Structure-based selection of human metabolite binding P4 pocket of
DRB1*15:01 and DRB1*15:03, with implications for multiple sclerosis

Maneesh K. Misra1 · Vincent Damotte1 · Jill A. Hollenbach1

Received: 25 July 2017 / Revised: 25 September 2017 / Accepted: 27 October 2017 / Published online: 20 January 2018
© Macmillan Publishers Ltd., part of Springer Nature 2018

Abstract

Binding of small molecules in the human leukocyte antigen (HLA) peptide-binding groove may result in conformational changes of bound peptide and an altered immune response, but previous studies have not considered a potential role for endogenous metabolites. We performed virtual screening of the complete Human Metabolite Database (HMDB) for docking to the multiple sclerosis (MS) susceptible DRB1*15:01 allele and compared the results to the closely related yet non-susceptible DRB1*15:03 allele; and assessed the potential impact on binding of human myelin basic peptide (MBP). We observed higher energy scores for metabolite binding to DRB1*15:01 than DRB1*15:03. Structural comparison of docked metabolites with DRB1*15:01 and DRB1*15:03 complexed with MBP revealed that PhenylalanineMBP92 allows binding of metabolites in the P4 pocket of DRB1*15:01 but ValineMBP89 abrogates metabolite binding in the P1 pocket. We observed differences in the energy scores for binding of metabolites in the P4 pockets of DRB1*15:01 vs. DRB1*15:03 suggesting stronger binding to DRB1*15:01. Our study confirmed that specific, disease-associated human metabolites bind effectively with the most polymorphic P4 pocket of DRB1*15:01, the primary MS susceptible allele in most populations. Our results suggest that endogenous human metabolites bound in specific pockets of HLA may be immunomodulatory and implicated in autoimmune disease.

Introduction

Multiple Sclerosis (MS) (OMIM: 126200, MIM: 142857) is a chronic inflammatory and demyelinating disorder of the central nervous system (CNS). Demyelinated lesions throughout the CNS, involving both the white and gray matter, are responsible for progressive neurological deficits, although extensive heterogeneity in disability exists between individuals. Genes within the human leukocyte antigen (HLA) region account for the largest component of the genetic risk for MS. The primary signal in the region maps to the DRB1 gene, and specifically to the DRB1*15:01 allele, in the class II segment of this locus [1].

This strong association is observed across multiple populations, including virtually all European-ancestry population examined to-date. Interestingly, DRB1*15:01 is also observed to be the primary predisposing allele in African American MS patients, despite the fact that the closely related DRB1*15:03 is much more common in this population [2]. Although DRB1*15:03 differs from DRB1*15:01 at only a single amino acid position 30 (His>Tyr), our recent work suggests that DRB1*15:03 may not play a role in susceptibility to MS in African Americans (Damotte et al., in prep).

The prevailing view of HLA-mediated autoimmune pathogenesis involves presentation of self-derived cellular proteins by HLA molecules to T cells, with resulting activation and immune responsiveness against the self-antigen [3]. In MS, several myelin-derived antigens have been proposed in disease pathogenesis. Both myelin basic protein (MBP) and myelin oligodendrocyte glycoprotein have been shown to bind to DRB1*15:01 [4, 5], and initiate an immune response against myelin. The ligands of HLA class II molecules are produced mainly from exogenous protein sources [6]. Subsequently, they are digested by proteases and loaded onto class II HLA molecules after internalization.
in a procedure catalyzed by an endosomal and lysosomal chaperone HLA-DM [7]. The processing and loading occurs in a devoted endosomal compartment [8]. The peptide-loaded HLA are subsequently transported to the cell surface and when recognized by CD4+ T cells, stimulate the Ag-specific immune response [9]. The peptide-binding groove of the class II HLA molecule consists of nine different structural pockets (P) from P1 to P9, which accommodate the antigen peptide side chains.

A peptide from the middle region of human MBP consisting of amino acids from position 84 to 102 has been shown to be immunodominant for human MBP-specific T cells in individual carriers of the HLA-DR2 (which includes DRB1*15:01) haplotype [10, 11]. Further, it has been suggested that MBP (residue 85–99)-specific T cell clones from individuals with the DR2 haplotype are restricted to DRB1*15:01 and the MBP (85–99) peptide was also discovered to bind with purified DRB1*15:01 [12, 13]. Subsequently, two residues of MBP (85–99), valine (Val) and phenylalanine (Phe) at position 89 and 90, respectively, were recognized as P1 (Val89) and P4 (Phe92) anchor residues for the binding in DRB1*15:01 [13]. DRB1*15:01 is distinguished by the presence of a large, predominantly hydrophobic P4 pocket, due to the presence of alanine (Ala) in the polymorphic DR 8-chain at amino acid position 71 (DRβ71) [14]. Thus the large phenylalanine residue of the MBP peptide at position 92 in P4 functions as a major anchor residue for HLA-DR2 binding [14]. In contrast, the presence of valine at amino acid position 86 in the DR 8-chain at the base of the P1 pocket gives rise to a smaller pocket, obligating residence of a smaller amino acid residue [14].

