

站在研发厄洛替尼的罗氏的角度看，埃克替尼的专利突破使其收入减少，这也是类似的拥有较强原研能力的药企的痛点。

子结构检索的应用场景	相似结构检索的应用场景
分子骨架相同，侧链稍有不同的分子结构很可能有同样的性能或功效，所以对于期刊或者专利中保护的化合物进行仿造的时候，可以通过具有同样分子骨架的结构式来进行	分子骨架相似的分子结构也很可能有同样的性能或功效，因此也可以通过相似分子结构式来仿造受保护的化合物
对于化学家来说保持分子骨架不变，只是对侧链进行修改，在工艺上比较容易实现	相对来说相似分子结构更容易找到创新点，突破专利的保护范围
但是专利文献通常通过化学结构通式对分子骨架的各种侧链进行保护，所以仿造的空间比较有限	但是分子骨架的变化，会导致原有分子结构的合成工艺没有太多的参考价值，需要研发新的合成工艺路线，因此对研发能力的要求比较高

组合结构式搜索的应用场景

patsnap

使用场景| :如何快速检索化学药物的制备方法专利或组合物专利？

一般方法如下：

Step 1 : 构建检索药物化合物结构式, “一键检索”所有专利；

Step 2 : 进入专利数据库, 确定包含该药物化合物的所有专利；

Step 3 : 按照申请日从旧到新排序所有专利, 查阅公开药物化合物的专利；

Step 4 : 选择制备方法专利, 阅读全文, 化学高亮重点制备方法信息, 确定可行的反应原料；

Step 5 : 导入反应原料和药物化合物结构, 在化合物数据库中进行多个化学结构组合检索；

Step 6 : 进入专利数据库, 查阅所有相关的制备方法专利。

The screenshot shows the patsnap chemical search interface. At the top, there's a navigation bar with 'chemical by patsnap', '搜索' (Search), 'LabBench', and '邮件提醒' (Email Reminder). Below this is a sub-navigation bar with '结构式搜索' (Structure Search), '组合结构式搜索' (Combination Structure Search), '性质搜索' (Property Search), and '批量搜索' (Batch Search). The '组合结构式搜索' tab is selected.

Below the navigation bar, there's a text box: '添加多个结构进行相似/子结构/配方搜索；或者搜索包含所添加结构的专利。' (Add multiple structures for similarity/substructure/formula search; or search for patents containing the added structures).

The main search area shows two chemical structures being added to a search query. The first structure is a complex organic molecule, and the second is a simpler molecule. They are separated by an 'AND' operator. Annotations with red boxes and arrows point to these elements:

- '点击添加结构式 (最多五个)' (Click to add structure (up to five)) points to the first structure.
- '根据应用场景使用布尔运算符' (Use Boolean operators according to the application scenario) points to the 'AND' operator.
- '点击添加结构' (Click to add structure) points to the second structure.

Below the search area, there's a row of filters: '全部 国家' (All Countries), '专利名称/摘要/权利要求/说明书' (Patent Name/Abstract/Claims/Description), '+ 添加关键词' (+ Add Keywords), and '搜索偏好设置' (Search Preference Settings).

At the bottom, there's a row of search buttons: '搜索专利' (Search Patents), '搜索配方' (Search Formulas), and '搜索化合物' (Search Compounds). Annotations with red boxes and arrows point to these buttons:

- '通过多个化合物限定检索专利, 结果更准确。(组合物专利, 工艺专利等)' (Limit patent search by multiple compounds, results are more accurate. (Combination patents, process patents, etc.)) points to the '搜索专利' button.
- '一键搜索包含查询化合物的专利中的主要化合物' (One-click search for main compounds in patents containing the query compound) points to the '搜索配方' button.
- '搜索两个化合物相似/子结构/超结构集中的公共化合物' (Search for common compounds in the similarity/substructure/superstructure set of two compounds) points to the '搜索化合物' button.

通过InchiKey (国际化合物标识) 进行化合物结构批量检索

chemical
by patsnap

搜索 LabBench 邮件提醒

?

Di

结构式搜索 组合结构式搜索 性质搜索 批量搜索

同时输入多个化合物InchiKey, 批量检索

请输入化学物质的InchiKey

① 支持在此输入多个InchiKey进行查询, 每个InchiKey之间请使用空格、换行或逗号进行分隔。

开始匹配

应用场景:当检索时结构, 如果遇到化合物多种固定结构统一检索时。在用子结构, 相似结构检索时, 数据较多不容易找到目标结构, 例如某一化合物检索时, 盐, 水溶剂等形式, 可以连同相应化合物统一检索。

性质搜索的应用场景

「专利+药品及化合物信息」多维度精确定位目标化合物

chemical by patSnap 搜索 LabBench 邮件提醒

结构式搜索 组合结构式搜索 **性质搜索** 批量搜索

专利信息+化合物结构信息+药品信息联合检索，缩小检索范围使结果更准确。

专利信息

专利名称/摘要/权利要求/说明书

当前申请(专利权)人

申请日

IPC 分类号

化合物结构理化信息

结构特征

分子式

酸碱性

分子量

药物(化合物)各国审批信息

批准特征

USFDA 申请公司/机构

USFDA 批准日期

USFDA 当前状态

药物临床信息

临床试验性能

主要责任人

首次收到日期

Structure properties	molecular formula
	molecular weight
	molecular species
	hydrogen bond donor count
	hydrogen bond acceptor count
	rotatable bond count
	exact mass
	monoisotopic mass
	topological polar surface area
	xlogP3
	heavy atom count
	isotope atom count
	defined atom stereocenter count
	undefined atom stereocenter count
	defined bond stereocenter count
	undefined bond stereocenter count
	formal charge
	complexity
	ACD Basic pKa
	ACD Acidic pKa
专利可检索字段	关键词 标题/摘要/权利要求/说明书
	申请人/发明人/专利权人
	CPC/IPC分类号
	申请日/公开日
药物批准信息可检索字段	申请人地址
	申请公司/机构/商品名/适应症/中文名称
	批准日期/橙皮书专利到期日期/非专利保护到期日/NMPA有效期等
	FDA当前状态/EMA是否仿制药/NMPA省市等
ClinicalTrials.gov临床试验数据可检索字段	主要责任人/适应症/主要测量指标
	首次收到日期/最后更新日期
	研究对象年龄/受试者
	临床研究阶段/临床试验状态/性别/研究类型

▲ 多维度【可检索字段】示意

调研抗病毒研究的重点化合物 (往期课程介绍)

patSnap

方法一：通过查找专利提取重要化合物，对于化合物进行调研分析



方法二：通过化学数据库的信息检索所有的化合物



方法三：通过化学数据库的临床信息的适应症，ATC代码解剖学上的分类等等来进行调研

TACD_ALL:(冠状病毒 or 具外套膜 \$WS 正链单股RNA or SARS病毒) or A61P31/12(抗病毒剂)

GILEAD

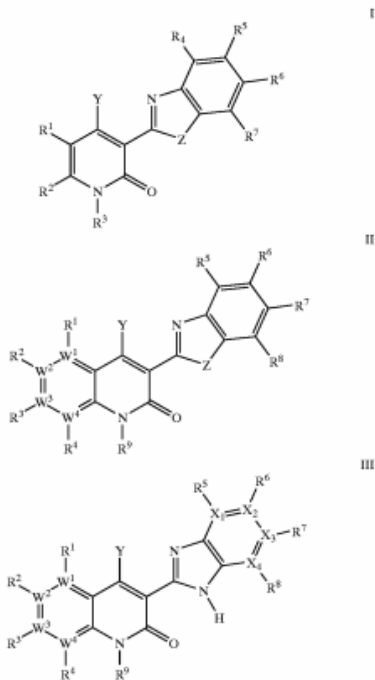
适应症：Ebola Virus , HIV Infections, Severe Acute Respiratory Syndrome

analytics.zhihuiya.com

(57)

ABSTRACT

Organic compounds having the structural formulas I, II, and III are provided where the variables have the values described herein and R¹ and R² in structure I join together to form a 5 to 7 membered substituted or unsubstituted ring including at least one O, N, or S atom, and Z is an O, S, NH or NR group in structures I and II.



