Supplementary Table S1. Site directed mutation of signal peptide of preproinsulin

Designation	Sequence
Wild type PPI SP sequence	MALWMRLLPLLALLALWGPDPAAA
9L mutant	MALWMRLL <u>L</u> LLLLLLWGPDPAAA
12L mutant	MALWMRLLLLL <u>L</u> LLLLWGPDPAAA
15L mutant	MALWMRLLLLLLL <u>L</u> LWGPDPAAA
9L/12L mutant	MALWMRLL <u>L</u> LLLLWGPDPAAA
9L/15L mutant	MALWMRLL <u>L</u> LLLLLLWGPDPAAA
12L/15L mutant	MALWMRLLLLL <u>L</u> LL <u>L</u> LWGPDPAAA
9L/12L/15L mutant	MALWMRLL <u>L</u> LL <u>L</u> LL <u>L</u> LWGPDPAAA

Supplementary Table S2. Details of subjects studied for pancreatic expression of antigen processing enzymes

Case ID	Group	Source	Donor Status	Age (years)	Duration
485 88	No diabetes control	UK Pancreas Biobank	Autopsy	2	
315 89	No diabetes control	UK Pancreas Biobank	Autopsy	9	
65 71	No diabetes control	UK Pancreas Biobank	Autopsy	40	
8579	No diabetes control	UK Pancreas Biobank	Autopsy	7	
540 91	No diabetes control	UK Pancreas Biobank	Autopsy	11	
PAN8	No diabetes control	UK Pancreas Biobank	Autopsy	19	
E560	Type 1 diabetes	UK Type 1 diabetes Biobank	Organ Donor	42	1.5y
Sc115	Type 1 diabetes	UK Type 1 diabetes Biobank	Autopsy	1	0 'Recent'
E124B	Type 1 diabetes	UK Type 1 diabetes Biobank	Autopsy	17	0 'Recent'
E375	Type 1 diabetes	UK Type 1 diabetes Biobank	Autopsy	11	unknown
11746	Type 1 diabetes	UK Type 1 diabetes Biobank	Autopsy	6	1 week
11713	Type 1 diabetes	UK Type 1 diabetes Biobank	Autopsy	3	3 months

Supplementary Table S3. Details of the staining protocol for each antibody employed in the current study. Where multiple antigens were stained on the same section, antibodies were applied sequentially (staining order: anti-TAP1 (overnight at 4°C) followed by anti-HLA-ABC (1h at room temperature) or anti-glucagon (1h at room temperature) then anti-insulin plus DAPI (1h at room temperature)).

Primary Antibody	Manufacturer and clone	Antigen Retrieval	Antibody Dilution	Incubation time with primary antibody	Secondary Detection System
HLA-ABC	Abcam C#ab70328 Mouse monoclonal EMR8-5	10mM citrate pH6.0	1/1000	1h at RT	Immunofluoroscence staining using anti-mouse IgG (H+L) Alexa Fluor™-conjugated secondary antibodies (1/400 for 1h)
Insulin	DakoC#A0564 Guinea-pig polyclonal	10mM citrate pH6.0	1/700	1h at RT	Immunofluoroscence staining using anti-guinea-pig IgG (H+L) Alexa Fluor TM -conjugated secondary antibodies (1/400 for 1h)
Glucagon	AbcamC#ab92517 Rabbitmonoclonal EP3070	10mM citrate pH6.0	1/4000	1h at RT	Immunofluoroscence staining using anti-rabbit IgG (H+L) Alexa Fluor TM -conjugated secondary antibodies (1/400 for 1h)
TAP1	Protein Tech C#11114-1-AP Rabbit polyclonal	10mM citrate pH6.0	1/200	o/n 4°C	Immunofluoroscence staining using anti-rabbit IgG (H+L) Alexa Fluor™ -conjugated secondary antibodies (1/400 for 1h)
ERAP	R & D Systems C#AF2334 Goat polyclonal	10mM Tris, 1mM EDTA pH9.0	1/200	o/n 4°C	Dako REAL TM Envision TM Detection System with anti-goat IgG HRP-conjugated secondary antibodies (1/800 for 1h).

Supplementary Table S4. Oligonucleotides used for CRISPR-mediated gene knockout.

	Forward	Reverse
Exon 2	GAAGCGGGCGCATCCCGGCGTTTT	GCCGGGATGCCGCCCGCTTCCGGTG
Exon 3	GCAGGAGGTTGATGTACTCCGTTTT	GGAGTACATCAACCTCCTGCCGGTG
Exon 10	GCAGCCTACATCTTCGGCCTGTTTT	AGGCCGAAGATGTAGGCTGCCGGTG
Exon 11	CAGGACAGGAAAACCGATGCGTTTT	GCATCGGTTTTCCTGTCCTGCGGTG

Supplementary Table S5. Eluted epitopes are not selectively derived from SP containing source protein nor their signal peptide region.

	HLA- A*02:01	HLA- A*11:01	HLA- A*24:02	HLA- B*18:01	HLA- B*38:01	HLA- B*39:06	AII HLA	Human proteome
Source proteins	458	667	190	510	378	259	2462	20226
Number of source proteins containing a signal peptide	26	29	17	25	35	21	153	3560
Frequency of signal peptide containing source proteins	5.68%	4.35%	8.95%	4.90%	9.26%	8.11%	6.21%	17.6%
Number of proteins with signal peptide-derived epitope	12	2	4	5	4	13	40	
Frequency of identified SP epitopes from SP-Protein	46.15%	6.90%	23.53%	20.00%	11.43%	61.90%	26.14%	

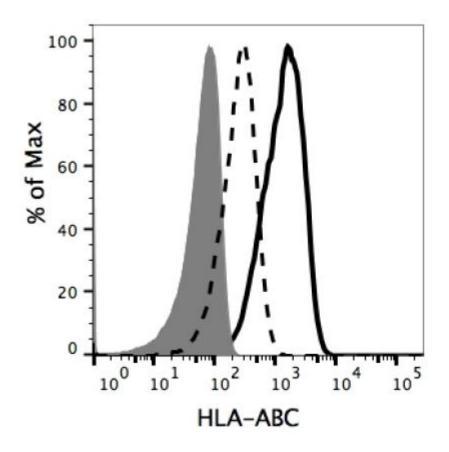
Source Proteins are those represented by at least one epitope with Mascot Score >40 within our HLA class I elution data (1). Uniprot (2) (access via http://www.uniprot.org/) retrieve/ID mapping function was used to identify signal peptide containing source proteins. SP, signal peptide.

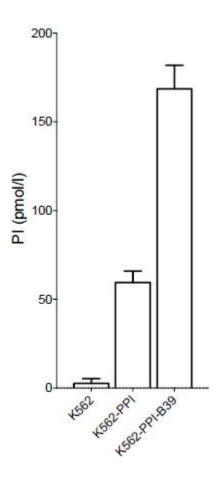
Supplementary Table S6. Putative PPI epitopes identified by in silico binding prediction algorithms.

Binding prediction were performed by SYFPEITHI (3) (access via : www.syfpeithi.de)) to identify putative PPI epitopes to compare eluted epitopes to those identified in silico. Epitopes derived from the PPI signal peptide are highlighted in bold and those identified by elution in green.

