Approach for Identifying Human Leukocyte Antigen (HLA)-DR Bound Peptides from Scarce Clinical Samples*†‡

Tina Heyder†§, Maxie Kohler‡, Nataliya K. Tarasova§, Sabrina Haag¶, Dorothea Rutishauser§, Natalia V. Rivera†, Charlotte Sandin||, Sohel Mia||, Vivianne Malmström||, Åsa M. Wheelock‡, Jan Wahlström‡, Rikard Holmdahl¶, Anders Eklund‡, Roman A. Zubarev§, Johan Grunewald‡‡, and A. Jimmy Ytterberg§¶**‡‡

Immune-mediated diseases strongly associating with human leukocyte antigen (HLA) alleles are likely linked to specific antigens. These antigens are presented to T cells in the form of peptides bound to HLA molecules on antigen presenting cells, e.g. dendritic cells, macrophages or B cells. The identification of HLA-DR-bound peptides presents a valuable tool to investigate the human immunopeptidome. The lung is likely a key player in the activation of potentially auto-aggressive T cells prior to entering target tissues and inducing autoimmune disease. This makes the lung of exceptional interest and presents an ideal paradigm to study the human immunopeptidome and to identify antigenic peptides.

Our previous investigation of HLA-DR peptide presentation in the lung required high numbers of cells (800 \times 10^6 bronchoalveolar lavage (BAL) cells). Because BAL from healthy nonsmokers typically contains 10–15 \times 10^6 cells, there is a need for a highly sensitive approach to study immunopeptides in the lungs of individual patients and controls.

In this work, we analyzed the HLA-DR immunopeptidome in the lung by an optimized methodology to identify HLA-DR-bound peptides from low cell numbers. We used an Epstein-Barr Virus (EBV) immortalized B cell line

From the †Respiratory Medicine Unit, Department of Medicine and Center for Molecular Medicine, Solna, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden; §Division of Physiological Chemistry I, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden; ¶Division of Medical Inflammation Research, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden; ||Rheumatology Unit, Department of Medicine, Solna, Karolinska Institutet, Stockholm, Sweden

Received May 6, 2016, and in revised form, July 19, 2016

Published, MCP Papers in Press, July 24, 2016, DOI 10.1074/mcp.M116.060764

Author contributions: T.H., M.K., D.R., V.M., R.A.Z., J.G., and A.J.Y. designed research; T.H., M.K., N.K.T., S.H., D.R., C.S., S.M., and A.J.Y. performed research; T.H. and R.M. contributed new reagents or analytic tools; T.H., A.M.W., J.W., and A.J.Y. analyzed data; T.H., S.H., N.V.R., A.M.W., J.W., R.H., A.E., R.A.Z., J.G., and A.J.Y. wrote the paper; A.E. performed bronchoscopies.

Autoimmune diseases are complex inflammatory disorders characterized by the immune system losing self-tolerance against own cells or tissue. The prevalence of such diseases is approximately five percent in Europe and North America and constitute the 10th leading cause of death worldwide (1, 2). The major histocompatibility complex (MHC)1, which encompasses the human leukocyte antigens (HLA), plays a central role in the genetic susceptibility to such diseases, predisposing individuals to e.g., type I diabetes, rheumatoid arthritis (RA) or multiples sclerosis (3–5). Each of these disease is likely to have at least a subset of peptides being presented by the

1 The abbreviations used are: MHC, major histocompatibility complex; APC, Antigen presenting cells; BAL, bronchoalveolar lavage; BCA, bicinchoninic acid; DM, n-Dodecyl β-D-maltoside; EBV, Epstein–Barr virus; FDR, false discovery rate; HCD, Higher-energy collisional dissociation; HLA, human leukocyte antigen; IC50, half maximal inhibitory concentration; IIP, idiopathic interstitial pneumonia; IPF, idiopathic pulmonary fibrosis; MudPIT, Multidimensional Protein Identification Technology; RA, rheumatoid arthritis; SCX, strong cationic exchange; WASOG, World Association for Sarcoidosis and Other Granulomatous Disorders; WAX, weak anionic exchange.

Molecular & Cellular Proteomics 15.9
3017

© 2016 by The American Society for Biochemistry and Molecular Biology, Inc.
This paper is available on line at http://www.mcponline.org
Optimization of the Methodology to Identify HLA-DR Peptides

Patient characteristics. Abbreviations: Diagnosis (Diagn.), chest radiographic stage (CXR stage), vital capacity (VC), forced vital capacity (FVC), forced expiratory volume in one second (FEV1), diffusion capacity of the lung for carbon monoxide (DLCO), patient 1–7 (PAT1–7), sarcoidosis (S), idiopathic pulmonary fibrosis (IPF), idiopathic interstitial pneumonia (IIP), female (F), male (M), smoker (sm), ex-smoker (ex-sm), not relevant (NR) and not analyzed (NA).

| Sample | Total cell number | Diagn. | HLA genotype | Age | Sex | Smoking status | Löfgren’s syndrome | CXR stage | VC | FVC | FEV1 | DLCO |
|--------|------------------|--------|--------------|-----|-----|---------------|-------------------|-----------|----|-----|------|------|
| BAL cells |
| PAT1 | $85 \times 10^6$ | S | DR4 | DR7 | 54 | M | sm | No | II | 64 | 65 | 62 | 54 |
| PAT2 | $50 \times 10^6$ | S | DR4 | DR13 | 31 | M | sm | No | III | 71 | 103 | 110 | 73 |
| PAT3 | $23 \times 10^6$ | S | DR7 | DR13 | 30 | M | sm | Yes | I | 103 | NA | 99 | 87 |
| PAT4 | $10 \times 10^6$ | S | DR13 | DR14 | 46 | M | sm | No | II | NA | NA | NA | NA |
| PAT5 | $2 \times 10 \times 10^6$ | IPF | DR8 | DR13 | 64 | F | sm | NR | NA | 87 | 88 | 86 | 65 |
| PAT6 | $2 \times 10 \times 10^6$ | IIP | Missing | 57 | F | ex-sm | NR | NA | NA | 80 | 84 | NA | NA |
| PAT7 | $4 \times 11.5 \times 10^6$ | S | DR3 | DR15 | 62 | M | ex-sm | No | IV | 106 | 98 | 103 | 78 |

HLA molecules, that are specific for the disease. The peptides presented by the HLA molecules are referred to as the immunopeptidome.

HLA and non-HLA genes are located in the MHC region on chromosome 6 and make up the largest polymorphic region in the human genome. These genes are key factors for the regulation and control of the homeostasis of the immune system. The function of HLA molecules is to present peptides on the cell surface to be recognized by distinct T cells in order to trigger an immune response when appropriate. Typically, peptides from endogenous proteins are presented on HLA class I molecules (HLA-A, -B and -C) and recognized by CD8+ T cells, whereas peptides from exogenous proteins are presented on HLA class II molecules (HLA-DR, -DQ and -DP) and are recognized by CD4+ T cells. However, this mode of detection is not always systematic as cross-presentation occurs (6). Moreover, the activation of T cells by recognition of specific peptides is a complex process, making it a crucial component for understanding the pathogenic mechanisms in inflammation and autoimmunity (7). The location where T cell activation takes place is an important constituent in this process. Recently, the lung has been suggested to play a central role in the activation of auto-aggressive T cells prior to entering target tissues and inducing autoimmune disease, as shown in an animal model for multiple sclerosis (8). This, as well as the occurrence of several T cell mediated lung disorders, make BAL cells from the lungs an ideal model system to identify antigenic peptides, i.e. its immunopeptidome, particularly under inflammatory conditions, as in the case of sarcoidosis. Sarcoidosis is a systemic, granulomatous disease most commonly affecting the lungs. The strong HLA-DR allele association with disease susceptibility (5, 9) and associated (oligo-) clonal expansion in the deep airways of T cells expressing specific T cell receptor segments (9, 10) suggest the presentation of distinct disease-associated antigens.

