Package 'nmrrr'

May 10, 2023

Type Package

Title Binning and Visualizing NMR Spectra in Environmental Samples

Version 1.0.0

Description A reproducible workflow for binning and visualizing NMR (nuclear magnetic resonance) spectra from environmental samples. The 'nmrrr' package is intended for post-processing of NMR data, including importing, merging and, cleaning data from multiple files, visualizing NMR spectra, performing binning/integrations for compound classes, and relative abundance calculations.
This package can be easily inserted into existing analysis workflows by users to help with analyzing and interpreting NMR data.

URL https://github.com/kaizadp/nmrrr

BugReports https://github.com/kaizadp/nmrrr/issues

License MIT + file LICENSE

Copyright Battelle Memorial Institute

Encoding UTF-8

RoxygenNote 7.2.3

Imports DescTools, ggplot2, tidyr, utils

Suggests covr, knitr, rmarkdown, testthat (>= 3.0.0), tibble

Config/testthat/edition 3

Depends R (>= 3.50)

LazyData true

VignetteBuilder knitr

NeedsCompilation no

Author Kaizad Patel [aut, cre] (<https://orcid.org/0000-0001-6437-7809>),
Allison Myers-Pigg [aut] (<https://orcid.org/0000-0002-6905-6841>),
Ben Bond-Lamberty [ctb] (<https://orcid.org/0000-0001-9525-4633>),
Morgan Barnes [ctb] (<https://orcid.org/0000-0001-8722-6552>)

Maintainer Kaizad Patel <kaizad.patel@pnnl.gov>

Repository CRAN

Date/Publication 2023-05-10 16:10:02 UTC

R topics documented:

`bins_CadeMenun2015` *NMR grouping bins from Cade-Menun (2015).*

Description

NMR grouping bins from Cade-Menun (2015), for 31P, using D2O as a solvent. (1) polyphosphate (-20 to -4); (2) diester (-1.5 to 2.0); (3) monoester (3.0 to 5.5); (4) orthophosphate (5.5 to 9.0); (5) phosphate (9.0 to 40.0)

Usage

```
bins_CadeMenun2015
```

Format

A data frame with 5 rows and 5 variables:

number Bin number

group Name of bin group

start ppm shift range, lower limit

stop ppm shift range, upper limit

description Description of the bin group

Note

The NMR spectrum can be split into several bins, based on chemical shift (ppm). Binsets are specific to nuclei and solvents and by definition are open on the left and closed on the right; for example, a bin of (0,1) includes 1 but *not* 0.

Source

B. Cade-Menun. "Improved peak identification in 31P-NMR spectra of environmental samples with a standardized method and peak library". Geoderma. doi:10.1016/j.geoderma.2014.12.016

See Also

bins_Clemente2012 bins_Lynch2019 bins_Mitchell2018

`bins_Clemente2012` *NMR grouping bins from Clemente et al. (2012).*

Description

NMR grouping bins from Clemente et al. (2012), using DMSO-D6 as solvent. (1) aliphatic polymethylene and methyl groups (0.6–1.3 ppm, "aliphatic1"); (2) aliphatic methyl and methylene near O and N (1.3–2.9 ppm, "aliphatic2"); (3) O-alkyl, mainly from carbohydrates and lignin (2.9–4.1 ppm); (4) alpha-proton of peptides (4.1–4.8 ppm); (5) aromatic and phenolic (6.2–7.8 ppm); and (6) amide, from proteins (7.8–8.4 ppm).

Usage

```
bins_Clemente2012
```

Format

A data frame with 6 rows and 5 variables:

number Bin number

group Name of bin group

start ppm shift range, lower limit

stop ppm shift range, upper limit

description Description of the bin group

Note

The NMR spectrum can be split into several bins, based on chemical shift (ppm). Binsets are specific to nuclei and solvents and by definition are open on the left and closed on the right; for example, a bin of (0,1) includes 1 but *not* 0.

