Antigen selection in the IgE response of allergic and nonallergic individuals

Isabell Dahlke, PhD,[a] David J. Nott, PhD,[a,b] John Ruhno, MBBS FRACP,[c] William A. Sewell, MBBS PhD,[d] and Andrew M. Collins, PhD[a] *Sydney, Australia*

Background: Affinity maturation within germinal centers should usually lead to an accumulation of replacement mutations in complementarity-determining regions (CDRs) of Ig genes as a result of antigen selection. A number of studies have suggested, but not statistically demonstrated, that antigen selection might not guide such an accumulation of replacement mutations in allergic IgE sequences. This has been suggested to reflect the nature of allergens themselves or of the allergic response.

Objective: We sought to investigate the role of antigen selection in the evolution of the IgE response by mean of analysis of Ig sequences derived from both allergic and nonallergic individuals.

Methods: IgE sequences were amplified from peripheral blood of allergic and nonallergic individuals by using seminested RT-PCR. Additional IgE and IgG sequences were obtained from public databases. Analysis considered replacement mutations in the CDRs as a proportion of total mutations and compared IgE sequences with IgG sequences.

Results: The nonallergic IgE sequences were significantly less mutated than both the allergic IgE ($P < .001$) and IgG ($P < .01$) sequences. There was a low proportion of replacement mutations in the CDRs of both nonallergic and allergic IgE sequences and a significantly increased proportion of such mutations in IgG sequences ($P < .001$).

Conclusions: IgE antibodies in both nonallergic and allergic individuals appear to accumulate few somatic point mutations as a consequence of antigen selection.

Clinical implications: Allergic and nonallergic IgE responses might often develop along a common pathway that is distinct from the conventional germinal center reaction through which the IgG response develops. (J Allergy Clin Immunol 2006;117:1477-83.)

Key words: *Human, B cells, IgE antibodies, allergy, repertoire development, antigen selection*

Abbreviations used
CDR: Complementarity-determining region
FR: Framework region
IGHV: Immunoglobulin heavy chain variable
IMGT: ImMunoGeneTics
M_V: Total number of IGHV gene mutations
R_{CDR}: IGHV replacement mutations in the CDR

A hallmark of allergic disease is the development of IgE antibodies directed against innocuous environmental antigens, and it is generally accepted that the IgE response emerges in parallel with that of the more abundant IgG and IgA isotypes; that is, within the germinal center, some antigen-selected B cells undergo class switching to the various isotypes. During expansion of the responding cell population, these switched cells are subjected to the targeted mutation process that gives rise to an accumulation of somatic point mutations within the variable regions of the Ig genes.[1] This mutation process leads to the production of higher affinity antibodies as a consequence of selection of mutations that lead to improved antigen binding.[2]

It has been argued that this process of selection should favor replacement mutations rather than silent mutations within the complementarity-determining regions (CDRs).[3] In contrast, the need to select against mutations that compromise the structural integrity of Igs should lead to selection pressures against many replacement mutations in sequences encoding the framework regions (FRs). Analysis of the distribution of replacement and silent mutations along immunoglobulin heavy chain variable (IGHV) gene sequences of antigen-selected B cells should therefore show higher than expected replacement/silent mutation ratios in the CDR and lower than expected ratios in the FR. Such selection pressures have generally been assumed to act on Ig genes in cells expressing antibodies of all isotypes.

The possibility that mutational features of antigen selection are absent from IgE sequences has been raised in a number of studies.[4-6] These claims have been made, however, after consideration of very few IgE sequences, and statistical confirmation of the claims have not been provided in any of these reports. In fact, the suitability of the analytic approach taken in these studies has been challenged.[7] In this study we report a rigorous approach to the statistical inference of antigen selection from the distribution of somatic point mutations. In addition, we

From [a]the School of Biotechnology and Biomolecular Sciences, [b]the School of Mathematics, and [d]the Garvan Institute of Medical Research and St Vincent's Clinical School, University of New South Wales, and [c]Royal North Shore Hospital.

