Published in final edited form as:

J Immunol Methods. 2012 November 30; 385(1-2): 35–44. doi:10.1016/j.jim.2012.07.024.

Identification of Novel Autoantigens by a Triangulation Approach

Tricia R. Cottrell, John C. Hall, Antony Rosen, and Livia Casciola-Rosen¹

Division of Rheumatology, School of Medicine, Johns Hopkins University, 5200 Eastern Avenue, MFL-CT Suite 5300, Baltimore, MD 21224

Tricia R. Cottrell: tcottre2@jhmi.edu; John C. Hall: jhall4@jhmi.edu; Antony Rosen: arosen@jhmi.edu; Livia Casciola-Rosen: lcr@jhmi.edu

Abstract

High titer autoantibodies, which are often associated with specific clinical phenotypes, are useful diagnostically and prognostically in systemic autoimmune diseases. In several autoimmune rheumatic diseases (e.g. myositis and Sjogren's syndrome), 20–40% of patients are autoantibody negative as assessed by conventional assays. The recent discovery of new specificities (e.g., anti-MDA5) in a subset of these autoantibody-negative subjects demonstrates that additional specificities await identification. In this manuscript, we describe a rapid multidimensional method to identify new autoantigens. A central foundation of this rapid approach is the use of an antigen source in which a pathogenic pathway active in the disease is recapitulated. Additionally, the method involves a modified serological proteome analysis strategy which allows confirmation that the correct gel plug has been removed prior to sending for sequencing. Lastly, the approach uses multiple sources of information to enable rapid triangulation and identification of protein candidates. Possible permutations and underlying principles of this triangulation strategy are elaborated to demonstrate the broad utility of this approach for antigen discovery.

1. Introduction

High titer autoantibodies are frequently observed in systemic autoimmunity, where they are often associated with specific phenotypic manifestations (Stinton and Fritzler, 2007, Bizzaro, 2007). These antibody specificities are therefore useful for diagnosis and prognosis, and provide insights into disease mechanisms. Paradoxically, the vast majority of antigen targets identified to date in systemic autoimmunity are ubiquitously expressed, a finding which is difficult to reconcile with the highly specific patterns of tissue damage observed in these diseases. For example, antibodies against aminoacyl tRNA synthetases are found only in patients with myositis (in which muscle and lung are the primary target organs), and yet these proteins perform an essential function in all cells. Similarly, topoisomerase I and centromere proteins are necessarily present in all cell types and are well defined autoantibody targets in systemic sclerosis, a disease in which the skin, lungs, and gastrointestinal tract are most commonly affected. In addition, many patients with systemic autoimmune diseases lack known autoantibody specificities (e.g., in myositis, ~20–40% of

© 2012 Elsevier B.V. All rights reserved.

¹Corresponding Author: Livia Casciola-Rosen, PhD, Division of Rheumatology, Johns Hopkins University School of Medicine, 5200 Eastern Avenue, Mason Lord Building, Center Tower, Suite 5200, Baltimore, MD 21224, Ph: (410) 550-1890; Fax: 410-550-2072.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

patients are autoantibody negative as assessed using currently available clinical assays (Mammen, 2011)).

In several autoimmune diseases, autoantibodies recognize post-translationally modified proteins, such as citrullinated proteins in rheumatoid arthritis (Darrah and Rosen, 2010, Eggleton et al., 2008) and proteins susceptible to granzyme B cleavage (Casciola-Rosen et al., 1999). It has been proposed that perturbations which result in the modification or enhanced expression of particular proteins may drive the immune response in some autoimmune diseases.

In the inflammatory myopathies, an increasing body of data supports a role for disease-specific changes in antigen expression within the target tissue in driving the disease-specific immune response. For example, expression of Mi-2 is increased specifically in dermatomyositis (DM) muscle, the disease in which Mi-2 autoantibodies are found (Casciola-Rosen et al., 2005). Additionally, Mi-2 is expressed at high levels in regenerating muscle cells, but only at low levels in differentiated muscle cells (Mammen et al., 2009). Similarly, 3-hydroxy-3-methylglutaryl-coenzyme A reductase expression is dramatically upregulated by statins both in cultured cells, as well as in muscle from patients with statin-induced necrotizing myopathy (of note, these patients make high titer autoantibodies to this enzyme) (Mammen et al., 2011). Antibodies targeting melanoma differentiation-associated protein 5 (MDA5) have also recently been identified in myositis patients previously considered antibody negative (Sato et al., 2009, Fiorentino et al., 2011). Interestingly, expression of this protein is highly induced by interferon- α (IFN α), a finding consistent with the strong type I IFN signature that has been well documented in the muscle of patients with DM (Greenberg et al., 2011).

These data suggest that screening relevant cells exposed to an appropriate perturbation *in vitro*, which mimics changes identified in the target tissue in that disease, may enable the discovery of novel autoantibodies. We have therefore focused on autoantibody discovery using cellular sources in which pathways known to be active in the target tissue *in situ* are recapitulated; examples described here are exposure to type I IFN, and muscle regeneration. While several methods of antigen identification exist, including screening of recombinant cDNA expression libraries (SEREX), peptide libraries, and multiple affinity protein profiling, few accommodate the complexity of protein dynamics observed *in vivo*, such as splice variants, post-translational modifications, cleavage, nucleic acid binding, inducible expression, and cryptic epitope exposure.

