Molecular characterization of paired immunoglobulin heavy and light chains of clonally expanded antibodies in the presence of a strong polyclonal background

Birgit Obermeier (obermeb@gmail.com)
Institute of Clinical Neuroimmunology, Ludwig-Maximilians University, D-81377 Munich, Germany
Eduardo Beltrán
Institute of Clinical Neuroimmunology, Ludwig-Maximilians University, D-81377 Munich, Germany
Reinhard Hohlfeld
Institute of Clinical Neuroimmunology, Ludwig-Maximilians University, D-81377 Munich, Germany
Klaus Dornmair
Institute of Clinical Neuroimmunology, Ludwig-Maximilians University, D-81377 Munich, Germany

Method Article

Keywords: Oligoclonal band antibodies, B cells, Immunoglobulin, Proteomics, Transcriptomics, neurological inflammatory diseases

Posted Date: October 20th, 2017

DOI: https://doi.org/10.1038/protex.2017.113

Abstract

We describe a protocol for the direct molecular characterization of expanded antibody species detectable in minute amounts in human cerebrospinal fluid or other body fluids and tissues against a huge background of irrelevant antibodies. Antibodies are purified by affinity chromatography and applied to high resolution 2-dimensional PAGE under non-reducing conditions. Individual heterodimers consisting of heavy and light chains are isolated and analyzed by mass spectrometry. Since this rarely yields full sequence coverage, we reveal the complete sequences by alignment of the peptides to the immunoglobulin transcriptome of the same patient. Recombinant antibodies may then be produced and used for antigen searches or other experiments. We have analyzed antigens of oligoclonal band antibodies from the cerebrospinal fluid of patients with neurological inflammatory diseases, but the technology, which takes about two weeks, may be applied to many disease-relevant \(auto)antibodies in human infectious and autoimmune diseases, as well as to B-cell malignancies.

Introduction

Clonally expanded antibodies are commonly observed in various conditions such as autoimmune diseases, malignancies, and infections. They may have beneficial effects by defending the organism against tumors or intruders such as bacteria, viruses, or parasites. On the contrary they may have detrimental effects by attacking self-tissue as occurring in autoimmune diseases. Expanded antibody populations are usually accompanied by a plethora of unrelated antibodies, which may differ only by a few amino acids. Therefore it is virtually impossible to isolate them by conventional biochemical tools. The only suited technique would be affinity chromatography which, however, requires that the antigens are known _a priori_, but this is almost never the case. Several methods have been developed to analyze Immunoglobulin \(Ig) chains of single B or plasma cells[1-4], but it is usually not known which particular cell secretes disease-relevant expanded antibodies and which cell produces irrelevant Ig chains. We have therefore developed a protocol that starts out with the analysis of expanded antibodies at the protein level. It allows us to selectively characterize individual expanded antibody species in the presence of other expanded antibody species and a high background of unrelated antibodies. We recently applied this new technology to a particularly prominent example of clonally expanded antibodies, namely the oligoclonal band \(OCB) antibodies[5]. OCBs are present in the cerebrospinal fluid \(CSF) of most patients with inflammatory diseases of the central nervous system \(CNS) such as multiple sclerosis \(MS)[6-9]. They can be visualized as prominent, distinct Ig bands in isoelectric focusing \(IEF) gels of CSF samples. OCBs serve as important diagnostic and prognostic laboratory markers for MS that may even guide therapies. Despite intense efforts, their target antigens have remained unknown^6. The main reason is that a distinct OCB species could not be separated from other OCBs and from the strong polyclonal antibody background present in the CSF. Thus, a typical IEF gel of a patient´s CSF shows several distinct bands of variable intensities on a strong background that smears throughout the entire pH-range. All these antibodies most likely will recognize different antigens, but they are chemically very similar, because the vast majority of their amino acids are identical. Our new technology[5] is based on previous

studies where we asked whether the B cells present in CSF might produce OCBs[10]. To this end, we compared the IgG proteomes and transcriptomes of antibodies and B cells from the CSF of MS patients. To investigate how CSF B cells were related to B cells in the brain we compared CSF IgG transcriptomes and proteomes to the CNS[11]. We identified significant overlaps in both studies, providing evidence that the B cells present in the CSF are indeed the sources of the OCB antibodies and that they are clonally related to the B cells in CNS. Technically, we analyzed the IgG proteomes of CSF samples by IEF gels and subsequent mass spectrometry. In parallel we analyzed the IgG transcriptomes of the same patients by cloning and sequencing mRNAs coding for IgG heavy \(IgG-H) and light \(IgG-L) chains. To correlate IgG transcriptomes and proteomes, we made use of the extensive somatic hypermutations \(SHM) and unique hypervariable complementarity determining region 3 \(CDR3) rearrangements of OCBs. SHM are introduced in response to sustained antigen stimulation and are observed in most expanded antibody species. SHM and CDR3 rearrangements are unique to each antibody species. For comparing the peptides revealed by mass spectrometry to the transcriptomes of the patients, we therefore focused on peptides that contained SHM and CDR3 regions because they are unique to each patient \(termed "characteristic peptides"). Thus it was possible to align characteristic peptides unequivocally to full-length Ig-chains which were identified by transcriptome analysis. This demonstrated that B cells present in CSF and brain of MS patients may produce the OCB antibodies[10,11]. However, it did not resolve matching IgG-H and IgG-L chains because the heterodimeric IgG molecules disassembled partly during IEF. With our new method[5] it is now possible to isolate individual heterodimeric OCB antibody species directly from IEF gels, and to produce them in large amounts as needed for antigen searches. We developed our protocol by investigating OCBs from patients with MS but the method will be applicable to any body fluid or tissue that contains B cells and expanded antibody populations even on a strong background of irrelevant antibodies. Obvious applications include e.g. autoimmune and infectious diseases, allergies, or hematopoietic malignancies[12-15]. Further, the method is not only applicable to expanded IgG antibody populations but also to other Ig families. For example, IgM populations have recently been found to be particularly interesting, though still enigmatic, in MS[4], and IgE species are known to play a pivotal role in allergy and asthma[16].

Reagents

· RNeasy Micro Kit \(Qiagen) · DEPC-treated water \(Invitrogen) · IGEPAL® CA-630 for molecular biology \(Sigma) **CAUTION** May develop turbidity or sediment during storage. A clear liquid can be obtained on heating to 40 ° C. · Random Primers \(Invitrogen) · dNTP Set, 100 mM each A,C,G,T \(GE Healthcare) · RNasin® Plus RNase Inhibitor \(Promega) · Prime RNase Inhibitor \(Eppendorf) · SuperScript® III Reverse Transcriptase \(Invitrogen) · Taq DNA Polymerase \(Roche) · Agarose \(Biozym) · Trizma® base \(Sigma) **CAUTION** It is irritating to eyes. · Boric acid \(Sigma) **CAUTION** It is toxic. · EDTA, Titriplex® III \(Merck) · Ethidiumbromide, 1% stock solution \(Carl Roth) **CAUTION** It has mutagenic properties and is irritating. · Glycerol, ROTIPURAN® \(Carl Roth) · Bromophenol Blue \(Serva) · Xylene Cyanole FF \(Bio-Rad) **CAUTION** It is irritating to eyes, respiratory system and skin. · 100 bp DNA ladder \(New England Biolabs) · EasyPure® DNA Purification Kit \(Biozym) · TOPO-TA Cloning® Kit \(Invitrogen) · Bacto™

