CD8+ T-Cell Epitopes Mapped on the Human Amyloid-Beta Precursor Protein Exhibit Highest Population Coverage in South Korea

Jesther Ryan G. Cinco and Arturo L. Gaitano III*

Immunoinformatics and Novel Synthetic Peptide Immunogens Research (INSPIRe) Group, Chemistry Department, College of Science, Adamson University, Manila, Philippines

*Corresponding Author: arturo.gaitano.research@gmail.com, arturo.gaitano@adamson.edu.ph

ABSTRACT

Alzheimer's disease (AD) is the prevalent form of dementia caused by the accumulation of neurotoxic amyloid-beta (Aß) plaques produced by the cleavage of the human Aß precursor protein (HAPP). Clearance of Aß can be promoted via the apoptosis of peripheral cells where HAPP is overexpressed. This can be achieved by the induction of an immune response via exposure to CD8+ T-cell epitopes from HAPP. Immunoinformatics was used to map these epitopes and the human leukocyte antigens (HLAs) that pair with them. Six immunogenic epitopes were obtained, with one epitope exhibiting "promiscuity"," binding to two different HLAs. A total of seven epitope-HLA pairs were identified with their respective binding free energies and dissociation constants: DTKEGILQY- HLA-A*26:01 (−13.3, 4.20E-10), TPDAVDKY- HLA-B*35:01 (−10.9, 2.10E-08), EVHHQKLVFF- HLA-A*26:01 (−11.1, 1.40E-08), KADKKAVIQHF- HLA-B*58:01 (−10.0, 8.40E-08), AEPQIAMF- HLA-B*44:02 (−9.90, 1.10E-07), AEPQIAMF- HLA-B*44:03 (−8.90, 5.70E-07), and LLPVNGEF- HLA-B*15:01 (−7.00, 1.20E-05). While most of the epitope-HLA pairs exhibited good binding evidenced by their low binding free energies and dissociation constants, AEPQIAMF- HLA-B*44:03 and LLPVNGEF- HLA-B*15:01 did not meet the required threshold. The epitopes are nonallergenic and nontoxic and exhibit significant sequence conservation across different isoforms. They also do not demonstrate significant cross-reactivity with other proteins in the human proteome. The identified epitope-HLA pairs showed highest population coverage in the East Asian region, specifically in South Korea and Japan. These results indicate that these CD8+ T-cell epitopes can be used for a potential peptide vaccine candidate especially in the identified regions.

Keywords: HAPP, CD8+ T-cell epitopes, immunoinformatics, peptide vaccines, Alzheimer's disease

INTRODUCTION

Alzheimer's disease (AD) is the prevalent form of dementia and manifests as a decline in cognitive and physical function and performance (Haque & Levey, 2019). The development and accumulation of neurotoxic amyloid-beta (Aß) plaques due to the imbalance in its production and clearance is the major factor implicated in its development (Selkoe & Hardy, 2016). These Aß plaques are produced by the cleavage of the human Aß precursor protein (HAPP) via proteolysis (Abubakar et al., 2022). Theoretically, the clearance of HAPP, especially before this cleavage into the Aß peptide, should lead into a disruption of its equilibrium and into its clearance, which is a potential therapeutic strategy against AD (Tarasoff-Conway et al., 2015). One way by which this may be achieved is via immunotherapy.

The overexpression of amyloid precursor protein (APP) in transgenic mice has been implicated in their development of AD, and some evidence shows similar correlation in humans (Howlett & Richardson, 2009). Immunotherapy may promote clearance of the Aß peptide from the brain interstitial fluid (central sink) and the cerebrospinal fluid (CSF sink) by decreasing the amount of HAPP in the circulatory system (peripheral sink), in accordance with the "three-sink therapeutic strategy" often utilized against AD (Schreiner et al., 2022). This may be achieved by eliciting the immune system to target the cells in the circulatory system where HAPP is overexpressed, forcing HAPP and related Aß peptides to move from the central and CSF sinks into the peripheral sink. One way by which this may be achieved is by promoting the CD8+ T-cell response.

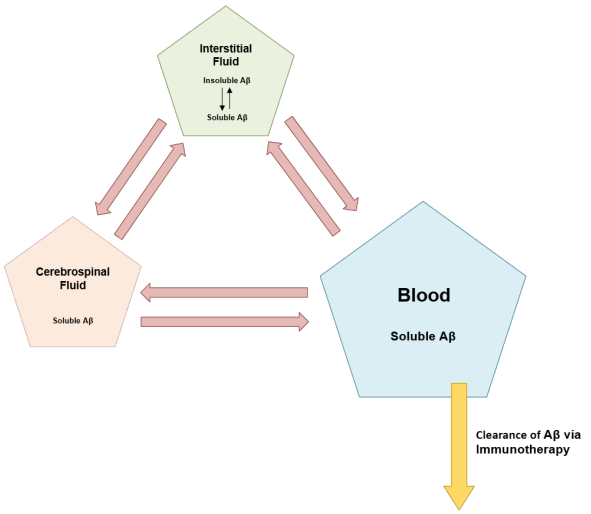

Figure 1. Three-sink therapeutic strategy focused on clearance of amyloid-beta in the circulatory system (blood). Aß = amyloid-beta.

The CD8+ T-cell response, also known as the cytotoxic T-cell response, is the part of the adaptive immune system responsible for the apoptosis, or programmed cell death, of target cells. Stimulating this response has been used as a therapeutic strategy against infectious diseases such as COVID-19 (Chen & John Wherry, 2020) as well as against other diseases like cancer (van der Leun et al., 2020). This immune response is elicited by certain sequences, known as CD8+ T-cell epitopes, that are found on protein antigens. These epitopes are short peptide sequences, usually 8–11 amino acids long, that are loaded into the peptide binding groove of MHC (major histocompatibility complex) Class I proteins to be eventually presented to T-cell receptors, stimulating the CD8+ T-cell response (Pishesha et al., 2022). Identification of these CD8+ T-cell epitopes is key to successful immunotherapy via this approach.

