A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis

Stephen W. Scally,¹ Jan Petersen,¹ Soi Cheng Law,² Nadine L. Dudek,¹ Hendrik J. Nei,¹ Khai Lee Loh,¹ Lakshmi C. Wijeyewickrema,¹ Sidonia B.G. Eckle,³ Jurgen van Heemst,⁴ Robert N. Pike,¹ James McCluskey,³ Rene E. Toes,⁴ Nicole L. La Gruta,³ Anthony W. Purcell,¹ Hugh H. Reid,¹ Ranjeny Thomas,² and Jamie Rossjohn¹

Rheumatoid arthritis (RA) is strongly associated with the human leukocyte antigen (HLA)-DRB1 locus that possesses the shared susceptibility epitope (SE) and the citrullination of self-antigens. We show how citrullinated aggrecan and vimentin epitopes bind to HLA-DRB1*04:01/04. Citrulline was accommodated within the electropositive P4 pocket of HLA-DRB1*04:01/04, whereas the electronegative P4 pocket of the RA-resistant HLA-DRB1*04:02 allomorph interacted with arginine or citrulline-containing epitopes. Peptide elution studies revealed P4 arginine–containing peptides from HLA-DRB1*04:02, but not from HLA-DRB1*04:01/04. Citrullination altered protease susceptibility of vimentin, thereby generating self-epitopes that are presented to T cells in HLA-DRB1*04:01+ individuals. Using HLA-II tetramers, we observed citrullinated vimentin- and aggrecan-specific CD4+ T cells in the peripheral blood of HLA-DRB1*04:01+ RA-affected and healthy individuals. In RA patients, autoreactive T cell numbers correlated with disease activity and were deficient in regulatory T cells relative to healthy individuals. These findings reshape our understanding of the association between citrullination, the HLA-DRB1 locus, and T cell autoreactivity in RA.

© 2013 Scally et al. This article is distributed under the terms of an Attribution–Noncommercial–Share Alike–No Mirror Sites license for the first six months after the publication date (see http://www.rupress.org/terms). After six months it is available under a Creative Commons License (Attribution–Noncommercial–Share Alike 3.0 Unported license, as described at http://creativecommons.org/licenses/by-nc-sa/3.0/).

Abbreviations used: ACPA, anticitrullinated protein antibody; FDR, false discovery rate; FMO, fluorescence minus one; HA, hemagglutinin; HLA, human leukocyte antigen; MFI, mean fluorescence staining intensity; OR, odds ratio; PAD, peptidyl arginine deiminases; PE, phycoerythrin; pHLA, peptide-HLA; RA, rheumatoid arthritis; SE, shared susceptibility epitope; T reg cell, regulatory T cell.

A.W. Purcell, H.H. Reid, R. Thomas, and J. Rossjohn contributed equally to this paper.
monomorphic. Notwithstanding some HLA disease associations, little is known about the nature of the HLA-bound self-peptides that are involved in autoimmunity, limiting development of specific immune intervention strategies aimed to inhibit or prevent such deleterious immune responses. Nevertheless, rheumatoid arthritis (RA) is arguably one of the best-described systems for understanding the genetic association between HLA-II alleles, autoimmunity, and self-peptide presentation (Raychaudhuri et al., 2012; Viatte et al., 2013).

RA is a systemic autoimmune disease, affecting ~1% of the population (Helmick et al., 2008). RA is characterized by inflammation of synovial tissues in the joints, pannus formation, and erosion of the bones (Klærskog et al., 2009). Like most human autoimmune diseases, multiple genes contribute to RA susceptibility and severity (Viatte et al., 2013). The most comprehensive genetic association exists with HLA-DRB1 genes and in particular the HLA-DR4 alleles. Specifically, the association has been mapped to a highly polymorphic N-terminal region of the HLA DRβ chain around positions 70–74 (Viatte et al., 2013). This region encodes a conserved positively charged residue at position 71 that is thought to dictate the nature of the amino acid that is accommodated in the P4 pocket of the antigen-binding groove (Hammer et al., 1995). Alleles having this shared conserved region of the DRβ 70–74 region are termed to have a shared susceptibility epitope (SE; Gregersen et al., 1987) and include the commonly occurring HLA DRB1*04:01, *04:04, and *01:01 molecules. Recently, a large haplotype association study involving >5,000 seropositive RA patients and 15,000 controls has attributed most of the DR-associated risk to positions 11, 13, 71, and 74 of the HLA-DRB1 polypeptide chain encoded by SE alleles (Raychaudhuri et al., 2012), strongly suggesting that this allele permits binding and presentation of autoantigenic peptides (Raychaudhuri et al., 2012), strongly suggesting that this alloepitope binds and presentation of autoantigenic peptides. In addition, HLA-DRB1*04+ individuals have accelerated CD4+ T cell telomere erosion and immunosenescence commencing early in life, relative to HLA-DRB1*04- individuals, regardless of the development of RA (Schönland et al., 2003). However, the molecular basis for the RA association with the SE remains unclear.

