T Cell Epitope Redundancy: Cross-conservation of the TCR face between Pathogens and Self and its Implications for Vaccines and Autoimmunity

Leonard Moise[1,2], Sarah Beseme[1], Ryan Tassone[2], Rui Liu[2], Farzana Kibria[1], Frances Terry[1], William Martin[1], Anne S. De Groot[1,2*]

1 EpiVax, Inc., 146 Clifford St., Providence, RI 02903 USA

2 Institute for Immunology and Informatics, University of Rhode Island, 80 Washington St., Providence, RI 02903 USA

*Corresponding author

Anne S. De Groot, MD, CEO/CSO, EpiVax Inc., 146 Clifford St., Providence, RI 02903; Director, Institute of Immunology and Informatics and Professor (Research) University of Rhode Island, 80 Washington St., Providence, RI 02903. Telephone: 401-272-2123. Fax: 401-272-7562. Email: annied@epivax.com

Summary

T cells are extensively trained on ‘self’ in the thymus and then move to the periphery, where they seek out and destroy infections and regulate immune response to self-antigens. T cell receptors (TCR) on T cells’ surface, recognize T cell epitopes, short linear strings of amino acids presented by antigen-presenting cells. Some of these epitopes activate T effectors, while others activate regulatory T cells. It was recently discovered that T cell epitopes that are highly conserved on their TCR face with human genome sequences are often associated with T cells that regulate immune response. These TCR-cross-conserved or ‘redundant epitopes’ are more common in proteins found in pathogens that have co-evolved with humans, than in other non-commensal pathogens. Epitope redundancy might be the link between pathogens and autoimmune disease. This article reviews recently published data and addresses epitope redundancy, the “elephant in the room” for vaccine developers and T cell immunologists.

Keywords

T cell epitope, Regulatory T cell, Cross-conservation, T cell receptor, TCR, TCR Degeneracy, Molecular mimicry, Autoimmune disease, Immunoinformatics, Guillain Barre Syndrome, Narcolepsy, Multiple Sclerosis, Vaccine, Off Target Effects, HIV, HCV, H7N9, Influenza Vaccine, Cancer Vaccine, Virus, Bacteria, Parasite.

Introduction

Like samurai warriors, T cells undergo extensive training on self-antigens in the thymus before venturing into the periphery to fight infections. Some T cells use their T cell receptors (TCR) to search and counter pathogens, while others find T cell epitopes that seem to be more likely to trigger regulation or suppression of immune response (**Figure 1**). However, by definition, naïve T cells that escape deletion in the thymus to war against pathogens in the periphery have T cell receptors that are trained on 'self' T cell epitopes. Some of the T cells become regulatory T cells (Tregs), while others become effector T cells (Teff).

How T cells develop to perform the function of 'killer' T cells, despite being trained on self-epitopes in the thymus, is one of the greatest mysteries of immunology. Deregulation or abnormal activation of these functions leads to immune intolerance, such as autoimmune diseases or allergy. While vaccines are safe for the large majority of people and generate herd immunity that protects vulnerable persons who cannot be immunized, some vaccines have been associated with the induction of autoimmune responses, as observed in the recent cases of narcolepsy that were associated with H1N1 influenza vaccination [1].

Computer analysis of T cell epitopes is now uncovering patterns of amino acids that may be associated with each of these T cell types. Taking advantage of this system, pathogens have found ways to circumvent immune detection, reducing the likelihood that they will be eliminated by our humoral and cellular immune system. They do this by removing Teff epitopes (immune escape) and, as has been discovered recently, by adopting epitopes that look like human epitopes (immune camouflage). There is much that we can learn from defining and understanding these redundant epitopes - such as T cell epitopes found in pathogens - about ourselves and our microbial enemies.

The development of safer vaccines might be dependent on the identification of which T cell epitopes present in the vaccine – and by extension, the pathogen – triggers this reaction against self. Computer algorithms for predicting and analyzing T cell epitopes have made it possible to discern common patterns that are important for regulating immune responses to self. They may be applied not only to design better vaccines but may also contribute to development of new treatments for autoimmune diseases. This article reviews the concept of epitope redundancy, describes the immunoinformatics tools to identify epitope redundancy, and discusses the relevance of this approach to design better vaccines and understand pathogen-related autoimmune diseases.

TCR specificity in T cell development

As a result of the selection steps in the thymus, conventional T cells and natural regulatory T cells (nTreg) emerge with an intermediate level of self-reactivity, one that exceeds the level leading to clonal deletion and falls below the level for negative selection (**Figure 2**). Little is known about T cell epitope specificity of nTregs. Some studies support a “buddy” hypothesis that TCRs of conventional and nTreg cells share the same specificity, which serves to prevent autoimmunity [2]. Consistent with this concept, the TCR repertoire of thymic-derived effector T (Teff) cells and Tregs has been shown to overlap to some degree [3]. Other TCR repertoire studies have not shown overlapping repertoires [4-6]. Further studies, focusing on the exact specificities of conventional Teff cells versus nTregs, are clearly needed, so as to better understand the implications of overlapping TCR repertoires. Based on current understanding of the maturation of T cell subsets, there appears to be a swords-edge balance between protective immune response to pathogens and regulation of autoimmunity in the periphery.

TCR degeneracy, Epitope Redundancy, Mimicry and Camouflage

T cells are activated when T cell epitopes, short linear strings of amino acids, are presented at surface of APCs in the context of human leukocyte antigen (HLA) molecules. The peptide-HLA complex presents a unique surface that is recognized by the TCR. The topic of T cell epitope cross-reactivity, believed to be due to TCR degeneracy or lack of fine specificity, is a concept that is familiar to infectious disease and autoimmune disease researchers. Conservation between T cell epitopes from different strains of a given pathogen is well known (e.g., influenza [7] and Dengue [8]).

In the realm of autoimmunity, similarity between self and pathogen has also been called molecular mimicry [9]; however this terminology does not allow for differentiation between different types of mimicry - structural (as in cross-reactive B cell epitopes), or mechanism-related (as in Epstein - Barr virus (EBV)'s adoption of IL-10 as a mediator of immune regulation [10]). In this article, we will specifically focus on T cell epitope redundancy, as defined below, which is one of the direct results of the TCR training on self in the thymus [11] and education of these T cells in the periphery [12]. Our focus on cross-reactivity of T cell epitopes is enabled by the development of accurate computational tools for the identification of these epitopes [13,14] and enlivened by reports associating autoimmune diseases to exposure to pathogens, such as the recent report linking narcolepsy to influenza exposure or vaccination [15].

Cross-reactivity occurs because TCRs are somewhat 'promiscuous' in that they can recognize remarkably similar (but not necessarily identical) epitopes that are defined by binding to the same HLA and presentation of similar TCR-facing residues to the TCR. Complex relationships between T cell epitopes have been described by Welsh and Selin [16], in which different epitopes (different in sequence) may be recognized by the same T cell. They have called this cross-conservation between pathogens *heterologous immunity*, as previous exposure to the conserved epitope can pre-define immune response to a novel pathogen. We define 'redundant' epitopes as non-self

linear sequences that bind to the same HLA and may be recognized by the same TCR as a self-epitope despite differences at HLA binding positions. Human commensal pathogens that have established a pattern of ‘hit and stay’ (such as herpes simplex virus) have more redundant sequences than pathogens that ‘hit and run’ (such as Ebola and Marburg) [17,18]. To what degree this redundancy may contribute to autoimmunity, is as yet unknown [19].

