



Interpreting T-Cell Cross-reactivity through Structure: Implications for TCR-Based Cancer Immunotherapy

Dinler A. Antunes^{1,2}, Maurício M. Rigo^{1,3}, Martiela V. Freitas¹, Marcus F. A. Mendes¹, Marialva Sinigaglia¹, Gregory Lizée⁴, Lydia E. Kavraki², Liisa K. Selin⁵, Markus Cornberg^{6,7} and Gustavo F. Vieira^{1,8*}

¹ Núcleo de Bioinformática do Laboratório de Imunogenética (NBLI), Department of Genetics, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil, ² Kavraki Lab, Department of Computer Science, Rice University, Houston, TX, United States, ³ Laboratório de Imunologia Celular e Molecular, Instituto de Pesquisas Biomédicas (IPB), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre, Brazil, ⁴ Lizée Lab, Department of Melanoma Medical Oncology – Research, The University of Texas M. D. Anderson Cancer Center, Houston, TX, United States, ⁵ Selin Lab, Department of Pathology, University of Massachusetts Medical School, Worcester, MA, United States, ⁶ Cornberg Lab, Department of Gastroenterology, Hepatology and Endocrinology, Hannover Medical School, Hannover, Germany, ⁷ German Center for Infection Research (DZIF), Partner-Site Hannover-Braunschweig, Hannover, Germany, ⁸ Programa de Pós-Graduação em Saúde e Desenvolvimento Humano, Universidade La Salle, Porto Alegre, Brazil

OPEN ACCESS

Edited by:

Cyrille J. Cohen,
Bar-Ilan University, Israel

Reviewed by:

Alessandro Poggi,
Ospedale Policlinico
San Martino, Italy
Zsolt Sebestyen,
University Medical Center
Utrecht, Netherlands

*Correspondence:

Gustavo F. Vieira
gustavo.vieira@unilasalle.edu.br

Specialty section:

This article was submitted
to Cancer Immunity
and Immunotherapy,
a section of the journal
Frontiers in Immunology

Received: 24 July 2017

Accepted: 12 September 2017

Published: 04 October 2017

Citation:

Antunes DA, Rigo MM, Freitas MV, Mendes MFA, Sinigaglia M, Lizée G, Kavraki LE, Selin LK, Cornberg M and Vieira GF (2017) Interpreting T-Cell Cross-reactivity through Structure: Implications for TCR-Based Cancer Immunotherapy. *Front. Immunol.* 8:1210. doi: 10.3389/fimmu.2017.01210

Immunotherapy has become one of the most promising avenues for cancer treatment, making use of the patient's own immune system to eliminate cancer cells. Clinical trials with T-cell-based immunotherapies have shown dramatic tumor regressions, being effective in multiple cancer types and for many different patients. Unfortunately, this progress was tempered by reports of serious (even fatal) side effects. Such therapies rely on the use of cytotoxic T-cell lymphocytes, an essential part of the adaptive immune system. Cytotoxic T-cells are regularly involved in surveillance and are capable of both eliminating diseased cells and generating protective immunological memory. The specificity of a given T-cell is determined through the structural interaction between the T-cell receptor (TCR) and a peptide-loaded major histocompatibility complex (MHC); i.e., an intracellular peptide–ligand displayed at the cell surface by an MHC molecule. However, a given TCR can recognize different peptide–MHC (pMHC) complexes, which can sometimes trigger an unwanted response that is referred to as T-cell cross-reactivity. This has become a major safety issue in TCR-based immunotherapies, following reports of melanoma-specific T-cells causing cytotoxic damage to healthy tissues (e.g., heart and nervous system). T-cell cross-reactivity has been extensively studied in the context of viral immunology and tissue transplantation. Growing evidence suggests that it is largely driven by structural similarities of seemingly unrelated pMHC complexes. Here, we review recent reports about the existence of pMHC “hot-spots” for cross-reactivity and propose the existence of a TCR interaction profile (i.e., a refinement of a more general TCR footprint in which some amino acid residues are more important than others in triggering T-cell cross-reactivity). We also make use of available structural data and pMHC models to interpret previously reported cross-reactivity patterns among virus-derived peptides. Our study provides further evidence that structural analyses of pMHC complexes can be used to assess the intrinsic likelihood of cross-reactivity among peptide-targets. Furthermore,

we hypothesize that some apparent inconsistencies in reported cross-reactivities, such as a preferential directionality, might also be driven by particular structural features of the targeted pMHC complex. Finally, we explain why TCR-based immunotherapy provides a special context in which meaningful T-cell cross-reactivity predictions can be made.

Keywords: T-cell cross-reactivity, peptide–MHC complex, cross-reactivity hot-spots, TCR-interacting surface, hierarchical clustering, TCR/pMHC, cancer immunotherapy

1. HYPOTHESIS AND THEORY

1.1. Cellular Immunity, Private Specificity, and T-Cell Cross-reactivity

Cellular immunity relies on T-cell lymphocytes and their ability to produce unique T-cell receptors (TCRs), while humoral immunity relies on B-cell lymphocytes and their ability to produce antibodies (also referred to as B-cell receptors) (1, 2). Combined, these two branches compose the adaptive immunity, a major “upgrade” in the evolution of the immune system, first seen in jawed vertebrates (1, 2). Different from more ancestral mechanisms of innate immunity, adaptive immunity allows creating specific immune responses to virtually any new pathogen encountered by the host organism. It also allows generating immunological memory, protecting the host against future encounters with the same pathogen (3). This new system was essential in facing the threat of viruses, which are incredibly diverse and evolve at an amazing rate (4). While antibodies can neutralize circulating viruses, cytotoxic T-cells can find and eliminate infected cells (i.e., the “hijacked factories” producing new viral particles). In fact, coevolution with viruses is a major factor shaping the complexity and diversity of the mechanisms involved in cellular immunity (5–7).

The key players in this system are the major histocompatibility complex (MHC) molecules, a diverse set of protein receptors capable of binding peptides derived from intracellular proteins and displaying them at the cell surface (5). This allows circulating cytotoxic T-cells to interact directly with these peptide–MHC (pMHC) complexes, using their TCRs. After a complex selection process in early stages of their development (8, 9), T-cells are able to recognize “non-self” pMHC complexes. For instance, a virus-infected cell displays at its surface MHC molecules loaded with virus-derived peptides. These non-self pMHC complexes can trigger a T-cell response that, in turn, eliminates the infected cell. Moreover, the recognition of these non-self pMHCs can generate immunological memory against this particular virus strain (3).

The efficiency of antiviral immunity, however, depends on the ability of an individual to produce and store a pool of memory T-cells (i.e., a T-cell *repertoire*) able to specifically recognize most of the hugely variable pMHC complexes displayed by cells in different tissues. It actually is quite a puzzling task, if one considers (i) the diversity of MHC allotypes of the host (i.e., the number of MHC protein variants in the human population), (ii) the genetic variability of viruses (i.e., peptide diversity), and (iii) the frequency of viral infections. The solution to this puzzle involves a combination of two important features of cellular immunity: (i) *somatic recombination* of TCR-encoding genes and (ii) *T-cell cross-reactivity*. Somatic recombination allows for a potential

combinatorial diversity of TCRs which exceeds 10^{20} (10, 11). Cross-reactivity allows optimizing the repertoire of T-cells for the recognition of most possible targets, despite the limited number of T-cells that can exist in a given individual, at a given time ($\approx 10^{11}$ in humans) (10, 12). Each newly generated T-cell has a unique TCR and is added to the diverse repertoire of circulating T-cells. If activated by a given pMHC, one T-cell generates an entire pool of clone cells (referred to as a *T-cell line*). All these clones display essentially the same TCR and, therefore, are specific to the same (*cognate*) pMHC. However, after being added to the memory pool, some of these T-cells can be recruited in an initial response to a different *heterologous* pMHC (e.g., the same MHC displaying the peptide of a different virus).

T-cell cross-reactivity is defined as the ability of a given T-cell to be activated by two or more heterologous pMHCs (12). This cross-reactivity can even mediate *heterologous immunity*, when a contact with one pathogen generates a partial immunity against a second (heterologous) pathogen (13). Heterologous immunity is a double-edged sword: it can be protective and desired for wide spectrum vaccine development (14, 15), but it can also mediate impaired cellular response, chronic infection and immunopathology (15–18). The stochastic nature of TCR specificity generation entails that each individual has a unique set of TCRs (referred to as *private specificity*) (13). In addition, given the size limit of the T-cell repertoire and the constant challenges with a variety of pathogens, the memory pool of an individual is ever changing (e.g., some T-cell lines expand, others are lost) (19, 20). In time, cross-reactive cells represent an important part of our memory repertoire, and our immunity against every new challenge is directly influenced by our *immunological history* (12, 19, 21–23). Note that there exist some known biases in the somatic recombination process, producing some TCR sequence combinations with higher frequency in a population (24). This phenomenon is referred to as *public TCR usage* and will be discussed later (see section 1.4).

Recent studies are corroborating the idea that T-cell cross-reactivity is the rule, rather than the exception (19, 22, 25, 26), and that structural features involved in specific TCR/pMHC interactions are the main features driving cross-reactive responses against heterologous targets (25, 27–29). Despite all the evidence accumulated in the context of viral immunity and tissue transplantation, integration of T-cell cross-reactivity into other fields of immunology and human health has been rather slow. This delay can be partially explained by the complexity of the mechanisms involved, as well as concerns about the reproducibility of experimental results characterizing T-cell cross-reactivity (26).

In a pioneering study, Wedemeyer and colleagues were able to collect T-cells recognizing a peptide derived from hepatitis

C virus (HCV), from the blood of healthy donors (30) who had no history of infection by HCV. This implied that these HCV-specific T-cells were probably cross-reactive memory cells previously triggered by a heterologous pathogen. In fact, the authors were able to identify a peptide from influenza A virus (IAV) having 77% of sequence similarity with the HCV-derived peptide used to expand the T-cells. They also showed that these cells were able to recognize both peptides, and that T-cells with the same specificity were generated in response to IAV infection. However, a later study by Kaspirowicz et al. (31) suggested that cross-reactivity between these heterologous peptides was rather weak and had a preferential *directionality* from HCV to IAV (i.e., T-cells primed with the HCV-derived peptide also recognize the IAV-derived peptide, but the opposite was usually not true) (31). More recent studies help clarify situations like this, showing that heterologous immunity between viruses is greatly influenced by private specificities and immunological history (19, 23, 32). Therefore, observed results are not solely determined by peptide sequence similarity, but also dependent on the particular T-cells dominating the response (*in vivo*), or the T-cell line selected for the experiments (*in vitro* or *ex vivo*) (24).

The rebirth of T-cell cross-reactivity as a major interest for human health, however, is coming from cancer research. For decades, immunologists have suggested that the same mechanisms involved in antiviral surveillance were also involved in detecting and eliminating cancer cells, which can display MHCs loaded with tumor-specific peptides (33). More recently, the field of cancer immunotherapy has grown as one of the most promising paths for cancer treatment, relying on the mechanisms of cellular immunity to provide personalized therapies that can eliminate tumors in different tissues and even generate protective memory (33–36). A number of TCR-based therapies were put forward, making use of the latest molecular biology technologies to enhance TCR affinity against tumor-specific peptides (37). Unfortunately, the excitement was tempered by safety concerns. These supposedly tumor-specific T-cells can present unexpected T-cell cross-reactivities in some individuals, attacking healthy tissues (38). In fact, off-target toxicity effects have been observed in recent clinical trials, with at least 5 deadly cases reported (39–41). Two of these cases were clearly linked to T-cell cross-reactivity between the targeted tumor-specific peptide (the melanoma-associated antigen MAGE-A3) and a Titin-derived peptide expressed in healthy cardiac cells (39, 42). The peptides involved have only 55% of sequence similarity, exemplifying the great challenge faced by current preclinical screenings. Later analysis using X-ray crystallography confirmed the structural similarity of the corresponding pMHC complexes as the molecular basis for the observed T-cell cross-reactivity (43).

