Title: HLA DR2b-binding peptides from human endogenous retrovirus envelope, Epstein-Barr virus and brain proteins in the context of molecular mimicry in multiple sclerosis.

Authors: Ranjan Ramasamy[1*], Fiyaz Mohammed[2], Ute-C. Meier[3]

[1] ID-FISH Technology Inc., 556 Gibraltar Drive, Miltipas, CA 95035, United States of America

[2] Cancer Immunology and Immunotherapy Centre, Institute of Immunology and Immunotherapy, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

[3]Department of Neuroscience and Trauma, Blizard Institute, 4 Newark St, Whitechapel, London E1 2AT, United Kingdom

[*] Corresponding author (RR)

Email: rjr200911@yahoo.com

Running Title: Molecular mimicry in multiple sclerosis

Abbreviations: ABP - α, β Crystallin; βSYN - β Synuclein; BLAST - Basic Local Alignment Search Tool; CNS – Central Nervous System; EAE- Experimental Autoimmune Encephalomyelitis; EBV – Epstein Barr Virus; EBNA1 - Epstein-Barr nuclear antigen 1; env – envelope; HERV – Human Endogenous Retrovirus; IEDB – Immune Epitope Data Base; MAG - Myelin-associated glycoprotein; MBP – Myelin Basic Protein; MOG - Myelin Oligodendrocyte Glycoprotein; MS – Multiple Sclerosis; MSRV - Multiple Sclerosis Associated Retrovirus; NCBI – National Center for Biotechnology Information; OSP – Oligodendrocyte Specific Protein; PLP - Proteolipid Protein; SMM - Stabilised Matrix Method; SYN1 - Syncytin-1; SYN2 - Syncytin-2; TCR – T cell receptor.

Abstract

The aetiology of multiple sclerosis (MS) is as yet poorly understood. Multiple mechanisms in different disease stages are responsible for immunopathology in MS. HLA Class II DR2b (DRB1*1501 β, DRA1*0101 α) is the strongest genetic risk factor for MS. Remnants of ancient retroviruses in the human genome, termed human endogenous retroviruses (HERV), and Epstein-Barr virus (EBV) infection are also associated with MS. *In silico* analyses of human endogenous retroviral envelope (HERV env) proteins and three myelin proteins that are principal targets of an autoimmune response in MS showed sequence similarities between potential T_H epitopes within pairs of viral and myelin peptides predicted to bind HLA DR2b. This led to the proposal that such molecular mimicry may potentially trigger MS. HLA DR2b binding characteristics of previously identified peptides from the three myelin proteins and HERV env proteins as well as additional *in silico* predicted peptides from other encephalitogenic brain proteins and EBV proteins were studied to further investigate molecular mimicry. Peptides containing potential T_H epitopes from the myelin oligodendrocyte glycoprotein and HERV env previously predicted to bind HLA DR2b as well as other pertinent potential HLA DR2b-restricted T_H epitopes were confirmed to bind HLA DR2b molecules. Molecular modelling of HLA DR2b in complex with high affinity peptides derived from MOG and HERV env proteins showed that their binding could occur in a similar manner to a HLA DR2b-binding peptide containing a known T_H epitope. A structurally related pair of peptides predicted to bind HLA DR2b from the EBV protein EBNA1 and β synuclein, a brain protein implicated in MS, were also shown to similarly bind HLA DR2b. The findings justify investigating CD4+ T cell responses to the identified peptides.

Key Words: autoimmunity; Epstein-Barr virus; HLA DR2b-peptide complex; human endogenous retroviruses; molecular mimicry; multiple sclerosis.

1. Introduction

Multiple sclerosis (MS) is an inflammatory autoimmune disease of the central nervous system (CNS) that involves progressive damage to the myelin sheath and axons leading to neurodegeneration [1 - 3]. Studies on MS patients and experimental allergic (autoimmune) encephalomyelitis (EAE) in rodents have implicated several CNS proteins, prominently myelin basic protein (MBP), myelin oligodendrocyte glycoprotein (MOG) and proteolipid protein (PLP), as targets of an autoimmune response in MS [1 - 3]. However, the aetiology of MS is not well understood. Cells of the innate immune system, $CD4^+$ helper T cells (T_H), $CD8^+$ cytotoxic T cells and antibodies are involved in the immunopathology of MS, while T_H among all types of antigen-specific cells, are considered to have the critical role in initiating an autoimmune process [1 - 3]. Environmental factors, *e.g.* vitamin D deficiency [4], and infection with Epstein Barr virus (EBV) [1-2, 5-8], have been implicated in predisposition to MS. Genome-wide association studies identified the HLA Class II allele DRB1*1501 β chain variant, which pairs with the relatively invariant DRA1*0101 α chain to form the HLA DR2b heterodimer in antigen-presenting cells (APCs), as the strongest genetic risk factor for MS [9]. The production of virions and expression of envelope protein (env) of a member of the genome-encoded human endogenous retrovirus W-family (HERV-W), termed the MS-associated retrovirus or MSRV [10], has also been implicated in MS [11 - 15]. However, the molecular mechanisms linking T_H cells to genetic elements in the aetiology of MS are not established. A molecular mimicry hypothesis has been advanced that epitopes in MSRV and possibly other HERV family env proteins that cross-react with epitopes in myelin proteins, and presented by HLA DR2b on APCs to T_H cells in an inflammatory milieu, provide the requisite link [16].

Regions of amino acid sequence homology have been demonstrated by BLAST analysis between MBP, MOG and PLP on one hand and MSRV env on the other [16, 17]. In addition, regions with amino acid sequence homologies are found between the myelin proteins and syncytin-1 (SYN1) [16], another HERV-W family-derived env protein that has

evolved to perform an essential role in forming the syncytiotrophoblast of the placenta [18]. SYN1 is 87% identical in amino acid sequence to MSRV env [16] and also more distantly related to syncytin-2 (SYN2), another essential fusogenic placental protein derived from a different HERV family termed HERV-FRD [19]. SYN2 also possesses regions of amino acid sequence homology with the three myelin proteins [16]. SYN1 has an additional fusogenic role in the development of myotubes from myoblasts [20] and possibly osteoclasts [21].

In silico analyses utilizing the Immune Epitope Data Base (IEDB) [22] to predict HLA DR2b-binding 15mer peptides showed amino acid sequence similarities between potential nonamer T_H epitopes within the 15mers from the HERV env proteins and all three myelin proteins that are predicted to bind to HLA DR2b with high (IC_{50}<50nM) or intermediate affinity (50nM<IC_{50}<500nM) [16]. Sequence similarities between a potential nonamer epitope in MOG and those in MSRV env, SYN1 and SYN2 were particularly prominent [16]. Interestingly, some predicted higher affinity HLA DR2b-binding peptides with sequence similarities lie within longer regions of sequence homology between myelin proteins and HERV env proteins whilst others do not [16]. Since SYN1 and SYN2 have evolved to perform essential physiological functions in humans, it is possible that T_H cells that react with them may be deleted in the thymus and/or regulatory T cells (T_{regs}) that dampen an immune response are selected against them. This may not apply to MSRV env which is not expected to be normally expressed during development. However MSRV env is expressed within innate immune cells in an inflammatory situation, *e.g.* during EBV infection [23], and is a potent stimulant of Toll-like receptor 4 present on macrophages and microglia [12, 24], leading to neuronal damage [12] and impaired functional maturation of myelin-producing oligodendrocytes [24]. While existing data are compatible with an initiating role for molecular mimicry between the MSRV env and myelin proteins in MS, it is unclear whether this extends to the related SYN1 and SYN2 molecules. Once MS has been initiated in the proposed manner [16], further damage could arise from T_H cells recognising other myelin

104 epitopes presented by different HLA Class II molecules as a result of epitope spreading [16,
105 25].

106 This study experimentally investigated HLA DR2b binding of 15mer peptides
107 derived from MBP, MOG, PLP and HERV env proteins earlier identified *in silico* as
108 potentially able to bind to HLA DR2b [16]. It also examined HLA DR2b-binding of peptides
109 from additional CNS proteins reported to be encephalitogenic [1] that had sequence
110 similarities to corresponding peptides present in HERV env proteins or EBV proteins that
111 elicit prominent human $CD4^+$ T cell responses [2]. Pertinent peptides that bound HLA DR2b
112 were also examined by molecular modelling of peptide–HLA DR2b complexes.

113

114 **2. Materials and Methods**

115 **2.1 Selection of CNS proteins for investigation**

116 The three myelin proteins previously used for *in silico* analysis of peptides
117 capable of binding to HLA DR2b [16] and six other CNS proteins reported to be
118 encephalitogenic [1, 26] selected for the present study are listed in Table 1.

119 **Table 1. CNS proteins selected for investigation**

Protein	**Abbreviation**	**NCBI sequence ID**
Myelin basic protein	MBP	P02686.3
Myelin oligodendrocyte glycoprotein	MOG	Q16653.2
Phospholipid protein	PLP	P60201.2
α, β Crystallin	ABP	ACP18852
Myelin-associated oligodendrocyte basic protein	MOPB	NP_001265251.1
Oligodendrocyte-specific protein	OSP	AAC25187
2’3’ Cyclic nucleotide 3’ phosphodiesterase	CNPase	P09543
Myelin-associated glycoprotein	MAG	AAH53347.1

β-Synuclein	βSYN	Q16143

Legend to Table 3. NCBI – National Centre for Biological Information, Bethesda, MD.

2.2 Selection of HERV and EBV proteins for investigation

The three HERV env proteins used previously for predicting HLA DR2b-restricted peptides through the IEDB *in silico* procedure [16] and six EBV proteins reported to elicit strong human $CD4^+$ T cell responses [2] were initially selected for the present study (Table 2).

Table 2. Virus-derived proteins selected for investigation

Virus protein	Abbreviation	NCBI sequence ID
HERV-W Syncytin-1	SYN1	Q9UQF0
HERV-FRD Syncytin-2	SYN2	NP_997465
HERV-W Multiple sclerosis-associated retrovirus envelope protein	MSRV env	AAK18189.1
Epstein-Barr nuclear antigen 1	EBNA1	YP_401677.1
Epstein-Barr nuclear antigen 2	EBNA2	ALV83014.1
Epstein-Barr nuclear antigen 3C	EBNA3C	CEQ33769.1
Epstein-Barr virus transactivator BZLF1	BZLF1	CAD53423
Epstein-Barr virus glycoprotein BZLF2	BZLF2	CEQ33770.1
Epstein-Barr virus envelope glycoprotein H	BXLF2	ATE89094.1

Legend to Table 3. NCBI – National Centre for Biological Information, Bethesda, MD.

2.3 Sequence homologies between CNS and virus-derived proteins

The additional CNS coding sequences obtained from the US National Center for Biotechnology Information (NCBI) data base were compared by pairwise Basic Local Alignment Search Tool (BLASTp) analysis online using default parameters

(https://www.ncbi.nlm.nih.gov/blast) against HERV env proteins, as previously described for MBP, MOG, PLP and the three HERV env proteins SYN1, SYN2 and MSRV env [16]. Each of the CNS proteins shown in Table 1 were also individually tested in BLASTp searches for regions of homology against all non-redundant protein sequences of EBV (human herpes virus 4 strain B95-8) with NCBI taxonomy ID 10377.

2.4 Prediction of peptides potentially binding to HLA DR2b molecules

Prediction of potential peptides binding to HLA DR2b molecules was performed as previously described [16] using the IEDB analysis resource (www.iedb.org) [22, 27, 28]. The default peptide length of 15 amino acids was used in the analysis but the results also show the core nonamers that are expected to bind to the HLA DR2b molecule and constitute the major portion of the T cell epitope [22, 27, 28]. The Stabilised Matrix Method (SMM) was used to rank the peptides according to their predicted binding affinities or IC_{50} which indicates the nM concentration of peptide expected to bind and achieve 50% saturation of the HLA DR2b molecules [22, 27, 28]. Sequence similarities between core nonamer sequences in 15mer peptides that were predicted to bind HLA DR2b with high or intermediate affinity (IC_{50}<500nM) were manually identified within pairs of proteins.

2.5 Determination of the binding affinity and stability of HLA DR2b-peptide complexes

Peptides (15mers) were synthesized by FMOC solid-phase chemistry and quality checked with matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) by ProImmune (Oxford, UK). Binding characteristics of the peptides to HLA DR2b were then determined by ProImmune using the cell-free REVEAL® MHC class II binding assay [29]. The REVEAL® assay measured the ability of a peptide to stabilize the MHC-peptide complex based on detection of the native conformation of the MHC-peptide complex with a specific monoclonal antibody [29]. After an initial incubation with peptide for

160 determining the proportion of MHC molecules binding the peptide to measure affinity, an
161 additional measurement was taken after a further 24h incubation at 37^0C to measure
162 stability. The stability provides information on whether peptide can be bound long enough to
163 serve as a T cell epitope. Affinity and stability indices were measured as a percentage of the
164 signal generated by the test peptide in comparison to a proprietary ProImmune positive
165 control peptide. A well characterised 15mer MBP peptide with the sequence
166 ENPVVHFFKNIVTPR (hereafter referred to as MBP_3) that is presented by HLA DR2b and
167 activates $CD4^+$ T cells [30] was chosen as the internal comparative standard in the assays.

168 Details of the two sets of 40 peptides from CNS and viral proteins that were
169 tested in the HLA DR2b binding assays are provided in Supplementary Table S1. The first
170 set of 40 contained sequence-related peptides derived from MBP (including the control
171 peptide MBP_3), MOG, PLP, SYN1 and MSRV env previously identified *in silico* using the
172 IEDB algorithm as being potentially important for molecular mimicry [16]. Staggered arrays
173 of 15mers were used to identify the best binding peptide in the REVEAL® assay. The first set
174 also contained a 15mer derived from EBV DNA polymerase shown to cross-react at the
175 $CD4^+$ T cell level with the control peptide MBP_3 on presentation by HLA DR2b [30]. A
176 different HLA DR2b-restricted MOG epitope shown previously to stimulate $CD4^+$ T cells to
177 produce IFNγ [31] was also included in the first set of peptides. Others in the first set were
178 four MSRV env 15mer peptides with the nonamer sequence TSVLVGPLV that exhibited
179 weaker sequence homology to MOG nonamer IVLPVLGPLV [16].

180 The second set of 40 peptides (Supplementary Table S1) were chosen to
181 replicate and further examine the binding characteristics of the more promising HLA DR2b-
182 binding peptides identified from first set. They were independently synthesised and tested in
183 REVEAL® binding assays. The second set additionally tested sequence-related pairs of
184 HLA DR2b-binding peptides identified through IEDB *in silico* analysis in EBNA1 and HERV
185 env proteins on one hand and different CNS proteins on the other. They included peptide
186 pairs from EBNA1 and βSYN as well as EBNA1 and OSP that had also been predicted to

bind HLA DR2b using a different algorithm in an unrelated study [32], and peptide pairs from MSRV env and OSP. The second set also included a different MBP peptide reported to be recognised by T_H cells in the context of HLA DR2b [33].

2.6 Modelling of 15mer peptides binding to HLA DR2b

Molecular modelling of the HLA DR2b-peptide complexes were performed using the *in silico* docking program HADDOCK (high ambiguity driven protein-protein docking) [34]. Coordinates for the HLA DR2b complex were retrieved from the Protein Data Bank entry 1YMM [35]. Initial coordinates for the DR2b-restricted peptide moieties were extracted from the crystal structure of the T cell receptor(TCR)/HLA DR2b/MBP_3-peptide complex (entry 1YMM), and then used to build models of peptides with the molecular builder tool in COOT [36]. Each HLA DR2b-restricted peptide was subsequently subjected to a short regularisation protocol to ensure that the geometry of the peptide residues conformed to known bond lengths and angles.

The docking procedure was driven using only ambiguous intermolecular restraints, which were defined based on previously determined HLA DR2b-peptide complexes [35, 37, 38]. These structures revealed that the MBP_3 peptide is bound in the HLA DR2b peptide-binding groove with peptide side chains P1, P4, P6 and P9 occupying pockets within the groove. Hence residues that line the P1, P4, P6 and P9 pockets of HLA DR2b were selected as active residues (comprised of E11α, F24α, F32α, W43α, F54α, N62α, D66α, R76α, R13β, F26β, D28β, Q70β, A71β, Y78β, D57β and W61β). For the peptide only the anchor residue side chains at P1, P4, P6 and P9 were defined as active residues. Passively involved residues were selected automatically. The 200 structures obtained after water refinement were analysed and ranked according to their HADDOCK score, a weighted sum of electrostatic, van der Waals, and restraint energy terms [34]. The

lowest energy structure solutions were visualised and analysed using Pymol (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC).

3. Results

3.1 Sequence homologies between CNS and EBV or HERV env proteins

Regions of sequence homologies between the three HERV env proteins and the three myelin proteins MBP, MOG and PLP observed in BLASTp analysis have been previously described [16]. BLASTp analysis of each of the selected CNS proteins against all the non-redundant protein sequences coded in whole EBV genome revealed only a single region of weak homology between α, β crystallin (ABP) and a 53 residue segment of the EBV protein EBNA4 (NCBI Protein ID P03203.3) with an E value of 0.95 (Supplementary Table S2). Pairwise BLASTp analysis of each of the six newly selected CNS proteins against the three HERV env proteins demonstrated regions of homology with E≤0.5 only between the pairs ABP and SYN2, ABP and MSRV env, and myelin-associated glycoprotein (MAG) and MSRV env (Supplementary Table S3).

3.2 Structurally related peptides in CNS and EBV or HERV env proteins predicted to bind to HLA DR2b molecules

IEDB analysis of MBP, MOG, PLP on one hand and the three HERV env proteins on the other, that identified 15mer peptides containing sequence-related nonamers predicted to bind to HLA DR2b were previously described [16]. Similar IEDB analysis was performed on the additional CNS proteins ABP, MAG, OSP and βSYN (Supplementary Table S4,) and on the EBV proteins EBNA1 and EBNA4 (Supplementary Table S5). ABP, MAG and EBNA4 were selected because of relevant regions of sequence homology identified in section 3.1, while EBNA1, OSP and βSYN were chosen because of the

prediction of pertinent HLA DR2b binding peptides in an independent study [32]. Examination of the IEDB results from the protein pairs ABP/SYN2, ABP/MSRV env and MAG/MSRV env, using data in Table S4 and published data in reference 16 did not identify 15mer peptides of potentially high or intermediate affinity of binding to HLA DR2b that also contained sequence-related nonamers. Nonamers of similar sequences were however found in predicted HLA DR2b-binding 15mers from the pair OSP/MSRV env. Analysis of IEDB results in Supplementary Tables S4 & S5 showed sequence-related pairs of potential HLA DR2b-binding peptides of high or intermediate affinity also in EBNA1/OSP and EBNA1/βSYN. The pairs of peptides containing structurally related, predicted HLA DR2b-binding nonamers, with their sequences shown in Supplementary Table S6, were subsequently investigated in HLA DR2b binding assays.

3.3 Experimental binding to HLA DR2b of CNS and viral peptides predicted *in silico* to bind HLA DR2b

The results of REVEAL binding assays on the selected peptides (Supplementary Table S6) showed that the pairs of peptides from MOG and the corresponding three HERV env proteins containing sequence-related nonamers previously predicted to engage HLA DR2b [16], and implicated in molecular mimicry, are able to bind HLA DR2b with comparable affinities and stabilities to the control MBP_3 peptide. A peptide from βSYN containing a predicted HLA DR2b-binding nonamer of sequence GVLYVGSKT and two similarly predicted peptides from EBNA1 with the related nonamer sequence VFVYGGSKT also bound to HLA DR2b with binding characteristics comparable to MBP_3 (Supplementary Table S6).

The results also show that some OSP peptides with similar nonamer sequences to EBNA1 and MSRV env peptides and with predicted *in silico* intermediate binding affinity are able to bind well to HLA DR2b. However, the corresponding 15mer viral peptides did not

bind strongly to HLA DR2b. For example, the OSP 15mer STTLRALAPRLMRRV which bound strongly had five identities in its predicted nonamer HLA DR2b-binding sequence (LRALAPRLM) to the corresponding nonamer (LRALLARSH) in two 15mer EBNA1 peptides that however only showed weak binding to HLA DR2b (Supplementary Table S6).

Peptides from the closely related signal sequences of SYN1 and MSRV that contained sequence-related nonamers to those in internal peptides of MBP (including the control peptide MBP_3) and PLP identified previously [16] did not bind strongly to HLA DR2b in the assays. Only one PLP peptide TASFFFLYGALLLAE that contained the nonamer sequence FFFLYGALL that was predicted to bind strongly to HLA DR2b [16] was confirmed to bind strongly to HLA DR2b. Four MSRV env peptides tested containing the nonamer sequence TSVLVGPLV with weaker sequence homology to the MOG nonamer IVLPVLGPLV did not bind to HLA DR2b.

Peptides from EBV DNA polymerase and a different MOG region that had been shown to be presented on HLA DR2b and stimulate $CD4^+$ T cells [30, 31] revealed significant binding affinity to HLA DR2b in the assay. A MBP peptide (GTLSKIFKLGGRDSR) containing a putative HLA DR2b-resticted T cell epitope [33] but with a weak predicted IC_{50} of 940nM based on IEDB analysis, only demonstrated marginal binding to HLA DR2b.

An exact correlation between the *in silico* predicted affinity (IC_{50}) and the experimentally determined affinity by the REVEAL® binding assay for HLA DR2b was not observed. For example, some PLP peptides with high predicted affinity (IC_{50}<1nM) showed poor experimental binding, while four SYN2 peptides with predicted intermediate affinities (IC_{50} of 130 to 149nM) had experimental binding comparable to MBP_3 (Supplementary Table S6).

Data on the best binding 15mer peptides with nonamers relevant for molecular mimicry, grouped together and compared with the binding of the control MBP peptide are listed in Table 3.

Table 3. Binding characteristics and Haddock scores of the best pairs of HLA DR2b-binding 15mer peptides containing sequence-related nonamers relevant to molecular mimicry

Homology Group	Peptide	Peptide Sequence	Relative affinity	Relative stability	HADDOCK model score
1. MOG & HERV env	MOG_4	ITLFV**IVPVLGPLV**A	151	110	-123.7±2.5
	MSRV env _5	MPW**TLPFLGPLA**AII	69	33	-158.1±2.1
	SYN1_2	MPW**ILPFLGPLA**AII	144	124	-152.5±4.4
	SYN2_5	KWFSW**VLPLTGPLV**S	348	181	-137.8±5.3
2. βSYN & EBNA1	β synuclein	EKTKE**GVLYVGSKT**R	91	95	-128.7±3.0
	EBNA1_2	VAG**VFVYGGSKT**SLY	118	43	-131.5±5.1
3. Control	MBP_3	ENPV**VHFFKNIVT**PR	100	100	-163.7±1.6

Legend to Table 3. Results show the experimentally determined relative affinity and stability of binding of peptides expressed as a percentage of that observed with the control MBP_3 peptide assigned values of 100. The nonamers sequence predicted to bind in the peptide-binding groove in HLA DR2b are shown in bold letters and underlined. The docking scores for the HADDOCK-derived lowest energy HLA DR2b-peptide complex models are shown.

3.4 Molecular models of sequence-related peptides binding to HLA DR2b

We employed *in silico* molecular docking strategies to understand the molecular mechanisms governing binding of the structurally related pairs of peptides by HLA DR2b and their potential recognition by TCR. To evaluate the feasibility of using such approaches we first modelled the binding of the control MBP_3 peptide ENPVVHFFKNIVTPR to HLA DR2b using HADDOCK and then compared with the available crystallographic structure (PDB entry 1YMM) [35, 37, 38]. The HLA DR2b-MBP_3 complex model corresponding to the lowest intermolecular energy (with a HADDOCK score of -163.7) shows substantial similarity with the published structure in terms of epitope conformation and docking mode (Figure 1A). Superposition of the MBP_3 peptides derived from the published and model complex

structures show that the main chain conformation is highly conserved (Figure 1A). In addition, similar to the published structure, the modelled MBP-3 peptide side chains at P1, P4, P6 and P9 serve as anchors slotting into the HLA DR2b antigen binding cleft (Figure 1B&C). Finally, in both the published and modelled complexes, the peptide was held in the HLA DR2b antigen-binding cleft by a conserved network of hydrogen bonding and non-polar interactions (Figure 1B&C). These observations justified the use of the HADDOCK docking approach to generate models of HLA DR2b bound to peptides that are relevant to MS.

To address the molecular mimicry hypothesis we generated models of HLA DR2b in complex with peptides of the highest affinity derived from MOG and the HERV env proteins MRSV env, SYN1 and SYN2 that are shown in Table 3. Superposition of the MOG, MSRV env, SYN1 and SYN2 peptides show that they all adopt a very similar back-bone conformation (Figure 2A). Similarly to the control MBP_3 peptide, the P1, P4, P6 and P9 peptide side chain positions serve as anchors inserting into the HLA DR2b antigen binding cleft (Figure 2B-D). The HLA DR2b-peptide interactions were remarkably conserved between the different complexes including the control HLA DR2b-MBP_3 complex. In addition, positions P-1, P2, P5, and P8 are predicted to be surface exposed in the HADDOCK derived HLA DR2b-peptide complex models, and therefore potentially involved in binding to the TCR. The chemical characteristics of these prominent solvent exposed residues were either identical or structurally related in the relevant pairs of peptides. Taken together, these findings support the molecular mimicry hypothesis between MOG and HERV env proteins in MS.

HADDOCK derived models of HLA DR2b in complex with the βSYN and EBNA1 peptides shown in Table 3 were also generated (Figure 3). These peptides adopted similar main chain conformations (Figure 3A) and mediated a conserved network of polar and non-polar interactions with side chains of HLA DR2b (Figure 3 B&C). As with comparisons between MOG and HERV env proteins, the most prominent surface exposed residues (at P-

1, P2, P5, and P8) and hence potential TCR contacts were mainly conserved or semi-conservatively substituted between the βSYN and EBNA1 peptide pair.

It is noteworthy that surface-exposed residues that can contact TCRs were however significantly different between the two sets of unrelated peptide pairs βSYN/EBNA1 and HERV env/MOG, and between each of these and MBP_3 (Table 3 and Figures 1-3).

The HADDOCK docking scores of the best binding 15mer peptides possessing the relevant sequence-related nonamer pairs are listed in Table 3. The HADDOCK scores do not correlate with experimentally measured REVEAL® binding affinities or stability indices for the peptides but the high negative values point towards energetically favourable binding to HLA DR2b molecules.

4. Discussion

More recent findings are pertinent to the originally proposed HERV-related molecular mimicry hypothesis [16]. EBV, which primarily infects B cells, has been further implicated as a necessary but not sufficient cause of MS, partly because of its increased and dysregulated expression in peripheral blood and brain [39 -43]. In addition, antibody titres to EBNA1 have lately been re-confirmed to be higher in MS patients compared to controls [44]. EBNA1 has recently been reported to promote alternative splicing of cellular genes [45]. Since EBNA1 is widely expressed in EBV infected cells [46], it is intriguing to speculate that its splicing activity has a role in the *trans* splicing that has been postulated to produce functional MSRV env molecules [17]. This adds to the many different mechanisms proposed to explain why EBV infections are a predisposition for MS [1-2, 5-8].

