The HLA-DP peptide repertoire from human respiratory syncytial virus is focused on major structural proteins with the exception of the viral polymerase

Elena Lorente, Eilon Barnea, Carmen Mir, Arie Admon, and Daniel López.

J Proteomics 2020 Jun 15;221:103759.

https://doi.org/10.1016/j.jprot.2020.103759

The HLA-DP peptide repertoire from human respiratory syncytial virus is focused on major structural proteins with the exception of the viral polymerase

Elena Lorente [1], Eilon Barnea [2], Carmen Mir [1], Arie Admon [2], and Daniel López [1,#]

[1] Unidad de Presentación y Regulación Inmunes, Instituto de Salud Carlos III, 28220 Majadahonda (Madrid), Spain

[2] Department of Biology, Technion-Israel Institute of Technology, 32000 Haifa, Israel

[#] Correspondence to: Dr. Daniel López. Unidad de Presentación y Regulación Inmunes. Centro Nacional de Microbiología. Instituto de Salud Carlos III. 28220 Majadahonda (Madrid), Spain. Tel: +34 91 822 37 08, FAX: +34 91 509 79 19, E-mail: ***dlopez@isciii.es***

Short title: Natural HLA-DP ligands from human respiratory syncytial virus

ACKNOWLEDGMENTS

This work was supported by the Spanish Ministry of Economy grants BIO2011-25636, SAF2014-58052, and "Acción Estratégica en Salud" to D.L., and by Israel Science Foundation, grant No. 1435/16 to A. A. The funding agencies had no role in the study design, data collection, analysis decision to publish, or preparation of the manuscript. The authors had no conflicting financial interests.

ABSTRACT

The recognition by specific T helper cells of viral antigenic peptides complexed with HLA class II molecules exposed on the surface of antigen presenting cells is the first step of the complex cascade of immunological events that generates the protective cellular and humoral immune responses. The HLA class II-restricted helper immune response is critical in the control and the clearance of human respiratory syncytial virus (HRSV) infection, a pathogen with severe health risk in pediatric, immunocompromised and elderly populations. In this study, a mass spectrometry analysis was used to identify HRSV ligands bound to HLA-DP class II molecules present on the surface of HRSV-infected cells. Among the thousands of cellular peptides bound to HLA class II proteins in the virus-infected cells, sixty-four naturally processed viral ligands, most of them included in complex nested set of peptides, were identified bound to HLA-DP molecules. These viral ligands arose from five of six major structural HRSV proteins: attachment, fusion, matrix, nucleoprotein, and phosphoprotein. In contrast, no HLA-DP ligands were identified from polymerase protein, the largest HRSV protein that includes half of the viral proteome. These findings have important implications for analysis of the helper immune response as for antiviral vaccine design.

SIGNIFICANCE

The existence of a supertype including five alleles that bind a peptide repertoire very similar make HLA-DP class II molecules an interesting target for the design of vaccines. Here, we analyze the HLA-DP-restricted peptide repertoire against the human respiratory syncytial virus, a pathogen that represents a high health risk in infected pediatric, immunocompromised and elderly populations. This repertoire is focused on major structural proteins with the exception of the viral polymerase.

INTRODUCTION

Newly synthesized HLA class II molecules are associated with the class II invariant chain in the Endoplasmic Reticulum from antigen presenting cells. These complexes are eventually transported to the Golgi apparatus and later on to specialized endosomal compartments, where the class II invariant chain is progressively proteolyzed until only the fragment known as the class II-associated invariant chain peptide (CLIP) remains bound in the HLA class II peptide-binding groove. This association prevents the binding of cellular or pathogen peptides from the endogenous pathways to HLA class II molecules. Later, these endosomal compartments fuse with late endosomes, which contain exogenous proteins and/or pathogen particles previously endocytosed by phagocytosis, endocytosis, or pinocytosis and processed by several resident cathepsins. In these resulting phagolysosomes, the release of CLIP exposes the antigen-binding site of the HLA class II molecules allowing the binding of processed peptides of different lengths (up to 30 residues), but with specific major anchor residues that can be deeply accommodated into specific pockets of the antigen recognition site of the HLA class II molecule. These HLA class II/peptide complexes are then transported to the cell membrane where they are exposed for T cell recognition massive. The interaction of the receptor of the $CD4^+$ T helper cells with pathogen peptides bound to HLA class II molecules from the antigen presenting cells triggers the activity of these T lymphocytes, initiating, regulating or suppressing the other components of the adaptive immune response. If the HLA class II-restricted helper response is defective or missing neither humoral nor cytotoxic immune components can be effectively activated, no antibodies are generated nor $CD8^+$ T lymphocytes can eliminate infected cells and thus, the pathogen evades the immune response with significant damage or even lethal results for the host.

Human respiratory syncytial virus (HRSV) [2] is a member of the *Pneumoviridae* family of the Mononegavirales order with a single-stranded, negative-sense RNA genome that codes for 11 proteins. This enveloped Orthopneumovirus causes repeated infections throughout life in people of all ages, although in healthy adults mild infections are generally reported. However, the health risk in infected pediatric, immunocompromised and elderly populations is much more severe. HRSV is the single most important cause of serious lower respiratory tract illnesses involving hospitalization for bronchiolitis and pneumonia in newborns and young children, with infection rates close to 70% in the first year of life [3]. At the age of 2-3 years nearly all children have been infected by HRSV [3]. Worldwide, more than 3.4 million hospital admissions each year are associated with HRSV disease, and the global mortality rate was estimated at more than a quarter of a million deaths in 2010, mainly in developing countries [4]. In addition, HRSV is a major nosocomial hazard in hospital or healthcare service units for patients of all ages [5], involving an important medical as well as economic impact.

The different immune mechanisms involved in HRSV disease, reinfection and protection are currently partially understood. It is well established that HRSV infection induces broad mucosal and systemic humoral and cellular responses. Studies analyzing $CD8^+$ and/or $CD4^+$ T-lymphocyte subpopulations showed that both effector MHC class I- and helper MHC class II-restricted cellular responses were absolutely necessaries in clearing HRSV infections [6]. Some HRSV epitopes that are restricted by different HLA class II molecules were identified using T cells from seropositive individuals [7],[8],[9],[10]. However, these experiments were performed only with overlapping synthetic peptides against the F protein or a short fragment of 21 residues from the G protein [7],[8],[9],[10]. In contrast, only one published study attempted to elucidate the possible array of viral ligands that are restricted by HLA class II molecules

in two HRSV patients [10]. These $CD4^+$ T cells were specific for two peptides from the matrix and attachment G proteins and restricted by two different HLA-DP alleles [10].

We are interested in the identification of viral ligands presented by several HLA class II molecules in HRSV-infected cells using mass spectrometry. In a previous study nineteen naturally processed HLA-DR ligands from HRSV-infected cells were identified [11]. In the current report we carried out a similar immunoproteomics analysis to identify the ligands bound to HLA-DP class II molecules, isolated from large amounts of uninfected or HRSV-infected cells without any methodological bias (e.g., selection of an individual protein, use of HLA consensus scoring algorithms, etc.). Several dozens of naturally processed, included in complex nested set of HLA-DP ligands from five different structural HRSV proteins, were identified. This analysis defines the nature of the HLA-DP class II-specific response against HRSV.

EXPERIMENTAL PROCEDURES

Synthetic peptides

Peptides were purchased from Peptide 2.0 (Chantilly, VA). The correct molecular mass and composition of the peptides at > 90% purity was established by quadrupole ion trap microHPLC.

HRSV infection of the human JY Epstein-Barr-transformed cell line

JY cells (IPD-IMGT/HLA; 10882) expressing HLA-DPA1*01:03/B1*02:01, and DPA1*01:03/B1*04:01 HLA class II complexes were incubated with the HRSV Long strain and assayed at different time points for the presence of HRSV antigens using flow cytometry, as previously described with either Epstein-Barr-transformed human B-cell lines [12] or other cell lines [13], to obtain a persistently infected JY-cell line [14].

