Human Leukocyte Antigen (HLA) B27 Allotype-Specific Binding and Candidate Arthritogenic Peptides Revealed through Heuristic Clustering of Data-independent Acquisition Mass Spectrometry (DIA-MS) Data*

Ralf B. Schittenhelm‡, Saranjah Sivaneswaran‡, Terry C. C. Lim Kam Sian‡, Nathan P. Croft‡, and Anthony W. Purcell‡¶

Expression of HLA-B27 is strongly associated with ankylosing spondylitis (AS) and other spondyloarthropathies. While this is true for the majority of HLA-B27 allotypes, HLA-B*27:06 and HLA-B*27:09 are not associated with AS. These two subtypes contain polymorphisms that are ideally positioned to influence the bound peptide repertoire. The existence of disease-inducing peptides (so-called arthritogenic peptides) has therefore been proposed that are exclusively presented by disease-associated HLA-B27 allotypes. However, we have recently demonstrated that this segregation of allotype-bound peptides is not the case and that many peptides that display sequence features predicted to favor binding to disease-associated subtypes are also capable of being presented naturally by protective alleles. To further probe more subtle quantitative changes in peptide presentation, we have used a combination of data-independent acquisition (DIA) and multiple reaction monitoring (MRM) mass spectrometry to quantify the abundance of 1646 HLA-B27 restricted peptides across the eight most frequent HLA-B27 allotypes (HLA-B*27:02–HLA-B*27:09). We utilized K means cluster analysis to group peptides with similar allelic binding preferences across the eight HLA-B27 allotypes, which enabled us to identify the most-stringent binding characteristics for each HLA-B27 allotype and further refined their existing consensus-binding motifs. Moreover, a thorough analysis of this quantitative dataset led to the identification of 26 peptides, which are presented in lower abundance by HLA-B*27:06 and HLA-B*27:09 compared with disease-associated HLA-B27 subtypes. Although these differences were observed to be very subtle, these 26 peptides might encompass the sought-after arthritogenic peptide(s). Molecular & Cellular Proteomics 15: 10.1074/mcp.M115.056358, 1867–1876, 2016.

Human leukocyte antigen (HLA) class I molecules present endogenous peptides derived from self-proteins or intracellular pathogens at the cell surface for scrutiny by CD8+ T lymphocytes. These peptide ligands are constrained within the peptide-binding cleft, which consists of six binding pockets (A-F) that accommodate the amino acid side chains of the peptides (1). Polymorphisms characterizing and differentiating HLA alleles and allotypes mostly cluster around these binding pockets, thereby shaping and defining their specific peptide cargo.

HLA-B27, which belongs to the family of HLA class I molecules, is an extremely polymorphic B-locus with more than 100 allelic variants known to date (2, 3). The most frequent allotypes with comprehensively described peptide repertoires are HLA-B*27:02–HLA-B*27:09 (4–18). Due to the architecture of their B pocket, the most dominant feature of all HLA-B27 bound peptides is the presence of an Arg residue at position P2 (19–21). Additional primary and secondary anchor positions are P3, PΩ-2, and PΩ, which mostly carry hydrophobic and aromatic amino acid residues that bind into the D, E, and F pockets, respectively (18, 20, 21). All other peptide positions are considered as nonanchor positions and tolerate a large variety of amino acid residues with notable differences between the various allotypes (summarized in (18). Ankylosing spondylitis (AS)1, a chronic inflammatory autoimmune disease, and other spondyloarthropathies are strongly

From the ‡Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia
Received February 9, 2016, and in revised form, February 9, 2016
Published, MCP Papers in Press, February 29, 2016, DOI 10.1074/mcp.M115.056358
Author contributions: R.B.S. and A.W.P. designed the research; R.B.S., S.S., T.C.L., and N.P.C. performed the research; R.B.S. and A.W.P. analyzed data; and R.B.S. and A.W.P. wrote the paper.

1 The abbreviations used are: AS, ankylosing spondylitis; CTL, cytotoxic T lymphocytes; DDA, data-dependent acquisition; DIA, data-independent acquisition; t; HLA, human leukocyte antigen; MRM, multiple reaction monitoring; RP-HPLC, reverse-phase high-performance liquid chromatography; SWATH, sequential window acquisition of all theoretical mass spectra.

Molecular & Cellular Proteomics 15.6 1867
linked to the expression of HLA-B27 (reviewed in (22, 23). Even though this association has been known for more than 40 years (24, 25), the pathogenic property of HLA-B27 remains still unknown and many hypotheses are currently being investigated to elucidate its apparent role in AS (reviewed in (26, 27). Due to the peptide-presenting function of HLA-B27 and other HLA molecules, one of the most popular hypotheses is the arthritogenic peptide theory. This theory assumes that cytotoxic T lymphocytes (CTL) primed against pathogenic HLA-B27-restricted antigen(s) cross-react on molecularly similar self-peptide(s), which subsequently leads to tissue damage and onset of autoimmunity. In support of this theory, onset of AS is often preceded by infection with enteric bacteria and HLA-B27-restricted CTLs have been identified in the synovial fluid of AS patients, which recognize both bacterial epitopes and self-peptides (28, 29). Moreover, two allelic variants of HLA-B27 (HLA-B*27:06 and HLA-B*27:09), which contain polymorphisms ideally placed to alter the presented peptide repertoire, are not associated with AS (30–34).