Citrullination, the conversion of arginine to citrulline, is a physiological process catalyzed by peptidyl arginine deiminases (PAD; Vossenaar and van Venrooij, 2004). This process is increased during inflammation, stress, and apoptosis, and expands the repertoire of presented epitopes after protein immunization (Klærskog et al., 2008). Citrullinated proteins and PAD (arising from inflammatory cells) are found in RA patient synovium (Vossenaar et al., 2004; Foulquier et al., 2007) and in RA- and non–RA-associated pneumonia (Bongartz et al., 2007). Moreover, expression of citrullinated proteins is up-regulated in the lung epithelial cells of healthy smokers relative to nonsmokers (Makrygiannakis et al., 2008). Consistent with this observation, smoking increases the risk of developing anticitrullinated protein antibody (ACPA)-positive RA, particularly in SE+ individuals (Padyukov et al., 2004; Klærskog et al., 2008). Numerous citrullinated autoantigens, of which most are ubiquitous proteins, have been identified in RA (Hill et al., 2003, 2008; Vossenaar et al., 2004; Vossenaar and van Venrooij, 2004; Klærskog et al., 2008; Law et al., 2012), with some showing cross reactivity with microbial antigens (Lundberg et al., 2008). Indeed, autoantibodies specific for citrullinated antigens are found in the serum of RA patients and are highly specific to the disease (van Gaalen et al., 2004; Klærskog et al., 2008; Klærskog et al., 2009). Over the last decade, this observation has led to a rapid clinical translation and adoption of ACPA reactivity as an important diagnostic tool, including the prediction of more erosive outcomes in RA (Klærskog et al., 2009; Klærskog et al., 2008; van Gaalen et al., 2004). ACPA may directly influence joint inflammation and erosion through local binding of citrullinated proteins (Kuhn et al., 2006; Harre et al., 2012). Moreover, HLA-DRB1 susceptibility alleles are strongly associated with ACPA-positive RA, strengthening the conclusion that the HLA-SE molecules restrict antigen presentation of citrullinated autoantigens (Huizinga et al., 2005; Klærskog et al., 2008, 2009; van Gaalen et al., 2004). However, despite the clinical utility of elucidating autoantibody responses toward them, the precise role of citrullinated antigens in the initiation and/or progression of RA has remained elusive.

RESULTS

Structural basis of citrullinated epitope presentation

Several citrullinated (cit) epitopes, including vimentin59–71 (GVYATR/citSSAVR/citSSAVR/cit; Snir et al., 2011), vimentin66–78 (SARVR/citSSVGVR; Hill et al., 2003; Law et al., 2012), fibrinogen-α-1 (QDFTRN/citINLKKN5S; Hill et al., 2008; Law et al., 2012), and aggrecan94–103 (VVLVATEG/R/CitVRVSAYQDK; Law et al., 2012; von Delwig et al., 2010) are associated with ACPA+ RA and the SE-encoded HLA alleles. To establish the basis of citrullination-dependent binding to the SE-HLA allomorphs (Fig. 1, a and b), we determined the high resolution structures of HLA-DRB1*04:01 complexed to vimentin59–71 epitopes that were citrullinated at position 64 (vimentin–64Cit59–71), as well as at positions 64, 69, and 71 (vimentin–64–69–71Cit59–71); the vimentin66–78 epitope that was citrullinated at position 71 (vimentin–71Cit66–78); and the aggrecan94–103 epitope that was citrullinated at positions 93 and 95 (aggrecan–93–95Cit89–103). This provided a broad perspective of how citrullination of epitopes enables HLA-DRB1*04:01 binding (Fig. 1; Fig. 2, a–c; and Table 1). The citrullinated epitopes were located within the Ag-binding cleft of HLA-DRB1*04:01, and all four structures adopted a very similar conformation and were similar to previously determined HLA-DR4 structures that bound noncitrullinated antigens (Dessen et al., 1997; Fig. 1, c and d; and Fig. 2, a–c). The vimentin–71Cit66–78 epitope bound in a linear, extended manner with P1-Val, P4-Cit, P6-Ser, and P9-Gly occupying the P1, P4, P6, and P9 pockets of HLA-DRB1*04:01, respectively, whereas P2-Arg, P5-Ser, P7-Val, P8-Pro, and P11-Arg represented potential TCR contact sites (Fig. 1 c). The P4-Cit bent back upon itself and adopted a constrained U-shaped conformation in which its aliphatic moiety packed...
against Phe26β, Tyr78β, and His13β of HLA-DRB1*04:01 (Fig. 3 a). Of the residues within the SE, positions 72 and 73 pointed away from the P4 pocket, whereas position 74 was orientated toward the pocket, packing against Phe26β, yet did not contact the P4-Cit. The citrullinated head group formed a direct H-bond to Lys71βN, the latter of which was stabilized by a salt bridge to Asp28β, and an H-bond to P5-SerO (Fig. 3 a). A P4-Arg could not be accommodated within this P4 pocket, as Lys71β would electrostatically repel the positively charged guanidinium head group and, moreover, there is insufficient space surrounding the P4 pocket to enable Lys71β or P4-Arg to adopt differing conformations, which is consistent with the peptide elution data (discussed below). Although there were some sequence differences between the vimentin59-71 and vimentin66-78 epitopes, which related to differing anchor residue interactions at the P1 (Val → Tyr) and P9 pockets (Gly → Arg; Fig. 2 a), the P4-Cit residues adopted essentially identical interactions within the P4 pocket (Fig. 3 b). Moreover, in the vimentin-64-69-71Cit59-71 epitope, the P4-Cit adopted a very similar conformation to that observed in the vimentin-64Cit59-71 epitope (Fig. 2 b and Fig. 3 c). While the C-terminally located P11-Cit of vimentin-64-69-71Cit59-71 was solvent exposed and mobile, the P9-Cit occupied the P9 pocket of HLA-DRB1*04:01 (Fig. 2 b). Here, the P9 pocket seemed equally well suited to accommodate P9-Arg or P9-Cit, with Tyr37β H-bonding to both moieties (not depicted). The ready accommodation of P9-Arg/P9-Cit within the P9 pocket was consistent with the similar thermal stability values for HLA-DRB1*04:01–vimentin-64Cit59-71 and HLA-DRB1*04:01–vimentin-64-69-71Cit59-71 (Tm of 66.7°C and 69.1°C, respectively; Table 2). The structure of the HLA-DRB1*04:01–aggrecan-93-95Cit89-103 complex showed that the positioning of the P4-Cit, and the immediate environment of the P4 pocket, was very similar to that of the HLA-DRB1*04:01–vimentin complexes, despite the differing hydrogen bonding network with Lys71β (Fig. 2 c and Fig. 3 d). In the HLA-DRB1*04:01–aggrecan-93-95Cit89-103 complex, the P2-Cit was highly solvent exposed (Fig. 2 c and Fig. 3 d), suggesting that citrullination of this position could potentially impact on TCR recognition. Hence, the P4 pocket of HLA-DRB1*04:01 was highly suited to preferentially

Figure 1. HLA-DRB1*04:01 in complex with Vimentin-71Cit66-78.
(a) Polymorphic residues involved in susceptibility to RA. The peptide-binding groove of an HLA-DR molecule is shown in cartoon representation with the α-chain colored in green and the β-chain colored in pink.

