CAS STNext® COFFEE LECTURE

ADJUST SUBSTANCE SEARCH PRECISION WITH ELEMENT INFORMATION IN CAS REGISTRY®

Jan Baur, ACS International

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Agenda

Element information search fields

AND versus (P) operator

Search examples

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Element information

- Every substance with a molecular formula in CAS REGISTRY® has searchable element information
- Element information can be used to
  - Retrieve relevant compounds with a common composition
  - Refine a large set of compounds
  - Create a subset for very broad structure searches
- Element information search options
  - Number of elements present
  - Specific elements present

Select element information search fields

Search field	Definition	Search Example
ELC.SUB	Total number of elements in a substance	5/ELC.SUB
ELC	Total number of elements in a single-component substance or in a single component of a multi-component substance	5/ELC
ATC	Total number of atoms in a single-component substance or in a single component of a multi-component substance	23/ATC
<Element Symbol>	Count for an individual elements per substance	10/C and 2/O
ELS	Specific elements in a substance	C/ELS AND O/ELS
ELF	Elements in a component formula. Elements must be in Hill order, separated by spaces	C Cl F Mn O/ELF
PG	Elements of a periodic group and select periods, e.g. A1 (alkaline metals) or LNTH (lanthanides), except C and H	A2/PG

Periodic group search index

																		A8
A1	**A2**												**A3**	**A4**	**A5**	**A6**	**A7**	He
Li	Be												B		N	O	F	Ne
Na	Mg		**B3**	**B4**	**B5**	**B6**	**B7**		**B8**		**B1**		**B2**	Al	Si	P	S	Cl	Ar
K	Ca	**T1**	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu		Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	**T2**	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag		Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	**T3**	La	Hf	Ta	W	Re	Os	Ir	Pt	Au		Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra		Ac																

			LNTH	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu		
			ACTN	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr		
			SHEL	Rf	Db	Sg	Bh	Hs	Mt										

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Expressions to retrieve the displayed compound

Iron nickel ruthenium strontium oxide
Fe10.6 Ni0.7 Ru0.7 Sr O19

Specific element count
`S 1/Fe AND 1/Ru`

Molecular formula (Hill order)
`S Fe.Ni.O.Ru.Sr/MF`

Element count for an entire substance
`S 5/ELC.SUB`

Element symbol (/ELS)
`S (Ru AND O AND Fe)/ELS`

Specific ratios per element
`S 10-12/Fe AND 0.7/Ni AND 0.7/Ru AND 1/Sr AND 19/O`

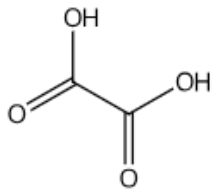
Element symbol (/ELS) and element count
`S (Ru AND O AND Fe)/ELS AND 5/ELC`

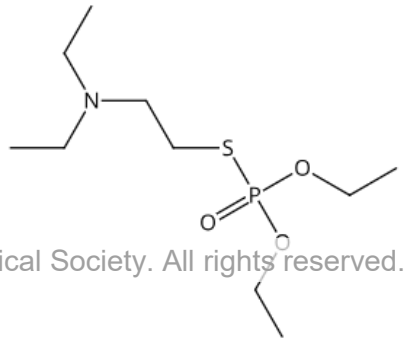
Operators: AND versus (P)

- **AND** searches in all parts of the molecular formula, i.e. in any component

```
=> s 2/H and 4/o and 1/S and 3/elc

L1       59810 2/H AND 4/O AND 1/S AND 3/ELC
```

```
CN   Tetram monooxalate
MF   C10 H24 N O3 P S . C2 H2 O4

CRN  144-62-7
CMF  C2 H2 O4

CM   2
CRN  78-53-5
CMF  C10 H24 N O3 P S
```

