

## List of POI characteristics and predictions

Target name/Identifier:	B56-alpha_human
No of AA:	486
MW:	56.2 kDa
Theor. pI:	6.27
Theor. Ext. coeff:	0.83
Description:	Q15172 (Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit alpha isoform, <i>H. sapiens</i> , Cytoplasm, Nucleus). Belongs to the phosphatase 2A regulatory subunit B family. Protein phosphatase 2A is one of the four major Ser/Thr phosphatases, and it is implicated in the negative control of cell growth and division. It consists of a common heteromeric core enzyme, which is composed of a catalytic subunit and a constant regulatory subunit, that associates with a variety of regulatory subunits. The B regulatory subunit might modulate substrate selectivity and catalytic activity. This gene encodes an alpha isoform of the regulatory subunit B56 subfamily
Signal peptide (see page 3) :	none
Transmembrane segments: (see page 4)	none
Predicted ECD: (see page 4)	none
Unique protease sies: (see page 5)	0 relevant for expression/purification purposes
Furin sites (see page 6):	2 potential
Glycosylation sites: (see page 7 + 8)	16 potential (O-glycosylation) 1 potential (N-glycosylation)
Antigenicity (see page 9):	16 antigenic determinants

## POI sequence

10            20            30            40            50            60  
MSSSSPPAGA ASAASISEK VDGFTRKSVR KAQRQKRSQG SSQFRSQQSQ AELHPLPQLK  
  
70            80            90            100          110          120  
DATSNQQQEL FCQKLQQCCI LFDFMDSVSD LKSKEIKRAT LNELVEYVST NRGVIVESAY  
  
130          140          150          160          170          180  
SDIVVKMISAN IFRTLPPPSDN PPDFDEEDEP TLEASWPHIQ LVYEFFLRLP ESPDPQPSIA  
  
190          200          210          220          230          240  
KRYIDQKFVQ QLLELFDSED PRERDFLKVTV LHRIYGKFQIN NIFLRFIYET  
  
250          260          270          280          290          300  
EHFNGVAELL EILGSIINGF ALPLKAHKQ FLMKVLPMPH TAKGLALFHQ QLAYCVVQFL  
  
310          320          330          340          350          360  
EKDTTLTEPV IRGLLKFWPK TCSQKEVMFL GEIEEILDVI EPTQFKKIEE PLFKQISKCV  
  
370          380          390          400          410          420  
SSSHFQVAER ALYFWNNYEI LSLIEENIDK ILPIMFASLY KISKEHWNPVALVYNVLK  
  
430          440          450          460          470          480  
TLMEMNGKLF DDLTSSYKAE RQREKKKELE REELWKLEE LKLKKALEKQ NSAYNMHSIL

SNTSAE

## Targetname/Identifier: B56-alpha\_human

### Uniprot blast

<https://www.uniprot.org/blast/>

Entry Variant viewer Feature viewer Genomic coordinates Publications External links History

Feature viewer [J. Download](#)

Molecule processing Chain Initiator methionine Sequence information Compositional bias Sequence conflict Domains Region PTM Modified residue Modified residue (large scale data) Antigenic sequences Antibody binding sequences Variants Proteomics PDBe 3D structure coverage AlphaFold

SOURCE IDENTIFIER METHOD RESOLUTION CHAIN POSITIONS LINKS

PDB 6NTS EM 2.63 Å B 1-486 PDBe - RCSB-PDB - PDBsum AlphaFold AF-Q15172-F1 Predicted 1-486 AlphaFold

Family & Domains' Features

Showing features for region<sup>1</sup>, compositional bias.