A2	9-mer	Score	A11	9-mer	Score	A24	9-mer	Score	B18	8-mer	Score	B38	9-mer	Score
1 6	RLLPLLALL	31	38	ALYLVCGER	18	6	RLLPLLALL	15	34	HLVEALYL	15	33	SHLVEALYL	21
2 2	ALWMRLLPL	28	81	ALEGSLQKR	18	17	WGPDPAAAF	13	36	VEALYLVC	14	5	MRLLPLLAL	18
3 34	HLVEALYLV	27	35	LVEALYLVC	16	39	LYLVCGERG	13	56	REAEDLQV	14	6	RLLPLLALL	14
4 60	DLQVGQVEL	25	75	GSLQPLALE	16	97	TSICSLYQL	13	61	LQVGQVEL	13	3	LWMRLLPLL	13
5 15	ALWGPDPAA	22	76	SLQPLALEG	16	3	LWMRLLPLL	12	32	GSHLVEAL	12	2	ALWMRLLPL	12
6 8	LPLLALLAL	20	32	GSHLVEALY	15	5	MRLLPLLAL	12	42	VCGERGFF	12	8	LPLLALLAL	12
7 3	LWMRLLPLL	19	80	LALEGSLQK	15	8	LPLLALLAL	12	58	AEDLQVGQ	12	31	CGSHLVEAL	12
8 5	MRLLPLLAL	19	62	QVGQVELGG	14	22	AAAFVNQHL	12	3	LWMRLLPL	11	60	DLQVGQVEL	12
9 81 10 7	ALEGSLQKR	19	84	GSLQKRGIV	14	31	CGSHLVEAL	12	6 9	RLLPLLAL	11	72 74	PGAGSLQPL	12
	LLPLLALLA	18	85	SLQKRGIVE	14	33	SHLVEALYL	12	_	PLLALLAL	11 11		AGSLQPLAL	12
11 76 12 22	SLQPLALEG AAAFVNQHL	18 17	96 2	CTSICSLYQ ALWMRLLPL	14 13	49 78	FYTPKTRRE QPLALEGSL	12 12	33 41	S H L V E A L Y L V C G E R G F	11	97 28	TSICSLYQL QHLCGSHLV	12 11
13 69	GGGPGAGSL	17	4	WMRLLPLLA	13	2	ALWMRLLPL	11	44	GERGFFYT	11	69	GGGPGAGSL	11
14 97	TSICSLYQL	17	7	LLPLLALLA	13	53	KTRREAEDL	11	66	VELGGGPG	11	78	QPLALEGSL	11
15 9	PLLALLALW	16	47	GFFYTPKTR	13	69	GGGPGAGSL	11	75	GSLQPLAL	11	94	QCCTSICSL	11
16 10	LLALLALW	16	10	LLALLALWG	12	74	AGSLQPLAL	11	82	LEGSLQKR	11	22	AAAFVNQHL	10
17 31	CGSHLVEAL	16	29	HLCGSHLVE	12	94	QCCTSICSL	11	92	VEQCCTSI	11	91	IVEQCCTSI	10
18 58	AEDLQVGQV	16	60	DLQVGQVEL	12	27	NQHLCGSHL	10	96	CTSICSLY	11	17	WGPDPAAAF	9
19 94	QCCTSICSL	16	98	SICSLYQLE	12	40	YLVCGERGF	10	1	MALWMELL	10	27	NQHLCGSHL	9
20 12	ALLALWGPD	15	12	ALLALWGPD	11	41	LVCGERGFF	10	18	GPDPAAAF	10	41	LVCGERGFF	9
21 13	LLALWGPDP	15	41	LVCGERGFF	11	60	DLQVGQVEL	10	43	CGERGFFY	10	53	KTRREAEDL	9
22 18	GPDPAAAFV	15	45	ERGFFYTPK	11	72	PGAGSLQPL	10	54	TRREAEDL	10	83	EGSLQKRGI	9
23 28	QHLCGSHLV	15	50	YTPKTRREA	11	83	EGSLQKRGI	10	101	SLYQLENY	10	40	YLVCGERGF	8
24 33	SHLVEALYL	15	53	KTRREAEDL	11	91	IVEQCCTSI	10	4	WMRLLPLL	9	55	RREAEDLQV	8
25 38	ALYLVEGER	15	97	TSICSLYQL	11	102	LYQLENYCN	10	7	LLPLLALL	9	54	TRREAEDLQ	7
23 30	ALILVEGEN	13	21	ISICSLIQL	11	102	LIQLENTON	10	,	LLFLLALL	-	34	INNEALDEQ	,
A2	10-mer	Score	A11	10-mer	Score	A24	10-mer	Score	B18	9-mer	Score	B38	10-mer	Score
1 2	ALWMRLLPLL	28	79	PLALEGSLQK	21	39	LYLVCGERGF	22	100	CSLYQLENY	15	5	MRLLPLLALL	19
1 2 2 7	ALW MRLLPLL LLPLLALLAL	28 28	79 75	PLALEGSLQK GSLQPLALEG	21 20	39 5	LYLVCGERGF MRLLPLLALL	22 14	100 60	CSLYQLENY DLQVGQVEL	15 14	5 59	MRLLPLLALL EDLQVGQVEL	19 14
1 2 2 7 3 4	ALWMRLLPLL LLPLLALLAL WMRLLPLLAL	28 28 24	79 75 84	PLALEGSLQK GSLQPLALEG GSLQKRGIVE	21 20 19	39 5 26	LYLVCGERGF MRLLPLLALL VNQHLCGSHL	22 14 13	100 60 31	CSLYQLENY DLQVGQVEL CGSHLVEAL	15 14 13	5 59 7	MRLLPLLALL EDLQVGQVEL LLPLLALLAL	19 14 13
1 2 2 7 3 4 4 29	ALWMRLLPLL LLPLLALLAL WMRLLPLLAL HLCGSHLVEA	28 28 24 24	79 75 84 32	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL	21 20 19 18	39 5 26 1	LYLVCGERGF MRLLPLLALL VNQHLCGSHL MALWMRLLPL	22 14 13 12	100 60 31 44	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP	15 14 13 13	5 59 7 73	MRLLPLLALL EDLQVGQVEL LLPLLALLAL GAGSLQPLAL	19 14 13 13
1 2 2 7 3 4 4 29 5 15	ALWMRLLPLL LLPLLALLAL WMRLLPLLAL HLCGSHLVEA ALWGPDPAAA	28 28 24 24 22	79 75 84 32 6	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLLPLLALLA	21 20 19 18 15	39 5 26 1 2	LYLVCGERGF MRLLPLLALL VNQHLCGSHL MALWMRLLPL ALWMRLLPLL	22 14 13 12	100 60 31 44 56	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG	15 14 13 13	5 59 7 73 1	MRLLPLLALL EDLQVGQVEL LLPLLALLAL GAGSLQPLAL MALWMRLLPL	19 14 13 13 12
1 2 2 7 3 4 4 29 5 15 6 1	ALW MRLLPLL LLPLLALLAL WM RLLPLLAL HLCGSHLVEA ALW GPDPAAA MALW MRLLPL	28 28 24 24 22 21	79 75 84 32 6 35	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLLPLLALLA LVEALYLVCG	21 20 19 18 15	39 5 26 1 2 7	LYLVCGERGF MRLLPLLALL VNQHLCGSHL MALWMRLLPL ALWMRLLPLL LLPLLALLAL	22 14 13 12 12	100 60 31 44 56 8	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLLALLAL	15 14 13 13 13	5 59 7 73 1 2	MRLLPLLALL EDLQVGQVEL LLPLLALLAL GAGSLQPLAL MALWMRLLPL ALWMRLLPLL	19 14 13 13 12 12
1 2 2 7 3 4 4 29 5 15 6 1 7 90	ALWMRLLPLL LLPLLALLAL WMRLLPLLAL HLCGSHLVEA ALWGPDPAAA MALWMRLLPL GIVEQCCTSI	28 28 24 24 22 21 21	79 75 84 32 6 35 85	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLPLLALLA LVEALYLVCG SLQKRGIVEQ	21 20 19 18 15 15	39 5 26 1 2 7 30	LYLVCGERGF MRLLPLLALL VNQHLCGSHL MALW MRLLPL ALW MRLLPLL LLPLLALLAL LCGSHLVEAL	22 14 13 12 12 12 12	100 60 31 44 56 8 36	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLLALLAL VEALYLVCG	15 14 13 13 13 12 12	5 59 7 73 1 2	MRLLPLLALL EDLQVGQVEL LLPLLALLAL GAGSLQPLAL MALWMRLLPL ALWMRLLPLL WMRLLPLLAL	19 14 13 13 12 12 12
1 2 2 7 3 4 4 29 5 15 6 1 7 90 8 85	ALWMRLIPLL LIPLLALLAL WMRLIPLLAL HLCGSHLVEA ALWGPDPAAA MALWMRLIPL GIVEQCCTSI SLQKRGIVEQ	28 28 24 24 22 21 21 20	79 75 84 32 6 35 85 98	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLIALIA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN	21 20 19 18 15 15 15	39 5 26 1 2 7 30 59	LYLVCGERGF MRLLPLLALL VNQHLCGSHL MALWMRLLPL ALW MRLLPL LLPLLALLAL LCGSHLVEAL EDLQVGQVEL	22 14 13 12 12 12 12 12	100 60 31 44 56 8 36 58	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLIALIAL VEALYLVCG AEDLQVGQV	15 14 13 13 13 12 12	5 59 7 73 1 2 4 32	MRLLPLLALL EDLQVGQVEL LLPLLALLAL GAGSLQPLAL MALWMRLLPL ALWMRLLPLL WMRLLPLL GSHLVEALYL	19 14 13 13 12 12 12 12
1 2 2 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6	ALWMRLPLL LIPLALLAL WMRLPLLAL HLCGSHLVEA ALWGPDPAAA MALWMRLLPL GIVEQCCTSI SLQKRGIVEQ RLLPLLALLA	28 28 24 24 22 21 21 20 19	79 75 84 32 6 35 85 98 44	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLLALLA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN GERGFFYTPK	21 20 19 18 15 15 15 15	39 5 26 1 2 7 30 59 77	LYLVCGERGF MRLLPLLALL VNQHLCGSHL MALWMRLLPL ALWMRLLPL LLPLLALLAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL	22 14 13 12 12 12 12 12 12 12	100 60 31 44 56 8 36 58 74	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL	15 14 13 13 13 12 12 12	5 59 7 73 1 2 4 32 33	MRLLPLLALL EDLQVGQVEL LLPLLALLAL GAGSLQPLAL MALWMRLLPL ALWMRLLPLL WMRLLPLL GSHLVEALYL SHLVEALYLV	19 14 13 13 12 12 12 12 12
1 2 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6 10 10	ALWMRLPLL LIPLLALLAL WMRLIPLLAL HLCGSHLVEA ALWGPDPAAA MALWMRLIPL GIVEQCCTSI SLQKRGIVEQ RLIPLLALLA LLALLALWGP	28 28 24 24 22 21 21 20 19	79 75 84 32 6 35 85 98 44	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLLPLLALLA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN GERGFFYTPK PLLALLALWG	21 20 19 18 15 15 15 15 14	39 5 26 1 2 7 30 59 77 90	LYLVCGERGF MRLIPLLALL VNQHLCGSHL MALWMRLIPL LLPLLALLAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI	22 14 13 12 12 12 12 12 12 12 12	100 60 31 44 56 8 36 58 74 78	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLLALLAL VEALYVCG AEDLQVGQV AGSLQPLAL QPLALEGSL	15 14 13 13 13 12 12 12 12	5 59 7 73 1 2 4 32 33 68	MRLLPLLALL EDLQVGQVEL LLPLLALLAL GAGSLQPLAL MALWMRLLPL ALWMRLLPLL WMRLLPLLAL GSHLVEALYL SHLVEALYLV LGGGPGAGSL	19 14 13 13 12 12 12 12 12 12
