

Identification of HLA-DR-bound peptides presented by human bronchoalveolar lavage cells in sarcoidosis

Jan Wahlström, Jörn Dengjel, Bengt Persson, Hüseyin Duyar, Hans-Georg Rammensee, Stefan Stevanović, Anders Eklund, Robert Weissert, Johan Grunewald

J Clin Invest. 2007;117(11):3576-3582. <https://doi.org/10.1172/JCI32401>.

Technical Advance

Sarcoidosis is an inflammatory disease of unknown etiology, most commonly affecting the lungs. Activated CD4⁺ T cells accumulate in the lungs of individuals with sarcoidosis and are considered to be of central importance for inflammation. We have previously shown that Scandinavian sarcoidosis patients expressing the HLA-DR allele DRB1*0301 are characterized by large accumulations in the lungs of CD4⁺ T cells expressing the TCR AV2S3 gene segment. This association afforded us a unique opportunity to identify a sarcoidosis-specific antigen recognized by AV2S3⁺ T cells. To identify candidates for the postulated sarcoidosis-specific antigen, lung cells from 16 HLA-DRB1*0301^{pos} patients were obtained by bronchoalveolar lavage. HLA-DR molecules were affinity purified and bound peptides acid eluted. Subsequently, peptides were separated by reversed-phase HPLC and analyzed by liquid chromatography–mass spectrometry. We identified 78 amino acid sequences from self proteins presented in the lungs of sarcoidosis patients, some of which were well-known autoantigens such as vimentin and ATP synthase. For the first time, to our knowledge, we have identified HLA-bound peptides presented *in vivo* during an inflammatory condition. This approach can be extended to characterize HLA-bound peptides in various autoimmune settings.

Find the latest version:

<https://jci.me/32401/pdf>





Identification of HLA-DR-bound peptides presented by human bronchoalveolar lavage cells in sarcoidosis

Jan Wahlström,¹ Jörn Dengjel,² Bengt Persson,^{3,4} Hüseyin Duyar,⁵ Hans-Georg Rammensee,² Stefan Stevanović,² Anders Eklund,¹ Robert Weissert,⁵ and Johan Grunewald¹

¹Lung Research Laboratory, Department of Medicine, Karolinska Institutet, Stockholm, Sweden. ²Department of Immunology, Institute for Cell Biology, University of Tübingen, Tübingen, Germany. ³Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm, Sweden. ⁴IFM Bioinformatics, Linköping University, Linköping, Sweden. ⁵Experimental Neuroimmunology Laboratory, Hertie Institute for Clinical Brain Research, Department of Neurology, University of Tübingen, Tübingen, Germany.

Sarcoidosis is an inflammatory disease of unknown etiology, most commonly affecting the lungs. Activated CD4⁺ T cells accumulate in the lungs of individuals with sarcoidosis and are considered to be of central importance for inflammation. We have previously shown that Scandinavian sarcoidosis patients expressing the HLA-DR allele DRB1*0301 are characterized by large accumulations in the lungs of CD4⁺ T cells expressing the TCR AV2S3 gene segment. This association afforded us a unique opportunity to identify a sarcoidosis-specific antigen recognized by AV2S3⁺ T cells. To identify candidates for the postulated sarcoidosis-specific antigen, lung cells from 16 HLA-DRB1*0301^{pos} patients were obtained by bronchoalveolar lavage. HLA-DR molecules were affinity purified and bound peptides acid eluted. Subsequently, peptides were separated by reversed-phase HPLC and analyzed by liquid chromatography–mass spectrometry. We identified 78 amino acid sequences from self proteins presented in the lungs of sarcoidosis patients, some of which were well-known autoantigens such as vimentin and ATP synthase. For the first time, to our knowledge, we have identified HLA-bound peptides presented in vivo during an inflammatory condition. This approach can be extended to characterize HLA-bound peptides in various autoimmune settings.

Introduction

Sarcoidosis is a granulomatous disease of unknown etiology. It affects both sexes and all races, with the highest prevalence rates reported for Swedes, Danes, and US African Americans. The peak incidence is in adults less than 40 years of age. Although a systemic disease, it most commonly affects the lungs. The majority of patients recover, but some develop chronic disease that may result in fibrosis and respiratory failure (1). Significant morbidity and mortality may also occur when there is multorgan involvement, in particular cardiac and neurological. Most patients undergo bronchoscopy with bronchoalveolar lavage (BAL) as a routine investigation. Therefore, cells from the site of inflammation, i.e., the lungs, are readily available for investigation. Activated CD4⁺ T cells with a Th1 cytokine profile are accumulated in the lungs and are considered to be of central importance for the inflammation (2).

Our group has previously shown that a subgroup of patients, defined by the presence of the HLA isotype and allele DRB1*0301 (formerly also known as HLA-DR17) and encompassing one-third of Scandinavian sarcoidosis patients, differs strikingly from other patients in several respects. HLA-DRB1*0301^{pos} patients have a distinct tissue-specific local immune response, as a particular TCR- α chain, variable gene segment 2.3 (AV2S3), is expressed by large numbers of CD4⁺ T cells in the lungs, but not in blood, of

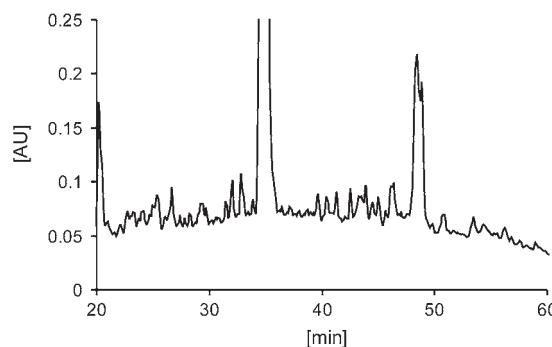
DRB1*0301^{pos} sarcoidosis patients (3). The same phenomenon is not found in DRB1*0301^{pos} healthy individuals or in those with other inflammatory disorders (3). The AV2S3⁺ BAL T cells associate with disease activity (4), and higher numbers of AV2S3⁺ cells at disease onset also correlated with a better prognosis, indicating that they may in fact have a protective role (5). Lung-restricted accumulations of CD4⁺AV2S3⁺ cells may also correlate with the presence of HLA-DR13, at least in HLA-DR13^{pos} patients carrying the HLA-DRB3*0101 allele (6). DRB3*0101 molecules are known to be structurally similar to and to be able to present similar antigenic peptides as DRB1*0301 molecules (7).

