Supplemental Data for

A single autoimmune T-cell receptor recognizes over a million different peptides

Linda Wooldridge[#], Julia Ekeruche-Makinde[#], Hugo A. van den Berg[#], Anna Skowera, John J. Miles, Mai Ping Tan, Garry Dolton, Mathew Clement, Sian Llewellyn-Lacey, David A. Price, Mark Peakman and Andrew K. Sewell

File prepared by LW. E-mail: wooldridgel@cardiff.ac.uk

[#]These authors contributed equally to this study.

SUPPLEMENTAL EQUATIONS

Estimation of $p\mathrm{EC}_{50}$

Functional sensitivity is expressed by the $p\mathrm{EC}_{50}$ of that peptide with respect to the TCR. This is defined as minus 1 times the base-10 logarithm of the 50% efficacy concentration. Accordingly, The read-out y (MIP1β by ELISA) is related to the incubation concentration C, as follows:

$$y = y_{\min} + \frac{y_{\max} - y_{\min}}{1 + 10^{\kappa(p\mathrm{EC}_{50} + \log C)}} \qquad \text{(S1)}$$

where $y_{\min}$, $y_{\max}$, and κ are parameters that were estimated using non-linear least squares. Parsimony was achieved through simultaneous fitting; i.e. $y_{\min}$, $y_{\max}$, and κ were assumed to have the same value for all peptides (κ is the steepness of the dose-response curve). Eqn (S1) is derived by assuming, first, that each pMHC molecule contributes a signal w_{ij} where i denotes the TCR clonotype and j the pMHC ligand, so that the combined signal generated by Z_j copies of the ligand present in the contact area between the $CD8^+$ T-cell and the C1R-A2 B-cell is given by $Z_j w_{ij}$; the quantity w_{ij} represents the functional sensitivity. Second, it is assumed that the $CD8^+$ T-cell is activated when the signal $Z_j w_{ij}$ exceeds a signalling threshold W_{act} (1), which is assumed to follow a log-logistic distribution. Third, it is assumed that Z_j is proportional to the pulsing concentration (the proportionality is unknown but is eliminated by studying the ratio relative to the EC_{50} of the index peptide; i.e. the difference in $p\mathrm{EC}_{50}$). Then, the EC_{50} is inversely proportional to w_{ij} and eqn (S1)

follows by assuming that the fraction of responding $CD8^+$ T-cells is proportional to the response above baseline ($y_{\min}$).

Theoretical curve

For comparison, the curve derived from TCR activation theory (1,2) is also shown (grey dashed line in Figure 4), based on the formula for the number of peptides with a relative functional sensitivity that is at least as strong as ω:

$$\mathcal{N}[p\mathrm{EC}_{50} - p\mathrm{EC}_{50}^{\mathrm{index}} > \omega] = \frac{N_0}{2}\left(\mathrm{erf}\left[\frac{\beta}{\alpha\sqrt{2}}\left(1+\alpha-(\omega-\gamma)\log 10 + \mathcal{W}_0\left(-e^{-1+(\omega-\gamma)\log 10}\right)\right)\right] - \mathrm{erf}\left[\frac{\beta}{\alpha\sqrt{2}}\left(1+\alpha-(\omega-\gamma)\log 10 + \mathcal{W}_{-1}\left(-e^{-1+(\omega-\gamma)\log 10}\right)\right)\right]\right) \quad \text{(S2)}$$

where erf denotes the error function, $\mathcal{W}$ denotes the Lambert W-function, N_0 is the number of MHC-anchorable peptides, α is a location parameter, β is a specificity parameter, and γ is an offset to account for the fact that the index peptide, rather than the absolutely optimal peptide, is used as reference.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1: Position degeneracy of the 1E6 TCR. MIP1β activation data for a set of peptides with the sequence ALWGPDPxAA, where x is 1 of the 20 natural proteogenic L-amino acids. 6×10^4 C1R-A2 B-cells were pulsed with peptides at various concentrations. After 2 hours, 3×10^4 1E6 $CD8^+$ T-cells were added and incubated overnight. Supernatant was harvested and assayed for MIP1β.

Figure S2: Recognition of peptides sampled at random from large fixed motif sets. A&B. The response of 1E6 $CD8^+$ T-cells to 30 peptides sampled at random from a large motif-restricted set:

RQWGPDP{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;

total set size = 225). Assays as in Figure S1. **A.** Selected titration curves are shown to demonstrate the range of functional sensitivities observed within the set of 30 peptides. Standard deviation from the mean of two replicates is shown. **B.** The functional sensitivities of all peptides tested are displayed relative to that of the index peptide ($p\mathrm{EC}_{50} - p\mathrm{EC}_{50}$ index) to control for variations in absolute values between assays. Peptides with a functional sensitivity equivalent to index have a value 0, peptides with a functional sensitivity greater than index have values >0 and peptides that are less immunogenic than index have values <0. Coloured bars match the key shown in panel A for individual peptides in the set. **C&D.** Details as for A&B, except that the motif is RQWGP{D/F}{P/F}xx{A/I/L/V} where x denotes any one of the 19 amino acids excluding cysteine and the total set size = 5776.

Figure S3: Recognition of 30 peptides sampled at random from large peptide sets (motif: RQxGPDxxxA; total set size = 19^4 or xQxGPDxxxV; total set size = 19^5). Assays as in Figure S1.The functional sensitivities of all peptides tested are displayed relative to that of the index peptide (pEC$_{50}$ – pEC$_{50}$ index).

Figure S4: Recognition of peptides selected by CPL-based importance sampling. A&B. 1E6 CD8$^+$ T-cell recognition of two sets of 30 peptides sampled from a CPL-based importance sampling set with effective size = 1.66×10^8 (calculated from the sampling entropy). Assays as described in Figure S1 legend.

Figure S5: *Mathematica* code. Peptides were sampled at random from larger motif-restricted or CPL-based importance sampling sets varying in total size from 225 to 1.66×10^8 individual peptides using Mathematica® (Wolfram Research Europe Ltd., Long Hanborough, UK). Displayed are the codes to generate the peptide samples as well as an example of the simultaneous non-linear curve fitting procedure. The workspace file is available upon request.