Heterologous immunity [20- 24] may also involve TCR recognition of similar HLA-binding epitopes that exhibit variability at TCR-facing positions. However, existing methods for the identification of heterologous epitopes, in which the epitopes may not be similar in sequence but trigger cross-reactive T cell responses, are cumbersome. Redundant epitopes comprise linear sequences that are identifiable using computational methods [13]. These computational methods enable large-scale comparisons to peptides across genomes (the human genome, a cancer genome, the intestinal microbiome, viral pathogens, and parasite genomes, to name just a few). A limitation of the computational approach at this time is the ability to identify cross-reactive TCR specificities, such as those found in heterologous immunity that are defined only by identical TCR-face sequence. Thus, unlike the category of cross-reactive epitopes that may be defined by intensive studies of heterologous immunity [25], T cell epitope redundancy is defined by the ability to examine it using in silico methods. Once the epitopes are defined, methods such as TCR clonotyping [26], high-resolution tetramer staining [27,28], and x-ray crystal structural determination may help to further elucidate the nature of the TCR-cross-reactive immune responses.

We have postulated that epitope redundancy might enable pathogens to camouflage themselves from immune response. Other means of immune escape by pathogens have been well described [29]. For example, Gram-negative bacteria can escape innate immunity through structural modifications of their PAMPs, thus escaping recognition by phagocytes [30]. *Plasmodium falciparum*, the parasite responsible for malaria, can also escape the human immune system by mutually exclusive gene expression of the var gene family, leading to antigenic variation [31]. Survival

within macrophages may be an immune evasion and persistence strategy of *C. glabrata* [32]. Viruses have also developed sophisticated means of escaping immune response. They can escape the first round of innate immunity by impairing monocyte differentiation into functional DCs [33,34] or develop “shape change” strategies, such as impairing non-self RNA recognition by replicating the self cap structures [35]. EBV produces a regulatory cytokine that is nearly indistinguishable from human IL-10 [36,37]. In response to attack by T cells, pathogens such as human immunodeficiency virus (HIV) and hepatitis C virus (HCV) have been demonstrated to avoid T-cell mediated response by the deletion of Teff epitopes [38-40].

Adding to this repertoire of escape mechanisms, we recently described “self-like” or redundant epitopes in certain viruses that exhibit on the TCR face patterns that are conserved with human Treg epitope sequences, potentially reducing immune response against the pathogen. We proposed a novel means of immune escape - immune camouflage – that can be discovered using immunoinformatics tools that compare T cell epitopes [17]. Further evidence is emerging that HIV [41], HCV [42], EBV, HSV, and avian influenza (H7N9) [43] have built themselves a very successful niche using bits and pieces of the human genome, which allow them to create an immune signature that is indistinguishable from self.

The impact of Immune camouflage through epitope redundancy in immune responses to pathogens may be difficult to discern at the population level, since the net effect of these influences may vary in different individuals due to previous exposures, vaccinations, and the HLA-restriction of cross-reactive immune responses. The identification of redundant epitopes that are restricted by multiple HLA, as described in the next section, has enabled our group to evaluate these epitopes *in vitro* and define patterns that may be associated with regulatory or auto-reactive immune response.

Redundancy and regulation of T cell response

It should be noted that cross-conservation of T cell epitopes from pathogens and autoantigens, and studies that show that these epitopes can be either immune-regulatory and/or immunostimulatory, is not a new discovery. For example, heat shock proteins have been the focus of intense research, as they are conserved in bacterial species and humans. Van Herwijnen et al. found a mycobacterial (Mtb) HSP70 peptide (B29) that is highly conserved with human to be an immunodominant peptide in BALB/c mice [44]. They adoptively transferred the Mtb B29-specific Tregs and were able to show that these mycobacterial epitope-specific Tregs suppressed ongoing experimentally induced arthritis in mice. In a separate study, peptides H161 and H167, derived from the most conserved region of HSP70 (both among species and among human HSP70 variants), showed strong recognition by PBMCs of healthy humans; T cell proliferation, gamma-interferon production and IL-10 secretion were reported [45]. More recently, peptides derived from HSP "BiP" were shown to include T cell proliferation, IL-10 expression, and to reduce inflammation in a murine model [46]. In a separate study, stress-induced presentation of HSP-derived peptides in mouse APC resulted in activation of CD4+ hybridoma T cells specific for the identified peptide [47]. The current explanation for this observation is that stress-induced augmented expression of inducible HSP70 leads to enhanced presentation of HSP70 peptides on MHC class II, leading to recognition by HSP70-specific T cells, which may subsequently dampen inflammation [48]. It is not clear whether these HSP-specific T cells are iTreg or nTreg, or even T eff, and how they evolve to adapt a particular phenotype. We have examined these epitopes for epitope redundancy, and did find that the epitopes are not only promiscuous, but they are also highly conserved (across HSP proteins) and that many HSP peptides are found in the thymic and lymph 'peptidomes' [49,50,51].

Epitope Mapping using Computational Tools

The process of identifying which pathogen sequences trigger T effector immune responses and which trigger regulatory T cell responses may have been simplified by the development of epitope-mapping tools. Most T cell epitope mapping tools evaluate the amino acid sequence of each input 9-mer peptide (derived from the set of overlapping 9-mer sequences in any protein antigen) to a set of binding coefficients that define the propensity of the sequence to bind to HLA. On an *overall* scale, proteins carrying more putative epitopes would expected to be more immunogenic while proteins carrying fewer putative epitopes would tend to be less immunogenic [14]. On a *local* scale, T-cell epitopes (class II-restricted in particular) are not randomly distributed throughout protein sequences but instead tend to "cluster" in specific regions, and are known as promiscuous T cell epitopes [52]. Promiscuous class II restricted T-cell epitope "clusters" can range from 9 to roughly 25 amino acids in length and, considering their affinity to multiple alleles and across multiple binding registers, can contain anywhere from 4 to 40 binding motifs. Many of the most reactive T-cell epitope clusters present a feature we have described as an "EpiBar" (Epitope Bar). An EpiBar is a single 9-mer frame containing binding motifs for at least four different HLA alleles (**Figure 3**). In retrospective evaluations of published promiscuous epitopes, we have found this pattern to be a signature feature of highly immunogenic, promiscuous class II epitopes [53,54].

Some promiscuous epitopes that have this feature and are highly redundant (either found in very prevalent proteins, or highly cross-conserved with the human genome) have been shown to trigger Tregs to respond and expand. This discovery has been true for epitopes discovered by our group (IgG T cell epitopes known as Tregitopes [12], HCV [42], H7N9 [43]) as well as for Treg epitopes discovered by other groups such as Edratide [55] and epitopes found in heat shock proteins: HSP60 [56,57,58], HSP 70 [48] and BiP [46].

Improving computational tools to identify epitope redundancy

To investigate further this newly discovered TCR-facing epitope redundancy on a very large scale, we built an algorithm called JanusMatrix that can identify peptides with similar MHC-binding propensities that are TCR-face homologous, in any genome, as compared to a reference epitope [13]. After the binding prediction analysis described above has been run on an antigen, JanusMatrix compares the 9-mer's TCR-facing residues to the proteins of the human genome and the human microbiome. Significant differences between the degree of cross reactivity with protein databases of human genome, human microbiome, and human pathogen sequences when comparing known T effector and T regulatory epitopes are identified by ratios [13]. JanusMatrix "Homology" score is calculated by determining the average depth of coverage in the target database for each epitope identified in the source sequence. Homology to non-binding peptides is assumed to be irrelevant.

