

# Title: The B Cell Repertoire in Multiple Sclerosis Reveals Molecular Mimicry between EBV EBNA1 and GlialCAM

**Authors:** Tobias V. Lanz<sup>1,2</sup>, R. Camille Brewer<sup>1</sup>, Peggy P. Ho<sup>3</sup>, Kevin M. Jude<sup>4</sup>, Daniel Fernandez<sup>5</sup>, Ricardo A. Fernandes<sup>4</sup>, Alejandro Gomez<sup>1</sup>, Gabriel Nadj<sup>1</sup>, Ryan D. Schubert<sup>6</sup>, Bianca Teegen<sup>7</sup>, Jeffrey E. Dunn<sup>8</sup>, Christopher B. Lock<sup>8</sup>, Lucas B. Kipp<sup>8</sup>, Victoria C. Cotham<sup>9</sup>, Beatrix M. Ueberheide<sup>9</sup>, Blake T. Aftab<sup>10</sup>, Michael R. Wilson<sup>6</sup>, Rachael J.M. Bashford-Rogers<sup>11</sup>, Michael Platten<sup>2,12</sup>, Raymond A. Sobel<sup>13</sup>, Christopher K. Garcia<sup>4</sup>, Lawrence Steinman<sup>3</sup>, William H. Robinson<sup>1,\*</sup>

## Affiliations:

<sup>1</sup>Division of Immunology and Rheumatology, Department of Medicine, Stanford University School of Medicine, 269 Campus Drive, Stanford, CA 94305, United States, and the VA Palo Alto Health Care System, 3801 Miranda Ave, Palo Alto, CA 94304, United States

<sup>2</sup>Department of Neurology, Medical Faculty Mannheim, University of Heidelberg, Theodor-Kutzer-Ufer 1-3, D-68167 Mannheim, Germany

<sup>3</sup>Department of Neurology and Neurological Sciences, Stanford University, Beckman Center for Molecular Medicine, 279 Campus Drive, Stanford, CA 94305, United States

<sup>4</sup>Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA 94305, United States

<sup>5</sup>Department of Chemistry, Engineering, & Medicine for Human Health, 290 Jane Stanford Way, Stanford, CA 94305, United States

<sup>6</sup>Department of Neurology, UCSF Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA 94158, United States

<sup>7</sup>Institute of Experimental Immunology, Euroimmun AG, 23560 Lübeck, Germany

<sup>8</sup>Division of Neuroimmunology, Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, United States

<sup>9</sup>Proteomics Resource Facility, Department of Pathology, University of Michigan, Ann Arbor, MI 48109, USA

<sup>10</sup>Preclinical Science and Translational Medicine, Atara Biotherapeutics, South San Francisco, CA, USA

<sup>11</sup>Wellcome Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

<sup>12</sup>DKTK Clinical Cooperation Unit Neuroimmunology and Brain Tumor Immunology, German Cancer Research Center (DKFZ), Heidelberg, Germany

<sup>13</sup>Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, and the VA Palo Alto Health Care System, 3801 Miranda Ave, Palo Alto, CA 94304, United States

\*Corresponding Author: William H. Robinson, Division of Immunology and Rheumatology, Department of Medicine, Stanford University School of Medicine, 269 Campus Drive, Stanford, CA 94305, United States, email: wrobins@stanford.edu

**Abstract (122 words):** In multiple sclerosis (MS) intrathecal B lymphocytes contribute to inflammation and secrete oligoclonal immunoglobulins. Rates of Epstein-Barr-Virus (EBV) infection are increased in MS but its role in MS pathobiology is unclear. Here we performed B cell repertoire sequencing of MS blood and cerebrospinal fluid (CSF) which revealed hallmarks of ongoing antigen-specific clonal expansion and intrathecal somatic hypermutation in CSF plasmablasts. Selected CSF-derived antibodies were tested against viruses associated with MS. Antibody MS39p2w174 binds Epstein-Barr-Virus (EBV) Nuclear Antigen 1 (EBNA1) and cross-reacts with the glial cellular adhesion molecule GlialCAM, facilitated by serine-phosphorylation. EBNA1 peptide immunization exacerbates the mouse model of MS. EBNA1/GlialCAM cross-reactive antibodies are prevalent in MS patients. Our results suggest that anti-EBNA1-antibodies cross-react with CNS GlialCAM to contribute to MS pathology.

**One Sentence Summary (114 Characters):** Clinically relevant molecular mimicry between EBV EBNA1 and CNS GlialCAM was identified by sequencing the CSF B cell repertoire in MS.

**Main Text:** In Multiple Sclerosis (MS), autoreactive B and T cells cause tissue-specific destruction of myelin in the central nervous system (CNS). Viral infections have long been theorized to contribute to MS by eliciting molecular mimicry to myelin proteins. Epidemiological data provides strong evidence for a contribution of Epstein Barr Virus (EBV) to MS pathology, with close to 100% of MS patients testing positive for anti-EBV antibodies vs. 93.6%-94.4% of healthy individuals (1–3). Infectious mononucleosis and MS share the same major risk allele HLA-DRB1\*15:01 and infectious mononucleosis as well as serum immunoglobulin reactivity against EBV nuclear antigen 1 (EBNA1) are independent and synergistic risk factors for MS (4). Immunodominant MS-associated epitopes of EBNA1 have been identified (5–9), and candidates for molecular mimicry have been proposed (8), but little is known beyond the descriptive identification of EBV/CNS cross-reactivities in CSF and sera of MS patients, and a mechanistic link between EBV infection and MS pathobiology has not been established.

In MS, the presence of oligoclonal bands (OCB) in cerebrospinal fluid (CSF) and successful clinical trials with B cell depleting therapies have emphasized the importance of B cells in MS pathology (10–12). B cells secrete autoantibodies and modulate T cell responses during antigen presentation (13). Many aspects of B cell immunology in MS are not well understood, including their phenotypic and functional characteristics in the inflamed CNS (14–16), and their antigen-specificity (17).

Here we show that a large fraction of B cells in the CSF are activated plasmablasts (PB), distinct from PB in peripheral blood with regard to activation and trafficking markers. To understand their antigen-specificity, we sequenced the single-cell paired-chain B cell receptor (BCR) / antibody repertoire of 1,689 B cells from CSF and 13,578 PB from blood of 9 MS patients and found substantially higher clonality and skewed immunoglobulin heavy chain V gene (IGHV gene) usage in the CSF, indicative of ongoing intrathecal somatic hypermutation and antigen-specific proliferation. 148 BCR sequences, representative of large clonal expansions, were expressed as recombinant monoclonal antibodies (mAbs) and tested for reactivity to viral proteins and peptides, including those from EBV. The CSF-derived mAb MS39p2w174 was discovered to bind EBNA1 within a region previously described to be associated with higher serum reactivity in MS patients (AA386-405) (5–7, 9). MS39p2w174 cross-reacts to the glial cell adhesion molecule GlialCAM, a type 1 membrane protein expressed on oligodendrocytes and astrocytes, in particular at the

astrocytic perivascular endfeet that maintain blood brain barrier integrity (18–20). GlialCAM aids the physiologic expression of aquaporin-4, an important B cell antigen in neuromyelitis optica (21). In addition, it enables cell-cell contacts by homo-oligomerization and is indispensable for glial chloride and water homeostasis as a beta-subunit of the MLC1 (22) and an auxiliary subunit of CLC2 (23). We mapped the EBNA1 and GlialCAM epitopes in detail, including a 2.5Å crystal structure of the mAb / EBNA1 peptide complex. We show that the unmutated germline sequence of MS39p2w174 is already prone to bind EBNA1, while additional mutations increase affinity to GlialCAM. The interaction with GlialCAM is facilitated by a serine phosphorylation N-terminal of the central epitope. We observed elevated anti-GlialCAM serum reactivity in MS patients in comparison to healthy individuals. Finally, we show that similar antibodies against GlialCAM are generated in mice upon immunization with EBNA1 AA386-405, which aggravates the mouse model of MS, experimental autoimmune encephalomyelitis (EAE).

## **IgG<sup>+</sup> plasmablasts dominate the CSF B cell compartment in MS patients**

CSF and blood samples were obtained from 9 MS patients during the initial onset of disease (clinically isolated syndrome, CIS, n=5) or an acute episode of relapsing-remitting MS (RRMS, n=4). All patients had a pleocytosis of >10 cells /  $\mu$ l CSF (table S1). Single B cells were sorted by flow cytometry. Approximately 30% of B cells in the CSF exhibit an activated plasmablast (PB) phenotype (CD19<sup>+</sup> CD20<sup>low</sup> CD27<sup>+</sup> CD38<sup>+</sup>), while only a small fraction of B cells in the blood are PB (~4%) (fig. 1A and fig. S1A,B). Conversely, naïve B cells are diminished in the CSF, while unswitched and switched memory as well as double negative B cell numbers are comparable in blood and CSF (fig. S1C). Profound phenotypic differences between blood and CSF were detected in PB but not in non-PB B cells: (i) Blood PB express high levels of the trafficking receptor  $\alpha$ 4 integrin, whereas PB in the CSF abate  $\alpha$ 4 expression over time (fig. 1B, fig. S1D). (ii) PB in the CSF express higher levels of HLA-DR than their blood-derived counterparts (fig. 1C,D), emphasizing their role in antigen presentation in the CSF and CNS. Similar differences were not seen in non-PB B cells. Of note, HLA-DR expression was independent of the HLA-DRB1\*15:01 genotype (HLA-DR15) (table S2 and fig. S1E,F). (iii) The predominant immunoglobulin (Ig) class within the PB compartment in CSF is IgG, while both IgG and IgA are the main classes in blood-derived PB (fig. 1E). In contrast, non-PB B cells express similar Ig-classes in CSF and blood (fig. 1F). These differences expand on similar findings in previous reports (24, 25) and suggest an elevated pathogenic role of PB in MS.

## **The B cell repertoire in CSF PB is highly clonal and skewed**

To gain a comprehensive overview of the intrathecal antigen-specific B cell response in MS, we sorted single PB from blood and single B cells from CSF of MS patients by flow cytometry and sequenced full-length paired heavy-chain (HC) and light-chain (LC) VDJ regions using our in-house plate-based single-cell sequencing technology (26). 13,578 paired sequences from blood PB and 1,689 from CSF B cells passed filter thresholds. In comparison to the repertoire of blood PB, the repertoire of CSF PB is significantly more clonal and largely dominated by IgG (fig. 1G, H), suggesting antigen-specific proliferation of just a few clones within the CSF. Of note, the CSF repertoire in acute viral encephalitis did not resemble a similar amount of clonal BCR sequences (not shown). Not surprisingly, non-PB B cells in the CSF are less clonal and express

IgA and IgM more frequently than PB (fig. S2A). While mutation counts in IGHV and IGLV genes did not differ significantly between PB in blood and CSF (fig. S2B), HC-CDR3 lengths are on average 0.94 amino acids longer in CSF PB, indicating ongoing intrathecal somatic hypermutation (fig. S2C). The repertoire in the CSF is skewed towards more preferential usage of five IGHV chains: IGHV4-59, IGHV4-39, IGHV4-34, IGHV1-2, and IGHV3-7 (fig. 1I, fig. S2D-F), which for the most part is in line with previous reports (27–30) and indicates that a select group of MS-related antigens in the CSF drive PB survival and proliferation in the CSF, with a few distinct immunoglobulin germline genes being predestined to bind those antigens.

