

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 ASAAISASEK VDGFTTRKSVR KAQRQKRSGG SSQFRSQGSQ AELHFLPQLK

70      80      90     100     110     120
DATSNEQQEL FCQKLQCCCI LFDPMDSVSD LKSKEIKRAT LNELVEYVST NRGVIVESAY

130     140     150     160     170     180
SDIVKMISAN IFRTLPPSDN PDFDPDEEDF TLEASWPHIQ LVYEFFLRLFL ESPDFQPSIA

190     200     210     220     230     240
KRYIDQKFVQ QLLELFDSED PRERDFLKTV LHRIFYGKFLG LRAFIKQKIN NIFLRFIYET

250     260     270     280     290     300
EHFNGVAELL EILGSIINGF ALPLKAEHKQ FILMKVLIPMH TAKGLALPHA QLAYCVVQFL

310     320     330     340     350     360
EKDITLTPEPV IRLGLKFWPK TCSQKEVMFL GEIEEILDVI EPTQFKKIEE PLFKQISKCV

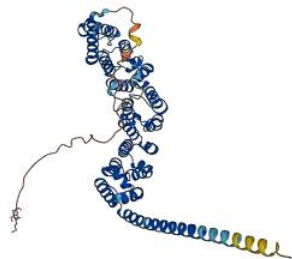
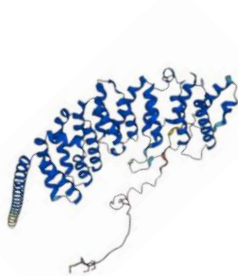
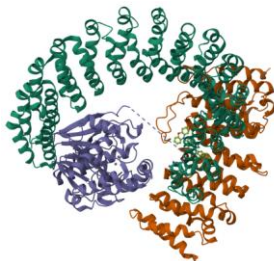
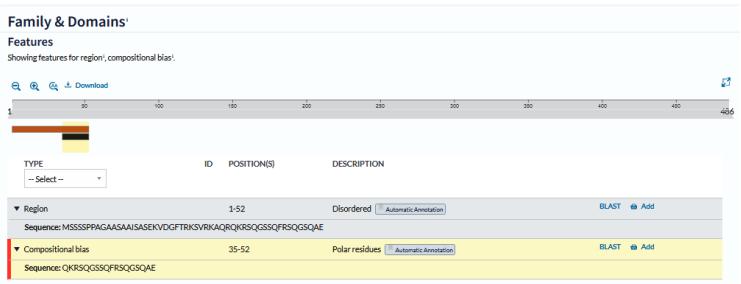
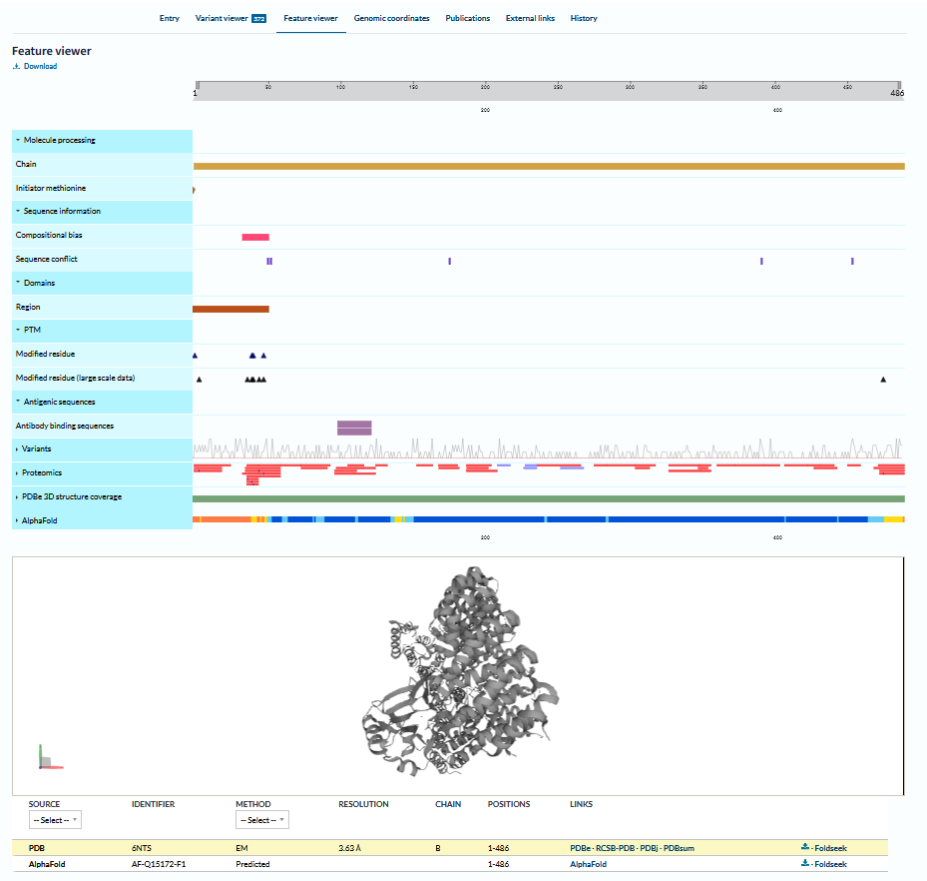
370     380     390     400     410     420
SSSHFQVAER ALYFWNNEYI LSLIEENIDK ILPIMFASLY KISKEHWNPT IVALVYNVLK

