What is a database?

- A collection of information organized in such a way that a computer program can quickly select desired pieces of data.

  You can think of a database as an electronic filing system.

SciFinder^n
A CAS SOLUTION

What do we need database?

- Data is easier to manager
  Too much data in one reference
  Too many references
  Too many language for reference

- Databases can better enforce data quality

- Leads to better data integration

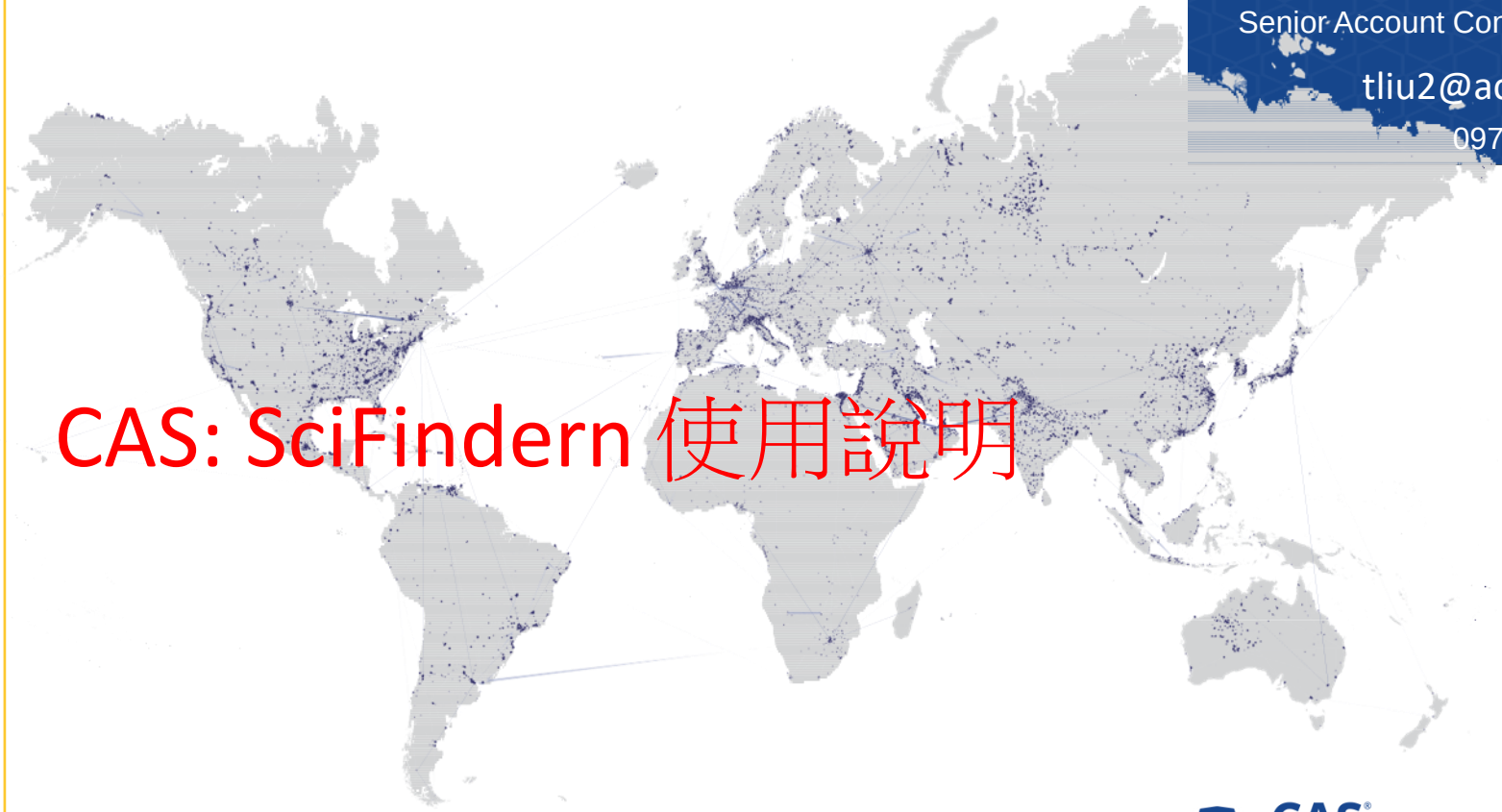
- Databases are used to store, organize and retrieve data, quickly!!

SciFinder^n
A CAS SOLUTION

議程：

時間	議題 / 議題大綱	主講人
40mins	**SciFinderN綜合介紹** 1.文獻專利查詢 (自動存取記錄, 專利物質引導, 文章引用地圖) 2. AI逆合成分析: SciFinder-n Synthesis Planning逆合成分析	劉全哲 美國化學協會資深客戶經理
10mins	中場休息	
40mins	1.化學結構查詢介紹 (高分子,有機金屬等…) 2.專利權利項結構搜尋 3.化學反應查詢 4.系統更新優化介紹	劉全哲 美國化學協會資深客戶經理

Chuan-Che Liu Ph.D.
Senior Account Consultant
tliu2@acs-i.org
0975765358



CAS: SciFindern 使用說明

CAS
A DIVISION OF THE
AMERICAN CHEMICAL SOCIETY

ABOUT US

Founded in 1907, chemists around the world understood the value to research in aggregating scientific information.

Today we are a global organization of expert scientists, technologists, and business leaders with a long and successful history of harnessing scientific information to support valuable research insights.

Approximately 1,600 staff members – including CAS scientists, speaking 50 languages among them (600+ Ph.D. Scientists)

CAS®
A DIVISION OF THE
AMERICAN CHEMICAL SOCIETY

CAS, Chemical Abstracts Service

- Coverage : CAplus, produced by CAS, contains more than 44 million chemistry and chemistry related research records categorized in 80 sections, with references in:

All areas of biochemistry	All areas of organic chemistry	All areas of macromolecular chemistry	All areas of applied chemistry	All areas of physical, inorganic, analytical chemistry
Agrochemical regulators	Amino acids	Cellulose, lignin, paper	Air pollution	Surface chemistry
Biochemical genetics	Biomolecules	Coatings, inks	Ceramics	Catalysts
Fermentation	Carbohydrates	Dyes, organic pigments	Essential oils, cosmetics	Phase equilibrium
Immunochemistry	Organometallic compounds	Synthetic elastomers	Fossil fuels	Nuclear phenomena
Pharmacology	Steroids	Textiles, fibers	Ferrous metals, alloys	Electrochemistry

Curating the highest quality chemistry content

arXiv.org

Aldrichimica ACTA

ACS chemical biology

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

division of polymer chemistry, inc American Chemical Society

JACS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

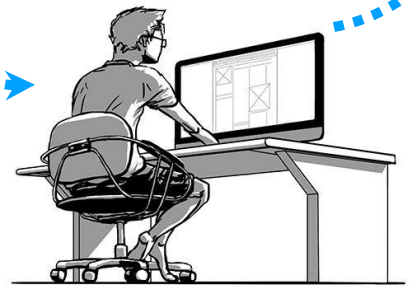
Cornell University Library

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

ACS Chemical Neuroscience

SCIFINDER^n A CAS SOLUTION

Data Platform



OUR PRODUCT

SciFinder^n
A CAS SOLUTION

OUR PRODUCT PORTFOLIO & CAS Service

CONTENT SERVICES

- Custom data sets
- Content licensing
- Data structure
- Data platforms

TECHNOLOGY SERVICES

- Data architecture
- Platform integration
- Search engines
- Custom solutions

KNOWLEDGE SERVICES

- Outsourced IP search
- Analytics
- Opportunity exploration
- Technical Assessment

PROFESSIONAL SERVICES

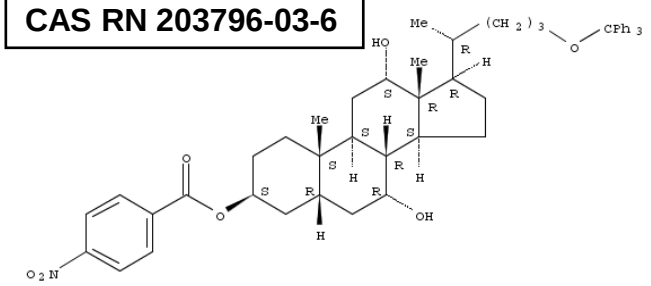
- Talent augmentation
- Consulting
- Technical writing
- Scientific analysis

SciFinderⁿ
A CAS SOLUTION

CAS scientists find the chemistry, and save you time!

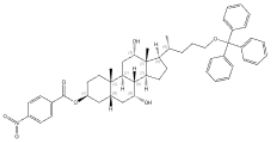
Compound 34: Diisopropyl azodicarboxylate (DIAD) (1.20 mL, 6.08 mmol) was added to triphenylphosphine (1.60 g, 6.08 mmol) in THF (100 mL) at 0 °C. and was stirred for half an hour during which time the yellow solution became a paste. Compound 14 (2.58 g, 4.06 mmol) and p-nitrobenzoic acid (0.81 g, 4.87 mmol) were dissolved in THF (50 mL) and added to the paste. The resulted mixture was stirred at ambient temperature overnight. Water (100 mL) was added and the mixture was made slightly basic by adding $NaHCO_3$ solution followed by extraction with EtOAc (3x50 mL). The combined extracts were washed with brine once and dried over anhydrous Na_2SO_4. The desired product (2.72 g, 85% yield) was obtained as white powder after SiO_2 chromatography (Et_2O/hexanes 1:2). m.p. 207-209 °C.; IR (KBr) 3434, 3056, 2940, 2868, 1722, 1608, 1529,1489, 1448, 1345 cm^{-1} ; ^1H NMR ($CDCl_3$, 300 MHz) δ 8.30-8.26 (m, 2 H), 8.21-8.16 (m, 2 H), 7.46-7.42 (m, 6 H), 7.31-7.18 (m, 9 H)5.33 (bs, 1 H), 4.02 (bs, 1 H), 3.90 (bs, 1 H), 3.09-2.97 (m, 2 H), 2.68 (td, J=14.95, 2.56 Hz, 1 H), 2.29-2.19 (m, 1 H), 2.07-1.06 (series of multiplets, 24 H), 1.01 (s, 3 H), 0.98 (d, J=6.6 Hz, 3 H), 0.70 (s, 3 H); ^{13}C NMR ($CDCl_3$, 75 MHz) δ 164.21, 150.56, 144.70, 136.79, 130.77, 64.22, 47.79, 46.79, 42.1 28.74, 27.71, 26.85, 26.3 (thioglycerol+Na^+ matrix

CAS RN 203796-03-6



Absolute stereochemistry.

203796-03-6



Absolute stereochemistry shown

$C_{50}H_{59}NO_7$

Cholane-3,7,12-triol, 24-(tripheny lmethoxy)-, 3-(4-nitrobenzoate), (3β, 5β,7α,12α)-

Role: Reactant, Synthetic Preparation, Reactant or Reagent, Preparation

SciFinder® A CAS SOLUTION

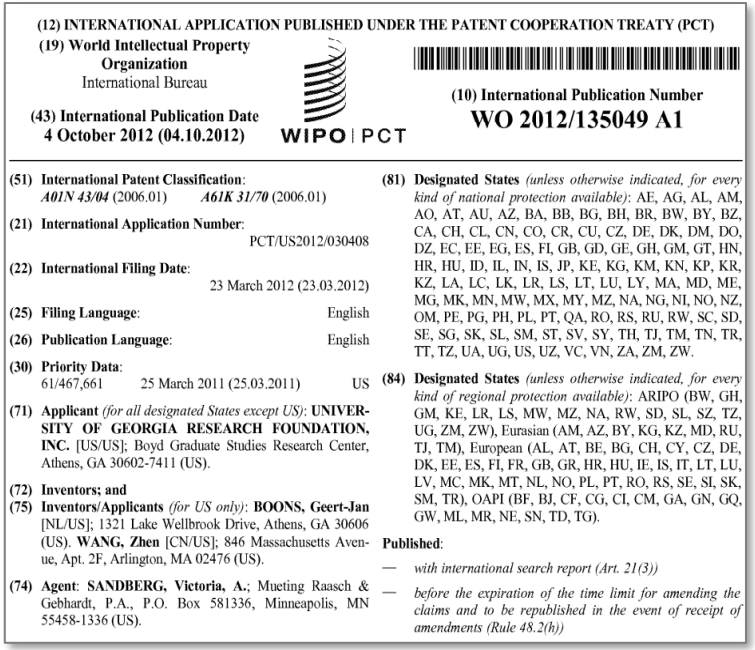
CONTENT TEAM

Content Providers

- 24/7/365 content intake
- ~30,000 documents daily, including pre-print
- Everything you need to know since the 1800s

CAS Acquisition Process

- Onsite Acquisitions Team – unique to CAS
- Proprietary processing technology
- Information painstakingly assembled so critical connections are made



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 October 2012 (04.10.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/135049 A1

(51) International Patent Classification:
A01N 43/04 (2006.01) *A61K 31/70* (2006.01)

(21) International Application Number:
PCT/US2012/030408

(22) International Filing Date:
23 March 2012 (23.03.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/467,661 25 March 2011 (25.03.2011) US

(71) Applicant *(for all designated States except US)*: **UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC.** [US/US]; Boyd Graduate Studies Research Center, Athens, GA 30602-7411 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **BOONS, Geert-Jan** [NL/US]; 1321 Lake Wellbrook Drive, Athens, GA 30606 (US). **WANG, Zhen** [CN/US]; 846 Massachusetts Avenue, Apt. 2F, Arlington, MA 02476 (US).