### **Clonal PB are the main source of oligoclonal bands**

To find out if the identified clonal PB are the source of intrathecal oligoclonal bands (OCB), we purified immunoglobulin from CSF samples and characterized them by mass spectrometry, searching the MS/MS spectra against the patient specific databases. For 87% of clonal sequences, peptides matching variable chain sequences unique to the patient could be identified, whereas only 40% of singleton sequences could be detected by mass spectrometry (fig. 1J). Highly abundant immunoglobulins, which were identified with ten or more peptide-spectral matches (PSM >10) likely correlate with OCB. These sequences aligned almost exclusively to clonally expanded B cell sequences (fig. 1K), suggesting that clonally expanded B cells are the source of the OCB. This correlation holds true for PB, which are more clonal than non-PB B cells (fig. S2G,H). Taken together, clonal PB are likely the main source of antibodies and OCB in the CSF of MS patients.

### **CSF-derived monoclonal antibodies bind EBV antigens**

A total of 148 sequences from the CSF repertoires were selected for recombinant expression, each one representative of a major clonal expansion (fig. S3). To test anti-viral reactivities of the selected mAbs, they were probed on a planar protein microarray representing 2 EBV viral lysates, 23 recombinant latent and lytic EBV proteins, 240 peptides spanning four prominent EBV proteins, as well as 7 lysates of other MS-associated viruses, including measles, rubella, and varicella-zoster virus (VZV) (31) (fig. 2A,B, fig. S4A, S5, S6A, tables S3-S4). One-third of

the expressed mAbs bound to EBV proteins and peptides and ~20% to other viruses, in particular to VZV and cytomegalovirus (CMV) (fig. 2A). Of note, half of the VZV-reactive antibodies cross-reacted to CMV and EBV, indicative of broader antigens common to herpes viruses.

Interestingly, we found mAbs in 6 out of 9 patients that bound to the EBV transcription factor EBNA1 (fig. 2A, fig. S4A), and mAbs binding to EBNA1 peptides in 8 out of 9 patients (fig. 2B, fig. S5). Anti-EBNA1-reactivity has been implicated in MS pathogenesis (32) and the region AA365-425 (“MS-associated epitope” in fig. 2B and fig. S6A) is known to elicit a stronger antibody response in MS patients than in healthy individuals (5–7, 9). Protein and peptide arrays revealed that our mAb MS39p2w174 binds EBNA1 within this region (AA386-405, fig. 2B, fig. S5). The interaction was verified by western blot analysis using full-length and truncated EBNA1 proteins (fig. 2C) and ELISA-based peptide scans spanning full-length EBNA1 (20mer peptides, 13AA overlap, fig. 2D). Alanine-scanning determined the proline-rich region AA394-399 as the central epitope (fig. 2E, fig. 6B). Taken together, we identified multiple mAbs directed against EBV including mAb MS39p2w174, which binds a well-described MS-associated epitope of EBNA1.

### Crystal structure reveals key residues of mAb-EBNA1 interaction

While the presence of antibodies against the broader EBNA1 region AA365-425 is well established in MS patients (5–7, 9), their relevance to MS pathology has remained elusive. Efforts to model its structure have remained hypothetical and did not further our understanding of the epitope’s characteristics and functions (9). To understand its immunogenicity and impact on MS pathology in detail, we solved the crystal structure of EBNA1 AA386-405 in complex with MS39p2w174 at a resolution of 2.5 Å (fig. 2F-J, fig. S6C-E, PDB ID: 7K7R). It confirmed close interactions of the peptide residues P394-P398 with all complementary determining regions (CDRs) but the very short LCDR2. Residues Tyr31 and Tyr38 on LCDR1 together with Trp38 on HCDR1 and Pro108, Pro109, and Tyr114 on HCDR3 create a hydrophobic cage for the peptide’s first two prolines Pro394 and Pro395 and the proximal side chain of Arg396 (fig. 2H,I). The C-terminal end of the antibody binding groove is wider and Pro398 is carried by a large aromatic tryptophan residue (Trp114 in HCDR1) on the bottom of the groove (fig. 2G,H). The central arginines Arg395 and Arg396 engage in close polar interactions (<3.1 Å) with residues



on HCDR2, HCDR3, and HC framework region 2. Contrary to the results of our alanine scan (fig. 2E, fig. S6B), Pro399 does not appear to interact directly with antibody side chains, and we assume that alanine at position 399 disrupts the conformation of the three prolines Pro398-Pro400 causing steric hindrance within the binding pocket.

The encoding IGHV gene of MS39p2w174 is IGHV3-7, one of the IGHV chains over-represented in CSF (fig. 1I). Interestingly, all but one of the residues that directly interact with EBNA1 are unmutated germline (GL) residues (IGHV3-7, IGHJ4, IGKV2-30, IGKJ1) (fig. S6F,G). We therefore hypothesized that the unmutated ancestor of MS39p2w174 might have an inert propensity to bind EBNA1 AA386-405. Indeed, we could show that GL binds to EBNA1 with only slightly lower affinity than MS39p2w174 ( $K_D$  MS39p2w174: 1.99 nM, GL: 4.19 nM) (fig. 2K,L).

### **Molecular mimicry between EBNA1 AA386-405 and GlialCAM**

Studying a mAb as opposed to patient-derived sera and CSF samples allows for direct identification of molecular mimicry with human proteins. We probed mAb MS39p2w174 on a HuProt protein microarray, which represents >16,000 proteins spanning >80% of the human proteome (33). Glial cell adhesion molecule (GlialCAM) was identified as the top binding partner to MS39p2w174 (fig. 3A). GlialCAM is a cell adhesion molecule that is almost exclusively expressed in the CNS (fig. S7A, [www.proteinatlas.org](http://www.proteinatlas.org)) (34), mainly in astrocytes and oligodendrocytes (18–20). In multiple sclerosis, it has been found to be decreased in acute and chronic MS plaques, but elevated in chronic-active plaques (Acc #: Q8N7I3) (35). GlialCAM AA337-385 was also identified as a binding partner for MS39p2w174 on a 49mer phage display representing the whole human proteome (36) (356 out of  $10^5$  reads, MS39p2w174 was the only GlialCAM-binding mAb in a set of 300 mAbs). Binding of MS39p2w174 to the intracellular domain (ICD, AA262-416) of GlialCAM was further confirmed by ELISA (fig. 3B) and western blot (fig. 3C). Affinity measurements with bio-layer interferometry revealed higher affinity of MS39p2w174 to GlialCAM ( $K_D$ : 190 pM) vs. EBNA1 ( $K_D$ : 1.99 nM). This is in contrast to the unmutated GL mAb, which binds GlialCAM with lower affinity ( $K_D$  GlialCAM: 10.46 nM,  $K_D$  EBNA1: 4.19 nM) (fig. 2K,L and fig. 3D,E). Evidently, while GL harbors a



propensity to bind to EBNA1, somatic hypermutation of the mature MS39p2w174 B cell increased its affinity to the CNS mimic GlialCAM by 2 orders of magnitude.

### **Phosphorylation at GlialCAM Ser376 enables MS39p2w174 binding**

The EBNA1 epitope AA386-405 is located between the protein's long N-terminal Gly-Ala-rich low-complexity region (AA90-380) and its highly structured DNA-binding domain (AA: 461-607, PDB: 1B3T). The above-mentioned GlialCAM region AA337-385 is located at the C-terminal end of the ICD and contains a proline-rich region that closely resembles the central epitope of EBNA1 (fig. 3F). MS39p2w174 detects both proteins on western blots under denaturing conditions (fig. 2C, fig. 3C), suggesting linear epitopes for both targets. This is in line with predictions that both epitopes are located in intrinsically disordered regions of the respective protein (Fig 3G, H). However, while MS39p2w174 binds the EBNA1 peptide AA386-405 with high affinity ( $K_D$ : 2.67 nM), its affinity to GlialCAM peptide AA370-389 is drastically lower ( $K_D$ : 302 nM). As the intracellular domain of GlialCAM is heavily phosphorylated (37–39) (fig. S7B, phosphosite.org (40)), and post-translational modifications often determine antibody-antigen interactions (41), we tested if phosphorylation at one of the 4 serine residues surrounding the central epitope region (residues Ser376, 377, 383, and 384) could increase binding affinity of MS39p2w174 to GlialCAM AA370-389. Indeed, phosphorylation at Ser376 facilitates MS39p2w174 interaction with the peptide ( $K_D$ : 6.1 nM) and additional phosphorylation of Ser377 further enhances binding affinity ( $K_D$ : 3.73 nM) (fig. 3I-K). In contrast, citrullination of arginine residues Arg373, 380, and 387 did not alter peptide binding to MS39p2w174 (fig. S7C). The important residue Arg397 in EBNA1 AA386-405, which engages in 2 hydrogen-bonds with Glu64 at HCDR2 (fig. 2 H,J) is replaced with alanine in GlialCAM AA370-389 (Ala381) (fig. 3F), which explains the decreased binding affinity between MS39p2w174 and GlialCAM peptide. Phosphorylation at position 376 likely enables binding by adding new polar interactions to the proximal LC, possibly with Arg36, a positively charged residue that is mutated from asparagine in GL (fig. 2H, fig. S6G). Taken together, we show that a post-translational phosphorylation enables cross-binding of MS39p2w174 to GlialCAM.

### **Anti-GlialCAM IgG reactivity is elevated in MS patients**

To test if the observed anti-GlialCAM reactivity of MS39p2w174 is part of a broader phenomenon, we tested the remaining 147 mAbs for reactivity against GlialCAM protein and the broader region AA315-395. We found 10 additional mAbs that bound the ICD and 7 that bound the extracellular domain (ECD) (fig. 3L). Two mAbs bound unphosphorylated GlialCAM AA370-389, interestingly, both also cross-reacted with EBNA1 as well as with the two early lytic EBV proteins BHRF1 and BLLF3 (fig. 3L). This shows that anti-GlialCAM reactivity is not an isolated phenomenon. Though we did not identify any additional mAbs in our collection with the exact same characteristics as MS39p2w174, antibodies against several GlialCAM epitopes are prevalent in the majority of patients.

We proceeded with testing a larger set of plasma samples from 36 MS patients and 20 healthy individuals for reactivity to EBNA1 and GlialCAM. As expected, all MS patients and the majority of healthy individuals showed plasma reactivity to EBNA1 protein (fig. 3M). Reactivity to EBNA1 AA386-405 and to GlialCAM was significantly higher in MS patients (fig. 3N,O), with the three most highly reactive samples overlapping between EBNA1 AA386-405 and GlialCAM (MS16, MS30, MS49). Reactivity against EBNA1 AA386-405 could be abrogated by blocking with GlialCAM, substantiating the presence of cross-reactive antibodies in the plasma of this patient (MS30, fig. 3P). This data shows that molecular mimicry between EBNA1 AA386-405 and GlialCAM AA370-389 is prevalent in a subset of MS patients.

