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

Jonathan P. Hutchinson[1,#], Ioannis Temponeras[2,#], Jonas Kuiper[3], Adrian Cortes[4], Justyna Korczynska[1], Semra Kitchen[5] and Efstratios Stratikos[2,6,*]

Affiliations:
[1]Medicinal Science and Technology, GlaxoSmithKline, Stevenage, Hertfordshire SG1 2NY, U.K.
[2]National Centre for Scientific Research "Demokritos", Athens 15341, Greece.
[3]Department of Ophthalmology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands. Center for Translational Immunology, University Medical Center Utrecht, University of Utrecht, Utrecht, Netherlands.
[4] Human Genetics, GlaxoSmithKline, Stevenage, Hertfordshire SG1 2NY, U.K.
[5] Adaptive Immunity Research Unit, GlaxoSmithKline, Stevenage, Hertfordshire SG1 2NY, U.K.
[6] Laboratory of Biochemistry, Department of Chemistry, National and Kapodistrian University of Athens, Panepistimiopolis Zografou 157 84, Greece.

[#] These authors contributed equally to the manuscript
[*] E-mail: stratos@rrp.demokritos.gr or estratikos@chem.uoa.gr.

Abstract

Objective: Polymorphic variation of immune system proteins can drive variability of individual immune responses. ER aminopeptidase 1 (ERAP1) generates antigenic peptides for presentation by MHC class I molecules. Coding single nucleotide polymorphisms (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.

Methods: We used phased genotype data to estimate ERAP1 allotype frequency in 2,504 individuals across five major human populations, generated highly pure recombinant enzymes corresponding to the 10 most common ERAP1 allotypes and systematically characterized their *in vitro* enzymatic properties.

Results: We find that ERAP1 allotypes possess a wide range of enzymatic activities, 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 sub-active ERAP1 gene. Enzymatic analysis revealed that ERAP1 allotypes can differ in both catalytic efficiency and substrate affinity, differences that can change intermediate accumulation in multi-step trimming reactions. Alterations in efficacy of an allosteric inhibitor that targets the regulatory site of the enzyme suggest that allotypic variation influences the communication between the regulatory and the active site.

Conclusion: 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 MHC haplotypes, to immune response variability and to predisposition to chronic inflammatory conditions.

Introduction

The Major Histocompatibility Complex molecules (MHC, Human Leukocyte Antigens in humans) are the most polymorphic human genes with tens of thousands different allomorphs identified to date (1). This polymorphic variation is considered to be important for enhancing population immune responses to existing and novel pathogens within the population and is often the result of balancing selection in host-pathogen interactions. MHC class I molecules (MHCI), 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.

MHCI bind their peptide cargo in the Endoplasmic Reticulum (ER) with the assistance of the peptide loading complex (3). Peptide candidates for binding onto MHCI enter the ER through the action of the *Transporter associated with antigen processing* (TAP) (4). While MHCI tend to bind peptides that are between 8-11 amino acids long (the majority of which are 9mers) many peptides that enter the ER can be substantially longer (5). 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 (6). As a result, the activity of these two enzymes can affect which peptides are finally presented by MHCI and shape the resulting adaptive immune responses.

The *ERAP1* gene is also unusually polymorphic and a variety of coding single nucleotide polymorphisms (SNPs) confer susceptibility to human disease, most notably chronic inflammatory conditions. Here, the genetic signature of the ERAP1 associations is the often epistasis with HLA class I alleles, which emphasizes the critical role of ERAP1 in HLA class I antigen presentation (7-10). The genetic association with *ERAP1* SNPs and inflammatory diseases such as *HLA-B27-associated Ankylosing Spondylitis*, *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 (11-14). Several of these ERAP1 SNPs have been described to affect the function of the enzyme (15, 16). Mechanisms proposed to underlie

this effect include direct interactions with the substrate (17), effects on the conformational dynamics of the enzyme (18) and effects on the protein expression level (19, 20), or combinations of these(11). However, not all of the possible combinations of the 9 most common coding SNPs (i.e. allotypes) (21) occur at equal frequency in the human population (10). Rather, these SNPs encode a rather limited palette of allotypes of ERAP1 that are maintained at high frequency (>1%) in human populations and suggests functional asymmetry between allotypes of ERAP1. This is supported by the fact that some ERAP1 allotypes are protective while others confer risk to inflammatory diseases (22). A deep understanding of the functional properties of full-length ERAP1 allotypes rather than individual SNPs is critical to unraveling the physiological impact on health and disease, but the mechanistic consequences of particular SNP combinations as they naturally occur across all common allotypes are not known.

Previous studies have described several ERAP1 allotypes. In most studies, but not all, these allotypes are defined as the combination of 9 coding SNPs at amino acid positions 56, 127, 276, 346, 349, 528, 575, 725 and 730. Ombrello et al reported ten common ERAP1 allotypes in three populations of European and East Asian Ancestry (n=160) (10). Reeves et al reported some additional distinct allotypes discovered in a small patient cohort of Ankylosing Spondylitis (n=72) (16). These allotypes however have been controversial and proposed to be rare by others (23, 24). In a recent study, Reeves et al also explored ERAP1 allotypes in a small patient cohort of oropharyngeal squamous cell carcinoma (n=25) (25). While these studies have contributed to our understanding of the emerging role of ERAP1 allotypes in disease, some controversy still remains on which ERAP1 allotypes are common in the population and systematic analysis of their functional differences has been lacking.

In this study, we used phased genotype data of the 1000 genome project (26) to define common ERAP1 allotypes in 2504 individuals of five major human populations. We then generated recombinant versions of the 10 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 that are due to effects on catalytic efficiency and substrate affinity. Allotype 10 is consistently a low-activity variant, suggesting that a significant number of people carry a sub-active ERAP1 protein and thus individuals that are homozygous for allotype 10 have minimal ERAP1 activity. Our findings suggest that ERAP1 allotypic variation has the potential to strongly synergize with MHCI

alleles, in an epitope dependent manner, to enhance immune system variability in natural human populations.

Experimental methods

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 YTAFTIPSI 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 (27) and inhibitor GSK849 was obtained, purified and characterized as described previously (17).

ERAP1 allotype estimation in the human population

Available phased genotype data for 9 coding SNPs at 5q15 in 2504 samples of 26 ethnic groups of European (EUR), African (AFR), East Asian (EAS), South Asian (SAS), and mixed American (AMR) ancestry were obtained from the 1000 Genomes Project Phase 3 (26). 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 (28)*.

Gene constructs, protein expression and purifications

The sequence corresponding to full length ERAP 1 allotype 2 with a non-cleavable C-terminal 6 His-tag was codon optimised for insect expression and the gene was chemically synthesised. The gene was then ligated with the BamHI and XhoI-linearised 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. The remaining ERAP1 allotypes were generated by site directed mutagenesis using allotype 2 construct as a DNA template. Protein expression was performed as previously described, with the exception that the C-terminal His tag was not removed (29). Concentrations of protein stock solutions were

determined spectrophotometrically using an extinction coefficient of 171,200 $M^{-1}cm^{-1}$ at 280 nm.

Enzymatic assays with dipeptide substrates

The enzymatic activity of ERAP1 was measured using the dipeptidic fluorigenic substrate Leucine-7-amido-4-methycoumarin (Leu-AMC) as previously described (30). Briefly, the change in fluorescence at 460 nm (excitation at 380 nm) was followed over time using a TECAN SPARK 10m plate reader. A standard curve of AMC was used to convert the signal to product concentration. For MM measurements the dipeptide substrate L-Leucine-para-nitroanilide (Leu-pNA) was used and the generation of pNA was followed by measuring the absorption at 405 nm as described previously (30). For experiments measuring the rate of hydrolysis versus enzyme or substrate concentrations, measurements using the dipeptidic fluorogenic substrate Leucine-7-amidocoumarin (Leu-AMC) were made in a buffer of 20 mM HEPES pH 7.0, 100 mM NaCl, 0.002% Tween 20, in deep well 384 well plates at a final volume of 50 uL/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 multichannel pipette and initiated by adding a final concentration of 25 µM 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 (AMC) 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 timecourses.

