



# Allotypic variation in antigen processing controls antigenic peptide generation from SARS-CoV-2 S1 spike glycoprotein

Received for publication, July 13, 2021, and in revised form, October 2, 2021. Published, Papers in Press, October 22, 2021, <https://doi.org/10.1016/j.jbc.2021.101329>

George Stamatakis<sup>1</sup>, Martina Samiotaki<sup>1</sup>, Ioannis Temponeras<sup>2</sup>, George Panayotou<sup>1</sup>, and Efstratios Stratikos<sup>2,3,\*</sup>

From the <sup>1</sup>Institute for Bioinnovation, Biomedical Sciences Research Center "Alexander Fleming", Vari, Attica, Greece; <sup>2</sup>National Centre for Scientific Research "Demokritos", Agia Paraskevi, Attica, Greece; <sup>3</sup>Department of Chemistry, National and Kapodistrian University of Athens, Panepistimiopolis Zographou, Athens, Greece

Edited by Craig Cameron

Population genetic variability in immune system genes can often underlie variability in immune responses to pathogens. Cytotoxic T-lymphocytes are emerging as critical determinants of both severe acute respiratory syndrome coronavirus 2 infection severity and long-term immunity, after either recovery or vaccination. A hallmark of coronavirus disease 2019 is its highly variable severity and breadth of immune responses between individuals. To address the underlying mechanisms behind this phenomenon, we analyzed the proteolytic processing of S1 spike glycoprotein precursor antigenic peptides across ten common allotypes of endoplasmic reticulum aminopeptidase 1 (ERAP1), a polymorphic intracellular enzyme that can regulate cytotoxic T-lymphocyte responses by generating or destroying antigenic peptides. We utilized a systematic proteomic approach that allows the concurrent analysis of hundreds of trimming reactions in parallel, thus better emulating antigen processing in the cell. While all ERAP1 allotypes were capable of producing optimal ligands for major histocompatibility complex class I molecules, including known severe acute respiratory syndrome coronavirus 2 epitopes, they presented significant differences in peptide sequences produced, suggesting allotype-dependent sequence biases. Allotype 10, previously suggested to be enzymatically deficient, was rather found to be functionally distinct from other allotypes. Our findings suggest that common ERAP1 allotypes can be a major source of heterogeneity in antigen processing and through this mechanism contribute to variable immune responses in coronavirus disease 2019.

Immune responses to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the pathogen behind coronavirus disease 19 (COVID-19), play critical roles in disease pathophysiology (1, 2). Appropriate innate and adaptive immune responses are necessary for viral clearance, whereas aberrant uncontrolled responses have a major impact on mortality (3, 4). Furthermore, long-term immunity after infection or vaccination is of critical importance for ending the current pandemic (5). Although initial analyses focused on antibody-dependent immunity, T-cell-mediated immunity

appears to be important for both viral clearance and long-term immunity, especially against emerging virus variants (6–8). As a result, detailed knowledge of how SARS-CoV-2 epitopes are generated and selected is crucial for both allowing a better understanding of long-term antiviral immune responses by individuals and for the optimization of vaccines against virus variants (9).

COVID-19 is characterized by significant variability in disease severity amongst individuals (10). As a result, several studies have focused in determining genetic predispositions to COVID-19 in an effort to establish better preventive and treatment measures (11, 12). Indeed, rare mutations in components of the innate immune response and in particular the interferon I pathway can predispose individuals to severe COVID-19 (13, 14). In addition, the role of common polymorphic variation in components of adaptive immunity, such as human leukocyte antigen (HLA) alleles, has been emerging as a major factor (15–17).

HLA molecules (also called major histocompatibility complex [MHC] molecules) bind small peptides generated intracellularly from antigenic proteins of pathogens and present them on the cell surface. The HLA class I–peptide complexes interact with specialized receptors on CD8<sup>+</sup> T cells, and successful recognition indicates an infected cell and initiates a molecular response that leads to cell lysis (18). Several enzymes play roles in generating peptide ligands for HLA molecules. Endoplasmic reticulum aminopeptidase 1 (ERAP1) is an endoplasmic reticulum (ER)-resident aminopeptidase that trims precursors of antigenic peptides to optimize them for binding onto HLA (19). Appropriate of aberrant trimming of antigenic peptide precursors by ERAP1 can alter the repertoire of peptides available for presentation by HLA and indirectly regulate adaptive immune responses (20). Thus, the generation of antigenic epitopes by ERAP1 can be critical for immune evasion by viruses (21, 22).

HLA molecules are highly polymorphic with thousands of different alleles discovered to date (23). This polymorphic variation allows for the binding of a vast variety of peptide sequences, ensuring sufficient presentation of epitopes from unknown pathogens on a population level but not necessarily an individual level. Thus, HLA variability can underlie differences in immune responses between individuals. ERAP1 is also

\* For correspondence: Efstratios Stratikos, [estratikos@chem.uoa.gr](mailto:estratikos@chem.uoa.gr).

## ERAP1 allotypes affect the generation of SARS-CoV-2 epitopes

polymorphic, and coding SNPs in the ERAP1 gene associate with predisposition to autoimmunity or cancer often in epistasis with HLA alleles (24) and can help shape the cellular immunopeptidome (25). ERAP1 SNPs are found in the population in limited combinations that define specific allotypes and have functional consequences in peptide trimming (26, 27). The exact interplay of ERAP1 and HLA polymorphic variation in determining antigen presentation is currently a subject of active research (21, 27–29).

Generation and loading of antigenic peptides on HLA occur inside the ER, where thousands of different peptides compete with each other. While *in vitro* enzymatic analysis of ERAP1 function has provided significant insights on the mechanism of the enzyme (30, 31), the large network of possible interactions between substrates and the large substrate-binding cavity of ERAP1 results in a complex landscape of poorly understood specificity determinants (30, 31). To address these issues, we previously devised a proteomic approach to concurrently analyze hundreds of trimming reactions and compared the function of different antigen-processing enzymes (32). Here, we applied this approach to analyze the relative effect of the ten most common ERAP1 allotypes (covering 99.9% of the European population) in processing antigenic peptide precursors derived from the S1 spike glycoprotein of SARS-CoV-2. Our analysis suggests that ERAP1 allotypes carry significant sequence biases and generate sufficiently different peptide repertoires, which, in combination with HLA peptide binding specificity, can underlie differences in adaptive immune responses that contribute to COVID-19 disease severity variability in natural populations.

### Results

To explore how different ERAP1 allotypes can process S1 spike glycoprotein and generate putative antigenic peptides, we utilized a library of 315 synthetic 15mers, spanning the sequence of the S1 spike glycoprotein, with 11-residue overlap between adjacent peptides. These 15mer peptides model possible antigenic peptide precursors that are generated in the cytosol and enter the ER, where they are further digested by ERAP1 to generate smaller products that bind and are presented by MHC class I. Thus, this mixture constitutes a useful tool for the systematic sampling of the entire sequence of the protein and has been used before to compare the activities of

ERAP1 to other enzymes in the antigen processing pathway (32). To digest the mixture, we used highly pure recombinant ERAP1 protein variants corresponding to naturally occurring allotypes. These allotypes are defined as combinations of nine SNPs in the ERAP1 gene as shown in Table 1 and constitute the most common allotypes in humans, covering 99.9% of European population and 94.1% of the global population (27).

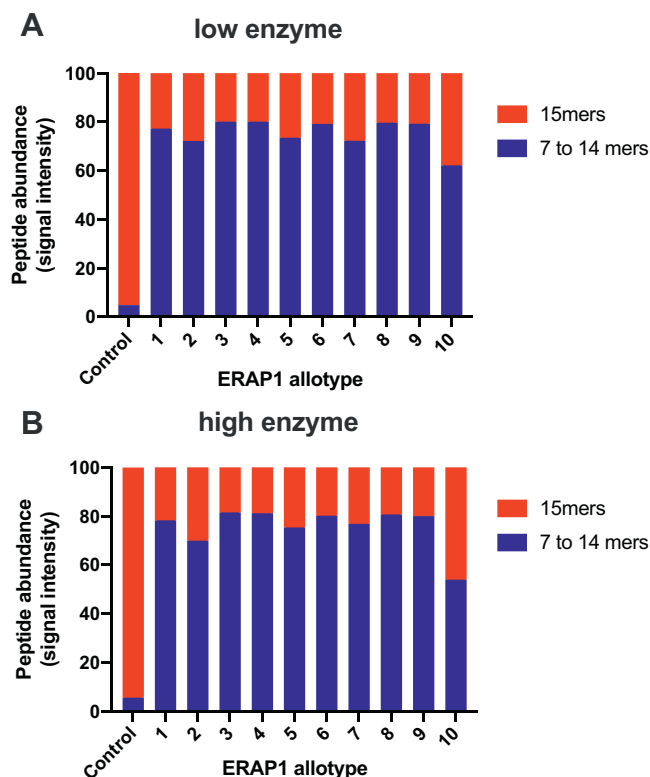
Since the trimming of antigenic peptide precursors by ERAP1 is a dynamic phenomenon, the peptide pool was mixed with recombinant ERAP1 at two concentrations (100 and 300 nM, henceforth called low-enzyme and high-enzyme condition) so as to provide us with insight on the kinetics of the reaction. After incubation, the digestion products were analyzed by LC–MS/MS using a custom search database generated by *in silico* digestions of the full S1 spike glycoprotein sequence (UniProt ID: P0DTC2). Three biological replicates for each reaction and three replicates of a negative control reaction were performed, totaling to 66 samples, and the identified peptides were filtered for robustness of detection as described in the *Experimental procedures* section. The unfiltered lists of identified peptides, along with relevant identification parameters, are shown separately for each reaction condition in Tables S1 and S2, respectively.

