Published in final edited form as:

J Proteome Res. 2017 January 06; 16(1): 122–136. doi:10.1021/acs.jproteome.6b00386.

Immunogenic HLA-DR-Presented Self-Peptides Identified Directly from Clinical Samples of Synovial Tissue, Synovial Fluid, or Peripheral Blood in Patients with Rheumatoid Arthritis or Lyme Arthritis

Qi Wang^{1,#}, Elise E. Drouin^{2,#}, Chunxiang Yao¹, Jiyang Zhang^{1,3}, Yu Huang¹, Deborah R. Leon¹, Allen C. Steere², and Catherine E. Costello^{1,*}

¹Center for Biomedical Mass Spectrometry, Department of Biochemistry, Boston University School of Medicine, Boston, Massachusetts 02118

²Center for Immunology and Inflammatory Diseases, Division of Rheumatology, Allergy and Immunology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114

³National University of Defense Technology, Changsha, Hunan Province, China

Abstract

HLA-DR-molecules are highly expressed in synovial tissue (ST), the target of the immune response in chronic inflammatory forms of arthritis. Here, we used LC-MS/MS to identify HLA-DR-presented self-peptides in cells taken directly from clinical samples: ST, synovial fluid mononuclear cells (SFMC), or peripheral blood mononuclear cells (PBMC) from five patients with rheumatoid arthritis (RA) and eight with Lyme arthritis (LA). We identified 1,593 non-redundant HLA-DR-presented peptides, derived from 870 source proteins. Sixty-seven percent of the peptides identified in SFMC and 55% of those found in PBMC were found in ST, but analysis of SFMC/PBMC also revealed new antigen-presented peptides. Peptides were synthesized and examined for reactivity with the patients' PBMC. To date, three autoantigens in RA and four novel autoantigens in LA, presented in ST and/or PBMC, were shown to be targets of T- and B-cell responses in these diseases; ongoing analyses may add to this list. Thus, immunoprecipitation and

*Corresponding Author. To whom correspondence should be addressed: Prof. Catherine E. Costello, Center for Biomedical Mass Spectrometry, Boston University School of Medicine, 670 Albany St., Rm. 511, Boston, MA 02118-2646. Tel.: 617-638-6490; Fax: 617-638-6491; cecmsms@bu.edu.

#These authors made equal contributions to the reported research.

Supporting Information

Supplementary Table 1: Lists of consensus peptides identified by at least two search engines.

Supplementary Table 2: Lists of unique peptides identified by consensus matches by at least two search engines.

Supplementary Table 3: Lists of peptides identified by only one program, passed manual inspection of tandem MS quality, and tested for immune responses.

Supplementary Table 4: Lists of the non-redundant peptides identified from each sample.

Supplementary Table 5: Lists of source proteins identified from each sample.

Supplementary labeled mass spectra: The labeled tandem mass spectra were listed for peptides identified by at least two protein database search engines, as well as peptides identified by one protein database search engine with the assignment approved upon manual inspection. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

LC-MS/MS can now identify hundreds of HLA-DR-presented self-peptides from individual patients' tissues or fluids with mixed cell populations. Importantly, identification of HLA-DR-presented peptides from SFMC or PBMC allows testing of more patients, including those early in the disease. Direct analysis of clinical samples facilitates identification of novel immunogenic T-cell epitopes. Data are available via ProteomeXchange (identifier PXD003051).

for TOC only

Keywords

HLA-DR; synovial tissue; synovial fluid mononuclear cells; peripheral blood mononuclear cells; Lyme arthritis; rheumatoid arthritis; autoantigens

INTRODUCTION

Human leukocyte antigen – antigen D related (HLA-DR) molecules play a central role in autoimmune diseases, as first reported by McDevitt in 1974,¹ HLA-DR molecules are highly expressed in synovial tissue (ST), the target of the immune response in chronic inflammatory forms of arthritis, including rheumatoid arthritis (RA) and Lyme arthritis (LA).^{2, 3} The greatest known genetic risk factors for these diseases are certain HLA-DR alleles.⁴ HLA-DR molecules present foreign or self-peptides to CD4+ T cells, leading to their activation.⁵ After activation, T cells can supply the necessary help to B cells to produce sustained, high-titer antibodies against the same antigen. In RA, 10 HLA-DR alleles, including the HLA-DRB1*0101 and 0401 alleles,^{6, 7} convey the greatest risk. These alleles code for the RA shared epitope, a stretch of five amino acids that contribute to the P4 binding pocket of the HLA-DR molecule. Anti-citrullinated protein antibodies (ACPA) which are the only known specific autoantibodies in RA, are found in approximately 60% of RA patients, whereas the remaining 40% lack this specific diagnostic marker. ACPA are significantly more frequent in patients who have the shared epitope.^{8–13}

LA, another type of chronic inflammatory arthritis, results from infection with the tick-borne spirochete, *Borrelia burgdorferi* (Bb).^{14, 15} Most patients with this late stage manifestation respond well to antibiotic therapy, and are designated as having antibiotic-responsive LA.¹⁶ However, some patients experience persistent synovitis that lasts for months or several years, despite treatment with 2–3 months of oral and intravenous (IV) antibiotic therapy, and their disease is termed antibiotic-refractory LA.¹⁷ Rather than persistent infection, this latter outcome is thought to result from excessive inflammation,^{18, 19} immune dysregulation,^{20, 21} and site-specific autoimmunity in joints.^{22–26} In LA, seven HLA-DR alleles have been identified, including the DRB1*0101, 0401, and 1501 alleles, that are more common in patients with antibiotic-refractory LA than in those with antibiotic-responsive LA.²⁷

In both RA and LA, synovial tissue shows marked synovial hypertrophy, vascular proliferation, mononuclear cell infiltration, and intense expression of HLA-DR molecules on synoviocytes and infiltrating cells.^{2, 3} The tissue is bathed in synovial fluid (SF), which contains a rich mixture of inflammatory cells, including large numbers of dendritic cells and macrophages; the latter also have increased expression of HLA-DR molecules.^{28, 29} However, until our earlier publication,³⁰ little was known about the specificities of self-peptides presented by HLA-DR molecules in ST in these diseases.

With the advent of highly sensitive, nanoflow LC-MS/MS-based peptide sequencing,³¹ it has become possible to identify peptides presented by HLA-DR molecules in individual patients' tissues or cells. In 1992, Hunt et al. were the first to use LC-MS/MS to identify class I major histocompatibility complex (MHC) molecules in tumor tissue culture systems.³² They have continued to lead this rapidly growing field, and the identification of MHC-presented peptides has been widely undertaken, primarily using cell lines.³³⁻⁴¹ Moreover, advances in mass spectrometry techniques and bioinformatics have improved the efficiency and accuracy for identification of MHC-presented molecules.^{42, 43}

As a next step, a few laboratories have undertaken the direct analysis of class II HLA-DR-presented peptides in clinical samples of mixed cell lines.⁴⁴⁻⁴⁹ In our initial study, we were able to identify 57–104 non-redundant, HLA-DR-presented self-peptides per patient from the ST of four patients, two with RA and two with LA.³⁰ However, large amounts of synovial tissue can only be obtained when patients undergo surgical procedures, such as synovectomy or joint replacement, and this limits the analysis to patients at the severe end of the spectrum, usually seen late in the illness.

In our current study, we increased the number of patients studied to thirteen, five with RA and eight with LA. Moreover, we demonstrated that HLA-DR-presented peptides could be identified in individual patients - not only from ST, but also from synovial fluid mononuclear cells (SFMC) and peripheral blood mononuclear cells (PBMC), which allow the analysis not only of a greater number of patients, but also of patients seen earlier in the illness. In addition, as our study progressed, the increasing sensitivity of new mass spectrometers resulted in the identification of greater numbers of HLA-DR-presented peptides from each sample. In the current study, we analyzed 18 new ST, PBMC and SFMC samples and we re-analyzed previous LC-MS/MS data to compare the early results with new data from SFMC and PBMC samples obtained from the same patients.

Altogether, from 22 samples that originated from 13 patients, we identified 1,593 non-redundant HLA-DR-presented, self-peptides derived from 870 source proteins. This experience shows that LC-MS/MS now has the sensitivity and specificity to identify large numbers of HLA-DR-presented self-peptides in tissues obtained from individual patients, and analysis of SFMC/PBMC also revealed new peptide antigens. To date, the identified HLA-DR-presented peptides from five of the eight patients have been synthesized and tested with the matching patient's PBMC. We determined that among these peptides were novel immunogenic T cell epitopes derived from the source proteins endothelial cell growth factor (ECGF),²² apolipoprotein B-100 (apoB-100),²³ matrix metalloproteinase-10 (MMP-10)²⁴ in LA; from annexin A2 in both RA and LA,²⁵ and from *N*-acetylglucosamine-6-sulfatase and

filamin A in RA.²⁶ Furthermore, we then showed that these epitopes or their source proteins were targets of T and B cell responses in many patients.^{22–26} Ongoing immunological analyses may reveal further epitopes.

EXPERIMENTAL PROCEDURES

Study Patients

The study was approved by the Human Investigations Committee at Massachusetts General Hospital, and all participants gave written informed consent. The five RA patients met the 2010 criteria of the American College of Rheumatology/ European League against Rheumatism (ACR/EULAR) definition for the diagnosis of rheumatoid arthritis,⁵⁰ and the eight LA patients met the criteria of the Centers for Disease Control and Prevention for the diagnosis of Lyme disease.⁵¹ All eight LA patients had antibiotic-refractory LA, defined as persistent arthritis for three months after the start of two months of appropriate oral antibiotic therapy and one month of IV antibiotic therapy.⁵² In LA patients, synovial tissue was obtained at the time of therapeutic, arthroscopic synovectomies, and, in RA patients, the tissue was collected at the time of synovectomy or joint replacement surgery. Sufficient quantities of SFMC were available for analysis in six patients and PBMC were available in four patients.

ST, SFMC, and PBMC Preparation

All the reagents were purchased from Sigma-Aldrich (St. Louis, MO) except where noted. ST samples (8–17 g) were prepared using the protocol described previously.³⁰ Heparinized SFMC and PBMC (~ 3 to 8.7×10^7 cells) were obtained by Ficoll-Hypaque separation (MP Biomedicals) and stored in liquid nitrogen prior to analysis. On the day of purification, SFMC and PBMC samples were quickly thawed by placement in a 37 °C water bath for two minutes. Cells were washed twice in phosphate buffered saline (PBS) (Life Technologies) by centrifuging at $800 \times g$ for 10 min. The pellet was re-suspended with 10 mL lysis buffer: 150 mM sodium chloride, 20 mM tris(hydroxymethyl)aminomethane•HCl (pH 8.0), 5 mM ethylenediaminetetraacetic acid disodium solution, 0.04% sodium azide, 1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride•HCl, 10 µg/mL leupeptin, 10 µg/mL pepstatin A, 5 µg/mL aprotinin, and 1% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate. Immunoaffinity purifications of HLA-DR complexes from SFMC and PBMC were performed using the protocol previously described for synovial tissue samples.³⁰ HLA-DR protein and the bound peptide complexes were eluted from L243-conjugated beads into seven fractions using 0.7 mL of 0.5% formic acid (FA) in each fraction. HLA-DR protein and the bound peptides were separated using Microcon 10,000-Da molecular weight cutoff filters (Millipore, Billerica, MA). The peptide fractions were dried under vacuum, followed by desalting using C₁₈ ZipTip (Millipore, Billerica, MA). The eluates from the seven desalted peptide fractions were combined, dried under vacuum, and reconstituted with 20 µL of 0.1% FA. For each LC-MS/MS run, 1 µL of this solution was used for synovial tissue samples, and 2 µL was used for SFMC or PBMC samples, without regard to which LC-MS/MS system was employed for the analysis.