JanusMatrix is now used by our group and our collaborators to identify significant cross-conservation with proteins contained within the human genome and the human microbiome. We found that immune response to T cell epitopes was inversely correlated with their degree of cross-reactivity with the human genome in a subsequent study of H7N9 peptides [43]. Examining more than 4,000 published human T cell epitopes (from the Immune Epitope Database, IEDB, http://www.iedb.org/), we recently determined that the JanusMatrix "homology" score is significantly higher ($p<0.0001$) for T cell epitopes shown to induce regulatory cytokine (IL-10) response (N=3,780 epitopes), and significantly lower ($p<0.01$) for T cell epitopes shown to induce effector cytokine (IL-4) response (N=1,433 epitopes) [14]. Thus large-scale analysis of published T cell epitopes in public databases has provided additional confirmation for the hypothesis that epitope-redundancy deviate the phenotype of T cell-mediated immune responses.

From redundant epitopes to Treg epitopes

Using JanusMatrix, we identified human-like T cell epitopes in the genomes of viruses and bacteria. Analyzing pathogens that co-evolved with humans (e.g. EBV, CMV, HIV), we identified a disproportionately large number of linear epitopes that possess TCR-facing amino acid sequences that are highly conserved with self [17]. In the following paragraphs, we provide a summary of Treg epitopes that were identified using JanusMatrix, in H7N9, HCV, and HIV.

- **Treg epitopes in H7N9 influenza**

Avian-origin H7N9 influenza is a novel influenza A virus that emerged in humans in China in 2013. Using JanusMatrix, we identified several H7N9 T cell epitopes with TCR-facing residues identical to those of multiple epitopes from human proteins [59]. These H7N9 human-like T cell epitopes possess low immunogenicity; in PBMCs derived from healthy donors, human Teff responses to individual H7N9 peptides was inversely correlated with the degree of the peptide's resemblance to self [43]. The human-like H7N9 version of an immunodominant hemagglutinin T cell epitope suppressed IFN-gamma responses when co-administered with peptides having less cross-reactivity with the human genome. T cells that respond to this epitope demonstrated a Treg phenotype [43]. In further as yet unpublished work, we have demonstrated that the Treg-activating epitopes may contribute to lower antibody responses in humanized mice, and modification of these epitopes may improve H7N9 vaccine efficacy.

- **HCV Treg epitopes**

Using JanusMatrix, we discovered a promiscuous class II epitope in HCV that is conserved with hundreds of human homologs. The epitope induces CD4+CD25+FoxP3+ Treg proliferation and function in peripheral blood leukocyte cultures from an HLA-diverse cohort of HCV-infected patients, but not in patients who spontaneously cleared HCV nor in non-infected individuals [42]. Human homologs of the HCV epitope stimulate Tregs in both HCV- and non-infected people,

suggesting that tolerance of this particular protein is promoted by activation of Tregs that recognize a common TCR-face. It is well known that HCV epitopes mutate over the course of infection to decrease MHC binding. This has been well-defined for CTL epitopes, but is less well defined for CD4+ T cell epitopes [60-62]. However, in the case of the HCV epitope described by Losikoff et al. [42], the HCV genomic sequences appears to contain the same TCR-facing residues as highly prevalent autologous T cell epitopes, so as to acquire the potential to drive Treg responses in an HLA-diverse population.

- **HIV Treg epitopes**

HIV also exhibits curious patterns of cross-reactivity by JanusMatrix analysis. In searching HIV envelope sequences for cross-conservation with the human genome, we recently uncovered a high frequency of human MHC molecule sequences that share a TCR-face with a highly conserved human epitope in the envelope protein (orange highlight, **Figure 4**). This protein is present on the cell surface of nearly every living cell, and the conserved epitope is a highly conserved motif, despite allelic variation. An IEDB search shows a closely related epitope has been described as capable of stimulating T cell proliferation [63] (the phenotype of the proliferating cells was not defined). If these CD4+ T cells are regulatory in nature, activation of the T cells by the HIV homolog epitope could promote HIV expansion [64], which may promote HIV viral persistence and increase HIV 'fitness', instead.

These three case studies are consistent with our previously published observation that that 'hit and stay' viruses like HCV and HIV escape protective immune responses by cross-reacting with Tregs [17] and suggests that Treg-activating HCV and HIV sequences may affect HIV and HCV vaccine efficacy.

- **Patterns may be common to Treg epitopes**

In our retrospective evaluations of the computational signature of these promiscuous, Treg-inducing epitopes, we find several common characteristics that may be useful for future studies. First, as suggested above, the T cell epitopes that induce Treg responses tend to be promiscuous epitopes (presented by multiple class II HLA DR alleles), and second, they tend to be found in extremely prevalent proteins (such as Immunoglobulin G for Tregitopes and Edratide (hCDR1), and heat shock proteins such as HSP 60 (Diapep277, [57]) and HSP 70 [65]. A third feature that is common to such epitopes, when found in pathogens, is that their sequence of amino acids that face the TCR face may be identical, or redundant, with many human genome sequences that are predicted to bind to the same MHC alleles [17]. This suggests that the prevalence of peptides (or the frequency with which circulating T cells may encounter peptides) that bind to the same MHC and present the same TCR face to the T cell may be important in terms of determining the phenotype of the immune responses.

Redundant epitopes in autoimmune diseases

Epitope redundancy may also cause the immune response to go awry, where cross-reactivity between pathogen and self triggers effector T cell response against self. Due to TCR cross-recognition between self and pathogen, exposure to a pathogen may trigger anti-self responses, leading to the initiation of an auto-immune condition. A number of examples have been published, such as conservation between certain EBV and autologous epitopes in nerve-associated proteins, potentially contributing to the development of Multiple Sclerosis (MS) [66] (**Figure 5**). Vaccination has also been reported to be associated with immune diseases [15,67,68]. Genes involved in TCR recognition of epitopes such as HLA-DRB1, CD4 and INF-gamma or TGFbeta1 have been involved in autoimmune diseases known to be associated with vaccination [69]. In the next section, we will focus on recent advances on selected autoimmune diseases that are believed to be

related to infection with pathogens, and explore whether epitope redundancy could explain many of these autoimmune-infectious disease linkages.

- **Guillain-Barré Syndrome**

Guillain-Barré syndrome (GBS) is the main cause of acute flaccid paralysis, characterized by a weakness in muscle contraction. In about 60% of the cases, an upper respiratory or gastrointestinal infection precedes the onset of the disease. Among the pathogens that trigger GBS, C*ampylobacter jejuni* (*C. jejuni*) is the most frequent (in 13% to 39% of cases), followed by cytomegalovirus (5% to 22%), EBV (1% to 13%), and Mycoplasma pneumoniae (5%) [70]. One of the most documented mechanisms for GBS is the presence of anti-gangliosides antibodies (anti-GM1 and anti-GD1), induced by similarity of these self-proteins with lipo-oligosaccharides present on *C. jejuni*'s LPS [71]. There is also evidence of an association between influenza virus and GBS. In 1976 and 2009 in the US, seasonal influenza strains were of swine origin and an increased relative risk of developing GBS was observed in both cases, whether the vaccine was adjuvanted or not [72,73]. Although these results were not confirmed in Europe [74], several studies have established a link between GBS and infection by the 2009 pH1N1 [75,76].