Trypton \(Becton Dickinson) · Bacto™ Yeast Extract \(Becton Dickinson) · Bacto™ Agar \(Becton Dickinson) · Ampicillin sodium salt \(Sigma) · 5-Bromo-4-chloro-3-indoyl ß-D-galactopyranoside \(X-Gal, Sigma) · DirectPrep 96 Miniprep Kit \(Qiagen) · NaH_2PO_4 x H_2O \(Carl Roth) · Na_2HPO_4 \(Carl Roth) · NaCl \(Merck) · Dynabeads® Protein G \(Invitrogen) · Na-acetate \(Merck) · SDS \(Carl Roth) **CAUTION** It is harmful. Avoid inhalation and skin contact. · MEGA-10 \(Merck) · N-Glycosidase F recombinant \(Roche) · Urea \(Sigma) · Thiourea \(Sigma) **CAUTION** It is harmful if swallowed. · Pharmalyte™, broad range pH 3-10 \(GE Healthcare) · Immobiline™ Dry Strip, pH 3-10, 24 cm \(GE Healthcare) · SDS-PAGE Molecular Weight Standard, broad range \(Bio-Rad) · Brilliant Blue R 250 \(Sigma) · Methanol \(Sigma) **CAUTION** It is flammable. · Acetic acid \(Sigma) **CAUTION** It is corrosive and irritating. · Ammonium bicarbonate \(NH_4HCO_3, Sigma) · DL-Dithiothreitol \(DTT, Sigma) **CAUTION** It is toxic and may cause irritation to skin and respiratory system. · Acetonitrile \(Sigma) **CAUTION** It is flammable and harmful. · Jodoacetamide \(Sigma) **CAUTION** It is toxic. · Trypsin modified from porcine pancreas \(Serva) · Trifluoroacetic acid \(TFA, Applied Biosystems) **CAUTION** It is harmful and corrosive. · alpha-Cyano-4-Hydroxycinnamic acid \(Bruker Daltonics) **CAUTION** It is harmful. **Reagent Setup** · TBE \(10x): 0.89 M Tris, 0.89 M Boric acid, 0.02 M EDTA, pH 8.0; Store at room temperature. · DNA sample buffer \(6x): 50% \(v/v) Glycerol, 0.02% \(w/v) Bromphenol blue, 0.02% \(w/v) Xylencyanol FF, 10 mM Tris, pH 7.5; Store at 4 ° C. · LB-Medium: 1% \(w/v) Bacto™ Trypton, 0.5% \(w/v), Bacto™ Yeast Extract, 0.5% \(w/v) NaCl; Autoclave and store at room temperature. · LB-Agar: LB-Medium with 1.5% \(w/v) Bacto™ Agar, supplied with 100 µg/ml Ampicillin; Autoclave and store plates at 4 ° C. · PBS \(10x): 18.8 mM NaH_2PO_4 x H_2O, 83.8 mM Na_2HPO_4, 1,5 M NaCl; Store at room temperature. Check pH after dilution to 1x. · Washing buffer for Dynabeads® Protein G: 0.1 M Na-acetate, 0.15 M NaCl, pH 5.0; store at room temperature. Check pH before use. · Elution buffer for Dynabeads® Protein G: 1% \(w/v) SDS in PBS; Store at room temperature. · Deglycosylation buffer: PBS with 1% \(w/v) MEGA-10, pH 7.2; Store at room temperature. **CRITICAL** In case of MEGA-10 precipitates heat buffer while stirring until detergent has dissolved completely. Let cool down to room temperature and check pH before use. · Rehydration buffer for IPG strips: 8 M Urea, 2 M Thiourea, 2% \(v/v) Pharmalyte™ 3-10, Bromphenol blue; Prepare a fresh solution for each experiment. · IEF sample buffer: 6 M Urea, 2 M Thiourea, 10% \(v/v) Glycerol, Bromphenol blue; Prepare a fresh solution for each experiment. · Equilibration buffer for 2D gel: 6 M Urea, 4% \(w/v) SDS, 30% \(w/v) Glycerol, 50 mM Tris, Bromphenol blue; Prepare a fresh solution for each experiment. · Coomassie staining: 0.1% \(w/v) Coomassie® Brilliant-Blue R-250, 40% Methanol, 10% Acetic acid; Store at room temperature. · Buffer for Trypsin digestion: 10 mM NH_4HCO_3, pH 8.5; Prepare a fresh solution for each experiment. · Matrix for MALDI: alpha-Cyano-4-Hydroxy-Cinnamic Acid \(5 mg/ml) in 50% Acetonitrile/0.2% TFA; Store at room temperature but not longer than 3 days.

Equipment

· Tabletop Microcentrifuge \(Eppendorf) · Thermocycler \(T personal, Biometra; GeneAmp PCR System 9600, Perkin Elmer) · Equipment for agarose gel electrophoresis · UV transilluminator \(Bachofer Laborgeräte) · Water bath \(MA6, Lauda) · Shaker 37 ° C \(HT, Infors) · Megafuge 1.0.R \(Heraeus

Instruments) · Microcentrifuge Tubes, DNA low binding \(Eppendorf) · Microcentrifuge Tubes, Protein low binding \(Eppendorf) · Magnet \(Invitrogen) · Platform Shaker \(Unimax 2010, Heidolph Instruments) · Thermomixer \(Comfort 5436, Eppendorf) · Incubator 37 ° C \(Jouan EB 53, Thermo Fisher Scientific) · D-tube^{TM} Dialyzer Mini \(MWCO 12-14 kDa, Merck) · Stirrer \(Ikamag RCT, IKA® Werke) · Power Supply \(EPS 500/400, Amersham Pharmacia) · 3100 OFFGEL Fractionator \(Agilent) · Loading cup for IEF, 8 x 2 mm, conical · Equipment for 2D gel electrophoresis · ZipTip® Pipette Tips \(Millipore) · MALDI Target \(Applied Biosystems) · Proteomics Analyzer 4700 MALDI-TOF/TOF \(Applied Biosystems) · “Data Explorer” Software · Program “Mascot Peptide Mass Fingerprint” \(Matrix Science) · Program “Mascot MS/MS Ions Search \(Matrix Science)

