

Package ‘tmhmm’

November 18, 2020

Type Package

Title Interface to TMHMM

Version 1.3

Author Richel J.C. Bilderbeek

Maintainer Richel J.C. Bilderbeek <richel@richelbilderbeek.nl>

Description Predicting the topology of membrane proteins is part of the field of computational biology.
‘TMHMM’ is a program to predict transmembrane helices in proteins.
‘TMHMM’ can be called from a website and/or locally using Bash.
This package allows to install and call TMHMM from R and parse its results.

License GPL-3

Encoding UTF-8

LazyData true

Imports ggplot2, rappdirs, stringr, tibble

Suggests curl, testthat (>= 2.1.0), knitr, rmarkdown

URL <https://github.com/richelbilderbeek/tmhmm/>

BugReports <https://github.com/richelbilderbeek/tmhmm/issues>

RoxygenNote 7.1.1

VignetteBuilder knitr

SystemRequirements TMHMM
(<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>)

NeedsCompilation no

Repository CRAN

Date/Publication 2020-11-18 08:50:02 UTC

R topics documented:

are_tmhs	2
check_tmhmm_installation	3

check_topology	4
default_params_doc	4
get_default_tmhmm_bin_path	5
get_default_tmhmm_folder	6
get_tmhmm_url	6
install_tmhmm	7
install_tmhmm_bin	7
is_locatome_line	8
is_on_travis	8
is_protein_name_line	9
is_tmh	9
is_tmhmm_bin_installed	10
is_tmhmm_installed	10
is_tmhmm_result	11
is_tmhmm_set_up	12
is_url_valid	12
locatome_to_df	13
plot_locatome	13
run_tmhmm	13
run_tmhmm_on_sequence	14
run_tmhmm_to_file	15
set_up_tmhmm	16
tally_locatome	16
tally_tmhs	17
tmhmm	17
tmhmm_self_test	18
uninstall_tmhmm	18

Index **20**

are_tmhs *Are the sequences TMHs?*

Description

Are the sequences TMHs?

Usage

are_tmhs(protein_sequences)

Arguments

protein_sequences
 one or more protein sequences

Author(s)

Richèl J.C. Bilderbeek

Examples

```
if (is_tmhmm_installed()) {  
  
  sequences <- c(  
    "VVIILTIAGNILVIMAVSLE",  
    "VVIILTIRGNILVIMAVSLE"  
  )  
  
  are_tmhs(sequences)  
}
```

check_tmhmm_installation

Checks the installation of TMHMM. Throws a helpful error message if incomplete, else does nothing

Description

Checks the installation of TMHMM. Throws a helpful error message if incomplete, else does nothing

Usage

```
check_tmhmm_installation(folder_name = get_default_tmhmm_folder())
```

Arguments

folder_name superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share by default, as can be obtained by [get_default_tmhmm_folder](#)

Value

Nothing

Author(s)

Richèl J.C. Bilderbeek

Examples

```
if (is_tmhmm_installed()) {  
  check_tmhmm_installation()  
}
```

check_topology	<i>Check if the topology is valid.</i>
----------------	--

Description

Check if the argument is of the same type as a predicted topology, as can be created with [run_tmhmm](#). Will **stop** if not.

Usage

```
check_topology(topology)
```

Arguments

topology	a tibble with two columns, named name for a protein's name and topology, for its predicted topology (e.g. ooooMMMiiii).
----------	---

Value

Nothing. Will **stop** with a helpful error message if the topology is invalid.

Author(s)

Richèl J.C. Bilderbeek

default_params_doc	<i>This function does nothing. It is intended to inherit its parameters' documentation.</i>
--------------------	---

Description

This function does nothing. It is intended to inherit its parameters' documentation.

Usage

```
default_params_doc(  
  download_url,  
  fasta_filename,  
  folder_name,  
  locatome,  
  protein_sequence,  
  protein_sequences,  
  tmhmm_bin_filename,  
  tmhmm_filename,  
  tmhmm_result,  
  topology,  
  verbose  
)
```

Arguments

download_url	the URL to download TMHMM from
fasta_filename	path to a FASTA file
folder_name	superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share by default, as can be obtained by get_default_tmhmm_folder
locatome	text similar to a FASTA file, with protein names and the locations of its amino acids
protein_sequence	a protein sequence
protein_sequences	one or more protein sequences
tmhmm_bin_filename	path to the TMHMM binary file. Its default path can be obtained by using get_default_tmhmm_bin_path
tmhmm_filename	filename to write the TMHMM results to
tmhmm_result	the results of a call to TMHMM, for example, c(">protein name", "ooooMMMMiiii")
topology	a tibble with two columns, named name for a protein's name and topology, for its predicted topology (e.g. ooooMMMMiiii).
verbose	set to TRUE for more output