1 2 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6 10 10 11 33	ALWMRLIPLL LLPLLALLAL WMRLIPLIAL HLCGSHLVEA ALWGPDPAAA MALWMRLIPL GIVEQCCTSI SLQKRGIVEQ RLIPLIALLA LLALLALLWGP SHLVEALYLV	28 28 24 24 22 21 21 20 19 19	79 75 84 32 6 35 85 98 44 9	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLIALLA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN GERGFFYTPK PLIALLALWG TSICSLYQLE	21 20 19 18 15 15 15 15 14 13	39 5 26 1 2 7 30 59 77 90 16	LYLVCGERGF MRLIPLIALL VNQHLCGSHL MALWMRLLPL ALWMRLLPL LLPLLALLAL CGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF	22 14 13 12 12 12 12 12 12 12 12 12	100 60 31 44 56 8 36 58 74 78 82	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLLALLAL VEALYVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG	15 14 13 13 13 12 12 12 12 12 12	5 59 7 73 1 2 4 32 33 68 93	MRLIPLIALL EDLQVGQVEL LIPLIALLAL GAGSLQPLAL MALW MRLLPL ALW MRLIPLIAL GSHLVEALYL SHLVEALYL LGGGPGAGSL EQCCTSICSL	19 14 13 13 12 12 12 12 12 12 12
1 2 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6 10 10 11 33 12 76	ALWMRILPLL LIPILALLAL WMRILPLIAL HLGGSHLVEA ALWGPDPAAA MALWMRILPL GIVEQCCTSI SLQKRGIVEQ RILPLIALLA LIALLALWGP SHLVEALIVV SLQPLALEGS	28 28 24 24 22 21 21 20 19 19 19	79 75 84 32 6 35 85 98 44 9	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RILPLIALIA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN GERGFFYTPK PLIALIALW TSICSLYQLE HLCGSHLVEA	21 20 19 18 15 15 15 15 14 13 13	39 5 26 1 2 7 30 59 77 90 16 21	LYLVCGERGF MRLIPLIALL VNQHLCGSHL MALWMRLLPL ALWMRLLPL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL	22 14 13 12 12 12 12 12 12 12 12 11 11	100 60 31 44 56 8 36 58 74 78 82 3	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLIALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLLPLL	15 14 13 13 13 12 12 12 12 12 12 12 12	5 59 7 73 1 2 4 32 33 68 93 26	MRLIPLIALL EDLQVGQVEL LIPLALIAL GAGSLQPLAL MALWMRLIPLI WMRLIPLIAL GSHLVEALYL SHLVEALYL LGGGPGAGSL VNQHLCGSHL	19 14 13 13 12 12 12 12 12 12 12 12 12
1 2 2 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6 10 10 11 33 12 76 13 96	ALWMRLIPIL LIPILALIAL WMRLIPILAL HLCGSHLVEA ALWGPDPAAA MALWMRLIPI GIVEQCCTSI SLQKRGIVEQ RLIPILALIA LLALIALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL	28 28 24 24 22 21 21 20 19 19 19	79 75 84 32 6 35 85 98 44 9 97 29	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLLALLA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN GERGFFYTPK PLLALLALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR	21 20 19 18 15 15 15 15 14 13 13 12	39 5 26 1 2 7 30 59 77 90 16 21 49	LYLVCGERGF MRLIPLIALI VNQHLCGSHL MALWMRLLPL LIPLIALIAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREA	22 14 13 12 12 12 12 12 12 12 11 11	100 60 31 44 56 8 36 58 74 78 82 3 5	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPILALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLLPLL MRLLPLLAL	15 14 13 13 13 12 12 12 12 12 12 11 11	5 59 7 73 1 2 4 32 33 68 93 26 28	MRLIPLIALL EDLQVGQVEL LIPLIALLIAL GAGSLQPLAL MALWMRLIPL ALWMRLIPLL WMRLIPLIAL GSHLVEALYL LGGGPGAGSL EQCCTSICSL VNQHLCGSHL QHLCGSHLVE	19 14 13 13 12 12 12 12 12 12 11 11
1 2 2 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6 10 10 11 33 12 76 13 96 14 5	ALWMRLIPLL LLPLLALLAL WMRLIPLIAL HLCGSHLVEA ALWGPDPAAA MALWMRLIPL GIVEQCCTSI SLQKRGIVEQ RLIPLLALLA LLALLALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPLIALL	28 28 24 24 22 21 21 20 19 19 19 19 19	79 75 84 32 6 35 85 98 44 9 97 29 47 53	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLIALIA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN GERGFFYTPK PLIALIALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ	21 20 19 18 15 15 15 15 14 13 13 12 12	39 5 26 1 2 7 30 59 77 90 16 21 49 52	LYLVCGERGF MRLIPLIALL VNQHLCGSHL MALWMRLLPL LLPLIALLAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREA	22 14 13 12 12 12 12 12 12 12 12 11 11 11	100 60 31 44 56 8 36 58 74 78 82 3 5	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLLPLL MRLLPLLAL RLLPLLALL	15 14 13 13 13 12 12 12 12 12 12 11 11 11	5 59 7 73 1 2 4 32 33 68 93 26 28 30	MRLIPLIALL EDLQVGQVEL LIPLIALIAL GAGSLQPLAL MALWMRLIPL ALWMRLIPLL WMRLIPLLAL GSHLVEALYL SHLVEALYL LGGGPGAGSL EQCCTSICSL VNQHLCGSHL LQHLCGSHLVE LCGSHLVEAL	19 14 13 13 12 12 12 12 12 12 12 11 11 11
1 2 7 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6 10 10 11 33 12 76 13 96 14 5 15 13	ALWMRILPLI LIPILALIAL WMRILPLIAL HLGGSHLVEA ALWGPDPAAA MALWMRILPL GIVEQCCTSI SLQKRGIVEQ RLIPILALLA LLALLALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPILALL LLALWGPDPA	28 28 24 24 22 21 20 19 19 19 19 19 18	79 75 84 32 6 35 85 98 44 9 97 29 47 53 2	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLLALLA LVEALYLVEQ SICSLYQLEN GERGFFYTPK PLLALLALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ ALWMRLIPLL	21 20 19 18 15 15 15 15 14 13 13 12 12 12	39 5 26 1 2 7 30 59 77 90 16 21 49 52 73	LYLVCGERGF MRLIPLIALL VNQHLCGSHL MALWMRLLPL ALWMRLLPL LLPLLALLAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREA GAGSLQPLAL	22 14 13 12 12 12 12 12 12 12 11 11 11 11	100 60 31 44 56 8 36 58 74 78 82 3 5 6 32	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLIALLAL VEALYVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLLPLL MRLLPLLAL GSHLVEALY	15 14 13 13 13 12 12 12 12 12 12 11 11 11	5 59 7 73 1 2 4 32 33 68 93 26 28 30 52	MRLIPLIALL EDLQVGQVEL LIPLIALLAL GAGSLQPLAL MALWMRLIPL ALWMRLIPL GSHLVEALYL SHLVEALYL SHLVEALYL LGGGPGAGSL EQCCTSICSL VNQHLCGSHL QHLCGSHLVEALYL PKTRREAEDL	19 14 13 13 12 12 12 12 12 12 12 11 11 11
1 2 2 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6 10 10 11 33 12 76 13 96 14 5 15 13 16 30	ALWMRLIPIL LIPILALIAI WMRLIPILAL HLCGSHLVEA ALWGPDPAAA MALWMRLIPI GIVEQCCTSI SLQKRGIVEQ RLIPILALIA LLALIALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPILALI LLALWGPDPA LCGSHLVEAL	28 28 24 24 22 21 20 19 19 19 19 19 19 17 18 18	79 75 84 32 6 35 85 98 44 9 97 29 47 53 2	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLALLA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN PLALLALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ ALWMRLIPLL ALLALWGPDP	21 20 19 18 15 15 15 15 14 13 13 12 12 12 11	39 5 26 1 2 7 30 59 77 90 16 21 49 52 73 93	LYLVCGERGF MRLIPLIALI VNQHLCGSHL MALWMRLLPL LIPLLALIAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF PAAAFVQHL FYTPKTRREAE DL GAGSLQPLAL EQCCTSISL	22 14 13 12 12 12 12 12 12 12 11 11 11 11	100 60 31 44 56 8 36 58 74 78 82 3 5 6 32 42	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPILALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLLPLL MRLLPLLAL RSHLVEALY VCGERGFFY	15 14 13 13 13 12 12 12 12 12 11 11 11 11	5 59 7 73 1 2 4 32 33 68 93 26 28 30 52 71	MRLIPLIALL EDLQVGQVEL LIPLIALLIAL GAGSLQPLAL MALWMRLIPL ALWMRLIPLIAL GSHLVEALYL LGGGPGAGSL EQCCTSICSL VNQHLCGSHL QHLCGSHLVEAL LCGSHLVEAL FXTRREAEDL GPGAGSLQPL	19 14 13 13 12 12 12 12 12 12 11 11 11 11