### **Immunization with EBNA1 AA386-405 aggravates EAE**

To assess the effect of an anti-EBNA1 AA386-405 response on neuroinflammation, we utilized the mouse model EAE. SJL mice were immunized with either scrambled control peptide (SPSRPGRSRSGSPFPQPSP, not binding to MS39p2w174) or EBNA1 AA386-405. Three weeks later, EAE was induced by a second immunization of the same respective peptides mixed with PLP AA139-151. Mice in the EBNA1 group generated a robust antibody response to both EBNA1 AA386-405 (fig. 4A) as well as GlialCAM ICD (fig. 4B). The antibody response to PLP AA139-151 was unaltered in both groups (fig. 4C). Clinically, the EBNA1 group showed more severe symptoms of paresis, most pronounced during the initial peak of disease and the subsequent relapses (fig. 4D). EBNA1 AA386-405 immunization promoted infiltration of inflammatory cells into the CNS (fig. 4E,F), and enhanced demyelination (fig. 4G,H). In addition

to the B cell response, EBNA1 AA386-405 induced a strong T cell response while the PLP AA139-151 specific T cell response remained comparable in both groups (fig. S8A). EBNA1 AA386-405 stimulated the secretion of B cell stimulatory Th1 cytokines IFN- $\gamma$ , TNF, IL-12 as well as IL-6 and IL-10 (fig. S8B-F) but suppressed the key Th17 cytokine IL-17 (fig. S8G). No robust T cell response against GlialCAM AA369-388 pSer375 could be detected (fig. S8A-G). Taken together, we can show that antibodies against GlialCAM are generated in response to immunization with EBNA1 AA386-405, mimicking human cross-specificity. Co-immunization with EBNA1 AA386-405 and PLP in the EAE-model enhanced central-nervous immune cell infiltration, and demyelination, two prominent features of human MS pathology.

## Discussion

Viral triggers of MS and other autoimmune diseases have long been the subject of intense research efforts, but evidence for their functional relevance is scarce (8). By utilizing paired-chain BCR repertoire sequencing, rational selection of clonal antibody sequences, and high-throughput proteomics, we identified a monoclonal antibody from the CSF of an MS patient that binds the well-described MS associated epitope AA386-405 on EBV EBNA1 with high affinity and found that it cross-reacts with the myelin glial cellular adhesion molecule GlialCAM. We demonstrate the presence of cross-reactive EBV EBNA1 and GlialCAM antibodies in a broader population of MS patients and that immunization of mice with this EBV EBNA1 epitope exacerbates autoimmune demyelination. Our findings demonstrate a mechanistic link between EBV infection and the pathobiology of MS.

Structure and binding data of MS CSF plasmablast mAb MS39p2w174 and its unmutated germline ancestor (GL) show that GL only binds EBNA1 AA486-405 with high affinity and somatic hypermutation of the VDJ region is required for a highly affine interaction with GlialCAM. A precursor of MS39p2w174 likely entered the CNS / CSF space and matured there towards increased anti-GlialCAM affinity, increasing its impact on neuroinflammation. This is in line with our analysis of the BCR repertoire, which shows a highly clonal and skewed repertoire in CSF, indicative of a select group of B cells undergoing somatic hypermutation, likely driven by perpetual activation by CNS antigens.

We show that phosphorylation of GlialCAM at Ser376 is necessary for the close interaction with MS39p2w174 by generating new polar interactions at the C-terminal end of the binding pocket. Epitopes altered by post-translational modifications (PTMs) have been described as targets for autoimmune responses, e.g. citrullinated proteins in rheumatoid arthritis and phosphorylated Ro/La in systemic lupus erythematosus (41–43). PTMs are highly variable and often tissue-specific and could therefore differ between CNS tissue and lymphatic organs, resulting in a lack of central tolerance against post-translationally modified proteins. Several kinases have been described as risk genes for MS, including MERTK, MAPK1 and 3, TYK2, which could contribute to alternative phosphorylation patterns in the CNS (44, 45).

We show that plasma IgG titers against EBNA1 AA386-405 and GlialCAM are higher in MS patients in comparison to healthy controls and demonstrate cross-reactivity. In addition, we were

able to show in *in-vivo* experiments that immunization with EBNA1 AA386-405 generated a strong humoral immune response against GlialCAM, stimulated T cells to secrete Th1 cytokines and B cell activating cytokines including IL6 and IL10, and aggravated EAE. Antigen-presentation and immunomodulation of T cells is a large part of the immunogenic role of B cells in MS. Cross-talk between B cells and follicular T helper cells (Tfh) in the CNS as well as in meningeal ectopic lymphoid follicles seems to have an outsized impact on neuroinflammation (15, 46). Our study does not yet provide a clear delineation of the impact of anti-GlialCAM antibodies vs. GlialCAM-specific cellular immunity on the course of the disease. However, our description of high numbers of activated PB in the CSF, expressing exceptionally high levels of HLA-DR, suggest that they present antigens and exchange inflammatory signals with Tfh.

Our large analysis of the single-cell paired-chain BCR repertoire from blood and CSF of MS patients reveals features of intrathecal expansion of a few clones and ongoing somatic hypermutation, strikingly different from BCR repertoires in patients with viral encephalitis. PB in the CSF of MS patients are highly clonal with longer CDR3 regions and skewing towards preferential use of IGHV4-39, IGHV4-34, IGHV4-59, IGHV1-2, and IGHV3-7. A dominance of IGHV4 sequences, and in particular IGHV4-34 and IGHV4-39 has been observed in autoimmune disorders including MS (27–30, 47, 48), but their relevance is not well understood. IGHV3-7 has been found repeatedly in CSF BCR repertoires and oligoclonal bands of MS patients (28, 30, 48–50), but selected CSF-derived mAbs from this family has not been studied in detail so far.

In summary, our study advances our understanding of MS pathobiology on several levels: (i) We present a detailed picture of phenotypic differences and the single-cell paired-chain B cell repertoire in MS CSF and blood, which emphasize an important inflammatory role of PB and show signs of ongoing intrathecal affinity maturation. (ii) Using selected CSF-derived mAbs, we show broad reactivity against viral antigens and against EBV antigens in particular. (iii) A detailed analysis of a CSF-derived mAb against EBNA1 AA386-405 showed molecular mimicry to GlialCAM, facilitated by phosphorylation at serine 376. (iv) Cross-reactive antibodies against the two antigens are elevated in MS patients. (v) In-vivo EAE experiments showed that immunization with EBNA1 AA386-405 generated anti-GlialCAM antibodies in mice and aggravated the disease. Together, our findings establish a mechanistic link based on molecular mimicry between EBV infection and the pathobiology of MS.

## References and Notes:

1. E. L. Thacker, F. Mirzaei, A. Ascherio, Infectious mononucleosis and risk for multiple sclerosis: A meta-analysis. *Ann. Neurol.* **59**, 499–503 (2006).
2. A. Ascherio, K. L. Munger, Environmental risk factors for multiple sclerosis. Part I: the role of infection. *Ann. Neurol.* **61**, 288–299 (2007).
3. S. Abrahamyan, B. Eberspächer, M.-M. Hoshi, L. Aly, F. Luessi, S. Groppa, L. Klotz, S. G. Meuth, C. Schroeder, T. Gräter, B. Tackenberg, F. Paul, F. Then-Bergh, T. Kümpfel, F. Weber, M. Stangel, A. Bayas, B. T. Wildemann, C. Heesen, U. Zettl, C. Warnke, G. Antony, N. Hessler, H. Wiendl, S. Bittner, B. Hemmer, R. Gold, A. Salmen, K. Ruprecht, *J. Neurol. Neurosurg. Psychiatry*, in press (available at <http://jnnp.bmj.com/lookup/doi/10.1136/jnnp-2020-322941>).
4. A. K. Hedström, J. Huang, A. Michel, J. Butt, N. Brenner, J. Hillert, T. Waterboer, I. Kockum, T. Olsson, L. Alfredsson, High Levels of Epstein-Barr Virus Nuclear Antigen-1-Specific Antibodies and Infectious Mononucleosis Act Both Independently and Synergistically to Increase Multiple Sclerosis Risk. *Front. Neurol.* **10**, 1368 (2019).
5. K. Ruprecht, B. Wunderlich, R. Gieß, P. Meyer, M. Loebel, K. Lenz, J. Hofmann, B. Rosche, O. Wengert, F. Paul, U. Reimer, C. Scheibenbogen, Multiple sclerosis: The elevated antibody response to Epstein–Barr virus primarily targets, but is not confined to, the glycine–alanine repeat of Epstein–Barr nuclear antigen-1. *J. Neuroimmunol.* **272**, 56–61 (2014).
6. N. Jafari, G. P. van Nierop, G. M. G. M. Verjans, A. D. M. E. Osterhaus, J. M. Middeldorp, R. Q. Hintzen, No evidence for intrathecal IgG synthesis to Epstein Barr virus nuclear antigen-1 in multiple sclerosis. *J. Clin. Virol.* **49**, 26–31 (2010).
7. J. Salzer, M. Nyström, G. Hallmans, H. Stenlund, G. Wadell, P. Sundström, Epstein-Barr virus antibodies and vitamin D in prospective multiple sclerosis biobank samples. *Mult. Scler.* **19**, 1587–1591 (2013).
8. K. Tengvall, J. Huang, C. Hellström, P. Kammer, M. Biström, B. Ayoglu, I. Lima Bomfim, P. Stridh, J. Butt, N. Brenner, A. Michel, K. Lundberg, L. Padyukov, I. E. Lundberg, E. Svenungsson, I. Ernberg, S. Olafsson, A. T. Dilthey, J. Hillert, L. Alfredsson, P. Sundström, P. Nilsson, T. Waterboer, T. Olsson, I. Kockum, Molecular mimicry between Anoctamin 2 and Epstein-Barr virus nuclear antigen 1 associates with multiple sclerosis risk. *Proc. Natl. Acad. Sci. U. S. A.* **5**, 201902623–201916960 (2019).
9. Z. Wang, P. G. Kennedy, C. Dupree, M. Wang, C. Lee, T. Pointon, T. D. Langford, M. W. Graner, X. Yu, Antibodies from Multiple Sclerosis Brain Identified Epstein-Barr Virus Nuclear Antigen 1 & 2 Epitopes which Are Recognized by Oligoclonal Bands. *J. Neuroimmune Pharmacol.* (2020), doi:10.1007/s11481-020-09948-1.
10. K. Hawker, P. O'Connor, M. S. Freedman, P. A. Calabresi, J. Antel, J. Simon, S. L. Hauser, E. Waubant, T. Vollmer, H. Panitch, J. Zhang, P. Chin, C. H. Smith, O. trial group, Rituximab in patients with primary progressive multiple sclerosis: results of a randomized double-blind placebo-controlled multicenter trial. *Ann. Neurol.* **66**, 460–471 (2009).



