JCI The Journal of Clinical Investigation

Two rheumatoid arthritis–specific autoantigens correlate microbial immunity with autoimmune responses in joints

Annalisa Pianta, … , Catherine E. Costello, Allen C. Steere

J Clin Invest. 2017. https://doi.org/10.1172/JCI93450.

Research Article **Autoimmunity**

In rheumatoid arthritis (RA), immunological triggers at mucosal sites, such as the gut microbiota, may promote autoimmunity that affects joints. Here, we used discovery-based proteomics to detect HLA-DR–presented peptides in synovia or peripheral blood mononuclear cells and identified 2 autoantigens, *N*-acetylglucosamine-6-sulfatase (GNS) and filamin A (FLNA), as targets of T and B cell responses in 52% and 56% of RA patients, respectively. Both GNS and FLNA were highly expressed in synovia. GNS appeared to be citrullinated, and GNS antibody values correlated with anti–citrullinated protein antibody (ACPA) levels. FLNA did not show the same results. The HLA-DR–presented GNS peptide has marked sequence homology with epitopes from sulfatase proteins of the *Prevotella* sp. and *Parabacteroides* sp., whereas the HLA-DR–presented FLNA peptide has homology with epitopes from proteins of the *Prevotella* sp. and *Butyricimonas* sp., another gut commensal. Patients with T cell reactivity with each self-peptide also had responses to the corresponding microbial peptides, and the levels were directly correlated. Furthermore, HLA-DR molecules encoded by shared-epitope (SE) alleles were predicted to bind these self- and microbial peptides strongly, and these responses were more common in RA patients with SE alleles. Thus, sequence homology between T cell epitopes of 2 self-proteins and a related order of gut microbes may provide a link between mucosal and joint immunity in patients with RA.

Find the latest version:

https://jci.me/93450/pdf



Two rheumatoid arthritis–specific autoantigens correlate microbial immunity with autoimmune responses in joints

Annalisa Pianta,[1] Sheila L. Arvikar,[1] Klemen Strle,[1] Elise E. Drouin,[1] Qi Wang,[2] Catherine E. Costello,[2] and Allen C. Steere[1]

[1]Center for Immunology and Inflammatory Diseases, Division of Rheumatology, Allergy and Immunology, Massachusetts General Hospital (MGH), Harvard Medical School, Boston, Massachusetts, USA.

[2]Center for Biomedical Mass Spectrometry, Boston University School of Medicine, Boston, Massachusetts, USA.

In rheumatoid arthritis (RA), immunological triggers at mucosal sites, such as the gut microbiota, may promote autoimmunity that affects joints. Here, we used discovery-based proteomics to detect HLA-DR–presented peptides in synovia or peripheral blood mononuclear cells and identified 2 autoantigens, *N*-acetylglucosamine-6-sulfatase (GNS) and filamin A (FLNA), as targets of T and B cell responses in 52% and 56% of RA patients, respectively. Both GNS and FLNA were highly expressed in synovia. GNS appeared to be citrullinated, and GNS antibody values correlated with anti–citrullinated protein antibody (ACPA) levels. FLNA did not show the same results. The HLA-DR–presented GNS peptide has marked sequence homology with epitopes from sulfatase proteins of the *Prevotella* sp. and *Parabacteroides* sp., whereas the HLA-DR–presented FLNA peptide has homology with epitopes from proteins of the *Prevotella* sp. and *Butyricimonas* sp., another gut commensal. Patients with T cell reactivity with each self-peptide also had responses to the corresponding microbial peptides, and the levels were directly correlated. Furthermore, HLA-DR molecules encoded by shared-epitope (SE) alleles were predicted to bind these self- and microbial peptides strongly, and these responses were more common in RA patients with SE alleles. Thus, sequence homology between T cell epitopes of 2 self-proteins and a related order of gut microbes may provide a link between mucosal and joint immunity in patients with RA.

Introduction

Rheumatoid arthritis (RA) is an HLA class II–associated autoimmune disease, in which arthritogenic T cells drive the progressive inflammation and destruction of synovial joints (1). Both genetic and environmental factors are thought to contribute to disease development and progression. The greatest genetic risk factor is HLA-DRB1–susceptibility alleles that share a 5–amino acid sequence in the B1 chain, termed the RA shared epitope (SE) (2). HLA-DRB1 SE alleles largely influence the development of seropositive RA, which is defined by positive tests for rheumatoid factor (RF) and/or anti–citrullinated protein antibodies (ACPAs) (3), the latter being the only known specific autoantibodies for this disease (4–6). RF and/or ACPAs may develop years before the onset of clinical arthritis (7–9), suggesting that autoimmunity may be triggered at sites other than joints in patients with RA.

Causative environmental factors are less well characterized. However, T cell epitope mimicry between microbial pathogens and self-proteins has been implicated as a possible factor in the induction or exacerbation of autoimmune disease (10–12). In addition, alteration of the oral or gut microbiota may affect mucosal immunity, inducing aberrant immune responses that affect joints in patients with RA (13, 14). Using high-throughput sequencing, Scher et al. showed that *Prevotella* species (spp.), including *P. copri*, in the gut microbiota were expanded in stool samples from patients with new-onset RA (NORA), suggesting that these organisms might have this role in RA pathogenesis (13). Moreover, a recent study in mice showed that gut dysbiosis contributes to arthritis development via the activation of autoreactive T cells in the intestine (15). Proposed mechanisms to link infection and autoimmunity include molecular mimicry between T cell microbial and host epitopes (16); infection-induced alteration and release of sequestered self-antigens (11); or nonspecific, infection-induced inflammatory responses that function as adjuvants in the induction of pathogenic autoimmunity (17, 18).

The identification of disease-relevant infectious or self-antigens has been challenging in any autoimmune disease, but current discovery-based methods offer innovative approaches to this problem. We have developed an approach for antigen detection in chronic inflammatory arthritides, in which HLA-DR–presented peptides (T cell epitopes) are identified directly from patients' inflamed synovial tissue, synovial fluid mononuclear cells (SFMCs), or peripheral blood mononuclear cells (PBMCs) by liquid chromatography–tandem mass spectrometry (LC-MS/MS) and tested for immunogenicity using patients' samples (19–24).

With this approach, we recently identified an HLA-DR–presented peptide from a 27-kDa protein of *P. copri* (*Pc*-p27), which stimulated T and B cell responses in approximately 40% of patients with RA, but not in patients with other rheumatic diseases or in healthy controls (HCs) (25). Using the same methodolo-

Conflict of interest: The authors have declared that no conflict of interest exists.
Submitted: February 15, 2017; **Accepted:** May 4, 2017.
Reference information: *J Clin Invest*. https://doi.org/10.1172/JCI93450.

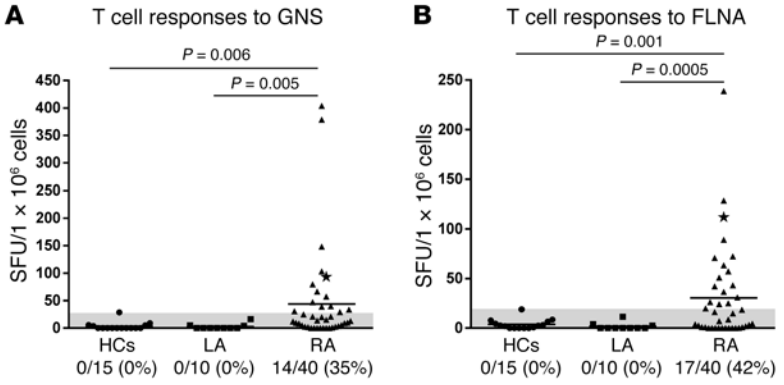


Figure 1. T cell reactivity to GNS and FLNA peptides in RA patients and comparison group subjects. In initial experiments, (**A**) PBMCs from patients with RA or LA or from HC subjects were stimulated with a pool of 4 peptides, including the single GNS HLA-DR–presented peptide isolated from the synovial tissue of patient RA1, and 3 predicted promiscuous HLA-DR–binding peptides from GNS (1 μM each). (**B**) PBMCs from patients and HCs were incubated with a pool of 4 peptides, including the single FLNA HLA-DR–presented peptide identified from the synovial tissue and PBMCs from patient RA1, and 3 predicted promiscuous HLA-DR–binding peptides from FLNA (1 μM each). In each assay, a positive control (phytohemagglutinin) and a negative control (no peptide) were included. The amount of IFN-γ secretion, as determined by an ELISpot assay, is shown. A positive response was defined as greater than 3 SD above the mean value for HCs (area above the shaded region). The values for patient RA1 are indicated with a star. Horizontal lines represent the mean values for each group. *P* values were determined by unpaired, 2-tailed *t* test with Welch's correction. SFU, spot-forming units per million PBMCs.

gy, we report here the identification of 2 previously unidentified autoantigens, *N*-acetylglucosamine-6-sulfatase (GNS) and filamin A (FLNA) that are targets of T and B cell responses and appear to be specific for RA. Both autoantigens are highly expressed in inflamed synovial tissue; they share homologous T cell epitopes with *Prevotella* and several other gut microbes, and they are targets of specific T and B cell responses in patients with RA, providing evidence that may link immune responses to microbial peptides from gut commensals and autoimmune responses affecting joints.