Serological proteome analysis (SERPA), which combines two-dimensional (2D) electrophoresis with immunoblotting and tandem mass spectrometry (Canelle et al., 2005), is currently one of the most useful techniques to define modified autoantigens. A major advantage of this method is the high sensitivity of both immunoblotting and tandem mass spectrometry, allowing the detection of relatively low abundance antigens within a complex protein mixture (Mallick and Kuster, 2010). However, this approach is significantly limited by the fact that selection of a low abundance protein spot on a 2D gel by alignment with a patient serum immunoblot is often insufficiently precise. Since numerous candidate proteins are generated by MS/MS sequencing (Duncan et al., 2010), it is critical that candidate antigens are rigorously confirmed to eliminate false positives. This often requires a substantial investment of time and resources before the correct protein is confirmed.

In this manuscript, we describe a rapid multidimensional method of autoantigen discovery using patient sera as probes. Using a combination of (i) cells subjected to a perturbation known to be present in the diseased tissue; (ii) a modified SERPA strategy which confirms that the relevant spot has been plucked from the gel before sequencing; and (iii) mRNA

expression profile data which defines molecules whose expression is enhanced by the perturbation, enables rapid autoantigen identification. Data is shown which details the identification process for 2 novel autoantigens whose expression is regulated in the context of 2 different pathways of pathogenic relevance.

2. Materials and methods

2.1. Sera

Frozen aliquots of de-identified sera from 141 myositis patients, obtained during routine blood draws at the Johns Hopkins Myositis Center, were used in the preliminary screens for this study. Patients were clinically defined according to the Bohan and Peter criteria (Bohan and Peter, 1975a, b). Normal sera from 50 healthy donors were also used. All samples were obtained, and studies were performed, in compliance with Institutional Review Board and Health Insurance Portability and Accountability Act regulations.

2.2. Cell culture

Human salivary gland (HSG) cells were a generous gift from Dr. Bruce Baum and were cultured in minimal essential medium (MEM; Invitrogen) with 10% heat-inactivated fetal calf serum (HI FCS), 1% L-glutamine, and 1% penicillin/streptomycin at 37°C with 5% CO₂. Human cervical epithelial cells (HeLa) were cultured in DMEM low glucose (Invitrogen) with 10% HI FCS, 1% L-glutamine, and 1% penicillin/streptomycin at 37°C with 7.5% CO₂. Second passage primary human keratinocytes (Lonza) were grown in Keratinocyte Basal Medium-2 (KBM-2) supplemented with the KGM-2 bullet kit (Lonza) without subculturing. For IFN treatment, cells were grown to 70% confluence before being cultured for an additional 24 hours in the absence or presence of 1000 units/mL IFN- α (Sigma). Normal human skeletal muscle cells were cultured and induced to differentiate into myotubes over a 2 week period as described (Casciola-Rosen et al., 2005).

2.3. Lysate preparation from cultured cells and human muscle biopsies

Culture dishes were kept on ice during two washes with cold PBS followed by scraping into lysis buffer (1% Nonidet P-40 (NP40); 20 mM Tris, pH 7.4; 150 mM NaCl; 1mM EDTA) containing a protease inhibitor cocktail. Human muscle biopsies were obtained from patients seen at the Johns Hopkins Myositis Center. Surgical procedures were performed for patient management, and the research tissue samples were in excess of that required for routine diagnostic purposes. Collection and use of the human biopsy specimens were approved by the Johns Hopkins Institutional Review Board. All patient samples were deidentified, with clinical and laboratory features linked only to the patient code. Detergent extracts were prepared in lysis buffer from frozen muscle biopsies obtained from DM and polymyositis (PM) patients, and from a control individual, as described (Casciola-Rosen et al., 2005). Patients were defined as having DM or PM if they had probable or definite disease according to the Bohan and Peter criteria (Bohan and Peter, 1975a, b). All lysates were diluted to working concentrations in gel application buffer (2% SDS; 4% glycerol; 40 mM Tris-HCl, pH 6.8; 0.002% bromophenol blue; 1% β -mercaptoethanol) and boiled for 3 min.

2.4. Immunoblotting

Primary immunoblotting screens using lysate pairs from HSG and HeLa cells incubated in the absence or presence of IFN- α or muscle cells at various stages of differentiation, were performed as follows. 5 μ g/lane of each lysate was electrophoresed on 7cm, 10% SDS-PAGE gels for 170Vh, then transferred to nitrocellulose membranes at 100V for 1 hour. Membranes were stained with ponceau (0.3% w/v in 3% trichloroacetic acid) and washed with distilled H₂O to identify molecular weight markers, followed by blocking for 1 hour at room temperature (RT) in 3% bovine serum albumin in wash buffer (WB) (10mM Tris-HCl,

pH 7.6; 0.5% NP40; 150mM NaCl). Membranes were then incubated with patient serum diluted in WB (1:5000) for 1hr at RT. After extensive washing with WB (5 × 5 mins/wash), membranes were subsequently incubated with a horseradish peroxidase-conjugated goat anti-human IgG secondary antibody (Jackson ImmunoResearch) at 1:10,000 dilution in WB for 1 hour. Membranes were then washed as above and incubated with enhanced chemiluminescent reagent (Pierce) followed by exposure to X-ray film. For serum M4, a second screen including lysates from HSGs, HeLas, and primary human keratinocytes treated with 1000 units/mL IFN- α was performed as described above.