LC-MS/MS

The methods used for each mass spectrometer are given below.

LTQ-Orbitrap XL mass spectrometer

A nanoAcquity UPLC system (Waters Corp., Milford, MA) was coupled to an LTQ-Orbitrap XL MS (Thermo Fisher Scientific, San Jose, CA) through a TriVersa NanoMate ion source (Advion, Ithaca, NY). Peptide enrichment was performed with a trapping column (180 $\mu\text{m} \times 20 \text{ mm}$, 5 μm 100 Å Symmetry C₁₈, Waters Corp) at a flow rate of 15 $\mu\text{L}/\text{min}$ for 1 min, and separation was achieved with a capillary column (150 $\mu\text{m} \times 10 \text{ cm}$, 1.7 μm 130 Å BEH C₁₈, Waters Corp). Buffer A contained 1% acetonitrile (ACN) and 0.1% formic acid in water, and buffer B contained 1% water and 0.1% formic acid in ACN. A linear gradient of buffer B from 2% to 40% over 52 min was used at a flow rate of 0.5 $\mu\text{L}/\text{min}$. The capillary voltage was set to 1.7 kV using the NanoMate, and the capillary temperature was set to 120 °C. The mass spectra were recorded over the range m/z 300–2000 at a resolution of 60,000 (the width of the peak at half its maximum height at m/z 300) at a scan rate of approximately 1.2 s/spectrum. Tandem MS was performed for the five most abundant, multiply-charged species in the mass spectra that had a signal intensity threshold of 8000 NL. The normalized collision energy for CID was set to 35%, and helium was used as the collision gas. MS/MS spectra were recorded with the LTQ XL linear ion trap. All spectra were recorded in profile mode. Two to five LC-MS/MS runs were performed for each sample, and all the RAW files for each sample were combined into one Mascot generic file using Proteome Discoverer 1.3 software (Thermo Fisher Scientific).

Q Exactive plus mass spectrometer

A nanoAcquity UPLC system was coupled to a Q Exactive plus MS (Thermo Fisher Scientific, San Jose, CA) through a TriVersa NanoMate ion source. The same peptide enrichment and separation columns, HPLC solvents and gradients, and NanoMate capillary voltage and temperature used on the LTQ-Orbitrap XL MS were applied on the Q Exactive plus MS. Survey MS scans were recorded over the range m/z 400–2000 with a resolution of 70,000 (at m/z 300). The 15 most abundant, multiply-charged ions were selected for higher-energy collisional induced dissociation (HCD) MS/MS with a resolution of 17,500, an isolation width of 1.5 m/z , and a normalized collision energy of 27%. All spectra were recorded in profile mode. One LC-MS/MS run was performed for each sample, and the RAW file for each sample was converted to a Mascot generic file using Proteome Discoverer 1.3 software (Thermo Fisher Scientific).

6550 iFunnel QTOF LC/MS

Two 1260 HPLC systems were coupled to a 6550 iFunnel QTOF MS through a chip cube interface (Agilent Technologies, Santa Clara, CA). A Polaris-HR-3C18 chip was used for peptide enrichment and separation. The flow rate on the enrichment column was set at 2 $\mu\text{L}/\text{min}$ with sample flush volume set as 4 μL . With the same solvents used on the Waters UPLC systems, a gradient of buffer B (from 2% to 5% over 0.1 min, 5% to 20% over 74.9 min, 20% to 40% over 10 min, 40% to 90% over 10 min, held at 90% buffer B for 5 min, returned to 2% buffer B and held for 10 min) was used at a flow rate of 0.3 $\mu\text{L}/\text{min}$ for

peptide separation. The capillary voltage was set to 1.7 to 1.95 kV, and the gas temperature was set to 225 °C. The survey mass spectra were recorded over the range m/z 295–1700 with acquisition rate of 8 spectra/second. Tandem mass spectra were recorded over the range m/z 50–1700 with an acquisition rate of 3 spectra/second. Precursor ion isolation was performed with a narrow isolation window of 1.3 m/z . Tandem MS was performed for the 20 most abundant, multiply-charged species in the mass spectra that had a signal intensity threshold of 5000 counts or a relative threshold of 0.001%. The collision energy for CID was set using the formula $[3.1 \times (m/z)/100 + 1]$ for doubly charged precursor ions and $[3.6 \times (m/z)/100 - 4.8]$ for triply and higher charged precursor ions. Active exclusion was enabled after recording one tandem spectrum and released after 0.15 min. All spectra were recorded in the centroid mode. One to two LC-MS/MS runs were performed for each sample, and all the data files for each sample were combined into one Mascot generic file using MassHunter Qualitative Analysis B.05.00 (Agilent Technologies, Santa Clara, CA).

Protein Database Searching

In our original study, the International Protein Index (IPI) database (72,155 protein sequences) was used to identify HLA-DR-presented peptides from ST samples from two patients each with RA or LA (identified here as patients RA1, RA2, LA1 and LA2).³⁰ Because the IPI database is now closed and no longer maintained, the spectra obtained for the samples from these four patients were reanalyzed using the UniProt protein database (95,500 entries, August 2010) and the findings were compared to our original study. Using the UniProt database, ~96–98% of the peptides identified were the same as those previously assigned using the IPI database. Since an SFMC sample from LA2, and SFMC and PBMC samples from RA2 were available for analysis in this study, the re-analyzed LC-MS/MS data from ST of RA1, RA2, LA1, and LA2 were included here for comparison of the HLA-DR-presented peptides repertoires in ST, SFMC, and PBMC. In the current study, HLA-DR-presented self-peptides were identified by searching the Mascot generic file from each sample against the UniProt human database (including both Swiss-Prot and TrEMBL databases, 95,500 entries, August 2010) concatenated with a decoy database. The decoy databases were generated by randomizing each protein sequence in the database using the Perl script decoy.pl (Matrix Science). The same protein database containing the decoy database was used by all three protein database search engines, Mascot 2.4.0 (Matrix Science), OMSSA Browser 2.1.1 (NCBI), and X!Tandem tandem-win-12-10-01-1 (The Global Proteome Machine Organization www.thegpm.org). For inclusion of a peptide sequence in our database, a consensus match was required (assignment of the sequence by >two search engines) or a match by only one program that was judged acceptable upon review by two experts. Cysteinylation of cysteine, deamidation of glutamine and asparagine, pyroglutamic acid from amino-terminal glutamine and glutamic acid, and oxidation of methionine were specified as the variable modifications in all searches. *S*-cysteinylation peptides have been identified in MHC class I- and class II-presented peptides^{53–56} and in our previous study.³⁰ In addition, the conversion of arginine to citrulline was included as a variable post-translational modification; peaks exhibiting the difference in molecular weight corresponding to arginine vs. citrulline (= 0.98402 Da) were searched for each arginine residue in the database.

For data acquired using the LTQ-Orbitrap XL MS, precursor ion error tolerance was set as 0.01 Da and fragment ion error tolerance was set as 0.5 Da; for data acquired using the Q Exactive plus MS, precursor ion error tolerance was set as 10 ppm and fragment ion error tolerance was set as 20 mmu; for data acquired using 6550 iFunnel QTOF MS, precursor ion error tolerance was set as 20 ppm and fragment ion error tolerance was set as 50 mmu. “No enzyme” was specified for all the protein database searches. In this study, we applied scores or E-values as threshold to filter the peptide identification results from three search engines: Mascot score 20; OMSSA e-value 0.01; and X!Tandem e-value 10. These thresholds were set at a lower stringency level to allow more candidate peptides into the second filter which was a: consensus matches by at least two search engines, accomplished by a Microsoft Access query. With the application of the two-level filter, we were able to obtain a sufficient list of HLA-DR-presented peptides with higher quality (shown by the low levels of decoy hits) for immunogenicity tests with limited amounts of patients’ cells.

Over the course of the study, UniProt human protein database was continuously updated. To validate the results obtained with the UniProt database presented herein, the spectra obtained from a recent PBMC sample from an RA patient were reanalyzed using a UniProt human database (Swiss-Prot database, 20,100 entries, January 2016) and 93% of the peptides identified with the updated database were the same as those assigned using the UniProt database used here.

Removal of redundant peptide sequences

The endogenous HLA-DR-presented peptides derived from the same source protein often contain overlapping sequences with different peptide lengths. For example, $^{116}\text{DHLKYVMLPVA}^{126}$, $^{116}\text{DHLKYVMLPVAD}^{127}$, and $^{116}\text{DHLKYVMLPVADQ}^{128}$ from HP protein (accession Q6PEJ8) were all identified from ST of LA8, and $^{116}\text{DHLKYVMLPVA}^{126}$ and $^{116}\text{DHLKYVMLPVAD}^{127}$ were considered redundant peptides of $^{116}\text{DHLKYVMLPVADQ}^{128}$. To address this issue, the lists of identified consensus peptides were submitted to software generated in-house that collates and sets aside redundant peptides that have amino acid sequences that overlap in a core sequence. Peptide sequences with at least six continuous overlapping amino acid residues were considered redundant, and the peptide with the longest amino acid sequence was output from the software program. This software was also used to compare the consensus peptides among different samples. Peptide sequences with at least six continuous overlapping amino acid residues were considered consensus.

Identification of all source proteins containing a specific peptide sequence

All of the consensus peptides identified by at least two protein database search engines, and peptides assigned by a single program whose identifications passed close manual inspection, were submitted to another software program generated in-house, that screens the peptide sequences against the UniProt human protein database (not including decoy sequences). All the accession numbers of the source proteins containing the screened peptide sequence are displayed as a text string concatenating the accession numbers by semicolons in Microsoft Excel. This search showed that a number of identified peptides could be derived from more than one source protein, the consensus source proteins among different samples were

correlated by matching the text strings containing protein accession numbers using a Microsoft Access query. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium⁵⁷ via the PRIDE partner repository with the dataset identifier PXD003051.