BAL cells mainly contain macrophages, which are antigen presenting cells and therefore offer an excellent epitome for method development aiming to identify specific peptides in healthy and diseased states. However, the handling and determination of the HLA immunopeptidome from BAL cells, requires highly sensitive approaches because healthy individuals have relatively low numbers of resident immune cells in the deep airways when compared with inflammatory lung disease patients.

Historically, methods for the identification of HLA associated peptide repertoires have been limited because of the complexity and variable nature of endogenous peptides, as well as the limitations in conventional technologies (11–18). That is, technologies used to analyze peptides associated to HLA class I complexes based on Edman degradation (11) have been subsequently improved to isolate and identify HLA-I peptides with high sensitivity using mass spectrometry (MS) (12), followed by isolation of HLA-II peptides (13). These advancements, though valuable and beneficial for the field of immunopeptidomics, have however still required relatively large amounts of cells or tissue, of at least $5 \times 10^6$ cells expressing $2 \times 10^6$ HLA molecules/cell (14), and these amounts are not always feasible to obtain from human samples.

In the present study, we established an improved method that enables the identification of HLA-DR bound peptides from low cell numbers. Specifically, we used cell populations consisting of Epstein-Barr (EBV)-immortalized B cells and bronchoalveolar lavage (BAL) cells obtained from patients with sarcoidosis to identify the HLA-DR subset of their immunopeptidome. Our optimized approach facilitates the analysis and mapping of the peptidome of individual patients and healthy controls and may reveal disease specific peptides (autoantigens). These peptides might give new insights into disease etiology and/or might be used as diagnostic tools or vaccination targets.

**EXPERIMENTAL PROCEDURES**

Patient Characteristics—Seven patients (PAT1–7) with lung diseases: sarcoidosis (S, n = 5), idiopathic pulmonary fibrosis (IPF, n = 1), and idiopathic interstitial pneumonia (IIP, n = 1) were included in...
pellets (106 cells) stored at serum (Sigma). Cells were harvested, washed with PBS twice and cell

Isolation of BAL Cells—BAL was collected during bronchoscopy as saline (PBS), instilled into the middle lobe bronchus and aspirated into a n ice-cooled flask. The BAL cells were collected from the BAL fluid using centrifugation for 10 min at 400

Sarcoidosis and other Granulomatous Diseases (WASOG) (19), where patients with IPF and IIP were diagnosed after a multidisciplinary conference at the clinic (20). The HLA-DR types were obtained for the following patients: PAT1: HLA-DRB1*04, DRB1*07 (DR4,DR7); PAT2: HLA-DRB1*04, DRB1*13 (DR4,DR13); PAT3: HLA-DRB1*07, DRB1*13 (DR7,DR13); PAT4: HLA-DRB1*13, DRB1*14 (DR13,DR14); PAT5: HLA-DRB1*08, DRB1*13 (DR8,DR13); PAT7 HLA-DRB1*03, DRB1*15 (DR3,DR15). Smokers (5 current and 2 exsmokers) were included in the study because BAL from smokers in general contains higher numbers of cells and would allow for multiple analysis of the same sample. The study was approved by the Regional Ethical Review Board in Stockholm, Sweden, and written informed consent was obtained from all subjects.

EBV-B Cell Line—An Epstein-Barr virus transformed human B cell line (EBV-B cells), Priess, homozygous for HLA-DRB1*04:01 (21), was cultured in RPMI 1640 medium (Sigma, Steinheim, Germany) supplemented with 1% HEPES (Sigma), 1% L-glutamine (Sigma), 100 U/ml penicillin, 100 μg/ml streptomycin (Sigma) and 10% fetal bovine serum (Sigma). Cells were harvested, washed with PBS twice and cell pellets (10^6 cells) stored at −80 °C.

Optimization of the Methodology to Identify HLA-DR Peptides

SOS-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting—Aliquots were taken at different steps during the isolation of membrane complexes and protein concentrations were determined by using the Pierce BCA protein assay kit (Thermo Fisher, Rockford, IL). 2.5 μg of each fraction were separated with 1D polyacrylamide gel electrophoresis (1D-PAGE) using 4–12% NuPAGE Bis-Tris gels (Invitrogen) according to the manufacturer’s recommendations. One SDS-PAGE gel was stained using the fluorescent LavaPurple Total Protein Stain (Fluorotechnics Ltd., Sydney, Australia), whereas a second gel was used to detect the HLA-DR complexes by Western blotting using an anti–HLA-DR primary antibody (dilution 2 μg/ml; TAL 1B5 ab20181, Abcam, Cambridge, England) and a goat anti-mouse secondary antibody labeled by a fluorescent probe (dilution 1:2500; Amersham Biosciences ECL Plex Goat-α-Mouse IgG-Cy3, GE Healthcare, Buckinghamshire, UK). The fluorescent signals from both 1D-PAGE and Western blot were detected by the Typhoon FLA 9000 laser scanner (GE Healthcare, Uppsala, Sweden).

Immunoprecipitation of HLA-DR Bound Peptides—anti-HLA-DR (L243) antibodies (25) were added to the solubilized HLA-DR complexes at a ratio of 40 μg antibody/50 μg protein and incubated overnight on an end-over-end tube rotator at 6 rpm, 4 °C. Resuspended Protein G Ultra Link Resin (Thermo Fisher Scientific) corresponding to 40 μl settled resin was added and the samples were incubated for 2 h at room temperature. The immunoprecipitates were collected by centrifugation at 84 × g. Nonspecifically bound molecules were removed by washing with 10 bed volumes of 150 mM NaCl, 50 mM ammonium bicarbonate, followed by 150 mM NaCl, 50 mM ammonium bicarbonate and 5% acetonitrile at room temperature. HLA-DR bound peptides were eluted using two times 5 bed volumes of 50% acetonitrile and 5% formic acid. The peptides were further purified using strong cationic exchange (SCX) by a single-step elution with 0.5 M ammonium formate, 2.5% formic acid, and 20% acetonitrile from PolySULFOETHYL A (5 μM, 200 Å; PolyLC Inc., Columbia, Maryland) packed in P10 tips. Prior to LC-MS/MS, the peptides where desalted using either ZipTips (C18, Merck Millipore Ltd, Ireland) or StageTips (C18, Thermo Scientific). A detailed description of the procedure for isolation of HLA-DR peptides has been added to the supplemental data (supplemental Procedures).