HERVs and their putative role in autoimmunity have been lately reviewed [47-49] and cross-reactive B cell epitopes in MOG and HERV-W env have been documented [50]. The presence of antibodies to HERV-W env proteins have recently been reported to differentiate MS from related neurological diseases [51, 52]. Recent data also show that

MSRV env is present in microglia associated with myelinated axons in MS lesions, MSRV env induces inflammatory myelin and neuron damaging activity in microglia *in vitro* and that antibodies to MSRV env are neuroprotective in MS patients [12]. These observations are pertinent to further examining a role for potential molecular mimicry between MSRV env and MOG in triggering MS.

Evidence that human GDP-L-fucose synthase peptides are recognised by $CD4^+$ T cells in the context of HLA DRB3 *0202 in MS patients, and that gut bacterial GDP-L-fucose synthase may be cross-reactive has led to a different proposal for molecular mimicry in MS [53]. RAS guanyl releasing protein 2 in peripheral memory B cells driving the proliferation of brain-infiltrating $CD4^+$ T_H1 in a HLA DR2b-restricted manner that then recognise epitopes from the same protein expressed in CNS cells has been proposed as another autoimmune mechanism explaining the association between MS and HLA DR2b [54].

The molecular mimicry hypothesis proposed previously [16] was supported by the *in silico* identification of sequence-related pairs of 15mer peptides predicted to bind HLA DR2b in myelin-associated MBP, MOG and PLP proteins on one hand and HERV env proteins on the other. Sequence homologies were particularly prominent between the predicted MOG and HERV env peptides [16] and the present study confirmed that these predicted peptides are indeed able to bind to HLA DR2b with characteristics comparable with the well-known T_H epitope in MBP_3. Such binding was not demonstrable for the HERV env peptides identified previously that were related in sequence to peptides from MBP and PLP [16]. However the corresponding MBP peptides and one PLP peptide have been shown to bind HLA DR2b.

The peptides containing sequence-related nonamers with potential T_H epitopes in MSRV env, SYN1, SYN2 and MOG have the capacity to bind to HLA DR2b molecules with similar binding topology to the well characterised MBP_3 peptide containing a T_H epitope. The molecular modelling suggests that potential surface exposed residues that

388 contact TCR are relatively conserved between the MOG and HERV env peptides which is
389 consistent with the proposed molecular mimicry hypothesis. The MOG peptide is located in
390 the predicted C terminal transmembrane domain of MOG. The corresponding HLA DR2b-
391 binding nonamers from MSRV env, SYN1 and SYN2 are also sited in predicted
392 transmembrane domains. A longer peptide from the transmembrane region of MOG, that
393 contained the MOG peptide identified in the present work, has independently been shown to
394 stimulate $CD4^+$ T cells from MS patients to proliferate and secrete IFNγ in a HLA DRB-
395 restricted manner [55]. It is possible that SYN1 and SYN2 may normally elicit tolerance as
396 they may be recognised as self-proteins. MSRV env on the other hand may function as a
397 foreign protein that can generate autoimmunity through molecular mimicry in an
398 inflammatory milieu, possibly driven by EBV infection, within the CNS or outside it as
399 previously discussed [16]. Studies on $CD4^+$ T cell responses to the peptides identified in this
400 study will help clarify the potential roles of MOG and the three HERV env proteins in the
401 immunopathogenesis of MS. It is relevant in this context that TCR recognition of MBP_3
402 bound to HLA DR2b has been shown to involve skewed binding, not typical of TCR binding
403 foreign peptide-Class II MHC complexes, which can result in potentially weaker interactions
404 that may permit autoimmune T cells to escape deletion in the thymus [37].

405 This study also identified a pair of sequence-related nonamers derived from
406 βSYN and EBNA1 that showed binding affinity and stability comparable to MBP_3 in the
407 REVEAL® assay for HLA DR2b. This pair of peptides had been independently predicted to
408 bind HLA DR2b [32]. Modelling of the βSYN and EBNA1 peptides with HLA DR2b revealed
409 binding to the peptide binding cleft similar to MBP_3 and relative conservation of the surface
410 exposed, potential TCR contact residues in the two peptides. This suggests the molecular
411 mimicry is possible between the βSYN and EBNA1 peptides. It is relevant in this context that
412 βSYN-reactive T_H cells have recently been proposed to be responsible for autoimmune
413 damage to CNS grey matter in the progressive stage of MS [26]. The possibility that EBNA1
414 generated, βSYN-reactive T_H cells induce additional autoimmune pathology, after the

potential initiation of MS by molecular mimicry between MOG and HERV env proteins, therefore justifies investigation. There is long standing evidence for HLA DR2b-dependent molecular mimicry between MBP_3 and an EBV polymerase peptide [30, 37, 38]. Their HLA DR2b binding was again confirmed but their precise role in the aetiology of MS remains to be elucidated.

Investigations on other pairs of potential HLA DR2b-binding peptides in EBNA1 and a variety of CNS proteins observed in an independent study [32] may also be useful because our study was limited to CNS proteins with high encephalitogenic potential and constrained by the number of peptide pairs that could be tested for HLA DR2b binding.

The HLA DR2a molecule is formed by pairing of the DRB5*0101 β chain variant, whose gene is closely linked to the DRB1*1501 gene in many individuals, with the relatively non-polymorphic DRA1*0101 α chain. The previous *in silico* based predictions failed to identify strong HLA DR2a binding pairs of potential sequence-related T_H cell epitopes in HERV env and myelin proteins MBP, MOG and PLP [16]. However, because of the close genetic linkage of the two β chain loci, investigating potential HLA DR2a-binding T_H epitopes in an extended set of encephalitogenic CNS proteins and EBV or HERV proteins is warranted because HLA DR2a and HLA DR2b molecules bind complementary sets of peptides through different binding motifs [56].

5. Conclusions

The results of the cell free HLA DR2b binding assays and molecular modelling show that sequence-related MOG and HERV env as well as βSYN and EBNA1 peptide pairs, with each set of pairs containing related potential T_H epitopes, are able to bind to HLA DR2b with similar affinity and conformation to a peptide MBP_3 containing an experimentally confirmed T_H epitope. Such pairs of sequence-related peptides are candidates for molecular mimicry in MS. However, definitive support for molecular mimicry will require detailed studies

441 on $CD4^+$ T_H cell responses to the identified peptides. Such investigations may also contribute
442 to the variety of immunomodulatory approaches presently being explored for treating MS
443 [12, 24, 57 - 64].

444

445 **Conflict of interest statement**

446 The authors declare no conflict of interest.

447

448 **Acknowledgements**

449 This research was supported by the US National Multiple Sclerosis Society (PP-
450 1711-29350). FM is funded by the Wellcome Trust grant 099266/Z/12/Z. The authors are
451 grateful to Professor Nick Willcox of the MRC Weatherall Institute of Molecular Medicine at
452 the University of Oxford for his continuing encouragement.

453

454 **Author contributions**

455 RR and UM initiated the project, FM performed the modelling studies, and RR
456 did the IEDB analysis, collation of data and drafting of the manuscript. All authors read and
457 approved the final manuscript.

458

459 **References**

460 1. M. Sospedra, R. Martin. Immunology of multiple sclerosis. Annu. Rev. Immunol. 23
461 (2005) 683–747, doi: 10.1146/annurev.immunol.23.021704.115707.
462 2. G.S. Taylor, H.M. Long, J.M. Brooks, A.B. Rickinson, A.D. Hislop. The immunology of
463 Epstein-Barr virus–induced disease. Annu. Rev. Immunol. 33 (2015) 787-821, doi:
464 10.1146/annurev-immunol-032414-112326.

3. C.A. Dendrou, L. Fugger, M.A. Friese. Immunopathology of multiple sclerosis. Nat. Rev. Immunol. 15 (2015) 545–558.
4. E. Kocovská, F. Gaughran, A. Krivoy, U-C Meier. Vitamin-D deficiency as a potential environmental risk factor in multiple sclerosis, schizophrenia, and autism. Front. Psychiatry 8 (2017) 47, doi: 10.3389/fpsyt.2017.00047.
5. M.F. Cusick, J.E. Libbey, R.S. Fujinami, Multiple sclerosis: autoimmunity and viruses, Curr. Opin. Rheumatol. 25 (2013) 496–501, http://dx.doi.org/10.1097/BOR.0b013e328362004d.
6. J.S. Tzartos, G. Khan, A. Vossenkamper, M. Cruz-Sadaba, S. Lonardi, E. Sefia, et al., Association of innate immune activation with latent Epstein-Barr virus in active MS lesions. Neurology 78 (2012) 15-23, doi: 10.1212/WNL.0b013e31823ed057.
7. S. Sisay, L. Lopez-Lozano, M. Mickunas, A. Quiroga-Fernández, J. Palace, G. Warnes, et al., Untreated relapsing remitting multiple sclerosis patients show antibody production against latent Epstein Barr Virus (EBV) antigens mainly in the periphery and innate immune IL-8 responses preferentially in the CNS. J. Neuroimmunol. 306 (2017) 40-45, doi: 10.1016/j.jneuroim.2017.02.017.
8. S.V. Ramagopalan, R. Dobson, U.C. Meier, G. Giovannoni. Multiple sclerosis: risk factors, prodromes, and potential causal pathways. Lancet Neurol. 9 (2010) 727-739. doi: 10.1016/S1474-4422(10)70094-6.
9. International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control Consortium, S. Sawcer, G. Hellenthal, M. Pirinen, C.C. Spencer, et al., Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis, Nature 476 (2012) 214–219, http://dx.doi.org/10.1038/nature10251.
10. H. Perron, B. Lalande, B. Gratacap, A. Laurent, O. Genoulaz, C. Geny, et al. Isolation of retrovirus from patients with multiple sclerosis. Lancet 337 (1991) 862–863.
11. T. Christensen. Human endogenous retroviruses in the aetiology of MS. Acta. Neurol. Scand. 136 Suppl. 201 (2017) 18-21.
12. D. Kremer, J. Gruchot, V. Weyers, L. Oldemeier, P. Göttle, L. Healy, et al. pHERV-W

493 envelope protein fuels microglial cell dependent damage of myelinated axons in
494 multiple sclerosis. Proc. Natl. Acad. Sci. USA 116 (2019) 15216–15225.
495 doi/10.1073/pnas.1901283116.
496 13. K.K. Nissen, M.J. Laska, B. Hansen, T. Terkelsen, P. Villesen, S. Bahrami, et al.
497 Endogenous retroviruses and multiple sclerosis-new pieces to the puzzle. BMC Neurol.
498 13 (111) (2013), http://dx.doi.org/10.1186/1471-2377-13-111.
499 14. A. Dolei, E. Uleri, G. Ibba, M. Caocci, C. Piu, C. Serra. The aliens inside human DNA:
500 HERV-W/MSRV/syncytin-1 endogenous retroviruses and neurodegeneration. J. Infect.
501 Dev. Ctries. 9 (2015) 577-587, doi: 10.3855/jidc.6916.
502 15. E. Morandi, R.E. Tarlinton, B. Gran, Multiple sclerosis between genetics and infections:
503 human endogenous retroviruses in monocytes and macrophages. Front. Immunol. 6
504 (647) (2015), http://dx.doi.org/10.3389/fimmu.2015.00647.
505 16. R. Ramasamy, B. Joseph, T. Whittall. Potential molecular mimicry between the human
506 endogenous retrovirus W family envelope proteins and myelin proteins in multiple
507 sclerosis. Immunol. Lett. 183 (2017) 79-85. dx.doi.org/10.1016/j.imlet.2017.02.003.
508 17. G.S. do Olival, T.S. Faria, L.H. Nali, A.C. de Oliveira, J. Casseb, J.E. Vidal, et al.
509 Genomic analysis of ERVWE2 locus in patients with multiple sclerosis: absence of
510 genetic association but potential role of human endogenous retrovirus type W elements
511 in molecular mimicry with myelin antigen. Front. Microbiol. 4 (2013)172, doi:
512 10.3389/fmicb.2013.00172.
513 18. S. Mi, X. Lee, X. Li, G.M. Veldman, H. Finnerty, L. Racie, Syncytin-1 is a captive
514 retroviral envelope protein involved in human placental morphogenesis. Nature 403
515 (2000) 785–789.
516 19. A.G. Lokossou, C. Toudic, B. Barbeau. Implication of human endogenous retrovirus
517 envelope proteins in placental functions. Viruses 6 (2014) 4609–4627,
518 http://dx.doi.org/10.3390/v6114609.
519 20. B. Bjerregard, I. Ziomkiewicz, A. Schulz, L.I. Larsson. Syncytin-1 in differentiating
520 human myoblasts: relationship to caveolin-3 and myogenin. Cell Tissue Res. 357

(2014) 355-362, doi: 10.1007/s00441-014-1930-9

21. K. Søe, T.L. Andersen, A.S. Hobolt-Pedersen, B. Bjerregaard, L.I. Larsson, J.M. Delaissé. Involvement of human endogenous retroviral syncytin-1 in human osteoclast fusion. Bone 48 (2011) 837-846.
22. Y. Kim, J. Ponomarenko, Z. Zhu, D. Tamang, P. Wang, J. Greenbaum, et al. Immune epitope database analysis resource. Nucleic Acids Res. 40 (2012) W525-W530, doi: 10.1093/nar/gks438.
23. E. Morandi, R.E. Tarlinton, B. Gran. Multiple sclerosis between genetics and infections: human endogenous retroviruses in monocytes and macrophages, Front. Immunol. 6 (2015) 647, http://dx.doi.org/10.3389/fimmu.2015.00647.
24. A. Madeira, I. Burgelin, H. Perron, F. Curtin, A.B. Lang, R. Faucard. MSRV envelope protein is a potent, endogenous and pathogenic agonist of human toll-like receptor 4: Relevance of GNbAC1 in multiple sclerosis treatment. J. Neuroimmunol. 291 (2016) 29-38, doi: 10.1016/j.jneuroim.2015.12.006
25. K. Raddassi, S.C. Kent, J. Yang, K. Bourcier, E.M. Bradshaw, V. Seyfert-Margolis, et al. Increased frequencies of myelin oligodendrocyte glycoprotein/MHC class II-binding CD4 cells in patients with multiple sclerosis, J. Immunol. 187 (2011) 1039–1046, http://dx.doi.org/10.4049/jimmunol.1001543.
26. D. Lodygin, M. Hermann, N. Schweingruber, C. Flügel-Koch, T. Watanabe, C. Schlosser C, et al. β-Synuclein-reactive T cells induce autoimmune CNS grey matter degeneration. Nature. 566 (2019) 503-508, doi: 10.1038/s41586-019-0964-2.
27. P. Wang, J. Sidney, C. Dow, B. Mothé, A. Sette, B. Peters. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS. Comput. Biol. 4 (2008) e1000048.
28. P. Wang, J. Sidney, Y. Kim, A. Sette, O. Lund, M. Nielsen, B. Peters. Peptide binding predictions for HLA DR, DP and DQ molecules. BMC. Bioinformatics 11 (2010) 568.
29. https://www.proimmune.com/ecommerce/page.php?page=reveal_class2. Accessed 24 May 2019.

549 30. K.W. Wucherpfennig, J.L. Strominger. Molecular mimicry in T cell-mediated
550 autoimmunity: viral peptides activate human T cell clones specific for myelin basic
551 protein. Cell 80 (1995) 695 – 705.
552 31. R. Weissert, J. Kuhle, K.L. de Graaf, W. Wienhold, M.M. Herrmann, C. Müller, et al.
553 High immunogenicity of intracellular myelin oligodendrocyte glycoprotein epitopes. J.
554 Immunol. 169 (2002) 548-556.
555 32. M. Tschochner, S. Leary, D. Cooper, K. Strautins, A. Chopra, H. Clark, et al. Identifying
556 patient-specific Epstein-Barr NuclearAntigen-1 genetic variation and potential
557 autoreactive targets relevant to multiple sclerosis pathogenesis. PLoS. One 11 (2016)
558 e0147567, doi:10.1371/journal.pone.0147567
559 33. M. Pette, K. Fujita, D. Wilkinson, D.M. Altmann, J. Trowsdale, G. Giegerich, et al.
560 Myelin autoreactivity in multiple sclerosis: recognition of myelin basic protein in the
561 context of HLA-DR2 products by T lymphocytes of multiple-sclerosis patients and
562 healthy donors. Proc. Natl. Acad. Sci. USA. 87 (1990) 7968-7672.
563 34. G.C.P. van Zundert, J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastritis, E.
564 Karaca, et al. The HADDOCK2.2 web server: user-friendly integrative modeling of
565 biomolecular complexes. J. Mol. Biol. 428 (2016) 720-725, doi:
566 10.1016/j.jmb.2015.09.014.
567 35. M. Hahn, M.J. Nicholson, J. Pyrdol, K.W. Wucherpfennig. Unconventional topology of
568 self peptide-major histocompatibility complex binding by a human autoimmune T cell
569 receptor. Nat. Immunol. 6 (2005) 490-496.
570 36. P. Emsley, K. Cowtan. Coot: model-building tools for molecular graphics. Acta.
571 Crystallogr. D Biol. Crystallogr. 60 (2004) 2126-2132.
572 37. M.J. Nicholson, M. Hahn, K.W. Wucherpfennig. Unusual features of self-peptide/MHC
573 binding by autoimmune T cell receptors. Immunity 23 (2005) 351–360,
574 doi:10.1016/j.immuni.2005.09.009.

575 38. K.J. Smith, J. Pyrdol, L. Gauthier, D.C. Wiley, K.W. Wucherpfennig. Crystal structure of
576 HLA-DR2 (DRA*0101, DRB1*1501) complexed with a peptide from human myelin basic
577 protein. J. Exp. Med. 188 (1998) 1511–1520.
578 39. C. Veroni, B. Serafini, B. Rosicarelli, C. Fagnani, F. Aloisi. Transcriptional profile and
579 Epstein-Barr virus infection status of laser-cut immune infiltrates from the brain of
580 patients with progressive multiple sclerosis. J. Neuroinflammation. 15 (2018)18, doi:
581 10.1186/s12974-017-1049-5.
582 40. M.P. Pender, P.A. Csurhes, J.M. Burrows, S.R. Burrows. Defective T-cell control of
583 Epstein-Barr virus infection in multiple sclerosis. Clin. Transl. Immunology. 6 (2017)
584 e126, doi: 10.1038/cti.2016.87.
585 41. A. Afrasiabi, G.P. Parnell, N. Fewings, S.D. Schibeci, M.A. Basuki, R. Chandramohan,
586 et al. Evidence from genome wide association studies implicates reduced control of
587 Epstein-Barr virus infection in multiple sclerosis susceptibility. Genome. Med. 11 (2019)
588 26, doi: 10.1186/s13073-019-0640-z.
589 42. B. Nourbakhsh, A. Rutatangwa, M. Waltz, M. Rensel, M. Moodley, J. Graves, et al.
590 Heterogeneity in association of remote herpesvirus infections and pediatric MS. Ann.
591 Clin. Transl. Neurol. 5 (2018)1222-1228, doi: 10.1002/acn3.636.
592 43. A. Hassani, J.R. Corboy, S. Al-Salam, G. Khan. Epstein-Barr virus is present in the
593 brain of most cases of multiple sclerosis and may engage more than just B cells. PLoS.
594 One 13 (2018) e0192109, https://doi.org/10.1371/journal.pone.0192109
595 44. S. Agostini, R. Mancuso, F.R. Guerini, S. D'Alfonso, C. Agliardi, A. Hernis, et al. HLA
596 alleles modulate EBV viral load in multiple sclerosis. J. Transl. Med. 16 (2018) 80, doi:
597 10.1186/s12967-018-1450-6.
598 45. S. Boudreault, V.E.S. Armero, M.S. Scott, J.P. Perreault, M. Bisaillon. The Epstein-Barr
599 virus EBNA1 protein modulates the alternative splicing of cellular genes. Virol. J. 16
600 (2019) 29, doi: 10.1186/s12985-019-1137-5.
601 46. J. McKenzie, A.G. El-Guindy AG. Epstein-Barr virus lytic cycle reactivation. Curr. Top.
602 Microbiol. Immunol. 391 (2015) 237-261, doi: 10.1007/978-3-319-22834-1_8.

603 47. N. Grandi, E. Tramontano. HERV envelope proteins: physiological role and pathogenic
604 potential in cancer and autoimmunity. Front. Microbiol. 9 (2018) 462, doi:
605 10.3389/fmicb.2018.00462.
606 48. V. Gröger, H. Cynis. Human endogenous retroviruses and their putative role in the
607 development of autoimmune disorders such as multiple sclerosis. Front. Microbiol. 9
608 (2018) 265, doi: 10.3389/fmicb.2018.00265.
609 49. G. Morris, M. Maes, M. Murdjeva, B.K. Puri. Do human endogenous retroviruses
610 contribute to multiple sclerosis, and if so, how? Mol. Neurobiol. 56 (2019) 2590-2605,
611 doi: 10.1007/s12035-018-1255-x.
612 50. V. de Luca, A. H. Martins, C. M. Romano, G.M. Pimenta, L.A. Peroni. Cross-reactivity
613 between myelin oligodendrocyte glycoprotein and human endogenous retrovirus W
614 protein: nanotechnological evidence for the potential trigger of multiple sclerosis.
615 Micron 120 (2019) 66-73, doi: 10.1016/j.micron.2019.02.005.
616 51. G. Arru, G. Mameli, G.A. Deiana, A.L. Rassu, R. Piredda, E. Sechi, et al. Humoral
617 immunity response to human endogenous retroviruses K/W differentiates between
618 amyotrophic lateral sclerosis and other neurological diseases. Eur. J. Neurol. 25
619 (2018)1076-e84, doi: 10.1111/ene.13648.
620 52. G. Arru, E. Sechi, S. Mariotto, A. Farinazzo, C. Mancinelli, D. Alberti, et al. Antibody
621 response against HERV-W env surface peptides differentiates multiple sclerosis and
622 neuromyelitis optica spectrum disorder. Mult .Scler. J. Exp. Transl. Clin. 3 (2017)
623 2055217317742425, doi: 10.1177/2055217317742425.
624 53. R. Planas, R. Santos, P. Tomas-Ojer, C. Cruciani, A. Lutterotti, W. Faigle, et al. GDP-L-
625 fucose synthase is a CD4+ T cell-specific autoantigen in DRB3*02:02 patients with
626 multiple sclerosis. Sci. Transl. Med. 10 (2018) 462 pii: eaat4301, doi:
627 10.1126/scitranslmed.aat4301.
628 54. I. Jelcic, F. Al Nimer, J. Wang, V. Lentsch, R. Planas, I. Jelcic, et al. Memory B cells
629 activate brain-homing, autoreactive CD4+ T cells in multiple sclerosis. Cell 175 (2018)
630 85-100 e23, doi: 10.1016/j.cell.2018.08.011.

631 55. A. Shetty, S.G. Gupta, M. Varrin-Doyer, M.S. Weber, T. Prod'homme, N. Molnarfi, et al.
632 Immunodominant T-cell epitopes of MOG reside in its transmembrane and cytoplasmic
633 domains in EAE. Neurol. Neuroimmunol. Neuroinflamm. 1 (2014) e22, doi:
634 10.1212/NXI.0000000000000022.
635 56. E.M. Scholz, M. Marcilla, X. Daura, D. Arribas-Layton, E.A. James, I. Alvarez. Human
636 leukocyte antigen (HLA)-DRB1*15:01 and HLA-DRB5*01:01 present complementary
637 peptide repertoires. Front. Immunol. 8 (2017) 984, doi: 10.3389/fimmu.2017.00984.
638 57. C.A. Dendrou, L. Fugger. Immunomodulation in multiple sclerosis: promises and
639 pitfalls. Curr. Opin. Immunol. 49 (2017) 37–43.
640 58. J. Chataway, K. Martin, K. Barrell, B. Sharrack, P. Stolt, D.C. Wraith, et al. Effects of
641 ATX-MS-1467 immunotherapy over 16 weeks in relapsing multiple sclerosis. Neurology
642 90 (2018) e955-e962, doi: 10.1212/WNL.0000000000005118.
643 59. N. Kaushansky, A. Kaminitz, H. Allouche-Arnon, A. Ben-Nun. Modulation of MS-like
644 disease by a multi epitope protein is mediated by induction of CD11c+CD11b+Gr1+
645 myeloid-derived dendritic cells. J. Neuroimmunol. 333 (2019) 476953. doi:
646 10.1016/j.jneuroim.2019.04.013.
647 60. S. Kasagi, D. Wang, P. Zhang, P. Zanvit, H. Chen, D. Zhang, et al. Combination of
648 apoptotic T cell induction and self-peptide administration for therapy of experimental
649 autoimmune encephalomyelitis. EBioMedicine. (2019) S2352-3964(19)30306-8, doi:
650 10.1016/j.ebiom.2019.05.005
651 61. C.J. Pickens, M.A. Christopher, M.A. Leon, M.M. Pressnall, S.N. Johnson, S. Thati, et
652 al. Antigen-drug conjugates as a novel therapeutic class for the treatment of antigen-
653 specific autoimmune disorders. Mol. Pharm. 16 (2019) 2452-2461, doi:
654 10.1021/acs.molpharmaceut.9b00063.
655 62. N. Ji, A. Somanaboeina, A. Dixit, K. Kawamura, N.J. Hayward, C. Self, et al. Small
656 molecule inhibitor of antigen binding and presentation by HLA-DR2b as a therapeutic
657 strategy for the treatment of multiple sclerosis, J. Immunol. 191 (2013) 5074–5084,
658 http://dx.doi.org/10.4049/jimmunol.1300407.

63. I. Zubizarreta, G. Flórez-Grau, G. Vila, R. Cabezón, C. España, M. Andorra, et al. Immune tolerance in multiple sclerosis and neuromyelitis optica with peptide-loaded tolerogenic dendritic cells in a phase 1b trial. Proc. Natl. Acad. Sci. USA. 116 (2019) 8463-8470, doi: 10.1073/pnas.1820039116.

64. N. Saligrama, F. Zhao, M.J. Sikora, W.S. Serratelli, R.A. Fernandes, D.M. Louis, et al. Opposing T cell responses in experimental autoimmune encephalomyelitis. Nature (2019) https://doi/org/10.1038/s41586-019-1467-x.

