Package ‘cleaver’

April 5, 2014

Version 1.0.0

Date 2013-07-24

Title Cleavage of polypeptide sequences

Maintainer Sebastian Gibb `<mail@sebastiangibb.de>`

Depends R (>= 3.0.0), methods, Biostrings (>= 1.29.8)

Imports IRanges

Suggests testthat, knitr, BiocStyle (>= 0.0.14), BRAIN, UniProt.ws (>= 2.1.4)

Description In-silico cleavage of polypeptide sequences. The cleavage rules are taken from: http://web.expasy.org/peptide_cutter/peptidecutter_enzymes.html

License GPL (>= 3)

URL https://github.com/sgibb/cleaver/

BugReports https://github.com/sgibb/cleaver/issues/

LazyLoad yes

VignetteBuilder knitr

biocViews Bioinformatics, Proteomics

Author Sebastian Gibb [aut, cre]

R topics documented:

`cleaver-package` *Cleavage of polypeptide sequences*

Description

This package cleaves polypeptide sequences. It provides only one function: `cleave`

Details

The cleavage rules are taken from: `http://web.expasy.org/peptide_cutter/peptidecutter_enzymes.html`

Package:	cleaver
License:	GPL (>= 3)
URL:	https://github.com/sgibb/cleaver/

Author(s)

Sebastian Gibb `<mail@sebastiangibb.de>`

References

`https://github.com/sgibb/cleaver/`

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; "Protein Identification and Analysis Tools on the ExPASy Server". (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005).
`http://web.expasy.org/peptide_cutter/peptidecutter_enzymes.html`

See Also

`cleave`

`cleave` *Cleavage of polypeptide sequences*

Description

This function cleaves polypeptide sequences.

Usage

```
## S4 method for signature character
cleave(x, enzym = "trypsin", missedCleavages = 0)

## S4 method for signature AAString
cleave(x, enzym = "trypsin", missedCleavages = 0)

## S4 method for signature AAStringSet
cleave(x, enzym = "trypsin", missedCleavages = 0)
```

Arguments

`x`	polypeptide sequences.
`enzym`	`character`, cleavage rule.
`missedCleavages`	`numeric`, number of missed cleavages.

Details

The cleavage rules are taken from: `http://web.expasy.org/peptide_cutter/peptidecutter_enzymes.html`

Cleavage rules (cleavage between P1 and P1'):

Rule name	P4	P3	P2	P1	P1'
`arg-c proteinase`	-	-	-	R	-
`asp-n endopeptidase`	-	-	-	-	D
`bnps-skatole-c`	-	-	-	W	-
`caspase1`	F,W,Y,L	-	H,A,T	D	not P,E,D,Q,K,R
`caspase2`	D	V	A	D	not P,E,D,Q,K,R
`caspase3`	D	M	Q	D	not P,E,D,Q,K,R
`caspase4`	L	E	V	D	not P,E,D,Q,K,R
`caspase5`	L,W	E	H	D	-
`caspase6`	V	E	H,I	D	not P,E,D,Q,K,R
`caspase7`	D	E	V	D	not P,E,D,Q,K,R
`caspase8`	I,L	E	T	D	not P,E,D,Q,K,R
`caspase9`	L	E	H	D	-
`caspase10`	I	E	A	D	-
`chymotrypsin-high`	-	-	-	F,Y	not P
	-	-	-	W	not M,P
`chymotrypsin-low`	-	-	-	F,L,Y	not P
	-	-	-	W	not M,P
	-	-	-	M	not P,Y
	-	-	-	H	not D,M,P,W
`clostripain`	-	-	-	R	-
`cnbr`	-	-	-	M	-
`enterokinase`	D,E	D,E	D,E	K	-
`factor xa`	A,F,G,I,L,T,V,M	D,E	G	R	-
`formic acid`	-	-	-	D	-