(74) Agent: **SANDBERG, Victoria, A.**; Mueting Raasch & Gebhardt, P.A., P.O. Box 581336, Minneapolis, MN 55458-1336 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

SciFinder^n
A CAS SOLUTION

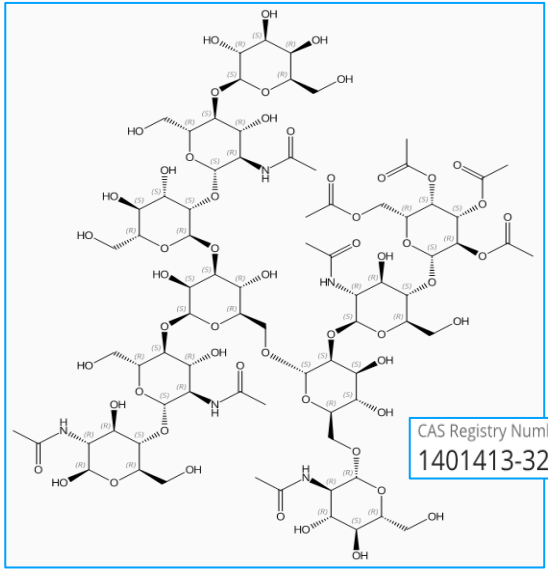
CONTENT TEAM

CAS Content Team

- At any given time ~1500 advanced degreed scientists and technologists covering 80 disciplines and 50 languages are working on CAS products and services
- Average experience: 20+ years
- ~$150,000,000 invested annually in content creation alone

Rigorous Hiring & Training Process

- Hiring processes includes actual testing of scientific knowledge – our own entrance exam…and hiring rate of ~5% of qualified applicants (lower than both Harvard and Oxford)
- Apprenticeship lasts 1-3 years
- Quality audits of content and workflow

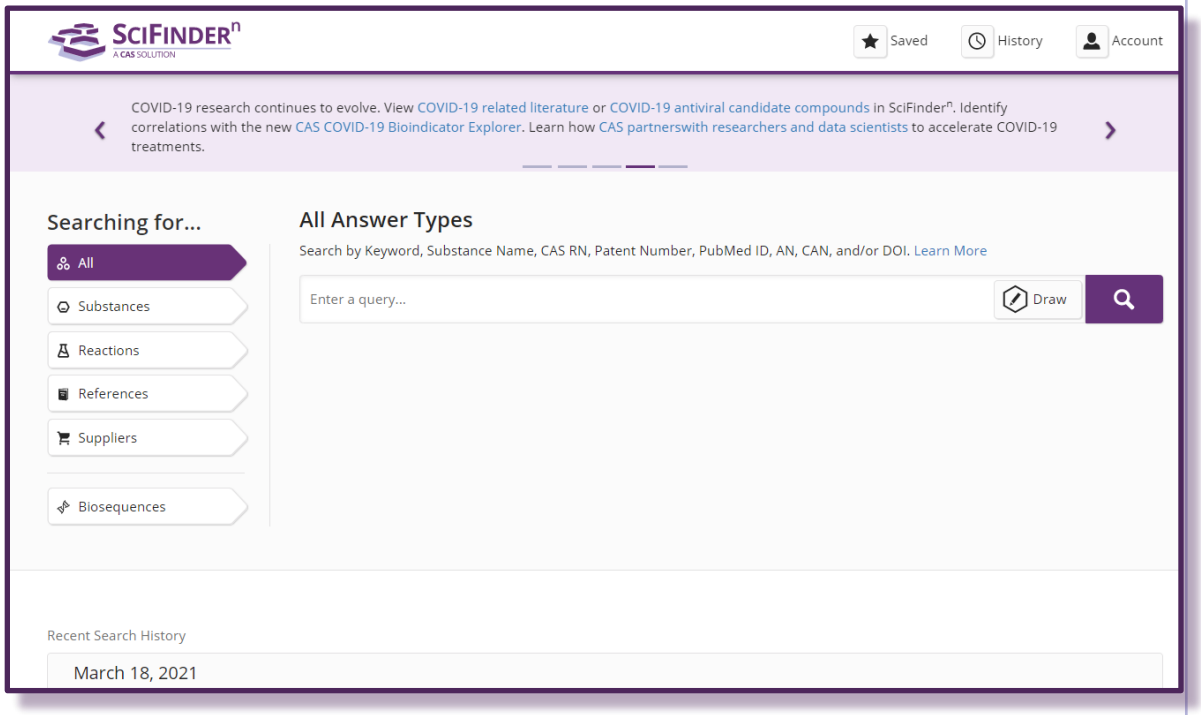


SciFinder^n
A CAS SOLUTION

Interface of Scifinder^n

Top 5 key searching function

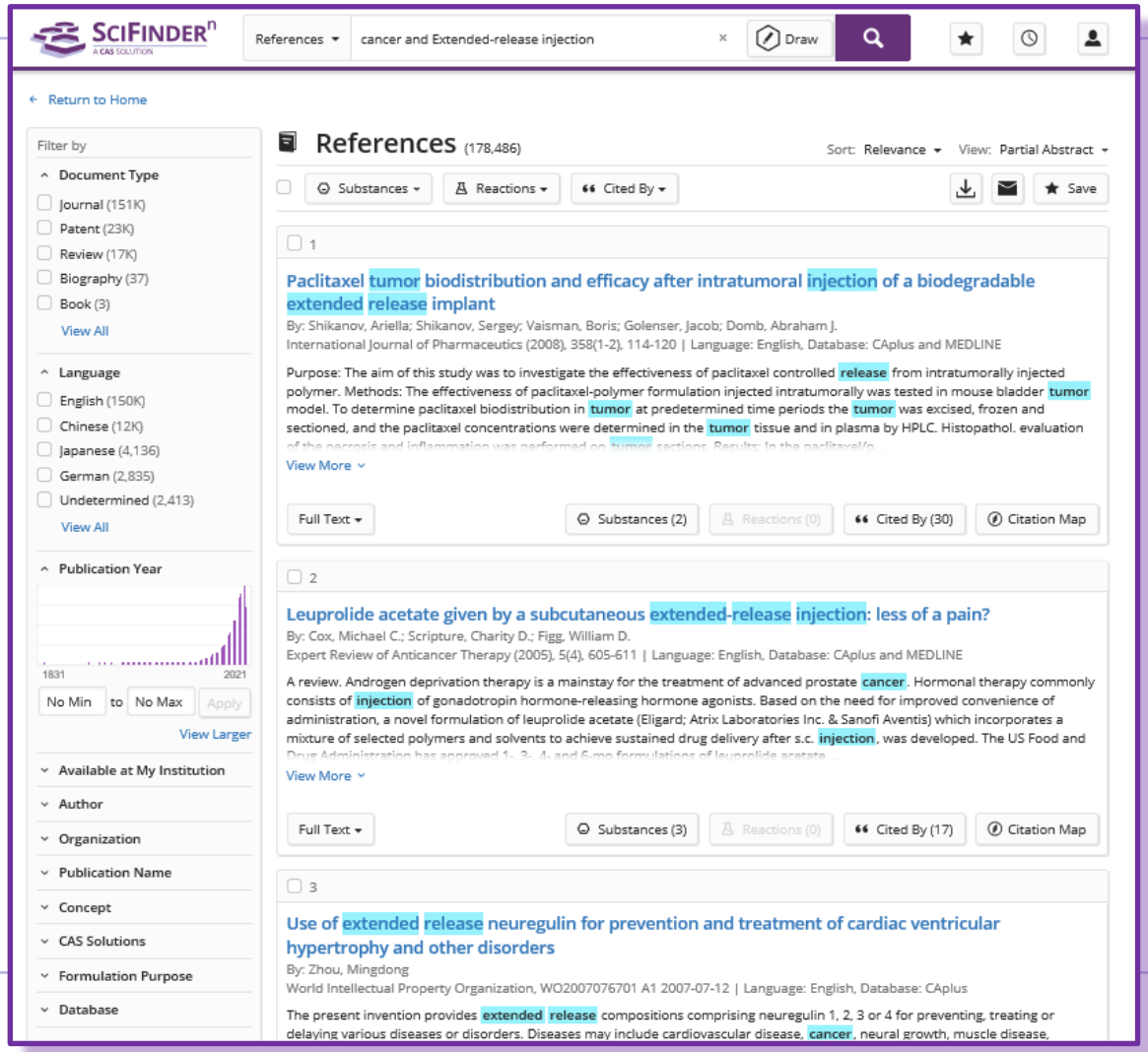
- Reference searching
- Reaction searching
- Substance searching
- Sequence searching



SciFinder^n
A CAS SOLUTION

Reference:

- An easy-to-read display lets user quickly browse reference
- Filters are available to further narrow results…by year, source type, and organization



Reference Detail Page:

Reference Detail

Substances (6) | Reactions (0)

Patent

Patent Information

Patent Number
WO2007076701

Publication Date
2007-07-12

Application Number
WO2006-CN3694

Application Date
2006-12-29

Kind Code
A1

Assignee
Zensun (Shanghai) Science & Technology Limited, China

Source
World Intellectual Property Organization

Database Information
AN: 2007:762994
CAN: 147:134439
CAplus

Language
English

Use of ventric

By: Zhou,

The presen various dis muscular schizophre failure and reducing t differentia cytoskelet

Keywords:

PATENT

Patent F

Patent

WO2007

AU20063

CA26349

CA28413

US20070

EP19815

CN10139

Concepts

Alzheimer disease

Animal gene
Modifier: NRG1, NRG2, NRG3 ar

Role: Biological Study, Unclassif Activity; Therapeutic Use

Anti-Alzheimer agents

Antidiabetic agents

Antitumor agents

Cardiac hypertrophy

Cardiomyocyte
Modifier: growth and differentia

Cardioprotective agents

Cardiovascular disease

Cell differentiation
Modifier: cardiomyocyte

Cell proliferation
Modifier: cardiomyocyte

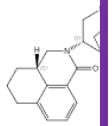
Controlled-release drug delive

Diabetes mellitus

Microspheres
Modifier: extended release neuregulin accomplished via

Substances (4)

135729-61-2

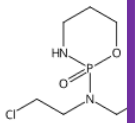


Absolute stereochemistry sh

$C_{19}H_{24}N_2O$
Palonosetron

Role: Pharmacological Activity Use, Biological Study, Uses

50-18-0



$C_7H_{15}Cl_2N_2O_2P$
Cyclophosphamide

Role: Adverse Effect, Includin Pharmacological Activity, The Biological Study, Uses

Formulations

Collagen Formulation: Anticancer Agents

View **Formulus®** Detail

Location: Article Page 3, 6
Purpose: Antitumor agents
Target: cancer

Component	Function	Amount Reported
Group: cyclic dinucleotides stock solution	-	-
cyclic dinucleotides	-	6.67 µg
endotoxin-free water	solvent	-
Phosphate-buffered saline solutions	buffer	-
collagen stock solution	carrier	3.5-5 mg/mL