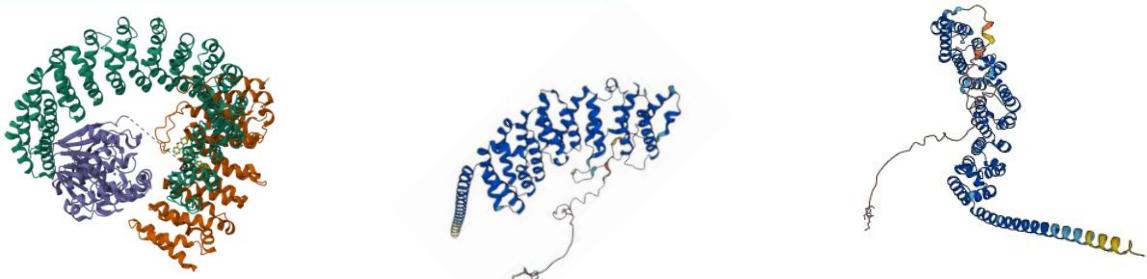
Type ID POSITION(S) DESCRIPTION

Region 1-52 Disordered Automatic Annotation BLAST Add

Sequence: MSSSSPPAAGAAASIAASEKVDGFTRKSVRIKAQKQKRSQGSQFQSQQAE

Compositional bias 35-52 Polar residues Automatic Annotation BLAST Add

Sequence: QKRSQGSSQFRSQGSQAE



## Targetname/Identifier: B56-alpha\_human

### Signal Sequence Analysis

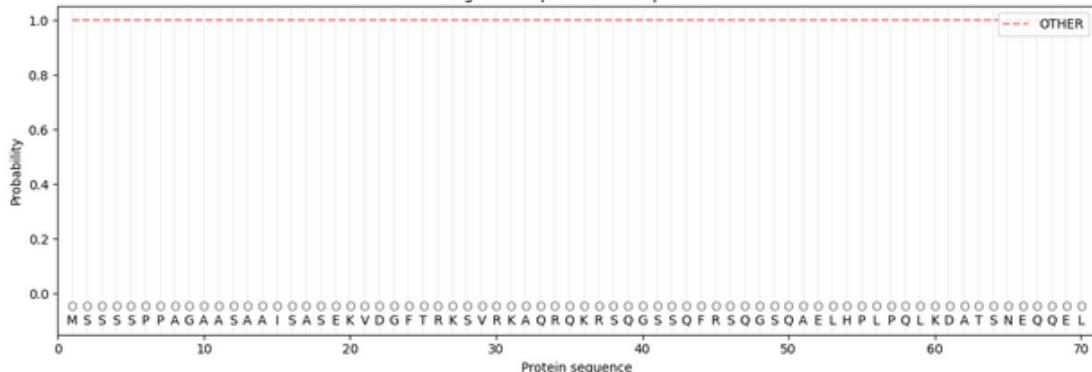
<https://services.healthtech.dtu.dk/services/SignalP-6.0/>

Sequence  
Prediction: Other

Protein type	Other	Signal Peptide (Sec/SPI)	Lipoprotein signal peptide (Sec/SPII)	TAT signal peptide (Tat/SPI)	TAT Lipoprotein signal peptide (Tat/SPII)	Pilin-like signal peptide (Sec/SPIII)
Likelihood	1	0	0	0	0	0

Download: [PNG](#) / [EPS](#) / [Tabular](#)

SignalP 6.0 prediction: Sequence



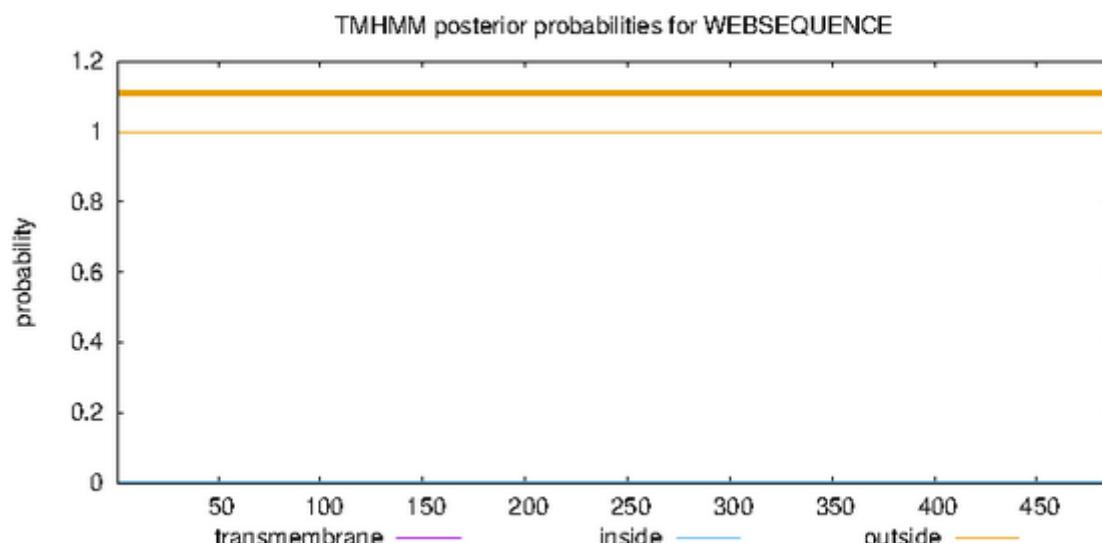
## Targetname/Identifier: B56-alpha\_human