Figure S6: The functional sensitivity of 1E6 CD8$^+$ T-cells to all peptide ligands tested.

Simultaneous curve fitting (as described in supplementary equations; eqn S1) was used to estimate functional sensitivity measured as pEC$_{50}$ for peptides sampled from: **A:** RQWGPDP{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 (set size 225; 30 peptides sampled at random); **B:** RQWGP{D/F}{P/F}xx{A/I/L/V} (set size 5776; 30 peptides sampled at random); **C:**

RQxGPDxxxA (set size 19^4; 30 peptides sampled at random); **D:** xQxGPDxxxV (set size 19^5; 30 peptides sampled at random) and **E&F:** two replicates of a biased sampling set (effective set size $1.66x10^8$, calculated from the sampling entropy); each set of 30 peptides was sampled with bias towards strong agonists, weighted based on the primary CPL scan. Values for each peptide are displayed in Table S3.

SUPPLEMENTAL TABLE LEGENDS

Table S1: Data matrix used for derivation of biased sampling set. First, raw data values generated from the decamer CPL scan (MIP1β levels in pg/ml) were inserted into the table. The following modifications were then made, based on results obtained with previous peptide screening of the 1E6 clone (data not shown): cysteine set to zero; small or negative values set to 5; position 2: double weight Q; position 3: assign 400 to A, E, K and N, turn V up to 400; position 4: tune L and W down to 50; position 5: tune all responses down to 50 except for P; position 6: tune all responses down to 50 except R &M; position 7: leave as original screen data; position 8: 500 for A, 200 for F, R and V; position 9: 200 for F; position 10: tune V down to 1000, add value for index peptide for A and increase weight of L to 2000.

Table S2: Normalized matrix used for derivation of biased sampling set. All MIP1β readings from Table S1 were given a value between 0-1.

Table S3: pEC_{50} values for all peptide ligands tested. Simultaneous curve fitting (as described in supplementary equations) was used to estimate functional sensitivity measured as pEC_{50} for every peptide tested (see Figure S6).

SUPPLEMENTAL REFERENCES

1. van den Berg, H. A., Rand, D. A., and Burroughs, N. J. (2001) *J Theor Biol* **209**, 465-486
2. van den Berg, H. A., and Rand, D. A. (2007) *Immunol Rev* **216**, 81-92

Figure S1

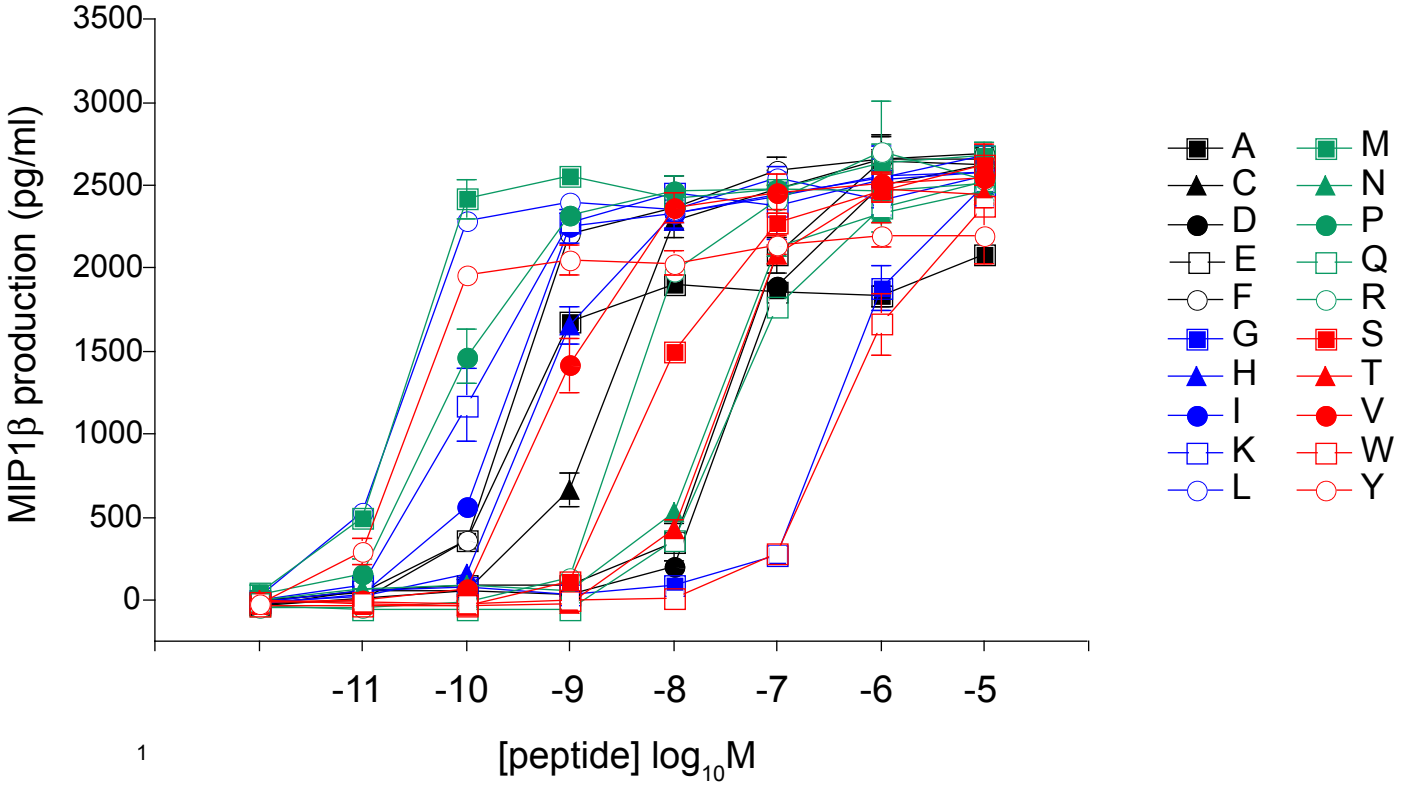
MIP1β production (pg/ml)
3500
3000
2500
2000
1500
1000
500
0
-11
-10
-9
-8
-7
-6
-5
[peptide] log10M
A
C
D
E
F
G
H
I
K
L
M
N
P
Q
R
S
T
V
W
Y

