CAS STNext® Coffee Lecture

# Indexing in U.S. patent full-text databases on CAS STNext®

Jan Baur, CAS



CAS
A division of the
American Chemical Society

# OVERVIEW




CAS
A division of the
American Chemical Society

# USPATFULL, USPAT2

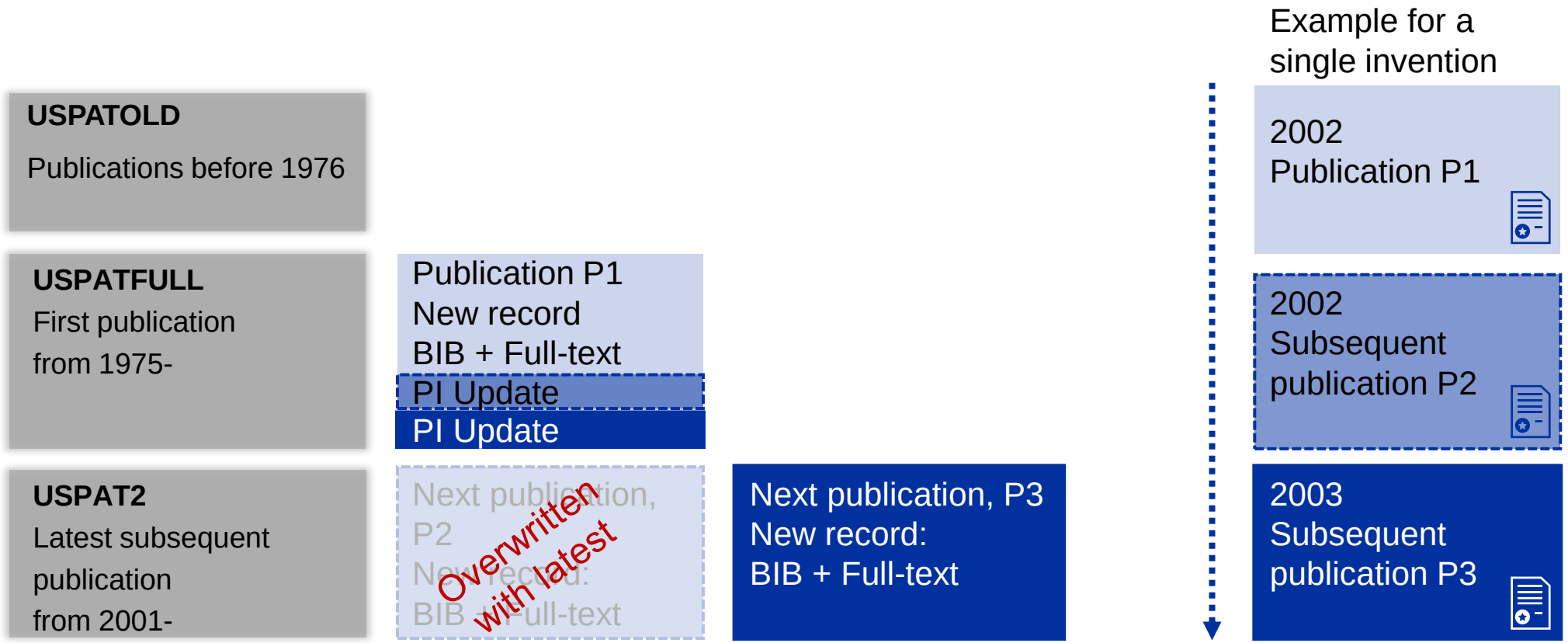
- Full text of granted US patents for original publications from 1975 to the present, US patent applications since 2001
- Companion file USPAT2 contains full texts of the latest publication, when the first publication is available in USPATFULL
- CAS value-add features
  - CAS Registry Number® indexing
  - Controlled Terms
  - Classification codes
- [1]CAS PatentPak®
- Ultimate Owner
- Numeric Property Search (NPS)
- Patent Classifications: CPC, IPC, NCL


