

**ANÁLISIS DE LOS INMUNOPEPTIDOMAS ASOCIADOS A  
HLA DE CLASE I: PAPEL DEL PROTEASOMA INTERMEDIO DE TIPO I  
Y ESTUDIO DE LOS SUBPEPTIDOMAS DEL SUPERTIPO B7**

**María Area Navarro**

**ADVERTIMENT.** L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  <https://creativecommons.org/licenses/?lang=ca>

**ADVERTENCIA.** El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <https://creativecommons.org/licenses/?lang=es>

**WARNING.** The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>



# **UAB**

## **Universitat Autònoma de Barcelona**

**Departament de Biologia Cel·lular, de Fisiologia i d'Immunologia**

**Facultat de Biociències**

**Memoria presentada por María Area Navarro para optar al grado de Doctor en  
Inmunología en el Programa de Doctorado en Inmunología Avanzada de la  
Universitat Autònoma de Barcelona**

### **ANÁLISIS DE LOS INMUNOPEPTIDOMAS ASOCIADOS A HLA DE CLASE I: PAPEL DEL PROTEASOMA INTERMEDIO DE TIPO I Y ESTUDIO DE LOS SUBPEPTIDOMAS DEL SUPERTIPO B7**

**Tesis realizada en el Laboratori d'Immunologia Cel·lular del Institut de Biotecnologia i  
Biomedicina de la Universitat Autònoma de Barcelona bajo la dirección de los  
doctores Iñaki Álvarez y José Ramón Palacio.**

**María Area Navarro**

Doctoranda

**Iñaki Álvarez Pérez**

Director y tutor

**José Ramón Palacio Cornide**

Director

**Octubre 2025**





# **CONTENIDOS**

|                                                                 |           |
|-----------------------------------------------------------------|-----------|
| <b>ABREVIATURAS</b>                                             | <b>11</b> |
| <b>RESUMEN</b>                                                  | <b>13</b> |
| RESUMEN DE TESIS                                                | 15        |
| THESIS RESUME                                                   | 19        |
| RESUM DE TESI                                                   | 23        |
| <b>INTRODUCCIÓN</b>                                             | <b>27</b> |
| COMPLEJO MAYOR DE HISTOCOMPATIBILIDAD                           | 29        |
| 1. GENÉTICA Y ESTRUCTURA                                        | 29        |
| 2. PROCESAMIENTO ANTIGÉNICO                                     | 33        |
| 3. HLA Y ENFERMEDADES                                           | 39        |
| PROTEASOMA                                                      | 41        |
| 1. SISTEMA UBIQUITÍN-PROTEASOMA                                 | 41        |
| 2. ESTRUCTURA DEL PROTEASOMA                                    | 43        |
| 3. ENSAMBLAJE DEL PROTEASOMA 20S                                | 46        |
| 4. SUBTIPOS DE PROTEASOMA                                       | 47        |
| SUPERTIPOS Y SUBPEPTIDOMAS DE HLA                               | 51        |
| 1. SUPERTIPO B7                                                 | 53        |
| <b>HIPÓTESIS</b>                                                | <b>55</b> |
| <b>OBJETIVOS</b>                                                | <b>59</b> |
| <b>MATERIALES Y MÉTODOS</b>                                     | <b>63</b> |
| CULTIVOS CELULARES                                              | 65        |
| 1. LÍNEAS CELULARES                                             | 65        |
| 2. CONDICIONES DE CULTIVO                                       | 65        |
| CONSTRUCCIONES GÉNICAS                                          | 67        |
| 1. DISEÑO DE <i>PRIMERS</i>                                     | 67        |
| 2. CONSTRUCCIONES GÉNICAS                                       | 68        |
| 3. EXTRACCIÓN DE RNA, RETROTRANSCRIPCIÓN Y GADPH                | 69        |
| 4. CLONAJE                                                      | 70        |
| 5. LIGACIÓN, TRANSFORMACIÓN Y SECUENCIACIÓN                     | 71        |
| ANTICUERPOS                                                     | 72        |
| CITOMETRÍA DE FLUJO                                             | 73        |
| 1. TINCIÓN CELULAR DE SUPERFICIE                                | 73        |
| 2. TINCIÓN INTRACELULAR                                         | 73        |
| 3. TOMA Y ANÁLISIS DE DATOS                                     | 74        |
| CARACTERIZACIÓN DEL PROTEASOMA INTERMEDIO DE TIPO I             | 75        |
| 1. TRANSFECCIÓN                                                 | 75        |
| 2. WESTERN BLOT                                                 | 76        |
| 3. PURIFICACIÓN DEL PROTEASOMA 20S                              | 77        |
| 4. ENSAYO DE ACTIVIDAD CATALÍTICA                               | 78        |
| 5. ESTRUCTURA DEL PROTEASOMA                                    | 78        |
| EVALUACIÓN DE LA EXPRESIÓN DE HLA-A*03:01 EN HEK-293 Y HEK-293F | 80        |
| 1. PCR CUANTITATIVA - SYBER GREEN                               | 80        |

|                                                                                                                                                        |            |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>EVALUACIÓN DE LA PRESENCIA DE SUBPEPTIDOMAS EN MOLÉCULAS DE HLA-I DEL SUPERTIPO B7</b>                                                              | <b>81</b>  |
| 1. PEPTIDOMAS DE HLA-B                                                                                                                                 | 81         |
| 2. ANÁLISIS DE ESTRUCTURAS CRISTALOGRÁFICAS                                                                                                            | 81         |
| 3. MUTAGÉNESIS DIRIGIDA                                                                                                                                | 81         |
| 4. TRANSFECCIÓN POR ELECTROPORACIÓN DE CÉLULAS EN SUSPENSIÓN                                                                                           | 82         |
| <b>PURIFICACIÓN DE LOS INMUNOPEPTIDOMAS DE HLA DE CLASE I</b>                                                                                          | <b>83</b>  |
| 1. PURIFICACIÓN DE INMUNOPEPTIDOMAS DE MOLÉCULAS DE HLA-I                                                                                              | 83         |
| 2. ESPECTROMETRÍA DE MASAS                                                                                                                             | 84         |
| 3. ANÁLISIS DE RESULTADOS Y ESTADÍSTICA                                                                                                                | 86         |
| <b>RESULTADOS</b>                                                                                                                                      | <b>87</b>  |
| <b>CAPÍTULO I: CARACTERIZACIÓN DEL PROTEASOMA INTERMEDIO DE TIPO I</b>                                                                                 | <b>89</b>  |
| 1. GENERACIÓN DE MODELOS CELULARES PARA LA CARACTERIZACIÓN DEL PROTEASOMA INTERMEDIO DE TIPO I                                                         | 91         |
| 2. ANÁLISIS DEL REPERTORIO PEPTÍDICO DE HEK-293F y HEK-293F-β5i-StII                                                                                   | 93         |
| 2.1. Localización celular                                                                                                                              | 94         |
| 2.2. Asignación alélica de los péptidos presentes en el inmunopeptidoma de HLA-I                                                                       | 95         |
| 2.3. Afinidad teórica de los péptidos presentes en el inmunopeptidoma de HLA-I                                                                         | 98         |
| 2.4. Análisis del residuo C-terminal de los péptidos presentes en el inmunopeptidoma de HLA-I                                                          | 99         |
| 3. ANÁLISIS DE LA EXPRESIÓN DE HLA-A*03:01 EN LAS LINEAS HEK-293 Y HEK-293F                                                                            | 100        |
| 3.1. Análisis de la expresión en membrana de las moléculas HLA-A*02, -A*03 y -B*07 en las líneas celulares HEK-293 y HEK-293F por citometría de flujo  | 100        |
| 3.2. Análisis de la expresión génica de las moléculas HLA-A*02:01, -A*03:01 y -B*07:02 en las líneas celulares HEK-293 y HEK-293F por PCR cuantitativa | 102        |
| 4. ANÁLISIS DEL REPERTORIO PEPTÍDICO DE HEK-293 Y HEK-293-β5i                                                                                          | 104        |
| 4.1. Localización celular                                                                                                                              | 106        |
| 4.2. Asignación alélica de los péptidos presentes en el inmunopeptidoma de HLA-I                                                                       | 107        |
| 4.3. Afinidad teórica de los péptidos presentes en el inmunopeptidoma de HLA-I                                                                         | 109        |
| 4.4. Análisis del residuo C-terminal de los péptidos presentes en el inmunopeptidoma de HLA-I                                                          | 110        |
| 5. PURIFICACIÓN DEL PROTEASOMA 20S                                                                                                                     | 113        |
| 6. EVALUACIÓN DE LA ACTIVIDAD CATALÍTICA                                                                                                               | 114        |
| 7. ESTRUCTURA                                                                                                                                          | 118        |
| <b>CAPÍTULO II: ANÁLISIS DEL REPERTORIO PEPTÍDICO DE MOLÉCULAS DE HLA-I PERTENECIENTES AL SUPERTIPO B7</b>                                             | <b>123</b> |
| <b>ANÁLISIS DE LA PRESENCIA DE SUBPEPTIDOMAS EN LAS MOLÉCULAS DE HLA-I PERTENECIENTES AL SUPERTIPO B7</b>                                              | <b>125</b> |
| 1. ANÁLISIS DE LOS INMUNOPEPTIDOMAS PRESENTADOS POR HLA-I DE MOLÉCULAS DEL SUPERTIPO B7                                                                | 125        |
| 1.2. Grado de solapamiento de los inmunopeptidomas                                                                                                     | 127        |
| 1.2. Residuos de anclaje                                                                                                                               | 129        |
| 1.3. Afinidad teórica de los ligandos de moléculas de HLA del supertipo B7                                                                             | 129        |
| 2. ANÁLISIS DE LA PRESENCIA DE SUBPEPTIDOMAS DE LAS MOLÉCULAS DEL SUPERTIPO B7 ANALIZADAS                                                              | 131        |
| 2.1. HLA-B*54:01 presenta dos subpeptidomas similares a HLA-B*51:01 y HLA-B*51:08                                                                      | 131        |

|                                                                                                                                                               |                                                                                                                                                                                                             |            |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 2.2.                                                                                                                                                          | El grupo alélico HLA-B*35 y la molécula HLA-B*53:01 presentan dos subpeptidomas con residuos polares y ácidos en P1 en el subpeptidoma que contiene Ala2                                                    | 133        |
| 2.3.                                                                                                                                                          | Otras moléculas del supertipo B7 no presentan subpeptidomas                                                                                                                                                 | 135        |
| 2.4.                                                                                                                                                          | La presencia de Thr45 y de Phe67 podría asociarse a un incremento de la presencia de péptidos con Ala2 y de subpeptidomas en moléculas del supertipo B7                                                     | 137        |
| <b>ANÁLISIS DE LA APORTACIÓN DE LAS POSICIONES 45 Y 67 A LA PRESENCIA DE Ala2 Y DE SUBPEPTIDOMAS EN EL REPERTORIO PEPTÍDICO DE MOLÉCULAS DEL SUPERTIPO B7</b> |                                                                                                                                                                                                             |            |
| 1.                                                                                                                                                            | GENERACIÓN DE MODELOS CELULARES PARA EVALUAR LA RELEVANCIA DE LOS RESIDUOS LOCALIZADOS EN POSICIONES 45 Y 67 EN LA CONFIGURACIÓN DEL REPERTORIO PEPTÍDICO EN LAS MOLÉCULAS HLA-B*51:01, -B*35:01 Y -B*07:02 | 141        |
| 2.                                                                                                                                                            | PURIFICACIÓN DE LOS PÉPTIDOS DE UNIÓN A HLA-I DE LOS DIFERENTES TRANSFECTANTES                                                                                                                              | 143        |
| 3.                                                                                                                                                            | ANÁLISIS DE LOS INMUNOPEPTIDOMAS DE LAS MOLÉCULAS HLA-B*51:01, -B*35:01 Y -B*07:02 Y DE SUS VARIANTES MUTADAS EN POSICIONES 45 Y 67                                                                         | 144        |
| 3.1.                                                                                                                                                          | Localización celular                                                                                                                                                                                        | 145        |
| 3.2.                                                                                                                                                          | Grado de solapamiento de los inmunopectidomas                                                                                                                                                               | 147        |
| 3.3.                                                                                                                                                          | Residuos de anclaje                                                                                                                                                                                         | 148        |
| 3.4.                                                                                                                                                          | Afinidad teórica                                                                                                                                                                                            | 150        |
| 4.                                                                                                                                                            | IMPLICACIÓN DE LAS POSICIONES 45 Y 67 EN LA CONFORMACIÓN DEL REPERTORIO PEPTÍDICO DE LAS MOLÉCULAS HLA-B*51:01, HLA-B*35:01 Y HLA-B*07:02                                                                   | 152        |
| 4.1.                                                                                                                                                          | Los residuos que ocupan las posiciones 45 y 67 afectan a la presencia de péptidos con Ala2 en los repertorios peptídicos de las moléculas de HLA-I analizadas                                               | 152        |
| 4.2.                                                                                                                                                          | Los residuos que ocupan las posiciones 45 y 67 afectan a la presencia de subpeptidomas en diferente grado en los repertorios peptídicos de las moléculas de HLA-I analizadas                                | 154        |
| <b>DISCUSIÓN</b>                                                                                                                                              |                                                                                                                                                                                                             | <b>163</b> |
| <b>CONCLUSIONES</b>                                                                                                                                           |                                                                                                                                                                                                             | <b>181</b> |
| <b>BIBLIOGRAFÍA</b>                                                                                                                                           |                                                                                                                                                                                                             | <b>187</b> |
| <b>ANEXOS</b>                                                                                                                                                 |                                                                                                                                                                                                             | <b>209</b> |



## **ABREVIATURAS**



**Aa:** Aminoácido  
**ACE:** Enzima Convertidora de Angiotensina  
**AMC:** 7-amino-4-metilcumarina  
**ATP:** Adenosín trifosfato  
**β2m:** β2-microglobulina  
**BCA:** *Bicinchoninic Acid Assay*  
**CLIP:** *Class II Associated li Peptide*  
**cTECs:** Células Epiteliales Tímicas Corticales (*Cortical Thymic Epithelial Cells*)  
**CTL:** *Cytotoxic T Cell*  
**DEPC:** Dietilpirocarbonato  
**DIP:** *Double Intermediate Proteasome*  
**DMEM:** *Dulbecco's Modified Eagle Medium*  
**DNA:** Ácido desoxirribonucleico (*Deoxyribonucleic acid*)  
**dNTPs:** *Deoxyribonucleotide triphosphate*  
**EBV:** *Epstein Barr virus*  
**EDTA:** *Ethylenediaminetetraacetic acid*  
**FBS:** Suero Fetal Bovino (*Fetal bovine serum*)  
**FDR:** *False Discovery Rate*  
**HEK:** *Human Embryonic Kidney Cells*  
**HLA:** Antígeno Leucocitario Humano (*Human Leukocyte Antigen*)  
**HPLC:** *High Performance Liquid Chromatography*  
**HSP:** *Heat Shock Protein*  
**IDE:** Enzima Degradadora de Insulina  
**INF:** Interferón  
**KO:** *Knock-out*  
**LB:** *Lysogeny Broth*  
**Li:** *Invariant chain*  
**M-MLV RT:** *Moloney Murine Leukemia Virus*  
**MHC:** Complejo Principal de Histocompatibilidad (*Major Histocompatibility Complex*)  
**mTECs:** Células Epiteliales Tímicas Medulares (*Medullar Thymic Epithelial Cells*)  
**PAC:** *Proteasome Assembly Chaperone*  
**PAGE:** *Polyacrylamide Gel Electrophoresis*  
**PBMCs:** *Peripheral Blood Mononuclear Cells*  
**PBS:** *Phosphate Buffered Saline*  
**PEI:** *Polyethyleneimine*  
**PFA:** Paraformaldehído  
**PLC:** *Peptide-Loading Complex*  
**POMP:** *Proteasome Maturation Protein*  
**PVDF:** *Polyvinylidene Difluoride Membranes*  
**RCF:** Fuerza Centrífuga Relativa (*Relative Centrifugal Force*)  
**RE:** Retículo Endoplasmático  
**RNA:** Ácido ribonucleico (*Ribonucleic acid*)  
**RPM:** Revoluciones por minuto  
**SDS:** *Sodium Dodecyl Sulfate*  
**SIP:** *Single Intermediate Proteasome*  
**TAP:** *Transporter Associated with Antigen Presentation*  
**TFA:** *Trifluoroacetic Acid*  
**TNF:** *Tumor Necrosis Factor*  
**TOP:** Oligopeptidasa Thimet  
**TPPII:** Tripeptidil Peptidasa  
**UPS:** Sistema Ubiquitín-Proteasoma





## **RESUMEN**



## **RESUMEN DE TESIS**

El Complejo Principal de Histocompatibilidad (MHC) o Antígeno Leucocitario Humano (HLA) en humanos, codifica glicoproteínas de membrana que comúnmente se agrupan en dos clases: MHC de clase I, que presentan péptidos derivados de proteínas citosólicas y nucleares a los linfocitos T CD8<sup>+</sup>, y MHC de clase II, que presentan péptidos derivados de proteínas extracelulares a los linfocitos T CD4<sup>+</sup>. Para poder ser presentadas por las moléculas de MHC-I, las proteínas deben ser degradadas y procesadas previamente en el citosol, acción que lleva a cabo principalmente el proteasoma, componente principal del Sistema Ubiquitín-Proteasoma (UPS). Se trata de un gran complejo catalítico muy conservado entre las especies cuyo objetivo son proteínas poliubiquitinadas en residuos de lisina. Una parte de los péptidos generados por el proteasoma son transportados al lumen del retículo endoplasmático, donde son cargados en moléculas de HLA de clase I y, posteriormente, presentados a los linfocitos T CD8<sup>+</sup>. El núcleo catalítico, o proteasoma 20S, está formado por el ensamblaje de 4 anillos heptaméricos, 2 anillos exteriores compuestos por subunidades estructurales  $\alpha$  y 2 anillos interiores por subunidades  $\beta$ . La actividad catalítica reside en las subunidades  $\beta$ 1,  $\beta$ 2 y  $\beta$ 5. Además del proteasoma constitutivo ( $\beta$ 1,  $\beta$ 2, y  $\beta$ 5), existen otros tipos de proteasoma que reemplazan las subunidades estándar por otras análogas. Entre ellos están: el inmunoproteasoma ( $\beta$ 1i,  $\beta$ 2i, y  $\beta$ 5i), el timoproteasoma ( $\beta$ 1i,  $\beta$ 2i, y  $\beta$ 5t) y el espermatoproteasoma ( $\alpha$ 4s). Más recientemente se ha descrito la presencia de proteasomas intermedios, compuestos por una mezcla de subunidades del proteasoma constitutivo y del inmunoproteasoma: el proteasoma intermedio de tipo I (*single intermediate proteasome*, SIP) ( $\beta$ 1,  $\beta$ 2, y  $\beta$ 5i) y el proteasoma intermedio de tipo II (*double intermediate proteasome*, DIP) ( $\beta$ 1i,  $\beta$ 2, y  $\beta$ 5i). La presencia de diferentes subunidades catalíticas produce una especificidad diferente y debería implicar cambios en los inmunopeptidomas asociados a MHC-I. La primera parte de esta tesis aborda el análisis del inmunopeptidoma de clase I generado por el SIP y la caracterización de la especificidad y la estructura de dicho proteasoma 20S. Para ello, se generaron modelos celulares en las líneas HEK-293 y HEK-293F con expresión exclusiva del proteasoma constitutivo o del SIP. El análisis del inmunopeptidoma puso de manifiesto diferencias en la composición de los mismos. Así, había un incremento de péptidos con residuos aromáticos y básicos y una disminución

de Ala y Val en la posición P $\Omega$  en la línea que expresaba el SIP. Mediante ensayos de actividad catalítica utilizando péptidos fluorogénicos se vio que el SIP resultaba más eficiente catalizando el enlace tras Trp y Arg, mientras que el proteasoma constitutivo era más eficiente catalizando tras Ala. Finalmente, la resolución de la estructura del SIP por microscopía criogénica (Cryo-EM) reveló que la subunidad  $\beta 5i$  presentaba un *pocket* S1 más amplio que la subunidad  $\beta 5$ , lo que favorecería que acomodara residuos aromáticos. Estos resultados demuestran que la presencia del SIP modifica los péptidos generados y presentados por HLA-I, lo que sugiere que pueden modificar el desarrollo de la respuesta inmune.

Asimismo, el análisis de los inmunopeptidomas generados por el SIP y por el proteasoma constitutivo en el modelo celular HEK-293F mostró la ausencia de péptidos con extremo C-terminal básico. En esta tesis se ha analizado la expresión del alelo *HLA-A\*03:01* (que une péptidos básicos en posición P $\Omega$ ) en la línea celular HEK-293F mediante qPCR (SYBER Green). Los niveles de mRNA de *HLA-A\*03:01* fueron inferiores en relación a los expresados en la línea HEK-293, pero se detectaron niveles considerables. Sin embargo, mediante citometría de flujo con anticuerpos específicos de alelo, se demostró la ausencia de expresión de *HLA-A\*03:01* en la superficie celular de la línea HEK-293F. Este resultado explica la ausencia de péptidos con extremo C-terminal básico en el inmunopeptidoma de estas células.

La segunda parte de esta tesis trata de dilucidar la presencia de subpeptidomas en moléculas de HLA-I pertenecientes al supertipo B7. El análisis *in silico* de los inmunopeptidomas de HLA-I descritos en la literatura de 13 moléculas pertenecientes a este supertipo muestra que diversos alotipos presentaban subpeptidomas: uno de ellos con Pro2 y residuos hidrofóbicos en P1 y otro formado por Ala2 y residuos polares y ácidos en P1 (*HLA-B\*51:01*, *-B\*51:08*, *-B\*54:01*, *HLA-B\*35:01*, *-B\*35:03*, *-B\*35:07* y *-B\*53:01*) mientras que otros no presentaban subpeptidomas (*HLA-B\*07:02*, *-B\*07:04*, *-B\*42:01*, *-B\*55:01* *-B\*55:02* y *-B\*56:01*). Además, datos publicados sugieren que la presencia de determinados residuos en las posiciones 45 y 67 (polimórficas en este supertipo) son relevantes en la presencia de Ala2 y, tal vez, de diferentes subpeptidomas. Para analizar el papel de dichos residuos, se generaron mutantes simples y dobles en las posiciones 45 y 67 de 3 moléculas representativas del supertipo

B7: HLA-B\*51:01, HLA-B\*35:01 y HLA-B\*07:02. Las moléculas que presentaban mayor porcentaje de Ala2 y diferentes subpeptidomas presentaban Thr45 y Phe67, y las moléculas que no presentaban subpeptidomas y contenían menor porcentaje de Ala2 presentaban Glu45 y Tyr67. Se realizó mutagénesis dirigida de estas posiciones con el fin de determinar la relevancia de estos residuos en la prevalencia de Ala2 y en la presencia o ausencia de subpeptidomas. El análisis de los inmunopeptidomas de estos transfectantes puso de manifiesto que estas posiciones no resultan fundamentales para la presencia de subpeptidomas, pero sí condicionan la abundancia de péptidos con Ala2 en el repertorio peptídico. En estructuras descritas en la literatura, Glu45 y Tyr67 interaccionan con Asn63, dando lugar a un *pocket* más restrictivo en tamaño que solo acomoda Pro2, sin embargo, Thr45 y Phe67 no interaccionan con Asn63, dando lugar a un *pocket* más laxo.



## **THESIS RESUME**

The Major Histocompatibility Complex (MHC), or Human Leukocyte Antigen (HLA) in humans, encodes membrane glycoproteins that are generally classified into two groups: MHC class I molecules, which present peptides derived from cytosolic and nuclear proteins to CD8<sup>+</sup> T lymphocytes, and MHC class II molecules, which present peptides derived from extracellular proteins to CD4<sup>+</sup> T lymphocytes. To be presented by HLA class I molecules, proteins must first be degraded and processed in the cytosol, a process carried out mainly by the proteasome, the central component of the Ubiquitin-Proteasome System (UPS). This is a large, highly conserved catalytic complex that targets polyubiquitinated proteins on lysine residues. A subset of peptides generated by the proteasome are transported to the lumen of the endoplasmic reticulum (ER), where they are loaded in HLA class I molecules and subsequently presented to CD8<sup>+</sup> T lymphocytes. The catalytic core, or 20S proteasome, is composed by four heptameric rings: two outer rings of  $\alpha$  subunits and two inner rings of  $\beta$  subunits. The catalytic activity resides in the  $\beta$ 1,  $\beta$ 2, and  $\beta$ 5 subunits, each of which present different substrate specificity. In addition to the constitutive proteasome ( $\beta$ 1,  $\beta$ 2, and  $\beta$ 5), other types of proteasomes exist in which the standard catalytic subunits are replaced by alternative counterparts. These include the immunoproteasome ( $\beta$ 1i,  $\beta$ 2i, and  $\beta$ 5i), the thymoproteasome ( $\beta$ 1i,  $\beta$ 2i, and  $\beta$ 5t), and the spermatoproteasome ( $\alpha$ 4s). More recently, intermediate proteasomes have been described, composed of constitutive and immune subunits: the type I intermediate proteasome (single intermediate proteasome, SIP) ( $\beta$ 1,  $\beta$ 2, and  $\beta$ 5i) and the type II intermediate proteasome (double intermediate proteasome, DIP) ( $\beta$ 1i,  $\beta$ 2, and  $\beta$ 5i). The incorporation of different catalytic subunits alters substrate specificity, which in turn is expected to affect the composition of MHC class I immunopeptidomes. The first part of this thesis is focused on analyzing the class I immunopeptidome generated by the SIP, and on characterizing the specificity and structure of this 20S proteasome. To this end, cell models were generated in HEK-293 and HEK-293F lines with exclusive expression of either the constitutive proteasome or the SIP. Immunopeptidome analysis revealed compositional differences: in the SIP-expressing line, there was an increase in peptides with aromatic and basic residues, and a decrease in Ala and Val at the C-terminal (P $\Omega$ ) position. Catalytic activity assays with

fluorogenic peptides showed distinct substrate preferences between SIP and the constitutive proteasome: SIP was more efficient at cleaving after Trp and Arg residues, while the constitutive proteasome was more efficient at cleaving after small hydrophobic residues such as Ala. Finally, the 3D structure of SIP resolved by cryo-electron microscopy (cryo-EM) revealed that the  $\beta 5i$  subunit presents a wider S1 pocket than  $\beta 5$ , what would favor the accommodation of aromatic residues. These results demonstrate that SIP modifies the peptides generated and thus shapes the HLA-I antigenic repertoire, potentially impacting immune responses.

Furthermore, the analysis of immunopeptidomes from SIP and constitutive proteasomes in the HEK-293F model revealed an absence of peptides with basic C-terminal residues. In this thesis, expression of the *HLA-A\*03:01* allele (which binds peptides with basic residues at the P $\Omega$  position) was examined in the HEK-293F line by SYBR Green qPCR. Although *HLA-A\*03:01* mRNA was detectable at lower levels than those in HEK-293, flow cytometry with allele-specific antibodies confirmed the absence of *HLA-A\*03:01* surface expression in HEK-293F cells. This result explains the lack of C-terminal basic peptides in the immunopeptidome of this line.

The second part of this thesis aims to elucidate the presence of subpeptidomes in HLA class I molecules belonging to the B7 supertype. In silico analysis of published HLA-I immunopeptidomes from 13 B7-supertype molecules showed that several allotypes present subpeptidomes: one with Pro2 and hydrophobic residues at P1, and other with Ala2 and polar or acidic residues at P1 (*HLA-B\*51:01*, *-B\*51:08*, *-B\*54:01*, *HLA-B\*35:01*, *-B\*35:03*, *-B\*35:07*, and *-B\*53:01*), while others did not (*HLA-B\*07:02*, *-B\*07:04*, *-B\*42:01*, *-B\*55:01*, *-B\*55:02*, and *-B\*56:01*). Published data further suggest that specific polymorphic residues at positions 45 and 67 within this supertype may influence the presence of Ala2 and, possibly, the presence of distinct subpeptidomes. To test this, single and double mutants were generated at positions 45 and 67 in three representative B7-supertype molecules: *HLA-B\*51:01*, *HLA-B\*35:01*, and *HLA-B\*07:02*. Molecules with higher Ala2 prevalence and distinct subpeptidomes typically contained Thr45 and Phe67, whereas those lacking subpeptidomes and showing reduced Ala2 contained Glu45 and Tyr67. Site-directed mutagenesis of these positions demonstrated



that, while they are not strictly essential for the presence or absence of subpeptidomes, they do influence the abundance of Ala2-containing peptides. Structural data from the literature suggest that Glu45 and Tyr67 interact with Asn63, creating a more restrictive pocket that accommodates only Pro2, whereas Thr45 and Phe67 do not interact with Asn63, resulting in a more permissive pocket.



## **RESUM DE TESI**

El Complex Principal d'Histocompatibilitat (MHC), o Antigen Limfocitari Humà (HLA) en humans, codifica glicoproteïnes de membrana que habitualment es classifiquen en dues classes: les molècules HLA de classe I, que presenten pèptids derivats de proteïnes citosòliques i nuclears als limfòcits T CD8<sup>+</sup>, i les molècules HLA de classe II, que presenten pèptids derivats de proteïnes extracel·lulars als limfòcits T CD4<sup>+</sup>. Per tal que els pèptids puguin ser presentats per les molècules HLA de classe I, les proteïnes han de ser prèviament degradades i processades al citosol, una funció que porta a terme principalment el proteasoma, el component central del Sistema Ubiquitina-Proteasoma (UPS). Es tracta d'un gran complex catalític altament conservat entre espècies, que reconeix proteïnes poliubiquitinades en residus de lisina. Una part dels pèptids generats pel proteasoma són transportats al lumen del reticle endoplasmàtic (RE), on són carregats a molècules HLA de classe I i, posteriorment presentats als limfòcits T CD8<sup>+</sup>.

El nucli catalític, o proteasoma 20S, està format estructuralment per l'assemblatge de quatre anells heptamèrics: dos anells exteriors compostos per subunitats  $\alpha$  i dos anells interiors formats per subunitats  $\beta$ . L'activitat catalítica resideix a les subunitats  $\beta 1$ ,  $\beta 2$  i  $\beta 5$ , cadascuna amb una especificitat de substrat diferent. A més del proteasoma constitutiu ( $\beta 1$ ,  $\beta 2$  i  $\beta 5$ ), existeixen altres tipus de proteasomes que substitueixen les subunitats estàndard per altres anàlogues. Entre ells hi ha l'immunoproteasoma ( $\beta 1i$ ,  $\beta 2i$  i  $\beta 5i$ ), el timoproteasoma ( $\beta 1i$ ,  $\beta 2i$  i  $\beta 5t$ ) i l'espermatoproteasoma ( $\alpha 4s$ ). Més recentment s'han descrit els proteasomes intermedis, formats per una barreja de subunitats del proteasoma constitutiu i de l'immunoproteasoma: el proteasoma intermedi de tipus I (*single intermediate proteasome*, SIP) ( $\beta 1$ ,  $\beta 2$  i  $\beta 5i$ ) i el proteasoma intermedi de tipus II (*double intermediate proteasome*, DIP) ( $\beta 1i$ ,  $\beta 2$  i  $\beta 5i$ ). La presència de diferents subunitats catalítiques confereix especificitats diferenciades i, per tant, hauria d'implicar canvis en els immunoepitidomes associats a MHC-I. El fet que cada subtipus de proteasoma presenti subunitats catalítiques diferents provoca diferències en l'eficiència catalítica i, conseqüentment, en els repertoris generats. La primera part d'aquesta tesi aborda l'anàlisi de l'immunoepitidoma de classe I generat pel SIP i la caracterització de l'especificitat i de l'estructura d'aquest proteasoma 20S. Per a això, es

van generat models cel·lulars en les línies HEK-293 i HEK-293F amb expressió exclusiva del proteasoma constitutiu o del SIP. L'anàlisi de l'immunopeptidoma va posar de manifest diferències en la seva composició: en la línia que expressava el SIP hi havia un increment de pèptids amb residus aromàtics i bàsics i una disminució d'Ala i Val en la posició P $\Omega$ . Mitjançant assajos d'activitat catalítica amb pèptids fluorogènics es va observar una diferent hidròlisi de substrats hidrofòbics entre el SIP i el proteasoma constitutiu: el SIP era més eficient catalitzant enllaços després de Trp i Arg, mentre que el proteasoma constitutiu era més eficient després de residus hidrofòbics petits, com Ala. Finalment, la resolució de l'estructura 3D del SIP per microscòpia criogènica (Cryo-EM) va revelar que la subunitat  $\beta 5i$  presentava un *pocket* S1 més ampli que la subunitat  $\beta 5$ , fet que favoria l'acomodació de residus aromàtics. Aquests resultats van demostrar que la presència del SIP modifica els pèptids generats i, en conseqüència, afecta el repertori antigènic presentat per HLA-I, la qual cosa suggeria que pot modular el desenvolupament de la resposta immune.

A més, l'anàlisi dels immunopeptidomes generats pel SIP i pel proteasoma constitutiu en el model HEK-293F va mostrar l'absència de pèptids amb residus bàsics en l'extrem C-terminal. En aquesta tesi s'ha analitzat l'expressió de l'al·lel *HLA-A\*03:01* (que uneix pèptids bàsics en posició P $\Omega$ ) en la línia HEK-293F mitjançant qPCR (SYBR Green). Els nivells de mRNA d'*HLA-A\*03:01* van ser inferiors als de la línia HEK-293, però es van detectar nivells considerables. No obstant això, mitjançant citometria de flux amb anticossos específics d'al·lel es va demostrar l'absència d'expressió d'*HLA-A\*03:01* a la superfície cel·lular de la línia HEK-293F. Aquest resultat explica l'absència de pèptids amb extrem C-terminal bàsic en l'immunopeptidoma d'aquestes cèl·lules.

La segona part d'aquesta tesi es centra en dilucidar la presència de subpeptidomes en molècules d'HLA-I pertanyents al supertipus B7. L'anàlisi *in silico* dels immunopeptidomes descrits a la literatura de 13 molècules d'aquest supertipus mostra que diversos alotips presenten subpeptidomes: un amb Pro2 i residus hidrofòbics en P1, i un altre amb Ala2 i residus polars o àcids en P1 (*HLA-B\*51:01*, *-B\*51:08*, *-B\*54:01*, *HLA-B\*35:01*, *-B\*35:03*, *-B\*35:07* i *-B\*53:01*), mentre que altres no en presenten (*HLA-B\*07:02*, *-B\*07:04*, *-B\*42:01*, *-B\*55:01*, *-B\*55:02* i *-B\*56:01*). A més, les dades suggereixen que la presència de determinats residus a les posicions 45 i 67

(polimòrfiques en aquest supertipus) és rellevant en la presència d'Ala2 i potser en la formació de subpeptidomes diferents. Per analitzar el paper d'aquests residus, es van generar mutants simples i dobles a les posicions 45 i 67 de tres molècules representatives del supertipus B7: HLA-B\*51:01, HLA-B\*35:01 i HLA-B\*07:02. Les molècules amb major percentatge d'Ala2 i subpeptidomes diferents solen contenir Thr45 i Phe67, mentre que les que no presenten subpeptidomes i mostren menor percentatge d'Ala2 contenen Glu45 i Tyr67. La mutagènesi dirigida d'aquestes posicions ha mostrat que, tot i que no són essencials per a la presència o absència de subpeptidomes, sí que condicionen l'abundància de pèptids amb Ala2 al repertori peptídic. En estructures descrites a la literatura, Glu45 i Tyr67 interaccionen amb Asn63, donant lloc a un *pocket* més restrictiu que només acull Pro2, mentre que Thr45 i Phe67 no interaccionen amb Asn63, generant un *pocket* més lax.



# **INTRODUCCIÓN**





## **COMPLEJO MAYOR DE HISTOCOMPATIBILIDAD**

El Complejo Principal de Histocompatibilidad (MHC), conocido como antígeno leucocitario humano (HLA) en humanos, es una región de DNA que codifica proteínas que juegan un papel fundamental en la regulación de la respuesta inmune innata y adaptativa, en el rechazo de trasplantes y en la patogénesis de numerosas enfermedades infecciosas y autoinmunes. Las moléculas de HLA son glicoproteínas de membrana cuya función principal es la presentación de péptidos endógenos y exógenos a los linfocitos T para que tenga lugar el reconocimiento y el desarrollo de la respuesta inmune<sup>1,2</sup>. Las moléculas de HLA que presentan antígenos a los linfocitos T se dividen en dos clases principales: HLA de clase I, que juegan un papel fundamental en la defensa contra patógenos intracelulares a través de la presentación de péptidos a los linfocitos T CD8<sup>+</sup>, y HLA de clase II, encargados de la defensa contra patógenos extracelulares gracias al reconocimiento de antígenos por parte de los linfocitos T CD4<sup>+</sup> 2-4.

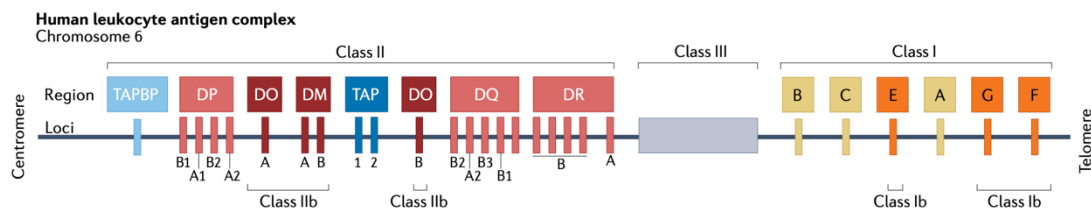
### **1. GENÉTICA Y ESTRUCTURA**

El MHC de los vertebrados se encuentra muy conservado entre las especies, sin embargo, existen variaciones en el número de genes y en las variantes alélicas de cada locus<sup>5</sup>. El sistema HLA se encuentra en una región del DNA de aproximadamente 4000 kb, localizada en el brazo corto del cromosoma 6 (6p21.3) y codifica aproximadamente 280 genes. Se organiza en tres regiones (**Figura 1**)<sup>1,3,6</sup>.

**Región de clase I:** es la región más distal o telomérica. Se encuentran codificadas las cadenas  $\alpha$  de los genes clásicos de clase I (*HLA-A*, *HLA-B* y *HLA-C*)<sup>3</sup>, junto con dos *clusters* de genes no clásicos de clase I: los genes que codifican las moléculas de *class I-like* (*HLA-E*, *HLA-F* y *HLA-G*) y los genes que codifican para ligandos del receptor NKG2D (*MIC-A* y *MIC-B*). Con la excepción de los genes *MIC*, todas las moléculas de clase I forman heterodímeros con la cadena  $\beta$ 2-microglobulina, codificada en el cromosoma 15<sup>1,6</sup>.

**Región de clase II:** se localiza en la región más centromérica y codifica ambas cadenas de los genes de las moléculas clásicas de clase II (*HLA-DP*, *HLA-DQ*, y *HLA-DR*), las moléculas no clásicas que intervienen en el reemplazamiento peptídico en los compartimentos vesiculares (*HLA-DM* y *HLA-DO*), proteínas del *Transporter Associated with Antigen Presentation* (*TAP-1* y *TAP-2*)<sup>1,2,6</sup> y genes de las subunidades de proteasoma *LMP2* (subunidad  $\beta 1i$ ) y *LMP7* (subunidad  $\beta 5i$ )<sup>7-9</sup>.

**Región de clase III:** se localiza entre las dos anteriores y codifica componentes del sistema del complemento, la hidroxilasa 21, proteínas tipo *heat shock protein* (HSP) y proteínas tipo *tumor necrosis factor* (TNF)<sup>1,6</sup>.

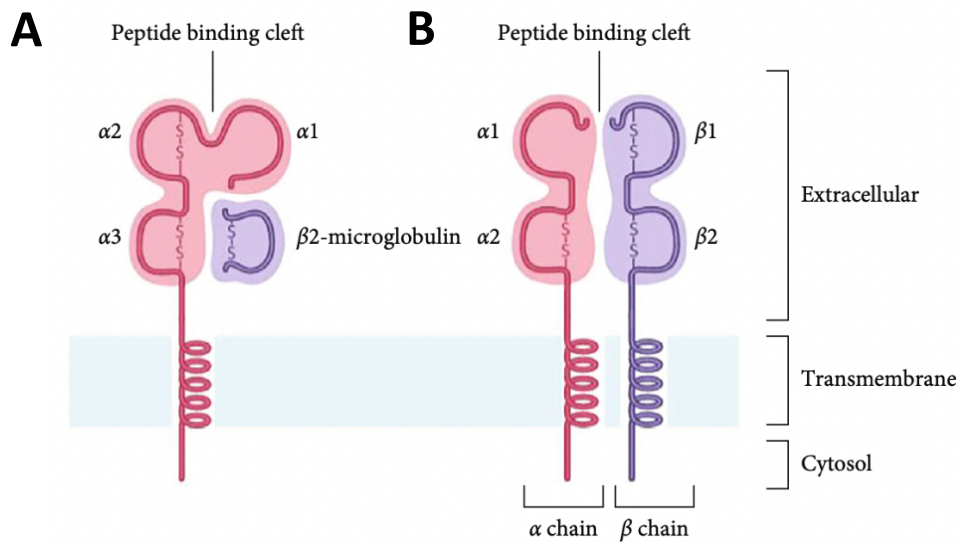


**Figura 1: Región HLA.** Se localiza en el brazo corto del cromosoma 6. La región centromérica codifica las moléculas clásicas de clase II, las moléculas no clásicas que intervienen en el intercambio peptídico, TAP y las subunidades  $\beta 1i$  y  $\beta 5i$  del proteasoma. La región telomérica codifica las cadenas  $\alpha$  de las moléculas clásicas de clase I, las moléculas no clásicas de clase I y ligandos del receptor NKG2D. La región situada entre estas dos codifica componentes del sistema del complemento, la hidroxilasa 21, proteínas tipo *heat shock protein* (HSP) y proteínas tipo *tumor necrosis factor* (TNF). Extraído de Pishesha et al. 2022, *Nature Reviews Immunology*. 751-764

Tanto los genes de clase I, como los de clase II, presentan secuencias reguladoras para garantizar la transcripción constitutiva y la regulación de la expresión de citoquinas, sin embargo, algunas de estas secuencias son compartidas por genes de HLA de la misma clase, dando lugar a una transcripción paralela. Además, los genes de HLA están regulados por trans-activadores específicos de cada clase (CITA y CIITA, respectivamente) y el factor nuclear kappa B (NF- $\kappa$ B). Estos, entre otros factores, controlan la transcripción y la función de las moléculas de HLA en cada tejido o tipo celular<sup>10</sup>.

Las moléculas de **HLA de clase I** clásicas son glicoproteínas de membrana presentes en la superficie de prácticamente todas las células nucleadas del organismo, encargadas de presentar péptidos endógenos (propios y no propios) a los linfocitos T CD8<sup>+</sup> <sup>11,12</sup>. Estructuralmente, están formadas por la unión no covalente de dos cadenas polipeptídicas: una cadena pesada  $\alpha$  (de aproximadamente 43 kDa), codificada en la región de clase I del MHC, y una cadena ligera o  $\beta$ 2-microglobulina (12 kDa) no polimórfica<sup>13</sup>. La cadena  $\alpha$  posee tres dominios extracelulares de aproximadamente 90 aminoácidos cada uno ( $\alpha$ 1,  $\alpha$ 2, y  $\alpha$ 3), una región transmembrana de 25 aminoácidos y una cola citoplasmática de 30 aminoácidos. Los dominios  $\alpha$ 1 y  $\alpha$ 2 de la cadena pesada se pliegan formando el sitio de unión al péptido<sup>14,15</sup>. Por su parte, el dominio  $\alpha$ 3 se asocia con la cadena ligera  $\beta$ 2-microglobulina, interacciona con CD8<sup>+</sup> y otorga estabilidad al complejo (**Figura 2A**)<sup>3</sup>.

Las moléculas de **HLA de clase II** se encuentran en la superficie de las células presentadoras de antígeno (APC), y de las células epiteliales tímicas, corticales (cTECs) y medulares (mTECs), y presentan péptidos a los linfocitos T CD4<sup>+</sup> <sup>2</sup>. Están formadas por dos cadenas polipeptídicas:  $\alpha$  (32-34 kDa) y  $\beta$  (29-32 kDa). Cada una de estas cadenas contiene dos dominios extracelulares ( $\alpha$ 1 y  $\alpha$ 2 y  $\beta$ 1 y  $\beta$ 2, respectivamente)<sup>16</sup>, una región transmembrana de 20-25 aminoácidos y una cola citoplasmática de 12-15 aminoácidos<sup>17,18</sup>. El surco de unión al péptido está formado por los dominios más distales,  $\alpha$ 1 y  $\beta$ 1. Los dominios proximales,  $\alpha$ 2 y  $\beta$ 2, se encuentran conservados (**Figura 2B**)<sup>3,19</sup>. El CD4<sup>+</sup> interacciona con un epítipo localizado en el dominio  $\beta$ 2.



**Figura 2: Estructura de las moléculas de HLA.** Tanto las moléculas de clase I como las de clase II son glicoproteínas de membrana que contienen una región extracelular, una región citosólica y un fragmento transmembrana. **A: HLA-I:** está formada por una cadena pesada  $\alpha$  y una cadena ligera  $\beta$ 2-microglobulina. La cadena  $\alpha$  posee tres dominios extracelulares:  $\alpha$ 1,  $\alpha$ 2, y  $\alpha$ 3. Los dominios  $\alpha$ 1 y  $\alpha$ 2 de la cadena pesada se pliegan formando el surco de unión al péptido. **B: HLA-II:** formada por dos cadenas  $\alpha$  y dos cadenas  $\beta$ . Cada una de estas cadenas contiene dos dominios:  $\alpha$ 1 y  $\alpha$ 2, y  $\beta$ 1 y  $\beta$ 2, respectivamente. El surco de unión al péptido está formado por la región distal de los dominios  $\alpha$ 1 y  $\beta$ 1. Extraído de Medhasi et al. 2022, *Journal of Immunology Research*

Los genes HLA son los más polimórficos del genoma humano<sup>3</sup>, se han identificado más de 22000 alelos<sup>20</sup>. Este polimorfismo se concentra en la región de unión al péptido, modulando el repertorio peptídico unido por moléculas de clase I y de clase II<sup>21</sup>.

La estructura del surco de unión al péptido está formada por dos dominios: de una sola cadena en las moléculas de clase I (cadena  $\alpha$ ), y de dos cadenas en las moléculas de clase II (cadena  $\alpha$  y cadena  $\beta$ ) (**Figura 2**). La estructura consiste en una lámina  $\beta$  de 8 cadenas  $\beta$  antiparalelas y 2 hélices  $\alpha$  que se encuentran suficientemente alejadas como para acomodar el péptido. La interacción entre el péptido y la molécula de MHC es mediante enlaces no covalentes, tales como interacciones hidrofóbicas, enlaces iónicos, puentes de hidrógeno, etc. En este surco se forman diferentes *pockets* donde se acomodan las cadenas laterales de determinados residuos del ligando peptídico<sup>22</sup>. Algunos *pockets* son más restrictivos, permitiendo la unión de un determinado aminoácido o un grupo de aminoácidos con características similares. En general, los *pockets* más conservados de HLA de clase I son aquellos que interaccionan con las posiciones P2 (*pocket B*) y P $\Omega$ <sup>23–26</sup>, mientras que en clase II con las posiciones P1, P4, P6, y P9<sup>24,25</sup>.

El surco de unión al péptido de las moléculas de HLA de clase I se encuentra cerrado en sus extremos por residuos de tirosina altamente conservados, dando lugar a una restricción por tamaño de 8-13 residuos<sup>27,28</sup>, aunque pueden llegar a unirse péptidos de mayor longitud, adoptando una conformación abultada en el centro<sup>11,29-31</sup>. Por otro lado, el surco de unión al péptido de las moléculas de HLA de clase II se encuentra abierto, pudiendo acomodar péptidos de mayor número de residuos, normalmente entre 13 y 18<sup>26</sup>, aunque también pueden ser mayores<sup>32</sup>. El hecho de que el surco se encuentre abierto hace que puedan unirse péptidos de diferente longitud que comparten el mismo *core* (9 aminoácidos), a estos grupos de péptidos se les conoce como familias peptídicas o *nested sets*<sup>26,32,33</sup>.

## 2. PROCESAMIENTO ANTIGÉNICO

La respuesta inmune adaptativa requiere del procesamiento y presentación antigénica. La procedencia de los péptidos presentados por las moléculas de MHC permite clasificar las dos clases principales que existen. En primer lugar, las moléculas de MHC de clase I presentan péptidos procedentes del catabolismo de proteínas citosólicas y nucleares que se cargan en el retículo endoplasmático (RE), mientras que las moléculas de clase II presentan péptidos derivados de proteínas extracelulares o de compartimentos endosomales<sup>2</sup>.

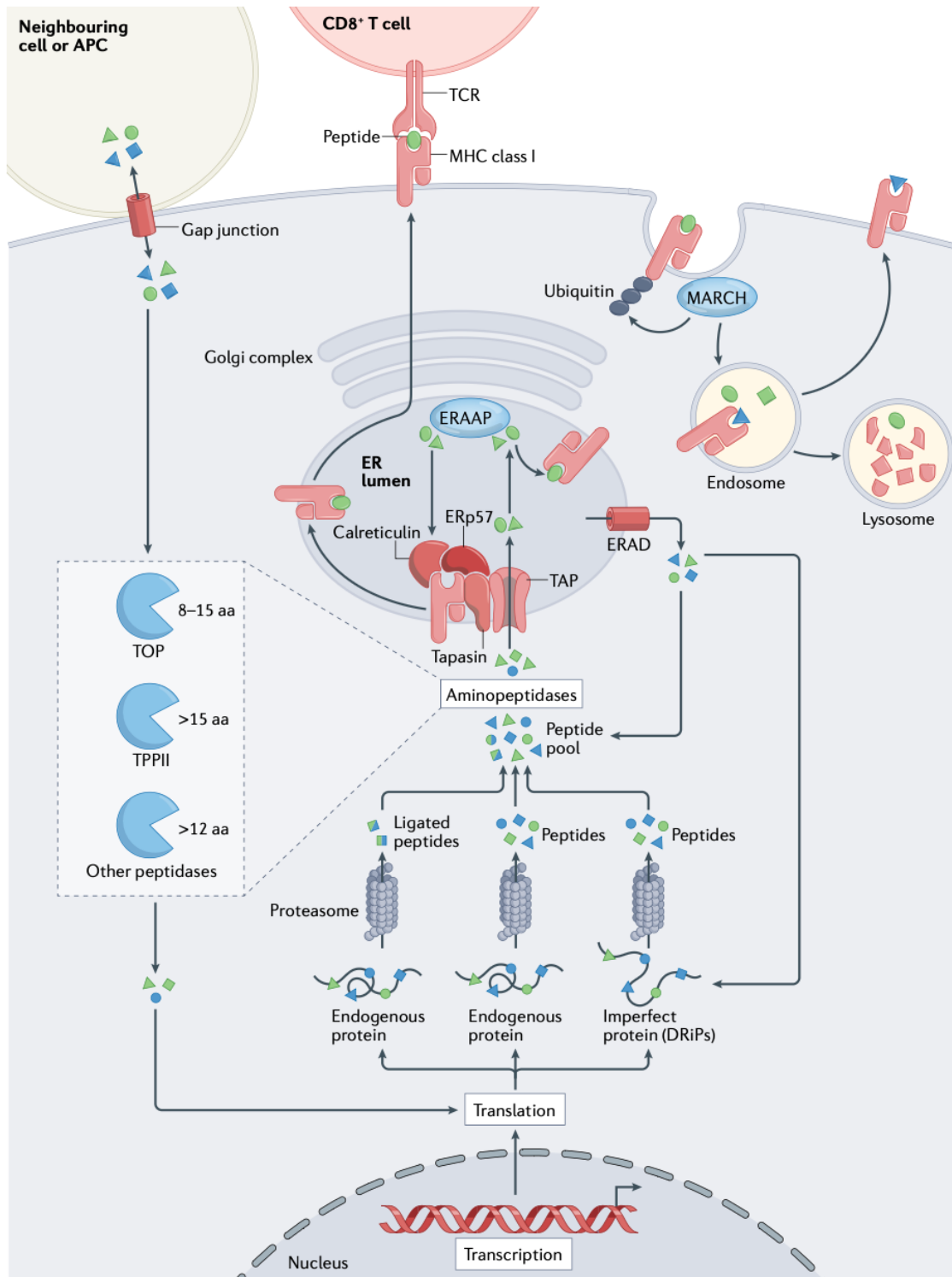
### 2.1. Ruta de procesamiento antigénico de clase I

Los ligandos de las moléculas de MHC clásicas son péptidos derivados de proteínas y versiones modificadas de estos (modificaciones postraduccionales, etc.), así como productos ribosomales defectuosos (DRiPs)<sup>34,35</sup>, mientras que las moléculas de MHC no clásicas unen otro tipo de moléculas, como lípidos<sup>36</sup>. La proteólisis es un proceso fundamental en la generación del inmunopeptidoma donde intervienen múltiples proteasas, mientras que, en el procesamiento de lípidos, glicolípidos, etc., intervienen otras enzimas<sup>36,37</sup>.

Las proteínas citosólicas se encuentran en un *turnover* constante, lo que implica su degradación constante. La mayor parte de la actividad proteolítica de la célula es llevada a cabo por el Sistema Ubiquitin-Proteasoma (UPS), aunque se ha descrito que también intervienen otras proteasas como la Tripeptidil Peptidasa (TPPII)<sup>38,39</sup>, la nardilisina<sup>40</sup>, la Oligopeptidasa Thimet (TOP)<sup>40,41</sup>, la enzima degradadora de insulina (IDE)<sup>42</sup> y la enzima convertidora de angiotensina (ACE)<sup>43</sup>, entre otras. En ausencia de patógenos, las proteínas marcadas para degradación por este sistema son proteínas propias del individuo que han sufrido errores en la traducción o plegamiento, entre otros motivos. En caso de infección, como en infecciones víricas, la síntesis de proteínas víricas contribuye al proteoma de la célula infectada y sirve como fuente de antígenos a las moléculas de MHC de clase I, asegurando que la infección se refleja en el inmunopeptidoma. Las proteínas de membrana y secretadas que no son capaces de atravesar el control del RE son enviadas al citosol en un proceso conocido como “dislocación” y también contribuyen al *pool* de antígenos presentados por las moléculas de clase I. Además, las proteínas del espacio extracelular también pueden servir como fuente de antígenos a las moléculas de MHC de clase I, en un fenómeno conocido como *cross-presentation*<sup>44</sup>.

Los productos de degradación generados por el proteasoma son péptidos de 2-24 residuos de longitud. El proteasoma genera directamente los extremos C-terminales de los ligandos de HLA-I<sup>45</sup>, mientras que el extremo amino de estos péptidos puede sufrir hidrólisis posterior por parte de proteasas presentes en el RE, como ERAAP<sup>46-49</sup>. El transporte de los péptidos se lleva a cabo mediante el transportador específico TAP (*Transporter Associated with Antigen Presentation*) un heterodímero formado por las subunidades TAP-1 y TAP-2 que transporta los péptidos mediante la hidrólisis de ATP y solo acepta los péptidos de longitud adecuada (8-16 aminoácidos) para unirse al surco de unión de las moléculas de clase I, preferentemente con residuo C-terminal hidrofóbico<sup>50,51</sup>, aunque también se han descrito péptidos independientes de TAP<sup>52</sup>. Por su parte, la molécula de MHC-I se ensambla en el RE: la cadena  $\alpha$  se sintetiza y se une a una chaperona que la mantiene parcialmente desplegada, la calnexina<sup>53</sup>, posteriormente se asocia la cadena  $\beta$ 2-microglobulina, lo que provoca la disociación de la calnexina y la unión de la calreticulina<sup>54,55</sup>. El complejo formado por la cadena  $\alpha$ , la

cadena  $\beta 2m$  y la calreticulina se une a la tapasina<sup>56</sup> y a la ERp57<sup>57,58</sup>, formando el *Peptide-Loading Complex* (PLC)<sup>53–58</sup>. La tapasina cataliza el intercambio de péptidos de baja afinidad por péptidos de mayor afinidad hasta que un péptido de alta afinidad es capaz de estabilizar la molécula de MHC-I<sup>59,60</sup>. Finalmente, el complejo p-MHC-I se transporta a la membrana celular donde presentará el péptido a los linfocitos T CD8<sup>+</sup> (Figura 3)<sup>11</sup>.



**Figura 3: Ruta de procesamiento antigénico de clase I.** Las proteínas endógenas y los productos de traducción defectuosa (DRiPs) son degradados principalmente por el proteasoma, generando péptidos de distinta longitud (2-24 aa). Estos péptidos pueden ser procesados adicionalmente por aminopeptidasas citosólicas. Posteriormente, los péptidos son transportados al retículo endoplásmico (ER) por el transportador asociado a procesamiento antigénico (TAP). Tras la formación del complejo de carga, PLC, la tapasina cataliza el intercambio de péptidos de baja afinidad por péptidos de alta afinidad. El complejo HLA-I-péptido se transporta a la superficie celular a través del aparato de Golgi para su reconocimiento por linfocitos T CD8<sup>+</sup>. Extraído de *Pishesha et al. 2022, Nature Reviews Immunology. 751-764*

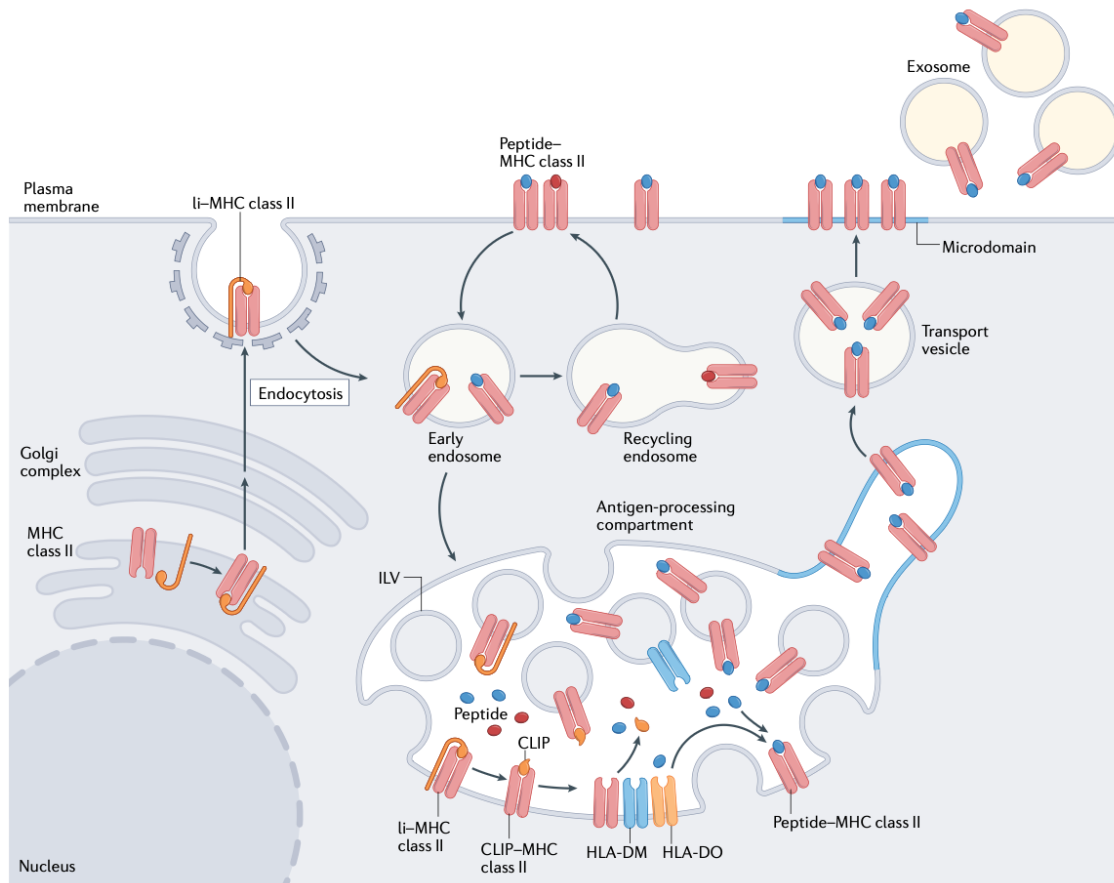


## 2.2. Ruta de procesamiento antigénico de clase II

Las células que expresan MHC de clase II adquieren antígenos a través de procesos en los que intervienen receptores de diversa afinidad, como la fagocitosis y la endocitosis. La ruta endocítica consistente en una serie de compartimentos interconectados, los cuales sirven de fuente de ligandos para las moléculas de MHC de clase II<sup>61</sup>. En condiciones de homeostasis, las proteínas extracelulares también se encuentran en *turnover* constante, y son internalizadas y enviadas a los compartimentos endosomales, por lo que, en ausencia de infección por un patógeno, los ligandos de MHC de clase II son las proteínas del propio individuo. Además, en la célula existen fenómenos, como la autofagia, que permiten que proteínas citosólicas sean procesadas y presentadas por moléculas de MHC de clase II<sup>44</sup>. Los compartimentos de la ruta endocítica son ricos en proteasas y presentan acidificación del medio. La bajada de pH del medio desestabiliza las proteínas, desplegándolas, facilitando así la proteólisis por parte de las proteasas endosomales, entre las que se encuentra la familia de las catepsinas<sup>62</sup>.

Las cadenas  $\alpha$  y  $\beta$  de las moléculas de MHC de clase II también se sintetizan y ensamblan en el RE, donde se asocian a la cadena invariante asociada al MHC de clase II (Ii) o CD74, una glicoproteína de membrana que actúa como pseudopéptido<sup>63</sup>. Tres complejos triméricos formados por ambas cadenas de MHC-II y Ii (trímero de trímeros) alcanzan los compartimentos endocíticos gracias a motivos de dileucinas localizados en la cadena invariante que aseguran la entrada de las moléculas de MHC-II en los compartimentos endocíticos<sup>16</sup>. La unión de las moléculas de MHC de clase II con compartimentos endosomales, conocidos como MIIC, garantiza que se expongan a los posibles péptidos ligandos<sup>64</sup>. En los compartimentos ácidos, diversas proteasas residentes, especialmente catepsinas L y S, hidrolizan específicamente Ii, excepto un fragmento denominado CLIP (*Class II Associated Ii Peptide*)<sup>65</sup>, que se mantiene asociado a la molécula de MHC-II y previene la unión de péptidos. HLA-DM actúa como molécula editora de péptidos<sup>66</sup>, permitiendo la sustitución de CLIP por un péptido de alta afinidad y estabiliza MHC-II. HLA-DM continúa intercambiando ligandos hasta que no es capaz de desplazar uno de ellos, dando como resultado un complejo MHC-II-péptido muy estable. El complejo p-MHC-II es transportado a la membrana plasmática por vía vesicular, donde interaccionará con linfocitos T CD4<sup>+</sup>. HLA-DO es otra molécula no polimórfica expresada

en células B y células epiteliales, aunque su función no está claramente definida<sup>67,68</sup> (Figura 4).



**Figura 4: Ruta de procesamiento antigénico de clase II.** Las moléculas de HLA-II se sintetizan en el retículo endoplásmico asociadas a la cadena invariante (Ii), que bloquea el surco de unión y dirige el complejo hacia compartimentos endosómicos especializados. Ii es degradada hasta dejar un fragmento denominado CLIP (*Class II-associated invariant chain peptide*), que permanece ocupando el surco de unión. La proteína HLA-DM facilita la liberación de CLIP y la carga de péptidos derivados de proteínas endocitadas y procesadas en los endosomas, mientras que HLA-DO modula la actividad de HLA-DM. El resultado es la formación de complejos HLA-II-péptido estables que son transportados a la superficie celular a través de vesículas, donde pueden ser reconocidos por linfocitos T CD4<sup>+</sup>. Adicionalmente, parte de estos complejos puede ser liberada en exosomas. Extraído de Pishesha et al. 2022, *Nature Reviews Immunology*. 751-764

### **3. HLA Y ENFERMEDADES**

Las moléculas de HLA se encargan de presentar péptidos, tanto propios como extraños, a los linfocitos T, lo que determina la activación de la respuesta inmune o la tolerancia. Este hecho hace que estas moléculas jueguen un papel central en la susceptibilidad y desarrollo diferentes enfermedades.

Las moléculas de HLA son los principales determinantes genéticos asociados al desarrollo de enfermedades autoinmunes. Esta asociación no se limita al desarrollo inicial, sino que también modula la gravedad, la evolución y la respuesta frente diversos tratamientos. Ejemplos de estas asociaciones son el alelo de clase I HLA-B\*27<sup>69</sup> y la espondilitis anquilosante y HLA-DR3 y la diabetes tipo I<sup>70</sup>.

En cáncer, el papel de las moléculas de HLA es doble. Por un lado, las moléculas de HLA-I son capaces de presentar péptidos derivados de proteínas mutadas o neoantígenos a los linfocitos T CD8<sup>+</sup>, activando la respuesta inmune. Por otro lado, su alteración constituye uno de los principales mecanismos de escape tumoral. La reducción de la expresión de moléculas de clase I o las modificaciones epigenéticas que afectan su regulación, son fenómenos frecuentes en numerosos tipos de cáncer y se asocian con un peor pronóstico y resistencia a la inmunoterapia<sup>71</sup>. Históricamente, las inmunoterapias contra cáncer se han basado en utilizar productos biológicos que activan células T contra tumores<sup>72</sup>. El análisis por espectrometría de masas de péptidos tumorales se ha utilizado para identificar péptidos de neoantígenos a partir de células o tejidos tumorales<sup>73</sup>. Los virus oncolíticos son capaces de inducir a las células tumorales a producir nuevos ligandos de moléculas del HLA de clase I, activando respuestas de linfocitos T CD8<sup>+</sup> <sup>74,75</sup>.

Asimismo, las moléculas de clase I tienen un rol mayoritario en la susceptibilidad a enfermedades infecciosas. Diversos estudios respaldan la asociación entre los alelos de HLA de clase I y de clase II con enfermedades infecciosas, mayoritariamente VIH, Hepatitis B y C. En la infección por VIH existen diferentes factores genéticos, incluyendo las moléculas de HLA, que se asocian con la susceptibilidad a sufrir la infección, la transmisión vertical madre-hijo, la transmisión horizontal, la ratio de progresión a SIDA y el riesgo de reacciones adversas hacia diferentes tratamientos<sup>76</sup>.

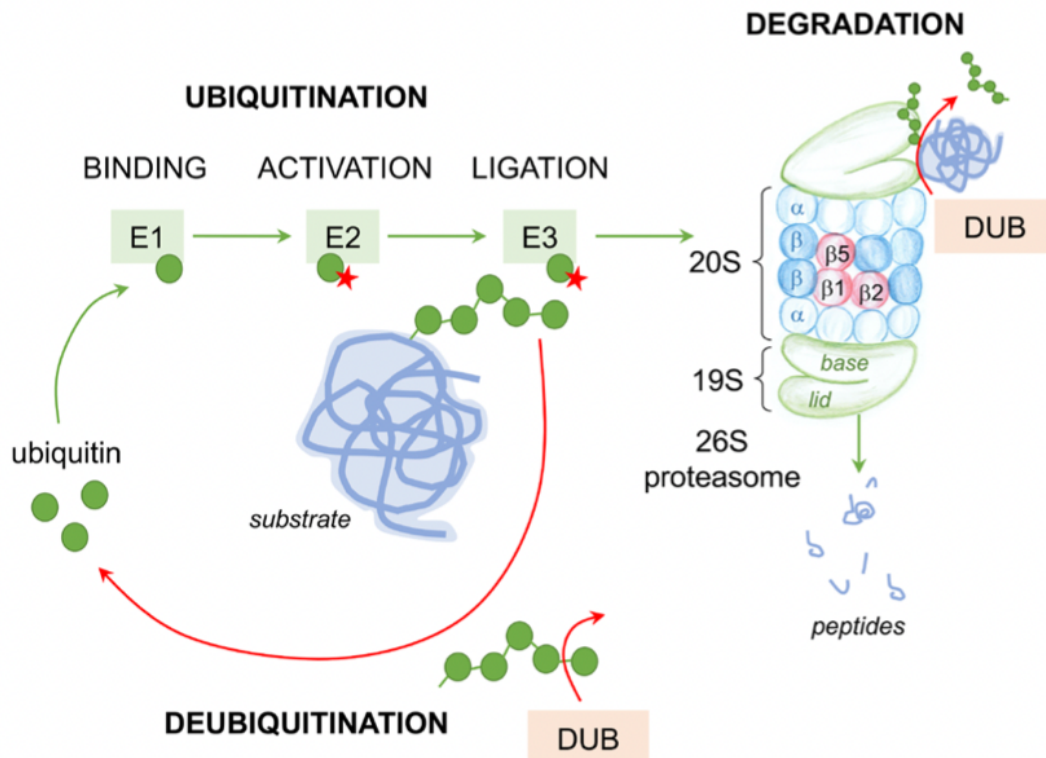
## **PROTEASOMA**

### **1. SISTEMA UBIQUITÍN-PROTEASOMA**

El Sistema Ubiquitín-Proteasoma (UPS) es el encargado de la degradación del 80-90% de las proteínas intracelulares<sup>77-79</sup>. El UPS es esencial en el mantenimiento de la homeostasis proteica, ya que regula numerosos procesos como la diferenciación celular, la apoptosis y la transducción de señales<sup>80</sup>. Además, como ya se ha indicado, es fundamental en la generación de péptidos antigénicos presentados por MHC-I<sup>81</sup>. El proteasoma 26S (de un tamaño de aproximadamente 2,5MDa) es el componente principal de este sistema. Es un gran complejo degradativo, multimérico y altamente conservado entre las especies, cuya función es degradar proteínas poliubiquitinadas en residuos de lisina<sup>82-85</sup>. Por su parte, la ubiquitina (9kDa) funciona como marca de degradación. La ubiquitinación consiste en la unión covalente entre la ubiquitina a uno o más residuos de lisina del sustrato por acción de las enzimas E1, E2 y E3. Brevemente, la enzima E1 activa la ubiquitina a través de una reacción dependiente de ATP que consiste en la formación de un enlace tioéster entre un residuo de glicina (Gly76) del extremo C-terminal de la ubiquitina y la cisteína del sitio activo de la enzima E1. Después, las enzimas E2 se encargan de transferir las ubiquitinas activadas al sustrato. Las enzimas E3 funcionan como nexo entre el sustrato y la enzima E2 con la ubiquitina activada y da la especificidad al proceso, ya que la ubiquitina transferida a la E2 se une por el extremo C-terminal a un grupo  $\epsilon$ -amino de la proteína diana gracias a las enzimas E3 (**Figura 5**)<sup>86</sup>.

Existen diferentes patrones, como la conjugación de monómeros de ubiquitina (monoubiquitinación) o di-, oligo- y poli-ubiquitinación. El patrón de ubiquitinación determina el destino de las proteínas. Además, dependiendo de los residuos de lisina donde se unen las cadenas de ubiquitina se distinguen los patrones K6-, K11-, K27-, K29-, K33-, K48-, y K63. Las enzimas deubiquitinas (DUBs) son capaces de eliminar las cadenas de ubiquitina de las proteínas a las que están unidas. La marca de degradación proteica por parte del proteasoma 26S es la poliubiquitinación en la K63<sup>87,88</sup>.

Después de que las proteínas ubiquitinadas se unan al proteasoma, este desensambla y libera las cadenas de ubiquitina, introduce el sustrato desubiquitinado en la cámara proteolítica y lo degrada. La actividad de las enzimas DUBs es importante en el reciclaje de la ubiquitina para realizar reacciones de ubiquitinación posteriores, para prevenir que el proteasoma colapse y para controlar el *turnover* de las proteínas<sup>78,89</sup>.

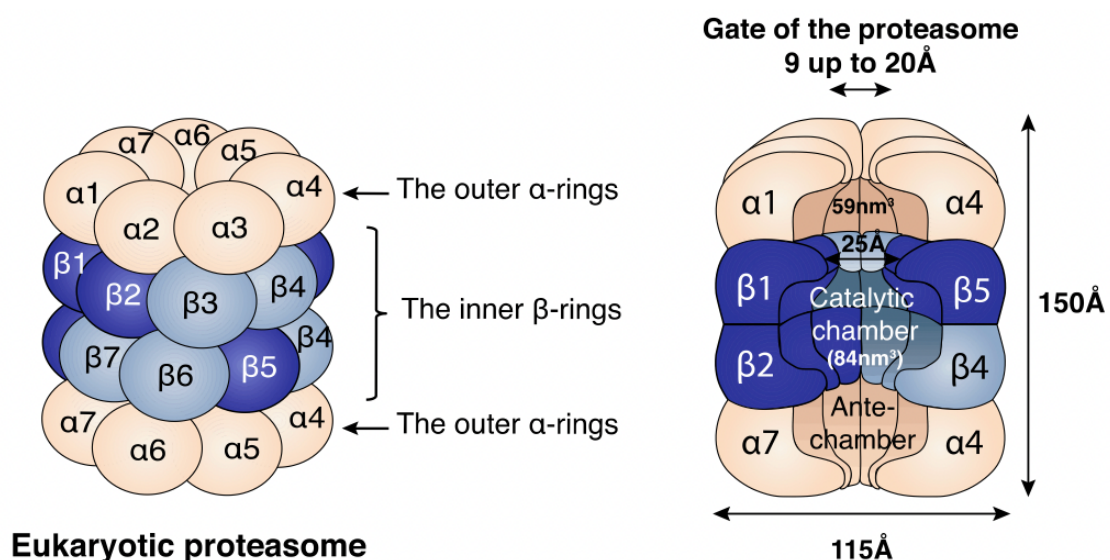


**Figura 5: Esquema de ubiquitinación y degradación por parte del Sistema Ubiquitín-Proteasoma.** La enzima E1 activa la ubiquitina y las enzimas E2 las transfieren al sustrato. Las enzimas E3 actúan como nexos. Posteriormente las proteínas marcadas para degradación ingresan en la cámara proteolítica del proteasoma a través del reconocimiento de la marca de ubiquitina por la partícula 19S y son degradadas en péptidos por este complejo. Las enzimas DUB se encargan de reciclar las moléculas de ubiquitina. Extraído de Heintz et al. 2021, *Cell Physiol Biochem*.