Source

JS Clemente et al. 2012. "Comparison of Nuclear Magnetic Resonance Methods for the Analysis of Organic Matter Composition from Soil Density and Particle Fractions." Environmental Chemistry doi:10.1071/EN11096

See Also

bins_Lynch2019 bins_Mitchell2018 bins_Hertkorn2013

`bins_Hertkorn2013` *NMR grouping bins from Hertkorn et al. (2013).*

Description

NMR grouping bins from Hertkorn et al. (2013), using MeOD as solvent. (1) aliphatics, HCCC (0.0-1.9); (2) acetate analogs and CRAM (carboxyl-rich alicyclic materials), HCX (1.9-3.1); (3) carbohydrate-like and methoxy, HCO (3.1-4.9); (4) olefinic HC=C (5.3-7.0); (5) aromatic (7.0-10.0).

Usage

```
bins_Hertkorn2013
```

Format

A data frame with 5 rows and 5 variables:

number Bin number

group Name of bin group

start ppm shift range, lower limit

stop ppm shift range, upper limit

description Description of the bin group

Note

The NMR spectrum can be split into several bins, based on chemical shift (ppm). Binsets are specific to nuclei and solvents and by definition are open on the left and closed on the right; for example, a bin of (0,1) includes 1 but *not* 0.

Source

N. Hertkorn et al. 2013. "High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter" Biogeosciences doi:10.5194/bg1015832013

See Also

`bins_Clemente2012` `bins_Lynch2019` `bins_Mitchell2018`

`bins_Lynch2019` *NMR grouping bins from Lynch et al. (2019).*

Description

NMR grouping bins from Lynch et al. (2019), using D2O as solvent. (1) methyl, methylene, and methane bearing protons (0.6–1.6 ppm) ; (2) unsaturated functional groups (1.6–3.2 ppm), including ketone, benzylic, and alicyclic-bearing protons; (3) unsaturated, heteroatomic compounds, including O-bearing carbohydrates, ethers, and alcohols (3.2–4.5 ppm); (4) conjugated, double bond functionalities, including aromatic, amide, and phenolic structures (6.5–8.5 ppm).

Usage

```
bins_Lynch2019
```

Format

A data frame with 4 rows and 5 variables:

number Bin number

group Name of bin group

start ppm shift range, lower limit

stop ppm shift range, upper limit

description Description of the bin group

Note

The NMR spectrum can be split into several bins, based on chemical shift (ppm). Binsets are specific to nuclei and solvents and by definition are open on the left and closed on the right; for example, a bin of (0,1) includes 1 but *not* 0.

Source

LM Lynch et al. 2019. "Dissolved Organic Matter Chemistry and Transport along an Arctic Tundra Hillslope." Global Biogeochemical Cycles doi:10.1029/2018GB006030

See Also

`bins_Clemente2012` `bins_Mitchell2018` `bins_Hertkorn2013`

`bins_Mitchell2018` *NMR grouping bins from Mitchell et al. (2018).*

Description

NMR grouping bins from Mitchell et al. (2018), using DMSO-D6 as solvent. (1) aliphatic polymethylene and methyl groups (0.6–1.3 ppm); (2) N- and O-substituted aliphatic (1.3–2.9 ppm); (3) O-alkyl (2.9–4.1 ppm); (4) alpha-proton of peptides (4.1–4.8 ppm); (5) anomeric proton of carbohydrates (4.8–5.2 ppm); (6) aromatic and phenolic (6.2–7.8 ppm); (7) amide (7.8–8.4 ppm).

Usage

```
bins_Mitchell2018
```

Format

A data frame with 7 rows and 5 variables:

number Bin number

group Name of bin group

start ppm shift range, lower limit

stop ppm shift range, upper limit

description Description of the bin group

Note

The NMR spectrum can be split into several bins, based on chemical shift (ppm). Binsets are specific to nuclei and solvents and by definition are open on the left and closed on the right; for example, a bin of (0,1) includes 1 but *not* 0.

