Published August 10, 2015, doi:10.4049/jimmunol.1501103

Crystal Structure of Insulin-Regulated Aminopeptidase with Bound Substrate Analogue Provides Insight on Antigenic Epitope Precursor Recognition and Processing

Anastasia Mpakali,* Emmanuel Saridakis,* Karl Harlos,† Yuguang Zhao,† Athanasios Papakyriakou,* Paraskevi Kokkala,*,‡ Dimitris Georgiadis,‡ and Efstratios Stratikos*

Aminopeptidases that generate antigenic peptides influence immunodominance and adaptive cytotoxic immune responses. The mechanisms that allow these enzymes to efficiently process a vast number of different long peptide substrates are poorly understood. In this work, we report the structure of insulin-regulated aminopeptidase, an enzyme that prepares antigenic epitopes for cross-presentation in dendritic cells, in complex with an antigenic peptide precursor analog. Insulin-regulated aminopeptidase is found in a semiclosed conformation with an extended internal cavity with limited access to the solvent. The N-terminal moiety of the peptide is located at the active site, positioned optimally for catalysis, whereas the C-terminal moiety of the peptide is stabilized along the extended internal cavity lodged between domains II and IV. Hydrophobic interactions and shape complementarity enhance peptide affinity beyond the catalytic site and support a limited selectivity model for antigenic peptide selection that may underlie the generation of complex immunopeptidomes. ***The Journal of Immunology*****, 2015, 195: 000–000.**

Cytotoxic adaptive immune responses rely on the recognition on the cell surface of complexes of MHC class I molecules (MHCI) with antigenic peptides. These peptides are generated inside the cell by proteolytic digestion that often includes the trimming of one or more amino acids from the N terminus of antigenic peptide precursors to generate the mature antigenic peptides (1). This trimming is performed by specialized intracellular aminopeptidases that generate or destroy MHCI peptide ligands and can therefore influence the generation of cytotoxic responses. In conventional Ag presentation, somatic cells degrade intracellular proteins and present derived antigenic peptides. In cross-presentation, specialized professional APCs such as dendritic cells (DCs) take up, process, and present external Ags onto MHCI to allow cross-priming of naive $CD8^+$ T cells (2). The cross-presentation pathway is key for the immune defense against many viruses, bacteria, and tumors and can help the immune system avoid immune evasion strategies. Furthermore, cross-presentation is important for inducing cytotoxic responses through vaccination, especially against tumors (3, 4).

Insulin-regulated aminopeptidase (IRAP; EC 3.4.11.3) is a transmembrane enzyme localized in intracellular vesicles, which has been shown to generate antigenic peptides for MHCI for cross-presentation (5, 6). IRAP (also known as placental leucine aminopeptidase and oxytocinase) also has other important biological functions, including the regulation of trafficking of the glucose transporter type 4, the control of oxytocin levels in pregnancy, and the regulation of brain oxytocin and vasopressin levels (7). The soluble domain of IRAP carries the aminopeptidase enzymatic activity and is highly homologous to two other intracellular aminopeptidases that are important for Ag processing, namely endoplasmic reticulum aminopeptidase (ERAP)1 and ERAP2. The three aminopeptidases have been recently classified under the oxytocinase subfamily of M1 aminopeptidases and characterized to be primarily responsible for antigenic peptide N-terminal trimming in the cell (8). ERAP1 has been shown to be a critical editor of antigenic peptides and its activity to influence the immunopeptidome and concomitant cellular cytotoxic responses, as well as immunodominance (9, 10). Similarly, IRAP can generate correct-length antigenic peptides from precursors in vitro, but in distinct patterns compared with ERAP1, suggesting differences in specificity and mechanism (11). Several studies have suggested that Ag processing can vary significantly between cells and inflammatory states, complicating vaccine design (12–14). As a result, understanding the mechanism of antigenic peptide recognition by intracellular aminopeptidases is of great importance. ERAP1 and ERAP2 have been shown to be polymorphic, and several single nucleotide polymorphisms (SNPs) in these enzymes have been associated with predisposition to disease, ranging from viral infections and cancer to autoimmunity. These SNPs have been shown to affect enzymatic

*National Center for Scientific Research Demokritos, Agia Paraskevi, Athens 15310, Greece; †Division of Structural Biology, Wellcome Trust Centre for Human Genetics, Oxford University, Oxford OX3 7BN, United Kingdom; and ‡Department of Chemistry, University of Athens, Athens 15771, Greece

ORCID: 0000-0002-3566-2309 (E.S.).

Received for publication May 13, 2015. Accepted for publication July 13, 2015.

This work was supported by the European Union (European Social Fund) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework: Research Funding Program of the General Secretariat for Research & Technology (Grant ERC-14). The research leading to these results has received support from the European Community's Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (Grant Agreement N°283570). K.H. was supported by the Medical Research Council, and the Wellcome Trust provided administrative support (Grant 075491/Z/04).

The atomic coordinates and structure factors presented in this article have been submitted to the Protein Data Bank under accession numbers 4Z7I and 5C97 for the ligand-bound and ligand-free structure, respectively.

Address correspondence and reprint requests to Efstratios Stratikos, National Center for Scientific Research Demokritos, Agia Paraskevi, Athens 15310, Greece. E-mail address: stratos@rrp.demokritos.gr or stratikos@gmail.com

The online version of this article contains supplemental material.

Abbreviations used in this article: DC, dendritic cell; ERAP, endoplasmic reticulum aminopeptidase; hPhe, homophenylalanine; IRAP, insulin-regulated aminopeptidase; PDB, Protein Data Bank; PEG, polyethylene glycol; SNP, single nucleotide polymorphism; TFA, trifluoroacetic acid.

activity, Ag generation, and cytotoxic responses [reviewed in (15–17)] and to form specific disease-associated functionally distinct allotypes (18, 19). Recently, two SNPs in IRAP have also been reported to associate with psoriasis and ankylosing spondylitis, suggesting that IRAP polymorphic variation may also contribute to the pathogenesis of autoimmunity (20, 21).

To gain insight on the mechanism and determinants of Ag selection in cross-presentation by IRAP, we crystallized IRAP in the absence and presence of a model antigenic peptide and solved the structures to 3.4Å and 3.3Å, respectively. The peptide model was built on electron density extending from the catalytic site of the enzyme, along the central cavity toward the base of domain IV. The overall conformation of IRAP, compared with the known structures of the homologous ERAP1, represents a semiclosed state, which forms an internal cavity with limited external solvent access. The peptide N terminus is anchored at the catalytic site, and its C-terminal half is lodged between domains II and IV, primarily through hydrophobic interactions and space complementarity. Structural and biochemical analysis provide insight on the substrate specificity of IRAP and on the mechanism by which it can process a very large variety of peptide sequences, but also highlight significant mechanistic differences from ERAP1, a finding that may underlie differences in Ag processing in different cells. Structural mapping and functional analysis of two disease-associated IRAP SNPs indicated the one best associated with autoimmunity (A609T) affects enzymatic activity and is located at an interdomain location similar to the location of a well-studied ERAP1 SNP (K528R).

Materials and Methods

Expression vector construction

pFastBac1_IRAP construct for insect cell expression. pFastBac1 vector was obtained from the construct pFastBac1_ERAP2N (22) after digestion with the restrictive endonucleases BssHII/NotI (New England Biolabs). The coding sequence of the soluble part of human IRAP gene along with the signal sequence from adenoviral E3/19K protein was obtained from the construct pcDNA6/myc-HisA (11) with two sequential steps of PCR, to add to the two termini of the cDNA the desired tags and recognition sites for BssHII and NotI. The product of the first PCR served as template for the second PCR. For both reactions, the Pfu Polymerase Kapa Taq HotStart (Kapa Biosystems) was used according to the manufacturer's instructions. The final PCR product contained, at the 5′ end of IRAP cDNA, a BssHII recognition site (5′-GCGCGC-3′) and the 12-aa lobster tropomyosin cDNA leader sequence 5′-AACTCCTAAAAAACCGCCACC-3′, whereas the 3′ end contained a linker sequence (5′-GGATCC-3′), a 3C protease recognition sequence (5′-CTGGAAGTGTTATTCCAGGGGCCC-3′), a 6His-tag (5′-CATCATCACCATCACCAC-3′), a stop codon (TAG), and the recognition site for NotI (5′-GCGGCCGC-3′). All the purification steps of the DNA fragments were performed with electroelution and phenol/chloroform extraction, followed by precipitation of DNA with ethanol/sodium acetate. The final PCR fragment was digested with BssHII/NotI and ligated to the digested pFastBac1 vector with T4 DNA ligase (Promega). The correct cloning was verified by diagnostic digests and sequencing (VBC). The sequences of the primers (5′-3′) are as follows: hsIRAP_FW1, TAAAAAACCGCCACCATGCGGTACATGATC; hsIRAP_FW2, TTGGCGCGCAACTCCTAAAAAACCGCCACC; hsIRAP_REV1, GTGGTGATGGTGATGATGGGATCCCAGCCACCATGTGAGACTTTT; and hsIRAP_REV2, ATAGTTTAGCGGCCGCCTAGTGGTGATGGTGATGATGGGA.

pURD_hsIRAP construct for mammalian expression. The pURD vector was digested with the restriction endonucleases AgeI/Acc65I under the appropriate conditions (23). The cDNA of hsIRAP gene was obtained from the construct pFastBac1_hsIRAP by PCR, using a pair of primers that add the recognition site of AgeI at 5′ end of the cDNA, and the BsiWi site at the 3′ terminus (FW, CTCGACCGGTGCAACAAATGGGAAATTG; REV, CTGAACGTACGCAGCCACCATGTGAGAC). For the PCR, the Pfu DNA polymerase KAPATaq HotStart (Kapa BIosystems) was used according to the manufacturer's directions. The PCR product was purified by electroelution and digested with AgeI/BsiWi restriction endonucleases. Vector and insert were ligated by T4 DNA ligase (Promega). Correct insertion was verified by diagnostic digests and sequencing (VBC). The final construct encodes also for a rhodopsin 1D4 tag at the C terminus that facilitates the purification of the expressed protein. All DNA purifications were performed with electroelution and phenol/chloroform extraction and precipitated with ethanol/sodium acetate.