1 2 2 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6 10 10 11 33 12 76 13 96 14 5 15 13 16 30 17 73	ALWMRILPLI LIPLLALLA WMRLIPLA HLCGSHLVEA ALWGPDPAAA MALWMRLIPL GIVEQCCTSI SLQKRGIVEQ RLIPLLALLA LLALLALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPLLALL LLALWGPDPA LCGSHLVEAL GGGSHLVEAL	28 28 24 24 22 21 21 20 19 19 19 19 19 18 18 18 17	79 75 84 32 6 35 85 98 44 9 7 29 47 53 2 12	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLIALIA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN GERGFFYTPK PLIALIALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ ALWMRLIPLI ALLALWGPDPAAA	21 20 19 18 15 15 15 14 13 13 12 12 12 11 11	39 5 26 1 2 7 30 59 77 90 16 21 49 52 73 93 40	LYLVCGERGF MRLIPLIALL VNQHLCGSHL MALWMRLIPL LIPLIALLAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREA PKTRREAEDL GAGSLQPLAL EQCCTSICSL YLVCGERGFF	22 14 13 12 12 12 12 12 12 12 11 11 11 11 11 11	100 60 31 44 56 8 36 58 74 78 82 3 5 6 32 42 95	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLLPLL MRLLPLLAL GSHLVEALY VCGERGFFY CCTSICSLY	15 14 13 13 13 12 12 12 12 12 12 11 11 11 11 11	5 59 7 73 1 2 4 32 33 68 93 26 28 30 52 71 96	MRLIPLIALL EDLQVGQVEL LIPPLIALLAL GAGSLQPLAL MALWMRLIPL ALWMRLIPLAL GSHLVEALYLV LGGGPGAGSL EQCCTSICSL VNQHLCGSHLV LCGSHLVEAL PKTRREAEDL CTSICSLYQL	19 14 13 13 12 12 12 12 12 12 11 11 11 11 11
1 2 2 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6 10 10 11 33 12 76 13 96 14 5 15 13 16 30 17 73 18 68	ALWMRLIPLL LIPLIALLIAL WMRLIPLAL HLCGSHLVEA ALWGPDPAA MALWMRLIPL GIVEQCCTSI SLQKRGIVEQ RLIPLIALLA LIALLALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPLIALL LIALWGPDPA LCGSHLVEAL LGGGPGAGSL	28 28 24 24 22 21 20 19 19 19 19 19 19 17 17 16	79 75 84 32 6 35 85 98 44 9 97 29 47 53 2 12 15	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLLALLA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN GERGFFYTPK PLLALLALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ ALWMRILPLL ALLALWGPDP ALWGPDPAAA HLVEALYLVC	21 20 19 18 15 15 15 14 13 13 12 12 12 11 11	39 5 26 1 2 7 30 59 77 90 16 21 49 52 73 93 40 68	LYLVCGERGF MRLIPLIALI VNQHLICGSHL MALWMRLLPL LUPLIALLAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVECCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREA PKTRREAEDL GAGSLQPLAL EQCCTSICSL YLVCGERGFF LGGGPGAGSL	22 14 13 12 12 12 12 12 12 12 11 11 11 11 11 11	100 60 31 44 56 8 36 58 74 78 82 3 5 6 32 42 95 9	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLLALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLLPLLAL RLIPLLALL GSHLVEALY VCGERGFFY CCTSICSLY PLLALLALW	15 14 13 13 12 12 12 12 12 12 11 11 11 11 11 11	5 59 7 73 1 2 4 32 33 68 93 26 28 30 52 71 96 16	MRLIPLIALI EDLQVGQVEL LIPLIALIAL GAGSLQPLAL MALWMRLIPLI ALWMRLIPLIAL GSHLVEALYL LGGGPGAGSL EQCCTSICSL VNQHLCGSHLVEALYL LGGGFLCSSHLVEALYL LGGGFLVEAL QHLCGSHLVEAL PKTRREAEDL GPGAGSLQPL CTSICSLYQL LWGPDPAAAF	19 14 13 13 12 12 12 12 12 12 11 11 11 11 11 11 10
1 2 7 7 3 4 4 29 5 15 6 1 7 90 8 85 9 6 10 10 11 33 12 76 13 96 14 5 15 13 16 30 17 73 18 68 19 80	ALWMRLIPIL LIPILALIAI WMRLIPILAL HLCGSHLVEA ALWGPDPAAA MALWMRLIPI GIVEQCCTSI SLQKRGIVEQ RLIPILALIA LLALIALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPILALI LLALWGPDPA LCGSHLVEAL GAGSLQPLAL LGGGPGAGSL LALEGSLQKR	28 28 24 22 21 21 20 19 19 19 19 19 19 17 18 18 17 17 16 16	79 75 84 32 6 35 85 98 44 9 97 29 47 53 2 12 15 34 41	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLALLA LVEALYIVCG SLQKRGIVEQ SICSLYQLEN PLALLALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ ALWMRLIPLL ALLALWGPDPAAM HLVEALYLVC LVCGERGFFY	21 20 19 18 15 15 15 14 13 13 12 12 12 11 11 11	39 5 26 1 2 7 30 59 77 90 16 21 49 52 73 93 40 68 71	LYLVCGERGF MRLIPLIALI VNQHLCGSHL MALWMRLLPL LIPLLALIAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREA PKTRREAEDL GGGSLQPLAL EQCCTSICSL YLVCGERGFF LGGGPGAGSL GPGAGSL	22 14 13 12 12 12 12 12 12 12 11 11 11 11 11 11	100 60 31 44 56 8 36 58 82 3 5 6 32 42 95 27	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPILALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLIPLL MRLIPLLAL RLLPLLALL GSHLVEALY VCGERGFFY CCTSICSLY PLALLALW NQHLCGSHL	15 14 13 13 12 12 12 12 12 12 12 11 11 11 11 11 11	5 59 7 73 1 2 4 32 33 68 93 26 28 30 52 71 96 16 21	MRLIPLIALL EDLQVGQVEL LIPLIALLIAL GAGSLQPLAL MALWMRLIPL ALWMRLIPLIAL GSHLVEALYL LGGGPGAGSL EQCCTSICSL VNQHLCGSHLVE LCGSHLVEAL PKTRREAEDL GPGAGSLQPL CTSICSLYQL LWGPDPAAAF PAAAFVNQHL	19 14 13 13 12 12 12 12 12 12 11 11 11 11 11 11 10 10
1 2 2 7 4 4 29 5 15 1 9 6 10 10 11 33 12 76 13 16 30 17 73 18 68 19 80 20 12	ALWMRILPLI LIPLIALIAM WMRLIPLIAL HLCGSHLVEA ALWGPDPAAA MALWMRILPL GIVEQCCTSI SLQKRGIVEQ RLIPLIALIA LLALIALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPLIALL LLALWGPDPA LCGSHLVEAL GAGSLQPLAL LGGGPGAGSL LALEGSLQKR ALLALWGPDP	28 28 24 24 22 21 20 19 19 19 19 19 19 18 18 17 17 16 16 15	79 75 84 32 6 35 85 98 44 9 97 29 47 53 2 12 15 34 41 100	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLIALIA LVEALYLVCG SICGSLYQLEN GERGFFYPK PLIALIALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ ALWMRLIPLI ALLALWGPDP ALWGPDPAAA HLVEALYLVC LVCGERGFFY CSLYQLENYC	21 20 19 18 15 15 15 14 13 12 12 12 11 11 11 11	39 5 26 1 2 7 30 59 77 90 16 21 49 52 73 93 40 68 71 82	LYLVCGERGF MRLIPLIALL VNQHLCGSHL MALWMRLLPL LLPLIALLAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREA PKTRREAEDL GAGSLQPLAL EQCCTSICSL YLVCGERGFF LGGGPGAGSL LGPGAGSLQPL LEGSLQKRGI	22 14 13 12 12 12 12 12 12 12 11 11 11 11 11 10 10	100 60 31 44 56 8 36 58 82 3 5 6 32 42 95 9	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL GPLALEGSL LEGSLQKRG LWMRLLPLL RLIPLLAL GSHLVEAL CCTSICSLY PLIALLALW NQHLCGSHL NQHLCGSHL	15 14 13 13 12 12 12 12 12 12 11 11 11 11 11 11 10 10	5 59 7 73 1 2 4 32 33 68 93 26 28 30 52 71 96 16 21 77	MRLIPLIALL EDLQVGQVEL LIPPLIALLIAL GAGSLQPLAL MALWMRLIPL ALWMRLIPLIAL GSHLVEALYLV LGGGPGAGSL EQCCTSICSL VNQHLCGSHLVE LCGSHLVEAL PKTRREAEDL GPGAGSLQPL CTSICSLYQL LWGPDPAAAF PAAAFNQHL LQPLALEGSL	19 14 13 13 12 12 12 12 12 12 11 11 11 11 11 10 10 10