[1]For subscribers only




CAS
A division of the American Chemical Society

# The U.S. Patent Fulltext cluster on STNext




1 record represents 1 publication




CAS
A division of the
American Chemical Society

# Added value in USPATFULL

## Search fields shown in parentheses

PatentPak PDF | PatentPak PDF+ | PatentPak Interactive
AN 2016:189913 USPATFULL Full-text
TI ISOXAZOLE HYDROXAMIC ACID COMPOUNDS AS LpxC INHIBITORS
IN FU, Jiping, Danville, CA, UNITED STATES
JIN, Xianming, San Ramon, CA, UNITED STATES
KARUR, Subramanian, Dublin, CA, UNITED STATES
LAPOINTE, Guillaume, San Francisco, CA, UNITED STATES
MADERA, Ann Marie, Dublin, CA, UNITED STATES
SWEENEY, Zachary Kevin, Redwood City, CA, UNITED STATES
PA Novartis AG, Basel, SWITZERLAND (non-U.S. corporation)
UO NOVARTIS AG
UOS Novartis
PI US 20160166548 A1 20160616
AI US 2015-14969930 A1 20151215 (14)
PRAI US 2014-62092402 20141216 (62)
DT Utility
FS APPLICATION
CLMN Number of Claims: 20
ECL Exemplary Claim: 1
DRWN No Drawings
LN.CNT 3441
CAS INDEXING IS AVAILABLE FOR THIS PATENT.
AB This invention pertains generally to compounds of Formula I as described herein and compositions containing such compounds, as well as methods of using such compounds to treat bacterial infections. In certain aspects, the invention provides methods and compositions comprising these compounds for treating infections caused by Gram-negative bacteria.

Links to annotated full-text in PatentPak (PPAK/FA)

Ultimate owner (/UO) and ultimate owner standardized

Indication that CAS indexing is available




CAS
A division of the American Chemical Society

# Added value in USPATFULL (cont.)

## Search fields shown in parentheses

```
CPC     CPCI    C07D0261-08 [I]; C07D0413-12 [I]; A61K0031-42 [I]; A61K0031-422 [I]; A61K0045-06 [I];
                C07D0413-06 [I]; Y02A0050-30; A61P0011-00 [I]; A61P0031-04 [I]; A61P0009-00 [I];
                A61K2300-00; A61K0031-42 [I], A61K2300-00 [I]; A61K0031-422 [I], A61K2300-00 [I];
                A61K0031-42 [I], A61K2300-00 [I]; A61K0031-422 [I], A61K2300-00 [I]
        CPCI-2  C07D0261-08 [I]; C07D0413-12 [I]; A61K0031-42 [I]; A61K0031-422 [I]; A61K0045-06 [I];
                C07D0413-06 [I]; Y02A0050-30; A61P0011-00 [I]; A61P0031-04 [I]; A61P0009-00 [I];
                A61K2300-00; A61K0031-42 [I], A61K2300-00 [I]; A61K0031-422 [I], A61K2300-00 [I];
                A61K0031-42 [I], A61K2300-00 [I]; A61K0031-422 [I], A61K2300-00 [I]
IPC     IPCI    A61K0031-42 [I]; A61K0045-06 [I]; A61K0031-422 [I]; C07D0261-08 [I]; C07D0413-12 [I]
        IPCI-2  A61K0031-42 [I]; A61K0045-06 [I]; A61K0031-422 [I]; C07D0261-08 [I]; C07D0413-12 [I];
                C07D0413-06 [I]
        IPCR    A61K0031-42 [I]; A61K0031-422 [I]; A61K0045-06 [I]; C07D0261-08 [I]; C07D0413-12 [I]
```