11. L. Kappos, D. Li, P. A. Calabresi, P. O'Connor, A. Bar-Or, F. Barkhof, M. Yin, D. Leppert, R. Glanzman, J. Tinbergen, S. L. Hauser, Ocrelizumab in relapsing-remitting multiple sclerosis: a phase 2, randomised, placebo-controlled, multicentre trial. *Lancet*. **378**, 1779–1787 (2011).
12. X. Montalban, S. L. Hauser, L. Kappos, D. L. Arnold, A. Bar-Or, G. Comi, J. de Seze, G. Giovannoni, H.-P. Hartung, B. Hemmer, F. Lublin, K. W. Rammohan, K. Selmaj, A. Traboulsee, A. Sauter, D. Masterman, P. Fontoura, S. Belachew, H. Garren, N. Mairon, P. Chin, J. S. Wolinsky, O. C. Investigators, Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. *N. Engl. J. Med.* **376**, 209–220 (2017).
13. A. Negron, R. R. Robinson, O. Stüve, T. G. Forsthuber, The role of B cells in multiple sclerosis: Current and future therapies. *Cell. Immunol.* **339**, 10–23 (2019).
14. T. V. Lanz, A.-K. Pröbstel, I. Mildenerger, M. Platten, L. Schirmer, Single-Cell High-Throughput Technologies in Cerebrospinal Fluid Research and Diagnostics. *Front. Immunol.* **10**, 1302 (2019).
15. D. Schafflick, C. A. Xu, M. Hartlehnert, M. Cole, T. Lautwein, A. Schulte-Mecklenbeck, J. Wolbert, M. Heming, S. G. Meuth, T. Kuhlmann, C. C. Gross, H. Wiendl, N. Yosef, G. M. zu Horste, Integrated single cell analysis of blood and cerebrospinal fluid leukocytes in multiple sclerosis. **192**, 2551.
16. A. Wang, O. Rojas, D. Lee, J. L. Gommerman, Regulation of neuroinflammation by B cells and plasma cells. *Immunol. Rev.*, 1–16 (2020).
17. R. Hohlfeld, K. Dornmair, E. Meinl, H. Wekerle, The search for the target antigens of multiple sclerosis, part 2: CD8<sup>+</sup> T cells, B cells, and antibodies in the focus of reverse-translational research. *Lancet Neurol.* (2016), , doi:10.1016/S1474-4422(15)00313-0.
18. A. Gilbert, X. E. Vidal, R. Estevez, M. Cohen-Salmon, A.-C. Boulay, Postnatal development of the astrocyte perivascular MLC1/GlialCAM complex defines a temporal window for the gliovascular unit maturation. *Brain Struct. Funct.* **7**, 12–41 (2019).
19. L. Favre-Kontula, A. Rolland, L. Bernasconi, M. Karmirantzou, C. Power, B. Antonsson, U. Boschert, GlialCAM, an immunoglobulin-like cell adhesion molecule is expressed in glial cells of the central nervous system. *Glia*. **56**, 633–645 (2008).
20. T. López-Hernández, M. C. Ridder, M. Montolio, X. Capdevila-Nortes, E. Polder, S. Sirisi, A. Duarri, U. Schulte, B. Fakler, V. Nunes, G. C. Scheper, A. Martínez, R. Estevez, M. S. van der Knaap, Mutant GlialCAM causes megalencephalic leukoencephalopathy with subcortical cysts, benign familial macrocephaly, and macrocephaly with retardation and autism. *Am. J. Hum. Genet.* **88**, 422–432 (2011).
21. M. Bugiani, M. Dubey, M. Breur, N. L. Postma, M. P. Dekker, T. Ter Braak, U. Boschert, T. E. M. Abbink, H. D. Mansvelder, R. Min, J. R. T. van Weering, M. S. van der Knaap, Megalencephalic leukoencephalopathy with cysts: the Glialcam-null mouse model. *Ann. Clin. Transl. Neurol.* **4**, 450–465 (2017).
22. T. López-Hernández, S. Sirisi, X. Capdevila-Nortes, M. Montolio, V. Fernández-Dueñas, G. C. Scheper, M. S. van der Knaap, P. Casquero, F. Ciruela, I. Ferrer, V. Nunes, R. Estevez, Molecular mechanisms of MLC1 and GLIALCAM mutations in megalencephalic



- leukoencephalopathy with subcortical cysts. *Hum. Mol. Genet.* **20**, 3266–3277 (2011).
23. E. Jeworutzki, T. López-Hernández, X. Capdevila-Nortes, S. Sirisi, L. Bengtsson, M. Montolio, G. Zifarelli, T. Arnedo, C. S. Müller, U. Schulte, V. Nunes, A. Martínez, T. J. Jentsch, X. Gasull, M. Pusch, R. Estevez, GlialCAM, a Protein Defective in a Leukodystrophy, Serves as a CIC-2 Cl<sup>−</sup> Channel Auxiliary Subunit. *Neuron*. **73**, 951–961 (2012).
  24. M. C. Kowarik, V. Grummel, S. Wemlinger, D. Buck, M. S. Weber, A. Berthele, B. Hemmer, Immune cell subtyping in the cerebrospinal fluid of patients with neurological diseases. *J. Neurol.* **261**, 130–143 (2014).
  25. A. Pröbstel, X. Zhou, R. Baumann, S. Wischnewski, M. Kutza, O. L. Rojas, K. Sellrie, A. Bischof, K. Kim, A. Ramesh, R. Dandekar, A. L. Greenfield, R. D. Schubert, J. E. Bisanz, S. Vistnes, K. Khaleghi, J. Landefeld, G. Kirkish, F. Liesche-starnecker, V. Ramaglia, S. Singh, E. B. Tran, P. Barba, K. Zorn, J. Oechtering, K. Forsberg, L. R. Shiow, R. G. Henry, J. Graves, B. A. C. Cree, S. L. Hauser, J. Kuhle, J. M. Gelfand, P. M. Andersen, J. Schlegel, P. J. Turnbaugh, P. H. Seeberger, J. L. Gommerman, M. R. Wilson, L. Schirmer, S. E. Baranzini, Gut microbiota – specific IgA + B cells traffic to the CNS in active multiple sclerosis. **7191** (2020).
  26. Y.-C. Tan, S. Kongpachith, L. K. Blum, C.-H. Ju, L. J. Lahey, D. R. Lu, X. Cai, C. A. Wagner, T. M. Lindstrom, J. Sokolove, W. H. Robinson, Barcode-enabled sequencing of plasmablast antibody repertoires in rheumatoid arthritis. *Arthritis Rheumatol. (Hoboken, N.J.)*. **66**, 2706–2715 (2014).
  27. A. Palanichamy, L. Apeltsin, T. C. Kuo, M. Sirota, S. Wang, S. J. Pitts, P. D. Sundar, D. Telman, L. Z. Zhao, M. Derstine, A. Abounasr, S. L. Hauser, H.-C. von Büdingen, Immunoglobulin class-switched B cells form an active immune axis between CNS and periphery in multiple sclerosis. *Sci. Transl. Med.* **6**, 248ra106-248ra106 (2014).
  28. H.-C. von Büdingen, T. C. Kuo, M. Sirota, C. J. van Belle, L. Apeltsin, J. Glanville, B. A. Cree, P.-A. Gourraud, A. Schwartzburg, G. Huerta, D. Telman, P. D. Sundar, T. Casey, D. R. Cox, S. L. Hauser, B cell exchange across the blood-brain barrier in multiple sclerosis. *J. Clin. Invest.* **122**, 4533–4543 (2012).
  29. W. H. Rounds, E. A. Salinas, T. B. Wilks, M. K. Levin, A. J. Ligocki, C. Ionete, C. A. Pardo, S. Vernino, B. M. Greenberg, D. W. Bigwood, E. M. Eastman, L. G. Cowell, N. L. Monson, MSPrecise: A molecular diagnostic test for multiple sclerosis using next generation sequencing. *Gene*. **572**, 191–197 (2015).
  30. J. Bankoti, L. Apeltsin, S. L. Hauser, S. Allen, M. E. Albertolle, H. E. Witkowska, H. C. Von Büdingen, In multiple sclerosis, oligoclonal bands connect to peripheral B-cell responses. *Ann. Neurol.* **75**, 266–276 (2014).
  31. S. Jarius, P. Eichhorn, D. Franciotta, H.-F. Petereit, G. Akman-Demir, M. Wick, B. T. Wildemann, The MRZ reaction as a highly specific marker of multiple sclerosis: re-evaluation and structured review of the literature. *J. Neurol.* **264**, 453–466 (2017).
  32. J. Y. Mescheriakova, G. P. van Nierop, A. A. van der Eijk, K. L. Kreft, R. Q. Hintzen, EBNA-1 titer gradient in families with multiple sclerosis indicates a genetic contribution.

- Neurol. Neuroimmunol. neuroinflammation.* **7**, 1–8 (2020).
33. J. S. Jeong, L. Jiang, E. Albino, J. Marrero, H. S. Rho, J. Hu, S. Hu, C. Vera, D. Bayron-Poueymiroy, Z. A. Rivera-Pacheco, L. Ramos, C. Torres-Castro, J. Qian, J. Bonaventura, J. D. Boeke, W. Y. Yap, I. Pino, D. J. Eichinger, H. Zhu, S. Blackshaw, Rapid identification of monospecific monoclonal antibodies using a human proteome microarray. *Mol. Cell. Proteomics* (2012), doi:10.1074/mcp.O111.016253.
  34. M. Uhlén, L. Fagerberg, B. M. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, Å. Sivertsson, C. Kampf, E. Sjöstedt, A. Asplund, I. Olsson, K. Edlund, E. Lundberg, S. Navani, C. A.-K. Szigartyo, J. Odeberg, D. Djureinovic, J. O. Takanen, S. Hober, T. Alm, P.-H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J. M. Schwenk, M. Hamsten, K. von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. von Heijne, J. Nielsen, F. Pontén, Proteomics. Tissue-based map of the human proteome. *Science*. **347**, 1260419 (2015).
  35. M. H. Han, S.-I. Hwang, D. B. Roy, D. H. Lundgren, J. V Price, S. S. Ousman, G. H. Fernald, B. Gerlitz, W. H. Robinson, S. E. Baranzini, B. W. Grinnell, C. S. Raine, R. A. Sobel, D. K. Han, L. Steinman, Proteomic analysis of active multiple sclerosis lesions reveals therapeutic targets. *Nature*. **451**, 1076–1081 (2008).
  36. B. O'Donovan, C. Mandel-Brehm, S. E. Vazquez, J. Liu, A. V Parent, M. S. Anderson, T. Kassimatis, A. Zekeridou, S. L. Hauser, S. J. Pittock, E. Chow, M. R. Wilson, J. L. DeRisi, High resolution epitope mapping of anti-Hu and anti-Yo autoimmunity by programmable phage display. *Brain Commun.* (2020), doi:10.1093/braincomms/fcaa059.
  37. M. C. Moh, C. Zhang, C. Luo, L. H. Lee, S. Shen, Structural and Functional Analyses of a Novel Ig-like Cell Adhesion Molecule, hepaCAM, in the Human Breast Carcinoma MCF7 Cells. *J. Biol. Chem.* **280**, 27366–27374 (2005).
  38. D. Britton, Y. Zen, A. Quaglia, S. Selzer, V. Mitra, C. Löbner, S. Jung, G. Böhm, P. Schmid, P. Prefot, C. Hoehle, S. Koncarevic, J. Gee, R. Nicholson, M. Ward, L. Castellano, J. Stebbing, H. D. Zucht, D. Sarker, N. Heaton, I. Pike, Quantification of pancreatic cancer proteome and phosphorylome: Indicates molecular events likely contributing to cancer and activity of drug targets. *PLoS One*. **9** (2014), doi:10.1371/journal.pone.0090948.
  39. J. H. Herskowitz, N. T. Seyfried, D. M. Duong, Q. Xia, H. D. Rees, M. Gearing, J. Peng, J. J. Lah, A. I. Levey, Phosphoproteomic analysis reveals site-specific changes in GFAP and NDRG2 phosphorylation in frontotemporal lobar degeneration. *J. Proteome Res.* **9**, 6368–6379 (2010).
  40. P. V. Hornbeck, I. Chabra, J. M. Kornhauser, E. Skrzypek, B. Zhang, PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics*. **4**, 1551–1561 (2004).
  41. M. G. Zavala-Cerna, E. A. Martínez-García, O. Torres-Bugarín, B. Rubio-Jurado, C. Riebeling, A. Nava, The clinical significance of posttranslational modification of autoantigens. *Clin. Rev. Allergy Immunol.* **47**, 73–90 (2014).
  42. N. Lingampalli, J. Sokolove, L. J. Lahey, J. D. Edison, W. R. Gilliland, V. M. Holers, K.