Pharmaceutical formulations include the organic compound or a pharmaceutically acceptable salt thereof and a pharmaceutically acceptable carrier. A method of treating a patient includes administering a pharmaceutical formulation according to the invention to a patient in need thereof.

- 在阅读专利原文时，可以通过“化学高亮”功能将原文中的化合物名称全部高亮显示，并且把鼠标移动到高亮文本时，对应的化学结构式就会从页面悬浮显示出来；文献阅读过程因此会变得十分容易。

US6759417 Heterocyclic Compounds

专利详情 图文对照 引用信息 同族专利 法律信息

化学高亮 添加到工作空间 引用分析

29 附图

Example 1

3-[2-(Methoxycarbonyl)acetylaminopyridine-4-carboxylic acid]

A solution of 3-aminopyridine-4-carboxylic acid (1.0 eq), methyl 2-(chlorocarbonyl)acetate (1.1 eq), and triethylamine (2.0 eq) in acetone was stirred overnight at room temperature. The solvent was removed in vacuo. The product was used without further purification. LC/MS m/z 239.2 (MH⁺), R.sub.t 1.40 minutes.

3-Benzimidazol-2-yl-4-hydroxyhydropyridino[3,4-b]pyridin-2-one

3-[2-(Methoxycarbonyl)acetylaminopyridine-4-carboxylic acid] (1.1 eq) was combined with 1,2-phenylenediamine (1.0 eq) and heated at 150.degree. C. for 3 hours. The crude product was purified by reversed-phase HPLC (DMSO/5% TFA). LC/MS m/z 279.3 (MH⁺), R.sub.t 1.73 minutes.

Example 2

4-Hydroxy-3-[5-methylbenzimidazol-2-yl]hydropyridino[3,4-b]pyridin-2-one

The title compound was synthesized as described in Example 1 using 3-phenylenediamine. The crude product was purified by reversed-phase HPLC.

Example 3

Ethyl 2-benzimidazol-2-ylacetate

A solution of 1,2-phenylenediamine (1.0 eq) and ethyl 3-ethoxy-3-imino-2-oxopropionate (1.0 eq) in THF at -78.degree. C. After 1 hour, 8-amino-4-cyanopyridine (0.8 eq) in THF was added. The resulting mixture was allowed to warm to room temperature overnight. The mixture was quenched with NH.sub.4Cl (aqueous saturated solution) and extracted with EtOAc. The organic layer washed with EtOAc and brine, dried over Na.sub.2SO.sub.4, filtered, and concentrated in vacuo to yield a brown solid. The crude material was purified by silica gel chromatography (5:1 EtOAc:hexane) to yield the desired product. LC/MS m/z 205.2 (MH⁺), R.sub.t 1.44 minutes.

Method A

2-Benzimidazol-2-yl-N-(4-cyano(3-pyridyl))acetamide

LIHMDS (2.5 eq) was added to ethyl 2-benzimidazol-2-ylacetate (1.0 eq) in THF at -78.degree. C. After 1 hour, 8-amino-4-cyanopyridine (0.8 eq) in THF was added. The resulting mixture was allowed to warm to room temperature overnight. The mixture was quenched with NH.sub.4Cl (aqueous saturated solution) and extracted with EtOAc. The organic layer washed with EtOAc and brine, dried over Na.sub.2SO.sub.4, filtered, and concentrated in vacuo to yield a brown solid. The crude material was purified by silica gel chromatography (5:1 EtOAc:hexane) to yield the desired product. LC/MS m/z 278.3 (MH⁺), R.sub.t 1.88 minutes.

4-Amino-3-benzimidazol-2-ylhydropyridino[3,4-b]pyridin-2-one

Chemical structure of 4-hydroxy-3-[5-methylbenzimidazol-2-yl]hydropyridino[3,4-b]pyridin-2-one. The structure shows a pyridine ring substituted with a methyl group at position 5 and a hydroxy group at position 4, and a Y-Z linkage at position 3. The Y-Z linkage is a benzimidazole ring system.

用户可以一键提取这篇专利中的所有化合物

- 打开其中一篇专利的详情页，在图文对照页面可以看到专利摘要部分的Markush结构通式，了解到此专利要保护的主要化学空间；

Chemical化合物二次提取

patsnap

从专利中批量提取化合物进行二次检索或导出

patSnap

搜索

工作空间

(INCHI_TITLE:(79487526) OR INCHI_ABST:(79487526) OR INCHI_CLAIMS:(79487526) OR ~ 88.6K)

帮助中心

AND 过滤关键词

共 88,597 条专利 | 分析

导出 设置邮件提醒 保存到工作空间

切换视图 复制检索式 保存检索式 申请日从旧到新

公开(公告)号	标题	申请日	优先权	扩展同族成员数量	[标]当前申请(专利权人)
☒ US4020049	制备聚酯树脂的方法	1971-04-02		10	GOODYEAR TIRE & RUBBER COMPANY(固特异)
☒ US3998100	锻炼过程和设备	1975-04-21		1	PIZATELLA ROBERT JOHNSON E A
☒ US4216141	肽的环化方法	1978-07-19		1	SALK INST FOR BIOLOGICAL STUDIES(萨克生物)
☒ EP0024224B1	合成1,1,2,2-四氢全氟醇及其酯的方法	1980-07-25			PROD CHIMS CHIMIE KUHLMANN
☒ US4328212	新型抗生素NCS-C及其制备方法	1980-07-28			KAYAKU ANTIBIOTICS
☒ EP0037723A2	激素基因组克隆的表达	1981-04-03			JN... Y OF

过滤项

申请(专利权)人

[标]当前申请(专利权)人

☐ MERCK 7503

☐ BRISTOL-MYERS SQUIBB 2505

☐ NOVARTIS 1862

☐ TAKEDA PHARMACEUTICAL COMPANY 1793

☐ PROCTER & GAMBLE 1375

☐ ASTRAZENECA 1189

☐ BOEHRINGER INGELHEIM INT 1158

☐ NOVOZYMES 1130

☐ PFIZER RES & DEV N 1123

☐ SANOFI 912

更多

自定义申请人组

管理申请人组

☐ 默沙东制药 7982

分类号

申请年

[标]当前申请(专利权)人类型

受理局

专利类型

国民经济行业分类号

100 条已勾选

导出

保存到工作空间

提取化学结构式

提取基因序列

搜索同族专利

取消全部勾选

多篇专利化合物提取

02

PART

化学检索画图工具应用

- 多种画图软件在智慧芽做图工具有兼容性
- 环状结构任意位置的检索
- 同族基团和广义原子简化Markush化学结构的检索
- R基团固定多类别取值检索

用户场景:

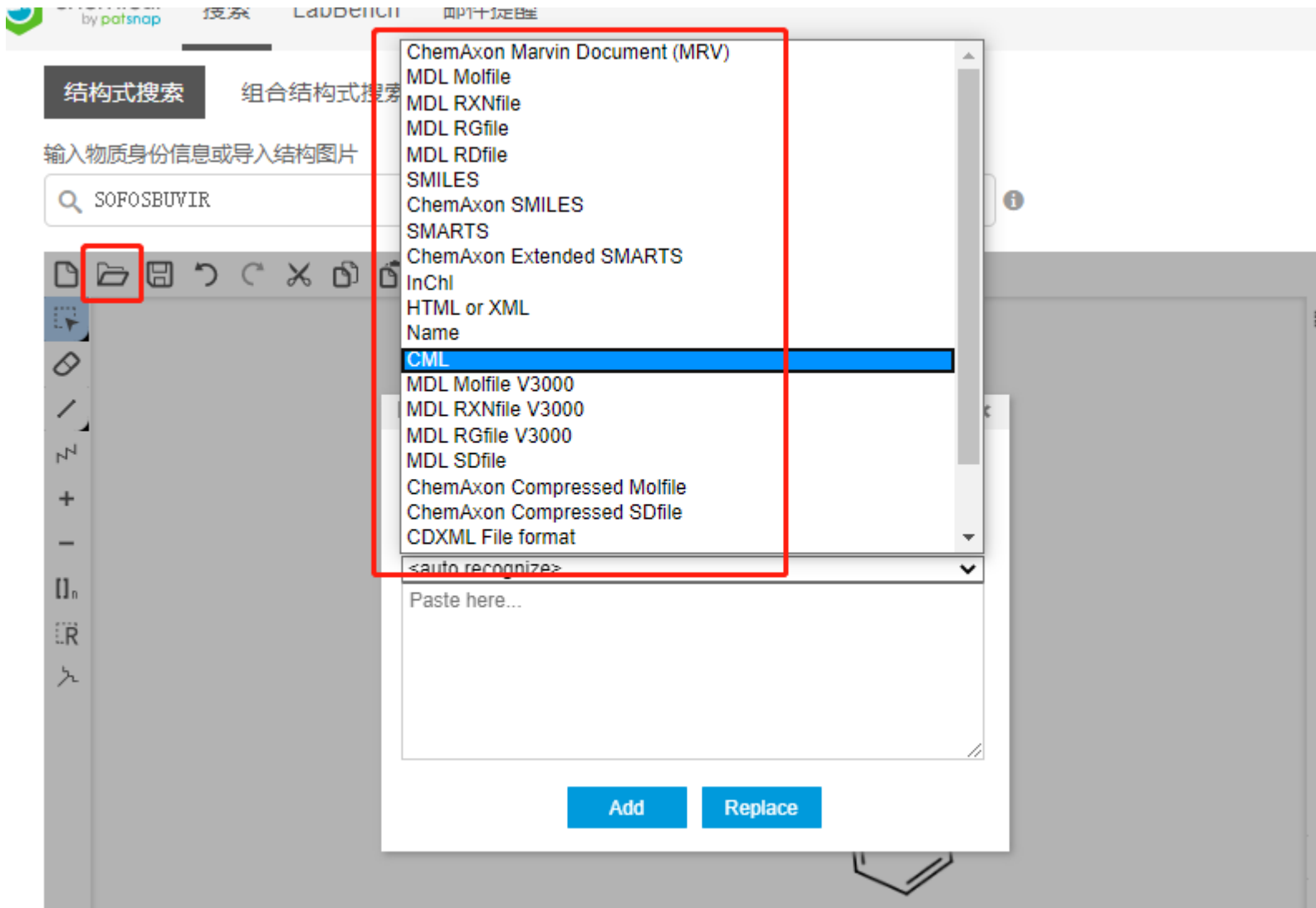
通过智慧芽的做图工具Marvin JS, 做出来的图可以保存下来, (MDL Molfile, SMARTS, INCHLKEY, PNG, JPEG, CML等) 下次检索时可以直接把文件或图片直接上传检索。其他工具的格式也支持兼容, 具体见结构保存相关格式

用户痛点:

多种画图软件及图片文本格式在画图工具上的兼容性, 画出来的结构无法保存, 需要每次重复进入

多种画图软件在智慧芽做图工具有兼容性

patSnap



上传结构文档支持多种格式

多种画图软件在智慧芽做图工具有兼容性

patSnap



- 1.支持输入化合物名称（包括中文名称），分子式，各种号码格式
- 2.支持上传图片，镜面匹配

多种画图软件在智慧芽做图工具有兼容性

patsnap

chemical.zhihuiya.com/search/#/

chemical by patsnap 搜索 LabBench 邮件提醒

结构式搜索 组合结构式搜索 性质搜索 批量搜索

输入物质身份信息或导入结构图片

+ 支持物质名称/分子式/SMILES等身份信息

Export

Structure Image

Format PNG Zoom mode PNG Background transparent Width 200 Height 200

☒ Show chiral flag ☒ Show valence errors ☐ Show lone pairs ☐ Index atoms ☒ Show atom maps ☒ Use CPK colors ☐ Show carbon labels

Implicit H Hetero and Terminal Display wireframe

Download

Export

Structure Image

Format ChemAxon Marvin Document (MRV) <cm1><MDOcument

MDL Molfile MDL RXNfile MDL RGfile MDL RDfile SMILES ChemAxon SMILES SMARTS ChemAxon Extended SMARTS InChI InChIKey Name CML MDL Molfile V3000 MDL RXNfile V3000 MDL RGfile V3000 MDL SDfile MDL SDfile (CTAB v2000) ChemAxon Compressed Molfile ChemAxon Compressed SDfile

Use Ctrl+

支持保存图片及其他格式文件

用户场景:

当环状结构上连接邻位，对位，间位，都是目标检索结构

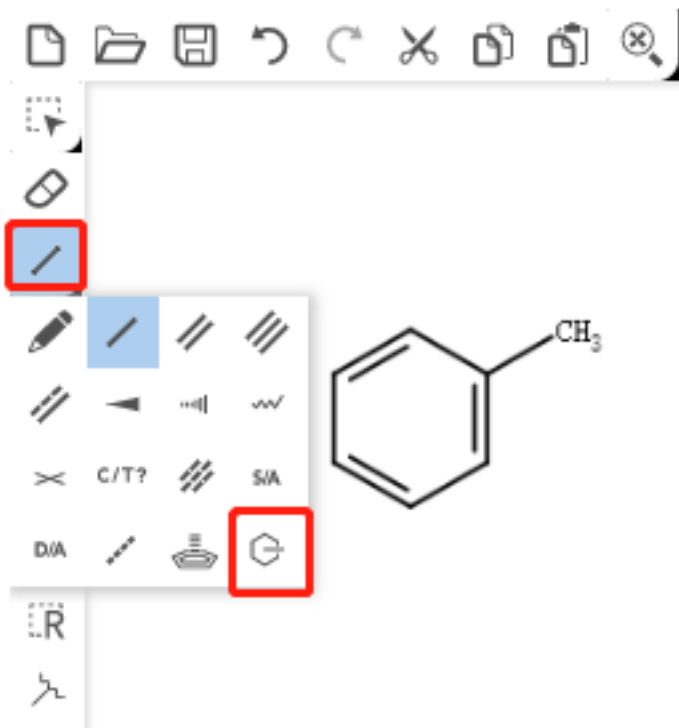
如甲苯上可能想要绘制对二甲苯，邻二甲苯，间二甲苯时，需要分别绘制三次，如果涉及长链结构，结构的异构不同，可能需要绘制更多结构的表达，为了简化此工作，智慧芽化学结构中设置了任意位点绘图。

用户痛点:

在描述每个环状结构上的取代基的时候，需要绘制出每个可能的位点，这往往要花一些时间，而且某些情况下用户不能够枚举出所有想要的位点。

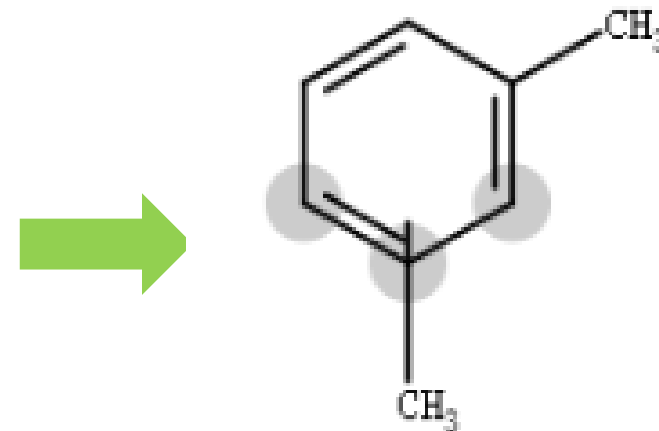
环状结构任意位置的检索

patsnap



1. 在绘图区左侧选定化合键位置工具

2. 鼠标框选需要检索的化合物可能存在的区域



3. 绘制完成

用户场景：

子结构检索是分子骨架相同，侧链稍有不同的分子结构，但是没有位置的限制，有时对于定向结构R基结构变化检索的话可以通过Marvin JS绘图工具可以绘制带有同族（广义）基团和原子的Markush提问结构式，并进行精确结构或者子结构搜索