## 2. ESTRUCTURA DEL PROTEASOMA

El proteasoma 20S o partícula *core* (CP) (700 kDa) es un complejo catalítico muy conservado a lo largo de la evolución y está presente en los dominios *Archea*, *Bacteria* y *Eucarya*<sup>90</sup>. Está formado por el ensamblaje de 28 proteínas dispuestas en 4 anillos heptaméricos en forma de barril. Los anillos exteriores están compuestos por 7 subunidades  $\alpha$  y forman un estrecho canal de entrada para los diferentes sustratos, ya que estas subunidades presentan extensiones N-terminales que controlan la entrada de estos. Por su parte, los anillos interiores están compuestos por 7 subunidades  $\beta$  y crean la cavidad donde tiene lugar la proteólisis<sup>91–93</sup>.

El proteasoma 20S de bacterias y arqueas presenta un solo tipo de subunidades  $\alpha$  y  $\beta$ , por lo que presenta 14 copias idénticas de subunidades  $\alpha$  y  $\beta$  y, por lo tanto, 14 subunidades catalíticamente activas en la cámara central, correspondientes a los dos anillos interiores de subunidades  $\beta$ <sup>94</sup>. En eucariotas, por el contrario, presenta 7 subunidades  $\alpha$  y 7 subunidades  $\beta$  diferentes, sin embargo, solo 3 subunidades  $\beta$  son catalíticamente activas:  $\beta 1$  (PSMB6, Y),  $\beta 2$  (PSMB7, X) y  $\beta 5$  (PSMB5, Y), dando lugar a 6 sitios catalíticamente activos (Figura 6)<sup>95</sup>.



**Figura 6: Estructura del proteasoma eucariota 20S.** El proteasoma 20S está formado por el ensamblaje de 28 subunidades dispuestas en 4 anillos heptaméricos. Los anillos exteriores están compuestos de 7 subunidades  $\alpha$  y los anillos interiores de 7 subunidades  $\beta$ , delimitando la cámara central donde tiene lugar la proteólisis por parte de las subunidades  $\beta 1$ ,  $\beta 2$  y  $\beta 5$ . Extraído de *Habib J et al. 2022, Cells*.

A pesar de que el mecanismo de acción de las tres subunidades  $\beta$  es idéntico, la especificidad de sustrato varía. Así, se ha descrito que la subunidad  $\beta 1$  cataliza el enlace tras residuos ácidos, la subunidad  $\beta 2$  tras residuos básicos y la subunidad  $\beta 5$  tras residuos hidrofóbicos, dando lugar a actividades tipo caspasa, tripsina y quimotripsina, respectivamente<sup>96</sup>. Esta especificidad se da principalmente por las propiedades del *binding pocket* de cada subunidad, concretamente de la posición 45. El carácter básico del *pocket* S1 de la subunidad  $\beta 1$  es concebido por la Arg45, que favorece el corte tras residuos ácidos en posición *P1*. En la subunidad  $\beta 2$ , la posición 45 está ocupada por Gly, mientras que en la subunidad  $\beta 5$  por Met, favoreciendo el corte tras residuos básicos e hidrofóbicos, respectivamente, aunque los residuos que se encuentran flanqueando también juegan un papel fundamental en la unión y catalización del sustrato<sup>97,98</sup>.

El proteasoma es una treonin-proteasa, el grupo hidroxilo de la Thr del extremo N-terminal de la subunidad  $\beta$  madura lleva a cabo un ataque nucleofílico al carbono carbonilo del enlace peptídico, dando lugar a un intermediario, donde el extremo N-terminal continúa unido al proteasoma por un enlace éster, después, una molécula de agua que se encuentra en la cámara catalítica hidroliza este intermediario, liberando el péptido<sup>93,98-100</sup>. Para evitar la catalización indiscriminada, las subunidades  $\beta 1$ ,  $\beta 2$  y  $\beta 5$  (y la subunidad  $\beta 6$ ) son sintetizadas como zimógenos con una extensión peptídica en el extremo N-terminal. Tras el corte del propéptido, la Thr se queda expuesta y, por lo tanto, la subunidad activa. La maduración de las subunidades  $\beta$  tiene lugar durante el ensamblaje del proteasoma<sup>101</sup>.



Como ya se ha mencionado, la generación de los péptidos de unión a moléculas de HLA-I tiene lugar por la proteólisis de proteínas presentes en el citosol por parte del proteasoma. Hasta hace dos décadas, se pensaba que los péptidos ligados de HLA-I se producían por la mera catalización de proteínas citosólicas por el UPS, sin embargo, en 2004 Hanada et al,<sup>102</sup> observaron un clon de CTL que reconocía un péptido de carcinoma renal compuesto por dos fragmentos de un péptido que originalmente se encontraban distantes. La existencia de este péptido híbrido requería la escisión de una secuencia de 40 aminoácidos y el *splicing* de ambos fragmentos. Esto, junto con experimentos posteriores, confirmaron que el proteasoma realizaba *splicing* alternativo<sup>103</sup> (Figura 7).

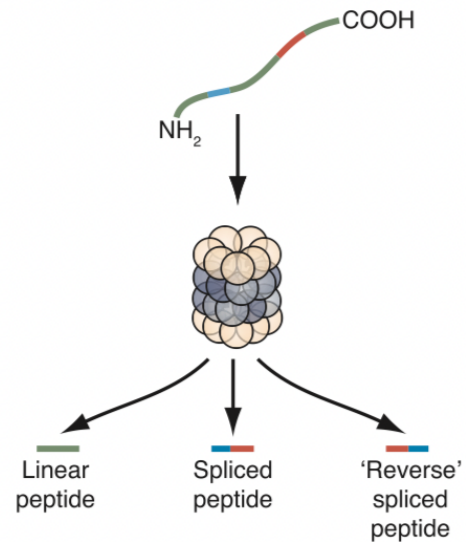


Figura 7: Esquema *splicing* alternativo por parte del proteasoma. Extraído de Vigneron N et al. 2004, *Science*. 587-590

El proteasoma 20S puede asociarse a diversas partículas reguladoras gracias a la unión de estas con las subunidades  $\alpha$ , las más comunes son: PA700 (19S), PA28 (11S), PA26 y PA200<sup>104,105</sup>. Estas pueden unirse a uno o ambos lados del proteasoma 20S, incluso dar lugar a proteasomas híbridos<sup>90</sup>, con dos partículas reguladoras diferentes a cada extremo del 20S y regulan el reconocimiento y acceso del sustrato al *core* catalítico 20S<sup>80</sup>. En eucariotas, la configuración más común de proteasoma es el 26S, formado por la unión del 20S y una o dos partículas reguladoras 19S<sup>78,82</sup>. La partícula reguladora 19S se trata de un complejo formado por una base y una tapa. La base contiene 6 subunidades ATP-asa diferentes (Rpt1 - Rpt6) y 4 subunidades no ATP-asa (Rpn1, Rpn2, Rpn10 y Rpn13). Las subunidades con actividad ATP-asa de la base son responsables de la hidrólisis de ATP que hace que los sustratos se desplieguen y accedan a la cavidad central. Rpn10 es la subunidad encargada de reconocer la marca de ubiquitina. La tapa está formada por 9 subunidades no ATP-asa, entre ellas Rpn11, que corta las cadenas de poliubiquitina del sustrato antes de su degradación. Normalmente, los anillos de

subunidades  $\alpha$  del *core* se encuentran cerrados y, durante muchos años, se ha pensado que su apertura requería la partícula reguladora 19S, sin embargo, existe evidencia de que el proteasoma 20S es capaz de degradar sustratos de forma independiente<sup>106,107</sup>

### 3. ENSAMBLAJE DEL PROTEASOMA 20S

El ensamblaje del proteasoma es un proceso muy sofisticado y altamente regulado que requiere la acción de 4 proteínas chaperonas conocidas como *Proteasome Assembly Chaperones* (PAC1, PAC2, PAC3, PAC4) y la *Proteasome Maturation Protein* (POMP)<sup>108</sup>. Estas proteínas contribuyen al correcto ensamblaje o plegamiento, pero se encuentran ausentes en la forma madura final. En primer lugar, las chaperonas PAC1-PAC2 y PAC3-PAC4 actúan como heterodímeros en la formación del anillo  $\alpha$ . Los últimos avances en este campo indican que las primeras subunidades en ensamblarse son  $\alpha_4$ ,  $\alpha_5$ ,  $\alpha_6$  y  $\alpha_7$  por acción de PAC3-PAC4, después, PAC1-PAC2 ensamblan las subunidades  $\alpha_1$ ,  $\alpha_2$  y  $\alpha_3$ , dando lugar a la formación de un anillo  $\alpha$  maduro, que servirá como molde a las subunidades  $\beta$  gracias a la acción de POMP<sup>109</sup>.

La formación del anillo de subunidades  $\beta$  comienza con el ensamblaje de la subunidad pro- $\beta_2$ , seguida de  $\beta_3$  y  $\beta_4$ . Cuando se ensamblan estas subunidades, el heterodímero PAC3-PAC4 se libera, y se unen las subunidades pro- $\beta_5$ , pro- $\beta_6$  y pro- $\beta_1$ . Recientemente se ha sugerido que los pro-péptidos de las subunidades catalíticas no existen exclusivamente para evitar la activación temprana de estas subunidades, sino que juegan un papel fundamental en el ensamblaje del resto de subunidades  $\beta$ . Los pro-péptidos de las subunidades  $\beta_2$  y  $\beta_5$  son esenciales para la incorporación de las subunidades  $\beta_3$  y  $\beta_6$ , respectivamente, mientras que el pro-péptido de  $\beta_5$  es además fundamental para la interacción con POMP<sup>110</sup>. Por último, se ensambla la subunidad  $\beta_7$ , formando medio proteasoma 20S, que, tras dimerizar, forma el proteasoma 20S maduro<sup>111</sup>. La formación del proteasoma 20S maduro requiere la liberación de las chaperonas PAC1-PAC2 y POMP, sin embargo, no está del todo claro si esta liberación es llevada a cabo por el proteasoma en sí mismo o si las chaperonas son liberadas intactas y después recicladas<sup>110</sup>.

#### 4. SUBTIPOS DE PROTEASOMA

Existen diversas formas de proteasoma en una misma célula<sup>112</sup>, y sus ratios cambian considerablemente entre tejidos y bajo distintos estímulos. El proteasoma estándar o constitutivo, está formado por las ya mencionadas subunidades catalíticas  $\beta 1$ ,  $\beta 2$  y  $\beta 5$  y se trata de un complejo fundamental para la célula. Sin embargo, bajo condiciones de estrés, o inducido por citoquinas proinflamatorias como IFN- $\gamma$  y TNF- $\alpha$ , se expresa el inmunoproteasoma, aunque su expresión es constitutiva en diversas células inmunes, como linfocitos B, T y células presentadoras de antígeno<sup>113</sup>. La característica principal de este subtipo de proteasoma es que las subunidades catalíticas  $\beta 1$ ,  $\beta 2$  y  $\beta 5$  son sustituidas por las subunidades catalíticas inmunes  $\beta 1i$  (PSMB9, LMP2),  $\beta 2i$  (PDMB10, MECL1) y  $\beta 5i$  (PSMB8, LMP7), respectivamente (**Figura 8**). Esta sustitución provoca un aumento de las actividades tripsina y quimotripsina, mientras que la actividad caspasa se ve disminuida<sup>97,114</sup>. Este cambio en el patrón de actividad catalítica provoca que el inmunoproteasoma sea más eficiente que el proteasoma constitutivo en la generación de péptidos de unión a MHC-I<sup>113,115</sup>. El inmunoproteasoma está involucrado en la eliminación de células tumorales y células infectadas por virus por parte de los linfocitos T CD8<sup>+</sup> <sup>110,112,116,117</sup>.

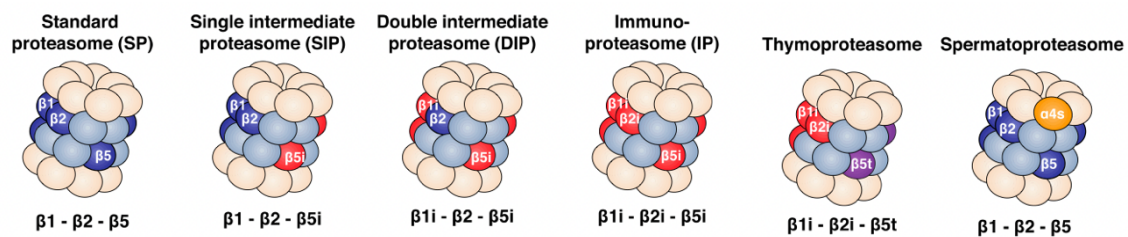
Durante años se pensó que las tres subunidades catalíticas presentes en el proteasoma debían ser o bien constitutivas, o bien inmunes, sin embargo, Dalhmann et al, describió en el año 2000 dos subtipos de proteasomas adicionales, formados por subunidades catalíticas constitutivas e inmunes, que representan un porcentaje variable de los proteasomas presentes en colon, intestino delgado, hígado y riñones<sup>118</sup>. Los dos subtipos adicionales son: el proteasoma intermedio de tipo I o *single intermediate proteasome* (SIP) que presenta las subunidades catalíticas  $\beta 1$ ,  $\beta 2$  y  $\beta 5i$  y el proteasoma intermedio de tipo II o *double intermediate proteasome* (DIP), que presenta las subunidades inmunes  $\beta 1i$ ,  $\beta 2$  y  $\beta 5i$  (**Figura 8**)<sup>110,119</sup>. Ambos proteasomas presentan actividades tripsina y quimotripsina intermedia entre el proteasoma constitutivo y el inmunoproteasoma, sin embargo, la actividad caspasa del proteasoma constitutivo y SIP es similar entre ellos, mientras que en DIP es menor. Este fenómeno se debe a que la

actividad caspasa está asociada a la subunidad  $\beta 1$ , presente en el proteasoma constitutivo y SIP, pero no en DIP<sup>110</sup>.

Al contrario que el proteasoma constitutivo, donde la subunidad  $\beta 1$  es la penúltima en incorporarse, el ensamblaje del inmunoproteasoma se ha definido con experimentos de *KO* de las diversas subunidades inmunes, y ponen de manifiesto que las subunidades  $\beta 1i$  y  $\beta 2i$  son las primeras en ensamblarse al anillo de subunidades  $\alpha$  y que, concretamente, la subunidad  $\beta 2i$  depende de la subunidad  $\beta 1i$ <sup>108,120,121</sup>. Asimismo, la subunidad  $\beta 5i$  puede incorporarse de forma independiente de  $\beta 4$ <sup>110,120</sup>, mientras que la subunidad  $\beta 5$  constitutiva requiere de  $\beta 4$  para su incorporación al anillo  $\beta$ . Además, la incorporación de  $\beta 5i$  es necesaria para el corte de los pro-péptidos de  $\beta 1i$  y  $\beta 2i$  y presenta mayor afinidad por POMP que  $\beta 5$ <sup>120,122</sup>. Todas estas características del ensamblaje del inmunoproteasoma hacen que este proceso sea 4 veces más rápido que el ensamblaje del proteasoma constitutivo. La existencia de proteasomas intermedios es consistente con las leyes de ensamblaje de este complejo. Se cree que el ensamblaje de estos proteasomas es dependiente de la subunidad  $\beta 5i$ , ya que se puede unir al *core* de forma independiente al resto de subunidades inmunes y es necesaria para procesar el pro-péptido de  $\beta 1i$  y  $\beta 2i$ . Sin embargo, estudios de ratones *KO* exponen que no se debe descartar la existencia de otros proteasomas intermedios<sup>123</sup>.

Adicionalmente, existen proteasomas específicos de tejido. Las células epiteliales tímicas corticales (cTECs) presentan un subtipo de proteasoma exclusivo, conocido como timoproteasoma<sup>117,124</sup>, que participa en la selección positiva de los timocitos en el timo, aunque su papel en este proceso no está del todo definido. Este subtipo presenta las subunidades catalíticas inmunes  $\beta 1i$  y  $\beta 2i$ , junto con  $\beta 5t$  (PSMB11, subunidad homóloga a  $\beta 5$  y  $\beta 5i$ )<sup>124</sup> (**Figura 8**). La expresión de esta subunidad está mediada por el factor de transcripción Foxn1<sup>125,126</sup> y su incorporación en lugar de  $\beta 5$  o  $\beta 5i$  reduce la actividad quimotríptica, ya que el *pocket* S1 de esta subunidad presenta más residuos hidrofílicos, reduciendo su capacidad de interaccionar con residuos hidrofóbicos<sup>97,117</sup>. El ensamblaje de este complejo sigue el mismo patrón que el inmunoproteasoma, diferenciando la subunidad  $\beta 5i$  por  $\beta 5t$ <sup>127</sup>.

El espermato proteasoma es un subtipo de proteasoma que, al contrario que los anteriores, presenta una subunidad  $\alpha$  diferente, la subunidad  $\alpha 4s$ , y se expresa específicamente en los testículos de mamíferos (**Figura 8**)<sup>128,129</sup>. Esta sustitución no afecta a su actividad, sin embargo, es necesaria para el correcto ensamblaje del proteasoma en los testículos de mamíferos, para la degradación de histonas dependiente de acetilación y para la correcta progresión de la meiosis en las células germinales masculinas<sup>130,131</sup>.



**Figura 8: Subtipos de proteasoma.** Extraído de *Habib J et al. 2022, Cells*.

La existencia de diversos subtipos de proteasoma constituye un mecanismo fundamental en la modulación de la diversidad de péptidos presentados por moléculas de HLA-I, lo que contribuye a la plasticidad del inmunopeptidoma en diferentes contextos fisiológicos. En el inmunoproteasoma y en los proteasomas intermedios, la subunidad  $\beta 5$  es reemplazada por la subunidad  $\beta 5i$ . Estudios recientes han demostrado que determinados epítomos tumorales son generados exclusivamente por proteasomas intermedios, evidenciando que las diferencias entre la subunidad  $\beta 5$  y la subunidad  $\beta 5i$  modifican el patrón de catalización del sustrato<sup>132</sup>. Por lo tanto, el análisis de estas diferencias funcionales y estructurales es esencial, ya que los proteasomas intermedios representan entre un 30-50% del proteasoma total de tejidos no linfoides como colon, intestino delgado, hígado y riñones<sup>133</sup>. Ignorar la contribución de estos proteasomas al inmunopeptidoma puede llevar a infravalorar el repertorio real de péptidos presentados que, a su vez, puede comprometer la identificación de epítomos relevantes para el desarrollo de terapias contra cáncer, enfermedades autoinmunes y vacunas.

En esta tesis, realizamos una caracterización exhaustiva del proteasoma intermedio de tipo I, incluyendo el análisis del inmunopeptidoma presentado por HLA-I de células que expresan el SIP en comparación con células que expresan el proteasoma constitutivo, la caracterización su actividad catalítica y la determinación de la estructura tridimensional de este complejo, con el fin de contribuir a la predicción de epítomos presentados por HLA-I en el futuro.

## **SUPERTIPOS Y SUBPEPTIDOMAS DE HLA**

El inmunopeptidoma es el repertorio de péptidos que se unen y presentan por moléculas de MHC de clase I, de clase II y no clásicas<sup>134–136</sup>. Cada alotipo de HLA presenta decenas de miles de péptidos, lo que teniendo en cuenta la gran variedad de moléculas de HLA que presenta un individuo, da lugar a inmunopeptidomas compuestos por millones de péptidos. El inmunopeptidoma es una muestra del proteoma celular, por lo que aporta información sobre el estado de esta<sup>137</sup>. Los p-MHC expuestos en la superficie de las células tisulares y APCs son examinados constantemente por los linfocitos T circulantes. Cuando un linfocito T detecta un péptido no propio o anormal, desencadena la respuesta inmune<sup>20</sup>.

La región de HLA es la región del genoma humano asociada con un mayor número de enfermedades y sus polimorfismos contribuyen a que exista variabilidad genética y diferencias en la susceptibilidad a enfermedades entre individuos. Es importante resaltar que diferencias mínimas, incluso modificaciones postraduccionales de un solo aminoácido, pueden tener un gran impacto en el reconocimiento por parte de células T y en el desarrollo de una enfermedad<sup>138–140</sup>. Uno de los problemas principales que se plantean a la hora de desarrollar vacunas contra diferentes patógenos es precisamente esta variabilidad. Por ello, se planteó diseñar péptidos capaces de ser reconocidos por varias moléculas de HLA. Este concepto se conoce como supermotivos de HLA, cada uno de los cuales corresponde a ligandos de HLA capaces de unirse a varios alelos. Las variantes alélicas de HLA que unen a péptidos que poseen un supermotivo de HLA concreto se denominan un supertipo de HLA. Este concepto se introdujo por primera vez por el grupo de Alessandro Sette. Así, un supertipo de HLA lo constituyen alelos de HLA con características de unión al péptido muy similares, por lo que las moléculas de HLA de un mismo supertipo presentan repertorios peptídicos parcialmente solapantes. Las moléculas de HLA-I se pueden clasificar en 10 supertipos: HLA-A01, -A02, -A03, -A24, -B07, -B08, -B27, -B44, -B58 y -B62<sup>141–143</sup>. A pesar de unir repertorios con motivos de anclaje similares, el hecho de pertenecer a un determinado supertipo no garantiza la presentación de un péptido en concreto por parte de los diferentes alotipos, ya que las diferencias en los inmunopeptidomas, es decir, los péptidos que cada alelo puede

presentar, pueden variar incluso entre alelos estrechamente relacionados dentro de un mismo supertipo<sup>144</sup>. Estas diferencias pueden explicar la presentación, o la no presentación, de epítomos inmunológicamente relevantes, ya sean derivados de patógenos, tumores o autoantígenos. En este contexto, polimorfismos puntuales en el surco de unión al péptido pueden alterar de manera directa la presentación antigénica, generando variaciones en los inmunopeptidomas. Comprender los determinantes estructurales y funcionales que originan estos repertorios peptídicos es esencial para explicar la amplitud y especificidad de la presentación antigénica, así como sus implicaciones en el reconocimiento inmunitario.

| Supertipo | P2                      | PΩ                                             |
|-----------|-------------------------|------------------------------------------------|
| A01       | Pequeños y alifáticos   | Aromáticos e hidrofóbicos grandes              |
| A02       | Pequeños y alifáticos   | Alifáticos e hidrofóbicos pequeños             |
| A03       | Pequeños y alifáticos   | Básicos                                        |
| A24       | Aromáticos y alifáticos | Aromáticos, alifáticos e hidrofóbicos          |
| B07       | Prolina                 | Aromáticos, alifáticos e hidrofóbicos          |
| B08       | Indefinido              | Aromáticos, alifáticos e hidrofóbicos          |
| B27       | Básicos                 | Aromáticos, alifáticos, básicos e hidrofóbicos |
| B44       | Ácidos                  | Aromáticos, alifáticos e hidrofóbicos          |
| B58       | Pequeños                | Aromáticos, alifáticos e hidrofóbicos          |
| B62       | Alifáticos              | Aromáticos, alifáticos e hidrofóbicos          |

Tabla 1: Especificidad de los supertipos de HLA-I<sup>143</sup>.



## 1. SUPERTIPO B7

El supertipo B7 está formado por diversos alelos que presentan péptidos con Pro2 y residuos alifáticos (A, L, I, M y V) o aromáticos (F, W y Y) en posición C-terminal<sup>143</sup>. Se ha descrito que HLA-B\*51:01 y HLA-B\*51:08 pueden unir, además de Pro, un elevado porcentaje de Ala en la posición *P2* del péptido<sup>145-147</sup>. Otros alotipos de este supertipo, como HLA-B\*07:02 muestra una presencia mucho menor de péptidos con Ala<sup>148</sup>. Además, en HLA-B\*51:01 y HLA-B\*51:08, los péptidos con Pro2 unen mayoritariamente residuos hidrofóbicos en *P1*, mientras que un alto porcentaje de los péptidos con Ala2 unen residuos ácidos, principalmente Asp en la posición *P1*. Por tanto, estas moléculas presentan repertorios con dos subpeptidomas diferenciados: péptidos con Pro2 y residuos hidrofóbicos en *P1* y péptidos con Ala2 y Asp<sup>149,150</sup>.

No se ha abordado si la presencia de subpeptidomas, entendidos como la modificación de la preferencia en la posición *P1* dependiendo del aminoácido en la posición *P2*, es habitual entre las demás moléculas del supertipo B7. Asimismo, tampoco se ha estudiado qué residuos favorecen que se incremente el porcentaje con péptidos con Ala2 y, si la presencia de este aminoácido es la causante de la formación de subpeptidomas, influyendo la preferencia de uno u otro aminoácido en la posición *P1*. Además, los determinantes estructurales responsables de esta diversidad no se han caracterizado completamente. Modelos estructurales previos han señalado a los residuos localizados en las posiciones 45 y 67 del surco de unión como posibles moduladores de la especificidad en *P2*, lo que podría contribuir a la aparición de subpeptidomas<sup>151,152</sup>.

En este contexto y con el objetivo de dilucidar estos mecanismos, la presente tesis aborda el estudio de subpeptidomas en moléculas HLA-I, analizando la presencia de subpeptidomas de 13 moléculas de HLA-I pertenecientes al supertipo B7 y el impacto sobre el repertorio peptídico que supone mutar las posiciones 45 y 67, dos residuos localizados en el surco de unión al péptido, de tres de estas moléculas: HLA-B\*51:01, -B\*35:01 y -B\*07:02, que podrían modular la formación de subpeptidomas y la especificidad del residuo localizado en *P1* y *P2*.



## **HIPÓTESIS**



El trabajo de esta tesis consiste en dos apartados diferenciados, cada uno de los cuales ha partido de una hipótesis postulada en base a los resultados previos obtenidos de la literatura:

1. Las diferencias estructurales de las subunidades  $\beta 5$  y  $\beta 5i$  condicionan la actividad catalítica del proteasoma constitutivo y del proteasoma intermedio de tipo I, respectivamente, por lo que ambos complejos generan inmunopeptidomas presentados por HLA de clase I diferentes.
2. La presencia de péptidos con Ala2 y de subpeptidomas entre los alotipos del supertipo B7 no se limita a HLA-B\*51:01 y HLA-B\*51:08. Los residuos 45 y 67 de la cadena  $\alpha$  de las moléculas de HLA de clase I son relevantes en la presencia de péptidos con alanina en posición 2 y en la presencia de subpeptidomas.



## **OBJETIVOS**





Los **objetivos principales** de la presente tesis son:

1. Caracterizar la especificidad del proteasoma intermedio de tipo I y su influencia en el inmunopeptidoma de HLA de clase I, así como determinar su estructura atómica.
2. Determinar la presencia de subpeptidomas en moléculas de HLA-I pertenecientes al supertipo B7. Describir la contribución de las posiciones 45 y 67 en la presencia de péptidos con Ala2 y la formación de subpeptidomas.

Los **objetivos específicos** son los siguientes:

1. Analizar comparativamente los inmunopeptidomas de HLA de clase I de células que expresan el proteasoma constitutivo y el proteasoma intermedio de tipo I.
2. Caracterizar la actividad catalítica del proteasoma intermedio de tipo I.
3. Determinar la estructura del proteasoma intermedio de tipo I.
4. Identificar la presencia de Ala2 y de subpeptidomas en los repertorios peptídicos disponibles en la literatura de moléculas de HLA-I pertenecientes al supertipo B7.
5. Evaluar la contribución de las posiciones 45 y 67 en la presencia de péptidos con Ala2 y de subpeptidomas en los repertorios peptídicos de HLA-B\*51:01, -B\*35:01 y -B\*07:02.



## **MATERIALES Y MÉTODOS**



## **CULTIVOS CELULARES**

### **1. LÍNEAS CELULARES**

- HEK-293: línea celular obtenida a partir de la transformación de células embrionarias de riñón humano con el adenovirus 5<sup>153,154</sup>. Presenta crecimiento en adherencia y es fácilmente cultivable y transfectable. Expresa exclusivamente el proteasoma constitutivo.
- HEK-293F: adaptación de la línea celular HEK-293 al crecimiento en suspensión y sin suplemento de suero fetal bovino (FBS). Está ampliamente utilizada por su fácil crecimiento a gran escala<sup>155</sup>.
- HMy2.C1R: línea celular linfoblastoide transformada con el virus Epstein-Barr Virus (EBV)<sup>156</sup>. Expresa mayoritariamente el proteasoma constitutivo, y algo de inmunoproteasoma<sup>157</sup>.

### **2. CONDICIONES DE CULTIVO**

La línea celular HEK-293 se cultivó en medio *Dulbecco's Modified Eagle Medium* (DMEM) (Gibco) suplementado con 10% de suero fetal bovino (FBS) (Gibco) en condiciones de esterilidad a 37°C, 5% CO<sub>2</sub>. Las células se cultivaron en placas de 96, 24, o 6 pocillos (ThermoFisher), *Flask* T25cm<sup>2</sup>, *Flask* T75cm<sup>2</sup> y *Flask* T175cm<sup>2</sup> (Corning) según las necesidades del cultivo y del experimento. Se realizaron pases celulares cada 3-4 días mediante tripsinización. Para ello, se descartó el medio de cultivo y se realizó un lavado con el volumen pertinente de *Phosphate Buffer Saline* (PBS). Se añadió el volumen adecuado de Tripsina-EDTA y se incubó durante 5 minutos a 37°C, 5% CO<sub>2</sub>. Tras descartar el exceso de células, se añadió medio fresco y se incubaron a 37°C, 5% CO<sub>2</sub>. Tras la transfección de estas células a modo de selección, o cuando el cultivo lo requiera, las células fueron cultivadas con medio DMEM 10% FBS 1mg/ml Geneticina (G418) (Gibco).

La línea HEK-293F se cultivó en medio Freestyle (Gibco) en condiciones de esterilidad en un incubador a 37°C, 8% CO<sub>2</sub> y agitación orbital (150 rpm). Las células se cultivaron en Matraces Erlenmeyer de 125 ml, 500 ml, 1L o 2L (Benchmark Scientific) según las necesidades del cultivo y del experimento. La concentración celular se comprobó cada dos días y se ajustó su concentración a 5x10<sup>5</sup> o 3x10<sup>5</sup> células/ml para mantenimiento, o a la concentración requerida por el experimento.

La línea HMy2.C1R (C1R) se cultivó en medio RPMI-1640 (Sigma-Aldrich) suplementado con 10% de FBS en condiciones de esterilidad en un incubador a 37°C, 5% CO<sub>2</sub>. Cuando se necesitaba disponer de grandes volúmenes de células, se utilizó el sistema de *roller bottles*, incubando las células a 37°C sin control de CO<sub>2</sub>. Cada dos días la concentración se ajustó a 5x10<sup>5</sup> o 3x10<sup>5</sup> células/ml para mantenimiento, o a la concentración requerida por el experimento.

*Pellets* secos. Las células se centrifugaron en tubos falcon (Corning) y se realizaron 3 lavados con PBS frío a 300 RCF durante 5 minutos a 4°C. Los pellets se sumergieron en N<sub>2</sub> líquido y se almacenaron a -80°C.

## **CONSTRUCCIONES GÉNICAS**

### **1. DISEÑO DE PRIMERS**

#### **1.1. Diseño de primers para el análisis por qPCR de los alelos HLA-A\*02:01, HLA-A\*03:01 y HLA-B\*07:02**

A partir de las secuencias genómicas de los alelos *HLA-A\*02:01*, *HLA-A\*03:01* y *HLA-B\*07:02* de la molécula de HLA de clase I, obtenidas de la base de datos GenBank<sup>158</sup>, se realizó un BLAST de las dos primeras, con tal de observar donde se encontraba la zona polimórfica entre estas. Los *primers* específicos de alelo se construyeron a partir de esta región polimórfica entre las bases 62 y 132, en la que había 9 *missmatches*, mediante la base de datos NCBI *Primer Designing Tool*<sup>159</sup>. Los *primers* seleccionados fueron:

- *HLA-A\*02:01*
  - Forward: 5'-CGCATATGACTCACCACGCT-3'
  - Reverse 5'-AGGTCAGTGTGATCTCCGA-3'
- *HLA-A\*03:01*
  - Forward: 5'-CACATATGACCCACCACCCC-3'
  - Reverse: 5'-GCCAGGTCAGTGTGATCTCC-3'
- *HLA-B\*07:02*
  - Forward: 5'-GCTCCCACTCCATGAGGTATTTCT-3'
  - Reverse 5'-TCTCTCGGTCAGTCTGTGCCTG-3'

#### **1.2. Diseño de primers para la mutagénesis de las moléculas del supertipo B7**

Los *primers* de las moléculas de HLA-I pertenecientes al Supertipo B7 se realizaron a partir de las secuencias genómicas de los alelos de clase I *HLA-B\*51:01*, *HLA-B\*35:01* y *HLA-B\*07:02* obtenidas de la base de datos GenBank.

Los *primers* empleados para realizar la mutagénesis dirigida de los alelos *HLA-B\*51:01*, *HLA-B\*35:01* y *HLA-B\*07:02*, se realizaron con el fin de mutar las posiciones 45, 67 y ambas juntas. En el caso de las moléculas *HLA-B\*51:01* y *HLA-B\*35:01* se sustituyó T45E y F67Y, mientras que la molécula *HLA-B\*07:02* se sustituyó E45T y Y67F.

## 2. CONSTRUCCIONES GÉNICAS

Se realizaron diferentes transfecciones sobre las líneas celulares HEK-293, HEK-293F y C1R con diferentes construcciones en el vector pcDNA3.1:

- **HEK-293-β5i-StII**: línea celular HEK-293 transfectada con el vector pcDNA3.1 que codifica la subunidad β5i del proteasoma junto con el “tag” Strep-tag II.
- **HEK-293F-β5-StII**: línea celular HEK-293F transfectada con el vector pcDNA3.1 que codifica la subunidad β5 del proteasoma junto con el “tag” Strep-tag II.
- **HEK-293F-β5i-StII**: línea celular HEK-293F transfectada con el vector pcDNA3.1 que codifica la subunidad β5i del proteasoma junto con el “tag” Strep-tag II.
- 
- **C1R-HLA-B\*51:01-StII**: línea celular HMy2.C1R transfectada con el vector pcDNA3.1 que codifica la molécula de *HLA-B\*51:01* junto con el “tag” Strep-tag II.
  - **C1R-HLA-B\*51:01-StII T45E, C1R-HLA-B\*51:01-StII F67Y, C1R-HLA-B\*51:01-StII T45E y F67Y**: misma línea celular, pero con cambios de aminoácido en las posiciones 45, 67 y en ambas realizadas por mutagénesis dirigida.
- **C1R-HLA-B\*35:01-StII**: línea celular HMy2.C1R transfectada con el vector pcDNA3.1 que codifica la molécula de *HLA-B\*35:01* junto con el “tag” Strep-tag II.
  - **C1R-HLA-B\*35:01-StII T45E, C1R-HLA-B\*35:01-StII F67Y, C1R-HLA-B\*35:01-StII T45E y F67Y**: misma línea celular, pero con cambios de aminoácido en las posiciones 45, 67 y en ambas realizadas por mutagénesis dirigida.



- **C1R-HLA-B\*07:02-StII**: línea celular HMy2.C1R transfectada con el vector pcDNA3.1 que codifica la molécula de HLA-B\*07:02 junto con el “tag” Strep-tag II.
  - **C1R-HLA-B\*07:02-StII E45T, C1R-HLA-B\*07:02-StII Y67F, C1R-HLA-B\*07:02-StII E45T y Y67F**: misma línea celular, pero con cambios de aminoácido en las posiciones 45, 67 y en ambas realizadas por mutagénesis dirigida.

### 3. EXTRACCIÓN DE RNA, RETROTRANSCRIPCIÓN Y GADPH

Se partió de pellets secos de  $10^7$  células almacenados a  $-80^{\circ}\text{C}$ . La extracción de mRNA se realizó utilizando el kit RNeasy Plus Mini kit (Qiagen) siguiendo las indicaciones del fabricante. Posteriormente se midió la concentración de RNA mediante Nanodrop (ThermoFisher) y el RNA obtenido se almacenó a  $-80^{\circ}\text{C}$ .

A partir del RNA previamente purificado se obtuvo el cDNA mediante la acción de la retrotranscriptasa *Moloney Murine Leukemia Virus* (M-MLV RT) (Invitrogen). Se generó una primera *mix* compuesta por Oligo(dT)s  $1\text{ }\mu\text{g}/\mu\text{l}$ , dNTPs 10mM y agua DEPC junto con la muestra, a  $65^{\circ}\text{C}$  durante 5 minutos en el Termociclador SimpliAmp (Thermofisher). La segunda *mix* constó de Buffer 5x, DTT 0,1M y RNase out 40U/  $\mu\text{l}$ , y se añadió al termociclador SimpliAmp durante 2 minutos a  $37^{\circ}\text{C}$ . Finalmente se añadió la retrotranscriptasa (M-MLV) durante 50 minutos a  $37^{\circ}\text{C}$  seguido de 15 minutos a  $70^{\circ}\text{C}$  en el Termociclador SimpliAmp. La concentración del cDNA obtenido se midió mediante Nanodrop y finalmente se almacenó a  $-20^{\circ}\text{C}$  hasta su uso.

La integridad del cDNA se comprobó mediante PCR en una *mix* en la que se incluyeron los *primers* del gen que codifica para la enzima GAPDH, Buffer 10x, dNTPs 2,5mM y agua de PCR, junto con la muestra y *Taq Polimerase* (ThermoFisher) en el termociclador SimpliAmp (ThermoFisher) (un ciclo de 3 minutos a  $95^{\circ}\text{C}$ , seguido de 28 ciclos de 30 s a  $95^{\circ}\text{C}$ , 30 s a  $65^{\circ}\text{C}$  y 30 s a  $72^{\circ}\text{C}$  y finalmente un ciclo de 7 minutos a  $72^{\circ}\text{C}$ ). La visualización del amplímero se realizó en un gel al 1% de agarosa (PanReacAppliChem) con Midori green (Cultek).

#### 4. CLONAJE

La molécula *HLA-B\*35:01* se obtuvo a partir de la amplificación por PCR del cDNA previamente obtenido por retrotranscripción de PBMCs de un donante en una reacción que incluía Buffer 5x, dNTPs 10mM, *primers* específicos, agua de PCR y *Phusion Polimerase* (ThermoFisher), junto con la muestra (un ciclo de 5 minutos a 95°C, seguido de 40 ciclos de 30 s a 95°C, 30 s a 60°C y un minuto y medio a 72°C y finalmente un ciclo de 7 minutos a 72°C) en el termociclador SimpliAmp (ThermoFisher). La molécula *HLA-B\*07:02* se obtuvo mediante el mismo procedimiento, a partir de la línea celular HEK-293. El vector de la molécula *HLA-B\*51:01* fue cedido amablemente por el Dr. López de Castro. Los genes de los alelos de HLA de clase I *HLA-B\*51:01*, *HLA-B\*35:01* y *HLA-B\*07:02* (junto con los mutantes de cada una de las moléculas) se clonaron en el vector de expresión pcDNA3.1-Strep-tag II mediante digestión con las enzimas de restricción HindIII y XhoI (ThermoFisher). Tanto el vector como los diferentes insertos se incubaron durante 2 horas a 37°C y posteriormente se purificaron mediante el kit de NZYGelpure (NZYtech) a partir de una banda de agarosa o de una reacción enzimática, dependiendo del caso.

La subunidad catalítica inmune  $\beta 5i$  del proteasoma ya se encontraba clonada en nuestro laboratorio en el vector pcDNA3.1-Strep Tag II mediante digestión con las enzimas de restricción HindIII y XhoI (ThermoFisher).

## 5. LIGACIÓN, TRANSFORMACIÓN Y SECUENCIACIÓN

La ligación entre los diferentes insertos y vectores se realizó con la enzima *T4 ligase* (ThermoFisher) diluida en *Buffer T4 Ligase Buffer* (ThermoFisher). Para ello, se llevaron a cabo diferentes relaciones vector-inserto (1:1, 1:5, 1:10) y se incubaron durante 2 horas RT. En todos los casos se realizó un control de religación. La transformación se realizó con bacterias competentes XL-1-Blue almacenadas a -80°C. Se añadió entre un 20-40% del producto de ligación a un volumen total de 50 µl de bacterias competentes y se incubaron 20 minutos en hielo. Posteriormente se realizó un choque térmico a 42°C durante 90 segundos. Se incubaron 5 minutos en hielo y se plaquearon en placas de LB-Agar (1% Triptona, 0,5% Extracto de Levadura, 1% NaCl, 1,5% Agar pH 7,2) con Ampicilina (100 µg/ml) (Sigma-Aldrich) mediante *glass beads*. Las placas se incubaron O/N a 37°C. Se seleccionaron 6 colonias de las placas y se cultivaron en 5 ml de medio de cultivo LB con 0,1 mg/ml de ampicilina. Se incubaron a 37°C en agitación 250 rpm O/N.

El DNA plasmídico se purificó con el kit NZYMiniprep (NZYtech) siguiendo las instrucciones del fabricante. La concentración del DNA obtenido se cuantificó con el Nanodrop, así como la pureza. El DNA obtenido se digirió de nuevo con las mismas enzimas de restricción utilizadas para el clonaje, y se corrió un gel de agarosa al 1%.

Aquellas construcciones que presentaban el inserto y vector adecuados se secuenciaron utilizando los *primers* T7 y BGH en el Servicio de Genómica y Bioinformática del IBB-UAB.

Finalmente, se seleccionó una de las construcciones correctas y se inoculó en un cultivo de 250 ml de LB con 0,1 mg/ml de ampicilina y se incubó a 37°C en agitación O/N. Posteriormente se purificó el DNA plasmídico utilizando el kit de Maxiprep de plásmidos HiPure PureLink (ThermoFisher). La concentración del DNA obtenido se cuantificó en el Nanodrop, así como la pureza.

## ANTICUERPOS

| <b>ANTICUERPOS PRIMARIOS</b>                                        |
|---------------------------------------------------------------------|
| Anti-HLA-I (W6/32) mouse monoclonal antibody (Hibridoma)            |
| Anti-HLA-B*07 (ME.1) mouse monoclonal antibody (Hibridoma)          |
| Anti-HLA-A*02 (PA2.1) mouse monoclonal antibody (Hibridoma)         |
| Anti-HLA-A*03 (GAP-A3) mouse monoclonal antibody (BD Pharmingen)    |
| Anti-HLA-I vacía (HC.10) mouse monoclonal antibody (Hibridoma)      |
| Anti-HLA-II (B8.11.2) mouse monoclonal antibody (Hibridoma)         |
| Anti-NWSHPQFEK Tag mouse monoclonal antibody (GenScript)            |
| Anti-Vinculin mouse monoclonal antibody (Invitrogen)                |
| Anti-β5 (A-10) mouse monoclonal antibody (Santa Cruz Biotechnology) |
| Anti-β5i (mouse monoclonal antibody (Santa Cruz Biotechnology)      |

**Tabla 2: Anticuerpos primarios utilizados**

| <b>ANTICUERPOS SECUNDARIOS</b>                                                               |
|----------------------------------------------------------------------------------------------|
| AlexaFluor488 anti-Mouse IgG, goat polyclonal antibody IgG (H+L) (Invitrogen)                |
| AlexaFluor488 anti-Rabbit IgG, goat polyclonal antibody IgG (H+L) (Invitrogen)               |
| AlexaFluor568 anti-Mouse IgG, goat polyclonal antibody IgG (H+L) (Invitrogen)                |
| ECL anti-Mouse IgG, Horseradish Peroxidase-Linked whole antibody from sheep (GE Healthcare)  |
| ECL anti-Rabbit IgG, Horseradish Peroxidase-Linked whole antibody from sheep (GE Healthcare) |
| Precision Plus Protein StrepTactin-HRP Conjugate 5,000x (Bio-Rad)                            |

**Tabla 3: Anticuerpos secundarios utilizados**

## **CITOMETRÍA DE FLUJO**

### **1. TINCIÓN CELULAR DE SUPERFICIE**

Se dispusieron  $2 \times 10^5$  células/pocillo en una placa de 96 pocillos de fondo cónico (ThermoFisher). Se centrifugó la placa a 840 RCF durante 2 minutos a 4°C y se descartó el sobrenadante. Las células se lavaron dos veces con 200 µl de PBS, 2% FBS frío y una vez con PBS frío. Posteriormente las células se resuspendieron con 50 µl de la dilución pertinente de anticuerpo primario en PBS, 2% FBS o de PBS, 2% FBS frío en el caso de los controles. Se incubaron 30 minutos a 4°C. Tras la incubación se repitieron los lavados anteriores, las células se incubaron con 50 µl de la dilución pertinente en PBS, 2% FBS de anticuerpo secundario y se incubaron 30 minutos 4°C en oscuridad. Finalmente se realizaron 3 lavados con PBS 1 frío y se resuspendieron en 300 µl de PBS frío y se traspasaron las células a tubos de citometría.

### **2. TINCIÓN INTRACELULAR**

Se dispusieron  $5 \times 10^5$  células/pocillo en una placa de 96 pocillos de fondo cónico (ThermoFisher). Se centrifugó la placa a 840 RCF durante 2 minutos a temperatura ambiente y se descartó el sobrenadante. Se realizó un lavado de las células con PBS y posteriormente se fijaron durante 20 minutos con 200 µl de una solución de paraformaldeído (PFA) al 4%. Transcurrido el tiempo se centrifugó la placa a 840 RCF durante 2 minutos a temperatura ambiente y se descartó el sobrenadante. Se lavaron las células dos veces con 200 µl de PBS y una vez con solución de tinción (PBS, 0,5% Triton X100, 2% FBS, 0,05%  $\text{NaN}_3$ ). Las células se incubaron con 50 µl de la dilución pertinente de anticuerpo primario en solución de tinción o con 50 µl de solución de tinción en el caso de los controles, durante 1 hora a temperatura ambiente. Se lavaron las células dos veces con 200 µl de solución de tinción y se incubaron con 50 µl de la dilución pertinente de anticuerpo secundario en solución de tinción, durante 1 hora a temperatura ambiente y oscuridad. Transcurrido el tiempo se lavaron las células dos

veces con 200 µl de solución de tinción y una vez con 200 µl de PBS. Finalmente, se resuspendieron en 300 µl de PBS y se traspasaron a tubos de citometría.

### 3. TOMA Y ANÁLISIS DE DATOS

Se emplearon los citómetros CytoFlex (Beckman Coulter Life Science) y CytoFlexLX (Beckman Coulter Life Science) para adquirir los datos, y el *software* FlowJo v10<sup>160</sup> para analizarlos. Se seleccionó la región 1 correspondiente a las células vivas a través de los parámetros de tamaño y complejidad celular (FSC/SSC). Posteriormente se seleccionaron los singletes a través de la relación área-altura (FSC-a/FSC-h). El grado de expresión del gen de interés se determinó mediante el análisis de la fluorescencia emitida en el canal del fluorocromo del anticuerpo primario o secundario.

# **CARACTERIZACIÓN DEL PROTEASOMA INTERMEDIO DE TIPO I**

## **1. TRANSFECCIÓN**

### **1.1. Transfección con PEI de células adherentes. HEK-293**

Se plaquearon  $2 \times 10^5$  células/pocillo en placas de 6 pocillos 24 horas antes del experimento en 2ml de medio DMEM 10% FBS y se incubaron 37°C, 5% CO<sub>2</sub>. El día del experimento, las células fueron privadas de suero mediante la sustitución del medio de cultivo por 1ml de medio DMEM no suplementado durante 30 minutos 37°C 5% CO<sub>2</sub>. La solución de transfección se generó diluyendo 1 µg DNA en 300 µl de DMEM sin suplementar y el PEI (Polietilenimina) (Polysciencies Hirschberg, Germany) en una relación PEI: DNA 3:1 en 300 µl de DMEM sin suplementar. Se añadió la solución del PEI sobre la del DNA, se resuspendió con el vórtex durante 15 segundos y se incubó 30 minutos a temperatura ambiente. La dilución de DNA y PEI se añadió gota a gota a cada pocillo de la placa, agitando esta levemente y se incubaron 4 horas a 37°C, 5% CO<sub>2</sub>. Transcurrido el tiempo de incubación se sustituyó el medio de cultivo por 2 ml de medio DMEM 10% FBS. 48 horas tras la transfección, se sustituyó el medio de cultivo por medio DMEM 10% FBS 1mg/ml G418 (Gibco). Para aquellos casos en los que se quisiera obtener una población pura se realizó una dilución límite: la concentración de células se ajustó a  $3 \times 10^5$  células/ml y se realizaron diluciones seriadas 1/10 en medio de cultivo DMEM 10% FBS 1 mg/ml G418 (Gibco). Se plaquearon 100 µl de las diluciones 30 células/ml y 3 células/ml en placas de 96 pocillos de fondo plano (ThermoFisher), que previamente contenían 100 µl de medio DMEM 10% FBS 1 mg/ml G418 (Gibco). Las placas se incubaron a 37°C 5% CO<sub>2</sub> comprobando periódicamente el crecimiento de los pocillos a través del viraje de color del medio por su acidificación. Los pocillos en los cuales crecieron células se escalaron a placas de 6 pocillos (ThermoFisher) y se testó la expresión del gen de interés mediante una tinción intracelular con un anticuerpo específico de la subunidad o contra el “tag” por citometría de flujo.

### **1.2. Transfección con PEI de células en suspensión. HEK-293F**

La concentración de las células se ajustó a  $5 \times 10^5$  células/ml en 10 ml de medio de cultivo Freestyle en un Matraz Erlenmeyer de 125 ml 24 horas antes de comenzar el experimento y se incubaron 37°C, 8% CO<sub>2</sub> en agitación orbital 150 rpm. La solución de transfección se generó diluyendo 10 µg DNA en 500 µl de medio Freestyle y el PEI en una relación PEI:DNA 3:1 en 500 µl de medio Freestyle. Se añadió la solución del PEI sobre la del DNA, se resuspendió con el vórtex durante 15 segundos y se incubó 15 minutos a temperatura ambiente. La dilución de DNA y PEI se añadió gota a gota al Erlenmeyer, mientras este se agitaba suavemente. Se añadió ácido valproico (Sigma-Aldrich) a una concentración final en el cultivo de 4mM y se incubaron las células 24 horas a 37°C, 8% CO<sub>2</sub> en agitación. Transcurridas 24 horas las células se lavaron con PBS a 1300 RCF durante 5 minutos a 4°C y se substituyó por medio de cultivo Freestyle suplementado con 1mg/ml G418 (Gibco). Se testó la expresión del gen de interés mediante una tinción intracelular con un anticuerpo específico de subunidad o contra el "tag" por citometría de flujo.

## **2. WESTERN BLOT**

Se realizó la lisis celular de pellets secos almacenados a -80°C con el volumen correspondiente de Buffer de Lisis (150 mM NaCl, 20 mM Tris-HCl, inhibidor de proteasas (Roche), 1% NP40, pH 7,4). Se centrifugó a 16000 RCF durante 10 minutos a 4°C. La cuantificación de proteínas se realizó con el kit de cuantificación BCA (ThermoFisher) y el lector de placas por colorimetría Viktor3 (PerkinElmer) (562 nm).

Unos 20 µg de proteína diluidos en Buffer RBS 5x (Glicerol 25%, SDS 7,5%, Tris 337 mM, DTT 250 mM, 0,25% azul de bromofenol) desnaturalizados a 95°C durante 5 minutos, se analizaron por electroforesis SDS-PAGE en geles de poliacrilamida al 14 ó 16%. Las proteínas separadas se transfirieron a una membrana de PVDF durante 1 hora a un amperaje de 300mA. Las membranas se bloquearon durante 1 hora a RT en agitación suave con 10 ml de tampón de bloqueo (PBS 0,1%, Tween20 5%, leche en polvo). Transcurrido el bloqueo, las membranas se lavaron tres veces durante 5 minutos en



agitación fuerte con 10 ml de PBS 0,1% Tween20. Se incubaron las membranas durante 1 hora RT o durante toda la noche a 4°C en agitación suave con 10 ml de la dilución pertinente de anticuerpo primario en tampón de bloqueo. Se realizaron tres lavados durante 5 minutos en agitación fuerte con 10 ml de PBS 0,1% Tween20. Se incubaron las membranas con 10 ml de la dilución pertinente de anticuerpo secundario durante 1 hora a RT en agitación suave. Se realizaron tres lavados durante 5 minutos en agitación fuerte con 10 ml de PBS 0,1% Tween20. Tras el último lavado, la membrana se incubó con la solución de revelado Clarity Western ECL Substrate (Bio-Rad) y las bandas específicas se detectaron por quimioluminiscencia a través del LI-COR Licor Odyssey FC 2800 (LICORbio). Las imágenes fueron analizadas mediante el *software* Fiji (ImageJ)<sup>161</sup>.

### **3. PURIFICACIÓN DEL PROTEASOMA 20S**

Todo el proceso se realizó a 4°C. Se lisaron aproximadamente  $10^9$  células en 20 ml de Buffer de lisis (Tris 50 mM, KCl 25 mM, pH 8). Se realizaron 4 rondas de congelación-descongelación del lisado en N<sub>2</sub> líquido. Posteriormente el lisado se homogeneizó mecánicamente con un *Potter*. El lisado se centrifugó a 2500 RCF durante 10 minutos a 4°C. El sobrenadante se centrifugó a 50000 RCF durante 3 horas a 4°C. El proteasoma se purificó en una columna de de Strep Tactin XT 4Flow Resin (Ibion Technologies) según las instrucciones del fabricante. Se recogieron fracciones de 1ml y se midió la concentración de proteína en un Nanodrop. La pureza se midió en un gel de acrilamida al 14% en condiciones desnaturalizantes.

Las fracciones que contenían proteasoma se cargaron en un filtro Vivaccon2 de 125 kDa (ThermoFisher). Se centrifugó a 2000 RCF a 4°C hasta que pasó todo el volumen de las muestras y se lavó cuatro veces con 2ml de Buffer Tris 10mM 1mM EDTA pH 7,5. Finalmente el proteasoma se eluyó dando la vuelta al Vivaccon2 y centrifugando 500 RCF 5 minutos 4°C. La concentración del proteasoma se determinó mediante el Nanodrop y para comprobar su pureza se realizó un gel de acrilamida al 14% en condiciones desnaturalizantes. El proteasoma se almacenó a -80°C.

#### **4. ENSAYO DE ACTIVIDAD CATALÍTICA**

Los sustratos fluorogénicos (Enzo Life Sciences) se añadieron a una concentración final de 100  $\mu$ M diluidos en Buffer Tris 50 mM, KCl 25 mM, pH 8, junto con 0,5  $\mu$ g de proteasoma purificado diluido en Buffer Tris 50 mM, KCl 25 mM y 10  $\mu$ l de Bortezomib (ThermoFisher) 10nM cuando correspondía en placas negras opacas de 96 pocillos de fondo plano (Sigma-Aldrich) en un volumen final de 200  $\mu$ l. Se incubó en un lector de placas Viktor3 (PerkinElmer) a 37°C y la fluorescencia se midió cada 12 min durante 4 horas (Excitación 350nm, Emisión 430 nm).

Cada experimento se realizó por triplicado en días distintos. Previamente, se llevó a cabo un ensayo de una sola lectura utilizando diferentes concentraciones de la molécula de aminometilcumarina (AMC) (Invitrogen) (1nM, 10 nM, 100 nM, 1 $\mu$ M, 10  $\mu$ M) diluida en Dimetilsulfóxido (DMSO) (Sigma-Aldrich) para establecer el control positivo del experimento.

#### **5. ESTRUCTURA DEL PROTEASOMA**

##### **5.1. Preparación de la muestra**

Se añadieron 3  $\mu$ l de muestra sobre una rejilla de oro perforada (UltrAuFoil R1.2/1.3, 300 mallas) tratada para ser hidrofílica mediante un tratamiento de plasma de 120 s en un limpiador de plasma Fischione 1070, operando al 34% de potencia con una mezcla gaseosa de 80% argón y 20% oxígeno. Las rejillas se secaron por *blotting* durante 3,0 s con una fuerza de blot de 3, a 6 °C y 100% de humedad, y luego se congelaron rápidamente en etano líquido utilizando un Vitrobot Mark IV (ThermoFisher). En los casos en los que el proteasoma 20S se cristalizó con el inhibidor Bortezomib, este se incubó durante 1h a 37°C y se siguió el mismo procedimiento mencionado.

Todas las imágenes se recolectaron en un microscopio electrónico Glacios (ThermoFisher) operando a un voltaje de aceleración de 200 kV, acoplado a un detector Falcon 4, equipado con un filtro de energía Selectris X, empleando una ranura de 10 eV en modo “zero-loss”. Las imágenes se adquirieron con una ampliación nominal de

165.000×, correspondiente a un tamaño de píxel de 0,73 Å. Las películas se registraron en un detector en modo de conteo con una tasa de dosis de electrones de 7,88 e<sup>-</sup>/píxel/s y un tiempo total de exposición de 3 s, para una dosis acumulada de 44,48 e<sup>-</sup>/Å<sup>2</sup>. Las adquisiciones se realizaron de forma semiautomática con el *software* SerialEM<sup>162</sup>.

## 5.2. Determinación de la estructura

La corrección del movimiento por parches de los fotogramas de las películas se realizó en CryoSPARC v.4.2.1<sup>163</sup>, y la corrección del CTF se efectuó con Ctfind4 dentro de RELION 5.0<sup>164</sup>. El programa crYOLO 1.9<sup>165</sup> se utilizó para la selección de partículas tras el entrenamiento de un modelo basado en un subconjunto de partículas seleccionadas manualmente. Finalmente, el postprocesamiento del mapa se llevó a cabo con EMready 2.1<sup>166</sup>. El modelado atómico de la estructura se realizó manualmente, basándose en el mapa de densidad de Cryo-EM, utilizando COOT<sup>167</sup>. El modelo final se refinó frente al mapa de densidad de Cryo-EM mediante Phenix.real\_space\_refine<sup>168</sup>

## **EVALUACIÓN DE LA EXPRESIÓN DE HLA-A\*03:01 EN HEK-293 Y HEK-293F**

Previamente se realizó una tinción extracelular con los anticuerpos específicos de alelo anti-HLA-I (W6/32), anti-HLA-A\*02 (PA2.1), anti-HLA-A\*03 (GAP-A3) y anti-HLA-B\*07 (ME.1) de las líneas celulares HEK-293 y HEK-293F y se analizó la expresión de las diferentes moléculas de HLA-I.

### **1. PCR CUANTITATIVA - SYBER GREEN**

El cDNA previamente obtenido por retrotranscripción del RNA de interés se dispuso en placas translúcidas de 96 pocillos y fondo cónico (Bio-rad), junto con el SsoAdvanced Universal SYBR Green Supermix (Bio-rad) y los *primers* específicos para los alelos de HLA de clase I: *HLA-A\*02:01*, *HLA-A\*03:01* y *HLA-B\*07:02*. La placa se cubrió con una película protectora de sellado para placas de PCR Microseal B (Bio-rad) y se centrifugó a 573 RCF durante 1 min. Finalmente se dispuso en el termociclador CFX Opus (Bio-rad) en el programa RT-PCR (30 segundos 95°C, 5 segundos a 95°C y 30 segundos a 65°C). Se realizaron 3 ensayos por triplicado en días distintos.

## **EVALUACIÓN DE LA PRESENCIA DE SUBPEPTIDOMAS EN MOLÉCULAS DE HLA-I DEL SUPERTIPO B7**

### **1. PEPTIDOMAS DE HLA-B**

Los peptidomas de *Immuno Epitope Database and Analysis Resource* (IEDB)<sup>169</sup> se extrajeron a día 02-15-2022. Los péptidos duplicados fueron eliminados y los péptidos con modificaciones post-traduccionales se consideraron como péptidos únicos no modificados. Se tuvieron en cuenta los peptidomas con 500 péptidos o más.

### **2. ANÁLISIS DE ESTRUCTURAS CRISTALOGRÁFICAS**

Las estructuras cristalográficas de las moléculas de HLA fueron descritas y depositadas previamente en el *Protein Data Bank* (PDB, rcsb.org)<sup>170</sup>. El análisis se llevó a cabo con 3 estructuras de HLA-B\*51:01 (1E27, 1E28, 4MJI<sup>171,172</sup>), 12 estructuras de HLA-B\*07:02 (3VCL<sup>173</sup>, 5EO, 5EO1<sup>174</sup>, 5WMN, 5WMO, 5WMP<sup>175</sup>, 6AT5<sup>176</sup>, 6UJ7, 6UJ8<sup>177</sup>, 7LG0, 7LGD, 7LGT<sup>178</sup>) y 18 de HLA-B\*35:01 (1A1N<sup>179</sup>, 1A9B, 1A9E<sup>180</sup>, 1XH3<sup>181</sup>, 1ZSD<sup>182</sup>, 2AXG<sup>183</sup>, 2CIK<sup>184</sup>, 2FYY, 2FZ3<sup>185</sup>, 2H6P<sup>186</sup>, 3LKN, 3LKO, 3LKP, 3LKQ, 3LKR<sup>187</sup>, 4PR5, 4PRA, 4PRN<sup>188</sup>). Las secuencias se alinearon y analizaron el *software* PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

### **3. MUTAGÉNESIS DIRIGIDA**

Se realizó mutagénesis dirigida de los alelos de clase I *HLA-B\*51:01*, *HLA-B\*35:01* y *HLA-B\*07:02* para mutar las posiciones 45, 67 y ambas juntas mediante PCR. En un volumen final de 20 µl se hizo una *mix* en la que se añadió Buffer 5x, dNTPs 10mM, 10 ng cDNA, *primers* específicos para la mutagénesis, agua de PCR y *Phusion Polimerase* 5 minutos a 98°C, 30 segundos a 98°C, seguido de 30 segundos a 64°C y 4 minutos y 7 minutos a 72°C. El producto de PCR se digirió con DpnI (ThermoFisher) durante 2 horas 37°C para degradar el DNA molde metilado.

#### **4. TRANSFECCIÓN POR ELECTROPORACIÓN DE CÉLULAS EN SUSPENSIÓN**

La transfección se realizó como se había descrito previamente<sup>189</sup>. Alrededor de  $10^7$  células HMy2.C1R se lavaron dos veces con PBS frío centrifugando a 300 RCF durante 5 minutos a 4°C. Se descartó el sobrenadante y las células se resuspendieron en 800  $\mu$ l de PBS frío y se añadieron a cubetas de electroporación (Bio-Rad). Se resuspendieron 10  $\mu$ l de DNA con las células y la electroporación se llevó a cabo con el electroporador Gene Pulser (Bio-Rad) (960  $\mu$ F, 250mA). Se incubaron 10 minutos en hielo y se resuspendieron en 10 ml de medio de cultivo DMEM 10% FBS y se traspasaron a *Flask* T75 cm<sup>2</sup>. Transcurridas 48 horas tras la transfección se centrifugaron las células y se resuspendieron en 10 ml de medio DMEM, 10% FBS, 1 mg/ml G418 y se plaquearon en placas de 96 pocillos de fondo plano (ThermoFisher). Las placas se incubaron a 37°C, 5% CO<sub>2</sub> comprobando periódicamente el crecimiento de los pocillos a través del viraje de color del medio por su acidificación. La expresión de los diferentes alelos se analizó mediante una tinción extracelular con el anticuerpo W6/32 por citometría flujo.

## **PURIFICACIÓN DE LOS INMUNOPEPTIDOMAS DE HLA DE CLASE I**

### **1. PURIFICACIÓN DE INMUNOPEPTIDOMAS DE MOLÉCULAS DE HLA-I**

#### **1.1. Purificación de péptidos de unión a moléculas de HLA-I mediante columna de Strep-Tactin**

Los péptidos asociados a moléculas de HLA-I se obtuvieron mediante cromatografía de afinidad, elución ácida y ultrafiltración. Se lisaron  $1 \times 10^9$  células con 20 ml de Buffer de lisis (50 mM Tris 150 mM NaCl 1mM EDTA 1% Triton X100, inhibidor de proteasas (Roche) pH 8) y se mantuvieron en rotación durante 2 horas a 4°C. El lisado se centrifugó a 2500 RCF 10 minutos 4°C. El sobrenadante del lisado se centrifugó a 50000 RCF 3 horas 4°C. El lisado se pasó dos veces por una columna de 400 µl de Strep Tactin XT 4Flow Resin (Ibion Technologies) previamente equilibrada con 5 ml de Buffer de lisis sin detergente. Se lavó la columna con Buffer de lisis sin detergente y con Tris 5mM. Finalmente, se eluyó con H<sub>2</sub>O, 0,1% TFA en eppendorfs lo-bind (ThermoFisher).

Las fracciones se cuantificaron en el Nanodrop y aquellas que contenían proteína se pasaron por un filtro Vivacon2 (ThermoFisher) de 2kDa. La fracción no retenida se secó en un SpeedVac hasta un volumen de 50 µl y se desaló utilizando columnas C18 (ThermoFisher) y el eluido se secó en el SpeedVac. Finalmente los péptidos se analizaron por espectrometría de masas. Para comprobar la pureza de la purificación se realizó una electroforesis en condiciones desnaturalizantes (SDS-PAGE) al 16% de acrilamida de la fracción retenida.

## **1.2. Purificación de péptidos de unión a moléculas de HLA-I mediante columna de W6/32**

La purificación de los repertorios peptídicos se realizó como se describió previamente con pequeñas modificaciones<sup>190,191</sup>. Los péptidos asociados a moléculas de HLA-I se obtuvieron mediante cromatografía de afinidad, elución ácida y ultrafiltración. Las células se lisaron en buffer de lisis (Tris 20 mM NaCl 150 mM 1% Igepal CA-630 (Sigma-Aldrich), inhibidor de proteasas (Roche) pH 7,4) y se mantuvieron en agitación 2 horas a 4°C. El lisado se centrifugó a 2500 RCF 10 minutos 4°C. El sobrenadante se centrifugó a 50000 RCF 3 horas 4°C. El lisado se pasó por una columna W6/32-sefarosa (GE Healthcare), previamente equilibrada con 5 ml de Buffer de lisis. Se lavó la columna con Buffer de Lavado (TrisHCl 50mM NaCl saturado EDTA 5mM 10% 0,5% Igepal CA-630 pH 7,5) y posteriormente con NET (TrisHCl 50mM NaCl 150mM EDTA 5mM pH 7,5). Finalmente se lavó con Buffer Tris 5 mM y se eluyó en H<sub>2</sub>O, 0,1% TFA en 10 fracciones de 2 ml en eppendorfs Lo-bind.

## **2. ESPECTROMETRÍA DE MASAS**

El análisis por espectrometría de masas se realizó siguiendo dos metodologías muy similares, según la incorporación de nuevos equipos. En ningún caso los resultados se vieron comprometidos por el uso de dos equipos diferentes.

### **2.1. Metodología 1**

El sistema de HPLC estaba formado por una columna de preconcentración  $\mu$ Pac (ThermoFisher) y una columna  $\mu$ Pac Neo (50 cm (ThermoFisher). El disolvente A fue ácido fórmico al 0,1% y el disolvente B fue ácido fórmico al 0,1% en acetonitrilo al 80%. La separación se llevó a cabo a 60 °C con un caudal de 250 nl/min bajo las siguientes condiciones de elución en gradiente: incremento lineal de 4% a 30% de B en 75 min, incremento lineal hasta 50% de B en 5 min, incremento lineal hasta 90% de B en 5 min y mantenimiento en 90% de B durante 10 min. Los datos se adquirieron en modo DDA. Cada ciclo de adquisición tuvo una duración máxima de 3s e incluyó un escaneo MS (300–1400 m/z) con una resolución de 120000 (FWHM), seguido de escaneos MS/MS



con resolución de 30000 (FWHM). Tanto los espectros de MS como los de MS/MS se adquirieron en el analizador Orbitrap. Se seleccionaron péptidos con cargas de 1 a 4 para fragmentación HCD con una energía de colisión HCD del 30%. Se aplicó una ventana de exclusión dinámica de 10s después de la primera fragmentación de cada ion precursor.

## **2.2. Metodología 2**

Los péptidos se disolvieron en ácido fórmico al 0,1% para el análisis por LC-MS/MS. Este análisis se realizó en un sistema Ultimate 3000 HPLC (ThermoFisher) acoplado en línea a un espectrómetro de masas Orbitrap Exploris 240 (ThermoFisher). El HPLC estaba equipado con una columna de preconcentración PepMap Neo C18 (300  $\mu$ m  $\times$  5 mm; (ThermoFisher)) y una columna PepMap RSLC C18 (75  $\mu$ m  $\times$  50 cm; (ThermoFisher)). Los disolventes A y B fueron, respectivamente, ácido fórmico al 0,1% y ácido fórmico al 0,1% en acetonitrilo al 80%. La cromatografía se realizó a 50 °C con un caudal de 250 nl/min. Se aplicó el siguiente gradiente: 4% de B durante 3 min, incremento lineal hasta 35% de B en 120 min, incremento lineal hasta 90% de B en 1 min y mantenimiento al 90% de B durante 5 min.

El espectrómetro de masas operó en modo DDA. Cada ciclo de adquisición consistió en un escaneo de prospección (350–1200 m/z) con resolución de 60.000 (FWHM) y hasta 25 escaneos MS/MS con resolución de 15000 (FWHM). Se seleccionaron péptidos con cargas de 1 a 5 para fragmentación HCD con una energía de colisión HCD del 30%. Se aplicó una ventana de exclusión dinámica de 25s después de la primera fragmentación de cada ion precursor.

## **2.3. Búsqueda de iones MS/MS, identificación de péptidos y deconvolución de motivos de unión**

Los archivos brutos se analizaron con un flujo de trabajo *Deep novo peptidome* en PEAKS Studio 12.0 (Bioinformatics Solutions) contra el proteoma de referencia humano (20.413 entradas), descargado de Uniprot el 11 de enero de 2024. Para el refinamiento de datos, se habilitaron las opciones “Mass Correction” y “Associate feature with Chimera”. No se seleccionó ninguna enzima y las tolerancias de MS y MS/MS se establecieron en 10 ppm

y 0,02 Da, respectivamente. Se consideraron las siguientes modificaciones variables: oxidación de M y formación de ácido pirroglutámico a partir de Q N-terminal. Los resultados se filtraron con un FDR < 1% a nivel de péptido y solo se consideraron las identificaciones por búsqueda en bases de datos (se excluyeron los resultados *de novo*).

### 3. ANÁLISIS DE RESULTADOS Y ESTADÍSTICA

En el análisis de los péptidos por espectrometría de masas se seleccionaron los péptidos de entre 8 y 13 aminoácidos y se eliminaron los péptidos duplicados y modificados post-traduccionalmente. La localización celular se definió a través de la base de datos Uniprot<sup>192</sup> y se clasificaron en los compartimentos: núcleo, citoplasma, membrana plasmática, retículo endoplasmático, vesículas, mitocondria, aparato de Golgi y secreción y según si pertenecían a la ruta citoplasmática (núcleo, citoplasma y mitocondria) o a la ruta endocítica (retículo endoplasmático, aparato de Golgi, vesículas y secreción).

El grado de solapamiento de los peptidomas analizados se calculó por pares de moléculas de HLA-I cuando se analizaron las moléculas del supertipo B7. La afinidad teórica de los diferentes péptidos se calculó con el *software* NetMHCpan4.1<sup>193</sup> y la asignación alélica con el *software* MHCmotifDecon – 1.0<sup>194</sup>. Finalmente, los residuos de anclaje de los diferentes péptidos se analizaron utilizando el *software* PRBAM<sup>195</sup> y GibbsCluster – 2.0<sup>196</sup>. En el cálculo de la afinidad teórica, la significancia estadística de los resultados obtenidos se evaluó mediante test t de Student para comparar dos variables independientes. Los análisis se realizaron considerando un nivel de significancia de  $p < 0.05$ . Los cálculos estadísticos fueron efectuados utilizando GraphPad *Software*, LLC, San Diego, California, USA.

La presencia de subpeptidomas se llevó a cabo analizando si los residuos localizados en diferentes posiciones se encontraban enriquecidos dependiendo del residuo localizado en P2. Para ello, se realizó un test Chi cuadrado y la corrección de Bonferroni de todas las posiciones, considerando un nivel de significancia de  $p < 0.0$

## **RESULTADOS**



# **CAPÍTULO I: CARACTERIZACIÓN DEL PROTEASOMA INTERMEDIO DE TIPO I**



El proteasoma genera los ligandos peptídicos de las moléculas de HLA-I. La coexistencia de diversos tipos de proteasoma en una misma célula sugiere que cada uno de ellos contribuye de forma diferencial en la generación de los péptidos presentados a los linfocitos T CD8<sup>+</sup>. En este capítulo se ha profundizado en la caracterización del proteasoma intermedio de tipo I:

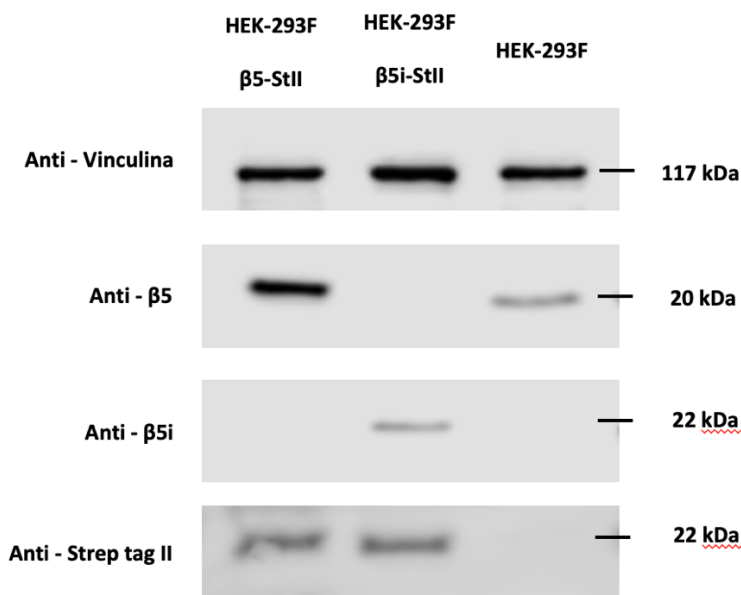
1. Se han analizado comparativamente los inmunopeptidomas de clase I generados por el proteasoma intermedio de tipo I y por el proteasoma constitutivo.
2. Se ha caracterizado la actividad catalítica de este complejo frente al proteasoma constitutivo.
3. Se ha determinado la estructura tridimensional del proteasoma intermedio de tipo I

#### **1. GENERACIÓN DE MODELOS CELULARES PARA LA CARACTERIZACIÓN DEL PROTEASOMA INTERMEDIO DE TIPO I**

Con el fin de estudiar comparativamente las diferencias entre el proteasoma constitutivo y el SIP, se generaron modelos celulares con expresión exclusiva de cada tipo de proteasoma. Para ello, se escogió la línea celular HEK-293F, una línea derivada de la línea HEK-293 adaptada al crecimiento en suspensión. La elección de esta línea se debió a que experimentos previos de nuestro grupo confirmaron que expresa exclusivamente el proteasoma constitutivo ([Anexo 1](#)) (Sholz, E, Tesis Doctoral: *Mecanismos de Procesamiento y Presentación Antigénica en Autoinmunidad*. Universidad Autónoma de Barcelona, 2017) y, además, su crecimiento en suspensión permitía obtener un mayor volumen de células en menor tiempo.

