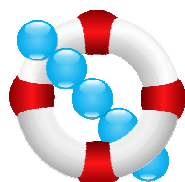


PROTEIN SEQUENCE ANALYSIS



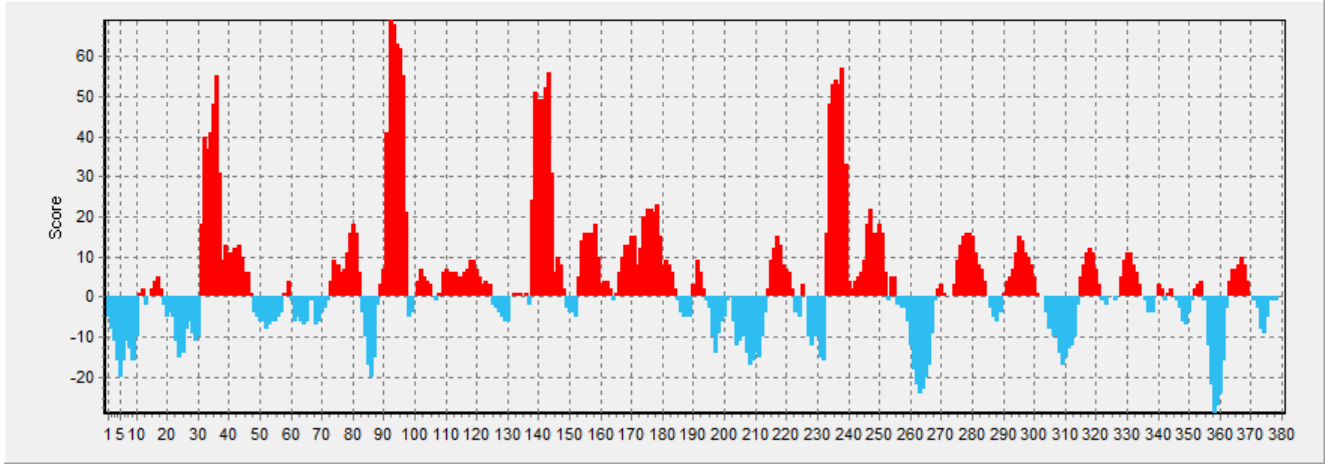
Performed using
EpiQuest™ 2013 v2.18

Order No	EXAMPLE
Date	May 20, 2013
Molecule	NS1
Acc. No	GenBank: CAA78918.1
Species of Origin	Dengue virus 2
Intended Species for Immunisation	Mouse
Sequence (FASTA)	MNSRSTSLSVSQVLVGIIVTLYLGMVQADSGCVVSWKNKELKCGSGIFVTDNVHTRTEQYKFQPESPSKL ASAIQKAHEEGICGIRSVTRLENLMWKQITSELNHILSENEVKLTIMTGDIKGIMQVGKRSLRPQPTEL YSWKTWGWKAKMLSTELHNQTFLLIDGPETAECPTNRAWNSLEVEDYGFVFTTNIWLRRLREKQDAFCDSK LMSAAIKDNRAVHADMGYWIESALNDTWKIEKASFIEVKSCHWPKSHTLWSNGVLESEMVI PKNFAGPKS QHNNRPGYHTQTAGPWHLGKLEMDDFDFCEGTTVVVTEDCGNRGP SLRRTTASGKLITWCCRCTLPLPLR YRGEDGCWYGMEIRPLKEKEENLVSSLVTA
Length	380
Region of Interest	Find immunodominant epitopes; whole molecule.

Comment: *This analysis demonstrates that use of EpiQuest-B (immunogenicity), EpiQuest-C (complexity) and Surface probability allows to correctly identify immunodominant epitopes in highly immunogenic protein. The immunodominant epitopes for NS1 are known, and correspond to those identified in this analysis.*

Predicted Immunodominant Epitopes

Whole Sequence overview:



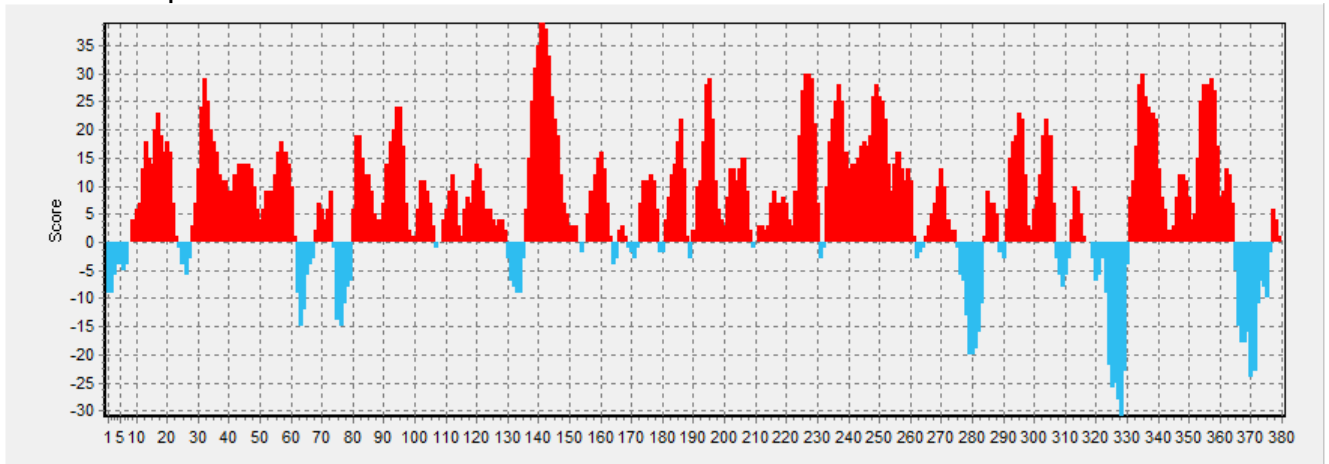
Positions and relative "weight" of Epitopes

MATRIX B7.1
 Frame 6
 Gap 0
 Threshold 0
 Peptide size 8 and above

Start	End	Length	Sequence	AGI	APB
89	97	9	TRLENLMWK	389	43
138	148	11	ELRYSWKTWGK	338	30
233	252	20	ALNDTWKIEKASFIEVKSCH	456	22
31	47	17	GCVVSWKNKELKCGSGI	362	21
165	184	20	GPETAECPTNRAWNSLEVE	249	12
73	82	10	AIQKAHEEGI	101	10
153	163	11	STELHNQTFLI	108	9
214	222	9	AAIKDNRAV	74	8
269	285	17	MVIPKNFAGPKSQHNNR	123	7
290	303	14	TQTAGPWHLGKLEM	92	6
363	370	8	IRPLKEKE	48	6

Analysis of Immunologic Complexity

Whole Sequence overview:



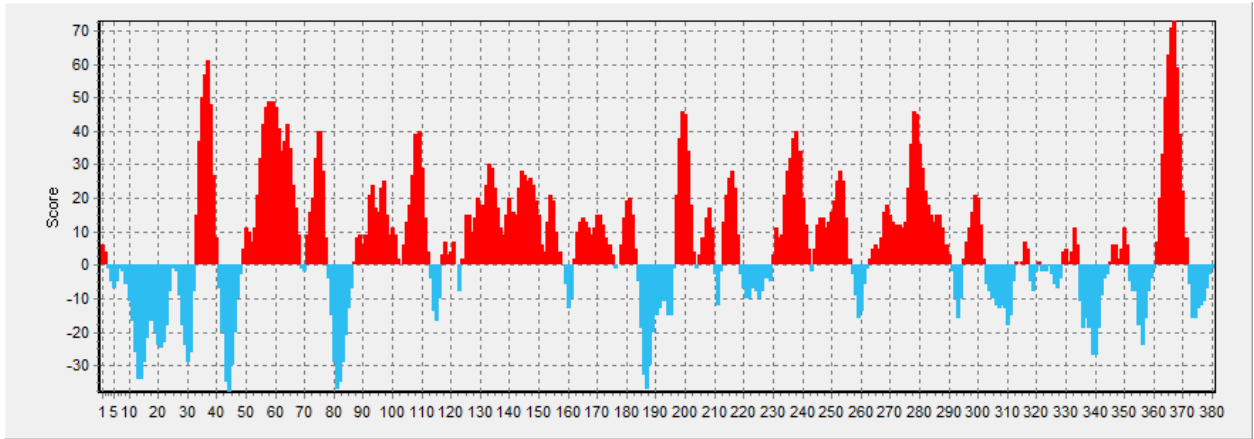
Positions of sequences with sufficient complexity

MATRIX C1.3
 Frame 6
 Gap 0
 Threshold 0
 Peptide size 8 and above

Start	End	Length	Sequence	AGI	APB
8	23	16	LSVSQVLVGIVTLYLG	197	12
28	61	34	ADSGCVVSWKNKELKCGSGIFVTDNVHTRTEQYK	428	12
80	106	27	EGICGIRSVTRLENLMWKQITSELNHI	287	10
108	129	22	SENEVKLTIMTGDIKGIMQVGK	141	6
136	153	18	PTELRYSWKTWGWKAKMLS	322	17
155	163	9	ELHNQTFLI	78	8
181	188	8	LEVEDYGF	92	11
190	208	19	VFTTNIWLRLREKQDAFCD	228	12
210	230	21	KLMSAAIKDNRAVHADMGYWI	235	11
233	261	29	ALNDTWKIEKASFIEVKSCHWPKSHTLWS	500	17
265	274	10	LESEMPIKN	57	5
291	306	16	QTAGPWHLGKLEMDFD	213	13
331	364	34	ASGKLITWCCRCTLPLRYRGEDGCWYGMEIR	501	14