Procedure

A. Analysis of Ig-V\(D)J-transcriptomes Three different techniques are applicable to generate Ig-H and Ig-L transcriptome databases \(**Fig. 2**). **Table 1** lists their advantages and disadvantages. Firstly, state of the art NGS techniques may be employed. A protocol that describes analysis of Ig repertoires has recently been published[17]. This method uses "unique molecular identifier" \(UMI) sequences which are introduced during generation of cDNA and therefore allow for precise quantification of individual mRNA molecules. This protocol is recommended because it reveals the entire Ig-H- and Ig-L-chain repertoires quantitatively, whereas Sanger-sequencing and single cell sequencing may reveal only a limited spectrum of the repertoires. Secondly, traditional transcriptome analysis has been outlined earlier[10]. A detailed methodology of this Sanger sequencing based method is available^19. These steps describe isolation of RNA, RT-PCR, cloning and sequence analysis of Ig-H and Ig-L chains. Thirdly, techniques for analyzing Ig chains from single B cells were recently advanced from traditional sequencing methods[2,3] to NGS techniques[18]. Although both Sanger-sequencing and NGS based methods directly will yield paired Ig-H and Ig-L chains, it will be uncertain how they will relate to distinct OCBs or other expanded Ig species. Moreover, it will not be clear whether or not a certain Ig-H and Ig-L pair will yield an OCB at all. This is an important issue, since only limited number of OCBs \(usually less than 20) are observed in a patient, whereas the number of different B and plasma cell clones in the CSF often exceeds thousands. Thus, only a small fraction of B or plasma cells will ultimately produce an OCB. Therefore, alignment of Ig-H and Ig-L chain databases to peptides identified by mass spectrometry is indispensable. **PAUSE POINT** PCR products and isolated plasmids can be kept for long-term storage at -80 ° C. **B. Purification of individual expanded IgG species** **B.1: IgG Purification by Protein G affinity chromatography** Steps 1 to 6 have been outlined earlier[10]. Detailed methodology is available[19]. For clarity, we include these steps in the protocol described here. 1. Thaw CSF supernatant on ice. 2. Check for precipitates by centrifugation at 17,500 g for 5 min at 4 ° C. In case of any precipitates transfer supernatant in a new microcentrifuge tube and keep on ice until further processing. 3. Wash Dynabeads® Protein G three times in 1 ml washing buffer. 4. Add required amount of Dynabeads® Protein G to CSF supernatant. Use 2 µl bead suspension for 1 µg of IgG. **CRITICAL STEP** Ensure that your Ig of interest binds to protein G as binding strength is subtype- and species-dependent. 5. Incubate for 60 min on a shaker at room temperature. **CRITICAL STEP** To ensure efficient mixing split CSF supernatant in 1 ml fractions and use 2 ml microcentrifuge

tubes each. 6. Wash Dynabeads® Protein G three times in 1 ml washing buffer. 7. Elute IgG by resuspending the beads in 10 µl 1% SDS followed by incubation for 2 min at 37 ° C. 8. Repeat step 16 with 15 µl of 1% SDS. **CRITICAL STEP** This step is required to elute residual IgG molecules still bound to Protein G after first round of elution. 9. Pool all eluates in a protein low bind 1.5 ml microcentrifuge tube. Final volume for purified IgG from 1 ml body fluid supernatant is thus 25 µl. **B.2a: Deglycosylation** 10. Incubate purified IgG at 95 ° C for 1 min. **CRITICAL STEP** By these means IgG molecules get unfolded so that glycosylation sites are better accessible for the enzyme. 11. Add 1 volume of deglycosylation buffer and mix. 12. Add 100 U/ml N-Glycosidase F recombinant. 13. Incubate for 3 h at 37 ° C. **B.2b: Dialysis** 14. Dialyze sample against 6 M Urea in rehydrated D-tube^{TM} Dialyzer Mini tubes for 2 h on a stirrer at room temperature. 15. Dialyze further against 6 M Urea for 15 min at 50 V in a flat bed gel electrophoresis chamber according to the manufacturer's recommendations to remove residual SDS. **CRITICAL STEP** Longer dialysis and higher voltages will cause more pronounced disassembly of H- and L-chains. **B.3: Isoelectric focusing** **CAUTION** Make sure that all steps are carried out under strictly non-reducing conditions. 16. Rehydrate Immobiline^{TM} Dry Strip \(24 cm, pH 3-10) in rehydration buffer overnight at room temperature according to the manufacturer's recommendations. **CAUTION** Must be carried out the day before performing isoelectric focusing. 17. Add 50 µl of IEF sample buffer to dialyzed IgG. 18. For isoelectric focusing use an Agilent 3100 OFFGEL Fractionator in the "in-gel" mode according to the manufacturer's recommendations but with two essential exceptions: 18.a Instead of rehydrating the IPG strips together with the eluate, place a loading cup at position pH 4.5 onto the rehydrated IPG strip and thus load the sample. **CRITICAL STEP** Applying the sample at other positions than 4.5 lead to pronounced disassembly of H- and L-chains \(if loaded at extreme pH values) or to insufficient migration of IgG molecules into gel matrix \(if applied at pH values near isoelectric points of IgG). **CAUTION** Capacity of loading cup is 200 µl. 18.b Modify the default focusing method to a slower voltage-increase and extended duration: for the first 30 min the voltage is limited to 500 V and then for 30 min to 1000 V. During electrophoresis allow maximum values of 8000 V, 50 µA and 200 mW until 120 kVh are reached. **CRITICAL STEP** Using different focusing methods than described above might lead to disassembly of H- and L-chains or insufficient focusing. If focusing is performed with other IEF equipments it might therefore be necessary to optimize the default focusing method. **B.4: Non-reducing SDS-PAGE** **CAUTION** Make sure that all steps are carried out under absolutely non-reducing conditions. 19. Shake IPG strip in equilibration buffer for 20 min at room temperature. 20. Set up second dimension of IgG separation according to standard methods and cool the chamber during electrophoresis overnight. **CRITICAL STEP** Use 9% acrylamide gels. Higher percentage gels will prevent IgG molecules from efficiently migrating into the gel matrix. Lower percentage gels are difficult to handle. **CRITICAL STEP** When starting electrophoresis, first apply 10 mA for 1 h to facilitate protein transmigration from IPG strip into SDS-PAGE gel. Then increase current to 30 mA. 21. Stain 2D gel with Coomassie-Blue for 2 h on a shaker at room temperature. 22. Destain with 10% Acetic acid. **C: Analysis of matched Ig-H and Ig-L chains** **C.1: In-gel digestion by Trypsin** 23. Excise individual spots containing H_2L_2 Heterodimers \(apparent molecular weight = 200 kDa) and transfer into 0.6 ml microcentrifuge tubes. **PAUSE POINT** Spots can be stored covered by H_2O at 4 ° C until further