How the swine influenza virus and the associated vaccines may have triggered GBS is not currently known. Although anti-ganglioside antibodies can recognize the influenza virus (probably due to glycosylation), no such antibodies were found in sera of vaccinated patients [77]. Although no studies have identified specific T cell epitopes involved, it is clear that their formation is T cell-dependent. Several populations of T-helper cells have been identified in cases of *C. Jejuni*-induced GBS [78], and immunoglobulin class-switching has been observed in transgenic mice that lack complex ganglioside when immunized with *C. Jejuni* LPS [79]. Evaluation of potential cross-reactivity between GBS-associated antigens in pathogens and self is both feasible and actionable using JanusMatrix.

- **Narcolepsy**

Narcolepsy is a neurological disorder characterized by a selective loss of hypocretin-secreting neurons that results in excessive daytime sleepiness [80]. Strong genetic associations with selected HLA alleles in narcolepsy provides support for an autoimmune basis of hypocretin-secreting neuron loss. Narcolepsy occurs almost exclusively in patients whose HLA is DQA1*01:02 and DQB1*06:02 [81,82]. Strong correlations between the onset of narcolepsy with *Streptococcus pyogenes* [83] and influenza A virus infections have been reported, with epitope redundancy suspected to play a major role in triggering the autoimmune reaction. A 6-9 fold increase in the risk of developing narcolepsy was observed following 2009-2010 H1N1 influenza pandemics in Europe [84,85] and China [86]. In these regions, the onset of narcolepsy was observed following vaccination with Pandemrix, an adjuvanted pH1N1 vaccine. A recent report has identified the presence of autoantibodies against ganglioside GM3 in Pandemrix-vaccinated children who developed narcolepsy with cataplexy [87], providing additional support for the theory that influenza-associated narcolepsy is autoimmune, and suggesting an underlying mechanism similar to GBS. Markers of inflammation (interferon gamma, CCL11, IL12) identified by Luminex assay have been associated with onset of narcolepsy [88].

In addition, a peptide from the surface-exposed region of influenza nucleoprotein A shared protein residues with a fragment of the extracellular domain of hypocretin receptor 2 in the host. Antibodies against these epitopes were present in the sera of individuals that previously received the Pandemrix vaccine [15], suggesting that the nucleoprotein A epitope is a redundant epitope. The presence of the adjuvant may have served as a catalyst to establish the autoimmunity in this particular case. The extent to which the relevant epitope is recognized by patients who do not have narcolepsy remains to be determined.

- **Multiple sclerosis**

Multiple sclerosis (MS) is the most common autoimmune disease, affecting more than 2 million people over the world. Demyelination of axons in the central nervous system contributes to the neurological disability associated with MS. Th1 and Th17 effector cells have been shown to play a major role in the pathogenesis of the disease [89], along with a deficit in Treg cells [90,91]. Viral or bacterial infections [92,93], commensal microbiota [94] have been associated to MS. Cross-conserved or redundant T cell epitopes may be involved in MS pathogenesis [95]. For example, a number of stimulatory, cross-reactive peptide sequences from environmental and human antigens were found in MS patient-derived anti-myelin TCR [96]. Furthermore, a peptide from *H. influenzae* mimicking a PLP-peptide has been shown to induce CNS disease in a mouse model [97].

- **Autoimmune Disease and the Gut Microbiome**

Links between bacterial infections and autoimmune diseases of the gut have been found [98,99]. However, the complexity of the human microbiome may make it difficult to deconvolute the many cross-reactive peptides that could be driving this link. Associations between specific gut microbes and disease [100] may make it possible to narrow down the possibilities, although it seems unlikely that a single gut microbe will be the culprit.

In conclusion, the impact of epitope redundancy on autoimmunity and infectious diseases may be difficult to decipher, since many factors can influence the outcome of infection and immune response. Focusing only on the T cell epitope component of the complex interaction between host and pathogen, T cell-related determinants of immune response include the breadth of T cell epitopes in the pathogen that can be presented to the immune system, and the individual HLA molecules that present these epitopes to the host. In addition, differences in immune responses to specific epitopes may be defined in advance due to previous exposure to the same pathogen (memory T cell response) and/or due to exposure to other pathogens that have similar, cross-

reactive epitopes (heterologous immunity). Rather than view these factors as impediments in the search for linkages between infectious diseases, autoimmunity, and tolerance, we view them as new frontiers that can be explored, applying recently developed computational tools such as JanusMatrix and T cell phenotyping techniques.

Expert Commentary

The development of new immunoinformatics tools have allowed to better define and understand the concept of T cell epitope redundancy between self and pathogens. Not only is epitope-redundancy relevant to host-pathogen immune responses, as it is involved in immune camouflage, but it is also relevant to autoimmunity. One of the major challenges to understanding mechanisms of immune tolerance and its evil twin, autoimmunity, is to identify the T cell epitopes that are recognized by regulatory and effector (inflammatory) T cell populations. Dissecting proteins for these linear determinants of T cell activation is essential to designing therapies that will reverse the course of autoimmune diseases and explain long-established linkages between autoimmunity and infections. Tools for pathogen-TCR-facing-self-wise analysis are now available, making it possible to investigate thse important topics. Further studies are necessary to answer these two important questions: (1) do highly TCR cross-conserved epitopes play a key role in tolerance to self; and (2) if epitope redundancy is also relevant to T effector epitopes, are epitopes with lower TCR-face cross-conservation associated with post-infection autoimmune diseases, such as Narcolepsy and Guillain-Barré syndrome? These are truly fundamental questions that are directly relevant to improving vaccines and safeguarding human health.

Five-year view

Biomaterial scientists have been adopting lessons learned from pathogens to evade immune responses to their products for decades. Can vaccinologists and immunologists, rheumatologists

and maybe also biologics developers learn about the immunopathogenesis of disease through explorations of epitope redundancy? To be more precise, perhaps we should learn from viruses, bacteria and parasites, by examining which host epitopes they choose to imitate. Why do they camouflage themselves with one epitope (sometimes multiple pathogens use the same 'self' epitope) and not the other? Do those [human] epitopes play a key role in the regulation of human immune response? Putting those special epitopes into biologics, may reduce immune responses to the drugs, improving patient outcomes. And taking them out of vaccines, may result in better vaccines. What has proven to be a weakness in host defense, may be the hope that emerges, leading to solutions for unmet medical needs.

It is time that we explore host-pathogen cross-reactivity at the molecular level, expanding the application of immunoinformatics analysis to new pathogens, and to pursue interdisciplinary studies of the immunopathogenesis of autoimmune disease and tolerance, with the goal of understanding the accelerating epidemic of autoimmune disease, and improving vaccine efficacy and global health for the 21st century.

Key issues

- T cell epitope cross-conservation, or ‘epitope redundancy’, is a consequence feature of TCR training on self antigens in the thymus.
- Epitope redundancy can be uncovered using computational tools that define epitopes, and new tools that allow exploration of TCR-conserved, similar MHC-binding epitopes in large sets of peptides (such as the human genome).
- Numerous studies point to epitope redundancy as a contributor to the development of auto-immune disease.
- New studies are revealing epitope-redundancy as a trigger for tolerance, potentially explaining the immunomodulatory effect of certain epitopes derived from Heat Shock Proteins, and other as-yet unexplained immunological observations.
- Applying computational tools to the human genome, to human pathogens, and to the human microbiome may uncover new relationships between infectious diseases, autoimmunity, and tolerance.