Site-directed mutagenesis

Mutagenesis reactions to generate the two SNPs in IRAP (A609T and I166M) were performed using the Quickchange II kit (Agilent Technologies), according to the manufacturer's instructions. Primers were designed with the Quickchange Primer Design tool (http://www.genomics.agilent.com) and were HPLC purified. The sequences of primers were (5′- 3′) as follows: A609T FW, GAAATGAGAACCATACTACACCCATCACCGAAGC; A609T REV, GCTTCGGTGATGGGTGTAGTATGGTTCTCATTTC; I166M FW, GTTTCCATGGGCACAGATGAGGCTTCCCACTG; and I166M REV, CAGTGGGAAGCCTCATCTGTGCCCATGGAAAC.

Stable cell line generation and protein expression in mammalian cells

Stable cell lines were created with the help of construct pURD_hsIRAP, because pURD is a stable cell line generation vector (23). Briefly, HEK 293S GnTI$^{(-)}$ cells (23, 24) were cotransfected with pURD_hsPLPAP and a PhiC31 integrase expression vector (pgk-phic31). The polyclonal cell population resulting from puromycin (2 mg/ml) selection was used for the protein production. The secreted protein has additional amino acids from the cloning vector, as follows: N-terminal ETG and C-terminal RTETSQVAPA sequences (last nine residues are Rhodopsin 1D4 tag). For protein purification, the conditioned media were passed through the anti-1D4 tag Ab (University of British Columbia) covalently linked to Sepharose beads at 4°C (CNBr-activated Sepharose 4 Fast Flow; GE Healthcare) and eluted with 1D4 peptide. Protein was further purified by size-exclusion chromatography (Superdex 200 16/60 column; GE Healthcare) in 10 mM Tris-HCl (pH 7.0) and 150 mM NaCl.

Baculovirus construction and protein expression from insect cells

DNA from construct pFastBac1_hsIRAP was transformed into DH10BAc competent cells for transposition into the bacmid, according to the instructions of Bac-to-Bac system (Invitrogen). White colonies containing the recombinant bacmid were selected for DNA isolation, after ensuring by restreak the true white color. One colony was selected and grew overnight, first in small scale and then in medium scale, in Luria Broth medium with the appropriate antibiotics (50 μg/ml kanamycin, 7 μg/ml gentamicin, and 10 μg/ml tetracycline). Next day, the culture was collected and centrifuged at 4000 rpm for 20 min. Cell pellet was resuspended in 50 mM Tris (pH 8.0) and 10 mM EDTA, in the presence of RNase A (100 μg/ml). Cells were then lysed with a buffer 0.2 N NaOH, 1% SDS, and followed the neutralization with 3 M potassium acetate (pH 5.5). After centrifugation to remove the debris, the DNA in the supernatant was precipitated with isopropanol. The pellet was washed with 70% EtOH, air dried, and resuspended in Tris-EDTA buffer (pH 8.0) for 3 d. The correct transposition was verified by PCR, according to the Bac-to-Bac system instructions.

The bacmid DNA was then transfected to insect Sf9 adherent cells, growing in SF900II serum-free medium (Life Technologies). For transfection, the Cellfectin reagent (Invitrogen) was used according to the manufacturer's instructions. Transfection was performed in six-well plates, and 2.5 μg/μl bacmid DNA was used per well. Cells were left growing for 4–6 d in the presence of gentamicin (50 μg/ml). After that period, the culture was collected and centrifuged at 1250 rpm, and the virus was isolated and filtered from the cell supernatant. Virus was further amplified by propagation first in Sf9 adherent cells (P1 virus stock) and then in Hi5 insect cells in suspension (P2 virus stock).

P2 virus stocks (1:50) for wild-type and both mutant hsIRAP proteins, respectively, were used to infect Hi5 insect cells in suspension. Cells were grown for 3 d and then collected and centrifuged twice, to remove cells and keep the supernatant with the expressed secreted protein. The supernatant was then extensively dialyzed against 10 mM sodium phosphate buffer (pH 8.0) and 100 mM sodium chloride. Next day, after adjusting NaP to 50 mM and NaCl to 300 mM, the dialysate was left for 2 h to bind to Ni-NTA agarose beads in the presence of 10 mM imidazole. Elution was performed using increasing concentrations of imidazole. Enzymatic activity was verified by the L-leucine 7-amido-4-methyl coumarin activity assay, whereas purity and concentration of the expressed protein were monitored by SDS-PAGE and Coomassie staining. Most of the protein was eluted at 150 mM imidazole, and therefore the corresponding fractions were pooled together and dialyzed against 10 mM HEPES (pH 7.0) and 100 mM sodium chloride. Protein was stocked at −80°C in aliquots with 10% glycerol.

Ligand synthesis, purification, and characterization

The synthesis of phosphinic pseudodecapeptide DG025 was performed on Rink amide lanterns (8 μmol/pin) by applying a standard solid-phase peptide synthesis Fmoc protocol. A 20% piperidine/dimethylformamide solution was used for the removal of Fmoc group in each cycle of the synthesis. Fmoc-protected amino acids (24 μmol/pin), diisopropylcarbodiimide (24 μmol/pin), and 1-hydroxybenzotriazole (24 μmol/pin) were used for the coupling steps, and each reaction was allowed to proceed for 5 h. For the introduction of phosphinic pseudodipeptidic sequence, the previously described building block Boc-(*R*)-homophenylalanine (hPhe)[PO(OAd)-CH_2]-(*R*,*S*)-LeuOH was used in the last coupling step (25). Coupling of the aforementioned building block (16 μmol/pin) was performed by using the same coupling reagents (24 μmol/pin of each reagent). Deprotection and removal of the final pseudodecapeptide from the solid support were performed by using a solution of trifluoroacetic acid (TFA)/H_2O/triisopropylsilane 95/2.5/2.5 over 2 h at room temperature. The solution of the deprotected peptide was concentrated in vacuo, and the crude product was precipitated in cold diethyl ether. DG025 was obtained after purification by analytical reverse-phase HPLC and characterized by mass spectroscopy [electrospray–mass spectrometry m/z (z = 1): calculated for $[C_{64}H_{93}N_{16}O_{12}P + H]^+$ 1309.70; found: 1309.92].

Enzymatic assays

The aminopeptidase activity of recombinant IRAP was measured by following the change in fluorescent signal produced upon digestion of the substrate L-leucine 7-amido-4-methyl coumarin (Sigma-Aldrich). The fluorescence was measured at 460 nm, whereas the excitation was set at 380 nm. Measurements were performed on a TECAN infinite M200 microplate fluorescence reader and on a QuantaMaster 4 spectrofluorimeter (Photon Technology International, Birmingham, NJ), as previously described (26, 27).

For analysis of the digestion of the LG_nL peptides, 20 μM peptide was incubated with 500 nM IRAP, at 37°C for 2 h, in a 50 mM HEPES (pH 7.0), 100 mM NaCl buffer. Reactions were terminated with the addition of 0.5% (v/v) TFA. The reactions were analyzed in a reverse-phase HPLC (chromolith C-18 column; Merck) by following the absorbance at 220 nm. A linear gradient elution system was used (solvent A: 0.05% TFA; solvent B: 0.05% TFA, 40% acetonitrile). The percentage of the substrate cleaved was calculated by integration of the area under each peptide peak, using appropriate standards. Specific activity was calculated using the GraphPad software, by fitting the results in a one-phase decay equation: $Y = Yo*e^{-kx}$, where x is the reaction time, Y_0 the substrate fraction left intact after $t = 0$ s (constrained as $Y = 1$), Y the substrate fraction left intact after the end of the reaction, and k the reaction rate constant.

Computational Methods

Preparation of the simulation systems. The initial model was based on the x-ray crystal structure of human IRAP complexed with the phosphinic decapeptide. Protein residues 159–1025 of chain A and the corresponding ligand and zinc ion were employed in the calculations. The simulation systems were prepared using the LEaP module of AMBER (v14) package (28). A disulfide bond between Cys^{828} and Cys^{835} was added, the protonation state of the zinc-bound His^{464} and His^{468} was set accordingly, and hydrogen atoms were added. The ff14SB force-field parameters were applied to the protein atoms, and a simple bonded model was employed for the zinc coordination sphere (29). Parameters for the nonstandard residues of the ligand were calculated using the ANTECHAMBER module of AMBER with AM1-BCC charges (30). The complex was solvated in a box of 30,036 TIP3P water molecules and 11 Na^+ ions to neutralize the system.