Note

This is an internal function, so it should be marked with @noRd. This is not done, as this will disallow all functions to find the documentation parameters

Author(s)

Richèl J.C. Bilderbeek

get_default_tmhmm_bin_path

Get the default path to the TMHMM binary

Description

Get the default path to the TMHMM binary

Usage

```
get_default_tmhmm_bin_path()
```

Value

the default path to the TMHMM binary

Author(s)

Richèl J.C. Bilderbeek

get_default_tmhmm_folder

Get the path to the folder where this package installs TMHMM by default

Description

Get the path to the folder where this package installs TMHMM by default

Usage

```
get_default_tmhmm_folder()
```

Value

the path to the folder where this package installs TMHMM by default

Author(s)

Richèl J.C. Bilderbeek

Examples

```
get_default_tmhmm_folder()
```

get_tmhmm_url

Get an URL to download TMHMM from. These URLs expire.

Description

Get an URL to download TMHMM from. These URLs expire.

Usage

```
get_tmhmm_url()
```

Author(s)

Richèl J.C. Bilderbeek

Examples

```
get_tmhmm_url()
```

install_tmhmm	<i>Install NetMHC2pan to a local folder</i>
---------------	---

Description

Install NetMHC2pan to a local folder

Usage

```
install_tmhmm(  
  download_url = get_tmhmm_url(),  
  folder_name = get_default_tmhmm_folder()  
)
```

Arguments

download_url	the URL to download TMHMM from
folder_name	superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share by default, as can be obtained by get_default_tmhmm_folder

Author(s)

Richèl J.C. Bilderbeek

Examples

```
install_tmhmm()
```

install_tmhmm_bin	<i>Install the NetMHC2pan binary to a local folder</i>
-------------------	--

Description

Install the NetMHC2pan binary to a local folder

Usage

```
install_tmhmm_bin(  
  download_url = get_tmhmm_url(),  
  folder_name = rappdirs::user_data_dir()  
)
```

Arguments

download_url the URL to download TMHMM from
 folder_name superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share
 by default, as can be obtained by [get_default_tmhmm_folder](#)

Author(s)

Richèl J.C. Bilderbeek

is_locatome_line *Is the line of text the locatome, as used within a FASTA filename?*

Description

Is the line of text the locatome, as used within a FASTA filename?

Usage

```
is_locatome_line(line)
```

Arguments

line line of text from a FASTA filename

Examples

```
# A valid locatome line
is_locatome_line("ooooMMMMiiii")

# An invalid locatome line
is_locatome_line("nonsense")
```

is_on_travis *Determines if the environment is Travis CI*

Description

Determines if the environment is Travis CI

Usage

```
is_on_travis()
```

Value

TRUE if run on Travis CI, FALSE otherwise

Author(s)

Richèl J.C. Bilderbeek

is_protein_name_line *Is the line of text the name of a protein, as used within a FASTA filename?*

Description

Is the line of text the name of a protein, as used within a FASTA filename?

Usage

```
is_protein_name_line(line)
```

Arguments

line line of text from a FASTA filename

Examples

```
# A valid protein name line
is_protein_name_line(">5H2A_CRIGR")

# An invalid protein name line:
# need the '>' before the protein name
is_protein_name_line("5H2A_CRIGR")

# An invalid protein name line
is_protein_name_line("nonsense")
```

is_tmh *Determine if the protein sequence contains at least one TMH.*

Description

Determine if the protein sequence contains at least one TMH.

Usage

```
is_tmh(protein_sequence)
```

Arguments

protein_sequence
 a protein sequence

Examples

```
if (is_tmhmm_installed()) {  
  # A TMH  
  is_tmh("VVIILTIAGNILVIMAVSLE")  
  
  # A protein without a TMH  
  is_tmh("VVIILTIRGNILVIMAVSLE")  
}
```

is_tmhmm_bin_installed

Measure if TMHMM binary is installed locally

Description

Measure if TMHMM binary is installed locally

Usage

```
is_tmhmm_bin_installed(folder_name = get_default_tmhmm_folder())
```

Arguments

folder_name superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share by default, as can be obtained by [get_default_tmhmm_folder](#)

Value

TRUE is TMHMM binary is installed locally, FALSE otherwise

Author(s)

Richèl J.C. Bilderbeek

is_tmhmm_installed

Measure if TMHMM is installed locally

Description

Measure if TMHMM is installed locally

Usage

```
is_tmhmm_installed(folder_name = get_default_tmhmm_folder())
```

Arguments

folder_name superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share by default, as can be obtained by [get_default_tmhmm_folder](#)

Value

TRUE is TMHMM is installed locally, FALSE otherwise

Author(s)

Richèl J.C. Bilderbeek

Examples

```
is_tmhmm_installed()
```

is_tmhmm_result	<i>Is the text a TMHMM result text?</i>
-----------------	---

Description

Is the text a TMHMM result text?