Results

Identification of naturally presented HLA-DR peptides (T cell epitopes). In a recent study (20), we identified HLA-DR–presented peptides in synovial tissue, SFMCs, and PBMCs from 5 patients with RA using LC-MS/MS, and the immunogenicity of the peptides was determined using patients' samples in enzyme-linked immunospot (ELISpot) assays. The findings from 1 patient (referred to here as RA1) were of particular interest. She had classic, seropositive RA, with severe symmetrical polyarthritis, a positive test for ACPAs, and 2 copies of SE alleles (HLA-DRB1*0401 and *0101). In this patient, an immunogenic HLA-DR–presented peptide derived from a *P. copri* protein (*Pc*-p27) was identified from the patient's PBMCs (25). We then showed that approximately 40% of patients with RA have T and/or B cell responses to *Pc*-p27 or to the whole *P. copri* organism (25).

In patient RA1, 2 immunogenic HLA-DR–presented human self-peptides derived from GNS and FLNA were also identified from her synovial tissue, and the same FLNA peptide was also found in her PBMCs (20). The HLA-DR–presented peptide derived from GNS was predicted to be promiscuous, binding to 24 of the 25 HLA-DR molecules modeled in the program TEPITOPE (26), and the FLNA-derived peptide was predicted to bind 9 of the 25 HLA-DR molecules. With both peptides, this included binding by HLA-DR molecules encoded by SE alleles *0101, *0401, *0404, and *0405. Neither the GNS protein, nor the FLNA protein, nor the *P. copri* protein had previously been noted to be antigens in RA.

T cell reactivity to GNS and FLNA peptides. To determine the immunogenicity of HLA-DR–presented peptides and their source proteins more broadly, we have developed a cohort of NORA patients seen prior to commencing therapy with disease-modifying antirheumatic drugs (DMARDs), which is a time when immune responses would be expected to be most robust. For comparison, we tested samples from patients with Lyme arthritis (LA) and from HC subjects. HLA-DR typing showed that 60% of the 40 RA patients had SE alleles, and 50% of the 10 LA patients and 42% of the 15 healthy subjects also had SE alleles. Nevertheless, since the patients and HCs had a range of different HLA-DR alleles, and since cell numbers are limited in human patients, our initial approach for determining T cell responses in multiple individuals consisted of pooling the original peptide with 3 additional peptides from the same protein that are predicted by the program TEPITOPE to be promiscuous HLA-DR binders (26). In addition, because of a limited number of cells, we did not include the testing of irrelevant control peptides in these experiments. However, we have previously shown that patients with RA do not have reactivity to peptides derived from endothelial cell growth factor (ECGF) or MMP-10 peptides (21, 23). These autoantigens in LA are irrelevant in RA.

When PBMCs from 40 patients with NORA were stimulated with the GNS peptides, we found that 14 of the 40 patients (35%) secreted levels of IFN-γ that were greater than 3 SD above the mean value for HCs ($P = 0.006$), as determined by an IFN-γ/IL-17 double-color ELISpot assay (Figure 1A). In comparison, PBMCs from patients with LA lacked reactivity to these peptides ($P = 0.005$). When FLNA peptides were used to stimulate PBMCs from the same set of patients and control subjects, 17 of the 40 patients with NORA (42%) had IFN-γ levels that were greater than 3 SD above the mean value for HCs ($P = 0.001$) and for patients with LA ($P = 0.0005$) (Figure 1B). In patients with RA, the predominant response to stimulation with both peptide sets was a Th1-type response with IFN-γ secretion, whereas PBMCs from only 3 RA patients secreted IL-17 (data not shown). Altogether, 21 of the 40 patients (52%) had T cell reactivity to GNS and/or FLNA peptides, and 10 (25%) had reactivity to both.

B cell reactivity to GNS and FLNA proteins. Since the role of $CD4^+$ T cells would likely be to help B cells produce autoantibodies against GNS or FLNA, we examined IgG levels for these proteins in serum samples from patients with RA and control group subjects. Since sera (but not PBMCs) were also available from patients with chronic RA (CRA), testing was done in 48 NORA patients and 53 CRA patients. Because the results were similar in both groups, they are presented together here.

Of the 101 patients with RA, 32 (32%) had IgG antibody responses against GNS that were greater than 3 SD above those

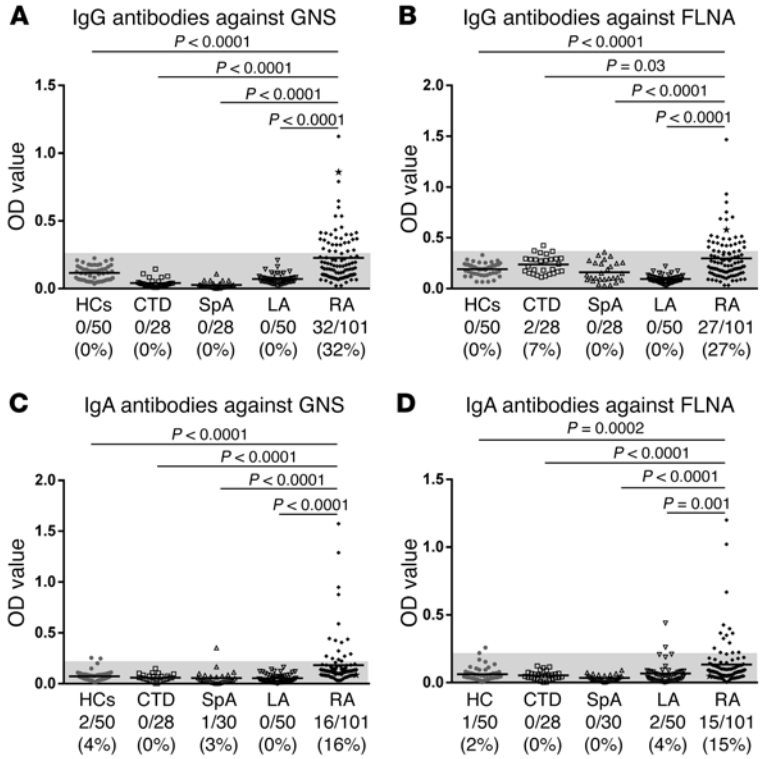


Figure 2. IgG and IgA responses to GNS and FLNA in RA patients and comparison group subjects. Serum samples from 259 patients with RA, patients with other forms of chronic inflammatory arthritis, and HCs were tested by ELISA for autoantibodies. (**A** and **C**) Plates were coated with the GNS protein and incubated with serum from patients or HCs. All serum samples were tested in duplicate for anti-GNS IgG (**A**) or IgA (**C**) antibody responses. (**B** and **D**) Plates were coated with the FLNA protein and incubated with serum from patients or control subjects. All serum samples were tested in duplicate for anti-FLNA IgG (**B**) or IgA (**D**) antibody responses. For all analyses, positivity was defined as greater than 3 SD above the mean value for HCs (area above the shaded region). Symbols represent values in individual patients, and horizontal lines show the mean values. Values for patient RA1 are indicated with a star. Only significant *P* values, determined by unpaired, 2-tailed *t* test with Welch's correction, are shown.

in HCs ($P < 0.0001$) (Figure 2A). In contrast, none of the 106 patients with other diseases, including those with LA, spondyloarthropathy (SpA), or connective tissue diseases (CTD), and none of the 50 HC subjects had positive IgG antibody responses against the protein (in each instance, $P < 0.0001$). Similarly, 27 of the 101 (27%) patients with RA had levels of IgG antibodies against FLNA that were greater than 3 SD above those in HCs ($P < 0.0001$), whereas only 2 patients with CTD had borderline positive IgG antibody responses against FLNA, and none of the other control subjects had positive responses (Figure 2B). Altogether, 48 (48%) of the 101 RA patients had IgG autoantibodies against GNS and/or FLNA, and 10 (10%) had IgG reactivity against both proteins.

Because autoimmune processes in RA may be triggered at mucosal sites, we also tested the levels of IgA antibodies against GNS and FLNA in serum samples from patients and control subjects. Of the 101 patients with RA, 16 (16%) had elevated IgA antibody responses against GNS that were greater than 3 SD above those in HCs ($P < 0.0001$) (Figure 2C). In contrast, of the 106 patients with other rheumatic diseases and the 50 HC subjects, only 1 patient with SpA had borderline positive IgA antibodies. Similarly, 15 (15%) of the 101 RA patients had FLNA IgA responses that were greater than 3 SD above those in HCs ($P = 0.0002$), whereas only 1 HC subject and 2 patients with LA had low-level positive responses (Figure 2D). Altogether, 21 (21%) of the 101 RA patients had IgA antibody responses against GNS and/or FLNA, and 10 (10%) had IgA antibody responses against both proteins.

When IgG and IgA responses were considered together, 48 of the 101 patients with RA (48%) had IgG antibody responses against GNS and/or FLNA; 21 (21%) had IgA responses against 1 or both of the proteins; and 56 (55%) had IgG and/or IgA responses against the proteins. Of the 14 patients who had T cell responses to GNS peptides, 11 (79%) had IgG and/or IgA antibody responses against the GNS protein. Among the 17 patients who had T cell reactivity to the FLNA peptides, 6 (35%) had IgG and/or IgA antibody responses against the FLNA protein. Thus, T and B cell concordance was greater with GNS than with FLNA.