Figure Legends

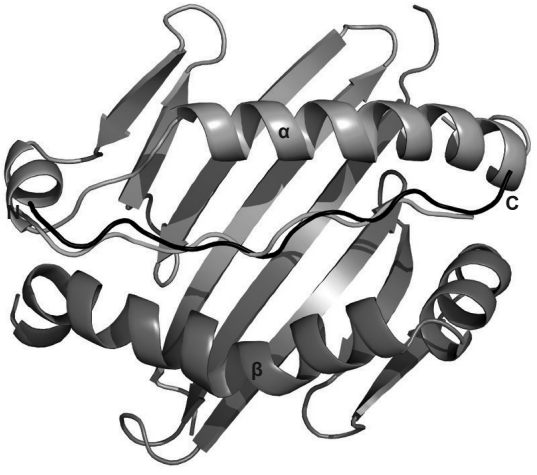
Figure 1 Comparison of the HLA DR2b-MBP_3 complex generated by HADDOCK with the reference structure. (A) Superposition of MBP_3 peptides bound to HLA DR2b in the reference (cyan) and modelled structures (black). (B) Ribbon representation of the published crystal structure of HLA DR2b bound to MBP peptide (MBP_3; ENPVVHFFKNIVTPR) (PDB entry 1YMM). (C) Ribbon representation of the lowest energy HLA DRb-MBP_3 complex model structure generated by HADDOCK. The HLA DR2b alpha and beta chains are depicted as pink and blue, respectively. For clarity only the peptide binding groove is highlighted. The peptide side chains (ball and stick format) and positions (red) are shown. Peptide residues P1, P4, P6 and P9 serve as anchor residues which slot into the antigen binding groove, whereas side chains at P-1, P2, P5 and P8 are surface exposed. HLA DR2b residues involved in stabilising peptide binding are also highlighted (ball and stick format). The black rectangle boxes correspond to the core 9-mer sequence for each peptide. Figure was generated with Pymol (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC)

Figure 2 Comparison of HADDOCK generated models of HLA DR2b in complex with peptides derived from myelin (MOG) and HERV W-family (MSRVenv, SYN1 and SYN2) associated proteins. (A) Superposition of MOG_4 (red), MSRVenv_5 (blue),

SYN1_2 (yellow) and SYN2_5 (green) peptides bound to HLA DR2b. (B) Ribbon representation of the lowest energy HLA DR2b-MOG_4 complex model structure. (C) Ribbon representation of the lowest energy HLA DRb-MSRV_5 complex model structure. (D) Ribbon representation of the lowest energy HLA DR2b-SYN1_2 complex model structure. (E) Ribbon representation of the lowest energy HLA DRb-SYN2_5 complex model structure. The HLA DR2b alpha and beta chains are depicted as pink and blue, respectively. For clarity only the peptide binding groove is highlighted. The peptide side chains (ball and stick format) and positions (red) are shown. Peptide residues P1, P4, P6 and P9 serve as anchor residues which insert into the antigen binding groove, whereas side chains at P-1, P2, P5 and P8 are surface exposed. HLA DR2b residues that contribute to peptide interactions are also highlighted (ball and stick format). The black rectangle boxes correspond to the core 9-mer sequence for each peptide.

Figure 3 Comparison of HADDOCK generated models of HLA DR2b in complex with peptides derived from a CNS (β-SYN) and an EBV (EBNA1) protein. (A) Superposition of βSYN (grey) and EBNA1_2 (orange) peptides bound to HLA DR2b. (B) Ribbon representation of the lowest energy HLA DR2b-βSYN complex model structure. (C) Ribbon representation of the lowest energy HLA DRb-EBNA1_2 complex model structure. The HLA DR2b alpha and beta chains are depicted as pink and blue, respectively. For clarity only the peptide binding groove is highlighted. The peptide side chains (ball and stick format) and positions (red) are shown. Peptide residues P1, P4, P6 and P9 serve as anchor residues which slot into the antigen binding groove, whereas side chains at P-1, P2, P5 and P8 are surface exposed. HLA DR2b residues involved in peptide binding are also highlighted (ball and stick format). The black rectangle boxes correspond to the core 9-mer sequence for each peptide.

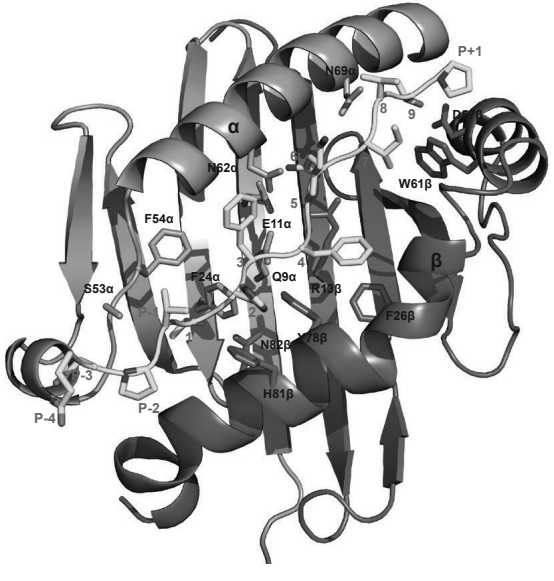
A



HLA DR2b-MBP_3 complex

B

Position	P-4	P-3	P-2	P-1	1	2	3	4	5	6	7	8	9	P+1	P+2
MBP_3	E	N	P	V	V	H	F	F	K	N	I	V	T	P	R

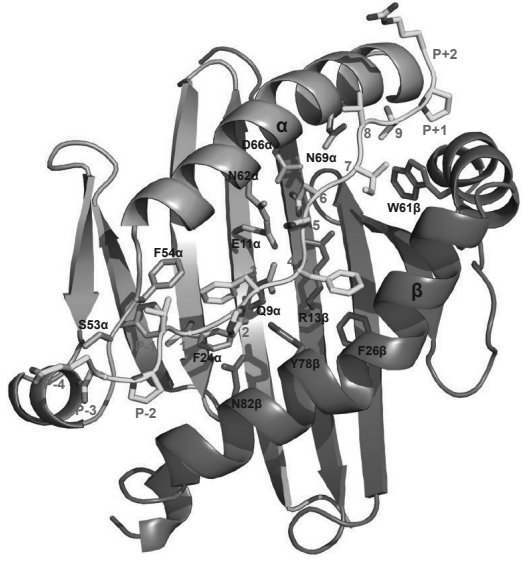


HLA DR2b-MBP_3 complex structure (PDB entry 1YMM)

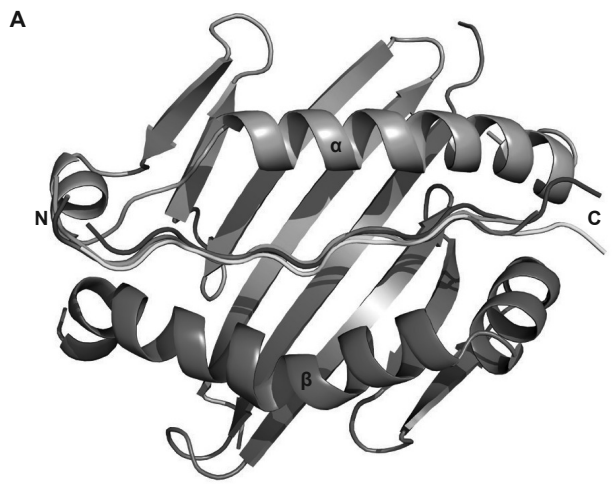
C

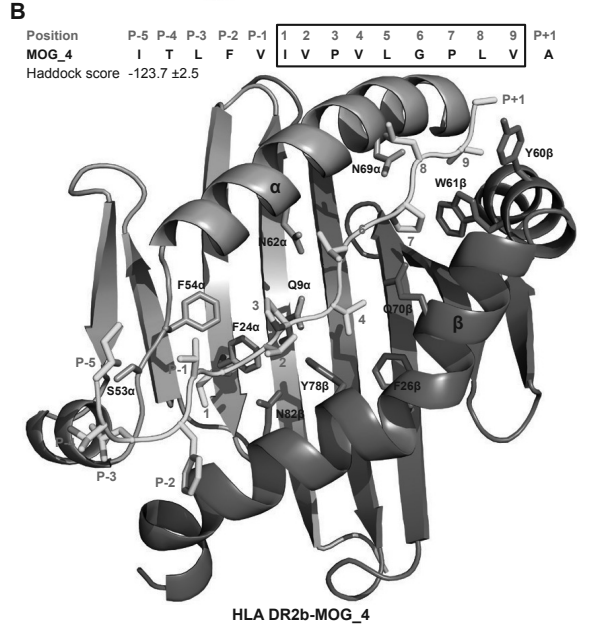
Position	P-4	P-3	P-2	P-1	1	2	3	4	5	6	7	8	9	P+1	P+2
MBP_3	E	N	P	V	V	H	F	F	K	N	I	V	T	P	R

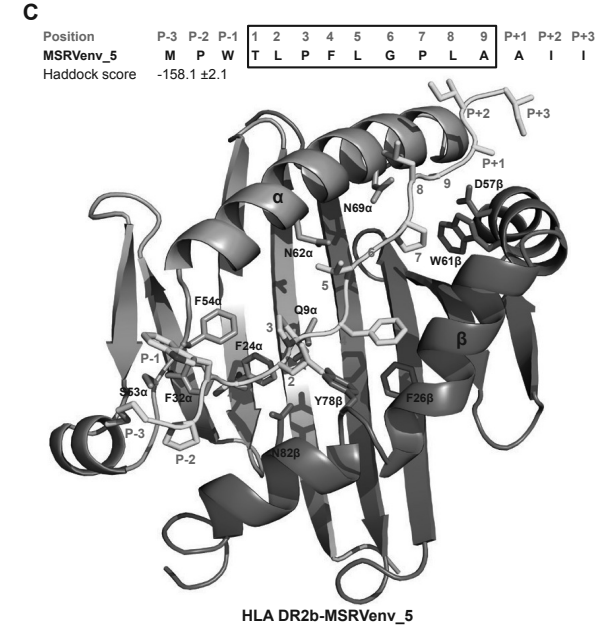
Haddock score -163.7 ±1.6

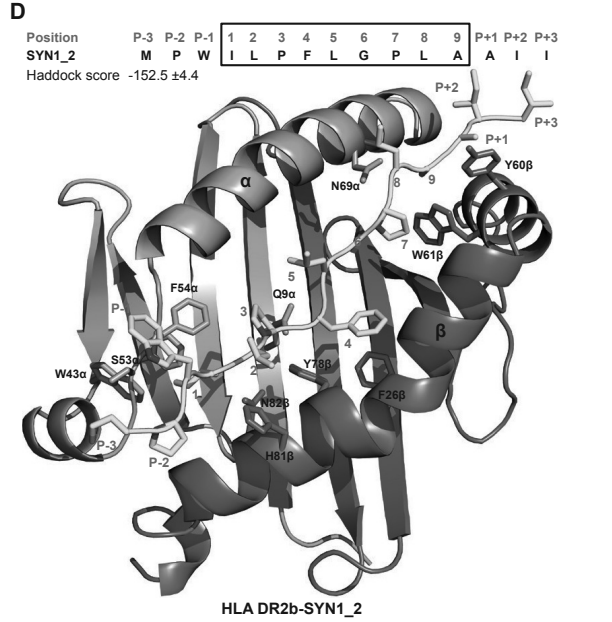


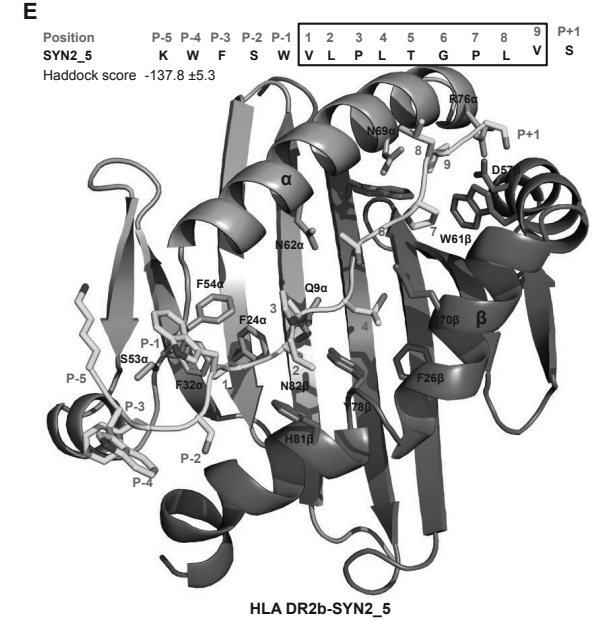
HLA DR2b-MBP_3 complex model

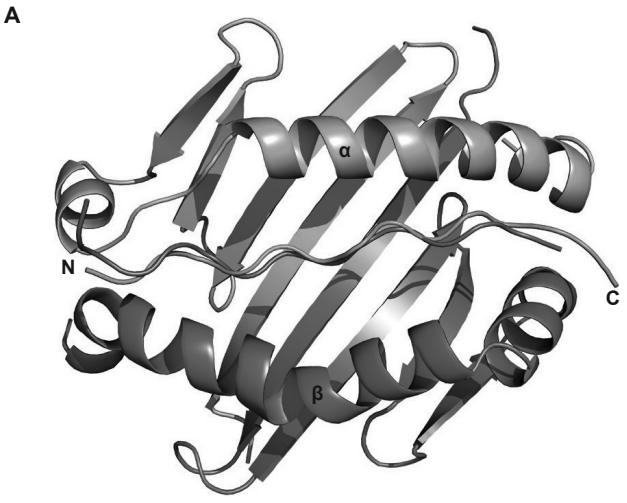

A
N
C
α
β


B
Position P-5 P-4 P-3 P-2 P-1 1 2 3 4 5 6 7 8 9 P+1
MOG_4 I T L F V I V P V L G P L V A
Haddock score -123.7 ±2.5
HLA DR2b-MOG_4


C
Position P-3 P-2 P-1 1 2 3 4 5 6 7 8 9 P+1 P+2 P+3
MSRVenv_5 M P W T L P F L G P L A A I I
Haddock score -158.1 ±2.1
HLA DR2b-MSRVenv_5


D
Position P-3 P-2 P-1 1 2 3 4 5 6 7 8 9 P+1 P+2 P+3
SYN1_2 M P W I L P F L G P L A A I I
Haddock score -152.5 ±4.4
HLA DR2b-SYN1_2

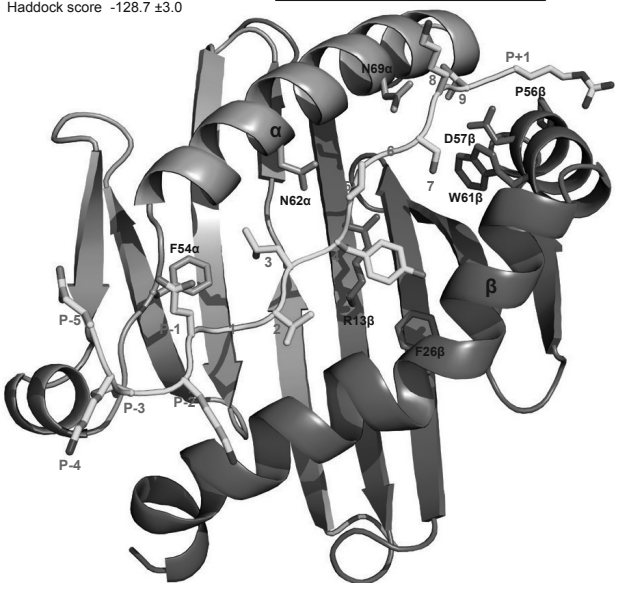

E
Position P-5 P-4 P-3 P-2 P-1 1 2 3 4 5 6 7 8 9 P+1
SYN2_5 K W F S W V L P L T G P L V S
Haddock score -137.8 ±5.3
HLA DR2b-SYN2_5



B

Position	P-5	P-4	P-3	P-2	P-1	1	2	3	4	5	6	7	8	9	P+1
β-synuclein	E	K	T	K	E	G	V	L	Y	V	G	S	K	T	R

Haddock score -128.7 ±3.0

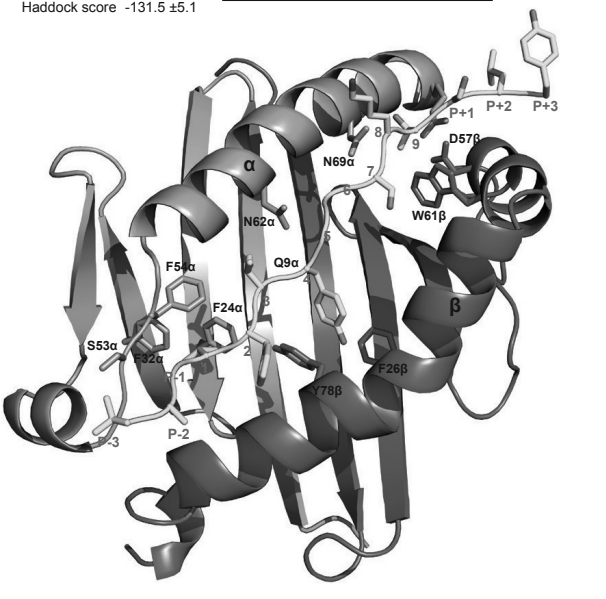


HLA-DR2b-β-synuclein complex model

C

Position	P-3	P-2	P-1	1	2	3	4	5	6	7	8	9	P+1	P+2	P+3
EBNA1_2	V	A	G	V	F	V	Y	G	G	S	K	T	S	L	Y

Haddock score -131.5 ±5.1



HLA-DR2b-EBNA1_2 complex model

Table S1. Details of Peptides Tested in the Two REVEAL DR2b Binding Assays

Peptide Set 1

Protein	Peptide Name	15mer Peptide Sequence	Amino Acid Residues	IC_{50}nM	SMM Rank	Core Nonamer	Reference
MBP	MBP_1	PVVHFFKNIVTPRTP	219-233	26	0.14	VHFFKNIVT	16
MBP	MBP_2	NPVVHFFKNIVTPRT	218-232	21	0.09	VHFFKNIVT	16
MBP	MBP_3	ENPVVHFFKNIVTPR	217-231	21	0.09	VHFFKNIVT	16
MBP	MBP_4	DENPVVHFFKNIVTP	216-230	22	0.1	VHFFKNIVT	16
MBP	MBP_5	QDENPVVHFFKNIVT	215-229	21	0.09	VVHFFKNIV	16
MOG	MOG_1	FVIVPVLGPLVALII	212-226	40	0.3	IVPVLGPLV	16
MOG	MOG_2	LFVIVPVLGPLVALI	211-225	38	0.27	IVPVLGPLV	16
MOG	MOG_3	TLFVIVPVLGPLVAL	210-224	41	0.31	IVPVLGPLV	16
MOG	MOG_4	ITLFVIVPVLGPLVA	209-223	42	0.33	IVPVLGPLV	16
MOG	MOG_5	KITLFVIVPVLGPLV	208-222	43	0.34	ITLFVIVPV	16
PLP	PLP_1	SFFFLYGALLLAEGF	77-91	88	1.04	FFLYGALLL	16
PLP	PLP_2	ASFFFLYGALLLAEG	76-90	52	0.46	FFFLYGALL	16
PLP	PLP_3	TASFFFLYGALLLAE	75-89	51	0.45	FFFLYGALL	16
PLP	PLP_4	GTASFFFLYGALLLA	74-88	50	0.43	FFFLYGALL	16
PLP	PLP_5	YGTASFFFLYGALLL	73-87	51	0.45	FFFLYGALL	16
SYNCYTIN1	SYN1_1	PWILPFLGPLAAIIL	446-460	39	0.28	ILPFLGPLA	16
SYNCYTIN1	SYN1_2	MPWILPFLGPLAAII	445-459	39	0.28	ILPFLGPLA	16
SYNCYTIN1	SYN1_3	WMPWILPFLGPLAAI	444-458	41	0.31	ILPFLGPLA	16
SYNCYTIN1	SYN1_4	QWMPWILPFLGPLAA	443-457	44	0.35	WMPWILPFL	16
SYNCYTIN1	SYN1_5	SQWMPWILPFLGPLA	442-456	44	0.35	ILPFLGPLA	16
MSRVenv	MSRVenv_1	LPFLGPLAAIIFLLL	449-463	296	5.5	LGPLAAIIF	16
MSRVenv	MSRVenv_2	TLPFLGPLAAIIFLL	448-462	294	5.46	LGPLAAIIF	16
MSRVenv	MSRVenv_3	WTLPFLGPLAAIIFL	447-461	297	5.52	LGPLAAIIF	16
MSRVenv	MSRVenv_4	MPWTLPFLGPLAAII	445-459	246	4.37	TLPFLGPLA	16
MSRVenv	MSRVenv_5	WMPWTLPFLGPLAAI	444-458	284	5.24	TLPFLGPLA	16
MSRVenv	MSRVenv_6	QWMPWTLPFLGPLAA	443-447	300	5.59	TLPFLGPLA	16
MSRVenv	MBP & PLP Homologue_1	LFTVLLPPFALTAPP	9-23	304	5.68	TVLLPPFAL	16
MSRVenv	MBP & PLP Homologue_2	TFLFTVLLPPFALTA	7-21	160	2.44	FTVLLPPFA	16
MSRVenv	MBP & PLP Homologue_3	HTFLFTVLLPPFALT	6-20	163	2.5	FTVLLPPFA	16

MSRVenv	MBP & PLP Homologue_4	YHTFLFTVLLPPFAL	5-19	164	2.52	FTVLLPPFA	16
MSRVenv	MBP & PLP Homologue_5	PYHTFLFTVLLPPFA	4-18	215	3.67	LFTVLLPPF	16
SYN1	MBP & PLP Homologue_6	PYHIFLFTVLLPSFT	4-18	239	4.21	HIFLFTVLL	16
SYN1	MBP & PLP Homologue_7	ALPYHIFLFTVLLPS	2-16	276	5.07	YHIFLFTVL	16
SYN1	MBP & PLP Homologue_8	MALPYHIFLFTVLLP	1-15	278	5.11	YHIFLFTVL	16
EBV DNA Polymerase		TGGVYHFVKKHVHES	627-641	235	4.12	VYHFVKKHV	30
MOG	Identified epitope	FLCLQYRLRGKLRAE	175-189	714	13.75	QYRLRGKLR	31
MSRVenv	Weaker homology to MOG_1	NTTSVLVGPLVSNLE	214-228	513	10	TSVLVGPLV	16
MSRVenv	Weaker homology to MOG_2	INTTSVLVGPLVSNL	213-227	632	12.25	TSVLVGPLV	16
MSRVenv	Weaker homology to MOG_3	EINTTSVLVGPLVSN	212-226	719	13.84	TSVLVGPLV	16
MSRVenv	Weaker homology to MOG_4	TEINTTSVLVGPLVS	211-225	725	13.93	TSVLVGPLV	16

Peptide Set 2

Protein	Peptide Name	15mer Peptide	Location in protein	IC_{50}nM	SMM Rank	Core Nonamer	Reference
MSRVenv	MSRVenv_1	LPFLGPLAAIIFLLL	449-463	296	5.5	LGPLAAIIF	16
MSRVenv	MSRVenv_2	TLPFLGPLAAIIFLL	448-462	294	5.46	LGPLAAIIF	16
MSRVenv	MSRVenv_3	WTLPFLGPLAAIIFL	447-461	297	5.52	LGPLAAIIF	16
MSRVenv	MSRVenv_4	MPWTLPFLGPLAAII	445-459	246	4.37	TLPFLGPLA	16
MSRVenv	MSRVenv_7	PWTLPFLGPLAAIIF	446-460	197	3.28	TLPFLGPLA	16
MOG	MOG_1	FVIVPVLGPLVALII	211-225	40	0.3	IVPVLGPLV	16
MOG	MOG_2	LFVIVPVLGPLVALI	210-224	38	0.27	IVPVLGPLV	16
MOG	MOG_3	TLFVIVPVLGPLVAL	209-223	41	0.31	IVPVLGPLV	16
MOG	MOG_4	ITLFVIVPVLGPLVA	208-222	42	0.33	IVPVLGPLV	16
SYN1	SYN1_1	PWILPFLGPLAAIIL	446-460	39	0.28	ILPFLGPLA	16
SYN1	SYN1_2	MPWILPFLGPLAAII	445-459	39	0.28	ILPFLGPLA	16
SYN1	SYN1_3	WMPWILPFLGPLAAI	444-458	41	0.31	ILPFLGPLA	16
SYN1	SYN1_4	QWMPWILPFLGPLAA	443-457	44	0.35	WMPWILPFL	16
SYN2	SYN2_1	WVLPLTGPLVSLLLL	482-496	308	5.78	VLPLTGPLV	16
SYN2	SYN2_2	SWVLPLTGPLVSLLL	481-495	133	1.88	VLPLTGPLV	16
SYN2	SYN2_3	FSWVLPLTGPLVSLL	480-494	130	1.83	VLPLTGPLV	16
SYN2	SYN2_4	WFSWVLPLTGPLVSL	479-493	142	2.07	VLPLTGPLV	16
SYN2	SYN2_5	KWFSWVLPLTGPLVS	478-492	149	2.21	VLPLTGPLV	16

SYN2	SYN2_6	WKWFSWVLPLTGPLV	477-491	147	2.17	FSWVLPLTG	16
MBP	MBP_3 as +ve control	ENPVVHFFKNIVTPR	217-231	21	0.09	VHFFKNIVT	16
MBP	MBP_6	GTLSKIFKLGGRDSR	282-296	940	17.59	IFKLGGRDS	33
SYN1	SYN1 homolog for MBP_1	YHIFLFTVLLPSFTL	5-19	95	1.18	FTVLLPSFT	16
SYN1	SYN1 homolog for MBP_2	LPYHIFLFTVLLPSF	3-17	274	5.03	YHIFLFTVL	16
SYN1	SYN1 homolog for MBP_3	ALPYHIFLFTVLLPS	2-16	276	5.07	YHIFLFTVL	16
SYN1	SYN1 homolog for MBP_4	MALPYHIFLFTVLLP	1-15	278	5.11	YHIFLFTVL	16
MSRVenv	MSRVenv homolog for MBP	LPYHTFLFTVLLPPF	3-17	738	14.17	YHTFLFTVL	16
OSP	EBNA1 homolog_1	STTLRALAPRLMRRV	190-204	149	2.21	LRALAPRLM	32
OSP	EBNA1 homolog_2	AGVLLILLALCALVA	123-137	33	0.2	LLILLALCA	32
EBNA1	OSP homolog_1	NIAEGLRALLARSHV	480-494	288	5.34	LRALLARSH	32
EBNA1	OSP homolog_2	AEGLRALLARSHVER	482-496	306	5.73	LRALLARSH	32
βSYN	β synuclein (EBNA1 homolog)	EKTKEGVLYVGSKTR	31-45	284	5.24	GVLYVGSKT	32
EBNA1	EBNA1_1 (β synuclein homolog)	AGVFVYGGSKTSLYN	505-519	488	9.48	VFVYGGSKT	32
EBNA1	EBNA1_2 (β synuclein homolog)	VAGVFVYGGSKTSLY	504-518	410	7.9	VFVYGGSKT	32
EBNA1	EBNA1_3 (β synuclein homolog)	WVAGVFVYGGSKTSL	503-517	429	8.28	VFVYGGSKT	32
MSRV	OSP homolog_1	PLAAIIFLLLFGPCI	454-468	85	0.99	IIFLLLFGP	n/a
MSRV	OSP homolog_2	LAAIIFLLLFGPCIF	455-469	69	0.73	FLLLFGPCI	n/a
OSP	MSRV homolog_1	GLPAILLLLTVLPCI	92-106	267	4.86	ILLLLTVLP	n/a
OSP	MSRV homolog_2	LPAILLLLTVLPCIR	93-107	268	4.88	ILLLLTVLP	n/a
OSP	MSRV homolog_3	PAILLLLTVLPCIRM	94-108	269	4.91	ILLLLTVLP	n/a
OSP	MSRV homolog_4	ILLLLTVLPCIRMGQ	96-110	425	8.21	LLLTVLPCI	n/a

Legend to Table S1

Column 1 - Protein origin of the concerned peptide

Column 2 - The peptide characteristics and nomenclature

Column 3 - Sequence of the 15mer peptide identified by IEDB analysis

Column 4 - 15mer Peptide sequence numbers corresponding to the coding sequences in the gene.