glutamyl endopeptidase	-	-	-	D	-
granzyme-b	I	E	P	D	-
hydroxylamine	-	-	-	N	G
iodosobenzoic acid	-	-	-	W	-
lysc	-	-	-	K	-
lysn	-	-	-	-	K
neutrophil elastase	-	-	-	A,V	-
ntcb	-	-	-	-	C
pepsin1.3	-	not H,K,R	not P	not R	F,L,W,Y
	-	not H,K,R	not P	F,L,W,Y	-
pepsin	-	not H,K,R	not P	not R	F,L
	-	not H,K,R	not P	F,L	-
proline endopeptidase	-	-	not H,K,R	P	not P
proteinase k	-	-	-	A,E,F,I,L,T,V,W,Y	-
staphylococcal peptidase i	-	-	not E	E	-
thermolysin	-	-	-	not D,E	A,F,I,L,M,V
thrombin	-	-	G	R	G
	A,F,G,I,L,T,V,M	A,F,G,I,L,T,V,W	P	R	not D,E
trypsin	-	-	-	K,R	not P
	-	-	W	K	P
	-	-	M	R	P

Exceptions:

Rule name	Enzyme name	P4	P3	P2	P1	P1’	P2’
trypsin	-	-	C,D	K	D	-	
		-	-	C	K	H,Y	-
		-	-	C	R	K	-
		-	-	R	R	H,R	-

Rule name	Enzyme name
arg-c proteinase	Arg-C proteinase
asp-n endopeptidase	Asp-N endopeptidase
bnps-skatole-c	BNPS-Skatole
caspase1	Caspase 1
caspase2	Caspase 2
caspase3	Caspase 3
caspase4	Caspase 4
caspase5	Caspase 5
caspase6	Caspase 6
caspase7	Caspase 7
caspase8	Caspase 8
caspase9	Caspase 9
caspase10	Caspase 10
chymotrypsin-high	Chymotrypsin-high specificity (C-term to [FYW], not before P)
chymotrypsin-low	Chymotrypsin-low specificity (C-term to [FYWML], not before P)

`clostripain`	Clostripain (Clostridiopeptidase B)
`cnbr`	CNBr
`enterokinase`	Enterokinase
`factor xa`	Factor Xa
`formic acid`	Formic acid
`glutamyl endopeptidase`	Glutamyl endopeptidase
`granzyme-b`	Granzyme B
`hydroxylamine`	Hydroxylamine
`iodosobenzoic acid`	Iodosobenzoic acid
`lysc`	LysC
`lysn`	LysN
`neutrophil elastase`	Neutrophil elastase
`ntcb`	NTCB (2-nitro-5-thiocyanobenzoic acid)
`pepsin1.3`	Pepsin (pH == 1.3)
`pepsin`	Pepsin (pH > 2)
`proline endopeptidase`	Proline-endopeptidase
`proteinase k`	Proteinase K
`staphylococcal peptidase i`	Staphylococcal Peptidase I
`thermolysin`	Thermolysin
`thrombin`	Thrombin
`trypsin`	Trypsin

Value

a `list` of splitted polypeptides if input `x` was a `character` or a `AAStringSetList` if input `x` was an `AAStringSet`).

Author(s)

Sebastian Gibb

References

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; "Protein Identification and Analysis Tools on the ExPASy Server". (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005).
`http://web.expasy.org/peptide_cutter/peptidecutter_enzymes.html`

See Also

`AAString`, `AAStringSet`, `AAStringSetList`

Examples

```
library("cleaver")

## Gastic juice peptide 1 (UniProtKB/Swiss-Prot: GAJU_HUMAN/P01358)
gaju <- "LAAGKVEDSD"
```

```
cleave(gaju, "trypsin")
# $LAAGKVEDSD
# [1] "LAAGK" "VEDSD"

cleave(gaju, "trypsin", missedCleavages=1)
# $LAAGKVEDSD
# [1] "LAAGKVEDSD"

cleave(gaju, "trypsin", missedCleavages=0:1)
# $LAAGKVEDSD
# [1] "LAAGK" "VEDSD" "LAAGKVEDSD"

cleave(gaju, "pepsin")
# $LAAGKVEDSD
# [1] "LAAGKVEDSD"
# (no cleavage)


## use AAStringSet
gaju <- AAStringSet("LAAGKVEDSD")

cleave(gaju)
# AAStringSetList of length 1
# [["LAAGKVEDSD"]] LAAGK VEDSD LAAGKVEDSD
```

Index