Determinations of immunogenicity of peptides and source proteins

For the translational component of the protocol, the non-redundant peptides identified by LC-MS/MS were synthesized and evaluated with the matching patient's PBMC using ELISpot assays. Immunogenic peptides were then tested for T cell reactivity by ELISpot assay, and the source proteins of the peptides were assayed for B cell responses by ELISA in many patients and control subjects. To date, the determinations of peptide immunogenicity have been completed for three RA and in five LA patients.^{22–26} Further such analyses are ongoing.

Statistics

All p values are two-tailed. A p value of 0.05 was considered statistically significant.

RESULTS

Patients

All five patients with RA were women, ages 37 to 70 (Table 1), a distribution which reflects the gender bias associated with RA. Three of the five patients had classic, seropositive RA. They had the HLA-DRB1*0101 or 0401 alleles, which code for the RA shared epitope; these three patients had ACPA and two had rheumatoid factor (RF). The remaining two RA patients had chronic seronegative RA. They did not have shared epitope alleles, and their tests were negative for ACPA and RF. In four patients, ST from a hip or knee was obtained when arthroscopic synovectomy or joint replacement surgery was performed three to 66 years after disease onset. For one patient (RA2), samples were available from all three sites (ST, SFMC and PBMC), and, for patient RA4, SFMC and PBMC samples (but not ST) were collected at the time of diagnosis, prior to therapy with disease modifying anti-rheumatic drugs (DMARDs).

The eight patients with antibiotic-refractory LA included four teenage boys, three adult men and one woman (Table 1). Four of the eight patients had the HLA-DR*0101, 0401 or 1501 alleles, which are known to be increased in frequency in patients with refractory LA.⁵⁸ All eight patients were treated with two months of oral antibiotics, seven also received one month of IV antibiotics, and seven subsequently received methotrexate. Due to incomplete resolution of synovitis, the eight patients underwent arthroscopic synovectomies of an affected knee six months to four years after the onset of arthritis, and ST samples were obtained during the surgeries. Additionally, SFMC samples were available from four patients and PBMC from two patients.

Number of Peptides and Source Proteins Identified from HLA-DR-Presented Self Peptides

Overall, 1,647 to 40,390 MS/MS spectra were generated per sample from the *in vivo* HLA-DR-presented peptides isolated from the 22 samples of ST, SFMC or PBMC from the 13

patients (Table 1). Using three different MS/MS spectra search engines, after applying the filters (Mascot score 20; OMSSA e-value 0.01; and X!Tandem e-value 10), 161 to 29,993 spectrum-to-peptide matches were obtained per sample. Searching protein databases without specific enzymatic cleavage generates larger numbers of false positive hits than the ones with specific cleavages. To overcome this difficulty and ensure high quality in peptide identification, we combined the results from multiple search engines, and only the peptides identified by consensus match (*i.e.*, two search programs) are reported in Table 1 and supplementary Tables 1 and 2; the decoy hits in the consensus peptides were below 1%. The spectra for consensus peptides were manually inspected and given priority for inclusion in immunological assays; this number ranged from 2 to 2,363 peptides per sample. As noted in Table 1, expert manual inspection of peptide MS/MS spectra assigned with high scores by only a single program led to the addition of 90 peptides for inclusion in the list for immunological assays (supplementary Table 3).

At this stringency level, 59 to 801 non-decoy, non-redundant HLA-DR-presented self-peptides were identified in the 12 individual ST samples; these peptides were derived from 42 to 493 different source proteins (Table 1, supplementary Table 4, and supplementary Table 5). Using this same methodology for the six SFMC and four PBMC samples, 2 to 71 non-redundant peptides derived from 2 to 49 source proteins were identified per patient from SFMC, and 7 to 34 non-redundant peptides derived from 7 to 24 source proteins were identified per patient from PBMC. The lengths of the MHC class II HLA-DR-presented peptides ranged from seven to 27 amino acids. Peptides with 13–17 amino acids, the typical length of HLA-DR class II-presented peptides, accounted for 72.3% of all the identified HLA-DR-presented peptides. Altogether, from all three sites, 1,593 non-redundant HLA-DR-presented, self-peptides were identified; these were derived from 870 source proteins.

In the translational component of the research, all non-redundant peptides were synthesized and tested for reactivity with the matching patient's PBMC. To date, this step has been completed for three patients with RA and five with LA (Table 1). One immunogenic HLA-DR-presented peptide was identified in ST from patient RA1, and one immunogenic peptide was identified in ST from patient RA2.^{25, 26} In this patient, an additional peptide was also found in both ST and PBMC in three different analyses. Similarly, among the five LA patients, at least one unique immunogenic HLA-DR-presented peptide was identified in each of three patients (LA1, LA4, and LA5).^{22–24} Four of the six immunogenic peptides were identified by consensus match, and two were identified by manual inspection of sequences identified by only one program.

Comparison of Results from Different LC-MS/MS Instruments in the Identification of HLA-DR-Presented Self Peptides

During the 7-year period of this study, newer high sensitivity LC-MS/MS instruments became available in the laboratory. Therefore, over the course of the study, HLA-DR-presented peptides samples were analyzed by four types of LC-MS/MS instruments (Table 2). The average numbers of unique-consensus peptides identified from LA or RA patients' ST was 239 using the QSTAR, 313 using the LTQ-Orbitrap XL, 571 using the 6550 Q-TOF, and 2,226 using the Q Exactive plus MS. The average numbers from patients' SFMC were

51 using the LTQ-Orbitrap XL, 6 using the 6550 Q-TOF, and 70 using the Q Exactive plus MS. From patients' PBMC, the average numbers were 15 using the LTQ-Orbitrap XL, and 50 using the Q Exactive plus MS. Thus, as the sensitivity and data handling capacity of the LC-MS/MS instruments improved during the study, the number of identified HLA-DR-presented peptides steadily increased.

Comparison of HLA-DR-presented Peptides and Source Protein Repertoires in ST, SFMC, and PBMC

The identified HLA-DR-presented peptides from ST of RA1, RA2, LA1, and LA2 using IPI database were reported in our previous report.³⁰ Since the SFMC sample from LA2, and SFMC and PBMC samples from RA2 were analyzed in this study, the original data from the four previous samples, which were reanalyzed using UniProt database, were included herein for comparison of the HLA-DR-presented peptides repertoires found in the three sample sites now being analyzed. In one patient each with RA or LA (RA2 and LA5), samples were available from all three compartments (ST, SFMC and PBMC), and this allowed for the degree of overlap in peptide presentation among these three sites to be determined for each of these patients (Figure 1A). For patient RA2, 53% of the HLA-DR-presented peptides identified in SFMC and 53% of those found in PBMC were shared with those in ST, but only 6% were identified at all three sites. Analysis of the source proteins from which the peptides were derived gave similar frequencies (Figure 1B). In comparison, patient LA5 had a smaller percentage of peptides shared between SFMC and ST (21%) than between PBMC and ST (36%), and the degree of overlap of source proteins between SFMC and ST (16%) was similar between PBMC and ST (22%).

When the degree of overlap was compared between pairs of sample types available from the five RA patients or eight LA patients, a similar picture was found. In general, more than half of the peptides identified in SFMC were also found in ST (Figure 1C). In RA, about half of the peptides identified in PBMC were also found in ST, whereas in LA, in which many more peptides were identified in ST, only about one-quarter of the PBMC peptides were also found in ST. There was a trend toward a lower percentage of source proteins than peptides being the same in SFMC or PBMC as in ST (Figure 1D). Thus, in both diseases, HLA-DR-presented peptides in SFMC or PBMC provided some reflection of the peptides presented in ST, but also revealed additional peptides.

Source Proteins for HLA-DR-presented Peptides

The complete lists of consensus peptides identified by at least two search engines from each site for each patient are given in supplemental Table 1. The non-redundant HLA-DR-presented peptides and their source proteins are given in supplemental Tables 2 and 3. The most common source proteins of peptides from ST that were found in at least four patients, regardless of disease, are listed in alphabetical order in Table 3. These peptides (in order of frequency) were derived from cathepsin S, collagen α -1(I) chain, syntenin-1, serine carboxypeptidase CPVL (probable), albumin, pro-low-density lipoprotein receptor-related protein 1, apolipoprotein A2, complement C3, cytochrome c oxidase subunit 5B (mitochondrial), filamin-A, monocyte differentiation antigen CD14, and stromelysin-1. Since fewer SFMC and PBMC samples were analyzed, the list from these sites shows the

source proteins identified in at least two patients rather than four patients. In SFMC, the list included keratin (type II cytoskeletal 1) and nuclease-sensitive element-binding protein 1. Keratin was commonly identified in both SFMC and PBMC. Because of the keratin type and the fact that the samples were not subjected to proteolysis during processing, it is unlikely that these peptides were sample-handling artifacts.

For most RA patients, the list of identified HLA-DR-presented peptides included entries for source proteins that are thought to serve as potential autoantigens in RA.^{59, 60} These included cartilage glycoprotein 39, enolase, fibrinogen, fibronectin, immunoglobulin, and vimentin (Table 4). For the most part, it is only the citrullinated forms of these proteins that have been reported to be autoantigens in RA. In addition, most of these proteins are highly expressed in synovial tissue, which probably explains why HLA-DR-presented peptides from these source proteins were also commonly identified in patients with LA.

Search for citrullinated HLA-DR-presented peptides

Protein citrullination is considered to be an important post-translational modification in RA. The strong association between RA shared-epitope alleles and ACPA-positive RA has been hypothesized to result from the binding and presentation of citrullinated peptides by HLA-DR molecules.¹¹ Therefore, the MS/MS spectra were searched carefully for evidence of this post-translational modification, which results in the gain of 0.984 D (NH₂ vs. OH) when arginine is converted to citrulline. The search algorithms assigned several observed MS/MS spectra that had incomplete fragment ion series to citrullinated peptides. However, in each of these cases, manual inspection and targeted tandem mass spectrometry yielded full sequence information that unambiguously defined that the peptide modification actually corresponded to hydrolysis of asparagine to aspartic acid or glutamine to glutamic acid.

For example, arginine conversion to citrulline at R8 of the peptide VNDTQFVRFDSDAASPRG, that corresponds to residues 27 to 44 in the source protein B6V6K3 MHC class I antigen (fragment), which was originally detected in the ST of patient RA3, was assigned as the citrullinated peptide by two protein database search programs during analysis of LC-MS/MS data acquired on the LTQ-Orbitrap XL MS. However, since conversion of arginine to citrulline and hydrolysis of asparagine to aspartic acid result in the same change in elemental composition and thus to the same peptide mass shift, the molecular weight change and partial sequence coverage should not be considered definitive. In this case, the signal for the MS/MS spectrum shown in Figure 2A was recorded using the low resolution ion trap of the LTQ-Orbitrap XL MS; the series of sequence-determining fragment ions was incomplete and left room for ambiguity. In Figure 2A, the value measured for the fragment ion at m/z 589.9 corresponds, within the 0.5 Da error tolerance, to values calculated for the y_{11}^{2+} ion for the peptide corresponding to MHC class I antigen residues 27–44 modified by either the arginine-to-citrulline conversion or hydrolysis of asparagine to aspartic acid. Therefore, despite the agreement of the two database searches, it was really not possible to distinguish between these two post-translational modifications on the basis of these data. Later, when we had the opportunity to apply targeted tandem MS on the Q Exactive plus MS, the full series of high accuracy, high abundance diagnostic fragment ions allowed unambiguous assignment of the modification as asparagine deamidation at N2

(Figure 2B). Similar results were obtained for each of the other tentative assignments of citrullination that had been based on low resolution MS/MS spectra containing incomplete ion series.

