Package 'Fiscore'

September 3, 2021

Type Package

Title Effective Protein Structural Data Visualisation and Exploration

Version 0.1.3

Author Auste Kanapeckaite

Maintainer Auste Kanapeckaite <auste.kan@algorithm379.com>

Description Collection of structural analysis tolls to compliment the research paper ``Fiscore: a novel approach to characterise protein topology and aid in drug discovery studies" ;<doi:10.1080/07391102.2020.1854859>. 'Fiscore' package allows to explore and identify new topologically and functionally relevant structural features by applying integrated Gaussian Mixture Models.

License GPL (>= 3)

Encoding UTF-8

VignetteBuilder knitr

RoxygenNote 7.1.1

Imports ggplot2, mclust, bio3d, stringr, plotly, methods,lattice, stats, dplyr, knitr, rmarkdown

NeedsCompilation no

Repository CRAN

Date/Publication 2021-09-02 22:50:11 UTC

R topics documented:

`B_plot_normalised` *B_plot_normalised*

Description

Function to plot B-factor normalised values per amino acid using a bar plot; NOTE: some PDB files have breakages in their amino acid sequence, that is some residues might be missing and the gaps will be reflected in the plot

Usage

```
B_plot_normalised(pdb_df)
```

Arguments

`pdb_df`	Requires a PDB data frame generated by PDB_prepare

Value

bar plot

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of B_plot_normalised
pdb_df<-read.table(path_to_processed_PDB)
B_plot_normalised(pdb_df)
```

`cluster_ID` *cluster_ID*

Description

Function to select an optimal number of clusters and a model to be fitted during the EM phase of clustering for Gaussian Mixture Models. The function provides summaries and helps to visualise clusters based on Fi-score using scatter plotting and dimension reduction plots.

Usage

```
cluster_ID(
  pdb_df,
  max_range = 20,
  secondary_structures = TRUE,
  clusters = NULL,
  modelNames = NULL
)
```

Arguments

pdb_df	data frame containing processed PDB file with Fi-score values
max_range	number of clusters to consider during model selection; default 20 clusters
secondary_structures	include information on secondary structure elements from PDB when plotting, default value is TRUE
clusters	number of clusters to test not based on the best BIC output, user also needs to supply modelNames
modelNames	can only be supplied when clusters are also specified, this option will model based on the user parameters

Value

A data frame object that contains a summary of clusters

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of cluster_ID
pdb_df<-read.table(path_to_processed_PDB)
head(cluster_ID(pdb_df))
```

---

density_plots *density_plots*

Description

Function plots a density plot set for phi/psi angle distributions, Fi-score and normalised B-factor. As well as 3D visualisation of angle distribution for every residue. The plots can be used for a quick assessment of the overall parameters.

Usage

```
density_plots(pdb_df, model_report)
```

Arguments

pdb_df	Requires a PDB data frame generated by PDB_prepare
model_report	Optional parameter to include data from cluster_ID

Value

multiple plots

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of density_plots
pdb_df<-read.table(path_to_processed_PDB)
density_plots(pdb_df)
```

Fiscore_secondary *Fiscore_secondary*

Description

Function plots a bar plot with a secondary structure element visualisation based on PDB file data; NOTE: NA refers to unidentified region, e.g., a likely disordered or unstructured region

Usage

```
Fiscore_secondary(pdb_df)
```

Arguments

pdb_df	Requires a PDB data frame generated by PDB_prepare

Value

bar plot

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of Fiscore_secondary
pdb_df<-read.table(path_to_processed_PDB)
Fiscore_secondary(pdb_df)
```

Fi_score_plot *Fi_score_plot*

Description

Function to plot Fi-score values per amino acid using a bar plot; NOTE: some PDB files have breakages in their amino acid sequence, that is some residues might be missing and the gaps will be reflected in the plot

Usage

```
Fi_score_plot(pdb_df)
```

Arguments

pdb_df	Requires a PDB data frame generated by PDB_prepare

Value

Bar plot

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of Fi_score_plot
pdb_df<-read.table(path_to_processed_PDB)
Fi_score_plot(pdb_df)
```

Fi_score_region *Fi_score_region*

Description

Function calculates combined Fi-score for a selected region; NOTE: some PDB files have breakages in their amino acid sequences and those values cannot be assessed. Moreover, values can be calculated either inclusively or not; include is set to FALSE by default

Usage

```
Fi_score_region(pdb_df, i, j, include = FALSE)
```

Arguments

pdb_df	Requires a PDB data frame generated by PDB_prepare
i	start residue for a region
j	end residue for a region
include	inclusive or not calculation, default FALSE

Value

Fi-score value

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of Fiscore_secondary
pdb_df<-read.table(path_to_processed_PDB)
Fi_score_region(pdb_df,900,925)
```

---

hydrophobicity_plot *hydrophobicity_plot*

Description

Function to plot amino acid sequence hydrophobicity profile using Kyte-Doolittle hydrophobicity scale; Reference= J. Mol. Biol. 157=105-132;1982. The Kyte-Doolittle scale is used for detecting hydrophobic regions in proteins. Regions with a positive value are hydrophobic ans those with negative values are hydrophylic. This scale can be used to identify both surface-exposed regions as well as transmembrane regions, depending on the window size used.