Recent work suggests that natural, synthetic or environmental small molecule compounds may efficiently occupy the pockets of HLA-DR [9, 15, 16]. An examination of HLA alleles important in type-2 diabetes mellitus showed that small molecules influence the immune response in an allele-specific way by transiently occupying a dimorphic pocket on the HLA-DR molecule, preventing the rapid inactivation of the heterodimer so that free peptides can reach the vacant peptide binding site [9, 15]. These studies suggest that small molecule binding within the HLA peptide binding groove can result in profound alterations in peptide binding and potentially immune responsiveness [9, 15, 16]. This notion is bolstered by robust models of HLA mediated drug hypersensitivity demonstrating that the binding of small molecules in the peptide binding groove results in either changes in conformation or causes a register shift of the bound peptide, leading to an aberrant immune response [17, 18]. It is important to consider that accommodation of a register shift is possible only for HLA class II, as the binding groove of the class I HLA molecule is restricted at both ends, with the N and C-terminal residues of the peptide acting as anchors [19]. Peptides cross-presented by different class I alleles must bind in the same register, though not necessarily in the same conformation [20].

Despite decades of research and development of multiple lines of disease modifying therapies for MS, no drugs have been successfully developed that directly inhibit the pathogenic T cells via inhibition of HLA-DRB1 mediated presentation of neuroantigens. However, recent studies suggest promise in structure-based selection and identification of small molecules for the alteration and inhibition of allele specific HLA class II antigen presentation [15, 16]. At the same time, it goes to follow that if exogenous molecules are capable of altering immune-responsiveness, naturally occurring endogenous small molecules may function to trigger the aberrant response observed in autoimmune diseases. Human metabolites are abundant, naturally occurring small molecules produced endogenously in the course of normal metabolic processes. Additionally, scores of small molecules have been shown to be derived from the human microbiome; each of the major metabolite classes have been observed and participate in a range of biological activities, including immune modulation and antibiosis [21].

In this study, we hypothesized that naturally occurring small molecules may bind the MS susceptibility allele DRB1*15:01, resulting in conformational changes or causing a register shift to the bound peptide and thus altering T-cell recognition and responsiveness, initiating an autoimmune response. We performed virtual screening of the complete Human Metabolite Database (HMDB) for docking to the susceptible DRB1*15:01 allele and compared the results to the closely related yet non-susceptible DRB1*15:03 allele; and assessed the potential impact on binding of human myelin basic peptide (MBP).

Results

Structural analysis and identification of pockets

Structural analysis of the MBP bound DRB1*15:01 and DRB1*15:03 protein structure shows a total of six chains: chain A, B, D and E represent the 8-chain of DRB1 protein and chain C and F belongs to MBP (Fig. 1a). The Chain A and D each consist of 180 amino acid residues individually, while each of chain B and E contains 191 amino acid residues. C and F chains each comprise 15 amino acid residues.

We identified a total of 9 pockets in the MBP peptide of DRB1*15:01 and DRB1*15:03 protein structures, Val 89 and Phe 92 of MBP peptide were recognized as P1 and P4 anchor residues for DRB1*15:01 and DRB1*15:03 protein.
structures, as previously reported [14]. The surface hydrophobicity of P1, P4, and P9 pockets are shown in Fig. 1b, illustrating the large hydrophobic P4 pocket that is specific to DRB1*1501.

Docking of human metabolites as MHC loading enhancers for DRB1 alleles

We provide energy score, van der Waals component, electrostatic component, polar solvation component and apolar solvation component, molecular weight, partition coefficient (xlogP), molecular charge, heavy atoms, polar contacts, nonpolar contacts, and number of conformations for the docked metabolites binding with DRB1*15:01 and DRB1*15:03 in Supplementary Tables S1, S2, and S3.

We selected the top 200 significant metabolites from the docking results in the P4 pocket for both DRB1*15:01 and DRB1*15:03 and ranked them on the basis of their energy score. We found that 78 and 81 metabolites bind individually with DRB1*15:01 (Supplementary Table S1) and DRB1*15:03 (Supplementary Table S2), respectively, whereas 79 metabolites bind commonly to both DRB1*15:01 and DRB1*15:03 (Supplementary Table S3). We observed overall higher energy scores for the binding of human metabolites with DRB1*15:01 than DRB1*15:03 (Fig. 2), suggesting generally stronger binding of metabolites to DRB1*15:01.

Determination of optimal pockets of the DRB1*15 alleles for binding to human metabolites

Structural comparison of the docked DRB1*15:01 and DRB1*15:03 metabolite complexes with DRB1*15:01 and DRB1*15:03 MBP complexes revealed that PhenylalanineMBP92 allows the binding of metabolites in the P4 pocket of DRB1*15:01 but ValineMBP89 in the P1 pockets does not (Table 1). We observed differences in the energy score even for the binding of the same metabolites in the P4 pockets of DRB1*15:01 and DRB1*15:03 (Table 1) or different metabolites in the P4 pockets of DRB1*15:01 and DRB1*15:03 (Table 1).

