# Advancing our knowledge of antigen processing with computational modelling, structural biology, and immunology

Steven Turner1,2, Jonathan W. Essex2, Tim Elliott3

1. Faculty of Medicine, University of Southampton, Southampton, UK
2. School of Chemistry, University of Southampton, Southampton UK
3. Oxford Cancer Centre for Immuno-oncology, and CAMS-Oxford Institute, Nuffield Department of Medicine, University of Oxford, UK.

## Abstract

Antigen processing is an immunological mechanism by which intracellular peptides are transported to the cell surface while bound to Major Histocompatibility Complex molecules, where they can be surveyed by circulating CD8+ or CD4+ T-cells, potentially triggering an immunological response. The antigen processing pathway is a complex multistage filter that refines a huge pool of potential peptide ligands derived from protein degradation into a smaller ensemble for surface presentation. Each stage presents unique challenges due to the number of ligands, the polymorphic nature of MHC and other protein constituents of the pathway and the nature of the interactions between them. Predicting the ensemble of displayed peptide antigens, as well as their immunogenicity, is critical for improving T cell vaccines against pathogens and cancer. Our predictive abilities have always been hindered by an incomplete empirical understanding of the antigen processing pathway. In this review, we highlight the role of computational and structural approaches in improving our understanding of antigen processing, including structural biology, computer simulation, and machine learning techniques, with a particular focus on the MHC-I pathway.

## Abbreviations

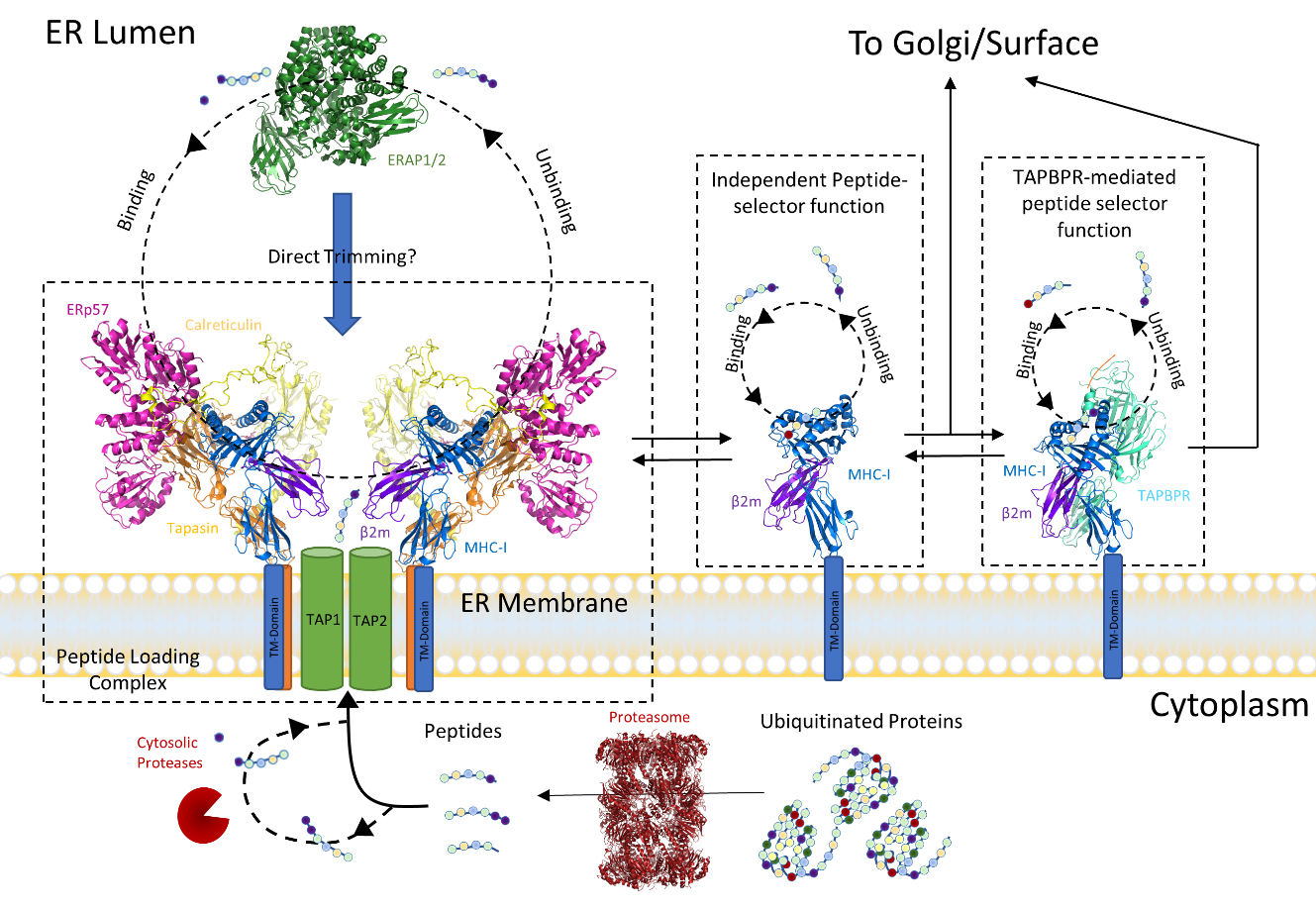
MD = Molecular Dynamics, MHC = Major Histocompatibility Complex, HLA = Human Leukocyte Antigen, TAP = Transporter associated with Antigen Processing, PBG = Peptide Binding Groove, β2m = Beta-2-microglobulin, PLC = Peptide Loading Complex, aa = amino acid

## Introduction

Activation of the T-cell mediated adaptive arm of the immune system relies upon transport of potential antigens within our cells to the cell surface. It is a longstanding challenge of the field to generate predictive models that determine which antigens successfully traffic to the cell surface, as these predictions can be used to evaluate potential neoantigen vaccine targets. In this review we focus on the MHC Class I (MHC-I) antigen processing pathway. However, the MHC-II pathway has similar limitations and challenges for both ligand prediction and mechanistic understanding of peptide-selection.