El CDS de los genes que codificaban las subunidades catalíticas **β5 (PSMB5)** y **β5i (PSMB8, LMP7)** se clonaron en el vector de expresión en mamíferos pcDNA3.1, junto con el “tag” Strep-tag II, mediante digestión con enzimas de restricción. Las construcciones pcDNA3.1-β5-StII y pcDNA3.1-β5i-StII se transfectaron sobre la línea celular HEK-293F y se seleccionaron con G418. La expresión de estas proteínas se analizó

por Western Blot con anticuerpos específicos de las subunidades  $\beta 5$  y  $\beta 5i$  y un anticuerpo específico contra el “tag” Strep tag II (**Figura 9**).



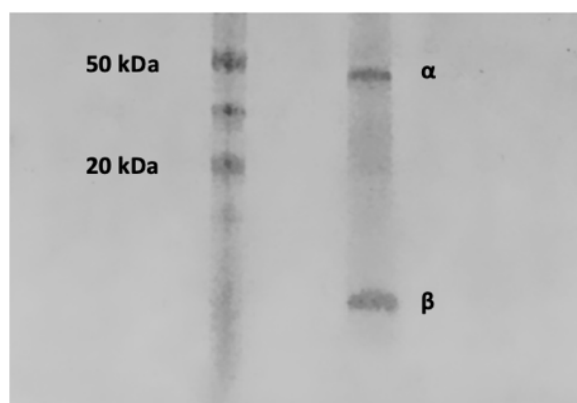
**Figura 9: Expresión proteica de las subunidades  $\beta 5$  y  $\beta 5i$  en los transfectantes generados.** Se cargaron 20  $\mu$ g de lisado total de las líneas celulares HEK-293F- $\beta 5$ -StII, HEK-293F- $\beta 5i$ -StII y HEK-293F en un gel de acrilamida 14%, se transfirieron a una membrana de PVDF y las proteínas se detectaron con anticuerpos primarios anti- $\beta 5$ , anti- $\beta 5i$  y anti-StII. El control de carga se realizó con el anticuerpo anti-vinculina. Las membranas se revelaron mediante quimioluminiscencia emitida por el anticuerpo secundario asociado a HRP. Se aprecia un ligero aumento del peso molecular de las subunidades  $\beta 5$  y  $\beta 5i$  en los transfectantes debido a la fusión de la proteína con el “tag”.

La subunidad  $\beta 5$  se detectó tanto en el transfectante HEK-293F- $\beta 5$ -StII como en el lisado control de HEK-293F. Sin embargo, la subunidad  $\beta 5i$  fue detectada específicamente en el transfectante HEK-293F- $\beta 5i$ -StII, sin evidencia de expresión residual de la subunidad endógena  $\beta 5$ , lo que confirmaba que este transfectante expresaba de forma exclusiva la subunidad  $\beta 5i$  y constituía un buen modelo para el estudio del SIP. Adicionalmente, se validó la expresión de la proteína “tag”, la cual se detectó en ambos transfectantes pero no en la línea parental. Cabe destacar que el peso molecular de estas proteínas quiméricas era ligeramente superior a las subunidades nativas debido a la fusión de la proteína Strep tag II (**Figura 9**).



## 2. ANÁLISIS DEL REPERTORIO PEPTÍDICO DE HEK-293F y HEK-293F-β5i-StII

Los ligandos peptídicos de HLA-I de células que expresan el proteasoma constitutivo y el SIP se purificaron a partir de pellets secos de **HEK-293F** y **HEK-293F-β5i-StII**, respectivamente. El inmunopeptidoma se purificó por inmunoafinidad con el anticuerpo monoclonal W6/32, específico de un determinante monomórfico de HLA-A, B y C, elución ácida y posterior ultrafiltración. El inmunopeptidoma se analizó por espectrometría de masas. La pureza del eluido de la columna de inmunoafinidad, que contenía las cadenas  $\alpha$  y  $\beta$ 2m de HLA-I junto con los péptidos unidos, se evaluó mediante un gel de acrilamida al 16%. La **Figura 10** muestra un experimento representativo de la purificación de los complejos p-MHC de las células HEK-293F-β5i-StII.



**Figura 10: Purificación de los complejos de HLA-I de HEK-293F-β5i-StII.** Se cargaron 10  $\mu$ g de la fracción retenida durante la ultrafiltración en un gel de acrilamida al 16% y se tiñó con *Coomassie Blue*. Se indican las bandas correspondientes a la cadena pesada  $\alpha$  (45 kDa) y a la cadena ligera  $\beta$ 2-microglobulina (12 kDa).

La búsqueda se realizó al 1% FDR y se identificaron un total de 3153 ligandos peptídicos de HLA-I en las células HEK-293F y 702 de HEK-293F-β5i-StII. Se tuvieron en cuenta los péptidos de entre 8 y 13 residuos de longitud y no se tuvieron en cuenta las modificaciones post-traduccionales, resultando en 3102 y 634 péptidos de HEK-293F y HEK-293F-β5i-StII, respectivamente, de los cuales 364 (10,8%) eran compartidos por ambos peptidomas. La mayoría de los péptidos que componían ambos peptidomas estaban compuestos por 9 residuos, seguidos de 10 residuos, 11 residuos, 8 residuos y 12 y 13 residuos, consistente con lo esperado para los inmunopeptidomas de HLA-I (**Figura 11**). El número medio de aminoácidos fue de 9,22 y 9,40 en HEK-293F y HEK-293F-β5i-StII, respectivamente (**Tabla 4**).

|                           | Total | Péptidos por número de residuos (%) |      |      |      |     |     | AAs promedio |
|---------------------------|-------|-------------------------------------|------|------|------|-----|-----|--------------|
|                           |       | 8                                   | 9    | 10   | 11   | 12  | 13  |              |
| HEK-293F                  | 3102  | 7,9                                 | 71,5 | 12,8 | 6,0  | 1,2 | 0,6 | 9,22         |
| HEK-293F- $\beta$ 5i-StII | 634   | 1,9                                 | 60,6 | 17,6 | 14,8 | 3,0 | 2,4 | 9,40         |

Tabla 4: Tamaño, por número de aminoácidos, de los inmunopeptidomas de HLA-I en las líneas celulares HEK-293F y del transfectante HEK-293F- $\beta$ 5i-StII.

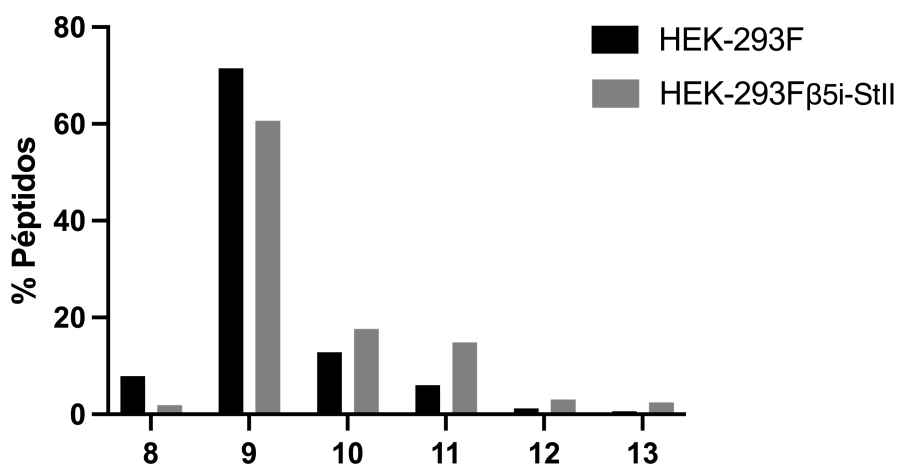


Figura 11: Distribución por tamaño de los péptidos asociados a HLA-I según el número de aminoácidos en las líneas HEK-293F y HEK-293F- $\beta$ 5i-StII. La mayoría de los péptidos eran nonámeros, seguidos de decámeros, octámeros y péptidos de 11, 12 y 13 residuos de longitud.

## 2.1. Localización celular

La localización subcelular de las proteínas parentales se definió mediante la búsqueda de éstas en la base de datos Uniprot<sup>192</sup>. En los casos en que las proteínas aparecían en más de un compartimento celular, se seleccionó aquel más común de acuerdo con su función biológica.

Los péptidos eluidos tanto de HEK-293F como HEK-293F- $\beta$ 5i-StII, provenían de 2142 y 502 proteínas, respectivamente, principalmente del compartimento citoplasmático y del núcleo (Figura 12A). Consecuentemente, la mayoría correspondían a la ruta citosólica, un 77,0% aproximadamente, frente el 23,0% correspondiente a la ruta endocítica (Figura 12B). No se observaron diferencias significativas entre ambas líneas celulares.

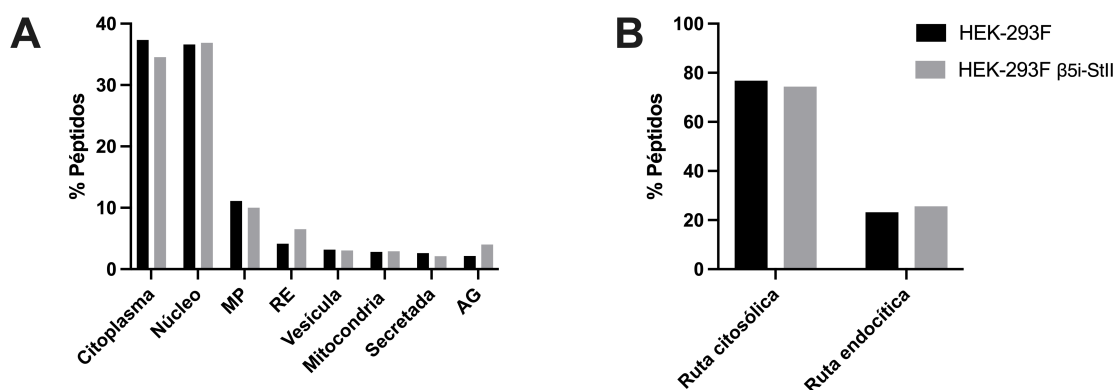


Figura 12: Localización celular de las proteínas parentales de los ligandos de HLA-I en las líneas celulares HEK-293F y HEK-293F-β5i-StII. (A) Porcentaje de péptidos localizados en cada compartimento celular: MP, membrana plasmática; RE, retículo endoplasmático; AG, aparato de Golgi. (B) Porcentaje de péptidos pertenecientes a la ruta citosólica (citoplasma, núcleo y mitocondrias) y endocítica (MP, RE, AG, vesículas y proteínas de secreción).

## 2.2. Asignación alélica de los péptidos presentes en el inmunopeptidoma de HLA-I

El conjunto de péptidos se analizó con el *software* MHCmotifDecon - 1.0<sup>194</sup>, un algoritmo diseñado para descomponer inmunopeptidomas complejos y asignar cada péptido al alelo más probable en base a su motivo de anclaje. La asignación alélica se realizó para los alelos de clase I -A\*02:01, -A\*03:01, -B\*07:02 y -C\*07:02, ya que el tipaje previo tanto de la línea HEK-293 como HEK-293F solicitado por nuestro grupo y realizado en el Laboratory Of Histocompatibility / Laboratory of Immunology Research and Applications to Diagnosis (LIRAD), Banc de Sang and Teixitis de Barcelona demuestra que estos son los alelos expresados por ambas líneas celulares.

Tanto el peptidoma de las células sin transfectar como las células HEK-293F-β5i-StII presentaron mayormente ligandos de HLA-A\*02:01, seguido de HLA-B\*07:02 y HLA-C\*07:02, sin embargo, la asignación alélica de los péptidos puso de manifiesto la escasa presencia de péptidos ligandos de HLA-A\*03:01. El *software* identificó 1909 (61,5%) péptidos ligandos de HLA-A\*02:01, 996 (32,1%) péptidos ligandos de -B\*07:02, 128 (4,1%) péptidos ligandos de -C\*07:02 y solo 25 (0,8%) péptidos ligandos de HLA-A\*03:01 en el inmunopeptidoma de HEK-293F. Se identificaron 44 (1,1%) péptidos como *trash*. Por su parte, en el inmunopeptidoma de HEK-293F-β5i-StII se identificaron 338 (53,3%) péptidos ligandos de HLA-A\*02:01, 238 (37,5%) péptidos ligandos de -B\*07:02, 37 (5,8%) péptidos ligandos de -C\*07:02, 7 (1,1%) de -A\*03:02 y 14 (2,2%) de péptidos se identificaron como *trash* (Figura 13).

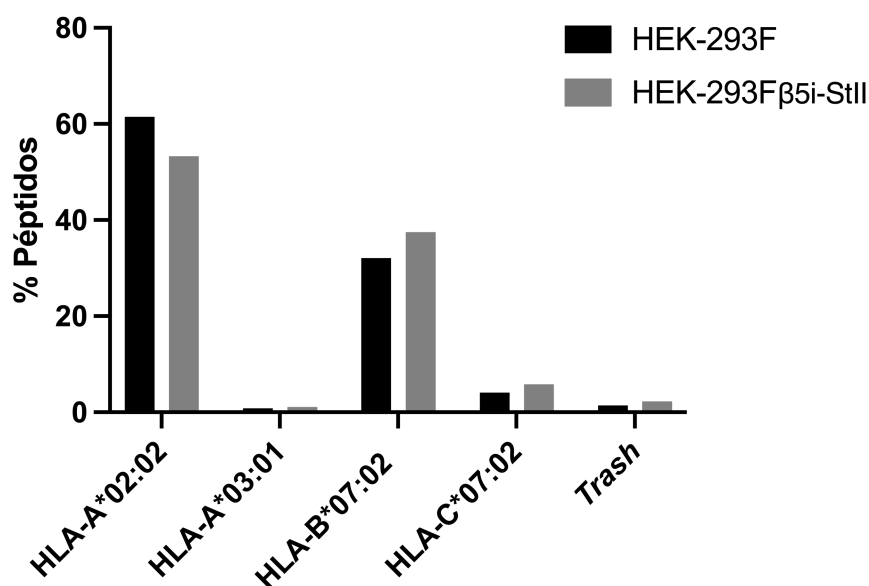
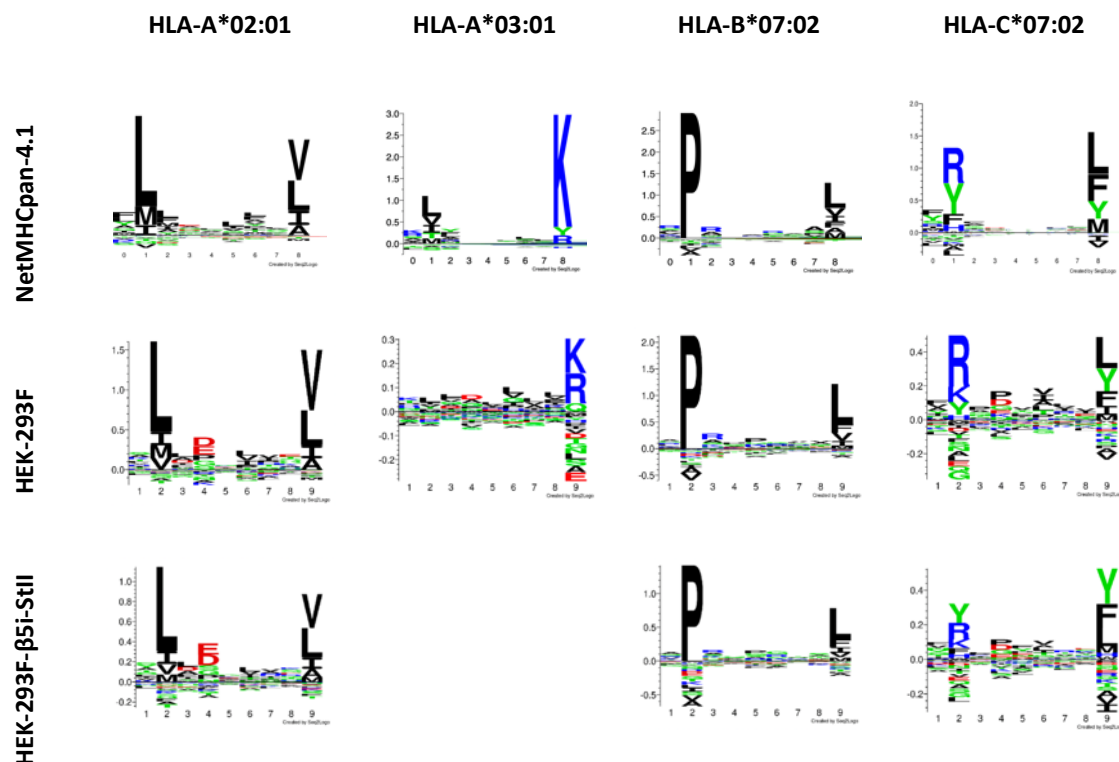


Figura 13: Asignación alélica de los ligandos de HLA-I en las líneas celulares HEK-293F y HEK-293F-β5i-StII según su afinidad teórica a través del *software* MHCMotifDecon - 1.0

Los motivos de unión que se generaron con el *software* MHCMotifDecon-1.0 eran consistentes con los motivos de unión disponibles en NetMHCpan-4.1<sup>193</sup> (Figura 14), lo que refuerza la fiabilidad de los resultados.

Las posiciones de anclaje de las moléculas de HLA-I presentes en esta línea celular son P2 y PΩ. Mientras que los motivos obtenidos en HLA-A\*02:01 y -B\*07:02 eran muy similares a los definidos por el *software*, en -C\*07:02, el motivo obtenido era un poco más laxo y en -A\*03:01 solo consideraba la posición PΩ. Esto es debido al bajo número de ligandos para -C\*07:02 y la probable ausencia de péptidos reales de A\*03:01, ya posiblemente sean contaminantes con residuos básicos en la posición C-terminal (Figura 14).



**Figura 14:** Logo de secuencias de los péptidos eluidos de HEK-293F y HEK-293F-β5i-StII y disponibles en NetMHCpan4.1. Los logos de secuencia se generaron con Seq2Logo. El logo de secuencias de los péptidos de HEK-293F-β5i-StII asociados a HLA-A\*03:01 no se muestra por la ausencia de éstos.

De manera adicional, los residuos de anclaje de los péptidos eluidos se identificaron con el software PRBAM (*Peptide-Based Anchor Motif program*)<sup>195</sup> desarrollado previamente por nuestro grupo. Esta herramienta identifica los motivos de anclaje de moléculas HLA-I y HLA-DR basándose en la abundancia de aminoácidos en cada posición y dando un valor de la desviación respecto a la media en el proteoma (DMP) para cada aminoácido en cada posición del péptido. Para simplificar los resultados, y dado que la mayoría de los péptidos identificados eran nonámeros, solo se consideraron estos para identificar los residuos de anclaje. Los residuos de anclaje se identificaron según los alelos de los que los péptidos eran ligandos.

En ambos modelos, HEK-293F y HEK-293F-β5i-StII, los ligandos de HLA-A\*02:01 presentaron preferencia de Leu, Ile y Met en P2 y de Val, Leu e Ile en PΩ. Los ligandos de -B\*07:02 estaban enriquecidos en Pro2 y Leu en PΩ. Los ligandos de -C\*07:02 estaban enriquecidos en Arg y Tyr en P2 y en Tyr y Phe en PΩ. Finalmente, los residuos de anclaje de los escasos ligandos de -A\*03:01 mostraron que estos eran probablemente contaminantes ([Anexo 2](#) y [Anexo 3](#)).

### 2.3. Afinidad teórica de los péptidos presentes en el inmunopeptidoma de HLA-I

Posteriormente, se calculó la afinidad teórica de los péptidos que componían ambos inmunopeptidomas con el *software* NetMHCpan-4.1<sup>193</sup>.

El análisis de la afinidad teórica de los péptidos eluidos en ambos modelos reveló diferencias en la presentación antigénica. Tanto en el inmunopeptidoma de HEK-293F como en el de HEK-293F- $\beta$ 5i-StII se observó que los péptidos presentaban alta afinidad por los alelos -A\*02:01 y -B\*07:02, muy similares entre ambos alelos, seguida de -C\*07:02, sin embargo, la afinidad por el alelo -A\*03:01 fue baja en ambos (Figura 15).

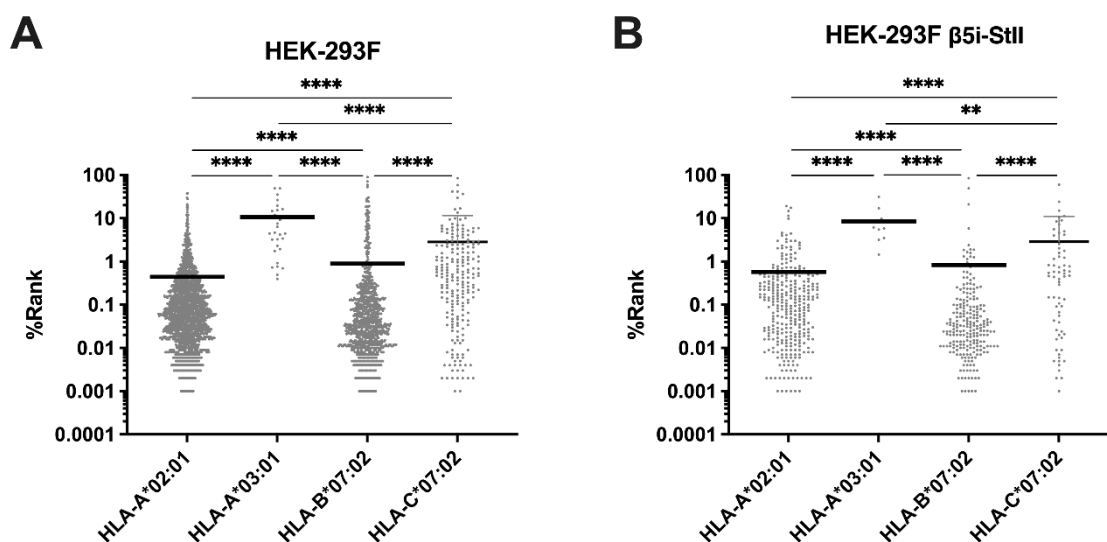
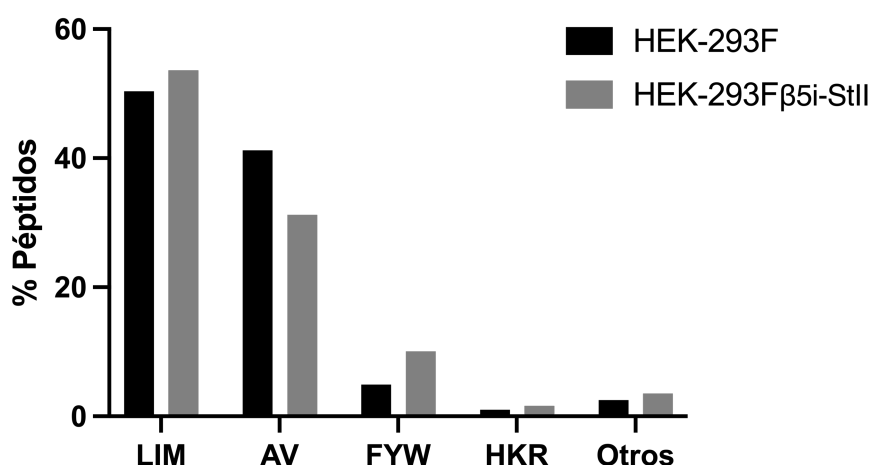


Figura 15: Afinidad teórica de los péptidos asignados a cada alotipo según el *software* NetMHCpan-4.1. (A) HEK-293F y (B) HEK-293F- $\beta$ 5i-StII. Los símbolos representan péptidos individuales; las líneas horizontales representan la media \* =  $0.01 < P < 0.1$ ; \*\* =  $0.001 < P < 0.01$ ; \*\*\* =  $0.0001 < P < 0.001$ ; \*\*\*\* =  $P < 0.0001$  según el test de Mann-Whitney U.

## 2.4. Análisis del residuo C-terminal de los péptidos presentes en el inmunopeptidoma de HLA-I

Existe un consenso que dictamina que el proteasoma genera los extremos C-terminales de los péptidos ligandos de HLA-I<sup>45</sup>. Los residuos C-terminales de los péptidos eluidos se agruparon según las características fisicoquímicas de los aminoácidos en esta posición en 5 grupos: residuos hidrofóbicos grandes (Leu, Ile y Met), residuos alifáticos pequeños (Ala y Val), residuos aromáticos (Trp, Tyr y Phe), residuos básicos (Arg, Lys y His) y otros.

El análisis de estos ligandos de clase I, tanto de células que expresan el proteasoma constitutivo, como de células que expresan el SIP, reveló una ausencia casi total de péptidos con extremo C-terminal básico: 1,0% y 1,6%, respectivamente. Asimismo, se observó una diferencia significativa en la proporción de péptidos con residuos C-terminales alifáticos pequeños y residuos aromáticos: mientras que el inmunopeptidoma de HEK-293F presentaba un 41,2% de péptidos acabados en Ala o Val, este porcentaje disminuía hasta un 31,2% en HEK-293F- $\beta$ 5i-StII. Por otro lado, el porcentaje de péptidos acabados en residuos aromáticos era superior en estas últimas, pasando de un 4,9% en HEK-293F a un 10,1% en HEK-293F- $\beta$ 5i-StII (**Figura 16**).



**Figura 16:** Residuo en posición C-terminal de los ligandos de HLA-I en las líneas celulares HEK-293F y HEK-293F- $\beta$ 5i-StII. Se muestra el porcentaje de cada grupo de residuos agrupados según sus propiedades fisicoquímicas.

Tal y como ya se ha mencionado, la línea celular HEK-293F es una adaptación de la línea celular HEK-293 al crecimiento en suspensión y cultivo sin suplemento de FBS. Al igual que la línea HEK-293, la utilización de la línea HEK-293F está muy extendida por ser fácilmente transfectable y está especialmente indicada cuando los experimentos requieren grandes volúmenes de células, ya que su crecimiento en suspensión permite obtener grandes cantidades de células en menor tiempo<sup>155</sup>. A pesar de que el tipaje previo de ambas líneas, HEK-293 y HEK-293F determinó que el genoma contenía los alelos de clase I -A\*02:01, -A\*03:01, -B\*07:02 y C\*07:02, el análisis del inmunopeptidoma de HLA-I de la línea HEK-293F detectó la ausencia de péptidos básicos en PΩ (**Figura 16**). Estos péptidos son presentados en superficie por la molécula HLA-A\*03:01<sup>153</sup>.

### **3. ANÁLISIS DE LA EXPRESIÓN DE HLA-A\*03:01 EN LAS LINEAS HEK-293 Y HEK-293F**

Dados los resultados descritos en los apartados previos, se decidió analizar la presencia de la molécula HLA-A\*03 en superficie, tanto en la línea HEK-293 como HEK-293F, mediante citometría de flujo y, posteriormente, la expresión de este alelo a nivel génico por PCR cuantitativa.

#### **3.1. Análisis de la expresión en membrana de las moléculas HLA-A\*02, -A\*03 y -B\*07 en las líneas celulares HEK-293 y HEK-293F por citometría de flujo**

Se realizó una tinción superficial con los anticuerpos W6/32, HC.10 y B8.11.2 con el fin de evaluar la expresión de las moléculas de HLA-I, HLA-I vacías y HLA-DR, respectivamente, en la línea parental HEK-293 y HEK-293F. Ambas líneas eran positivas para la tinción con el anticuerpo W6/32, ligeramente positivas con el anticuerpo HC.10 y prácticamente negativas para la tinción con el anticuerpo B8.11.2, lo que demuestra que la expresión de las moléculas de HLA en membrana de ambas líneas es principalmente de HLA-I en configuración estable (**Figura 17**).



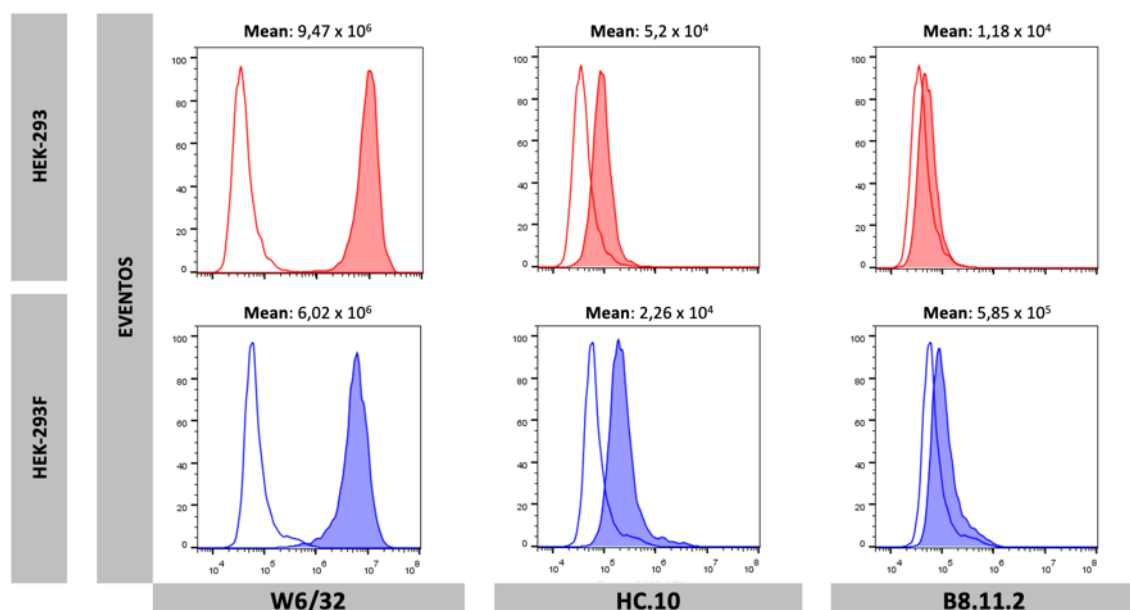


Figura 17: Análisis por citometría de flujo de las moléculas HLA-I estables (W6/32), HLA-I vacías (HC.10) y HLA-DR (B8.11.2) en las líneas celulares HEK-293 (rojo) y HEK-293F (azul). Los histogramas de fondo blanco representan el control de segundo anticuerpo. La **Mean** hace referencia a la media de intensidad de fluorescencia asociada a cada marcador.

Posteriormente se evaluó la expresión de los diferentes alelos de clase I detectados en el tipaje, excepto *HLA-C\*07:02*, utilizando anticuerpos específicos para dichos alelos: A\*02 (PA2.1), A\*03 (GAP-A3), y B\*07 (ME.1) (Figura 18).

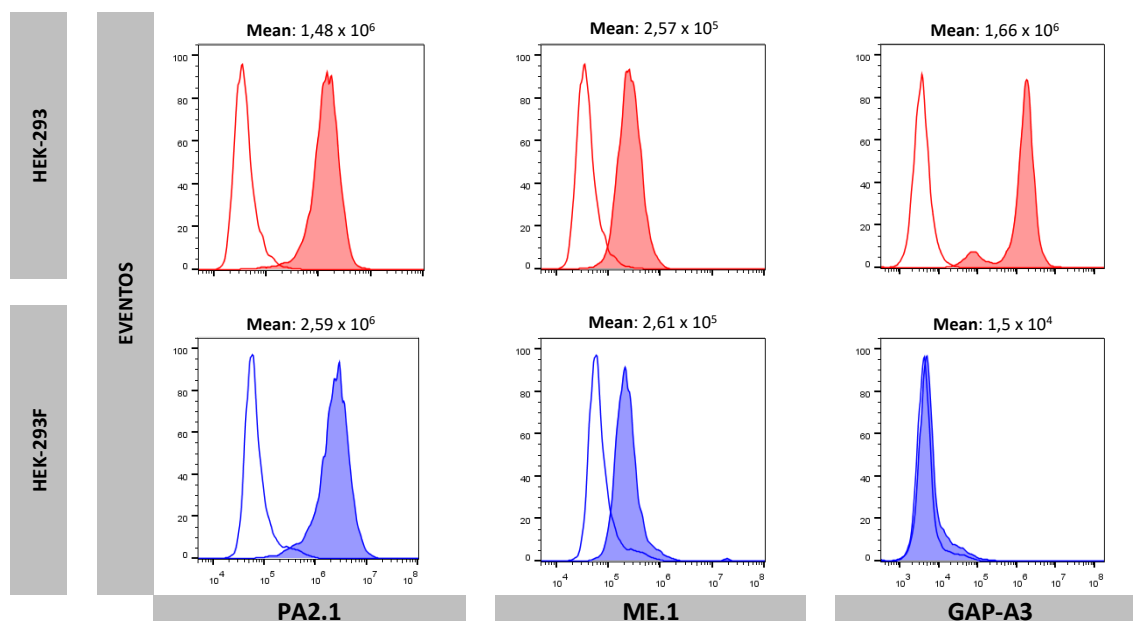


Figura 18: Expresión superficial de los diferentes alelos de HLA de clase I en células HEK-293 y HEK-293F. Análisis por citometría de flujo de las moléculas HLA-A\*02 (PA2.1), HLA-A\*03 (GAP-A3) y HLA-B\*07 (ME.1) en las líneas celulares HEK-293 (rojo) y HEK-293F (azul). Los histogramas de fondo blanco representan el control de segundo anticuerpo. La **Mean** hace referencia a la media de intensidad de fluorescencia asociada a cada marcador.

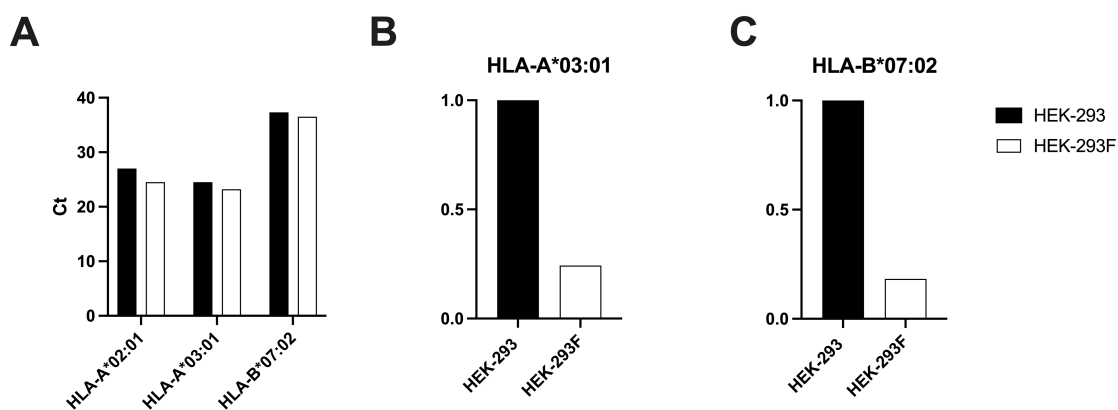
La tinción con el anticuerpo PA2.1 ( $\alpha$ -HLA-A\*02) reveló que ambas líneas celulares expresaban moléculas de HLA-A\*02 en superficie. Asimismo, la tinción con el anticuerpo ME.1 ( $\alpha$ -HLA-B\*07) reveló que ambas líneas también expresaban moléculas de HLA-B\*07, aparentemente en menor proporción. En cuanto al anticuerpo GAP-A3, específico de las moléculas de clase I HLA-A\*03, se observó que este alelo se expresaba en las células HEK-293, sin embargo, había una ausencia total en las células HEK-293F (**Figura 18**).

### **3.2. Análisis de la expresión génica de las moléculas HLA-A\*02:01, -A\*03:01 y -B\*07:02 en las líneas celulares HEK-293 y HEK-293F por PCR cuantitativa**

Dados los resultados, se decidió evaluar la expresión de los alelos *HLA-A\*02:01*, *-A\*03:01* y *-B\*07:02* a nivel transcripcional mediante qPCR para analizar si la ausencia de HLA-A\*03:01 se debía a falta de transcripción génica o a eventos posteriores a la generación de mRNA. Para ello, se amplificaron los genes de *HLA-A\*02:01*, *HLA-A\*03:01* y *HLA-B\*07:02* con *primers* específicos mediante qPCR.

En ambas líneas celulares, la expresión más elevada de mRNA correspondió a HLA-A\*03:01, ya que la Ct era menor que la correspondiente a -A\*02:01 y, sobre todo, a -B\*07:02 (**Figura 19A**).

La normalización de la expresión de mRNA de -A\*03:01 y -B\*07:02 frente a la expresión de -A\*02:01 en ambas líneas celulares, mostró que la expresión de A\*03:01 en HEK-293F correspondía a 0,24 (24,3%) veces la expresión de A\*03:01 en HEK-293 (**Figura 19B**), mientras que la expresión de HLA-B\*07:02 era de 0,18 (18,3%) veces en HEK-293F respecto HEK-293 (**Figura 19C**). Por tanto, la cuantificación por qPCR mostró una disminución relativa del mRNA de HLA-A\*03:01 presente en la célula, pero que no explica la ausencia de esta proteína en la superficie celular.

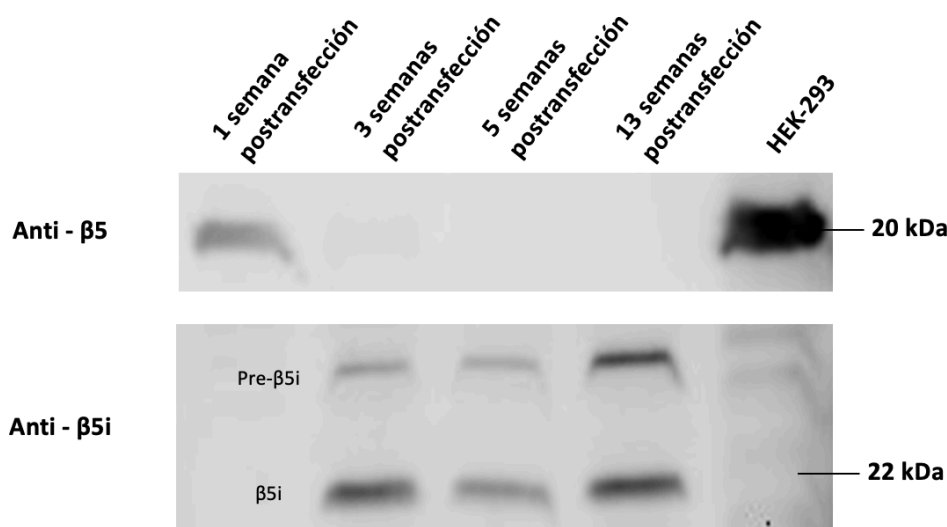


**Figura 19: (A):** Cuantificación por qPCR de los niveles de mRNA de alelos A\*02:01, A\*03:01 y B\*07:02 en células HEK-293 y HEK-293F. Se tomó el Ct (número de ciclos necesario para amplificar un producto génico hasta llegar a la pendiente requerida) como medida de abundancia. Menor Ct indica mayor expresión a nivel génico. **(B):** Expresión relativa de HLA-A\*03:01 respecto de la expresión de HLA-A\*02:01 en las células HEK-293 y en las células HEK-293F. **(C)** Expresión relativa de HLA-B\*07:02 respecto de la expresión de HLA-A\*02:01 en las células HEK-293 y en las células HEK-293F. Los datos muestran la expresión de los dos alelos de clase I en relación con el gen de *HLA-A\*02:01*.

Por tanto, estos datos muestran que, aunque hay expresión a nivel de mRNA de las moléculas de HLA-A\*03:01 y HLA-B\*07:02 en la línea celular HEK-293F, tal y como se detectó por qPCR, no hay expresión a nivel de proteína de HLA-A\*03, razón por la cual no se observaron péptidos ligados de HLA-A\*03:01 en experimentos previos de peptidómica y que, por este motivo, la línea celular HEK-293F no es equivalente a la línea HEK-293 para analizar el inmunopeptidoma de HLA-I y para futuros estudios funcionales.

#### 4. ANÁLISIS DEL REPERTORIO PEPTÍDICO DE HEK-293 Y HEK-293- $\beta$ 5i

El hecho de que la ausencia de ligandos de HLA-A\*03:01 se detectó en la línea HEK-293F y en el transfectante de la subunidad  $\beta$ 5i sugiere que la ausencia de ligandos de HLA-A\*03:01 es debida a una característica intrínseca de la línea celular y no a un artefacto de la transfección. Por tanto, se decidió generar un nuevo modelo celular de proteasoma intermedio de tipo I en la línea celular **HEK-293** siguiendo la misma metodología previamente descrita. Brevemente, la subunidad  $\beta$ 5i del proteasoma se transfectó sobre la línea celular HEK-293 mediante un protocolo estándar de transfección con PEI y selección con G418. La expresión residual de  $\beta$ 5 y la expresión de  $\beta$ 5i se analizó por Western Blot con anticuerpos específicos anti- $\beta$ 5 y anti- $\beta$ 5i durante diversas semanas hasta que la subunidad  $\beta$ 5i desplazó por completo la subunidad constitutiva  $\beta$ 5 (Figura 20).



**Figura 20: Expresión de las subunidades  $\beta$ 5 y  $\beta$ 5i en el transfectante HEK-293- $\beta$ 5i.** Se analizaron por Western Blot los lisados celulares de la línea parental HEK-293 y del transfectante HEK-293- $\beta$ 5i a diferentes tiempos post-transfección. Se cargaron 10  $\mu$ g de lisado total en un gel de acrilamida al 14% y se reveló con los anticuerpos primarios anti- $\beta$ 5 y anti- $\beta$ 5i. En el caso de la subunidad  $\beta$ 5i, se observan dos bandas correspondientes a la subunidad procesada y sin procesar.

La purificación de los inmunopeptidomas de HLA-I se llevó a cabo siguiendo el mismo protocolo previamente descrito. Se purificaron un total de 4493 ligandos peptídicos de HLA-I en la línea celular HEK-293 y 2937 péptidos en el transfectante HEK-293-β5i, tras aplicar una FDR del 1%. Se tuvieron en cuenta los péptidos de 8-13 residuos de longitud y no se tuvieron en cuenta las modificaciones postraduccionales, dando lugar a 4175 y 2705 péptidos de HEK-293 y HEK-293 β5i, respectivamente, de los cuales 1429 (26,2%) eran compartidos por ambos peptidomas. La distribución por longitud siguió el mismo patrón previamente observado: los péptidos de 9 residuos de longitud eran mayoritarios, seguidos de decámeros y péptidos de 11, 8, 12 y 13 residuos de longitud (**Figura 21**). El número medio de aminoácidos fue de entre 9,16 y 9,08 (**Tabla 5**). No se apreciaron diferencias significativas.

|                  | Total | Péptidos por número de residuos (%) |      |      |      |     |     | AAs promedio |
|------------------|-------|-------------------------------------|------|------|------|-----|-----|--------------|
|                  |       | 8                                   | 9    | 10   | 11   | 12  | 13  |              |
| HEK-293          | 4175  | 5,9                                 | 60,5 | 17,6 | 10,1 | 2,0 | 1,0 | 9,16         |
| HEK-293 β5i-StII | 2705  | 5,9                                 | 59,0 | 18,6 | 9,2  | 2,4 | 1,1 | 9,08         |

Tabla 5: Tamaño, por número de aminoácidos, de los inmunopeptidomas de HLA-I en las líneas celulares HEK-293 y HEK-293-β5i.

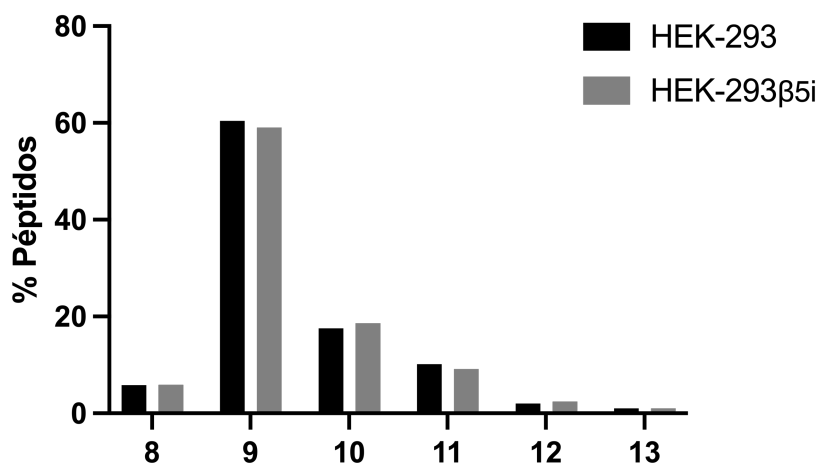
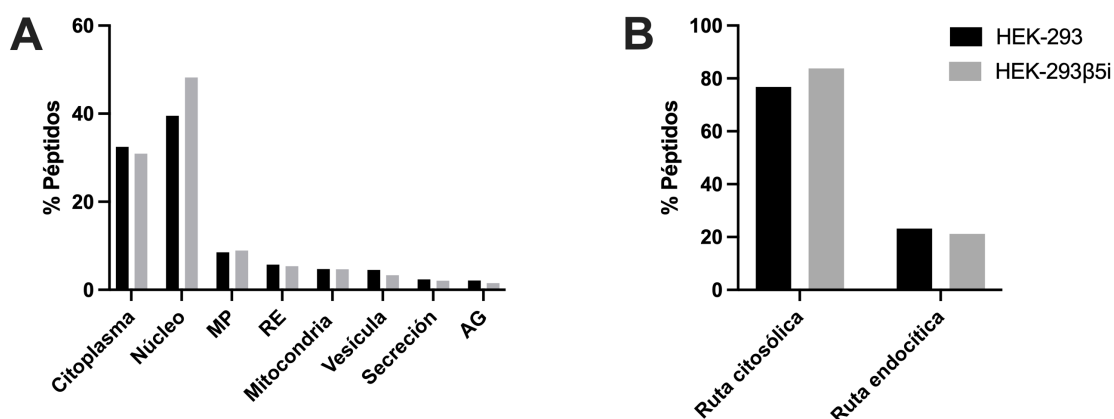


Figura 21: Distribución por tamaño de los péptidos asociados a HLA-I según el número de aminoácidos en las líneas HEK-293 y HEK-293-β5i. La mayoría de los péptidos eran nonámeros, seguidos de decámeros, octámeros y péptidos de 11, 12 y 13 residuos de longitud.

#### 4.1. Localización celular

La localización celular de las proteínas parentales se obtuvo consultando la base de datos UniProt, priorizando el compartimento principal según su función cuando aparecían en más de un compartimento. Los péptidos eluidos en las líneas celulares HEK-293 y HEK-293- $\beta$ 5i provenían de 2665 proteínas y 1628 proteínas, respectivamente, localizadas principalmente en el núcleo y el citosol (**Figura 22A**). Al igual que los peptidomas en HEK-293F, la mayoría de los péptidos, entre 77-83%, derivaron de la ruta endógena (citosol, núcleo y mitocondrias), mientras que el 23-21% correspondió a la ruta endocítica (membrana plasmática, retículo endoplasmático, aparato de Golgi, vesículas y proteínas secretadas) (**Figura 22B**). No se apreciaron diferencias significativas en la localización celular de los inmuno-peptidomas de HEK-293 y HEK-293F parentales o con la subunidad  $\beta$ 5i.



**Figura 22:** Localización celular de las proteínas parentales de los ligandos de HLA-I en las líneas celulares HEK-293 y HEK-293- $\beta$ 5i. **(A)** Porcentaje de péptidos localizados en cada compartimento celular: MP, membrana plasmática; RE, retículo endoplasmático; AG, aparato de Golgi. **(B)** Porcentaje de péptidos pertenecientes a la ruta citosólica (citoplasma, núcleo y mitocondrias) y endocítica (MP, RE, AG, vesículas y proteínas de secreción).

#### 4.2. Asignación alélica de los péptidos presentes en el inmunopeptidoma de HLA-I

La asignación alélica se realizó con el *software* MHCMotifDecon – 1.0. En el inmunopeptidoma de HEK-293 se identificaron 1518 (36,4%) ligandos de HLA-A\*02:01, 812 (19,4%) de -A\*03:01, 1536 (36,8%) de -B\*07:02 y 247 (5,9%) de -C\*07:02. Se identificaron 62 (1,5%) péptidos como *trash*. Por otro lado, en el inmunopeptidoma de HEK-293  $\beta$ 5i se identificaron 691 (25,5%) ligandos de HLA-A\*02:01, 742 (27,4%) de -A\*03:01, 1060 (39,2%) de -B\*07:02 y 199 (7,4%) de -C\*07:02. Se identificaron 13 (0,5%) péptidos como *trash* (**Figura 23**). Por tanto, los inmunopeptidomas de clase I de ambas líneas celulares, a diferencia de lo que se observó en la línea HEK-293F, contenían ligandos peptídicos de HLA-A\*03:01, así como de -A\*02:01, -B\*07:02 y -C\*07:02.

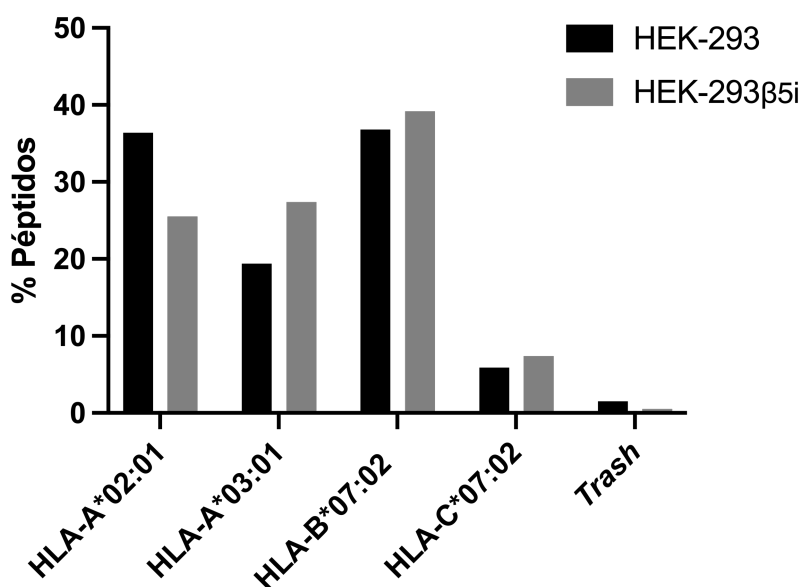


Figura 23: Asignación alélica de los ligandos de HLA-I en las líneas celulares HEK-293 y HEK-293- $\beta$ 5i según su afinidad teórica a través del *software* MHCMotifDecon - 1.0

Por su parte, el análisis de los motivos de anclaje mostró que la posición *P2* de los ligandos de HLA-A\*02:01 estaba muy conservada en Leu y la *PΩ* en Val y Leu. La *PΩ* de los ligandos de HLA-A\*03:01 estaba muy conservada en Lys. Los ligandos de HLA-B\*07:02 presentaron la *P2* muy conservada en Pro, así como la *PΩ* en Leu. Los ligandos de HLA-C\*07:02 presentaron *P2* muy conservada en residuos básicos y la *PΩ* en Tyr, Phe y Ile. De nuevo, los motivos de unión que se generaron con el *software* MHCMotifDecon-1.0 eran consistentes con los motivos de unión disponibles en NetMHCpan-4.1, reforzando la fiabilidad de los resultados (Figura 24).

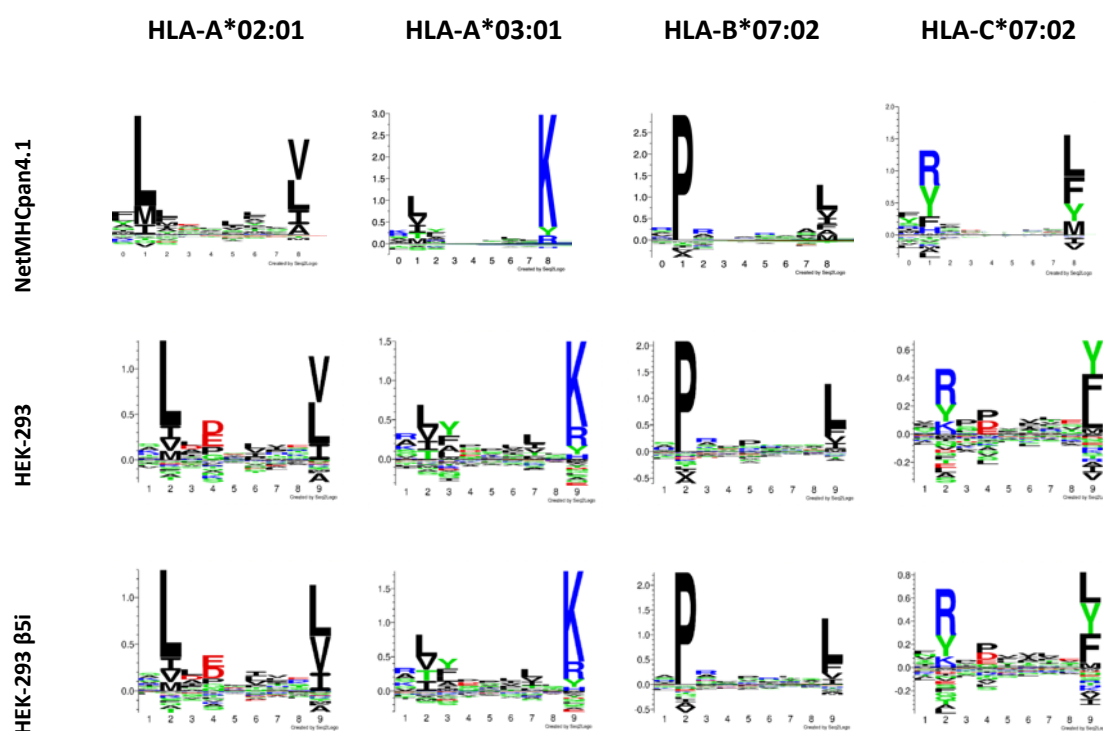


Figura 24: Logo de secuencias de los péptidos eluidos de HEK-293 y HEK-293-β5i y disponibles en NetMHCpan4.1. Los logos de secuencia se generaron con Seq2Logo.



Asimismo, los residuos de anclaje se identificaron con el *software* PRBAM. Tal y como se hizo con la línea HEK-293F, solo se consideraron los nonámeros y según el alelo al que se unía cada péptido. En ambos inmunopeptidomas analizados se observó enriquecimiento de Leu, Ile y Met en P2 y Val, Leu e Ile en PΩ de los ligandos de HLA-A\*02:01. Asimismo, se observó un enriquecimiento de Val, Leu e Ile en P2 y residuos básicos en los ligandos de HLA-A\*03:01. Los ligandos de HLA-B\*07:02 estaban enriquecidos en Pro2 y Leu en PΩ, además, los péptidos ligandos de este alelo del inmunopeptidoma de HEK-293 β5i mostraron un enriquecimiento de Phe en PΩ. Finalmente, los ligandos de HLA-C\*07:02 estaban enriquecidos en Arg y Tyr en P2 y Lys, Phe y Tyr en PΩ ([Anexo 4](#) y [Anexo 5](#)).

#### **4.3. Afinidad teórica de los péptidos presentes en el inmunopeptidoma de HLA-I**

Al igual que para las células HEK-293F y HEK-293F-β5i-StII, la afinidad teórica se calculó para los alelos HLA-A\*02:01, -A\*03:01, -B\*07:02 y -C\*07:02 con el *software* NetMHCpan-4.1. La [Figura 25](#) muestra que la afinidad teórica de los péptidos asignados a cada alelo era similar entre los diferentes alotipos. La diferencia de afinidad entre los péptidos de HLA-A\*02:01 y -B\*07:02 con los de HLA-A\*03:01 se redujo en comparación con el análisis en las células HEK-293F y HEK-293F-β5i-StII. Por tanto, los ligandos asignados a HLA-A\*03:01 en los inmunopeptidomas de las células HEK-293 y HEK-293 β5i-StII son de alta afinidad.

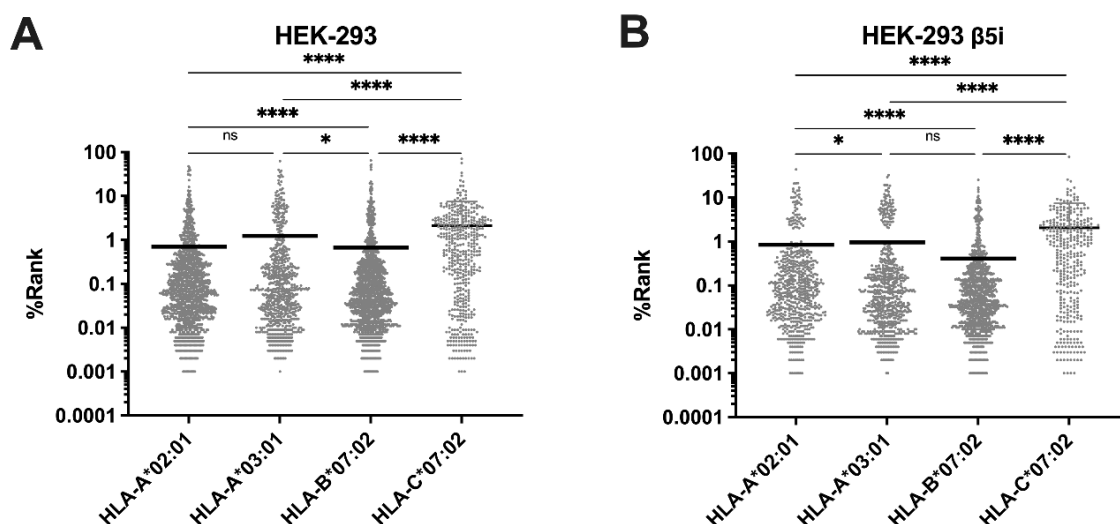


Figura 25: Afinidad teórica de los péptidos asignados a cada alotipo según el software NetMHCpan-4.1. (A) HEK-293 y (B) HEK-293-β5i. Los símbolos representan un péptido individual y las líneas horizontales representan la mediana de afinidad. Los símbolos representan péptidos individuales; las líneas horizontales representan la media \* =  $0.01 < P < 0.1$ ; \*\* =  $0.001 < P < 0.01$ ; \*\*\* =  $0.0001 < P < 0.001$ ; \*\*\*\* =  $P < 0.0001$  según el test de Mann-Whitney U.

#### 4.4. Análisis del residuo C-terminal de los péptidos presentes en el inmunopeptidoma de HLA-I

El análisis del inmunopeptidoma de las líneas HEK-293 y HEK-293-β5i, reveló que no existía ausencia de péptidos con extremo C-terminal básico, tal y como se había observado en la línea HEK-293F. Sin embargo, se observó un mayor porcentaje de péptidos con residuos básicos en C-terminal en las células con expresión del SIP ya que estos representaban un 18,4% en HEK-293 y un 26,3% en HEK-293 β5i.

Por otro lado, se confirmaron los datos observados en las células HEK-293F y HEK-293F-β5i en referencia a los residuos hidrofóbicos: el proteasoma constitutivo generó un repertorio enriquecido en péptidos con residuos hidrofóbicos pequeños en C-terminal respecto al SIP, 22,1% y 11,9%, respectivamente. El porcentaje de péptidos con residuos C-terminal aromático fue superior en células que expresaban el SIP, un 11,1% frente al 8,1% en las células que expresaban el proteasoma constitutivo (Figura 26).

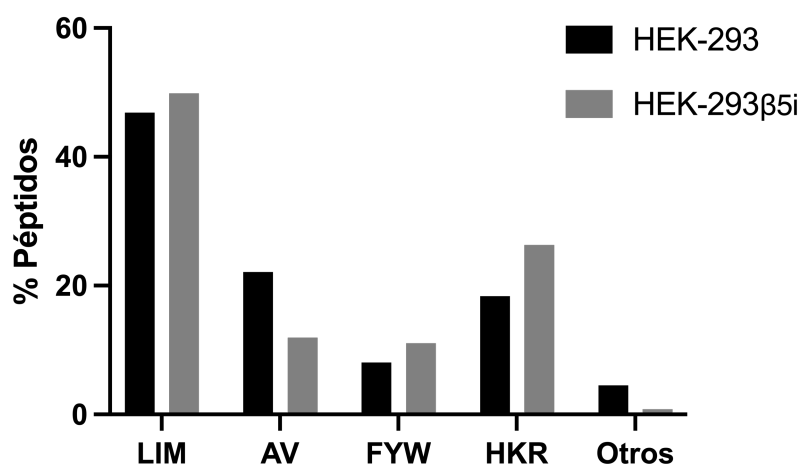
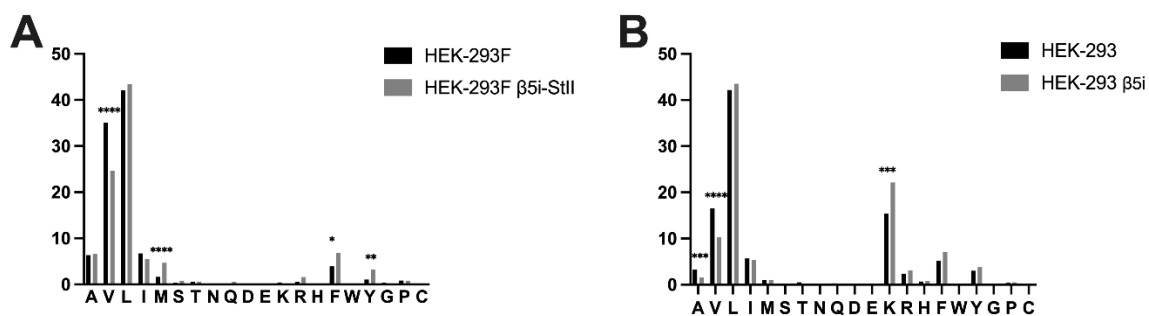


Figura 26: Residuo en posición C-terminal de los ligandos de HLA-I en las líneas celulares HEK-293 y HEK-293-β5i. Se muestra el porcentaje de cada grupo de residuos agrupados según sus propiedades fisicoquímicas.

Dados los resultados obtenidos en el análisis de los residuos localizados en posición C-terminal tanto en la línea HEK-293 como HEK-293F, se analizaron los residuos localizados en esta posición de manera individual, y se evaluó si existían diferencias significativas entre la línea parental y la línea transfectada. La **Figura 27A** muestra que los péptidos acabados en Val estaban significativamente incrementados en la línea parental respecto al transfectante HEK-293F-β5i-StII, mientras que los péptidos acabados en Met, Phe y Tyr estaban significativamente incrementados en el transfectante respecto la línea parental. El mismo fenómeno se observó cuando se compararon los residuos C-terminales de la línea adherente HEK-293 y el transfectante de β5i sobre la misma. Los péptidos acabados en Ala y Val estaban significativamente incrementados en la línea parental, mientras que el transfectante de HEK-293-β5i mostró un incremento significativo de residuos básicos, como Lys (**Figura 27B**). Las diferencias de los residuos básicos no se pudieron analizar en la línea 293F porque no se detectaron residuos C-terminales básicos al no expresar HLA-A\*03:01.



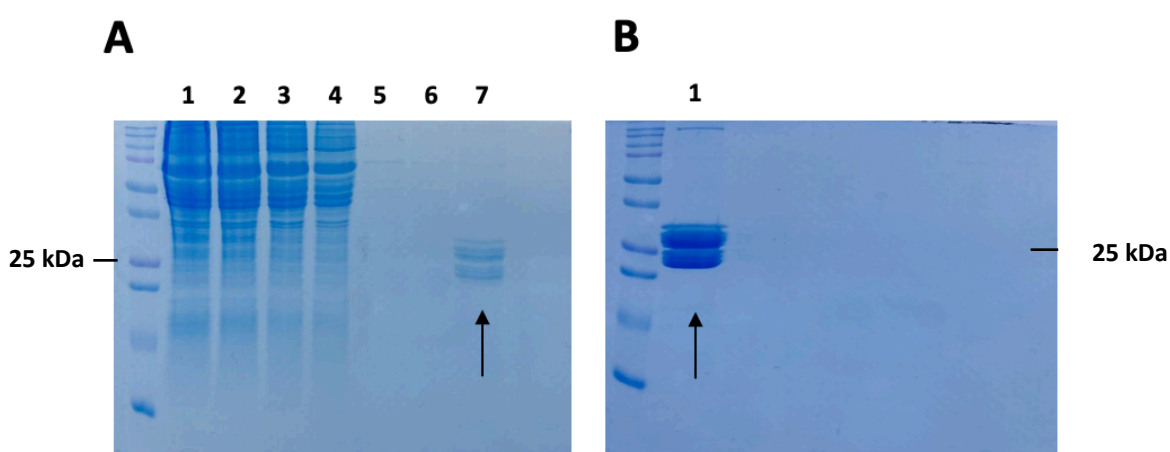
**Figura 27: Residuos localizados en posición C-terminal de los ligandos de HLA-I en (A) HEK-293F y HEK-293F-β5i-StII y (B) HEK-293 y HEK-293-β5i.** Se realizó un test chi cuadrado con la corrección de Bonferroni para evidenciar las diferencias entre la línea parental y transfectada con la subunidad β5i. \* =  $0.01 < P < 0.1$ ; \*\* =  $0.001 < P < 0.01$ ; \*\*\* =  $0.001 < P < 0.01$ ; \*\*\*\* =  $P < 0.0001$

Estos resultados muestran claramente uno de los aspectos propuestos en la hipótesis, el proteasoma intermedio de tipo I, que contiene la subunidad β5i, genera un inmunopeptidoma asociado a HLA-I diferente al del proteasoma constitutivo. Esto sugiere que dichos proteasomas presentan unas especificidades catalítica diferenciadas, siendo más eficiente catalizando el enlace tras residuos hidrofóbicos grandes y básicos en el caso del proteasoma intermedio, mientras que la subunidad β5 es más eficiente en su corte tras residuos hidrofóbicos pequeños.

## 5. PURIFICACIÓN DEL PROTEASOMA 20S

Con el fin de dilucidar si las diferencias en los inmunopeptidomas de clase I se debían a diferencias en la actividad catalítica entre los dos proteasomas, se evaluó la misma con digestiones *in vitro* de péptidos fluorogénicos con proteasoma 20S purificado. Además, para dilucidar si la estructura del complejo catalítico pudiera explicar las diferencias en los repertorios, se planteó abordar el estudio de la estructura tridimensional del proteasoma 20S intermedio de tipo I.

Para ello, tanto el proteasoma constitutivo, como el SIP, se purificaron a partir de pellets secos de **HEK-293F- $\beta$ 5-StII** y **HEK-293F- $\beta$ 5i-StII**, respectivamente. Se purificaron mediante cromatografía de afinidad por el “tag” Strep-tag II. La pureza de los proteasomas purificados se analizó mediante un gel SDS-PAGE 14%, lo que mostró un enriquecimiento significativo del complejo (**Figura 28A**). Sin embargo, se detectaron varias bandas correspondientes a proteínas contaminantes. Para aumentar la pureza, se realizó un segundo paso de purificación que consistió en ultrafiltración de tamaño de corte de 125 kDa. De nuevo, la pureza se evaluó mediante un gel SDS-PAGE 14% mostrando unos proteasomas purificados de alta pureza sin que se detectaran bandas contaminantes significativas por tinción con *Coomassie blue* (**Figura 28B**).



**Figura 28: Purificación del proteasoma 20S intermedio de tipo I. (A) Purificación con columna de afinidad por StII.** Carril 1: lisado total, carril 2: paso 2 de la purificación, carril 3: paso 3 de la purificación, carril 4: flow through de la columna de afinidad, carril 5: lavado 1, carril 6: lavado 2, carril 7: eluido de la columna de afinidad. **(B) Filtración por tamaño (125kDa).** Carril 1: eluido del centricon de 125 kDa.

## 6. EVALUACIÓN DE LA ACTIVIDAD CATALÍTICA

La actividad catalítica del proteasoma intermedio de tipo I se evaluó mediante un ensayo enzimático con péptidos fluorogénicos bajo condiciones controladas de temperatura.

Como se ha señalado previamente, el proteasoma presenta tres actividades catalíticas, cada una asociada a una subunidad catalítica distinta: caspasa ( $\beta 1$ ), tripsina ( $\beta 2$ ) y quimotripsina ( $\beta 5$ ). Las actividades caspasa y tripsina se evaluaron con dos péptidos fluorogénicos distintos:

- Actividad caspasa: **Z-LLE-AMC**
- Actividad tripsina: **Boc-LRR-AMC**

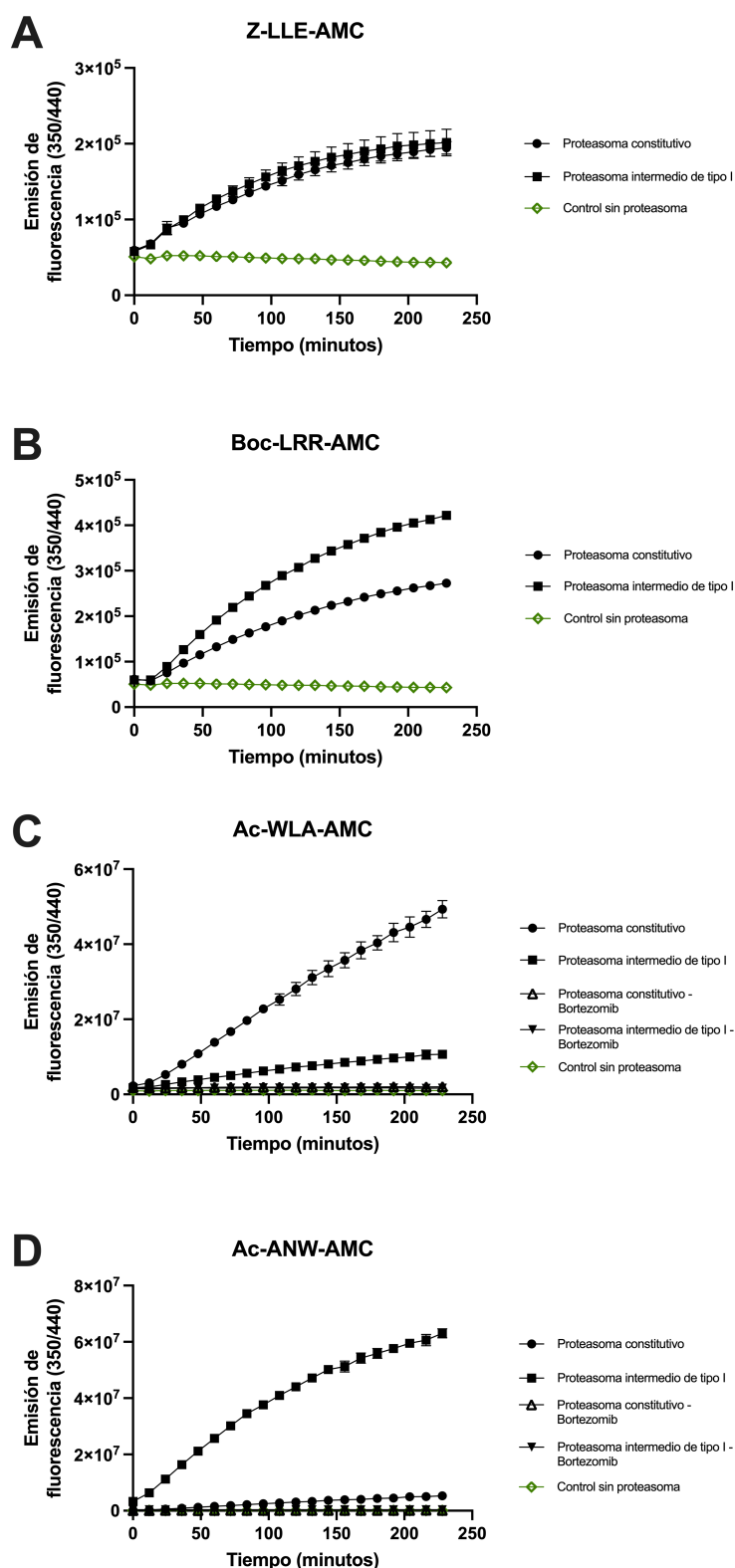
Puesto que el proteasoma constitutivo y el SIP se diferencian en la subunidad  $\beta 5/\beta 5i$ , asociada a la actividad quimotripsina, y en base a los resultados del análisis de los inmunopeptidomas previo, se seleccionaron específicamente dos péptidos fluorogénicos para reflejar mejor las características observadas en los repertorios:

- Un péptido con un residuo hidrofóbico pequeño (Ala) en el extremo C-terminal: **Ac-WLA-AMC**
- Un péptido con un residuo aromático (Trp) en el extremo C-terminal: **Ac-ANW-AMC**

Adicionalmente, se incluyó Bortezomib, un inhibidor selectivo del proteasoma, para confirmar la especificidad de los resultados obtenidos. Este ensayo se realizó por triplicado y en tres días distintos utilizando diferentes proteasomas purificados para asegurar la especificidad de los resultados ([Anexo 6](#)).

Los resultados obtenidos mostraron que la actividad caspasa ( $\beta 1$ ) era similar en ambos complejos ([Figura 29A](#)). En contraste, la actividad triptica ( $\beta 2$ ) se encontró incrementada en el SIP en comparación con el proteasoma constitutivo ([Figura 29B](#)). Además, el análisis de la hidrólisis de péptidos fluorogénicos hidrofóbicos reveló diferencias claras en la actividad quimotriptica entre ambos complejos. Así, el sustrato que presentaba Ala en posición C-terminal fue hidrolizado de manera más eficiente por el proteasoma constitutivo ([Figura 29C](#)). Por el contrario, el péptido con Trp en posición C-terminal

mostró ser hidrolizado más eficientemente por el SIP (**Figura 29D**). Estos resultados demuestran que las diferentes subunidades catalíticas del proteasoma intermedio de tipo I confieren especificidad de sustrato diferenciada respecto al proteasoma constitutivo, que se traduce en un incremento en la actividad tipo tripsina y una clara preferencia por sustratos con residuos aromáticos en el extremo C-terminal.