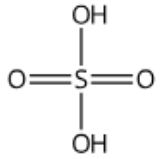


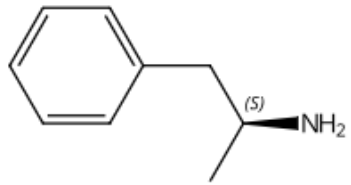


(P) restricts the search to the same component

```
=> s 2/H (p) 4/o (p) 1/S (p) 3/elc

L2       37098 2/H (P) 4/O (P) 1/S (P) 3/ELC
```

```
CN   Maxiton sulfate
MF   C9 H13 N . 1/2 H2 O4 S

CM   1
CRN  7664-93-9
CMF  H2 O4 S

CM   2
CRN  51-64-9
CMF  C9 H13 N
```





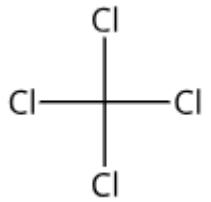
Use Case 1: Halogenated compounds

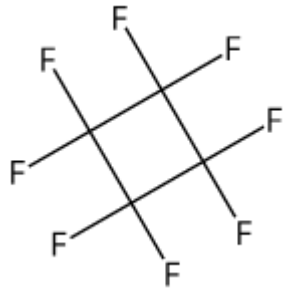
Task: Retrieve halogenated compounds in CAS REGISTRY containing only carbon and halogen and no other elements. Multi-component substances and polymers are acceptable.

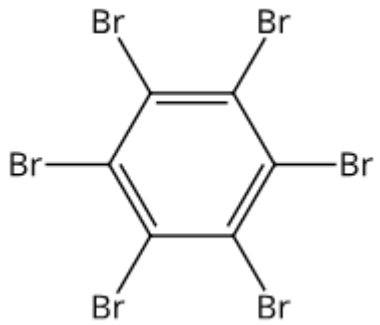
C and X in the same component

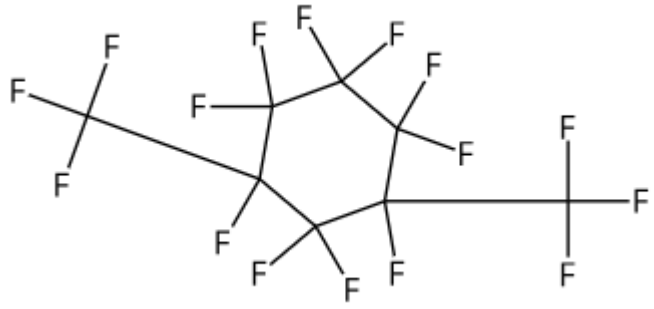
2 elements in the entire substance

```
=> S (C (P) X)/ELS and  2/ELC.SUB

      208739849 C/ELS
       97187296 X/ELS
       93122254 (C (P) X)/ELS
          522368 2/ELC.SUB
L1          6898 (C (P) X)/ELS AND  2/ELC.SUB
```









FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Use Case 2: Magnesium substances in formulations

Task: Broadly define chemically specified magnesium substances present in the CAS FORMULATIONS database

```
=> S MG/ELS
```

All substances which contain Mg

```
=> S L1 NOT ALLOY/CI
```

Removing alloys

```
       1332049 ALLOY/CI
L2      141062 L1 NOT ALLOY/CI
```

```
=> S L2 NOT 1/ELC.SUB
```

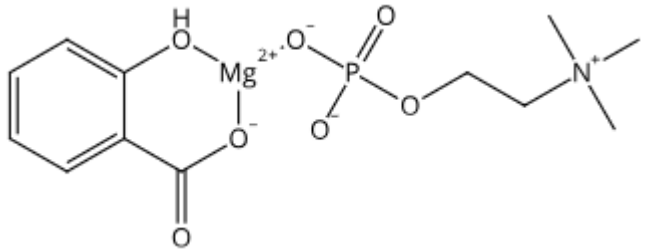
Removing magnesium as an element

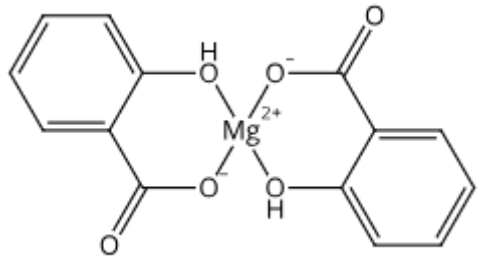
```
         36371 1/ELC.SUB
L3      140803 L2 NOT 1/ELC.SUB
```

```
=> S L3 AND CASFORM?/LC
```

