



Common allotypes of ER aminopeptidase 1 have substrate-dependent and highly variable enzymatic properties

Received for publication, December 1, 2020, and in revised form, February 11, 2021. Published, Papers in Press, February 20, 2021,

<https://doi.org/10.1016/j.jbc.2021.100443>

Jonathan P. Hutchinson^{1,‡}, Ioannis Temponeras^{2,‡}, Jonas Kuiper^{3,4} , Adrian Cortes⁵, Justyna Korczynska¹, Semra Kitchen⁵, and Efstratios Stratikos^{2,7,*}

From the ¹Medicinal Science and Technology, GlaxoSmithKline, Stevenage, Hertfordshire, UK; ²Protein Chemistry Laboratory, National Centre for Scientific Research “Demokritos”, Athens, Greece; ³Department of Ophthalmology, ⁴Center for Translational Immunology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands; ⁵Human Genetics, ⁶Adaptive Immunity Research Unit, GlaxoSmithKline, Stevenage, Hertfordshire, UK; ⁷Laboratory of Biochemistry, Department of Chemistry, National and Kapodistrian University of Athens, Panepistimiopolis Zografou, Athens, Greece

Edited by Peter Cresswell

Polymorphic variation of immune system proteins can drive variability of individual immune responses. Endoplasmic reticulum aminopeptidase 1 (ERAP1) generates antigenic peptides for presentation by major histocompatibility complex class I molecules. Coding SNPs in *ERAP1* have been associated with predisposition to inflammatory rheumatic disease and shown to affect functional properties of the enzyme, but the interplay between combinations of these SNPs as they exist in allotypes has not been thoroughly explored. We used phased genotype data to estimate ERAP1 allotype frequency in 2504 individuals across five major human populations, generated highly pure recombinant enzymes corresponding to the ten most common ERAP1 allotypes, and systematically characterized their *in vitro* enzymatic properties. We find that ERAP1 allotypes possess a wide range of enzymatic activities, up to 60-fold, whose ranking is substrate dependent. Strikingly, allotype 10, previously associated with Behçet’s disease, is consistently a low-activity outlier, suggesting that a significant percentage of individuals carry a subactive *ERAP1* gene. Enzymatic analysis revealed that ERAP1 allotypes can differ in both catalytic efficiency and substrate affinity, differences that can change intermediate accumulation in multistep trimming reactions. Alterations in efficacy of an allosteric inhibitor that targets the regulatory site suggest that allotypic variation influences the communication between the regulatory and the active site. Our work defines the wide landscape of ERAP1 activity in human populations and demonstrates how common allotypes can induce substrate-dependent variability in antigen processing, thus contributing, in synergy with major histocompatibility complex haplotypes, to immune response variability and predisposition to chronic inflammatory conditions.

Major histocompatibility complex (MHC) molecules (human leukocyte antigens [HLAs] in humans) are the most

polymorphic human genes with tens of thousands of different allomorphs identified to date (1). MHC class I (MHCI) molecules bind small protein fragments (peptides) that originate from normal cellular proteins or pathogen proteins and then translocate to the cell surface to present their cargo to cytotoxic T-lymphocytes (2). Polymorphic variation in MHCI predominantly affects the structure of the binding groove and allows the presentation of a large variety of peptide sequences.

MHCIs bind their peptide cargo in the endoplasmic reticulum (ER) with the assistance of the peptide-loading complex (3). While MHCI tend to bind peptides that are between 8 and 11 amino acids long (the majority of which are 9mers), many peptides that enter the ER can be substantially longer (4). Two ER-resident aminopeptidases, ER aminopeptidase 1 and ER aminopeptidase 2 (ERAP1 and ERAP2), catalytically process these precursor peptides and define the peptide pool that is available for binding onto MHCI (5).

The *ERAP1* gene is also polymorphic, and a variety of coding SNPs confer susceptibility to human disease, most notably chronic inflammatory conditions, often in epistasis with HLA class I alleles, which emphasize the critical role of ERAP1 in antigen presentation (6–9). The genetic association of inflammatory diseases, such as *HLA-B27-associated ankylosing spondylitis* (AS), *HLA-B51-associated Behçet’s disease*, and *HLA-A29-associated birdshot uveitis*, led to the hypothesis that these conditions are driven by pathogenic changes in antigen presentation as a direct result of alterations in substrate preferences or activity of ERAP1 (10–13). Several ERAP1 SNPs have been described to affect the function of the enzyme (14, 15). Mechanisms proposed to underlie this effect include direct interactions with the substrate (16), effects on conformational dynamics (17), protein expression level (18, 19), or combinations of these (10). However, not all the possible combinations of the nine most common coding SNPs (*i.e.*, allotypes) (20) occur at equal frequency in the human population (9). Rather, these SNPs encode a limited palette of allotypes that are maintained at high frequency (>1%) in populations, which suggests functional asymmetry between ERAP1 allotypes. This is supported by the fact that some ERAP1 allotypes are

[‡] These authors contributed equally to the manuscript.

* For correspondence: Efstratios Stratikos, estratikos@chem.uoa.gr, stratos@rrp.demokritos.gr.

Enzymatic properties of common ERAP1 allotypes

protective, whereas others confer risk to inflammatory diseases (21). A deep understanding of the functional properties of ERAP1 allotypes rather than individual SNPs is critical to unraveling their physiological impact on disease.

Previous studies have described several ERAP1 allotypes. Most studies defined ERAP1 allotypes as the combination of nine coding SNPs at amino acid positions 56, 127, 276, 346, 349, 528, 575, 725, and 730. Ombrello *et al.* (9) reported 10 common ERAP1 allotypes in three populations of European and East Asian Ancestry (n = 160). Reeves *et al.* reported some additional distinct allotypes discovered in small patient cohorts of AS (n = 72) (15) and oropharyngeal squamous cell carcinoma (n = 25) (22). These allotypes however have been controversial and proposed to be rare by others (23, 24). While these studies have contributed to our understanding of the emerging role of ERAP1 allotypes in disease, controversy remains on which ERAP1 allotypes are common, and systematic analysis of their functional differences has been lacking.

We used phased genotype data from the 1000 Genomes Project (25) to define common ERAP1 allotypes in 2504 individuals of five major human populations. We generated recombinant versions of the ten most common allotypes and comprehensively characterized their *in vitro* enzymatic properties. We find a complex landscape of large substrate-dependent enzymatic activity differences between allotypes because of effects on catalytic efficiency and substrate affinity. Our findings suggest that ERAP1 allotypic variation has the potential to strongly synergize with MHCII alleles, in an epitope-dependent manner, to enhance immune system variability in natural human populations.

Results

Analysis of the Genome Aggregation Database that contains 125,748 exome sequences (26) using a 1% frequency cutoff to qualify a coding genomic missense variant as a polymorphism revealed only ten amino acid positions as polymorphic, namely 12, 56, 127, 276, 346, 349, 528, 575, 725, and 730, consistent

with a previous study (9). Since however position 12 lies in the signal peptide that is normally excised after translocation of ERAP1 into the ER and thus does not appear in the mature protein, we focused our analysis on the remaining nine positions. These nine SNPs could be theoretically organized in up to 2⁹ discreet allotypes. To define which ERAP1 allotypes are common in human populations, we exploited available phased (*i.e.*, ordered along one chromosome) genotype data from 5q15 of the 1000 Genomes Project phase 3 (27). The frequency of the nine ERAP1 SNPs in different populations is shown in Table S1. Correlations between individual SNPs that indicate linkage disequilibrium are shown in Table S2 and are generally consistent with previous studies (9). The population frequency of the most common ERAP1 allotypes is shown in Table 1. An analogous analysis using data from the UK biobank revealed highly similar results (Table S3) (28). Strikingly, although ten common allotypes constitute 99.9% of all allotypes in the European population, some populations have additional allotypes not reported before. Overall, we were able to identify at least six additional allotypes that have frequencies of over 0.5% in at least one population (Table 1; allotypes 11–16). Our analysis not only confirmed previous results in a larger setting but also revealed that there is significant population variability between ERAP1 allotypes. Regardless, given the extensive use of allotypes 1 to 10 in the literature and their near-complete coverage of the global population (>94% globally, >99.9% in the European population), we proceeded with the functional characterization of these allotypes. Since individuals carry two copies of the *ERAP1* gene, we also analyzed the combinations of allotypes present in the 2504 samples (Fig. 1A and Tables S4 and S5). The most common combination was that of allotypes 8 and 2 followed by the 2-2 homozygous, which cumulatively account for almost 20% of the global population. Interestingly, the combination 8-2 was found to be about twice as frequent (11%) than predicted from a random distribution (25.6% × 21.8% = 5.6%). We in addition analyzed the prevalence of two SNPs in the homologous ERAP2, namely rs2549782 and