Source

P Mitchell et al. 2018. "Nuclear Magnetic Resonance Analysis of Changes in Dissolved Organic Matter Composition with Successive Layering on Clay Mineral Surfaces." Soil Systems doi:10.3390/soils2010008

See Also

`bins_Clemente2012` `bins_Lynch2019` `bins_Hertkorn2013`

`bins_ss_Baldock2004` *NMR grouping bins from Baldock et al. (2004).*

Description

NMR grouping bins from Baldock et al. (2004), for solid-state NMR. (1) alkyl C (0-45); (2) methoxyl C and N-alkyl C (45-60); (3) O-alkyl C (60-95); (4) di-O-alkyl C (95-110); (5) aromatic C (110-145); (6) phenolic C (145-165); (7) amide and carboxyl C (165-215)

Usage

```
bins_ss_Baldock2004
```

Format

A data frame with 5 rows and 5 variables:

number Bin number

group Name of bin group

start ppm shift range, lower limit

stop ppm shift range, upper limit

description Description of the bin group

Note

The NMR spectrum can be split into several bins, based on chemical shift (ppm). Binsets are specific to nuclei and solvents and by definition are open on the left and closed on the right; for example, a bin of (0,1) includes 1 but *not* 0.

Source

J. Baldock et al. "Cycling and composition of organic matter in terrestrial and marine ecosystems". Marine Chemistry. doi:10.1016/j.marchem.2004.06.016

See Also

`bins_Clemente2012` `bins_Lynch2019` `bins_Mitchell2018`

`bins_ss_Clemente2012` *NMR grouping bins from Clemente et al. (2012) - ss.*

Description

NMR grouping bins from Clemente et al. (2012), for solid-state NMR. (1) alkyl C (0-50); (2) O-alkyl C (60-93); (3) anomeric C (95-110); (4) aromatic C (110-160); (5) carboxyl-carbonyl C (160-200)

Usage

```
bins_ss_Clemente2012
```

Format

A data frame with 5 rows and 5 variables:

number Bin number

group Name of bin group

start ppm shift range, lower limit

stop ppm shift range, upper limit

description Description of the bin group

Note

The NMR spectrum can be split into several bins, based on chemical shift (ppm). Binsets are specific to nuclei and solvents and by definition are open on the left and closed on the right; for example, a bin of (0,1) includes 1 but *not* 0.

Source

JS Clemente et al. 2012. "Comparison of Nuclear Magnetic Resonance Methods for the Analysis of Organic Matter Composition from Soil Density and Particle Fractions." Environmental Chemistry doi:10.1071/EN11096

See Also

`bins_Clemente2012` `bins_ss_Preston2009` `bins_Mitchell2018`

`bins_ss_Preston2009` *NMR grouping bins from Preston et al. (2009).*

Description

NMR grouping bins from Preston et al. (2009), for solid-state NMR. (1) alkyl C (0-50); (2) methoxyl C (50-60); (3) O-alkyl C (60-93); (4) di-O-alkyl C (93-112); (5) aromatic C (112-140); (6) phenolic C (140-165); (7) carboxyl C (165-190)

Usage

```
bins_ss_Preston2009
```

Format

A data frame with 5 rows and 5 variables:

number Bin number

group Name of bin group

start ppm shift range, lower limit

stop ppm shift range, upper limit

description Description of the bin group

Note

The NMR spectrum can be split into several bins, based on chemical shift (ppm). Binsets are specific to nuclei and solvents and by definition are open on the left and closed on the right; for example, a bin of (0,1) includes 1 but *not* 0.

Source

C. Preston et al. 2009. "Chemical Changes During 6 Years of Decomposition of 11 Litters in Some Canadian Forest Sites. Part 1. Elemental Composition, Tannins, Phenolics, and Proximate Fractions". Ecosystems. doi:10.1007/s1002100992660

See Also

`bins_Clemente2012` `bins_Lynch2019` `bins_Mitchell2018`

nmr_assign_bins *Assign compound classes using the chosen binset*

Description

Assign group (bin name) to each row of the data based on the ppm column.

Usage

```
nmr_assign_bins(dat, binset)
```

Arguments

dat	Input dataframe. This could be spectral data, or peak picked data. Must include a 'ppm' column for compound class assignment
binset	A binset; e.g. bins_Clemente2012, bins_Hertkorn2013, etc., or a similarly-structured data frame

Value

The input data with a new group column whose entries are drawn from the binset. Entries will be NA if a ppm value does not fall into any group.

