

## The Importance of Hydrophobic Peptides in Peptide Pools for T-Cell Activation

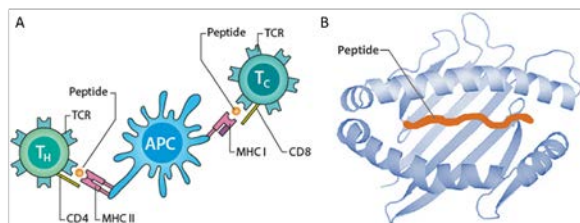
Sebastian Pfeil<sup>1</sup>, Maren Eckey<sup>1</sup>, Pavlo Holenya<sup>1</sup>, Florian Kern<sup>1,2</sup>, Karsten Schnatbaum<sup>1</sup>

<sup>1</sup>JPT Peptide Technologies GmbH, Hermann-Dorner-Allee 23, 12489 Berlin/Germany. <sup>2</sup> Brighton and Sussex Medical School, Brighton, UK

Peptide pools are essential tools for measuring antigen-specific T-cell responses in immune monitoring and other applications. Most peptide pools contain hydrophobic peptides that arise, for example, when overlapping peptides cover the entire sequence of hydrophobic antigens. However, some manufacturers sometimes omit difficult hydrophobic peptides, potentially reducing the T-cell stimulation efficacy of the resulting peptide pool. Here, we analyzed the hydrophobicity of the entire human and several viral peptidomes, revealing that human MHC-binding peptides are on average more hydrophobic than the average human peptide. This emphasizes the importance of hydrophobic peptides in antigen recognition by the human immune system. Since 2004 JPT has pioneered the manufacturing of such peptides and was the first company to offer and further develop the PepMix™ format, which ensures the presence of every peptide in overlapping peptide pools. Based on 20 years of experience in this area, JPT ensures the inclusion of even the most difficult-to-synthesize, hydrophobic peptide sequences in all peptide pools, despite the challenges in synthesizing and purifying such peptides.<sup>1,2</sup> This commitment guarantees 100% coverage of all target protein sequences in protein-spanning peptide pools, promotes effective T-cell stimulation, and maximizes research outcomes.

### Introduction

The recognition of peptides that are presented by class-I and class-II MHC molecules is a key step in T-cell activation. MHC molecules have a peptide binding groove in which peptides are embedded, and only peptides whose side-chains fit the binding groove can be accommodated. Concomitant interaction of the T-cell receptor with parts of the MHC-molecule as well as with certain amino acids of the embedded peptide triggers T-cell activation (Figure 1).



**Figure 1:** A: MHC class I and II molecules with bound peptides interacting with a complementary T-cell receptor (TCR). This highly simplified diagram shows an antigen presenting cell (APC), cytotoxic T-cell (Tc, CD8 positive), helper T-cell (Th, CD4 positive), and major histocompatibility complex class-I and II molecules (MHC). B: Typical structure of the peptide-binding groove of human MHC I.

While it is obvious that the sequence of a peptide is the most decisive determinant for its ability to activate T-cells, it is less clear how specific physicochemical properties like its polarity, for example, affect this ability. The polarity of a peptide is mainly determined by its amino acid composition. While a higher proportion of polar amino acids (like Lysine or Arginine) makes peptides more polar (hydrophilic), a higher proportion of hydrophobic amino acids (like Phenylalanine, Leucine, Isoleucine or Valine) makes peptides more unpolar (hydrophobic). Several models have been developed to quantitatively describe the polarity of peptides (see Table 2 in the appendix for examples). It is known that especially very hydrophobic peptides are difficult to synthesize and purify. One main reason is that so-called “difficult

sequences” (which often contain a high number of hydrophobic amino acids), tend to form secondary structures within the peptide and therefore have a high aggregation potential. This compromises effective amino acid coupling and deprotection during solid-phase peptide synthesis. In addition, hydrophobic aggregating peptides suffer from low solubility in aqueous and often even organic solvents, making such peptides difficult to purify and handle<sup>2</sup>. JPT has decade-long experience in the synthesis of hydrophobic peptides and addresses this challenge with optimized protocols, including but not limited to the use of optimized coupling protocols, pseudoproline building blocks and solubility tags. On the contrary, some peptide manufacturers simply omit hydrophobic peptides from antigen-spanning peptide pools when experiencing difficulties in preparing them.