HLA-bound peptide isolation

HLA-bound peptides were isolated from $4x10^{10}$ uninfected or HRSV-infected cells. Several previous analyses have shown that $4x10^{10}$ virus-infected human cells are enough to identify HLA-bound viral peptide pools [13],[15],[14],[11],[16]. Experiments were performed as unique experiments due to their laborious nature. The uninfected human cells were used as negative control to discriminate viral and cellular peptides (included in proteome databases as well as unknown peptides, whose parental proteins may not be included in current human databases) and to exclude erroneous assignments of viral peptides. The cells were lysed in 1% CHAPS (Sigma), 20 mM Tris/HCl buffer, and 150 mM NaCl, pH 7.5 in the presence of the cOmplete™, Mini Protease Inhibitor Cocktail (Merck KGaA, Darmstadt, Germany). After centrifugation, the supernatant was passed first through a control precolumn containing CNBr-activated Sepharose 4B (GE Healthcare, Buckinghamshire, UK) to remove non-specific peptides and proteins. The HLA-DP/peptide complexes were isolated via affinity

chromatography from the soluble cell extract fraction with the B7.21 mAb (specific for monomorphic pan-HLA-DP determinant) [17]. The HLA-bound peptides were eluted at 4ºC with 0.1% aqueous trifluoroacetic acid (TFA), dissociated from the HLA molecules, and concentrated by ultra-filtration with a Centricon 3 filter (Amicon, Beverly, MA), as previously described [14].

Electrospray-ion trap mass spectrometry analysis

Peptide mixtures recovered after the ultra-filtration step were desalted using Micro-Tip reversed-phase columns (C_{18}, 200 µl, Harvard Apparatus, Holliston, MA) [14]. Each C_{18} tip was activated with 80% acetonitrile in 0.1% TFA, washed with 0.1% TFA, and then loaded with the peptide mixture. The tip was then washed with an additional volume of 0.1% TFA, and the peptides were subsequently eluted with 80% acetonitrile in 0.1% TFA. Finally, the peptide samples were concentrated to approximately 20 µl using vacuum centrifugation [15],[14].

The HLA class II peptides recovered from immunoprecipitated HLA-DP-specific mAb, were analyzed by nanoLC-MS/MS using a Q-Exactive-Plus mass spectrometer that was fitted with an Ultimate 3000 RSLC nanocapillary UHPLC (Thermo Fisher Scientific, Waltham, MA), using the same parameters previously described [16].

The peptides were resolved on homemade 3.5 micron Reprosil C18-Aqua (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) capillary columns (75 micron ID and about 30 cm long [18] with a 5-28% acetonitrile linear gradient for 2 h in the presence of 0.1% formic acid at a flow rate of 0.15 µL/min. The dynamic exclusion was set to 20 sec and the automatic gain control value for the full MS was set to $3x10^6$. The selected masses were fragmented from the survey scan of mass-to-charge ratio (m/z) 300-1,800 AMU at resolution 70,000. The 10 most intense masses were selected for fragmentation by higher-energy collisional dissociation (HCD) from each full mass

spectrum. No fragmentation was performed for peptides with unassigned precursor ion charge states or charge states of five and above. MS/MS spectra were acquired with a resolution of 17,500 at m/z 200. The target value of the MS/MS was set to $1x10^5$ and the isolation window to 1.8 m/z. The maximum injection time was set to 100 ms and normalized collision energy to 25 eV. The peptide match option was set to Preferred.

Database searches

Raw mass spectrometry data were processed using Peaks 8.5 (Bioinformatics Solutions Inc., Waterloo, Canada) for peak-list generation from the nanoLC-MS/MS data. The peaks were identified using three different software programs (Peaks 8.5, Maxquant 1.4.1.2, and Proteome Discoverer 2.1) using the combination of human, and HRSV parts of the UniProt/Swiss-Prot database (Uniprot-proteome 27.4.1; 2018), which included 20267, and 11 proteins, respectively. As the human database utilized is 1200-fold greater than the corresponding HRSV database thus, the probability of erroneous viral assignments is statistically insignificant. The search was not limited by enzymatic specificity and both the peptide mass and the fragment ion tolerances were set at 5 and 20 ppm, respectively. This search was not limited by any methodological bias (e.g., individual protein selection or HLA consensus scoring algorithm use). The following variable modifications were also analyzed: acetylation of N-terminal residue, cysteinilation (C), oxidation (M, P, H, F, Y, and W), methylation (N- and C-terminus) and phosphorylation (Y, S, and T). The identified peptides were selected when their -10logP score was > 21 from Peaks, their score was > 80 from Maxquant, and Xcorr > 1.5 for singly, > 2.5 for doubly, > 3.2 for triply, and > 4.3 for quadruply charged peptides from Proteome Discoverer. In addition, the maximum false discovery rate (FDR) was set to 1%. No viral peptides were found in a search of the reversed databases. Also, no viral peptides were found in a search on the sample of uninfected cells. When the MS/MS spectra fitted more than one peptide, only the highest scoring

peptide was selected. Moreover, several synthetic peptides, which showed low scores with the three bioinformatics tools utilized, were used to confirm the assigned sequences. The mass spectrometry data have been deposited to the MassIVE repository (https://massive.ucsd.edu) with the data set identifier MSV000083591.

In silico binding prediction of HLA II ligands

The predicted binding of each peptide to HLA-DP class II molecules was calculated using the artificial neural network-based alignment method NetMHCIIpan (version 3.2) (available in http://www.cbs.dtu.dk/services/NetMHCIIpan/), and the MHC-II binding prediction from IEDB (available in http://tools.iedb.org/mhcii/).

RESULTS

The HLA-DP ligandome from uninfected and HRSV-infected human cells.

The HLA-DP-bound peptide pools were isolated from large numbers of either uninfected or HRSV-infected human JY cells. These peptide mixtures were subsequently separated by reversed-phase capillary HPLC and analyzed using mass spectrometry. Using Peaks bioinformatics tool, the MS analysis over a human proteome database resolved 2631 fragmentation spectra as peptidic sequences of 612 human cellular proteins bound to HLA-DP molecules (Supplementary Table 1). As was previously described in HLA class II antigen processing, multiple nested sets of peptides with the same core and various N- and C-terminally extended amino acids were identified (Supplementary Table 1).

The peptide size of HLA-DP ligands followed a normal distribution (Fig. 1A), with an average length of 16 residues (8-30), within the standard HLA class II size range. No length differences were found between uninfected and HRSV-infected samples. Figure 1B shows the predicted organelle location of these 612 cellular proteins based on the Gene Ontology database (http://www.geneontology.org). Extracellular (17%), membrane (20%), cytosol (20%), and nucleus (18%) were the more relevant compartments. Other organelles associated with the endocytic pathway as vesicles, Golgi and Endoplasmic Reticulum were the source of 7%, 4% and 5% of parental proteins, respectively. In total, the 53% of parental proteins were derived from the five compartments of the endo-lysosomal pathway, as usual in HLA class II molecules. No significant differences were found between proteins identified from uninfected and HRSV-infected samples.

Increased identification of ligands bound to HLA-DP class II molecules using both MaxQuant and Peaks software workflows.

Different proteomics software workflows obtain different number of peptides or proteins identified and also different number of missing values (for example, [19],[20]). These differences may have little relevance to the final conclusions when thousands or tens of thousands of tryptic peptides are identified in classical proteomics analysis using one particular bioinformatics tool. However, in viral immunoproteomics, when only few dozens of HLA class I or II ligands are identified in each study, the election of software engine may be very important. Thus, three popular proteomics tools used in the identification of HLA ligands were compared.