Figure S2

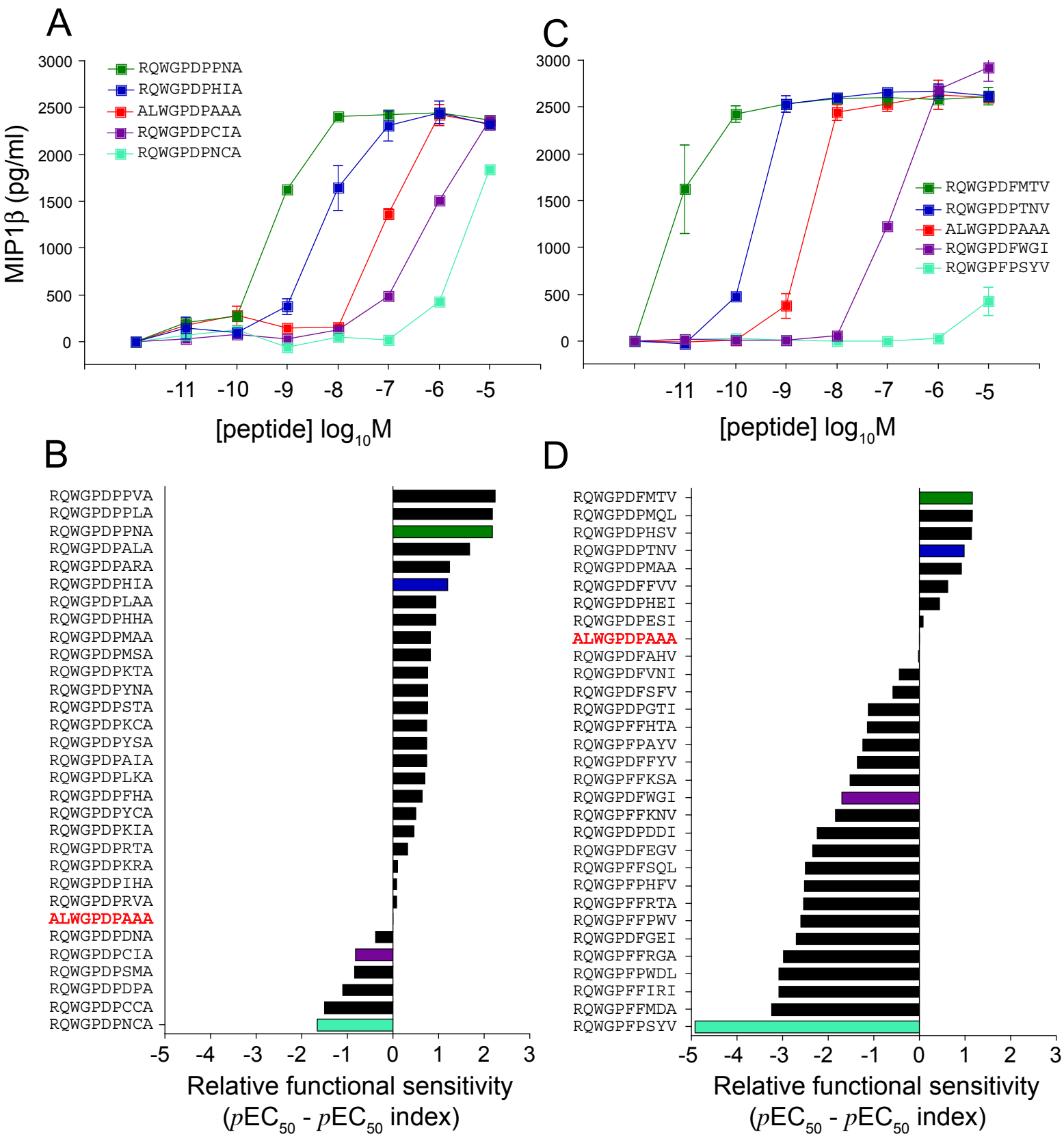
A
C
MIP1β (pg/ml)
RQWGPDPPNA
RQWGPDPHIA
ALWGPDPAAA
RQWGPDPCIA
RQWGPDPNCA
RQWGPDFMTV
RQWGPDPTNV
ALWGPDPAAA
RQWGPDFWGI
RQWGPFPSYV
[peptide] log10M
B
D
RQWGPDPPVA
RQWGPDPPLA
RQWGPDPPNA
RQWGPDPALA
RQWGPDPARA
RQWGPDPHIA
RQWGPDPLAA
RQWGPDPHHA
RQWGPDPMAA
RQWGPDPMSA
RQWGPDPKTA
RQWGPDPYNA
RQWGPDPSTA
RQWGPDPKCA
RQWGPDPYSA
RQWGPDPAIA
RQWGPDPLKA
RQWGPDPFHA
RQWGPDPYCA
RQWGPDPKIA
RQWGPDPRTA
RQWGPDPKRA
RQWGPDPIHA
RQWGPDPRVA
ALWGPDPAAA
RQWGPDPDNA
RQWGPDPCIA
RQWGPDPSMA
RQWGPDPDPA
RQWGPDPCCA
RQWGPDPNCA
RQWGPDFMTV
RQWGPDPMQL
RQWGPDPHSV
RQWGPDPTNV
RQWGPDPMAA
RQWGPDFFVV
RQWGPDPHEI
RQWGPDPESI
ALWGPDPAAA
RQWGPDFAHV
RQWGPDFVNI
RQWGPDFSFV
RQWGPDPGTI
RQWGPFFHTA
RQWGPFPAYV
RQWGPDFFYV
RQWGPFFKSA
RQWGPDFWGI
RQWGPFFKNV
RQWGPDPDDI
RQWGPDFEGV
RQWGPFFSQL
RQWGPFPHFV
RQWGPFFRTA
RQWGPFFPWV
RQWGPDFGEI
RQWGPFFRGA
RQWGPFPWDL
RQWGPFFIRI
RQWGPFFMDA
RQWGPFPSYV
Relative functional sensitivity
(pEC50 - pEC50 index)