**Figura 29:** Representación gráfica de un ensayo de evaluación de las actividades catalíticas del proteasoma constitutivo y del SIP al hidrolizar diferentes sustratos fluorogénicos. Media y desviación estándar del ensayo por triplicado. Se tomaron medidas de la emisión de fluorescencia (350/440nm) cada 12 minutos durante 4 horas. **A:** Evaluación de la actividad caspasa mediante la hidrólisis de un péptido fluorogénico con extremo C-terminal ácido (Glu): Z-LLE-AMC. **B:** Evaluación de la actividad triptica mediante la hidrólisis de un péptido fluorogénico con extremo C-terminal básico (Arg): Boc-LRR-AMC. **C:** Evaluación de la actividad quimotriptica mediante la hidrólisis de un péptido fluorogénico con extremo C-terminal hidrofóbico pequeño (Ala): Ac-WLA-AMC. **D:** Evaluación de la actividad quimotrica mediante la hidrólisis de un péptido fluorogénico con extremo C-terminal aromático (Trp): Ac-ANW-AMC.



La actividad catalítica se cuantificó a partir de la velocidad inicial ( $V_0$  RFU/min) de la reacción<sup>197</sup>, que permite comparar la eficiencia catalítica del proteasoma constitutivo y del SIP frente los diferentes sustratos fluorogénicos. Se calculó para los tres ensayos realizados. La velocidad relativa (%) se calculó tomando como referencia de 100% la  $V_0$  del proteasoma intermedio de tipo I (3) frente el sustrato Ac-ANW-AMC. De acuerdo con la bibliografía, el SIP y el proteasoma constitutivo presentan actividad tipo caspasa muy similar. Asimismo, se confirmó que la actividad tripsina del SIP era superior que la actividad tripsina del proteasoma constitutivo (Tabla 6). Los resultados obtenidos demostraron que el proteasoma constitutivo presentaba mayor eficiencia catalítica frente al sustrato Ac-WLA-AMC, mostrando una pendiente más pronunciada respecto el SIP. Por el contrario, este mostró una actividad catalítica mayor en la hidrólisis del sustrato Ac-ANW-AMC. Estas tendencias se observaron en los tres ensayos realizados (Tabla 6).

|             |                        | Proteasoma constitutivo (1) | Proteasoma constitutivo (2) | Proteasoma constitutivo (3) | Proteasoma intermedio de tipo I (1) | Proteasoma intermedio de tipo I (2) | Proteasoma intermedio de tipo I (3) |
|-------------|------------------------|-----------------------------|-----------------------------|-----------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Z-LLE-AMC   | $V_0$ RFU/min          | $1,3 \times 10^3$           | $1,03 \times 10^3$          | $9,8 \times 10^2$           | $2,2 \times 10^3$                   | $1,2 \times 10^3$                   | $1,9 \times 10^3$                   |
|             | Velocidad relativa (%) | 0,34                        | 0,27                        | 0,26                        | 0,58                                | 0,31                                | 0,50                                |
| Boc-LRR-AMC | $V_0$ RFU/min          | $1,1 \times 10^3$           | $1,2 \times 10^3$           | $1,3 \times 10^3$           | $1,9 \times 10^3$                   | $1,3 \times 10^3$                   | $2,4 \times 10^3$                   |
|             | Velocidad relativa (%) | 0,29                        | 0,31                        | 0,34                        | 0,48                                | 0,34                                | 0,63                                |
| Ac-WLA-AMC  | $V_0$ RFU/min          | $1,02 \times 10^5$          | $2,25 \times 10^5$          | $1,83 \times 10^5$          | $3,00 \times 10^4$                  | $1,88 \times 10^4$                  | $4,60 \times 10^4$                  |
|             | Velocidad relativa (%) | 26,70                       | 58,90                       | 47,9                        | 7,85                                | 4,92                                | 12,04                               |
| Ac-ANW-AMC  | $V_0$ RFU/min          | $8,02 \times 10^3$          | $2,62 \times 10^4$          | $2,29 \times 10^4$          | $4,36 \times 10^4$                  | $1,40 \times 10^5$                  | $3,82 \times 10^5$                  |
|             | Velocidad relativa (%) | 2,10                        | 6,86                        | 5,99                        | 11,41                               | 36,65                               | 100                                 |

**Tabla 6:**  $V_0$  (Unidades Relativa de Fluorescencia (RFU)) y velocidad relativa (%) de las actividades catalíticas del proteasoma constitutivo e intermedio de tipo I catalizando los sustratos fluorogénicos Z-LLE-AMC, Boc-LRR-AMC, Ac-WLA-AMC y Ac-ANW-AMC. La  $V_0$  se calcula como la pendiente de la fase lineal de emisión de fluorescencia. La velocidad relativa se calcula tomando como 100% la  $V_0$  del proteasoma intermedio de tipo I (3) frente el sustrato Ac-ANW-AMC. Cada valor hace referencia al promedio de triplicados. Se observa que el proteasoma intermedio de tipo I es más eficiente catalizando el péptido Boc-LRR-AMC, mientras que la catalización del péptido Z-LLE-AMC es muy similar por ambos complejos. El proteasoma constitutivo es más eficiente catalizando el péptido Ac-WLA-AMC y el proteasoma intermedio de tipo I es más eficiente catalizando el péptido Ac-ANW-AMC.

Asimismo, se calculó tanto la  $V_0$  (RFU/min), como la velocidad relativa (%) del ensayo en presencia de Bortezomib, demostrando que la actividad catalítica calculada corresponde al proteasoma (Tabla 7).

|            |                        | Proteasoma constitutivo (1) | Proteasoma constitutivo (2) | Proteasoma constitutivo (3) | Proteasoma intermedio de tipo I (1) | Proteasoma intermedio de tipo I (2) | Proteasoma intermedio de tipo I (3) |
|------------|------------------------|-----------------------------|-----------------------------|-----------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Ac-WLA-AMC | V <sub>0</sub> RFU/min | -1,48 x 10 <sup>2</sup>     | 1,12 x 10 <sup>3</sup>      | -3,68 x 10 <sup>2</sup>     | 1,3 x 10 <sup>3</sup>               | 4,37 x 10 <sup>2</sup>              | 3,84 x 10 <sup>3</sup>              |
|            | Velocidad relativa (%) | -0,04                       | 0,29                        | -0,09                       | 0,34                                | 0,11                                | 1,00                                |
| Ac-ANW-AMC | V <sub>0</sub> RFU/min | 7,28                        | 3,28 x 10 <sup>2</sup>      | 1,4 x 10                    | -2,6 x 10                           | -3,16 x 10 <sup>2</sup>             | -3,5 x 10                           |
|            | Velocidad relativa (%) | 1,91 x 10 <sup>-3</sup>     | 0,09                        | 3,66 x 10 <sup>-3</sup>     | -6,81 x 10 <sup>-3</sup>            | -0,08                               | 9,16 x 10 <sup>-3</sup>             |

Tabla 7: V<sub>0</sub> (Unidades Relativa de Fluorescencia (RFU/min)) y velocidad relativa (%) de la actividad catalítica del proteasoma constitutivo e intermedio de tipo I catalizando los sustratos fluorogénicos Ac-WLA-AMC y Ac-ANW-AMC en presencia de Bortezomib. La V<sub>0</sub> se calcula como la pendiente de la fase lineal de emisión de fluorescencia. La velocidad relativa se calcula tomando como 100% la V<sub>0</sub> del proteasoma intermedio de tipo I (3) frente Ac-ANW-AMC. Cada valor hace referencia al promedio de triplicados.

En conjunto, los resultados confirmaron que el proteasoma intermedio de tipo I poseía un perfil catalítico distinto al del proteasoma constitutivo, siendo más eficiente tras residuos básicos y aromáticos y menos activo tras residuos alifáticos pequeños.

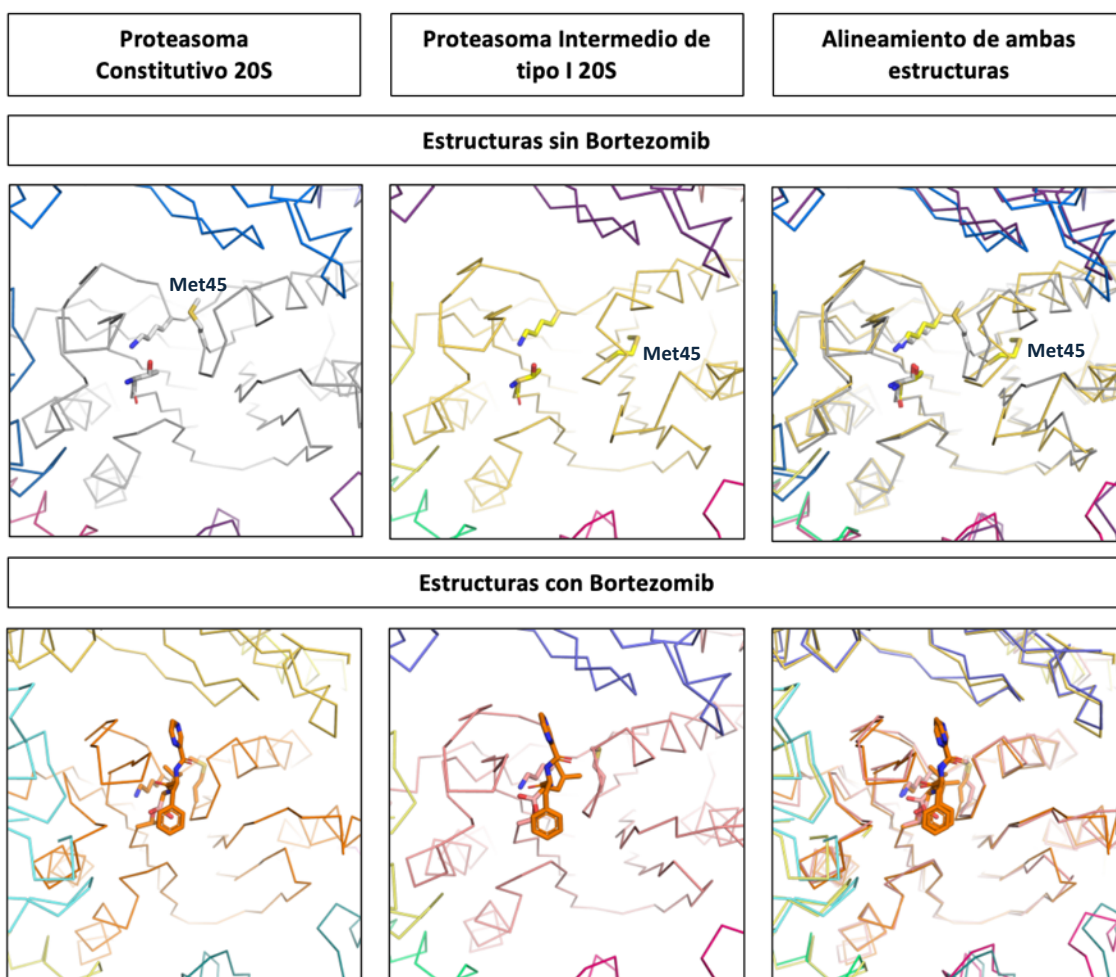
## 7. ESTRUCTURA

La resolución de la estructura tridimensional del proteasoma intermedio de tipo I se realizó en colaboración con el laboratorio del Dr. David Reverter del Institut de Biotecnologia i Biomedicina (IBB-UAB).

Para ello, el SIP se analizó en el microscopio criogénico (Cryo-EM) localizado en el Sincrotrón ALBA. Las imágenes obtenidas se procesaron y se logró definir la estructura tridimensional del SIP sin ningún inhibidor y en complejo con el inhibidor selectivo del proteasoma Bortezomib a una resolución de 2,94Å y 3,11Å, respectivamente. El análisis de la estructura obtenida se centró en el centro activo generado en la subunidad β5i. La **Figura 30** muestra que esta subunidad tiene una estructura similar a la subunidad β5 del proteasoma constitutivo (6RGQ<sup>198</sup>) aunque con diferencias que modifican el *pocket* S1 y que explican la diferente actividad catalítica. Así, la principal diferencia se observó en el residuo Met45, que se encuentra hacia el interior del *pocket* S1 de unión al sustrato en el proteasoma constitutivo y dirigido hacia el exterior en la subunidad β5i. La unión del inhibidor provocó que la Met45 adquiriera esta conformación cerrada en la subunidad β5i, similar al *pocket* de la subunidad β5 en ambos proteasomas.

Este estrechamiento se produce también cuando otro sustrato interacciona con la subunidad  $\beta 5i$  del inmunoproteasoma (6E5B<sup>199</sup>) (**Figura 31**).

Esta disposición diferencial de la Met45 en la subunidad  $\beta 5i$  en ausencia de inhibidor provocó que el *pocket* S1 fuera más amplio en comparación con la subunidad  $\beta 5$ , lo que permitió que la subunidad  $\beta 5i$  acomodara preferentemente residuos aromáticos, tal y como se había observado al analizar la actividad catalítica del proteasoma constitutivo y del SIP y, consecuentemente, se vio reflejado en los inmunopeptidomas de clase I presentados por células que expresan exclusivamente uno de los complejos (**Figura 32**).



**Figura 30:** Estructuras cristalográficas del proteasoma constitutivo y del proteasoma intermedio de tipo I solos y junto con el inhibidor Bortezomib. Entrada PDB del proteasoma constitutivo:6RGQ.

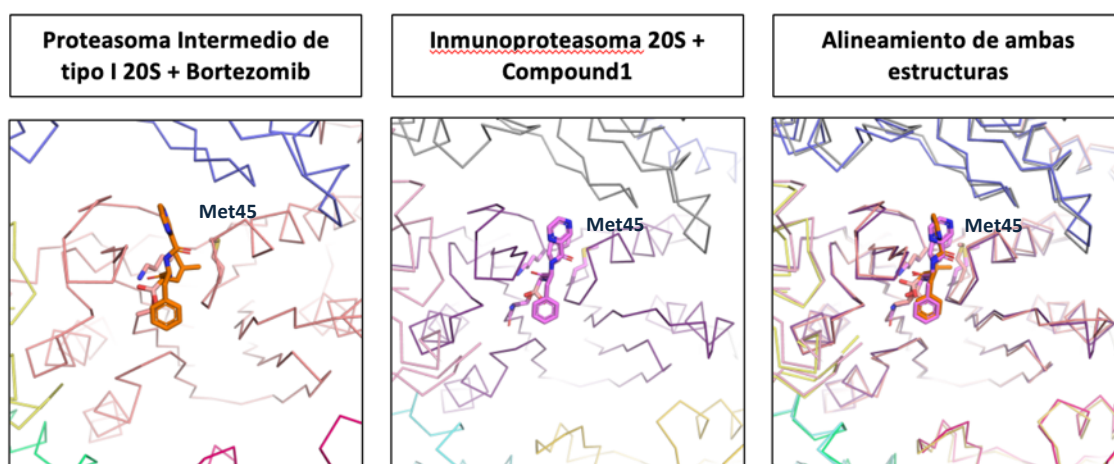


Figura 31: Estructura cristalográfica de del proteasoma intermedio de tipo I con el inhibidor Bortezomib y del inmunoproteasoma con el Compound 1. Entrada PDB del inmunoproteasoma 6E5B.

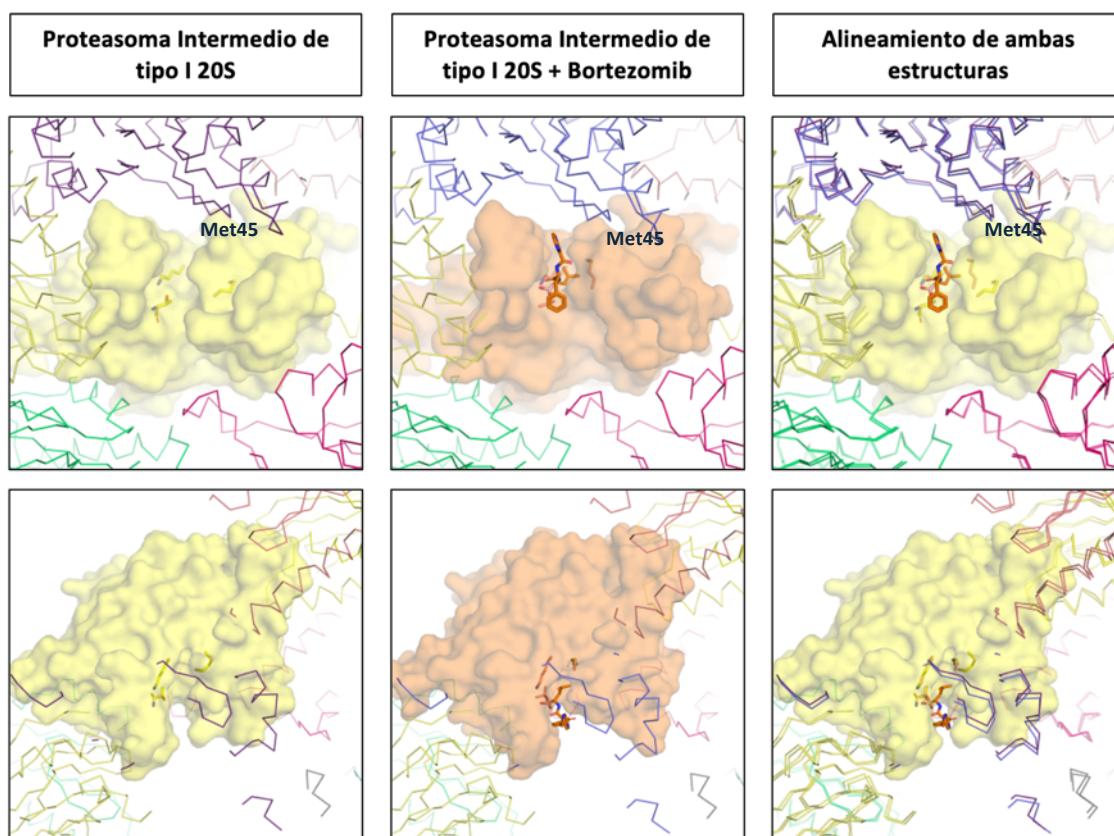


Figura 32: Estructura cristalográfica del proteasoma intermedio de tipo I solo y junto con el inhibidor Bortezomib.

Por lo tanto, los datos aportados en este capítulo mostraron que la estructura tridimensional de la subunidad  $\beta 5i$  del SIP y la subunidad  $\beta 5$  del proteasoma constitutivo presentan diferencias en la disposición de la Met45 que condiciona la especificidad de sustrato, haciendo el *pocket* de la subunidad  $\beta 5i$  más amplio y propenso a acomodar residuos aromáticos. Este fenómeno se tradujo en una actividad catalítica tipo quimotripsina diferenciada entre ambos complejos, siendo el SIP más eficiente catalizando el enlace tras residuos aromáticos y menos tras residuos alifáticos pequeños. En última instancia, estas observaciones quedaron reflejadas en los repertorios peptídicos presentados por HLA-I, en los que las células que expresaban el SIP presentaban mayor proporción de péptidos con residuos C-terminales aromáticos.



**CAPÍTULO II: ANÁLISIS DEL REPERTORIO**  
**PEPTÍDICO DE MOLÉCULAS DE HLA-I**  
**PERTENECIENTES AL SUPERTIPO B7**





El supertipo B7 está formado por moléculas que unen Pro2 y un aminoácido aromático o alifático en PΩ<sup>141</sup>, aunque también toleran residuos con Ala2<sup>145, 146,147</sup>. Así, las moléculas HLA-B\*51:01 y HLA-B\*51:08 unen un repertorio con un elevado porcentaje de péptidos con Ala2, el alotipo HLA-B\*07:02 presenta muy pocos péptidos con Ala2<sup>148</sup>. Se ha descrito que las moléculas HLA-B\*51:01 y HLA-B\*51:08, pertenecientes a este supertipo, presentan dos subpeptidomas: uno de ellos formado por Pro2 y residuos hidrofóbicos en P1, y otro por Ala2 y Asp1<sup>149,150,200</sup>.

En este capítulo:

1. Se ha estudiado si la presencia de subpeptidomas está a B\*51 o si se extiende a otras moléculas.
2. Se ha analizado la influencia de los residuos en posiciones 45 y 67 en la especificidad del *pocket* B y en la presencia de subpeptidomas.

## **ANÁLISIS DE LA PRESENCIA DE SUBPEPTIDOMAS EN LAS MOLÉCULAS DE HLA-I PERTENECIENTES AL SUPERTIPO B7**

### **1. ANÁLISIS DE LOS INMUNOPEPTIDOMAS PRESENTADOS POR HLA-I DE MOLÉCULAS DEL SUPERTIPO B7**

Se analizaron los repertorios de diferentes moléculas de HLA-B pertenecientes al supertipo B7 a partir de la base de datos IEDB<sup>169</sup>. Se tuvieron en cuenta solo los peptidomas que contenían más de 500 péptidos de entre 8 y 13 residuos de longitud. Bajo esta premisa, se seleccionaron los peptidomas de 13 moléculas: **HLA-B\*07:02, -B\*07:04, -B\*35:01, -B\*35:03, -B\*35:07, -B\*42:01, -B\*51:01, -B\*51:08, -B\*53:01, -B\*54:01, -B\*55:01, -B\*55:02 y -B\*56:01 (Figura 33).**

| AA Pos. | 10         | 20          | 30         | 40         | 50         | 60          | 70         | 80          | 90         | 100        |
|---------|------------|-------------|------------|------------|------------|-------------|------------|-------------|------------|------------|
| B*07:02 | GSHSMRYFYT | SVSRPGRGEP  | RFISVGIVDD | TQFVRFDSDA | ASPREEPRAP | WIEQEGPEYW  | DRNTQIYKAQ | AQTDRESLRN  | LRGYYNQSEA | GSHTLQSMYG |
| B*07:04 | -----      | -----       | -----      | -----      | -----      | -----       | -----      | -----       | -----      | -----      |
| B*42:01 | -----      | -----       | -----      | -----      | -----      | -----       | -----      | -----       | -----      | -----      |
| B*55:01 | -----      | AM-----     | --A-----   | -----      | -----      | -----       | -----      | -----       | -----      | ---W-T---  |
| B*55:02 | -----      | AM-----     | --A-----   | -----      | -----      | -----       | -----      | -----       | -----      | ---W-T---  |
| B*56:01 | -----      | AM-----     | --A-----   | -----      | -----      | -----       | -----      | -----       | -----      | ---W-T---  |
| B*35:01 | -----      | AM-----     | --A-----   | -----      | ---T-----  | -----       | -----F-TN  | T--Y-----   | -----      | ---II-R--- |
| B*35:03 | -----      | AM-----     | --A-----   | -----      | ---T-----  | -----       | -----F-TN  | T--Y-----   | -----      | ---II-R--- |
| B*35:07 | -----      | AM---V----- | --A-----   | -----      | ---T-----  | -----       | -----F-TN  | T--Y-----   | -----      | ---II-R--- |
| B*53:01 | -----      | AM-----     | --A-----   | -----      | ---T-----  | -----       | -----F-TN  | T--Y--N--I  | ALR-----   | ---II-R--- |
| B*51:01 | -----      | AM-----     | --A-----   | -----      | ---T-----  | -----       | -----F-TN  | T--Y--N--I  | ALR-----   | ---W-T---  |
| B*51:08 | -----      | AM-----     | --A-----   | -----      | ---T-----  | -----       | -----F-TN  | T--Y--N--I  | ALR-----   | ---W-T---  |
| B*54:01 | -----      | AM-----     | --A-----   | -----      | ---G-----  | -V-----     | -----      | -----       | -----      | ---W-T---  |
| AA Pos. | 110        | 120         | 130        | 140        | 150        | 160         | 170        | 180         | 190        | 200        |
| B*07:02 | CDVGPDRLL  | RGHDPYAYDG  | KDYIALNEDL | RSWTAADTAA | QITQRKWEAA | REAEQRRAYL  | EGECVEWLR  | YLENGKDKLE  | RADPPKTHVT | HHPISDHEAT |
| B*07:04 | -----      | -----       | -----      | -----      | -----      | -----D----- | -----      | -----       | -----      | -----      |
| B*42:01 | -----      | --N-----    | -----      | -----      | -----      | -V--D-----  | -T-----    | -----T----- | -----      | -----      |
| B*55:01 | --L-----   | --N-L-----  | -----      | S-----     | -----      | -----L----- | -T-----    | -----ET-Q   | -----      | -----      |
| B*55:02 | --L-----   | --N-L-----  | -----      | S-----     | -----      | -V--L-----  | -T-----    | -----ET-Q   | -----      | -----      |
| B*56:01 | --L-----   | --N-L-----  | -----      | S-----     | -----      | -V--L-----  | -L-----    | -----ET-Q   | -----      | -----      |
| B*35:01 | --L-----   | -----S----- | -----      | S-----     | -----      | -V--L-----  | --L-----   | -----ET-Q   | -----      | ---V-----  |
| B*35:03 | --L-----   | -----F----- | -----      | S-----     | -----      | -V--L-----  | --L-----   | -----ET-Q   | -----      | ---V-----  |
| B*35:07 | --L-----   | -----S----- | -----      | S-----     | -----      | -V--L-----  | --L-----   | -----ET-Q   | -----      | ---V-----  |
| B*53:01 | --L-----   | -----S----- | -----      | S-----     | -----      | -V--L-----  | --L-----   | -----ET-Q   | -----      | ---V-----  |
| B*51:01 | -----      | --N-----    | -----      | S-----     | -----      | -----L----- | --L-----   | H---ET-Q    | -----      | ---V-----  |
| B*51:08 | -----      | --N-----    | -----      | S-----     | -----      | -V--D-----  | --L-----   | H---ET-Q    | -----      | ---V-----  |
| B*54:01 | --L-----   | --N-L-----  | -----      | S-----     | -----      | -V--L-----  | -T-----    | -----ET-Q   | -----      | -----      |

**Figura 33: Alineamiento de secuencias de los 13 alotipos del supertipo B7 analizados.** Las secuencias primarias de las moléculas de HLA-I analizadas fueron alineadas. Se indican los residuos variables respecto a HLA-B\*07:02.

El número de péptidos que componía cada peptidoma fue de entre 624 y 12033, siendo la mayoría de ellos nonámeros (49,5% - 77,9%), seguido de octámeros y péptidos de entre 10 y 13 residuos de longitud. Las moléculas de HLA-B\*51:01 y HLA-B\*51:08 mostraron el mayor porcentaje de octámeros (27,4% y 13,6%, respectivamente) y el menor porcentaje de péptidos de 10 o más residuos. Por otro lado, las moléculas HLA-B\*07:02, -B\*07:04, -B\*55:02 y -B\*56:01 mostraron más de un 20% de péptidos de 10 residuos y más de un 11% de péptidos de 11 residuos de longitud. El número medio de residuos de los péptidos que componían cada peptidoma fue similar en todas las moléculas de HLA-I analizadas (entre un 9,3% y 9,7%) excepto el grupo de alelo B\*51, con una media de 9 residuos ([Tabla 8](#)).

| HLA-B* | Total | Péptidos por número de residuos (%) |      |      |      |     |     | AAs promedio |
|--------|-------|-------------------------------------|------|------|------|-----|-----|--------------|
|        |       | 8                                   | 9    | 10   | 11   | 12  | 13  |              |
| 07:02  | 12033 | 7,8                                 | 49,5 | 22,0 | 12,3 | 5,7 | 2,8 | 9,67         |
| 07:04  | 1480  | 5,7                                 | 61,2 | 20,0 | 10,3 | 1,9 | 0,8 | 9,44         |
| 35:01  | 1954  | 5,3                                 | 57,4 | 17,1 | 11,6 | 5,5 | 3,0 | 9,64         |
| 35:03  | 1527  | 4,8                                 | 76,4 | 10,4 | 6,1  | 1,8 | 0,6 | 9,26         |
| 35:07  | 1933  | 3,3                                 | 68,4 | 14,5 | 9,7  | 2,7 | 1,5 | 9,44         |
| 42:01  | 3428  | 5,9                                 | 59,4 | 17,5 | 13,0 | 2,9 | 1,3 | 9,52         |
| 51:01  | 5157  | 27,4                                | 55,9 | 8,5  | 4,1  | 2,4 | 1,9 | 9,04         |
| 51:08  | 624   | 13,6                                | 77,9 | 4,2  | 3,0  | 1,3 | 0,0 | 9,00         |
| 53:01  | 2204  | 4,5                                 | 63,4 | 17,6 | 9,6  | 3,6 | 1,3 | 9,48         |
| 54:01  | 1441  | 11,2                                | 52,3 | 16,5 | 12,3 | 4,4 | 3,3 | 9,56         |
| 55:01  | 1548  | 11,3                                | 55,8 | 19,8 | 9,0  | 2,4 | 1,7 | 9,40         |
| 55:02  | 1549  | 3,4                                 | 54,4 | 22,7 | 13,4 | 4,1 | 2,0 | 9,66         |
| 56:01  | 2113  | 4,9                                 | 51,5 | 24,3 | 13,5 | 3,7 | 2,0 | 9,66         |

Tabla 8: Tamaño, por número de aminoácidos, de los inmunopeptidomas de las moléculas de HLA-I del supertipo B7 seleccionadas.

## 1.2. Grado de solapamiento de los inmunopeptidomas

El grado de solapamiento de los diferentes repertorios peptídicos se analizó por pares de alotipos. Los 13 inmunopeptidomas estaban compuestos por un total de 27039 péptidos, de los cuales 6134 (22,7%) eran compartidos por al menos dos alotipos, y 2445 (9,0%) eran compartidos por al menos 3 moléculas de HLA-I. Por norma general, los peptidomas de diferentes subtipos del mismo grupo alélico (HLA-B\*07:02 con -B\*07:04; -B\*35:01 con -B\*35:03 y -B\*35:07; -B\*51:01 con -B\*51:08; y -B\*55:01 con -B\*55:02) presentaron mayor grado de solapamiento, aunque el mayor solapamiento se dio entre las moléculas HLA-B\*55:02 y -B\*56:01, que compartían 1.207 péptidos, equivalentes a un 57,1% y 77,9% del peptidoma, respectivamente ([Tabla 9](#)).

|             | 12033         | 1480          | 1954          | 1527          | 1933          | 3428         | 5157         | 624         | 2204        | 1441          | 1548          | 1549          | 2113    |
|-------------|---------------|---------------|---------------|---------------|---------------|--------------|--------------|-------------|-------------|---------------|---------------|---------------|---------|
| Alotipo HLA | B*07:02       | B*07:04       | B*35:01       | B*35:03       | B*35:07       | B*42:01      | B*51:01      | B*51:08     | B*53:01     | B*54:01       | B*55:01       | B*55:02       | B*56:01 |
| B*07:02     |               | 822           | 228           | 314           | 199           | 1693         | 203          | 33          | 181         | 82            | 286           | 230           | 271     |
| B*07:04     | 55,5/<br>6,8  |               | 25            | 59            | 41            | 878          | 32           | 9           | 36          | 19            | 120           | 80            | 85      |
| B*35:01     | 11,7/<br>1,9  | 1,3/<br>1,7   |               | 246           | 865           | 115          | 153          | 3           | 452         | 68            | 9             | 22            | 25      |
| B*35:03     | 20,6/<br>2,6  | 3,9/<br>4,0   | 16,1/<br>12,6 |               | 371           | 237          | 93           | 12          | 322         | 45            | 34            | 39            | 55      |
| B*35:07     | 10,3/<br>1,7  | 2,1/<br>2,8   | 44,7/<br>44,3 | 19,2/<br>24,3 |               | 169          | 60           | 3           | 613         | 28            | 21            | 31            | 42      |
| B*42:01     | 49,4/<br>14,1 | 25,6/<br>59,3 | 3,4/<br>5,9   | 6,9/<br>15,5  | 4,9/<br>8,7   |              | 129          | 49          | 149         | 66            | 238           | 191           | 209     |
| B*51:01     | 3,9/<br>1,7   | 0,6/<br>2,2   | 3,0/<br>7,8   | 1,8/<br>6,1   | 1,2/<br>3,1   | 2,5/<br>3,8  |              | 337         | 178         | 215           | 100           | 99            | 130     |
| B*51:08     | 5,3/<br>0,3   | 1,4/<br>0,6   | 0,5/<br>0,2   | 1,9/<br>0,8   | 0,5/<br>0,2   | 7,9/<br>1,4  | 54,0/<br>6,5 |             | 34          | 30            | 23            | 25            | 25      |
| B*53:01     | 8,2/<br>1,5   | 1,6/<br>2,4   | 20,5/<br>23,1 | 14,6/<br>21,1 | 27,8/<br>31,7 | 6,8/<br>4,3  | 8,1/<br>3,5  | 1,5/<br>5,4 |             | 37            | 32            | 36            | 41      |
| B*54:01     | 5,7/<br>0,7   | 1,3/<br>1,3   | 4,7/<br>3,5   | 3,1/<br>2,9   | 1,9/<br>1,4   | 4,6/<br>1,9  | 14,9/<br>4,2 | 2,1/<br>4,8 | 2,6/<br>1,7 |               | 226           | 329           | 425     |
| B*55:01     | 18,5/<br>2,4  | 7,8/<br>8,1   | 0,6/<br>0,5   | 2,2/<br>2,2   | 1,4/<br>1,1   | 15,4/<br>6,9 | 6,5/<br>1,9  | 1,5/<br>3,7 | 2,1/<br>1,5 | 14,6/<br>15,7 |               | 710           | 777     |
| B*55:02     | 14,8/<br>1,9  | 5,2/<br>5,4   | 1,4/<br>1,1   | 2,5/<br>2,6   | 2,0/<br>1,6   | 12,3/<br>5,6 | 6,4/<br>1,9  | 1,6/<br>4,0 | 2,3/<br>1,6 | 21,2/<br>22,8 | 45,8/<br>45,9 |               | 1207    |
| B*56:01     | 12,8/<br>2,3  | 4,0/<br>5,7   | 1,2/<br>1,3   | 2,6/<br>3,6   | 2,0/<br>2,2   | 9,9/<br>6,1  | 6,2/<br>2,5  | 1,2/<br>4,0 | 1,9/<br>1,9 | 20,1/<br>29,5 | 36,8/<br>50,2 | 57,1/<br>77,9 |         |

**Tabla 9: Solapamiento de los peptidomas de los diferentes alotipos analizados del supertipo B7.**

**Nota:** Por encima de la línea diagonal se muestran los números absolutos de los péptidos compartidos. Debajo de la diagonal, el porcentaje de los péptidos compartidos (el número a la izquierda de la barra hace referencia al alotipo en la izquierda, y el número a la derecha de la barra hace referencia al alotipo indicado arriba)

**Nota:** Los números sombreados en gris claro indican que los péptidos compartidos más de un 10% del peptidoma total en al menos un alotipo. El sombreado gris oscuro indica que los péptidos compartidos representan más de un 20% del peptidoma total en al menos un alotipo.

### 1.2. Residuos de anclaje

Como ya se ha mencionado, las moléculas pertenecientes al supertipo B7 unen preferentemente Pro2 y residuos aromáticos, alifáticos o hidrofóbicos en P $\Omega$ , aunque también se han identificado residuos que presentan Ala2. Los residuos de anclaje del peptidoma de las 13 moléculas analizadas se determinaron utilizando el *software* PRBAM ([Anexo 7](#)). La mayoría de estas moléculas unían preferentemente Pro2, mientras que los péptidos con Ala2 se toleraban en diferente grado, conformando entre un 2,6% del peptidoma en -B\*55:02 y -B\*56:01 hasta un 54,3% en -B\*51:08 ([Tabla 10](#)).

|             | Total | Pro2         | Ala2         | Otros        |
|-------------|-------|--------------|--------------|--------------|
| HLA-B*07:02 | 12033 | 9956 (82,7%) | 338 (2,8%)   | 1739 (14,5%) |
| HLA-B*07:04 | 1480  | 1233 (83,3%) | 58 (3,9%)    | 189 (12,8%)  |
| HLA-B*35:01 | 1954  | 1322 (67,7%) | 297 (15,2%)  | 335 (17,1%)  |
| HLA-B*35:03 | 1527  | 1251 (81,9%) | 177 (11,6%)  | 99 (6,5%)    |
| HLA-B*35:07 | 1933  | 1373 (71%)   | 399 (20,6%)  | 161 (8,3%)   |
| HLA-B*42:01 | 3428  | 2987 (87,1%) | 100 (2,9%)   | 341 (9,9%)   |
| HLA-B*51:01 | 5157  | 2566 (49,8%) | 1552 (30,1%) | 1039 (20,1%) |
| HLA-B*51:08 | 624   | 285 (45,7%)  | 339 (54,3%)  | 0 (0%)       |
| HLA-B*53:01 | 2204  | 1600 (72,6%) | 436 (19,8%)  | 168 (7,6%)   |
| HLA-B*54:01 | 1441  | 1038 (72,0%) | 145 (10,1%)  | 258 (17,9%)  |
| HLA-B*55:01 | 1548  | 1378 (89,0%) | 42 (2,7%)    | 128 (8,3%)   |
| HLA-B*55:02 | 1549  | 1407 (90,8%) | 40 (2,6%)    | 102 (6,6%)   |
| HLA-B*56:01 | 2113  | 1873 (88,6%) | 54 (2,6%)    | 186 (8,8%)   |

Tabla 10: Preferencia de aminoácido en posición P2 en las moléculas del supertipo B7 analizadas.

### 1.3. Afinidad teórica de los ligandos de moléculas de HLA del supertipo B7

La afinidad teórica de los diferentes péptidos de unión a las moléculas de HLA de clase I pertenecientes al supertipo B7 analizadas se evaluó utilizando el *software* NetMHCpan4.1. Las moléculas disponibles para analizar con este programa informático fueron: **HLA-B\*07:02, -B\*35:01, -B\*35:03, -B\*51:01, -B\*53:01 y -B\*54:01**. La afinidad teórica de los péptidos con Pro2 a su molécula de HLA correspondiente fue notablemente superior que la de aquellos péptidos con Ala2 ([Figura 34](#)).

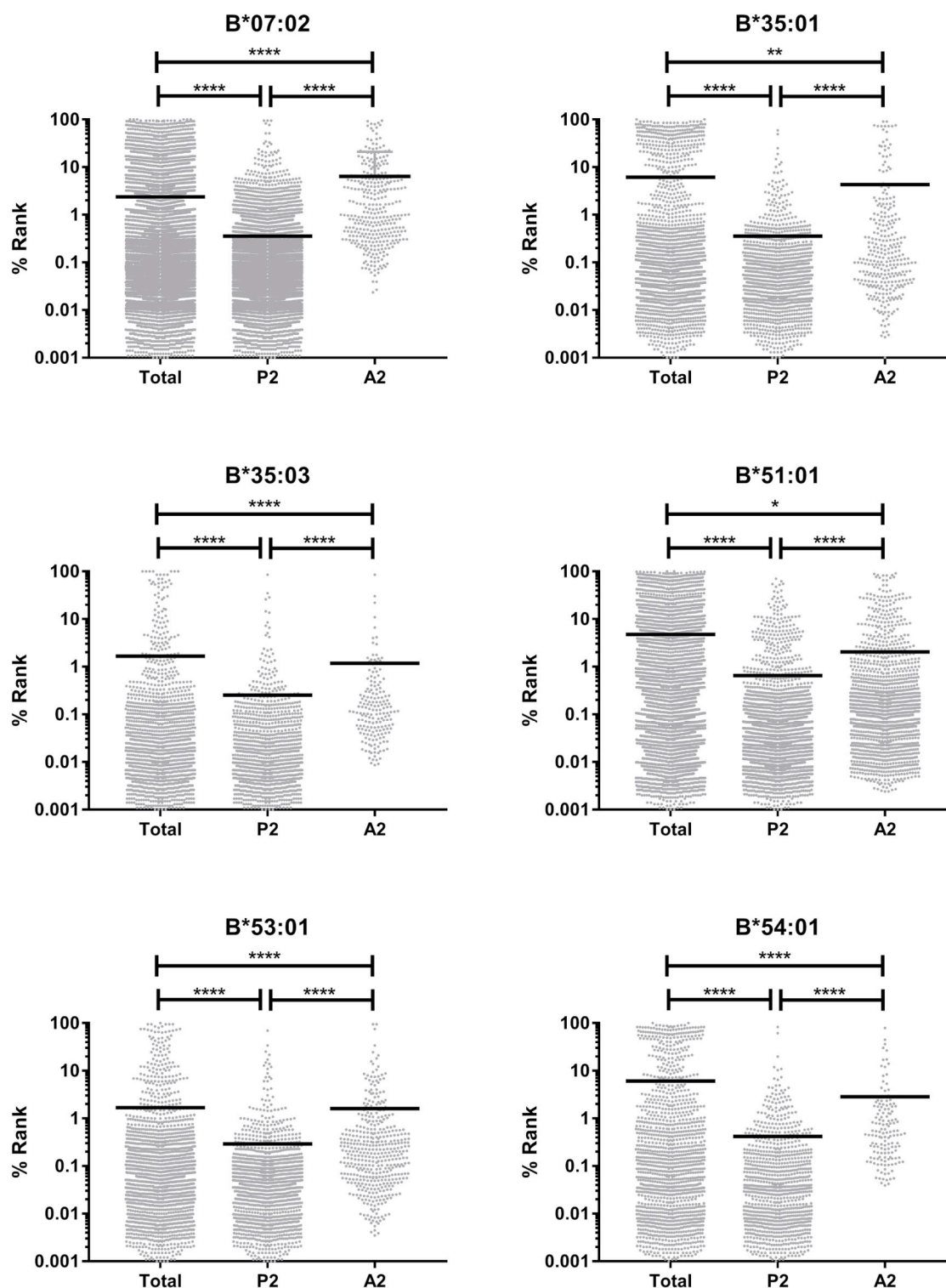


Figura 34: Afinidad teórica de las moléculas de HLA pertenecientes al supertipo B7 analizadas. Se muestra el % de péptidos totales, el subpeptidoma con Pro2 (P2) y el subpeptidoma con Ala2 (A2) respecto la molécula de HLA indicada. Los símbolos representan péptidos individuales; las líneas horizontales representan la media  $* = 0.01 < P < 0.1$ ;  $** = 0.001 < P < 0.01$ ;  $*** = 0.001 < P < 0.01$ ;  $**** = P < 0.0001$  según el test de Mann-Whitney U.

## **2. ANÁLISIS DE LA PRESENCIA DE SUBPEPTIDOMAS DE LAS MOLÉCULAS DEL SUPERTIPO B7 ANALIZADAS**

Tras caracterizar los repertorios peptídicos de las moléculas de HLA-I pertenecientes al supertipo B7 escogidas según los criterios previamente mencionados, se analizó si los repertorios de dichas moléculas presentaban subpeptidomas.

### **2.1. HLA-B\*54:01 presenta dos subpeptidomas similares a HLA-B\*51:01 y HLA-B\*51:08**

El repertorio peptídico de la molécula **HLA-B\*51:01** fue descrito a partir de 1620 péptidos, identificando dos subpeptidomas distintos. Esta molécula une preferentemente residuos con Pro2, sin embargo, es capaz de unir un elevado porcentaje de péptidos con Ala2<sup>146</sup>. Los autores describieron que, en este alotipo, la preferencia de aminoácido en *P2* dictaminaba los residuos localizados en *P1*, dando lugar a dos subpeptidomas diferenciados: uno formado por péptidos con Pro2 y residuos alifáticos y aromáticos en *P1*, y otro formado por péptidos con Ala2 y Asp1. En esta tesis, ese análisis se expandió con los 5157 ligandos de esta molécula disponibles en la base de datos IEDB, obteniendo resultados muy similares a los descritos previamente en la bibliografía (**Figura 35**). Asimismo, el peptidoma de la molécula **HLA-B\*51:08**, formado por 624 péptidos, fue previamente analizado<sup>145</sup> y mostró características similares al peptidoma de HLA-B\*51:01 (**Figura 35**).

Con el fin de determinar si la presencia de subpeptidomas es exclusiva de moléculas del grupo de alelo B\*51, este análisis se expandió al resto de moléculas de HLA-I del supertipo B7 previamente seleccionadas. El peptidoma de la molécula **HLA-B\*54:01** estaba formado por 1441 péptidos, de los cuales 1038 (72%) contenían Pro2, 145 (10,1%) Ala2 y 258 (17,9%) otros residuos en *P2*. Tal y como se describió en las moléculas HLA-B\*51:01 y HLA-B\*51:08, el subpeptidoma compuesto por péptidos con Pro2 presentaba residuos alifáticos o aromáticos en *P1*. Del mismo modo, el subpeptidoma compuesto por péptidos con Ala2 estaba significativamente enriquecido en residuos ácidos en *P1*. Sin embargo, contenían más Glu1 que Asp1 (**Figura 35**).

Por lo tanto, la presencia de dos subpeptidomas, uno de ellos con Pro2 y preferencia de residuos hidrofóbicos en *P1* y otro con Ala2 y preferencia de residuos ácidos en *P1*, está presente en al menos una molécula más del supertipo B7, diferente a los alotipos B\*51:01 y B\*51:08.

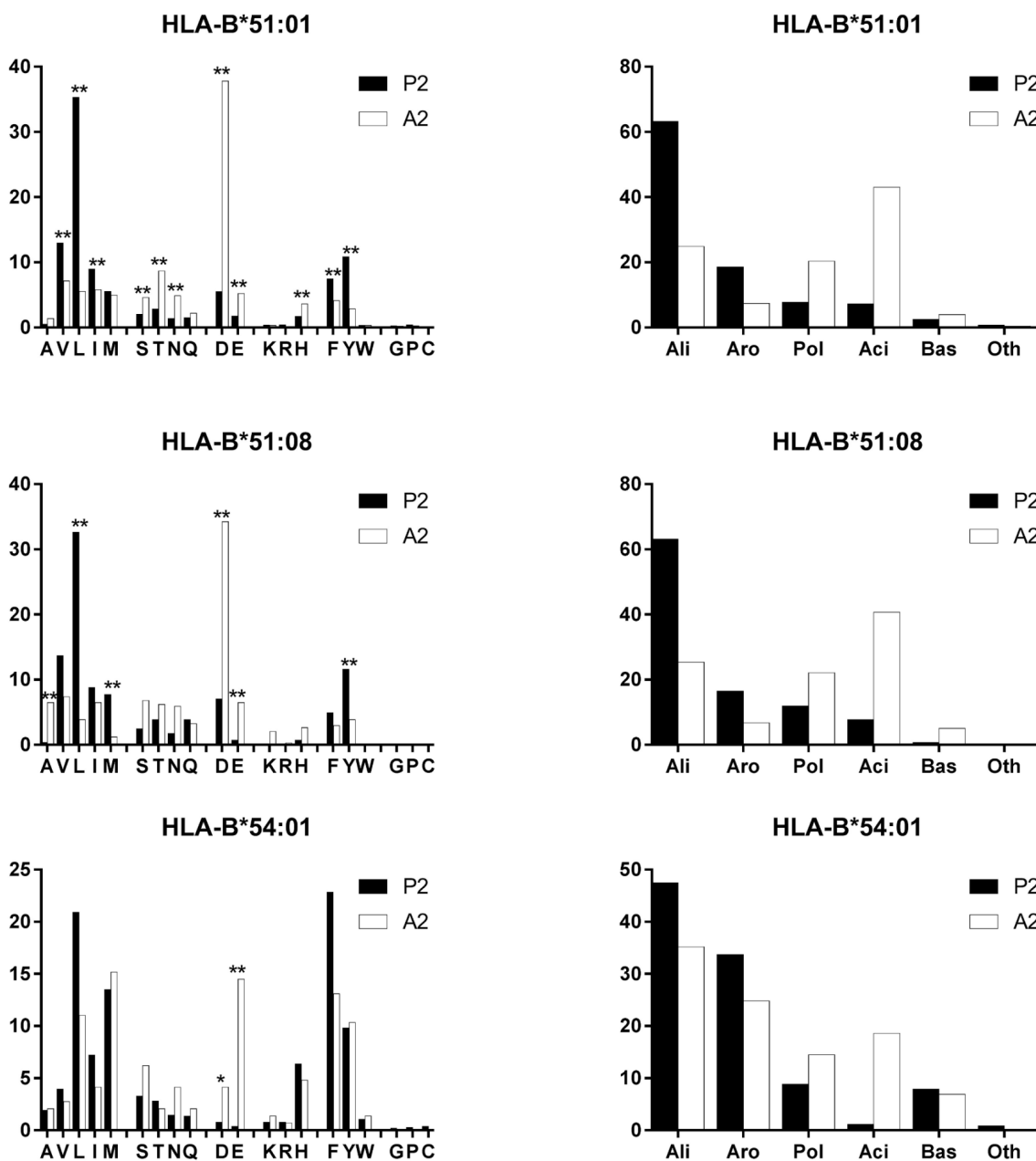


Figura 35: Residuos localizados en *P1* de los ligandos de HLA-B\*51:01, HLA-B\*51:08 y HLA-B\*54:01 con Prolina en *P2* (P2) y Alanina en *P2* (A2). Se utiliza el código de una sola letra para identificar los aminoácidos. En las gráficas de la izquierda, se muestra la frecuencia de cada residuo en posición 1 según si la posición 2 está ocupada por Prolina (negro) o Alanina (blanco). Las diferencias significativas se determinan mediante un test Chi cuadrado y la corrección de Bonferroni como \* < 0.05 y \*\* < 0.01. En las gráficas de la derecha, los residuos se agrupan según sus características fisicoquímicas.



## 2.2. El grupo alélico HLA-B\*35 y la molécula HLA-B\*53:01 presentan dos subpeptidomas con residuos polares y ácidos en P1 en el subpeptidoma que contiene Ala2

Este análisis se amplió con las moléculas del grupo alélico B\*35 disponibles en la base de datos IEDB y la molécula HLA-B\*53:01, que al igual que las moléculas previamente analizadas, presentaron dos subpeptidomas diferenciados.

Las peptidomas de las moléculas de clase I **HLA-B\*35:01, -B\*35:03, -B\*35:07 y -B\*53:01** presentaron preferencia de Pro2, entre un 67,7 y un 81,9%, así como de Ala2, entre un 11,6 y un 20,6% (**Tabla 8**). El análisis de estos peptidomas reveló que los péptidos que presentaban Pro2 tenían preferencia por residuos aromáticos y alifáticos en *P1*, mientras que los péptidos que presentaban Ala2 presentaron mayormente residuos polares y ácidos en *P1*. En el caso de estas moléculas, los residuos polares (Ser, Thr, Asn y Gln) eran preferidos frente los residuos ácidos (Asp y Glu) (**Figura 36**).

Se concluyó, por tanto, que las moléculas de clase I HLA-B\*35:01, -B\*35:03, -B\*35:07 y -B\*53:01 presentaban dos subpeptidomas: uno de ellos con Pro2 y residuos alifáticos y aromáticos en *P1*, y otro con Ala2 y residuos polares y ácidos en *P1*.

Cabe destacar que el subpeptidoma formado por péptidos con Ala2, presentado por la molécula HLA-B\*51:01, se encontraba enriquecido en residuos polares en *P1*. Esta tendencia también se observó en los subpeptidomas de las moléculas HLA-B\*51:08 y HLA-B\*54:01, lo que sugiere que los tres subpeptidomas caracterizados en estas moléculas del supertipo B7 podrían agruparse en dos subpeptidomas: uno formado por péptidos con Pro2 y residuos aromáticos y alifáticos en *P1*, y otro formado por Ala2 y residuos polares y ácidos en *P1*.

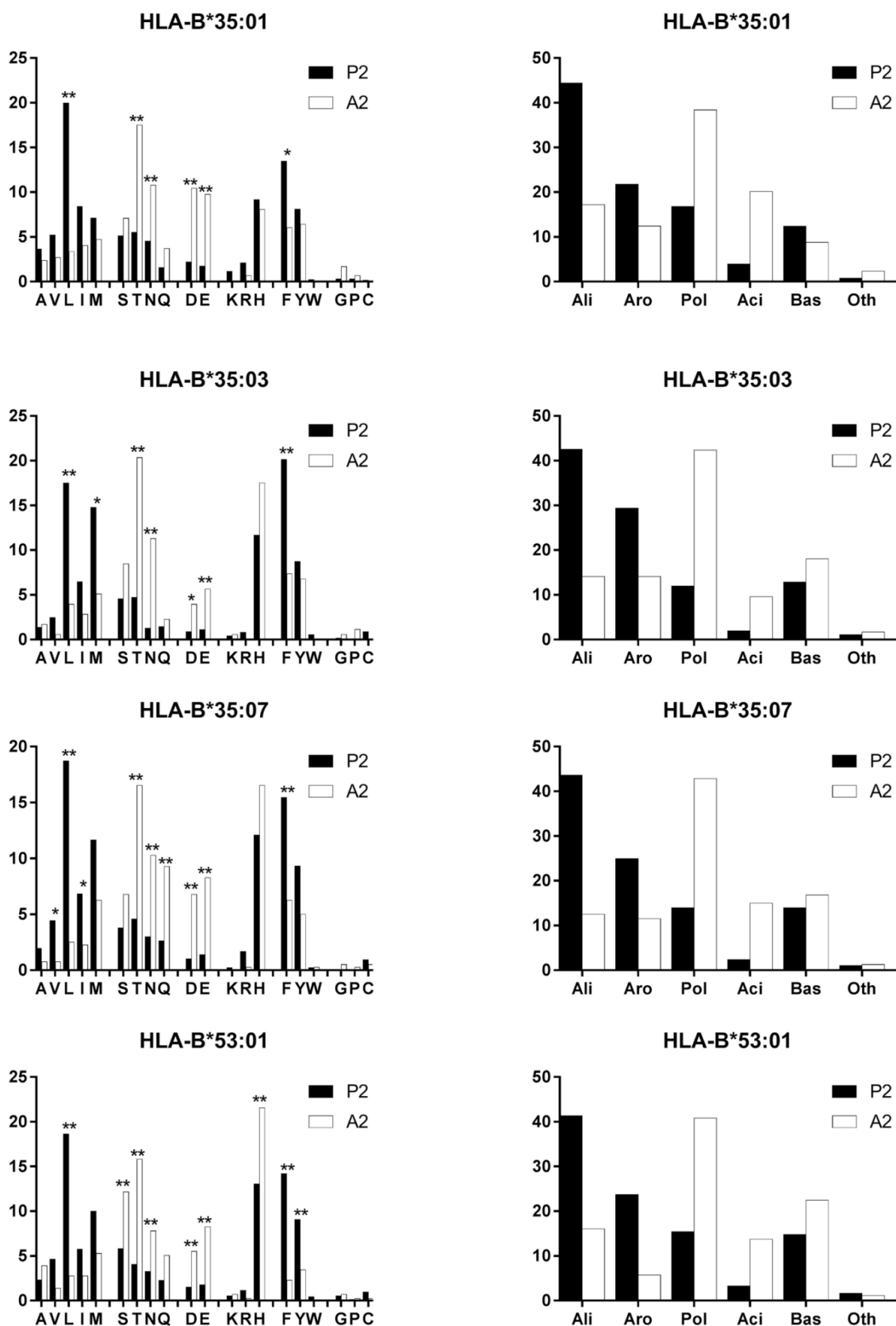


Figura 36: Residuos localizados en P1 de los ligandos de HLA-B\*35:01, HLA-B\*35:03, HLA-B\*35:07 y HLA-B\*53:01 con Prolina en P2 (P2) y Alanina en P2 (A2). Se utiliza el código de una sola letra para identificar los aminoácidos. En las gráficas de la izquierda, se muestra la frecuencia de cada residuo en posición 1 según si la posición 2 está ocupada por Prolina (negro) o Alanina (blanco). Las diferencias significativas se determinan mediante un test Chi cuadrado y la corrección de Bonferroni como \* < 0.05 y \*\* < 0.01. En las gráficas de la derecha, los residuos se agrupan según sus características fisicoquímicas.

### 2.3. Otras moléculas del supertipo B7 no presentan subpeptidomas

El análisis del peptidoma del resto de moléculas de HLA-B del supertipo B7 (**HLA-B\*07:02, -B\*07:04, -B\*42:01, -B\*55:01, -B\*55:02 y -B\*56:01**) reveló que no presentan subpeptidomas, solo presentaron Pro favorecida en *P2*. A pesar de que los péptidos con Ala2 no estaban favorecidos, algunos ligandos de estas moléculas sí contenían Ala en *P2* (con una abundancia entre 2,6% y 3,9%). Se analizó el residuo en *P1* de aquellos péptidos que contenían Ala2 y se comparó con aquellos que tenían Pro2. A pesar de que la Pro y la Gly estaban incrementadas en aquellos péptidos con Ala2, y la Ser y Arg mostraron una tendencia a estar incrementadas en los péptidos con Pro2, no se detectaron diferencias significativas, descartando la presencia de subpeptidomas como en los otros alotipos descritos previamente (**Figura 37**).

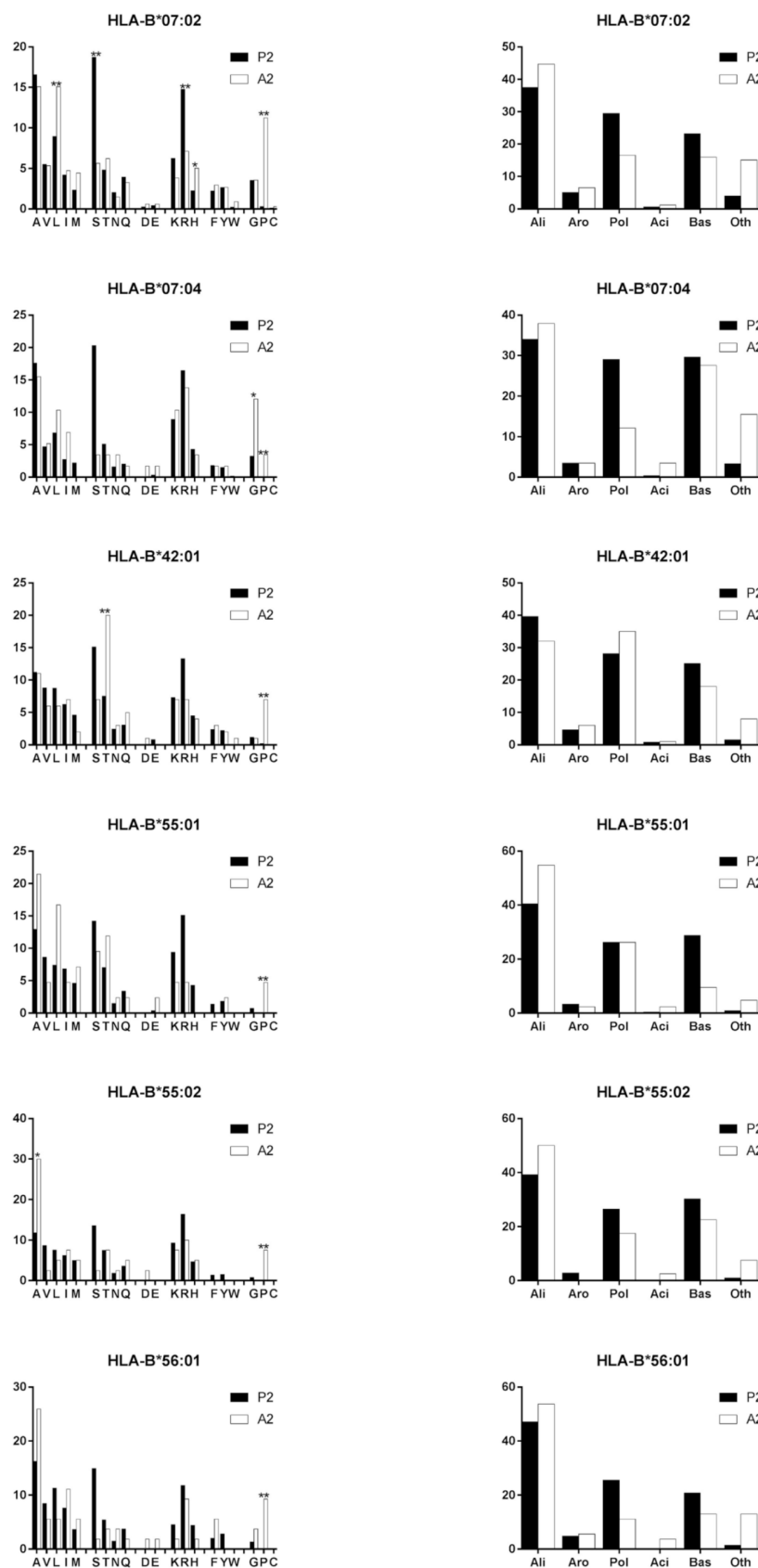


Figura 37: Residuos localizados en P1 de los ligandos de HLA-B\*07:02, HLA-B\*07:04, HLA-B\*42:01, HLA-B\*55:01, HLA-B\*55:02 y HLA-B\*56:01 con Prolina en P2 (P2) y Alanina en P2 (A2). Se utiliza el código de una sola letra para identificar los aminoácidos. En las gráficas de la izquierda, se muestra la frecuencia de cada residuo en posición 1 según si la posición 2 está ocupada por Prolina (negro) o Alanina (blanco). Las diferencias significativas se determinan mediante un test Chi cuadrado y la corrección de Bonferroni como \* < 0.05 y \*\* < 0.01. En las gráficas de la derecha, los residuos se agrupan según sus características fisicoquímicas.

#### **2.4. La presencia de Thr45 y de Phe67 podría asociarse a un incremento de la presencia de péptidos con Ala2 y de subpeptidomas en moléculas del supertipo B7**

El análisis de las secuencias de las diferentes moléculas de HLA-B pertenecientes al supertipo B7 mostró que aquellas que estaban desfavorecidas en la presencia de péptidos con Ala2 y no presentaban subpeptidomas contenían Glu45, mientras que aquellas donde los péptidos con Ala2 estaban favorecidos y sí presentaban subpeptidomas contenían Thr45, excepto HLA-B\*54:01 que contenía Gly45. Además, la posición 67 se encontraba ocupada por una Tyr en las moléculas con Glu45, mientras que en aquellas con Thr45 contenían Phe67, excepto HLA-B\*54:01 que, junto con Gly45, contenía Tyr67 (**Figura 33**) (**Tabla 11**). Para explicar estructuralmente el efecto de estos polimorfismos, se analizaron 3 cristales de HLA-B\*51:01, 18 cristales de HLA-B\*35:01 y 12 cristales de HLA-B\*07:02. En los cristales de la molécula de HLA-B\*07:02 se vio que Glu45 interaccionaba directamente con Tyr67 y Asn63 (residuo no polimórfico presente en todos los alotipos analizados) (**Figura 38A**). Esta interacción no tenía lugar cuando los residuos que ocupan las posiciones **45** y **67** eran Thr y Phe, respectivamente, como es el caso de las moléculas de HLA-B\*35:01 y HLA-B\*51:01 (**Figura 38B** y **Figura 38C**, respectivamente). La interacción de Glu45 con Asn63 cambia la orientación de este último residuo comparado con la orientación que toma en los alotipos con subpeptidomas, lo que provoca que el *pocket* sea mucho más rígido y limita la unión de residuos diferentes a Pro en *P2*.

| Molécula de HLA | Posición |    |
|-----------------|----------|----|
|                 | 45       | 67 |
| B*07:02         | E        | Y  |
| B*07:04         | -        | -  |
| B*42:01         | -        | -  |
| B*55:01         | -        | -  |
| B*55:02         | -        | -  |
| B*56:01         | -        | -  |
| B*35:01         | T        | F  |
| B*35:03         | T        | F  |
| B*35:07         | T        | F  |
| B*53:01         | T        | F  |
| B*51:01         | T        | F  |
| B*51:08         | T        | F  |
| B*54:01         | G        | -  |

Tabla 11: Residuos en posiciones 45 y 67 en las moléculas de HLA-I analizadas

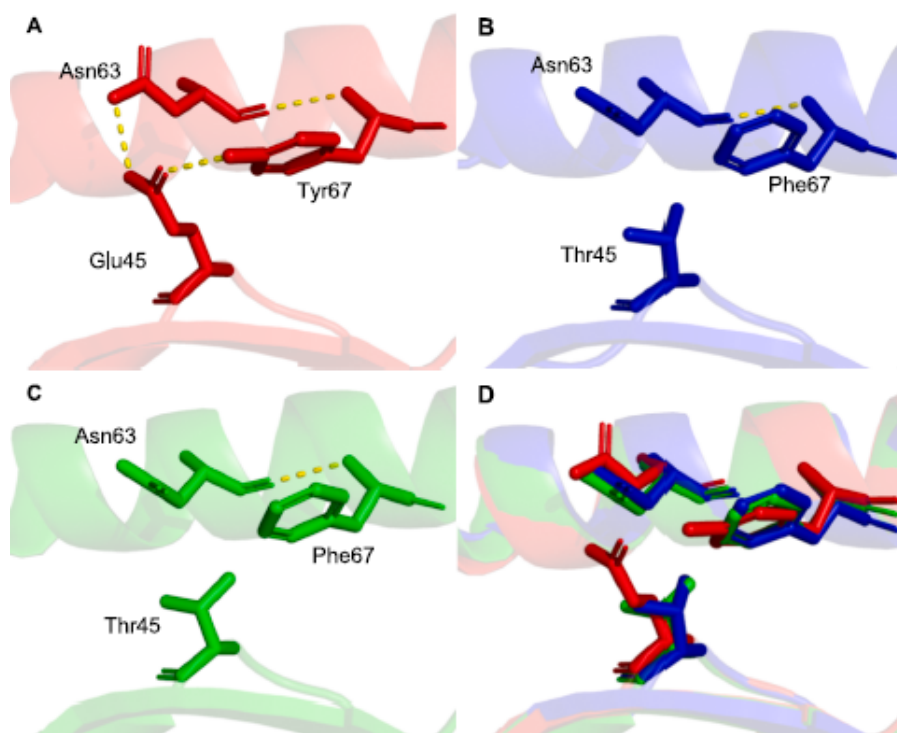


Figura 38: Representación de los residuos 45, 63 y 67 en: A: HLA-B\*07:02, -B: B\*35:01 y C: -B\*51:01. D: superposición de HLA-B\*07:02 (rojo), -B\*35:01 (azul) y -B\*51:01 (verde). Los residuos 45, 63 y 67 se marcan como “sticks”. Las uniones polares se representan en forma de línea discontinua amarilla. Entradas PDB: 1E27, 1A1N y 6UJ7, respectivamente.

Por tanto, el análisis de los repertorios peptídicos muestra la presencia de 3 grupos de alotipos con diferentes patrones entre los inmunopeptidomas de las moléculas del supertipo B7:

1. Los alotipos **HLA-B\*07:02**, **-B\*07:04**, **-B\*42:01**, **-B\*55:01**, **-B\*55:02** y **-B\*56:01** presentan inmunopeptidomas con péptidos con Ala2 desfavorecidos y sin diferencias en la posición *P1* entre los péptidos con Pro2 o Ala2.
2. Los alotipos **HLA-B\*51:01**, **-B\*51:08** y **-B\*54:01** presentan inmunopeptidomas con péptidos con Ala2 favorecidos y dos subppetidomas: uno de ellos donde los residuos con Ala2 están favorecidos con residuos ácidos en *P1* (así como residuos polares), y otro donde los residuos con Pro2 están favorecidos con residuos hidrofóbicos.
3. Los alotipos **HLA-B\*35:01**, **-B\*35:03**, **-B\*35:07** y **-B\*53:01** presentan inmunopeptidomas con péptidos con Ala2 favorecidos y dos subppetidomas: uno de ellos donde los residuos con Ala2 están favorecidos con residuos polares en *P1* (así como residuos ácidos), y otro donde los residuos con Pro2 están favorecidos con residuos hidrofóbicos.

Estos dos últimos grupos podrían agruparse en un único grupo, aunque presentan algunas diferencias.

## **ANÁLISIS DE LA APORTACIÓN DE LAS POSICIONES 45 Y 67 A LA PRESENCIA DE Ala2 Y DE SUBPEPTIDOMAS EN EL REPERTORIO PEPTÍDICO DE MOLÉCULAS DEL SUPERTIPO B7**

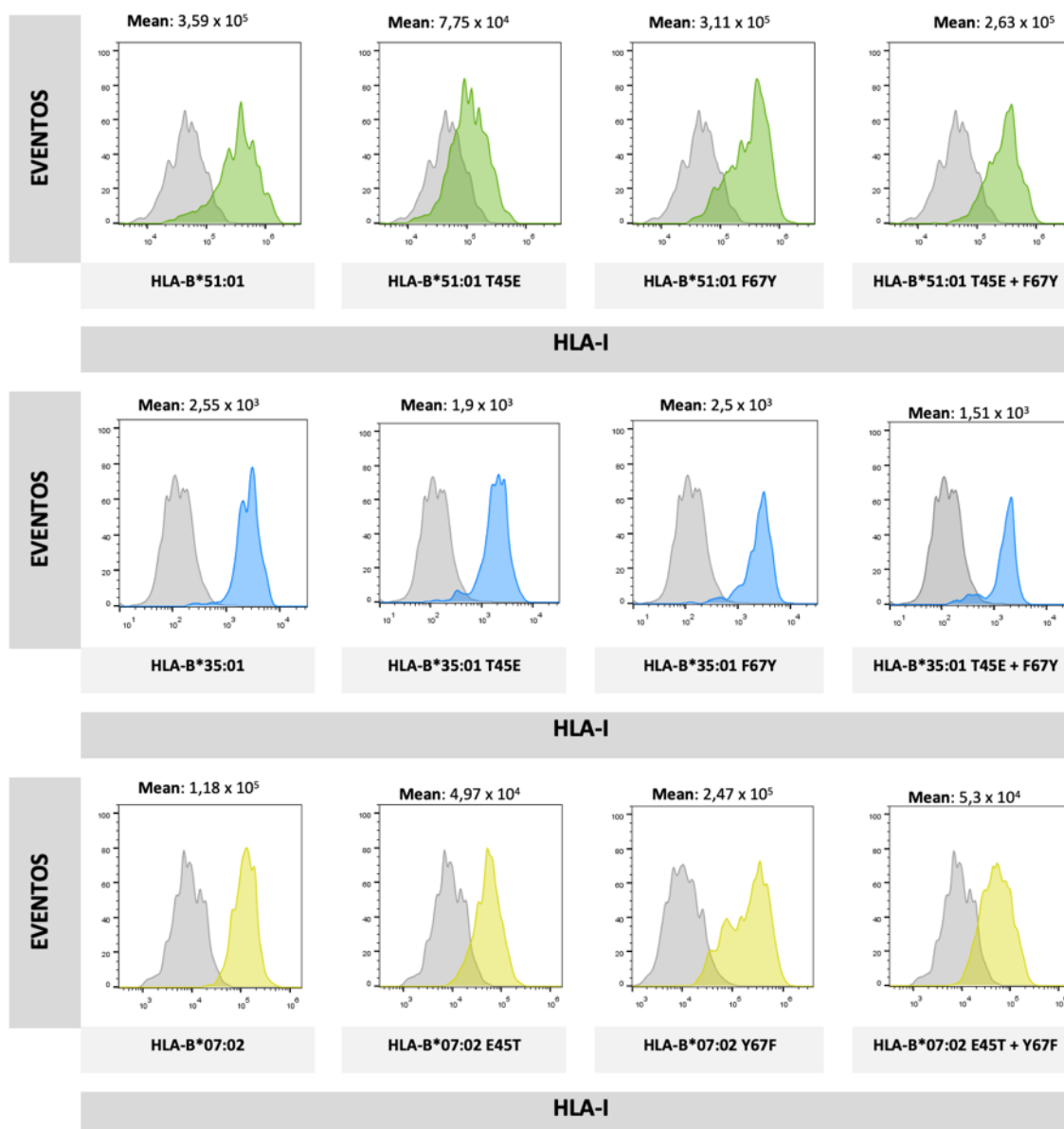
Para demostrar experimentalmente la importancia de los residuos localizados en las posiciones 45 y 67 en el incremento de péptidos con Ala2 en el inmunopeptidoma y analizar su papel en la presencia de subpeptidomas, se escogieron tres moléculas de HLA-I del supertipo B7 representativas de cada grupo descrito en el apartado anterior: **HLA-B\*51:01**, **-B\*35:01** y **-B\*07:02**, y se mutaron las posiciones 45 y 67 para estudiar el efecto sobre los inmunopeptidomas presentados.

Concretamente, las moléculas HLA-B\*51:01 y HLA-B\*35:01 fueron mutadas mediante mutagénesis dirigida para sustituir la Thr en posición 45 por Glu (**T45E**), mientras que la molécula HLA-B\*07:02 se mutó para sustituir el Glu en posición 45 por Thr (**E45T**). Por su parte, la posición 67 fue mutada para sustituir Phe por Tyr en las moléculas HLA-B\*51:01 y HLA-B\*35:01 (**F67Y**) y la Tyr por Phe en la molécula HLA-B\*07:02 (**Y67F**). Asimismo, se generaron los dobles mutantes (**T45E+F67Y**, en HLA-B\*51:01 y HLA-B\*35:01 y **E45T+Y67F**, en HLA-B\*07:02).



**1. GENERACIÓN DE MODELOS CELULARES PARA EVALUAR LA RELEVANCIA DE LOS RESIDUOS LOCALIZADOS EN POSICIONES 45 Y 67 EN LA CONFIGURACIÓN DEL REPERTORIO PEPTÍDICO EN LAS MOLÉCULAS HLA-B\*51:01, -B\*35:01 Y -B\*07:02**

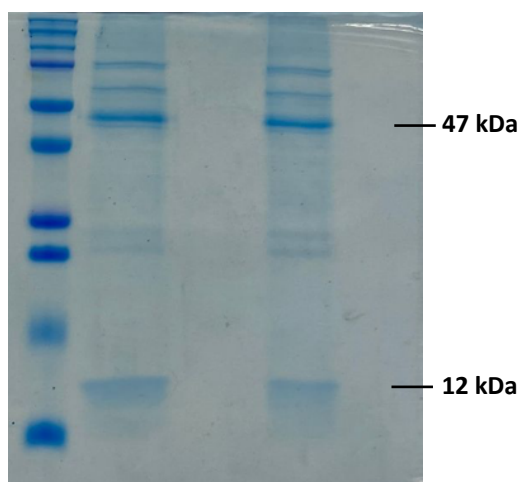
Los genes de cada molécula y sus mutantes se clonaron en el vector de expresión en eucariotas pcDNA3.1 junto con el “tag” Strep-tag II y se transfectó en la línea HMy2.C1R (C1R), que tiene muy baja expresión de moléculas de HLA-I endógenas. Las células positivas se seleccionaron con G418. Cuando fue necesario, se realizó una dilución límite para obtener clones puros. Los niveles de expresión en la superficie celular se midieron por citometría de flujo con el anticuerpo W6/32 (**Figura 39**). Todos los transfectantes mostraron un alto nivel de expresión, excepto el transfectante C1R-B\*51:01 T45E, del cual no se pudo obtener un transfectante con alta expresión a pesar de realizar múltiples intentos.