因此我们将为您实现通过同族基团（也称为广义基团）和广义原子来定义Markush提问结构，并进行结构式搜索的功能。例如图示的提问结构中Alkyl表示具有如下特征的取代基：

- 1, 最少一个碳原子；
- 2, 只包含碳原子和氢原子；
- 3, 只包含单键；
- 4, 不包含环键。

我们将支持数十种类似的广义基团和原子来进行结构式搜索。为了帮助他们轻松的跟踪相关化合物的更新情况，我们提供了邮件提醒的功能，点击几下鼠标，用户便可以通过定期的邮件提醒来跟踪他们感兴趣的化合物的相关动态

用户痛点：

在描述每个R基团的时候，需要绘制出每个可能的取代基，这往往要花一些时间，而且某些情况下用户不能够枚举出所有想要的取代基。

通过同族基团和广义原子简化Markush化学结构的检索

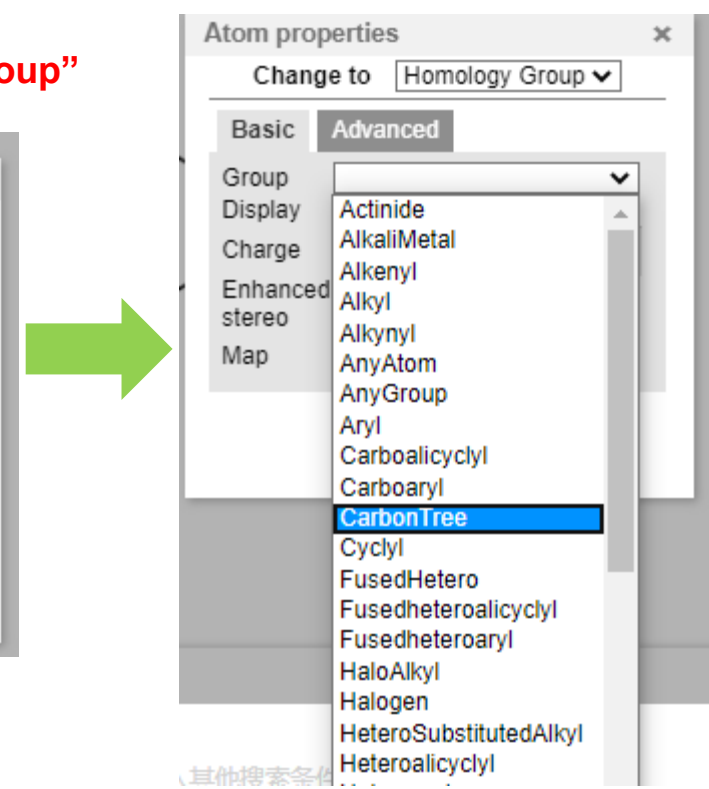
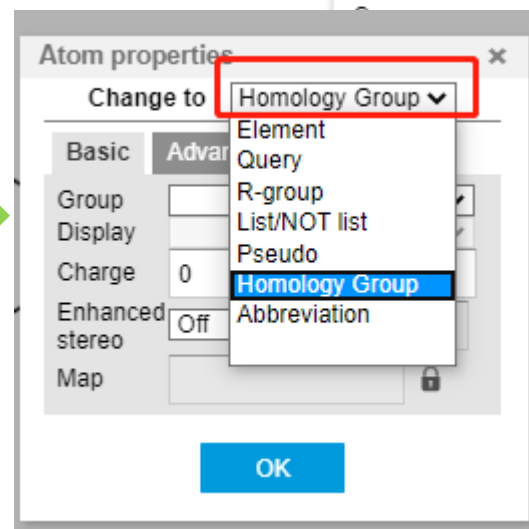
patSnap

输入物质身份信息或导入结构图片

🔍 + 支持物质名称/分子式/SMILES等身份信息



3. 选择Change to "Homology Group"



4. Group Display选择同族基团和广义原子简化Markush化学结构，绘制完成

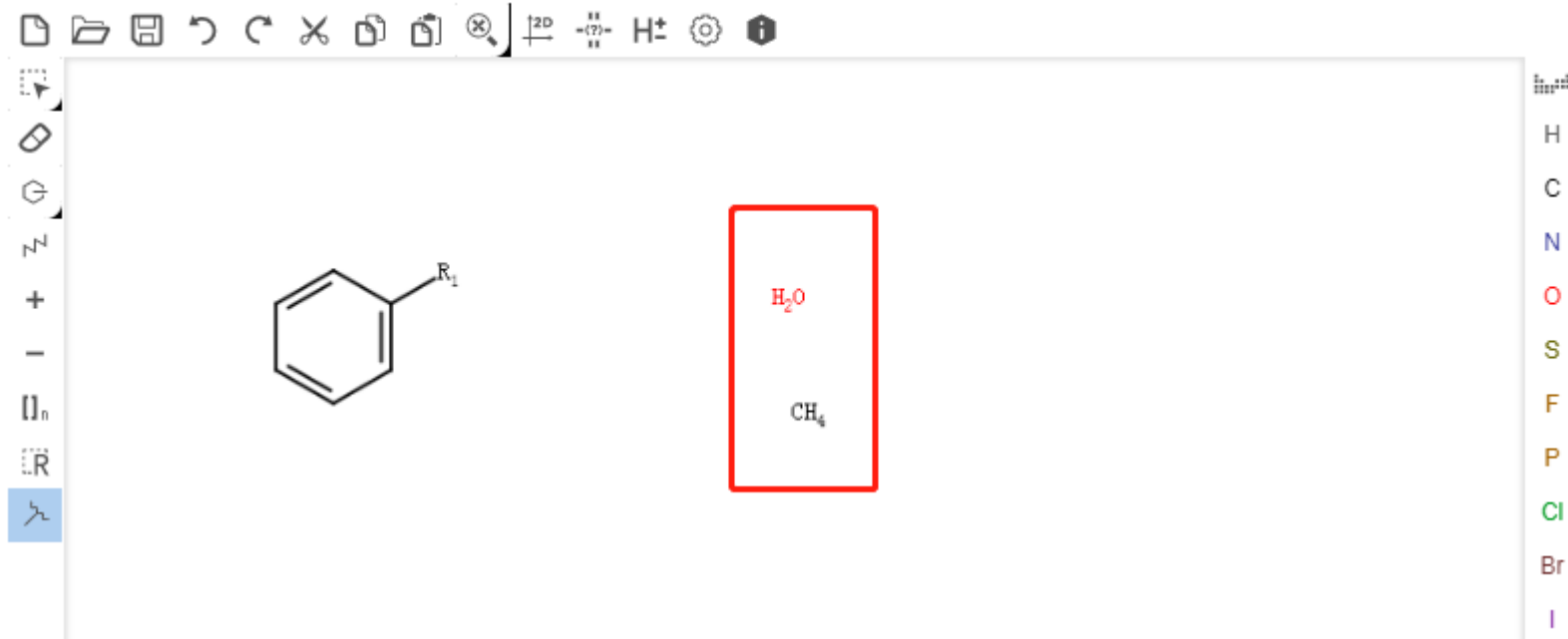
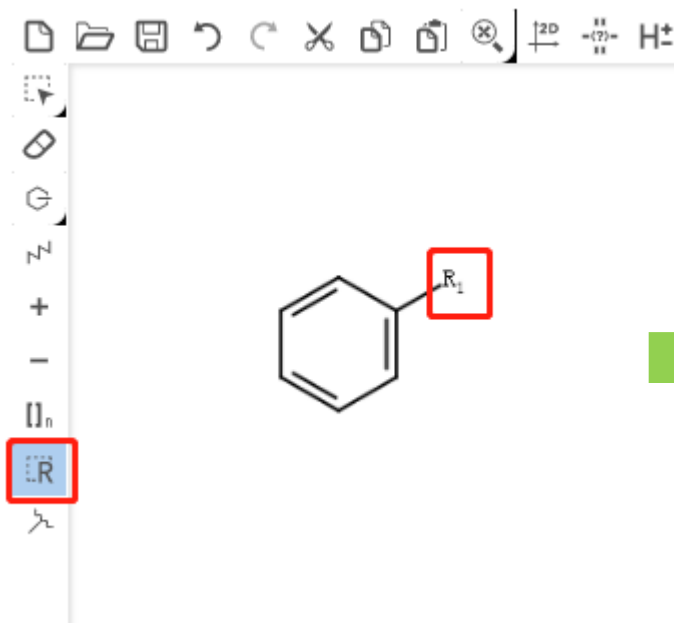
用户场景：