To evaluate the relative progress of each digestion, we first compared the total peptide abundance present after each reaction (Fig. 1). We grouped the peptides into two categories: (i) 15mers that correspond to the undigested peptides (substrates) and (ii) 7 to 14mers that correspond to the products of the digestions. For both reaction conditions (low enzyme and high enzyme), 70 to 80% of peptide signal came from digestion products for most allotypes, with relatively small differences between most allotypes. About 20% of 15mers were still present after digestion for both reaction conditions, likely representing peptides that are resistant to ERAP1 trimming. Allotype 10 was slightly less efficient, with 54 to 62% of peptide signal originating from product peptides. Still, this difference is much less pronounced than observed in previous *in vitro* experiments using single peptides that suggested that allotype 10 can be as much as 60-fold less active for some peptides (27). This finding suggests that although allotype 10 may be less active for some substrates, it is still able to operate with good efficiency in complex substrate mixtures.

To gain insight on the specific differences between allotypes, we generated heatmap plots of all reactions and performed

**Table 1**  
Amino acid composition at polymorphic positions for the ten most common ERAP1 allotypes

Allotype	Amino acid composition at each SNP								
	56	127	276	346	349	528	575	725	730
1	E	P	I	G	M	K	D	R	Q
2	E	R	I	G	M	K	D	R	Q
3	E	R	I	G	M	K	D	R	E
4	E	R	I	G	M	R	D	R	E
5	E	R	I	D	M	R	D	R	E
6	E	P	I	G	M	R	D	R	E
7	K	P	I	G	M	R	D	R	E
8	E	P	M	G	M	R	D	R	E
9	E	P	M	G	M	R	N	R	E
10	E	P	I	G	V	R	N	Q	E



**Figure 1. Relative peptide abundance in each sample, calculated from the total signal intensity, shown for the initial 15mer peptides and the produced 7 to 14mers, for both enzyme concentrations used.** Control reaction was performed by incubating the peptides in the absence of enzyme. The small percentages of 7 to 14mers identified in the control reaction are likely contaminants because of incomplete synthesis of the 15mer peptide pool and were not further interpreted.

cluster analysis as described in the *Experimental procedures* section (Fig. 2). For both enzyme concentrations, the pool of 15mer peptides was digested significantly as demonstrated by the reduction of intensity of detected 15mers (Fig. 2, A and C). Some differences were, however, evident between ERAP1 allotypes. This was most evident for allotype 10, which spared clusters of peptides, possibly because of its lower enzymatic activity as demonstrated before (27). Still, allotype 10 was able to efficiently trim large peptide clusters suggesting that its lower activity may be sequence specific.

In terms of production of 7 to 14mers, there were again apparent differences in the patterns between allotypes (Fig. 2, B and D). At both enzyme concentrations, allotype 10 appeared not only to underproduce several peptides but also to overproduce some peptide clusters. Similar, although less pronounced, phenomena were also evident for other allotypes. In particular, allotypes 4, 6, 8, and 9 appear to be the most efficient in producing a variety of different peptide products and tend to group together in the cluster analysis. Overall, while all allotypes were able to digest the 15mer substrates, significant differences in patterns are evident suggesting some differences in sequence preferences between allotypes.

Peptide length is one of the most important parameters for binding onto MHC class I molecules, and ERAP1 has been shown to show significant preferences for substrate length

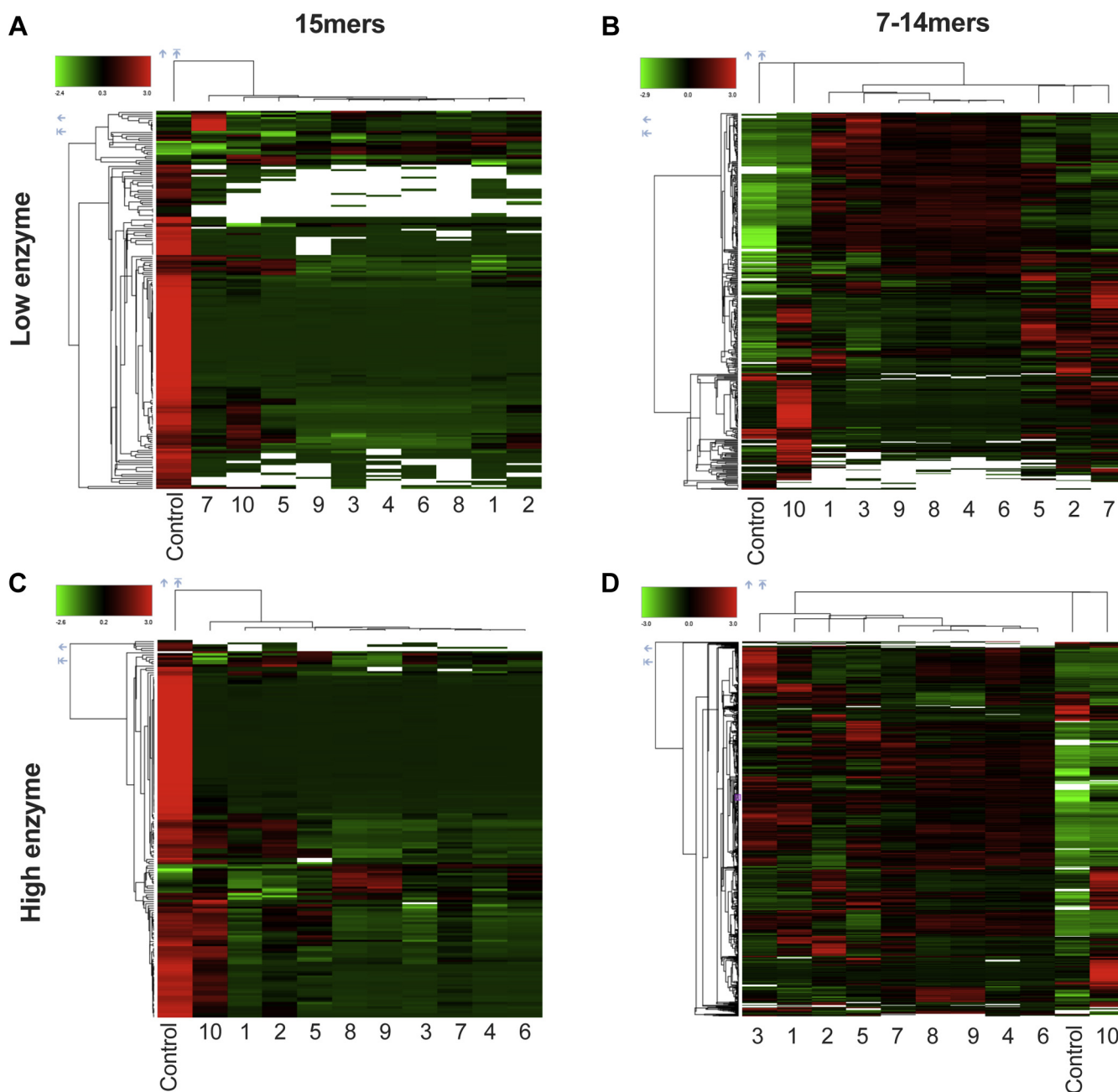
(30, 33). Most of the known MHC class I ligands are in the range of 8 to 12 amino acids, and the majority are 9mers. We thus analyzed the length distribution of produced peptides from each ERAP1 allele and for each of the two reaction conditions (Fig. 3). For this and following comparisons, we restricted analysis only for peptides that were detected in replicate measurements to be statistically significant ( $p < 0.05$ ) compared with the control reaction.

For the low-enzyme condition, the distribution of the length of peptides produced by all allotypes formed two apparent “peaks” around 9 and 13 amino acids long, respectively (Fig. 3A). The peak around 13 amino acids long was much more pronounced suggesting that sequential digestion was still limited for most peptides. Under this condition, allotypes 5, 7, and 10 lagged behind in terms of producing 9mer products.

The overall length distribution was reversed when a higher enzyme concentration was used (Fig. 3B). Overall, about double the number of peptides were identified when a higher amount of enzyme was used, consistent with faster sequential digestion of the 15mer substrates. In that condition, the majority of the produced peptides were 9mers for all allotypes, suggesting that all ERAP1 allotypes have the inherent capability to produce peptides with appropriate lengths for MHC class I binding. This is consistent with previous reports that have suggested that ERAP1 utilizes a “molecular ruler” mechanism that involves a regulatory allosteric binding site (30, 33). Interestingly, allotypes 5 and 7 that appeared to lag behind in terms of generating 9mers were now just as effective as the remaining allotypes, suggesting that 9mers can accumulate efficiently for those allotypes, albeit more slowly. In contrast, allotype 10 was again an outlier and produced a lower number of 9mer peptides compared with all the other allotypes.