Mapping human metabolites on the MS phenotype

We performed data mining using metabolites we identified that bind either DRB1*15:01 or both DRB1*15:01 and DRB1*15:03 as input in ingenuity pathway analysis (IPA) to screen for human metabolites that have been previously associated with MS. We identified several human metabolites previously associated with MS and/or neurodegeneration (Table 2).
Fig. 2: Heat map of binding energy scores of human metabolites with DRB1*15:01 or both DRB1*15:01 and DRB1*15:03. The 157 metabolites are listed on the y-axis and binding energy scores of metabolites with DRB1*15:01 or both DRB1*15:01 and DRB1*15:03 are shown on the x-axis. Each cell represents the binding energy score in kcal/mol on a heatmap color scale.
Discussion

HLA-DRB1 is the primary susceptibility locus in MS [1]. It is likely that the crucial role of this molecule is mediated by presentation of peptide antigens to myelin-reactive T cells, which are subsequently responsible for the stimulation, proliferation, and production of pathogenic cytokines. Recent studies have identified small molecules that bind to the peptide-binding groove of DR1 and DR3, and act as a catalyst to enhance peptide loading [9, 15, 16]. We used an in silico molecular docking algorithm to virtually screen the complete HMDB for docking to the MS predisposing DRB1*15:01 allele and compared the results to the closely related yet non-susceptible DRB1*15:03 allele; and evaluated the potential impact on binding to MBP peptide. Although small molecules have been evaluated with respect to antigen presentation by the HLA class II molecule (DR1 and DR3), the present study is the first to screen a large database of naturally occurring human metabolites using a molecular docking algorithm to predict interaction with the peptide-binding groove of HLA-DR2.

We found considerably higher energy scores for the binding of human metabolites with DRB1*15:01 than DRB1*15:03 (Fig. 2), suggesting stronger binding of metabolites with DRB1*15:01 (DR2) the primary MS associated allele. Further, our results confirm that phenylalanine $^{MBP92}$ allows the binding of metabolites in the P4 pocket of DRB1*15:01 (DR2) but valine $^{MBP99}$ prohibits metabolite binding in P1 pockets. The most commonly accepted explanation for this is that the predominantly hydrophobic P4 pocket (Fig. 1b), occupied by a phenylalanine $^{MBP92}$ of the MBP peptide, is the critical feature of the DR2 peptide-binding site and the existence of an alanine residue at the polymorphic DR$\beta$71 position generates the essential space for the binding of an aromatic side chain in the P4 pocket [13, 22]. The presence of two aromatic residues of the P4 pocket of DR2, Phe26 Phe and Phe78 Tyr, of which β26 is polymorphic, also facilitate the binding of aromatic side chains by the P4 pocket [14]. The occurrence of alanine at DR$\beta$71 is exceptional for DBR1 alleles and has been observed distinctively for some DR2 alleles including DRB1*15:01 and DRB1*15:03; other DRB1 alleles encode arginine, glutamic acid, lysine, threonine, glycine and serine at this position (IMGT/HLA-database version 3.27.0). DRB1*15:01 is the most frequently observed DR2 allele in individuals with European ancestry, who also have augmented risk for the development of MS. Large-scale GWAS confirms that the main susceptibility signal maps to the DRB1 gene in the HLA class II segment, and elucidates up to 10.5% of the genetic variance causal for MS susceptibility [23]. DRB1*15:01 has the greatest effect with an average odds ratios of 3.08, and all additional DRB1 associations emerge to account for less than 2% of the residual variance [23]. There is significant variation in the distribution of DRB1*15 haplotypes across different worldwide populations, and these tend to correlate with the incidence of MS [24]. It would be interesting to examine whether metabolomic profiles differ similarly between world populations, and consider their relationship to HLA class II frequency distributions.

Previous studies reported the binding of adamantyl derivatives and plant alkaloid Cepharanthine in the P1 pockets of DRB1*01:01 (DR1) [9], and DRB1*03:01 (DR3) [16]. HLA-DR1 (HLA-DRA, DRB1*0101) and DR4 (HLA-DRA, DRB1*0401) have large P1 pockets due to the presence of Gly at DR$\beta$86 position, while the arginine residue at position DR$\beta$74 is responsible for the larger P1 size for DR3 (HLA-DRA, DRB1*0301); thus, the P1 pockets may accommodate metabolite and peptide binding in DR1, DR3 and DR4. Meanwhile, the P1 pocket of DR2 (HLA-DRB, DRB1*1501) is too small (Val at DR$\beta$86) to accommodate small molecules such as human metabolites, whereas the greater size of the P4 pocket due to alanine at DR$\beta$71 allows the binding of metabolites in the P4 pockets of both DRB1*15:01 and DRB1*15:03. The polymorphism at position 71 in DR2 (alanine) emerges to be most significant in terms of producing the accessible space for the P4 aromatic side chain (Phe) of the MBP peptide [14]. The HLA-DR2 P4 pocket has an inclination for aromatic as well as aliphatic residues [14]. The P4 pocket of HLA-DR2 binding is detected with analogues of the MBP peptide, which contains a substitution by histidine, arginine, lysine,