Supported by a grant from the National Health and Medical Research Council.
Disclosure of potential conflict of interest: The authors have declared that they have no conflict of interest.
Received for publication June 14, 2005; revised December 19, 2005; accepted for publication December 28, 2005.
Available online March 31, 2006.

Reprint requests: Andrew Collins, PhD, School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW 2052, Australia. E-mail: a.collins@unsw.edu.au.

0091-6749/$32.00

doi:10.1016/j.jaci.2005.12.1359

report sensitive techniques for the study of the IgE sequences that have allowed us to investigate IgE sequences from nonallergic individuals for the first time. We present results suggesting that the IgE response in both allergic and nonallergic individuals is fundamentally different in character from the IgG response. Although biases in the distribution of replacement mutations are readily apparent in sets of IgG sequences, neither allergic nor nonallergic IgE sequences show a similar degree of mutational evidence of antigen selection.

METHODS

Patient samples

After informed consent, peripheral blood was collected from 2 individuals with no clinical symptoms or history of allergy or parasitic disease. They were determined to be nonallergic on the basis of negative serum allergen-specific IgE (RAST) assay and skin prick test results to an extensive array of important local allergens (house dust mite, cockroach, cat epithelia, dog epithelia, *Alternaria alternata*, *Aspergillus fumigatus*, *Plantain lanceolata*, *Lolium perenne*, couch grass, cod, potato, sesame, peanut, hazelnut, soybean, wheat, yeast, milk, and egg), and they had total serum IgE concentrations of 21 IU/mL and 93 IU/mL, which are within the reference range (<190 IU/mL). Blood was also collected from an individual with atopic dermatitis who had multiple severe allergic sensitivities and a total IgE concentration of greater than 5000 IU/mL. RAST assays demonstrated IgE reactivity to the house dust mite *Dermatophagoides pteronyssinus* (>100 kUA/L), rye grass pollen (>100 kUA/L), cat dander (34.1 kUA/L), *A alternata* (4.9 kUA/L), peanut (24.4 kUA/L), shrimp (5.8 kUA/L), egg white (8.6 kUA/L), and egg yolk (7.1 kUA/L). All work performed was approved by the University of New South Wales Human Research Ethics Committee.

RNA extraction, cDNA synthesis, and PCR amplification

Mononuclear cells were prepared from peripheral blood samples by means of density gradient centrifugation, and total cellular RNA was extracted by using the method of Chomczynski and Sacchi.[8] Synthesis of cDNA was performed with Superscript II Reverse Transcriptase (Invitrogen, Carlsbad, Calif), according to the manufacturer's protocol. In brief, 1 μg of total RNA, 500 ng of oligo(dT)$_{12\text{-}18}$ primers (Promega, Madison, Wis), 0.2 mM dNTP Mix (Promega), 5 mM dithiothreitol, 40 units of RNase Inhibitor (Promega), and 100 units of Superscript II were incubated for 2.5 hours at 42°C and then for 15 minutes at 70°C. IgE heavy chain gene rearrangements were then amplified by using a seminested PCR. For the first PCR (PCR1), specific primers for each IGHV gene family (IGHV1-IGHV6) were used as forward primers in separate reactions, together with an outer IgE constant region 1 (CH1) reverse primer (Cε1 reverse, GCT ACT AGT TTT GTT GTC GAC CCA GTC).[9] The gene family–specific primers were as follows: IGHV1 forward, CAC TCC CAG GTG CAG CTG GTG CAG TCT GG; IGHV2 forward, CAG GTC ACC TTG AAG GAG TCT GG; IGHV3 forward, GTC CAG TGT GAG GTG CAG CTG GTG GAG TCT GG; IGHV4 forward, GTC CTG TCC CAG GTG CAG CTG CAG GAG TCG GG; IGHV5 forward, GTC TGT GCC GAG GTG CAG CTG CTG CAG TCT GG; and IGHV6 forward, GTC CTG TCA CAG GTA CAG CTG CAG CAG TCA GG. In the second PCR aliquots of PCR1 products were amplified by a combination of the same IGHV gene family forward primers and an inner IgE CH1 reverse primer (GGG GAA GAC GGG TGG GCT CTG TGT GG). PCR amplification was performed with approximately 40 ng of cDNA, 0.6 μM of each primer, 1.5 mM $MgCl_2$, 0.2 units of *Taq* polymerase (Promega), and a buffer supplied by the manufacturer. In a control PCR the β-actin gene was amplified by using β-actin forward (CCA ACT GGG ACG ACA TG) and reverse (CAG GGA TAG CAC AGC CT) primers. Cycling in both PCRs was 94°C for 5 minutes, followed by 94°C for 30 seconds, 54°C for 50 seconds, and 72°C for 1.5 minutes for 35 cycles and then 72°C for 10 minutes on a Tpersonal 48 cycler (Biometra, Gottingen, Germany). PCR products were purified by means of agarose gel electrophoresis and a gel extraction kit (Qiagen, Chatsworth, Calif) and then ligated into the pGEM-T Easy Vector (Promega) and transformed into *Escherichia coli* JM109 cells (Promega), according to the manufacturer's protocol. Random colonies were picked, their plasmids were prepared (Qiagen), and the inserts were sequenced.