A key parameter for antigenicity is the sequence of the presented peptides. In this context, it is important to know if the differences between generated peptides are just kinetic or also qualitative (*i.e.*, sequence). To gain insight on whether different ERAP1 allotypes generate different peptide sequences, we calculated, for each peptide identified, the number of ERAP1 allotypes that were able to produce it (Fig. 4). About double the number of different peptides were produced by the high-enzyme condition, as expected from the enhanced sequential trimming of the 15mer substrates when more enzyme is present. The distribution was found to be U shaped, in contrast to the expected bell shaped if peptides were processed in a completely random manner by unrelated protease activities, suggesting specific commonalities and differences in sequence specificity between allotypes. For both reaction conditions, most peptides were produced by all allotypes, indicating that all allotypes have a core sequence preference that shapes the products of the digestion. Strikingly, the second most populous category were peptides that were uniquely produced only by a single allotype (Fig. 4, A and B). This finding suggests that ERAP1 allotypes may carry substrate biases that can influence the repertoire of produced peptides. The generation of allotype-unique peptides suggests that

## ERAP1 allotypes affect the generation of SARS-CoV-2 epitopes



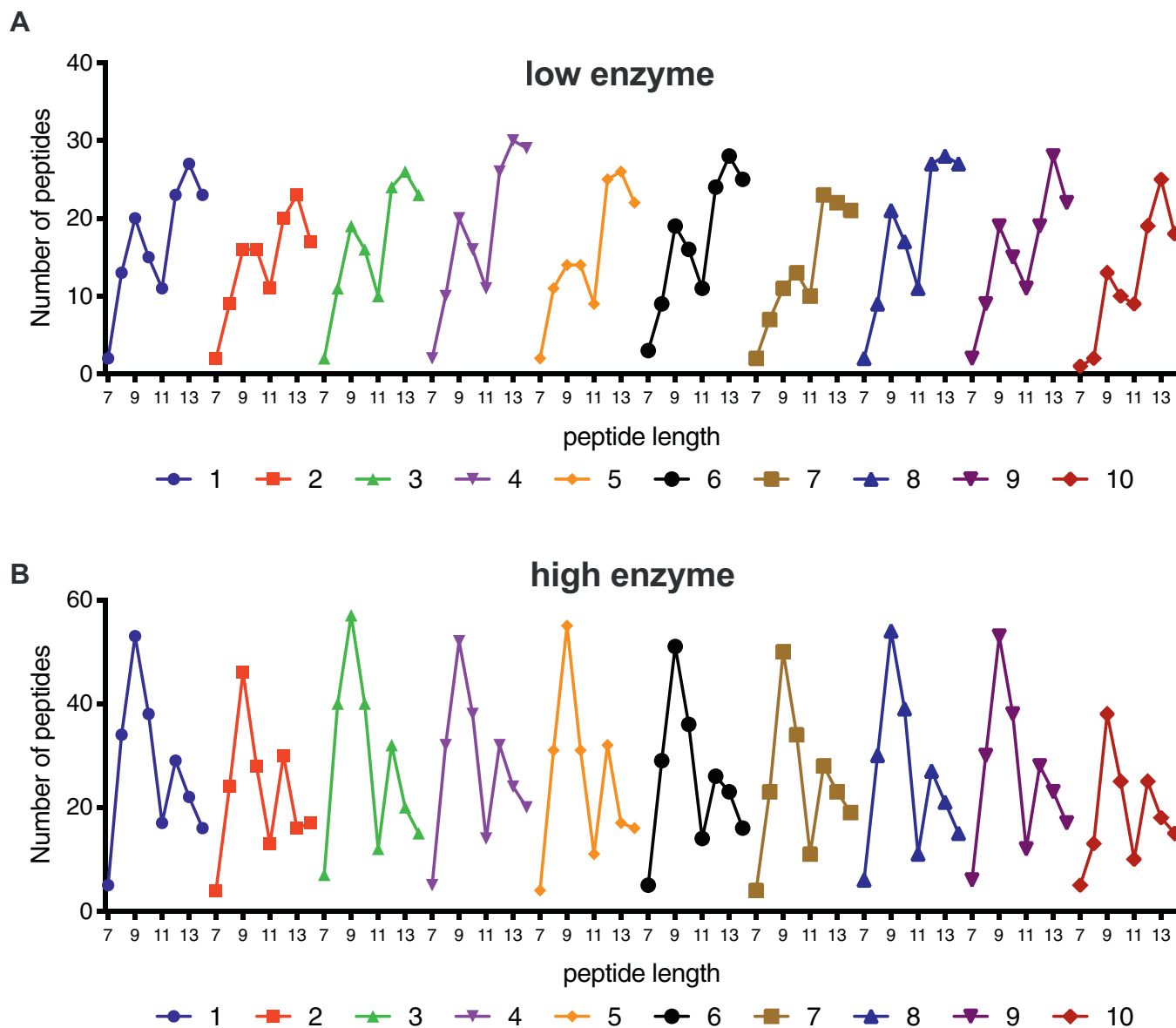
**Figure 2. Heatmap plots of peptides detected in each reaction condition.** Heatmap plots showing trimming of the pool of 15mer peptides (panels A and C) by all ERAP1 allotypes (allotypes 1–10) and generation of smaller peptide products 7 to 14 amino acids long (panels B and D) under two different conditions containing 100 and 300 nM enzyme, respectively (low-enzyme condition, panels A and B; high-enzyme condition, panels C and D). Samples are organized after cluster analysis performed by Proteome Discoverer 2.5 (indicated by *connecting lines*). Control: peptides incubated in the absence of enzyme. ERAP1, endoplasmic reticulum aminopeptidase 1.

different allotypes can potentially drive different immune responses by generating unique antigenic peptides.

To better understand the commonality of produced peptides for each allotype, we adapted the analysis shown in Figure 4 to report separately for each allotype (Fig. 5). For most allotypes and both reaction conditions, the distribution of produced peptides formed a trend toward common peptides, that is, most of the peptides produced by each allotype were also produced by several others. This was most clear for allotypes 6, 8, and 9. Several allotypes, however, namely allotypes 1 to 5, had a secondary peak around peptides that were either

unique or produced by one to three allotypes in total. A striking outlier was, again, allotype 10. In particular, in the high-enzyme condition, it was the only allotype in that the uniqueness distribution was reversed, and many of the peptides produced were either unique to allotype 10 or produced by few others. This observation suggests that there could be some sequence bias between allotypes, with allotype 10 being the most extreme example.

To further explore the similarities and differences between allotypes in a more unbiased manner, we performed principal component analysis (PCA) (Fig. 6). With a single exception

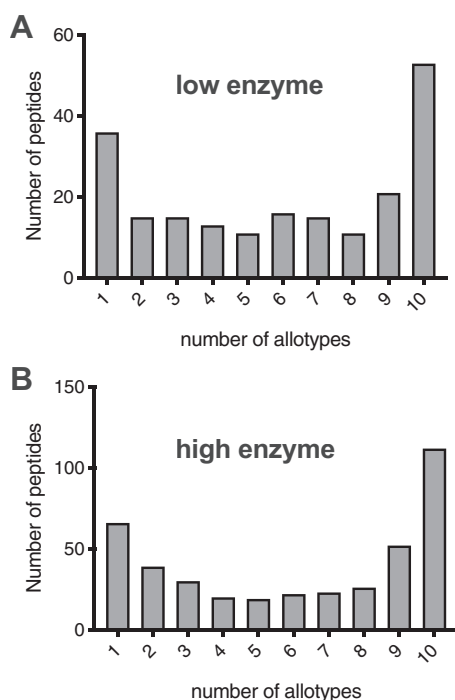


**Figure 3. Distribution of peptide length in 7 to 14mers produced after enzymatic digestion by each of the ten ERAP1 allotypes, shown for both reaction conditions.** ERAP1, endoplasmic reticulum aminopeptidase 1.

(one replicate for allotype 9), all replicates clustered together validating the reproducibility of our analyses. For the trimming of the 15mer substrates when using the low-enzyme condition, all allotypes and biological replicates cluster together, away from the control reactions (Fig. 6A). At the high-enzyme condition however, allotype 10 diverged from the other allotypes (Fig. 6C). The PCA for the generation of 7 to 14mers revealed a greater distribution between allotypes. For the low-enzyme condition, allotypes 1, 3, 4, 6, 8, and 9 formed a cluster. Allotype 10 was a strong outlier, followed by 7 and allotypes 2 and 5. For the high-enzyme condition, allotype 10 was again an outlier, followed by 2, 5, 7 and allotype 3. Overall, PCA revealed that the main difference between allotypes lie on the peptides produced and that allotype 10 consistently appears to differentiate from other allotypes, although intermediate differentiation is observable for some of the other allotypes.

