Expanding the Cancer Neoantigen Peptide Repertoire beyond In silico Tools

Amit Jain¹, Jackwee Lim²*

¹National Cancer Centre Singapore, 11 Hospital Crescent, Singapore 169610
²Singapore Immunology Network, Agency for Science, Technology and Research, 8A Biomedical Grove, Immunos, Singapore 138648

*Correspondence should be addressed to Jackwee Lim; lim_jack_wee@immunol.a-star.edu.sg

Received date: January 08, 2021, Accepted date: February 23, 2021

Copyright: © 2021 Jain A, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Commentary

CD8⁺ cytotoxic T cells recognise and kill cancer cells that present immunogenic peptides bound to the cell surface major histocompatibility complex class I (MHC-I) molecules. The immunogenicity of these peptides derives from them being recognised as non-self after their parent proteins are intracellularly processed and presented as peptide-major histocompatibility complex class I (pMHC-I) complexes. pMHC specific T cell receptor (TCR) recognition then leads to cytotoxic T cell response. Generally, robust pMHC-I binding is needed for a chance encounter with a pMHC specific TCR bearing CD8⁺ T cell derived from in vivo T cell evolution and thymic selection. As such, pMHC binding and stability are the key starting points towards understanding the T cell response. With current technologies, peptide and MHC interaction may be deduced from either direct pMHC binding/ stability assays or in silico prediction tools, and each method has its advantages and disadvantages (Table 1). Nonetheless, the ideal method should not exclude peptides that are potentially immunogenic. Hence there is continued need for more accurate high-throughput methods that assess the natural physico-biochemical interaction between peptide and MHC molecules.

Lim has recently developed the EZ MHC-I assay using the pMHC-I single chain trimer (SCT) molecule to enable a direct interrogation of the MHC ligandome predicted in silico or derived from patients' samples. It is an assay based on empty MHC-I protein fragmentation to rapidly characterize bound peptides for affinity and stability [1]. This will exclude predicted binders which do not stabilize the MHC-I molecule, identify missed hits, and potentially enable neoantigen discovery with better characterized peptides. Here, we describe the challenges of neoantigen selection, share missed hits identified by the EZ MHC-I assay, present the EZ MHC-I assay in greater details, and propose future SCT-based applications towards CD8⁺ T cell specific neoantigen discovery.

The hunt for immunogenic peptides

Immunogenic T cell antigens have been well characterized in the context of infectious diseases where the entire antigen from viral proteins is “non-self” and hence immunogenic. On the other hand, the identification of “immunogenic” cancer antigens or neoantigens is more challenging [2,3]. While in methylcholanthrene-induced mouse models, several immunogenic peptide sequences have been identified [4], human cancer mutations are largely patient specific and therefore bespoke, notwithstanding intratumoral heterogeneity. Broadly, cancer antigens comprise of cancer specific overexpressed proteins, viral proteins in virus associated cancers, and specific peptides derived from nonsynonymous mutations, deletions, or translocations. Several bed to bench clinical translational studies of cancer immunotherapy have unravelled that stable pMHC-T cell interactions are requisite for T cell cytotoxicity against cancer [5,6]. However, identifying stable pMHC molecules is experimentally laborious and in silico predictions remain under scrutiny [7-9]. Nonetheless, tremendous efforts to characterize neoantigen pools with bonafide T cell responses in patients responding to immune checkpoint inhibitors have led to clinical development of therapeutic cancer peptide vaccines by companies such as Gritstone Oncology and Genocea [10,11]. Presently, potential neoantigens are identified using paired normal
genomic sequencing of both DNA and RNA. This creates a large pool of mutation specific peptides which are then mapped to the patient specific human leukocyte antigen (HLA). Mass spectrometry and a large literature of experimental data have enabled the development of more accurate in silico prediction algorithms to identify peptides that are likely to bind stably in a MHC allelic groove [12]. Furthermore, the discovery of immune checkpoint receptors, CTLA-4 and PD-1 [13] that abrogate T cell responses against cancer has led to a reinvigoration of the study of pMHC binding to identify therapeutic targets [14], especially in the context of patients treated with immune checkpoint inhibitors to relieve CD8+ T cell exhaustion [15,16]. However, not all cancer mutations are similarly immunogenic, and immunogenicity is possibly driven by the duration of MHC peptide binding and pMHC-TCR bond conformation to trigger a T cell response [17,18]. Hence careful characterization of stable pMHC-I molecules can lead to the identification of neoantigens more likely to trigger tumor-specific T cell mediated antitumor immune response, and potentially drive the success of neoantigen derived cancer vaccine therapy [19-22].