Usage

```
is_tmhmm_result(tmhmm_result)
```

Arguments

tmhmm_result the results of a call to TMHMM, for example, c(">protein name", "ooooMMMMiiii")

Details

A TMHMM result text, is similar to a regular FASTA file: it shows the protein names and the locations of the amino acids (instead of the amino acids).

Here is an example of a TMHMM result text:

```
>protein name
oMMiii
```

In this example, protein name is the name of the protein. The protein consists out of 6 amino acids, of which the first is located outside the cell, the next two are in the membrane, and the last three are inside the (cytosol of) the cell. Note that this is a biologically impossible protein: it takes around 21 amino acids to span the lipid membrane.

Value

TRUE if the text is a TMHMM result text

Examples

```
tmhmm_result <- c(
  ">protein name",
  "ooooMMMMiiiiii"
)
is_tmhmm_result(tmhmm_result)
```

is_tmhmm_set_up	<i>Determine if TMHMM is set up</i>
-----------------	-------------------------------------

Description

Determine if TMHMM is set up

Usage

```
is_tmhmm_set_up(folder_name = get_default_tmhmm_folder())
```

Arguments

folder_name superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share by default, as can be obtained by [get_default_tmhmm_folder](#)

Value

TRUE if TMHMM is set up, FALSE otherwise

Author(s)

Richèl J.C. Bilderbeek

is_url_valid	<i>Is the download URL valid?</i>
--------------	-----------------------------------

Description

The download link expires after 4 hours.

Usage

```
is_url_valid(download_url = get_tmhmm_url())
```

Arguments

download_url the URL to download TMHMM from

locatome_to_df	<i>Convert a multi-line locatome to a tidy data frame</i>
----------------	---

Description

Convert a multi-line locatome to a tidy data frame

Usage

```
locatome_to_df(locatome)
```

Arguments

locatome	text similar to a FASTA file, with protein names and the locations of its amino acids
----------	---

plot_locatome	<i>Plot the locatome</i>
---------------	--------------------------

Description

Plot the locatome

Usage

```
plot_locatome(locatome)
```

Arguments

locatome	text similar to a FASTA file, with protein names and the locations of its amino acids
----------	---

run_tmhmm	<i>Run TMHMM</i>
-----------	------------------

Description

Run TMHMM

Usage

```
run_tmhmm(fasta_filename, folder_name = get_default_tmhmm_folder())
```

Arguments

fasta_filename path to a FASTA file
 folder_name superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share by default, as can be obtained by [get_default_tmhmm_folder](#)

Value

a character vector with the locatome

Author(s)

Richèl J.C. Bilderbeek

Examples

```
if (is_tmhmm_installed()) {
  fasta_filename <- system.file("extdata", "tmhmm.fasta", package = "tmhmm")
  locatome <- run_tmhmm(fasta_filename)
  message(locatome, sep = "\n")
}
```

run_tmhmm_on_sequence *Run TMHMM directly on a protein sequence*

Description

Run TMHMM directly on one protein sequence

Usage

```
run_tmhmm_on_sequence(
  protein_sequence,
  folder_name = get_default_tmhmm_folder()
)
```

Arguments

protein_sequence
 a protein sequence, with the amino acids as capitals, for example 'MEILCED-NTSLSSIPNSL'
 folder_name superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share by default, as can be obtained by [get_default_tmhmm_folder](#)

Value

the topology. The topology is a character string with the same length as the protein sequence. The topology consists of the characters i ('inside'), I ('inside'), m ('membrane'), M ('membrane'), o ('outside') and O ('outside')

Examples

```
if (is_tmhmm_installed()) {  
  protein_sequence <- paste0(  
    "QEKNSALLTAVVIIITIAIGNILVIMAVSLEKKLQATNYFLM",  
    "SLAIADMLLGFLVMPVSMILTILYGYRWP"  
  )  
  run_tmhmm_on_sequence(protein_sequence)  
}
```

run_tmhmm_to_file	<i>Creates a FASTA-like file, that has the locations of the amino acids.</i>
-------------------	--

Description

Creates a FASTA-like file, that has the locations of the amino acids.