(b) Amino Acid No
<table>
<thead>
<tr>
<th></th>
<th>11</th>
<th>20 / 61</th>
<th>70</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRB1*04:01</td>
<td>VKHECHFFNG</td>
<td>WNSQKDLLEQ</td>
<td>KRAVDTYCR</td>
<td></td>
</tr>
<tr>
<td>DRB1*04:04</td>
<td>------</td>
<td>-------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>DRB1*04:02</td>
<td>------</td>
<td>-------</td>
<td>------</td>
<td></td>
</tr>
</tbody>
</table>

(c) Vimentin-71Cit66-78

(d) Side view of the bound vimentin-71Cit66-78 peptide. The peptide’s 2Fo-Fc electron density map is shown in blue and contoured to 1σ, showing unambiguous density for the peptide. Peptide residues are labeled and numbered, with Citrulline71 occupying the P4 pocket.

Residues Val11β, His13β, Lys71β, and Ala74β are represented as sticks and correspond to the residues present in HLA-DR401, the HLA with the highest risk associated with RA. (b) Sequence alignment of the three HLA-DRB1*04 alleles used in this study showing amino acid polymorphisms. “-” indicates residue conserved with that of HLA-DRB1*04:01:01. Val11β, His13β are conserved in all three alleles (not depicted). (c) HLA-DRB1*04:01 in complex with vimentin-71Cit66-78. The vimentin-71Cit66-78 peptide is bound in the peptide-binding groove, with carbons colored in yellow, nitrogens colored in blue, and oxygens colored in red. The α and β chains are shown in cartoon representation, and colored in green and pink, respectively. (d) Side view of the bound vimentin-71Cit66-78 Peptide.
HLA DRβ polymorphisms and RA susceptibility

HLA DRβ polymorphisms are closely associated with RA disease susceptibility (Raychaudhuri et al., 2012; Viatte et al., 2013). For example, although the HLA-DRB1*04:01 allele is strongly associated with RA susceptibility (odds ratio [OR] 4.44), HLA-DRB1*04:08, *04:05, *04:04, and *10:01 allomorphs are, by comparison, marginally less associated (OR > 4.22), whereas allomorphs such as HLA-DRB1*04:02 and *13:01 are considered RA resistant/protective (OR 1.43 and 0.59, respectively; van der Woude et al., 2010; Raychaudhuri et al., 2012; Viatte et al., 2013). These differing associations are associated with polymorphic differences mapping to positions 11, 13, 71, and 74 (Fig. 1, a and b; Raychaudhuri et al., 2012; Viatte et al., 2013). To establish the differing hierarchies of RA disease susceptibility, we determined the structures of HLA-DRB1*04:04 and HLA-DRB1*04:02 in complex with vimentin–Cit71,66-78 (Fig. 2, d and e; and Table 1). HLA-DRB1*04:01 differs from HLA-DRB1*04:04 by 2 aa, of which a K→R polymorphism maps to position 71. Thus, the P4 pocket remains positively charged within HLA-DRB1*04:04, thereby disfavoring P4-Arg at this position. The P4-Cit of vimentin–Cit71,66-78 in DRB1*04:04 occupied a similar position to that observed in HLA-DRB1*04:01, but was in a more extended conformation (Fig. 3 e). Instead P4-Cit pointed toward and directly contacted Gln70 and Ala74 and H-bonded to Arg71 of HLA-DRB1*04:04, the latter of which occupied a very similar position to Lys71 (Fig. 3 e). Thus, the similarity of the P4 pockets of HLA-DRB1*04:01 and HLA-DRB1*04:04 provided a basis for the similar disease association of these allomorphs. The disease-associated effect of the polymorphisms at positions 11 and 13 in the DRβ chain is less clear. Position 11 resides within the P6 pocket, packing against His13β, the latter of which formed van der Waals contacts with the aliphatic moiety of P4-Cit. Therefore, a His13βSer polymorphism, as observed in the protective HLA-DRB1*13:01

Figure 2. Side view of epitopes bound to HLA-DR4. (a) HLA-DRB1*04:01 bound to vimentin–64Cit59-71. (b) HLA-DRB1*04:01 bound to vimentin–64-69-71Cit59-71. (c) HLA-DRB1*04:01 bound to aggrecan–93-95Cit89-103. (d) HLA-DRB1*04:04 bound to vimentin–71Cit59-71. (e) HLA-DRB1*04:02 bound to vimentin–71Cit59-71. The peptide’s 2Fo-Fc electron density map is shown in blue and contoured to 1.α. Peptide residues are labeled and numbered.
allomorph (Raychaudhuri et al., 2012; Viatte et al., 2013) is likely to impact the packing of the P4 residue. Regardless, a key difference between the HLA-DRB1*04:01 and HLA-DRB1*04:02 allomorphs is that the latter possesses Asp70 and Glu71, which enabled it to bind P4-Arg and P4-Cit (Tm of 77.1°C and 84.3°C, respectively; Table 2). Accordingly, we determined the structures of HLA-DRB1*04:02 in complex with Vimentin-71Cit66-78 and Vimentin66-78 (Fig. 2, e and f). The presence of Glu71, which caused a slight adjustment of neighboring residues in comparison to the HLA-DRB1*04:01 complex, enabled a direct H-bond and salt bridge to be formed with P4-Arg and P4-Cit, respectively (Fig. 4, a and b). In addition, Asp70 reoriented to form a salt bridge with P4-Arg. Hence, P4-Arg can be readily accommodated in some of the RA-protective HLA-DRB1 allomorphs due to the conversion toward a more electronegative P4 pocket (Fig. 4, c and d).