Using Proteome Discoverer, MaxQuant, and Peaks bioinformatics tools, 44, 58, and 59 fragmentation spectra presented in the HRSV-infected HLA-DP-bound peptide pool, but absent from the control uninfected pool, were resolved with high confidence parameters as HRSV peptides, respectively (Table 1 and Fig. 2). Additionally, a human proteome database search failed to identify any of these spectra as human protein fragments, verifying thus the viral origin of these peptides. Most of these peptides were identified by 2 or 3 proteomics software workflows (Table 1, Fig. 2, and Supplementary Fig. 1). Only 2 and 3 HLA-DP ligands were identified exclusively by MaxQuant and Peaks bioinformatics tools, respectively (Table 1, Fig. 2, and Supplementary Fig. 1). In contrast, all peptides identified using Proteome Discovered were also detected with at least one of the other two proteomics tools utilized (Table 1, Fig. 2, and Supplementary Fig. 1). In total, combining MaxQuant and Peaks tools 64 fragmentation spectra were assigned as viral sequences (Table 1, Fig. 2, and Supplementary Fig. 1). In addition, the theoretical assignments of six of the HLA-DP ligands, which ranked low scores with Maxquant or Peaks bioinformatics tools, were newly confirmed by MS/MS spectrum identification of the corresponding synthetic peptide (Supplementary Fig. 2).

In summary, the use of two software engines as MaxQuant and Peaks increased the identification of HLA-DP ligands from HRSV-infected cells.

The antigen processing generated multiple different viral HLA-DP ligands in human HRSV-infected cells.

The different HLA-DP viral ligands identified were derived from five different HRSV structural proteins. For example, in the N-terminus of the viral nucleoprotein (N) a complex nested set of 3 minimal cores (N_{2-10}, N_{6-14}, and N_{15-27}) and their C- and N-terminally extended peptides were detected (Table 1 and Fig. 3). In addition STGDSIDTPNYDVQK (N_{28-42}), and IEIESRKSYK ($N_{127-136}$) viral sequences were also identified (Table 1 and Fig. 3). Other three additional MS/MS spectra were assigned to amino acid sequences of the same HRSV N protein ($N_{140-150}$, $N_{343-365}$, and $N_{380-391}$), which were also the minimal core of other three N-terminally extended peptides (Table 1 and Fig. 3). Thus, 16 different viral N protein (individual or included in several nested sets of ligands) were presented by HLA-DP molecules in HRSV-infected cells (Fig. 3).

Similarly to N protein, multiple individual and nested sets of ligands from the fusion (F), matrix (M), attachment G, and phosphoprotein (P) proteins were generated by antigen processing in the HRSV-infected cells (Table 1 and Fig. 3). In summary, 15, 13, 11, and 9 different ligands derived from HRSV F, M, G, and P proteins were bound to HLA-DP molecules in the virus-infected cells, respectively (Fig. 3). Thus, these results indicate that a total of 64 HLA-DP ligands, 58 of which were included in some of different nested sets of peptides from the F, G, M, N and P proteins, were processed and presented in HRSV-infected cells.

Theoretical binding affinity of the HRSV peptides to HLA-DP molecules.

Prediction of HLA peptide binding by bioinformatics tools is usually employed to select potential candidates for viral ligands. Thus, to analyze the accuracy of the current algorithms, prediction of the HLA-DP A1*01:03/DPB1*0201, and -DP A1*01:03/ DPB1*0401 heterodimers binding of the HRSV ligands, identified by mass

spectrometry analysis, was carried out using two different computational approaches, the IEDB prediction software and the NetMHCIIpan neural network-based alignment. These analyses showed that only the HRSV peptides from G protein were predicted by both bioinformatics tools as HLA-DP high affinity ligands. In contrast, both IEDB and NetMHCIIpan analytical software's failed to predict the binding to HLA-DP DPA1*01:03/ B1*0201, or DPA1*01:03/B1*0401 heterodimers of all the ligands detected by mass spectrometry from F, M, N, and P proteins. In addition, both IEDB and NetMHCIIpan bioinformatics tools failed to predict the binding of all HLA-DP-bound ligands detected by mass spectrometry to the different HLA-DR class II molecules expressed by the JY cell line (Suplementary Table 2).

A significant fraction of the viral proteome was monitored in association with the HLA-DP class II molecules.

The different ligands bound to HLA-DP molecules, identified by mass spectrometry, covered between 30 residues from G protein to 115 amino acids from N protein (Table 2 and Fig. 3). These HLA-DP ligands represent between the 10% from G protein to the 29% from M and N proteins (Table 2 and Fig. 3). These high percentages of HLA-DP-restricted immune coverage suggest a high cellular immune pressure focused on these five proteins. In addition, as the Long strain HRSV proteome contains 4,402 residues, thus the 64 ligands presented by HLA-DP molecules represent the 9% of the viral proteome (Table 2).

In addition, in a previous immunoproteomics analysis 19 viral ligands, which were presented by several frequent HLA-DR class II molecules in the same HRSV-infected cells used in the current report, were identified [11]. These naturally processed HLA-DR ligands were derived from the viral F, M, and NS1 proteins with immune protein coverage of 13%, 7%, and 16%, respectively (Table 2). Thus globally, the HLA class II-restricted immune coverage ranged between the 10% from G protein to the

36% from M, and the global proteome was increased to 10% since different HLA-DR and -DP ligands were partially overlapping (Table 2).

Conservation of ligands among several HRSV strains.

HRSV exists as a single serotype, but presents two different antigenic subgroups, A and B [2]. Most viral proteins are highly conserved between these subgroups with an 88-99% of amino acid identity. The greater differences are found in G, M2-2, and SH proteins with only 61-71% amino acid identity. Within subgroups, the percentage of nucleotide and amino acid identity between different virus strains is much higher for all proteins (97-100% amino acid identity). Specifically, the percentage of amino acid identity between HRSV Long strain and the strains of HRSV subtype B for F, G, M, N, and P proteins was 95%, 63%, 96%, 99%, and 92%, respectively.

Thus, the sequences of the 64 HLA-DP ligands bound to HLA-DP class II molecules identified by mass spectrometry from HRSV Long strain (Fig. 3) were aligned with homologs derived from other two representative strains of HRSV subtype A, and three of HRSV strains B. Half of the HLA class II HRSV ligands were fully conserved among these virus strains, whereas the others showed at least one single residue change (Table 3). These results were in accordance with the percentage of amino acid identity of the proteins of antigenic subgroups of this virus.

All HLA-DP ligands from the F protein are accessible to direct interaction with HLA class II molecules.

The accessibility of the 15 F protein HLA-DP class II ligands that were identified by mass spectrometry from HRSV-infected cells was analyzed using the crystal structures of HRSV F proteins available in the Protein Data Bank as 4ZYP [21] and 3RRT [22]. All these HLA-DP ligands from HRSV F protein were located in the exposed

regions of the respective crystal structure, and were thus accessible to direct interaction with HLA-DP molecules.

As the crystal structure of G protein was not available, similar analysis to F protein could not be performed.

DISCUSSION

Different issues could be derived from the results reported in the current study about the HLA-DP-restricted response against HRSV.