The association between DRB1*0301 and lung-restricted expansions of CD4⁺ AV2S3⁺ T cells provides a unique opportunity to identify a sarcoidosis-specific antigen(s). Our hypothesis is that the AV2S3⁺ cells have recognized and proliferated in response to such an antigen. The aim of this study was to investigate which antigens are presented in vivo in the lungs of patients with sarcoidosis, in particular in HLA-DRB1*0301^{pos} patients. Technological advances allow the analysis of HLA-presented peptides from as few as 10⁹ cells. Using state-of-the-art mass spectrometers, HLA class I- as well as class II-presented peptides from human tumor specimens have been characterized, highlighting the possibility of identifying HLA-bound peptides presented in vivo (8, 9). In order to identify antigenic peptides presented in vivo in the lungs of sarcoidosis patients, we obtained cells by means of BAL and isolated HLA-DR-bound peptides. We were able to characterize 230 amino acid sequence tags by mass spectrometry (MS). Subsequently, database searches were applied to identify the corresponding human source proteins. Seventy-eight unique peptide sequences could be assigned.

Nonstandard abbreviations used: AV2S3, α chain variable gene segment 2.3; BAL, bronchoalveolar lavage; LC, liquid chromatography; MS, mass spectrometry; MS/MS, tandem MS; RP, reversed phase.

Conflict of interest: The authors have declared that no conflict of interest exists.

Citation for this article: *J. Clin. Invest.* 117:3576–3582 (2007). doi:10.1172/JCI32401.

**Figure 1**

RP-HPLC chromatogram of HLA-DR-bound peptides purified from BAL cells using the mAb L243 and recorded at 214 nm.

Results

HLA-DR-bound peptides and corresponding source proteins from BAL cells of sarcoidosis patients. To our knowledge, naturally presented HLA-bound peptides from patients suffering from chronic inflammatory diseases have thus far not been isolated and identified. As the detection limit of modern mass spectrometers lies routinely in the low femtomolar, high attomolar range, one would need a minimum of 10^9 cells to detect peptides presented only once per cell, taking 50% unspecific losses during the purification procedure into account and not addressing ionization properties important for mass spectrometric analysis. These cell numbers can be reached by combining BAL cells of several sarcoidosis patients. We collected BAL cells of 16 DRB1*0301^{pos} sarcoidosis patients, which were directly frozen at -80°C (see Methods) and stored for further processing. Subsequently, the combined cells (approximately 8×10^8 cells) were used to isolate HLA-DR-bound peptides according to standard protocols (10).

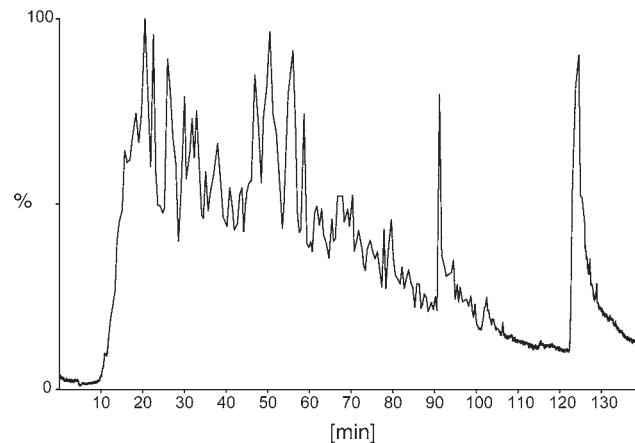
Most commonly, MALDI MS is used to generate peptide mass fingerprints for protein identification. This kind of analysis lacks sequence information. However, as HLA-bound peptides can derive from virtually every protein within a cell, resulting peptide samples represent very complex multiprotein mixtures. Thus, peptides have to be sequenced in order to be identified unambiguously. We decided to use nanocapillary liquid chromatography-MS (LC-MS) analysis to generate such data. One advantage of common LC-MS systems is that they can acquire data in an automated fashion, greatly reducing analysis time. The mass spectrometer used was a quadrupole orthogonal acceleration time-of-flight (Q-TOF) type of instrument yielding high quality MS and tandem MS (MS/MS) data, which is reflected by mass accuracies of less than ± 0.2 Da in the MS as well as the MS/MS mode.

In order to maximize the number of mass spectrometric sequencing events and the number of identified peptides, a reversed phase-HPLC (RP-HPLC) separation of isolated peptides was performed prior to LC-MS analysis (Figure 1). Sixty fractions of 150 μl each were collected, and subsequently every other peptide-containing fraction was combined to perform 2 data-dependent LC-MS experiments (Figure 2). As it is not uncommon to sequence only 1 single peptide from 1 specific source protein, general rules for mass spectrometric proteome analysis, as for example, the identification of at least 2 peptides per protein (11), do not apply for HLA peptide analysis (12). This has the consequence that data interpretation is still routinely done manually and that the number of unambigu-

ously identified peptides is rather low as compared with the number of mass spectrometric sequencing events. In our case, a total of 1368 sequencing events yielded 230 amino acid sequence tags of at least 4 consecutive amino acid residues. Subsequently, database searches were applied to identify peptides and corresponding endogenous (human) source proteins, since endogenous peptides can be expected to constitute a majority of peptides presented by HLA molecules (12, 13). A nonspecific search, including known sequences of microbes, would yield too many false-positive hits, greatly complicating data interpretation. Seventy-eight significant matches ($P < 0.05$; MASCOT software) were obtained (Table 1), leaving the remaining sequence tags unsolved.

We also used FASTA searches as described in Methods (FASTA searches are performed on 1 stretch of consecutive amino acids, in contrast with MASCOT searches, which can be done on a combination of consecutive amino acids identified either from the amino- or the carboxy terminus, in combination with the peptide mass; peptide numbers below refer to those in Table 1). The longest segments with 100% residue identity with known proteins were found for peptide 71 with 10 consecutive residues identical with transferrin (UniProtKB accession number P02787) and for peptide 49 with 10 consecutive residues identical with C4b-binding protein α chain (UniProtKB accession number P04003). Among peptides with nine consecutive identities to known human sequences, we noticed 4 findings of potential interest: peptide 35 with vimentin (UniProtKB accession number P08670), peptide 61 with HLA class II histocompatibility antigen (MHC class II antigen DRB3*1; UniProtKB accession number P79483), peptide 64 with lysyl-tRNA synthetase (UniProtKB accession number Q15046), and peptide 54 with fibronectin (UniProtKB accession number P02751).