The main biological phenomenon that drives cytotoxic T-lymphocyte responses is not intracellular processing but rather antigen presentation by MHC class I molecules (HLA in humans). Which peptides are presented depends on the combination of their availability in the ER and their ability to bind onto the particular HLA alleles that are expressed in each individual. To evaluate the HLA-binding ability of produced peptides, we utilized a well-established binding prediction server, NetMHCpan-4.1 (34). Over 10,000 HLA alleles have been discovered to date, having different preferences for binding peptides (35). In order to get a reasonable representation of the chances of each generated peptide to be presented, we restricted our analysis to a subset of the most common alleles in the human population, namely HLA-A\*01:01, HLA-A\*02:01, HLA-A\*03:01, HLA-A\*24:02, HLA-A\*26:01, HLA-B\*07:02, HLA-B\*08:01, HLA-B\*27:05,

## ERAP1 allotypes affect the generation of SARS-CoV-2 epitopes



**Figure 4. Frequency distribution of produced 7 to 14mer peptides from both reaction conditions.** The number of peptides produced by a specific number of ERAP1 allotypes is shown. For both reaction conditions, the majority of peptides are produced by all ten allotypes, but a large number of peptides are unique to a specific allotype. ERAP1, endoplasmic reticulum aminopeptidase 1.

HLA-B\*39:01, HLA-B\*40:01, HLA-B\*58:01, and HLA-B\*15:01. Each of the produced 8 to 12mers from every ERAP1 allotype was scored for each of the aforementioned alleles, and the best predicted percent rank was plotted for each allotype (Fig. 7). A percent rank below 2 (cyan region) is considered to be sufficient to promote binding and presentation by the particular allele.

The majority of produced 8 to 12mers were not predicted to be binders for the HLA alleles tested. This however does not mean that they would not be able to bind well onto HLA alleles not tested here. Not surprisingly, the high-enzyme condition produced a much greater number of peptides predicted to bind onto the HLA alleles tested, possibly because the higher degree of digestion produced more peptides of the appropriate length (such as 9mers, refer to Fig. 3B). Overall, all ERAP1 allotypes produced a significant number of peptides that are predicted to bind onto at least one of the HLA allele tested. This observation included the subactive allotype 10, which although is less efficient in generating peptides overall, it appears that it is efficient enough to generate many possible candidates for HLA presentation. Several of the peptides predicted to bind were produced by most or all allotypes (indicated by horizontal dotted lines in Fig. 7A), although some peptides were also produced uniquely by a single allotype (indicated by red highlighting in Fig. 7A). Some putative HLA binders produced by the low-enzyme condition were not detected in the high-enzyme condition, consistent with the reported ability of ERAP1 to destroy antigenic peptides by overtrimming (36). Overall, this analysis shows that although

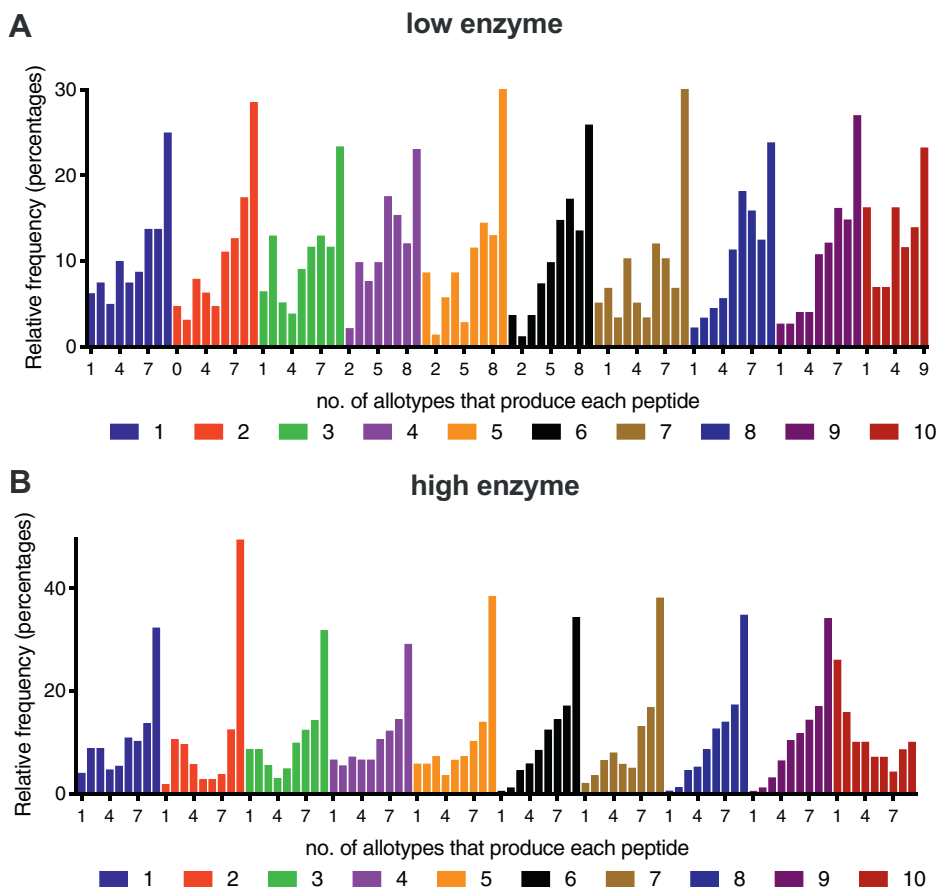
all ERAP1 allotypes are able to generate putative HLA ligands, the exact peptide sequences can be significantly different.

During the last few months, intensive research on the adaptive immune responses toward SARS-CoV-2 has allowed the identification of S1 spike glycoprotein epitopes presented by cells of infected individuals (7, 9, 37–41). To correlate the capacity of different ERAP1 allotypes to generate mature epitopes from this antigen, we compared SARS-CoV-2 S1 spike glycoprotein epitopes deposited in the Immune Epitope Database (<http://www.iedb.org/>) (35) with peptides generated during our *in vitro* digestions. Since even transient generation of an epitope could allow for HLA binding, we combined data from both digestions. We identified 20 epitopes that were produced at least from one allotype, 19 of which had published human HLA restrictions (Table 2). Half of those epitopes were generated by all tested ERAP1 allotypes. However, several epitopes were not produced by some allotypes; most notable allotype 10 did not produce seven of the epitopes. In addition, some epitopes were produced by only a small number of allotypes: epitope SWMESEFRV was produced by allotypes 5 and 10 and epitope NATRFASVY by allotypes 2, 7, and 10. Overall, allotype 10 presented a unique fingerprint in terms of its ability to generate known SARS-CoV-2 epitopes, being the least efficient in generating seven epitopes but at the same time generating two epitopes that were not produced by most other ERAP1 allotypes. Overall, our *in vitro* digestions revealed a significant degree of heterogeneity in producing viral antigen epitopes and highlighted a potential unique fingerprint for allotype 10.

## Discussion

### Generated peptide repertoire by ERAP1 versus HLA-restricted presentation

It is well established that adaptive immune responses are dependent on cell-surface HLA-restricted antigen presentation. Which peptides are presented depend on two major factors: (i) The binding preferences of the HLA alleles expressed by the cell and (ii) the availability of peptides with suitable sequence and length. Binding of peptides onto HLA is well studied, with many thousands of HLA alleles identified to date, having different preferences for peptides (42). In addition, several specialized editing chaperones in the ER help select and optimize peptide binding (43). In contrast, the mechanisms that control the availability of suitable peptides are less understood and depend on the cellular proteome and the proteolytic cascades that sample it. In particular, proteolytic enzymes like ERAP1 have been shown to influence the immunopeptidome of cells, often in profound manners (25). Indeed, ERAP1 SNPs have been shown to both associate with HLA-dependent autoimmunity and cancer (44), often in epistasis with HLA (45), and to functionally affect the peptide products (46, 47). However, functionally relevant ERAP1 SNPs exist in the population in particular combinations, allotypes, and recent analysis suggested that SNPs can synergize to differentiate allotype functional properties (27). Thus, in order to understand and predict antigen presentation, we need to



**Figure 5. Frequency distribution of generated peptides (7–14mers) produced by one or several different ERAP1 allotypes.** The distribution is shown for both reaction conditions and per allotype (1–10). Peptides produced by all ten allotypes are not plotted. ERAP1, endoplasmic reticulum aminopeptidase 1.

understand the interplay and synergisms between allotypic variation in antigen processing and HLA restriction.