Classically, the identification of virus-derived peptides presented by HLA molecules was usually addressed using partially overlapping synthetic peptides covering selected domains or full sequence of pathogen proteins or even the whole proteome when the virus is small enough [23]. Moreover in the last years, prediction of HLA peptide binding by bioinformatics algorithms emerged as a cheap and valuable tool to select potential candidates for high affinity ligands from complete pathogen proteomes. Nevertheless, the accuracy of the results obtained against the natural infection by these computational tools is highly dependent on the precision of the algorithm. In fact, in the present study, a broad inconsistency between the experimental data and the two bioinformatics tools utilized was found. Only the two cores from HRSV G protein showed high binding score by both NetMHCIIpan and IEDB tools. In contrast, these two bioinformatics algorithms failed to predict binding of the other 53 natural HLA-DP ligands included in 24 cores. It is remarkable that of the three previously described HLA-DP-restricted epitopes [8],[10], only those corresponding from HRSV G protein, but not those included in the viral F and M protein sequences, showed theoretical high binding affinity using both bioinformatics tools. Similarly, we found that only 54 of the 395 bacterial or viral HLA-DP-restricted ligands and epitopes included in the IEDB database showed theoretical high binding affinity using the same *in silico* binding prediction algorithms. In contrast, in a previous study using similar HRSV-infected cells both bioinformatics tools predicted properly that most ligands identified by mass spectrometry analysis were high affinity peptides from HLA-DR molecules [11]. This striking discrepancy indicates that for some HLA-DP but not HLA-DR class II alleles, these predictors must be refined by training with new sequences obtained experimentally by both functional and mass spectrometry assays in order to improve

their preciseness. In addition to HLA peptide binding, currently computed in these algorithms, the different elements involved in the HLA class II-restricted antigen processing and presentation pathway as the class II invariant chain, the accessory molecule HLA-DM, the endocytic activity, the lysosomal pH, or the activity of the different proteases involved, must be included by IEDB or NetMHCIIpan bioinformatics tools.

In comparison with the HLA-DR and -DQ, HLA-DP displays a remarkable feature as presenting molecule: the existence of a DP supertype including five HLA-DP class II molecules that bind a peptide repertoire very similar [24]. These HLA-DP supertype molecules are found in >90% of the human population [24]. Thus, most of HLA-DP-restricted virus peptides described in the current study are likely be universal ligands for most of the different HLA-DP molecules.

In the immunoproteomics analysis of HLA ligandomes mainly host-derived peptides are usually identified with a very small fraction of pathogen peptide sequences. In the current study, in the analysis with Peaks bioinformatics tool 58 viral peptides and 2631 human-derived sequences were found. Thus, the proportion of foreign peptides presented on HLA-DP molecules was 2.2%. Previously, in a similar immunoproteomics study 19 HRSV sequences and 728 host-derived ligands bound to HLA-DR molecules were identified [11]. Thus, the ratio non-self/self ligands was 2.6%, very similar to the detected in the current study. In summary, about 2-3% of total HLA class II ligands were derived from HRSV proteins.

For years, two different hypotheses (reviewed in [25]) have tried to explain how the MHC class II ligands are generated. The first is the “cut/trim first, bind later” model, which proposes an initial proteolytic activity of lysosomal proteases on the protein antigens and later, the resulting processed peptides bind to HLA class II molecules. In contrast, the second “bind first, cut/trim later” model suggests that HLA class II molecules scan the protein antigens. If the interaction is sufficient, the HLA class II

molecule protects the ligand, and the lysosomal proteases can only trim the protruding ends that are unprotected by HLA class II molecules to release the final antigen ligand. Both models could not necessarily be mutually exclusive according to other study, which suggests that autoantigens follow the “cut/trim first, bind later” model whereas pathogen-derived proteins could be processed according the second model [26]. In the “bind first, cut/trim later” model the ligands must be exposed to allow the direct interaction with HLA class II molecules. The HRSV F protein crystal structure determines that all HLA-DP class II ligands from this protein identified by mass spectrometry in the current report were located in the exposed regions of this viral protein. As the crystal structure of HRSV G protein was not available, and the HRSV M, N, and P crystal structures were individually resolved, but not in its complex form of virion, similar analysis to F protein could not be performed. In addition, ten of the eleven HLA-DR class II ligands identified in a similar study using HRSV-infected cells were also located in the exposed regions of the HRSV F protein [11]. In summary, only one of the 26 HLA class II-restricted viral ligands identified from F protein in HRSV-infected cells was not accessible to solvent and direct interaction with HLA class II molecules. These data would be compatible with the “bind first, cut/trim later” model for HLA class II pathogen-derived epitopes.

Five of six largest structural proteins from HRSV virus: F, G, M, N and P were source of HLA-DP ligands. The exception was the viral polymerase (L), the largest HRSV protein whose 2165 amino acids represent the 49% of the viral proteome. Previously, HLA-DR class II-restricted ligands were not found in this L protein [11]. In addition, only one HLA class I-restricted ligand with a weak T-cell response was identified from this polymerase [13],[14]. These data suggest that polymerase is not relevant for antiviral cellular immune response against HRSV, a fact that has significant implications in the design of recombinant vaccines against this virus. In addition, the humoral immune response against HRSV is focused on the F protein and no antibodies

against L protein were previously described in both humans and animal models. Thus, the viral polymerase seems to be totally irrelevant to the adaptive immune response against HRSV, and its gene could be excluded from any future vaccine construct, greatly simplifying its design.

Lastly, the analysis of the HLA-DP-restricted virus peptide repertoire described in this report together with the previous immunoproteomics studies analyzing the HLA class I [13],[27],[14] and HLA-DR [11] responses have additional implications for HRSV vaccine development. The previous HLA class I- and DR-restricted immune responses identified were mainly targeted to N, M, and F proteins [13],[27],[14],[11]. As these three viral proteins in addition to HRSV G, and P proteins were the source of HLA-DP ligands thus, a recombinant vaccine expressing these five proteins against which both the antiviral HLA class I and II immune responses are mainly targeted might be the most complete immunogen for HRSV protection.

CONCLUSIONS

Using high-throughput mass spectrometry analysis we identified sixty-four naturally processed HRSV ligands bound to HLA-DP class II molecules. This complex pool of viral ligands arose from five of six major structural HRSV proteins: attachment, fusion, matrix, nucleoprotein, and phosphoprotein, but not from the polymerase protein, the largest HRSV protein that includes half of the viral proteome. The existence of a DP supertype including five HLA-DP class II molecules that bind a very similar peptide repertoire makes these HLA class II molecules an interesting target for the design of vaccines.

HIGHLIGHTS

- High-throughput mass spectrometry analysis to identify HRSV ligands bound to HLA-DP class II molecules.
- Sixty-four naturally processed viral ligands, most of them included in complex nested set of peptides, were identified bound to HLA-DP molecules.
- The viral ligands arose from five of six major structural HRSV proteins: attachment, fusion, matrix, nucleoprotein, and phosphoprotein.
- No viral ligands were identified from polymerase protein, the largest HRSV protein that includes half of the viral proteome.

BIBLIOGRAPHY

[1] K. L. Rock, E. Reits, J. Neefjes. Present Yourself! By MHC Class I and MHC Class II Molecules. *Trends Immunol* 2016; 37:724-37.

[2] P. L. Collins, R. M. Chanock, B. R. Murphy. Respiratory Syncytial Virus. In *Fields Virology*. D. M. Knipe and P. M. Howley, eds. Lippincott Williams & Wilkins, Philadelphia, PA., 2013, pp. 1086-124.

[3] P. Glezen and F. W. Denny. Epidemiology of acute lower respiratory disease in children. *N Engl J Med* 1973; 288:498-505.

[4] R. Lozano, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; 380:2095-128.

[5] C. B. Hall. Nosocomial respiratory syncytial virus infections: the "Cold War" has not ended. *Clin Infect Dis* 2000; 31:590-6.

[6] B. S. Graham, L. A. Bunton, P. F. Wright, D. T. Karzon. Role of T lymphocyte subsets in the pathogenesis of primary infection and rechallenge with respiratory syncytial virus in mice. *J Clin Invest* 1991; 88:1026-33.

[7] M. E. Levely, C. A. Bannow, C. W. Smith, J. A. Nicholas. Immunodominant T-cell epitope on the F protein of respiratory syncytial virus recognized by human lymphocytes. *J Virol* 1991; 65:3789-96.

[8] G. M. Van Bleek, M. C. Poelen, M. R. van der, H. F. Brugghe, H. A. Timmermans, C. J. Boog, P. Hoogerhout, H. G. Otten, C. A. van Els. Identification of immunodominant epitopes derived from the respiratory syncytial virus fusion protein that are recognized by human CD4 T cells. *J Virol* 2003; 77:980-8.