### Antigen processing and HLA-DR4 peptide repertoire

To examine the propensity of the differentially RA-associated HLA-DR4 alleles to tolerate P4-Arg residues, we generated T2 cell lines (class II–deficient) that expressed HLA-DM and HLA-DRB1*04:01, *04:02, or *04:04. Accordingly, in contrast to previous studies on HLA-DR4-binding motifs (Hammer et al., 1993; Sette et al., 1993; Hammer et al., 1995), our data arises from a large number of novel naturally processed and presented peptides identified, using a common platform, from cells that express a single HLA-DR molecule that sample peptides from the same parental cell proteome. Our approach enabled an in-depth analysis of the repertoire of peptides bound to each HLA-DR allele. Over 1000 high confidence peptides were identified for each DR allele, elucidating HLA-II binding motifs for HLA-DRB1*04:01 (n = 1058), HLA-DRB1*04:04 (n = 1797) and HLA-DRB1*04:02 (n = 1239).

These endogenous peptide sequences determined from multiple peptide elution experiments were identified with high confidence using strict bioinformatic criteria that included the removal of common contaminants (Dudek et al., 2012). The motifs generated using this approach were in general agreement with previously determined motifs (Hammer et al., 1995; Sette et al., 1993), specifically exhibiting significantly different specificities at P1 and P4 (Fig. 5 a).

---

**Table 1. Data collection and refinement statistics**

<table>
<thead>
<tr>
<th>Structure</th>
<th>Space group</th>
<th>Cell dimensions (Å)</th>
<th>Resolution (Å)</th>
<th>Total no. observations</th>
<th>No. unique observations</th>
<th>Multiplicity</th>
<th>R&lt;sub&gt;merge&lt;/sub&gt;</th>
<th>R&lt;sub&gt;p.i.m&lt;/sub&gt;</th>
<th>〈I/〈I〉</th>
<th>Completeness (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR401Vim-71Cit&lt;sub&gt;66-78&lt;/sub&gt;</td>
<td>C22&lt;sub&gt;2&lt;/sub&gt;</td>
<td>67.9, 177.8, 76.7</td>
<td>62.73-2.30 (2.42-2.30)</td>
<td>120743 (17589)</td>
<td>20836 (2991)</td>
<td>5.8 (5.9)</td>
<td>14.7 (49.5)</td>
<td>6.7 (22.3)</td>
<td>9.5 (3.4)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>DR401Vim-64Cit&lt;sub&gt;59-71&lt;/sub&gt;</td>
<td>C22&lt;sub&gt;2&lt;/sub&gt;</td>
<td>67.1, 183.4, 77.3</td>
<td>48.84-2.41 (2.54-2.41)</td>
<td>88651 (13122)</td>
<td>18858 (2717)</td>
<td>4.7 (4.8)</td>
<td>15.8 (49.3)</td>
<td>8.2 (25.4)</td>
<td>7.3 (3.0)</td>
<td>99.9 (100)</td>
</tr>
<tr>
<td>DR401Vim-69-71Cit&lt;sub&gt;59-71&lt;/sub&gt;</td>
<td>C22&lt;sub&gt;2&lt;/sub&gt;</td>
<td>67.2, 183.6, 77.4</td>
<td>91.26-2.20 (2.32-2.20)</td>
<td>109458 (16044)</td>
<td>23984 (3509)</td>
<td>4.6 (4.6)</td>
<td>15.5 (60)</td>
<td>8 (31.2)</td>
<td>7.6 (2.6)</td>
<td>97.5 (98.3)</td>
</tr>
<tr>
<td>DR401Agg-93-103Cit&lt;sub&gt;89-103&lt;/sub&gt;</td>
<td>C22&lt;sub&gt;2&lt;/sub&gt;</td>
<td>67.1, 1825.775</td>
<td>62.97-1.95 (2.06-1.95)</td>
<td>343643 (49187)</td>
<td>34301 (4325)</td>
<td>6.5 (4.2)</td>
<td>12.2 (45.7)</td>
<td>5.1 (24.4)</td>
<td>11.3 (2.6)</td>
<td>97.5 (85.8)</td>
</tr>
<tr>
<td>DR402Vim-71Cit&lt;sub&gt;66-78&lt;/sub&gt;</td>
<td>C22&lt;sub&gt;2&lt;/sub&gt;</td>
<td>66.4, 1825.778</td>
<td>62.43-1.70 (1.79-1.70)</td>
<td>222923 (18256)</td>
<td>222984 (3509)</td>
<td>6.4 (6.5)</td>
<td>10.5 (58.8)</td>
<td>4.4 (24.3)</td>
<td>12.3 (2.9)</td>
<td>99.8 (99.8)</td>
</tr>
<tr>
<td>DR402Vim-71Cit&lt;sub&gt;66-78&lt;/sub&gt;</td>
<td>C22&lt;sub&gt;2&lt;/sub&gt;</td>
<td>67.0, 1829.77</td>
<td>66.4-2.0 (2.11-2.0)</td>
<td>334634 (49187)</td>
<td>334634 (49187)</td>
<td>7.0 (7.1)</td>
<td>12.2 (49.7)</td>
<td>5 (20.1)</td>
<td>12.5 (3.5)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>DR402Vim-71Cit&lt;sub&gt;66-78&lt;/sub&gt;</td>
<td>C22&lt;sub&gt;2&lt;/sub&gt;</td>
<td>67.4, 183.0, 77.5</td>
<td>45.75-1.65 (1.74-1.65)</td>
<td>228336 (33111)</td>
<td>228336 (33111)</td>
<td>7.2 (7.1)</td>
<td>10.0 (47.2)</td>
<td>4.0 (18.9)</td>
<td>10.6 (3.3)</td>
<td>100 (100)</td>
</tr>
</tbody>
</table>

---

1R<sub>p.i.m</sub> = Σ(|F<sub>o</sub>| − |F<sub>c</sub>|) / Σ|F<sub>c</sub>|
2R<sub>merge</sub> = Σ(||F<sub>o</sub>| − |F<sub>c</sub>|) / Σ|F<sub>c</sub>|
3<|imath>θ| = Σ(|θ<sub>o</sub> − θ<sub>c</sub>| / Σθ<sub>c</sub>)
4R<sub>p.i.m</sub> = Σ(|θ<sub>o</sub> − θ<sub>c</sub>| / Σθ<sub>c</sub>)
5Values in parentheses refer to the highest resolution bin.