Table 1
List of ERAP1 allotypes and their frequency in different populations

Allotype	Amino acid composition at each SNP									Population allotype frequency (%)					
	56	127	276	346	349	528	575	725	730	ALL	AFR	AMR	EUR	EAS	SAS
1	E	P	I	G	M	K	D	R	Q	6.1	0.5	12.1	14.0	0.0	7.6
2	E	R	I	G	M	K	D	R	Q	25.6	28.1	22.6	14.4	42.6	18.4
3	E	R	I	G	M	K	D	R	E	8.5	11.4	5.8	6.6	2.7	14.5
4	E	R	I	G	M	R	D	R	E	3.2	3.0	8.9	1.0	2.4	2.4
5	E	R	I	D	M	R	D	R	E	6.2	8.4	3.9	6.3	0.0	10.9
6	E	P	I	G	M	R	D	R	E	5.8	4.0	10.2	9.2	0.7	6.5
7	K	P	I	G	M	R	D	R	E	7.2	2.9	8.5	4.1	17.1	5.0
8	E	P	M	G	M	R	D	R	E	21.8	17.4	14.0	21.5	28.5	26.7
9	E	P	M	G	M	R	N	R	E	0.1	0.0	0.4	0.4	0.0	0.0
10	E	P	I	G	V	R	N	Q	E	9.6	3.5	12.4	22.4	5.8	6.5
Subtotal										94.1	79.2	98.8	99.9	99.8	98.5
11	E	P	I	D	M	R	D	R	E	2.7	9.9	0.7	0.1	0.0	0.0
12	K	P	I	D	M	R	D	R	E	1.4	5.2	0.1	0.0	0.0	0.0
13	E	P	I	G	V	R	N	R	E	0.8	3.0	0.1	0.0	0.0	0.0
14	E	P	M	G	M	R	N	Q	E	0.4	1.7	0.0	0.0	0.0	0.0
15	E	R	M	G	M	R	D	R	E	0.2	0.5	0.0	0.0	0.1	0.1
16	E	P	M	G	M	R	D	R	Q	0.12	0.0	0.0	0.0	0.0	0.6
Subtotal										5.6	20.3	0.9	0.1	0.1	0.7
Total										99.7	99.5	99.7	100	99.9	99.2

AFR, African; ALL, all populations cumulatively; AMR, American; EAS, East Asian; EUR, European; SAS, South Asian. Bold indicates membership in the ALL population. Italics indicates subtotals and totals.

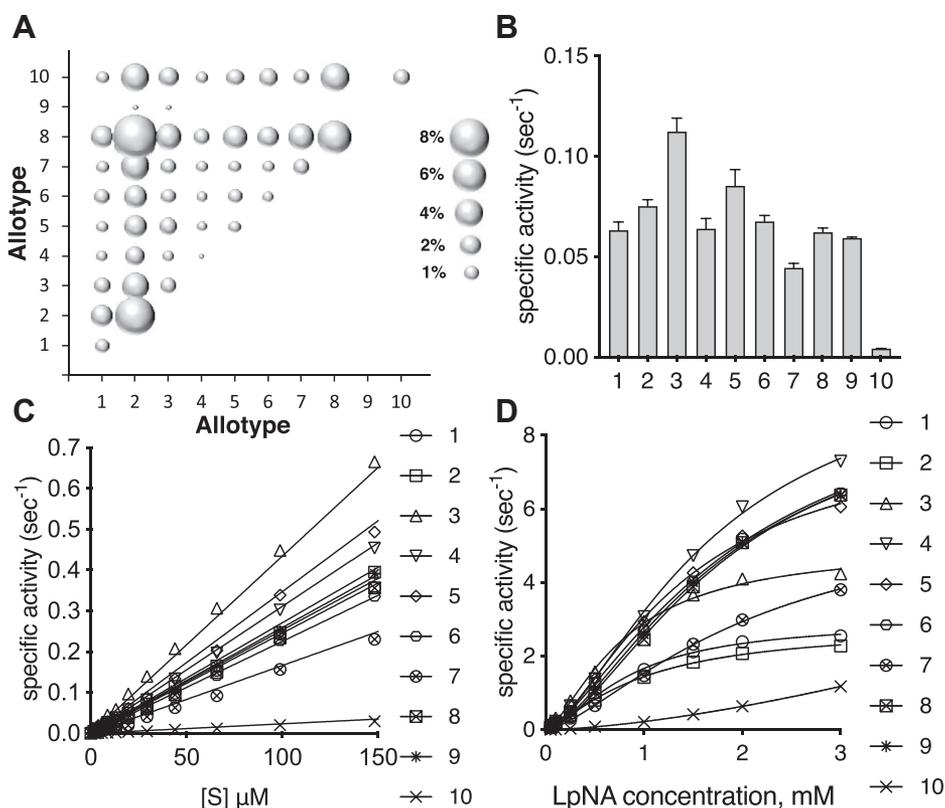


Figure 1. Population distribution and enzymatic activity of common ERAP1 allotypes. A, bubble chart showing the relative prevalence of endoplasmic reticulum aminopeptidase 1 allotype combinations in the global population based on data from 1000 genomes. B, specific activity of endoplasmic reticulum aminopeptidase 1 allotypes 1 to 10 for the hydrolysis of dipeptidic substrate L-leucine-7-amido-4-methylcoumarin. C, specific activity of L-leucine-7-amido-4-methylcoumarin hydrolysis versus substrate concentration. D, Michaelis–Menten analysis of L-leucine-para-nitroanilide hydrolysis by each allotype.

rs2248374 (29, 30). Of the four possible combinations, we were able to detect only two in our population sample, the [G,A] allotype (defined as allotype A in (30)) in 44.7% of the samples and the [T,G] allotype (defined as allotype B in (30)) in 55.3% of the samples. Since however the [G] allele of the rs2248374 leads to no detectable expression of ERAP2 (30), our results suggest that 31.3% of individuals in the database carry two [G] allotypes and should therefore have no ERAP2 protein expression. We have previously shown that the co-occurrence of functional ERAP2 protein is dependent on the ERAP1 haplotype background (10) in Dutch and Spanish controls; therefore, we were interested to see if the ERAP1 and ERAP2 allotypes correlate in this global data set. We calculated the distribution of ERAP1 allotypes in each group of individuals carrying a particular ERAP2 allotype (Table S6), and this revealed some interesting correlations. The [8,8] ERAP1 combination was found in 19.2% of individuals that carry the [A,A] ERAP2 allotype, which is much more frequent compared with the whole population (5.8%). In addition, the [2,2] ERAP1 allotype combination was found in 22.3% of individuals carrying the [B,B] ERAP2 allotype (8.2% in the whole population). The correlation of these specific combinations of ERAP1 and ERAP2 allotypes may be functionally relevant as discussed later in this article.

To characterize the enzymatic properties of the common ERAP1 allotypes, we generated recombinant ERAP1 variants corresponding to each allotype as listed in Table 1. The SNPs that define allotypes 1 to 10 are scattered throughout the

structure of ERAP1, away from the catalytic center (Fig. S1) and can be generally categorized into two groups: (a) SNPs that lie on the outside of the central cavity of the enzyme that normally accommodates the substrate (Fig. S1A) and (b) SNPs that line the interior surface of this cavity and may make direct interactions with the peptide substrate (Fig. S1B).