### TMHMM result

<https://services.healthtech.dtu.dk/services/TMHMM-2.0/>

### TMHMM result

```
# WEBSEQUENCE Length: 486
# WEBSEQUENCE Number of predicted TMHs: 0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.0761799999999999
# WEBSEQUENCE Exp number, first 60 AAs: 0
# WEBSEQUENCE Total prob of N-in: 0.00106
WEBSEQUENCE      TMHMM2.0      outside      1      486
```



## Targetname/Identifier: B56-alpha\_human

### Potential Cleavage Sites

[https://web.expasy.org/peptide\\_cutter/](https://web.expasy.org/peptide_cutter/)

These enzymes cleave the sequence:

Name of enzyme	No. of cleavage
Arg-C proteinase	21
Asp-N endopeptidase	20
Asp-N endopeptidase + N-terminal Glu	67
BNPS-Skatole	5
CNBr	10
Chymotrypsin-high specificity (C-term to [FYW], not before P)	47
Chymotrypsin-low specificity (C-term to [FYWML], not before P)	117
Clostripain	21
Formic acid	20
Glutamyl endopeptidase	47
Hydroxylamine	3
Iodosobenzoic acid	5
LysC	42
LysN	42
NTCB (2-nitro-5-thiocyanobenzoic acid)	6
Pepsin (pH1.3)	114
Pepsin (pH>2)	135
Proline-endopeptidase [*]	1
Proteinase K	256
Staphylococcal peptidase I	41
Thermolysin	152
Trypsin	63

These chosen enzymes do not cut:

Caspase1  
Caspase10  
Caspase2  
Caspase3  
Caspase4  
Caspase5  
Caspase6  
Caspase7  
Caspase8  
Caspase9  
Enterokinase  
Factor Xa  
GranzymeB  
Thrombin  
Tobacco etch virus protease

Targetname/Identifier: B56-alpha\_human

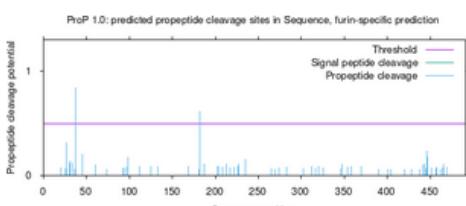
## Potential Furin or Propeptide Cleavage Sites

<https://services.healthtech.dtu.dk/services/ProP-1.0/>

Signal peptide cleavage site predicted: none

Propeptidase cleavage sites predicted:     $\text{acs}(\text{a})/\text{lys}(\text{b}) = 3$