In response to this critical need, new computational approaches are being developed and tested to improve our capacity to screen for potentially dangerous cross-reactivities. Some of these methods involve assessing peptide sequence similarity, while also accounting for protein tissue expression and MHC binding (44, 45). Others are based on pMHC structural similarity (46–48) or some combination of previously mentioned features (49, 50). Despite the incredible challenge at hand and the current limitations of these computational methods, encouraging results are

being reported. For instance, some of these methods can predict the previously mentioned cross-reactivity between the peptides derived from MAGE-A3 and Titin. A better understanding of the mechanisms underlying T-cell cross-reactivity, as well as the relationship between structural features of pMHC complexes and the activation of T-cell clones, is of utmost importance to further improve these computational methods. In turn, such progress will allow us to provide useful predictions that can be directly translated to the clinic.

In the following sections we attempt to connect the dots between the current understanding of pMHC structure and the goal of making safer TCR-based immunotherapies. First, we review structural aspects of the TCR/pMHC interaction and introduce the idea of structural clustering of pMHC complexes (section 1.2). Then we apply clustering methods to both available crystallographic data and modeled pMHC complexes, providing further evidence that pMHC structural information is essential to understand T-cell cross-reactivity (section 1.3). Next, we review how structural features of the pMHC complex can actually shape the TCR repertoire (section 1.4). Going one step further, we hypothesize how the same features might be shaping different patterns of cross-reactivity: they can be responsible for weak cross-reactivity among similar peptide-targets (section 1.5), or, conversely, drive cross-reactive responses among completely unrelated peptide-targets (section 1.6). Finally, we consider the implications of our work for T-cell cross-reactivity prediction and discuss why cancer immunotherapy provides a special context in which meaningful progress can be made (section 1.7).

1.2. Structural Analyses Can Uncover Key Features for T-Cell Activation

For simplicity, we usually talk about cross-reactivity of TCRs that recognize different peptides, but it is important to keep in mind that the TCR does not recognize the peptide itself; it recognizes the combined surface of the pMHC complex (51). Therefore, observed cross-reactivities between peptides are linked to their presentation in the “context” of a particular MHC. Even if two different MHCs are capable of binding the same peptide, which is not common, the resulting pMHC complexes will most likely be different (52). In fact, this is one of the causes for rejection in (allogeneic) tissue transplantation (26, 53). In this study, we focus on cross-reactivity between peptides presented by the same class I MHC. However, cross-reactivity involving different MHCs has also been reported (53, 54), and the discussion presented here can also be extended to that context.

Studies using X-ray crystallography have greatly contributed to the current understanding of the TCR/pMHC interaction, which was recently reviewed by Degauque et al. (26). The TCR structure contains flexible loops that can come in contact with the *TCR-interacting surface* of the pMHC (i.e., the “face” of the pMHC complex exposed to TCR interaction; see Figures S1A–C in Supplementary Material). These loops include the complementarity-determining regions (CDRs), which are the most variable regions of the TCR structure and the result of the previously mentioned somatic recombination. Despite the structural flexibility

of these loops and the possibility of local conformational changes (25, 55), there is a conserved binding mode for the TCR/pMHC complex. Most times, the CDRs corresponding to the α chain of the TCR will interact with the amino-terminal portion of the peptide, while the β chain CDRs will interact with the carboxi-terminal of the peptide, at a particular angle (26, 51) (Figure S1D in Supplementary Material). Note that the general docking mode of a TCR to its cognate pMHC is referred to as the *TCR footprint* (51). Although the mechanisms are still open for debate, recent studies suggest that the orientation of the TCR footprint is guided by genetically imprinted biases (on the TCR) to recognize conserved MHC amino acid residues (i.e., *germline bias*) (26, 29). However, with the accumulation of crystal structures and evidence from new experimental approaches, one can also see that different TCRs establish different interaction networks, and that some interactions on the pMHC surface seem more important than others to trigger recognition by a particular T-cell (24, 29). These special contacts have been previously referred to as *hot-spots* for T-cell cross-reactivity (25, 29, 56).

In previous work, our group described an *in silico* approach to evaluate the structural similarity of pMHC complexes (46, 48). We used hierarchical clustering as a tool to group pMHC complexes according to the similarity of their TCR-interacting surfaces. We also used available crystal structures as a reference to implement a method to model pMHC complexes for which no structural data were available (52, 57). Combining these methods, we were able to reproduce experimentally observed cross-reactivity patterns for a dataset of 28 naturally occurring variants of an HCV-derived peptide used for vaccine development (CINGVCWTV) (46). We also applied these methods to predict potential cross-reactivities between this HCV vaccine peptide and a dataset of non-related virus-derived peptides, in the context of a particular human MHC (HLA-A*02:01) (46). Our predictions were later confirmed by *in vitro* and *ex vivo* experiments (47), highlighting the prospecting potential of our methods. One of the detected cross-reactive peptides, derived from Epstein-Barr virus (LLWTLVLL), shared no sequence similarity with the vaccine peptide. Notwithstanding, both peptides show remarkably similar TCR-interacting surfaces when bound to HLA-A*02:01 (46, 47).

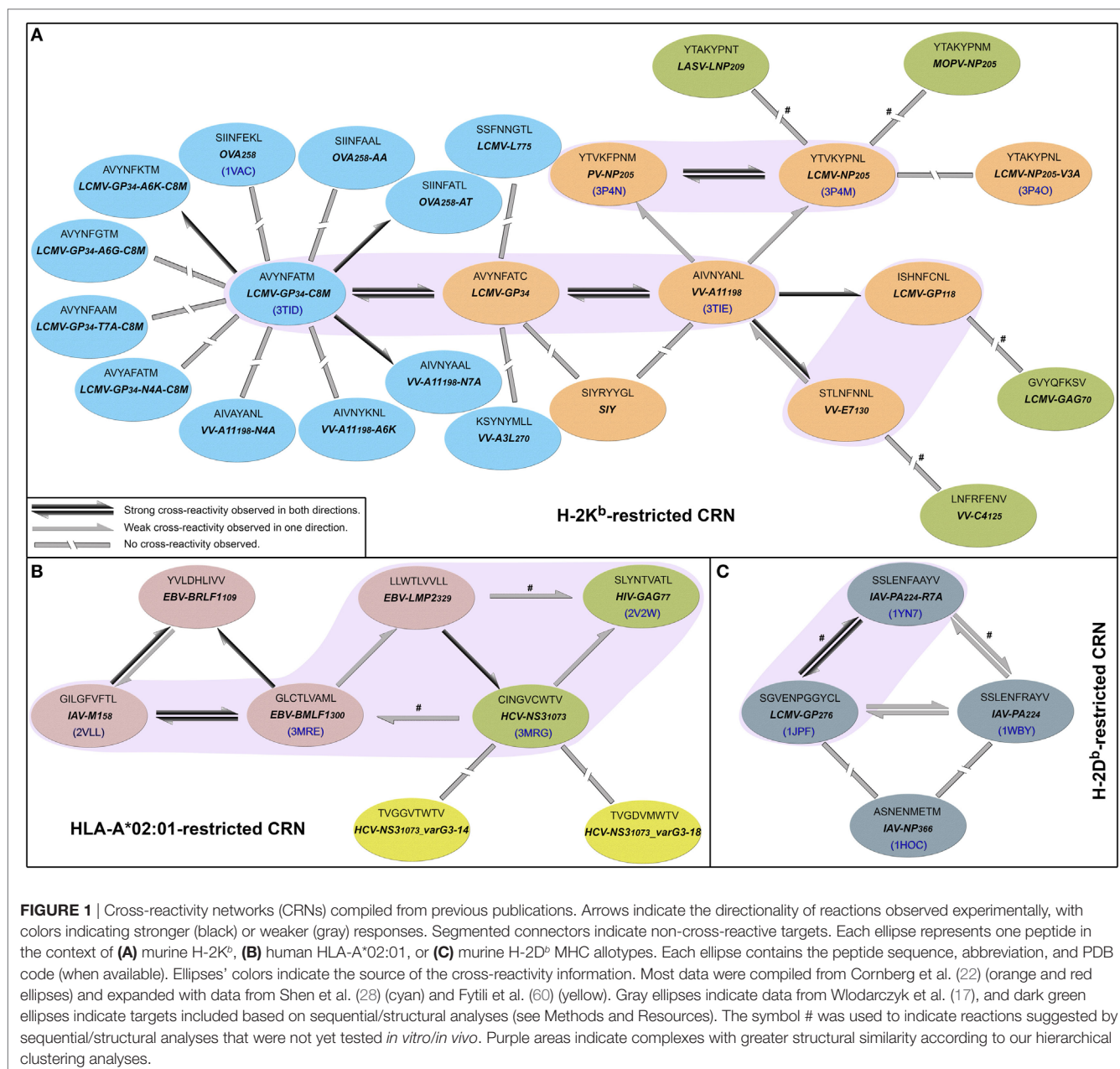
1.3. Structural Similarity of pMHC Complexes Can Reveal Their Likelihood for T-Cell Cross-reactivity

In 2010, Cornberg et al. (22) described *cross-reactivity networks* involving virus-derived peptides, within both human and murine memory T-cell pools (CD8⁺/CD44^{hi}). They used as a reference a peptide derived from vaccinia virus (VV), corresponding to a 9-mer sequence starting at position 198 of the A11 protein (hereafter denoted by VV-A11₁₉₈). Using this VV-derived peptide, which is displayed by the murine MHC H-2K^b, the authors were able to activate three different memory T-cell populations that also recognized peptides from lymphocytic choriomeningitis virus (LCMV-GP₃₄, LCMV-GP₁₁₈, and LCMV-NP₂₀₅). Therefore, VV-A11₁₉₈ could be seen as a cross-reactivity “hub,” connected to all these LCMV-derived peptides (Figure 1A). The concept of

cross-reactivity networks is interesting in highlighting how broad these T-cell cross-reactivities can be (25), sometimes involving completely unrelated targets. In this sense, graphical representations of such networks have been used in previous works (13, 50, 58, 59). However, it is extremely important to keep in mind that despite providing a nice way to visually summarize cross-reactivity relationships, the topology of these networks might not correspond to the cross-reactivities observed for a particular T-cell line. In other words, the “real” topology of the network in terms of T-cell activation depends on which T-cell is used to test these peptide-targets. In this study, we use cross-reactivity networks to summarize the information from previous studies, as a reference to analyze structural data and discuss cross-reactivity patterns (Figure 1). In our representation, each node describes a given peptide, and only peptides displayed by the same MHC are included in a given network (i.e., MHC-restricted network). Note that this is a schematic representation of the known relationships among peptides that are relevant to our discussion, and not a complete picture of known cross-reactivities; it is not expected to reflect the patterns observed in any particular T-cell assay. Additional information on all peptides included in our analysis can be found in Table S1 in Supplementary Material.

In their original study, Cornberg et al. (22) suggested that observed cross-reactivity patterns present a within-individual variation driven by private specificities and immunological history. For instance, the authors were able to collect VV-A11₁₉₈-specific T-cells from mice previously immunized with LCMV (i.e., LCMV-immune mice). Note that if the donor had no previous contact with VV-derived peptides, these VV-A11₁₉₈-specific T-cells should be cross-reactive cells primarily expanded *in vivo* by recognizing some LCMV-derived target. These cells were further expanded *in vitro* with the cognate (VV-A11₁₉₈) peptide and challenged with different peptides derived from LCMV, VV and pichinde virus (PV). Interestingly, these VV-A11₁₉₈-specific T-cells presented cross-reactivity with LCMV-GP₃₄, LCMV-GP₁₁₈, LCMV-NP₂₀₅, and PV-NP₂₀₅ (22) (Figure 2A). However, cross-reactivity against another VV-derived peptide (VV-E7₁₃₀) was not observed. On the other hand, a very different pattern was observed when the authors performed a similar experiment, but expanding VV-A11₁₉₈-specific T-cells from VV-immune mice instead of LCMV-immune mice (Figure 2B). In this case, cross-reactivity with VV-E7₁₃₀ and LCMV-GP₃₄ was observed, but no cross-reactivity was observed with LCMV-GP₁₁₈, LCMV-NP₂₀₅, and PV-NP₂₀₅. These contrasting results suggest the use of a different T-cell population with a different specificity (22). They also suggest a greater structural similarity between VV-A11₁₉₈ and LCMV-GP₃₄, since this cross-reactivity was observed for both LCMV-immune and VV-immune background. In fact, structural similarity between these targets was later confirmed by Shen et al. (28), which solved the crystal structures of VV-A11₁₉₈ and LCMV-GP₃₄-C8M bound to H-2K^b (PDB codes 3TIE and 3TID, respectively).

