

PROTEIN SEQUENCE ANALYSIS



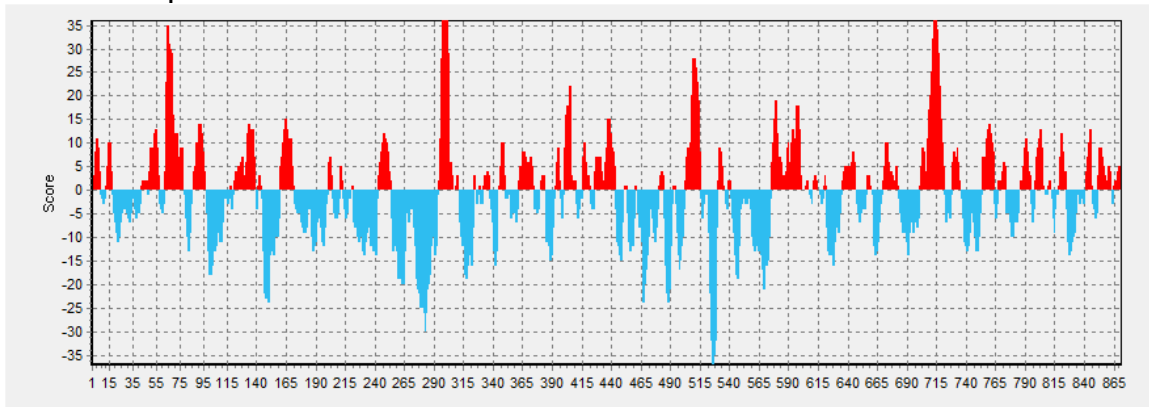
Performed using
EpiQuest™ 2013 v2.18

Order No	TEST
Date	May 02, 2013
Molecule	cadherin 1
Acc. No	GenBank: AAA61489.1
Species of Origin	Xenopus laevis
Intended Species for Immunisation	Rabbit
Sequence (FASTA)	MGLKRPWLLGAVVLLTLIQVQGGLAEWTQCRMGFSKEKYSFLVPKNLETDKALGRVIFNSCEGPVRIQFA SKDPNFEIHKDGTVYIKNPAKMKDNRKTFRVLAWETKGHVYSTNITLKREGHRHRQDLFSGKHSHPKSE TGLKRQKRDWVIPPPIIVSENEKGFPPKRIVQIKSSYAKEVKVYYSITGQGADTPPEGVFAIGREDGWLNV TRPLDREAIIDNYVLFSSHAVSSNGANVEDPMEIIKVVQDQNDNDPVFTQSVFEGSVPEGSKPGTAVMTVSA TDADDSVDMYNGVITYSILNQEPKEPTNKMFTIHSEGLISVLTTGLDREKNPVYTLTIQAADGEFGKDR TTTATALIVVMDTNDNPPVFDPTQYTAKVPENEVGYEVARLTVTDEDIEGTDANNAVYKIIKGNENYFS IQTDGTGNIIGLLKTVKGLDYELKKQYILSVIVTNKANFSVPLQTSTATVTVSVEDVNEAPIFLPPVKEVSV SEDLPSGQVVATYTAQDPDKEQNQKITIVIGNDPAGWVSVNKDNGIVTGNGLDRESKFVNLNNTYKVIIL AADSGSPSATGTGTLVNLNLLDVNDNGPFLEPQQESFCQKDPGFRVFTIIDRDLSPNTYPYKAELTGESNE NWTAIIVTDKILELRPKKELEIGQYDVMITLLDSFGLSNVTKLHITICQCDGDKMQCEEKAAIAGGLGISA IVGILGGILALLLLLLLLLLLFFVRRKKVVKEPLLPPEDETRDNVFSYDEEGGGEEDQDFDLSQLHRGLDAR PDVIRNDVAPVLAAPQYRPRPANPDEIGNFIDENLNAADNDPTAPPYDSSLVFDYEGSGSEAASLSSLNS SNSDLQDYALNDWGPFRFTKLADMYGGDED
Length	871
Region of Interest	Non-specified