Each class of MHC has a distinct processing pathway for peptide acquisition and presentation onto the cell surface. MHC-I derives its peptides from proteasomal degradation of recycled and aberrantly synthesised cellular proteins (1), which are transported into the endoplasmic reticulum by the ABC transporter TAP. TAP associates with other protein subunits of the peptide-loading complex including ERp57, tapasin, calreticulin, MHC-I heavy chain and β2m, for which a low-resolution cryo-EM structure of the whole complex has been published (2, 3). After complete assembly of the PLC, peptide ligands bind to the MHC-I α1-α2 domain peptide binding groove. Peptide binding triggers MHC-I dissociation from the PLC, as a heterotrimeric MHC-I complex that is subsequently released into the secretory pathway via COPII vesicles. Dissociation of low-affinity peptide before trafficking past the medial Golgi results in recycling of MHC-I back to the PLC through reglycosylation of MHC-I heavy chain by UGT1, allowing only optimally bound peptide-MHC complexes to survive to surface presentation (4-6). Further editing of the bound peptide can occur through interaction with TAPBPR, an ER-resident soluble tapasin homologue. Peptide editing to sample and select high-affinity peptide can occur at three distinct stages of the antigen processing pathway, tapasin mediated within the PLC, independently after PLC dissociation, and TAPBPR mediated (Figure 1). The adaptive response is restricted in responding to only that which is presented on the cell surface, highlighting the importance of understanding or controlling which peptide ligands survive the underlying antigen selection processes.

Figure 1 – MHC-I Antigen Processing Pathway

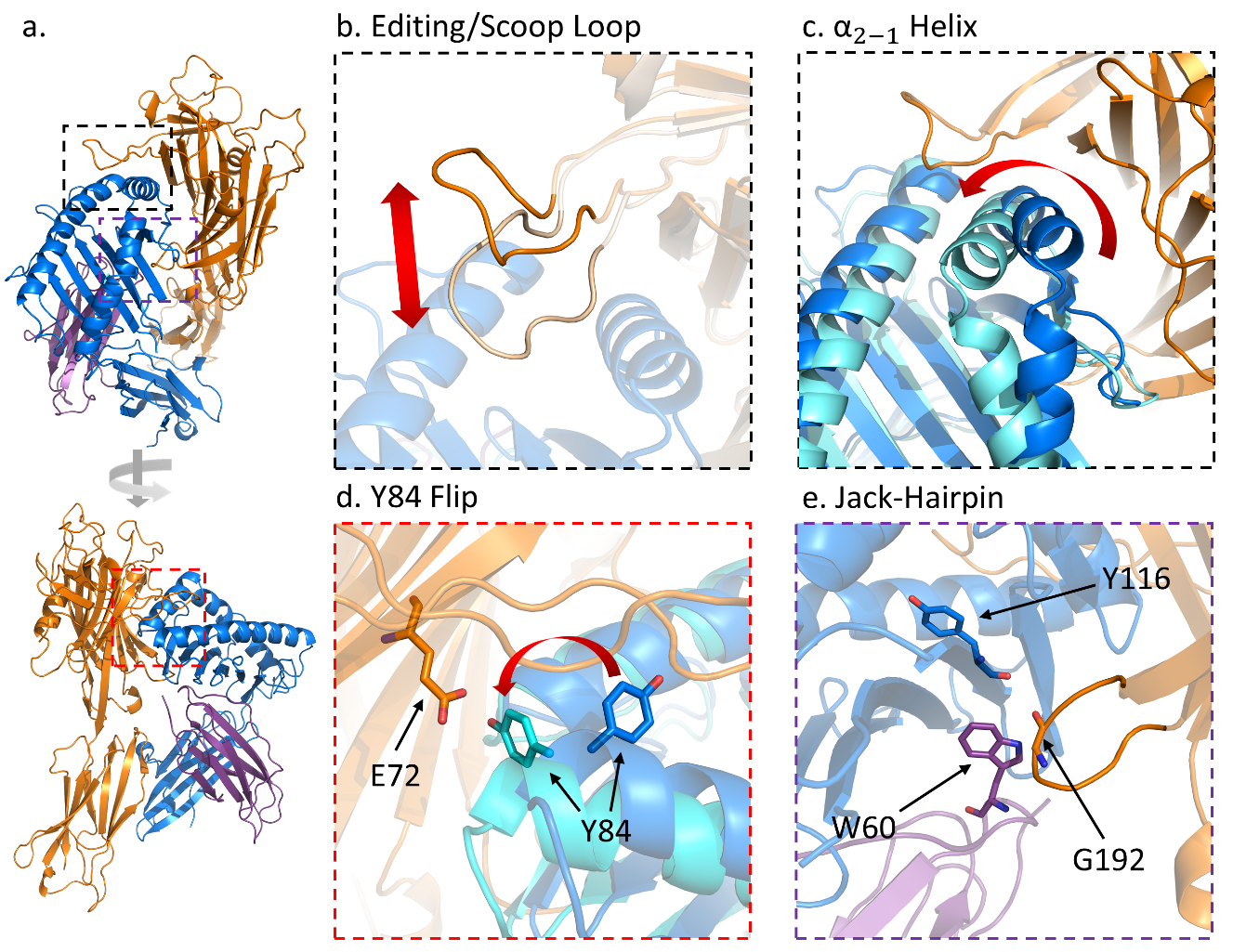


*Antigen Processing Pathway for MHC-I. Proteins are recycled into shorter peptides via the proteasome and further trimmed by cytosolic aminopeptidases. TAP1/2 transports peptides of optimal length and sequence preference into the ER, where they either directly bind MHC-I complexed with the PLC or dissociate for further trimming by ERAP1/2. MHC-I with bound peptide ligand dissociates from the PLC and can subsequently exchange independent of the PLC or have peptide exchange further catalysed by TAPBPR, depending on MHC-I allele and peptide ligand.*