Restricting to substances present in CAS Formulations. The result set contains single and multi-component substances

```
       2080377 CASFORM?/LC
L4        1218 L3 AND CASFORM?/LC
```





FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Use Case 3: Alkoxylated cetyl alcohols

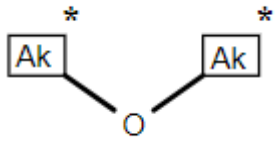
Task: Find structurally defined alkoxylated cetyl alcohols. Use element information and a structure search within this subset.

Components consisting of C, H and O

One oxygen only

More than 16 carbon atoms

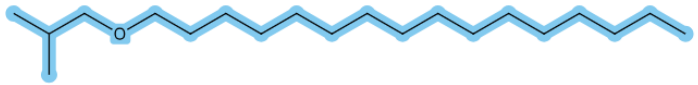
```
=> s c h o/elf (p) 1/o (p) c>16

L1      262175 C H O/ELF (P) 1/O (P) C>16
```

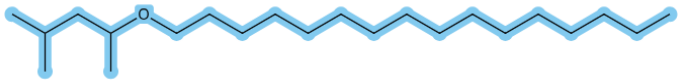



```
Saturated : 1  3
Element Counts
C: Exact(16) : 1
```

```
=> s l2 sub=l1 sss ful

L3          67 SEA SUB=L1 SSS FUL L2
```

Subset search of the generic ether query structure





FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Use Case 4: Alkylbenzene sulfonates

Task: Retrieve alkylbenzene sulfonates.

More than 12 carbon atoms

No further elements in the component formula

3 or 4 oxygen and 1 sulfur

(P) restricts retrieval to the same component formula

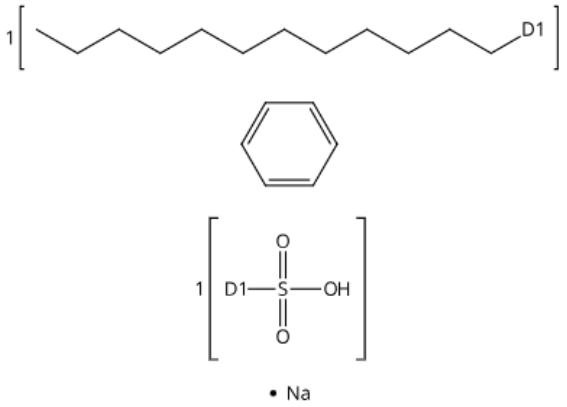
```
=> s c>12 (p) c h o s/elf (p) 3-4/o (p) 1/s

L1      255248 C>12 (P) C H O S/ELF (P) 3-4/O (P) 1/S
```

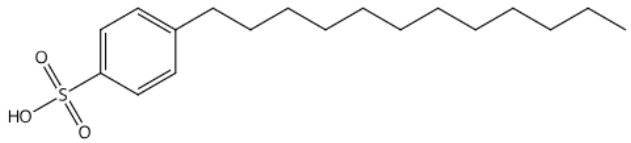
Ring identifier for benzene ring

Number of rings is one

```
=> s L1 and 46.150.18/rid and 1/nr

L2       38562 L1 AND 46.150.18/RID AND 1/NR
```

Incompletely defined substances are easily retrieved as well





FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society

Need support?

CAS Customer Center
1-800-753-4227
help@CAS.ORG

Monday 8:00 AM to Friday 12:00 midnight
U.S. Eastern Time

EMEA Customer Center
(+49) 721 95 883 155
EMEAhelp@CAS.org

Monday-Friday
8:30 AM to 5:00 PM CET

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

CAS
A division of the American Chemical Society