430     440     450     460     470     480
TLMEMNGKLF DDLTSSYKAE RQREKKKELE REELWKKLEE LKLLKKALEKQ NSAYNMHSIL

SNTSAE
```

Targetname/Identifier: B56-alpha_human

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



Targetname/Identifier: B56-alpha_human

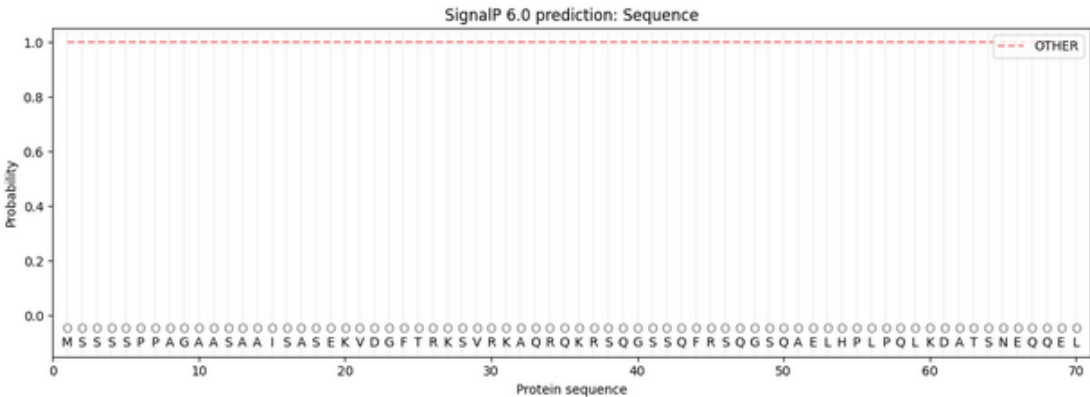
Signal Sequence Analysis

<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](#)



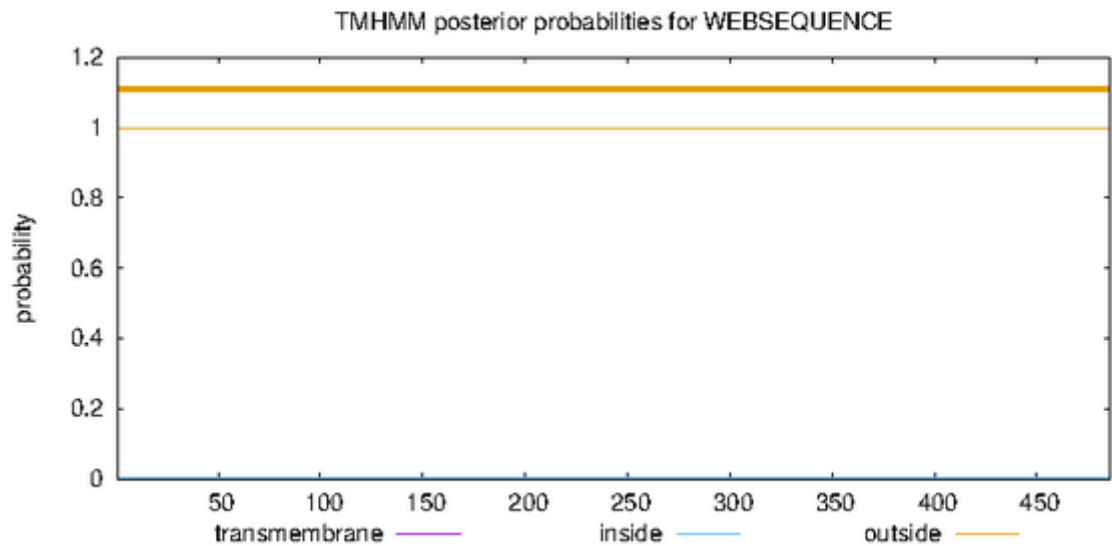
Targetname/Identifier: B56-alpha_human

TMHMM result

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

TMHMM result

