Package 'IsoSpecR'

October 12, 2022

Encoding UTF-8

Type Package

Title The IsoSpec Algorithm

Version 2.1.3

Date 2020-10-29

Author Mateusz Krzysztof Lacki and Michal Startek

Maintainer Matteo Lacki <matteo.lacki@gmail.com>

Description IsoSpec is a fine structure calculator used for obtaining the most probable masses of a chemical compound given the frequencies of the composing isotopes and their masses. It finds the smallest set of isotopologues with a given probability. The probability is assumed to be that of the product of multinomial distributions, each corresponding to one particular element and parametrized by the frequencies of finding these elements in nature. These numbers are supplied by IUPAC - the International Union of Pure and Applied Chemistry. See: Lacki, Valkenborg, Startek (2020) <DOI:10.1021/acs.analchem.0c00959> and Lacki, Startek, Valkenborg, Gambin (2017) <DOI:10.1021/acs.analchem.6b01459> for the description of the algorithms used.

License BSD_2_clause + file LICENCE

URL http://matteolacki.github.io/IsoSpec/

Depends R (>= 3.0.0)

Imports Rcpp (>= 0.12.0)

Suggests testthat

LazyData no

LinkingTo Rcpp

NeedsCompilation yes

SystemRequirements C++14

RoxygenNote 6.1.1

Repository CRAN

Date/Publication 2020-10-31 00:40:06 UTC

R topics documented:

`custom_isotopes_example` *An example of how to add your own elements.*

Description

This can be used, for instance, with isotopically labelled molecules.

Usage

```
custom_isotopes_example()
```

`IsoSpecify` *Calculate the isotopic fine structure peaks.*

Description

`IsoSpecify` is a wrapper around `Rinterface` that calls the C++ implementation of the IsoSpec algorithm. Given a molecular formula, it will calculate the smallest set of infinitely resolved peaks (isotopologues) that jointly is `p` probable, where `p` is provided by the user.

Usage

```
IsoSpecify(molecule, stopCondition, isotopes = NULL,
  showCounts = FALSE, trim = TRUE, algo = 0, step = 0.25)
```

Arguments

`molecule`	A named integer vector, e.g. `c(C=2,H=6,O=1)`, containing the chemical formula of the substance of interest.
`stopCondition`	A numeric value between 0 and 1.
`isotopes`	A named list of isotopic information required for IsoSpec. The names must be valid element symbols, see `isotopicData` for examples. Each enlisted object should be a `data.frame` containing columns `element` (specifying the symbol of the element), `mass` (specifying the mass of the isotope), `abundance` (specyfying the assumed frequency of finding that isotope).
`showCounts`	Logical. If `TRUE`, then we output matrix contains additionally counts of isotopes for each isotopologue.

`trim` Logical. If `FALSE`, then we output matrix contains additionally isotopologues that otherwise would get trimmed in order to find the smalles possible p-set. Therefore, switching to `FALSE` results in a slightly larger set then the optimal p-set.

`algo` An integer: 0 - use standard IsoSpec algoritm, where `stopCondition` specifies the probability of the optimal p-set, 1 - use a version of algorithm that uses priority queue. Slower than 0, but does not require sorting. 2 - use a threshold version of the algorithm, where `stopCondition` specifies the height of the pruned peaks. 3 - for the threshold version of IsoSpec with `stopCondition` being the percentage of the highest peak below which isotopologues get pruned.

`step` The percent of the the percentile of isotopologues in the current isolayer, specyfying the cutoff for the next isolayer. It has been optimised and better not change the default value.

Value

A numeric matrix containing the masses, the logarithms of probability, and, optionally, counts of isotopologues. Attention: this matrix does not have to be sorted. Sorting it would also compromise the linear complexity of our algorithm.

Examples

```
library(IsoSpecR)
res <- IsoSpecify( molecule = c(C=10,H=22,O=1), stopCondition = .9999 )
print(res)
```

`isotopicData` *Data on isotope masses, abundances and other.*

Description

A list of data frames or table data frames (dplyr like), containing different information on isotopes.

Usage

```
isotopicData
```

Format

A list of 6 tbl_df's or data frames, each constaining:

element The symbol of an element from Mendeleev's periodic table.

isotope String composed of the nucleon number and the symbol of element.

mass Isotope's Mass in Daltons.

abundance The abundance of the isotopes. In case of enviPat data abundances do not sum to one. In case of all other, they do.

ratioC As in enviPat reference manual: "Maximum number of atoms of an element for one C-atom in a molecule, based on 99.99 % of case molecules".

Source

R Package enviPat and Commission on Isotopic Abundances and Atomic Weights, CIAAW, https://www.ciaaw.org/index.htm.

Index