Molecular dynamics simulations. All molecular dynamics simulations were performed with the GPU version of PMEMD program using periodic boundary conditions (31). A time step of 2.0 fs was used, and the SHAKE algorithm was employed to constrain the bonds connecting hydrogen atoms. The temperature was controlled using a Langevin thermostat with a collision frequency of 2.0 ps^{-1}, and the pressure was regulated at 1 bar using the Berendsen weak-coupling algorithm with a relaxation time of 2.0 ps. Electrostatic interactions were evaluated by means of the Particle Mesh Ewald method with a real space cutoff of 9.0 Å and a direct sum tolerance of 10^{-6}. The system was initially minimized with harmonic positional restraints of 10 Kcal•mol^{-1}•$Å^{-2}$ force constant on the protein backbone atoms. The temperature was then increased from 10 K to 300 K as a linear function of time over the course of 100 ps. The restraints were then removed over four rounds of 50-ps in the isotherman–isobaric (NPT) ensemble by reducing their strength (5.0, 2.0, 1.0, 0.5 Kcal•mol^{-1}•$Å^{-2}$), and an additional 4.5 ns at constant isotropic pressure and temperature of 300 K was carried out. Two independent production runs of 30 ns each were carried out in the NPT ensemble, one without any restraint and another with restraints on the first six residues of the ligand. Snapshots were collected every 2 ps, and the trajectories were processed with the CPPTRAJ module of AmberTools (v14). The structures were clustered using a hierarchical agglomerative approach with a minimum distance between clusters of 1.0 Å, after mass-weighted, root-mean-square deviation fitting of the last four residues of the ligand atoms.

Crystallization and collection of diffraction data

Crystallization trials were performed by sitting drop vapor diffusion in 96-well plates (Greiner Bio-One, Stonehouse, U.K.), using a Cartesian Technologies Microsys MIC4000 liquid-handling robot, incubated at 21°C and periodically inspected using a TAP Biosystems storage vault (TAP, Royston, U.K.).

Purified IRAP at a concentration of 10 mg/ml in 150 mM NaCl and 10 mM HEPES buffer (pH 7.4) was screened for crystallization against various commercially available screens. The protein stock and reservoir solution was mixed at a 1:1 ratio to form a drop of volume 200 nL. Usable crystals were obtained at several conditions of the Morpheus Screen (Molecular Dimensions) (32). In all cases, data were collected at the I03 beamline at the Diamond Light Source UK, equipped with a Pilatus3 6M pixel detector, at a wavelength of 0.976 Å. Data were merged and scaled using the xia2 package (33).

The best dataset collected from an unsoaked IRAP crystal was from the following Morpheus screen condition: 10% (w/v) polyethylene glycol (PEG) of mean m.w. 4000, 20% (v/v) glycerol, 53.4 mM bicine, 46.6 mM Trizma base (pH of the buffer mixture is 8.5), and 0.02 M each of sodium L-glutamate, DL-alanine, glycine, DL-lysine, and DL-serine. The crystal belonged to space group $P2_1$ with a = 69.00 Å, b = 260.15 Å, c = 73.36 Å, and β = 111.61°. The resulting dataset (collected at 100 K) displayed useful data to 3.37 Å resolution, and 5% of the reflections were flagged for R_{free} calculations.

The structure was solved by molecular replacement with MOLREP (34), using the highly homologous ERAP1 open structure (Protein Data Bank [PDB] code 3QNF) as a search model. Two protein molecules were found in the asymmetric unit. Refinement was performed using programs Refmac (35) for the earlier stages and Phenix.refine (36) thereafter. After release of an IRAP structure at a higher resolution (3.02 Å) by a different group (37) (PDB code 4P8Q), the molecular replacement and refinement of our structure were repeated using that as a model, leading to better refinement statistics (R and R_{free} of 21.6 and 28.5%, respectively), but a virtually undistinguishable structure from that obtained from the ERAP1 model. Twenty-five molecules of *N*-acetyl-D-glucosamine, one α-D-mannose, and one water molecule were also modeled.

The best data for the IRAP/DG025 complex were collected from crystals obtained from the following Morpheus screen reservoir condition: 10% (w/v) PEG of mean m.w. 20,000, 20% (v/v) PEG monomethyl ether of mean m.w. 550, 53.4 mM bicine, 46.6 mM Trizma base (the pH of the buffer mixture was 8.5), and 0.02 M each of the following monosaccharides: D-glucose, D-mannose, D-galactose, L-fucose, D-xylose, and *N*-acetyl-D-glucosamine. Data were collected, merged, and scaled together from two IRAP/ligand complex crystals growing in the same drop, soaked for 3 d with 6.5 μM DG025 ligand. The crystals belonged to space group $P2_1$ with a = 68.52 Å, b = 256.35 Å, c = 73.06 Å, and β = 111.58°. The resulting dataset (collected at 100 K) displayed useful data to 3.31 Å resolution, and 5% of the reflections were flagged for R_{free} calculations.

The structure was solved by molecular replacement with MOLREP, using the IRAP structure (PDB code 4P8Q) as a search model. Two protein molecules were found in the asymmetric unit (referred to as chains A and B), with one ligand molecule bound to each. Initial stages of refinement were performed using Refmac, and later stages using alternate cycles of Phenix.refine and inspection by Coot (38). The ligand was built and refined by iterations of manual building by Coot and refinement in Phenix. Noncrystallographic symmetry restraints between the two chains were applied. The entire backbone of the ligand, plus some weaker electron density for most of its side chains, was clearly identifiable in chain A. Density was missing for the side chain of a phenylalanine at position 9, which was therefore not modeled. In addition, extended weak density made the side chain of terminal lysine 10 difficult to assign; therefore, a positioning reasonably compatible with both the electron density and chemical considerations was chosen. On the contrary, only the first five residues of the ligand were assignable to electron density at chain B. The structure converged to R and R_{free} of 21.1 and 27.1%, respectively. In chain A, no density was visible before residue 159 and between residues 640 and 647. In chain B, no density was visible before residue 160, for residue 597, and between residues 639 and 646 and 664 and 666. We also included in the model 27 molecules of *N*-acetyl-D-glucosamine, 1 α-D-mannose, and 16 water molecules.

Crystals of IRAP cocrystalized with DG025 were also obtained by adding a 5-fold excess of the ligand during the last concentration step (15 μM protein, 80 μM ligand) and incubating for 1 h at room temperature before further concentrating the protein to 10 mg/ml. These crystals were obtained under the same crystallization conditions and also diffracted to 3.3Å, but produced

a poorer electron density map (R/R_{free} = 20.3/28.6%). The resulting crystal structure from those crystals was essentially indistinguishable from the one generated from the DG025-soaked IRAP crystals, with a root-mean-square deviation of 0.317Å (all main chain atoms) between the two structures.

Results

Designing a transition state analog of a model antigenic epitope

Although the crystal structures of ERAP1 and ERAP2 have been very valuable in promoting our understanding of intracellular antigenic peptide processing, the lack of a structure with a peptidic ligand has limited our understanding on how this family of aminopeptidases can process a large number of antigenic peptide precursors with diverse sequences and at the same time impose some selection bias that may result in affecting immunodominance. Toward this goal, we designed a model nonhydrolyzable antigenic peptide precursor based on the antigenic epitope SRHHAFSFR from the aggrecan protein, a critical component for cartilage structure and function of the joints, normally presented by the HLA-B*27:05 MHC class I allele. For this model peptide, the first amino acid was hPhe, an optimal amino acid for recognition by IRAP to enhance binding affinity (26), and the Arg residues were substituted with Lys due to synthetic difficulties. The pseudopeptide contains a phosphinic group in the position of the first peptide bond that would be expected to act as a transition state analog for members of the M1 family of aminopeptidases. This change is expected to stabilize the peptide against the enzymatic activity of IRAP and enhance affinity for the active site. Furthermore, being a transition state analog, it can allow the detailed analysis of the catalytic cycle of the enzyme. The structure of the pseudopeptide is shown in Fig. 1. The affinity of this peptide for IRAP was calculated to be in the low nM range (7.7 nM; Fig. 1) and about two orders of magnitude higher than for previously reported phosphinic dipeptide that contains the same two first amino acids, indicating that the presence of additional amino acids contributes to binding onto the enzyme (25).

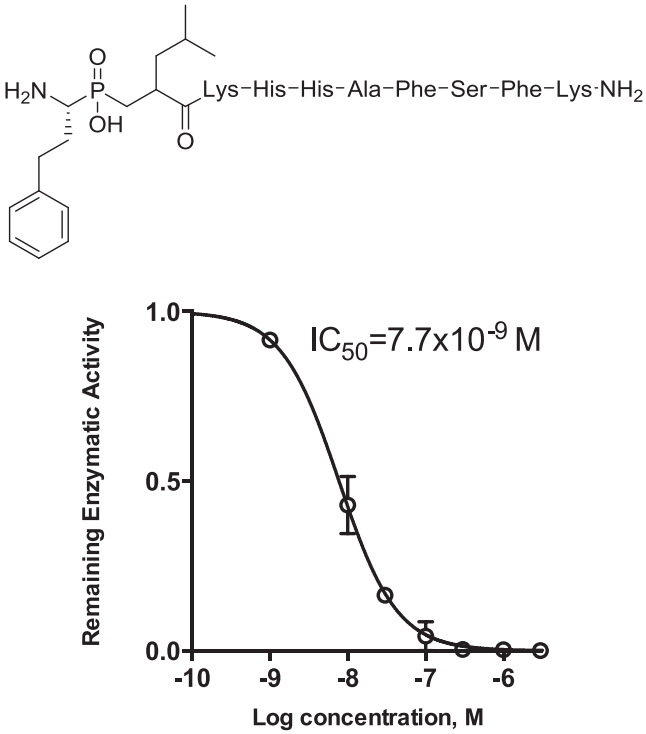


FIGURE 1. *Top*, Chemical structure of the DG025 peptide analog co-crystalized with IRAP. *Bottom*, Titration of DG025 inhibits IRAP activity with an IC_{50} of 7.7 nM.