```
# WEBSEQUENCE Length: 486
# WEBSEQUENCE Number of predicted TMs: 0
# WEBSEQUENCE Exp number of AAs in TMs: 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/

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

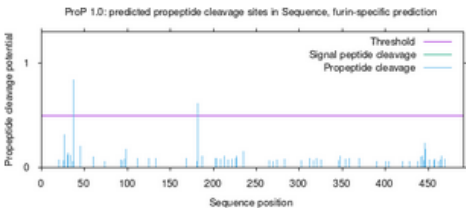
Potential Furin or Propeptide Cleavage Sites

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

```
488 Sequence
MSSSPPAGASASISSEKVDGATKSVKKAQKQMSQSSSQMSQSSQKELHPLPQLKDATSNEQQLPCQKLQCCZ 80
LMDPMSVSOLKSEKSKMATLNELVYVSTNGVVEASYSQVKNLSANLPTLPSSQWPDPRDEPTLEASIPHSIQ 160
LVYEPFLAPLESPPQPSQASRYIQQQLLEPSSDPREKPKITVLHNTIGPLGLRNPDKQINIDPLRYPTST 240
EHWNVQVALLLEISGSDGAPLSEKEMDPLHWLSDPTKAGLALPHAQLAYCVVQPLEKDTTTERVINGLLLPMPK 320
TCSQEVNPLGEEZEILDVIEPTPKKKEEPLKQISKCVSSHPQVABALYFNNVYVLSLEENIDKILPDMASLY 400
KISKEHNPITVALVYVNLKLEHNVGKLPDOLSSYKAEKQKKEKLEZEELKKLEELKKALEKQNSATYVMSIL 480
SNTSSE 560
.....F..... 60
.....P..... 160
..... 240
..... 320
..... 400
..... 480
..... 560

Signal peptide cleavage site predicted: none
Propeptide cleavage sites predicted: Arg(R)/Lys(K): 2
```

