

Chemical Substance Names in CA Search®: Chemical Abstracts® (File 399)

A specific chemical substance can be searched by the CAS® Registry Number for the substance, or by a non-systematic, common or trade name of the substance. These types of substance terms occur in different parts of the online records and, if used individually, give varying degrees of comprehensiveness to a search strategy. The highest possible recall for information on specific chemical substances is achieved using a combination of CAS Registry Number(s) and the various names used to designate the substance.

You can locate CAS Registry Numbers through a search of CHEMSEARCH™ (File 398), which includes all chemical substances registered by CAS since 1957.

Thus comprehensive retrieval of information about formaldehyde is achieved by the following search, e.g.,

```
?S FORMALDEHYDE OR RN=50-00-0
```

```
      74103 FORMALDEHYDE (SEE ?IGNOTE)
      68160 RN=50-00-0 (SEE ?IGNOTE)
S1 107948 FORMALDEHYDE OR RN=50-00-0
```

```
?T1/TI, ID, DE/1, 20
```

```
1/TI, ID, DE/1
```

```
DIALOG(R)File 399: CA SEARCH(R)
```

```
(c) 2009 American Chemical Society. All rights reserved.
```

Quick determination of formaldehyde residue in pharmaceutical intermediates by homemade microspectrometer

Identifiers: formaldehyde residue quick detn pharmaceutical intermediate microspectrometer

Descriptors:

Absorption spectroscopy ...

quick detn. of **formaldehyde** residue in pharmaceutical intermediates using microspectrometer

CAS Registry Numbers:

50-00-0 analysis, quick detn. of **formaldehyde** residue in pharmaceutical intermediates using microspectrometer

4338-98-1 10138-04-2 66357-35-5P quick detn. of **formaldehyde** residue in pharmaceutical intermediates using microspectrometer

```
1/TI, ID, DE/20
```

```
DIALOG(R)File 399: CA SEARCH(R)
```

```
(c) 2009 American Chemical Society. All rights reserved.
```

Preparation of chitosan based scaffolds using supercritical carbon dioxide

Identifiers: chitosan **formaldehyde** scaffold supercrit carbon dioxide tissue engineering

Descriptors:

Crosslinking... Hydrogels... Pore size distribution... Pore structure... Porosity... Schiff bases... Supercritical fluids...

Surface area... Thermal stability... Tissue engineering ...
 prepn. of chitosan-**formaldehyde** based scaffolds using supercrit.
 carbon dioxide
 Tissue engineering ...
 tissue engineering scaffolding; prepn. of chitosan-**formaldehyde**
 based scaffolds using supercrit. carbon dioxide
 Pore size ...
 vol.; prepn. of chitosan-**formaldehyde** based scaffolds using
 supercrit. carbon dioxide
CAS Registry Numbers:
 124-38-9 biological studies, prepn. of chitosan-**formaldehyde** based
 scaffolds using supercrit. carbon dioxide
 76755-58-3P prepn. of chitosan-**formaldehyde** based scaffolds using
 supercrit. carbon dioxide

Segmentation of chemical substance names. The names of chemical substances are often compound words composed of chemically significant words or word fragments that describe structure and composition. Some segmentation has been performed by document analysts in creating the keyword phrases that make up the Identifier field. This segmentation, however, is provided to reflect the author's emphasis on certain substituents and is present only in identifiers. To provide more comprehensive segmentation for the substance names in all Dialog chemical information files, an algorithmic segmentation procedure has been developed. In the CA SEARCH files, algorithmic segmentation has been applied to chemical substance names in the **Descriptor**, **Identifier**, and Title fields. This algorithmic segmentation is different from, and is provided in addition to, any segmentation performed by a document analyst for keyword phrases. Thus chemical substance names have been doubly indexed by both the compound words and their chemically significant segments.

For example, the chemical substance name "benzoylnitroacetanilide" can be searched by the following terms:

```

BENZOYLNITROACETANILIDE
BENZOYL
      NITRO
            ACET
                  ANILIDE
                        ACETANILIDE
                              NITROACETANILIDE
  
```

Segmentation of chemical names allows searching by any single segment to retrieve all compound names that contain that segment either as a complete word or as part of a longer, compound word. For example, the statement SELECT NITRO retrieves all records with substance names that include NITRO as part of the name (e.g., nitrobenzene, benzoylnitroacetanilide, trnitrotoluene, etc.).

Recombined segments. As shown in the example above, the basic segments of a chemical name are supplemented with additional index terms created by recombining the segments from right to left. For the name "benzoylnitroacetanilide," the additional terms created by recombining segments are ACETANILIDE and NITROACETANILIDE.

Limiting to un-segmented words. A chemical substance search term may be a complete term in the original context for one of three reasons:

- It completely describes the chemical substance.
- It was emphasized by the author, and was, therefore, segmented manually in the Identifier (keyword) field or in the Descriptor (index modification) field of the record.
- It was part of a systematic name, but offset by punctuation.

In some cases, it is useful to be able to search a term used as a full word, and not as a segment of a longer word. In CA SEARCH, the /FF suffix is used to search a full word in the Basic Index. For example to retrieve records that discuss acetanilide and exclude occurrences of nitroacetanilide or benzoynitroacetanilide, enter

SELECT ACETANILIDE/FF.

Chemical segments and proximity operations. Chemical segments are indexed in the Descriptor (/DE), Identifier (/ID), and Title (/TI) fields and can be searched with or without using the field suffixes in the Basic Index.

(T) operator. Segments derived from the same original term have term (T) proximity to one another. Use the (T) operator to "recombine" terms that have been generated from a longer word. Term order within a word is not maintained so that SELECT

CHLORO(T)BENZENE retrieves the same records as SELECT BENZENE(T)CHLORO.

(W) or (nW) operator. Use the (W) operator with segments to retrieve all items retrieved by the (T) operator, plus those records in which the segments are in adjacent words, in the order specified. Thus SELECT CHLORO(W)BENZENE retrieves additional records containing such phrases as "chloromethylstyrene diisopropenylbenzene."

(N) or (nN) operator. Use the (N) operator with segments to retrieve all items in which the segments are in adjacent words, in any order. The (N) operator retrieves all of the items retrieved by the (W) operator, plus additional records containing such phrases as "benzene chloromethane."

(S) operator. Use the (S) operator with segments to retrieve all items in which the segments are anywhere within the title, a descriptor phrase, or an identifier phrase. The (S) operator retrieves all of the items retrieved by the (N) operator, plus additional records in which the two segments may be separated by multiple words, but both are in the title, in a single descriptor, or in a single identifier.

The difference in retrieval using the various operators with the same two terms is illustrated below:

Set	Items	Description
S1	19765	CHLORO(T)BENZENE
S2	22414	CHLORO(W)BENZENE
S3	23850	CHLORO(N)BENZENE
S4	30006	CHLORO(S)BENZENE

DIALOG Corporate Headquarters
2250 Perimeter Park Drive
Suite 300
Morrisville, NC 27560
1-800-3 DIALOG

DIALOG Europe, Middle East, Africa
St. Andrews House
18-20 St. Andrew Street
London EC4A 3 AG, UK
+44 (0)20 7832 1729

Contact the Dialog Knowledge Center
Within North America 1 800 3 DIALOG (334 2564)
Outside North America 00 800 33 DIALOG (33 34 2564)
Email: Customer@dialog.com

DSACAS073009