Crystal structures of IRAP in the presence and absence of pseudopeptide DG025

In a previous developed mammalian expression system, IRAP was found to be hyperglycosylated (11), something that interferes with obtaining good quality crystals for x-ray diffraction. Expression of soluble human IRAP in insect cells resulted in highly active enzyme that was appropriate for biochemical studies, but did not generate good quality crystals. For this reason, we turned to a mammalian expression system that utilizes permanently transfected glycosylation-deficient HEK 293S GnTI(−) cells. Using this expression system, we were able to obtain diffraction quality crystals of IRAP and soak them with DG025. Crystals diffracted to 3.4 Å and 3.3 Å, respectively. Molecular replacement using as a search models the domains I–II and III–IV from the homologous ERAP1 (PDB code 3QNF) allowed the solution of the former crystal structure. In the meantime, a peptide-free structure of IRAP was independently published by another group (PDB code 4P8Q http://www.rcsb.org/pdb/home/home.do) (37), and that structure was used as a search model for solving the IRAP–DG025 complex structure. Crystallographic data and refinement statistics are shown in Table I (37). Both the empty IRAP and IRAP–DG025 complex structures have been deposited to the PDB database (PDB accession codes 5C97 and 4Z7I, respectively). Because our empty IRAP structure was found to be essentially identical to the published one (PDB code 4P8Q, RMSD = 0.439 Å), we focus our analysis on the novel aspects of ligand recognition by IRAP. All comparisons between the empty and ligand-bound IRAP structure throughout this study apply for the 4P8Q structure also. IRAP was also cocrystallized with DG025, resulting in a slightly lower quality electron density map, but essentially the same structure (including the configuration of the ligand) as from the soaked crystals (RMSD = 0.317 Å).

The overall domain organization of IRAP is independent of ligand binding and distinct from ERAP1

IRAP consists of four structural domains, as shown in Fig. 2. Domains II and IV are positioned away from each other so as to form an extended internal cavity that contains the enzymatic active site and has limited access to the external solvent. Compared with the known structures of ERAP1 and ERAP2 (3MDJ, 2YD0, 3QNF, and 3SE6), IRAP was crystallized in an intermediate state between more open ERAP1 (3MDJ, 3QNF) and the closed conformations found in ERAP1 and ERAP2 (2YD0 and 3SE6) (39–41). Surprisingly, the IRAP–DG025 complex was found to be in an identical conformation, with the same interdomain angle in the two IRAP structures (Fig. 2, *top*; 57˚), which is smaller compared with the open conformation of ERAP1 (65–67˚) and larger than the closed conformation of ERAP1 (53˚).

The peptide is anchored at the catalytic site and lodged between domains II and IV of IRAP

The 10-mer peptide was built based on the $|F_o\text{-}F_c|$ unbiased difference map along the internal cavity of IRAP, extending from the catalytic site in domain II toward the hinge region between domains II and IV (Fig. 3A). Very strong density was evident at the N terminus of the peptide, which was easily attributed to canonical binding to the first two residues attached to the phosphinic group, similar to a previously solved phosphinic tripeptide inhibitor in ERAP2 (25). The density was poorer moving toward the C terminus of the peptide, but of sufficient quality to trace the backbone and side chains of the first 6 aa and the backbone up to the C-terminal amino acid. Several refinement cycles improved the map; nevertheless, the side chain of

Table I. Crystallographic data and refinement statistics

PDB Entry Code	4Z7I	5C97
Space group	$P2_1$	$P2_1$
Cell	a = 68.52 Å b = 256.35 Å c = 73.06 Å β = 111.58°	a = 69.00 Å b = 260.15 Å c = 73.36 Å β = 111.61°
Data collection		
Temperature (°K)	100	100
Resolution (Å)	128.18–3.31 (3.40–3.31)*	68.27–3.37 (3.46–3.37)*
Completeness	99.9 (100)*	99.8 (99.6)*
Redundancy	13.8 (13.1)*	6.8 (7.0)*
R_{merge} (%)	7.6 (32.9)*	5.8 (72.0)*
I/σ (I)	20.1 (2.5)*	16.5 (2.7)*
Unique reflections	34,796	33,870
Refinement		
Refinement program	Phenix.refine	Phenix.refine
Resolution (Å)	128.18–3.31 (3.41–3.31)*	68.20–3.37 (3.46–3.37)*
Unique reflections used	34,758 (2,734)*	33,823 (2,643)*
R_{factor} (%)	21.1	21.6
R_{free} (%)	27.1	28.5
RMSD from ideal bond lengths (Å)	0.004	0.014
RMSD from ideal angles (°)	0.991	2.191
Ramachandran statistics		
Non-Gly/Pro residues in most favored regions	81.8%	83.3%
Non-Gly/Pro residues in additionally allowed regions	17.2%	16.0%
Non-Gly/Pro residues in disallowed regions	0.9%	0.7%

*Values in parentheses are for the outermost shell.
R_{factor}, residual factor; R_{free}, unbiased residual factor; R_{merge}, symmetry-related residual factor; RMSD, root-mean-square deviation.

residue 9 (Phe) was missing, and the side chain of residue 10 (Lys) was diffused. As a result, the former was left as a stub alanine residue, and the latter build in the diffused density to account for the only visually possible interaction with a nearby Glu residue (Glu1008).

The orientation of the peptide ligand appears to be defined by two major sets of interactions with IRAP. The canonical orientation of the two N-terminal residues of DG025 is defined by the transition state nature of the ligand and the hydrophobic and aromatic interactions of the hPhe residue in the S1 specificity pocket of the enzyme (and in particular with Phe^{544}). The remaining of the peptide extends toward the base of the cavity, where it is sandwiched between domains II and IV (Fig. 3B). Interestingly, the peptide is found at a location with so limited conformational freedom to exclude entrance by simple diffusion, suggesting that the initial binding may have occurred in a more open conformation of IRAP. Overall, although the peptide interacts with 16 residues of IRAP, no deep specificity pockets are evident that could drive substrate selection in a manner analogous to recognition of antigenic peptide anchor residues by MHC class I molecules (Fig. 4A, Supplemental Fig. 1). Most of the interacting residues belong to domain II, whereas the C-terminal moiety of the peptide interacts with residues of domain IV (Fig. 4A).

The gradually poor density seen for the peptide side chains beyond residue 6 prompted us to examine the possibility of multiple concurrent binding configurations for the C-terminal end of the peptide in the crystal. Molecular dynamics calculations indicated low df for the N terminus of the peptide with an increasing structural heterogeneity approaching the C terminus (Fig. 4B). Even when the first six residues were restrained to the conformation found in the crystal structure, the C-terminal residues demonstrated a large degree of structural plasticity (Fig. 4C). The weaker density toward the C terminus of the peptide, the lack of deep binding pockets, and the molecular dynamics simulations all corroborate the idea that the C-terminal end of the peptide is not specifically recognized by IRAP and that the structure is the result of opportunistic interactions and shape complementarity in the limited space in between domains II and IV. This is in sharp contrast to what has been proposed for the homologous ERAP1, for which a specific recognition of the C terminus of the peptide-ligand has been proposed based on both functional and structural data (39, 42, 43). To further investigate this, we tested whether IRAP has some of the enzymatic properties that are unique to ERAP1 and have been associated to recognition of the C terminus of the peptide. ERAP1 has been shown to trim a polyglycine series of peptides with preference for peptides >10 aa long (39). Hydrolysis of the same peptide series by IRAP indicated that, in contrast to ERAP1, peptides of intermediate length (6–8 aa) are trimmed faster by IRAP (Fig. 5A). Furthermore, ERAP1 has been shown to be activated in the presence of small peptidic products (39, 44). Indeed, we observed a 3- to 4-fold activation of ERAP1 by the SIINFEKL peptide. In contrast, IRAP was not activated, but was rather inhibited by the same peptide consistent with the expected product inhibition of an enzyme (Fig. 5B). We conclude that IRAP does not specifically recognize the C terminus of the peptide substrate and that this observation may associate with the different properties of this enzyme compared with ERAP1. This finding suggests that IRAP generates antigenic peptides with different preferences compared to ERAP1, something consistent with previous observations (11). The repercussions of this to our understanding of cross-presentation versus direct presentation will have to be investigated further.