1 2 7 4 4 29 5 15 6 1 7 9 6 10 10 11 33 12 76 13 96 14 5 15 13 16 30 88 19 80 20 12 21 57	ALWMRLIPLL LIPLLALLIA WMRLIPLA HLCGSHLVEA ALWGPDPAA MALWMRLIPL GIVEQCCTSI SLQKRGIVEQ RLIPLIALLA LIALLALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPLIALL LIALWGPDPA LCGSHLVEAL GAGSLQPLAL GAGSLQPLAL LGGGPGAGSL LALEGSLQKR ALLALWGPDP EAEDLQVGQV	28 28 24 24 22 21 20 19 19 19 19 19 17 17 16 16 16 15	79 75 84 32 6 35 85 98 44 9 97 29 47 53 2 12 15 34 41 1000 4	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLLALLA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN GERGFFYTPK PLLALLALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ ALWMRLLPLL ALLALWGPDP ALWGPDPAAA HLVEALYLVC LVCGERGFFY CSLYQLENYC	21 20 19 18 15 15 15 15 14 13 13 12 12 12 11 11 11 11 11	39 5 26 1 2 7 30 59 77 90 16 21 49 52 73 93 40 68 71 82 96	LYLVCGERGF MRLIPLIALI VNQHLICGSHL MALWMRLLPL LLPLLALLL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREA PKTRREAEDL GAGSLQPLAL EQCCTSICSL YLVCGERGFF LGGGPGAGSL GPGAGSLQPL LEGSLQKRGI CTSICSLYQL	22 14 13 12 12 12 12 12 12 12 11 11 11 11 11 11	100 60 31 44 56 8 36 58 74 78 82 2 3 5 6 32 42 95 9 27 40 41	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLLALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLLPLLAL RLIPLLAL RSHLVEALY VCGERGFFY CCTSICSLY PLLALLALW NQHLCGSHL VLVCGERGFF LVCGERGFF	15 14 13 13 13 12 12 12 12 12 12 11 11 11 11 11 11 10 10	5 59 7 73 1 2 4 32 33 68 93 26 28 30 52 71 96 16 21 77 40	MRLIPLIALI EDLQVGQVEL LIPLIALIAL GAGSLQPLAL MALWMRLIPL ALWMRLIPLIAL GSHLVEALYL UMGREPFAGSL EQCTSICSL VNQHLCGSHL QHLCGSHLVEALYL LGGGPGAGSLQPL CTSICSLVQL LWGPDPAAAF PAAAFVNQHL LQPLALEGSL LVLCERGFF	19 14 13 12 12 12 12 12 12 11 11 11 11 11 10 10 9
1 2 7 7 3 4 4 29 5 15 1 7 90 8 85 9 6 10 10 11 33 112 76 13 96 14 5 13 16 30 17 73 18 68 19 80 20 12 21 57 71	ALWMRLIPIL LIPILALIAI WMRLIPILAL HLCGSHLVEA ALWGPDPAAA MALWMRLIPI GIVEQCCTSI SLQKRGIVEQ RILPILALIA LLALIALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPILALI LLALWGPDPA LCGSHLVEAL GAGSLQPLAL LGGGPGAGSL LALEGSLQKR ALLALWGPDPA LCGSHLVEAL GAGSLQPLAL LGGGPGAGSL	28 28 24 24 22 21 20 19 19 19 19 19 19 18 18 17 17 16 16 15 15	79 75 84 32 6 35 85 98 44 9 7 29 47 53 2 12 15 34 41 100 4 7	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLALLA LVEALYIVCG SLQKRGIVEQ SICSLYQLEN PLALLALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ ALWMRLIPLL ALLALWGPDPAAA HLVEALYLVC LVCGERGFFY CSLYQLENYC WMRLIPLLAL LIPLLALLAL	21 20 19 18 15 15 15 14 13 12 12 12 11 11 11 11 11 11	39 5 26 1 2 7 30 59 77 90 16 21 49 52 73 93 40 68 71 82 96 4	LYLVCGERGF MRLIPLIALI VNQHLCGSHL MALWMRLLPLI LLPLLALIAL LCGSHLVEAL EDLQVGQVEL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREAEDL GGGSLQPLAL EQCCTSICSL YLVCGERGFF LGGGPGAGSLQPL LEGSLQKRGI CTSICSLYQL WMRLIPLIAL	22 14 13 12 12 12 12 12 12 12 11 11 11 11 11 10 10 10 9	100 60 31 44 56 8 36 58 74 78 82 3 5 6 32 42 95 9 27 40 41 66	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLLALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLIPLL MRLIPLLAL RLIPLLALL GSHLVEALY VCGERGFFY CCTSICSLY PLLALLALW NQHLCGSHL YLVCGERGF LVCGERGFF LVCGERGFF	15 14 13 13 12 12 12 12 12 12 11 11 11 11 11 10 10 10	5 59 7 73 1 2 4 32 33 68 93 26 28 30 52 71 16 21 77 40 54	MRLIPLIALL EDLQVGQVEL LIPLIALLIAL GAGSLQPLAL MALWMRLIPL ALWMRLIPLI WMRLIPLIAL GSHLVEALYLV LGGGPGAGSL EQCCTSICSL VNQHLCGSHLVEAL PKTRREAEDL GPGAGSLQPL CTSICSLYQL LWGPDPAAAF PAAAFVNQHL LQPLALEGSL YLVCGERGFF TRREAEDLQV	19 14 13 12 12 12 12 12 12 12 11 11 11 11 11 10 10 10 9 9
1 2 7 7 3 4 4 29 5 15 1 9 6 10 10 11 33 12 76 13 9 6 14 5 15 13 18 68 19 80 20 12 21 57 22 71 23 14	ALWMRLIPLL LLPLLALLAL WMRLIPLAL HLCGSHLVEA ALWGPDPAAA MALWMRLIPL GIVEQCCTSI SLQKRGIVEQ RLLPLLALLA LLALLALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPLLALL LLALWGPDPA LCGSHLVEAL GAGSLQPLAL LGGGPGAGSL LALEGSLQRR ALLALWGPDP EAEDLQVGQV GPGAGSLQPL LALWGPDPAA	28 28 24 24 24 22 21 20 19 19 19 19 18 18 18 17 17 16 15 15 15	79 75 84 32 6 35 85 98 44 9 97 29 47 53 2 12 15 34 41 1000 4 7 25	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPILALLA LVEALYLVCG SLQKRGIVEQ SICSLYQLEN GERGFFYTPK PLIALLALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ ALWMRLIPLL ALLALWGPDP ALWGPDPAAA HLVEALYLVC LVCGERGFFY CSLYQLENYC WMRLIPLLA LIPILALLAL FVNQHLCGSH	21 20 19 18 15 15 15 14 13 13 12 12 11 11 11 11 11 10 10	39 5 26 1 2 7 30 59 77 90 16 21 49 52 73 93 40 68 71 82 96 4	LYLVCGERGF MRLIPLIALL VNQHLCGSHL MALWMRLLPL LLPLLALLAL LCGSHLVEAL EDLQVGQVEL LQPLALEGSL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREA PKTRREAEDL GAGSLQPLAL EQCCTSICSL YLVCGERGFF LGGGPGAGSL CGGPGAGSL LEGSLQKRGI CTSICSLYQL WMRLIPLIAL GSHLVEALYL	22 14 13 12 12 12 12 12 12 12 11 11 11 11 11 10 10 10 9 9	100 60 31 44 56 8 36 58 74 78 82 3 5 6 32 42 95 9 27 40 41 66 2	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLLPLL RLLPLLAL GSHLVEAL VCGERGFFY CCTSICSLY PLLALLALW NQHLCGSHL YLVCGERGF LVCGERGF LVCGERGFG LVCGERGFGAALWMRLLPL	15 14 13 13 13 12 12 12 12 12 12 11 11 11 11 11 10 10 10 9	5 59 7 73 1 2 4 32 33 68 93 26 28 30 52 71 96 16 21 77 40 54 39	MRLIPLIALL EDLQVGQVEL LIPPLIALLIAL GAGSLQPLAL MALWMRLIPL ALWMRLIPLIAL GSHLVEALYLV LGGGPGAGSL EQCCTSICSL VNQHLCGSHLVE LCGSHLVEAL PKTRREAEDL GPGAGSLQPL CTSICSLYQL LWGPDPAAAF PAAAFNQHL LQPLALEGSL YLVCGERGFF TRREAEDLQ LYLVCGERGFF	19 14 13 13 12 12 12 12 12 12 11 11 11 11 10 10 9 9 8
1 2 7 7 3 4 4 29 5 15 1 7 90 8 85 9 6 10 10 11 33 112 76 13 96 14 5 15 13 16 30 17 73 18 68 19 80 20 12 21 57 12 71	ALWMRLIPIL LIPILALIAI WMRLIPILAL HLCGSHLVEA ALWGPDPAAA MALWMRLIPI GIVEQCCTSI SLQKRGIVEQ RILPILALIA LLALIALWGP SHLVEALYLV SLQPLALEGS CTSICSLYQL MRLIPILALI LLALWGPDPA LCGSHLVEAL GAGSLQPLAL LGGGPGAGSL LALEGSLQKR ALLALWGPDPA LCGSHLVEAL GAGSLQPLAL LGGGPGAGSL	28 28 24 24 22 21 20 19 19 19 19 19 19 18 18 17 17 16 16 15 15	79 75 84 32 6 35 85 98 44 9 7 29 47 53 2 12 15 34 41 100 4 7	PLALEGSLQK GSLQPLALEG GSLQKRGIVE GSHLVEALYL RLIPLALLA LVEALYIVCG SLQKRGIVEQ SICSLYQLEN PLALLALWG TSICSLYQLE HLCGSHLVEA GFFYTPKTRR KTRREAEDLQ ALWMRLIPLL ALLALWGPDPAAA HLVEALYLVC LVCGERGFFY CSLYQLENYC WMRLIPLLAL LIPLLALLAL	21 20 19 18 15 15 15 14 13 12 12 12 11 11 11 11 11 11	39 5 26 1 2 7 30 59 77 90 16 21 49 52 73 93 40 68 71 82 96 4	LYLVCGERGF MRLIPLIALI VNQHLCGSHL MALWMRLLPLI LLPLLALIAL LCGSHLVEAL EDLQVGQVEL GIVEQCCTSI LWGPDPAAAF PAAAFVNQHL FYTPKTRREAEDL GGGSLQPLAL EQCCTSICSL YLVCGERGFF LGGGPGAGSLQPL LEGSLQKRGI CTSICSLYQL WMRLIPLIAL	22 14 13 12 12 12 12 12 12 12 11 11 11 11 11 10 10 10 9	100 60 31 44 56 8 36 58 74 78 82 3 5 6 32 42 95 9 27 40 41 66	CSLYQLENY DLQVGQVEL CGSHLVEAL GERGFFYTP REAEDLQVG LPLLALLAL VEALYLVCG AEDLQVGQV AGSLQPLAL QPLALEGSL LEGSLQKRG LWMRLIPLL MRLIPLLAL RLIPLLALL GSHLVEALY VCGERGFFY CCTSICSLY PLLALLALW NQHLCGSHL YLVCGERGF LVCGERGFF LVCGERGFF	15 14 13 13 12 12 12 12 12 12 11 11 11 11 11 10 10 10	5 59 7 73 1 2 4 32 33 68 93 26 28 30 52 71 16 21 77 40 54	MRLIPLIALL EDLQVGQVEL LIPLIALLIAL GAGSLQPLAL MALWMRLIPL ALWMRLIPLI WMRLIPLIAL GSHLVEALYLV LGGGPGAGSL EQCCTSICSL VNQHLCGSHLVEAL PKTRREAEDL GPGAGSLQPL CTSICSLYQL LWGPDPAAAF PAAAFVNQHL LQPLALEGSL YLVCGERGFF TRREAEDLQV	19 14 13 12 12 12 12 12 12 12 11 11 11 11 11 10 10 10 9 9