The proteins to which the identified peptides were matched could be classified into the following categories: cytoskeletal proteins (β -actin, filamin A, syntenin, vimentin), enzymes (24-dehydrocholesterol reductase, alkaline phosphatase, ATP synthase, glutaredoxin, hexosaminidase B, hydroxyacyl-coenzyme A dehydrogenase, lysyl-tRNA synthetase, pronapsin A), serum proteins (albumin, α -1-antitrypsin, α -2-macroglobulin, apolipoprotein A-II, ceruloplasmin, complement component 2, complement component 4-binding protein- α , complement component 5, fibronectin,

**Figure 2**

Total ion current of peptides eluted from HLA-DR molecules and separated by RP-HPLC. Individual peptides were subjected to fragmentation in a collision cell.

**Table 1**

HLA-DR-bound peptides from pooled BAL cells of DRB1*0301-positive sarcoidosis patients

| Sequence -3-2-1 1 2 3 4 5 6 7 8 9 +1+2+3 | Gene symbol | Gene | Entrez Gene ID | SYFPEITHI score | Δ Mass | No. |
|--|------------------------------------|---|-------------------|--------------------|-----------|-----|
| Strong HLA-DRB1*0301-binding peptides | | | | | | |
| S S K F Q V D N N N R L | A2M | α-2-Macroglobulin | 2 | 33 | -0.11 | 1 |
| E L R V A P E E H P V L L T | ACTB | Actin, β | 60 | 20 | -0.08 | 2 |
| N E L R V A P E E H P V L L T | ACTB | Actin, β | 60 | 20 | -0.13 | 3 |
| Y N E L R V A P E E H P V L L T | ACTB | Actin, β | 60 | 20 | -0.09 | 4 |
| T P T L V E V S R N L G K V G | ALB | Albumin | 213 | 21 | -0.10 | 5 |
| V P P V Q V S P L I K L G R Y S A L | ATP5I | ATP synthase, subunit E | 521 | 21 | -0.26 | 6 |
| H A F I L Q D T K A L H Q V | C2 | Complement component 2 | 717 | 30 | -0.09 | 7 |
| H A F I L Q D T K A L H Q V F E | C2 | Complement component 2 | 717 | 30 | -0.10 | 8 |
| H D G H D D D V I D I E D D L | CANX | Calnexin | 821 | 25 | 0.09 | 9 |
| S G T L V L L Q G A R G F A | CD14 | CD14 molecule | 929 | 22 | -0.08 | 10 |
| N Y R I D T I N L F P A | CP | Ceruloplasmin | 1356 | 21 | -0.05 | 11 |
| N K N Y R I D T I N L F P | CP | Ceruloplasmin | 1356 | 21 | -0.05 | 12 |
| Q D I I P F G N N P I F R | DHCR24 | 24-Dehydrocholesterol reductase | 1718 | 24 | -0.11 | 13 |
| K S T I T L D G G V L V H | FABP4 | Fatty acid-binding protein 4 | 2167 | 25 | -0.16 | 14 |
| E T V I T V D T K A A G K G K | FLNA | Filamin A, α | 2316 | 29 | -0.06 | 15 |
| E T V I T V D T K A A G K G K | FLNA | Filamin A, α | 2316 | 29 | -0.15 | 16 |
| K V N V D E V G G E A L G R L | HBB | Hemoglobin, β | 3043 | 21 | -0.04 | 17 |
| K V N V D E V G G E A L G R L | HBB | Hemoglobin, β | 3043 | 21 | -0.13 | 18 |
| L V V Y P W T Q R F | HBB | Hemoglobin, β | 3043 | 20 | -0.11 | 19 |
| K V N V D A V G G E A L G R L | HBD | Hemoglobin, δ | 3045 | 20 | -0.14 | 20 |
| S V R F D S D V G E Y | HLA-DRB1 | MHC class II, DR β 1 | 3123 | 23 | 0.03 | 21 |
| F L R F D S D V G E Y | HLA-DRB3 | MHC class II, DR β 3 | 3125 | 23 | -0.02 | 22 |
| K D L F K A V D A A L K K | KNG1 | Kininogen 1 | 3827 | 20 | -0.17 | 23 |
| A F Q G D E S V R I D A | LRP1 | Low-density lipoprotein-related protein 1 | 4035 | 20 | -0.02 | 24 |
| E Q A F Q G D E S V R I D A | LRP1 | Low-density lipoprotein-related protein 1 | 4035 | 20 | 0.08 | 25 |
| E K N I M L Y K G S G L W S | MRC1 | Mannose receptor, C type 1 | 4360 | 21 | -0.04 | 26 |
| V E Q G L L D K P V F S F | NAPS1 | Napsin A aspartic peptidase | 9476 | 21 | -0.09 | 27 |
| K K L K E K R Q E Q L O K | NFE2L2 | Nuclear factor (erythroid-derived 2)-like 2 | 4780 | 24 | -0.17 | 28 |
| K P E P E L V Y E D L R G S V T F H | PIGR | Polymeric immunoglobulin receptor | 5284 | 28 | -0.07 | 29 |
| V A R I V G N S G L N I Y N L | CTSA | Cathepsin A | 5476 | 20 | -0.10 | 30 |
| I T S I V K D S S A A R N G L | SDCBP | Syndecan binding protein | 6386 | 24 | -0.14 | 31 |
| A I F F L P D E G K L Q | SERPINA1 | α-1-Antitrypsin | 5265 | 24 | -0.04 | 32 |
| K A V L T I D E K G T E A A | SERPINA1 | α-1-Antitrypsin | 5265 | 28 | -0.08 | 33 |
| D S L P L V D T H S K R T | VIM | Vimentin | 7431 | 25 | -0.08 | 34 |
| D S L P L V D T H S K R T L L | VIM | Vimentin | 7431 | 25 | -0.14 | 35 |