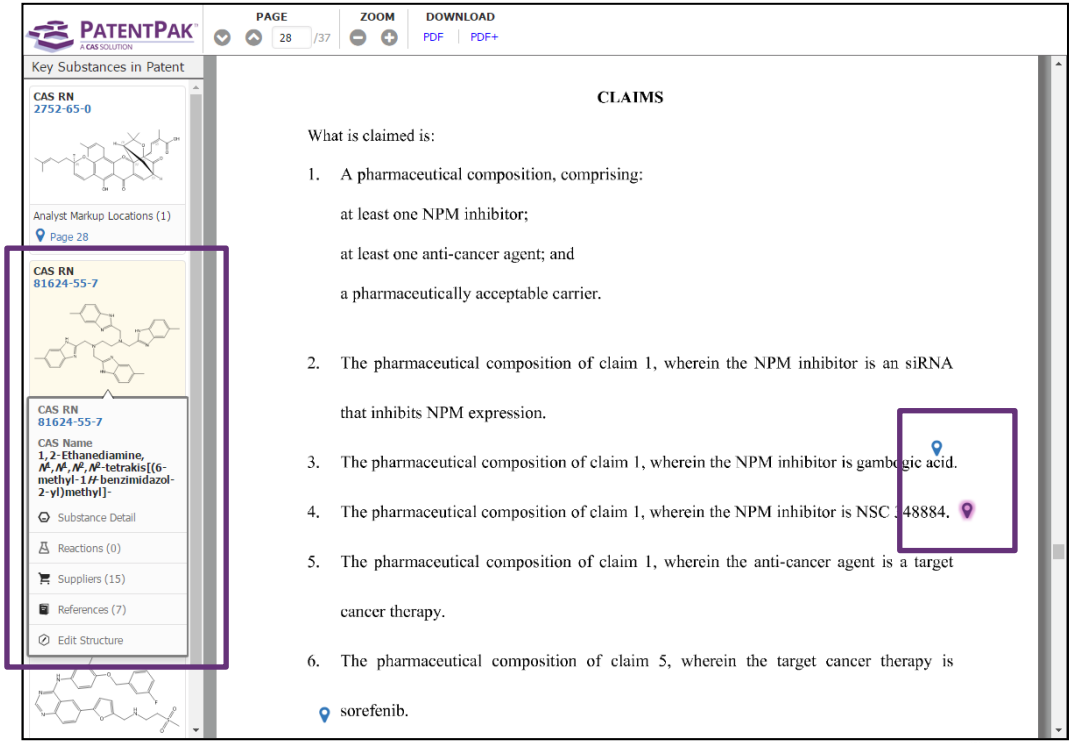
Stingel Formulation: Anticancer Agents

View **Formulus®** Detail

Location: Article Page 3, 6

PatentPak:
Why waste time slogging through dense patent material with direct access to and understanding of the chemistry within the document

Patent chemistry is fully annotated with structures, nomenclature and more!



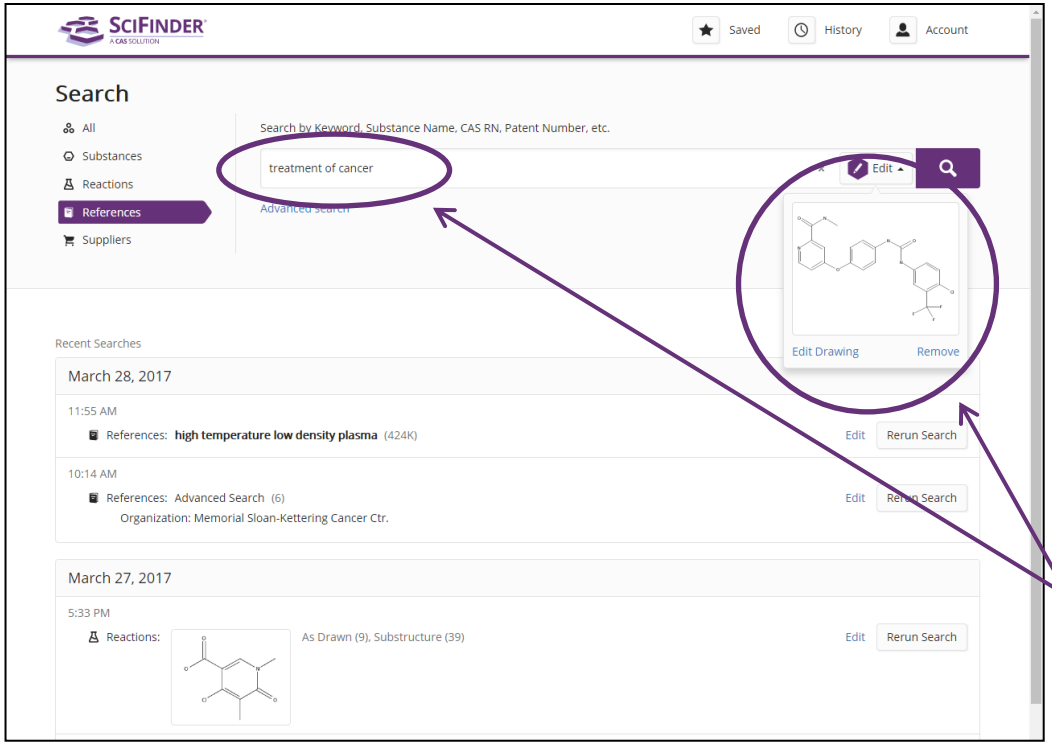
Important chemistry locations are identified by CAS expert scientists

SciFinder^n
A CAS SOLUTION

Get to actionable results more quickly with a streamlined new interface

Streamlined search navigation saves you clicks

Immediately run past searches

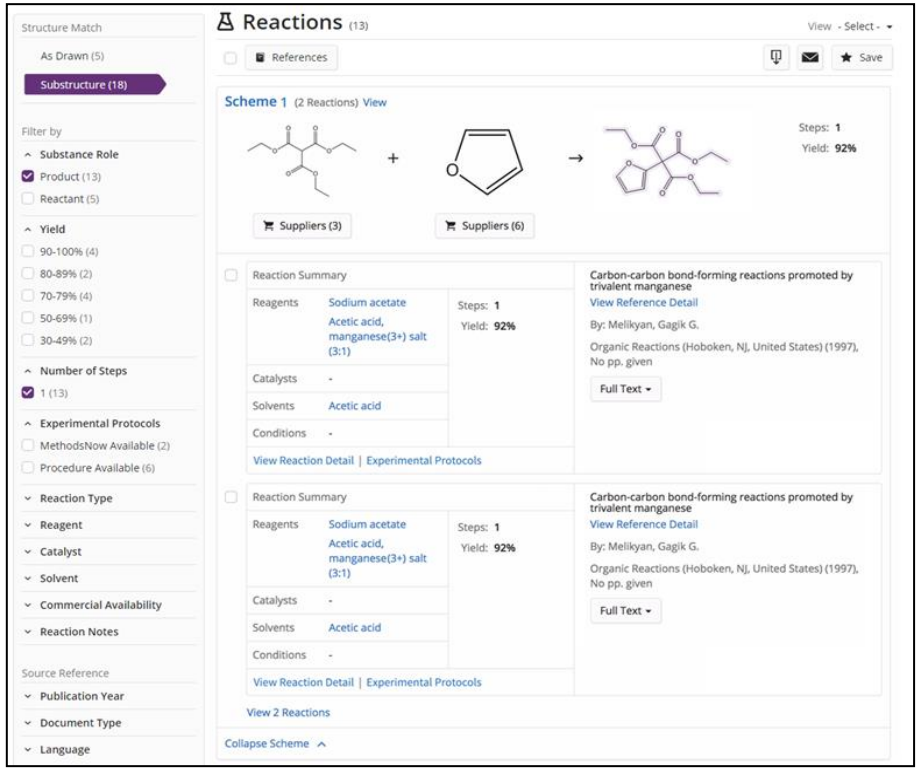


More conveniently search structures and text

SciFinder^n
A CAS SOLUTION

Reaction: Information presented to facilitate rapid understanding

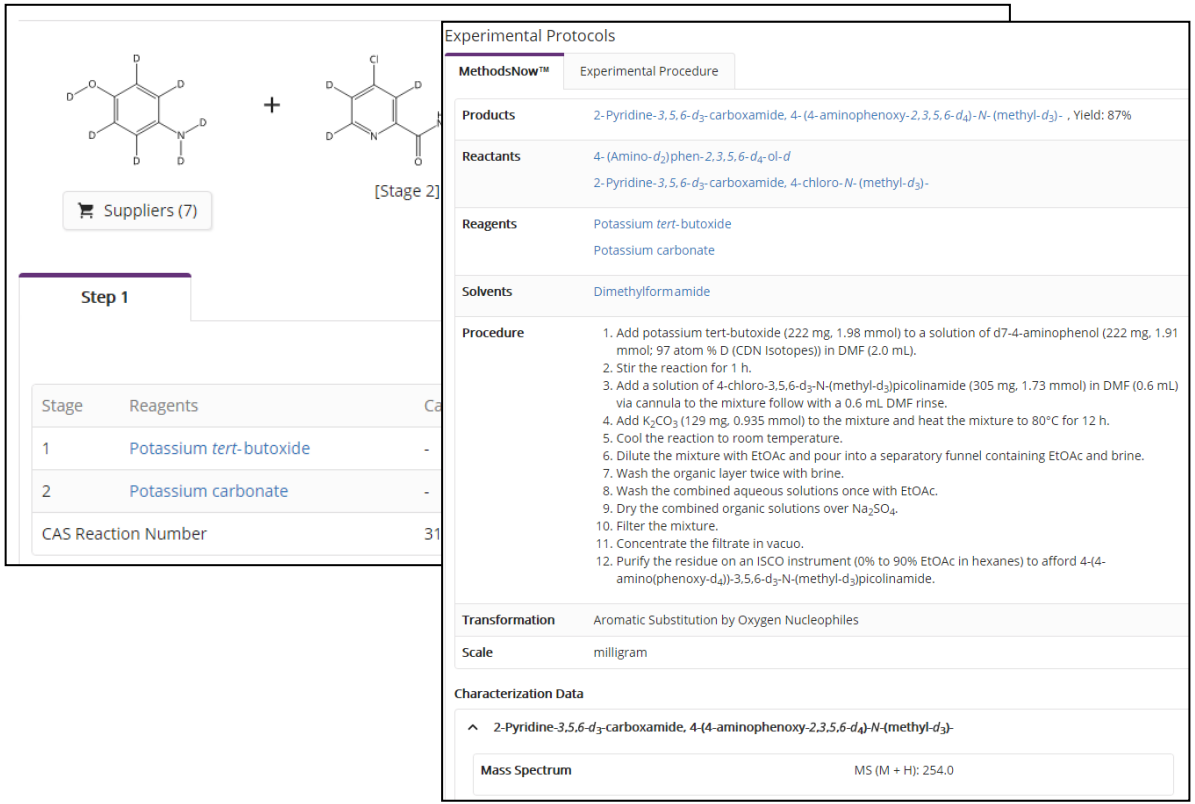
Powerful filtering capabilities allow rapid focus



Intuitive information layouts fosters quick comprehension

SciFinder^n A CAS SOLUTION

Go right to the lab with step-by-step synthetic procedures extracted and summarized from the literature



Key chemical components of the reaction are identified with links to additional information