Name	Pos	Context	Score	Pred
Sequenze	10	A\$AS\$AK\$IV	0.878	-
Sequenze	14	KVQD\$K\$K\$IV	0.878	-
Sequenze	27	WDG\$T\$K\$IV	0.817	-
Sequenze	30	PTK\$K\$K\$IV	0.818	-
Sequenze	31	TIV\$K\$K\$IV	0.848	-
Sequenze	34	SVKA\$K\$IV	0.820	-
Sequenze	36	KA\$K\$K\$IV	0.862	-
Sequenze	37	KA\$K\$K\$IV	0.862	"Prop"
Sequenze	45	W\$K\$K\$IV	0.862	-
Sequenze	46	W\$K\$K\$IV	0.862	-
Sequenze	74	Q\$K\$K\$IV	0.874	-
Sequenze	92	D\$S\$D\$K\$IV	0.876	-
Sequenze	94	V\$D\$D\$K\$IV	0.868	-
Sequenze	97	L\$K\$K\$K\$IV	0.888	-
Sequenze	98	K\$K\$K\$K\$IV	0.875	-
Sequenze	112	Y\$V\$T\$M\$IV	0.889	-
Sequenze	124	W\$V\$T\$M\$IV	0.889	-
Sequenze	133	Z\$A\$N\$M\$IV	0.891	-
Sequenze	148	V\$V\$P\$M\$IV	0.890	-
Sequenze	181	P\$Q\$S\$K\$IV	0.881	-
Sequenze	183	Q\$P\$S\$K\$IV	0.818	"Prop"
Sequenze	187	K\$Y\$D\$K\$IV	0.876	-
Sequenze	202	P\$D\$E\$K\$IV	0.888	-
Sequenze	204	S\$D\$P\$R\$IV	0.889	-
Sequenze	208	E\$R\$P\$R\$IV	0.889	-
Sequenze	213	W\$R\$P\$R\$IV	0.886	-
Sequenze	217	L\$H\$Y\$G\$IV	0.873	-
Sequenze	222	G\$K\$P\$G\$IV	0.881	-
Sequenze	226	G\$K\$A\$K\$IV	0.887	-
Sequenze	227	L\$A\$P\$K\$IV	0.813	-
Sequenze	235	Z\$N\$H\$K\$IV	0.812	-
Sequenze	245	G\$A\$P\$K\$IV	0.888	-
Sequenze	249	P\$K\$A\$K\$IV	0.888	-
Sequenze	251	W\$K\$A\$K\$IV	0.888	-
Sequenze	253	Z\$P\$K\$A\$K\$IV	0.879	-
Sequenze	302	V\$V\$O\$K\$IV	0.888	-
Sequenze	311	L\$T\$P\$V\$IV	0.888	-
Sequenze	316	V\$V\$G\$K\$IV	0.888	-
Sequenze	320	L\$K\$P\$K\$TC	0.891	-
Sequenze	325	P\$K\$T\$C\$IV	0.872	-
Sequenze	346	I\$P\$T\$K\$IV	0.883	-
Sequenze	347	E\$P\$T\$K\$IV	0.869	-
Sequenze	350	W\$P\$T\$K\$IV	0.888	-
Sequenze	352	L\$W\$Q\$K\$IV	0.889	-
Sequenze	370	M\$Q\$V\$K\$IV	0.884	-
Sequenze	390	ZE\$Z\$N\$K\$IV	0.881	-
Sequenze	401	M\$A\$S\$Y\$IV	0.858	-
Sequenze	404	S\$LY\$K\$K\$IV	0.857	-
Sequenze	420	L\$V\$V\$K\$K\$IV	0.881	-
Sequenze	428	W\$W\$W\$K\$K\$IV	0.881	-
Sequenze	435	D\$S\$S\$K\$K\$IV	0.886	-
Sequenze	436	E\$S\$S\$K\$K\$IV	0.886	-
Sequenze	443	T\$K\$A\$K\$K\$IV	0.804	-
Sequenze	445	A\$K\$E\$K\$K\$IV	0.885	-
Sequenze	446	E\$K\$O\$K\$K\$IV	0.833	-
Sequenze	447	R\$K\$E\$K\$K\$IV	0.778	-
Sequenze	451	K\$K\$E\$K\$K\$IV	0.871	-
Sequenze	455	E\$K\$E\$K\$K\$IV	0.887	-
Sequenze	457	M\$E\$E\$K\$K\$IV	0.881	-
Sequenze	461	E\$E\$E\$K\$K\$IV	0.881	-
Sequenze	464	L\$E\$L\$K\$K\$IV	0.879	-
Sequenze	465	E\$E\$K\$K\$K\$AL	0.110	-
Sequenze	469	L\$K\$A\$K\$K\$AL	0.883	-