Based on this apparent protective role of HLA-B*27:06 and HLA-B*27:09 in disease pathogenesis, a number of analyses focused on comparing the peptide repertoires of disease-associated and nonassociated HLA-B27 subtypes with the goal to identify putative arthritogenic peptides, which were expected to be qualitatively absent from nondisease-associated subtypes (11–14, 17, 18). However, such a peptide has not been described yet, and a peptide feature that stratifies disease- and nondisease-associated subtypes does not appear to exist. In agreement of this notion, we could recently show that a peptide, which is capable of binding to each disease-associated HLA-B27 subtype, will also bind to the nondisease-associated allelotypes HLA-B*27:06 and HLA-B*27:09 (18). This observation challenged the arthritogenic peptide theory in such that absolute binding preferences do not explain disease association. However, quantitative changes in peptide presentation between disease-associated and nonassociated subtypes may be relevant for disease pathogenesis by exceeding (or falling below) a threshold required for the activation of autoreactive T cells and, thus, for the onset of autoimmunity. Therefore, we sought to quantify the abundance of known HLA-B27 specific peptides across the eight most frequent HLA-B27 alleles (HLA-B*27:02–HLA-B*27:09) with the goal to specifically identify candidate arthritogenic peptides that are expected to be presented in lower abundance by HLA-B*27:06 and HLA-B*27:09 compared with disease-associated HLA-B27 subtypes.

Data-independent acquisition (DIA) mass spectrometry—often referred to as SWATH-MS on SCIEX TripleTOF® instruments—is the state-of-the-art method to quantify thousands of peptides across multiple samples (35, 36). A mass spectrometer operated in DIA/SWATH mode uses wide ms1 isolation windows (typically 20–25 Da) to scan sequentially and repeatedly through a given mass range. All precursor ions isolated within each window are simultaneously fragmented in the collision cell resulting in multipeptide-derived ms2 spectra. Even though these spectra are too complex to be interrogated by conventional search engine algorithms, the continuous acquisition of all fragment ions over the chromatographic run time allows the extraction of quantifiable ms2 traces for virtually every peptide of interest. However, a faithful and accurate evaluation of DIA/SWATH-MS data requires prior knowledge of spectral and chromatographic information for all peptides of interest, although significant efforts are currently being made to develop software packages that extract sequence information directly from DIA/SWATH data (for example DIA Umpire: (37)).

Despite unique and compelling advantages of DIA/SWATH technology, multiple reaction monitoring (MRM)-MS still excels in unparalleled sensitivity (38, 39). This level of sensitivity is achieved by analyzing only one fragment ion (transition) at a time on a triple quadrupole mass spectrometer, which combines excellent ion transmission efficiencies with dual ion selectivity processes to significantly reduce background noise in favor of the analyte of interest. However, MRM-MS is severely restricted by a relatively low number of peptides that can be analyzed in one acquisition.

In this study, we used a combination of SWATH- and MRM-MS to quantify the abundance of 1646 epitopes across eight different HLA-B27 allelotypes (HLA-B*27:02-HLA-B*27:09), including the non-AS associated subtypes HLA-B*27:06 and HLA-B*27:09. Computational algorithms were utilized to group all quantified peptides into distinct clusters with similar binding preferences across the HLA-B27 allelotypes investigated. A detailed analysis of these cluster-specific consensus-binding motifs provided a means to attain the most stringent binding characteristics for each HLA-B27 allele and further refined our understanding of the intricate mechanisms underlying HLA-peptide binding. Of note, a thorough comparison of this comprehensive quantitative dataset revealed 26 putative arthritogenic peptides, whose immunological potential will be analyzed in subsequent studies.

**Experimental Procedures**

**Experimental Design and Statistical Rationale**—We recently described the comprehensive analysis of the peptide repertoires of the eight most frequent HLA-B27 alleles (HLA-B*27:02–HLA-B*27:09) by LC-MS/MS (18). In brief, HLA-B27 bound peptides were isolated from stable C1R transfectants by immunoaffinity capture, separated by reverse-phase high-performance liquid chromatography (RP-HPLC), and sequenced on a 5600® TripleTOF® (SCIEX, Framingham, USA) mass spectrometer. At least four independent immunoprecipitations (biological replicates) were conducted to establish comprehensive immunopeptidomes for each HLA-B27 allomorph.

To obtain sufficient material for the eight SWATH-MS and 152 MRM-MS acquisitions presented in this study, the remaining aliquots of these subtype-specific peptide eluates were combined to form one biological pool for each of the eight distinct samples (HLA-B*27:02–HLA-B*27:09). By its nature, SWATH-MS eliminates technical variation allowing comprehensive interrogation of such pooled biological samples.
Quantification of HLA-B27 Peptides by SWATH- and MRM-MS

Data-Independent Acquisition (SWATH-MS)—Samples were loaded for 15 min with an isocratic flow (5 μl/min) of loading buffer (0.1% formic acid, 2% acetonitrile) onto a trap column (Eksigent, Dublin, USA; 200 μm × 0.5 mm ChromXP C18-CL 3 μm 120Å) using an Eksigent NanoUltra chipLC system. A microfluidic analytical column (Eksigent; 75 μm × 15 cm ChromXP C18-CL 3 μm 120Å) was switched in-line to separate the peptides by increasing concentrations of 80% acetonitrile/0.1% formic acid at a flow of 300 nl/min for 85 min.

The peptides were analyzed on a 5600+ TripleTOF mass spectrometer (SCIEX) operated in positive ion and SWATH-MS mode. In each cycle, a 200 ms precursor scan (range: 300–1800 Da) preceded the acquisition of 28 sequential SWATH-MS scans with an accumulation time of 100 ms each resulting in a total cycle time of 3 s. Each of the 28 SWATH-MS acquisitions scanned a 25 Da window between 300 to 1000 Da.

Raw SWATH data files have been deposited at PeptideAtlas (www.peptideatlas.org) with the dataset identifier PASS00763 (password: DT4743iu).