Identification of Peptides in the Flow Through from Immunoprecipitation—Peptides in the flow through at the immunoprecipitation step from two replicates of the EBV-B cells and two replicates from PAT7 were identified. These peptides are naturally occurring peptides or generated by in vivo and/or ex vivo proteinolysis, and referred to as flow through peptides. The flow through was passed through a 10 kDa cut off filter (Pall Corporation, Ann Arbor, MI) and desalted using C18 SEF Pack columns (Waters Corporation, Milford, MA) according to the manufacturer’s recommendations. Peptides were further purified with SCX self-packed columns as described above, to remove remaining detergents, and desalted using StageTips (C18, Thermo Scientific) prior to LC-MS/MS.

Mass Spectrometry, Peptide Identification, and Peptide Filtering—The purified peptides were analyzed by nanoLC-MS/MS using an Easy-nLC chromatographic system or a NanoUltimate 3000 coupled on-line to an LTQ Orbitrap Velos (PAT1), a Q Exactive (PAT2–6), or a Q Exactive Plus (EBV-B cells, PAT7) mass spectrometer (all Thermo Fisher Scientific). The peptides in sample PAT1–6 were separated using 10 cm long fused silica tip columns (OD 360 μm and ID 75 μm, 100 Å precol- umn (Thermo Scientific) together with a heated (55 °C) 25 cm long PicoFrit® self-pack analytical column (360 μm OD, 75 μm ID, 10 μm
Optimization of the Methodology to Identify HLA-DR Peptides

Our optimized procedure to isolate HLA-DR peptides from low cell numbers included cell lysis, immunoprecipitation of HLA-DR, elution of HLA-DR bound peptides, desalting and peptide concentration, and identification by mass spectrometry followed by peptide filtering. The specific steps imple-

Experimental Design and Statistical Rationale—HLA-DR peptides from EBV-B cells were isolated in technical replicates of $10^6$ cells, using at least four replicates to evaluate different conditions in the procedure. In addition, HLA-DR peptides from BAL cells were analyzed from two replicates of $10^6$ cells (PAT5) and two replicates of $10^6$ cells together with one replicate of $5 	imes 10^5$ cells (PAT6) and four replicates of $10^5$ cells (PAT7). PAT1–4 were analyzed in single replicates. The FDR for identification was calculated using the decoy database searching in Mascot. The Mann-Whitney test was used to compare the yields of different conditions for HLA-DR peptide elution (GraphPad Prism v5, www.graphpad.com). The Kolmogorov-Smirnov test was used to compare the different peptide length distributions (using “Real statistics using Excel,” www.real-statistics.com).

Evaluation of the Efficiency of Specific Steps During the Purification of HLA-DR Peptides—Certain steps in the procedure to identify HLA-DR bound peptides were evaluated by comparing the following conditions: the effect of freeze-thawing on the stability of HLA-DR peptide complexes; incubation time with anti-HLA-DR antibody during immunoprecipitation (2 h versus overnight) and the difference in immunoprecipitating HLA-DR peptide complexes directly from total cellular lysate compared with enriched membrane complexes. Two negative controls, immunoprecipitation with protein-G Sepharose only (i.e. no antibody) and immunoprecipitation with an irrelevant antibody (anti-tetanus) (28), were also included. Experiments were performed in four replicates per condition of each replicate containing $10^6$ cells.

Sequence Analysis—Binding motifs were predicted using the NNAlign prediction algorithm (v1.4) (www.cbs.dtu.dk/services/NNAlign) (29). Identified peptides in PAT7 and EBV B cells using the final protocol together with their matching sets of flow through peptides (i.e. peptides identified in the IP flow through) as negative control, were submitted together with a set of 20,000 randomly selected 1mers from the complete human proteome database as a control. Identified peptides were assigned a value of 1 and the randomly generated control peptides were assigned 0. The parameters were set as follows: peptide length: 9; flanking amino acids for NNAlign/)( 29). Identified peptides longer than eight amino acids from PAT7 and EBV-B cells were submitted together with their matching flow through peptides to the NetMHCIIpan 3.1 server (www.cbs.dtu.dk/services/NetMHCIIpan) to predict the binding strength between the peptide and HLA-DR molecule (30). Peptides with an IC50 of $< 500$ nM or belonging to the top 2% of the binders (comparing to a set of 200,000 random natural peptides) were considered to be weak binders, whereas peptides with an IC50 of $< 50$ nM or belonging to the top 0.5% of the binders were considered to be strong binders. For simplicity, binders were defined as having either strong or weak binding. The protein functions were predicted by the Panther algorithm (www.pantherdb.org) (31).

RESULTS

Our optimized procedure to isolate HLA-DR peptides from low cell numbers included cell lysis, immunoprecipitation of HLA-DR, elution of HLA-DR bound peptides, desalting and peptide concentration, and identification by mass spectrometry followed by peptide filtering. The specific steps imple-

Experimental Evaluation—HLA-DR peptides from EBV-B cells were isolated in technical replicates of $10^6$ cells, using at least four replicates to evaluate different conditions in the procedure. In addition, HLA-DR peptides from BAL cells were analyzed from two replicates of $10^6$ cells (PAT5) and two replicates of $10^6$ cells together with one replicate of $5 	imes 10^5$ cells (PAT6) and four replicates of $10^5$ cells (PAT7). PAT1–4 were analyzed in single replicates. The FDR for identification was calculated using the decoy database searching in Mascot. The Mann-Whitney test was used to compare the yields of different conditions for HLA-DR peptide elution (GraphPad Prism v5, www.graphpad.com). The Kolmogorov-Smirnov test was used to compare the different peptide length distributions (using “Real statistics using Excel,” www.real-statistics.com).

Evaluation of the Efficiency of Specific Steps During the Purification of HLA-DR Peptides—Certain steps in the procedure to identify HLA-DR bound peptides were evaluated by comparing the following conditions: the effect of freeze-thawing on the stability of HLA-DR peptide complexes; incubation time with anti-HLA-DR antibody during immunoprecipitation (2 h versus overnight) and the difference in immunoprecipitating HLA-DR peptide complexes directly from total cellular lysate compared with enriched membrane complexes. Two negative controls, immunoprecipitation with protein-G Sepharose only (i.e. no antibody) and immunoprecipitation with an irrelevant antibody (anti-tetanus) (28), were also included. Experiments were performed in four replicates per condition of each replicate containing $10^6$ cells.

Sequence Analysis—Binding motifs were predicted using the NNAlign prediction algorithm (v1.4) (www.cbs.dtu.dk/services/NNAlign) (29). Identified peptides in PAT7 and EBV B cells using the final protocol together with their matching sets of flow through peptides (i.e. peptides identified in the IP flow through) as negative control, were submitted together with a set of 20,000 randomly selected 1mers from the complete human proteome database as a control. Identified peptides were assigned a value of 1 and the randomly generated control peptides were assigned 0. The parameters were set as follows: peptide length: 9; flanking amino acids for NNAlign/)( 29). Identified peptides longer than eight amino acids from PAT7 and EBV-B cells were submitted together with their matching flow through peptides to the NetMHCIIpan 3.1 server (www.cbs.dtu.dk/services/NetMHCIIpan) to predict the binding strength between the peptide and HLA-DR molecule (30). Peptides with an IC50 of $< 500$ nM or belonging to the top 2% of the binders (comparing to a set of 200,000 random natural peptides) were considered to be weak binders, whereas peptides with an IC50 of $< 50$ nM or belonging to the top 0.5% of the binders were considered to be strong binders. For simplicity, binders were defined as having either strong or weak binding. The protein functions were predicted by the Panther algorithm (www.pantherdb.org) (31).