Predicted probability of surface exposure

Whole Sequence overview:



Positions of the best sequences

MATRIX EM3.2
 Frame 6
 Gap 0
 Threshold 0
 Peptide and
 size 8 above

Start	End	Length	Sequence	AGI	APB
33	40	8	VVSWKNKE	303	37
49	68	20	VTDNVHTRTEQYKFQPESPS	570	28
87	112	26	SVTRLENLMWKQITSELNHILSENEV	395	15
124	158	35	IMQVGKRSLRPQPTELRYSWKTWGKAKMLSTELHN	585	16
162	175	14	LIDGPETAECPNTN	142	10
230	242	13	IESALNDTWKIEK	261	20
244	256	13	SFIEVKSCHWPKS	198	15
263	290	28	GVLESEMVIPKNFAGPKSQHNNRPGYHT	469	16
295	302	8	PWHLGKLE	91	11
361	371	11	MEIRPLKEKEE	445	40

Analysis and comments

Selected Epitopes

#	Start	End	Length	Sequence	IMM	CMP	SE	RESULT
E1	31	47	17	GCVVSWKNKELKCGSGI	Excellent	Good	High	OK
E2	73	82	10	AIQKAHEEGI	Average	Poor	Unclear	
E3	89	97	9	TRLENLMWK	Excellent	Good	High	OK
E4	108	124	17	SENEVKLTIMTGDIKGI	Poor	Average	Low	
E5	138	148	11	ELRYSWKTWGK	Excellent	Good	High	OK
E6	153	163	11	STELHNQTFLI	Average	Good	Low	
E7	165	184	20	GPETAECPTNRAWNSLEVE	Average	Poor	Unclear	
E8	214	222	9	AAIKDNRAV	Average	Good	Low	
E9	233	252	20	ALNDTWKIEKASFIEVKSCH	Excellent	Good	Unclear	
E10	269	285	17	MVIPKNFAGPKSQHNNR	Poor	Average	High	
E11	290	303	14	TQTAGPWHLGKLEM	Poor	Good	High	
E12	327	335	9	RTTTASGKL	Poor	Poor	Low	
E13	363	370	8	IRPLKEKE	Poor	Poor	High	

Potential Cross-Reactivity for antibodies

There are two species cross-reactivity that has to be controlled for these epitopes:

- 1) Homology to other human molecules (since the virus detection will likely be in human cell/tissues)
- 2) Homology to mouse proteins, since a very close homology may affect negatively immune response to the epitope.

Epitope 5.

Homology to human proteins:

Protein Name	Score	Value
ref NP_001186851.1 sialate O-acetyltransferase isoform 2 [Ho...	26	1.3
gb AAG15386.1 AF300796_1 sialic acid-specific 9-O-acetyles...	26	1.3
emb CAB70771.1 hypothetical protein [Homo sapiens]	26	1.3
ref NP_733746.1 sialate O-acetyltransferase isoform 1 precu...	26	1.3
sp Q9HAT2.1 SIAE_HUMAN RecName: Full=Sialate O-acetylster...	26	1.3
gb AAG14897.1 AF303378_1 sialic acid-specific acetyltera...	26	1.3
gb AAH68450.1 Sialic acid acetyltransferase [Homo sapiens]	26	1.3
gb EAW67591.1 sialic acid acetyltransferase [Homo sapiens]	26	1.3
dbj BAG51622.1 unnamed protein product [Homo sapiens]	26	1.3
gb EAW82603.1 hCG1817378 [Homo sapiens]	25	3.0
ref NP_001186982.1 melanocyte protein PMEL isoform 2 prec...	24	6.1