| Haplotype     | Human metabolites                  | Energy score | DRB1 binding pocket | MBP residue and binding position |
|---------------|-----------------------------------|--------------|---------------------|----------------------------------|
|               |                                   |              | P4 pocket | P1 pocket | Phenylalanine $^{MBP92}$ | Valine $^{MBP99}$ |
| DRB1*15:01    | Myo-inositol hexakisphosphate     | −85.92       | Yes       | No        | Yes                 | No             |
| DRB1*15:03    | Myo-inositol hexakisphosphate     | −55.03       | Yes       | No        | Yes                 | No             |
| DRB1*15:01    | Pyridinoline                      | −60.69       | Yes       | No        | Yes                 | No             |
| DRB1*15:03    | Mesobilirubinogen                 | −45.32       | Yes       | No        | Yes                 | No             |
glutamine or asparagine at P4; substitution by aspartic acid is not tolerated [13]. Human metabolites appear to bind more strongly with P4 pockets of DRB1*15:01 than DRB1*15:03 due to the presence of a tyrosine residue at DRβ30 position, the only amino acid difference between the two alleles. The tyrosine side chains have a neutral charge, in contrast to the positively charged side chains of histidine at position 30 of DRB1*15:03, which may facilitate the stronger binding of metabolites.

Numerous studies have reported the association of the P4 pocket, the most polymorphic pocket of the HLA-DR binding site, in augmented risk to other autoimmune diseases [25–28], which further strengthens our observation. For example, DRB1*04:02, which has been associated with predisposition to pemphigus vulgaris [26, 27], contrasts from the rheumatoid arthritis linked DRB1*04:04 allele at only three positions: DRβ67, —DRβ70, and —DRβ71 [25, 28]. In DRB1*04:02, DRβ70, and —71 of the P4 pocket carry negative charges due to residues aspartic acid and glutamic acid, respectively [26, 27]. In contrast, glutamine and lysine/arginine residues are present respectively at DRβ70 and —71 positions in the rheumatoid arthritis correlated DRB1*04:01 and DRB1*04:04 alleles.

The binding of human metabolites in the P4 pockets of DRB1*15:01 potentially have profound effects on T cell proliferation and antigen presentation. One possible scenario for altered T-cell recognition related to bound metabolites in DRB1*15:01 is shown in Fig. 3a. In this model the bound metabolite causes a register shift in the 9-mer segment of MBP peptide from the normal bound position (Fig. 3b). HLA class II molecules consists of an open binding groove, which permits higher flexibility and bound peptides may be up to 14–20 amino acids [29]. The peptide-binding segment of the HLA class II molecule binds specific 9mer residues, which is recognized as the peptide binding register [29]. This 9mer segment is located within the binding groove of the HLA class II molecule, while the residues of the remaining peptide are present exterior to the groove. As the peptide occupies the binding groove, the anchor residues at locations P1 or P4 of the peptide register act as primary HLA contacts, interacting with residues in the pockets of the HLA molecule. The binding pockets differ among HLA molecules, which are encoded by distinctive alleles and are accountable for binding the peptide within the groove. The pocket residues are mainly obscured deep inside the binding groove. In a register shift, the MBP peptide and human metabolite complex may transiently occupy a pocket in the HLA peptide-binding groove and shift forward the next residues, sliding along the groove until the trimolecular complex of HLA-Peptide-Metabolite

| Phenotypes                                      | Human metabolites                                                                 | Pockets |
|------------------------------------------------|-----------------------------------------------------------------------------------|---------|
| Multiple Sclerosis                             | L-tryptophan, glutamic acid, D-sphingosine, sphingosine-1-phosphate, Cysteinyl-Glycine | P4      |
| Neurodegeneration                              | D-sphingosine, NAD+, sphingosyl beta-glucoside                                    | P4      |
| Progressive motor neuropathy                   | D-sphingosine, escitalopram, glutathione, L-tryptophan, loperamide, sphingosine-1-phosphate | P4      |
| Hereditary central nervous system demyelinating disease | glutathione, psychosine                                                          | P4      |
| Cell death of central nervous system cells     | dityrosine, glutathione, NAD+                                                    | P4      |
| Cell death of neuroglia                        | glutathione, NAD+                                                                | P4      |
| Apoptosis of motor neurons                     | glutathione                                                                       | P4      |
| Apoptosis of cerebellar macroneurons           | sphingosine-1-phosphate                                                          | P4      |
| Astrogligenesis of neural stem cells           | AICAR                                                                             | P4      |
| Discharge of axons                             | AICAR                                                                             | P4      |
| Invasion of nervous tissue cell lines          | S-adenosylmethionine                                                             | P4      |
| Mitogenesis of neurons                         | sphingosine-1-phosphate                                                          | P4      |
| Migration of astrocytes                        | sphingosine-1-phosphate                                                          | P4      |
| Conduction of nerves                           | glutathione                                                                       | P4      |
| Neuritogenesis of neurons                      | sphingosine-1-phosphate                                                          | P4      |
| Prevention of Multiple Sclerosis               | Escitalopram                                                                     | P4      |
| Proliferation of neuronal cells                | AICAR, hesperidin, nicotinamide adenine dinucleotide phosphate, sphingosine-1-phosphate, sphingosyl beta-glucoside | P4      |
| Morphology and Synapse Formation of Cerebellar Purkinje Cells | Myo-Inositol hexakisphosphat                                                       | P4      |
that a transient occupation of certain human metabolites in the P4 pocket HLA-DRB1*15:01 may enhance peptide presentation and potentially resulting in a change of the balance of Treg vs. conventional T-cell phenotype.