D. Deane, W. H. Robinson, Combination of anti-citrullinated protein antibodies and rheumatoid factor is associated with increased systemic inflammatory mediators and more rapid progression from preclinical to clinical rheumatoid arthritis. *Clin. Immunol.* **195**, 119–126 (2018).

43. A. G. Terzoglou, J. G. Routsias, S. Avrameas, H. M. Moutsopoulos, A. G. Tzioufas, Preferential recognition of the phosphorylated major linear B-cell epitope of La/SSB 349-368aa by anti-La/SSB autoantibodies from patients with systemic autoimmune diseases. *Clin. Exp. Immunol.* **144**, 432–439 (2006).
44. I. M. S. G. Consortium, W. T. C. C. C. 2, S. J. Sawcer, G. Hellenthal, M. Pirinen, C. C. A. Spencer, N. A. Patsopoulos, L. Moutsianas, A. Dilthey, Z. Su, C. Freeman, S. E. Hunt, S. Edkins, E. Gray, D. R. Booth, S. C. Potter, A. Goris, G. Band, A. B. Oturai, A. Strange, J. Saarela, C. Bellenguez, B. Fontaine, M. Gillman, B. Hemmer, R. Gwilliam, F. Zipp, A. Jayakumar, R. Martin, S. Leslie, S. Hawkins, E. Giannoulidou, S. D'Alfonso, H. Blackburn, F. Martinelli Boneschi, J. Liddle, H. F. Harbo, M. L. Perez, A. Spurkland, M. J. Waller, M. P. Mycko, M. Ricketts, M. Comabella, N. Hammond, I. Kockum, O. T. McCann, M. Ban, P. Whittaker, A. Kempainen, P. Weston, C. Hawkins, S. Widaa, J. Zajicek, S. Dronov, N. Robertson, S. J. Bumpstead, L. F. Barcellos, R. Ravindrarajah, R. Abraham, L. Alfredsson, K. Ardlie, C. Aubin, A. Baker, K. Baker, S. E. Baranzini, L. Bergamaschi, R. Bergamaschi, A. Bernstein, A. Berthele, M. Boggild, J. P. Bradfield, D. Brassat, S. A. Broadley, D. Buck, H. Butzkueven, R. Capra, W. M. Carroll, P. Cavalla, E. G. Celius, S. Cepok, R. Chiavacci, F. Clerget-Darpoux, K. Clysters, G. Comi, M. Cossburn, I. Cournu-Rebeix, M. B. Cox, W. Cozen, B. A. C. Cree, A. H. Cross, D. Cusi, M. J. Daly, E. Davis, P. I. W. de Bakker, M. Debouverie, M. B. D'hooghe, K. Dixon, R. Dobosi, B. Dubois, D. Ellinghaus, I. Elovaaara, F. Esposito, C. Fontenille, S. Foote, A. Franke, D. Galimberti, A. Ghezzi, J. Glessner, R. Gomez, O. Gout, C. Graham, S. F. A. Grant, F. R. Guerini, H. Hakonarson, P. Hall, A. Hamsten, H.-P. Hartung, R. N. Heard, S. Heath, J. Hobart, M. Hoshi, C. Infante-Duarte, G. Ingram, W. Ingram, T. Islam, M. Jagodic, M. Kabesch, A. G. Kermode, T. J. Kilpatrick, C. Kim, N. Klopp, K. Koivisto, M. Larsson, M. Lathrop, J. S. Lechner-Scott, M. A. Leone, V. Leppä, U. Liljedahl, I. L. Bomfim, R. R. Lincoln, J. Link, J. Liu, A. R. Lorentzen, S. Lupoli, F. Macciardi, T. Mack, M. Marriott, V. Martinelli, D. Mason, J. L. McCauley, F. Mentch, I.-L. Mero, T. Mihalova, X. Montalban, J. Mottershead, K.-M. Myhr, P. Naldi, W. Ollier, A. Page, A. Palotie, J. Pelletier, L. Piccio, T. Pickersgill, F. Piehl, S. Pobywajlo, H. L. Quach, P. P. Ramsay, M. Reunanen, R. Reynolds, J. D. Rioux, M. Rodegher, S. Roesner, J. P. Rubio, I.-M. Rückert, M. Salvetti, E. Salvi, A. Santaniello, C. A. Schaefer, S. Schreiber, C. Schulze, R. J. Scott, F. Sellebjerg, K. W. Selmaj, D. Sexton, L. Shen, B. Simms-Acuna, S. Skidmore, P. M. A. Sleiman, C. Smestad, P. S. Sorensen, H. B. Söndergaard, J. Stankovich, R. C. Strange, A.-M. Sulonen, E. Sundqvist, A.-C. Syvänen, F. Taddeo, B. Taylor, J. M. Blackwell, P. Tienari, E. Bramon, A. Tourbah, M. A. Brown, E. Tronczynska, J. P. Casas, N. Tubridy, A. Corvin, J. Vickery, J. Jankowski, P. Villoslada, H. S. Markus, K. Wang, C. G. Mathew, J. Wason, C. N. A. Palmer, H.-E. Wichmann, R. Plomin, E. Willoughby, A. Rautanen, J. Winkelmann, M. Wittig, R. C. Trembath, J. Yaouanq, A. C. Viswanathan, H. Zhang, N. W. Wood, R. Zuvich, P. Deloukas, C. Langford, A. Duncanson, J. R. Oksenberg, M. A. Pericak-Vance, J. L. Haines, T. Olsson, J. Hillert, A. J. Iverson, P. L. De Jager, L. Peltonen, G. J. Stewart, D. A. Hafler, S. L.

- Hauser, G. McVean, P. Donnelly, A. Compston, Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*. **476**, 214–219 (2011).
45. C. Gasperi, T. F. M. Andlauer, A. Keating, B. Knier, A. Klein, V. Pernpeintner, P. Lichtner, R. Gold, F. Zipp, F. Then Bergh, M. Stangel, H. Tumani, B. Wildemann, H. Wiendl, A. Bayas, T. Kümpfel, U. K. Zettl, R. A. Linker, U. Ziemann, M. Knop, C. Warnke, M. A. Friese, F. Paul, B. Tackenberg, A. Berthele, B. Hemmer, Genetic determinants of the humoral immune response in MS. *Neurol. Neuroimmunol. neuroinflammation*. **7** (2020), doi:10.1212/NXI.0000000000000827.
46. L. Bell, A. Lenhart, A. Rosenwald, C. M. Monoranu, F. Berberich-Siebelt, Lymphoid Aggregates in the CNS of Progressive Multiple Sclerosis Patients Lack Regulatory T Cells. *Front. Immunol.* **10**, 1–18 (2020).
47. R. J. M. Bashford-Rogers, K. G. C. Smith, D. C. Thomas, Antibody repertoire analysis in polygenic autoimmune diseases. *Immunology*. **33**, 197 (2018).
48. B. Obermeier, R. Mentele, J. Malotka, J. Kellermann, T. Kümpfel, H. Wekerle, F. Lottspeich, R. Hohlfeld, K. Dornmair, Matching of oligoclonal immunoglobulin transcriptomes and proteomes of cerebrospinal fluid in multiple sclerosis. *Nat. Med.* **14**, 688–693 (2008).
49. V. Singh, M. P. Stoop, C. Stingl, R. L. Luitwieler, L. J. Dekker, M. M. Van Duijn, K. L. Kreft, T. M. Luider, R. Q. Hintzen, Cerebrospinal-fluid-derived immunoglobulin G of different multiple sclerosis patients shares mutated sequences in complementarity determining regions. *Mol. Cell. Proteomics* (2013), doi:10.1074/mcp.M113.030346.
50. M. Colombo, M. Dono, P. Gazzola, S. Roncella, A. Valetto, N. Chiorazzi, G. L. Mancardi, M. Ferrarini, Accumulation of Clonally Related B Lymphocytes in the Cerebrospinal Fluid of Multiple Sclerosis Patients. *J. Immunol.* **164**, 2782–2789 (2000).
51. A. J. Thompson, B. L. Banwell, F. Barkhof, W. M. Carroll, T. Coetzee, G. Comi, J. Correale, F. Fazekas, M. Filippi, M. S. Freedman, K. Fujihara, S. L. Galetta, H.-P. Hartung, L. Kappos, F. D. Lublin, R. A. Marrie, A. E. Miller, D. H. Miller, X. Montalban, E. M. Mowry, P. S. Sorensen, M. Tintore, A. L. Traboulsee, M. Trojano, B. M. J. Uitdehaag, S. Vukusic, E. Waubant, B. G. Weinshenker, S. C. Reingold, J. A. Cohen, Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet. Neurol.* **17**, 162–173 (2018).
52. Y.-C. Tan, L. K. Blum, S. Kongpachith, C.-H. Ju, X. Cai, T. M. Lindstrom, J. Sokolove, W. H. Robinson, High-throughput sequencing of natively paired antibody chains provides evidence for original antigenic sin shaping the antibody response to influenza vaccination. *Clin. Immunol.* **151**, 55–65 (2014).
53. L. K. Blum, R. R. L. Cao, A. J. Sweatt, M. Bill, L. J. Lahey, A. C. Hsi, C. S. Lee, S. Kongpachith, C.-H. Ju, R. Mao, H. H. Wong, M. R. Nicolls, R. T. Zamanian, W. H. Robinson, Circulating plasmablasts are elevated and produce pathogenic anti-endothelial cell autoantibodies in idiopathic pulmonary arterial hypertension. *Eur. J. Immunol.* (2018) (available at <http://doi.wiley.com/10.1002/eji.201747460>).
54. R. C. Edgar, UPARSE: Highly accurate OTU sequences from microbial amplicon reads.

*Nat. Methods* (2013), doi:10.1038/nmeth.2604.