**Figura 39: Expresión de los diferentes alelos de HLA-I con y sin mutaciones puntuales en la línea celular C1R.** Análisis por citometría de flujo de la expresión de los diferentes alelos de HLA-B de clase I transfectados en la línea celular C1R con el anticuerpo W6/32 (anti-HLA-A, -B, -C). Se toma como control la línea celular C1R sin transfectar (gris). Los histogramas verdes hacen referencia al alelo HLA-B\*51:01 junto con los tres mutantes (T45E, F67Y, T45E + F67Y), el color azul hace referencia al alelo HLA-B\*35:01 junto con los tres mutantes (T45E, F67Y, T45E + F67Y) y el color amarillo hace referencia al alelo HLA-B\*07:02 junto con los tres mutantes (E45T, Y67F, E45T + Y67F). La **Mean** hace referencia a la media de intensidad de fluorescencia asociada a cada marcador.

## 2. PURIFICACIÓN DE LOS PÉPTIDOS DE UNIÓN A HLA-I DE LOS DIFERENTES TRANSFECTANTES

Se purificaron los ligandos peptídicos de las diferentes moléculas de HLA-B mediante cromatografía de afinidad, elución ácida y ultrafiltración. Para comprobar la pureza de la purificación, se realizó un gel de acrilamida al 16% con la fracción de la purificación correspondiente al HLA-I (**Figura 40**).



**Figura 40: Purificación de los péptidos ligandos de HLA-I de C1R-HLA-B\*51:01.** La purificación del inmunopeptidoma se llevó a cabo mediante una columna de afinidad, elución ácida y ultrafiltración. Se cargaron 10 µg de la fracción correspondiente al HLA-I de dicha ultrafiltración en un gel de acrilamida al 16% y se tiñó con *Coomassie Blue*. Las bandas correspondían a la cadena pesada  $\alpha$  (47 kDa) y a la cadena ligera  $\beta$ 2-microglobulina (12 kDa).

### **3. ANÁLISIS DE LOS INMUNOPEPTIDOMAS DE LAS MOLÉCULAS HLA-B\*51:01, -B\*35:01 Y -B\*07:02 Y DE SUS VARIANTES MUTADAS EN POSICIONES 45 Y 67**

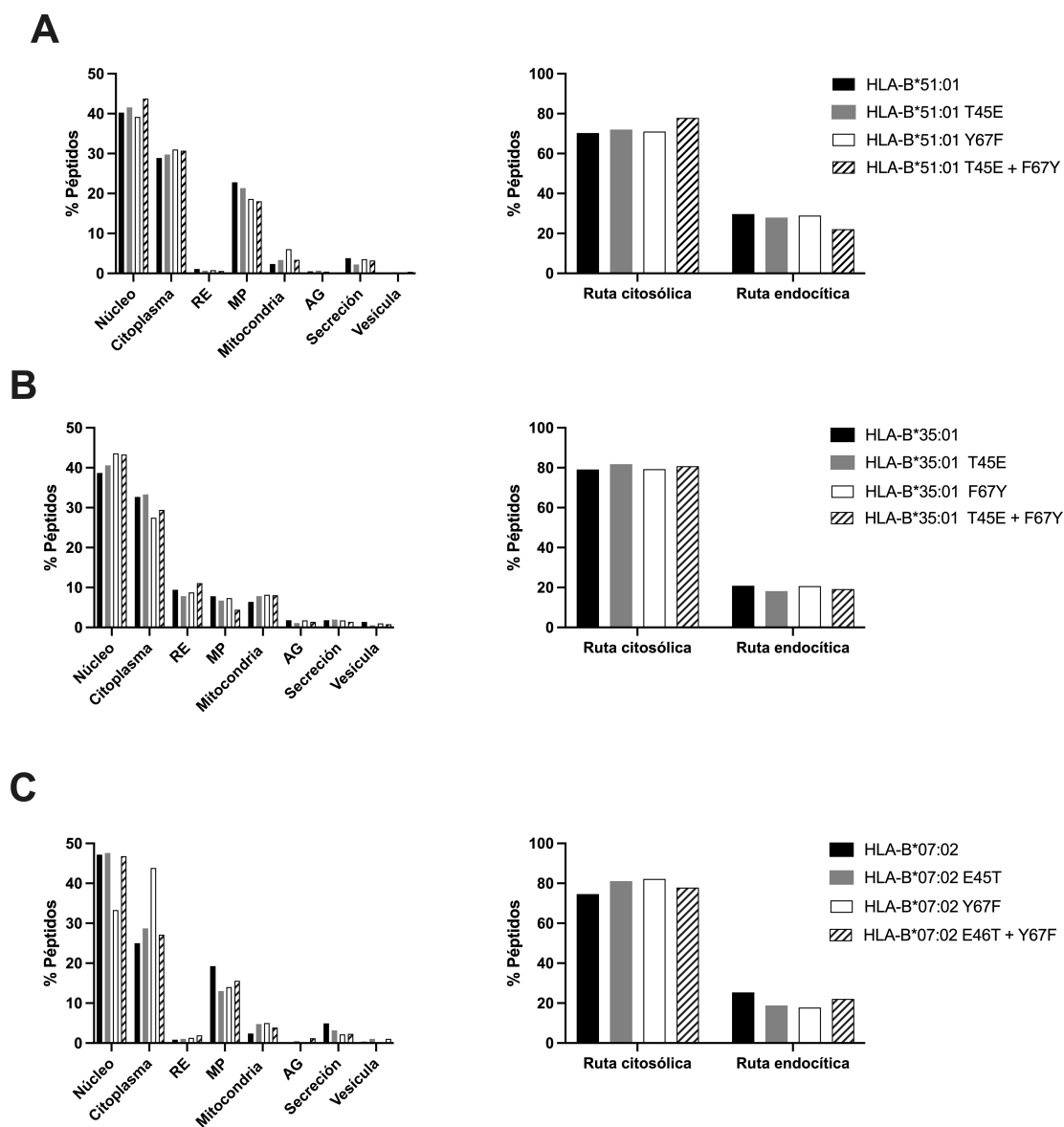
Se tuvieron en cuenta aquellos péptidos de 8-12 residuos de longitud, ya que los péptidos de 13 residuos identificados constituían <1% y muy probablemente se trataba de contaminantes. Bajo esta premisa, los peptidomas obtenidos de los diferentes transfectantes estaban compuestos por un número de péptidos entre 504 y 1478, excepto el peptidoma de HLA-B\*51:01 T45E, compuesto por 251 péptidos ([Tabla 12](#)). El análisis del número de residuos de cada uno de los péptidos obtenidos reveló que la mayoría de los péptidos eran nonámeros (55,5% - 82,2%) excepto HLA-B\*51:01 T45E, HLA-B\*51:01 F67Y y HLA-B\*51:01 T45E + F67Y, donde el porcentaje de octámeros (56,0%, 45,5% y 52,8%, respectivamente) fue superior que el de nonámeros (35,7%, 40,5% y 43,8%, respectivamente). La molécula HLA-B\*51:01 mostró el menor porcentaje de péptidos de 11 residuos de longitud (0,4%) frente a la molécula HLA-B\*35:01 F67Y que presentó el mayor porcentaje de péptidos de esta longitud (7,6%). La molécula HLA-B\*51:01 T45E + F67Y mostró el menor porcentaje de péptidos de 12 residuos de longitud (0,2%), mientras que la molécula HLA-B\*51:01 F67Y mostró el mayor porcentaje de estos péptidos (3,7%). El número medio de residuos de los péptidos que componían cada peptidoma fue similar en todas las moléculas de HLA-I analizadas, entre un 9,1% y 9,5%, excepto los transfectantes de HLA-B\*51:01, en los que el número medio de aminoácidos fue de entre 8,5% y 8,8% ([Tabla 12](#)).

| HLA-B*            | Total | Péptidos por número de residuos (%) |      |      |      |      | AAs promedio |
|-------------------|-------|-------------------------------------|------|------|------|------|--------------|
|                   |       | 8                                   | 9    | 10   | 11   | 12   |              |
| 51:01             | 504   | 39,9                                | 55,4 | 3,8  | 0,4  | 0,6  | 8,7          |
| 51:01 T45E        | 251   | 56,0                                | 35,7 | 4,0  | 2,8  | 1,2  | 8,7          |
| 51:01 F67Y        | 1493  | 45,5                                | 40,5 | 5,6  | 4,7  | 3,7  | 8,8          |
| 51:01 T45E + F67Y | 1298  | 52,8                                | 43,8 | 2,5  | 0,9  | 0,2  | 8,5          |
| 35:01             | 1172  | 6,5                                 | 69,9 | 14,0 | 6,9  | 2,7  | 9,3          |
| 35:01 T45E        | 771   | 4,1                                 | 82,2 | 8,7  | 4,0  | 0,9  | 9,5          |
| 35:01 F67Y        | 1232  | 3,6                                 | 72,5 | 14,4 | 7,6  | 2,0  | 9,3          |
| 35:01 T45E + F67Y | 833   | 6,5                                 | 74,0 | 11,8 | 5,7  | 2,0  | 9,2          |
| 07:02             | 1460  | 14,1                                | 65,4 | 14,1 | 5,3  | 1,1  | 9,1          |
| 07:02 E45T        | 943   | 10,3                                | 70,0 | 13,0 | 5,9  | 0,7  | 9,2          |
| 07:02 Y67F        | 637   | 10,2                                | 72,5 | 11,3 | 5,02 | 0,94 | 9,1          |
| 07:02 T45E + Y67F | 1478  | 16,0                                | 65,8 | 11,6 | 5,8  | 0,9  | 9,1          |

Tabla 12: Tamaño, por número de aminoácidos, de los inmunopeptidomas de HLA-I obtenidos por espectrometría de masas de las moléculas pertenecientes al supertipo B7 analizadas y los mutantes de estas.

### 3.1. Localización celular

La localización de las proteínas parentales se definió mediante la búsqueda de estas en la base de datos Uniprot. En los casos en que las proteínas aparecían en más de un compartimento celular, se seleccionó aquel más común de acuerdo con su función biológica. La mayoría de las proteínas parentales provenían principalmente del citoplasma y del núcleo, y, consecuentemente, correspondían a la ruta endocítica, entre un 74% y 82% (Figura 41).



**Figura 41: Localización celular de las proteínas parentales de las que proceden los péptidos de:** **A:** HLA-B\*51:01 (negro), HLA-B\*51:01 T45E (gris), HLA-B\*51:01 F67Y (blanco) y HLA-B\*51:01 T45E + F67Y (rayado). **B:** HLA-B\*35:01 (negro), HLA-B\*35:01 T45E (gris), HLA-B\*35:01 F67Y (blanco) y HLA-B\*35:01 T45E + F67Y (rayado). **C:** HLA-B\*07:02 (negro), HLA-B\*07:02 E45T (gris), HLA-B\*07:02 Y67F (blanco) y HLA-B\*07:02 E45T + Y67F (rayado). A la izquierda se muestra el porcentaje de péptidos según el compartimento celular. RE: retículo endoplasmático, MP: membrana plasmática, AG: aparato de Golgi. A la derecha se muestra el porcentaje de péptidos según si pertenecen a la ruta citoplasmática (citoplasma, núcleo y mitocondrias) o endocítica (MP, RE, AG, vesículas y proteínas de secreción).

### 3.2. Grado de solapamiento de los inmunopeptidomas

Al igual que en el análisis *in silico*, el grado de solapamiento de los diferentes repertorios peptídicos se analizó por pares de moléculas. Los 12 inmunopeptidomas analizados estaban compuestos por un total de 12074 péptidos (Tabla 13).

|                         | 504           | 251           | 1493          | 1298                       | 1172          | 771           | 1232          | 833                        | 1460          | 943           | 637           | 1478                       |
|-------------------------|---------------|---------------|---------------|----------------------------|---------------|---------------|---------------|----------------------------|---------------|---------------|---------------|----------------------------|
| Alotipo<br>HLA          | 51:01         | 51:01<br>T45E | 51:01<br>F67Y | 51:01<br>T45E<br>+<br>F67Y | 35:01         | 35:01<br>T45E | 35:01<br>F67Y | 35:01<br>T45E<br>+<br>F67Y | 07:02         | 07:02<br>E45T | 07:02<br>Y67F | 07:02<br>E45T<br>+<br>Y67F |
| 51:01                   |               | 78            | 283           | 257                        | 6             | 1             | 5             | 3                          | 9             | 14            | 4             | 26                         |
| 51:01<br>T45E           | 15,5/<br>31,0 |               | 176           | 120                        | 3             | 3             | 0             | 2                          | 7             | 2             | 0             | 3                          |
| 51:01<br>F67Y           | 19,0/<br>56,2 | 70,1/<br>11,8 |               | 519                        | 7             | 1             | 6             | 2                          | 18            | 32            | 16            | 48                         |
| 51:01<br>T45E +<br>F67Y | 19,8/<br>51,0 | 47,8/<br>9,2  | 40,0/<br>34,8 |                            | 5             | 2             | 10            | 1                          | 28            | 26            | 14            | 39                         |
| 35:01                   | 0,5/<br>1,2   | 0,3/<br>1,2   | 0,6/<br>0,5   | 0,4/<br>0,4                |               | 366           | 469           | 478                        | 40            | 57            | 27            | 87                         |
| 35:01<br>T45E           | 0,1/<br>0,2   | 0,4/<br>1,2   | 0,1/<br>0,1   | 0,3/<br>0,2                | 47,5/<br>31,2 |               | 420           | 385                        | 39            | 22            | 21            | 26                         |
| 35:01<br>F67Y           | 0,4/<br>1,0   |               | 0,5/<br>0,4   | 0,8/<br>0,8                | 38,1/<br>40,0 | 28,9/<br>46,2 |               | 356                        | 114           | 47            | 39            | 80                         |
| 35:01<br>T45E +<br>F67Y | 0,4/<br>0,6   | 0,2/<br>0,8   | 0,2/<br>0,1   | 0,1/<br>0,1                | 57,4/<br>40,8 | 46,2/<br>46,2 | 42,7/<br>28,9 |                            | 24            | 41            | 17            | 57                         |
| 07:02                   | 0,6/<br>1,8   | 0,5/<br>2,8   | 1,2/<br>1,2   | 1,9<br>2,2                 | 2,7/<br>3,4   | 2,7/<br>5,1   | 7,8/<br>9,3   | 1,6/<br>2,9                |               | 253           | 250           | 334                        |
| 07:02<br>E45T           | 1,5/<br>2,8   | 0,2/<br>0,8   | 3,3/<br>2,1   | 2,8/<br>2,0                | 6,0/<br>4,9   | 2,3/<br>2,9   | 5,0/<br>3,8   | 4,4/<br>4,9                | 26,8/<br>17,3 |               | 334           | 623                        |
| 07:02<br>Y67F           | 0,6/<br>0,8   |               | 2,5/<br>1,1   | 2,2/<br>1,1                | 4,2/<br>2,3   | 3,3/<br>2,7   | 6,1/<br>3,2   | 2,7/<br>2,0                | 39,2/<br>17,1 | 52,4/<br>35,4 |               | 378                        |
| 07:02<br>E45T +<br>Y67F | 1,8/<br>5,1   | 0,2/<br>1,2   | 3,2/<br>3,2   | 2,6/<br>3,0                | 5,9/<br>7,4   | 1,7/<br>3,4   | 5,4/<br>6,5   | 3,9/<br>6,5                | 22,6/<br>22,8 | 42,2/<br>66,1 | 25,6/<br>59,3 |                            |

Tabla 13: Solapamiento de los inmunopeptidomas de las moléculas HLA-B\*51:01, HLA-B\*35:01 y HLA-B\*07:02 y sus mutantes correspondientes.

Nota: Por encima de la línea diagonal se muestran los números absolutos de los péptidos compartidos. Debajo de la diagonal, el porcentaje de los péptidos compartidos (el número encima de la barra hace referencia al alotipo en la izquierda, y el número debajo de la barra hace referencia al alotipo indicado arriba)

Nota: Los números sombreados en gris claro indican que los péptidos compartidos más de un 10% del peptidoma total en al menos un alotipo. El sombreado gris oscuro indica que los péptidos compartidos representan más de un 20% del peptidoma total en al menos un alotipo.

De acuerdo con lo esperado, los diferentes peptidomas correspondientes a un mismo grupo alélico de HLA presentaron mayor grado de solapamiento: HLA-B\*51:01 con sus mutantes correspondientes HLA-B\*51:01 T45E, HLA-B\*51:01 F67Y y HLA-B\*51:01 T45E + F67Y; HLA-B\*35:01 con HLA-B\*35:01 T45E, HLA-B\*35:01 F67Y y HLA-B\*35:01 T45E + F67Y y HLA-B\*07:02 con HLA-B\*07:02 E45T, HLA-B\*07:02 Y67F y HLA-B\*07:02 E45T + Y67F (**Tabla 13**).

### 3.3. Residuos de anclaje

Los residuos de anclaje de los inmunopeptidomas de los diferentes transfectantes se analizaron con el *software* GibbsCluster – 2.0<sup>196</sup>, que presenta la ventaja de que no requiere seleccionar las moléculas sobre las que predecir el motivo de unión. Esta característica permitió predecir los residuos de anclaje de las moléculas mutadas, ya que en otros *softwares* es necesario seleccionar la molécula y no permiten añadir moléculas mutadas.

El análisis de los residuos de anclaje de las moléculas HLA-B\*51:01, HLA-B\*35:01 y HLA-B\*07:02 mostraron que la P2 estaba predominantemente ocupada por Pro, seguido de Ala. La P2 de los diferentes transfectantes también estaba ocupada principalmente por Pro, sin embargo, el porcentaje de Ala2 varió entre la molécula nativa y las moléculas mutadas en las diferentes posiciones. La posición P2 del transfectante HLA-B\*51:01 T45E no estaba muy definida, sin embargo, probablemente se debiera al escaso número de péptidos eluidos a partir de esta línea celular (**Figura 42**) (**Tabla 14**).

Asimismo, los residuos de anclaje se calcularon con el *software* PRBAM, obteniendo resultados muy similares a los descritos con el GibbsCluster – 2.0, aportando robustez a los mismos (**Anexo 8**). Se tuvieron en cuenta solo los no números para este análisis. Los resultados obtenidos muestran como la presencia de Ala2 disminuye en los mutantes simples y doble de las moléculas HLA-B\*51:01 y -B\*35:01 respecto la molécula nativa, mientras que en -B\*07:02 aumenta.



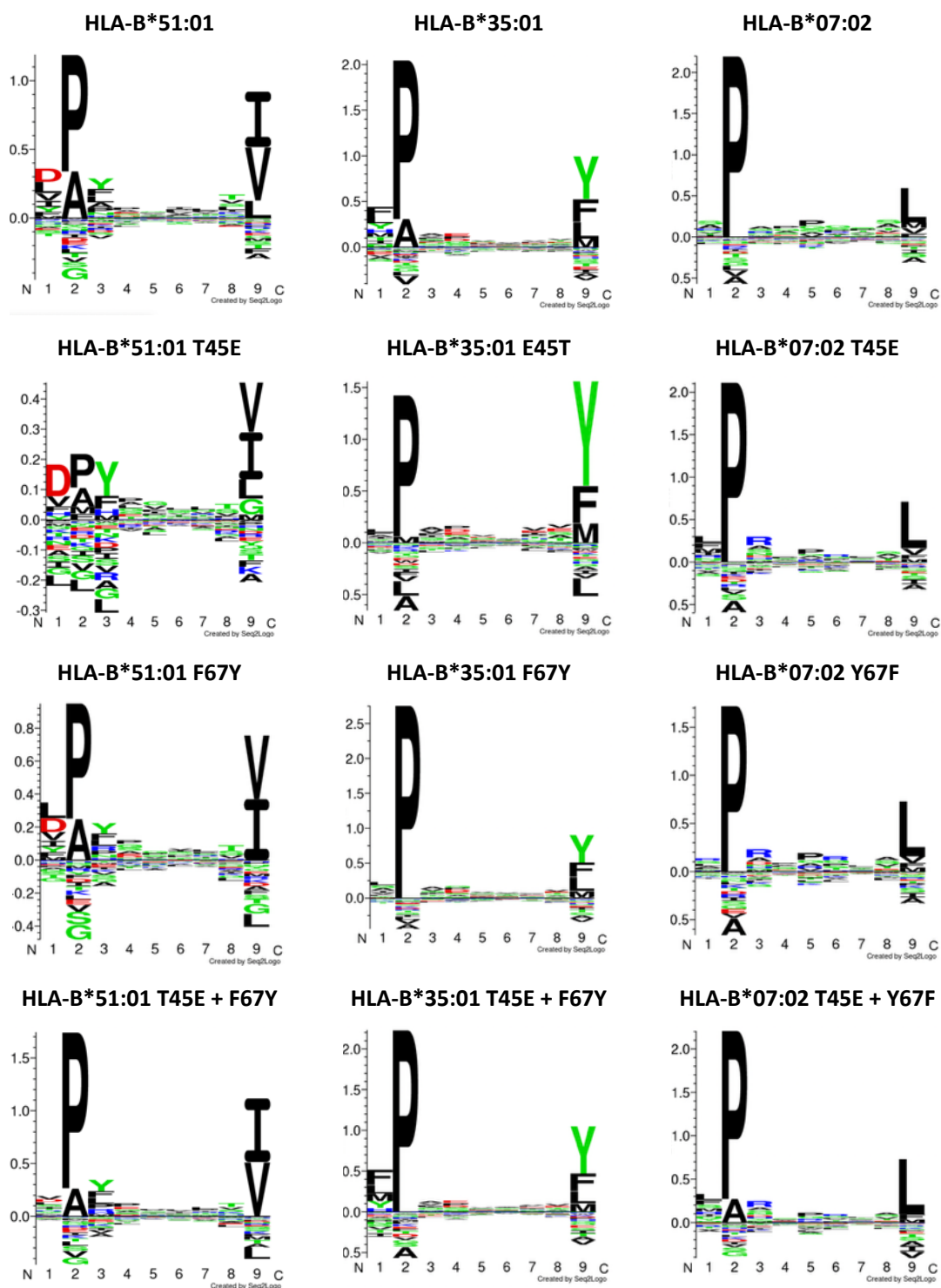


Figura 42: Logo de secuencias de los peptidomas de HLA-B\*51:01, HLA-B\*35:01 y HLA-B\*07:02 y sus variantes mutadas. Los logos de secuencia se generaron con Seq2Logo.

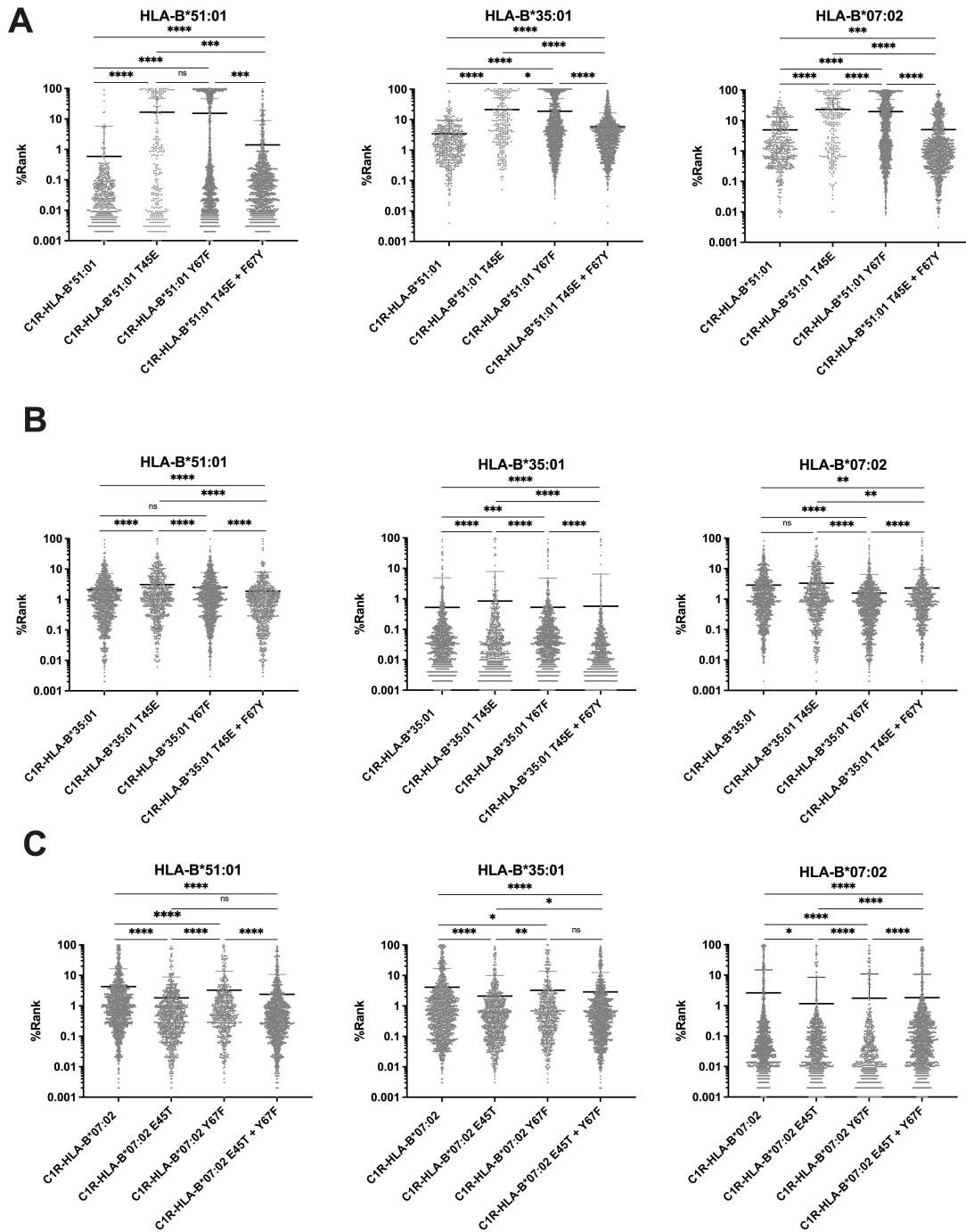
| HLA-B*                   | Total | Pro2         | Ala2         | Otros        |
|--------------------------|-------|--------------|--------------|--------------|
| <b>51:01 Literatura</b>  | 5157  | 2566 (49,8%) | 1552 (30,1%) | 1039 (20,1%) |
| <b>51:01</b>             | 504   | 345 (68,5%)  | 126 (25%)    | 33 (6,5%)    |
| <b>51:01 T45E</b>        | 252   | 96 (38,1%)   | 45 (17,9%)   | 111 (44,0%)  |
| <b>51:01 F67Y</b>        | 1494  | 749 (50,1%)  | 294 (19,7%)  | 451 (30,2%)  |
| <b>51:01 T45E + F67Y</b> | 1298  | 1010 (77,8%) | 160 (12,3%)  | 128 (9,9%)   |
| <b>35:01 Literatura</b>  | 1954  | 1322 (67,7%) | 297 (15,2%)  | 335 (17,1%)  |
| <b>35:01</b>             | 1172  | 963 (82,2%)  | 154 (13,1%)  | 55 (4,7%)    |
| <b>35:01 T45E</b>        | 771   | 630 (81,7%)  | 42 (5,4%)    | 99 (12,8%)   |
| <b>35:01 F67Y</b>        | 1232  | 1180 (95,8%) | 14 (1,1%)    | 38 (3,1%)    |
| <b>35:01 T45E + F67Y</b> | 833   | 780 (93,6%)  | 31 (3,7%)    | 22 (2,6%)    |
| <b>07:02 Literatura</b>  | 12033 | 9956 (82,7%) | 338 (2,8%)   | 1739 (14,5%) |
| <b>07:02</b>             | 1460  | 1283 (87,9%) | 44 (3,0%)    | 133 (9,1%)   |
| <b>07:02 E45T</b>        | 943   | 839 (89,0%)  | 44 (4,7%)    | 60 (6,4%)    |
| <b>07:02 Y67F</b>        | 637   | 558 (87,6%)  | 37 (5,8%)    | 42 (6,6%)    |
| <b>07:02 E45T + Y67F</b> | 1478  | 1240 (83,9%) | 139 (9,4%)   | 99 (6,7%)    |

Tabla 14: Preferencia de aminoácido en posición P2 en las moléculas HLA-B\*51:01, HLA-B\*35:01 y HLA-B\*07:02 y de los mutantes HLA-B\*51:01 T45E, HLA-B\*51:01 F67Y, HLA-B\*51:01 T45E + F67Y; HLA-B\*35:01 T45E, HLA-B\*35:01 F67Y, HLA-B\*35:01 T45E + F67Y Y HLA-B\*07:02 E45T, HLA-B\*07:02 Y67F, HLA-B\*07:02 E45T + Y67F.

### 3.4. Afinidad teórica

La **Figura 43** muestra la afinidad teórica de los péptidos eluidos de cada uno de los transfectantes para las moléculas HLA-B\*51:01, -B\*35:01 y -B\*07:02.

La afinidad teórica de los péptidos eluidos de HLA-B\*51:01 fue superior para HLA-B\*51:01 que para HLA-B\*35:01 y HLA-B\*07:02. Sin embargo, la afinidad teórica de los péptidos eluidos de los mutantes simples fue muy similar para los tres alotipos analizados, siendo mucho más baja en los tres casos. La afinidad teórica del doble mutante fue mayor que la de los mutantes simples pero menor que de la molécula nativa (**Figura 43A**). Se observó la misma tendencia en los péptidos eluidos de HLA-B\*35:01. La afinidad teórica fue superior para HLA-B\*35:01 que para -B\*51:01 y -B\*07:02, sin embargo, la afinidad teórica de los mutantes simples no fue tan drástica (**Figura 43B**). En contraposición, la afinidad teórica de los péptidos eluidos de los transfectantes HLA-B\*07:02 se mantuvo muy similar para las tres alotipos, siendo superior para HLA-B\*07:02 (**Figura 43C**).



**Figura 43: Afinidad teórica de las moléculas HLA-B\*51:01, -B\*35:01 y -B\*07:02 y los mutantes simples y dobles en posiciones 46 y 67. A.** Afinidad de los diferentes transfectantes de HLA-B\*51:01 para las moléculas HLA-B\*51:01, -B\*35:01 y -B\*07:02. **B:** Afinidad de los diferentes transfectantes de HLA-B\*35:01 para las moléculas HLA-B\*51:01, -B\*35:01 y -B\*07:02. **C:** Afinidad de los diferentes transfectantes de HLA-B\*07:02 para las moléculas HLA-B\*51:01, -B\*35:01 y -B\*07:02. Los símbolos representan péptidos individuales; las líneas horizontales representan la media \* = 0.01 < P < 0.1; \*\* = 0.001 < P < 0.01; \*\*\* = 0.0001 < P < 0.0001 según el test de Mann-Whitney U.

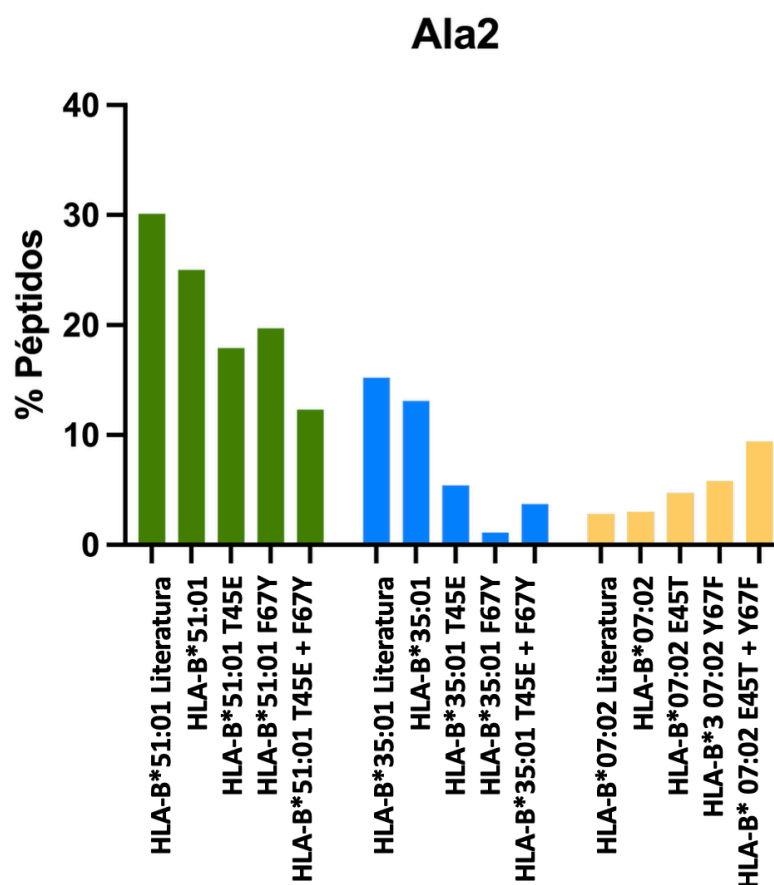
#### 4. IMPLICACIÓN DE LAS POSICIONES 45 Y 67 EN LA CONFORMACIÓN DEL REPERTORIO PEPTÍDICO DE LAS MOLÉCULAS HLA-B\*51:01, HLA-B\*35:01 Y HLA-B\*07:02

##### 4.1. Los residuos que ocupan las posiciones 45 y 67 afectan a la presencia de péptidos con Ala2 en los repertorios peptídicos de las moléculas de HLA-I analizadas

En primer lugar, se comparó el peptidoma obtenido por espectrometría de masas de las moléculas HLA-B\*51:01, HLA-B\*35:01 y HLA-B\*07:02 con el peptidoma de estas moléculas disponible en la literatura, con el fin de confirmar que los resultados eran similares y se habían purificado inmuno-peptidomas similares. La purificación de los ligandos de estas moléculas y su posterior análisis revelaron que la P2 se encontraba ocupada mayormente por Pro, y en menor medida por Ala, de manera muy similar a los ligandos descritos en la bibliografía. Según la literatura, el peptidoma de la molécula HLA-B\*51:01 presenta un 30,1% de péptidos con Ala2, mientras que, en nuestros resultados, este porcentaje se tradujo en un 25,0%. Por otro lado, el porcentaje de péptidos con Ala2 de los peptidomas de las moléculas HLA-B\*35:01 y HLA-B\*07:02 disponibles en la base de datos IEDB era de un 15,2% y 2,8%, respectivamente. Nuestros resultados mostraron un 13,1% y un 3,0% de péptidos con Ala2, respectivamente ([Tabla 14](#)), resultados muy similares en ambos casos.

El peptidoma de la molécula **HLA-B\*51:01** presentó un 25,0% de péptidos con Ala2. Este porcentaje disminuyó hasta un 17,9% cuando se mutó la posición 45 (T45E), a un 19,7% cuando se mutó la posición 67 (F67Y) y a un 12,3% en el doble mutante (T45E + F67Y). El peptidoma del alotipo **HLA-B\*35:01** presentó un 13,1% de péptidos con Ala2 que disminuyó hasta un 5,4% cuando se mutó la posición 45 (T45E), a un 1,1% cuando se mutó la posición 67 (F67Y) y a un 3,7% cuando se mutaron ambas (T45E+F67Y). Por su parte, el porcentaje de péptidos con Ala2 en **HLA-B\*07:02** fue de un 3,0%, que aumentó a un 4,7% cuando se mutó la posición 45 (E45T), un 5,8% cuando se mutó la posición 67 (Y67F) y un 9,4% cuando se mutaron ambas (E45T+Y67F), revelando que cuando las posiciones 45 y 67 estaban ocupadas por Thr y Phe respectivamente, la posición 2 era más laxa, permitiendo unir residuos diferentes a Pro ([Tabla 14](#)) ([Figura 44](#)). Por tanto, las mutaciones puntuales realizadas en las posiciones 45 y 67 produjeron cambios significativos en el porcentaje de péptidos con Ala2. Las mutaciones T45E, F67Y y

T45E+F67Y en las moléculas **HLA-B\*51:01** y **HLA-B\*35:01** provocaron una reducción significativa de la frecuencia de péptidos con Ala en P2. De manera inversa, la mutación de las posiciones E45T, Y67F, y E45T+Y67F en la molécula **HLA-B\*07:02** incrementó la proporción de péptidos con Ala2, lo que respalda el papel clave de estos residuos en la modulación de la especificidad del *pocket* B (**Tabla 14**).



**Figura 44: Frecuencia de péptidos con Ala2 de los inmunoepitopos de los diferentes transfectantes.** Se muestra el porcentaje de péptidos con Ala2 descritos en la literatura para las moléculas HLA-B\*51:01, B\*35:01 y B\*07:02, así como los transfectantes de estas moléculas y los diferentes mutantes. En verde se muestra HLA-B\*51:01, en azul HLA-B\*35:01 y en amarillo HLA-B\*07:02.

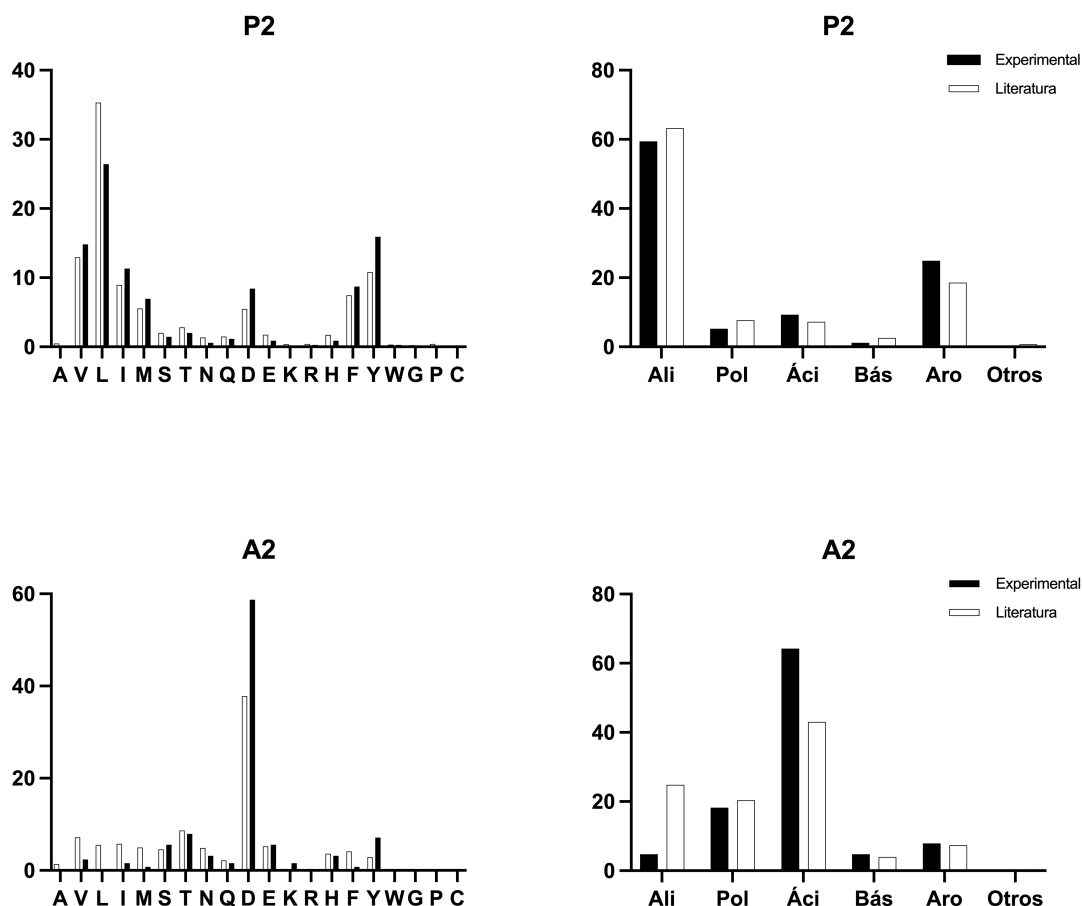
Estos resultados refuerzan las conclusiones obtenidas a partir de las estructuras cristalográficas (**Figura 38**), donde se mostró que la Thr45, presente en las moléculas HLA-B\*51:01 y HLA-B\*35:01 no interaccionaba con Phe67 y Asn63, lo que se traducía en un *pocket* más amplio y permisivo para la acomodación de residuos diferentes a Pro, principalmente Ala. De manera inversa, el residuo Glu45, presente en HLA-B\*07:02, interaccionaba con Tyr67 y Asn63, dando lugar a un *pocket* estrecho en la posición *P2* que restringía y limitaba la interacción de otros residuos que no fueran Pro2.

Al sustituir Thr45 por Glu45 en las moléculas HLA-B\*51:01 y HLA-B\*35:01, se observó que el *pocket* se volvía más restrictivo, similar al *pocket* de HLA-B\*07:02, lo que limitó la acomodación de residuos distintos a Pro en *P2*. Por el contrario, el cambio de Glu45 por Thr45 en HLA-B\*07:02, provocó que el *pocket* se volviera más amplio y permisivo, permitiendo la acomodación de otros residuos distintos a Pro. Estos resultados refuerzan la idea de que las interacciones entre las posiciones 45 y 67 son determinantes en la conformación del repertorio peptídico presentado por cada alelo.

#### **4.2. Los residuos que ocupan las posiciones 45 y 67 afectan a la presencia de subpeptidomas en diferente grado en los repertorios peptídicos de las moléculas de HLA-I analizadas**

Con el fin de confirmar experimentalmente la presencia de subpeptidomas en las moléculas HLA-B\*51:01, -B\*35:01 y -B\*07:02 se analizó la posición *P2* de los peptidomas identificados en esta tesis y su influencia sobre la posición *P1*. De nuevo, para aportar mayor robustez a los resultados obtenidos, estos peptidomas se compararon con los peptidomas descritos en la bibliografía.

El análisis del peptidoma de la molécula **HLA-B\*51:01** reveló que los péptidos con Pro2 mostraban un claro enriquecimiento en residuos aromáticos (F, Y) y alifáticos (V, L, I) en *P1*, mientras que los péptidos con Ala2 estaban enriquecidos en residuos ácidos (D) en *P1* (**Figura 45**). Por lo tanto, tal y como indica la bibliografía, la molécula HLA-B\*51:01 presentaba dos subpeptidomas diferenciados, uno de ellos formado por Pro2 y residuos alifáticos y aromáticos en *P1* y otro formado por Ala2 y residuos ácidos en *P1*.



**Figura 45: Residuos localizados en P1 de los ligandos de HLA-B\*51:01 con Prolina en P2 (P2) y Alanina en P2 (A2).** Se utiliza el código de una sola letra para identificar los aminoácidos. En las gráficas de la izquierda, se muestra la frecuencia de cada residuo en posición 1 según si la posición 2 está ocupada por Pro (arriba) o Ala (abajo) según los resultados obtenidos (negro) o los datos disponibles en la literatura (blanco). En las gráficas de la derecha, los residuos se agrupan según sus características fisicoquímicas.

Por su parte, el análisis del peptidoma de **HLA-B\*35:01** mostró que los péptidos en los que la **P2** se encontraba ocupada por Pro, la **P1** se encontraba favorecida por residuos aromáticos (F, Y) y alifáticos (V, L, I, M), mientras que en los péptidos con Ala2, la **P1** se encontraba favorecida por residuos polares (S, T, N, Q) y ácidos (D, E), demostrando experimentalmente la existencia de dos subpeptidomas en la molécula HLA-B\*35:01, uno de ellos formado por Pro2 y residuos aromáticos y alifáticos en **P1**, y otro formado por Ala2 y residuos polares y ácidos en **P1** (**Figura 46**), tal y como se observó previamente *in silico*.

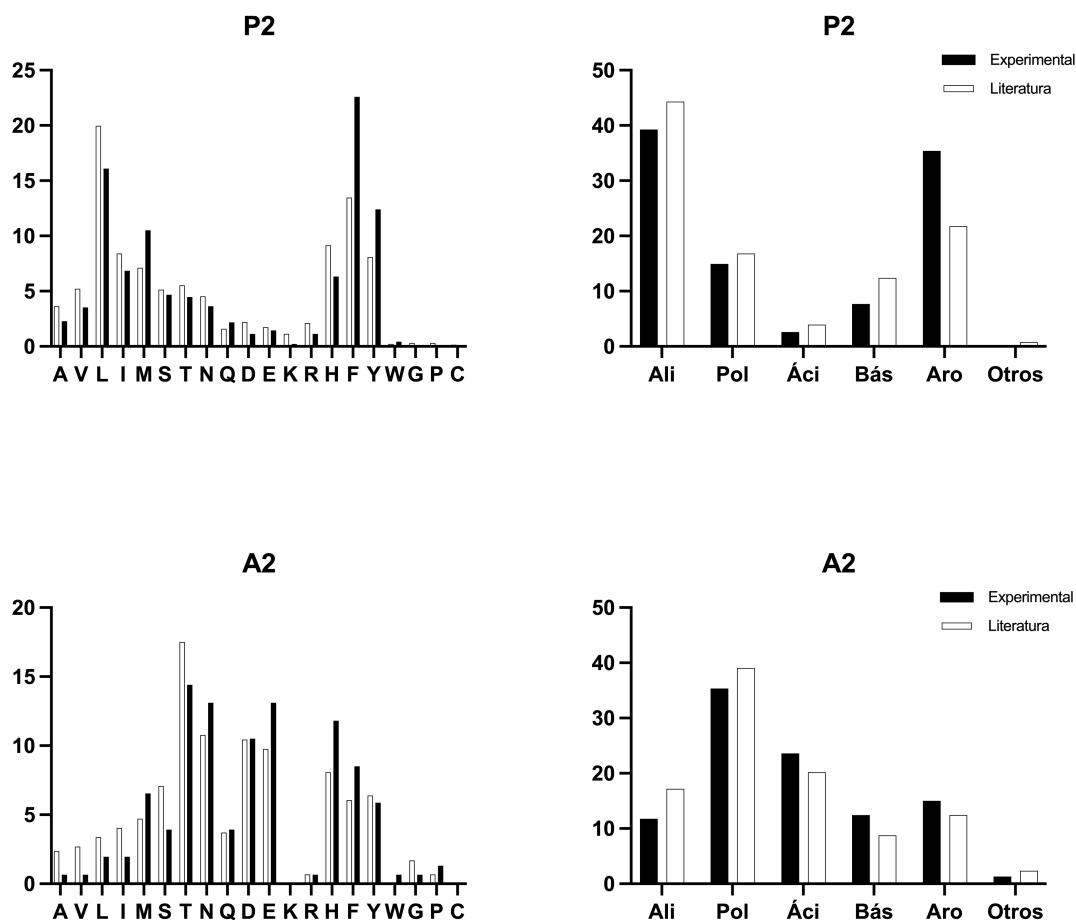
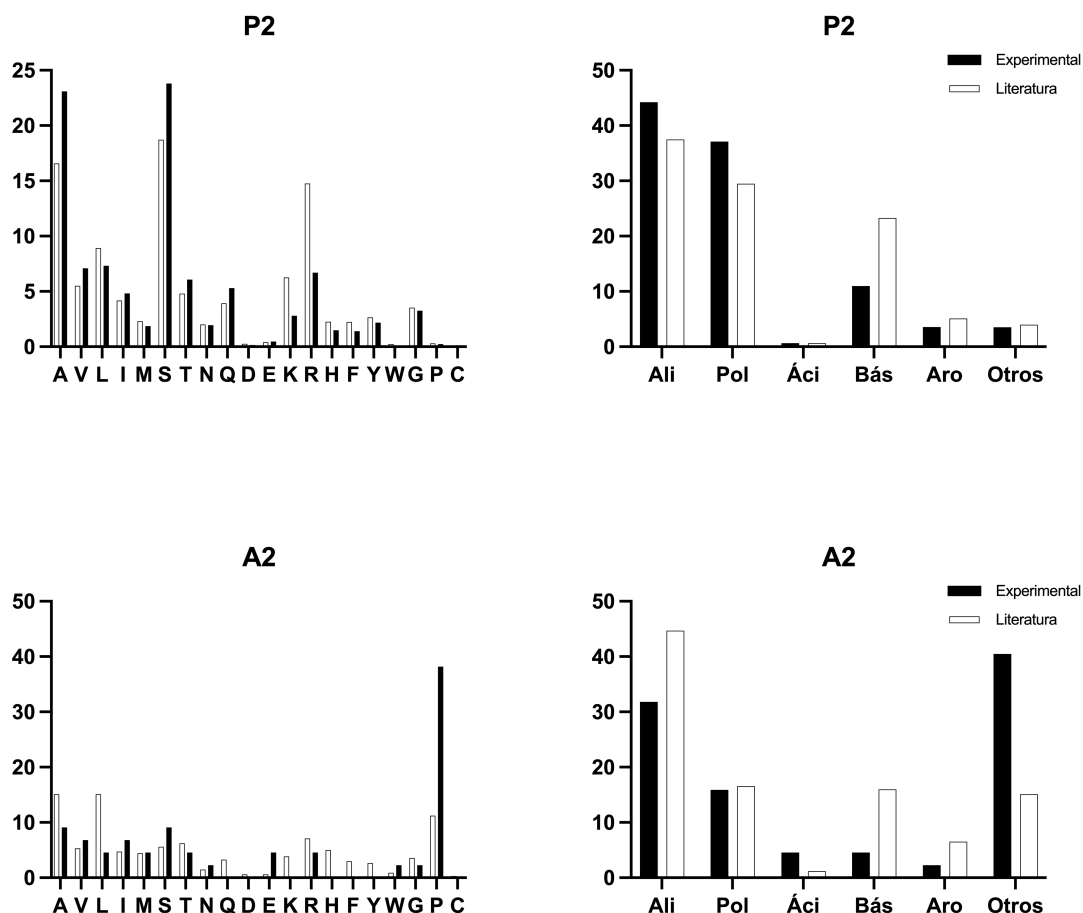


Figura 46: Residuos localizados en P1 de los ligandos de HLA-B\*35:01 con Prolina en P2 (P2) y Alanina en P2 (A2). Se utiliza el código de una sola letra para identificar los aminoácidos. En las gráficas de la izquierda, se muestra la frecuencia de cada residuo en posición 1 según si la posición 2 está ocupada por Pro (arriba) o Ala (abajo) según los resultados obtenidos (negro) o los datos disponibles en la literatura (blanco). En las gráficas de la derecha, los residuos se agrupan según sus características fisicoquímicas.

Por el contrario, la molécula **HLA-B\*07:02** no presentaba subpeptidomas. Aunque el análisis del peptidoma obtenido por espectrometría de masas reveló un incremento de Ala1 y Ser1 cuando la posición 2 estaba ocupada por Pro, no fue significativo, confirmando los resultados anteriores. Sin embargo, se detectó un incremento significativo de Pro1 cuando la posición 2 estaba ocupada por alanina (Figura 47).

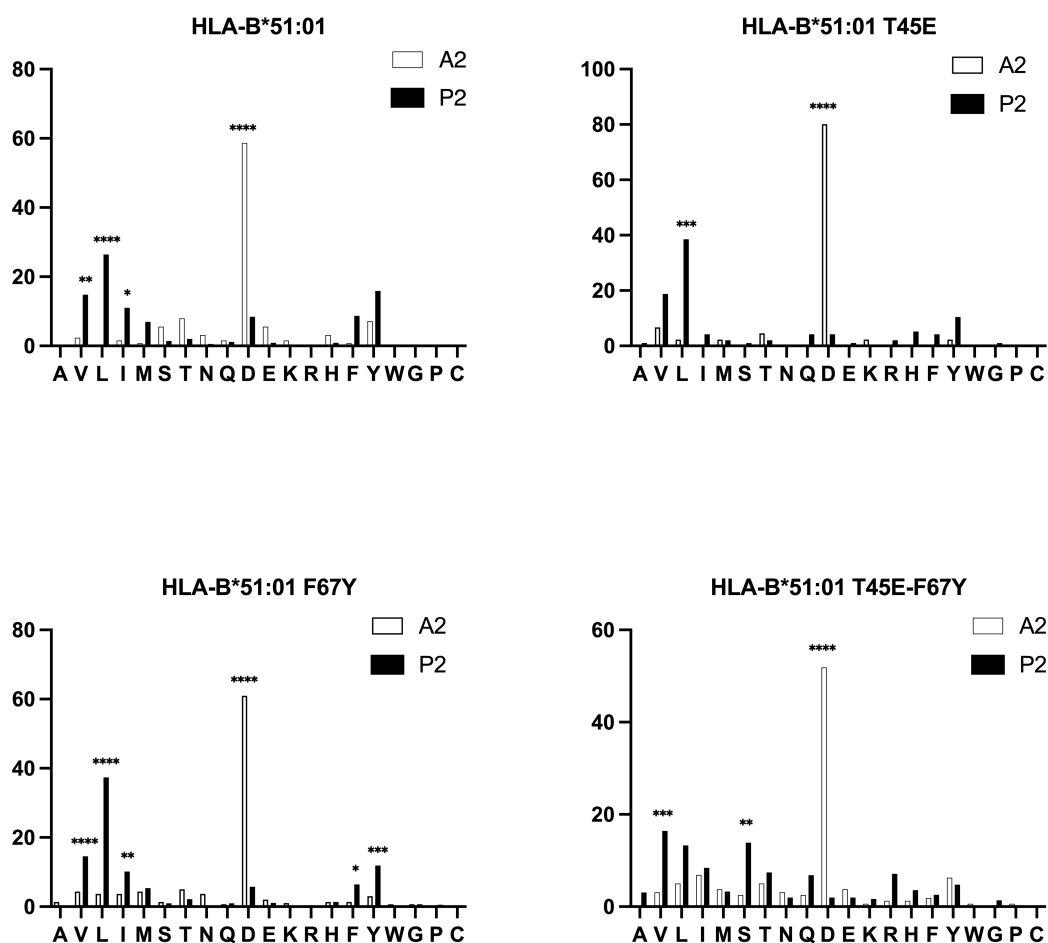




**Figura 47: Residuos localizados en P1 de los ligandos de HLA-B\*07:02 con Prolina en P2 (P2) y Alanina en P2 (A2).** Se utiliza el código de una sola letra para identificar los aminoácidos. En las gráficas de la izquierda, se muestra la frecuencia de cada residuo en posición 1 según si la posición 2 está ocupada por Pro (arriba) o Ala (abajo) según los resultados obtenidos (negro) o los datos disponibles en la literatura (blanco). En las gráficas de la derecha, los residuos se agrupan según sus características fisicoquímicas.

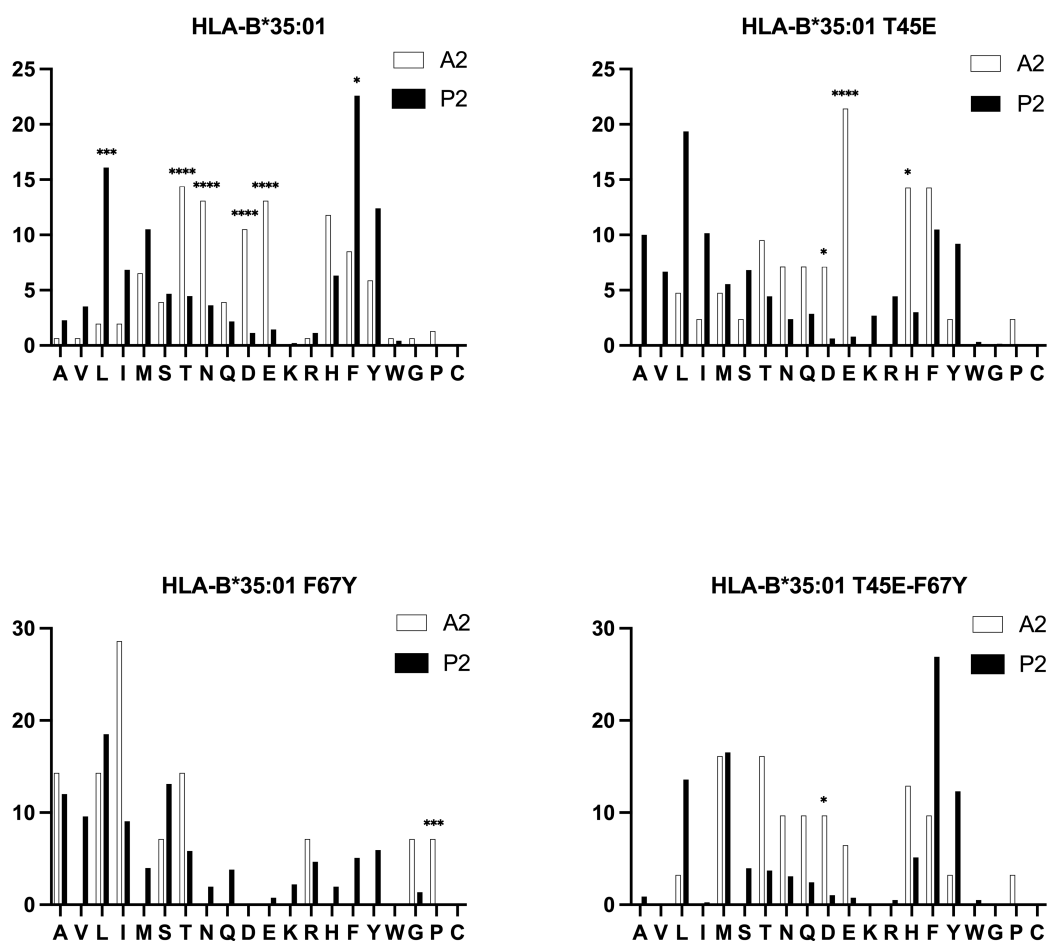
Una vez confirmada experimentalmente la presencia de subpeptidomas en HLA-B\*51:01 y HLA-B\*35:01, así como la ausencia en HLA-B\*07:02, se evaluó la contribución de los residuos localizados en las posiciones 45 y 67 al repertorio peptídico. Para ello, se analizó la posición *P2* de los diferentes mutantes generados para cada molécula y el efecto de sobre la posición *P1* con el fin de determinar su influencia en la presencia o ausencia de subpeptidomas.

En primer lugar, se analizaron los transfectantes referentes a la molécula **HLA-B\*51:01**. La mutación en posición 45 no provocó cambios significativos en los subpeptidomas descritos: cuando la posición 2 estaba ocupada por Ala, había una clara preferencia de Glu, mientras que cuando la posición 2 estaba ocupada por Pro, había preferencia de Leu. La mutación de la posición 67 provocó el mismo fenómeno, no se vieron alterados los subpeptidomas descritos, de hecho, se vieron favorecidos los residuos aromáticos, así como residuos alifáticos adicionales en los péptidos con Pro2. Finalmente, la molécula con ambas posiciones no mostró diferencias significativas respecto la molécula original (**Figura 48**).



**Figura 48:** Residuos localizados en P1 de los ligandos de HLA-B\*51:01, HLA-B\*51:01 T45E, HLA-B\*51:01 F67Y, HLA-B\*51:01 T45E + F67Y con Prolina en P2 (P2) y Alanina en P2 (A2). Se utiliza el código de una sola letra para identificar los aminoácidos. Se muestra la frecuencia de cada residuo en posición 1 según si la posición 2 está ocupada por Prolina (negro) o Alanina (blanco). Las diferencias significativas se determinan mediante un test Chi cuadrado y la corrección de Bonferroni como \* < 0.05, \*\* < 0.01, \*\*\* < 0.001 y \*\*\*\* < 0.0001. En las gráficas de la derecha, los residuos se agrupan según sus características fisicoquímicas.

La mutación en posición 45 (T45E) de la molécula **HLA-B\*35:01** tampoco afectó a la presencia de subpeptidomas. Al igual que la molécula sin mutar, se observó que existían dos subpeptidomas, uno de ellos con Pro2 y residuos alifáticos y aromáticos en *P1*, y otro formado por Ala2 y residuos polares y ácidos en *P1*. La mutación en posición 67 provocó ciertas diferencias en el peptidoma de esta molécula. Así, se observó un incremento de Leu, Ile y Ala en los péptidos con Ala2, pero no se detectó el incremento esperado de residuos ácidos en *P1*. Sin embargo, el escaso número de péptidos con Ala2 dificultan la interpretación de este resultado. Finalmente, la mutación de ambas posiciones a la vez no mostró cambios significativos en el peptidoma de esta molécula respecto a la molécula sin mutaciones (**Figura 49**).



**Figura 49:** Residuos localizados en *P1* de los ligandos de HLA-B\*35:01, HLA-B\*35:01 T45E, HLA-B\*35:01 F67Y, HLA-B\*35:01 T45E + F67Y con Prolina en *P2* (P2) y Alanina en *P2* (A2). Se utiliza el código de una sola letra para identificar los aminoácidos. Se muestra la frecuencia de cada residuo en posición 1 según si la posición 2 está ocupada por Prolina (negro) o Alanina (blanco). Las diferencias significativas se determinan mediante un test Chi cuadrado y la corrección de Bonferroni como \* < 0.05, \*\* < 0.01, \*\*\* < 0.001 y \*\*\*\* < 0.0001. En las gráficas de la derecha, los residuos se agrupan según sus características fisicoquímicas.

Por lo tanto, las mutaciones en posiciones 45 y 67 provocaron una disminución del porcentaje de péptidos con Ala2, es decir, hicieron que el *pocket* fuera más restrictivo y limitó la capacidad de que se uniera un residuo diferente de Pro. Sin embargo, a pesar de que los péptidos con Ala2 mermaron, estos seguían presentando subpeptidomas. Los resultados con el mutante F67Y son más confusos. Sin embargo, el limitado número de péptidos y el hecho de que los mutantes T45E y T45E+F67Y no muestren diferencias, sugieren que un mayor número de péptidos anularía estas diferencias. Por lo tanto, la mutación de las posiciones 45 y 67 no afectó a la presencia de subpeptidomas en las moléculas HLA-B\*51:01 y HLA-B\*35:01, sino a la restricción del *pocket*.

La molécula **HLA-B\*07:02** no presentaba subpeptidomas, pero tanto la mutación en posición 45, como la mutación en posición 67 y ambas juntas, provocaron ciertos cambios en el peptidoma. Estos cambios son complejos de analizar, pero en general, aunque solo se vieron cambios significativos en el doble mutante con Ser1 y Thr1, se observó una tendencia en los péptidos con Ala2, donde la posición 1 estaba favorecida con residuos polares y en el caso de la molécula HLA-B\*07:02 Y67F, también con residuos ácidos (**Figura 50**). Las características de estos peptidomas hacen que se asemejen a los subpeptidomas previamente identificados en B\*35:01, donde los péptidos con Pro2 estaban enriquecidos en residuos alifáticos en *P1*, y los péptidos con Ala2 estaban favorecidos con residuos polares y ácidos en *P1*. Por lo tanto, la mutación de las posiciones 45 y 67 en la molécula HLA-B\*07:02 provocó un aumento del porcentaje de péptidos con Ala2 y, además, estos péptidos presentaron características típicas de los subpeptidomas previamente identificados.

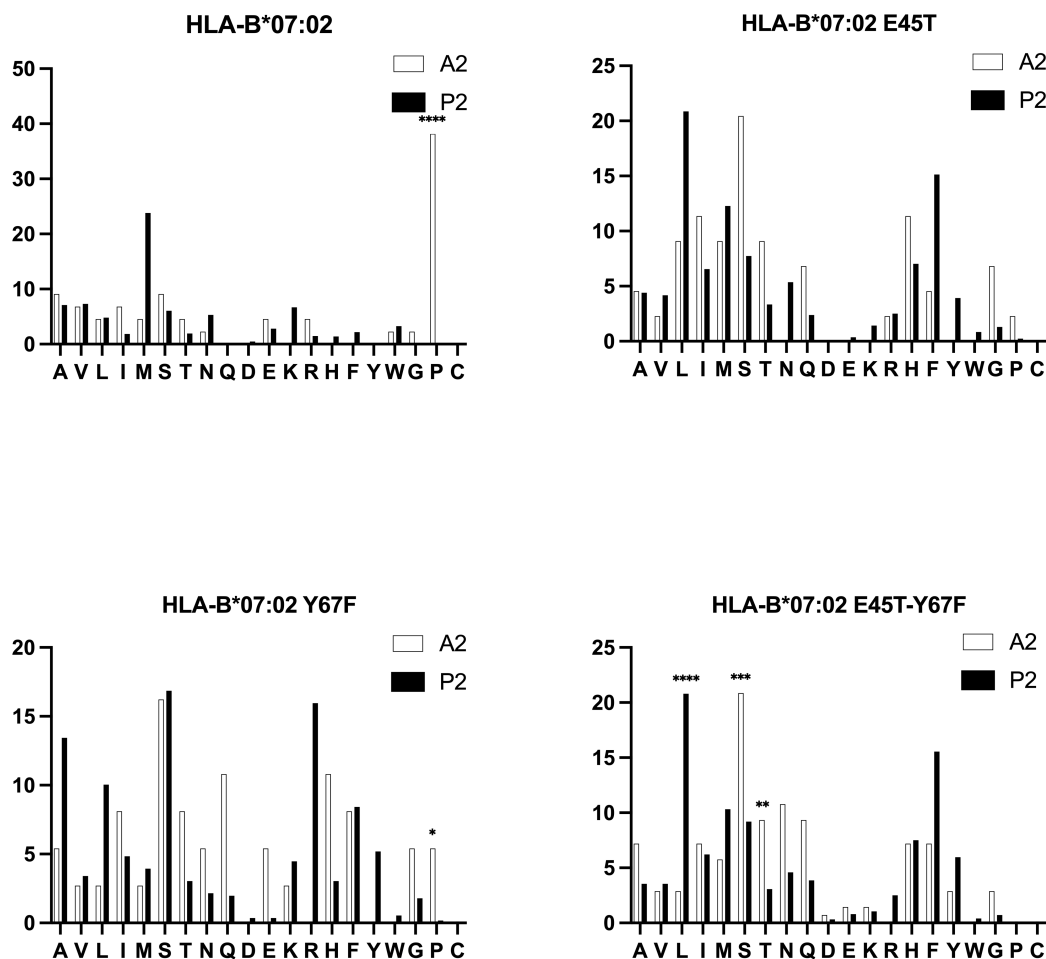


Figura 50: Residuos localizados en P1 de los ligandos de HLA-B\*07:02, HLA-B\*07:02 E45T, HLA-B\*07:02 Y67F, HLA-B\*07:02 E45T + Y67F con Prolina en P2 (P2) y Alanina en P2 (A2). Se utiliza el código de una sola letra para identificar los aminoácidos. Se muestra la frecuencia de cada residuo en posición 1 según si la posición 2 está ocupada por Prolina (negro) o Alanina (blanco). Las diferencias significativas se determinan mediante un test Chi cuadrado y la corrección de Bonferroni como \* < 0.05, \*\* < 0.01, \*\*\* < 0.001 y \*\*\*\* < 0.0001. En las gráficas de la derecha, los residuos se agrupan según sus características fisicoquímicas..

En conclusión, las sustituciones T45E y F67Y en HLA-B\*51:01 y HLA-B\*35:01 redujeron significativamente la frecuencia de péptidos con Ala2, aunque los péptidos restantes conservaron los subpeptidomas característicos de estos alotipos (Figura 48 y Figura 49). De forma inversa, la sustitución E45T y Y67F en HLA-B\*07:02 incrementó la proporción de péptidos con Ala2, los cuales en este caso sí modificaron la presencia de subpeptidomas, presentando un perfil similar a los descritos en HLA-B\*35:01 (Figura 50). En conjunto, estos resultados indican que, si bien las posiciones 45 y 67 solo parecen determinar la existencia de subpeptidomas en HLA-B\*07:02, sí desempeñan un papel crítico en la especificidad de anclaje de la P2 en los tres alotipos analizados.



## **DISCUSIÓN**





La respuesta inmune depende en gran medida de la degradación proteica. Parte de los péptidos generados en este proceso son posteriormente presentados a los linfocitos T CD8<sup>+</sup> y T CD4<sup>+</sup> a través de las moléculas de HLA de clase I y de clase II, respectivamente. El proteasoma 26S es el componente principal del UPS, un complejo catalítico que degrada proteínas intracelulares previamente poliubiquitinadas en residuos de lisina. Estructuralmente está compuesto por la partícula *core* o 20S, formada por el ensamblaje de 28 proteínas dispuestas en 4 anillos heptaméricos y una o dos partículas reguladoras 19S. La actividad catalítica reside en 3 subunidades del proteasoma 20S:  $\beta 1$ ,  $\beta 2$  y  $\beta 5$ , y cada una de ellas presenta una especificidad de sustrato distinta. Bajo condiciones de estrés, como en el caso de una infección, el proteasoma constitutivo ( $\beta 1$ ,  $\beta 2$ ,  $\beta 5$ ) es sustituido por el inmunoproteasoma ( $\beta 1i$ ,  $\beta 2i$ ,  $\beta 5i$ ). Además, existen proteasomas específicos de tejido, como el timoproteasoma ( $\beta 1i$ ,  $\beta 2i$ ,  $\beta 5t$ ), específico de las cTECs y el espermatoproteasoma ( $\alpha 4s$ ), específico de las líneas germinales masculinas<sup>201</sup>. Asimismo, existen proteasomas intermedios que combinan subunidades catalíticas tanto constitutivas como inmunes: SIP ( $\beta 1$ ,  $\beta 2$  y  $\beta 5i$ ) y DIP ( $\beta 1i$ ,  $\beta 2$  y  $\beta 5i$ ), lo que les otorga una actividad catalítica diferenciada del proteasoma constitutivo y del inmunoproteasoma. La existencia de los proteasomas intermedios fue sugerida por Dahlmann B. et al,<sup>118</sup> a principios de siglo y una década más tarde, fue confirmada por Guillaume et al,<sup>110</sup>. Durante mucho tiempo, se consideró que el proteasoma 20S era únicamente el núcleo catalítico del 26S, sin embargo, estudios posteriores demostraron que este complejo es capaz de llevar a cabo la proteólisis de proteínas independiente a la partícula reguladora 19S, de acuerdo principalmente a criterios estructurales: proteínas mal plegadas que son capaces de entrar en la cámara proteolítica. De hecho, en algunos casos la degradación de estas proteínas es más eficiente por parte del proteasoma 20S que por el 26S<sup>202</sup>. Sin embargo, el proteasoma 20S es incapaz de degradar otras proteínas completas *in vitro*.

El procesamiento antigénico que genera los péptidos presentados en la superficie celular a los linfocitos T CD8<sup>+</sup> requiere de varios pasos y cada uno de ellos influye en la generación del repertorio peptídico<sup>203</sup>. El proteasoma genera extensiones amino-terminales, pero también puede generar directamente ligandos peptídicos de las moléculas de HLA-I, tal y como se ha visto en experimentos *in vitro*<sup>107,204–206</sup>. El estudio

exhaustivo del papel que tienen los distintos tipos de proteasoma en la generación de ligandos peptídicos específicos, que pueden ser diana del ataque citotóxico, es un requisito para entender el proceso y plantear estrategias para modular, bien incrementando o disminuyendo, la respuesta inmune específica en procesos tumorales o autoinmunes, respectivamente.

En esta tesis, se ha realizado un análisis de la contribución al repertorio peptídico presentado por HLA-I del SIP, así como la caracterización exhaustiva del complejo 20S del mismo a nivel enzimático y estructural. El análisis del inmunopeptidoma de clase I de células con expresión exclusiva del SIP condujo a analizar la actividad catalítica de este complejo con 4 péptidos fluorogénicos que permitieran reflejar las características diferenciales observadas en el repertorio generado por el SIP frente al proteasoma constitutivo. Finalmente, por primera vez se ha determinado la estructura tridimensional de este complejo mediante Cryo-EM, con el objetivo de correlacionar los datos obtenidos en los repertorios y en los ensayos de actividad catalítica con la estructura tridimensional de este complejo.

La contribución del SIP al repertorio peptídico se analizó en modelos celulares. Se generaron dos modelos de estudio de este complejo en dos líneas celulares: HEK-293- $\beta 5i$  y HEK-293F- $\beta 5i$ -StII, a este último se le fusionó un “tag” en el extremo C-terminal de la subunidad  $\beta 5i$  para facilitar la purificación del complejo. El análisis comparativo entre el inmunopeptidoma de la línea parental HEK-293F y el transfectante HEK-293F- $\beta 5i$ -StII mostró que las características generales de los péptidos eluidos de ambas líneas eran concordantes con lo esperado para los repertorios de clase I, no obstante, se observaron diferencias en los residuos localizados en posición P $\Omega$ . El proteasoma constitutivo y el SIP se diferencian en la sustitución de la subunidad  $\beta 5$  por  $\beta 5i$  en este último. Esta subunidad confiere un incremento de la actividad tipo quimotripsina al complejo, respecto de su homóloga  $\beta 5$ <sup>207</sup>. Esta sustitución es coherente con su papel en contextos fisiológicos donde se requiere una mayor eficiencia en la generación de péptidos compatibles con la presentación por HLA-I, especialmente durante respuestas inmunes inducidas por citoquinas como el IFN- $\gamma$ .

La comparación de los inmunopeptidomas de HLA-I en células con expresión exclusiva de proteasoma constitutivo o SIP mostró que las primeras presentaban más péptidos con residuos hidrofóbicos pequeños (Ala, Val) mientras que las segundas favorecían la presentación de péptidos con residuos aromáticos (Phe, Tyr, Trp) en PΩ. Este hecho se observó tanto en la línea HEK-293 como en la línea HEK-293F (**Figura 16 y Figura 26**).

Se evaluó la actividad catalítica de ambos proteasomas a través de un ensayo de digestión con péptidos fluorogénicos. Este es un ensayo ampliamente utilizado para caracterizar dicha actividad<sup>208–210</sup>. Para ello se seleccionaron péptidos con los que evaluar las actividades caspasa, tripsina y en el caso de la actividad quimotríptica, se seleccionaron dos péptidos: uno acabado en un residuo pequeño (Ala) y otro en un residuo aromático (Trp). Los datos de las especificidades de los dos tipos de proteasoma se correspondieron con la presencia de los diversos péptidos observada en los repertorios peptídicos (**Figura 29**).

Finalmente, la resolución estructural del complejo permitió explicar tanto las particularidades del repertorio peptídico como las diferencias detectadas en su actividad catalítica. La disposición espacial de la Met45 fue diferente entre la subunidad β5 y β5i, abriendo o cerrando el *pocket* y, de esta forma, favoreciendo el procesamiento de residuos alifáticos pequeños o aromáticos, respectivamente (**Figura 30**).