Out of the 25 pMHCs included in our H-2K^b-restricted network (Figure 1A), at the time of our analysis, only 6 had their structure determined by experimental methods. Using our previously described structure-based approach (46), we performed a hierarchical clustering of these 6 crystallographic structures



(Figure S2 in Supplementary Material). Supported by multiscale bootstrap resampling with the R package *pvclust* (61), the clustering agreed with experimental data. The cross-reactive targets VV-A11₁₉₈ and LCMV-GP₃₄ fall in the same cluster; the same is observed for the highly cross-reactive targets LCMV-NP₂₀₅ and PV-NP₂₀₅. These four targets are closer to one another than to the non-cross-reactive target OVA₂₅₈. Finally, the most different structure in this analysis contained the non-cross-reactive escape variant LCMV-NP₂₀₅-V3A (21, 62).

To expand our analysis, we used the pMHC modeling method implemented in DockTope (52, 57), obtaining the structures of other complexes previously tested by Cornberg et al. (22) (Figure 1A). We also included in this analysis two unrelated

peptides, VV-C4₁₂₅ and LCMV-GAG₇₀, as putative non-cross-reactive controls (Table S1 in Supplementary Material). Our expanded hierarchical clustering reflects the greater structural similarity between VV-A11₁₉₈ and LCMV-GP₃₄, since both complexes fall in the same cluster, with the edge presenting the lowest height and the highest *p*-values (Figure 3). Peptides LCMV-GP₁₁₈ and VV-E7₁₃₀, which are cross-reactive with VV-A11₁₉₈, fall in the next branch, followed by a cluster with the other cross-reactive targets (LCMV-NP₂₀₅ and PV-NP₂₀₅). All these cross-reactive targets were grouped into a bigger cluster (see edge 5 in Figure 3), apart from all the non-cross-reactive targets. As discussed by Cornberg et al. (22), these cross-reactivities could not be easily predicted with peptide sequence

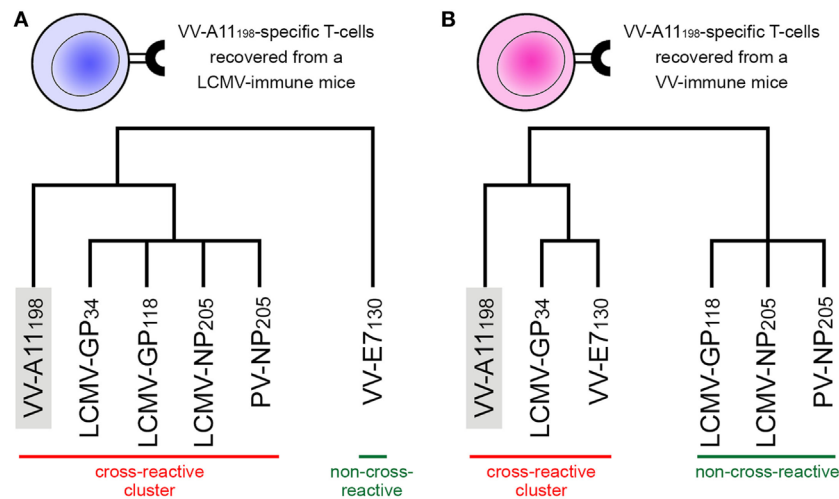


FIGURE 2 | Schematic representation of experimentally observed cross-reactivity patterns. Two alternative dendrograms were drawn to represent alternative outcomes observed in experiments previously performed by Cornberg et al. (22). **(A)** VV-A11₁₉₈-specific T-cells recovered from mice previously immunized with lymphocytic choriomeningitis virus (LCMV) recognize the cognate peptide (indicated by the gray box) as well as three other peptides derived from LCMV and one derived from pichinde virus (PV). We can represent these connections as a “cross-reactivity-cluster” in our dendrogram, as indicated in red. Another peptide derived from vaccinia virus (VV-E7₁₃₀), however, is not recognized. **(B)** VV-A11₁₉₈-specific T-cells recovered from mice previously immunized with vaccinia virus (VV) recognize the cognate peptide (gray box) as well as the other VV-derived peptide (VV-E7₁₃₀) and one LCMV-derived peptide (LCMV-GP₃₄). However, in this experiment, no cross-reactivity was observed against peptides LCMV-GP₁₁₈, LCMV-NP₂₀₅, and PV-NP₂₀₅ (indicated by the green bar). Although targeting the same VV-derived peptide, the alternative cross-reactivity patterns described in panels **(A,B)** reflect the use of different T-cell lines in each experiment (indicated as a blue or pink T-cell). Note that cross-reactivity between VV-A11₁₉₈ and LCMV-GP₃₄ was observed in both experiments, suggesting higher structural similarity of these peptides when displayed by H-2K^b. All peptides involved in these experiments are restricted to the murine MHC H-2K^b. This is a schematic representation, and the heights of the edges in the dendrogram do not capture the actual “distances” among the peptide-targets. Additional information on the presented peptides can be found in Table S1 in Supplementary Material.

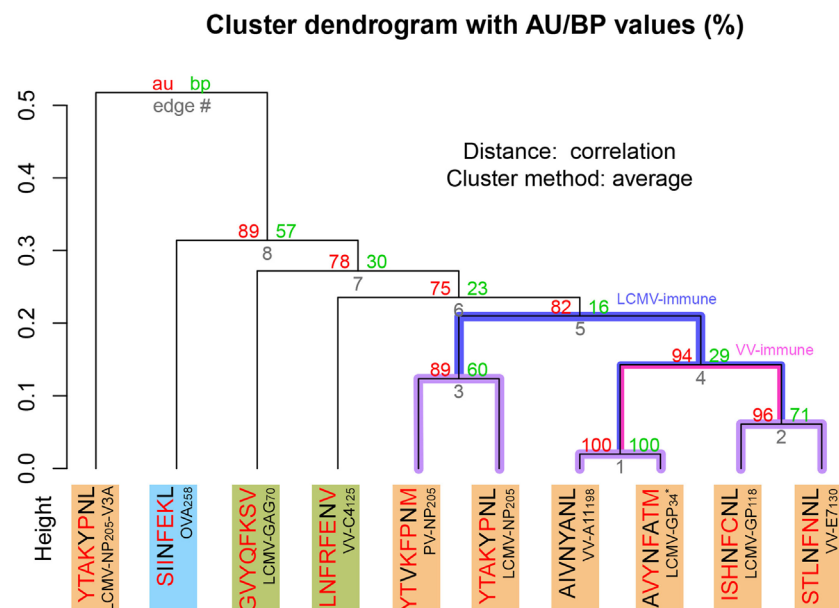


FIGURE 3 | Extended H-2K^b-restricted clustering. Structure-based hierarchical clustering performed with pvclust (61). Each putative cluster is represented by a specific edge (gray numbers), in order of increasing heights (y axis). Cluster confidence is measured with two p-values, approximately unbiased (AU), and bootstrap probabilities (BP). Lines highlighted in purple indicate structures with greater structural similarity (as represented in **Figure 1**). Lines highlighted in blue and pink indicate putative cross-reactivity thresholds for different memory T-cells (see **Figure 2**). Each peptide target is colored according to **Figure 1**. Peptide abbreviation and sequence are provided, with red amino acids indicating changes in relation to VV-A11₁₉₈. *Crystal structure 3TID was used to represent LCMV-GP₃₄, despite presenting a C8M exchange, as indicated by its sequence (see Methods and Resources).

similarity, since all these peptides share less than 50% of their amino acid residues. For instance, sequence similarity between VV-A11₁₉₈ and LCMV-GP₃₄ is only 37.5%, the same as between VV-A11₁₉₈ and the non-cross-reactive target OVA₂₅₈. In spite of that, our results show that this cluster of cross-reactivity involving peptides from three different viruses could be predicted by an *in silico* analysis of the corresponding pMHC structures (see edge 5 in **Figure 3**).

Cross-reactivity was indeed observed among these 6 peptides in the context of H-2K^b (22, 28). These pMHC complexes also present structural similarities, being clustered together in our structure-based hierarchical clustering. However, there was no experimental evidence of one T-cell population able to recognize all six peptides (22). As already discussed, cross-reactivity patterns depend on the specific T-cell population tested. Assuming our clustering correctly captures the relationships among these pMHCs, in terms of structural similarity, we can make some inferences about the T-cells used in the aforementioned experiments. We can say that T-cells from LCMV-immune mice are more cross-reactive, and we can visually represent them with a higher threshold in our clustering analysis (defining the blue cluster in **Figure 3**). Such threshold would correctly predict most of the observed cross-reactivities, with the exception of VV-E7₁₃₀ (which was not recognized). On the other hand, T-cells from VV-immune mice can be represented with a lower threshold (defining the pink cluster in **Figure 3**), since VV-A11₁₉₈-specific T-cells recognize neither NP₂₀₅ peptides. The exception in this case, would be LCMV-GP₁₁₈. These exceptions cannot be predicted considering the information provided by the pMHC structures, since they are most likely driven by TCR variability and private specificities. In spite of that, our data suggest a correlation between pMHC structural similarity and the probability to find cross-reactivity among pMHC targets; that is, the higher the similarity, the higher the likelihood of observing cross-reactive responses. Although cross-reactivity between LCMV-GP₁₁₈ and VV-E7₁₃₀ was not observed using the VV-A11₁₉₈-specific or VV-E7₁₃₀-specific T-cells (22), the similarity of these pMHC complexes (Figure S3 in Supplementary Material) suggests that this cross-reactivity should be observed using another T-cell population; maybe with LCMV-GP₁₁₈-specific T-cells.

1.4. Structural Features of the pMHC Can Shape the TCR Repertoire

More than a decade ago, Turner and colleagues (63) described differences in the T-cell population stimulated by a featureless peptide (referred to as a “vanilla” peptide), and a peptide having a prominent feature exposed to the TCR (hereafter referred to as a “spicy” peptide). The authors used a peptide derived from the polymerase acidic protein of influenza A virus as an example of spicy peptide (IAV-PA₂₂₄, see Table S1 in Supplementary Material). This peptide has an arginine at position 7 (P7), which becomes an exposed feature when displayed by the murine MHC molecule H-2D^b (Figure S4 in Supplementary Material). Immunization with this peptide triggered the expansion of a very diverse pool of T-cells, including cells with high affinity to the target pMHC.

Comparing the response across different animals, the authors noticed great variability in TCR usage. In other words, in each animal the response was dominated by TCRs with unique CDR sequences (i.e., shaped by private specificity).

Surprisingly, opposite results were observed when using a vanilla peptide. Immunization with a mutated version of IAV-PA₂₂₄, replacing the arginine at P7 with an alanine (IAV-PA₂₂₄-R7A), triggered the expansion of a much less diverse T-cell population. In this case, similar CDR sequences were observed for different individuals (i.e., public TCR usage). The same results were observed with a wild-type vanilla peptide (IAV-NP₃₆₆). Therefore, structural features of the pMHC complex can shape the composition of the TCR repertoire during a cellular immune response. A pMHC displaying a vanilla peptide has a TCR-interacting surface dominated by the (self) MHC; given the negative selection of T-cells, very few available TCRs can recognize this complex. This could explain the observation of a less diverse population and the use of public TCRs, sharing a germline bias to interact with the MHC. In addition, we could expect such TCRs to be more cross-reactive, since they rely mostly on (self) MHC features for the recognition. On the other hand, a spicy peptide offers a more evident discerning feature that various TCRs can recognize (in slightly different ways). Given their “focus” on this outstanding feature, we could expect such TCRs to be intrinsically less cross-reactive and they should be incapable (or impaired) to recognize pMHCs lacking such feature.