RESULTS

Our optimized procedure to isolate HLA-DR peptides from low cell numbers included cell lysis, immunoprecipitation of HLA-DR, elution of HLA-DR bound peptides, desalting and peptide concentration, and identification by mass spectrometry followed by peptide filtering. The specific steps imple-
mented to improve the overall performance compared with traditional methods (15, 16, 23) are outlined below and illustrated in Fig. 1.

**Isolation of HLA-DR membrane complexes**—The isolation of HLA-DR peptides from low cell numbers requires a very meticulous sample preparation and a minimal contamination...
Optimization of the Methodology to Identify HLA-DR Peptides

Fig. 2. Isolation of HLA-DR membrane complexes. A, SDS-PAGE of fractions from different steps during the isolation of HLA-DR membrane complexes. The fractionation shows a simplification of the protein profile from total cell lysate (TC1), total cell lysate after removing cell debris and intact cells (TC2), crude membranes (CM) to membrane complexes (MC). B, Western blot of the four fractions using an anti-HLA-DR antibody to evaluate the HLA-DR content in each fraction. The insert shows the quantification of the intensity of each band in the Western blot.

with interfering substances. Initial problems during the optimization of the procedure described by Wahlström et al. (23) to the current instrumental set up, where the peptides were separated on columns packed in the emitters and without pre-columns, were associated with frequent loss of spray, most likely due to detergents and polymers. Changing the detergent from the commonly used CHAPS (15–18, 23), to the MS-compatible n-Dodecyl β-d-maltoside (DM) (24) increased the spray stability. Moreover, improved washing of protein-G beads prior to elution (e.g. by adding acetonitrile to wash buffer), reduced the amount of polymer peaks in the spectra.

To limit the nonspecific protein binding, we isolated crude membranes prior to the extraction of HLA-DR complexes (24). By using this prepurification step, the HLA-DR concentration was increased 4.2 times, whereas the amounts of other potentially unrelated proteins were reduced (Fig. 2). The procedure yielded on average 73 μg protein/million BAL cells and 120 μg protein/million EBV-B cells after cell lysis, whereof 11% ended up in the fraction with membrane complexes.

Immunoprecipitation of HLA-DR Bound Peptides—The immunoprecipitation step was improved by using smaller volumes and a simpler procedure. The amount of antibody and protein-G beads was reduced to a ratio of 4:5 protein/antibody (w/w) and 1:1 antibody/protein-G beads (w/v), respectively. Acetonitrile at high concentration was added to the acidic elution buffers, which have been used in most procedures, to improve the elution yield and reduce the required elution volume. Similar to Haag et al. (32), the peptides were purified using strong cationic exchange (SCX), which here was performed manually using in-house packed microcolumns. Clean-up with either C18 ZipTips or StageTips were used for the final peptide purification. The peptides were identified using standard on-line nanoLC-MS/MS. The recovery from SCX was tested by isolating the peptides in the flow through using weak anionic exchange (WAX). Very few peptides matching the source proteins were identified in the WAX fractions, which suggests efficient recovery by the SCX step. Because including the data from the WAX fractions in the database search increased the false discovery rate (FDR), the WAX purification was discontinued (data not shown).

MS Identification and Filtering—With the exception of two samples identifying only a few peptides, samples from EBV-B cells (final version of the procedure) had an FDR of 0.00–2.56% and 0.00–3.33% for the identity and homology thresholds, respectively (as calculated by Mascot). To ensure high quality of the peptide identifications, the scores were recalculated using Percolator and required to fulfill two out of three criteria: the protein had to pass the Mascot significance score threshold (MudPIT) in at least one analysis, and either have a peptide score higher than the significance threshold (>13) or have at least four consecutive b- or y-ions. To further reduce the number of false positive identifications and peptides originating from in vivo and/or ex vivo proteolysis, and hence not belonging to the HLA-DR peptidome, peptide groups were formed using overlapping sequences of at least four amino acids (and therefore likely to be part of nested sets). Only groups with at least two peptides were kept. As a final filter, all peptides identified among the flow through peptides (i.e. peptides identified in the IP flow through) were also removed. The filtering was done using ClusterMHCII, an algorithm written in Perl.

Validation of Methodology—In a proof of principle experiment, cells with a set of known peptides were used for the optimized procedure to isolate HLA-DR peptides. To generate these cells, 5 × 10^6 EBV-B cells were pre-incubated with two synthesized peptides. The peptides were from the same sequence of α-enolase with the exception of one having citrulline instead of arginine at a specific position (TSDKGLFX-AAVPSGAS, where X was either Cit or Arg). Both peptides have been shown to bind to HLA-DR in a competitive assay and have been able to elicit T cell responses (Sandin et al. §§). Both peptides could be identified in the experiments where they had been included in the mixture and absent where they had not (supplemental Fig. S1).

Evaluation of Specific Steps During the Isolation of HLA-DR Peptides from EBV-B Cells—HLA-DR peptides were analyzed

§§ Sandin, C., Gerstner, C., James, E., Uchtenhagen, H., Pieper, J., Rieck, M., Reed, E., Zubarev, R. A., Ytterberg, A. J., Sandalova, T., Buckner, J. H., Adnane Achour, A. and Malmström, V. Qualitative differences in T-cell responses towards the Rheumatoid Arthritis-associated autoantigen alpha-enolase when restricted by HLA-DRB1*04:01 versus *01:01 and *04:04. Manuscript submitted for review.
Fig. 3. Evaluation of specific steps during the isolation of HLA-DR peptides from EBV B cells. A, Panel A presents the identified scans per million cells comparing freeze-thawing membrane complexes (M*), short incubation time with anti-HLA-DR antibody (M2h) and purifying peptides from membrane complexes (M) compared with a total cell lysate (T). The different symbols in M indicate 3 different sets of 4 replicates. B, In panel B the number of unique peptide sequences per million cells is plotted for all conditions. C, Presents the overlap of peptides and source proteins between HLA-DR peptides and flow through peptides in purifying peptides from a total cell lysate and crude membranes. D, Presents the top 5 ranked proteins in flow through peptides and the corresponding ranks of the same protein in the HLA-DR peptide data set identified from total cell lysate and crude membranes.

from $10 \times 10^6$ EBV-B cells and different steps in the procedure were evaluated, each in four replicates. In total, 507 unique peptide sequences, 104 core sequences and 75 proteins were identified as HLA-DR peptides (supplemental Table S2), whereas 494 peptides were identified in the flow through fraction (supplemental Table S3). To compare different conditions, both the number of unique sequences per million cells and the number of identified scans per million cells (so-called spectral counting) were calculated. The relative protein abundances in the fraction of flow through peptides were determined by ranking the proteins using spectral counting (supplemental Table S8) while merging the results from both replicates, but without requiring overlapping peptides. Two negative controls were tested: immunoprecipitation without antibody (i.e. protein-G Sepharose only) and an irrelevant antibody. In both cases, only a few peptides were identified, all of which were removed by the filtering process. These peptides matched proteins highly abundant in the flow through fraction, such as heterogeneous nuclear ribonucleoproteins (HNRNP) A2/B1, keratin and 60S ribosomal protein L28.

