

Identifier (/ID, /IF) Field in CA Search® (File 399)

The Identifier field in CA SEARCH is composed of uncontrolled terms from the keyword phrases that appear in the subject indexes of individual weekly issues of *Chemical Abstracts*.

Keyword phrases were originally selected by Chemical Abstracts Service document analysts from the title of the document, the abstract, or the source document to reflect author emphasis. Currently, computer analysis of the title, abstract and parts of the source document is the primary way that the keyword phrases are generated. Document analysts follow some guidelines, and these are incorporated into the computer algorithms in the construction and content of the keyword phrases, but do not attempt to rigidly control the content or format of keyword phrases. Familiarity with the guidelines, however, can help the searcher construct successful search strategies.

Keyword phrases (Dialog Identifiers) are not sentences. There is not necessarily any syntactical relationship between or among keywords in a keyword phrase. Several keyword phrases may be required to completely describe the subjects and substances of interest in an abstract.

CA Section Titles. Words in the CA Section Title, described in the separate document **CA Section Title (/SH) and CA Section Code (SC=) Fields in CA SEARCH**, are not routinely repeated in keyword phrases. For example, documents appearing in CA Section 14, "Toxicology", do not routinely include the keyword "toxicology," e.g.,

```
?S TOXICOLOGY
  S1 423439 TOXICOLOGY (SEE ?BROAD)
?S S1/ID
  S2      307 S1/ID
?S S1/SH
  S3 421240 S1/SH
```

Singular and noun forms used. Identifiers are generally in the singular form, except for certain Greek and Latin plurals in common use in the literature, e.g., spectra, bacteria, data, and algae. The noun form of a word is preferred to the adjectival form, e.g., **herbicide acetanilide, fungicide dithiocarbamate**. Articles, conjunctions, and prepositions are not used in Identifiers.

Punctuation marks and hyphenated terms. Punctuation marks, e.g., periods, hyphens, commas, parentheses, and brackets are not used in Identifiers. Normally hyphenated words are given as two words, e.g., X ray, half life, surface active, etc. Over the time span of the database, some hyphenated words have become single words, as used in Identifiers, e.g., **crosslinking agent siloxane, nonnewtonian liq dispersion model**. Because these words may still be hyphenated in other fields of the record, search both forms for comprehensive results, e.g.,

```
?S CROSSLINKING OR CROSS(W)LINKING
  101007 CROSSLINKING (SEE ?IGNOTE)
  77089  CROSS
  4721   LINKING
  2962   CROSS(W)LINKING
  S4 101845 CROSSLINKING OR CROSS(W)LINKING

?T/ID
  4/ID/1
DIALOG(R)File 399: CA SEARCH(R)
(c)American Chemical Society. All rights reserved.
  IDENTIFIERS: photothermog material alkoxy titanium
crosslinking agent, acid anhydride polymer binder
crosslinking agent, sulfone compd antifoggant
photothermog material
```

Abbreviations. Abbreviations are frequently employed to make records more compact. *CAS Standard Abbreviations & Acronyms*, available from CAS, forms the basis for the abbreviations (with certain restrictions discussed below) that are acceptable in Identifiers. Since no punctuation marks are used, all acceptable abbreviations appear without a period. Words that would have a single-letter abbreviation are usually spelled out, e. g., " second, " not "s." Abbreviations that are themselves words, e.g., "at" for "atomic," are restricted and appear only occasionally. Words formed by adding prefixes to words that are normally abbreviated, are also abbreviated, e. g., **photodegrdn, nonaq**, etc.

The Dialog system uses the related term feature to incorporate cross references in the Basic Index to frequently used abbreviations. To verify the presence of an abbreviation for a term, use the EXPAND command with the term enclosed in parentheses, e.g., EXPAND (AQUEOUS). If an abbreviation is shown, use both the term and its abbreviation in the search strategy to ensure comprehensive results, e.g.,

```
?E (AQUEOUS)
Ref  Items Type  RT  Index-term
R1   112265      1  *AQUEOUS
R2   25882      R    AQ (AQUEOUS)

?S R1 OR R2
      112265  AQUEOUS
      25882   AQ
S5   121636  "AQUEOUS" OR "AQ"
```

Element symbols. Element names, not symbols, are used in Identifiers, e.g., **iron nickel tellurium system, tungsten sepn iridium**, etc.