The N-terminal specificity of IRAP

IRAP has been suggested to process antigenic peptides in specialized endosomal compartments of cross-presenting cells such as DCs. In contrast, ERAP1 and ERAP2 operate together either in the endoplasmic reticulum or in endosomes and have been proposed to form a functional dimer (45). It has been previously shown that ERAP1 shows preference for hydrophobic N termini, ERAP2 for positive charged N termini, and that IRAP combines that specificity, presumably due to its independent function (26). Our peptidic ligand contains a hydrophobic residue (hPhe) that interacts with the hydrophobic base of the S1 pocket of IRAP and in particular with

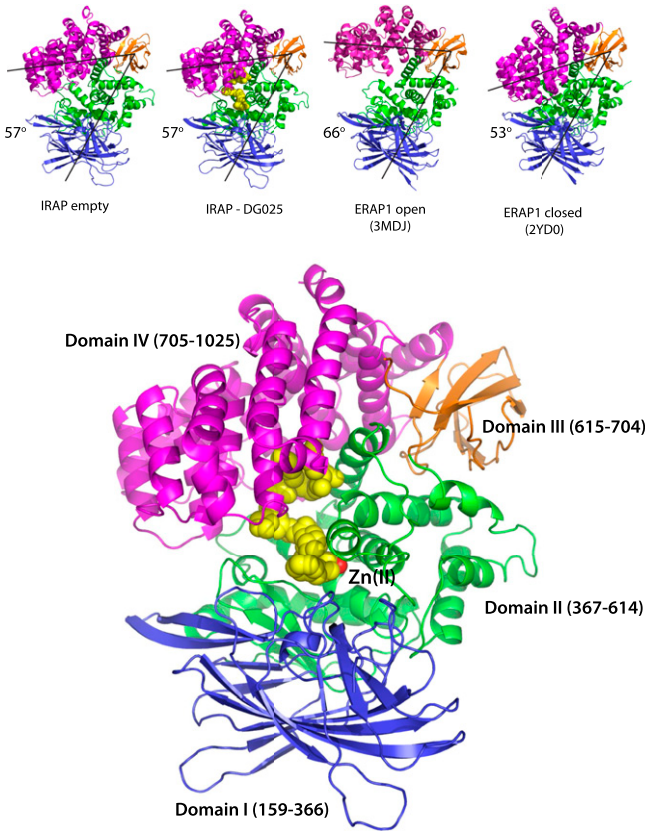


FIGURE 2. Overall structure of the IRAP–DG025 complex. Structural domains are color coded and indicated. DG025 ligand is shown as yellow spheres, and the active site Zn(II) atom as a red sphere. *Top*, Comparison of the domain organization of empty and ligand-bound IRAP with the two known conformations of ERAP1. The interdomain angle is shown, defined as the angle of the centers of mass of domains I and II, domain III, and domain IV.

Phe^{544} consistent with the ability of IRAP to recognize hydrophobic amino acids. To understand the ability of IRAP to recognize positively charged amino acids, we modeled an Arg residue in the IRAP active site, according to the orientation of the Lys residue cocrystallized with ERAP2 (PDB code 3SE6) (41). The arginine residue extends through the S1 pocket and comes in close proximity to Glu^{541}, a nonconserved residue between ERAP1, ERAP2, and IRAP (Fig. 6A). This residue has been shown before by site-directed mutagenesis to affect IRAP preferences (26), and we therefore conclude that it is indeed the key residue affording IRAP its unique S1 preferences.

The GAMEN motif

It has been reported recently that the GAMEN motif of IRAP adopts a different conformation compared with other aminopeptidases of the same family, allowing the better recognition of cyclic peptides (37). Comparison of the empty and peptide-bound IRAP structures shows that this altered orientation of the GAMEN motif is not significantly affected by peptide binding, indicating that it is probably a rigid characteristic of the active site of IRAP. The distinct configuration of the GAMEN motif in IRAP generates additional space around residues 3 and 4 of the bound linear peptide, which could be used for the accommodation of bulkier side chains, possibly affording a broader selectivity for peptide substrates. Still, we found no interactions between residues of the GAMEN motif and DG025, indicating that this structural motif is not a major determinant for linear peptide binding by IRAP (Fig. 6B).

Catalytic Tyr^{549} switches conformation upon ligand binding to stabilize the transition state

M1 aminopeptidases use a tyrosine residue in the active site to stabilize the transition state tetrahedral intermediate. Tyr^{438} in ERAP1 has been proposed to respond to the conformational change of the protein and to reorient toward the Zn(II) atom, facilitating catalysis in the active form of the enzyme (39, 40). The equivalent residue in IRAP, Tyr^{549}, was found at an intermediate orientation in comparison with the open and closed structures of ERAP1, consistent with the intermediate conformation of IRAP compared with the two known conformations of ERAP1 (Fig. 6C). In the presence of the peptidic transition-state analog, however, Tyr^{549} was further reoriented to interact with one of the oxygen atoms of the phosphinic group, which is the equivalent to the oxygen atoms of the substrate in the transition state. This tyrosine conformational switch is probably key to the catalytic cycle of the enzyme and accompanies substrate binding and catalysis cycles. The finding that Tyr^{549} can reorient due to substrate presence in the absence of any protein conformational change suggests that, at least in IRAP, the correct orientation of this

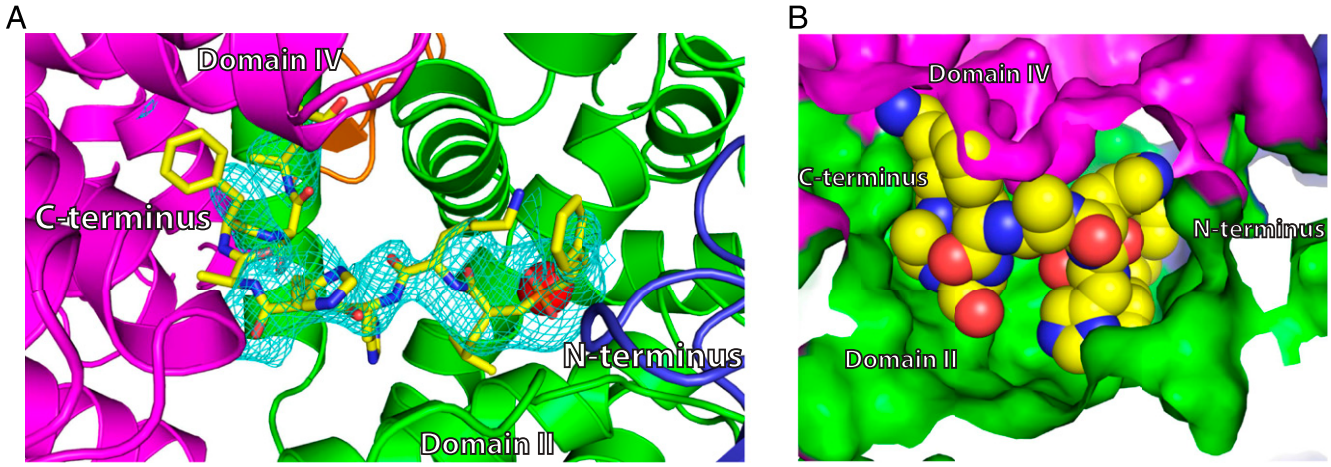


FIGURE 3. (**A**) Electron density interpreted to belong to bound ligand. Cyan mesh indicates the $|F_o\text{-}F_c|$ unbiased electron density extending from the catalytic site of IRAP along the internal cavity toward domain IV. Modeled ligand is shown in yellow. Active site zinc atom is shown as a red sphere. (**B**) Shape complementarity between the bound peptide and the internal cavity of IRAP. IRAP is shown as surface representation colored by domain (domain II is green; domain IV is magenta). Bound peptide is shown in spheres (carbon = yellow, oxygen = red, nitrogen = blue). Note the shape complementarity between the C-terminal half of the peptide and the space in between domains II and IV of IRAP.

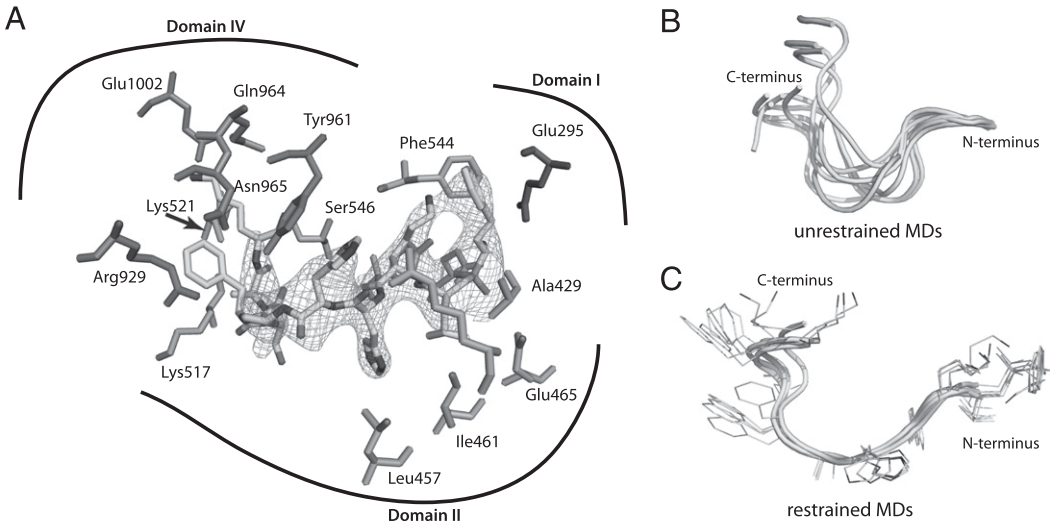


FIGURE 4. (**A**) Refined $2F_o$-F_c electron density map (gray mesh) is shown for the region of bound peptide. Interacting amino acids are shown as sticks, and the three interacting IRAP domains are indicated. (**B** and **C**) Representative structures of the DG025 ligand from (A) 30-ns unrestrained molecular dynamics simulations and (B) 30-ns molecular dynamics simulations, in which the first six residues were restrained. The structures were extracted after agglomerative hierarchical clustering of 15,000 snapshots, which resulted in eight and five clusters, respectively.

catalytic residue is dependent on both the protein conformation and substrate occupation. It is therefore possible that the Tyr conformational switch model proposed for ERAP1 may have to be re-evaluated to account for further conformational rearrangements during the catalytic cycle.

