

Table S3

I: RQWGPD{A/C/D/F/H/I/K/L/M/N/P/R/S/V/Y}
{A/C/G/H/I/K/L/M/N/P/Q/R/S/T/V}A motif

Number	Peptide Sequence	pEC ₅₀
1	RQWGPDNCA	5.484
2	RQWGPDCCA	5.651
3	RQWGPDPCA	6.047
4	RQWGPDPSMA	6.319
5	RQWGPDPCIA	6.326
6	RQWGPDNDNA	6.777
Index	ALWGPDPA	7.151
7	RQWGPDPRVA	7.236
8	RQWGPDPIHA	7.24
9	RQWGPDPKRA	7.257
10	RQWGPDPRTA	7.473
11	RQWGPDPKIA	7.612
12	RQWGPDPCYA	7.659
13	RQWGPDPFHA	7.791
14	RQWGPDPLKA	7.854
15	RQWGPDPAIA	7.882
16	RQWGPDPCYA	7.884
17	RQWGPDPKCA	7.887
18	RQWGPDPRTA	7.907
19	RQWGPDPCYA	7.912
20	RQWGPDPKTA	7.916
21	RQWGPDPCSA	7.969
22	RQWGPDPCMA	7.979
23	RQWGPDPCHA	8.093
24	RQWGPDPLAA	8.096
25	RQWGPDPCIA	8.36
26	RQWGPDPCRA	8.384
27	RQWGPDPCLA	8.823
28	RQWGPDPCNA	9.328
29	RQWGPDPCPLA	9.335
30	RQWGPDPCPA	9.395

II: RQWGPD{D/F}{P/F}xx{A/I/L/V} motif

Number	Peptide Sequence	pEC ₅₀
1	RQWGPDPSYV	3.723
2	RQWGPDFFMDA	5.401
3	RQWGPDFFIRI	5.561
4	RQWGPDFFWDL	5.564
5	RQWGPDFFRGA	5.658
6	RQWGPDFFGEI	5.949
7	RQWGPDFFPWV	6.048
8	RQWGPDFFRTA	6.102
9	RQWGPDFFPHFV	6.121
10	RQWGPDFFSQL	6.147
11	RQWGPDFFEGV	6.306
12	RQWGPDFFDDI	6.391
13	RQWGPDFFKNV	6.795
14	RQWGPDFFWGI	6.949
15	RQWGPDFFKSA	7.114
16	RQWGPDFFFYV	7.28
17	RQWGPDFFPAYV	7.404
18	RQWGPDFFHTA	7.498
19	RQWGPDPGTI	7.51
20	RQWGPDFFSFV	8.05
21	RQWGPDFFVNI	8.206
22	RQWGPDFFAHV	8.625
Index	ALWGPDPA	8.639
23	RQWGPDFFESI	8.716
24	RQWGPDFFHEI	9.086
25	RQWGPDFFVTV	9.253
26	RQWGPDPCMA	9.556
27	RQWGPDPTNV	9.622
28	RQWGPDPHSV	9.774
29	RQWGPDPMQL	9.802
30	RQWGPDFFMTV	9.808

III: RQxGPDxxx motif

Number	Peptide Sequence	pEC ₅₀
1	RQVGPDIVVA	Null
2	RQPGPDQINA	2.538
3	RQEGPDYFEA	2.567
4	RQPGPDNNMA	2.891
5	RQAGPDYQA	2.901
6	RQPGPDQRLA	2.92
7	RQTGPDLDYA	2.926
8	RQEGPDSEVA	2.996
9	RQDGPDRDA	3.099
10	RQYGPDVWQA	3.12
11	RQKGPDAEGA	3.479
12	RQIGPDDQLA	3.536
13	RQEGPDMDPA	3.84
14	RQHGPDDQEA	3.894
15	RQPGPDMSA	4.302
16	RQLGPDSSHA	4.409
17	RQRGPDQMAA	4.502
18	RQSGPDLLA	4.548
19	RQSGPDRIA	4.756
20	RQDGPDPPTA	4.771
21	RQDGPDPQRA	4.784
22	RQPGPDNNYA	4.922
23	RQPGPDLSA	5.076
24	RQSGPDPHRA	5.694
25	RQMGPDWSEA	5.701
26	RQNGPDQMSA	5.924
27	RQMGPDPEPA	5.944
Index	ALWGPDPA	6.49
28	RQSGPDHATA	6.932
29	RQMGPDFFLA	7.642
30	RQPGPDFFPA	8.12