It is easier to understand this analogy of the spicy feature having in mind some prominent structure that is specific to the peptide, as the examples mentioned earlier and in the next section. However, the TCR/pMHC interaction can be influenced by more subtle features, as recently described by Song et al. (24). They performed a comprehensive evaluation of the T-cell response to the peptide IAV-M1₅₈, displayed by HLA-A*02:01, using the next-generation sequencing of TCRs. In addition, they resolved the crystal structures of two selected TCR/pMHC complexes. IAV-M1₅₈ has been described as a vanilla peptide, since most of its side chains are buried when displayed by HLA-A*02:01. In turn, it was suggested that the lack of recognizable peptide features would lead to a very narrow T-cell response (i.e., lack of TCR diversity among stimulated T-cells). However, Song et al. (24) observed that the IAV-M1₅₈:HLA-A*0201 complex can actually be recognized by a broad range of TCRs; most of them sharing the same V β domain. They were also able to identify a conserved structural feature that seemed to be required for the recognition of this peptide. Interestingly, it was not something “prominent,” and it was not exactly a feature of the peptide alone. In fact, the authors describe a unique exposed pocket between the peptide and the MHC, with which very different TCRs are able to interact. In other words, this particular pocket is a recognizable structural feature that is specific to the IAV-M1₅₈:HLA-A*0201 complex. As a result, in the context of our discussion, we can describe IAV-M1₅₈:HLA-A*0201 as a spicy complex. The lack of a prominent peptide feature might facilitate the selection of some public TCRs, as indeed observed experimentally (24). But the pMHC-specific pocket allows the selection of a broad TCR repertoire, in the same way as for spicy peptides. Once again, these findings highlight the fact that in most cases we

cannot discuss T-cell activation or T-cell cross-reactivity only in terms of peptide-targets, since the key features for recognition might come from the unique combined structure of the pMHC complex.

1.5. Local Structural Differences among pMHC Complexes Can Account for Limited Cross-reactivity and Lack of Reciprocity

In a recent study, Wlodarczyk et al. (17) described a weak cross-reactivity between IAV-PA₂₂₄:H-2D^b and a heterologous complex displaying a peptide derived from lymphocytic choriomeningitis virus (LCMV-GP₂₇₆:H-2D^b, see Table S1 in Supplementary Material). Since crystal structures are available for both complexes, we can visually compare their TCR-interacting surfaces (**Figures 4A–C**). Notably, LCMV-GP₂₇₆:H-2D^b differs from IAV-PA₂₂₄:H-2D^b by not having the featured arginine at P7. As expected, using our structure-based hierarchical clustering, we can see greater proximity (i.e., structural similarity) between LCMV-GP₂₇₆:H-2D^b and IAV-PA₂₂₄:R7A:H-2D^b, than between these complexes and the wild-type (IAV-PA₂₂₄:H-2D^b) or the non-cross-reactive complex IAV-NP₃₆₆:H-2D^b (Figure S5 in Supplementary Material). As described by Wlodarczyk et al. (17), cross-reactivity between GP₂₇₆:H-2D^b and IAV-PA₂₂₄:H-2D^b was weak and showed a preferential directionality. From the pool of T-cells recognizing GP₂₇₆:H-2D^b (primer) it was possible to extract T-cells that also recognize IAV-PA₂₂₄:H-2D^b (i.e., heterologous challenge). However, the reverse experiment was not successful.

Taken together, these results allow us to postulate that immunization with IAV-PA₂₂₄ stimulates a pool of T-cells dominated by clones with high specificity to the spicy feature (in this case, a peptide feature: the R at P7). By challenging with a heterologous peptide that lacks this prominent feature, we would most likely fail to find a T-cell clone that can also recognize the heterologous vanilla peptide-target (e.g., LCMV-GP₂₇₆). However, by using the vanilla peptide as a primer, we would start from a population of T-cells that is less diverse (i.e., dominated by public TCRs) but more cross-reactive. These TCRs are primarily engaging with (self) MHC structural features; some of these clones might also recognize the heterologous spicy peptide (IAV-PA₂₂₄), regardless of the prominent amino acid residue at P7. We believe this recognition might involve some adjustment of the CDR loops around the center of the peptide, as recently discussed by Adams et al. (29). Naturally, some TCRs will not be able to undergo such adjustment and will not show cross-reactivity. We also hypothesize that the “stronger” the spicy feature (or the combination of diverging features), the stronger the directionality and the lower the likelihood of cross-reactivity. Conversely, we believe stronger cross-reactivity should be observed between very similar pMHC complexes, regardless of directionality. For instance, stronger cross-reactivity should be observed between LCMV-GP₂₇₆:H-2D^b and the mutated IAV-PA₂₂₄:R7A:H-2D^b, than with the wild-type (**Figure 1C**).

Additional examples supporting this theory can also be found in the context of human MHCs. By the time of our analysis, out of the 9 virus-derived peptides included in our

HLA-A*02:01-restricted network (**Figure 1B**), only 4 had available crystal structures. We modeled the remaining complexes and performed a hierarchical clustering (Figure S6 in Supplementary Material). As expected, the cross-reactive peptide-targets EBV-BMLF1₃₀₀, IAV-M1₅₈, HCV-NS3₁₀₇₃, HIV-GAG₇₇, and EBV-LMP2₃₂₉ were clustered together (see edge 5 in Figure S6 in Supplementary Material). These last two structures were actually the most similar pair of structures inside this cluster, in agreement with previous clustering results from our group (46).

Two non-cross-reactive variants of HCV-NS3₁₀₇₃ derived from HCV genotype 3, previously referred to as G3-14 and G3-18 (46, 60), fell in separate branches. Despite being the outermost branch of the main cluster (see edge 6 in Figure S6 in Supplementary Material), the small distance between G3-14 and the cross-reactive targets suggest that cross-reactivity with this HCV-derived escape variant might be observed depending on the T-cell population tested. Interestingly, the complex presenting EBV-BRLF1₁₀₉ falls in the same branch as G3-18, which is far from its cross-reactive target (EBV-BMLF1₃₀₀). This HCA result was due to a negatively charged spot in the surface of the EBV-BRLF1₁₀₉:HLA-A*0201 complex, which was not seen in its cross-reactive counterparts (**Figures 4D–F**). If we remove from our analysis this negatively charged spot, EBV-BRLF1₁₀₉ is clustered with EBV-BMLF1₃₀₀ (**Figure 5**). Note that we have had access to a yet unpublished crystal structure of EBV-BRLF1₁₀₉:HLA-A*0201, recently resolved by the team of Dr. Lawrence Stern (UMass Medical School, MA, USA), which confirms the existence of the outstanding negatively charged spot observed in our model (Song I, personal communication, June 2017).

Similar to the situation described for IAV-PA₂₂₄, cross-reactivities involving EBV-BRLF1₁₀₉ feature several peculiarities. For instance, they are not observed for most T-cell populations and normally respect a given directionality, from EBV-BMLF1₃₀₀ to EBV-BRLF1₁₀₉ (22). EBV-BRLF1₁₀₉-specific T-cells recovered from EBV-immune individuals and expanded *in vitro* in the presence of the cognate peptide present higher affinity/avidity in TCR/pMHC interaction. Note that these cells are not cross-reactive with EBV-BMLF1₃₀₀. On the other hand, EBV-BMLF1₃₀₀-specific T-cells expanded *in vitro* in the presence of the cognate peptide might also recognize EBV-BRLF1₁₀₉ (22). Further expansion of this population with the heterologous peptide (i.e., EBV-BRLF1₁₀₉) produces (cross-reactive) EBV-BRLF1₁₀₉-specific T-cells with lower affinity/avidity in TCR/pMHC interaction (data not shown).

It is known that TCRs usually interact with pMHCs using a “canonical” binding mode (25, 51, 65, 66), but it was shown that a given TCR can preferentially use distinct amino acid residues to come in contact with different complexes (67) or even modify its CDR loops to accommodate different peptides (68). Therefore, it can be argued that immunization with a spicy peptide (such as IAV-PA₂₂₄ or EBV-BRLF1₁₀₉) will trigger a highly polyclonal T-cell response, with a broad spectrum of TCR specificities. Some of these are less specific to the homologous target, and more cross-reactive with other peptides, probably by establishing an interaction “focused” on surface regions that are shared among these targets (**Figure 4**). On the other hand, some of these cells present higher affinity/avidity with this homologous peptide, by establishing an

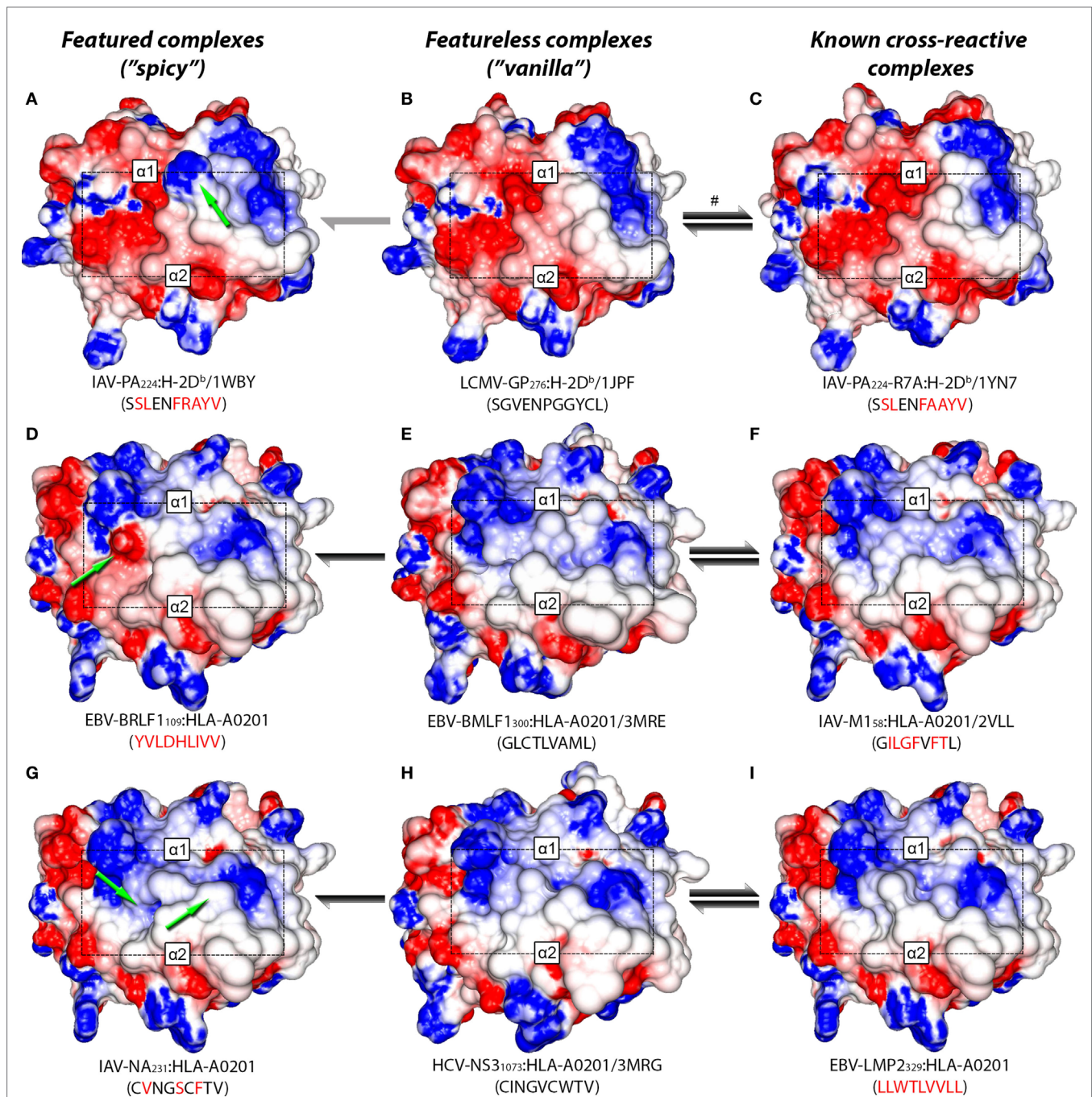
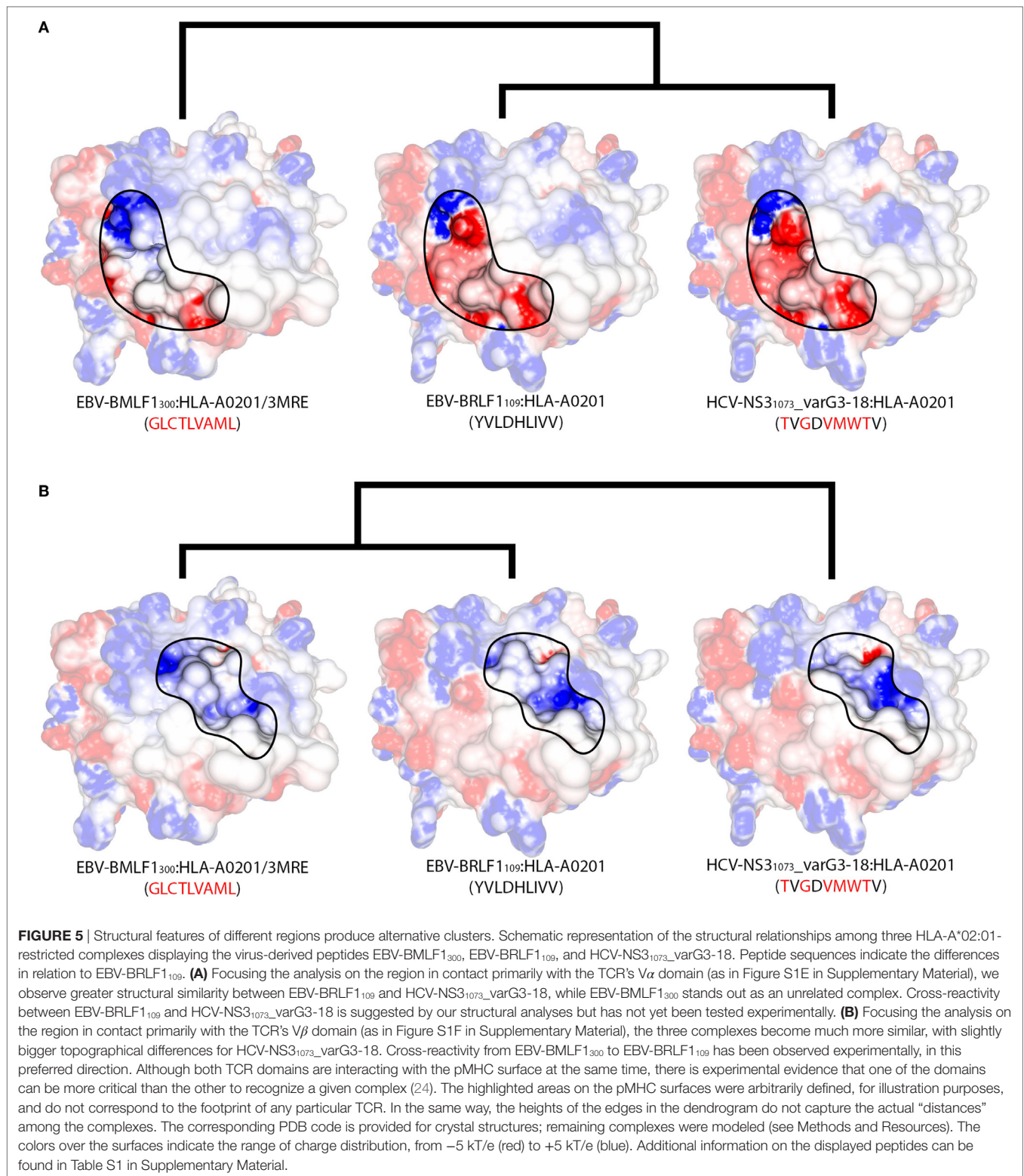