Figure S3

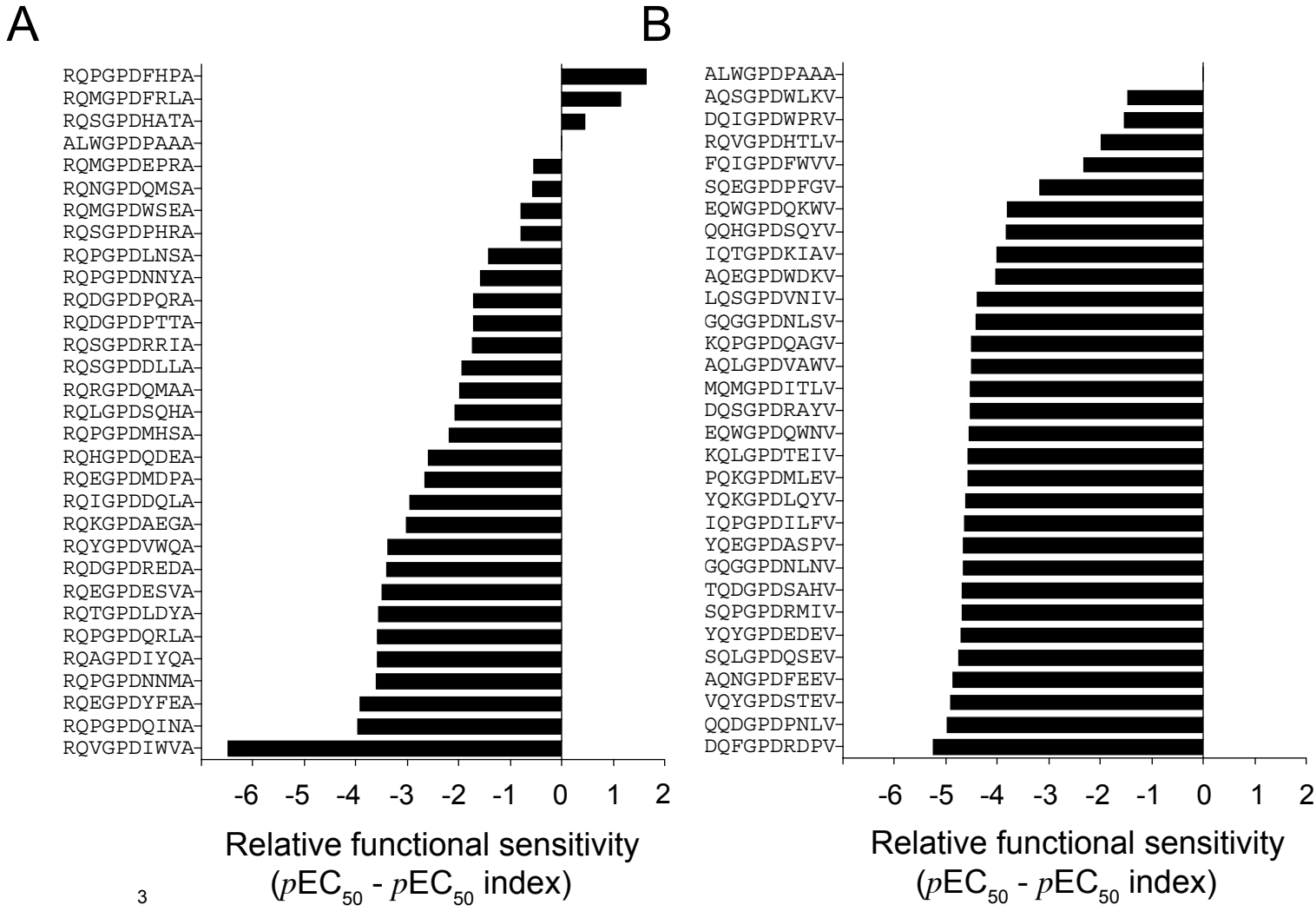
A
RQPGPDFHPA
RQMGPDFRLA
RQSGPDHATA
ALWGPDPAAA
RQMGPDEPRA
RQNGPDQMSA
RQMGPDWSEA
RQSGPDPHRA
RQPGPDLNSA
RQPGPDNNYA
RQDGPDPQRA
RQDGPDPTTA
RQSGPDRRIA
RQSGPDDLLA
RQRGPDQMAA
RQLGPDSQHA
RQPGPDMHSA
RQHGPDQDEA
RQEGPDMDPA
RQIGPDDQLA
RQKGPDAEGA
RQYGPDVWQA
RQDGPDREDA
RQEGPDESVA
RQTGPDLDYA
RQPGPDQRLA
RQAGPDIYQA
RQPGPDNNMA
RQEGPDYFEA
RQPGPDQINA
RQVGPDIWVA
-6 -5 -4 -3 -2 -1 0 1 2
Relative functional sensitivity
(pEC$_{50}$ - pEC$_{50}$ index)
B
ALWGPDPAAA
AQSGPDWLKV
DQIGPDWPRV
RQVGPDHTLV
FQIGPDFWVV
SQEGPDPFGV
EQWGPDQKWV
QQHGPDSQYV
IQTGPDKIAV
AQEGPDWDKV
LQSGPDVNIV
GQGGPDNLSV
KQPGPDQAGV
AQLGPDVAWV
MQMGPDITLV
DQSGPDRAYV
EQWGPDQWNV
KQLGPDTEIV
PQKGPDMLEV
YQKGPDLQYV
IQPGPDILFV
YQEGPDASPV
GQGGPDNLNV
TQDGPDSAHV
SQPGPDRMIV
YQYGPDEDEV
SQLGPDQSEV
AQNGPDFEEV
VQYGPDSTEV
QQDGPDPNLV
DQFGPDRDPV
-6 -5 -4 -3 -2 -1 0 1 2
Relative functional sensitivity
(pEC$_{50}$ - pEC$_{50}$ index)