## MHC-I peptide-selector function

The mechanism of peptide binding to MHC-I is not well understood. Crystal structures of MHC-I alleles across HLA-A, B and C allotypes show remarkable structural similarity, whilst significant variation in sequence preference and intrinsic peptide selector function i.e., the ability to independently select high affinity peptide in the absence of protein cofactors, can still exist (7). Evaluation of the peptide binding groove composition of MHC-I alleles, both in sequence and crystal structure, has been historically insufficient to completely predict the success of antigen presentation for all given ligands (8). As we only have crystal structures of the peptide-bound form of MHC-I, which exists after peptide selection and editing has already occurred, the relevant conformations and dynamics of the peptide selector function may only be observed in the elusive peptide-deficient state. The conceptual existence of a peptide-receptive intermediate state or ensemble of states responsible for the peptide-selector function was suggested through kinetic studies years ago (9, 10). The peptide selector function has been investigated through many simulation-based studies over the last 20+ years, as only through computational perturbation of existing peptide-bound structures can we create an atomistic model for a peptide-deficient conformation and potentially sample the peptide-receptive state. Our capacity to extensively sample the conformational ensemble of protein systems has improved dramatically through the creation of large-scale High-Performance computing systems and the development of enhanced-sampling methods, yet the consistency of reported simulations on the peptide-deficient state of MHC-I have failed to recover a definitive peptide-selection mechanism to date (11, 12). This peptide-selector function is generally considered under two conditions, an independent selector function by which certain MHC-I alleles are capable of binding high-affinity peptide-ligand without external cofactors, and a Tapasin/TAPBPR mediated peptide-selector function in which these protein cofactors provide some form of assistance in selecting high-affinity peptide (7). Several mechanisms have been proposed to describe the critical features of MHC-I peptide-selector function as a result of both MD simulation and biochemical experiments. These implicated mechanisms include the α2-1 helix (11, 13), 310 helix (14), Y84 flip (13, 15), jack-hairpin (16, 17) and Tapasin scoop-loop/editing loop (17-19), each depicted graphically in Figure 2 where possible. Yet none of these mechanisms provide a satisfactory solution to explain the variable intrinsic peptide selector function between alleles.

FIGURE 2 – Proposed Mechanisms of Peptide Editing and Tapasin Dependency



*(A) Cartoon representations of peptide-free MHC-I heavy chain (Blue) complexed with β2-microglobulin (purple) and tapasin (orange), visualised at different angles. (B-E) Magnification of mechanisms proposed to control peptide editing for HLA alleles either through biochemical assay or molecular dynamic simulation. Alternative conformations of MHC-I heavy chain and tapasin shown in teal and cream respectively. Amino acid side chains have been shown where relevant to the mechanism, with corresponding motions depicted with a red arrow.*

The prevailing systems model for peptide-selection presents the peptide-binding groove (PBG) opening rate as a peptide-dependent parameter, and PBG closing rate as an allele-dependent parameter (10), indicating certain alleles are able to efficiently capture peptide with a fast closing rate whilst other alleles require external help from protein cofactors such as Tapasin and TAPBPR. Despite these experimental quantitative measures of rate kinetics, molecular dynamics has only identified large-scale motions of protein dynamics (e.g. α2-1 helix, Figure 2c) that correlate with the kinetic measurements of tapasin dependency. These mechanisms lack an exact atomistic pathway that could be used to validate the effect as an allele-specific mechanism and not the result of insufficient sampling in simulations generating spurious correlations rather than causative mechanisms. This is not surprising given the almost insurmountable complexity of large-protein systems that make correlations easy to find but causative elements difficult to pin down, especially for convoluted allosteric pathways.

A crystal structure of the peptide-deficient form of MHC-I would be valuable to infer the mechanism of peptide binding. However, the instability of the peptide-deficient complex has made this impossible to date. Certain groups have derived experimental workarounds for this problem such as disulfide stabilization of flexible regions (20) or addition of shortened peptide ligands (21, 22) to derive crystal structures. These modifications appear to function by restraining the MHC-I F-pocket (C-terminal peptide binding site) alpha-helices which are highly implicated as a critical component of peptide-selector function. Whilst modifications have the potential to distort the relevant dynamic properties, peptide binding/selection is retained and T-Cell receptor activation shows no significant difference for disulfide-bound A\*02:01, which is reassuring (23). These disulfide crystal structures have identified two distinct conformations specific to the HLA-A\*02:01 allele localised around Tyr116 side chain reorientations, presumed to be representative of a peptide-locked and peptide-open state (21). Allele sets B\*44:02/B\*44:05 and B\*27\*05/B\*27:09 are the most common representative alleles for investigating intrinsic peptide-selector function as they differ significantly in ability to independently select high affinity peptide, and interestingly each pair only differ from one another by a single aa at position 116, further implicating the importance of this aa/region for peptide selection (7, 24).

## Tapasin & TAPBPR

MHC-I alleles are highly variable in their dependency on the PLC cofactor tapasin to present peptide onto the cell surface (7). Both Tapasin and its structural homologue TAPBPR (25) function as peptide-editors, modifying the repertoire of bound peptide ligands on MHC-I (26, 27). Both tapasin and TAPBPR modify the peptidome of MHC-I in an allele-specific manner, with the intrinsic ability of alleles to select high affinity peptide negatively correlating with their tapasin dependency, and variably to tapasin binding (28). Generally, these editors standardize the surface presentation across HLA-A and HLA-B alleles, with less effect on HLA-C (7, 29, 30). The recent influx of structural data has spurred a great number of interpretations for the mechanistic function of peptide-editing by these ER-resident chaperones. McShan et al. (31) have identified the existence of a minor conformational subpopulation of MHC-I alleles predicted to be the TAPBPR receptive conformation, however NMR provides no data on the atomistic structural conformation and was only able to implicate the α2 helix as correlated with the dynamic shift. It may be important to note that mouse alleles implicated the α3 domain as related to the conformational transition, whereas this was not observed in the two human alleles (31). Given the NMR-derived exchange rates to this novel conformation is on the order of milliseconds, it is possible that previous MD simulations have failed to capture the conformational change as conventional MD usually samples on low microsecond timescales. This indicates a need for enhanced sampling MD methods to improve our ability to sample rare events such as these. Whether this mechanism is consistent across multiple HLA alleles beyond the A\*02:01 and A\*01:01 representative set is yet to be seen.