| R E T N L D S L P L V D T H S K R T L L | VIM | Vimentin | 7431 | 25 | -0.16 | 36 |
| Weak HLA-DRB1*0301-binding peptides (score 18–19) and other HLA class II-binding peptides | | | | | | |
| G N R I A Q W Q S F Q L | A2M | α-2-Macroglobulin | 2 | 18 | -0.04 | 37 |
| G N R I A Q W Q S F Q L E G | A2M | α-2-Macroglobulin | 2 | 18 | -0.01 | 38 |
| Y K F Q N A L L V R Y T | ALB | Albumin | 213 | 18 | -0.13 | 39 |
| L G E Y K F Q N A L L V R Y T | ALB | Albumin | 213 | 18 | -0.08 | 40 |
| Q L G E Y K F Q N A L L V R Y T K | ALB | Albumin | 213 | 18 | -0.13 | 41 |
| E Q L G E Y K F Q N A L L V R Y T K | ALB | Albumin | 213 | 18 | -0.08 | 42 |
| F L K K Y L Y E I A R R H P Y | ALB | Albumin | 213 | 17 | -0.11 | 43 |
| N N Y Q A Q S A V P L R H | ALPL | Alkaline phosphatase, liver/bone/kidney | 249 | 10 | -0.06 | 44 |
| H N N Y Q A Q S A V P L R H E | ALPL | Alkaline phosphatase, liver/bone/kidney | 249 | 10 | 0.04 | 45 |
| V S Q Y F Q T V T D Y G K D L | AP0A2 | Apolipoprotein A-II | 336 | 19 | 0.05 | 46 |
| P V P P L P E Y G G K V R Y G | ATP5F1 | ATP synthase, subunit B1 | 515 | 18 | -0.08 | 47 |
| P T F K F E D P K F E V I E K P Q A | ATP5J | ATP synthase, subunit F6 | 522 | 18 | -0.02 | 48 |
| F G P I Y N Y K D T I V F K | C4BPA | Complement component 4-binding protein, α | 722 | 16 | -0.10 | 49 |
| K P V Y T P D Q S V K V R V | C5 | Complement component 5 | 727 | 16 | -0.12 | 50 |
| D K P V Y T P D Q S V K V K | C5 | Complement component 5 | 727 | 16 | -0.07 | 51 |
| D K P V Y T P D Q S V K V R V | C5 | Complement component 5 | 727 | 16 | -0.14 | 52 |
| N A D P L K V Y P P L K | CD74 | CD74 molecule (invariant chain) | 972 | 9 | -0.07 | 53 |
| T P G T E Y Y V V S I V A L N G R E | FN1 | Fibronectin 1 | 2335 | 14 | -0.04 | 54 |
| I Q D Y L Q O Q L T G A R T V P R | GLRX | Glutaredoxin | 2745 | 19 | -0.11 | 55 |
| E F F H E A F S G Q I L | HADHB | Hydroxyacyl-coenzyme A dehydrogenase, β subunit | 3032 | 10 | -0.02 | 56 |
| V K E P V A V L K A N R V W G | HEXB | Hexosaminidase B | 3074 | 19 | -0.18 | 57 |
| V K E P V A V L K A N R V W G A L | HEXB | Hexosaminidase B | 3074 | 19 | -0.20 | 58 |
| E A Q G A L A N I A V A D K A N | HLA-DRA | MHC class II, DR α | 3122 | 11 | -0.07 | 59 |
| I K E E H V I I Q A E F | HLA-DRA | MHC class II, DR α | 3122 | 15 | -0.08 | 60 |
| R P V A E S W N S Q K D L L E | HLA-DRB3 | MHC class II, DR β 3 | 3125 | 17 | 0.00 | 61 |
| D T L Y L Q M N S L R A E D T | IGHM | Ig heavy constant μ | 3507 | 18 | 0.06 | 62 |
| G Q A Y D V D F T P P F R | KARS | Lysyl-tRNA synthetase | 3735 | 18 | -0.01 | 63 |
| G P T E E I R A L H A A I G G I P | NAPS1 | Lysyl-tRNA synthetase | 3735 | 18 | 0.07 | 64 |
| G P T E E I R A L H A A I G G | NAPS1 | Napsin A aspartic peptidase | 9476 | 12 | -0.13 | 65 |
| T K V Y T V D L G R T V T I | PIGR | Napsin A aspartic peptidase | 9476 | 12 | -0.09 | 66 |
| D T K V Y T V D L G R T V T I | PIGR | Polymersic immunoglobulin receptor | 5284 | 12 | -0.12 | 67 |
| N P A N P A I L S E A S A P I P H | SDCBP | Syndecan-binding protein | 6386 | 14 | -0.08 | 68 |
| L T E G P S P G D F R Y S | SFTPA1B | Surfactant, pulmonary-associated protein A1B | 6435 | 13 | 0.04 | 69 |
| D P Q T F Y Y A V A V V K K D S G | TF | Transferrin | 7018 | 17 | -0.05 | 70 |
| L E I F K Q A S A F S R A S | TGFBI | TGF, β-induced 68 kDa | 7045 | 18 | -0.03 | 71 |
| N S L Y L Q M N S L R A E D T | | Ig heavy chain V region | 18 | 0.06 | 73 | |
| D R Y F H N Q E E F | | MHC class II | 14 | 0.11 | 74 | |
| Y L Q M N S L R A E D | | Ig heavy chain V region | 13 | -0.03 | 75 | |
| Y L Q M N S L R A E D T | | Ig heavy chain V region | 13 | 0.02 | 76 | |
| N T L Y L Q F N S L R A E D T | | Ig heavy chain V region | 18 | 0.02 | 77 | |
| R V V T E L G R P S A E Y W N S Q K D L L E | HLA-D class II antigen DR1 β chain | Ig heavy chain V region | 19 | 0.01 | 78 | |

Peptides and corresponding source proteins were identified by tandem MS. Peptides are arranged according to the HLA-DRB1*0301-binding motif (<http://www.syfpeithi.de>), indicated by score. Peptides with a score equal to or higher than 20 were regarded as strong binders. Anchor amino acids are in bold.