FIGURE 4 | Structural similarity between cross-reactive complexes. TCR-interacting surfaces of selected pMHC complexes were computed with Grasp2 (64). The rows correspond to different datasets of cross-reactive complexes. The central column indicates the reference (cognate) complex in each row (**B,E,H**). The left column indicates a known complex with limited cross-reactivity (**A,D,G**), while the right column indicates a highly cross-reactive complex (**C,F,I**). MHC heavy chain domains $\alpha 1$ and $\alpha 2$ are indicated in each complex, as well as the region corresponding to the peptide (black rectangle). Colors indicate the range of the electrostatic potential over the surface, from -5 kT/e (red) to $+5$ kT/e (blue). Complex information and peptide sequence are depicted below each pMHC. For crystal structures, the corresponding PDB ID is also provided. Complexes with no published crystal structure were modeled (see Methods and Resources). Peptide sequences in each line indicate mutations in relation to the corresponding reference peptide (central column). Green arrows highlight "spicy" features of peptides with "limited" cross-reactivity (left column). Black and gray arrows indicate the intensity and preferred directionality of cross-reactivities observed *in vitro*. The symbol # indicates a cross-reactivity that is suggested by our structural analyses, but that was not yet tested experimentally.

interaction "focused" on unique features of its surface (Figure 5). In turn, cross-reactivity between a spicy and a vanilla peptide depends on which T-cell populations are being tested.

We here hypothesize that, despite different TCRs can share a similar TCR footprint or even interact with the same pMHC amino acid residues, each TCR has a specific "interaction profile."



That is, some TCR/pMHC interactions are more important than others for triggering the T-cell response, and this interaction profile is specific to each TCR (Figures S1E,F in Supplementary Material). Knowing the specific hot-spots of a cognate pMHC,

i.e., the aforementioned “focus” of the TCR, would be key to predict cross-reactivity against heterologous pMHC targets. Moreover, although we tend to think of these hot-spots as pMHC amino acid residues, we need to expand this concept to

account for more subtle features of the TCR/pMHC interaction (e.g., pockets, hydrogen bonds, van der Waals contacts, and coordination of water molecules) (24).

1.6. T-Cell Cross-reactivity Can Be Triggered by High-Affinity Interactions with Specific Structural Features of the pMHC Complex

We previously suggested that T-cells expanded in response to a vanilla peptide should be intrinsically more cross-reactive, since they are focused on patterns shared across different pMHC complexes. Conversely, T-cells expanded in response to a spicy peptide are expected to be less cross-reactive in general, since most heterologous peptides would lack the spicy feature that is the focus of the response. However, these cells should still be cross-reactive with peptides having the spicy feature, in some cases regardless of other evident differences.

In fact, studies in cancer immunotherapy show that mutations leading to increased affinity of a given TCR-peptide interaction can actually increase cross-reactivity (38–40, 69). We hypothesize that although not changing the overall TCR footprint, such mutations can change the interaction profile of the TCR. In other words, the enhanced peptide-specific interaction becomes much more important for T-cell activation than the additional pMHC interactions, and any heterologous pMHC sharing the structural feature recognized by this enhanced TCR can become a cross-reactive target.

Further evidence for this hypothesis comes from a recent publication by Adams et al. (29). Using a carefully designed experimental approach, the authors investigated cross-reactive peptides showing limited sequence identity with the reference cognate peptide (restricted to H-2K^d). Despite apparent sequence diversity among peptides recognized by the probe TCR, closer analysis revealed a repeated focus on structurally and chemically similar elements of the peptides. For instance, the authors describe a preferred interaction with hydrophobic amino acid residues at P7; particularly phenylalanine. The authors refer to this amino acid residue as a peptide hot-spot for cross-reactivity, which in combination with some germline-mediated interactions greatly constrains the actual pool of potential cross-reactive pMHC targets (for the probe TCR). They also relate this description of the TCR/pMHC interaction with a more general feature of protein-protein interactions: a few energetically important contacts (usually in the center), surrounded by weaker and more diverse peripheral interactions. In the context of our discussion, we could see the phenylalanine at P7 as a spicy feature of the cognate peptide and the most important contact in the interaction profile of the probe TCR.

As mentioned earlier, we have previously described cross-reactivity between peptides with no sequence similarity, but with remarkably similar TCR-interacting surfaces (Figures 4G–I). The results described by Song et al. (24) provide an interesting example in which even greater variability can be anticipated. If the main feature for TCR recognition is a pocket defined by the peptide in the MHC cleft (e.g., IAV-M1₅₈:HLA-A*0201), we can expect that such “pocket-specific” T-cells will be cross-reactive

to other pMHC complexes having a similar pocket, maybe regardless of other differences in the TCR-interacting surface. For instance, it is possible for a completely unrelated pMHC (with a different peptide sequence and/or MHC allotype) to have a very similar pocket and, therefore, be a cross-reactive target for IAV-M1₅₈-specific T-cells.

As also discussed by Adams et al. (29), the implications of such “hot-spots” for cross-reactivity prediction are clear. A superficial look at the sequence diversity of cross-reactive peptides might suggest a completely promiscuous recognition, even considering a single TCR. The picture becomes even more complex if on top of that we start considering different pools of T-cells or the *in vivo* response of different individuals, which adds variability given to private specificity and immunological history. This complex picture helps understand the challenge of comparing results from different studies and drawing general conclusions about T-cell cross-reactivity. On the other hand, the characterization of cross-reactivity hot-spots and TCR-specific interaction profiles should allow us to focus our research and make progress for meaningful cross-reactivity predictions.

In fact, Arber et al. (56) published a study that goes in this very direction. They combined T-cell assays and computational analysis to evaluate T-cell cross-reactivity of different clones in the context of cancer immunotherapy. Based on IFN- γ production against a panel of alanine-exchanged variants of the cognate peptide, they defined T-cell-specific sequence motifs. These motifs were meant to capture T-cell-specific cross-reactivity hot-spots; they were later used for a sequence-based screening of potential cross-reactive targets in the human proteome. A number of positive hits were selected and tested experimentally, confirming that one T-cell line was much safer (i.e., less cross-reactive) than the other. The scope of this screening was still limited, not accounting for structural information of the pMHC or other potentially relevant features (14, 24, 70). Nevertheless, it provides us with an example of the type of framework that would be required for T-cell-specific prediction of potential cross-reactive targets.

1.7. Conclusions and Implications for Cancer Immunotherapy

Several immunotherapy trials are currently underway in a number of different tumor types to target tumor-associated peptides (71), including the melanoma-associated antigens MAGE-A3 and MART-1. These tumor antigens are expressed by multiple tumor types (39) but are not expressed by most normal tissues. Since MART-1 is highly expressed in both melanoma and normal melanocytes, MART-1 TCR-based therapies have led to antitumor responses concurrent with vitiligo and melanocyte destruction in the eye and inner ear, side effects that could be relieved with steroid administration (72). However, more severe safety issues with other TCR-based therapies have raised major concerns about this approach (33, 73, 74). As mentioned earlier, fatal adverse events were reported following adoptive transfer of TCR-transduced T-cells targeting complexes displaying the MAGE-A3 peptide (39–42). In two of these patients, unexpectedly severe cardiac toxicity was attributed to recognition of a completely unrelated peptide. This heterologous peptide-target was derived from the

self protein Titin and displayed by HLA-A*01:01 at the surface of healthy cardiac cells (43). As discussed by Stone et al. (38), T-cell cross-reactivity becomes specially relevant in the context of affinity-enhanced TCRs. Approaches like this are becoming more popular through the use of chimeric antigen receptors (CARs) (71). However, as reported by van den Berg et al. (41), severe off-target reactions can occur even without TCR-affinity enhancement. And this adds a layer of concern on top of toxicity and autoimmunity that might occur even with the use of autologous tumor infiltrating T-cells (72, 75). Moreover, as highlighted in our review, T-cell cross-reactivity seems to be rather the rule than the exception. Therefore, despite all mechanisms of central and peripheral tolerance (76), off-target toxicity mediated by T-cell cross-reactivity must be a concern in any TCR-based immunotherapy. However, the risk for off-target toxicity will differ depending on which specific form of therapy is being used.