[9] V. Yusibov, V. Mett, V. Mett, C. Davidson, K. Musiychuk, S. Gilliam, A. Farese, T. Macvittie, D. Mann. Peptide-based candidate vaccine against respiratory syncytial virus. *Vaccine* 2005; 23:2261-5.

[10] W. L. de, S. Yuksel, A. H. Brandenburg, J. P. Langedijk, K. Sintnicolaas, G. M. Verjans, A. D. Osterhaus, R. L. de Swart. Identification of a common HLA-DP4-restricted T-cell epitope in the conserved region of the respiratory syncytial virus G protein. *J Virol* 2004; 78:1775-81.

[11] E. Lorente, A. Barriga, E. Barnea, C. Mir, J. A. Gebe, A. Admon, D. Lopez. Structural and Nonstructural Viral Proteins Are Targets of T-Helper Immune Response against Human Respiratory Syncytial Virus. *Mol Cell Proteomics* 2016; 15:2141-51.

[12] C. R. Bangham and A. J. McMichael. Specific human cytotoxic T cells recognize B-cell lines persistently infected with respiratory syncytial virus. *Proc Natl Acad Sci U S A* 1986; 83:9183-7.

[13] S. Infantes, E. Lorente, E. Barnea, I. Beer, J. J. Cragnolini, R. García, F. Lasala, M. Jiménez, A. Admon, D. López. Multiple, non-conserved, internal viral ligands naturally presented by HLA-B27 in human respiratory syncytial virus-infected cells. *Mol Cell Proteomics* 2010; 9:1533-9.

[14] C. Johnstone, E. Lorente, A. Barriga, E. Barnea, S. Infantes, F. A. Lemonnier, C. S. David, A. Admon, D. Lopez. The Viral Transcription Group Determines the HLA Class I Cellular Immune Response Against Human Respiratory Syncytial Virus. *Mol Cell Proteomics* 2015; 14:893-904.

[15] E. Lorente, S. Infantes, E. Barnea, I. Beer, R. Garcia, F. Lasala, M. Jimenez, C. Vilches, F. A. Lemonnier, A. Admon, D. López. Multiple viral ligands naturally presented by different class I molecules in transporter antigen processing-deficient vaccinia virus-infected cells. *J Virol* 2012; 86:527-41.

[16] E. Lorente, A. J. Martin-Galiano, E. Barnea, A. Barriga, C. Palomo, J. Garcia-Arriaza, C. Mir, P. Lauzurica, M. Esteban, A. Admon, D. Lopez. Proteomics analysis reveals that structural proteins of the virion core and involved in gene expression are the main source for HLA class II ligands in vaccinia virus-infected cells. *J Proteome Res* 2019; 18:900-11.

[17] A. J. Watson, R. DeMars, I. S. Trowbridge, F. H. Bach. Detection of a novel human class II HLA antigen. *Nature* 1983; 304:358-61.

[18] Y. Ishihama, J. Rappsilber, J. S. Andersen, M. Mann. Microcolumns with self-assembled particle frits for proteomics. *J Chromatogr A* 2002; 979:233-9.

[19] C. Ramus, A. Hovasse, M. Marcellin, A. M. Hesse, E. Mouton-Barbosa, D. Bouyssie, S. Vaca, C. Carapito, K. Chaoui, C. Bruley, J. Garin, S. Cianferani, M. Ferro, D. A. Van, O. Burlet-Schiltz, C. Schaeffer, Y. Coute, P. A. Gonzalez de. Benchmarking quantitative label-free LC-MS data processing workflows using a complex spiked proteomic standard dataset. *J Proteomics* 2016; 132:51-62.

[20] T. Valikangas, T. Suomi, L. L. Elo. A comprehensive evaluation of popular proteomics software workflows for label-free proteome quantification and imputation. *Brief Bioinform* 2018; 19:1344-55.

[21] M. S. Gilman, S. M. Moin, V. Mas, M. Chen, N. K. Patel, K. Kramer, Q. Zhu, S. C. Kabeche, A. Kumar, C. Palomo, T. Beaumont, U. Baxa, N. D. Ulbrandt, J. A. Melero, B. S. Graham, J. S. McLellan. Characterization of a Prefusion-Specific Antibody That Recognizes a Quaternary, Cleavage-Dependent Epitope on the RSV Fusion Glycoprotein. *PLoS Pathog* 2015; 11:e1005035.

[22] J. S. McLellan, Y. Yang, B. S. Graham, P. D. Kwong. Structure of respiratory syncytial virus fusion glycoprotein in the postfusion conformation reveals preservation of neutralizing epitopes. *J Virol* 2011; 85:7788-96.

[23] W. J. Schulze zur, G. M. Lauer, C. L. Day, A. Y. Kim, K. Ouchi, J. E. Duncan, A. G. Wurcel, J. Timm, A. M. Jones, B. Mothe, T. M. Allen, B. McGovern, L. Lewis-Ximenez, J. Sidney, A. Sette, R. T. Chung, B. D. Walker. Broad repertoire of the CD4+ Th cell response in spontaneously controlled hepatitis C virus infection includes dominant and highly promiscuous epitopes. *J Immunol* 2005; 175:3603-13.

[24] J. Sidney, A. Steen, C. Moore, S. Ngo, J. Chung, B. Peters, A. Sette. Five HLA-DP molecules frequently expressed in the worldwide human population share a common HLA supertypic binding specificity. *J Immunol* 2010; 184:2492-503.

[25] E. S. Trombetta and I. Mellman. Cell biology of antigen processing *in vitro* and *in vivo*. *Annu Rev Immunol* 2005; 23:975-1028.

[26] A. Kim, I. Z. Hartman, B. Poore, T. Boronina, R. N. Cole, N. Song, M. T. Ciudad, R. R. Caspi, D. Jaraquemada, S. Sadegh-Nasseri. Divergent paths for the selection of immunodominant epitopes from distinct antigenic sources. *Nat Commun* 2014; 5:5369.

[27] S. Infantes, E. Lorente, J. J. Cragnolini, M. Ramos, R. García, M. Jiménez, S. Iborra, M. Del Val, D. López. Unusual viral ligand with alternative interactions is presented by HLA-Cw4 in human respiratory syncytial virus-infected cells. *Immunol Cell Biol* 2011; 89:558-65.

FIGURES

Figure 1. Size and intracellular distribution of HLA-DP cellular ligandome.

A: Peptide size distribution of self-derived ligands. B: Cellular location of the HLA-DR-associated peptide parental proteins; extracellular (E), membrane (Mb), vesicles: endosome/lysosome (V), Golgi (G), endoplasmic reticulum (ER), cytosol (C), mitochondria (Mt) or nuclear (N).

Figure 2. Overlap of HLA-DP class II ligands between the proteomics software workflows included in this study.

Venn diagram showing absolute numbers of detected HLA-DP ligands shared between the different bioinformatics tools analyzed.

Figure 3. Naturally processed peptides from the different HRSV proteins identified by mass spectrometry.

Diagram of identified ligands bound to HLA-DP molecules from HRSV N, F, M, G, and P proteins, respectively. Minimal cores (red boxes), ant their N-, C-, or N- and C-extended ligands (blue boxes) are depicted in the lower section of each viral protein represented at scale.

SUPPORTING INFORMATION:

Supplementary Table 1. Identification of the human HLA-DP ligands in uninfected and HRSV-infected cell extracts by mass spectrometry.

Supplementary Table 2. Theoretical affinity to HLA-DR class II molecules of the HRSV HLA-DP and -DR ligands detected by MS/MS analysis in the HRSV persistently infected cells.

Supplementary Figure 1. Identification of the HLA-DP ligands in HRSV-infected cell extracts by mass spectrometry.