SWATH-MS data analysis—All SWATH-MS data were evaluated with the open source software Skyline (v2.6.0.7176; (40)). The six most abundant b- or y-fragment ions of each peptide were selected for monitoring based on a spectral library, which contained information of all HLA-B27 peptides identified in our previous study (18). A custom-made retention time predictor based on retention time alignment of 15 intrinsic HLA-B27 peptides was used to correct retention time drift between analyses (Supplemental Fig. 1).

SWATH-MS data were imported into Skyline and the performance of the peak picking was visually assessed for each peptide and, if necessary, manually adjusted. A peptide identification was considered successful if all of the following criteria were met in at least one sample: (i) the deviation between the peptide’s observed and predicted retention time was less than 5 min, (ii) the dot product, which delineates the similarity between the library spectrum and the observed transition chromatograms, was greater than 0.5, (iii) the mass error was smaller than 5 ppm, and (iv) the maximal peak height of a served transition chromatograms, was greater than 0.8, (iii) the mass error was smaller than 5 ppm, and (iv) the maximal peak height of a served transition chromatograms, was greater than 0.8.

MRM-MS Acquisition—A QTRAP® 5500 mass spectrometer (SCIEX) was coupled to an Eksigent NanoUltra chipLC system. A sample was loaded via a microfluidic trap column (Eksigent; 75 μm × 120Å) onto a microfluidic analytical column (Eksigent; 75 μm × 15 cm ChromXP C18-CL 3 μm 120Å) and the peptides were separated by increasing concentrations of 98% acetonitrile/0.1% formic acid at a flow rate of 300 n/min for 90 min. The QTRAP® 5500 was operated in MRM mode (unit retention time resolution). The k-means algorithm of the TM4:MeV software (http://www.tm4.org/mev.html; (43)) was used to group the peptides according to their relative expression levels into 30 distinct clusters using the default parameters.

RESULTS

Quantification of B27-Specific Peptides across the Eight Most Frequent HLA-B27 Allotypes—We utilized a combination of SWATH-MS, a DIA method accessible on TripleTOF® 5600+ instruments (SCIEX), and MRM-MS to comprehensively and accurately quantify HLA-B27-specific peptides across the eight most frequent HLA-B27 allotypes (HLA-B*27:02–HLA-B*27:09) including the nondisease-associated subtype HLA-B*27:06 and HLA-B*27:09 (see Fig. 1 for experimental workflow). Equal amounts of subtype-specific HLA-B27 peptide eluates were pooled and analyzed on a TripleTOF® 5600+ mass spectrometer (SCIEX) operated in SWATH-MS mode. The data were evaluated in Skyline (40) based on a comprehensive library containing spectral information of peptides isolated from all HLA-B27 allotypes and a custom-made retention time predictor to align peak picking across each sample (IRT-B27 retention time predictor; Supplemental Fig. 1). In total, we analyzed 3306 peptides encompassing 3277 high confidence HLA-B27-bound peptides from our previous studies (18) and 29 selected HLA-Cw4-specific peptides, which are present in all the eluates due to low level expression of HLA-Cw4 on the parental C1R cell line (18, 44; Supplemental Table 2).

Because of the expected low abundance of certain peptides in some samples, the performance of the peak-picking algorithm was manually assessed for all peptides, and the peak boundaries were, if necessary, manually corrected to ensure faithful quantitative measurements. Only peptide assignments that fulfilled all of the following stringent evaluation criteria in at least one sample were considered for subsequent analyses: (i) the deviation between the observed and predicted retention time was less than 5 min, (ii) the dot product was greater than 0.8, (iii) the mass error was smaller than 5 ppm, and (iv) the maximal peak height exceeded 500 arbitrary intensity units. Using this strategy, quantitative measurements of 1095 HLA-B27-specific peptides were obtained across the eight different HLA-B27 subtypes and subsequently normalized to correct for differ-
This comprehensive SWATH-MS dataset contained 92 peptides, which were recently quantified across all HLA-B27 allotypes by MRM-MS (18). To validate the accuracy of
our SWATH-MS data, we compared the relative expression levels of these peptides to the published MRM-MS results (Supplemental Fig. 5A). Of note, 78 peptides (85%) showed an average variation of less than 20% across all HLA-B27 allotypes, confirming highly quantitative accuracy within the SWATH-MS dataset (Supplemental Fig. 5B). In addition, four peptides with both doubly and triply charged precursor ions were present in our SWATH-MS dataset. As expected, quantification of the differently charged precursor ions resulted in nearly identical relative expression levels across all HLA-B27 allotypes (Supplemental Fig. 5C).

In comparison to SWATH-MS, which results in MRM-like transition chromatograms, MRM-MS is a more focused and sensitive approach due to enhanced detection and use of target-optimized collision energies. Therefore, the 2211 peptides that did not pass our stringent SWATH-MS evaluation criteria were reanalyzed in 152 separate MRM-MS runs on a QTRAP® 5500 mass spectrometer (SCIEX) operated in MRM mode and evaluated in Skyline using the same criteria as described for the SWATH-MS acquisitions. Pursuing this strategy, quantitative data of an additional 551 HLA-B27-specific peptides was obtained (Supplemental Table 3: Quantification results).

Taken together, we were able to quantify the abundance of 1646 peptides across the eight most frequent HLA-B27 allotypes using a combination of SWATH- and MRM-MS.

Cluster Analyses Confirm and Extend the Allele Specific-Binding Motifs of the Different HLA-B27 Allotypes—We applied cluster analysis to group peptides with similar relative binding patterns across the different HLA-B27 allotypes. Based on adjusted “figure of merit” graphs (Supplemental Fig. 6), which propose appropriate input parameters for the k-Means algorithm (45), we opted to group the 1646 peptides into 30 distinct clusters, which appeared to be sufficient to adequately separate the peptides according to their allotype-specific binding patterns (Fig. 2; Supplemental Fig. 7). The resulting clusters contained between 17 and 128 peptides, and we determined the peptide length distributions and the consensus-binding motifs for each cluster (Supplemental Fig. 8, Supplemental Fig. 9).