name	Pos	Context	Score	Pred
Sequence	10	ACISASEK	0.070	-
Sequence	16	KVDGATK	0.067	-
Sequence	27	VDPGATK	0.117	-
Sequence	30	PKMSVKA	0.110	-
Sequence	31	TKMSVKA	0.140	-
Sequence	34	SVKKAQK	0.120	-
Sequence	36	KKAQKQK	0.062	-
Sequence	37	KKAQKQK	0.041	*prop*
Sequence	48	QSSSQMS	0.103	-
Sequence	60	HPLPQLK	0.105	-
Sequence	74	QELPCQK	0.056	-
Sequence	92	DSVSOLK	0.076	-
Sequence	94	VDSLSIK	0.060	-
Sequence	97	LKSEKIK	0.080	-
Sequence	98	KSEKIK	0.175	-
Sequence	112	ETVSTNK	0.069	-
Sequence	126	AYSEVNI	0.060	-
Sequence	133	ISANIPK	0.091	-
Sequence	160	VYEPFLK	0.090	-
Sequence	181	KPSIAK	0.062	-
Sequence	182	QPSIAK	0.110	*prop*
Sequence	187	KRYIQQK	0.106	-
Sequence	202	MSDPPK	0.086	-
Sequence	204	SDPPKPK	0.069	-
Sequence	209	REMDPLK	0.077	-
Sequence	213	LKTVLHK	0.106	-
Sequence	217	LHMYGK	0.073	-
Sequence	222	GKPLGK	0.081	-
Sequence	226	GLKAPK	0.067	-
Sequence	227	LKAPDK	0.113	-
Sequence	236	INMPLK	0.191	-
Sequence	260	GPALPK	0.060	-
Sequence	269	PLKAEK	0.060	-
Sequence	274	MQQLNK	0.071	-
Sequence	283	IPWTKK	0.079	-
Sequence	302	VVQLKIK	0.060	-
Sequence	312	LTPVKIK	0.060	-
Sequence	316	VMSLKK	0.060	-
Sequence	320	LKPNPK	0.091	-
Sequence	325	PKTCQK	0.072	-
Sequence	346	IEPTQK	0.063	-
Sequence	347	ETPTQK	0.109	-
Sequence	394	IEEPLK	0.081	-
Sequence	398	LKQISK	0.069	-
Sequence	370	MPQVAK	0.064	-
Sequence	390	IEEIDK	0.061	-
Sequence	401	MPASLYK	0.060	-
Sequence	404	SLYISK	0.057	-
Sequence	420	LVYVNLK	0.061	-
Sequence	426	LHEVNGK	0.061	-
Sequence	438	DLSSYK	0.060	-
Sequence	441	SSYKAEK	0.100	-
Sequence	443	YKAEKQK	0.106	-
Sequence	445	AKQREK	0.060	-
Sequence	446	EQQREK	0.110	-
Sequence	447	MQREKKK	0.170	-
Sequence	451	KKKELEK	0.071	-
Sequence	455	EEELKIK	0.067	-
Sequence	457	EEELKIK	0.062	-
Sequence	461	KKLEELK	0.060	-
Sequence	464	LEELKIK	0.070	-
Sequence	465	EEELKIK	0.110	-
Sequence	469	LKALAKK	0.063	-



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				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	3	3	0.929668				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	4	4	0.940091				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	5	5	0.910834				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	12	12	0.962965				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	16	16	0.962826				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	18	18	0.916524				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	25	25	0.845743				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	28	28	0.937804				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	38	38	0.921968				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	41	41	0.846378				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	42	42	0.838653				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	46	46	0.845417				#POSITIVE
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				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	87	87	0.0251052				
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	0.0936428				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	118	118	0.0751394				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	121	121	0.0506107				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	128	128	0.321726				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	134	134	0.423809				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	138	138	0.796482				#POSITIVE
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	151	151	0.161204				
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	0.0257042				
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	0.0920047				
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	0.0945963				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	362	362	0.0141773				
SEQUENCE	netOGlyc-4.0.0.13	CARBOHYD	363	363	0.0479051				
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				#POSITIVE
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				

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

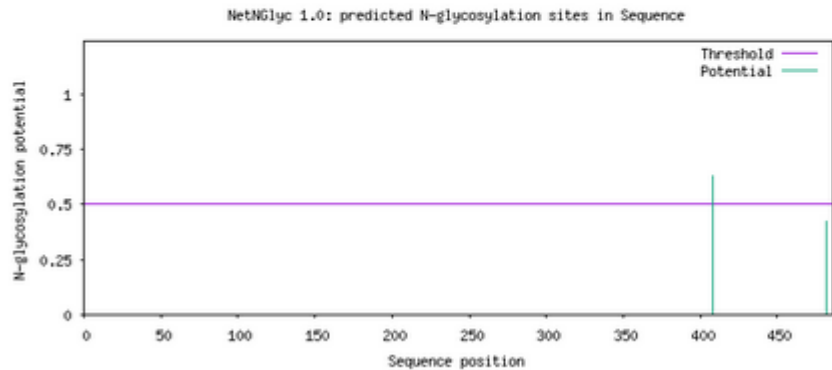
SignalP output is explained at https://services.healthtech.dtu.dk/services/SignalP-4.1/output.php