Column 5 - IC_{50} determined by IEDB-SMM analysis

Column 6 - Rank determined by IEDB-SMM analysis

Column 7 - HLA DR2b binding nonamer sequence predicted by IEDB-SMM analysis

Column 8 - Previous literature referring to the peptide

n/a	not applicable

Table S2. Region of sequence homology identified in BLASTp analysis of the selected CNS proteins against all non-redundant protein sequences in the EBV proteome taxid 10377

1. ABP (Query) vs whole EBV PROTEOME (subject)

RecName: Full=Epstein-Barr nuclear antigen 4; Short=EBNA-4; Short=EBV nuclear antigen 4; AltName: Full=Epstein-Barr nuclear antigen 3B; Short=EBNA-3B; Short=EBV nuclear antigen 3B

Sequence ID: P03203.3Length: 938Number of Matches: 1
Related Information
Gene-associated gene details
Identical Proteins-Identical proteins to YP_401670.1
Range 1: 571 to 623GenPeptGraphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps
22.3 bits(46)	0.95	Compositional matrix adjust.	17/53(32%)	25/53(47%)	2/53(3%)

```
Query  125  PADVDPLT--ITSSLSSDGVLTVNGPRKQVSGPERTIPITREEKPAVTAAPKK  175

             P ++ PLT   TS LSS         P   V   +     T++ +P  TAAP++

Sbjct  571  PLEIQPLTSPTTSQLSSSAPSCAQTPWPVVQPSQTPDDPTKQSRPPETAAPRQ  623
```

Table S3 – Regions of sequence homologies identified in BLASTp analysis of the selected CNS proteins against the three HERV env proteins MSRV env, SYN1 and SYN2

1. ABP (Query) vs SYN2 (Subject)

Score	Expect	Method	Identities	Positives	Gaps
25.4 bits(54)	0.001	Composition-based stats.	18/53(34%)	26/53(49%)	7/53(13%)

```
Query  6     HHPWIHRPFFPFHSPSRLFDQFFGEHLLESDLF----PTSTSLSPFYLRPPSF  54

             H+ + H+P FP   P+  F Q  G  L +S  F    P+S S   F+ RP  +

Sbjct  163   HNQFRHQPRFP-KPPNITFPQ--GTLLDKSSRFCQGRPSSCSTRNFWFRPADY  21
```

2. ABP (Query) vs MSRV env (Subject)

Score	Expect	Method	Identities	Positives	Gaps
16.9 bits(32)	0.50	Composition-based stats.	7/31(23%)	15/31(48%)	0/31(0%)

```
Query  116  REFHRKYRIPADVDPLTITSSLSSDGVLTVN  146

             +EF  + R+P ++D  +  S    +   T +

Sbjct  35   QEFLWRTRLPGNIDAPSYRSLSKGNSTFTAH  65
```

3. MAG (Query) vs MSRV env (Subject)

Score	Expect	Method	Identities	Positives	Gaps
19.6 bits(39)	0.34	Compositional matrix adjust.	17/71(24%)	32/71(45%)	5/71(7%)

```
Query  142
NIVVPPEVVAGTEVEVSCMVPDNCPELRPELSWLGHEGLGEPAVLGRLREDEGTWVQVSL  201

+ +VPP  +   +   + +VP    +  P L ++   G     VLGRL    G+    +

Subject 291

SFLVPPMTIYTEQDLYNHVVPKPHNKRVPILPFVIRAG-----VLGRLGTGIGSITTSTQ  345
```

Table S4. Results of the IEDB analysis of the CNS proteins ABP, MAG, OSP and βSYN showing peptides with potential to bind HLA DR2b

1. ABP

Start	End	Core nonamer	15mer peptide	IC_{50}	Percentile rank
9	23	FFPFHSPSR	WIHRPFFPFHSPSRL	394	7.59
19	33	FDQFFGEHL	SPSRLFDQFFGEHLL	409	7.88
21	35	FDQFFGEHL	SRLFDQFFGEHLLES	423	8.16
20	34	FDQFFGEHL	PSRLFDQFFGEHLLE	425	8.21
22	36	FDQFFGEHL	RLFDQFFGEHLLESD	433	8.37
8	22	HRPFFPFHS	PWIHRPFFPFHSPSR	449	8.69
42	56	FYLRPPSFL	TSLSPFYLRPPSFLR	467	9.04
18	32	SRLFDQFFG	HSPSRLFDQFFGEHL	483	9.37
10	24	FFPFHSPSR	IHRPFFPFHSPSRLF	495	9.62
11	25	FFPFHSPSR	HRPFFPFHSPSRLFD	511	9.96
43	57	YLRPPSFLR	SLSPFYLRPPSFLRA	529	10.32
48	62	PSFLRAPSW	YLRPPSFLRAPSWFD	554	10.83
12	26	FFPFHSPSR	RPFFPFHSPSRLFDQ	573	11.17
44	58	YLRPPSFLR	LSPFYLRPPSFLRAP	578	11.26
49	63	FLRAPSWFD	LRPPSFLRAPSWFDT	600	11.68
46	60	YLRPPSFLR	PFYLRPPSFLRAPSW	610	11.85
50	64	FLRAPSWFD	RPPSFLRAPSWFDTG	628	12.18
5	19	WIHRPFFPF	IHHPWIHRPFFPFHS	632	12.25
41	55	LSPFYLRPP	STSLSPFYLRPPSFL	635	12.3
6	20	HRPFFPFHS	HHPWIHRPFFPFHSP	645	12.48
45	59	YLRPPSFLR	SPFYLRPPSFLRAPS	655	12.66
7	21	HRPFFPFHS	HPWIHRPFFPFHSPS	662	12.78
128	142	LTITSSLSS	VDPLTITSSLSSDGV	672	12.98
38	52	STSLSPFYL	FPTSTSLSPFYLRPP	684	13.19

127	141	LTITSSLSS	DVDPLTITSSLSSDG	703	13.53
125	139	PLTITSSLS	PADVDPLTITSSLSS	707	13.61
126	140	LTITSSLSS	ADVDPLTITSSLSSD	721	13.87
51	65	FLRAPSWFD	PPSFLRAPSWFDTGL	740	14.21
52	66	FLRAPSWFD	PSFLRAPSWFDTGLS	744	14.28
129	143	LTITSSLSS	DPLTITSSLSSDGVL	754	14.45
39	53	STSLSPFYL	PTSTSLSPFYLRPPS	755	14.46
35	49	TSTSLSPFY	SDLFPTSTSLSPFYL	810	15.42
36	50	STSLSPFYL	DLFPTSTSLSPFYLR	818	15.56
37	51	STSLSPFYL	LFPTSTSLSPFYLRP	821	15.61
112	126	ISREFHRKY	GFISREFHRKYRIPA	858	16.25
1	15	IAIHHPWIH	MDIAIHHPWIHRPFF	865	16.35
47	61	PSFLRAPSW	FYLRPPSFLRAPSWF	924	17.32
24	38	FDQFFGEHL	FDQFFGEHLLESDLF	971	18.09
114	128	FHRKYRIPA	ISREFHRKYRIPADV	988	18.37
139	153	VLTVNGPRK	SDGVLTVNGPRKQVS	992	18.43
138	152	VLTVNGPRK	SSDGVLTVNGPRKQV	993	18.45
40	54	LSPFYLRPP	TSTSLSPFYLRPPSF	996	18.49
23	37	FDQFFGEHL	LFDQFFGEHLLESDL	1006	18.62
137	151	VLTVNGPRK	LSSDGVLTVNGPRKQ	1028	18.94
140	154	VLTVNGPRK	DGVLTVNGPRKQVSG	1036	19.07
78	92	VKHFSPEEL	NLDVKHFSPEELKVK	1041	19.14
113	127	FHRKYRIPA	FISREFHRKYRIPAD	1043	19.17
77	91	VKHFSPEEL	VNLDVKHFSPEELKV	1050	19.27
75	89	VNLDVKHFS	FSVNLDVKHFSPEEL	1056	19.35
136	150	GVLTVNGPR	SLSSDGVLTVNGPRK	1083	19.74
76	90	VKHFSPEEL	SVNLDVKHFSPEELK	1109	20.1
79	93	VKHFSPEEL	LDVKHFSPEELKVKV	1110	20.12
87	101	KVKVLGDVI	EELKVKVLGDVIEVH	1153	20.74
86	100	KVKVLGDVI	PEELKVKVLGDVIEV	1199	21.36
108	122	EHGFISREF	QDEHGFISREFHRKY	1220	21.64

115	129	FHRKYRIPA	SREFHRKYRIPADVD	1227	21.75
116	130	FHRKYRIPA	REFHRKYRIPADVDP	1247	22.01
85	99	KVKVLGDVI	SPEELKVKVLGDVIE	1262	22.21
13	27	FFPFHSPSR	PFFPFHSPSRLFDQF	1273	22.36
84	98	LKVKVLGDV	FSPEELKVKVLGDVI	1312	22.86
88	102	KVKVLGDVI	ELKVKVLGDVIEVHG	1364	23.53
110	124	ISREFHRKY	EHGFISREFHRKYRI	1364	23.53
131	145	LTITSSLSS	LTITSSLSSDGVLTV	1399	23.96
130	144	LTITSSLSS	PLTITSSLSSDGVLT	1491	25.07
14	28	FFPFHSPSR	FFPFHSPSRLFDQFF	1601	26.35
2	16	IAIHHPWIH	DIAIHHPWIHRPFFP	1626	26.64
111	125	REFHRKYRI	HGFISREFHRKYRIP	1680	27.27
107	121	EHGFISREF	RQDEHGFISREFHRK	1775	28.31
106	120	EHGFISREF	ERQDEHGFISREFHR	1867	29.3
109	123	EHGFISREF	DEHGFISREFHRKYR	1887	29.5
3	17	IAIHHPWIH	IAIHHPWIHRPFFPF	1901	29.64
124	138	ADVDPLTIT	IPADVDPLTITSSLS	1923	29.86
53	67	FLRAPSWFD	SFLRAPSWFDTGLSE	2097	31.59
27	41	EHLLESDLF	FFGEHLLESDLFPTS	2107	31.69
26	40	EHLLESDLF	QFFGEHLLESDLFPT	2124	31.86
28	42	EHLLESDLF	FGEHLLESDLFPTST	2218	32.72
65	79	MRLEKDRFS	LSEMRLEKDRFSVNL	2262	33.12
105	119	EHGFISREF	EERQDEHGFISREFH	2270	33.19
104	118	ERQDEHGFI	HEERQDEHGFISREF	2391	34.24
66	80	MRLEKDRFS	SEMRLEKDRFSVNLD	2453	34.76
25	39	EHLLESDLF	DQFFGEHLLESDLFP	2459	34.8
89	103	KVKVLGDVI	LKVKVLGDVIEVHGK	2475	34.94
34	48	LFPTSTSLS	ESDLFPTSTSLSPFY	2502	35.15
132	146	TSSLSSDGV	TITSSLSSDGVLTVN	2503	35.16
141	155	VLTVNGPRK	GVLTVNGPRKQVSGP	2508	35.2
62	76	EMRLEKDRF	DTGLSEMRLEKDRFS	2521	35.3

63	77	MRLEKDRFS	TGLSEMRLEKDRFSV	2547	35.5
64	78	MRLEKDRFS	GLSEMRLEKDRFSVN	2550	35.52
54	68	FLRAPSWFD	FLRAPSWFDTGLSEM	2585	35.8
118	132	FHRKYRIPA	FHRKYRIPADVDPLT	2634	36.2
71	85	DRFSVNLDV	EKDRFSVNLDVKHFS	2651	36.33
117	131	FHRKYRIPA	EFHRKYRIPADVDPL	2656	36.37
80	94	VKHFSPEEL	DVKHFSPEELKVKVL	2763	37.2
67	81	MRLEKDRFS	EMRLEKDRFSVNLDV	2806	37.51
68	82	DRFSVNLDV	MRLEKDRFSVNLDVK	2840	37.76
15	29	FPFHSPSRL	FPFHSPSRLFDQFFG	2871	37.98
142	156	VLTVNGPRK	VLTVNGPRKQVSGPE	2928	38.41
81	95	VKHFSPEEL	VKHFSPEELKVKVLG	2938	38.48
158	172	ITREEKPAV	TIPITREEKPAVTAA	3023	39.08
157	171	ITREEKPAV	RTIPITREEKPAVTA	3073	39.43
159	173	ITREEKPAV	IPITREEKPAVTAAP	3073	39.43
156	170	ITREEKPAV	ERTIPITREEKPAVT	3079	39.47
70	84	DRFSVNLDV	LEKDRFSVNLDVKHF	3128	39.78
91	105	LGDVIEVHG	VKVLGDVIEVHGKHE	3159	39.99
4	18	PWIHRPFFP	AIHHPWIHRPFFPFH	3238	40.5
29	43	LLESDLFPT	GEHLLESDLFPTSTS	3271	40.71
30	44	LLESDLFPT	EHLLESDLFPTSTSL	3275	40.73
90	104	KVKVLGDVI	KVKVLGDVIEVHGKH	3306	40.93
92	106	VIEVHGKHE	KVLGDVIEVHGKHEE	3439	41.78
16	30	SRLFDQFFG	PFHSPSRLFDQFFGE	3464	41.93
155	169	TIPITREEK	PERTIPITREEKPAV	3538	42.38
69	83	DRFSVNLDV	RLEKDRFSVNLDVKH	3626	42.89
17	31	SRLFDQFFG	FHSPSRLFDQFFGEH	3666	43.13
72	86	VNLDVKHFS	KDRFSVNLDVKHFSP	3945	44.7
133	147	LSSDGVLTV	ITSSLSSDGVLTVNG	3947	44.71
83	97	PEELKVKVL	HFSPEELKVKVLGDV	3958	44.77
73	87	VNLDVKHFS	DRFSVNLDVKHFSPE	4002	45

93	107	VIEVHGKHE	VLGDVIEVHGKHEER	4063	45.32
31	45	SDLFPTSTS	HLLESDLFPTSTSLS	4159	45.81
94	108	VIEVHGKHE	LGDVIEVHGKHEERQ	4206	46.05
59	73	FDTGLSEMR	SWFDTGLSEMRLEKD	4301	46.55
32	46	SDLFPTSTS	LLESDLFPTSTSLSP	4309	46.59
58	72	DTGLSEMRL	PSWFDTGLSEMRLEK	4590	47.93
56	70	FDTGLSEMR	RAPSWFDTGLSEMRL	4595	47.95
33	47	SDLFPTSTS	LESDLFPTSTSLSPF	4685	48.37
57	71	FDTGLSEMR	APSWFDTGLSEMRLE	4685	48.37
74	88	VNLDVKHFS	RFSVNLDVKHFSPEE	4710	48.49
95	109	VIEVHGKHE	GDVIEVHGKHEERQD	4806	48.91
135	149	LSSDGVLTV	SSLSSDGVLTVNGPR	4809	48.92
134	148	TSSLSSDGV	TSSLSSDGVLTVNGP	4948	49.55

2. MAG

Start	End	Core Nonamer	15mer Peptide	IC_{50}	Percentile Rank
6	20	FWIMISASR	ALPLFWIMISASRGG	120	1.64
7	21	FWIMISASR	LPLFWIMISASRGGH	121	1.66
8	22	FWIMISASR	PLFWIMISASRGGHW	126	1.75
1	15	MIFLTALPL	MIFLTALPLFWIMIS	149	2.21
9	23	IMISASRGG	LFWIMISASRGGHWG	160	2.44
5	19	FWIMISASR	TALPLFWIMISASRG	164	2.52
520	534	FAILIAIVC	VVAFAILIAIVCYIT	178	2.81
4	18	LTALPLFWI	LTALPLFWIMISASR	183	2.92
521	535	LIAIVCYIT	VAFAILIAIVCYITQ	187	3.02
10	24	IMISASRGG	FWIMISASRGGHWGA	214	3.66
523	537	LIAIVCYIT	FAILIAIVCYITQTR	221	3.8
53	67	VWYFNSPYP	AVVHGVWYFNSPYPK	224	3.87
223	237	FEGYASMDV	NTTLQFEGYASMDVK	225	3.89
464	478	GLVLTSILT	ERSGLVLTSILTLRG	225	3.89

465	479	GLVLTSILT	RSGLVLTSILTLRGQ	227	3.93
222	236	TLQFEGYAS	PNTTLQFEGYASMDV	228	3.96
524	538	LIAIVCYIT	AILIAIVCYITQTRR	228	3.96
354	368	LTIFKEKQI	DPILTIFKEKQILST	229	3.98
522	536	LIAIVCYIT	AFAILIAIVCYITQT	230	4
353	367	LTIFKEKQI	PDPILTIFKEKQILS	232	4.04
52	66	GVWYFNSPY	PAVVHGVWYFNSPYP	234	4.09
55	69	VWYFNSPYP	VHGVWYFNSPYPKNY	237	4.16
352	366	LTIFKEKQI	NPDPILTIFKEKQIL	238	4.19
224	238	FEGYASMDV	TTLQFEGYASMDVKY	241	4.25
355	369	LTIFKEKQI	PILTIFKEKQILSTV	241	4.25
54	68	VWYFNSPYP	VVHGVWYFNSPYPKN	243	4.3
225	239	FEGYASMDV	TLQFEGYASMDVKYP	247	4.39
351	365	ILTIFKEKQ	SNPDPILTIFKEKQI	258	4.64
226	240	FEGYASMDV	LQFEGYASMDVKYPP	260	4.69
411	425	VLLLESHCA	FAPVLLLESHCAAAR	263	4.77
517	531	VVAFAILIA	VGAVVAFAILIAIVC	266	4.83
412	426	VLLLESHCA	APVLLLESHCAAARD	274	5.03
357	371	EKQILSTVI	LTIFKEKQILSTVIY	275	5.05
518	532	FAILIAIVC	GAVVAFAILIAIVCY	276	5.07
356	370	LTIFKEKQI	ILTIFKEKQILSTVI	277	5.09
514	528	VGAVVAFAI	IGPVGAVVAFAILIA	277	5.09
410	424	VLLLESHCA	EFAPVLLLESHCAAA	278	5.11
515	529	VGAVVAFAI	GPVGAVVAFAILIAI	279	5.14
409	423	VLLLESHCA	VEFAPVLLLESHCAA	285	5.27
56	70	VWYFNSPYP	HGVWYFNSPYPKNYP	287	5.31
97	111	CTLLLSNVS	LRNCTLLLSNVSPEL	287	5.31
463	477	GLVLTSILT	SERSGLVLTSILTLR	290	5.37
408	422	PVLLLESHC	SVEFAPVLLLESHCA	292	5.41
98	112	CTLLLSNVS	RNCTLLLSNVSPELG	302	5.63
462	476	GLVLTSILT	YSERSGLVLTSILTL	303	5.66

166	180	LRPELSWLG	PELRPELSWLGHEGL	308	5.78
131	145	LDIVNTPNI	EHSVLDIVNTPNIVV	310	5.82
132	146	IVNTPNIVV	HSVLDIVNTPNIVVP	316	5.96
162	176	ELRPELSWL	PDNCPELRPELSWLG	317	5.98
163	177	LRPELSWLG	DNCPELRPELSWLGH	320	6.04
164	178	LRPELSWLG	NCPELRPELSWLGHE	322	6.08
94	108	LRNCTLLLS	DLGLRNCTLLLSNVS	329	6.24
133	147	IVNTPNIVV	SVLDIVNTPNIVVPP	329	6.24
2	16	LTALPLFWI	IFLTALPLFWIMISA	331	6.29
165	179	LRPELSWLG	CPELRPELSWLGHEG	333	6.34
466	480	LTSILTLRG	SGLVLTSILTLRGQA	338	6.46
467	481	LTSILTLRG	GLVLTSILTLRGQAQ	338	6.46
519	533	FAILIAIVC	AVVAFAILIAIVCYI	342	6.54
95	109	CTLLLSNVS	LGLRNCTLLLSNVSP	350	6.7
317	331	LSVMYAPWK	TVGLSVMYAPWKPTV	352	6.74
318	332	LSVMYAPWK	VGLSVMYAPWKPTVN	362	6.96
96	110	CTLLLSNVS	GLRNCTLLLSNVSPE	368	7.08
246	260	VEAIEGSHV	NSSVEAIEGSHVSLL	369	7.1
247	261	VEAIEGSHV	SSVEAIEGSHVSLLC	376	7.24
359	373	EKQILSTVI	IFKEKQILSTVIYES	383	7.37
358	372	EKQILSTVI	TIFKEKQILSTVIYE	385	7.41
516	530	VVAFAILIA	PVGAVVAFAILIAIV	385	7.41
316	330	LSVMYAPWK	RTVGLSVMYAPWKPT	386	7.43
26	40	ISAFEGTCV	MPSSISAFEGTCVSI	395	7.61
245	259	VEAIEGSHV	MNSSVEAIEGSHVSL	400	7.71
28	42	ISAFEGTCV	SSISAFEGTCVSIPC	401	7.73
27	41	ISAFEGTCV	PSSISAFEGTCVSIP	403	7.76
468	482	ILTLRGQAQ	LVLTSILTLRGQAQA	403	7.76
196	210	LLHFVPTRE	WVQVSLLHFVPTREA	404	7.79
461	475	VYSERSGLV	VYSERSGLVLTSILT	404	7.79
67	81	PVVFKSRTQ	KNYPPVVFKSRTQVV	408	7.86

11	25	IMISASRGG	WIMISASRGGHWGAW	410	7.9
25	39	ISAFEGTCV	WMPSSISAFEGTCVS	410	7.9
68	82	PVVFKSRTQ	NYPPVVFKSRTQVVH	411	7.92
24	38	PSSISAFEG	AWMPSSISAFEGTCV	412	7.94
315	329	LSVMYAPWK	NRTVGLSVMYAPWKP	413	7.96
487	501	RNLYGAKSL	ICTARNLYGAKSLEL	415	8.01
69	83	PVVFKSRTQ	YPPVVFKSRTQVVHE	420	8.1
134	148	IVNTPNIVV	VLDIVNTPNIVVPPE	422	8.14
135	149	IVNTPNIVV	LDIVNTPNIVVPPEV	428	8.26
488	502	RNLYGAKSL	CTARNLYGAKSLELP	429	8.28
197	211	LHFVPTREA	VQVSLLHFVPTREAN	434	8.39
244	258	VEAIEGSHV	EMNSSVEAIEGSHVS	434	8.39
360	374	EKQILSTVI	FKEKQILSTVIYESE	436	8.43
243	257	VEMNSSVEA	VEMNSSVEAIEGSHV	440	8.51
489	503	RNLYGAKSL	TARNLYGAKSLELPF	452	8.75
513	527	VGAVVAFAI	KIGPVGAVVAFAILI	457	8.85
496	510	ELPFQGAHR	AKSLELPFQGAHRLM	462	8.95
497	511	ELPFQGAHR	KSLELPFQGAHRLMW	467	9.04
198	212	LHFVPTREA	QVSLLHFVPTREANG	471	9.14
314	328	TVGLSVMYA	DNRTVGLSVMYAPWK	472	9.15
495	509	ELPFQGAHR	GAKSLELPFQGAHRL	482	9.36
525	539	LIAIVCYIT	ILIAIVCYITQTRRK	488	9.48
228	242	FEGYASMDV	FEGYASMDVKYPPVI	491	9.54
498	512	ELPFQGAHR	SLELPFQGAHRLMWA	491	9.54
66	80	PVVFKSRTQ	PKNYPPVVFKSRTQV	493	9.58
199	213	LHFVPTREA	VSLLHFVPTREANGH	503	9.79
12	26	IMISASRGG	IMISASRGGHWGAWM	507	9.87
456	470	VYSERSGLV	SEREFVYSERSGLVL	509	9.92
512	526	VGAVVAFAI	AKIGPVGAVVAFAIL	512	9.98
511	525	IGPVGAVVA	WAKIGPVGAVVAFAI	521	10.16
238	252	IVEMNSSVE	YPPVIVEMNSSVEAI	523	10.19

239	253	IVEMNSSVE	PPVIVEMNSSVEAIE	526	10.26
526	540	LIAIVCYIT	LIAIVCYITQTRRKK	529	10.32
405	419	LSVEFAPVL	FNLSVEFAPVLLLES	533	10.41
403	417	LSVEFAPVL	TAFNLSVEFAPVLLL	538	10.52
404	418	LSVEFAPVL	AFNLSVEFAPVLLLE	541	10.58
3	17	LTALPLFWI	FLTALPLFWIMISAS	546	10.68
611	625	LTEELAEYA	YTLTEELAEYAEIRV	547	10.7
612	626	LAEYAEIRV	TLTEELAEYAEIRVK	553	10.81
237	251	IVEMNSSVE	KYPPVIVEMNSSVEA	556	10.87
236	250	VIVEMNSSV	VKYPPVIVEMNSSVE	567	11.05
455	469	REFVYSERS	ESEREFVYSERSGLV	569	11.09
485	499	ARNLYGAKS	RVICTARNLYGAKSL	569	11.09
362	376	EKQILSTVI	EKQILSTVIYESELQ	576	11.23
65	79	YPKNYPPVV	YPKNYPPVVFKSRTQ	581	11.32
428	442	LCVVKSNPE	VQCLCVVKSNPEPSV	583	11.36
457	471	VYSERSGLV	EREFVYSERSGLVLT	603	11.73
458	472	VYSERSGLV	REFVYSERSGLVLTS	605	11.76
57	71	VWYFNSPYP	GVWYFNSPYPKNYPP	607	11.8
429	443	LCVVKSNPE	QCLCVVKSNPEPSVA	607	11.8
195	209	VQVSLLHFV	TWVQVSLLHFVPTRE	615	11.94
469	483	ILTLRGQAQ	VLTSILTLRGQAQAP	619	12.02
402	416	LSVEFAPVL	ATAFNLSVEFAPVLL	620	12.04
427	441	LCVVKSNPE	TVQCLCVVKSNPEPS	627	12.17
99	113	CTLLLSNVS	NCTLLLSNVSPELGG	631	12.23
470	484	ILTLRGQAQ	LTSILTLRGQAQAPP	634	12.29
231	245	DVKYPPVIV	YASMDVKYPPVIVEM	639	12.37
100	114	CTLLLSNVS	CTLLLSNVSPELGGK	641	12.41
551	565	PVLFSSDFR	GDNPPVLFSSDFRIS	646	12.5
425	439	VQCLCVVKS	RDTVQCLCVVKSNPE	647	12.51
550	564	PVLFSSDFR	AGDNPPVLFSSDFRI	647	12.51
229	243	MDVKYPPVI	EGYASMDVKYPPVIV	651	12.59