We first characterized the ERAP1 allotypes using well-established small dipeptide substrates. The specific activity for the hydrolysis of the substrate L-leucine-7-amido-4-methylcoumarin (Leu-AMC) is shown in Figure 1B. There was approximately a twofold spread in specific activities for allotypes 1 to 9, but allotype 10 was found to be at least 10-fold less active (Fig. 1B and Fig. S2). The relationship between specific activity and substrate concentration was found to be linear up to 150 μM substrate, which allowed the calculation of the $k_{\text{cat}}/K_{\text{M}}$ ratio for Leu-AMC (Fig. 1C and Table S7). Allotype 3 was found to be the most active of all, and allotype 10 was 18-fold less active. To obtain full Michaelis–Menten (MM) analysis, we employed the chromogenic dipeptide substrate L-leucine-para-nitroanilide (Leu-pNA) (31). Data fit best to an allosteric MM model as previously demonstrated (31, 32) allowing us to calculate the enzymatic parameters (Fig. 1D and Table S7). This analysis demonstrated that the changes in specific activity between allotypes are both because of changes in affinity for the substrate (K_{half}) and to changes in maximal catalytic efficiency (V_{max}). Similar to Leu-AMC, allotype 10 was much less active in hydrolyzing Leu-pNA, which

Enzymatic properties of common ERAP1 allotypes

unfortunately precluded reliable calculation of K_{half} and V_{max} for this allotype.

Since ERAP1 trims long N-terminally extended peptide precursors of antigenic peptides, we turned to more physiologically relevant longer peptides. Recently determined cocrystal structures of ERAP1 with bound 15mer and 10mer peptide analogs revealed that the peptides are processed in a large internal cavity, while making interactions with residues that line that cavity, which can drive selectivity (16). We measured the rate of N-terminus hydrolysis of two peptides of a similar backbone sequence as the cocrystallized analogs, namely the 15mer LLRIQRGPGRFVTI and the 10mer LLKHHAFSFK (Fig. 2, A and B). Like the results with the small substrates, we recorded a significant variation in trimming rates. Although allotype 10 was again the least efficient, processing rate differences were less pronounced compared with smaller substrates. Notably, allotype 10 was about as efficient as allotype 9 in trimming the 15mer. Interestingly, the pattern between the two peptides was significantly different. Allotypes 4 and 6 trimmed the 15mer the fastest, whereas allotypes 5 and 7 trimmed the 10mer the fastest. These results are consistent with a complex landscape of peptide–enzyme interactions that drive specificity and suggest that the effect of allotype variation may be substrate dependent (16, 33).

To better understand the mechanism behind the variation of trimming rates, we utilized a recently developed assay suitable for MM analysis that follows the trimming of a 9mer antigenic epitope with the sequence YTAFTIPSI (Fig. 2C) (34). Using this assay, we calculated k_{cat} and K_{M} for each ERAP1 allotype (Table S8). The turnover rate (k_{cat}), a measure of the maximal catalytic rate of hydrolysis, varied by only about twofold between allotypes 1 and 9 but was about 10-fold reduced for allotype 10, indicating that this allotype is catalytically deficient. The Michaelis constant (K_{M}), a measure of how well the enzyme can recognize the substrate, varied between allotypes by up to sixfold with allotypes 1 and 2 having the highest affinity for the substrate. The ratio $k_{\text{cat}}/K_{\text{M}}$, a measure of the overall catalytic efficiency of the enzyme, varied up to 60-fold, indicating that changes in k_{cat} and K_{M} can synergize to enhance differences in trimming rates for particular substrates (Fig. 2D). Notably, allotype 10 had a 60-fold lower catalytic efficiency compared with allotype 2 because of the combined effect of lower substrate affinity and lower turnover rate. Thus, we conclude that polymorphic variation in ERAP1 can influence peptide trimming rates by affecting both catalytic efficiency and substrate recognition.

ERAP1 trimming of antigenic epitope precursors in the ER often includes multiple trimming steps and which

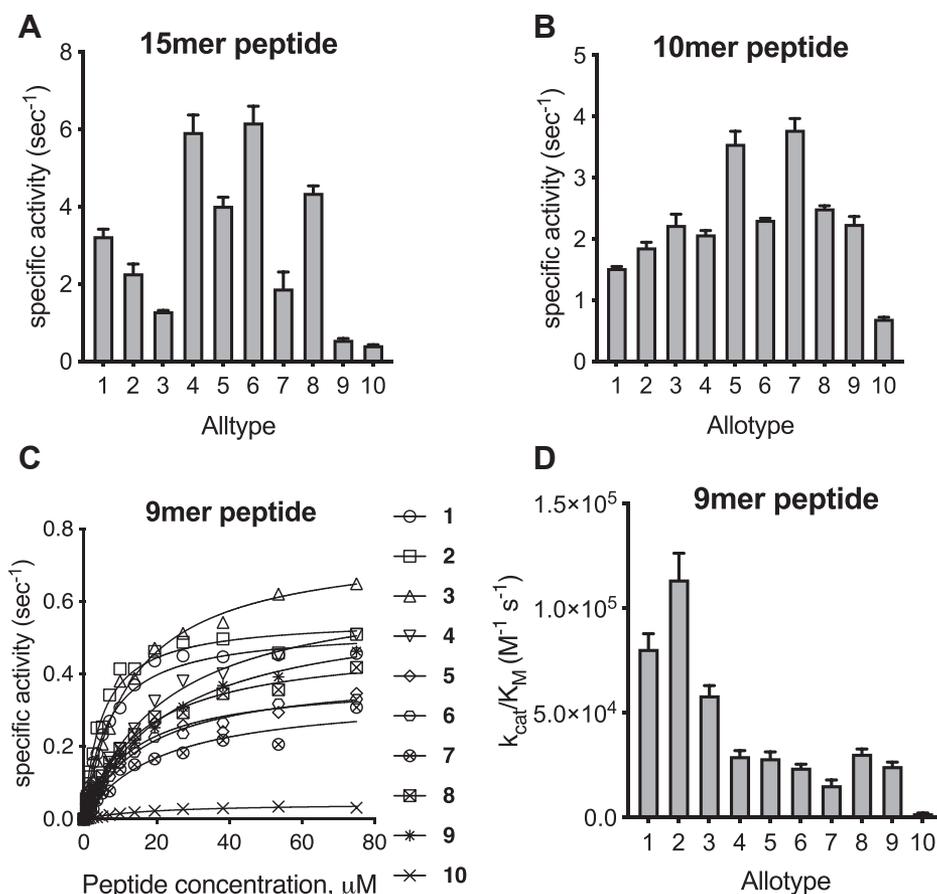


Figure 2. Hydrolysis of peptidic substrates by endoplasmic reticulum aminopeptidase 1 (ERAP1) allotypes. A, specific activity of the hydrolysis of 15mer peptide with the sequence LLRIQRGPGRFVTI. B, specific activity of the hydrolysis of 10mer peptide with the sequence LLKHHAFSFK. C, Michaelis–Menten analysis of the hydrolysis of the 9mer peptide YTAFTIPSI by ERAP1 allotypes. D, catalytic efficiency ($k_{\text{cat}}/K_{\text{M}}$) of each ERAP1 allotype for the trimming of the 9mer epitope YTAFTIPSI.

intermediates accumulate can affect which peptides will bind onto MHC I (35). To evaluate the effect of ERAP1 allotype variation on sequential trimming reactions, we followed the generation of the ovalbumin epitope SIINFEKL from the starting 14mer extended epitope GLEQLESIIINFEKL (Fig. 3). In all cases, the 14mer was catabolized and the 8mer produced but with variable efficiencies. As before, allotype 10 was less active in trimming, and it was necessary to use at a 10-fold higher concentration to follow the reaction. In all reactions, all possible intermediates were detected (Fig. 3A). However, we observed significant differences both between intermediate accumulation and on the overall rate of mature epitope generation (Fig. 3B). Specifically, allotypes 1 to 3 accumulated the mature epitope the fastest, whereas allotypes 5 and 7 were less efficient, in part because of slower trimming of the initial 14mer. This comes as a sharp contrast to their ability to trim the 10mer peptide shown in Figure 3B. Interestingly, allotype 5 accumulated significant amounts of the 11mer intermediate.

Several other allotypes (4 and 6–10) accumulated the 12mer intermediate. Since peptides of 10 to 12 residues can be immunogenic, accumulation of distinct intermediates by different allotypes could contribute to differences in immune responses between individuals.