Crystallography and mutational studies of TAPBPR have generated several conflicting hypotheses in recent years. Initial reports indicated the TAPBPR scoop-loop/editing-loop inserted into the MHC-I F-pocket, functioning as a competitive inhibitor for peptides (19, 32, 33). NMR studies subsequently showed no interaction between the F-pocket and TAPBPR editing-loop, instead indicating the editing-loop functioned as a kinetic trap for peptides, enhancing peptide affinity, in direct opposition to the prior hypothesis (18). Two very recent papers similarly showed peptide loading was controlled by N-terminal peptide-binding rather than C-terminal peptide binding in the A\*02:01 allele, using steered-MD and Markov state models (16, 34). This similarly suggests allosteric mechanisms for TAPBPR may be more relevant to peptide-selection as N-terminal binding would not be directly inhibited by F-pocket occupation by the editing-loop. As only one allele was simulated, this may yet be an allele specific-feature with different alleles utilising either a C-terminal or N-terminal loading pathway, potentially altered by TAPBPR or tapasin. It was further shown through correlation analyses of these simulations that aa 116 modulates allosteric communication between the F and A pockets, although an exact atomistic mechanism was not identified. This further implicated aa 116 as a critical mediator of tapasin/TAPBPR dependency, but as an allosteric coordinator rather than through direct peptide interactions as part of F-pocket dynamics. We also note that the TAPBPR jack-hairpinhas been proposed to function as a peptide-loading sensor, with the positioning of Trp60 correlating with the peptide-loading state (35). Whilst interesting, it is not yet clear how reorganisation of the region below the PBG mechanistically alters either peptide or TAPBPR binding, and limited research has been performed on the tapasin jack-hairpin effect. However, the localisation of the interaction directly below the heavily implicated aa 116 (Figure 2e) would indicate the potential to sense an allosteric shift along the peptide binding groove due to close localisation.

The recent study by Lan et al. (15) focused on the corresponding shorter tapasin editing-loop and showed allele-specific differences in functional effect following mutation of loop residues. The evolutionarily conserved hydrophobic Leu18 aa on the tapasin editing-loop appears critical for all observed alleles, but the auxiliary aa Lys16 and Ser14 have variable effects on different alleles, potentially indicating multiple distinct mechanisms of peptide-editing, but simultaneously highlighting the importance of the tapasin editing loop in mediating peptide-selector function. The simplest interpretation of these results suggests these sidechains directly insert into the binding pocket as acidity of the F-pocket correlated with effect of the polar/charged auxiliary sidechains (Figure 2b). However, similar to the TAPBPR NMR results (18), tapasin showed no clear variation in F-pocket dynamics upon mutation of the tapasin editing-loop. This reinforces the concept that tapasin-mediated peptide editing, whilst highly dependent on the tapasin editing-loop, occurs allosterically and that tapasin dependency is a function of the hydrogen bonding network generated by an “acidic” environment and not the direct binding affinity of the editing-loop for the F-pocket. A further suggestion of tapasin mechanism of action is the tapasin-mediated flip of Tyr84 (Figure 2d) which was observed in the TAPBPR-MHC crystal structure (36) and supported through NMR line broadening peak changes of tapasin. Whether this mechanism is a prerequisite for tapasin scoop-loop action, either through direct-entry or allosteric, or an alternative feature of the peptide selection mechanism is unknown. Capturing either an editing-loop bound conformation or identification of the allosteric mechanism through MD simulations should help deconvolute these two models. Unfortunately tapasin-bound MHC-I MD simulations have only been performed to identify potential mechanisms of peptide-selection and not analysed in the context of controlling tapasin-dependency (37).

## ERAPs

ERAP1 and ERAP2 are highly polymorphic aminopeptidases that trim peptides within the ER to optimal lengths for MHC-I binding, modifying the cellular surface immunopeptidome indirectly by altering the available pool of peptides. This function has an intrinsic duality in its ability to both destroy and create antigens, depending on the allotype and peptide-dependent extent of trimming (38). The mechanism of peptide trimming, commonly defined as the “molecular ruler” model, is well studied for soluble peptide ligand, leading to length-specific trimming of peptides by ERAP1. The mechanism suggests molecular interactions at both the active site and an undefined distal regulatory site induce a conformational transition to a more catalytically active conformation. X-ray crystal structures have identified two main states defined as “open” and “closed”, with SAXS data suggesting ligand binding induces transition to a more closed-like conformation (39). As crystal structures are not necessarily representative of the ensemble of states, MD has been used to generate a more representative set of structures to better relate to low-resolution SAXS, data which further validated the existing two-state model (39, 40). A key development is the identification of multiple distal sites that can bind the c-terminus of peptides or associated ligands leading to increased catalytic activity (40).

ERAP1 has been further shown to cleave MHC-bound peptide (41), though no mechanism consistent with structural data has yet been proposed. MD investigations into this ERAP function are hindered by the lack of an ERAP-peptide-MHC heterotrimeric complex. However, MD-based MHC-I peptide dissociation studies have shown N-terminal peptide dissociation mechanisms (42) which, combined with experimental observations of covalently bonded peptide showing trimming susceptibility (43), could suggest a potential mechanism not requiring a closely complexed ERAP-MHC structure. Recent evidence suggests bound-peptide trimming occurs at a rate slower than peptide dissociation Thus whilst the ability of ERAP trim bound peptide is supported, the biological relevance remains unclear (42, 44).