Go directly to the lab with step-by-step instructions

Additional useful information is included

SciFinder^n
A CAS SOLUTION

Retrosynthesis

Powered by ChemPlanner®

Overview Steps Scoring

Save

Step Key

Plan Opti

Synthetic
Predicted
Break & P
Edit Plan

Plan Infor

Estimated
Overall Pr
(USD per

Commerc
A, B, C, D,

Overview Steps Scoring

A ⇒ B + C *Stereoselective*
Maximum Yield: -
Evidence (3)
Alterr

B ⇒
Avera
Evide
Alterr

C ⇒
Maxir
Evide
Alterr

D ⇒
Maxir
Evide
Alterr

F ⇒
Maxir
Evide

Overview Steps Scoring

Scoring Profiles

Complexity Reduction

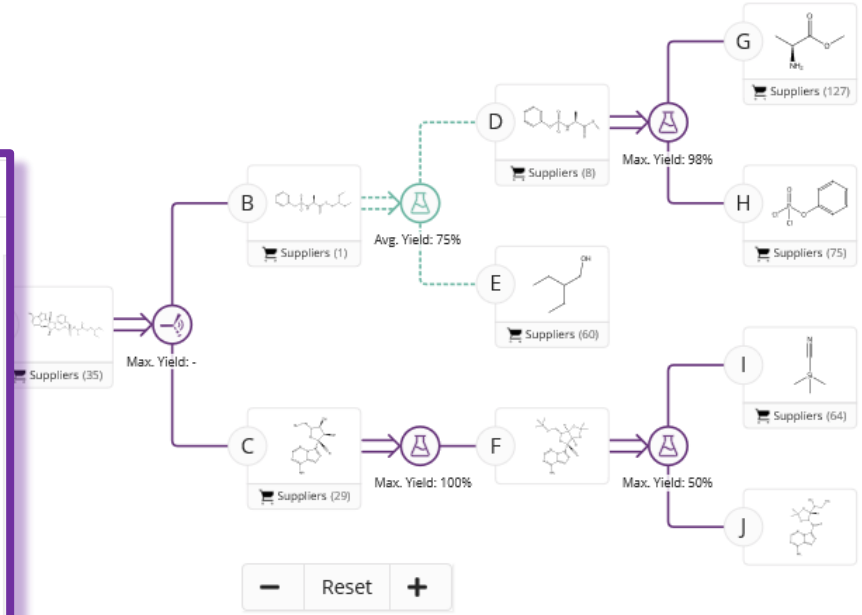
Convergence

Evidence

Yield

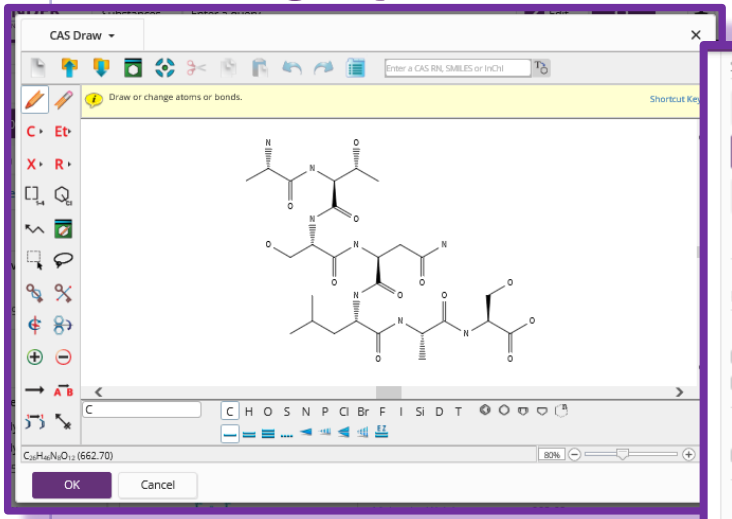
Atom Efficiency

Apply Reset Scoring



Feedback

Searching by Substances



Structure Match

As Drawn (1)

Substructure (101)

Similarity (1.2M)

Analyze Structure Precision

Filter by

- Commercial Availability
  - Available (10)
  - Not Available (91)
- Reaction Role
  - Product (2)
- Reference Role
  - Analytical Study (20)
  - Biological Study (75)
  - Preparation (45)
  - Process (3)
  - Properties (101)
  - View All
- Stereochemistry
  - Absolute Stereo Match (101)
- Number of Components
- Substance Class
- Isotopes
- Metals
- Molecular Weight
- Regulatory Information
- Bioactivity Indicator

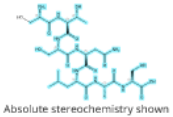
Substances (101)

Sort: Number of References: Descending View: Full

References Reactions Suppliers Save

1

156847-82-4



Absolute stereochemistry shown

$C_{26}H_{46}N_8O_{13}$

L-Seryl-L-threonyl-L-seryl-L-asparaginyl-L-leucyl-L-alanyl-L-serine

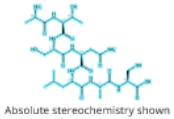
Protein/Peptide Sequence
Sequence Length: 7

247 References | 1 Reaction | 2 Suppliers

Key Physical Properties	Value	Condition
Molecular Weight	678.69	-
Boiling Point (Predicted)	1332.0±65.0 °C	Press: 760 Torr
Density (Predicted)	1.396±0.06 g/cm³	Temp: 20 °C; Press: 760 Torr
pKa (Predicted)	3.02±0.10	Most Acidic Temp: 25 °C

2

192433-87-7



Absolute stereochemistry shown

$C_{26}H_{46}N_8O_{12}$

L-Alanyl-L-threonyl-L-seryl-L-asparaginyl-L-leucyl-L-alanyl-L-serine

Protein/Peptide Sequence
Sequence Length: 7

141 References | 0 Reactions | 1 Supplier

Key Physical Properties	Value	Condition
Molecular Weight	662.69	-
Boiling Point (Predicted)	1257.6±65.0 °C	Press: 760 Torr
Density (Predicted)	1.357±0.06 g/cm³	Temp: 20 °C; Press: 760 Torr
pKa (Predicted)	3.02±0.10	Most Acidic Temp: 25 °C

3

167635-77-0

Key Physical Properties	Value	Condition

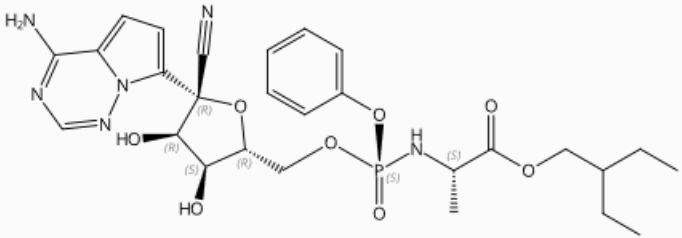
Substances Detail Page

Substance Detail (1 of 1)

Prev Next

References (488) | Reactions (325) | Suppliers (37) | Save

CAS Registry Number
1809249-37-3



Absolute stereochemistry shown

$C_{27}H_{35}N_6O_8P$

L-Alanine, *N*-[(*S*)-hydroxyphenoxyphosphinyl]-, 2-ethylbutyl ester, 6-ester with 2-*C*-(4-aminopyrrolo[2,1-*f*][1,2,4]triazin-7-yl)-2,5-anhydro-D-altrononitrile

Key Physical Properties	Value	Condition
Molecular Weight	602.58	-
Density (Predicted)	1.47±0.1 g/cm³	Temp: 20 °C; Press: 760 Torr
pKa (Predicted)	12.00±0.70	Most Acidic Temp: 25 °C

Expand All |

- Other Names and Identifiers
- Predicted Properties
- Predicted Spectra
- Bioactivity Indicators
  - Anti-infective agents
    - Antiviral agents (183)
- Target Indicators
  - Enzymes
    - Alanine aminotransferase (12)
    - Angiotensin-converting enzyme 2 (29)
    - Proteinase 3CLpro (11)
    - RNA-dependent RNA polymerase (30)
  - Glycoproteins
    - Viral spike glycoproteins (30)
  - Viral proteins
    - Viral spike glycoproteins (30)
- Regulatory Information
  - **Confidential Business Information:** Public
  - Regulatory Synonyms (2)
  - Details by Country/International & Other Lists
- Additional Details

Markush Detail Page

← Return to Home

Patent Markush Match

As Drawn (71)

Substructure (1,314)

Filter by

Patent Office

- World Intellectual Property Organization (27)
- United States (20)
- Japan (13)
- China (5)
- European Patent Organization (4)

View All

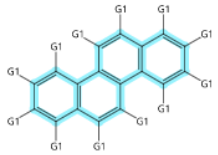
Patent Markush (71)

Sort: Relevance

References | Save

1

US20130214259



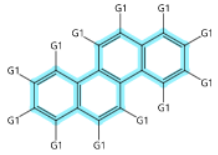
Patent claim 29

PATENTPAK | Full Text

There are no notes to display for this structure.

2

WO2009008352



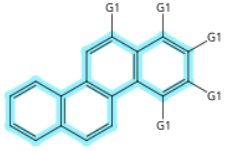
Patent claim 1

PATENTPAK | Full Text

There are no notes to display for this structure.

3

WO2002088025



Patent claim 2

PATENTPAK | Full Text

There are no notes to display for this structure.

PATENTPAK | PAGE 79 /100 | ZOOM | DOWNLOAD PDF PDF+

Key Substances in Patent

CAS RN 5122-94-1

Analyst Markup Locations (2)

Page 79

Page 82

CAS RN 591-18-4

Analyst Markup Locations (1)

Page 79

CAS RN 1762-87-4

Analyst Markup Locations (1)

Page 79

CAS RN 73183-34-3

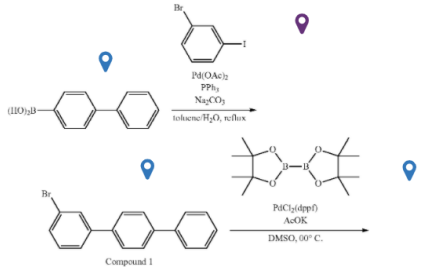
Analyst Markup Locations (1)

Page 79

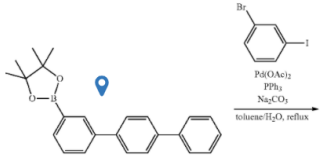
CAS RN 1401577-23-8

dures, and the like shown in the following Examples can be appropriately modified so far as the gist of the present inven-

[0281] A compound (H-1) was synthesized according to the following scheme.