The increase in computing power from the 20th century to the current day has enabled more and more computational tools to be used for biochemical research. These tools have been used for in silico drug discovery for *Mycobacterium tuberculosis* (MTb; Macalino et al., 2020), elucidation of the mechanisms of action of peptide immunogens against malaria (Duay et al., 2023), and vaccine design and epitope mapping via immunoinformatics (De Groot et al., 2020). The mapping of CD8+ T-cell epitopes to be used as potential peptide immunogens can be accelerated via the use of these immunoinformatics tools.

This in silico study provides a biochemical, not a clinical, exploration of the mapping of CD8+ T-cell epitopes in the protein sequence of HAPP via immunoinformatics for potential use as peptide immunogens. These identified epitope sequences were validated via molecular modelling and docking. Relevant safety parameters such as allergenicity, toxicity, and cross-reactivity were assessed. The broadness of efficacy was also estimated via sequence conservation and population coverage analysis, and an overview of molecular interactions was done via residual analysis. The identified CD8+ T-cell epitopes may potentially serve as a foundation for subsequent in vivo experimentation.

MATERIALS AND METHODS

CD8+ T-Cell Epitope Prediction From HAPP Sequence

Mapping of CD8+ T-cell epitopes was done on the HAPP sequence obtained from the National Center for Biotechnology Information (NCBI) protein database (Sayers et al., 2022). HAPP isoform A with accession no. NP_000475.1 was chosen since it is the longest isoform (770 amino acids long) among the three major isoforms that result from alternative splicing of a single gene (Müller et al., 2017). In addition, it is expressed in peripheral systems, specifically in thrombocytes in the blood as well as in the central nervous system (CNS) specifically in the endothelial cells of the blood-brain barrier (Miura et al., 2020). The epitopes were predicted via the NetMHCCons web server, which utilizes a combination of the tools NetMHCpan, NetMHC, and PickPocket (Karosiene et al., 2012). This tool generates predicted epitopes along with their proteasomal cleavage, transporter associated with antigen processing (TAP) transport, and MHC Class-I binding. The resulting epitopes were then filtered based on various criteria such as the presence of the signaling peptide, sites of posttranslational modifications (PTMs), proteasome scores, TAP scores, and MHC IC_{50} (half-maximal inhibitory concentration) scores (Juan et al., 2022).

Assessment of Antigenicity, Allergenicity, and Toxicity

The remaining epitopes are further filtered for various relevant parameters using other web servers. Antigenicity, which is crucial for peptide immunogenicity, was determined via the VaxiJen web server (Doytchinova & Flower, 2007). Allergenicity, measured to ensure that epitopes to be used will not cause unwanted allergic reactions, was assessed via AllergenFP (Dimitrov et al., 2014). Peptide toxicity is another parameter that had to be assessed and was measured via the ToxinPred server (Gupta et al., 2015).

Determination of Cross-Reactivity With the Human Proteome

Once the set of nonallergenic, nontoxic, and antigenic epitopes was determined, they were subjected to BLAST-p (Basic Local Alignment Search Tool for proteins) analysis in the NCBI web server (Stover & Cavalcanti, 2017). This was done to ensure that the identified epitope sequences are not present in other components of the human proteome. This prevents the immune system from attacking these proteins which are not the target of the immunotherapy (Rathore & St. John, 2020).

Analysis of Population Coverage

The population coverage of the identified epitope–human leukocyte antigen (HLA) pairs was analyzed via the population coverage tool (Bui et al., 2006) of the Immune Epitope Database (IEDB; Martini et al., 2020). The HLAs exist as different alleles and are therefore not uniformly distributed across populations (Gao et al., 2019). As such, certain epitope-HLA pairs might be more useful in some regions while other pairs might be more useful in other regions. Ideally, population coverage will be as high as possible, but in cases where it is not, utility of the epitope will be dependent on its specific population coverage.

Molecular Modeling and Docking

Many HLAs have their primary structures elucidated but not their tertiary structures. In these cases, the amino acid sequences for identified HLAs were obtained from the Immuno Polymorphism Database international ImMunoGeneTics project (IPD-IMGT/HLA; Robinson et al., 2013). The tertiary structures of these HLAs were generated via SWISS-MODEL (Waterhouse et al., 2018) and validated using MolProbity, which is built into its web server (Williams et al., 2018). Models that are sufficiently Ramachandran favored were chosen. The interactions between the epitopes and the HLAs were modelled via the GalaxyPepDock

web server (Lee et al., 2015), and the binding energies and dissociation constants were determined via PRODIGY (Xue et al., 2016).

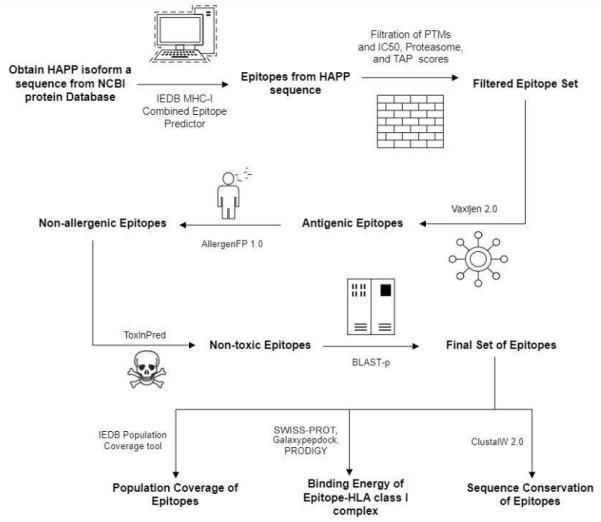

Figure 2. Overview of the study workflow. BLAST-p = Basic Local Alignment Search Tool for proteins; HAPP = human amyloid-beta precursor protein; HLA = human leukocyte antigen; IC_{50} = half-maximal inhibitory concentration; IEDB = Immune Epitope Database; MHC = major histocompatibility complex; NCBI = National Center for Biotechnology Information; PTMs = posttranslational modifications; TAP = transporter associated with antigen processing.