## Targetname/Identifier: B56-alpha\_human

### O-GalNAc (Mucin Type) Glycosylation Sites in mammalian proteins

<https://services.healthtech.dtu.dk/services/NetOGlyc-4.0/>

```
##gff-version 2
##source-version NetOGlyc 4.0.0.13
##date 24-5-3
##type Protein
#seqname source feature start end score strand frame comment
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 2 2 0.895696 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 3 3 0.929668 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 4 4 0.940091 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 5 5 0.910834 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 12 12 0.962965 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 16 16 0.962826 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 18 18 0.916524 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 25 25 0.845743 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 28 28 0.937804 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 38 38 0.921968 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 41 41 0.846378 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 42 42 0.838653 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 46 46 0.845417 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 49 49 0.339024 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 63 63 0.13659 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 64 64 0.561966 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 87 87 0.0251852 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 89 89 0.0228268 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 93 93 0.199126 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 100 100 0.172741 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 109 109 0.339422 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 110 110 0.0036428 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 118 118 0.0751394 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 121 121 0.0586107 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 128 128 0.321726 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 134 134 0.423889 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 138 138 0.796482 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 151 151 0.161284 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 155 155 0.0934532 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 172 172 0.0284876 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 178 178 0.282604 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 198 198 0.0733517 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 209 209 0.0257842 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 240 240 0.0157936 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 255 255 0.0471928 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 281 281 0.0920847 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 304 304 0.0297872 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 305 305 0.0287558 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 307 307 0.0928565 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 321 321 0.0355398 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 323 323 0.0323146 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 343 343 0.0535596 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 357 357 0.0760477 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 361 361 0.0045963 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 362 362 0.0141773 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 363 363 0.0479851 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 382 382 0.00359155 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 398 398 0.0595215 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 403 403 0.0648861 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 410 410 0.146634 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 421 421 0.0401934 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 434 434 0.258633 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 435 435 0.347652 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 436 436 0.516116 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 472 472 0.194274 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 478 478 0.201444 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 481 481 0.150218 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 483 483 0.0613976 .
SEQUENCE netOGlyc-4.0.0.13 CARBOHYD 484 484 0.0741123 . #POSITIVE
```

## Targetname/Identifier: B56-alpha\_human

### N-linked glycosylation sites in human proteins

<https://services.healthtech.dtu.dk/services/NetNGlyc-1.0/>

```
#####
#
```

Warning! This sequence may not contain a signal peptide!!

Proteins without signal peptides are unlikely to be exposed to  
the N-glycosylation machinery and thus may not be glycosylated  
(in vivo) even though they contain potential motifs.

SignalP-NN euk predictions are as follows:

#	name	Cmax	pos	?	Ymax	pos	?	Smax	pos	?	Smean	?	D	?	?
Sequence		0.150	24	0.127	24	0.138	1	0.109	0.117	N	0.450		SignalP-noTM		

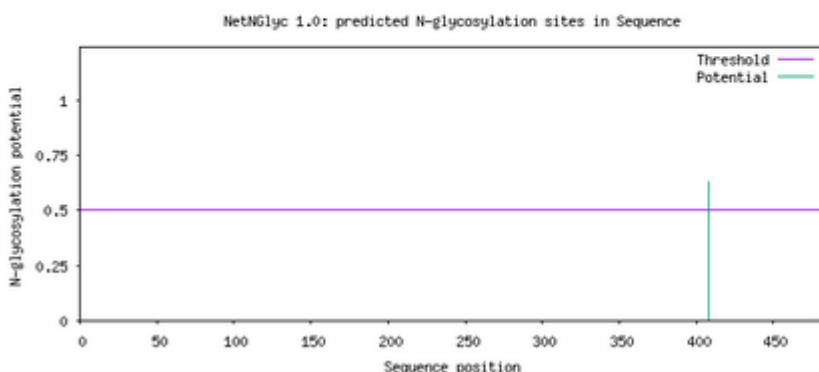
SignalP output is explained at <https://services.healthtech.dtu.dk/services/SignalP-4.1/output.php>

```
#####
#
```