Enzymatic assays with peptides

Trimming of peptides LLRIQRGPGRAFVTI and LLKHHAFSFK was performed as described previously (17). Briefly, 20 µM peptide and 1 nM enzyme at a final assay volume of 200 µL were mixed in assay buffer of 20 mM Heres pH 7.0, 100 mM NaCl, 0.002% Tween 20 and incubated for 30 min at 37 ^{o}C. Reactions were carried out in 3 replicates for each allotype, stopped by freezing and were stored at -80 ^{o}C until analyzed by HPLC chromatography. 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 well 384 well plates at a

final assay volume of 25 uL/well and temperature of 20°C. For the Michaelis-Menten kinetics a range of concentrations 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 minutes incubation by addition of 25 µL of 0.75 % TFA in water containing 5 µM 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 (MRM), as described previously (31). 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 Michaelis-Menten equation as described previously (31). To measure inhibition by GSK849, a 3-fold dilution series of inhibitor was made in DMSO 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 uM. Final ERAP1 concentrations varied according to allotype as follows: 0.5 nM allotype 3; 1 nM allotypes 1,2,4,5,6,8 and 9; 2 nM allotype 7; 10 nM allotype 10. After incubation for 60 minutes, reactions were stopped and measured as described above. 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 % turnover, from which product concentration was calculated. Data were normalized to % enzymatic activity between high (uninhibited) and low (100 µM DG13) controls and fitted to a 4 parameter logistic expression as described previously (31).

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 well 384 well plates at a final assay volume of 50 µL/well and temeprature 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 GLEQLESIINFEKL (at 25 µM final concentration). Wells were stopped at increasing time points by the addition of 50 µL of 0.75% TFA in water containing 5 µM Ac-YTAFTIPSI as internal standard. Mass signal intensities corresponding to the 7 ovalbumin peptide species (14-mer GLEQLESIINFEKL, 13-mer LEQLESIINFEKL, 12-mer EQLESIINFEKL, 11-

mer QLESIINFEKL, 10-mer LESIINFEKL, 9-mer ESIINFEKL, 8-mer SIINFEKL) and internal standard (Ac-YTAFTIPSI) were measured using the Rapidfire device as described above. 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.

Results

To address the question which ERAP1 allotypes are common in human populations we turned to public databases that now contain large volumes of human genetic data. Analysis of the Genome Aggregation Database that contains 125,748 exome sequences (32) using a 1% frequency cutoff to qualify a coding genomic missense variant as a polymorphism, revealed only 10 amino acid positions as polymorphic, namely 12, 56, 127, 276, 346, 349, 528, 575, 725 and 730, consistent with a previous study (10). 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 9 positions. These 9 SNPs could be theoretically organized in up to 2^9 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 (33). The frequency of the nine ERAP1 SNPs in different populations is shown in Supplemental Table 1a. Correlations between individual SNPs that indicate linkage disequilibrium are shown in Supplemental Table 1b and are generally consistent with previous studies (10). 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 (Supplemental Table 2) (34). Overall, our analysis is in good agreement with Ombrello *et al* (10). Strikingly, although 10 common allotypes constitute 99.9% of all allotypes in the European population, some populations had additional allotypes not reported before. Specifically, the 10 allotypes (number 1-10 in Table 1 and in (10)) covered less than 80% of the African population, which included additional common allotypes. Interestingly, the allotype 10 frequency in the AFR population is about half of the allele frequency of rs2287987-C (349V) in the same population although this SNP is only present in allotype 10. This apparent contradiction however is resolved by the observation that 349V frequently resides in uncharacterized allotypes within the AFR population. Overall, we were able to identify at least 6 additional allotypes that had frequencies of over 0.5% in at least one population (Table 1, allotypes 11 to 16). Our analysis

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 1-10 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 10 allotypes. Since individuals carry two copies of the ERAP1 gene, we also analyzed the combinations of allotypes present in the 2,504 samples (Table 1b and supplemental Table 3). The most common combination was that of allotype 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% x 21.8%=5.6%). We additionally analyzed the prevalence of two SNPs in the homologous ERAP2, namely rs2549782 and rs2248374. The former is a missense coding SNP that changes a single amino acid near the catalytic center of the enzyme and has been shown to affect ERAP2 activity and specificity (35). The latter affects ERAP2 expression levels through splicing a transcript variant susceptible to nonsense-mediated RNA decay (36). 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 (36)) in 44.7% of the samples and the [T,G] allotype (defined as allotype B in (36)) in 54.3% of the samples. Since however the [G] allele of the rs2248374 leads to no detectable expression of ERAP2 (36) our results suggest that 54.3% of individuals in the database carry an allotype that results to no ERAP2 protein expression under steady state conditions. Because we previously showed that the co-occurrence of functional ERAP2 protein is dependent on the ERAP1 haplotype background (11) in Dutch and Spanish controls, we were interested to see if the ERAP1 and ERAP2 allotypes correlate in this global data set. Thus, we calculated the distribution of ERAP1 allotypes in each group of individuals carrying a particular ERAP2 allotype (Supplemental Table 4). This analysis 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 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 manuscript.

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

	Amino acid composition at each SNP									Population allotype frequency %					
Allotype	**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
Sub-Total										*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
Sub-total										5.6	20.3	0.9	0.1	0.1	0.7
TOTAL										*99.7*	*99.5*	*99.7*	*100*	*99.9*	*99.2*

Table 1b: Distribution of combinations of ERAP1 allotypes in the global population

	% in population (global)									
Allotype	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***
1	0.9									
2	2.2	8.2								
3	1.1	3.5	1.1							
4	0.5	1.7	0.5	0.1						
5	0.6	2.4	1.4	0.4	0.6					
6	1.0	2.4	0.9	0.4	0.9	0.4				
7	0.7	4.1	1.0	0.6	0.6	0.6	1.1			
8	2.4	11.0	3.2	1.1	2.6	2.0	3.3	5.8		
9	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	
10	1.7	4.0	1.8	0.6	1.3	1.6	1.0	4.1	0.0	1.2

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 1-10 are scattered throughout the structure of ERAP1, away from the catalytic center (Figure 1) 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 (Figure 1A) and b) SNPs that line the interior surface of this cavity and may make

direct interactions with the peptide substrate (Figure 1B). All variants were produced and purified to homogeneity as described in the methods section.