## Predicting MHC-Peptide surface presentation

Neural networks have been the predominant solution to predicting MHC epitope presentation for over a decade, leveraging MHC sequence information and trained against experimental measures of either *in vitro* peptide-MHC binding affinity or *in vivo* eluted ligand mass spectrometry, a method that detects peptides on the cell surface. NetMHCpan remains the state-of-the-art methodology for antigen presentation prediction (45), and is most commonly used in recent vaccine target prediction studies (46, 47). Fewer than 30% of predicted neoepitopes elicited CD8+ responses in these referenced studies, suggesting real-world application of these predictive algorithms still has significant room for improvement. A number of papers have evaluated prediction performance of different neural networks for both MHC-I (8, 48, 49) and MHC-II (50, 51), and we also highlight the weekly updated performance evaluation of common prediction methods on the IEDB dataset (52). Bhattacharya et al. (53) specifically identified poor performance covariates for predictive methods, indicating number of training examples, insufficient number of binding peptides relative to non-binding peptides and poor peptide sequence diversity as principal causes for model inaccuracy. Subsequent improvement of training datasets identified complexity of the peptide-motifs as a significant source of error that additional data did not resolve, suggesting some alleles have harder to learn allelic preferences than others (54). Incorporation of empirical features such as peptide cleavage, gene expression and gene presentation bias have resulted in positive prediction improvements which may help to resolve these sources of error (54, 55). We are not yet aware of any machine learning prediction algorithms that utilise ERAP processing, tapasin effect or calreticulin/UGT1 recycling as direct training descriptors.

An alternative approach to machine learning predictions are mechanistic models that explicitly model biochemical processes in the antigen processing pathway. Boulanger et al. have presented a mechanistic model that relates intracellular peptide supply to surface presentation abundance, utilising experimental measures of intracellular peptide supply, peptide off-rate as training features to predict peptide surface presentation (56, 57). This model incorporates explicitly defined intermediate states in the peptide-selection process and fits rate parameters to experimental data. These parameters representing rates of distinct processes in the antigen processing pathway are predicted using a Markov chain Monte-Carlo sampling of the multidimensional parameter space and scoring the resulting parameter set using a maximum likelihood estimation. Features of tapasin-dependency have already been incorporated into such models using experimental measures (57), but other influential processes such as ERAP cleavage, calreticulin/UGT1 recycling and TAPBPR selector function can be readily incorporated. The dependency on experimental data to fit these models limits their application to comparisons of small numbers of peptides/allele pairs. Development of generalisable predictive models for some stages of the explicitly modelled pathway greatly reduce the requirements for experimental classification of peptide/allele pairs for these approaches but will require a complete empirical understanding of the mechanistic processes that control them. A combination of machine learning and empirical models built into the mechanistic modelling framework appears an attractive future approach to predict antigen presentation but needs to be reinforced through an improved empirical understanding of the mechanistic processes that govern antigen processing.

## T-Cell Interactions and Binding

We highlight the ongoing challenge in prediction of T-cell activation against neoantigens for vaccine immunotherapies. Prediction of successful antigen binding and trafficking to the cell surface is only one side of the coin, as once an antigen has made it to the cell surface, it still needs to be able to activate a corresponding T-cell receptor (TCR). The machine learning prediction models described previously rely on an absolute correlation between binding affinity/surface presentation and T-cell priming. Whilst binding and surface presentation is a necessary prerequisite for T-cell activation, the response is not guaranteed for a given peptide (58). A large number of direct immunogenicity prediction software have been created to fill this gap using physicochemical features of the MHC-I-TCR interaction (59-61). Despite the improvement of such predictive models, a complete understanding of what features trigger a T-cell response remains elusive. A more general result derived from the Tumour Neoantigen Selection Alliance (TESLA) reviews 25 teams predicting neoantigens on experimentally validated epitopes, in which the best scoring team identified 16 out of 37 validated neoantigens in the top 20 from a pool of 608 potential neoantigens predicted by the consortium (62). Despite this positive result, 18 neoantigens were not identified at all or not predicted as immunogenic, indicating we are still missing a large quantity of potentially therapeutic vaccine targets. As the samples used in this study only encompass 13 common MHC-I alleles, the generalisability of these methods to less common alleles remains of interest. The difficulty of this approach lies in the relatively small quantity of potential training data as well as a lack of direct understanding of what physiochemical features of the MHC-I complex determine if it will be immunogenic or not.

MD simulations have also been used to predict potential neoantigens since the early 1990s (63), with many neoantigen identification workflows performing MD as either a final validation or filtration step to select optimal vaccination targets (64-66). This is often used in combination with peptide docking software such as HADDOCK (67), AutoDock Vina (68) and ADCP (69) many of which are evaluated in this review (70). The computational expense of these approaches makes them inefficient for screening large numbers of peptide ligands, hence the common use of neural networks to refine the potential pool of ligands prior to simulation-based validation. Computational evaluation of epitope binding affinity is conventionally done either using docking scores, qualitatively through molecular descriptors (e.g. RMSD) or binding free energy calculations using MM/GBSA/PBSA (71). Peptides are a large ligand to work with using these computational approaches. The high dimensionality of the potential configurational space of the peptide means that MD simulations and unrestrained docking approaches are unlikely to sufficiently sample the relevant conformations for binding. Each docking method generally utilises some heuristic restraint that limits the conformational space of the peptide, which will of course introduce some potential for error if this restraint indirectly prevents sampling of the correct binding mode (72). Many approaches utilise the anchor pockets of the MHC-I peptide-binding groove or the backbone conformation of peptides in crystal structures to generate heuristic restraints for peptides. These approaches are typically good for alleles with large numbers of representative crystal structures and for scoring peptides that bind in the typical binding pose (73). It is important to note that certain peptides have non-conventional binding poses which will result in poorer scoring accuracy (74).

## Future Directions

There has been a large influx of improvements to machine-learning methods to improve our ability to predict neoantigen vaccine targets However it is evident that we are lacking significant understanding of critical features of the antigen processing pathway that directly control peptide supply and selection. Many current approaches rely on the hope that predictive models can intuit potentially complex features of protein dynamics from sequence and structural data without time resolution. The extent of success is quite remarkable given these constraints, yet if we hope to identify more comprehensive pools of vaccine targets, we believe an improved empirical understanding of the mechanistic function of antigen processing will be invaluable for the improvement of predictive models. Whilst prediction of antigen presentation and immunogenicity is the current focus of the field, it is important to note that discovering potential antigens is only one side of the therapeutic potential of antigen processing. Many diseases exhibit correlations with certain ERAP and MHC alleles, suggesting that in some cases we may not need to trigger an immune reaction but instead suppress or modulate an active response, specifically in the case of numerous autoimmune conditions (75). We therefore hope to observe a new field develop in curation of the antigen surface peptidome, empowered by our continually developing understanding of antigen processing.