Comment: *this is an example of finding epitope in molecule with a number of low complexity regions and many related proteins (other cadherins) with possible cross-reactivity.*

Predicted Immunodominant Epitopes

Whole Sequence overview:



Positions and relative "weight" of Epitopes

Matrix B7.1 SORTED BY ABP*
 Frame 6
 Gap 0
 Threshold 0
 Peptide size 10 and above

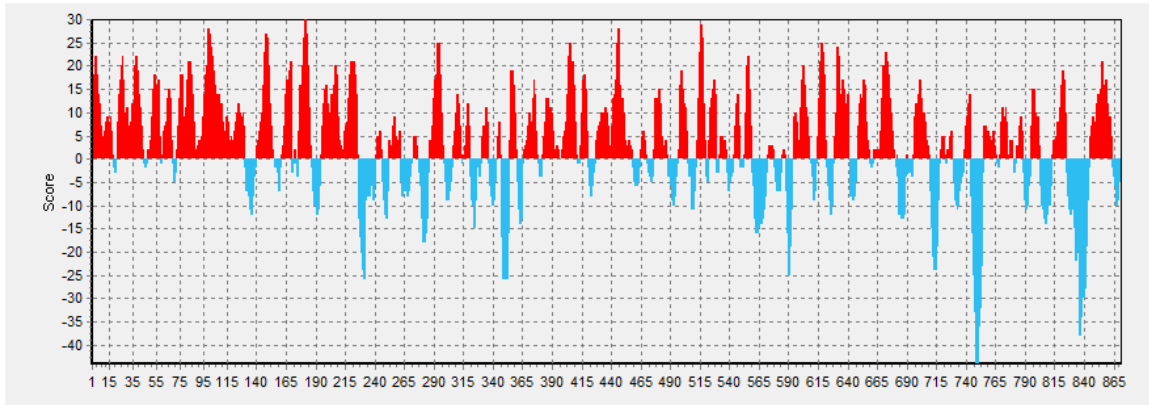
Start	End	Length	Sequence	AGI	APB
700	720	21	AIVGILGGILALLLLLLLLLLL	365	17
62	78	17	EGPVRIQFASKDPNFEI	263	15
502	514	13	TYTAQDPDKEQEQ	200	15
294	310	17	ITYSILNQEPKEPTNKM	247	14
400	409	10	GTDNAWVYK	105	10
159	171	13	ENEKGFPPKRVQ	112	8
753	763	11	EEDQDFDLSQL	94	8
85	96	12	YIKNPAKMKDNR	89	7
243	253	11	IIKVQDQNDND	79	7
426	442	17	GNIGLLKTVKGLDYELK	128	7
575	606	32	LVLNLLDVNDNGPFLEPQQESFCQKDPGFRVF	249	7
121	138	18	GHRHRQDLFSGKSHHPK	118	6
361	373	13	MDTNDNPPVFDPT	72	5
635	646	12	IVTDKILELRPK	55	4
669	682	14	VTKLHITICQCDGD	66	4
785	794	10	PQYRPRPANP	45	4
851	862	12	ALNDWGRFRTKL	53	4

Matrix B7.1 SORTED BY POSITION IN SEQUENCE
 Frame 6
 Gap 0
 Threshold 0
 Peptide size 10 and above

Start	End	Length	Sequence	AGI	APB
62	78	17	EGPVRIQFASKDPNFEI	263	15
85	96	12	YIKNPAKMKDNR	89	7
121	138	18	GHRHRQDLFSGKSHHPK	118	6
159	171	13	ENEKGPFPRIVQ	112	8
243	253	11	IIKVQDQNDND	79	7
294	310	17	ITYSILNQEPKEPTNKM	247	14
361	373	13	MDTNDNPPVFDPT	72	5
400	409	10	GTDRAWNAVYK	105	10
426	442	17	GNIGLLKTVKGLDYELK	128	7
502	514	13	TYTAQDPDKEQNQ	200	15
575	606	32	LVLNLLDVNDNGPFLEPQQESFCQKDPGFRVF	249	7
635	646	12	IVTDKILELRPK	55	4
669	682	14	VTKLHITICQCDGD	66	4
700	720	21	AIVGILGGILALLLLLLLLLLL	365	17
753	763	11	EEDQDFDLSQL	94	8
785	794	10	PQYRPRPANP	45	4
851	862	12	ALNDWGPFRFTKL	53	4