Numbers and special symbols. Roman numerals may be present in Identifiers, but are not used to indicate valence, e.g., **group IIIB nitroxide ESR**. Valence is indicated by Arabic numerals, e.g., **iron 2 sulfate calcination**. Superscript and subscript Arabic numerals appear on the line, e.g., **vitamin D₃ metab lipid**. Special symbols, i.e., Greek letters, are represented by mnemonic, alphabetic strings, e.g., **beta prodn neutron decay, gamma ray vulcanization**.

Taxonomic names. The genus name alone, without species name, is usually given, e.g., **Aspergillus genetics choline sulfate**.

Chemical substance names. Locants and isomeric designations, when separated from a substance name by punctuation, are not used in Identifiers. Thus the author name "2,6-dichloropyridine" in a title, would appear as "dichloropyridine" in the Identifier field. The exclusion of locants and isomeric designations tends to make the Identifiers more generic (less specific) than in the original document.

Classes of compounds are generally referred to in the singular form, e.g., **haloanthanthrone, amination dichloropyridine, polyester fiber elasticity**, etc.

Each ligand and central atom or parent anion of an inorganic substance is entered as a separate word, e.g., **nitrosyl chloro aquo ruthenate**. When the name of an organic ligand is too long to be used as an entry with the central metal atom, the ligand name appears alone.

Classical group names such as alkali metal, rare earth, chalcogen, are used when all or most of the elements in the group are discussed in the indexed document, e.g., **memory switching chalcogenide film**.

Radical groups are entered by name, not abbreviation, e.g., **propyl alc acetyl**.

When there is no author emphasis on a particular component or substituent of a chemical substance, the substance name is entered unsegmented, e.g., **fluoropropene copolymer blend, benzoylnitroacetanilide reagent iron detn**, etc. CAS does segment some names to reflect the author's emphasis on certain substituents or ring components, e.g., **phenyl imidazothiazole anthelmintic, pentane bromo dielec relaxation, pyridine dichloro hydrogenation**, etc.

Segmentation. Algorithmic segmentation of chemical substance terms in the Identifier field provides additional search terms. For details, please see the separate document on [Chemical Substance Names in CA SEARCH](#).

Searching Identifiers. Identifiers can be searched as a complete phrase or by individual words of the phrase using proximity operators. Searching by words with (S) or (N) proximity operators is recommended because the words in an Identifier may have been rotated to provide additional index entries in the printed *Chemical Abstracts*, and because only unique Identifiers are retained in the online file, e.g.,

```
?S IRIDIUM(N)TUNGSTEN(S)ALLOY/ID
```

```

      14843  IRIDIUM/ID
      73864  TUNGSTEN/ID
      249057  ALLOY/ID
S20      6  IRIDIUM(N)TUNGSTEN(S)ALLOY/ID

```

```
?T/ID
```

```

6/ID/1
DIALOG(R)File 399: CA SEARCH(R)
(c) American Chemical Society. All rights reserved.
  IDENTIFIERS: thorium doped iridium tungsten alloy
weld property, grain structure weldability laser arc
GTA

```

```
?T
```

```

6/ID/2
DIALOG(R)File 399: CA SEARCH(R)
(c) American Chemical Society. All rights reserved.
  IDENTIFIERS: rare earth hafnium tungsten iridium
alloy cathode

```

Un-segmented Identifiers. Use the /IF suffix to search a term as a complete, unsegmented Identifier. For example SELECT BENZENE/IF retrieves records in which the Identifiers refer to the substance benzene rather than a substituted benzene. Use the /IF suffix also with a term that contains segments, but could also be a recombined term from a longer chemical name. For example SELECT CHLOROBENZENE/IF retrieves records in which the Identifiers refer to the substance chlorobenzene, but generally excludes those records that contain Identifiers that refer to dichlorobenzene, trichlorobenzene, etc.

Dialog (www.dialog.com), a pioneer of online searching in advance of the internet, is the global leader in providing mission-critical information that drives research in science, engineering, business and intellectual property. The online services provided by Dialog — offered through the Dialog® and DataStar® brands — give users the ability to search through more than 1.5 billion unique records from the world's most authoritative publishers. Sources include in-depth, deep web repositories of scientific, technical and biomedical data, patents and trademarks, and thousands of industry, news and company information sources. Librarians, researchers and information users from corporate, professional and government organizations in more than 100 countries prize Dialog for depth of content, precision searching and the ability to provide complete solutions.

Dialog, LLC, part of the ProQuest (www.proquest.com) family of brands, is headquartered in Morrisville, North Carolina, with offices in Sunnyvale, California, and London.

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

DSACAS080309