We first characterized the difference between performing the immunoprecipitation using total cellular lysate (T) and crude membranes (M). For condition T, the median yield was 4.50 scans/million cells (3.75 unique sequences/million cells) compared with 11.95 scans/million cells for M (9.85 unique sequences/million cells; Fig. 3A). To validate the reduced complexity of M compared with T shown by SDS-PAGE (CM and TC in Fig. 2), we estimated the relative “protein” abundance by spectral counting (by merging the replicates, but without requiring overlaps) and compared the protein ranks with the flow through peptides identified in the flow through. Out of the 5 highest ranked proteins in the flow through fraction (HNRNP A2/B1, HNRNP A1, vimentin, histone 3.2 and histone 3.1), vimentin was identified in T (rank 69 out of 80) and histone 3.2 and histone 3.1 in M though with low ranks (113.5 and 128 out of 137, respectively). HNRNP were found in both fractions, but with much higher ranks in T compared with M (rank 33 versus 58 and 49.5 versus 128, in T and M, respectively; Fig. 3D). Also when comparing peptides present in the three sets, T showed a larger overlap with flow through peptides (15.5%) than M did (7.5%; Fig. 3C). On the protein level, the numbers increase to 30.2% for T and 30.3% for M (Fig. 3C).

We next evaluated the effect of freeze-thawing the membrane complexes once during the procedure (M* versus M in Fig. 3A) and using a shorter incubation time with the anti-HLA-DR antibody (M2h versus M, Fig. 3A). Both changes reduced the number of scans/million cells from medians of 11.95 for M to 2.0 for M*, and 2.80 for M2h. Similar changes were observed comparing the number of unique sequences per million cells, which yielded the median values of 1.75, 2.45 and 9.85 (Fig. 3B). Based on these results, the crude membrane preparations with freeze thawing and overnight incubation with the anti-HLA-DR antibody were included in the final protocol. Though the median for M was higher compared with the other three conditions, the differences did not reach statistical significance ($p$ value).

Identification of HLA-DR Peptides from BAL Cells—HLA-DR peptides from BAL cells were identified using the final protocol. In total, 1434 unique peptide sequences belonging to 384 core sequences from 215 proteins were iden-
tified (supplemental Table S4). In three patients (PAT5–7) the immunopeptidome was investigated in multiple technical replicates. HLA-DR bound peptides were isolated from two replicates of $10^6$ cells (PAT5), two replicates of $10^6$ cells together with $50 \times 10^6$ cells (PAT6) and four replicates of $10^6$ cells (PAT7). On average we identified 9.4 peptides/million BAL cells (PAT1–7; Fig. 4A and 4B) and the overlaps of peptides between experimental replicates were 36.8 and 42.0% for PAT5 and PAT6, respectively. For PAT7, where four replicates were analyzed, 21.6% of the peptides were identified in all fractions, 40.8% were identified in at least three replicates, and 65.6% were identified in at least two replicates. Source proteins from the identified peptides were classified according to function (gene ontology), and the six largest groups in both data sets (EBV-B and BAL) were binding, catalytic activity, enzyme regulatory activity, receptor activity, structural molecule activity and nucleic acid binding transcription factor activity (supplemental Fig. S2). When comparing with the proteins identified by the flow through peptides, the largest differences were found in the groups for receptor activity and enzyme regulatory activity (8.9 versus 3.8% and 8.9 versus 1.6%, for BAL and flow through respectively). For EBV-B cells the greatest differences are found in the groups for catalytic activity, structural molecule activity, receptor activity, protein binding activity (10.8%, 7.3 versus 18.1%, 10.9 versus 4.8% and 1.8 versus 0% for EBV-B cells and flow through respectively).

Peptides present in the flow through of immunoprecipitation were also identified from two replicates of PAT7 (flow through peptides). In total, 2579 peptides were identified (supplemental Table S5) and the most abundant proteins were several actins and histones.

Specificity of Identified Peptides—In order to evaluate to what degree the identified peptides were likely to be HLA-DR bound peptides, specific characteristics, such as peptide length distribution, HLA-DR binding motifs and peptide binding strength were determined in silico. The combined HLA-DR peptides identified from BAL cells of PAT7 and EBV-B cells (condition M, Fig. 3A and 3B) both show narrow distributions compared with flow through and tryptic peptides (Fig. 5A–5B). HLA-DR peptides from BAL had a median length of 15 amino acid residues (S.D. 2.6). The peptide length distribution for EBV-B cells was slightly higher with a median peptide length of 16 (S.D. 2.7). The distributions of flow through peptides in both BAL and EBV-B cells showed a much wider distribution with medians of 15 (S.D. 5.2) for BAL and 14 (S.D. 5.4) for EBV-B cells. The peptide length distributions for both BAL and EBV-B cells were statistically different from their corresponding flow through peptides ($p = 2.5 \times 10^{-7}$ and $p = 2.3 \times 10^{-5}$ for BAL and EBV-B cells, respectively, using the Kolmogorov-Smirnov test). The peptide length distribution for HLA-DR derived peptides was also compared with a large data set of tryptic peptides generated from the human THP-1 cell line (33). The tryptic data set was generated from 12 biological replicates with multiple technical replicates (in total 30 LC-MS analyses), and consisted of 19,316 peptides (13,446 unique sequences) from 2765 proteins. The tryptic data set showed a much wider distribution with a median of 14 (S.D. 5.6). Also the peptide length distributions for THP-1 were statically different from the HLA-DR peptides ($p = 2.025 \times 10^{-4}$ and $p = 6.0 \times 10^{-3}$ for BAL and EBV-B cells, respectively).

Binding motifs of HLA-DR peptides and flow through peptides identified from EBV-B cells (supplemental Tables S2 and S3, on peptides identified using condition M) and BAL cells (PAT7; supplemental Table S4 and S5) were generated. The predicted motifs for the HLA-DR peptides were largely consistent with motifs previously generated by Nielsen and colleagues (34) (Fig. 6A, 6C, 6E, 6F, and 6G). In the case of PAT7, which is heterozygous regarding the HLA-DR variants (DR3, DR15), we compared the motif with two motifs previously generated by Nielsen and colleagues (34) for single alleles. Though the subclasses of the HLA-DR alleles were not known in this case, it was assumed that the patient had the most common variants among northern Europeans: HLA-DR1*0301, DR1*1501 (35). The binding motif for PAT7 was consistent with a combination of the motif for both alleles (Fig. 6C versus 6F and 6G). Flow through peptides from both EBV-B and BAL cells did not result in a motif with conserved amino acids (i.e. binding anchors) in the core sequence (Fig. 6B and 6D).