Table S3 cont

IV: xQxGPDxxxV motif

Number	Peptide Sequence	pEC ₅₀
1	DQFGPDRDPV	2.731
2	QQDGPDPNLV	2.991
3	VQYGPDESTV	3.064
4	AQNGPDFEEV	3.109
5	SQLGPDQSEV	3.226
6	YQYGPDEDEV	3.261
7	SQPGPDRMIV	3.29
8	TQDGPDSAHV	3.294
9	GQGGPDNLNV	3.308
10	YQEGPDASPV	3.32
11	IQPGPDILFV	3.342
12	YQKGPDLQYV	3.37
13	PQKGPDMLEV	3.413
14	KQLGPDTEIV	3.417
15	EQWGPDQWNV	3.437
16	DQSGPDRAYV	3.45
17	MQMGPDITLV	3.457
18	AQLGPDVAWV	3.464
19	KQPGPDQAGV	3.477
20	GQGGPDNLNV	3.556
21	LQSGPDVNIV	3.577
22	AQEGPDWDKV	3.949
23	IQTGPDKIAV	3.975
24	QQHGPDSQYV	4.147
25	EQWGPDQKWV	4.166
26	SQEGPDFPGV	4.792
27	FQIGPDFWVV	5.654
28	RQVGPDHTLV	5.999
29	DQIGPDWPRV	6.446
30	AQSGPDWLKV	6.512
Index	ALWGPDPAAA	7.976

Va: Biased sampling set (1st)

Number	Peptide Sequence	pEC ₅₀
1	GVYGPFMKGM	5.741
2	MQILPDFFKL	6.293
3	IQLGPDWISA	6.337
4	YVVGPFVKGL	6.487
5	VQQGPFAMNL	6.771
6	KLWGPDFRNA	7.095
7	SVVGPSWKQA	7.323
8	LQFGPFFMNM	7.343
9	KQKLPPFFTA	7.395
10	WLYGPFFMPV	7.523
11	HVQGWFFLT	7.617
12	WQEGPDWPFT	8.197
13	LTYGPFVKLL	8.416
14	MVNGPDWKTM	8.442
15	RQFGPFWPHL	8.582
16	RSFGVDFPVI	8.586
Index	ALWGPDPAAA	8.593
17	KQIGPFWKGI	8.601
18	KQYGPFFLIA	8.726
19	YQWGPFFLPA	8.78
20	RQWGPDFLRI	8.807
21	GLFGPDEKLI	9.043
22	KTMGPFLMRI	9.12
23	WQWGPFFLKV	9.158
24	WQYWPDFMKA	9.305
25	YQGPFAFPTM	9.461
26	KLQGPDFKSI	9.521
27	LQLGPDFKTA	9.749
28	KQFGPDFPTA	10.346
29	RQFGPDFPKL	10.467
30	YQFGPDFPIA	10.749

Vb: Biased sampling set (2nd)

Number	Peptide Sequence	pEC ₅₀
1	IQYGPFMFGA	3.254
2	MQFGPFFGPA	3.265
3	AIFGPFFIKL	4.093
4	WQWGGDWKFL	4.114
5	TQKGPFWIQA	4.279
6	IQFGPFFLPA	4.703
7	YLAGPAFKRA	4.81
8	LQWGPFFTIL	4.91
9	ITWGNDFKVA	4.939
10	FQYGPFLPRT	5.242
11	MQFGPFWPGA	5.346
12	LQEGPDWPKA	5.353
13	LTWGPDFWFA	5.365
14	WQQGPDRMNI	5.44
15	HQFGPDRMTL	5.553
16	HIWGPDWKTI	5.679
17	RVLGPDFVKV	5.931
18	AQYGPFWPTI	6.063
19	IIWGNDFMVL	6.618
20	LVYGPFFKII	6.646
21	RSWLPDFLTI	6.86
22	YLFGPDLMKV	7.186
Index	ALWGPDPAAA	7.381
23	LQVGPDFMKA	7.615
24	MQVGPDFMNI	7.716
25	MVFGPDFKSA	7.895
26	KQWGPDFPLA	8.203
27	RQYGPFFAST	8.564
28	YQYGPDFINA	8.754
29	MQYGPDFPRT	8.892
30	RQFGPDFPTI	10.357