

DEPARTAMENT DE BIOLOGIA CEL·LULAR, FISIOLOGIA I IMMUNOLOGIA
UNIVERSITAT AUTÒNOMA DE BARCELONA

Programa de Doctorat en Immunologia

**MHC I MALALTIES AUTOIMMUNITÀRIES:
CONTRIBUCIÓ ESPECÍFICA DELS AL·LELS DE SUSCEPTIBILITAT**

Carol Guitart

Bellaterra, novembre de 2014

DEPARTAMENT DE BIOLOGIA CEL·LULAR, FISIOLOGIA I IMMUNOLOGIA
UNIVERSITAT AUTÒNOMA DE BARCELONA

Memòria presentada per

Carol Guitart Marín

Per a optar al grau de

DOCTORA EN IMMUNOLOGIA

Tesi realitzada en el laboratori d'Immunologia Cel·lular de l'Institut de Biotecnologia i Biomedicina de la Universitat Autònoma de Barcelona, sota la direcció de la Dra. Dolores Jaraquemada i la Dra. Montserrat Carrascal.

La doctoranda

Les directores

CAROL GUITART MARÍN

Dra. DOLORES JARAQUEMADA
Catedràtica d'Immunologia, UAB

Dra. MONTSERRAT CARRASCAL
Científic Titular, IIBB-CSIC

Bellaterra, novembre de 2014

INDEX

INDEX

INTRODUCCIÓ	11
1. El complex principal d'histocompatibilitat	13
1.1. Organització genòmica del MHC	13
1.2. Les molècules d' MHC com a presentadores d'antigen	15
1.3. Les molècules d' MHC de classe II	18
1.3.1. Síntesi de la molècula	19
1.3.2. Caracterització del motiu d'unió	21
2. Tolerància en el context MHC	23
2.1. Tolerància central	24
2.2. Tolerància perifèrica	24
2.3. Trencament de la tolerància	26
3. Autoimmunitat i malalties autoimmunitàries	27
3.1. Malalties autoimmunitàries de la tiroide	27
3.1.2. Malaltia de Graves-Basedow (GB)	29
 Objectius	 35
 Materials i mètodes	 39
1. Línies cel·lulars	41
1.1. Cultius de cèl·lules en suspensió	41
1.2. Cultius de cèl·lules adherents.....	41
2. Teixit de ratolí Balb/c	42
2.1. Mostres de timus i melsa de ratolí Balb/c estàndards	42
2.2. Mostres de timus i melsa de ratolí Balb/c transgènic	42
3. Transfeccions	41
3.1. Transfecció de la línia HT93	42
4. Citometria de flux	44
4.1. Marcatge superficial	44
4.2. Marcatge intracel·lular	45
5. Elució dels lligands naturals d'MHC de classe II	45
5.1. Obtenció d'anticossos anti-HLA-DR	47
5.1.1. Producció i purificació de l'anticòs B8.11.2	47
5.1.2. Determinació de la concentració d'anticòs	48
5.1.3. Determinació de l'activitat de l'anticòs B8.11.2	48
5.1.4. Determinació de la puresa de l'anticòs per SDS-PAGE	49

5.2. Obtenció dels anticossos anti-I-A ^d i anti-I-E ^d	50
5.3. Preparació de les columnes per la cromatografia d'afinitat	50
5.4. Immunopurificació de pèptids a partir de pellets secs provinents de línies cel·lulars.....	52
5.5. Immunopurificació de pèptids a partir de teixit congelat de ratolí Balb/c	52
5.6. Elució dels complexos de mhc de classe II-pèptid	53
5.6.1. Separació dels pèptids de les cadenes α i β per ultrafiltració.....	53
5.6.2. Gels de proteïnes dels retinuts del centricó	54
5.7. Dessalat de les mostres per C18	54
5.8. Fraccionament de les mostres per cromatografia líquida d'intercanvi catiònic	55
6. Identificació dels repertoris de pèptids associats a MHC mitjançant espectrometria de masses	56
7. Identificació i validació els espectres de fragmentació	56
7.1. Cerca base de dades i avaluació automàtica	56
7.2. Avaluació manual	57
8. Anàlisi predictiva d'afinitat	57
9. Determinació dels motius d'ancoratge	58
10. Establiment del protocol de simulació pel modelatge de les molècules I-A ^d i I-E ^d	59
10.1. Molècula I-A ^d	60
10.2. Molècula I-E ^d	62
Objectiu específic 1 i antecedents	63
Capítol 1: repertori peptídic presentat per HLA-DR3: revisió del motiu d'ancoratge	67
1.1. Processament de les mostres	69
1.2. Repertori de pèptids associats a HLA-DR3	69
1.3. Descripció del motiu d'ancoratge D'HLA-DR3	74
1.3.1. Identificació dels nonàmers d'unió a HLA-DR3	74
1.3.2. Selecció dels nonàmers.....	76
1.3.3. Generació de les matrius per la caracterització del motiu d'unió a DR3	76

1.3.4. Control del sistema de generació de les matrius	78
1.3.5. Comparació dels resultats amb les dades prèviament publicades	79
1.4. Predicció d'afinitat teòrica dels lligands d'HLA-DR3	81
1.5. Observacions finals	82

Capítol 2: influència de les xaperones HLA-DM i cadena invariant (Ii)

en els repertoris peptídics presentats per HLA-DR3	85
2.1. Processament de les mostres	87
2.2. Repertoris de pèptids associats a HLA-DR3 en cèl·lules transfectades	
RIN-M5F: DR3, DR3-II, DR3-DM, DR3-DM-II	87
2.2.1 RINDR3	87
2.2.2. RINDR3DM	90
2.2.3. RINDR3II	92
2.2.4. RINDR3DMII	94
2.3. Comparació general entre els repertoris	99
2.3.1. Comparació dels repertoris comuns i únics entre les línies	
transfectants	100
2.3.2. La via citosòlica és la ruta de processament preferent per a	
HLA-DR3	100
2.3.3. Predicció d'afinitat dels lligands d'HLA-DR3	102
2.3.4. El patró de processament en DR3 està lligat a extrems terminals	103
2.4. Comparació dels repertoris únics a les cèl·lules transfectants DR3	107
2.5. Comparació dels repertoris obtinguts a RIN5mf transfectada amb DR3	
i amb DR4.....	109
2.6. Observacions finals	112

Capítol 3: Presentació del TSHR, autoantígen dominant a la malaltia de

Graves-Basedow.....	115
3.1. Obtenció d'un model cel·lular epitelial amb expressió d'HLA-DR	
i TSHR.....	117
3.2. Processament de les mostres	118
3.3. Repertoris de pèptids associats a HLA-DR en cèl·lules transfectades	
HT93	118
3.3.1. Descripció del repertori associat a HT93CIITA	119
3.3.2. Descripció del repertori associat a HT93DT	123

3.4. Predicció d'afinitat dels lligands d'HLA-DR	129
3.5. Patró de processament	130
3.6. Comparació entre ambdós repertoris	132
3.7. Observacions finals	133
 Objectiu específic 2 i antecedents	 135
Capítol 4: descripció del motiu d'ancoratge dels al·lells de classe II d'H-2 ^d	139
4.1. Processament de les mostres	141
4.2. Repertori de pèptids associats a I-A ^d i I-E ^d	141
4.3. Comparació de les seqüències de les cadenes β dels diferents al·lells	149
4.4. Motiu d'ancoratge per la molècula I-A ^d	151
4.4.1 Selecció dels nonàmers4.4.2. generació de les matrius per la caracterització del motiu d'unió a I-A ^d	151
4.4.2. Generació de les matrius per la caracterització del motiu d'unió a I-A ^d	152
4.4.3. Comparació dels resultats amb les dades prèviament publicades	154
4.5. Motiu d'ancoratge per la molècula I-E ^d	155
4.5.1 Selecció dels nonàmers	155
4.5.2. Generació de les matrius per la caracterització del motiu d'unió a I-A ^d	154
4.5.3. Comparació dels resultats amb les dades prèviament publicades.....	157
4.6. Modelatge de les molècules I-A ^d i I-E ^d	158
4.7. Re-evaluació del motiu d'ancoratge per la molècula I-E ^d	164
4.8. Predicció d'unió de pèptids naturals i comparació amb epítops definits de cèl·lules T	165
4.9. Observacions finals	167
 Capítol 5: Estudi comparatiu entre els repertoris peptídics de timus i melsa en ratolins Balb/c	 169
5.1. Absència de diferències als repertoris peptídics de ratolins Balb/c i ratolins transgènics per al TSHR	171
5.2. Estudi del repertori peptídic presentat en timus de Balb/c	174
5.2.1. Repertoris peptídics obtinguts per timus Balb/c	174
5.2.2. Origen i via de processament del repertori peptídic en timus Balb/c	175

5.2.3. Patró de processament dels pèptids aïllats de timus Balb/c	176
5.3. Estudi del repertori peptídic presentat en melsa de Balb/c	177
5.3.1. Repertoris peptídics	177
5.3.2. Distribució de la mida	178
5.3.3. Origen i via de degradació del repertori peptídic en melsa Balb/c	179
5.3.4. Patró de processament dels pèptids aïllats en melsa Balb/c	179
5.4. Comparació entre el repertori peptídic presentat en timus i melsa de Balb/c	180
5.4.1. Timus i melsa no comparteixen un ampli repertori peptídic	181
5.4.2. Comparació de les rutes de processament de pèptids entre teixits	182
5.5. Identificació de pèptids derivats d'antígens perifèrics (TRAS) presentats per MHC en timus de ratolins balb/c	183
5.6. Observacions finals	184
Conclusions	185
Bibliografia	189

1. EL COMPLEX PRINCIPAL D'HISTOCOMPATIBILITAT (MHC)

1.1. Organització genòmica del MHC

El complex major d'histocompatibilitat és un complex poligènic, amb gens polimòrfics i codominants, que presenta la més alta densitat de gens en mamífers. En humans s'anomena HLA (*Human Leucocyte Antigens*) i en ratolí s'anomena complex H-2. La distribució dels seus gens és diferent en ambdues espècies però mantenen una certa organització en tres regions, una per cada classe d'MHC en humans (figura 1) i quatre regions en ratolins.

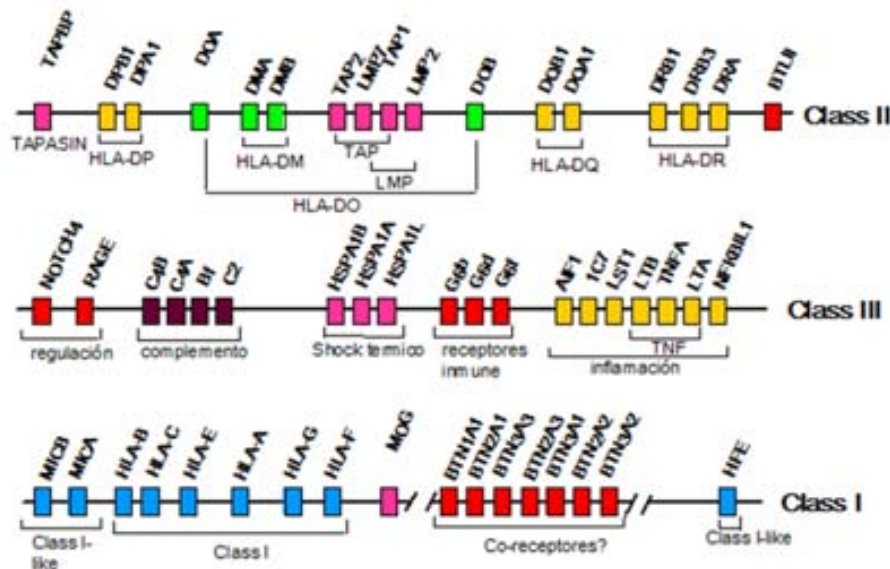


Figura 1: Estructura genètica de l'HLA. Extret de la web <http://epidemiologiamolecular.com>

Molècules HLA

- Les molècules de classe I vénen codificades pels loci A, B i C, donant lloc a les molècules clàssiques de classe I, HLA-A, HLA-B i HLA-C juntament amb gens no clàssics com HLA-E, HLA-G, MICA i MICB, entre d'altres. L'heterodímer HLA està format per una cadena anomenada α que s'uneix de manera no covalentment a una proteïna extracel·lular anomenada β_2 -microglobulina (β_2 -m). La cadena α està formada per tres dominis extracel·lulars denominats: $\alpha 1$, $\alpha 2$ i $\alpha 3$; una regió transmembrana i una petita cua citoplasmàtica. La β_2 -m està unida no covalentment al domini $\alpha 3$.

Entre els dominis $\alpha 1$ i $\alpha 2$ es forma una cavitat on s'uneix el pèptid (figura 2).

- Les molècules de classe II vénen codificades pels loci DP, DQ i DR en humans, donant lloc a les molècules clàssiques: HLA-DR, DP i DQ, no clàssiques com DM i DO, i altres com el transportador associat amb processament antigènic (TAP) o les dues subunitats específiques del immunoproteasoma, totes involucrades en la generació de complexos HLA-pèptid. Les molècules del MHC de classe II són glicoproteïnes de membrana, formades per l'associació no covalent de dues cadenes, α i β , de 32-34 kDa i 29-32 kDa, respectivament (figura 2).

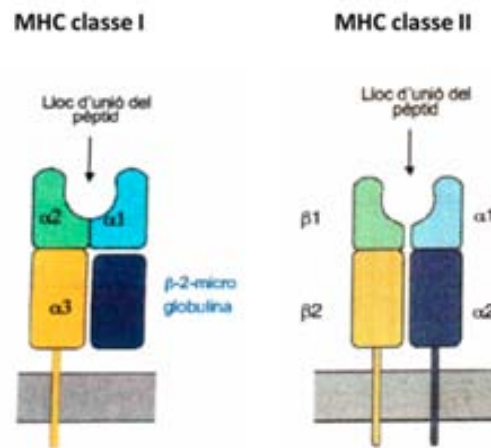


Figura 2: Comparació entre les heterodímers que formen la molècula de classe I i classe II. Extret de Janeway et al: Immunobiologia (Masson, Barcelona)

- Les molècules de classe III venen codificades per les regions C4, C2 i BF que originen diverses proteïnes del complement així com $\text{TNF}\alpha$ i $\text{TNF}\beta$ [1]

Els gens que codifiquen per la cadena pesada de les molècules de classe I, i les cadenes α i β de les molècules de classe II, es localitzen al braç curt del cromosoma 6 humà [2], juntament amb altres gens que codifiquen per molècules importants en el funcionament del sistema immunitari (figura 1). El gen que codifica per la $\beta 2m$ està localitzat al cromosoma 15 humà [3].

Sistema H-2

El complex major d'histocompatibilitat en el ratolí (sistema H-2) es localitza en el braç curt del cromosoma 17 i inclou diversos loci. Com en ocorre amb l'HLA,, els gens H-2 s'hereten en bloc (com haplotips), són codominants i es segreguen seguint les lleis de Mendel.

- Les molècules de classe I vénen codificades pels loci K, L i D, donant lloc a les molècules H-2K, H-2L i H-2d. A l'igual que ens humans, aquests gens sintetitzen la cadena α que també s'uneix a la β_2 -microglobulina (codificada pel cromosoma 2).
- Les molècules de classe II vénen codificades pels loci IA i IE (similars als gens DQ i DR en l'humà) contenen els gens A α /A β i E α /E β que codifiquen els antígens MHC de classe II (I-A i I-E).
- Dins del complex H-2 també està inserida una regió S on es troben els gens corresponents als antígens de classe III (C2 i C4).

1.2. Les molècules d'MHC com a presentadores d'antigen

Els grups de gens polimòrfics del MHC de classe I i II, codifiquen dos grups de proteïnes estructuralment diferents però homòlogues, la funció de les quals és unir pèptids per presentar-los a les cèl·lules T. Les molècules de classe I presenten pèptids a les cèl·lules CD8⁺ mentre que les de classe II presenten pèptids a les cèl·lules CD4⁺.

Estudis de resolució d'estructures cristal·lines de les molècules de MHC de classe I i II han permès determinar les característiques estructurals i funcionals d'aquestes. Les molècules de classe I i II tenen una estructura semblant que els permet realitzar la seva funció essencial de presentació de pèptids a les cèl·lules T. Ambdues tenen un solc extracel·lular format per dues hèlix α i 8 fulles β plegades dins el qual es disposa el pèptid, un parell de dominis tipus immunoglobulina (Ig) que contenen llocs d'unió pels coreceptors CD4 i CD8, i

estan ancorades a la cèl·lula mitjançant dominis transmembrana i citoplasmàtics [4-8] (figura 3)

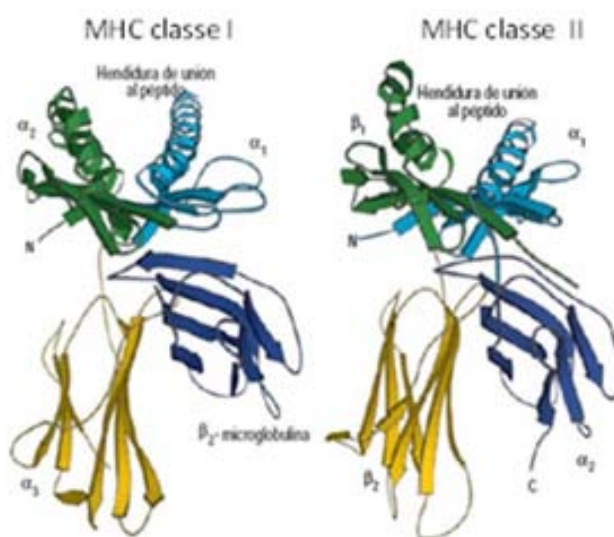


Figura 3: Estructura tridimensional de les molècules de MHC I i MHC II. L'heterodímer de classe I consisteix en una cadena α que travessa la membrana, associada de manera no covalent a la β_2 microglobulina. El solc d'unió al pèptid està format pels dominis α_1 i α_2 de la cadena. La molècula de classe II està formada per dues cadenes α i dues β formades cadascuna per dos dominis. Els dominis α_1 i β_1 formen el solc d'unió al pèptid. Extret de Janeway et al: Immunobiologia (Masson, Barcelona)

En ambdues tipus de molècules MHC, els solcs d'unió estan constituïts per diferents subcavitats (pockets) que serveixen com a punts d'ancoratge pels pèptids que s'associen. Dins del solc, el pèptid adopta una conformació estesa, d'hèlix de poliprolina tipus II, amb les cadenes laterals d'alguns residus encarades cap al solc d'unió i amb d'altres exposades al solvent per tal de ser reconegudes [9]. A ambdós tipus de molècules les regions més polimòrfiques es troben en el solc d'unió al pèptid i pròxims a aquest i són les responsables de les diferents especificitats peptídiques. D'aquesta manera la gran variabilitat d'aminoàcids d'aquesta regió fa que diferents molècules de MHC uneixin i presentin pèptids diferents i siguin reconeguts per diferents TCR. Ambdós molècules (classe I i II) són inestables quan no tenen un pèptid unit.

La interacció del complex trimolecular s'estabilitza per l'establiment de ponts d'hidrogen entre residus conservats del solc de les molècules de MHC i la cadena principal o els grups terminals del pèptid associat. En les molècules

MHC de classe I, aquestes interaccions involucren principalment a residus agrupats en els extrems del solc i els grups N- i C-terminals lliures dels pèptids [10] que tenen un tamany predominant de 9 aminoàcids [11].

D'altra banda, les molècules de MHC II estableixen ponts d'hidrogen conservats amb residus localitzats al llarg de tota la cadena principal del pèptid. En aquestes molècules els extrems del solc són oberts i els pèptids que s'hi uneixen poden sobresortir pels dos costats. Com a conseqüència, els pèptids associats a classe II presenten un rang de longitud major i més variable d'entre 10-22 residus [12, 13]. Els pèptids tenen una regió central d'unió a MHC o "core" composta per 9 aminoàcids [9] i uns residus en ambdós extrems d'aquest "core" que sobresurten del solc d'unió a l'antigen. Aquests extrems s'anomenen PFRs (peptide flanking residues) i poden interaccionar amb l'heterodímer MHC II augmentant l'afinitat de la unió i estabilitzant la conformació del complex [14, 15]. Les posicions relatives P1, 4, 6, 7 i 9 dels pèptids són les que habitualment interaccionen amb els residus de les butxaques dels solcs d'unió. El conjunt de residus d'ancoratge de cada variant al·lèlica és diferent i s'anomena motiu d'unió (figura 4).

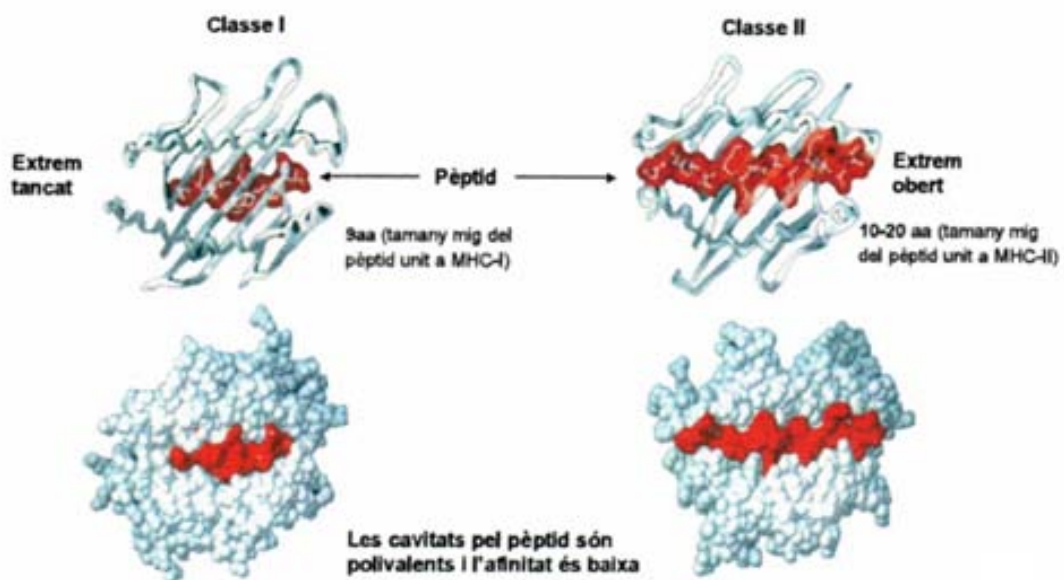


Figura 4: Distribució dels aminoàcids polimòrfics dins el solc d'unió al pèptid en MHC de classe I i en classe II

El polimorfisme del solc genera variacions en les característiques bioquímiques dels *pockets* en funció de la variant al·lèlica, és a dir, cada al·lel té uns butxaques que uneixen preferentment un determinat patró d'aminoàcids. Suri et al. varen demostrar que una variació en un únic aminoàcid en la cadena beta als al·lells de MHC-II, I-A^d i I-Ag7, feia que aquests al·lells presentessin pèptids amb característiques diferencials tot i tenir les mateixes proteïnes disponibles per al seu processament [16]. És important conèixer els repertoris peptídics de MHC-II en funció de l'al·lel que el presenta (caracterització dels motius d'unió) a l'hora d'entendre la resposta immunitària o de dissenyar estratègies de vacunació o immunoteràpia.

1.3. Les molècules d'MHC de classe II

Les molècules de classe II són glicoproteïnes de membrana formades per dues cadenes, α i β , associades no covalentment. Al contrari que les de classe I que són ubiquïes, les molècules de classe II s'expressen només en cèl·lules presentadores d'antigen professionals (APC) i uneixen pèptids que presenten a cèl·lules T CD4⁺. Uneixen pèptids dins d'un rang de mida d'entre 9 i 22 aminoàcids tot i que són capaces d'unir pèptids de més de 30 residus ja que, a diferència de les de classe I, gràcies a que tenen els extrems del lloc d'unió oberts [13, 17, 18]. La principal font de generació d'aquests pèptids són proteïnes externes o de membrana [17] que són internalitzades i degradades per la via endo-lisosomal (endocítica) [19, 20]. En molts casos els pèptids presentats per molècules de classe II formen famílies de pèptids solapats o "nested sets" que són seqüències amb la mateixa regió central però que varien en els extrems N o C-terminals, suggerint la possibilitat que es puguin generar per proteòlisi post unió al MHC (figura 5). Aquesta característica és típica dels pèptids de classe II.

P01903	HLA class II histocompatibility antigen DR alpha chain	GALANIAVDKANLE QGALANIAVDKANLE AQQALANIAVDKANLE EAQQALANIAVDKANLE AQQALANIAVDKANLEI EAQQALANIAVDKANLEI
--------	--	---

Figura 5: Exemple de seqüències “nested set” d’un pèptid aïllat de la línia HT93DT. Els *nested sets* són famílies de pèptids amb la mateixa regió nucli (*core*, en vermell) que varien en la longitud dels seus dos extrems

1.3.1. Síntesi de la molècula

Les cadenes α i β del MHC-II es sintetitzen en els ribosomes adherits al reticle endoplasmàtic (RE) i es transfereixen durant la seva síntesi cap a la llum del RE, on es pleguen i s'acoblen entre si. El procés de formació i maduració del complex està regulat per la xaperona calnexina (IP90), que s'associa als seus components [21] evitant la sortida cap a la ruta secretora dels complexos en període de formació. Aquest dímer però, es inestable i per estabilitzar-se requereix la unió d’una altra xaperona, la cadena invariant (Ii), que a més evita la unió prematura de pèptids residents al reticle endoplasmàtic [22]. Aquesta facilita la formació d’un complex multimèric format per tres heterodímers $I_i(\alpha\beta)_3$.(figura 6)

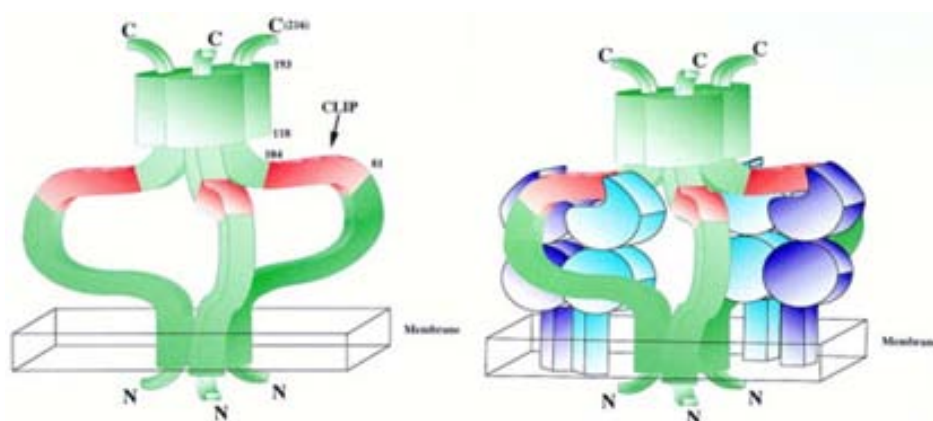


Figura 6. Models del trímer de la cadena Ii (esquerra) i del complex MHC II-Ii(dreta). La porció N-terminal del domini lumenal de la cadena invariant s’esten, permetent que CLIP pugui interactuar amb el solc d’unió de classe II. [22]

Aquest nonàmer es dirigeix directament a superfície a través de la via secretora, on es internalitza [23] gràcies a un senyal d'internalització del domini citoplasmàtic de la li o bé va directament cap a la via endocítica a través del Golgi gràcies a un senyal situat també a la regió citoplasmàtica de li [24]. Els compartiments endosòmics on s'acumulen més complexos MHC-II es denominen MIIC (MHC class II compartment) i són sobretot MVB (cossos multivesiculars) o MLB (cossos multilaminars). Aquestes vesícules són de pH baix i corresponen a fases tardanes de la ruta endosòmica, molt properes als lisosomes i entren en contacte o es fusionen amb endosomes entrants que contenen proteïnes o pèptids en el seu interior (figura 7). En aquests compartiments la li és degradada de forma seqüencial per l'acció de les proteases de la via endocítica [25].

Figura 7: Via de processament d'antigen per MHC de classe II. [26]

El primer fragment que es genera és el LIP22 degut a la pèrdua del domini de trimerització situat a l'extrem C-terminal. L'absència d'aquest domini produeix la desestabilització del nonàmer donant lloc a tres complexos LIP22- $\alpha\beta$. Seguidament es genera LIP10 a causa de la degradació de la resta de l'extrem C-terminal fins a la regió CLIP (pèptid de la cadena invariant associat a classe II). Amb això, s'eliminen els residus glicosilats situats en aquesta regió el que permet l'últim pas degradatiu realitzat, en funció del tipus cel·lular, per les

cathepsines S, L o F [27]. L'acció d'aquestes proteases elimina el domini citoplasmàtic i transmembrana deixant el fragment CLIP, protegit pel solc d'unió, unit al dímer $\alpha\beta$ [28]. Un cop format el complex $\alpha\beta$ -CLIP, els pèptids procedents de la degradació de material a la via endocítica competeixen amb el pèptid CLIP pel solc d'unió. L'intercanvi de CLIP per un altre pèptid es realitza amb la col·laboració de la xaperona HLA-DM, que estabilitza el dímer $\alpha\beta$ permetent l'alliberament de CLIP i realitzant a més, una funció d'editor peptídica de la qualitat i quantitat dels pèptids seleccionats [29]. Els dímers en absència de pèptid no són estables i seran degradats en els compartiments lisosomals, en canvi, les molècules carregades seran transportades a la superfície per a la presentació del pèptid. En absència d'infecció, els pèptids presentats per molècules de MHC-II deriven de proteïnes autòlogues, ja siguin de la pròpia cèl·lula o de l'espai extracel·lular. La biosíntesi de les molècules de MHC de classe II condiona així el repertori que presenta, especialitzant-se en la presentació de pèptids procedents de proteïnes degradades en la ruta endocítica, majoritàriament proteïnes extracel·lulars captades per APCs o components internalitzats de la membrana plasmàtica.

Tot i així, hi ha rutes alternatives que aporten variabilitat al repertori peptídica presentat per MHC II i que en alguns casos poden ser molt rellevants [30]. L'autofàgia, un procés relacionat amb l'estrès cel·lular i constitutiu en cèl·lules epitelials tímiques, és un mecanisme utilitzat per les APCs per presentar components citosòlics per la via de classe II. També ho és la fagocitosi de cossos apoptòtics de cèl·lules veïnes [31] [32] [33]. D'altra banda, s'ha descrit que algunes proteases solubles o de membrana són capaços de generar lligands per MHC-II a l'espai extracel·lular que, si tenen alta afinitat per la molècula, podrien competir pels lligands presentats per les molècules de MHC-II de superfície, desplaçant-los in situ o podrien ser internalitzats, incorporant-se a la ruta clàssica de MHC-II [34].

1.3.2. Caracterització del motiu d'unió

L'existència de motius d'unió específics d'al·lel es va fer evident a principis dels anys 90 a partir de la caracterització dels repertoris peptídics naturals de MHC

[35-38]. L'estratègia que utilitzaven estava basada en la seqüenciació en "pool" d'Edman, la qual cosa implicava la seqüenciació directa de barreges de pèptids. Els complexos MHC-pèptid es purificaven per immunoprecipitació i mitjançant elució àcida i ultracentrifugació es separaven els pèptids de les molècules d'MHC. Per disminuir la complexitat de la mostra es fraccionaven els pèptids per cromatografia líquida HPLC. Els diferents cicles de la seqüenciació d'Edman permetien determinar els aminoàcids més freqüents en cada una de les posicions de la seqüència peptídica, caracteritzant així la presència de residus específics en aquestes posicions. Aquests estudis són més eficients amb els pèptids de classe I que amb els de classe II degut a la major homogeneïtat de mida d'aquests (8-11aa), però tot i així la seqüenciació d'Edman també s'ha utilitzat per identificar lligands de classe II molt abundants [13, 18, 36, 39].

La regió peptídica d'unió al MHC-II és de 9 aminoàcids, però hi ha posicions més determinants que d'altres en la interacció. Les interaccions entre cadenes laterals dels lligands poden augmentar l'afinitat d'unió i l'especificitat dels lligands o interferir amb els residus del pocket disminuint la unió [40]. La posició d'ancoratge P1 es considera la més rellevant en la interacció pèptid-MHC. El motiu d'unió a aquest *pocket*, tot i que amb diferències en la freqüència d'aminoàcids que l'ocupen, està bastant conservat entre tots els al·lels d'HLA-DR. P1 està associat a aminoàcids no polars com Val, Leu, Ile, Met o Ala i amb residus aromàtics, grans i també apolars com la Phe, Trp o Tyr. Aquest coneixement i els avenços tecnològics de l'espectrometria de masses, van permetre l'estudi i caracterització massiva de lligands de classe II i la caracterització dels motius d'unió de molts dels al·lels d'HLA-II.

La informació corresponent a cada motiu d'unió pot presentar-se en forma de matriu d'afinitat, on s'indica la importància relativa de la presència dels diferents aminoàcids a cada una de les posicions. Aquestes matrius permeten calcular l'afinitat teòrica d'un pèptid per un determinat al·lel de MHC. Per calcular l'índex d'afinitat teòric d'un determinat pèptid s'assignen valors a cadascun dels residus de la seqüència i la suma d'aquests valors es correlaciona amb l'afinitat d'unió a MHC d'aquest pèptid. Aquest càlcul es basa en la hipòtesis que la unió

d'un pèptid a MHC correlaciona amb el resultat net de tots els efectes de les cadenes laterals d'aquest, i té en compte que la majoria dels efectes de les cadenes laterals depenen de la seva posició relativa a l'aminoàcid ancorat a P1, més que de la seva posició absoluta [40]. La importància de les matrius com a eina de predicció de lligands d'HLA ha estat demostrada per diferents grups [40-42], i s'ha vist que els valors numèrics obtinguts (scores) correlacionen parcialment amb l'afinitat d'unió dels lligands a HLA-DR. Això converteix les matrius en una eina important alhora de predir possibles epítops de cèl·lules T [43].

Altres treballs demostraven que els perfils d'especificitat de les butxaques eren independents de la resta del solc d'HLA-DR [44]. Això és important en la determinació dels perfils de les butxaques, que poden ser compartits entre al·lels sempre i quan tinguin topografies idèntiques [45]. La principal conseqüència de tot això és que amb un petit nombre de perfils es poden construir un gran nombre de matrius.

2. TOLERÀNCIA EN EL CONTEXT MHC

El sistema immunològic ha desenvolupat mecanismes molt diversos per defensar l'organisme d'agents patògens i situacions anormals, minimitzant el dany causat per les respostes de defensa. Aquesta capacitat comporta l'educació dels elements cel·lulars del sistema immunitari, mitjançant la generació de tolerància als antígens propis durant l'ontogènia dels limfòcits, i la necessitat del manteniment d'aquesta tolerància una vegada els limfòcits madurs arriben a perifèria i als òrgans limfàtics secundaris.

La tolerància immunològica és per tant, la manca de resposta per part del sistema immunitari en front a antígens propis (autoantígens) i requereix dos processos: la maduració i selecció en del limfòcits T al timus, destinats a generar un repertori limfoide funcional i no autorreactiu (tolerància central) i els basats en la contenció de respostes autoimmunes per part del repertori madur de cèl·lules T (tolerància perifèrica).

Així els principals mecanismes d'adquisició i manteniment de la tolerància són: deleció, anergia i ignorància clonal. Cada un d'ells pot intervenir a nivell central

(timus o medul·la òssia per a les cèl·lules T i B respectivament) o perifèric (òrgans limfàtics secundaris i altres teixits). Existirien a més altres mecanismes: supressió i interaccions idiotípiques, que si bé tindrien un paper en la regulació de la resposta immune, s'ha proposat que puguin també intervenir en el manteniment de la tolerància

2.1. Tolerància central

La tolerància central és el principal mecanisme d'eliminació de limfòcits autoreactius per selecció clonal eliminació per apoptosi. Pels limfòcits T, consisteix en l'eliminació en el timus, dels limfòcits immadurs que reconeixen complexes MHC/pèptid propis amb elevada o mitjana afinitat; aquest procés es coneix amb el nom de selecció negativa.

L'expressió en el timus d'antígens perifèrics específics de teixit (TRAs) [46-48] juga un paper clau a l'hora de modelar el repertori de cèl·lules T autoreactives, ja que s'ha vist que els nivells d'expressió de determinats antígens correlacionen amb la susceptibilitat a patir malalties autoimmunitàries específiques d'òrgan [46, 47]. També les característiques de la unió de pèptids específica de cada al·lel determinarà el repertori de cèl·lules T a la perifèria [49]. Tot i la seva efectivitat, per alguns antígens la tolerància central té certes limitacions i per tant cèl·lules T autoreactives poden escapar de la selecció. Alguns antígens no s'arriben a expressar mai al timus mentre que altres s'expressen més tard al llarg de la vida d'un individu, un cop ja està format el repertori limfocitari. A més a més, l'organisme està constantment exposat a antígens ambientals no patogènics en front dels quals ha d'haver-hi tolerància, així com organismes comensals dins els nostres sistema respiratori o digestiu. Aquestes limitacions de la tolerància central fa necessari un sistema eficient de manteniment de la tolerància en perifèria (figura 7).

2.2. Tolerància perifèrica

El resultat del procés d'educació tímica és que les cèl·lules T madures que arriben a la perifèria són cèl·lules T autorreactives de baixa afinitat pels autoantígens però que poden presentar una afinitat elevada pels antígens

forans [50] [51] per això és freqüent trobar anticossos i cèl·lules T circulants amb la capacitat de reconèixer antigens propis perifèrics com la peroxidasa tiroidea (TPO) [52], el col·lagen [53]. Aquesta sortida a perifèria de cèl·lules T madures que reconeixen pèptids propis amb baixa afinitat, posa de manifest la necessitat d'altres mecanismes responsables del manteniment de la tolerància a perifèria, per evitar les respostes autorreactives que es podrien donar quan els antigens propis són presentats a les cèl·lules T en processos inflamatoris derivats de la resposta contra patògens. La ignorància immunològica, la inducció d'anergia i la regulació de la resposta immunitària son part dels mecanismes que sinèrgicament mantenen l'equilibri entre tolerància i immunitat a la perifèria.

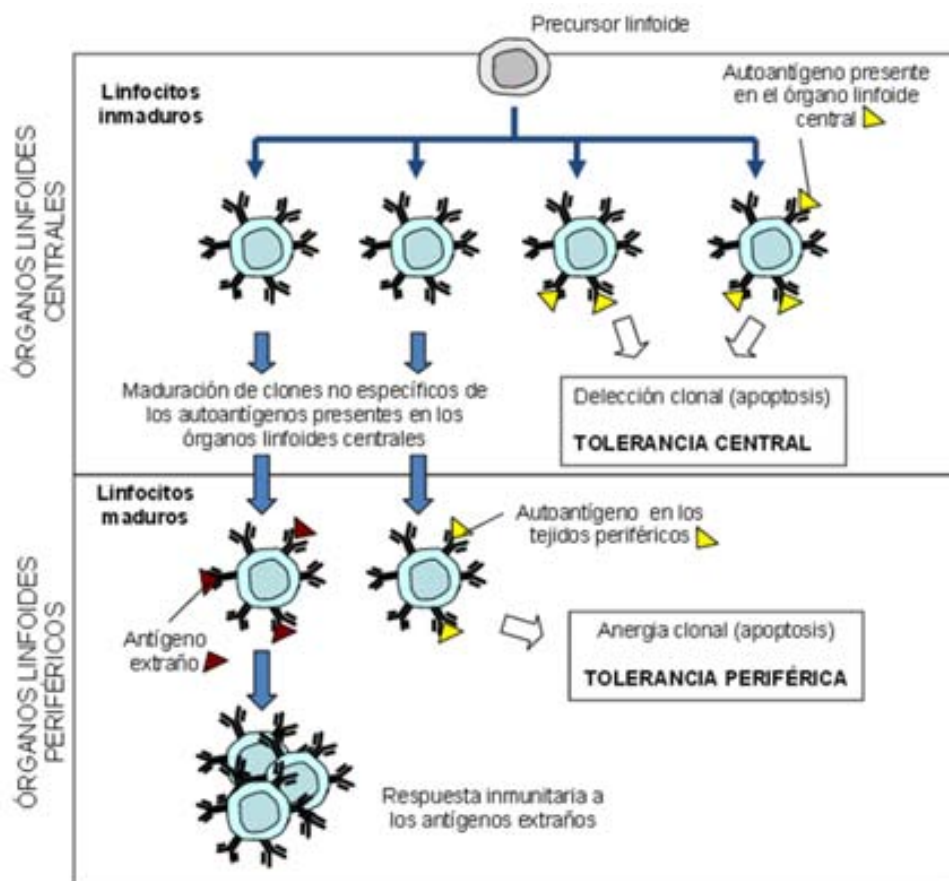


Figura 7: Mecanismes generals d'autotolerància en limfòcits

2.3. Trencament de la tolerància

Els mecanismes de tolerància a antígens tissulars presenten diferents nivells de control i quan el balanç entre tolerància i immunitat s'altera en individus susceptibles, es produeixen les malalties autoimmunitàries.

Hi ha diverses circumstàncies que expliquen que es produeixi un trencament de la tolerància:

(i) **Mimetisme molecular**; els limfòcits podrien reconèixer un antígen no propi que fos semblant en estructura o seqüència a un antígen propi de manera que la resposta inicial contra l'antígen del patògen desencadenaria una resposta autoimmunitària per reacció creuada [54]

(ii) **Regulació anòmla de citocines**; citocines alliberades en grans quantitats a la perifèria com a resultat d'una infecció podrien activar cèl·lules T autoreactives

(iii) **Epítops críptics**; epítops que no es generen en condicions normals o que no estan a la vista del sistema immunitari, però que podrien presentar-se en condicions favorables i iniciar una resposta de les cèl·lules T específiques.

En tot cas, hi ha evidències suficients que indiquen que al menys el manteniment de les malalties autoimmunitàries específiques d'òrgan poden ser conseqüència de la presentació inadequada d'autoantígens en el teixit diana [55, 56]. El fet que les cèl·lules T *naïves* convencionals no entrin als teixits [57] va suggerir que el primer contacte entre l'autoantigen i la cèl·lula T 'convencional' havia de tenir lloc en un gangli limfàtic i via cèl·lules presentadores d'antigen professionals (APC). Una resposta desencadenada per lesions locals (infecció vírica, canvi de la dieta,) iniciaria un procés inflamatori inespecífic que permetria que les APCs captessin els autoantígens i es dirigissin al gangli limfàtic. En aquest context el paper de l'òrgan diana seria l'expansió i manteniment de la resposta a través de l'expressió de molècules de MHC-II i la presentació d'antigen a les cèl·lules T i fins i tot la formació d'estructures de fol·licles limfàtics en el propi teixit diana [58].

3. AUTOIMMUNITAT I MALALTIES AUTOIMMUNITÀRIES

Les malalties autoimmunitàries es consideren multifactorials. Existeixen molts factors genètics de predisposició, en la majoria dels casos gens del MHC o gens relacionats. S'ha observat una tendència a l'augment de la incidència d'algunes malalties amb l'edat i sovint són més comuns en les dones que en els homes, indicant la influència de factors hormonals. La incidència de les malalties autoimmunitàries també és diferent entre els diferents grups ètnics i també influeixen factors ambientals com la localització geogràfica o determinats elements de la dieta.

Les respostes autoimmunitàries solen ser processos lents que acaben sent crònics. Durant el desenvolupament d'aquestes respostes es produeix una diversificació de l'especificitat i el ventall de determinants reconeguts per les cèl·lules T autoreactives augmenta i canvia amb el temps. Aquesta ampliació de la resposta s'anomena expansió d'epítops i s'ha vist dins d'un mateix autoantigen o bé cap a altres autoantígens [59].

Les malalties autoimmunitàries es poden classificar en dos grans grups:

1. Malalties específiques d'òrgan, en les quals la resposta autoimmunitària va dirigida contra antígens que s'expressen en un determinat òrgan, teixit o tipus cel·lular. Aquestes malalties s'associen a la producció d'autoanticossos específics i a la presència de cèl·lules T específiques d'antígens de l'òrgan diana;
2. Malalties sistèmiques en les quals l'antigen contra el qual va dirigida la resposta es troba distribuït àmpliament per tot l'organisme. Aquestes malalties s'associen generalment a una elevada producció d'anticossos i a la formació d'immunocomplexes circulants.

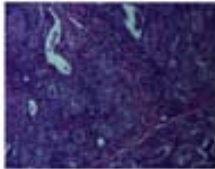
3.1. Malalties autoimmunitàries de la tiroide: Autoimmunitat endocrina

Les cèl·lules endocrines són cèl·lules altament especialitzades dedicades a la producció d'hormones que se secreten al flux sanguini per distribuir-se a diferents parts del cos. Les malalties autoimmunitàries endocrines es donen com a conseqüència del trencament de la tolerància cap a cèl·lules endocrines.

Exemples d'autoimmunitat endocrina serien la diabetis tipus I (T1D) o les malalties autoimmunitàries de la tiroide (AITD).

Les malalties autoimmunitàries de la tiroide (AITD, autoimmune thyroid disease) són malalties autoimmunitàries específiques d'òrgan en les quals la resposta immunològica va dirigida contra antígens propis de la glàndula. Aquestes malalties s'associen a la producció d'autoanticossos específics (anti-tiroglobulina, anti-TPO i anti-receptor de la TSH) i a la presència de cèl·lules T i B infiltrants, reactives contra antígens de la tiroide. Tot i que la presència d'autoanticossos és un tret característic d'aquestes malalties està demostrada la seva presència en la població sana. La prevalença d'autoanticossos de tiroide varia en funció de la ètnia, la regió i de factors ambientals [60-62].

Les AITD afecten a més d'un 5% de la població i representen més del 30% de les malalties autoimmunitàries específiques d'òrgan. Els tipus de malalties autoimmunitàries que afecten més freqüentment a la glàndula tiroide són: la Tiroiditis de Hashimoto (HT) que dona lloc a hipofunció de la glàndula i la malaltia de Graves-Basedow (GB) que causa la hiperfunció de la tiroide (Taula 1).

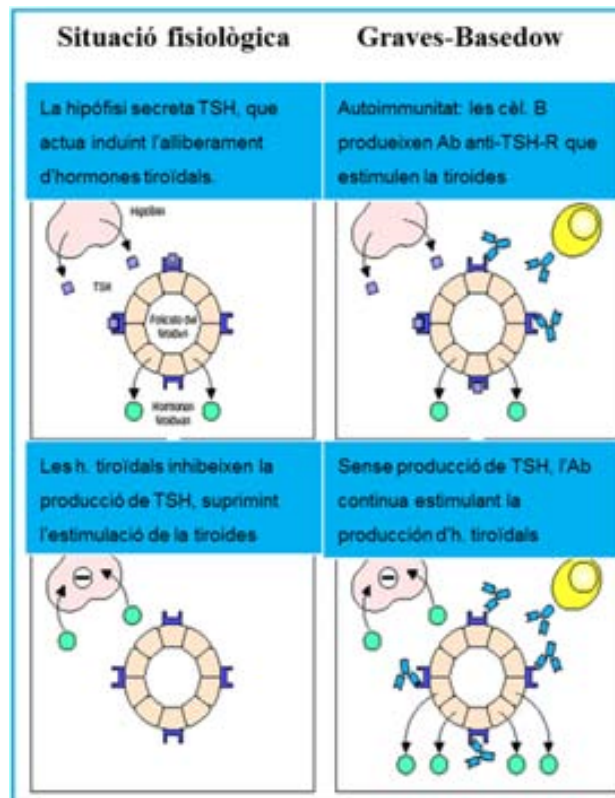
	Graves- Basedow	Tiroiditis de Hashimoto
		
Infiltració limfocitària	+ / ++ Infiltració difusa CD4>CD8 Fol·licles limfoides secundaris	++++ Infiltració massiva CD8>CD4 Fol·licles limfoides secundaris
Autoanticossos: anti-tiroglobulina anti-TPO anti-receptor de TSH-R	+ ++ +++	++ / +++ +++ + / -
Dany al teixit	Hipertròfia i hiperplàsia	Atròfia i fibrosi
Predisposició genètica HLA - classe II	DR3	DR5

Taula 1: Comparació entre la malaltia de Graves-Basedaw i la Tiroidis de Hashimoto, a nivell d'infiltració limfocitària, dany tissular, producció d'autoanticossos i al·lel d'HLA associat.

Tot i que la etiologia d'aquestes malalties no es coneix, existeixen evidències que indiquen una influència important de la genètica en el desenvolupament de les AITDs. Està acceptat que les AITD es desenvolupen en individus genèticament predisposats, en conjunció amb exposicions a determinats factors mediambientals. Durant l'última dècada hi ha hagut grans progressos per entendre la contribució de factors genètics a l'etiologia de les AITDs. S'ha identificat i caracteritzat gens de susceptibilitat a les AITD, alguns específics per GB o HT i d'altres que confereixen susceptibilitat a ambdues malalties. El primer gen de susceptibilitat identificat va ser el locus HLA-DR i posteriorment es van identificar altres gens com CTLA-4, CD40, PTPN22, tiroglobulina i el receptor de la TSH (TSHR) [63]. També s'ha relacionat amb la susceptibilitat a la malaltia de GB polimorfismes d'altres gens codificats a la regió intermèdia de l'MHC, que intervenen en la resposta immunològica, com la proteïna d'estrès tèrmic hsp70 [64] o citocines com $TNF\beta$ [65]

3.1.2. Malaltia de Graves-Basedow (GB)

Hiperfunció de la glàndula tiroide degut a la presència d'anticossos agonistes anti-receptor de la TSH (TSHR) que estimulen les cèl·lules fol·liculars del tiroides, mimetitzant l'efecte de la TSH i provocant una secreció excessiva d'hormones tiroïdals [66, 67] (figura 8). Un percentatge elevat de pacients, aproximadament un 75%, també tenen autoanticossos anti-tiroglobulina i anti-peroxidasa tiroïdal (TPO).



Janeway et al.: Immunobiologia, © Masson, Barcelona, 2000

Figura 8: Durant la malaltia de Graves-Basedow queda interrompuda la regulació per retroalimentació de la producció d'hormona tiroïdal. Normalment les hormones tiroïdals es produeixen en resposta a la TSH i limiten la seva pròpia producció inhibint la producció de TSH per la hipòfisi (dibuix esquerre). En GB els autoanticossos són agonistes del receptor de la TSH i per tant estimulen la producció d'hormones tiroïdals (dibuix dret). Imatge extreta de Janeway, Immunobiologia, Masson, Bcelona 2000.

Les cèl·lules epitelials són les dianes del dany tisular. Aquestes cèl·lules es caracteritzen per presentar (i) un elevat grau de diferenciació, (ii) expressió de molècules específiques del teixit que són els autoantígens més comuns i (iii) una sobreexpressió de classe I i expressió ectòpica de classe II, li i HLA-DM. L'expressió d'ambdues xaperones donava lloc a la presència de complexes MHC-pèptid amb conformació estable a la superfície d'aquestes cèl·lules i la seva expressió augmentava amb l'addició d'IFN γ . En aquest estudi, el nivell de DM detectat en les cèl·lules fol·liculars tiroïdals era menor que el detectat en APCs [68].

Aquesta malaltia es considera la tiroidopatia més freqüent, amb una incidència aproximadament de 4/10000 per any [61, 69]. Afecta principalment dones entre 30 i 40 anys. Es caracteritza per una hiperplàsia de la glàndula de la tiroides

amb hiperfunció i tirotoxicosis. Les manifestacions clíniques de la malaltia són: pèrdua de pes, taquicàrdia, hiperquinesia (falta d'atenció, inquietud), goll difós, i nivells elevats de T4 i/o T3.

Histològicament s'observa una hipertròfia i hiperplàsia de l'epiteli fol·licular, els fol·licles augmenten de mida i les cèl·lules fol·liculars es fan més altes. La llum del fol·licle disminueix i conté menys col·loide. En el pol apical de les cèl·lules fol·liculars s'observen vacuoles de reabsorció. Destaca la presència d'un infiltrat difús format per cèl·lules T ($CD4^+ > CD8^+$), B i cèl·lules plasmàtiques. Poden produir-se també manifestacions extratiroïdals, sent la més freqüent l'oftalmopatia infiltrativa (GO) que afecta entre un 25-50% dels pacients amb GD [70], però també poden haver-hi afectacions cardiovasculars, gastrointestinals, dèrmiques, neuromusculars i reproductives que li poden conferir a la malaltia un caràcter sistèmic. L'etiologia d'aquestes manifestacions extratiroïdals no es coneix gaire bé, se suposa que l'oftalmopatia es desenvolupa com a conseqüència d'una reacció creuada entre els autoantígens tiroïdals i els orbitals. Cada cop hi ha més evidències que el TSHR també juga un paper important: els nivells d'anticossos anti-TSHR correlacionen amb la clínica de GO [71] i amb l'expressió de TSHR en teixits orbitals de pacients amb GB. Aquestes dades fan pensar que les cèl·lules T TSHR-específiques podrien ser reclutades a l'òrbita on s'activarien i secretarien citocines capaces d'induir la proliferació de fibroblasts i la producció de glicosaminoglicans.

S'ha establert l'associació de GB amb l'al·lel HLA-DR3 [72]. La freqüència de DR3 entre els pacients amb GB és del 40-50%. Dins la població caucàsica l'al·lel HLA-DQA1*0501, en desequilibri de lligament amb DR3, també està associat a GB [73, 74]. Altres estudis demostren que la presència d'aquest al·lel per sí sol també confereix susceptibilitat en pacients que no expressen DR3 [73]. Més recentment s'ha demostrat que la presència d'Arg a la posició 74 de la cadena $\beta 1$ (Arg $\beta 74$) del DR s'associa significativament amb la susceptibilitat a GD. Segons els autors, l'associació es deu a la presència de l'Arg $\beta 74$ més que a DR3, amb el qual hi hauria una associació secundària, ja que moltes variants de DR3 contenen Arg $\beta 74$.

Receptor de la tirotropina (TSHR)

Fisiologia. El receptor de la TSH s'expressa a nivells molts baixos a la membrana baso-lateral de les cèl·lules fol·liculars tiroïdals (10^3 - 10^4 molècules per cèl·lula). Pertany a la família de receptors acoblats a proteïna G i pot interaccionar amb diferents membres d'aquesta família. A través d'aquestes combinacions, el TSHR exerceix les diferents funcions de control del creixement i de la funció de la tiroide [75].

Estructura proteica. Aquest receptor és una glicoproteïna de 764 aminoàcids formada per dues subunitats unides per ponts disulfur: la subunitat A que és la regió extracel·lular on hi ha el lloc d'unió de la TSH i la subunitat B formada per set dominis transmembrana i una cua citoplasmàtica [76, 77]. La principal diferència respecte altres receptors d'hormones és que una proporció variable dels TSHR es presenta a la superfície cel·lular amb les subunitats, A i B separades (figura 9). Un cop expressat a la membrana dels tiròcits, la cadena del TSHR es trenca en la subunitat A i B. Aquest trencament produeix l'escissió d'una regió "pèptid C" i a vegades s'allibera la subunitat A. Aquest procés pot generar fragments proteics estimuladors de la resposta autoimmunitària [78].

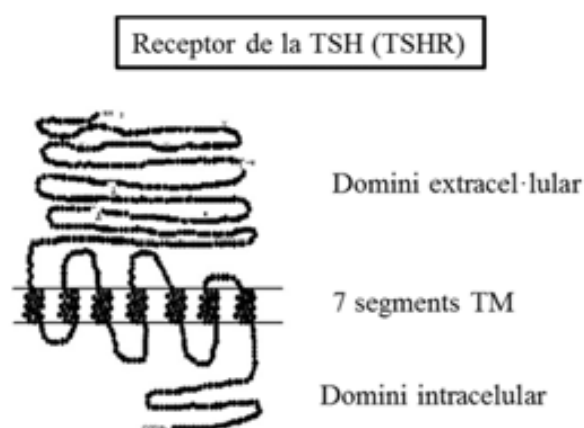


Figura 9: Estructura del receptor de TSHR: Domini extracel·lular (subunitat A), 7 segments transmembrana i un domini intracel·lular (ambdós formen la subunitat B).

Hi ha indicis de que la subunitat A podria estar implicada en la generació d'anticossos estimuladors (TSAb) [79]. La subunitat A lliure és capaç de neutralitzar *in vitro* l'activitat dels TSAb [80]. La immunització de ratolins BALB/c

amb un adenovirus amb la subunitat A indueix hipertoroidisme a 70 % dels animals a diferencia del 10% quan s'utilitzen adenovirus que codifiquen un TSHR mutat per evitar el trencament [81, 82].

OBJECTIUS

Objectiu general: Caracterització dels repertoris de pèptids presentats per molècules de classe II associades a autoimmunitat en humans i en ratolí

- Objectiu específic 1: Presentació de pèptids per l'al·lel HLA-DR17 (DRB1*0301/ DRA1*0101), al·lel preferentment associat a malalties autoimmunitàries

Capítol 1: Repertori peptídic presentat per HLA-DR17 (DR3): revisió del motiu d'ancoratge

Capítol 2: Influència de xaperones HLA-DM i cadena invariant (Ii) en els repertoris presentats per HLA-DR3

Capítol 3: Presentació del TSHR, autoantígen dominant en la malaltia de Graves-Basedow

- Objectiu específic 2: Presentació de pèptids per les molècules de classe II de l'haplotip H-2^d, un al·lel associat a malalties autoimmunitàries tipus Th2

Capítol 4: Descripció del motiu d'ancoratge dels al·lels de classe II d'H-2^d: I-A^d i I-E^d

Capítol 5: Estudi comparatiu entre els repertoris peptídics de timus i melsa en ratolins Balb/c

MATERIALS I MÈTODES

1. LINIEES CEL·LULARS

1.1. Cultius de cèl·lules en suspensió

La línia **CDG** es va generar al nostre laboratori a partir d'un donant infantil de diabetis tipus 1 que era homozigot per a HLA-DR3, i va ser transformada amb el virus EB. Les cèl·lules es van créixer en suspensió amb medi RPMI (Sigma-Aldrich) suplementat amb 2 mM L-glutamina, 0.1 mg/ml de estreptomicina, 100 U/ml penicil·lina i 10% de sèrum boví fetal (FBS) (Gibco) prèviament descomplementat a 56°C durant 30 minuts (anomenat medi RPMI complet a partir d'ara). Es van usar flascons de poliestirè T175 en posició vertical. Les cèl·lules es recollien per centrifugació a 500xg durant 10 minuts i després de dos rentats amb PBS, els botons cel·lulars (*pellets* secs) es submergien directament en N₂ líquid i s'emmagatzemaven a -80°C fins a ser processats.

1.2. Cultius de cèl·lules adherents

La línies cel·lulars **RINm5F**, generades a partir d'un insulinoma de rata [83] es van cultivar en monocapa, amb medi de cultiu RPMI (Sigma-Aldrich) complet. Es van utilitzar quatre línies diferents RINm5F [84]: RINm5F-DRB1*0301 (DR3), RINm5F-DRB1*0301-Cadena invariant (DR3li), RINm5F-DRB1*0301-HLA-DM (DR3DM) i RINm5F-DRB1*0301-Cadena invariant- HLA-DM (DR3liDM) .

La línia cel·lular **HT93**, derivada de cèl·lules fol·liculars humanes del tiroides, va ser obtinguda per transformació amb el virus SV40 (Belfiore A et al 1991). Les diferents línies HT93 transfectades així com les HT93 van ser cultivades en medi DMEM (Sigma-Aldrich) suplementat amb 10% de FBS .

Les cèl·lules, que creixien en monocapa, es van cultivar en flascons de poliestirè de T175 en posició horitzontal i es diluïen cada 2-3 dies mitjançant un procés de tripsinització (0.025% de tripsina, (Sigma, St Louis, MO, USA) i dos rentats amb PBS. Quan estaven al voltant de un 70-80% de confluència, les

cèl·lules es centrifugaven, es congelaven els botons cel·lulars en N₂ líquid i s'emmagatzemaven a -80°C fins al moment de ser usats.

Totes les línies cel·lulars es mantingueren en cultiu en un incubador a 37°C i en una atmosfera del 5% de CO₂. Es van créixer 1×10^7 cèl·lules de cada línia cel·lular per cada experiment d'aïllament i identificació de pèptids. Tots els experiments es van fer per duplicat. Com a control de l'expressió de HLA-DR, a les diferents línies se'n mesurava l'expressió per citometria amb un anticòs anti-HLA-DR (veure més endavant) abans de ser congelades i emmagatzemades en *pellets* secs.

2. TEIXIT DE RATOLÍ BALB/C

2.1. Mostres de timus i melsa de ratolí balb/c estàndards

Els ratolins Balb/c van ser proporcionats pel laboratory Charles River S.A. a través del Servei d'Estabulari de la UAB, on eren sacrificats. Seguidament els timus i melses eren extrets, rentats amb PBS, submergits en N₂ líquid i emmagatzemats a -80 °C. Es va emmagatzemar 40 timus i 40 melses.

2.2. Mostres de timus i melsa de ratolí Balb/c transgènic

Per l'obtenció de timus i melses de ratolins Balb/c transgènics per al TSHR humà, vam contar amb la col·laboració dels Drs. S. McLaglan i B. Rapoport de l'Institut de Recerca del Cedars-Sinai de Los Angeles, (USA), que ens van enviar teixit congelat provinent d'aquests animals transgènics que ells havien generat al seu laboratori [85].

3. TRANSFECCIONS

3.1. Transfecció de la línia HT93

Tots els transfectants van ser generats per electroporació i amb 60 µl de DNA.

HT93-CIITA

Es va partir d'aproximadament 6×10^6 de cèl·lules HT93 en cultiu. Un cop tripsinitzades i rentades amb PBS, es van resuspendre en 17 µl d'una dilució que contenia 3.4µg/µl del vector pcDNAI (neo^R)-CIITA amb PBS (cedit pel Dr. Antonio Celada, Universitat de Barcelona, Barcelona) [86], purificat i linealitzat amb BglII. La mescla es va electroporar a 250 Volts, 950 microfarads i 24 ohms en un electroporador *Electro cell manipulator* 600 (BTX, USA). Després de 10 minuts amb gel, les cèl·lules es va resuspendre en 15 ml de medi DMEM suplementat amb 10% de FBS i L-glutamina. Es van transferir a flascs de 75 cm² i es van cultivar a 37°C, 5% CO₂ durant 48 hores. Aquestes cèl·lules es van mantenir en cultiu amb canvis de medi setmanals, fins arribar a una concentració suficient per criopreservar 3 alíquotes de 2×10^6 cèl·lules. El resultat de la transfecció es va comprovar analitzant l'expressió superficial de HLA-DR induïda per l'expressió de CIITA per citometria de flux, en un citòmetre FACS Calibur (Becton Dickinson, USA), al Servei de Citometria de l'IBB.

HT93-TSHR

Es va seguir el mateix procediment que l'usat per l'obtenció de la línia cel·lular HT93-CIITA. En aquest cas, el vector recombinant utilitzat va ser el pcDNA3.1-TSHR-LAMP, cedit pels Drs. S.McLachlan i B.Rapoport (Pichurin et al, 2006). Aquest vector conté un senyal de transport a lisosomes (LAMP) que afavoreix l'acumulació del TSHR en la via endocítica i per tant la generació de pèptids accessibles a HLA-DR. A més, el plàsmid de TSHR conté un gen de resistència a la geneticina per a la selecció de transfectants.