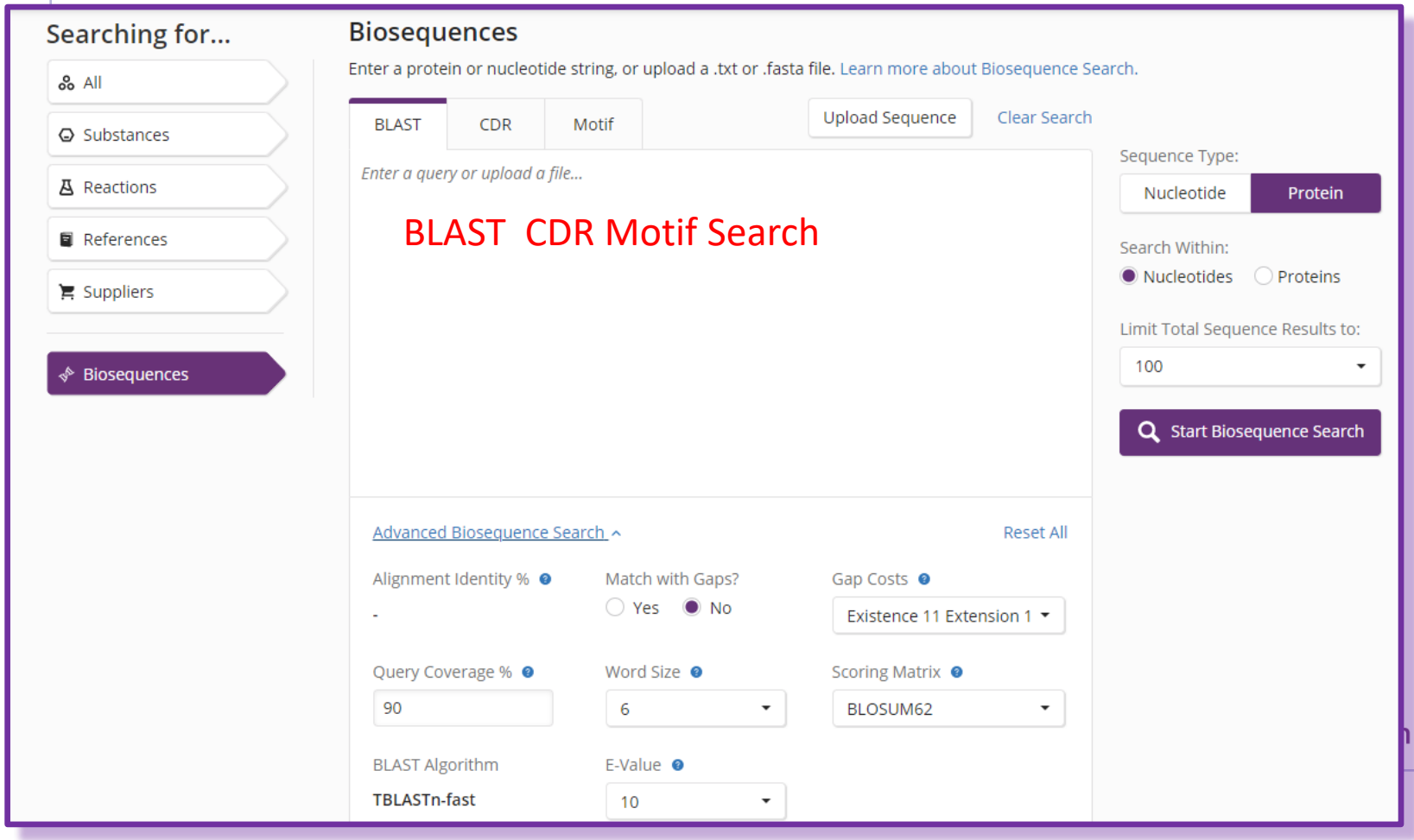


Compound 1



Compound 2

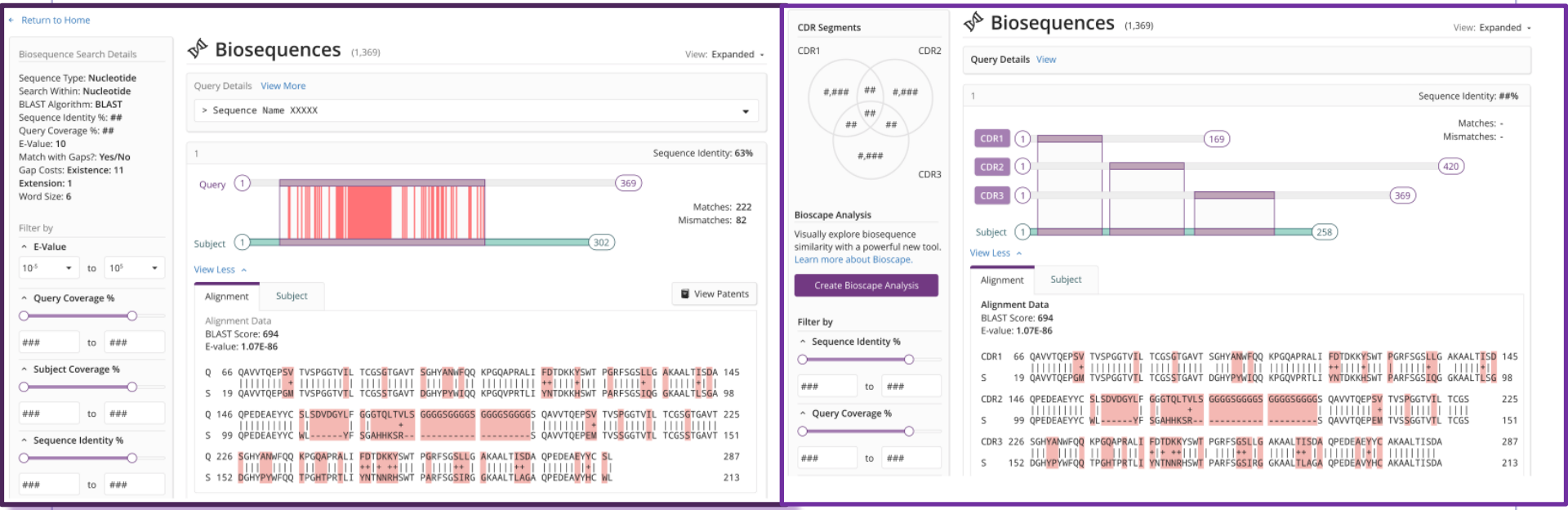
New Content & Capabilities - Sequence Search



New Content & Capabilities - Sequence Search

BLAST Display

CDR Display



SciFinder^n
A CAS SOLUTION

Chuan-Che Liu Ph.D.
Senior Account Consultant
tliu2@acs-i.org

SciFinderN 文獻查詢

CAS
A DIVISION OF THE
AMERICAN CHEMICAL SOCIETY

Find References

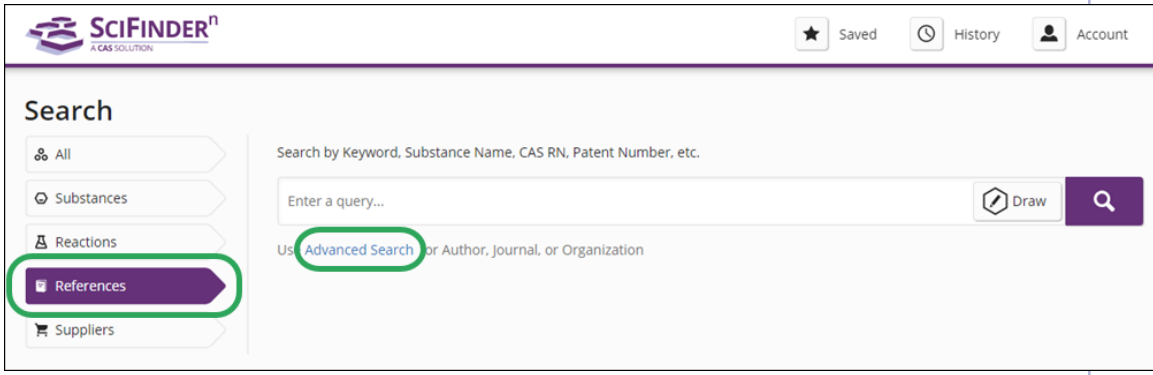
Find document references that match your query, which can be **search terms**, a **chemical structure**, or both:
•If you enter a search term **or** chemical structure query, SciFinder^n retrieves associated references, such as journal articles and patents.
•If you enter **both** a search term query and a structure query, the search results must match both criteria (queries are joined by the logical AND).

Find ADVANCED **References** by:
•**Author and Keywords**
•**A Specific Author**

Find **References Associated with a Company or Organization**

Find a **Reference Using Journal Information**

Find **Patent References** by:
•**Assignee**
•**Inventor**



SciFinder^n
A CAS SOLUTION

Find References Using Search Terms

Search terms may be:

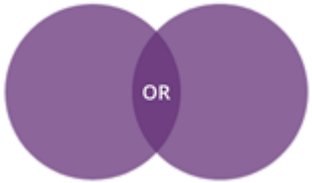
- **Research Topic/Keyword/Concept** (e.g., analgesics)
- **Substance Name** (e.g., ibuprofen, ß-amyloid)
  Note: If your search term contains a Greek letter, see Greek Letters Used in Searching to insert the correct character.
- **CAS Registry Number** (with dashes, e.g., 51146-57-7)
- **Accession Number** (e.g., 1986:230471)
- **PubMed ID Number** (e.g., 15980585)
- **Digital Object Identifier (DOI)** (e.g., 10.1093/nar/gki470)
- **Patent Number** (no spaces, e.g., US4571400)
- **Patent Application Number** (with dash, e.g., US1984-682902)

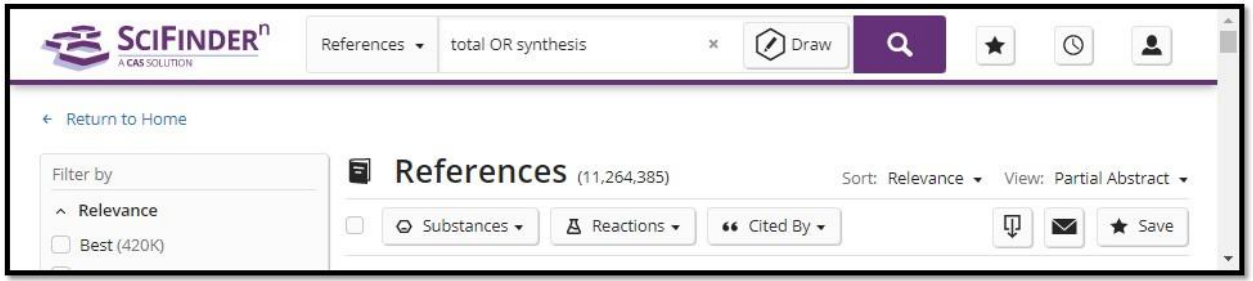
Note: You may enter multiple numbers separated by a space, no commas or other punctuation. The search field has a 2000-character limit.

SciFinder^n
A CAS SOLUTION

There are a total of five basic Boolean search operators:

- OR
  - The OR operator will provide results that contain either of the keywords
  - Will provide results that contain either of the terms
  - This is the default search when no operators are present in the search query
  - total OR synthesis





Examples:

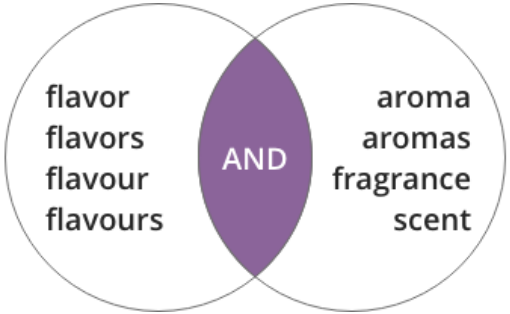
- ("flavor" **or** "extract") and ("turmeric" **or** "curcumin")
- (flavor not dye) **or** extract
- 13463-67-7 **or** 7664-41-7

SciFinder n
A CAS SOLUTION

•AND

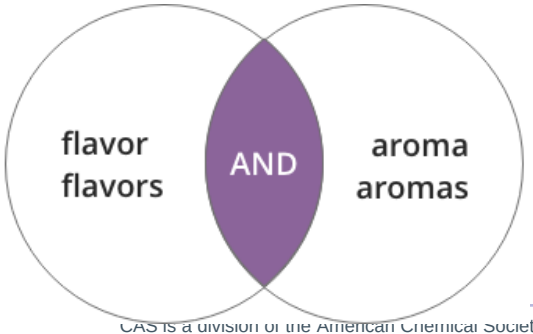
The **and** operator returns any combination of the terms' plurals, synonyms, or alternate spellings.

Example: The query **flavor and aroma** returns only results that **contain plurals/synonyms/alternate spellings of both terms**.



Entering terms in quotation marks to create a bound phrase **narrows the term to include**.

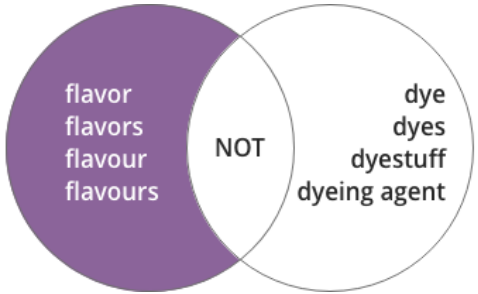
Example: The query **"flavor" and "aroma"** returns only results that **contain the exact spelling (includes plurals) of both terms**.



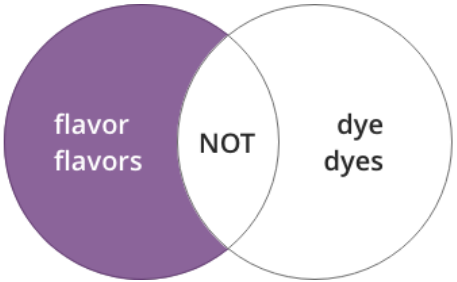
SciFinder^n
A CAS SOLUTION

•NOT

The **not** operator excludes any combination of the terms' plurals, synonyms, or alternate spellings.
Example: The query **flavor not dye** returns results that **contain plurals/synonyms/alternate spellings of flavor** but **excludes** those that **contain plurals/synonyms/alternate spellings of both flavor and dye**.



Entering terms in quotation marks to create a bound phrase **narrows the term to exclude**.
Example: The query **"flavor" not "dye"** returns results that **contain flavor(s)** but **excludes** those that **contain both flavor(s) and dye(s)**.