Supplementary Figure 2. Comparison between experimentally detected ligands and their synthetic peptides.

Table 1

Summary of the HLA-DP ligands that were detected by MS/MS analysis in the HRSV-infected cells

Experimental mass [a]	ΔMass [b]	m/z	Sequence [c]	Protein	Position	MQ [d]	PK [e]	PD [f]	Theoretical binding affinity [g]	
									DPA1*01:03/ DPB1*02:01	DPA1*01:03/ DPB1*04:01
1276.74	0.8	2	LIKQELDKYK	F	78-87	+	+			
1503.90	1.4	2	VKLIKQELDKYK	F	76-87	+	+	+		
2145.20	0.4	2	IKQELDKYKNAVTELQLL	F	79-96	+	+			
1630.82	-1.7	2	NAVTELQLLM*QSTPA	F	88-102	+	+			
1486.84	1.4	2	TNKAVVSLSNGVSVL [h]	F	174-188	+	+	+		
1587.88	4.6	2	TNKAVVSLSNGVSVLT	F	174-189	+	+			
1361.76	-1.2	3	DLKNYIDKQLL [h]	F	194-204	+	+	+		
1670.96	-0.5	2	DLKNYIDKQLLPIV	F	194-207	+	+			
1785.00	1.0	3	DLKNYIDKQLLPIVN	F	194-208	+	+	+		
1913.10	-0.12	3	DLKNYIDKQLLPIVNK	F	194-209	+	+	+		
2041.16	-1.2	3	DLKNYIDKQLLPIVNKQ	F	194-210	+	+	+		
1661.86	4.4	2	ISNIETVIEFQQKN	F	214-227	+	+	+		
1303.61	0.9	2	INDM*PITNDQK	F	261-271	+	+	+		
1214.59	1.3	2	YYVNKQEGKS	F	457-466	+	+			
1986.01	0.8	3	TVSVGNTLYYVNKQEGKS	F	449-466	+	+	+		

879.51	-1.3	2	ISSGLYKL	G	26-33	+	+		++	++
1053.46	0.9	1	NDFHFEVF	G	161-168	+	+		++	++
1732.70	0.0	3	NDFHFEVFNFVPC*	G	161-173	+	+	+	++	++
1819.73	-0.9	2/3	NDFHFEVFNFVPC*S	G	161-174	+	+	+	++	++
1392.65	4.7	2	KPNNDFHFEVF	G	158-168	+	+	+	++	++
1506.69	3.9	2	KPNNDFHFEVFN	G	158-169	+	+	+	++	++
1849.88	-0,9	2	KPNNDFHFEVFNFVP	G	158-172	+			++	++
2071.89	1.7	3	KPNNDFHFEVFNFVPC*	G	158-173		+	+	++	++
2158.92	2.3	3	KPNNDFHFEVFNFVPC*S	G	158-174	+	+	+	++	++
2272.01	0.4	3	KPNNDFHFEVFNFVPC*SI [h]	G	158-175	+		+	++	++
1828.89	0.2	2	KPPNKPNNDFHFEVF	G	154-168	+	+	+	++	++
1377.62	2.4	2	M*ETYVNKLHEG	M	1-11	+	+			
1829.81	2.6	2	M*ETYVNKLHEGSTYT	M	1-15	+	+			
3290.49	2.7	3	M*ETYVNKLHEGSTYTAAVQYNVLEKDDD	M	1-28		+			
2495.18	0.2	3	KLHEGSTYTAAVQYNVLEKDDD	M	7-28	+	+	+		
1637.81	1.0	2	DLNTLENITTTEFK	M	173-186	+	+	+		
1879.95	2.3	2	NKDLNTLENITTTEFK	M	171-186	+	+	+		
2135.12	2.1	3	VRNKDLNTLENITTTEFK	M	169-186	+	+	+		
1002.61	-0,5	2	KIIPYSGLL	M	193-201	+		+		
1472.86	1.1	2	AITNAKIIPYSGLL	M	188-201	+		+		
1585.94	-2.3	2	AITNAKIIPYSGLLL	M	188-202	+	+	+		
1586.90	0.8	2	NAITNAKIIPYSGLL	M	187-201	+	+	+		
1291.71	0.8	2	VITVTDNKGAFK	M	203-214	+	+			
1404.80	2.8	2	LVITVTDNKGAFK	M	202-214	+	+	+		
1028.59	3.4	2	A*LSKVKLND	N	2-10	+	+	+		
1484.86	1.5	2	A*LSKVKLNDTLNK	N	2-14	+	+	+		
1599.88	1.4	2	A*LSKVKLNDTLNKD	N	2-15	+	+	+		

3004.66	4.4	4	A*LSKVKLNDTLNKDQLLSSSKYTIQR	N	2-27		+	
1043.60	1.8	2	VKLNDTLNK	N	6-14	+	+	+
1902.04	1.5	3	VKLNDTLNKDQLLSSSK	N	6-22	+	+	+
2563.40	1.8	4	VKLNDTLNKDQLLSSSKYTIQR	N	6-27	+	+	+
1537.81	2.7	2	DQLLSSSKYTIQR [h]	N	15-27	+	+	+
1638.74	3.7	2	STGDSIDTPNYDVQK	N	28-42	+	+	+
1251.68	-1.1	2	IEIESRKSYK	N	127-136	+	+	
1323.61	-0.6	2	KEM*GEVAPEYR	N	140-150	+	+	+
1583.73	2.0	2	M*LKEM*GEVAPEYR	N	138-150	+	+	+
2629.27	-0.4	3	NQDLYDAAKAYAEQLKENGVINY	N	343-365	+	+	+
3041.50	3.2	3/4	GTPRNQDLYDAAKAYAEQLKENGVINY	N	339-365		+	+
1420.69	1.0	2	HQLNPKDNDVEL	N	380-391	+	+	+
2633.32	-0.2	3	DLTAEELEAIKHQLNPKDNDVEL	N	369-391	+	+	+
1119.63	1.0	3	IKGKFTSPKD	P	24-33		+	
1508.80	-0.6	2	YQRKPLVSFKED	P	79-90	+	+	
2995.58	4.5	4	YQRKPLVSFKEDPIPSDNPFSKLYK [h]	P	79-103	+	+	+
2704.46	2.5	4	RKPLVSFKEDPIPSDNPFSKLYK	P	81-103		+	+
1748.86	2.2	2	EDPIPSDNPFSKLYK	P	89-103	+	+	+
1504.79	2.6	3	PIPSDNPFSKLYK	P	91-103	+	+	
1014.55	1.8	2	VVASAGPTSAR [h]	P	153-163	+	+	
1129.57	1.5	2	VVASAGPTSARD	P	153-164	+	+	+
1641.84	1.1	3	VVASAGPTSARDGIRDA	P	153-169	+		

[a] The monoisotopic ion mass in amu.

[b] The difference between nominal and experimentally detected monoisotopic ions in ppm.

[c] Asterisks indicate acetylation N-terminal residue, oxidation of Met or cysteinylation of Cys.

[d, e, and f] Identified peptides with MaxQuant, Peaks, and Proteome Discoverer bioinformatic tools, respectively.

[g] As calculated in the NetMHCIIpan-3.0 and the Immuno Epitope Database servers. The threshold for strong (++) binding peptides was 2 of the percentile rank.

[h] The bioinformatic assignment was confirmed with the MS/MS spectrum of corresponding synthetic peptide (Supplementary Figure 1).

Table 2

Distribution of HLA class II immune coverage in the HRSV proteome

Protein [a]	Nº of residues [b]	Nº of residues of HLA-DP ligands detected by MS/MS	% of HLA-DP immune coverage [c]	% of HLA-DR immune coverage [d]	% of HLA class II immune coverage [e]
F	574	103	18	16	26
G	298	30	10	0	10
M	256	74	29	7	36
N	391	115	29	0	29
P	241	52	22	0	22
NS1	139	0	0	13	13
Proteome	4402	374	9	3	10

[a] The abbreviations used were F (Fusion protein), G (Attachment protein), M (Matrix protein), N (Nucleoprotein), P (Phosphoprotein), and NS1 (Non-structural protein 1).