### Role of ERAP1 allotypic variation in the biology of antigen presentation

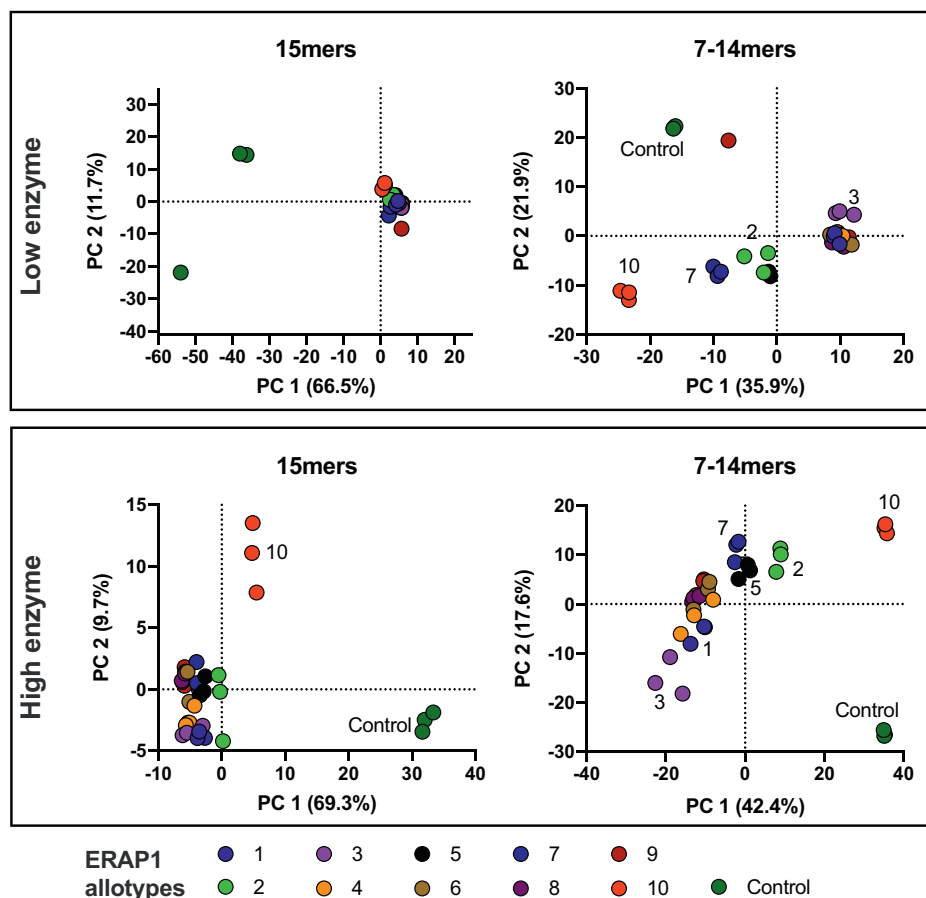
The results presented here clearly suggest that ERAP1 allotypic variation, expressed as ten common ERAP1 allotypes, translates to the generated repertoire of peptides that can be available for HLA binding. While all allotypes have the capacity to produce a common product “core,” they differentiate in their capacity to produce different sets of peptides that have variable affinities for HLA alleles. In this context, ERAP1 allotypes parallel the capacity of HLA for defining the immunopeptidome in qualitative but not absolute patterns and likely complement and synergize with HLA haplotypes to shape the antigen presentation “preferences” of each cell. In this context, a more fundamental biological role for ERAP1 can emerge. While initial studies on ERAP1 highlighted its importance for the generation of particular epitopes (48–50), its role in destroying some antigenic epitopes later emerged as equally important (51, 52). However, more global immunopeptidomic analyses later demonstrated that the effect of ERAP1 can be limited (36, 53, 54) or more focused on antigenic peptide destruction (36, 55). Thus, it is possible that the main

biological function of ERAP1 is not to be an obligate generator of antigenic peptides as initially suspected but rather lies in its allotypic variability. Synergism between ERAP1 trimming specificity and HLA binding preferences could enhance the variability of HLA presentation, thus contributing to the variability of adaptive immune responses in the human population. In this context, the term “ERAP1-dependent and HLA-restricted” may be a more appropriate way to describe the breadth of antigen presentation for a given cell or individual that carry a limited set of ERAP1 allotypes and HLA haplotypes.

### Allotype 10 is not a loss-of-function allele but produces a different fingerprint of peptides

In a previous study, we demonstrated that allotype 10 had a significantly lower enzymatic activity for many substrates tested, and this was due to both reduction in catalytic efficiency as well as substrate affinity (27). That study however was performed with a limited number of substrates analyzed one at a time. Since however ERAP1–substrate interactions can be complex and include competition between substrates and phenomena such as substrate inhibition (31), concurrent analysis of digestion of multiple peptides can provide more accurate insight and better emulate the subcellular

## ERAP1 allotypes affect the generation of SARS-CoV-2 epitopes



**Figure 6. Principal component analysis for the trimming of the initial 15mers and the generation of 7 to 14mers for both reaction conditions.** All biological triplicates (three per allotype) are depicted. The two major components are shown.

environment where ERAP1 normally functions. Indeed, when analyzing the concurrent digestion of hundreds of substrates, allotype 10 was found to be only about 2-fold less enzymatically efficient. Still, allotype 10 was found to produce peptides with distinctly different patterns and constitute a functional outlier compared with the other allotypes, suggesting mechanical differences in substrate selection. Thus, the notion that individuals homozygous for allotype 10 carry just a subactive ERAP1 enzyme should be revised since it appears to only apply to particular substrates. Rather, homozygous individuals appear to carry an ERAP1 variant that exerts different specificity pressures on the peptide repertoire and could result to particular changes in presented peptides.

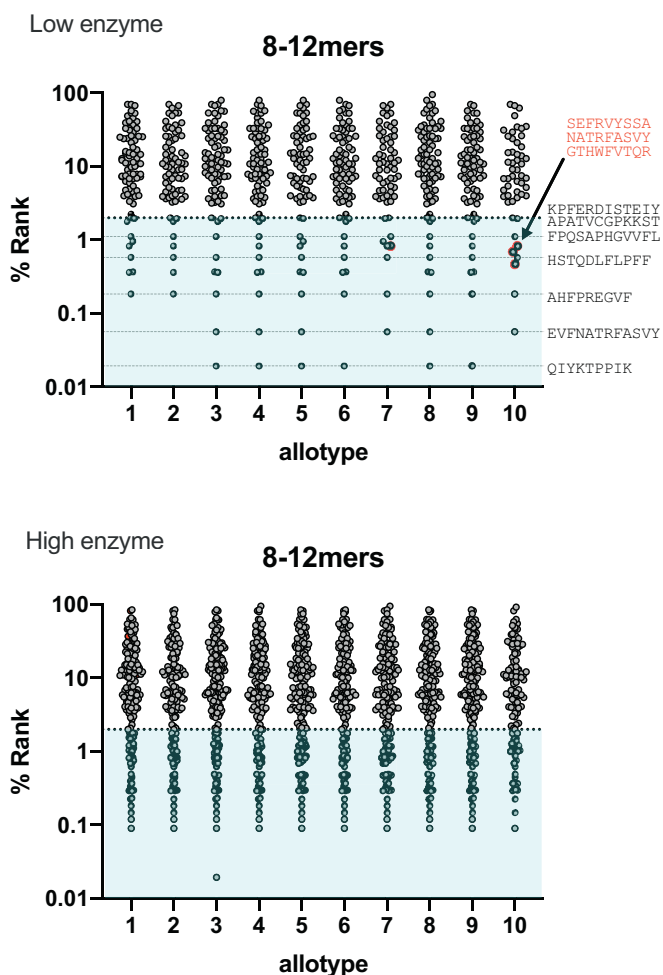
The structural basis of the distinct behavior of allotype 10 is not clear. Some insight however can be derived by comparing the SNPs present in this allotype to the SNPs in other allotypes. In particular, allotype 10 contains a combination of two SNPs that are unique for this allotype, namely Val349 and Gln725. Both these residues were found to either directly or indirectly interact with substrate analogs in recently determined crystal structures of ERAP1 (30). Val349 is located adjacent to the active site and although a conservative substitution (Met to Val) could influence substrate recognition. Gln725 is a nonconservative substitution (Arg to Gln) and lies near the hinge domain of ERAP1 and adjacent to the

C-terminal residue of a 10mer peptide substrate analog that has been crystallized with ERAP1 (30). The SNP Arg725 makes a salt bridge with Asp766, which interacts with the C-terminal Lys residue of that peptide. Furthermore, Gln725 lies in a region that undergoes significant structural reconfigurations when ERAP1 changes conformation from the “open” state to its “closed” state and could thus influence turnover rates (56). It is therefore possible that the combination of those two SNPs in allotype 10 synergizes to affect catalytic rates for different substrates, resulting in an apparent different specificity. Additional structural studies will be necessary to better understand the unique behavior of this ERAP1 allotype.

### ERAP1 allotypic variation and COVID-19

The significant variability of individuals to susceptibility to COVID-19 has been a hallmark of the ongoing pandemic (10). While age appears to be the most important factor, the genetic blueprint of infected individuals also appears to predispose some to severe COVID-19, but the exact genetic factors responsible are poorly understood. Studies have shown that rare mutations as well as polymorphic variability in immune system components can be critical. The role of rare mutations in components of the innate immune response and in particular the interferon I pathway (13, 14) has been described, but





**Figure 7. Scatter plots showing the predicted affinity of produced peptides (8–12mers) by both reaction conditions.** For common HLA alleles (HLA-A\*01:01, HLA-A\*02:01, HLA-A\*03:01, HLA-A\*24:02, HLA-A\*26:01, HLA-B\*07:02, HLA-B\*08:01, HLA-B\*27:05, HLA-B\*39:01, HLA-B\*40:01, HLA-B\*58:01, and HLA-B\*15:01) as calculated by the NetMHCpan 4.1 Server (34).

the exact role of natural polymorphic variation in components of adaptive immunity, such as HLA alleles, is only now emerging, possibly because of difficulties in isolating specific genetic associations from complex polymorphic genetic systems by population studies (15–17). The existence and potential role of ERAP1 allotypes have only recently been recognized (26) and thus have not been studied in the context of COVID-19 susceptibility. Our results, however, suggest that ERAP1 allotypic variation could play, in tandem to HLA alleles, important roles in determining anti-SARS-CoV-2 immune responses and by extension, susceptibility to severe COVID-19. For example, allotype 10 homozygous individuals, which could represent up to 4.8% of the population in Europe (27), could be defective in producing epitopes from the S1 glycoprotein that are presented from particular HLA alleles. Conversely, our analysis suggests that particular antigenic epitopes may be produced by only specific ERAP1 allotypes. Thus, particular combinations of ERAP1 allotypes with HLA alleles could lead to either effective or defective adaptive immune responses against SARS-CoV-2. It is therefore likely important to include the ERAP1 allotypic state in genetic

analyses that focus on HtLA associations with predisposition to severe COVID-19.