Table 2
Synthesized peptides

| Protein | Peptide sequence | No. ^a |
|-----------------------|--------------------|------------------|
| Lysyl-tRNA-synthetase | GPEGQAYDVDFTPPFR | 64 |
| Lysyl-tRNA-synthetase | GQAYDVDFTPPFR | 63 |
| ATP synthase | PVPPLEYGGKVRYG | 47 |
| ATP synthase | PTFKFEDPKFEVIEKPQA | 48 |
| ATP synthase | VPPVQVSPLIKLGRYSAL | 6 |
| β -actin | NELRVAPEEHPVLLT | 3 |
| Kininogen | KDLFKAVIDAALKK | 23 |
| Vimentin | DSLPLVDTHSKRTLL | 35 |

Peptides from known autoantigens were synthesized in order to confirm the identity of natural eluted BAL peptides. ^aNumbers are from Table 1.

kininogen, transferrin immunoglobulin), cell surface receptors (CD14 molecule, low-density lipoprotein-related protein 1, macrophage mannose receptor, polymeric immunoglobulin receptor), proteins related to antigen processing and presentation (calnexin, invariant chain), and various others (surfactant pulmonary associated protein A1, TGF- β -induced, fatty acid-binding protein 4, hemoglobin β chain cathepsin A).

Verification of identified peptides. The identified source proteins were analyzed for possible known autoantigens, and indeed, several well-described autoantigens were presented as HLA-bound peptides on the cell surface of BAL cells. Three vimentin peptides (14–16), 3 β -actin peptides (17), 2 lysyl-tRNA synthetase peptides (18, 19), 3 ATP synthase peptides (20), and 1 kininogen peptide (21) could be identified (Table 1). A subset of these peptides was synthesized in order to prove their identity (Table 2). Mass spectrometric analysis was performed on the synthetic peptides, and MS/MS spectra were compared with spectra of the corresponding BAL peptides. In all cases, spectra of the synthetic peptides matched the ones of their eluted natural counterparts, proving the correct interpretation of the recorded mass spectrometric data. Figure 3 shows as an example the MS/MS spectra of the peptide vimentin_{429–443} DSLPLVDTHSKRTLL (no. 35).

Discussion

In this study, using what we believe is a novel approach in the study of chronic inflammatory diseases, we have identified a number of potentially interesting self peptides that are expressed and presented by antigen-presenting cells for T cells at the site of inflammation in a genetically and clinically well-defined subset of patients with sarcoidosis. Several of these derived from normally abundant proteins such as serum proteins. Other source proteins, such as some intracellular enzymes and cytoskeletal components, could better be anticipated to serve as targets for autoimmune attacks (22). Thus, for the first time, to our knowledge, we have analyzed peptides presented *in vivo* in the human lung, demonstrating the feasibility of studying the molecular basis of possibly autoreactive T cells in chronic inflammatory diseases such as sarcoidosis.

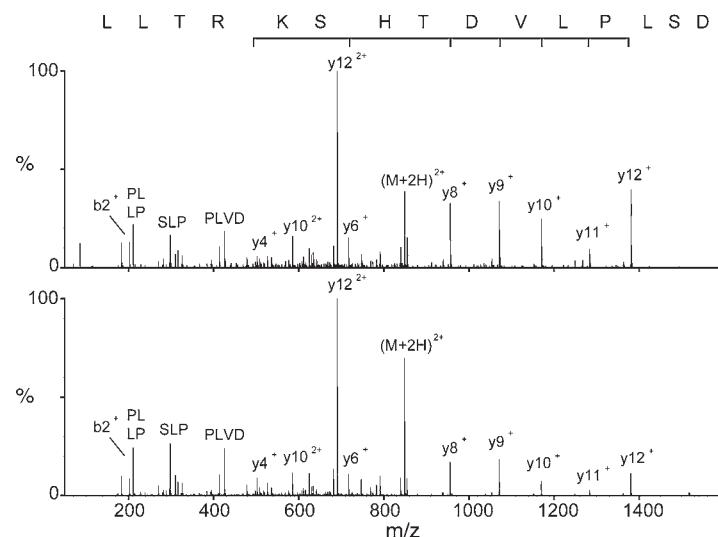
Figure 3

Examples of MS/MS spectra. Collision-induced decay (CID) mass spectra of the natural eluted (upper) and the synthetic (lower) peptide DSLPLVDTHSKRTLL ($(M+2H)^{2+}$ =847.9, vimentin_{429–443}).

Sensitivities of modern mass spectrometers lie routinely in the high attomolar region, allowing the analysis of HLA-bound peptides from a minimal number of cells. These technical advances open up the possibility of characterizing HLA-bound peptides from human specimens, which are only available in a limited amount. We have combined BAL cells of 16 sarcoidosis patients yielding approximately 8×10^8 cells and were able to identify 78 unique HLA-DR-presented peptides. For sequence identification, we used 2 complementary strategies – FASTA searches showing complete sequence identity for peptides identified from MS and MASCOT searches based upon total peptide mass and partial sequence identity. In most cases, both methods gave the same results, strengthening the findings, but in some cases, only the MASCOT search could identify the protein. This was mainly due to the fact that the sequence was too short to be identified with FASTA using our thresholds.

We were able to identify several peptides from proteins that have previously been reported to be autoantigens in various diseases. Vimentin has previously been implicated as an autoantigen in pulmonary inflammation. Levels of anti-vimentin antibodies were previously shown to be significantly higher in patients with idiopathic pulmonary fibrosis and nonspecific interstitial pneumonia than in healthy controls (23). Vimentin has also been described as an autoantigen in rheumatic disease (14–16). Interestingly, it has been reported that pulmonary fibrosis is more common in RA patients who have anti-vimentin antibodies (24). Vimentin is a cytoskeletal protein, also known as intermediate filament polypeptide, and is present in fibroblasts (25) but also in multinucleated and mononuclear cells from peripheral giant cell granulomas (26). Vimentin is also present in sarcoid granulomas, particularly in the multinucleated giant cells' asteroid bodies, which are made up predominantly of vimentin (27).