The binding strength was predicted for peptides longer than 8 amino acids from EBV-B cells purified from membrane complexes and BAL cells from PAT7 using the NetMHCIIpan 3.1 Server (data not shown) (30). From the total 450 peptides...
for EBV-B cells, 67% were predicted to be binders. For PAT7, the binding strength of the 341 identified peptides longer than 8 amino acids were predicted for both HLA alleles variants DRB1*03:01 and DRB1*15:01. The majority (55%) of the peptides were predicted to bind at least one allele. In contrast, only 8.7% and 15.6% were predicted to be binders in the sets of flow through peptides for EBV-B cells and BAL cells, respectively (data not shown).

Comparison to Other Studies—The achieved yields of identified HLA-DR peptides were further compared with recent studies in the literature (Fig. 7). On average we identified 10.9 peptides/million EBV-B cells and 9.4 peptides/million BAL cells (with a median of 9.9 and 6.9 peptides/million cells, for EBV B and BAL cells, respectively). The total average yield from all experiments (EBV-B cells and BAL cells) was 10.1 peptides/million cells (with a median of 7.30). A study on BAL
Optimization of the Methodology to Identify HLA-DR Peptides

Fig. 7. **Comparison of yields with other studies.** The graph presents the achieved yields of identified HLA-DR bound peptides from EBV-B cells and BAL cells compared with other studies. The total average yield from all experiments (EBV-B cells and BAL cells) achieved 10.1 peptides/million cells (with a median of 7.30). Wahlström et al. (23) identified 0.1 peptides/million cells from a cell pool of $800 \times 10^6$ BAL cells. Adamopoulou et al. identified on average 0.8 peptides/million cells (with a median of 0.7) using dendritic cells (DCs).

cells in 2007 (23) showed a yield of 0.1 identified peptides/million cells in a single pool of $800 \times 10^6$ BAL cells and a recent study on dendritic cells (17) identified on average 0.8 peptides/million cells (with a median of 0.7) in four samples. Compared with the two former studies, the yield of identified HLA-DR peptides in our work is increased 100- and 13-fold, respectively.

**DISCUSSION**

By implementing a series of significant modifications to conventional techniques used in immunopeptidomics, we established an optimized method to identify HLA-DR bound peptides using low cell numbers—that is, $10 \times 10^6$ cells. This cell number corresponds to what can be obtained from the lungs of a single healthy subject. Thus, our proposed method is a valuable tool to study individual immunopeptidomes of the lung in health and disease, and therefore suitable to reveal disease-specific peptides. The proposed method consists of several modifications, including isolation of crude membranes, the use of a MS-compatible detergent, elution using organic solvents, identification of endogenous peptides in the flow through of the immunoprecipitation step to use as a negative control, and peptide filtering after database search using an in-house developed software.

The validation and optimization of the proposed method was achieved with an EBV-B cell line using $10 \times 10^6$ cells to mimic the available number of BAL cells from healthy controls to evaluate several steps in the methodology. We could achieve the highest median number of peptides per million cells using a set up that includes enrichment of HLA-DR peptides from membranes after overnight incubation with anti-HLA-DR antibody. Furthermore, we applied our optimized approach to human material using BAL cells with different cell counts ($10^8$–$10^9$ cells). In addition, we included matching negative controls by identifying peptides present in the flow through from immunoprecipitation, which were then used during filtering by the ClusterMHCII software in order to generate “true” HLA-DR peptides. The applied changes offered a tremendous improvement to attain peptide identification of HLA-DR complexes.

In total, we identified 507 peptides (from 75 proteins) in the EBV-B cells and 1434 peptides (from 215 proteins) in the BAL samples with an average yield of 10.1 peptides/million cells. Our achieved yields represent a great improvement when compared with two previous studies on BAL cells (23) and dendritic cells from the thymus (17), that offered a yield of 100- and 13-fold lower, respectively. Though, it should be acknowledged that some of the increase is likely because of the improvement in the sensitivity of the instrumentation and using a database searching strategy instead of de novo sequencing. To identify hundreds to thousands of HLA-I peptides, at least 1 g of tissue is required (14), whereas $10 \times 10^6$ cells yield protein amounts in the order of one milligram. Previous studies on human tissue (15, 16, 18) make it difficult to compare yields in terms of cell count. However, after removing redundancies in these data sets and only counting unique sequences once, the number of identified peptides turned out to be lower or of the same order of magnitude as in this study.

Our method is material economical because it requires only $10 \times 10^6$ cells for peptide identification. Additionally, our method offers a higher likelihood of reproducibility given that current techniques lacked the inclusion of negative controls and utilization of replicates, making reproducibility challenging. Furthermore, the described procedure shows a large improvement in speed and it currently takes less than a week from cell lysis to identification of peptides. The peptides identified in the flow through are the most abundant peptides in the sample, and although unlikely, originate from HLA-DR, these peptides are most likely to be nonspecific. Because both sets of peptides can be isolated and identified from the same sample, the flow through peptides constitute an excellent negative control, for the ClusterMHCII software.

Current methodologies for peptide and protein identification frequently require high-throughput identification that demands automatic evaluation with a minimum of manual intervention. Thus, a statistical evaluation, which is a standard practice for analyzing mass spectrometry data in proteomics, is necessary. The high mass accuracy of modern mass spectrometers makes it possible to use stringent error tolerances during the database searching, which significantly reduces the FDR. Peptides such as HLA-DR and endogenous peptides are in general more challenging to identify than the tryptic peptides, which are analyzed during routine proteomic
analysis. Multiply-protonated tryptic peptides with one strongly basic chemical group \(i.e.,\) have either lysine or arginine at the C terminus, and the much less basic free amine at the N terminus) generates MS/MS spectra of higher quality and therefore yield better scores as opposed to endogenous peptides (which might contain more than one strongly basic amino acid, and/or with the basic residue(s) at other positions than at the C- or N terminus). Most MS/MS search engines, such as Mascot, are "protein oriented", \(i.e.,\) are based on determining whether the protein is present in the sample. In our experiments, where a single or only a few peptides per proteins were identified, lower and sometimes sub-threshold scores were achieved when compared with "normal" proteomics experiments. Also given, that there are two sources of nonrelevant peptides: (1) false positive identification where the number of false positives can be evaluated and controlled and (2) presence of nonrelevant peptides: (1) false positive identification where the number can be reduced by applying more stringent conditions during the isolation and/or filtering the peptide lists post-acquisition. We eliminated the former source of nonrelevant peptides by using Percolator adjusted Mascot scores (where a score of 13 corresponds to an FDR < 0.05), and addressed the latter source by utilizing more stringent washing during immunoprecipitation, and by filtering the peptide lists after acquisition, which resulted in peptide identification of higher quality.

HLA-DR derived peptides frequently showed patterns of overlapping sequences, \(i.e.,\) sharing the core-binding motif at few sites per protein and therefore low sequence coverage. Non-HLA binding ("nonspecific") peptides would either be present in the sample at the time of cell lysis or be generated during the procedure through proteolysis by proteases in the sample, and would be co-purified during the affinity purification. Thus to obtain reliable results, nonspecific peptides need to be either removed during the sample preparation procedure or be taken into account during the analysis process. Following this, database searching is also a crucial step for authentic peptide identification.