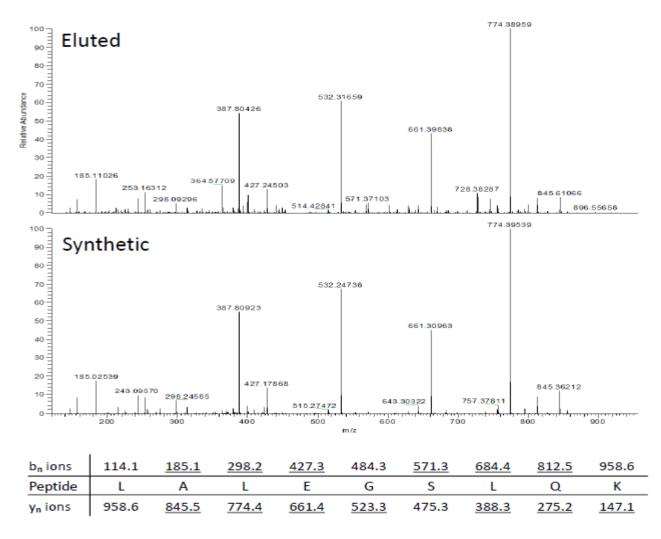
Supplementary Figure S1. Representative Expression of HLA class I and secretion of proinsulin. Expression of HLA class I and secretion of proinsulin by the generated K562-PPI (dashed line) and K562-PPI-HLA (solid line) cell lines (isotype control grey shaded). Single cell cloning was performed to select for the K562-PPI-HLA cell line with best combination of HLA class I expression (left) and Proinsulin expression (right). Data is representative, with B*3906 shown, and similar data obtained for A*11:01, B*1801 and B*38:01.