## Perspectives

* Prediction of antigen processing and T-cell activation is a critical step in the development of vaccines against numerous pathogens and cancer. The identification of novel targets for immunotherapeutic intervention in a patient-specific approach could drastically improve treatment outcomes.
* Prediction of antigen processing and T-cell activation is currently possible but limited by insufficient understanding mechanistic differences between MHC-I and ERAP alleles resulting in incomplete identification of potential vaccine targets.
* We need to improve our empirical understanding of each stage of the antigen processing pathway to allow both mechanistic and machine learning models to better capture the complex biochemical features of peptide supply, selection, editing and T-cell activation.

## Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

## Author Contribution

S.T. prepared the text and figures. T.E. and J.W.E. commented on the draft.

## References

1. Schubert U, Anton LC, Gibbs J, Norbury CC, Yewdell JW, Bennink JR. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. Nature. 2000;404(6779):770-4.

2. 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.

3. Fisette O, Schröder GF, Schäfer LV. Atomistic structure and dynamics of the human MHC-I peptide-loading complex. Proceedings of the National Academy of Sciences. 2020;117(34):20597-606.

4. Zhang W, Wearsch PA, Zhu Y, Leonhardt RM, Cresswell P. A role for UDP-glucose glycoprotein glucosyltransferase in expression and quality control of MHC class I molecules. Proc Natl Acad Sci U S A. 2011;108(12):4956-61.

5. Neerincx A, Hermann C, Antrobus R, van Hateren A, Cao H, Trautwein N, et al. TAPBPR bridges UDP-glucose:glycoprotein glucosyltransferase 1 onto MHC class I to provide quality control in the antigen presentation pathway. eLife. 2017;6:e23049.

6. Neerincx A, Hermann C, Antrobus R, van Hateren A, Cao H, Trautwein N, et al. TAPBPR bridges UDP-glucose:glycoprotein glucosyltransferase 1 onto MHC class I to provide quality control in the antigen presentation pathway. Elife. 2017;6.

7. Bashirova AA, Viard M, Naranbhai V, Grifoni A, Garcia-Beltran W, Akdag M, et al. HLA tapasin independence: broader peptide repertoire and HIV control. Proceedings of the National Academy of Sciences. 2020;117(45):28232-8.

8. Zhang L, Udaka K, Mamitsuka H, Zhu S. Toward more accurate pan-specific MHC-peptide binding prediction: a review of current methods and tools. Briefings in Bioinformatics. 2012;13(3):350-64.

9. Springer S, Doring K, Skipper JC, Townsend AR, Cerundolo V. Fast association rates suggest a conformational change in the MHC class I molecule H-2Db upon peptide binding. Biochemistry. 1998;37(9):3001-12.

10. Elliott T. How do peptides associate with MHC class I molecules? Immunology Today. 1991;12(11):386-8.

11. Bailey A, Dalchau N, Carter R, Emmott S, Phillips A, Werner JM, et al. Selector function of MHC I molecules is determined by protein plasticity. Scientific Reports. 2015;5(1):14928.

12. Ostermeir K, Springer S, Zacharias M. Coupling between side chain interactions and binding pocket flexibility in HLA-B\*44:02 molecules investigated by molecular dynamics simulations. Molecular Immunology. 2015;63(2):312-9.

13. McShan AC, Natarajan K, Kumirov VK, Flores-Solis D, Jiang J, Badstubner M, et al. Peptide exchange on MHC-I by TAPBPR is driven by a negative allostery release cycle. Nat Chem Biol. 2018;14(8):811-20.

14. Mage MG, Dolan MA, Wang R, Boyd LF, Revilleza MJ, Robinson H, et al. A structural and molecular dynamics approach to understanding the peptide-receptive transition state of MHC-I molecules. Mol Immunol. 2013;55(2):123-5.

15. Lan H, Abualrous ET, Sticht J, Fernandez LMA, Werk T, Weise C, et al. Exchange catalysis by tapasin exploits conserved and allele-specific features of MHC-I molecules. Nature Communications. 2021;12(1):4236.

16. Xu H, Song K, Da LT. Dynamics of peptide loading into major histocompatibility complex class I molecules chaperoned by TAPBPR. Phys Chem Chem Phys. 2022;24(20):12397-409.

17. Beerbaum M, Ballaschk M, Erdmann N, Schnick C, Diehl A, Uchanska-Ziegler B, et al. NMR spectroscopy reveals unexpected structural variation at the protein-protein interface in MHC class I molecules. J Biomol NMR. 2013;57(2):167-78.

18. McShan AC, Devlin CA, Morozov GI, Overall SA, Moschidi D, Akella N, et al. TAPBPR promotes antigen loading on MHC-I molecules using a peptide trap. Nat Commun. 2021;12(1):3174.

19. Ilca FT, Neerincx A, Hermann C, Marcu A, Stevanović S, Deane JE, et al. TAPBPR mediates peptide dissociation from MHC class I using a leucine lever. eLife. 2018;7:e40126.

20. Wieczorek M, Abualrous ET, Sticht J, Alvaro-Benito M, Stolzenberg S, Noe F, et al. Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. Front Immunol. 2017;8:292.

21. Anjanappa R, Garcia-Alai M, Kopicki JD, Lockhauserbaumer J, Aboelmagd M, Hinrichs J, et al. Structures of peptide-free and partially loaded MHC class I molecules reveal mechanisms of peptide selection. Nat Commun. 2020;11(1):1314.

22. Glithero A, Tormo J, Doering K, Kojima M, Jones EY, Elliott T. The crystal structure of H-2D(b) complexed with a partial peptide epitope suggests a major histocompatibility complex class I assembly intermediate. J Biol Chem. 2006;281(18):12699-704.