130	144	LDIVNTPNI	SEHSVLDIVNTPNIV	653	12.62
486	500	RNLYGAKSL	VICTARNLYGAKSLE	653	12.62
504	518	HRLMWAKIG	QGAHRLMWAKIGPVG	653	12.62
230	244	DVKYPPVIV	GYASMDVKYPPVIVE	655	12.66
553	567	PVLFSSDFR	NPPVLFSSDFRISGA	655	12.66
426	440	LCVVKSNPE	DTVQCLCVVKSNPEP	660	12.74
552	566	PVLFSSDFR	DNPPVLFSSDFRISG	664	12.82
503	517	HRLMWAKIG	FQGAHRLMWAKIGPV	673	12.99
490	504	LYGAKSLEL	ARNLYGAKSLELPFQ	680	13.12
319	333	LSVMYAPWK	GLSVMYAPWKPTVNG	686	13.23
365	379	TVIYESELQ	ILSTVIYESELQLEL	687	13.25
505	519	HRLMWAKIG	GAHRLMWAKIGPVGA	698	13.43
363	377	TVIYESELQ	KQILSTVIYESELQL	716	13.79
232	246	DVKYPPVIV	ASMDVKYPPVIVEMN	718	13.82
200	214	LHFVPTREA	SLLHFVPTREANGHR	720	13.85
491	505	LYGAKSLEL	RNLYGAKSLELPFQG	722	13.88
129	143	VLDIVNTPN	FSEHSVLDIVNTPNI	728	13.98
413	427	VLLLESHCA	PVLLLESHCAAARDT	733	14.08
227	241	FEGYASMDV	QFEGYASMDVKYPPV	736	14.14
366	380	TVIYESELQ	LSTVIYESELQLELP	738	14.17
459	473	VYSERSGLV	EFVYSERSGLVLTSI	740	14.21
168	182	LRPELSWLG	LRPELSWLGHEGLGE	743	14.26
577	591	ERRLLGLRG	LGSERRLLGLRGEPP	743	14.26
364	378	TVIYESELQ	QILSTVIYESELQLE	744	14.28
115	129	FRGDLGGYN	YYFRGDLGGYNQYTF	747	14.33
414	428	VLLLESHCA	VLLLESHCAAARDTV	756	14.48
240	254	IVEMNSSVE	PVIVEMNSSVEAIEG	757	14.5
578	592	ERRLLGLRG	GSERRLLGLRGEPPE	761	14.56
167	181	LRPELSWLG	ELRPELSWLGHEGLG	763	14.6
193	207	WVQVSLLHF	EGTWVQVSLLHFVPT	773	14.78
501	515	AHRLMWAKI	LPFQGAHRLMWAKIG	782	14.94

192	206	WVQVSLLHF	DEGTWVQVSLLHFVP	802	15.26
194	208	WVQVSLLHF	GTWVQVSLLHFVPTR	814	15.48
406	420	SVEFAPVLL	NLSVEFAPVLLLESH	817	15.53
471	485	ILTLRGQAQ	TSILTLRGQAQAPPR	831	15.78
494	508	AKSLELPFQ	YGAKSLELPFQGAHR	844	16.02
191	205	WVQVSLLHF	EDEGTWVQVSLLHFV	847	16.07
58	72	VWYFNSPYP	VWYFNSPYPKNYPPV	854	16.18
508	522	WAKIGPVGA	RLMWAKIGPVGAVVA	855	16.19
499	513	LPFQGAHRL	LELPFQGAHRLMWAK	875	16.52
137	151	IVNTPNIVV	IVNTPNIVVPPEVVA	878	16.57
89	103	LLGDLGLRN	SRLLGDLGLRNCTLL	880	16.6
361	375	EKQILSTVI	KEKQILSTVIYESEL	882	16.63
509	523	IGPVGAVVA	LMWAKIGPVGAVVAF	887	16.71
320	334	LSVMYAPWK	LSVMYAPWKPTVNGT	889	16.74
506	520	LMWAKIGPV	AHRLMWAKIGPVGAV	897	16.88
70	84	PVVFKSRTQ	PPVVFKSRTQVVHES	912	17.11
136	150	IVNTPNIVV	DIVNTPNIVVPPEVV	913	17.13
71	85	PVVFKSRTQ	PVVFKSRTQVVHESF	931	17.45
502	516	HRLMWAKIG	PFQGAHRLMWAKIGP	941	17.61
507	521	LMWAKIGPV	HRLMWAKIGPVGAVV	942	17.62
574	588	ERRLGSERR	ERRLGSERRLLGLRG	946	17.68
88	102	LLGDLGLRN	RSRLLGDLGLRNCTL	990	18.4
116	130	LGGYNQYTF	YFRGDLGGYNQYTFS	990	18.4
117	131	LGGYNQYTF	FRGDLGGYNQYTFSE	992	18.43
51	65	AVVHGVWYF	RPAVVHGVWYFNSPY	1000	18.54
233	247	DVKYPPVIV	SMDVKYPPVIVEMNS	1001	18.55
575	589	ERRLLGLRG	RRLGSERRLLGLRGE	1003	18.58
248	262	VEAIEGSHV	SVEAIEGSHVSLLCG	1008	18.65
500	514	ELPFQGAHR	ELPFQGAHRLMWAKI	1012	18.71
101	115	LLSNVSPEL	TLLLSNVSPELGGKY	1021	18.84
484	498	ICTARNLYG	PRVICTARNLYGAKS	1031	18.99

549	563	SAGDNPPVL	SAGDNPPVLFSSDFR	1046	19.22
579	593	LLGLRGEPP	SERRLLGLRGEPPEL	1051	19.29
190	204	TWVQVSLLH	REDEGTWVQVSLLHF	1058	19.39
332	346	MVAVEGETV	NGTMVAVEGETVSIL	1078	19.67
249	263	VEAIEGSHV	VEAIEGSHVSLLCGA	1079	19.69
510	524	IGPVGAVVA	MWAKIGPVGAVVAFA	1079	19.69
554	568	VLFSSDFRI	PPVLFSSDFRISGAP	1080	19.7
576	590	ERRLLGLRG	RLGSERRLLGLRGEP	1081	19.71
401	415	TAFNLSVEF	RATAFNLSVEFAPVL	1093	19.89
482	496	ICTARNLYG	APPRVICTARNLYGA	1094	19.9
235	249	DVKYPPVIV	DVKYPPVIVEMNSSV	1097	19.95
333	347	MVAVEGETV	GTMVAVEGETVSILC	1101	19.99
580	594	LLGLRGEPP	ERRLLGLRGEPPELD	1103	20.02
481	495	RVICTARNL	QAPPRVICTARNLYG	1110	20.12
369	383	ELQLELPAV	VIYESELQLELPAVS	1113	20.16
212	226	LGCQASFPN	GHRLGCQASFPNTTL	1120	20.28
86	100	LLGDLGLRN	QGRSRLLGDLGLRNC	1128	20.39
483	497	ICTARNLYG	PPRVICTARNLYGAK	1131	20.43
340	354	VSILCSTQS	GETVSILCSTQSNPD	1135	20.49
267	281	LTWMRDGTV	PPPLLTWMRDGTVLR	1136	20.51
460	474	VYSERSGLV	FVYSERSGLVLTSIL	1136	20.51
439	453	AFELPSRNV	EPSVAFELPSRNVTV	1139	20.55
370	384	LQLELPAVS	IYESELQLELPAVSP	1150	20.69
371	385	LQLELPAVS	YESELQLELPAVSPE	1158	20.81
213	227	LGCQASFPN	HRLGCQASFPNTTLQ	1161	20.85
220	234	TLQFEGYAS	SFPNTTLQFEGYASM	1161	20.85
29	43	ISAFEGTCV	SISAFEGTCVSIPCR	1166	20.92
30	44	ISAFEGTCV	ISAFEGTCVSIPCRF	1176	21.05
407	421	LSVEFAPVL	LSVEFAPVLLLESHC	1182	21.13
221	235	TLQFEGYAS	FPNTTLQFEGYASMD	1184	21.16
331	345	MVAVEGETV	VNGTMVAVEGETVSI	1186	21.19

118	132	LGGYNQYTF	RGDLGGYNQYTFSEH	1188	21.21
91	105	LLGDLGLRN	LLGDLGLRNCTLLLS	1192	21.26
372	386	LQLELPAVS	ESELQLELPAVSPED	1193	21.28
268	282	LTWMRDGTV	PPLLTWMRDGTVLRE	1194	21.29
85	99	SRLLGDLGL	FQGRSRLLGDLGLRN	1205	21.44
341	355	VSILCSTQS	ETVSILCSTQSNPDP	1206	21.45
440	454	AFELPSRNV	PSVAFELPSRNVTVN	1212	21.53
87	101	LLGDLGLRN	GRSRLLGDLGLRNCT	1229	21.77
269	283	LTWMRDGTV	PLLTWMRDGTVLREA	1231	21.8
329	343	VNGTMVAVE	PTVNGTMVAVEGETV	1255	22.12
330	344	MVAVEGETV	TVNGTMVAVEGETVS	1256	22.14
438	452	AFELPSRNV	PEPSVAFELPSRNVT	1264	22.24
338	352	VSILCSTQS	VEGETVSILCSTQSN	1268	22.28
47	61	PDELRPAVV	PDELRPAVVHGVWYF	1269	22.3
337	351	TVSILCSTQ	AVEGETVSILCSTQS	1272	22.35
339	353	VSILCSTQS	EGETVSILCSTQSNP	1291	22.59
373	387	LQLELPAVS	SELQLELPAVSPEDD	1312	22.86
581	595	LLGLRGEPP	RRLLGLRGEPPELDL	1317	22.92
472	486	ILTLRGQAQ	SILTLRGQAQAPPRV	1322	22.99
119	133	LGGYNQYTF	GDLGGYNQYTFSEHS	1324	23.02
441	455	AFELPSRNV	SVAFELPSRNVTVNE	1329	23.09
111	125	KYYFRGDLG	LGGKYYFRGDLGGYN	1348	23.33
437	451	VAFELPSRN	NPEPSVAFELPSRNV	1356	23.43
90	104	LLGDLGLRN	RLLGDLGLRNCTLLL	1359	23.46
473	487	GQAQAPPRV	ILTLRGQAQAPPRVI	1379	23.71
241	255	IVEMNSSVE	VIVEMNSSVEAIEGS	1395	23.91
201	215	LHFVPTREA	LLHFVPTREANGHRL	1402	24
431	445	LCVVKSNPE	LCVVKSNPEPSVAFE	1404	24.02
430	444	LCVVKSNPE	CLCVVKSNPEPSVAF	1411	24.11
49	63	AVVHGVWYF	ELRPAVVHGVWYFNS	1415	24.16
50	64	AVVHGVWYF	LRPAVVHGVWYFNSP	1450	24.57

138	152	PNIVVPPEV	VNTPNIVVPPEVVAG	1503	25.21
454	468	REFVYSERS	NESEREFVYSERSGL	1508	25.27
527	541	IVCYITQTR	IAIVCYITQTRRKKN	1529	25.51
139	153	PNIVVPPEV	NTPNIVVPPEVVAGT	1546	25.71
112	126	FRGDLGGYN	GGKYYFRGDLGGYNQ	1559	25.86
219	233	TTLQFEGYA	ASFPNTTLQFEGYAS	1607	26.42
480	494	RVICTARNL	AQAPPRVICTARNLY	1667	27.11
114	128	FRGDLGGYN	KYYFRGDLGGYNQYT	1676	27.23
367	381	TVIYESELQ	STVIYESELQLELPA	1693	27.4
556	570	VLFSSDFRI	VLFSSDFRISGAPEK	1700	27.48
48	62	AVVHGVWYF	DELRPAVVHGVWYFN	1706	27.55
210	224	LGCQASFPN	ANGHRLGCQASFPNT	1712	27.61
209	223	GHRLGCQAS	EANGHRLGCQASFPN	1734	27.86
271	285	WMRDGTVLR	LTWMRDGTVLREAVA	1738	27.9
572	586	ERRLGSERR	ESERRLGSERRLLGL	1743	27.96
312	326	TVGLSVMYA	GQDNRTVGLSVMYAP	1753	28.07
265	279	PPPLLTWMR	SNPPPLLTWMRDGTV	1766	28.21
571	585	ERRLGSERR	YESERRLGSERRLLG	1769	28.24
311	325	DNRTVGLSV	YGQDNRTVGLSVMYA	1777	28.34
266	280	LTWMRDGTV	NPPPLLTWMRDGTVL	1799	28.58
555	569	PVLFSSDFR	PVLFSSDFRISGAPE	1802	28.61
92	106	DLGLRNCTL	LGDLGLRNCTLLLSN	1811	28.71
113	127	FRGDLGGYN	GKYYFRGDLGGYNQY	1811	28.71
102	116	LLSNVSPEL	LLLSNVSPELGGKYY	1874	29.37
211	225	LGCQASFPN	NGHRLGCQASFPNTT	1891	29.54
214	228	QASFPNTTL	RLGCQASFPNTTLQF	1895	29.59
570	584	ERRLGSERR	KYESERRLGSERRLL	1902	29.65
234	248	DVKYPPVIV	MDVKYPPVIVEMNSS	1908	29.7
368	382	TVIYESELQ	TVIYESELQLELPAV	1908	29.7
215	229	QASFPNTTL	LGCQASFPNTTLQFE	1926	29.89
60	74	YPKNYPPVV	YFNSPYPKNYPPVVF	1962	30.26

270	284	WMRDGTVLR	LLTWMRDGTVLREAV	1973	30.36
121	135	LGGYNQYTF	LGGYNQYTFSEHSVL	1982	30.45
61	75	YPKNYPPVV	FNSPYPKNYPPVVFK	2012	30.73
474	488	GQAQAPPRV	LTLRGQAQAPPRVIC	2038	31
313	327	TVGLSVMYA	QDNRTVGLSVMYAPW	2045	31.06
589	603	LDLSYSHSD	EPPELDLSYSHSDLG	2060	31.2
569	583	ERRLGSERR	EKYESERRLGSERRL	2078	31.39
93	107	GLRNCTLLL	GDLGLRNCTLLLSNV	2106	31.68
493	507	LYGAKSLEL	LYGAKSLELPFQGAH	2111	31.73
202	216	LHFVPTREA	LHFVPTREANGHRLG	2117	31.78
62	76	YPKNYPPVV	NSPYPKNYPPVVFKS	2126	31.88
557	571	FRISGAPEK	LFSSDFRISGAPEKY	2130	31.92
282	296	EAVAESLLL	EAVAESLLLELEEVT	2157	32.17
559	573	FRISGAPEK	SSDFRISGAPEKYES	2163	32.23
63	77	YPKNYPPVV	SPYPKNYPPVVFKSR	2169	32.28
558	572	FRISGAPEK	FSSDFRISGAPEKYE	2171	32.3
591	605	LSYSHSDLG	PELDLSYSHSDLGKR	2194	32.51
108	122	GKYYFRGDL	SPELGGKYYFRGDLG	2198	32.54
284	298	LLLELEEVT	VAESLLLELEEVTPA	2207	32.62
479	493	PRVICTARN	QAQAPPRVICTARNL	2214	32.68
285	299	LLLELEEVT	AESLLLELEEVTPAE	2231	32.83
19	33	GAWMPSSIS	GGHWGAWMPSSISAF	2239	32.91
109	123	KYYFRGDLG	PELGGKYYFRGDLGG	2240	32.92
590	604	LSYSHSDLG	PPELDLSYSHSDLGK	2241	32.93
21	35	GAWMPSSIS	HWGAWMPSSISAFEG	2243	32.94
20	34	GAWMPSSIS	GHWGAWMPSSISAFE	2269	33.18
103	117	LLSNVSPEL	LLSNVSPELGGKYYF	2271	33.19
283	297	LLLELEEVT	AVAESLLLELEEVTP	2284	33.31
59	73	WYFNSPYPK	WYFNSPYPKNYPPVV	2291	33.37
140	154	NIVVPPEVV	TPNIVVPPEVVAGTE	2298	33.43
568	582	YESERRLGS	PEKYESERRLGSERR	2324	33.66

321	335	YAPWKPTVN	SVMYAPWKPTVNGTM	2326	33.68
251	265	HVSLLCGAD	AIEGSHVSLLCGADS	2348	33.87
528	542	VCYITQTRR	AIVCYITQTRRKKNV	2348	33.87
492	506	LYGAKSLEL	NLYGAKSLELPFQGA	2349	33.88
286	300	LLLELEEVT	ESLLLELEEVTPAED	2351	33.89
120	134	LGGYNQYTF	DLGGYNQYTFSEHSV	2354	33.92
254	268	VSLLCGADS	GSHVSLLCGADSNPP	2369	34.05
529	543	IVCYITQTR	IVCYITQTRRKKNVT	2387	34.21
252	266	VSLLCGADS	IEGSHVSLLCGADSN	2397	34.29
342	356	VSILCSTQS	TVSILCSTQSNPDPI	2419	34.47
343	357	VSILCSTQS	VSILCSTQSNPDPIL	2434	34.61
442	456	AFELPSRNV	VAFELPSRNVTVNES	2455	34.77
453	467	REFVYSERS	VNESEREFVYSERSG	2478	34.96
475	489	GQAQAPPRV	TLRGQAQAPPRVICT	2487	35.03
255	269	VSLLCGADS	SHVSLLCGADSNPPP	2488	35.04
253	267	VSLLCGADS	EGSHVSLLCGADSNP	2505	35.18
536	550	KNVTESPSF	TRRKKNVTESPSFSA	2527	35.35
560	574	FRISGAPEK	SDFRISGAPEKYESE	2541	35.46
110	124	KYYFRGDLG	ELGGKYYFRGDLGGY	2564	35.64
537	551	VTESPSFSA	RRKKNVTESPSFSAG	2605	35.97
78	92	HESFQGRSR	TQVVHESFQGRSRLL	2623	36.11
476	490	GQAQAPPRV	LRGQAQAPPRVICTA	2627	36.14
79	93	HESFQGRSR	QVVHESFQGRSRLLG	2641	36.26
452	466	VNESEREFV	TVNESEREFVYSERS	2669	36.47
538	552	VTESPSFSA	RKKNVTESPSFSAGD	2679	36.56
123	137	YTFSEHSVL	GYNQYTFSEHSVLDI	2721	36.87
592	606	LSYSHSDLG	ELDLSYSHSDLGKRP	2736	37
18	32	GAWMPSSIS	RGGHWGAWMPSSISA	2813	37.56
44	58	PDELRPAVV	FDFPDELRPAVVHGV	2833	37.71
45	59	PDELRPAVV	DFPDELRPAVVHGVW	2866	37.95
122	136	YTFSEHSVL	GGYNQYTFSEHSVLD	2875	38.01

80	94	HESFQGRSR	VVHESFQGRSRLLGD	2889	38.12
242	256	IVEMNSSVE	IVEMNSSVEAIEGSH	2913	38.3
17	31	WGAWMPSSI	SRGGHWGAWMPSSIS	2935	38.46
276	290	LREAVAESL	DGTVLREAVAESLLL	2940	38.49
83	97	RSRLLGDLG	ESFQGRSRLLGDLGL	2970	38.72
588	602	LDLSYSHSD	GEPPELDLSYSHSDL	2982	38.8
77	91	HESFQGRSR	RTQVVHESFQGRSRL	2984	38.81
277	291	LREAVAESL	GTVLREAVAESLLLE	2984	38.81
13	27	ISASRGGHW	MISASRGGHWGAWMP	2994	38.88
322	336	WKPTVNGTM	VMYAPWKPTVNGTMV	2999	38.91
539	553	VTESPSFSA	KKNVTESPSFSAGDN	3009	38.99
443	457	AFELPSRNV	AFELPSRNVTVNESE	3039	39.19
350	364	PDPILTIFK	QSNPDPILTIFKEKQ	3070	39.41
334	348	MVAVEGETV	TMVAVEGETVSILCS	3071	39.42
393	407	ENQYGQRAT	VAENQYGQRATAFNL	3087	39.52
530	544	YITQTRRKK	VCYITQTRRKKNVTE	3099	39.59
610	624	LTEELAEYA	SYTLTEELAEYAEIR	3118	39.72
335	349	MVAVEGETV	MVAVEGETVSILCST	3131	39.8
540	554	VTESPSFSA	KNVTESPSFSAGDNP	3153	39.95
278	292	LREAVAESL	TVLREAVAESLLLEL	3164	40.02
128	142	HSVLDIVNT	TFSEHSVLDIVNTPN	3166	40.03
14	28	ISASRGGHW	ISASRGGHWGAWMPS	3177	40.1
582	596	LLGLRGEPP	RLLGLRGEPPELDLS	3180	40.12
593	607	LSYSHSDLG	LDLSYSHSDLGKRPT	3216	40.35
82	96	SFQGRSRLL	HESFQGRSRLLGDLG	3228	40.43
275	289	LREAVAESL	RDGTVLREAVAESLL	3275	40.73
273	287	GTVLREAVA	WMRDGTVLREAVAES	3287	40.81
22	36	WMPSSISAF	WGAWMPSSISAFEGT	3293	40.85
392	406	AENQYGQRA	CVAENQYGQRATAFN	3307	40.93
299	313	VYACLAENA	EDGVYACLAENAYGQ	3322	41.02
105	119	VSPELGGKY	SNVSPELGGKYYFRG	3325	41.04

124	138	YTFSEHSVL	YNQYTFSEHSVLDIV	3330	41.07
84	98	SRLLGDLGL	SFQGRSRLLGDLGLR	3371	41.34
436	450	SVAFELPSR	SNPEPSVAFELPSRN	3381	41.4
274	288	GTVLREAVA	MRDGTVLREAVAESL	3383	41.41
300	314	VYACLAENA	DGVYACLAENAYGQD	3388	41.45
76	90	QVVHESFQG	SRTQVVHESFQGRSR	3405	41.57
216	230	QASFPNTTL	GCQASFPNTTLQFEG	3416	41.63
272	286	GTVLREAVA	TWMRDGTVLREAVAE	3446	41.82
573	587	ERRLGSERR	SERRLGSERRLLGLR	3472	41.98
141	155	PNIVVPPEV	PNIVVPPEVVAGTEV	3542	42.41
374	388	LQLELPAVS	ELQLELPAVSPEDDG	3543	42.41
64	78	PKNYPPVVF	PYPKNYPPVVFKSRT	3544	42.42
298	312	VYACLAENA	AEDGVYACLAENAYG	3548	42.44
107	121	LGGKYYFRG	VSPELGGKYYFRGDL	3589	42.68
395	409	GQRATAFNL	ENQYGQRATAFNLSV	3600	42.74
391	405	AENQYGQRA	WCVAENQYGQRATAF	3613	42.82
23	37	WMPSSISAF	GAWMPSSISAFEGTC	3636	42.96
394	408	GQRATAFNL	AENQYGQRATAFNLS	3639	42.97
125	139	YTFSEHSVL	NQYTFSEHSVLDIVN	3659	43.1
609	623	LTEELAEYA	DSYTLTEELAEYAEI	3698	43.32
43	57	PDELRPAVV	RFDFPDELRPAVVHG	3724	43.47
81	95	SFQGRSRLL	VHESFQGRSRLLGDL	3739	43.56
308	322	AYGQDNRTV	ENAYGQDNRTVGLSV	3743	43.57
607	621	SYTLTEELA	TKDSYTLTEELAEYA	3752	43.64
297	311	VYACLAENA	PAEDGVYACLAENAY	3774	43.77
310	324	DNRTVGLSV	AYGQDNRTVGLSVMY	3795	43.89
608	622	LTEELAEYA	KDSYTLTEELAEYAE	3812	43.98
583	597	LLGLRGEPP	LLGLRGEPPELDLSY	3849	44.18
41	55	RFDFPDELR	PCRFDFPDELRPAVV	3865	44.27
397	411	GQRATAFNL	QYGQRATAFNLSVEF	3867	44.28
42	56	PDELRPAVV	CRFDFPDELRPAVVH	3895	44.42

542	556	VTESPSFSA	VTESPSFSAGDNPPV	3906	44.48
104	118	VSPELGGKY	LSNVSPELGGKYYFR	3918	44.55
169	183	LSWLGHEGL	RPELSWLGHEGLGEP	3981	44.89
73	87	VFKSRTQVV	VFKSRTQVVHESFQG	3982	44.89
189	203	LREDEGTWV	LREDEGTWVQVSLLH	3984	44.91
279	293	EAVAESLLL	VLREAVAESLLLELE	4013	45.06
375	389	LQLELPAVS	LQLELPAVSPEDDGE	4027	45.14
170	184	LSWLGHEGL	PELSWLGHEGLGEPA	4070	45.36
544	558	FSAGDNPPV	ESPSFSAGDNPPVLF	4138	45.7
545	559	FSAGDNPPV	SPSFSAGDNPPVLFS	4203	46.04
389	403	AENQYGQRA	EYWCVAENQYGQRAT	4245	46.25
46	60	LRPAVVHGV	FPDELRPAVVHGVWY	4256	46.31
546	560	FSAGDNPPV	PSFSAGDNPPVLFSS	4259	46.32
543	557	FSAGDNPPV	TESPSFSAGDNPPVL	4294	46.51
562	576	FRISGAPEK	FRISGAPEKYESERR	4316	46.62
478	492	GQAQAPPRV	GQAQAPPRVICTARN	4319	46.64
587	601	PELDLSYSH	RGEPPELDLSYSHSD	4333	46.7
396	410	GQRATAFNL	NQYGQRATAFNLSVE	4346	46.76
323	337	WKPTVNGTM	MYAPWKPTVNGTMVA	4407	47.06
531	545	QTRRKKNVT	CYITQTRRKKNVTES	4501	47.5
280	294	EAVAESLLL	LREAVAESLLLELEE	4555	47.76
106	120	LGGKYYFRG	NVSPELGGKYYFRGD	4619	48.06
172	186	WLGHEGLGE	LSWLGHEGLGEPAVL	4684	48.37
309	323	DNRTVGLSV	NAYGQDNRTVGLSVM	4770	48.75
325	339	WKPTVNGTM	APWKPTVNGTMVAVE	4792	48.85
432	446	VKSNPEPSV	CVVKSNPEPSVAFEL	4836	49.05
324	338	WKPTVNGTM	YAPWKPTVNGTMVAV	4847	49.09
75	89	QVVHESFQG	KSRTQVVHESFQGRS	4902	49.35
532	546	QTRRKKNVT	YITQTRRKKNVTESP	4941	49.52
399	413	TAFNLSVEF	GQRATAFNLSVEFAP	4975	49.67
257	271	VSLLCGADS	VSLLCGADSNPPPLL	4996	49.77