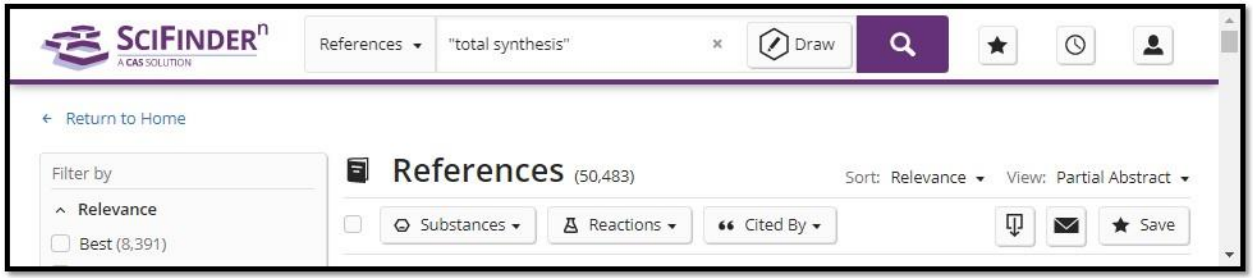


SciFinder^n
A CAS SOLUTION

- “”

  - Quotation marks means you are searching for a specific phrase
  - Will provide results that contain the exact phrase and will not return references with uses of these keywords on their own
  - “total synthesis”

"term"

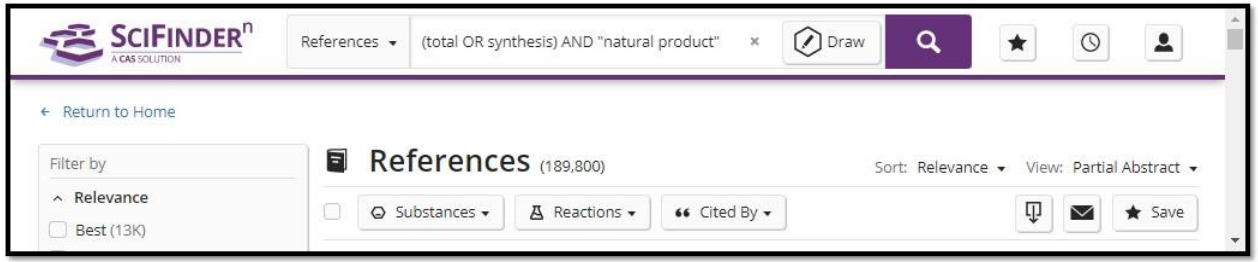


Entering terms in **quotation marks** **creates a bound phrase that specifies an exact spelling** (plurals accepted, but no alternative spellings or synonyms) **and side-by-side relationship**.

SciFinder n
A CAS SOLUTION

• ()

- Parentheses can be added to allow the user to combine Boolean operators
- (total OR synthesis) AND “natural product”



Entering terms and operators in **parentheses** **creates an expression that functions as a single unit** that interacts with other terms.

Examples:

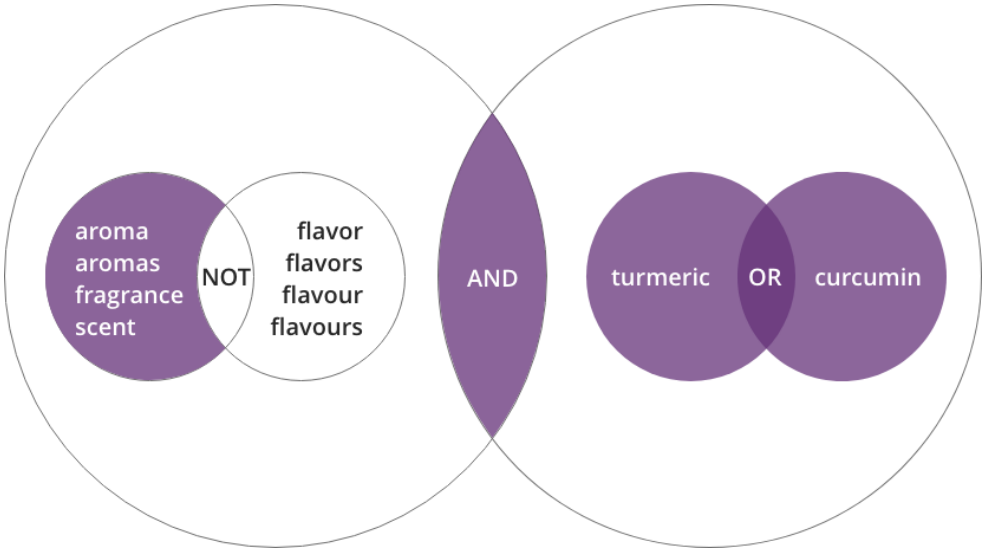
- **(**extract not flavor**)** and **(**"turmeric" or "curcumin"**)**
- **((**turmeric or flavonoids**)** and immune cells**)** not inflammatory

SciFinder n
A CAS SOLUTION

Boolean Modifier: ( )

Enclosing **terms and a Boolean operator within parentheses creates a Boolean expression that functions as a single unit/condition** that can interact with other terms, expressions, phrases, and operators.

Example: The query **(aroma not flavor) and ("turmeric" or "curcumin")** returns only results that **contain plurals/synonyms/alternate spellings of aroma, but not plurals/synonyms/alternate spellings of flavor and also contain turmeric, curcumin, or both**.



SciFinder^n
A CAS SOLUTION

Syntactic Reference Text Wildcard Searching * ?

You may search references using the asterisk (* - match 0 or more characters) and question mark (? - match 0 or 1 character) wildcard characters.

Wildcard queries search a reference's:

•Title

•Abstract

•Keywords

•Substances/Medline Chemicals

•CAplus Concepts

Notes:

•Wildcards **contained within quotes are ignored**.

•Wildcards **do not find two terms that a separated by a space** (for example, *water?based* will find *water-based* but not *water based*).

The following wildcard characters will be supported:

- * - Used to match 0 or more characters
- ? - Used to match 0 or 1 character

Wildcard characters can be used within a term (infix) or at the end of a term (terminal).

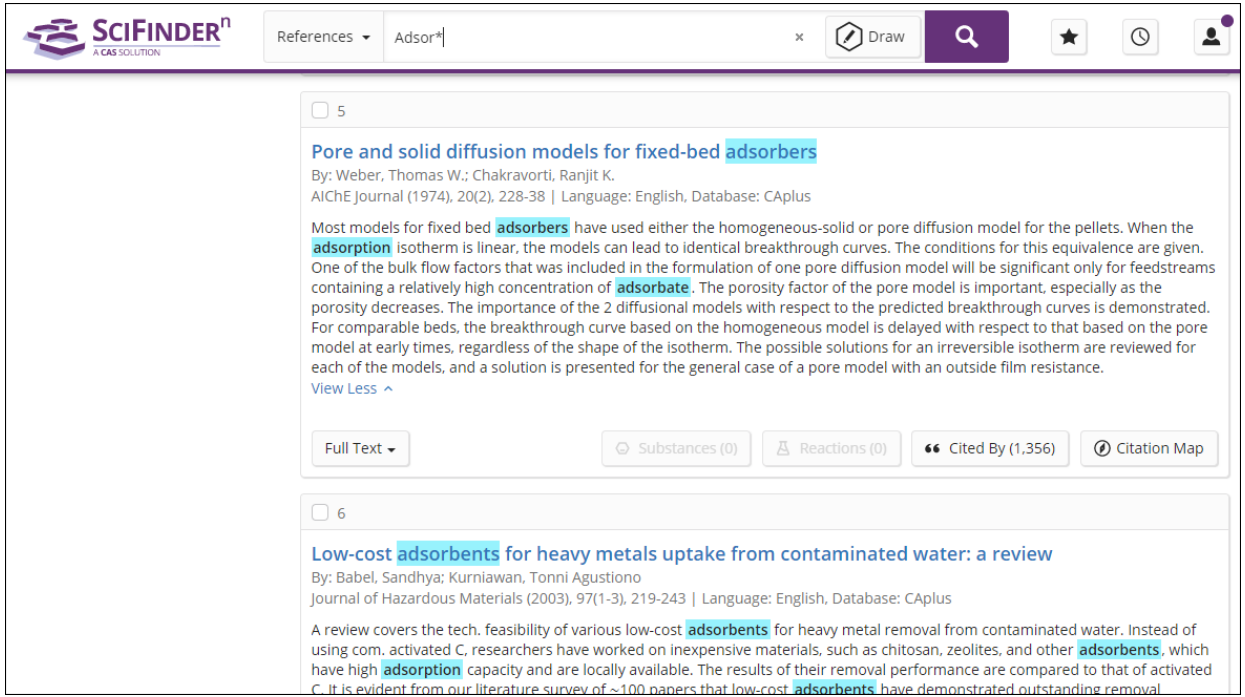
- Leading wildcard characters are not supported and will be ignored in a query.

SciFinder^n
A CAS SOLUTION

Examples

The query **Adsor*** will match terms such as:

- **Adsorbents**
- **Adsorbers**
- **Adsorption**

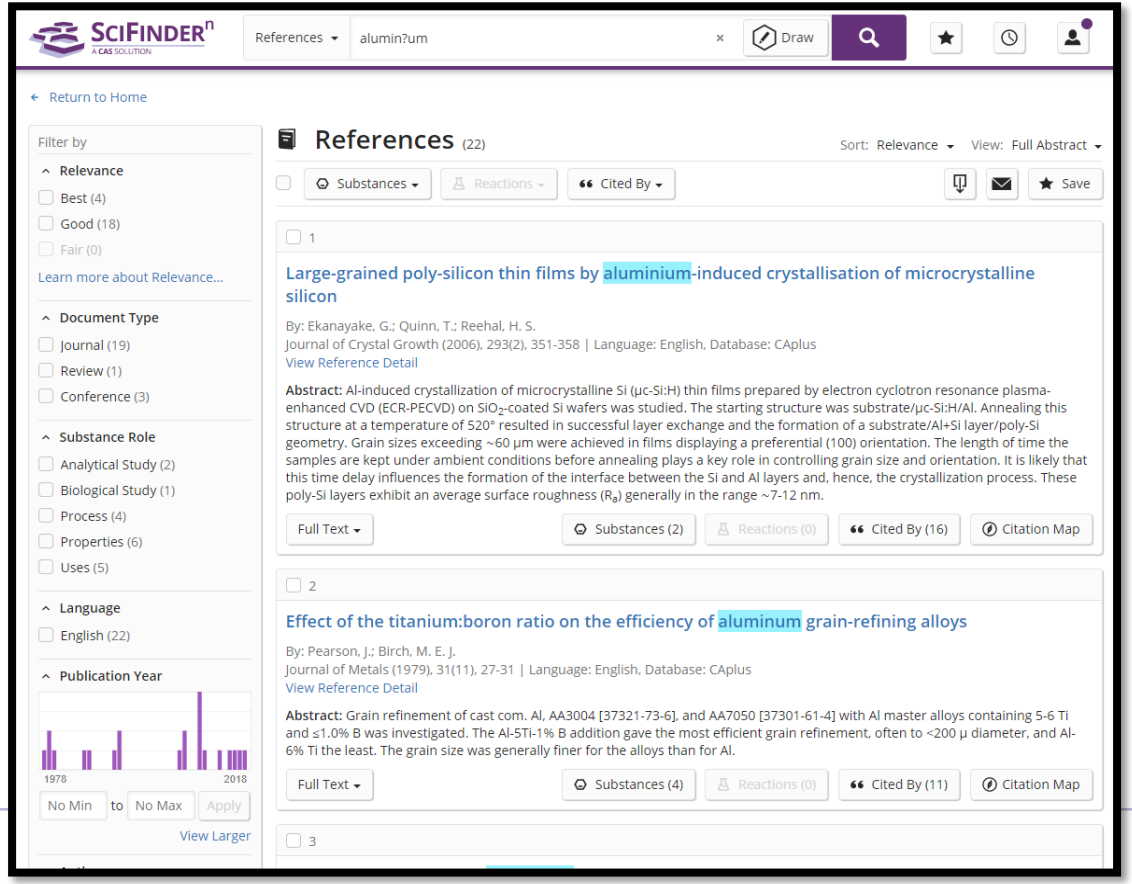


SciFinder n
A CAS SOLUTION

Examples

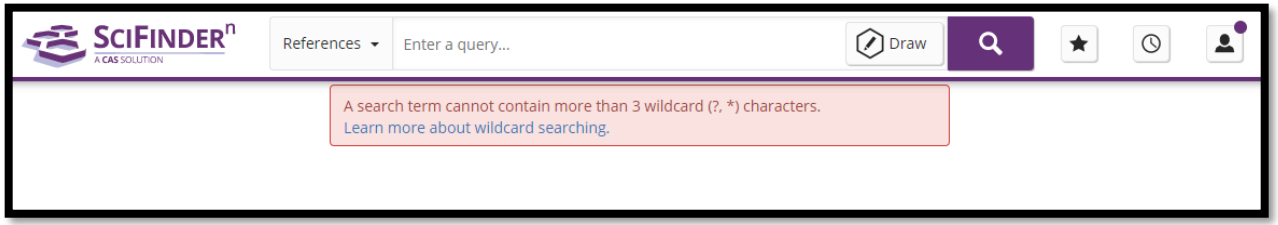
The query alumin?um will hit on terms such as

- Aluminum
- Aluminium
- Etc.