HAPP Sequence Conservancy

The sequences of all available isoforms of HAPP in the NCBI RefSeq database were analyzed for sequence conservation across isoforms. This was done to ensure the efficacy of the potential peptide immunogens against all possible isoforms of HAPP. This was done using the ClustalW alignment tool (Larkin et al., 2007), with the alignment results viewed using Jalview. High sequence conservancy ensures broad utility of the peptide immunogen.

Residue Interaction Analysis

The interactions between specific residues of the epitopes and the HLAs provide a glimpse into the nature and mechanisms of peptide immunogenicity. Residue interaction analysis was done via the iCn3D web server (Wang et al., 2020) to visualize these interactions both in the models of the protein complexes and in a 2D interaction map.

RESULTS AND DISCUSSION

The MHC Class-I binding analysis obtained from the NetMHCCons tool of the IEDB resulted in 82,242 epitopes; these were then subjected to filtering via the MHC IC_{50} score. Epitopes that had a score of >500 nM (recommended by the IEDB) signified poor binding with the specified HLA restricted protein and were thus removed, leading to the number of viable epitopes going down to 1,130. Using information found in the UniProt web server (The UniProt Consortium, 2017), the residues in HAPP with posttranscriptional modifications were identified. The specific peptides and residues for each type of PTM are found in Table 1.

Predicted epitopes containing residues that have these PTMs are removed from the list of viable candidates, since the presence of PTMs alters receptor binding (Houde et al., 2010). Doing so narrows down the number of viable

Table 1. Residues Containing Posttranslational Modifications (PTMs) in HAPP

PTM	Residue Number
Signal peptide	1–17
Modified	198, 206, 217, 262, 336, 441, 497, 729, 730, 743, 757, 761, 762
Glycosylated	542, 571, 633, 651, 652, 656, 659, 663, 667, 681
Disulfide bridge[a]	38-62
	73-117
	98-105
	133-187
	144-174
	158-186
	291-341
	316-337

[a]The "-" signifies the disulfide bond formed between the two cysteine residues.

epitopes to 604. These remaining epitopes were filtered further using a proteasome score and TAP score threshold of ≥1.0 to signify that the epitope would most likely be cleaved by the proteasome with that exact sequence (Nielsen et al., 2005) and that the epitope will be transported into the HLA for antigen presentation (Peters et al., 2020). This narrowed down the number of viable epitopes to 40. Finally, testing for antigenicity, allergenicity, toxicity, and cross-reactivity yielded the final set of six viable CD8+ T-cell epitopes. One of these epitopes is promiscuous (^{30}AEPQIAMF^{37}), meaning it can bind to more than one HLA. Promiscuity of CD8+ T-cell epitopes is potentially advantageous since it allows one epitope to bind to more HLAs. Increased promiscuity therefore often leads to increased population coverage. This makes potential immunotherapeutics based on these epitopes more efficient and gives the epitopes themselves greater utility. The observed promiscuity gives a total of seven epitope-HLA pairs, summarized in Table 2. The locations of these epitopes within the tertiary structure of HAPP is highlighted in Figure 3, which shows that some of these identified CD8+ T-cell epitopes can be found in disordered regions like random coils (C and F), but most are found in relatively ordered regions like α-helices (A, B, and D), and some are found in a mix of both regions (E). Additionally, the breakdown of the number of epitopes that get filtered out after each step is shown in Figure 4.

Due to the inherent polymorphism of the HLA, there will be inherent differences in the HLA allele distribution among certain populations. Globally, the epitope set exhibited only 35.75% population coverage. However, among the different regions, it was found that this set of alleles had the highest population coverage in East Asia, covering

51.44%. Specifically, the highest population coverage was observed in South Korea at 56.98%. Central Africa, Europe, North Africa, North America, Oceania, South Africa, South America, and West Indies all had ^{30}AEPQIAMF^{37} as the highest coverage epitope. ^{425}KADKKAVIQHF^{435} was prevalent in Central America, East Africa, South Asia, and Southeast Asia, while ^{606}LLPVNGEF^{613} is dominant for East Asia and Northeast Asia. ^{64}DTKEGILQY^{72} and ^{682}EVHHQKLVFF^{691} both exhibit highest coverage in Southwest Asia. Lastly, ^{371}TPDAVDKY^{378} was found to have the highest coverage in West Africa.

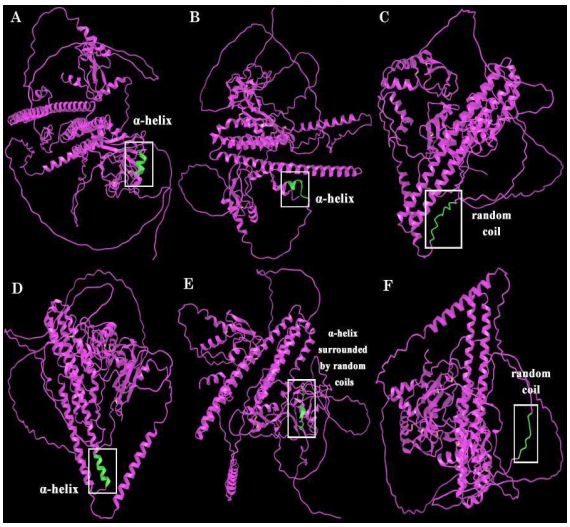

Figure 3. Locations of the viable CD8+ T-cell epitopes on HAPP: A) ^{64}DTKEGILQY^{72}, B) ^{371}TPDAVDKY^{378}, C) ^{682}EVHHQKLVFF^{691}, D) ^{425}KADKKAVIQHF^{435}, E) ^{30}AEPQIAMF^{37}, and F) ^{606}LLPVNGEF^{613}. Viewed via iCn3D. HAPP = human amyloid-beta precursor protein.