Identification of Immunogenic peptides and source proteins

In five of the eight patients for whom sample testing has been completed, one to two immunogenic peptides (T cell epitopes) were identified from the peptides identified from each patient. When these peptides or their source proteins were tested for reactivity with PBMC or serum samples from many patients, we learned that 10–50% of ~50 RA or ~100 LA patients had T and/or B cell reactivity with each immunogenic peptide or its source protein. We have reported the immunological data and key MS/MS data for each individual autoantigen.^{22–26} Here, in an effort to integrate and make available the full mass spectrometry dataset, we present the complete list of peptides identified from each patient that correspond to assignments indicated on the MS/MS spectra that are included in the Supplementary Material.

The peptides found to be immunogenic were derived from the source proteins endothelial cell growth factor (ECGF),²² apolipoprotein B-100 (apoB-100),²³ matrix metalloproteinase-10 (MMP-10)²⁴ in LA; from annexin A2 in both RA and LA,²⁵ and from *N*-acetylglucosamine-6-sulfatase and filamin A in RA.²⁶ The MS/MS spectra and peptide sequences determined for these autoantigens are shown in Figure 3.

The immunogenic HLA-DR-presented peptide derived from filamin A was identified in both ST and PBMC, whereas the peptides from ECGF, apoB-100, annexin A2, and *N*-acetylglucosamine-6-sulfatase were identified only in ST. Although annexin A2 had been previously shown to serve as an autoantigen in several rheumatic diseases, including about 10% of patients with RA,⁶¹ it had not been previously identified as an autoantigen in LA.²⁵ Likewise, none of the other three autoantigens in LA (ECGF, MMP-10, and apoB-100) nor the two in RA (*N*-acetylglucosamine 6-sulfatase and filamin A) had been previously identified as autoantigens in those diseases. Moreover, for proteins in this group that had been reported to be immunogenic prior to our investigation, the precise epitopes had not previously been determined.

In 14 patients with antibiotic-refractory LA, ST was available for histopathologic correlations with autoantibody responses. Autoantibody reactivity with endothelial cell growth factor correlated strongly with obliterative microvascular lesions;³⁶ anti-apoB-100 autoantibodies correlated with greater numbers and activation of endothelial cells and more synovial fibroblast proliferation,²³ anti-MMP-10 autoantibodies correlated with measures of cell proliferation, including greater synovial lining layer thickness and greater numbers of synovial fibroblasts²⁴ and anti-annexin A2 autoantibodies correlated with greater synovial fibroblast proliferation.²⁵ Thus, the identification of T cell epitopes led to the identification of B cell responses that correlated with pathophysiological consequences.

DISCUSSION

In this study, we identified large numbers of HLA-DR-presented peptides directly from clinical samples of ST, SFMC or PBMC of patients with RA or LA. The challenge inherent in this approach exceeds that presented by the already difficult task of identifying HLA-DR-presented peptides in tissue culture systems, in which large numbers of a single cell type can be grown and harvested for analysis. Nevertheless, it was the impressive success that had been realized from tissue culture studies which motivated us to undertake the direct identification from clinical samples that are reported herein. Tandem mass spectrometry was first used to identify MHC molecules in well-defined tissue culture systems;³² for such analyses, at least 5×10^8 cells expressing 2×10^5 MHC molecules per cell was regarded as the minimum sample requirement.⁴³ Current reports from the Hunt group and other groups continue to improve and exploit this strategy very effectively to identify very large numbers of MHC- or HLA-bound peptides. These studies have identified both MHC Class I and Class II antigenic peptides in various tumor types, and have probed the mechanisms underlying the generation of HLA-presented peptides.^{37–39} However, the repertoire of HLA-DR-presented peptides in tissue culture systems contains many peptides derived from source proteins in the tissue culture medium.^{40, 41} It is thus quite different from direct analysis of clinical samples, as we have found.³⁰ Previously, *in vivo* presented peptides have been identified in patients' spleen,⁴⁴ colon,⁴⁵ kidney,⁴⁶ pooled bronchiolavage fluid,⁴⁷ thyroid,⁴⁸ or thymus;⁴⁹ these studies, on average, identified ~20–40 peptides per sample, and in some instances, samples from different patients needed to be pooled to identify peptides. An important exception was the identification of thousands of peptides, including some cancer-related peptides, presented by soluble HLA complexes that were immunoprecipitated from the plasma of patients and healthy controls.³³

In our original study of two patients each with RA or LA, in which samples were analyzed using the QSTAR mass spectrometer, we identified >80 non-redundant peptides in each ST sample.³⁰ In comparison, when we applied the same type of strategy to the sample obtained from our most recent patient, use of the Q Exactive plus MS led to the identification of 801 non-redundant peptides, a substantial increase. Even with recent increases in sensitivity and improvements in data analysis tools, identification of HLA-DR-presented peptides by LC-MS/MS still offers technical challenges. First, not all HLA-DR complexes are quantitatively recovered from a given sample. Additionally, some of the isolated peptides have low ionization efficiency and/or are not selected for MS/MS analysis, and not all spectra can be assigned by the search algorithms, due to low signal intensity, incomplete component resolution, and/or the presence of unusual PTMs. However, as shown here, the performance of MS/MS instruments and search algorithms continues to move forward.⁶²

In the current study, we took advantage of the greater sensitivity now available by analyzing not only ST, but also SFMC and PBMC as additional sources of *in vivo* HLA-DR-presented peptides that can be determined in individual patients. In general, the most HLA-DR-presented peptides were identified in ST, an inflamed tissue with high HLA-DR expression; an intermediate number of peptides was identified in SFMC, and the lowest number was identified in PBMC. Yet, 67% of the peptides identified in SFMC were also found in ST, and 55% of the peptides identified in PBMC were found in ST. A number of unique peptides

were only identified in the two additional sample sources. Therefore, the analyses of SFMC and PBMC each provided insights of HLA-DR-presented peptides in ST, but also revealed additional HLA-DR-presented peptides not found in ST.

An analysis of the cellular location of the source proteins of the HLA-DR-presented peptides identified from ST, SF and PB showed similar frequencies of both extracellular and intracellular proteins, and the highest percentages of proteins were from multiple locations. Thus, HLA-DR molecules expressed on cells from ST-, SF- and PB-presented peptides were derived not only from extracellular proteins, which are known to enter the endosomal MHC class II pathway, but also intracellular proteins, which are generally expected to be presented by the secretory MHC class I pathway. A possible explanation for this finding is that APCs randomly endocytose the contents of their local microenvironment, including remnants of dying cells, or these cells undergo autophagy,⁶³ an intracellular degradation pathway required for maintaining cellular and energy homeostasis. As a result, proteins typically defined as intracellular may enter the extracellular pathway, resulting in their presentation by MHC class II molecules.

Our data indicate that a more complete picture in an individual patient can be obtained by analyzing tissue or fluids from more than one site. The addition of SFMC and/or PBMC to the analytical scheme has three advantages. First, a more complete picture of an individual patient can be obtained by analyzing samples from multiple sites. Second, since ST can usually only be obtained from patients who undergo surgical procedures, frequently late in the disease, use of SFMC and/or PBMC allows evaluation of patients early in the disease when immune responses may be more robust. Third, given that ST is usually not available, one may obtain at least some information about events in ST, the target of the immune response in chronic inflammatory forms of arthritis. Thus, the testing of SFMC or PBMC opens up the possibility of testing a broader spectrum of patients at earlier stages of their diseases, and increasing the yield of information about immune reactivity in individual patients.

In RA patients, HLA-DR-presented peptides were identified from many source proteins previously reported to be autoantigens in RA, including vimentin, enolase, and fibrinogen. Interestingly, peptides from many of these same source proteins were frequently found in the joints of LA patients, where they are not known to be immunogenic. In RA, many proteins, such as vimentin, enolase, fibrinogen, type 2 collagen, fibronectin, calreticulin, have been shown to contain citrulline, the non-standard amino acid that results from the enzymatic post-translational modification of arginine.^{11, 64–70} Antibodies to these or other citrullinated proteins, known as ACPA, are strongly associated with ACPA-positive RA, particularly in patients with HLA-DR alleles that encode for the RA shared epitope,^{11–13} a highly homologous amino acid sequence at positions 70–74 (Q K/R RAA) in the B1 chain of the HLA-DR molecule. These residues contribute to the P4 anchoring pocket, giving it an overall positive charge. A disease-linked peptide containing citrulline (neutral R-group) has been shown to have greater affinity for the HLA shared epitope in the P4 pocket, as compared to the unmodified analog that contains arginine (positively charged R-group); this citrullinated peptide has been reported to induce T cell activation in DR4-IE transgenic mice.¹¹

Therefore, we searched our MS/MS spectra carefully for evidence of peptides that could contain arginine-to-citrulline conversions. Our search algorithms assigned several peptides as forms containing citrulline residues. However, deamidation at asparagine results in the same peptide mass change as arginine-to-citrulline conversion, and the initial MS/MS spectra contained incomplete series of fragment ions that lacked diagnostic peaks between the residues in question. Closer manual inspection of the initially obtained spectra and targeted tandem mass spectra recorded on the Q Exactive plus MS, the most advanced instrument employed in this study, determined that, for each of these peptides, the one-Dalton gain was actually due to deamidation at asparagine. The example shown in Figure 2 illustrates the MS/MS spectra obtained for the modified peptide ²⁷VNDTQFVRFDSAASPRG⁴⁴ that is derived from B6V6K3 MHC class I antigen (fragment); similar results were found for a number of other peptides derived from different source proteins. Asparagine deamidation, in addition to being the product of biological processes, can also result from exposure of the sample to high temperature and/or low pH,⁷¹ the very conditions used to elute peptides from the MHC molecules, and thus this modification could, in some cases, be an artifact generated during sample processing.