Correlation of antibody responses against P. copri, GNS, and FLNA. Using these same serum samples (25), we have previously tested IgG and IgA antibody responses against 2 RA-associated bacteria, *P. copri*, a gut microbe, and *Porphyromonas gingivalis*, a periodontal pathogen (27). Antibody responses against *P. copri* were found in 32% of RA patients, but were absent in patients with other CTDs, SpA, or LA, as well as in healthy subjects (25). Therefore, using these data, we correlated IgG and IgA antibody responses against these 2 organisms with the GNS and FLNA antibody responses determined here.

In patients with RA, the levels of anti-GNS IgG and IgA antibodies strongly correlated with *P. copri* antibody responses ($P = 0.002$ and $P < 0.0001$, respectively), and we found a similar correlation between anti-FLNA IgG and IgA antibody responses and *P. copri* antibodies ($P < 0.0001$ and $P < 0.0001$) (Figure 3A). In contrast, anti-GNS and anti-FLNA IgG or IgA levels did not correlate with *P. gingivalis* antibody responses (Figure 3B). Additionally, we observed no correlations among these parameters in healthy subjects. Thus, in RA patients, the higher the IgG or IgA antibody responses against *P. copri*, the greater the autoantibody responses against these autoantigens.

Testing of citrullinated GNS and FLNA proteins. Because citrullinated autoantigens are thought to play a central role in RA, particularly in patients with SE alleles, we investigated whether autoantibody responses against GNS or FLNA were greater when these proteins were citrullinated. For this purpose, the native proteins

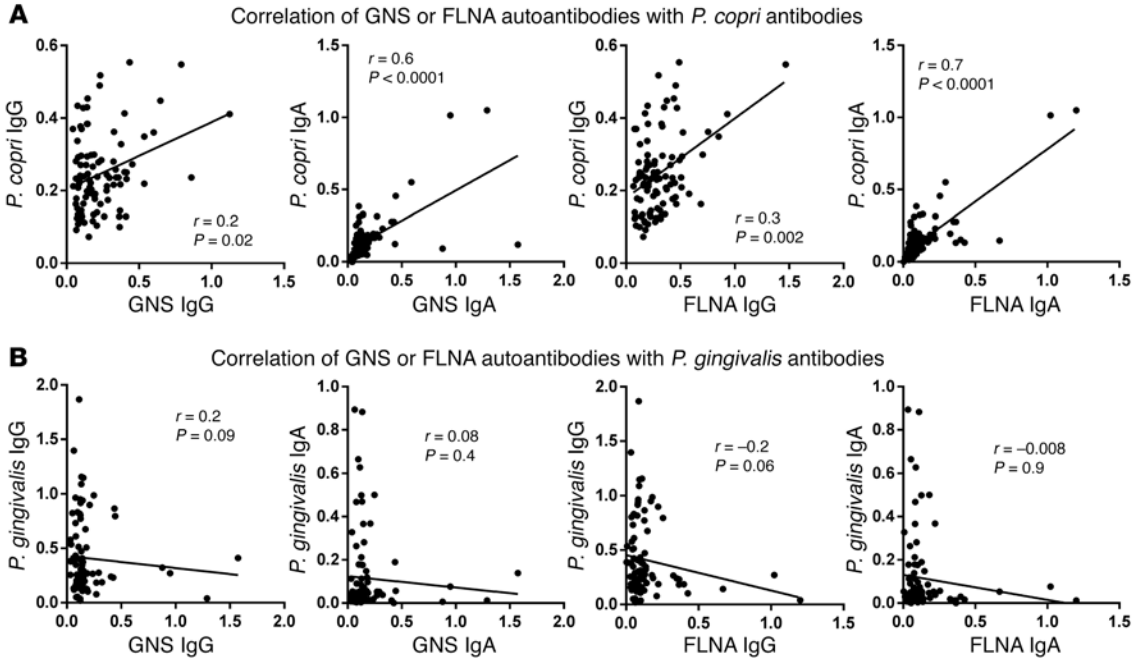


Figure 3. Autoantibody correlations with *P. copri* and *P. gingivalis* antibodies. Correlations between anti-GNS or anti-FLNA antibodies (IgG or IgA) and antibodies against *P. copri* (**A**) or *P. gingivalis* (**B**) in 101 RA patients. The *r* and *P* values for the corresponding statistical comparisons were determined by Spearman's correlation test.

were citrullinated in vitro using recombinant human (rh) peptidylarginine deiminase 4 (PAD4) enzyme. Using samples from 46 RA patients in whom a sufficient amount of serum still remained, IgG antibody responses were higher against citrullinated GNS compared with responses against the uncitrullinated protein ($P = 0.005$), whereas the responses were negative against both forms of the protein in 15 HC subjects (Figure 4A). Moreover, the magnitude of anti–citrullinated GNS antibody responses correlated with ACPA levels in these patients ($P = 0.03$) (Figure 4B). In contrast, IgG antibody responses against citrullinated and uncitrullinated FLNA were not significantly different in the 46 patients (Figure 4C), and the levels of anti–citrullinated FLNA antibodies did not correlate with ACPA levels (Figure 4D). These results suggest that the GNS protein, but not the FLNA protein, may be citrullinated in vivo in patients with RA.

Utility of GNS and FLNA autoantibody evaluation in the diagnosis of RA. In our patient cohort, 70 (69%) of the 101 NORA and CRA patients were seropositive for ACPAs and/or RF, which are standard, commercially available autoantibody determinations for support of the diagnosis of RA. Among the 31 patients who did not have a positive test for ACPAs and/or RF, 13 had a positive test for IgG and/or IgA GNS autoantibodies and 9 had a positive test for IgG and/or IgA FLNA autoantibodies. Taken together, 17 of the 31 seronegative patients (55%) had such autoantibodies, 15 of whom could be identified with the IgG test alone. Overall, when autoantibody responses against GNS and FLNA were combined with standard autoantibody determinations, 87 (86%) of the 101 patients with RA had a positive test result for support of the diagnosis, and only 14 (14%) lacked a specific marker for RA.

GNS and FLNA protein levels in serum and joints. For a self-protein to become the target of autoimmune responses in RA patients' inflamed joints, one would predict that the protein would be present at high concentrations there. For this purpose, we measured GNS and FLNA protein concentrations in serum samples from the 101 patients with RA and in synovial fluid (SF) from 17 patients for whom such samples were available. The levels of GNS were higher in the serum of RA patients than were GNS levels in the control groups ($P \leq 0.002$) (Figure 5A), and in RA patients, the levels of this protein tended to be higher in SF than in serum. Similarly, FLNA protein levels were significantly higher in the serum of RA patients than in HCs ($P < 0.0001$), but in RA patients, FLNA protein levels in SF and serum were similar (Figure 5B).

To gain further insight into the protein abundance and distribution, synovial tissues from 10 patients, 4 with RA and 3 each with LA or osteoarthritis (OA), were stained for expression of GNS and FLNA, using immunohistologic methods. GNS showed a fine, reticular pattern in and around endothelial cells in 3 of the 4 patients with RA, but not in those with LA or OA (Figure 6). We observed that FLNA was intensely expressed in the tunica muscularis around blood vessels and in large or elongated cells, presumably synoviocytes, and it was also faintly expressed in the extracellular matrix. We detected FLNA expression in all RA patients, lesser staining in 2 of the LA patients, but no staining in the OA patients (Figure 6). Thus, in RA, these 2 proteins were present in inflamed synovial tissue, particularly around blood vessels, where they could become targets of autoimmune responses.

Sequence homology between T cell epitopes of microbial and self-peptides. In an effort to determine whether T cell epitope mimicry may play a role in linking *Prevotella* reactivity with GNS and FLNA autoimmune responses, the sequence of each of the 2 self-peptides isolated from patient RA1 was used first to search for regions of similarity with any microbial protein using the