2.5. 2-Dimensional (2D) electrophoresis

IFN- α treated cells (section 2.2) and muscle cells at day 4 of differentiation were washed twice with cold PBS followed by two washes with cold 200 mM Sucrose, 10mM Tris HCl, pH7.4 and lysed directly into isoelectric focusing (IEF) buffer (8 M urea, 2 M thiourea, 4% CHAPS, 1% dithiothreitol and protease inhibitors). First dimension focusing was performed using Bio-Rad ReadyStrip IPG Strips and Protean IEF Cell per manufacturer's instructions. 7cm IEF strips were used to determine the optimum pH range for the protein of interest while 17cm IEF strips were used for sequencing. Briefly, each 7cm IEF strip was loaded with 130 μ l consisting of 100 μ g lysate, 1.5% 2-hydroxyethyl disulfide, 0.1% bromophenol blue and 0.4% ampholytes (pH dependent on IEF strip pH range). For 17cm strips, 300 μ g lysate was used in a total volume of 300 μ l. Up to 12 IEF strips can be run simultaneously; these can subsequently be processed immediately for the second dimension gel (see below), or they can be stored at -80°C. Prior to running the second dimension gels, IEF strips were incubated for 20 min at RT in 10 mg/ml dithiothreitol, followed by 20 min in 25 mg/ml iodoacetamide, both in equilibration buffer (6M urea; 0.375M Tris-HCl, pH 8.8; 2% SDS; 20% glycerol). IEF strips were then loaded onto 10% SDS PAGE gels (16 cm or 7 cm gels, depending on the size of the IPG strips), covered with 1% low melt agarose (Promega) and electrophoresed as described above (16 cm gels were electrophoresed for 850 volt hours, and transferred onto nitrocellulose membranes for 4 hr at 400 mA). To ensure consistency, IEF strips run in parallel were used for 2D immunoblotting (section 2.6) and corresponding protein sequencing (section 2.7) experiments.

2.6. 2D Immunoblotting and antigen localization

2D gels (section 2.5) were transferred to nitrocellulose membranes, stained with ponceau to visualize resolved protein spots, and scanned at high resolution for alignment with the subsequent immunoblots. Immunoblots were performed as described (section 2.4). During X-ray film exposure, precise documentation of the membrane perimeter on the film is critical for subsequent alignment of the exposed film with a printout of the ponceau scan. This enables the immunoblotted spot of interest to be aligned with the corresponding ponceau-stained protein spot. Once the protein of interest is located on the ponceau scan, nearby spot patterns can be used for orientation to find the same spot on the preparative gel for plucking and protein identification (section 2.7).

2.7. Protein identification

Preparative SDS-PAGE second dimension gels were made using solutions and procedures that minimize exposure to airborne contaminants. All glassware and equipment were only used for proteomics experiments and were washed twice with 0.1% SDS (w/v) followed by extensive rinsing with ultrapure distilled water and 100% methanol to minimize protein contamination. After performing 1st and 2nd dimension gel runs as described (section 2.5), gels were fixed in destain (14% v/v acetic acid; 24% v/v methanol) for 15 min, incubated for 1 hour in Gel Code Blue Stain (Thermo Scientific) and then destained in ultrapure distilled water for 2–4 hours. Destained gels were placed on a light box to facilitate visualization of stained protein spots, and gel plugs containing the protein spot of interest and a nearby

“blank” gel plug containing no visible protein were carefully removed from the preparative gel using a sterile pipette tip. The gel plugs were immediately washed twice in 50% methanol for 10 min each with shaking and stored at -20°C . The remaining preparative gel was washed 3×5 min in running buffer (25mM Tris-HCl; 200mM glycine; 0.1% (w/v) SDS), and proteins were transferred to nitrocellulose membrane for immunoblotting (section 2.6) to verify successful plucking. Gel plugs confirmed to contain the protein of interest were sent to the Johns Hopkins Mass Spectrometry and Proteomics Facility for liquid chromatography tandem mass spectrometry (LC/MS/MS) peptide sequencing.

2.8. Microarray

Total RNA was extracted using the Trizol reagent method (Invitrogen) from HSGs cultured for 24 hours in the absence or presence of IFN- α at [1000 units/ml], and from muscle cells at various stages of differentiation (days 0–5). All samples were collected in triplicate. Additional purification was performed on RNeasy columns (Qiagen), and the quality of total RNA samples was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies). Biotin-labeled, complementary RNA (cRNA) was prepared from total RNA according to the chip manufacturer’s protocol (Illumina). cRNA was hybridized to Illumina HumanHT12 v3 Expression BeadChips, and signal was detected with streptavidin-Cy3 (IFN treatment). cRNA was hybridized to Illumina Sentrix HumanRef-8 Expression BeadChips, and signal was detected with streptavidin-Cy3 (muscle cell differentiation). All signal intensity quantification was performed using an Illumina BeadStation 500GX Genetic Analysis Systems scanner.

2.9. Analytical methods and statistical analysis for microarray data

IFN treatment - A single intensity (expression) value for each Illumina probe on the array was obtained using Illumina BeadStudio software with standard settings and no background correction. The expression values for all the probes for each sample were scaled to have median 256 (2^8) and then log (base 2) transformed. Genes (i.e., Illumina probes) considered to be differentially expressed between two groups of samples were those satisfying the following criteria: (i) Welch t-test p-values less than or equal to 0.01 (10^{-2}) (Pan, 2002); (ii) a Benjamini-Hochberg false discovery rate (FDR) less than or equal to 0.25 (Benjamini and Hochberg, 1995); (iii) a fold change above 2.0 (calculated using geometric means); and (iv) the expression value of the gene is above the Illumina BeadStudio calculated background for all 3 samples in the group.