Our study corroborates the idea that structural similarity among pMHC complexes is one of the main features driving the likelihood of cross-reactive T-cell responses. Cross-reactivity is very likely to be observed between two structurally identical complexes, for most T-cell lines recognizing one of the complexes, and in both directions. On the other hand, finding a T-cell line capable of recognizing two completely different pMHC complexes is highly unlikely. However, in most cases, two complexes will have common features but also different ones. In this situation, cross-reactivity can only be assessed by the level of pMHC structural similarity, as an intrinsic likelihood. However, its occurrence, intensity and directionality will be driven by the specific T-cell population stimulated by the first target and selectively expanded after heterologous challenges.

In the context of polyclonal T-cell populations, this outcome is mostly a consequence of private specificity and immunological history (19, 20, 32). Therefore, predicting patient cross-reactivity in response to immunization, infection or tissue transplantation is very challenging. Even knowing the peptide-targets and the MHC alleles of the patient, and having the perfect tools to estimate intrinsic cross-reactivity probabilities, we would still lack information on the available T-cell repertoire and the interaction profile of the dominating T-cell line. On the other hand, some problems in cancer immunotherapy offer a much more constrained scenario. In the context of TCR-based immunotherapies, researchers know which TCR is being used to recognize the tumor-derived peptide-target and can ensure that this will be the dominating population during treatment. By narrowing our analysis to a particular therapeutic T-cell line, we can limit the scope of cross-reactivity to structural features of the targeted pMHC; more specifically, to hot-spots that are the focus of the therapeutic TCR.

Therefore, we advocate that an important goal of structural analyses in the field of immunotherapy should be the characterization of the TCR-specific recognition profile. This profile should be a refinement of a more general TCR footprint, highlighting which pMHC structural features are more important for triggering this particular T-cell. In turn, this information can be used to guide large-scale *in silico* screenings, based on a combination of structural and sequential information. Currently, no tool can perform such screenings in a personalized fashion, especially

when considering the diversity of MHC alleles in the human population (5). However, T-cell cross-reactivity prediction will soon be enabled by advances in both pMHC structural modeling and TCR sequence analyses.

On the pMHC side, the combination of new modeling methods (57, 77) and structural clustering approaches (48, 78, 79) will allow considering structural information for larger datasets, regardless of whether experimental data are available. On the TCR side, recent reports have shown exciting results in the identification of conserved CDR motifs that can be directly linked to TCR specificity (11, 80). In time, we should be able to define T-cell-specific interaction profiles based on the sequence of the CDR regions of the TCR of interest.

Finally, better understanding of all subtle structural features relevant to TCR/pMHC engagement (11, 24) and their contributions to TCR binding affinity (37, 81–83) will also facilitate efforts toward TCR engineering and rational design (37, 84–86). The TCR-specific interaction profile can inform computer-aided efforts to increase TCR affinity to tumor-specific peptides, while reducing the risk for off-target toxicity. Hopefully, the combined use of these new technologies will soon allow researchers to predict and validate potentially dangerous cross-reactivities in the early stages of therapy development, guiding additional procedures to achieve safer TCR-based immunotherapies. Despite the overall complexity of the subject, urgent needs in cancer immunotherapy are pushing the discussion forward and should pave the way for many additional contributions to other areas of human health.

2. METHODS AND RESOURCES

2.1. Experimental Data on Cross-reactivity Networks

Cross-reactivity networks depicted in **Figure 1** were compiled from previously published experiments. Most data were made available by Cornberg et al. (22), who first presented these networks. The authors also described an escape variant of LCMV-NP₂₀₅ with a V3A substitution (21), suggested its sequence similarity with peptides from old world arenaviruses (MOPV-NP₂₀₅ and LASV-LNP₂₀₉) and finally solved its 3D crystal structure in the context of H-2K^b (62). This study with murine cross-reactivities was further explored by Shen et al. (28). The murine H-2D^b-restricted network was depicted with data from Wlodarczyk et al. (17).

Cornberg et al. (22) also described a human HLA-A*02:01-restricted network. We expanded this network by including a cross-reactive target prospected through structural *in silico* analysis (46) and already confirmed experimentally (47), as well as two non-cross-reactive targets described by Fytli et al. (60). These tested non-cross-reactive targets were included both in human and murine cross-reactivity networks to provide further experimental information to guide our structure-based analysis.

A careful verification of peptides' information was performed to determine the correct protein name and peptide position, providing an updated reference for future studies (Table S1 in Supplementary Material). Curated information from Uniprot (87) was used as the main reference, and GenBank (88) was also

consulted. References to the Immune Epitope Database (IEDB) (89), the Protein Data Bank (PDB) (90), and the CrossTope Database (91) were also provided, when available.

2.2. Crystal Structures

Crystal structures were obtained from the Protein Data Bank (PDB) (90) and revised as needed using the PyMOL Viewer (92). The resulting pMHC structure was submitted to a short energy minimization with the Gromacs 4.5.1 package (93).

Note that 3TID is referred to as the crystal structure of LCMV-GP₃₄:H-2K^b complex, despite presenting an amino acid exchange at P8 (LCMV-GP₃₄-C8M). According to the authors who described the structure (28), this exchange has no significant impact on TCR/pMHC interactions and this C8M variant was used in previous studies as an “equivalent” to the wild-type sequence. Here, sequence divergence between LCMV-GP₃₄ and LCMV-GP₃₄-C8M is indicated in **Figure 3**, but 3TID was considered as the crystal structure of LCMV-GP₃₄ for all structure-based analyses.

2.3. Modeled Structures

Peptide–MHC complexes without published crystal structures were predicted using the DockTope webserver (57). Briefly, a reference crystal structure of the MHC allotype of interest (without its ligand) was used as a receptor (“MHC_donor”) for a molecular docking with Autodock Vina 1.1.2 (94). The input ligand structure was produced by mutating a peptide structure obtained in the context of the same MHC allotype (“Peptide_pattern”). The resulting pMHC structure was then refined through a full atom energy minimization step with the Gromacs 4.5.1 package (93). A new docking search was performed with only the peptide side chains being flexible. This automated approach for pMHC structure prediction was largely validated against available crystal structures (57).

2.4. Electrostatic Potential Calculation and Image Analysis

Electrostatic potential over the TCR-interacting surface of pMHCs (for both crystals and models) was calculated using Delphi (95), through the molecular viewer software GRASP2 (64). Automated scripts were used to prepare the structures for this analysis, allowing all pMHCs to be observed in the same fixed orientation. Images of the TCR-interacting surfaces were saved and imported to the ImageJ 1.46r software (National Institute of Health, USA, <https://imagej.nih.gov/ij/>). Using preexisting classes from ImageJ, our team adapted a plugin to import RGB values from predetermined regions over the pMHC surface (as in Figure S1D in Supplementary Material), following a previously described protocol (46, 47). Values were exported as “csv” tables and used as input for hierarchical cluster analysis.

2.5. Hierarchical Cluster Analysis

In this study, hierarchical clustering was used as a tool to assess structure-based similarity among pMHC complexes. Input values were extracted from the images of the TCR-interacting surfaces (see section 2.4). Hierarchical clustering was performed

with pvclust (61), an R package for assessing the uncertainty in hierarchical clustering. The “average” linkage method was used with “correlation” distance, and the number of bootstrap replications was set to 10,000. Results were plotted as dendrograms with approximately unbiased (AU) and bootstrap probabilities (BP) p-values. BP values are calculated by normal bootstrap resampling, and AU values are computed through multiscale bootstrap resampling, which is considered a better approximation to unbiased p-value (61). SEs for AU p-values were obtained with seplot, presenting values lower than 0.01 for all clusterings performed.

AUTHOR CONTRIBUTIONS

DA, GV, MC, and LS suggested the initial idea behind this work. DA, MR, MS, and GV conceived the experiments. DA selected the dataset and MS curated the information on selected peptides. DA, MR, and MM conducted the modeling and clustering experiments. MF adapted the ImageJ plugin and helped with the extraction of the values for clustering. LK revised clustering experiments and algorithmic choices. DA, MR, GV, MS, MC, and LS analyzed and interpreted the results. GL contributed with the applications to immunotherapy and the review of related literature. DA wrote the manuscript. All the authors reviewed and approved the final manuscript.

ACKNOWLEDGMENTS

The authors thank the *Centro Nacional de Supercomputação* (CESUP/UFRGS) for allowing access to its computational resources. The authors also thank Inyoung Song and Dr. Lawrence Stern, from the University of Massachusetts Medical School (Worcester, MA, USA), for sharing the crystal structure of EBV-BRLF1₁₀₉:HLA-A*0201 before publication. Finally, the authors thank Dr. Didier Devaurs for his helpful comments on the final manuscript.

FUNDING

This work has been supported in part by *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq/Brazil) and *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES/Brazil). This work was also partially supported by the Cancer Prevention & Research Institute of Texas (CPRIT), under award number RP170508.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at <http://journal.frontiersin.org/article/10.3389/fimmu.2017.01210/full#supplementary-material>.

TABLE S1 | List of all studied complexes. Identification and source information for each peptide and MHC, as well as corresponding access codes to relevant databases.

FIGURE S1 | The TCR-interacting surface and the proposed TCR interaction profiles. **(A)** Top view of a pMHC complex depicting the MHC-receptor as *cartoon* (gray) and the peptide–ligand as *sticks* (pink). **(B)** Top view of the same

pMHC complex, depicting the exposed surface of the MHC (gray) and the exposed surface of the peptide (pink). **(C)** The combined surface of the pMHC complex, with the colors indicating the range of charge distribution over the surface from -5 kT/e (red) to $+5$ kT/e (blue). This is the “face” of the pMHC exposed for TCR recognition, referred to as the *TCR-interacting surface*. **(D)** The TCR binds to the pMHC in a conserved orientation: the TCR's variant domain $V\alpha$ primarily interacts with the N-terminal portion of the peptide, while the $V\beta$ domain primarily interacts with the C-terminal portion of the peptide. This area of TCR/pMHC interaction, in a particular docking angle, is referred to as the *TCR footprint*. **(E)** Schematic representation of a *TCR-specific interaction profile* over the pMHC surface. Colored boxes indicate “hot-spots” for cross-reactivity (green) and secondary contacts that also contribute to TCR binding affinity (yellow). **(F)** Schematic representation of a different interaction profile, displayed by a different TCR that still shares the same general TCR footprint. Both depicted profiles are simplified schematic representations and do not represent known interactions of a any particular TCR.

FIGURE S2 | Crystal-based H-2K^b-restricted clustering. Structure-based hierarchical clustering performed with pvclust (61). Each putative cluster is represented by a specific edge (gray numbers), in order of increasing heights (y axis). Cluster confidence is measured with two p-values, approximately unbiased (AU) and bootstrap probabilities (BP). Lines highlighted in purple indicate structures with greater structural similarity (as represented in **Figure 1**). Peptide abbreviation and corresponding PDB code for each crystal structure (in blue) are provided. *Crystal structure 3TID was used to represent LCMV-GP₃₄, despite presenting a C8M exchange (see Methods and Resources).

FIGURE S3 | TCR-interacting surfaces of predicted cross-reactive targets. Regions with positive (blue) and negative (red) charges are represented with a scale from -5 to $+5$ kT/e. Information on the corresponding peptide and MHC restriction is provided below each complex. Amino acid exchanges in relation to LCMV-GP₁₁₈ are indicated. Great structural similarity is observed between these

two complexes, both in terms of topography and electrostatic potential over the TCR-interacting surface. Note that other subtle structural differences might exist but are not well captured by this representation of the complexes.

FIGURE S4 | Specific interaction with a prominent peptide amino acid.

(A) Surface of the IAV-PA₂₂₄:H-2D^b complex (spicy peptide) according to a crystal structure obtained in the absence of the TCR (PDB code 1WBY). **(B)** Surface of the same complex according to a crystal structure obtained in the presence of the TCR (PDB code 3PQY). **(C)** Cartoon depiction of the 3PQY structure highlighting TCR amino acid residues that interact directly with a prominent arginine at the peptide (R7), forming a negatively charged cavity. Side chain of amino acid R7 is depicted in *ball and stick*. TCR and MHC domains are indicated, and green arrows highlight the location of amino acid residue R7. Electrostatic potentials were computed with Grasp2 (64).