ERAP1 is an emerging pharmacological target for cancer immunotherapy and the control of inflammatory autoimmunity, including rheumatic conditions such as AS (36, 37). Given the wide distribution of common allotypes in the population, it is crucial to know if inhibitors with clinical potential can effectively inhibit all allotypes. We performed inhibition titrations using Leu-AMC and inhibitors DG013A and GSK849, both shown to be active in cellular assays. DG013A is a potent transition state analog that targets the active site of the enzyme (38). GSK849 targets a regulatory site of ERAP1 and while it is an activator for small substrates, it inhibits long peptide hydrolysis by interfering with C-terminus recognition (34). DG013A was able to inhibit all 10 allotypes with high potency

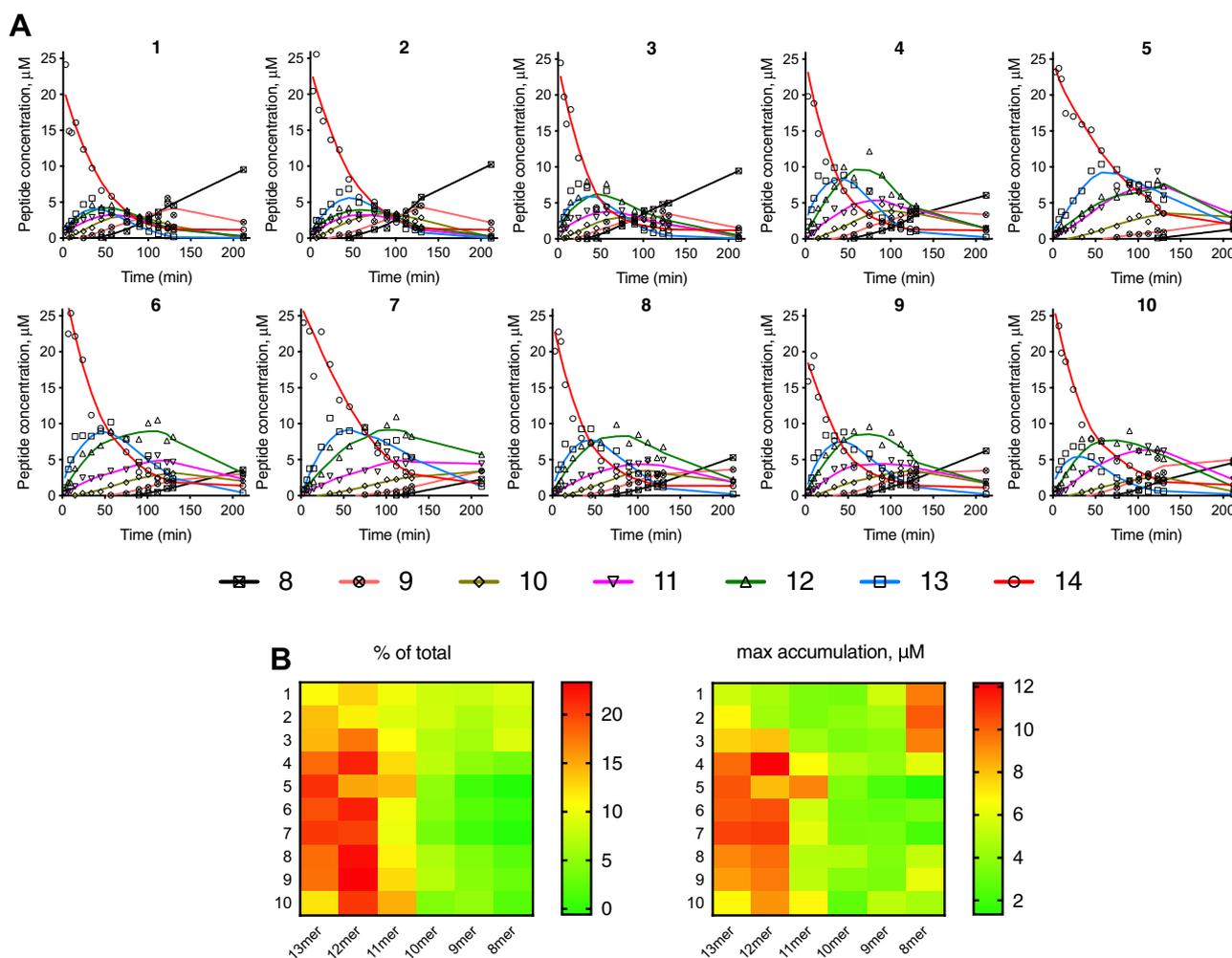


Figure 3. Sequential trimming of peptide substrates by ERAP1 allotypes. A, trimming of the 14mer peptide GLEQLESIIINFEKL, precursor to the ovalbumin epitope SIINFEKL, by endoplasmic reticulum aminopeptidase 1 allotypes. Each peptide is quantified by mass spectrometry, and its concentration is plotted as a function of reaction time. Lines correspond to locally weighted scatterplot smoothing (LOWESS) and are only used as a visual aid to follow peptide accumulation. Allotype 10 was used at a 10-fold higher concentration. B, left, heat map indicating accumulation of each peptide intermediate as a fraction of the total amount of peptides in the reaction; right, heat map showing the maximal accumulation of each peptide intermediate during the reaction.

Enzymatic properties of common ERAP1 allotypes

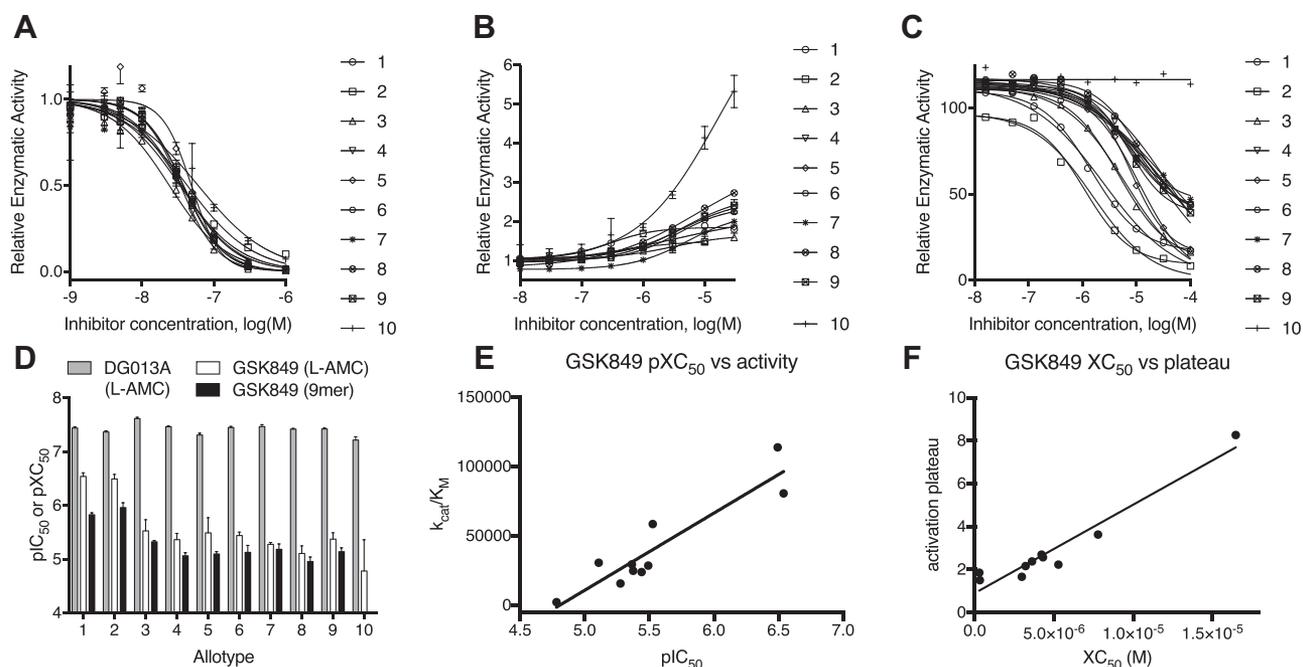


Figure 4. Effect of inhibitors DG013A and GSK849 on the activity of endoplasmic reticulum aminopeptidase 1 (ERAP1) allotypes. *A*, effect of titration of DG013A on the hydrolysis of Leu-AMC by ERAP1 allotypes. *B*, effect of titration of GSK849 on the hydrolysis of Leu-AMC. *C*, ability of GSK849 to inhibit hydrolysis of a 9mer peptide by ERAP1 allotypes. *D*, comparison of pIC_{50} or pXC_{50} values from the titrations shown in panels *A*–*C*. *E*, correlation between allotype specific activity and pXC_{50} value for GSK849. *F*, correlation between maximal activity enhancement by GSK849 and XC_{50} of each allotype. Leu-AMC, L-leucine-7-amido-4-methylcoumarin.