Published November 4, 2013

JEM Vol. 210, No. 12

2573
citrullinated peptides may also contribute to their antigenicity. To establish this, we expressed recombinant vimentin and citrullinated it using the PAD2 enzyme. We compared in vitro cathepsin L digestion patterns of native and citrullinated vimentin and observed relative protection of the vimentin59-71 epitope when the antigen was citrullinated at positions 64, 69 and 71 (Fig. 5 b). Similar differences in cleavage patterns were observed using synthetic peptides encompassing the native and citrullinated vimentin 57–71 region (not shown). This suggests that citrullination not only facilitates binding of autoantigenic epitopes to RA-associated HLA allotypes but that the modification of arginine residues also alters protease cleavage whereas P4-Arg was absent in all the peptides bound to HLA-DRB1*04:01 and HLA-DRB1*04:04, arginine for HLA-DRB1*04:02-bound peptides was better tolerated in this position (Fig. 5 a). These data are consistent with HLA-DRB1*04:01/04 disfavoring P4-Arg in vitro (Fig. 3) and not being selected at all in vivo (Fig. 5 a). In contrast HLA-DRB1*04:02 has a propensity to bind P4-Arg in vitro (Fig. 4 b) and is permissive to P4-Arg containing peptides in vivo (Fig. 5 a) with 1.7% of naturally selected peptides containing a P4-Arg. In addition to satisfying the binding requirements to HLA-DRB1*04:01 and other RA-associated HLA-DR allotypes, we hypothesized that differential processing of citrullinated peptides may also contribute to their antigenicity.
patterns protecting regions of the antigen normally degraded in APCs. Thus, citrullination has a double-edged effect, both permitting SE binding and preventing degradation of post-translationally modified epitopes that can be presented to autoreactive T cells in the context of the SE.

Ex vivo T cell analysis using HLA DR4 tetramers

Next, we aimed to identify circulating citrullinated epitope-specific CD4+ T cells. We recruited 20 HLA-DRB1*04:01+ RA patients and 6 HLA-matched healthy controls, with the RA patients possessing a range of disease durations, disease activity, and treatments (Table 3). We generated phycoerythrin (PE)-labeled HLA-DRB1*04:01 tetramers complexed with either: control influenza hemagglutinin (HA)306-318, vimentin-64Cit59-71, or aggrecan-93-95Cit89-103 peptides. We demonstrated that gating based on PE fluorescence-minus-one (FMO) staining reliably gates HA-specific T cells in immunized mice without background in saline-treated mice (unpublished data), and then showed specificity of the T cells using tetramers labeled with different fluorochromes (Tung et al., 2007; Fig. 6 a and not depicted). Although we determined the median absolute number of CD4+ T cells to be 7 × 10^4 in healthy controls and 10.2 × 10^4/ml blood in RA patients, the median number of HA, cit-vimentin, or cit-aggrecan HLA-DRB1*04:01 tetramer+ cells ranged from 47 to 80/ml in RA patients and 30–40/ml in healthy controls—a frequency of ~1/2,000 CD4+ T cells. There was no significant difference between RA patients and healthy controls in the number of CD4+ or tetramer+ T cells/ml (Mann-Whitney test compared RA patients and controls for each specificity; Fig. 6 b). However, the number of vimentin-64Cit59-71 (spearman r = 0.76; P < 0.05) or aggrecan-93-95Cit89-103 tetramer+ T cells (spearman r = 0.76; P < 0.05), but not the total number of CD4+ T cells, was correlated with RA disease activity score

![Figure 4](https://example.com/figure4.png)
and effector-memory T cells. Human PB CD4+ T cells can be subdivided into resting CD45RAFoxp3CD25+ and activated CD45RAFoxp3CD25hi suppressive populations (regulatory T [T reg] cells), and Foxp3CD45RA− and Foxp3CD45RA+ and CD45RA− nonsuppressive populations, which each have potential for proinflammatory cytokine production upon stimulation ex vivo (Miyara et al., 2009, 2011). We stained PBMCs similarly, substituting CD45RO, which identifies a reciprocal population to CD45RA. Among
the total CD4+ cells in PB of HLA-DRB1*04:01+ RA patients relative to healthy control donors, the proportion of resting (Fig. 6 e) and activated (Fig. 6 f) T reg cells was significantly reduced and the proportion of FoxP3− effector/memory (Fig. 6 h) cells tended to be increased (P = 0.05). Vimentin-64Cit59-71 and, in most cases, aggrecan-93-95Cit89-103-specific T cells were significantly less likely to be resting (Fig. 6 e) or activated (Fig. 6 f) T reg cells and significantly more likely to have a FoxP3− CD45RO− naive (Fig. 6 g) or CD45RO+ effector memory (Fig. 6 h) phenotype in RA than healthy control PBs. HA-specific T cells were also significantly more likely to have an effector memory phenotype in RA than healthy control PB (Fig. 6 h). These data indicate that the HLA-DRB1*0401 SE permits the selection and/or peripheral expansion of low numbers of CD4+ T cell populations specific for vimentin-64Cit59-71 and aggrecan-93-95Cit89-103 self-antigens unrelated to a history of RA. This is consistent with recent findings in healthy HLA-DRB1*0401+ individuals, where self-antigen–specific CD4+ T cells were observed in preenriched samples, despite the donors not suffering from autoimmune disease (Su et al., 2013). The enrichment in naive and effector/memory T cells and paucity of T reg cells among antigen–specific CD4+ T cells, further indicates that T cell regulatory capacity is deficient among CD4+ T cells, including autoreactive CD4+ T cells in HLA-DRB1*0401+ RA patients relative to HLA-DRB1*0401+ healthy controls.

