Package 'isopat'

February 20, 2015

Type Package

Title Calculation of isotopic pattern for a given molecular formula

Version 1.0

Date 2011-04-05

Author Martin Loos

Maintainer Martin Loos <loosmart@eawag.ch>

Description The function calculates the isotopic pattern (fine
structures) for a given chemical formula.

License GPL-2

LazyLoad yes

Repository CRAN

Date/Publication 2012-10-29 08:59:03

NeedsCompilation no

R topics documented:

---

isopat-package   *Calculation of isotopic pattern of a chemical formula.*

Description

The function calculates the isotopic pattern (fine structures) for a given chemical formula. It returns
a list of accurate masses, abundances and isotopic compositions. The elements used and their
isotopic characteristics can be defined by the user.

Details

Package:	isopat
Type:	Package
Version:	1.0
Date:	2011-04-05
License:	GPL-2
LazyLoad:	yes

Starting from a monoisotopic peak of a given molecule, the isotopic compositions and the concomitant abundances of isotopologues are changed iteratively towards less abundant isotopes over a number of generations 1,...,i,...,n. At each generation level i, a number i of isotopes contained in the monoisotopic peak have been exchanged for less abundant isotopes over all possible combinatorial changes of size i. Thus, generations of isotopic compositions are derived from an initial monoisotopic peak (i=0) while progressing to less abundant isotope compositions of a molecule.

Author(s)

Martin Loos

Maintainer: Martin Loos <loosmart@eawag.ch>

See Also

isopattern iso_list

Examples

```
## calculate isotopic peaks for coffein-D9 #####################################
data(iso_list);
isopattern(iso_list,"C8H1N4O2D9",1e-10);
```

---

isopattern	*Isotopic pattern calculation*

Description

The function calculates the isotopic pattern (fine structures) for a given chemical formula. It returns a list of accurate masses, abundances and isotopic compositions. The isotopic characterisitics of all elements contained in a molecule can be defined by the user as argument to the function.

Usage

```
isopattern(iso_list, compound, limit)
```

Arguments

iso_list
: a list or dataframe with four columns, specifying the isotopic composition of the elements found in the chemical formula of the compound (see dataset iso_list contained in this package): Column one contains the element name, e.g. C. Column two contains the isotope names of the element, e.g. 12C and 13C. Columns three and four contain the exact masses and abundances of the isotopes, respectively. Individual isotopes (e.g. for labelled compounds) must be defined in the element name column (cp. iso_list data set).

compound
: character string of a molecular formula for the compound for which an isotopic pattern needs to be calculated. All counts of all elements must be provided, i.e. C1H4 is a valid argument whereas CH4 is not. Element names must be followed by numbers (atom counts of that element), except for numbers in square brackets indicating individual isotopes defined in the element name column of iso_list, e.g. [14]C or [18]O. For example, [13]C2C35H67N1O13 is the molecular formula of erythromycin labeled at two C-positions with ^{13}C; C37H67N1O13 is the molecular formula of the unlabeled compound.

limit
: abundance limit; isotopic compositions equal to or below that abundance are dropped from further calculations.

Details

Starting from a monoisotopic peak of a given molecule (e.g. C2H6, with atoms of all elements set to those isotopes with highest abundance), the isotopic compositions and the concomitant abundances of isotopologues are iteratively altered towards less abundant isotopes. For a first iteration (generation i=1), [12]C2[1]H6 would hence be changed to both [12]C1[13]C1[1]H6 and [12]C2[1]H5[2]H1. Similarly, the latter two isotopologues then again lead to two combinations each for C and H at the second generation i=2. At each generation level i, a number i of isotopes contained in the monoisotopic peak have been exchanged for less abundant isotopes over all possible combinatorial changes of size i. Abundances are checked for <= the limit argument. If below limit, the concerned isotopologues are not changed forward to the next generation i+1. Different combinatorial changes carried from one generation to the next can eventually lead to the same isotopic composition at a given generation, causing double occurrences of some peaks in a generation. Therefore, peaks are checked against double isotopologues at each generation level.

The methodology used for updating abundances and masses when progressing from generation i to i+1 resembles that of Li et al. (2008). However, the updating does not strictly follow increasing nucleon numbers. Instead, generations of isotopic compositions are derived from an initial monoisotopic peak (i=0) via exchanging more abundant for less abundant isotopes.

Value

peaks
: A peak matrix with rows corresponding to single peaks; sorted by increasing masses. First column = mass. Second column = abundance. Third column = generation number. Fourth column = limit reached? Rest columns = isotopic composition of that isotopologue peak.

Note

Too high values (e.g.1E-5) for the limit may prevent the calculation of isotopologues for molecules containing both (1) several abundant isotopes of one element (e.g.[35]Cl and [37]Cl) and (2) many atom counts for the latter (e.g. hypothetical Cl500 in examples). On the other hand, too low values for the limit may lead to the unnecessary calculation of peaks with very little abundance. The user is requested to find a trade-off, possibly by comparing peak lists derived from different limit settings.

The maximum number of peaks that can be calculated is currently restricted to 5e+05.

Author(s)

Martin Loos, loosmart@eawag.ch

References

Li, L., Kresh, J., Karabacak, N., Cobb, J., Agar, J. and Hong, P. (2008). A Hierarchical Algorithm for Calculating the Isotopic Fine Structures of Molecules. Journal of the American Society for Mass Spectrometry, 19, 1867–1874.

See Also

iso_list

Examples

```
## firstly, load list of isotope data #########################################
data(iso_list);
iso_list;

## calculate isotopic peaks for coffein-D9 ####################################
peaks<-isopattern(iso_list,"C8H1N4O2D9",1e-10);
peaks;
# plot with y = absolute abundance
plot(peaks[,1],peaks[,2],type="h",xlab="mass",ylab="absolute abundance");
# plot with y = relative abundance scaled to the most intensive peak
peaks2<-peaks;
peaks2[,2]<-peaks2[,2]/max(peaks2[,2]);
plot(peaks2[,1],peaks2[,2],type="h",xlab="mass",ylab="relative abundance");
# plot with y = relative abundance scaled to the monoisotopic peak
# which, in this case, makes no difference
peaks2<-peaks;
peaks2[,2]<-peaks2[,2]/peaks2[1,2];
plot(peaks2[,1],peaks2[,2],type="h",xlab="mass",ylab="relative abundance");

## calculate isotopic peaks for the hypothetical compound Cl500 ###############
# use a too low value for limit -> plot shows only two peaks:
system.time(peaks<-isopattern(iso_list,"Cl500",1e-10));
plot(peaks[,1],peaks[,2],type="h",xlab="mass",ylab="absolute abundance");
# reset limit to a very low value for this "difficult" compound, then plot again:
system.time(peaks<-isopattern(iso_list,"Cl500",1e-100));
plot(peaks[,1],peaks[,2],type="h",xlab="mass",ylab="absolute abundance");
```

iso_list   *Isotope masses and abundances*

Description

Data frame of isotope masses and abundances for a number of elements.

Usage

```
data(iso_list)
```

Format

A data frame with four columns:

element name of element

isotope isotopes of an element

mass mass of an isotope of an element

abundance abundance of an isotope of an element

Details

The data frame (or any other data frame with that format) serves as argument for the isopattern function. The first column of the data frame also contains names of specific [isotopes] used for labeled compounds.

Source

www.iupac.org/publications/pac/2003/pdf/7506x0683.pdf

References

isopattern

Examples

```
data(iso_list)
```

Index