(pIC_{50} between 7.2 and 7.6) (Fig. 4*A*). GSK849 acted as an apparent activator of small substrate hydrolysis as previously reported (34), and its efficacy varied significantly between allotypes (pXC_{50} between 4.8 and 6.5) (Fig. 4*B*). GSK849 was, however, an inhibitor of the more physiologically relevant 9mer substrate YTAFTIPSI (Fig. 4*C*). A comparison of the pIC_{50} and pXC_{50} values for the two substrates is shown in Figure 4*D*. GSK849 was most active against allotypes 1 and 2 and least active *versus* allotype 10. Overall, there was a positive correlation between allotype activity and the ability of GSK849 to inhibit (Fig. 4*E*). In addition, there was a positive correlation between the pXC_{50} value and the fractional activation observed (Fig. 4*F*). This surprising finding suggests that the regulatory, but not the catalytic site of ERAP1, is sensitive to the allotypic state of the enzyme.

The high variability in enzymatic activity between ERAP1 allotypes suggests that individuals carrying different combinations of allotypes can have an even larger variability of ERAP1 activities. Assuming no specific interactions between the two alleles, the total enzymatic activity of an individual would be expected to be the sum of the activities of each allele. A caveat in this analysis is that SNPs and therefore allotypes may affect gene expression or protein turnover, thus affecting the steady-state protein levels. Indeed, two recent reports suggested that SNPs can affect ERAP1 expression to some degree, which could either exacerbate or ameliorate allotype activity variation (18, 19). Since however, existing studies were limited to effects of specific SNPs and not allotypes, and effects on expression levels were relatively small, the potential effect of altered levels of expression is not examined here. To calculate the expected total

activity of the two alleles carried by individuals, we utilized the measurements of activity for the 9mer substrate since we had obtained reliable catalytic efficiency measurements (k_{cat}/K_M , Fig. 2*D* and Table S8). A plot of expected specific activity *versus* population frequency is shown in Figure 5*A*. A bubble chart showing the population frequency and expected enzymatic activity for each possible combination of ERAP1 allotype is shown in Figure 5*B*. Individuals carrying different common combinations of allotypes are expected to possess a wide range, of about 60-fold, of total ERAP1 activity. Most common allotype combinations fall within a more limited range, about 10-fold (cyan region, Fig. 5*A*). Homozygous individuals of allotype 2 are quite frequent in the global population and would be expected to feature the highest ERAP1 activity (*red* region, Fig. 5*A*). Combinations of 2 with 8 are also very common and have moderate-to-high enzyme activity (red region, Fig. 5*A*). Several moderately active combinations of allotype 9 are rare in the population, as is the [4,4] homozygous. Finally, homozygous individuals of allotype 10 are found in $\sim 1.2\%$ of the global population and should have the lowest ERAP1 activity, being functional knockouts for some peptide substrates. Significant variability in both activity and frequency distribution was found in different populations (Fig. S3), a phenomenon that may signify local host–pathogen balancing selection pressures, a notion previously suggested for individual SNPs (39).

Discussion

A wealth of genetic association studies has linked ERAP1 polymorphic variation to the existence of coding SNPs (7, 40), which have been reported to affect enzyme function, thus

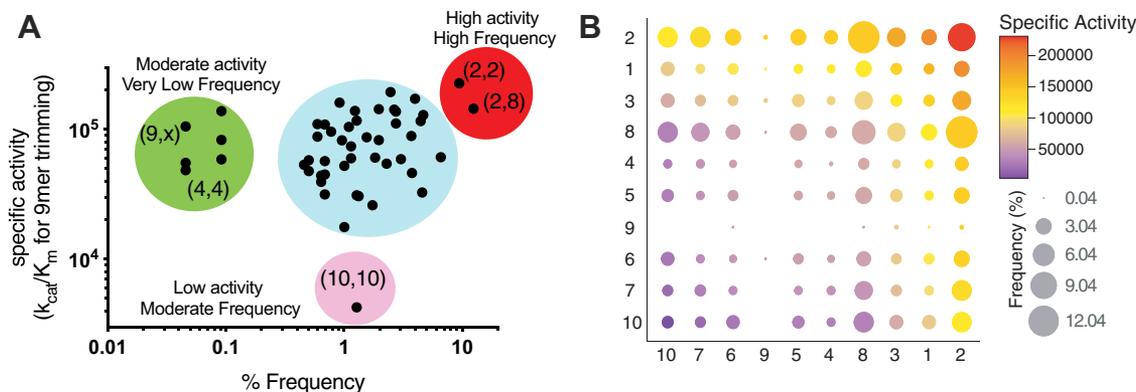


Figure 5. The landscape of endoplasmic reticulum aminopeptidase 1 activities in the population. A, scatterplot showing the correlation between total estimated endoplasmic reticulum aminopeptidase 1 activity and frequency of allotype combinations found in individuals. Specific allotype combinations found in individuals are indicated in parenthesis. B, bubble chart showing the frequency of allotype combinations color coded by their estimated activity. Allotypes have been clustered based on activity (highest activity top right, lowest activity bottom left).

providing mechanistic support for observed epistasis between ERAP1 and HLA (14, 15, 22, 29). In general, functional effects of single SNPs are relatively modest but complex and hint at possible synergism between SNPs that is poorly understood. This is highlighted when common allotypes harbor SNPs that protect along with SNPs that predispose to disease (21). Thus, a deep understanding of the complex patterns of disease susceptibility observed in genetic-association studies requires a detailed knowledge of the functional differences between naturally occurring allotypes and not just SNPs. Our results here suggest that SNPs can synergize to affect function in a substrate-dependent manner. This is probably best highlighted for allotype 10, which is up to 60-fold less active than other allotypes. Such a strong functional change has not been reported before for individual SNPs in ERAP1 and is probably the result of synergism. Specifically, allotype 10 carries the SNP 528R that has been reported to associate with decreased risk to AS and psoriasis and to reduce enzyme activity by about twofold (13, 14, 20, 41). In addition, it carries the SNP 575N, which has been reported to affect enzymatic activity depending on the polymorphic state of position 528 (42). These two SNPs may synergize along with two additional SNPs unique to this allotype: (i) 349V, which is relatively close to the catalytic site although it is a conservative alteration that has not been reported to affect activity in isolation and (ii) 725Q, which has been demonstrated to reduce enzymatic activity (43) and lies in the interface between domains II and IV and could induce changes in the dynamics of the conformational change of ERAP1, similar to position 528 (44). It is possible that synergism between those four SNPs could underlie the greatly reduced catalytic efficiency of allotype 10. Strikingly, this reduction appears to be partially substrate dependent. Trimming of the 10mer substrate was less affected by this allotype compared with the other substrates we tested. This could be due to unproductive interactions between the C-terminal side chain of this peptide and 725R as observed in a recent crystal structure, making the substitution 725Q favorable for this particular substrate (16). This observation highlights a motif

that can be seen throughout our results. While some changes in activity between allotypes are consistent, they can be influenced by the substrate. This is remarkably reminiscent of the effects of polymorphic variation in MHC molecules. Changes in the shape and dynamics of the peptide-binding groove affect the binding of different peptides both thermodynamically and kinetically and contribute to the variability of immune responses (45). It appears that ERAP1 allotypes operate in the same general principle as MHC haplotypes. They induce variability of substrate processing, in a substrate-dependent manner, thus contributing, in tandem with MHC polymorphic variation, to immune response variability within the population.

While it is difficult to dissect the mechanism that underlies the role of each SNP to the activity of each allotype without additional structural information, some insight can still be extracted. Previous studies demonstrated that 528K results in higher enzymatic activity possibly because of effects on the conformational dynamics of the enzyme (44). Accordingly, allotypes 1 to 3, all of which carry this SNP, are among the most active in our assays. The polymorphism 730Q has been suggested to affect activity because of changes in interaction with the C-terminal moiety of the peptide (7, 44). Accordingly, allotypes 1 and 2, which carry this SNP, are more active *versus* the 9mer substrate that carries a hydrophobic C-terminal side chain; allotypes 3 to 10 that have 730E at that location, a negative charge, would be expected to be worse in interacting with this substrate, all are less active compared with 1 and 2. Furthermore, allotypes that carry the 127P polymorphism (also reported to be associated with AS (21)) tend to have lower activity, although no specific effects of this SNP have been reported before (7). This polymorphic location lies on a putative substrate exit channel and the reduced structural flexibility in the mouth of this channel because of the proline residue may reduce the kinetics of product-substrate exchange leading to a slower apparent activity (16). Finally, the sensitivity of the regulatory site to inhibition appears to be allotype-dependent, a notion that is consistent with a previous

Enzymatic properties of common ERAP1 allotypes

study from our group that suggested that the regulatory site communicates with the active site through effects on the conformational dynamics of the enzyme (32). Further structural analysis of these common ERAP1 allotypes will be invaluable in dissecting synergism between SNPs that underlie functional changes in allotypes.