The identified HLAs currently do not have experimentally validated tertiary structures in the Protein Data Bank (RCSB-PDB). As such, primary sequences were obtained from IPD-IMGT/HLA, and models were generated via SWISS-MODEL. The validity of the generated models was tested via the 95% minimum percentage requirement of Ramachandran-favored residues, which are illustrated in Figure 5. Ramachandran - favored residues are found within the green areas of the plots, and generated models that are energetically stable must have an overwhelming majority of their residues within this area (Fowler et al., 2020). All models met this criterion and are therefore valid to use for molecular modelling, docking, and residual analysis.

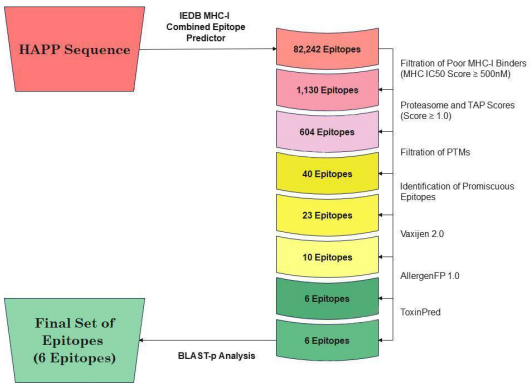

Figure 4. Results of CD8+ epitope filtration from the HAPP sequence. BLAST-p = Basic Local Alignment Search Tool for proteins; HAPP = human amyloid-beta precursor protein; IC_{50} = half-maximal inhibitory concentration; IEDB = Immune Epitope Database; MHC = major histocompatibility complex; PTMs = posttranslational modifications; TAP = transporter associated with antigen processing.

The validated HLA models were docked with the identified epitopes to form the epitope-HLA complexes shown in Figure 6. All of these generated complexes exhibit the signature CD8+ T-cell epitope inserting itself into the peptide binding groove of the HLA composed of two helical structures (Bekker & Kamiya, 2021). This enables the HLAs to load the epitopes that underwent proteasomal cleavage and TAP transport

Table 2. Final Set of Nonallergenic, Nontoxic, Non-Cross-Reactive, Viable CD8+ T-Cell Epitopes

Epitope	MHC Binding Allele	Start	End	Length	MHC IC_{50}	Proteasome Score	TAP Score
DTKEGILQY	HLA-A*26:01	64	72	9	10.5	1.35	1.13
TPDAVDKY	HLA-B*35:01	371	378	8	77	1.21	1.08
EVHHQKLVFF	HLA-A*26:01	682	691	10	216.5	1.54	1.06
KADKKAVIQHF	HLA-B*58:01	425	435	11	117.5	1.25	1.03
AEPQIAMF	HLA-B*44:02	30	37	8	316.1	1.31	1.12
	HLA-B*44:03	30	37	8	444.5	1.31	1.12
LLPVNGEF	HLA-B*15:01	606	613	8	446.9	1.35	1.03

Note. MHC = major histocompatibility complex; IC_{50} = half-maximal inhibitory concentration; TAP = transporter associated with antigen processing.

into its peptide binding groove, allowing for its subsequent presentation on the surface of cells and eventually to bind with CD8+ T-cell receptors (Dhatchinamoorthy et al., 2021).

The different isoforms of HLAs allow for the diversity of responses that CD8+ T-cells can utilize against different pathogens (Reina-Campos et al., 2021). This polymorphism allows them to interact with different epitopes to present to T-cell receptors. In order for these epitopes to be presented, they must form a stable complex with their identified HLA, characterized by an exergonic binding energy and low dissociation constants (Sun et al., 2023). However, a threshold of 500 nM is observed to affirm strong binding of the epitope-HLA pair (Koyanagi et al., 2010; Paul et al., 2013).

The summary of these results is shown in Table 4. While all pairs exhibited negative binding energies (exergonic) and low dissociation constants, ^{30}AEPQIAMF^{37} - HLA-B*44:03 and ^{606}LLPVNGEF^{613} - HLA-B*15:01 failed to attain a dissociation constant equal or lower than 500 nM. In addition, the number and types of interactions between each epitope-HLA pair

Table 3. Population Coverage of the Set of Seven Epitope-HLA Pairs Identified by the Study

Population/Area	Coverage	Population/Area	Coverage
World	35.75%	South Africa	25.69%
Central Africa	31.05%	South America	16.33%
Central America	1.40%	South Asia	30.25%
East Africa	23.68%	Southeast Asia	27.36%
East Asia	51.44%	Southwest Asia	24.88%
Europe	44.18%	West Africa	35.66%
North Africa	28.75%	West Indies	41.32%
North America	40.44%	Korea, South[a]	56.98%
Northeast Asia	30.11%	Japan	50.24%
Oceania	12.95%		

Note. HLA = human leukocyte antigen.

[a]The epitope set exhibits the highest population coverage in this region.

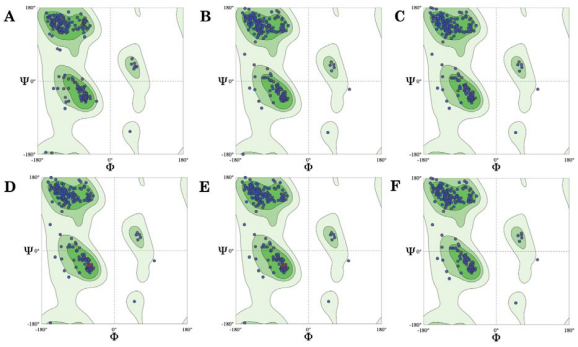

Figure 5. Ramachandran plots of generated models of the different HLAs: A) HLA-A*26:01, B) HLA-B*35:01, C) HLA-B*58:01, D) HLA-B*44:02, E) HLA-B*44:03, and F) HLA-B*15:01. HLA = human leukocyte antigen.

were also analyzed and displayed on Figure 7 as 2D interaction maps. These can be cross-referenced with the calculated values of ΔG and K_D to paint a coherent picture of the correlation between the two sets of data.