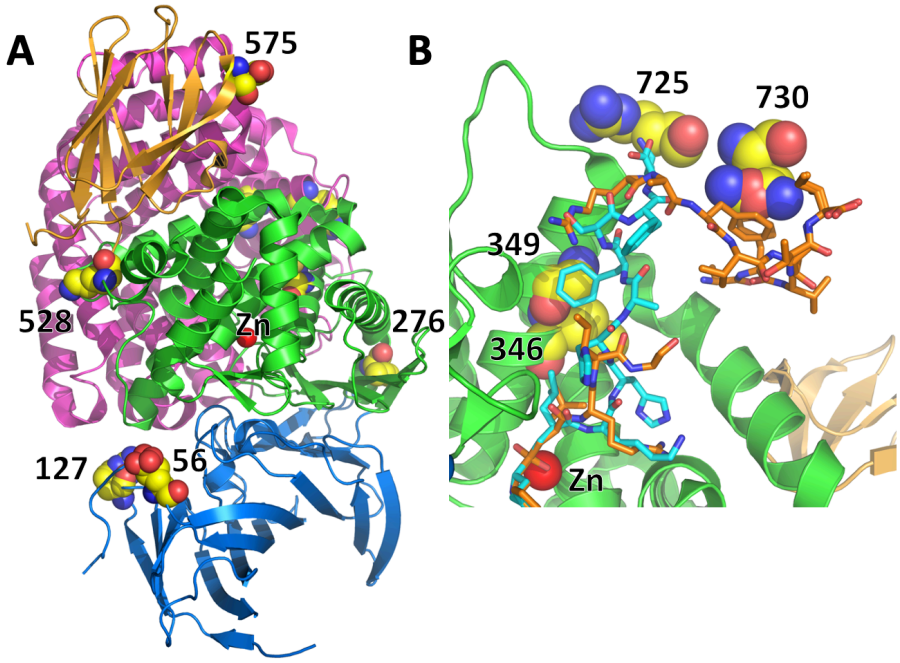

Figure 1: Schematic representations of ERAP1 crystal structure (PDB codes 6RYF and 6RQX) indicating the positions of 9 SNPs. Structure is colored by domain (domain I in blue, domain II in green, domain III in orange and domain IV in magenta). Active site Zn(II) atom is shown as a red sphere). Panel A, polymorphic residues that lie on the outside of the protein are indicated by spheres (carbon=yellow, oxygen=red, nitrogen=blue). Panel B, polymorphic residues that lie in the inside of the substrate binding cavity of the enzyme are shown as spheres (carbon=yellow, oxygen=red, nitrogen=blue). Peptide substrates co-crystallized with ERAP1 are shown in stick representation (blue=10mer, orange=15mer).

We first characterized the ERAP1 allotypes using well established small model dipeptide substrates. The specific activity for the hydrolysis of the substrate Leu-AMC is shown in Figure 2A. There was about a 2-fold spread in specific activities for allotypes 1-9, similar to the differences reported for single SNPs (15, 37). Strikingly, allotype 10 was found to be at least 10-fold less active than the other allotypes (Figure 2A). Rate of hydrolysis was found to be linear with enzyme concentration and this allowed for a more accurate determination of specific activity that confirmed the previous result (Figure 2B and Table 2). Furthermore, the relationship between specific activity and substrate concentration was found to be linear up to 150μM substrate which allowed the calculation of the k_{cat}/K_M ratio for Leu-AMC (Figure 2C and Table 2). Allotype 3 was found to be the most active of all and 10 the least active, being 18-fold less active than allotype 3. Due to spectral interferences using

the Leu-AMC substrate at high concentrations we were unable to perform a reliable full Michaelis-Menten analysis using this substrate (38). Thus, we turned to the chromogenic dipeptide substrate Leu-pNA, for which we were able to obtain full MM curves. Data fit best to an allosteric MM model as previously demonstrated (29, 38) allowing us to calculate relevant enzymatic parameters (Figure 2D and Table 2). This analysis demonstrated that the changes in specific activity between allotypes are due both to changes in affinity for the substrate (indicated by the K_{half} parameter) and to changes in maximal catalytic efficiency (indicated by the V_{max} parameter). As previously shown for Leu-AMC, allotype 10 was much less active in hydrolyzing Leu-pNA, which unfortunately precluded reliable calculation of K_{half} and V_{max} for this allotype.

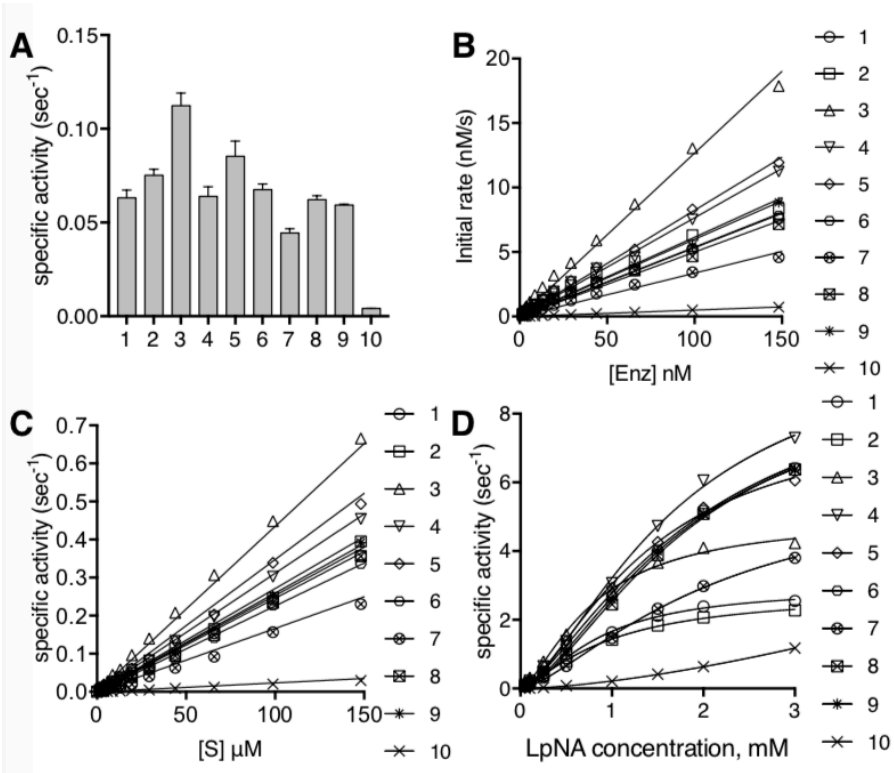

Figure 2: Panel A, specific activity of ERAP1 allotypes 1-10 for the hydrolysis of dipeptidic substrate Leu-AMC. **Panel B**, initial rates of hydrolysis of Leu-AMC by ERAP1 allotypes plotted against enzyme concentration. **Panel C**, specific activity of Leu-AMC hydrolysis versus substrate concentration. **Panel D**, Michaelis-Menten analysis of Leu-pNA hydrolysis by each allotype.

Table 2: Enzymatic parameters of each ERAP1 allotype versus small dipeptide substrates.

	Leu-AMC substrate			Leu-pNA substrate		
Allotype	**Sp. Act. (s^{-1})**	**k_{cat}/K_M (M^{-1} s^{-1})**	**V_{max} (s^{-1})**	**Hill coeff**	**K_{half} (mM)**	**K_{prime} (mM)**
1	0.0505±0.0011	2638 ± 32	2.95 ± 0.1	1.58 ± 0.09	0.85 ± 0.05	0.77 ± 0.08
2	0.0558±0.0025	2529 ± 18	2.8 ± 0.48	1.35 ± 0.31	0.95 ± 0.31	0.94 ± 0.42
3	0.1202±0.0026	4296 ± 49	4.98 ± 0.16	1.49 ± 0.08	0.8 ± 0.05	0.72 ± 0.08
4	0.0736±0.0019	3130 ± 15	10.47 ± 0.56	1.53 ± 0.07	1.7 ± 0.13	2.26 ± 0.19
5	0.0795±0.0014	3519 ± 29	8.07 ± 0.3	1.51 ± 0.06	1.4 ± 0.08	1.65 ± 0.12
6	0.0513±0.001	2273 ± 12	9.94 ± 0.79	1.43 ± 0.09	1.94 ± 0.23	2.57 ± 0.31
7	0.0311±0.0013	1706 ± 33	6.13 ± 0.35	1.46 ± 0.06	2.12 ± 0.17	2.99 ± 0.24
8	0.0463±0.0016	2492 ± 17	9.45 ± 0.35	1.6 ± 0.05	1.87 ± 0.1	2.72 ± 0.15
9	0.0568±0.0017	2747 ± 18	9.33 ± 0.32	1.55 ± 0.05	1.81 ± 0.09	2.52 ± 0.13
10	0.0049±0.0001	240 ± 5	nd	1.53 ± 0.08	nd	nd