Analysis of Immunologic Complexity

Whole Sequence overview:



Positions of sequences with sufficient complexity

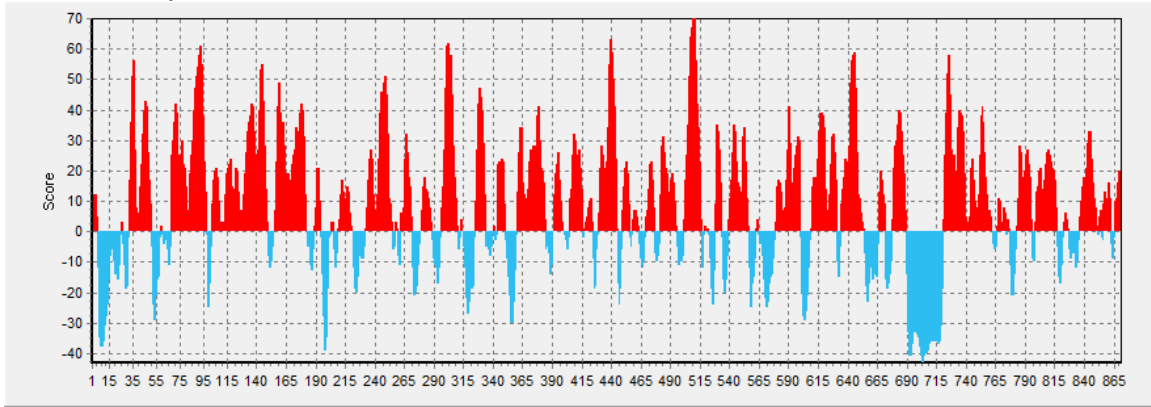
Matrix C1.3
 Frame 6
 Gap 0
 Threshold 0
 Peptide size 10 and above

Start	End	Length	Sequence	AGI	APB
1	18	18	MGLKRPWLLGAVVLLTLI	175	9
21	43	23	QGGLAEWTQCRMGSKEKYSFLV	282	12
47	58	12	LETDKALGRVIF	118	9
72	129	58	KDPNFEIHKDGTVYIKNPAKMKDNRKTFRVLAWETKGHVYSTNITLK REGHRHRQDLF	648	11
139	154	16	SETGLKRQKRDWVIPP	190	11
175	185	11	SYAKEVKVYYS	204	18
194	225	32	PPEGVFAIGREDGWLNVTRPLDREAIDNYVLF	356	11
251	261	11	DNDPVFTQSVF	46	4
286	297	12	SVDMYNGVITYS	164	13
366	377	12	NPPVFDPTQYTA	93	7
381	410	30	ENEVGYEVARLTVTDEDIEGTDNAVYKI	262	8
427	457	31	NIGLLKTVKGLDYELKKQYILSVIVTNKANF	297	9
475	486	12	VNEAPIFLPPVK	92	7
593	607	15	QESFCQKDPGFRVFT	131	8
628	641	14	SNENWTAIVTDKIL	186	13
647	657	11	KELEIGQYDVM	106	9
661	679	19	LDSFGLSNVTKLHITICQC	183	9
695	708	14	GLGISAIVGILGGI	117	8
755	765	11	DQDFDLSQLHR	55	5
768	779	12	DARPDVIRNDVA	66	5