**An in silico challenge: predicting unconventional peptide binding in MHC-I groove**

Early neoantigen discovery based on early proofs of peptide antigens presented on cell surfaces propelled the development of pMHC-I/TCR related applications [23,24]. Currently, prediction of neoantigens by bioinformatics remains a popular approach prior to experimental validation. Indeed, prediction algorithms such as NetMHCpan, MHCSeqNet, MHCflurry and NetMHCstab have become highly reliable in predicting MHC-peptide binding affinity or stability [11,12,25-27]. This initially propelled the clinical development of neoantigen vaccines by both academia [28-30] and companies such as Neon, BioNTech, and subsequently also in combination with immune checkpoint inhibitors such as in the clinical trials run by Moderna and Merck [31,32]. However, the suboptimal nature of in silico prediction has also driven companies such as Gritstone Oncology and Genoece to develop patient data-trained deep learning tools or in vitro cellular methods, respectively to improve the early identification of potential neoantigens [10,11]. Hence further characterization of these neoantigens for their stability will favor more optimized therapeutics. Indeed recent studies on experimentally validated neoantigen have shown that both binding affinity and complex stability are key parameters of a pMHC molecule to stimulate patient tumor infiltrating lymphocytes and mount an immunogenic response [18,33]. However, urea denaturing methods to measure actual pMHC-I stability are tedious and predicting the energetic stability of a pMHC complex can be computationally inaccurate, a grand challenge to consider the entropic and enthalpic factors even for the protein folding community [34,35]. Indeed, a large majority of predicted neoantigens do not elicit T-cell responses as only a small fraction is capable of presentation as cell surface pMHC and subsequent recognition by the rare TCR. Therefore, actual measurement of pMHC-I affinity and stability can potentially improve the reliability of predicted peptides [7,36]. Indeed, an additional 60% of predicted epidermal growth factor receptor (EGFR) mutated peptide candidates was found using the EZ MHC-I assay, which remain to be validated in future patient studies (Figure 1).

<table>
<thead>
<tr>
<th>Resources</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| *In silico* algorithms e.g. NetMHCpan, NetMHCstab, MHCSeqNet and MHCflurry. | • Fairly accurate and fast  
• Free public access  
• High-throughput | • Lack of biophysical characterization  
• Restricted binding motifs  
• Numerous candidates  
• Missed hits |
| Peptide-based mass spectrometry | • Natural processing of protein antigen and presentation of peptide antigens  
• Limited multiplexing | • Peptide recovery challenge  
• Complex data analysis  
• Mixed results immunogenicity in humans |
| *In vitro* assays e.g. ELISA, ProlImmune REVEAL and Immunitrack NeoScreen. | • Early biophysical characterization  
• Unrestricted peptide types  
• High-throughput | • Multiple-step irreproducibility  
• Sample preparation challenge  
• Hypothetical candidates |
| EZ MHC-I assay | • Early biophysical characterization  
• Unrestricted peptide types  
• Direct band visualization of MHC-I protein  
• Shortest hands-on time among current *in vitro* assays  
• High-throughput | • Limited to MHC-I alleles  
• Production of suitable MHC-I molecule  
• Hypothetical candidates |

**Table 1: Advantages and disadvantages of different T cell epitope prediction resources.**

An in silico challenge: predicting unconventional peptide binding in MHC-I groove
Favoring biophysical characterization with easier pMHC-I production

To date, the human MHC gene, also known as HLA has been strongly associated with hundreds of disease and thus play a pivotal role in disease susceptibility genetic testing [37]. Several structural studies have also advanced our understanding of different immunogenic peptides bound within the MHC-I/II protein groove, including the development of automated modeling albeit selected MHC-I allotypes [38]. X-ray crystallography has further revealed peptides of 10/11 amino acids, which can bind the MHC-I groove in either a zig-zag or bulging manner, while anchoring its N and C termini into the A and F pockets, respectively [39]. Moreover, several high-resolution structures have also shown non-canonical peptide lengths of up to 16-mer exiting out at the F pocket [40-43]. These suggest that longer peptide lengths can adopt unconventional binding modes within the enclosed groove of MHC-I protein, previously thought to be unique to the more opened groove of MHC-II protein. Also, longer peptide precursors of MHC-I do undergo trimming by endoplasmic reticulum aminopeptidases 1 and 2 along the transporter associated with antigen processing machinery to be shorten to a more preferable 9- to 13-mer and adopt multiple bulging conformations in the MHC-I binding groove [44-46]. Taken together, the binding of peptide in the MHC-I groove is not always flat and thus unpredictable in silico.

To enable real biophysical characterization, different pMHC-I molecules have to be produced using recombinant methods [1,47-51]. Stable pMHC-I molecules are traditionally used in cytometry for probing CD8+ T cells but unsuitable reagents for peptide-exchange. Thus Bakker et al. used a photolabile peptide cleavable by UV irradiation to make empty MHC-I molecules to enable peptide exchange [51]. Similarly, advances in peptide-exchange technologies include peptide deficient MHC-I/TAPasin binding protein related complexes and thermal exchangeable pMHC-I molecules to overcome possible photodamage using UV cleavable pMHC-I molecules [49,50]. More importantly, to...
encourage biophysical characterization, secreted peptide exchangeable SCT proteins were successfully made using the mellitin-based baculovirus expression vector system (mBEVS, mellitin leader: MKFLKVLALVMVVYISIYIA) [1]. A major technical advantage of a secreted SCT is the rapid purification of functional pMHC-I protein in hours instead of several days using the traditional E. coli. system [52,53]. Hence the time-saving mBEVS method is more likely to create more SCT fusion analogues compared to current tedious in vitro refolding methods, and favor biophysical characterization of pMHC-I molecules [47,54].