Our analysis of HLA interactions with human metabolites could pave the way for similar investigations in clinical allo-transplantations and other autoimmune diseases, such as rheumatoid arthritis (RA), celiac disease (CD), systemic lupus erythmatosis, ankylosing spondylitis, dermatomyositis, and type I diabetes mellitus. Our approach might be helpful to weight or group together various HLA genes and alleles that are involved in susceptibility to different diseases. The findings of this study could also be useful in resolving the conflicts regarding the association of HLA polymorphisms with RA susceptibility association with anti-cyclic citrullinated peptide (ACPP) autoantibody. In some studies, ACCP antibodies are associated with the presence of HLA class II alleles; especially the DRB1 shared epitope alleles in European-ancestry RA patients [37, 38], whereas subsequent studies have failed to replicate this association [39]. The identification of specific pathogenic metabolites in RA could help to explain some of these discrepancies. Similarly, Lee and colleagues have recently suggested that the genetic contribution to prognosis in Crohn’s disease is largely independent of the contribution to disease susceptibility [40]. The detection of highly immunogenic metabolite-HLA loads in individuals could provide insight regarding prognosis as well as new therapeutic opportunities. Finally, donor-recipient matching for immunogenic metabolites that occupy specific pockets in HLA molecules could be a strategy employed in HLA mismatched donor-recipient pairs to improved transplant outcome and avoid graft rejections, providing a potential

Fig. 3 a: Hypothetical Model for the binding of human metabolites with DRB1*15:01 and DRB1*15:03 protein structures. In this model, the 9-mer segments of MBP are present deep inside the binding groove of DRB1*15:01 and DRB1*15:03, and the remaining residues are present outside the peptide-binding groove. b: Three different potential registers for Human MBP85–99 binding to DRB1*15:01 and DRB1*15:03. The 9mer residues from position 89 to 97 of MBP85–99 peptide plays a crucial role in metabolite binding as well as T cell recognition and responsiveness. In the 1ST register, the red color represents the HLA contact residues of pockets P1, P4, P6 and P9, and the black color shows the TCR contact residues of pockets P2, P3, P5, and P7. The TCR contact residues allow the interaction of the DRB1-metabolite-MBP complex with TCR for presentation to CD4+ T-cells. When a metabolite occupies the binding groove of the DRB1 molecule it may shift the P1 residue and all subsequent residues until the DRB1-metabolite complex is stabilized. Alternatively, the presence of one of the high scoring metabolites in the P4 pocket of DRB1*15:01 may enhance the risk of MS by augmenting antigen loading. This possibility is supported by previous reports suggesting that a transient occupation of HLA-DR pocket by an organic compound stabilizes the peptide-receptive conformation allowing rapid antigen loading [9].

Interestingly, our IPA analysis revealed that specific metabolites that bind the P4 pocket of DRB1*15:01, L-tryptophan, glutamic acid, D-sphingosine, sphingosine-1-phosphate, cysteinyl-glycine, and NAD+ have been previously associated with MS and neurodegeneration (Table 2). Similarly, the metabolomic signatures of tryptophan and glutamic acid have been suggested to be associated with disease severity in multiple sclerosis [30]. The sphingolipids are a class of biologically active lipids; sphingosine 1-phosphate receptors 1 and 3 have been shown to be upregulated in multiple sclerosis lesions [31]. A homocysteinemia regulator cysteinylglycine has been linked with physiological and pathological conditions in cerebrovascular and multiple sclerosis patients [32]. More recently, human metabolites NAD(+), myo-inositol hexakisphosphate and glutathione have been correlated, respectively with neurodegeneration [33], regulation of morphology and synapse formation of cerebellar purkinje cells [34], and metabolism impairment in MS [35].

Recently, a study conducted in autoimmune Goodpasture disease has shown that HLA-DR15-3135-145 tetramer+ T cells in HLA-DR15 transgenic mice display a conventional T-cell phenotype, which secretes pro-inflammatory cytokines and may be associated with augmented disease risk [36]. Meanwhile, HLADR1-3135-145 tetramer+ T cells in HLA-DR1 and HLA-DR15/DR1 transgenic mice are mainly CD4+Foxp3+ regulatory T cells (Treg cells) expressing tolerogenic cytokines, which is linked with disease protection [36]. This study provides a mechanistic model for the relative abundance of self-epitope specific Treg cells, which leads to either protection or risk of autoimmunity. The same concept may be applicable to the transient occupation of certain human metabolites in the P4 pocket HLA-DRB1*15:01, altering peptide presentation and potentially resulting in a change of the balance of Treg vs. conventional T-cell phenotype.