The reliable detection of citrullinated peptides may indeed present one of the largest challenges in this field. Nevertheless, we and others have been able to detect such peptides using the mass spectrometric methods that we are employing for retrieval and identification of HLA-presented antigens. Two recent articles reported a number of citrullinated proteins from RA patients' synovial tissue/fluid using LC-MS/MS,^{72, 73} and a paper that appeared while this manuscript was under review reports the detection of citrullinated peptides spiked into samples of cells from which HLA-presented peptides were retrieved and analyzed by LC/MS/MS.⁷⁴ From the five RA patients evaluated here, we detected HLA-DR-presented peptides that originated from source proteins in these lists, *e.g.*, vimentin, fibrinogen, or fibronectin. However, manual search of our LC-MS/MS data according to these lists did not result in the detection of any citrullinated HLA-DR-presented peptides.

The detection of HLA-presented citrullinated peptides may present one of the largest challenges in this field. We and others have been able to detect such peptides using the mass spectrometric methods that we are employing for retrieval and identification of HLA-presented antigens. A paper, which is now in press at *Mol. Cell. Proteomics*, recently appeared online and reports the analysis of HLA-presented peptides from 5–10 × 10⁶ EBV and BAL cells. The lists of thousands of identified peptides include large numbers of deamidated peptides (N,Q) but the authors only report detection of citrullinated peptides when α-enolase peptide standards with this modification were spiked into the samples as a control.⁷⁴ In the spiked samples, the citrullinated peptides could be clearly detected. (Because of sample limitations, we cannot perform a spiking experiment and repeat the analyses of the samples reported here.)

Considerations discussed by van Beers et al. address the difficulty of detecting intact proteins or large polypeptides that bear citrullination.⁷² Since there is considerable heterogeneity in site occupancy, the occurrence of each individual citrullinated epitope is low. The ionization efficiency of neutral citrullinated peptides is inherently lower than that of the equivalent peptide that contains a basic arginine residue. Therefore, it seems

understandable that the likelihood of detection of any single citrullinated epitope among HLA-presented peptides is even lower than the likelihood for detecting the intact modified protein. Furthermore, at least some of the anti-citrullinated antibodies may have specificity for overall conformation rather than for short linear epitopes presented by HLA-DR molecules.

Based on five patients studied here and the current methodology, we cannot make a definitive statement about the existence of citrullinated T cell epitopes in RA. It could happen that a given protein contains citrullinated B cell epitopes, but the T cell epitope identified was not citrullinated, or no single T cell epitope was citrullinated at a detectable level. For example, although we and others have been able to detect *in vitro* citrullinated peptides, it remains possible that the abundances of individual HLA-presented citrulline-containing peptides in the pools released from immunoprecipitated complexes were lower than the MS detection limit. However, the situation may change as a wider range of patients are tested and more diverse and more sensitive analytical strategies can be applied. We intend to pursue this topic in future studies, using still higher sensitivity instruments, ion mobility as an additional separation mode, and multiple reaction monitoring for selected candidate epitopes.

In conclusion, we have shown that LC-MS/MS is now sensitive enough to identify large numbers of HLA-DR-presented peptides directly from clinical samples of synovial tissue in individual patients with chronic inflammatory arthritis, and analyses of SFMC/PBMC may also reveal new peptide antigens. Moreover, discovery-based LC-MS/MS identification of HLA-DR-presented peptides from an individual patient's tissues or fluids provided peptide antigens for the identification of novel autoimmune responses in larger patient cohorts. Of the three autoantigens identified in RA patients, only one had been described previously, and all of the four autoantigens identified in LA patients were novel. In our experience, a principal advantage of autoantigen identification starting with an immunogenic T cell epitope is that it often focuses the search on more broadly applicable T and B cell responses in many patients that result in sustained high-titer autoantibodies, and these responses correlate with pathophysiological consequences. Furthermore, the identification of *in vivo* HLA-DR-presented peptides from PBMC, which can be readily obtained from any patient, offers a precision, personalized approach for the identification of potentially pathogenic immune responses in patients with immune-mediated diseases.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to the participating patients and controls and to the clinicians who obtained the samples for our analyses: Drs. John M. Aversa and Dennis Burke who helped by obtaining synovial tissue at the time of the patients' arthroscopic synovectomies, Dr. Sheila L. Arvikar for help with patient care, and Gail McHugh for laboratory support. We thank Thermo Fisher for loan of the Q-Exactive plus mass spectrometer.

Funding Sources

This research is supported by NIH grants [1] P41 GM104603, S10 RR020946 and S10 OD010724 and NIH-NHLBI contract HHSN268201000031C (to C.E.C.) and NIH grant R01 AI110175 and grants from the following: the Ounsworth-Fitzgerald Foundation; American College of Rheumatology, Mathers Foundation; English, Bonter, Mitchell Foundation; Littauer Foundation; Eshe Fund; and the Lillian B. Davey Foundation (to A.C.S.).

[1] The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ABBREVIATIONS

ACPA	anti-citrullinated protein antibodies
ACR	American College of Rheumatology
apoB-100	apolipoprotein B-100
Bb	<i>Borrelia burgdorferi</i>
DMARDs	disease modifying anti-rheumatic drugs
ECGF	endothelial cell growth factor
EULAR	European League against Rheumatism
HCD	higher-energy collision-induced dissociation
HLA-DR	human leukocyte antigen – antigen D related
IPI	International Protein Index
IV	intravenous
LA	Lyme arthritis
MHC	major histocompatibility complex
MMP-10	matrix metalloproteinase-10
PBMC	peripheral blood mononuclear cell
PBS	phosphate buffered saline
RA	rheumatoid arthritis
RF	rheumatoid factor
SF	synovial fluid
SFMC	synovial fluid mononuclear cell
ST	synovial tissue
STRAP	Software Tool for Rapid Annotation of Proteins

References

1. McDevitt HO, Bodmer WF. HL-A, immune-response genes, and disease. *Lancet*. 1974; 1:1269–1275. [PubMed: 4134154]
2. Takemura S, Braun A, Crowson C, Kurtin PJ, Cofield RH, O'Fallon WM, Goronzy JJ, Weyand CM. Lymphoid neogenesis in rheumatoid synovitis. *J Immunol*. 2001; 167:1072–1080. [PubMed: 11441118]
3. Steere AC, Duray PH, Butcher EC. Spirochetal antigens and lymphoid cell surface markers in Lyme synovitis. Comparison with rheumatoid synovium and tonsillar lymphoid tissue. *Arthritis Rheum*. 1988; 31:487–495. [PubMed: 3258751]
4. Plenge RM. Rheumatoid arthritis genetics: 2009 update. *Curr Rheumatol Rep*. 2009; 11:351–356. [PubMed: 19772830]
5. Stefanova I, Dorfman JR, Tsukamoto M, Germain RN. On the role of self-recognition in T cell responses to foreign antigen. *Immunol Rev*. 2003; 191:97–106. [PubMed: 12614354]
6. 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; 30:1205–1213. [PubMed: 2446635]
7. Holoshitz J. The rheumatoid arthritis HLA-DRB1 shared epitope. *Curr Opin Rheumatol*. 2010; 22:293–298. [PubMed: 20061955]
8. Berglin E, Johansson T, Sundin U, Jidell E, Wadell G, Hallmans G, Rantapaa-Dahlqvist S. Radiological outcome in rheumatoid arthritis is predicted by presence of antibodies against cyclic citrullinated peptide before and at disease onset, and by IgA-RF at disease onset. *Ann Rheum Dis*. 2006; 65:453–458. [PubMed: 16176994]
9. de Vries-Bouwstra JK, Goekoop-Ruiterman YP, Verpoort KN, Schreuder GM, Ewals JA, Terwiel JP, Ronda HK, Kerstens PJ, Toes RE, de Vries RR, Breedveld FC, Dijkmans BA, Huizinga TW, Allaart CF. Progression of joint damage in early rheumatoid arthritis: association with HLA-DRB1, rheumatoid factor, and anti-citrullinated protein antibodies in relation to different treatment strategies. *Arthritis Rheum*. 2008; 58:1293–1298. [PubMed: 18438829]
10. Pratesi F, Petit Teixeira E, Sidney J, Michou L, Puxeddu I, Sette A, Cornelis F, Migliorini P. HLA shared epitope and ACPA: just a marker or an active player? *Autoimmun Rev*. 2013; 12:1182–1187. [PubMed: 23958703]
11. Hill JA, Southwood S, Sette A, Jevnikar AM, Bell DA, Cairns E. Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. *J Immunol*. 2003; 171:538–541. [PubMed: 12847215]
12. Snir O, Widhe M, von Spee C, Lindberg J, Padyukov L, Lundberg K, Engstrom A, Venables PJ, Lundeberg J, Holmdahl R, Klareskog L, Malmstrom V. Multiple antibody reactivities to citrullinated antigens in sera from patients with rheumatoid arthritis: association with HLA-DRB1 alleles. *Ann Rheum Dis*. 2009; 68:736–743. [PubMed: 18635594]
13. Huizinga TW, Amos CI, van der Helm-van Mil AH, Chen W, van Gaalen FA, Jawaheer D, Schreuder GM, Wener M, Breedveld FC, Ahmad N, Lum RF, de Vries RR, Gregersen PK, Toes RE, Criswell LA. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. *Arthritis Rheum*. 2005; 52:3433–3438. [PubMed: 16255021]
14. Steere AC. Lyme disease. *N Engl J Med*. 2001; 345:115–125. [PubMed: 11450660]
15. Steere AC, Schoen RT, Taylor E. The clinical evolution of Lyme arthritis. *Ann Intern Med*. 1987; 107:725–731. [PubMed: 3662285]
16. Steere AC, Levin RE, Molloy PJ, Kalish RA, Abraham JH 3rd, Liu NY, Schmid CH. Treatment of Lyme arthritis. *Arthritis Rheum*. 1994; 37:878–888. [PubMed: 8003060]
17. Steere AC, Angelis SM. Therapy for Lyme arthritis: strategies for the treatment of antibiotic-refractory arthritis. *Arthritis Rheum*. 2006; 54:3079–3086. [PubMed: 17009226]
18. Shin JJ, Glickstein LJ, Steere AC. High levels of inflammatory chemokines and cytokines in joint fluid and synovial tissue throughout the course of antibiotic-refractory Lyme arthritis. *Arthritis Rheum*. 2007; 56:1325–1335. [PubMed: 17393419]