### Limitations of the study

While our study highlights the potential importance of ERAP1 allotypes in immune response variability between individuals, it has some important limitations that need to be taken into account when interpreting results. The *in vitro* nature of the digestions may not be an optimal surrogate for cellular antigen processing, either because of missing components (such as other peptidases) or because of cellular compartmentalization. In addition, digestions were analyzed at only a single time point, which although is representative of the kinetics of antigen processing, limits interpretation of sequential generation and destruction of antigenic epitopes. HLA restriction is only indirectly simulated and thus does not take into account kinetic components of peptide competition for binding inside the ER. Furthermore, although antigenic peptide destruction is evident when comparing the two reaction conditions, it is difficult to analyze statistically without an explicit set of 9mer peptides in the initial reaction. Finally, although our approach of concurrently analyzing hundreds of peptide-trimming reactions proved invaluable in identifying difference between trimming patterns of ERAP1 allotypes, the number of peptides used was not sufficient to allow for the explicit identification of sequence motifs in trimming preferences. The latter however may not be readily feasible given the large substrate cavity of ERAP1 that allows for a complex landscape of peptide–enzyme interactions (30).

### Cytotoxic responses after vaccination—different responses depending on ERAP1 allotype?

Although vaccines developed for COVID-19 had as a primary goal the induction of robust antibody responses, the generation of potent and long-lasting cellular adaptive responses will be critical for controlling disease severity and long-term immunity (57–59). Furthermore, some peptide epitopes from SARS-CoV-2 antigens may be immunodominant and shape long-term immunity for a large percentage of individuals. Our results here suggest that the polymorphic variability in peptide epitope generation by ERAP1 allotypes could also play a role, in tandem to HLA restriction, in eliciting and sustaining vaccine-induced cellular immunity. Further studies aiming at defining SARS-CoV-2 immunodominant epitopes shared across individuals will be necessary to test this hypothesis.

### Conclusions

We demonstrate that ERAP1 allotypes common in the population demonstrate significant differences in their ability to process antigenic epitope precursors derived from the S1 spike glycoprotein of SARS-CoV-2. A clear outlier is allotype 10, a common ERAP1 variant (present in more than 20% of Europeans) previously suggested to be subactive, which we find to still be able to generate HLA ligands but with distinct patterns. Our findings suggest that ERAP1 allotypes can be a

## ERAP1 allotypes affect the generation of SARS-CoV-2 epitopes

**Table 2**

List of known SARS-CoV-2 epitopes identified to be produced by the enzymatic digestion of the S1 glycoprotein peptide pool

Sequence	Produced by ERAP1 allotype										Possible MHC restriction	Reference	
	1	2	3	4	5	6	7	8	9	10			
ALNTLVKQL												HLA-A*02:01	(37)
QPYRVVLSF												HLA-B*07:02	(7)
RLDKVEAEV												HLA-A*02:01	(9)
GTHWFVTQR												HLA-A*11:01	(38)
NQKLIANQF												HLA-B*15:01	(39)
TPINLVRDL												HLA-B*07:02	(9)
IIAYTMSL												H2-Kb	(40)
NESLIDLQEL												HLA-B*40:01	(7)
IANQFNSAI												HLA-B*51:01	(7)
LGAENSVAY												HLA-B*35:01	(7)
NATRFASVY												HLA-B*35:01	(7)
NITRFQTL												HLA-B*08:01	(7)
NVYADSFVIR												HLA-A*68:01	(7)
RLQSLQTY												HLA-B*15:01	(7)
TLADAGFIK												HLA-A*03:01	(7)
VYYDPKVF												HLA-A*24:02	(7)
SWMESEFRV												HLA-A*24:02	(9)
VLTESNKKF												HLA-B*15:01	(9)
KPFERDISTEY												HLA-B*40:01	(41)
EVFNATRFASVY												HLA-A*26:01	(7)

The ability of each ERAP1 allotype to produce each of the epitopes is indicated by a *grayed box*.

major contributor to heterogeneity in antigen presentation, and in conjunction with HLA-allele binding specificity, contribute to variable immune responses to disease including COVID-19.

### Experimental procedures

#### Materials

Two PepMix SARS-CoV-2 peptide mixtures were purchased by JBT Peptide Technologies GmbH, dissolved in dimethyl sulfoxide, and stored at  $-80\text{ }^{\circ}\text{C}$ . The two peptide collections (158 and 157 peptides, respectively) were mixed at equimolar concentrations and diluted in buffer containing 10 mM Hepes, pH 7.0, and 100 mM NaCl to a final concentration of 48  $\mu\text{M}$ .

#### Protein expression and purification

Recombinant ERAP1 allotypes 1 to 10 have been described previously (27). Briefly, all ERAP1 variants were produced by insect cell culture after infection with recombinant baculovirus and purified to homogeneity by nickel-nitrilotriacetic acid chromatography and size-exclusion chromatography. Enzymes were stored in aliquots at  $-80\text{ }^{\circ}\text{C}$  with 10% glycerol until needed.

#### Enzymatic reactions

Enzymatic reactions were performed in triplicate in a total volume of 50  $\mu\text{l}$  in 10 mM Hepes, pH 7, and 150 mM NaCl. Freshly thawed enzyme stocks were added to each reaction to

final concentrations of 100 nM or 300 nM (low-enzyme condition and high-enzyme condition, respectively). Reactions were incubated at  $37\text{ }^{\circ}\text{C}$  for 2 h, stopped by the addition of 7.5  $\mu\text{l}$  of a 10% TFA solution, flash frozen in liquid nitrogen, and stored at  $-80\text{ }^{\circ}\text{C}$  until analyzed by LC-MS/MS.

#### LC-MS/MS analysis

Enzymatic reaction samples were directly injected on a PepSep C18 column (250  $\times$  0.75 mm, 1.9  $\mu\text{m}$ ) and separated using a gradient of buffer A (0.1% formic acid in water), 7% buffer B (0.1% formic acid in 80% acetonitrile) to 35% for 40 min followed by an increase to 45% in 5 min and a second increase to 99% in 0.5 min and then kept constant for 4.5 min. The column was equilibrated for 15 min prior to the subsequent injection. A full MS was acquired using a Q Exactive HF-X Hybrid Quadrupole-Orbitrap mass spectrometer, in the scan range of 350 to 1500  $m/z$  using a resolving power of 120 K with an automatic gain control of  $3 \times 10^6$  and maximum injection time of 100 ms, followed by MS/MS scans of the 12 most abundant ions, using a resolving power of 15 K with an automatic gain control of  $1 \times 10^5$  and maximum injection time of 22 ms and a normalized collision energy of 28 and a dynamic exclusion of 30 s.

#### Database search

The generated raw files were processed by the Proteome Discoverer software (Thermo) (version 2.4) using a workflow for the precursor-based quantification, using SequestHT

with Multi Peptide Search and Percolator validation. The Minora algorithm was used for the quantification. The SPIKE\_SARS2.fasta was used as database, and the search was performed in an unspecific mode (no-enzyme specificity was selected). The minimum peptide length was six amino acids. Precursor mass tolerance was 10 ppm, and fragment mass tolerance was 0.02 Da. Fixed Value PSM (peptide-spectrum match) Validator was employed for the validation of the peptides, and only high confident PSMs were used. The peptides identified were filtered based on their unambiguity, a Sequest HT Xcorr above 2, lack of modifications, and a number of PSM above 5.

### Statistical analysis

Three biological replicas of each haplotype were compared against the negative control reaction using the Proteome Discovery pipeline for pairwise ratio protein abundance calculation and the background-based *t* test.

### Data availability

All data described are available in the article and associated supporting information. Numerical values used for generation of graphs are available upon request to the corresponding author (Efstratios Stratikos; E-mail: [stratos@rrp.demokritos.gr](mailto:stratos@rrp.demokritos.gr) or [estratikos@chem.uoa.gr](mailto:estratikos@chem.uoa.gr)). The MS proteomics data have been deposited to the ProteomeXchange Consortium *via* the PRIDE (60) partner repository with the dataset identifier PXD027006 (<http://www.ebi.ac.uk/pride/archive/>).