当R基可以为多个取代基的一种，都是目标检索结构

如母链结构上，如果想要某一固定结构被OH,NH₂,SH, C₂H₅等等多设计的结构。都可以一次性检索。而智慧芽子结构，并不能限定位置，也不能限定结构的类型，所以用R基固定位置，固定设计结构，可以达到精准检索的需求

用户痛点：

在描述每个R基可以为多个取代基的一种的时候，需要绘制出每个可能的取代基，这往往要花一些时间，而且某些情况下用户不能够枚举出所有想要的取代基。

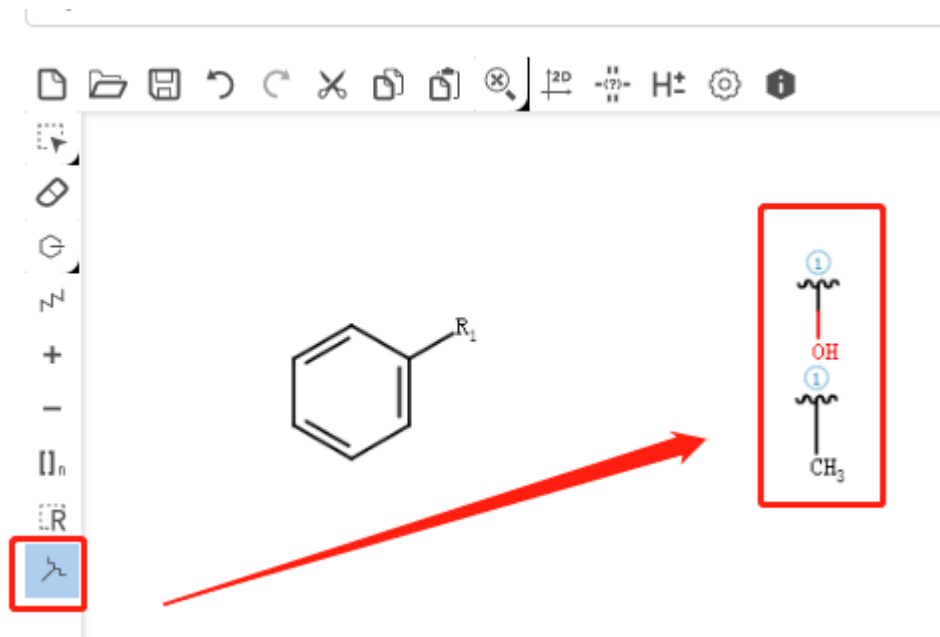


1.在绘图区左侧选择R结构，绘制在结构上某一或多个位点上

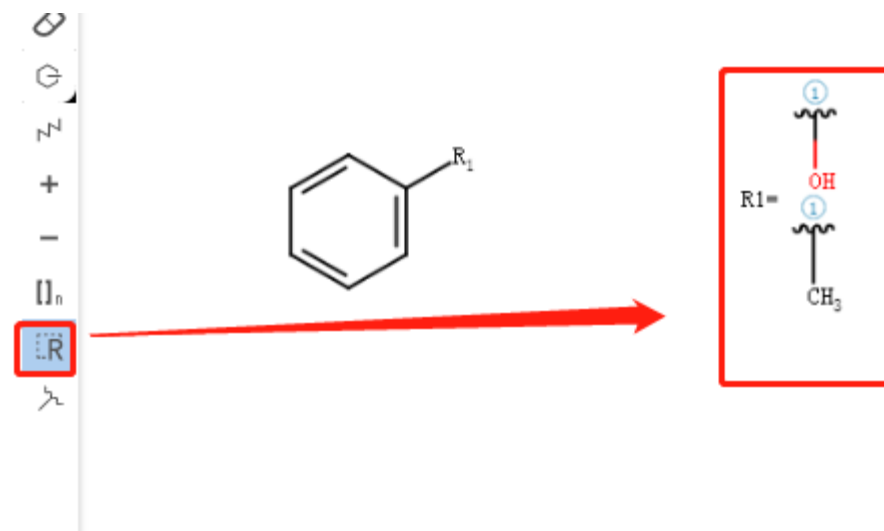
2.绘制出可能R1存在的取代基



下一页



3.  R基团连接绘制在取代基上



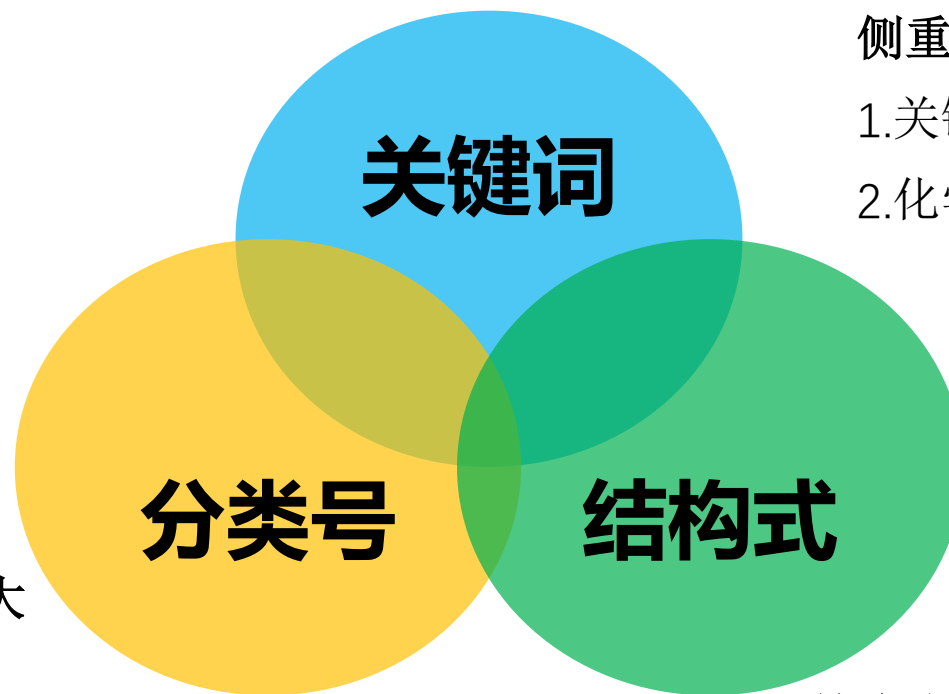
3.  R结构选取取代基，绘制完成

化学与专利数据库的巧妙联合搜索

03

PART

- 检索式构建(实现查全, 多类技术信息运算, 去噪音)
- 快速确定原研药公司专利布局策略
- 风险排查巧查全



侧重于检索准确，可能结果不全

1. 关键词无法完全罗列
2. 化学结构无法表达

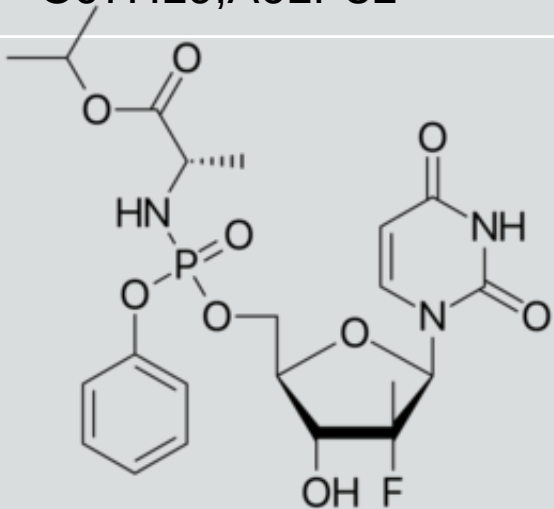
侧重于检索全面，可能噪音较大

1. IPC每五年修订一次
2. 化学领域IPC分类比较特殊，如限定结构里的双键含量等。

检索全面，兼顾全面和检索准确

1. 化学别名，商品名，代码
2. 马库什结构，具体结构, OCR识别图片
3. 化合物其他信息：临床，审批，工艺等

构建检索式和选择检索策略（往期课程介绍）

	检索要素1:索非布韦	检索要素2: 丙肝治疗
关键词拓展	索非布韦 or 索氟布韦 or 索磷布韦 or (嘧啶 \$WS 呋喃 \$WS 磷酸基 \$WS 氨基 \$WS 丙酸 \$WS 酯) (同时可以利用关键词助手及机器翻译 检索)	AND (丙肝 and 抑制剂) or (丙 型肝炎 and 抑制剂) or (HCV and 抑制剂)
关键词 (中日英德法)	Sovaldi or Sofosbuvir	
IPC	C07H19,A61P31	A61P31
化学结构式		