Les cèl·lules HT93 es van resuspendre amb 75 µl (0.8µg/µl) del vector i es va seguir el protocol d'electroporació en les mateixes condicions mencionades abans. Després de 48 hores de cultiu, es va eliminar el sobrenedant i es va afegir medi de selecció que contenia 1mg/ml de geneticina (G418, Invitrogen USA). Es van deixar en cultiu durant dues setmanes fins que es va observar la presència de cèl·lules que havien sobreviscut i crescut en presència del medi

de selecció. Els transfectants generats es van analitzar mesurant l'expressió intracel·lular del TSHR per citometria de flux.

Dobles transfectants HT93-CIITA-TSHR (HT93 DT)

Per a l'obtenció del doble transfectant CIITA⁺ TSHR⁺ es va seguir el mateix protocol explicat per a les transfeccions simples, però afegint una barreja dels DNA dels dos plàsmids. La relació CIITA:TSHR va ser 10:1 i les quantitats de cada plàsmid que es va afegir a la barreja van ser 17 µl (3.4µg/µl) del vector pcDNA1(neo^R)-CIITA i 1.25 µl (0.8µg/µl) del pcDNA3.1-TSHR-LAMP. Aquesta relació es basava en que només el plàsmid de TSHR conté un gen de selecció. Un cop les cèl·lules es mantenien en cultiu i no perdien expressió, es van separar les cèl·lules DR⁺ per separació magnètica i les resultants que expressaven més eTSHR foren clonades per dilució límit. Dels clons obtinguts, es van seleccionar el clon 2.1, que tenia l'expressió més alta d'HLA-DR i eTSHR. Aquest clon va ser l'escollit per a la comparació amb els altres transfectants. El resultat de la transfecció es va analitzar per citometria de flux, mitjançant una tinció superficial i una tinció intracel·lular.

4. CITOMETRIA DE FLUX

El fenotipat de les línies cel·lulars es va realitzar mitjançant immunofluorescència directa per citometria de flux de cèl·lules en suspensió. Per a l'anàlisi es va utilitzar un citòmetre de flux FACscan (Becton-Dickinson, San Jose, USA). Depenent de la molècula a caracteritzar es va realitzar un marcatge de superfície (CLIP i HLA-DR) o un marcatge citoplasmàtic (TSHR, li, HLA-DM).

4.1. Marcatge superficial

Aproximadament 2×10^5 cèl·lules s'utilitzaven per cada marcatge. Es feien dos rentats amb PBS-2% FCS y es centrifugaven 5 minuts a 500 x g. Les cèl·lules s'incubaren amb 50 µl de sobrenedant de l'anticòs primari (molècula a

analitzar) durant 30 minuts a 4°C .Els botons cel·lulars van ser resuspesos amb 50 µl d'una dilució 1/200 de l'anticòs secundari (FITC goat anti-mouse Ig (H+L), Invitrogen) i es va incubar a 4°C durant 20 minuts en foscor. Seguidament , es van fer dos rentats amb PBS, es van resuspendre en un volum final de 400 µl de PBS i es va analitzar per citometria de flux. Els anticossos primaris utilitzats van ser CerCLIP (CLIP, Dr. P. Cresswell) [87] i B8.11.2 (Dr F. Koning Leiden University, Països Baixos).

4.2. Marcatge intracel·lular

Es repartien en dos tubs 2×10^5 cèl·lules (mostra i control positiu) i es feien dos rentats amb PBS. Les cèl·lules es fixaven en una solució de formaldehid al 3,7% durant 20 min. a temperatura ambient, es centrifugen i es descarta el sobrenedant. Es realitzaren dos rentats amb PBS i es van resuspendre els botons cel·lulars amb l'anticòs primari i tritó X-100 al 0,5% (Sigma). Per al control negatiu es va usar 50 µl de PBS amb 2% de FBS i 0,5% tritó X-100 i es va deixar incubar 45 minuts a temperatura ambient.

Les cèl·lules es van rentar dos cops amb PBS i es va incubar seguidament amb l'anticòs secundari (Alexa Fluor® 488 donkey anti-mouse IgG (H+L), Invitrogen) durant 45 minuts a temperatura ambient i sense llum. Passat el temps estimat, es van fer dos rentats amb PBS i es van resuspendre les cèl·lules en un volum final de 400 µl de PBS. Seguidament, es va analitzar per citometria de flux.

Els anticossos primaris utilitzats van ser: 50 µl d'anticòs (IgG anti-human TSHR de ratolí, AbD Seritec) diluït 1/200 en PBS 2% de FBS; VIC-Y1 (anti-Ii, Dr. W. Knapp, IFI, Viena, Austria) [88] i 5C1 (HLA-DMα).

5. ELUCIÓ DELS LLIGANDS NATURALS D'MHC DE CLASSE II

L'aïllament de pèptids a partir de botons cel·lulars així com de teixit es va fer mitjançant cromatografia d'afinitat, elució àcida i ultrafiltració. Posteriorment les fraccions peptídiques es fraccionaven per cromatografia líquida de intercanvi iònic i a continuació s'analitzaven per espectrometria de masses en el laboratori de Proteòmica CSIC/UAB, (IIBB-CSIC, IDIBAPS, Facultat de Medicina, UAB).El protocol que es va utilitzar per la purificació i fraccionament

dels pèptids va ser semblant al que es feia servir en el nostre laboratori per les línies en cultiu [89, 90] però amb algunes modificacions (figura M1). El protocol es va optimitzar amb l'objectiu d'aconseguir minimitzar les interferències degudes a l'ús del detergent Nonidet P40 (Surfact-Amps NP40) i augmentar així el nombre de seqüències identificades. El NP40 ha estat el detergent més utilitzat clàssicament per a la solubilització de les molècules de MHC i per a la seva posterior purificació. No obstant això, aquest detergent era difícil d'eliminar, tot i realitzar diverses etapes de purificació i rentats, i interferia amb l'anàlisi espectromètric. Per aquest motiu, es va decidir utilitzar com a substitut el detergent CHAPS, un detergent zwitteriònic més fàcil d'eliminar de les mostres. A més, es va eliminar el rentat de la columna amb el buffer de lisi, deixant els rentats lliures de CHAPS abans de l'elució (veure més endavant). A més de les millores en el processament de la mostra, en aquest treball s'ha utilitzat un espectòmetre de masses d'alta resolució (LTQ-Orbitrap) que ens ha permès augmentar la sensibilitat de les anàlisis, disminuir el nombre de falsos positius i incrementar notablement la velocitat de processament de les dades.

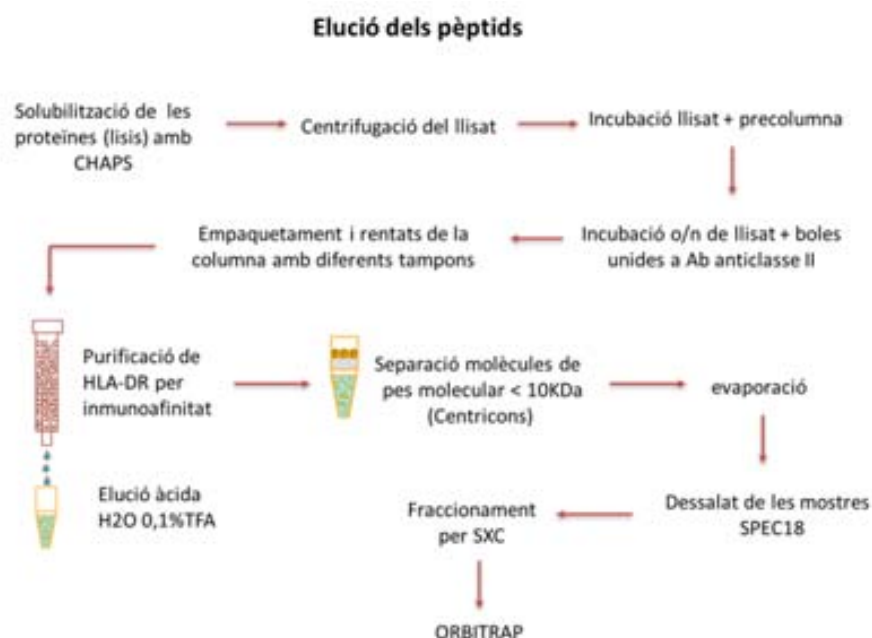


Figura M1: Protocol de purificació i identificació de pèptids associats a MHC.

5.1. Obtenció d'anticossos anti-HLA-DR

5.1.1. Producció i purificació de l'anticòs B8.11.2

Per purificar les molècules de HLA-DR de les línies cel·lulars, es va utilitzar l'anticòs B8.11.2. Aquest anticòs s'obté a partir del corresponent hibridoma que va ser cedit per Dr F. Koning (Leiden University, Països Baixos). El hibridoma B8.11.2 produeix un anticòs monoclonal IgG2b específic per DR β , que reconeix un epítot monomòrfic de HLA-DR, comú per a tots els al·lels d'HLA-DR. Aquest anticòs no reconeix altres isotips d'HLA de classe II.

El cultiu de l'hibridoma B8.11.2 es va realitzar en flasks de 175 utilitzant medi RPMI complet. Les cèl·lules es creixien fins a obtenir un volum de 500 ml per flascó i quan aproximadament el 80% de les cèl·lules havien mort es centrifugaven a 600xg durant 6 minuts per eliminar els residus cel·lulars i les cèl·lules restants. A continuació el sobrenedant es filtrava amb el sistema Sterifil Aseptic System (Millipore, USA) usant filtres de 0,45 μ m de diàmetre de porus i es guardava a -20°C fins el moment de la purificació.

Per la purificació de l'anticòs es van utilitzar columnes de proteïna A sefarosa (Protein A SepharoseTM 4 Fast Flow, Amersham Biosciences). El sobrenedant de B8.11.2 es diluïa amb la mateixa quantitat de tampó d'acoblament (tampó fosfat 0.1M pH 7.0). La columna es rentava prèviament també amb 50 ml de tampó d'acoblament i llavors es feia passar dos cops el sobrenedant per la columna a 4°C, per retenir així la major part d'anticòs possible. Un cop passat el sobrenedant es feia un rentat amb 30ml de tampó d'acoblament i després s'eluïa l'anticòs amb 3-5 volums de columna d'àcid cítric 0.1M pH 3.0. Es recollien al·lquotes d'1ml controlant el moment en el que el pH de l'eluït de la columna virava a àcid. Aquestes fraccions es neutralitzaven amb Tris-HCl 1.5M pH 8. Per regenerar la columna es feia un rentat amb tampó fosfat 0.1M pH 7.0 i es guardava amb tampó d'acoblament amb 0.05% d'azida sòdica per evitar contaminacions o amb etanol al 20%.

La presència de proteïna a les fraccions es va determinar mitjançant la tècnica de Bradford (explicada en el següent apartat). Totes les fraccions que contenen proteïna es van ajuntar en una única mostra. La mostra es dialitza en una membrana de diàlisi (Medicell Internacional) enfront de bicarbonat sòdic 0,1M, NaCl 0,5M, pH 8,3. Es van realitzar quatre canvis de 500 ml de tampó cada quatre hores. El material obtingut es va al·luitar i es va congelar a -20°C fins al seu ús.

5.1.2 Determinació de la concentració d'anticòs

La concentració de l'anticòs es va determinar mitjançant la tècnica de Bradford. Per a això es va usar el reactiu Bio-Rad protein assay (Bio-Rad) i es van seguir les instruccions del proveïdor. L'absorbància es va mesurar a 620 nm. Per determinar la concentració, les lectures obtingudes es van extrapolar davant d'una recta patró de concentracions conegudes d'albumina sèrica bovina (BSA).

5.1.3 Determinació de l'activitat de l'anticòs B8.11.2

Es va realitzar la titulació de l'anticòs B8.11.2 per citometria de flux, seguint els mateixos passos explicats en la secció 4.1 per a les tincions superficials. En aquest cas, es van realitzar incubacions de cèl·lules limfoblastoides d'elevada expressió d'HLA-DR davant dilucions de l'anticòs purificat que anaven des 1/50 fins 1/3200 (figura M2)

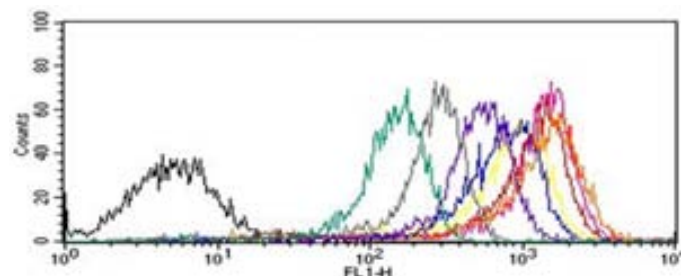


Figura M2: Histograma d'una titulació de l'anticòs B8.11.2. Negre: marcatge amb un anticòs secundari en absència de B811.2; Groc: dilució 1/25; Rosa: dilució 1/50; Roig: dilució 1/100; Taronja: dilució 1/200; Blau: dilució 1/400; Morat: dilució 1/800; Gris: dilució 1/1600; Verd: dilució 1/3200.

5.1.4 Determinació de la puresa de l'anticòs per SDS-PAGE

Es van realitzar gels de SDS-PAGE, segons el mètode de Laemmli [91], per determinar la puresa de l'anticòs purificat. Es van agafar 5-10 μ l de l'anticòs i es van barrejar amb el mateix volum de tampó de càrrega (Tris 0,125 M, glicerol 10%, SDS 4%, β -mercaptoetanol 10%, blau de bromofenol 0,002%). La barreja es va bullir durant 5 minuts. Les proteïnes es van separar en un gel de poliàcrilamida al 12%, a un voltatge constant de 200V fins que el front d'electroforesi va arribar al final del gel. El gel es va tenyir amb blau de Coomassie (metanol 20%, 10% àcid acètic, 0,1% Blau Coomassie) durant 30 min a 1h. Es va rentar el gel amb aigua destil·lada dues vegades i es va destenyir usant una solució decolorant (20% Metanol, 10% àcid acètic) durant mínim 2 hores. Un exemple del resultat, on es visualitzen les cadenes que formen l'anticòs es mostra a la figura M3.

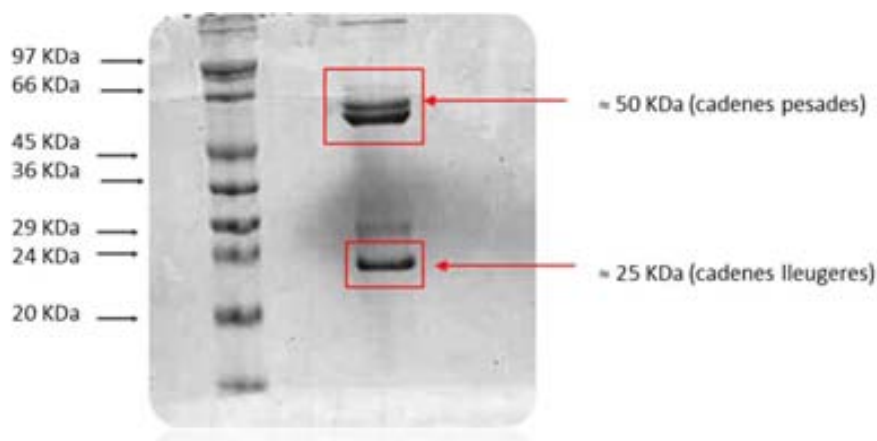


Figura M3: Gels de SDS-PAGE que ens mostra les cadenes pesades i lleugeres de l'anticòs B811.2

Un resum dels passos seguits per l'obtenció d'anticòs purificat es mostra a la figura M4.

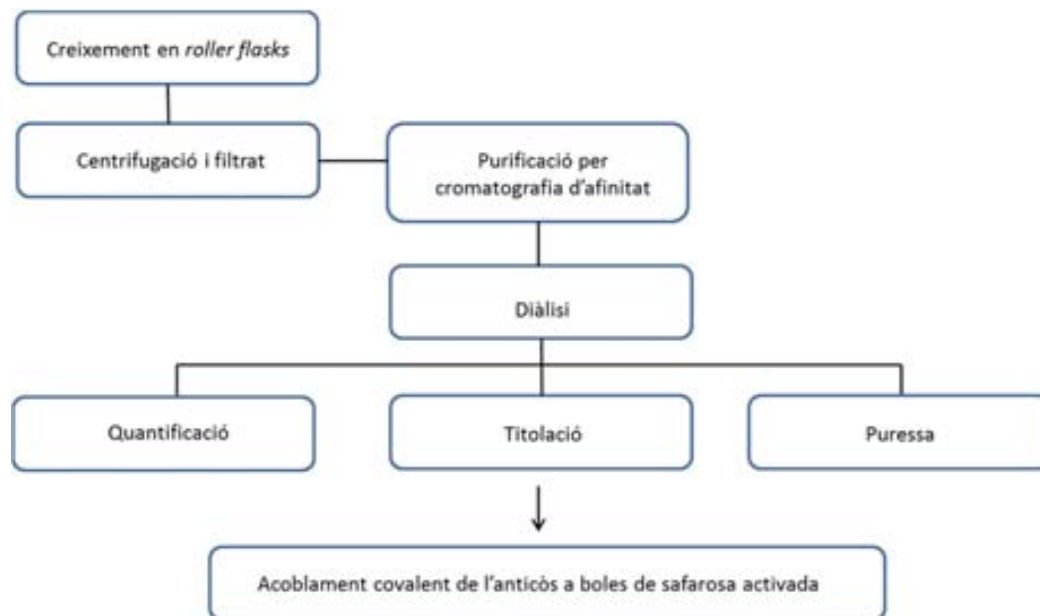


Figura M4: Esquema dels passos seguits per l'obtenció d'anticòs HLA-DR, des del creixement de l'hibridoma fins l'acoblament a boles de safarosa activades, per l'elució de pèptids.

5.2. Obtenció dels anticossos anti-IA^d i anti-IE^d

Els anticossos de classe II monoclonals de ratolí específics per Balbc/c que es van utilitzar van ser l'anticòs MK-D6 (anti-IA^d) i l'anticòs 14-4-4S (anti-IE^d). Ambdós van ser proporcionats pels doctors S. McLaghlan i B. Rapoport, del Autoimmune Disease Unit, de l'institut de Recerca Cedars-Sinai de Los Angeles (USA)

5.3. Preparació de les columnes per la cromatografia d'afinitat

Per preparar les columnes per la cromatografia d'afinitat es va utilitzar safarosa CL-4B activada per CNBr (CNBr-activated sepharoseTM 4B, Amersham Biosciences). Es van fer servir columnes d'1 ml amb 5 mg/ml d'anticòs de B811.2 i columnes de 1 ml amb 2 mg/ml d'anticòs MK-D6 i 14-4-4S. El mètode que es va seguir per acoblar l'anticòs a les boles de safarosa va ser el mateix per tots els anticossos. En tots els casos es va preparar una columna amb safarosa unida amb anticòs i una altra columna, que anomenavem precolumna, que només portava safarosa (sense anticòs) i que es feia servir prèviament a la d'afinitat.

Per activar la safarosa i generar els èsters de cianogen reactius, les boles (aproximadament 0.7 g/mL columna) van ser resuspeses en 50 ml d'HCl 1mM, agitant per evitar la formació d'agregats de sefarosa i deixant en rotació durant 30 minuts (Affinity Chromatography: Principles and Applications 2013).

Seguidament, les boles es netejaven en un filtre de vidre (Glassfilter 3, Duran R Schott, Mainz. Alemanya) amb 200 ml de HCl 1 mM durant uns 20 minuts, es resuspenien amb tampó d'acoblament (NaHCO_3 0,1 M, NaCl 0,5 M, pH 8,3) i es centrifugaven 2 minuts a 500xg. A continuació, s'enrasava amb 20 ml de tampó d'acoblament i el volum resultant se repartia en dos tubs. Es van centrifugar durant 2 minuts a 500xg i es va eliminar el sobrenedant. Un dels tubs amb safarosa es resuspenia amb la solució d'anticòs monoclonal B.8.11.2 o d'anticòs MK-D6 o 14-4-4S per l'acoblament, la resta de la sefarosa es resuspenia amb igual volum de solució tampó d'acoblament- per generar la precolumna. Les mescles es van incubar durant 4 hores a temperatura ambient en rotació. Seguidament es van centrifugar les boles i es va recollir els sobrenedats per determinar la seva concentració proteica, mitjançant la lectura de l'absorbància a 280nm, com a mesura de comprovació de la unió de l'anticòs a les boles de safarosa.

Tant la precolumna com la columna amb l'anticòs acoblat es van resuspendre en tampó de bloqueig (Tris-HCl 0,1M pH 8) per inactivar els radicals lliures que puguin unir-se de forma inespecífica a altres proteïnes i es van centrifugar 2 minuts a 500xg. Aquest procés de rentat es va repetir tres cops. Per bloquejar els radicals actius de la safarosa amb capacitat d'unió a proteïna, es van incubar durant 2 hores a temperatura ambient i en rotació amb el tampó de bloqueig. Es van rentar les safaroses en el filtre de vidre alternant 30 ml de rentat pH 4 (NaAC 0,1M, NaCl 0,5M, pH 4) i tot seguit, sense deixar-la assecar, amb el tampó de rentat pH8 (Tris-HCl 0,1M, NaCl 0,5M). Aquest procés de rentat es va repetir tres vegades. Finalment, tant les columnes de safarosa acoblada a l'anticòs per a la immunopurificació com la safarosa de precolumna es van resuspendre amb tampó de rentat pH 8 amb 0,02% d'azida sòdica i es van guardar a 4 ° per al seu ús posterior.

5.4. Immunopurificació de pèptids a partir de pellets secs provinents de línies cel·lulars

Els *pellets* cel·lulars emmagatzemats a -80°C es dissolien amb 50 ml de buffer d'homogenització consistent en 50 mM Tris HCl (pH 7.6), 150 mM NaCl, 1% CHAPS (Sigma-Aldrich) i un *cocktail* d'inhibidors de proteases (Complete Protease Inhibitor Cocktail, Roche Diagnostics, Basel, Switzerland) i es deixava llsant durant 2 hores a 4°C amb agitació. A continuació, es centrifugava 10 min a 2500 rpm i a 4°C per eliminar les restes cel·lulars i s'ultracentrifugava (Centrifuga Beckman, model L8 60M, rotor 60Ti) a 100.000g durant 1 hora i a 4°C.

El extracte proteic s'incubava primerament amb les boles de sefarosa bloquejades amb Tris (precolumna) a una relació de 1 ml de lisat / 5 ml de suspensió durant 3 h a 4°C en rotació, per descartar les proteïnes solubles que poden unir-se de manera no específica a la safarosa. Un cop transcorregut el temps d'incubació, s'eliminaven les boles de safarosa per centrifugació, i el llisat s'incubava durant tota la nit a 4°C en rotació amb les boles de safarosa unides a l'anticòs anti-MHC. L'endemà el llisat, que contenia els complexos pèptid-MHC-safarosa, es dipositava per gravetat dins una columna de polipropilè que retenia les boles i deixava passar el sobrenedant. El sobrenedant resultant es tornava a passar un segon cop a través de les boles i finalment es congelava (*flow throw*).

Per disminuir les interaccions no específiques i la concentració de sal i eliminar el detergent restant, es realitzaven dues sèries de rentats a les boles per a cada mostra. Primer un rentat amb 250 ml de 50 mM Tris HCl (pH 7,6) i NaCl 150 mM, seguit d'un segon rentat amb 200 ml de Tris-HCl 5 mM pH 7,6.

5.5. Immunopurificació de pèptids a partir de teixit congelat de ratolí Balb/c

Les mostres de timus i les de melsa es van processar en paral·lel i al mateix temps.

Les mostres, que es trobaven emmagatzemades a -80°C , es llisaven mecànicament amb 50 ml de buffer d'homogenització. Es van seguir els mateixos passos abans explicats fins la obtenció del sobrenedant, després de la incubació amb l'anticòs 14-4-4S. Aquest sobrenedant, es va utilitzar per fer una segona immunopurificació, incubant-lo durant tota la nit amb safarosa conjugada amb l'anticòs Mkd6, el qual es tornà a dipositar en una nova columna. Seguidament, es feien els mateixos passos explicats en l'apartat anterior.

5.6. Elució dels complexos de MHC de classe II-pèptid

En ambdós tipus de mostres (teixits i botons cel·lulars) el procés d'elució va ser el mateix: una elució àcida que separava els complexos MHC-pèptid. Per cada mostra van obtenir dos eluits (anti-IA^d i anti-IE^d).

Els complexos MHC-pèptid es van eluir amb 8 ml de 0,1% àcid trifluoroacètic (TFA) i es van recollir en tubs Eppendorf (Protein LoBind Tube 5 ml, Eppendorf). Aquestes fraccions contenien els pèptids i les cadenes α i β l'HLA-DR separades. Es va mesurar el pH de les fraccions obtingudes, seleccionant per a l'aïllament dels pèptids units a MHCII totes les fraccions àcides i dos amb caràcter més bàsic: l'última abans de detectar el canvi de pH a causa de l'arribada del TFA i la primera del tampó de regeneració de la columna.

5.6.1. Separació dels pèptids de les cadenes α i β per ultrafiltració

Les fraccions es van ultrafiltrar utilitzant filtres de diàmetre de porus de 10 KDa (Centricon 10.000 MWCO, Millipore Laboratories, Bedford, MA, USA) centrifugant a 5000g a 4°C el temps necessari per passar tot el volum eluït (Centrifuga Beckman J20-XP, rotoR JA 25.50). Tot el material filtrat es va concentrar per centrifugació al buit en un Speed Vac (Thermo Fisher Scientific, Waltham, MA). A mesura que es concentraven la mostra en els diferents eppendorfs, el contingut s'ajuntava, acumulant tot l'eluït en un únic tub.

Els tubs buits es van rentar amb 200 µl d'1% TFA i es van afegir al tub final per intentar perdre la mínima quantitat possible de mostra. El volum final de mostra va quedar reduït concentrat fins a aproximadament 200 µl.

5.6.2. Gels de proteïnes dels retinuts del centricó

L'eficiència de la immunoprecipitació i elució de la mostra es va verificar mitjançant la visualització en un gel de SDS-PAGE de les cadenes α i β de l'MHC retingudes durant la ultrafiltració (figura M5).

Els gels es feien al 12% de poliàcrilamida i després es tenyien amb blau de *coomassie* per visualitzar les bandes.

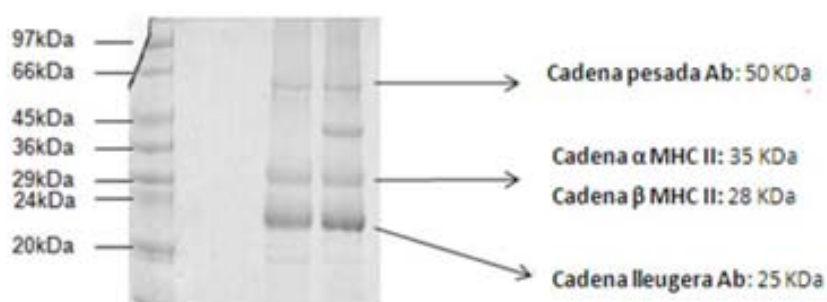


Figura M5: Gel amb el retinut de dos eluïts diferents. A la dreta, es mostren les bandes corresponents a marcador de pes molecular. S'indica amb les fletxes corresponents on es situen les cadenes de MHC.

5.7. Dessalat de les mostres per C18

Les mostres concentrades es dilueixen en 1'3 ml d'1% de TFA i es dessalaven. Pel dessalat es va usar cartutxos d'extracció en fase sòlida C18 (3 ml, 15 mg suport, Varian, Lake Forest, USA). Es va equilibrar la columna (cartutxo C18) amb 9 ml d'H₂O/Acetonitril 1:1 i 3 ml de 0.1% TFA. Es va carregar la mostra a la columna dos cops seguits. Els rentats es van realitzar amb 2 ml de 0.1% TFA i 9 ml d'aigua MilliQ. La elució es va realitzar en dos passos, afegint cada vegada 500 µl de 0.25% TFA en ACN/H₂O milliQ 1:1. El pas dels tampons d'equilibrat, la mostra i l'eluït per la columna es va realitzar a baixa velocitat per facilitar la interacció del pèptid amb la matriu del C18 i per tal de reduir la

pèrdua de lligands durant el procés. Les mostres dessalades es van concentrar per centrifugació al buit fins a un volum aproximat de 10 µl.

5.8. Fraccionament de les mostres per cromatografia líquida d'intercanvi catiònic

Les mostres a separar per cromatografia líquida d'intercanvi catiònic (SCX) es van resuspendre en 500 µl de dissolvent A (0.1% àcid fòrmic, 30% acetonitril). Es va usar una columna Polysulfoethyl (PolyLC inc, Columbia, USA) de 200 x 2.1 mm amb un loop d'injecció de 500 µl per al fraccionament.. Es va treballar a 200 µl /min usant un gradient lineal de 0 a 70% del solvent B (àcid fòrmic al 0,1%, 30% d'acetonitril NH₄Cl 500 mM), durant 35 min i després a 100% durant 12 min. L'anàlisi cromatogràfica es va monitoritzar a 220 nm (UV) de longitud d'ona. Es van recol·lectar fraccions de quatre minuts entre els 4 i 49 minuts. Les fraccions es van evaporar, es va concentrar fins a 5 µl i després es va diluir a 20 µl amb àcid fòrmic 1%, metanol 5%. Només les fraccions compreses entre els minuts 12 i 44 de la SCX es van analitzar a l'espectrometre de masses. La resta es va descartar després de comprovar que el nombre de pèptids a aquestes fraccions era molt baix, disminuint així el nombre de fraccions a analitzar per MS i per tant augmentat la velocitat d'obtenció de les dades (figura M6).

Temps	%B	Flux
0	3	200 µl/min
5	3	200 µl/min
35	70	200 µl/min
40	70	200 µl/min
42	100	200 µl/min
47	100	200 µl/min
49	3	200 µl/min

Figura M6: Gradient de fase B aplicat en el fraccionament de les mostres per SCX en funció del temps a un flux constant de 200 µl per minut .

6. IDENTIFICACIÓ DELS REPERTORIS DE PÈPTIDS ASSOCIATS A MHC MITJANÇANT ESPECTROMETRIA DE MASSES

Es va realitzar l'anàlisi d'espectrometria de masses (MS) usant un espectròmetre de masses d'alta resolució model LTQ-Orbitrap XL, (Thermo Fisher Scientific, San Jose, CA, EUA) equipat amb una interfase de microelectrosprai de Protana (Protana A/S, Dinamarca). Les anàlisis es van fer al laboratori de Proteòmica CSIC/UAB. El sistema d'HPLC es componia d'una bomba de càrrega, una bomba binària de nanoflux, un microinjector de temperatura controlada i una vàlvula de commutació. La separació es va dur a terme a 0,4 l/min usant un gradient lineal d'acetonitril de 3 a 40% en 120 min (dissolvent A: àcid fòrmic al 0,1% B: acetonitril-0,1% d'àcid fòrmic). L'instrument operava sempre en mode positiu amb un voltatge d'esprai de 2 kV. El rang d'escàner per al full MS va ser m/z 400-2000. L'anàlisi d'MS es va realitzar en mode *data dependent*, realitzant un escaneig de l'escombrat complet seguit per 8 espectres de fragmentació (MS/MS) per a la seqüenciació dels senyals més abundants. La resolució de l'Orbitrap es va establir a 60000 full MS. Per minimitzar la selecció redundant d'ions precursors, l'exclusió dinàmica es va fixar a 1 amb una finestra de temps de 5 minuts.

7. IDENTIFICACIÓ I VALIDACIÓ DELS ESPECTRES DE FRAGMENTACIÓ

A partir dels espectres obtinguts per MS/MS es van identificar les seqüències enfrontant-les a una base de dades de proteïnes humanes, ratolí o rata (depenent de l'origen de la mostra). Totes les seqüències es van validar mitjançant un sistema automàtic i una avaluació final manual.

7.1. Cerca base de dades i evaluació automàtica

Els espectres de fragmentació MS/MS es van identificar utilitzant Protein Discover v1.4, (ThermoFisher, San Jose, CA, USA) contra una base de dades combinada Target/Decoy. Aquesta base de dades consisteix en la base de seqüències de proteïnes humanes o de ratolí (depenent de la mostra) d'UniProt (Swissprot 2013_07) a la qual s'han afegit les seqüències reverses de la

mateixa base de dades. Aquesta estratègia *target/decoy* es basa en el principi de que les identificacions incorrectes tenen la mateixa probabilitat de trobar-se a la base de dades *target* com a la *decoy* [92, 93] mentre que els espectres assignats correctament donaran valors elevats amb la seva seqüència corresponent de la base de dades *target* però no amb la base de dades *decoy*.

Els paràmetres de cerca van ser els següents: tolerància de massa pèptid: 20 ppm; tolerància fragment: 0,8 Da; cap restricció de l'enzim; modificacions dinàmiques, l'oxidació de metionina (16 Da). Les identificacions de seqüències peptídiques correctes van ser evaluades usant el paràmetre Percolator [94] amb un FDR menor del 0,1% en el cas de les mostres RIN (capítol 2) i timus/melses de ratolí (capítol 4 i 5), i amb un FDR de 1% per a les mostres CDG (capítol 1) i per les transfectants HT93 (capítol 3).

7.2. Avaluació manual

Com a sistema de control, tots els espectres únics i, almenys un dels espectres que formaven part de cada *nested set*, es van avaluar manualment per confirmar la correcta assignació dels fragments iònics. En l'avaluació manual es va comprovar la qualitat de l'espectre i l'assignació dels ions per part del programa informàtic. A més a més, es van eliminar:

- Totes les seqüències peptídiques derivades de les mescles de pèptids utilitzades per a l'avaluació del sistema de RP-HPLC (calibradors).
- Pèptids amb seqüències inferior a 9 aminoàcids, per no considerar-se dins del rang acceptable d'unió a MHC de classe II.
- Contaminants de la pell: es van detectar un gran nombre de pèptids de queratines procedents de l'epidermis, probablement derivades de la manipulació de les mostres.

8. Anàlisi predictiva d'afinitat

A partir de les seqüències obtingudes per espectrometria de masses, es van generar tots els possibles nonàmers que presenten un residu vàlid en P1 utilitzant el programa *LaiaMotifs* (<https://code.google.com/p/mhc-laia-motifs/>).

Els possible *cores* generats es van avaluar un per un, segons la seva homologia amb els motius d'unió descrits en SYFPEITHY (www.syfpeithi.de/), a excepció de HLA-DR3 i per l'al·lel H-2^d (I-A^d, I-E^d) que vam utilitzar el motius que havíem descrit al nostre laboratori (veure resultats capítol 1 i 4). Utilitzant aquesta aproximació, i degut a que en aquesta base de dades només hi ha introduïdes seqüències confirmades, s'asseguren que només es llisten els lligands naturals rellevants per la molècula de MHC. Es van considerar les posicions P1, P4, P6 i P9 del *core* peptídic. De les diferents combinacions de *cores* resultants per a un mateix pèptid, es va escollir la seqüència que millor complia amb els motius d'unió dels al·lells d'HLA-DR de cada mostra i se li va assignar un valor predictiu en funció del grau d'homologia. Així, els *cores* amb tres o més coincidències amb el motiu es van considerar HB, IB per a dues coincidències i la resta de seqüències es van considerar LB. Les seqüències sense cap homologia amb el motiu d'unió, també es van considerar LB, ja que s'havien identificat d'el·luïts naturals dels al·lells corresponents.

9. DETERMINACIÓ DELS MOTIUS D'ANCORATGE

Per determinar els motius d'ancoratge de l'al·lel HLA-DRB1*0301 i dels al·lells d'H-2, I-A^d i I-E^d, un cop es tenien totes les seqüències peptídiques validades es seleccionaven les seqüències úniques i s'establia la seqüència comú de cada família peptídica (*nested set*), per tal de no esbiaixar els resultats degut a la presència de gran nombre de pèptids en aquestes famílies peptídiques. Aquestes seqüències seleccionades eren les que s'utilitzaven per identificar els nonàmers que formaven els *cores* i generar les matrius de freqüència de cada aminoàcid a cadascuna de les 9 posicions dels *cores*.

Per determinar els *cores* dins de cada seqüència es va fer un programa *ad hoc*. Aquest programa, *Laia motifs* (figura M7), va ser creat pel Dr Joaquin Abián, del LP CSIC-UAB i està disponible a la web (<http://code.google.com/p/mhc-laia-motifs/>) (Muixi et al, 2008). El programa permet identificar dins de cada seqüència tots els nonàmers que podien constituir el *core*, fixant els residus permesos en P1. A partir de la llista de possibles *cores*, es generava una matriu de freqüències de cada aminoàcid en cadascuna de les 9 posicions, que

després es corregia generant una segona matriu normalitzant les freqüències respecte a la freqüència de cadascun dels aminoàcids en el proteoma de vertebrats.

Una vegada s'ha obtingut aquesta matriu, es van limitar el nombre de *cores* fixant característiques específiques de cada al·lel. D'aquesta manera s'escullen els *cores* més probables de cada pèptid i s'obté la matriu final a partir d'un sol *core* per seqüència.

Una descripció més detallada dels passos realitzats per determinar els motius d'ancoratge concrets pels al·lells DR3 i I-A^d i I-E^d, es troba a la secció de resultats del capítol 1 i 4 respectivament.



Figura M7: Programa Laia Motifs. Definint els possibles aminoàcids admesos a P1 i la mida, el programa selecciona els possibles cores dins un grup de seqüències. El programa també permet definir un segon residu d'ancoratge en una altra posició

10. ESTABLIMENT DEL PROTOCOL DE SIMULACIÓ PEL MODELATGE DE LES MOLÈCULES I-A^d i I-E^d

Aquesta part de l'estudi va ser realitzada en col·laboració amb el Dr X. Daura del Grup de Biologia Computacional de l'IBB (Universitat Autònoma de Barcelona).

Es van seleccionar pèptids d'alta i de baixa afinitat del repertori obtingut per I-A^d i I-E^d (capítol 4) i es va fer un modelatge utilitzant un protocol de simulació similar al presentat en un treball anterior del nostre laboratori [95].

10.1. Molècula I-A^d

L'estructura d'I-A^d en complex amb un pèptid d'hemaglutinina de la influença (HA) es va obtenir del Protein Data Bank (entrada 2IAD del PDB) [96, 97]. La configuració inicial del sistema per a la modelització de les interaccions de l'I-A^d amb els pèptids escollits per al modelatge incloïa els residus Ile27-Ala109 de la I-A^d α i His33-Arg120 de la I-A^d β (seqüències d'acord amb UniProt [UniProt Consorci 2014] entrades P04228 i P01921, respectivament), que constitueixen el domini d'unió de l'I-A^d amb pèptids, i els nou residus de l'HA que ocupen les posicions P1 a P9 en el solc d'unió d'I-A^d.

El fragment es va transformar en un pèptid nona-glicina (nou residus de glicina de P1 a P9) mitjançant l'eliminació de les cadenes laterals dels vuit residus del pèptid diferents a la glicina. La simulació utilitzada per al modelatge genera els àtoms de les cadenes laterals corresponents al pèptid en estudi a través d'una tècnica de dinàmica molecular anomenada *slow-growth* en condicions de “no equilibri”, transformant així el pèptid nona-glicina (derivat de l'HA) en el pèptid escollit, dins del solc d'unió a I-A^d. A més, el mètode permet l'assignació d'un score d'afinitat. Per a cada pèptid escollit del repertori, es va generar un esquelet de nona-glicines, amb la resta d'àtoms de les cadenes laterals corresponents a la seqüència d'interès definits com a “*dummies*”, en el camp de força biomolecular del paquet de simulació GROMOS96 [98].

Un àtom “dummy” és un àtom que no té interaccions amb la resta del sistema. Per la resta d'àtoms del sistema, es van assignar paràmetres estàndard del camp de força 54B7 [99] corresponents a condicions d'entorn de buit, i que inclouen l'apantallament de les interaccions càrrega-càrrega amb la finalitat de mimetitzar un dels efectes bàsic del solvent aquós. L'estat de protonació dels residus ionitzables es va triar per mimetitzar un pH neutre. Després d'equilibrar el sistema en el buit a 10 K, es va realitzar una simulació de dinàmica molecular de *slow-growth* en condicions de “no equilibri” (200 ps). Es va triar una temperatura baixa per limitar l'extensió de la dinàmica de la proteïna sota la condició no fisiològica de buit.

Durant la simulació, els àtoms dummies es van transformar progressivament en àtoms regulars (és a dir, en àtoms amb els paràmetres de camp de força estàndard) utilitzant un potencial de anomenat de *soft-core* per evitar singularitats en la integració de les equacions en els passos de temps inicials, mentre que les posicions dels àtoms de l'esquelet de l'I-A^d es van restringir amb un potencial harmònic. Durant el procés, els àtoms que van materialitzant-se (és a dir, els que inicialment eren àtoms dummies) es van acomodant en l'ambient del solc d'unió sota l'única influència del camp de força. La simulació proporciona el treball implicat en la transformació del pèptid nona-glicina en el pèptid d'interès en les condicions especificades [100].

Per acumular estadística sobre aquest procés de no equilibri, la simulació es va repetir 1000 vegades amb diferents distribucions de velocitat inicials. El canvi d'energia lliure associat a la transformació es va estimar a partir dels valors de treball per mitjà de l'equació de Jarzinsky [101]:

$$\Delta G = -kBT \ln \langle \exp(-W/kBT) \rangle$$

on kB és la constant de Boltzmann, T és la temperatura absoluta, W és el treball, i els parèntesis angulars indiquen una mitjana dels valors de treball sobre les 1000 simulacions. Per obtenir un score que es pugués utilitzar per classificar la unió de diferents pèptids, la transformació anterior (nona-glicina pel pèptid en estudi) necessitava ser realitzada també per al pèptid aïllat en buit. El score d'unió es pot aleshores calcular com:

$$S_{\text{bind}} \equiv \Delta \Delta G_{\text{bind}} = \Delta G_{\text{groove}} - \Delta G_{\text{vacuum}}$$

on ΔG_{groove} es refereix a la transformació dins del solc d'unió de l'I-A^d i ΔG_{vacuum} es refereix a la transformació del pèptid aïllat. Tenint en compte que, a causa de les condicions de la transformació, aquest resultat no inclou els efectes de la solvatació i l'entropia, el valor obtingut correspon a l'afinitat entàlpica relativa entre l'I-A^d i el pèptid en estudi.

10.2. Molècula I-E^d

Com l'estructura cristal·logràfica d'I-E^d no estava disponible, es va generar un model per al domini d'unió d'aquesta molècula utilitzant I-E^k com a plantilla. L'estructura d'I-E^k en complex amb un fragment de β -2-globina unida covalentment (HB) es va obtenir del PDB (entrada 1FNG) [102].

Per a la realització del model vàrem utilitzar els residus Ile26-Pro106 de la I-E^k α i Arg30-Pro118 de la I-E^k β (seqüències d'acord amb les entrades d'UniProt P04224 i Q31163, respectivament), que constitueixen el domini d'unió de l'I-E^k a pèptid, i els nou residus d'HB que ocupaven les posicions P1 a P9 en el solc d'unió. Donat el nivell d'identitat entre les seqüències d'ambdues molècules, que difereixen en només 18 residus del domini d'unió de la cadena β (seqüència I-E^d β basada en l'entrada UniProt P01915), es va assumir que ambdues molècules presenten el mateix plegament. Per modelar les 18 cadenes laterals diferents en I-E^d i permetre l'adaptació estructural de la resta de la proteïna a l'allotjament d'aquests canvis, es va utilitzar el mateix procediment explicat anteriorment. En aquest cas, la transformació de I-E^k en I-E^d tenia igualment associada una mesura del treball relatiu a la perturbació, que vàrem utilitzar per a classificar els 1000 models generats. El model amb el valor de treball associat més baix es va utilitzar per a realitzar el modelatge de les interaccions d'I-E^d amb els pèptids indicats a la Taula 1, seguint de nou el mateix protocol explicat per I-A^d.

OBJECTIU ESPECÍFIC 1:

**PRESENTACIÓ DE PÈPTIDS PER HLA-DR17 (DRB1*0301/DRA1*0101),
AL·LEL PREFERENTEMENT ASSOCIAT A MALALTIES AUTOIMMUNITÀRIES**

CAPITOLS:

1. Repertori peptídic presentat per HLA-DR17 (DR3): revisió del motiu d'ancoratge
2. Influència de xaperones HLA-DM i cadena invariant (Ii) en els repertoris presentats per HLA-DR3
3. Presentació del TSHR, autoantígen dominant a la malaltia de Graves-Basedow

ANTECEDENTS

El primer objectiu d'aquest part era estudiar el motiu d'ancoratge de la molècula HLA-DR*0301 (DRA1*0101,DRB1*0301). Tot i existir informació sobre el motiu d'ancoratge d'unió a l'al·lel DR3, es va considerar interessant fer un nou estudi a partir de pèptids naturals associats ja que, les tècniques i eines d'espectrometria de masses estandaritzades al nostre laboratori actualment són molt més sensibles que les que es van fer servir aleshores [103, 104] [103-105]. Els mètodes de predicció de la unió d'un pèptid a un determinat MHC són importants per poder identificar aquells segments d'una proteïna que seran presentats a les cèl·lules T, i per tant ens poden ser útils a l'hora de generar vacunes peptídiques o pel desenvolupament de teràpies antigen-específiques.

Durant els últims anys, s'han desenvolupat molts mètodes computacionals diferents per predir la unió MHC-pèptid, des de anàlisi simple de seqüències [43, 45, 106], a matrius quantitatives [42], o matrius virtuals [45] etc. L'ús de llibreries peptídiques expressades per fags [40, 107], juntament amb l'estudi estructural dels complexos HLA-DR-pèptid aporten també una base per la caracterització de les interaccions MHC-pèptid. Unanue et al. van crear un programa computacional per predir epítops peptídics presentats per molècules de MHC de classe II diabetogèniques a partir d'una base de dades de pèptids naturals [108]. Aquest algoritme els permetia assignar valors d'unió alts i baixos a pèptids que s'unien amb alta o baixa afinitat respectivament. Basant-nos en aquesta idea, al nostre laboratori es va desenvolupar un estudi similar pels al·lells DR10 i DR8 [90, 109] i que en aquest treball utilitzarem per la descripció de DR3.

Com a segon objectiu, hem volgut analitzar la importància de les xaperones DM i la cadena invariant (Ii) en la presentació de pèptids per DR3. Un estudi similar es va realitzar al nostre grup amb DR4 [89] i ens servirà per comparar-ho amb aquest al·lel i també amb altres al·lells d'HLA-DR. Es dona el cas que DR3 és l'al·lel associat a més malalties autoimmunitàries i volem identificar possibles característiques que puguin explicar aquesta associació preferencial. Així doncs, l'objectiu es esbrinar si l'expressió d'Ii i DM afecta al repertori peptídic de

DR3 produint variacions en el seu contingut i en l'origen dels pèptids i si aquestes xaperones poden afectar a l'estabilitat, seqüència i afinitat dels pèptids de cada repertori.

Conèixer el motiu d'ancoratge de DR3 així com les característiques d'aquest al·lel en quan a la presentació de pèptids a partir d'eluits naturals ens pot ser una eina important per futurs anàlisis d'epítops d'antígens diana desconeguts en malalties immunitàries i ens ajudaria a entendre els mecanismes de presentació d'aquesta molècula .

Entre la susceptibilitat a malalties associades a HLA-DR3 es troben les autoimmunitats endocrines com ara la diabetis tipus 1 i la malaltia de Graves – Basedow (GB). Per aquest motiu, el nostre tercer objectiu fou crear un model que mimetitzara el que ocorre a GB, a partir de cèl·lules epitelials del tiroides (HT93) heterozigotes per DR3, que foren transfectades amb TSHR i un activador dels gens de classe II (expressats ectòpicament a les cèl·lules durant la malaltia). En estudis previs al nostre laboratori, havíem demostrat l'associació *in vivo* de pèptids de tiroglobulina (Tg) a molècules d'HLA-DR en teixit tiroïdal afectat per la malaltia de Graves-Basedow [95]. També s'havia demostrat que almenys un d'aquests pèptids era capaç d'induir tiroïditis en un model experimental utilitzant ratolins transgènics per HLA-DR3 [110]. Tenint en compte que l'antigen principal de GB és el receptor de la tirotropina (TSHR), el nostre objectiu fou estudiar els pèptids del TSHR que s'uneixen a l'HLA-DR3 *in vivo* i conèixer a més, si les característiques observades en els altres dos tipus cel·lulars DR3 dels objectius anteriors ocorren també amb aquesta línia amb expressió de DR induïda.

CAPITOL 1: REPERTORI PEPTÍDIC PRESENTAT PER HLA-DR3: REVISIÓ DEL MOTIU D'ANCORATGE

1.1. PROCESSAMENT DE LES MOSTRES

L'anàlisi del repertori peptídic de DR3 es va fer per seqüenciació dels pèptids eluïts de la línia CDG, homozigota per a HLA-DR3 (HLA-DR17: DRA1*0101; DRB1*0301), de la qual es van créixer 1000 milions de cèl·lules (veure materials i mètodes).

1.2. REPERTORI DE PÈPTIDS ASSOCIATS A HLA-DR3 (DRA1*0101; B1*0301)

Es van identificar 246 espectres de fragmentació MS/MS mitjançant la cerca en una base de dades humana, utilitzant el sistema de recerca *Protein Discover* v1.4, (ThermoFisher, San Jose, CA, USA). La validació de les assignacions s'explica detalladament a l'apartat 7 del capítol de Materials i mètodes. En aquest cas, el FDR utilitzat fou de 1%. El nombre de lligands finals obtinguts va ser 201, que pertanyien a 121 proteïnes diferents (taula 1.1).

Mostra	Espectres identificats	Seqüències finals	Proteïnes(*)	% nesteds
CDG	246	201	121	32

(*) Proteïnes: s'indica el nombre de grups de proteïnes calculat per Sequest (ver material i mètodes)

Taula 1.1: Resum del nombre de seqüències obtingudes per a la línia cel·lular CDG.

A la taula 1.2, es mostren les 201 seqüències dels pèptids associats a classe II identificats per la mostra CDG.

Taula 1.2: Seqüències dels pèptids associats a classe II de la mostra CDG.

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
P16885	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-2	IYDPNLAFLRF	11	C
Q15029	116 kDa U5 small nuclear ribonucleoprotein component	LELAKQDVVLNYPM	14	N
Q9H0U6	39S ribosomal protein L18, mitochondrial	STSSEPAAKPEVD	13	MIT
P39019	40S ribosomal protein S19	PGVTVKDVNQQE	12	N
P61313	60S ribosomal protein L15	EVILIDPFHK	10	C
P62888	60S ribosomal protein L30	PGDSDIIRSMPEQTGEK	17	N
P62899	60S ribosomal protein L31	KNLQTVNVNVDEN	11	C
P62899	60S ribosomal protein L31	VTYVPVTTFKNLQTVNVND	18	
P62899	60S ribosomal protein L31	YVPVTTFKNLQTVNVNVDEN	18	
P62899	60S ribosomal protein L31	VTYVPVTTFKNLQTVNVNVDEN	20	
P32969	60S ribosomal protein L9	VSEKGTVQQADE	12	C
P11021	78 kDa glucose-regulated protein	YGSAGPPPTGEEDTAEKDEL	20	C
P60709	Actin, cytoplasmic 1	AGFAGDDAPR	10	C
P60709	Actin, cytoplasmic 1	EIVRDIKEKL	10	
P60709	Actin, cytoplasmic 1	DSYVGDEAQSK	11	
P60709	Actin, cytoplasmic 1	GIVMDSGDGVTH	12	
P60709	Actin, cytoplasmic 1	YASGRTTGIVMD	12	
P60709	Actin, cytoplasmic 1	YASGRTTGIVMDSGD	15	
P60709	Actin, cytoplasmic 1	MATAASSSSLEK	12	
P60709	Actin, cytoplasmic 1	MATAASSSSLEKS	13	
P60709	Actin, cytoplasmic 1	EQEMATAASSSSLEKS	16	
P60709	Actin, cytoplasmic 1	FEQEMATAASSSSLEK	16	
O00767	Acyl-CoA desaturase	TTITAPPSR	9	ER
Q12802	A-kinase anchor protein 13	DTSSLPTVIMR	11	C
Q96DE5	Anaphase-promoting complex subunit 16	SVSDLAPPR	9	C
P20073	Annexin A7	AIVDVVANR	9	C
P46013	Antigen KI-67	EVVDTTASTKR	11	N
P04114	Apolipoprotein B-100	IEGNLIFDPNNYLPK	15	S
P04114	Apolipoprotein B-100	IEGNLIFDPNNYLPKE	16	
P05089	Arginase-1	INTPLTTTSGNLHGQP	16	C
P05089	Arginase-1	INTPLTTTSGNLHGQPV	18	
Q9ULA0	Aspartyl aminopeptidase	FELFPSLSHLLVD	14	C
Q8WWZ7	ATP-binding cassette sub-family A member 5	DEADILADRKAVISQ	15	END
Q8WWZ7	ATP-binding cassette sub-family A member 5	DEADILADRKAVISQG	16	
Q92499	ATP-dependent RNA helicase DDX1	GYLPNQLFRFT	11	N
P15907	Beta-galactoside alpha-2,6-sialyltransferase 1	DAVLRFNAGAPTANFQ	15	G
P15907	Beta-galactoside alpha-2,6-sialyltransferase 1	DAVLRFNAGAPTANFQQ	16	
P27824	Calnexin	DEDAPAKIPDEE	12	ER
P25774	Cathepsin S	DNKGIDSDASYPYKAMD	17	LIS
P25774	Cathepsin S	VLKEAVANKGPVSVGVD	17	
P25774	Cathepsin S	DVLKEAVANKGPVSVGVD	18	
O00299	Chloride intracellular channel protein 1	FSAYIKNSNPALNDN	15	N
Q00610	Clathrin heavy chain 1	ATAGIIGVNR	10	C
P09496	Clathrin light chain A	ISLKQAPLVH	10	C
Q7Z460	CLIP-associating protein 1	TSVGSPSNTIGR	12	C
Q15517	Corneodesmosin	GILNPSQPGQSSSSQT	17	S
P20674	Cytochrome c oxidase subunit 5A, mitochondrial	NELGISTPEELGLDKV	16	MIT
P15954	Cytochrome c oxidase subunit 7C, mitochondrial	ATPFLVVRHQLLKT	14	MIT
Q8NF50	Dedicator of cytokinesis protein 8	EVTDVDSVVGR	11	C
Q02413	Desmoglein-1	VVTGNMGSNNDKVG	14	MB
Q02413	Desmoglein-1	VVTGNMGSNNDKVGDF	15	
Q9P265	Disco-interacting protein 2 homolog B	TADGVPVSSR	10	C
Q16531	DNA damage-binding protein 1	EVQDTSGGTTALR	13	N
Q9H324	DnaJ homolog subfamily C member 5	TTQLTADSHPSYHTD	15	MB
Q9H324	DnaJ homolog subfamily C member 5	TTQLTADSHPSYHTDG	16	
Q9H324	DnaJ homolog subfamily C member 5	ETTQLTADSHPSYHTDG	17	
Q7L591	Docking protein 3	YAGGPSGVAR	10	C
Q86TM6	E3 ubiquitin-protein ligase synoviolin	AVTDAIMSR	9	RE
P24534	Elongation factor 1-beta	GFGDLKSPAGLQVLND	16	C
P14625	Endoplasmic	SGYLLPDTKAYGDR	14	ER
Q96AE4	Far upstream element-binding protein 1	YYAQTSPQGMPPQHPAPQGG	20	C
Q5D862	Filaggrin-2	PVLKNPDDPDVTD	13	C
P21333	Filamin-A	AEIVEGENHTY	11	C
P85037	Forkhead box protein K1	SSAPVVVTR	9	N
Q96RU3	Formin-binding protein 1	DYSSILQKF	9	MB
P04075	Fructose-bisphosphate aldolase A	SESLFVSNHAY	11	C

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
P30279	G1/S-specific cyclin-D2	ETSLPTAEKL	10	N
P13284	Gamma-interferon-inducible lysosomal thiol reductase	LDFFGNGPPVNYK	13	LIS
P13284	Gamma-interferon-inducible lysosomal thiol reductase	LDFFGNGPPVNYKTG	15	
P13284	Gamma-interferon-inducible lysosomal thiol reductase	ALDFFGNGPPVNYKTG	16	
P06744	Glucose-6-phosphate isomerase	TPILVDGKDVMPPE	13	C
P06744	Glucose-6-phosphate isomerase	TPILVDGKDVMPPEV	14	
P06744	Glucose-6-phosphate isomerase	TPILVDGKDVMPPEVN	15	
P15104	Glutamine synthetase	LNETGDEPFQYKN	13	C
P04406	Glyceraldehyde-3-phosphate dehydrogenase	GYSNRVVDLMAHMASKE	17	C
Q9NU53	Glycoprotein integral membrane protein 1AN	EEVVEIDGKQVQKQ	14	MB
Q9NU53	Glycoprotein integral membrane protein 1AN	EEVVEIDGKQVQKQDV	16	
Q9NU53	Glycoprotein integral membrane protein 1AN	EEVVEIDGKQVQKQDVT	17	
P27144	GTP:AMP phosphotransferase AK4, mitochondrial	FSNKITPIQSKEAY	14	MIT
P62826	GTP-binding nuclear protein Ran	APPEVVMDDPALAAQYEH	17	N
P62826	GTP-binding nuclear protein Ran	APPEVVMDDPALAAQYEH	18	
P54652	Heat shock-related 70 kDa protein 2	LNKSINPDEAVAYG	14	C
P54652	Heat shock-related 70 kDa protein 2	LNKSINPDEAVAYGAA	16	
P69905	Hemoglobin subunit alpha	EYGAEALERM	10	C
P69905	Hemoglobin subunit alpha	GEYGAEALERM	11	
P69905	Hemoglobin subunit alpha	AGEYGAEALERM	12	
P69905	Hemoglobin subunit alpha	LASVSTVLTSKY	12	
P69905	Hemoglobin subunit alpha	VLSPADKTNVKA	12	
P69905	Hemoglobin subunit alpha	VLSPADKTNVKAA	13	
P69905	Hemoglobin subunit alpha	GKVGAGAHGEYGAEALERM	18	
Q13151	Heterogeneous nuclear ribonucleoprotein A0	AEGDLIEHF	9	N
Q9BUJ2	Heterogeneous nuclear ribonucleoprotein U-like protein 1	YGSYSNGTQGGTSTQ	15	N
Q9BUJ2	Heterogeneous nuclear ribonucleoprotein U-like protein 1	DYGSYSNGTQGGTSTQ	16	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	SGNFGGSRNMGGP	13	N
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	IKDDEKEAEEGEDDRD	16	N
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	IKDDEKEAEEGEDDRDSANGEDDS	24	
Q60814	Histone H2B type 1-K	PEPAKSAPAP	10	C
Q86X55	Histone-arginine methyltransferase CARM1	MASPMISPTNTMHYGS	16	N
P04439	HLA class I histocompatibility antigen, A-3 alpha chain	AAVVVPSGEEQRYT	14	MB
P04439	HLA class I histocompatibility antigen, A-3 alpha chain	VDDTQFVRFDSDAASQ	16	
P04439	HLA class I histocompatibility antigen, A-3 alpha chain	VDDTQFVRFDSDAASQR	17	
P10319	HLA class I histocompatibility antigen, B-58 alpha chain	DTAAQITQR	9	MB
P10319	HLA class I histocompatibility antigen, B-58 alpha chain	EAGSHIQR	9	
P10319	HLA class I histocompatibility antigen, B-58 alpha chain	TAADTAAQITQR	12	
P10319	HLA class I histocompatibility antigen, B-58 alpha chain	VDDTQFVRFDSDAASPR	17	
Q31612	HLA class I histocompatibility antigen, B-73 alpha chain	AAVVVPSGEEQRYT	14	MB
P04233	HLA class II histocompatibility antigen gamma chain CD74	PSSGLGVTKQDLGPVPM	17	MB
P04233	HLA class II histocompatibility antigen gamma chain CD74	RMATPLLMQALPM	13	
P04233	HLA class II histocompatibility antigen gamma chain CD74	MATPLLMQALPMGALPQ	17	
P04233	HLA class II histocompatibility antigen gamma chain CD74	RMATPLLMQALPMGALP	17	
P04233	HLA class II histocompatibility antigen gamma chain CD74	RMATPLLMQALPMGALPQ	18	
P01909	HLA class II histocompatibility antigen, DQ alpha 1 chain	IVADHVASYGVN	12	MB
P01909	HLA class II histocompatibility antigen, DQ alpha 1 chain	DIVADHVASYGVN	13	
P01909	HLA class II histocompatibility antigen, DQ alpha 1 chain	EDIVADHVASYGVN	14	
P01903	HLA class II histocompatibility antigen, DR alpha chain	ANIAVDKANLEIMT	14	MB
P01903	HLA class II histocompatibility antigen, DR alpha chain	LANIAVDKANLEIM	14	
P01903	HLA class II histocompatibility antigen, DR alpha chain	LANIAVDKANLEIMT	15	
P01903	HLA class II histocompatibility antigen, DR alpha chain	ALANIAVDKANLEIMT	16	
P24001	Interleukin-32	LAELEDDFKEGYLE	14	S
P24001	Interleukin-32	LAELEDDFKEGYLET	15	
P24001	Interleukin-32	LAELEDDFKEGYLETV	16	
Q04760	Lactoylglutathione lyase	IEILNPNKMATLM	13	C
Q99538	Legumain	VPKDYTGEDVTPQN	14	LYS
P19397	Leukocyte surface antigen CD53	IYLIHNNF	9	MB
P11279	Lysosome-associated membrane glycoprotein 1	LNTILPDARDPAFK	14	LYS
P14174	Macrophage migration inhibitory factor	PMFIVNTNVPR	11	C
P14174	Macrophage migration inhibitory factor	PMFIVNTNVPRASVPD	16	
P14174	Macrophage migration inhibitory factor	IVNTNVPRASVPDGFSEL	19	
P40926	Malate dehydrogenase, mitochondrial	AKVAVLGASGGIGQPLSL	18	MIT
Q86WK9	Membrane progesterin receptor alpha	EFWHYSFFF	9	MB
Q9UBU8	Mortality factor 4-like protein 1	TSETPQPPR	9	C
Q9UBU8	Mortality factor 4-like protein 1	STSETPQPPR	10	
P35579	Myosin-9	IDQINTDLNLSHAQ	16	C
Q14697	Neutral alpha-glucosidase AB	AVAAVAARR	9	ER

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
Q15758	Neutral amino acid transporter B(0)	GPAGDATVASEKESVM	16	MB
P67809	Nuclease-sensitive element-binding protein 1	PPAENSSAPEAEQGG	15	N
P67809	Nuclease-sensitive element-binding protein 1	PPAENSSAPEAEQGGAE	17	
P67809	Nuclease-sensitive element-binding protein 1	KETKAADPPAENSSAPEAEQGGGA	25	
P12270	Nucleoprotein TPR	EVVQSSPVER	10	C
Q99733	Nucleosome assembly protein 1-like 4	EEVHDLERKY	10	C
P19021	Peptidyl-glycine alpha-amidating monooxygenase	IPVDEEAFVIDFKPRASMD	19	MB
P19021	Peptidyl-glycine alpha-amidating monooxygenase	IPVDEEAFVIDFKPRASMDT	20	
P05121	Plasminogen activator inhibitor 1	APEEIMDRPFLF	13	S
P05121	Plasminogen activator inhibitor 1	APEEIMDRPFLFVV	15	
P13796	Plastin-2	SEQSSVGTQHSY	12	C
P13796	Plastin-2	VPAVVIDMSGLEK	14	
P13796	Plastin-2	VPAVVIDMSGLEKDD	15	
P13796	Plastin-2	VPAVVIDMSGLEKDD	16	
Q15366	Poly(rC)-binding protein 2	VRLSSETGGMGSS	13	C
Q4VXU2	Polyadenylate-binding protein 1-like	YVGDLHPDVTEAM	13	C
P01189	Pro-opiomelanocortin	YGGFMTSEKSTQPLVT	16	S
Q9UL46	Proteasome activator complex subunit 2	NVADLTSLR	9	C
P28066	Proteasome subunit alpha type-5	GPQLFHMDPSGTFVQ	15	C
Q15084	Protein disulfide-isomerase A6	DVELDDLKDEL	12	ER
Q15084	Protein disulfide-isomerase A6	IDLSDELDDLKDEL	16	
O15297	Protein phosphatase 1D	TTASVVIIR	9	N
P06454	Prothymosin alpha	DLKEKKEVV	9	N
P06454	Prothymosin alpha	AVDTSSEITTK	11	
P62491	Ras-related protein Rab-11A	TRSIQVDGKTIKAQ	14	END
P62491	Ras-related protein Rab-11A	TRSIQVDGKTIKAQI	15	
Q9H0U4	Ras-related protein Rab-1B	IRTIELDGTIKLQ	14	END
P52566	Rho GDP-dissociation inhibitor 2	DGPVVTPDKAPNVV	14	N
P56182	Ribosomal RNA processing protein 1 homolog A	LAGNEQVTR	9	N
P35637	RNA-binding protein FUS	PPSAKAAID	9	C
P35637	RNA-binding protein FUS	GPGGGPGGSHMGNGY	16	
P35637	RNA-binding protein FUS	GPGGGPGGSHMGNGYD	17	
P35637	RNA-binding protein FUS	GPGGGPGGSHMGNGYD	18	
Q92854	Semaphorin-4D	YTQIVVDRTQALDG	14	MB
P10124	Serglycin	DLGQHGLEEDFM	12	C
P10124	Serglycin	LPDSQDLGQHGLEEDFM	18	
P10124	Serglycin	LPDSQDLGQHGLEEDFML	19	
P52630	Signal transducer and activator of transcription 2	NTVDEVYVSR	10	C
P62306	Small nuclear ribonucleoprotein F	GVEEEEDGEMRE	13	C
Q92922	SWI/SNF complex subunit SMARCC1	GVPPPPAPGPPASAAP	16	N
Q8TAQ2	SWI/SNF complex subunit SMARCC2	AVQGNLLPASPLPD	15	N
Q8TAQ2	SWI/SNF complex subunit SMARCC2	PTAPSPGTVTPVPPQ	16	
O00560	Syntenin-1	ITSIVKDSSAARNGLL	16	C
Q86V81	THO complex subunit 4	DAQLDAYNARMDTS	14	C
P02786	Transferrin receptor protein 1	SVIIVDKNGRLV	12	MB
P02786	Transferrin receptor protein 1	NSVIIVDKNGRLV	13	
P02786	Transferrin receptor protein 1	SVIIVDKNGRLVY	13	
P02786	Transferrin receptor protein 1	NSVIIVDKNGRLVY	14	
P02786	Transferrin receptor protein 1	NSVIIVDKNGRLVYL	15	
P02786	Transferrin receptor protein 1	SVIIVDKNGRLVYL	15	
P02786	Transferrin receptor protein 1	DSAQNSVIIVDKNGRLV	17	
P02786	Transferrin receptor protein 1	DSAQNSVIIVDKNGRLVY	18	
P61586	Transforming protein RhoA	VADIEVDGKQVE	12	MB
Q92973	Transportin-1	ESQSPDTTIQR	11	C
Q08J23	tRNA (cytosine(34)-C(5))-methyltransferase	DPAGVHPPR	9	N
P07437	Tubulin beta chain	DSDLQLDRISVYYNEA	16	C
P07437	Tubulin beta chain	GDSDLQLDRISVYYNE	16	
P07437	Tubulin beta chain	GDSDLQLDRISVYYNEA	17	
P21580	Tumor necrosis factor alpha-induced protein 3	RPIIVISDKMLRS	13	C
Q94966	Ubiquitin carboxyl-terminal hydrolase 19	TTVDESQVTR	11	ER
P18206	Vinculin	ISPMVMADAKAVAGN	14	C
P18206	Vinculin	ISPMVMADAKAVAGNI	15	
Q16864	V-type proton ATPase subunit F	NEIEDTFRQF	10	C
Q96NC0	Zinc finger matrin-type protein 2	AVMGFSFGSTKKS	15	N
Q07352	Zinc finger protein 36, C3H1 type-like 1	TPAGGGFPR	9	C
Q92785	Zinc finger protein ubi-d4	LKEKASIYQNNSS	14	C
Q8TCW7	Zona pellucida-like domain-containing protein 1	NTQLASSAAIS	12	MB

Uniprot AN: Identificació de la proteïna en SwissProt; Proteïna: proteïna origen de la qual prové el pèptid; Longitud: mida del pèptid en nombre d'aminoàcids; Localització: MB (membrana plasmàtica); Lis (lisosomes); End (endosomes); S (secretada); EM (matriu extracel·lular); ER (Reticle endoplàsmic); G (Golgi); MIT (mitochondria) C (citòsol); N (nuclear).