Patent classifications

```
CHEMICAL ABSTRACTS INDEXING    COPYRIGHT 2025 ACS on STN
---------------------------
OS
                        PATENT      KIND   DATE
                  ----------------- ---- --------
      CA 165:91706  * US  20160166548  A1  20160616
* CA Indexing for this record included
CC    28-6 (Heterocyclic Compounds (More Than One Hetero Atom))
ST    isoxazole hydroxamic acid LpxC inhibitor Gram neg bacteria infection; antibacterial Gram neg
      bacteria isoxazole hydroxamic acid LpxC inhibitor
IT    Antimicrobial agents
        (co-drugs; prepn. of isoxazole hydroxamic acid compds. as LpxC inhibitors for treating
        infections caused by Gram-neg. bacteria)
IT    Achromobacter
      Achromobacter xylosoxidans
```

CA section indexing (/CC). Help section shows the section hierarchy

Non-controlled keyword indexing (/ST)

Indexing terms (/IT) either show CAS Lexicon expressions (/CT) or CAS Registry numbers (/RN), including text modifiers




CAS
A division of the American Chemical Society

# Added value in USPATFULL (cont.)

## Search fields shown in parentheses

IT Hydroxamic acids
(prepn. of isoxazole hydroxamic acid compds. as LpxC inhibitors for treating infections caused by Gram-neg. bacteria)

IT **157971-99-8**
(LpxC; prepn. of isoxazole hydroxamic acid compds. as LpxC inhibitors for treating infections caused by Gram-neg. bacteria)

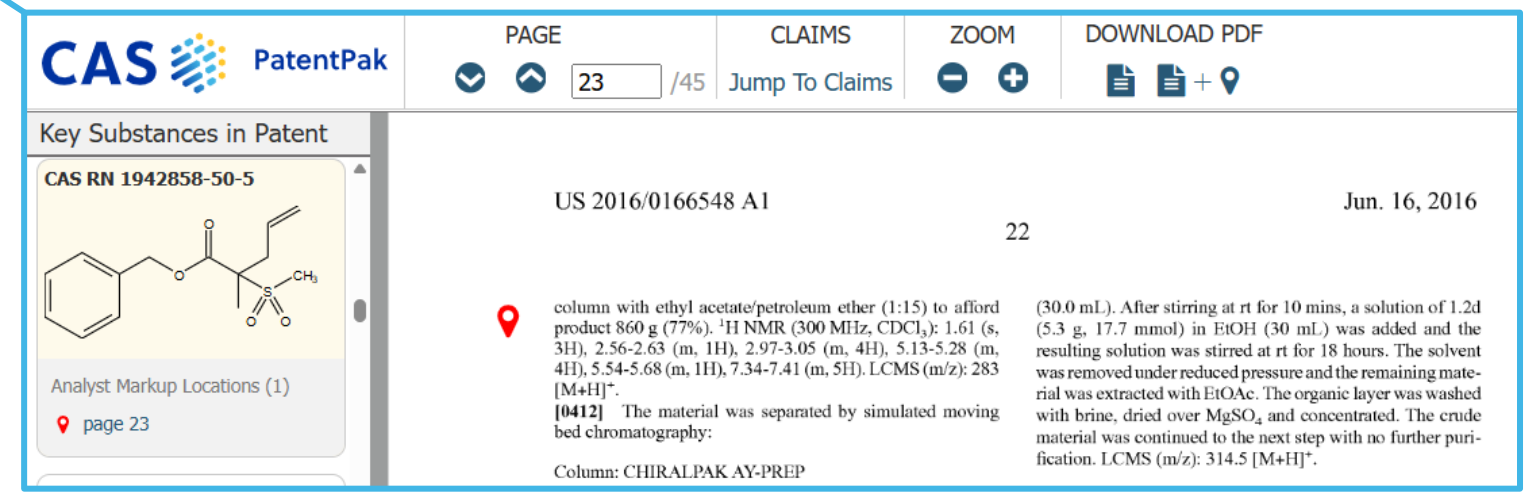
Registry Number indexing (RN), including text modifiers