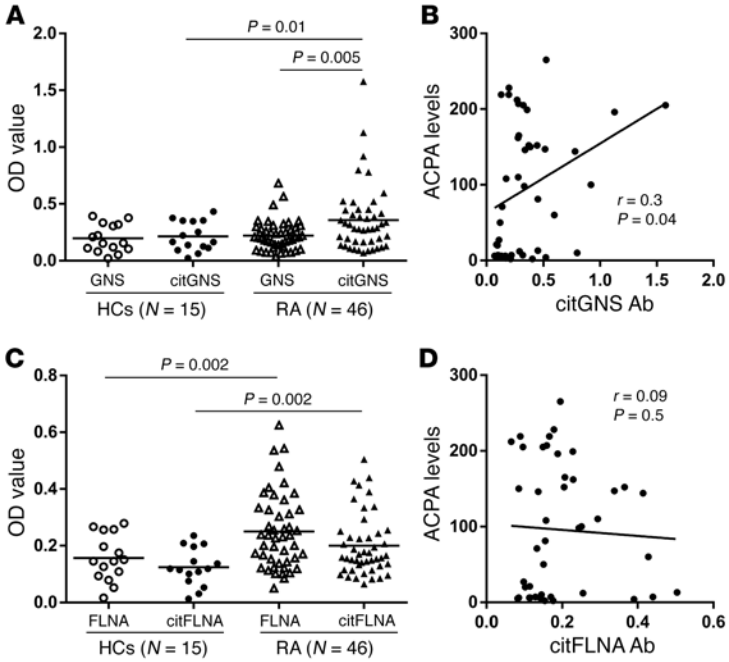


Figure 4. Autoantibody responses to citrullinated GNS and FLNA, and correlations with ACPAs. Serum samples from 46 patients with RA and 15 healthy individuals were tested for IgG antibody responses against citrullinated versus uncitrullinated GNS or FLNA. Plates were coated with GNS (**A**) or FLNA (**C**), with or without citrullination, incubated with serum from patients or HC subjects, and tested in duplicate. Symbols represent values for individual patients, and horizontal lines indicate the mean values. In **A** and **C**, only significant *P* values, calculated by an unpaired, 2-tailed *t* test with Welch's correction, are shown. (**B**) Correlation between IgG antibody responses against citrullinated GNS or citrullinated FLNA (**D**) and ACPA levels in the 46 patients with RA. The *r* and *P* values shown in **B** and **D** were determined by Spearman's correlations. citGNS, citrullinated GNS; citFLNA, citrullinated FLNA.

microbial protein database in BLASTP (Basic Local Alignment Search Tool, protein) (https://blast.ncbi.nlm.nih.gov/Blast.cgi). For both self-peptides, *Prevotella* spp. peptides were among the top sequences producing significant alignment, specifically in areas predicted to be in the HLA-DR–binding groove. Therefore, we refined the search for sequence similarity by screening only microbial sequences from *Prevotellaceae* (NCBI Entrez Genome taxid:171552). To evaluate sequence homology, self- and microbial peptides were aligned using the program Clustal Omega (28).

The peptide derived from GNS had 67% sequence homology with a peptide from the *Prevotella* arylsulfatase protein (WP_062433009) (Figure 7A), which was predicted by CELLO software to have a periplasmic location (29). Importantly, the major area of homology for this microbial peptide was restricted to amino acids predicted to be in the HLA-DR–binding groove (Figure 7A). For the *Prevotella* peptide, 5 of the 9 amino acids were identical to one of the predicted binding registers (P1 = the first F) in the GNS peptide. Moreover, the peptides shared amino acid identity at the P1, P4, and P6 sites, which are critical for peptide binding, as well as in the flanking regions at each end of the peptide, which also influence peptide binding (Figure 7A). The FLNA peptide had 80% identity with a peptide derived from an uncharacterized *Prevotella* protein (WP_028897633) (Figure 7A), which was predicted to have an extracellular location (a secreted protein). Moreover, the major area of homology was again found in the HLA-DR–binding groove, where 7 of the 9 amino acids were identical in the *Prevotella* and FLNA peptides, and the remaining 2 amino acids had conserved properties (Figure 7A).

For comparison, we analyzed the GNS and FLNA sequences for homology with *P. gingivalis*, a periodontal pathogen of interest in RA, using *Porphyromonadaceae* (taxid:171551) as the reference database in the BLASTP search. However, we found no homology between *P. gingivalis* and GNS or FLNA sequences. *P. gingivalis* also stimulates antibody responses in the subgroup of RA patients who have periodontal disease (27, 30), but there is little overlap between RA patients with *P. gingivalis* antibodies and those with *P. copri* antibodies (25).

Instead, among the *Porphyromonadaceae*, the GNS epitope had partial sequence similarity with a peptide from the periplasmic protein *N*-acetylgalactosamine-6-sulfatase of the *Parabacte-*

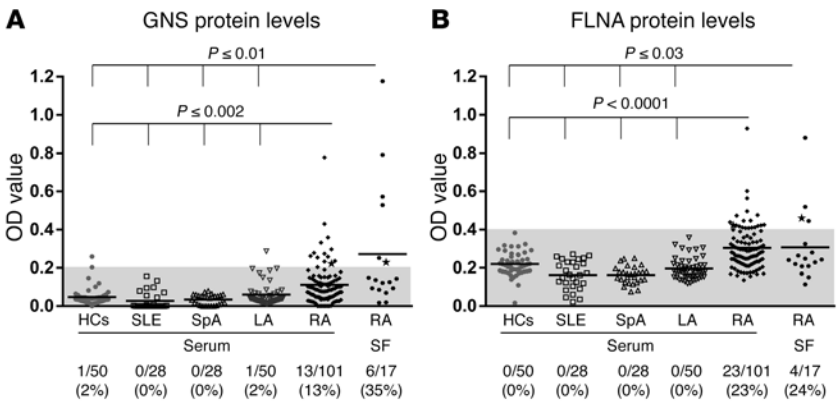


Figure 5. GNS and FLNA protein levels in RA patients and comparison group subjects. GNS and FLNA protein concentrations were measured in serum and SF samples from patients with RA, serum samples from patients with CTD, SpA, or LA, and serum samples from HCs. (**A**) GNS protein concentrations and (**B**) FLNA protein concentrations are shown, as measured by ELISA assay. For both analyses, positivity was defined as greater than 3 SD above the mean value for HC subjects (area above the shaded region). Symbols represent values for individual patients, and horizontal lines indicate the mean values. The values for patient RA1 are indicated with a star. Only significant *P* values, determined by unpaired, 2-tailed *t* test with Welch's correction, are shown. SLE, systemic lupus erythematosus.

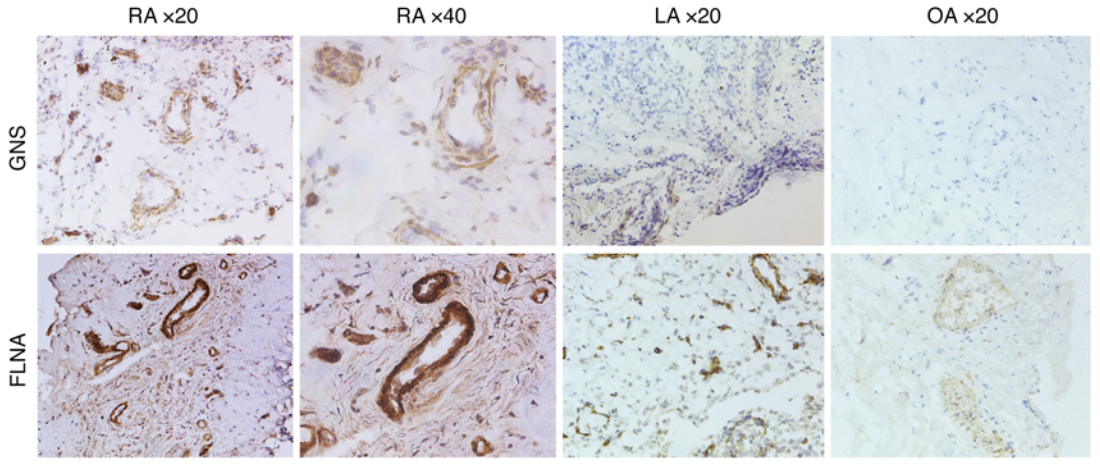


Figure 6. Immunohistochemical staining of synovial tissue for GNS and FLNA. Representative synovial tissue images from 1 patient with RA, 1 with LA, and 1 with OA are shown for expression of GNS or FLNA protein. Brown color indicates specific staining of GNS or FLNA self-proteins, and purple indicates hematoxylin staining. Images were taken at ×20 magnification, and ×40 magnification was used for the RA patient to highlight the staining around blood vessels.

roides sp. (WP_046148720) (Figure 7A). Thus, as with the homology between the GNS peptide and the peptide from the *Prevotella* arylsulfatase protein, the *Parabacteroides* protein was also a sulfatase. These enzymes are key in the adaptation and persistence of human commensal bacteria in the gut (31, 32). Moreover, the *Parabacteroides* peptide had 4 amino acids identical to the GNS peptide and 1 amino acid with conserved properties. These included amino acids with shared identity in the P1 through P4 and P6 sites. A similar evaluation of the FLNA peptide showed sequence homology with a predicted cytoplasmic uncharacterized protein of the *Butyricimonas* sp. (WP_065219401.1), another gut commensal. The *Butyricimonas* peptide shared identity with 6 of 9 amino acids in one of the predicted HLA-DR registers (P1 = F) of the FLNA peptide, and two of the three remaining amino acids had conserved properties (Figure 7A). *Prevotella*, *Parabacteroides*, and *Butyricimonas* are each members of the Bacteroidetes phylum, one of the two major phyla of gut commensal organisms.