19. Strle K, Shin JJ, Glickstein LJ, Steere AC. Association of a Toll-like receptor 1 polymorphism with heightened Th1 inflammatory responses and antibiotic-refractory Lyme arthritis. *Arthritis Rheum.* 2012; 64:1497–1507. [PubMed: 22246581]
20. Shen S, Shin JJ, Strle K, McHugh G, Li X, Glickstein LJ, Drouin EE, Steere AC. Treg cell numbers and function in patients with antibiotic-refractory or antibiotic-responsive Lyme arthritis. *Arthritis Rheum.* 2010; 62:2127–2137. [PubMed: 20506317]
21. Vudattu NK, Strle K, Steere AC, Drouin EE. Dysregulation of CD4+CD25(high) T cells in the synovial fluid of patients with antibiotic-refractory Lyme arthritis. *Arthritis Rheum.* 2013; 65:1643–1653. [PubMed: 23450683]
22. Drouin EE, Seward RJ, Strle K, McHugh G, Katchar K, Londono D, Yao C, Costello CE, Steere AC. A novel human autoantigen, endothelial cell growth factor, is a target of T and B cell responses in patients with Lyme disease. *Arthritis Rheum.* 2013; 65:186–196. [PubMed: 23044924]
23. Crowley JT, Drouin EE, Pianta A, Strle K, Wang Q, Costello CE, Steere AC. A highly expressed human protein, apolipoprotein B-100, serves as an autoantigen in a subgroup of patients with Lyme disease. *J Infect Dis.* 2015; doi: 10.1093/infdis/jiv1310
24. Crowley JT, Strle K, Drouin EE, Pianta A, Arvikar SL, Wang Q, Costello CE, Steere AC. Matrix metalloproteinase-10 is a target of T and B cell responses that correlate with synovial pathology in patients with antibiotic-refractory Lyme arthritis. *J Autoimmun.* 2016; 69:24–37. [PubMed: 26922382]
25. Pianta A, Drouin EE, Crowley JT, Arvikar S, Strle K, Costello CE, Steere AC. Annexin A2 is a target of T and B cell responses associated with synovial fibroblast proliferation in patients with antibiotic-refractory Lyme arthritis. *Clin Immunol.* 2015; 160:336–341. [PubMed: 26187145]
26. Pianta A, Drouin EE, Wang Q, Arvikar S, Costello CE, Steere AC. Identification of *N*-acetylglucosamine-6-sulfatase and filamin A as novel targets of autoimmune T and B cell responses in rheumatoid arthritis. *Ann Rheum Dis.* 2015; 74S:112.
27. Steere AC, Klitz W, Drouin EE, Falk BA, Kwok WW, Nepom GT, Baxter-Lowe LA. Antibiotic-refractory Lyme arthritis is associated with HLA-DR molecules that bind a *Borrelia burgdorferi* peptide. *J Exp Med.* 2006; 203:961–971. [PubMed: 16585267]
28. Moret FM, Hack CE, van der Wurff-Jacobs KM, de Jager W, Radstake TR, Lafeber FP, van Roon JA. Intra-articular CD1c-expressing myeloid dendritic cells from rheumatoid arthritis patients express a unique set of T cell-attracting chemokines and spontaneously induce Th1, Th17 and Th2 cell activity. *Arthritis Res Ther.* 2013; 15:R155. [PubMed: 24286358]
29. Katchar K, Drouin EE, Steere AC. Natural killer cells and natural killer T cells in Lyme arthritis. *Arthritis Res Ther.* 2013; 15:R183. [PubMed: 24286535]
30. Seward RJ, Drouin EE, Steere AC, Costello CE. Peptides presented by HLA-DR molecules in synovia of patients with rheumatoid arthritis or antibiotic-refractory Lyme arthritis. *Mol Cell Proteomics.* 2011; 10 M110 002477.
31. Biemann K. Appendix 5. Nomenclature for peptide fragment ions (positive ions). *Methods Enzymol.* 1990; 193:886–887. [PubMed: 2074849]
32. Hunt DF, Henderson RA, Shabanowitz J, Sakaguchi K, Michel H, Sevilir N, Cox AL, Appella E, Engelhard VH. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science.* 1992; 255:1261–1263. [PubMed: 1546328]
33. Bassani-Sternberg M, Barnea E, Beer I, Avivi I, Katz T, Admon A. Soluble plasma HLA peptidome as a potential source for cancer biomarkers. *Proc Natl Acad Sci U S A.* 2010; 107:18769–18776. [PubMed: 20974924]
34. Alvarez-Navarro C, Cragolini JJ, Dos Santos HG, Barnea E, Admon A, Morreale A, Lopez de Castro JA. Novel HLA-B27-restricted epitopes from *Chlamydia trachomatis* generated upon endogenous processing of bacterial proteins suggest a role of molecular mimicry in reactive arthritis. *J Biol Chem.* 2013; 288:25810–25825. [PubMed: 23867464]
35. Bassani-Sternberg M, Pletscher-Frankild S, Jensen LJ, Mann M. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol Cell Proteomics.* 2015; 14:658–673. [PubMed: 25576301]

36. Londono D, Cadavid D, Drouin EE, Strle K, McHugh G, Aversa JM, Steere AC. Antibodies to endothelial cell growth factor and obliterative microvascular lesions in the synovium of patients with antibiotic-refractory lyme arthritis. *Arthritis Rheumatol.* 2014; 66:2124–2133. [PubMed: 24623727]
37. Cobbold M, De La Pena H, Norris A, Polefrone JM, Qian J, English AM, Cummings KL, Penny S, Turner JE, Cottine J, Abelin JG, Malaker SA, Zarling AL, Huang HW, Goodyear O, Freeman SD, Shabanowitz J, Pratt G, Craddock C, Williams ME, Hunt DF, Engelhard VH. MHC class I-associated phosphopeptides are the targets of memory-like immunity in leukemia. *Sci Transl Med.* 2013; 5:203ra125.
38. Li Y, Depontieu FR, Sidney J, Salay TM, Engelhard VH, Hunt DF, Sette A, Topalian SL, Mariuzza RA. Structural basis for the presentation of tumor-associated MHC class II-restricted phosphopeptides to CD4+ T cells. *J Mol Biol.* 2010; 399:596–603. [PubMed: 20417641]
39. Guasp P, Alvarez-Navarro C, Gomez-Molina P, Martin-Esteban A, Marcilla M, Barnea E, Admon A, Lopez de Castro JA. The Peptidome of Behcet's Disease-Associated HLA-B*51:01 Includes Two Subpeptidomes Differentially Shaped by Endoplasmic Reticulum Aminopeptidase 1. *Arthritis Rheumatol.* 2016; 68:505–515. [PubMed: 26360328]
40. Chicz RM, Urban RG, Gorga JC, Vignali DA, Lane WS, Strominger JL. Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. *J Exp Med.* 1993; 178:27–47. [PubMed: 8315383]
41. Lippolis JD, White FM, Marto JA, Luckey CJ, Bullock TN, Shabanowitz J, Hunt DF, Engelhard VH. Analysis of MHC class II antigen processing by quantitation of peptides that constitute nested sets. *J Immunol.* 2002; 169:5089–5097. [PubMed: 12391225]
42. Caron E, Espona L, Kowalewski DJ, Schuster H, Ternette N, Alpizar A, Schittenhelm RB, Ramarathinam SH, Lindestam Arlehamn CS, Chiek Koh C, Gillet LC, Rabsteyn A, Navarro P, Kim S, Lam H, Sturm T, Marcilla M, Sette A, Campbell DS, Deutsch EW, Moritz RL, Purcell AW, Rammensee HG, Stevanovic S, Aebersold R. An open-source computational and data resource to analyze digital maps of immunopeptidomes. *Elife.* 2015; 4doi: 10.7554/eLife.07661
43. Caron E, Kowalewski DJ, Chiek Koh C, Sturm T, Schuster H, Aebersold R. Analysis of Major Histocompatibility Complex (MHC) Immunopeptidomes Using Mass Spectrometry. *Mol Cell Proteomics.* 2015; 14:3105–3117. [PubMed: 26628741]
44. Gordon RD, Young JA, Rayner S, Luke RW, Crowther ML, Wordsworth P, Bell J, Hassall G, Evans J, Hinchliffe SA, et al. Purification and characterization of endogenous peptides extracted from HLA-DR isolated from the spleen of a patient with rheumatoid arthritis. *Eur J Immunol.* 1995; 25:1473–1476. [PubMed: 7539763]
45. Oshitani N, Hato F, Kitagawa S, Maeda K, Higuchi K, Matsumoto T, Arakawa T. Analysis of intestinal HLA-DR bound peptides and dysregulated immune responses to enteric flora in the pathogenesis of inflammatory bowel disease. *Int J Mol Med.* 2003; 11:99–104. [PubMed: 12469227]
46. Dengjel J, Nastke MD, Gouttefangeas C, Gitsioudis G, Schoor O, Altenberend F, Muller M, Kramer B, Missiou A, Sauter M, Hennenlotter J, Wernet D, Stenzl A, Rammensee HG, Klingel K, Stevanovic S. Unexpected abundance of HLA class II presented peptides in primary renal cell carcinomas. *Clin Cancer Res.* 2006; 12:4163–4170. [PubMed: 16857787]
47. Wahlstrom J, Dengjel J, Persson B, Duyar H, Rammensee HG, Stevanovic S, Eklund A, Weissert R, Grunewald J. Identification of HLA-DR-bound peptides presented by human bronchoalveolar lavage cells in sarcoidosis. *J Clin Invest.* 2007; 117:3576–3582. [PubMed: 17975675]
48. Muixi L, Carrascal M, Alvarez I, Daura X, Marti M, Armengol MP, Pinilla C, Abian J, Pujol-Borrell R, Jaraquemada D. Thyroglobulin peptides associate in vivo to HLA-DR in autoimmune thyroid glands. *J Immunol.* 2008; 181:795–807. [PubMed: 18566446]
49. Collado JA, Alvarez I, Ciudad MT, Espinosa G, Canals F, Pujol-Borrell R, Carrascal M, Abian J, Jaraquemada D. Composition of the HLA-DR-associated human thymus peptidome. *Eur J Immunol.* 2013; 43:2273–2282. [PubMed: 23719902]
50. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd, Birnbaum NS, Burmester GR, Bykerk VP, Cohen MD, Combe B, Costenbader KH, Dougados M, Emery P, Ferraccioli G, Hazes JM, Hobbs K, Huizinga TW, Kavanaugh A, Kay J, Kvien TK, Laing T, Mease P, Menard HA, Moreland LW, Naden RL, Pincus T, Smolen JS, Stanislawski-Biernat E,