FIGURE S5 | Crystal-based H-2D^b-restricted clustering. Structure-based hierarchical clustering performed with pvclust. Each putative cluster is represented by a specific edge (gray numbers), in order of increasing heights (y axis). Cluster confidence is measured with two p-values, approximately unbiased (AU) and bootstrap probabilities (BP). Peptide abbreviation and the respective PDB code for each crystal structure (in blue) are provided. Lines highlighted in purple indicate structures with greater structural similarity (as represented in **Figure 1**).

FIGURE S6 | Extended HLA-A*02:01-restricted clustering. Structure-based hierarchical clustering performed with pvclust. Each putative cluster is represented by a specific edge (gray numbers), in order of increasing heights (y axis). Cluster confidence is measured with two p-values, approximately unbiased (AU) and bootstrap probabilities (BP). Abbreviation of crystal structures includes their PDB code (in blue), while “Mod” indicates modeled structures. Lines highlighted in purple indicate structures with greater structural similarity (as represented in **Figure 1**).

REFERENCES

- Dzik JM. The ancestry and cumulative evolution of immune reactions. *Acta Biochim Pol* (2010) 57(4):443–66.
- Hirano M. Evolution of vertebrate adaptive immunity: immune cells and tissues, and AID/APOBEC cytidine deaminases. *Bioessays* (2015) 37(8):877–87. doi:10.1002/bies.201400178
- Welsh RM, Selin LK, Szomolanyi-Tsuda E. Immunological memory to viral infections. *Annu Rev Immunol* (2004) 22:711–43. doi:10.1146/annurev.immunol.22.012703.104527
- Lauring AS, Frydman J, Andino R. The role of mutational robustness in RNA virus evolution. *Nat Rev Microbiol* (2013) 11(5):327–36. doi:10.1038/nrmicro3003
- Vandiedonck C, Knight JC. The human Major Histocompatibility Complex as a paradigm in genomics research. *Brief Funct Genomic Proteomic* (2009) 8(5):379–94. doi:10.1093/bfpg/elp010
- Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR, et al. Antagonistic coevolution accelerates molecular evolution. *Nature* (2010) 464(7286):275–8. doi:10.1038/nature08798
- Kubinak JL, Ruff JS, Hyzer CW, Slev PR, Potts WK. Experimental viral evolution to specific host MHC genotypes reveals fitness and virulence trade-offs in alternative MHC types. *Proc Natl Acad Sci U S A* (2012) 109(9):3422–7. doi:10.1073/pnas.1112633109
- Sohn SJ, Thompson J, Winoto A. Apoptosis during negative selection of autoreactive thymocytes. *Curr Opin Immunol* (2007) 19(5):510–5. doi:10.1016/j.coi.2007.06.001
- Elhanati Y, Murugan A, Callan CG, Mora T, Walczak AM. Quantifying selection in immune receptor repertoires. *Proc Natl Acad Sci U S A* (2014) 111(27):9875–80. doi:10.1073/pnas.1409572111
- Zarnitsyna VI, Evavold BD, Schoettle LN, Blattman JN, Antia R. Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire. *Front Immunol* (2013) 4:485. doi:10.3389/fimmu.2013.00485
- Dash P, Fiore-Gartland AJ, Hertz T, Wang GC, Sharma S, Souquette A, et al. Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature* (2017) 547(7661):89–93. doi:10.1038/nature22383
- Welsh RM, Selin LK. No one is naive: the significance of heterologous T-cell immunity. *Nat Rev Immunol* (2002) 2(6):417–26. doi:10.1038/nri820
- Welsh RM, Che JW, Brehm MA, Selin LK. Heterologous immunity between viruses. *Immunol Rev* (2010) 235(1):244–66. doi:10.1111/j.0105-2896.2010.00897.x
- Vieira GF, Chies JA. Immunodominant viral peptides as determinants of cross-reactivity in the immune system – can we develop wide spectrum viral vaccines? *Med Hypotheses* (2005) 65(5):873–9. doi:10.1016/j.mehy.2005.05.041
- Welsh RM, Fujinami RS. Pathogenic epitopes, heterologous immunity and vaccine design. *Nat Rev Microbiol* (2007) 5(7):555–63. doi:10.1038/nrmicro1709
- Selin LK, Cornberg M, Brehm MA, Kim SK, Calcagno C, Ghersi D, et al. CD8 memory T cells: cross-reactivity and heterologous immunity. *Semin Immunol* (2004) 16(5):335–47. doi:10.1016/j.smim.2004.08.014
- Włodarczyk MF, Kraft AR, Chen HD, Kenney LL, Selin LK. Anti-IFN- γ and peptide-tolerization therapies inhibit acute lung injury induced by cross-reactive Influenza A-specific memory T cells. *J Immunol* (2013) 190(6):2736–46. doi:10.4049/jimmunol.1201936
- Cornberg M, Kenney LL, Chen AT, Waggoner SN, Kim SK, Dienes HP, et al. Clonal exhaustion as a mechanism to protect against severe immunopathology and death from an overwhelming CD8 T cell response. *Front Immunol* (2013) 4:475. doi:10.3389/fimmu.2013.00475
- Cornberg M, Wedemeyer H. Hepatitis C virus infection from the perspective of heterologous immunity. *Curr Opin Virol* (2016) 16:41–8. doi:10.1016/j.coviro.2016.01.005
- Gil A, Yassai MB, Naumov YN, Selin LK. Narrowing of human Influenza A virus-specific T cell receptor α and β repertoires with increasing age. *J Virol* (2015) 89(8):4102–16. doi:10.1128/JVI.03020-14
- Cornberg M, Chen AT, Wilkinson LA, Brehm MA, Kim SK, Calcagno C, et al. Narrowed TCR repertoire and viral escape as a consequence of heterologous immunity. *J Clin Invest* (2006) 116(5):1443–56. doi:10.1172/JCI27804
- Cornberg M, Clute SC, Watkin LB, Saccoccio FM, Kim SK, Naumov YN, et al. CD8 T cell cross-reactivity networks mediate heterologous immunity in human EBV and murine vaccinia virus infections. *J Immunol* (2010) 184(6):2825–38. doi:10.4049/jimmunol.0902168

23. Gil A, Kenney LL, Mishra R, Watkin LB, Aslan N, Selin LK. Vaccination and heterologous immunity: educating the immune system. *Trans R Soc Trop Med Hyg* (2015) 109(1):62–9. doi:10.1093/trstmh/tru198
24. Song I, Gil A, Mishra R, Ghersi D, Selin LK, Stern LJ. Broad TCR repertoire and diverse structural solutions for recognition of an immunodominant CD8(+) T cell epitope. *Nat Struct Mol Biol* (2017) 24(4):395–406. doi:10.1038/nsmb.3383
25. Birnbaum ME, Mendoza JL, Sethi DK, Dong S, Glanville J, Dobbins J, et al. Deconstructing the peptide-MHC specificity of T cell recognition. *Cell* (2014) 157(5):1073–87. doi:10.1016/j.cell.2014.03.047
26. Degauque N, Brouard S, Soullillou JP. Cross-reactivity of TCR repertoire: current concepts, challenges, and implication for allotransplantation. *Front Immunol* (2016) 7:89. doi:10.3389/fimmu.2016.00089
27. Yin Y, Li Y, Mariuzza RA. Structural basis for self-recognition by autoimmune T-cell receptors. *Immunol Rev* (2012) 250(1):32–48. doi:10.1111/imr.12002
28. Shen ZT, Nguyen TT, Daniels KA, Welsh RM, Stern LJ. Disparate epitopes mediating protective heterologous immunity to unrelated viruses share peptide-MHC structural features recognized by cross-reactive T cells. *J Immunol* (2013) 191(10):5139–52. doi:10.4049/jimmunol.1300852
29. Adams JJ, Narayanan S, Birnbaum ME, Sidhu SS, Blevins SJ, Gee MH, et al. Structural interplay between germline interactions and adaptive recognition determines the bandwidth of TCR-peptide-MHC cross-reactivity. *Nat Immunol* (2016) 17(1):87–94. doi:10.1038/ni.3310
30. Wedemeyer H, Mizukoshi E, Davis AR, Bennink JR, Rehmann B. Cross-reactivity between hepatitis C virus and Influenza A virus determinant-specific cytotoxic T cells. *J Virol* (2001) 75(23):11392–400. doi:10.1128/JVI.75.23.11392-11400.2001
31. Kasprócz V, Ward SM, Turner A, Grammatikos A, Nolan BE, Lewis-Ximenez L, et al. Defining the directionality and quality of influenza virus-specific CD8+ T cell cross-reactivity in individuals infected with hepatitis C virus. *J Clin Invest* (2008) 118(3):1143–53. doi:10.1172/JCI33082
32. Che JW, Selin LK, Welsh RM. Evaluation of non-reciprocal heterologous immunity between unrelated viruses. *Virology* (2015) 482:89–97. doi:10.1016/j.virol.2015.03.002
33. Lizée G, Overwijk WW, Radvanyi L, Gao J, Sharma P, Hwu P. Harnessing the power of the immune system to target cancer. *Annu Rev Med* (2013) 64:71–90. doi:10.1146/annurev-med-112311-083918
34. Yee C, Lizée G, Schueneman AJ. Endogenous T-cell therapy: clinical experience. *Cancer J* (2015) 21(6):492–500. doi:10.1097/PPO.0000000000000158
35. Rouce RH, Sharma S, Huynh M, Heslop HE. Recent advances in T-cell immunotherapy for haematological malignancies. *Br J Haematol* (2017) 176(5):688–704. doi:10.1111/bjh.14470
36. Menon S, Shin S, Dy G. Advances in cancer immunotherapy in solid tumors. *Cancers (Basel)* (2016) 8(12):1–21. doi:10.3390/cancers8120106
37. Pierce BG, Hellman LM, Hossain M, Singh NK, Vander Kooi CW, Weng Z, et al. Computational design of the affinity and specificity of a therapeutic T cell receptor. *PLoS Comput Biol* (2014) 10(2):e1003478. doi:10.1371/journal.pcbi.1003478
38. Stone JD, Harris DT, Kranz DM. TCR affinity for pMHC formed by tumor antigens that are self-proteins: impact on efficacy and toxicity. *Curr Opin Immunol* (2015) 33:16–22. doi:10.1016/j.coi.2015.01.003
39. Linette GP, Stadtmayer EA, Maus MV, Rapoport AP, Levine BL, Emery L, et al. Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma. *Blood* (2013) 122(6):863–71. doi:10.1182/blood-2013-03-490565
40. Morgan RA, Chinnasamy N, Abate-Daga D, Gros A, Robbins PF, Zheng Z, et al. Cancer regression and neurological toxicity following anti-MAGE-A3 TCR gene therapy. *J Immunother* (2013) 36(2):133–51. doi:10.1097/CJI.0b013e3182829903
41. van den Berg JH, Gomez-Eerland R, van de Wiel B, Hulshoff L, van den Broek D, Bins A, et al. Case report of a fatal serious adverse event upon administration of T cells transduced with a MART-1-specific T-cell receptor. *Mol Ther* (2015) 23(9):1541–50. doi:10.1038/mt.2015.60
42. Cameron BJ, Gerry AB, Dukes J, Harper JV, Kannan V, Bianchi FC, et al. Identification of a Titin-derived HLA-A1-presented peptide as a cross-reactive target for engineered MAGE A3-directed T cells. *Sci Transl Med* (2013) 5(197):197ra103. doi:10.1126/scitranslmed.3006034
43. Raman MC, Rizkallah PJ, Simmons R, Donnellan Z, Dukes J, Bossi G, et al. Direct molecular mimicry enables off-target cardiovascular toxicity by an enhanced affinity TCR designed for cancer immunotherapy. *Sci Rep* (2016) 6:18851. doi:10.1038/srep18851
44. Haase K, Raffegerst S, Schendel DJ, Frishman D. Expitope: a web server for epitope expression. *Bioinformatics* (2015) 31(11):1854–6. doi:10.1093/bioinformatics/btv068
45. Jaravine V, Raffegerst S, Schendel DJ, Frishman D. Assessment of cancer and virus antigens for cross-reactivity in human tissues. *Bioinformatics* (2017) 33(1):104–11. doi:10.1093/bioinformatics/btw567
46. Antunes DA, Rigo MM, Silva JP, Cibulski SP, Sinigaglia M, Chies JA, et al. Structural in silico analysis of cross-genotype-reactivity among naturally occurring HCV NS3-1073-variants in the context of HLA-A*02:01 allele. *Mol Immunol* (2011) 48(12–13):1461–7. doi:10.1016/j.molimm.2011.03.019
47. Zhang S, Bakshi RK, Suneetha PV, Fytily P, Antunes DA, Vieira GF, et al. Frequency, private specificity, and cross-reactivity of preexisting hepatitis C virus (HCV)-specific CD8+ T cells in HCV-seronegative individuals: implications for vaccine responses. *J Virol* (2015) 89(16):8304–17. doi:10.1128/JVI.00539-15
48. Mendes MF, Antunes DA, Rigo MM, Sinigaglia M, Vieira GF. Improved structural method for T-cell cross-reactivity prediction. *Mol Immunol* (2015) 67(2 Pt B):303–10. doi:10.1016/j.molimm.2015.06.017
49. Dhanik A, Kirshner JR, MacDonald D, Thurston G, Lin HC, Murphy AJ, et al. In-silico discovery of cancer-specific peptide-HLA complexes for targeted therapy. *BMC Bioinformatics* (2016) 17:286. doi:10.1186/s12859-016-1150-2
50. Moise L, Gutierrez AH, Bailey-Kellogg C, Terry F, Leng Q, Abdel Hady KM, et al. The two-faced T cell epitope: examining the host-microbe interface with JanusMatrix. *Hum Vaccin Immunother* (2013) 9(7):1577–86. doi:10.4161/hv.24615
51. Adams JJ, Narayanan S, Liu B, Birnbaum ME, Kruse AC, Bowerman NA, et al. T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity* (2011) 35(5):681–93. doi:10.1016/j.immuni.2011.09.013
52. Antunes DA, Vieira GF, Rigo MM, Cibulski SP, Sinigaglia M, Chies JA. Structural allele-specific patterns adopted by epitopes in the MHC-I cleft and reconstruction of MHC:peptide complexes to cross-reactivity assessment. *PLoS One* (2010) 5(4):e10353. doi:10.1371/journal.pone.0010353
53. D'Orsogna LJ, Roelen DL, Doxiadis II, Claas FH. TCR cross-reactivity and allorecognition: new insights into the immunogenetics of allorecognition. *Immunogenetics* (2012) 64(2):77–85. doi:10.1007/s00251-011-0590-0
54. van den Heuvel H, Heutinck KM, van der Meer-Prins EMW, Yong SL, van Miert PPMC, Anholts JDH, et al. Allo-HLA cross-reactivities of Cytomegalovirus-, Influenza-, and Varicella Zoster virus-specific memory T cells are shared by different healthy individuals. *Am J Transplant* (2017) 17(8):2033–44. doi:10.1111/ajt.14279
55. Ayres CM, Scott DR, Corcelli SA, Baker BM. Differential utilization of binding loop flexibility in T cell receptor ligand selection and cross-reactivity. *Sci Rep* (2016) 6:25070. doi:10.1038/srep25070
56. Arber C, Feng X, Abhyankar H, Romero E, Wu MF, Heslop HE, et al. Survivin-specific T cell receptor targets tumor but not T cells. *J Clin Invest* (2015) 125(1):157–68. doi:10.1172/JCI75876
57. Rigo MM, Antunes DA, Vaz de Freitas M, Fabiano de Almeida Mendes M, Meira L, Sinigaglia M, et al. DockTope: a web-based tool for automated pMHC-I modelling. *Sci Rep* (2015) 5:18413. doi:10.1038/srep18413
58. Selin LK, Włodarczyk MF, Kraft AR, Nie S, Kenney LL, Puzone R, et al. Heterologous immunity: immunopathology, autoimmunity and protection during viral infections. *Autoimmunity* (2011) 44(4):328–47. doi:10.3109/08916934.2011.523277
59. Petrova GV, Naumova EN, Gorski J. The polyclonal CD8 T cell response to influenza M158-66 generates a fully connected network of cross-reactive clonotypes to structurally related peptides: a paradigm for memory repertoire coverage of novel epitopes or escape mutants. *J Immunol* (2011) 186(11):6390–7. doi:10.4049/jimmunol.1004031
60. Fytily P, Dalekos GN, Schlaphoff V, Suneetha PV, Sarrazin C, Zauner W, et al. Cross-genotype-reactivity of the immunodominant HCV CD8 T-cell epitope NS3-1073. *Vaccine* (2008) 26(31):3818–26. doi:10.1016/j.vaccine.2008.05.045