The first eight cluster groups contained peptides that are mainly presented by a single HLA-B27 allotype (Fig. 2;
Supplemental Fig. 7). As expected, the binding motifs of these clusters were similar to previously determined canonical consensus-binding motifs (4–18). However, the consensus-binding motifs detailed in these studies are defined by analyzing all peptides eluted from a specific HLA-B27 allotype. Conversely, peptides within any of the first eight cluster groups are predominantly presented by only one HLA-B27 allotype and therefore, must contain feature(s) that are favored by this allotype but less tolerated by the remaining subtypes. Indeed, a thorough comparison of these cluster-specific binding motifs revealed several allotype-favored peptide features for each HLA-B27 subtype and provided a means to attain the most stringent binding characteristics for each allotype (Table I).

The remaining 22 clusters are composed of peptides that are abundantly presented by more than one HLA-B27 allotype (Fig. 2; Supplemental Fig. 7). Of note, the consensus-binding motifs of these clusters show a mixture of some but not all allotype-favored peptide features of the respective HLA-B27 subtypes (Table I). For example, cluster 29 represents peptides that are highly presented by HLA-B*27:03, HLA-B*27:04, and HLA-B*27:06. Their consensus-binding motif shows a preferred usage of His, Arg, or Lys at P1 (HLA-B*27:03-favored); Leu at P1–2 (HLA-B*27:04-favored); and Gly, Ser, or Ala at P1–1 (HLA-B*27:06-favored). These observations suggest that the combination of allotype-specific features present in a given peptide dictates (at least to some extent) its abundance levels across the different HLA-B27 allotypes. Moreover, these features can provide a means to estimate the relative abundance level of a given peptide bound by each HLA-B27 allotype.

Identification of Candidate Arthritogenic Peptides—To identify possible arthritogenic peptides, we screened our quantitative dataset for peptides that are presented in less amounts by HLA-B*27:06 and HLA-B*27:09 compared with disease-associated HLA-B27 subtypes. A total of 26 peptides fulfilled this criterion (Table II; Supplemental Table 5). The majority of these peptides carry a Tyr or Phe residue at their C terminus but do not contain other obvious similarities in addition to the conserved Arg at P2. To assess the degree by which these peptides are less presented by HLA-B*27:06 and HLA-B*27:09, we calculated for each peptide the quantitative difference to the disease-associated subtype with the lowest relative expression level (Supplemental Fig. 10). On average, this difference was 3.9% (±4.4%) with individual values varying between less than 1% to more than 20% (Supplemental Fig. 10). This reflects the complexity of the dataset and is consistent with our view that qualitative differences are rarely seen between HLA-B27 allotypic peptide ligands but rather there are subtle to quite substantial quantitative changes in peptide abundance. To highlight this, we performed an identical calculation across all 1646 quantified peptides and obtained an average quantitative difference of 6.7% (±7.5%) between the three lowest relative expression levels on different HLA B27 allotypes. Only two peptides considered as candidate arthritogenic peptides (ARYVFQSENTFD and ARVLLVPDNTF) exceeded this error region in such that the quantitative difference between HLA-B*27:06 or HLA-B*27:09 to a disease-associated subtype was significantly greater than the average quantitative difference between the three lowest relative expression levels of a random peptide (Supplemental Fig. 10). However, the noise from allotypic binding preferences almost certainly masks the candidate peptides and even subtle quantitative differences are sufficient to cause a differential T cell response leading to the onset and progression of AS. Therefore, we believe each of these 26 peptides might represent the potential arthritogenic self-peptide and should be analyzed in future studies.
For this work, we generated such a spectral library in the open-source software Skyline (40) based on all HLA-B27 peptides identified in our previous study (18). Instead of spiking in peptides for retention time alignment, we exploited the presence of intrinsic peptides that were identiﬁed in each of the different HLA-B27 DDA acquisitions. Based on these peptides, we developed a custom-made “iRT-B27” retention time predictor (Supplemental Fig. 1) that allowed us to quantify 1095 peptides by SWATH-MS across all HLA-B27 allotypes.