马库什化合物、合成方法、晶型，制剂等专利如何查找？

(往期课程介绍)

patSnap

	马库什	合成方法	晶型	制剂
关键词	检索结果 页面限制 权力要求:CLMS:(R1)	限制关键词为方法，工艺，步骤等 例：TA:(方法 OR 工艺 OR 步骤 OR 制备 OR 合成)	限制晶体关键词或达到的效果。 例：TA:(晶型 OR "晶形" OR "多晶" OR "混晶" OR 单晶 or 稳定性 or 生物利用度 or 疗效)	限制剂的关键词 例：片剂，丸剂或针剂等
分类号	-	IPC:C07C69/54...丙烯酸酯；甲基丙烯酸酯		A61K9/00以特殊物理形状为特征的医药配制品 A61K9/20·丸剂、锭剂或片剂 A61K9/52持续释放型或间断释放型
绘图区及相应数据	通过结构 绘制区画出对应的 结构	1.化合物内合成工艺信息 2.组合结构搜索 例：制备由乙烯制备甲基丙烯酸酯：1.在组合结构搜索里限定原料及合成产物，也可以限制合成用到的催化剂，或排除一些无关的化合物		
权利要求类型	✓	✓		

检索式构建 (实现查全, 多类技术信息运算, 去噪音)

patsnap

检索式: (结构式检索结果 or 关键词检索结果) and 制剂

图例中编辑组合式检索式: (S1 OR S2) AND S3

简单搜索

高级搜索

批量处理

语义搜索

图像搜索

扩展搜索

分类号搜索

专利类型

发明申请

发明专利

实用新型

外观设计

全部数据库 (116/116)

知识产权五局

美国 (US)

中国 (CN)

欧洲 (EP)

日本 (JP)

韩国 (KR)

主要国家/地区

世界知识...

奥地利 (AT)

澳大利亚 (AU)

字段搜索

标题/摘要/权利要求

片剂 or 丸剂 or 针

ANC

[标]当前申请(专利权)人

例如: 华为 OR 中

OR

IPC分类号

A61K9

ANC

主要字段

可输入关键词, 公

ANC

主要字段

可输入关键词, 公

+ 添加字段

检索式中的逻辑运算优先顺序为NOT>AND>OR, 您也可以

TAC_ALL:(片剂 or 丸剂 or 针剂 OR "制剂" OR "注射剂" OR

编辑检索式

编辑组合检索式

当前检索共 4.14M 条专利

简单搜索

高级搜索

批量处理

语义搜索

图像搜索

扩展搜索

分类号搜索

法律搜索

化学搜

切换到字段搜索

LED封装... x

项目 3 x

培训1 x

项目 2 x

冠状病毒-... x

LDM x

索非布韦 x

AND | OR | NOT | \$W | \$PRE | \$WS | [] | " " | ()

编号 检索式

S1 (INCHI_TITLE:(63829444) OR INCHI_ABST:(63829444) OR INCHI_CLAIMS:(63829444) OR

S2 TAC_ALL:(索非布韦 or 索磷布韦 or 索磷布韦 or (嘧啶 \$WS 咪唑 \$WS 磷酸基 \$WS 氨基 \$WS 丙酸 \$WS

S3 TAC_ALL:(片剂 or 丸剂 or 针剂 OR "制剂" OR "注射剂" OR "固体制剂" OR "口服") OR IPC:(A61K9)

提示: 可以使用'S1'代替检索式以及进行检索式合并, 例如'S1 and (S2 or S3)'

清空

化学结构式检索结果

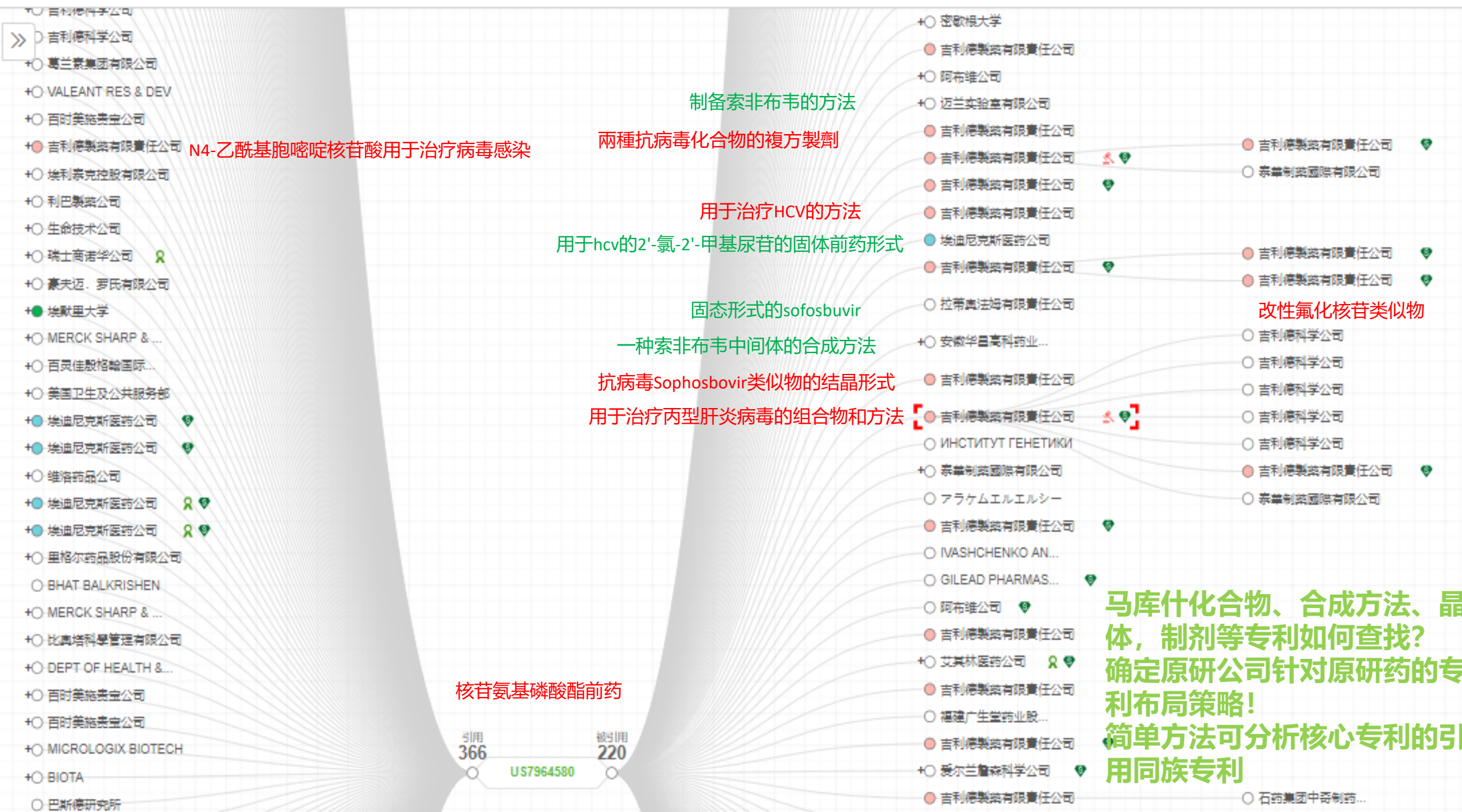
关键词检索结果

检索要素2: 制剂

搜索包含机器翻译数据

清空

搜索



SG

US

AU

PT

2007

2008

2009

2011

优先权

树状图关系清晰明了，快速分析企业的全球布局策略更加便捷

接续案

分案

PatSnap同族是由智慧芽自主深度加工的同族数据，覆盖全球数据库，其同族归类规则为：专利之间直接或间接有至少一个相同优先权，并且包含专利的所有分案、PCT申请、接续案。