Since ERAP1 is responsible for trimming N-terminally extended peptide precursors of antigenic peptides, it is important to explore how the different allotypes process more physiologically relevant long peptides. Indeed, in a recent study, co-crystal structures of ERAP1 with bound 15mer and 10mer peptide analogues revealed that the peptides are trapped, before hydrolysis, in a large internal cavity of the enzyme, while making interactions with residues that line that cavity, and these interactions can drive selectivity (17). To explore how polymorphic variation in the 10 common ERAP1 allotypes can affect trimming of large peptides we measured the rate of N-terminus hydrolysis of two peptides of a similar backbone sequence as the analogues co-crystalized with the enzyme, namely the 15mer LLRIQRGPGRAFVTI and the 10mer LLKHHAFSFK (Figure 3A and B). Similar to the results with the small dipeptidic substrates we recorded a significant variation in trimming rates. Although allotype 10 was again the least efficient in trimming the peptides, the difference from the other allotypes was greatly reduced compared to the small 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 and while allotypes 4 and 6 trimmed the 15mer the fastest, allotypes 5 and 7 trimmed the 10mer the fastest. This results are consistent with a complex landscape of peptide enzyme interactions that drive specificity as previously proposed and suggest that the effect of allotype variation may be substrate-dependent (17, 39).

To better understand the mechanism behind the variation of trimming rates, we utilized a recently developed assay that is suitable for Michaelis-Menten analysis of peptide trimming and follows the trimming of the N-terminal residue of a 9mer antigenic epitope with

the sequence YTAFTIPSI (31). Using this assay we calculated k_{cat} and K_M for each ERAP1 allotype (Figure 3C and Table 3). The turnover rate (k_{cat}), a measure of the maximal catalytic rate of hydrolysis for a specific substrate, varied by only about 2-fold between allotypes 1-9, but was about 10-fold reduced for allotype 10, indicating that this allotype is catalytically inefficient. The Michaelis constant (K_M), a measure of how well the enzyme can recognize and bind the substrate, also varied between allotypes by up to 6-fold with allotypes 1 and 2 being the best in binding the substrate. The ratio k_{cat}/K_M, a measure of the overall catalytic efficiency of the enzyme, varied by a very large margin and up to almost 60-fold, indicating that apparently independent changes in k_{cat} and K_M can synergize to enhance differences in trimming rates for each substrate. Notably, allotype 10 had a 60-fold lower catalytic efficiency compared to allotype 2 because of the combined effect of lower substrate affinity and lower turnover number. Thus, we conclude that polymorphic variation in ERAP1 can greatly affect peptide trimming rates by combining variation in catalytic efficiency with differences in substrate recognition. The latter would suggest that the effect of different allotypes may be strongly substrate dependent.

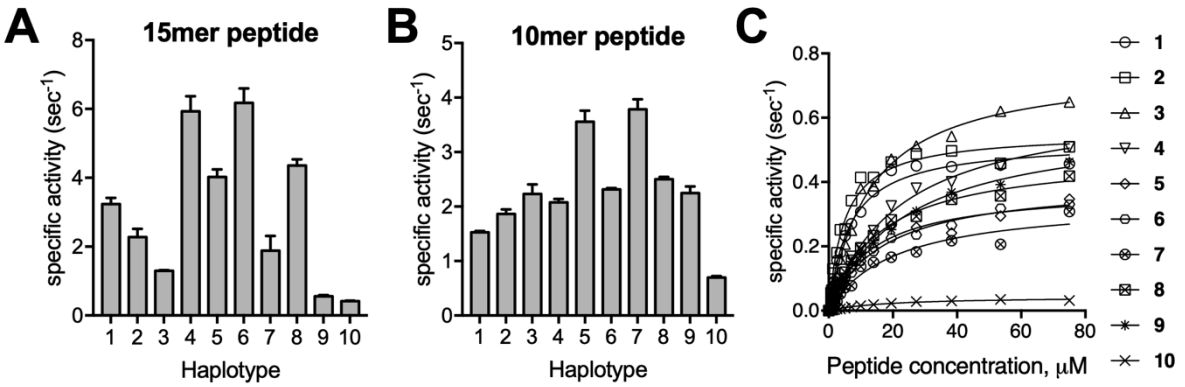

Figure 3: Hydrolysis of peptidic substrates by ERAP1 allotypes. **Panel A,** specific activity of the hydrolysis of 15mer peptide with the sequence LLRIQRGPGRAFVTI. **Panel B,** specific activity of the hydrolysis of 10mer peptide with the sequence LLKHHAFSFK. **Panel C,** Michaelis-Menten analysis of the hydrolysis of the 9mer peptide YTAFTIPSI by ERAP1 allotypes.

Table 3: Calculated parameters from Michaelis-Menten analysis of trimming of the 9mer peptide YTAFTIPSI by ERAP1 allotypes.

	9mer peptide		
Allotype	**k_{cat} (s^{-1})**	**K_M (μM)**	**k_{cat}/K_M (M^{-1} s^{-1})**
1	0.527 ± 0.014	6.5 ± 0.6	80523 ± 7251
2	0.553 ± 0.016	4.9 ± 0.5	113718 ± 12448
3	0.761 ± 0.021	13 ± 1	58480 ± 4601
4	0.655 ± 0.022	22.2 ± 1.8	29463 ± 2537
5	0.4 ± 0.015	14 ± 1.3	28478 ± 2938
6	0.403 ± 0.01	16.8 ± 1.1	23915 ± 1648
7	0.389 ± 0.027	25 ± 3.6	15567 ± 2466
8	0.489 ± 0.013	16 ± 1.1	30614 ± 2207
9	0.589 ± 0.018	23.8 ± 1.6	24717 ± 1826
10	0.051 ± 0.002	23.8 ± 1.4	2134 ± 144

ERAP1 trimming of antigenic epitope precursors in the ER often includes multiple trimming steps before the product is at optimal length for binding onto MHCI. Therefore, which products will accumulate can affect which peptides will bind onto MHCI. To evaluate the effect of ERAP1 allotype variation on multiple trimming reactions we followed the generation of the ovalbumin epitope SIINFEKL from the starting 14mer extended epitope GLEQLESIINFEKL (Figure 4). In all cases the 14mer was catabolized and the 8mer produced but with variable efficiencies. As observed before, allotype 10 was less active in trimming and it was necessary to use at a 10-fold higher concentration to generate appreciable amounts of the SIINFEKL peptide. In all reactions we were able to detect the generation of all possible intermediates (Figure 4A). However, we observed significant differences both between which intermediates accumulated and on the overall rate of generation of the mature epitope (Figure 4B,C). Specifically, allotypes 1-3 accumulated the mature epitope the fastest, whereas allotypes 5 and 7 appeared inefficient in producing the mature epitope to a large degree because of slower trimming of the initial 14mer substrate. 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 10-12 residues can often bind onto MHCI and be immunogenic, distinct accumulation of trimming intermediates by different allotypes could be a contributing factor to differences in immune responses between individuals.