**DISCUSSION**

Given the central role of posttranslational modifications of proteins in regulating essential physiological processes, surprisingly little is known regarding the molecular basis underlying their impact on immunity (Petersen et al., 2009). Nevertheless, some recent advances, particularly in the area of T cell–mediated autoimmunity, have demonstrated the capacity of T cells to recognize HLA-restricted posttranslationally modified antigens. For such reactivity, an individual’s immunogenetics and the antigens themselves are closely associated with the pathogenesis of diseases, including type 1 diabetes (Mannering et al., 2005), celiac disease (Abadie et al., 2011; Broughton et al., 2012), and RA (Law et al., 2012; Snir et al., 2011; van Delwijk et al., 2010).

The association between the HLA-DRB1 locus and RA has been known for over 25 yr, leading to the shared epitope hypothesis (Gregersen et al., 1987). Further, there is a clear association between these shared–epitope alleles and citrullination, where several citrullinated epitopes are identified in RA patients. Moreover, a large haplotype association study attributed most of the HLA-DR–associated risk to positions 11, 13, 71, and 74 of the HLA DRB polypeptide chain in RA, strongly suggesting that this allotype permits binding and presentation of autoantigenic peptides (Raychaudhuri et al., 2012). Our findings provide a comprehensive structural portrait of the association between RA, HLA-DRB1, and citrullinated peptides. Namely, we describe seven, high resolution, crystal structures of HLA DR4–Ag complexes of direct relevance to RA. We show how four RA-associated citrullinated epitopes (vimentin-64Cit59-71, vimentin-64-69-71Cit59-71, vimentin-71Cit66-78, and aggrecan-93-95Cit89-103) bound to HLA-DRB1*04:01, an allele that is strongly associated with RA susceptibility. These four structures show that the mode of binding of the P4-Citrulline residue within the P4 pocket of HLA-DRB1*04:01 is conserved, in which P4-Cit contacted positions 13 and 71 of the SE motif. These structures provided a clear and general explanation as to (a) why P4-Arg could not be accommodated within the P4 pocket of HLA-DRB1*04:01 and (b) how P4-Cit was accommodated within the electropositive P4 pocket of HLA-DRB1*04:01. Namely, the P4 pocket of HLA-DRB1*04:01 is highly suited to preferentially accommodate citrulline over the corresponding Arg, with Lys71β of the SE playing a key discriminatory role.

Next, we examined how HLA DRB polymorphisms can impact RA hierarchies of disease susceptibility. First, we determined the structure of HLA-DRB1*04:04 in complex with vimentin–Cit71Cit71, strongly suggesting that this allotype permits binding and presentation of autoantigenic peptides (Raychaudhuri et al., 2012). Our findings provide a clear and general explanation as to (a) why P4-Arg could not be accommodated within the P4 pocket of HLA-DRB1*04:01 and (b) how P4-Cit was accommodated within the electropositive P4 pocket of HLA-DRB1*04:01. Namely, the P4 pocket of HLA-DRB1*04:01 is highly suited to preferentially accommodate citrulline over the corresponding Arg, with Lys71β of the SE playing a key discriminatory role.

## Table 3. Characteristics of HLA-DRB1*04:01+ RA patients used for optimization, enumeration and phenotypic studies

<table>
<thead>
<tr>
<th>Demographic and clinical details</th>
<th>RA patients (n = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD)</td>
<td>58 (13.7)</td>
</tr>
<tr>
<td>Female sex, n (%)</td>
<td>15 (75)</td>
</tr>
<tr>
<td>Disease duration (y), median (i.q.r)</td>
<td>2 (1.25-5.75)</td>
</tr>
<tr>
<td>ACPA+, n (%)</td>
<td>15 (75)</td>
</tr>
<tr>
<td>Ever smokers, n (%)</td>
<td>17 (85)</td>
</tr>
<tr>
<td>Disease activity (DAS4v-CRP), median (i.q.r)</td>
<td>2.38 (1.58-3.0)</td>
</tr>
</tbody>
</table>

## Treatment

<table>
<thead>
<tr>
<th>Treatment</th>
<th>RA patients (n = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methotrexate, n (%)</td>
<td>13 (65)</td>
</tr>
<tr>
<td>Hydroxychloroquine, n (%)</td>
<td>13 (65)</td>
</tr>
<tr>
<td>Sulfasalazine, n (%)</td>
<td>8 (40)</td>
</tr>
<tr>
<td>Low dose prednisone, n (%)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Leflunomide, n (%)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Multiple antirheumatic drugs, n (%)</td>
<td>14 (70)</td>
</tr>
<tr>
<td>No treatment, n (%)</td>
<td>1 (5)</td>
</tr>
</tbody>
</table>
numbers in the CD4+ T cell peripheral blood repertoire of *HLA-DRB1*04:01+ RA patients and healthy controls. Our data imply that the exclusion of P4-Arg and acceptance of P4-Cit by HLA-DRB1*04:01 leads to the presentation of peptides that can interact with the corresponding autoreactive T cell repertoire to increase selection and/or expansion of autoreactive CD4+ T cells. T cells of highest self-reactivity escaping the affinity threshold for deletion in the thymus are found among the natural T reg cell population (Hsieh et al., 2004). Indeed, the autoreactive T cells in *HLA-DRB1*04:01+ healthy controls were enriched in resting and activated Foxp3+ T reg cells, and MFI reflecting tetramer binding avidity was higher in healthy controls than in RA patients, whose autoreactive T cells were relatively deficient in T reg cells. The correlation between antigen-specific T cell frequency and RA disease activity suggests disproportionate peripheral expansion or survival of effector/memory cells relative to T reg cells in RA patients, potentially due to antigen presenting cell activation or IL-2 availability, on which RA genetic background and inflammation impact (Li et al., 2013; Pettit et al., 2000; Viatte et al., 2013). These data are consistent with previously reported expansion of total CD4+ T cells in RA PB, correlated with disease activity (Scarsi et al., 2010; Sempere-Ortells et al., 2009). Further, the higher proportion of FoxP3+ autoreactive effector/memory T cells in RA patients indicates higher cytokine production potential in response to presentation of citrullinated autoantigens at sites of inflammation including the lung of smokers and RA joints.
which has been demonstrated ex vivo (Law et al., 2012; Makrygiannakis et al., 2008; Wegner et al., 2010). This recognition is amplified by the protection of regions of citrullinated antigens from prolylase, thereby promoting the presentation of citrullinated self-epitopes to autoreactive T cells. Collectively, our findings have reshaped our understanding of the association between citrullination, the HLA-DRB1 locus, autoreactive T cells, and their regulation in RA.