It was previously reported that there is a bias towards hydrophobic amino acids at T-cell receptor contact residues in published T-cell epitopes.<sup>3</sup> Also, published studies have shown that the peptide binding grooves of a number of human MHC molecules, for example, HLA-A\*02:01, HLA-B\*07:02, HLA-C\*07:01, and HLA-DRB1\*01:01, exhibit a preference for peptides with hydrophobic residues. These hydrophobic pockets have a critical role in peptide binding and presentation to T-cells.<sup>4-7</sup>

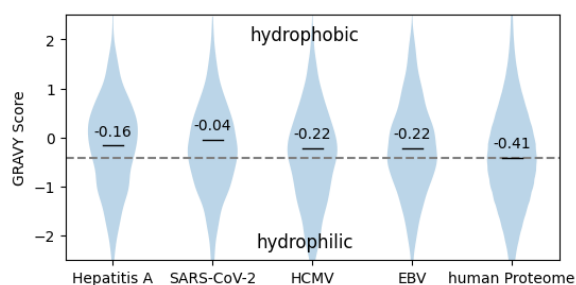
This application note further examines if increased peptide hydrophobicity makes a peptide more likely to be recognized by T-cells, and if, as a result of this, particular attention must be paid to the presence of hydrophobic peptides in protein-spanning peptide pools.

### Results

We analysed a range of known and predicted T-cell-stimulating peptides in regards to their hydrophobicity. First, we dissected the human proteome into overlapping 9-amino-acid-peptides (9/8 scan) *in silico*. This resulted in approximately 30 million peptides. To predict the hydrophobicity of these peptides we calculated their GRAVY (grand average of hydropathy) score.<sup>8</sup> The GRAVY score is widely used to estimate the hydrophobicity of peptides and is the sum of hydropathy values assigned to each amino acid, divided by the length of the sequence. It ranges from -4.5 to +4.5 with higher

PepMix™ Peptide Pools

values representing higher hydrophobicity. The mean GRAVY score of all *in silico*-generated human proteome-derived peptides was -0.41. In contrast to this, all 8-10 amino acid-long T-cell epitopes listed in the Immune Epitope Database (IEDB, ~216,000 peptides)<sup>9</sup> had a mean GRAVY score of 0.04. Thus human MHC-binding peptides were on average slightly more hydrophobic than the human proteome as a whole (difference in GRAVY score: 0.45 on a scale of -4.5 to +4.5). This was also true for several extensively researched viral proteomes analysed in the same way: the proteomes of Hepatitis A, SARS-CoV-2, EBV, and HCMV all had higher average GRAVY scores than the human proteome (Figure 2).



**Figure 2:** The Gravy scores of the proteomes of common human viruses compared to the human proteome (negative: hydrophilic; positive: hydrophobic). For this analysis, proteomes were divided into peptides with 9 amino acids in length and 8 amino acids overlap. The lines represent the mean Gravy score values with the actual values indicated. Dotted line: mean Gravy score of the human proteome.

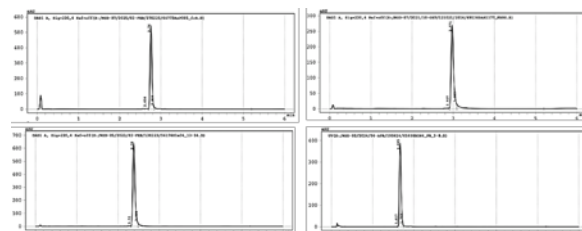
These interesting observations might suggest that non-human viral peptides, on average, have a higher chance to be presented in the human MHC. However, it will be important to consider individual T-cell-stimulating peptides to investigate this further. We, therefore, in the following had a more detailed look at individual T-cell-stimulating peptides listed in the IEDB database.<sup>9</sup>

The peptides shown in Table 1 represent examples for frequently published viral epitopes in the context of HLA-A\*02:01 (IEDB database). The fact that they are highly hydrophobic (GRAVY scores above 1.5) is a clear indication for the importance of hydrophobic peptides in antigen-specific T cell responses.