Previous studies such as (16, 17, 23) have used partial \textit{de novo} sequencing followed by database searching using sequence tags whereas another study (15) has used three different algorithms to search the data and only accepted peptides identified by at least two search engines. Also, several of the studies have manually validated all MS/MS spectra or relied on the Mascot MS/MS search engine for the database searching. In our proposed method, the sequences used were from the UniProt human complete proteome database that is similar to the IPI database. Although we did not use \textit{de novo} sequencing, we were able to determine the maximum length of the consecutive b- or y-ion series for each identified peptide and minimize the FDR (as calculated by Mascot) by optimizing the error tolerance for precursor and fragment masses. Nevertheless, the integration of complementary fragmentation techniques followed by computational \textit{de novo} sequencing are possible future improvements.

In order to achieve high-throughput identification of HLA-DR peptides and to prevent subjective interpretations of the data, it was necessary to use a score, based on statistics \(e.g.,\) Mascot scoring) and to evaluate the results automatically, thus minimizing the number of spectra that need to be manually validated. In the proposed method, we exploited Mascot scoring and filtered the data using the ClusterMHCII software, rendering the analysis a statistical-based approach that can be performed routinely in an automated manner.

Regarding source proteins identified in EBV-B cells and BAL cells, many of the findings included proteins involved in antigen processing and presentation \(e.g.,\) HLA-A, HLA-B, HLA-C, and CatS), as well as proteins that likely originate from autophagy \(e.g.,\) Hsp70, glyceraldehyde-3-phosphate dehydrogenase and lysosomal proteases). These proteins were also found in a previous study using other cell lines (36). One peptide from a protein of particular interest identified in BAL has previously been shown to stimulate T cell proliferation (37).

The HLA-DR eluted peptides showed narrow length distributions with a typical length of HLA-II derived peptides, in agreement with published results (15–18, 23, 38). Interestingly, BAL showed a narrower distribution compared with EBV-B cells. This may be explained either by that the EBV-B cells are a cell line that has not been exposed to the same environment as the BAL cells \(i.e.,\) only media) or that lung macrophages (BAL cells) and B cells have indeed narrower peptide length distributions. In contrast, flow through peptides and tryptic peptides have wider length distributions. The predicted binding motif of peptides identified from membrane complexes of EBV-B cells is largely consistent with previously determined motifs for the same HLA-DRB1*0401 allele variant (34). Also calculated binding strengths predicted the majority of identified peptides to be binders to the HLA-DR allele type. This suggests that the identified peptides are indeed true HLA-DR derived peptides.

The main limitation in our study is the large sample-to-sample variation in peptides yields. Although the procedure is further improved, it is still laborious and the high number of samples processed during a short period might have impact on the reproducibility. For primary cells as a whole, the biological samples showed somewhat consistent yields, even though only a single or a few replicates of each patient sample were analyzed. Also, most abundant peptides from the same subset of proteins were identified in all or most of the replicates regardless of the yield, underscoring a reproducible identification. Another limitation is that because we require two overlapping peptides to be considered as a core sequence, any HLA-DR peptide present in the data set once would be omitted. However, because all data has been saved, it is possible to re-inquire any time for additional peptides of interest.
Optimization of the Methodology to Identify HLA-DR Peptides

CONCLUSIONS

In the presented study, we were able to optimize the procedure for HLA-DR peptide elution to obtain from individual samples the same or larger yield using fewer cells compared with recent studies. We have also evaluated individual steps in the process using a B cell line with multiple replicates, making it possible to compare the efficiency of our procedure with previously published techniques. For identifying biomarkers, particularly of lung diseases, it is essential to be able to work with biological material down to $10^6$ cells because healthy nonsmokers usually yield about $10^6$ cells, whereas smokers and/or patients with inflammatory lung disease typically yield in excess of $50 \times 10^6$ cells. The results achieved using the proposed method demonstrate that it is possible to study the HLA-DR immunopeptidome of both healthy and diseased individuals. Identified peptides could be utilized as biomarkers to help setting diagnosis or follow-ups on disease progression. Additionally, if high enough sensitivity is achieved, disease specific agents may also be identified. Identification and characterization of T cell responses to such antigens may help clarify the immune pathogenesis of chronic inflammatory diseases, such as autoimmune disease, and provide a rational basis for novel treatment strategies.

Acknowledgments—We thank the research nurses Heléne Blomqvist, Margitha Dahl and Gunnar de Forest as well as biomedical analyst Benita Dahlberg for assistance with sample collection and preparation. We also extend our thanks to Prof. Stephan Stefanovic for generously introducing the approaches to immunopeptide elution utilized in his laboratory and to Prof. Morten Nielsen for help with NNAlign software.

* This study was supported by the Swedish Foundation for Strategic Research (through the PRIIMI project), Knut and Alice Wallenberg foundation, King Gustaf V’s 80-year foundation, The Swedish Research Council, The Swedish Heart Lung Foundation, The King Oscar II Jubilee Foundation, The Stockholm County Council, The Swedish Association for Chest Physicians, the EU Innovative Medicine initiative BeTheCure grant and Karolinska Institutet.

** This article contains supplemental material.

†† These authors share senior authorship.

E-mail addresses: Tina Heyder tina.heyder@ki.se; Maxie Kohler maxie.kohler@web.de; Nataliya K. Tarasova nataliya.tarasova@ki.se; Sabrina Haag sabrina_haag@dfci.harvard.edu; Dorothea Rutishauser dorothea.rutishauser@gmail.com; Natalia V. Rivera natalia.valeria.rivera@ki.se; Charlotte Sandin charlotte.sandin@gmail.com; Sohel Mia sohel.mia@ki.se; Vivianne Malmström vivianne.malmstrom@ki.se; Åsa M. Wheelock asa.wheelock@ki.se; Jan Wahlström jan.wahlstrom@ki.se; Rikard Holmdahl rikard.holmdahl@ki.se; Anders Eklund anders.eklund@ki.se; Roman A. Zubarev roman.zubarev@ki.se; Johan Grunevald johan.grunevald@ki.se; A. Jimmy Ytterberg jimmy.ytterberg@ki.se.