References

Reference annotations

** Of interest*

*** Of considerable interest*

**1. Ahmed SS, Volkmuth W, Duca J, et al. Antibodies to influenza nucleoprotein cross-react with human hypocretin receptor 2. Sci Transl Med 2015;7(294):294ra105

One of several recent publications seeking to explain the relationship between the uptick in narcolepsy cases following the emergence of H1N1 2009 swine-origin influenza and observed differences between H1N1 vaccine-associated narcolepsy and selected vaccines. This publication identifies cross-reactivity between the hypocretin receptor, and cross-reactivity with M1 protein, as the likely link, and shows a plausible association with HLA DQ, as described previously.

2. Zhao J, Zhao J, Fett C, et al. IFN-γ- and IL-10-expressing virus epitope-specific Foxp3(+) T reg cells in the central nervous system during encephalomyelitis. J Exp Med 2011;208(8):1571-7

3 . Fazilleau N1, Bachelez H, Gougeon ML, et al. Cutting edge: size and diversity of CD4+CD25high Foxp3+ regulatory T cell repertoire in humans: evidence for similarities and partial overlapping with CD4+CD25- T cells. J Immunol 2007;179(6):3412-6

4 . Pacholczyk R, Ignatowicz H, Kraj P, et al. Origin and T cell receptor diversity of Foxp3+CD4+CD25+ T cells. Immunity 2006;25(2):249-59

5. Wong J, Obst R, Correia-Neves M, et al. Adaptation of TCR repertoires to self-peptides in regulatory and nonregulatory CD4+ T cells. J Immunol 2007;178(11):7032-41

6. Hsieh CS, Zheng Y, Liang Y, et al. An intersection between the self-reactive regulatory and nonregulatory T cell receptor repertoires. Nat Immunol 2006;7(4):401-10.

* 7. Moise L, Tassone R, Latimer H, et al. Immunization with cross-conserved H1N1 influenza CD4+ T-cell epitopes lowers viral burden in HLA DR3 transgenic mice. Hum Vaccin Immunother 2013;9(10):2060-8

In vivo (murine) validation of the concept that T cell epitopes can protect against severe influenza. This observation was also validated by Richard Webby, who performed an analogous study in ferrets and showed that both CD4+ T helper and CD8+ T effector T cells were critically important to protection against emerging H1N1 influenza

8. Friberg H, Bashyam H, Toyosaki-Maeda T, et al. Cross-reactivity and expansion of dengue-specific T cells during acute primary and secondary infections in humans. Sci Rep 2011;1:51

9. Shahrizaila N, Yuki N. Guillain-barré syndrome animal model: the first proof of molecular mimicry in human autoimmune disorder. J Biomed Biotechnol 2011;2011:829129

*10. Slobedman B, Barry BA, Spencer JV, et al. Virus-Encoded Homologs of Cellular Interleukin-10 and Their Control of Host Immune Function. J. Virol 2009;83(19) 9618-29

An important mechanism of viral immune escape; it suggests that IL-10,-like self proteins, is actively engaged in the modulation of the immune response to viral infection.

11. Yarkoni S, Kaminitz A, Sagiv Y, et al. I Involvement of IL-2 in homeostasis of regulatory T cells: the IL-2 cycle. Bioessays 2008;30(9):875-88

**12. De Groot AS, Moise L, McMurry JA, et al. Activation of natural regulatory T cells by IgG Fc-derived peptide "Tregitopes". Blood 2008;112(8):3303-11

The first publication to describe the discovery of Tregitopes, which are promiscuous Class II restricted T cell epitopes that are found in IgG and shown to modulate immunogenicity.

**13. Moise L, Gutierrez AH, Bailey-Kellogg C, et al. The two-faced T cell epitope: examining the host-microbe interface with JanusMatrix. Hum Vaccin Immunother 2013;9(7):1577-86

Describes a new algorithm for the identification of Treg epitopes, called JanusMatrix.

14. Moise L, Gutierrez A, Kibria F, et al. iVAX: An integrated toolkit for the selection and optimization of antigens and the design of epitope-driven vaccines. Hum Vaccin Immunother 2015;11(9):2312-21

15. Ahmed SS, Schur PH, MacDonald NE, Steinman L. Narcolepsy, 2009 A(H1N1) pandemic influenza, and pandemic influenza vaccinations: what is known and unknown about the neurological disorder, the role for autoimmunity, and vaccine adjuvants. J Autoimmun 2014;50:1-11

16. Chen AT, Cornberg M, Gras S, et al. Loss of anti-viral immunity by infection with a virus encoding a cross-reactive pathogenic epitope. PLoS Pathog 2012;8(4):e1002633

*17. He L, De Groot AS, Gutierrez AH, et al. Integrated assessment of predicted MHC binding and cross-conservation with self reveals patterns of viral camouflage. BMC Bioinformatics 2014;15 Suppl 4:S1

An article demonstrating the application of JanusMatrix to cross-reactivity between viral pathogens and self.

18. De Groot AS, Moise L, Liu R, et al. Immune camouflage: relevance to vaccines and human immunology. Hum Vaccin Immunother 2014;10(12):3570-5

19. Cusick MF, Libbey JE, Fujinami RS. Molecular mimicry as a mechanism of autoimmune disease. Clin Rev Allergy Immunol 2012;42(1):102-11

20. Sharma S, Thomas PG. The two faces of heterologous immunity: protection or immunopathology. J Leukoc Biol 2014;95(3):405-16

21. Weber CA, Mehta PJ, Ardito M, et al. T cell epitope: friend or foe? Immunogenicity of biologics in context. Adv Drug Deliv Rev 2009;61(11):965-76

22. Page KR, Scott AL, Manabe YC. The expanding realm of heterologous immunity: friend or foe? Cell Microbiol 2006;8(2):185-96.

23. Che JW, Selin LK, Welsh RM. Evaluation of non-reciprocal heterologous immunity between unrelated viruses. Virology 2015;482:89-97

24. Shen ZT, Nguyen TT, Daniels KA, et al. 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

25. Birnbaum ME, Mendoza JL, Sethi DK, et al. Deconstructing the peptide-MHC specificity of T cell recognition. Cell 2014;157(5):1073-87

*26. Madi A, Shifrut E, Reich-Zeliger S, et al. T-cell receptor repertoires share a restricted set of public and abundant CDR3 sequences that are associated with self-related immunity. Genome Res 2014;24(10):1603-12

Observations that begin to uncover the links between shared T cell receptor sequences, pathogen T cell epitopes, and self epitopes.

*27. Su LF, Kidd BA, Han A, et al. Virus-specific CD4(+) memory-phenotype T cells are abundant in unexposed adults. Immunity 2013 21; 38(2):373-83Publication that was concurrent with the development of JanusMatrix and validates JanusMatrix findings.