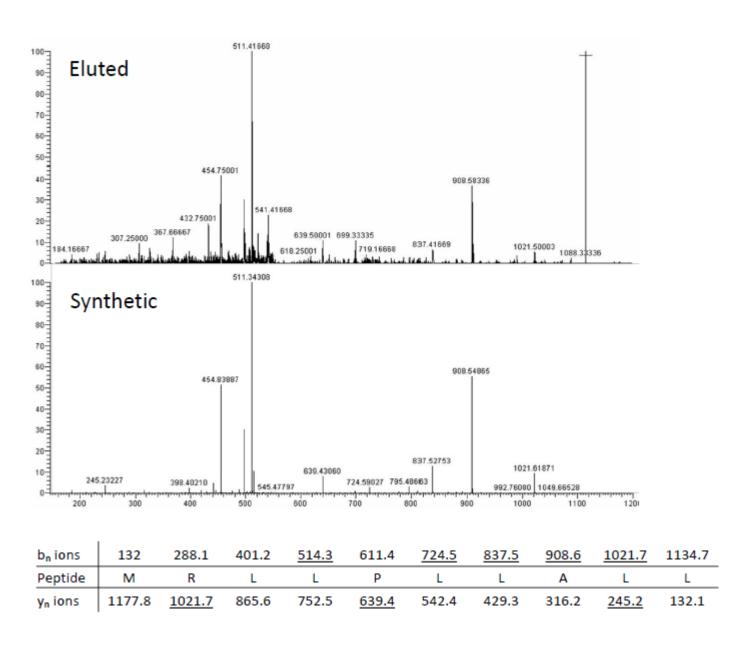


Supplementary Figure S2. PPI epitope discovery Tandem mass spectrometry analysis of collision-induced dissociation revealing the tandem mass spectrum of a PPI peptide (A) HLA-A*1101 Epitope PPI₈₀₋₈₈ (LALEGSLQQK), (B) HLA-B*3801 Epitope PPI₅₋₁₄ (MRLLPLLALL) and (C) HLA-B*3801 Epitope PPI₃₃₋₄₁ (SHLVEALYL). The correct identity of the peptide was proven by tandem mass spectrometry of the synthetic compound. The table lists the amino acid sequence of the peptide with the expected b- and y-fragment ions (fragment ions extending from the amino- and carboxyl terminus respectively). Observed fragment ions are underlined.

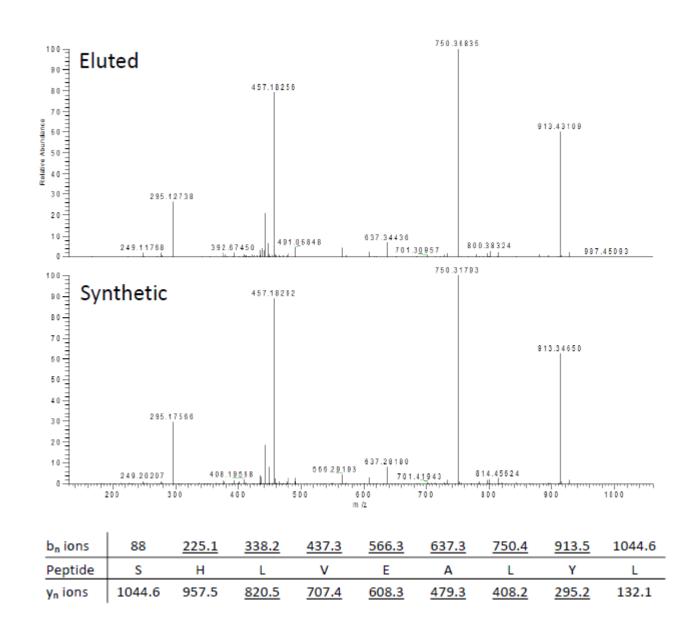
Supplementary Figure 2 A) HLA-A*1101 Epitope PPI₈₀₋₈₈



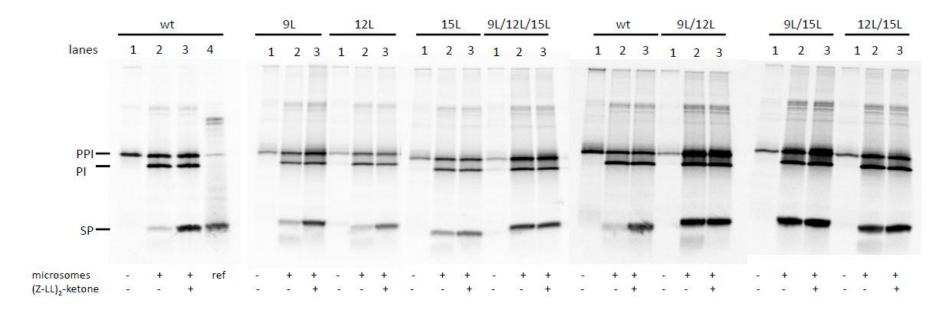
Supplementary Figure 2 B) HLA-B*3801 Epitope PPI₅₋₁₄



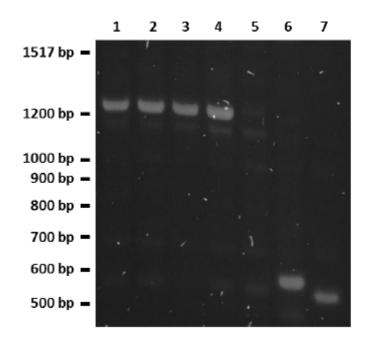
Supplementary Figure 2 C) HLA-B*3801 Epitope PPI₃₃₋₄₁



Supplementary Figure S3. Processing of the PPI signal peptide in microsomes. *In vitro* translation of wt PPI mRNA or mutant PPI mRNA (P9L, A12L, A15L) in the absence (lanes 1) or in the presence of ER-derived microsomes (lanes 2 and 3) and SPP inhibitor (Z-LL)₂-ketone (lanes 3). Microsomes were isolated and analyzed by SDS-PAGE, and radiolabeled proteins visualized by phosphorimaging. Lane 4, *in vitro*-translated reference. Images are representative images of n=2, apart from wt where n=5. Equal translocation efficiency and PPI precursor availability for processing was controlled by comparing the amount of proinsulin between conditions.