3. OSP

Start	End	Core Nonamer	15mer Peptide	IC_{50}	Percentile Rank
123	137	LLILLALCA	AGVLLILLALCALVA	33	0.2
124	138	LLILLALCA	GVLLILLALCALVAT	34	0.21
122	136	LLILLALCA	LAGVLLILLALCALV	35	0.23
121	135	LLILLALCA	QLAGVLLILLALCAL	39	0.28
120	134	LAGVLLILL	AQLAGVLLILLALCA	42	0.33
66	80	ILILPGYVQ	KPLVDILILPGYVQA	56	0.52
68	82	ILILPGYVQ	LVDILILPGYVQACR	56	0.52
67	81	ILILPGYVQ	PLVDILILPGYVQAC	57	0.54
65	79	DILILPGYV	CKPLVDILILPGYVQ	58	0.56
125	139	LLILLALCA	VLLILLALCALVATI	69	0.73
69	83	ILILPGYVQ	VDILILPGYVQACRA	70	0.75
126	140	LLILLALCA	LLILLALCALVATIW	74	0.81
86	100	VLGLPAILL	IAASVLGLPAILLLL	96	1.2
87	101	VLGLPAILL	AASVLGLPAILLLLT	96	1.2
88	102	VLGLPAILL	ASVLGLPAILLLLTV	100	1.27
85	99	VLGLPAILL	MIAASVLGLPAILLL	104	1.35
84	98	SVLGLPAIL	LMIAASVLGLPAILL	111	1.48
6	20	VVGFVTSFV	LQVVGFVTSFVGWIG	114	1.54
5	19	VVGFVTSFV	CLQVVGFVTSFVGWI	118	1.61
3	17	VVGFVTSFV	ATCLQVVGFVTSFVG	123	1.69
4	18	VVGFVTSFV	TCLQVVGFVTSFVGW	126	1.75
127	141	LLALCALVA	LILLALCALVATIWF	144	2.11
190	204	LRALAPRLM	STTLRALAPRLMRRV	149	2.21
150	164	VSFGYSLYA	TIVSFGYSLYAGWIG	159	2.42
189	203	LRALAPRLM	VSTTLRALAPRLMRR	159	2.42
188	202	LRALAPRLM	NVSTTLRALAPRLMR	164	2.52
187	201	TLRALAPRL	ENVSTTLRALAPRLM	169	2.62
2	16	CLQVVGFVT	VATCLQVVGFVTSFV	172	2.68

70	84	ILILPGYVQ	DILILPGYVQACRAL	173	2.71
7	21	VGFVTSFVG	QVVGFVTSFVGWIGV	177	2.79
191	205	LRALAPRLM	TTLRALAPRLMRRVP	179	2.83
90	104	VLGLPAILL	VLGLPAILLLLTVLP	185	2.97
153	167	YSLYAGWIG	SFGYSLYAGWIGAVL	185	2.97
154	168	YSLYAGWIG	FGYSLYAGWIGAVLC	196	3.26
152	166	YSLYAGWIG	VSFGYSLYAGWIGAV	197	3.28
64	78	VDILILPGY	HCKPLVDILILPGYV	198	3.31
151	165	YSLYAGWIG	IVSFGYSLYAGWIGA	199	3.33
71	85	ILILPGYVQ	ILILPGYVQACRALM	203	3.42
89	103	VLGLPAILL	SVLGLPAILLLLTVL	212	3.61
79	93	LMIAASVLG	QACRALMIAASVLGL	220	3.78
80	94	LMIAASVLG	ACRALMIAASVLGLP	224	3.87
78	92	ALMIAASVL	VQACRALMIAASVLG	225	3.89
81	95	LMIAASVLG	CRALMIAASVLGLPA	234	4.09
9	23	VTSFVGWIG	VGFVTSFVGWIGVIV	252	4.51
146	160	IVSFGYSLY	HRETTIVSFGYSLYA	252	4.51
8	22	VTSFVGWIG	VVGFVTSFVGWIGVI	256	4.59
92	106	ILLLLTVLP	GLPAILLLLTVLPCI	267	4.86
93	107	ILLLLTVLP	LPAILLLLTVLPCIR	268	4.88
94	108	ILLLLTVLP	PAILLLLTVLPCIRM	269	4.91
147	161	VSFGYSLYA	RETTIVSFGYSLYAG	273	5.01
82	96	LMIAASVLG	RALMIAASVLGLPAI	276	5.07
148	162	VSFGYSLYA	ETTIVSFGYSLYAGW	280	5.16
129	143	LLALCALVA	LLALCALVATIWFPV	300	5.59
128	142	LLALCALVA	ILLALCALVATIWFP	303	5.66
10	24	VTSFVGWIG	GFVTSFVGWIGVIVT	307	5.76
116	130	RRAQLAGVL	KYRRAQLAGVLLILL	307	5.76
149	163	VSFGYSLYA	TTIVSFGYSLYAGWI	309	5.8
91	105	ILLLLTVLP	LGLPAILLLLTVLPC	316	5.96
119	133	LAGVLLILL	RAQLAGVLLILLALC	322	6.08
118	132	LAGVLLILL	RRAQLAGVLLILLAL	352	6.74
117	131	LAGVLLILL	YRRAQLAGVLLILLA	360	6.92

192	206	LRALAPRLM	TLRALAPRLMRRVPT	370	7.12
77	91	RALMIAASV	YVQACRALMIAASVL	380	7.32
195	209	LMRRVPTYK	ALAPRLMRRVPTYKR	381	7.34
196	210	MRRVPTYKR	LAPRLMRRVPTYKRA	382	7.35
95	109	LTVLPCIRM	AILLLLTVLPCIRMG	393	7.57
193	207	LRALAPRLM	LRALAPRLMRRVPTY	406	7.82
96	110	LLLTVLPCI	ILLLLTVLPCIRMGQ	425	8.21
197	211	MRRVPTYKR	APRLMRRVPTYKRAA	425	8.21
198	212	MRRVPTYKR	PRLMRRVPTYKRAAR	432	8.35
155	169	YSLYAGWIG	GYSLYAGWIGAVLCL	454	8.78
145	159	ETTIVSFGY	AHRETTIVSFGYSLY	465	9.01
156	170	YSLYAGWIG	YSLYAGWIGAVLCLV	470	9.11
76	90	VQACRALMI	GYVQACRALMIAASV	477	9.26
199	213	MRRVPTYKR	RLMRRVPTYKRAARL	482	9.36
11	25	FVGWIGVIV	FVTSFVGWIGVIVTT	488	9.48
12	26	FVGWIGVIV	VTSFVGWIGVIVTTS	510	9.94
13	27	FVGWIGVIV	TSFVGWIGVIVTTST	521	10.16
130	144	LCALVATIW	LALCALVATIWFPVC	526	10.26
112	126	VAKYRRAQL	PGVAKYRRAQLAGVL	550	10.75
194	208	LAPRLMRRV	RALAPRLMRRVPTYK	560	10.94
72	86	YVQACRALM	LILPGYVQACRALMI	562	10.98
83	97	LMIAASVLG	ALMIAASVLGLPAIL	568	11.07
74	88	VQACRALMI	LPGYVQACRALMIAA	571	11.13
73	87	VQACRALMI	ILPGYVQACRALMIA	598	11.64
75	89	VQACRALMI	PGYVQACRALMIAAS	601	11.69
161	175	VLCLVGGCV	GWIGAVLCLVGGCVI	608	11.82
113	127	RRAQLAGVL	GVAKYRRAQLAGVLL	635	12.3
114	128	RRAQLAGVL	VAKYRRAQLAGVLLI	652	12.6
115	129	RRAQLAGVL	AKYRRAQLAGVLLIL	671	12.96
162	176	VLCLVGGCV	WIGAVLCLVGGCVIL	722	13.88
42	56	LDELGSKGL	RKLDELGSKGLWADC	724	13.91
41	55	LDELGSKGL	CRKLDELGSKGLWAD	733	14.08
40	54	LDELGSKGL	TCRKLDELGSKGLWA	734	14.09

97	111	LTVLPCIRM	LLLLTVLPCIRMGQE	742	14.24
163	177	VLCLVGGCV	IGAVLCLVGGCVILC	744	14.28
131	145	LVATIWFPV	ALCALVATIWFPVCA	753	14.43
38	52	IPTCRKLDE	IPTCRKLDELGSKGL	758	14.51
39	53	LDELGSKGL	PTCRKLDELGSKGLW	776	14.84
132	146	LVATIWFPV	LCALVATIWFPVCAH	792	15.1
160	174	IGAVLCLVG	AGWIGAVLCLVGGCV	799	15.21
14	28	IGVIVTTST	SFVGWIGVIVTTSTN	800	15.23
1	15	CLQVVGFVT	MVATCLQVVGFVTSF	820	15.59
164	178	VLCLVGGCV	GAVLCLVGGCVILCC	842	15.98
186	200	TTLRALAPR	GENVSTTLRALAPRL	853	16.16
133	147	LVATIWFPV	CALVATIWFPVCAHR	879	16.59
157	171	YAGWIGAVL	SLYAGWIGAVLCLVG	888	16.73
201	215	MRRVPTYKR	MRRVPTYKRAARLPT	896	16.86
15	29	IGVIVTTST	FVGWIGVIVTTSTND	932	17.47
200	214	MRRVPTYKR	LMRRVPTYKRAARLP	932	17.47
111	125	VAKYRRAQL	EPGVAKYRRAQLAGV	942	17.62
63	77	LVDILILPG	YHCKPLVDILILPGY	968	18.04
144	158	ETTIVSFGY	CAHRETTIVSFGYSL	1011	18.69
109	123	VAKYRRAQL	GQEPGVAKYRRAQLA	1044	19.19
98	112	LTVLPCIRM	LLLTVLPCIRMGQEP	1045	19.2
110	124	VAKYRRAQL	QEPGVAKYRRAQLAG	1078	19.67
27	41	WVVTCGYTI	TNDWVVTCGYTIPTC	1112	20.15
26	40	WVVTCGYTI	STNDWVVTCGYTIPT	1127	20.37
28	42	WVVTCGYTI	NDWVVTCGYTIPTCR	1185	21.17
24	38	DWVVTCGYT	TTSTNDWVVTCGYTI	1195	21.31
25	39	WVVTCGYTI	TSTNDWVVTCGYTIP	1201	21.39
57	71	GLYHCKPLV	VMATGLYHCKPLVDI	1286	22.52
56	70	GLYHCKPLV	CVMATGLYHCKPLVD	1301	22.71
108	122	EPGVAKYRR	MGQEPGVAKYRRAQL	1320	22.96
158	172	IGAVLCLVG	LYAGWIGAVLCLVGG	1325	23.03
16	30	IGVIVTTST	VGWIGVIVTTSTNDW	1345	23.29
58	72	GLYHCKPLV	MATGLYHCKPLVDIL	1368	23.58

17	31	IGVIVTTST	GWIGVIVTTSTNDWV	1410	24.1
44	58	LDELGSKGL	LDELGSKGLWADCVM	1458	24.67
159	173	IGAVLCLVG	YAGWIGAVLCLVGGC	1465	24.75
204	218	YKRAARLPT	VPTYKRAARLPTEVL	1484	24.99
55	69	TGLYHCKPL	DCVMATGLYHCKPLV	1488	25.03
43	57	LDELGSKGL	KLDELGSKGLWADCV	1490	25.06
59	73	GLYHCKPLV	ATGLYHCKPLVDILI	1500	25.17
62	76	CKPLVDILI	LYHCKPLVDILILPG	1534	25.57
165	179	LCLVGGCVI	AVLCLVGGCVILCCA	1605	26.4
202	216	PTYKRAARL	RRVPTYKRAARLPTE	1618	26.54
134	148	LVATIWFPV	ALVATIWFPVCAHRE	1685	27.32
203	217	PTYKRAARL	RVPTYKRAARLPTEV	1774	28.3
54	68	CVMATGLYH	ADCVMATGLYHCKPL	1775	28.31
135	149	LVATIWFPV	LVATIWFPVCAHRET	1847	29.09
143	157	ETTIVSFGY	VCAHRETTIVSFGYS	1923	29.86
60	74	CKPLVDILI	TGLYHCKPLVDILIL	2020	30.82
18	32	IVTTSTNDW	WIGVIVTTSTNDWVV	2111	31.73
142	156	AHRETTIVS	PVCAHRETTIVSFGY	2214	32.68
19	33	IVTTSTNDW	IGVIVTTSTNDWVVT	2238	32.9
166	180	VLCLVGGCV	VLCLVGGCVILCCAG	2379	34.14
52	66	VMATGLYHC	LWADCVMATGLYHCK	2423	34.51
46	60	SKGLWADCV	ELGSKGLWADCVMAT	2430	34.57
47	61	SKGLWADCV	LGSKGLWADCVMATG	2469	34.88
51	65	CVMATGLYH	GLWADCVMATGLYHC	2478	34.96
100	114	LTVLPCIRM	LTVLPCIRMGQEPGV	2556	35.58
30	44	WVVTCGYTI	WVVTCGYTIPTCRKL	2607	35.98
99	113	LTVLPCIRM	LLTVLPCIRMGQEPG	2693	36.66
29	43	WVVTCGYTI	DWVVTCGYTIPTCRK	2751	37.11
179	193	AQAFGENVS	AGDAQAFGENVSTTL	2802	37.48
53	67	VMATGLYHC	WADCVMATGLYHCKP	2849	37.82
180	194	AQAFGENVS	GDAQAFGENVSTTLR	2850	37.83
61	75	CKPLVDILI	GLYHCKPLVDILILP	3061	39.35
20	34	IVTTSTNDW	GVIVTTSTNDWVVTC	3095	39.56

103	117	IRMGQEPGV	LPCIRMGQEPGVAKY	3189 40.17
102	116	IRMGQEPGV	VLPCIRMGQEPGVAK	3214 40.33
45	59	SKGLWADCV	DELGSKGLWADCVMA	3217 40.35
101	115	IRMGQEPGV	TVLPCIRMGQEPGVA	3307 40.93
50	64	LWADCVMAT	KGLWADCVMATGLYH	3519 42.27
104	118	IRMGQEPGV	PCIRMGQEPGVAKYR	3577 42.61
183	197	FGENVSTTL	QAFGENVSTTLRALA	3684 43.24
185	199	VSTTLRALA	FGENVSTTLRALAPR	3829 44.08
49	63	SKGLWADCV	SKGLWADCVMATGLY	3928 44.6
182	196	FGENVSTTL	AQAFGENVSTTLRAL	3933 44.63
137	151	IWFPVCAHR	ATIWFPVCAHRETTI	4087 45.45
181	195	FGENVSTTL	DAQAFGENVSTTLRA	4188 45.96
48	62	SKGLWADCV	GSKGLWADCVMATGL	4576 47.86
171	185	VILCCAGDA	GGCVILCCAGDAQAF	4598 47.96
184	198	VSTTLRALA	AFGENVSTTLRALAP	4650 48.2
170	184	VILCCAGDA	VGGCVILCCAGDAQA	4672 48.31
178	192	AQAFGENVS	CAGDAQAFGENVSTT	4681 48.35
169	183	VILCCAGDA	LVGGCVILCCAGDAQ	4806 48.91
177	191	AQAFGENVS	CCAGDAQAFGENVST	4833 49.04
168	182	CVILCCAGD	CLVGGCVILCCAGDA	4882 49.26

4. βSYN

Start	End	Core Nonamer	15mer Peptide	IC_{50}	Percentile Rank
34	48	VLYVGSKTR	KEGVLYVGSKTREGV	279	5.14
31	45	GVLYVGSKT	EKTKEGVLYVGSKTR	284	5.24
32	46	VLYVGSKTR	KTKEGVLYVGSKTRE	285	5.27
33	47	VLYVGSKTR	TKEGVLYVGSKTREG	286	5.29
35	49	VLYVGSKTR	EGVLYVGSKTREGVV	377	7.26
2	16	FMKGLSMAK	DVFMKGLSMAKEGVV	702	13.51
36	50	VLYVGSKTR	GVLYVGSKTREGVVQ	762	14.58
1	15	FMKGLSMAK	MDVFMKGLSMAKEGV	780	14.91
30	44	AEKTKEGVL	AEKTKEGVLYVGSKT	851	16.13

3	17	LSMAKEGVV	VFMKGLSMAKEGVVA	1081	19.71
37	51	VLYVGSKTR	VLYVGSKTREGVVQG	1136	20.51
4	18	LSMAKEGVV	FMKGLSMAKEGVVAA	1396	23.93
5	19	LSMAKEGVV	MKGLSMAKEGVVAAA	1596	26.3
6	20	LSMAKEGVV	KGLSMAKEGVVAAAE	1953	30.17
45	59	VQGVASVAE	REGVVQGVASVAEKT	2593	35.87
1	15	GNIAAATGL	FSGAGNIAAATGLVK	2608	35.99
46	60	VQGVASVAE	EGVVQGVASVAEKTK	2635	36.21
44	58	VQGVASVAE	TREGVVQGVASVAEK	2649	36.32
2	16	IAAATGLVK	SGAGNIAAATGLVKR	2662	36.42
43	57	VVQGVASVA	KTREGVVQGVASVAE	2687	36.62
10	24	LVKREEFPT	ATGLVKREEFPTDLK	2836	37.73
7	21	IAAATGLVK	IAAATGLVKREEFPT	2844	37.79
3	17	IAAATGLVK	GAGNIAAATGLVKRE	2846	37.8
9	23	LVKREEFPT	AATGLVKREEFPTDL	2882	38.07
11	25	LVKREEFPT	TGLVKREEFPTDLKP	2920	38.35
56	70	TKEQASHLG	AEKTKEQASHLGGAV	2947	38.54
47	61	VQGVASVAE	GVVQGVASVAEKTKE	3034	39.15
8	22	LVKREEFPT	AAATGLVKREEFPTD	3111	39.67
4	18	IAAATGLVK	AGNIAAATGLVKREE	3198	40.23
5	19	IAAATGLVK	GNIAAATGLVKREEF	3733	43.52
53	67	KTKEQASHL	ASVAEKTKEQASHLG	3969	44.82
55	69	TKEQASHLG	VAEKTKEQASHLGGA	4011	45.06
7	21	LSMAKEGVV	GLSMAKEGVVAAAEK	4041	45.21
54	68	TKEQASHLG	SVAEKTKEQASHLGG	4080	45.42
42	56	GVVQGVASV	SKTREGVVQGVASVA	4342	46.75
8	22	LSMAKEGVV	LSMAKEGVVAAAEKT	4662	48.26

Legend to Table S4

The analysis of binding of the peptides to HLA DR2b was done using the Immune Epitope Data Base or IEDB (www. iedb.org). The default peptide length of 15 amino acids was used in the analysis but the sequence of the core nonamer peptide that is expected to bind to the HLA DR molecule and constitute the major portion of

the T cell epitope is shown (column 3) for each 15 mer sequence (column 4). The starting and ending amino acid residue number for each nonamer in the protein sequence are shown in columns 1 and 2 respectively. The analysis was done using the Stabilised Matrix Method (SMM) where the peptides are ranked according to their predicted binding affinities or IC_{50} (column 5) which indicates the concentration of peptide in nM expected to achieve 50% saturation of the HLA DR molecule. Therefore a lower IC_{50} shows a higher affinity. As a guide, peptides with IC_{50} values <50 nM are considered to bind with high affinity, between 50nM to 500 nM with intermediate affinity and between 500nM to 5000 nM with low affinity. For each peptide, a percentile rank (columns 6) is generated by comparing the peptide's score against the scores of five million random 15 mers selected from the SWISSPROT protein database. Therefore smaller percentile rank values, typically <10, also indicate higher affinity and specificity of binding to the HLA DR molecule.

Table S5. Results of the IEDB analysis of EBNA1 and EBNA4 showing peptides with potential to bind HLA DR2b

1. EBNA1

Start	End	Core nonamer	15mer peptide	IC50	Percentile rank
558	572	FMVFLQTHI	IVCYFMVFLQTHIFA	90	1.08
557	571	FMVFLQTHI	SIVCYFMVFLQTHIF	91	1.1
559	573	FMVFLQTHI	VCYFMVFLQTHIFAE	91	1.1
556	570	IVCYFMVFL	ESIVCYFMVFLQTHI	97	1.21
560	574	FMVFLQTHI	CYFMVFLQTHIFAEV	97	1.21
561	575	FMVFLQTHI	YFMVFLQTHIFAEVL	197	3.28
562	576	FMVFLQTHI	FMVFLQTHIFAEVLK	248	4.41
480	494	LRALLARSH	NIAEGLRALLARSHV	288	5.34
482	496	LRALLARSH	AEGLRALLARSHVER	306	5.73
579	593	LVMTKPAPT	IKDLVMTKPAPTCNI	317	5.98
481	495	LRALLARSH	IAEGLRALLARSHVE	318	6
483	497	LRALLARSH	EGLRALLARSHVERT	323	6.11
479	493	AEGLRALLA	ENIAEGLRALLARSH	326	6.18
576	590	IKDLVMTKP	KDAIKDLVMTKPAPT	327	6.2
577	591	LVMTKPAPT	DAIKDLVMTKPAPTC	331	6.29
580	594	LVMTKPAPT	KDLVMTKPAPTCNIR	332	6.31
578	592	LVMTKPAPT	AIKDLVMTKPAPTCN	336	6.41
554	568	IVCYFMVFL	LRESIVCYFMVFLQT	390	7.51
553	567	IVCYFMVFL	PLRESIVCYFMVFLQ	393	7.57
504	518	VFVYGGSKT	VAGVFVYGGSKTSLY	410	7.9
503	517	VFVYGGSKT	WVAGVFVYGGSKTSL	429	8.28
555	569	IVCYFMVFL	RESIVCYFMVFLQTH	448	8.67
502	516	VFVYGGSKT	TWVAGVFVYGGSKTS	451	8.73
505	519	VFVYGGSKT	AGVFVYGGSKTSLYN	488	9.48
501	515	GVFVYGGSK	GTWVAGVFVYGGSKT	549	10.73
552	566	ESIVCYFMV	GPLRESIVCYFMVFL	649	12.55

484	498	LRALLARSH	GLRALLARSHVERTT	661	12.76
514	528	YNLRRGTAL	KTSLYNLRRGTALAI	675	13.03
515	529	YNLRRGTAL	TSLYNLRRGTALAIP	695	13.39
512	526	TSLYNLRRG	GSKTSLYNLRRGTAL	752	14.41
563	577	MVFLQTHIF	MVFLQTHIFAEVLKD	805	15.33
516	530	YNLRRGTAL	SLYNLRRGTALAIPQ	808	15.38
506	520	FVYGGSKTS	GVFVYGGSKTSLYNL	815	15.5
513	527	YNLRRGTAL	SKTSLYNLRRGTALA	823	15.64
485	499	LRALLARSH	LRALLARSHVERTTD	863	16.32
581	595	LVMTKPAPT	DLVMTKPAPTCNIRV	940	17.59
582	596	LVMTKPAPT	LVMTKPAPTCNIRVT	969	18.06
530	544	LTPLSRLPF	QCRLTPLSRLPFGMA	1059	19.4
531	545	LTPLSRLPF	CRLTPLSRLPFGMAP	1088	19.81
564	578	THIFAEVLK	VFLQTHIFAEVLKDA	1196	21.32
507	521	VFVYGGSKT	VFVYGGSKTSLYNLR	1242	21.95
517	531	LRRGTALAI	LYNLRRGTALAIPQC	1272	22.35
565	579	THIFAEVLK	FLQTHIFAEVLKDAI	1283	22.49
478	492	AEGLRALLA	FENIAEGLRALLARS	1344	23.28
476	490	IAEGLRALL	PKFENIAEGLRALLA	1378	23.7
511	525	KTSLYNLRR	GGSKTSLYNLRRGTA	1404	24.02
477	491	AEGLRALLA	KFENIAEGLRALLAR	1427	24.29
518	532	LRRGTALAI	YNLRRGTALAIPQCR	1444	24.5
529	543	LTPLSRLPF	PQCRLTPLSRLPFGM	1487	25.02
566	580	THIFAEVLK	LQTHIFAEVLKDAIK	1542	25.66
527	541	QCRLTPLSR	AIPQCRLTPLSRLPF	1565	25.92
510	524	KTSLYNLRR	YGGSKTSLYNLRRGT	1574	26.03
528	542	LTPLSRLPF	IPQCRLTPLSRLPFG	1576	26.05
508	522	YGGSKTSLY	FVYGGSKTSLYNLRR	1592	26.24
509	523	KTSLYNLRR	VYGGSKTSLYNLRRG	1595	26.29
532	546	LSRLPFGMA	RLTPLSRLPFGMAPG	1617	26.53
533	547	LSRLPFGMA	LTPLSRLPFGMAPGP	1671	27.16

520	534	LRRGTALAI	LRRGTALAIPQCRLT	2020	30.82
500	514	VAGVFVYGG	EGTWVAGVFVYGGSK	2092	31.54
534	548	LSRLPFGMA	TPLSRLPFGMAPGPG	2223	32.76
551	565	ESIVCYFMV	PGPLRESIVCYFMVF	2330	33.72
568	582	FAEVLKDAI	THIFAEVLKDAIKDL	2494	35.09
522	536	LAIPQCRLT	RGTALAIPQCRLTPL	2558	35.6
524	538	LAIPQCRLT	TALAIPQCRLTPLSR	2593	35.87
523	537	LAIPQCRLT	GTALAIPQCRLTPLS	2699	36.71
567	581	FAEVLKDAI	QTHIFAEVLKDAIKD	2708	36.77
550	564	RESIVCYFM	QPGPLRESIVCYFMV	2781	37.33
475	489	ENIAEGLRA	NPKFENIAEGLRALL	2853	37.86
498	512	GTWVAGVFV	TDEGTWVAGVFVYGG	2881	38.06
486	500	RALLARSHV	RALLARSHVERTTDE	2939	38.48
499	513	GTWVAGVFV	DEGTWVAGVFVYGGS	2971	38.72
535	549	RLPFGMAPG	PLSRLPFGMAPGPGP	3084	39.5
521	535	LAIPQCRLT	RRGTALAIPQCRLTP	3122	39.74
536	550	RLPFGMAPG	LSRLPFGMAPGPGPQ	3146	39.9
569	583	FAEVLKDAI	HIFAEVLKDAIKDLV	3170	40.06
525	539	QCRLTPLSR	ALAIPQCRLTPLSRL	3239	40.5
519	533	LRRGTALAI	NLRRGTALAIPQCRL	3305	40.92
526	540	QCRLTPLSR	LAIPQCRLTPLSRLP	3531	42.34
400	414	RRPFFHPVG	PGRRPFFHPVGEADY	4096	45.5
474	488	ENIAEGLRA	SNPKFENIAEGLRAL	4105	45.54
497	511	GTWVAGVFV	TTDEGTWVAGVFVYG	4179	45.91
549	563	PGPLRESIV	PQPGPLRESIVCYFM	4196	46
570	584	VLKDAIKDL	IFAEVLKDAIKDLVM	4211	46.08
496	510	GTWVAGVFV	RTTDEGTWVAGVFVY	4241	46.23
600	614	GVDLPPWFP	FDDGVDLPPWFPPMV	4300	46.54
574	588	KDAIKDLVM	VLKDAIKDLVMTKPA	4302	46.55
601	615	LPPWFPPMV	DDGVDLPPWFPPMVE	4424	47.14
571	585	VLKDAIKDL	FAEVLKDAIKDLVMT	4429	47.17