La mida dels pèptids presentava una distribució normal i un rang d'entre 9 i 25 aminoàcids (figura 1.1). Tot i que un 90% dels pèptids tenien la mida estàndard de classe II, entre 10 i 20 aminoàcids [13, 17, 18], un 10% presentava una mida de 9 aminoàcids poc comú en altres al·lels que s'havien descrit al nostre laboratori. Una gran part de les seqüències identificades (32%) formaven part de *nested sets*, tret característic dels pèptids presentats per les molècules de classe II [35, 111]. La massa mitjana dels pèptids fou de 1364 Da.

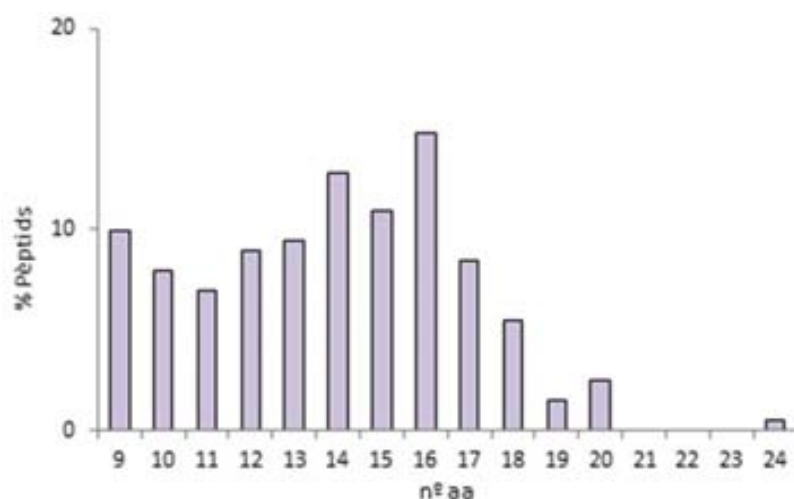


Figura 1.1: Distribució en funció de la mida dels pèptids aïllats d'HLA-DR3

Els pèptids provenien de 121 proteïnes diferents i de diversos compartiments cel·lulars (taula 1.2). Per evitar redundàncies, s'usava la localització més comú per aquelles proteïnes que es podien trobar en diferents compartiments. Fins al moment s'havia descrit que la majoria de pèptids aïllats d'altres al·lels d'HLA-DR provenien de proteïnes exògenes, de membrana o eren residents a la via endocítica [112] [90] [84] amb una presència minoritària de pèptids generats per degradació al citosol, però en aquest cas, la ruta citosòlica era més abundant (63%) que l'endocítica, fet que no s'havia descrit per altres al·lels de HLA-DR (figura 1.2).

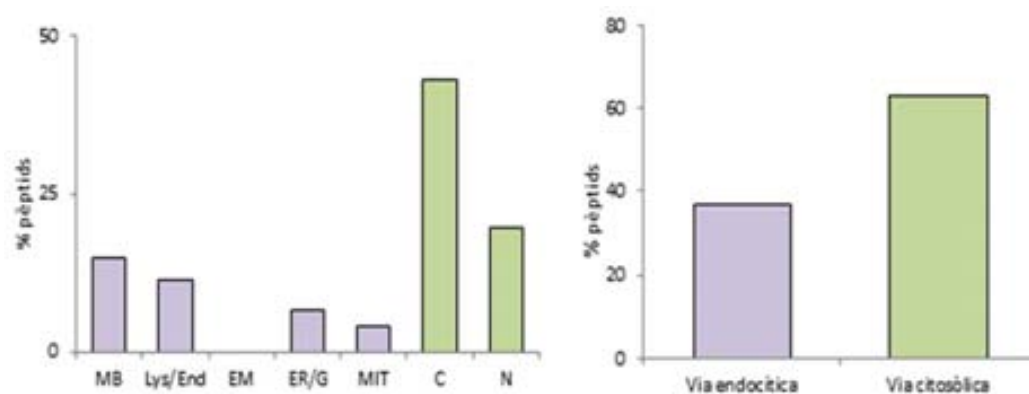


Figura 1.3: Localització de les proteïnes origen dels pèptids aïllats de DR3 en la línia homozigota CDG. Gràfic en funció de la seva localització cel·lular: MB (membrana plasmàtica); Lis/End (lisosomes o endosomes); EM (matriu extracel·lular); ER/G (Reticle endoplàsmic o Golgi); MIT (mitocòndria) C (citosol); N (nuclear). Al gràfic de l'esquerra, les localitzacions s'agrupen en funció de la seva via de degradació, via endocítica (barres morades) o citosòlica (barres verdes).

Dins de la via endocítica (37%), trobarem proteïnes relacionades amb la presentació i processament d'antigen, com ara la cathepsina S, CLIP-associating protein 1, proteïnes de la cadena alfa d'HLA de classe I (A-3, B-58, B-73). També CD74, HLA-DQ, HLA-DR, IL-32 i la xaperona endoplasmica, implicada en el processament i el transport de les proteïnes secretades.

D'altra banda, a la via citosòlica es van identificar dos proteïnes que formen part del proteosoma: la subunitat 2 del complex activador del proteosoma i la subunitat *alpha* tipus 5. També 10 lligands diferents d'actina citoplasmàtica que pertanyien a 5 regions diferents de la proteïna parental. A més, es van identificar diferents tipus de ribonucleoproteïnes i de les subunitats 40S i 60S dels ribosomes.

1.3. DESCRIPCIÓ DEL MOTIU D'ANCORATGE D'HLA-DR3 (DRA1*0101; B1*0301)

1.3.1. Identificació dels nonàmers d'unió a HLA-DR3

El mètode que es va utilitzar per determinar els motius d'ancoratge de la molècula d'HLA-DR3, es basava en la idea que els perfils de les butxaques del

groove de classe II, on s'uneixen els principals residus d'ancoratge dels pèptids són gairebé independents de la resta del solc d'unió del pèptid [113]. Així, un cop s'ha determinat el perfil d'unió d'una butxaca al seu corresponent residu d'ancoratge *in vitro*, aquest pot ser compartit per altres al·lels sempre i quan els aminoàcids que formen part de la butxaca siguin idèntics [45]. Existeixen una gran varietat d'al·lels de DRB1 i fins al juny de 2014 s'havien descrit 12.010 a la base d'IMGT/HLA [114]. És important remarcar però que tot i el gran polimorfisme d'aquests gens, hi ha regions que no presenten gaires modificacions i per tant si aquestes zones més conservades formen part d'alguna butxaca d'unió al pèptid, es pot assumir que aquells al·lels idèntics uniran en aquella posició els mateixos tipus d'aminoàcids. Aquest és el cas de la regió que forma part de la butxaca on es situa la posició 1 del pèptid, que en molts al·lels és idèntic.

Comparació de les seqüències de les cadenes β dels diferents al·lels

Per conèixer quins eren els aminoàcids de la cadena β de la molècula d'HLA-DR implicats en la formació de les butxaques d'unió al pèptid [90] es van alinear les seqüències de DRB1*0101 i DRB1*0301. A la figura 1.3 es mostren en diferents colors els aminoàcids que formen les diferents butxaques d'interacció amb els residus d'ancoratge dels pèptids. La regió de la cadena β que forma la butxaca on es col·loca la posició 1 del pèptid (P1), del residu 80 al 91, és una zona comú a molts al·lels d'DRB1, i aquest és el cas de DRB1*0301, que només té un aminoàcid diferent a DRB1*0101, un al·lel que s'ha estudiat àmpliament, del qual es coneix el motiu d'ancoratge i que es considera de referència.

Per tant, es va assumir que DRB1*0301 unia a P1 els mateixos residus que DRB1*0101, és a dir, aminoàcids amb cadenes laterals hidrofòbiques o aromàtiques, preferentment Phe, Ile, Leu, Val, Tyr, Trp, Ala, Met (anomenats aminoàcids clàssics)

(<http://www.syfpeithi.de/Scripts/MHCServer.dll/FindYourMotif.htm>).

A més, vam observar que a la butxaca on es col·loca la posició 4 del pèptid (P4, en blau a la figura) hi havia dos residus bàsics: lisina (K) en posició 71 i arginina (R) en posició 74. Per contra, a la butxaca on es col·loca la posició 9 del pèptid (P9, en roig a la figura) vam trobar dos residus àcids: àcid glutàmic (E) i aspàrtic (D). Aquest fet ens podia indicar la preferència per residus àcids en P4 i bàsics en P9.

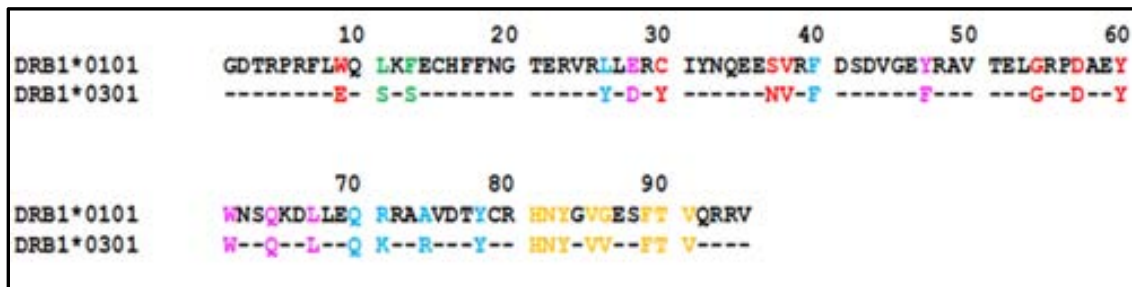


Figura 1.3: Seqüències de les cadenes β dels al·lels DRB1*0101 i *0301. Amb diferents colors s'indiquen els aminoàcids implicats en la formació de les butxaques P1 (taronja), P4 (blau), P6 (verd), P7 (rosa) i P9 (roig), implicades en la unió amb el pèptid.

1.3.2. Selecció dels nonàmers

Un cop sabíem els aminoàcids que podien localitzar-se preferentment a la posició P1 del solc d'unió, es van definir tots els possibles nonàmers amb aquest P1 dins de cada seqüència identificada a les cèl·lules CDG. Per fer això de forma automàtica vam utilitzar el programa LaiaMotifs [90] que seleccionava tots els possibles nonàmers d'un conjunt de seqüències, definint els residus de la posició 1. Aquest programa va ser dissenyat pel Dr. Joaquin Abián del laboratori de Proteòmica CSIC-UAB, i està disponible a <http://code.google.com/p/mhc-laia-motifs/>. El programa *Laia Motifs*, calculava també una matriu de freqüències i una segona matriu de freqüències normalitzades respecte la freqüència mitja de cada aminoàcid dins el proteoma humà. Els resultats es visualitzaven en una fulla d'excel i també gràficament.

1.3.3. Generació de les matrius per la caracterització del motiu d'unió a DR3

En funció del nombre de residus aromàtics i hidrofòbics i de la seqüència, cada un dels pèptids podien generar diferent nombre de *cores* de 9 aminoàcids. Seguint aquest model es van obtenir 501 nonàmers, a partir dels quals es va generar la primera matriu de freqüències, calculant la freqüència de cada un dels 20 aminoàcids a cadascuna de les 9 posicions (figura 1.4).

A partir d'aquí, s'anaven generant matrius cada vegada més restringides en funció dels resultats que s'anaven obtenint.

A la primera matriu (M1), es va fixar la posició 1 amb els aminoàcids clàssics abans esmentats i es va comprovar si a les posicions P4 i P9 s'obtenia un percentatge alt d'aminoàcids àcids i bàsics respectivament, degut al caràcter bàsic i àcid dels residus de les butxaques per a aquestes posicions que havíem observat amb l'alineament de seqüències. Així doncs, vam obtenir un 15% d'àcid aspàrtic (Asp) a P4 mentre que a P9 no vam observar un percentatge alt d'aminoàcids bàsics. Per tant, vam generar una nova matriu (M2), en la qual es fixava P1 com abans i també P4, on s'incloïen tots els aminoàcids menys els bàsics Lys, His, Arg. Seguint aquest model, el nombre de nonàmers es va reduir a 441 nonàmers, amb un 17% d'Asp a P4.

Dels 441 nonàmers, 276 eren únics (es van eliminar repeticions) i d'aquests es seleccionava el *core* que millor s'ajustava a la matriu 2 per a cada seqüència peptídica. Quan més d'un *core* era acceptat per a una única seqüència, es seleccionava el que tenia més residus en comú amb la matriu 2. Els nonàmers resultants ($n = 124$) van ser utilitzats per generar una matriu 3 (M3), fixant una vegada més els residus clàssics per a P1. A continuació, la llista original de pèptids es va comparar amb la llista de *cores* finals generats a la matriu 3. Els pèptids que es van quedar fora de la primera anàlisi a causa de les restriccions en P1 o P4, es van comparar de forma manual amb el motiu de la matriu 3. Així, es va obtenir una matriu de freqüència final (matriu 4) on el residu dominant a P4 era l'àcid aspàrtic amb un 35% de freqüència, però on a més trobàvem Asn en un percentatge menor però no menyspreable (13%) i de Ser (11%). A la posició P1 també va destacar la dominància de Isoleucina (28%) per sobre de la resta d'aminoàcids clàssics per a la P1. A la posició P9 vam

obtenir un 20% d'arginina (R), resultat que lliga perfectament amb el caràcter bàsic de la butxaca (figura 1.4).

M1	1	2	3	4	5	6	7	8	9
A	15	7	5	7	8	7	6	8	5
C	0	0	0	0	0	0	0	0	0
D	0	13	12	15	8	5	8	3	4
E	0	5	5	4	5	3	7	6	5
F	9	3	4	5	3	4	4	2	4
G	0	5	6	4	8	6	7	7	4
H	0	1	1	5	2	3	2	3	2
I	22	9	9	7	4	6	2	5	2
K	0	2	3	5	5	7	3	4	4
L	11	4	3	4	4	6	5	5	6
M	16	7	9	7	7	11	12	10	10
N	0	6	6	9	9	9	8	6	7
P	0	7	9	7	8	8	8	6	5
Q	0	5	4	2	6	4	6	7	8
R	0	2	4	3	5	2	5	8	9
S	0	4	5	5	8	6	5	4	5
T	0	6	6	4	4	5	3	5	5
V	18	12	8	6	5	8	6	6	5
W	0	0	0	0	0	0	0	0	0
Y	8	2	1	2	1	2	4	4	9

M2	1	2	3	4	5	6	7	8	9
A	15	6	5	7	7	7	6	9	6
C	0	0	0	0	0	0	0	0	0
D	0	11	9	17	9	6	8	3	4
E	0	6	5	5	6	3	7	6	5
F	9	3	4	6	3	4	5	2	5
G	0	5	5	5	8	5	7	8	4
H	0	1	2	0	2	3	2	2	3
I	21	11	10	8	3	5	2	4	2
K	0	2	4	0	6	7	3	4	4
L	11	5	3	4	5	6	5	4	7
M	17	7	10	8	7	12	11	11	10
N	0	6	6	10	8	8	8	7	7
P	0	7	9	9	8	8	9	6	5
Q	0	6	3	2	5	4	6	7	6
R	0	2	4	0	5	2	3	9	10
S	0	3	5	6	9	6	5	4	5
T	0	6	7	4	4	5	3	6	4
V	18	10	8	7	5	6	6	7	4
W	0	0	0	0	0	0	0	0	0
Y	9	2	1	3	1	2	4	3	11

M3	1	2	3	4	5	6	7	8	9
A	14	4	5	5	6	11	7	10	7
C	0	0	0	0	0	0	0	0	0
D	0	6	8	33	4	2	7	2	4
E	0	7	6	2	8	4	11	6	4
F	14	2	3	2	3	0	5	0	7
G	0	7	8	4	10	2	7	7	6
H	0	5	0	0	2	5	6	5	0
I	27	8	7	4	5	6	2	6	2
K	0	3	2	0	3	11	1	4	3
L	11	5	4	3	4	5	3	5	6
M	7	2	12	3	10	7	7	2	7
N	0	7	2	14	2	8	7	6	6
P	0	6	15	5	9	8	8	8	4
Q	0	8	4	1	6	7	9	8	2
R	0	2	2	0	8	1	1	6	20
S	0	5	6	11	6	7	3	4	4
T	0	4	7	4	6	6	8	9	5
V	15	13	8	5	5	6	6	7	2
W	0	0	0	0	0	0	0	0	0
Y	12	5	1	4	3	3	1	4	10

M4	1	2	3	4	5	6	7	8	9
A	13	4	6	5	6	12	7	8	7
C	0	0	0	0	0	0	0	0	0
D	0	6	6	35	5	2	8	0	4
E	0	7	5	3	6	3	11	8	4
F	15	2	3	1	2	0	5	0	7
G	0	8	7	3	11	1	6	5	6
H	0	3	0	2	2	5	6	5	1
I	28	9	7	4	6	5	1	6	2
K	0	3	2	2	3	12	1	4	3
L	12	5	5	3	4	5	3	6	5
M	5	3	13	3	10	8	8	3	7
N	0	5	2	13	2	9	7	7	5
P	0	6	14	3	8	9	9	8	3
Q	0	6	5	1	9	6	10	10	3
R	0	2	2	0	8	1	0	7	20
S	0	5	6	11	7	7	3	5	3
T	0	5	7	4	6	7	7	7	5
V	15	13	9	4	4	6	7	7	3
W	0	0	0	0	0	0	0	0	0
Y	12	6	1	3	1	1	1	4	10

Figura 1.4: Seqüències de les diferents matrius obtingudes en la generació del motiu d'ancoratge de DR3. M1: Matriu amb els 397 nonàmers possibles; M2: Matriu seleccionant la posició 4 (343 nonàmers); M3: Matriu generada seleccionant els cores per a cada pèptid que s'ajustava a la matriu 2. (91 nonàmers); M4: Matriu final generada seleccionant un únic nonàmer per pèptid.

1.3.4. Control del sistema de generació de les matrius

Per comprovar que el sistema per generar les matrius no era subjectiu i que el fet de seleccionar determinades posicions des d'un principi no esbiaixava els

resultats obtinguts, es van generar altres matrius seguint diferents criteris, fixant per exemple P1 i P9 amb tots els residus no àcids. Els resultats que es van obtenir van ser molt semblants als que hem obtingut fixant P4, de manera que s'obtenia un alt percentatge d'Asp a la posició P4 i Arg en P9. La gran majoria dels *cores* seleccionats a la matriu final eren els mateixos, seguint el criteri que seguíssim (figures no mostrades).

1.3.5. Comparació dels resultats amb les dades prèviament publicades

La figura 1.5 mostra el motiu d'ancoratge que es va obtenir per DRB1*0301 en aquest treball. Es van considerar com a residus principals aquells que es veien incrementats significativament després de la correcció de Bonferroni. Els residus secundaris eren aquells que estaven augmentats significativament únicament abans de la correcció de Bonferroni.

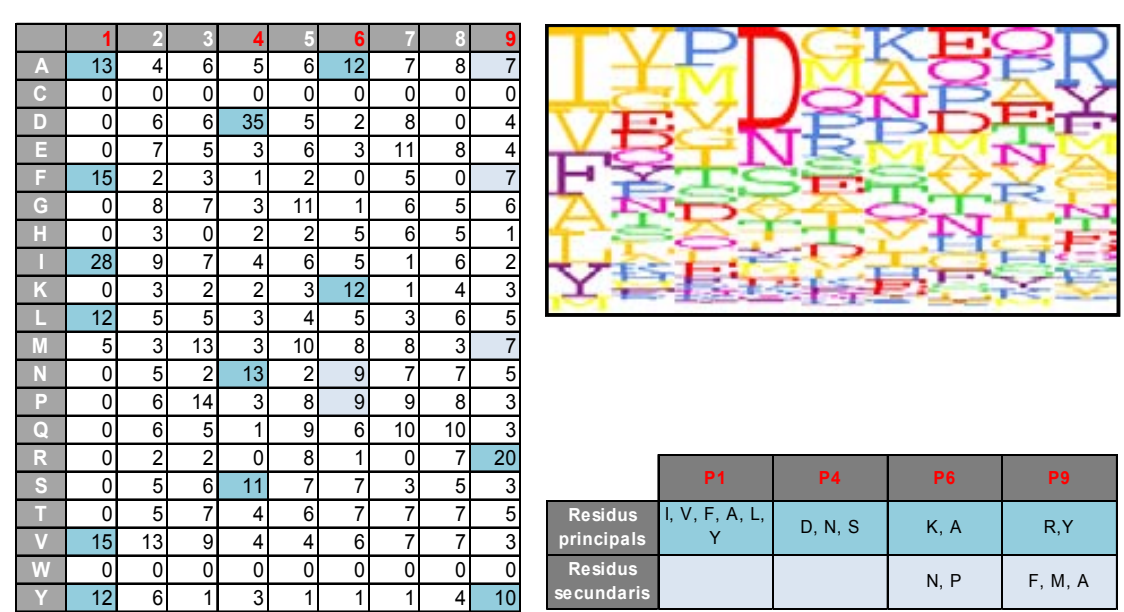


Figura 1.5: Motiu d'ancoratge de DRB1*0301. Dreta: Representació esquemàtica dels percentatges de cada aminoàcid en cada una de les posicions, de P1 a P9, normalitzats respecte la freqüència de cada aminoàcid dins el conjunt del proteoma. Esquerra, dalt: Representació gràfica dels aminoàcids que es troben a cada posició, on la mida es correlaciona amb el grau de percentatge obtingut. Esquerra, baix, motiu d'ancoratge de DR3 on es mostren els residus principals (blau) i secundaris (gris).

A la base de dades SYFPEITHI (<http://www.syfpeithi.de/>) hi ha descrit un motiu d'ancoratge per a DRB1*0301 (figura 1.6), a partir de tres articles publicats entre 1992 i 1994 [103-105]. Aquests treballs es van realitzar mitjançant una anàlisi de seqüències en pool i a partir d'un nombre baix de seqüències.

	P1	P4	P6	P9
Residus principals	L, I, F, M, V	D	K, R	Y, L, F
Residus secundaris			E, Q, N	

Figura 1.6: Motiu d'ancoratge de DRB1*0301 publicat a *Syfpeithi*

A continuació es comparen els resultats que es van obtenir en aquest treball i els resultats anteriorment publicats.

Posició 1 (P1): Els resultats per aquesta posició concorden, i mostren una clara dominància de Isoleucina (I). En el nostre cas no trobem Metionina (M) amb un freqüència significativa, però sí Alanina (A). Per tant els residus permesos a la P1 serien: Ile, Val, Phe, Ala, Leu i Tyr.

Posició 4 (P4): Tots els treballs de caracterització de motius d'ancoratge de variants de DR3 descriuen la posició 4 com una de les més importants. Els nostres resultats mostren preferència clara d'Asp en aquesta posició, d'acord amb el que s'havia afirmat però a diferència del motiu descrit en el SYFPEITHI, també altres aminoàcids són permesos com Asparragina (N) i Serina (S).

Posició 6 (P6): A pesar de no trobar aminoàcids àcids a la cadena β de DR3, els nostres resultats i els de la bibliografia coincideixen en la presència afavorida en P6 de lisina (K). Segons les nostres dades, basades en pèptids naturals, a més estaria permesa l'alanina (A), asparragina (N) i prolina (P).

Posició 9 (P9): En aquesta posició els nostres resultats mostren com a residu principal la tirosina (Y), com també apareixia a la bibliografia (*Syfpeithi*). A la cadena β de DR3, trobem dos residus àcids a la butxaca que uneix residus del pèptid en posició P9. Aquest fet concorda amb la presència d'arginina (R) com

a residu principal en el nostre motiu. A més, també hi estarien permesos la fenilalanina (F), metionina (M) y alanina (A).

1.4. PREDICCIÓ D'AFINITAT TEÒRICA DELS LLIGANDS D'HLA-DR3

El mètode de predicció d'afinitat teòrica dels pèptids a l'al·lel corresponent d'HLA-DR es descriu al capítol de materials i mètodes.

El 44% dels lligands seqüenciats van mostrar una afinitat teòrica alta (HB) per l'al·lel DR3. El 32% i 24% restant van correspondre a lligands amb afinitats teòriques intermèdies (IB) o baixes (LB), respectivament (Figura 1.7). En analitzar les afinitats dels pèptids corresponents a les diferents rutes de degradació, tant els pèptids processats per la via de degradació endocítica com per la citosòlica van presentar un percentatge major de lligands HB (Figura 1.7) però amb diferències. A la via endocítica, les diferències de percentatges eren similar a les de l'afinitat global obtinguda: 21% d'HB, 10% d'IB, 4% de LB. Per la via citosòlica, les diferències entre les tres afinitats eren mínimes: 23% d'HB, 22% d'IB, 20% de LB.

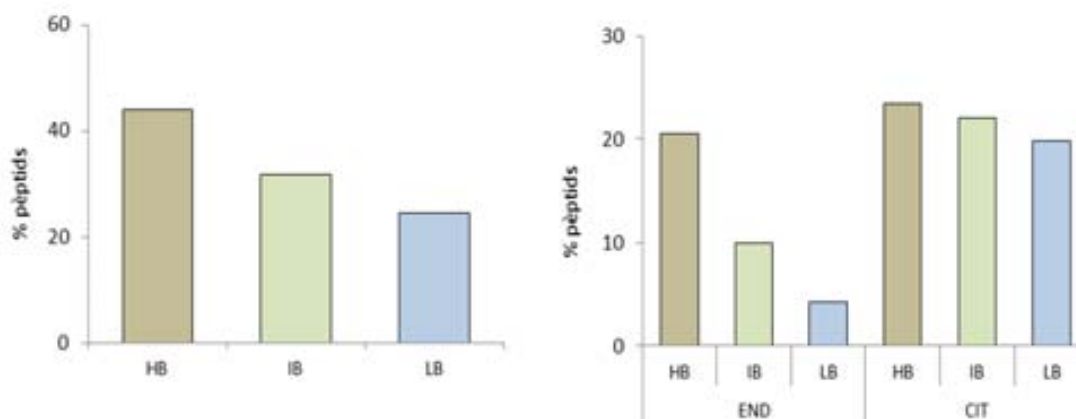


Figura 1.7. Esquerra: Comparació de la predicció d'afinitat d'unió teòrica dels pèptids associats a HLA-DR3. Dreta: Predicció d'afinitat d'unió dels pèptids a la molècula d'HLA-DR en funció de la via de degradació proteica.

1.5 OBSERVACIONS FINALS

Aquest estudi ha permès fer una caracterització acurada del motiu d'ancoratge de la molècula d'HLA-DR3 (DRA1*0101/DRB1*0301) a partir d'un nombre elevat de pèptids naturals, enlloc de pèptids sintètics o de llibreries peptídiques. L'ús de lligands naturals és preferible a l'ús de llibreries peptídiques ja que són productes fisiològics del processament i presentació d'antígens. A més, tenen l'avantatge que han estat seleccionats per les cèl·lules presentadors d'antigen durant el seu processament, on un gran nombre de fenòmens influencien la selecció dels pèptids que seran presentats [108]. S'han identificat un total de 201 lligands naturals de l'al·lel HLA-DR3 mitjançant espectrometria de masses d'alta resolució LTQ-Orbitrap, a partir d'una línia homozigota generada en el nostre laboratori.

Tot i l'elevat polimorfisme de la cadena β d'HLA-DR, la regió compresa entre els residus β 81- β 92 que uneix l'aminoàcid que ocupa la posició 1 del pèptid, no presentava gaires variacions entre HLA-DR1 i HLA-DR3. Per tant, es va poder assumir que aquestes dues molècules unien en posició 1 residus hidrofòbics o aromàtics. Aquesta característica es va utilitzar per determinar els cores dels pèptids, tal i com s'havia fet en treballs anteriors del nostre grup, on es definien els motius d'ancoratge per HLA-DR10 i per HLA-DR8 [90, 112]. A més, es va detectar l'existència de residus bàsics a la butxaca on es col·loca la posició 4 del pèptid (P4): lisina (K) en posició 71 i arginina (R) en posició 74, i a la butxaca on es col·loca la posició 9 del pèptid (P9) trobàvem per contra dos residus àcids. La primera matriu generada fixant sols P1, ja ens donava un percentatge destacat d'àcid aspàrtic en posició 4, i la matriu final ens indicava que aquest era l'aminoàcid principal però no l'únic, com s'havia descrit fins ara [103-105]. D'altra banda, s'ha determinat que el residu amb un percentatge més alt a la posició P9 és l'arginina, fet que concorda amb el caràcter àcid de la butxaca que havíem observat però que no estava descrit al motiu publicat a *Syfypeithi*.

Els nostres resultats no divergeixen en gran mesura amb els anteriorment publicats, però sí que aconsegueixen una descripció més acurada del motiu: la

posició P4, tot i preferir àcid glutàmic, també permet altres residus i per contra, la posició P9 mostra preferència per residus bàsics, fet no descrit fins al moment.

A la literatura trobàvem que els lligands descrits per als al·lels de HLA-DR provenien majoritàriament de proteïnes de membrana o de compartiments membranosos com endosomes, lisosomes, el reticle endoplasmàtic o l'aparell de Golgi [84, 112], característica intrínseca descrita pels repertoris de HLA de classe II. És interessant fer notar que en aquest estudi s'ha identificat la via citosòlica com la via principal de generació de lligands per HLA-DR3. Fins el moment, no s'havia descrit un patró diferent per l'al·lel DR3, ni per cap altre al·lel d'HLA-DR, tot i haver-se descrit el motiu per altres mètodes.

**CAPITOL 2: INFLUÈNCIA DE LES XAPERONES HLA-DM I
CADENA INVARIANT (Ii) EN ELS REPERTORIS PEPTIDICS
PRESENTATS PER HLA-DR3**

2.1. Processament de les mostres

L'aïllament del repertori peptídic associat a les quatre línies transfectades endocrines RIN-m5F es va fer en dos experiments consecutius, seguint el mateix protocol descrit al capítol 1 i partint de 1000 milions de cèl·lules per cada preparació.

2.2. Repertoris de pèptids associats a HLA-DR3 en cèl·lules transfectades RIN-m5F: DR3, DR3-Ii, DR3-DM, DR3-DM-Ii

Es van seguir els mateixos passos detallats al capítol 1 per a la identificació i validació dels espectres obtinguts, utilitzant el sistema de recerca *Protein Discover* v1.4, (ThermoFisher, San Jose, CA, USA). En aquest cas, el FDR utilitzat fou de 0'1%. Un resum de les seqüències obtingudes es mostra a la taula 2.1.

Mostra	Espectres identificats	Seqüències no redundants	Seqüències finals	Proteïnes(*)	% nesteds
RINDR3	160	102	90	76	16
RINDR3DM	136	98	82	51	38
RINDR3Ii	78	65	60	44	27
RINDR3DMIi	430	198	176	87	51

(*) Proteïnes: s'indica el nombre de grups de proteïnes calculat per Sequest (ver material i mètodes)

Taula 2.1: Resum del nombre de seqüències obtingudes per a les quatre línies transfectants.

2.2.1 RINDR3

A les cèl·lules transfectades amb DR3 en absència de les dues xaperones (RINDR3), es van identificar 90 lligands diferents de longitud entre 9 i 25 aminoàcids, corresponents a 76 proteïnes diferents. A la taula 2.2, es mostra el llistat de les seqüències identificades de l'eluït de RINDR3 i un resum de les

seues característiques. El percentatge de *nested sets* en aquesta línia era baix (16%), comparat amb altres cèl·lules estudiades. La massa mitjana dels pèptids era de 1431 Da.

La distribució de la mida era més heterogènia de l'esperat per a MHC-II, amb dos pics màxims de 13 i 16 aminoàcids, tal i com es pot observar a la figura 2.1 on es compara amb la mida de les altres línies RIN.

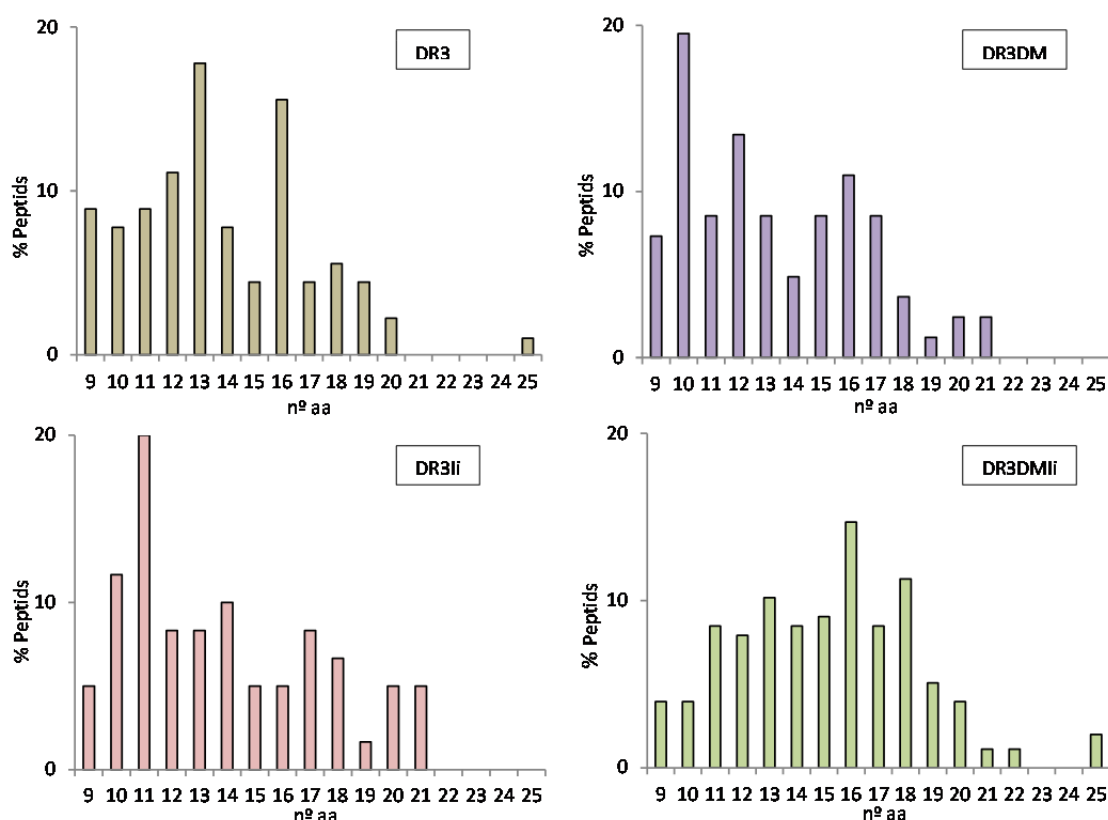


Figura 2.1: Comparació de les mides dels pèptids corresponents a les quatre línies RIN transfectades: RINDR3 (marró), RINDR3DM (blau), RINDR3li (rosa) i RINDR3DMli (verd).

Un 71% de les proteïnes eren residents a la via citosòlica (figura 2.7A), com ara l'actina citoplasmàtica, la miosina, la calmodulina, les histones 1.4 i 4 i enzims com la triosefosfat isomerasa, glutamina sintetasa, Map3k9 (quinasa) o la disulfid isomerasa A6. El repertori associat a RINDR3, va generar un pèptid de la proteïna HLA-DR (IQAEFYLNPDQSGEF) que també es va aïllar a la línia RINDR3DMli. També es va trobar un pèptid de la proteïna A1h, que forma part de la família de molècules altament conservades de MHC classe I no clàssic de rata.

Taula 2.2: Seqüències dels pèptids associats a classe II de la mostra RINDR3

Uniprot AN	Proteïna	Seqüència	Tamany	Localització	Terminació
P19945	60S acidic ribosomal protein P0	TTAAPAAAAAPAK	13	C	C-ter
P62902	60S ribosomal protein L31	KNLQTVNVN	11	C	C-ter
P62902	60S ribosomal protein L31	YVPVTFKLNQTVNVN	18		
P62902	60S ribosomal protein L31	VTYVPVTFKLNQTVNVN	20		
Q9YQ3	A1h	QLMFGCDLGS DGSLLR	16	MB	m
P60711	Actin cytoplasmic 1	EITALAPSTMK	11	C	C-ter
P11030	Acyl-CoA-binding protein	QATVGDVNTDRPGLDLK	18	END	m
Q63495	Advanced glycosylation end product-specific receptor	RVRVYQIPGKP	11	MB	m
F1M3T0	ARF GTPase-activating protein GIT1	KAAKQLVTITRE	13	C	C-ter
P21571	ATP synthase-coupling factor 6	NFEDPKFEVLDPQS	15	MIT	C-ter
Q641Y8	ATP-dependent RNA helicase DDX1	LHLGYLPNQLF	11	C	C-ter
D4ABV5	Calmodulin	IREADIDGGDQVN	13	C	C-ter
D4ABV5	Calmodulin	MIREADIDGGDQVN	14		
P11240	Cytochrome c oxidase subunit 5A	ISTPEELGLDKV	12	MIT	C-ter
P11240	Cytochrome c oxidase subunit 5A	GISTPEELGLDKV	13		
P11240	Cytochrome c oxidase subunit 5A	LGISTPEELGLDKV	14		
P11240	Cytochrome c oxidase subunit 5A	NELGISTPEELGLDKV	16		
Q9QXU8	Cytoplasmic dynein 1 light intermediate chain 1	GGGAAGASTSLPP	13	C	C-ter
Q6AYI1	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	AAAPMIGYPMPTGYSQ	16	N	C-ter
Q4QQR7	Dsp protein	SEISDPDDTKG	13	MB	m
Q5EB83	Eftud2 protein	LELAKQDVVLNYPM	14	C	C-ter
P62630	Elongation factor 1-alpha 1	GIGTVPVGRVE	11	N	m
B5DEN5	Eukaryotic translation elongation factor 1 beta 2	VQSM DVAAFNKI	12	N	N-ter
B5DEN5	Eukaryotic translation elongation factor 1 beta 2	GFGDLKTPAGLQVLND	16		C-ter
B5DF22	Fam53a protein	PRKPVSTG	9	N	m
P09606	Glutamine synthetase	LNETGDEPFQYKN	13	C	C-ter
P14659	Heat shock-related 70 kDa protein 2	KCQEVINWLDNRNQMAEKDEYEHKQKELERV	30	C	C-ter
D4A962	Heterogeneous nuclear ribonucleoprotein U-like 1	DYGGYSGTQGGTSTQ	16	N	C-ter
P15865	Histone H1.4	ASGPPVSELITK	12	N	N-ter
P62804	Histone H4	ISGLIYEETR	10	N	m
P62804	Histone H4	DNIQGITKPAIR	12		m
P01903	HLA class II histocompatibility antigen, DR alpha chain	IQAIFYLNPDQSGEF	15	MB	N-ter
O70592	HnRNP protein	YSGSGSGSGYGSR	13	N	C-ter
B2GV69	Hnrnpa2b1 protein	SGNFGGSRNMGGP	13	N	C-ter
P97519	Hydroxymethylglutaryl-CoA lyase, mitochondrial	MLTAVLHEV	9	MIT	m
G3V8L7	Integrin alpha M	VGGLVLLAL	9	MB	C-ter
M0RCF3	Integrin beta	AILLIGLATLLIWKLITI	19	MB	C-ter
Q45GW3	IQ and ubiquitin-like domain-containing protein	DEVVSLGSATDSHEQE	16	NA	m
P11497-2	Isoform 2 of Acetyl-CoA carboxylase 1	RKKDLVKTMRR	11	C	m
Q498T2-2	Isoform 2 of Chromatin target of PRMT1 protein	YMAQTD PETND	11	N	C-ter
Q498T2-2	Isoform 2 of Chromatin target of PRMT1 protein	DAYMAQTD PETND	13		
P02688-2	Isoform 2 of Myelin basic protein	GKGRGLSLSRFSWGA	15	MB	m
Q5M9H1	Leucine-rich repeat-containing protein 41	MEATSREAPAKS	13	C	N-ter
Q62667	Major vault protein	FGLLGLSDGQPPAQK	16	C	C-ter
Q62667	Major vault protein	FSTAFGLLGLSDGQPPAQK	20		
Q0QF43	Malate dehydrogenase	AKVAVLGASGGIGQPLSL	18	C	N-ter
Q0QF43	Malate dehydrogenase	AKVAVLGASGGIGQPLSLL	19		
F1LMU0	Myosin-4	QSKIEDEQALGM	12	C	m
O09017	Nuclear receptor subfamily 2 group F member 6	FGRLLRLPALRAVPASL	18	N	Cter
D4AA57	Olfactory receptor	LVVASVIGLFLS	13	MB	m
Q4V8F6	Pcbp2 protein	VRLSSETGGMGSS	13	N	C-ter
P01194	Pro-opiomelanocortin	YGGFMTSEKSQTPLVT	16	S	C-ter
Q6AYR1	Protein 100911212	GQQPTSQAPAPA	12	C	C-ter
D4ABZ3	Protein AatK	VLSPSAAGPAGGQPRAMDS	19	MB	m
E9PU15	Protein Dicer1	CDPDELEEKIKLE	14	ER	N-ter
Q63081	Protein disulfide-isomerase A6	IDLSVDVLDDEKDEL	16	ER	C-ter

Uniprot AN	Proteïna	Seqüència	Tamany	Localització	Terminació
D3Z914	Protein Inip	PVLPRLDPE	9	N	C-ter
M0R8J3	Protein Jmy	SQQPSEIKTQ	10	N	m
Q6IFW7	Protein Krt28	NARLAADDF	9	C	m
Q6IFW7	Protein Krt28	LQIDNARLAADDF	13		
D3Z8D7	Protein LOC100361854	KRRNNGRAKK	10	N	N-ter
D4A6G6	Protein LOC100362339	PGVTVKDVNQQE	12	N	N-ter
D3ZFA8	Protein LOC100362366	QVTQPTVGMNFKTPRGAV	18	N	C-ter
F1LRK5	Protein LOC311352	EAGSVDQVNF	10	NA	m
D3ZE63	Protein LOC679748	NAANVGWNGSTFA	13	NA	C-ter
D3ZE63	Protein LOC679748	YDMNAANVGWNGSTFA	16		
D4A510	Protein LOC685179	PTAPSPGTVTPVPPPQ	16	N	C-ter
F1LRA7	Protein Map3k9	DGYKQWSSAPN	12	C	m
D3ZT14	Protein Pikfyve	MLADYRARGGRIQSK	15	END	m
D3ZTN0	Protein Prkdc	LDQLSNPDM	9	N	m
M0R5Y8	Protein RGD1565048	VSEKGTVQQPDE	12	MIT	C-ter
B2RZB7	Protein Snrpd1	FILPDSLPLDTL	12	N	m
D4A4T8	Protein Socs7	RSQGITHHTRMEHY	14	C	m
F1M8W3	Protein Sp3	VTVSGNETME	10	N	C-ter
D3ZSS1	Protein Taf15	VQGLGEGVSTDQVGEF	16	C	m
D3ZJS3	Protein Tomm6	LARNLSDIDLMAQPGV	17	MIT	C-ter
D4A175	Protein Tri1	VDDMLAAGL	9	MIT	m
M0RCY7	Protein Ufsp1	CELQAAGWVGWQKVNH	16	C	C-ter
Q672J3	Protein Zbtb1	VGVDERRGLE	10	N	m
Q5BJV1	Rasgrf1 protein	AIFRLKKTWLKVSK	14	C	m
Q5XI66	Ss18 protein	DQGQYGNVQQ	10	N	C-ter
Q63564	Synaptic vesicle glycoprotein 2B	RLISGIGIGSLP	13	C	m
D3ZUW7	Synaptonemal complex protein 2	ESKKLLTLTKNIVKISKK	19	N	m
P28480	T-complex protein 1 subunit alpha	AMAVRRVLKRDLEKRIAK	17	C	m
Q920J4	Thioredoxin-like protein 1	GSDPDFQPELSGAGSR	16	C	N-ter
P62329	Thymosin beta-4	PLPSKETIEQEKQAGES	17	C	C-ter
Q69BT7	Trafficking protein particle complex subunit 4	NKLALAEVAKAGT	14	END	C-ter
P48500	Triosephosphate isomerase	FVDIINAKQ	9	C	C-ter
P48500	Triosephosphate isomerase	LVGGASLKPEF	11		C-ter
F1M0E9	Uncharacterized protein	PPAENSSAPEAQGGAE	17	N	C-ter

Uniprot AN: Identificació de la proteïna en SwissProt; Proteïna: proteïna origen de la qual prové el pèptid; Longitud: mida del pèptid en nombre d'aminoàcids; Localització: MB (membrana plasmàtica); Lis (lisosomes); End (endosomes); S (secretada); EM (matriu extracel·lular); ER (Reticle endoplàsmic); G (Golgi); MIT (mitochondria) C (citòsol); N (nuclear)

2.2.2. RINDR3DM

Per a la línia RINDR3DM, es van identificar 82 lligands diferents corresponents a 51 proteïnes (taula 2.3). El percentatge de *nested sets* en aquesta línia fou del 38% (el més alt després de la triple transfectant). La distribució de la mida dels pèptids presentava un alt nombre de pèptids petits, on més d'un 50% del total de les seqüències tenien una mida d'entre 9-13 aminoàcids (figura 2.1). De fet, la mida màxima era de 10 aminoàcids (20% del total). La massa mitjana dels pèptids per a DR3DM era de 1300 Da.

El 80% del repertori corresponia novament a proteïnes residents al citosol i associades a nucli. Dins d'aquesta via, s'han identificat fins a 15 seqüències corresponents a sis tipus d'histones diferents. Altres proteïnes com la vimentina, l'actina citoplasmàtica i diferents enzims també formaven part del repertori associat al citoplasma/nucli (figura 2.7B). Al repertori de la línia DR3DM no es van trobar pèptids procedents de proteïnes de l'MHC o associades a la immunitat.

Taula 2.3: Seqüències dels pèptids associats a classe II de la mostra RINDR3DM

Uniprot AN	Proteïna	Seqüència	Tamany	Localització	Terminació
P04644	40S ribosomal protein S17	IIEVDPDTKEM	11	C	m
P62853	40S ribosomal protein S25	YTRNTKGGDAPAGEDA	17	C	C-ter
P20280	60S ribosomal protein L21	KEPELLEPIPYEFMA	15	C	C-ter
P20280	60S ribosomal protein L21	GKEPELLEPIPYEFMA	16		
P62752	60S ribosomal protein L23a	ALDVANKIGII	11	N	C-ter
P62902	60S ribosomal protein L31	KNLQTVNVN DEN	11	C	C-ter
P62902	60S ribosomal protein L31	VTYVPVTTFKNL	12		
P62902	60S ribosomal protein L31	YVPVTTFKNLQTVNVN DEN	18		
P62902	60S ribosomal protein L31	VTYVPVTTFKNLQTVNVN DEN	20		
P62902	60S ribosomal protein L31	LVTVVPVTTFKNLQTVNVN DEN	21		
P60711	Actin, cytoplasmic 1	AGFAGDDAPR	10	C	N-ter
P60711	Actin, cytoplasmic 1	RVAPEEHPVL	10		m
P60711	Actin, cytoplasmic 1	EITALAPSTMK	11		m
P60711	Actin, cytoplasmic 1	SYELPDGQVITIGNER	16		m
P60711	Actin, cytoplasmic 1	IVMDSGDGVTHTVPIYEG	18		m
P60711	Actin, cytoplasmic 1	DLYANTVLSGGTTMYPGIADR	21		m
Q09073	ADP/ATP translocase 2	TDAAVSFAKD	10	MB	N-ter
P04764	Alpha-enolase	NVINGGSHAGNK	12	C	m
P21571	ATP synthase-coupling factor 6, mitochondrial	PKFEVLDKPQS	11	MIT	C-ter
P62161	Calmodulin	IREADIDGDGQVN	13	C	C-ter
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	GISTPEELGLDKV	13	MIT	C-ter
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	NELGISTPEELGLDKV	16		
P62630	Elongation factor 1-alpha 1	FSDYPPLGRF	10	C	C-ter
P09606	Glutamine synthetase	LNETGDEPFQYKN	13	C	C-ter
Q07439	Heat shock 70 kDa protein 1A/1B	ILMGDKSENVQDL	13	N	m
P15865	Histone H1.4	VQTKGTGASGSF	12	N	m
P02262	Histone H2A type 1	LGKVTIAQGGVL	12	N	C-ter
P02262	Histone H2A type 1	LGKVTIAQGGVLPNIQ	16		
P02262	Histone H2A type 1	LGKVTIAQGGVLPNIQA	17		
A9UMV8	Histone H2A.J	LGRVTIAQGGVLPNIQ	16	N	C-ter
A9UMV8	Histone H2A.J	LGRVTIAQGGVLPNIQA	17		
A9UMV8	Histone H2A.J	VTIAQGGVLPNIQAVLLPK	19		
A9UMV8	Histone H2A.J	VLLPKKTES	9		C-ter
A9UMV8	Histone H2A.J	AVLLPKKTES	10		
A9UMV8	Histone H2A.J	AGLQFPVGR	9		N-ter
Q00715	Histone H2B type 1	AMGIMNSFVNDIFER	15		m
P84245	Histone H3.3	IAQDFKTDL	9	N	m
P84245	Histone H3.3	LVGLFEDTNL	10		m
P84245	Histone H3.3	RFQSAAGALQEA	13		m
P62804	Histone H4	TVTAMDVVYALK	12	N	C-ter

Uniprot AN	Proteïna	Seqüència	Tamany	Localització	Terminació
D3ZBS2	Inter-alpha-trypsin inhibitor heavy chain H3	TPLTSMVVTKE	12	S	C-ter
D3ZBS2	Inter-alpha-trypsin inhibitor heavy chain H3	FVTPLTSMVVTKEPDN	16		
B0LPN4	Ryanodine receptor 2	GKVVEPDMSAGF	12	MB	m
Q9R1R4	Tudor domain-containing protein 7	KNMAESADDFVED	13	C	m
D3ZJ96	Ubiquitin carboxyl-terminal hydrolase 28	NHWCSYLKGDIAENL	15	N	C-ter*
Q9ERC1	Isoform 2 of Unconventional myosin-XVI	SMLTDVRHFLSSGGDVN	17	C	m
Q4AE70	Histone-arginine methyltransferase CARM1	MASPMISPTNTMHYGS	16	N	c-ter
P09951-2	Isoform IB of Synapsin-1	AAPVASPAAP	10	G	N-ter
Q00438-2	Isoform PYBP1 of Polypyrimidine tract-binding protein 1	LVSNLNPERVTPQSL	15	N	m
O88204	Low-density lipoprotein receptor-related protein 3	PGPVPDPVP	10	MB	m
O88954	MAGUK p55 subfamily member 3	RAVLNVHDTVAQKNFDP	17	NA	m
P04636	Malate dehydrogenase, mitochondrial	YDIAHTPGVA	10	MIT	m
P04636	Malate dehydrogenase, mitochondrial	AKVAVLGASGGIGQPLS	17		
P04636	Malate dehydrogenase, mitochondrial	AKVAVLGASGGIGQPLSL	18		N-ter
Q5FVM4	Non-POU domain-containing octamer-binding protein	FSVFGQVERA	10	N	m
Q63525	Nuclear migration protein nudC	FIGGEEGMAEKL	12	N	m
P62961	Nuclease-sensitive element-binding protein 1	PPAENSAPAEQEG	14	N	c-ter
P62961	Nuclease-sensitive element-binding protein 1	PPAENSAPAEQGG	15		
P62961	Nuclease-sensitive element-binding protein 1	PPAENSAPAEQGGAE	17		
P16617	Phosphoglycerate kinase 1	LEGKVLPGVDALSNV	15	C	C-ter
Q63429	Polyubiquitin-C	MQIFVKTLTG	10	C	N-ter
Q6NX65	Programmed cell death protein 10	ILQTFKTVA	9	C	C-ter
P01194	Pro-opiomelanocortin	YGGFMTSEKSQTPLVT	16	EM	C-ter
D4AB70	Protein Edem3	DLNQLQDQ	9	ER	C-ter
G3V829	Protein Fubp3	FYGQTLGQAQHSQEQ	16	C	C-ter
D3ZSP1	Protein LOC100361838	GDNITLLQSVSN	12	N	C-ter
D4A6G6	Protein LOC100362339	PGVTVKDVNQEQ	12	N	N-ter
D3ZJT4	Protein RGD1564469	TTAAPAAAAAPAK	13	N	C-ter
Q63528	Replication protein A 32 kDa subunit	STVDDHFKSTDAE	14	N	C-ter
Q9JHY2	Sideroflexin-3	FTVTDPRNLL	10	N	N-ter
Q5XI66	Ss18 protein	YDQGGYGYNYQQ	11	N	C-ter
P63281	SUMO-conjugating enzyme UBC9	VAVPTKNPDGTMNL	14	N	C-ter
Q7TPB1	T-complex protein 1 subunit delta	KIDDVVNTR	9	C	C-ter
Q5BK48	Tetratricopeptide repeat protein 5	WGVLLIGDSVA	10	N	C-ter*
Q5XFX0	Transgelin-2	SQAGMTGYGMPRQIL	15	S	C-ter
P48500	Triosephosphate isomerase	NAAKLPADTE	10	C	m
P48500	Triosephosphate isomerase	VGGASLKPEF	10		c-ter
P48500	Triosephosphate isomerase	LVGGASLKPEF	11		
P48500	Triosephosphate isomerase	FLVGGASLKPEF	12		
P48500	Triosephosphate isomerase	ASQPDVDGFLVGGASLKPEF	20		
P31000	Vimentin	FADLSEAANR	10	C	m
G3V8C3	Vimentin	YVTTSTRYSLGSA	14		N-ter

Uniprot AN: Identificació de la proteïna en SwissProt; Proteïna: proteïna origen de la qual prové el pèptid; Longitud: mida del pèptid en nombre d'aminoàcids; Localització: MB (membrana plasmàtica); Lis (lisosomes); End (endosomes); S (secretada); EM (matriu extracel·lular); ER (Reticle endoplàsmic); G (Golgi); MIT (mitochondria) C (citòsol); N (nuclear)

2.2.3. RINDR3li

Per a la línia RINDR3li, es van identificar 60 lligands diferents corresponents a 44 proteïnes diferents (taula 2.4). La distribució de la mida també es desviava de la distribució habitual, presentant un alt nombre de pèptids petits, i novament

un 50% del total de les seqüències tenien una mida d'entre 9-13 aminoàcids (figura 2.1). En aquest cas, la mida més freqüent fou de 11 aminoàcids (20% del total). Un 27% dels pèptids s'agrupaven en *nested sets*. La massa mitjana dels pèptids era de 1446 Da.

Un 81% de les proteïnes formaven part de la via citosòlica, el percentatge més alt de les quatre línies transfectades (figura 2.7C). Dintre d'aquesta via trobem l'actina, l'anexina, miosina, histones i diferents enzims. Només 8 proteïnes de les 44 proteïnes s'associaven a la via endocítica, cap d'elles associada a MHC. Sorprenentment, no es va identificar cap pèptid de la cadena invariant (li, CD74).