- Symmons D, Tak PP, Upchurch KS, Vencovsky J, Wolfe F, Hawker G. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum.* 2010; 62:2569–2581. [PubMed: 20872595]
51. Wharton M, Chorba TL, Vogt RL, Morse DL, Buehler JW. Case definitions for public health surveillance. *MMWR Recomm Rep.* 1990; 39:1–43.
 52. Steere AC, Glickstein L. Elucidation of Lyme arthritis. *Nat Rev Immunol.* 2004; 4:143–152. [PubMed: 15040587]
 53. Chen W, Yewdell JW, Levine RL, Bennink JR. Modification of cysteine residues in vitro and in vivo affects the immunogenicity and antigenicity of major histocompatibility complex class I-restricted viral determinants. *J Exp Med.* 1999; 189:1757–1764. [PubMed: 10359579]
 54. Haque MA, Hawes JW, Blum JS. Cysteinylation of MHC class II ligands: peptide endocytosis and reduction within APC influences T cell recognition. *J Immunol.* 2001; 166:4543–4551. [PubMed: 11254711]
 55. Meadows L, Wang W, den Haan JM, Blokland E, Reinhardus C, Drijfhout JW, Shabanowitz J, Pierce R, Agulnik AI, Bishop CE, Hunt DF, Goulmy E, Engelhard VH. The HLA-A*0201-restricted H-Y antigen contains a posttranslationally modified cysteine that significantly affects T cell recognition. *Immunity.* 1997; 6:273–281. [PubMed: 9075928]
 56. Pierce RA, Field ED, den Haan JM, Caldwell JA, White FM, Marto JA, Wang W, Frost LM, Blokland E, Reinhardus C, Shabanowitz J, Hunt DF, Goulmy E, Engelhard VH. Cutting edge: the HLA-A*0101-restricted HY minor histocompatibility antigen originates from DFFRY and contains a cysteinylated cysteine residue as identified by a novel mass spectrometric technique. *J Immunol.* 1999; 163:6360–6364. [PubMed: 10586024]
 57. Vizcaino JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D, Dienes JA, Sun Z, Farrah T, Bandeira N, Binz PA, Xenarios I, Eisenacher M, Mayer G, Gatto L, Campos A, Chalkley RJ, Kraus HJ, Albar JP, Martinez-Bartolome S, Apweiler R, Omenn GS, Martens L, Jones AR, Hermjakob H. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol.* 2014; 32:223–226. [PubMed: 24727771]
 58. Behera AK, Hildebrand E, Szafranski J, Hung HH, Grodzinsky AJ, Lafyatis R, Koch AE, Kalish R, Perides G, Steere AC, Hu LT. Role of aggrecanase 1 in Lyme arthritis. *Arthritis Rheum.* 2006; 54:3319–3329. [PubMed: 17009305]
 59. Hoffman, M., Lundberg, K., Steiner, G. Autoantibodies in rheumatoid arthritis. In: Hochberg, MC, Silman, AJ, Smolen, JS, Weinblatt, ME., Weisman, MH., editors. *Rheumatology*. sixth. Elsevier; Philadelphia: 2015. p. 750-757.
 60. Davidson, A., Bridges, SLBJ. Autoimmunity. In: St. Clair, EW, Pisetsky, DS., Haynes, BF., editors. *Rheumatoid Arthritis*. Lippincott Williams & Wilkins; Philadelphia: 2004. p. 197-213.
 61. Salle V, Maziere JC, Smail A, Cevallos R, Maziere C, Fuentes V, Tramier B, Makdassi R, Choukroun G, Vittecoq O, Goeb V, Ducroix JP. Anti-annexin II antibodies in systemic autoimmune diseases and antiphospholipid syndrome. *J Clin Immunol.* 2008; 28:291–297. [PubMed: 18322784]
 62. Zhang B, Pirmoradian M, Chernobrovkin A, Zubarev RA. DeMix workflow for efficient identification of cofragmented peptides in high resolution data-dependent tandem mass spectrometry. *Mol Cell Proteomics.* 2014; 13:3211–3223. [PubMed: 25100859]
 63. Kroemer G. Autophagy: a druggable process that is deregulated in aging and human disease. *J Clin Invest.* 2015; 125:1–4. [PubMed: 25654544]
 64. Vossenaar ER, Despres N, Lapointe E, van der Heijden A, Lora M, Senshu T, van Venrooij WJ, Menard HA. Rheumatoid arthritis specific anti-Sa antibodies target citrullinated vimentin. *Arthritis Res Ther.* 2004; 6:R142–150. [PubMed: 15059278]
 65. Kinloch A, Tatzer V, Wait R, Peston D, Lundberg K, Donatien P, Moyes D, Taylor PC, Venables PJ. Identification of citrullinated alpha-enolase as a candidate autoantigen in rheumatoid arthritis. *Arthritis Res Ther.* 2005; 7:R1421–1429. [PubMed: 16277695]
 66. Takizawa Y, Suzuki A, Sawada T, Ohsaka M, Inoue T, Yamada R, Yamamoto K. Citrullinated fibrinogen detected as a soluble citrullinated autoantigen in rheumatoid arthritis synovial fluids. *Ann Rheum Dis.* 2006; 65:1013–1020. [PubMed: 16449316]

67. Burkhardt H, Sehnert B, Bockermann R, Engstrom A, Kalden JR, Holmdahl R. Humoral immune response to citrullinated collagen type II determinants in early rheumatoid arthritis. *Eur J Immunol.* 2005; 35:1643–1652. [PubMed: 15832289]
68. van Beers JJ, Willemze A, Stammen-Vogelzangs J, Drijfhout JW, Toes RE, Pruijn GJ. Anti-citrullinated fibronectin antibodies in rheumatoid arthritis are associated with human leukocyte antigen-DRB1 shared epitope alleles. *Arthritis Res Ther.* 2012; 14:R35. [PubMed: 22339947]
69. Ling S, Cline EN, Haug TS, Fox DA, Holoshitz J. Citrullinated calreticulin potentiates rheumatoid arthritis shared epitope signaling. *Arthritis Rheum.* 2013; 65:618–626. [PubMed: 23233327]
70. Spengler J, Lugonja B, Ytterberg AJ, Zubarev RA, Creese AJ, Pearson MJ, Grant MM, Milward M, Lundberg K, Buckley CD, Filer A, Raza K, Cooper PR, Chapple IL, Scheel-Toellner D. Release of Active Peptidyl Arginine Deiminases by Neutrophils Can Explain Production of Extracellular Citrullinated Autoantigens in Rheumatoid Arthritis Synovial Fluid. *Arthritis Rheumatol.* 2015; 67:3135–3145. [PubMed: 26245941]
71. Pace AL, Wong RL, Zhang YT, Kao YH, Wang YJ. Asparagine deamidation dependence on buffer type, pH, and temperature. *J Pharm Sci.* 2013; 102:1712–1723. [PubMed: 23568760]
72. van Beers JJ, Schwarte CM, Stammen-Vogelzangs J, Oosterink E, Bozic B, Pruijn GJ. The rheumatoid arthritis synovial fluid citrullinome reveals novel citrullinated epitopes in apolipoprotein E, myeloid nuclear differentiation antigen, and beta-actin. *Arthritis Rheum.* 2013; 65:69–80. [PubMed: 23044660]
73. Ytterberg AJ, Joshua V, Reynisdottir G, Tarasova NK, Rutishauser D, Ossipova E, Haj Hensvold A, Eklund A, Skold CM, Grunewald J, Malmstrom V, Jakobsson PJ, Ronnelid J, Padyukov L, Zubarev RA, Klareskog L, Catrina AI. Shared immunological targets in the lungs and joints of patients with rheumatoid arthritis: identification and validation. *Ann Rheum Dis.* 2015; 74:1772–1777. [PubMed: 24817415]
74. Heyder T, Kohler M, Tarasova NK, Haag S, Rutishauser D, Rivera NV, Sandin C, Mia S, Malmstrom V, Wheelock AM, Wahlstrom J, Holmdahl R, Eklund A, Zubarev RA, Grunewald J, Ytterberg AJ. Approach for Identifying Human Leukocyte Antigen (HLA)-DR Bound Peptides from Scarce Clinical Samples. *Mol Cell Proteomics.* 2016; 15:3017–3029. [PubMed: 27452731]

Figure 1. Degree of overlap in HLA-DR-presented peptides and their source proteins identified in the ST, SFMC or PBMC of RA or LA patients

The Venn diagrams depict the number of shared peptides or source proteins identified in the ST, SFMC or PBMC of individual RA or LA patients or in the entire RA and/or LA cohort. (A) Shared peptides or (B) shared source proteins identified in patients RA2 or LA5. (C) Shared peptides or (D) shared source proteins identified in the RA, LA, or RA and LA cohorts. All peptide comparisons were based upon consensus-matched, non-redundant identified peptides using in-house generated software. The consensus on source proteins matched between cohorts was assembled by matching the source protein accession numbers using a Microsoft Access query.

Figure 2. High abundance, high accuracy peptide fragment ions can provide full sequence coverage and clearly distinguish asparagine deamidation from arginine citrullination (A) CID MS² spectrum, recorded using the ion trap on the LTQ-Orbitrap XL MS for analysis of the mixture of MHC class II-presented peptides recovered from the synovial tissue of patient RA3. The spectrum was assigned to a modified form of the peptide ²⁷VNDTQFVRFDSDAASPRG⁴⁴ derived from the source protein B6V6K3 MHC class I antigen (fragment). No sequence-defining fragment ions were observed between the second and eighth residues; nevertheless, the post-translational modification of arginine conversion to citrulline at **R8** was assigned by two protein database search programs. (B) HCD MS² spectrum of the same component, recorded with the Q Exactive plus MS, allowed unambiguous assignment of asparagine deamidation at **N2** as the post-translational modification on the peptide sequence. Note: ° loss of H₂O; * loss of NH₃.

OMSSA, and X!Tandem from RA2 ST (the same peptide was also identified consistently by Mascot, OMSSA, and X!Tandem from RA2 PB);²⁶ and (F) ²²²FEPFFMMIATPAPH²³⁵ from *N*-acetylglucosamine-6-sulfatase was consistently identified by Mascot, OMSSA, and X!Tandem from RA2 ST.²⁶ Note: ° loss of H₂O; * loss of NH₃.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Clinical characteristics, number of MS/MS spectra, peptide matches, decoy hits, non-redundant peptides, and source proteins identified from each patient’s HLA-DR molecules in synovial tissue, synovial fluid or peripheral blood