processing. 24. Proceed with Trypsin digestion as previously described[10,19]. **C.2: MALDI-TOF/TOF mass spectrometry** 25. Desalt samples after Trypsin digestion using ZipTip® Pipettes according to the recommendations of the manufacturer. 26. Mix sample and MALDI-matrix 1:1 and spot 1 µl onto a MALDI target. 27. Let dry at room temperature. 28. Measure samples using the Proteomics Analyzer 4700 \(MALDI-TOF/TOF) mass spectrometer with a 355 nm Nb-YAG-laser in positive reflector mode with 20 kV acceleration voltage. For MS/MS measurements perform CID. **C.3: Alignment of peptide identified by mass spectrometry to the Ig-V\(D)J trancriptome database** 29. Analyze MS and MS/MS spectra with “Data Explorer” Software and copy peaklist in program “Mascot Peptide Mass Fingerprint” or “Mascot MS/MS Ions Search”, respectively. 30. Search for matching peptides with the patient specific Ig-V\(D)J-transcriptome \(generated in step A) using the following parameters: 30.a for MS spectra: Enzyme: Trypsin; Allowed missed cleavages: 1; Fixed modifications: Carbamidomethyl \(C); Variable modifications: Oxidation \(M), \(W); Peptide tolerance: +/- 100 ppm; Mass values: MH+; Monoisotopic; 30.b for MS/MS spectra: Enzyme: Trypsin; Allowed missed cleavages: 1; Fixed modifications: Carbamidomethyl \(C); Variable modifications: Oxidation \(M), \(W); Peptide tolerance: +/- 100 ppm; Peptide charge: 1+; MS/MS tol: +/- 0.8 Da; Monoisotopic; Precursor: mass of interest. **C.4: Semi-quantitative analysis of MALDI-TOF data** 31. To distinguish between relevant peaks from expanded antibodies and irrelevant peaks from polyclonal antibody background only consider matching peptides whose peak intensities are above the following thresholds: · mass below 1.5 kDa: peaks exceeding background for factor of 5 · mass range 1.5 – 2.5 kDa: peaks exceeding background for factor of 7 · mass range 2.5 kDa – 4.0 kDa: peaks exceeding background for factor of 5 · mass above 4.0 kDa: peaks exceeding background for factor of 3 32. In order to unequivocally determine matching H- and L-chains of one expanded antibody out of the pool of “dominant” peptides from step 31 the following criteria must be fulfilled: 32a. Three or more peptides belong to only one H- and only one L-chain in the patient-specific Ig-V\(D)J transcriptome, respectively. 32b. At least one of three peptides contains SHM or comprises the hypervariable CDR3 region, i.e. is characteristic. 32c. Peak intensities of peptides from V\(D)J-regions are about as high as the intensities of peptides from the constant Ig-regions. 33. Obtain the full-length sequence of Ig-chains with the help of the Ig-V\(D)J-transcriptome. **CRITICAL STEP** For H/L-identification do not consider spots with significant contaminations of peptides from other Ig-chains or spots where many peptides are detected that can be aligned to different Ig-chains.

Timing

A. Transcriptome analysis: 1-2 weeks B.1 IgG Purification: 90 min B.2a Deglycosylation: 3 h B.2b Dialysis: 2 h 15 min B.3 Isoelectric focusing: overnight B.4 Non-reducing SDS-PAGE: overnight C.1 In-gel digestion by Trypsin: overnight C.2 MALDI-TOF/TOF: variable, depending on number of spots C.3 and C.4 Analysis of MALDI-TOF data: variable, depending on number of mass spectra to analyze

Troubleshooting

Step 4: Protein G does not bind Binding strength of protein G differs depending Ig of interest on subtype and species. You might use Dynabeads® Protein A instead or prepare your own Dynabeads® with specific antibody against particular Ig-subtype or species using the Dynabeads® Antibody Coupling Kit. Test binding and elution with a commercial antibody of the subtype of interest before proceeding with clinical samples. **Step 18a: Sample exceeds loading capacity for IEF** You may concentrate your sample by TCA or Acetone precipitation and solubilize the pellet directly in IEF sample buffer. **Step 22: No spots on 2-d gel** Preselect clinical samples for high content of Ig as detection limit of Coomassie-Blue is limited \(about 500 ng per spot). Gels may be stained with silver after Coomassie-Blue staining, but be aware that only some protocols are compatible with subsequent analysis by mass spectrometry[18]. Quality of mass spectromery data from silver-stained spots are mostly inferior. **Step 22: Smear on 2-d gel** This indicates that the clinical sample mainly contains polyclonal antibody populations. Pre-select clinical samples for monoclonally or oligoclonally expanded antibodies. When working with CSF samples, pre-select for high Ig quotients. **Step 22: Disassembly of H-/L- chains** Make sure that the equipment used for entire IgG analysis is free of any contaminating reducing agents. Check with a small aliquot of sample on 1-d SDS-PAGE whether Ig-H- and L-chains are still linked after dialysis and prior to IEF. Make sure that sample loading for IEF is carried out between pH 4.5 and pH 5. Loading at pH < 4 or pH > 9 leads to pronounced disassembly by reduction of disulfide bonds. **Limitations** There are two obvious limitations to this method: Firstly, it is evident that the RNA recovered from tissue or body fluid needs to be of sufficient quality. Specifically, samples must not be fixed by aldehydes. Secondly, sufficient amounts of antibody must be available for mass spectrometry. A typical MALDI experiment as employed in our experiments[5] requires about 1 µg protein contained in the spot to be analyzed. As a rule of thumb, this corresponds to a visible spot on a Coomassie Blue stained gel. Although there are more sensitive mass spectrometery equipment available, a limitation is the visualization of the H_2L_2 spots after 2-d electrophoresis. A more intricate limitation is the requirement that clonally expanded B cells and antibodies are indeed present in the tissue or body fluid to be analyzed. The more limited their heterogeneity, the easier it will be to identify and characterize the expanded cells and antibodies from the samples. This, however, will depend on the disease or condition to be studied. We note that it is not necessary to use purified B cells or antibodies, because the Ig transcripts are exclusively amplified by Ig-specific PCR primers, and antibodies will be purified by affinity chromatography.

Anticipated Results

In contrast to our earlier studies[10,11] our new protocol now allows us to characterize matching IgG-H and IgG-L chains from single spots on non-reducing 2-dimensional \(2-d) gels. Basis is the co-purification of disulfide-linked IgG-H and IgG-L chains, concomitant analysis by mass spectrometry **\(Fig. 1)**. See figure in Figures section. We first purified the Ig molecules by Protein A or G affinity chromatography \(step #1), and introduced a degycosylation step to reduce the heterogeneity of bands on the gels \(step #2). The most important improvement, however, was to modify IEF conditions in order to prevent pronounced dissociation of IgG-H and IgG-L chains \(step #3). Previously[10,11], significant disassembly of the heterodimers was observed. This disassembly was probably caused by the high electrical field