Mapping and functional characterization of IRAP disease-associated SNPs

Polymorphic variation in ERAP1 has been repeatedly associated with disease predisposition, most notably autoimmunity, and shown to affect peptide processing and Ag presentation. Recently, two coding single-nucleotide polymorphisms in IRAP have been associated with predisposition to the autoimmune diseases psoriasis and ankylosing spondylitis, namely rs2303138 coding for the A609T change and rs61752351 coding for the I166M change (20, 21). The latter association was not further supported by linkage dissociation studies that indicated that the association may be due to nearby ERAP1 or ERAP2 SNPs (20). To gain insight on the possible effect of those SNPs on IRAP function, we mapped them on the structure and performed enzymatic analysis (Fig. 7). The I166M polymorphism is located in domain I of the enzyme at the interface with domain II and makes contact with loop 534–538 that is adjacent to the S1 pocket of the enzyme. The equivalent region is found disordered in the open conformation of ERAP1, but folds into an extension of helix-5 in the closed conformation and, therefore, has been proposed to be part of the activation mechanism for ERAP1 (39). Enzymatic analysis indicated a small, but not statistically significant, reduction in enzymatic activity due to that polymorphism, supporting the poor association of this SNP with ankylosing spondylitis (20). In contrast, polymorphism A609T interacts with the hinge domain III of IRAP, similarly to SNP K528R in ERAP1, which has been repeatedly validated for disease association and effects on activity (15, 46). We find that the A609T mutation leads to an almost 2-fold reduction in activity, a finding that provides functional validation for the disease association of this SNP (Fig. 7).

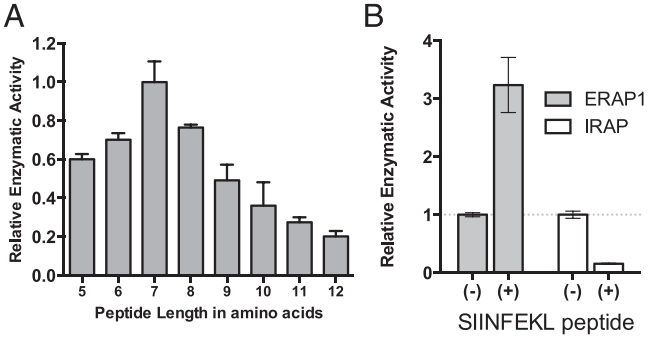


FIGURE 5. (**A**) Relative hydrolysis rate of the N-terminal leucine residue of a series of poly-glycine peptides of the sequence LG_nL, where *n* is 3–10. The *y*-axis indicates the peptide length (5–12 aa). (**B**) Rate of hydrolysis of Leu-AMC substrate in the presence or absence of 100 μM SIINFEKL peptide.

Discussion

Proteolysis of the intracellular proteome can give rise to a very large number of peptides, many of which are transported into the endoplasmic reticulum or specialized endosomal compartments for loading onto MHCI. As a result, aminopeptidases that are tasked with further processing these peptides have to be able to recognize a very large number of sequences effectively. At the same time, many studies have demonstrated that intracellular aminopeptidases such as ERAP1 can greatly affect qualitative aspects of the immunopeptidome, imposing selection bias on final antigenic epitopes. An in vitro study from our group has suggested that IRAP shares this property because it was able to generate antigenic peptides from almost all tested precursors, but with greatly variable efficiencies (11). The IRAP–peptide structure described in this study can serve as a working model to help us understand this property. The peptide N terminus binds at the catalytic site, and the remaining sequence extends along the internal cavity toward domain IV. The C-terminal half of the peptide is stabilized between domains II and IV, in a configuration that is dominated by hydrophobic shape–complementarity interactions. Apart from the expected S1 pocket interactions, we find no evidence for critical anchor residues or deep specificity pockets as it is often seen in MHCI–peptide complexes. Given the very limited polymorphic variation in IRAP, any highly specific interaction with peptide substrate side chains would bias the generated antigenic epitopes to a level that may be immunologically unacceptable because it could exclude immunologically important epitopes. At the same time, the weaker, opportunistic interactions between the peptide and domains II and IV would be expected to

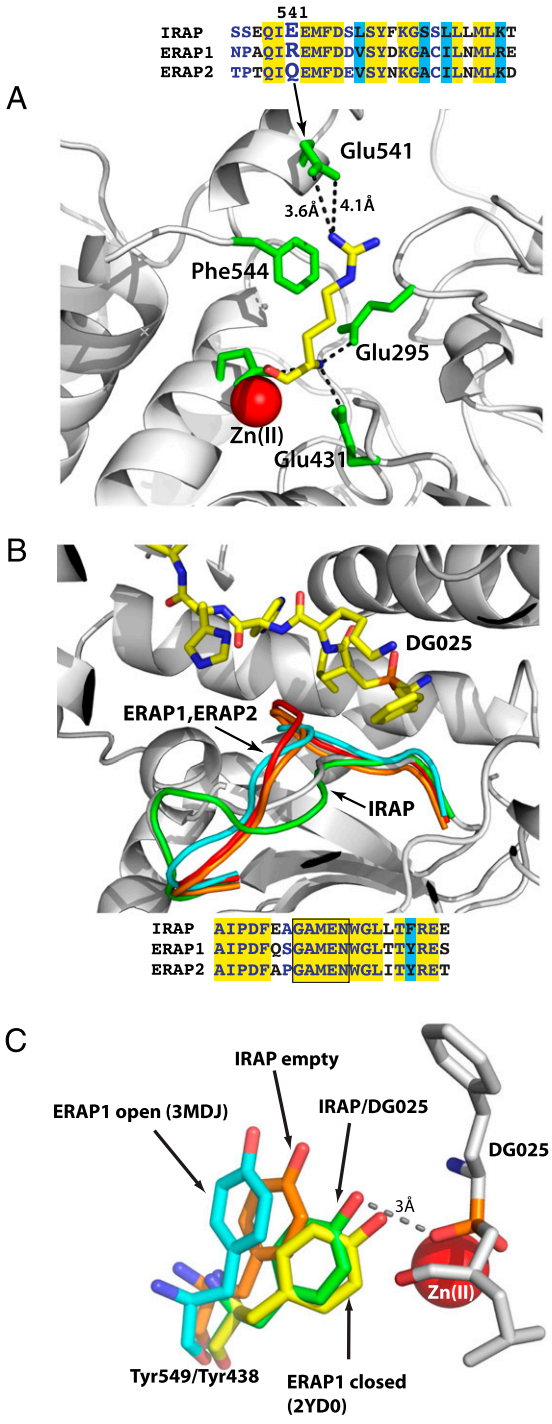


FIGURE 6. (**A**) Model of an arginine residue in the S1 pocket of IRAP. The amino acid was placed in same orientation as the lysine residue found in the crystal structure of ERAP2 (PDB code 3SE6). The arginine side chain can approach Glu^{541} for salt-bridge interactions. The sequence alignment of the region around Glu^{541} for IRAP, ERAP1, and ERAP2 indicating the lack of conservation for that amino acid is also shown. (**B**) Comparison of the configuration of the GAMEN motif in IRAP, ERAP1, and ERAP2. The IRAP–DG025 structure is shown in light gray, and the GAMEN loops of empty IRAP (green, 4Z7I), open conformation ERAP1 (cyan, 3MDJ), and closed conformation ERAP1 (red, 2YD0) and ERAP2 (orange, 4JBS) are shown. Note the lack of interaction between the IRAP GAMEN motif and bound peptide. A sequence alignment showing the location of the GAMEN motif in the three enzymes is also shown. (**C**) Orientation of the active site Tyr in IRAP and ERAP1 structures. In the open and closed conformations of ERAP1, Tyr^{438} occupies two extreme positions (cyan and yellow), whereas the equivalent amino acid in IRAP, Tyr^{549}, is found in an intermediate position (orange and green). Tyr^{549} in IRAP rotates toward the phosphinic peptide bound in the active site.