Figure S4

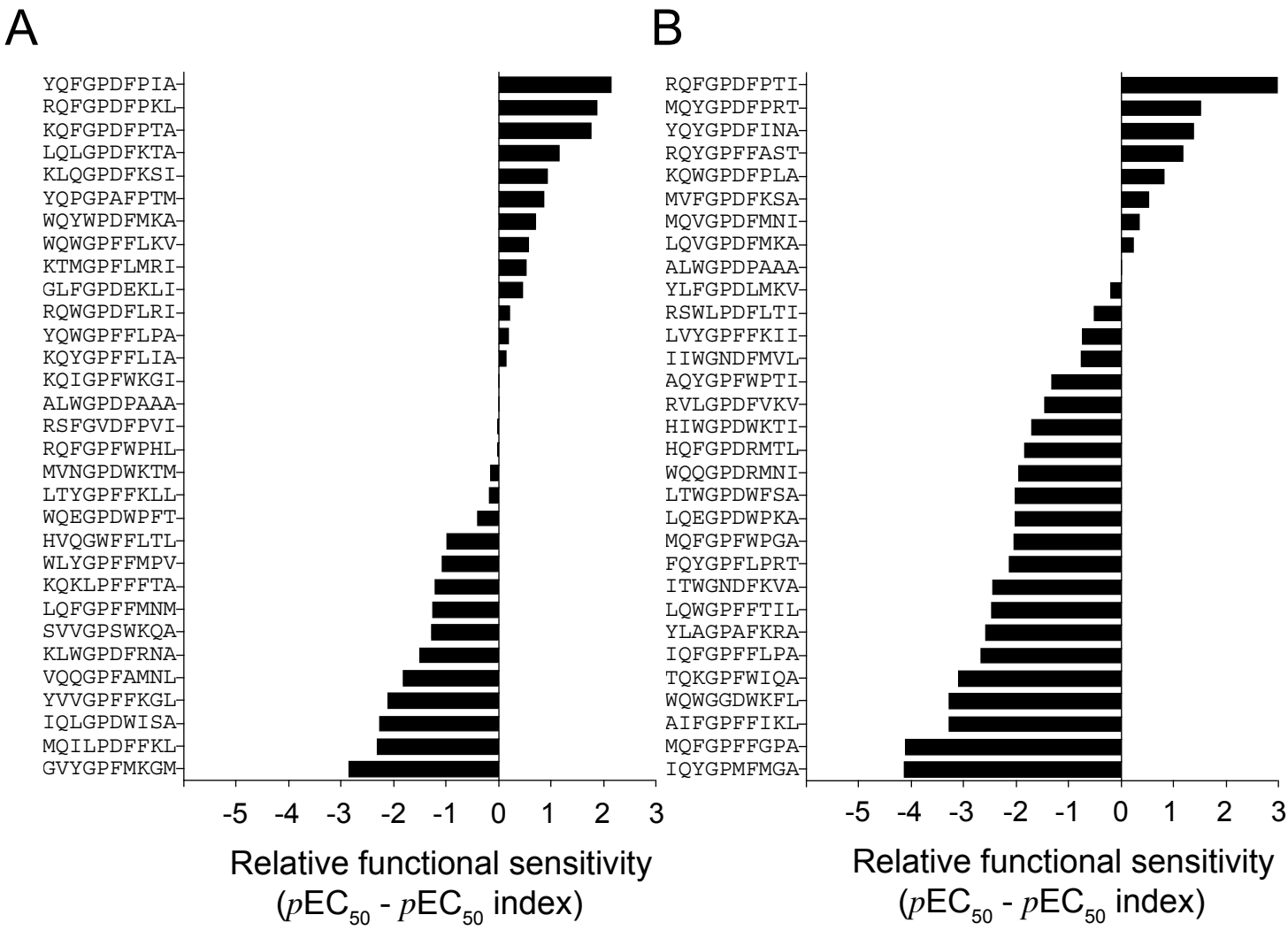
A
YQFGPDFPIA
RQFGPDFPKL
KQFGPDFPTA
LQLGPDFKTA
KLQGPDFKSI
YQPGPAFPTM
WQYWPDFMKA
WQWGPFFLKV
KTMGPFLMRI
GLFGPDEKLI
RQWGPDFLRI
YQWGPFFLPA
KQYGPFFLIA
KQIGPFWKGI
ALWGPDPAAA
RSFGVDFPVI
RQFGPFWPHL
MVNGPDWKTM
LTYGPFFKLL
WQEGPDWPFT
HVQGWFFLTL
WLYGPFFMPV
KQKLPFFFTA
LQFGPFFMNM
SVVGPSWKQA
KLWGPDFRNA
VQQGPFAMNL
YVVGPFFKGL
IQLGPDWISA
MQILPDFFKL
GVYGPFMKGM
-5 -4 -3 -2 -1 0 1 2 3
Relative functional sensitivity
(pEC$_{50}$ - pEC$_{50}$ index)
B
RQFGPDFPTI
MQYGPDFPRT
YQYGPDFINA
RQYGPFFAST
KQWGPDFPLA
MVFGPDFKSA
MQVGPDFMNI
LQVGPDFMKA
ALWGPDPAAA
YLFGPDLMKV
RSWLPDFLTI
LVYGPFFKII
IIWGNDFMVL
AQYGPFWPTI
RVLGPDFVKV
HIWGPDWKTI
HQFGPDRMTL
WQQGPDRMNI
LTWGPDWFSA
LQEGPDWPKA
MQFGPFWPGA
FQYGPFLPRT
ITWGNDFKVA
LQWGPFFTIL
YLAGPAFKRA
IQFGPFFLPA
TQKGPFWIQA
WQWGGDWKFL
AIFGPFFIKL
MQFGPFFGPA
IQYGPMFMGA
-5 -4 -3 -2 -1 0 1 2 3
Relative functional sensitivity
(pEC$_{50}$ - pEC$_{50}$ index)