快速确定原研药公司专利布局策略

natsnap

AND 过滤关键词

已选过滤项

AND [标]当前申请(专利权)人×
吉利德製藥有限责任公司
吉利德科学公司

过滤项

申请(专利权)人

[标]当前申请(专利权)人

☒ 吉利德科学公司 28
☒ 吉利德製藥有限責任公司 20
☐ 墨席业有限公司 2
☐ BJORNSON KYLA 1
☐ MISHRA RUCHIKA 1
☐ BONDY STEVEN S 1
☐ TRENKLE JAMES D 1
☐ CHOU CHIEN HUNG 1
☐ 美商基利法瑪席特有限公司 1
☐ VIVIAN RANDALL W 1

更多

分类号

IPC CPC LOC UPC

IPC小组

☐ A61K31/7072 14
.....含有两个.....
☐ A61P31/14 14
.....用于RNA病毒的[7]
☐ A61K45/06 13
.....尤化学特性之有效成.....
☐ A61K31/4188 11
.....与杂环耦合.....
☐ A61K31/519 8

48 组PatSnap同族 (共907条)

分析 3D专利地图分析 英策 (生成报告)

切换视图

复制检索式 保存检索式

2.选择去重方式为patsnap同族去重

3.申请日从旧到新

申请日从旧到新

导出 设置邮件提醒 保存到工作空间

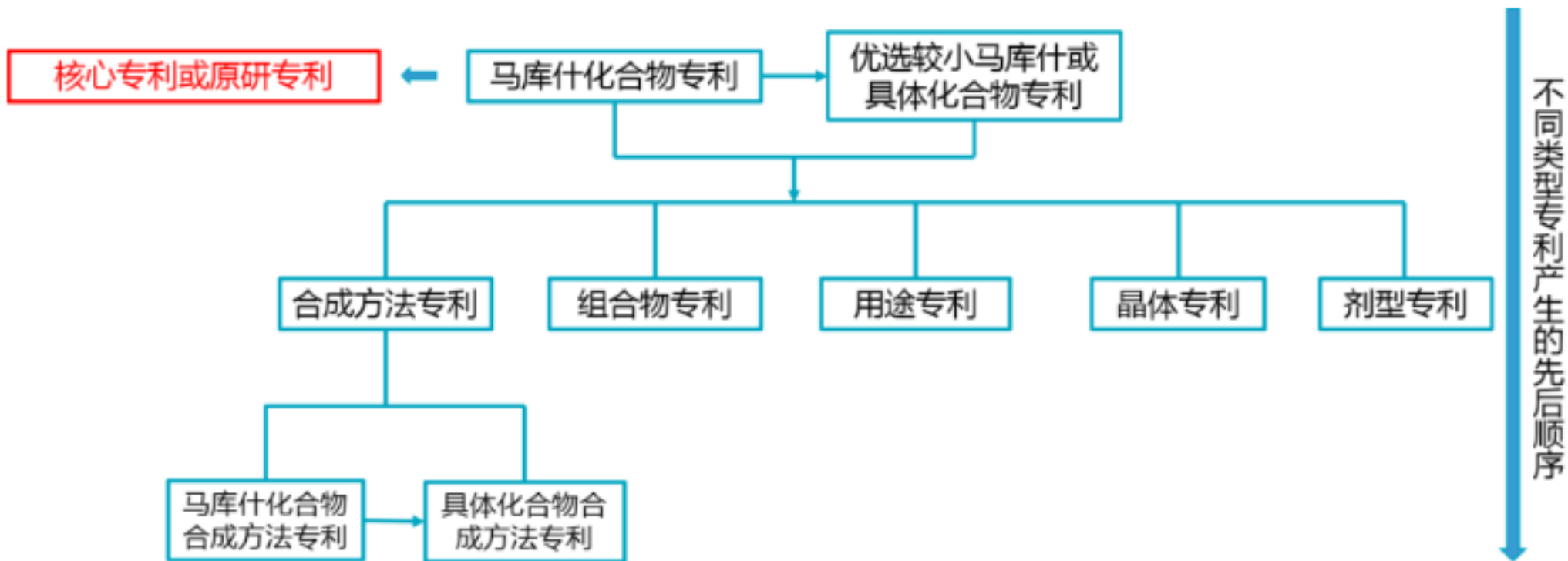
公开(公告)号	标题	申请日	原始申请(专利权)人	当前申请(专利权)人	公开(公告)日	专利价值	法律状态/事件
CN1816558A	修饰的氟化核苷类似物	2004-04-21	法莫赛特股份有限公司	吉利德制药有限责任公司	2006-08-09	\$3,520,000	授权 无效程序 口头审理
CN109776637A	核苷氨基磷酸酯前药	2008-03-26	吉利德制药有限责任公司	吉利德制药有限责任公司	2019-05-21	\$3,660,000	实质审查
TW201902477A	抗病毒化合物	2010-05-12	美商基利法瑪席特有限公司 GILEAD PHARMASSET LLC	美商基利法瑪席特有限公司(GILEAD PHARMASSET LLC)	2019-01-16	\$4,360,000	授权
ES2648803T3	生产sofosbuvir的方法	2010-05-20	GILEAD PHARMASSET LLC	GILEAD PHARMASSET LLC	2018-01-08	\$4,630,000	授权
WO2011146817A1	杂环黄病毒科病毒抑制剂	2011-05-20	GILEAD SCIENCES, INC. BJORNSON, KYLA BONDY, STEVEN, S. +6	GILEAD SCIENCES, INC. BJORNSON, KYLA BONDY, STEVEN, S. +6	2011-11-24	-	PCT进入指定国
US9090642	制备非对映体纯的氨基磷酸酯前药的方法	2011-07-19	CHO, AESOP CLARKE, MICHAEL O'NEIL HANRAHAN WOLCKENHAUER, SCOTT ALAN	GILEAD SCIENCES, INC.	2015-07-28	\$1,940,000	授权 权利转移

了解专利布局

在线帮助

快速确定原研药公司专利布局策略







(((INCHI_TITLE:(171170827 OR 171410645 OR 163201903 OR 172574819 OR



AND 过滤关键词



1,064 组PatSnap同族 (共4,518条)

分析



设置邮件

切换视图



复制检索式

保存检索式



公开(公告)号

标题

申请日

1

CN1816558A

修饰的氟化核苷类似物

2004-04-21

如果做印度专利风险排查工作是只通过受理局查查询吗？下页PPT，我们来看看区别

2

CN101166750A

治疗用咪唑并嘧啶类化合物和噻吩并嘧啶类化合物

2005-10-28

3

WO2007020193A2

Antiviral phosphoramidates of 4'-substituted pronucleotides

2006-08-03

在我们查询专利时不管通过什么的高超方法，总是会发现有漏检。不光是风险排查，或其他的查全工作中，都建议这种方法的应用，与查漏验证

The screenshot displays the Patsnap web application interface. On the left, there's a sidebar with filters for 'AND 过滤关键词', '1,064 组PatSnap', and various filters like '申请(专利权)人', '分类号', '申请年', '[标]当前申请(专利权)人类型', and '受理局'. The main area shows a list of patent results. A modal window titled '保存到工作空间' (Save to Workspace) is open, showing a list of patent options. The option '当前选中的专利及其PatSnap同族专利 (10,998)' is highlighted with a red box. A red arrow points from this option to a '保存到工作空间' button in the top right corner of the main interface. The button is also highlighted with a red box. Below the modal, there are two toggle switches: '专利状态发生变化时发邮件通知我' (OFF) and '有新专利时自动更新至该文件夹' (OFF). At the bottom of the modal are '保存到工作空间' and '取消' buttons.