Our approach might be helpful to weight or group together various HLA genes and alleles that are involved in susceptibility to different diseases. The findings of this study could also be useful in resolving the conflicts regarding the association of HLA polymorphisms with RA susceptibility association with anti-cyclic citrullinated peptide (ACPP) autoantibody. In some studies, ACCP antibodies are associated with the presence of HLA class II alleles; especially the DRB1 shared epitope alleles in European-ancestry RA patients [37, 38], whereas subsequent studies have failed to replicate this association [39]. The identification of specific pathogenic metabolites in RA could help to explain some of these discrepancies. Similarly, Lee and colleagues have recently suggested that the genetic contribution to prognosis in Crohn’s disease is largely independent of the contribution to disease susceptibility [40]. The detection of highly immunogenic metabolite-HLA loads in individuals could provide insight regarding prognosis as well as new therapeutic opportunities. Finally, donor-recipient matching for immunogenic metabolites that occupy specific pockets in HLA molecules could be a strategy employed in HLA mismatched donor-recipient pairs to improved transplant outcome and avoid graft rejections, providing a potential
applicability of our study in terms of clinical allograft transplantation. Although finding the exact match of donor-recipient pairs for immunogenic metabolites could be challenging, donor-recipient pairs with low metabolite mismatch loads or without highly immunogenic metabolite mismatches could be considered.

In summary, we have shown that naturally occurring human metabolites bind to DRB1*15:01, the primary MS susceptible allele in most populations, and may potentially impact MS risk. An improved understanding of the kinetics of HLA binding of human metabolites may allow development of small molecule based immunomodulatory drugs antagonistic to metabolites that enhance antigen loading or alter presentation. Identification of our high scoring metabolites in metabolomics studies in MS may also be informative in identification of suitable MS biomarkers. A limitation of our study is that we did not test the docked DRB1-Metabolite complex in the laboratory. This is an outstanding point of uncertainty and will need to be addressed in future studies to fully elucidate the role of the metabolites identified in silico in T-cell proliferation and antigen presentation. A structure-based strategy could determine structural features at the peptide/HLA interface and recognize molecules that may stimulate or suppress TCR signaling in response to particular peptide antigens. The present study suggests that the combination of structure directed virtual screening and the notion that human metabolites targeted to certain HLA pockets may be immunomodulatory contains extensive applicability to adaptive immunity and autoimmune conditions.

**Methods**

**Selection and preparation of DRB1 structure**

The X-ray crystal structure for the DRB1*15:01 complex with bound peptide from human MBP with a resolution of 2.6 Å was obtained from the protein data bank under accession number 1BX2 [14]. We prepared the target (DRB1*15:01 and DRB1*15:03) and binding site residue with N-acetyl-D-glucosamine (NAG) for docking with human metabolites using the UCSF Chimera software package (version 1.11.2) [41]. DRB1*15:01 exon-2 sequence differs from DRB1*15:03 at amino acid position 30 from tyrosine (Tyr) to histidine (His). The in silico DRB1*15:03 structure was obtained from the known structure of DRB1*15:01 by mutating the residue Tyr<sup>30</sup> > His [30].

**Virtual screening and molecular docking of DRB1*15:01, DRB1*15:03, and human metabolites**

The HMDB version 3.6 consists of 41,993 metabolites derived from various sources such as exogenous, endogenous, food, microbial, toxic, and drug metabolites. We screened the complete HMDB library (version 3.6) for docking with DRB1*15:01 and DRB1*15:03. The schematic representation of the overall study design and molecular docking pipeline is shown in Fig. 4.

Virtual screening was performed using the UCSF Chimera version 1.11.2 [41], Clipper version 1.6.0 [42], and

---

**Fig. 4** Schematic representation of overall study design (left panel) and docking pipeline (right panel)

---
Dockblaster version 1.6.0 softwares [43] as previously described [41–43]. UCSF Chimera 1.11.2 was used for the selection, visualization and preparation of target and ligand and the visualization and analysis of DRB1-metabolite complexes. Virtual Screening for DRB1*15:01 and DRB1*15:03 was conducted on a structure obtained in molecular dynamics simulations of the complex with a bound peptide from MBP. Clipper 1.6.0 was used for the identification and analysis of different peptide binding pockets in DRB1*15:01 and DRB1*15:03.

Subsequently, we conducted independent screenings to probe each pocket from P1 to P9 of both DRB1*15:01 and DRB1*15:03 with HMDB (version 3.6) using Dockblaster 1.6.0 [43], including the calculation of energy score, van der Waals component, electrostatic component, polar solvation component, molecular charge, heavy atoms, polar contacts (3.3 A), nonpolar contacts (3.3–4.5 A), and the number of conformations for docked DRB1 and metabolite complexes. The Dockblaster 1.6.0 results were used to select the top ranked metabolites and investigate the distribution of their scores in assessing the point beyond which their scores become indistinguishable. On this basis we selected metabolites from the docking results in each pocket and ranked them on the basis of their energy score. We used the following criteria to compute the energy score of docked metabolite with DRB1*15:03 and DRB1*15:03 in kcal/mol:

\[
\text{Energy Score} = \text{Energy of Electrostatic interaction (ES)} + \text{Energy of van der Waals interaction (VdW)} + \text{Desolvation energy}
\]

Desolvation energy was calculated by combining the energy of polar solvation component and apolar solvation component.