Figure S5

Supplementary Material: *Mathematica* Code

```
AmNum[AmAc_] :=
 Module[{},
  First[First[
    Position[{"A", "C", "D", "E", "F", "G", "H", "I", "K", "L", "M", "N", "P",
       "Q", "R", "S", "T", "V", "W", "Y"}, AmAc]]]];

GenerateSample[mat_, merN_, samN_] :=
 Module[{res, matnorm, biases, cnt, cntstop}, cntstop = 500 samN; cnt = 0;
  res = {1};
While[Length[Union[res]] < samN && cnt < cntstop, Print["Attempt"]; cnt++; res
= Transpose[Table[RandomChoice[mat[[p]] -> {"A", "C", "D", "E", "F", "G", "H",
"I", "K", "L", "M", "N", "P", "Q", "R", "S", "T", "V", "W", "Y"}, samN], {p,
merN}]]];
  Print["There are ", Length[Union[res]]," distinct peptides in the sample."];
  matnorm = Transpose[Table[Normalize[mat[[p]], Total], {p, merN}]];
  biases = Normalize[Exp[-Map[Total,Log[Table[matnorm[[AmNum[res[[i, p]]], p]],
{i, samN}, {p, merN}]]]],Total];
Transpose[Append[Transpose[res], N[biases]]] // MatrixForm]

EffectiveSampleSize[mat_, merN_] := Module[{matnorm, p, HH, Htot},
  (* actually, effective population size *)
  matnorm = Table[Normalize[mat[[p]], Total], {p, merN}];
  HH = Table[0, {p, 1, merN}]; Htot = 0;
  For[p = 1, p < merN + 1, p++,
   HH[[p]] = Sum[EntroTerm[matnorm[[p, a]]], {a, 1, 20}];
   Htot += HH[[p]];
   ];
  Exp[Htot]
  ]

EntroTerm[x_] := Module[{}, If[x > 0, x Log[1/x], 0]]

MakeWeightMatrix[motif_] :=
 Module[{MM}, MM = Table[0, {i, 10}, {j, 20}]; motif[[1]];
  For[i = 1, i < 11, i++,
   MM[[i, First /@
StringPosition["ACDEFGHIKLMNPQRSTVWY",StringSplit[motif[[i]]]]]] = 1]; MM]
```

Example of simultaneous fitting:

```
NumPep = 21; inits = Table[-6, {i, 3 + NumPep}]; inits[[1]] = 0;
inits[[2]] = 3000; inits[[3]] = 1; Print[inits]; pars = Array[a, 3 + NumPep];
PickPar[pars_, y_, ymax_] := Sum[pars[[3 + i]]*If[i == y, 1, 0], {i, ymax}];
model[pars_, SetInd_, NumPep_] := pars[[1]] + (pars[[2]] - pars[[1]])/(1 +
10^{(-pars[[3]])*(x - PickPar[pars, SetInd, NumPep])}); fitresult =
FindFit[data, model[pars, y, NumPep], Table[pars[[i]] (1 - j) + inits[[i]] j,
{i, 3 + NumPep}, {j, 0, 1}], {x, y},MaxIterations -> 1000]

NumPerSet = 10; MyPlotRange = {{-14, -3}, {-150, 2000}}; plts =
 Array[b, NumPep]; Do[
 plts[[i]] = Plot[model[pars, i, NumPep] /. fitresult, {x, -14, -1},PlotStyle -
> Black, PlotRange -> MyPlotRange], {i, 1, NumPep}]; tmp_dat1 =
 Show[plts[[Range[NumPep]]], PlotRange -> MyPlotRange, Frame -> True]; tmp_dat2
=ListPlot[data[[Range[NumPerSet NumPep], {1, 3}]], Frame -> True, PlotStyle ->
Black, PlotRange -> MyPlotRange]; Show[tmp_dat1, tmp_dat2]
```