required for the polyacryamid IEF gels. The new procedure also caused some disassembly, but to a much lesser extent, so that enough heterodimeric antibody was left for analysis. To separate complexes consisting of two IgG-H- and two IgG-L-chains \(H_2L_2) from complexes where one or both light chains were disassembled \(H_2L and H_2) and from single chains, we added a second dimension to gel electrophoresis \(step #4), i.e. we ran a non-reducing SDS-PAGE after the IEF gels. This procedure yielded single spots, which we cut from the gels, in-gel digested with trypsin \(step #5), and analyzed by mass spectrometry \(step #6). We typically identified several peptides from the V\(nDn)J-regions of both chains, but these peptides did not cover the entire length of the chains. Therefore we created in parallel IgG transcriptome databases which contained full length sequences of the IgG-H and IgG-L chains of each patient \(step #7). Such databases can be obtained from unfractionated body fluid or tissue by Sanger sequencing of many cDNA clones[5,10] \(**Fig. 2**, **left panel**) or by NGS[5,17] \(**Fig. 2**, **middle panel**). Another possibility is the analysis of many single cells \(**Fig. 2**, **right panel**). However, this will not tell whether or not a particular cell is producing an expanded antibody, therefore correlation to proteomes is still required for identification of disease-related clones. **Fig. 2** gives an overview on the methods and **Table 1** lists their advantages and disadvantages. See figure in Figures section. See figure in Figures section. The characteristic peptides identified by mass spectrometry can then be aligned to the full-length sequences identified by cDNA cloning \(**Fig. 1**, step #8), and distinguished peptides from expanded OCB antibodies from background signals by a semi-quantitative analysis \(step #9). This analysis allowed us to identify one IgG-H and one IgG-L chain from a single spot on the 2-d gel. We note that a high sequence coverage by mass spectrometry is not required. It is sufficient that one or few characteristic peptides for each chain can be identified, so that the peptide\(s) may be attributed unequivocally to a particular Ig full-length sequence from the transcriptome database. Such an alignment will reveal all missing parts of the sequences. Thus, identification of few peptides that cover only parts of the sequences will permit the reconstruction of the entire chain due to the unique amino acids introduced by SHM and V\(D)J recombination. This will allow the recombinant expression of the antibodies that were expanded in the patients´ tissues or body fluids. Finally, these antibodies may be used to search for their antigens or for other purposes. Here, we show exemplarily the proteomic analysis of the monoclonal antibody 145.2c11 \(eBioscience) **\(Fig. 3)**. See figure in Figures section. First, for comparison, we show a "traditional" IEF gel where disulfide bonds were reduced and no deglycosylation steps were performed \(**Fig. 3a**). The proteins focus between pH 5 and 7. Our improved protocol results in a 2-d gel that shows 7 bands migrating at 200 kDa corresponding to the H_2L_2 heterodimer, 7 bands migrating at 150 kDa, corresponding to the H_2L, and 7 bands at 100 kDa, corresponding to H_2 \(**Fig. 3b**). Free Ig-L chains migrate at 25 kDa. Several spots are visible and the proteins focus between pH 7 and 8, presumably due to lack of negatively charged carbohydrates, such as sialic acid and glucuronic acid. The insert shows the spots that were analyzed by mass spectrometry \(**Fig. 3b**). We identified five peptides from the Ig-H, and four peptides from the Ig-L chain \(**Fig. 3c**). As expected, no peptides from the Ig-L chains were detected from the H2 complex \(*Fig. 3d). We applied this protocol to CSF samples from a patient with neuroborreliosis and four patients with MS[5]. We identified and cloned seven matching Ig-H and Ig-L chains representing OCB antibodies. After recombinant production in HEK

cells, we identified the antigens of the neuroborreliosis antibody and of three MS antibodies. **Comparison with other methods** An established method to identify matching pairs of Ig-H and Ig-L chains is cloning mRNA/cDNA from single B or plasma cells that have been isolated by flow cytometry[1-3]. Meanwhile a highly efficient and quantitative next generation sequencing \(NGS) based protocol is available^17. However this approach may not relate a particular cell to an expanded antibody species. Thus, it will be impossible to say which particular cell out of the thousands of B or plasma cells in a clinical sample will produce a disease-related expanded antibody and which will contribute to the polyclonal background antibody population. There is currently no protein-based method that allows the characterization of clonally expanded antibodies in the presence of other expanded antibodies and a strong polyclonal background unless specific antigens are already known. The reasons are obvious: antibodies are very similar in structure, in particular if they belong to the same \(sub)families. The theoretical variability of antibodies is almost infinite due to the introduction of random amino acids by recombination and SHM, but this relates predominantly to the small antigen combining sites. However, the overall structure and the large conserved regions are identical for all antibodies of the same family. Therefore classical biochemical methods \(such as ion exchange-, gel-, reversed phase-chromatography etc.) are not suited to isolate one distinct antibody species out of a vast number of others. Only affinity chromatography may be suited, but it evidently requires a priori knowledge of the antigen. Exactly this is usually not known in most biomedically interesting conditions. On the contrary, it is hoped that expanded antibodies may be utilized as tools for identifying enigmatic antigens.

References

1. Wardemann, H. et al. Predominant autoantibody production by early human B cell precursors. Science 301, 1374-1377 \(2003). 2. Owens, G.P. et al. Antibodies produced by clonally expanded plasma cells in multiple sclerosis cerebrospinal fluid. Ann. Neurol. 65, 639-649 \(2009). 3. von Budingen, H.C. et al. Clonally expanded plasma cells in the cerebrospinal fluid of MS patients produce myelin-specific antibodies. Eur. J. Immunol. 38, 2014-2023 \(2008). 4. Beltran, E. et al. Intrathecal somatic hypermutation of IgM in multiple sclerosis and neuroinflammation. Brain 137, 2703-2714 \(2014). 5. Brandle, S.M. et al. Distinct oligoclonal band antibodies in multiple sclerosis recognize ubiquitous self-proteins. Proc. Natl. Acad. Sci. U. S. A. 113, 7864-7869 \(2016). 6. Hohlfeld, R. et al. The search for the target antigens of multiple sclerosis, part 2: CD8+ T cells, B cells, and antibodies in the focus of reverse-translational research. Lancet Neurol. 15, 317-331 \(2016). 7. Stangel, M. et al. The utility of cerebrospinal fluid analysis in patients with multiple sclerosis. Nat. Rev. Neurol. 9, 267-276 \(2013). 8. Awad, A. et al. Analyses of cerebrospinal fluid in the diagnosis and monitoring of multiple sclerosis. J. Neuroimmunol. 219, 1-7 \(2010). 9. Antel, J. & Bar-Or, A. Roles of immunoglobulins and B cells in multiple sclerosis: from pathogenesis to treatment. J. Neuroimmunol. 180, 3-8 \(2006). 10. Obermeier, B. et al. Matching of oligoclonal immunoglobulin transcriptomes and proteomes of cerebrospinal fluid in multiple sclerosis. Nat. Med. 14, 688-693 \(2008). 11. Obermeier, B. et al. Related B cell clones that populate the CSF and CNS of patients with multiple sclerosis produce CSF immunoglobulin. J. Neuroimmunol. 233, 245-248 \(2011). 12. Ali, Y.M. et al. Monoclonal gammopathy in systemic lupus erythematosus. Lupus 16, 426-429

\(2007). 13. Konstantinopoulos, P.A. et al. Protein electrophoresis and immunoglobulin analysis in HIV-infected patients. Am. J. Clin. Pathol. 128, 596-603 \(2007). 14. Mitus, A.J. et al. Monoclonal and oligoclonal gammopathy after bone marrow transplantation. Blood 74, 2764-2768 \(1989). 15. Davies, J.M. & O'Hehir, R.E. Immunogenetic characteristics of immunoglobulin E in allergic disease. Clin. Exp. Allergy 38, 566-578 \(2008). 16. Gould, H.J. & Sutton, B.J. IgE in allergy and asthma today. Nat. Rev. Immunol. 8, 205-217 \(2008). 17. Turchaninova, M.A. et al. High-quality full-length immunoglobulin profiling with unique molecular barcoding. Nat. Protoc. 11, 1599-1616 \(2016). 18. Imkeller, K. et al. sciReptor: analysis of single-cell level immunoglobulin repertoires. BMC Bioinformatics 17, 67 \(2016). 19. Obermeier, B. et al. Analysis of patient-specific immunoglobulin proteomes and transcriptomes by PCR cloning and mass spectrometry doi:10.1038/nprot.2008.87 \(2008)

Acknowledgements

We thank Josef Kellermann and Friedrich Lottspeich for valuable discussion, and Monica Zobawa, Reinhard Mentele, Joachim Malotka, and Ingrid Eiglmeier for expert technical assistance. This work was supported by grants from Deutsche Forschungsgemeinschaft \(SFB 571-A1, CRC-128-A5, CRC-128-B8, and SyNergy EXC 1010 to KD and RH).