查询专利的结果，可以保存至工作空间，选择保存Patsnap同族（可以调取检索结果内的所有同族专利）关注一下数据变化，在前面的同族里有介绍到，是有优先权关系，所以同族里专利相关性高

≡ patsnap 搜索 工作空间

支持搜索关键词、公司名、人名、专利号等

< 索非布韦 设置 工作空间小贴士

文件夹 过滤 层级

在当前文件夹搜索 高级

申请年 受理局

- ☐ 美国 2301
- ☐ WIPO 世界知识产权组织 1033
- ☐ 欧洲 956
- ☐ 中国 840
- ☐ 日本 643
- ☐ 澳大利亚 563
- ☐ 加拿大 386
- ☐ 中国香港 359
- ☐ 韩国 308
- ☐ 以色列 290
- ☒ 印度 270
- ☐ 墨西哥 222
- ☐ 新西兰 215

公开(公告)号 标题 发明人

1	● CN1816558A	修饰的氟化核苷类似物	J·克拉克
	☐ ● CN101166750A	治疗用咪唑并嘧啶类化合物和噻吩并嘧啶类化合物	亚拉加达·S·巴布 普兰·钱德 吴敏 普拉文·L·科田 V·萨蒂斯·库马 亚亚·埃尔·卡坦 阿吉特·K·戈斯
3	● KR1020070100237A	치료용 푸로피리미딘 및 티에노피리미딘	바부,알라가다,에: 찬드,푸란 우,민 코티안,프라빈,엘. 쿠마르,브이,사티. 린,츄 싱 엘 카틴

印度的数据增加了，初始受理局的数据45，通过我们把同族里相应的印度数据打开后，现在为270件，相信在风险排查中会是常用的一种方式

化学工作者还有哪些可以做的功课呢？

batsnap

是导航预警吗？

发现新的时机，新的化合物，新的用途，新的专利公开？ ...

<  消炎药(头孢) 

文件夹

过滤

层级

▼ 工艺

↻ 化学合成 (4918)

↻ 中间体合成 (975)

↻ 提纯、精制 (2307)

↻ 晶型 (21175)

▼ 制剂

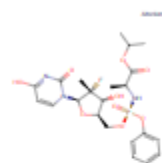
↻ 口服 (17372)

↻ 注射 (23032)

↻ 其他 (11645)

↻ 组合物 (14071)

邮件提醒



子结构搜索
SOFOSBUVIR

标题 SOFOSBUVIR-Substructure Search Email Alert

类型 新专利 (名称/摘要/权利要求)
新的相似化合物 (> 相似分数)
新的子结构
新的超结构
批准信息更新
临床信息更新

发送频率 每周一

邮件提醒

文件夹 (提纯、精制)

2307条专利

标题

提纯、精制

类型

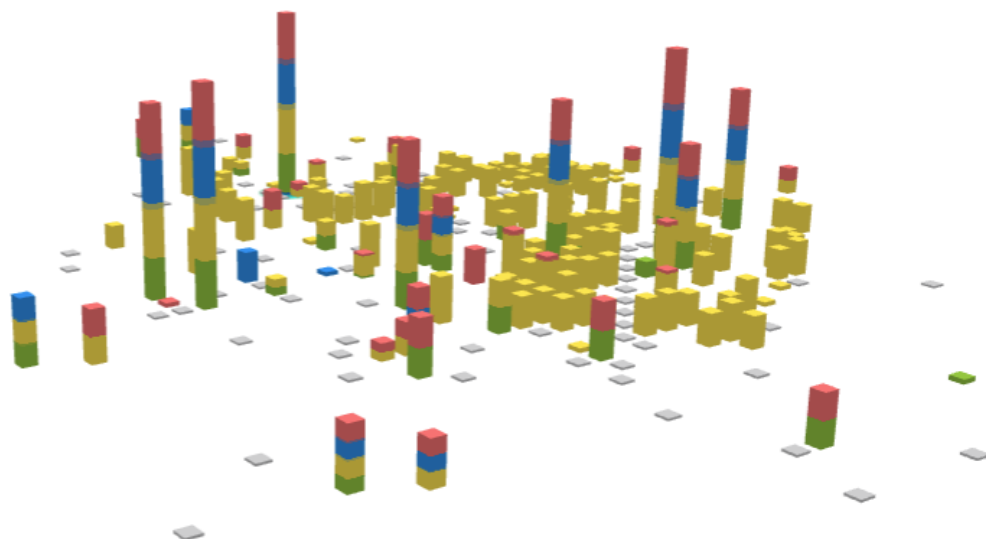
- ☒ 全选
- ☒ 专利更新 (当有符合检索条件的新专利公开)
- ☒ 首次公开提醒 (当专利为首次公开，此前没有任)
- ☒ 法律状态变更 (当专利法律状态变更时发送提醒)
- ☒ 同族专利更新 (当专利家族加入新成员时发送提醒)
- ☒ 被引用更新 (当专利被引用时发送提醒)
- ☒ 过期专利提醒 (当专利过期时发送提醒)
- ☒ 专利诉讼 (当专利诉讼信息被更新时发送提醒)
- ☒ 专利许可 (当专利许可信息更新时发送提醒)
- ☒ 备注更新提醒 (当有备注更新时发送提醒)

化学工作者还有哪些可以做的我呢?

patsnap

是专利分析吗?

筛选化合物, 筛选技术空白点, 筛选新的市场?



- 通过ChemScape 3D视图可以多维度对化合物和相关专利状况进行观察, 如:
 - 最核心的几个化合物的专利数量对比
 - 主要专利持有人的专利分布
 - 各个化合物的主要应用领域, 技术领域, 诉讼情况, 药品审批状态, 临床研发状态
 - 等等

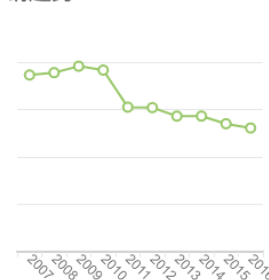
通过Analysis统计分析视图, 可以了解:

- 相关化合物的研发趋势, 专利申请量的逐年变化
- 主要的专利持有人排名
- 主要的治疗领域和技术领域
- 主要的专利审批地区
- 被大量专利保护的重要化合物排名

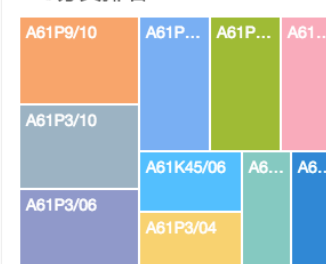
1-50 个化学物质, 总计4685个化学物质

Tanimoto系数 0.8 ↓ 专利数 50

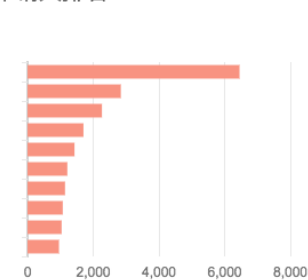
申请趋势



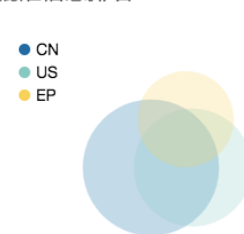
IPC分类排名



申请人排名



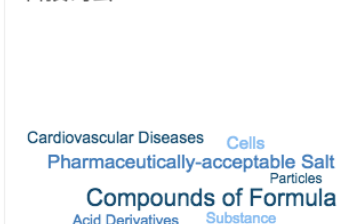
批准信息排名



CPC分类排名



科技词云



拥有超过**1.4亿**项专利，**1.7亿**种化学结构，集专利、文献、药品、临床、反应数据为一体

多维度数据

4400 万

化学专利文献

100 万

反应数据

1.4 亿

专利数据

420 万

非专利文献

13 万余

临床数据

12.5 万余

药品审批数据

4000 多万

有权利要求类型的专利

2.0 万

专利补充保护