Taula 2.4: Seqüències dels pèptids associats a classe II de la mostra RINDR3li

Uniprot AN	Proteïna	Seqüència	Tamany	Localització	Terminació
P17074	40S ribosomal protein S19	PGVTVKDVNQQE	12	N	N-ter
P62752	60S ribosomal protein L23a	ALDVANKIGII	11	N	C-ter
P62890	60S ribosomal protein L30	PGDSDIIRSMPEQTGEK	17	N	C-ter
P62902	60S ribosomal protein L31	KNLQTVNVNVDEN	11	N	C-ter
P62902	60S ribosomal protein L31	YVPVTTFKNLQTVNVNVDEN	18		
P62902	60S ribosomal protein L31	VTYVPVTTFKNLQTVNVNVDEN	20		
P62902	60S ribosomal protein L31	LVTYVPVTTFKNLQTVNVNVDEN	21		
P60711	Actin, cytoplasmic 1	AGFAGDDAPR	10	C	N-ter
P60711	Actin, cytoplasmic 1	EITALAPSTMK	11		m
P60711	Actin, cytoplasmic 1	VAPEEHPVLLTEAPLNPK	18		m
P07150	Annexin A1	TPAQFDADELRL	11	N	m
P21571	ATP synthase-coupling factor 6	PKFEVLDPKQS	11	N	m
A120K8	Beta-actin	ATAASSSSLEKS	12	N	C-ter
P70645	Bleomycin hydrolase	AQNVGTTHDLL	11	N	m
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	ISTPEELGDKV	12	MIT	C-ter
P11240	Cytochrome c oxidase subunit 5A	GISTPEELGDKV	13		
P11240	Cytochrome c oxidase subunit 5A	NELGISTPEELGDKV	16		
P62630	Elongation factor 1-alpha 1	GIGTVPVGRVE	11	C	m
P62630	Elongation factor 1-alpha 1	IGGIGTVPVGR	11		
C0JPT7	Filamin alpha	AEIVEGENHTY	11	C	m
P09606	Glutamine synthetase	LNETGDEPFQYKN	13	C	C-ter
A9UMV8	Histone H2A type 1	VLLPKKTES	9	N	C-ter
A9UMV8	Histone H2A type 1	AGLQFPVGR	9		N-ter
P0C0S7	Histone H2A.Z	GDEELDSLK	10	N	m
P62804	Histone H4	ISGLIYEETR	10	N	m
Q6P6G9	Hnrpa1 protein	SSGPYGGGGQYFAKP	15	N	C-ter
Q6P6G9	Hnrpa1 protein	NQGGYGGSSSSSYGSGR	18		
P53565	Homeobox protein cut-like 1	IVPMAKPAKPSVPP	14	N	m
Q5BJX3	Hs2st1 protein	QSDIWKME	10	MB	C-ter
D3ZBS2	Inter-alpha-trypsin inhibitor heavy chain H3	FVTPLTSMVVTKE	14	S	C-ter
P63095	Guanine nucleotide-binding protein G(s) subunit alpha isoforms short	ILLGAGESGK	11	MB	N-ter
B6CZ61	Leucine-rich repeat-containing protein 51	PVLTFFNLSVLYL	14	C	m
B5DEJ9	LOC691036 protein	MVNLLVPLDTSKN	13	MB	m
G3V9Y1	Myosin, heavy polypeptide 10	TESKTSVDNETQPPQSE	17	C	C-ter
G3V9Y1	Myosin, heavy polypeptide 10	DTEKTSVDNETQPPQSE	18		
G3V9Y1	Myosin, heavy polypeptide 10	DDTESKTSVDNETQPPQSE	19		
G3V9Y1	Myosin, heavy polypeptide 10	DDTESKTSVDNETQPPQSE	20		
G3V9Y1	Myosin, heavy polypeptide 10	SDDTESKTSVDNETQPPQSE	21		

Uniprot AN	Proteïna	Seqüència	Tamany	Localització	Terminació
Q5FVM4	Non-POU domain-containing octamer-binding protein	FSVFGQVERA	10	N	m
P62961	Nuclease-sensitive element-binding protein 1	PPAENSSAPEAEQGG	15	N	C-ter
P62961	Nuclease-sensitive element-binding protein 1	PPAENSSAPEAEQGGAE	17		
Q6AXS5	Plasminogen activator inhibitor 1 RNA-binding protein	TDKSSASAPDVEDPEAFPALA	21	C	C-ter
Q9EPH8	Polyadenylate-binding protein 1	YVGDLHPDVTEAM	13	N	N-ter
P01194	Pro-opiomelanocortin	YGGFMTSEKSTPLVT	16	S	C-ter
D3ZEV0	Protein LOC100912427	AADPPAENSSAPEAEQGGAE	20	N	N-ter
D3ZTH8	Protein LOC689899	LAPDYDALDVANKIGII	17	C	N-ter
B2RZ53	Protein Rab36	VSAKTGENVK	10	C	m
B5DE50	Protein Snrpd2	SVIVVLRNPLIAGK	14	NA	C-ter
B5DE50	Protein Snrpd2	GDSVIVVLRNPLIAGK	16		
M0R7E6	Protein Srrt	YRDLAPDDVDFF	13	C	C-ter
Q5U1Z0	Rab3 GTPase-activating protein non-catalytic subunit	ASQLLVLTGQRL	12	C	m
Q4V898	RNA-binding motif protein, X chromosome	SAPSGPVRSSSGMGG	15	N	m
Q9R050	Single-stranded DNA-binding protein 3	NYSPSMTMSV	10	N	C-ter
Q9JJ09	Sodium-dependent phosphate transport protein 2B	TKLIQLDKKVI	12	C	m
P48500	Triosephosphate isomerase	LVGGASLKPEF	11	C	C-ter
F1LUE4	Uncharacterized protein	VSVQRTQAPAVATT	14	N	C-ter
P0CD95	Uncharacterized protein C17orf67	ALTLLAFSSEASPILTE	17	S	N-ter
G3V8C3	Vimentin	YVTSTRTYSLGSA	14	C	N-ter
F1LMC2	Xin actin-binding repeat-containing protein 2	DMGDRRGVWTE	11	N	m

Uniprot AN: Identificació de la proteïna en SwissProt; Proteïna: proteïna origen de la qual prové el pèptid; Longitud: mida del pèptid en nombre d'aminoàcids; Localització: MB (membrana plasmàtica); Lis (lisosomes); End (endosomes); S (secretada); EM (matriu extracel·lular); ER (Reticle endoplàsmic); G (Golgi); MIT (mitochondria) C (citòsol); N (nuclear)

2.2.4 RINDR3DMli

La línia RINDR3DMli és la que va generar un major nombre d'identificacions. Es van identificar un total de 176 lligands diferents corresponents a 91 proteïnes (taula 2.5). A més, un 48% dels pèptids s'agrupaven en *nested sets*. Aquesta línia va mostrar una distribució irregular de les mides dels pèptids, amb un pic del 15% corresponent a 16 aminoàcids (figura 2.1). La majoria de pèptids estaven dins d'un rang entre 9 a 22 aminoàcids, tot i que es van identificar 4 pèptids amb 25 aminoàcids. Aquest patró és semblant a les RINDR3. La massa mitjana fou de 1595 Da.

En aquest cas, només el 67% de les proteïnes es localitzaven a la ruta citosòlica (figura 2.7D). Entre les proteïnes d'origen citosòlic/nuclear vam trobar un alt nombre de ribonucleoproteïnes, l'actina, l'arginasa, histones i enzims diversos (disulfid isomerasa, malat deshidrogenasa).

Dins de les proteïnes associades a la via endocítica, vam trobar HLA-DR, concretament 7 pèptids diferents corresponents a 2 grups de *nested sets* diferents. A més, només en aquesta línia vam trobar pèptids que corresponien

a la cadena invariant CD74, agrupats també en 2 *nested sets* diferents: un grup contenia part de la seqüència corresponent a la cadena invariant que es situa a la part interna de la proteïna (4 pèptids) i que coincidia amb CLIP, i l'altre grup es localitzava a la part C-terminal de la proteïna (5 pèptids) (figura 2.5).

HLA class II histocompatibility antigen gamma chain CD74	
LEDPSSGLGVTKQDLGPVPM	RMATPLLMQALPMGALPQ
EDPSSGLGVTKQDLGPVPM	RMATPLLMQALPMGALP
DPSSGLGVTKQDLGPVPM	RMATPLLMQALPM
PSSGLGVTKQDLGPVPM	MATPLLMQALPMGALP
VTQDLGPVPM	

Figura 2.5: Pèptids aïllats corresponents a l'antigen CD74. En color verd: *nested set* compost per 5 pèptids diferents que es troben a la part C-terminal de la proteïna. En color gris: *nested set* compost per 4 pèptids diferents que formen part de la seqüència d'li i es localitzaven a la part interna de la proteïna.

Taula 2.5: Seqüències dels pèptids associats a classe II de la mostra RINDR3li

Uniprot AN	Proteïna	Seqüència	Tamany	Localització	Terminació
P62752	60S ribosomal protein L23a	YDALDVANKIGII	13	N	C-ter
P60711	Actin, cytoplasmic 1	AGFAGDDAPR	10	C	N-ter
P60711	Actin, cytoplasmic 1	VAPEEHPVLLTEAPLNPK	18		m
P11030	Acyl-CoA-binding protein	QATVGDVNTDRPGLDLK	18	ER	m
P04764	Alpha-enolase	NVINGGSHAGNK	12	C	m
P04764	Alpha-enolase	YGKDATNVGDEGG	13		m
Q9JHZ1	Androgen receptor-related apoptosis-associated protein CBL27	SPAPLLPGDLLGGGSDSI	18	MB	C-ter
P07824	Arginase-1	RDVDPGEHY	9	C	m
Q4V8H5	Aspartyl aminopeptidase	ELFPSVSRNLLVD	13	C	C-ter
D4A3E3	AT rich interactive domain 1A	QQQPQTQQPQLQ	12	N	N-ter
P29419	ATP synthase subunit e, mitochondrial	VPPVQVSPLIKFGRYSAL	18	MIT	N-ter
Q6PDU7	ATP synthase subunit g, mitochondrial	YIGEIIKRGIVGYDV	16	MIT	C-ter
Q06647	ATP synthase subunit O, mitochondrial	FSKLVRRPPVQVY	12	MIT	N-ter
P21571	ATP synthase-coupling factor 6, mitochondrial	NFEDPKFEVLDKPQS	15	MIT	C-ter
Q641Y8	ATP-dependent RNA helicase DDX1	GYLPNQLFRTF	11	N	C-ter
Q8R2H3	Barttin	YPKITFVPADS	11	MB	N-ter
P08081	Clathrin light chain A	SVLISLKQAPLVH	13	C	C-ter
P08082	Clathrin light chain B	SVLMSLKQTPLSR	13	C	C-ter
Q498T2	Chromatin target of PRMT1 protein	YMAQTDPEPND	11	N	C-ter
Q498T2	Chromatin target of PRMT1 protein	AYMAQTDPEPND	11		
Q498T2	Chromatin target of PRMT1 protein	LDAYMAQTDPEPND	14		
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	ISTPEELGLDKV	12	MIT	C-ter
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	GISTPEELGLDKV	13		
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	GISTPEELGLDKVK	14		
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	ELGISTPEELGLDKV	15		
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	NELGISTPEELGLDKV	16		
P11240	Cytochrome c oxidase subunit 5A, mitochondrial	ELRPTLNELGISTPEELGLDKV	22		
Q6AYI1	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	FGGSRTGPLS	10	N	N-ter
Q6AYI1	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	AAAPMIGYPMPTGYSQ	16		C-ter

Uniprot AN	Proteïna	Seqüència	Tamany	Localització	Terminació
P63036	DnaJ homolog subfamily A member 1	EPGDIIIIVLDQKDHAVFT	18	N	m
Q4FZY0	EF-hand domain-containing protein D2	MELKLMMKLG	12	MB	m
F7EP67	Endothelin-converting enzyme 1	YPNFIMDPKELDKVFN	16	MB	m
B5DEN5	Eukaryotic translation elongation factor 1 beta 2	VQSM DVAAFNKI	12	N	C-ter
B5DEN5	Eukaryotic translation elongation factor 1 beta 2	GFGLKTPAGLQVLND	16		N-ter
Q7TSL3	F-box only protein 11	DSAPPIESNTLQHN	14	N	C-ter
Q5PQK2	Fusion, derived from t(1216) malignant liposarcoma	YGGQQQSYGQQQ	12	C	N-ter
Q5PQK2	Fusion, derived from t(1216) malignant liposarcoma	GAYPTQPGQGSQQ	14		
Q5PQK2	Fusion, derived from t(1216) malignant liposarcoma	SGYGGQQSYGQQQ	14		
Q5PQK2	Fusion, derived from t(1216) malignant liposarcoma	SSYGGQQSGGYGQQ	14		
Q5PQK2	Fusion, derived from t(1216) malignant liposarcoma	SYGAYPTQPGQGSQQ	16		
Q5PQK2	Fusion, derived from t(1216) malignant liposarcoma	ATQSYGAYPTQPGQGSQQ	19		
Q04827	G1/S-specific cyclin-D2	ATDFKFAMY	9	N	m
P09606	Glutamine synthetase	LNETGDEPFQYKN	13	C	C-ter
P04797	Glyceraldehyde-3-phosphate dehydrogenase	VMGVNHEKY	9	C	m
P62828	GTP-binding nuclear protein Ran	APPEVMDPALAAQYE	16	N	C-ter
P62828	GTP-binding nuclear protein Ran	APPEVMDPALAAQYEH	17		
P82995	Heat shock protein HSP 90-alpha	EMPPLEGDDTSRMEVD	18	C	C-ter
P82995	Heat shock protein HSP 90-alpha	EEMPPLEGDDTSRMEVD	19		
P14659	Heat shock-related 70 kDa protein 2	KSINPDEAVAYGAA	14		m
D4A962	Heterogeneous nuclear ribonucleoprotein U-like 1	DYGGYSGSTGGTSTQ	16	N	C-ter
A7VIC2	Heterogeneous nuclear ribonucleoproteins A2/B1	SGNFGGSRNMGGP	13	C	C-ter
A7VIC2	Heterogeneous nuclear ribonucleoproteins A2/B1	NYGPGSGSGGYGG	15		
A7VIC2	Heterogeneous nuclear ribonucleoproteins A2/B1	SGNFGGSRNMGGPYG	15		
A7VIC2	Heterogeneous nuclear ribonucleoproteins A2/B1	GSRNMGGPYGGNYGP	16		
A7VIC2	Heterogeneous nuclear ribonucleoproteins A2/B1	YGGGNYGPGSGSGG	16		
A7VIC2	Heterogeneous nuclear ribonucleoproteins A2/B1	SGNFGGSRNMGGPYGGG	17		
A7VIC2	Heterogeneous nuclear ribonucleoproteins A2/B1	SGNFGGSRNMGGPYGGNYGP	21		
Q8VHV7	Heterogeneous nuclear ribonucleoprotein H	ENSSDFQSNIA	11	N	C-ter
Q00715	Histone H2B type 1	AMGIMNSFVNDIFER	15	N	m
D3ZJ08	Histone H3	ELLIRKLPF	9	N	m
P62804	Histone H4	DNIQGITKPAIR	12	N	m
P62804	Histone H4	TVTAMDVVYALK	12		C-ter
P04233	HLA class II histocompatibility antigen gamma chain CD74	VTKQDLGPVPM	11	MB	C-ter
P04233	HLA class II histocompatibility antigen gamma chain CD74	PSSGLGVTKQDLGPVPM	17		
P04233	HLA class II histocompatibility antigen gamma chain CD74	DPSSGLGVTKQDLGPVPM	18		
P04233	HLA class II histocompatibility antigen gamma chain CD74	EDPSSGLGVTKQDLGPVPM	19		
P04233	HLA class II histocompatibility antigen gamma chain CD74	LEDPSGLGVTKQDLGPVPM	20		
P04233	HLA class II histocompatibility antigen gamma chain CD74	MATPLLQALPMGALP	16		m
P04233	HLA class II histocompatibility antigen gamma chain CD74	RMATPLLQALPMGALP	17		
P04233	HLA class II histocompatibility antigen gamma chain CD74	RMATPLLQALPM	13		
P04233	HLA class II histocompatibility antigen gamma chain CD74	RMATPLLQALPMGALPQ	18		
P01903	HLA class II histocompatibility antigen, DR alpha chain	ANIAVDKANLEIMT	14	MB	m
P01903	HLA class II histocompatibility antigen, DR alpha chain	LANIAVDKANLEIMT	15		
P01903	HLA class II histocompatibility antigen, DR alpha chain	ALANIAVDKANLEIMT	16		
P01903	HLA class II histocompatibility antigen, DR alpha chain	ALANIAVDKANLEIMTK	17		
P01903	HLA class II histocompatibility antigen, DR alpha chain	LANIAVDKANLEIM	14		
P01903	HLA class II histocompatibility antigen, DR alpha chain	LANIAVDKANLEIMTK	16		
P01903	HLA class II histocompatibility antigen, DR alpha chain	IQAIFYLNPDQSGEF	15		N-ter
P01903	HLA class II histocompatibility antigen, DR alpha chain	IQAIFYLNPDQSGEFM	16		
O70592	HnRNP protein	SNYGPMKGSFSGG	13	N	C-ter
O70592	HnRNP protein	GSFSGGRSSGSP	12		C-ter
O70592	HnRNP protein	GSFSGGRSSGSPYG	14		
O70592	HnRNP protein	GSFSGGRSSGSPYGGG	16		
O70592	HnRNP protein	GSFSGGRSSGSPYGGGYG	18		
O70592	HnRNP protein	GRSSGSPYGGGYG	13		
O70592	HnRNP protein	GGYSGSGSGGYG	13		C-ter
O70592	HnRNP protein	YGGGYSGSGSGGY	14		
O70592	HnRNP protein	YGGGYSGSGSGGYG	15		
O70592	HnRNP protein	GSPYGGGYSGSGSGG	16		
O70592	HnRNP protein	GGGPGYGNQGGGYGGGGG	19		m
O70592	HnRNP protein	GGGPGYGNQGGGYG	14		

Uniprot AN	Proteïna	Seqüència	Tamany	Localització	Terminació
Q6P6G9	Hnrpa1 protein	GGNFGGRSSGPGYGGGGQYF	19	N	C-ter
Q6P6G9	Hnrpa1 protein	GGNFGGRSSGPGYGG	14		
Q6P6G9	Hnrpa1 protein	RSSGPGYGGGGQY	12		
Q6P6G9	Hnrpa1 protein	RSSGPGYGGGGQYF	13		
Q6P6G9	Hnrpa1 protein	NQGGYGGSSSSSYGSG	17		C-ter
Q6P0K8	Junction plakoglobin	LHSRNEGAT	10	C	m
P02688	Myelin basic protein	GKGRGLSLSRFSWGA	15	MB	m
Q62667	Major vault protein	FGLLGLSDGQPPAQK	16	C	C-ter
Q62667	Major vault protein	FSTAFGLLGLSDGQPPAQK	20		
Q0QF43	Malate dehydrogenase	AKVAVLGASGGIGQPLSL	18	C	N-ter
Q0QF43	Malate dehydrogenase	AKVAVLGASGGIGQPLSLL	19		
P49186	Mitogen-activated protein kinase 9	RPKYPGIKFEE	11	N	m
Q9QYW3	MOB-like protein phocein	EEVQNSVSGESEA	13	G	C-ter
Q9JJ40	Na(+)/H(+) exchange regulatory cofactor NHE-RF3	LRIEKDTDGH	11	MB	N-ter
P19804	Nucleoside diphosphate kinase B	GETNPADSKPGTIRGD	16	C	m
Q2YDU3	OTU domain-containing protein 5	GPGGPGGGPGDALGAT	16	NA	N-ter
Q9Z254	PDZ domain-containing protein GIPC1	AFDMISQRSSGHPGSG	17	C	m
P14925	Peptidyl-glycine alpha-amidating monooxygenase	EEAFVIDFKPRASMD	15	MB	N-ter
P14925	Peptidyl-glycine alpha-amidating monooxygenase	EEAFVIDFKPRASMDT	16		
P14925	Peptidyl-glycine alpha-amidating monooxygenase	LPVDEEAFVIDFKPRASMD	19		
P14925	Peptidyl-glycine alpha-amidating monooxygenase	LPVDEEAFVIDFKPRASMDT	20		
E9PTD2	Phosphatidylinositol-binding clathrin assembly protein	TMAPPVMAYPATTP	14	N	C-ter
D4AEH3	Proteasome (Prosome, macropain) 26S subunit, non-ATPase, 7	PELAVQKVVVHPLVL	15	C	N-ter
P34064	Proteasome subunit alpha type-5	FGGVDEKGP	9	C	m
Q6AYR1	Protein 100911212	MAPPSTGANPY	11	C	C-ter
Q6AYR1	Protein 100911212	TMAPPSTGANPY	12		
Q6AYR1	Protein 100911212	TMAPPSTGANPYAR	14		
Q6AYR1	Protein 100911212	TMAPPSTGANPYA	13		
Q6AYR1	Protein 100911212	QGYTQPGPGYR	11		C-ter
Q6AYR1	Protein 100911212	NRPPFGQGYTQPGPGYR	17		
Q6AYR1	Protein 100911212	SGQPQQLPAQPPQ	13		m
Q6AYR1	Protein 100911212	SGQPQQLPAQPPQQ	14		
Q6AYR1	Protein 100911212	SGQPQQLPAQPPQQYQ	16		
Q6AYR1	Protein 100911212	SGQPQQLPAQPPQQY	15		
Q6AYR1	Protein 100911212	SGQPQQLPAQPPQQYQA	17		
Q6AYR1	Protein 100911212	SGQPQQLPAQPPQQYQAS	18		
Q6AYR1	Protein 100911212	SASYSQQTGPQQPQQF	16		m
D3ZXL1	Protein Arih1	ALQRYLFYCNRYMNMHMSLRFEHKL	25	C	m
E9PTI1	Protein Atp6v1h	SEQPQTAAARS	11	END	C-ter
Q63081	Protein disulfide-isomerase A6	DVELDDLEKDEL	12	ER	C-ter
Q63081	Protein disulfide-isomerase A6	IDLSDELDDLEKDEL	16		
Q63081	Protein disulfide-isomerase A6	DIDLSDELDDLEKDEL	17		
Q63081	Protein disulfide-isomerase A6	DDIDLSDELDDLEKDEL	18		
M0RBT0	Protein Gsg2	RLMRKENKNCWGEYHPYNNVLWLHY	25	N	C-ter*
D4ABT8	Protein Hnrnpul2	LEEEDEEPPPAQALG	15	N	N-ter
D3Z914	Protein Inip	PVLPRLDPE	9	N	C-ter
D4A6G6	Protein LOC100362339	PGVTVKDVNQQE	12	ER	N-ter
D4A6G6	Protein LOC100362339	PGVTVKDVNQQEFVRALA	18		
D3ZFA8	Protein LOC100362366	QVTQPTVGMNFKTPRGAV	18	ER	C-ter
M0RB03	Protein LOC100911677	AQPATYRAQPSVSLG	15	N	m
M0RB03	Protein LOC100911677	SSQPAAYVAQPAAY	17		m
D3ZEV0	Protein LOC100912427	PPAENSSAPEAEQGGAE	17	N	C-ter
D3ZEV0	Protein LOC100912427	AADPPAENSSAPEAEQGGAE	20		
D4A510	Protein LOC685179	PTAPSPGTVPVPPQ	16	N	C-ter
D4A510	Protein LOC685179	AAAPATVHGLAVAPASVA	18		m
F1LPX5	Protein Plscr2	IQFPLDLVKMKAVMLG	17	C	C-ter
Q5XFW8	Protein SEC13 homolog	ASITEGQQNEQ	11	END	C-ter
D4AAM0	Protein Tnp03	RLAVIFRHTNPVENGQTHP	20	N	m
Q70453	Putative heme oxygenase 3	DDRAENTQFVKDF	13	NA	N-ter

Uniprot AN	Proteïna	Seqüència	Tamany	Localització	Terminació
Q9JIS1	Regulating synaptic membrane exocytosis protein 2	ELEDSRISLLR	11	MB	m
D3ZXA2	Ribosomal protein L15	EVILIDPFHK	10	C	m
B5DFE9	Smarca4 protein	GSEHASSPVPASGPSSGPQM	20	N	m
Q88457	Sodium channel protein type 11 subunit alpha	GEGGALPTSQH	11	MB	m
Q6AYL5	Splicing factor 3B subunit 4	ADAPPPPSAPNPVSSLG	18	N	m
Q5XI66	Ss18 protein	DQGQYGNYYQ	10	N	C-ter
F1M953	Stress-70 protein, mitochondrial	ASEAIKGAVVG	11	MIT	N-ter
F1M953	Stress-70 protein, mitochondrial	ASEAIKGAVVGIDLG	15		
O54772	matrix-associated actin-dependent regulator of chromatin subfa	SPGSRMPMAGLQVGPAGS	19	N	m
O54772	matrix-associated actin-dependent regulator of chromatin subfa	EQVLGIRLT	9		C-ter
Q56A18	matrix-associated actin-dependent regulator of chromatin subfa	VEEPPTDPVPEDEKKE	16	N	C-ter
Q56A18	matrix-associated actin-dependent regulator of chromatin subfa	SATVEEPPTDPVPEDEKKE	19		
Q56A18	matrix-associated actin-dependent regulator of chromatin subfa	TSDSNTGSESNSATVEEPTD	21		
Q9JI92	Syntenin-1	ITSIVKDSAAARNGLL	16	MB	m
P62329	Thymosin beta-4	PLPSKETIEQEQAGES	17	C	C-ter
P69897	Tubulin beta-5 chain	GDSDLQLDRISVYVNEA	17	C	N-ter
Q5I0L3	Tyrosine--tRNA ligase, mitochondrial	IGKRNFYIHK	10	M	C-ter
F1M3Q4	Uncharacterized protein	LAIDPGDSDIIRSMPEQTGEK	22	RE	C-ter
F1LTG5	Uncharacterized protein	ASGGGVPTDEEQATG	15	MIT	m
F1LTG5	Uncharacterized protein	ASGGGVPTDEEQATGLER	18		
D3Z8W1	Uncharacterized protein	YVPVTTFKNLQTVNVN	18	N	C-ter
D3Z8W1	Uncharacterized protein	VTYVPVTTFKNLQTVNVN	20		
M0RAD0	Uncharacterized protein	EKKVVKKEKYGKIAAYR	18	NA	m
M0RB94	Uncharacterized protein	KMQFQSQRQTERDRETDQRERERE	25	NA	C-ter
Q5XI52	Vezatin	LQELKSVLGF	10	MB	m
A1L1J6	Zinc finger protein 652	CPRVFNTRWYLEKHMNVTHRRMQIC	25	N	m

Uniprot AN: Identificació de la proteïna en SwissProt; Proteïna: proteïna origen de la qual prové el pèptid; Longitud: mida del pèptid en nombre d'aminoàcids; Localització: MB (membrana plasmàtica); Lis (lisosomes); End (endosomes); S (secretada); EM (matriu extracel·lular); ER (Reticle endoplàsmic); G (Golgi); MIT (mitocòndria) C (citòsol); N (nuclear)

2.3 COMPARACIÓ GENERAL ENTRE ELS REPERTORIS

Per estudiar l'efecte de la l'expressió de la cadena invariant i HLA-DM en els pèptids associats a DR3, es van comparar els repertoris de les quatre línies cel·lulars.

2.3.1 Comparació dels repertoris comuns i únics entre les línies transfectants

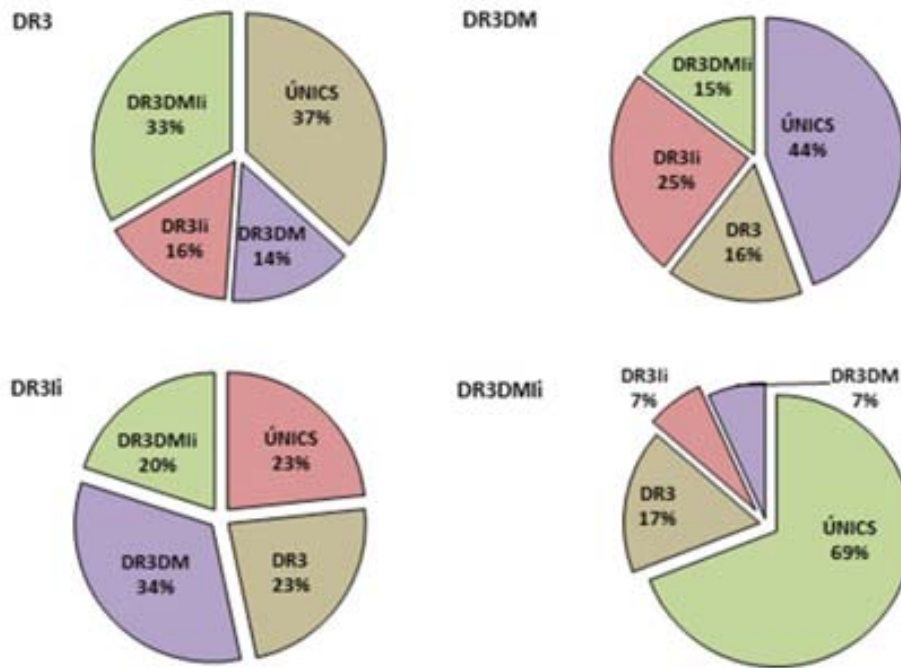


Figura 2.6: Gràfic de percentatges on es mostra quina part del repertori peptídic es compartit i quina és única per les diferents línies transfectades RINDR3. Els colors corresponen a cada una de les cèl·lules, DR3 (marró), DR3DM (morat) DR3li (vermell) i DR3DMli (verd).

A la figura 2.6 observen el percentatge de pèptids que són únics per cada línia i quins són compartits amb les altres línies.

- RINDR3 : En absència de les xaperones, es van identificar 90 pèptids dels que 33 eren únics (37%) i no els trobaven als repertoris de les altres 3 línies. Un 33% del repertori el trobàvem a la línia DR3DMli, i un 16% i 14% a les línies DR3li i DR3DM respectivament.
- RINDR3DM: Es van identificar 81 pèptids dels quals 36 eren únics (44%) i no es trobaven als repertori de les tres altres línies. Un 15% del repertori contenia pèptids en comú amb la línia DR3DMli, i un 25% i 16% amb DR3li i DR3 respectivament.
- RINDR3li: Dels 60 pèptids identificats, 14 eren únics (23%) i no els trobaven als repertoris de les altres 3 línies. Un 20% del repertori era

compartit amb la línia DR3DMli, i un 34% i 23% amb DR3DM i DR3 respectivament.

- RINDR3DMli: Per aquesta línia es va identificar 176 pèptids dels quals 121 eren únics (69%) i no els trobaven als altres repertoris. Un 17% del repertori conté pèptids en comú amb la línia DR3, i només un 7% i 7% del total els trobaven també als repertoris de DR3DM i DR3li respectivament.

Per tant, la línia RINDR3DMli era la que presentava un major repertori de pèptids únics, seguida de la línia DM. La línia DR3li va ser la que presentava un repertori propi menor.

2.3.2 La via citosòlica és la ruta de processament preferent per a HLA-DR3

Es va comparar la localització subcel·lular de les proteïnes que generaven els lligands de HLA-DR3, agrupant-les segons la seva localització i la seva ruta de processament.

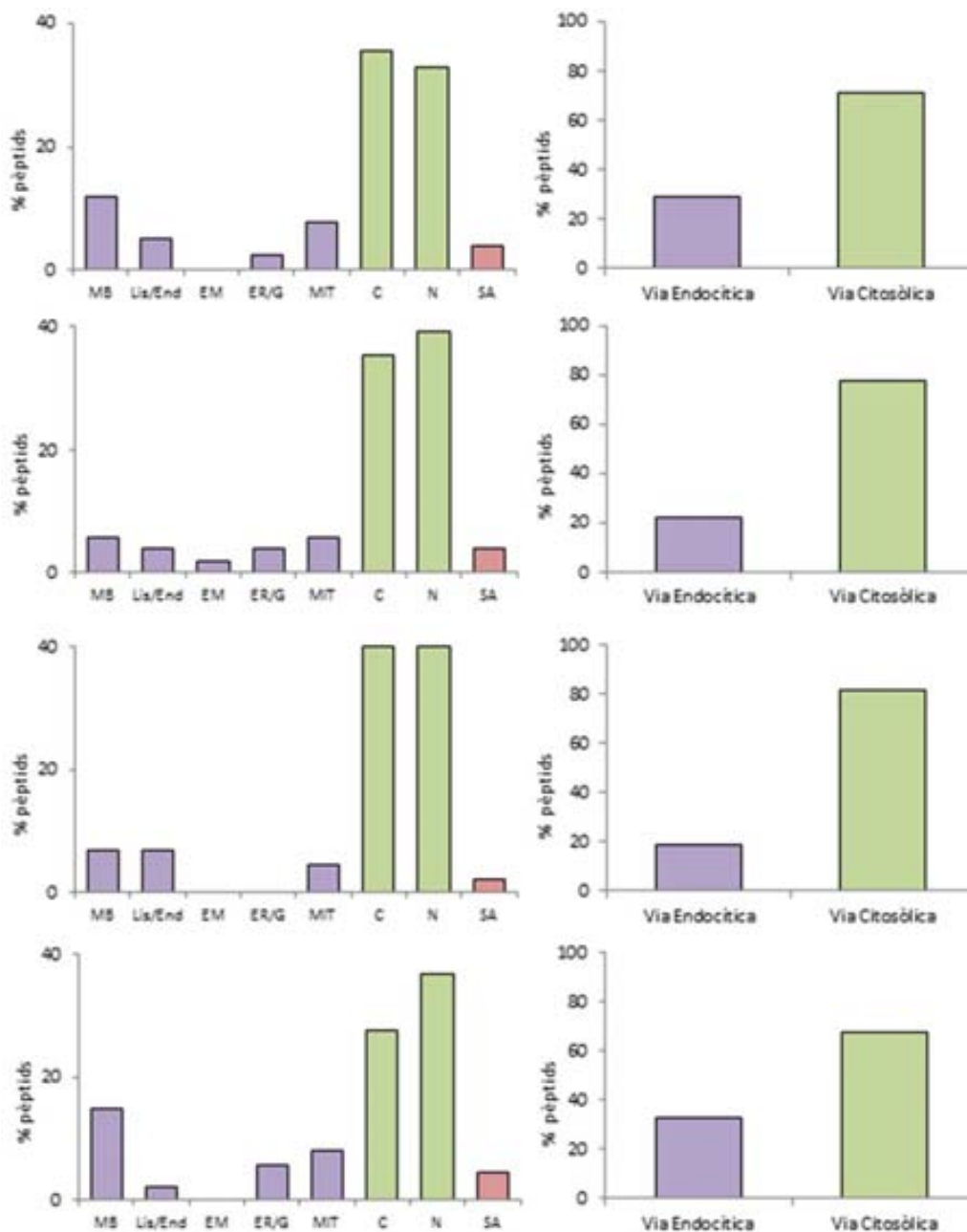


Figura 2.7. Localització de les proteïnes origen dels pèptids aïllats de DR3. A) RINDR3; B) RINDR3DM; C) RINDR3li; D) RINDR3DMli. Esquerra: Gràfic en funció de la seva localització cel·lular: MB (membrana plasmàtica); Lis/End (lisosomes o endosomes); EM (matriu extracel·lular); ER/G (Reticle endoplàsmic o Golgi); MIT (mitochondria); C (citoplasm); N (nuclear); SA (sense assignació). Dreta: les localitzacions s'agrupen en funció de la seva via de processament, via endocítica (barres morades); C: via citosòlica (barres verdes).

Per a les quatre línies DR3, la via citosòlica fou la via de degradació majoritària. En aquest cas, més un 71% dels pèptids pertanyen a aquesta ruta de processament en el cas de RINDR3, un 67% en el cas de RINDR3DMi, un 80% en el cas de DR3DM i superant el 81% en el cas de DR3li (figura 2.7). Aquest fet només s'havia observat en cèl·lules DR4 en absència de xaperones (RINDR4) [89]. Per tant i per DR3, sembla ser que l'expressió o absència de les xaperones no influenciava la procedència de las proteïnes origen dels pèptids. A la figura 2.10 es representa aquestes dades per cada repertori analitzat de les quatre línies transfectades. Aquestes dades coincideixen amb la tendència observada en la línia cel·lular CDG, utilitzada per l'estudi del motiu de DR3 (capítol 1).

2.3.3 Predicció d'afinitat dels lligands d'HLA-DR3

El mètode de predicció d'afinitat teòrica dels pèptids es va fer basant-nos en el motiu que s'havia descrit per HLA-DR3 al capítol 1.

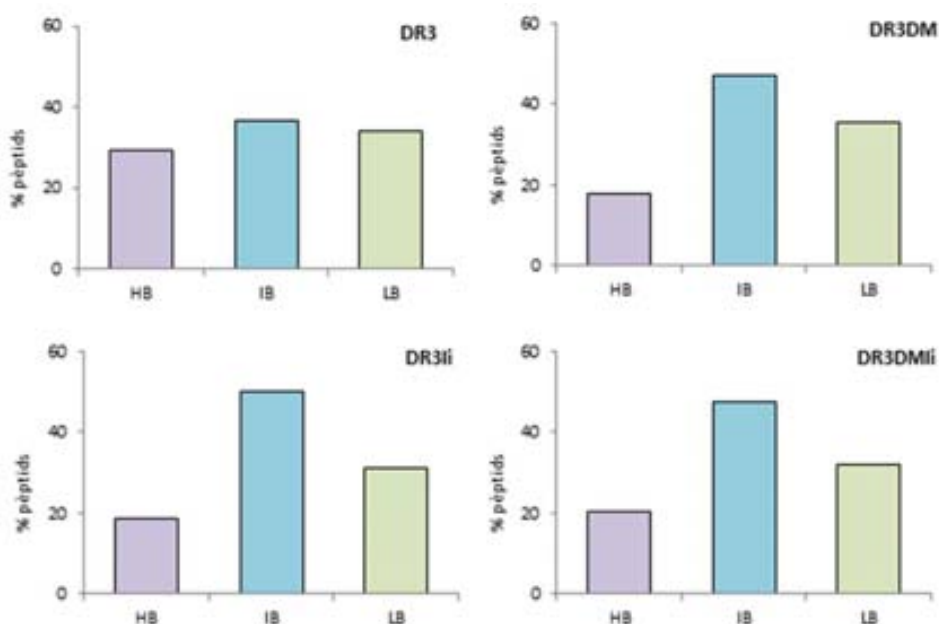


Figura 2.8 : Comparació de la predicció d'afinitat d'unió teòrica dels pèptids associats a HLA-DR3.

Els resultats entre les quatre línies fou similar, amb un domini dels epítops d'afinitat intermèdia (IB): 37% per RINDR3, 47% per RINDR3DM, 50% per

RINDR3li i 47% per RINDR3DMli (figura 2.8). A les RINDR3 es on vam trobar menys diferències entre les afinitats i a més, un percentatge més alt de HB (29%) en comparació amb les altres línies.

2.3.4. El patró de processament en DR3 està lligat a extrems terminals

Es va analitzar el possible patró de processament dels pèptids aïllats segons la localització d'aquests en la proteïna parental a les quatre línies DR3 (figura 2.9). En el cas dels pèptids agrupats en *nested sets*, només comptàvem un sol cop la localització. Si la proteïna generava dos tipus de pèptids amb processament diferent, llavors comptàvem les dues localitzacions. Es va obtenir que un 60% aproximadament dels pèptids presentaven patrons terminals, sent C-terminal el majoritari (40-50% del total).

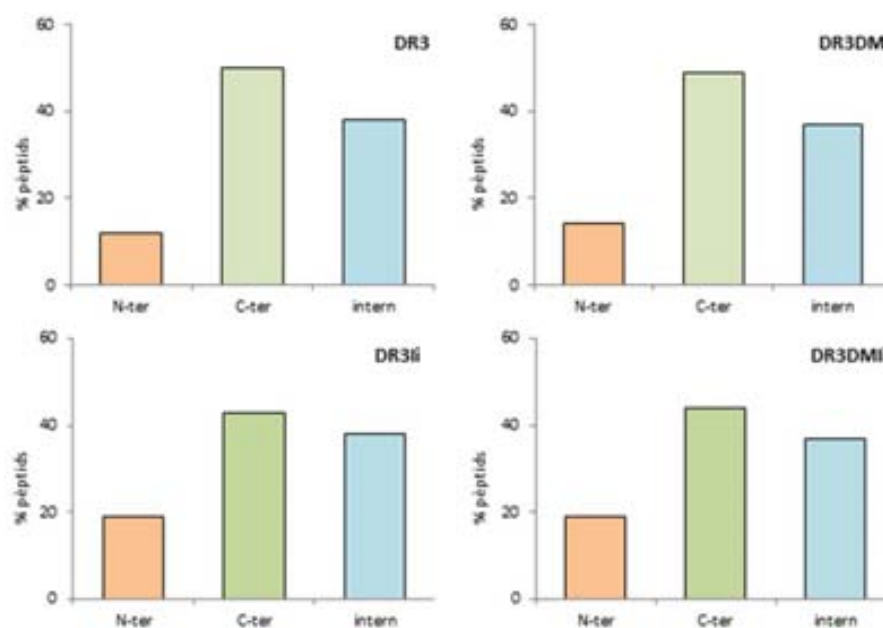


Figura 2.9: Patró de processament dels pèptids aïllats de les línies DR3. Gràfic en funció del patró de processament dels pèptids depenent si són N-terminal (taronja), C-terminal (verd) o interns (blau).

A més, analitzàrem aquest patrons, relacionats amb la ruta de processament de la proteïna, amb la via de degradació a la que pertanyia: la via citosòlica era la majoritària per als pèptids amb patró C-terminal menys en les cèl·lules DR3li on el percentatge de pèptids interns dins de la via citosòlica era superior (37%

interns, 30% C-ter). A la via endocítica trobàvem dues situacions: per a DR3 i DR3DMii trobàvem pràcticament la mateixa freqüència de seqüències internes com per de pèptids C-terminal i per a DR3DM i DR3li els pèptids majoritaris eren C-ter. (figura 2.10).

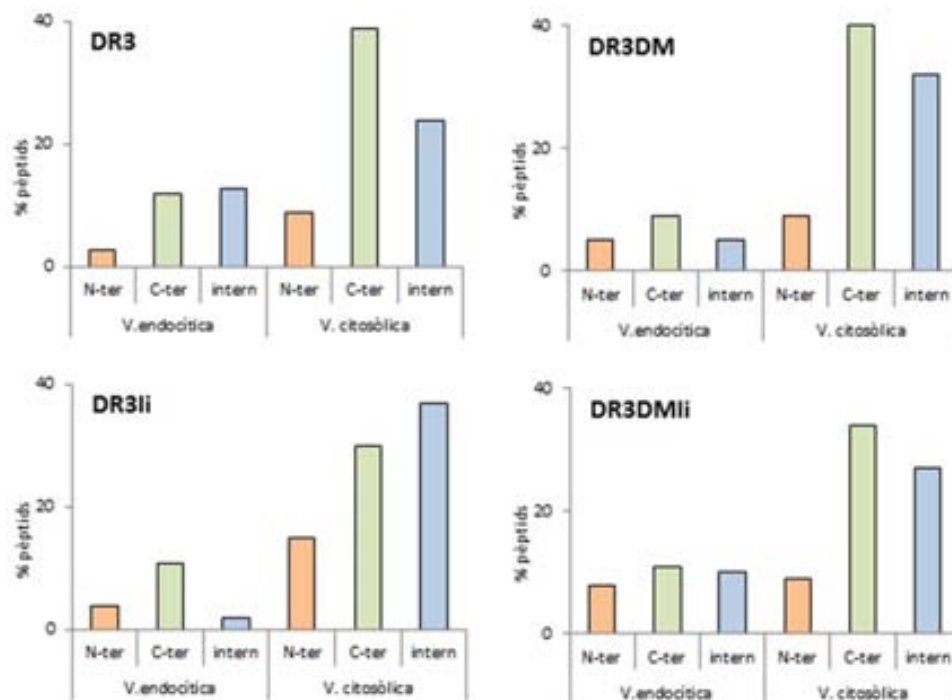


Figura 2.10: Gràfic comparatiu per les quatre línies RINDR3 on es relaciona quin patró de processament de les pèptids (N-ter, C-ter o intern) trobàvem respecte la seva via de degradació (via endocítica o citosòlica).

Es va analitzar el repertori de pèptids de HLA-DR3 en la línia CDG (homozigota per DR3) estudiada al capítol 1. En aquestes cèl·lules, es va observar que un 51% dels pèptids es localitzaven als extrems N-terminal (19%) i C-terminal (32%) i un 49% dels pèptids a la part interna de la proteïna parental. Si analitzàvem a més, a quina via de degradació pertanyien aquests pèptids, comprovàvem que la majoria de pèptids C-terminal pertanyien a la via citosòlica (figura 2.11).

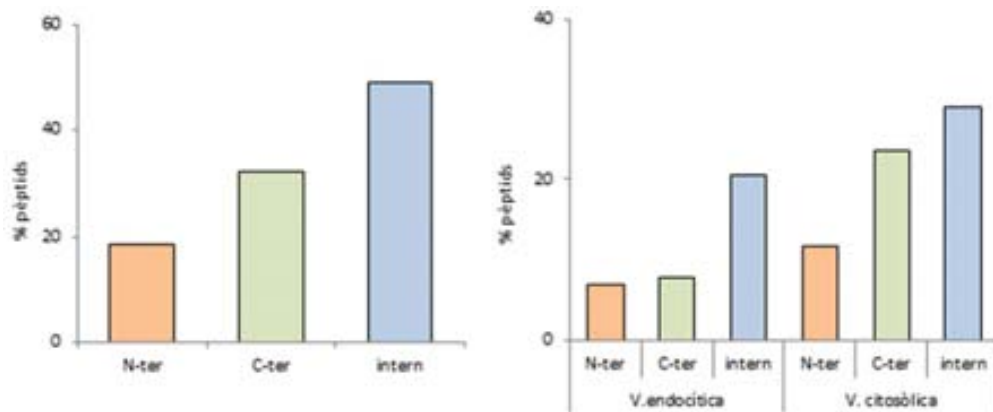


Figura 2.11: Patró de processament dels pèptids aïllats de DR3 en la línia homozigota CDG. Esquerra: Gràfic en funció de la localització dels pèptids respecte la proteïna parental depenent si són N-terminal (taronja), C-terminal (verd) o intern (blau). Al gràfic de la dreta, les localitzacions s'agrupen en funció de la seva via de processament, via endocítica i via citosòlica.

En alguns casos, vam observar que una mateixa proteïna podia generar dos pèptids que provenien de dos localitzacions diferents, com ara la miosina 9 de la qual es va obtenir un pèptid amb processament C-terminal (EEVDGKADGAEAKPAE) i un amb patró intern (IDQINTDLNLERSHAQ).

Per validar el resultat en diferents línies i teixits, es va fer una anàlisi global amb diferents línies i teixits (DR3 i d'altres que expressaven altres al·lels de DR) que s'havien caracteritzat i estudiat al nostre laboratori. En alguns casos, les línies o teixits eren heterozigots i llavors en eixe cas, s'escollia els pèptids que mostraven afinitat únicament per un dels dos al·lels que expressava.

1) Línies cel·lulars: una línia cel·lular LCL-DR8 [90], una línia LCL-DR10 [109], una línia LCL-DR4 [115] una línia EVB-DR3 (capítol 1), una línia heterozigota per DR3 (HT93CIITA, capítol 3). A la figura 2.12, es mostra un gràfic comparatiu dels patrons obtinguts per aquestes mostres.

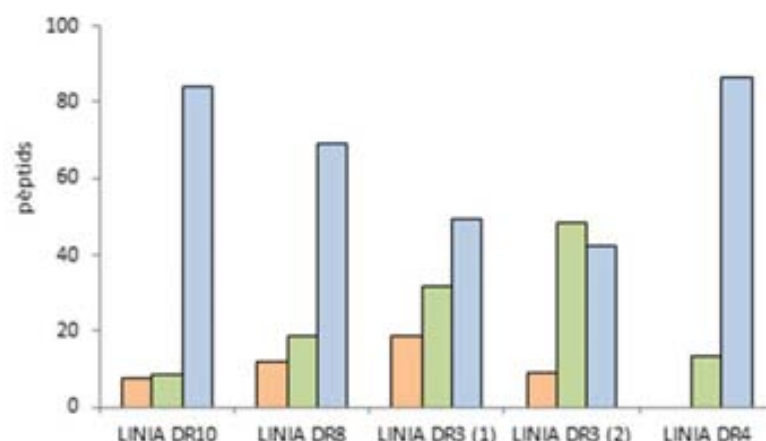


Figura 2.12: Patró de processament dels pèptids aïllats en diferents línies cel·lulars, en funció de si són N-terminal (taronja), C-terminal (verd) o interns (blau). Línia DR3 (1) CDG (capítol 1). Línia DR3 (2) línia heterozigota per DR3 estudiada al capítol 3.

2) Teixits: una melsa heterozigota DRB1*0301 *0801 (escollirem els dos al·lels), una melsa heterozigota DRB1*0405 *0407 (escollirem *0405), una melsa heterozigota DRB1*0701 *1501 (escollirem *0701) una melsa homozigota DR3 i dos timus DR3 (un homozigot i un heterozigot) dels quals es van unir les dades per tal d'augmentar el nombre (Tesis Dr. JA Collado, UAB 2013). A la figura 2.13 es mostra un gràfic comparatiu dels patrons obtinguts per aquestes mostres.

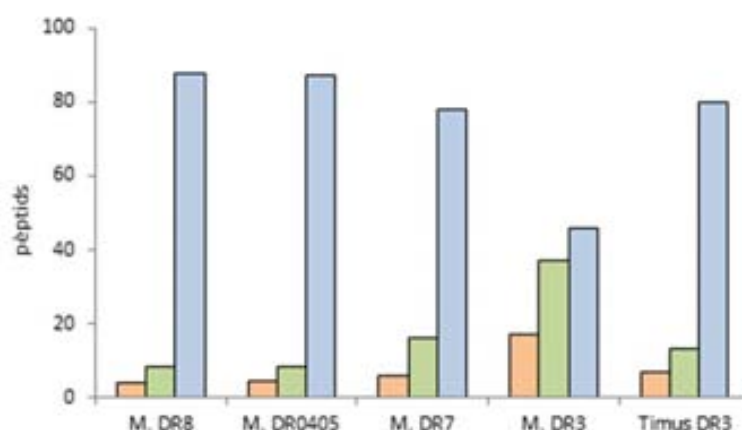


Figura 2.13: Patró de processament dels pèptids aïllats en melsa (M) i en timus DR3, en funció de si són N-terminal (taronja), C-terminal (verd) o interns (blau).

Per a les línies no DR3, aproximadament un 70-80% eren pèptids corresponents a seqüències internes de les proteïnes corresponents. El mateix

patró s'observà per a les diferents melses no DR3 i pel timus DR3. Per contra, en el cas de la melsa DR3, hi ha un 54% de seqüències terminals i com ja havíem vist a l'apartat anterior, el mateix ocorre amb les dues línies DR3, on el balanç de pèptids que pertanyen a seqüències terminals de la proteïna és major, sent l'extrem C-terminal el principal.

2.4. COMPARACIÓ DELS REPERTORIS ÚNICS A LES CÈL·LULES TRANSFECTANTS DR3

Amb l'objectiu de conèixer el comportament dels pèptids únics de cada línia cel·lular, aquestes seqüències es van analitzar per separat i es van comparar amb els pèptids comuns.

A diferència del apartat 2.4, les comparacions es van fer a nivell de pèptid i no de proteïna parental. A la figura 2.14 es comparen les vies de degradació per cada tipus de repertori. Observàrem que la via citosòlica, com era d'esperar, era la més abundant. Tot i així, cal destacar que per cèl·lules DR3DMli, la via endocítica suposava un 34% del qual un 29% corresponia al repertori únic. Així mateix, per la línia DR3DM la via endocítica suposava un 18% en total, del qual un 16% corresponia al repertori únic. En les altres dues línies, tot i que el repertori únic presentava un percentatge més alt de pèptids d'origen citosòlic, el balanç entre aquests i el del repertori compartit no era tan acusat.

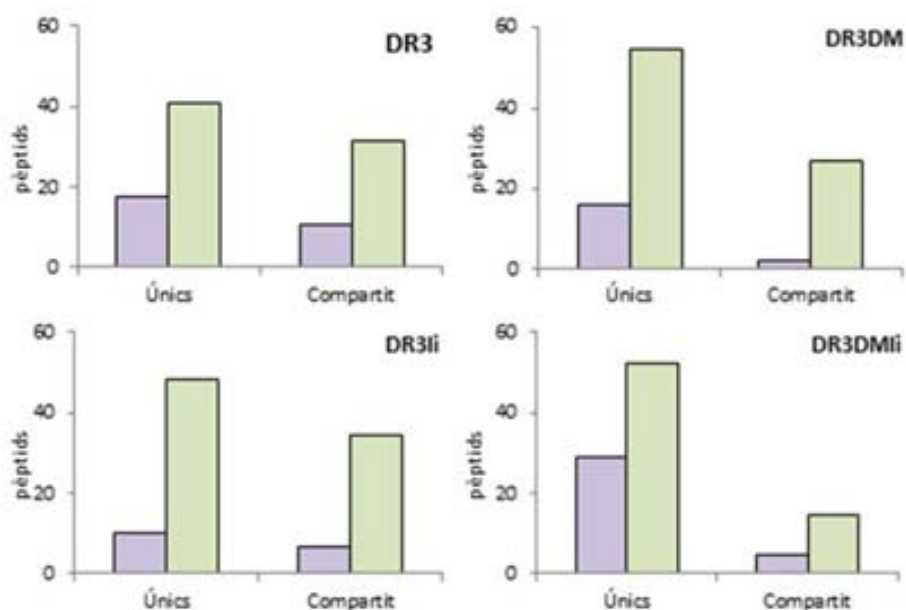


Figura 2.14: Repertoris únics i compartits a les cèl·lules transfectants DR3. Els repertoris s'agrupen segons la via de degradació a la que es localitzen: endocítica (lila) i citosòlica (verd).

Respecte a la localització del pèptid a la proteïna parental (figura 2.15), observarem que en tots els casos el patró C-ter va ser el més abundant, menys en el cas del repertori propi de DR3 en absència de les dues xaperones, on el pèptid localitzat en la part interna de la proteïna era lleugerament superior (29%) als c-terminals (27%).

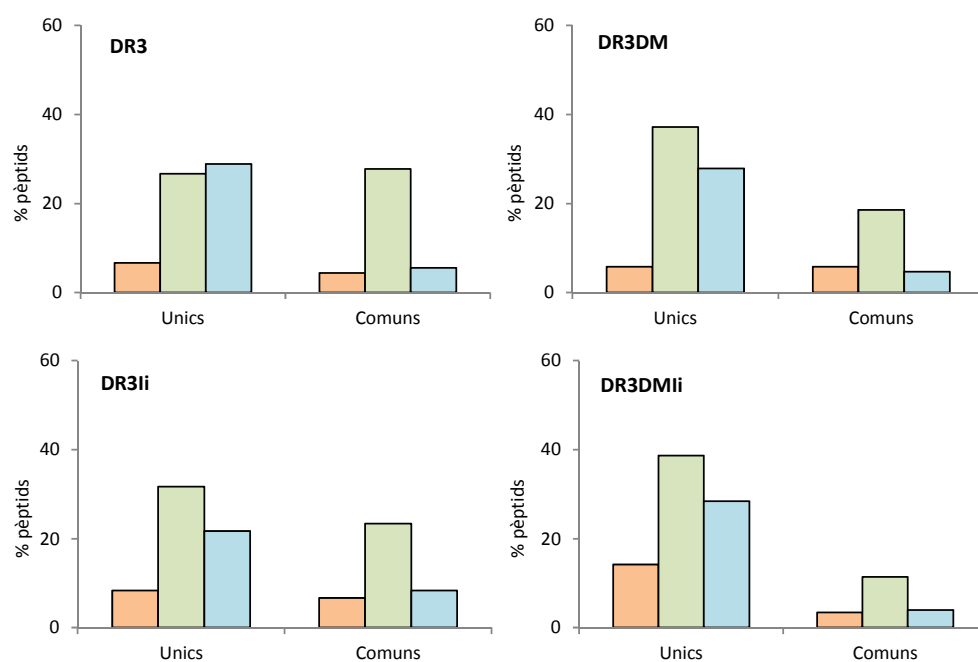


Figura 2.15: Patró de processament dels pèptids en funció de si són N-terminal (taronja), C-terminal (verd) o interns (blau) pels pèptids únics i comuns de cada línia transfectant.

Finalment, ens vàrem fixar en l'afinitat del repertori únic on trobàrem més heterogeneïtat. Per a DR3DMli i per a DR3DM, el patró va ser IB>LB>HB tal i com havíem obtingut per al repertori total i per DR3 (capítol 1). Per DR3 i DR3li, el patró fou LB>IB>HB. A més, el repertori propi de DR3 en absència de xaperones, presentava el percentatge més alt de pèptids d'alta afinitat (22%) en comparació amb les altres 3 línies (figura 2.16).

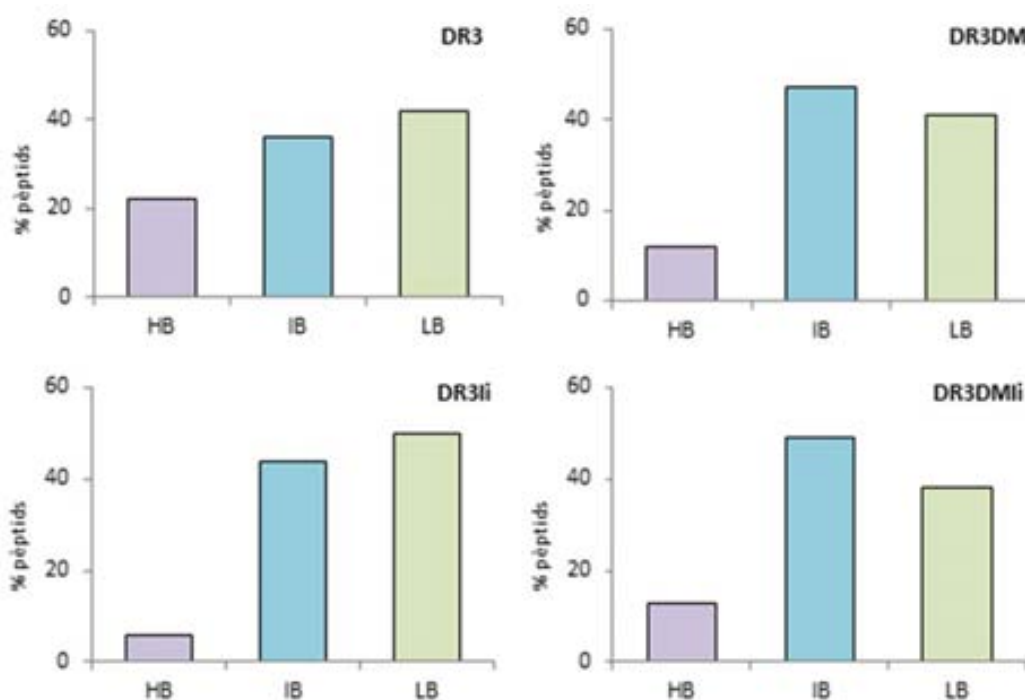


Figura 2.16: Comparació de la predicció d'afinitat d'unió teòrica dels repertoris únics de cada línia associats a HLA-DR3.

2.5. COMPARACIÓ DELS REPERTORIS OBTINGUTS A RIN5MF TRANSFECTADA AMB DR3 I AMB DR4

En un treball anterior del nostre grup realitzat per la Dra. A. Muntasell i continuat per la Dra. M. Carrascal (Tesis DRA M. Carrascal) es va estudiar la influència que exercien les xaperones li i DM en la presentació dels repertoris peptídics en DR4. La comparació de seqüències dels quatre repertoris indicava que DR4 en les cèl·lules DR4liDM s'associava a un conjunt de pèptids, el 60% dels quals eren dependents de la presència d'li i DM. La presència d'li en

absència de DM, també condicionava els lligands de DR4 tant pel predomini del CLIP com per la resta d'epítops identificats, que semblaven exclusius del repertori d'aquestes cèl·lules. En absència d'ambdues xaperones, el repertori estava format majoritàriament per seqüències d'extrems N- o C-terminals de proteïnes citosòliques, la meitat dels quals no respectaven el motiu d'unió a DR4 i eren dependents de l'absència d'I κ i i DM.

Si comparem els resultats que hem obtingut per DR3 amb els que es van obtenir amb les cèl·lules DR4, veiem que la línia DR4 en absència d'ambdues xaperones, presenta efectivament un repertori més semblant al que segueixen totes les DR3, on la via citosòlica seria la majoritària amb un 64% de pèptids (figura 2.17).

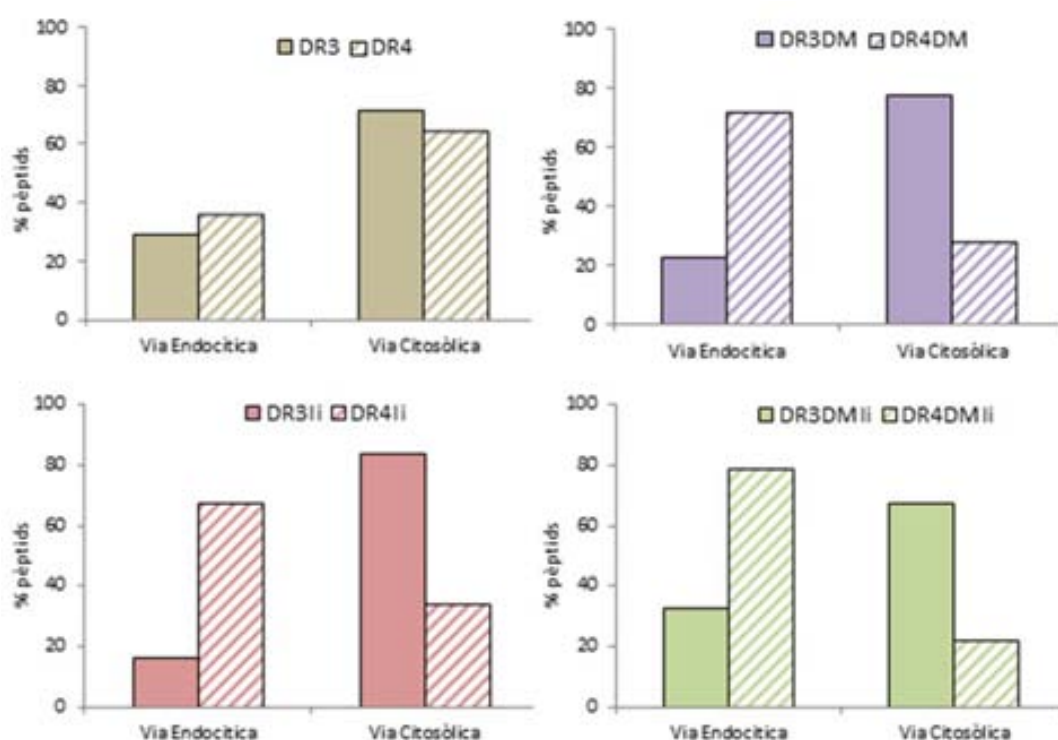


Figura 2.17: Comparació de la localització de les proteïnes origen dels pèptids aïllats en les diferents línies RIN transfectades DR3 i DR4 en funció de la seva via de processament

També es va comparar la localització dels pèptids respecte la proteïna d'origen en ambdós al·lells (figura 2.18).

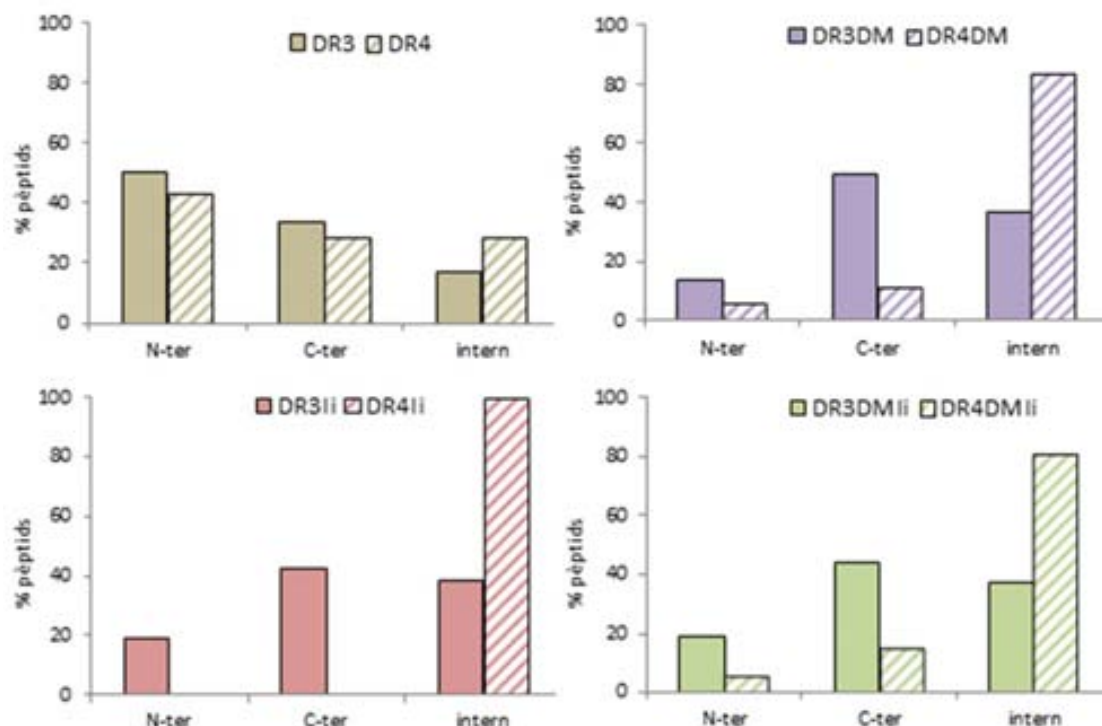


Figura 2.18: Patró de processament dels pèptids respecte la proteïna parental procedents de les diferents línies RIN DR3 i DR4

La línia DR4 en absència de xaperones presentava un patró semblat al que hem descrit per DR3: Dels 12 pèptids aïllats a aquesta línia, sols 2 eren seqüències internes de la proteïna parental, per tant, un hi havia un 71% dels pèptids que es localitzaven en els extrems N o C-terminals (Tesis M. Carrascal 2004). Per contra, la DR4DM, DR4li i DR4DMli, presenten un patró semblant descrit per les línies DR4, DR8 i DR10, on la gran majoria dels pèptids tenen una localització interna respecte la proteïna parental.