**Challenges of developing the EZ MHC-I assay**

However, identifying a peptide exchangeable SCT, which is suitable for EZ MHC-I assay can be challenging. A primary limitation is the wrong choice of a peptide can often result in insoluble SCT inclusion bodies, an indication of misfolded protein. Therefore, different peptides interacting weakly with the A and F pockets of the MHC-I groove were screened and evaluated for secreted soluble SCT protein. Additionally, these peptides unless cleaved, are covalently tethered to the N-terminal of the human β2-microglobulin chain to stabilize the original pMHC-I molecule [48]. More importantly, the tethered peptides should readily dissociate from the MHC-I molecule when cleaved as previously described [1]. For example, known HLA-A*02:01 epitopes KILGRVFFV/ KLTKLITI and HLA-A*02:07 epitope FLPSDYFPSV were found non-exchangeable and thus unsuitable for EZ MHC-I assay. Nonetheless, suitable SCT proteins were successfully produced for EZ MHC-I assay [1].

To encourage actual physico-biochemical measurement of stable MHC-I peptide binding, we have also significantly reduced the time spent in pMHC-I binding assays by eliminating traditional enzyme-linked immunosorbent assay (ELISA) methodology. Here, blocking and washing steps in standard ELISA were removed. Instead the EZ MHC-I assay is a de novo approach of direct protein fragmentation [1]. EZ MHC-I assay is developed based on a combination of unfavorable observations; First, emptied MHC-I proteins were previously known to destabilize and dissociate into α-heavy and β2-microglobulin light chains [55]. Second, fusion protein when destabilized will partially dissociate into α–heavy and β2-microglobulin light chains [56]. Third, enterokinase has been reported to be a non-specific protease in some cases [57,58]. Taken together, these unfavorable observations were successfully incorporated to make a SCT fusion protein, which results in enterokinase-induced fragments in the absence of a rescue peptide.

**In search for more neoantigens**

Pipeline to identify neontigen-specific T cells in blood and tumor samples still remain challenging. Presently, neoantigens due to non-synonymous mutation can be identified using next generation sequencing techniques but still require filtering of tumor DNA against germline DNA and the subsequent identification of private neoantigens unique to different patients. However, the latter requires selecting peptides that are either naturally processed or presented in tumor cells. Moreover, to date, peptide selection using in silico algorithms can still generate a high number of predicted candidates, especially for cancers with high mutational load. Nonetheless, immunogenicity validation of these numerous peptide candidates has been successful in the neoantigen discovery in melanoma and glioblastoma [29,59], but is costly. Thus reducing the peptide pool while improving the quality of peptide candidates are relevant to identifying more neoantigens. However, biophysical characterization of a large number of peptides can be technically laborious and also require the MHC-I molecule. In this commentary, the EZ MHC-I assay uses the pMHC-I SCT molecules and offers a fast and hassle-free approach to screen large peptide libraries which form stable pMHC-I molecules prior to expensive patient sample screening. Besides the EZ MHC-I assay, the rapid mBEVS pMHC-I SCT protein production can also attract more users and promote SCT-based applications for detecting antigenic CD8+ T cells. The C-terminal end, away from the peptide binding groove of the SCT molecule still remains highly modificable. Possible modifications include the additional of a BirA recognition sequence for biotinylation for streptavidin tagging [60], coiled-coil motif for multimerization [61] and the incorporation of clickable chemical groups for bioconjugation to oligonucleotides [62] or as fusion protein [50,63]. Hence the feasible manipulation of the C-terminal end in the pMHC-I SCT molecule would undoubtedly create many tools for cellular cytometry, cellular assays and imaging studies to identify undiscovered tumor neoantigens (Figure 1).

**Conclusion**

Moving forward, the success of neoantigen discovery and cancer vaccine largely requires a pool of predicted MHC peptides with qualities of good affinity and stability. This commentary sheds light on possible missed hits, which remain unaccounted for in silico due to non-conventional environmental factors and opens doors for the EZ MHC-I assay or similar experimental binding/stability assays. Additionally, the use of mBEVS for rapid pMHC-I SCT protein production bearing different C-terminal modifications may create new technologies to unveil anti-tumor CD8+ cytotoxic T cells.

**Acknowledgements**

Authors acknowledge financial support by the Singapore Immunology Network (SIgN) core grant from A*STAR. Authors also acknowledge the National Cancer Centre,
Singapore for financial support by the NCCRF Grant NCCSPG-YR2016-JAN-18.

References

1. Lim J. Destabilizing single chain major histocompatibility complex class I protein for repurposed enterokinase proteolysis. Scientific Reports. 2020 Sep 10;10(1):1-0.


43. Josephs TM, Grant EJ, Gras S. Molecular challenges