#####

Name: Sequence          Length: 486
MSSSSPPAGAAASAAISASEKVDGFTKRSVRKAQRQKRSQGSQFRSQGSQAELHPLPQLKDATSNEQQELFCQKLQCCCI      80
LFDPMDSVSLKSKEIKRATLNELVEYVSTNRRGVIVESAYSQIVKMISANIFRTLPPSDNPOFDPEDEPTLEASWPHIQ      160
LVYEFFLRFLSPDFQPSIAKRYIDQKFKVQQLLELFDSEDPREDFLKTVLHRIYQKFLGLRAFIRKQINNIFLRFIYET      240
EHFNGVAELLEILGSIINGFALPLKAHKKQFLMKVLIIPMHTAKGLALFHAQLAYCVVQFLEKDTTLTEPVIRGLLKFWPK      320
TCSQKEVMFLGEIEEILDVIEPTQFKKIEEPLFKQISKCVSSSHFQVAERALYFWMNEYILSLIEENIDKILPIMFASLY      400
KISKEHWNPTIVALVYNVLKTLMEMNGKLFDDLTSSYKAERQREKKKELEREELWKKLEELKLKALEKQNSAYNMHSIL      480
SNTSAE
.....
.....
.....
.....
.....N.....
.....

(Threshold=0.5)

-----
SeqName      Position  Potential  Jury      N-Glyc
              agreement result
-----
Sequence     408 NPTI    0.6317    (8/9)    +   WARNING: PRO-X1.
Sequence     482 NTSA    0.4229    (6/9)    -
-----
```



[Graphics in PostScript](#)

Targetname/Identifier: B56-alpha_human

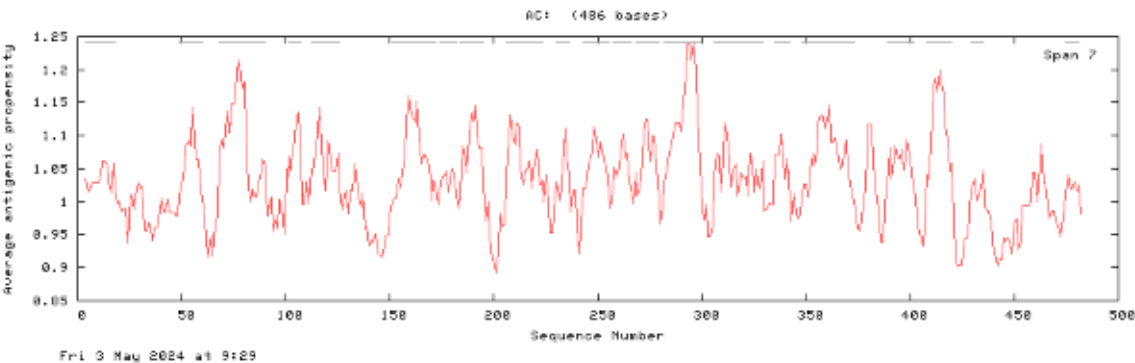
Antigenic Plot

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

Your sequence is 486 residues long

Average antigenic propensity for this protein is 1.0317

Antigenic plot for sequence



There are 21 antigenic determinants in your sequence:

n	Start Position	Sequence	End Position
1	4	SSPPAGAASAAISAS	18
2	49	SQAELHPLQLK	60
3	68	QELFCQKLQQCCILDFMDSVSD	90
4	101	LNELVEY	107
5	112	RGVIVESAYSDEVKM	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	AKGLALFHAQLAYCVVQF	299
15	306	LTEPVIRGLLKFWPKTCSQKEVMF	329
16	335	EILDVIEP	342
17	348	IEEPLFKQISKCVSSSHFQVAERALY	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