**MATERIALS AND METHODS**

**Mammalian expression vector construction.** The extracellular domains of the HLA-DR4 (DRA*01:01/DRB1*04:01, *04:02, and *04:04) α and β chains were cloned into the pHISec (Aricescu et al., 2006) vector for expression in HEK 293S (GnTI−) cells (Reeves et al., 2002). Constructs contained C-terminal enterokinase cleavable fos/jun zippers to promote dimerization. The β chain also contained a BirA site for biotinylation and tetramer generation and a Histidine tag for IMAC purification. HLA-DR4 was expressed with the class II–associated invariant chain peptide (CLIP) covalently attached via a Factor Xa cleavable linker to the N terminus of the β chain and is preceded by a Strep-II tag (IBA; Göttingen) for purification.

**Expression and purification.** The HLA-DR4CLIP construct was transiently expressed in HEK 293S (GnTI−) cells and soluble protein was purified from the culture medium. In brief, culture medium was concentrated and buffer exchanged using the Cogen M1 TFF system (Merck Millipore) into 10 mM Tris, pH 8.0, and 500 mM NaCl. The proteins were then purified using IMAC via Ni Sepharose 6 Fast Flow (GE Healthcare) and size exclusion in HEK 293S (GnTI−) cells (Reeves et al., 2002). Constructs contained C-terminal enterokinase cleavable fos/jun zippers to promote dimerization. The β chain also contained a BirA site for biotinylation and tetramer generation and a Histidine tag for IMAC purification. HLA-DR4 was expressed with the class II–associated invariant chain peptide (CLIP) covalently attached via a Factor Xa cleavable linker to the N terminus of the β chain and is preceded by a Strep-II tag (IBA; Göttingen) for purification.

**Thermal stability assays.** Thermal stability assays of HLA-DR4 peptide complexes were performed using a Real-Time Detector instrument (Corbett RotorGene 3000). In brief, HLA-DR4 peptide complexes were prepared at either 10 or 20 µM in 10 mM Tris, pH 8.0, 150 mM NaCl. SYPRO orange (Invitrogen) was added to monitor unfolding, samples were heated at 16 h for 6 h at 21°C before peptide exchange. HLA-DR4 was subsequently loaded with test peptides by incubating for 16 h at 37°C in a 20-fold excess of peptide in 100 mM 150 mM NaCl. HLA-DR4CLIP was cleaved with Factor Xa for 6 h at 21°C using IMAC via Ni Sepharose 6 Fast Flow (GE Healthcare) and size exclusion in HEK 293S (GnTI−) cells (Reeves et al., 2002). Constructs contained C-terminal enterokinase cleavable fos/jun zippers to promote dimerization. The β chain also contained a BirA site for biotinylation and tetramer generation and a Histidine tag for IMAC purification. HLA-DR4 was expressed with the class II–associated invariant chain peptide (CLIP) covalently attached via a Factor Xa cleavable linker to the N terminus of the β chain and is preceded by a Strep-II tag (IBA; Göttingen) for purification.

**Crystallization and structure determination.** Crystal trays were set up at 16°C. Crystals were then buffer exchanged into 10 mM Tris, pH 7.6, and 50 mM NaCl, followed by removal of the mother liquor of 100 mM BTP, pH 7.3, 22–28% (vol/vol) PEG3350, and 0.2 M KNO3, fractionated to wash 11:1 ratio. Platelet-like crystals typically grew within 5 d. Crystals were flash frozen in 16% (vol/vol) ethylene glycol before data collection. Data were collected at the MX1 or MX2 beamlines at the Australian Synchrotron and processed using the program Mosflm. The structures were determined by molecular replacement using the program Phaser and subsequently refined using Phenix and iterations of manual refinement using Coot. The structures were validated using MOLPROBITY.

**Human subjects.** 20 patients who fulfilled the 1987 American College of Rheumatology (ACR) criteria for RA (Aletaha et al., 2010) and 6 ACPA−/SE− healthy controls were included. All individuals provided peripheral blood (PB) samples, although the yield was insufficient for all assays in some cases. Patient demographic details are outlined in Table 3. Disease activity scores (DAS4vCRP) were determined on the day of blood sampling for the study. HLA-DR genotyping was performed at Queensland Health Pathology Services. The study was approved by the Metro South and University of Queensland Human Research Ethics Committee, and informed consent was obtained from each individual.

**Tetramer generation.** HLA-DR4 peptide samples were buffer exchanged into 10 mM Tris, pH 8.0, and biotinylated as described previously (Broughton et al., 2012). The percentage of biotinylation was determined by native gel electrophoresis and complexation with avidin. Tetramers were generated by the addition of streptavidin-PE (BD) or streptavidin-Brilliant Violet (BV421; BioLegend) in an 8:1 molar ratio.