Table 4 reveals that the strongest binding energy and binding affinity were observed in the 64DTKEGILQY72 - HLA-A*26:01 pair, while the weakest binding exists in the 606LLPVNGEF613 - HLA-B*15:01 pair. It is reasonable to assume that this means more interactions present in the first complex compared with the second (Alrosan et al., 2022).

Table 4. Binding Energy and Dissociation Constants of Identified Epitope-HLA Pairs at 37°C

Epitope-HLA Complex	Binding Energy (ΔG, kcal/mol)	Dissociation Constant (K_D, M)
^{64}DTKEGILQY^{72} - HLA-A*26:01[a]	−13.3	4.20E-10
^{371}TPDAVDKY^{378} - HLA-B*35:01	−10.9	2.10E-08
^{682}EVHHQKLVFF^{691} - HLA-A*26:01	−11.1	1.40E-08
^{425}KADKKAVIQHF^{435} - HLA-B*58:01	−10.0	8.40E-08
^{30}AEPQIAMF^{37} - HLA-B*44:02	−9.90	1.10E-07
^{30}AEPQIAMF^{37} - HLA-B*44:03	−8.90	5.70E-07
^{606}LLPVNGEF^{613} - HLA-B*15:01	−7.00	1.20E-05

Note. HLA = human leukocyte antigen.

[a]The epitope-HLA complex that exhibited the strongest binding and greatest stability.

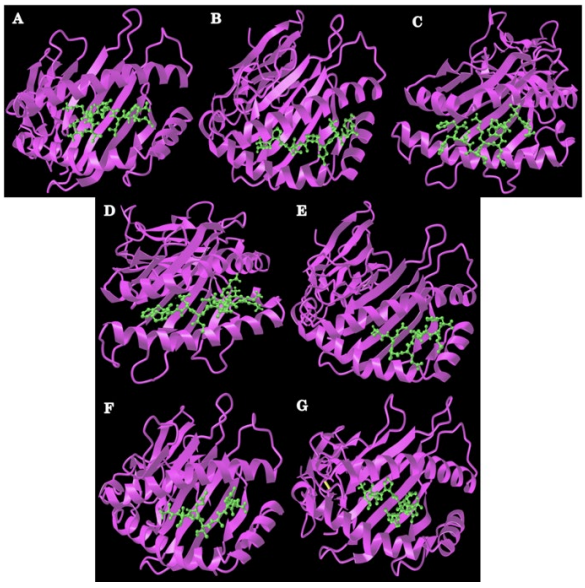

Figure 6. Generated models of the epitope-HLA complexes: A) ^{64}DTKEGILQY^{72} - HLA-A*26:01, B) ^{371}TPDAVDKY^{378} - HLA-B*35:01, C) ^{682}EVHHQKLVFF^{691} - HLA-A*26:01, D) ^{425}KADKKAVIQHF^{435} - HLA-B*58:01, E) ^{30}AEPQIAMF^{37} - HLA-B*44:02, F) ^{30}AEPQIAMF^{37} - HLA-B*44:03, and G) ^{606}LLPVNGEF^{613} - HLA-B*15:01. Generated via GalaxyPepDock and viewed via iCn3D. HLA = human leukocyte antigen.

This difference is easily noticeable from Table 5, since ^{606}LLPVNGEF^{613} - HLA-B*15:01 has a lower number of interactions across the board compared with ^{64}DTKEGILQY^{72} - HLA-A*26:01. However, it is important to note that this relationship is not absolute, since ^{425}KADKKAVIQHF^{435} - HLA-B*58:01 has more contacts while ^{682}EVHHQKLVFF^{691} - HLA-A*26:01 has more of all other types of interaction compared with ^{64}DTKEGILQY^{72} - HLA-A*26:01, but both had lower binding energies and dissociation constants. Other factors might need to be taken into consideration such as the peptide length, the energetics within the peptide itself, or differences in binding efficiency exhibited by different protein-protein complexes (Day et al., 2012).

The increased length of ^{682}EVHHQKLVFF^{691} and ^{425}KADKKAVIQHF^{435} compared to ^{64}DTKEGILQY^{72} could have affected the efficiency of binding since the HLA Class I binding groove is known to be restricted due

to the helices along the groove closing in by the ends (Liao & Arthur, 2011). This could have exposed the peptide to different parts of the HLA protein outside of the binding groove, explaining the increase in contacts. Another consideration could be the dihedral angle strain of the peptide; the residues along the binding grooves could be forcing the peptide into a conformation with high angle strain increasing the instability of the binding (Gloaguen et al., 2020) and therefore making the free energy of binding more positive.

One advantage of this set of epitopes being identified is that two of the HLAs that bind with these epitopes are among the most common alleles found in the South Korean population. These are HLA-B*58:01 and HLA-B*15:01 (Jung et al., 2023). This could partly explain the fact that it exhibits the highest population coverage in that region. However, the dominant allele for South Korea and East Asia in general, HLA-A*24:02, does not exist in this epitope set. This may explain why despite it having the highest population coverage in that region, it is still only at 56.98%. If an epitope in the set was able to strongly bind with HLA-A*24:02, then the population coverage of the potential vaccine could be expanded to cater to most people in South Korea, as well as East Asia. While a definitive percentage cannot be ascertained, it would be expected that the percentage would increase past the current estimate. Globally however, East Asia is still only a small portion of the world, but the potential vaccine could help bring forth new theories or studies in dealing with AD. In addition, the epitope - HLA pairs containing the relevant HLAs, ^{425}KADKKAVIQHF^{435} - HLA-B*58:01 and especially ^{606}LLPVNGEF^{613} - HLA-B*15:01, have weaker binding energies than other candidates, which could affect their immunogenicity. Despite these drawbacks, this CD8+ T-cell epitope set could be used as potential peptide immunogens for HAPP especially in South Korea.