812	824	13	PTAPPYDSLIVFD	109	8
844	862	19	DLDQDYSALNDWGPFRFTKL	223	11

Predicted probability of surface exposure

Whole Sequence overview:



Positions of best sequences

Host EM3.2
 Frame 6
 Gap 0
 Threshold 0
 Peptide size 7 and above

Start	End	Length	Sequence	AGI	APB
32	50	19	MGFSKEKYSFLVPKNLETD	527	27
67	96	30	IQFASKDPNFEIHKDGTVYIKNPAKMKDNR	941	31
102	148	47	LAWETKGHVYSTNITLKREGHRHRQDLFSGKHSHPKSETGLKRQKR	959	20
154	182	29	PIIVSENEKGFPPKRIVQIKSSYAKEVKV	766	26
208	220	13	LNVTRPLDREAID	108	8
232	254	23	NGANVEDPMEIIKVDQNDNDP	518	22
262	270	9	EGSVPEGSK	137	15
278	287	10	VSATDADDSV	105	10
296	309	14	YSILNQEPKEPTNK	496	35
325	332	8	TGLDREKN	251	31
343	349	7	DGEFGKD	135	19
360	383	24	VMDTNDNPPVFDPTQYTAKVPENE	549	22
390	399	10	RLTVTDEDIE	133	13
404	415	12	WNAVYKIIKGNE	243	20
417	423	7	NYFSIQT	41	5
429	444	16	GLLKTVKGLDYELKKQ	490	30
469	476	8	TVSVEDVN	97	12
481	494	14	FLPPVKEVSVSEDL	241	17
501	515	15	ATYTAQDPDKEQNQK	628	41
539	555	17	GNGNLDRESKFVLNNTY	311	18

578	599	22	NLLDVNDNGPFLEPQQESFCQK	429	19
608	630	23	IIDRDLSPTYPYKAELTGESNE	480	20
634	653	20	AIVTDKILELRPKKELEIGQ	479	23
677	688	12	CQCDGDKMQCEE	311	25
720	761	42	LFVRRKKVVKEPLLPEDETRDNVFSYDEEGGGEEDQDFDLS	934	22
766	772	7	GLDARPD	43	6
782	795	14	LAAPQYRPRPANPD	253	18
798	816	19	GNFIDENLNAADNDPTAPP	332	17
835	850	16	LSSLNSSNSDLQDYS	273	17
865	871	7	MYGGDED	83	11

Analysis and comments

Selected Epitopes

#	Start	End	Length	Sequence	IMM	CMP	SE	RESULT
E1	62	78	17	EGPVRIQFASKDPNFEI	Good	Average	Good	GO
E2	85	96	12	YIKNPAKMKDNR	Poor	Good	Good	
E3	121	138	18	GHRHRQDLFSGKSHSHPK	Poor	Average	Good	
E4	159	171	13	ENEKGPFPRIVQ	Average	Average	Good	GO
E5	243	253	11	IIKVQDQNDND	Poor	Poor	Good	
E6	294	310	17	ITYSILNQEPKEPTNKM	Average	Good	Good	GO
E7	361	373	13	MDTNDNPPVFDPT	Poor	Average	Excellent	
E8	400	409	10	GTDAWNAVYK	Average	Good	Excellent	GO
E9	426	442	17	GNIGLLKTVKGLDYELK	Poor	Good	Poor	
E10	502	514	13	TYTAQDPDKEQNG	Good	Poor	Excellent	
E11	575	606	32	LVLNLLDVNDNGPFLEPQQESFCQKDPGFRVF	Poor	Good*	Good	
E12	635	646	12	IVTDKILELRPK	Poor	Average	Good	
E13	669	682	14	VTKLHITICQCDGD	Poor	Average	Poor	
E14	700	720	21	AIVGILGGILALLLLLLLLLLL	Good	Negative	Poor	
E15	753	763	11	EEDQDFDLSQL	Average	Average	Good	GO
E16	785	794	10	PQYRPRPANP	Poor	Poor	Good	
E17	851	862	12	ALNDWGPFRFTKL	Poor	Good	Poor	