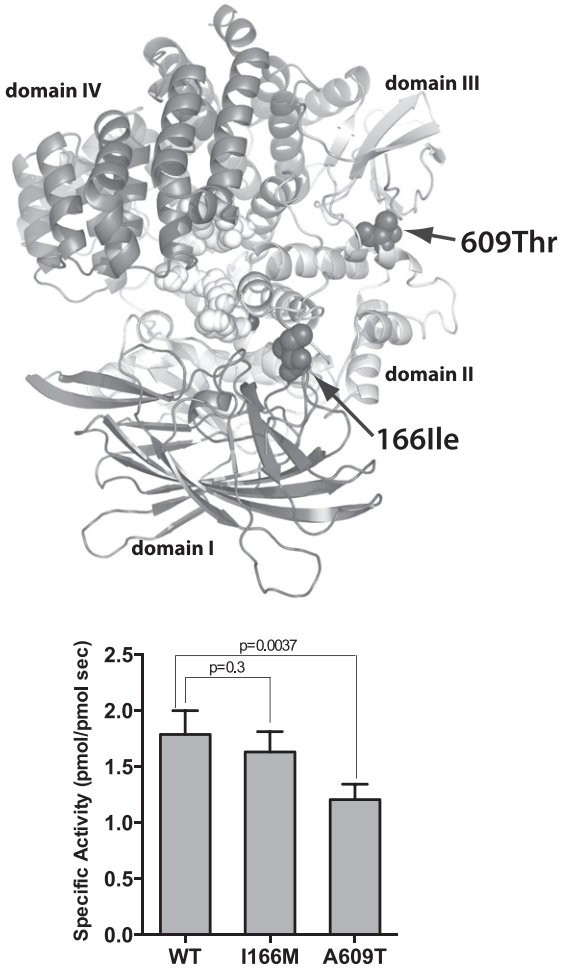


FIGURE 7. *Top*, Mapping of two coding autoimmunity-associated SNPs in IRAP (shown in spheres). Ile^{166} is located in domain I, but makes interactions with domain II near the S1 specificity pocket. Thr^{609} is located in the interface of domain II and hinge-domain III. *Bottom*, specific activity of the two IRAP variants toward the hydrolysis of the fluorigenic substrate L-leucine 7-amido-4-methyl coumarin. Only the A609T variant shows statistically significant deviation compared with the wild type.

afford at least some level of specificity that would bias the generation of the immunopeptidome. The contribution of the C-terminal half of the peptide to the recognition by IRAP can be gauged by comparing the affinity of the inhibitor dipeptide DG002 (IC_{50}~300 nM) (25) to the affinity of the 10-mer DG025 (IC_{50} = 7.7 nM). The transition state analog moiety (DG002) is, as expected, crucial for binding, but the rest of the peptide can enhance binding by almost two orders of magnitude, indicating that the visualized interactions, although individually weak, are cumulatively important. We propose that nonspecific, opportunistic interactions between the peptide and IRAP is an appropriate solution to the problem of having to deal with a vast number of different sequences, but to also confer some degree of selectivity.

The Ag generation function of IRAP has been only found to apply in the case of cross-presentation, whereas ERAP1/ERAP2 play roles in both direct and cross-presentation. It is unclear why a separate enzyme is necessary for this function and what the repercussions of this are for the generation of the immunopeptidome of cross-presenting cells. IRAP can mimic the combined N-terminal selectivity of ERAP1 and ERAP2, but, at least in vitro, generates antigenic peptides in distinct patterns. A recent study showed that DCs, macrophages, and monocytes demonstrate distinct Ag-processing

activities leading to differences in the generation of HIV epitopes and resulting CTL responses (14). Our structural and biochemical analysis suggests that ERAP1 and IRAP have important differences on the mechanism of antigenic peptide recognition. IRAP has different length selection properties, is not activated by small peptides, and appears not to specifically recognize the C terminus of the peptide like ERAP1 has been proposed to do. These differences in mechanism may lead to changes in the mature epitopes produced through distinct pathways of cross-presentation as well as between cross-presentation and direct presentation. The immunological repercussions of this observation are not clear at this point. Further structural analysis of complexes between antigenic peptides and IRAP, ERAP1, or ERAP2 may be necessary to elaborate how each enzyme recognizes its peptide substrates.

Coding SNPs in ERAP1 have been repeatedly associated with predisposition to disease, most notably autoimmunity. The most commonly discovered SNP is K528R, a position located in the interface of domains II and III that has been shown to reduce ERAP1 activity and affect intracellular Ag processing as well as the immunopeptidome (18, 19, 44, 46). Two recent studies have implicated two IRAP SNPs in the pathogenesis of autoimmune disease (20, 21). SNP I166M was found to only weakly associate after linkage disequilibrium correction for SNPs in ERAP1 and ERAP2, whereas the association with SNP A609T is more robust. Our analysis suggests that A609T affects enzymatic activity in a manner similar to the ERAP1 K528R and is located in an equivalent position in the structure. This result provides functional validation for this disease association and suggests that further functional studies on a possible role of IRAP activity on the pathogenesis of autoimmunity are warranted.

In summary, we have analyzed two structures of IRAP, an aminopeptidase important for cross-presentation and adaptive immune response. To our knowledge, the IRAP–DG025 complex is the first reported crystal structure of an Ag-processing aminopeptidase with a model antigenic peptide precursor, and thus represents an important first step for understanding antigenic epitope selection and its role in the pathogenesis of human disease.

Acknowledgments

We thank Dr. Irene M. Mavridis for valuable advice on solving the crystal structures. We also thank the staff of Beamline I03 at the Diamond Light Source Synchrotron for technical support.

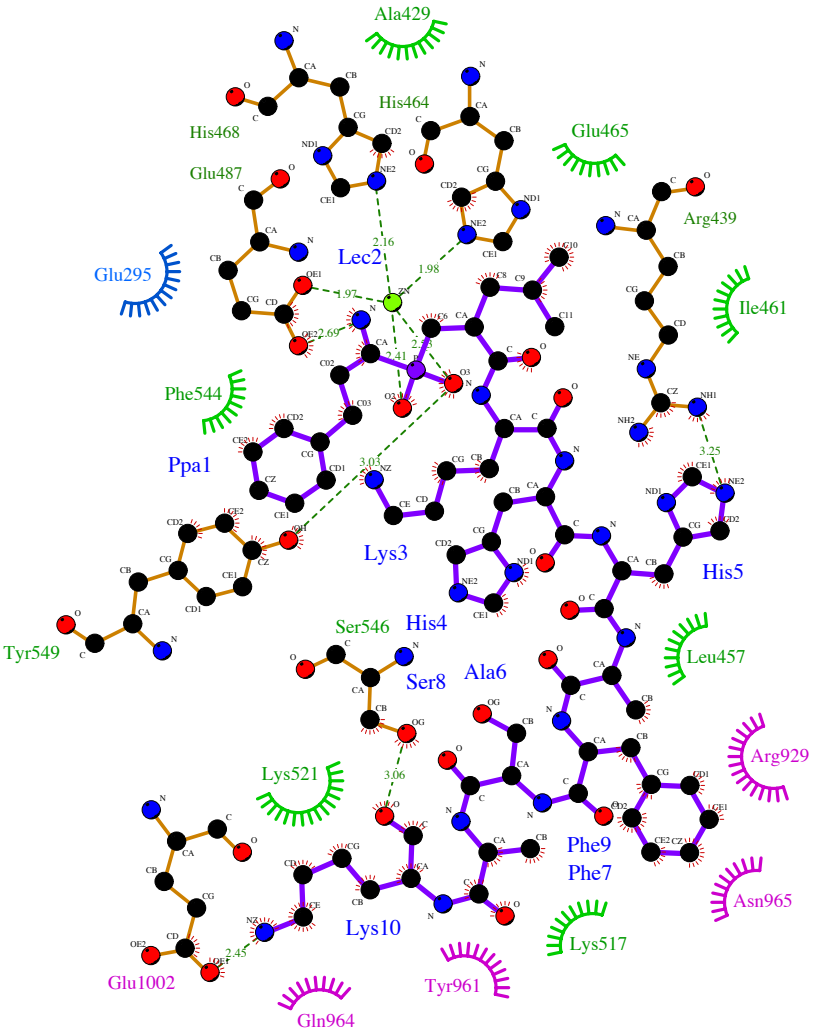
Disclosures

The authors have no financial conflicts of interest.