Muscle cell differentiation - Preliminary analysis of the scanned data was performed using Illumina BeadStudio software. Any gene consistently below a background threshold level of $D=0.98$ for all samples was eliminated from further analysis. Bead averaged data were normalized by Z transformation. Briefly, microarray data were $\log_{10}$ transformed and data were standardized to the mean expression of the dataset and represented in terms of a Z score.

2.10. Antigen identity confirmation

Embryonic myosin light chain 4 (MYL4) cDNA was cloned from myoblasts at day 4 of differentiation, and was sequence verified. IFIT3 cDNA (Origene) and MYL4 cDNA were transiently transfected into HeLa cells using Lipofectamine (Invitrogen), per the manufacturer’s protocol.

Mouse monoclonal anti-IFIT3 and anti-MYL4 antibodies were obtained from BD Transduction Labs and Novus, respectively. Cell lysates were transferred onto nitrocellulose membranes (section 2.4) and blocked in Tris buffered saline/Tween 20 (TBST; 10mM Tris-HCl, pH 7.4; 150mM NaCl; 0.05% (v/v) Tween 20) with 5% (w/v) powdered milk for 1 hr

at RT. Membranes were then incubated with anti-IFIT3 or anti-MYL4 antibodies (diluted 1:3,000 or 1:1,000, respectively, in TBST/5% milk) overnight at 4°C. Membranes were washed for 5 × 5 min with TBST/5% milk at RT, followed by incubation with a 1:10,000 dilution of horseradish peroxidase conjugated anti-mouse secondary antibody in TBST/5% milk for 1 hr at RT. After 5 × 5 min washes using TBST buffer, membranes were incubated for 5 min with enhanced chemiluminescent reagent (Pierce) and then exposed to X-ray film.

³⁵S-methionine-labeled IFIT3 protein was synthesized using the TNT T7 Coupled Reticulocyte Lysate System (Promega) for *in vitro* transcription and translation (IVTT) using IFIT3 cDNA per the manufacturer's instructions. Immunoprecipitations using ³⁵S-methionine-labeled IFIT3 protein were performed as described, with all steps carried out at 4°C (Fiorentino et al., 2011). Briefly, 1 µl of IVTT product was diluted in lysis buffer with protease inhibitors, 1 µl of patient serum was added and the mixture was rotated for 1 hr. 35 µl of immobilized Protein A agarose (Pierce) was then added, and the mixture was rocked for 20 mins. Immunoprecipitates were subsequently washed 4 times with lysis buffer, electrophoresed on 10% SDS-PAGE gels and visualized by fluorography.

3. Results and Discussion

3.1. Detection of IFN-α induced autoantigens, and selection of a novel specificity

To screen for new autoantibodies targeting IFN-α-induced proteins, lysates made from untreated and IFN-α treated cells were immunoblotted with myositis patient sera. A representative screening blot performed on untreated and IFN-α treated HSGs is shown (Fig. 1A). Several different patterns were observed: some sera immunoblotted proteins present exclusively in IFN-α treated cell lysates (e.g., the 60 kDa protein blotted by M4); others blotted proteins present at baseline, but upregulated by IFN-α (e.g., the 52 kDa protein blotted by M1 and M3); some blotted proteins whose expression remained unchanged (e.g., M2 and M5) and some did not blot anything (e.g., M6). We focused on sera that detected novel specificities induced by IFN-α treatment; hence, of the DM patient sera shown in Fig 1A, only M4 was of interest (the 52 kDa band blotted by M1 and M3 is Ro52 (TRIM21), an autoantigen which is known to be targeted in patients with myositis). Of note, we found antibodies recognizing IFN-α induced proteins in 16/141 (11%) of the myositis sera; IFN-inducible specificities were never detected with healthy control sera (data not shown).

A subsequent screen was performed to determine if the 60 kDa antigen blotted by serum M4 could be detected in other cell types treated with IFN-α. As shown in Fig 1B, the 60 kDa protein was also detected in HeLa cells and primary human keratinocytes upon treatment with IFN-α, suggesting that local type I IFN signaling may induce autoantigen expression by a variety of cell types. Moreover, this enabled the use of data from multiple IFN-treated cell types for antigen identification.

3.2. Antigen identification strategy

We next pursued the identification of the 60 kDa protein using 2D gel electrophoresis and LC/MS/MS peptide sequencing. A critical aspect of this methodology is precise localization of the protein of interest by immunoblotting a replicate 2D gel (Fig. 2A). It is important to scan the ponceau stained membrane prior to immunoblotting because careful alignment of the immunoblot with a printout of the scanned ponceau image allows identification of the antigenic protein spot to be plucked. The surrounding pattern of spots visible only on the ponceau scan is critical for locating the spot of interest on an identically run preparative 2D gel stained with colloidal blue (Fig. 2B). Of note, spot identification can be a challenging step if the antigen of interest is relatively low abundance (hence faintly stained, as highlighted for the M4 antigen in Fig. 2B), which may result in inaccurate plucking. Once

removed, the gel plug is sent for peptide sequencing by LC/MS/MS and a list of candidate proteins is generated by BLAST searching all of the peptide sequences generated against a mammalian protein sequence database (MASCOT). Although only a tiny gel plug containing the spot of interest is used for sequencing, the high sensitivity of the LC/MS/MS frequently results in lists of ~10–20 candidate proteins from the detected peptides, even after subtracting contaminants sequenced from a nearby “blank” gel plug. Pursuing candidate lists obtained from inaccurate spot selections therefore wastes significant time and resources to eliminate spurious candidates.