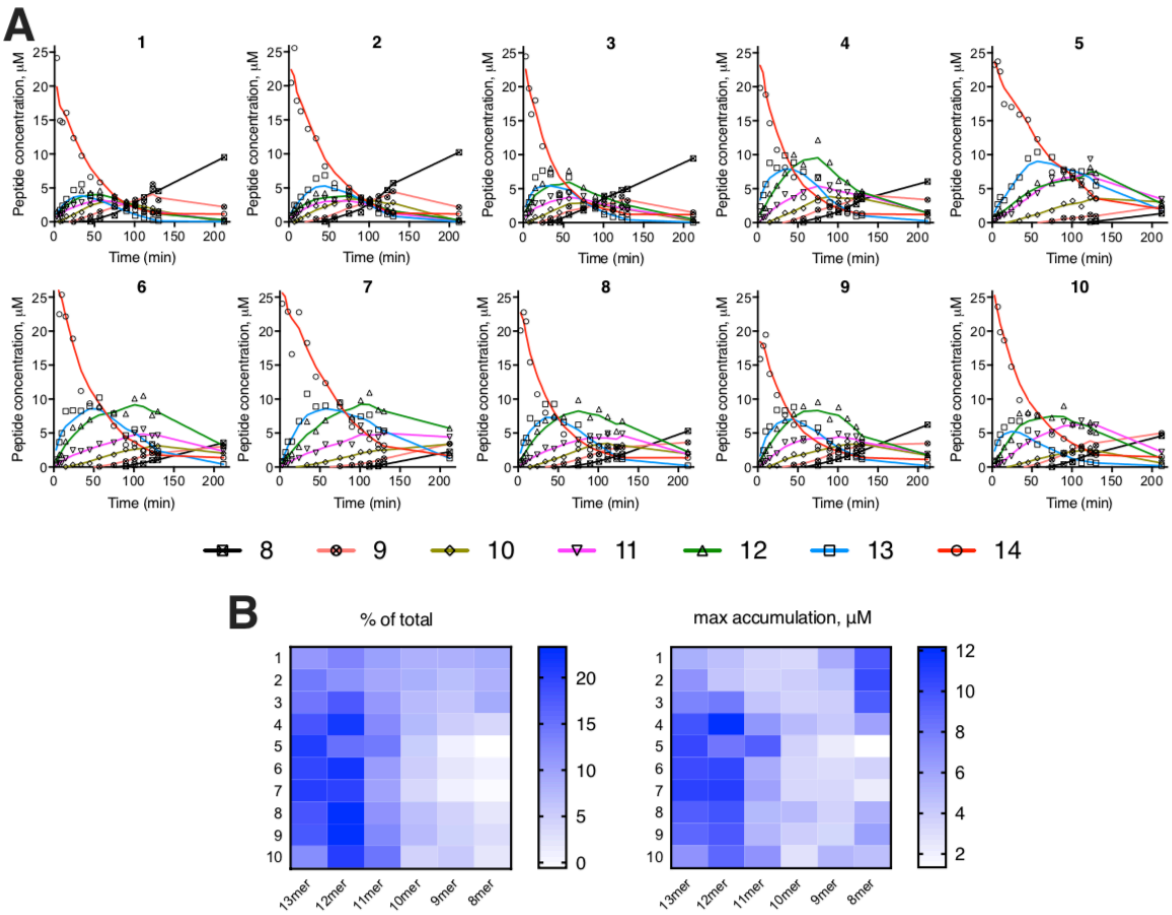

Figure 4: Panel A, trimming of the 14mer peptide GLEQLESIINFEKL, precursor to the ovalbumin epitope SIINFEKL, by ERAP1 allotypes. Each peptide is quantified by Mass Spectrometry and its concentration is plotted as a function of reaction time. Lines correspond to locally weighted scatterplot smoothing (LOWESS) and are only used as a visual aid to follow peptide accumulation. Allotype 10 was used at a 10-fold higher concentration. **Panel B,** left: heatmap indicating accumulation of each peptide intermediate as a fraction of the total amount of peptides in the reaction; right: heatmap showing the maximal accumulation of each peptide intermediate during the reaction.

ERAP1 is an emerging pharmacological target for cancer immunotherapy and the control of inflammatory disease of autoimmune etiologies, including rheumatic conditions such as ankylosing spondylitis (40, 41). Given the wide distribution of common allotypes in the human population, it is crucial to know if inhibitors with clinical potential are able to inhibit all allotypes with efficacy. To explore this, we performed inhibition titrations using the small substrate Leu-AMC and inhibitors DG013A and GSK849, both shown to be active in cellular assays. DG013A is a transition state analogue that targets the active site of the enzyme with high potency (27). GSK849 is a natural product that targets the regulatory site of ERAP1 and while is an apparent activator for small substrates, it inhibits peptide hydrolysis by interfering with C-terminus recognition (31). DG013A was able to inhibit all 10 allotypes

with high potency (pIC_{50} between 7.2 and 7.6) (Figure 5A). GSK849 acted as an apparent activator of small substrate hydrolysis as previously reported (31) and its activity varied significantly between allotypes (pXC_{50} between 4.8 and 6.5) (Figure 5B). GSK849 was, however, an inhibitor of the more physiologically relevant 9mer substrate YTAFTIPSI hydrolysis (Figure 5C). A comparison of the pIC_{50} and pXC_{50} values for the two substrates are shown in Figure 5D. Interestingly, 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 GSK849's ability to inhibit (Figure 5E). Additionally, there was a positive correlation between the pXC_{50} value and the fractional activation observed (Figure 5F). This surprising finding suggested that the regulatory, but not the catalytic site of ERAP1, is sensitive to the allotypic state of the enzyme.

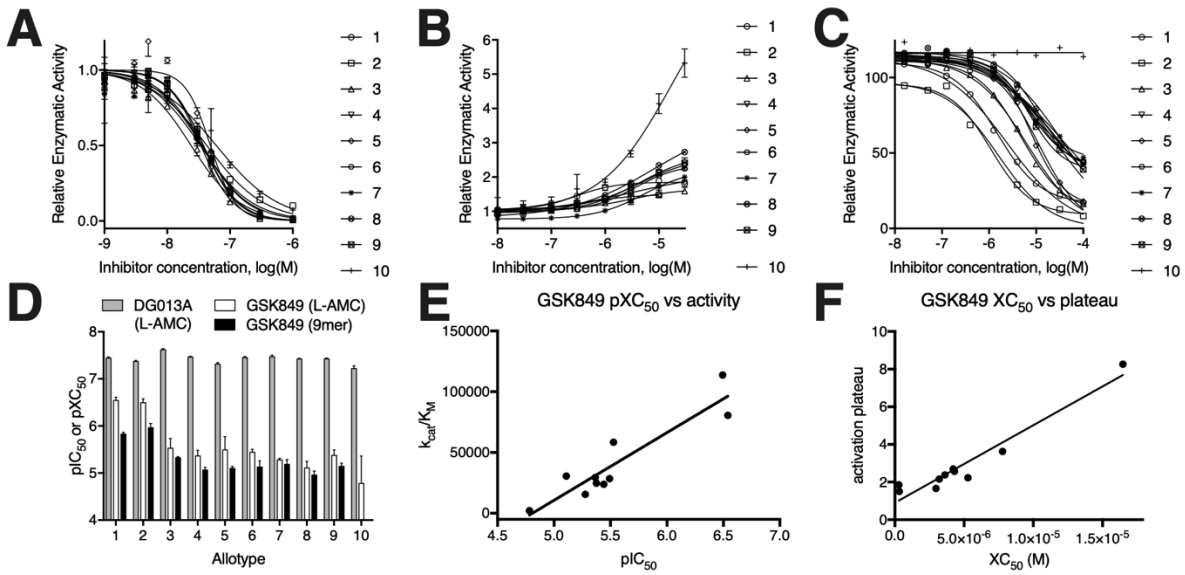

Figure 5: Effect of inhibitors DG013A and GSK849 on the activity of ERAP1 allotypes. **Panel A,** effect of titration of DG013A on the hydrolysis of Leu-AMC by ERAP1 allotypes. **Panel B,** effect of titration of GSK849 on the hydrolysis of Leu-AMC. **Panel C,** ability of GSK849 to inhibit hydrolysis of a 9mer peptide by ERAP1 allotypes. **Panel D,** comparison of pIC_{50} or pXC_{50} values from the titrations shown in panels A-C. **Panel E,** correlation between allotype specific activity and pIC_{50} value for GSK849. **Panel F**, correlation between maximal activity enhancement by GSK849 and XC_{50} of each allotype.