55. E. Alamyar, V. Giudicelli, S. Li, P. Duroux, M. P. L. Immunome, 2012, IMGT/HighV-QUEST: the IMGT® web portal for immunoglobulin (IG) or antibody and T cell receptor (TR) analysis from NGS high throughput and deep .... *imgt.igh.cnrs.fr* (available at [http://imgt.igh.cnrs.fr/PDF/ImmunomeRes/8\\_1-15\\_2012.pdf](http://imgt.igh.cnrs.fr/PDF/ImmunomeRes/8_1-15_2012.pdf)).
56. R. C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* (2004), doi:10.1093/nar/gkh340.
57. S. Guindon, J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
58. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).
59. M. Bern, Y. J. Kil, C. Becker, Byonic: advanced peptide and protein identification software, Curr Protoc Bioinformatics. *Curr Protoc Bioinforma. Chapter 13*, Unit13 (2012).
60. W. H. Robinson, C. DiGennaro, W. Hueber, B. B. Haab, M. Kamachi, E. J. Dean, S. Fournel, D. Fong, M. C. Genovese, H. E. N. de Vegvar, K. Skriver, D. L. Hirschberg, R. I. Morris, S. Muller, G. J. Pruijn, W. J. van Venrooij, J. S. Smolen, P. O. Brown, L. Steinman, P. J. Utz, Autoantigen microarrays for multiplex characterization of autoantibody responses. *Nat. Med.* **8**, 295–301 (2002).
61. S. Kuerten, T. V. Lanz, N. Lingampalli, L. J. Lahey, C. Kleinschnitz, M. Mäurer, M. Schroeter, S. Braune, T. Ziemssen, P. P. Ho, W. H. Robinson, L. Steinman, Autoantibodies against central nervous system antigens in a subset of B cell–dominant multiple sclerosis patients. *Proc. Natl. Acad. Sci.*, 202011249 (2020).
62. V. G. Tusher, R. J. Tibshirani, G. Chu, Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci.* **98**, 5116–5121 (2001).
63. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
64. G. Tickle, I. J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonnrhein, C., Bricogne, STARANISO. *Glob. Phasing Ltd. Cambridge*, (2018).
65. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
66. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
67. D. Liebschner, P. V Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. D, Struct. Biol.* **75**, 861–877



(2019).

68. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 352–367 (2012).
69. L. Schrodinger, The PyMOL molecular graphics system (2010) (available at [http://http://scholar.google.comjavascript:void\(0\)\)](http://http://scholar.google.comjavascript:void(0))).
70. T. V. Lanz, S. K. Williams, A. Stojic, S. Iwantscheff, J. K. Sonner, C. Grabitz, S. Becker, L.-I. Böhrer, S. R. Mohapatra, F. Sahm, G. Küblbeck, T. Nakamura, H. Funakoshi, C. A. Opitz, W. Wick, R. Diem, M. Platten, Tryptophan-2,3-Dioxygenase (TDO) deficiency is associated with subclinical neuroprotection in a mouse model of multiple sclerosis. *Sci. Rep.* **7**, 41271 (2017).
71. T. V. Lanz, C. A. Opitz, P. P.-K. Ho, A. Agrawal, C. Lutz, M. Weller, A. L. Mellor, L. Steinman, W. Wick, M. Platten, Mouse mesenchymal stem cells suppress antigen-specific TH cell immunity independent of indoleamine 2,3-dioxygenase 1 (IDO1). *Stem Cells Dev.* **19**, 657–668 (2010).

**Acknowledgments:** We thank Wolfgang Wick for support with sample collection, Ben Bell for advice on crystallization, and George Harauz for insightful discussions on poly-proline motifs.

**Funding:** This work was supported by NIH R01 AR063676 and U19 AI110491 to W.H.R., the Juvenile Diabetes Research Foundation and Lupus Research Alliance Funding to W.H.R. and T.V.L., the German Research Foundation (DFG, LA3657/1) to T.V.L., and Atara to L.S. and P.P.H.. The mass spectrometric experiments are in part supported by the NYU Grossman School of Medicine and a shared instrumentation grant from the NIH, 1S10OD010582-01A1, for the purchase of an Orbitrap Fusion Lumos. **Author contributions:** Conceptualization, T.V.L., L.S., W.H.R.; Methodology, T.V.L., R.C.B., P.P.H., K.M.J., D.F., R.A.F., A.G., R.D.S., B.T., V.C.C., B.M.U.; Software, T.V.L., R.C.B., D.F., K.M.J., V.C.C.; Validation, T.V.L., R.C.B., K.M.J., B.M.U., R.J.M.B.-R., R.A.S., C.K.G., L.S., W.H.R., Formal Analysis, T.V.L., R.C.B., K.M.J., D.F., A.G., V.C.C., B.M.U.; Investigation, T.V.L., R.C.B., P.P.H., D.F., G.N., R.D.S., B.T., V.C.C.; Resources, T.V.L., P.P.H., D.F., B.T., J.E.D., C.B.L., L.B.K., B.M.U., M.R.W., M.P., R.A.S., C.K.G., L.S., W.H.R.; Data Curation, T.V.L., R.C.B., K.M.J., D.F., V.C.C., B.M.U.; Writing – Original Draft, T.V.L.; Writing – Review & Editing, T.V.L., R.C.B., P.P.H., W.H.R.; Visualization, T.V.L., R.C.B., K.M.J.; Supervision, T.V.L., W.H.R.; Project Administration, T.V.L., W.H.R.; Funding Acquisition, T.V.L., P.P.H., B.T.A., L.S., W.H.R. **Competing interests:** W.H.R. owns equity in, serves as a consultant to, and is a member of the Board of Directors of Atreca, Inc. L.S. owns equity in and serves as a consultant to Atreca, Inc. (add competing interests: patents, financial holdings, professional affiliations, advisory positions, and board memberships). **Data and materials availability:** Genomic data: (GEO link). Mass spectrometry data: www.massive.ucsd.edu Acc #: MSV000086829 (dataset is not yet public. Reviewers can access it under Username: MSV000086829\_reviewer, Password: reviewersonly), structural data: www.rcsb.org PDB ID: 7K7R.



## Supplementary Materials:

### Materials and Methods

**Study design and human Subjects.** Patient samples were collected at the university hospitals at Stanford, CA, and Heidelberg, Germany. Relapsing remitting MS (RRMS) was diagnosed according to the 2017 McDonald criteria (51). All patients had elevated CSF white blood cell counts  $\geq 10$  cells/ $\mu$ l and blood contaminated CSF samples were excluded by visual and microscopic inspection. Peripheral blood and CSF were obtained at the time of clinical onset (CIS) or during an acute episode of the disease. All but one patient had not received any MS-specific treatment prior to sample collection (table S1). All experimental protocols were approved by the ethical review boards of Stanford University (IRB# 34529) and the University of Heidelberg (IRB# S-466/2015), respectively. Written informed consent was obtained from each patient.

**Cell preparation, antibody staining, and flow cytometric cell sorting.** CSF was centrifuged immediately after lumbar puncture and cells were counted. PBMCs were isolated from heparin blood by density gradient centrifugation using Ficoll PLUS media (Cytiva, Marlborough, MA). Cells were stained according to standard protocols with antibodies against the following cell surface markers: CD19 (clone HIB19), CD20 (clone L27), CD38 (clone HB7) (all BD Biosciences, Franklin Lakes, NJ), CD3 (clone OKT3), CD27 (clone O323), IgM (clone MHM-88), IgD (clone IA6-2), HLA-DR (clone L243),  $\alpha 4$  integrin (clone 9F10) (all BioLegend, San Diego, CA), IgA (clone IS11-8E10) (Miltenyi Biotec, Bergisch Gladbach, Germany), and Sytox blue (Thermo Fisher Scientific, Waltham, MA). Single cells were sorted with a FACS Aria II cell sorter (BD Biosciences) into 96-well PCR plates (BioRad, Hercules, CA). FlowJo Version 10.7.1 (BD) and R version 3.6.1 was used to evaluate flow cytometry data.

**Single-cell BCR repertoire sequencing.** BCR repertoire sequencing was carried out using our in-house developed plate-bound single-cell sequencing technology as described previously (26, 52, 53). Briefly, reverse transcription with oligo-dT was carried out in separate wells, attaching unique well-ID barcodes by template switching activity of Maxima Reverse Transcriptase (Thermo Scientific). Barcoded cDNA from each plate was pooled and amplified in 3 consecutive PCRs while attaching plate-specific barcodes and sequencing adapters. PCRs were carried out separately for HC of IgG, IgA, and IgM, as well as for  $\kappa$  for  $\lambda$  LC, and separate libraries were

generated from each, gel-purified, cleaned with Ampure XP beads (Beckman Coulter, Brea, CA), and sequenced on an Illumina MiSeq (Illumina, San Diego, CA) with 2 x 330 paired-end reads.

**Sequence analysis.** The MiSeq FASTQ workflow was used for Fastq generation and plate demultiplexing. R version 3.6.1 was used for custom analyses. Paired reads of sequences that passed quality thresholds were stitched and separated by plate and well IDs. Similar reads sharing the same plate and well IDs were clustered into operational taxonomic units (OTUs) (54). Consensus sequences were aligned to germline variable-chain immunoglobulin sequences with IMGT HighV-QUEST v1.3.1 (55), which reports V, D, and J germline genes, HC and LC CDR3-lengths, and non-silent mutation counts and locations. Clonal expansions were defined based on sharing the same HC and LC V and J genes and exhibiting >70% amino acid identity within the HC and LC CDR3s. Percent clonality represents the percent of all sequences that fulfill these criteria. To compare clonality between blood and CSF, the number of blood sequences were down-sampled at random to the sequence number of the corresponding CSF sample and means were calculated from 250 iterations. To calculate IGHV, IGLV, IGHJ, and IGLJ gene usages, percent abundance of each particular gene was calculated in blood and CSF PB of each patient and means were calculated across patients. Genes that were present in less than three CSF samples were excluded from this analysis. While our sequencing method preferentially captures PB sequences due to higher amounts of immunoglobulin mRNA (fig. S2I), enough non-PB B cell sequences passed filter thresholds to compare the non-PB repertoire to the PB repertoire in 7 patients (fig. S2A). Patient samples MS12, for which <5 non-PB sequences were sorted, and MS39, for which only PB were sorted, were excluded from this analysis. For phylogenetic analysis, sequences were binned according to their HC V-gene. Concatenated LC and HC were then aligned with Muscle (56) and clustered with PhyML (57) using maximum-likelihood clustering. Each tree-partition was rooted by their HC V-gene. Trees were drawn using iTol (58).