We therefore developed a rapid, inexpensive quality control step to confirm that the correct protein spot is plucked before sequencing, and have confirmed that this dramatically improves the efficiency and accuracy of antigen identification. This is accomplished by transferring the proteins remaining after plucking to a nitrocellulose membrane, and immunoblotting with the patient serum. If the spot of interest is accurately plucked, there will be a halo of signal around the removed spot (as is shown in the boxed area, Fig. 2C). In those cases where it is challenging to accurately identify which spot to pluck, we recommend performing several plucks (including multiple gels and a variety of cell types when possible), and performing a post-pluck immunoblot on each. This strategy rapidly and inexpensively identifies the best gel plug, and only this is then sent for sequencing. LC/MS/MS peptide sequencing of a confirmed pluck of the M4 antigen from IFN- α treated HeLa lysate resulted in a list of twelve candidate proteins in the 50–75 kDa range that were absent from a nearby “blank” gel plug (Table 1). As expected, most of these proteins also had a predicted isoelectric point (pI) similar to that observed for the antigen of interest (estimated to be ~5.2 from the 2D immunoblot).

The triangulation approach requires at least one additional piece of information about the protein of interest to allow multiple different paths of pursuit to converge on the correct antigen identity. Fortunately, in the case of antigen M4, we were able to take advantage of expression profiling data from IFN- α treated cells performed for an unrelated project. We compared the list generated by proteomic sequencing of the antigen from IFN- α treated HeLa lysate (Table 1) to a list of genes which were induced by at least two-fold in IFN- α treated HSGs (Table 2). This quickly enabled us to focus on a single leading candidate present on both lists: IFN-induced protein with tetratricopeptide repeats 3 (IFIT3).

3.3. Confirmation of antigen identity

Several strategies were used to definitively demonstrate that the protein targeted by high titer serum autoantibodies of patient M4 is IFIT3. (i) Immunoblots performed with a monoclonal anti-IFIT3 antibody confirmed that IFIT3 is strongly induced in IFN- α treated HeLa cells and that it comigrates with the 60 kDa M4 protein (Fig. 3A, lanes 1,2,5 & 6); (ii) Both the M4 patient serum and the anti-IFIT3 monoclonal antibody immunoblot a 60 kDa protein expressed in HeLa cells transiently transfected with IFIT3 cDNA (Fig. 3A, lanes 4 & 8), but absent in untransfected cells (lanes 3&7); and (iii) ^{35}S -methionine labeled IFIT3 (^{35}S -IFIT3) was generated by IVTT using IFIT3 cDNA. Serum M4 immunoprecipitated ^{35}S -IFIT3 (Fig. 3B), whereas other myositis patient sera known to be anti-IFIT3 antibody negative (as assessed by immunoblot of IFN- α induced cell lysates, data not shown), did not. Healthy control sera also did not immunoprecipitate ^{35}S -methionine-IFIT3 (data not shown).

3.4. Validation of the above approach using muscle cells at successive stages of differentiation to screen for new myositis autoantibodies: identification of MYL4

We sought to validate the approach described above by using similar methods to detect myositis autoantigens whose expression is enhanced in the context of additional pathways

active in diseased tissue. Specifically, since myositis autoantigen expression has been demonstrated to be enhanced in differentiating muscle cells (Casciola-Rosen et al., 2005), we used these as a novel source for autoantigen discovery. Myositis serum S293 immunoblotted a 28 kDa band that was robustly expressed in muscle cells at day 4 (or thereafter) of differentiation, but was absent from myoblasts (day 0 cultures), as well as HeLa, HSG and keratinocyte lysates (Fig. 4A and data not shown). Lysates made from myoblasts harvested at day 4 of differentiation were electrophoresed on 2D gels (1st dimension, pI 3–10; 2nd dimension, 12% SDS-PAGE). Duplicate runs were done: the first gel was transferred to nitrocellulose and used for immunoblotting with serum S-293 to identify the spot corresponding to the 28 kDa protein. The second, replicate gel was stained with colloidal coomassie blue immediately after SDS-PAGE, and the 28 kDa spot was plucked and processed for peptide sequencing, after confirming by post-pluck blotting that the correct spot had been plucked.

To aid in identifying the real target autoantigen from the resulting list of candidate proteins, we used microarrays to obtain gene expression patterns during myoblast differentiation (days 0–5). In the case of the unidentified 28 kDa antigen, there were 8 candidate proteins from the 2D proteomic analysis. Two of the top candidates were known muscle-specific proteins; embryonic myosin light chain 4 (MYL4) and troponin I, skeletal slow twitch (TNNT1). When these were searched in our myoblast differentiation gene expression database, only 1 was absent at day 0, but was dramatically increased after 3–4 days of differentiation. This antigen was MYL4, a muscle-specific, differentiation stage-specific 28 kDa protein (Fig. 4B), making it the top candidate.

We validated that serum S293 does have antibodies to MYL4 in the following ways. Immunoblots performed using a monoclonal anti-MYL4 antibody confirmed that the kinetics of expression during myoblast differentiation were identical to those seen using S293 serum (Fig. 4C). Furthermore, the 28 kDa band blotted by both the monoclonal antibody and S293 co-migrated (Fig. 4D). Lastly, we cloned the cDNA for MYL4 from day 4 myoblast cultures and transiently transfected it into HeLa cells. The 28 kDa band was detected by immunoblotting in transfected lysates only, using both the monoclonal antibody and serum S-293 (Fig. 4D). Taken together, the above approaches confirm that the 28 kDa antigen blotted by S293 is MYL4.