Lysyl-tRNA synthetase belongs to a family consisting of 20 different aminoacyl tRNA synthetases, each responsible for coupling of a particular amino acid to its corresponding tRNA. Anti-synthetase antibodies are common in patients with polymyositis and dermatomyositis and particularly so in patients with concurrent interstitial lung disease (ILD). Another peptide was derived from H⁺/K⁺ ATP synthase, i.e., ATP synthase. This is a well-known antigen in autoimmune gastritis. However, it is also an autoantigen





targeted by serum antibodies in celiac disease (20). Celiac disease is caused by intolerance to ingested gluten proteins. The precise mechanisms whereby antigen-specific responses, initially directed against gluten proteins (such as gliadin), spread to self antigens are not known. There is a strong genetic linkage to specific HLA class II antigens, in particular HLA-DQ2, which is expressed by approximately 95% of patients with celiac disease (28). HLA-DQ2 molecules present gluten peptides to T cells that drive the inflammatory response and provide help for B cells making anti-gliadin antibodies. Interestingly, in a study of immune reactivity in serum, it was found that nearly 40% of sarcoidosis patients had antibodies to ATP synthase (24%) or to gliadin (15%) (29). Similarly, another study found that approximately 40% of sarcoidosis patients were sensitized to gluten (30). One explanation behind a shared immune reactivity between sarcoidosis and celiac disease could be shared HLA alleles, since DQ2 and DRB1*0301 are strongly linked (31). One proposed pathogenic mechanism of celiac disease is that inflammation of the intestinal mucosa may result in unmasking of cryptic antigens normally not exposed to the immune system. These antigens may be identical to or share epitopes with antigens present in other organs. Similarly, we propose that the pulmonary inflammation in sarcoidosis also may lead to epitope spreading to additional antigens.

A future task is to explore the antigenic potential of the identified peptides. It will be of particular interest to investigate whether any of these peptides will stimulate the CD4⁺AV2S3⁺ cells that accumulate in the lungs of DRB1*0301^{pos} sarcoidosis patients. Detailed immunological analyses of these AV2S3⁺ T cells regarding their TCR- α chain nucleotide sequences (32) and their expression of activation markers (33) together indicate that AV2S3⁺ T cells have recognized and proliferated in response to a sarcoidosis-associated antigen.

In conclusion, we have used a generic approach in a distinct clinical setting in order to identify peptides presented by HLA-DR molecules at the site of inflammation, the lung. DRB1*0301^{pos} sarcoidosis patients, in which the immune response strongly indicates the presence of a specific antigen, provided a unique opportunity to identify such a disease-specific antigen(s). Several peptides derived from known autoantigens were characterized, which are currently under investigation to decipher their role in sarcoidosis. The used approach allows the identification of *in vivo*-presented HLA-bound peptides in diverse disease settings provided that patient specimens are available. Thus, it is now possible to study the molecular machinery responsible *in vivo* for autoreactive T cells. In the future, such T cells may become targets for new, antigen-specific therapeutic approaches.

Methods

Patient characterization. All patients included in this study were consecutive patients referred to the Division of Respiratory Medicine (Karolinska University Hospital, Stockholm, Sweden) for investigation. All patients had a clinical picture in accordance with pulmonary sarcoidosis, as determined by symptoms (such as fatigue, dyspnea on exertion, and dry coughing), chest radiography, and pulmonary function tests, and the diagnosis was established using the criteria of the World Association of Sarcoidosis and other Granulomatous Disorders (WASOG) (1). Thus, patients diagnosed with sarcoidosis had either a positive biopsy (showing noncaseating granuloma), a BAL CD4/CD8 ratio of more than 3.5, or Löfgren syndrome (defined as bilateral hilar lymphadenopathy, erythema nodosum, and/or ankle arthritis; ref. 34). Written informed consent was obtained from

all subjects, and the study was approved by the Regional Ethical Review Board in Stockholm.

Peptides were derived from BAL cells of 16 HLA-DRB1*0301^{pos} sarcoidosis patients. Eight of the patients presented with Löfgren syndrome. One patient had a typical clinical picture, chest x-ray findings, and pulmonary function test results but neither a positive biopsy, Löfgren syndrome, nor an elevated BAL CD4/CD8 ratio. Since other differential diagnoses were excluded, this patient was included in the study. There were 9 male and 7 female patients. Their median age was 31 years, with a range of 25–73, 8 were nonsmokers (including 2 ex-smokers who quit smoking more than 1 year before the study), and 8 were smokers. Nine patients presented with chest radiographic stage I, 3 had stage II, and 4 had stage III. Only 1 of the patients was treated with oral steroids at the time of bronchoscopy and BAL. BAL cell concentration was 235.8×10^6 cells/l (75.6×10^6 to $1,057 \times 10^6$ cells/l) and differential count showed 82.7% (48.0%–97.4%) alveolar macrophages and 14.4% (0.9%–49.8%) lymphocytes. Nonseparated BAL cells were stored at -80°C (see *Handling of BAL fluid*).

Bronchoscopy and BAL. This was performed as described previously (35). After premedication with morphine-hyoscine (Morfin-skopolamin; Meda) intramuscularly and topically applied lidocaine (Xylocaine; AstraZeneca), bronchoscopy was performed with a flexible fiberoptic bronchoscope. BAL was performed by wedging the bronchoscope in one of the subsegments of the middle lobe. For BAL, five 50-ml aliquots of warmed PBS were instilled and aspirated. The fluid was collected in a silicone-treated bottle kept on ice, which was immediately transported to the laboratory.

Handling of BAL fluid. BAL fluid was strained through a Dacron net (Millipore), and the volume of recovered fluid was measured. Cells were pelleted by centrifugation at 400 g, 4°C , for 10 minutes, and the supernatants were poured off. The cell pellets were resuspended in RPMI medium (Sigma-Aldrich). Cells were counted in a Bürker chamber, and total cell viability was determined by Trypan blue exclusion. Cells to be used for peptide elution were pelleted and directly frozen at -80°C . Smears for differential counts were prepared by cytocentrifugation at 50 g for 3 minutes (Cytospin 2 Shandon; Southern Products Ltd.), whereupon cells were stained with May-Grünwald-Giemsa, and 500 cells were counted.

Immunostaining and flow cytometry. Phenotypic characterization of BAL lymphocytes was performed using monoclonal R-phycoerythrin-Cy5-conjugated CD4 antibodies, R-phycoerythrin-conjugated CD8 antibodies, and FITC-conjugated CD3 antibodies obtained from Dako. Staining of cells was performed as previously described (36). After antibody incubations and washings, cells were fixed in PBS with 1% formaldehyde. The samples were analyzed in a flow cytometer (FACSCalibur; BD). Lymphocytes were identified and gated by forward and side light scattering properties. In the CD3⁺ lymphocyte population, the percentages of positively labeled cells in the CD4⁺ and CD8⁺ subsets were determined, whereafter the BAL CD4/CD8 ratio was calculated. Isotype-matched negative control antibodies always stained less than 1% of lymphocytes.