**Peptide identification with Mass Spectrometry.** Immunoglobulins were purified from 1.5ml of CSF samples with Protein A (ThermoFisher Scientific). The purified IgGs were reduced with 0.02 M dithiothreitol at 57°C for 1 hour, alkylated with 0.05 M iodoacetamide at room temperature (RT) in the dark, and digested with trypsin overnight at RT. Peptides were extracted

and desalted as previously described (59). An aliquot of the peptide mixtures was loaded onto an Acclaim PepMap 100 precolumn (75 $\mu$ m  $\times$  2cm, C18, 3 $\mu$ m, 100Å) in-line with an EASY-Spray, PepMap column (75 $\mu$ m  $\times$  50cm, C18, 2 $\mu$ m, 100Å) with a 5 $\mu$ m emitter using the autosampler of an EASY-nLC 1000 (Thermo Scientific). The peptides were gradient eluted into a Lumos Fusion Tribrid (Thermo Scientific) mass spectrometer using a 120min gradient from 5% to 35% solvent B (90% acetonitrile, 0.5% acetic acid), followed by 10 minutes from 35% to 45% solvent B and 10 min from 45 to 100% B. High resolution full MS spectra were acquired with a resolution of 120,000, an AGC target of 4e5, a maximum ion time of 50 ms, and scan range of 400 to 1800 m/z. Following each full MS scan as many data-dependent HCD MS/MS spectra were acquired in the orbitrap as possible in a 3 second cycle time. Monoisotopic precursor selection (MIPS) was set to peptide, precursors with a charge state of 2 – 7 and minimum intensity of 5e4 were selected for MS/MS. Dynamic exclusion was set to 60 seconds after a single selection. All MS/MS spectra were collected using the following instrument parameters: resolution of 30,000, an AGC target of 1e5, maximum ion time of 120 ms, two microscans, 1.6 m/z isolation window, and Normalized Collision Energy (NCE) of 32.

The MS/MS spectra were searched against the respective peptide specific database including common contaminant proteins using the search engine Byonic (59). The search parameters were set to trypsin allowing two missed cleavages, fixed modification of carbamidomethyl on cysteine, variable modification of oxidation on methionine and deamidation on glutamine and asparagine. Peptides mapping to variable regions of IgG were manually verified. In order to include only sequence-specific peptides, peptides that aligned to non-immunoglobulin or constant-region sequences were excluded from the analysis, as were peptides that aligned to the repertoire of multiple patients. Included were peptides that aligned to one variable sequence in a single patient. Peptides that aligned to more than one variable sequence in a single patient were included if all matching sequences were exact matches or clonally related, in which case the peptide was counted as representative for all matches. Counts of identical or non-identical peptide spectral matches (PSM) per sequence were tallied for each sequence. Sequences that had >1 or >10 matching peptides were presented as percentage of all sequences (fig. 1J,K, fig. S2G,H). The mass spectrometry files are accessible at MassIVE ([massive.ucsd.edu](http://massive.ucsd.edu)) under accession number

MSV000086829 (dataset is not yet public. Reviewers can access it under Username: MSV000086829\_reviewer, Password: reviewersonly).

**Selection and recombinant expression of mAbs.** Representative antibodies of the largest clonal B cell expansions from CSF were chosen for recombinant expression. In patients with more than 10 large clonal expansions, sequences were preferentially chosen based on their usage of one of the 11 most abundant IGHV genes in the CSF (fig. S3). HC and LC variable sequences were custom generated (IDT, Coralville, IA, USA), and cloned into pFuse vectors (Invivogen, San Diego, CA, USA), containing human IgG constant region or kappa or lambda constant regions, respectively. Fab HC were expressed in in-house plasmids, containing the constant-region C1 up to Cys103. Plasmids were transfected into Expi293T cells using Expifectamine (ThermoFisher Scientific). Medium was harvested after 4 and 7 days of culture. mAbs and Fabs were purified with protein A and protein G resins, respectively (ThermoFisher Scientific). Antibody concentrations were measured with a nanodrop spectrophotometer (ThermoFisher Scientific) and hIgG quantitation ELISAs (Bethyl Laboratories, Inc, Montgomery, TX, USA) and checked for purity on SDS protein gels with Coomassie staining.

**Protein expression and purification.** Different versions of EBNA1 were obtained from the following sources: full-length AA1-641 (Abcam, Cambridge, MA, USA), AA328-641 (Virion Serion, Würzburg, Germany), AA408-641 (ProspecBio, Rehovot, Israel). GlialCAM versions: full-length AA34-416 (OriGene, Rockville, MD, USA), ECD AA34-234 (Novoprotein, Summit, NJ, USA), and ICD AA262-416 with N-terminal His-Tag was cloned into a pet30(+) vector, expressed in BL21 E. coli to an OD of 600nm, and induced with IPTG (Sigma Aldrich, St. Louis, MO) for 3h at 37°C. Cell pellets were disrupted by sonication and proteins were purified with cOmplete His-Tag Purification Resins (Roche Life Science, Penzberg, Germany), followed by size-exclusion purification (Cytiva). For all other used peptides and proteins see tables S3-S5.

**Planar protein microarrays.** Protein microarrays were generated as described previously (60, 61)(<https://web.stanford.edu/group/antigenarrays/>). In brief, peptides, recombinant proteins, and lysates were diluted at the indicated concentrations in a 1:1 solution of PBS/water and protein printing buffer (ArrayIt Corporation Sunnyvale, CA) (tables S3-S5), aliquoted on 384-well plates, and printed on SuperEpoxy Slides using a NanoPrint LM210 system (ArrayIt). Two independent quadruplicates of each analyte were spotted, and some proteins were used in several

versions / preparations from different sources (table S3). Ready-made HuProt Arrays version 3.1 were obtained from CDI labs (Baltimore, MD). Arrays were circumscribed with a hydrophobic marker, blocked overnight at 4 °C in PBS containing 3% FCS and 0.1% Tween-20, and incubated with individual mAbs at a concentration of 1 µg/ml for 1h at 4°C, then washed twice for 20 min in blocking buffer on a rotating shaker. Arrays were then incubated with Cy-3-conjugated secondary goat anti-human IgG (0.8 µg/mL) (Jackson ImmunoResearch) for 1h at 4 °C, then washed twice for 30 min in blocking buffer, twice for 30 min in PBS, and twice for 15 s in water. Arrays were spun dry and scanned with a GenePix 4000B scanner (Molecular Devices). Median pixel intensities for each fluorescent spot were determined with GenePix Pro-3.0 software (Molecular Devices). Z-scores for each row of antigens were calculated for viral antigens, raw intensities were analyzed for GlialCAM arrays. Heatmaps were generated with Morpheus software (The Broad Institute; <https://software.broadinstitute.org/morpheus>).

**ELISA.** Cytokine ELISA sets were used according to manufacturers' instructions: mouse IL-6, IL-10, IL-12, IFN-γ, and TNF (BD Biosciences), and IL-17A (ThermoScientific). For protein and peptide ELISAs, MaxiSorp 384-well plates (ThermoFisher Scientific) were coated with 1 µg/ml peptide or protein in carbonate-bicarbonate buffer at 4°C overnight, then washed 6x with TBST (PBS + 0.05 % Tween20), blocked with blocking buffer (PBS + 1% BSA) for 1h, and mAbs were applied at 1 µg/ml in blocking buffer. Human and mouse plasma samples were diluted 1:100, CSF samples 1:20, and T cell supernatants 1:4 in blocking buffer. For serum ELISAs with blocking of serum IgG, diluted serum was incubated with 0.2 mg/ml of scrambled peptide, EBNA1 AA385-405, EBNA1 AA328-641, or GlialCAM ICD for 2h at RT before applying it to the plate. After overnight incubation at 4°C, plates were washed again 6x with TBST, secondary antibody HRP-conjugated goat anti-human IgG (Jackson ImmunoResearch, Inc., West Grove, PA), was applied for 1h at RT, and after 6 additional washes with TBST, plates were developed with TMB substrate (ThermoFisher Scientific), stopped with 1N sulfuric acid, and read on a SpectraMax Paradigm Microplate Reader (Molecular Devices, San Jose, CA).

**Western blotting.** Western blots were run according to standard protocols. Briefly, purified proteins were boiled in Laemmli-buffer with 10% beta-mercaptoethanol for 5 minutes, run on 4-12% Criterion XT Bis-Tris protein Gels (Bio-Rad, Hercules, CA), then transferred onto a

nitroglycerine membrane using a Trans-Blot Turbo semi-dry transfer system (Bio-Rad) and stained with MS39p2w174 at 10 $\mu$ g/ml or with mouse anti-EBNA1 antibody (Biorbyt, Cambridge, UK) or mouse anti-GlialCAM antibody (R&D systems, Minneapolis, MN) and secondary HRP-conjugated goat anti-human IgG and anti-mouse IgG (Jackson ImmunoResearch).

**Bio-Layer Interferometry.** Association and dissociation constants of mAbs to proteins and peptides were measured with bio-layer interferometry on an Octet QK device (Fortebio / Sartorius, Fremont, CA) according to standard protocols. For peptide kinetics, biotinylated peptides were bound to high precision streptavidin (SAX) biosensors (peptide concentration in solution: 100 nM) and mAbs MS39p2w174 and GL was probed as analytes in concentrations ranging from 10 – 270 nM. For protein kinetics, mAbs were bound to anti-hIgG Fc Capture (AHC) biosensors (mAb concentration in solution: 20 nM), and proteins were probed as analytes in concentrations ranging from 1.56 – 125 nM. Data was analyzed with BLI analysis software (Fortebio / Sartorius, version 7.1). Buffer controls were subtracted, and curves were fitted globally for each group consisting of all concentrations of the same ligand. KD values + SD as well as association / dissociation curves were reported and plotted with GraphPad prism (version 8.4).

**Prediction of Protein Disorder.** Order and disorder along the amino acid sequences of EBNA1 and GlialCAM were analyzed with PONDR (Predictor of Natural Disordered Regions, WSU Research Foundation) (62), using the VSL2 algorithm.

**Crystallization of antibody-antigen complexes.** EBNA1 AA386-405 20mer peptides (>98% purity) (Sigma Aldrich) were mixed with MS39p2w174-Fab (15 mg/ml) in a 1:7.5 molar ratio and incubated overnight. Crystals for MS39p2w174-Fab + EBNA1 AA386-405 grew in 0.48M Sodium Citrate, 0.72M Sodium/Potassium Phosphate, and 3% MPD (v/v) in 0.1M HEPES, pH 6.9 (fig. S6C). Data were collected at beamline SSRL 12-2 at the Stanford Linear Accelerator (SLAC), and processed and scaled using XDS/aimless and Staraniso (63, 64). Crystals belonged to space group I222 ( $a = 119.66 \text{ \AA}$ ,  $b = 137.56 \text{ \AA}$ ,  $c = 179.00 \text{ \AA}$ ,  $\alpha = \beta = \gamma = 90^\circ$ ) and contained two Fab-peptide complexes per asymmetric unit (fig. S6D). Phaser was used for molecular replacement (65) with the model structure 4LRI (PDB), stripped of all CDR loops. Loops were re-constructed with Coot (66) and structures were refined with phenix.refine (67, 68)



in iterations with Coot. Measurements and figure design was done with Pymol v2.1 (69). The structure was deposited in the protein data bank (PDB, [www.rcsb.org](http://www.rcsb.org)) with PDB ID: 7K7R.