[b] Number residues of each protein or the complete viral proteome.

[c] Number residues of ligands/number of total residues as a percentage.

[d] Number residues of HLA-DR ligands/number of total residues as a percentage from [11].

[e] Number residues of HLA-DR and -DP class II ligands/number of total residues as a percentage.

Table 3
Conservation of HLA-DP viral ligands in several HRSV strains

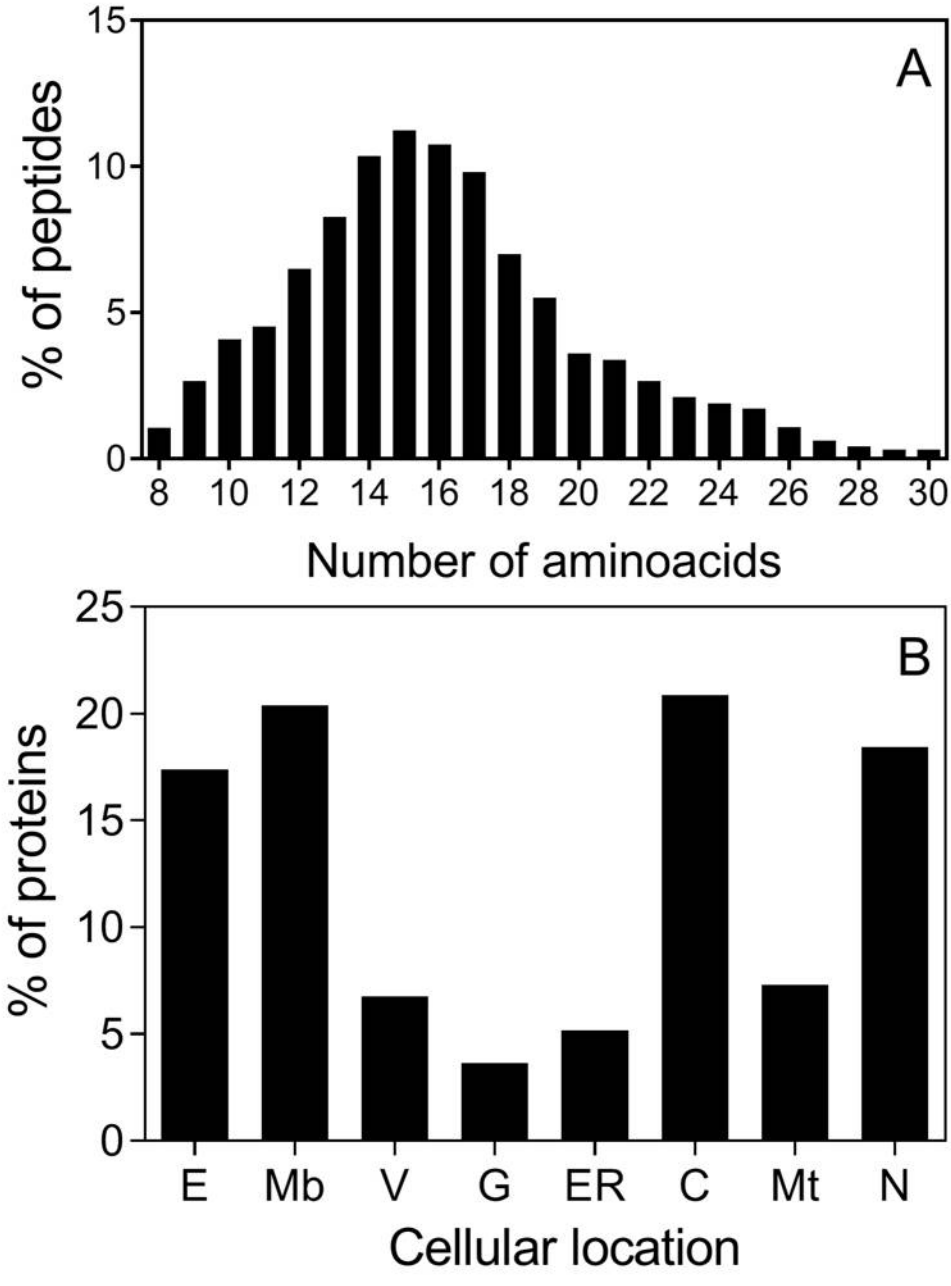
Strain [a,b]	F 76-87	F 88-102	F 174-189	F 194-210	F 214-227	F 261-271
Long	VKLIKQELDKYK	NAVTELQLLMQSTPA	TNKAVVSLSNGVSVLT	DLKNYIDKQLLPIVNKQ	ISNIETVIEFQQKN	INDMPITNDQK
S2	------------	S--------------	----------------	-----------------	--------------	-----------
A2	------------	--------------P	----------------	-----------------	--------------	-----------
9320	------------	---------T-N---	----------------	---S--NN-------Q-	--------------	-----------
18537	------------	-----------N---	----------------	-----NNR-------Q-	------------M-	-----------
B1	------------	-----------N---	----------------	-----NN--------Q-	--------------	-----------

Strain	F 449-466	G 26-33	G 154-175	M 1-11	M 7-28	M 169-186
Long	TVSVGNTLYYVNKQEGKS	ISSGLYKL	KPPNKPNNDFHFEVFCSI	METYVNKLHEG	KLHEGSTYTAAVQYNVLEKDDD	VRNKDLNTLENITTTEFK
S2	------------------	---C----	------------------	-----------	----------------------	------------------
A2	------------------	---C----	---S--------------	-----------	----------------------	------------------
9320	-------------L---N	---C--R-	N--K--KD-Y--------	-----------	----------------------	-------S----A-----
18537	-------------L---N	---C--R-	N--K--KD-Y--------			
B1	-------------L---N	---C--R-	N--K--KD-Y--------	-----------	----------------------	-------S----A-----