PPAK

**1942858-42-5** P, Pg 21

**1942858-50-5** P, Pg 23

HITPPAK display only shows PatentPak links for searched substances







CAS
A division of the American Chemical Society

# USE CASES




CAS
A division of the
American Chemical Society

# Scenario 1: Find methods details by searching the entire patent text

```
=> fil reg
=> s fsgsgsgts/sqsp
L1        10772 FSGSGSGTS/SQSP

=> fil caplu
=> s l1 and ultrafiltrat?/bi,clm
L2           23 L1 AND ULTRAFILTRAT?/BI,CLM

=> fil uspatfull
=> s l2
L3          883 L1 AND ULTRAFILTRAT?/BI,CLM

=> s l2 and ultrafiltrat? (s) flow diagram
L4            5 L2 AND ULTRAFILTRAT? (S) FLOW DIAGRAM

=> d kwic 1-tot
```

Search for a subsequence in Registry

In the context of ultrafiltration, 23 entries are found in CAplus

Many more patents are retrieved by searching the full text in USPATFULL.

It's possible to search for the filtration method in flow diagrams




CAS
A division of the American Chemical Society

# Scenario 1: Find methods details by searching the entire patent text (cont.)

```
=> d kwic 1-tot
DETD
TABLE 21
       TG-1101 Manufacturing Process Flow Diagram

       Vial of TG-1101 Working Cell Bank
       ↓
       Inoculum Expansion
       (Shake flasks, 50 L cellbag
. . .  ↓Exchange Membrane Chromatography (AEX)
       (Mustang Q)
       ↓
       Viral Filtration (VF)
       (VIRESOLVE.sup.® viral filter)
       ↓
       Ultrafiltration/Diafiltration (UFDF)
       (30 kDa NMWCO)
       ↓
       Formulation, Filtration
       (Polysorbate 80 addition; 0.2 µm filtration)
. . .  ↓
```