El patró definitiu que s'observa a la comparació ens mostra que totes les línies cel·lulars DR3⁺ es comporten com les cèl·lules DR4 sense xaperones, respecte a l'origen de les proteïnes parentals i a la dominància de pèptids c-terminals.

Però el resultat que es va obtenir més rellevant fou que observant el patrons a nivell de teixits DR3, aquesta tendència es confirmava en melsa però no al timus.

2.6. OBSERVACIONS FINALS

Els resultats d'aquest capítol van permetre caracteritzar el repertori de pèptids associats a DR3, generats per la línia endocrina RINm5F i aprofundir en la identificació dels lligands més abundants presentats. Les dades presentades demostren que ambdues xaperones no afecten les característiques específiques dels repertoris associats a DR3 en les diferents cèl·lules, al contrari del que ocorria en els estudis amb DR4. En altres estudis basats en la reactivitat de cèl·lules T davant de cèl·lules transfectades que expressaven una o dues de les xaperones específiques o mitjançant l'ús d'APCs aïllades dels ratolins KO per Ii, DM o ambdues molècules [116] [117] [118] [119] també s'observaren diferències degudes a l'expressió de les xaperones.

Aquests resultats prenen rellevància si a més els comparem amb els obtinguts al nostre grup per DR4 on en absència de les dues xaperones, el repertori s'associava majoritàriament a pèptids derivats de proteïnes citoplasmàtiques. La coexpressió de la cadena invariant en absència de DM mostrava la presència de CLIP com a pèptid majoritari i afavoria l'associació de DR4 amb pèptids exògens i de la superfície cel·lular. La presència de DM independentment de l'expressió de Ii afavoria l'associació amb lligands endògens derivats de proteïnes de la superfície cel·lular o de la via endocítica i generava un repertori molt similar al de la triple transfectant.

Aquestes dades ens indiquen que la molècula d'HLA-DR3 té unes característiques diferencials que donen lloc a repertoris peptídics singulars i que podrien estar relacionades amb la susceptibilitat d'aquest al·lel amb diferents malalties autoimmunes.

Les molècules DR3 expressades en les cèl·lules endocrines tant en presència com en absència d'Ii i DM estan associades a un repertori de pèptids homogeni en contingut i en localització bàsicament nuclear i citosòlica.

D'acord amb el que es coneix per altres models experimentals [120] [121] [122] en les cèl·lules amb expressió d'Ii i absència d'HLA-DM, CLIP va ser el pèptid majoritari que es va presentar a la superfície. Així ocorria amb les cèl·lules RINDR4Ii, que presentaven en la seva superfície majoritàriament el pèptid CLIP a més d'una varietat d'epítops que suggerien l'existència de diferències d'especificitat relacionades amb l'absència de la funció editora de la xaperona HLA-DM, mentre que en DR3 no es detectava Ii. D'altra banda, Ghosh et al [5], realitzaren un estudi amb cèl·lules DR3 deficientes en la presentació d'antígens [123] i sense DM on revelaren que CLIP ocupava el lloc d'unió al pèptid de manera idèntica a com ho feia DR1 amb el pèptid HA 306-318 [5] el que suggeria que CLIP estabilitzava tant DR3 com ho podia fer un pèptid antigènic.

En el cas de les cèl·lules que co-expressaven DR3 i DM, els nostres resultats ens indicarien que en almenys el 50-60% dels casos, donada la naturalesa dels pèptids, la càrrega dels pèptids sobre DR3 es donaria en compartiments que no han de contenir DM, i per tant, el transport podria estar interferit.

D'altra banda, a totes aquestes característiques diferencials cal afegir que el patró de processament en DR3 està lligat a extrems terminals, tal i com hem observat a les quatre línies transfectades. Novament, sembla ser que l'acció d'ambdues xaperones no causa efecte en el processament pròpiament dit.

Per tant es pot concloure que ni la cadena invariant ni DM tindrien els efectes sobre el repertori de DR3 que s'esperaven segons la funció canònica de les dues xaperones i aquest fet ens podria indicar que DR3 seria menys sensible a l'acció d'ambdues.

CAPÍTOL 3: PRESENTACIÓ DEL TSHR, AUTOANTÍGEN DOMINANT A LA MALALTIA DE GRAVES-BASEDOW

3.1. OBTENCIÓ D'UN MODEL CEL·LULAR EPITELIAL AMB EXPRESSIÓ D'HLA-DR I TSHR

El principal autoantígen en la malaltia de Graves-Basedow és el receptor de la tirotropina (TSHR). Malgrat això, la seva expressió és molt baixa i fins al moment no s'ha pogut aïllar cap pèptids presentat per HLA-DR d'aquesta proteïna. Es va voler obtenir cèl·lules epitelials humanes d'origen tiroïdal que expressaren HLA-DR i TSHR, i així mimetitzar el que ocorria a la malaltia. Per això, es va escollir la línia cel·lular HT93 (DRB1*0301,*0701), generada per transformació de cèl·lules fol·liculars del tiroides amb SV40 [124].

Es van obtenir 3 línies diferents de cèl·lules transfectants (veure materials i mètodes): dues transfectants simples HT93-CIITA i HT93-TSHR, i una doble transfectant amb els dos gens: HT93-CIITA-TSHR (HT93DT) (figura 3.1). El gen CIITA és un coactivador de l'expressió de diversos gens que codifiquen proteïnes necessàries per a la presentació d'antígens restringits via MHC de classe II, inclosos HLA-DR, DP i DQ, la cadena invariant (Ii), HLA-DM i HLA-DO [86] [125]. Per a l'expressió del TSHR, es va utilitzar un vector recombinant que contenia el cDNA del domini extracel·lular del TSHR (eTSHR), definit com la part immunològica de l'autoantigen (Pichurin PN et al 2006) i un senyal de transport a lisosomes (LAMP) que afavoreix l'acumulació del eTSHR en la via endocítica i per tant la generació de pèptids accessibles a HLA-DR.

Les línies HT93 sense transfectar (*mock transfection*) i la línia HT93-eTSHR ens van servir com a control de l'expressió d'HLA-DR i TSHR, la qual es validava durant el creixement de les diferents línies mitjançant citometria de flux (figura 3.1).

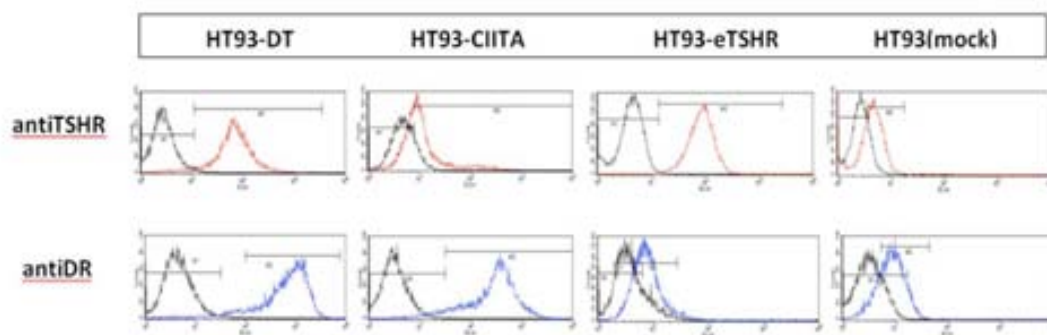


Figura 3.1: Histogrames que mostren l'expressió dels gens transfectats a les diferents línies cel·lulars. Roig: anticòs anti-TSHR; Blau: anticòs anti-DR.

3.2. PROCESSAMENT DE LES MOSTRES

L'aïllament del repertori peptídic associat a les línies HT93CIITA i HT93DT es va fer en dos experiments consecutius, seguint el mateix protocol descrit al capítol 1 i partint de 1000 milions de cèl·lules per cada preparació.

3.3. REPERTORIS DE PÈPTIDS ASSOCIATS A HLA-DR EN CÈL·LULES TRANSFECTADES HT93

Es van identificar 482 i 795 espectres per HT93CIITA i HT93DT respectivament, mitjançant la cerca en la base de dades en Bioworks v4.1, (ThermoFisher, San Jose, CA) amb un FDR a l'1%. Es van seguir els mateixos passos que al capítol 1 per a la identificació i validació dels espectres obtinguts. Un resum comparatiu de les seqüències obtingudes es mostra a la taula 3.1

Mostra	Espectres identificats	Seqüències finals	Proteïnes (*)	% nesteds
HT93CIITA	482	153	82	33
HT93DT	795	292	166	22

(*) s'indica el nombre de grups de proteïnes calculat per Sequest (ver material i mètodes)

Taula 3.1: Resum comparatiu del nombre de seqüències obtingudes per a les línies transfectants HT93CIITA i HT93DT.

3.3.1. Descripció del repertori associat a HT93CIITA

Es van identificar 153 lligands corresponents a 82 proteïnes diferents (taula 3.2). Un 33% del repertori s'agrupava en *nested sets*. La distribució de la mida dels pèptids comprenia un rang d'entre 9-25 aminoàcids, trobant un percentatge major de pèptids entre 11 i 18 aminoàcids (figura 3.2). La massa mitjana dels pèptids fou de 1590 Da.

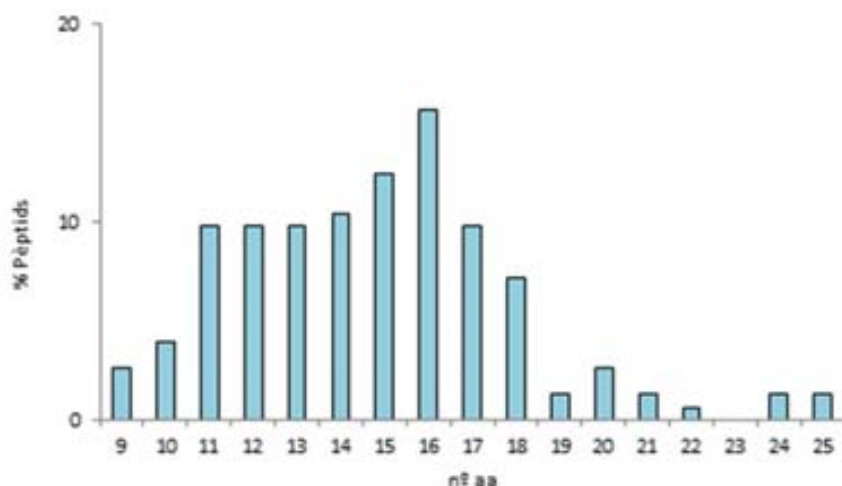


Figura 3.2: Distribució de les mides dels pèptids aïllats de la línia HT93CIITA

Taula 3.2: Seqüències dels pèptids associats a classe II de la mostra HT93CIITA. Uniprot AN: Identificació de la proteïna en SwissProt; Proteïna: proteïna origen de la qual prové el pèptid; Longitud: mida del pèptid en nombre d'aminoàcids; Localització: MB (membrana plasmàtica); Lis (lisosomes); End (endosomes); S (secretada); EM (matriu extracel·lular); ER (Reticle endoplàsmic); G (Golgi); MIT (mitochondria) C (citosol); N (nuclear).

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
Q15029	116 kDa U5 small nuclear ribonucleoprotein component	LAKQDVVLNYPM	12	N
Q15029	116 kDa U5 small nuclear ribonucleoprotein component	LELAKQDVVLNYPM	14	
Q9Y3D3	28S ribosomal protein S16, mitochondrial	SQKTDAEATDTEATET	16	MIT
P0CW22	40S ribosomal protein S17-like	QVTQPTVGMNFK	12	C
P0CW22	40S ribosomal protein S17-like	QVTQPTVGMNFKTPRGPV	18	
P39019	40S ribosomal protein S19	PGVTVKDVNQQE	12	N
P15880	40S ribosomal protein S2	VSVQRTQAPAVATT	14	C
P63220	40S ribosomal protein S21	AEVDKVTGRF	10	C
P10809	60 kDa heat shock protein, mitochondrial	DARALMLQGVDLLADA	16	MIT
P62750	60S ribosomal protein L23a	YDALDVANKIGII	13	C
P62888	60S ribosomal protein L30	PGDSDIIRSMPEQTGEK	17	C
P62888	60S ribosomal protein L30	AIIDPGDSDIIRSMPEQTGEK	21	
P62899	60S ribosomal protein L31	KNLQTVNVNVDEN	11	C
P62899	60S ribosomal protein L31	FKNLQTVNVNVDEN	12	
P62899	60S ribosomal protein L31	YVPVTTFKNLQTVNVNVDEN	18	
P62899	60S ribosomal protein L31	VTYVPVTTFKNLQTVNVNVDEN	20	
P62899	60S ribosomal protein L31	LVTYVPVTTFKNLQTVNVNVDEN	21	

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
Q02878	60S ribosomal protein L6	ALTNGIYPHKLVF	13	C
P32969	60S ribosomal protein L9	VSEKGTVQQADE	12	C
P11021	78 kDa glucose-regulated protein	SAGPPPTGEEDTAEKDEL	18	ER
P11021	78 kDa glucose-regulated protein	YGSAGPPPTGEEDTAEKDEL	20	
P60709	Actin, cytoplasmic 1	AGFAGDDAPR	10	C
P60709	Actin, cytoplasmic 1	MATAASSSSLEKS	13	
Q06278	Aldehyde oxidase	QIRNMASLGHI	12	C
P01009	Alpha-1-antitrypsin	GKVVNPTQK	9	EM
P08133	Annexin A6	DEYLKTTGKPIE	12	C
P08133	Annexin A6	DEYLKTTGKPIEAS	14	
Q9ULA0	Aspartyl aminopeptidase	FELFPSLSHNLVD	14	C
Q92499	ATP-dependent RNA helicase DDX1	GYLPNQLFRTF	11	N
Q92499	ATP-dependent RNA helicase DDX1	LHLGYLPNQLFRTF	14	
P04280	Basic salivary proline-rich protein 1	GPPPPPGKPQ	10	S
P35221	Catenin alpha-1	IMAQLPQEKA	11	C
Q14011	Cold-inducible RNA-binding protein	DSYDSYATHNE	11	N
P02452	Collagen alpha-1(I) chain	TVIEYKTTKTS	11	EM
P02452	Collagen alpha-1(I) chain	LPIIDVAPLDVGAPD	15	
P02452	Collagen alpha-1(I) chain	RLPIIDVAPLDVGAP	15	
P02452	Collagen alpha-1(I) chain	RLPIIDVAPLDVGAPD	16	
P02452	Collagen alpha-1(I) chain	TVIEYKTTKTSRLPII	16	
P02452	Collagen alpha-1(I) chain	SRLPIIDVAPLDVGAPD	17	
P02452	Collagen alpha-1(I) chain	TVIEYKTTKTSRLPIID	17	
P02452	Collagen alpha-1(I) chain	RLPIIDVAPLDVGAPDQE	18	
P02452	Collagen alpha-1(I) chain	TSRLPIIDVAPLDVGAPD	18	
P02452	Collagen alpha-1(I) chain	SRLPIIDVAPLDVGAPDQE	19	
P02452	Collagen alpha-1(I) chain	KTSRLPIIDVAPLDVGAPD	19	
P02452	Collagen alpha-1(I) chain	TSRLPIIDVAPLDVGAPDQE	20	
P08572	Collagen alpha-2(IV) chain	IPQKIAVQPGTVGPQ	15	EM
P08572	Collagen alpha-2(IV) chain	IPQKIAVQPGTVGPQG	16	
P08572	Collagen alpha-2(IV) chain	GIPQKIAVQPGTVGPQG	17	
P08572	Collagen alpha-2(IV) chain	GIPQKIAVQPGTVGPQGR	18	
Q15517	Corneodesmosin	GILNPSQPGQSSSSSQT	17	S
Q08554	Desmocollin-1	KPVDPDGPENGPP	13	MB
Q02413	Desmoglein-1	VVTGNMGSKNDKVGDF	15	MB
P24534	Elongation factor 1-beta	VQSMDDVAAFNKI	12	C
P24534	Elongation factor 1-beta	YVQSMDDVAAFNKI	13	
P24534	Elongation factor 1-beta	GFGDLKSPAGLQVLND	16	
Q96124	Far upstream element-binding protein 3	YGQTLGQAQAHSEQE	15	N
Q96124	Far upstream element-binding protein 3	FYGQTLGQAQAHSEQE	16	
P02751	Fibronectin	APSNLRFATTNPSL	15	EM
P02751	Fibronectin	APITGYRIVYSPSVEGS	17	
P21333	Filamin-A	AEIVEGENHTY	11	C
Q12789	General transcription factor 3C polypeptide 1	AEESDLSRQY	10	N
P15104	Glutamine synthetase	LNETGDPEFYQKN	13	MIT
P27144	GTP:AMP phosphotransferase AK4, mitochondrial	FSNKITPIQSKEAY	14	MIT
P08107	Heat shock 70 kDa protein 1A/1B	AQGPKGGSGSGPTIEVD	18	C
P08107	Heat shock 70 kDa protein 1A/1B	TPSYVAFTDTERL	13	
P11142	Heat shock cognate 71 kDa protein	GILNVSVDKSTGKE	15	C
P11142	Heat shock cognate 71 kDa protein	GILNVSVDKSTGKEN	16	
P69905	Hemoglobin subunit alpha	VLSPADKTNVKAA	13	C
P52597	Heterogeneous nuclear ribonucleoprotein F	YGAGYSQGNMGGYD	15	N
P55795	Heterogeneous nuclear ribonucleoprotein H2	FLNSTAGTSGGAYDHSY	17	N
P31942	Heterogeneous nuclear ribonucleoprotein H3	SVGRMGMGNNYSGGYG	16	N
Q9BUJ2	Heterogeneous nuclear ribonucleoprotein U-like protein 1	YGSYSGNTQGGTSTQ	15	N
Q9BUJ2	Heterogeneous nuclear ribonucleoprotein U-like protein 1	DYGSYSGNTQGGTSTQ	16	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	NQQPSNYGPMK	11	N
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	NQQPSNYGPMKSG	13	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	NQQPSNYGPMKSGNFG	16	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	GSRNMGGPYGGGNYGPGG	18	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	GPYGGGNYGPGGSGGSGYGGRSRY	25	

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	DDEKEAEEGEDDRDSANGEDDS	22	N
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	IKDDEKEAEEGEDDRDSANGEDDS	24	
Q86X55	Histone-arginine methyltransferase CARM1	MASPMISPTNTMHYGS	16	N
Q81ZD2	Histone-lysine N-methyltransferase 2E	NASKPTPAKV	10	N
P01909	HLA class II histocompatibility antigen, DQ alpha 1 chain	EDIVADHVASYGVN	14	MB
P01903	HLA class II histocompatibility antigen, DR alpha chain	GALANIAVDKANLE	14	MB
P01903	HLA class II histocompatibility antigen, DR alpha chain	AQGALANIAVDKANLE	16	
P01903	HLA class II histocompatibility antigen, DR alpha chain	EAQGALANIAVDKANLE	17	
P01903	HLA class II histocompatibility antigen, DR alpha chain	AQGALANIAVDKANLEI	17	
P01903	HLA class II histocompatibility antigen, DR alpha chain	EAQGALANIAVDKANLEI	18	
Q86Y23	Hornerin	GSXGSGWSSSRGPY	13	EM
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	TPLTSMVVTKEPE	12	S
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	VTPLTSMVVTKEPE	13	
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	TPLTSMVVTKEPEDN	14	
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	FVTPLTSMVVTKEPE	14	
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	VTPLTSMVVTKEPEDN	15	
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	FVTPLTSMVVTKEPEDN	16	
P00338	L-lactate dehydrogenase A chain	IVSGKDYNTANSKL	15	C
P19105	Myosin regulatory light chain 12A	LTTMGDRFTDE	11	EM
P35580	Myosin-10	DTEKTSQDVNETQPPQSE	18	C
P35579	Myosin-9	GKADGAEAKPAE	12	C
P35579	Myosin-9	EEVDGKADGAEAKPAE	16	
P41227	N-alpha-acetyltransferase 10	AEDENGKIVGY	11	N
P67809	Nuclease-sensitive element-binding protein 1	PPAENSSAPEAEQEG	14	C
P67809	Nuclease-sensitive element-binding protein 1	PPAENSSAPEAEQGGAE	17	
P67809	Nuclease-sensitive element-binding protein 1	GKETKAADPPAENSSAPEAEQGGAE	25	
Q99733	Nucleosome assembly protein 1-like 4	EEVHDLERKY	10	N
Q8N138	ORM1-like protein 3	GMVIFLHTV	9	ER
P05121	Plasminogen activator inhibitor 1	APEEIIIMDRPFLFVV	15	S
Q15366	Poly(rC)-binding protein 2	VRLSSETGGMGSS	13	N
Q6NZI2	Polymerase I and transcript release factor	SLLDKIIGA	9	ER
P01189	Pro-opiomelanocortin	YGGFMTSEKSTPLVT	16	S
Q15084	Protein disulfide-isomerase A6	VELDDLKDEL	11	MIT
Q15084	Protein disulfide-isomerase A6	DVELDDLKDEL	12	
Q15084	Protein disulfide-isomerase A6	SDVELDDLKDEL	13	
Q15084	Protein disulfide-isomerase A6	IDLSDELDDLKDEL	16	
P07237	Protein disulfide-isomerase	GSEIRLAKVDATEE	14	MIT
P07237	Protein disulfide-isomerase	EGSEIRLAKVDATEE	15	
P07237	Protein disulfide-isomerase	AEGSEIRLAKVDATEE	16	
P07237	Protein disulfide-isomerase	AEGSEIRLAKVDATEES	17	
P07237	Protein disulfide-isomerase	AEGSEIRLAKVDATEESD	18	
P05109	Protein S100-A8	MLTELEKALNSIIDV	15	C
Q92734	Protein TFG	YTQPGPGYR	9	C
Q92734	Protein TFG	QGYTQPGPGYR	11	
Q92734	Protein TFG	GQGYTQPGPGYR	12	
Q92734	Protein TFG	NRPPFGQGYTQPGPGY	16	
Q92734	Protein TFG	NRPPFGQGYTQPGPGYR	17	
Q94979	Protein transport protein Sec31A	KVVLTLQANKLGV	12	N
Q14568	Putative heat shock protein HSP 90-alpha A2	VDTGIGMTKADLINN	15	C
Q14568	Putative heat shock protein HSP 90-alpha A2	IVDTGIGMTKADLINN	16	
Q14568	Putative heat shock protein HSP 90-alpha A2	LVDTGIGMTKADLINN	16	
Q14568	Putative heat shock protein HSP 90-alpha A2	VDTGIGMTKADLINNL	16	
Q14568	Putative heat shock protein HSP 90-alpha A2	VDTGIGMTKADLINNLG	17	
A5PLK6	Regulator of G-protein signaling protein-like	FQGQLSPEEMLQCD	14	MB
Q9H299	SH3 domain-binding glutamic acid-rich-like protein 3	DGKRIQYQLVDISQDN	16	C
P29353	SHC-transforming protein 1	PSYVNVQNLDKARQ	14	C
P29353	SHC-transforming protein 1	DDPSYVNVQNLDKARQ	16	
P29353	SHC-transforming protein 1	DDPSYVNVQNLDKARQA	17	
P37108	Signal recognition particle 14 kDa protein	VLLSEQFLTELTRL	15	C
Q8TAQ2	SWI/SNF complex subunit SMARCC2	AVQGNLLPSASPLPD	15	N
Q86V81	THO complex subunit 4	LDAYNARMDS	11	N
Q86V81	THO complex subunit 4	AQLDAYNARMDS	13	
Q86V81	THO complex subunit 4	DAQLDAYNARMDS	14	

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
P02786	Transferrin receptor protein 1	NPGGVVAYSKAATVTGK	17	MB
P02766	Transthyretin	STTAVVTNPKE	11	C
P60174	Triosephosphate isomerase	FLVGGASLKPEFVDIINAKQ	20	C
P67936	Tropomyosin alpha-4 chain	RIQLVEEELDRAQER	15	C
P07437	Tubulin beta chain	EEYPDRIIMNTF	11	MB
P08670	Vimentin	FGGPGTASRPS	11	C
P08670	Vimentin	FGGPGTASRPSS	12	
P08670	Vimentin	FGGPGTASRPSSS	13	
P08670	Vimentin	YVTTSTRTYSLGSA	14	
P18206	Vinculin	ISPMVMDAKAVAGNI	15	C
P16989	Y-box-binding protein 3	GKEAKAGEAPTENPAPPTQQSSAE	24	C
Q92785	Zinc finger protein ubi-d4	EKASIYQNNSS	12	C

Uniprot AN: Identificació de la proteïna en SwissProt; Proteïna: proteïna origen de la qual prové el pèptid; Longitud: mida del pèptid en nombre d'aminoàcids; Localització: MB (membrana plasmàtica); Lis (lisosomes); End (endosomes); S (secretada); EM (matriu extracel·lular); ER (Reticle endoplàsmic); G (Golgi); MIT (mitocòndria) C (citosol); N (nuclear).

Els pèptids es van classificar en funció de la localització de la proteïna parental. El 67% dels pèptids seqüenciats procedien de la ruta degradació citosòlica, mentre que el 33% restant derivaven de la ruta de degradació endocítica (figura 3.3). Aquesta tendència ja l'havíem observat amb la resta de línies cèl·lules HLA-DR3.

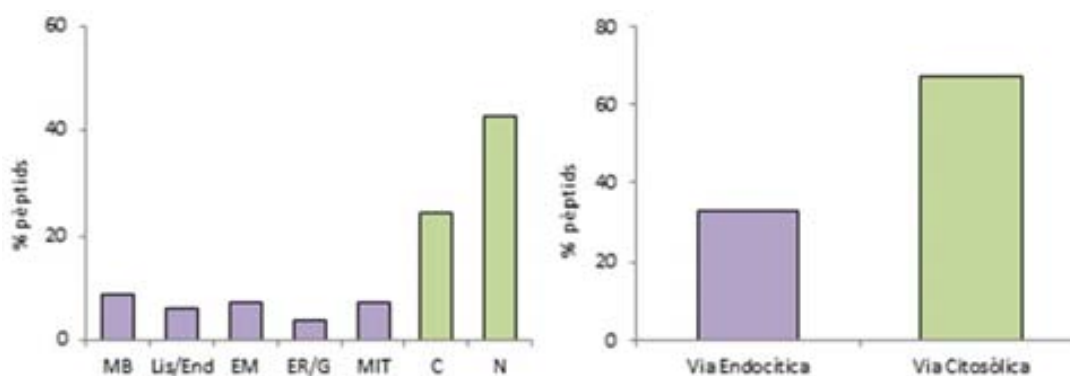


Figura 3.3: Localització de les proteïnes origen dels pèptids aïllats de HT93CIITA en funció de la seva localització cel·lular: MB (membrana plasmàtica); Lis/End (lisosomes o endosomes); EM (matriu extracel·lular); ER/G (Reticle endoplàsmic o Golgi); MIT (mitocòndria) C (citosol); N (nuclear). Al gràfic de l'esquerra, les localitzacions s'agrupen en funció de la seva via de processament, via endocítica (barres morades); C: via citosòlica (barres verdes).

Dins de la via endocítica, vam trobar proteïnes relacionades amb la immunitat i amb l'expressió de CIITA, com ara pèptids derivats de HLA-DR i HLA-DQ. Un

fet a destacar fou la identificació de 9 pèptids de diferent mida que pertanyien a un *nested set* de la proteïna del col·lagen alfa tipus I, procedent de la matriu extracel·lular.

En la via de degradació citosòlica, les proteïnes del citoplasma suposaven el 43% del repertori peptídic total i les nuclears el 24%. Els pèptids d'aquesta ruta que es van identificar eren principalment procedents de proteïnes que formen part de les subunitats dels ribosomes 40S i 60S, i diferents tipus de ribonucleoproteïnes (F, H2, H3, U, A2/B1, C1/C2). També trobarem pèptids específiques de tiroides com la proteïna TFG, associada al carcinoma papil·lar de tiroides i la transtiretina, proteïna d'unió a hormones del tiroides que transporta la tiroxina (T4) des del torrent sanguini fins al cervell.

3.3.2. Descripció del repertori associat a HT93DT

Es van identificar 292 lligands corresponents a 166 proteïnes diferents (taula 3.3). Aquesta mostra presentava un 22% del seu repertori agrupat en *nested sets*. La distribució de la mida dels pèptids comprenia un rang d'entre 9-25 aminoàcids, no havent tampoc una dominància clara d'una mida determinada (figura 3.4). La massa mitjana dels pèptids fou de 1491 Da.

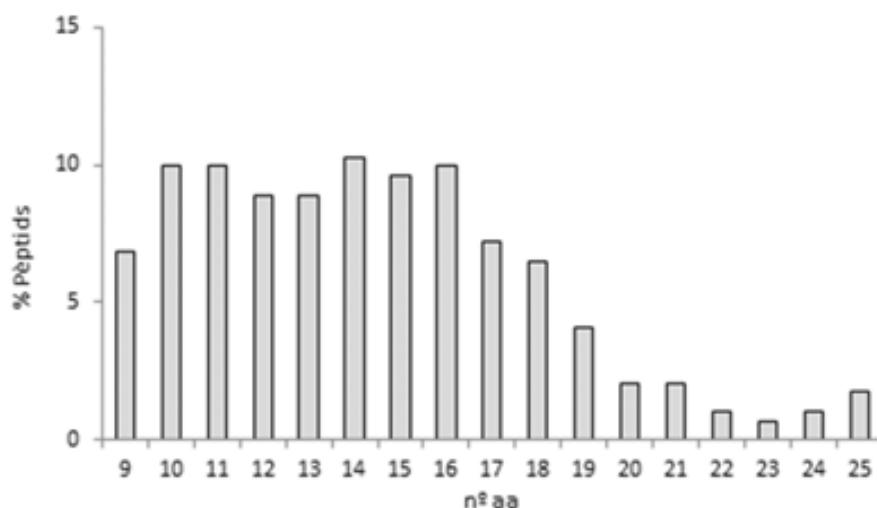


Figura 3.4: Distribució de les mides dels pèptids aïllats de la línia HT93CIITA-TSHR

Taula 3.3: Seqüències dels pèptids associats a classe II de la mostra HT93CIITA.

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
Q15029	116 kDa U5 small nuclear ribonucleoprotein component	AKQDVVLNYPM	11	N
Q15029	116 kDa U5 small nuclear ribonucleoprotein component	LELAQDQVVLNYPM	14	
Q9Y3D3	28S ribosomal protein S16, mitochondrial	SQKTDAAETDTEATET	16	MIT
Q9Y2Q9	28S ribosomal protein S28, mitochondrial	VGTESGSESGSSNAKEPKTR	20	MIT
P62249	40S ribosomal protein S16	PSKGPLQSVQVFR	14	C
P0CW22	40S ribosomal protein S17-like	QVTQPTVGMNFK	12	N
P0CW22	40S ribosomal protein S17-like	QVTQPTVGMNFKTPR	15	
P0CW22	40S ribosomal protein S17-like	QVTQPTVGMNFKTPRGPV	18	
P39019	40S ribosomal protein S19	PGVTVKDVNQOE	12	N
P15880	40S ribosomal protein S2	VQRTQAPAVATT	12	C
P15880	40S ribosomal protein S2	VSVQRTQAPAVATT	14	
P62851	40S ribosomal protein S25	AQVIYTRNTKGGDAPAAGEDA	21	C
P62081	40S ribosomal protein S7	LTGKDVNFEPFQQL	15	C
P46781	40S ribosomal protein S9	NAKKGQGGAGAGDDEED	18	C
Q9H0D6	5'-3' exoribonuclease 2	NSPGSQVASNPR	12	N
P84098	60S ribosomal protein L19	PNETNEIANANSR	13	C
P62750	60S ribosomal protein L23a	YDALDVANKIGII	13	C
P62750	60S ribosomal protein L23a	LAPDYDALDVANKIGII	17	
P83731	60S ribosomal protein L24	AITGASLADIMAK	13	C
P47914	60S ribosomal protein L29	RTQAPTQKASE	10	C
P47914	60S ribosomal protein L29	AAAPASVPAQAPK	13	
P47914	60S ribosomal protein L29	AQAAAPASVPAQAPK	15	
P47914	60S ribosomal protein L29	DQTKAQAAAPASVPAQAPK	19	
P62888	60S ribosomal protein L30	PGDSDIIRSMEQTEGK	17	C
P62899	60S ribosomal protein L31	NEDESPNKLY	11	C
P62899	60S ribosomal protein L31	KNLQTVNVNVDEN	11	
P62899	60S ribosomal protein L31	YVPVTTFKNLQTVNVNVDEN	18	
P62899	60S ribosomal protein L31	VTYVPVTTFKNLQTVNVNVDEN	20	
P62899	60S ribosomal protein L31	LVTYVPVTTFKNLQTVNVNVDEN	21	
Q02878	60S ribosomal protein L6	ALTNGIYPHKLVF	13	C
P62917	60S ribosomal protein L8	AVVGVVAGGGR	11	C
P32969	60S ribosomal protein L9	VSEKGTVQQADE	12	C
P11021	78 kDa glucose-regulated protein	SAGPPPTGEEDTAEKDEL	18	ER
P11021	79 kDa glucose-regulated protein	YGSAGPPPTGEEDTAEKDEL	20	
P60709	Actin, cytoplasmic 1	AGFAGDDAPR	10	C
P60709	Actin, cytoplasmic 1	MATAASSSSLEKS	13	
P54819	Adenylate kinase 2, mitochondrial	FLLDGFPRTV	10	MIT
P01009	Alpha-1-antitrypsin	TEEDFHVD	9	ER
Q9GZV1	Ankyrin repeat domain-containing protein 2	RTGQVEIVEHFLSLGL	16	C
P53680	AP-2 complex subunit sigma	VLNEYFHNV	9	MB
O14791	Apolipoprotein L1	ALADGVQKV	9	S
Q9ULA0	Aspartyl aminopeptidase	FELFPSLSHNLLVD	14	C
Q8WWM7	Ataxin-2-like protein	EENYGVKTTY	10	MB
Q8WWM7	Ataxin-2-like protein	GAPHPQVQMLH	12	
P56385	ATP synthase subunit e, mitochondrial	VPPVQVSPLIKLGRYSAL	18	MIT
Q92499	ATP-dependent RNA helicase DDX1	LHLGYLPNQL	10	N
Q92499	ATP-dependent RNA helicase DDX1	LGYPNQLFRFT	12	
Q92499	ATP-dependent RNA helicase DDX1	LHLGYLPNQLFRFT	14	
Q00571	ATP-dependent RNA helicase DDX3X	DYRQSSGASSSSFS	16	N
P31944	Caspase-14	KDSPQTIPTYTDAL	14	N
P35221	Catenin alpha-1	IMAQLPQEQA	11	C
P07858	Cathepsin B	LPASFDAREQ	10	LYS
Q9Y696	Chloride intracellular channel protein 4	FLDGNEMTL	9	MB
Q9ULV3	Cip1-interacting zinc finger protein	LLGPPPVGV	9	N
P09497	Clathrin light chain B	MSLKQTPLSR	10	N
P00451	Coagulation factor VIII	HQIALRMEVLGCEAQDL	17	S
P02452	Collagen alpha-1(I) chain	PIIDVAPLDVGAPD	14	EM
P02452	Collagen alpha-1(I) chain	LPIIDVAPLDVGAPD	15	
P02452	Collagen alpha-1(I) chain	RLPIIDVAPLDVGAPD	16	
P02452	Collagen alpha-1(I) chain	SRLPIIDVAPLDVGAPD	17	
P02452	Collagen alpha-1(I) chain	RLPIIDVAPLDVGAPDQE	18	
P02452	Collagen alpha-1(I) chain	TSRLPIIDVAPLDVGAPD	18	
P02452	Collagen alpha-1(I) chain	SRLPIIDVAPLDVGAPDQE	19	
P02452	Collagen alpha-1(I) chain	TSRLPIIDVAPLDVGAPDQE	20	
Q03692	Collagen alpha-1(X) chain	SEYVHSSFSGLVAPM	16	EM
P01024	Complement C3	VPVTVTVHD	9	S

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
P00751	Complement factor B	FKVKDMENLEDVFYQIMIDESQSLS	24	S
Q15021	Condensin complex subunit 1	SLAGDVALQQL	11	C
Q15517	Corneodesmosin	GILNPSQPGQSSSSSQT	17	S
P01040	Cystatin-A	IPGGLSEAKPATPEIQE	17	C
P15954	Cytochrome c oxidase subunit 7C, mitochondrial	ATPFLVVRHQLLKT	14	ER
Q9H147	Deoxynucleotidyltransferase terminal-interacting protein 1	EEIQTVFNKY	10	N
P81605	Dermcidin	SSLLEKGLD	9	S
Q08554	Desmocollin-1	KPVDPDGPENGPP	13	MB
Q02413	Desmoglein-1	VVTGNMGSDNDKVGDF	15	MB
Q9UQ16	Dynamin-3	TLIDLPGITKV	11	END
Q9H223	EH domain-containing protein 4	DEFSEAIKAF	10	MB
P24534	Elongation factor 1-beta	VQSM DVAAFNKI	12	C
P24534	Elongation factor 1-beta	YVQSM DVAAFNKI	13	
P24534	Elongation factor 1-beta	GFGDLKSPAGLQVLN	15	
P24534	Elongation factor 1-beta	GFGDLKSPAGLQVLND	16	
Q99613	Eukaryotic translation initiation factor 3 subunit C	SLDQPTQTV	9	C
O14980	Exportin-1	VLIDYQRNV	9	N
P55060	Exportin-2	AESIVVHTY	9	MB
Q8IWF2	FAD-dependent oxidoreductase domain-containing protein 2	NEASGLYQMF	10	ER
Q96124	Far upstream element-binding protein 3	QTLGQAQAHSQEQ	13	N
Q96124	Far upstream element-binding protein 3	YGQTLGQAQAHSQEQ	15	
Q96124	Far upstream element-binding protein 3	FYGGTLGQAQAHSQEQ	16	
P02671	Fibrinogen alpha chain	FSSANNRNDNTYNRVSED	17	S
P02675	Fibrinogen beta chain	PYKQGFGNVATNTD	14	S
P02679	Fibrinogen gamma chain	FLSTYQTKVD	10	S
P02751	Fibronectin	APITGYRIVYSPSVEGS	17	EM
Q5D862	Filaggrin-2	PVLKNPDDPDVTD	13	C
P21333	Filamin-A	AEIVEGENHTY	11	C
P21333	Filamin-A	IQDNHDGTYTV	11	
Q8TF40	Folliculin-interacting protein 1	EENAVDVVKQY	10	C
P13284	Gamma-interferon-inducible lysosomal thiol reductase	LLDVPTAAV	9	S
P13284	Gamma-interferon-inducible lysosomal thiol reductase	LLLDVPTAAVQA	12	
O75496	Geminin	YMAELIERL	9	N
Q12789	General transcription factor 3C polypeptide 1	AEESDLRQY	10	N
P15104	Glutamine synthetase	LNETGDEPFQYKN	13	MIT
P04406	Glyceraldehyde-3-phosphate dehydrogenase	DAGAGIALNDH	11	C
P00739	Haptoglobin-related protein	YAEVGRVGYVSG	12	S
P00739	Haptoglobin-related protein	AGSAFAVHDLIED	13	
P08107	Heat shock 70 kDa protein 1A/1B	AQGPKGSGSGPTIEVD	18	C
P04792	Heat shock protein beta-1	GPEAAKSDATAK	13	
P04792	Heat shock protein beta-1	AQLGGPEAAKSDATAK	17	
P54652	Heat shock-related 70 kDa protein 2	TPSYVAFTDTERLIGDA	17	C
Q9BYK8	Helicase with zinc finger domain 2	AEQTALMAKY	10	N
Q13151	Heterogeneous nuclear ribonucleoprotein A0	GGYGGGGYGGSSF	14	N
P09651	Heterogeneous nuclear ribonucleoprotein A1	SSGPYGGGGQYF	12	N
P09651	Heterogeneous nuclear ribonucleoprotein A1	NQGGYGGSSSSSY	14	
P09651	Heterogeneous nuclear ribonucleoprotein A1	SSGPYGGGGQYFAKP	15	
P09651	Heterogeneous nuclear ribonucleoprotein A1	RNQG YGGSSSSSY	15	
P09651	Heterogeneous nuclear ribonucleoprotein A1	SSGPYGGGGQYFAKPR	16	
P09651	Heterogeneous nuclear ribonucleoprotein A1	GGYGGSSSSSYGSGR	16	
P09651	Heterogeneous nuclear ribonucleoprotein A1	NQGGYGGSSSSSYGSG	17	
P09651	Heterogeneous nuclear ribonucleoprotein A1	NQGGYGGSSSSSYGSGR	18	
P09651	Heterogeneous nuclear ribonucleoprotein A1	RNQG YGGSSSSSYGSG	18	
P09651	Heterogeneous nuclear ribonucleoprotein A1	NNQSSNF GPMKGGNF GGR	18	
P09651	Heterogeneous nuclear ribonucleoprotein A1	NQGGYGGSSSSSYGSGRR	19	
P09651	Heterogeneous nuclear ribonucleoprotein A1	RNQG YGGSSSSSYGSGR	19	
P09651	Heterogeneous nuclear ribonucleoprotein A1	NQGGYGGSSSSSYGSGRRF	20	
P09651	Heterogeneous nuclear ribonucleoprotein A1	AKPRNQG YGGSSSSSYGSG	21	
P09651	Heterogeneous nuclear ribonucleoprotein A1	AKPRNQG YGGSSSSSYGSGR	22	
P51991	Heterogeneous nuclear ribonucleoprotein A3	SPYGGGYSGGGSGGY	16	N
P51991	Heterogeneous nuclear ribonucleoprotein A3	SSGSPYGGGYSGGGSGGY	19	
P51991	Heterogeneous nuclear ribonucleoprotein A3	SPYGGGYSGGGSGGYGSR	19	
P51991	Heterogeneous nuclear ribonucleoprotein A3	GSPYGGGYSGGGSGGYGSR	20	
P51991	Heterogeneous nuclear ribonucleoprotein A3	SGSPYGGGYSGGGSGGYGSR	21	
P51991	Heterogeneous nuclear ribonucleoprotein A3	SPYGGGYSGGGSGGYGSR	21	
P51991	Heterogeneous nuclear ribonucleoprotein A3	SSGSPYGGGYSGGGSGGYGSR	22	
P51991	Heterogeneous nuclear ribonucleoprotein A3	RSSGSPYGGGYSGGGSGGYGSR	23	
P52597	Heterogeneous nuclear ribonucleoprotein F	GAGYSGQNSMGGYD	14	N
P52597	Heterogeneous nuclear ribonucleoprotein F	YGAGYSGQNSMGGYD	15	

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
P14866	Heterogeneous nuclear ribonucleoprotein L	YGGGSEGGAPK	12	N
P14866	Heterogeneous nuclear ribonucleoprotein L	YGGGSEGGAPK	13	
P14866	Heterogeneous nuclear ribonucleoprotein L	VKMAAAGGGGGGGRY	15	
O43390	Heterogeneous nuclear ribonucleoprotein R	YQDTYQQQWK	10	N
Q9BUJ2	Heterogeneous nuclear ribonucleoprotein U-like protein 1	YGSYSNTQGGTSTQ	15	N
Q9BUJ2	Heterogeneous nuclear ribonucleoprotein U-like protein 1	DYGSYSNTQGGTSTQ	16	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	GNFGGSRNMGGP	12	N
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	NMGGPYGGGNYGP	13	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	SGNFGGSRNMGGP	13	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	GGGNGFGPGPSNFR	15	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	YGPGGSGSGGYGGR	15	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	YGPGGSGSGGYGGRS	16	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	NYGPGGSGSGGYGGR	16	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	NYGPGGSGSGGYGGRS	17	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	GGNYGPGSGSGSGGYGGR	18	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	NQQPSNYGPMKSGNFGGS	18	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	GGPYGGGNYGPGSGSGSGGYGGRS	24	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	MGGPYGGGNYGPGSGSGSGGYGGR	24	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	NMGGPYGGGNYGPGSGSGSGGYGGR	25	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	MGGPYGGGNYGPGSGSGSGGYGGRS	25	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	GGGNGFGPGPSNFRGGSDGYGSGR	25	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	GGYGGGPGYGNQQ	14	
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	GGYGGGPGYGNQGGG	16	
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	YSYPARVPPPPPIAR	15	N
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	DDEKEAEEGEDDRDSANGEDDS	22	
Q92522	Histone H1x	GAPAAATAPAPTAHKA	16	N
P62805	Histone H4	ISGLIYEETR	10	N
Q86X55	Histone-arginine methyltransferase CARM1	MASPMISPTNTMHYGS	16	N
P01889	HLA class I histocompatibility antigen, B-7 alpha chain	ALGFYPAEITL	11	MB
P01889	HLA class I histocompatibility antigen, B-7 alpha chain	SDSAQGSQVSLTA	13	
P04233	HLA class II histocompatibility antigen gamma chain CD74	MATPLLMAQALPMGALPQ	17	RE/G
P01909	HLA class II histocompatibility antigen, DQ alpha 1 chain	EDIVADHVASYG	12	MB
P01909	HLA class II histocompatibility antigen, DQ alpha 1 chain	EDIVADHVASYGVN	14	
P01903	HLA class II histocompatibility antigen, DR alpha chain	GALANIAVDKANLE	14	MB
P01903	HLA class II histocompatibility antigen, DR alpha chain	QGALANIAVDKANLE	15	
P01903	HLA class II histocompatibility antigen, DR alpha chain	AQALANIAVDKANLE	16	
P01903	HLA class II histocompatibility antigen, DR alpha chain	EAQALANIAVDKANLE	17	
P01903	HLA class II histocompatibility antigen, DR alpha chain	AQALANIAVDKANLEI	17	
P01903	HLA class II histocompatibility antigen, DR alpha chain	EAQALANIAVDKANLEI	18	
Q86Y23	Hornerin	GGSGWSSSRGPY	13	C
P01857	Ig gamma-1 chain C region	YTQKSLSLSPG	11	S
P01857	Ig gamma-1 chain C region	HYTQKSLSLSPG	12	
P01857	Ig gamma-1 chain C region	NYKTPPVLDSD	12	
P01834	Ig kappa chain C region	NALQSGNSQESVTEQD	16	EM
P01834	Ig kappa chain C region	NALQSGNSQESVTEQDSKD	19	
Q5TA45	Integrator complex subunit 11	DELEIKAYY	9	C
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	TPLTSMVVTKE	12	S
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	VTPLTSMVVTKE	13	
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	TPLTSMVVTKEPDN	14	
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	FVTPLTSMVVTKE	14	
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	VTPLTSMVVTKEPDN	15	
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	FVTPLTSMVVTKEPDN	16	
Q9H410	Kinetochore-associated protein DSN1 homolog	TEVKVEPMY	10	N
Q04760	Lactoylgutathione lyase	IEILNPNKMATLM	13	C
Q6PKG0	La-related protein 1	SLPTTVPEPNYRNT	15	C
O60341	Lysine-specific histone demethylase 1A	YTLPRQATPGVPAQQSPSM	19	N
P14174	Macrophage migration inhibitory factor	IVNTNVPRASVPDGLSEL	19	S
Q14764	Major vault protein	AQLELEVSKAQQLAEV	16	C
P40926	Malate dehydrogenase, mitochondrial	AKVAVLGASGGIGQPLSL	18	MIT
P43243	Matrin-3	GNLGAGNGLQGPR	14	N
P50579	Methionine aminopeptidase 2	YEVEIDGKTY	10	N
Q15014	Mortality factor 4-like protein 2	NEVVAGIKEY	10	C
P35579	Myosin-9	GKADGAEAKPAE	12	C
P35579	Myosin-9	GSDEEVDGKADGAEAKPAE	19	
P41227	N-alpha-acetyltransferase 10	AEDENGKIVGY	11	C
P55769	NHP2-like protein 1	TEADVNPAY	10	N
Q15233	Non-POU domain-containing octamer-binding protein	AAPGAEFAPNKR	12	N

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
P57740	Nuclear pore complex protein Nup107	AEDLFNRY	9	C
P67809	Nuclease-sensitive element-binding protein 1	PPAENSSAPEAEQGG	14	C
P67809	Nuclease-sensitive element-binding protein 1	PPAENSSAPEAEQGGG	15	
P67809	Nuclease-sensitive element-binding protein 1	PPAENSSAPEAEQGGGA	16	
P67809	Nuclease-sensitive element-binding protein 1	PPAENSSAPEAEQGGAE	17	
P67809	Nuclease-sensitive element-binding protein 1	GKETKAADPPAENSSAPEAEQGGAE	25	
Q99733	Nucleosome assembly protein 1-like 4	EEVHDLERKY	10	N
Q96RD7	Pannexin-1	EENISEVKSYS	10	MB
Q9Y2H2	Phosphatidylinositol phosphatase SAC2	LSLENLEKI	9	N
Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein	TDKSSASAPDVDDPEAFPALA	21	C
Q15366	Poly(rC)-binding protein 2	VRLSSETGGMGSS	13	C
P11940	Polyadenylate-binding protein 1	EAAQKAVNSATGVPTV	16	C
Q9H361	Polyadenylate-binding protein 3	AAAAAATPAVR	11	N
Q6NZI2	Polymerase I and transcript release factor	SLLDKIIGA	9	ER
P02545	Prelamin-A/C	SVGGSGGSGFGDNLVTRSY	19	N
Q92841	Probable ATP-dependent RNA helicase DDX17	SSQSSSQFSGIGR	14	N
Q15084	Protein disulfide-isomerase A6	VELDDLKDEL	11	MIT
Q15084	Protein disulfide-isomerase A6	DVELDDLKDEL	12	
Q15084	Protein disulfide-isomerase A6	SDVELDDLKDEL	13	
Q15084	Protein disulfide-isomerase A6	IDLSQVELDDLKDEL	16	
Q3MHD2	Protein LSM12 homolog	SQAQQPQKEAALSS	14	NA
P06703	Protein S100-A6	ALALIYNEAL	10	C
Q92734	Protein TFG	QGYTQPGPGYR	11	C
Q92734	Protein TFG	NRPPFGQGYTQPGP	14	
Q92734	Protein TFG	NRPPFGQGYTQPGPGY	16	
Q92734	Protein TFG	NRPPFGQGYTQPGPGYR	17	
Q94979	Protein transport protein Sec31A	KVVLQANKLGV	12	ER
P60468	Protein transport protein Sec61 subunit beta	PGPTPSGTNVGSSGR	15	ER
P60468	Protein transport protein Sec61 subunit beta	PGPTPSGTNVGSSGRSP	17	
P60468	Protein transport protein Sec61 subunit beta	PGPTPSGTNVGSSGRSPS	18	
P60468	Protein transport protein Sec61 subunit beta	PGPTPSGTNVGSSGRSPSKAVAA	23	
P00491	Purine nucleoside phosphorylase	MASIPLDKAS	11	C
Q14568	Putative heat shock protein HSP 90-alpha A2	VDTGIGMTKADLNN	15	C
Q14568	Putative heat shock protein HSP 90-alpha A2	IVDTGIGMTKADLNN	16	
Q14568	Putative heat shock protein HSP 90-alpha A2	LVDTGIGMTKADLNN	16	
Q14568	Putative heat shock protein HSP 90-alpha A2	VDTGIGMTKADLNNLG	17	
Q14568	Putative heat shock protein HSP 90-alpha A2	IVDTGIGMTKADLNNLG	18	
Q14568	Putative heat shock protein HSP 90-alpha A2	LVDTGIGMTKADLNNLG	18	
A5PLK6	Regulator of G-protein signaling protein-like	FQQQLSPEMLQCD	14	MB
Q9UKL0	REST corepressor 1	DEAPVLDVRY	10	N
P38159	RNA-binding motif protein, X chromosome	SAPSGPVRSSSGMG	14	N
P38159	RNA-binding motif protein, X chromosome	SAPSGPVRSSSGMG	15	
P42696	RNA-binding protein 34	VLFEINTDSVHL	11	N
Q9Y580	RNA-binding protein 7	TMDNMTSSAQIQR	14	N
P35637	RNA-binding protein FUS	PPSAKAAID	9	N
P35637	RNA-binding protein FUS	GPGGGPGGSHMGNGYGD	17	
P35637	RNA-binding protein FUS	GPGGGPGGSHMGNGYD	18	
P02810	Salivary acidic proline-rich phosphoprotein 1/2	RPQGPPQGQSPQ	12	S
Q9UL54	Serine/threonine-protein kinase TAO2	ILLSEPLVKL	11	C
P62714	Serine/threonine-protein phosphatase 2A catalytic subunit beta	MELDDTLKY	9	C
P02768	Serum albumin	KSLHTLFGD	9	S
P02768	Serum albumin	TLRETYGEMAD	11	
P02768	Serum albumin	KLVAASQAALGL	12	
P02768	Serum albumin	TVATLRETYGEMAD	14	
P37108	Signal recognition particle 14 kDa protein	ESEQFLTELTRL	12	C
P37108	Signal recognition particle 14 kDa protein	VLLSEQLTELTRL	15	
P42226	Signal transducer and activator of transcription 6	VERDQPLPTP	10	C
P62304	Small nuclear ribonucleoprotein E	GDNITLLQSVSN	12	C
Q13813	Spectrin alpha chain, non-erythrocytic 1	EEALILDNKY	10	C
P63208	S-phase kinase-associated protein 1	YLDIKLLDV	10	N
P23246	Splicing factor, proline- and glutamine-rich	GMGPGTPAGYGR	12	N
P23246	Splicing factor, proline- and glutamine-rich	GMGPGTPAGYGRG	13	
P23246	Splicing factor, proline- and glutamine-rich	FGQGGAGPVGGQGP	15	
P16949	Stathmin	HEAEVLKQL	9	C
P38646	Stress-70 protein, mitochondrial	ASEAIKAVVVGIDLG	15	MIT
Q92922	SWI/SNF complex subunit SMARCC1	GVPPPPAPGPPASAAP	16	N
O00186	Syntaxin-binding protein 3	ALGTDAGQKV	11	MB
Q9ULW0	Targeting protein for Xklp2	AEVEAQKAQQL	11	N

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
Q86V81	THO complex subunit 4	LDAYNARMDTS	11	N
Q86V81	THO complex subunit 4	AQLDAYNARMDTS	13	
Q86V81	THO complex subunit 4	DAQLDAYNARMDTS	14	
Q56UQ5	TPT1-like protein	QETSFTKEAY	10	EM
Q15370	Transcription elongation factor B polypeptide 2	VMKPKQDSGSSANEQAVQ	17	N
Q96QR8	Transcriptional activator protein Pur-beta	RGGSGGGEESEGEVED	19	N
P02786	Transferrin receptor protein 1	NPGGYVAYSKAATVTG	16	MB
Q8IW70	Transmembrane protein 151B	PPGSAAGESAAGGGGGGGPGVSEE	25	MB
P07437	Tubulin beta chain	EEYDPDRIMNTF	11	C
Q14157	Ubiquitin-associated protein 2-like	SQTSSIPQKPQTN	13	NA
Q9NWX9	Uncharacterized protein C14orf119	SEPDFVAKFY	10	MIT
Q08A18	Uncharacterized protein C2orf54	FASQPEAALR	10	NA
P08670	Vimentin	YVTTSTRYSL	11	C
P08670	Vimentin	FGGPGTASRPSSS	13	
P08670	Vimentin	YVTTSTRYSLGSA	14	
P08670	Vimentin	SLYASSPGGVYATR	14	
P08670	Vimentin	MFGGPGTASRPSSS	14	
P08670	Vimentin	DGQVINETSQHDDLE	16	
P18206	Vinculin	ISPMVMDAKAVAGNI	15	C
Q16864	V-type proton ATPase subunit F	NEIEDTFRQF	10	C
Q7Z2W4	Zinc finger CCCH-type antiviral protein 1	ATDLGGTSQAGTSQR	15	C
Q7Z2W4	Zinc finger CCCH-type antiviral protein 1	NNNADGVATDITSTR	15	

Uniprot AN: Identificació de la proteïna en SwissProt; Proteïna: proteïna origen de la qual prové el pèptid; Longitud: mida del pèptid en nombre d'aminoàcids; Localització: MB (membrana plasmàtica); Lis (lisosomes); End (endosomes); S (secretada); EM (matriu extracel·lular); ER (Reticle endoplàsmic); G (Golgi); MIT (mitochondria) C (citòsol); N (nuclear).

Com per a les cèl·lules HT93CIITA, els pèptids es van classificar en funció de la localització de la proteïna parental. El 67% dels pèptids seqüenciats procedien de la ruta citosòlica, mentre que el 33% restant derivaven de la ruta de degradació endocítica (figura 3.5). Aquesta tendència es característica d'HLA-DR3 (capítols 1 i 2) i s'observa també a la línia HT93CIITA.

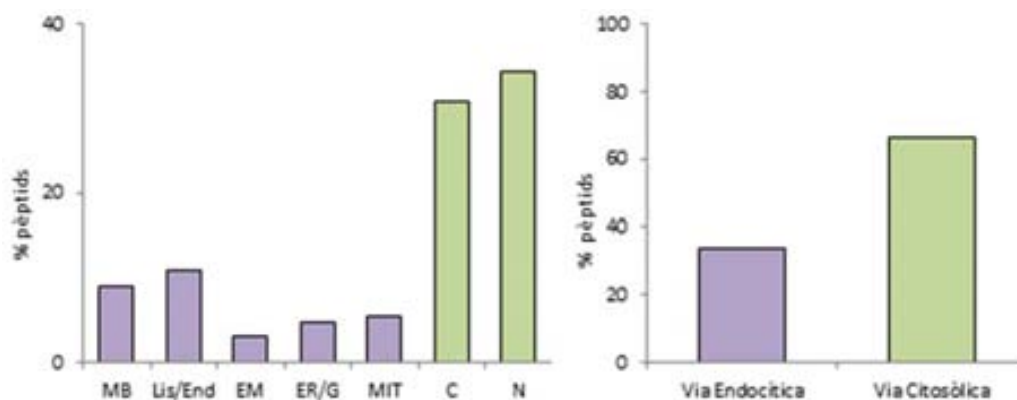


Figura 3.5: Dreta: Localització de les proteïnes origen dels pèptids aïllats de HT93CIITA en funció de la seva localització cel·lular: MB (membrana plasmàtica); Lis/End (lisosomes o endosomes); EM (matriu extracel·lular); ER/G (Reticle endoplàsmic o Golgi); MIT (mitochondria) C (citòsol); N (nuclear). Al gràfic de l'esquerra, les localitzacions s'agrupen en funció de la seva via de processament, via endocítica (barres morades); C: via citosòlica (barres verdes).

Dins de la via endocítica, vam trobar proteïnes associades a la membrana relacionades amb immunitat i l'expressió de CIITA: pèptids derivats de HLA-DR, DQ, CD74, HLA-classe I i de les cadenes gamma i kappa de les immunoglobulines. També en va identificar la proteïna tiol reductasa, que facilita el desplegament de les proteïnes destinades a la degradació lisosomal, jugant per tant un paper important en el processament d'antigen. A més, s'identificaren proteïnes de la via alternativa del sistema del complement, com C3 i el factor B. No obstant això, no és va aconseguir detectar cap pèptid derivat de TSHR.

Per la via citosòlica, les proteïnes del citoplasma suposaven el 31% del repertori peptídic total i les nuclears el 34%. D'aquestes, cal destacar 9 tipus diferents de ribonucleoproteïnes, on vàrem trobar famílies peptídiques composades per 17 pèptids diferents, localitzats en la part C-terminal de la proteïna, com en el cas de ribonucleoproteïna nuclear A2/B1.

3.4. PREDICCIÓ D'AFINITAT DELS LLIGANDS D'HLA-DR

Per fer la predicció d'afinitat dels pèptids a cada al·lel d'HLA-DR, es va utilitzar el programa *Laia Motifs*, tal i com es descriu al capítol 1, relacionant el grau de similitud de cada *core* amb els motius d'unió descrits a la literatura. En aquest cas per l'al·lel HLA-DR3 vam utilitzar el motiu descrit al capítol 1 i per HLA-DR7 el motiu definit a SYFPEITHY.

Al tractar-se d'una línia heterozigota, a cada *core* generat per pèptid (o *nested*) se li va assignar l'afinitat teòrica corresponent de cada al·lel i es va seleccionar el que mostrava una afinitat millor. En el cas d'obtenir un mateix grau d'afinitat, com que no eren capaços de discriminar a quin s'unia, el *core* es tenia en compte pels dos l'al·lells.

Per ambdues línies cel·lulars, la predicció teòrica va donar resultats semblants: predomini de pèptids amb afinitat intermèdia (IB). Aquest predomini d'unió intermèdia coincideix amb el que s'havia descrit al capítol 2 per lligands de DR3. El patró per HLA-DR7 era similar i quasi idèntic entre ambdues línies cel·lulars (figura 3.6).

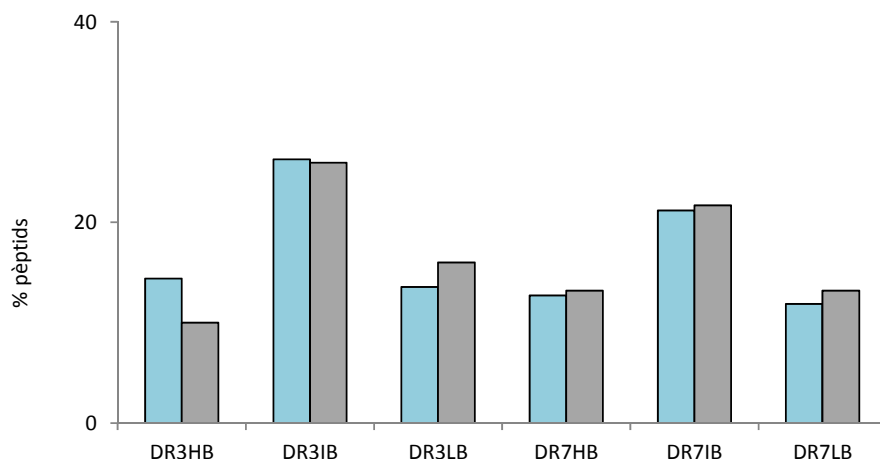


Figura 3.6: Comparació de l'anàlisi de l'afinitat d'unió teòrica dels pèptids a HLA-DR3 i HLA-DR7 entre la línia HT93CIITA (barres blaves) i HT93DT (barres gris)

3.5. PATRÓ DE PROCESSAMENT

Es va analitzar el possible patró de processament dels pèptids segons la seva localització respecte a la proteïna parental. Primer, es va fer un anàlisi del repertori total de pèptids de cada cèl·lula (figura 3.7). Així, vam observar que un 57% dels pèptids d'HT93CIITA corresponien a seqüències terminals (48% de C-ter, 9% de N-ter,) i 43% a seqüències internes. Per HT93DT, un 60% dels pèptids provenien d'una localització C-ter (48%) o N-ter (12%).