Usage

```
hydrophobicity_plot(pdb_df, window = 3, weight = 100, model = "exponential")
```

Arguments

pdb_df	Requires a pdb data frame generated by PDB_prepare
window	Size of a window between 3 and 21, default is 21
weight	Relative weight of the window edges compared to the window center in percent; default=100
model	Weight variation model either "linear" or "exponential", if the relative weight at the edges is selected to be < 100 percent; default="linear"

Value

Scaled line graph

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of hydrophobicity_plot
pdb_df<-read.table(path_to_processed_PDB)
hydrophobicity_plot(pdb_df)
```

PDB_prepare *PDB_prepare*

Description

Function to prepare a PDB file after it was pre-processed to generate Fi-score and normalised B factor values as well as secondary structure designations

Usage

```
PDB_prepare(file_name)
```

Arguments

file_name	PDB file name to load that was split into chains, e.g. '6KZ5_A.pdb'

Value

returns a processed data frame with Fi-score 'Fi_score', normalised B factor values 'B_normalised' and secondary structure designations

Examples

```
path_to_processed_PDB<- system.file("extdata", "3nf5_A.pdb", package="Fiscore")
# you can call PDB_prepare with the set path
head(PDB_prepare(path_to_processed_PDB))
```

PDB_process *PDB_process*

Description

Function to preprocess and inspect a PDB file

Usage

```
PDB_process(file_name, path = "split_PDB")
```

Arguments

file_name	PDB file name to load, e.g. '6KZ5.pdb'
path	location where to transfer split PDB files, default will create a new directory in your working directory

Value

generates split chain PDB files in the default or selected directory and then returns the names of the files

Examples

```
path_to_PDB_file<- system.file("extdata", "3nf5.pdb", package="Fiscore")
# basic usage of PDB_process calls the selected path to load a large file
```

`phi_psi_3D` *phi_psi_3D*

Description

Function plots a 3D scatter plot with a secondary structure element visualisation based on PDB file data; NOTE: NA refers to unidentified region, e.g., a disordered region. The plot includes information, such as phi and psi dihedral angles as well as normalised B-factor values.

Usage

```
phi_psi_3D(pdb_df)
```

Arguments

`pdb_df`	Requires a PDB data frame generated by PDB_prepare

Value

Interactive plot

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of phi_psi_3D
pdb_df<-read.table(path_to_processed_PDB)
phi_psi_3D(pdb_df)
```

phi_psi_bar_plot *phi_psi_bar_plot*

Description

Function to plot PDB file dihedral angle distribution per amino acid using a bar plot. NOTE: some PDB files have breakages in their amino acid sequences; that is, some residues might be missing and it will be reflected in the plot via empty spaces

Usage

```
phi_psi_bar_plot(pdb_df)
```

Arguments

pdb_df	Requires a PDB data frame generated by PDB_prepare

Value

Bar plot

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of phi_psi_bar_plot
pdb_df<-read.table(path_to_processed_PDB)
phi_psi_bar_plot(pdb_df)
```

phi_psi_interactive *phi_psi_interactive*

Description

Function plots a scatter plot with a secondary structure element visualisation based on the PDB file data; NOTE: NA refers to unidentified region, e.g., a disordered region

Usage

```
phi_psi_interactive(pdb_df)
```

Arguments

pdb_df	Requires a PDB data frame generated by PDB_prepare

Value

Interactive plot

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of phi_psi_interactive
pdb_df<-read.table(path_to_processed_PDB)
phi_psi_interactive(pdb_df)
```

---

phi_psi_plot *phi_psi_plot*

Description

Function to plot a PDB file phi/psi angle density plot

Usage

```
phi_psi_plot(pdb_df)
```

Arguments

pdb_df	Requires a PDB data frame generated by PDB_prepare

Value

2D phi/psi angle distribution plot with residues binned

Examples

```
path_to_processed_PDB<- system.file("extdata", "pdb_df.tabular", package="Fiscore")
# basic usage of phi_psi_plot
pdb_df<-read.table(path_to_processed_PDB)
phi_psi_plot(pdb_df)
```

Index