28. Liu J, Cao S, Peppers G, et al. Clonotype-specific avidity influences the dynamics and hierarchy of virus-specific regulatory and effector CD4(+) T-cell responses. Eur J Immunol 2014;44(4):1058-68

29. Lucas M, Karrer U, Lucas A, et al. Viral escape mechanisms-escapology taught by viruses. Int J Exp Pathol 2001;82(5):269-86

30. Matsuura M. Structural Modifications of Bacterial Lipopolysaccharide that Facilitate Gram-Negative Bacteria Evasion of Host Innate Immunity. Front Immunol 2013;4:109

31. Ukaegbu UE, Zhang X, Heinberg AR, et al. A Unique Virulence Gene Occupies a Principal Position in Immune Evasion by the Malaria Parasite Plasmodium falciparum. PLoS Genet 2015;11(5):e1005234

32. Kasper L, Seider K, Hube B. Intracellular survival of Candida glabrata in macrophages: Immune evasion and persistence. FEMS Yeast Res 2015;15(5):pii: fov042

33. Granato M, Lacconi V, Peddis M, et al. Hepatitis C virus present in the sera of infected patients interferes with the autophagic process of monocytes impairing their in-vitro differentiation into dendritic cells. Biochim Biophys Acta 2014;1843(7):1348-55

34. Li L, Liu D, Hutt-Fletcher L, et al. Epstein-Barr virus inhibits the development of dendritic cells by promoting apoptosis of their monocyte precursors in the presence of granulocyte macrophage-colony-stimulating factor and interleukin-4. Blood 2002;99(10):3725-34

35. Menachery VD, Yount BL Jr, Josset L, et al, Agnihothram S, Katze MG, Baric RS. Attenuation and restoration of severe acute respiratory syndrome coronavirus mutant lacking 2'-o-methyltransferase activity. J Virol 2014;88(8):4251-64

36 Vieira P, de Waal-Malefyt R, Dang MN, et al. Isolation and expression of human cytokine synthesis inhibitory factor cDNA clones: homology to Epstein-Barr virus open reading frame BCRFI. Proc Natl Acad Sci USA 1991;88(4):1172-6

37. Sin SH, Dittmer DP. Cytokine homologs of human gamma herpes viruses. J Interferon Cytokine Res 2012;32(2):53-9

38. Larrubia JR, Moreno-Cubero E, Lokhande MU, et al. Adaptive immune response during hepatitis C virus infection. World J Gastroenterol 2014;20(13):3418-30

39. Song H, Pavlicek JW, Cai F, et al. Impact of immune escape mutations on HIV-1 fitness in the context of the cognate transmitted/founder genome. Retrovirology 2012;9:89

40. Vider-Shalit T, Sarid R, Maman K, et al. Viruses selectively mutate their CD8 CTL epitopes–a large-scale immunomic analysis. Bioinformatics 2009;25(12):i39-44

41. Moise L, Terry F, Gutierrez AH, et al. Smarter vaccine design will circumvent regulatory T cell-mediated evasion in chronic HIV and HCV infection. Front Microbiol 2014;6:5:502

**42. Losikoff PT, Mishra S, Terry F, et al. HCV epitope, homologous to multiple human protein sequences, induces a regulatory T cell response in infected patients. J Hepatol 2015;62(1):48–55

Description and validation of a Treg epitope that is highly conserved in the HCV genome.

** 43. Liu R, Moise L, Tassone R, et al. H7N9 T-cell epitopes that mimic human sequences are less immunogenic and may induce Treg-mediated tolerance. Human Vaccin Immunother 2015;11(9):2241-52

Description and validation of Treg epitopes located in highly conserved sequences in the H7N9 genome.

*44. van Herwijnen MJ, Wieten L, Van der zee R, et al. Regulatory T cells that recognize a ubiquitous stress-inducible self-antigen are long-lived suppressors of autoimmune arthritis. Proc Natl Acad Sci USA 2012;109(35):14134-9

Heat-shock protein (HSP) -specific T cells appear to recognize highly promiscuous T cell epitopes in the ubiquitous HSP protein; these T cells may represent HSP-Tregitopes.

45. de Jong H, Koffeman EC, Meerding JM, et al. T cell recognition of naturally presented epitopes of self-heat shock protein 70. Cell Stress Chaperones 2014;19(4):569-78

*46. Shoda H, Fujio K, Sakurai K, et al. Autoantigen BiP-Derived HLA-DR4 Epitopes Differentially Recognized by Effector and Regulatory T Cells in Rheumatoid Arthritis. Arthritis Rheumatol 2015;67(5):1171-81

Similar observation to #45 above: Heat-shock protein (HSP)-specific T cells appear to recognize highly promiscuous T cell epitopes in another ubiquitous HSP protein; these T cells may represent BiP-Tregitopes.

47. Wieten L, van der Zee R, Spiering R, et al. A novel heat-shock protein coinducer boosts stress protein Hsp70 to activate T cell regulation of inflammation in autoimmune arthritis. Arthritis Rheum 2010;62(4):1026-35

**48. van Eden W, van der Zee R, Prakken B. Heat-shock proteins induce T-cell regulation of chronic inflammation. Nat Rev Immunol 2005;5(4):318-30Early evidence that autologous epitopes could modulate immune response; recognition of these T cell epitopes may be impaired in some forms of autoimmune disease.

49. Adamopoulou E, Tenzer S, Hillen N, et al. Exploring the MHC-peptide matrix of central tolerance in the human thymus. Nat Commun 2013;4:2039

50. Espinosa G, Collado JA, Scholz E, et al. Peptides presented by HLA class I molecules in the human thymus. J Proteomics 2013;94:23-36

51. Clement CC, Cannizzo ES, Nastke MD, et al. An expanded self-antigen peptidome is carried by the human lymph as compared to the plasma. PLoS One 2010;5(3):e9863

52. Nascimento EJ, Mailliard RB, Khan AM, et al. Identification of conserved and HLA promiscuous DENV3 T-cell epitopes. PLoS Negl Trop Dis 2013;7(10):e2497

53. Zhang S, Desrosiers J, Aponte-Pieras JR, et al. Human immune responses to H. pylori HLA Class II epitopes identified by immunoinformatic methods. PLoS One 2014;9(4):e94974

54. Moise L, McMurry JA, Pappo J, et al. Identification of genome-derived vaccine candidates conserved between human and mouse-adapted strains of H. pylori. Hum Vaccin 2008;4(3):219-23

*55. Sthoeger Z, Zinger H, Sharabi A, et al. The tolerogenic peptide, hCDR1, down-regulates the expression of interferon-α in murine and human systemic lupus erythematosus. PLoS One 2013;8(3):e60394

An IgG-derived Tregitope, found in the ubiquitous Immunoglobulin G protein, showing evidence of tolerance induction in murine and human SLE. This peptide is highly homologous to Tregitope 029 (unpublished, De Groot and Martin EpiVax Inc)

56. Zanin-Zhorov A, Cahalon L, Tal G, et al. Heat shock protein 60 enhances CD4+ CD25+ regulatory T cell function via innate TLR2 signaling. J Clin Invest 2006;116(7):2022-32

57. Kamphuis S, Kuis W, de Jager W, et al. Tolerogenic immune responses to novel T-cell epitopes from heat-shock protein 60 in juvenile idiopathic arthritis. Lancet 2005;366(9479):50-6

58. Huurman VA, van der Meide PE, Duinkerken G, et al. Immunological efficacy of heat shock protein 60 peptide DiaPep277 therapy in clinical type I diabetes. Clin Exp Immunol 2008;152(3):488-97