**Mouse Immunization, EAE, and histology.** All animal experiments were performed in accordance with state and federal guidelines and regulations and approved by the Stanford Institutional Animal Care and use Committee. 8-week-old female SJL/J mice were purchased from The Jackson Laboratory (Bar Harbor, ME) and immunized s.c. with 200 µg/mouse of EBNA1 AA386-405 (peptide sequence: SQSSSSGSPRRPPPGRRPF) or scrambled control peptide (peptide sequence: SPSRPGRSRSRGSPFPQPSP) (10 mice per group), mixed with 100 µg/mouse of CpG (ODN1826, Invivogen, San Diego, CA) in 100 µl/mouse incomplete Freund's adjuvant (BD Difco, Franklin Lakes, NJ). 3 weeks later EAE was induced by s.c. immunization with 200 µg/mouse of PLP AA139-151, mixed again with the same peptides as in the first immunization, in 100µl of complete Freund's adjuvant, supplemented with 200 µg/mouse of mycobacterium tuberculosis (strain H37 RA, BD Difco). Serum samples were obtained by retro-orbital blood draws 3 days before the 1st and 2nd immunizations (day -24 and day -3), and during termination of the experiment (day 50). Mice were weighed daily, and disease severity was assessed according to a 5-point standard scoring system: 0, no clinical signs; 1, loss of tail tone; 2, hind limb weakness; 3, complete hind limb paralysis; 4, hind limb and forelimb paralysis; 5, moribund or dead. Mice were euthanized on day 50 post induction of EAE by deep anesthesia with i.p. injections of 0.01 ml/g body weight 7.2% Xylazine (Bayer Healthcare, Leverkusen, Germany) and 10.8% Ketamine (Pfizer, New York City, NY). Lymph nodes and spleens were extracted, and mice were then perfused with 10 ml PBS and 20 ml 4% paraformaldehyde (PFA) (Electron Microscopy Sciences, Hatfield, PA). Brains and spinal cords were extracted, stored in 4% PFA overnight, followed by 30% sucrose in PBS. Tissues were embedded in paraffin, sectioned, and stained for H&E as well as Luxol fast blue according to standard protocols. Infiltration of inflammatory cells into the spinal cord on H&E slides was assessed by a blinded investigator by counting lesions of infiltrating cells per slide, taking lesion size into account: 1: small infiltrate (<10 cells), 2: medium infiltrate (<100 cells), 3: large infiltrate (>100 cells). Demyelination was assessed by a blinded investigator by according to a histological score (70): 0.5, single demyelinated spot; 1, several spots; 2, confluent sites of demyelination; 3, extensive demyelination, less than half of a spinal cord; 4, demyelination of



more than half of the spinal cord; and 5, extensive demyelination affecting >85 % of the total white matter of the spinal cord.

**T cell proliferation.** Mouse T cell proliferation assays were set up as described previously (71). Briefly, spleens and lymph nodes were mechanically dissociated and after red blood cell lysis cells were seeded into flat-bottom 96-well plates at  $5 \times 10^5$  cells / well with the indicated peptides at 50 µg/ml. Cells were pulsed with  $^3\text{H}$ -methylthymidine after 48h, harvested 24h later, and read with a beta-plate reader (Wallac, Monza, Italy). Supernatants for cytokine ELISAs were obtained from unpulsed cells after 72h.

**Statistics.** GraphPad Prism version 8.4.1 and R version 3.6.1 were used for statistical analyses. Statistical tests used are indicated in the respective methods section or in the figure legends.

**Figures 1 – 4**

**Fig. 1: B cell phenotype and repertoire differences in MS blood and CSF.** (A-F): Flow cytometry data. (A) PB as percent of all B cells in MS blood and CSF. (B) VLA-4 expression and (C) HLA-DR expression in non-PB (red) and PB (blue) in blood vs. CSF. (C) Representative flow cytometry data (patient MS37) of HLA-DR expression in non-PB (red) and PB (blue) in blood (upper panel) and CSF (lower panel). (E,F) Immunoglobulin classes in (E) PB and (F) non-PB in blood (red) vs. CSF (blue). (A-F) p values according to unpaired Student's t tests. (G-I) Single-cell BCR repertoire sequencing data, (G) clonality, percent of clonal sequences are shown in blood vs. CSF, larger numbers of sequences in blood were down-sampled to match CSF sequence numbers, p values according to Mann-Whitney U test. (H) Overview of individual PB BCR repertoires, showing clonality, size of individual clonal expansions, and immunoglobulin classes in blood (upper panel) vs. CSF (lower panel), numbers indicate number of sequences, inner circle: colored wedges represent clonal expansions and grey area represents singleton antibody sequences, outer circle: immunoglobulin classes, red: IgG, blue: IgA, green: IgM, sequence locations in outer circle correspond to inner circle. (I) IGHV gene distribution in blood vs. CSF PB, p according to Student's T tests, Holm-Sidak adjusted p-values: \* < 0.05, \*\* < 0.01, \*\*\* < 0.0001. (J,K) Mass spectrometry data of purified immunoglobulins from CSF samples in singleton BCR B cells vs. clonally expanded B cells, (J) percent of VDJ sequences that could be uniquely identified with mass spectrometry in the respective groups (PSM cutoff: 1), (K) percent of VDJ sequences that were highly abundant in CSF ( $\geq 10$  PSM). (J,K) p values according to Mann-Whitney U test.

**Fig. 2: MS CSF mAb reactivity to EBV proteins and interaction of MS39p2w174 with EBNA1 AA386-405.** (A) Heatmaps showing mAb reactivities (z-scores) to viral lysates and EBV proteins and (B) to EBNA1 peptides. Selected mABs are shown with highest reactivities to respective antigens. IE: immediate early, E: early, and L: late lytic stage, red: mAB MS39p2w174 and antigen EBNA1 / MS-associated region. (C) Western blot of recombinant EBNA1 (full-length and truncated proteins), stained with commercial anti-C-terminal EBNA1 antibody (top panel) and MS39p2w174 (bottom panel). (D,E) ELISA data, (D) overlapping peptide scan of MS39p2w174 binding EBNA1 peptides (20mers, 13AA overlap), and (E) Alanine-scan, EBNA1 AA386-405 logo representation showing the contribution of each residue to binding of MS39p2w174. (F-J) Crystal structure of MS39p2w174 in complex with EBNA1

AA386-405. (F) Cartoon and stick representation, showing EBNA1 AA393-401 in the binding groove. Additional peptide residues are truncated for better visualization. HC: red/brown colors, LC: blue/cyan, CDR loop colors correspond to annotations in G. (G) View of the binding groove from the top. Surface representation of the Fab with EBNA1 AA386-402 in stick representation. (H-J) Cartoon and stick representation outlining close interactions. Major H-bondforming residues are represented as sticks. H-bonds  $< 3.1 \text{ \AA}$  are represented as black dashed lines. (I) Magnification of peptide in hydrophobic cage, (J) magnification of region around Arg396 to emphasize polar contacts of HC residues with Arg396 and Arg397. (K,L) (E) Bio-layer interferometry measurement of MS39p2w174 (red) and germline (GL, blue) affinity to EBNA1 full-length protein. (K) KD in nM, (L) association and dissociation curves. P according to unpaired Student's t test.

**Fig. 3: Molecular Mimicry between EBNA1 and GlialCAM.** (A) Heatmap showing top 16 results of HuProt array for MS39p2w174, compared to 3 control mAbs, sorted from top to bottom by the ratio of MS39p2w174 / average of controls (left column, min: 89, max: 911). Raw counts are shown in the four columns on the right (min: 1, max: 36450). (B) ELISA showing binding of MS39p2w174 and two control mAbs to recombinant proteins EBNA1 AA328-641 as well as GlialCAM AA34-416 and A262-416. (C) Western blot of recombinant GlialCAM (full-length, ICD, and ECD), stained with commercial anti-GlialCAM antibody (top panel) and MS39p2w174 (bottom panel). (D,E) Bio-layer interferometry affinity measurement of MS39p2w174 (red) and germline (GL, blue) to GlialCAM (full-length protein). (D) KD in  $\mu\text{M}$ , (E) association and dissociation curves, p according to unpaired Student's t test. (F) Logo plot, showing alignment of amino acid sequences of EBNA1 AA386-405 and GlialCAM AA370-389 and pointing out the central epitope region. (G,H) Prediction of disorder with PONDR for (G) EBNA1 and (H) GlialCAM. High scores indicate disorder, red areas: epitope regions. (I) ELISA data showing binding of MS39p2w174 to EBNA1 AA386-405 and GlialCAM AA370-389 non-phosphorylated and phosphorylated at the indicated serine residues, pSer / Sp: phosphorylated serine. (J,K) Bio-layer interferometry measurement of MS39p2w174 affinity to GlialCAM 20mer peptides. pSer: phosphorylated Serine residues, p according to unpaired Student's t test. (L) Heatmap showing mAb reactivities (MFI) to GlialCAM proteins, peptides, and phosphorylated peptides, as well as cross-reactivities to EBNA1 and other EBV proteins. ICD: intracellular domain, ECD: extracellular domain, pSer: phosphorylated serine residues. (M-O)

ELISA data showing human plasma reactivities against (M) EBNA1 full-length protein, (N) EBNA1 AA386-405, and (O) GlialCAM protein, p according to unpaired Student's t test. (P) Plasma reactivity of patient MS30 to EBNA1 AA386-405, blocked with indicated proteins or peptides, p according to unpaired Student's t test.

**Fig. 4: Anti-EBNA1 AA386-405 antibodies exacerbate autoimmune-mediated demyelination in vivo.** (A-C) ELISA data, mouse plasma IgG response against (A) EBNA1 AA386-405, (B) GlialCAM ICD, and (C) PLP on the indicated timepoints pre and post EAE induction for scrambled peptide immunized mice (blue) and EBNA1 386-405 immunized mice (red). Fold induction over first data point is shown, each data point represents mean value of a three independent IgG titer measurements per mouse. Mean + SD are shown per group, p according to unpaired Student's t-test, Holm-Sidak adjusted. (D) EAE scores of mice immunized with scrambled peptide control (blue) and EBNA1 AA386-405 (red), \*  $p < 0.05$  (Mann-Whitney U test). (E-H) Spinal cord histology, (E) representative H&E stained spinal cords from scrambled peptide group (top panel) and EBNA1 AA386-405 group (bottom panel). (F) Statistical evaluation of H&E score, p according to unpaired Student's t-test. (G) Representative Luxol Fast Blue stained spinal cords from scrambled peptide group (top panel) and EBNA1 AA386-405 group (bottom panel). (H) Statistical evaluation of Luxol Fast Blue score, p according to unpaired Student's t-test. (E,G) scale bars left images: 200  $\mu\text{m}$ , right images: 50  $\mu\text{m}$ .

## Supplementary figures S1 – S8

**Supplementary figure S1:** Analysis of B cell phenotypes in blood and CSF

**Supplementary figure S2:** Extended BCR repertoire data.

**Supplementary figure S3:** Phylogenetic trees of B cells from blood and CSF.

**Supplementary figure S4:** CSF mAb reactivity to EBV and GlialCAM antigens.

**Supplementary figure S5:** CSF mAb reactivity to EBV peptides.

**Supplementary figure S6:** mAb reactivity to EBV peptides and structural data for EBNA1 AA386-405 / MS39p2w174 complex.

**Supplementary figure S7:** Extended characteristics of GlialCAM AA370-389.

990 **Supplementary figure S8:** T cell response in mice immunized with EBNA1 AA386-405.

991

992 **Supplementary tables S1 – S5**

993 **Supplementary table S1:** Patient collective.

994 **Supplementary table S2:** Genotype of full HLA locus.

995 **Supplementary table S3:** Viral proteins and lysates.

996 **Supplementary table S4:** Viral peptides.

997 **Supplementary table S5:** GlialCAM proteins and peptides.