Se ha descrito que la hidrólisis de proteínas por parte del proteasoma genera péptidos de 2-24 residuos de longitud, los cuales pueden ser degradados posteriormente por peptidasas intracelulares o transportados al lumen del RE para su unión a las moléculas de HLA-I<sup>45</sup>. Los complejos péptido-HLA son transportados a la membrana plasmática, donde desempeñan un papel fundamental en la inmunovigilancia. El papel del proteasoma en la presentación antigénica se describió hace décadas<sup>211</sup>. Mediante un modelo basado en un péptido derivado de ovalbúmina (OVA), cuya generación era dependiente de ubiquitinación, se analizaron los péptidos generados y se demostró que el proteasoma 20S puede generar epítomos antigénicos sin necesidad de otros componentes del sistema<sup>212</sup>. La inhibición del proteasoma condujo al bloqueo de la producción de péptidos y de su presentación por MHC-I, evidenciando la relevancia del proteasoma en la producción de ligandos de clase I. Además, el uso de inhibidores

demostró que el proteasoma es el responsable de generar el extremo C-terminal de los péptidos, pero que no siempre produce directamente el extremo N-terminal, que puede ser procesado por aminopeptidasas del RE como ERAP1 y ERAP2<sup>213</sup>.

Los distintos tipos de proteasoma contienen diferentes subunidades catalíticas, lo que conlleva diferentes especificidades de sustrato y sugiere que moldean el repertorio antigénico presentado por HLA-I. Sin embargo, la composición de proteasomas de una célula tiene múltiples implicaciones, no solo afectando a la generación del repertorio presentado por HLA-I, sino también a otros procesos fisiológicos, como la transducción de señales por la vía NFκB y la degradación proteica para la homeostasis celular<sup>214</sup>. Recientemente se ha descrito que los proteasomas intermedios se expresan en las células β-pancreáticas y que las subunidades β1i y β5i están reguladas positivamente en condiciones de baja concentración de IL-1β<sup>215</sup>. Se ha postulado que el SIP, al igual que el DIP, aumentan la variabilidad de péptidos de unión a HLA-I, aumentando el repertorio antigénico presentado por MHC de clase I para incluir péptidos que puedan ser procesados por células dendríticas (que contienen tanto inmunoproteasoma como proteasomas intermedios) y células tumorales (que contienen tanto proteasoma constitutivo como proteasomas intermedios). Debido a su actividad catalítica diferenciada del resto de complejos, los proteasomas intermedios pueden generar péptidos únicos, incluyendo péptidos derivados de antígenos tumorales, como los antígenos restringidos a HLA-A\*02 MAGE-A10<sub>254–262</sub>, y MAGE-C2<sub>191–200</sub>, que son producidos exclusivamente por el DIP<sup>132</sup>, y MAGE-A3<sub>271–279</sub>, por SIP<sup>119</sup>. Por tanto, los péptidos producidos por los proteasomas intermedios podrían constituir *targets* valiosos para inmunoterapia.

Estudios previos de la actividad catalítica del SIP y del DIP han mostrado que ambos presentan actividad tipo tripsina y quimotripsina intermedia entre el proteasoma constitutivo y el inmunoproteasoma<sup>215</sup>. El SIP conserva actividad caspasa similar al proteasoma constitutivo, mientras que el DIP la muestra reducida. Esto se debe a que la actividad caspasa está asociada a la subunidad β1, presente en el proteasoma constitutivo y SIP, pero no el DIP<sup>119</sup>.

En nuestros datos, la actividad tríptica está incrementada en el SIP respecto al proteasoma constitutivo. Esto se ha visto reflejado en el inmunopeptidoma de células con expresión única de ambos proteasomas (**Figura 26 y Figura 27**). En ambos proteasomas se encuentra presente la subunidad  $\beta 2$ . Sin embargo, este incremento en la actividad tríptica ha sido ya descrito, tanto con péptidos fluorogénicos, como con digestiones con péptidos sintéticos largos (Sholz, E, Tesis Doctoral: *Mecanismos de Procesamiento y Presentación Antigénica en Autoinmunidad*. Universidad Autónoma de Barcelona, 2017). El hecho de que la subunidad  $\beta 2$  se encuentre presente en ambos proteasomas y se modifique su actividad sugiere que las especificidades de los distintos proteasomas es un fenómeno más complejo que la simple actividad independiente de las subunidades por separado, produciéndose fenómenos alostéricos en un complejo enzimático con varios centros activos trabajando simultáneamente.

La selectividad de las subunidades catalíticas se define por la interacción de la  $P1$  de la cadena lateral del sustrato y el *pocket* S1 de unión al sustrato de la subunidad catalítica, concretamente, al residuo 45 de esta última. En la subunidad  $\beta 1$ , el *pocket* S1 está formado por Thr20, Thr31, Thr35, **Arg45**, Ala49, Gln53 y Asp120. La Arg45, cargada positivamente, permite que se acomoden residuos ácidos. El análisis estructural de la subunidad  $\beta 1$  y  $\beta 1i$  en un modelo murino muestra que existen diferencias entre estas subunidades que explican la diferencia de catálisis de sustratos, como las sustituciones: R45L y T31F en  $\beta 1i$  respecto a  $\beta 1$ . Estas sustituciones favorecen la unión de residuos hidrofóbicos pequeños en lugar de residuos ácidos, lo que a su vez favorece la actividad quimotripsina ( $\beta 1i$ ) frente la actividad caspasa ( $\beta 1$ ), limitando la producción de péptidos con residuos C-terminales ácidos, que no se suelen unir a HLA y aumentando los residuos hidrofóbicos<sup>216</sup>.

En la subunidad  $\beta 2$ , el *pocket* S1 está formado por los residuos Cys31, His35, **Gly45**, Asp53 y Cys129. El residuo Gly45 está separado 12 Å, resultando en un *pocket* considerablemente más grande que los *pockets* responsables de las actividades quimotripsina y caspasa. El *pocket* S1 de la subunidad  $\beta 2i$  es idéntico al de la subunidad  $\beta 2$ , excepto por la sustitución de Asp53 por Glu en la subunidad  $\beta 2i$ <sup>216</sup>.

En cuanto a la subunidad  $\beta 5$ , los residuos que dan lugar a la actividad tipo quimotripsina son: Ala20, **Met45**, Ala49 y Cys52. La Met45 sobresale del *pocket* 6,9 Å de la Thr1, permitiendo acomodar residuos hidrofóbicos como alanina, valina o tirosina<sup>216</sup>. El estudio de Huber et al, (2012) demostró, mediante el análisis estructural comparativo entre el proteasoma constitutivo y el inmunoproteasoma de un modelo murino, que, aunque los residuos determinantes de la especificidad al *pocket* S1 no varían entre las subunidades  $\beta 5$  y  $\beta 5i$ , este varía en tamaño, ya que la Met45 adopta diferentes conformaciones que resultan en un *pocket* S1 más espacioso en la subunidad  $\beta 5i$ , y significativamente más pequeño en la subunidad  $\beta 5$ . La formación de un *pocket* S1 más grande está favorecido por interacciones de van der Waals de la Met45 con Gln53, mientras que en la subunidad  $\beta 5$ , la Ser53 no puede estabilizar la Met45. Como resultado, la subunidad  $\beta 5$  tiene un *pocket* S1 mucho más pequeño. Por lo tanto, el *pocket* S1 de la subunidad  $\beta 5i$  acomoda preferentemente residuos hidrofóbicos como Phe, Tyr y Trp, mientras que el *pocket* S1 de la subunidad  $\beta 5$  acomoda preferentemente residuos hidrofóbicos pequeños como Ala y Val. Estas observaciones han sido confirmadas con experimentos de hidrólisis de sustratos fluorogénicos<sup>217</sup>. En 2016, Huber et al, confirmaron los resultados que muestran que la Met45 adopta una conformación abierta independientemente de la especie, indicando que la reorientación de la Met45 es inherente a la subunidad  $\beta 5i$  y que este cambio conformacional está codificando entre los residuos 1-138<sup>218</sup>. Nuestros resultados reflejan la capacidad del *pocket* S1 de la subunidad  $\beta 5i$  del proteasoma intermedio de tipo I humano para acomodar residuos voluminosos y confirman que la sustitución de la subunidad  $\beta 5$  por  $\beta 5i$  influye de manera específica en la selección y procesamiento de péptidos con características estructurales determinadas, en concordancia con su papel en la generación de ligandos para la presentación por HLA-I y con los resultados obtenidos tras analizar el inmunopeptidoma de nuestro modelo celular de SIP. Sin embargo, se debe tener en cuenta que los residuos que se encuentran flanqueando también influyen en la catalización de los sustratos.

Hasta la fecha, no se había resuelto la estructura tridimensional del proteasoma intermedio de tipo I. Huber et al, realizaron la cristalografía de un proteasoma constitutivo de levadura con la sustitución de la subunidad  $\beta 5$  por la subunidad  $\beta 5i$

humana<sup>218</sup>. Por este motivo, y con el fin de confirmar los datos observados en el inmunopeptidoma y el ensayo de actividad catalítica, se decidió determinar la estructura tridimensional de este complejo. En este trabajo, por primera vez se ha resuelto la estructura tridimensional del proteasoma 20S intermedio de tipo I humano mediante Cryo-EM en su forma Apo y junto con el inhibidor Bortezomib a una resolución de 2,94Å y 3,11Å, respectivamente. De acuerdo con la bibliografía, la disposición espacial de la Met45 varía entre el proteasoma constitutivo y el SIP, siendo más abierta en este último. Sin embargo, la unión del inhibidor selectivo de esta subunidad Bortezomib provocó cambios en el *pocket*, haciéndolo más estrecho y similar al *pocket* de la subunidad β5 (**Figura 30**). Nuestros datos de la estructura del SIP muestran a nivel estructural la presencia de un número inferior de péptidos con residuos alifáticos pequeños y mayor con residuos aromáticos en el extremo C-terminal. Además, las estructuras obtenidas en esta tesis confirman los datos descritos por Huber et al, en 2016 en el que se muestra que esta disposición diferencial de la subunidad β5i está conservada y es independiente de la especie.

La presencia de proteasomas intermedios en células tumorales incrementa la diversidad de antígenos presentados por estas células y tiene implicaciones en terapias anti-tumorales. Guillaume et al, demostró que las células tumorales que expresan proteasoma constitutivo junto con un 2-8% de proteasoma intermedio de tipo II eran atacadas por CTLs contra antígenos producidos exclusivamente por el proteasoma intermedio de tipo II, demostrando que una baja proporción del proteasoma adecuado produce los antígenos necesarios para desarrollar una respuesta eficaz<sup>110</sup>. En un contexto clínico, la identificación de péptidos generados exclusivamente por proteasomas intermedios abre nuevas vías para el desarrollo de inmunoterapias, pero sería adecuado determinar el *pool* de proteasomas presentes en la célula, ya que de éste depende, al menos en parte, que tenga lugar el reconcomiendo por parte de los CTLs. Un número variable de antígenos relevantes para el rechazo de un tumor, son exclusivamente producidos por el inmunoproteasoma, sobre todo aquellos que dependen de la presentación por parte de las células dendríticas (que contienen principalmente inmunoproteasoma). Sin embargo, las células tumorales que no se localizan en focos inflamatorios tienden a no expresar el inmunoproteasoma, lo que

implicaría el no reconocimiento por parte de células T<sup>219</sup>. Otro trabajo demuestra que las células tumorales presentan, en la mayoría de los casos, proteasomas intermedios que producen los mismos antígenos que el inmunoproteasoma (junto con antígenos adicionales), lo que asegura la respuesta contra estos tumores, incluso cuando no son expuestos a INF- $\gamma$ . Los antígenos tumorales MAGE-C2<sub>336-44</sub> y MAGE-C2<sub>42-50</sub> son producidos eficientemente por el inmunoproteasoma, sin embargo, en la línea celular de melanoma EB81-MEL, que no contiene inmunoproteasoma, en ausencia de INF- $\gamma$  estos antígenos son reconocidos por CTLs, ya que son producidos por proteasomas intermedios<sup>133</sup>. Estos hallazgos sugieren que las inmunoterapias dirigidas contra antígenos producidos por el inmunoproteasoma deberían ser efectivas contra tumores no expuestos a condiciones inflamatorias, siempre que expresen proteasomas intermedios, evidenciando que el *pool* de proteasomas presente en la célula modula la diversidad y especificidad del inmunopeptidoma de clase I. Esta plasticidad posibilita que las células adapten la respuesta inmunitaria ante diferentes situaciones, como infecciones virales o desarrollo de tumores, favoreciendo la generación de epítomos que puedan ser reconocidos por los linfocitos T CD8<sup>+</sup> eficazmente. A pesar de estos datos, haría falta determinar las características de los inmunopeptidomas del DIP y del inmunoproteasoma para cuantificar las posibilidades de presentar péptidos comunes y asumir que la presencia de proteasomas intermedios pueden suplir la ausencia de inmunoproteasoma en cáncer.

La línea celular HEK (*Human Embryonic Kidney cells*)-293 es una línea de origen humano obtenida en 1973 a partir de la inmortalización por la integración de 4 kbp del adenovirus 5 en el cromosoma 19 de células embionarias de riñón<sup>153,154</sup>. A pesar de su origen epitelial, su maquinaria bioquímica es capaz de llevar a cabo la mayor parte del plegamiento y procesamiento postraducciona necesario para generar proteínas maduras y funcionales a partir de un amplio espectro de ácidos nucleicos<sup>220</sup>. A partir de esta línea, se han desarrollado diversas líneas adicionales. Las líneas celulares HEK-293 y HEK-293T son muy utilizadas en la expresión y purificación de proteínas debido a su fácil manipulación y transfección. La línea HEK-293F presenta la ventaja que su crecimiento en suspensión permite obtener grandes volúmenes de células en poco tiempo<sup>155</sup>.



El análisis comparativo del inmunopeptidoma de las líneas celulares HEK-293F y HEK-293F-β5i-StII reveló la ausencia de péptidos con extremo C-terminal básico, sin embargo, en el análisis del inmunopeptidoma de HEK-293 y HEK-293 β5i sí se detectaron estos péptidos. Además, la línea HEK-293 había sido previamente utilizada por nuestro grupo en experimentos de peptidómica, en los cuales se detectaron péptidos con extremo C-terminal básico, motivo por el cual se consideró que la ausencia de péptidos con residuos básicos en PΩ podía ser una característica intrínseca de la línea HEK-293F. Estos péptidos son presentados en la superficie celular por la molécula HLA-A\*03:01. Para confirmar esta hipótesis, se evaluó la presencia de esta molécula tanto en la línea HEK-293 como HEK-293F mediante una tinción extracelular con anticuerpos específicos de alelo. Los resultados obtenidos indicaron la ausencia de la molécula HLA-A\*03:01 en la membrana de la línea HEK-293F (**Figura 18**), lo cual explica por qué en esta línea no se detectaron péptidos con extremo C-terminal básico, en contraste con la línea HEK-293, donde sí se detectaron. Sin embargo, el tipaje solicitado por nuestro grupo mostró que ambas líneas contenían a nivel genómico los alelos de clase I: *HLA-A\*02:01*, *A\*03:01*, *-B\*07:02* y *-C\*07:02*; por lo que se procedió a realizar un análisis por qPCR con *primers* específicos para estos alelos (excepto *HLA-C\*07:02*). Los resultados confirmaron la expresión a nivel de mRNA de los tres alelos en ambas líneas celulares (**Figura 19**), lo que sugería la existencia de un fenómeno post-transcripcional que impedía la correcta expresión de la molécula HLA-A\*03:01 en la membrana celular de la línea HEK-293F.

A pesar de que todas estas líneas celulares derivan de la misma línea parental, existen cambios genómicos y transcriptómicos debido a la inestabilidad genómica de HEK-293, y estos cambios son más evidentes en las líneas de crecimiento en suspensión<sup>155,221</sup>. La organización cromosómica de la línea HEK-293 se encuentra en constante evolución debido a eventos de traslocación y alteraciones en el número de copias, lo que sugiere que el cultivo prolongado y la sub-clonación de células conducen a una deriva cariotípica. En nuestro trabajo, se ha trabajado con el pase 5 de las células HEK-293F obtenidas del ATCC, por lo que la posibilidad de que la pérdida de HLA-A\*03:01 sea debida al cultivo celular es muy baja. Además, se ha reportado una regulación negativa de genes implicados en la matriz extracelular, locomoción y adhesión de las líneas derivadas respecto la línea parental HEK-293, mientras que los procesos de metabolismo

de aminoácidos y pequeñas moléculas se encuentran regulados positivamente. También se han observado diferentes patrones de expresión de genes involucrados en la organización celular, como la morfogénesis, el citoesqueleto y la membrana<sup>155</sup>. En definitiva, la línea parental HEK-293 difiere significativamente de sus líneas derivadas en la expresión de proteínas de membrana y procesos de adhesión celular, lo que hace que la línea HEK-293F no constituya un modelo idóneo de análisis del inmunopeptidoma. Estos resultados se relacionan con los resultados obtenidos en el análisis de los residuos de anclaje de los modelos celulares generados para caracterizar el SIP. Tal y como ya se ha mencionado, los residuos de anclaje clave en las moléculas de HLA-I presentes en estas líneas son P2 y PΩ; el análisis de los residuos de anclaje en la línea HEK-293 mostró que en PΩ se unían tanto residuos alifáticos como básicos, sin embargo, en HEK-293F, no se detectaron los residuos básicos (**Figura 14 y Figura 24**). Asimismo, otro trabajo realizado en nuestro grupo con la línea HEK-293 ha mostrado resultados similares. La transfección de esta línea con la proteína Spike de Sars-CoV-2 provocó la pérdida de expresión en membrana de la molécula HLA-A\*03:01, demostrando, de nuevo, que esta línea presenta inestabilidad en el genoma y que, además, esta inestabilidad se traduce en la pérdida específica de la molécula HLA-A\*03:01 en superficie. Estos resultados ponen de manifiesto la necesidad de evaluar la expresión de los diferentes alelos de HLA previo al análisis del inmunopeptidoma, con el fin de no menospreciar el repertorio antigénico y obviar péptidos no presentados por el hecho de no expresar el alelo del cual son ligandos.

Este trabajo se ha realizado en líneas celulares *in vitro*, por lo que la validación de los resultados obtenidos en modelos *in vivo* resulta fundamental para confirmar la relevancia fisiológica de los datos. Los experimentos llevados a cabo en modelos celulares permiten realizar la caracterización del SIP, pero no reproducen la complejidad ofrecida por un modelo fisiológico. La validación *in vivo* permitiría evaluar de manera más realista la relevancia inmunológica de los péptidos generados por el SIP. Además, en un contexto inflamatorio, la interacción con el microambiente, incluyendo células y citocinas, podría modular la expresión y actividad de los distintos subtipos de proteasoma. Futuras investigaciones podrían enfocarse en la caracterización de los proteasomas intermedios a partir de la purificación de estos de tejidos sanos y

tumorales, con el fin de caracterizarlos y evaluar su función en un microambiente tumoral. Además, resulta imprescindible explorar la respuesta inmune mediada por CTL frente a péptidos exclusivamente generados por los proteasomas intermedios, especialmente de tejidos tumorales inducidos y no inducidos con IFN- $\gamma$ , de acuerdo con el trabajo de Van den Eyde. Sin embargo, los datos obtenidos aportan información acerca del SIP y su contribución al repertorio antigénico, lo que en un futuro podría contribuir a la identificación de antígenos y desarrollo de inmunoterapias.

De cara al futuro, este análisis debe expandirse al DIP. Durante la presente tesis doctoral, se ha tratado de obtener un transfectante sobre la línea celular HEK-293 que expresara el DIP ( $\beta 1i$ ,  $\beta 2$ ,  $\beta 5i$ ) pero a pesar de llevar a cabo múltiples aproximaciones con el fin de conseguirlo, no se logró el desplazamiento total de la subunidad  $\beta 1$  por la subunidad inmune  $\beta 1i$ .

Los repertorios peptídicos, tanto de clase I como de clase II, están determinados en gran medida por la estructura de las moléculas de HLA, que permiten, o no, acomodar un péptido dado en el sitio de unión. Estos genes son extremadamente polimórficos, y gran parte de este polimorfismo se concentra en la región de unión al péptido, lo que condiciona la unión de éstos. A pesar del elevado número de alelos de HLA-I, las moléculas de HLA pueden clasificarse en supertipos, grupos de moléculas que unen inmunopeptidomas con motivos de anclaje similares. Cada supertipo une un supermotivo, es decir, refleja los residuos de anclaje principales reconocidos por las moléculas de HLA pertenecientes a un mismo supertipo. Este término fue acuñado por primera vez por el grupo de Alessandro Sette, definiendo 9 supertipos de HLA<sup>141</sup>. Esta clasificación ha demostrado ser útil en la caracterización e identificación de epítomos promiscuos de células T de diversas enfermedades, como sarampión-paperas-rubéola, SARS, EBV, VIH, virus del herpes humano asociado al sarcoma de Kaposi, VHC, VHB, VPH, influenza, *P. falciparum*, LCMV, virus de Lassa, *F. tularensis* y también antígenos de cáncer<sup>143,222</sup>. Asimismo, la clasificación en supertipos ha sido utilizada en los algoritmos de predicción de péptidos candidatos con capacidad de unión a HLA-I alterada y en estudios de asociación de enfermedades<sup>223,224</sup>. Es importante destacar que la

clasificación de supertipos se basa en los residuos de anclaje de las diferentes moléculas de HLA, pero no asegura el reconocimiento por parte de células T, ya que en la literatura existen cientos de péptidos de alta afinidad por una molécula de HLA que no son reconocidos por células T.

El concepto de supertipo ha sido fundamental para explicar cómo alelos distintos pueden compartir patrones de unión al péptido y, en consecuencia, presentar repertorios parcialmente solapantes. Sin embargo, trabajos posteriores demostraron que existía un nivel adicional de complejidad ya que, dentro de un mismo supertipo, se observaron repertorios diferenciales. A estos conjuntos de péptidos se les denominó subpeptidomas. En esta tesis, el concepto de subpeptidoma se refiere a las diferencias que se producen en una posición (*P1*) cuando se presentan diferentes residuos en otra posición (*P2*). La existencia de subpeptidomas permite explicar cómo polimorfismos puntuales en el surco de unión al péptido condicionan la unión de un mismo epítipo, lo que resulta fundamental para comprender la susceptibilidad a enfermedades. Dentro del supertipo B7, las moléculas HLA-B\*51:01 y HLA-B\*51:08 unen péptidos con Pro2 o Ala2 y residuos hidrofóbicos en *PQ*. Trabajos anteriores muestran que los repertorios de estas moléculas presentan dos subpeptidomas, uno con Pro2 y residuos hidrofóbicos en *P1*, y otro con Ala2 y Asp1<sup>145–147</sup>.

En esta tesis, se ha confirmado que la presencia de subpeptidomas en moléculas de HLA pertenecientes al supertipo B7 no es una característica exclusiva del grupo alélico B\*51, sino que se extiende a más moléculas de este supertipo. Para ello, se ha realizado el análisis sistemático de 13 moléculas de HLA pertenecientes al supertipo B7 disponibles en la base de datos IEDB. Se ha analizado el residuo en posición 2, así como el impacto de este sobre la posición 1, demostrando *in silico* la existencia de 2 subpeptidomas, uno de ellos con Pro2 y residuos hidrofóbicos en *P1* y otro con Ala2 y residuos polares y ácidos en *P1*. Además, el análisis de cristales descritos en la literatura de tres moléculas representativas de este supertipo (HLA-B\*51:01, HLA-B\*35:01 y HLA-B\*07:02) puso de manifiesto que las posiciones 45 y 67 podrían estar involucradas en la especificidad del *pocket*<sup>151,152</sup>. Por ello, la segunda parte del Capítulo II de resultados trata de explicar experimentalmente los hallazgos publicados<sup>225</sup>. Se generaron modelos celulares de dichas moléculas en la línea celular C1R mutando las posiciones 45 y 67 y ambas juntas

con el fin de analizar el impacto de estas posiciones en el repertorio antigénico, así como estudiar experimentalmente si estas posiciones definen la presencia o ausencia de subpeptidomas. Los datos experimentales confirmaron los resultados *in silico* sobre que estas posiciones modulaban la promiscuidad del *pocket* en relación con aceptar otros residuos, además de la Pro. Además, la mutación de estas posiciones clave mostró que los residuos ocupados en estas posiciones no afectan a la presencia de subpeptidomas de la misma forma en todas las moléculas, no influyendo en HLA-B\*51:01 ni HLA-B\*35:01, pero sí en HLA-B\*07:02.

Los resultados obtenidos en este trabajo permiten comprender cómo pequeñas variaciones en la secuencia de HLA pueden modular el repertorio de péptidos presentados. Esta modulación tiene implicaciones directas en la respuesta inmune frente a patógenos, antígenos tumorales y autoantígenos. Los datos generados podrían contribuir a explicar diferencias en susceptibilidad a enfermedades, progresión de infecciones virales y eficacia de estrategias de inmunoterapia basadas en epítomos específicos. Asimismo, podrían definir mutaciones de los patógenos para eludir la presentación antigénica dependiendo del alotipo expresado en el hospedador.

Existen diferentes enfermedades asociadas al supertipo B7 y la base de estas asociaciones podrían ser diferencias sutiles en los repertorios antigénicos. La enfermedad de Behçet está estrechamente relacionada con el alelo HLA-B\*51:01<sup>146</sup>. La molécula HLA-B\*35:01 está asociada a una rápida progresión de la infección por VIH<sup>226</sup>, así como en hepatitis C<sup>227</sup>. HLA-B\*07:02 está fuertemente asociado a la susceptibilidad a esclerosis múltiple<sup>228</sup>.

Pequeñas variaciones en la secuencia del péptido influyen en que un epítomo llegue a ser presentado por moléculas HLA-I. Los subpeptidomas no solo describen variaciones del repertorio, sino que pueden predecir la probabilidad de presentación de epítomos clínicamente relevantes, por lo que su estudio constituye una ventaja en la generación de vacunas basadas en epítomos. Aunque el concepto de supertipo ha permitido identificar repertorios parcialmente solapantes, el concepto de subpeptidoma ha añadido un nivel de complejidad, lo que implica que epítomos derivados de patógenos que cumplen con residuos de anclaje descritos para un grupo alélico dado, pueden no

ser presentados de forma eficiente por todos los alelos que lo integran, debido a la preferencia diferencial de cada subpeptidoma. En infecciones como VIH o hepatitis C, donde la presión selectiva favorece la aparición de mutaciones para evadir el sistema inmune, priorizar epítopos que encajen de forma robusta en los subpeptidomas de los alelos más frecuentes podría incrementar la eficacia de la respuesta inmune. De forma similar, en patógenos con alta variabilidad antigénica, como el virus de la influenza o SARS-CoV-2, tener en cuenta los subpeptidomas permitiría diseñar vacunas multiepitópicas que aseguren la presentación de al menos un conjunto de péptidos clave en la mayoría de los individuos. Tener en cuenta los subpeptidomas permite seleccionar de manera más precisa los epítopos con mayor probabilidad de ser presentados, optimizando la respuesta inmune deseada.

En conjunto, los resultados de esta parte de la tesis ponen de manifiesto que la existencia de subpeptidomas en moléculas de HLA-I del supertipo B7 constituye un determinante clave en la conformación del repertorio antigénico y abre la puerta a futuros estudios. Comprender cómo las posiciones críticas del surco de unión al péptido modulan la unión de éstos proporciona una base sólida para explicar asociaciones alélicas con enfermedades, así como para optimizar el diseño de vacunas y de inmunoterapias. De este modo, este trabajo no solo aporta una caracterización detallada de los subpeptidomas en el contexto experimental, sino que también los posiciona como un concepto para tener en cuenta en la especificidad de la inmunidad adaptativa.

Una de las principales limitaciones de este estudio es que gran parte de este se ha realizado *in silico* y mediante espectrometría de masas, que, a pesar de ser una herramienta muy potente, puede introducir sesgos y dejar algunos ligandos fuera del análisis, aunque con la mejora de los equipos y los motores de búsqueda, esto es cada vez menos probable. Algunos transfectantes analizados presentan un número de péptidos limitado, lo cual dificulta el análisis. El análisis del porcentaje de Ala2 en algunos transfectantes resultó estadísticamente no significativo a pesar de que gráficamente las diferencias eran claras. Este fenómeno se debe a que la mutación de las posiciones 45 y 67 condujo a una notable reducción en la proporción de péptidos con Ala2 en HLA-B\*51:01 y HLA-B\*35:01, lo que impedía obtener un resultado

estadísticamente significativo, sin embargo, gráficamente, la presencia de péptidos ácidos en el repertorio con Ala2 es evidente, indicando que, a pesar de estas mutaciones, se conserva la existencia de subpeptidomas en las moléculas HLA-B\*51:01 y HLA-B\*35:01. Este trabajo se basa en un enfoque descriptivo, por lo que resulta fundamental validar funcionalmente las observaciones realizadas y definir si realmente estos epítomos son reconocidos diferencialmente por los linfocitos T. La mutación de las posiciones 45 y 67 puso de manifiesto su importancia en la especificidad del *pocket* de unión al péptido, sin embargo, no se debe descartar la implicación de posiciones adicionales. Futuros trabajos podrían enfocarse en un punto de vista más clínico, definiendo los subpeptidomas presentados por estas moléculas en muestras de pacientes, y evaluar funcionalmente la respuesta específica de los linfocitos T CD8<sup>+</sup>. Asimismo, este estudio podría ampliarse a otros supertipos de HLA donde aún no se ha descrito la presencia de subpeptidomas.

En definitiva, tanto la caracterización del proteasoma intermedio de tipo I, como el análisis de los repertorios de moléculas del supertipo B7 amplían la comprensión de la variabilidad en la presentación de péptidos por HLA-I y sus consecuencias biológicas.





## **CONCLUSIONES**



1. El inmunopeptidoma de HLA-I de células que expresan exclusivamente el SIP contiene un menor número de ligandos con residuos alifáticos pequeños y mayor con residuos básicos y aromáticos en la posición P $\Omega$  que el inmunopeptidoma de células con expresión exclusiva del proteasoma constitutivo.
2. Los ensayos *in vitro* con péptidos fluorogénicos y proteasoma 20S purificado demuestran que el proteasoma constitutivo presenta actividad quimotríptica diferente del SIP, siendo el proteasoma constitutivo más eficiente catalizando el enlace tras Ala y el proteasoma intermedio de tipo tras Trp. Además, el SIP tiene una actividad tríptica más alta que el proteasoma constitutivo. Estas actividades catalíticas son compatibles con los repertorios peptídicos de HLA-I purificados en células con expresión exclusiva de dichos complejos.
3. La subunidad  $\beta 5i$  del SIP presenta un *pocket* de unión al sustrato más amplio que la subunidad  $\beta 5$ , al tener la Met45 diferente disposición en el centro activo. Esto explica las diferencias en la actividad quimotríptica y los repertorios peptídicos.
4. La línea celular HEK-293F no presenta en superficie el alelo de clase I *HLA-A\*03:01* a pesar de que este se expresa a nivel de mRNA. Por este motivo, esta línea celular no es equivalente a la línea parental HEK-293 para ser utilizada como modelo de procesos de presentación antigénica de péptidos presentados por dicho alelo.

5. La presencia de subpeptidomas en el supertipo B7 de HLA-I no es exclusiva de las moléculas del grupo alélico B\*51. Los alotipos HLA-B\*51:01, -B\*51:08 y -B\*54:01 presentan inmunopeptidomas con péptidos con Ala2 favorecidos y dos subpeptidomas: uno de ellos con Ala2 y residuos ácidos en *P1* (así como residuos polares), y otro con Pro2 y residuos hidrofóbicos en *P1*. Los alotipos HLA-B\*35:01, -B\*35:03, -B\*35:07 y -B\*53:01 presentan inmunopeptidomas con péptidos con Ala2 favorecidos y dos subpeptidomas: uno con Ala2 y residuos polares en *P1* (así como residuos ácidos), y otro con Pro2 y residuos hidrofóbicos en *P1*. Los alotipos HLA-B\*07:02, -B\*07:04, -B\*42:01, -B\*55:01 -B\*55:02 y -B\*56:01 presentan inmunopeptidomas con péptidos con Ala2 desfavorecidos y sin diferencias en la posición *P1* entre los péptidos con Pro2 o Ala2.
6. Las posiciones 45 y 67 del *pocket* B afectan en diferente grado a la generación de subpeptidomas en moléculas de HLA-I, pero juegan un papel fundamental en la especificidad del sitio de unión al péptido.





## **BIBLIOGRAFÍA**





1. *Basic Immunology and Its Clinical Application*. vol. 1444 (Springer Singapore, 2024).
2. Kumánovics, A., Takada, T., & Lindahl, K. F. (2003). Genomic organization of the mammalian MHC. *Annual review of immunology*, 21, 629–657.
3. Medhasi, S., & Chantratita, N. (2022). Human Leukocyte Antigen (HLA) System: Genetics and Association with Bacterial and Viral Infections. *Journal of immunology research*, 2022, 9710376
4. Klein, J., & Sato, A. (1998). Birth of the major histocompatibility complex. *Scandinavian journal of immunology*, 47(3), 199–209.
5. Kelley, J., Walter, L., & Trowsdale, J. (2005). Comparative genomics of major histocompatibility complexes. *Immunogenetics*, 56(10), 683–695.
6. Li, D., Sun, K., Zhao, Y., Lin, A., Li, S., Jiang, Y., & Feng, J. (2017). Polymorphism in the major histocompatibility complex (MHC class II B) genes of the Rufous-backed Bunting (*Emberiza jankowskii*). *PeerJ*, 5, e2917.
7. Martinez, C. K., & Monaco, J. J. (1991). Homology of proteasome subunits to a major histocompatibility complex-linked LMP gene. *Nature*, 353(6345), 664–667.
8. Glynne, R., Powis, S. H., Beck, S., Kelly, A., Kerr, L. A., & Trowsdale, J. (1991). A proteasome-related gene between the two ABC transporter loci in the class II region of the human MHC. *Nature*, 353(6342), 357–360.
9. Pishesha, N., Harmand, T. J., & Ploegh, H. L. (2022). A guide to antigen processing and presentation. *Nature reviews. Immunology*, 22(12), 751–764.
10. Downs, I., Vijayan, S., Sidiq, T., & Kobayashi, K. S. (2016). CITA/NLRC5: A critical transcriptional regulator of MHC class I gene expression. *BioFactors (Oxford, England)*, 42(4), 349–357.
11. Hunt, D. F., Henderson, R. A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A. L., Appella, E., & Engelhard, V. H. (1992). Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science (New York, N.Y.)*, 255(5049), 1261–1263.
12. Rammensee, H. G., Falk, K., & Rötzschke, O. (1993). Peptides naturally presented by MHC class I molecules. *Annual review of immunology*, 11, 213–244.
13. Goodfellow, P. N., Jones, E. A., Van Heyningen, V., Solomon, E., Bobrow, M., Miggiano, V., & Bodmer, W. F. (1975). The beta2-microglobulin gene is on chromosome 15 and not in the HL-A region. *Nature*, 254(5497), 267–269.
14. Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., & Wiley, D. C. (1987). Structure of the human class I histocompatibility antigen, HLA-A2. *Nature*, 329(6139), 506–512.
15. Cresswell, P., Turner, M. J., & Strominger, J. L. (1973). Papain-solubilized HL-A antigens from cultured human lymphocytes contain two peptide fragments. *Proceedings of the National Academy of Sciences of the United States of America*, 70(5), 1603–1607.

16. Cresswell P, Roche PA. Invariant chain-MHC class II complexes: always odd and never invariant. *Immunol Cell Biol.* 2014 Jul;92(6):471-2.
17. Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., & Wiley, D. C. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, 364(6432), 33–39.
18. Cresswell P. (1994). Assembly, transport, and function of MHC class II molecules. *Annual review of immunology*, 12, 259–293.
19. Knight, S. C., & Stagg, A. J. (1993). Antigen-presenting cell types. *Current opinion in immunology*, 5(3), 374–382.
20. Admon A. (2023). The biogenesis of the immunopeptidome. *Seminars in immunology*, 67, 101766.
21. Blum, J. S., Wearsch, P. A., & Cresswell, P. (2013). Pathways of antigen processing. *Annual review of immunology*, 31, 443–473.
22. Engelhard V. H. (1994). Structure of peptides associated with class I and class II MHC molecules. *Annual review of immunology*, 12, 181–207.
23. Madden, D. R., Gorga, J. C., Strominger, J. L., & Wiley, D. C. (1991). The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. *Nature*, 353(6342), 321–325.
24. Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A., & Wilson, I. A. (1992). Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. *Science (New York, N.Y.)*, 257(5072), 919–927.
25. Silver, M. L., Guo, H. C., Strominger, J. L., & Wiley, D. C. (1992). Atomic structure of a human MHC molecule presenting an influenza virus peptide. *Nature*, 360(6402), 367–369.
26. Rudensky, Y., Preston-Hurlburt, P., Hong, S. C., Barlow, A., & Janeway, C. A., Jr (1991). Sequence analysis of peptides bound to MHC class II molecules. *Nature*, 353(6345), 622–627.
27. Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R., & Wiley, D. C. (1991). Identification of self peptides bound to purified HLA-B27. *Nature*, 353(6342), 326–329.
28. Schumacher, T. N., De Bruijn, M. L., Vernie, L. N., Kast, W. M., Melief, C. J., Neefjes, J. J., & Ploegh, H. L. (1991). Peptide selection by MHC class I molecules. *Nature*, 350(6320), 703–706.
29. Guo, H. C., Jardetzky, T. S., Garrett, T. P., Lane, W. S., Strominger, J. L., & Wiley, D. C. (1992). Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. *Nature*, 360(6402), 364–366.

30. DiBrino, M., Parker, K. C., Shiloach, J., Knierman, M., Lukszo, J., Turner, R. V., Biddison, W. E., & Coligan, J. E. (1993). Endogenous peptides bound to HLA-A3 possess a specific combination of anchor residues that permit identification of potential antigenic peptides. *Proceedings of the National Academy of Sciences of the United States of America*, 90(4), 1508–1512.
31. Corr, M., Boyd, L. F., Frankel, S. R., Kozlowski, S., Padlan, E. A., & Margulies, D. H. (1992). Endogenous peptides of a soluble major histocompatibility complex class I molecule, H-2Lds: sequence motif, quantitative binding, and molecular modeling of the complex. *The Journal of experimental medicine*, 176(6), 1681–1692.
32. Chicz, R. M., Urban, R. G., Gorga, J. C., Vignali, D. A., Lane, W. S., & Strominger, J. L. (1993). Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. *The Journal of experimental medicine*, 178(1), 27–47.
33. Rudensky AY, Preston-Hurlburt, P., al-Ramadi, B. K., Rothbard, J., & Janeway, C. A., Jr (1992). Truncation variants of peptides isolated from MHC class II molecules suggest sequence motifs. *Nature*, 359(6394), 429–431.
34. Whitehead, G. G., Makino, S., Lien, C. L., & Keating, M. T. (2005). fgf20 is essential for initiating zebrafish fin regeneration. *Science (New York, N.Y.)*, 310(5756), 1957–1960.
35. Cardinaud S, Starck SR, Chandra P, Shastri N. The synthesis of truncated polypeptides for immune surveillance and viral evasion. *PLoS One*. 2010 Jan 21;5(1):e8692.
36. Brigl, M., & Brenner, M. B. (2004). CD1: antigen presentation and T cell function. *Annual review of immunology*, 22, 817–890.
37. Van Kaer, L., Wu, L., & Joyce, S. (2016). Mechanisms and Consequences of Antigen Presentation by CD1. *Trends in immunology*, 37(11), 738–754.
38. Geier, E., Pfeifer, G., Wilm, M., Lucchiari-Hartz, M., Baumeister, W., Eichmann, K., & Niedermann, G. (1999). A giant protease with potential to substitute for some functions of the proteasome. *Science (New York, N.Y.)*, 283(5404), 978–981.
39. Reits, E., Neijssen, J., Herberts, C., Benckhuijsen, W., Janssen, L., Drijfhout, J. W., & Neefjes, J. (2004). A major role for TPP1 in trimming proteasomal degradation products for MHC class I antigen presentation. *Immunity*, 20(4), 495–506.
40. Kessler, J. H., Khan, S., Seifert, U., Le Gall, S., Chow, K. M., Paschen, A., Bres-Vloemans, S. A., de Ru, A., van Montfoort, N., Franken, K. L., Benckhuijsen, W. E., Brooks, J. M., van Hall, T., Ray, K., Mulder, A., Doxiadis, I. I., van Swieten, P. F., Overkleeft, H. S., Prat, A., Tomkinson, B., ... Melief, C. J. (2011). Antigen processing by nardilysin and thimet oligopeptidase generates cytotoxic T cell epitopes. *Nature immunology*, 12(1), 45–53.
41. York, I. A., Mo, A. X., Lemerise, K., Zeng, W., Shen, Y., Abraham, C. R., Saric, T., Goldberg, A. L., & Rock, K. L. (2003). The cytosolic endopeptidase, thimet oligopeptidase, destroys antigenic peptides and limits the extent of MHC class I antigen presentation. *Immunity*, 18(3), 429–440.
42. Parmentier, N., Stroobant, V., Colau, D., de Diesbach, P., Morel, S., Chapiro, J., van Endert, P., & Van den Eynde, B. J. (2010). Production of an antigenic peptide by insulin-degrading enzyme. *Nature immunology*, 11(5), 449–454.

43. Shen, X. Z., Lukacher, A. E., Billet, S., Williams, I. R., & Bernstein, K. E. (2008). Expression of angiotensin-converting enzyme changes major histocompatibility complex class I peptide presentation by modifying C termini of peptide precursors. *The Journal of biological chemistry*, 283(15), 9957–9965.
44. Norbury, C. C., Basta, S., Donohue, K. B., Tschärke, D. C., Princiotta, M. F., Berglund, P., Gibbs, J., Bennink, J. R., & Yewdell, J. W. (2004). CD8+ T cell cross-priming via transfer of proteasome substrates. *Science (New York, N.Y.)*, 304(5675), 1318–1321.
45. Heemels, M. T., & Ploegh, H. (1995). Generation, translocation, and presentation of MHC class I-restricted peptides. *Annual review of biochemistry*, 64, 463–491.
46. York, I. A., Chang, S. C., Saric, T., Keys, J. A., Favreau, J. M., Goldberg, A. L., & Rock, K. L. (2002). The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8-9 residues. *Nature immunology*, 3(12), 1177–1184.
47. Serwold, T., Gonzalez, F., Kim, J., Jacob, R., & Shastri, N. (2002). ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature*, 419(6906), 480–483.
48. Saric, T., Chang, S. C., Hattori, A., York, I. A., Markant, S., Rock, K. L., Tsujimoto, M., & Goldberg, A. L. (2002). An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nature immunology*, 3(12), 1169–1176.
49. Saveanu, L., Carroll, O., Lindo, V., Del Val, M., Lopez, D., Lepelletier, Y., Greer, F., Schomburg, L., Fruci, D., Niedermann, G., & van Endert, P. M. (2005). Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. *Nature immunology*, 6(7), 689–697.
50. Howard J. C. (1995). Supply and transport of peptides presented by class I MHC molecules. *Current opinion in immunology*, 7(1), 69–76.
51. Parcej, D., & Tampé, R. (2010). ABC proteins in antigen translocation and viral inhibition. *Nature chemical biology*, 6(8), 572–580.
52. Del Val, M., Antón, L. C., Ramos, M., Muñoz-Abad, V., & Campos-Sánchez, E. (2020). Endogenous TAP-independent MHC-I antigen presentation: not just the ER lumen. *Current opinion in immunology*, 64, 9–14.
53. Rajagopalan, S., & Brenner, M. B. (1994). Calnexin retains unassembled major histocompatibility complex class I free heavy chains in the endoplasmic reticulum. *The Journal of experimental medicine*, 180(1), 407–412.
54. Sadasivan, B., Lehner, P. J., Ortmann, B., Spies, T., & Cresswell, P. (1996). Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP. *Immunity*, 5(2), 103–114.
55. Wearsch, P. A., Peaper, D. R., & Cresswell, P. (2011). Essential glycan-dependent interactions optimize MHC class I peptide loading. *Proceedings of the National Academy of Sciences of the United States of America*, 108(12), 4950–4955.

56. Ortmann, B., Copeman, J., Lehner, P. J., Sadasivan, B., Herberg, J. A., Grandea, A. G., Riddell, S. R., Tampé, R., Spies, T., Trowsdale, J., & Cresswell, P. (1997). A critical role for tapasin in the assembly and function of multimeric MHC class I-TAP complexes. *Science (New York, N.Y.)*, 277(5330), 1306–1309.
57. Morrice, N. A., & Powis, S. J. (1998). A role for the thiol-dependent reductase ERp57 in the assembly of MHC class I molecules. *Current biology : CB*, 8(12), 713–716.
58. Hughes, E. A., & Cresswell, P. (1998). The thiol oxidoreductase ERp57 is a component of the MHC class I peptide-loading complex. *Current biology : CB*, 8(12), 709–712.
59. Fisette, O., Wingbermühle, S., Tampé, R. *et al.* Molecular mechanism of peptide editing in the tapasin–MHC I complex. *Sci Rep* 6, 19085 (2016).
60. Townsend, A., Ohlén, C., Bastin, J., Ljunggren, H. G., Foster, L., & Kärre, K. (1989). Association of class I major histocompatibility heavy and light chains induced by viral peptides. *Nature*, 340(6233), 443–448.
61. Dongre, A. R., Kovats, S., deRoos, P., McCormack, A. L., Nakagawa, T., Paharkova-Vatchkova, V., Eng, J., Caldwell, H., Yates, J. R., 3rd, & Rudensky, A. Y. (2001). In vivo MHC class II presentation of cytosolic proteins revealed by rapid automated tandem mass spectrometry and functional analyses. *European journal of immunology*, 31(5), 1485–1494.
62. Scarcella, M., d'Angelo, D., Ciampa, M., Tafuri, S., Avallone, L., Pavone, L. M., & De Pasquale, V. (2022). The Key Role of Lysosomal Protease Cathepsins in Viral Infections. *International journal of molecular sciences*, 23(16), 9089.
63. Rock, K. L., Reits, E., & Neefjes, J. (2016). Present Yourself! By MHC Class I and MHC Class II Molecules. *Trends in immunology*, 37(11), 724–737.
64. Neefjes J. (1999). CIIV, MIIC and other compartments for MHC class II loading. *European journal of immunology*, 29(5), 1421–1425.
65. Mosyak, L., Zaller, D. M., & Wiley, D. C. (1998). The structure of HLA-DM, the peptide exchange catalyst that loads antigen onto class II MHC molecules during antigen presentation. *Immunity*, 9(3), 377–383.
66. Anders, A. K., Call, M. J., Schulze, M. S., Fowler, K. D., Schubert, D. A., Seth, N. P., Sundberg, E. J., & Wucherpfennig, K. W. (2011). HLA-DM captures partially empty HLA-DR molecules for catalyzed removal of peptide. *Nature immunology*, 12(1), 54–61.
67. Denzin, L. K., Fallas, J. L., Prendes, M., & Yi, W. (2005). Right place, right time, right peptide: DO keeps DM focused. *Immunological reviews*, 207, 279–292.
68. Guce, A. I., Mortimer, S. E., Yoon, T., Painter, C. A., Jiang, W., Mellins, E. D., & Stern, L. J. (2013). HLA-DO acts as a substrate mimic to inhibit HLA-DM by a competitive mechanism. *Nature structural & molecular biology*, 20(1), 90–98.
69. Grahame, R., Calin, A., Tudor, M., Kennedy, L., & Perrin, A. (1975). Ankylosing rheumatoid arthritis. *Rheumatology and rehabilitation*, 14(1), 25–30.

70. Schober, E., Schernthaner, G., & Mayr, W. R. (1981). HLA-DR antigens in insulin-dependent diabetes. *Archives of disease in childhood*, 56(3), 227–229.
71. Hazini, A., Fisher, K., & Seymour, L. (2021). Deregulation of HLA-I in cancer and its central importance for immunotherapy. *Journal for immunotherapy of cancer*, 9(8), e002899.
72. Garrido F. (2019). HLA Class-I Expression and Cancer Immunotherapy. *Advances in experimental medicine and biology*, 1151, 79–90.
73. Schachner, L. F., Phung, W., Han, G., Darwish, M., Bell, A., Mellors, J. S., Srzentic, K., Huguet, R., Blanchette, C., & Sandoval, W. (2022). High-Throughput, Quantitative Analysis of Peptide-Exchanged MHCI Complexes by Native Mass Spectrometry. *Analytical chemistry*, 94(42), 14593–14602.
74. Pourchet, A., Fuhrmann, S. R., Pilonis, K. A., Demaria, S., Frey, A. B., Mulvey, M., & Mohr, I. (2016). CD8(+) T-cell Immune Evasion Enables Oncolytic Virus Immunotherapy. *EBioMedicine*, 5, 59–67.
75. Liu DH, Mou FF, An M, Xia P. Human leukocyte antigen and tumor immunotherapy (Review). *Int J Oncol*. 2023 Jun;62(6):68.
76. Blackwell, J. M., Jamieson, S. E., & Burgner, D. (2009). HLA and infectious diseases. *Clinical microbiology reviews*, 22(2), 370–385.77.
77. Elliott Koshland, M. *SHEER LUCK MADE ME AN IMMUNOLOGIST*. (1996).
78. Hershko, A., & Ciechanover, A. (1998). The ubiquitin system. *Annual review of biochemistry*, 67, 425–479.
79. Heintz, L., & Meyer-Schwesinger, C. (2021). The Intertwining of Autophagy and the Ubiquitin Proteasome System in Podocyte (Patho)Physiology. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology*, 55(S4), 68–95.
80. Tomko, R. J., Jr, & Hochstrasser, M. (2013). Molecular architecture and assembly of the eukaryotic proteasome. *Annual review of biochemistry*, 82, 415–445.
81. Rock, K. L., & Goldberg, A. L. (1999). Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annual review of immunology*, 17, 739–779.
82. Rechsteiner, M., Hoffman, L., & Dubiel, W. (1993). The multicatalytic and 26 S proteases. *The Journal of biological chemistry*, 268(9), 6065–6068.
83. Bochtler, M., Ditzel, L., Groll, M., Hartmann, C., & Huber, R. (1999). The proteasome. *Annual review of biophysics and biomolecular structure*, 28, 295–317.
84. Wilk, S., & Orlowski, M. (1980). Cation-sensitive neutral endopeptidase: isolation and specificity of the bovine pituitary enzyme. *Journal of neurochemistry*, 35(5), 1172–1182.
85. Harris J. R. (1968). Release of a macromolecular protein component from human erythrocyte ghosts. *Biochimica et biophysica acta*, 150(3), 534–537.

86. Coux, O., Tanaka, K., & Goldberg, A. L. (1996). Structure and functions of the 20S and 26S proteasomes. *Annual review of biochemistry*, 65, 801–847.
87. Pickart C. M. (2001). Mechanisms underlying ubiquitination. *Annual review of biochemistry*, 70, 503–533.
88. Swatek, K. N., & Komander, D. (2016). Ubiquitin modifications. *Cell research*, 26(4), 399–422.
89. Clague, M. J., Urbé, S., & Komander, D. (2019). Breaking the chains: deubiquitylating enzyme specificity begets function. *Nature reviews. Molecular cell biology*, 20(6), 338–352.
90. Voges, D., Zwickl, P., & Baumeister, W. (1999). The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annual review of biochemistry*, 68, 1015–1068.
91. Grziwa, A., Baumeister, W., Dahlmann, B., & Kopp, F. (1991). Localization of subunits in proteasomes from *Thermoplasma acidophilum* by immunoelectron microscopy. *FEBS letters*, 290(1-2), 186–190.
92. Pühler, G., Weinkauff, S., Bachmann, L., Müller, S., Engel, A., Hegerl, R., & Baumeister, W. (1992). Subunit stoichiometry and three-dimensional arrangement in proteasomes from *Thermoplasma acidophilum*. *The EMBO journal*, 11(4), 1607–1616.
93. Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., & Huber, R. (1995). Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science (New York, N.Y.)*, 268(5210), 533–539.
94. Vimer, S., Ben-Nissan, G., Morgenstern, D., Kumar-Deshmukh, F., Polkinghorn, C., Quintyn, R. S., Vasil'ev, Y. V., Beckman, J. S., Elad, N., Wysocki, V. H., & Sharon, M. (2020). Comparative Structural Analysis of 20S Proteasome Ortholog Protein Complexes by Native Mass Spectrometry. *ACS central science*, 6(4), 573–588.
95. Marques, A. J., Palanimurugan, R., Matias, A. C., Ramos, P. C., & Dohmen, R. J. (2009). Catalytic mechanism and assembly of the proteasome. *Chemical reviews*, 109(4), 1509–1536.
96. Unno, M., Mizushima, T., Morimoto, Y., Tomisugi, Y., Tanaka, K., Yasuoka, N., & Tsukihara, T. (2002). The structure of the mammalian 20S proteasome at 2.75 Å resolution. *Structure (London, England : 1993)*, 10(5), 609–618.
97. Vigneron, N., & Van den Eynde, B. J. (2014). Proteasome subtypes and regulators in the processing of antigenic peptides presented by class I molecules of the major histocompatibility complex. *Biomolecules*, 4(4), 994–1025.
98. Heinemeyer, W., Fischer, M., Krimmer, T., Stachon, U., & Wolf, D. H. (1997). The active sites of the eukaryotic 20 S proteasome and their involvement in subunit precursor processing. *The Journal of biological chemistry*, 272(40), 25200–25209.
99. Seemüller, E., Lupas, A., Stock, D., Löwe, J., Huber, R., & Baumeister, W. (1995). Proteasome from *Thermoplasma acidophilum*: a threonine protease. *Science (New York, N.Y.)*, 268(5210), 579–582.

100. Groll, M., & Huber, R. (2003). Substrate access and processing by the 20S proteasome core particle. *The international journal of biochemistry & cell biology*, 35(5), 606–616
101. Chen, P., & Hochstrasser, M. (1996). Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly. *Cell*, 86(6), 961–972.
102. Hanada, K., Yewdell, J. W., & Yang, J. C. (2004). Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature*, 427(6971), 252–256.
103. Vigneron, N., Stroobant, V., Chapiro, J., Ooms, A., Degiovanni, G., Morel, S., van der Bruggen, P., Boon, T., & Van den Eynde, B. J. (2004). An antigenic peptide produced by peptide splicing in the proteasome. *Science (New York, N.Y.)*, 304(5670), 587–590.
104. Stadtmueller BM, Hill CP. Proteasome activators. *Mol Cell*. 2011 Jan 7;41(1):8-19.
105. Raule, M., Cerruti, F., Benaroudj, N., Migotti, R., Kikuchi, J., Bachi, A., Navon, A., Dittmar, G., & Cascio, P. (2014). PA28 $\alpha\beta$  reduces size and increases hydrophilicity of 20S immunoproteasome peptide products. *Chemistry & biology*, 21(4), 470–480.
106. Ben-Nissan, G., & Sharon, M. (2014). Regulating the 20S proteasome ubiquitin-independent degradation pathway. *Biomolecules*, 4(3), 862–884.
107. Alvarez, I., Sesma, L., Marcilla, M., Ramos, M., Marti, M., Camafeita, E., & de Castro, J. A. (2001). Identification of novel HLA-B27 ligands derived from polymorphic regions of its own or other class I molecules based on direct generation by 20 S proteasome. *The Journal of biological chemistry*, 276(35), 32729–32737.
108. Murata, S., Yashiroda, H., & Tanaka, K. (2009). Molecular mechanisms of proteasome assembly. *Nature reviews. Molecular cell biology*, 10(2), 104–115.
109. Zhang, H., Zhou, C., Mohammad, Z. *et al*. Structural basis of human 20S proteasome biogenesis. *Nat Commun* **15**, 8184 (2024).
110. Guillaume, B., Chapiro, J., Stroobant, V., Colau, D., Van Holle, B., Parvizi, G., Bousquet-Dubouch, M. P., Théate, I., Parmentier, N., & Van den Eynde, B. J. (2010). Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 107(43), 18599–18604.
111. Hirano, Y., Kaneko, T., Okamoto, K., Bai, M., Yashiroda, H., Furuyama, K., Kato, K., Tanaka, K., & Murata, S. (2008). Dissecting beta-ring assembly pathway of the mammalian 20S proteasome. *The EMBO journal*, 27(16), 2204–2213.
112. Groettrup, M., Kirk, C. J., & Basler, M. (2010). Proteasomes in immune cells: more than peptide producers?. *Nature reviews. Immunology*, 10(1), 73–78.
113. Aki, M., Shimbara, N., Takashina, M., Akiyama, K., Kagawa, S., Tamura, T., Tanahashi, N., Yoshimura, T., Tanaka, K., & Ichihara, A. (1994). Interferon-gamma induces different subunit organizations and functional diversity of proteasomes. *Journal of biochemistry*, 115(2), 257–269.



114. Gaczynska, M., Rock, K. L., & Goldberg, A. L. (1993). Gamma-interferon and expression of MHC genes regulate peptide hydrolysis by proteasomes. *Nature*, 365(6443), 264–267.
115. Witt, E., Zantopf, D., Schmidt, M., Kraft, R., Kloetzel, P. M., & Krüger, E. (2000). Characterisation of the newly identified human Ump1 homologue POMP and analysis of LMP7(beta 5i) incorporation into 20 S proteasomes. *Journal of molecular biology*, 301(1), 1–9.
116. Chapiro, J., Claverol, S., Piette, F., Ma, W., Stroobant, V., Guillaume, B., Gairin, J. E., Morel, S., Burlet-Schiltz, O., Monsarrat, B., Boon, T., & Van den Eynde, B. J. (2006). Destructive cleavage of antigenic peptides either by the immunoproteasome or by the standard proteasome results in differential antigen presentation. *Journal of immunology (Baltimore, Md. : 1950)*, 176(2), 1053–1061.
117. Murata, S., Sasaki, K., Kishimoto, T., Niwa, S., Hayashi, H., Takahama, Y., & Tanaka, K. (2007). Regulation of CD8+ T cell development by thymus-specific proteasomes. *Science (New York, N.Y.)*, 316(5829), 1349–1353.
118. Dahlmann, B., Ruppert, T., Kuehn, L., Merforth, S., & Kloetzel, P. M. (2000). Different proteasome subtypes in a single tissue exhibit different enzymatic properties. *Journal of molecular biology*, 303(5), 643–653.
119. Abi Habib, J., Lesenfans, J., Vigneron, N., & Van den Eynde, B. J. (2022). Functional Differences between Proteasome Subtypes. *Cells*, 11(3), 421.
120. Griffin, T. A., Nandi, D., Cruz, M., Fehling, H. J., Kaer, L. V., Monaco, J. J., & Colbert, R. A. (1998). Immunoproteasome assembly: cooperative incorporation of interferon gamma (IFN-gamma)-inducible subunits. *The Journal of experimental medicine*, 187(1), 97–104.
121. Groettrup, M., Standera, S., Stohwasser, R., & Kloetzel, P. M. (1997). The subunits MECL-1 and LMP2 are mutually required for incorporation into the 20S proteasome. *Proceedings of the National Academy of Sciences of the United States of America*, 94(17), 8970–8975.
122. Kingsbury, D. J., Griffin, T. A., & Colbert, R. A. (2000). Novel propeptide function in 20 S proteasome assembly influences beta subunit composition. *The Journal of biological chemistry*, 275(31), 24156–24162.
123. Kincaid, E. Z., Che, J. W., York, I., Escobar, H., Reyes-Vargas, E., Delgado, J. C., Welsh, R. M., Karow, M. L., Murphy, A. J., Valenzuela, D. M., Yancopoulos, G. D., & Rock, K. L. (2011). Mice completely lacking immunoproteasomes show major changes in antigen presentation. *Nature immunology*, 13(2), 129–135.
124. Tomaru, U., Ishizu, A., Murata, S., Miyatake, Y., Suzuki, S., Takahashi, S., Kazamaki, T., Ohara, J., Baba, T., Iwasaki, S., Fugo, K., Otsuka, N., Tanaka, K., & Kasahara, M. (2009). Exclusive expression of proteasome subunit {beta}5t in the human thymic cortex. *Blood*, 113(21), 5186–5191.
125. Ripen, A. M., Nitta, T., Murata, S., Tanaka, K., & Takahama, Y. (2011). Ontogeny of thymic cortical epithelial cells expressing the thymoproteasome subunit β5t. *European journal of immunology*, 41(5), 1278–1287.

126. Uddin, M. M. *et al.* Foxn1- $\beta$ 5t transcriptional axis controls CD8<sup>+</sup> T-cell production in the thymus. *Nat Commun* 8, (2017).
127. Murata, S., Takahama, Y., Kasahara, M., & Tanaka, K. (2018). The immunoproteasome and thymoproteasome: functions, evolution and human disease. *Nature immunology*, 19(9), 923–931.
128. Zhang Z, Clawson A, Realini C, Jensen CC, Knowlton JR, Hill CP, Rechsteiner M. Identification of an activation region in the proteasome activator REG $\alpha$ . *Proc Natl Acad Sci U S A*. 1998 Mar 17;95(6):2807-11.
129. Uechi, H., Hamazaki, J., & Murata, S. (2014). Characterization of the testis-specific proteasome subunit  $\alpha$ 4s in mammals. *The Journal of biological chemistry*, 289(18), 12365–12374.
130. Zhang, Z. H., Jiang, T. X., Chen, L. B., Zhou, W., Liu, Y., Gao, F., & Qiu, X. B. (2021). Proteasome subunit  $\alpha$ 4s is essential for formation of spermatoproteasomes and histone degradation during meiotic DNA repair in spermatocytes. *The Journal of biological chemistry*, 296, 100130.
131. Gómez-H, L., Felipe-Medina, N., Condezo, Y. B., Garcia-Valiente, R., Ramos, I., Suja, J. A., Barbero, J. L., Roig, I., Sánchez-Martín, M., de Rooij, D. G., Llano, E., & Pendas, A. M. (2019). The PSMA8 subunit of the spermatoproteasome is essential for proper meiotic exit and mouse fertility. *PLoS genetics*, 15(8), e1008316.
132. Burov, A., Funikov, S., Vagapova, E., Dalina, A., Rezvykh, A., Shyrokov, E., Lebedev, T., Grigorieva, E., Popenko, V., Leonova, O., Spasskaya, D., Spirin, P., Prassolov, V., Karpov, V., & Morozov, A. (2021). A Cell-Based Platform for the Investigation of Immunoproteasome Subunit  $\beta$ 5i Expression and Biology of  $\beta$ 5i-Containing Proteasomes. *Cells*, 10(11), 3049.
133. Guillaume, B., Stroobant, V., Bousquet-Dubouch, M. P., Colau, D., Chapiro, J., Parmentier, N., Dalet, A., & Van den Eynde, B. J. (2012). Analysis of the processing of seven human tumor antigens by intermediate proteasomes. *Journal of immunology (Baltimore, Md. : 1950)*, 189(7), 3538–3547.
134. Falk, K., Rötzschke, O., Stevanović, S., Jung, G., & Rammensee, H. G. (1991). Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*, 351(6324), 290–296.
135. Rammensee, H. G., Friede, T., & Stevanović, S. (1995). MHC ligands and peptide motifs: first listing. *Immunogenetics*, 41(4), 178–228.
136. Purcell, A. W., Ramarathnam, S. H., & Ternette, N. (2019). Mass spectrometry-based identification of MHC-bound peptides for immunopeptidomics. *Nature protocols*, 14(6), 1687–1707.
137. Shapiro, I. E., & Bassani-Sternberg, M. (2023). The impact of immunopeptidomics: From basic research to clinical implementation. *Seminars in immunology*, 66, 101727.
138. Dendrou, C. A., Petersen, J., Rossjohn, J., & Fugger, L. (2018). HLA variation and disease. *Nature reviews. Immunology*, 18(5), 325–339.

139. Matzaraki, V., Kumar, V., Wijmenga, C., & Zhernakova, A. (2017). The MHC locus and genetic susceptibility to autoimmune and infectious diseases. *Genome biology*, 18(1), 76.
140. Trowsdale, J. & Knight, J. C. Major histocompatibility complex genomics and human disease. *Annu Rev Genomics Hum Genet* 14, 301–323 (2013).
141. Sette, A., & Sidney, J. (1999). Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics*, 50(3-4), 201–212.
142. Lund, O., Nielsen, M., Kesmir, C., Petersen, A. G., Lundegaard, C., Worning, P., Sylvester-Hvid, C., Lamberth, K., Røder, G., Justesen, S., Buus, S., & Brunak, S. (2004). Definition of supertypes for HLA molecules using clustering of specificity matrices. *Immunogenetics*, 55(12), 797–810.
143. Sidney, J., Peters, B., Frahm, N. *et al.* HLA class I supertypes: a revised and updated classification. *BMC Immunol* 9, 1 (2008).
144. Chen, L., Shi, H., Koftori, D., Sekine, T., Nicastrì, A., Ternette, N., & Bowness, P. (2020). Identification of an Unconventional Subpeptidome Bound to the Behçet's Disease-associated HLA-B\*51:01 that is Regulated by Endoplasmic Reticulum Aminopeptidase 1 (ERAP1). *Molecular & cellular proteomics : MCP*, 19(5), 871–883.
145. Guasp, P., Barnea, E., González-Escribano, M. F., Jiménez-Reinoso, A., Regueiro, J. R., Admon, A., & López de Castro, J. A. (2017). The Behçet's disease-associated variant of the aminopeptidase ERAP1 shapes a low-affinity HLA-B\*51 peptidome by differential subpeptidome processing. *The Journal of biological chemistry*, 292(23), 9680–9689.
146. Guasp, P. *et al.* The Peptidome of Behçet's Disease-Associated HLA-B\*51:01 Includes Two Subpeptidomes Differentially Shaped by Endoplasmic Reticulum Aminopeptidase 1. *Arthritis and Rheumatology* 68, 505–515 (2016).
147. Gfeller, D., Guillaume, P., Michaux, J., Pak, H. S., Daniel, R. T., Racle, J., Coukos, G., & Bassani-Sternberg, M. (2018). The Length Distribution and Multiple Specificity of Naturally Presented HLA-I Ligands. *Journal of immunology (Baltimore, Md. : 1950)*, 201(12), 3705–3716.
148. Mukherjee, S., Sanchez-Bernabeu, A., Demmers, L. C., Wu, W., & Heck, A. J. R. (2021). The HLA Ligandome Comprises a Limited Repertoire of O-GlcNAcylated Antigens Preferentially Associated With HLA-B\*07:02. *Frontiers in immunology*, 12, 796584.
149. Sakaguchi, T., Ibe, M., Miwa, K., Kaneko, Y., Yokota, S., Tanaka, K., & Takiguchi, M. (1997). Binding of 8-mer to 11-mer peptides carrying the anchor residues to slow assembling HLA class I molecules (HLA-B\*5101). *Immunogenetics*, 45(4), 259–265.
150. Falk, K., Rötzschke, O., Takiguchi, M., Gnau, V., Stevanović, S., Jung, G., & Rammensee, H. G. (1995). Peptide motifs of HLA-B51, -B52 and -B78 molecules, and implications for Behçet's disease. *International immunology*, 7(2), 223–228.
151. Zhang, C., Anderson, A., & DeLisi, C. (1998). Structural principles that govern the peptide-binding motifs of class I MHC molecules. *Journal of molecular biology*, 281(5), 929–947.

152. Wei, X., Wang, S., Li, Z., Li, Z., Qu, Z., Wang, S., Zou, B., Liang, R., Xia, C., & Zhang, N. (2021). Peptidomes and Structures Illustrate Two Distinguishing Mechanisms of Alternating the Peptide Plasticity Caused by Swine MHC Class I Micropolymorphism. *Frontiers in immunology*, *12*, 592447.
153. Thomas, P., & Smart, T. G. (2005). HEK293 cell line: a vehicle for the expression of recombinant proteins. *Journal of pharmacological and toxicological methods*, *51*(3), 187–200.
154. Shaw, G., Morse, S., Ararat, M., & Graham, F. L. (2002). Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, *16*(8), 869–871.
155. Malm, M., Saghaleyni, R., Lundqvist, M. *et al.* Evolution from adherent to suspension: systems biology of HEK293 cell line development. *Sci Rep* **10**, 18996 (2020).
156. Zemmour, J., Little, A. M., Schendel, D. J., & Parham, P. (1992). The HLA-A,B "negative" mutant cell line C1R expresses a novel HLA-B35 allele, which also has a point mutation in the translation initiation codon. *Journal of immunology (Baltimore, Md. : 1950)*, *148*(6), 1941–1948.
157. Montserrat, V., Galocha, B., Marcilla, M., Vázquez, M., & López de Castro, J. A. (2006). HLA-B\*2704, an allotype associated with ankylosing spondylitis, is critically dependent on transporter associated with antigen processing and relatively independent of tapasin and immunoproteasome for maturation, surface expression, and T cell recognition: relationship to B\*2705 and B\*2706. *Journal of immunology (Baltimore, Md. : 1950)*, *177*(10), 7015–7023.
158. Benson, D. A. *et al.* GenBank. *Nucleic Acids Res* **41**, D36–D42 (2012).
159. Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics*, *13*, 134.
160. BD Life Sciences. FlowJo™ Software v10.8.1. (2023).
161. Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods*, *9*(7), 671–675.
162. Mastronarde, D. N. SerialEM: A Program for Automated Tilt Series Acquisition on Tecnai Microscopes Using Prediction of Specimen Position. *Microscopy and Microanalysis* **9**, 1182–1183 (2003).
163. Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature methods*, *14*(3), 290–296.
164. Scheres S. H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of structural biology*, *180*(3), 519–530.

165. Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., Quentin, D., Roderer, D., Tacke, S., Siebolds, B., Schubert, E., Shaikh, T. R., Lill, P., Gatsogiannis, C., & Raunser, S. (2019).
166. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Communications biology*, 2, 218166. He, J., Li, T. & Huang, S.-Y. Improvement of cryo-EM maps by simultaneous local and non-local deep learning. *Nat Commun* 14, 3217 (2023).
167. Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 4), 486–501.
168. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., ... Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta crystallographica. Section D, Structural biology*, 75(Pt 10), 861–877.
169. Vita, R., Blazeska, N., Marrama, D., IEDB Curation Team Members , Duesing, S., Bennett, J., Greenbaum, J., De Almeida Mendes, M., Mahita, J., Wheeler, D. K., Cantrell, J. R., Overton, J. A., Natale, D. A., Sette, A., & Peters, B. (2025). The Immune Epitope Database (IEDB): 2024 update. *Nucleic acids research*, 53(D1), D436–D443.
170. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic acids research*, 28(1), 235–242.
171. Maenaka, K., Maenaka, T., Tomiyama, H., Takiguchi, M., Stuart, D. I., & Jones, E. Y. (2000). Nonstandard peptide binding revealed by crystal structures of HLA-B\*5101 complexed with HIV immunodominant epitopes. *Journal of immunology (Baltimore, Md. : 1950)*, 165(6), 3260–3267.
172. Motozono, C., Kuse, N., Sun, X., Rizkallah, P. J., Fuller, A., Oka, S., Cole, D. K., Sewell, A. K., & Takiguchi, M. (2014). Molecular basis of a dominant T cell response to an HIV reverse transcriptase 8-mer epitope presented by the protective allele HLA-B\*51:01. *Journal of immunology (Baltimore, Md. : 1950)*, 192(7), 3428–3434.
173. Brennan, R. M., Petersen, J., Neller, M. A., Miles, J. J., Burrows, J. M., Smith, C., McCluskey, J., Khanna, R., Rossjohn, J., & Burrows, S. R. (2012).
174. The impact of a large and frequent deletion in the human TCR  $\beta$  locus on antiviral immunity. *Journal of immunology (Baltimore, Md. : 1950)*, 188(6), 2742–2748. Du, V. Y. *et al.* HIV-1–Specific CD8 T Cells Exhibit Limited Cross-Reactivity during Acute Infection. *The Journal of Immunology* 196, 3276–3286 (2016).
175. Rowntree, L. C., Nguyen, T. H. O., Halim, H., Purcell, A. W., Rossjohn, J., Gras, S., Kotsimbos, T. C., & Mifsud, N. A. (2018). Inability To Detect Cross-Reactive Memory T Cells Challenges the Frequency of Heterologous Immunity among Common Viruses. *Journal of immunology (Baltimore, Md. : 1950)*, 200(12), 3993–4003.

176. Chan, K. F., Gully, B. S., Gras, S., Beringer, D. X., Kjer-Nielsen, L., Cebon, J., McCluskey, J., Chen, W., & Rossjohn, J. (2018). Divergent T-cell receptor recognition modes of a HLA-I restricted extended tumour-associated peptide. *Nature communications*, 9(1), 1026.
177. Hwang, M. S., Miller, M. S., Thirawatananond, P., Douglass, J., Wright, K. M., Hsiue, E. H., Mog, B. J., Aytenfisu, T. Y., Murphy, M. B., Aitana Azurmendi, P., Skora, A. D., Pearlman, A. H., Paul, S., DiNapoli, S. R., Konig, M. F., Bettegowda, C., Pardoll, D. M., Papadopoulos, N., Kinzler, K. W., Vogelstein, B., ... Gabelli, S. B. (2021). Structural engineering of chimeric antigen receptors targeting HLA-restricted neoantigens. *Nature communications*, 12(1), 5271.
178. Lineburg, K. E., Grant, E. J., Swaminathan, S., Chatzileontiadou, D. S. M., Szeto, C., Sloane, H., Panikkar, A., Raju, J., Crooks, P., Rehan, S., Nguyen, A. T., Lekieffre, L., Neller, M. A., Tong, Z. W. M., Jayasinghe, D., Chew, K. Y., Lobos, C. A., Halim, H., Burrows, J. M., Riboldi-Tunncliffe, A., ... Gras, S. (2021). CD8<sup>+</sup> T cells specific for an immunodominant SARS-CoV-2 nucleocapsid epitope cross-react with selective seasonal coronaviruses. *Immunity*, 54(5), 1055–1065.e5.
179. Smith, K. J., Reid, S. W., Stuart, D. I., McMichael, A. J., Jones, E. Y., & Bell, J. I. (1996). An altered position of the alpha 2 helix of MHC class I is revealed by the crystal structure of HLA-B\*3501. *Immunity*, 4(3), 203–213.
180. Menssen, R., Orth, P., Ziegler, A., & Saenger, W. (1999). Decamer-like conformation of a nona-peptide bound to HLA-B\*3501 due to non-standard positioning of the C terminus. *Journal of molecular biology*, 285(2), 645–653.
181. Probst-Kepper, M., Hecht, H. J., Herrmann, H., Janke, V., Ocklenburg, F., Klempnauer, J., van den Eynde, B. J., & Weiss, S. (2004). Conformational restraints and flexibility of 14-meric peptides in complex with HLA-B\*3501. *Journal of immunology (Baltimore, Md. : 1950)*, 173(9), 5610–5616.182.
182. Miles, J. J. *et al.* CTL Recognition of a Bulged Viral Peptide Involves Biased TCR Selection. *The Journal of Immunology* 175, 3826–3834 (2005).
183. Tynan, F. E., Elhassen, D., Purcell, A. W., Burrows, J. M., Borg, N. A., Miles, J. J., Williamson, N. A., Green, K. J., Tellam, J., Kjer-Nielsen, L., McCluskey, J., Rossjohn, J., & Burrows, S. R. (2005). The immunogenicity of a viral cytotoxic T cell epitope is controlled by its MHC-bound conformation. *The Journal of experimental medicine*, 202(9), 1249–1260.
184. Hourigan, C. S., Harkiolaki, M., Peterson, N. A., Bell, J. I., Jones, E. Y., & O'Callaghan, C. A. (2006). The structure of the human allo-ligand HLA-B\*3501 in complex with a cytochrome p450 peptide: steric hindrance influences TCR allo-recognition. *European journal of immunology*, 36(12), 3288–3293.
185. Miles, J. J., Borg, N. A., Brennan, R. M., Tynan, F. E., Kjer-Nielsen, L., Silins, S. L., Bell, M. J., Burrows, J. M., McCluskey, J., Rossjohn, J., & Burrows, S. R. (2006). TCR alpha genes direct MHC restriction in the potent human T cell response to a class I-bound viral epitope. *Journal of immunology (Baltimore, Md. : 1950)*, 177(10), 6804–6814.