dbj BAH13223.1	unnamed protein product [Homo sapiens]	24	6.1
gb EAW96855.1	silver homolog (mouse), isoform CRA_d [Homo...	24	6.1
gb EAW96852.1	silver homolog (mouse), isoform CRA_a [Homo...	24	6.1
gb AAB19181.1	Pmel 17 [Homo sapiens]	24	6.1
gb AAB31176.1	melanoma antigen gp100 [human, breast cance...	24	6.1
ref NP_008859.1	melanocyte protein PMEL isoform 3 preprop...	24	6.1
sp P40967.2 PMEL_HUMAN	RecName: Full=Melanocyte protein PM...	24	6.1
gb AAC60634.1	gp100 [Homo sapiens]	24	6.1
gb AAB00386.1	melanocyte protein Pmel 17 [Homo sapiens]	24	6.1
gb AAH01414.1	Silver homolog (mouse) [Homo sapiens]	24	6.1
gb AAP35866.1	silver homolog (mouse) [Homo sapiens]	24	6.1
gb EAW96853.1	silver homolog (mouse), isoform CRA_b [Homo...	24	6.1
dbj BAG52619.1	unnamed protein product [Homo sapiens]	24	6.1
gb AAA18479.1	me20m [Homo sapiens]	24	6.1
gb AAA60121.1	Pmel 17 protein [Homo sapiens]	24	6.1
prf 2124412A	melanoma-associated antigen	24	6.1
ref NP_001186983.1	melanocyte protein PMEL isoform 1 prec...	24	6.1
gb EAW96854.1	silver homolog (mouse), isoform CRA_c [Homo...	24	6.1
gb AAA35930.1	melanocyte-specific secreted glycoprotein, ...	24	6.1

Homology to mouse proteins:

Mus musculus	(mouse) [rodents] taxid 10090		
ref NP_068682.2	melanocyte protein PMEL precursor [Mus mu...	27	0.46
dbj BAB28486.1	unnamed protein product [Mus musculus]	27	0.46
gb AAH82555.1	Silver [Mus musculus]	27	0.46
dbj BAE20900.1	unnamed protein product [Mus musculus]	27	0.46
gb EDL24601.1	silver, isoform CRA_a [Mus musculus]	27	0.46
gb EDL24602.1	silver, isoform CRA_a [Mus musculus]	27	0.46
sp Q60696.1 PMEL_MOUSE	RecName: Full=Melanocyte protein PM...	27	0.46
gb AAA69538.1	pmel17 protein [Mus musculus]	27	0.46
gb AAI20736.1	Sfrs17b protein [Mus musculus]	25	1.6
gb AAI31918.1	Sfrs17b protein [Mus musculus]	25	1.6
ref NP_001075425.1	A-kinase anchor protein 17B [Mus muscu...	25	1.6
sp A2A3V1.2 AK17B_MOUSE	RecName: Full=A-kinase anchor prot...	25	1.6
emb CAA85268.1	sialoadhesin [Mus musculus]	25	2.2
gb EDL28299.1	sialic acid binding Ig-like lectin 1, sialo...	25	2.2
emb CAA85290.1	sialoadhesin [Mus musculus]	25	2.2
gb AAB95641.1	sialoadhesin [Mus musculus]	25	2.2
sp Q62230.2 SN_MOUSE	RecName: Full=Sialoadhesin; AltName: ...	25	2.2
gb AAI41337.1	Sialic acid binding Ig-like lectin 1, sialo...	25	2.2
gb AAI41336.1	Sialic acid binding Ig-like lectin 1, sialo...	25	2.2
ref NP_035556.3	sialoadhesin precursor [Mus musculus]	25	2.2
gb EDL28300.1	sialic acid binding Ig-like lectin 1, sialo...	25	2.2
dbj BAC28536.1	unnamed protein product [Mus musculus]	23	7.2
gb AAD55976.1	cytosolic sialic acid 9-O-acetylerase [M...	23	7.3
dbj BAC29164.1	unnamed protein product [Mus musculus]	23	7.3
gb AAH07136.1	Sialic acid acetylerase [Mus musculus]	23	7.3
dbj BAE37942.1	unnamed protein product [Mus musculus]	23	7.3
dbj BAC26026.1	unnamed protein product [Mus musculus]	23	7.3
ref NP_035864.2	sialate O-acetylerase precursor [Mus m...	23	7.3
sp P70665.3 SIAE_MOUSE	RecName: Full=Sialate O-acetyler... ...	23	7.3
gb AAC52880.1	sialic acid-specific 9-O-acetylerase [Mu...	23	7.3
emb CAA67214.1	sialic acid-specific 9-O-acetylerase [M...	23	7.3
dbj BAC26049.1	unnamed protein product [Mus musculus]	23	7.3
dbj BAE39252.1	unnamed protein product [Mus musculus]	23	7.3
gb AAB07813.1	sialic-acid O-acetylerase [Mus musculus]	23	7.3
gb EDL25438.1	sialic acid acetylerase [Mus musculus]	23	7.3

Final Selection:

In the order of preference, E5, E3, E1.
The first choice is epitope E5.

Comment: This analysis demonstrates that use of EpiQuest-B (immunogenicity), EpiQuest-C (complexity) and Surface probability allows to correctly identify immunodominant epitopes in highly immunogenic protein. The immunodominant epitopes for NS1 are known, and correspond to those identified in this analysis.