**Tetramer staining.** Initial staining optimization was required as cells were rare, and HLA-DR4 tetramer staining intensity was low. PBMCs from HLA-DRB1*04:01 RA patients and healthy controls were thawed from frozen aliquots, stained with 4.2 µg/ml PE-labeled tetramers; aque live-dead discriminator (Invitrogen); FITC-labeled anti-CD11c, -CD14, -CD16, and -CD19, and APC/Cy7-labeled anti-CD4 in the presence of 50 nM dasatinib (Selleckchem). Live CD4+ T cells were gated and non–T cell lineage+ cells were excluded, and then enriched with anti-PE immunomagnetic beads (MACS; Miltenyi Biotec). The HLA-DR4 tetramer gate was set for CD4+ T cells based on PE fluorescence minus one (FMO) staining. Inclusion of 50 nM dasatinib (Selleckchem) during staining markedly increased the detection of tetramer+ T cells. Whereas immunomagnetic enrichment with anti-PE-beads (MACS; Miltenyi Biotec) after staining reduced the number of cells required for acquisition by the flow cytometer, it underestimated the frequency and skewed the phenotype of tetramer+ T cells. Following these optimization experiments, immunomagnetic enrichment was not used; each sample of PBMC was divided into three, and each stained with one PE-labeled tetramer, aque live-dead, anti-CD14-PerCP/Cy5.5, anti-CD4-APC/Cy7, anti–CD45RO-Pacific blue, anti–CD25-PE/Cy7, anti–Foxp3-Alexa Fluor 488, and anti–CD28-APC (BioLegend and BD), and then analyzed using a Gallios flow cytometer and Kaluza software (Beckman Coulter). HLA-DR4 tetramer gating based on PE FMO staining was kept constant for the entire study. The frequency of CD4+CD14− tetramer+ cells/ml blood was calculated based on cell number determined by addition of TruCOUNT beads (BD).

**Mice and immunization.** 1-A4/−/− C57BL/6 mice expressing a chimeric class II transgene containing the α1B1 domains of human DRβ1*0101-B1*0401 on a mouse IEd backbone (DR04:01-IEd mice) were obtained from Taconic and bred and housed under specific pathogen–free conditions at University of Queensland. Experiments were approved by the University of Queensland Animal Ethics Committee. Draining lymph nodes of mice immunized with 1 µg Fluvax 2012 (CSL) or saline-treated mice were removed 4 wk later and stained with 7-AAD, FITC-labeled anti-CD11c, CD14, CD16, and CD19, CD4-APC, and PE-labeled DRB1*04:01-HA348-358 tetramer in the presence of dasatinib. Cells were analyzed using a Gallios flow cytometer and Kaluza software. Live CD4+ T cells were gated and non–T cell lineage+ cells were excluded. Gates for the pHLA-II tetramer staining were set based on PE FMO staining.

**Statistical analysis.** The Kruskal-Wallis test with Dunns’ Multiple Comparision Test compared multiple means. Significance is indicated as *, P < 0.05; **, P < 0.01; ***, P < 0.0001. All error bars represent SEM.

**Preparation and digestion of citrullinated vimentin.** Recombinant human vimentin was generated as an N-terminal hexahistidine fusion
protein in S9 insect cells using the Bac-to-Bac Baculovirus Expression System (Invitrogen). Affinity-purified protein was incubated with rabbit skeletal PAD (Sigma-Aldrich) for 16 h at 37°C in 0.1 M Tris-HCl (pH 7.6) containing 10 mM CaCl2 and 5 mM dithiothreitol. Citrullinated and control proteins were further purified using Nickel Sepharose (GE). Human cathpspin L was expressed in Pichia pastoris (system donated by D. Brömme, University of British Columbia, Vancouver, Canada), purified and titrated using E-64 as described previously (Brömme et al., 1999). Cathpspin L was preincubated by incubation in 0.1 M acetic acid, 1 mM EDTA, and 10 mM cysteine, pH 5.0 for 30 min at room temperature, and 2 nmol cathpspin L was used to digest the vimentin proteins at pH 5.0. At the indicated times, samples were acidified and desalted using a C18 Zip-tip. Samples were eluted with 50% (vol/vol) acetonitrile/0.1% (vol/vol) formic acid, concentrated and separated on an Eksigent Ultra cHiPLC system using a gradient of 5–80% (vol/vol) Acetonitrile for 90 min, and analyzed online using an AB SCIEX 5600+ TripleTOF high resolution mass spectrometer.

Repetitive analysis of HLA-DR4 allomorphs. T2-DRB1*04:01, *04:02, and *04:04 cells expressing DM were generated via retroviral transduction of the parental T2 line as previously described (Pang et al., 2010). Cells were expanded in RPMI-10% FCS and pellets of 106 cells snap frozen in liquid nitrogen. Cells were ground under cryogenic conditions and resuspended in lysis buffer (0.5% IGEPAL, 50 mM Tris, pH 8.0, 150 mM NaCl and protease inhibitors) as previously described (Dudek et al., 2012; Illing et al., 2012). Cleared lysates were passed over a protein A precolumn followed by an affinity column cross-linked with a monoclonal antibody specific for HLA-DR (LB3.1). Peptide–MHC complexes were eluted from the column by acidification with 10% (vol/vol) acetic acid. Peptides were isolated using reversed-phase HPLC (Chromolith C18 Speed Rod; Merck) on an Akta Ettan HPLC system (GE HealthCare). Fractions were concentrated and analyzed using an AB SCIEX 5600+ TripleTOF high-resolution mass spectrometer as previously described (Dudek et al., 2012). Acquired data were searched against the human proteome (Uniprot/Swissprot v2012_7) excluding 10 mM CaCl2 and 5 mM dithiothreitol. Citrullinated and control proteins: correlation with rheumatoid arthritis association.

We thank Helen Pahau for recruitment and venesection of patients and controls, and staff at the MX1 and MX2 beamlines of the Australian synchrotron for assistance with data collection.

This research was supported by the National Health and Medical Research Council of Australia (NHMRC) and the Australian Research Council. A.W. Purcell was supported by an NHMRC Senior Research Fellowship; N.L. La Gruta was supported by a Sylvia and Charles Viertel Senior Medical Research Fellowship; A.W. Purcell was supported by an NHMRC SRF; J. Rossjohn was supported by an NHMRC Australia Fellowship; R. Thomas was supported by an ARC Future Fellowship and Arthritis Queensland.

There was no commercial support for this work. However we wish to declare that R. Thomas has filed a provisional patent surrounding technology for targeting Dcs for antigen-specific tolerance, and is a director of a spin-off company that is commercializing vaccines that target Dcs to suppress autoimmune diseases.

Submitted: 13 June 2013
Accepted: 4 October 2013

REFERENCES
http://dx.doi.org/10.1083/jem.181.5.1847

Downloaded from jem.rupress.org on October 13, 2015