**Table 1:** Overview of the 10 most frequently published HLA-A\*02:01-presented viral epitopes in IEDB that have a GRAVY score of > 1.5; PMID: unique PubMed ID.

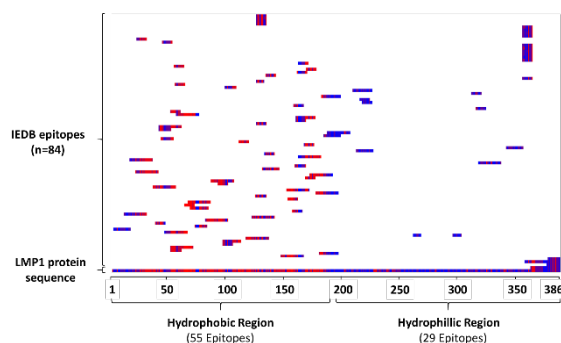
Peptide Sequence	Organism	Gravy Score	PMID Entries
NLVPMVATV	Human cytomegalovirus (HCMV)	1.58	365
GILGFVFTL	Influenza A virus (H1N1)	2.27	278
GLCTLVAML	Epstein-Barr virus (EBV, HHV4)	2.30	169
KLVALGINAV	Hepatitis C virus (HCV)	1.63	110
CLGGLTMTV	Epstein-Barr virus (EBV, HHV4)	2.05	105
ELAGIGILTV	Vaccinia virus (VACV)	1.76	104
LLFGYPVYV	Primate T-lymphotropic virus (PTLV-1)	1.57	91
WLSLLVPFV	Hepatitis B virus (HBV)	2.14	54
LLFNILGGWV	Hepacivirus hominis	1.77	49
YVLDHLIVV	Epstein-Barr virus (EBV, HHV4)	1.86	43

JPT has extensive experience in the synthesis of viral epitopes. We achieved high purities even with difficult-to-synthesize and hydrophobic peptides. As an example, Figure 3 shows the top 4 listed IEDB epitopes in Table 1 (NLVPMVATV, GILGFVFTL, GLCTLVAML, KLVALGINAV). The corresponding LCMS profiles of the purified peptides demonstrates the high quality of the preparation.



**Figure 3:** LCMS profiles demonstrating high purity for the peptides NLVPMVATV (top left), GILGFVFTL (top right), GLCTLVAML (bottom left), and KLVALGINAV (bottom right).

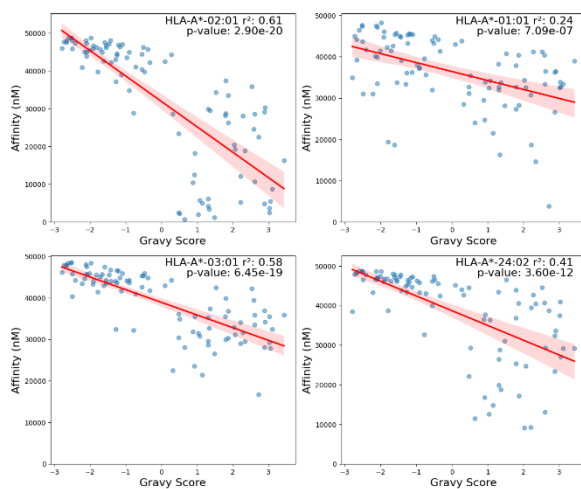
To further explore the importance of peptide hydrophobicity for MHC presentation we investigated the Latent Membrane Protein 1 (LMP1) of the Epstein-Barr virus (EBV, HHV4). This very hydrophobic protein is one of only a few viral proteins that are expressed in malignancies caused by this highly endemic oncogenic human virus. As such, it is of great interest as a T-cell antigen. A detailed examination of published LMP1 epitopes (IEDB) revealed that these are predominantly contained in hydrophobic sequence regions (n=55 published epitopes) (Figure 4, left). A smaller number of epitopes (n=29 published epitopes) is found in the more hydrophilic sequence stretches (Figure 3, right).