455	469	GRRKKGGWF	DGGRRKKGGWFGKHR	4460	47.31
495	509	TTDEGTWVA	ERTTDEGTWVAGVFV	4737	48.61
459	473	KGGWFGKHR	RKKGGWFGKHRGQGG	4772	48.76
572	586	VLKDAIKDL	AEVLKDAIKDLVMTK	4888	49.29
473	487	FENIAEGLR	GSNPKFENIAEGLRA	4988	49.72

2. EBNA4

Start	End	Core Nonamer	15mer Peptide	IC_{50}	Percentile Rank
117	131	DLRPLGSLF	SMLQSDLRPLGSLFL	67	0.7
118	132	LRPLGSLFL	MLQSDLRPLGSLFLE	68	0.72
119	133	LRPLGSLFL	LQSDLRPLGSLFLEQ	68	0.72
120	134	LRPLGSLFL	QSDLRPLGSLFLEQN	69	0.73
121	135	LRPLGSLFL	SDLRPLGSLFLEQNL	72	0.78
123	137	LRPLGSLFL	LRPLGSLFLEQNLNI	162	2.48
156	170	IVKQRRWKL	KKPLPIVKQRRWKLL	177	2.79
157	171	IVKQRRWKL	KPLPIVKQRRWKLLS	178	2.81
680	694	LLRQWAPAT	GPATMLLRQWAPATM	189	3.07
682	696	LLRQWAPAT	ATMLLRQWAPATMQT	189	3.07
681	695	LLRQWAPAT	PATMLLRQWAPATMQ	190	3.11
159	173	IVKQRRWKL	LPIVKQRRWKLLSSC	193	3.18
158	172	IVKQRRWKL	PLPIVKQRRWKLLSS	196	3.26
155	169	LPIVKQRRW	RKKPLPIVKQRRWKL	200	3.35
683	697	LLRQWAPAT	TMLLRQWAPATMQTP	200	3.35
122	136	LRPLGSLFL	DLRPLGSLFLEQNLN	201	3.38
679	693	TMLLRQWAP	TGPATMLLRQWAPAT	203	3.42
164	178	WKLLSSCRS	QRRWKLLSSCRSWRM	209	3.53
624	638	LRPIPMRPL	WPMPLRPIPMRPLRM	210	3.56
626	640	LRPIPMRPL	MPLRPIPMRPLRMQP	211	3.58
625	639	LRPIPMRPL	PMPLRPIPMRPLRMQ	212	3.61
571	585	IQPLTSPTT	PLEIQPLTSPTTSQL	221	3.8

572	586	IQPLTSPTT	LEIQPLTSPTTSQLS	222	3.82
570	584	IQPLTSPTT	DPLEIQPLTSPTTSQ	235	4.12
165	179	LSSCRSWRM	RRWKLLSSCRSWRMG	236	4.14
568	582	EIQPLTSPT	GPDPLEIQPLTSPTT	236	4.14
569	583	IQPLTSPTT	PDPLEIQPLTSPTTS	236	4.14
623	637	LRPIPMRPL	QWPMPLRPIPMRPLR	237	4.16
622	636	PMPLRPIPM	RQWPMPLRPIPMRPL	241	4.25
768	782	LRQLLTGGV	TKQILRQLLTGGVKK	257	4.62
234	248	TAFLMARRA	KIETAFLMARRARSL	265	4.81
767	781	LRQLLTGGV	PTKQILRQLLTGGVK	266	4.83
769	783	LRQLLTGGV	KQILRQLLTGGVKKG	266	4.83
766	780	ILRQLLTGG	GPTKQILRQLLTGGV	270	4.94
235	249	LMARRARSL	IETAFLMARRARSLS	274	5.03
161	175	IVKQRRWKL	IVKQRRWKLLSSCRS	278	5.11
770	784	LRQLLTGGV	QILRQLLTGGVKKGR	292	5.41
236	250	LMARRARSL	ETAFLMARRARSLSA	318	6
166	180	LSSCRSWRM	RWKLLSSCRSWRMGY	320	6.04
237	251	LMARRARSL	TAFLMARRARSLSAE	320	6.04
167	181	LSSCRSWRM	WKLLSSCRSWRMGYR	333	6.34
628	642	LRPIPMRPL	LRPIPMRPLRMQPIP	362	6.96
253	267	FDLVSSGNT	YTLFFDLVSSGNTLY	364	7
627	641	LRPIPMRPL	PLRPIPMRPLRMQPI	364	7
254	268	FDLVSSGNT	TLFFDLVSSGNTLYA	367	7.05
238	252	LMARRARSL	AFLMARRARSLSAER	377	7.26
255	269	FDLVSSGNT	LFFDLVSSGNTLYAI	380	7.32
251	265	YTLFFDLVS	ERYTLFFDLVSSGNT	398	7.67
631	645	MRPLRMQPI	IPMRPLRMQPIPFNH	406	7.82
168	182	LSSCRSWRM	KLLSSCRSWRMGYRT	409	7.88
630	644	MRPLRMQPI	PIPMRPLRMQPIPFN	409	7.88
276	290	FIEFVGWLC	KNRVSFIEFVGWLCK	410	7.9
105	119	VIQLVHAVY	APVIQLVHAVYDSML	411	7.92

252	266	FDLVSSGNT	RYTLFFDLVSSGNTL	416	8.03
277	291	FIEFVGWLC	NRVSFIEFVGWLCKK	419	8.08
629	643	MRPLRMQPI	RPIPMRPLRMQPIPF	424	8.18
811	825	APIFYPPVL	DKIVQAPIFYPPVLQ	434	8.39
813	827	APIFYPPVL	IVQAPIFYPPVLQPI	444	8.59
278	292	FIEFVGWLC	RVSFIEFVGWLCKKD	445	8.61
275	289	RVSFIEFVG	TKNRVSFIEFVGWLC	454	8.78
160	174	IVKQRRWKL	PIVKQRRWKLLSSCR	460	8.91
810	824	KIVQAPIFY	SDKIVQAPIFYPPVL	460	8.91
279	293	FIEFVGWLC	VSFIEFVGWLCKKDH	466	9.02
148	162	AIRKKPLPI	RHRCQAIRKKPLPIV	473	9.17
812	826	APIFYPPVL	KIVQAPIFYPPVLQP	473	9.17
149	163	IRKKPLPIV	HRCQAIRKKPLPIVK	476	9.24
814	828	APIFYPPVL	VQAPIFYPPVLQPIQ	487	9.46
163	177	WKLLSSCRS	KQRRWKLLSSCRSWR	516	10.06
233	247	TAFLMARRA	QKIETAFLMARRARS	530	10.34
684	698	LLRQWAPAT	MLLRQWAPATMQTPP	532	10.38
106	120	VHAVYDSML	PVIQLVHAVYDSMLQ	547	10.7
162	176	WKLLSSCRS	VKQRRWKLLSSCRSW	548	10.72
261	275	GNTLYAIWI	SSGNTLYAIWIGLGT	548	10.72
104	118	VIQLVHAVY	QAPVIQLVHAVYDSM	551	10.77
103	117	VIQLVHAVY	TQAPVIQLVHAVYDS	563	10.99
102	116	VIQLVHAVY	PTQAPVIQLVHAVYD	564	11.01
515	529	VMATLLPPV	VMEQRVMATLLPPVP	565	11.02
257	271	LVSSGNTLY	FDLVSSGNTLYAIWI	574	11.19
150	164	AIRKKPLPI	RCQAIRKKPLPIVKQ	575	11.21
151	165	IRKKPLPIV	CQAIRKKPLPIVKQR	586	11.41
264	278	YAIWIGLGT	NTLYAIWIGLGTKNR	590	11.48
231	245	ETAFLMARR	QNQKIETAFLMARRA	610	11.85
265	279	WIGLGTKNR	TLYAIWIGLGTKNRV	610	11.85
685	699	LLRQWAPAT	LLRQWAPATMQTPPR	623	12.1

232	246	TAFLMARRA	NQKIETAFLMARRAR	634	12.29
514	528	EQRVMATLL	EVMEQRVMATLLPPV	642	12.43
516	530	VMATLLPPV	MEQRVMATLLPPVPQ	651	12.59
517	531	VMATLLPPV	EQRVMATLLPPVPQQ	651	12.59
573	587	IQPLTSPTT	EIQPLTSPTTSQLSS	657	12.69
260	274	GNTLYAIWI	VSSGNTLYAIWIGLG	663	12.8
633	647	PLRMQPIPF	MRPLRMQPIPFNHPV	663	12.8
518	532	VMATLLPPV	QRVMATLLPPVPQQP	668	12.89
656	670	ITPYKPTWA	QVEITPYKPTWAQIG	669	12.91
574	588	IQPLTSPTT	IQPLTSPTTSQLSSS	670	12.93
632	646	PLRMQPIPF	PMRPLRMQPIPFNHP	679	13.1
655	669	ITPYKPTWA	PQVEITPYKPTWAQI	685	13.21
657	671	ITPYKPTWA	VEITPYKPTWAQIGH	691	13.32
771	785	LRQLLTGGV	ILRQLLTGGVKKGRP	691	13.32
267	281	WIGLGTKNR	YAIWIGLGTKNRVSF	693	13.35
266	280	WIGLGTKNR	LYAIWIGLGTKNRVS	698	13.43
653	667	VEITPYKPT	QTPQVEITPYKPTWA	717	13.8
654	668	ITPYKPTWA	TPQVEITPYKPTWAQ	722	13.88
262	276	YAIWIGLGT	SGNTLYAIWIGLGTK	728	13.98
107	121	VHAVYDSML	VIQLVHAVYDSMLQS	731	14.04
259	273	GNTLYAIWI	LVSSGNTLYAIWIGL	736	14.14
263	277	YAIWIGLGT	GNTLYAIWIGLGTKN	736	14.14
256	270	FDLVSSGNT	FFDLVSSGNTLYAIW	742	14.24
268	282	WIGLGTKNR	AIWIGLGTKNRVSFI	760	14.55
239	253	LMARRARSL	FLMARRARSLSAERY	761	14.56
809	823	DKIVQAPIF	TSDKIVQAPIFYPPV	780	14.91
414	428	IVTDFSVIK	RAIVTDFSVIKAIEE	788	15.04
101	115	PVIQLVHAV	NPTQAPVIQLVHAVY	789	15.06
772	786	LRQLLTGGV	LRQLLTGGVKKGRPS	791	15.08
258	272	GNTLYAIWI	DLVSSGNTLYAIWIG	794	15.12
807	821	DKIVQAPIF	SGTSDKIVQAPIFYP	794	15.12

475	489	WQPLPGPQV	LEPWQPLPGPQVTAV	811	15.43
808	822	DKIVQAPIF	GTSDKIVQAPIFYPP	811	15.43
806	820	DKIVQAPIF	GSGTSDKIVQAPIFY	826	15.69
765	779	KQILRQLLT	QGPTKQILRQLLTGG	833	15.82
472	486	LEPWQPLPG	QARLEPWQPLPGPQV	844	16.02
473	487	WQPLPGPQV	ARLEPWQPLPGPQVT	846	16.06
153	167	IRKKPLPIV	AIRKKPLPIVKQRRW	848	16.08
147	161	CQAIRKKPL	VRHRCQAIRKKPLPI	855	16.19
240	254	LMARRARSL	LMARRARSLSAERYT	856	16.21
474	488	WQPLPGPQV	RLEPWQPLPGPQVTA	857	16.23
476	490	WQPLPGPQV	EPWQPLPGPQVTAVL	860	16.27
888	902	VVILENVGQ	SHSPVVILENVGQGQ	860	16.27
108	122	VHAVYDSML	IQLVHAVYDSMLQSD	861	16.29
666	680	HIPYQPTPT	WAQIGHIPYQPTPTG	863	16.32
889	903	VVILENVGQ	HSPVVILENVGQGQQ	865	16.35
781	795	SLKLQAALE	KKGRPSLKLQAALER	874	16.5
887	901	VVILENVGQ	ASHSPVVILENVGQG	875	16.52
886	900	PVVILENVG	GASHSPVVILENVGQ	876	16.53
783	797	LKLQAALER	GRPSLKLQAALERQA	892	16.78
824	838	VMGQGGSPT	LQPIQVMGQGGSPTA	902	16.96
447	461	VVLQRPPTQ	SQAPTVVLQRPPTQQ	911	17.1
446	460	TVVLQRPPT	ESQAPTVVLQRPPTQ	912	17.11
782	796	LKLQAALER	KGRPSLKLQAALERQ	916	17.19
448	462	VVLQRPPTQ	QAPTVVLQRPPTQQE	917	17.21
784	798	LKLQAALER	RPSLKLQAALERQAA	920	17.26
815	829	APIFYPPVL	QAPIFYPPVLQPIQV	921	17.27
823	837	IQVMGQGGS	VLQPIQVMGQGGSPT	922	17.29
124	138	LFLEQNLNI	RPLGSLFLEQNLNIE	925	17.35
665	679	IGHIPYQPT	TWAQIGHIPYQPTPT	926	17.36
449	463	VVLQRPPTQ	APTVVLQRPPTQQEP	927	17.38
411	425	IVTDFSVIK	KKCRAIVTDFSVIKA	930	17.43

125	139	LFLEQNLNI	PLGSLFLEQNLNIEE	932	17.47
152	166	IRKKPLPIV	QAIRKKPLPIVKQRR	932	17.47
825	839	VMGQGGSPT	QPIQVMGQGGSPTAM	933	17.49
412	426	IVTDFSVIK	KCRAIVTDFSVIKAI	935	17.52
667	681	HIPYQPTPT	AQIGHIPYQPTPTGP	942	17.62
827	841	VMGQGGSPT	IQVMGQGGSPTAMAA	962	17.95
154	168	LPIVKQRRW	IRKKPLPIVKQRRWK	975	18.16
109	123	VHAVYDSML	QLVHAVYDSMLQSDL	977	18.19
826	840	VMGQGGSPT	PIQVMGQGGSPTAMA	977	18.19
785	799	LKLQAALER	PSLKLQAALERQAAA	992	18.43
97	111	VNPTQAPVI	FVDVNPTQAPVIQLV	1017	18.78
170	184	LSSCRSWRM	LSSCRSWRMGYRTHN	1038	19.1
169	183	LSSCRSWRM	LLSSCRSWRMGYRTH	1040	19.13
890	904	VVILENVGQ	SPVVILENVGQGQQQ	1044	19.19
787	801	LKLQAALER	LKLQAALERQAAAGW	1047	19.23
336	350	AMNIEAPRL	YARGQAMNIEAPRLP	1048	19.25
174	188	RMGYRTHNL	RSWRMGYRTHNLKVN	1056	19.35
413	427	IVTDFSVIK	CRAIVTDFSVIKAIE	1059	19.4
98	112	VNPTQAPVI	VDVNPTQAPVIQLVH	1063	19.46
96	110	VNPTQAPVI	RFVDVNPTQAPVIQL	1065	19.48
126	140	LFLEQNLNI	LGSLFLEQNLNIEEF	1065	19.48
173	187	RMGYRTHNL	CRSWRMGYRTHNLKV	1077	19.66
410	424	RAIVTDFSV	TKKCRAIVTDFSVIK	1080	19.7
95	109	VNPTQAPVI	PRFVDVNPTQAPVIQ	1086	19.79
94	108	FVDVNPTQA	QPRFVDVNPTQAPVI	1087	19.8
763	777	TKQILRQLL	GQQGPTKQILRQLLT	1089	19.83
337	351	AMNIEAPRL	ARGQAMNIEAPRLPD	1094	19.9
805	819	SDKIVQAPI	PGSGTSDKIVQAPIF	1095	19.91
27	41	VTQVGSEPI	GDQGNVTQVGSEPIS	1096	19.93
26	40	NVTQVGSEP	YGDQGNVTQVGSEPI	1098	19.96
749	763	MQLALRAPA	VLPTPMQLALRAPAG	1100	19.98

750	764	QLALRAPAG	LPTPMQLALRAPAGQ	1105	20.05
764	778	KQILRQLLT	QQGPTKQILRQLLTG	1106	20.06
211	225	ATTYSAGIV	GTRHATTYSAGIVQI	1110	20.12
230	244	NQKIETAFL	DQNQKIETAFLMARR	1113	20.16
335	349	YARGQAMNI	AYARGQAMNIEAPRL	1114	20.17
338	352	AMNIEAPRL	RGQAMNIEAPRLPDD	1117	20.23
28	42	VTQVGSEPI	DQGNVTQVGSEPISP	1119	20.25
29	43	VTQVGSEPI	QGNVTQVGSEPISPE	1125	20.35
280	294	FIEFVGWLC	SFIEFVGWLCKKDHT	1125	20.35
751	765	QLALRAPAG	PTPMQLALRAPAGQQ	1132	20.45
30	44	VTQVGSEPI	GNVTQVGSEPISPEI	1140	20.56
212	226	ATTYSAGIV	TRHATTYSAGIVQIP	1146	20.64
752	766	QLALRAPAG	TPMQLALRAPAGQQG	1146	20.64
417	431	FSVIKAIEE	VTDFSVIKAIEEEHR	1153	20.74
269	283	IGLGTKNRV	IWIGLGTKNRVSFIE	1155	20.76
136	150	FIWMCMTVR	NIEEFIWMCMTVRHR	1160	20.83
209	223	HATTYSAGI	DEGTRHATTYSAGIV	1164	20.89
175	189	RMGYRTHNL	SWRMGYRTHNLKVNS	1169	20.96
418	432	IKAIEEEHR	TDFSVIKAIEEEHRK	1173	21.01
210	224	ATTYSAGIV	EGTRHATTYSAGIVQ	1176	21.05
816	830	APIFYPPVL	APIFYPPVLQPIQVM	1179	21.09
339	353	AMNIEAPRL	GQAMNIEAPRLPDDP	1185	21.17
145	159	RHRCQAIRK	MTVRHRCQAIRKKPL	1186	21.19
663	677	WAQIGHIPY	KPTWAQIGHIPYQPT	1195	21.31
450	464	VVLQRPPTQ	PTVVLQRPPTQQEPG	1197	21.34
375	389	MESLKNIPQ	KSGMESLKNIPQTLP	1204	21.43
374	388	MESLKNIPQ	DKSGMESLKNIPQTL	1206	21.45
664	678	IGHIPYQPT	PTWAQIGHIPYQPTP	1216	21.59
137	151	FIWMCMTVR	IEEFIWMCMTVRHRC	1226	21.73
415	429	FSVIKAIEE	AIVTDFSVIKAIEEE	1226	21.73
172	186	RMGYRTHNL	SCRSWRMGYRTHNLK	1235	21.85

373	387	MESLKNIPQ	EDKSGMESLKNIPQT	1239	21.91
376	390	MESLKNIPQ	SGMESLKNIPQTLPY	1243	21.97
372	386	KSGMESLKN	EEDKSGMESLKNIPQ	1245	21.99
171	185	WRMGYRTHN	SSCRSWRMGYRTHNL	1247	22.01
127	141	LFLEQNLNI	GSLFLEQNLNIEEFI	1270	22.32
135	149	FIWMCMTVR	LNIEEFIWMCMTVRH	1272	22.35
86	100	VHTRQPRFV	DPLDVHTRQPRFVDV	1276	22.4
511	525	VMEQRVMAT	KDDEVMEQRVMATLL	1294	22.62
512	526	EQRVMATLL	DDEVMEQRVMATLLP	1297	22.66
918	932	DIAVSSPSS	MLGLGDIAVSSPSSS	1309	22.82
87	101	VHTRQPRFV	PLDVHTRQPRFVDVN	1312	22.86
788	802	LERQAAAGW	KLQAALERQAAAGWQ	1314	22.89
513	527	EQRVMATLL	DEVMEQRVMATLLPP	1315	22.9
281	295	FIEFVGWLC	FIEFVGWLCKKDHTH	1316	22.91
519	533	VMATLLPPV	RVMATLLPPVPQQPR	1321	22.97
84	98	DVHTRQPRF	GDDPLDVHTRQPRFV	1332	23.13
668	682	HIPYQPTPT	QIGHIPYQPTPTGPA	1332	23.13
85	99	VHTRQPRFV	DDPLDVHTRQPRFVD	1339	23.22
620	634	PMPLRPIPM	APRQWPMPLRPIPMR	1340	23.23
669	683	HIPYQPTPT	IGHIPYQPTPTGPAT	1355	23.41
789	803	LERQAAAGW	LQAALERQAAAGWQP	1368	23.58
146	160	CQAIRKKPL	TVRHRCQAIRKKPLP	1369	23.59
621	635	PMPLRPIPM	PRQWPMPLRPIPMRP	1378	23.7
138	152	FIWMCMTVR	EEFIWMCMTVRHRCQ	1386	23.8
790	804	LERQAAAGW	QAALERQAAAGWQPS	1392	23.88
634	648	LRMQPIPFN	RPLRMQPIPFNHPVG	1395	23.91
753	767	QLALRAPAG	PMQLALRAPAGQQGP	1403	24.01
88	102	VHTRQPRFV	LDVHTRQPRFVDVNP	1411	24.11
791	805	LERQAAAGW	AALERQAAAGWQPSP	1416	24.17
917	931	MLGLGDIAV	DMLGLGDIAVSSPSS	1422	24.24
913	927	MLGLGDIAV	KQERDMLGLGDIAVS	1439	24.44

912	926	DMLGLGDIA	AKQERDMLGLGDIAV	1442	24.48
229	243	NQKIETAFL	SDQNQKIETAFLMAR	1447	24.54
213	227	ATTYSAGIV	RHATTYSAGIVQIPR	1460	24.69
250	264	YTLFFDLVS	AERYTLFFDLVSSGN	1474	24.87
914	928	MLGLGDIAV	QERDMLGLGDIAVSS	1474	24.87
110	124	HAVYDSMLQ	LVHAVYDSMLQSDLR	1486	25.01
619	633	QWPMPLRPI	AAPRQWPMPLRPIPM	1487	25.02
420	434	IKAIEEEHR	FSVIKAIEEEHRKKK	1492	25.08
330	344	LAYARGQAM	EEIDLAYARGQAMNI	1494	25.1
495	509	GVQVHGSML	SMQGVQVHGSMLDLL	1499	25.16
134	148	IEEFIWMCM	NLNIEEFIWMCMTVR	1500	25.17
915	929	MLGLGDIAV	ERDMLGLGDIAVSSP	1500	25.17
419	433	IKAIEEEHR	DFSVIKAIEEEHRKK	1519	25.4
658	672	ITPYKPTWA	EITPYKPTWAQIGHI	1519	25.4
496	510	GVQVHGSML	MQGVQVHGSMLDLLE	1520	25.41
916	930	MLGLGDIAV	RDMLGLGDIAVSSPS	1520	25.41
178	192	YRTHNLKVN	MGYRTHNLKVNSFES	1537	25.61
328	342	IDLAYARGQ	TNEEIDLAYARGQAM	1562	25.89
659	673	ITPYKPTWA	ITPYKPTWAQIGHIP	1575	26.04
416	430	FSVIKAIEE	IVTDFSVIKAIEEEH	1578	26.07
242	256	ARSLSAERY	ARRARSLSAERYTLF	1583	26.14
228	242	NQKIETAFL	ISDQNQKIETAFLMA	1628	26.66
100	114	VNPTQAPVI	VNPTQAPVIQLVHAV	1635	26.74
270	284	GLGTKNRVS	WIGLGTKNRVSFIEF	1657	26.99
227	241	NQKIETAFL	RISDQNQKIETAFLM	1659	27.02
247	261	LSAERYTLF	SLSAERYTLFFDLVS	1672	27.18
421	435	IKAIEEEHR	SVIKAIEEEHRKKKA	1673	27.19
920	934	DIAVSSPSS	GLGDIAVSSPSSSET	1675	27.21
248	262	YTLFFDLVS	LSAERYTLFFDLVSS	1678	27.25
331	345	LAYARGQAM	EIDLAYARGQAMNIE	1683	27.3
919	933	DIAVSSPSS	LGLGDIAVSSPSSSE	1695	27.43

775	789	VKKGRPSLK	LLTGGVKKGRPSLKL	1698	27.46
181	195	NLKVNSFES	RTHNLKVNSFESGGD	1702	27.5
329	343	LAYARGQAM	NEEIDLAYARGQAMN	1709	27.58
776	790	KKGRPSLKL	LTGGVKKGRPSLKLQ	1717	27.67
307	321	KPWLRAHPV	PKAAKPWLRAHPVAI	1737	27.89
308	322	KPWLRAHPV	KAAKPWLRAHPVAIP	1739	27.92
494	508	GVQVHGSML	ESMQGVQVHGSMLDL	1747	28
249	263	YTLFFDLVS	SAERYTLFFDLVSSG	1756	28.1
332	346	LAYARGQAM	IDLAYARGQAMNIEA	1782	28.39
921	935	DIAVSSPSS	LGDIAVSSPSSSETS	1787	28.45
660	674	YKPTWAQIG	TPYKPTWAQIGHIPY	1795	28.53
182	196	VNSFESGGD	THNLKVNSFESGGDN	1798	28.57
309	323	KPWLRAHPV	AAKPWLRAHPVAIPY	1803	28.62
176	190	RMGYRTHNL	WRMGYRTHNLKVNSF	1804	28.64
179	193	NLKVNSFES	GYRTHNLKVNSFESG	1818	28.78
216	230	AGIVQIPRI	TTYSAGIVQIPRISD	1832	28.94
217	231	AGIVQIPRI	TYSAGIVQIPRISDQ	1832	28.94
140	154	WMCMTVRHR	FIWMCMTVRHRCQAI	1859	29.21
218	232	AGIVQIPRI	YSAGIVQIPRISDQN	1859	29.21
243	257	ARSLSAERY	RRARSLSAERYTLFF	1880	29.44
493	507	GVQVHGSML	EESMQGVQVHGSMLD	1882	29.46
408	422	TKKCRAIVT	KSTKKCRAIVTDFSV	1901	29.64
777	791	KKGRPSLKL	TGGVKKGRPSLKLQA	1902	29.65
520	534	VMATLLPPV	VMATLLPPVPQQPRA	1918	29.81
177	191	RMGYRTHNL	RMGYRTHNLKVNSFE	1931	29.94
833	847	TAMAASAVT	GGSPTAMAASAVTQA	1936	29.98
778	792	KKGRPSLKL	GGVKKGRPSLKLQAA	1943	30.06
678	692	TMLLRQWAP	PTGPATMLLRQWAPA	1945	30.08
635	649	PLRMQPIPF	PLRMQPIPFNHPVGP	1954	30.18
139	153	WMCMTVRHR	EFIWMCMTVRHRCQA	1985	30.48
487	501	LLHEESMQG	TAVLLHEESMQGVQV	1990	30.53