Name	Sequence	Length	486
MSSSPPAGAASAAISASEKVDGFTRKSVRKAQRQKRSQGSSQFRSQGSQAEHLHPLPQLKDATSNEQQELFCQKLQQCCI		80	
LFDFMDSVSOLKSKEIKRATLNLVEVYSTNRGVIVESAYSDIVKMISANIRRTLPPSDNPQDFPEEDEPTLEASHPHIQ		160	
LVYEFFFLRFLESPOFQPSIAKRYIDQKFVQQLLELFQDSEDPRERDFLKTVLHRIYGKFLGLRAFIKRQINNIFLRFIYET		240	
EHFNGVAELLEILGSIINGFALPLKAEHKKQFLMKVLIPMHTAKGLALFHAQLAYCVVQFLEKDTTLTEPVIRGLLKFWPK		320	
TCSQKEVMFLGEIEEILDVIEPTQFKKIEEPLFKQISKCVSSSHFQVAERALYFVINNEYILSLIEENIDKILPIMFASLY		400	
KISKEHN <del>N</del> PTI <del>V</del> VALVYNVLKTLMEMNGKLFDDLTSSYKAERQREKKELEREELWKKLEELKLKKALEKQNSAYNMHSIL		480	
SNTSAE		560	
.....		80	
.....		160	
.....		240	
.....		320	
.....		400	
.....N.....		480	
.....		560	

(Threshold=0.5)

SeqName	Position	Potential	Jury	N-Glyc	agreement result
Sequence	488	NPTI	0.6317	(8/9)	+
Sequence	482	NTSA	0.4229	(6/9)	-

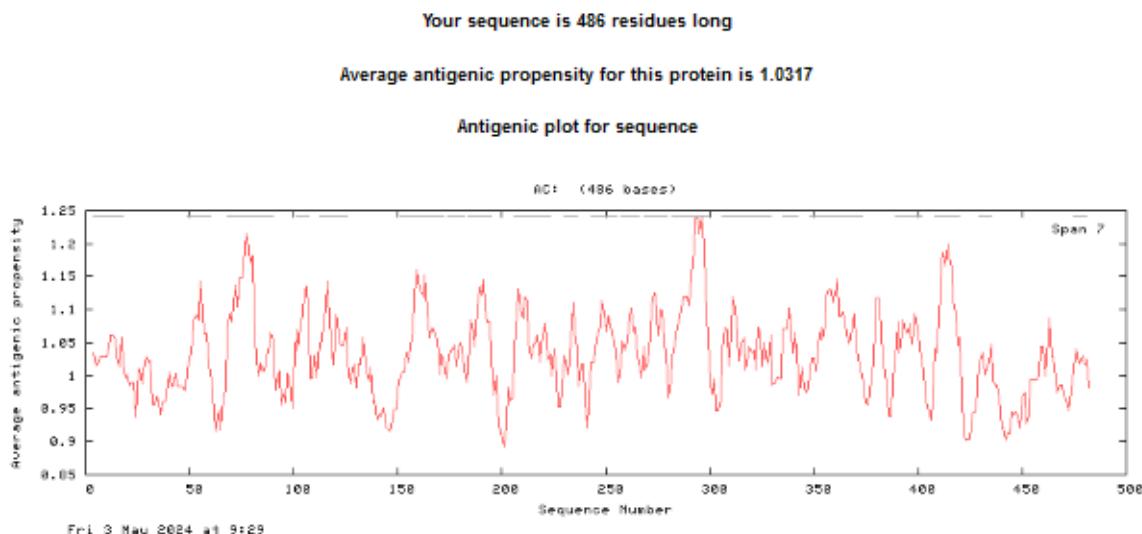


[Graphics in PostScript](#)

## Targetname/Identifier: B56-alpha\_human

### Antigenic Plot

<http://imed.med.ucm.es/Tools/antigenic.pl>



There are 21 antigenic determinants in your sequence:

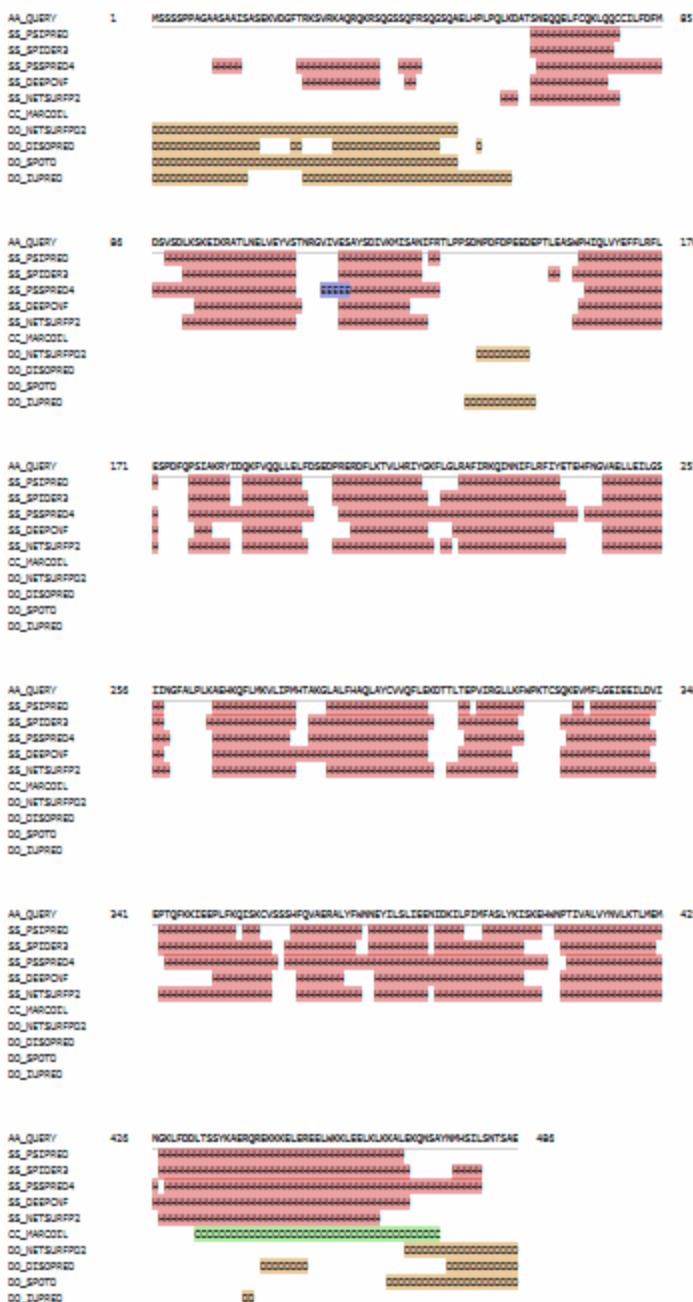
n	Start Position	Sequence	End Position
1	4	SSPPAGAAASAIAS	18
2	49	SQAEHLPLPQLK	60
3	68	QELFCQKLQQCCILDFMDSVSD	90
4	101	LNELVEY	107
5	112	RGVIVESAYSDIVKM	126
6	151	TLEASWPHIQLVYEFFLRFLES	172
7	174	DFQPSIAKR	182
8	184	IDQKFVQQLLEL	195
9	206	FLKTVLHRIYGKFLGLRA	223
10	230	NNIFLRF	236
11	243	FNGVAELLEILGS	255
12	257	INGFALPLKA	266
13	268	HKQFLMKVLIPM	279
14	282	AKGLALFHQLAYCVVQF	299
15	306	LTEPVIRGLLKFWPKTCSQKEVMF	329
16	335	EILDVIEP	342
17	348	IIEPLFKQISKCVSSSHFQVAERALY	373
18	389	DKILPIMFASLYKI	402
19	408	NPTIVALVYNVLK	420
20	429	LFDDLTS	435
21	475	NMHSILS	481

## Targetname/Identifier: B56-alpha\_human

### MPI Toolkit – Quick2D

<https://toolkit.tuebingen.mpg.de/tools/quick2d>

Protein ID: Q\_3264644



SS = α-helix β-strand π-helix    CC = Coiled Coils    TM = Transmembrane    DO = Disorder