Usage

```
run_tmhmm_to_file(fasta_filename, tmhmm_filename)
```

Arguments

fasta_filename path to a FASTA file

tmhmm_filename filename to write the TMHMM results to

Examples

```
if (is_tmhmm_installed()) {  
  
  # Write results to a temporary file  
  tmhmm_filename <- tempfile()  
  run_tmhmm_to_file(  
    fasta_filename = system.file(  
      "extdata", "tmhmm.fasta", package = "tmhmm"  
    ),  
    tmhmm_filename = tmhmm_filename  
  )  
  
  # Result is written to 'tmhmm_filename', cleaning up  
  file.remove(tmhmm_filename)  
}
```

set_up_tmhmm	<i>Set up TMHMM</i>
--------------	---------------------

Description

Set up TMHMM

Usage

```
set_up_tmhmm(folder_name = get_default_tmhmm_folder())
```

Arguments

folder_name	superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share by default, as can be obtained by get_default_tmhmm_folder
-------------	---

Author(s)

Richèl J.C. Bilderbeek

Examples

```
set_up_tmhmm()
```

tally_locatome	<i>Tally each location in a locatome</i>
----------------	--

Description

Tally each location in a locatome

Usage

```
tally_locatome(locatome)
```

Arguments

locatome	text similar to a FASTA file, with protein names and the locations of its amino acids
----------	---

Author(s)

Richèl J.C. Bilderbeek

Examples

```
tally_locatome("iMMoo")
```

tally_tmhs	<i>Count the number of transmembrane helices in a topology</i>
------------	--

Description

Count the number of transmembrane helices in a topology

Usage

```
tally_tmhs(topology)
```

Arguments

topology a tibble with two columns, named name for a protein's name and topology, for its predicted topology (e.g. ooooMMMiiii).

Value

a [tibble](#) with the number of TMHs per protein

Examples

```
if (is_tmhmm_installed()) {  
  tally_tmhs(  
    locatome_to_df(  
      c(  
        ">my_protein",  
        "FASTASTIC"  
      )  
    )  
  )  
}
```

tmhmm	<i>tmhmm: estimate the topoplogy of membrane proteins</i>
-------	---

Description

Each amino acid that is part a membrane protein can be located either in the cytosol, membrane or at the exterior. 'tmhmm' estimate this topology

Author(s)

Richèl J.C. Bilderbeek

Examples

```
if (is_tmhmm_installed()) {  
  fasta_filename <- system.file("extdata", "tmhmm.fasta", package = "tmhmm")  
  locatome <- run_tmhmm(fasta_filename)  
  
  # Show as text  
  message(locatome, sep = "\n")  
  
  # Show as data frame  
  message(locatome_to_df(locatome))  
  
  # Show as plot  
  plot_locatome(locatome)  
}
```

tmhmm_self_test	<i>Self-test the TMHMM installation</i>
-----------------	---

Description

Self-test the TMHMM installation

Usage

```
tmhmm_self_test()
```

Author(s)

Richèl J.C. Bilderbeek

Examples

```
if (is_tmhmm_installed()) {  
  tmhmm_self_test()  
}
```

uninstall_tmhmm	<i>Uninstall TMHMM</i>
-----------------	------------------------

Description

Uninstall TMHMM. Will stop if TMHMM is not installed.

Usage

```
uninstall_tmhmm(folder_name = get_default_tmhmm_folder())
```

Arguments

folder_name superfolder of TMHMM. The superfolder's name is /home/[user_name]/.local/share by default, as can be obtained by [get_default_tmhmm_folder](#)

Value

Nothing

Author(s)

Richèl J.C. Bilderbeek

Examples

```
uninstall_tmhmm()
```

Index

are_tmhs, 2

check_tmhmm_installation, 3

check_topology, 4

default_params_doc, 4

get_default_tmhmm_bin_path, 5, 5

get_default_tmhmm_folder, 3, 5, 6, 7, 8,
10–12, 14, 16, 19

get_tmhmm_url, 6

install_tmhmm, 7

install_tmhmm_bin, 7

is_locatome_line, 8

is_on_travis, 8

is_protein_name_line, 9

is_tmh, 9

is_tmhmm_bin_installed, 10

is_tmhmm_installed, 10

is_tmhmm_result, 11

is_tmhmm_set_up, 12

is_url_valid, 12

locatome_to_df, 13

plot_locatome, 13

run_tmhmm, 4, 13

run_tmhmm_on_sequence, 14

run_tmhmm_to_file, 15

set_up_tmhmm, 16

stop, 4

tally_locatome, 16

tally_tmhs, 17

tibble, 17

tmhmm, 17

tmhmm_self_test, 18

uninstall_tmhmm, 18