**Figure 4:** Alignment of the LMP1-based epitope sequences obtained from IEDB to the LMP1 protein sequence. The complete LMP1 sequence (amino acid 1-386) is located at the bottom. The 84 epitopes published in IEDB are shown above, arranged according to their position in the protein. The Jalview colour scheme hydrophobic<sup>11</sup> was used, where red represents hydrophobic amino acids and blue represents non-hydrophobic amino acids.

Finally, we performed binding predictions on peptides derived from a scan through the LMP1 protein (using the NetMHCpan 4.1 prediction algorithm at <https://services.healthtech.dtu.dk>). Because HLA-A\*02:01 is the most abundant HLA-allele in people of European (50%) and East Asian (30-40%) descent and actually worldwide,<sup>10</sup> we initially focused our analysis on this allele. This showed that hydrophobic peptides had a much higher predicted binding affinity to HLA-A\*02:01 compared with hydrophilic peptides. More specifically, the peptide GRAVY score<sup>8</sup> for the

peptides showed a strong positive correlation with the predicted peptide/MHC binding affinity ( $r^2=0.61$ , Figure 5, top left). Note that we consider correlations strong if  $r \geq 0.6$  ( $r^2 \geq 0.36$ ), and moderate if  $r \geq 0.4$  ( $r^2 \geq 0.16$ ). And please note that the predicted affinity is given in nM IC50, and thus lower numbers represent higher affinity. The preference of HLA-A\*02:01 for hydrophobic peptides might at least in part be explained by the fact that its anchor amino acids are hydrophobic.<sup>12</sup> However, also for peptides binding to HLA-A\*03:01 we observed a clear preference for hydrophobic peptides. (correlation  $r^2=0.58$ ) This means that the observed effect cannot be explained by anchor amino acid hydrophobicity alone, since the main anchor amino acid of HLA-A\*03:01 is the very polar Lysine.<sup>12</sup> A clear preference for hydrophobic peptides was also shown for HLA-A\*24:02 ( $r^2=0.41$ ) and HLA-A\*01:01 ( $r^2=0.24$ ), where the observed correlations between predicted peptide binding and GRAVY score were strong and moderate, respectively.



**Figure 5:** HLA-A\*02:01, HLA-A\*01:01, HLA-A\*03:01 and HLA-A\*24:02 binding prediction of LMP1-derived peptides. The predicted affinity (nM, NetMHC-4.1) is plotted against the GRAVY score of all peptides from a 15/11 scan through LMP1. The Pearson's correlation coefficient (depicted as  $r^2$ ) is shown at the top of each plot.

### Summary

Here, we present the results of an analysis of known and predicted T-cell-stimulating peptides in regards to their hydrophobicity. The analysis shows that the 8-10 amino-acid-long T-cell stimulating peptides listed in IEDB are on average more hydrophobic than the average peptide from the human proteome. The average peptide from several common virus proteomes was also more hydrophobic than the average human peptide. In agreement with this, the most frequently published viral epitopes in the IEDB database are characterized by strong hydrophobicity. The most common HLA-allele worldwide, HLA-A\*02:01, showed a clear preference for binding hydrophobic amino acids, as indicated by the fact that overall peptide hydrophobicity was strongly positively correlated with the predicted HLA-A\*02:01 binding affinity. For three other HLA alleles analysed in addition the correlation was moderate to strong. Based on our findings, it can be concluded that special attention must be paid to the successful chemical synthesis of hydrophobic peptides, because their inclusion in protein-spanning peptide pools is essential for complete epitope coverage.