V gene sequencing and V gene sequence alignment

Cloned IgE heavy chain inserts were sequenced with a Dye Terminator Kit (Applied Biosystems, Foster City, Calif) with 200 ng of plasmid and 3.2 pmol of M13 forward (GTT TTC CCA GTC ACG ACG) or M13 reverse (CAT GAT TAC GCC AAG CTA TT) primer. Sequences were analyzed on an ABI 3730 automated sequencer (Applied Biosystems) by using the following cycling conditions: 96°C for 10 seconds, 50°C for 5 seconds, and 60°C for 4 minutes for 25 cycles.

IgG sequences and additional IgE sequences were identified from the ImMunoGeneTics (IMGT) Ig sequences database.[10] One hundred fifty-two IgG sequences were selected after exclusion of disease-related sequences, nonproductive sequences, and sequences including ambiguous nucleotides. One hundred twenty-eight IgE sequences from allergic individuals were identified in the IMGT[10] and Genbank databases. The IgE sequences were isolated from individuals with allergic rhinitis,[11,12] dermatitis,[4,13] peanut allergy,[14] and asthma.[5,15,16] All IgE and IgG sequences used in this study are detailed in Tables E1 and E2, respectively, which are available in the Online Repository at www.jacionline.org. All sequences were complete for the region beginning at codon 10 and ending at the 3′ codon of the FR3 region (codon 104) by using the IMGT numbering system.[17]

IgE and IgG sequences were aligned with the IMGT VQUEST program,[10] and all silent and replacement mutations in the sequences, from codon 10 to codon 104, were noted. Where more than one mutation affected a codon, each mutation was treated as an independent event, and all possible mutational pathways to the final sequence were computed. The number of replacement mutations scored for that codon was then calculated as the mean of the number of replacement mutations seen in the various pathways. For analysis of antigen selection, CDRs were defined to encompass codons included in the CDRs of both the IMGT[17] and Kabat[18] definitions. The CDR1 region was therefore defined to include nucleotides from codon 27 to codon 40, and the CDR2 region was defined to include those from codon 55 to codon 74.

Statistical analysis of somatic point mutations was conducted by using Pearson goodness-of-fit tests, with P values calculated by using a simulation-based method for testing the level of total mutations in the different groups. We also examined the way that the probability of replacement mutations in CDR depends on group (IgG, allergic IgE, or nonallergic IgE) and on the total number of mutations by fitting a quasibinomial model.[19] The distribution of replacement mutations was also compared with mutations seen in a model of the mutation process in which patterns of mutation are determined entirely by the location of mutational hotspots. These hotspots were determined by reference to trinucleotide mutability scores, as previously described,[20] and the relative mutability of the FR and