**Supporting information**—This article contains supporting information.

**Acknowledgments**—We extend our gratitude to GlaxoSmithKline and in particular Drs Semra Kitchen and Jonathan P. Hutchinson for providing recombinant enzymes that allowed this study to be performed in a timely manner (27). This research was supported by internal funds of the National Centre for Scientific Research “Demokritos” and the project “The Greek Research Infrastructure for Personalised Medicine (pMedGR)” (MIS 5002802), which is implemented under the Action “Reinforcement of the Research and Innovation Infrastructure,” funded by the Operational Programme “Competitiveness, Entrepreneurship and Innovation” (NSRF 2014-2020) and cofinanced by Greece and the European Union (European Regional Development Fund).

**Author contributions**—M. S. and E. S. conceptualization; G. S., M. S., I. T., G. P., and E. S. methodology; G. S. and G. P. formal analysis; G. S., M. S., I. T., and E. S. investigation; G. S., M. S., and E. S. data curation; E. S. writing—original draft; G. S., M. S., and G. P. writing—review and editing; G. S. and E. S. visualization; E. S. supervision.

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: COVID-19, coronavirus disease 19; ER, endoplasmic reticulum; ERAP1, endoplasmic reticulum aminopeptidase 1; HLA, human leukocyte antigen; MHC,

major histocompatibility complex; PCA, principal component analysis; PSM, peptide-spectrum match; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

### References

- Brodin, P. (2021) Immune determinants of COVID-19 disease presentation and severity. *Nat. Med.* **1**, 28–33
- St John, A. L., and Rathore, A. P. S. (2020) Early insights into immune responses during COVID-19. *J. Immunol.* **205**, 555–564
- Song, P., Li, W., Xie, J., Hou, Y., and You, C. (2020) Cytokine storm induced by SARS-CoV-2. *Clin. Chim. Acta* **509**, 280–287
- Li, K., Hao, Z., Zhao, X., Du, J., and Zhou, Y. (2020) SARS-CoV-2 infection-induced immune responses: Friends or foes? *Scand. J. Immunol.* **92**, e12895
- Radbruch, A., and Chang, H.-D. (2021) A long-term perspective on immunity to COVID. *Nature* **595**, 359–360
- Leslie, M. T. (2020) Cells found in coronavirus patients “bode well” for long-term immunity. *Science* **368**, 809–810
- Tarke, A., Sidney, J., Kidd, C. K., Dan, J. M., Ramirez, S. I., Yu, E. D., Mateus, J., da Silva Antunes, R., Moore, E., Rubiro, P., Methot, N., Phillips, E., Mallal, S., Frazier, A., Rawlings, S. A., *et al.* (2021) Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. *Cell Rep. Med* **2**, 100204
- Ledford, H. (2021) How “killer” T cells could boost COVID immunity in face of new variants. *Nature* **590**, 374–375
- Saini, S. K., Hersby, D. S., Tamhane, T., Povlsen, H. R., Amaya Hernandez, S. P., Nielsen, M., Gang, A. O., and Hadrup, S. R. (2021) SARS-CoV-2 genome-wide T cell epitope mapping reveals immunodominance and substantial CD8+ T cell activation in COVID-19 patients. *Sci. Immunol.* **6**, eabf7550
- Pereira, N. L., Ahmad, F., Byku, M., Cummins, N. W., Morris, A. A., Owens, A., Tuteja, S., and Cresci, S. (2021) COVID-19: Understanding inter-individual variability and implications for precision medicine. *Mayo Clin. Proc.* **96**, 446–463
- Fricke-Galindo, I., and Falfán-Valencia, R. (2021) Genetics insight for COVID-19 susceptibility and severity: A review. *Front. Immunol.* **12**, 622176
- Zeberg, H., and Pääbo, S. (2020) The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. *Nature* **587**, 610–612
- Zhang, Q., Bastard, P., Liu, Z., Le Pen, J., Moncada-Velez, M., Chen, J., Ogishi, M., Sabli, I. K. D., Hodeib, S., Korol, C., Rosain, J., Bilguvar, K., Ye, J., Bolze, A., Bigio, B., *et al.* (2020) Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **370**, eabd4570
- Lee, J. S., and Shin, E. C. (2020) The type I interferon response in COVID-19: Implications for treatment. *Nat. Rev. Immunol.* **20**, 585–586
- Shkurnikov, M., Nersisyan, S., Jankevic, T., Galatenko, A., Gordeev, I., Vechorko, V., and Tonevitsky, A. (2021) Association of HLA class I genotypes with severity of coronavirus disease-19. *Front. Immunol.* **12**, 641900
- Saulle, I., Vicentini, C., Clerici, M., and Biasin, M. (2021) Antigen presentation in SARS-CoV-2 infection: The role of class I HLA and ERAP polymorphisms. *Hum. Immunol.* **82**, 551–560
- Secolin, R., de Araujo, T. K., Gonsales, M. C., Rocha, C. S., Naslavsky, M., Marco, L. De, Bicalho, M. A. C., Vazquez, V. L., Zatz, M., Silva, W. A., and Lopes-Cendes, I. (2021) Genetic variability in COVID-19-related genes in the Brazilian population. *Hum. Genome* **8**, 15
- Rock, K. L., and Goldberg, A. L. (1999) Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annu. Rev. Immunol.* **17**, 739–779
- Weimershaus, M., Evnouchidou, I., Saveanu, L., and van Endert, P. (2013) Peptidases trimming MHC class I ligands. *Curr. Opin. Immunol.* **25**, 90–96
- Hammer, G. E., Kanaseki, T., and Shastri, N. (2007) The final touches make perfect the peptide-MHC class I repertoire. *Immunity* **26**, 397–406
- Reeves, E., Wood, O., Ottensmeier, C. H., King, E. V., Thomas, G. J., Elliott, T., and James, E. (2019) HPV epitope processing differences correlate with ERAP1 allotype and extent of CD8(+) T-cell tumor infiltration in OPSCC. *Cancer Immunol. Res.* **7**, 1202–1213