Figure S6

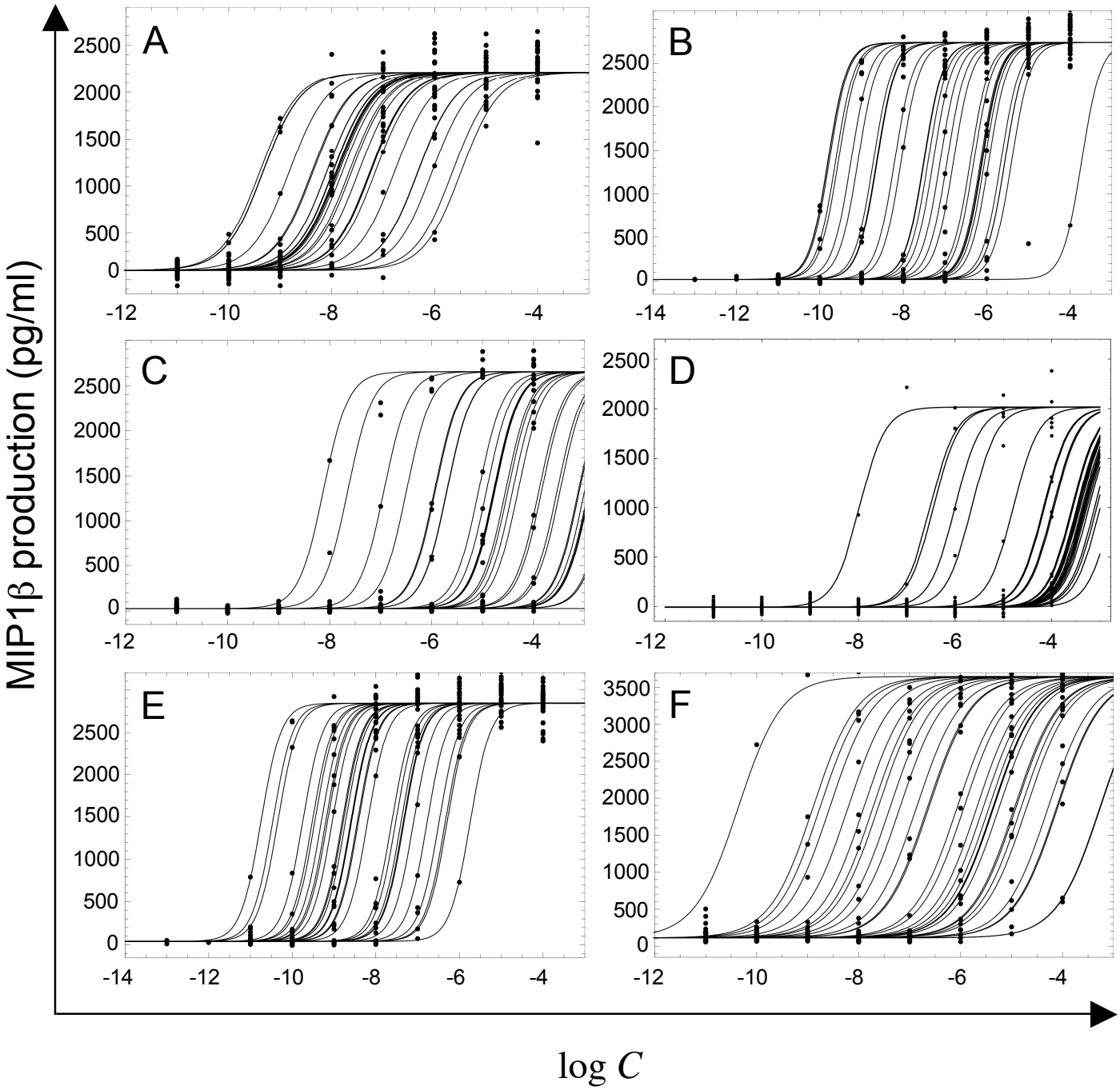

Table S1

Amino Acid

Peptide Position

	A	C	D	E	F	G	H	I	K	L
1	302.74	0	5	5	70.66	365.48	512.89	361.83	1528.32	891.17
2	5	0	5	5	5	110.86	5	246.09	5	949.64
3	400	0	5	400	2856.85	5	110.25	441.62	400	607.31
4	5	0	5	5	5	2234.31	5	5	5	50
5	5	0	5	5	5	5	50	5	5	5
6	398.38	0	3064.57	5	2271.47	5	5	5	5	5
7	121.83	0	15.84	10	3044.47	10	224.16	10	10	254.01
8	500	0	5	61.52	200	46.90	307.61	553.71	2861.12	1811.57
9	518.38	0	5	5	200	913.71	364.87	1497.26	1252.39	360
10	2248.32	0	5	5	5	5	5	1944.97	5	2000

	M	N	P	Q	R	S	T	V	W	Y
1	821.12	5	5	5	1863.35	194.31	176.04	236.35	991.07	1061.12
2	61.52	66.40	5	4851.17	5	275.94	546.40	835.74	5	5
3	31.68	400	116.95	421.52	5	5	5	400	2348.22	2248.93
4	5	5	5	5	5	5	5	5	50	5
5	5	50	2711.88	5	5	5	5	50	50	5
6	50	5	5	5	50	5	5	5	5	5
7	103.55	10	10	10	140.71	10	91.37	10	1242.64	28.02
8	2301.93	65.18	3363.05	5	200	5	5	200	18.88	276.55
9	31.07	908.83	1254.82	383.15	557.36	1237.16	1270.66	1195.74	5	15.84
10	345.99	5	17.06	5	5	5	854.01	1000	5	5

Table S2

Amino Acid

Peptide Position

	A	C	D	E	F	G	H	I	K	L
1	0.095 103	0	0.001 571	0.001 571	0.022 197	0.114 812	0.161 12	0.113 664	0.480 108	0.279 951
2	0.000 988	0	0.000 988	0.000 988	0.000 988	0.021 915	0.000 988	0.048 647	0.000 988	0.187 726
3	0.088 775	0	0.001 11	0.088 775	0.634 043	0.001 11	0.024 469	0.098 013	0.088 775	0.134 785
4	0.002 237	0	0.002 237	0.002 237	0.002 237	0.999 46	0.002 237	0.002 237	0.002 237	0.022 366
5	0.001 842	0	0.001 842	0.001 842	0.001 842	0.001 842	0.018 424	0.001 842	0.001 842	0.001 842
6	0.103 851	0	0.798 888	0.001 303	0.592 14	0.001 303	0.001 303	0.001 303	0.001 303	0.001 303
7	0.036 76	0	0.004 779	0.003 017	0.918 637	0.003 017	0.067 639	0.003 017	0.003 017	0.076 645
8	0.092 952	0	0.000 93	0.011 437	0.037 181	0.008 72	0.057 187	0.102 936	0.531 891	0.336 778
9	0.145 742	0	0.001 406	0.001 406	0.056 23	0.256 89	0.102 585	0.420 958	0.352 111	0.101 215
10	0.586 658	0	0.001 305	0.001 305	0.001 305	0.001 305	0.001 305	0.507 504	0.001 305	0.521 862