Although ERAP2 has been described as an accessory aminopeptidase that supplements ERAP1 activity, recent genetic association data have pointed to important roles in both cancer immunotherapy and autoimmunity (46–48). Because the rs2248374 polymorphism leads to lack of protein expression (30), the effects on functional ERAP2 activity in the cell can be pronounced since individuals homozygous for the G allele express no full-length enzymatically active ERAP2, and heterozygous individuals express half the canonical protein amount. Although ERAP2 has different specificity than ERAP1 (49, 50), since they are the only known ER-resident aminopeptidases, they cumulatively define the spectrum of aminopeptidase activity in the ER. From the genetic and enzymatic analysis, it appears that the highest ERAP2 expression allotype combination is found more frequently in individuals that carry an intermediate activity ERAP1 allotype (allotype 8), whereas individuals with reduced ERAP2 protein expression are more often homozygous for the most active ERAP1 allotype (allotype 2). This is in line with previous genetic analysis of co-occurrence of ERAP1–ERAP2 haplotypes (10). Thus, it appears that some balancing selection may exist that attempts to normalize total aminopeptidase activity in the ER.

Although our *in vitro* analysis has the advantage of allowing the accurate determination of fundamental molecular properties, how these properties translate to changes in antigen presentation by cells is not always straightforward. Antigen presentation is primarily controlled by peptide binding onto MHCI, and this process may mask some changes in ERAP1 activity. Indeed, recent studies have suggested that the potential of ERAP1 in influencing antigen presentation is limited by this phenomenon (51, 52). Even so, many recent cell-based studies using proteomic approaches have provided strong support that changes in ERAP1 activity because of polymorphic variation are translated to changes in antigen presentation (53). Also, although the expression of ERAP1 mediated by splice interfering variants (18) is in linkage disequilibrium with variants that encode distinct allotypes, genetic association studies have shown that for some (HLA-associated) conditions the disease risk is primarily mapped to enzymatic activity (10). Thus, although MHCI binding is the dominant filter, changes in ERAP1 activity are highly relevant to antigen presentation and probably underlie part of known disease associations.

In summary, we provide a current allotype and genotype analysis for the *ERAP1* gene in human populations and a detailed enzymatic characterization of the 10 most common ERAP1 allotypes. Our results suggest that individual SNPs synergize to shape allotype enzymatic properties by affecting both catalytic efficiency and substrate affinity. Our analysis defines a two-order of magnitude-wide landscape of ERAP1 genotype activities in human populations and suggests that

ERAP1/2 genotypes operate in tandem with MHCI haplotypes to generate the necessary plurality in antigen presentation that supports the observed variability of immune responses between individuals.

Experimental procedures

Materials

Leu-pNA and Leu-AMC were purchased from Sigma–Aldrich. Peptide with the sequence LLKHHAFSFK was purchased from Genecust. Peptide LLRIQRGPGRAFVTI was purchased from JPT Peptide Technologies. Peptide YTAF-TIPSI was purchased from BioPeptide Inc, and ovalbumin peptides were from CRB Discovery. All peptides were HPLC purified to >95% purity and confirmed by mass spectrometry to have the correct mass. Inhibitor DG013A was synthesized as described previously (38), and inhibitor GSK849 was obtained, purified, and characterized as described previously (16).

ERAP1 allotype estimation in the human population

Available phased genotype data for nine coding SNPs at 5q15 in 2504 samples of 26 ethnic groups of European, African, East Asian, South Asian, and mixed American ancestry were obtained from the 1000 Genomes Project phase 3 (25). Phased genotypes were used to estimate the allotypes on each chromosome based on 9-SNP haplotypes, which occur in >1% of all populations. Combinations of the estimated allotype frequencies were plotted using the R package *ggplot2* (54).

Gene constructs, protein expression, and purifications

The sequences corresponding to full-length ERAP1 allotypes with a noncleavable C-terminal 6 His tag were codon optimized for insect expression, and the genes were chemically synthesized (GenScript Biotech). Each gene was then ligated with the BamHI and XhoI-linearized pFASTBAC1 vector using T4 ligase. The product of the ligation reaction was transformed into the competent cells, and the positive clones were selected by single-colony screening and DNA sequencing. The positive plasmid was further verified by digestion with BamHI and XhoI. Protein expression was performed as previously described, with the exception that the C-terminal His tag was not removed (32). Concentrations of protein stock solutions were determined spectrophotometrically using an extinction coefficient of $171,200 \text{ M}^{-1} \text{ cm}^{-1}$ at 280 nm. Protein purity and integrity was estimated by SDS-PAGE and size-exclusion chromatography (using a TSK G3000SW or an Agilent AdvanceBio SEC 300A column) and was over 95% (Figs. S4 and S5). All protein variants were found to be monomeric and monodispersed. Stability toward aggregation was tested by size-exclusion chromatography after repeat freeze–thaw cycles. All ERAP1 allotypes were validated for the presence of the appropriate combinations of SNPs by DNA sequencing (Fig. S6).

Enzymatic assays with dipeptide substrates

The enzymatic activity of ERAP1 was measured using the dipeptidic fluorogenic substrate Leu-AMC as previously

described (55). Briefly, the change in fluorescence at 460 nm (excitation at 380 nm) was followed over time using a TECAN SPARK 10 m plate reader. A standard curve of aminocoumarin was used to convert the signal to product concentration. For MM measurements, the dipeptide substrate Leu-pNA was used, and the generation of pNA was followed by measuring the absorption at 405 nm as described previously (55). For experiments measuring the rate of hydrolysis *versus* enzyme or substrate concentrations, measurements using the dipeptidic fluorogenic substrate Leu-AMC were made in a buffer of 20 mM Hepes (pH 7.0), 100 mM NaCl, 0.002% Tween 20, in deep 384-well plates at a final volume of 50 μ l/well at 20 °C, using a Tecan M1000 plate reader with 360 nm excitation and 460 nm emission (5 nm bandpass on both monochromators). Measurements where enzyme concentration was varied were set up using a multi-channel pipette and initiated by adding a final concentration of 25 μ M Leu-AMC. Measurements where substrate concentration was varied were set up using a Hewlett Packard D300 digital dispenser and initiated by adding a final concentration of 25 nM ERAP1 (allotypes 1–9) or 250 nM ERAP1 (allotype 10). An aminocoumarin standard curve was used to convert fluorescence intensity to product concentration, after which initial rates were obtained by linear fit of the early region of the time courses.

Enzymatic assays with peptides

Trimming of peptides LLRIQRGPGRAFVTI and LLKHHAFFSK was performed as described previously (16). Briefly, 20 μ M peptide and 1 nM enzyme at a final assay volume of 200 μ l were mixed in assay buffer of 20 mM Hepes (pH 7.0), 100 mM NaCl, 0.002% Tween 20, and incubated for 30 min at 37 °C. Reactions were carried out in three replicates for each allotype, stopped by freezing, and were stored at –80 °C until analyzed by HPLC. For trimming assays using the YTAFTIPSI substrate, measurements were made in an assay buffer of 20 mM Hepes (pH 7.0), 100 mM NaCl, 0.002% Tween 20, in deep 384-well plates at a final assay volume of 25 μ l/well and temperature of 20 °C. For the MM kinetics, a range of concentration of YTAFTIPSI peptide was dispensed using a Hewlett Packard D300 digital dispenser, and reactions were initiated by addition of a final concentration of 1 nM ERAP1 (allotypes 1–9) or 10 nM ERAP1 (allotype 10). Reactions were stopped after 60 min of incubation by addition of 25 μ l of 0.75% TFA in water containing 5 μ M Ac-YTAFTIPSI as internal standard. Mass signal intensities corresponding to the product (TAFTIPSI) and internal standard (Ac-YTAFTIPSI) were measured on a Rapidfire autosampler (Agilent) equipped with a C18 solid phase cartridge, coupled to a Sciex 4000 Q-trap MS (AB Sciex), using multiple reaction monitoring, as described previously (34). Integrated product intensity signals were normalized to the respective internal standard intensity signals, before conversion to product concentration using a TAFTIPSI standard curve. Turnover did not exceed 35% at any substrate concentration. Data were fitted to the MM equation as described previously (34). To measure inhibition by GSK849, a threefold dilution series of inhibitor was made in dimethyl sulfoxide, and 250 nl of each concentration point was dispensed to a 384-well assay plate

using an Echo acoustic dispenser. ERAP1 was added followed by YTAFTIPSI substrate at a final concentration of 5 μ M. Final ERAP1 concentrations varied according to allotypes as follows: 0.5 nM allotype 3; 1 nM allotypes 1, 2, 4, 5, 6, 8, and 9; 2 nM allotype 7; and 10 nM allotype 10. After incubation for 60 min, reactions were stopped and measured as described previously. Calibration curves were measured for substrate (YTAFTIPSI) and product (TAFTIPSI) and used to correct for the difference in detection sensitivity between the two analytes. The corrected intensity data were used to calculate percent turnover, from which product concentration was calculated. Data were normalized to percent enzymatic activity between high (uninhibited) and low (100 μ M DG13) controls and fitted to a four-parameter logistic expression as described previously (34).