## ERAP1 allotypes affect the generation of SARS-CoV-2 epitopes

22. Kim, S., Lee, S., Shin, J., Kim, Y., Evnouchidou, I., Kim, D., Kim, Y.-K., Kim, Y.-E., Ahn, J.-H., Riddell, S. R., Stratikos, E., Kim, V. N., and Ahn, K. (2011) Human cytomegalovirus MicroRNA MiR-US4-1 inhibits CD8<sup>+</sup> T cell responses by targeting the aminopeptidase ERAP1. *Nat. Immunol.* **12**, 984–991
23. Robinson, J., Barker, D. J., Georgiou, X., Cooper, M. A., Flicek, P., and Marsh, S. G. E. (2020) IPD-IMGT/HLA database. *Nucleic Acids Res.* **48**, D948–D955
24. Kirino, Y., Bertsias, G., Ishigatsubo, Y., Mizuki, N., Tugal-Tutkun, I., Seyahi, E., Ozyazgan, Y., Sacli, F. S., Erer, B., Inoko, H., Emrence, Z., Cakar, A., Abaci, N., Ustek, D., Satorius, C., et al. (2013) Genome-wide association analysis identifies new susceptibility loci for Behcet's disease and epistasis between HLA-B\*51 and ERAP1. *Nat. Genet.* **45**, 202–207
25. López de Castro, J. A. (2018) How ERAP1 and ERAP2 shape the peptidomes of disease-associated MHC-I proteins. *Front. Immunol.* **9**, 2463
26. Ombrello, M. J., Kastner, D. L., and Remmers, E. F. (2015) Endoplasmic reticulum-associated amino-peptidase 1 and rheumatic disease: Genetics. *Curr. Opin. Rheumatol.* **27**, 349–356
27. Hutchinson, J. P., Temponeras, I., Kuiper, J., Cortes, A., Korczynska, J., Kitchen, S., and Stratikos, E. (2021) Common allotypes of ER aminopeptidase 1 have substrate-dependent and highly variable enzymatic properties. *J. Biol. Chem.* **296**, 100443
28. Reeves, E., Edwards, C. J., Elliott, T., and James, E. (2013) Naturally occurring ERAP1 haplotypes encode functionally distinct alleles with fine substrate specificity. *J. Immunol.* **191**, 35–43
29. Roberts, A. R., Appleton, L. H., Cortes, A., Vecellio, M., Lau, J., Watts, L., Brown, M. A., and Wordsworth, P. (2017) ERAP1 association with ankylosing spondylitis is attributable to common genotypes rather than rare haplotype combinations. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 558–561
30. Giastas, P., Mpakali, A., Papakyriakou, A., Lelis, A., Kokkala, P., Neu, M., Rowland, P., Liddle, J., Georgiadis, D., and Stratikos, E. (2019) Mechanism for antigenic peptide selection by endoplasmic reticulum aminopeptidase 1. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 26709–26716
31. Evnouchidou, I., Kamal, R. P., Seregin, S. S., Goto, Y., Tsujimoto, M., Hattori, A., Voulgari, P. V., Drosos, A. A., Amalfitano, A., York, I. A., and Stratikos, E. (2011) Coding single nucleotide polymorphisms of endoplasmic reticulum aminopeptidase 1 can affect antigenic peptide generation *in vitro* by influencing basic enzymatic properties of the enzyme. *J. Immunol.* **186**, 1909–1913
32. Stamatakis, G., Samiotaki, M., Mpakali, A., Panayotou, G., and Stratikos, E. (2020) Generation of SARS-CoV-2 S1 spike glycoprotein putative antigenic epitopes *in vitro* by intracellular aminopeptidases. *J. Proteome Res.* **19**, 4398–4406
33. Chang, S. C., Momburg, F., Bhutani, N., and Goldberg, A. L. (2005) The ER aminopeptidase, ERAP1, trims precursors to lengths of MHC class I peptides by a “molecular ruler” mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 17107–17112
34. Reynisson, B., Alvarez, B., Paul, S., Peters, B., and Nielsen, M. (2020) NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **48**, W449–W454
35. Vita, R., Mahajan, S., Overton, J. A., Dhanda, S. K., Martini, S., Cantrell, J. R., Wheeler, D. K., Sette, A., and Peters, B. (2019) The immune epitope database (IEDB): 2018 update. *Nucleic Acids Res.* **47**, D339–D343
36. Admon, A. (2019) ERAP1 shapes just part of the immunopeptidome. *Hum. Immunol.* **80**, 296–301
37. Shomuradova, A. S., Vagida, M. S., Sheetikov, S. A., Zornikova, K. V., Kiryukhin, D., Titov, A., Peshkova, I. O., Khmelevskaya, A., Dianov, D. V., Malasheva, M., Shmelev, A., Serdyuk, Y., Bagaev, D. V., Pivnyuk, A., Shcherbinin, D. S., et al. (2020) SARS-CoV-2 epitopes are recognized by a public and diverse repertoire of human T cell receptors. *Immunity* **53**, 1245–1257.e5
38. Kared, H., Redd, A. D., Bloch, E. M., Bonny, T. S., Sumatoh, H., Kairi, F., Carbajo, D., Abel, B., Newell, E. W., Bettinotti, M. P., Benner, S. E., Patel, E. U., Littlefield, K., Laeyendecker, O., Shoham, S., et al. (2021) SARS-CoV-2-Specific CD8<sup>+</sup> T cell responses in convalescent COVID-19 individuals. *J. Clin. Invest.* **131**, e145476
39. Minervina, A. A., Komech, E. A., Titov, A., Koraichi, M. B., Rosati, E., Mamedov, I. Z., Franke, A., Efimov, G. A., Chudakov, D. M., Mora, T., Walczak, A. M., Lebedev, Y. B., and Pogorely, M. V. (2021) Longitudinal high-throughput Tcr repertoire profiling reveals the dynamics of t-cell memory formation after mild covid-19 infection. *Elife* **10**, 1–17
40. Hassert, M., Geerling, E., Stone, E. T., Steffen, T. L., Feldman, M. S., Dickson, A. L., Class, J., Richner, J. M., Brien, J. D., and Pinto, A. K. (2020) MRNA induced expression of human angiotensin-converting enzyme 2 in mice for the study of the adaptive immune response to severe acute respiratory Syndrome coronavirus 2. *PLoS Pathog.* **16**, e1009163
41. [preprint] Snyder, T. M., Gittelman, R. M., Klingler, M., May, D. H., Osborne, E. J., Taniguchi, R., Zahid, H. J., Kaplan, I. M., Dines, J. N., Noakes, M. N., Pandya, R., Chen, X., Elasad, S., Robins, H. S., et al. (2020) Magnitude and dynamics of the T-cell response to SARS-CoV-2 infection at both individual and population levels. *MedRxiv*. <https://doi.org/10.1101/2020.07.31.20165647>
42. Robinson, J., Halliwell, J. A., Hayhurst, J. D., Flicek, P., Parham, P., and Marsh, S. G. E. (2015) The IPD and IMGT/HLA database: Allele variant databases. *Nucleic Acids Res.* **43**, D423–D431
43. Thomas, C., and Tampe, R. (2019) MHC I chaperone complexes shaping immunity. *Curr. Opin. Immunol.* **58**, 9–15
44. Reeves, E., Islam, Y., and James, E. (2020) ERAP1: A potential therapeutic target for a myriad of diseases. *Expert Opin. Ther. Targets* **24**, 535–544
45. Cortes, A., Pulit, S. L., Leo, P. J., Pointon, J. J., Robinson, P. C., Weisman, M. H., Ward, M., Gensler, L. S., Zhou, X., Garchon, H. J., Chiochia, G., Nossent, J., Lie, B. A., Forre, O., Tuomilehto, J., et al. (2015) Major histocompatibility complex associations of ankylosing spondylitis are complex and involve further epistasis with ERAP1. *Nat. Commun.* **6**, 7146
46. Alvarez-Navarro, C., and Lopez de Castro, J. A. (2014) ERAP1 structure, function and pathogenetic role in ankylosing spondylitis and other MHC-associated diseases. *Mol. Immunol.* **57**, 12–21
47. Stratikos, E., Stamogiannos, A., Zervoudi, E., and Fruci, D. (2014) A role for naturally occurring alleles of endoplasmic reticulum aminopeptidases in tumor immunity and cancer predisposition. *Front. Oncol.* **4**, 363
48. Serwold, T., Gonzalez, F., Kim, J., Jacob, R., and Shastri, N. (2002) ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature* **419**, 480–483
49. Saric, T., Chang, S. C., Hattori, A., York, I. A., Markant, S., Rock, K. L., Tsujimoto, M., and Goldberg, A. L. (2002) An IFN-Gamma-Induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat. Immunol.* **3**, 1169–1176
50. Hammer, G. E., Gonzalez, F., James, E., Nolla, H., and Shastri, N. (2007) In the absence of aminopeptidase ERAAP, MHC class I molecules present many unstable and highly immunogenic peptides. *Nat. Immunol.* **8**, 101–108
51. Keller, M., Ebstein, F., Bürger, E., Textoris-Taube, K., Gorny, X., Urban, S., Zhao, F., Dannenberg, T., Sucker, A., Keller, C., Saveanu, L., Krüger, E., Rothkötter, H. J., Dahlmann, B., Henklein, P., et al. (2015) The proteasome immunosubunits, PA28 and ER-aminopeptidase 1 protect melanoma cells from efficient MART-126-35-specific T-cell recognition. *Eur. J. Immunol.* **45**, 3257–3268
52. James, E., Bailey, I., Sugiyarto, G., and Elliott, T. (2013) Induction of protective antitumor immunity through attenuation of ERAAP function. *J. Immunol.* **190**, 5839–5846
53. Barnea, E., Melamed Kadosh, D., Haimovich, Y., Satumira, N., Dorris, M. L., Nguyen, M. T., Hammer, R. E., Tran, T. M., Colbert, R. A., Taurog, J. D., and Admon, A. (2017) The human Leukocyte antigen (HLA)-B27 peptidome *in vivo*, in spondyloarthritis-susceptible HLA-B27 transgenic rats and the effect of Erap1 deletion. *Mol. Cell. Proteomics* **16**, 642–662
54. Koumantou, D., Barnea, E., Martin-Esteban, A., Maben, Z., Papakyriakou, A., Mpakali, A., Kokkala, P., Pratsinis, H., Georgiadis, D., Stern, L. J. L., Admon, A., and Stratikos, E. (2019) Editing the immunopeptidome of melanoma cells using a potent inhibitor of endoplasmic reticulum aminopeptidase 1 (ERAP1). *Cancer Immunol. Immunother.* **68**
55. Komov, L., Kadosh, D. M., Barnea, E., Milner, E., Hender, A., and Admon, A. (2018) Cell surface MHC class I expression is limited by the

- availability of peptide-receptive “empty” molecules rather than by the supply of peptide ligands. *Proteomics* **18**, e1700248
56. Stratikos, E., and Stern, L. J. (2013) Antigenic peptide trimming by ER aminopeptidases—insights from structural studies. *Mol. Immunol.* **55**, 212–219
  57. Ewer, K. J., Barrett, J. R., Belij-Rammerstorfer, S., Sharpe, H., Makinson, R., Morter, R., Flaxman, A., Wright, D., Bellamy, D., Bittaye, M., Dold, C., Provine, N. M., Aboagye, J., Fowler, J., Silk, S. E., *et al.* (2021) T cell and antibody responses induced by a single dose of ChAdOx1 NCoV-19 (AZD1222) vaccine in a phase 1/2 clinical trial. *Nat. Med.* **27**, 270–278
  58. Sahin, U., Muik, A., Derhovanessian, E., Vogler, I., Kranz, L. M., Vormehr, M., Baum, A., Pascal, K., Quandt, J., Maurus, D., Brachtendorf, S., Lörks, V., Sikorski, J., Hilker, R., Becker, D., *et al.* (2020) COVID-19 vaccine BNT162b1 elicits human antibody and TH1 T cell responses. *Nature* **586**, 594–599
  59. Reynolds, C. J., Pade, C., Gibbons, J. M., Butler, D. K., Otter, A. D., Menacho, K., Fontana, M., Smit, A., Sackville-West, J. E., Cutino-Moguel, T., Maini, M. K., Chain, B., Noursadeghi, M., Brooks, T., Semper, A., *et al.* (2021) Prior SARS-CoV-2 infection rescues B and T cell responses to variants after first vaccine dose. *Science* **80**, eabh1282
  60. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., Perez, E., Uszkoreit, J., Pfeuffer, J., Sachsenberg, T., Yilmaz, S., *et al.* (2019) The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450