### Materials & Methods

#### Bioinformatics Analyses

Published LMP1 epitopes were obtained from the IEDB database (<https://www.iedb.org/>, downloaded on 06/2024), aligned using Clustal Omega (<https://www.ebi.ac.uk/jdispatcher/msa/clustalo>) and colored using Jalview (Version: 2.11.4.1) The binding predictions were carried out using NetMHC 4.1 (<https://services.healthtech.dtu.dk/services/NetMHC-4.1/>), only considering the alleles HLA-A\*02:01, HLA-A\*01:01, HLA-A\*03:01 and HLA-A\*24:02. GRAVY score calculation, correlation analysis and graphical representation were carried out using Python programming language.

#### References

- Schnatbaum K, Holenya P, Pfeil S, Drosch M, Eckey M, Reimer U, Wenschuh H, Kern F. An Overview of Peptides and Peptide Pools for Antigen-Specific Stimulation in T-Cell Assays. *Methods Mol Biol.* 2024;2768:29-50. doi: 10.1007/978-1-0716-3690-9\_3. PMID: 38502386.
- Mueller LK, Baumruck AC, Zhdanova H, Tietze AA. Challenges and Perspectives in Chemical Synthesis of Highly Hydrophobic Peptides. *Front Bioeng Biotechnol.* 2020 Mar 4;8:162. doi: 10.3389/fbioe.2020.00162. PMID: 32195241; PMCID: PMC7064641.
- Chowell D, Krishna S, Becker PD, Cocita C, Shu J, Tan X, Greenberg PD, Klavinskis LS, Blattman JN, Anderson KS. TCR contact residue hydrophobicity is a hallmark of immunogenic CD8+ T cell epitopes. *Proc Natl Acad Sci U S A.* 2015 Apr 7;112(14):E1754-62. doi: 10.1073/pnas.1500973112. Epub 2015 Mar 23. PMID: 25831525; PMCID: PMC4394253.
- Madden DR. The three-dimensional structure of peptide-MHC complexes. *Annu Rev Immunol.* 1995;13:587-622. doi: 10.1146/annurev.iy.13.040195.003103. PMID: 7612235.
- Sidney J, Peters B, Frahm N, Brander C, Sette A. HLA class I supertypes: a revised and updated classification. *BMC Immunol.* 2008 Jan 22;9:1. doi: 10.1186/1471-2172-9-1. PMID: 18211710; PMCID: PMC2245908.
- Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanović S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics.* 1999 Nov;50(3-4):213-9. doi: 10.1007/s002510050595. PMID: 10602881.
- Sette A, Sidney J. HLA supertypes and supermotifs: a functional perspective on HLA polymorphism. *Curr Opin Immunol.* 1998 Aug;10(4):478-82. doi: 10.1016/s0952-7915(98)80124-6. PMID: 9722926.
- Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol.* 1982 May 5;157(1):105-32. doi: 10.1016/0022-2836(82)90515-0. PMID: 7108955.
- Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A, Peters B. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* 2018 Oct 24. doi: 10.1093/nar/gky1006. PMID: 30357391; PMCID: PMC6324067 Middleton D, Menchaca L, Rood H, Komerofsky R. New allele frequency database: <http://www.allele-frequencies.net>. *Tissue Antigens.* 2003 May;61(5):403-7. doi: 10.1034/j.1399-0039.2003.00062.x. PMID: 12753660.
- Middleton D, Menchaca L, Rood H, Komerofsky R. New allele frequency database: <http://www.allele-frequencies.net>. *Tissue Antigens.* 2003 May;61(5):403-7. doi: 10.1034/j.1399-0039.2003.00062.x. PMID: 12753660.
- Procter JB, Carstairs GM, Soares B, Mourão K, Ofoegbu TC, Barton D, Lui L, Menard A, Sherstnev N, Roldan-Martinez D, Duce S, Martin DMA, Barton GJ. Alignment of Biological Sequences with Jalview. *Methods Mol Biol.* 2021;2231:203-224. doi: 10.1007/978-1-0716-1036-7\_13. Erratum in: *Methods Mol Biol.* 2021;2231:C1. doi: 10.1007/978-1-0716-1036-7\_18. PMID: 33289895; PMCID: PMC7116599.
- Bassani-Sternberg M, Chong C, Guillaume P, Solleder M, Pak H, Gannon PO, Kandalaf LE, Coukos G, Gfeller D. Deciphering HLA-I motifs across HLA peptidomes improves neo-antigen predictions and identifies allosteric regulating HLA specificity. *PLoS Comput Biol.* 2017 Aug 23;13(8):e1005725. doi: 10.1371/journal.pcbi.1005725. PMID: 28832583; PMCID: PMC5584980.