**TABLE I**

Cluster properties

| Cluster | # of peptide | Highest expression levels on | HLA-B27 allotype-favored peptide features |
|---------|--------------|------------------------------|------------------------------------------|
|         |              | B*27:02                      | B*27:03                                  |
| 1       | 96           | B*27:02                      | R @ P1 W @ P0                           |
| 2       | 128          | B*27:03                      | H, K, R @ P1 L @ P3 V @ P0              |
| 3       | 64           | B*27:04                      | L @ P2 I @ P0                           |
| 4       | 71           | B*27:05                      | Q @ P0 S, G @ P0 I @ P0                 |
| 5       | 91           | B*27:06                      | N @ P0                                 |
| 6       | 57           | B*27:07                      | F @ P0                                 |
| 7       | 87           | B*27:08                      | V @ P0                                 |
| 8       | 66           | B*27:09                      | A, E, T, V @ P0                         |
| 9       | 83           | B*27:02 + B*27:03 + B*27:04 | W @ P0                                 |
| 10      | 29           | B*27:02                      | V @ P0                                 |
| 11      | 40           | B*27:02 + B*27:04 + B*27:06 | L @ P2                                 |
| 12      | 43           | B*27:05                      | I @ P0                                 |
| 13      | 23           | B*27:02 + B*27:06 + B*27:08 | S, G @ P0                               |
| 14      | 21           | B*27:07 + B*27:08 + B*27:09 | A, E, T, V @ P0                         |
| 15      | 41           | B*27:03 + B*27:04 + B*27:09 | V @ P0                                 |
| 16      | 43           | B*27:03 + B*27:07 + B*27:08 | L @ P2                                 |
| 17      | 60           | B*27:03 + B*27:05 + B*27:09 | I @ P0                                 |
| 18      | 49           | B*27:03 + B*27:05 + B*27:09 | L @ P0                                 |
| 19      | 47           | B*27:03 + B*27:05 + B*27:06 | H, K, R @ P1                           |
| 20      | 42           | B*27:03 + B*27:05 + B*27:06 | L @ P0                                 |
| 21      | 24           | B*27:03 + B*27:05 + B*27:06 | H, K, R @ P1                           |
| 22      | 21           | B*27:03 + B*27:05 + B*27:09 | A, E, T, V @ P0                         |
| 23      | 27           | B*27:04 + B*27:05 + B*27:06 | L @ P2                                 |
| 24      | 21           | B*27:04 + B*27:05 + B*27:06 | I @ P0                                 |
| 25      | 38           | B*27:04 + B*27:05 + B*27:09 | L @ P2                                 |
| 26      | 43           | B*27:05 + B*27:09 + B*27:09 | S, G @ P0                               |
| 27      | 45           | B*27:05 + B*27:09 + B*27:09 | I @ P0                                 |
| 28      | 73           | B*27:05 + B*27:09 + B*27:09 | L @ P2                                 |
| 29      | 64           | B*27:05 + B*27:09 + B*27:09 | S, G @ P0                               |
| 30      | 17           | B*27:05 + B*27:09 + B*27:09 | A, E, T, V @ P0                         |
strategy, quantitative measurements of additional 551 peptides, which samples in only 152 individual MRM-MS runs resulting in quantification measurements and the inability to unambiguously determine the presence (or absence) of a peptide in the background region, which inevitably leads to erroneous software might pick an incorrect peak or an unspecific signal sensitivity threshold of detection. In those cases, the analysis dance levels are very low and/or only marginally exceed the sensitivity. Thus, it is not surprising that several individual peptides within these 22 clusters show unexpected primary sequences, which are not in agreement with the allotype-favored peptide features of the respective cluster. Nevertheless, this cluster analysis describes and highlights an easy, empirical, and universal approach to attain the most stringent binding characteristics for a given HLA molecule and by-

### Table II

| Peptide                | Length | PΩ |
|------------------------|--------|----|
| GRVLLPEGGITA           | 12     | A  |
| HRAFSVFLF             | 9      | F  |
| SRVAEMFLF             | 9      | F  |
| TRYDLYHTF             | 9      | F  |
| ERLKDLFGKF           | 10     | F  |
| QRWDEAFRKF           | 10     | F  |
| ARVLLVPONTF          | 11     | F  |
| ARVVFQSENTF          | 11     | F  |
| IRFPLMTIEEF           | 11     | F  |
| QRIDLVPPFT           | 11     | F  |
| SRWKIPSSW            | 9      | W  |
| TRGDLSIRW            | 10     | W  |
| SRLGVSDVITW          | 12     | W  |
| HRAPPIGY             | 9      | Y  |
| IRLVPLFY            | 9      | Y  |
| LFHPIHILAY           | 9      | Y  |
| QRYLVTYY             | 9      | Y  |
| QRLKIMEYY            | 9      | Y  |
| YQOWLKEEY           | 9      | Y  |
| FRFGSLLLGYY         | 10     | Y  |
| MRTIALLEY            | 10     | Y  |
| SRFHHTPSLY          | 10     | Y  |
| ARTTINEOY           | 11     | Y  |
| ARVGSLSWNSY          | 11     | Y  |
| GRFGYLEQQEY         | 11     | Y  |
| KRGFGPEGSVELY       | 13     | Y  |

MRM-MS - in comparison to SWATH-MS - is a more targeted approach and excels in unparalleled sensitivity but is limited by the number of peptides that can be analyzed in one run. We therefore opted to utilize MRM-MS to reanalyze only those peptides that were not identified by SWATH-MS. In addition, we did not measure the peptides of the iRT-B27 retention time predictor in each MRM-MS run but relied on a predetermined, MRM-gradient-specific regression curve of the iRT-B27 peptides for retention time alignment (correlation coefficient R² = 0.9776; Supplemental Fig. 2). Pursuing this strategy, we were able to analyze 2211 peptides across eight samples in only 152 individual MRM-MS runs resulting in quantitative measurements of additional 551 peptides, which constitute ~25% of the peptides that escaped quantification by SWATH-MS.

Quantifying peptide ligands across various HLA molecules is hampered by the low abundance of some shared HLA-associated peptides and that not all peptides are expected to be present in each sample. Despite significant advances in MS instrumentation during the last decade, faithful quantitative measurements are not possible for peptides whose abundance levels are very low and/or only marginally exceed the sensitivity threshold of detection. In those cases, the analysis software might pick an incorrect peak or an unspecific signal in the background region, which inevitably leads to erroneous quantification measurements and the inability to unambiguously determine the presence (or absence) of a peptide in a given sample. To attenuate this issue, we manually assessed the performance of the peak-picking algorithm of all peptides and corrected the peak boundaries as necessary. Moreover, we applied strict evaluation criteria to our comprehensive datasets and annotated all peptides, which did not fulfill these criteria in some samples, accordingly in Supplemental Table 3.

Through the utility of the TM4:MeV software suite (43), we grouped all quantified peptides into 30 different clusters according to their relative binding to the eight HLA-B27 allotypes. Instead of accepting missing values in our cluster analysis, we opted to use all quantified values irrespective of whether a peptide assignment in a given sample passed our stringent evaluation criteria or not. This decision is based on the assumption that a peptide, which failed our evaluation criteria in a sample, is most likely either absent or of extremely low abundance. Either way, the quantified (background) signal is sufficiently low to not perturb the results of the cluster analysis.