Figures

Figure 1: Flowchart for the characterization of individual antibodies. Start
usually a diagnostic body fluid or tissue, but other sample formats from an
possible as well. The only requirements are that expanded antibody species
producing B cells are contained. The analysis of Ig proteins (left panel) sta
all proteins contained in the sample. Antibodies are purified from the samp
protein G affinity chromatography **(step 1)**. Then the antibodies are deglyc
more homogeneous preparation **(step 2)**. Subsequently a modified 2-d gel-
performed. The first dimension is an improved IEF electrophoresis using p
(step 3), and the second dimension is SDS-PAGE **(step 4)**. It is essential th
step 4 are performed under strictly non-reducing conditions to prevent disa
H and Ig-L chain heterodimers (H_2L_2) as far as possible. Spots that contain
are excised from the stained 2-d gels and in-gel digested with trypsin **(step**
peptides can then be analyzed by mass spectrometry **(step 6)**, yielding pept
a subset also peptide sequences. The Ig transcriptome is determined in para
this end, RNA from B-lineage cells is transcribed into cDNA by reverse tra
cDNA transcripts are amplified, sequenced and analyzed, as described in d
Ig peptides obtained in steps 1 through 6 are then aligned to the patient or s
V(D)J-transcriptome database which has been established in step 7 **(step 8**
quantitative analysis is performed, which reveals the full-length sequences
and Ig-L chains **(step 9)**.

Figure 1

Flowchart for the characterization of individual antibodies

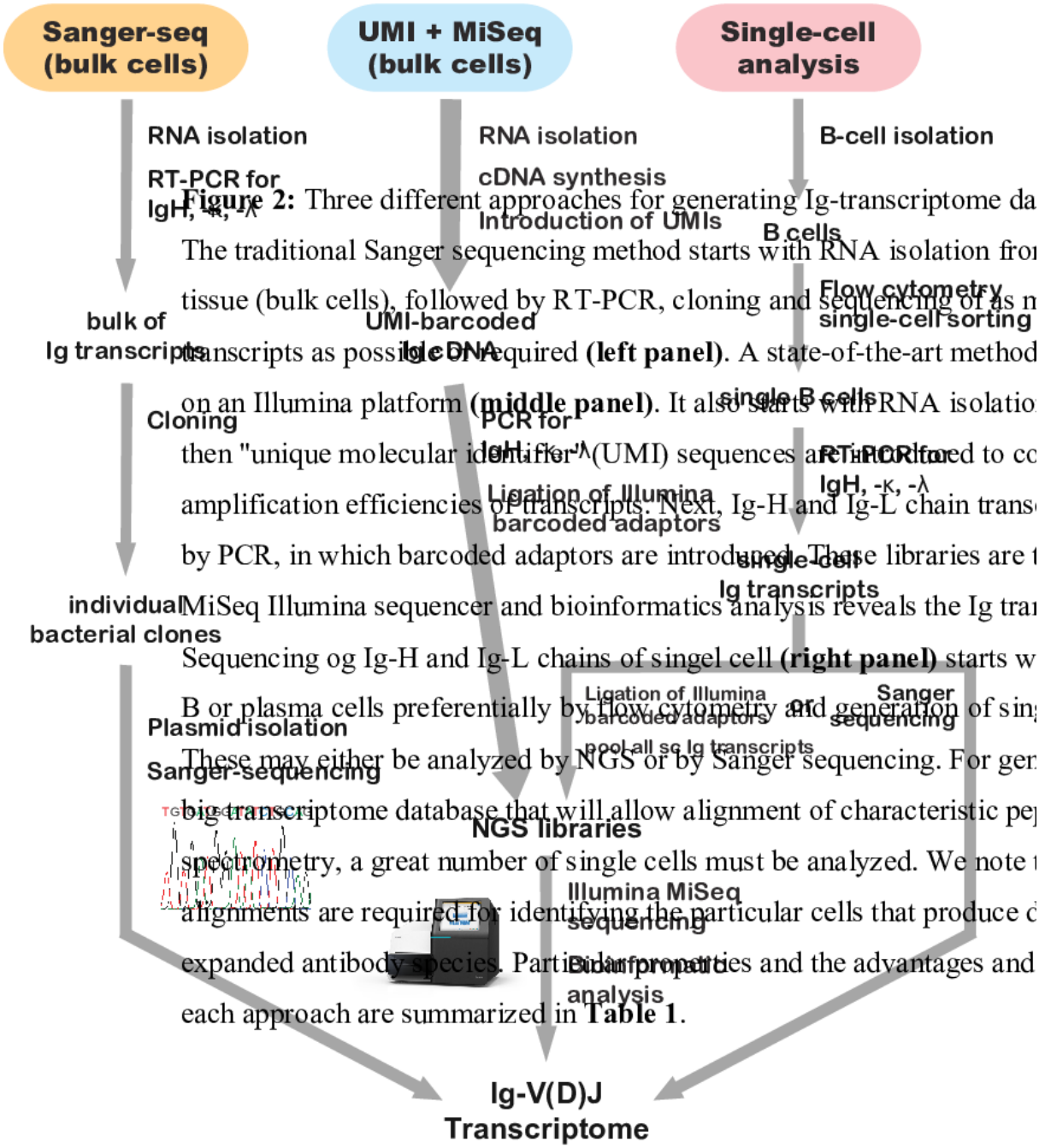

Figure 2: Three different approaches for generating Ig-transcriptome da

The traditional Sanger sequencing method starts with RNA isolation fro tissue (bulk cells), followed by RT-PCR, cloning and sequencing of as n transcripts as possible are required **(left panel)**. A state-of-the-art method on an Illumina platform **(middle panel)**. It also starts with RNA isolatio then "unique molecular identifier" (UMI) sequences are introduced to cc amplification efficiencies of transcripts. Next, Ig-H and Ig-L chain trans by PCR, in which barcoded adaptors are introduced. These libraries are t MiSeq Illumina sequencer and bioinformatics analysis reveals the Ig tra Sequencing og Ig-H and Ig-L chains of singel cell **(right panel)** starts w B or plasma cells preferentially by flow cytometry and generation of sin These may either be analyzed by NGS or by Sanger sequencing. For ger big transcriptome database that will allow alignment of characteristic pe spectrometry, a great number of single cells must be analyzed. We note alignments are required for identifying the particular cells that produce d expanded antibody species. Particular properties and the advantages and each approach are summarized in **Table 1**.