Author(s)

Kaizad Patel

Examples

```
sdir <- system.file("extdata", "kfp_hysteresis", "spectra_mnova", package = "nmrrr")
spec <- nmr_import_spectra(path = sdir, method = "mnova")
nmr_assign_bins(spec, bins_Clemente2012)
```

nmr_import_peaks *Import and process picked peaks data*

Description

Process data of peaks picked with NMR software.

Usage

```
nmr_import_peaks(path, method, pattern = "*.csv$", quiet = FALSE)
```

Arguments

path	Directory where the peaks data are saved
method	Format of input data, depending on how the data were exported. "multiple columns": data are in split-column format, obtained by pasting "peaks table" in MNova. "single column": data are in single-column format, exported from MNova as "peaks script".
pattern	Filename pattern to search for (by default "*.csv$")
quiet	Print diagnostic messages? Logical

Value

A dataframe with columns describing sample ID, ppm, intensity, area, group name.

Author(s)

Kaizad Patel

Examples

```
sdir <- system.file("extdata", "kfp_hysteresis", "peaks_mnova_multiple", package = "nmrrr")
nmr_import_peaks(path = sdir, method = "multiple columns")
```

---

nmr_import_spectra *Import and process NMR spectral data*

Description

Imports multiple spectra files and then combines and cleans the data.

Usage

```
nmr_import_spectra(path, method, pattern = "*.csv$", quiet = FALSE)
```

Arguments

path	Directory where the spectra files are saved
method	Software used for initial processing of NMR spectra (before using this package). Available options include "mnova" and "topspin".
pattern	Filename pattern to search for (by default "*.csv$")
quiet	Print diagnostic messages? Logical

Value

A link{data.frame} with data from all files found, concatenated and sorted.

Author(s)

Kaizad Patel

Examples

```
sdir <- system.file("extdata", "kfp_hysteresis", "spectra_mnova", package = "nmrrr")
nmr_import_spectra(path = sdir, method = "mnova")
```

---

nmr_plot_spectra *Plot NMR spectra*

Description

Plot NMR spectra, with line-brackets denoting binned regions. Uses spectra data processed in MestreNova or TopSpin.

Usage

```
nmr_plot_spectra(
  dat,
  binset,
  label_position = 100,
  mapping = aes(x = ppm, y = intensity),
  stagger = 10
)
```

Arguments

dat	Processed spectral data, output from (a) nmr_import_spectra and nmr_assign_bins; or (b) nmr_import_peaks
binset	A binset; e.g. bins_Clemente2012, bins_Hertkorn2013, etc., or a similarly-structured data frame
label_position	y-axis position for bin labels
mapping	An aesthetic mapping generated by aes: aes(x = ..., y = ...)
stagger	How much to stagger the labels, numeric; same units as label_position

Value

A ggplot object.

Author(s)

Kaizad Patel

Examples

```
sdir <- system.file("extdata", "kfp_hysteresis", "spectra_mnova", package = "nmrrr")
spec <- nmr_import_spectra(path = sdir, method = "mnova")
library(ggplot2)
p_aes <- aes(x = ppm, y = intensity)
p <- nmr_plot_spectra(spec, bins_Clemente2012, 5, p_aes, stagger = 0.5)
p + ylim(0, 6)
```

---

nmr_relabund *Compute relative abundance for each sample*

Description

Compute relative abundance of compound classes for each sample.

Usage

```
nmr_relabund(dat, method)
```

Arguments

dat	Processed spectral data, output from (a) nmr_import_spectra and nmr_assign_bins; or (b) nmr_import_peaks
method	The method for calculating relative abundance. Options include (a) "AUC", integrating the spectral region within each bin; (b) "peaks", adding areas of peaks if a peak-picked file is provided.

Value

A data.frame with columns describing relative contributions of compound classes. Compound classes are determined by selecting the desired binset.

Author(s)

Kaizad Patel

Examples

```
sdir <- system.file("extdata", "kfp_hysteresis", "peaks_mnova_multiple", package = "nmrrr")
peaks <- nmr_import_peaks(path = sdir, method = "multiple columns")
peaks <- nmr_assign_bins(peaks, bins_Clemente2012)
nmr_relabund(peaks, "peaks")
```

Index