The high variability in enzymatic activity between ERAP1 allotypes that are common in the population suggests that individuals carrying different combinations of allotypes can have an even larger variability of ERAP1 activities in the cell. Assuming that no specific interactions exist between the two alleles, the total enzymatic activity of an individual would

be expected to be the sum of the activity of each allele they carry. 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 in the cell, complicating analysis. Indeed, two recent reports suggested that SNPs can affect ERAP1 expression to a certain degree something that could either exacerbate or ameliorate allotype activity variation (19, 20). 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 taken into account here. Thus, to get insight on how changes in allotype activity can affect the distribution of total ERAP1 activity in the population we calculated the expected total activity of the two alleles carried by individuals. We utilized the measurements of activity for the 9mer substrate since we were able to obtain full MM analyses for this substrate and thus obtain the most reliable catalytic efficiency measurements (k_{cat}/K_M, Figure 3C and Table 3). A plot of expected specific activity versus population frequency is shown in Figure 6A. A bubble chart showing the population frequency and expected enzymatic activity for each possible combination of ERAP1 allotype is shown in Figure 6B. Individuals carrying different common combinations of allotypes are expected to possess a wide area of total ERAP1 activity, covering a range of about 60-fold. Most common allotype combinations fall within a more limited breadth of activities, about 10-fold (cyan region, Figure 6A). Homozygous individuals of allotype 2 are quite frequent in the global population and would be expected to feature the highest total ERAP1 activity (red region, Figure 6A). Combinations of 2 with 8 are also very common and should feature moderate-to-high enzyme activity (magenta region, Figure 6A). 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 ~1.2% of the global population and should result in the lowest ERAP1 activity and may constitute functional knockouts, at least for some peptide substrates. Analysis on sub-populations (Supplemental Figure 1) suggested that significant variability in both activity and frequency distribution is found in different populations, a phenomenon that may signify local host-pathogen balancing selection pressures, a notion previously suggested for individual SNPs (42).

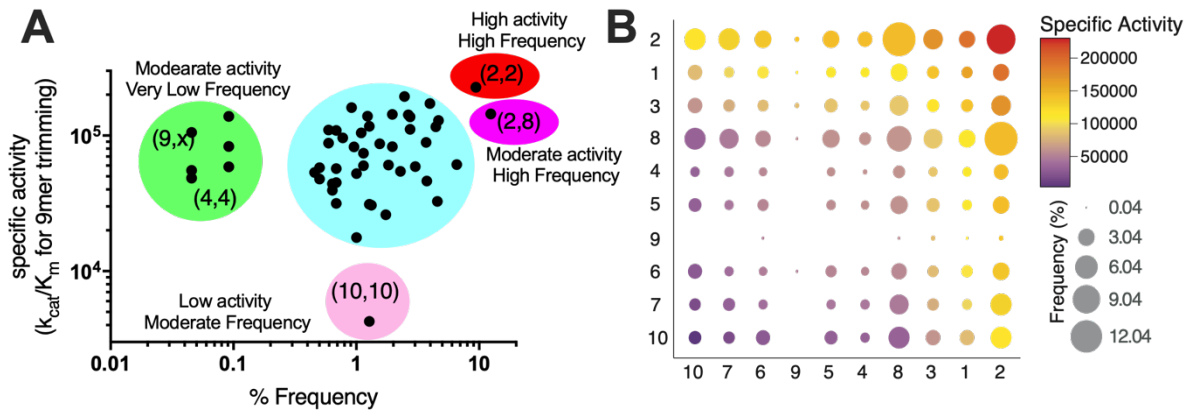

Figure 6: The landscape of ERAP1 activities in the population. **Panel A,** scatterplot showing the correlation between total estimated ERAP1 activity and frequency of allotype combinations found in individuals. Specific allotype combinations are indicated in parenthesis. **Panel B,** bubble chart showing the frequency of allotype combinations color-coded by their estimated activity. Allotypes have been clustered based on activity (highest activity top right, lowest bottom left).

Discussion

A wealth of genetic association studies has linked ERAP1 polymorphic variation to the existence of coding SNPs (8, 43). Follow-up studies extended this work by demonstrating that these SNPs can affect enzyme function, thus providing mechanistic support for observed epistasis between ERAP1 and HLA (15, 16, 25, 35). The effects on enzyme function have been complex and hinted a possible synergism between SNPs that is poorly understood. Accordingly, it has been suggested that most SNPs affect the conformational dynamics of the enzyme and thus influence activity indirectly (18). This notion is consistent with the observation that all 9 SNPs lie away from the catalytic site. Furthermore, functional studies focusing on single SNPs have indicated relatively modest effects on activity and very little is known on the functional interplay of disease-associated SNP combinations in naturally occurring ERAP1 allotypes. This is especially true in cases in which common allotypes harbor SNPs that protect along with SNPs that predispose to disease (22). Thus, a better 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. Complex synergism between SNPs is likely and has already been suggested before and thus it is more physiologically relevant to study ERAP1 allotypes that exist in the population rather than single SNPs in unknown backgrounds that could influence

their effect on enzyme function. Our results here suggest that SNPs can synergize to affect function and that the effect is strongly substrate dependent. This is probably best highlighted by the case of allotype 10 which was found to be up to 60-fold less active than other common allotypes. Such a strong functional change has not been reported before for individual SNPs in ERAP1 and is probably the result of synergism of several SNPs. Specifically, allotype 10 carries the SNP 528R that has been reported by several groups to associate with decreased risk to Ankylosing Spondylitis and Psoriasis and to result in an enzyme with about 2-fold reduced activity (14, 15, 21, 37). In addition, it carries the SNP 575N which has been reported to affect enzymatic activity depending on the polymorphic state of position 528 (44). These two SNPs may synergize in reducing the activity of allotype 10 along with two additional SNPs unique to this allotype: i) 349V, that changes an amino acid that is relatively close to the catalytic site although it is a conservative alteration that has not been reported to affect activity and ii) 725Q, that has been demonstrated to reduce enzymatic activity (45) and lies in the interface between domains II and IV, adjacent to a poorly resolved loop in existing crystal structures and could induce changes in the dynamics of ERAP1's conformational change, similar to position 528 (46). It is possible that synergism between those four SNPs could underlie the greatly reduced catalytic efficiency of allotype 10. Strikingly however, this reduction appears to be partially substrate-dependent. Trimming of the 10mer substrate was less affected by this allotype compared to 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 substrate (17). This observation highlights a motif that can be seen throughout our results. While some changes in activity between activities of 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 peptide sequences both thermodynamically and kinetically and that way contribute to the variability of immune responses (47). 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 specific insights can still be extracted. Several studies have previously demonstrated that 528K results in higher enzymatic activity and proposed that this is due to effects on the conformational dynamics of the enzyme (46). Accordingly, allotypes 1-3 which all carry this SNP, are generally found to be amongst the most active in most assays. The polymorphism 730Q has been suggested to affect activity due to changes in interaction with the C-terminal moiety of the peptide (8, 46). 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 3-10 that carry the 730E at that location, a negative charge, would be expected to be worse in interacting with this substrate, all are less active compared to 1 and 2. Furthermore, allotypes that carry the 127P polymorphism (also reported to be associated with AS (22)) tend to have lower activity, although no specific effects of this SNP have been reported before (8). This polymorphic location has been suggested to lie on a putative substrate exit channel based on a recent ERAP1-peptide analogue crystal structure (17). It is thus possible that the reduced structural flexibility in the mouth of this cannel due to the proline residue may reduce the kinetics of product-substrate exchange leading to a slower apparent activity. Finally, the sensitivity of the regulatory site to inhibition appears to be allotype dependent, a notion that is consistent with a previous study from our group that suggested that the regulatory site communicates with the active site though effects on the conformational dynamics of the enzyme (29). Unfortunately, crystal structures of the common ERAP1 allotypes are not available, making structural explanations for observed functional differences difficult, especially in the view of possible synergism of particular SNPs. Additional structural information on these allotypes would be invaluable in dissecting the mechanistic differences that are behind the functional changes.