23. Saini SK, Tamhane T, Anjanappa R, Saikia A, Ramskov S, Donia M, et al. Empty peptide-receptive MHC class I molecules for efficient detection of antigen-specific T cells. Science Immunology. 2019;4(37):eaau9039.

24. Abualrous ET, Fritzsche S, Hein Z, Al-Balushi MS, Reinink P, Boyle LH, et al. F pocket flexibility influences the tapasin dependence of two differentially disease-associated MHC Class I proteins. European Journal of Immunology. 2015;45(4):1248-57.

25. Teng MS, Stephens R, Pasquier LD, Freeman T, Lindquist JA, Trowsdale J. A human TAPBP (TAPASIN)-related gene,TAPBP-R. European Journal of Immunology. 2002;32(4):1059-68.

26. Wearsch PA, Cresswell P. Selective loading of high-affinity peptides onto major histocompatibility complex class I molecules by the tapasin-ERp57 heterodimer. Nat Immunol. 2007;8(8):873-81.

27. Morozov GI, Zhao H, Mage MG, Boyd LF, Jiang J, Dolan MA, et al. Interaction of TAPBPR, a tapasin homolog, with MHC-I molecules promotes peptide editing. Proc Natl Acad Sci U S A. 2016;113(8):E1006-15.

28. van Hateren A, Carter R, Bailey A, Kontouli N, Williams AP, Kaufman J, et al. A mechanistic basis for the co-evolution of chicken tapasin and major histocompatibility complex class I (MHC I) proteins. J Biol Chem. 2013;288(45):32797-808.

29. Ilca FT, Drexhage LZ, Brewin G, Peacock S, Boyle LH. Distinct Polymorphisms in HLA Class I Molecules Govern Their Susceptibility to Peptide Editing by TAPBPR. Cell Rep. 2019;29(6):1621-32 e3.

30. Williams AP, Peh CA, Purcell AW, Mccluskey J, Elliott T. Optimization of the MHC Class I Peptide Cargo Is Dependent on Tapasin. Immunity. 2002;16(4):509-20.

31. Mcshan AC, Devlin CA, Overall SA, Park J, Toor JS, Moschidi D, et al. Molecular determinants of chaperone interactions on MHC-I for folding and antigen repertoire selection. Proceedings of the National Academy of Sciences. 2019;116(51):25602-13.

32. Hafstrand I, Sayitoglu EC, Apavaloaei A, Josey BJ, Sun R, Han X, et al. Successive crystal structure snapshots suggest the basis for MHC class I peptide loading and editing by tapasin. Proc Natl Acad Sci U S A. 2019;116(11):5055-60.

33. Sagert L, Hennig F, Thomas C, Tampé R. A loop structure allows TAPBPR to exert its dual function as MHC I chaperone and peptide editor. eLife. 2020;9:e55326.

34. Lin M, Xu H, Shi Y, Da L-T. Antigenic Peptide Loading into Major Histocompatibility Complex Class I Is Driven by the Substrate N-Terminus. CCS Chemistry. 2022;4(3):910-25.

35. Hee CS, Beerbaum M, Loll B, Ballaschk M, Schmieder P, Uchanska-Ziegler B, et al. Dynamics of free versus complexed beta2-microglobulin and the evolution of interfaces in MHC class I molecules. Immunogenetics. 2013;65(3):157-72.

36. Jiang J, Natarajan K, Boyd LF, Morozov GI, Mage MG, Margulies DH. Crystal structure of a TAPBPR–MHC I complex reveals the mechanism of peptide editing in antigen presentation. Science. 2017;358(6366):1064-8.

37. Fisette O, Wingbermühle S, Tampé R, Schäfer LV. Molecular mechanism of peptide editing in the tapasin–MHC I complex. Scientific Reports. 2016;6(1):19085.

38. Koumantou D, Barnea E, Martin-Esteban A, Maben Z, Papakyriakou A, Mpakali A, et al. Editing the immunopeptidome of melanoma cells using a potent inhibitor of endoplasmic reticulum aminopeptidase 1 (ERAP1). Cancer Immunol Immunother. 2019;68(8):1245-61.

39. Maben Z, Arya R, Georgiadis D, Stratikos E, Stern LJ. Conformational dynamics linked to domain closure and substrate binding explain the ERAP1 allosteric regulation mechanism. Nat Commun. 2021;12(1):5302.

40. Maben Z, Arya R, Rane D, An WF, Metkar S, Hickey M, et al. Discovery of Selective Inhibitors of Endoplasmic Reticulum Aminopeptidase 1. J Med Chem. 2020;63(1):103-21.

41. Chen H, Li L, Weimershaus M, Evnouchidou I, Van Endert P, Bouvier M. ERAP1-ERAP2 dimers trim MHC I-bound precursor peptides; implications for understanding peptide editing. Scientific Reports. 2016;6(1):28902.

42. Mavridis G, Arya R, Domnick A, Zoidakis J, Makridakis M, Vlahou A, et al. A systematic re-examination of processing of MHCI-bound antigenic peptide precursors by endoplasmic reticulum aminopeptidase 1. J Biol Chem. 2020;295(21):7193-210.

43. Li L, Batliwala M, Bouvier M. ERAP1 enzyme-mediated trimming and structural analyses of MHC I-bound precursor peptides yield novel insights into antigen processing and presentation. J Biol Chem. 2019;294(49):18534-44.

44. Mavridis G, Mpakali A, Zoidakis J, Makridakis M, Vlahou A, Kaloumenou E, et al. The ERAP1 active site cannot productively access the N-terminus of antigenic peptide precursors stably bound onto MHC class I. Sci Rep. 2021;11(1):16475.

45. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Research. 2020;48(W1):W449-W54.

46. Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, et al. An immunogenic personal neoantigen vaccine for patients with melanoma. Nature. 2017;547(7662):217-21.

47. Sahin U, Derhovanessian E, Miller M, Kloke BP, Simon P, Lower M, et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. Nature. 2017;547(7662):222-6.