T cell responses to homologous microbial and self-peptides. To address whether patients had reactivity to these self-epitopes and the corresponding microbial epitopes, we performed ELISpot assays with each of these peptides using PBMCs from the 24 patients with NORA in whom sufficient numbers of cells remained and from 10 HCs. When cells were stimulated with the GNS peptide or each of the 2 corresponding microbial peptides (1 derived from *Prevotella* and the other from *Parabacteroides*), we found Th1 cell reactivity to all 3 peptides. Of the 24 RA patients, 8 (33%) had T cell reactivity to the GNS peptide, 9 (38%) showed responses to the *Prevotella* peptide, and 6 (25%) had reactivity to the *Parabacteroides* peptide, all of which were responses that were 3 SD or more above the mean values for HCs (Figure 7B). Of the 8 patients who had reactivity to the GNS peptide, 7 also had responses to the microbial peptides. When PBMCs were incubated with the FLNA peptide, 9 of the 24 patients with RA (38%) had T cell responses, 10 (42%) showed reactivity to the corresponding *Prevotella* peptide, and 7 (29%) had responses to the *Butyricimonas* peptide (Figure 7B). Furthermore, all 9 patients with reactivity to the FLNA peptide also had responses to the microbial peptides. Thus, except for 1 patient, the same patients who had reactivity to the GNS and/or FLNA peptides also had responses to the corresponding microbial peptides. Additionally, among the microbial peptides, we observed a trend toward a higher percentage of patients who had reactivity to the *Prevotella* peptides than to the other gut commensals.

Moreover, when the magnitude of the T cell responses to each self-peptide was correlated with that of the corresponding microbial peptides, the responses to the GNS or FLNA peptide strongly correlated with reactivity to each of the 2 microbial peptides (in each instance, $P < 0.0001$) (Figure 7, C and D). Therefore, the stronger the response to the microbial peptides, the greater the response to the self-peptide. In contrast, PBMCs from 10 HC subjects did not show a correlation with any of the self- or microbial peptides (Figure 7, C and D).

Of the 24 RA patients tested with the single peptides of GNS and FLNA (Figure 7), 17 were initially analyzed for T cell reactivity to pools of 4 peptides derived from these proteins (Figure 1). Only 4 of the 17 patients responded to the pool of peptides and not to the single peptide, suggesting that the majority of patients had reactivity to the single peptide epitope. Only 2 of the 17 patients responded to the single peptides, but failed to respond to the peptide pools.

Using an in silico prediction method (Immune Epitope Database [IEDB] Analysis Resource tool; http://tools.immuneepitope.org/mhcii/), the GNS peptide and the corresponding microbial peptides were predicted to bind HLA-DR molecules encoded by SE alleles with significantly higher affinity than non-SE alleles (Figure 8A). In addition, there was a trend toward greater affinity of SE binding of the FLNA peptide and the corresponding microbial peptides (Figure 8B). Of the 24 patients with RA, 15 (62%) had SE alleles. Consistent with the IEDB binding predictions, 9 of 11 patients (82%) with self- and microbial T cell reactivity had SE alleles compared with 5 of 13 patients (38%) without T cell responses to these antigens ($P = 0.05$). Thus, patients who had reactivity to the self-peptides often responded to the corresponding microbial peptides; the magnitude of the self-responses showed a significant correlation with the microbial responses, and these were more frequent in patients with SE alleles.

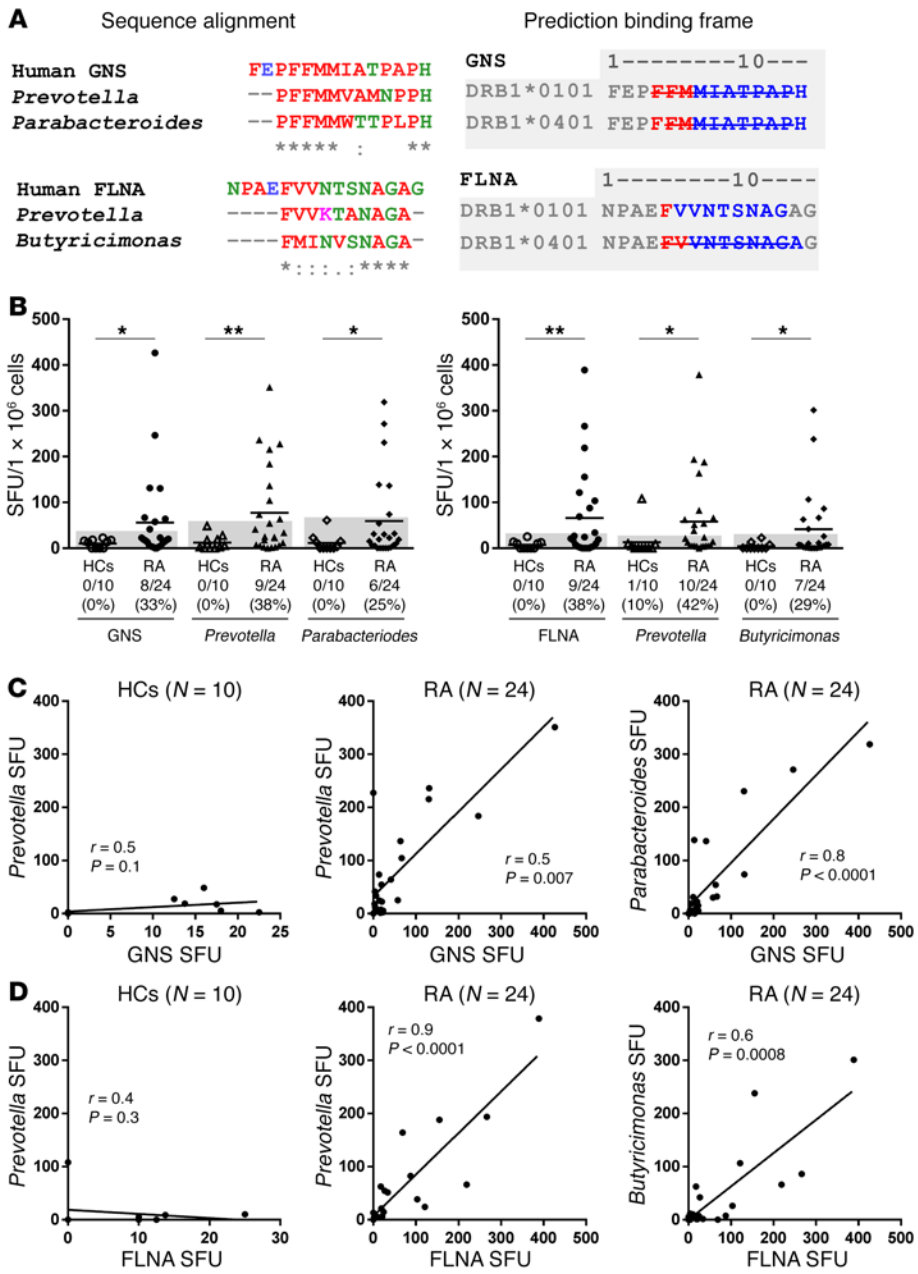


Figure 7. Sequence homology between self- and microbial peptides. (**A**) Sequence alignment of the self- and corresponding microbial peptides is shown (Clustal Omega), and the predicted binding frames of the self-peptides are given for the HLA-DRB1*0101 and *0401 molecules. Red residues indicate the P1 position (TEPITOPE predicted 3 binding registers for GNS [both HLA-DRB1*0101 and *0401], 2 for FLNA [HLA-DRB1*0401], and 1 for FLNA [HLA-DR*0101]), and blue residues indicate positions P2 through P9. The line through the amino acid residues indicates that the peptide-binding register contains an amino acid with an R-group that may not interact favorably with one of the MHC-binding pockets. (**B**) PBMCs from 24 RA patients and 10 HCs were incubated with 1 of the 2 self-peptides (GNS or FLNA) or each of the 2 corresponding microbial peptides (1 μM each). In each assay, a positive control (phytohemagglutinin) and a negative control (no peptide) were included. The amount of IFN-γ secretion is shown, as determined by ELISpot assay. A positive response was defined as greater than 3 SD above the mean value for the HCs (area above the shaded region). Horizontal lines represent the mean values for each group. $*P < 0.05$ and $**P < 0.005$, by unpaired, 2-tailed *t* test with Welch's correction. (**C**) Correlations between the T cell reactivity to the GNS peptide and the 2 corresponding microbial peptides, 1 derived from the *Prevotella* arylsulfatase protein and the other from the *Parabacteroides* GNS protein. (**D**) Correlations between the T cell reactivity to the FLNA peptide and the 2 corresponding microbial peptides derived from 2 hypothetical proteins, 1 from the *Prevotella* sp. and the other from the *Butyricimonas* sp. *P* and *r* values shown in **C** and **D** were calculated using Spearman's correlation test.

Discussion

RA is an HLA class II–associated autoimmune disease in which mucosal immunity, often resulting from interaction with oral or gut microbes or from inhaled antigens in the lung, is hypothesized to cause autoimmune phenomena leading to joint inflammation and damage. However, the factors linking mucosal immunity to autoimmunity in joints have been unclear. In this study in which HLA-DR–presented peptides were identified directly from patients' synovial tissue or PBMCs, 2 previously unidentified self-antigens, GNS and FLNA, were shown to be targets of T and B cell responses in 52% and 56% of RA patients, respectively. Importantly, the GNS and FLNA HLA-DR–presented T cell epitopes have considerable sequence homology with *Prevotella* epitopes and with similar epitopes from several related gut commensals belonging to the same order, particularly in areas predicted to be in the HLA-DR–binding groove. Moreover, T cell responses to the corresponding microbial and self-peptides were strongly correlated, suggesting that T cell epitope mimicry may provide a potential link between mucosal immunity and immune responses in affected joints. Moreover, GNS and FLNA autoantibodies correlated with *P. copri* antibody responses. This finding might also be due to cross-reactive *Prevotella* and host protein B cell epitopes, but we currently lack the recombinant microbial proteins to test this hypothesis directly. Alternatively, this correlation might simply be a reflection of T cell help resulting from T cell epitope mimicry.