REFERENCES

1. (2000) Autoimmune disease. Nature biotechnology 18 Suppl., IT7–9
2. Fridkis-Harel, M. (2013) Design of peptide immunotherapies for MHC Class-II-associated autoimmune disorders. Clin. Dev. Immunol. 2013, 826191, 2013
3. Chun-Lai, T., Padyukov, L., Dhaliwal, J. S., Lundstrom, E., Yahya, A., Muhammad, N. A., Klareskog, L., Alfredsson, L., Larsson, P. T., Murad, S., and Malaysian Epidemiological Investigation of and Rheumatoid Arthritis Study, G. (2011) Shared epitope alleles remain a risk factor for autoantibody positive rheumatoid arthritis in three Asian ethnic groups. PLoS ONE 6, e21069
4. Patsopoulos, N. A., Barcellos, L. F., Hintzen, R. Q., Schaefer, C., van Duijn, C. M., Noble, J. A., Raj, T., Immgen Anzgene Gourraud, P. A., Stranger, B. E., Oksenberg, J., Olsson, T., Taylor, B. V., Sawcer, S., Hafer, D. A., Carrington, M., De Jager, P. L., and de Bakker, P. I. (2013) Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. PLoS Genet. 9, e1003926
5. Chun-Lai, T., Mair, S., Dhaliwal, J. S., D’Souza, R., Niel, M., Cederlund, K., Hillert, J., and Eklund, A. (2010) Different HLA-DRB1 allele distributions in distinct clinical subgroups of sarcoidosis patients. Respir. Rev. 11, 25
6. Segura, E., and Amigorena, S. (2015) Cross-presentation in mouse and human dendritic cells. Adv. Immunol. 127, 1–31
7. Kamradt, T., and Mitchison, N. A. (2001) Tolerance and autoimmunity. New Engl. J. Med. 344, 655–664
8. Gillard, F., Sie, C., Streu, K., Ulaganathan, V. K., Schlager, C., Lodygin, D., Heckelsmiller, K., Nietzsche, K., Ellwart, J., Klinkert, W., Lottaz, C., Nosov, M., Brinkmann, V., Spang, R., Lehrah, H., Vingeron, M., Wekerle, H., Flugel-Koch, C., and Flugel, A. (2012) T cells become licensed in the lung to enter the central nervous system. Nature 488, 675–679
9. (1999) Statement on sarcoidosis. Joint Statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS) and the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) adopted by the ATS Board of Directors and by the ERS Executive Committee, February 1999. Am. J. Resp. Critical Care Med. 160, 736–755
10. Chen, E. S., and Moller, D. R. (2011) Sarcoidosis–scientific progress and clinical challenges. Nat. Rev. Rheumatol. 7, 457–467
11. Falk, K., Rotzschke, O., Stevanovic, S., Jung, G., and Rammensee, H. G. (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature 351, 290–296
12. Hunt, D. F., Henderson, R. A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevill, N., Cox, A., L., Appella, E., and Engelhard, V. H. (1992) Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. Science 255, 1261–1263
13. Gordon, R. D., Young, J. A., Rayner, S., Luke, R. W., Crowther, M. L., Wordsworth, P., Bell, J., Hassal, G., Evans, J., Hinchcliffe, S. A., and et al. (2015) Purification and characterization of endogenous peptides extracted from HLA-DR isolated from the spleen of a patient with rheumatoid arthritis. Eur. J. Immunol. 25, 1473–1476
14. Caron, E., Kowalewski, D. J., Chieh Koh, C., Sturm, T., Schuster, H., and Aebersold, R. (2015) Analysis of major histocompatibility complex (MHC) immunopeptidomes using mass spectrometry. Mol. Cell. Proteomics 14, 3105–3117
15. Seward, R. J., Drouin, E. E., Steere, A. C., and Costello, C. E. (2011) Peptides presented by HLA-DR molecules in synovia of patients with rheumatoid arthritis or antibiotic-refractory Lymph arthritis. Mol. Cell. Proteomics 10, M110 002477
16. Fissolo, N., Haag, S., de Graaf, K. L., Drews, O., Stevanovic, S., Rammensee, H. G., and Weisssert, R. (2009) Naturally presented peptides on major histocompatibility complex I and II molecules eluted from central nervous system of multiple sclerosis patients. Mol. Cell. Proteomics 8, 2090–2101
17. Adamopoulou, E., Tenzer, S., Hifen, N., Klug, P., Rota, I. A., Tietz, S., Gebhardt, M., Stevanovic, S., Schild, H., Tolosa, E., Melms, A., and Stoeckle, C. (2013) Exploring the MHC-epitope matrix of central tolerance in the human thymus. Nat. Commun. 4, 2039
18. Collado, J. A., Alvarez, I., Ciudad, M. T., Espinosa, G., Canals, F., Pujol-Borrell, R., Carrascal, M., Abian, J., and Jaraquemada, D. (2013) Composition of the HLA-DR-associated human thymus peptidome. Eur. J. Immunol. 43, 2273–2282
19. Hunninghake, G. W., Costabel, U., Ando, M., Baughman, R., Cordier, J. F., do Bois, R., Eklund, A., Kitaichi, M., Lynch, J., Rizzato, G., Rose, C., Selroos, O., Semenzato, G., and Sharma, O. P. (1999) ATS/ERS/WASOG statement on sarcoidosis. American Thoracic Society/Euro
Optimization of the Methodology to Identify HLA-DR Peptides

patients have a strong bias toward citrullinated autoantigen recognition. J. Exp. Med. 210, 445–455
29. Andreatta, M., Schafer-Nielsen, C., Lund, O., Buus, S., and Nielsen, M. (2011) NNAlign: a web-based prediction method allowing non-expert end-user discovery of sequence motifs in quantitative peptide data. PLoS ONE 6, e26781
30. Karosiene, E., Rasmussen, M., Blicher, T., Lund, O., Buus, S., and Nielsen, M. (2013) NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isoforms, HLA-DR, HLA-DP and HLA-DQ. Immunogenetics 65, 711–724
31. Mi, H., Muruganujan, A., and Thomas, P. D. (2013) PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res. 41, D377–D386
32. Haag, S., Tuncel, J., Thordardottir, S., Mason, D. E., Yau, A. C. Y., Dobritzsch, D., Bäcklund, J., Peters, E. C., and Holmdahl, R. (2015) Positional identification of RT1-B (HLA-DQ) as susceptibility locus for autoimmune arthritis. J. Immunol. 194, 2539–2550
33. Tarasova, N. K., Ytterberg, A. J., Lundberg, K., Zhang, X. M., Harris, R. A., and Zubarev, R. A. (2015) Proteomics reveals a role for attachment in monocyte differentiation into efficient proinflammatory macrophages. J. Proteome Res. 14, 3940–3947
34. Rapin, N., Hoef, I., Lund, O., and Nielsen, M. (2010) The MHC motif viewer: a visualization tool for MHC binding motifs. Curr. Protoc. Immunol. Chapter 18, Unit 18 17
35. Berlin, M., Fogdell-Hahn, A., Olerup, O., Eklund, A., and Grunewald, J. (1997) HLA-DR predicts the prognosis in Scandinavian patients with pulmonary sarcoidosis. Am. J. Respiratory Critical Care Med. 156, 1621–1625
36. Dengjie, J., Schoor, O., Fischer, R., Reich, M., Kraus, M., Muller, M., Kreyberg, K., Altenberend, F., Brandenburg, J., Kalbacher, H., Brock, R., Driessen, C., Rammensee, H. G., and Stevanovic, S. (2005) Autoantibody promotes MHC class II presentation of peptides from intracellular source proteins. Proc. Natl. Acad. Sci. U.S.A. 102, 7922–7927
37. Wahlstrom, J., Dengjie, J., Wingqvist, O., Targoff, I., Persson, B., Duyar, H., Rammensee, H. G., Eklund, A., Weissert, R., and Grunewald, J. (2009) Autoimmune T cell responses to antigenic peptides presented by bronchoalveolar lavage cell HLA-DR molecules in sarcoidosis. Clin. Immunol. 133, 353–363
38. Rammensee, H. G. (1995) Chemistry of peptides associated with MHC class I and class II molecules. Curr. Opin. Immunol. 7, 85–96
39. Rapin, N., Hoef, I., Lund, O., and Nielsen, M. (2008) MHC motif viewer. Immunogenetics 60, 759–765