**Ultrafiltration** was found in flow diagram (table 21) of US11965032 B1

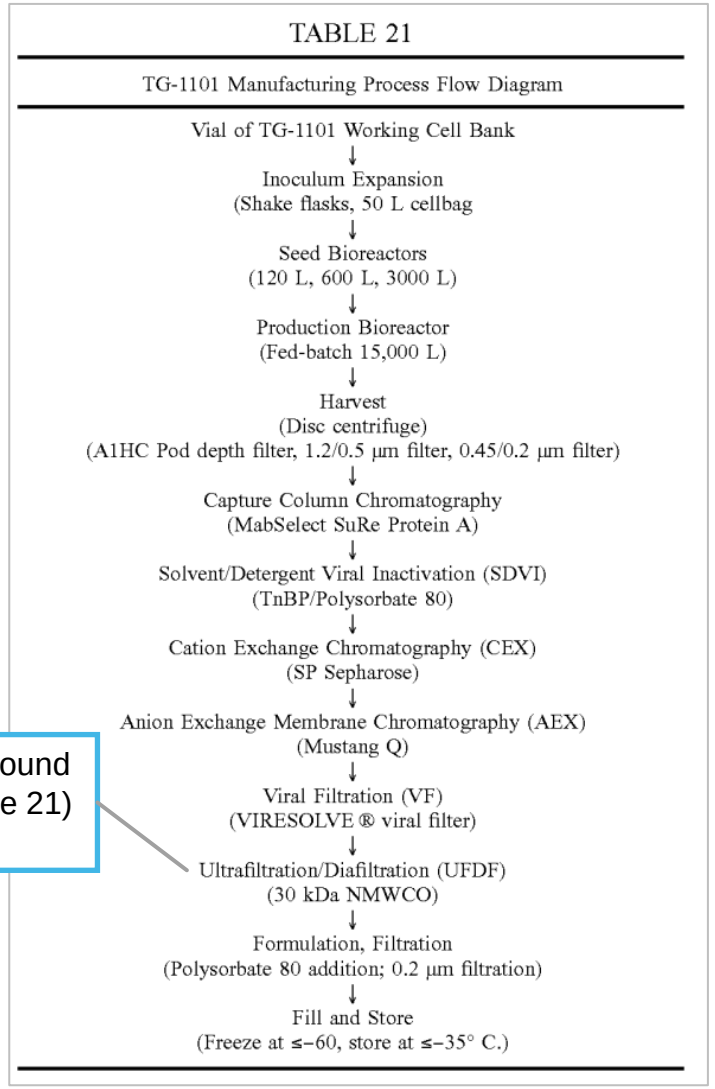


TABLE 21

TG-1101 Manufacturing Process Flow Diagram

Vial of TG-1101 Working Cell Bank
↓
Inoculum Expansion
(Shake flasks, 50 L cellbag
↓
Seed Bioreactors
(120 L, 600 L, 3000 L)
↓
Production Bioreactor
(Fed-batch 15,000 L)
↓
Harvest
(Disc centrifuge)
(A1HC Pod depth filter, 1.2/0.5 µm filter, 0.45/0.2 µm filter)
↓
Capture Column Chromatography
(MabSelect SuRe Protein A)
↓
Solvent/Detergent Viral Inactivation (SDVI)
(TnBP/Polysorbate 80)
↓
Cation Exchange Chromatography (CEX)
(SP Sepharose)
↓
Anion Exchange Membrane Chromatography (AEX)
(Mustang Q)
↓
Viral Filtration (VF)
(VIRESOLVE ® viral filter)
↓
Ultrafiltration/Diafiltration (UFDF)
(30 kDa NMWCO)
↓
Formulation, Filtration
(Polysorbate 80 addition; 0.2 µm filtration)
↓
Fill and Store
(Freeze at ≤−60, store at ≤−35° C.)




CAS
A division of the American Chemical Society

# Scenario 1: Find methods details by searching the entire patent text (cont.)

```
=> d kwic 1-tot
. . .
IT    3008624-25-4P 3008624-26-5P
        (amino acid sequence; anti-CD20 antibody compns.)
```

```
SEQ      1 QIVLSQSPAI LSASPGEKVT MTCRASSSVS YMHWYQQKPG SSPKPWIYAT
        51 SNLASGVPAR FSGSGSGTSY SFTISRVEAE DAATYYCQQW TFNPPTFGGG
       101 TRLEIKRTVA APSVFIFPPS DEQLKSGTAS VVCLLNNFYP REAKVQWKVD
       151 NALQSGNSQE SVTEQDSKDS TYSLSSTLTL SKADYEKHKV YACEVTHQGL
       201 SSPVTKSFNR GEC
```

Peptide **3008624-26-5** was retrieved via substance indexing

# Scenario 2: Search claims of application and grant

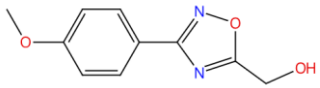
```
=> fil caplus
=> set sfield bi clm
=> s 5372-39-4 and (A61P0031-04+nt/CPC,IPC
     or (anti? (1t) (?microb? or ?bact?)))
     and (escherichia or pseudomonas)
L1             0 5372-39-4/BI,CLM AND (. . .
=> s 5372-39-4 and (escherichia or pseudomonas)/bi,clm
L2             0 5372-39-4/BI,CLM AND (ESCHERICHIA OR
                 PSEUDOMONAS)/BI,CLM

=> fil uspatfull
=> s 5372-39-4/rn and (A61P0031-04+nt/CPC,IPC or (anti? (1t)
     (?microb? or ?bact?))) and (escherichia or pseudomonas)
L3             2 5372-39-4/RN AND (A61P0031-04+NT/CPC,IPC . . .

=> d bib hitstr kwic 1-2
```

**(1T)** retrieves **antibacterial** and **anti-bacterial**

Set the standard search fields to BI and CLM

Using patent classifications and efficient text queries to search for antibacterials and limit to bacterial genera

RN 5372-39-4:




2 patents are found in USPATFULL

**Why didn't we find these patents in CAplus?**




CAS
A division of the American Chemical Society

# Scenario 2: Search claims of application and grant (cont.)

```
=> d bib hitstr kwic 1-2
L3   ANSWER 1 OF 2  USPATFULL on STN
PatentPak PDF
AN     2025:107232  USPATFULL Full-text
TI     [3-(4-METHOXYPHENYL)-1,2,4-OXADIAZOL-5-YL]METHANOL AS
       AN ANTITUMOR AND ANTIMICROBIAL COMPOUND
IN     ALI, MAI MOSTAFA KHALAF, AL-AHSA, SAUDI ARABIA
. . .
USPA   KING FAISAL UNIVERSITY, AL-AHSA, SAUDI ARABIA
PI     US 20250091999      A1  20250320
AI     US 2024-18429111    A1  20240131 (18)
RLI    Division of Ser. No. US 2023-18368138, filed on 14 Sep 2023, ABANDONED
DT     Utility
FS     APPLICATION
CAS INDEXING IS AVAILABLE FOR THIS PATENT.
```

This is the substance of interest




CAS
A division of the American Chemical Society

# Scenario 2: Search claims of application and grant (cont.)

**=> d bib hitstr kwic 1-2**

. . .

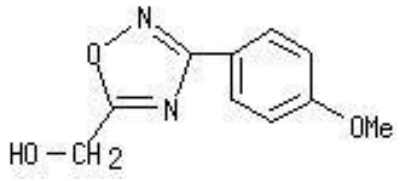
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

IT **5372-39-4P**, [3-(4-Methoxyphenyl)-1,2,4-oxadiazol-5-yl]methanol
(prepn. of [3-(4-methoxyphenyl)-1,2,4-oxadiazol-5-yl]methanol as an
antitumor and antimicrobial compd.)

RN 5372-39-4 USPATFULL

CN 1,2,4-Oxadiazole-5-methanol, 3-(4-methoxyphenyl)- (CA INDEX NAME)

STR 5372-39-4







CAS
A division of the American Chemical Society

# Scenario 2: Search claims of application and grant (cont.)

```
=> d bib hitstr kwic 1-2
. . .
CLM    What is claimed is:
 .  .  .  9, wherein the microbial infection is caused by one or more
       Gram-negative bacterial strains selected from the group consisting of
       Escherichia coli and Pseudomonas aeruginosa; one or more
       Gram-positive bacterial strains selected from the group consisting of
       Bacillus cereus and Staphylococcus aureus; one or.  .  .
```

**Why didn't we find this patent in CAplus?**

The US application claims the infectious bacterial species, the US grant does not...



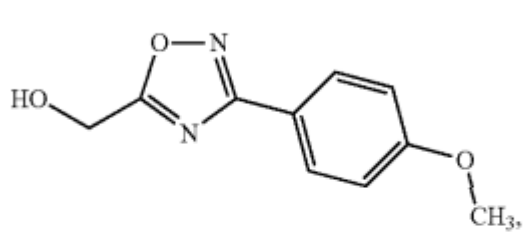

CAS
A division of the American Chemical Society

# Scenario 2: Search claims of application and grant (cont.)

US 2025/0091999 A1

**1-4**. (canceled)

**5**. A method of treating cancer in a patient comprising administering to the patient a therapeutically effective amount of [3-(4-methoxyphenyl)-1,2,4-oxadiazol-5-yl] methanol having the formula I:




I

wherein the cancer is prostate cancer or osteosarcoma.

**6-7**. (canceled)

**8**. A method of treating a microbial infection in a patient comprising administering to a patient in need thereof a therapeutically effective amount of a [3-(4-methoxyphenyl)-1,2,4-oxadiazol-5-yl]methanol compound.

**9**. The method of treating the microbial infection of claim **8**, wherein the microbial infection is caused by one or more bacteria or fungi.

**10**. The method of treating the microbial infection of claim **9**, wherein the microbial infection is caused by one or more Gram-negative bacterial strains selected from the group consisting of *Escherichia coli* and *Pseudomonas*

US 12,180,172 B1

We claim:

**1**. A method of making a [3-(4-methoxyphenyl)-1,2,4-oxadiazol-5-yl]methanol compound, the method comprising:

adding carbonyldiimidazole (CDI) to a solution of [5-methyl-2-(propan-2-yl)phenoxy]acetic acid in acetonitrile to obtain a reaction mixture;

adding N'-hydroxy-4-methoxybenzene-1-carboximidamide to the reaction mixture with reflux until reaction completion to obtain a precipitate; and

obtaining the [3-(4-methoxyphenyl)-1,2,4-oxadiazol-5-yl]methanol compound.