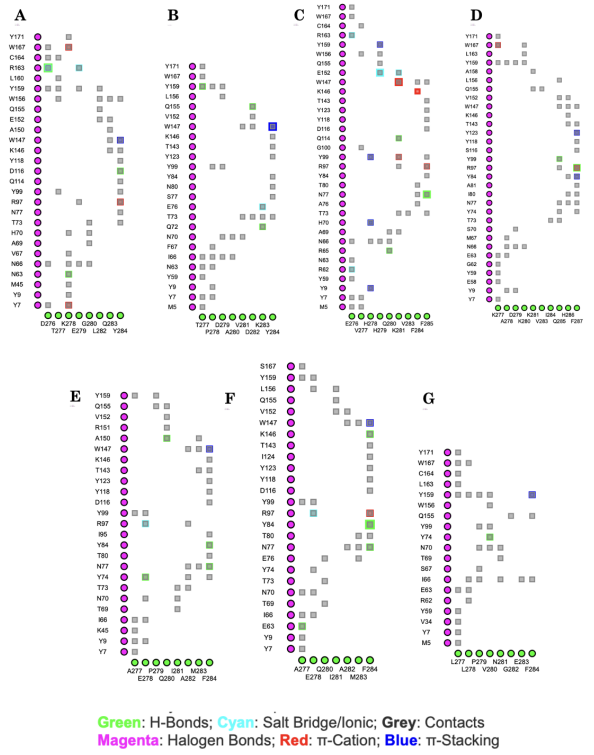

Figure 7. 2D interaction maps of the epitope-HLA complexes: A) ^{64}DTKEGILQY^{72} - HLA-A*26:01, B) ^{371}TPDAVDKY^{378} - HLA-B*35:01, C) ^{682}EVHHQKLVFF^{691} - HLA-A*26:01, D) ^{425}KADKKAVIQHF^{435} - HLA-B*58:01, E) ^{30}AEPQIAMF^{37} - HLA-B*44:02, F) ^{30}AEPQIAMF^{37} - HLA-B*44:03, and G) ^{606}LLPVNGEF^{613} - HLA-B*15:01. Generated via iCn3D. HLA = human leukocyte antigen.

CONCLUSION

The proteolysis of HAPP into the Aß peptide leads to the formation of Aß plaques that are implicated in the development of AD. Clearance of Aß may be achieved via apoptosis of peripheral cells where HAPP is overexpressed. This may be done via a CD8+ T-cell response induced by exposure to CD8+ T-cell epitopes.

Table 5. Summary of the Number and Types of Interactions in Epitope-HLA Complexes

Epitope-HLA Complex	Number of Interactions					
	Hydrogen Bonds	Salt Bridges	Contacts	Halogen bonds	∏-Cation	∏-Stacking
^{64}DTKEGILQY^{72} - HLA-A*26:01[a]	3	1	48	0	3	1
^{371}TPDAVDKY^{378} - HLA-B*35:01	3	1	40	0	0	1
^{682}EVHHQKLVFF^{691} - HLA-A*26:01	4	4	48	0	4	4
^{425}KADKKAVIQHF^{435} - HLA-B*58:01	2	0	51	0	2	2
^{30}AEPQIAMF^{37} - HLA-B*44:02	4	1	41	0	0	1
^{30}AEPQIAMF^{37} - HLA-B*44:03	4	1	39	0	1	1
^{606}LLPVNGEF^{613} - HLA-B*15:01	1	0	32	0	0	1

Note. HLA = human leukocyte antigen

Immunoinformatic analysis mapped these potential CD8+ T-cell epitopes found on HAPP, revealing six nonallergenic, nontoxic, and non-cross-reactive possible epitopes that bind with seven HLAs. This epitope-HLA set demonstrated the highest population coverage in South Korea. Molecular modelling shows the signature binding of the identified epitopes within the peptide-binding groove of the HLAs. All epitope-HLA pairs exhibited exergonic binding energy and low dissociation constants, representing the stability of these complexes for eventual presentation to T-cell receptors. However, the final two epitope-HLA pairs in the set did not meet the required threshold for strong binding. Residual analysis shows significant, but not absolute, correlation between the number and types of interactions between the epitopes and their respective HLAs. While these results show promise in tackling the battle against AD, these epitopes only remain as "potential" vaccine candidates and must be put through extensive testing and scrutiny. Even if in silico tests were done to check for potential harm to the human body like allergenicity, toxicity, and cross-reactivity, the side effects that arise can only be determined through rigorous in vivo experimentation. Overall, immunoinformatics-based epitope mapping was able to provide potential peptide immunogens that may be used to elicit a CD8+ T-cell response against HAPP especially in South Korea.

REFERENCES

Abubakar, M. B., Sanusi, K. O., Ugusman, A., Mohamed, W., Kamal, H., Ibrahim, N. H., Khoo, C. S., & Kumar, J. (2022 Mar. 30). Alzheimer's disease: An update and insights into pathophysiology. *Frontiers in Aging Neuroscience*, *14*, 742408. https://doi.org/10.3389/fnagi.2022.742408

Alrosan, M., Tan, T.-C., Easa, A. M., Gammoh, S., & Alu'datt, M. H. (2022). Molecular forces governing protein-protein interaction: Structure-function relationship of complexes protein in the food industry. *Critical Reviews in Food Science and Nutrition*, *62*(15), 4036–4052. https://doi.org/10.1080/10408398.2021.1871589

Bekker, G.-J., & Kamiya, N. (2021). N-terminal-driven binding mechanism of an antigen peptide to human leukocyte antigen-A*2402 elucidated by multicanonical molecular dynamic-based dynamic docking and path sampling simulations. *The Journal of Physical Chemistry B*, *125*(49), 13376–13384. https://doi.org/10.1021/acs.jpcb.1c07230

Bui, H.-H., Sidney, J., Dinh, K., Southwood, S., Newman, M. J., & Sette, A. (2006). Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*, *7*, 153. https://doi.org/10.1186/1471-2105-7-153