References

1. Weimershaus, M., I. Evnouchidou, L. Saveanu, and P. van Endert. 2013. Peptidases trimming MHC class I ligands. *Curr. Opin. Immunol.* 25: 90–96.
2. Rock, K. L., and L. Shen. 2005. Cross-presentation: underlying mechanisms and role in immune surveillance. *Immunol. Rev.* 207: 166–183.
3. Fehres, C. M., W. W. Unger, J. J. Garcia-Vallejo, and Y. van Kooyk. 2014. Understanding the biology of antigen cross-presentation for the design of vaccines against cancer. *Front. Immunol.* 5: 149.
4. Shortman, K., M. H. Lahoud, and I. Caminschi. 2009. Improving vaccines by targeting antigens to dendritic cells. *Exp. Mol. Med.* 41: 61–66.
5. Saveanu, L., O. Carroll, M. Weimershaus, P. Guermonprez, E. Firat, V. Lindo, F. Greer, J. Davoust, R. Kratzer, S. R. Keller, et al. 2009. IRAP identifies an endosomal compartment required for MHC class I cross-presentation. *Science* 325: 213–217.
6. Saveanu, L., J. Babdor, M. Lawand, and P. van Endert. 2013. Insulin-regulated aminopeptidase and its compartment in dendritic cells. *Mol. Immunol.* 55: 153–155.
7. Hattori, A., and M. Tsujimoto. 2013. Endoplasmic reticulum aminopeptidases: biochemistry, physiology and pathology. *J. Biochem.* 154: 219–228.
8. Tsujimoto, M., and A. Hattori. 2005. The oxytocinase subfamily of M1 aminopeptidases. *Biochim. Biophys. Acta* 1751: 9–18.
9. Hammer, G. E., F. Gonzalez, E. James, H. Nolla, and N. Shastri. 2007. In the absence of aminopeptidase ERAAP, MHC class I molecules present many unstable and highly immunogenic peptides. *Nat. Immunol.* 8: 101–108.
10. York, I. A., M. A. Brehm, S. Zendzian, C. F. Towne, and K. L. Rock. 2006. Endoplasmic reticulum aminopeptidase 1 (ERAP1) trims MHC class I-presented peptides in vivo and plays an important role in immunodominance. *Proc. Natl. Acad. Sci. USA* 103: 9202–9207.
11. Georgiadou, D., A. Hearn, I. Evnouchidou, A. Chroni, L. Leondiadis, I. A. York, K. L. Rock, and E. Stratikos. 2010. Placental leucine aminopeptidase efficiently generates mature antigenic peptides in vitro but in patterns distinct from endoplasmic reticulum aminopeptidase 1. *J. Immunol.* 185: 1584–1592.
12. Weimershaus, M., S. Maschalidi, F. Sepulveda, B. Manoury, P. van Endert, and L. Saveanu. 2012. Conventional dendritic cells require IRAP-Rab14 endosomes for efficient cross-presentation. *J. Immunol.* 188: 1840–1846.
13. Segura, E., A. L. Albiston, I. P. Wicks, S. Y. Chai, and J. A. Villadangos. 2009. Different cross-presentation pathways in steady-state and inflammatory dendritic cells. *Proc. Natl. Acad. Sci. USA* 106: 20377–20381.
14. Dinter, J., P. Gourdain, N. Y. Lai, E. Duong, E. Bracho-Sanchez, M. Rucevic, P. H. Liebesny, Y. Xu, M. Shimada, M. Ghebremichael, et al. 2014. Different antigen-processing activities in dendritic cells, macrophages, and monocytes lead to uneven production of HIV epitopes and affect CTL recognition. *J. Immunol.* 193: 4322–4334.
15. Alvarez-Navarro, C., and J. A. López de Castro. 2014. ERAP1 structure, function and pathogenetic role in ankylosing spondylitis and other MHC-associated diseases. *Mol. Immunol.* 57: 12–21.
16. Fierabracci, A., A. Milillo, F. Locatelli, and D. Fruci. 2012. The putative role of endoplasmic reticulum aminopeptidases in autoimmunity: insights from genomic-wide association studies. *Autoimmun. Rev.* 12: 281–288.
17. Stratikos, E., A. Stamogiannos, E. Zervoudi, and D. Fruci. 2014. A role for naturally occurring alleles of endoplasmic reticulum aminopeptidases in tumor immunity and cancer pre-disposition. *Front. Oncol.* 4: 363.
18. Reeves, E., A. Colebatch-Bourn, T. Elliott, C. J. Edwards, and E. James. 2014. Functionally distinct ERAP1 allotype combinations distinguish individuals with ankylosing spondylitis. *Proc. Natl. Acad. Sci. USA* 111: 17594–17599.
19. Reeves, E., C. J. Edwards, T. Elliott, and E. James. 2013. Naturally occurring ERAP1 haplotypes encode functionally distinct alleles with fine substrate specificity. *J. Immunol.* 191: 35–43.
20. International Genetics of Ankylosing Spondylitis Consortium (IGAS). 2013. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat. Genet.* 45: 730–738.
21. Cheng, H., Y. Li, X. B. Zuo, H. Y. Tang, X. F. Tang, J. P. Gao, Y. J. Sheng, X. Y. Yin, F. S. Zhou, C. Zhang, et al. 2014. Identification of a missense variant in LNPEP that confers psoriasis risk. *J. Invest. Dermatol.* 134: 359–365.
22. Evnouchidou, I., J. Birtley, S. Seregin, A. Papakyriakou, E. Zervoudi, M. Samiotaki, G. Panayotou, P. Giastas, O. Petrakis, D. Georgiadis, et al. 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.
23. Zhao, Y., T. Malinauskas, K. Harlos, and E. Y. Jones. 2014. Structural insights into the inhibition of Wnt signaling by cancer antigen 5T4/Wnt-activated inhibitory factor 1. *Structure* 22: 612–620.
24. Reeves, P. J., N. Callewaert, R. Contreras, and H. G. Khorana. 2002. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl. Acad. Sci. USA* 99: 13419–13424.
25. Zervoudi, E., E. Saridakis, J. R. Birtley, S. S. Seregin, E. Reeves, P. Kokkala, Y. A. Aldhamen, A. Amalfitano, I. M. Mavridis, E. James, et al. 2013. Rationally designed inhibitor targeting antigen-trimming aminopeptidases enhances antigen presentation and cytotoxic T-cell responses. *Proc. Natl. Acad. Sci. USA* 110: 19890–19895.
26. Zervoudi, E., A. Papakyriakou, D. Georgiadou, I. Evnouchidou, A. Gajda, M. Poreba, G. S. Salvesen, M. Drag, A. Hattori, L. Swevers, et al. 2011. Probing the S1 specificity pocket of the aminopeptidases that generate antigenic peptides. *Biochem. J.* 435: 411–420.
27. Evnouchidou, I., M. J. Berardi, and E. Stratikos. 2009. A continuous fluorigenic assay for the measurement of the activity of endoplasmic reticulum aminopeptidase 1: competition kinetics as a tool for enzyme specificity investigation. *Anal. Biochem.* 395: 33–40.
28. Case, D. A., T. E. Cheatham, III, T. Darden, H. Gohlke, R. Luo, K. M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods. 2005. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26: 1668–1688.
29. Papakyriakou, A., G. A. Spyroulias, E. D. Sturrock, E. Manessi-Zoupa, and P. Cordopatis. 2007. Simulated interactions between angiotensin-converting enzyme and substrate gonadotropin-releasing hormone: novel insights into domain selectivity. *Biochemistry* 46: 8753–8765.
30. Wang, J., R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. 2004. Development and testing of a general amber force field. *J. Comput. Chem.* 25: 1157–1174.
31. Salomon-Ferrer, R., A. W. Gotz, D. Poole, S. Le Grand, and R. C. Walker. 2013. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* 9: 3878–3888.
32. Gorrec, F. 2009. The MORPHEUS protein crystallization screen. *J. Appl. Cryst.* 42: 1035–1042.
33. Winter, G. 2010. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Cryst.* 43: 186–190.
34. Vagin, A., and A. Teplyakov. 2010. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* 66: 22–25.

35. Murshudov, G. N., A. A. Vagin, and E. J. Dodson. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53: 240–255.
36. Adams, P. D., P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66: 213–221.
37. Hermans, S. J., D. B. Ascher, N. C. Hancock, J. K. Holien, B. J. Michell, S. Y. Chai, C. J. Morton, and M. W. Parker. 2015. Crystal structure of human insulin-regulated aminopeptidase with specificity for cyclic peptides. *Protein Sci.* 24: 190–199.
38. Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60: 2126–2132.
39. Nguyen, T. T., S. C. Chang, I. Evnouchidou, I. A. York, C. Zikos, K. L. Rock, A. L. Goldberg, E. Stratikos, and L. J. Stern. 2011. Structural basis for antigenic peptide precursor processing by the endoplasmic reticulum aminopeptidase ERAP1. *Nat. Struct. Mol. Biol.* 18: 604–613.
40. Kochan, G., T. Krojer, D. Harvey, R. Fischer, L. Chen, M. Vollmar, F. von Delft, K. L. Kavanagh, M. A. Brown, P. Bowness, et al. 2011. Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming. *Proc. Natl. Acad. Sci. USA* 108: 7745–7750.
41. Birtley, J. R., E. Saridakis, E. Stratikos, and I. M. Mavridis. 2012. The crystal structure of human endoplasmic reticulum aminopeptidase 2 reveals the atomic basis for distinct roles in antigen processing. *Biochemistry* 51: 286–295.
42. Chang, S. C., F. Momburg, N. Bhutani, and A. L. Goldberg. 2005. The ER aminopeptidase, ERAP1, trims precursors to lengths of MHC class I peptides by a "molecular ruler" mechanism. *Proc. Natl. Acad. Sci. USA* 102: 17107–17112.
43. Gandhi, A., D. Lakshminarasimhan, Y. Sun, and H. C. Guo. 2011. Structural insights into the molecular ruler mechanism of the endoplasmic reticulum aminopeptidase ERAP1. *Sci. Rep.* 1: 186.
44. Evnouchidou, I., R. P. Kamal, S. S. Seregin, Y. Goto, M. Tsujimoto, A. Hattori, P. V. Voulgari, A. A. Drosos, A. Amalfitano, I. A. York, and E. Stratikos. 2011. Cutting edge: 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.
45. Evnouchidou, I., M. Weimershaus, L. Saveanu, and P. van Endert. 2014. ERAP1-ERAP2 dimerization increases peptide-trimming efficiency. *J. Immunol.* 193: 901–908.
46. Sanz-Bravo, A., J. Campos, M. S. Mazariegos, and J. A. López de Castro. 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.



Supplemental Figure 1: LIGPLOT diagram showing interactions between the peptide and amino acids in IRAP. Amino acids are colored by domain according to the color-coding of Figure 2.