**Sequence alignment**

In order to understand whether the human metabolites are binding in the P1 or P4 pockets, we performed sequence alignment of both DRB1*15:01-metabolite complex and DRB1*15:03-metabolite complex individually with the crystal structure of DRB1*15:01 and DRB1*15:03 with UCSF Chimera version 1.11.2 [41]. This comparison of each protein structure prior to and after the metabolite binding through sequence alignments allowed assessment of changes in the dynamics of DRB1*15:01 and DRB1*15:03 protein structures attributable to metabolite binding.

**Acknowledgements** This study was supported by grants from the National Institutes of Health (U19NS095774, R01AI128775). The authors wish to thank Professor Jorge Oksenberg for his support and valuable comments on this work.

**Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no conflict of interest.

**References**

1. Hollenbach JA, Ok senberg JR. The immunogenetics of multiple sclerosis: A comprehensive review. J Autoimmun. 2015;64:13–25.

2. Maiers M, Gragert L, Klitz W. High-resolution HLA alleles and haplotypes in the United States population. Hum Immunol. 2007;68:779–88.

3. Ji N, Somanaboeina A, Dixit A, Kawamura K, Hayward NJ, Self C, et al. Small molecule inhibitor of antigen binding and presentation by HLA-DR2b as a therapeutic strategy for the treatment of multiple sclerosis. J Immunol. 2013;191:5074–84.

4. Krosgaard M, Wucherpfennig KW, C annella B, Hansen BE, Sveigaard A, Pyrdol J, et al. Visualization of myelin basic protein (MBP) T cell epitopes in multiple sclerosis lesions using a monoclonal antibody specific for the human histocompatibility leukocyte antigen (HLA)-DR2-MBP 85-99 complex. J Exp Med. 2000;191:395–412.

5. Vergelli M, Kalbus M, Rojo SC, Hemmer B, Kalbacher H, Tranquill L, et al. T cell response to myelin basic protein in the context of the multiple sclerosis-associated HLA-DR15 haplotype: peptide binding, immunodominance and effector functions of T cells. J Neuroimmunol. 1997;77:195–203.

6. Boes M, Ploegh HL. Translating cell biology in vitro to immunity in vivo. Nature. 2004;430:264–71.

7. Sloan VS, Cameron P, Porter G, Gammon M, Amaya M, Mellins E, et al. Mediation by HLA-DM of dissociation of peptides from HLA-DR. Nature. 1995;375:802–6.

8. Watts C. Antigen processing in the endocytic compartment. Curr Opin Immunol. 2001;13:26–31.

9. Hopner S, Dickhaut K, Hofstatter M, Kramer H, Ruckerl D, Soderhall JA, et al. Small organic compounds enhance antigen loading of class II major histocompatibility complex proteins by targeting the polymorphic P1 pocket. J Biol Chem. 2006;281:38535–42.

10. Ota K, Matsui M, Milford EL, Mackin GA, Weiner HL, Hafler DA. T-cell recognition of an immunodominant myelin basic protein epitope in multiple sclerosis. Nature. 1990;346:183–7.

11. Pette M, Fujita K, Wilkinson D, Altmann DM, Trowsdale J, Giegerich G, et al. Myelin autoreactivity in multiple sclerosis: recognition of myelin basic protein in the context of HLA-DR2 products by T lymphocytes of multiple-sclerosis patients and healthy donors. Proc Natl Acad Sci USA. 1990;87:7968–72.

12. Valli A, Sette A, Kappos L, Oseroff C, Sidney J, Miescher G, et al. Binding of myelin basic protein peptides to human histocompatibility leukocyte antigen class II molecules and their recognition by T cells from multiple sclerosis patients. J Clin Invest. 1993;91:616–28.

13. Wucherpfennig KW, Sette A, Southwood S, Oseroff C, Matsui M, Strominger JL, et al. Structural requirements for binding of an immunodominant myelin basic protein peptide to DR2 isotypes and for its recognition by human T cell clones. J Exp Med. 1994;179:279–90.

14. Smith KJ, Pyrdol J, Gauthier L, Wiley DC, Wucherpfennig KW. Crystal structure of HLA-DR2 (DRA*0101, DRB1*1501) complexed with a peptide from human myelin basic protein. J Exp Med. 1998;18:1511–20.

15. Michels AW, Ostrov DA, Zhang L, Nakayama M, Fuse M, McDaniel K, et al. Structure-based selection of small molecules to
alter allele-specific MHC class II antigen presentation. J Immunol. 2011;187:5921–30.

16. Li CW, Mencconi F, Osman R, Mezei M, Jacobson EM, Concepcion E, et al. Identifying a Molecular Blocking Antigen Presentation in Autoimmune Thyroiditis. J Biol Chem. 2016;291:4079–90.

17. Illing PT, Vivian JP, Dudek NL, Kostenko L, Chen Z, Bharadwaj M, et al. Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. Nature. 2012;486:554–8.

18. Pavlos R, Mallal S, Ostrov D, Buus S, Metushi I, Peters B, et al. T cell-mediated hypersensitivity reactions to drugs. Annu Rev Med. 2015;66:439–54.

19. Madden DR. The three-dimensional structure of peptide-MHC complexes. Annu Rev Immunol. 1995;13:587–622.

20. Li Y, Li H, Martin R, Mariuzza RA. Structural basis for the binding of an immunodominant peptide from myelin basic protein in different registers by two HLA-DR2 proteins. J Mol Biol. 2000;304:177–88.