The presence of clusters that contained peptides, which were exclusively presented by only one HLA-B27 allotype, enabled us to identify the most stringent binding characteristics for each HLA-B27 allotype (Table I). As expected, most of these allotype-favored peptide features are present in previously determined canonical consensus-binding motifs (4–18). However, not all features identified by canonical consensus-binding motifs represent allotype-favored peptide features, which highlights the importance of allotype-specific cluster analyses. Of note, allotype-favored peptide features also suggest the preferred HLA-B27 subtype of generally tolerated features. For example, all HLA-B27 allotypes are known to permit binding of peptides carrying a C-terminal Leu residue. However, it appears that HLA-B*27:06 is the most-suitable subtype to accommodate this peptide feature.

The remaining 22 clusters, which contained peptides that are abundantly presented by more than one HLA-B27 allotype, showed a mixture of allotype-favored peptide features of the respective HLA-B27 subtypes (Table I). This observation suggests that (allotype-favored) peptide features represent a key element in determining the abundance levels of peptides across different HLA allotypes. However, not all individual peptides within these 22 clusters show appropriate peptide features implying that other factors such as (i) secondary structure formations, (ii) conformational changes, (iii) availability and prevalence of source proteins, and (iv) differential interactions with components of the peptide-loading complex act in concert to regulate peptide binding and abundance. Thus, it is not surprising that several individual peptides within these 22 clusters show unexpected primary sequences, which are not in agreement with the allotype-favored peptide features of the respective cluster. Nevertheless, this cluster analysis describes and highlights an easy, empirical, and universal approach to attain the most stringent binding characteristics for a given HLA molecule and by-
passes the requirement for complex models and prediction algorithms, which often fail to produce unambiguous results. Unraveling the determinant selection of closely related HLA molecules will lead to a better understanding of a panel of clinically relevant topics such as allograft reactions, hypersensitivity reactions and other autoimmune diseases.

A detailed analysis of nested sets of N- and C-terminally extended peptides suggested that most peptides are able to anchor with their termini into the binding cleft of HLA-B27 subtypes resulting in central bulging rather than terminal protrusion. This notion, which also has been proposed previously (15, 18), is based on the observation that the abundance levels of peptides across the HLA-B27 subtypes only correlate with the presence of allotype-specific feature(s) if these peptides anchor with their termini into the binding cleft. While this is true for nearly all peptides that are extended by only one or two amino acids, some unusually long peptides (extensions of more than three amino acids resulting in peptides longer than 15 amino acids) appear to have similar abundance distribution patterns as their base peptide. This observation may suggest the usage of identical anchor residues and implies that the termini of these peptides stick out from the binding cleft resulting in terminal protrusions. This is particularly evident in case of the nested set surrounding the base peptide ARLALSPVPSH, which is mainly presented by HLA-B*27:05 (Supplemental Table 4). As expected, ARLALSPVPSHW, which carries a one amino acid C-terminal Trp extension, is predominantly presented by HLA-B*27:02. An additional Met extension at the C terminus (ARLALSPVPHWWM), which is a HLA-B*27:08-favored feature, completely abolishes presentation by HLA-B*27:02 in favor of HLA-B*27:08. All further C-terminally extended peptides (ARLALSPVPHWVMVA and ARLALSPVPHWVMWVMA) show a nearly identical distribution pattern across the HLA-B27 allotypes as ARLALSPVPHWWM. Although additional experiments such as binding and refolding assays have to be carried out to support this notion, we speculate that the vast majority of peptides bind with their termini into the peptide-binding groove resulting in central bulging.

The main goal of this study was the identification of peptides that are less presented by HLA-B*27:06 and HLA-B*27:09 compared with disease-associated HLA-B27 subtypes. We identified 26 peptides fulfilling this criterion and B*27:09 compared with disease-associated HLA-B27 subtypes that are less presented by HLA-B*27:06 and HLA-B*27:08-favored feature, completely abolishes presentation by HLA-B*27:02. An additional Met extension at the C terminus (ARLALSPVPHWWM), which is a HLA-B*27:08-favored feature, completely abolishes presentation by HLA-B*27:02 in favor of HLA-B*27:08. All further C-terminally extended peptides (ARLALSPVPHWVMVA and ARLALSPVPHWVMWVMA) show a nearly identical distribution pattern across the HLA-B27 allotypes as ARLALSPVPHWWM. Although additional experiments such as binding and refolding assays have to be carried out to support this notion, we speculate that the vast majority of peptides bind with their termini into the peptide-binding groove resulting in central bulging.

In summary, we utilized a combination of targeted mass spectrometry techniques to identify 26 candidate arthritogenic peptides whose immunogenic potential will be analyzed in further immunological and structural studies.

* This work was supported by project grants 1022509 and 1085017 from the National Health and Medical Research Council of Australia (NH&MRC). A.W.P. is an NH&MRC Senior Research Fellow. R.B.S. acknowledges support from the Swiss National Science Foundation. S This article contains supplemental material. ¶ To whom correspondence should be addressed: E-mail: anthony.purcell@monash.edu.