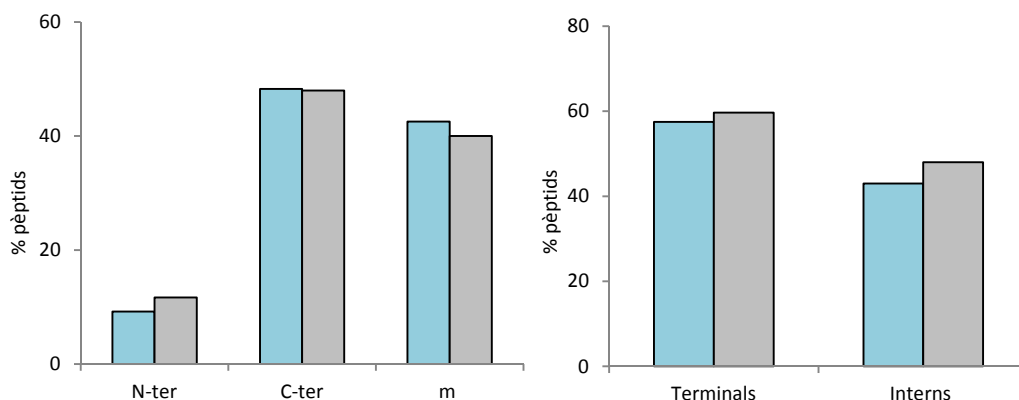


Figura 3.7: Comparació del patró de processament entre la línia HT93CIITA (barres blaves) i HT93DT (barres gris).

El següent anàlisi es va fer tenint en compte l'assignació de cada pèptid a l'al·lel DR3 o DR7 i els resultats d'afinitat presentats a l'apartat 3.5.

En el cas de HT93CIITA, HLA-DR3 unia majoritàriament pèptids C-ter independents de l'afinitat d'unió del pèptids, sent IB l'afinitat majoritària amb un 14%, un 11% van ser HB i 4% de LB. Un 25% eren lligands interns: i 11% eren HB, 13% IB i 2% de LB

Per DR7, es van obtenir majoritàriament pèptids de seqüències internes amb alta (13%), intermèdia (9%) i baixa afinitat (2%). També s'hi trobaren pèptids C-ter, un 16% repartit entre HB i IB (figura 3.8).

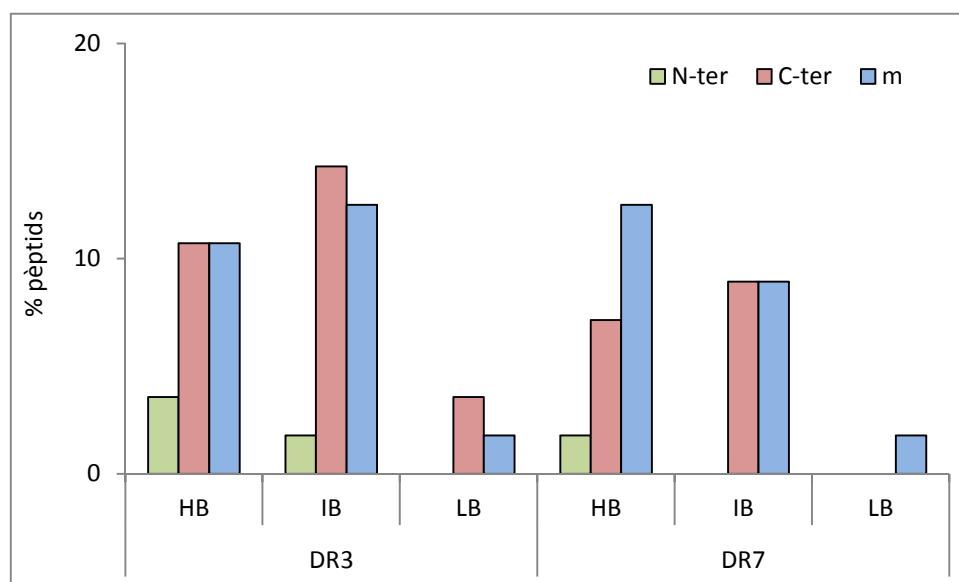


Figura 3.8: Gràfic on es relaciona el patró de processament dels lligands de HT93CIITA amb l'afinitat per l'al·lel al que s'uneix. N-ter: verd; C-ter: vermell; interns: blau

En el cas de HT93DT (figura 3.9), per HLA-DR3 obteníem un 28% de pèptids C-ter on un 13% eren IB, un 9% eren HB i un 6% LB. Un percentatge lleugerament inferior obteníem de lligands interns: 22%. La proporció menor eren pèptids N-ter: 1% HB, 5% IB, 1% LB.

Per HLA-DR7, un 21% eren seqüències internes: un 10% eren HB, un 9% corresponien a lligands IB i 2% eren LB. També trobàvem pèptids C-ter, amb

un percentatge semblant que per l'al·lel DR3 (17%): 8% d'HB i 87% IB. Finalment, sols un 7% eren pèptids N-ter, dels quals un 4% eren HB i un 3%IB.

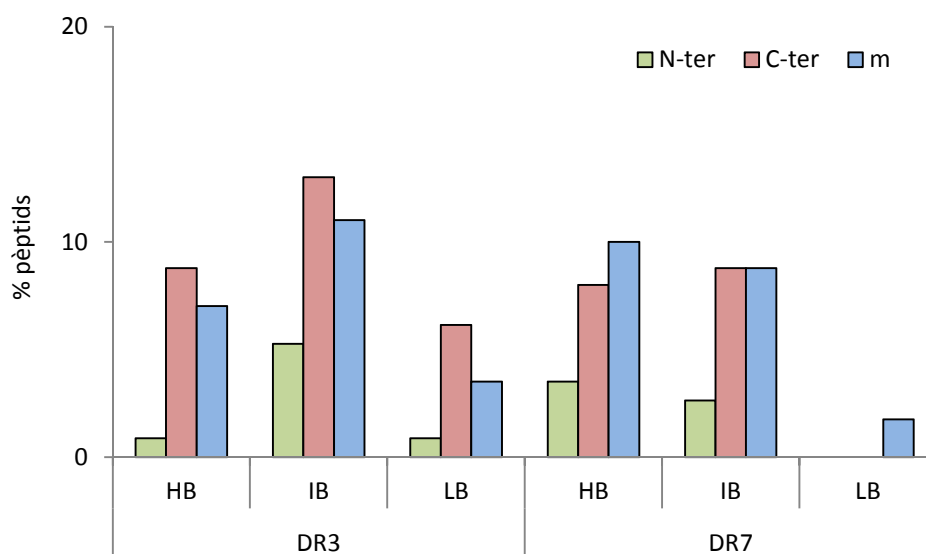


Figura 3.9: Gràfic on es relaciona el patró de processament dels lligands de HT93DT amb l'afinitat per l'al·lel al que s'uneix. N-ter: verd; C-ter: vermell; interns: blau

En conjunt les dues cèl·lules HT93CIITA i HT93DT presenten patrons semblants, on l'al·lel és el que fa les diferències. Per DR3 els pèptids són principalment IB i el patró dominant és C-ter. Per DR7, els pèptids dominants són interns i majoritàriament d'HB.

3.6. COMPARACIÓ ENTRE AMBDÓS REPERTORIS

Es van comparar els repertoris de pèptids entre ambdues línies. Respecte al nombre de seqüències peptídiques idèntiques, 88 seqüències eren comuns entre HT93CIITA i HT93DT, el que significava un 57% per a HT93CIITA i un 30% per a HT93DT. Pel que fa a les proteïnes parentals compartien 54, un 65% del repertori per CIITA i un 32% per a DT (figura 3.10)

Aquestes dades confirmen que, tal i com s'esperava, gran part del repertori present a les cèl·lules CIITA està present a les DT, encara que en aquestes últimes hi ha un repertori més ampli i diferenciat.

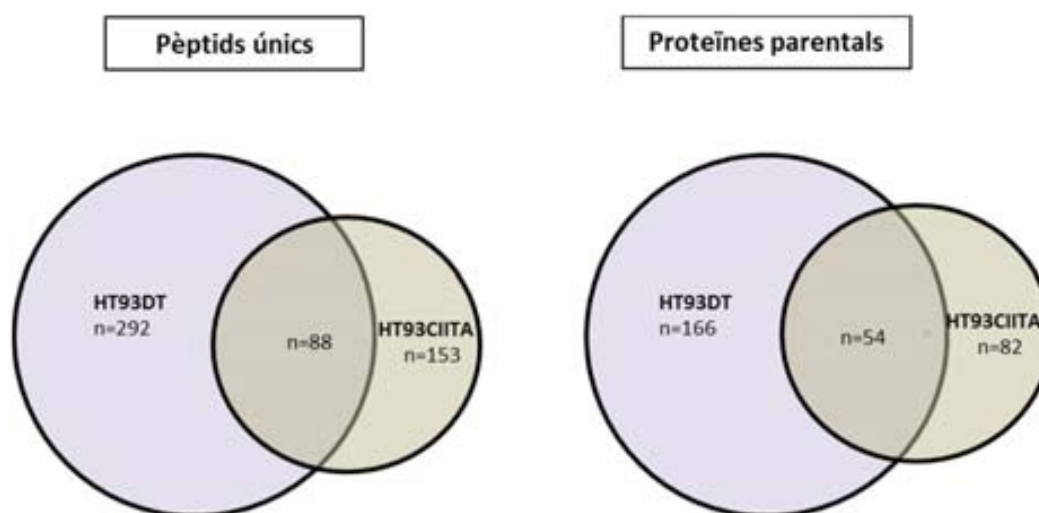


Figura 3.10: Diagrama de Venn que ens mostra el nombre de pèptids (esquerra) i proteïnes (dreta) que comparteixen ambdues línies.

3.7. OBSERVACIONS FINALS

El repertori peptídic associat a les línies transfectants HT93CIITA i HT93DT va mostrar unes característiques pràcticament iguals, tant pel que fa al tipus de pèptids presentats com a les proteïnes de les quals deriven. Així, el TSHR, en les condicions que s'expressava als transfectants, no va generar pèptids que poguessin ser detectats pel nostre sistema analític. Això podia ser degut a diverses raons:

- 1) Que el nostre model no tingués una expressió de TSHR suficient com per a ser presentat per la molècula d'HLA en les nostres condicions. Això no seria esperable ja que hem demostrat que el TSHR s'expressa a nivells molt alts (caracterització per citometria de flux, figura 3.1), molt més que a les cèl·lules del tiroide on els nivells d'expressió són mínims (ref).
- 2) Que la derivació del receptor cap als lisosomes (degut al LAMP unit al TSHR), inicialment pensada per afavorir la trobada dels pèptids amb HLA-DR i evitar la seva secreció, incrementi la degradació del TSHR i per tant ja no quedin seqüències reconeixibles per HLA en el lloc de càrrega dels pèptids.

3) HLA-DR3 es comporta en aquestes cèl·lules de la mateixa manera que en els estudis presentats anteriors (capítol 1 i 2), és a dir, presentant pèptids citosòlics i C-terminals preferentment.

4) Que el TSHR es comporti de manera diferent degut a la seva mida i diferent conformació extracel·lular del TSHR (eTSHR) i per tant, no generi pèptids presentables si no és a partir del processament del receptor sencer. Tot i així es va descartar la transfecció amb el TSHR sencer degut a la seva complexitat (7 segments transmembrana i 2 unitats addicionals, extracel·lular i intracel·lular).

En conclusió, no es van detectar pèptids derivats de TSHR presentats per HLA en aquest model. Una alternativa a aquest estudi seria la utilització d'un sistema lliure de cèl·lules (Cell-free System). En aquest mètode es fa una digestió *in vitro* de la proteïna amb diferents catepsines (B, H, S) i a diferents pHs (Hartman IZ et al, 2010) com a pas previ a la incubació amb la molècula d'HLA-DR. Aquesta tècnica s'està posant a punt per part de l'estudiant predoctoral Teresa Ciudad, on ja s'ha testat amb tiroglobulina i on s'esperen obtenir resultats positius.

No obstant això, hem demostrat que DR3 es comporta igual en el cas de ser una molècula d'expressió induïda que en el cas de tenir una expressió natural a la cèl·lula o de ser transfectada, com els altres casos presentats (capítol 1 i 2, respectivament): preferència per pèptids citosòlics i i localitzats a l'extrem c-terminal de les proteïnes.

OBJECTIU ESPECÍFIC 2:

PRESENTACIÓ DE PÈPTIDS PER LES MOLÈCULES DE CLASSE II DE L'HAPLOTIP H-2^d, ASSOCIAT A MALALTIES AUTOIMMUNITÀRIES TIPUS Th2

CAPITOLS:

4. Descripció del motiu d'ancoratge dels al·lels de classe II d'H-2^d: I-A^d i I-E^d
5. Estudi comparatiu entre els repertoris peptídics de timus i melsa en ratolins Balb/c

ANTECEDENTS

Els ratolins Balb/c mostren tendència a desenvolupar respostes immunes de tipus Th2 [126] També són propensos a algunes malalties, incloent el càncer de mama i altres tumors, així com malalties autoimmunes de tipus Th2, com ara la malaltia de Graves [127-130]. Degut a que la malaltia de Graves humana és causada directament per una resposta humoral i es va associar amb l'al·lèrgia estacional (nivells sèrics d'Ig E elevats, producció de citocines Th2) en general s'accepta que és una malaltia de tipus Th2. Els estudis realitzats en alguns models animals induïts de la malaltia de Graves recolzen encara més aquesta teoria [131].

Un requisit essencial per a l'estudi de qualsevol model de malalties mediatades per cèl·lules T és la identificació d'epítops de patògens, que necessita una clara definició de motius d'unió a pèptids dels al·lells de MHC implicats. Per tal de caracteritzar els motius associats als al·lells de MHC-II de la soca Balb/c, I-A^d i I-E^d, es farà servir el mètode establert al nostre laboratori basat en l'aïllament dels pèptids per immunoafinitat i l'anàlisi per espectròmetria de masses. A més a més, analitzarem les seqüències dels pèptids amb eines bioinformàtiques i es faran servir tècniques de modelatge i simulació biomoleculars per estudiar les interaccions entre els pèptids seleccionats i les molècules MHC.

Per una altra banda, ens estudis recents al nostre laboratori s'havia comparat el repertori de pèptids associats a HLA-DR en un òrgan secundari (melsa), comparat amb els d'un òrgan primari (timus), amb l'objectiu de revelar algunes característiques específiques d'aquells pèptids que indueixen tolerància i aquells que la mantenen a la perifèria. (Tesis doctoral JA Collado, UAB 2013). Així doncs, es va voler fer un estudi similar amb aquest teixits en ratolins Balb/c, per tal d'identificar similituds i diferències entre ambdós repertoris i comparar-ho amb el que ocorria en humans.

La molècula estructuralment homologa a HLA-DR en ratolí Balb/c és I-E i per tant, l'estudi comparatiu es basarà amb els resultats obtinguts per a aquesta

molècula. Tot i així, ja que disposem de les dades, farem un estudi addicional amb la molècula I-A^d, que es considera en realitat funcionalment homologa a DR, per detectar també diferències i semblances amb ambdós repertoris.

CAPÍTOL 4: DESCRIPCIÓ DEL MOTIU D'ANCORATGE DELS AL·LELS DE CLASSE II D'H-2^d: I-A^d i I-E^d

4.1. Processament de les mostres

L'anàlisi del repertori peptídic dels al·lels H-2^d es va fer per seqüenciació dels pèptids aïllats a partir de 40 mostres de timus i 40 de melses senceres de ratolins Balb/c. La identificació de pèptids es va realitzar mitjançant anàlisi per espectrometria de masses de alta resolució tal como s'indica a material i mètodes.

El protocol que es va utilitzar per a la purificació i fraccionament dels pèptids va ser semblant al que es feia servir per les línies en cultiu [84, 132] però amb algunes modificacions (veure materials i mètodes).

4.2. Repertori de pèptids associats a I-A^d i I-E^d

Es van identificar un total de 371 i 310 espectres de MS/MS per I-A^d i I-E^d, respectivament. Les seqüències repetides van ser excloses de l'anàlisi i les seqüències amb menys de 9 aminoàcids (13 per I-A^d i 29 per I-E^d) van ser eliminades perquè estaven per sota la mida canònica mínima per a la unió amb MHC-II. Així finalment, es van considerar 247 lligands únics naturals per I-A^d, corresponents a 148 proteïnes i 187 lligands per I-E^d, corresponents a 122 proteïnes. Només 41 proteïnes eren comuns entre ambdós repertoris. Una llista completa de les seqüències obtingudes i acceptades es mostren a les taules 4.1 i 4.2. Per a les dues molècules, el 70% del repertori estava compost per pèptids individuals i el 30% per pèptids que s'agrupava en *nested sets* que contenien entre 2-14 pèptids diferents.

Taula 4.1: Seqüències dels pèptids associats a classe II de la mostra I-E^d

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
B7ZW12	1110005A23Rik protein	IVTSSAGTGTTDEAK	17	N
B7ZW12	1110005A23Rik protein	GIVTSSAGTGTTDEAK	18	
P62259	14-3-3 protein epsilon	EAAENSLVAYKAASD	15	C
P62889	60S ribosomal protein L30	PGDSDIIRSMPEQTGEK	17	C
P62900	60S ribosomal protein L31	KNLQTVNVVDEN	11	C
P20029	78 kDa glucose-regulated protein	SGGPPPTGEEDTSEKDEL	18	END
F8WGM8	Actin, cytoplasmic 2	ISKQEYDESGPS	12	C
Q6IR43	Akap12 protein	SESNVEEMAQAASQAN	17	C
Q61838	Alpha-2-macroglobulin	LAYLTSESSRPT	12	S
D9IFT4	Alpha-dystroglycan	TPAREGAMSAQLGYVP	16	C
Q61649	Alpha-globin mRNA.	VLSGEDKSNIK	11	C
Q61649	Alpha-globin mRNA.	VLSGEDKSNIA	12	
Q61649	Alpha-globin mRNA.	VLSGEDKSNIAA	13	
A8DUV1	Alpha-globin	MFASFPTTKT	10	C
A8DUV1	Alpha-globin	IASVSTVLTSKY	12	
A7YL62	Apolipoprotein A-II	FSSLMNLEEKPAAPAA	15	EM
G3UZM8	Apolipoprotein E	ELEEQLGPVAEETR	14	S
G3UZM8	Apolipoprotein E	ELEEQLGPVAEETRA	15	
Q1L6K5	Apoptosis-inducing factor short isoform 2	ASSGSSGGKMDNSVLVL	17	C
O08691	Arginase-2, mitochondrial	DLVEVNPPLATSE	13	MIT
P97450	ATP synthase-coupling factor 6, mitochondrial	PKFEVIDKPGS	11	MIT
E9QAM0	Bec1-1	LLFLSPFLTKGILR	16	MB
Q8BTJ4	Bis(5'-adenosyl)-triphosphatase enpp4	LYEESHGIVANSMDYD	15	MB
Q9EQR4	Calcium-activated chloride channel regulator 2	IQAYNEAGLTSEVSN	15	MB
Q9EQR4	Calcium-activated chloride channel regulator 2	VEKESMVLGVTFDSAA	16	
Q9EQR4	Calcium-activated chloride channel regulator 2	IQAYNEAGLTSEVSNIA	17	
P26231	Catenin alpha-1	EQVEAAVEALSSDPA	15	C
Q6LAF6	Cathepsin B	VSNSVKIEIMAEIYKNGPV	18	END
Q545Y5	CD74 antigen	MATPLLMRPM	10	END
Q9ET41	Coatamer protein alpha subunit	DGPVRGIDF	9	C
Q99LL6	Col1a1 protein	GPAGPQQGR	9	EM
F6UF12	Collagen alpha-1(XX) chain	LDITVHNVLDFPQLDTLAPL	20	S
Q02105	Complement C1q subcomponent subunit C	FNSVVTNPQGHYNP	14	S
Q02105	Complement C1q subcomponent subunit C	FNSVVTNPQGHYNPS	15	
Q05CK6	Ctcf protein	TKIKKKKKKKKKKKK	16	N
P12787	Cytochrome c oxidase subunit 5A, mitochondrial	ISTPEELGDKV	12	MIT
P12787	Cytochrome c oxidase subunit 5A, mitochondrial	GISTPEELGDKV	13	
P12787	Cytochrome c oxidase subunit 5A, mitochondrial	LGISTPEELGDKV	14	
P12787	Cytochrome c oxidase subunit 5A, mitochondrial	NELGISTPEELGDKV	16	
P19536	Cytochrome c oxidase subunit 5B, mitochondrial	ASGGGVPTDEEQATGLER	18	MIT
P19536	Cytochrome c oxidase subunit 5B, mitochondrial	ASGGGVPTDEEQATGLERE	19	
P19536	Cytochrome c oxidase subunit 5B, mitochondrial	ASGGGVPTDEEQATGLEREIM	21	
P43024	Cytochrome c oxidase subunit 6A1, mitochondrial	NPHVNPLPTGYEDE	14	MIT
P48771	Cytochrome c oxidase subunit 7A2, mitochondrial	LAMAAFPKKQN	11	MIT
Q921Y3	Ddx5 protein	YDSTQQYGSNNVANM	14	N
Q9JME4	E25B protein	YQTIEENIKIFEED	14	G
O70251	Elongation factor 1-beta	VQSMDDVAFAFNKI	12	N
A8JYK8	Epoxide hydrolase 2C	DSLQMMCELSQHDFDLIESCQV	23	END
Q8R3F3	Fn1 protein	EPGTEYTIYVIALKNNQ	17	END
P15105	Glutamine synthetase	LNETGDEPFQYKN	13	C
P14435	H-2 class II histocompatibility antigen, A-F alpha chain	EDDIEADHVGFGY	13	MB
Q31135	H2-Ab1 protein	VMLEMTPHQ	9	MB
Q31135	H2-Ab1 protein	VLVMLEMTPH	10	
Q31135	H2-Ab1 protein	VLVMLEMTPHQ	11	
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	NQQPSNYGPMK	11	N
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	NQQPSNYGPMKSG	13	
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	SGNFGGSRNMGGP	13	
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	SNYGPMSGNFGGS	14	
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	SGSYNDFGNYNQQP	14	
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	NQQPSNYGPMKSGNFG	16	
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	NQQPSNYGPMKSGNFGGS	18	
A2AL13	Heterogeneous nuclear ribonucleoprotein A3	YGGGYSGGGSGGYG	15	N
A2AL13	Heterogeneous nuclear ribonucleoprotein A3	YGGGYSGGGSGGYGS	16	
A2AL13	Heterogeneous nuclear ribonucleoprotein A3	YNEGGNFGGGNYGGGNYNDFGNY	24	
H3BJ59	Heterogeneous nuclear ribonucleoprotein K	VPDSSGPERILSI	13	N
Q8VDM6	Heterogeneous nuclear ribonucleoprotein U-like protein 1	DYGGYSGTQGGTSTQ	16	N
Q9Z204	Heterogeneous nuclear ribonucleoproteins C1/C2	DLDDYDFQRD	9	N
Q1WWK3	Hist1h1b protein	AGGYDVEKNNS	11	N
Q1WWK3	Hist1h1b protein	VQTKGTGASGSF	12	
Q1WWK3	Hist1h1b protein	ALAAGGYDVEKNN	13	

Uniprot AN	Proteína	Seqüência	Longitud	Localització
P43274	Histone H1.4	SETAPAAPAAPAPAEKTPV	19	N
3274;Q1WW	Histone H1.4	GTLVQTKGTGASGS	14	
P10853	Histone H2B type 1-F/J/L	PEPAKSAPAP	10	N
P10853	Histone H2B type 1-F/J/L	STITSREIQT	11	
P01868	Ig gamma-1 chain C region secreted form	DVEVHTAQQT	10	S
P01863	Ig gamma-2A chain C region, A allele	NVEVHTAQTT	10	S
P01863	Ig gamma-2A chain C region, A allele	VNNNEVHTAQTT	12	
P01863	Ig gamma-2A chain C region, A allele	FVNNNEVHTAQTT	13	
P01863	Ig gamma-2A chain C region, A allele	YKNTPEVLDSGSGY	14	
P01863	Ig gamma-2A chain C region, A allele	SSGVHTFPAVL	11	
P01863	Ig gamma-2A chain C region, A allele	SGVHTFPAVLQ	11	
P01863	Ig gamma-2A chain C region, A allele	SSGVHTFPAVLQ	12	
P01863	Ig gamma-2A chain C region, A allele	SSGVHTFPAVLQSDLY	16	
P01863	Ig gamma-2A chain C region, A allele	SSGVHTFPAVLQSDLYTLS	19	
B2RWU2	Immunodeficiency virus type 1 enhancer binding protein 2	SDAEESDGEDGDDNDDDDDD	21	N
Q3V1M1	Immunoglobulin superfamily member 10	LTPAYYGR	9	S
P13020-2	Isoform 2 of Gelsolin	TASDFISKMQYPRQTQ	16	S
O88569-3	Isoform 3 of Heterogeneous nuclear ribonucleoproteins A2/B1	GNFGGSRNMGGP	12	C
Q99M83	Low-voltage-activated calcium channel alpha1.3 subunit	GFGSSAPEPQ	10	MB
B9EHV0	Lrrc7 protein	EYVQQPSKNIADKL	14	NA
Q9CQC8	Masparidin	SRLTLNCONSYVEPHKIDIPV	22	C
G5E8Z2	MCG113797	VSLQPENPVVSGAAVTLAIP	20	N
G3UW34	MCG130981	VSEKGTVQQADE	12	N
Q1XIR8	MHC class I heavy chain H2-K	DMELVETRPAGDG	13	MB
Q1XIR8	MHC class I heavy chain H2-K	LTQDMELVETRPAGDG	16	
Q9XRN9	MHC class II H2-E alpha chain	ALANIAVDKA	10	MB
Q9XRN9	MHC class II H2-E alpha chain	ALANIAVDKANLD	13	
Q9XRN9	MHC class II H2-E alpha chain	FEAQGALANIAVD	13	
Q9XRN9	MHC class II H2-E alpha chain	FEAQGALANIAVDK	14	
Q9XRN9	MHC class II H2-E alpha chain	FEAQGALANIAVDKA	15	
Q9XRN9	MHC class II H2-E alpha chain	SFEAQGALANIAVDK	15	
Q9XRN9	MHC class II H2-E alpha chain	ASFEAQGALANIAVD	15	
Q9XRN9	MHC class II H2-E alpha chain	ASFEAQGALANIAVDK	16	
Q9XRN9	MHC class II H2-E alpha chain	ASFEAQGALANIAVDKA	17	
Q31154	MHC H2-K-d transplantation antigen H2-Kd	EMELVETRPAGDG	13	MB
Q31154	MHC H2-K-d transplantation antigen H2-Kd	LTQEMELVETRPAGDG	16	
Q31154	MHC H2-K-d transplantation antigen H2-Kd	EPPSSTKTNTVIAVPVVLG	20	
Q6A0C8	MKIAA0113 protein	LNKALEEALSQASPSPP	19	C
P28665	Murine globulin-1	IPVEFSMVPMAKM	13	S
P28665	Murine globulin-1	IPVEFSMVPMAKML	14	
P28665	Murine globulin-1	IPVEFSMVPMAKMLI	15	
H3BL13	N-acyl ethanolamine-hydrolyzing acid amidase	EAAVYTLAKTPLIA	14	END
H3BL13	N-acyl ethanolamine-hydrolyzing acid amidase	EAAVYTLAKTPLIADV	16	
Q0IJ77	Npas3 protein	KLLPLPAITSQ	12	N
E9PUQ9	Piezo-type mechanosensitive ion channel component 1	KRRLWLASRFR	11	MB
Q6P1B3	PILR alpha-associated neural protein	GAPAFQLNRIPLVN	14	MB
F6VRE5	Piwi-like protein 4	YRDGVGNQGLKAVLE	15	N
F6QU68	Pleckstrin homology-like domain family B member 2	MGGKIKTWKRWVFDNRNKR	20	MB
Q4FK16	Plscr1 protein	IDQLLVHQJIE	11	NA
B2KX1	Polymorphic-like 2 (Drosophila)	PPQAIKVPQ	9	NA
Q61423	Potassium voltage-gated channel subfamily A member 4	GYAAQARARERER	13	MB
P01193	Pro-opiomelanocortin	YGGFMTSEKSTPLVT	16	S
A2AQ92	Protein Duox1	STGQGGPRGVYAF	13	MB
D3Z263	Protein Rtca	KDMAAAVRCIRKE	14	N
P27005	Protein S100-A8	DINSDNAINF	10	S
P27005	Protein S100-A8	PSELEKALSNI	11	
G3UYL5	Protein Snx21	DDDAEGLSSRLSGTSL	16	C
F6RYN6	Protein TBATA	PIGDPQSNRNPQL	13	N
E9Q1L0	Protein Vmn2r80	VKVGTFVPSA	10	MB
Q6PEM8	Proton-coupled folate transporter	LAMLMAISGIFNSIY	14	MB
Q8BPV5	Putative uncharacterized protein	TENEVEPVDARPAAD	15	END
Q8BPV5	Putative uncharacterized protein	TENEVEPVDARPAADR	16	
Q3UWH6	Putative uncharacterized protein	MDASHPSLQ	9	END
Q3UWH6	Putative uncharacterized protein	MEMNAFGDMTNEEFR	15	
Q3UFP0	Putative uncharacterized protein	TVAAPQLKTRPS	12	NA
Q9CV17	Putative uncharacterized protein	LQIDNARLAADD	13	C
Q9CTT4	Putative uncharacterized protein	ADNQTRDTQL	10	MIT
Q3TZN6	Putative uncharacterized protein	YPNVNIHNF	9	C
Q3V2J6	Putative uncharacterized protein	FINREANLQALIATG	15	NA
Q3V2J6	Putative uncharacterized protein	FLNREANLQALIATGG	16	
Q3V2J6	Putative uncharacterized protein	FINREANLQALIATGGD	17	
Q8BWM7	Putative uncharacterized protein	IAEMFLSNHIPLRIG	15	END

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
Q3U804	Putative uncharacterized protein	TQIMFETFNTPAM	13	C
Q3U804	Putative uncharacterized protein	TPAMYVAIQAVLS	13	
Q3U804	Putative uncharacterized protein	NTPAMYVAIQAVLS	14	
Q3U804	Putative uncharacterized protein	DFEQEMATAASSSSL	15	
Q3U804	Putative uncharacterized protein	DFEQEMATAASSSLE	16	
Q3U775	Putative uncharacterized protein	RIALENYLAALQSD	14	EM
Q3U775	Putative uncharacterized protein	RIALENYLAALQSDPP	16	
Q3UBR4	Putative uncharacterized protein	YSMGMHTYQVG	11	END
Q8BP77	Putative uncharacterized protein	SEILSDPSDDTKG	13	MB
Q3TGJ9	Putative uncharacterized protein	TASDFISKMQYPRQT	15	END
Q3UZN3	Putative uncharacterized protein	EVELQPVVE	9	NA
Q3UW00	Putative uncharacterized protein	LGPQRVSGL	9	C
Q8C2K0	Putative uncharacterized protein	IQNRQAQIEVVPSAS	15	C
P61226	Ras-related protein Rap-2b	KVVVLGSGG	9	END
Q569U7	Repin1 protein	LLIKKKKKK	10	C
Q5HZH2	Ribosome biogenesis protein TSR3 homolog	LDAFAEEVGAAALRASV	16	C
E0CY03	Selenoprotein P	ISYIVVNHQGSPTS	13	EM
O88697	Serine/threonine-protein kinase 16	FQKGDSDVALAVQNEL	15	C
G3UWL2	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	QHVKSALASVIM	12	END
Q921I1	Serotransferrin	LPEGTTPEKYLGAEYMQSVG	20	S
P07724	Serum albumin	KSLHTLFGD	9	S
P07724	Serum albumin	APQVSTPTLVEAARN	15	
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	GSITMQNIGAMSS	13	MB
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	GSITMQNIGAMSSYL	15	
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	SITMQNIGAMSS	12	
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	SITMQNIGAMSSYL	14	
Q8R1S9	Sodium-coupled neutral amino acid transporter 4	LPAAFYKLKLVKK	12	MB
O88307	Sortilin-related receptor	SDELTIVYKVNQLQ	13	MB
Q9QYT0	Squalene synthase	FSASEFEDP	9	MB
Q8CFZ0	SUMO-conjugating enzyme UBC9	VAVPTKNPDGTMNL	14	N
P09671	Superoxide dismutase [Mn], mitochondrial	DVTTQVALQPAKFNNGG	18	MIT
P09671	Superoxide dismutase [Mn], mitochondrial	GDVTTQVALQPAKFNNGG	19	
Q8VHA8	Synovial sarcoma associated SS18-delta	RPPQQGPPQ	9	C
Q6LBR0	T10 class I MHC gene (exon 5)	CPWSCGHQCCCGGFCE	16	NA
Q9D0B5	Thiosulfate sulfurtransferase/rhodanese-like domain-containing protein 3	FSCLAGVRSKKAM	13	NA
Q62WX2	Thymosin, beta 4, X chromosome	IEQEKQAGES	10	C
Q9QXE5	Thymus-specific serine protease	MAPMRPSDPS	11	C
Q9QXE5	Thymus-specific serine protease	DMAPMRPSDPSL	13	
Q9Z0U1	Tight junction protein ZO-2	DAPRSVLGRVKIF	13	MB
Q9WVA4	Transgelin-2	LINSLYPEGQAPVK	14	C
Q5R1C0	TRAV5-4	VLLDKKAKRFSL	12	NA
Q3UHC0	Trinucleotide repeat-containing gene 6C protein	ARTMQQPPVQ	10	N
O08712	Tumor necrosis factor receptor superfamily member 11B	LEMIGNQVQSVK	12	S
A2AKJ2	Vimentin	FGSGTSSRPSSN	13	C
P21614	Vitamin D-binding protein	SSSTFEQVNLQVKE	14	S

Uniprot AN: Identificació de la proteïna en SwissProt; Proteïna: proteïna origen de la qual prové el pèptid; Longitud: mida del pèptid en nombre d'aminoàcids; Localització: MB (membrana plasmàtica); Lis (lisosomes); End (endosomes); S (secretada); EM (matriu extracel·lular); ER (Reticle endoplàsmic); G (Golgi); MIT (mitochondria) C (citòsol); N (nuclear).

Taula 5.1: Seqüències dels pèptids associats a classe II de la mostra I-A^d

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
P62259	14-3-3 protein epsilon	EAAENSLVAYKAAAD	15	C
Q8K4S1	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase epsilon-1	DQECVFQAQSKWKG	14	C
B7ZMS6	2500003M10Rik protein	YMAQTDPETND	11	NA
Q3UHX2	28 kDa heat- and acid-stable phosphoprotein	SPPEIDAQLQAEKQKANE	18	N
Q62041	57-KD Calcium-binding protein (MCaBP)	DEKEDEEESPGQAKDEL	18	END
F8WGM8	Actin, cytoplasmic 2	KQEYDESGPS	10	C
F8WGM8	Actin, cytoplasmic 2	SKQEYDESGPS	11	
F8WGM8	Actin, cytoplasmic 2	ISKQEYDESGPS	12	
D3Z563	Acyl-CoA-binding protein	QATVGDVNTDRPGLDLK	18	MIT
E9QLQ9	Adenomatous polyposis coli protein	IHIPGLRNSSTSPVSKKGP	21	C
Q6IR43	Akap12 protein	ESNVEEMAQAAESQ	14	C
Q6IR43	Akap12 protein	SESNVEEMAQAAESQ	15	
Q6IR43	Akap12 protein	ESNVEEMAQAAESQAN	16	
Q6IR43	Akap12 protein	SESNVEEMAQAAESQAN	17	
Q6IR43	Akap12 protein	AQSSTEIPLQAESGGTGE	18	
Q6IR43	Akap12 protein	AQSSTEIPLQAESGGTGE	19	
Q6IR43	Akap12 protein	IPASESNVEEMAQAAESQAN	20	
Q6IR43	Akap12 protein	SESNVEEMAQAAESQANDVG	20	

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
Q61247	Alpha-2-antiplasmin	ELSEAGVEAAAATS	14	S
Q61247	Alpha-2-antiplasmin	MELSEAGVEAAAATS	15	
Q61838	Alpha-2-macroglobulin	MQQTLLTDQAVDK	14	S
A8DUV1	Alpha-globin	LASVSTVLTSKYR	13	C
D3Z1T0	Alpha-ketoglutarate-dependent dioxygenase alkB homolog 6	IMPHEDGPL	9	N
F7ANV6	Annexin	DIEQSIKSETSG	12	MB
Q07076	Annexin A7	TQMYQKTLSTMIASDT	16	N
B7STB7	Annexin	ENQEQEYVQAVKSYKGGPG	19	MB
Q9JL80	Anti-myosin immunoglobulin light chain variable region	ASNVESGVPARFSGSG	16	NA
A7YL62	Apolipoprotein A-II	MNLEKPAPAA	11	EM
E9Q1Y3	Apolipoprotein B-100	EPVNVNLGLEVNDVADKP	18	ER
G3UZM8	Apolipoprotein E	ELEEQLGPVAETRA	15	EM
P41233	ATP-binding cassette sub-family A member 1	NPSQYGITAFNHP	13	MB
P41233	ATP-binding cassette sub-family A member 1	ENPSQYGITAFNHP	14	
P09803	Cadherin-1	PRKDLEIGEYK	11	MB
Q9EQR4	Calcium-activated chloride channel regulator 2	IQAYNEAGLTSEVSN	15	MB
Q6GQS1	Calcium-binding mitochondrial carrier protein ScaMC-3	MWWKQLVAGAVAGAVS	16	MIT
Q9Z121	C-C motif chemokine 8	VLLKLIPLR	9	S
Q8VEI1	Cct6a protein	KDEPIDLFMVEIME	14	END
Q64692	CMP-N-acetylneuraminate-poly-alpha-2,8-sialyltransferase	DFITMNPVSVQRAFGGF	17	G
Q99LL6	Col1a1 protein	GPAGPQGP	9	EM
Q68G78	Csda protein	GEMKDGVP	9	N
P12787	Cytochrome c oxidase subunit 5A, mitochondrial	LGISTPEELGLDKV	14	MIT
P12787	Cytochrome c oxidase subunit 5A, mitochondrial	ELGISTPEELGLDKV	15	
P12787	Cytochrome c oxidase subunit 5A, mitochondrial	NELGISTPEELGLDKV	16	
P19536	Cytochrome c oxidase subunit 5B, mitochondrial	ASGGGVPTDEEQATGLER	18	MIT
P19536	Cytochrome c oxidase subunit 5B, mitochondrial	ASGGGVPTDEEQATGLERE	19	
P19536	Cytochrome c oxidase subunit 5B, mitochondrial	ASGGGVPTDEEQATGLEREIM	21	
Q8C9B9	Death-inducer obliterator 1	IKISSVHKRL	10	C
E9PZW0	Desmoplakin	VTGYNDPETGNI	12	C
Q9DD18	D-tyrosyl-tRNA(Tyr) deacylase 1	AEGDVSSEREP	11	C
Q9JME4	E25B protein	YQTIEENIKIFEED	14	G
Q9JME4	E25B protein	YQTIEENIKIFEEDA	15	
Q5SU72	E3 ubiquitin/ISG15 ligase TRIM25	AVVEQFLQAEQAR	13	C
Q5SU72	E3 ubiquitin/ISG15 ligase TRIM25	AVVEQFLQAEQARTP	15	
Q5SU72	E3 ubiquitin/ISG15 ligase TRIM25	AVVEQFLQAEQARTPV	16	
MOQWK5	Elongation factor 1-beta	GFGLKTPAGLQVLND	16	N
Q9CXC9	ETS translocation variant 5	PLQMPKMMPEQYPSE	16	N
Q99K47	Fibrinogen, alpha polypeptide	VIEKAQQIQALQSN	14	S
Q99K47	Fibrinogen, alpha polypeptide	VIEKAQQIQALQSNV	15	
Q9WVA7	FRAT3	PPRVLAALP	9	NA
P15105	Glutamine synthetase	LNETGDEPFQYKN	13	MIT
V9GZG5	Glutamine synthetase	LLNETGDEPFQYKN	14	
Q8VEB4	Group XV phospholipase A2	RQDTEGLVEAMTPP	14	END
Q8VEB4	Group XV phospholipase A2	RQDTEGLVEAMTPPG	15	
P14435	H-2 class II histocompatibility antigen, A-F alpha chain	EDDIEADHVG	11	MB
P14435	H-2 class II histocompatibility antigen, A-F alpha chain	EDDIEADHVG	12	
P14435	H-2 class II histocompatibility antigen, A-F alpha chain	EDDIEADHVG	13	
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	GNFGGSRNMGGP	12	N
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	SGNFGGSRNMGGP	13	
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	YDNYGGGNYGSGS	13	
B7ZP22	Heterogeneous nuclear ribonucleoprotein A2/B1	NQQPSNYGPMKSGN	14	
A2AL13	Heterogeneous nuclear ribonucleoprotein A3	YGGGYSGGGSGGYG	15	N
A2AL13	Heterogeneous nuclear ribonucleoprotein A3	YGGGYSGGGSGGYGS	16	
G3XA10	Heterogeneous nuclear ribonucleoprotein U	GAAGAADAGAMEEEE	15	N
G3XA10	Heterogeneous nuclear ribonucleoprotein U	NGAAGAADAGAMEEEE	16	
G3XA10	Heterogeneous nuclear ribonucleoprotein U	RPAMEPGNGSLDLGGDAAG	19	
G3XA10	Heterogeneous nuclear ribonucleoprotein U	LGEENGAAGAADAGAMEEEE	20	
Q8VDM6	Heterogeneous nuclear ribonucleoprotein U-like protein 1	DYGGYSTGGTSTQ	16	N
Q1WWK3	Hist1h1b protein	AAGGYDVEKNNS	12	N
Q1WWK3	Hist1h1b protein	GTLVQTKGTGASGS	14	
Q1WWK3	Hist1h1b protein	SETAPAETAAPAPVEK	16	
Q91VF2	Histamine N-methyltransferase	ASAKILILVSGT	13	N
P10853	Histone H2B type 1-F/J/L	PEPAKSAPAP	10	N
P70351	Histone-lysine N-methyltransferase EZH1	DMKERYRELTMSD	14	N
P01863	Ig gamma-2A chain C region, A allele	SSGVHTFPAVLQ	12	S
P01788	Ig heavy chain V region H8	IYYCARDYYGBS	12	MB
P01837	Ig kappa chain C region	TDQDSKDYTSMS	13	S
P01668	Ig kappa chain V-III region PC 7210	ASNLESIGIPARFSGS	15	MB
P01668	Ig kappa chain V-III region PC 7210	ASNLESIGIPARFSGSG	16	
P01668	Ig kappa chain V-III region PC 7210	AASNLESIGIPARFSGSG	17	

Uniprot AN	Proteína	Seqüência	Longitud	Localització
P01872	Ig mu chain C region secreted form	LPQEKYVTSAPMPGAPG	19	
K7TRG2	IgM heavy chain variable region	YLQMNTLRAEDSAT	14	S
K7TRG2	IgM heavy chain variable region	LYLQMNTLRAEDSAT	15	
K7THH7	IgM heavy chain variable region	VYLQMNTLRAEDTGIY	16	
D3Z2U5	Inactive serine/threonine-protein kinase VRK3	WKKQFLPL	10	N
B2RWT2	Integrin alpha E, epithelial-associated	DFLLVAAPFY	10	MB
Q61703	Inter-alpha-trypsin inhibitor heavy chain H2	DPSKLTEVQSIITATSANT	19	S
Q8R412	Interferon alpha-inducible protein 27-like protein 2A	MMSAAAIANGGGVAAAGSL	18	MB
Q3TAS6-2	Isoform 2 of ER membrane protein complex subunit 10	TALRPAAPGPAPAPTEA	17	ER
P42703	Leukemia inhibitory factor receptor	LIVYWKPLPI	10	S
B9EHV0	Lrrc7 protein	EYVQQPSKNIADKL	14	NA
Q8VCD7	Lysine-specific demethylase 4C	RQRVLSSRLKN	11	N
A11314	Macrophage-expressed gene 1 protein	DNQNSQNTVTASAGIA	16	MB
K3W4N9	MAGUK p55 subfamily member 4	AMRIVCLVKNQ	11	C
G3UW34	MCG130981	VSEKGTVQQADE	12	NA
Q8VCB2	Mediator of RNA polymerase II transcription subunit 25	RPQNPGANPQLR	12	N
Q1XIR8	MHC class I heavy chain H2-K	TQDMELVETRPAGDG	15	MB
Q1XIR8	MHC class I heavy chain H2-K	LTQDMELVETRPAGDG	16	
Q1XIR8	MHC class I heavy chain H2-K	DLTQDMELVETRPAGDG	17	
Q9XRN7	MHC class II H2-E alpha chain	FEAQGALANIAVDK	14	MB
Q9XRN7	MHC class II H2-E alpha chain	FEAQGALANIAVDKA	15	
Q9XRN7	MHC class II H2-E alpha chain	SFEAQGALANIAVDK	15	
Q9XRN7	MHC class II H2-E alpha chain	ASFEAQGALANIAVDK	16	
Q9XRN7	MHC class II H2-E alpha chain	ASFEAQGALANIAVDKA	17	
Q9XRN7	MHC class II H2-E alpha chain	ALANIAVDKA	10	
Q9XRN7	MHC class II H2-E alpha chain	FEAQGALANIAVD	13	
Q9XRN7	MHC class II H2-E alpha chain	SFEAQGALANIAVD	14	
Q9XRN7	MHC class II H2-E alpha chain	ASFEAQGALANIAVD	15	
Q9XRN7	MHC class II H2-E alpha chain	ASFEAQGALANIAVDKAN	18	
Q31154	MHC H2-K-d transplantation antigen H2-Kd	TQEMELVETRPAGDG	15	MB
Q31154	MHC H2-K-d transplantation antigen H2-Kd	LTQEMELVETRPAGD	15	
Q31154	MHC H2-K-d transplantation antigen H2-Kd	LTQEMELVETRPAGDG	16	
Q31154	MHC H2-K-d transplantation antigen H2-Kd	ELTQEMELVETRPAGDG	17	
Q3TZK4	Mitochondrial inner membrane protein	SSVTLQTITAQNAAVQ	16	MIT
G5E889	Mitochondrial translation optimization 1 homolog (S. cerevisiae), isoform CRA_b	LKSIKFSSKWK	13	MIT
Q99JJ3	Myh9 protein	LLAEKTISAKYAE	15	C
O09111	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 11, mitochondrial	YFDPKIQLPEDD	13	MIT
Q8R0W6	NEDD4 family-interacting protein 1	APPPYSSITAESAA	14	END
Q6R891	Neurabin-2	NKERMEKLEGYWGE	14	C
H3BK79	Nucleobindin-1	GVLDQELEALFTKE	15	G
H3BK79	Nucleobindin-1	DGVLDQELEALFTKE	16	
Q7TSZ8	Nucleus accumbens-associated protein 1	HKVLLRRL	9	N
Q9ERY0	Odorant receptor	VLLSLHLVL	10	MB
Q3U1F9	Phosphoprotein associated with glycosphingolipid-enriched microdomains 1	DPTLTFFEEISAMYSSVN	17	MB
Q3U1F9	Phosphoprotein associated with glycosphingolipid-enriched microdomains 1	DPTLTFFEEISAMYSSVKNPG	20	
P97290	Plasma protease C1 inhibitor	ELTESGVAAAAASA	14	S
P97290	Plasma protease C1 inhibitor	LELTESGVAAAAASA	15	
Q9WU78	Programmed cell death 6-interacting protein	FEKMPVPSVQQSL	13	C
Q9WU78	Programmed cell death 6-interacting protein	LFKMPVPSVQQSL	14	
Q9WU78	Programmed cell death 6-interacting protein	LFKMPVPSVQQSLA	15	
P01193	Pro-opiomelanocortin	YGGFMTSEKSTPLVT	16	S
Q922R8	Protein disulfide-isomerase A6	IDLSDVELDDLEKDEL	16	ER
D3Z5B9	Protein ERGIC-53	NNPAIVVIGNNGQIN	15	END
Q8CB11	Protein Gm10129	VAVGLVVD	9	MB
A2AFI6	Protein Gm364	LALFIPAGYMSAK	13	MB
F6SVV1	Protein Gm9493	AQLRELNITAAKEIE	15	C
Q3TLH4	Protein PRRC2C	LPTSAAERISALESQPA	17	NA
D3Z263	Protein Rtca	KDMAAAVRCIRKE	14	N
P27005	Protein S100-A8	PSELEKALSNI	11	S
F6RYN6	Protein TBATA	PIGDQSNRNPQL	13	N
P26350	Prothymosin alpha	VVEEAENGRDAPAN	14	N
P26350	Prothymosin alpha	ADNEVDEEEEGGEE	15	
P26350	Prothymosin alpha	QEAADNEVDEEEEGGEE	17	
P26350	Prothymosin alpha	NGRDAPANGNAQNEENGE	18	
P26350	Prothymosin alpha	GRDAPANGNAQNEENGEQE	19	
P26350	Prothymosin alpha	NGRDAPANGNAQNEENGEQE	20	
Q5XFY7	Ptpcr protein	SAEEPEHAANGSAPAPTQ	19	EM
P63054	Purkinje cell protein 4	APETERAAVAIQSQ	14	C
Q8BPV5	Putative uncharacterized protein	TENEVEPVDARPAAD	15	ER
Q8BPV5	Putative uncharacterized protein	TENEVEPVDARPAADR	16	
Q8BPV5	Putative uncharacterized protein	VPANTENEVEPVDARPAADR	20	

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
Q3UU71	Putative uncharacterized protein	VPAGQAANIVAHRSQEM	17	MB
Q8C094	Putative uncharacterized protein	RPKYPGKIFEE	11	C
Q8BNW4	Putative uncharacterized protein	ILSAMLISAFFQK	13	N
Q8BNW4	Putative uncharacterized protein	GILSAMLISAFFQK	14	
Q3TZN6	Putative uncharacterized protein	YPNVNIHNF	9	C
Q3V2J6	Putative uncharacterized protein	FLNREANLQALIATG	15	NA
Q3V2J6	Putative uncharacterized protein	FINREANLQALIATGG	16	
Q3V2J6	Putative uncharacterized protein	FINREANLQALIATGGD	17	
Q3UD36	Putative uncharacterized protein	VETRDGGVINETSQ	14	C
Q3TUW1	Putative uncharacterized protein	LEGDAMNIE	9	C
Q8BRJ5	Putative uncharacterized protein	PPAENSSAPEAEQGGAE	17	C
Q3U804	Putative uncharacterized protein	ISKQEYDESGP	11	C
Q3U804	Putative uncharacterized protein	TPAMYVAIQAVLS	13	
Q3U804	Putative uncharacterized protein	DFEQEMATAASSS	13	
Q3U804	Putative uncharacterized protein	NTPAMYVAIQAVLS	14	
Q3U804	Putative uncharacterized protein	DFEQEMATAASSS	14	
Q3U804	Putative uncharacterized protein	LDFEQEMATAASSS	15	
Q3U804	Putative uncharacterized protein	DFEQEMATAASSSSL	15	
Q3U804	Putative uncharacterized protein	FEQEMATAASSSSLE	15	
Q3U804	Putative uncharacterized protein	DFEQEMATAASSSSLE	16	
Q3U804	Putative uncharacterized protein	LDFEQEMATAASSSSLE	17	
Q3TGE1	Putative uncharacterized protein	DREVGIPPEQSLE	13	N
Q3V2J3	Putative uncharacterized protein	ILILLILLILLILLGR	15	C
Q3U775	Putative uncharacterized protein	IAENYLAALQSD	13	EM
Q3U775	Putative uncharacterized protein	IAENYLAALQSDPP	15	
Q3U775	Putative uncharacterized protein	RIALENYLAALQSDPP	16	
Q3UBR4	Putative uncharacterized protein	YSMGMTYQVG	11	END
Q3UGI7	Putative uncharacterized protein	ATSYGQPPTGYSTAPQA	19	N
Q3TGJ9	Putative uncharacterized protein	VPFDAATLHTSTAMA	15	END
Q3UW00	Putative uncharacterized protein	LGPQRVSGL	9	C
Q9CXX1	Putative uncharacterized protein	IDIYEQVGTSAAMDSP	15	C
Q3TGD3	Putative uncharacterized protein	HSQDQDAEDGALQ	13	N
Q8C2K0	Putative uncharacterized protein	IQNRQAQIEVVPSAS	15	END
Q8C2K0	Putative uncharacterized protein	IQNRQAQIEVVPSASA	16	
Q8CEW9	Putative uncharacterized protein	MLQPSGPPADKPEEN	15	NA
Q8BYD3	Putative uncharacterized protein	LPMWAPLPG	9	N
F6ZN61	Rho guanine nucleotide exchange factor 1	SSIREPLSSSE	12	MB
Q5HZH2	Ribosome biogenesis protein TSR3 homolog	LDFAEEVGAAALRASV	16	C
D6RG56	Round spermatid basic protein 1	FVLKKIKKKKKK	13	N
Q8K1R7	Serine/threonine-protein kinase Nek9	STVTEAPIAVVTSRT	15	N
Q921I1	Serotransferrin	TTPEKYLGAEYMQS	14	S
Q921I1	Serotransferrin	GTTPEKYLGAEYMQSVG	17	
Q921I1	Serotransferrin	LPEGTTPEKYLGAEYMQ	17	
Q921I1	Serotransferrin	LPEGTTPEKYLGAEYMQS	18	
Q921I1	Serotransferrin	LPEGTTPEKYLGAEYMQSVG	20	
P07724	Serum albumin	APQVSTPTLVEAARN	15	S
P07724	Serum albumin	APQVSTPTLVEAARNL	16	
P07724	Serum albumin	APQVSTPTLVEAARNLG	17	
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	GSITMQNIGAMSS	13	MB
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	GSITMQNIGAMSSY	14	
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	GSITMQNIGAMSSYL	15	
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	SITMQNIGAMSS	12	
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	SITMQNIGAMSSY	13	
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	ITMQNIGAMSSYL	13	
Q8CFE6	Sodium-coupled neutral amino acid transporter 2	SITMQNIGAMSSYL	14	
P32037	Solute carrier family 2, facilitated glucose transporter member 3	KDAGVQEPIYATIGAG	16	MB
O88307	Sortilin-related receptor	SGADAAVIQAARSTDV	16	MB
O88307	Sortilin-related receptor	SGADAAVIQAARSTDVA	17	
O88307	Sortilin-related receptor	SSGADAAVIQAARSTDVA	18	
Q9Z1W9	STE20/SPS1-related proline-alanine-rich protein kinase	ATSAPAPAPAPAPAA	15	N
A2AKJ5	Syntenin-1	EDLKVDKVIQAQTAYS	16	C
O55054	T4 surface glycoprotein	QSTAITAYKSEGES	14	MB
O55054	T4 surface glycoprotein	GFQSTAITAYKSEGE	15	
O55054	T4 surface glycoprotein	FQSTAITAYKSEGES	15	
Q9R2D0	TGF-beta-related neurotrophic receptor-3	PALRLQPRLPILSF	14	EM
Q6ZWX2	Thymosin, beta 4, X chromosome	IEQEKQAGES	10	C
Q0V930	Tln1 protein	LNFEEQILEAAKSIAA	16	C
F7AT44	Toll-interacting protein	TPTQQQQQLDAQAAQQ	18	C
F7AT44	Toll-interacting protein	TPTQQQQQLDAQAAQQQL	19	
Q3UDK1	TRAF-type zinc finger domain-containing protein 1	RTMNNVASCNRLNL	15	N
Q9CQH7	Transcription factor BTF3 homolog 4	FDEASKNEAN	10	N
Q9CQH7	Transcription factor BTF3 homolog 4	NFDEASKNEAN	11	
Q99KA2	Tuba1b protein	KDYEEVGVDSEVEGEG	15	C
Q99KA2	Tuba1b protein	EKDYEEVGVDSEVEGEG	16	
Q60805	Tyrosine-protein kinase Mer	EKWEETELDQLFSG	14	MB

Uniprot AN	Proteïna	Seqüència	Longitud	Localització
D3Z0S6	U1 small nuclear ribonucleoprotein A	MPGQMPPAQP	10	N
Q9R1R9	UBE-1a	GELAAREIQAVTGNS	15	NA
Q3UY15	Uncharacterized protein	SHALLSGQEHTECK	15	NA
D3YU45	Uncharacterized protein	YDALDVANKIGII	13	NA
A2AKJ2	Vimentin	FGSGTSSRPSS	12	C
A2AKJ2	Vimentin	FGSGTSSRPSSN	13	
Q8R539	WD-repeat protein p122	LLVYNLDLGGHLMV	15	NA
Q8R205	Zinc finger CCCH domain-containing protein 10	GGTGGGGSTGSAPP	14	MB
F6TRQ3	Zinc transporter 1	APDQEETNTLVANTSNS	17	MB
F6TRQ3	Zinc transporter 1	APDQEETNTLVANTSNSN	18	

Uniprot AN: Identificació de la proteïna en SwissProt; Proteïna: proteïna origen de la qual prové el pèptid; Longitud: mida del pèptid en nombre d'aminoàcids; Localització: MB (membrana plasmàtica); Lis (lisosomes); End (endosomes); S (secretada); EM (matriu extracel·lular); ER (Reticle endoplàsmic); G (Golgi); MIT (mitochondria) C (citosol); N (nuclear).

El repertori associat a I-A^d presentava una distribució de la mida dels pèptids que concordava amb el que s'ha descrit per a lligands de classe II en humans (Figura 4.1, barres morades), amb una mida entre 9 i 21 aminoàcids i una longitud mitjana de 15 aminoàcids, dins del rang de grandària definida MHC-II [12, 13, 17]. En canvi, el repertori per I-E^d va mostrar un major nombre de petits pèptids (figura 4.1, barres marrons).

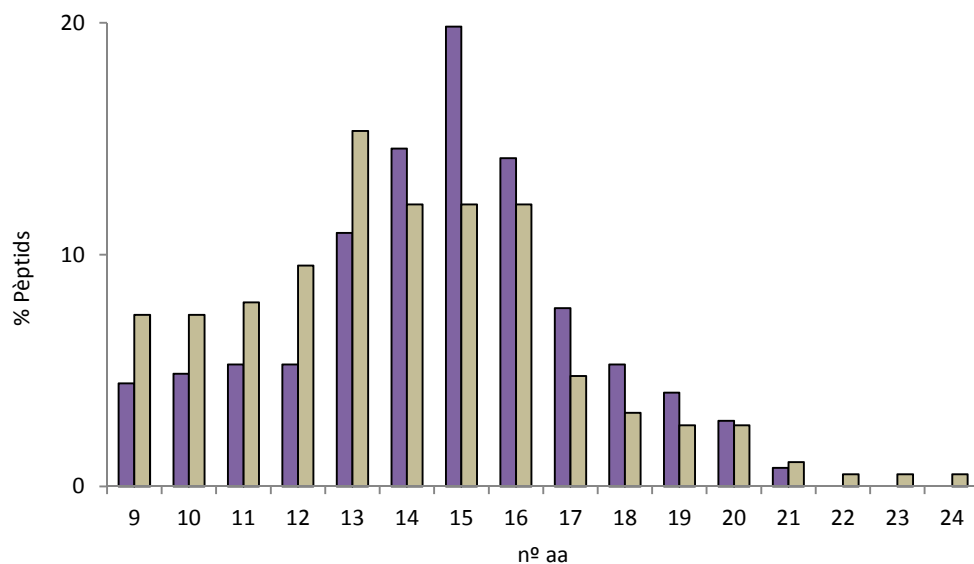


Figura 4.1: Distribució de mida de les lligands obtinguts per a I-A^d (morat) i I-E^d (marró)

Com es va fer pels al·lels de HLA-DR, els pèptids es van classificar en funció del compartiment cel·lular on es troba la proteïna parental i després van ser agrupats segons les seves vies de degradació citosòlica o endocítica [133]. Per

ambdues molècules, la freqüència de pèptids derivats de proteïnes degradades a través de la via endocítica va ser lleugerament superior respecte a la dels pèptids generats per la via de degradació citosòlica (53 i 56% pèptids a la via endocítica per I-A^d i I-E^d, respectivament, davant de 47 i 44 % pels pèptids de degradació citosòlica). Per tant, l'alt predomini de la via endocítica per a la generació de pèptids MHC-II observats en la majoria d'al·lels d'HLA-DR [115] [90, 109] no era tan evident entre els pèptids associats a les molècules de classe II d'H-2^d (figura 4.2.).

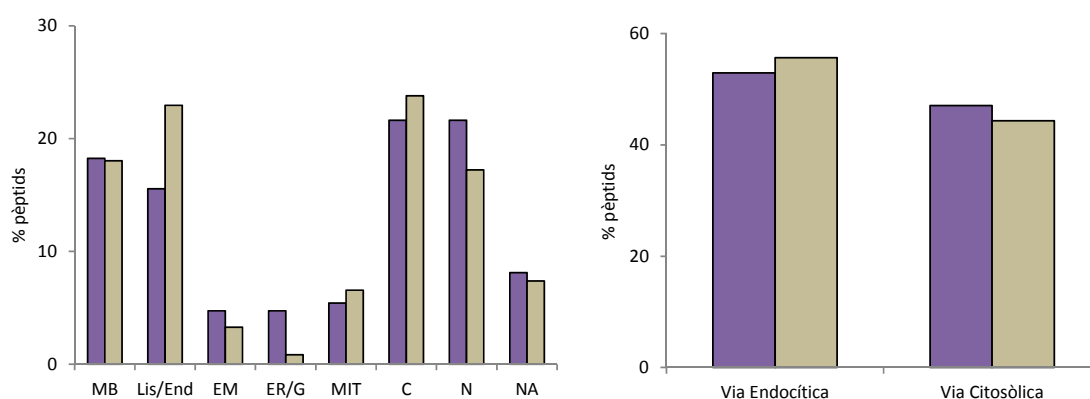


Figura 4.2: Esquerra: Localització cel·lular de les proteïnes parentals dels lligands d'I-A^d (morat) i I-E^d (marró); membrana (MB), matriu extracel·lular (EM), Reticle endoplasmàtic i Golgi (ER/G), endosoma/ lisosoma (Lis/End), citosol (C), mitocondria (MIT), nuclis (N) sense assignació (NA). Dreta: Percentatge mitjà dels lligands d'I-E^d i I-A^d classificats segons la seva ruta degradativa. Ruta endocítica: MB, Lis / End, EM, ER/G, MIT. Ruta citosòlica: C i N

4.3. Comparació de les seqüències de les cadenes β dels diferents al·lels

Per determinar els motius d'ancoratge de la molècula d'HLA-DR, van utilitzar el mateix mètode descrit en el capítol 1 per la determinació del motiu de HLA-DR3 a les cèl·lules CDG. Com ja vam comentar, ens vam basar en l'idea que dins la seqüència dels dominis polimòrfics d'MHC-II, els perfils de les butxaques d'ancoratge dels pèptids són independents de la resta del solc d'unió del pèptid [44]. Així, un cop determinat el perfil d'unió d'una d'aquestes butxaques per un al·lel, aquest pot ser compartit per altres al·lels sempre i quan els aminoàcids que formen part de la butxaca siguin idèntics [45].

Així el primer que es va fer va ser analitzar quins eren els aminoàcids de la molècula de MHC de classe II implicats en la formació de les butxaques d'unió al pèptid. Per aplicar aquest enfocament, es van alinear les seqüències del domini $\beta 1$ de les molècules I-A^d, I-E^d i I-E^k amb la d'HLA-DR1 i una seqüència consens HLA-DQ usant *Clustal W2*.

A la figura 4.3 es mostren en diferents colors els aminoàcids que formen les diferents butxaques d'interacció de les seqüències de les cadenes α i β de les diferents molècules estudiades amb els pèptids. La regió de la cadena β que forma la butxaca on es col·loca la posició 1 del core peptídica (P1), del residu 80 al 91, comparteix seqüència amb molts al·lels.