186. Archbold, J. K., Macdonald, W. A., Miles, J. J., Brennan, R. M., Kjer-Nielsen, L., McCluskey, J., Burrows, S. R., & Rossjohn, J. (2006). Alloreactivity between disparate cognate and allogeneic pMHC-I complexes is the result of highly focused, peptide-dependent structural mimicry. *The Journal of biological chemistry*, 281(45), 34324–34332.
187. Gras, S., Kedzierski, L., Valkenburg, S. A., Laurie, K., Liu, Y. C., Denholm, J. T., Richards, M. J., Rimmelzwaan, G. F., Kelso, A., Doherty, P. C., Turner, S. J., Rossjohn, J., & Kedzierska, K. (2010). Cross-reactive CD8+ T-cell immunity between the pandemic H1N1-2009 and H1N1-1918 influenza A viruses. *Proceedings of the National Academy of Sciences of the United States of America*, 107(28), 12599–12604.
188. Liu, Y. C., Chen, Z., Neller, M. A., Miles, J. J., Purcell, A. W., McCluskey, J., Burrows, S. R., Rossjohn, J., & Gras, S. (2014). A molecular basis for the interplay between T cells, viral mutants, and human leukocyte antigen micropolymorphism. *The Journal of biological chemistry*, 289(24), 16688–16698.
189. Alvarez, I., Martí, M., Vázquez, J., Camafeita, E., Ogueta, S., & López de Castro, J. A. (2001). The Cys-67 residue of HLA-B27 influences cell surface stability, peptide specificity, and T-cell antigen presentation. *The Journal of biological chemistry*, 276(52), 48740–48747.
190. Scholz, E., Mestre-Ferrer, A., Daura, X., García-Medel, N., Carrascal, M., James, E. A., Kwok, W. W., Canals, F., & Álvarez, I. (2016). A Comparative Analysis of the Peptide Repertoires of HLA-DR Molecules Differentially Associated With Rheumatoid Arthritis. *Arthritis & rheumatology (Hoboken, N.J.)*, 68(10), 2412–2421.
191. Scholz, E. M., Marcilla, M., Daura, X., Arribas-Layton, D., James, E. A., & Alvarez, I. (2017). Human Leukocyte Antigen (HLA)-DRB1\*15:01 and HLA-DRB5\*01:01 Present Complementary Peptide Repertoires. *Frontiers in immunology*, 8, 984.
192. The UniProt Consortium, UniProt: the Universal Protein Knowledgebase in 2025, *Nucleic Acids Research*, Volume 53, Issue D1, 6 January 2025, Pages D609–D617,
193. Reynisson, B., Alvarez, B., Paul, S., Peters, B., & Nielsen, M. (2020). NetMHCpan-4.1 and NetMHCIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic acids research*, 48(W1), W449–W454.
194. Kaabinejadian, S., Barra, C., Alvarez, B., Yari, H., Hildebrand, W. H., & Nielsen, M. (2022). Accurate MHC Motif Deconvolution of Immunopeptidomics Data Reveals a Significant Contribution of DRB3, 4 and 5 to the Total DR Immunopeptidome. *Frontiers in immunology*, 13, 835454.
195. Mestre-Ferrer A, Scholz E, Humet-Alsius J, Alvarez I. PRBAM: a new tool to analyze the MHC class I and HLA-DR anchor motifs. *Immunology*. 2019;156(2):187-198.
196. Andreatta, M., Alvarez, B., & Nielsen, M. (2017). GibbsCluster: unsupervised clustering and alignment of peptide sequences. *Nucleic acids research*, 45(W1), W458–W463.
197. Kisselev, A. F., & Goldberg, A. L. (2005). Monitoring activity and inhibition of 26S proteasomes with fluorogenic peptide substrates. *Methods in enzymology*, 398, 364–378.

198. Toste Rêgo, A., & da Fonseca, P. C. A. (2019). Characterization of Fully Recombinant Human 20S and 20S-PA200 Proteasome Complexes. *Molecular cell*, 76(1), 138–147.e5.
199. Ladi, E., Everett, C., Stivala, C. E., Daniels, B. E., Durk, M. R., Harris, S. F., Huestis, M. P., Purkey, H. E., Staben, S. T., Augustin, M., Blaesse, M., Steinbacher, S., Eidenschenk, C., Pappu, R., & Siu, M. (2019). Design and Evaluation of Highly Selective Human Immunoproteasome Inhibitors Reveal a Compensatory Process That Preserves Immune Cell Viability. *Journal of medicinal chemistry*, 62(15), 7032–7041.
200. Sakaguchi, T., Ibe, M., Miwa, K., Kaneko, Y., Yokota, S., Tanaka, K., & Takiguchi, M. (1997). Binding of 8-mer to 11-mer peptides carrying the anchor residues to slow assembling HLA class I molecules (HLA-B\*5101). *Immunogenetics*, 45(4), 259–265.
201. Uechi, H., Hamazaki, J., & Murata, S. (2014). Characterization of the testis-specific proteasome subunit  $\alpha 4s$  in mammals. *The Journal of biological chemistry*, 289(18), 12365–12374.
202. Kumar Deshmukh, F., Yaffe, D., Olshina, M. A., Ben-Nissan, G., & Sharon, M. (2019). The Contribution of the 20S Proteasome to Proteostasis. *Biomolecules*, 9(5), 190.
203. Paradela, A., Alvarez, I., García-Peydró, M., Sesma, L., Ramos, M., Vázquez, J., & López De Castro, J. A. (2000). Limited diversity of peptides related to an alloreactive T cell epitope in the HLA-B27-bound peptide repertoire results from restrictions at multiple steps along the processing-loading pathway. *Journal of immunology (Baltimore, Md. : 1950)*, 164(1), 329–337.
204. Yagüe J, Alvarez I, Rognan D, Ramos M, Vázquez J, de Castro JA. An N-acetylated natural ligand of human histocompatibility leukocyte antigen (HLA)-B39. Classical major histocompatibility complex class I proteins bind peptides with a blocked NH(2) terminus in vivo. *J Exp Med*. 2000 Jun 19;191(12):2083-92.
205. Ramos, M., Alvarez, I., Sesma, L., Logean, A., Rognan, D., & López de Castro, J. A. (2002). Molecular mimicry of an HLA-B27-derived ligand of arthritis-linked subtypes with chlamydial proteins. *The Journal of biological chemistry*, 277(40), 37573–37581.
206. Sesma, L., Alvarez, I., Marcilla, M., Paradela, A., & López de Castro, J. A. (2003). Species-specific differences in proteasomal processing and tapasin-mediated loading influence peptide presentation by HLA-B27 in murine cells. *The Journal of biological chemistry*, 278(47), 46461–46472.
207. Vigneron, N., Abi Habib, J., & Van den Eynde, B. J. (2015). The capture proteasome assay: A method to measure proteasome activity in vitro. *Analytical biochemistry*, 482, 7–15.
208. Rut, W., & Drag, M. (2016). Human 20S proteasome activity towards fluorogenic peptides of various chain lengths. *Biological chemistry*, 397(9), 921–926.
209. Bonet-Costa, V., Sun, P. Y., & Davies, K. J. A. (2019). Measuring redox effects on the activities of intracellular proteases such as the 20S Proteasome and the Immuno-Proteasome with fluorogenic peptides. *Free radical biology & medicine*, 143, 16–24.
210. Liggett, A., Crawford, L. J., Walker, B., Morris, T. C., & Irvine, A. E. (2010). Methods for measuring proteasome activity: current limitations and future developments. *Leukemia research*, 34(11), 1403–1409. 211.



211. Michalek, M. T., Grant, E. P., Gramm, C., Goldberg, A. L., & Rock, K. L. (1993). A role for the ubiquitin-dependent proteolytic pathway in MHC class I-restricted antigen presentation. *Nature*, 363(6429), 552–554
212. Dick, L. R., Aldrich, C., Jameson, S. C., Moomaw, C. R., Pramanik, B. C., Doyle, C. K., DeMartino, G. N., Bevan, M. J., Forman, J. M., & Slaughter, C. A. (1994). Proteolytic processing of ovalbumin and beta-galactosidase by the proteasome to a yield antigenic peptides. *Journal of immunology (Baltimore, Md. : 1950)*, 152(8), 3884–3894.
213. Kloetzel, P. M., & Ossendorp, F. (2004). Proteasome and peptidase function in MHC-class-I-mediated antigen presentation. *Current opinion in immunology*, 16(1), 76–81.
214. Pakjoo, M., Ahmadi, S.E., Zahedi, M. *et al.* Interplay between proteasome inhibitors and NF- $\kappa$ B pathway in leukemia and lymphoma: a comprehensive review on challenges ahead of proteasome inhibitors. *Cell Commun Signal* **22**, 105 (2024).
215. Khilji, M. S., Verstappen, D., Dahlby, T., Burstein Prause, M. C., Pihl, C., Bresson, S. E., Bryde, T. H., Keller Andersen, P. A., Klindt, K., Zivkovic, D., Bousquet-Dubouch, M. P., Tyrberg, B., Mandrup-Poulsen, T., & Marzec, M. T. (2020). The intermediate proteasome is constitutively expressed in pancreatic beta cells and upregulated by stimulatory, low concentrations of interleukin 1  $\beta$ . *PLoS one*, 15(2), e0222432.
216. Huber, E. M., Basler, M., Schwab, R., Heinemeyer, W., Kirk, C. J., Groettrup, M., & Groll, M. (2012). Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity. *Cell*, 148(4), 727–738.
217. Loy, C. A., & Trader, D. J. (2024). Primed for Interactions: Investigating the Primed Substrate Channel of the Proteasome for Improved Molecular Engagement. *Molecules (Basel, Switzerland)*, 29(14), 3356.
218. Huber, E. M., Basler, M., Schwab, R., Heinemeyer, W., Kirk, C. J., Groettrup, M., & Groll, M. (2012). Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity. *Cell*, 148(4), 727–738.
219. Woods, K., Knights, A. J., Anaka, M., Schittenhelm, R. B., Purcell, A. W., Behren, A., & Cebon, J. (2016). Mismatch in epitope specificities between IFN $\gamma$  inflamed and uninfamed conditions leads to escape from T lymphocyte killing in melanoma. *Journal for immunotherapy of cancer*, 4, 10.
220. Thomas, P., & Smart, T. G. (2005). HEK293 cell line: a vehicle for the expression of recombinant proteins. *Journal of pharmacological and toxicological methods*, 51(3), 187–200.
221. Lin, Y. C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., Chen, J., Speleman, F., Lambrechts, D., Van de Peer, Y., Tavernier, J., & Callewaert, N. (2014). Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nature communications*, 5, 4767.
222. Alvarez, M. G., Postan, M., Weatherly, D. B., Albareda, M. C., Sidney, J., Sette, A., Olivera, C., Armenti, A. H., Tarleton, R. L., & Laucella, S. A. (2008). HLA Class I-T cell epitopes from trans-sialidase proteins reveal functionally distinct subsets of CD8 $^{+}$  T cells in chronic Chagas disease. *PLoS neglected tropical diseases*, 2(9), e288.

223. Zhang, G. L., Bozic, I., Kwoh, C. K., August, J. T., & Brusic, V. (2007). Prediction of supertype-specific HLA class I binding peptides using support vector machines. *Journal of immunological methods*, 320(1-2), 143–154.
224. Zhang, G. L., DeLuca, D. S., Keskin, D. B., Chitkushev, L., Zlateva, T., Lund, O., Reinherz, E. L., & Brusic, V. (2011). MULTIPRED2: a computational system for large-scale identification of peptides predicted to bind to HLA supertypes and alleles. *Journal of immunological methods*, 374(1-2), 53–61.
225. Tirado-Herranz, A., Guasp, P., Pastor-Moreno, A., Area-Navarro, M., & Alvarez, I. (2023). Analysis of the different subpeptidomes presented by the HLA class I molecules of the B7 supertype. *Cellular immunology*, 387, 104707.
226. Lobos, C. A., Chatzileontiadou, D. S., Sok, B., Almedia, C. A., Halim, H., D'Orsogna, L., & Gras, S. (2024). Molecular insights into the HLA-B35 molecules' classification associated with HIV control. *Immunology and cell biology*, 102(1), 34–45.
227. Mehra, Simmi; Tiwari, Aseem Kumar; Rajvanshi, Chhavi; Chauhan, Rajni; Sud, Randhir<sup>1</sup>; Soin, Arvinder Singh<sup>2</sup>. An Analysis of Association of Human Leukocyte Antigen and Liver Disease Patients in a Tertiary Care Hospital, India. *Global Journal of Transfusion Medicine* 6(1):p 65-71, Jan–Jun 2021.
228. Agostini, S., Mancuso, R., Guerini, F. R., D'Alfonso, S., Agliardi, C., Hernis, A., Zanzottera, M., Barizzone, N., Leone, M. A., Caputo, D., Rovaris, M., & Clerici, M. (2018). HLA alleles modulate EBV viral load in multiple sclerosis. *Journal of translational medicine*, 16(1), 80.



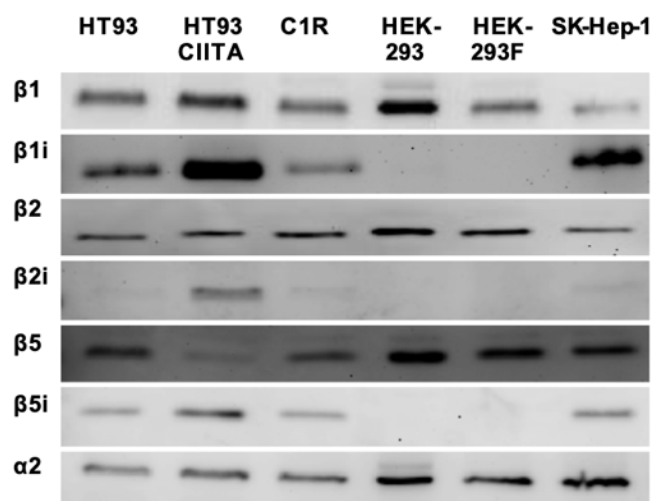


## **ANEXOS**



|                |                                                                                                                                                                                               |
|----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Anexo 1</b> | Expresión de las subunidades catalíticas $\beta 1$ , $\beta 1i$ , $\beta 2$ , $\beta 2i$ , $\beta 5$ y $\beta 5i$ en diferentes líneas celulares                                              |
| <b>Anexo 2</b> | Residuos de anclaje del inmunopeptidoma de HEK-293F calculados con el <i>software</i> PRBAM                                                                                                   |
| <b>Anexo 3</b> | Residuos de anclaje del inmunopeptidoma de HEK-293F- $\beta 5i$ -StII calculados con el <i>software</i> PRBAM                                                                                 |
| <b>Anexo 4</b> | Residuos de anclaje del inmunopeptidoma de HEK-293 calculados con el <i>software</i> PRBAM                                                                                                    |
| <b>Anexo 5</b> | Residuos de anclaje del inmunopeptidoma de HEK-293 $\beta 5i$ calculados con el <i>software</i> PRBAM                                                                                         |
| <b>Anexo 6</b> | Ensayos de evaluación de la actividad catalítica del proteasoma                                                                                                                               |
| <b>Anexo 7</b> | Residuos de anclaje de las moléculas del supertipo B7 disponibles en IEDB calculados con el <i>software</i> PRBAM                                                                             |
| <b>Anexo 8</b> | Residuos de anclaje de los transfectantes de HLA-B*51:01, -B*35:01 y -B*07:02 calculados con el <i>software</i> PRBAM                                                                         |
| <b>Anexo 9</b> | <b><i>Analysis of the different subpeptidomes presented by the HLA class I molecules of the B7 supertype.</i></b> Cell Immunol . 2023 May ; 387: 104707.<br>doi:10.1016/j.cellimm.2023.104707 |

**Anexo 1:** Expresión de las subunidades catalíticas  $\beta 1$ ,  $\beta 1i$ ,  $\beta 2$ ,  $\beta 2i$ ,  $\beta 5$  y  $\beta 5i$  en diferentes líneas celulares



**Expresión de las subunidades catalíticas  $\beta 1$ ,  $\beta 1i$ ,  $\beta 2$ ,  $\beta 2i$ ,  $\beta 5$  y  $\beta 5i$  en diferentes líneas celulares.** Se realizó un extracto proteico de las líneas celulares HT93, HT-93-CIITA, HMy2.C1R, HEK-293, HEK-293F y SK-Hep-1. Alrededor de 20 ug de proteína total, de cada una de las muestras, se analizaron por Western Blot usando anticuerpos específicos contra las subunidades catalíticas del proteasoma constitutivo o del inmunoproteasoma. Como control de carga se usó el mAb MCP21, específico de la subunidad  $\alpha 2$  (Sholz, E, Tesis Doctoral: *Mecanismos de Procesamiento y Presentación Antigénica en Autoinmunidad*. Universidad Autónoma de Barcelona, 2017)



**Anexo 2:** Residuos de anclaje del inmunopeptidoma de HEK-293F calculados con el software PRBAM

- **HLA-A\*02:01**



- HLA-A\*03:01



- HLA-B\*07:02



- HLA-C\*07:02



**Anexo 3:** Residuos de anclaje del inmunopeptidoma de HEK-293F- $\beta$ 5i-StII calculados con el *software* PRBAM

- **HLA-A\*02:01**



- HLA-A\*03:01



- HLA-B\*07:02



- HLA-C\*07:02





# **Anexo 4:** Residuos de anclaje del inmunopeptidoma de HEK-293 calculados con el *software* PRBAM

## - **HLA-A\*02:01**



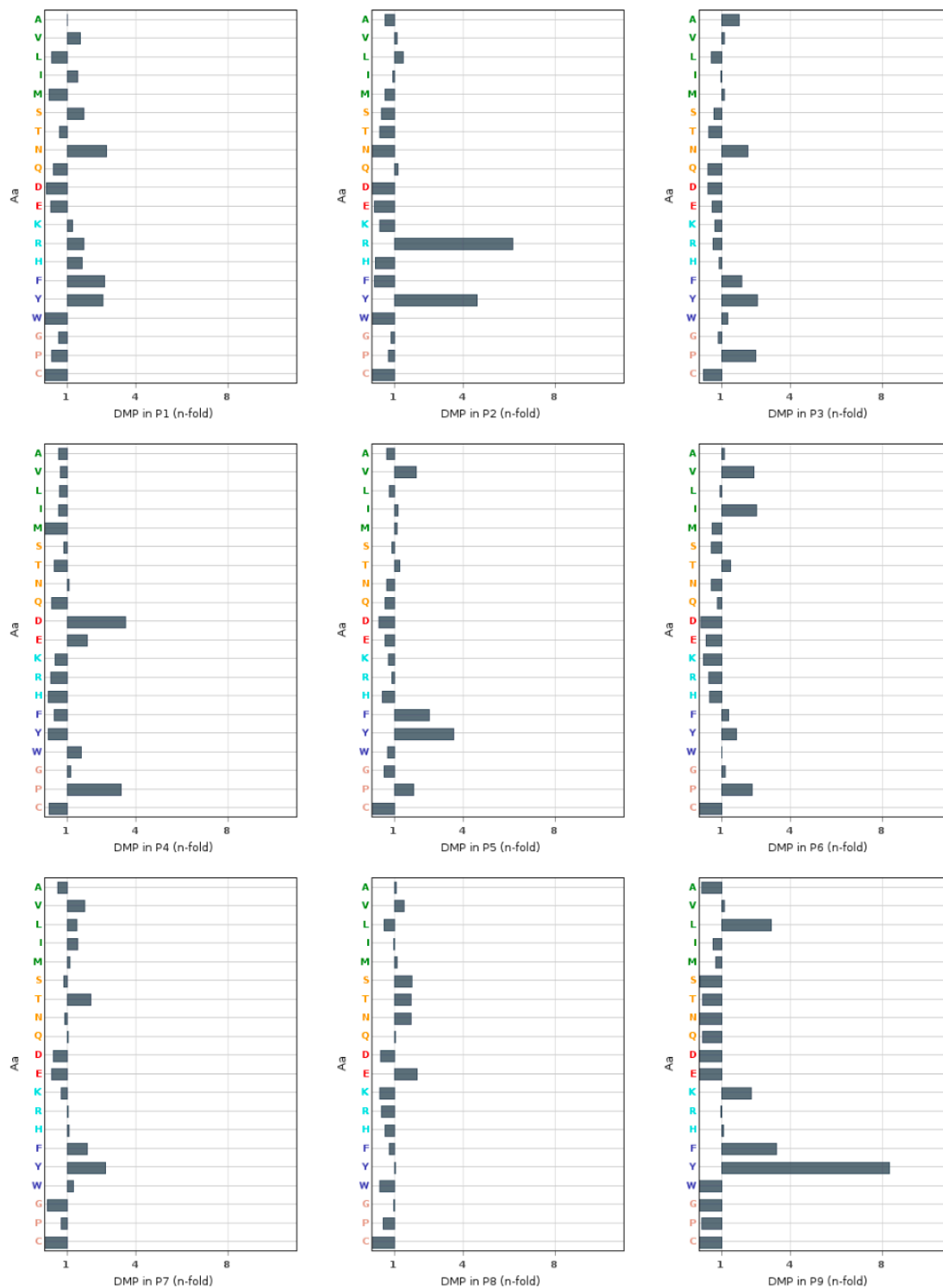
- HLA-A\*03:01



- HLA-B\*07:02



- HLA-C\*07:02



**Anexo 5:** Residuos de anclaje del inmunopeptidoma de HEK-293-β5i calculados con el *software* PRBAM

- **HLA-A\*02:01**



- HLA-A\*03:01



- HLA-B\*07:02

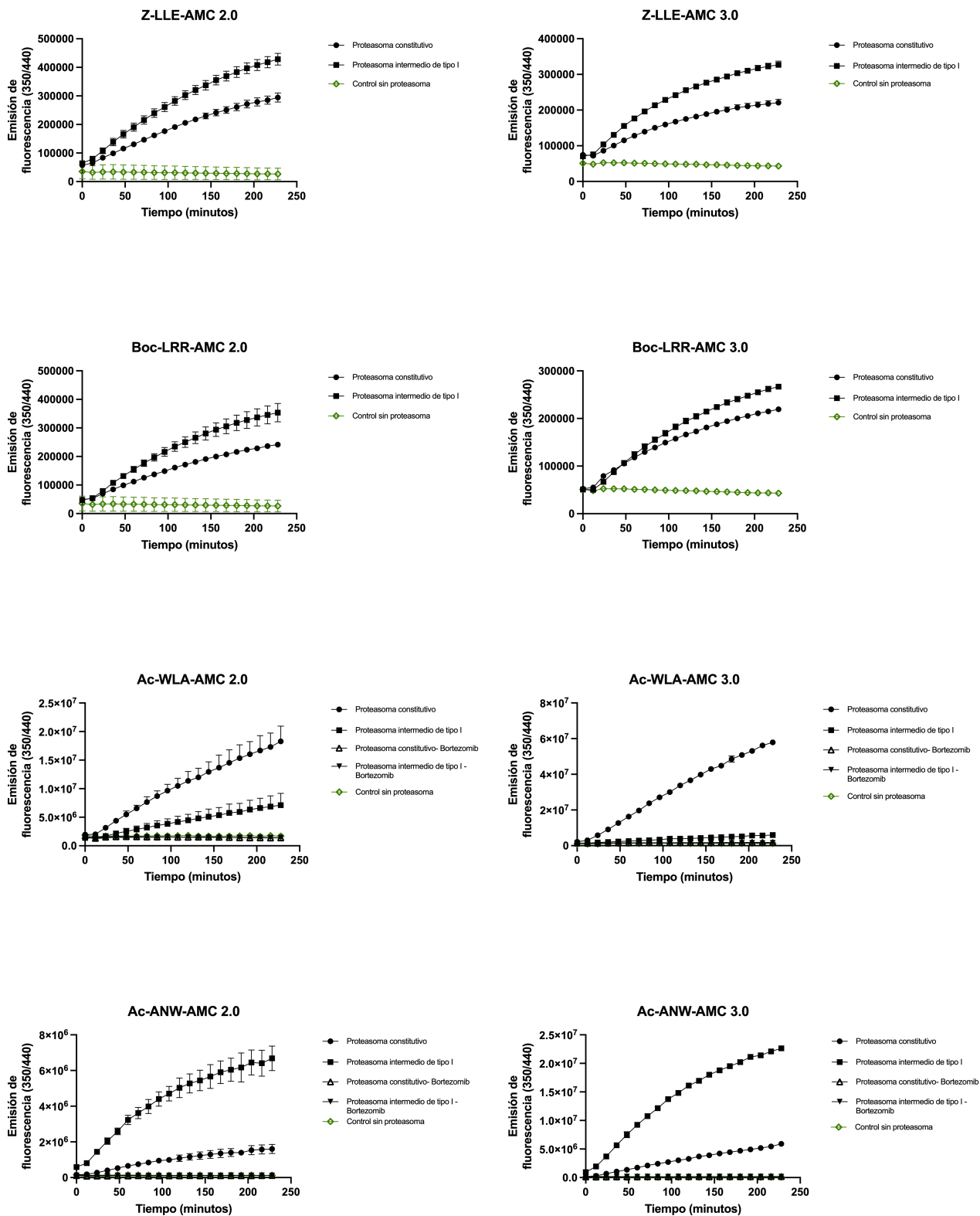


- HLA-C\*07:02



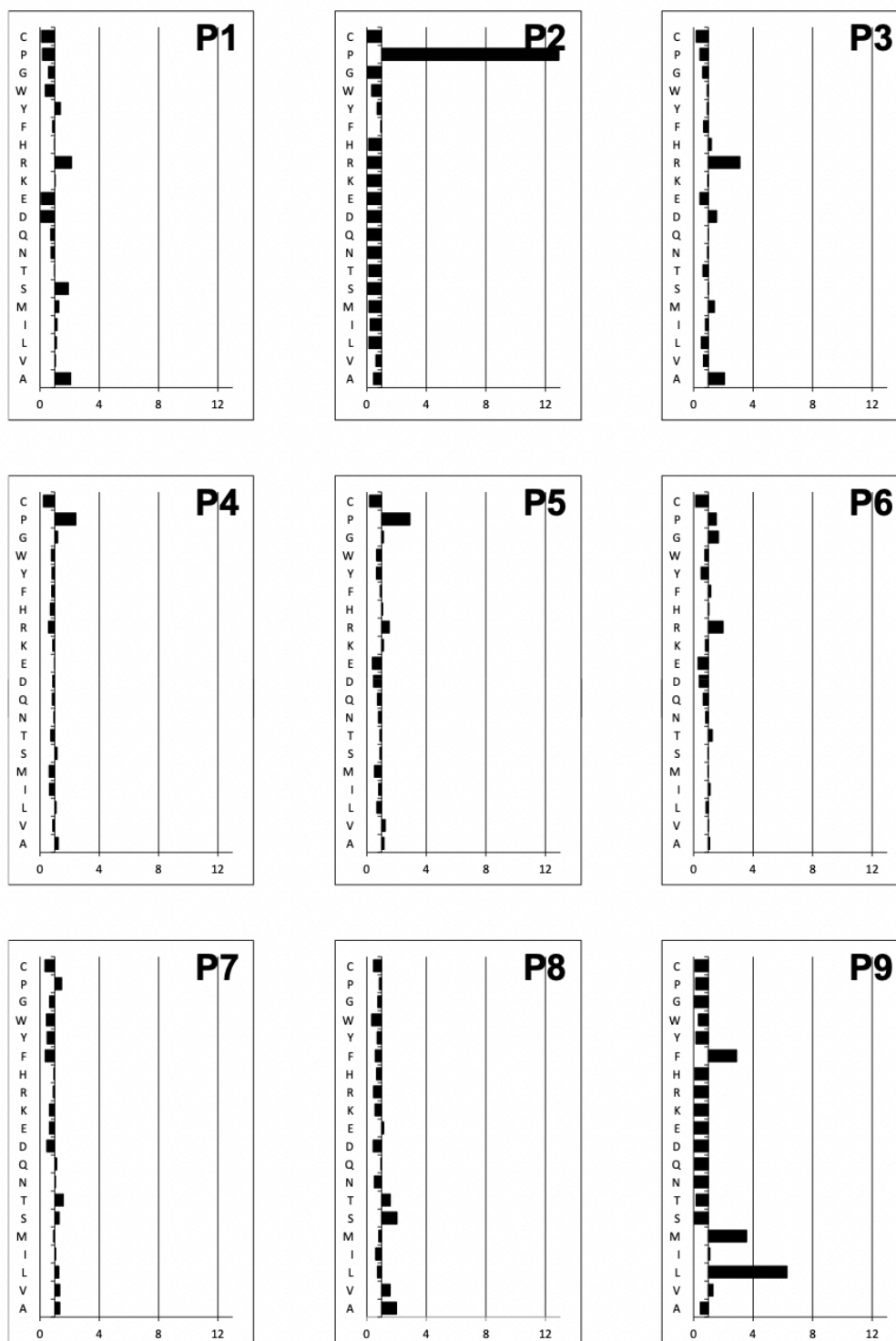


Anexo 6: Ensayos de evaluación de actividad catalítica del proteasoma

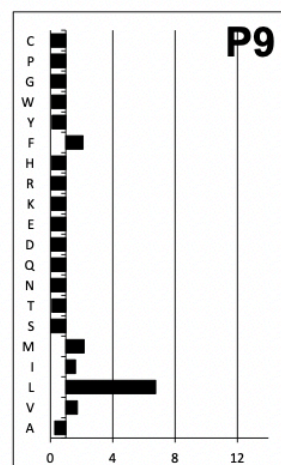
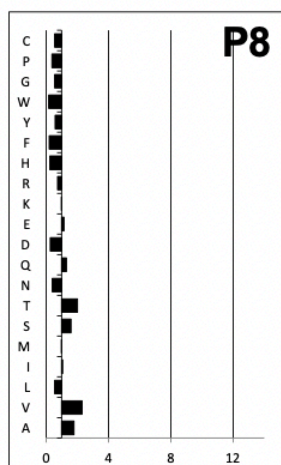
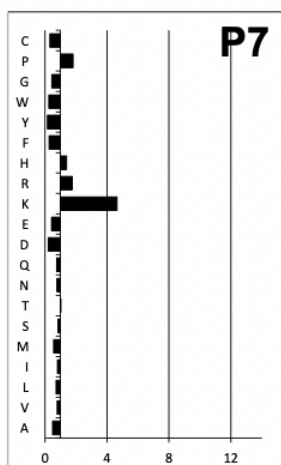
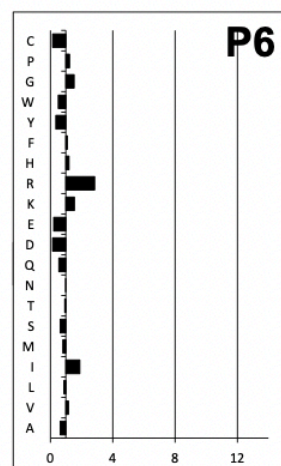
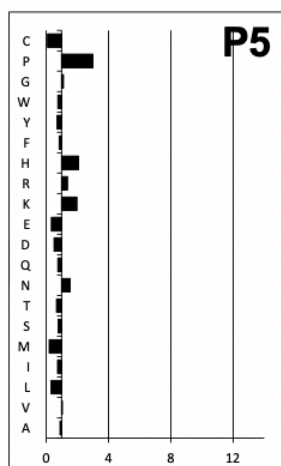
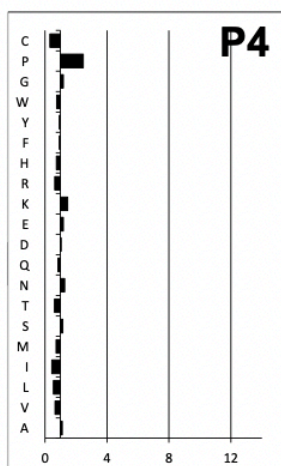
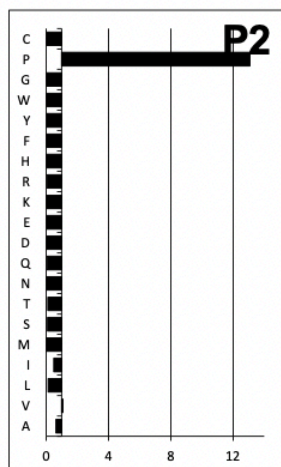
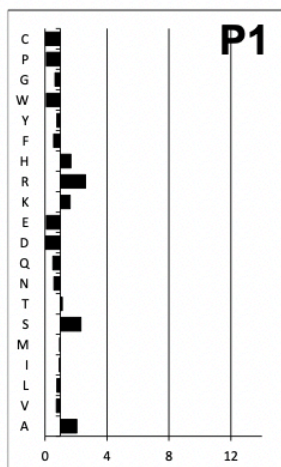


**Anexo 7:** Residuos de anclaje de las moléculas de HLA-I del supertipo B7 disponibles en la literatura calculados con el *software* PRBAM.

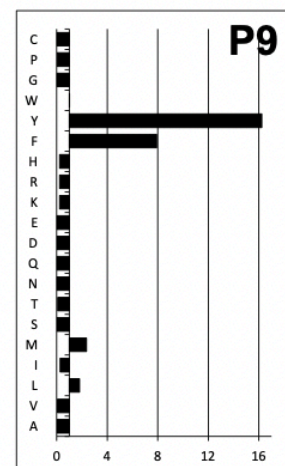
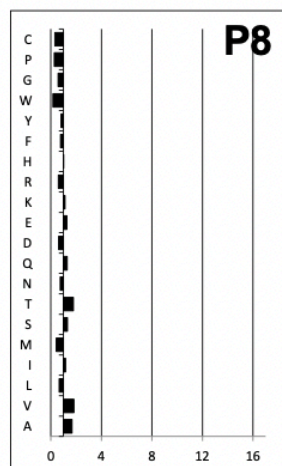
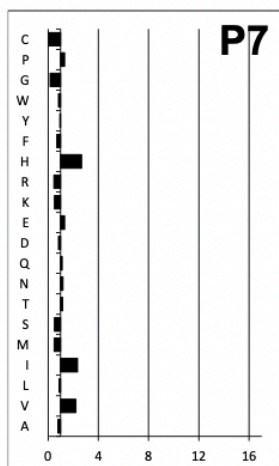
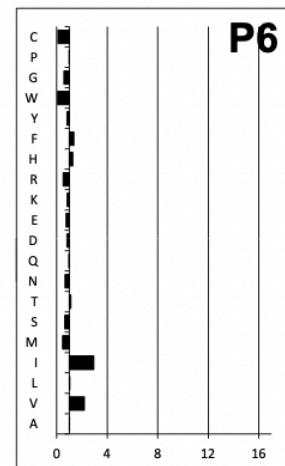
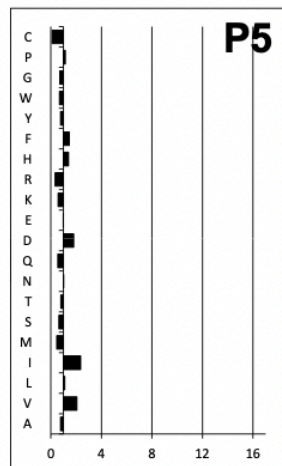
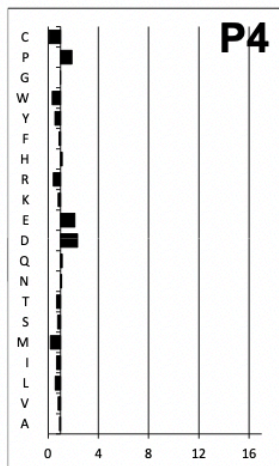
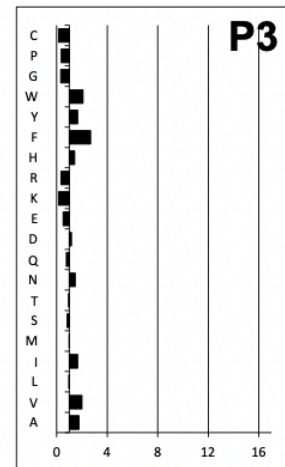
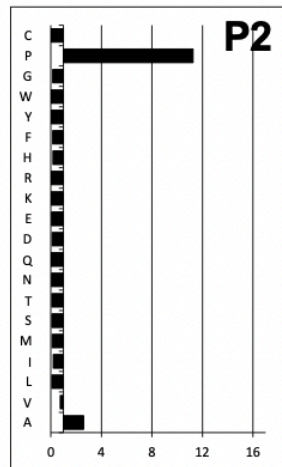
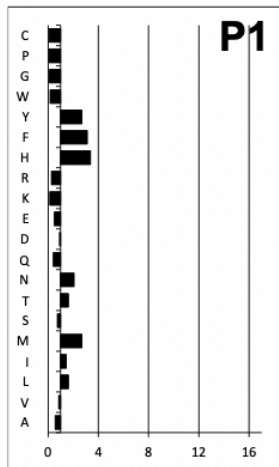
- **HLA-B\*07:02**



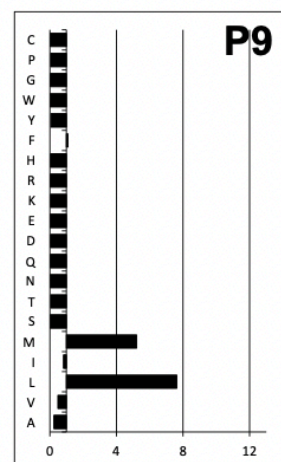
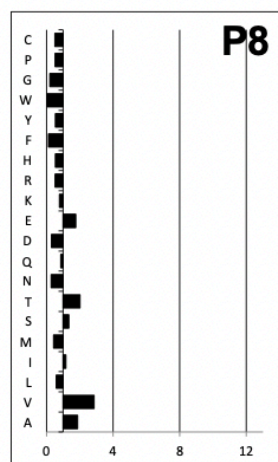
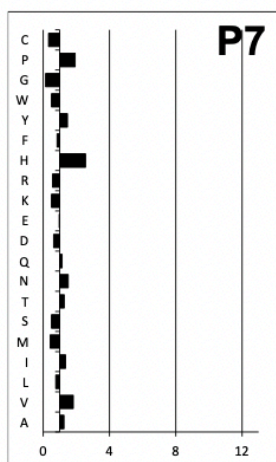
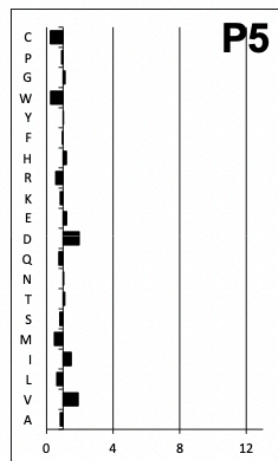
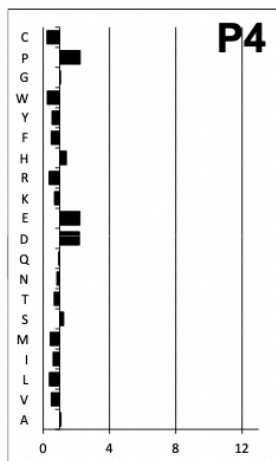
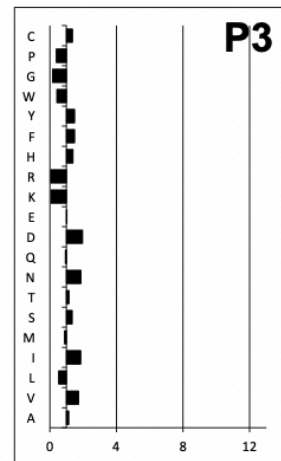
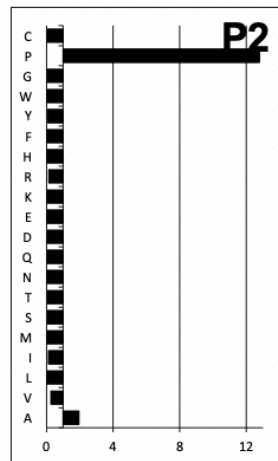
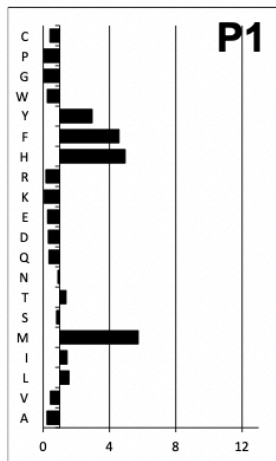
- HLA-B\*07:04



- **HLA-B\*35:01**

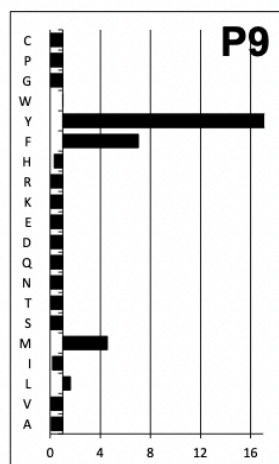
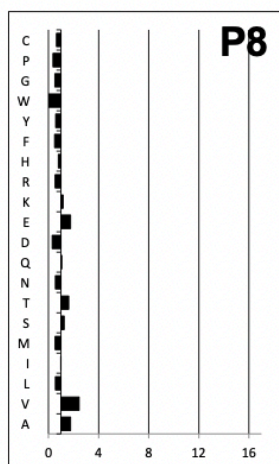
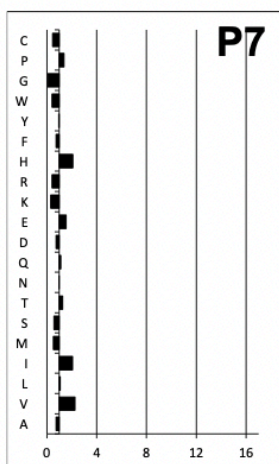
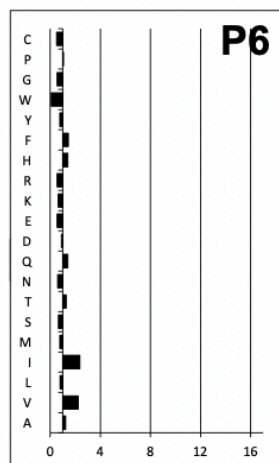
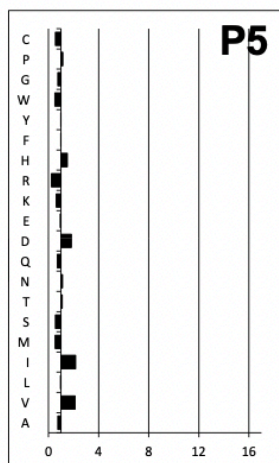
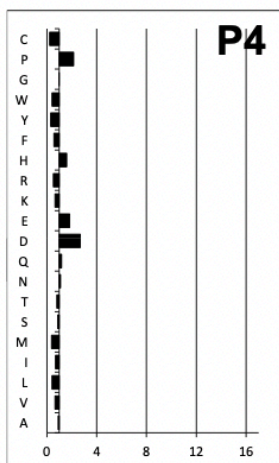
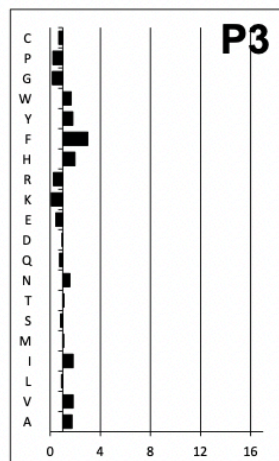
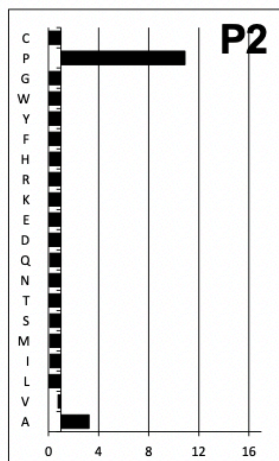
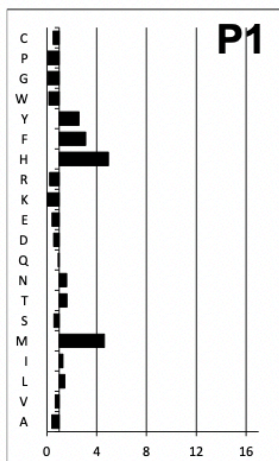


- HLA-B\*35:03

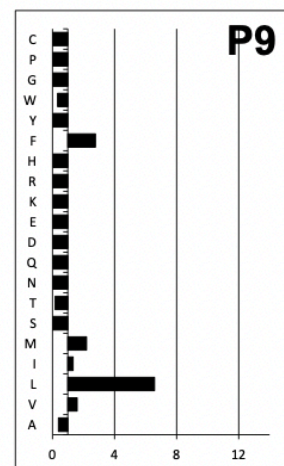
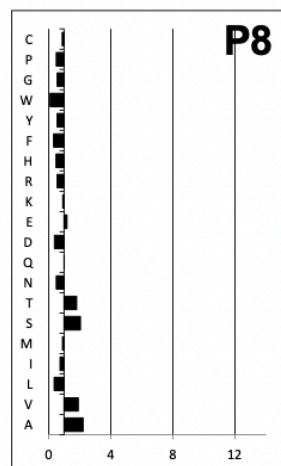
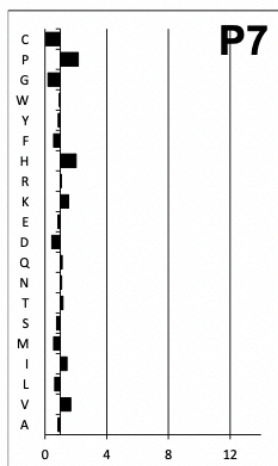
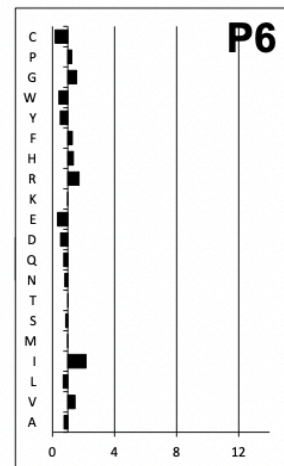
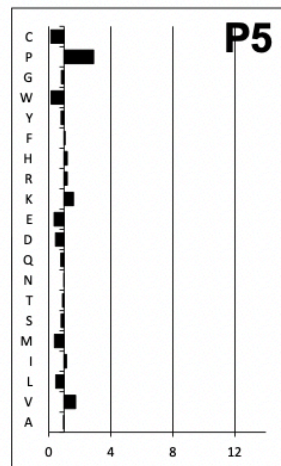
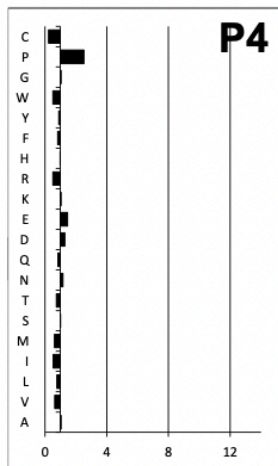
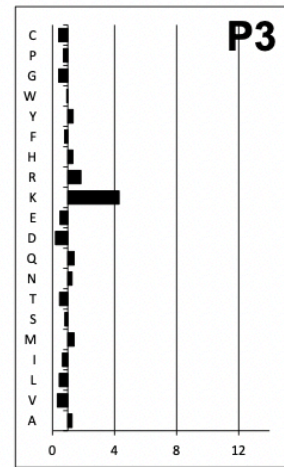
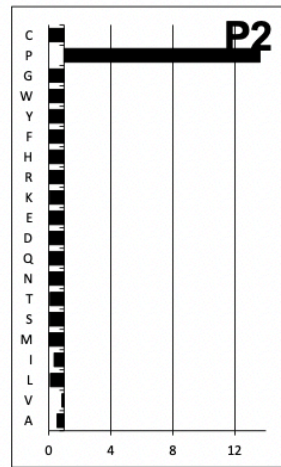
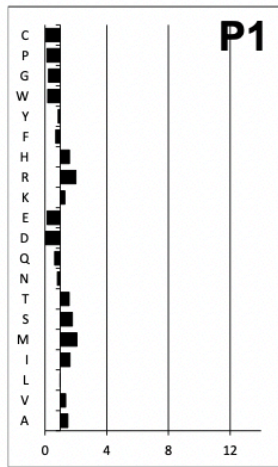




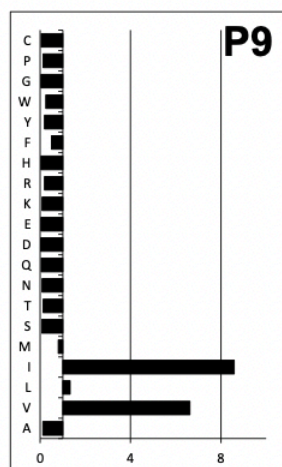
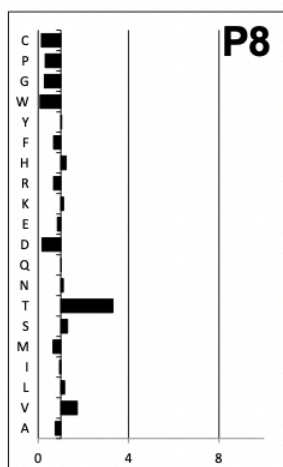
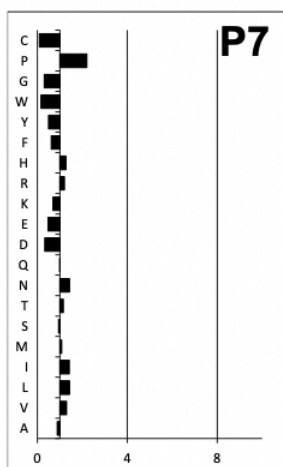
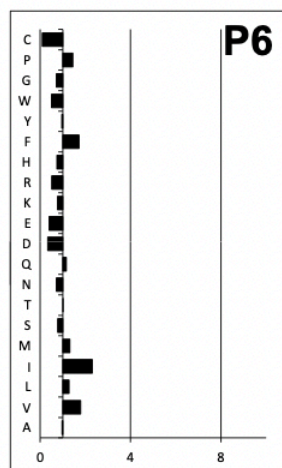
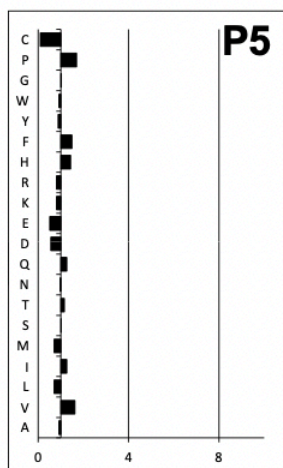
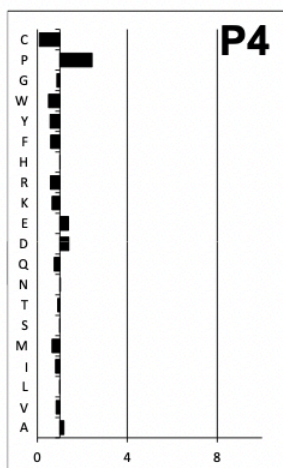
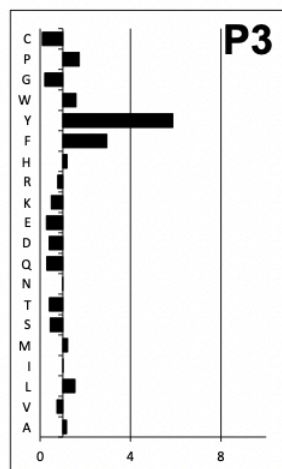
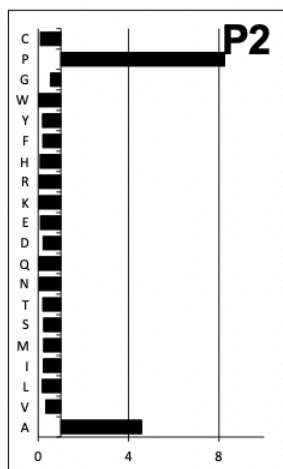
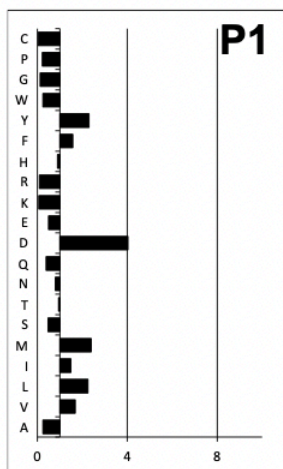
- HLA-B\*35:07



- HLA-B\*42:01

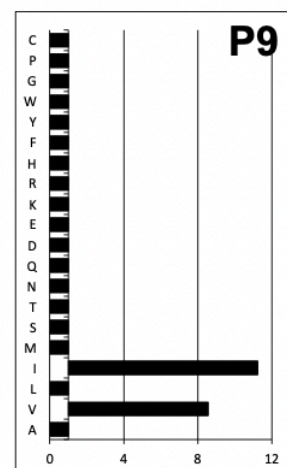
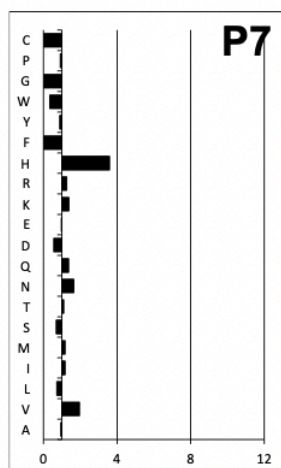
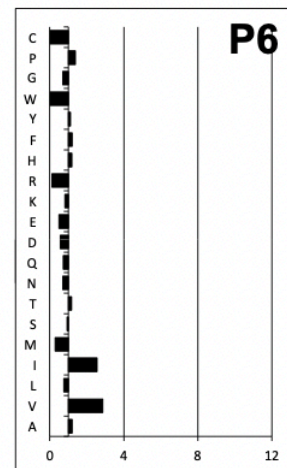
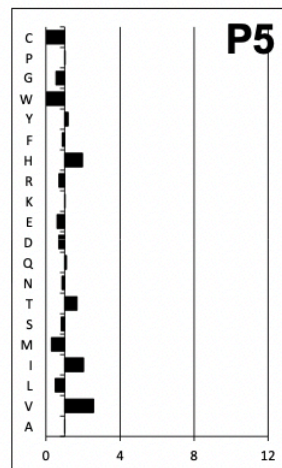
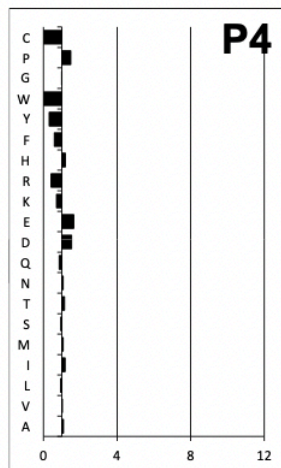
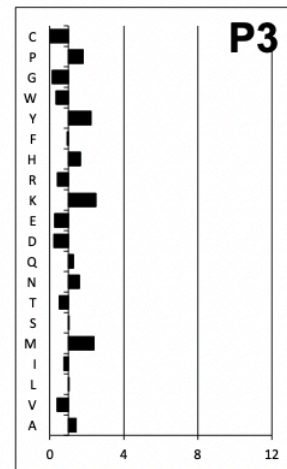
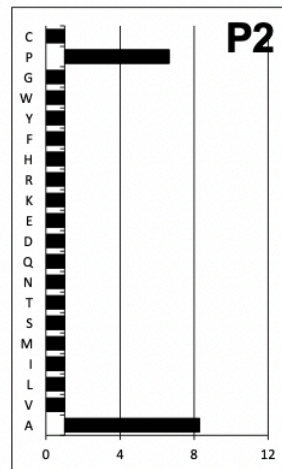
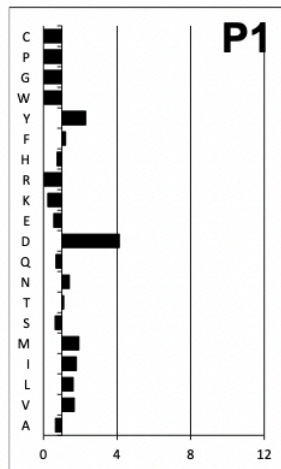


- HLA-B\*51:01

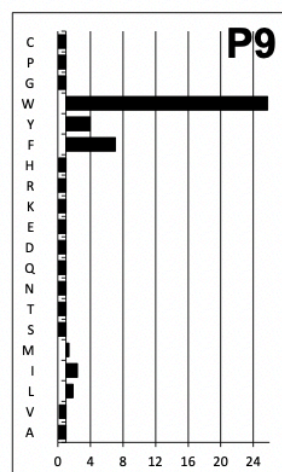
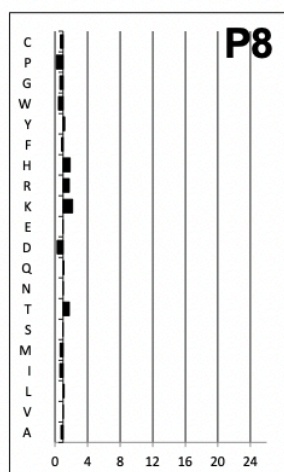
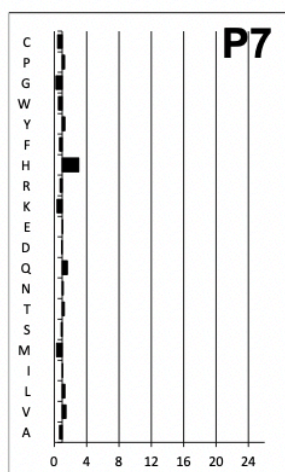
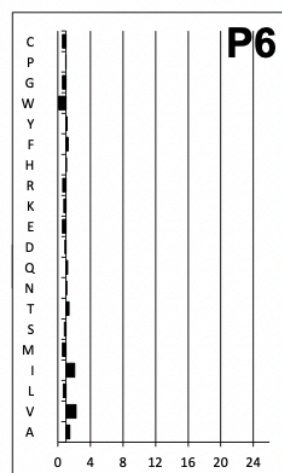
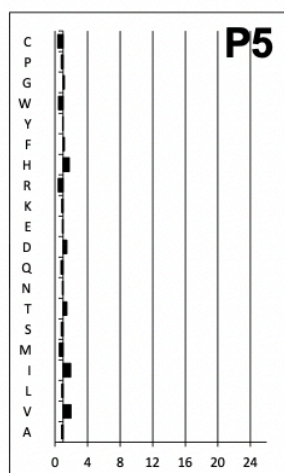
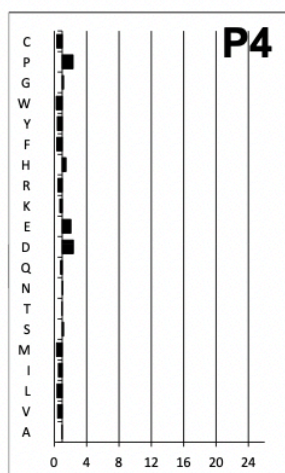
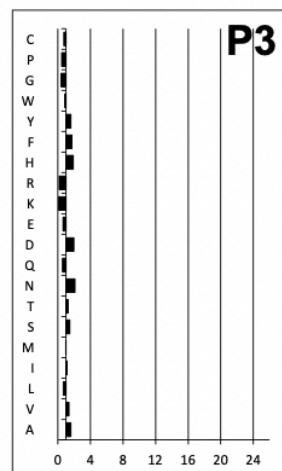
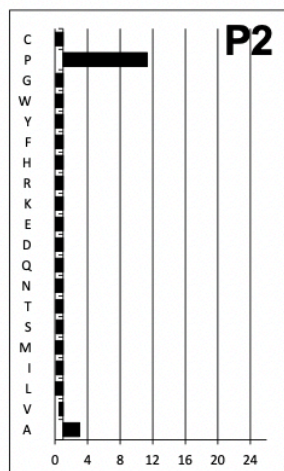
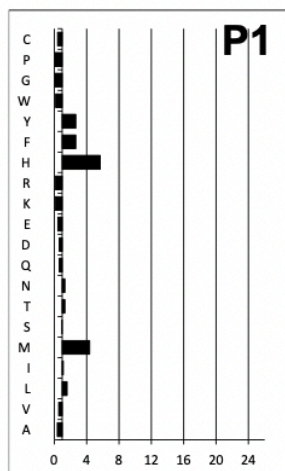




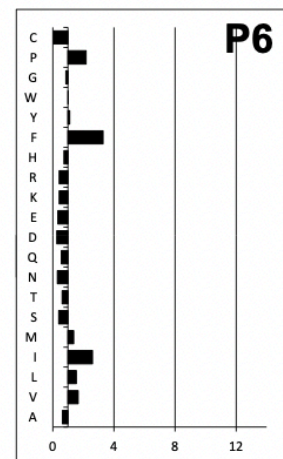
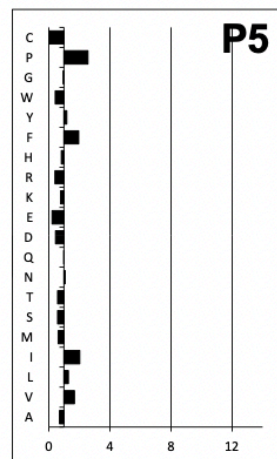
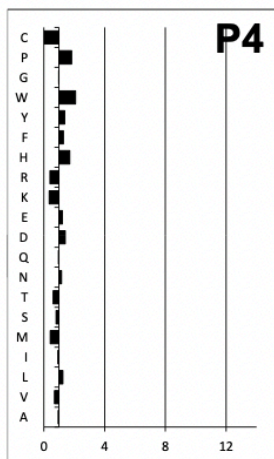
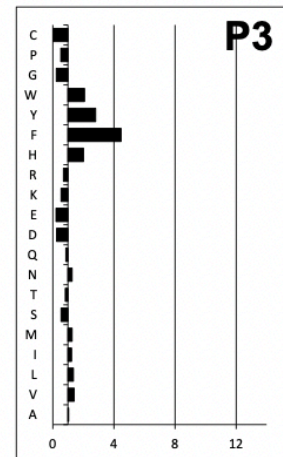
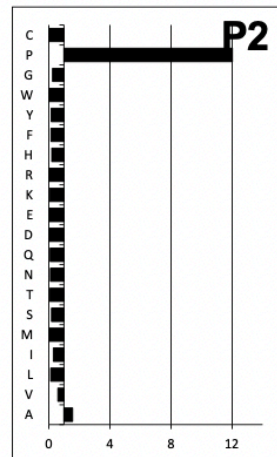
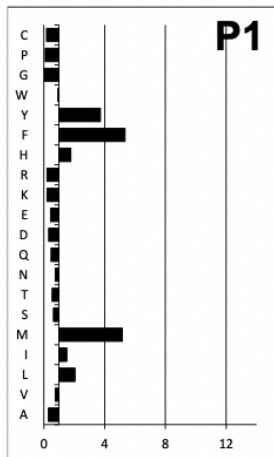
- **HLA-B\*51:08**



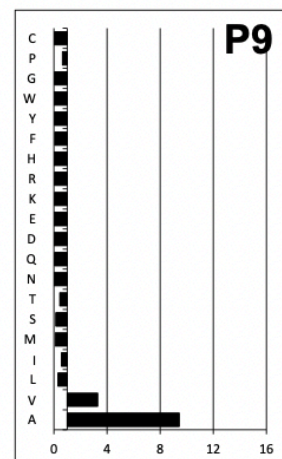
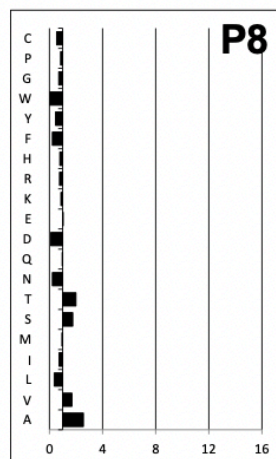
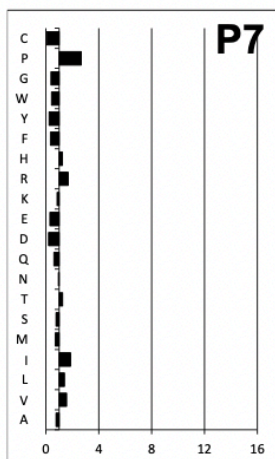
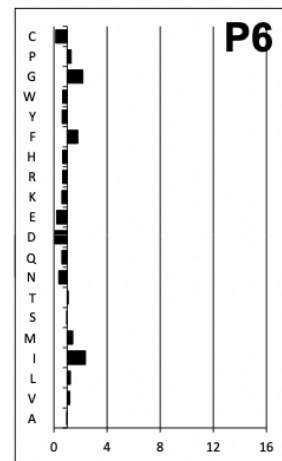
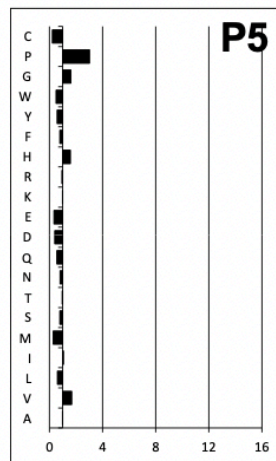
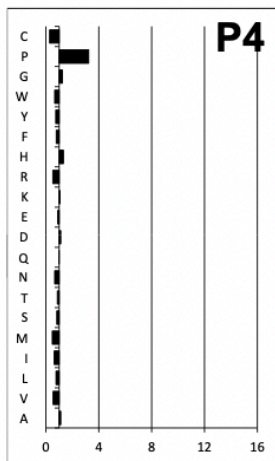
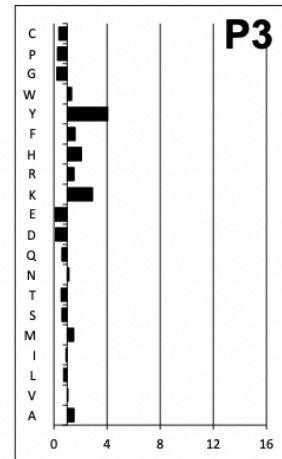
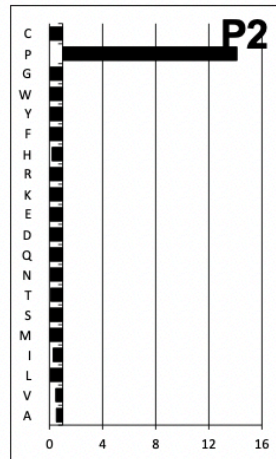
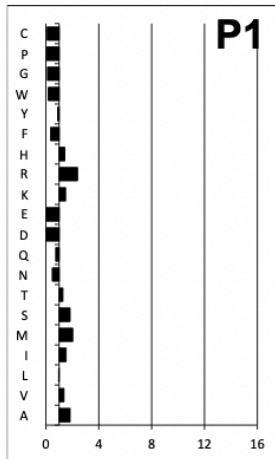
- HLA-B\*53:01



- HLA-B\*54:01

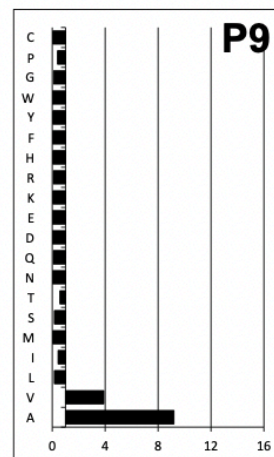
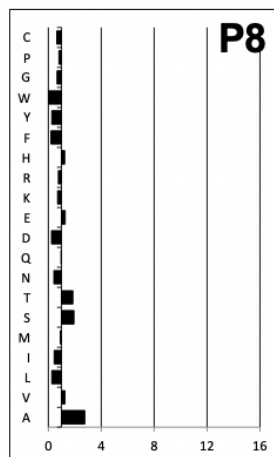
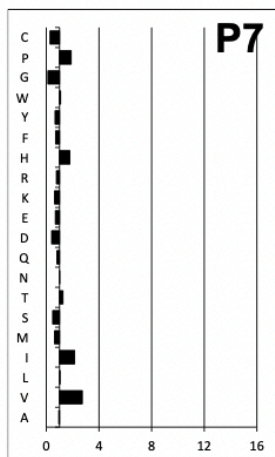
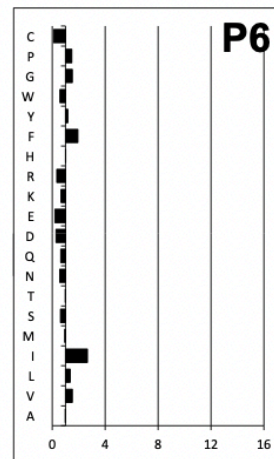
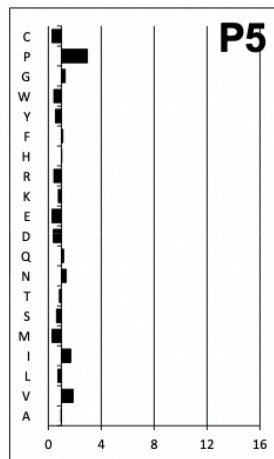
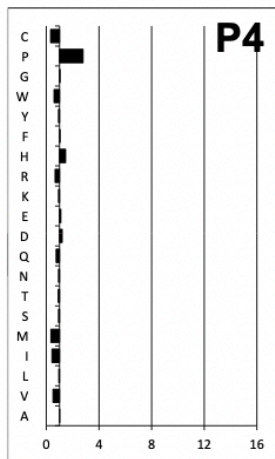
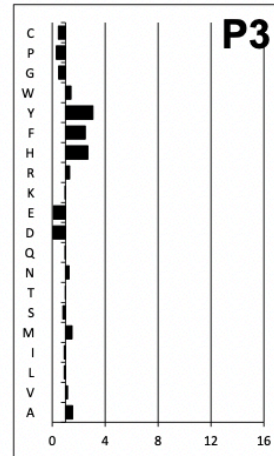
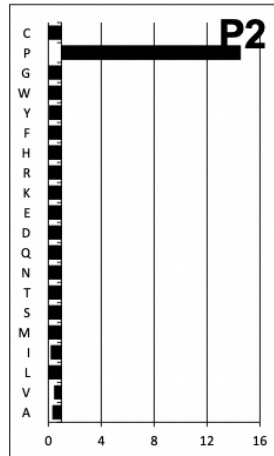
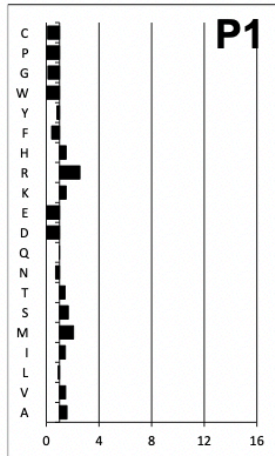


- **HLA-B\*55:01**

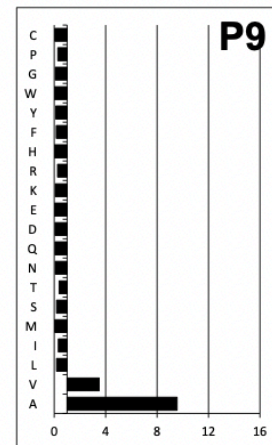
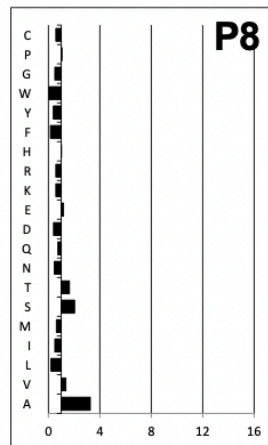
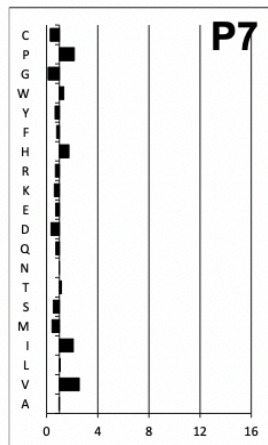
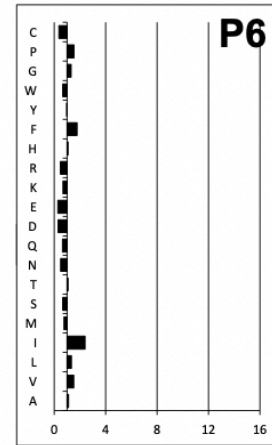
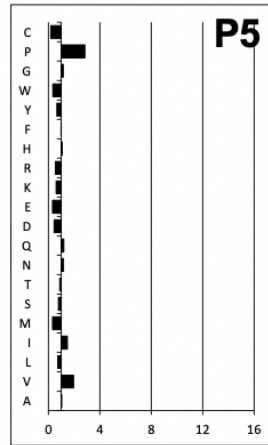
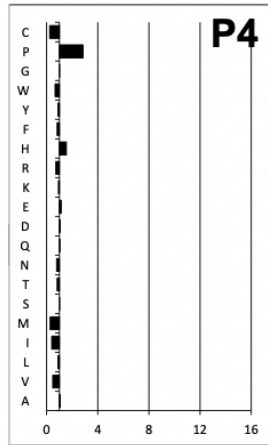
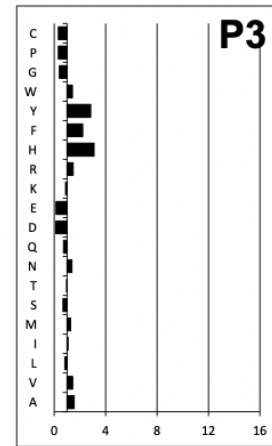
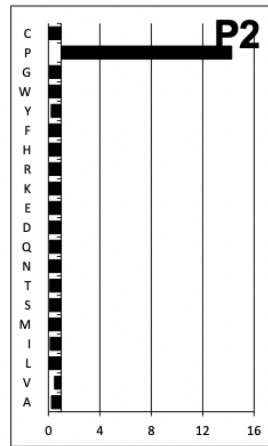
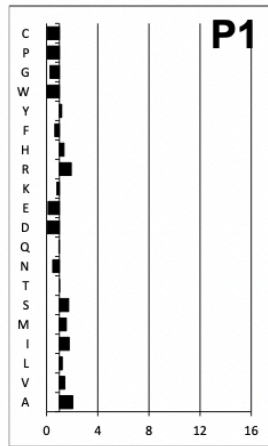




- HLA-B\*55:02



- HLA-B\*56:01



**Anexo 8:** Residuos de anclaje de los transfectantes de HLA-B\*51:01, -B\*35:01 y -B\*07:02 calculados con el *software* PRBAM

- **HLA-B\*51:01**



- HLA-B\*51:01 T45E





- HLA-B\*51:01 F67Y



- HLA-B\*51:01 T45E + F67Y



- HLA-B\*35:01



- HLA-B\*35:01 T45E



- HLA-B\*35:01 F67Y



- HLA-B\*35:01 T45E + F67Y



- HLA-B\*07:02



- HLA-B\*07:02 E45T





- HLA-B\*07:02 Y67F



- HLA-B\*07:02 E45T + Y67F



**Anexo 9:**

***Analysis of the different subpeptidomes presented by the HLA class I molecules of the B7 supertype.*** Cell Immunol . 2023 May ; 387: 104707.

doi:10.1016/j.cellimm.2023.104707.