**PepMix™ Peptide Pools**

13. Wimley WC, White SH. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol.* 1996 Oct;3(10):842-8. doi: 10.1038/nsb1096-842. PMID: 8836100.
14. Zhao G, London E. An amino acid "transmembrane tendency" scale that approaches the theoretical limit to accuracy for prediction of transmembrane helices: relationship to biological hydrophobicity. *Protein Sci.* 2006 Aug;15(8):1987-2001. doi: 10.1110/ps.062286306. PMID: 16877712; PMCID: PMC2242586.

**The Author**


**Dr. Sebastian Pfeil**  
sebastian.pfeil@jpt.com  
JPT Peptide Technologies GmbH

Dr. Sebastian Pfeil is engaged in the R&D and Data Management department at JPT Peptide Technologies, focusing on bioinformatic development of peptide

products for clinical research and developing and adapting company-wide databases and production tools. He utilizes his skills in data science and research to drive innovation in scientific and immunological research.

**The Company**

**JPT Peptide Technologies** utilizes a quality management system that is compliant to ISO 9001:2015 standards and provides its customers with innovative peptide solutions for cellular and humoral immune monitoring, seromarker discovery & validation, vaccine target discovery, peptide lead identification & optimization, targeted proteomics, and enzyme profiling. Have a look at the [World of Peptide Pools](#).

**Contact us for further information!**

Email: peptide@jpt.com  
Phone: +49-30-322980-7878

Please visit us online at <https://www.jpt.com> and take a look at our [PepMix™ Peptide Pools](#) !

**Appendix**

**Table 2:** All natural amino acids with their hydrophobicity scales. The amino acids are classified here as follows: red: hydrophobic, blue: non-hydrophobic. In addition to the frequently used Kyte-Doolittle (KD) scale,<sup>8</sup> on which GRAVY score calculations are based, other scales like the Wimley and White (WW)<sup>13</sup> or the transmembrane tendency (TT)<sup>14</sup> hydrophobicity scales were developed.

Aminoacid	GRAVY score Hydrophobicity <sup>4</sup>	WW Hydrophobicity <sup>9</sup>	TT Hydrophobicity <sup>10</sup>
Isoleucine	4.5	0.31	1.97
Valine	4.2	-0.07	1.46
Leucine	3.8	0.56	1.82
Phenylalanine	2.8	1.13	1.98
Cysteine	2.5	0.24	-0.30
Methionine	1.9	0.23	1.40
Alanine	1.8	-0.17	0.38
Glycine	-0.4	-0.01	-0.19
Threonine	-0.7	-0.14	-0.32
Serine	-0.8	-0.13	-0.53
Tryptophan	-0.9	1.85	1.53
Tyrosine	-1.3	0.94	0.49
Proline	-1.6	-0.45	-1.44
Histidine	-3.2	-0.96	-1.44
Glutamic acid	-3.5	-2.02	-2.90
Glutamine	-3.5	-0.58	-1.84
Aspartic acid	-3.5	-1.23	-3.27
Asparagine	-3.5	-0.42	-1.62
Lysine	-3.9	-0.99	-3.46
Arginine	-4.5	-0.81	-2.57