DNA extraction and genotyping of HLA. DNA was extracted from whole blood using the “salting-out” method (37). Olerup SSP PCR (Sequence Specific Primers for PCR) DR low resolution typing kits were used for analysis of HLA-DR types.

Elution and analysis of HLA-DR-bound peptides presented by BAL cells. HLA-DR-associated peptides were isolated from frozen total BAL cell pellets according to standard protocols (38) using the HLA-DR-specific antibody L243 (39). In brief, BAL cells were thawed at 4°C in PBS/1 mM 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) in the presence of protease inhibitors (Complete Protease Inhibitor Cocktail Tablet; Roche Molecular Biochemicals) and homogenized in a Dounce homogenizer. Subsequently, HLA-DR molecules and bound peptides were isolated with the solid-phase bound monoclonal antibody by immuno-



affinity chromatography. Precipitates were treated with 0.1% trifluoroacetic acid, ultrafiltrated by a 10-kDa ultrafilter, and separated by gradient elution from an RP column (HPLC, SMART system, μRP SC C2/C18 column, 100 × 2.1 mm; Amersham Pharmacia Biotech). A binary gradient of 10%–60% B within 100 minutes was performed, applying a flow rate of 150 μl/min. Solvent A was 0.1% (vol/vol) trifluoroacetic acid water, and solvent B contained 0.08% trifluoroacetic acid in 80% (vol/vol) acetonitrile water. The peptides were fractionated in 150 μl aliquots. Subsequently, every other peptide-containing HPLC fraction was pooled and analyzed by nanocapillary RP-HPLC (UltimateHPLC system; Dionex) directly coupled to the ESI source of a hybrid Q-TOF tandem mass spectrometer (Micromass; Waters). Samples were loaded onto a C 18 precolumn for concentration and desalting. After loading, the precolumn was placed in line for separation by a fused-silica microcapillary column (75 μm i.d. × 250 mm) packed with 5 μm C18 RP material (Dionex). Solvent A was 4 mM ammonium acetate/water. Solvent B was 2 mM ammonium acetate in 80% acetonitrile/water. Both solvents were adjusted to pH 3.0 with formic acid. Samples were run through a binary gradient of 15%–60% B within 120 minutes, applying a flow rate of 200 μl/min reduced to 300 nL/min by the Ultimate split-system. A gold-coated glass capillary (PicoTip; New Objective) was used for introduction into the micro-ESI source. In MS/MS experiments, sequence information was obtained by manual interpretation of fragment spectra using computer-assisted database (NCBIInr, nonredundant protein database) searching tools (MASCOT; Matrix Science) (40). MASCOT searches were performed using *Homo sapiens* as taxonomy, no enzymatic specificity, ±0.3 Da for MS tolerance, and ±0.2 Da for MS/MS tolerance. MS/MS spectra from synthetic peptides were recorded and compared with spectra from the corresponding eluted natural peptides to prove their identity.

FASTA database searches. Peptide sequences from mass-spectrometric analyses were searched using FASTA (41), with generous expect value parameters

in order to allow for the detection of 100% identical peptides of length down to 6 residues. Due to unambiguous assignments of Ile/Leu and Lys/Gln, both residue alternatives at applicable positions were investigated in the FASTA searches. The databases used were UniProtKB (Swiss-Prot + TrEMBL), version 8.3 (42) and an in-house database of all genomes in the public domain, GenomeLPG (Bresell and Hedlund, unpublished observations).

Acknowledgments

This work was supported by the Foundation for Sarcoidosis Research/American Thoracic Society (S-05-001), the Torsten and Ragnar Söderberg Foundation, the NIH (1 R21 HL077579-01), the Swedish Research Council (2006-5423), the Swedish Heart-Lung Foundation, the King Oscar II Jubilee Foundation, the Carl Trygger Foundation, and the Stockholm County Council and Karolinska Institutet as well as the German Research Foundation (DFG We 1947/5-1). J. Dengjel is supported by the European Molecular Biology Organization (long-term fellowship). R. Weissert was holding a Heisenberg fellowship (DFG We 1947/4-2).

Received for publication April 16, 2007, and accepted in revised form August 15, 2007.

Address correspondence to: Jan Wahlström, Lung Research Laboratory, L4:01, Karolinska University Hospital, Solna, 171 76 Stockholm, Sweden. Phone: 46-8-5177-0663; Fax: 46-8-5177-5451; E-mail: jan.wahlstrom@ki.se.

Jörn Dengjel's present address is: Center of Experimental Bioinformatics, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark.