Although ERAP2 has been described as an accessory aminopeptidase that supplements ERAP1 activity for some epitope precursors, recent genetic association data have pointed to important roles for ERAP2 in both cancer immunotherapy and autoimmunity (48-50). Because the rs2248374 polymorphism that leads to lack of ERAP2 protein expression (36), 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 (51, 52), since there are no other ER-resident aminopeptidases reported to date, ERAP1 and ERAP2 cumulatively define the spectrum of aminopeptidase activity in the ER. From the combination of our 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 that carry ERAP2 allotypes that lead to no 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 (11). Thus, it appears that some balancing selection may exist that attempts to normalize total aminopeptidase activity in the ER. It may be worthwhile that future population-wide genetic association studies were to combine the ERAP1/2 genotype activity results described here to examine possible correlations between total ER aminopeptidase activity and predisposition to disease.

Although our *in vitro* analysis has the advantage of allowing the accurate determination of fundamental molecular properties, it is limited by its *in vitro* nature since changes in enzymatic activity of ERAP1 are not necessarily directly translated to changes in antigen presentation in cells. Antigen presentation is primarily controlled by peptide binding by the MHCI and selection of peptides by the MHCI may mask some changes in activity in ERAP1. Indeed, some recent studies have suggested that the potential of ERAP1 in influencing antigen presentation is limited by this phenomenon (53, 54). Additionally, effects on protein expression can further complicate interpretation since if a low activity allotype is expressed at a higher level, overall activity can remain the same. However, reported effects on protein expression are relatively small and therefore we did not take them into account for our analysis (19, 20). Most importantly, several recent cell-based studies using proteomic approaches have provided strong support that changes in ERAP1 activity due to polymorphic variation are to a large extent translated to changes in antigen presentation (55). Also, although the expression of ERAP1 mediated by splice interfering variants (19) is in linkage disequilibrium with variants that encode enzymatically distinct allotypes, genetic association studies have shown that for some (HLA-associated) conditions the disease risk is primarily mapped to enzymatic activity (11). Thus, although MHCI binding is the dominant filter, changes in ERAP1 activity are relevant to antigen presentation and probably underlie 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.

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). Jonas J. W. Kuiper is supported by a VENI award from the Netherlands Organization for Scientific Research (N.W.O. project number 016.186.006).

References

1. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. IPD-IMGT/HLA Database. Nucleic Acids Res. 2020;48(D1):D948-D55.
2. Rock KL, Goldberg AL. Degradation of cell proteins and the generation of MHC class I-presented peptides. Annu Rev Immunol. 1999;17:739-79.
3. Blees A, Januliene D, Hofmann T, Koller N, Schmidt C, Trowitzsch S, et al. Structure of the human MHC-I peptide-loading complex. Nature. 2017;551(7681):525-8.
4. Abele R, Tampe R. Modulation of the antigen transport machinery TAP by friends and enemies. FEBS Lett. 2006;580(4):1156-63.
5. Cascio P, Hilton C, Kisselev AF, Rock KL, Goldberg AL. 26S proteasomes and immunoproteasomes produce mainly N-extended versions of an antigenic peptide. EMBO J. 2001;20(10):2357-66.
6. Weimershaus M, Evnouchidou I, Saveanu L, van Endert P. Peptidases trimming MHC class I ligands. Curr Opin Immunol. 2013;25(1):90-6.
7. Strange A, Capon F, Spencer CC, Knight J, Weale ME, Allen MH, et al. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. Nat Genet. 2010;42(11):985-90.
8. Lopez de Castro JA, Alvarez-Navarro C, Brito A, Guasp P, Martin-Esteban A, Sanz-Bravo A. Molecular and pathogenic effects of endoplasmic reticulum aminopeptidases ERAP1 and ERAP2 in MHC-I-associated inflammatory disorders: Towards a unifying view. Mol Immunol. 2016;77:193-204.
9. Cortes A, Pulit SL, Leo PJ, Pointon JJ, Robinson PC, Weisman MH, et al. Major histocompatibility complex associations of ankylosing spondylitis are complex and involve further epistasis with ERAP1. Nat Commun. 2015;6:7146.

10. Ombrello MJ, Kastner DL, Remmers EF. Endoplasmic reticulum-associated amino-peptidase 1 and rheumatic disease: genetics. Curr Opin Rheumatol. 2015;27(4):349-56.
11. Kuiper JJW, Setten JV, Devall M, Cretu-Stancu M, Hiddingh S, Ophoff RA, et al. Functionally distinct ERAP1 and ERAP2 are a hallmark of HLA-A29-(Birdshot) Uveitis. Hum Mol Genet. 2018;27(24):4333-43.
12. Chen L, Fischer R, Peng Y, Reeves E, McHugh K, Ternette N, et al. Critical Role of Endoplasmic Reticulum Aminopeptidase 1 in Determining the Length and Sequence of Peptides Bound and Presented by HLA-B27. Arthritis Rheumatol. 2014;66(2):284-94.
13. Guasp P, Alvarez-Navarro C, Gomez-Molina P, Martin-Esteban A, Marcilla M, Barnea E, et al. The Peptidome of Behcet's Disease-Associated HLA-B*51:01 Includes Two Subpeptidomes Differentially Shaped by Endoplasmic Reticulum Aminopeptidase 1. Arthritis Rheumatol. 2016;68(2):505-15.
14. Sanz-Bravo A, Campos J, Mazariegos MS, Lopez de Castro JA. Dominant Role of the ERAP1 Polymorphism R528K in Shaping the HLA-B27 Peptidome Through Differential Processing Determined by Multiple Peptide Residues. Arthritis Rheumatol. 2015;67(3):692-701.
15. Evnouchidou I, Kamal RP, Seregin SS, Goto Y, Tsujimoto M, Hattori A, et al. 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. 2011;186(4):1909-13.
16. Reeves E, Colebatch-Bourn A, Elliott T, Edwards CJ, James E. Functionally distinct ERAP1 allotype combinations distinguish individuals with Ankylosing Spondylitis. Proc Natl Acad Sci U S A. 2014;111(49):17594-9.
17. Giastas P, Mpakali A, Papakyriakou A, Lelis A, Kokkala P, Neu M, et al. Mechanism for antigenic peptide selection by endoplasmic reticulum aminopeptidase 1. Proc Natl Acad Sci U S A. 2019;52(116):26709-16.
18. Papakyriakou A, Stratikos E. The Role of Conformational Dynamics in Antigen Trimming by Intracellular Aminopeptidases. Frontiers in Immunology. 2017;8(946).
19. Hanson AL, Cuddihy T, Haynes K, Loo D, Morton CJ, Oppermann U, et al. Genetic Variants in ERAP1 and ERAP2 Associated With Immune-Mediated Diseases Influence Protein Expression and the Isoform Profile. Arthritis Rheumatol. 2018;70(2):255-65.
20. Costantino F, Talpin A, Evnouchidou I, Kadi A, Leboime A, Said-Nahal R, et al. ERAP1 Gene Expression Is Influenced by Nonsynonymous Polymorphisms Associated With Predisposition to Spondyloarthritis. Arthritis Rheumatol. 2015;67(6):1525-34.
21. Goto Y, Hattori A, Ishii Y, Tsujimoto M. Reduced activity of the hypertension-associated Lys528Arg mutant of human adipocyte-derived leucine aminopeptidase (A-LAP)/ER-aminopeptidase-1. FEBS Lett. 2006;580(7):1833-8.
22. Alvarez-Navarro C, Lopez de Castro JA. ERAP1 structure, function and pathogenetic role in ankylosing spondylitis and other MHC-associated diseases. Mol Immunol. 2014;57(1):12-21.
23. Wordsworth BP, Vecellio M, Roberts AR. Reply to Reeves et al.: No evidence for rare ERAP1 haplotypes and haplotype combinations in ankylosing spondylitis. Proc Natl Acad Sci U S A. 2017;114(9):E1577.
24. Roberts AR, Appleton LH, Cortes A, Vecellio M, Lau J, Watts L, et al. ERAP1 association with ankylosing spondylitis is attributable to common genotypes rather than rare haplotype combinations. Proc Natl Acad Sci U S A. 2017;114(3):558-61.