	Rheumatoid Arthritis						Lyme Arthritis							
	RA1	RA2	RA3	RA4	RA5		LA1	LA2	LA3	LA4	LA5	LA6	LA7	LA8
Patient														
Gender	F	F	F	F	F		M	M	M	M	F	M	M	M
Age (years)	70	37	45	63	68		12	43	18	38	51	14	16	15
Arthritis Duration (years)	66	3	7	1	3		1.1	1.6	0.5	1.6	4	0.83	0.83	0.75
Anti-citrullinated protein antibodies	neg	pos	pos	pos	neg		neg	neg	neg	neg	neg	neg	neg	neg
Rheumatoid Factor	neg	pos	pos	neg	neg		neg	neg	neg	neg	neg	neg	neg	neg
HLA-DR*B1 genotype	0402/1104	0101/0401	0101/0401	0401/1501	0801/1501		0101/0801	1001/1101	0301/1501	1101/1302	0301/0305	0103/0301	0403/1501	0401/1001
Synovial tissue [‡] , No. of														
MS/MS spectra	1647 ¹	1712 ¹	10220 ²		10375 ²		2237 ¹	1664 ¹	9637 ²	11940 ²	14141 ²	5892 ²	6090 ²	29454 ³
spectrum-to-peptide matches [‡]	1031	1323	2538		1852		1082	696	4516	2531	2453	621	1214	20066
consensus matched peptides	248	292	471		145		282	135	481	315	399	95	285	732
Decoy hits	1	0	1		0		0	0	0	0	0	0	0	3
consensus non-decoy non-redundant peptides	95	94	172		68		95	57	167	115	135	36	115	271
peptides tested for immune responses identified by only one program	18	8	ND		ND		12	2	10	18	12	ND	ND	ND
Immunogenic peptides	1	2	N/A		NYD		1	0	0	1	1	NYD	NYD	NYD
Source proteins	80	78	126		52		79	42	136	106	109	36	84	176
Synovial fluid mononuclear cells, No. of														
Cells tested		8.7×10 ⁷		3.6×10 ⁷				6.4×10 ⁷			3.4×10 ⁷		5.9×10 ⁷	6.2×10 ⁷
MS/MS spectra		20979 ²	5120 ⁴	16169 ²				14244 ²			13030 ²		1000 ³	7307 ⁴
spectrum-to-peptide matches [‡]		2461	254	1145				1694			1380		120	162
consensus matched peptides		30	51	66				2			107		6	45
Decoy hits		1	0	0				0			0		0	0
consensus non-decoy non-redundant peptides		15	32	37				2			61		5	26
peptides tested for immune responses identified by only one program		ND	ND	ND				ND			10		ND	ND

	Rheumatoid Arthritis						Lyme Arthritis					
		0	0	0	0		0				NYD	NYD
Immunogenic peptides											NYD	NYD
Source proteins		14	23	32			2				40	22
Peripheral blood mononuclear cells, No. of												
Cells tested		6.3×10 ⁷								8.0×10 ⁷		
MS/MS spectra		20931 ²	6227 ⁴	11060 ²						3616 ⁴		
spectrum-to-peptide matches [‡]		1897	372	1216						161		
consensus matched peptides		22	60	7						40		
Decoy hits		0	0	0						0		
consensus non-decoy non-redundant peptides		14	31	7						28		
peptides tested for immune responses identified by only one program		ND	ND	ND						ND		
Immunogenic peptides		1	1	0						NYD		
Source proteins		11	24	7						12		

¹LC-MS/MS data acquired with a Sciex QSTAR mass spectrometer;

²Thermo Fisher LTQ-Orbitrap XL mass spectrometer;

³Agilent 6550 Q-TOF mass spectrometer; and

⁴Thermo Fisher Q Exactive plus mass spectrometer

^QResults from a second time immune purification from the tissue cell lysate

[‡]RA1 to RA5 and LA1 to LA5 purified from 7–10 grams of ST; LA6, LA7, and LA8 purified from 15, 14.5, and 17 grams of ST, respectively.

[‡]Spectrum-to-peptide matches after applying the filters (Mascot score ≥20; OMSSA e-value ≤0.01; and X!Tandem e-value ≤10)

N/A: not available; NYD: not yet done; ND: not done (sufficient information already available from consensus peptides)

Table 2

Comparison of the spectra generated and identified by the four different mass spectrometers.

	Synovial tissue	Synovial fluid	Peripheral blood
QSTAR	N=4 *		
# MS/MS spectra [†]	1815		
# spectrum-to-peptide matches	1033		
# consensus matched peptides	239		
# Decoy hits	0		
# non-decoy non-redundant peptides	95		
# Source Proteins	70		
LTQ-Orbitrap XL	N=7	N=4	N=3
# MS/MS spectra	9756	16106	19379
# spectrum-to-peptide matches	2246	1670	1498
# consensus matched peptides	313	51	15
# Decoy hits	0	0	0
# non-decoy non-redundant peptides	121	31	11
# Source Proteins	93	24	9
6550 iFunnel Q-TOF	N=3	N=1	
# MS/MS spectra	19633	1000	
# spectrum-to-peptide matches	13106	120	
# consensus matched peptides	571	6	
# Decoy hits	3	0	
# non-decoy non-redundant peptides	202	5	
# Source Proteins	141	5	
Q Exactive plus	N=1	N=3	N=2
# MS/MS spectra	10936	6125	4922
# spectrum-to-peptide matches	9927	255	267
# consensus matched peptides	2226	70	50
# Decoy hits	3	0	0
# non-decoy non-redundant peptides	765	38	30
# Source Proteins	479	28	18

* Number of different samples analyzed. Only data for the first pass analysis are included.

[†] All numbers given indicate the average number of spectra identified based upon the number of samples analyzed.

Table 3

The most common source proteins for the identified HLA-DR-presented peptides from three sites in RA and LA patients

Synovial tissue *							
Protein Name	Accession	RA#	LA#	Protein Name	Accession	RA#	LA#
Albumin	P02768	2	5	Filamin-A	P21333	2	4
Alpha-1B-glycoprotein	P04217	1	3	Fructose-bisphosphate aldolase A	P04075	2	2
Alpha-2-macroglobulin	P01023	2	3	Hemopexin	P02790	1	3
Alpha-enolase	P06733	1	3	HLA class I histocompatibility antigen, A-2 alpha chain	P01892	2	2
Angiotensinogen	P01019	1	3	Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	1	3
Annexin A2	P07355	1	3	Legumain	Q99538	1	3
Apolipoprotein A-II	P02652	3	3	Lysosome-associated membrane glycoprotein 1	P11279	0	4
C4b-binding protein alpha chain	P04003	1	3	Macrophage mannose receptor 1	Q5VSK2	1	3
Calreticulin	P27797	1	3	Monocyte differentiation antigen CD14	P08571	2	4
Cathepsin S	P25774	3	5	Olfactomedin-like protein 2B	Q68BL8	0	4
Caveolin-1	Q03135	1	3	Perilipin-3	O60664	0	4
Chondroadherin	O15335	1	3	Phosphatidylcholine-sterol acyltransferase	P04180	2	2
Coagulation factor XIII A chain	P00488	2	2	Phosphoglycerate kinase 1	P00558	3	1
Cofilin-1	P23528	2	3	Plexin domain-containing protein 2	Q6UX71	3	2
Collagen alpha-1(I) chain	P02452	3	5	Profilin-1	P07737	2	3
Collagen alpha-1(XD) chain	P12107	2	2	Prolow-density lipoprotein receptor-related protein 1	Q07954	3	4
Collagen alpha-2(I) chain	P08123	2	3	Proteoglycan 4	Q92954	1	3
Collagen alpha-2(V) chain	P05997	2	3	Rho GDP-dissociation inhibitor 2	P52566	1	4
Complement C1q subcomponent subunit B	P02746	0	4	60S ribosomal protein L30	P62888	3	1
Complement C1s subcomponent	P09871	1	4	60S ribosomal protein L6	Q02878	2	2
Complement C3	P01024	3	3	Serine carboxypeptidase CPVL (probable)	Q9H3G5	3	4
Complement C5	P01031	1	4	SPARC-like protein 1	Q14515	0	4
C-reactive protein	P02741	2	3	Stabilin-1	Q9NY15	2	3
Cytochrome c oxidase subunit 5B, mitochondrial	P10606	2	4	Stromelysin-1	P08254	2	4
Cytochrome c oxidase subunit 8A, mitochondrial	P10176	2	2	Syntenin-1	O00560	2	6
Decorin	P07585	1	4	Talin-1	Q9Y490	0	5
Epitidmval secretory protein E1	P61916	1	3	Vascular cell adhesion protein 1	P19320	0	4

Synovial tissue [*]					
Protein Name	Accession	RA#	LA#	Protein Name	Accession RA# LA#
ER-Golgi intermediate compartment protein 1	Q969X5	1	3	Vitroneectin	P04004 2 3
Fibrinogen beta chain	P02675	3	2		

Synovial fluid [†]					
Protein Name	Accession	RA#	LA#	Protein Name	Accession RA# LA#
Albumin	P02768	1	1	Keratin, type I cytoskeletal 9	P35527 1 1
Alpha-2-macroglobulin	P01023	1	1	Keratin, type II cytoskeletal 1	P04264 1 2
Annexin A6	P08133	1	1	Lysosomal alpha-mannosidase	O00754 1 1
C4b-binding protein alpha chain	P04003	1	1	MHC class I antigen	P01889 1 1
Cathepsin S	P25774	1	1	MHC class II antigen	P04229 1 1
Cofilin-1	P23528	1	1	Nuclease-sensitive element-binding protein 1	P67809 1 2
Complement C3	P01024	1	1	Plasminogen	P00747 1 1
Fibrinogen alpha chain	P02671	1	1	Plexin domain-containing protein 2	Q6UX71 1 1
Ficolin-1	O00602	0	2	Stromelysin-1	P08254 1 1
Glutathione S-transferase P	P09211	1	1	T-cell surface glycoprotein CD4	P01730 1 1
Hemoglobin subunit alpha	P69905	1	1	Toll-like receptor 2	O60603 1 1
Histone H1	Q02539	0	2		

Peripheral Blood [‡]					
Protein Name	Accession	RA#	LA#	Protein Name	Accession RA# LA#
Alpha-1-antitrypsin	P01009	0	2	Keratin, type I cytoskeletal 9	P35527 1 2
Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	1	1	Keratin, type II cytoskeletal 1	P04264 1 1
Keratin, type I cytoskeletal 10	P13645	1	1	Nuclease-sensitive element-binding protein 1	P67809 1 1

The numbers in "RA#" and "LA#" columns indicate the number of patients from whom each source protein was identified.

^{*} Source proteins identified in ST in at least four of the twelve patients.

[†] Source proteins identified in SF in at least two of six patients tested.

[‡] Source proteins identified in PB in at least two of four patients tested.

Table 4

Source proteins found both in rheumatoid arthritis and Lyme arthritis that are known or proposed to serve as autoantigens in rheumatoid arthritis*

	<u>Rheumatoid Arthritis</u> Patient	<u>Lyme Arthritis</u> Patient
Calreticulin	RA1	LA3/LA4/LA8
Cartilage glycoprotein 39	RA2	LA1/LA8
Enolase	RA4 [‡] /RA5	LA2/LA5 [‡] /LA6/LA7/LA8
Fibrinogen -	alpha RA2 [‡] /RA3/RA5	LA3/LA4/LA5/LA7 [‡] /LA8 [‡]
	beta RA2/RA3/RA5	LA1/LA4/LA6/LA7 [‡] /LA8
Fibronectin	RA1/RA2/RA3/RA5	LA1/LA2/LA3/LA4/LA5/LA6/LA7 [‡] /LA8
Immunoglobulin	RA2/RA3	LA1/LA3/LA4/LA8
Vimentin	RA1/RA2/RA3/RA5	LA1/LA3/LA7/LA8 [‡]

* Unless marked otherwise, peptides from these source proteins were identified only in synovial tissue.

[‡] Peptide(s) from this source protein identified in synovial tissue and synovial fluid

[‡] Peptide(s) from this source protein identified only in synovial fluid