61. Suzuki R, Shimodaira H. Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* (2006) 22(12):1540–2. doi:10.1093/bioinformatics/btl117
62. Chen AT, Cornberg M, Gras S, Guillonnet C, Rossjohn J, Trees A, et al. Loss of anti-viral immunity by infection with a virus encoding a cross-reactive pathogenic epitope. *PLoS Pathog* (2012) 8(4):e1002633. doi:10.1371/journal.ppat.1002633
63. Turner SJ, Kedzierska K, Komodromou H, La Gruta NL, Dunstone MA, Webb AI, et al. Lack of prominent peptide-major histocompatibility complex features limits repertoire diversity in virus-specific CD8+ T cell populations. *Nat Immunol* (2005) 6(4):382–9. doi:10.1038/nri1175
64. Petrey D, Honig B. GRASP2: visualization, surface properties, and electrostatics of macromolecular structures and sequences. *Methods Enzymol* (2003) 374:492–509. doi:10.1016/S0076-6879(03)74021-X
65. Garcia KC, Adams JJ, Feng D, Ely LK. The molecular basis of TCR germline bias for MHC is surprisingly simple. *Nat Immunol* (2009) 10(2):143–7. doi:10.1038/nri1219
66. Gras S, Burrows SR, Turner SJ, Sewell AK, McCluskey J, Rossjohn J. A structural voyage toward an understanding of the MHC-I-restricted immune response: lessons learned and much to be learned. *Immunol Rev* (2012) 250(1): 61–81. doi:10.1111/j.1600-065X.2012.01159.x
67. Santori FR, Holmberg K, Ostrov D, Gascoigne NR, Vukmanović S. Distinct footprints of TCR engagement with highly homologous ligands. *J Immunol* (2004) 172(12):7466–75. doi:10.4049/jimmunol.172.12.7466
68. Mazza C, Auphan-Anezin N, Gregoire C, Guimezanes A, Kellenberger C, Roussel A, et al. How much can a T-cell antigen receptor adapt to structurally distinct antigenic peptides? *EMBO J* (2007) 26(7):1972–83. doi:10.1038/sj.emboj.7601605
69. Holler PD, Chlewicki LK, Kranz DM. TCRs with high affinity for foreign pMHC show self-reactivity. *Nat Immunol* (2003) 4(1):55–62. doi:10.1038/nr1863
70. Frankild S, de Boer RJ, Lund O, Nielsen M, Kesmir C. Amino acid similarity accounts for T cell cross-reactivity and for “holes” in the T cell repertoire. *PLoS One* (2008) 3(3):e1831. doi:10.1371/journal.pone.0001831
71. Wang RF, Wang HY. Immune targets and neoantigens for cancer immunotherapy and precision medicine. *Cell Res* (2017) 27(1):11–37. doi:10.1038/cr.2016.155
72. Dudley ME, Wunderlich JR, Robbins PF, Yang JC, Hwu P, Schwartzentruber DJ, et al. Cancer regression and autoimmunity in patients after clonal repopulation with antitumor lymphocytes. *Science* (2002) 298(5594):850–4. doi:10.1126/science.1076514
73. Tan MP, Gerry AB, Brewer JE, Melchiori L, Bridgeman JS, Bennett AD, et al. T cell receptor binding affinity governs the functional profile of cancer-specific CD8+ T cells. *Clin Exp Immunol* (2015) 180(2):255–70. doi:10.1111/cei.12570
74. Perica K, Varela JC, Oelke M, Schneck J. Adoptive T cell immunotherapy for cancer. *Rambam Maimonides Med J* (2015) 6(1):e0004. doi:10.5041/RMMJ.10179
75. Weber JS, Yang JC, Atkins MB, Disis ML. Toxicities of immunotherapy for the practitioner. *J Clin Oncol* (2015) 33(18):2092–9. doi:10.1200/JCO.2014.60.0379
76. Xing Y, Hogquist KA. T-cell tolerance: central and peripheral. *Cold Spring Harb Perspect Biol* (2012) 4(6):1–15. doi:10.1101/cshperspect.a006957
77. Khan JM, Ranganathan S. pDOCK: a new technique for rapid and accurate docking of peptide ligands to Major Histocompatibility Complexes. *Immunome Res* (2010) 6:S2. doi:10.1186/1745-7580-6-S1-S2
78. Richter S, Wenzel A, Stein M, Gabdoulline RR, Wade RC. webPIPSA: a web server for the comparison of protein interaction properties. *Nucleic Acids Res* (2008) 36(Web Server issue):W276–80. doi:10.1093/nar/gkn181
79. Bryant DH, Moll M, Chen BY, Fofanov VY, Kavraki LE. Analysis of substructural variation in families of enzymatic proteins with applications to protein function prediction. *BMC Bioinformatics* (2010) 11:242. doi:10.1186/1471-2105-11-242
80. Glanville J, Huang H, Nau A, Hatton O, Wagar LE, Rubelt F, et al. Identifying specificity groups in the T cell receptor repertoire. *Nature* (2017) 547(7661): 94–8. doi:10.1038/nature22976
81. Borrmann T, Cimons J, Cosiano M, Purcaro M, Pierce BG, Baker BM, et al. ATLAS: a database linking binding affinities with structures for wild-type and mutant TCR-pMHC complexes. *Proteins* (2017) 85(5):908–16. doi:10.1002/prot.25260
82. Hoffmann T, Marion A, Antes I. DynaDom: structure-based prediction of T cell receptor inter-domain and T cell receptor-peptide-MHC (class I) association angles. *BMC Struct Biol* (2017) 17(1):2. doi:10.1186/s12900-016-0071-7
83. Riley TP, Singh NK, Pierce BG, Weng Z, Baker BM. Computational modeling of T cell receptor complexes. *Methods Mol Biol* (2016) 1414:319–40. doi:10.1007/978-1-4939-3569-7_19
84. Zoete V, Irving M, Ferber M, Cuendet MA, Michielin O. Structure-based, rational design of T cell receptors. *Front Immunol* (2013) 4:268. doi:10.3389/fimmu.2013.00268
85. Malecek K, Grigoryan A, Zhong S, Gu WJ, Johnson LA, Rosenberg SA, et al. Specific increase in potency via structure-based design of a TCR. *J Immunol* (2014) 193(5):2587–99. doi:10.4049/jimmunol.1302344
86. Sharma P, Kranz DM. Recent advances in T-cell engineering for use in immunotherapy. *F1000Res* (2016) 5:1–12. doi:10.12688/f1000research.9073.1
87. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res* (2017) 45(D1):D158–69. doi:10.1093/nar/gkw1099
88. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res* (2016) 44(D1):67–72. doi:10.1093/nar/gkv1276
89. Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, et al. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res* (2015) 43(Database issue):D405–12. doi:10.1093/nar/gku938
90. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Res* (2000) 28(1):235–42. doi:10.1093/nar/28.1.235
91. Sinigaglia M, Antunes DA, Rigo MM, Chies JA, Vieira GF. CrossTope: a curate repository of 3D structures of immunogenic peptide: MHC complexes. *Database (Oxford)* (2013) 2013:bat002. doi:10.1093/database/bat002
92. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015). Available from: <https://pymol.org/citing>
93. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* (2013) 29(7):845–54. doi:10.1093/bioinformatics/btt055
94. 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. doi:10.1002/jcc.21334
95. Li L, Li C, Sarkar S, Zhang J, Witham S, Zhang Z, et al. DelPhi: a comprehensive suite for DelPhi software and associated resources. *BMC Biophys* (2012) 5:9. doi:10.1186/2046-1682-5-9

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Antunes, Rigo, Freitas, Mendes, Sinigaglia, Lizée, Kavraki, Selin, Cornberg and Vieira. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.