REFERENCES

1. Garrett, T. P., Saper, M. A., Bjorkman, P. J., Strominger, J. L., and Wiley, D. C. (1989) Specificity pockets for the side chains of peptide antigens in HLA-Aw68. Nature 342, 692–696
2. Khan, M. A., Mathieu, A., Sorrentino, R., and Akkoc, N. (2007) The patho- genetic role of HLA-B27 and its subtypes. Autoimmun. Rev. 6, 183–189
3. Khan, M. A. (2013) Polymorphism of HLA-B27: 105 subtypes currently known. Curr. Rheumatol. Rep. 15, 362
4. Rötzschke, O., Falk, K., Stevanovic, S., Gnau, V., Jung, G., and Ram- mensee, H. G. (1994) Dominant aromatic/aliphatic C-terminal anchor in HLA-B2702 and B2705 peptide motifs. Immunogenetics 39, 74–77
5. Colbert, R. A., Rowland-Jones, S. L., McMichael, A. J., and Frelinger, J. A. (1994) Differences in peptide presentation between B27 subtypes: The importance of the P1 side chain in maintaining high affinity peptide binding to B2703. Immunity 1, 121–130
6. Boisgérault, F., Tieng, V., Stolzenberg, M. C., Dulphy, N., Khalil, I., Ta- mouza, R., Charro, D., and Toubert, A. (1996) Differences in endoge- nous peptides presented by HLA-B2702 and B2705 allelic variants. Implications for susceptibility to spondylarthropathies. J. Clin. Investiga- tion 98, 2764–2770
7. Fiorillo, M. T., Meadova, L., D’Amato, M., Shabanowitz, J., Hunt, D. F., Appella, E., and Sorrentino, R. (1997) Susceptibility to ankylosing spondylitis correlates with the C-terminal residue of peptides presented by various HLA-B27 subtypes. Eur. J. Immunol. 27, 368–373
8. Griffin, T. A., Yuan, J., Friede, T., Stevanovic, S., Ariyoshi, K., Rowland- Jones, S. L., Rammensee, H. G., and Colbert, R. A. (1997) Naturally occurring A-pocket polymorphism in HLA-B2703 increases the dependence on an accessory anchor residue at P1 for optimal binding of nonamer peptides. J. Immunol. 159, 4887–4897
9. Tieng, V., Dulphy, N., Boisgérault, F., Tamouza, R., Charro, D., and Toub- ert, A. (1997) HLA-B2707 peptide motif: Tyr C-terminal anchor is not shared by all disease-associated subtypes. Immunogenetics 47, 103–105
10. Marti, M., Alvarez, I., Montserrat, V., and Lopez de Castro, J. A. (2001) Large sharing of T-cell epitopes and natural ligands between HLA-B27 subtypes (B2702 and B2705) associated with spondyloarthrosis. Tissue Antigens 58, 351–362
11. Sesma, L., Montserrat, V., Lamas, J. R., Marina, A., Vazquez, J., and Lopez de Castro, J. A. (2002) The peptide repertoires of HLA-B27 subtypes differentially associated to spondylarthropathy (B2704 and B2706) differ by specific changes at three anchor positions. J. Biolog. Chem. 277, 16744–16749
12. Ramos, M., Paradelo, A., Vazquez, M., Marina, A., Vazquez, J., and Lopez de Castro, J. A. (2002) Differential association of HLA-B2705 and B2709 to ankylosing spondylitis correlates with limited peptide subsets but not with altered cell surface stability. J. Biolog. Chem. 277, 28749–28756
13. Lopez de Castro, J. A., Alvarez, I., Marcilla, M., Paradelo, A., Ramos, M., Sesma, L., and Vázquez, M. (2004) HLA-B27: A registry of constitutive peptide ligands. Tissue Antigens 63, 424–445
14. Gómez, P., Montserrat, V., Marcilla, M., Paradelo, A., and de Castro, J. A. (2006) B2707 differs in peptide specificity from B2705 and B2704 as much as from HLA-B27 subtypes not associated to spondyloarthrits.

Molecular & Cellular Proteomics 15.6
Quantification of HLA-B27 Peptides by SWATH- and MRM-MS

Eur. J. Immunol. 36, 1867–1881
15. Ben Dror, L., Barnea, E., Beer, I., Mann, M., and Admon, A. (2010) The HLA-B2705 peptidome. Arthritis Rheumatism 62, 420–429

16. Infantes, S., Lorente, E., Barnea, E., Beer, I., Barriga, A., Lasala, F., Jimenez, M., Admon, A., and Lopez, D. (2013) Natural HLA-B2705 protein ligands with galactose as amino anchor: Implications for HLA-B27 association with spondyloarthropathy. J. Biol. Chem. 288, 10882–10896

17. Garcia-Medel, N., Sanz-Bravo, A., Alvarez-Navarro, C., Gomez-Molina, P., Barnea, E., Marcilla, M., Admon, A., and de Castro, J. A. (2014) Peptide handling by HLA-B27 subtypes influences their biological behavior, association with ankylosing spondylitis and susceptibility to endoplasmic reticulum aminopeptidase 1 (ERAP1). Mol. Cell. Proteomics 13, 3367–3380

18. Schittenhelm, R. B., Sian, T. C., Wilmann, P. G., Dudek, N. L., and Purcell, A. W. (2015) Revisiting the arthritogenic peptide theory: Quantitative not qualitative changes in the peptide repertoire of HLA-B27 allotypes. Arthritis Rheumatol. 67, 702–713

19. Madden, D. R., Gorga, J. C., Strominger, J. L., and Wiley, D. C. (1991) The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. Nature 353, 321–325

20. Jardetzky, T. S., Lane, W. S., Robinson, P. A., Madden, D. R., and Wiley, D. C. (1991) Identification of self peptides bound to purified HLA-B27. Nature 353, 326–329

21. Madden, D. R., Gorga, J. C., Strominger, J. L., and Wiley, D. C. (1992) The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. Cell 70, 1035–1048

22. Breban, M., Miceli-Richard, C., Zinovieva, E., Monnet, D., and Said-Nahal, R. (2006) The genetics of spondyloarthropathies. Joint. Bone, Spine 73, 355–362

23. Vegvar, A., Szabo, Z., Szanto, S., Glant, T. T., Mikecz, K., and Szekanecz, Z. (2009) The genetic background of ankylosing spondylitis. Joint. Bone, Spine 76, 623–628