PDB	UniProt	Name	10	20	30	40	50
model	P01904	H2-Ea	IKEEHT- Y Q A	EFYLLPDKRG	EFMFD F DGDE	I FHVDIEKSE	TI W RLEEF A K
1FNG	P04224	H2-Ea	IKEEHT- Y Q A	EFYLLPDKRG	EFMFD F DGDE	I FHVDIEKSE	TI W RLEEF A K
4AH2	P01903	HLA-DRA	IKEEHV- Y Q A	EFYLNPDQSG	EFMFD F DGDE	I FHVDM A KKE	TV W RLEEFGR
2IAD	P04228	H2-Aa	IEADHVG F Y G T	TVYQSPGDIG	QY T HEFDGDE	L FYVDLDKKK	TV W R L PEFGQ
2NNA	Q30063	HLA-DQA1	IVADHVAS Y G V	NLYQSYGPSG	QYS H EPDGDE	E FYVDLERKE	TV W QLPLFR R

model	P01904	H2-Ea	60	70	80
1FNG	P04224	H2-Ea	F A S F E A Q G A L	AN I AVDKAN L	D V M K ER S N N T PD
4AH2	P01903	HLA-DRA	F A S F E A Q G A L	AN I AVDKAN L	D V M K ER S N N T PD
2IAD	P04228	H2-Aa	L I L F E P Q G G L	Q N I A E K I N L	G I L T K R S N F T PA
2NNA	Q30063	HLA-DQA1	F R R F D P Q F A L	T N I A V L K I N L	N V I K R S N S T AA

PDB	UniProt	Name	10	20	30	40	50
model	P01915	H2-Eb1-d	RDTRPR F L Y	VT S E C H F Y N G	TQHVR L L R F	IYNRE N L R F	DSDVGE Y R A V
1FNG	Q31163	H2-Eb1-k	RDSRPWF L Y	CK S E C H F Y N G	TQVR L L R F	FYNLE N L R F	DSDVGE Y R A V
4AH2	P04229	HLA-DRB1*01	GDTRPR F L W Q	L K E C H F Y N G	TERVR L L E R C	IYNQE S V R F	DSDVGE Y R A V
2IAD	P01921	H2-Ab1-d	GNSERHF V V Q	K G E C Y T N G	TQIR L V R Y	IYNRE E V R Y	DSDVGE Y R A V
2NNA	O19714	HLA-DQB1	RDSPED F V Y Q	K G M C Y F T N G	TERVR L V R Y	IYNRE E A R F	DSDVGE Y R A V

model	P01915	H2-Eb1-d	60	70	80	90
1FNG	Q31163	H2-Eb1-k	TELGR P DAEN	W N SQ P E L ED	ARAS V DT V CR	H N Y E I S DK F L VRR
4AH2	P04229	HLA-DRB1*01	TELGR P DAEN	W N SQ P E L EQ	KRAE V DT V CR	H N Y E I F DN F L VPR
2IAD	P01921	H2-Ab1-d	TELGR P DAEY	W N SQ K D L EQ	RAA V DT V CR	H N Y G V G ES F T VQR
2NNA	O19714	HLA-DQB1	TP L GP P AEY	W N SQ K E V LER	TRA L DT V CR	H N Y Q L E L R TT LQR

Figura 4.3. Seqüències de les cadenes α i β dels al·lels DRB1*0101, DQ, I-A^d, I-E^d, I-E^k. Amb diferents colors s'indiquen els aminoàcids implicats en la formació de les butxaques P1 (groc), P4 (verd), P6 (blau), P7 (taronja) i P9 (roig), implicades en la unió amb el pèptid. Es pot donar el cas que un mateix aminoàcid estigui implicat en dues butxaques (marcatge amb dos colors).

Vam trobar un 63% d'identitat entre $\beta 1$ de DR1 i I-A^d i un 10% entre DR1 i I-E^d. Amb la seqüència d'HLA-DQ, hi havia 70% d'identitat amb I-A^d. Si ens fixàvem en la figura 4.3, veiem que els aminoàcids que formen les diferents butxaques d'interacció no coincideixen al 100% amb DR1, i tampoc entre I-A^d i I-E^d.

La seqüència de la butxaca d'unió en la cadena β per P1 en el cas d'I-A^d implicava els aminoàcids en posició: 82, 86, 87 i 90. En I-E^d eren les posicions 82, 85, 86, 89, 90. En DR1, 82, 83, 85, 86, 89, 90 i 91.

Si es comparava I-A^d i DR1, s'hi trobaven 3 posicions en comú amb un canvi rellevant en la posició β 86: incorporació de prolina en lloc de glicina, fet que podia impedir la formació d'alguns enllaços d'hidrogen. Per tant, no podíem fer servir la comparació de seqüències per predir la butxaca P1 en I-A^d com a referència única. En lloc d'això, es va utilitzar addicionalment l'estructura del cristall definida per I-A^d [97].

En el cas d'I-E^d, cinc posicions eren comunes per la formació de P1 amb DR1 amb tres canvis conservatius de residus: β 85 (Val→Ile), β 86 (Gly→Ser) and β 90 (Thr→Leu). Per aquest motiu, es va assumir que els pèptids units a I-E^d presentarien aproximadament el mateix tipus de residus a P1 (Tyr, Phe, Trp, Leu, Ile, Ala, Val o Met) que HLA-DR1 i a tots els altres al·lels HLA-DRB1.

4.4. MOTIU D'ANCORATGE PER LA MOLECULA I-A^d

4.4.1 Selecció dels nonàmers

L'estructura publicada d'I-A^d per Scott et [97] mostrava que la butxaca d'unió al residu P1 era tan gran com ho era per a la majoria de les molècules DR, però que preferia residus d'una mida menor. D'altra banda, les butxaques d'unió dels residus P4, P6 i P9 presenten cavitats més petites que podien acomodar residus petits i no carregats. També van descriure que H-2 I-A^d era estructuralment similar a les molècules DQ, especialment perquè totes dues comparteixen a la cadena β una protuberància en la base del domini d'unió al pèptid (la base de la butxaca pel residu P4), el que suggereix que aquesta protuberància podia ser una característica conservada entre I-A^d i HLA-DQ i podria ser una posició rellevant per a la unió dels pèptids amb aquesta molècula.

La Prolina en $\beta 86$ d'I-A^d suggeria diferències en el contingut de P1, de manera que els residus "clàssics" de P1 no eren suficients per definir el *core*. D'altra banda, l'estructura cristal·lina indicava que P1 no era la posició més important per l'ancoratge del pèptid al solc d'unió. El mètode de predicció, per tant, es va basar en l'estructura cristal·logràfica publicada i en la comparació de les seqüències del domini $\beta 1$, mitjançant la fixació de la posició P1 i P4 per definir finalment el *core*, usant *LaiaMotifs*.

Es van utilitzar dos conjunts de dades per definir els motius MHC. La primera anàlisi es va realitzar amb totes les seqüències identificades per a cada un dels al·lels. Per al segon anàlisi, els conjunts de pèptids que formaven *nested sets* van ser substituïts per un únic pèptid que contenia els residus comuns a totes les seqüències. Aquest segon conjunt de dades es va utilitzar per a la confirmació de la primera anàlisi.

4.4.2. Generació de les matrius per la caracterització del motiu d'unió a I-A^d

Per fixar la posició P1, es van seleccionar els aminoàcids amb una relació pes molecular i volum petit-mitjà, no polars i no carregats, (Ala, Val, Leu, Ile, Met, Gli, Ser, Thr, Asn, Pro, Gln) i a més es fixà la posició P4 amb els quatre aminoàcids no polars més petits (Ala, Val, Leu, Ile). De 247 pèptids es van obtenir 325 nonàmers diferents (P1-P9) per construir la primera matriu de freqüència (matriu 1). A partir d'aquí, s'anaven generant matrius cada vegada més restringides en funció dels resultats que s'anaven obtenint (veure capítol 1).

La Matriu 1 (M1) va mostrar una dominància de la metionina en P1 i d'alanina en els altres tres posicions d'ancoratge (46, 21 i 19% per P4, P6 i P9, respectivament). Els residus aminoacídics que ocupaven P6 i P9 eren petits, com s'esperava. Dels 325 nonàmers, 185 eren únics (es van eliminar repeticions derivades del pèptids que formaven *nested sets*). D'aquests es va seleccionar el *core* que millor s'ajustava a la matriu 1 per a cada pèptid i/o *nested set*: Quan més d'un *core* era acceptable per a una única seqüència, es seleccionava el que tenia més residus amb comú amb la matriu 1. Els

nonàmers resultants (n = 85) van ser utilitzats per generar una matriu 2 (M2), fixant sols els residus per a P1. A continuació, la llista original de pèptids es va comparar amb la llista de cores

final. Els pèptids que es van quedar fora de la primera anàlisi a causa de les restriccions en P1 o P4, es van comparar de forma manual amb el motiu de la matriu 2. Al final, es va obtenir una matriu de freqüència final (M3) que defineix el millor motiu d'unió per al I-A^d. El nombre final de cores únics era de 126. Els resultats van suggerir que les butxaques P4 i P9 són més rellevants que P1 i P6, que semblen ser menys restrictiu (figura 4.4).

M1	1	2	3	4	5	6	7	8	9
A	12	12	11	46	13	21	16	14	19
C	0	0	0	0	0	1	0	1	0
D	0	1	5	0	3	3	3	3	7
E	0	9	10	0	9	4	4	4	3
F	0	3	0	0	1	1	2	2	1
G	5	5	6	0	5	3	5	6	7
H	0	0	0	0	2	1	2	1	3
I	5	12	4	24	6	12	9	9	2
K	0	1	2	0	1	3	1	3	2
L	4	6	8	15	5	5	3	5	6
M	29	5	8	0	6	3	10	2	4
N	11	5	8	0	8	9	7	4	8
P	6	4	5	0	5	2	4	6	5
Q	16	15	14	0	14	4	12	10	7
R	0	5	2	0	6	3	5	5	4
S	3	4	2	0	3	8	8	9	13
T	4	6	8	0	5	8	3	5	3
V	4	4	4	15	5	5	3	5	7
W	0	0	0	0	0	0	0	3	0
Y	0	4	3	0	2	2	3	5	0

M2	1	2	3	4	5	6	7	8	9
A	10	7	4	29	8	28	13	7	22
C	0	0	0	0	2	0	0	2	0
D	0	0	3	0	4	1	3	4	4
E	0	13	16	0	12	5	5	4	2
F	0	3	0	0	3	0	5	2	0
G	1	2	5	0	7	6	7	2	9
H	0	0	0	0	4	0	6	0	2
I	1	13	3	34	3	5	2	10	7
K	0	2	4	0	3	3	3	4	3
L	5	6	6	19	6	4	2	10	2
M	31	7	4	0	7	8	7	0	4
N	11	7	10	0	4	8	6	7	6
P	12	5	8	0	4	1	5	7	6
Q	19	8	14	0	12	6	10	8	8
R	0	6	3	0	6	2	9	7	7
S	3	6	3	0	4	5	5	5	13
T	2	6	7	0	6	9	4	4	3
V	4	4	7	18	3	5	6	4	2
W	0	0	0	0	0	0	0	9	0
Y	0	6	2	0	2	6	4	2	0

M3	1	2	3	4	5	6	7	8	9
A	9	8	3	19	8	21	10	6	20
C	0	0	0	0	1	0	0	1	0
D	0	1	4	4	5	4	4	4	4
E	0	14	10	2	11	7	6	6	2
F	0	1	0	0	2	0	4	2	0
G	4	3	8	3	5	4	7	4	9
H	0	0	0	0	5	0	7	0	2
I	3	10	6	26	1	6	3	9	4
K	0	3	4	1	3	5	2	5	2
L	7	4	5	15	3	3	2	7	3
M	28	10	8	3	8	9	5	3	3
N	9	6	8	0	3	7	4	7	8
P	10	7	8	1	6	3	8	10	6
Q	18	10	13	6	10	4	9	9	9
R	0	4	2	0	7	0	8	6	9
S	5	3	3	1	5	5	5	5	14
T	2	5	7	2	6	10	5	4	3
V	4	2	5	12	4	8	5	1	2
W	0	0	0	4	4	0	0	7	0
Y	0	7	4	2	3	5	6	4	0

Figura 4.4: Seqüències de les diferents matrius obtingudes en la generació del motiu d'ancoratge de I-A^d. Les matrius mostren una representació esquemàtica dels percentatges de cada aminoàcid en cada una de les posicions, de P1 a P9, normalitzats respecte la freqüència

de cada aminoàcid dins el conjunt del proteoma M1: Matriu seleccionant la posició 4 amb els 325 nonàmers possibles; M2: Matriu generada seleccionant els cores per a cada pèptid que s'ajustava a la matriu 1 (85 nonàmers); M3: Matriu final generada seleccionant un únic nonàmer per pèptid.

4.4.3. Comparació dels resultats amb les dades prèviament publicades

La figura 4.5 mostra el motiu d'ancoratge que es va obtenir per I-A^d en aquest treball. Es van considerar com a residus principals aquells que es veien incrementats significativament després de la correcció de Bonferroni. Els residus secundaris eren aquells que estaven augmentats significativament únicament abans de la correcció de Bonferroni.

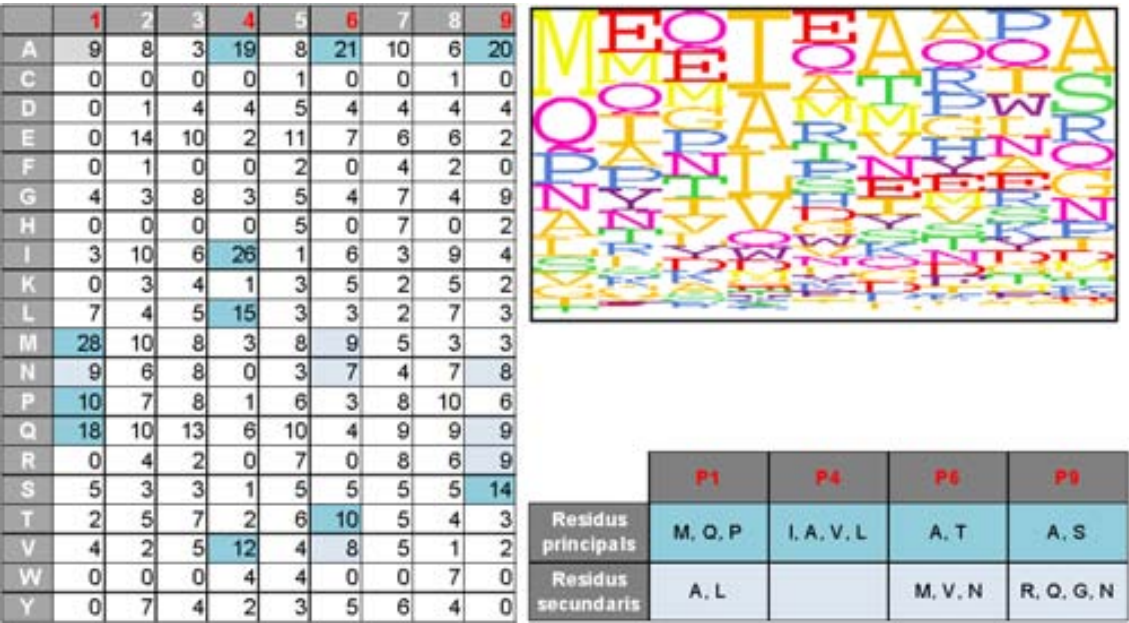


Figura 4.5: Motiu d'ancoratge d'I-A^d. Dreta: Matriu final amb la representació esquemàtica dels percentatges de cada aminoàcid en cada una de les posicions. Esquerra, dalt: Representació gràfica dels aminoàcids que es troben a cada posició, on la mida es correlaciona amb el grau de percentatge obtingut. Esquerra, baix, motiu d'ancoratge d'I-A^d on es mostren els residus principals (blau) i secundaris (gris).

A la base de dades SYFPEITHI (<http://www.syfpeithi.de/>) hi ha descrit un motiu d'ancoratge per a I-A^d (figura 4.6), a partir de quatre articles publicats entre 1987 i 1997 [97] [97] [134, 135].

	P1	P4	P6	P9
Residus acceptats	S, T, Y, E, V, W, M, L, I, H, F, A, C, D	V, L, I, A	A, V	

Figura 4.6: Motiu d'ancoratge d'I-A^d publicat a *Syfy*peithi

Si es comparàvem els resultats, la posició P1 era molt poc restrictiva respecte als resultats obtinguts per nosaltres i, a més, no assignaven cap residu per la posició P9. Les posicions P4 i P6 eren idèntiques a les obtingudes pel nostre sistema, tot i que per nosaltres P6 podria admetre a més treonina, metionina i valina.

4.5. MOTIU D'ANCORATGE PER LA MOLECULA I-E^d

4.5.1 Selecció dels nonàmers

La comparació de les seqüències corresponents a la butxaca d'unió a P1 (en groc a la figura 4.3) mostra tres residus diferents entre HLA-DR1 i I-E^d, corresponent a canvis entre aminoàcids amb la mateixa càrrega i mida similar: β85 (Val→Ile) β86 (Gly→Ser) i β90 (Thr→Leu). Per tant, vam considerar que els pèptids que s'unien en P1 en l'al·lel I-E^d havien de permetre els mateixos residus aromàtics o hidrofòbics (Tyr, Phe, Trp, Leu, Ile, Ala, Val o Met) que en DR1 i els altres al·lells d'HLA-DRB1.

4.5.2. Generació de les matrius per la caracterització del motiu d'unió a I-A^d

Per construir la matriu de freqüències (M1) es va fixar la posició P1 seleccionant tots els cores de pèptids amb un residu aromàtic o hidrofòbic a P1. De 187 pèptids, es van obtenir 362 diferents nonàmers P1-P9. La primera matriu (M1) mostrar predomini (27%) de la metionina en P1, però no hi trobarem un domini evident de qualsevol altre aminoàcid en qualsevol altra posició d'ancoratge.

D'altra banda, el domini β1 d'I-E^d contenia dos residus àcids a la butxaca d'unió de P9 (E a la posició 9 i D en la posició 57 de la seqüència de β1 regió (figura 4.3). Així, es va re-calcular la matriu de freqüència fixant els residus de P1 i

acceptant aminoàcids de caràcter no àcid en P9 (és a dir, els 20 aminoàcids excepte l'àcid aspàrtic i glutàmic), obtenint 336 nonàmers

La nova matriu (M2) va mostrar un patró molt similar: la metionina tornava a ser dominant a P1. Es van seguir els mateixos passos que per I-A^d i es van seleccionar els cores que s'ajustaven millor a la matriu 2. Els cores seleccionats (102) es van utilitzar per generar una matriu 3 (M3), de nou fixant la posició P1. A continuació, es van analitzar la llista original de pèptids contra la matriu 3, incloent l'anàlisi manual dels pèptids sense un residu P1 correcte, generant la matriu final (figura 4.7)

M1	1	2	3	4	5	6	7	8	9
A	14	8	10	8	8	14	6	9	8
C	0	0	1	0	1	1	1	1	0
D	0	3	2	2	3	2	3	7	4
E	0	11	7	8	7	7	5	3	4
F	12	3	3	5	3	4	2	3	5
G	0	6	6	5	5	4	6	7	4
H	0	2	0	4	3	3	3	3	1
I	13	4	4	6	6	5	5	7	6
K	0	2	2	2	4	5	4	5	5
L	11	6	5	4	7	5	6	5	6
M	24	9	14	12	12	6	10	4	9
N	0	7	9	5	7	4	9	5	7
P	0	6	6	5	8	6	8	10	7
Q	0	11	7	10	6	9	7	5	13
R	0	1	1	5	4	2	5	5	5
S	0	7	6	5	4	6	6	5	6
T	0	5	8	3	4	6	4	8	5
V	12	5	5	6	4	6	6	4	3
W	2	0	0	1	1	0	0	2	0
Y	11	4	3	5	3	5	3	3	3

M2	1	2	3	4	5	6	7	8	9
A	14	7	10	7	8	15	5	9	9
C	0	0	0	0	0	0	0	0	0
D	0	2	2	1	2	2	3	8	0
E	0	13	6	8	7	6	4	3	0
F	12	4	4	6	4	4	3	2	5
G	0	5	4	3	5	3	7	5	4
H	0	3	0	5	3	4	4	4	2
I	14	4	4	7	7	6	5	8	6
K	0	2	1	1	3	3	2	4	5
L	10	6	6	4	8	5	7	6	8
M	28	9	16	13	12	8	12	4	9
N	0	6	10	7	8	4	10	5	7
P	0	5	6	5	8	4	7	10	8
Q	0	13	7	11	8	8	7	6	14
R	0	1	1	4	2	2	4	3	4
S	0	6	5	5	4	7	6	5	6
T	0	6	9	3	4	6	4	7	5
V	13	5	5	6	4	6	7	4	4
W	1	0	0	1	0	0	0	2	0
Y	9	5	3	5	3	6	3	3	3

M3	1	2	3	4	5	6	7	8	9
A	10	6	8	8	6	8	5	6	11
C	0	0	0	0	0	0	0	0	0
D	0	3	3	3	3	2	4	6	0
E	0	12	9	12	6	4	4	4	0
F	10	8	3	5	4	4	0	3	7
G	0	5	3	4	4	2	8	6	1
H	0	0	0	2	5	3	5	5	2
I	15	1	7	3	1	6	4	10	4
K	0	2	2	2	3	2	1	3	6
L	11	4	3	7	7	2	6	4	9
M	26	3	14	17	14	14	6	0	14
N	0	10	7	5	9	7	12	10	5
P	0	7	7	3	5	8	8	12	2
Q	0	9	11	10	7	11	9	7	21
R	0	4	2	1	4	3	6	4	4
S	0	8	7	6	6	6	8	3	4
T	0	8	7	2	3	5	5	7	5
V	12	4	4	8	9	9	5	5	4
W	3	0	0	0	0	0	0	0	0
Y	13	5	3	3	3	4	5	5	2

M4	1	2	3	4	5	6	7	8	9
A	10	6	10	7	6	9	6	6	12
C	0	0	0	0	0	0	0	0	0
D	0	3	3	3	3	2	4	4	1
E	0	12	9	13	8	6	4	4	2
F	9	8	3	5	4	2	0	3	8
G	0	5	3	4	4	2	8	8	1
H	0	0	0	2	3	3	5	5	2
I	18	1	5	4	3	4	4	10	4
K	0	2	2	2	3	2	1	4	7
L	11	4	3	7	6	2	6	4	11
M	22	3	11	14	14	14	6	0	11
N	0	10	9	3	10	7	12	9	3
P	0	6	7	3	6	8	7	11	2
Q	0	10	12	10	7	12	8	7	22
R	0	4	2	1	2	3	6	4	2
S	0	9	6	6	6	6	7	3	4
T	0	8	7	3	3	5	5	7	5
V	13	4	4	7	8	9	6	5	3
W	3	0	0	0	0	0	0	0	0
Y	14	5	3	5	3	4	5	6	2

Figura 4.7: Seqüències de les diferents matrius obtingudes en la generació del motiu d'ancoratge de I-E^d. M1: Matriu amb els 362 nonàmers possibles; M2: Matriu seleccionant la posició 9 (336 nonàmers); M3: Matriu generada seleccionant els cores per a cada pèptid que s'ajustava a la matriu 2. (102 nonàmers); M4: Matriu final generada seleccionant un únic nonàmer per pèptid.

4.5.3. Comparació dels resultats amb les dades prèviament publicades

La figura 4.8 mostra el motiu d'ancoratge que es va obtenir per I-E^d en aquest treball.

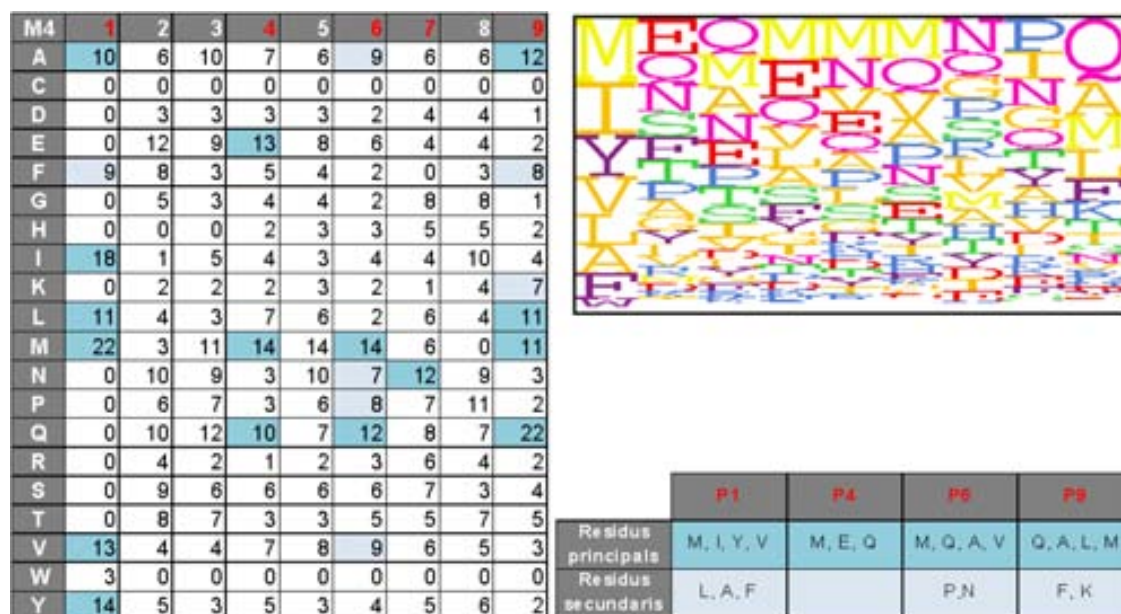


Figura 4.8: Motiu d'ancoratge d'I-E^d. Dreta: Matriu final amb la representació esquemàtica dels percentatges de cada aminoàcid en cada una de les posicions. Esquerra, dalt: Representació gràfica dels aminoàcids que es troben a cada posició, on la mida es correlaciona amb el grau de percentatge obtingut. Esquerra, baix, motiu d'ancoratge d'I-E^d on es mostren els residus principals (blau) i secundaris (gris).

Tal i com s'ha explicat per I-A^d, a la base de dades SYFPEITHI (<http://www.syfpeithi.de/>) hi ha descrit un motiu d'ancoratge per a I-E^d (figura 4.9), a partir de l'article publicat per [Schild et al.](#) en 1995 [99].

	P1	P4	P6	P9
Residus acceptats	W, Y, F, I, L, V	K, R, I, V	I, L, V, G	K, R

Figura 4.9: Motiu d'ancoratge d'I-E^d publicat a *Syfpeithi*

Si comparàvem ambdós motius, en aquest cas si s'observaven diferències importants principalment a la posició P4. Per aquest motiu, es va realitzar el modelatge de les molècules I-A^d i I-E^d i així comprovar la validesa dels nostres resultats.

4.6. MODELATGE DE LES MOLÈCULES I-A^d i I-E^d

Aquesta part va ser realitzada pel Dr X.Daura de l'Unitat de Bioinformàtica de l'IBB.

Es van seleccionar pèptids d'alta afinitat (HB) i de baixa afinitat (LB) del repertori obtingut per I-A^d i I-E^d (taula 4.3) i es va fer un modelatge utilitzant un protocol de simulació similar al presentat en treball anteriors al nostre laboratori[95]. L'establiment del protocol de simulació per ambdues molècules s'explica detalladament al capítol de materials i mètodes.

Els *scores* obtinguts de la unió de pèptids seleccionats a les molècules I-A^d i I-E^d a partir de l'anàlisi computacional basat en mètodes de simulació de dinàmica molecular, es donen a la Taula 4.3. Aquests *scores* proporcionen una estimació de l'afinitat relativa dels pèptids seleccionats per I-A^d o I-E^d (Taula 4.3), segons les seves energies d'interacció. L'anàlisi va ometre les energies de solvatació i efectes entròpics, ja que una constitució detallada d'aquests factors conduiria a temps computacionals inassequibles per aquest estudi. Malgrat aquestes aproximacions, per I-A^d els *scores* computacionals obtinguts van ser capaços de distingir clarament els pèptids predits com d'alta afinitat i dels pronosticats com de baixa afinitat, amb valors de més de dues vegades inferior (major afinitat) en el primer grup.

Molècula	Proteïna	Pèptid Id	Pèptid	core	Afinitat	score de simulació (kJ/mol)
I-Ad	IgM heavy chain variable region	PA1	YLQMNTLRAEDSAT	MNTLRAEDS	HB	-127
	Sodium-coupled neutral amino acid transporter 2	PA2	GSITMQNIGAMSS	MQNIGAMSS	HB	-143
	Fibrinogen alpha polypeptide	PA3	VIEKAQQIQALQSN	AQQIQALQS	HB	-133
	MHC class I heavy chain H2-K	PA4	TQDMELVETRPAGDG	MELVETRPA	HB	-129
	Neurabin-2	PA5	NKERMEKLEGYWGE	MEKLEGYWG	LB	-51
	Ig mu chain C region secreted form	PA6	EKYVTSAPMPEPGAPG	VTSAPMPEP	LB	-45
	TRAF-type zinc finger domain-containing protein 1	PA7	RTMNNVASNRLNL	NNVASNRL	LB	-78
I-Ed	Calcium-activated chloride channel regulator 2	PE1	VEKESMVGLVTFDSAA	VEKESMVGL	HB	-75 (-90)
	MHC class II H2-E alpha chain	PE2	ASFEAQGALANIAVD	ASFEAQGAL	HB	-50 (-58)
	Ig gamma-2A chain C region	PE3	SSGVHTFPAVLQSDLY	FPAVLQSDL	HB	-49 (-54)
	Cytochrome c oxidase subunit 6A1 mitochondrial	PE4	NPHVNPPLPTGYEDE	VNPPLPTGYE	LB	-118 (-102)
	Glutamine synthetase	PE5	LNETGDEPFQYKN	LNETGDEPF	LB	-107 (-103)
	Protein Snx21	PE6	DDDAEGLSSRLSGTSL	AEGLSSRLS	LB	-79 (-79)

Taula 4.3. Els residus que constitueixen el nucli d'unió i una afinitat d'unió qualitativa (HB / LB) s'han predit usant les matrius es mostren a les Figures 5 i 7. Les puntuacions de la simulació s'han obtingut per als nuclis predits utilitzant un enfocament de dinàmica molecular. Una puntuació més negativa és indicativa d'una major afinitat; els números entre parèntesis corresponen als càlculs amb un segon model d'I-E^d.

La concordança entre els càlculs físics obtinguts pel modelatge i les prediccions basades en l'estadística (dels repertoris peptídics) descrites anteriorment, proporcionava un suport addicional per al motiu descrit per I-A^d (Figura 4.5).

Els scores d'afinitat calculats per I-E^d van donar, en canvi, una correlació negativa amb les prediccions basades en la matriu. En aquest cas, el model I-E^d s'havia generat agafant l'estructura cristal·logràfica de la I-E^k en complex amb un pèptid d'HB com a plantilla. Per provar la dependència dels resultats en el model específic I-E^d, que s'utilitzà en els càlculs, es van fer servir unes noves estimacions d'afinitat dels pèptids utilitzant un segon model I-E^d del conjunt de 1.000 models generats, donant essencialment el mateix resultat (Taula 4.3).

La manca de coincidència entre la matriu que s'havia predit i aquests resultats basats en l'estructura, suggeria errors en el model estructural I-E^d generat i/o en el motiu predit a partir de la matriu desenvolupada. Per avaluar-ho més a fons, es va realitzar una anàlisi detallada de les interaccions a les butxaques d'unió en les estructures de models de complexos generats per cadascun dels sis pèptids a la Taula 4.3.

Per augmentar la diversitat dels residus en cada posició, es va escollir un nou pèptid de la literatura (PE7: LFRKDIAAK) [136] per ser modelat amb I-E^d

utilitzant el mateix protocol que per als anteriors. Els resultats d'aquesta anàlisi es donen a la Taula 4.4, on se subratllen les interaccions que es poden considerar específiques, és a dir, interaccions dels grups hidrofòbics a curta distància, les interaccions *pi* de grups aromàtics (incloent interaccions catió-pi) d'enllaç d'hidrogen dels residus polars i ponts de sal de residus ionitzables. Aquestes dades van permetre una primera descripció preliminar de les possibles preferències d'unió a les cinc butxaques des d'un punt de vista estructural.

Com es va predir anteriorment a partir de les consideracions observades a nivell de seqüència, la butxaca P1 semblava preferir els residus hidrofòbics grans, tot i ser menys profunda que en DR1. Les interaccions principals es realitzaven amb la cadena alfa, en particular, les fenilalanines 24, 32 i 54. Les leucines en els pèptids PE5 i PE7 interactuen també amb Trp α 43, i la fenilalanina de PE3 interactua, a més, amb Phe β 89.

Tot i que no hi havia tirosina en la posició 1 als pèptids escollits, la manera d'unió del residu Phe en PE3 indicava que la substitució per Tyr permetria un enllaç d'hidrogen amb Ser β 86 a la part inferior de la cavitat, pel que podria ser potencialment un aminoàcid amb una alta afinitat per ocupar aquesta posició (i per tant, es podria acceptar).

La butxaca P4 semblava preferir la unió de residus hidròfobs de mida mitjana, sent la leucina l'aminoàcid amb un major nombre d'interaccions específiques en el nostre conjunt, és a dir, amb α 24 Phe, Phe i Tyr β 26 β 78. Tot i que la cavitat es trobava envoltada per una sèrie de residus polars i carregats, als models no existia una distància i orientació apropiada per a la formació d'enllaços d'hidrogen o ponts salins amb els residus polars i carregats dels pèptids analitzats. Dels pèptids analitzats, només la cadena lateral de lisina podria establir un enllaç d'hidrogen amb la columna vertebral de Ser β 13. A més, la cavitat no semblava adequada per a residus àcids, a causa de la proximitat de tres àcids glutàmics (α 11, β 14 i β 28). Això es trobava en contradicció amb el motiu de la matriu, on s'havia predit la possibilitat de trobar àcid glutàmic (E) en aquesta posició. De fet, DR1 conserva els tres residus d'àcids esmentats anteriorment i se sabia que no tenien afinitat per l'àcid glutàmic a la P4.

L'enriquiment en la metionina està, en canvi, en línia amb les xarxes d'interacció observats en el model estructural.

La butxaca P6 semblava clarament preferir aminoàcids polars no carregats. Les principals interaccions (ponts d'hidrogen) s'estableixen amb $\alpha 11$ Glu, Asn $\alpha 62$ i $\alpha 66$ Asp. Residus no polars com ara metionina i isoleucina interactuen amb $\beta 11$ Val i Phe $\beta 30$ i podrien representar una alternativa. Per contra, els residus àcids no estarien afavorits degut a la proximitat de Glu $\alpha 11$ i $\alpha 66$ Asp. Qualitativament, aquest patró no està en desacord amb el motiu de la matriu descrit a la Figura 7.

En el model per I-E^d, P7 no forma una cavitat ben definida. De fet, s'ha suggerit que P7 no és rellevant per a la unió de pèptids en aquesta molècula.

L'anàlisi de la butxaca P9 suggereix una afinitat per aminoàcids petits polars, aminoàcids carregats (especialment positivament) i aminoàcids no polars de mida mitjana. La serina en aquesta posició establiria la coexistència dels enllaços d'hidrogen amb Asn $\alpha 69$ i Asp $\beta 57$ i, alternativament, podria també formar ponts d'hidrogen amb Arg $\alpha 76$. L'afinitat dels residus no polars era deguda principalment per $\alpha 73$ Met, Phe $\beta 30$ i $\beta 61$ Trp. Un àcid glutàmic en aquesta posició pot establir un enllaç d'hidrogen amb Asn $\beta 37$, però la proximitat de Glu i Asp $\beta 57$ $\beta 9$ fa que no sigui una bona opció. Finalment, la lisina combinava fortes interaccions polars amb $\beta 9$ Glu, Asp $\beta 57$ i $\beta 37$ Asn i una interacció catió- π amb Trp $\beta 61$. Una vegada més aquesta posició no estaria qualitativament en desacord amb el motiu de la matriu descrita inicialment per I-E^d, però no obstant això hi ha diferències en la classificació dels aminoàcids que es consideren principals i secundaris.

Pocket	I-Ed	Pèptids						
		PE1	PE2	PE3	PE4	PE5	PE6	PE7
P1		Val	Ala	Phe	Val	Leu	Ala	Leu
	Ile α 7			2 sc				
	Phe α 24	<u>10 sc</u>	<u>6 sc</u>	<u>6 sc</u>	<u>9 sc</u>	<u>6 sc</u>	<u>5 sc</u>	<u>4 sc</u>
	Ile α 31			4 sc		1 sc		4 sc
	Phe α 32	<u>6 sc</u>	<u>5 sc</u>	<u>6 sc</u>	<u>4 sc</u>	<u>6 sc</u>	<u>4 sc</u>	<u>6 sc</u>
	Trp α 43	3 sc	1 sc	<u>8 sc</u>	2 sc	<u>4 sc</u>	1 sc	<u>4 sc</u>
	Ala α 52	3 bs	3 bs	3 bs	1 sc	2 bs	3 bs	3 bs
	Ser α 53	6 bs	6 bs	6 bs	6 bs	6 bs	6 bs	6 bs
	Phe α 54	<u>9 bs</u>	<u>9 bs</u>	<u>9 bs</u>	<u>9 bs</u>	<u>7 bs</u>	<u>8 bs</u>	<u>6 bs</u>
	Glu α 55			1 sc	1 sc			
	His β 81		2 sc	2 sc	4 sc	4 sc	2 sc	2 sc
	Asn β 82	6 sc	2 sc	10 bs	6 sc	7 bs	3 sc	7 bs
	Ile β 85	4 sc	4 sc	4 sc	4 sc	4 sc	4 sc	4 sc
	Ser β 86			6 bs		4 bs		6 bs
	Phe β 89	1 sc	1 sc	<u>7 sc</u>	1 sc	2 sc	1 sc	3 sc
	Leu β 90			2 sc				
P4		Glu	Glu	Val	Leu	Thr	Leu	Lys
	Ile α 7	1 sc						5 bs
	Ile α 8							5 bs
	Gln α 9	10 bs	7 sc	7 sc	9 bs	7 sc	9 bs	11 bs
	Glu α 11		1 sc					
	Phe α 22	4 sc	4 sc	4 sc	4 sc	4 sc	4 sc	4 sc
	Phe α 24	7 sc	4 sc	4 sc	<u>6 sc</u>	4 sc	<u>6 sc</u>	6 sc
	Phe α 54	1 sc					1 sc	1 sc
	Asn α 62	3 sc	3 sc	3 sc	3 sc	3 sc	3 sc	3 sc
	Ser β 13	6 bs	3 sc	3 sc	6 bs	3 sc	6 bs	<u>5 bs</u>
	Glu β 14	2 bs						5 bs
	Cys β 15	1 sc			1 sc		1 sc	5 bs
	Phe β 26	6 sc	5 sc	4 sc	<u>6 sc</u>	4 sc	<u>6 sc</u>	5 sc
	Glu β 28	1 sc	4 sc	2 sc	1 sc	2 sc	1 sc	
	Tyr β 47		1 sc					
	Ser β 74		2 sc	1 sc	2 sc	3 sc	1 sc	
	Tyr β 78	11 bs	10 sc	<u>12 sc</u>	<u>7 sc</u>	14 bs	<u>10 sc</u>	12 sc
	Cys β 79							1 sc
	Asn β 82	3 sc			3 sc		3 sc	3 sc
P6		Met	Gln	Gln	Thr	Asp	Ser	Ile
	Gln α 9	4 sc	3 sc	3 sc	1 sc	1 sc	1 sc	3 sc
	Glu α 11	3 sc	<u>4 sc</u>	<u>4 sc</u>	3 sc	3 sc	3 sc	4 sc
	Asn α 62	7 bs	10 bs	11 bs	<u>9 bs</u>	10 bs	9 bs	7 bs
	Val α 65	1 sc	5 bs	5 bs	5 bs	7 bs	4 bs	1 sc
	Asp α 66	1 sc	6 bs	7 bs	<u>7 bs</u>	7 bs	<u>7 bs</u>	2 sc
	Asn α 69	2 sc	5 sc	5 sc	4 sc	4 sc	2 sc	3 sc
	Val β 11	<u>3 sc</u>	4 bs	4 bs	3 sc	3 sc	3 sc	<u>3 sc</u>
	Ser β 13	3 sc	1 sc	1 sc				2 sc
	Glu β 28	4 sc	1 sc	1 sc				4 sc
	Phe β 30	<u>6 sc</u>	8 sc	8 sc	6 sc	5 sc	2 sc	<u>6 sc</u>
	Trp β 61	1 sc	3 sc	3 sc	1 sc			1 sc

Pocket	I-Ed	Pèptids						
		PE1	PE2	PE3	PE4	PE5	PE6	PE7
P7		Val	Gly	Ser	Gly	Glu	Arg	Ala
	Gln α 9						1 sc	
	Val α 65	1 sc					1 sc	
	Asn α 69	1 sc		1 sc		1 sc	3 sc	
	Ser β 13						3 sc	
	Phe β 26						6 sc	
	Glu β 28					4 sc	<u>5 sc</u>	
	Phe β 30	2 sc		2 sc		7 sc	5 sc	4 sc
	Phe β 40						1 sc	
	Tyr β47					<u>7 sc</u>	7 sc	
	Trp β61	<u>13 sc</u>		16 sc		13 sc	13 sc	14 sc
	Ile β67	<u>4 sc</u>		4 sc		2 sc	7 bs	4 sc
	Asp β70						4 bs	
	Ala β71					1 sc	6 bs	
	Ser β 74						3 sc	
	Tyr β 78						1 sc	
P9		Leu	Leu	Leu	Glu	Phe	Ser	Lys
	Val α 65		1 sc	2 bs	2 bs	2 bs	2 bs	2 bs
	Asn α69	11 bs	11 bs	11 bs	11 bs	11 bs	<u>11 bs</u>	11 bs
	Leu α 70				1 bb			
	Val α72	3 sc	3 sc	3 sc	3 sc	3 sc	3 sc	3 sc
	Met α73	<u>3 sc</u>	<u>2 sc</u>	<u>3 sc</u>	4 sc	<u>7 bs</u>	2 sc	2 sc
	Arg α76	5 sc	6 sc	6 sc	6 sc	7 sc	<u>6 sc</u>	7 sc
	Glu β9	4 sc	4 sc	4 sc	4 sc	4 sc		<u>2 sc</u>
	Val β 11					1 sc		
	Phe β30	<u>5 sc</u>	<u>5 sc</u>	<u>5 sc</u>	4 sc	<u>11 bs</u>		<u>11 bs</u>
	Ile β 31					2 bs		
	Asn β37	3 sc	3 sc	3 sc	<u>5 sc</u>	6 sc	3 sc	<u>5 sc</u>
	Leu β38	1 sc	1 sc	1 sc				6 bs
	Leu β 53					1 bb		1 bb
	Pro β 56	1 bb						
	Asp β57	8 bs	6 bs	5 bs	4 sc	5 bs	<u>4 sc</u>	<u>5 bs</u>
	Asn β 60	2 sc						
	Trp β61	<u>11 sc</u>	<u>9 sc</u>	<u>8 sc</u>	6 sc	<u>9 sc</u>	6 sc	<u>11 sc</u>

Taula 4.4: Interaccions de curt abast (≤ 5 Å) entre les cadenes laterals del pèptid en les posicions P1, P4, P6, P7 i P9 i els residus I-E^d de cadascun dels 6 pèptids que es mostren a la Taula 4.3. El pèptid LFRKDIAAK (Berkower et al, 1986) també s'ha inclòs en l'anàlisi (PE7, del que s'havia obtingut una puntuació de simulació de -85 kJ / mol). Els elements de la taula indiquen el nombre d'àtoms del residu de proteïna que estan dins de 5 Å de la cadena lateral dels pèptids. La següent nomenclatura s'utilitza per indicar si el contacte és amb la columna vertebral i/o amb els àtoms de la cadena lateral del residu de la proteïna: bb (backbone), sc (cadena lateral), bs (columna vertebral i de la cadena lateral). Les cadenes laterals dels residus mostrats en negreta es troben a la superfície de la butxaca d'unió en l'estructura cristal·logràfica. Es subratllen les interaccions específiques (enllaços d'hidrogen, ponts de sal, les interaccions catió- π i interaccions hidrofòbiques estretes).

4.7. REEVALUACIÓ DEL MOTIU D'ANCORATGE PER LA MOLECULA I-E^d

Amb les noves dades proporcionades per l'estructura del modelatge d'I-E^d es va re-definir el motiu, establint les següents pautes:

- Fixació de P1 amb aminoàcids clàssics: I, V, L, N, A, Y, W
- Fixació de P9 més restrictiva, amb aminoàcids petits polars (G, S, T), bàsics (K, R) i no polars (A, V, L, I, M, F).

Es van seguir els mateixos passos explicats a l'apartat 4.6 obtenint diferents matrius i finalment una matriu final que definia un nou motiu més acurat per I-E^d tal i com es mostra a la figura 4.10.

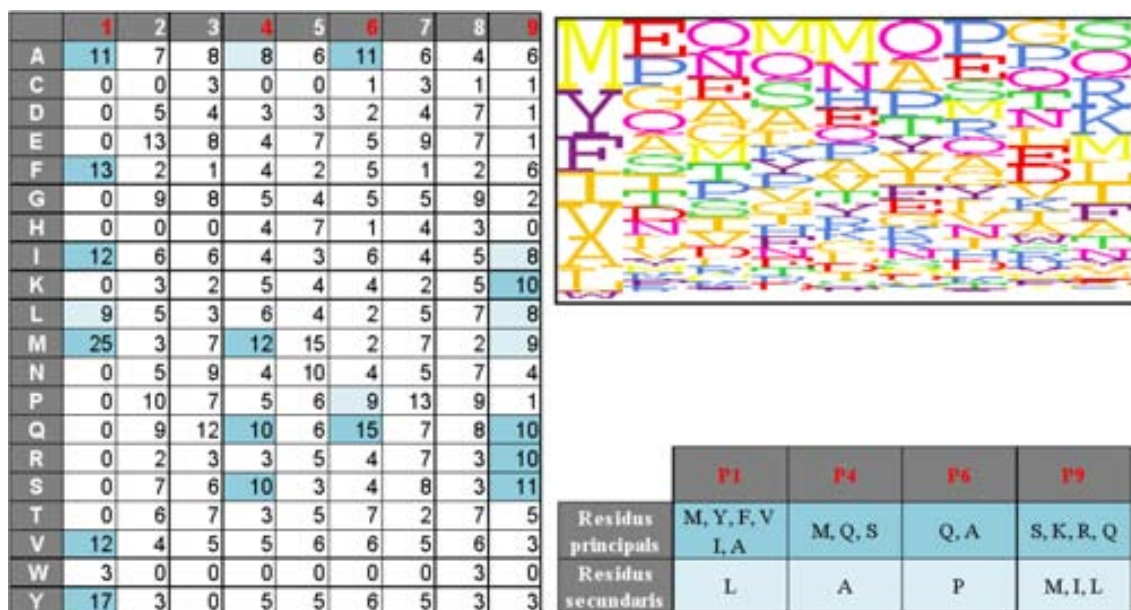


Figura 4.10: Motiu final d'ancoratge d'I-E^d a partir dels resultats obtinguts pel modelatge de la molècula. Dreta: Matriu final amb la representació esquemàtica dels percentatges de cada aminoàcid en cada una de les posicions. Esquerra, dalt: Representació gràfica dels aminoàcids que es troben a cada posició, on la mida es correlaciona amb el grau de percentatge obtingut. Esquerra, baix, motiu d'ancoratge d'I-E^d on es mostren els residus principals (blau) i secundaris (gris).

Comparant aquests resultats amb els obtinguts pel modelatge de l'estructura d'I-E^d, trobàvem més coincidències que amb la matriu anterior descrita (figura 4.7).

- Posició 1 (P1): Hi ha un canvi en el percentatge dels residus preferits, i per tant, considerem ara com a residus principals: Met, Tyr, Phe, Val, Ile, Ala i com a secundari Leu.

- Posició 4 (P4): Es conserven els residus Met i Gln, que no es contradient amb el modelatge, però s'elimina l'àcid glutàmic que semblava no estar afavorit en aquesta posició. Segons el resultat del modelatge, aquesta posició admetria residus hidròfobs i per tant, l'aparició d'Ala com a residu principal estaria acceptada.

- Posició 6 (P6): Aquesta posició es veuria restringida i acceptaria Gly, Ala, Pro, aminoàcids ja presents al primer motiu descrit i que no estaven en desacord amb l'estructura segons els resultats del modelatge.

- Posició 9 (P9): S'obtenen com a residus principals: Ser amb una freqüència del 11% i Lys, Arg, Gly amb un 10% cadascun. Aquesta butxaca és la que presenta els principals canvis, junt a la butxaca P4, amb la incorporació d'Arg com a residu principal i l'augment de Lys, ja existent al model anterior però ara com a residu dominant acceptat. Aquesta dominància bàsica per P9 coincideix amb el descrit a la literatura (Schild et al. 1995) i amb l'estructura que havíem predit segons el modelatge.

4.8. PREDICCIÓ D'UNIÓ DE PÈPTIDS NATURALS I COMPARACIÓ AMB EPÍTOPS DEFINITS DE CÈL·LULES T

Per a validar els motius, es van analitzar tots els pèptids de cada repertori, calculant la predicció d'unió de cada pèptid a I-A^d o I-E^d. El mètode de predicció d'afinitat teòrica dels pèptids per a aquestes molècules es descriu al capítol de materials i mètodes.

El patró de freqüències per cada una de les categories (alta afinitat: HB; afinitat intermèdia: IB; i baixa afinitat: LB) entre els pèptids d'I-A^d va mostrar un clar domini d'afinitat teòrica alta (HB), seguit d'IB. En canvi, els pèptids d'I-E^d mostraven un patró diferent, ja que la majoria eren IB, seguit de pèptids HB. En ambdós casos, els LB constituïen el percentatge menor (figura 4.11).

Quan s'analitzaven els pèptids en funció de la via de degradació, els pèptids d'I-A^d de la via endocítica mostraven el mateix patró, mentre que els pèptids citosòlics mostraven igual freqüència per HB, IB i LB. En contrast, tots els pèptids d'I-E^d van mostrar el mateix patró (IB > HB > LB) independentment de la via de degradació de la qual provenien, (Figura 4.11).

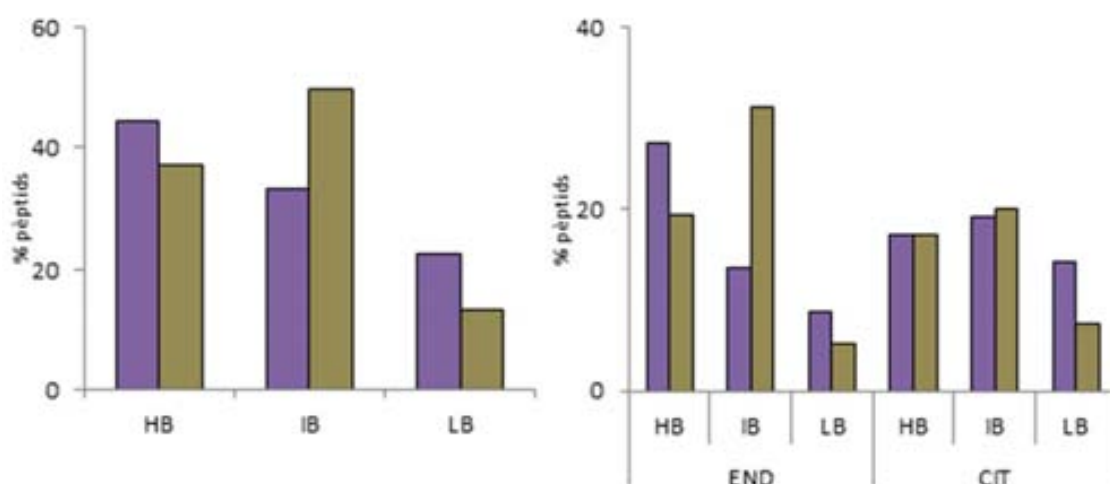


Figura 4.11. Esquerra: Comparació de la predicció d'afinitat d'unió teòrica dels pèptids associats a I-A^d (morat) i I-E^d (marró). Dreta: Predicció d'afinitat d'unió dels pèptids en funció de la via de degradació proteica.

La següent validació dels motius es va fer analitzant de la mateixa forma un nombre de epítops ja definits que presenten I-A^d o I-E^d a cèl·lules T i que figuren a la base de dades IEDB (<http://www.iedb.org/>). El nombre de epítops T analitzats van ser 116 per I-A^d i 89 per I-E^d (figura 4.12). La predicció d'unió d'aquets epítops a les molècules I-A^d o I-E^d va mostrar patrons molt similars a la dels pèptids naturals, amb major freqüència de HB per a I-A^d i de IB per a I-E^d, mantenint la freqüència de LB al voltant de 20% en I-A^d i menys de 10% en I-E^d. (Figura 4.12).

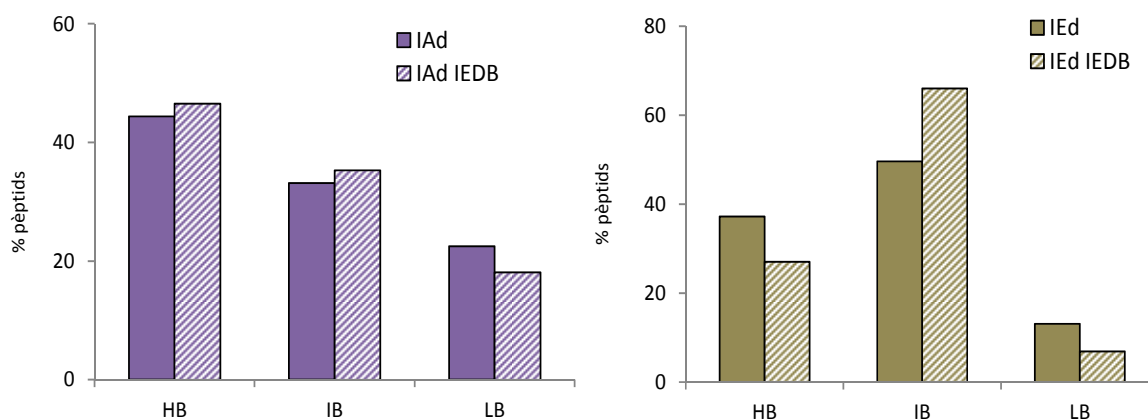


Figura 4.12. Comparació de la predicció d'afinitat d'unió a I-A^d (esquerra) i I-E^d (dreta). Les barres en colors sòlids representen els pèptids naturals i les barres ratllades representen els epítops T associats a I-A^d (morat) i I-E^d (marró) que figuren a la base de dades IEDB.

4.9. OBSERVACIONS FINALS

Aquesta és la primera anàlisi de pèptids associats de forma natural als al·lels de classe II H-2^d I-A^d i I-E^d a partir de mostres de teixits de ratolí sense tractament *ex-vivo* [137]. Hem seqüenciat 247 i 187 pèptids naturals d'I-A^d i I-E^d respectivament per espectrometria de masses d'alta resolució. Estudis previs que caracteritzen els al·lels de H-2^d de classe II s'havien dut a terme principalment utilitzant pèptids no naturals (pèptids sintètics) que no imitaven l'entorn *in vivo* o amb un nombre molt limitat de pèptids naturals. Aquesta aproximació té la gran avantatge que els pèptids utilitzats han estat seleccionats per les cèl·lules presentadors d'antigen durant el seu processament, on un gran nombre de fenòmens influencien la selecció dels pèptids que seran presentats [108].

En molècules humanes de classe II, la majoria de les proteïnes que generen els lligands són derivats dels compartiments de membrana, incloent la membrana plasmàtica, aparell de Golgi, mitoncòndries, endosomes, reticle endoplasmàtic i vesícules de secreció. La proporció de pèptids citosòlics o nuclears sempre sol ser relativament baixa, a excepció del que s'ha descrit per HLA-DR3 al capítol 1, 2 i 3. Per a les molècules de ratolí de classe H-2^d, al

voltant de la meitat dels pèptids provenen de la degradació de proteïnes d'origen citosòlic o nuclear.

Les molècules d'I-A són estructuralment similars a l'humà HLA-DQ. L'estructura del model d'HLA-DQ, confirmada per l'estructura cristal·logràfica [138] presenta diferències amb HLA-DR a nivell de seqüència de les butxaques d'unió dels pèptids (figura 4.3), el que probablement li donaria a DQ una importància funcional. De fet, el motiu I-A^d que hem descrit estaria d'acord amb les característiques d'unió de pèptids en DQ publicades com ara (i) la posició P1 no es tan important com ho és per als al·lels de HLA-DR [139] i (ii) la rellevància de les butxaques 4 i 9 en la unió al pèptid. El motiu d'I-A^d definit ha estat a més, ben confirmat amb la modelització i per l'anàlisi d'epítops de cèl·lules T definits per I-A^d (IDBD).

En el cas d'I-E^d, el motiu d'unió a pèptid va ser menys ben definit inicialment, però gràcies al modelatge de la molècula a partir del cristall descrit per I-E^k, es va poder definir un motiu més acurat. Fins ara, les dades havien suggerit que I-E^d era molt similar a altres al·lels d'IE, com ara I-E^k i altres [99] amb sols residus bàsics dominants en P9 [140]. Estudis previs sobre la unió dels epítops de cèl·lules T definits per a I-Ed, havien mostrat també aquesta tendència en P9 i també en altres posicions: el motiu "tres elements bàsics" definit per Sette et al, basant-se en que la seqüència I-E^d conté dos residus àcids a la butxaca d' ancoratges en posició P9. Els nostres resultats mostraven que tot i què a la posició P9 trobàvem principalment residus bàsics, no eren els únics admesos possibles: la serina i la glutamina tindrien estadísticament possibilitats de ser admesos i estructuralment, serien vàlids per aquesta posició.

CAPÍTOL 5: ESTUDI COMPARATIU ENTRE ELS REPERTORIS PEPTÍDICS DE TIMUS I MELSA EN RATOLINS BALB/C

5.1. ABSÈNCIA DE DIFERÈNCIES ALS REPERTORIS PEPTÍDICS DE RATOLINS BALB/C I RATOLINS TRANSGÈNICS PER AL TSHR

Es va partir de timus i melses de ratolins Balb/c estàndards i de ratolins transgènics per a l'eTSHR humà. L'única diferència entre els ratolins transgènics i els estàndards era la introducció d'un transgen humà (eTSHR) d'expressió a tiroides. Els objectius d'aquest estudi eren en primer lloc comprovar si l'expressió de l'eTSHR afectava els repertoris presentats al timus i la melsa i veure si eren capaços de detectar pèptids derivats d'aquest receptor als animals transgènics..

No es va detectar pèptids derivats de l'eTSHR en els repertoris peptídics analitzats. Seguidament, es van buscar les possibles diferències entre els pèptids identificats a cada tipus de animal.

La distribució de mides dels lligands I-E^d i I-A^d seqüenciats no va mostrar diferències importants entre ratolins estàndards i transgènics. Més del 90% dels pèptids seqüenciats en ambdós animals tenien una mida d'entre 10 i 22 aminoàcids (figura 5.1). Tot i observar petites variacions en els percentatges detallats, ambdues situacions segueixen el patró canònic de la mida de pèptids associats a MHC de classe II descrit.

Es va comparar l'origen i la via de processament del repertori peptídic en timus i melses dels estàndards i transgènics (figura 5.2). Pel que respecta a la localització de les proteïnes parentals dels repertoris de pèptids associats a I-E^d al timus, tant en wt com en tg, la via citosòlica era la que presentava un percentatge més elevat, 52% i 59% respectivament (figura 5.2 A).

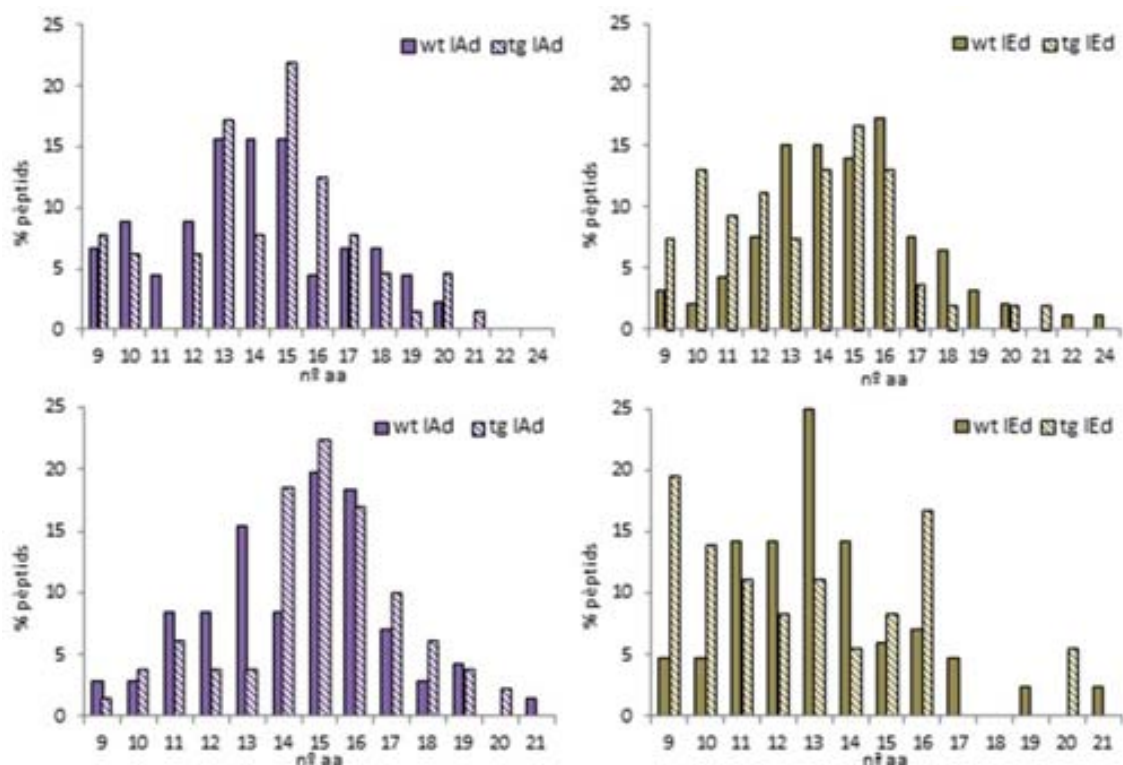


Figura 5.1: Distribució de mida de les seqüències obtingudes en les mostres de timus (A) i melsa (B) de ratolins balb/c estàndars (wt) i transgènics (tg) tant per la molècula I-A^d (esquerra) com per a I-E^d (dreta)

Pel que fa als timus I-A^d, en els ratolins estàndars hi havia un equilibri entre els pèptids provinents de la via endocítica i els de la citosòlica (figura 5.2 A). En els transgènics, aquest equilibri quedava lleugerament desplaçat i un 53% provenien de la via endocítica (figura 5.2 B). El patró final era equivalent.