To follow the trimming of ovalbumin epitope precursors, measurements were made in an assay buffer of 50 mM Hepes (pH 7.0), 100 mM NaCl, 0.002% Tween 20, in deep 384-well plates at a final assay volume of 50 μ l/well and temperature of 20 °C. ERAP1 (allotypes 1–9 at 10 nM final concentration, allotype 10 at 100 nM final concentration) were mixed with 14-mer peptide GLEQLESIIINFEKL (at 25 μ M final concentration). Wells were stopped at increasing time points by the addition of 50 μ l of 0.75% TFA in water containing 5 μ M Ac-YTAFTIPSI as internal standard. Mass signal intensities corresponding to the seven ovalbumin peptide species (14-mer GLEQLESIIINFEKL, 13-mer LEQLESIIINFEKL, 12-mer EQLESIIINFEKL, 11-mer QLESIIINFEKL, 10-mer LESIIINFEKL, 9-mer ESIIINFEKL, and 8-mer SIIINFEKL) and internal standard (Ac-YTAFTIPSI) were measured using the Rapidfire device as described previously. Signals at each peptide mass were divided by the respective internal standard signal. Normalized data were converted to peptide concentration using calibration curves for each peptide species.

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 or estratikos@chem.uoa.gr).

Supporting information—This article contains [supporting information](#).

Acknowledgments—This research has been funded in part by internal funds of the National Centre for Scientific Research “Demokritos” and the UK Biobank Resource (application no. 20361).

Author contributions—J. P. H. helped design the study as well as designed and performed kinetic analysis experiments, processed data, and interpreted results; I. T. designed and performed enzymatic analysis, processed data, and interpreted results; J. K. and A. C. supervised and performed genetic analysis and interpreted the results; J. K. supervised and performed construct design and protein production and validation; S. K. helped conceive the study, supervised experiments, and interpreted results; E. S. helped conceive and

Enzymatic properties of common ERAP1 allotypes

design the study, supervised experiments, interpreted results, and wrote the article with help from J. P. H., S. K., and J. K. All authors have edited and approved the final version of the article.

Funding and additional information—J. K. is supported by a VENI award from the Netherlands Organization for Scientific Research (NWO project number 016.186.006).

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

Abbreviations—The abbreviations used are: AS, ankylosing spondylitis; ER, endoplasmic reticulum; ERAP, endoplasmic reticulum aminopeptidase; HLAs, human leukocyte antigens; Leu-AMC, L-leucine-7-amido-4-methylcoumarin; Leu-pNA, L-leucine-para-nitroanilide; MHC, major histocompatibility complex; MHCI, MHC class I molecules; MM, Michaelis–Menten.

References

1. 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
2. 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
3. Bles, A., Janulienė, D., Hofmann, T., Koller, N., Schmidt, C., Trowitzsch, S., Moeller, A., and Tampe, R. (2017) Structure of the human MHC-I peptide-loading complex. *Nature* **551**, 525–528
4. Cascio, P., Hilton, C., Kisselev, A. F., Rock, K. L., and Goldberg, A. L. (2001) 26S proteasomes and immunoproteasomes produce mainly N-extended versions of an antigenic peptide. *EMBO J.* **20**, 2357–2366
5. Weimershaus, M., Evnouchidou, I., Saveanu, L., and van Endert, P. (2013) Peptidases trimming MHC class I ligands. *Curr. Opin. Immunol.* **25**, 90–96
6. Strange, A., Capon, F., Spencer, C. C., Knight, J., Weale, M. E., Allen, M. H., Barton, A., Band, G., Bellenguez, C., Bergboer, J. G., Blackwell, J. M., Bramon, E., Bumpstead, S. J., Casas, J. P., Cork, M. J., et al. (2010) A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat. Genet.* **42**, 985–990
7. Lopez de Castro, J. A., Alvarez-Navarro, C., Brito, A., Guasp, P., Martin-Esteban, A., and Sanz-Bravo, A. (2016) Molecular and pathogenic effects of endoplasmic reticulum aminopeptidases ERAP1 and ERAP2 in MHC-I-associated inflammatory disorders: Towards a unifying view. *Mol. Immunol.* **77**, 193–204
8. 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
9. 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
10. Kuiper, J. J. W., Setten, J. V., Devall, M., Cretu-Stancu, M., Hiddingh, S., Ophoff, R. A., Missotten, T., Velthoven, M. V., Den Hollander, A. I., Hoyng, C. B., James, E., Reeves, E., Cordero-Coma, M., Fonollosa, A., Adan, A., et al. (2018) Functionally distinct ERAP1 and ERAP2 are a hallmark of HLA-A29-(birdshot) uveitis. *Hum. Mol. Genet.* **27**, 4333–4343
11. Chen, L., Fischer, R., Peng, Y., Reeves, E., McHugh, K., Ternette, N., Hanke, T., Dong, T., Elliott, T., Shastri, N., Kollnberger, S., James, E., Kessler, B., and Bowness, P. (2014) Critical role of endoplasmic reticulum aminopeptidase 1 in determining the length and sequence of peptides bound and presented by HLA-B27. *Arthritis Rheumatol.* **66**, 284–294
12. Guasp, P., Alvarez-Navarro, C., Gomez-Molina, P., Martin-Esteban, A., Marcilla, M., Barnea, E., Admon, A., and Lopez de Castro, J. A. (2016) The peptidome of Behcet's disease-associated HLA-B*51:01 includes two subpeptidomes differentially shaped by endoplasmic reticulum aminopeptidase 1. *Arthritis Rheumatol.* **68**, 505–515
13. Sanz-Bravo, A., Campos, J., Mazariegos, M. S., and Lopez de Castro, J. A. (2015) Dominant role of the ERAP1 polymorphism R528K in shaping the HLA-B27 peptidome through differential processing determined by multiple peptide residues. *Arthritis Rheumatol.* **67**, 692–701
14. 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
15. Reeves, E., Colebatch-Bourn, A., Elliott, T., Edwards, C. J., and James, E. (2014) Functionally distinct ERAP1 allotype combinations distinguish individuals with Ankylosing Spondylitis. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 17594–17599
16. 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.* **52**, 26709–26716
17. Papakyriakou, A., and Stratikos, E. (2017) The role of conformational dynamics in antigen trimming by intracellular aminopeptidases. *Front. Immunol.* **8**, 946
18. Hanson, A. L., Cuddihy, T., Haynes, K., Loo, D., Morton, C. J., Oppermann, U., Leo, P., Thomas, G. P., Le Cao, K. A., Kenna, T. J., and Brown, M. A. (2018) Genetic variants in ERAP1 and ERAP2 associated with immune-mediated diseases influence protein expression and the isoform profile. *Arthritis Rheumatol.* **70**, 255–265
19. Costantino, F., Talpin, A., Evnouchidou, I., Kadi, A., Leboime, A., Said-Nahal, R., Bonilla, N., Letourneur, F., Leturcq, T., Ka, Z., van Endert, P., Garchon, H. J., Chiochia, G., and Breban, M. (2015) ERAP1 gene expression is influenced by nonsynonymous polymorphisms associated with predisposition to spondyloarthritis. *Arthritis Rheumatol.* **67**, 1525–1534
20. Goto, Y., Hattori, A., Ishii, Y., and Tsujimoto, M. (2006) Reduced activity of the hypertension-associated Lys528Arg mutant of human adipocyte-derived leucine aminopeptidase (A-LAP)/ER-aminopeptidase-1. *FEBS Lett.* **580**, 1833–1838
21. 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
22. 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
23. Wordsworth, B. P., Vecellio, M., and Roberts, A. R. (2017) Reply to Reeves et al.: No evidence for rare ERAP1 haplotypes and haplotype combinations in ankylosing spondylitis. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1577
24. 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
25. Genomes Project, C., Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G. A., and Abecasis, G. R. (2015) A global reference for human genetic variation. *Nature* **526**, 68–74
26. Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., et al. (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443
27. 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

28. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., *et al.* (2015) UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779
29. Evnouchidou, I., Birtley, J., Seregin, S., Papakyriakou, A., Zervoudi, E., Samiotaki, M., Panayotou, G., Giastas, P., Petrakis, O., Georgiadis, D., Amalfitano, A., Saridakis, E., Mavridis, I. M., and Stratikos, E. (2012) A common single nucleotide polymorphism in endoplasmic reticulum aminopeptidase 2 induces a specificity switch that leads to altered antigen processing. *J. Immunol.* **189**, 2383–2392
30. Andres, A. M., Dennis, M. Y., Kretzschmar, W. W., Cannons, J. L., Lee-Lin, S. Q., Hurler, B., Program, N. C. S., Schwartzberg, P. L., Williamson, S. H., Bustamante, C. D., Nielsen, R., Clark, A. G., and Green, E. D. (2010) Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. *PLoS Genet.* **6**, e1001157
31. Nguyen, T. T., Chang, S. C., Evnouchidou, I., York, I. A., Zikos, C., Rock, K. L., Goldberg, A. L., Stratikos, E., and Stern, L. J. (2011) Structural basis for antigenic peptide precursor processing by the endoplasmic reticulum aminopeptidase ERAP1. *Nat. Struct. Mol. Biol.* **18**, 604–613
32. Giastas, P., Neu, M., Rowland, P., and Stratikos, E. (2019) High-resolution crystal structure of endoplasmic reticulum aminopeptidase 1 with bound phosphonic transition-state analogue inhibitor. *ACS Med. Chem. Lett.* **10**, 708–713
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. Liddle, J., Hutchinson, J. P., Kitchen, S., Rowland, P., Neu, M., Cecconie, T., Holmes, D. S., Jones, E., Korczynska, J., Koumantou, D., Lea, J. D., Nickels, L., Pemberton, M., Phillipou, A., Schneck, J. L., *et al.* (2020) Targeting the regulatory site of ER aminopeptidase 1 leads to the discovery of a natural product modulator of antigen presentation. *J. Med. Chem.* **63**, 3348–3358
35. Lorente, E., Barriga, A., Johnstone, C., Mir, C., Jimenez, M., and Lopez, D. (2013) Concerted *in vitro* trimming of viral HLA-B27-restricted ligands by human ERAP1 and ERAP2 aminopeptidases. *PLoS One* **8**, e79596
36. Stratikos, E. (2014) Modulating antigen processing for cancer immunotherapy. *Oncoimmunology* **3**, e27568
37. 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
38. Zervoudi, E., Saridakis, E., Birtley, J. R., Seregin, S. S., Reeves, E., Kokkala, P., Aldhamen, Y. A., Amalfitano, A., Mavridis, I. M., James, E., Georgiadis, D., and Stratikos, E. (2013) Rationally designed inhibitor targeting antigen-trimming aminopeptidases enhances antigen presentation and cytotoxic T-cell responses. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 19890–19895
39. Cagliani, R., Riva, S., Biasin, M., Fumagalli, M., Pozzoli, U., Lo Caputo, S., Mazzotta, F., Piacentini, L., Bresolin, N., Clerici, M., and Sironi, M. (2010) Genetic diversity at endoplasmic reticulum aminopeptidases is maintained by balancing selection and is associated with natural resistance to HIV-1 infection. *Hum. Mol. Genet.* **19**, 4705–4714
40. Agrawal, N., and Brown, M. A. (2014) Genetic associations and functional characterization of M1 aminopeptidases and immune-mediated diseases. *Genes Immun.* **15**, 521–527
41. Evans, D. M., Spencer, C. C., Pointon, J. J., Su, Z., Harvey, D., Kochan, G., Oppermann, U., Dilthey, A., Pirinen, M., Stone, M. A., Appleton, L., Moutsianas, L., Leslie, S., Wordsworth, T., Kenna, T. J., *et al.* (2011) Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat. Genet.* **43**, 761–767
42. Martin-Esteban, A., Gomez-Molina, P., Sanz-Bravo, A., and Lopez de Castro, J. A. (2014) Combined effects of ankylosing spondylitis-associated ERAP1 polymorphisms outside the catalytic and peptide-binding sites on the processing of natural HLA-B27 ligands. *J. Biol. Chem.* **289**, 3978–3990
43. 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
44. Stamogiannos, A., Koumantou, D., Papakyriakou, A., and Stratikos, E. (2015) Effects of polymorphic variation on the mechanism of endoplasmic reticulum aminopeptidase 1. *Mol. Immunol.* **67**, 426–435
45. Ayres, C. M., Abualrous, E. T., Bailey, A., Abraham, C., Hellman, L. M., Corcelli, S. A., Noe, F., Elliott, T., and Baker, B. M. (2019) Dynamically driven allostery in MHC proteins: Peptide-dependent tuning of class I MHC global flexibility. *Front. Immunol.* **10**, 966
46. de Castro, J. A. L., and Stratikos, E. (2019) Intracellular antigen processing by ERAP2: Molecular mechanism and roles in health and disease. *Hum. Immunol.* **80**, 310–317
47. Lim, Y. W., Chen-Harris, H., Mayba, O., Lianoglou, S., Wuster, A., Bhangale, T., Khan, Z., Mariathasan, S., Daemen, A., Reeder, J., Haverty, P. M., Forrest, W. F., Brauer, M., Mellman, I., and Albert, M. L. (2018) Germline genetic polymorphisms influence tumor gene expression and immune cell infiltration. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E11701–E11710
48. Kuiper, J. J., Van Setten, J., Ripke, S., Van, T. S. R., Mulder, F., Missotten, T., Baarsma, G. S., Francioli, L. C., Pulit, S. L., De Kovel, C. G., Ten Dam-Van Loon, N., Den Hollander, A. I., Huis in het Veld, P., Hoyng, C. B., Cordero-Coma, M., *et al.* (2014) A genome-wide association study identifies a functional ERAP2 haplotype associated with birdshot chorioretinopathy. *Hum. Mol. Genet.* **23**, 6081–6087
49. Zervoudi, E., Papakyriakou, A., Georgiadou, D., Evnouchidou, I., Gajda, A., Poreba, M., Salvesen, G. S., Drag, M., Hattori, A., Swevers, L., Vourloumis, D., and Stratikos, E. (2011) Probing the S1 specificity pocket of the aminopeptidases that generate antigenic peptides. *Biochem. J.* **435**, 411–420
50. Mpakali, A., Giastas, P., Mathioudakis, N., Mavridis, I. M., Saridakis, E., and Stratikos, E. (2015) Structural basis for antigenic peptide recognition and processing by endoplasmic reticulum (ER) aminopeptidase 2. *J. Biol. Chem.* **290**, 26021–26032
51. Admon, A. (2019) ERAP1 shapes just part of the immunopeptidome. *Hum. Immunol.* **80**, 296–301
52. Komov, L., Kadosh, D. M., Barnea, E., Milner, E., Hendler, 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
53. de Castro, J. A. L. (2018) How ERAP1 and ERAP2 shape the peptidomes of disease-associated MHC-I proteins. *Front. Immunol.* **9**, 2463
54. Ginestet, C. (2011) ggplot2: Elegant graphics for data analysis. *J. R. Stat. Soc. A Stat.* **174**, 245
55. Stamogiannos, A., Maben, Z., Papakyriakou, A., Mpakali, A., Kokkala, P., Georgiadis, D., Stern, L. J., and Stratikos, E. (2017) Critical role of interdomain interactions in the conformational change and catalytic mechanism of endoplasmic reticulum aminopeptidase 1. *Biochemistry* **56**, 1546–1558