59. De Groot AS, Ardito M, Terry F, et al. Low immunogenicity predicted for emerging avian-origin H7N9: implication for influenza vaccine design. Hum Vaccin Immunother 2013; 9:950-6

60. Petrovic D, Dempsey E, Doherty DG, et al. Hepatitis C virus--T-cell responses and viral escape mutations. Eur J Immunol 2012;42(1):17-26

61. Norris PJ, Stone JD, Anikeeva N, et al. Antagonism of HIV-specific CD4+ T cells by C-terminal truncation of a minimum epitope. Mol Immunol 2006;43(9):1349-57

62. Harcourt GC, Garrard S, Davenport MP, Edwards A, Phillips RE. HIV-1 variation diminishes CD4 T lymphocyte recognition. J Exp Med. 1998;188(10):1785-93

63. Atassi H, Atassi MZ. HIV envelope protein is recognized as an alloantigen by human DR-specific alloreactive T cells. Hum Immunol 1992;34(1):31-8

64. Sanjuán R, Nebot MR, Peris JB, et al. Immune activation promotes evolutionary conservation of T-cell epitopes in HIV-1. PLoS Biol 2013;11(4):e1001523

65. Wachstein J, Tischer S, Figueiredo C, et al. HSP70 enhances immunosuppressive function of CD4(+)CD25(+)FoxP3(+) T regulatory cells and cytotoxicity in CD4(+)CD25(-) T cells. PLoS One 2012;7(12):e51747

*66. Rist MJ, Hibbert KM, Croft NP, et al. T Cell Cross-Reactivity between a Highly Immunogenic EBV Epitope and a Self-Peptide Naturally Presented by HLA-B*18:01+ Cells. J Immunol 2015;194(10):4668-75

This study shows T cell cross-reactivity between self and a common human virus and the razor thin edge between tolerance and autoimmunity.

67. Vaarala O, Vuorela A, Partinen M, et al. Antigenic differences between AS03 adjuvanted influenza A (H1N1) pandemic vaccines: implications for Pandemrix-associated narcolepsy risk. PLoS One 2014;9(12):e114361

68. Oliveira AC, Maria Henrique da Mota L, Dos Santos-Neto LL, et al. Occurrence of Autoimmune Diseases Related to the Vaccine against Yellow Fever. Autoimmune Dis. 2014;2014:473170

69. McGarvey PB, Suzek BE, Baraniuk JN, et al. In silico analysis of autoimmune diseases and genetic relationships to vaccination against infectious diseases. BMC Immunol 2014;15:61

70. Hadden RD, Karch H, Hartung HP, et al. Preceding infections, immune factors, and outcome in Guillain-Barré syndrome. Neurology 2001;56(6):758-65

71. Yu RK, Usuki S, and Ariga T. Ganglioside Molecular Mimicry and Its Pathological Roles in Guillain-Barré Syndrome and Related Diseases. Infect Immun 2006;74(12): 6517–27

72. Marks JS, Halpin TJ. Guillain-Barré syndrome in recipients of A/New Jersey influenza vaccine. JAMA 1980;243(24):2490-4

73. Polakowski LL, Sandhu SK, Martin DB, et al. Chart-confirmed guillain-barre syndrome after 2009 H1N1 influenza vaccination among the Medicare population, 2009-2010. Am J Epidemiol 2013;178(6):962-73

74. Romio S, Weibel D, Dieleman JP, et al. Guillain-Barré syndrome and adjuvanted pandemic influenza A (H1N1) 2009 vaccines: a multinational self-controlled case series in Europe. PLoS One 2014;9(1):e82222

75. Stowe J, Andrews N, Wise L, et al. Investigation of the temporal association of Guillain-Barre syndrome with influenza vaccine and influenza-like illness using the United Kingdom general practice research database. Am J Epidemiol 2009;169:382–8

76. Verity C, Stellitano L, Winstone AM, et al. Guillain-Barre syndrome and H1N1 influenza vaccine in UK children. Lancet 2011;378:1545-6

77. Wang DJ, Boltz DA, McElhaney J, et al. No evidence of a link between influenza vaccines and Guillain-Barre syndrome-associated anti-ganglioside antibodies. Influenza Other Respir Viruses 2012;6(3):159-66

78. Li S, Jin T, Zhang HL, et al. Circulating Th17, Th22, and Th1 cells are elevated in the Guillain-Barré syndrome and downregulated by IVIg treatments. Mediators Inflamm 2014;2014:740947

79. Bowes T, Wagner ER, Boffey J, et al. Tolerance to self gangliosides is the major factor restricting the antibody response to lipopolysaccharide core oligosaccharides in Campylobacter jejuni strains associated with Guillain-Barré syndrome. Infect Immun 2002;70(9):5008-18

80. Savvidou A, Knudsen S, Olsson-engman M, et al. Hypocretin deficiency develops during onset of human narcolepsy with cataplexy. Sleep 2013;36(1):147-8

**81. Mignot E, Hayduk R, Black J, et al. HLA DQB1*0602 is associated with cataplexy in 509 narcoleptic patients. Sleep 1997;20(11):1012-20

An early and important contribution to the field, linking HLA (and hence T cells, and TCR redundancy) to narcolepsy.

82. Watson NF, Ton TG, Koepsell TD, et al. Does narcolepsy symptom severity vary according to HLA-DQB1*0602 allele status? Sleep 2010;33(1):29-35

83. Aran A, Lin L, Nevsimalova S, et al. Elevated anti-streptococcal antibodies in patients with recent narcolepsy onset. Sleep 2009;32(8):979-83

84. Wijnans L, Lecomte C, de Vries C, et al. The incidence of narcolepsy in Europe: before, during, and after the influenza A(H1N1)pdm09 pandemic and vaccination campaigns. Vaccine 2013;31(8):1246-54

85. Nohynek H, Jokinen J, Partinen M, et al. AS03 adjuvanted AH1N1 vaccine associated with an abrupt increase in the incidence of childhood narcolepsy in Finland. PLoS One 2012;7(3):e33536

86. Han F, Lin L, Warby SC, et al. Narcolepsy onset is seasonal and increased following the 2009 H1N1 pandemic in China. Ann Neurol 2011;70(3):410-7

87. Saariaho AH, Vuorela A, Freitag TL, et al. Autoantibodies against ganglioside GM3 are associated with narcolepsy-cataplexy developing after Pandemrix vaccination against 2009 pandemic H1N1 type influenza virus. J Autoimmun. 2015;63:68-75

88. Lecendreux M, Libri V, Jaussent I, Mottez E, Lopez R, Lavault S, Regnault A, Arnulf I, Dauvilliers Y. Impact of cytokine in type 1 narcolepsy: Role of pandemic H1N1 vaccination? J Autoimmun. 2015;60:20-31

89. Komiyama Y, Nakae S, Matsuki T, et al. IL-17 plays an important role in the development of experimental autoimmune encephalomyelitis. J Immunol 2006;177(1):566–73

90 . Viglietta V, Baecher-Allan C, Weiner HL, et al. Loss of functional suppression by CD4+CD25+ regulatory T cells in patients with multiple sclerosis. J Exp Med 2004;199(7):971–9

91. Elyaman W, Khoury SJ, Scott DW, et al. Potential application of Tregitopes as immunomodulating agents in multiple sclerosis. Neurol Res Int 2011;2011:256460

92. Wucherpfennig KW, Strominger JL. Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. Cell 1995;80:695–705

*93. Tejada-Simon MV, Zang YC, Hong J, et al. Cross-reactivity with myelin basic protein and human herpesvirus-6 in multiple sclerosis. Ann Neurol 2003;53:189–97

This study links self/pathogen T cell cross-reactivity to B cell responses with implications for pathogenicity.