25. Reeves E, Wood O, Ottensmeier CH, King EV, Thomas GJ, Elliott T, et al. HPV Epitope Processing Differences Correlate with ERAP1 Allotype and Extent of CD8(+) T-cell Tumor Infiltration in OPSCC. Cancer Immunol Res. 2019;7(7):1202-13.
26. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.
27. Zervoudi E, Saridakis E, Birtley JR, Seregin SS, Reeves E, Kokkala P, et al. Rationally designed inhibitor targeting antigen-trimming aminopeptidases enhances antigen presentation and cytotoxic T-cell responses. Proc Natl Acad Sci U S A. 2013;110(49):19890-5.
28. Ginestet C. ggplot2: Elegant Graphics for Data Analysis. J R Stat Soc a Stat. 2011;174:245-.
29. Giastas P, Neu M, Rowland P, Stratikos E. High-Resolution Crystal Structure of Endoplasmic Reticulum Aminopeptidase 1 with Bound Phosphinic Transition-State Analogue Inhibitor. ACS Med Chem Lett. 2019;10(5):708-13.
30. Stamogiannos A, Maben Z, Papakyriakou A, Mpakali A, Kokkala P, Georgiadis D, et al. Critical Role of Interdomain Interactions in the Conformational Change and Catalytic Mechanism of Endoplasmic Reticulum Aminopeptidase 1. Biochemistry. 2017;56(10):1546-58.
31. Liddle J, Hutchinson JP, Kitchen S, Rowland P, Neu M, Cecconie T, et al. Targeting the Regulatory Site of ER Aminopeptidase 1 Leads to the Discovery of a Natural Product Modulator of Antigen Presentation. J Med Chem. 2020;63(6):3348-58.
32. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434-43.
33. Song P, Li W, Xie J, Hou Y, You C. Cytokine storm induced by SARS-CoV-2. Clin Chim Acta. 2020;509:280-7.
34. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3):e1001779.
35. Evnouchidou I, Birtley J, Seregin S, Papakyriakou A, Zervoudi E, Samiotaki M, et al. A common single nucleotide polymorphism in endoplasmic reticulum aminopeptidase 2 induces a specificity switch that leads to altered antigen processing. J Immunol. 2012;189(5):2383-92.
36. Andres AM, Dennis MY, Kretzschmar WW, Cannons JL, Lee-Lin SQ, Hurle B, et al. Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. PLoS Genet. 2010;6(10):e1001157.
37. Evans DM, Spencer CC, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nat Genet. 2011;43(8):761-7.
38. Nguyen TT, Chang SC, Evnouchidou I, York IA, Zikos C, Rock KL, et al. Structural basis for antigenic peptide precursor processing by the endoplasmic reticulum aminopeptidase ERAP1. Nat Struct Mol Biol. 2011;18(5):604-13.
39. Chang SC, Momburg F, Bhutani N, Goldberg AL. 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. 2005;102(47):17107-12.
40. Stratikos E. Modulating antigen processing for cancer immunotherapy. Oncoimmunology. 2014;3(1):e27568.

41. Reeves E, Islam Y, James E. ERAP1: a potential therapeutic target for a myriad of diseases. Expert Opin Ther Targets. 2020;24(6):535-44.
42. Cagliani R, Riva S, Biasin M, Fumagalli M, Pozzoli U, Lo Caputo S, et al. Genetic diversity at endoplasmic reticulum aminopeptidases is maintained by balancing selection and is associated with natural resistance to HIV-1 infection. Hum Mol Genet. 2010;19(23):4705-14.
43. Agrawal N, Brown MA. Genetic associations and functional characterization of M1 aminopeptidases and immune-mediated diseases. Genes Immun. 2014;15(8):521-7.
44. Martin-Esteban A, Gomez-Molina P, Sanz-Bravo A, Lopez de Castro JA. 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. 2014;289(7):3978-90.
45. Reeves E, Edwards CJ, Elliott T, James E. Naturally occurring ERAP1 haplotypes encode functionally distinct alleles with fine substrate specificity. J Immunol. 2013;191(1):35-43.
46. Stamogiannos A, Koumantou D, Papakyriakou A, Stratikos E. Effects of polymorphic variation on the mechanism of Endoplasmic Reticulum Aminopeptidase 1. Mol Immunol. 2015;67(2 Pt B):426-35.
47. Ayres CM, Abualrous ET, Bailey A, Abraham C, Hellman LM, Corcelli SA, et al. Dynamically Driven Allostery in MHC Proteins: Peptide-Dependent Tuning of Class I MHC Global Flexibility. Front Immunol. 2019;10:966.
48. de Castro JAL, Stratikos E. Intracellular antigen processing by ERAP2: Molecular mechanism and roles in health and disease. Hum Immunol. 2019;80(5):310-7.
49. Lim YW, Chen-Harris H, Mayba O, Lianoglou S, Wuster A, Bhangale T, et al. Germline genetic polymorphisms influence tumor gene expression and immune cell infiltration. Proc Natl Acad Sci U S A. 2018;115(50):E11701-E10.
50. Kuiper JJ, Van Setten J, Ripke S, Van TSR, Mulder F, Missotten T, et al. A genome-wide association study identifies a functional ERAP2 haplotype associated with birdshot chorioretinopathy. Hum Mol Genet. 2014;23(22):6081-7.
51. Zervoudi E, Papakyriakou A, Georgiadou D, Evnouchidou I, Gajda A, Poreba M, et al. Probing the S1 specificity pocket of the aminopeptidases that generate antigenic peptides. Biochem J. 2011;435(2):411-20.
52. Mpakali A, Giastas P, Mathioudakis N, Mavridis IM, Saridakis E, Stratikos E. Structural Basis for Antigenic Peptide Recognition and Processing by Endoplasmic Reticulum (ER) Aminopeptidase 2. J Biol Chem. 2015;290(43):26021-32.
53. Admon A. ERAP1 shapes just part of the immunopeptidome. Hum Immunol. 2019;80(5):296-301.
54. Komov L, Kadosh DM, Barnea E, Milner E, Hendler A, Admon A. 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. 2018;18(12):e1700248.
55. de Castro JAL. How ERAP1 and ERAP2 Shape the Peptidomes of Disease-Associated MHC-I Proteins. Frontiers in Immunology. 2018;9:2463.