Chen, Z., & John Wherry, E. (2020). T cell responses in patients with COVID-19. *Nature Reviews Immunology*, *20*(9), 529–536. https://doi.org/10.1038/s41577-020-0402-6

Day, E. S., Cote, S. M., & Whitty, A. (2012). Binding efficiency of protein–protein complexes. *Biochemistry*, *51*(45), 9124–9136. https://doi.org/10.1021/bi301039t

De Groot, A. S., Moise, L., Terry, F., Gutierrez, A. H., Hindocha, P., Richard, G., Hoft, D. F., Ross, T. M., Noe, A. R., Takahashi, Y., Kotraiah, V., Silk, S. E., Nielsen, C. M., Minassian, A. M., Ashfield, R., Ardito, M., Draper, S. J., & Martin, W. D. (2020). Better epitope discovery, precision immune engineering, and accelerated vaccine design using immunoinformatics tools. *Frontiers in Immunology,* *11*, 442. https://doi.org/10.3389/fimmu.2020.00442

Dhatchinamoorthy, K., Colbert, J. D., & Rock, K. L. (2021). Cancer immune evasion through loss of MHC Class I antigen presentation. *Frontiers in Immunology*, *12*, 636568. https://doi.org/10.3389/fimmu.2021.636568

Dimitrov, I., Naneva, L., Doytchinova, I., & Bangov, I. (2014). AllergenFP: Allergenicity prediction by descriptor fingerprints. *Bioinformatics*, *30*(6), 846–851. https://doi.org/10.1093/bioinformatics/btt619

Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, *8*(1), 4. https://doi.org/10.1186/1471-2105-8-4

Duay, S. S., Yap, R. C. Y., Gaitano, A. L., Santos, J. A. A., & Macalino, S. J. Y. (2023). Roles of virtual screening and molecular dynamics simulations in discovering and understanding antimalarial drugs. *International Journal of Molecular Sciences*, *24*(11), Article 11. https://doi.org/10.3390/ijms24119289

Fowler, N. J., Sljoka, A., & Williamson, M. P. (2020). A method for validating the accuracy of NMR protein structures. *Nature Communications*, *11*(1), 6321. https://doi.org/10.1038/s41467-020-20177-1

Gao, J., Zhu, C., Zhu, Z., Tang, L., Liu, L., Wen, L., & Sun, L. (2019). The human leukocyte antigen and genetic susceptibility in human diseases. *Journal of Bio-X Research*, *2*(03), 112–120. https://doi.org/10.1097/JBR.0000000000000044

Gloaguen, E., Mons, M., Schwing, K., & Gerhards, M. (2020 Nov. 25). Neutral peptides in the gas phase: Conformation and aggregation issues. *Chemical Reviews*, *120*(22), 12490–12562. https://doi.org/10.1021/acs.chemrev.0c00168

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., & Raghava, G. P. S. (2015). Peptide toxicity prediction. *Methods in Molecular Biology (Clifton, N.J.)*, *1268*, 143–157. https://doi.org/10.1007/978-1-4939-2285-7_7

Haque, R. U., & Levey, A. I. (2019). Alzheimer's disease: A clinical perspective and future nonhuman primate research opportunities. *Proceedings of the National Academy of Sciences*, *116*(52), 26224–26229. https://doi.org/10.1073/pnas.1912954116

Houde, D., Peng, Y., Berkowitz, S. A., & Engen, J. R. (2010). Post-translational modifications differentially affect IgG1 conformation and receptor binding. *Molecular & Cellular Proteomics*, *9*(8), 1716–1728. https://doi.org/10.1074/mcp.M900540-MCP200

Howlett, D. R., & Richardson, J. C. (2009). The pathology of APP transgenic mice, a model of Alzheimer's disease or simply overexpression of APP? *Histology and Histopathology*. https://digitum.um.es/digitum/handle/10201/35972

Juan, A. A., Palma, K. C., Suarez, M., & Herrera-Ong, L. D. M. (2022). Immunoinformatics-based identification of highly conserved cytotoxic T-cell epitopes in polyprotein pp220 of African swine fever virus. *Biomedical and Biotechnology Research Journal (BBRJ)*, *6*(3), 319. https://doi.org/10.4103/bbrj.bbrj_79_22

Jung, K., Kim, J. G., Shin, S., Roh, E. Y., Hong, Y. J., & Song, E. Y. (2023). Allele and haplotype frequencies of 11 HLA loci in Koreans by next-generation sequencing. *HLA*, *101*(6), 602–612. https://doi.org/10.1111/tan.14980

Karosiene, E., Lundegaard, C., Lund, O., & Nielsen, M. (2012). NetMHCcons: A consensus method for the major histocompatibility complex class I predictions. *Immunogenetics*, *64*(3), 177–186. https://doi.org/10.1007/s00251-011-0579-8

Koyanagi, M., Kerns, J. A., Chung, L., Zhang, Y., Brown, S., Moldoveanu, T., Bix, M. (2010). Diversifying selection and functional analysis of interleukin-4 suggests antagonism-driven evolution at receptor-binding interfaces. *BMC Evolutionary Biology*, *10*(1), 223.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, *23*(21), 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Lee, H., Heo, L., Lee, M. S., & Seok, C. (2015). GalaxyPepDock: A protein–peptide docking tool based on interaction similarity and energy optimization. *Nucleic Acids Research*, *43*(W1), W431–W435. https://doi.org/10.1093/nar/gkv495

Liao, W. W., & Arthur, J. W. (2011). Predicting peptide binding to major histocompatibility complex molecules. *Autoimmunity Reviews*, *10*, 469–473.

Macalino, S. J. Y., Billones, J. B., Organo, V. G., & Carrillo, M. C. O. (2020). In silico strategies in tuberculosis drug discovery.