1. [No authors listed]. 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. Respir. Crit. Care Med.* **160**:736–755.
2. Newman, L.S., Rose, C.S., and Maier, L.A. 1997. Sarcoidosis. *N. Engl. J. Med.* **336**:1224–1234.
3. Grunewald, J., et al. 1994. T-cell receptor variable region gene usage by CD4+ and CD8+ T cells in bronchoalveolar lavage fluid and peripheral blood of sarcoidosis patients. *Proc. Natl. Acad. Sci. U. S. A.* **91**:4965–4969.
4. Planck, A., Eklund, A., and Grunewald, J. 2003. Markers of activity in clinically recovered human leukocyte antigen-DR17-positive sarcoidosis patients. *Eur. Respir. J.* **21**:52–57.
5. Grunewald, J., Berlin, M., Olerup, O., and Eklund, A. 2000. Lung T-helper cells expressing T-cell receptor AV2S3 associate with clinical features of pulmonary sarcoidosis. *Am. J. Respir. Crit. Care Med.* **161**:814–818.
6. Grunewald, J., et al. 2002. Lung restricted T cell receptor AV2S3+ CD4+ T cell expansions in sarcoidosis patients with a shared HLA-DRB β chain conformation. *Thorax* **57**:348–352.
7. Texier, C., et al. 2001. Complementarity and redundancy of the binding specificity of HLA-DRB1, -DRB3, -DRB4 and -DRB5 molecules. *Eur. J. Immunol.* **31**:1837–1846.
8. Dengjel, J., et al. 2006. Unexpected abundance of HLA class II presented peptides in primary renal cell carcinomas. *Clin. Cancer Res.* **12**:4163–4170.
9. Kruger, T., et al. 2005. Lessons to be learned from primary renal cell carcinomas: novel tumor antigens and HLA ligands for immunotherapy. *Cancer* **103**:1826–1835.
10. Immunol. Immunother. **54**:826–836.
11. Dengjel, J., Rammensee, H.G., and Stevanovic, S. 2005. Glycan side chains on naturally presented MHC class II ligands. *J. Mass Spectrom.* **40**:100–104.
12. Steen, H., and Mann, M. 2004. The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* **5**:699–711.
13. Hillen, N., and Stevanovic, S. 2006. Contribution of mass spectrometry-based proteomics to immunology. *Expert Rev Proteomics* **3**:653–664.
14. Planz, C., et al. 2001. A naturally processed rat major histocompatibility complex class I-associated viral peptide as target structure of borna disease virus-specific CD8+ T cells. *J. Biol. Chem.* **276**:13689–13694.
15. Senecal, J.L., Oliver, J.M., and Rothfield, N. 1985. Anticytoskeletal autoantibodies in the connective tissue diseases. *Arthritis Rheum.* **28**:889–898.
16. Kurki, P., Helve, T., and Virtanen, I. 1983. Antibodies to cytoplasmic intermediate filaments in rheumatic diseases. *J. Rheumatol.* **10**:558–562.
17. Alcover, A., et al. 1984. Antibodies to vimentin intermediate filaments in sera from patients with systemic lupus erythematosus. *Arthritis Rheum.* **27**:922–928.
18. Boulassel, M.R., Tomasi, J.P., Deggouj, N., and Gersdorff, M. 2000. Identification of beta-actin as a candidate autoantigen in autoimmune inner ear disease. *Clin. Otolaryngol. Allied Sci.* **25**:535–541.
19. Schnabel, A., Hellmich, B., and Gross, W.L. 2005. Interstitial lung disease in polymyositis and dermatomyositis. *Curr. Rheumatol. Rep.* **7**:99–105.
20. Love, L.A., et al. 1991. A new approach to the classification of idiopathic inflammatory myopathy: myositis-specific autoantibodies define useful homogeneous patient groups. *Medicine (Baltimore)* **70**:360–374.
21. Stet, Y.N., and Komatireddy, G.R. 2001. Antiphospholipid antibodies in systemic lupus erythematosus and the antiphospholipid syndrome. *Front. Biosci.* **6**:E207–E212.
22. Lim, P.L., and Zouali, M. 2006. Pathogenic autoantibodies: emerging insights into tissue injury. *Immunol. Lett.* **103**:17–26.
23. Yang, Y., et al. 2002. Detection of antivimentin antibody in sera of patients with idiopathic pulmonary fibrosis and non-specific interstitial pneumonia. *Clin. Exp. Immunol.* **128**:169–174.
24. Uchida, K., Kondo, H., and Kashiwazaki, S. 1987. The clinical significance of antibodies to vimentin intermediate filaments in patients with rheumatoid arthritis [In Japanese]. *Ryumachi* **27**:356–363.
25. Lazarides, E. 1980. Intermediate filaments as mechanical integrators of cellular space. *Nature* **283**:249–256.
26. Carvalho, Y.R., Loyola, A.M., Gomez, R.S., and Araujo, V.C. 1995. Peripheral giant cell granuloma. An immunohistochemical and ultrastructural study. *Oral Dis.* **1**:20–25.
27. Cain, H., and Kraus, B. 1983. Immunofluorescence microscopic demonstration of vimentin filaments in asteroid bodies of sarcoidosis. A comparison with electron microscopic findings. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **42**:213–226.
28. James, S.P. 2005. Prototypic disorders of gastrointestinal mucosal immune function: celiac disease and Crohn's disease. *J. Allergy Clin. Immunol.* **115**:25–30.
29. Papadopoulos, K.I., Sjoberg, K., Lindgren, S., and Hallengren, B. 1999. Evidence of gastrointestinal immune reactivity in patients with sarcoidosis. *J. Intern. Med.* **245**:525–531.
30. McCormick, P.A., et al. 1988. Altered gastroin-



testinal immune response in sarcoidosis. *Gut*. **29**:1628–1631.

31. Candore, G., Lio, D., Colonna Romano, G., and Caruso, C. 2002. Pathogenesis of autoimmune diseases associated with 8.1 ancestral haplotype: effect of multiple gene interactions. *Autoimmun. Rev.* **1**:29–35.

32. Grunewald, J., Hultman, T., Bucht, A., Eklund, A., and Wigzell, H. 1995. Restricted usage of T cell receptor V α 1/J α 1 gene segments with different nucleotide but identical amino acid sequences in HLA-DR3+ sarcoidosis patients. *Mol. Med.* **1**:287–296.

33. Katchar, K., Wahlstrom, J., Eklund, A., and Grunewald, J. 2001. Highly activated T-cell receptor AV2S3(+) CD4(+) lung T-cell expansions in pulmonary sarcoidosis. *Am. J. Respir. Crit. Care Med.* **163**:1540–1545.

34. Grunewald, J., and Eklund, A. 2007. Sex-Specific Manifestations of Lofgren's Syndrome. *Am. J. Respir. Crit. Care Med.* **175**:40–44.

35. Eklund, A., and Blaschke, E. 1986. Relationship between changed alveolar-capillary permeability and angiotensin converting enzyme activity in serum in sarcoidosis. *Thorax*. **41**:629–634.

36. Wahlstrom, J., Dahlen, B., Ihre, E., Wigzell, H., Grunewald, J., and Eklund, A. 1998. Selective CD8+ T cells accumulate in the lungs of patients with allergic asthma after allergen bronchoprovocation. *Clin. Exp. Immunol.* **112**:1–9.

37. Zetterquist, H., et al. 1997. Report from the HLA class II typing by PCR-SSP Multicentre Study. *Eur. J. Immunogenet.* **24**:191–199.

38. Duyar, H., et al. 2005. Peptide motif for the rat MHC class II molecule RT1.Da: similarities to the multiple sclerosis-associated HLA-DRB1*1501 molecule. *Immunogenetics*. **57**:69–76.

39. Lampson, L.A., and Levy, R. 1980. Two populations of Ia-like molecules on a human B cell line. *J. Immunol.* **125**:293–299.

40. Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. **20**:3551–3567.

41. Pearson, W.R., and Lipman, D.J. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U. S. A.* **85**:2444–2448.

42. Bairoch, A., et al. 2005. The Universal Protein Resource (UniProt). *Nucleic Acids Res.* **33**:D154–D159.