The presence of natural IgM, IgG, or IgA autoantibodies has been reported in the serum of normal control subjects and is characterized by broad reactivity to self- and microbial antigens (33, 34). However, our findings cannot be explained simply by nonspecific physiological responses. First, GNS and FLNA autoanti-

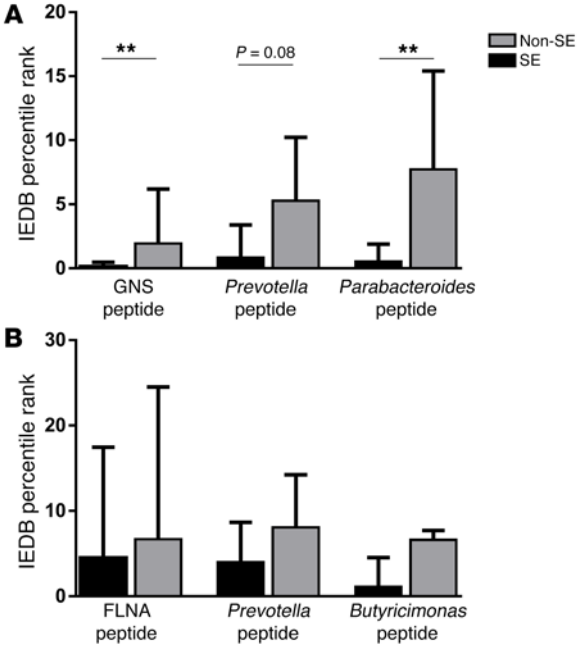


Figure 8. HLA-DR–binding prediction for self- and microbial peptides. Prediction analyses were performed using the IEDB Analysis Resource consensus tool. A low percentile rank indicates good peptide binders. (**A**) MHC class II–binding prediction of the GNS peptide and the corresponding microbial peptides. (**B**) MHC class II–binding prediction of the FLNA peptide and the corresponding microbial peptides. SE alleles: *0101, *0102, *0401, *0404, and *1001. Non-SE alleles: *0103, *0301, *0403, *0803, *1101, *1201, *1302, *1501, and *1601. Data represent median values with interquartile ranges. $^{**}P < 0.005$, by Mann-Whitney *U* test.

bodies were increased specifically in a large subgroup of patients with RA, and not in patients with other rheumatic diseases or in HCs. Second, these self-autoantibodies correlated strongly with *P. copri* antibodies, but not with *P. gingivalis* antibodies. Third, patients with RA, but not healthy subjects, often had both T and B cell responses to these self-antigens. For control of the microbiota, the generation of specific, high-affinity IgA antibody responses requires T and B cell interactions (35), whereas the "natural" antibody pool consists primarily of low-affinity polyreactive IgA antibodies, whose production is independent of T cell help (36–38).

Currently, the only known antibody responses that are specific for RA are directed against citrullinated proteins, in particular against enolase, vimentin, and fibronectin (39). Moreover, there is hypercitrullination of a range of proteins in the joints of many patients with RA (40). In our study, the GNS protein appeared to be citrullinated in vivo, and antibodies against it correlated with ACPA levels, whereas levels of antibodies against the FLNA protein, which did not appear to be citrullinated, did not correlate with ACPA levels. Moreover, HLA-DR molecules encoded by RA SE alleles bound the GNS peptide and, to a lesser degree, the FLNA peptide with higher affinity than did non-SE alleles, and SE alleles largely influence the generation of ACPAs (3).

Both FLNA and GNS self-antigens, which are often highly expressed in RA synovial tissue, represent good autoimmune targets. GNS is an enzyme located in lysosomes and is involved in the degradation and recycling of different molecules, such as glycosaminoglycans (GAGs) (https://ghr.nlm.nih.gov/gene/GNS), a major component of joint cartilage and other soft connective tissues (41). FLNA is a ubiquitous, fundamental protein for building the cell cytoskeleton and organizing the extracellular matrix (42). In addition, FLNA is required for cell-cell contact in vascular development (43). Consistent with its function, we found that FLNA protein was especially prominent in the tunica muscularis around blood vessels in RA synovial tissue and was also faintly expressed in the extracellular matrix. However, it is not yet clear why the GNS protein was seen only in a reticular pattern around blood vessels. Nevertheless, the important point is that both host proteins would be available for HLA-DR presentation in synovial tissue and both represent potential targets for autoimmune responses.

High-throughput sequencing of stool samples in 2 studies of RA patients showed gut dysbiosis (13, 14), and 1 study reported overexpansion of *Prevotella* spp., particularly *P. copri* (13). However, little is known about potential bowel pathology in RA. The major limitation of our study is the lack of specific information about pathology or immune responses in the bowel. Our study of patients with NORA was originally intended as a study of mouth flora and periodontal disease (30), and, therefore, the collection of stool samples was not a part of our protocol. Nevertheless, the finding of gut dysbiosis favoring *Prevotella* (13), the identification of *P. copri* as an immune-relevant bacterium in RA (25), and the finding here of specific T and B cell responses to *Prevotella* and corresponding GNS and FLNA epitopes suggest that mucosal immune responses in the gut may be a part of the disease in a sizable subgroup of RA patients. Although DMARD therapy may resolve dysbiosis (14) and *P. copri* overexpansion (13), immune reactivity to the organism, once triggered, appeared to persist or even increase in patients with CRA who were taking these medications. In contrast, patients with SpA or postinfectious, antibiotic-refractory LA, which may also be treated with DMARDs, did not have reactivity to this organism (25).

Taken together, our studies highlight a possible mechanism linking gut and joint inflammation. We hypothesize that dysbiosis or low-grade gut inflammation may compromise the mucosal barrier and result in leakage of commensal organisms, leading to activation of lymphocytes targeting microbial antigens. Here, we provide evidence that T cell epitopes of a related order of gut microbes, particularly *Prevotella* spp., may cross-react with self-epitopes of highly expressed proteins in joints, especially in patients with SE alleles. T cells activated by microbial peptides at the mucosal surface may then home to inflamed joints and cross-react with homologous self-antigens. In addition, as we have previously reported (25), commensal organisms or their remnants or microbe-activated antigen-presenting cells (APCs) may occasionally home to joints, where they further amplify joint inflammation and restimulate self-reactive T cells. Thus, an initial trigger in the gut mucosa may activate cross-reactive microbial immune responses that have the potential to shift to an autoimmune phenotype at sites where the homologous self-antigens are expressed. Even though proof of cross-reactivity will require the cloning of single T cells, the prominence of homologous epitopes among gut microbes and epitopes from 2 self-proteins in synovial tissue, the greater affinity of HLA-DR molecules encoded by SE alleles in binding these epitopes, and the demonstration that RA

patients often have T cell responses to corresponding microbial and self-epitopes suggest the potential for these immune responses to have a role in RA pathogenesis. We do not think that these mechanisms are the entire explanation for the marked inflammatory and proliferative responses in the RA synovial lesion. Rather, we propose that these mechanisms play a role in linking gut and joint immune responses and in amplifying synovial inflammation in a large subgroup of RA patients.

Although knowledge of bowel pathology is not yet clear in patients with RA, the identification of immune responses to *P. copri*, GNS, and FLNA has practical implications for the diagnosis and treatment of the disease. First, these immune responses appear to be specific for RA. As shown here, the addition of GNS and FLNA to standard autoantibody determinations (ACPAs and RF) increased the percentage of seropositive patients from 69% to 86%, thereby improving the diagnostic support for some patients with seronegative RA, who currently lack specific markers. Furthermore, the inclusion of citrullinated GNS in a test for ACPAs may increase the sensitivity of that test. These determinations have increased in importance, because earlier treatment with DMARDs may improve long-term outcomes, whereas a lack of diagnostic markers for seronegative patients may delay appropriate DMARD treatment, leading to less favorable clinical outcomes (44, 45). Finally, in addition to DMARDs, the identification of patients with these antibody responses may allow the testing and development of adjunctive forms of treatment, such as brief, targeted antibiotic regimens or diet alterations. Ultimately, the identification of pathogenic T cell epitopes in synovial tissue, as was done here, may make it possible to engineer blocking peptides, which would limit autoimmune stimulation and ameliorate these presumably disadvantageous autoimmune responses.

Methods

Patients and control subjects. The 101 patients in this study met the American College of Rheumatology and the European League Against Rheumatism Collaborative Initiative criteria for the diagnosis of RA (46). Of these 101 patients, 49 had NORA for 12 months or less and had not yet been treated with DMARDs. The remaining 52 patients had CRA of more than 1 year's duration (usually for many years), often treated with DMARDs. For isolation of HLA-DR–presented peptides, synovial tissue and PBMCs were obtained from patient RA1, who was undergoing arthroscopic synovectomy. To test the implicated peptides and their source proteins for immunoreactivity in additional patients, PBMCs and serum samples were collected from patients with NORA, but only serum samples were available from patients with CRA.