Per la melsa I-E^d, la via endocítica era la majoritària en ambdós tipus d'animals i en el cas de la melsa I-A^d es donava una situació semblant on la via endocítica era també la majoritària (figura 5.2C i 5.2D)

Per tant no es van observar diferències importants entre els repertoris d'animals estàndars i transgènics i per tal de donar més validesa estadística als resultats, es va decidir que, a l'igual que per a la definició del motiu d'unió (capítol 4), no diferenciaríem entre ratolins transgènic i estàndars per fer la comparació entre timus i melsa (a partir d'ara els anomenarem Balb/c a tots dos animals)..

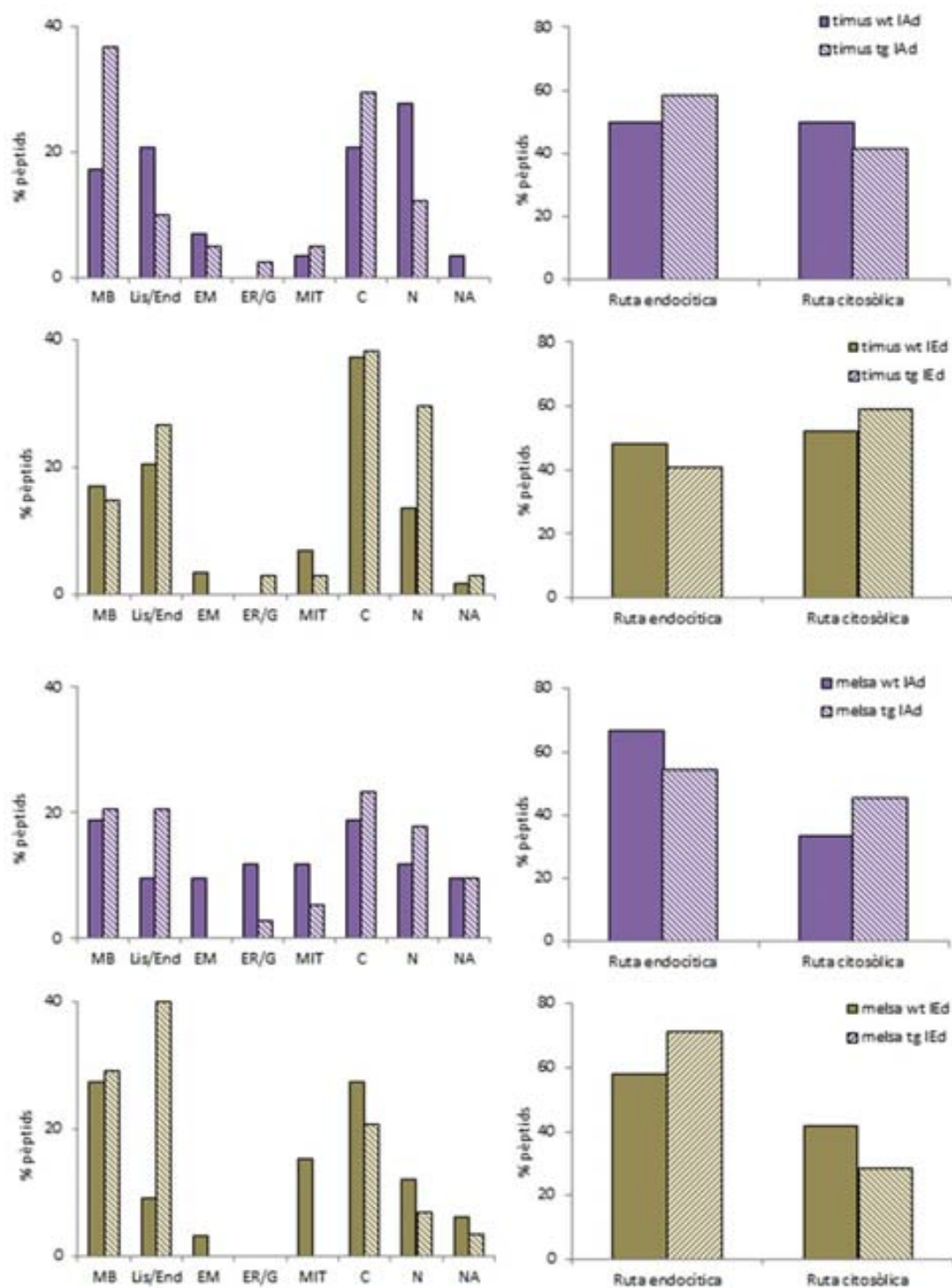


Figura 5.2: Gràfiques comparatives del repertori peptídic de les mostres de timus i de melses entre ratolins estàndards (wt) i transgènics (tg). Es representa les rutes de processament dels pèptids (esquerra) i la localització de les proteïnes parentals dels pèptids (dreta) per la molècula I-E^d (marró) i la molècula I-A^d (lila). Membrana (MB), matriu extracel·lular (EM), Reticle endoplasmàtic i Golgi (ER/G), endosoma / lisosoma (Lis/End), citosol (C), mitocondria (MIT), nuclis (N) sense assignació (NA). Via endocítica: MB, Lis / End, EM, ER/G, MIT. Via citosòlica: C i N

5.2. ESTUDI DEL REPERTORI PEPTÍDIC PRESENTAT EN TIMUS DE BALB/C

5.2.1. Repertoris peptídics obtinguts per timus Balb/c

Per la molècula d'I-A^d es van obtenir 97 lligands corresponents a 63 proteïnes diferents. En el cas d'I-E^d, 135 lligands diferents identificats corresponien a 81 proteïnes. Un resum comparatiu del nombre de seqüències obtingudes dels timus es mostra a la taula 5.1.

	Espectres identificats	Seqüències no redundants	Seqüències Finals	Proteïnes (*)	% nesteds
timus IAd	126	105	97	63	12
timus IEd	172	152	135	81	14

(*) Proteïnes: s'indica el nombre de grups de proteïnes calculat per Sequest (ver material i mètodes)

Taula 5.1: Resum comparatiu del nombre de seqüències obtingudes per a timus Balb/c

La mida dels pèptids seqüenciats seguia una distribució normal en totes les mostres (Figura 5.3). Per a I-E^d, s'obtenia una mitjana de 14 aminoàcids (d'entre 10 a 24 aminoàcids) i una massa mitjana de 744 Da. En el cas de I-A^d, la mitjana fou de 15 aminoàcids (d'entre 10 i 21 aminoàcids) i la massa mitjana de 719 Da. Per tant, els pèptids seqüenciats en timus i per ambdues molècules estaven dins del rang estàndard de la mida dels lligands d'HLA-DR i en comparació amb humans, els pèptids presentaven una massa mitjana més petita.

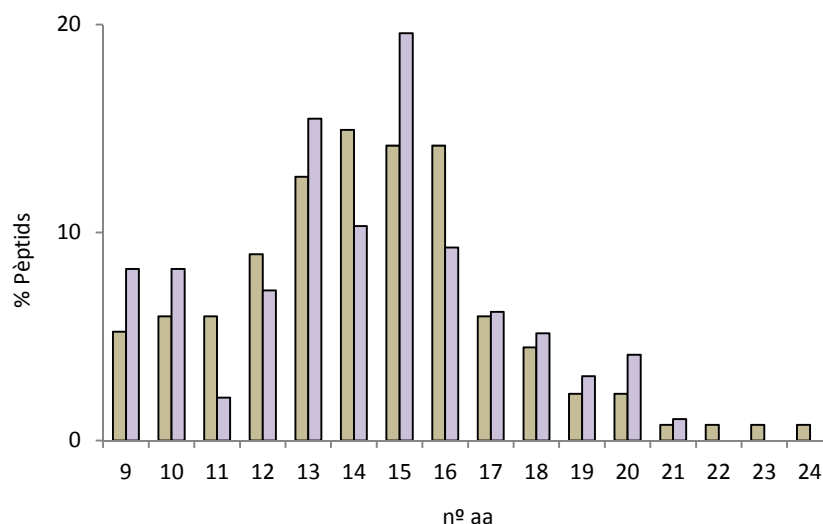


Figura 5.3: Distribució de mida de les lligands obtinguts en a les mostres tímiques per I-E^d (marró) i I-A^d (morat)

5.2.2. Origen i via de processament del repertori peptídic en timus Balb/c

Com ens capítols anteriors, els pèptids es van classificar en funció de la localització de la proteïna parental i de la seva via de degradació (figura 5.4). Per a I-E^d, el 51% dels pèptids seqüenciats procedien de la via degradació citosòlica, mentre que el 49% restant derivaven de proteïnes degradades endocítiques. Per a I-A^d, s'invertia la dominància: un 52% de pèptids pertanyien a la via endocítica i un 48% a la via citosòlica, encara que les diferències eren poc importants. Les proteïnes de citosol (33%) foren les majoritàries en I-E^d i les de membrana (27%) junt amb les de citosol (25%) ho van ser a I-A^d.

Segons el que ha estat publicat recentment al nostre laboratori, en timus d'humans (Collado et al, 2013) la via endocítica fou la majoritària per tots els al·lels estudiats.

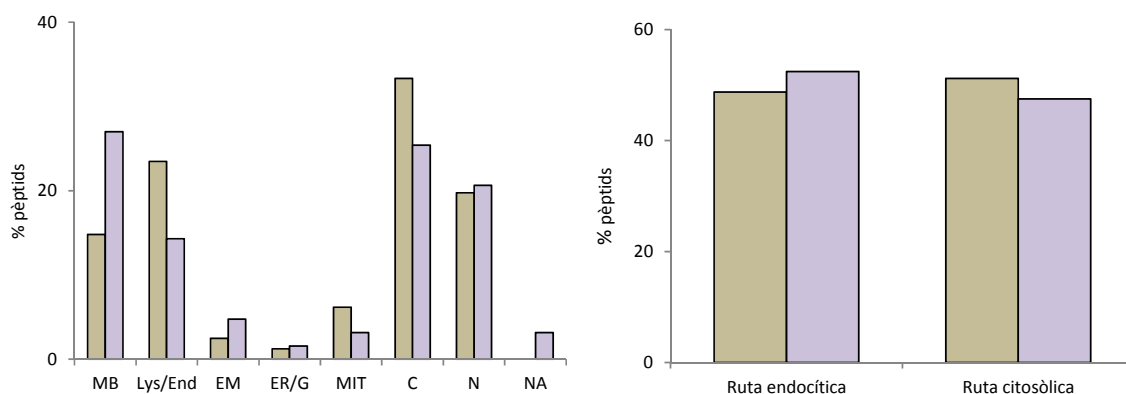


Figura 5.4: Esquerra: Localització cel·lular de les proteïnes parentals dels lligands d'I-E^d (marró) i I-A^d (morat); Dreta: Percentatge dels lligands d'I-E^d i I-A^d classificats segons la seva via degradativa.

Dins del repertori peptídic per a I-E^d, trobàvem pèptids procedents de proteïnes residents de la via endocítica i relacionades amb el sistema immunitari com ara Catepsina B, CD74, *H-2 class II histocompatibility antigen, A-F alpha chain*, *Ig gamma-1 chain C*, *Ig gamma-2A chain C*, *MHC class I heavy chain H2-K*, *MHC class II H2-E alpha chain*. Entre les proteïnes citosòliques, que suposaven el percentatge més alt, hi trobàvem la timosina beta 4 i proteïnes d'ensamblatge dels ribosomes. També es van trobar un alt nombre de ribonucleoproteïnes i histones diferents.

Al repertori d'I-A^d, entre les proteïnes de la via endocítica, també es van trobar pèptids relacionats amb proteïnes de la immunitat com *H-2 class II histocompatibility antigen A-F*, *Ig gamma-2A chain*, *Ig heavy chain V*, *Ig kappa chain C*, *MHC class I heavy chain H2-K*, *MHC class II H2-E alpha chain*. Pel que feia a la via citosòlica, també trobàvem la timosina beta 4 i diferents cinases com l'Akap 12, amb un *nested set* format per 6 pèptids diferents.

5.2.3. Patró de processament dels pèptids aïllats de timus Balb/c

Es va analitzar el patró de processament dels pèptids segons si es generaven a partir dels extrems terminals o de la part interna de la proteïna parental (figura 5.5). Per I-E^d vam obtenir que un 32% dels pèptids presentaven patrons terminals: N-terminal (4%) i C-terminal (33%) mentre que un 63% eren seqüències internes.

Per I-A^d, un 8% del repertori provenien de la regió N-terminals, un 11% de C-terminal i un 79% provenien de la part interna. Per la via citosòlica com per l'endocítica, tant en I-E^d com en I-A^d, la majoritària dels pèptids eren seqüències internes. Tanmateix, en el repertori d'I-Ed trobàvem més seqüències terminals que en el d'I-A^d.

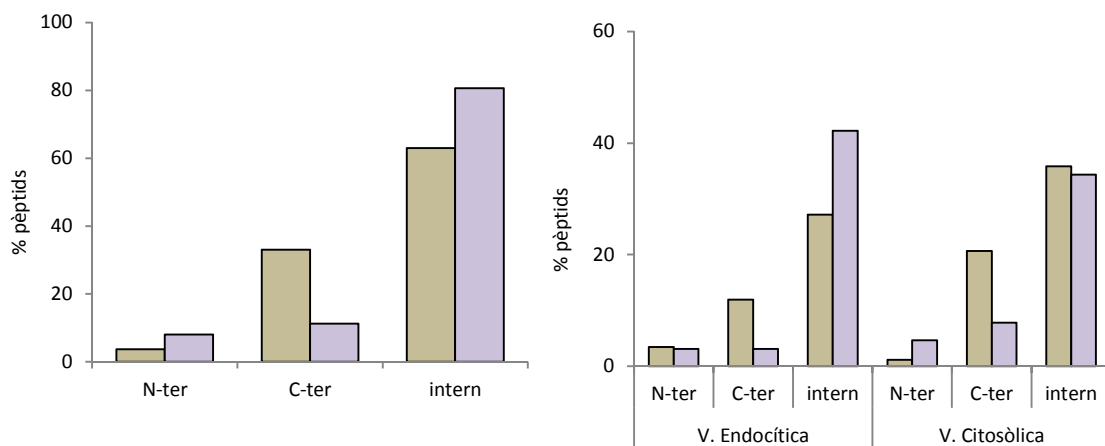


Figura 5.5: Patró de processament dels pèptids aïllats de timus en I-E^d (marró) i I-A^d (morat). Gràfic en funció de la localització dels pèptids respecte la proteïna parental, depenent si eren N-terminal (N-ter), C-terminal (C-ter) o de la part interna. Al gràfic de l'esquerra, les localitzacions s'agrupaven en funció de la seva via de degradació, via endocítica i via citosòlica

5.3. ESTUDI DEL REPERTORI PEPTÍDIC PRESENTAT EN MELSA DE BALB/C

5.3.1. Repertoris peptídics

Per I-A^d, es van obtenir 175 lligands corresponents a 107 proteïnes diferents. Per I-E^d es van obtenir 71 lligands diferents que corresponien a 57 proteïnes. Un resum comparatiu del nombre de seqüències obtingudes per als timus es troben a la taula 5.2.

	Espectres identificats	Seqüències no redundants	Seqüències Finals	Proteïnes (*)	% nesteds
melsa IAd	245	181	175	107	18
melsa IEd	138	89	71	57	11

(*) Proteïnes: s'indica el nombre de grups de proteïnes calculat per Sequest (ver material i mètodes)

Taula 5.2: Resum comparatiu del nombre de seqüències obtingudes per a la melsa

5.3.2. Distribució de la mida

La mida dels pèptids seqüenciats presentava una distribució normal en totes les mostres (Figura 5.6). En el cas de I-E^d, amb una mida mitjana de 15 aminoàcids (d'entre 10 a 21 aa) i una massa mitjana de 774 Da. En el cas de I-A^d, la mida mitjana també fou de 13 aminoàcids i la massa mitjana fou de 667 Da.

Per a I-E^d, s'observà un percentatge alt de pèptids petits que no s'observà en I-A^d, que sseguia millor la tendència descrita per a pèptids de MHC. Per tant, els pèptids seqüenciats en melsa i per ambdues molècules es trobaven dins del rang de mida estàndard dels lligands de MHC-II encara que en comparació amb humans, els pèptids presentaven una massa mitjana més petita com ja ocorria amb el cas del timus.

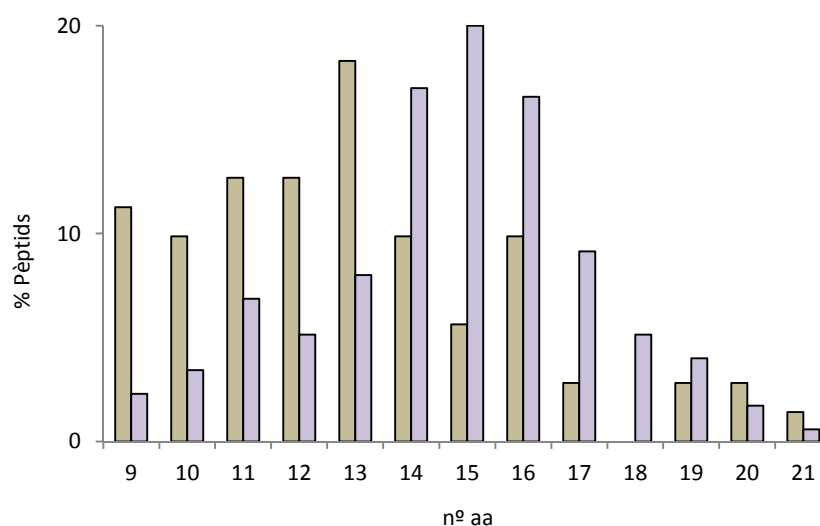


Figura 5.6: Distribució de mida dels lligands obtinguts a les mostres de melsa per a I-E^d (marró) i I-Ad (morat)

5.3.3. Origen i via de degradació del repertori peptídic en melsa Balb/c

Els pèptids es van classificar en funció de la localització de la proteïna parental i de seva via de degradació (figura 5.7). En el cas de I-E^d, el 64% dels pèptids seqüenciats procedien de la via degradació endocítica, mentre que el 36% restant derivaven de proteïnes degradades al citosol. Per a I-A^d, les dades eren semblants, i la via endocítica també fou la predominant amb un 54 % del total.

Les proteïnes de membrana (25%) i de compartiments vesiculars (25%) foren les majoritàries en I-E^d tot i que també un alt percentatge (23%) provenien del citosol. En el cas de I-A^d, les proteïnes citosòliques (26%) foren les més abundants.

Al nostre laboratori es va descriure (Tesis Dr. JA Collado, UAB 2013), que en humans i per la melsa, la via endocítica seria la majoritària a l'igual que ocorre amb I-E^d.

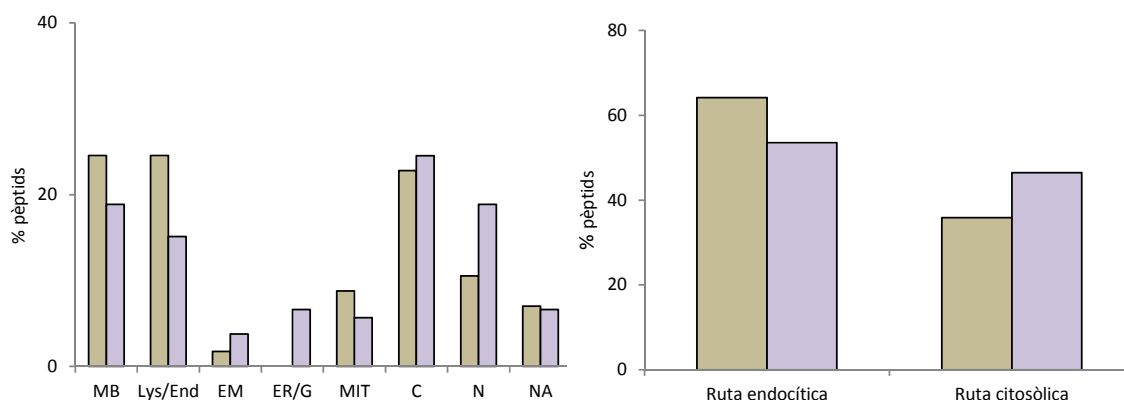


Figura 5.7: Esquerra: Localització cel·lular de les proteïnes parentals dels lligands d'I-E^d (marró) i I-A^d (morat) en melsa Balb/c. Dreta: Percentatge mitjà dels lligands d'I-A^d i I-E^d classificats segons la seva via degradativa.

5.3.4. Patró de processament dels pèptids aïllats en melsa Balb/c

Es va analitzar el possible patró de processament dels pèptids segons si generaven a partir dels extrems terminals o de la part interna de la proteïna parental (figura 5.8). Per I-E^d vam obtenir que un 43% dels pèptids presentaven patrons N-terminal (11%) i C-terminal (32%) i un 57% presentava localització

interna. A més, analitzàrem aquest patró relacionats amb la via de processament. Per I-A^d, un 9% presentava N-terminals, un 26% C-ter i un 66% una localització interna.

Tant per I-E^d com per I-A^d, no es van observar diferències en relació a la via de degradació i la localització del pèptid respecte la proteïna parental. Per ambdues molècules, els pèptids interns eren els majoritaris en les dues vies de degradació.

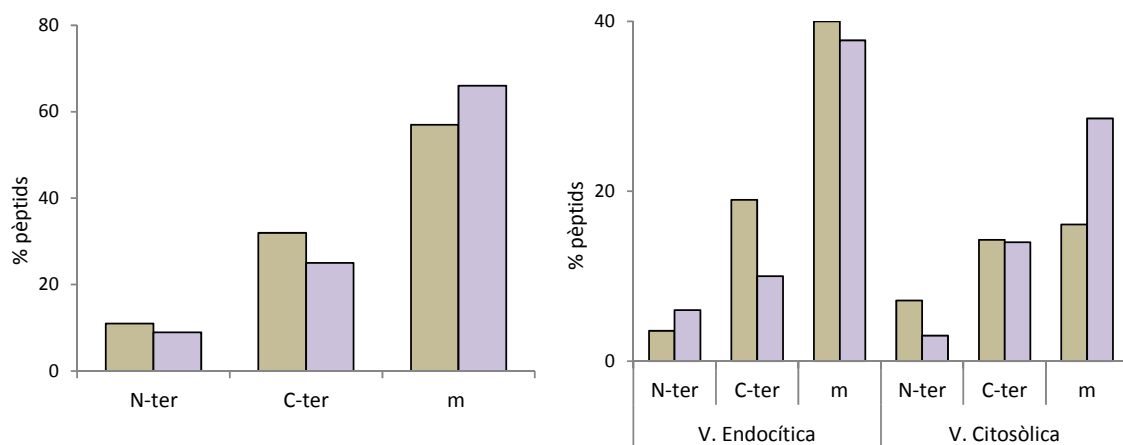


Figura 5.8: Patró de processament dels pèptids aïllats de melsa en I-A^d (morat) i I-E^d (marró). Gràfic en funció de la localització dels pèptids respecte la proteïna parental, depenent si eren N-terminal (N-ter), C-terminal (C-ter) o de la part interna. Al gràfic de l'esquerra, les localitzacions s'agrupaven en funció de la seva via de degradació, via endocítica i via citosòlica

5.4. COMPARACIÓ ENTRE EL REPERTORI PEPTÍDIC PRESENTAT EN TIMUS I MELSA DE BALB/C

La comparació del repertori de pèptids associats a MHC de classe II en un òrgan secundari com la melsa, comparat amb el d'un òrgan primari com el timus, hauria de revelar algunes característiques específiques d'aquells pèptids que indueixen tolerància i que la mantindrien a la perifèria. A fi d'identificar similituds i diferències entre ambdós repertoris, es va fer una anàlisi comparativa global dels pèptids associats a les molècules I-E^d i I-A^d en ambdós teixits.

5.4.1. En Balb/c, timus i melsa no comparteixen un ampli repertori peptídic

Es va analitzar el nombre de proteïnes compartides entre ambdós teixits. Es va observar que el timus i la melsa compartien un nombre relativament baix de proteïnes, fet que ja havíem observat al nostre laboratori amb aquests teixits en humans i en diferents al·lels (Collado et al, no publicat)

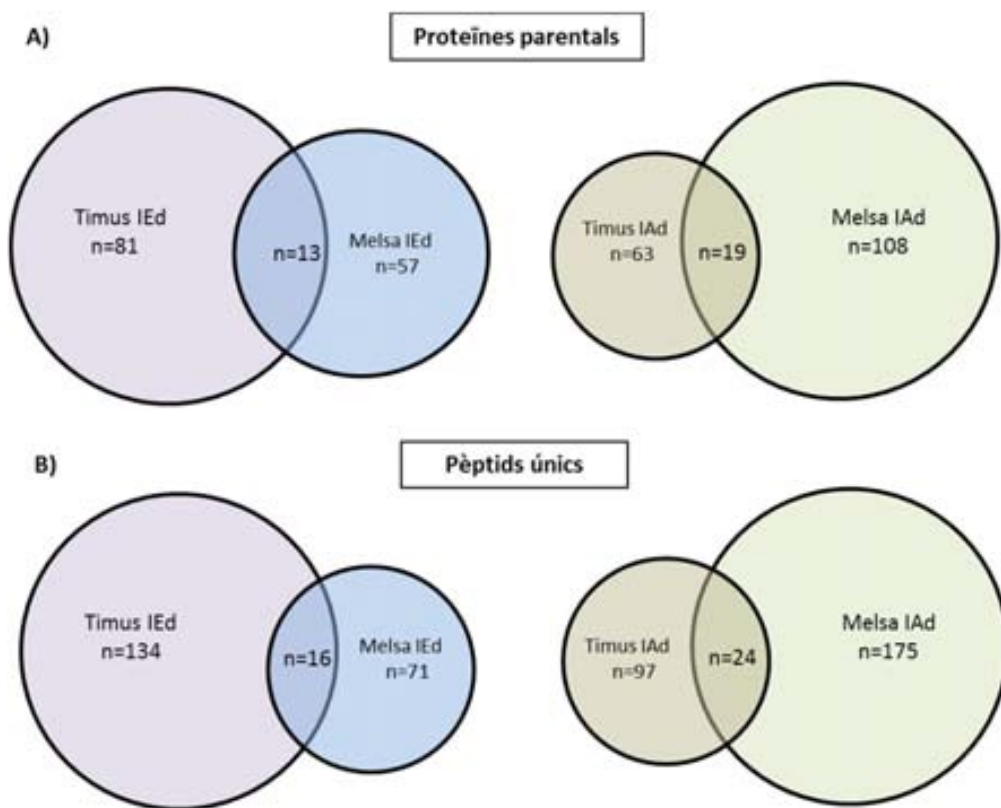


Figura 5.9. Diagrames de Venn que representen els repertoris obtinguts de les mostres de timus i melsa de ratolí balb/c. A). Representació de les proteïnes parentals seqüenciades en els diferents teixits. per a I-E^d (esquerra) i per a I-A^d (dreta); B) Representació dels pèptids idèntics seqüenciats en ambdós teixits per a I-E^d (esquerra) i per a I-A^d (dreta)

En el cas de I-E^d, de 57 proteïnes parentals en melsa i 81 en timus, només 13 eren comuns per a tots dos teixits. En el cas de I-A^d, de 108 proteïnes parentals en melsa i 63 en timus, només 19 eren comuns per a tots dos teixits (figura 5.9A)

Respecte al nombre de seqüències peptídiques idèntiques entre ambdós teixits de 134 pèptids en timus I-E^d i 71 en melsa, només 16 eren comuns per a tots dos teixits, el que suposa només un 12% dels lligands identificats en timus i un 22,5 % en la melsa.

En el cas de I-A^d, de 175 pèptids únics en melsa i 97 en timus, 24 eren comuns entre els dos teixits, es a dir, un 14% dels lligands en melsa i 25% en timus. (figura 5.9B)

5.4.2. Comparació de les rutes de processament de pèptids entre teixits

La localització cel·lular de les proteïnes parentals dels lligands peptídics per I-E^d va mostrar diferències a tenir en compte: a la melsa, la via endocítica és la majoritària mentre que al timus, les dues rutes semblen contribuir de manera equitativa en la producció de lligands. Un alt percentatge de proteïnes residents al citosol són degradades al timus, contribuint a l'augment d'aquesta via de degradació envers de la melsa (figura 5.10A).

En I-A^d, el percentatge de pèptids generats a la via endocítica i citosòlica era gairebé idèntic i alhora, idèntic entre timus i melsa (figura 5.10B).

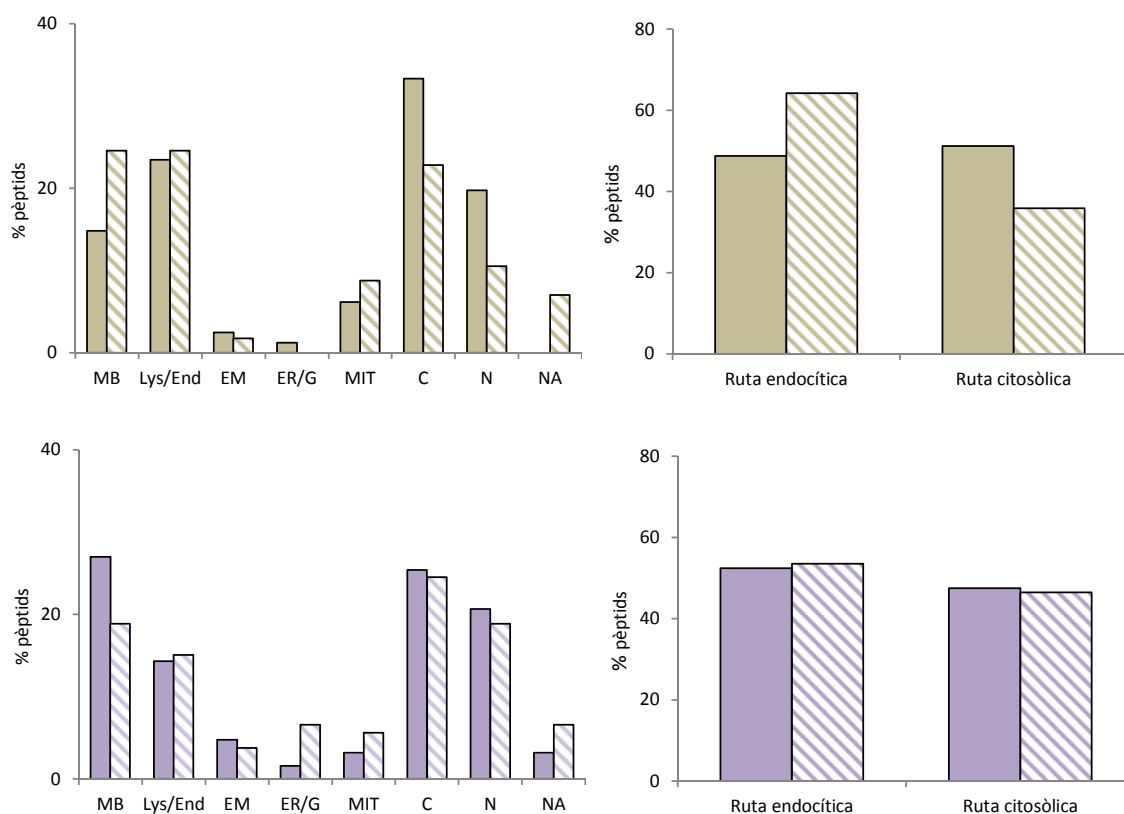


Figura 5.10: Esquerra: Localització cel·lular de les proteïnes parentals dels lligands d'I-E^d (marró) i I-A^d (morat) en timus (barres opaques) i melsa (barres ratllades) de Balb/c;. Dreta: Percentatge de lligands d'I-E^d i I-A^d classificats segons la seva via degradativa.

5.5. IDENTIFICACIÓ DE PÈPTIDS DERIVATS D'ANTÍGENS PERIFÈRICS (TRAs) PRESENTATS PER MHC EN TIMUS DE RATOLINS BALB/C

Entre els pèptids identificats com lligands d'MHC en timus de ratolí Balb/c es va identificar un pèptid derivat d'una proteïna amb expressió restringida a teixits perifèrics: la proteïna TBATA (thymus, brain and testes associated) amb funció en la diferenciació espermàtica (figura 5.10) segons la base de dades Genevestigator (<https://www.genevestigator.com/gv/doc/biomed/freetools.jsp>). Degut a que aquest gen no es trobava a la base de dades, es va agafar el seu homòleg en humà per tal de fer la cerca.

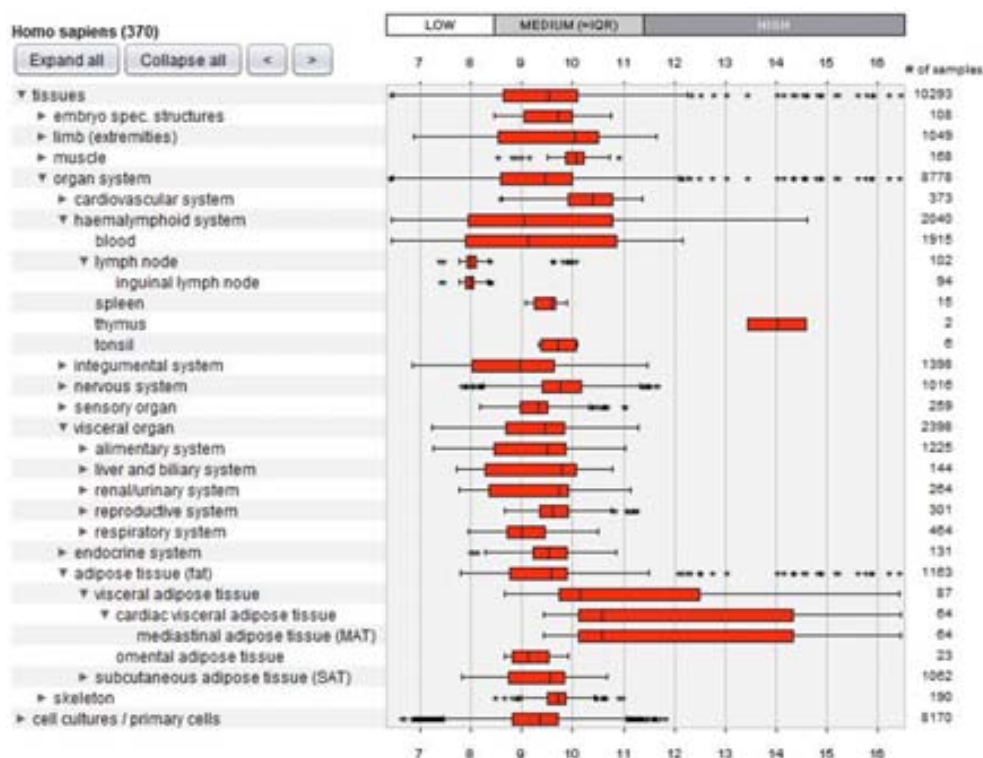


Figura 5.10: Expressió tissular del gen TBATA segons la base de dades Genevestigator.-

5.6. OBSERVACIONS FINALS

Com a aproximació a un estudi del repertori tolerogènic en mostres de timus i melsa de ratolí Balb/c es va estudiar el repertori peptídic associat a MHC, per fer una anàlisi comparativa d'ambdós repertoris. Un estudi similar s'havia realitzat ja al nostre laboratori en teixit humà, i vam creure interessant veure si els resultats eren comparables amb ratolí Balb/c, al·lel lligat a malalties de tipus Th2. Aquest estudi ens ha servit per una banda, per obtenir dades sobre els repertoris a ambdós teixits i per fer una comparació entre ambdues molècules H-2^d i a més, poder extrapol·lar els resultats amb el que ocorre en humans, concretament amb la molècula I-E^d anàleg a DR.

La diferència més important comparant amb les dades obtingudes a humans va ser que el repertori de la melsa presentava majoritàriament proteïnes secretades, sent aquesta localització (endocítica) la de major aportació al repertori (64%), mentre que en timus, la via endocítica i la via citosòlica aportarien aproximadament la mateixa quantitat. En humans, tant a timus com a melsa la via majoritària seria l'endocítica.

CONCLUSIONS

CONCLUSIONS

1. El motiu d'unió a HLA-DR3 basat en pèptids naturals ens indica que la posició P4 és la més important en la interacció, amb l'àcid glutàmic com a residu preferencial. A més, la posició P9 mostra preferència per residus bàsics, fet no descrit fins al moment.
2. Les xaperones DM i cadena invariant no afecten al repertori peptídic associat a DR3 tal i com s'esperava segons la funció canònica de les dues molècules. Aquest fet ens mostra que DR3 seria menys sensible a l'acció d'ambdues i ens podria explicar part de la susceptibilitat d'aquest al·lel per diverses enfermetats.
3. Les molècules DR3 expressades en les cèl·lules endocrines tant en presència com en absència d'I i DM estan associades a un repertori de pèptids homogeni en contingut i en localització bàsicament nuclear i citosòlica.
4. HLA-DR3 es comporta igual en el cas de ser una molècula d'expressió induïda que en el cas de tenir una expressió natural a la cèl·lula o de ser transfectada: preferència per pèptids citosòlics i localitzats a l'extrem c-terminal de les proteïnes.
5. No es van detectar pèptids derivats de l'eTSHR en els repertoris peptídics analitzats: ni en cèl·lules epitelials tiroïdals humanes transfectades ni en ratolins Balb/c transgènics.
6. A diferència amb el que ocorre amb humans, les dades obtingudes al repertori de la melsa del ratolí Balb/c presentava majoritàriament proteïnes secretades, sent aquesta localització (endocítica) la de major aportació al repertori (64%), mentre que en timus, la via endocítica i la via citosòlica aportarien aproximadament la mateixa quantitat.

BIBLIOGRAFIA

1. Campbell, R.D., *Map of the human MHC*. Immunol Today, 1993. **14**(7): p. 349-352.
2. Dunham, I., et al., *Molecular mapping of the human major histocompatibility complex by pulsed-field gel electrophoresis*. Proc Natl Acad Sci, 1987. **4**(20): p. 7237-41.
3. Goodfellow, P.N., *The beta2-microglobulin gene is on chromosome 15 and not in the HL-A region*. Nature, 1975. **254**(5497): p. 267-9.
4. Brown, J.H., et al., *Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1*. Nature, 1993. **364**(6432): p. 33-9.
5. Ghosh, P., et al., *The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3*. Nature, 1995. **378**(6556): p. 457-62.
6. Stern, L.J., et al., *Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide*. Nature, 1994. **368**(6468): p. 215-21.
7. Engelhard, V.H., *Structure of peptides associated with class I and class II MHC molecules*. Annu Rev Immunol, 1994. **12**: p. 181-207.
8. Bjorkman, P.J., et al., *Structure of the human class I histocompatibility antigen, HLA-A2*. Nature, 1987. **329**(6139): p. 506-12.
9. Jardetzky, T.S., et al., *Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides*. Proc Natl Acad Sci U S A, 1996. **93**(2): p. 734-8.
10. Madden, D., *The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC*. Cell, 1992. **70**(6): p. 1035-48.
11. Jardetzky, T.S., et al., *Identification of self peptides bound to purified HLA-B27*. Nature, 1991. **353**(6342): p. 326-9.
12. Rudensky, A. and P. Preston-Hurlburt, *Sequence analysis of peptides bound to MHC class II molecules*. Nature, 1991. **353**: p. 622-627.
13. Chicz, R.M., et al., *Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size*. Nature, 1992. **358**(6389): p. 764-8.
14. Sant'Angelo, D.B., et al., *Recognition of core and flanking amino acids of MHC class II-bound peptides by the T cell receptor*. Eur J Immunol, 2002. **32**(9): p. 2510-20.
15. Lovitch, S.B., Z. Pu, and E.R. Unanue, *Amino-terminal flanking residues determine the conformation of a peptide-class II MHC complex*. J Immunol, 2006. **176**(5): p. 2958-68.
16. Suri, A., et al., *Specificity of peptide selection by antigen-presenting cells homozygous or heterozygous for expression of class II MHC molecules: The lack of competition*. Proc Natl Acad Sci, 2003. **100**(9): p. 5330-5.
17. Chicz, R.M., et al., *Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles*. J Exp Med, 1993. **178**(1): p. 27-47.
18. Rudensky, A., et al., *Sequence analysis of peptides bound to MHC class II molecules*. Nature, 1991. **353**(6345): p. 622-7.
19. Cresswell, P., *Assembly, transport, and function of MHC class II molecules*. Annu Rev Immunol, 1994. **12**: p. 259-93.
20. Watts, C., *Capture and processing of exogenous antigens for presentation on MHC molecules*. Annu Rev Immunol, 1997. **15**: p. 821-50.
21. Anderson, K.S. and P. Cresswell, *A role for calnexin (IP90) in the assembly of class II MHC molecules*. Embo J, 1994. **13**(3): p. 675-82.

22. Cresswell, P., *Invariant chain structure and MHC class II function*. Cell, 1996. **84**(4): p. 505-7.
23. Warmerdam, P.A., E.O. Long, and P.A. Roche, *Isoforms of the invariant chain regulate transport of MHC class II molecules to antigen processing compartments*. J Cell Biol, 1996. **133**(2): p. 281-91.
24. Bakke, O. and B. Dobberstein, *MHC class II-associated invariant chain contains a sorting signal for endosomal compartments*. Cell, 1990. **63**(4): p. 707-16.
25. Villadangos, J.A., *MHC class II expression is regulated in dendritic cells independently of invariant chain degradation*. Immunity, 2001. **14**(6): p. 739-49.
26. Rocha, N. and J. Neefjes, *MHC class II molecules on the move for successful antigen presentation*. Embo J, 2008. **27**(1): p. 1-5.
27. Manoury, B., et al., *Asparagine endopeptidase can initiate the removal of the MHC class II invariant chain chaperone*. Immunity, 2003. **18**(4): p. 489-98.
28. Neumann, J., N. Schach, and N. Koch, *Glycosylation signals that separate the trimerization from the mhc class II-binding domain control intracellular degradation of invariant chain*. J Biol Chem, 2001. **276**(16): p. 13469-75.
29. Lovitch, S.B., S. Petzold, and E.R. Unanue, *Cutting Edge: H-2DM Is Responsible for the Large Differences in Presentation Among Peptides Selected by I-Ak During Antigen Processing*. J Immunol, 2003. **171**(5): p. 2183-6.
30. Collado, J.A., et al., *The Repertoires of Peptides Presented by MHC-II in the Thymus and in Peripheral Tissue: A Clue for Autoimmunity?* Front Immunol, 2013. **4**(442).
31. Klein, L., C. Münz, and J.D. Lünemann, *Autophagy-mediated antigen processing in CD4(+) T cell tolerance and immunity*. FEBS Lett, 2010. **584**(7): p. 1405-10.
32. Münz, C., *Enhancing immunity through autophagy*. Annu Rev Immunol, 2009. **27**: p. 423-49.
33. Giancchetti, E., D.V. Delfino, and A. Fierabracci, *Recent insights on the putative role of autophagy in autoimmune diseases*. Autoimmun Rev, 2014. **13**(3): p. 231-41.
34. Santambrogio, L., et al., *Extracellular antigen processing and presentation by immature dendritic cells*. Proc Natl Acad Sci, 1999. **96**(26): p. 15056-61.
35. Falk, K., et al., *Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules*. Nature, 1991. **351**(6324): p. 290-6.
36. Falk, K., et al., *Pool sequencing of natural HLA-DR, DQ, and DP ligands reveals detailed peptide motifs, constraints of processing, and general rules*. Immunogenetics, 1994. **39**(4): p. 230-42.
37. Corr, M., et al., *Endogenous peptides of a soluble major histocompatibility complex class I molecule, H-2Lds: sequence motif, quantitative binding, and molecular modeling of the complex*. J Exp Med, 1992. **176**(6): p. 1681-92.
38. Corr, M., et al., *H-2Dd exploits a four residue peptide binding motif*. J Exp Med, 1993. **178**(6): p. 1877-92.
39. Verreck, F.A., et al., *Identification of an HLA-DQ2 peptide binding motif and HLA-DPw3-bound self-peptide by pool sequencing*. Eur J Immunol, 1994. **24**(2): p. 375-9.
40. Hammer, J., et al., *Precise prediction of major histocompatibility complex class II-peptide interaction based on peptide side chain scanning*. J Exp Med, 1994. **180**(6): p. 2353-8.

41. Marshall, K.W., et al., *Prediction of peptide affinity to HLA DRB1*0401*. J Immunol, 1995. **154**(11): p. 5927-33.
42. Marshall, K.W., et al., *Prediction of peptide affinity to HLA DR molecules*. Biomed Pept Proteins Nucleic Acids, 1995. **1**(3): p. 157-62.
43. Hammer, J., *New methods to predict MHC-binding sequences within protein antigens*. Curr Opin Immunol, 1995. **7**(2): p. 263-9.
44. Hammer, J., et al., *Peptide binding specificity of HLA-DR4 molecules: correlation with rheumatoid arthritis association*. J Exp Med, 1995. **181**(5): p. 1847-55.
45. Sturniolo, T., et al., *Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices*. Nat Biotechnol, 1999. **17**(6): p. 555-61.
46. Egwuagu, C.E., P. Charukamnoetkanok, and I. Gery, *Thymic expression of autoantigens correlates with resistance to autoimmune disease*. J Immunol, 1997. **159**(7): p. 3109-12.
47. Liu, H., et al., *Mice resistant to experimental autoimmune encephalomyelitis have increased thymic expression of myelin basic protein and increased MBP specific T cell tolerance*. J Neuroimmunol, 2001. **115**(1-2): p. 118-26.
48. Sospedra, M., et al., *Transcription of a broad range of self-antigens in human thymus suggests a role for central mechanisms in tolerance toward peripheral antigens*. J Immunol, 1998. **161**(11): p. 5918-29.
49. Gebe, J.A., E. Swanson, and W.W. Kwok, *HLA class II peptide-binding and autoimmunity*. Tissue Antigens, 2002. **59**(2): p. 78-87.
50. Fairchild, P.J. and D.C. Wraith, *Lowering the tone: mechanisms of immunodominance among epitopes with low affinity for MHC*. Immunol Today, 1996. **17**(2): p. 80-85.
51. Mason, M.D. and M.F. Pera, *Immunohistochemical and biochemical characterisation of the expression of a human embryonal carcinoma cell proteoglycan antigen in human germ cell tumours and other tissues*. Eur J Cancer, 1992.
52. Dayan, C.M., et al., *Autoantigen recognition by thyroid-infiltrating T cells in Graves disease*. Proc Natl Acad Sci U S A, 1991. **88**(16): p. 7415-9.
53. Lacour, M., et al., *Type II collagen-specific human T cell lines established from healthy donors*. Eur J Immunol, 1990. **20**(4): p. 931-4.
54. Conant, S.B. and R.H. Swanborg, *MHC class II peptide flanking residues of exogenous antigens influence recognition by autoreactive T cells*. Autoimmun Rev, 2003. **2**(1): p. 8-12.
55. Bottazzo, G.F., et al., *Role of aberrant HLA-DR expression and antigen presentation in induction of endocrine autoimmunity*. Lancet, 1983. **2**(8359): p. 1115-9.
56. Hanafusa, T., et al., *Aberrant expression of HLA-DR antigen on thyrocytes in Graves' disease: relevance for autoimmunity*. Lancet, 1983. **2**(8359): p. 1111-5.
57. Zinkernagel, R.M., et al., *Antigen localisation regulates immune responses in a dose- and time-dependent fashion: a geographical view of immune reactivity*. Immunol Rev, 1997. **156**: p. 199-209.
58. Armengol, M.P., et al., *Thyroid autoimmune disease: demonstration of thyroid antigen-specific B cells and recombination-activating gene expression in chemokine-containing active intrathyroidal germinal centers*. Am J Pathol, 2001. **159**(3): p. 861-73.

59. Thrasyvoulides, A. and P. Lymberi, *Evidence for intramolecular B-cell epitope spreading during experimental immunization with an immunogenic thyroglobulin peptide*. Clin Exp Immunol, 2003. **132**(3): p. 401-7.
60. O'Leary, P.C., et al., *Investigations of thyroid hormones and antibodies based on a community health survey: the Busselton thyroid study*. Clin Endocrinol (Oxf), 2006. **64**(1): p. 97-104.
61. Hollowell, J.G., et al., *Serum TSH, T(4), and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III)*. J Clin Endocrinol Metab, 2002. **87**(2): p. 489-99.
62. Prentice, L.M., et al., *Geographical distribution of subclinical autoimmune thyroid disease in Britain: a study using highly sensitive direct assays for autoantibodies to thyroglobulin and thyroid peroxidase*. Acta Endocrinol (Copenh), 1990. **123**(5): p. 493-8.
63. Jacobson, E.M. and Y. Tomer, *The genetic basis of thyroid autoimmunity*. Thyroid, 2007. **17**(10): p. 949-61.
64. Ratanachaiyavong, S., et al., *Heat shock protein 70 (HSP70) and complement C4 genotypes in patients with hyperthyroid Graves' disease*. Clin Exp Immunol, 1991. **84**(1): p. 48-52.
65. Badenhoop, K., et al., *Tumor necrosis factor beta gene polymorphisms in Graves' disease*. J Clin Endocrinol Metab, 1992. **74**(2): p. 287-91.
66. Rees Smith, B., S.M. McLachlan, and J. Furmaniak, *Autoantibodies to the thyrotropin receptor*. Endocr Rev, 1988. **9**(1): p. 106-21.
67. Rapoport, B., et al., *The thyrotropin (TSH) receptor: interaction with TSH and autoantibodies*. Endocr Rev, 1998. **19**(6): p. 673-716.
68. Catalfamo, M., et al., *HLA-DM and invariant chain are expressed by thyroid follicular cells, enabling the expression of compact DR molecules*. Int Immunology, 1999. **11**(2): p. 269-77.
69. Vanderpump, M.P., et al., *The incidence of thyroid disorders in the community: a twenty-year follow-up of the Wickham Survey*. Clin Endocrinol (Oxf), 1995. **43**(1): p. 55-68.
70. Ludgate, M. and G. Baker, *Unlocking the immunological mechanisms of orbital inflammation in thyroid eye disease*. Clin Exp Immunol, 2002. **127**(2): p. 193-8.
71. Gerding, M.N., et al., *Association of thyrotrophin receptor antibodies with the clinical features of Graves' ophthalmopathy*. Clin Endocrinol (Oxf), 2000. **52**(3): p. 267-71.
72. Tomer, Y. and T.F. Davies, *Searching for the autoimmune thyroid disease susceptibility genes: from gene mapping to gene function*. Endocr Rev, 2003. **24**(5): p. 694-717.
73. Yanagawa, T., et al., *Human histocompatibility leukocyte antigen-DQA1*0501 allele associated with genetic susceptibility to Graves' disease in a Caucasian population*. J Clin Endocrinol Metab, 1993. **76**(6): p. 1569-74.
74. Barlow, A.B., et al., *Association of HLA-DQA1*0501 with Graves' disease in English Caucasian men and women*. Clin Endocrinol (Oxf), 1996. **44**(1): p. 73-7.
75. Laugwitz, K.L., et al., *The human thyrotropin receptor: a heptahelical receptor capable of stimulating members of all four G protein families*. Proc Natl Acad Sci U S A, 1996. **93**(1): p. 116-20.
76. Furmaniak, J. and B.R. Smith, *The structure of thyroid autoantigens*. Autoimmunity, 1990. **7**(1): p. 63-80.

77. Loosfelt, H., et al., *Two-subunit structure of the human thyrotropin receptor*. Proc Natl Acad Sci U S A, 1992. **89**(9): p. 3765-9.
78. Tanaka, K., et al., *The shed thyrotropin receptor is primarily a carboxyl terminal truncated form of the A subunit, not the entire A subunit*. Mol Cell Endocrinol, 1999. **150**(1-2): p. 113-9.
79. Nagayama, Y., et al., *Binding domains of stimulatory and inhibitory thyrotropin (TSH) receptor autoantibodies determined with chimeric TSH-lutropin/chorionic gonadotropin receptors*. J Clin Invest, 1991. **88**(1): p. 336-40.
80. Schwarz-Lauer, L., et al., *Evidence for a simplified view of autoantibody interactions with the thyrotropin receptor*. Thyroid, 2002. **12**(2): p. 115-20.
81. Chen, C.R., et al., *The thyrotropin receptor autoantigen in Graves disease is the culprit as well as the victim*. J Clin Invest, 2003. **111**(12): p. 1897-904.
82. McLachlan, S.M., Y. Nagayama, and B. Rapoport, *Insight into Graves' hyperthyroidism from animal models*. Endocr Rev, 2005. **26**(6): p. 800-32.
83. Gadzar, F.A., et al., *Continuous, clonal, insulin- and somatostatin-secreting cell lines established from a transplantable rat islet cell tumor*. Proc Natl Acad Sci, 1980. **77**(6): p. 3519-23.
84. Muntasell, A., et al., *HLA-DR4 molecules in neuroendocrine epithelial cells associate to a heterogeneous repertoire of cytoplasmic and surface self peptides*. J Immunol, 2002. **169**(9): p. 5052-60.
85. Pichurin, P., et al., *Targeted Expression of the human thyrotropin receptor A-subunit to the mouse thyroid: insight into the overcoming the lack of response to A-subunit Adenovirus immunization*. J Immunology, 2006. **176**: p. 668-676.
86. Chang, C.H., et al., *Class II transactivator (CIITA) is sufficient for the inducible expression of major histocompatibility complex class II genes*. J Exp Med, 1994. **180**(4): p. 1367-74.
87. Denzin, L., et al., *Assembly and intracellular transport of HLA-DM and correction of the class II antigen-processing defect in T2 cells*. Immunity, 1994. **1**(7): p. 595-606.
88. Quaranta, V., et al., *A human Ia cytoplasmic determinant located on multiple forms of invariant chain (gamma, gamma 2, gamma 3)*. J Immunology, 1984. **132**(4): p. 1900-5.
89. Muntasell, A. and M. Carrascal, *Dissection of the HLA-DR4 peptide repertoire in endocrine epithelial cells: strong influence of invariant chain and HLA-DM expression on the nature of ligands*. J Immunology, 2004. **173**(2): p. 1085-93.
90. Muixí, L., et al., *The peptide-binding motif of HLA-DR8 shares important structural features with other type 1 diabetes-associated alleles*. Genes and Immunity, 2011: p. 1-9.
91. Laemmli, U.K., *Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4*. Nature, 1970. **227**(5259): p. 680-885.
92. Elias, J. and S. Gygi, *Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry*. Nat Methods, 2007. **4**(3): p. 207-214.
93. Elias, J.E., *Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations*. Nat Methods, 2005. **2**(9): p. 667-675.
94. Brosch, M., *Accurate and Sensitive Peptide Identification with Mascot Percolator*. Journal of Proteome Research, 2009. **8**: p. 3176-3181.
95. Muixí, L., et al., *Thyroglobulin peptides associate in vivo to HLA-DR in autoimmune thyroid glands*. J.Immunol., 2008. **181**(1): p. 795-807.

96. Bernstein, F.C., et al., *The Protein Data Bank: A Computer-based Archival File For Macromolecular Structures*. J. Mol. Biol, 1997. **112**: p. 535-542.
97. Scott, C.A., et al., *Crystal structures of two I-Ad-peptide complexes reveal that high affinity can be achieved without large anchor residues*. Immunity, 1998. **8**(3): p. 319-329.
98. Scott, W.R.P., et al., *The GROMOS biomolecular simulation program package*. J. Phys. Chem. A 1999. **103**: p. 3596-3607.
99. Schild, H., et al., *Natural ligand motifs of H-2E molecules are allele specific and illustrate homology to HLA-DR molecules*. Int Immunol. , 1995. **7**(12): p. 1957-65.
100. Gunsteren, W.F.v., X. Daura, and A.E. Mark, *Computation of free energy*. Helv. Chim. Acta, 2002. **85**: p. 3113-3129.
101. Jarzynski, C., *Nonequilibrium equality for free energy differences*. Phys. Rev. Lett, 1997. **78**: p. 2690-2693.
102. Kersh, G.J., et al., *Structural and functional consequences of altering a peptide MHC anchor residue*. J Immunology, 2001. **166**: p. 3345-3354.
103. Geluk, A., et al., *Functional analysis of DR17(DR3)-restricted mycobacterial T cell epitopes reveals DR17-binding motif and enables the design of allele-specific competitor peptides*. J Immunology, 1992. **149**(9): p. 2864-71.
104. Geluk, A., et al., *HLA-DR3 molecules can bind peptides carrying two alternative specific submotifs*. J Immunology, 1994. **152**(12): p. 5742-5748.
105. Malcherek, G., et al., *Natural peptide ligand motifs of two HLA molecules associated with myasthenia gravis*. Int Immunology, 1993. **5**(10): p. 1229-37.
106. Rammensee, H., et al., *SYFPEITHI: database for MHC ligands and peptide motifs*. Immunogenetics, 1999. **50**(3-4): p. 213-9.
107. Hammer, J., et al., *Promiscuous and allele-specific anchors in HLA-DR-binding peptides*. Cell, 1993. **74**(1): p. 197-203.
108. Chang, K.Y., A. Suri, and E.R. Unanue, *Predicting peptides bound to I-Ag7 class II histocompatibility molecules using a novel expectation-maximization alignment algorithm*. Proteomics, 2007. **7**(3): p. 367-77.
109. Alvarez, I., et al., *The rheumatoid arthritis-associated allele HLA-DR10 (DRB1*1001) shares part of its repertoire with HLA-DR1 (DRB1*0101) and HLA-DR4 (DRB*0401)*. Arthritis Rheum. , 2008. **58**(6): p. 1630-9.
110. Kong, Y.C., et al., *Efficacy of HLA-DRB1*03:01 and H2E transgenic mouse strains to correlate pathogenic thyroglobulin epitopes for autoimmune thyroiditis*. J Autoimmunology, 2011. **37**(2): p. 63-70.
111. Rammensee, H.G., T. Friede, and S. Stevanović, *MHC ligands and peptide motifs: first listing*. Immunogenetics, 1995. **41**(4): p. 178-228.
112. Alvarez, I., et al., *The rheumatoid arthritis-associated allele HLA-DR10 (DRB1*1001) shares part of its repertoire with HLA-DR1 (DRB1*0101) and HLA-DR4 (DRB*0401)*. Arthritis Rheum, 2008. **58**(6): p. 1630-1639.
113. Hammer, J., *New methods to predict MHC-binding sequences within protein antigens*. Curr Opin Immunol, 1995. **7**: p. 263-269.
114. Robinson, J., et al., *IMGT/HLA and IMGT/MHC: Sequence Databases for the Study of the Major Histocompatibility Complex*. Nucleic Acids Research 2003. **31**(1): p. 311-314.
115. Muntasell, A., et al., *HLA-DR4 molecules in neuroendocrine epithelial cells associate to a heterogeneous repertoire of cytoplasmic and surface self peptides*. J Immunology, 2002. **169**(9): p. 5052-60.

116. Viville, S., et al., *Mice lacking the MHC class II-associated invariant chain*. Cell, 1993. **72**(4): p. 635-48.
117. Tourne, S., et al., *Functionality of major histocompatibility complex class II molecules in mice doubly deficient for invariant chain and H-2M complexes*. Proc Natl Acad Sci U S A., 1997. **94**(17): p. 9255-60.
118. Kenty, G., et al., *MHC class II expression in double mutant mice lacking invariant chain and DM functions*. J Immunology, 1998. **160**(2): p. 606-614.
119. Lightstone, L., et al., *In the absence of the invariant chain, HLA-DR molecules display a distinct array of peptides which is influenced by the presence or absence of HLA-DM*. Proc Natl Acad Sci, 1997. **94**(11): p. 5772-7.
120. Koonce, C.H., *DM loss in k haplotype mice reveals isotype-specific chaperone requirements*. J Immunology, 2003. **170**(7): p. 3751-61.
121. Riberdy, J.M. and P. Cresswell, *HLA-DR molecules from an antigen-processing mutant cell line are associated with invariant chain peptides*. Nature, 1992. **360**: p. 474-77.
122. Sette, A., et al., *Invariant chain peptides in most HLA-DR molecules of an antigen-processing mutant*. Science, 1992. **258**: p. 1801-4.
123. Mellins, E., et al., *Defective processing and presentation of exogenous antigens in mutants with normal HLA class II genes*. Nature, 1990. **343**(6253): p. 71-74.
124. Belfiore, A., *De novo HLA Class II and enhanced HLA Class I molecule expression in SV40 transfected human thyroid epithelial cells*. 1991. **4**(3).
125. LeibundGut-Landmann, S., *Mini-review: Specificity and expression of CIITA, the master regulator of MHC class II genes*. Eur J Immunol, 2004. **34**(6): p. 1513-25.
126. Watanabe, H., et al., *Innate immune response in Th1- and Th2-dominant mouse strains*. Shock, 2004. **22**(5): p. 460-6.
127. Chen, C.R., et al., *Susceptibility rather than resistance to hyperthyroidism is dominant in a thyrotropin receptor adenovirus-induced animal model of Graves' disease as revealed by BALB/c-C57BL/6 hybrid mice*. Endocrinology, 2004. **145**(11): p. 4927-4933.
128. Bentvelzen, P., et al., *Genetic transmission of viruses that incite mammary tumor in mice*. Proc Natl Acad Sci U S A, 1970. **67**(1): p. 377-384.
129. Ebbesen, P., *Reticulosarcoma and amyloid development in BALB/c mice inoculated with syngeneic cells from young and old donors*. J Natl Cancer Inst 1971. **47**(6): p. 1241-1245.
130. Roderick, T.H., S.H. Langley, and E.H. Leiter, *Some unusual genetic characteristics of BALB/c and evidence for genetic variation among BALB/c substrains*. Curr Top Microbiol Immunol, 1985. **122**: p. 9-18.
131. Kaithamana, S., et al., *Induction of experimental autoimmune Graves' disease in BALB/c mice*. J Immunol 1999. **163**: p. 5157-5164.
132. Muntasell, A., et al., *Dissection of the HLA-DR4 peptide repertoire in endocrine epithelial cells: strong influence of invariant chain and HLA-DM expression on the nature of ligands*. J Immunol, 2004. **173**(2): p. 1085-93.
133. Kumar, G. and S. Ranganathan, *Network analysis of human protein location*. BMC Bioinform. , 2010. **11**(Suppl 7): p. S9.
134. Sette, A., *I-Ad-binding peptides derived from unrelated protein antigens share a common structural motif*. J Immunol, 1988. **141**(1): p. 45-8.
135. Sette, A., et al., *Structural characteristics of an antigen required for its interaction with Ia and recognition by T cells*. Nature, 1987. **328**(6129): p. 395-9.

136. Berkower, I., G. Buckenmeyer, and J. Berzofsky, *Molecular mapping of a histocompatibility-restricted immunodominant T cell epitope with synthetic and natural peptides: implications for T cell antigenic structure*. J Immunol, 1986. **136**(7): p. 2498-503.
137. Bozzacco, L., et al., *Mass spectrometry analysis and quantitation of peptides presented on the MHC II molecules of mouse spleen dendritic cells*. J Proteome Res. , 2011. **10**(11): p. 5016-30.
138. Paliakasis, K., et al., *Novel structural features of the human histocompatibility molecules HLA-DQ as revealed by modeling based on the published structure of the related molecule HLA-DR*. J Struct Biol., 1996. **117**(2): p. 145-63.
139. James, A.E., et al., *The Binding of Antigenic Peptides to HLA-DR Is Influenced by Interactions between Pocket 6 and Pocket 9*. J Immunol., 2009. **183**(5): p. 3249–3258.
140. Rammensee, H., et al., *SYFPEITHI: database for MHC ligands and peptide motifs*. Immunogenetics, 1999. **50**(34): p. 213-9.