48. Chen Z, Min MR, Ning X. Ranking-Based Convolutional Neural Network Models for Peptide-MHC Class I Binding Prediction. Front Mol Biosci. 2021;8:634836.

49. Nibeyro G, Girotti R, Prato L, Moron G, Luján HD, Fernandez EA. MHC-I binding affinity derived metrics fail to predict tumor specific neoantigen immunogenicity. bioRxiv. 2022:484285.

50. Moore TV, Nishimura MI. Improved MHC II epitope prediction - a step towards personalized medicine. Nat Rev Clin Oncol. 2020;17(2):71-2.

51. Nielsen M, Lund O, Buus S, Lundegaard C. MHC class II epitope predictive algorithms. Immunology. 2010;130(3):319-28.

52. Trolle T, Metushi IG, Greenbaum JA, Kim Y, Sidney J, Lund O, et al. Automated benchmarking of peptide-MHC class I binding predictions. Bioinformatics. 2015;31(13):2174-81.

53. Bhattacharya R, Sivakumar A, Tokheim C, Guthrie VB, Anagnostou V, Velculescu VE, et al. Evaluation of machine learning methods to predict peptide binding to MHC Class I proteins. bioRxiv. 2017:154757.

54. Sarkizova S, Klaeger S, Le PM, Li LW, Oliveira G, Keshishian H, et al. A large peptidome dataset improves HLA class I epitope prediction across most of the human population. Nature Biotechnology. 2020;38(2):199-209.

55. Koşaloğlu-Yalçın Z, Lee J, Greenbaum J, Schoenberger SP, Miller A, Kim YJ, et al. Combined assessment of MHC binding and antigen abundance improves T cell epitope predictions. iScience. 2022;25(2):103850.

56. Boulanger DSM, Eccleston RC, Phillips A, Coveney PV, Elliott T, Dalchau N. A Mechanistic Model for Predicting Cell Surface Presentation of Competing Peptides by MHC Class I Molecules. Front Immunol. 2018;9:1538.

57. Dalchau N, Phillips A, Goldstein LD, Howarth M, Cardelli L, Emmott S, et al. A peptide filtering relation quantifies MHC class I peptide optimization. PLoS Comput Biol. 2011;7(10):e1002144.

58. Harndahl M, Rasmussen M, Roder G, Dalgaard Pedersen I, Sorensen M, Nielsen M, et al. Peptide-MHC class I stability is a better predictor than peptide affinity of CTL immunogenicity. Eur J Immunol. 2012;42(6):1405-16.

59. Calis JJ, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, et al. Properties of MHC class I presented peptides that enhance immunogenicity. PLoS Comput Biol. 2013;9(10):e1003266.

60. Schmidt J, Smith AR, Magnin M, Racle J, Devlin JR, Bobisse S, et al. Prediction of neo-epitope immunogenicity reveals TCR recognition determinants and provides insight into immunoediting. Cell Rep Med. 2021;2(2):100194.

61. Li G, Iyer B, Prasath VBS, Ni Y, Salomonis N. DeepImmuno: deep learning-empowered prediction and generation of immunogenic peptides for T-cell immunity. Briefings in Bioinformatics. 2021;22(6):bbab160.

62. Wells DK, van Buuren MM, Dang KK, Hubbard-Lucey VM, Sheehan KCF, Campbell KM, et al. Key Parameters of Tumor Epitope Immunogenicity Revealed Through a Consortium Approach Improve Neoantigen Prediction. Cell. 2020;183(3):818-34 e13.

63. Rognan D, Scapozza L, Folkers G, Daser A. Molecular Dynamics Simulation of MHC-Peptide Complexes as a Tool for Predicting Potential T Cell Epitopes. Biochemistry. 1994;33(38):11476-85.

64. da Costa AS, Fernandes TVA, Bello ML, de Souza TLF. Evaluation of potential MHC-I allele-specific epitopes in Zika virus proteins and the effects of mutations on peptide-MHC-I interaction studied using in silico approaches. Comput Biol Chem. 2021;92:107459.

65. Ghobadi Z, Mahnam K, Shakhsi-Niaei M. In-silico design of peptides for inhibition of HLA-A\*03-KLIETYFSK complex as a new drug design for treatment of multiples sclerosis disease. Journal of Molecular Graphics and Modelling. 2022;111:108079.

66. Olotu FA, Soliman MES. Immunoinformatics prediction of potential B-cell and T-cell epitopes as effective vaccine candidates for eliciting immunogenic responses against Epstein-Barr virus. Biomed J. 2021;44(3):317-37.

67. Spiliotopoulos D, Kastritis PL, Melquiond ASJ, Bonvin AMJJ, Musco G, Rocchia W, et al. dMM-PBSA: A New HADDOCK Scoring Function for Protein-Peptide Docking. Frontiers in Molecular Biosciences. 2016;3.

68. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455-61.

69. Zhang Y, Sanner MF. AutoDock CrankPep: combining folding and docking to predict protein-peptide complexes. Bioinformatics. 2019;35(24):5121-7.

70. Weng G, Gao J, Wang Z, Wang E, Hu X, Yao X, et al. Comprehensive Evaluation of Fourteen Docking Programs on Protein–Peptide Complexes. Journal of Chemical Theory and Computation. 2020;16(6):3959-69.

71. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res. 2000;33(12):889-97.

72. Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang JZH, et al. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. Chemical Reviews. 2019;119(16):9478-508.

73. Antunes DA, Devaurs D, Moll M, Lizee G, Kavraki LE. General Prediction of Peptide-MHC Binding Modes Using Incremental Docking: A Proof of Concept. Sci Rep. 2018;8(1):4327.

74. Marzella DF, Parizi FM, van Tilborg D, Renaud N, Sybrandi D, Buzatu R, et al. PANDORA: A Fast, Anchor-Restrained Modelling Protocol for Peptide: MHC Complexes. Front Immunol. 2022;13:878762.

75. Matzaraki V, Kumar V, Wijmenga C, Zhernakova A. The MHC locus and genetic susceptibility to autoimmune and infectious diseases. Genome Biol. 2017;18(1):76.