24. Schlioster, L., Terasaki, P. I., Bluestone, R., and Pearson, C. M. (1973) High association of an HLA antigen, W27, with ankylosing spondylitis. New Eng. J. Med. 288, 704–706

25. Brewerton, D. A., Hart, F. D., Nicholls, A., Caffrey, M., James, D. C., and Sturrock, R. D. (1973) Ankylosing spondylitis and HLA-A 27. Lancet 1, 904–907

26. López de Castro, J. A. (2007) HLA-B27 and the pathogenesis of spondyloarthropathies. Immunol. Lett. 108, 27–33

27. Bowness, P. (2015) Hla-B27. Annu. Rev. Immunol. 33, 29–48

28. Hermann, E., Yu, D. T., Meyer zum Büschenfelde, K. H., and Fleischer, B. (1993) HLA-B27-restricted CD8 T cells from synovial fluids of patients with reactive arthritis and ankylosing spondylitis. J. Immunol. 142, 646–650

29. Faust, H., von Goessel, H., May, E., and Märker-Hermann, E. (2003) HLA-B27-restricted T cells from patients with ankylosing spondylitis recognize peptides from B*2705 that are similar to bacteria-derived proteins. J. Autoimmun. 20, 169–176

30. López-Larrea, C., Sujirachato, K., Menha, N. K., Chiewsilp, P., Isarangkura, D., Kang, U., Dominguez, O., Coto, E., Pená, M., and Setién, F. (1995) HLA-B27 subtypes in Asian patients with ankylosing spondylitis. Evidence for new associations. Tissue Antigens 45, 169–176

31. D’Amato, M., Fiorillo, M. T., Galeazzi, M., Martinetti, M., Amoroso, A., and Sorrentino, R. (1995) Frequency of the new HLA-B*2709 allele in ankylosing spondylitis patients and healthy individuals. Disease Markers 12, 215–217

32. Ren, E. C., Koh, W. H., Sim, D., Boey, M. L., Wee, G. B., and Chan, S. H. (1997) Possible protective role of HLA-B*2706 for ankylosing spondylitis. Tissue Antigens 49, 67–69

33. Nasution, A. R., Mardjadi, A., Kunmartini, S., Surjadinana, N. G., Setyo-hadi, B., Sudarsono, D., Lardy, N. M., and Feltkamp, T. E. (1997) HLA-B27 subtypes positively and negatively associated with spondyloarthropathy. J. Rheumatol. 24, 1111–1114

34. Paladini, F., Taccari, E., Fiorillo, M. T., Cauì, A., Passiu, G., Mathieu, A., Punzi, L., Lapadula, G., Scarpa, R., and Sorrentino, R. (2005) Distribution of HLA-B27 subtypes in Sicaria and correlation with their association with spondyloarthropathies. Arthritis Rheumatism 52, 3319–3321

35. Gilleit, L. C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., Bonner, R., and Aebersold, R. (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: A new concept for consistent and accurate proteome analysis. Mol. Cell. Proteomics 11, O111.016717

36. Tate, S., Larsen, B., Bonner, R., and Gingeras, A. C. (2013) Label-free quantitative proteomics trends for protein-protein interactions. J. Proteins 81, 91–101

37. Tsou, C. C., Antonovom, D., Larsen, B., Tucholska, M., Choi, H., Gingeras, A. C., and Nesvizhskii, A. I. (2015) DIA-Umpire: Comprehensive computational framework for data-independent acquisition proteomics. Nat. Methods 12, 258–264, 257 p following 264

38. Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008) Selected reaction monitoring for quantitative proteomics: A tutorial. Mol. Syst. Biol. 4, 222

39. Doer, A. (2013) Mass spectrometry-based targeted proteomics. Nat. Methods 10, 23

40. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. Nat. Biotechnol. 28, 966–968

41. Croft, N. P., Smith, A. S., Wong, Y. C., Tan, C. T., Dudek, N. L., Fleisch, I. E., Lin, L. C., Tscharke, D. C., and Purcell, A. W. (2013) Kinetics of antigen expression and epitope presentation during virus infection. PLoS Pathogens 9, e1003129

42. Croft, N. P., de Verteuil, D. A., Smith, A. S., Wong, Y. C., Schittenhelm, R. B., Tscharke, D. C., and Purcell, A. W. (2015) Simultaneous quantification of viral antigen expression kinetics using data-independent (DIA) mass spectrometry. Mol. Cell. Proteomics 14, 1361–1372

43. Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturm, A., Snuffin, M., Rezantsev, A., Popov, P., Rytsov, A., Kosukovich, E., Borisyovski, V., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. (2003) TM4: A free, open-source system for microarray data management and analysis. BioTechniques 34, 374–378

44. Schittenhelm, R. B., Dudek, N. L., Croft, N. P., Ramarathnamin, S. H., and Purcell, A. W. (2014) A comprehensive analysis of constitutive naturally processed and presented HLA-C*04:01 (Cw4)-specific peptides. Tissue Antigens 83, 174–179

45. Yeung, K. Y., Haynor, D. R., and Ruzzo, W. L. (2001) Validating clustering for gene expression data. Bioinformatics 17, 309–318

46. Escher, C., Reiter, L., MacLean, B., Ossola, R., Herzog, F., Chilton, J., MacCoss, M. J., and Rinner, O. (2012) Using IRT, a normalized retention time for more targeted measurement of peptides. Proteomics 12, 1111–1121

47. García, F., Marina, A., and López de Castro, J. A. (1997) Lack of carboxyl-terminal tyrosine distinguishes the B*2706-bound peptide repertoire from those of B*2704 and other HLA-B27 subtypes associated with ankylosing spondylitis. Tissue Antigens 49, 215–221