21. Donia MS, Fischbach MA. HUMAN MICROBIOTA. Small molecules from the human microbiota. Science. 2015;349:1254766.

22. Vogt AB, Kropshofer H, Kalbacher H, Kalbus M, Rammensee HG, Coligan JE, et al. Ligand motifs of HLA-DRB5*0101 and DRB1*1501 molecules delineated from self-peptides. J Immunol. 1994;153:1665–73.

23. George MF, Briggs FB, Shao X, Gianfrancesco MA, Kockum I, Harbo HF, et al. Multiple sclerosis risk loci and disease severity in 7,125 individuals from 10 studies. Neurol Genet. 2016;2:e87.

24. Risco J, Maldonado H, Luna L, Osada J, Ruiz P, Juarez A, et al. Latitudinal prevalence gradient of multiple sclerosis in Latin America. Mult Scler. 2011;17:1055–9.

25. Hammer J, Gallazzi F, Bono E, Karr RW, Guenot J, Valsasnini P, et al. Peptide binding specificity of HLA-DR4 molecules: correlation with rheumatoid arthritis association. J Exp Med. 1995;181:1847–55.

26. Scharf SJ, Friedmann A, Brautbar C, Szafer F, Steinman L, Horn G, et al. HLA class II allelic variation and susceptibility to pemphigus vulgaris. Proc Natl Acad Sci USA. 1988;85:3504–8.

27. Ahmed AR, Yunis EJ, Khatri K, Wagner R, Notani G, Awdeh Z, et al. Major histocompatibility complex haplotype studies in Ashkenazi Jewish patients with pemphigus vulgaris. Proc Natl Acad Sci USA. 1990;87:7658–62.

28. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum. 1987;30:1205–13.

29. Mohan JF, Unanue ER. Unconventional recognition of peptides by T cells and the implications for autoimmunity. Nat Rev Immunol. 2012;12:721–8.

30. Villoslada P, Alonso C, Agirrezabala I, Kotelnikova E, Zubizarreta I, Pulido-Valderivas I, et al. Metabolomic signatures associated with disease severity in multiple sclerosis. Neurol Neuroimmunol Neuroinflamm. 2017;4:e321.

31. Van Doorn R, Van Horssen S, Verzijl D, Witte M, Ronken E, Van Het Hof B, et al. Sphingosine 1-phosphate receptor 1 and 3 are upregulated in multiple sclerosis lesions. Glia. 2010;58:1465–76.

32. Ulivelli M, Priora R, Di Giuseppe D, Coppo L, Summa D, Margaritis A, et al. Homocysteinemia control by cysteine in cerebral vascular patients after methionine loading test: evidences in physiological and pathological conditions in cerebro-vascular and multiple sclerosis patients. Amino Acids. 2016;48:1477–89.

33. Verdin E. NAD(+/-) in aging, metabolism, and neurodegeneration. Science. 2015;350:1208–13.

34. Fu C, Xu J, Li RJ, Crawford JA, Khan AB, Ma TM, et al. Inositol hexakisphosphate kinase-3 regulates the morphology and synapse formation of cerebellar purkinje cells via spectrin/adducin. J Neurosci. 2015;35:11056–67.

35. Carvalho AN, Lim JL, Nijland PG, Witte ME, Van Horssen J. Glutathione in multiple sclerosis: more than just an antioxidant? Mult Scler. 2014;20:1425–31.

36. Ooi JD, Petersen J, Tan YH, Huynh M, Willett ZJ, Ramarathnam SH, et al. Dominant protection from HLA-linked autoimmunity by antigen-specific regulatory T cells. Nature. 2017;545:243–7.

37. van Gaalen FA, van Aken J, Huizinga TW, Schreuder GM, Breedveld FC, Zanelli E, et al. Association between HLA class II genes and autoantibodies to cyclic citrullinated peptides (CCPs) influences the severity of rheumatoid arthritis. Arthritis Rheum. 2004;50:2113–21.

38. Irigoyen P, Lee AT, Wener MH, Li W, Kern M, Baltiwalla F, et al. Regulation of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: contrasting effects of HLA-DR3 and the shared epitope alleles. Arthritis Rheum. 2005;52:3813–8.

39. Xue Y, Zhang J, Chen YM, Guan M, Zheng SG, Zou HJ. The HLA-DRB1 shared epitope is not associated with antibodies against cyclic citrullinated peptide in Chinese patients with rheumatoid arthritis. Scand J Rheumatol. 2008;37:183–7.

40. Lee JC, Biasci D, Roberts R, Gearry RB, Mansfield JC, Ahmad T, et al. Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn’s disease. Nat Genet. 2017;49:262–8.

41. Pettersen EF, Goddard TD, Huang CS, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605–12.

42. Coleman RG, Sharp KA. Protein pockets: inventory, shape, and comparison. J Chem Inf Model. 2010;50:589–603.

43. Irwin JJ, Shoichet BK, Mysinger MM, Huang N, Colizzi F, Wassam P, et al. Automated docking screens: a feasibility study. J Med Chem. 2009;52:5712–20.