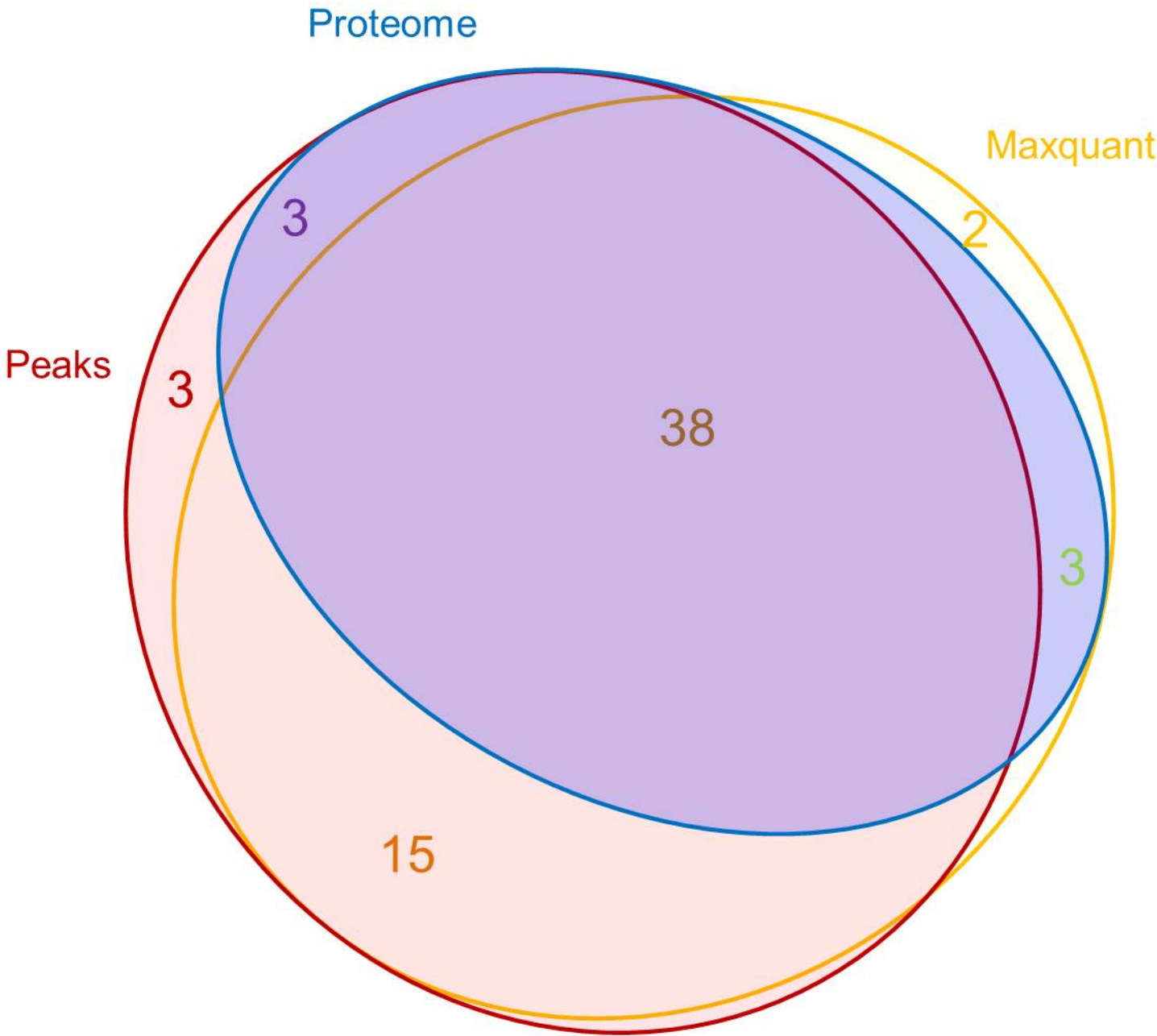
Strain	M 187-201	M 202-214	N 2-10	N 6-14	N 15-27	N 28-42	N 127-136	N 138-150
Long	NAITNAKIIPYSGLL	LVITVTDNKGAFK	ALSKVKLND	VKLNDTLNK	DQLLSSSKYTIQR	STGDSIDTPNYDVQK	IEIESRKSYK	MLKEMGEVAPEYR
S2	---------------	-------------	---------	---------	-------------	---------------	----------	-------------
A2	---------------	-------------	---------	---------	-------------	---------------	----------	-------------
9320	----------A--V	-------------	---------	---------	-------------	----N----------	----------	-------------
18537			---------	---------	-------------	----N----------	----------	-------------
B1	----------A--V	-------------	---------	---------	-------------	----N----------	----------	-------------

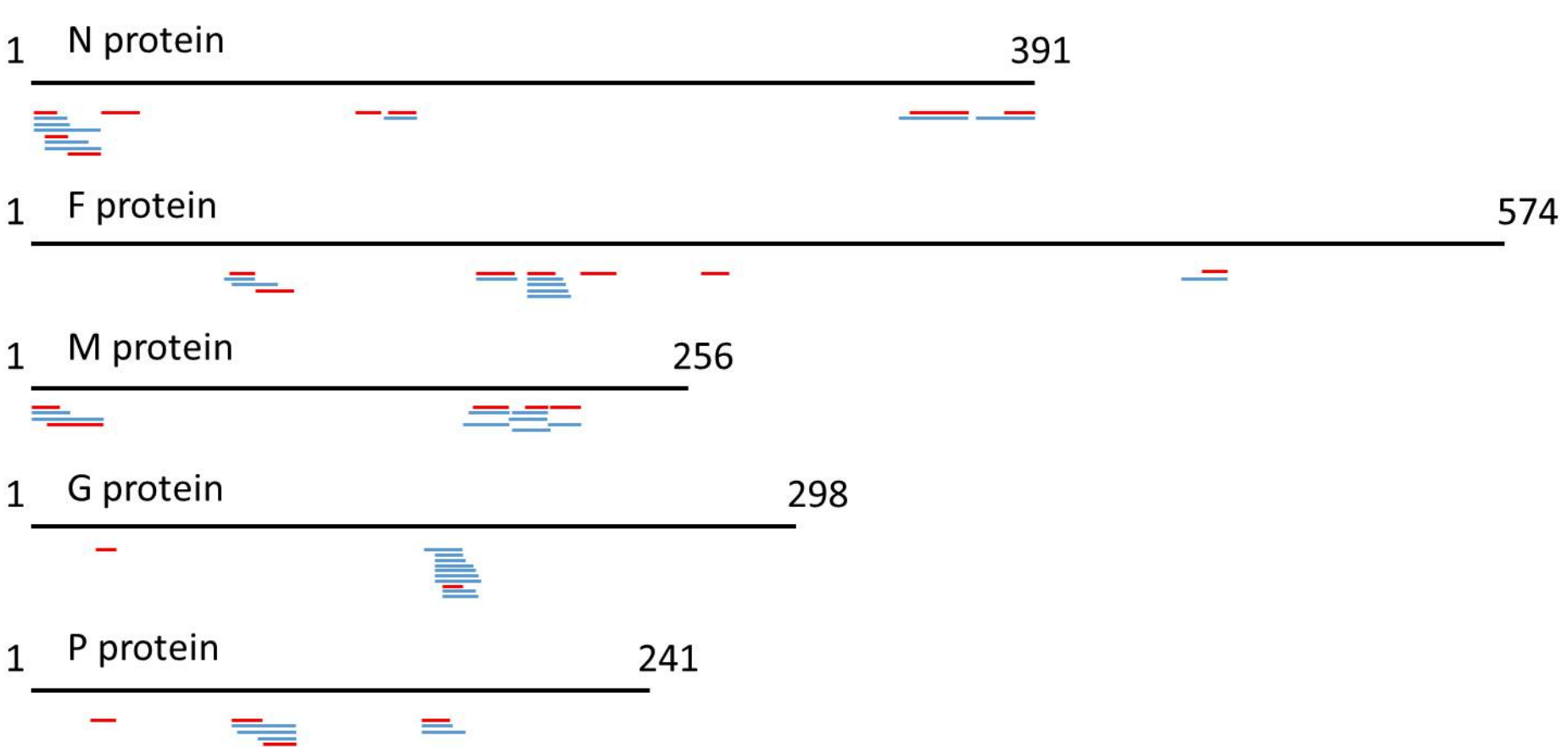
Strain	N 339-365	N 369-391	P 24-33	P 79-90	P 91-103	P 153-169
Long	GTPRNQDLYDAAKAYAEQLKENGVINY	DLTAEELEAIKHQLNPKDNDVEL	IKGKFTSPKD	YQRKPLVSFKED	PIPSDNPFSKLYK	VVASAGPTSARDGIRDA
S2	--------------------------	-----------------------	----------	----------	-T----------	-----------------
A2	--------------------------	-----------------------	----------	----------	-T----------	-----------------
9320	--------------------------	----------------ED----	-----A-S--	-P----------	LT----------	-----------------
18537	--------------------------	----------------ED----	-----A-S--	-P----D-----	LT----------	-----------------
B1	--------------------------	----------------ED----	-----A-S--	-P----------	LT----------	-----------------

[a] Long, S2 and A2 strains are representative of HRSV subtype A, whereas 9320, 18537 and B1 are HRSV B strains.
[b] Dashed lines mean conserved residues in comparison to HRSV Long strain.



Lorente et al. Figure 1


Proteome
Maxquant
Peaks
3
2
3
38
3
15


1
N protein
391
1
F protein
574
1
M protein
256
1
G protein
298
1
P protein
241