Figure 2

Figure 2

Three different approaches for generating Ig-transcriptome databases

Figure 3: Analysis of Ig-H and Ig-L chains by conventional as compared to improved gel electrophoresis and mass spectrometry, using monoclonal antibody 145.2c11 as an example. **a:** Conventional 1-dimensional IEF gel with sample loaded at pH 10. This results in thirteen distinct bands between pH 5 and 7 (photo taken from ref. 10). **b:** Improved 2-dimensional ge electrophoresis (see steps 1 through 4 in **Fig. 1**). For the first dimension an IEF gel was run with deglycosylated monoclonal antibody 145.2c11 in the range of pH 4 and 10 with sample loaded at pH 4.5. The second dimension is a non-reducing SDS-PAGE. We indicate molecular masses and the positions of H2L2 heterodimers and of complexes where one or bo light chains were disassembled (H2L and H2). Most spots were detected between pH 7 and 8 (see insert). Free light chains were detected between 20 and 30 kDa. 6 single spots of H2L2, H2L, and H2 complexes were excised as indicated in the insert, the proteins were digested with trypsin and then analyzed by mass spectrometry. **c:** Amino Acid sequences of the Ig-H and Ig-L V-Regions of monoclonal antibody 145.2c11. Peptides identified by mass spectrometry are numbered and highlighted in blue and green for H- and L-chain, respectively. **d:** Overview of peptides recovered from the spots indicated in **Fig. 3b**. Five peptides from the Ig-H and four from the Ig-L chains were detected in spots 1 to 4 comprised of H2L2 and H2L complexes (+). Peptides from the Ig-L chains were not detected in spots 5 and 6 (-), only peptides from the Ig-H chains (+), thus indicating the presence of H2 complexes.

Figure 3

Analysis of Ig-H and Ig-L chains by conventional as compared to improved gel electrophoresis and mass spectrometry, using monoclonal antibody 145.2c11 as an example

	Bulk cells		Single cell analysis	
	Sanger-seq	UMI + MiSeq	MiSeq	Sanger-seq
Sequencing target	Individual bacterial clones	mRNA/cDNA libraries of the entire Ig repertoire	mRNA/cDNA amplicons of H+L chains of single B cells	mRNA/cDNA amplicons of H+L chain of only one single B cell
Multiplexing	Not possible.	Whole Ig repertoire from many samples. UMI provides quantitative information.	H+L sequences of many single cells from 1 sample	Not possible.
Sample preparation	Simple but highly time-consuming. Can be done in a basic lab.	Complex protocol including many steps. Lab must have access to a MiSeq sequencer.	Complex protocol including many steps. Access to single-cell sorter and MiSeq sequencer required.	Many steps. Access to single-cell sorter required.
Sequence data analysis	Simple	High bioinformatics knowledge and equipment required.	High bioinformatics knowledge and equipment required.	Simple
PROs	Low sequencing error rate. High accuracy is achieved due to many combined reads of each nucleotide.	Millions of Ig mRNA libraries sequenced per run. Very low cost per sample. Use of UMI corrects for PCR and sequencing error.	Information of paired H+L chains of all sorted single B cells.	Information of the paired H+L chains from one sorted single B cell.
CONs	Only 1 sequence sequenced at a time per run. High cost per sample.	Overrepresentation of reads coming from plasma cells with high Ig mRNA synthesis vs. naïve or memory B cells.	Limited capacity. Incomplete B cell repertoire due to loss of many cells during sorting and mRNA/cDNA	As CONs of single cell analysis by MiSeq. In addition very time-consuming and high cost per sample.

Figure 4

Table 1 Overview on the properties of the different approaches to generate Ig transcriptome databases