* 593-606

Potential Cross-Reactivity for antibodies

Epitope E1

Xenopus laevis (platanna, ...) [frogs & toads] taxid 8355		
gb AAA61489.1 cadherin [Xenopus laevis]	58	2e-11
prf 2019237A cadherin	58	2e-11
ref NP_001165703.1 cadherin-1 precursor [Xenopus laevis]	58	2e-11

sp P30944.2 CADH1_XENLA RecName: Full=Cadherin-1; AltName:...	58	2e-11
gb AAA93116.1 E-cadherin [Xenopus laevis]	58	2e-11
gb AAH43888.1 LOC398471 protein, partial [Xenopus laevis]	22	11
ref NP_001082440.1 uncharacterized protein LOC398471 [Xen...	22	11
gb AAH88719.1 LOC398471 protein [Xenopus laevis]	22	11
ref NP_001129638.1 patched 2 [Xenopus laevis]	22	16
gb AAH70995.1 Unknown (protein for MGC:79874) [Xenopus la...	22	16

No risk of x-reactivity

Epitope E4

Xenopus laevis (platanna, ...) [frogs & toads] taxid 8355		
emb CAA53206.1 E-cadherin [Xenopus laevis]	45	2e-07
gb AAA61489.1 cadherin [Xenopus laevis]	45	2e-07
prf 2019237A cadherin	45	2e-07
ref NP_001165703.1 cadherin-1 precursor [Xenopus laevis]	45	2e-07
sp P30944.2 CADH1_XENLA RecName: Full=Cadherin-1; AltName:...	45	2e-07
gb AAA93116.1 E-cadherin [Xenopus laevis]	45	2e-07
pdb 1L3W A Chain A, C-Cadherin Ectodomain	38	4e-05
gb AAC16910.1 C-cadherin [Xenopus laevis]	38	4e-05
sp P33148.2 CADHF_XENLA RecName: Full=EP-cadherin; AltName:...	38	4e-05
sp P33152.3 CADHB_XENLA RecName: Full=Blastomere cadherin;...	38	4e-05
gb AAH68940.1 Xb-cad protein [Xenopus laevis]	38	4e-05
ref NP_001080946.1 EP-cadherin precursor [Xenopus laevis]	38	4e-05
emb CAA45252.1 EP-cadherin [Xenopus laevis]	38	4e-05
ref NP_001089045.1 blastomere cadherin [Xenopus laevis]	38	4e-05
emb CAA55292.1 XB-cadherin [Xenopus laevis]	38	4e-05
gb AAI68523.1 Unknown (protein for IMAGE:5570937) [Xenopu...	23	2.5
ref NP_001080560.1 ciliary neurotrophic factor receptor p...	21	12
gb AAH43961.1 Cntfr-prov protein [Xenopus laevis]	21	12
sp P79883.1 CADH4_XENLA RecName: Full=Cadherin-4; AltName:...	21	12
pdb 3K6D A Chain A, Crystal Structure Of Xenopus Laevis T-...	21	17
ref NP_001088960.1 PDZ binding kinase [Xenopus laevis]	21	17
gb AAH88936.1 LOC496340 protein [Xenopus laevis]	21	17
gb AAI30125.1 LOC100037019 protein [Xenopus laevis]	21	17
gb AAI55969.1 LOC100127343 protein [Xenopus laevis]	21	17