. . .

**6**. The method of making the [3-(4-methoxyphenyl)-1,2,4-oxadiazol-5-yl]methanol compound of claim **1**, wherein the [3-(4-methoxyphenyl)-1,2,4-oxadiazol-5-yl]methanol compound is obtained in a 87% yield.

**7**. A method of making a [3-(4-methoxyphenyl)-1,2,4-oxadiazol-5-yl]methanol compound, the method comprising:

refluxing a solution of 3-(4-chlorophenyl)-5-{[5-methyl-2-(propan-2-yl)phenoxy]methyl}-1,2,4-oxadiazole in acetonitrile until reaction completion to obtain a precipitate; and

obtaining the [3-(4-methoxyphenyl)-1,2,4-oxadiazol-5-yl]methanol compound.

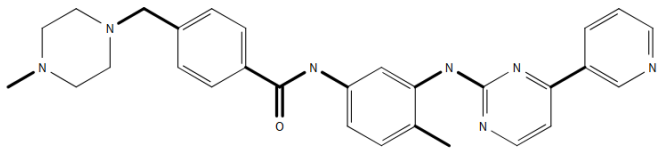
The grant was published first, thus only the B document was indexed in CAplus, which claims the method only. Patent claims in CAplus are available for the basic patent.

CAS
A division of the American Chemical Society

# Scenario 3: Numeric property search

## Gleevec formulations with claimed dosage information, a family structure search in Registry was already executed

Structure query



```
=> fil uspatfull
=> s l2
L3          4596 L2
```

1-50 mg per kg and day is the default unit, but also mg per day will be found.

```
=> s l3 and 1-50/doa (s) (administ? or dos?)/clm
L4            79 L3 AND 1 MG/KG/DAY - 50 MG/KG/DAY /DOA
                 (S) (ADMINIST? OR DOS?)/CLM
```

The NPS search for dosage per day (DOA) itself cannot be restricted to specific text fields but the text refinement allows us to focus on claims.

(S) forces the NPS hit to occur in the same claim sentence

```
=> d bib hit 1-tot
```

HIT will show the full hit claim




CAS
A division of the American Chemical Society

# Scenario 3: Numeric property search (cont.)

## Gleevec formulations with claimed dosage information, a family structure search in Registry was already executed

```
=> d bib hit 1-tot
L4   ANSWER x OF 79  USPATFULL on STN
PI     US 20050119352      A1  20050602
       US 7361691          B2  20080422
CLM    What is claimed is:
       19. The method of claim 1, wherein the oncogenic kinase modulator is
       administered at a dosage from about 10 mg/day to about 2000 mg/day.

IT    4707-32-8, β-Lapachone   4707-32-8D, β-Lapachone, derivs. and
      analogs 152459-95-5, Imatinib
        (cell cycle checkpoint activator combined with oncogenic kinase
        modulator for treating cancer)
```

The mentioned kinase modulator is named in claim 6

"6. The method of claim 5, wherein the Bcr-Abl signal transduction pathway modulator is imatinib."

The hit dosage range of 10-2000 mg/day overlaps with our search of 1-50 mg/kg

Indexing terms show Imatinib, the oncogenic kinase modulator of interest.

**(!)** Often, compounds are not named explicitly in the respective claim, so searching it will miss relevant hits. Often only its pharmacological or formulation-related function is articulated, e.g. 'modulator' or 'active ingredient'.




CAS
A division of the American Chemical Society

# Connections
# that matter

**Jan Baur**
Manager, Customer Success Specialists, EMEA
jbaur@acs-i.org




CAS
A division of the
American Chemical Society