For other comparison groups, serum samples were collected from 28 patients with CTD, including systemic lupus, mixed CTD, scleroderma, and Sjögren's syndrome; 28 patients with SpA; and 50 patients with LA. Additionally, serum samples and PBMCs were collected from 15 healthy hospital personnel who did not have a history of RA or other autoimmune diseases, and serum samples were obtained from 40 healthy blood bank donors. HLA-DR typing was performed on blood samples from all RA or LA patients and from healthy subjects at the American Red Cross Laboratory in Dedham, Massachusetts, USA.

ELISpot T cell assay. A detailed description of the methods for the isolation and identification of HLA-DR–presented peptides is given in our previous publications (19, 20). In the current study, all nonredundant HLA-DR–presented peptides identified from patient RA1 were synthesized (Mimotopes) and tested for reactivity in sets of 3 (1 μM of each peptide) using the patient's PBMCs in IFN-γ ELISpot assays. In this initial screen, 2 self-peptides derived from GNS and FLNA were implicated; the remaining 138 peptides showed no reactivity. For testing in larger numbers of patients and control subjects, who would have a range of HLA-DR genotypes, the original peptides identified from the patient (shown in bold) and 3 additional promiscuous peptides that were predicted to bind 16 or more HLA-DR molecules were synthesized and HPLC purified at the MGH Core Facility. The GNS peptide sequences were as follows: ^{47}PNVVLLLTDDQDE^{59}, ^{222}**FEPFFMMIATPAPH**^{235}, ^{451}TYACVRTMSALWNLQ^{465}, ^{504}NYRLMMLQSCSGPTC^{518}. The FLNA peptide sequences were as follows: ^{73}DGLRLIALLEVLSQKK^{88}, ^{118}SIKLVSIDSKAIVDG^{132}, ^{416}VEVVIQDPMGQKG^{428}, ^{2446}**NPAEFVVNTSNAGAG**^{2460}.

To assess microbial sequence homology with the GNS peptide ^{222}FEPFFMMIATPAPH^{235}, the following microbial peptides were tested: ^{231}KKPFFMMVAMNPPH^{245}, which was derived from the arylsulfatase protein of the *Prevotella* spp. (WP_062433009.1) and ^{247}DVPFFMMWTTPLPH^{261}, derived from the GNS of the *Parabacteroides* sp. (WP_046148720.1). To evaluate the microbial sequence homology for the FLNA peptide ^{2446}NPAEFVVNTSNAGAG^{2460}, the following microbial peptides were tested: ^{133}QGFVVKTANAGAL^{146}, which was derived from an uncharacterized protein of the *Prevotella* sp. (WP_028897633.1) and ^{90}ANFMINVSNAGAL^{103}, derived from an uncharacterized protein of the *Butyricimonas* sp. (WP_065219401.1).

The RA patients' PBMCs were stimulated with the pool of peptides or single peptides and analyzed for reactivity to a human IFN-γ/IL-17 Double-Color ELISpot Kit (Cellular Technology Ltd.). All peptides (1 μM) were tested in duplicate wells, as were positive (phytohemagglutinin) and negative (no antigen) control samples. After 5 days, cells were transferred to ELISpot plates coated with IFN-γ/IL-17 antibodies and incubated overnight. Images of wells were captured using an ImmunoSpot Series 3B Analyzer (Cellular Technology Ltd.), and spots were counted using ImmunoSpot software. A positive T cell response was defined as 3 SD above the mean value for HC subjects.

ELISA for serum IgG and IgA anti-GNS and anti-FLNA autoantibodies. ELISA plates were coated with 0.5 μg/ml rhGNS (Novoprotein) or rhFLNA (Novusbio) overnight at 4°C. Subsequent incubations and washes were performed at room temperature. After washing with PBS containing 0.05% Tween-20 (PBST), the plates were blocked with blocking buffer (5% nonfat dry milk in PBST) for 1 hour. Afterwards, 100 μl of each patient's serum sample (diluted 50-fold) was added in duplicate wells for 1.5 hours, followed by HRP-conjugated goat anti-human IgG (sc-2453; Santa Cruz Biotechnology Inc.) or HRP-conjugated goat anti-human IgA (STAR141P; Bio-Rad), and then tetramethylbenzidine (TMB) substrate (BD). For interplate standardization, 2 control samples were included with each assay.

In vitro citrullination assay. rhGNS (2 μg) or rhFLNA (2 μg) was incubated with rhPAD4 (400 ng), which was provided by Maximilian Koenig from the laboratory of Felipe Andrade (Johns Hopkins School of Medicine, Baltimore, Maryland, USA) in 1 M Tris (pH 7.6) in the presence of 200 mM $CaCl_2$. A negative control reaction was performed, substituting 200 mM $CaCl_2$ with 200 mM EDTA. Incubation was performed at 37°C for 3 hours. Protein citrullination was determined by anti–modified citrulline immunoblotting, according to the manufacturer's recommendations (EMD Millipore).

Quantification of GNS and FLNA protein levels in serum and joint fluid. ELISA plates were coated with 5 μg/ml capture antibody (sc-161669 for GNS and sc-58764 for FLNA; Santa Cruz Biotechnology Inc.) overnight at 4°C. Plates were then washed with PBS and blocked with PBST containing 5% milk (blocking buffer) for 1 hour. Afterwards, serum or joint fluid samples (diluted 1:10 in blocking buffer) were added in duplicate and incubated for 2 hours at room temperature. Next, a detection antibody (5 μg/ml) (SAB1410557 for GNS from Sigma-Aldrich; sc-28284 for FLNA from Santa Cruz Biotechnology Inc.) was added for 2 hours. The plates were then incubated with goat anti-rabbit IgG-HRP (sc-2030; Santa Cruz Biotechnology Inc.), followed by TMB substrate. For interplate standardization, 2 control samples were included on each plate.

Immunohistochemistry. Fresh-frozen synovial tissue samples were stained for GNS and FLNA proteins. After blocking, the sections were incubated with a rabbit polyclonal antibody against GNS (SAB1410557; Sigma-Aldrich) and a mouse monoclonal antibody against FLNA (sc-17749; Santa Cruz Biotechnology Inc.) at 4°C overnight. For negative controls, nonspecific rabbit or mouse IgG antibodies (Sigma-Aldrich) were used. The following day, the sections were incubated with a biotinylated anti-rabbit (HK336-5R; Biogenex) or anti-mouse (HK335-5M; Biogenex) secondary antibody, peroxidase-streptavidin, and then a diaminobenzidine substrate. The slides were counterstained with Mayer's hematoxylin and mounted with Permount Mounting Medium (Fisher Scientific). Microscopic images were obtained with a Zeiss widefield microscope.

In silico determinations of HLA-DR–binding affinity. The affinity of HLA-DR binding of the GNS, FLNA, and related microbial peptides was determined using the T cell epitope – MHC class II molecules binding prediction tool from the IEDB Analysis Resource. This tool uses different methods to predict MHC class II epitopes, including a consensus approach that combines NN-align, SMM-align, and combinatorial library methods. For the analysis, all HLA-DR alleles found in the RA cohort were analyzed, including 5 SE alleles and 9 non-SE alleles.

Statistics. Categorical data were analyzed by Fisher's exact test, and quantitative data were analyzed using an unpaired, 2-tailed *t* test with Welch's correction. Correlations were determined using Spearman's correlation test. All analyses were performed using GraphPad Prism 6 (GraphPad Software). All *P* values are 2-tailed. *P* values of 0.05 or less were considered statistically significant.

Study approval. The present study was reviewed and approved by the Human Investigations Committee at MGH (2008 to 2014). All patients and control subjects provided written informed consent prior to their participation in the study.

Author contributions

AP designed and performed the experiments and analyzed the data. SLA and ACS enrolled and cared for patients in the study, handled clinical data collection, and performed clinical correlations. QW and CEC performed mass spectrometric analyses. EED, KS, and ACS provided advice and data analysis interpretation. All authors contributed to the preparation of the manuscript.

Acknowledgments

We thank Deborah Collier and Marcy Bolster (MGH, Boston, MA) for help with patient care; Maximilian Koenig and Felipe Andrade (Johns Hopkins University, Baltimore, MD) for providing the human PAD4 enzyme for citrullination assays; Gunnlaugur Petur Nielsen (MGH, Boston, MA) for help with the description of synovial immunohistology; Dennis Burke (MGH, Boston, MA) and John Aversa (Yale University, New Haven, CT) for help with obtaining synovial tissue; Mandakolathur Murali (MGH, Boston, MA) for RF and ACPA analyses; the American Red Cross laboratory for HLA-DR typing of patients; and Katherine Sulka (MGH, Boston, MA) for help with the preparation of samples for HLA-DR typing. This work was supported by grants from the American College of Rheumatology Innovative Grant Program "Within our Reach, Finding a Cure for RA;" the Ounsworth-Fitzgerald Foundation; the Mathers Foundation; the English, Bonter, Mitchell Foundation; the Littauer Foundation; the Lillian B. Davey Foundation; the Eshe Fund (to ACS); and the NIH (P41 GM104603, S10 RR020946, and S10 OD010724, to CEC). SLA received support from a Scientist Development Award from the Rheumatology Research Foundation.