3.5 Expression of IFIT3 and MYL4 in muscle biopsy lysates

We used lysates made from muscle biopsies (the target tissue in myositis) to examine expression of IFIT3 and MYL4 *in vivo*. Lysates were made from control, DM and polymyositis (PM) muscle. IFIT3 was detected in DM muscle only, and was absent from PM and control muscle (Fig. 5, lanes 4–6), consistent with the well-documented type 1 IFN signature in DM muscle (Greenberg et al., 2005). Of note, IFIT3 expression was also found to be enhanced (though at lower levels) during muscle differentiation (absent at day 0, present at day 4, Fig. 5 lanes 1 & 3). MYL4 was robustly expressed in both DM and PM muscle biopsy lysates but was not detected in control muscle (Fig. 5). This expression pattern is consistent with immunohistochemical data showing extensive muscle cell regeneration in both DM and PM (Casciola-Rosen et al., 2005), but not in control muscle.

3.6 Principles of the triangulation strategy

The triangulation approach described here has 3 main features which improve the efficiency of autoantigen identification (Fig. 6): (i) Use of antigen sources which express pathways shown to be active in the diseased tissue *in vivo* (e.g., specific cell or tissue types, differentiation stages, activated signaling pathways, etc). This may allow the detection of novel autoantigens of mechanistic importance missed by other screening strategies; (ii)

Confirmation of spot selection by post-pluck immunoblotting prior to LC/MS/MS peptide sequencing; and (iii) Use of data from additional methods (e.g. biochemical properties of the antigen, microarray analysis of gene expression) to filter the proteomic data, prior to rigorous experimental confirmation of antigen identity. This filtering step narrows the number of candidates and limits the expense and time required for confirmation. As were used in this study, an antibody specific for the protein candidate and/or a cDNA expression plasmid are very useful in providing unequivocal evidence of antigen identity. Additional methods of antigen identity confirmation may include knocking down expression of the protein with specific siRNAs, screening a knockout cell line, comparison of protease specific cleavage patterns, or 2D immunoblotting to confirm exact co-localization using the patient serum and a protein specific antibody.

A potential disadvantage of this strategy is that identification of proteins by 2D gel and LC/MS/MS sequencing can be challenging for very large and poorly soluble proteins (Mallick and Kuster, 2010), such as those in cell membranes, although sequencing of cleavage fragments instead of the full length protein may circumvent this limitation. Post-translational modifications also may pose significant challenge, as the modified residues can result in peptide fragments that are difficult to identify, especially in the case of glycosylation (Duncan et al., 2010). Furthermore, filtering by changes in gene expression is not suited to autoantigens generated by post-translational modification. Evaluating novel antigens for post-translational modifications prior to sequencing permits the identification of modifications that should be included in the analysis of the LC/MS/MS data and potentially allows for their removal prior to sequencing. Finally, if a novel antigen is a protein that has not been previously identified, it may be absent from the mammalian database and thus may also be missed by this approach.

These studies emphasize that current technologies enable rapid and efficient autoantigen discovery, and allow novel autoantigens expressed in pathways relevant to disease pathogenesis to be defined. Such autoantigen targets – when induced by immune effector or tissue repair pathways - may be of particular importance in amplifying and sustaining injury in these diseases. These molecules may provide a new class of probes for diagnosis and monitoring.

Acknowledgments

We thank the Johns Hopkins Mass Spectrometry and Proteomics Facility and the Johns Hopkins Microarray Core Facility for excellent services. These studies were supported by NIH grants AR-44684 (to LCR), and P30 AR053503 and R37 DE12354 (to AR), and by the Donald and Dorothy Stabler Foundation.

Abbreviations

2D	2-dimensional
DM	dermatomyositis
HSG cells	human salivary gland epithelial cells
IEF	isoelectric focusing
IFN	interferon
IFIT3	IFN-induced protein with tetratricopeptide repeats 3
IVTT	in vitro transcription and translation
LC/MS/MS	liquid chromatography tandem mass spectrometry

MDA5	melanoma differentiation-associated protein 5
MYL4	embryonic myosin light chain 4
PM	polymyositis
SERPA	serological proteome analysis
TNNI	troponin I (skeletal, slow)