Potential cross-reactivity with other cadherins

Epitope E6

Xenopus laevis (platanna, ...) [frogs & toads] taxid 8355		
emb CAA53206.1 E-cadherin [Xenopus laevis]	60	5e-12
gb AAA61489.1 cadherin [Xenopus laevis]	60	5e-12
prf 2019237A cadherin	60	5e-12
ref NP_001165703.1 cadherin-1 precursor [Xenopus laevis]	60	5e-12
sp P30944.2 CADH1_XENLA RecName: Full=Cadherin-1; AltName:...	60	5e-12
gb AAA93116.1 E-cadherin [Xenopus laevis]	60	5e-12
ref NP_001165703.1 cadherin-1 precursor [Xenopus laevis]	14	11454
sp P30944.2 CADH1_XENLA RecName: Full=Cadherin-1; AltName:...	14	11454
gb AAA93116.1 E-cadherin [Xenopus laevis]	14	11454
sp P33152.3 CADHB_XENLA RecName: Full=Blastomere cadherin;...	31	0.012
gb AAH68940.1 Xb-cad protein [Xenopus laevis]	31	0.012
ref NP_001089045.1 blastomere cadherin [Xenopus laevis]	31	0.012
emb CAA55292.1 XB-cadherin [Xenopus laevis]	31	0.012
pdb 1L3W A Chain A, C-Cadherin Ectodomain	28	0.11
gb AAC16910.1 C-cadherin [Xenopus laevis]	28	0.11
sp P33148.2 CADHF_XENLA RecName: Full=EP-cadherin; AltName:...	28	0.11
ref NP_001080946.1 EP-cadherin precursor [Xenopus laevis]	28	0.11
emb CAA45252.1 EP-cadherin [Xenopus laevis]	28	0.11
ref NP_001087430.1 phosphatidylinositol glycan anchor bio...	24	2.0

Potential cross-reactivity with other cadherins

Epitope 8

Accession	Protein Name	Score	Value
Xenopus laevis (platanna, ...) [frogs & toads] taxid 8355			
emb CAA53206.1 	E-cadherin [Xenopus laevis]	36	2e-04
gb AAA61489.1 	cadherin [Xenopus laevis]	36	2e-04
prf 2019237A	cadherin	36	2e-04
ref NP_001165703.1 	cadherin-1 precursor [Xenopus laevis]	36	2e-04
sp P30944.2 CADH1 XENLA	RecName: Full=Cadherin-1; AltName:...	36	2e-04
gb AAA93116.1 	E-cadherin [Xenopus laevis]	36	2e-04
sp P33152.3 CADHB XENLA	RecName: Full=Blastomere cadherin;...	26	0.22
gb AAH68940.1 	Xb-cad protein [Xenopus laevis]	26	0.22
ref NP_001089045.1 	blastomere cadherin [Xenopus laevis]	26	0.22
emb CAA55292.1 	XB-cadherin [Xenopus laevis]	26	0.22
pdb 1L3W A	Chain A, C-Cadherin Ectodomain	24	1.0
gb AAC16910.1 	C-cadherin [Xenopus laevis]	24	1.0
sp P33148.2 CADHF XENLA	RecName: Full=EP-cadherin; AltName:...	24	1.0
ref NP_001080946.1 	EP-cadherin precursor [Xenopus laevis]	24	1.0
emb CAA45252.1 	EP-cadherin [Xenopus laevis]	24	1.0
sp P20310.1 CADHN XENLA	RecName: Full=Neural cadherin-1; S...	24	1.4
ref NP_001165707.1 	neural cadherin-2 precursor [Xenopus l...	24	1.4
sp P33147.1 CADHO XENLA	RecName: Full=Neural cadherin-2; S...	24	1.4
emb CAA40867.1 	N-cadherin [Xenopus laevis]	24	1.4
ref NP_001086490.1 	transmembrane protein 53-B [Xenopus la...	22	4.7
sp Q6DJC8.1 TM53B XENLA	RecName: Full=Transmembrane protei...	22	4.7
gb AAH75254.1 	MGC84808 protein [Xenopus laevis]	22	4.7
ref NP_001088199.1 	cadherin 13, H-cadherin (heart) precu...	22	6.5
gb AAH84115.1 	LOC495024 protein [Xenopus laevis]	22	6.5

Final Selection:

Epitope 1 is the best choice as 1) relatively strong, 2) likely exposed at the surface of native molecule, 3) antibodies raised against it will likely be specific to Cadherin 1 only without cross-reactivity to other