Address correspondence to: Annalisa Pianta, Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, CNY 149/8301, 55 Fruit Street, Boston, Massachusetts 02114, USA. Phone: 617.726.1530; E-mail: apianta@mgh.harvard.edu.

EED's present address is: Agenus Inc, Lexington, Massachusetts, USA.

QW's present address is: Waters Corporation, Milford, Massachusetts, USA.

1. Li Y, et al. Deficient activity of the nuclease MRE11A induces T cell aging and promotes arthritogenic effector functions in patients with rheumatoid arthritis. *Immunity*. 2016;45(4):903–916.
2. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum*. 1987;30(11):1205–1213.
3. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. *N Engl J Med*. 2011;365(23):2205–2219.
4. Nishimura K, et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. *Ann Intern Med*. 2007;146(11):797–808.
5. Schellekens GA, de Jong BA, van den Hoogen FH, van de Putte LB, van Venrooij WJ. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. *J Clin Invest*. 1998;101(1):273–281.
6. Suzuki K, et al. High diagnostic performance of ELISA detection of antibodies to citrullinated antigens in rheumatoid arthritis. *Scand J Rheumatol*. 2003;32(4):197–204.
7. del Puente A, Knowler WC, Pettitt DJ, Bennett PH. The incidence of rheumatoid arthritis is predicted by rheumatoid factor titer in a longitudinal population study. *Arthritis Rheum*. 1988;31(10):1239–1244.
8. Majka DS, et al. Duration of preclinical rheumatoid arthritis-related autoantibody positivity increases in subjects with older age at time of disease diagnosis. *Ann Rheum Dis*. 2008;67(6):801–807.
9. Nielen MM, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum*. 2004;50(2):380–386.
10. Bachmaier K, Neu N, de la Maza LM, Pal S, Hessel A, Penninger JM. Chlamydia infections and heart disease linked through antigenic mimicry. *Science*. 1999;283(5406):1335–1339.
11. Miller SD, et al. Persistent infection with Theiler's virus leads to CNS autoimmunity via epitope spreading. *Nat Med*. 1997;3(10):1133–1136.
12. Zhao ZS, Granucci F, Yeh L, Schaffer PA, Cantor H. Molecular mimicry by herpes simplex virus-type 1: autoimmune disease after viral infection.

Science. 1998;279(5355):1344–1347.
13. Scher JU, et al. Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. *Elife*. 2013;2:e01202.
14. Zhang X, et al. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat Med*. 2015;21(8):895–905.
15. Maeda Y, et al. Dysbiosis contributes to arthritis development via activation of autoreactive T cells in the intestine. *Arthritis Rheumatol*. 2016;68(11):2646–2661.
16. Hemmer B, et al. Predictable TCR antigen recognition based on peptide scans leads to the identification of agonist ligands with no sequence homology. *J Immunol*. 1998;160(8):3631–3636.
17. Horwitz MS, Bradley LM, Harbertson J, Krahl T, Lee J, Sarvetnick N. Diabetes induced by Coxsackie virus: initiation by bystander damage and not molecular mimicry. *Nat Med*. 1998;4(7):781–785.
18. Ehl S, et al. Viral and bacterial infections interfere with peripheral tolerance induction and activate CD8+ T cells to cause immunopathology. *J Exp Med*. 1998;187(5):763–774.
19. Seward RJ, Drouin EE, Steere AC, Costello CE. Peptides presented by HLA-DR molecules in synovia of patients with rheumatoid arthritis or antibiotic-refractory Lyme arthritis. *Mol Cell Proteomics*. 2011;10(3):M110.002477.
20. Wang Q, et al. Immunogenic HLA-DR-presented self-peptides identified directly from clinical samples of synovial tissue, synovial fluid, or peripheral blood in patients with rheumatoid arthritis or Lyme arthritis. *J Proteome Res*. 2017;16(1):122–136.
21. Drouin EE, et al. A novel human autoantigen, endothelial cell growth factor, is a target of T and B cell responses in patients with Lyme disease. *Arthritis Rheum*. 2013;65(1):186–196.
22. Crowley JT, et al. A highly expressed human protein, apolipoprotein B-100, serves as an autoantigen in a subgroup of patients with Lyme disease. *J Infect Dis*. 2015;212(11):1841–1850.
23. Crowley JT, et al. Matrix metalloproteinase-10 is a target of T and B cell responses that correlate with synovial pathology in patients with antibiotic-refractory Lyme arthritis. *J Autoimmun*. 2016;69:24–37.
24. Pianta A, et al. Annexin A2 is a target of autoimmune T and B cell responses associated with synovial fibroblast proliferation in patients with antibiotic-refractory Lyme arthritis. *Clin Immunol*. 2015;160(2):336–341.
25. Pianta A, et al. Evidence for immune relevance of Prevotella copri, a gut microbe, in patients with rheumatoid arthritis. *Arthritis Rheumatol*. 2017;69(5):964–975.
26. Sturniolo T, et al. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol*. 1999;17(6):555–561.
27. Mikuls TR, et al. Periodontitis and Porphyromonas gingivalis in patients with rheumatoid arthritis. *Arthritis Rheumatol*. 2014;66(5):1090–1100.
28. Sievers F, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 2011;7:539.
29. Yu CS, Lin CJ, Hwang JK. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci*. 2004;13(5):1402–1406.
30. Arvikar SL, et al. Clinical correlations with Porphyromonas gingivalis antibody responses in patients with early rheumatoid arthritis. *Arthritis Res Ther*. 2013;15(5):R109.
31. Benjdia A, Martens EC, Gordon JI, Berteau O. Sulfatases and a radical S-adenosyl-L-methionine (AdoMet) enzyme are key for mucosal foraging and fitness of the prominent human gut symbiont, Bacteroides thetaiotaomicron. *J Biol Chem*. 2011;286(29):25973–25982.
32. Ulmer JE, et al. Characterization of glycosaminoglycan (GAG) sulfatases from the human gut symbiont Bacteroides thetaiotaomicron reveals the first GAG-specific bacterial endosulfatase. *J Biol Chem*. 2014;289(35):24289–24303.
33. Coutinho A, Kazatchkine MD, Avrameas S. Natural autoantibodies. *Curr Opin Immunol*. 1995;7(6):812–818.
34. Dighiero G. Natural autoantibodies, tolerance, and autoimmunity. *Ann N Y Acad Sci*. 1997;815:182–192.
35. Kroese FG, Butcher EC, Stall AM, Lalor PA, Adams S, Herzenberg LA. Many of the IgA producing plasma cells in murine gut are derived from self-replenishing precursors in the peritoneal cavity. *Int Immunol*. 1989;1(1):75–84.
36. Macpherson AJ, Gatto D, Sainsbury E, Harriman GR, Hengartner H, Zinkernagel RM. A primitive T cell-independent mechanism of intestinal mucosal IgA responses to commensal bacteria. *Science*. 2000;288(5474):2222–2226.
37. Suzuki K, Ha SA, Tsuji M, Fagarasan S. Intestinal IgA synthesis: a primitive form of adaptive immunity that regulates microbial communities in the gut. *Semin Immunol*. 2007;19(2):127–135.
38. Kubinak JL, Round JL. Do antibodies select a healthy microbiota? *Nat Rev Immunol*. 2016;16(12):767–774.
39. Klareskog L, Rönnelid J, Lundberg K, Padyukov L, Alfredsson L. Immunity to citrullinated proteins in rheumatoid arthritis. *Annu Rev Immunol*. 2008;26:651–675.
40. Romero V, et al. Immune-mediated pore-forming pathways induce cellular hypercitrullination and generate citrullinated autoantigens in rheumatoid arthritis. *Sci Transl Med*. 2013;5(209):209ra150.
41. Wang JY, Roehrl MH. Glycosaminoglycans are a potential cause of rheumatoid arthritis. *Proc Natl Acad Sci U S A*. 2002;99(22):14362–14367.
42. Kim H, McCulloch CA. Filamin A mediates interactions between cytoskeletal proteins that control cell adhesion. *FEBS Lett*. 2011;585(1):18–22.
43. Feng Y, et al. Filamin A (FLNA) is required for cell-cell contact in vascular development and cardiac morphogenesis. *Proc Natl Acad Sci U S A*. 2006;103(52):19836–19841.
44. Kuriya B, et al. Earlier time to remission predicts sustained clinical remission in early rheumatoid arthritis--results from the Canadian Early Arthritis Cohort (CATCH). *J Rheumatol*. 2014;41(11):2161–2166.
45. van Nies JA, Tsonaka R, Gaujoux-Viala C, Fautrel B, van der Helm-van Mil AH. Evaluating relationships between symptom duration and persistence of rheumatoid arthritis: does a window of opportunity exist? Results on the Leiden early arthritis clinic and ESPOIR cohorts. *Ann Rheum Dis*. 2015;74(5):806–812.
46. Aletaha D, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum*. 2010;62(9):2569–2581.