Published in final edited form as:

*Cell Immunol.* 2023 May ; 387: 104707. doi:10.1016/j.cellimm.2023.104707.

## Analysis of the different subpeptidomes presented by the HLA class I molecules of the B7 supertype

Adrián Tirado-Herranz<sup>1,2,†</sup>, Pablo Guasp<sup>3,4,†</sup>, Alba Pastor-Moreno<sup>1,2</sup>, María Area-Navarro<sup>1,2</sup>, Iñaki Alvarez<sup>1,2,\*</sup>

<sup>1</sup>Immunology Unit, Department of Cell Biology, Physiology and Immunology, Autonomous University of Barcelona, 08193 Bellaterra, Spain

<sup>2</sup>Institute of Biotechnology and Biomedicine, Autonomous University of Barcelona, 08193 Bellaterra, Spain

<sup>3</sup>Immuno-Oncology Service, Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

<sup>4</sup>Hepatopancreatobiliary Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA

### Abstract

MHC-I molecules of the HLA-B7 supertype preferentially bind peptides with proline at position 2. HLA-B\*51:01 and B\*51:08 present two predominant subpeptidomes, one with Pro2 and hydrophobic residues at P1, and another with Ala2 and Asp enriched at position 1. Here, we present a meta-analysis of the peptidomes presented by molecules of the B7 supertype to investigate the presence of subpeptidomes across different allotypes. Several allotypes presented subpeptidomes differing in the presence of Pro or another residue at P2. The Ala2 subpeptidomes preferred Asp1 except in HLA-B\*54:01, where ligands with Ala2 contained Glu1. Sequence alignment and the analysis of crystal structures allowed us to propose positions 45 and 67 of the MHC heavy chain as relevant for the presence of subpeptidomes. Deciphering the principles behind the presence of subpeptidomes could improve our understanding of antigen presentation in other MHC-I molecules.

### Keywords

HLA; antigen presentation; immunopeptidome; B7 supertype; peptide

\* **Correspondence:** Iñaki Álvarez PhD. Phone: 34-93-581 3237. Fax: 34-93-581 2011. Inaki.Alvarez@uab.cat, Postal address: Cellular Immunology laboratory, Institut de Biotecnologia i Biomedicina, Campus de Bellaterra, 08193, Bellaterra, Barcelona, Spain.

<sup>†</sup>These authors have contributed equally to this work.

#### Author Contributions

AT and PG: conceptualization, methodology, data curation. AP and MA; methodology, data curation. IA: conceptualization, writing original draft preparation, supervision and validation.

#### Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## 1 Introduction

Human leukocytes antigens class I molecules (HLA-I) bind peptides derived from cytosolic or nuclear proteins and present them to cytotoxic T lymphocytes (CTL), which are able to discriminate between “self” and “non-self” peptides and subsequently kill infected or transformed cells. HLA-I are heterodimers composed of a heavy and variable  $\alpha$  chain encoded in the major histocompatibility complex (MHC) region and a light and constant chain, beta-2-microglobulin ( $\beta_2m$ ), encoded outside the MHC. The distal extracellular  $\alpha_1$  and  $\alpha_2$  domains of the heavy chain form the peptide binding site, composed by two  $\alpha$  helices and a  $\beta$  sheet of 8 anti-parallel  $\beta$  strands, forming six different pockets that accommodate the side chains of some residues of the peptide [1]. In HLA-I, the most restrictive pockets are those that bind the side chains of the second and last peptide positions (P2 and P $\Omega$ , respectively) [2]. Three *loci* encode the heavy chain of the classical HLA-I molecules: *HLA-A*, *HLA-B* and *HLA-C*. *HLA* genes are the most polymorphic genes in the genome, and most of their polymorphisms are located in the peptide binding site, making different HLA molecules bind different peptidomes with specific anchor motifs [2]. The interaction of HLA-I with specific residues of the peptide restricts the number of potential ligands. Thus, the correct identification of the anchor motifs from the different HLA-I allotypes is crucial to properly predict which peptides from a given protein have the potential to bind to each HLA molecule and be presented to CTLs.

HLA-B molecules have been grouped in 6 different supertypes based on the anchor motifs of their peptide ligands, called B7, B8, B27, B44, B58 and B62 supertypes [3-7]. HLA-I molecules belonging to the same supertype bind partially overlapping peptidomes. The B7 supertype is composed of HLA-B molecules that bind Pro at the P2 peptide position and aromatic, aliphatic and hydrophobic residues at their P $\Omega$  [3].

HLA-B\*51:01 and -B\*51:08, two members of the B7 supertype, bind peptides with Pro2 or Ala2 and hydrophobic residues at P $\Omega$  [8-11]. Previous work reported that the peptide repertoires bound to these allotypes are composed by two different subpeptidomes: one with Pro2 and a preference for hydrophobic residues at P1, and one containing Ala2 with preference for Asp1 [12], and subsequently confirmed [13]. More recently, it was described the presence of subpeptidomes in other HLA-I peptide pools [13, 14], showing that it is not unique to the -B\*51 subtypes.

The presence or absence of subpeptidomes and the different features that can show these subpeptidomes among close related HLA-I molecules may explain the presentation or not of putative relevant T cell epitopes from pathogens or self-antigens. In addition, the knowledge of the structural features explaining the presence of subpeptidomes may be useful to predict which allotypes are able to bind specific peptides.

Here, we present a meta-analysis of the peptidomes of 13 HLA-I molecules belonging to the -B7 supertype: we analyzed the presence or absence of subpeptidomes and the effect of the amino acid variation at the anchor position 2 on the residue at position 1. The presence of subpeptidomes varied among the HLA-B7 allotypes. Sequence alignment of the HLA-I molecules strongly suggests that residues Thr45 and Phe67 are important for the

presence of subpeptidomes in the B7 supertype molecules. Positions 45 and 67 are part of the Pocket B [15] and have been described to alter peptide presentation [16]. Finally, the analysis of multiple crystal structures of -B\*07:01, -B\*35:01 and -B\*51:01 showed that the orientation of residue Asn63 varied depending on the presence of Glu45-Tyr67 (in -B\*07:01) or Thr45-Phe67 (in -B\*35:01 and -B\*51:01).

## 2 Materials and methods

### HLA-B peptidomes

Peptidomes from Immuno Epitope Database and Analysis Resource (IEDB) were extracted at 02-15-2022. Duplicated peptides in any HLA-B peptidome were eliminated and considered only once. Peptides with or without post-translational modifications were treated as unique non-modified peptides. Only peptidomes from HLA-B molecules with more than 500 peptides were included in this study.

#### 2.1 Identification of the HLA-I anchor motifs

To determine the HLA-I anchor motifs, peptidomes were analyzed with the Peptide Repertoire-Based Anchor Motif (PRBAM) program [17]. In order to simplify the output, only nonamers were analyzed to define the anchor motifs of these allotypes.

#### 2.2 MHC binding affinities of HLA-I ligands.

The theoretical affinity of the peptides for HLA-I molecules was calculated using the NetMHC4.0 1.1 Server (<http://www.cbs.dtu.dk/services/NetMHC/>) [18, 19].

#### 2.3 Statistical analysis.

Differences in residue frequencies were analyzed by chi-square test with Bonferroni correction as previously described [12].

Differences in MHC binding affinities were analyzed by the Mann-Whitney U test. Different P values were considered.

#### 2.4 Analysis of crystal structures

Crystal structures of different HLA molecules were previously reported and deposited on the Protein Data Bank (PDB, [rcsb.org](https://www.rcsb.org/)) [20]. To perform the analysis, three crystals structures of HLA-B\*51:01: 1E27, 1E28, 4MJI [21, 22], twelve of HLA-B\*07:02: 3VCL [23], 5EO, 5EO1 [24], 5WMN, 5WMO, 5WMP [25], 6AT5 [26], 6UJ7, 6UJ8 [27], 7LG0, 7LGD, 7LGT [28] and eighteen of HLA-B\*35:01: 1A1N [29], 1A9B, 1A9E [30], 1XH3 [31], 1ZSD [32], 2AXG [33], 2CIK [34], 2FYY, 2FZ3 [35], 2H6P [36], 3LKN, 3LKO, 3LKP, 3LKQ, 3LKR [37], 4PR5, 4PRA, 4PRN [38] were chosen. Sequences were aligned and analyzed using the PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

### 3 Results

#### 3.1 Features of the B7 supertype peptidomes

We obtained the peptide repertoires of different HLA-B molecules belonging to the B7 supertype from IEDB [39]. Peptidomes with more than 500 peptides ranging from 8 to 13 residues long were considered. Following this criterium, we selected 13 peptidomes, derived from HLA-B\*07:02, -B\*07:04, -B\*35:01, -B\*35:03, -B\*35:07, -B\*42:01, -B\*51:01, -B\*51:08, -B\*53:01, -B\*54:01, -B\*55:01, -B\*55:02 and -B\*56:01. Figure 1 shows the sequence alignment of these proteins. The number of reported peptides from each molecule ranged from 624 to 12033 peptides (Table 1). As expected, nonamers comprised the majority of each peptidome (from 49.5% to 77.9%), with the percentage of octamers and peptides 10-13 residues long varying among different allotypes. HLA-B\*51:01 and -B\*51:08 contained the highest proportion of octamers (27.4% and 13.6%, respectively) and the lowest number of peptides of 10 or more residues. In contrast, HLA-B\*07:02, -B\*07:04, -B\*55:02 and -B\*56:01 bound more than 20% of 10-mer and more than 10% of 11-mer peptides. The average number of residues in the analyzed peptidomes was similar (ranging from 9.3 to 9.7), except for the two B\*51 alleles, with an average number of 9.0 residues.

The B7 supertype is characterized by the presence of Pro at P2 and aromatic, aliphatic and hydrophobic residues at PΩ [3]. We analyzed the degree of overlap between the peptidomes of the different members of the B7 supertype. The 13 peptide pools were composed of a total of 27039 peptides, of which 6134 (22.7%) were shared by at least two allotypes and 2445 (9.0%) by at least three HLA-I molecules. The degree of overlap was analyzed by pairs of HLA-B molecules (Table 2). In general, peptidomes from different subtypes of the same allele group (HLA-B\*07:02 with -B\*07:04; -B\*35:01 with -B\*35:03 and -B\*35:07; -B\*51:01 with -B\*51:08; and -B\*55:01 with -B\*55:02) shared a higher number of peptides than with other molecules. Nevertheless, in this dataset, the highest overlap was found between HLA-B\*55:02 and -B\*56:01, sharing 1207 peptides (77.9% and 57.1% of the -B\*55:02 and -B\*56:01 peptide repertoires, respectively) (Table 2).

#### 3.2 Anchor motifs of the B7 supertype HLA molecules

All 13 molecules bind preferentially peptides with Pro2, while peptides with Ala2 are tolerated to different degree, ranging from 2.6% of the peptidome in B\*55:02 and B\*56:01 to 54.3% in B\*51:08 (Table 3). To further characterize the anchor motifs of these molecules, their peptidomes were analyzed with the PRBAM software. This software identifies the anchor motifs of HLA-I and HLA-DR molecules based on the amino acid abundance at each position. The program returns a deviation from the mean in the proteome (DMP) value for each amino acid in each peptide position. The resulting anchor motifs for nonamers are shown in Supplementary Figure 1. Seven allotypes presented peptidomes with favored Ala2 (the detected abundance of Ala at P2 is higher than its abundance in the human proteome): HLA-B\*35:01, -B\*35:03, -B\*35:07, -B\*51:01, -B\*51:08, -B\*53:01 and -B\*54:01 (Supplementary Figure 1).



### 3.3 HLA-B\*54:01 presents, as B\*51:01 and B\*51:08, two subpeptidomes with favored acid residues in P1, but with Glu1 preferred to Asp1 in the subpeptidome containing Ala2

The peptide repertoire of HLA-B\*51:01 was previously described based on the identification of 1620 eluted peptides, and two different subpeptidomes were identified based on the presence of Pro or Ala at P2. Peptides with Pro2 contained mostly aliphatic and aromatic residues at P1, while peptides with Ala2 showed preferentially Asp at P1 [12]. We expanded the analysis with 5157 ligands from the IEDB obtaining similar results, with peptides with Pro2 presenting aliphatic or aromatic residues at P1 and peptides with Ala1 containing mainly Asp1 (Figure 2).

The HLA-B\*51:08 peptidome, containing 624 unique peptides, was analyzed previously [40]. As shown in Figure 2, the HLA-B\*51:08 peptide repertoire is also composed by two subpeptidomes with the same features than those of HLA-B\*51:01.

The IEDB peptidome of HLA-B\*54:01 consisted of 1441 peptides, from which 1038 (72.0%) contained Pro2, 145 (10.1%) Ala2 and 258 (17.9%) other amino acids at P2. As in HLA-B\*51:01 and -B\*51:08, the Ala2 subpeptidome from HLA-B\*54:01 showed an increase in acidic residues at P1. Interestingly, peptides contained more Glu1 than Asp1. Peptides with Pro2 presented aliphatic or aromatic residues in P1.

Thus, the existence of two subpeptidomes, one with Pro2 associated to a prevalence of hydrophobic residues at P1 and a second one with Ala2 and acidic residues at P1, is present in at least other molecule of the B7 supertype different to the B\*51 subtypes.

### 3.4 HLA-B35 subtypes and HLA-B\*53:01 show two subpeptidomes with polar and acidic residues favored at P1 in the subpeptidome containing Ala2

The HLA-B\*35:01, -B\*35:03, -B\*35:07 and -B\*53:01 peptidomes also presented, in addition to Pro2, Ala increased at P2 (Table 3 and Supplementary Figure 1). The analysis of these peptidomes indicated that the subpeptidomes containing Pro2 presented a preference for aliphatic or aromatic residues at P1, while the subpeptidome with Ala2 preferred for polar and acidic residues. In this case, polar residues (Ser, Thr, Asn and Gln) were preferred at P1 to Asp and Glu (Figure 3). In conclusion, HLA-B\*35:01, -B\*35:03, -B\*35:07 and -B\*53:01 present two subpeptidomes: one with Pro2 and preference for hydrophobic residues at P1, and another one presenting Ala2 and polar and acidic residues favored at P1.

### 3.5 Other HLA-B molecules of the B7 supertype do not present apparent subpeptidomes

The peptide repertoires of the other HLA-B molecules of the B7 supertype obtained from the IEDB (HLA-B\*07:02, -B\*07:04, -B\*42:01, -B\*55:01, -B\*55:02 and -B\*56:01) only had Pro favored at P2. Although peptides with Ala2 were not favored in these peptidomes, some ligands of these molecules contained Ala at P2 (with an abundance ranging from 2.6% to 3.9% of the total peptides). We analyzed the residue usage at P1 of these peptides and compared it to that from Pro2 peptides. As shown in Figure 4, we were not able to detect significant differences at P1, contrary to what was observed in other members of the B7 supertype, although Pro and Gly were increased in the peptides containing Ala2 and Ser and Arg tended to be increased in the peptides with Pro2.

### 3.6 Peptidomes containing different amino acids than Pro or Ala at P2 show similar features to the Ala2 subpeptidome

A considerable number of peptides contained different residues than Pro or Ala at P2 (Table 3). To evaluate whether these peptides are similar to the Pro2 or Ala2 subpeptidomes, or instead form a third subpeptidome on their own, we compared this peptide pool with those containing Pro2 or Ala2. In general, this peptide pool was similar to those with the subpeptidome of Ala2 in reference to the P1 composition (Supplementary Figure 2 and Supplementary Figure 3).

### 3.7 The presence of Thr45 and Phe67 instead of Glu45 and Tyr67 is associated with the presence of subpeptidomes in the B7 supertype molecules

The  $\alpha$  chain sequence alignment of the members of the B7 supertype indicated that molecules that do not present subpeptidomes contain Glu45, while the alleles that present subpeptidomes contain Thr45, except for HLA-B\*54:01 which has Gly45 (Table 4). In addition, Phe67 is present in the alleles with subpeptidomes, except for HLA-B\*54:01 which, like the allotypes without apparent subpeptidomes, contains Tyr67.

The analysis of multiple crystal structures of HLA-B\*07:02 showed that Glu45 directly interacts with Tyr67 and Asn63 (a non-polymorphic residue present in all the allotypes analyzed) (Figure 5A and Supplementary figure 4). This interaction does not exist when Thr and Phe are present at positions 45 and 67, respectively, as is the case of -B\*35:01 and -B\*51:01 (Figure 5B and 5C and Supplementary figure 4). The interaction of Glu45 with Asn63 changes the orientation of the latter compared to the orientation of this residue in the allotypes that contain different subpeptidomes (Figure 5 and Supplementary figure 4).

### 3.8 Residue 116 defines different residue preferences at the peptide C-terminal position

We can identify three main groups based on the peptide C-terminal residue usage: 1) HLA-B\*35:01, -B\*35:07 and -B\*53:01 present mainly aromatic residues at P $\Omega$ ; 2) HLA-B\*54:01, -B\*55:01, -B\*55:02 and -B\*56:01 bind peptides with Ala or Val at P $\Omega$ ; 3) the rest of the analyzed allotypes bind peptides with Leu or Met at P $\Omega$  (HLA-B\*07:02, -B\*07:04, -B\*35:03 and -B\*42:01) or Leu or Val (HLA-B\*51:01 and -B\*51:08) at P $\Omega$  (Supplementary Figure 1). The analysis of the HLA residues that interact with the F pocket revealed that the presence of different residues at position 116 is associated with the preference for specific amino acids at P $\Omega$ . Thus, the allotypes that bind peptides with aromatic residues at the C-terminus have Ser116, the ones that bind Ala or Val contain Leu116, and the molecules binding peptides with Leu, Met or Val at C-terminus with as the most abundant amino acids have Tyr116, except for HLA-B\*35:03 that has Phe116 (both aromatic residues) (Table 4).

### 3.9 Theoretical binding affinity of ligands to HLA molecules of the B7 supertype

The theoretical binding affinity of the peptides bound to different members of the B7 supertype was evaluated with NetMHC 4.1, a powerful tool that calculates the affinity of peptides to HLA-I molecules. We analyzed all the molecules included in this analysis that were available to study in NetMHC4.1. These molecules were: HLA-B\*07:02, -B\*35:01, -B\*35:03, -B\*51:01, -B\*53:01 and -B\*54:01. The predicted affinity of the peptides with Pro2 to their corresponding HLA-B molecule was consistently higher than those with Ala2

(Figure 6). Interestingly, in the case of -B\*07:02, the subpeptidome containing Ala2 showed a lower theoretical affinity than the total immunopeptidome. In the other HLA-I molecules, which present preferences by acidic or polar residues in P1, the theoretical affinity was similar or higher than that of the total peptidome (Figure 6).

## 4 Discussion

It has been described that some HLA alleles show multiple specificities for their ligands [13]. In this study, we analyzed the peptidomes presented by 13 allotypes belonging to the HLA-B7 supertype. We found that, although all the B7 molecules can bind Pro and Ala (in addition to other less abundant residues) at P2, several members of the B7 supertype, but not all, present two subpeptidomes in which these amino acids are favored at P2, a feature previously described in HLA-B\*51:01 and -B\*51:08 [12, 40]. We analyzed the effect of the presence of these amino acids at P2 on the residues favored at P1 and analyzed the structural features involved in this effect. Although the presence of subpeptidomes have been previously reported in different HLA-I molecules [13, 14], to our knowledge, this is the first study that systematically characterizes the presence of subpeptidomes across different members of the B7 supertype and provides a structural explanation for the existence of such subpeptidomes.

The peptidomes presented by the analyzed members of the B7 supertype showed different degrees of overlap, being in general higher among the subtypes of the same allele group. However, the most similar peptidomes were those from HLA-B\*55:02 and -B\*56:01. Both alleles contain Val152, a position described to affect peptide selection at the internal positions [40, 41], which could explain the lower overlap with the peptidome presented by HLA-B\*55:01, which carries Glu152. HLA-B\*55 and -B\*56, together with HLA-B\*54, compose the broad serotype B22. Nevertheless, the peptidome presented by HLA-B\*54:01 shows a relatively low overlap with those of HLA-B\*55:01, -B\*55:02 and -B\*56:01. This could be explained by the polymorphisms Glu45Gly and Ile52Val in HLA-B\*54:01 compared to -B\*55 and -B\*56.

The analyzed peptides come both from transfected cells expressing only one HLA-B allotype and from cells expressing different HLA-A. In the last case the HLA-B molecule assignment was made based on algorithms. We cannot discard that the use of algorithms could introduce a bias in the assignment. Similarly, although the peptides were obtained by similar methodology, the large differences in the size of the peptidomes may have introduced some bias in the anchor motif, as it is expected that the peptide repertoires with a lower number of peptides are composed of the more abundant and with a higher affinity for the HLA-B molecule.

The presence of subpeptidomes among the members of the B7 supertype correlates with their ability to bind other residues than Pro2 at P2. By performing sequence alignments we were able to identify polymorphisms Glu45Thr and Tyr67Phe as important residues for the presence of subpeptidomes, except for HLA-B\*54:01, that carry Gly45 and Tyr67. The crystal structures of HLA-B\*07:02 show that Glu45 interacts with both Tyr67 and Asn63, a residue conserved among all the analyzed alleles. On the other hand, Thr45, present

in the alleles that bind subpeptidomes, is not able to interact with Asn63 and Phe67, as shown by the structures of HLA-B\*51:01 and -B\*35:01. We propose that the interactions mediated by Glu45 reorient Asn63 and, overall, change the structure of the Pocket B, making it more rigid, limiting its ability to bind residues different than Pro. Additionally, peptide bonds involving Pro are rigid and can exist in the cis-configuration due to the cyclized nature of this amino acid. This characteristic could modify the interactions of the P1 residue with the Pocket A when followed by Pro, altering the HLA specificities when binding ligands with Pro at P2 compared to other residues at that position. As described elsewhere [13], the crystal structures of HLA-B\*51:01 bound to a ligand with Pro2 or with Ala2 showed that, when Ala is present at P2, the P1 residue can interact with Arg62 favoring the presence of acidic and polar residues at that position, while in ligands with Pro2 this interaction is impeded, favoring the presence of non-polar residues at P1. Previously, it was described that Endoplasmic Reticulum Aminopeptidase 1 (ERAP1) is responsible for the differential residue usage at P1 between the two subpeptidomes of HLA-B\*51:01 [12]. ERAP1 trims peptides in the endoplasmic reticulum (ER) to the correct length so that they can be loaded onto the HLA molecules [42, 43]. This process occurs with different efficiency depending on the nature of the P1 residue; moreover, ERAP1 is unable to trim peptide bonds involving Pro [44]. Thus, peptides that carry an ERAP1-susceptible residue at P1 followed by an amino acid other than Pro would be degraded by the enzyme. On the other hand peptides with ERAP1-resistant P1 residues, such as Asp, a residue predominant at P1 in the Ala2 subpeptidome, will remain in the ER long enough to bind the HLA molecule and be presented on the cell surface. Since ERAP1 cannot degrade the X-Pro bond, even peptides with ERAP1-susceptible residues at P1, such as aliphatic residues, will not be degraded by the enzyme when followed by Pro and will bind HLA-B\*51:01. The peptidomes presented by cells carrying ERAP1 variants with different activity [40] and by cells lacking ERAP1 [45] support this mechanism but do not fully explain the different usage at P1 between subpeptidomes, suggesting that the HLA molecules themselves exert different binding selections among the subpeptidomes. The effects of ERAP1 activity on the peptidomes presented by the different members of the B7 supertype has not yet been characterized, but its role on shaping the different subpeptidomes of these alleles could be similar to the one described for HLA-B\*51:01.

The predictive affinity to HLA of the Pro2 subpeptidome was systemically higher than that of the Ala2 subpeptidome across the different members of the B7 supertype. These data are in agreement with previous data on HLA-B51 allotypes [12]. In the case of HLA-B\*07:02, the theoretical affinity of the total peptidome presented a higher affinity than the Ala2-containing subpeptidome. The other 5 allotypes analyzed, which presented subpeptidomes with acid or polar residues favored in P1 did not present this lower affinity in the Ala2-containing subpeptidome. Thus, although -B\*07:02 can bind peptides with Ala2, the absence of polar or acidic residues at P1 may produce a lower affinity. We cannot discard that the presence of two subpeptidomes, one of them with more peptides than the other, may result in the underestimation of the affinity of peptides belonging to the less abundant subpeptidome.

In conclusion, this work indicates that the study of the peptidomes presented by an HLA-I molecules is complex, and that a thorough analysis is required, at least among the members

of the B7 supertype, to consider differences between subpeptidomes. A detailed analysis of these subpeptidomes is necessary to understand the rules underlying HLA ligand selection and, consequently, to correctly predict which peptides have the potential to bind to and be presented by the different HLA alleles.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding

This work was supported by Grant RTI2018-097414-B-I00 from the Spanish Ministry of Economy and Competitiveness to I.A.

## Abbreviations:

|              |                                               |
|--------------|-----------------------------------------------|
| <b>HLA</b>   | human leukocyte antigens                      |
| <b>CTL</b>   | cytotoxic T lymphocytes                       |
| <b>MHC</b>   | major histocompatibility complex              |
| <b>β2m</b>   | Beta-2 microglobulin                          |
| <b>IEDB</b>  | Immuno Epitope Database and Analysis Resource |
| <b>PRBAM</b> | Peptide Repertoire-Based Anchor Motif         |
| <b>PDB</b>   | Protein Data Bank                             |

## 10 References

- [1]. Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC, Structure of the human class I histocompatibility antigen, HLA-A2, Nature, 329 (1987) 506–512. [PubMed: 3309677]
- [2]. Falk K, Rotzschke O, Stevanovic S, Jung G, Rammensee HG, Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules, Nature, 351 (1991) 290–296. [PubMed: 1709722]
- [3]. Sidney J, Peters B, Frahm N, Brander C, Sette A, HLA class I supertypes: a revised and updated classification, BMC Immunol, 9 (2008) 1. [PubMed: 18211710]
- [4]. Hertz T, Yanover C, Identifying HLA supertypes by learning distance functions, Bioinformatics, 23 (2007) e148–155. [PubMed: 17237084]
- [5]. Kanguane P, Sakharkar MK, Rajaseger G, Bolisetty S, Sivasekari B, Zhao B, Ravichandran M, Shapshak P, Subbiah S, A framework to sub-type HLA supertypes, Front Biosci, 10 (2005) 879–886. [PubMed: 15569626]
- [6]. Lund O, Nielsen M, Kesmir C, Petersen AG, Lundegaard C, Worning P, Sylvester-Hvid C, Lamberth K, Roder G, Justesen S, Buus S, Brunak S, Definition of supertypes for HLA molecules using clustering of specificity matrices, Immunogenetics, 55 (2004) 797–810. [PubMed: 14963618]
- [7]. Wang M, Claesson MH, Classification of human leukocyte antigen (HLA) supertypes, Methods Mol Biol, 1184 (2014) 309–317. [PubMed: 25048132]

- [8]. Kikuchi A, Sakaguchi T, Miwa K, Takamiya Y, Rammensee HG, Kaneko Y, Takiguchi M, Binding of nonamer peptides to three HLA-B51 molecules which differ by a single amino acid substitution in the A-pocket, *Immunogenetics*, 43 (1996) 268–276. [PubMed: 9110930]
- [9]. Sakaguchi T, Ibe M, Miwa K, Kaneko Y, Yokota S, Tanaka K, Takiguchi M, Binding of 8-mer to 11-mer peptides carrying the anchor residues to slow assembling HLA class I molecules (HLA-B\*5101), *Immunogenetics*, 45 (1997) 259–265. [PubMed: 9002446]
- [10]. Sakaguchi T, Ibe M, Miwa K, Yokota S, Tanaka K, Schonbach C, Takiguchi M, Predominant role of N-terminal residue of nonamer peptides in their binding to HLA-B\* 5101 molecules, *Immunogenetics*, 46 (1997) 245–248. [PubMed: 9211752]
- [11]. Falk K, Rotzschke O, Takiguchi M, Gnau V, Stevanovic S, Jung G, Rammensee HG, Peptide motifs of HLA-B51, -B52 and -B78 molecules, and implications for Behcet's disease, *International immunology*, 7 (1995) 223–228. [PubMed: 7734418]
- [12]. Guasp P, Alvarez-Navarro C, Gomez-Molina P, Martin-Esteban A, Marcilla M, Barnea E, Admon A, Lopez de Castro JA, The Peptidome of Behcet's Disease-Associated HLA-B\*51:01 Includes Two Subpeptidomes Differentially Shaped by Endoplasmic Reticulum Aminopeptidase 1, *Arthritis Rheumatol*, 68 (2016) 505–515. [PubMed: 26360328]
- [13]. Gfeller D, Guillaume P, Michaux J, Pak HS, Daniel RT, Racle J, Coukos G, Bassani-Sternberg M, The Length Distribution and Multiple Specificity of Naturally Presented HLA-I Ligands, *Journal of immunology*, 201 (2018) 3705–3716.
- [14]. Sarkizova S, Klaeger S, Le PM, Li LW, Oliveira G, Keshishian H, Hartigan CR, Zhang W, Braun DA, Ligon KL, Bachireddy P, Zervantonakis IK, Rosenbluth JM, Ouspenskaia T, Law T, Justesen S, Stevens J, Lane WJ, Eisenhaure T, Lan Zhang G, Clauser KR, Hacohen N, Carr SA, Wu CJ, Keskin DB, A large peptidome dataset improves HLA class I epitope prediction across most of the human population, *Nature biotechnology*, 38 (2020) 199–209.
- [15]. Saper MA, Bjorkman PJ, Wiley DC, Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution, *Journal of molecular biology*, 219 (1991) 277–319. [PubMed: 2038058]
- [16]. Carreno BM, Winter CC, Taurog JD, Hansen TH, Biddison WE, Residues in pockets B and F of HLA-B27 are critical in the presentation of an influenza A virus nucleoprotein peptide and influence the stability of peptide - MHC complexes, *International immunology*, 5 (1993) 353–360. [PubMed: 8494822]
- [17]. Mestre-Ferrer A, Scholz E, Humet-Alsius J, Alvarez I, PRBAM: a new tool to analyze the MHC class I and HLA-DR anchor motifs, *Immunology*, 156 (2019) 187–198. [PubMed: 30408168]
- [18]. Andreatta M, Nielsen M, Gapped sequence alignment using artificial neural networks: application to the MHC class I system, *Bioinformatics*, 32 (2016) 511–517. [PubMed: 26515819]
- [19]. Nielsen M, Lundegaard C, Worning P, Lauemoller SL, Lamberth K, Buus S, Brunak S, Lund O, Reliable prediction of T-cell epitopes using neural networks with novel sequence representations, *Protein Sci*, 12 (2003) 1007–1017. [PubMed: 12717023]
- [20]. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE, The Protein Data Bank, *Nucleic acids research*, 28 (2000) 235–242. [PubMed: 10592235]
- [21]. Maenaka K, Maenaka T, Tomiyama H, Takiguchi M, Stuart DI, Jones EY, Nonstandard peptide binding revealed by crystal structures of HLA-B\*5101 complexed with HIV immunodominant epitopes, *Journal of immunology*, 165 (2000) 3260–3267.
- [22]. Motozono C, Kuse N, Sun X, Rizkallah PJ, Fuller A, Oka S, Cole DK, Sewell AK, Takiguchi M, Molecular basis of a dominant T cell response to an HIV reverse transcriptase 8-mer epitope presented by the protective allele HLA-B\*51:01, *Journal of immunology*, 192 (2014) 3428–3434.
- [23]. Brennan RM, Petersen J, Neller MA, Miles JJ, Burrows JM, Smith C, McCluskey J, Khanna R, Rossjohn J, Burrows SR, The impact of a large and frequent deletion in the human TCR beta locus on antiviral immunity, *Journal of immunology*, 188 (2012) 2742–2748.
- [24]. Du VY, Bansal A, Carlson J, Salazar-Gonzalez JF, Salazar MG, Ladell K, Gras S, Josephs TM, Heath SL, Price DA, Rossjohn J, Hunter E, Goepfert PA, HIV-1-Specific CD8 T Cells Exhibit Limited Cross-Reactivity during Acute Infection, *Journal of immunology*, 196 (2016) 3276–3286.



- [25]. Rowntree LC, Nguyen THO, Halim H, Purcell AW, Rossjohn J, Gras S, Kotsimbos TC, Mifsud NA, Inability To Detect Cross-Reactive Memory T Cells Challenges the Frequency of Heterologous Immunity among Common Viruses, *Journal of immunology*, 200 (2018) 3993–4003.
- [26]. Chan KF, Gully BS, Gras S, Beringer DX, Kjer-Nielsen L, Cebon J, McCluskey J, Chen W, Rossjohn J, Divergent T-cell receptor recognition modes of a HLA-I restricted extended tumour-associated peptide, *Nature communications*, 9 (2018) 1026.
- [27]. Hwang MS, Miller MS, Thirawatananond P, Douglass J, Wright KM, Hsiue EH, Mog BJ, Aytenfisu TY, Murphy MB, Aitana Azurmendi P, Skora AD, Pearlman AH, Paul S, DiNapoli SR, Konig MF, Bettgowda C, Pardoll DM, Papadopoulos N, Kinzler KW, Vogelstein B, Zhou S, Gabelli SB, Structural engineering of chimeric antigen receptors targeting HLA-restricted neoantigens, *Nature communications*, 12 (2021) 5271.
- [28]. Lineburg KE, Grant EJ, Swaminathan S, Chatzileontiadou DSM, Szeto C, Sloane H, Panikkar A, Raju J, Crooks P, Rehan S, Nguyen AT, Lekieffre L, Neller MA, Tong ZWM, Jayasinghe D, Chew KY, Lobos CA, Halim H, Burrows JM, Riboldi-Tunncliffe A, Chen W, D'Orsogna L, Khanna R, Short KR, Smith C, Gras S, CD8(+) T cells specific for an immunodominant SARS-CoV-2 nucleocapsid epitope cross-react with selective seasonal coronaviruses, *Immunity*, 54 (2021) 1055–1065 e1055. [PubMed: 33945786]
- [29]. Smith KJ, Reid SW, Stuart DI, McMichael AJ, Jones EY, Bell JI, An altered position of the alpha 2 helix of MHC class I is revealed by the crystal structure of HLA-B\*3501, *Immunity*, 4 (1996) 203–213. [PubMed: 8624811]
- [30]. Menssen R, Orth P, Ziegler A, Saenger W, Decamer-like conformation of a nona-peptide bound to HLA-B\*3501 due to non-standard positioning of the C terminus, *Journal of molecular biology*, 285 (1999) 645–653. [PubMed: 9878435]
- [31]. Probst-Kepper M, Hecht HJ, Herrmann H, Janke V, Ocklenburg F, Klempnauer J, van den Eynde BJ, Weiss S, Conformational restraints and flexibility of 14-meric peptides in complex with HLA-B\*3501, *Journal of immunology*, 173 (2004) 5610–5616.
- [32]. Miles JJ, Elhassen D, Borg NA, Silins SL, Tynan FE, Burrows JM, Purcell AW, Kjer-Nielsen L, Rossjohn J, Burrows SR, McCluskey J, CTL recognition of a bulged viral peptide involves biased TCR selection, *Journal of immunology*, 175 (2005) 3826–3834.
- [33]. Tynan FE, Elhassen D, Purcell AW, Burrows JM, Borg NA, Miles JJ, Williamson NA, Green KJ, Tellam J, Kjer-Nielsen L, McCluskey J, Rossjohn J, Burrows SR, The immunogenicity of a viral cytotoxic T cell epitope is controlled by its MHC-bound conformation, *The Journal of experimental medicine*, 202 (2005) 1249–1260. [PubMed: 16275762]
- [34]. Hourigan CS, Harkiolaki M, Peterson NA, Bell JI, Jones EY, O'Callaghan CA, The structure of the human allo-ligand HLA-B\*3501 in complex with a cytochrome p450 peptide: steric hindrance influences TCR allo-recognition, *European journal of immunology*, 36 (2006) 3288–3293. [PubMed: 17109469]
- [35]. Miles JJ, Borg NA, Brennan RM, Tynan FE, Kjer-Nielsen L, Silins SL, Bell MJ, Burrows JM, McCluskey J, Rossjohn J, Burrows SR, TCR alpha genes direct MHC restriction in the potent human T cell response to a class I-bound viral epitope, *Journal of immunology*, 177 (2006) 6804–6814.
- [36]. Archbold JK, Macdonald WA, Miles JJ, Brennan RM, Kjer-Nielsen L, McCluskey J, Burrows SR, Rossjohn J, Alloreactivity between disparate cognate and allogeneic pMHC-I complexes is the result of highly focused, peptide-dependent structural mimicry, *The Journal of biological chemistry*, 281 (2006) 34324–34332. [PubMed: 16963442]
- [37]. Gras S, Kedzierski L, Valkenburg SA, Laurie K, Liu YC, Denholm JT, Richards MJ, Rimmelzwaan GF, Kelso A, Doherty PC, Turner SJ, Rossjohn J, Kedzierska K, Cross-reactive CD8+ T-cell immunity between the pandemic H1N1-2009 and H1N1-1918 influenza A viruses, *Proceedings of the National Academy of Sciences of the United States of America*, 107 (2010) 12599–12604. [PubMed: 20616031]
- [38]. Liu YC, Chen Z, Neller MA, Miles JJ, Purcell AW, McCluskey J, Burrows SR, Rossjohn J, Gras S, A molecular basis for the interplay between T cells, viral mutants, and human leukocyte antigen micropolymorphism, *The Journal of biological chemistry*, 289 (2014) 16688–16698. [PubMed: 24759101]

- [39]. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A, Peters B, The Immune Epitope Database (IEDB): 2018 update, *Nucleic acids research*, 47 (2019) D339–D343. [PubMed: 30357391]
- [40]. Guasp P, Barnea E, Gonzalez-Escribano MF, Jimenez-Reinoso A, Regueiro JR, Admon A, Lopez de Castro JA, The Behcet's disease-associated variant of the aminopeptidase ERAP1 shapes a low-affinity HLA-B\*51 peptidome by differential subpeptidome processing, *The Journal of biological chemistry*, 292 (2017) 9680–9689. [PubMed: 28446606]
- [41]. Marino SR, Lin S, Maiers M, Haagenson M, Spellman S, Klein JP, Binkowski TA, Lee SJ, van Besien K, Identification by random forest method of HLA class I amino acid substitutions associated with lower survival at day 100 in unrelated donor hematopoietic cell transplantation, *Bone marrow transplantation*, 47 (2012) 217–226. [PubMed: 21441965]
- [42]. Saric T, Chang SC, Hattori A, York IA, Markant S, Rock KL, Tsujimoto M, Goldberg AL, An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides, *Nat Immunol*, 3 (2002) 1169–1176. [PubMed: 12436109]
- [43]. Serwold T, Gonzalez F, Kim J, Jacob R, Shastri N, ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum, *Nature*, 419 (2002) 480–483. [PubMed: 12368856]
- [44]. Hearn A, York IA, Rock KL, The specificity of trimming of MHC class I-presented peptides in the endoplasmic reticulum, *Journal of immunology*, 183 (2009) 5526–5536.
- [45]. Guasp P, Lorente E, Martin-Esteban A, Barnea E, Romania P, Fruci D, Kuiper JW, Admon A, Lopez de Castro JA, Redundancy and Complementarity between ERAP1 and ERAP2 Revealed by their Effects on the Behcet's Disease-associated HLA-B\*51 Peptidome, *Mol Cell Proteomics*, 18 (2019) 1491–1510. [PubMed: 31092671]



**Highlights:**

- - Molecules of the HLA B7 supertype differ in the prevalence of Ala2 in addition of the canonical Pro2.
- - We found three different patterns in the B7 supertype molecules: i) non-presence of subpeptidomes, ii) subpeptidomes with Pro2 and Ala2 with acid residues favored in P1, and iii) subpeptidomes with Pro2 and Ala2 with polar residues favored in P1.
- - B\*54:01 presents a subpeptidome where Ala2 peptides favor Glu1 over Asp1 (favored in B\*51:01 and B\*51:08).
- - Polymorphisms in position 45 and 67 of the  $\alpha$ -chain are important for the presence of subpeptidomes in the B7 supertype allotypes.

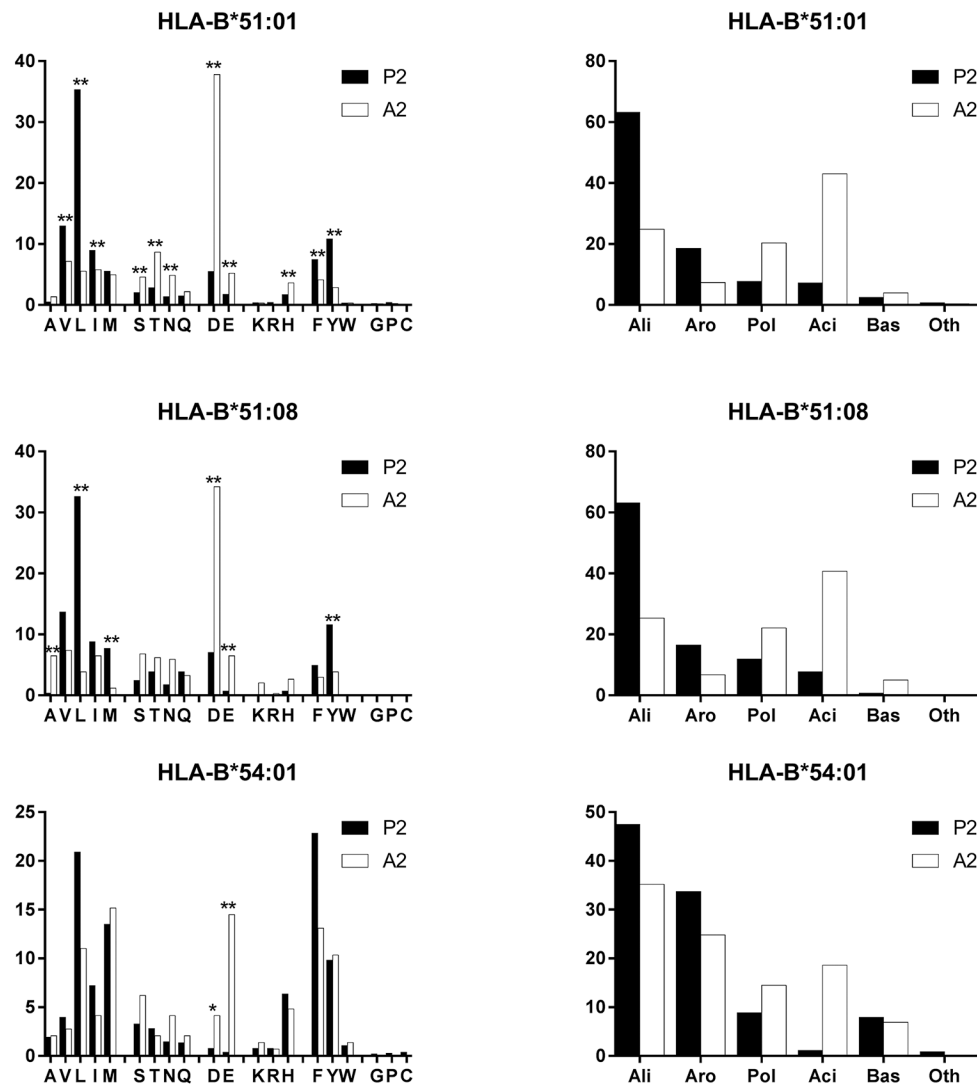
| AA Pos. | 10         | 20         | 30          | 40         | 50          | 60         | 70         | 80         | 90         | 100            |
|---------|------------|------------|-------------|------------|-------------|------------|------------|------------|------------|----------------|
| B*07:02 | GSHSMRYFYT | SVSRPGRGEP | RFISVGIVDD  | TQFVRFDSDA | ASPREEPRAP  | WIEQEGPEYW | DRNTQIYKAQ | AQTDRESLRN | LRGYYNQSEA | GSHTLQSMYG     |
| B*07:04 | -----      | -----      | -----       | -----      | -----       | -----      | -----      | -----      | -----      | -----          |
| B*42:01 | -----      | -----      | -----       | -----      | -----       | -----      | -----      | -----      | -----      | -----          |
| B*55:01 | -----      | AM-----    | -----A----- | -----      | -----       | -----      | -----      | -----      | -----      | -----W-T-----  |
| B*55:02 | -----      | AM-----    | -----A----- | -----      | -----       | -----      | -----      | -----      | -----      | -----W-T-----  |
| B*56:01 | -----      | AM-----    | -----A----- | -----      | -----       | -----      | -----      | -----      | -----      | -----W-T-----  |
| B*35:01 | -----      | AM-----    | -----A----- | -----      | -----T----- | -----      | -----F-TN  | T--Y-----  | -----      | -----II-R----- |
| B*35:03 | -----      | AM-----    | -----A----- | -----      | -----T----- | -----      | -----F-TN  | T--Y-----  | -----      | -----II-R----- |
| B*35:07 | -----      | AM--V----- | -----A----- | -----      | -----T----- | -----      | -----F-TN  | T--Y-----  | -----      | -----II-R----- |
| B*53:01 | -----      | AM-----    | -----A----- | -----      | -----T----- | -----      | -----F-TN  | T--Y--N--I | ALR-----   | -----II-R----- |
| B*51:01 | -----      | AM-----    | -----A----- | -----      | -----T----- | -----      | -----F-TN  | T--Y--N--I | ALR-----   | -----W-T-----  |
| B*51:08 | -----      | AM-----    | -----A----- | -----      | -----T----- | -----      | -----F-TN  | T--Y--N--I | ALR-----   | -----W-T-----  |
| B*54:01 | -----      | AM-----    | -----A----- | -----      | -----G----- | -V-----    | -----      | -----      | -----      | -----W-T-----  |

| AA Pos. | 110       | 120         | 130         | 140        | 150        | 160         | 170        | 180             | 190        | 200         |
|---------|-----------|-------------|-------------|------------|------------|-------------|------------|-----------------|------------|-------------|
| B*07:02 | CDVGPDGRL | RGHDQYAYDG  | KDYIALNEDL  | RSWTAADTAA | QITQRKWEAA | REAEQRRAYL  | EGECVWELRR | YLENGKDKLE      | RADPPKTHVT | HPISDHEAT   |
| B*07:04 | -----     | -----       | -----       | -----      | -----      | -----D----- | -----      | -----           | -----      | -----       |
| B*42:01 | -----     | --N-----    | -----       | -----      | -----      | -V--D-----  | -T-----    | -----T-----     | -----      | -----       |
| B*55:01 | --L-----  | --N-L-----  | -----S----- | -----      | -----      | -V--L-----  | -T-----    | -----ET-Q-----  | -----      | -----       |
| B*55:02 | --L-----  | --N-L-----  | -----S----- | -----      | -----      | -V--L-----  | -T-----    | -----ET-Q-----  | -----      | -----       |
| B*56:01 | --L-----  | --N-L-----  | -----S----- | -----      | -----      | -V--L-----  | -L-----    | -----ET-Q-----  | -----      | -----       |
| B*35:01 | --L-----  | -----S----- | -----S----- | -----      | -----      | -V--L-----  | -L-----    | -----ET-Q-----  | -----      | -----V----- |
| B*35:03 | --L-----  | -----F----- | -----S----- | -----      | -----      | -V--L-----  | -L-----    | -----ET-Q-----  | -----      | -----V----- |
| B*35:07 | --L-----  | -----S----- | -----S----- | -----      | -----      | -V--L-----  | -L-----    | -----ET-Q-----  | -----      | -----V----- |
| B*53:01 | --L-----  | -----S----- | -----S----- | -----      | -----      | -V--L-----  | -L-----    | -----ET-Q-----  | -----      | -----V----- |
| B*51:01 | -----     | --N-----    | -----S----- | -----      | -----      | -----L----- | -L-----    | H-----ET-Q----- | -----      | -----V----- |
| B*51:08 | -----     | --N-----    | -----S----- | -----      | -----      | -V--D-----  | -L-----    | H-----ET-Q----- | -----      | -----V----- |
| B*54:01 | --L-----  | --N-L-----  | -----S----- | -----      | -----      | -V--L-----  | -T-----    | -----ET-Q-----  | -----      | -----       |

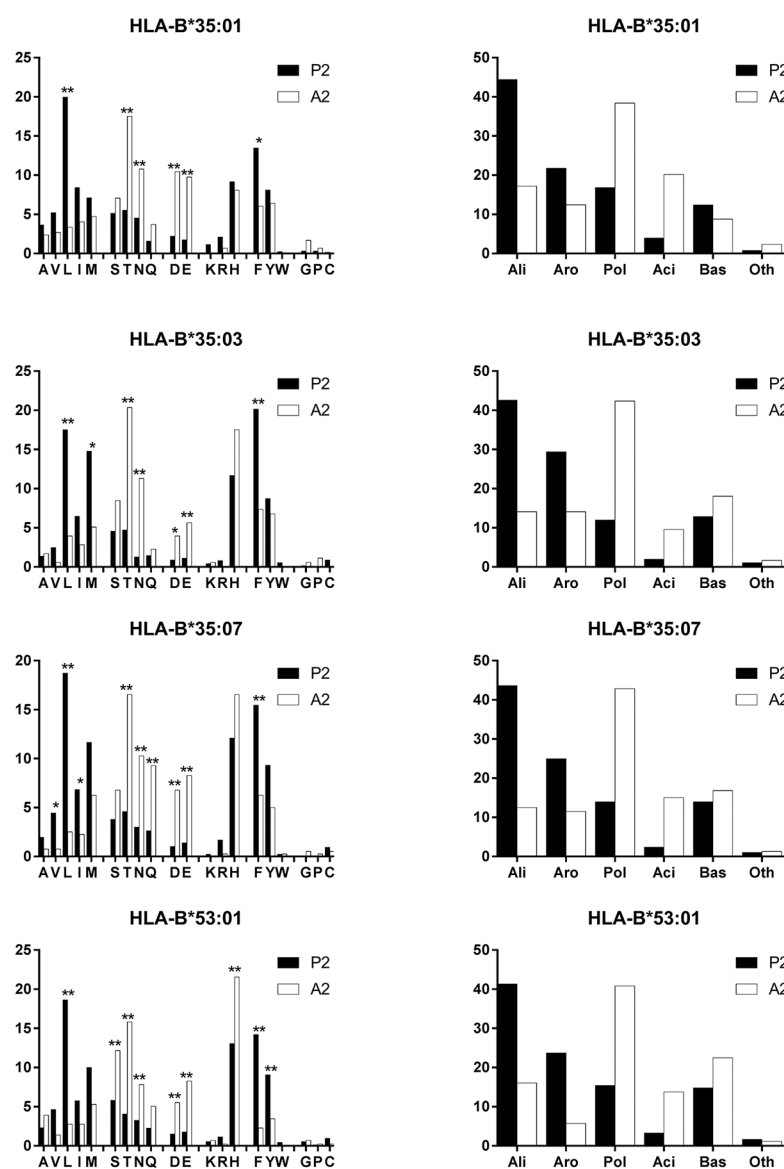
**Figure 1.**

Sequence alignment of the 13 B7 supertype allotypes. Primary sequences of the HLA-I molecules analyzed in this work were aligned and variable residues with regard to HLA-B\*07:02 are indicated.

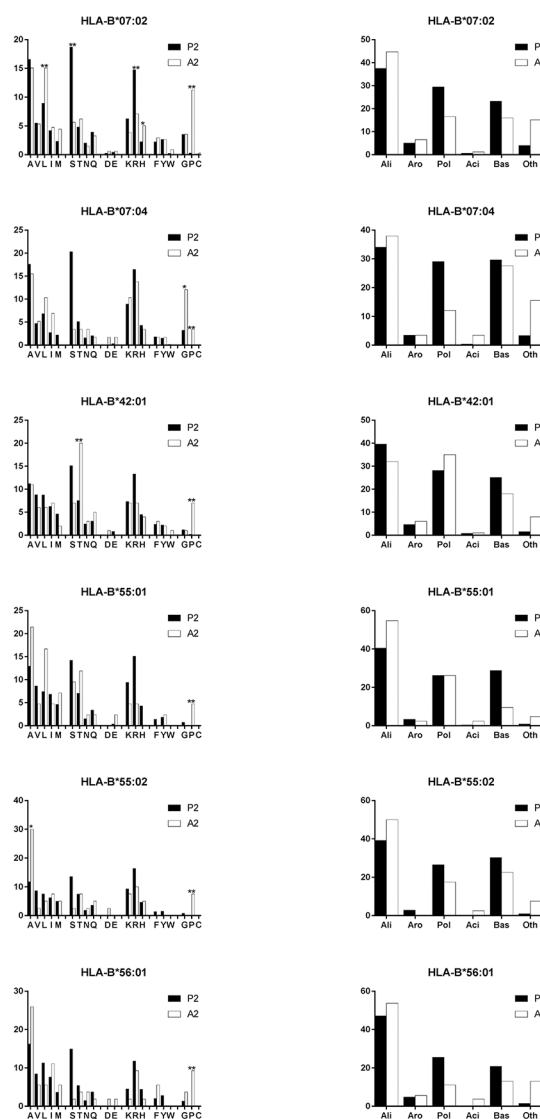


**Figure 2. Differential residue usage at the N-terminal position (P1) between HLA-B\*51:01, -B\*51:08 and -B\*54:01 ligands with Pro at position 2 (P2) or Ala at position 2 (A2).**

The single-letter code for amino acids is used. Left panels, Position 1 residue frequencies among peptides with Pro2 (P2, black) or Ala2 (A2, white). Residues are sorted according to their chemical features. Statistically significant differences, as determined by the chi-square test with Bonferroni correction were indicated as \* < 0.05 and \*\* < 0.01. Right panels, frequency of residues at P1 grouped by their chemical nature.

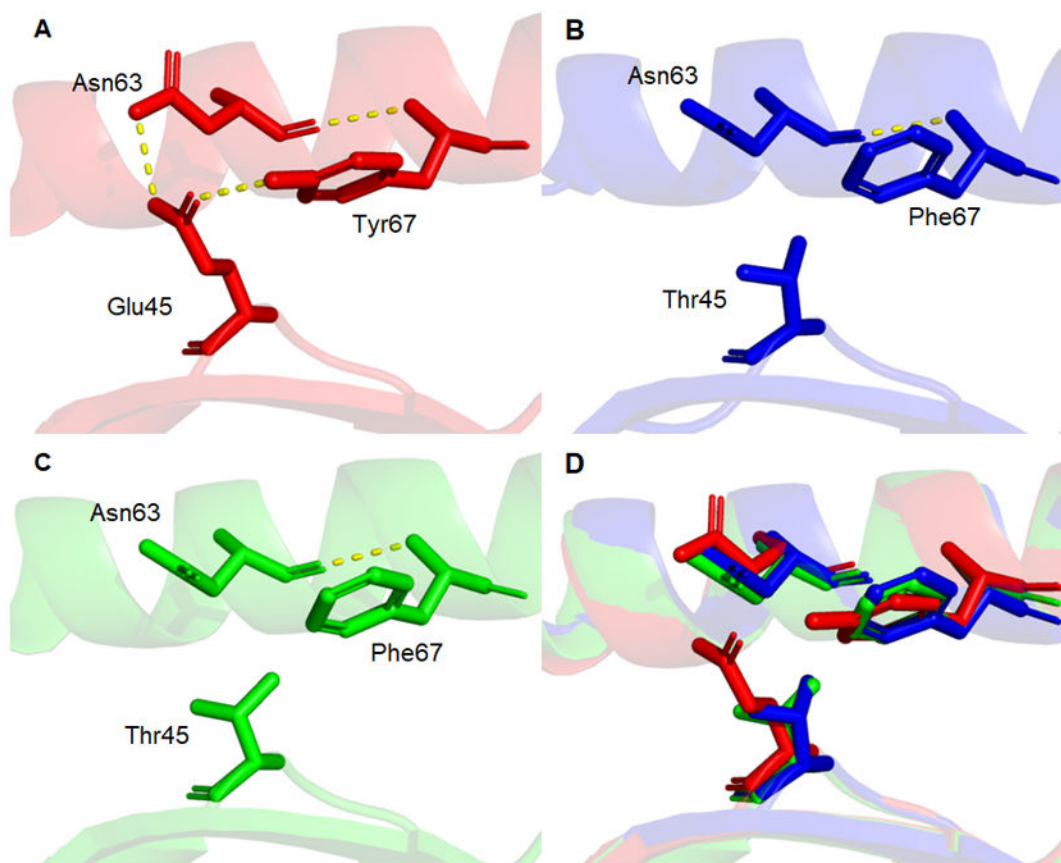


**Figure 3. Differential residue usage at the N-terminal position (P1) between HLA-B\*35:01, -B\*35:03, -B\*35:07 and -B\*53:01 ligands with Pro at position 2 (P2) or Ala at position 2 (A2).** The single-letter code for amino acids is used. Left panels, Position 1 residue frequencies among peptides with Pro2 (P2, black) or Ala2 (A2, white). Residues are sorted according to their chemical features. Statistically significant differences, as determined by the chi-square test with Bonferroni correction were indicated as \*  $< 0.05$  and \*\*  $< 0.01$ . Right panels, frequency of residues at P1 grouped by their chemical nature.



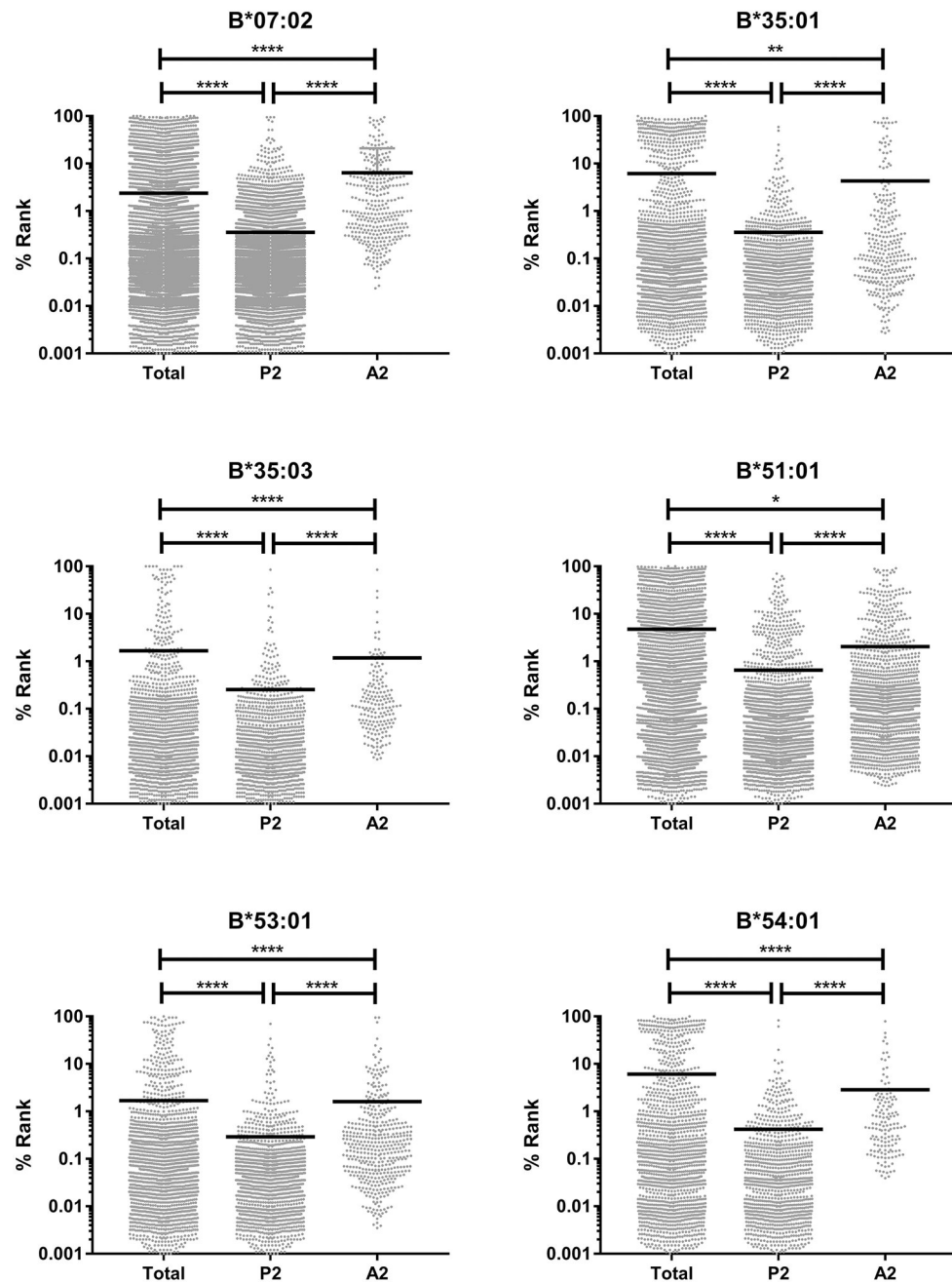
**Figure 4. Differential residue usage at the N-terminal position (P1) between HLA-B\*07:02, -B\*07:04, -B\*42:01, -B\*55:01, -B\*55:02 and -B\*56:01 ligands with Pro at position 2 (P2) or Ala at position 2 (A2).**

The single-letter code for amino acids is used. Left panels, Position 1 residue frequencies among peptides with Pro2 (P2, black) or Ala2 (A2, white). Residues are sorted according to their chemical features. Statistically significant differences, as determined by the chi-square test with Bonferroni correction were indicated as \*  $< 0.05$  and \*\*  $< 0.01$ . Right panels, frequency of residues at P1 grouped by their chemical nature.



**Figure 5: A-C. Representation of residues 45, 63 and 67 in HLA-B\*07:02, -B\*35:01 and -B\*51:01.**

HLA in HLA-B\*07:02 (A), -B\*35:01 (B) and -B\*51:01 (C). D. Superposition of HLA-B\*07:02 (red), -B\*35:01 (blue) and B\*51:01 (green). Residues 46, 63 and 67 are represented as sticks. Polar contacts are represented as dotted yellow lines. PDB entries: 1E27, 1A1N and 6UJ7.



**Figure 6. Theoretical binding affinity to HLA of the subpeptidomes of the members of the B7 supertype.**

The %Rank of the peptides belonging to the whole peptidome (total), the Pro2 subpeptidome (P2) and the Ala2 subpeptidome (A2) to their respective HLA molecule is indicated.

Symbols represent individual peptides; horizontal lines represent the median. \* =  $0.01 < P < 0.1$ ; \*\* =  $0.001 < P < 0.01$ ; \*\*\* =  $0.001 < P < 0.01$ ; \*\*\*\* =  $P < 0.0001$  by Mann-Whitney U test.

**Table 1.**

General features of the B7 supertype peptidomes.

| HLA-B* | Total | Peptides by number of residues (%) |       |       |       |      |      | AAs average |
|--------|-------|------------------------------------|-------|-------|-------|------|------|-------------|
|        |       | 8                                  | 9     | 10    | 11    | 12   | 13   |             |
| 07:02  | 12033 | 7.83                               | 49.49 | 21.94 | 12.29 | 5.68 | 2.78 | 9.67        |
| 07:04  | 1480  | 5.74                               | 61.22 | 20.00 | 10.34 | 1.89 | 0.81 | 9.44        |
| 35:01  | 1954  | 5.27                               | 57.37 | 17.14 | 11.67 | 5.53 | 3.02 | 9.64        |
| 35:03  | 1527  | 4.78                               | 76.36 | 10.35 | 6.09  | 1.83 | 0.59 | 9.26        |
| 35:07  | 1933  | 3.31                               | 68.39 | 14.49 | 9.67  | 2.69 | 1.45 | 9.44        |
| 42:01  | 3428  | 5.86                               | 59.42 | 17.53 | 12.95 | 2.92 | 1.31 | 9.52        |
| 51:01  | 5157  | 27.38                              | 55.85 | 8.45  | 4.07  | 2.39 | 1.86 | 9.04        |
| 51:08  | 624   | 13.62                              | 77.88 | 4.17  | 3.04  | 1.28 | 0.00 | 9.00        |
| 53:01  | 2204  | 4.54                               | 63.43 | 17.56 | 9.62  | 3.58 | 1.27 | 9.48        |
| 54:01  | 1441  | 11.24                              | 52.26 | 16.52 | 12.28 | 4.44 | 3.26 | 9.56        |
| 55:01  | 1548  | 11.30                              | 55.81 | 19.77 | 9.04  | 2.39 | 1.68 | 9.40        |
| 55:02  | 1549  | 3.42                               | 54.36 | 22.66 | 13.43 | 4.13 | 2.00 | 9.66        |
| 56:01  | 2113  | 4.92                               | 51.49 | 24.33 | 13.54 | 3.74 | 1.99 | 9.66        |



Table 2.

Immunopeptidome overlap.

| Peptide Number |              | 12033       | 1480        | 1954        | 1527        | 1933        | 3428       | 5157       | 624       | 2204      | 1441        | 1548        | 1549        | 2113 |
|----------------|--------------|-------------|-------------|-------------|-------------|-------------|------------|------------|-----------|-----------|-------------|-------------|-------------|------|
|                | HLA allotype |             |             |             |             |             |            |            |           |           |             |             |             |      |
| 12033          | B*07:02      |             | 822         | 228         | 314         | 199         | 1693       | 203        | 33        | 181       | 82          | 286         | 230         | 271  |
| 1480           | B*07:04      | 55.5 \ 6.8  |             | 25          | 59          | 41          | 878        | 32         | 9         | 36        | 19          | 120         | 80          | 85   |
| 1954           | B*35:01      | 11.7 \ 1.9  | 1.3 \ 1.7   |             | 246         | 865         | 115        | 153        | 3         | 452       | 68          | 9           | 22          | 25   |
| 1527           | B*35:03      | 20.6 \ 2.6  | 3.9 \ 4.0   | 16.1 \ 12.6 |             | 371         | 237        | 93         | 12        | 322       | 45          | 34          | 39          | 55   |
| 1933           | B*35:07      | 10.3 \ 1.7  | 2.1 \ 2.8   | 44.7 \ 44.3 | 19.2 \ 24.3 |             | 168        | 60         | 3         | 613       | 28          | 21          | 31          | 42   |
| 3428           | B*42:01      | 49.4 \ 14.1 | 25.6 \ 59.3 | 3.4 \ 5.9   | 6.9 \ 15.5  | 4.9 \ 8.7   |            | 129        | 49        | 149       | 66          | 238         | 191         | 209  |
| 5157           | B*51:01      | 3.9 \ 1.7   | 0.6 \ 2.2   | 3.0 \ 7.8   | 1.8 \ 6.1   | 1.2 \ 3.1   | 2.5 \ 3.8  |            | 337       | 178       | 215         | 100         | 99          | 130  |
| 624            | B*51:08      | 5.3 \ 0.3   | 1.4 \ 0.6   | 0.5 \ 0.2   | 1.9 \ 0.8   | 0.5 \ 0.2   | 7.9 \ 1.4  | 54.0 \ 6.5 |           | 34        | 30          | 23          | 25          | 25   |
| 2204           | B*53:01      | 8.2 \ 1.5   | 1.6 \ 2.4   | 20.5 \ 23.1 | 14.6 \ 21.1 | 27.8 \ 31.7 | 6.8 \ 4.3  | 8.1 \ 3.5  | 1.5 \ 5.4 |           | 37          | 32          | 36          | 41   |
| 1441           | B*54:01      | 5.7 \ 0.7   | 1.3 \ 1.3   | 4.7 \ 3.5   | 3.1 \ 2.9   | 1.9 \ 1.4   | 4.6 \ 1.9  | 14.9 \ 4.2 | 2.1 \ 4.8 | 2.6 \ 1.7 |             | 226         | 329         | 425  |
| 1548           | B*55:01      | 18.5 \ 2.4  | 7.8 \ 8.1   | 0.6 \ 0.5   | 2.2 \ 2.2   | 1.4 \ 1.1   | 15.4 \ 6.9 | 6.5 \ 1.9  | 1.5 \ 3.7 | 2.1 \ 1.5 | 14.6 \ 15.7 |             | 710         | 777  |
| 1549           | B*55:02      | 14.8 \ 1.9  | 5.2 \ 5.4   | 1.4 \ 1.1   | 2.5 \ 2.6   | 2.0 \ 1.6   | 12.3 \ 5.6 | 6.4 \ 1.9  | 1.6 \ 4.0 | 2.3 \ 1.6 | 21.2 \ 22.8 | 45.8 \ 45.9 |             | 1207 |
| 2113           | B*56:01      | 12.8 \ 2.3  | 4.0 \ 5.7   | 1.2 \ 1.3   | 2.6 \ 3.6   | 2.0 \ 2.2   | 9.9 \ 6.1  | 6.2 \ 2.5  | 1.2 \ 4.0 | 1.9 \ 1.9 | 20.1 \ 29.5 | 36.8 \ 50.2 | 57.1 \ 77.9 |      |

Note: Above the diagonal, absolute numbers of shared peptides are indicated. Under the diagonal, percentage that the shared peptide represents on the total number of peptides in the peptidomes (the number before backslash correspond to the allotype indicated in the left, and the number after backslash corresponds to the allotype indicated above).

Note: Light grey shows that shared peptides represent more than 20% of the total peptidome of at least one of the pair. Dark shows that shared peptides represent more than 20% of the total peptidome of at least one of the pair.

**Table 3.**

Amino acid presence at the peptide position 2 among the peptidomes of the B7 supertype molecules.

|                    | Total | P2           | A2           | Others       |
|--------------------|-------|--------------|--------------|--------------|
| <b>HLA-B*07:02</b> | 12033 | 9956 (82.7%) | 338 (2.8%)   | 1739 (14.5%) |
| <b>HLA-B*07:04</b> | 1480  | 1233 (83.3%) | 58 (3.9%)    | 189 (12.8%)  |
| <b>HLA-B*35:01</b> | 1954  | 1322 (67.7%) | 297 (15.2%)  | 335 (17.1%)  |
| <b>HLA-B*35:03</b> | 1527  | 1251 (81.9%) | 177 (11.6%)  | 99 (6.5%)    |
| <b>HLA-B*35:07</b> | 1933  | 1373 (71.0%) | 399 (20.6%)  | 161 (8.3%)   |
| <b>HLA-B*42:01</b> | 3428  | 2987 (87.1%) | 100 (2.9%)   | 341 (9.9%)   |
| <b>HLA-B*51:01</b> | 5157  | 2566 (49.8%) | 1552 (30.1%) | 1039 (20.1%) |
| <b>HLA-B*51:08</b> | 624   | 285 (45.7%)  | 339 (54.3%)  | 0 (0.0%)     |
| <b>HLA-B*53:01</b> | 2204  | 1600 (72.6%) | 436 (19.8%)  | 168 (7.6%)   |
| <b>HLA-B*54:01</b> | 1441  | 1038 (72.0%) | 145 (10.1%)  | 258 (17.9%)  |
| <b>HLA-B*55:01</b> | 1548  | 1378 (89.0%) | 42 (2.7%)    | 128 (8.3%)   |
| <b>HLA-B*55:02</b> | 1549  | 1407 (90.8%) | 40 (2.6%)    | 102 (6.6%)   |
| <b>HLA-B*56:01</b> | 2113  | 1873 (88.6%) | 54 (2.6%)    | 186 (8.8%)   |

**Table 4.**

Residues at positions 45, 67 and 116 in the HLA-I molecules analyzed.

| HLA MOLECULE   | POSITION |    |     | MAIN PEPTIDE<br>RESIDUE AT PΩ |
|----------------|----------|----|-----|-------------------------------|
|                | 45       | 67 | 116 |                               |
| <b>B*07:02</b> | E        | Y  | Y   | Leu or Met                    |
| <b>B*07:04</b> | -        | -  | -   | Leu or Met                    |
| <b>B*42:01</b> | -        | -  | -   | Leu or Met                    |
| <b>B*55:01</b> | -        | -  | L   | Ala or Val                    |
| <b>B*55:02</b> | -        | -  | L   | Ala or Val                    |
| <b>B*56:01</b> | -        | -  | L   | Ala or Val                    |
|                |          |    |     |                               |
| <b>B*35:01</b> | T        | F  | S   | Aromatic                      |
| <b>B*35:03</b> | T        | F  | F   | Leu or Met                    |
| <b>B*35:07</b> | T        | F  | S   | Aromatic                      |
| <b>B*53:01</b> | T        | F  | S   | Aromatic                      |
|                |          |    |     |                               |
| <b>B*51:01</b> | T        | F  | -   | Leu or Val                    |
| <b>B*51:08</b> | T        | F  | -   | Leu or Val                    |
| <b>B*54:01</b> | G        | -  | L   | Ala or Val                    |