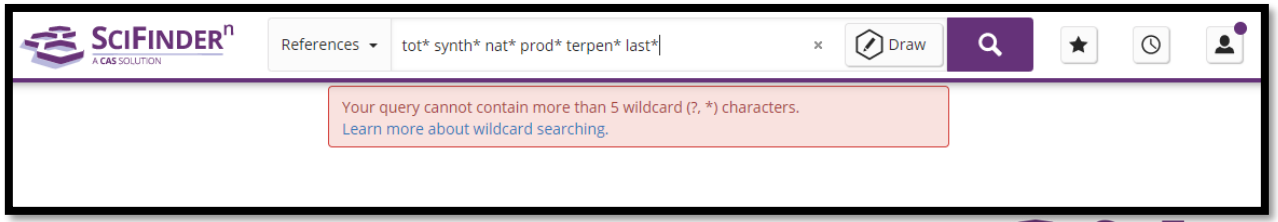
Exceptions:

- A term must have at least 3 non-wildcard characters for a wildcard to be searched
  - If a term has less than 3 non-wildcard characters, any wildcard terms will be searched literally.
- A search term cannot contain more than 3 wildcard characters
  - Currently, searchers will be provided with an error message for this scenario.



- In a subsequent build, terms with more than 1 wildcard character will be searched literally and the user will be provided with an informational message.

A search cannot contain more than 5 wildcard terms.



SciFinder
A CAS SOLUTION

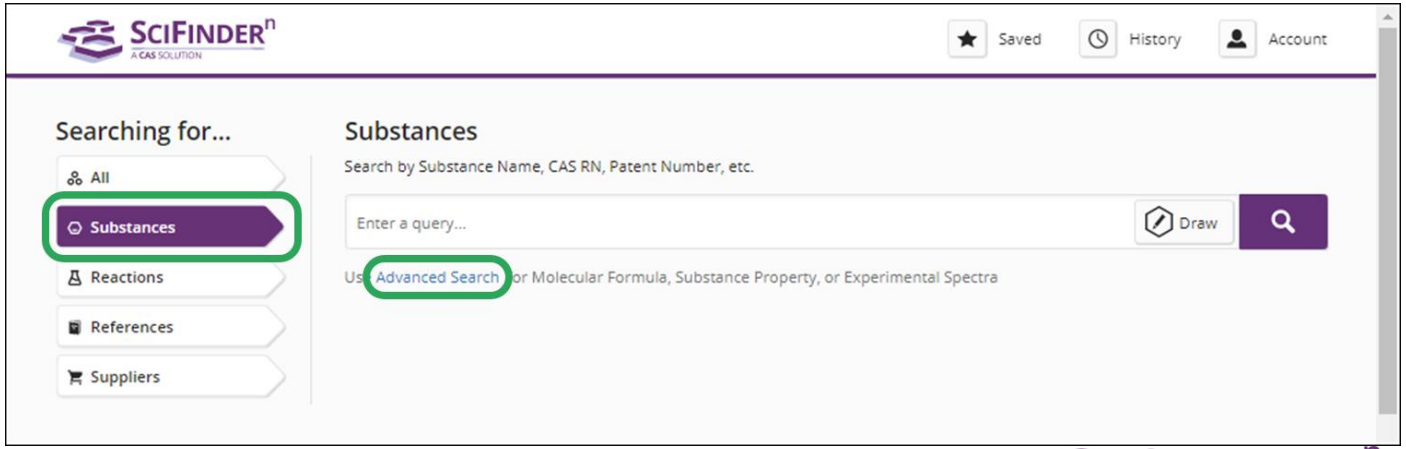
Chuan-Che Liu Ph.D.
Senior Account Consultant
tliu2@acs-i.org

Substance Searching

CAS
A DIVISION OF THE
AMERICAN CHEMICAL SOCIETY

Find Substances

- Find substances that match your query, which can be substance names, CAS Registry Numbers, document identifiers, or a chemical structure:
- **Find Substances by Substance Name, Registry Number, or Document Identifier**
- **Find Substances by Chemical Structure**
- ***Note:*** If **both search term and chemical structure queries are present**, substances will **only match the chemical structure query (search term query is ignored)**.
- Click **Advanced Search** to find substances by **molecular formula**, **substance property**, and **experimental spectra**.



SciFinder n
A CAS SOLUTION

Find Substances by Name, Registry Number, or Document Identifier

Search terms may be:

•**Substance Name:**

- **Example:** benoxaprofen
  Note: Wildcard searching using an asterisk only works for single-search-term, single-word substance name searches (e.g., benoxa*). Partial substance names without a wildcard character will not return results.
- **Example:** methyl ethyl ketone
  Note: For an exact match, use double quotes around the substance name. Searching multiple word substance names without quotes also returns results matching part of the name. For example, searching the above example without double quotes returns methyl ethyl ketone, methyl, ethyl ketone, and ethyl.
- **Example:** ß-amyloid
  Note: If your search term contains a Greek letter, see Greek Letters Used in Searching to insert the correct character. to insert the correct character.

•**CAS Registry Number** (with or without dashes, e.g., 51146-57-7, 51146577)
Note: Square brackets are accepted around RNs (e.g., [51146-57-7], but not single or double quotes.

•**Document Identifier:**

- **Patent Number** (no spaces, e.g., US4571400)
- **Accession Number** (e.g., 1986:230471)
- **PubMed ID Number** (e.g., 15980585)
- **CAS Accession Number (CAN):** document number in CA Plus (e.g.,148:486341)

SciFinder^n
A CAS SOLUTION

Wildcard searching for substance (Lactobacillus*)

【常見益生菌乳酸菌中英文名對照表 / 市售益生菌乳酸菌中英文名對照】

乳酸桿菌屬 Lactobacillus	簡稱	關於
嗜酸乳桿菌（Lactobacillus acidophilus）	A菌	如優格等發酵乳製品就含有豐富的嗜酸乳桿菌。這種益生菌能協助免疫系統保持腸道細菌的平衡。女性多攝食嗜酸乳桿菌有助於抑制白色念珠菌造成的陰道炎。嗜酸乳桿菌還能調節體內的膽固醇濃度，能在小腸中產生可對抗病原微生物的有益物質。
乾酪乳桿菌（Lactobacillus casei ）	C菌	相當耐酸，能有效地通過胃酸膽鹼的考驗，而有規模的進入腸道定殖。
約氏乳桿菌（Lactobacillus johnsonii ）	LJ菌	健康新生兒之消化道中分離純化，屬人體原生菌種。Salivarius意指“腺體型”最初是由人類腺體中被發現，是人體消化道中的原生菌種，亦是美國食品藥品監督管理局（FDA）表列安全菌種之一。
副乾酪乳桿菌（Lactobacillus paracasei ）	LP菌	耐胃酸及膽鹽，在腸道中定殖效果良好，能促進體內Th1細胞激素分泌，抑制Th2細胞所造成的敏感免疫反應，達到免疫系統平衡。對於異位性皮膚炎等過症狀可能有療效。
鼠李糖乳桿菌GG株（Lactobacillus rhamnosus GG）	LGG菌	是當前世界上研究最多的益生菌，也是首批被證實能夠在人體腸道存活並定殖的益生菌之一。可在血清中增加足夠的細胞間白素-10，降低引起局部性敏感免疫反應的細胞激素形成。能促進益菌生長、降低對乳品或食物的過敏、治療不明原因或急性腹瀉功能等。
洛德乳桿菌（Lactobacillus reuteri ）	R菌	是少數在成人與嬰兒體內皆可發現到的乳酸菌之一，可幫助寶寶腸道細胞的生長，促進益菌繁殖。

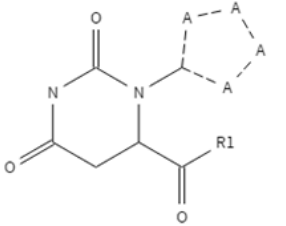
乳酸桿菌屬 Lactobacillus	簡稱	關於
發酵乳桿菌（Lactobacillus Fermentum）	LF菌	常見於發酵動物和植物材料，用作益生菌的商品化發酵乳桿菌菌株包括PCC， ME-3 和 CECT5716。

Find Patent Markush Structures

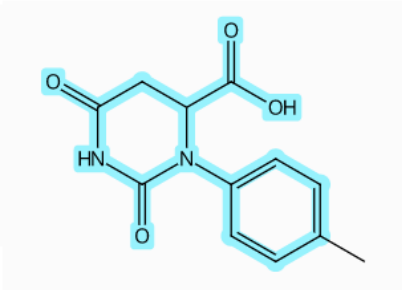
Use a **Patent Markush** search to find structures (including generic structures) matching your query within patent references.

For example:

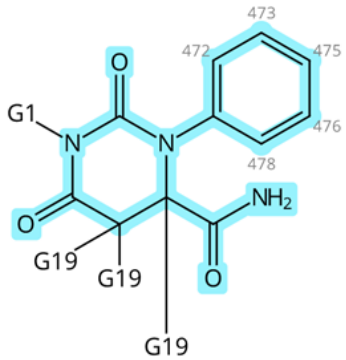
Query



Substructure Match

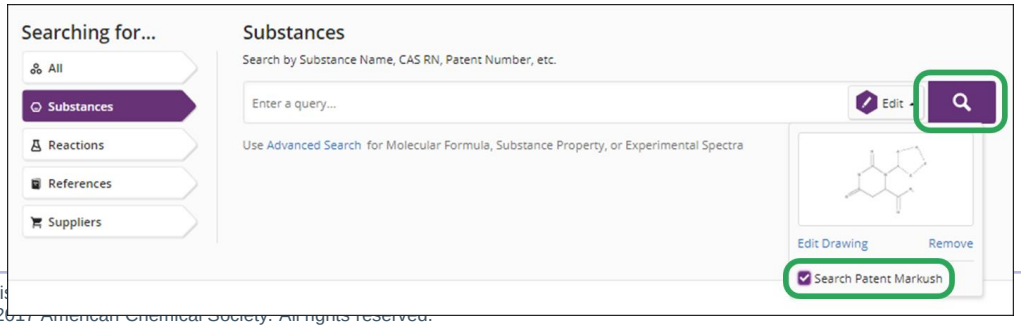


Markush Match

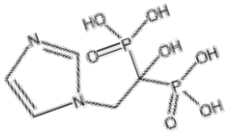


To conduct a Patent Markush search, click the **Substances** search type. Draw the query using a structure editor, and then check the box for **Search Patent Markush**.

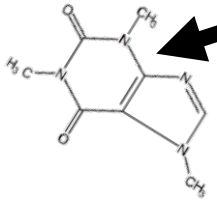
Click the **magnifying glass** to submit the query.