832	846	TAMAASAVT	QGGSPTAMAASAVTQ	2009	30.71
834	848	TAMAASAVT	GSPTAMAASAVTQAP	2009	30.71
831	845	PTAMAASAV	GQGGSPTAMAASAVT	2011	30.73
762	776	PTKQILRQL	AGQQGPTKQILRQLL	2023	30.85
407	421	TKKCRAIVT	AKSTKKCRAIVTDFS	2057	31.18
111	125	VHAVYDSML	VHAVYDSMLQSDLRP	2064	31.24
215	229	AGIVQIPRI	ATTYSAGIVQIPRIS	2069	31.3
485	499	LLHEESMQG	QVTAVLLHEESMQGV	2089	31.49
142	156	TVRHRCQAI	WMCMTVRHRCQAIRK	2103	31.65
486	500	LLHEESMQG	VTAVLLHEESMQGVQ	2103	31.65
383	397	TLPYNPTVY	NIPQTLPYNPTVYGR	2119	31.8
779	793	KKGRPSLKL	GVKKGRPSLKLQAAL	2127	31.89
180	194	NLKVNSFES	YRTHNLKVNSFESGG	2134	31.96
382	396	TLPYNPTVY	KNIPQTLPYNPTVYG	2139	32.01
381	395	PQTLPYNPT	LKNIPQTLPYNPTVY	2165	32.24
241	255	ARSLSAERY	MARRARSLSAERYTL	2177	32.35
214	228	ATTYSAGIV	HATTYSAGIVQIPRI	2183	32.41
406	420	TKKCRAIVT	DAKSTKKCRAIVTDF	2204	32.59
497	511	VQVHGSMLD	QGVQVHGSMLDLLEK	2226	32.79
404	418	STKKCRAIV	KSDAKSTKKCRAIVT	2265	33.15
113	127	DSMLQSDLR	AVYDSMLQSDLRPLG	2269	33.18
478	492	WQPLPGPQV	WQPLPGPQVTAVLLH	2277	33.25
193	207	VHPVLVTAT	GGDNVHPVLVTATLG	2283	33.3
194	208	VHPVLVTAT	GDNVHPVLVTATLGC	2291	33.37
891	905	VVILENVGQ	PVVILENVGQGQQQT	2291	33.37
294	308	IREWFRQCT	THIREWFRQCTGRPK	2301	33.46
579	593	SQLSSSAPS	SPTTSQLSSSAPSCA	2303	33.48
405	419	TKKCRAIVT	SDAKSTKKCRAIVTD	2309	33.53
116	130	LQSDLRPLG	DSMLQSDLRPLGSLF	2329	33.71
661	675	WAQIGHIPY	PYKPTWAQIGHIPYQ	2330	33.72
488	502	LLHEESMQG	AVLLHEESMQGVQVH	2341	33.8

786	800	LKLQAALER	SLKLQAALERQAAAG	2346	33.85
195	209	VHPVLVTAT	DNVHPVLVTATLGCD	2366	34.03
305	319	PKAAKPWLR	GRPKAAKPWLRAHPV	2372	34.08
477	491	WQPLPGPQV	PWQPLPGPQVTAVLL	2372	34.08
445	459	PESQAPTVV	PESQAPTVVLQRPPT	2377	34.12
821	835	LQPIQVMGQ	PPVLQPIQVMGQGGS	2378	34.13
662	676	WAQIGHIPY	YKPTWAQIGHIPYQP	2384	34.18
310	324	WLRAHPVAI	AKPWLRAHPVAIPYD	2394	34.26
387	401	PTVYGRPAV	TLPYNPTVYGRPAVF	2395	34.27
112	126	DSMLQSDLR	HAVYDSMLQSDLRPL	2405	34.35
389	403	PTVYGRPAV	PYNPTVYGRPAVFDR	2452	34.75
822	836	IQVMGQGGS	PVLQPIQVMGQGGSP	2455	34.77
385	399	TLPYNPTVY	PQTLPYNPTVYGRPA	2459	34.8
388	402	PTVYGRPAV	LPYNPTVYGRPAVFD	2459	34.8
114	128	MLQSDLRPL	VYDSMLQSDLRPLGS	2461	34.82
219	233	IVQIPRISD	SAGIVQIPRISDQNQ	2488	35.04
384	398	TLPYNPTVY	IPQTLPYNPTVYGRP	2508	35.2
143	157	RHRCQAIRK	MCMTVRHRCQAIRKK	2520	35.29
296	310	FRQCTGRPK	IREWFRQCTGRPKAA	2526	35.34
885	899	ASHSPVVIL	GGASHSPVVILENVG	2533	35.39
295	309	FRQCTGRPK	HIREWFRQCTGRPKA	2542	35.47
89	103	VHTRQPRFV	DVHTRQPRFVDVNPT	2564	35.64
748	762	TPMQLALRA	QVLPTPMQLALRAPA	2564	35.64
144	158	RHRCQAIRK	CMTVRHRCQAIRKKP	2571	35.69
577	591	TTSQLSSSA	LTSPTTSQLSSSAPS	2601	35.93
183	197	VNSFESGGD	HNLKVNSFESGGDNV	2602	35.94
580	594	SQLSSSAPS	PTTSQLSSSAPSCAQ	2605	35.97
754	768	QLALRAPAG	MQLALRAPAGQQGPT	2627	36.14
128	142	LFLEQNLNI	SLFLEQNLNIEEFIW	2641	36.26
141	155	TVRHRCQAI	IWMCMTVRHRCQAIR	2645	36.29
746	760	VLPTPMQLA	PPQVLPTPMQLALRA	2656	36.37

581	595	SQLSSSAPS	TTSQLSSSAPSCAQT	2657	36.37
377	391	MESLKNIPQ	GMESLKNIPQTLPYN	2663	36.42
578	592	SQLSSSAPS	TSPTTSQLSSSAPSC	2663	36.42
747	761	VLPTPMQLA	PQVLPTPMQLALRAP	2668	36.46
311	325	WLRAHPVAI	KPWLRAHPVAIPYDD	2671	36.49
378	392	MESLKNIPQ	MESLKNIPQTLPYNP	2685	36.6
274	288	RVSFIEFVG	GTKNRVSFIEFVGWL	2688	36.63
244	258	LSAERYTLF	RARSLSAERYTLFFD	2689	36.63
484	498	VLLHEESMQ	PQVTAVLLHEESMQG	2697	36.69
426	440	HRKKKAART	IEEEHRKKKAARTEQ	2703	36.73
424	438	EHRKKKAAR	KAIEEEHRKKKAART	2705	36.75
425	439	HRKKKAART	AIEEEHRKKKAARTE	2706	36.76
327	341	IDLAYARGQ	LTNEEIDLAYARGQA	2708	36.77
386	400	NPTVYGRPA	QTLPYNPTVYGRPAV	2731	36.95
451	465	VVLQRPPTQ	TVVLQRPPTQQEPGP	2744	37.06
828	842	VMGQGGSPT	QVMGQGGSPTAMAAS	2753	37.13
427	441	HRKKKAART	EEEHRKKKAARTEQP	2756	37.15
297	311	FRQCTGRPK	REWFRQCTGRPKAAK	2757	37.16
390	404	PTVYGRPAV	YNPTVYGRPAVFDRK	2759	37.17
115	129	MLQSDLRPL	YDSMLQSDLRPLGSL	2764	37.21
340	354	AMNIEAPRL	QAMNIEAPRLPDDPI	2772	37.27
99	113	VNPTQAPVI	DVNPTQAPVIQLVHA	2776	37.29
129	143	LFLEQNLNI	LFLEQNLNIEEFIWM	2800	37.46
780	794	KKGRPSLKL	VKKGRPSLKLQAALE	2800	37.46
184	198	VNSFESGGD	NLKVNSFESGGDNVH	2802	37.48
90	104	VHTRQPRFV	VHTRQPRFVDVNPTQ	2823	37.64
273	287	RVSFIEFVG	LGTKNRVSFIEFVGW	2829	37.68
510	524	VMEQRVMAT	EKDDEVMEQRVMATL	2851	37.84
835	849	TAMAASAVT	SPTAMAASAVTQAPT	2860	37.91
428	442	HRKKKAART	EEHRKKKAARTEQPR	2872	37.99
290	304	HTHIREWFR	KKDHTHIREWFRQCT	2947	38.54

829	843	VMGQGGSPT	VMGQGGSPTAMAASA	2949	38.56
326	340	EEIDLAYAR	PLTNEEIDLAYARGQ	2961	38.64
892	906	VVILENVGQ	VVILENVGQGQQQTL	2978	38.77
132	146	IEEFIWMCM	EQNLNIEEFIWMCMT	2984	38.81
670	684	HIPYQPTPT	GHIPYQPTPTGPATM	3009	38.99
131	145	NLNIEEFIW	LEQNLNIEEFIWMCM	3016	39.03
192	206	VHPVLVTAT	SGGDNVHPVLVTATL	3017	39.04
492	506	MQGVQVHGS	HEESMQGVQVHGSML	3019	39.05
291	305	HTHIREWFR	KDHTHIREWFRQCTG	3050	39.27
409	423	RAIVTDFSV	STKKCRAIVTDFSVI	3052	39.28
272	286	GLGTKNRVS	GLGTKNRVSFIEFVG	3063	39.36
818	832	PVLQPIQVM	IFYPPVLQPIQVMGQ	3071	39.42
333	347	YARGQAMNI	DLAYARGQAMNIEAP	3073	39.43
820	834	LQPIQVMGQ	YPPVLQPIQVMGQGG	3076	39.45
636	650	PIPFNHPVG	LRMQPIPFNHPVGPT	3083	39.49
745	759	VLPTPMQLA	PPPQVLPTPMQLALR	3129	39.79
744	758	VLPTPMQLA	TPPPQVLPTPMQLAL	3139	39.85
245	259	LSAERYTLF	ARSLSAERYTLFFDL	3140	39.86
31	45	VTQVGSEPI	NVTQVGSEPISPEIG	3164	40.02
830	844	QGGSPTAMA	MGQGGSPTAMAASAV	3168	40.04
817	831	PIFYPPVLQ	PIFYPPVLQPIQVMG	3181	40.13
208	222	GTRHATTYS	CDEGTRHATTYSAGI	3192	40.19
306	320	KPWLRAHPV	RPKAAKPWLRAHPVA	3218	40.36
226	240	RISDQNQKI	PRISDQNQKIETAFL	3223	40.4
32	46	VTQVGSEPI	VTQVGSEPISPEIGP	3233	40.46
133	147	IEEFIWMCM	QNLNIEEFIWMCMTV	3233	40.46
298	312	FRQCTGRPK	EWFRQCTGRPKAAKP	3256	40.62
303	317	PKAAKPWLR	CTGRPKAAKPWLRAH	3257	40.62
334	348	YARGQAMNI	LAYARGQAMNIEAPR	3265	40.67
819	833	LQPIQVMGQ	FYPPVLQPIQVMGQG	3279	40.76
38	52	ISPEIGPFE	EPISPEIGPFELSAA	3286	40.8

302	316	PKAAKPWLR	QCTGRPKAAKPWLRA	3296	40.86
498	512	VHGSMLDLL	GVQVHGSMLDLLEKD	3297	40.87
617	631	APRQWPMPL	ETAAPRQWPMPLRPI	3297	40.87
652	666	VEITPYKPT	HQTPQVEITPYKPTW	3309	40.95
220	234	IVQIPRISD	AGIVQIPRISDQNQK	3315	40.98
618	632	QWPMPLRPI	TAAPRQWPMPLRPIP	3345	41.17
509	523	DDEVMEQRV	LEKDDEVMEQRVMAT	3348	41.19
793	807	LERQAAAGW	LERQAAAGWQPSPGS	3358	41.26
288	302	HTHIREWFR	LCKKDHTHIREWFRQ	3367	41.32
881	895	GGASHSPVV	EMPHGGASHSPVVIL	3370	41.34
651	665	QVEITPYKP	PHQTPQVEITPYKPT	3428	41.71
489	503	LHEESMQGV	VLLHEESMQGVQVHG	3446	41.82
191	205	DNVHPVLVT	ESGGDNVHPVLVTAT	3447	41.83
39	53	IGPFELSAA	PISPEIGPFELSAAS	3450	41.85
452	466	VVLQRPPTQ	VVLQRPPTQQEPGPV	3453	41.86
743	757	PQVLPTPMQ	PTPPPQVLPTPMQLA	3457	41.89
380	394	PQTLPYNPT	SLKNIPQTLPYNPTV	3468	41.95
677	691	ATMLLRQWA	TPTGPATMLLRQWAP	3479	42.01
185	199	VNSFESGGD	LKVNSFESGGDNVHP	3512	42.23
882	896	GGASHSPVV	MPHGGASHSPVVILE	3539	42.39
922	936	IAVSSPSSS	GDIAVSSPSSSETSN	3544	42.42
341	355	AMNIEAPRL	AMNIEAPRLPDDPII	3557	42.5
41	55	IGPFELSAA	SPEIGPFELSAASED	3562	42.53
196	210	PVLVTATLG	NVHPVLVTATLGCDE	3618	42.85
40	54	IGPFELSAA	ISPEIGPFELSAASE	3624	42.88
587	601	SCAQTPWPV	SSAPSCAQTPWPVVQ	3640	42.98
282	296	GWLCKKDHT	IEFVGWLCKKDHTHI	3644	43
304	318	PKAAKPWLR	TGRPKAAKPWLRAHP	3646	43.01
774	788	GGVKKGRPS	QLLTGGVKKGRPSLK	3648	43.02
271	285	IGLGTKNRV	IGLGTKNRVSFIEFV	3655	43.07
586	600	SCAQTPWPV	SSSAPSCAQTPWPVV	3677	43.2

883	897	GGASHSPVV	PHGGASHSPVVILEN	3713	43.41
588	602	SCAQTPWPV	SAPSCAQTPWPVVQP	3719	43.44
42	56	IGPFELSAA	PEIGPFELSAASEDD	3741	43.56
287	301	LCKKDHTHI	WLCKKDHTHIREWFR	3754	43.65
301	315	GRPKAAKPW	RQCTGRPKAAKPWLR	3764	43.71
589	603	SCAQTPWPV	APSCAQTPWPVVQPS	3767	43.73
755	769	QLALRAPAG	QLALRAPAGQQGPTK	3817	44.01
868	882	RAKIEAYTE	PSKRAKIEAYTEPEM	3822	44.03
880	894	GGASHSPVV	PEMPHGGASHSPVVI	3835	44.11
879	893	EMPHGGASH	EPEMPHGGASHSPVV	3860	44.24
197	211	PVLVTATLG	VHPVLVTATLGCDEG	3862	44.25
246	260	LSAERYTLF	RSLSAERYTLFFDLV	3925	44.59
289	303	HTHIREWFR	CKKDHTHIREWFRQC	3950	44.72
379	393	LKNIPQTLP	ESLKNIPQTLPYNPT	3977	44.87
728	742	VPRQRPRGA	QVPPVPRQRPRGAPT	3995	44.97
708	722	VPRQRPRGA	EVPPVPRQRPRGAPT	4008	45.04
792	806	LERQAAAGW	ALERQAAAGWQPSPG	4046	45.24
391	405	TVYGRPAVF	NPTVYGRPAVFDRKS	4057	45.29
709	723	VPRQRPRGA	VPPVPRQRPRGAPTP	4082	45.43
729	743	VPRQRPRGA	VPPVPRQRPRGAPTP	4082	45.43
671	685	HIPYQPTPT	HIPYQPTPTGPATML	4143	45.73
686	700	LRQWAPATM	LRQWAPATMQTPPRA	4148	45.76
869	883	IEAYTEPEM	SKRAKIEAYTEPEMP	4159	45.81
852	866	RRGVGPMPP	TRERRGVGPMPPTDI	4208	46.06
423	437	IKAIEEEHR	IKAIEEEHRKKKAAR	4301	46.55
853	867	VGPMPPTDI	RERRGVGPMPPTDIP	4332	46.7
710	724	VPRQRPRGA	PPVPRQRPRGAPTPT	4350	46.78
730	744	VPRQRPRGA	PPVPRQRPRGAPTPT	4350	46.78
773	787	GGVKKGRPS	RQLLTGGVKKGRPSL	4369	46.88
585	599	LSSSAPSCA	LSSSAPSCAQTPWPV	4379	46.93
463	477	VGPLSVQAR	PGPVGPLSVQARLEP	4388	46.97

469	483	ARLEPWQPL	LSVQARLEPWQPLPG	4493	47.47
422	436	IKAIEEEHR	VIKAIEEEHRKKKAA	4535	47.66
462	476	VGPLSVQAR	EPGPVGPLSVQARLE	4542	47.7
726	740	PVPRQRPRG	PPQVPPVPRQRPRGA	4577	47.86
706	720	PVPRQRPRG	PPEVPPVPRQRPRGA	4605	47.99
221	235	IVQIPRISD	GIVQIPRISDQNQKI	4618	48.05
854	868	VGPMPPTDI	ERRGVGPMPPTDIPP	4636	48.14
855	869	VGPMPPTDI	RRGVGPMPPTDIPPS	4638	48.15
1	15	MKKAWLSRA	MKKAWLSRAQQADAG	4646	48.18
461	475	VGPLSVQAR	QEPGPVGPLSVQARL	4651	48.2
555	569	HDQLLPAPG	STEPVHDQLLPAPGP	4693	48.41
576	590	TTSQLSSSA	PLTSPTTSQLSSSAP	4695	48.42
556	570	HDQLLPAPG	TEPVHDQLLPAPGPD	4762	48.72
727	741	VPRQRPRGA	PQVPPVPRQRPRGAP	4762	48.72
707	721	VPRQRPRGA	PEVPPVPRQRPRGAP	4775	48.77
490	504	EESMQGVQV	LLHEESMQGVQVHGS	4800	48.88
499	513	VHGSMLDLL	VQVHGSMLDLLEKDD	4817	48.96
575	589	TSPTTSQLS	QPLTSPTTSQLSSSA	4837	49.05
464	478	VGPLSVQAR	GPVGPLSVQARLEPW	4914	49.39
293	307	HTHIREWFR	HTHIREWFRQCTGRP	4936	49.5
557	571	HDQLLPAPG	EPVHDQLLPAPGPDP	4943	49.53
470	484	LEPWQPLPG	SVQARLEPWQPLPGP	4962	49.61
856	870	VGPMPPTDI	RGVGPMPPTDIPPSK	4971	49.66
582	596	LSSSAPSCA	TSQLSSSAPSCAQTP	4987	49.72

Legend to Table S5

The analysis of binding of the peptides to HLA DR2b was done using the Immune Epitope Data Base or IEDB (www. iedb.org). The default peptide length of 15 amino acids was used in the analysis but the sequence of the core nonamer peptide that is expected to bind to the HLA DR molecule and constitute the major portion of the T cell epitope is shown (column 3) for each 15 mer sequence (column 4). The starting and ending amino acid residue number for each nonamer in the protein sequence are shown in columns 1 and 2 respectively. The

analysis was done using the Stabilised Matrix Method (SMM) where the peptides are ranked according to their predicted binding affinities or IC_{50} (column 5) which indicates the concentration of peptide in nM expected to achieve 50% saturation of the HLA DR molecule. Therefore a lower IC_{50} shows a higher affinity. As a guide, peptides with IC_{50} values <50 nM are considered to bind with high affinity, between 50nM to 500 nM with intermediate affinity and between 500nM to 5000 nM with low affinity. For each peptide, a percentile rank (columns 6) is generated by comparing the peptide's score against the scores of five million random 15 mers selected from the SWISSPROT protein database. Therefore smaller percentile rank values, typically <10, also indicate higher affinity and specificity of binding to the HLA DR molecule.

Table S6 - Results of the HLA DR2b binding assays on peptides

Peptide No.	MOG		Affinity	Stability In	Predicted IC_{50}	Affinity vs MBP Cont (%)	Stability Index vs MBP Cont%
1 MOG_1	FVIVPVLGPLVALII	*	10.8	1.1	40	54	52
2 MOG_2	LFVIVPVLGPLVALI	*	22	1	38	111	48
3 MOG_3	TLFVIVPVLGPLVAL	*	25.9	2	41	130	95
4 MOG_4	ITLFVIVPVLGPLVA	*	30	2.3	42	151	110
5 MOG_5	KITLFVIVPVLGPLV		19.1	1.4	43	96	67
	Syncytin 1						
1 SYN1_1	PWILPFLGPLAAIIL	*	27.9	2.1	39	140	100
2 SYN1_2	MPWILPFLGPLAAII	*	28.6	2.6	39	144	124
3 SYN1_3	WMPWILPFLGPLAAI	*	14	0.6	41	70	29
4 SYN1_4	QWMPWILPFLGPLAA	*	20.3	1.6	44	102	76
5 SYN1_5	SQWMPWILPFLGPLA		6.9	0	44	35	0
	MSRV						
1 MSRV_1	LPFLGPLAAIIFLLL	*	0.1	0	296	1	0
2 MSRV_2	TLPFLGPLAAIIFLL	*	2.8	0	294	14	0
3 MSRV_3	WTLPFLGPLAAIIFL	*	12.5	1.1	297	63	52
4 MSRV_4	PWTLPFLGPLAAIIF		11.9	0.8	197	60	38
5 MSRV_5	MPWTLPFLGPLAAII	*	13.8	0.7	246	69	33
6 MSRV_6	WMPWTLPFLGPLAAI		9.2	0	284	46	0
7 MSRV_7	QWMPWTLPFLGPLAA		15	0	300	75	0
	Syncytin 2						
1 SYN2_1	WVLPLTGPLVSLLLL		12.7	1	308	64	48
2 SYN2_2	SWVLPLTGPLVSLLL		21.4	1.6	133	108	76
3 SYN2_3	FSWVLPLTGPLVSLL		15.3	1.3	130	77	62
4 SYN2_4	WFSWVLPLTGPLVSL		26.2	2.8	142	132	133
5 SYN2_5	KWFSWVLPLTGPLVS		69.3	3.8	149	348	181
6 SYN2_6	WKWFSWVLPLTGPLV		0.2	0	147	1	0
	β Synuclein & EBNA1						
1 βSYN	EKTKEGVLYVGSKTR		18.2	2	284	91	95
2 EBNA1_1	AGVFVYGGSKTSLYN		21	0.5	488	106	24
3 EBNA1_2	VAGVFVYGGSKTSLY		23.4	0.9	410	118	43
4 EBNA1_3	WVAGVFVYGGSKTSL		3.5	0.2	429	18	10
	MBP						
1 MBP_1	PVVHFFKNIVTPRTP		11.8	4.3	26	59	205
2 MBP_2	NPVVHFFKNIVTPRT		15.8	0	21	79	0
3 MBP_3	ENPVVHFFKNIVTPR	*	19.9	2.1	21	100	100

4 MBP_4	DENPVVHFFKNIVTP		29.2	0	22	147	0
5 MBP_5	QDENPVVHFFKNIVT		19.6	0	21	98	0
	OSP&EBNA1						
1 OSP homolog to EBNA1 _1	STTLRALAPRLMRRV		183.9	23.3	149	924	1110
2 EBNA1 homolog to OSP_1	NIAEGLRALLARSHV		1.8	0	288	9	0
3 EBNA1 homolog to OSP_2	AEGLRALLARSHVER		2.8	0	306	14	0
4. OSP homolog to EBNA1_2	AGVLLILLALCALVA		0.3	0	33	2	0
Protein origin & peptide characteristic	**Other tested peptides**						
MBP	GTLSKIFKLGGRDSR		0.6	0	940	3	0
MOG 175-189	FLCLQYRLRGKLRAE		7.4	0	13.75	37	0
EBV DNA POL	TGGVYHFVKKHVHES		16.7	0	4.12	84	0
PLP	GTASFFFLYGALLLA		0	0	0.43	0	0
PLP	YGTASFFFLYGALLL		0	0	0.45	0	0
PLP	TASFFFLYGALLLAE		35.9	3.9	0.45	180	185
PLP	ASFFFLYGALLLAEG		1.4	0	0.46	7	0
PLP	SFFFLYGALLLAEGF		No synthesis	No synthes	1.04	n.d	n.d
MSRVenv homologous to MBP/PLP	LFTVLLPPFALTAPP		2.1	0	5.68	11	0
MSRVenv homologous to MBP/PLP	TFLFTVLLPPFALTA		2.3	0	2.44	12	0
MSRVenv homologous to MBP/PLP	HTFLFTVLLPPFALT		0.4	0	2.5	2	0
MSRVenv homologous to MBP/PLP	YHTFLFTVLLPPFAL		0	0	2.52	0	0
MSRVenv homologous to MBP/PLP	PYHTFLFTVLLPPFA		0	0	3.67	0	0
MSRVenv homologous to MBP/PLP	LPYHTFLFTVLLPPF		0	0	738	0	0
SYN1 homologous to MBP/PLP	YHIFLFTVLLPSFTL		2.4	0	95	12	0
SYN1 homologous to MBP/PLP	PYHIFLFTVLLPSFT		0	0	4.21	0	0
SYN1 homologous to MBP/PLP	LPYHIFLFTVLLPSF		0.2	0	274	1	0
SYN1 homologous to MBP/PLP	ALPYHIFLFTVLLPS	*	0.6	0	5.07	3	0
SYN1 homologous to MBP/PLP	MALPYHIFLFTVLLP	*	4.5	0	5.11	23	0
MSRVenv weaker homology to MOG	NTTSVLVGPLVSNLE		0.5	0.6	10	3	29
MSRVenv weaker homology to MOG	INTTSVLVGPLVSNL		0.1	0	12.25	1	0
MSRVenv weaker homology to MOG	EINTTSVLVGPLVSN		0.4	0	13.84	2	0
MSRVenv weaker homology to MOG	TEINTTSVLVGPLVS		0	0	13.93	0	0
MSRVenv homologous to OSP	PLAAIIFLLLFGPCI		0	0	85	0	0
MSRVenv homologous to OSP	LAAIIFLLLFGPCIF		0.1	0	69	1	0
OSP homologous to MSRVenv	GLPAILLLLTVLPCI		15.2	1.2	267	76	57
OSP homologous to MSRVenv	LPAILLLLTVLPCIR		5.1	0.3	268	26	14

OSP homologous to MSRVenv	PAILLLLTVLPCIRM	9.6	0.7	269	48	33
OSP homologous to MSRVenv	ILLLLTVLPCIRMGQ	43.1	4.9	425	217	233

Legend to Table S6

Control MBP_3 peptide (yellow highlight) was the internal standard against which relative per cent affinity and stability index of other peptides were measured
n.d = not done
1st column shows the names and characteristics of the peptides
2nd column shows amino acid sequence of the tested peptides with the nonamer predicted to bind HLA DR2b in red letters
3rd column * shows mean affinity & stability index from two independent REVEAL® binding assays. All others values of affinity and stability are from a single assay.
4th column shows the raw affinity index in the REVEAL® assay
5th column shows the raw stability index in the REVEAL® assay
6th column shows the IC_{50} for HLA DR2b binding predicted in the IEDB-SMM analysis
7th column shows the percent relative affinity compared with the internal control peptide MBP_3
8th column shows the percent relative stability compared with the internal control peptide MBP_3