Molecules, *25*(3), Article 3. https://doi.org/10.3390/molecules25030665

Martini, S., Nielsen, M., Peters, B., & Sette, A. (2020). The Immune Epitope Database and Analysis Resource Program 2003–2018: Reflections and outlook. *Immunogenetics*, *72*(1–2), 57–76. https://doi.org/10.1007/s00251-019-01137-6

Miura, S., Yoshihisa, A., Misaka, T., Yamaki, T., Kojima, T., Toyokawa, M.,…Kitazume, S. (2020). Amyloid precursor protein 770 is specifically expressed and released from platelets. *Journal of Biological Chemistry*, *295*(38), 13194–13201.

Müller, U. C., Deller, T., & Korte, M. (2017). Not just amyloid: Physiological functions of the amyloid precursor protein family. *Nature Reviews Neuroscience*, *18*(5), 281–298.

Nielsen, M., Lundegaard, C., Lund, O., & Keşmir, C. (2005). The role of the proteasome in generating cytotoxic T-cell epitopes: Insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics*, *57*(1), 33–41. https://doi.org/10.1007/s00251-005-0781-7

Paul, S., Weiskopf, D., Angelo, M. A., Sidney, J., Peters, B., & Sette, A. (2013). HLA Class I alleles are associated with peptide-binding repertoires of different size, affinity, and immunogenicity. *The Journal of Immunology*, *191*(12), 5831–5839.

Peters, B., Nielsen, M., & Sette, A. (2020). T cell epitope predictions. *Annual Review of Immunology*, *38*, 123–145. https://doi.org/10.1146/annurev-immunol-082119-124838

Pishesha, N., Harmand, T. J., & Ploegh, H. L. (2022). A guide to antigen processing and presentation. *Nature Reviews Immunology*, *22*(12), 751–764. https://doi.org/10.1038/s41577-022-00707-2

Rathore, A. P., & St. John, A. L. (2020). Cross-reactive immunity among flaviviruses. *Frontiers in Immunology*, *11*, 334. https://doi.org/10.3389/fimmu.2020.00334

Reina-Campos, M., Scharping, N. E., & Goldrath, A. W. (2021). CD8+ T cell metabolism in infection and cancer. *Nature Reviews Immunology*, *21*(11), 718–738. https://doi.org/10.1038/s41577-021-00537-8

Robinson, J., Halliwell, J. A., McWilliam, H., Lopez, R., & Marsh, S. G. E. (2013). IPD—the Immuno Polymorphism Database. *Nucleic Acids Research*, *41*(D1), D1234–D1240. https://doi.org/10.1093/nar/gks1140

Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y.,…Sherry, S. T. (2022). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, *50*(D1), D20–D26. https://doi.org/10.1093/nar/gkab1112

Schreiner, T. G., Menéndez-González, M., & Popescu, B. O. (2022). The "cerebrospinal fluid sink therapeutic strategy" in Alzheimer's disease—From theory to design of applied systems. *Biomedicines*, *10*(7), Article 7. https://doi.org/10.3390/biomedicines10071509

Selkoe, D. J., & Hardy, J. (2016). The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Molecular Medicine*, *8*(6), 595–608. https://doi.org/10.15252/emmm.201606210

Stover, N. A., & Cavalcanti, A. R. O. (2017). Using NCBI BLAST. *Current Protocols Essential Laboratory Techniques*, *14*, 11.1.1–11.1.34. Retrieved August 28, 2024, from https://currentprotocols.onlinelibrary.wiley.com/doi/abs/10.1002/cpet.8

Sun, Y., Young, M. C., Woodward, C. H., Danon, J. N., Truong, H. V., Gupta, S., Winters, T. J., Font-Burgada, J., Burslem, G. M., & Sgourakis, N. G. (2023). Universal open MHC-I molecules for rapid peptide loading and enhanced complex stability across HLA allotypes. *Proceedings of the National Academy of Sciences*, *120*(25), e2304055120. https://doi.org/10.1073/pnas.2304055120

Tarasoff-Conway, J. M., Carare, R. O., Osorio, R. S., Glodzik, L., Butler, T., Fieremans, E., Axel, L., Rusinek, H., Nicholson, C., Zlokovic, B. V., Frangione, B., Blennow, K., Ménard, J., Zetterberg, H., Wisniewski, T., & de Leon, M. J. (2015). Clearance systems in the brain—Implications for Alzheimer disease. *Nature Reviews Neurology*, *11*(8), 457–470. https://doi.org/10.1038/nrneurol.2015.119

The UniProt Consortium. (2017). UniProt: The universal protein knowledgebase. *Nucleic Acids Research*, *45*(D1), D158–D169. https://doi.org/10.1093/nar/gkw1099

van der Leun, A. M., Thommen, D. S., & Schumacher, T. N. (2020). CD8+ T cell states in human cancer: Insights from single-cell analysis. *Nature Reviews Cancer*, *20*(4), 218–232. https://doi.org/10.1038/s41568-019-0235-4

Wang, J., Youkharibache, P., Zhang, D., Lanczycki, C. J., Geer, R. C., Madej, T., Phan, L., Ward, M., Lu, S., Marchler, G. H., Wang, Y., Bryant, S. H., Geer, L. Y., & Marchler-Bauer, A. (2020). iCn3D, a web-based 3D viewer for sharing 1D/2D/3D representations of biomolecular structures. *Bioinformatics (Oxford, England)*, *36*(1), 131–135. https://doi.org/10.1093/bioinformatics/btz502

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303. https://doi.org/10.1093/nar/gky427

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S., & Richardson, D. C. (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science: A Publication of the Protein Society*, *27*(1), 293–315. https://doi.org/10.1002/pro.3330

Xue, L. C., Rodrigues, J. P., Kastritis, P. L., Bonvin, A. M., & Vangone, A. (2016). PRODIGY: A web server for predicting the binding affinity of protein–protein complexes. *Bioinformatics*, *32*(23), 3676–3678. https://doi.org/10.1093/bioinformatics/btw514