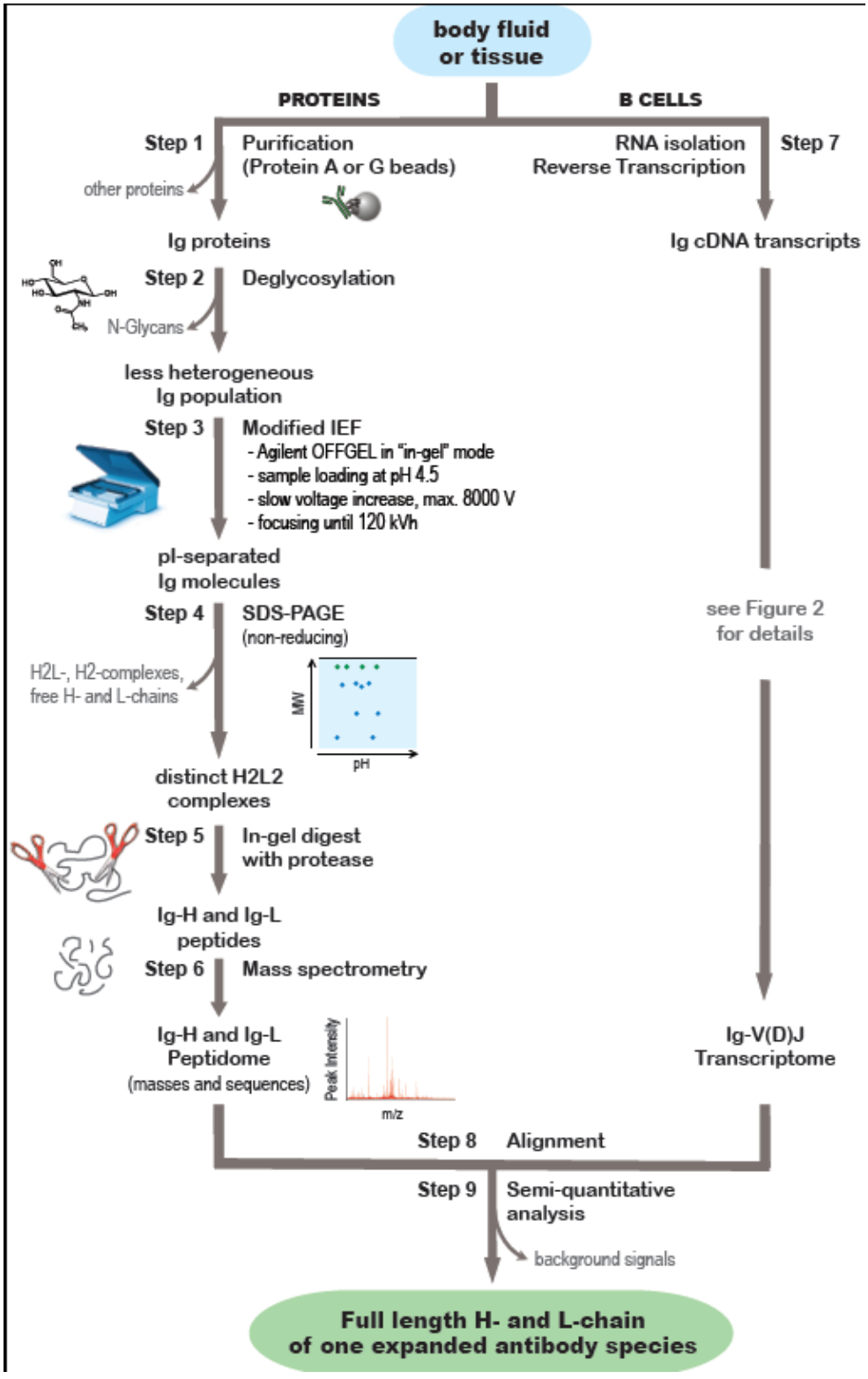

Figure 5

Figure 1 Flowchart for the characterization of individual antibodies

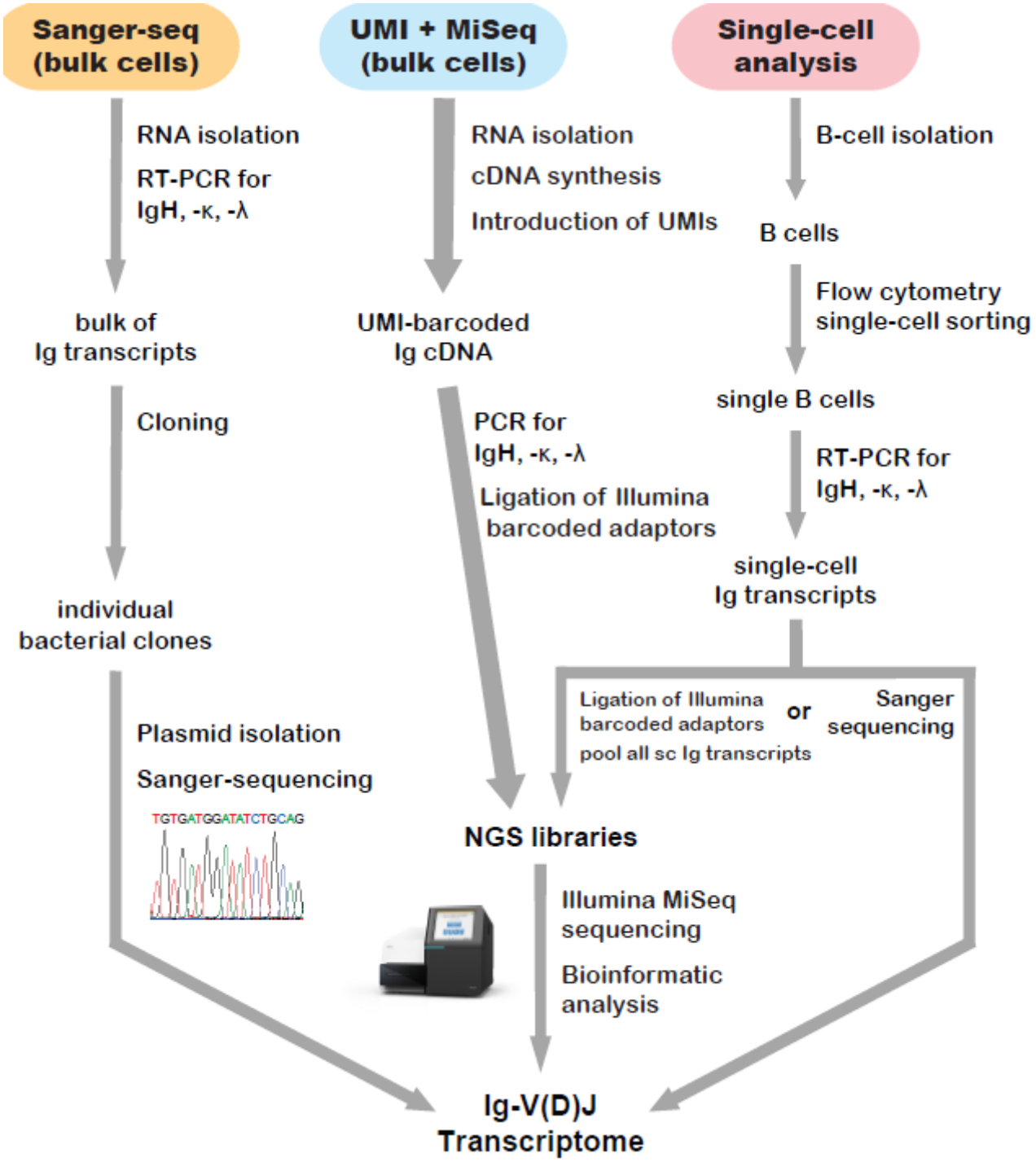

Figure 6

Figure 2 Three different approaches for generating Ig-transcriptome databases

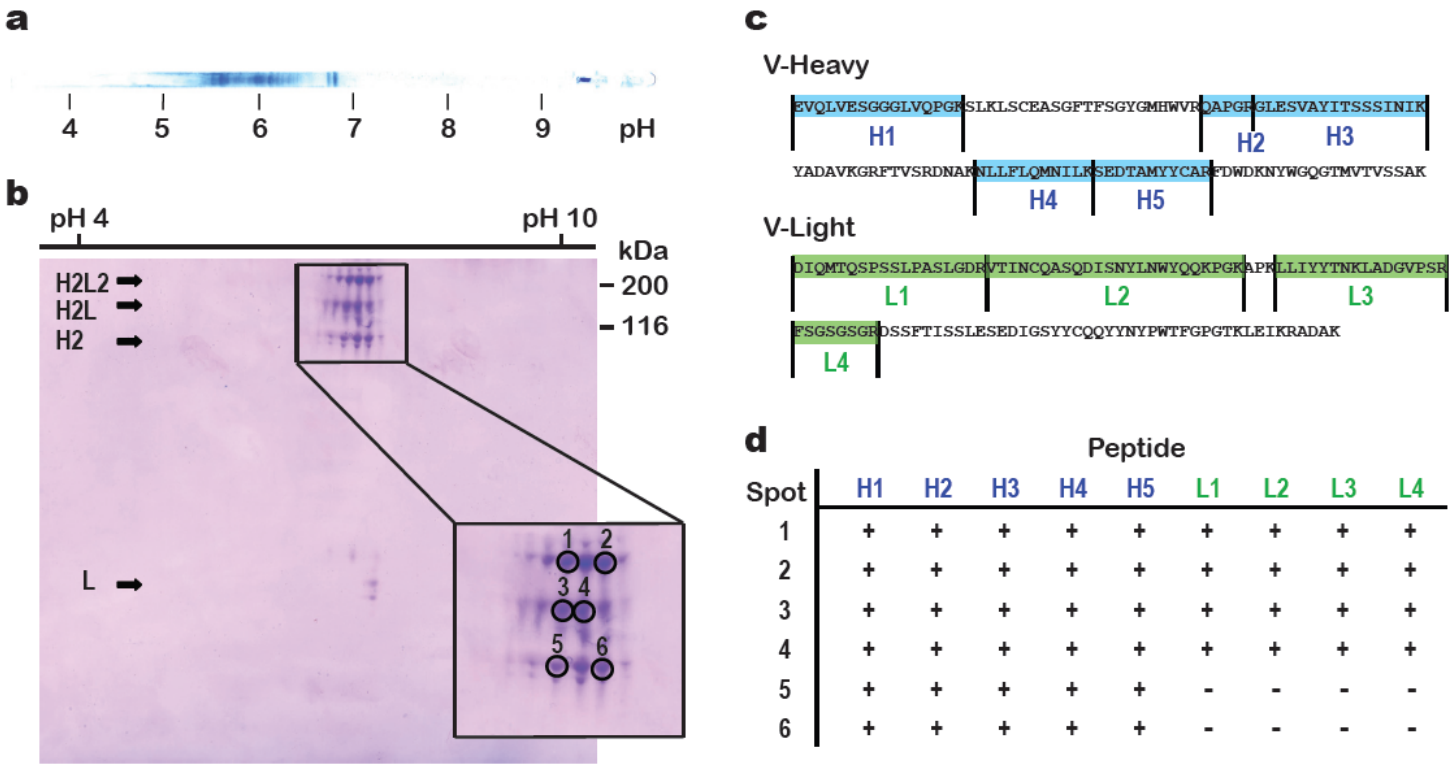

Figure 7

Figure 3 Analysis of Ig-H and Ig-L chains by conventional as compared to improved gel electrophoresis and mass spectrometry, using monoclonal antibody 145.2c11 as an example

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplement0.pdf