94. Berer K, Mues M, Koutrolos M, et al. Commensal microbiota and myelin autoantigen cooperate to trigger autoimmune demyelination. Nature 2011;479(7374):538-41

95. Zhang X, Tang Y, Sujkowska D, et al. Degenerate TCR recognition and dual DR2 restriction of autoreactive T cells: implications for the initiation of the autoimmune response in multiple sclerosis. Eur J Immunol 2008;38(5):1297-309

96. Reynolds CJ, Sim MJ, Quigley KJ, et al. Autoantigen cross-reactive environmental antigen can trigger multiple sclerosis-like disease. J Neuroinflammation 2015;12:91

97. Croxford JL, Olson JK, Anger HA, et al. Initiation and exacerbation of autoimmune demyelination of the central nervous system via virus-induced molecular mimicry: implications for the pathogenesis of multiple sclerosis. J Virol 2005;79:8581–90

98. Ng SC, Benjamin JL, McCarthy NE, et al. Relationship between human intestinal dendritic cells, gut microbiota, and disease activity in Crohn's disease. Inflamm Bowel Dis 2011;17(10):2027-37

99. Kawamoto S, Maruya M, Kato LM, et al. Foxp3(+) T cells regulate immunoglobulin a selection and facilitate diversification of bacterial species responsible for immune homeostasis. Immunity 2014;41(1):152-65

100. Moya-Pérez A, Neef A, Sanz Y. Bifidobacterium pseudocatenulatum CECT 7765 Reduces Obesity-Associated Inflammation by Restoring the Lymphocyte-Macrophage Balance and Gut Microbiota Structure in High-Fat Diet-Fed Mice. PLoS One 2015;10(7):e0126976

101. Greidinger EL, Zang YJ, Jaimes K, et al. CD4+ T cells target epitopes residing within the RNA-binding domain of the U1-70-kDa small nuclear ribonucleoprotein autoantigen and have restricted TCR diversity in an HLA-DR4-transgenic murine model of mixed connective tissue disease. J Immunol 2008;180(12):8444-54

102. Veeraraghavan S, Renzoni EA, Jeal H, et al. Mapping of the immunodominant T cell epitopes of the protein topoisomerase I. Ann Rheum Dis 2004;63(8):982-7

103. Danke NA, Koelle DM, Yee C, et al. Autoreactive T cells in healthy individuals. J Immunol 2004;172(10):5967-72

104. Wucherpfennig KW, Yu B, Bhol K, et al. Structural basis for major histocompatibility complex (MHC)-linked susceptibility to autoimmunity: charged residues of a single MHC binding pocket confer selective presentation of self-peptides in pemphigus vulgaris. Proc Natl Acad Sci U S A 1995;92(25):11935-9

Figure Legends

Figure 1. Network of T-cell activation by pathogen epitope

Pathogens are taken by APCs which will present epitopes on their MHC molecules to initiate the activation of T cells. Trained in the thymus, T cells can recognize epitopes that are conserved on the TCR face between pathogens and autologous (self) proteins, resulting in the activation of potentially regulatory pathways (A). Some of these epitopes can also be unique to this pathogen, activating T cells that engage the immune response (B). This response may be limited due to lack of previous exposure. Pathogen epitopes can also be conserved in the human microbiome (C). The subsequent response can be either regulatory or effector, depending on the exposure.

Figure 2. Induction of the TCR repertoire

In the thymus, CD4+ T cells are exposed to self-antigens. TCR interaction with MHC II/self antigen complexes will determine the fate of these T-cells. Those that express a TCR that binds with a strong interaction with self will be removed by apoptosis (**negative selection**), as well as those with no interaction with MHC II/antigen complex (**death by neglect**). T cells with the weakest interaction with self-antigens will become conventional T cells (**Tconv**). Natural regulatory T cell (**nTreg**) arise from a population that interact with self moderately and express FoxP3. A population of T cells that interact more strongly with self but lack FoxP3 expression is called **potential Treg**. Conversely, a population of Tconv with a stronger affinity for self will have the potential to express FoxP3, thus they are called **unstable/Treg-like Tconv**. These cells are likely to become induced Tregs (**iTreg**).

Figure 3. T cell epitope mapping and immunogenicity scoring

Representation of a Typical EpiBar for the Tetanus toxin peptide AA 1234-1248. Z score indicates potential of a 9-mer frame to bind to a given HLA allele. Scores ≥ 1.64 are considered "hits"

because they fall in the 95th percentile of binding likelihood, following a normal distribution (intermediate blue). Peptides with scores in the 99th percentile are even more likely to bind to the given allele, and are shaded dark blue. Scores ranked in the 90th percentile (shown but not highlighted) are considered elevated; all other scores are masked for simplicity. 9-mer frames containing four or more alleles scoring above 1.64 are referred to colloquially as an “EpiBars” and are highlighted in yellow (see frame 1237: YKKMEAVKL). This band-like pattern is characteristic of promiscuous epitopes. The Cluster Score represents the deviation in aggregate epitope content relative to the random expectation for a peptide of similar length; Cluster Scores above +10 are considered significant.

Figure 4. JanusMatrix analysis of HIV-1 Env

Epitope network depicting the degree of similarity between the Treg-inducing 'human-like' HIV-1 Env peptide (green diamond). From this peptide, three 9-mer frames were predicted to bind HLA (grey squares). JanusMatrix identified cross-conserved 9-mers in the human genome (dark blue triangles), that were linked to their source human proteins (circles). Some of these human sequences were found in multiple human source proteins (orange highlights), indicating that this epitope is highly redundant.

Figure 5. Graphic representation of JanusMatrix networks for epitopes identified in auto-immune diseases

Epitopes identified in autoantigens involved in various autoimmune diseases were tested in JanusMatrix: **A**. Ribonucleoprotein in Mixed Connective Tissue Disease (MTCD) [101]; **B**. Topoisomerase in Systemic Scleroderma [102]; **C**. NY-ESO-1 in cancer [103]; **D**. Glutamic acid decarboxylase 65 (GAD) in type 1 diabetes [103]; **E**. Desmoglein 3 in Pemphigus Vulgaris [104]. The limited networks found for each of these auto-antigens indicates a high likelihood of inflammatory response. Green diamonds represent the original peptide analyzed in JanusMatrix; Grey squares

are the 9-mer frames predicted to bind a certain HLA; Dark blue triangles are 9-mers found in the human genome that shares TCR contact residues; Blue circles represent the source protein of these human 9-mer frames.

Financial and competing interests disclosure

ADG and WM are founders and majority owners of EpiVax, Inc. a biotechnology company that provides access to immunoinformatics tools and designs vaccines for commercial clients. LM holds options at EpiVax, Inc. Authors SB, FK, and FT are employees of EpiVax, Inc. Due to this relationship with EpiVax, the six authors acknowledge that there is a potential conflict of interest inherent in the publication of this manuscript, and assert that they made an effort to reduce or eliminate that conflict where possible. The authors thank Guilhem Richard and Genevieve De Groot for their contribution to the illustrations. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.