References

- Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995; 57:289.
- Bizzaro N. Autoantibodies as predictors of disease: the clinical and experimental evidence. *Autoimmun Rev*. 2007; 6:325–333. [PubMed: 17537376]
- Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med*. 1975a; 292:344–347. [PubMed: 1090839]
- Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). *N Engl J Med*. 1975b; 292:403–407. [PubMed: 1089199]
- Canelle L, Bousquet J, Pionneau C, Deneux L, Imam-Sghiouar N, Caron M, Joubert-Caron R. An efficient proteomics-based approach for the screening of autoantibodies. *J Immunol Methods*. 2005; 299:77–89. [PubMed: 15914192]
- Casciola-Rosen L, Andrade F, Ulanet D, Wong WB, Rosen A. Cleavage by granzyme B is strongly predictive of autoantigen status: implications for initiation of autoimmunity. *J Exp Med*. 1999; 190:815–826. [PubMed: 10499920]
- Casciola-Rosen L, Nagaraju K, Plotz P, Wang K, Levine S, Gabrielson E, Corse A, Rosen A. Enhanced autoantigen expression in regenerating muscle cells in idiopathic inflammatory myopathy. *J Exp Med*. 2005; 201:591–601. [PubMed: 15728237]
- Darrah E, Rosen A. Granzyme B cleavage of autoantigens in autoimmunity. *Cell Death Differ*. 2010; 17:624–632. [PubMed: 20075942]
- Duncan MW, Aebersold R, Caprioli RM. The pros and cons of peptide-centric proteomics. *Nat Biotechnol*. 2010; 28:659–664. [PubMed: 20622832]
- Eggleton P, Haigh R, Winyard PG. Consequence of neo-antigenicity of the ‘altered self’. *Rheumatology (Oxford)*. 2008; 47:567–571. [PubMed: 18316337]
- Fiorentino D, Chung L, Zwerner J, Rosen A, Casciola-Rosen L. The mucocutaneous and systemic phenotype of dermatomyositis patients with antibodies to MDA5 (CADM-140): a retrospective study. *J Am Acad Dermatol*. 2011; 65:25–34. [PubMed: 21531040]
- Greenberg SA, Pinkus JL, Pinkus GS, Burleson T, Sanoudou D, Tawil R, Barohn RJ, Saperstein DS, Briemberg HR, Ericsson M, Park P, Amato AA. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. *Ann Neurol*. 2005; 57:664–678. [PubMed: 15852401]
- Greenberg SA, Higgs BW, Morehouse C, Walsh RJ, Won Kong S, Brohawn P, Zhu W, Amato A, Salajegheh M, White B, Kiener PA, Jallal B, Yao Y. Relationship between disease activity and type 1 interferon- and other cytokine-inducible gene expression in blood in dermatomyositis and polymyositis. *Genes Immun*. 2011; 67:53–63.
- Mallick P, Kuster B. Proteomics: a pragmatic perspective. *Nat Biotechnol*. 2010; 28:695–709. [PubMed: 20622844]
- Mammen AL. Autoimmune myopathies: autoantibodies, phenotypes and pathogenesis. *Nat Rev Neurol*. 2011; 7:343–354. [PubMed: 21654717]
- Mammen AL, Casciola-Rosen LA, Hall JC, Christopher-Stine L, Corse AM, Rosen A. Expression of the dermatomyositis autoantigen Mi-2 in regenerating muscle. *Arthritis Rheum*. 2009; 60:3784–3793. [PubMed: 19950298]

- Mammen AL, Chung T, Christopher-Stine L, Rosen P, Rosen A, Doering KR, Casciola-Rosen LA. Autoantibodies against 3-hydroxy-3-methylglutaryl-coenzyme A reductase in patients with statin-associated autoimmune myopathy. *Arthritis Rheum.* 2011; 63:713–721. [PubMed: 21360500]
- Pan W. A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments. *Bioinformatics.* 2002; 18:546–554. [PubMed: 12016052]
- Sato S, Hoshino K, Satoh T, Fujita T, Kawakami Y, Fujita T, Kuwana M. RNA helicase encoded by melanoma differentiation-associated gene 5 is a major autoantigen in patients with clinically amyopathic dermatomyositis: Association with rapidly progressive interstitial lung disease. *Arthritis Rheum.* 2009; 60:2193–2200. [PubMed: 19565506]
- Stinton LM, Fritzler MJ. A clinical approach to autoantibody testing in systemic autoimmune rheumatic disorders. *Autoimmun Rev.* 2007; 7:77–84. [PubMed: 17967730]

Highlights

- Enhancement of antigen discovery with disease relevant antigen sources.
- Modification of the serological proteome analysis method of antigen identification.
- Confirmation of antigen acquisition prior to proteomic sequencing.
- Improved efficiency of antigen discovery with a triangulation approach.

Fig. 1. Patient sera detect novel myositis autoantigens in IFN- α treated cells

(A) Immunoblot screen of untreated and IFN- α treated HSG cell lysates with six myositis patient sera (M1–M6). Note that patient M4 has high titer autoantibodies targeting a 60 kDa protein only detectable after IFN- α treatment (lanes 7 & 8). Migration of molecular weight standards are shown down the left side. (B) IFN- α treatment induces expression of the 60 kDa M4 antigen in HSG, HeLa and primary keratinocyte cultures. Vinculin blot reflects the amount of protein loaded in each lane.

Fig. 2. Immunoblot confirmation of protein spot selection increases antigen identification efficiency

(A) 2D immunoblot of IFN- α treated lysate using patient serum M4. First dimension: pI 3–10 IEF; second dimension, 10% SDS-PAGE. The area outlined by the white box is magnified in the lower right corner of each image. (B) Identical replicate preparative gel stained with gel code blue. The spot to be plucked (marked with the arrow) is identified by aligning the ponceau stained transfer and the immunoblot obtained in A. (C) Immunoblot confirming accurate removal of a gel plug containing the antigen of interest from the preparative gel. Note the halo of signal visible around the removed gel plug.

Fig. 3. Confirmation that the 60 kDa M4 antigen is IFIT3

(A) Immunoblots were performed with lysates made from HeLa cells incubated in the absence (lanes 1 & 5) or presence (lanes 2 & 6) of IFN- α , or transiently transfected with IFIT3 cDNA (lanes 4 & 8) or vector only (lanes 3 & 7). Equal amounts of protein were loaded in each gel lane. Immunoblots were performed with monoclonal anti-IFIT3 antibody (lanes 1–4) or M4 serum (lanes 5–8). (B) ^{35}S -methionine labeled IFIT3 is immunoprecipitated by the M4 serum (lane 1), but not anti-IFIT3 antibody negative myositis sera (M7–M9, lanes 2–4).