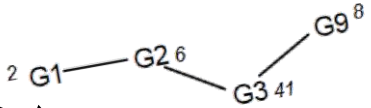
Patent documents are evaluated for specific and Markush structures



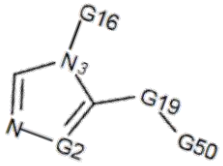
Specific Structures



Patent Document*



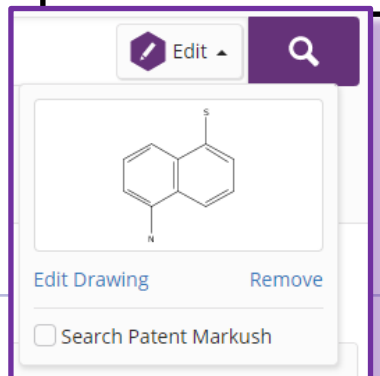
Markush Structures

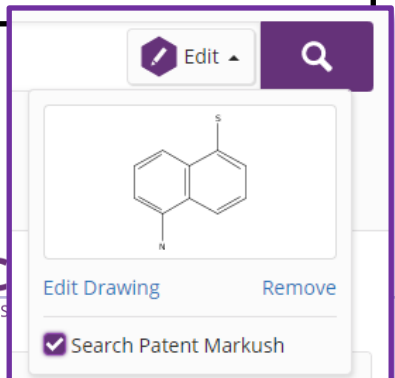


CAS REGISTRYSM
- Specific substances
- Structure-searchable

MARPAT
- Generic substances
- Structure-searchable

CAplusSM Record





Find Reactions by Name, Registry Number, or Document Identifier

Search terms may be:

•**Substance Name:**

- **Example:** benoxaprofen
  Note: Wildcard searching using an asterisk only works for single-search-term, single-word substance name searches (e.g., benoxa*).
- **Example:** methyl ethyl ketone
- **Example:** ß-amyloid
  Note: If your search term contains a Greek letter, see Greek Letters Used in Searching to insert the correct character.

•**CAS Registry Number** (with or without dashes, e.g., 51146-57-7, 51146577)

Notes:

- Square brackets are accepted around RNs (e.g., [51146-57-7], but not single or double quotes.
- You may enter multiple CAS RNs separated by a space, no commas or other punctuation. The search field has a 2000-character limit.

•**Document Identifier:**

- **Patent Number** (no spaces, e.g., US4571400)
- **Accession Number** (e.g., 1986:230471)
- **PubMed ID Number** (e.g., 15980585)
- **CAS Accession Number (CAN):** document number in CA Plus (e.g.,148:486341)

SciFinder^n
A CAS SOLUTION

Find Biosequences

There are three biosequence search types:

- **BLAST (Basic Local Alignment Search Tool):** Search for proteins as well as nucleotides using a set of local alignment algorithms (BLASTn, MegaBlast, BLASTp, tBLASTn, BLASTx).

- **CDR (Complementarity-Determining Region):** Search for antibody and t-cell receptors.

- **Motif:** Search for short patterns in DNA, RNA, or proteins with queries enabled for additional variability.

Find Biosequences - BLAST

Find biosequences that match your query, which can be a protein/nucleotide string or a .txt/.fasta file.

1.Enter or copy and paste a protein/nucleotide string, or upload a sequence file:

- **Single sequence:** .txt file
- **Multiple sequences:** .fasta file

Note: The **maximum number** of sequences is **100**.

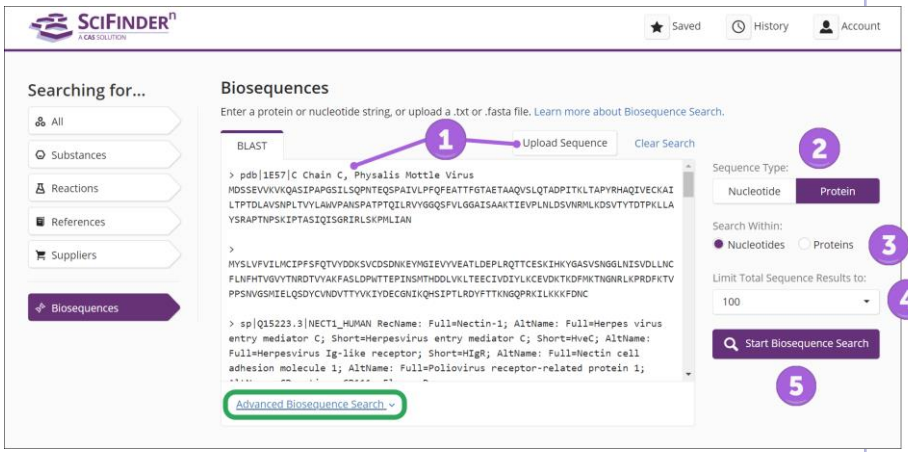
2.Select the appropriate **Sequence Type** based on your query.

3.Select the relevant **Search Within** option.

4.You may a select **total sequence result limit** of 10 to 20,000 (default is 100).

5.Click the **Start Biosequence Search** button.

Click **Advanced Biosequence Search** to find sequences using parameters such as sequence identity percentage, gap costs, and BLAST algorithm.



SciFinder n
A CAS SOLUTION

Find Biosequences - BLAST

Advanced Biosequence Search has default values based on the selected **Sequence Type** and **Search Within** options.

Sequence Type / Search Within	Options (Default Value Shown)										
	Adjust Parameters for Short Sequences	Sequence Identity %	Match with Gaps	Gap Costs	Query Coverage %	Word Size	Reward for Match, Penalty for Mismatch	Scoring Matrix	BLAST Algorithm	E-Value	Exclude Low-Complexity Regions
Nucleotide / Nucleotides	Available	80	No	Existence 5, Extension 2	90	11	2, -3	—	BLASTn	10	No
Nucleotide / Proteins	N/A	—	No	Existence 11, Extension 1	90	6	—	BLOSUM62	BLASTx-fast	10	—
Protein / Nucleotides	N/A	—	No	Existence 11, Extension 1	90	6	—	BLOSUM62	TBLASTn-fast	10	—
Protein / Proteins	Available	—	No	Existence 11, Extension 1	90	3	—	BLOSUM62	BLASTp	10	No

SciFinder^n
A CAS SOLUTION

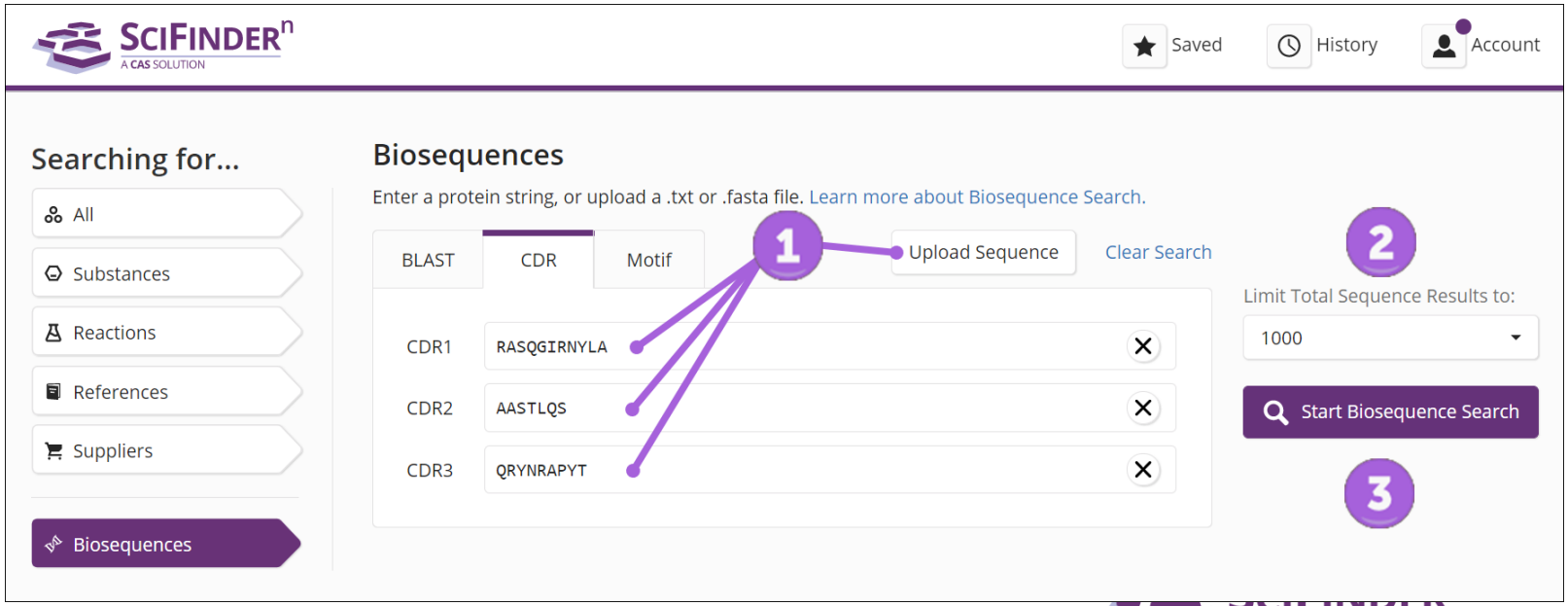
Find Biosequences - CDR

Find biosequences that match your query, which can be a protein string or a .txt/.fasta file.

1.Enter, **copy and paste**, or **upload** a file for **up to three CDRs**.

2.You may a select **total sequence result limit** of 10 to 20,000 (default is 100).

3.Click the **Start Biosequence Search** button.



Find Biosequences - Motif

Find biosequences that match your query, which can be a protein/nucleotide string.

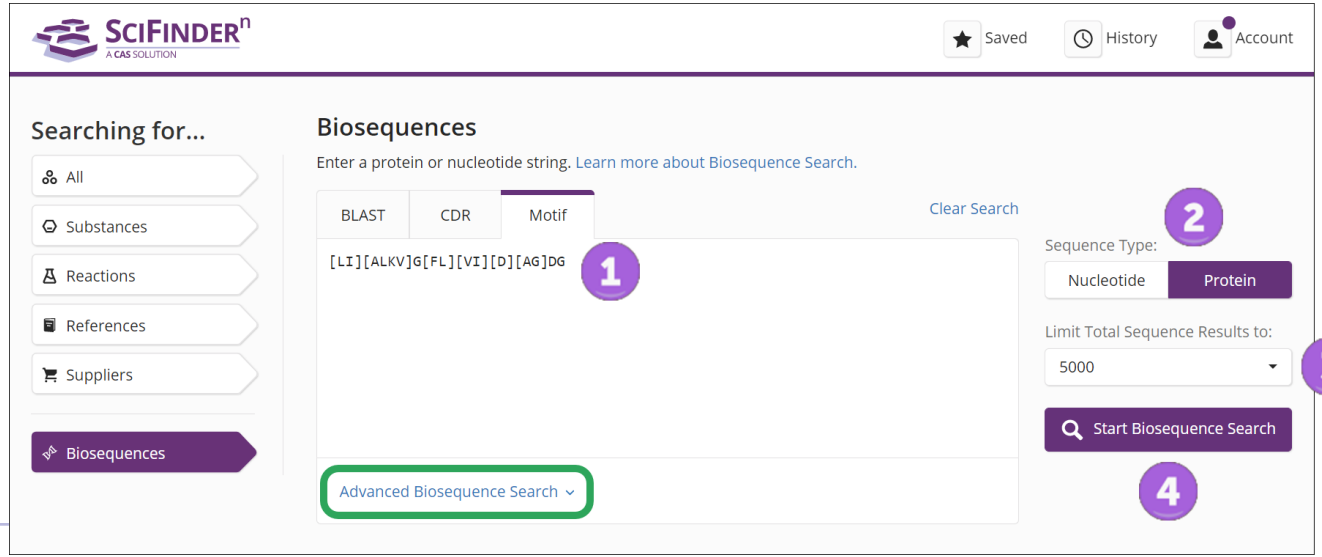
1.Enter or copy and paste a protein/nucleotide string.

2.Select the appropriate **Sequence Type** based on your query.

3.You may a select **total sequence result limit** of 10 to 20,000 (default is 100).

4.Click the **Start Biosequence Search** button.

Click **Advanced Biosequence Search** to find sequences using query coverage % and e-value.



Thank you!

SciFinder^n
A CAS SOLUTION