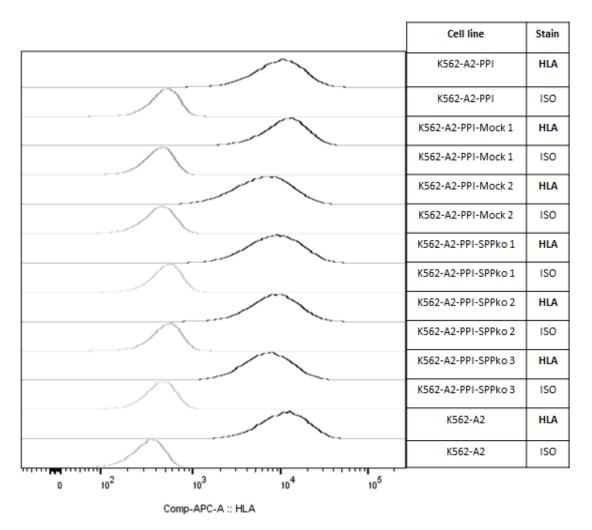


Supplementary Figure S4. Validation of SPP knockout (truncation) by PCR. Within the K562-A2-PPI cells line the SPP gene was truncated using CRISPR-Cas9 directed double targeting of each exon at the start and end of the gene. Simultaneous targeting leads to truncation of the gene in effect leading to functional gene knockout. Length of the gene was assessed using PCR with SPP specific primers on cDNA generated from isolated mRNA from the cell lines. WT (lane 1 and 2) and mock transfected (lane 3 and 4) cell lines harbour full length SPP, whereas reduction (lane 6 and 7) in size or abrogation (lane 5) was observed for cell lines with effective CRISPR-Cas9 targeting of SPP.

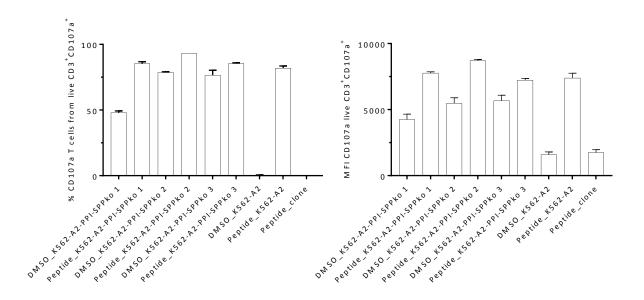


Lane	Cell line
1	K562-A2
2	K562-A2-PPI
3	K562-A2-PPI-Mock 1
4	K562-A2-PPI-Mock 2
5	K562-A2-PPI-SPPko 1
6	K562-A2-PPI-SPPko 2
7	K562-A2-PPI-SPPko 3

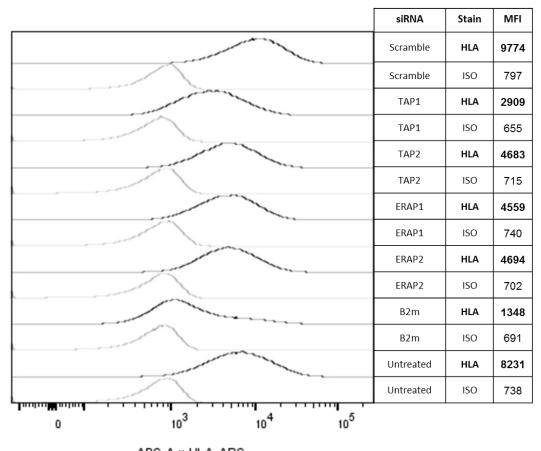
Supplementary Figure S5. HLA class I surface expression on Mock and SPPko cell lines. Cell lines were stained with anti-pan HLA class I antibody (HLA, black line) and corresponding isotype control (ISO, grey line) antibody. Levels of surface HLA expression are similar for each experimental condition (WT, mock and SPP knockout).



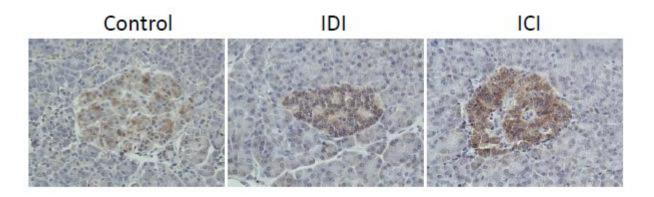
Supplementary Figure S6. Pulsing SPPko cells with cognate peptide rescues their phenotype. K562-A2 and K562-A2-PPI-SPPko cell lines were pulsed for 1 hour with 10uM/ml peptide or peptide diluent (DMSO) prior to co-culture with T cell clone specific for PPI₁₅₋₂₄-HLA-A0201. T cell activation was assessed by CD107a expression. For each SPPko cell line, peptide pulsing leads to increased T cell activation as evidenced by both increases in percentage expression of CD107a (left) and median fluorescence intensity of CD107a (right) with comparable levels in the pulsed WT K562-A2 cell line. Three independent experiments.



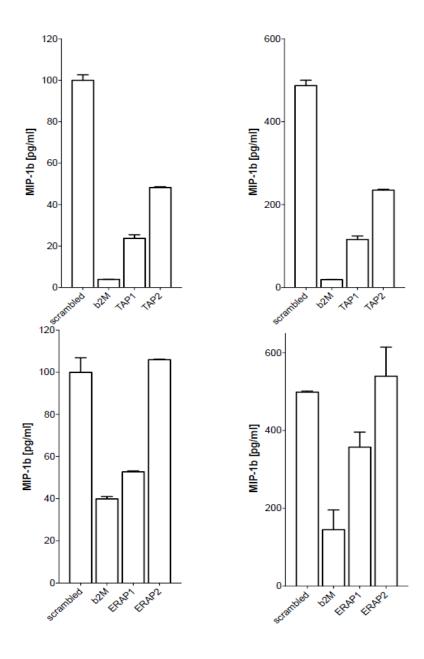
Supplementary Figure S7. HLA class I expression in siRNA experiments. Surface HLA-ABC (W6/32 clone) expression on b2M, ERAP1, ERAP2, TAP1, TAP2, scramble knockdown treated and untreated K562-A24-PPI respectively (HLA, black line) and isotype control (ISO, grey line). Median Fluorescence Intensity (MFI) of HLA staining is shown.



Supplementary Figure S8. Representative immunohistochemistry staining of ERAP. Representative immunohistochemistry staining of ERAP on pancreas samples from patients with type 1 diabetes (n=3) and a representative control sample. ERAP is expressed in all islet cells irrespective of sample source and pancreas samples from type 1 diabetes patients show similar expression in insulin-deficient islets (IDI) and insulin-containing islets (ICI). These data suggest that ERAP expression is not significantly altered in the islets of patients with type 1 diabetes.



Supplementary Figure S9. MIP-1 β in siRNA experiments. MIP-1 β production of HLA-A2402-restricted PPI₃₋₁₁ specific CD8 T cell clone 4C6 upon coculture with θ 2M, TAP1, TAP2, genes (top panel) and ERAP1, ERAP2 gene (bottom panel) knockdown in K562-A24-PPI cells. Two independent experiments.



References

- 1. Eichmann M, de Ru A, van Veelen PA, Peakman M, Kronenberg-Versteeg D: Identification and characterisation of peptide binding motifs of six autoimmune disease-associated human leukocyte antigen-class I molecules including HLA-B*39:06. Tissue antigens 2014;84:378-388
- 2. The UniProt C: UniProt: the universal protein knowledgebase. Nucleic Acids Res 2017;45:D158-D169
- 3. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S: SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 1999;50:213-219