Fig. 4. Confirmation that the 28 kDa antigen blotted by S293 serum is MYL4

(A & C) Immunoblots were performed with lysates made from muscle cells harvested at various times during differentiation, using serum S293 (panel A) and monoclonal anti-MYL4 antibody (panel C). Equal amounts of lysate were electrophoresed in each gel lane. (B) Gene expression profiles of muscle cells harvested at days 0 – 5 of differentiation show that MYL4 gene expression is absent at day 0, but is robust by day 4, a pattern consistent with the protein expression in the same lysates. TNNI gene expression remained low throughout differentiation. Gene expression values are represented as a Z-score. (D) Untransfected HeLa cell lysates (denoted “MYL4 –”), or HeLa cells transiently transfected with MYL4 cDNA (denoted “MYL4 +”, middle lane of each set), or muscle cells at day 4 of differentiation were immunoblotted with monoclonal anti-MYL4 antibody or S293 serum. MYL4 (28 kDa) was detected in transfected HeLa cells and the day 4 muscle cell lysates by both. Equal protein amounts were electrophoresed in all gel lanes. (A, C & D) Vertical lines between lanes indicate where lanes on the original autorads have been spliced out for clarity of presentation.

Fig. 5. IFIT3 and MYL4 are detected in myositis muscle biopsy lysates, but not in control muscle
 Lysates made from undifferentiated myoblasts (day 0) incubated in the absence or presence of IFN- α (lanes 1 & 2), muscle cells at day 4 of differentiation (lane 3) and lysed muscle biopsies from patients with PM, DM or control individuals (lanes 3–6) were immunoblotted with monoclonal antibodies against IFIT3 or MYL4. Equal amounts of protein were loaded in each gel lane.

- Step 1. Screen for antigens specifically present in a disease relevant antigen source
- ↓
- Step 2. 2D immunoblot confirmation of antigen acquisition followed by proteomic sequencing
- ↓
- Step 3. Use sequencing data with secondary information to triangulate antigen identity and confirm rigorously

Fig. 6. Summary of the triangulation approach to antigen identification

This strategy can be adapted for a variety of antigen identification investigations. Efficient antigen identification is facilitated by focusing on a disease relevant antigen source, confirming accurate plucking of the protein spot of interest by immunoblotting the remaining 2D gel, and combining multiple pieces of information about the antigen to efficiently narrow the candidate pool prior to rigorous confirmation of antigen identity.

Table 1

Candidate proteins generated by LC/MS/MS peptide sequencing of a 2D gel spot containing the 60 kDa M4 antigen, listed in order of strength and restricted to those with a predicted MW of 50–75 kDa.

GI Number	Protein ^a	MW (kDa) ^b	pI ^b	Gene Symbol
gi 31542947	60 kDa heat shock protein, mitochondrial	61.1	5.7	HSPD1
gi 1777782	Interferon-induced protein with tetratricopeptide repeats 3	55.9	5.1	IFIT3
gi 28317	Keratin 10	58.8	5.1	KRT10
gi 20070228	Nucleobindin 1	53.9	5.1	NUCB1
gi 35655	Protein Disulfide Isomerase	57.1	4.8	P4HB
gi 624873	Rab GDI alpha	50.4	5.2	GDI1
gi 547754	Cytokeratin-2e	65.9	8.1	KRT2
gi 2343085	Programmed cell death protein 4	51.7	5.0	PDCD4
gi 12803709	Keratin 14	51.7	5.1	KRT14
gi 1136741	T complex protein 1 subunit theta	58.6	5.8	CCT8
gi 2160719	Myc box-dependent-interacting protein 1	64.7	5.0	BIN1
gi 121570	78 kDa glucose-regulated protein	72.3	5.1	HSPA5

^aOnly proteins predicted to be 50–75 kDa are shown.

^bPredicted based on amino acid sequence.

GI = GenInfo Identifier; aa = amino acids; MW = molecular weight; kDa = kilodaltons; pI = isoelectric point

Table 2

Genes whose transcripts are induced by 2 fold with IFN- α treatment of human salivary gland (HSG) cells with predicted MW of 50–75 kDa.

Gene Symbol	Fold Induction	Length (aa)	MW (kDa) ^a
IFI44	28.2	444	50.5
IFI44L	30.1	452	51.3
BCL2L13	2.04	485	52.7
WARS	3.45	471	53.2
TRIM38	2.52	465	53.4
TRIM21	5.97	475	54.2
IRF7	13.73	503	54.3
MUC13	2.56	511	54.5
IFIT2	12.88	472	54.6
SERPING1	2.2	500	55.2
IFIT1	27	478	55.4
IFIT5	2.43	482	55.8
IFIT3	31.2	490	56.0
LAP3	3.94	519	56.2
TRIM5	3.46	493	56.3
TRIM22	8.9	498	56.9
TRIM69	2.9	500	57.4
BTN3A1	2.09	513	57.7
OASL	8.33	514	59.2
EIF2AK2	3.34	551	62.1
REC8	2.25	547	62.6
SLC15A3	6.93	581	63.6
BTN3A3	2.09	584	65.0
LGALS3BP	4.04	585	65.3
UNC93B1	2.33	597	66.6
GBP1	4.16	592	67.9
TRIM25	2.13	630	71.0

^aPredicted based on amino acid sequence.

aa = amino acids; MW = molecular weight; kDa = kilodaltons