	M	N	P	Q	R	S	T	V	W	Y
1	0.257 945	0.001 571	0.001 571	0.001 571	0.585 352	0.061 042	0.055 301	0.074 245	0.311 333	0.333 339
2	0.012 162	0.013 125	0.000 988	0.958 982	0.000 988	0.054 548	0.108 012	0.165 209	0.000 988	0.000 988
3	0.007 03	0.088 775	0.025 957	0.093 552	0.001 11	0.001 11	0.001 11	0.088 775	0.521 159	0.499 123
4	0.002 237	0.002 237	0.002 237	0.002 237	0.002 237	0.002 237	0.002 237	0.002 237	0.022 366	0.002 237
5	0.001 842	0.018 424	0.999 297	0.001 842	0.001 842	0.001 842	0.001 842	0.018 424	0.018 424	0.001 842
6	0.013 034	0.001 303	0.001 303	0.001 303	0.013 034	0.001 303	0.001 303	0.001 303	0.001 303	0.001 303
7	0.031 246	0.003 017	0.003 017	0.003 017	0.042 458	0.003 017	0.027 57	0.003 017	0.374 954	0.008 455
8	0.427 936	0.012 117	0.625 202	0.000 93	0.037 181	0.000 93	0.000 93	0.037 181	0.003 51	0.051 411
9	0.008 734	0.255 52	0.352 796	0.107 723	0.156 703	0.347 829	0.357 249	0.336 184	0.001 406	0.004 453
10	0.090 279	0.001 305	0.004 45	0.001 305	0.001 305	0.001 305	0.222 838	0.260 931	0.001 305	0.001 305

Table S3

I: RQWGPDP{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	RQWGPDPNCA	5.484
2	RQWGPDPCCA	5.651
3	RQWGPDPDPA	6.047
4	RQWGPDPSMA	6.319
5	RQWGPDPCIA	6.326
6	RQWGPDPDNA	6.777
Index	ALWGPDPAAA	7.151
7	RQWGPDPRVA	7.236
8	RQWGPDPIHA	7.24
9	RQWGPDPKRA	7.257
10	RQWGPDPRTA	7.473
11	RQWGPDPKIA	7.612
12	RQWGPDPYCA	7.659
13	RQWGPDPFHA	7.791
14	RQWGPDPLKA	7.854
15	RQWGPDPAIA	7.882
16	RQWGPDPYSA	7.884
17	RQWGPDPKCA	7.887
18	RQWGPDPSTA	7.907
19	RQWGPDPYNA	7.912
20	RQWGPDPKTA	7.916
21	RQWGPDPMSA	7.969
22	RQWGPDPMAA	7.979
23	RQWGPDPHHA	8.093
24	RQWGPDPLAA	8.096
25	RQWGPDPHIA	8.36
26	RQWGPDPARA	8.384
27	RQWGPDPALA	8.823
28	RQWGPDPPNA	9.328
29	RQWGPDPPLA	9.335
30	RQWGPDPPVA	9.395

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

Number	Peptide Sequence	pEC_{50}
1	RQWGPFPSYV	3.723
2	RQWGPFFMDA	5.401
3	RQWGPFFIRI	5.561
4	RQWGPFPWDL	5.564
5	RQWGPFFRGA	5.658
6	RQWGPDFGEI	5.949
7	RQWGPFFPWV	6.048
8	RQWGPFFRTA	6.102
9	RQWGPFPHFV	6.121
10	RQWGPFFSQL	6.147
11	RQWGPDFEGV	6.306
12	RQWGPDPDDI	6.391
13	RQWGPFFKNV	6.795
14	RQWGPDFWGI	6.949
15	RQWGPFFKSA	7.114
16	RQWGPDFFYV	7.28
17	RQWGPFPAYV	7.404
18	RQWGPFFHTA	7.498
19	RQWGPDPGTI	7.51
20	RQWGPDFSFV	8.05
21	RQWGPDFVNI	8.206
22	RQWGPDFAHV	8.625
Index	ALWGPDPAAA	8.639
23	RQWGPDPESI	8.716
24	RQWGPDPHEI	9.086
25	RQWGPDFFVV	9.253
26	RQWGPDPMAA	9.556
27	RQWGPDPTNV	9.622
28	RQWGPDPHSV	9.774
29	RQWGPDPMQL	9.802
30	RQWGPDFMTV	9.808

III: RQxGPDxxxA motif

Number	Peptide Sequence	pEC_{50}
1	RQVGPDIWVA	Null
2	RQPGPDQINA	2.538
3	RQEGPDYFEA	2.567
4	RQPGPDNNMA	2.891
5	RQAGPDIYQA	2.901
6	RQPGPDQRLA	2.92
7	RQTGPDLDYA	2.926
8	RQEGPDESVA	2.996
9	RQDGPDREDA	3.099
10	RQYGPDVWQA	3.12
11	RQKGPDAEGA	3.479
12	RQIGPDDQLA	3.536
13	RQEGPDMDPA	3.84
14	RQHGPDQDEA	3.894
15	RQPGPDMHSA	4.302
16	RQLGPDSQHA	4.409
17	RQRGPDQMAA	4.502
18	RQSGPDDLLA	4.548
19	RQSGPDRRIA	4.756
20	RQDGPDPTTA	4.771
21	RQDGPDPQRA	4.784
22	RQPGPDNNYA	4.922
23	RQPGPDLNSA	5.076
24	RQSGPDPHRA	5.694
25	RQMGPDWSEA	5.701
26	RQNGPDQMSA	5.924
27	RQMGPDEPRA	5.944
Index	ALWGPDPAAA	6.49
28	RQSGPDHATA	6.932
29	RQMGPDFRLA	7.642
30	RQPGPDFHPA	8.12

IV: xQxGPDxxxV motif

Number	Peptide Sequence	pEC_{50}
1	DQFGPDRDPV	2.731
2	QQDGPDPNLV	2.991
3	VQYGPDSTEV	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	GQGGPDNLSV	3.556
21	LQSGPDVNIV	3.577
22	AQEGPDWDKV	3.949
23	IQTGPDKIAV	3.975
24	QQHGPDSQYV	4.147
25	EQWGPDQKWV	4.166
26	SQEGPDPFGV	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	YVVGPFFKGL	6.487
5	VQQGPFAMNL	6.771
6	KLWGPDFRNA	7.095
7	SVVGPSWKQA	7.323
8	LQFGPFFMNM	7.343
9	KQKLPFFFTA	7.395
10	WLYGPFFMPV	7.523
11	HVQGWFFLTL	7.617
12	WQEGPDWPFT	8.197
13	LTYGPFFKLL	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	YQPGPAFPTM	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	IQYGPMFMGA	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	LTWGPDWFSA	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