IgG5, the Tenth Member of the Human Immunoglobulin Family

Zeinab Dalloul[1], Iman Dalloul[1], Christelle Oblet[1], Virginie Pascal[1], Alexis Saintamand[1], François Boyer[1], Brice Laffleur, Stéphanie Durand-Panteix[1], Jean-Claude Aldigier[1], Yannic Danger[2], Yolla El Makhour, Jeanne Cook-Moreau[1] & Michel Cogné[1]

(Article en préparation)

Les immunoglobulines constituent des éléments clés de l'immunité adaptative. Outre les IgM et IgD ancestrales, un répertoire de régions constantes étendues, avec les propriétés effectrices supplémentaires des anticorps IgG, IgA et IgE, a évolué chez les mammifères. Ceci s'est terminé chez l'Homme avec 9 classes et sous-classes différentes: IgM, IgD, IgG1, IgG2, IgG3, IgG4, IgE, IgA1 et IgA2. La structure et la fonction des sous-classes d'IgG varient fortement en raison de changements de séquence au niveau des résidus importants pour l'activation du complément, du recrutement de la cytotoxicité cellulaire et du repliement de la molécule.

Les cassures chromosomiques qui permettent le switch chez l'Homme se situent dans une région relativement proche du gène constant concerné, la longueur du fragment ciblé par les recombinaisons étant inférieure à 10kb (la plus longue région « switch » humaine, Sγ1, mesure 5 kilobases) et se termine à quelques centaines de bases de la séquence codante du CH1 de Cγ1. Le gène Ψγ est très particulier car aucune séquence répétitive n'est observée au sein des 10 kilobases qui le précèdent. Cette absence de région switch « canonique », et l'absence d'évidence de son expression l'ont ainsi fait classer parmi les pseudogènes, alors que l'ensemble de sa structure exon/intron semble pourtant conforme à celle d'un gène fonctionnel. Ce gène CΨγ est précédé par le *super-enhancer* 3'RR1 qui se trouve à environ 16 kb du début de l'exon CH1 Ψγ. Des cassures affectent ce superenhancer 3'RR du locus IgH lors de recombinaisons suicides du locus (LSR), mais la distance importante qui les sépare du gène Ψγ n'autorise sans doute pas un épissage normal de l'exon VDJ sur le gène Ψγ après une recombinaison de ce type. Comme le laboratoire s'intéresse aux recombinaisons LSR, nous avons décidé d'explorer aussi la présence de cassures éventuelles affectant la région située immédiatement en amont de Ψγ. En effet, tels évènements créeraient un locus recombiné très voisine de celle d'un switch classique vers les gènes Cγ1, 2, 3 ou 4. Nous avons donc recherché des événements de type « CSR Ψγ » joignant Sμ à des sites

accépteurs situés dans les 10 kb immédiatement flanquant en 5 ’de Ψγ (dans une région cible donc située entre le superenhancer 3’RR et le gène Ψγ). Cette exploration nous a en effet permis d’identifier une certaine diversité et une fréquence non négligeable de telles jonctions. Notre méthode de recherche a reposé sur plusieurs essais de PCR avec différents oligonucléotides, soit situés dans la région spécifique en amont de Ψγ, soit en utilisant une stratégie consensuelle ancrée sur une séquence conservée au sein de la région constante des 5 gènes Cγ humains. Parmi toutes les ruptures en amont d'un gène Cγ, nous évaluons globalement à environ 10^{-5} la fréquence de celles survenant dans les 10 kb adjacents en amont de Ψγ.

La question suivante était de savoir si telles recombinaisons peuvent permettre la production de transcrits fonctionnels V_HDJ_H-CΨγ. Afin de les rechercher malgré leur « dilution » parmi tous les abondants transcrits V_HDJ_H-Cγ fonctionnels, nous avons fait le design d’un essai de RACE-PCR dédié à l'amplification et la validation spécifiques de la séquence Ψγ. Grâce à cet essai, nous avons pu caractériser un certain nombre de transcrits. Leur existence est également retrouvée au sein de données RNAseq. Par les 2 méthodes, l’abondance de ces transcrits est bien inférieure (avec un facteur de l’ordre de 10^5) à celle des transcrits classiques V_HDJ_H-Cγ de chaîne IgH. Le répertoire de ces transcrits V_HDJ_H-CΨγ se révèle cependant diversifié et l'utilisation de VH ne diffère pas de manière significative de celle des autres transcrits Cγ. Il est à noter que dans notre analyse, le niveau de SHM dans ces transcrits était inférieur à celui des transcrits Cγ classiques. Par rapport à un faible nombre de transcrits fonctionnels VDJ-Ψγ, des transcrits non fonctionnels (non épissés) provenant de cette région du locus IgH ont également été identifiés, ce qui indique que les méthodes capables de détecter les rares transcrits fonctionnels nous imposent de descendre à un tel niveau de sensibilité que le bruit de fond des transcrits germinaux du locus non réarrangés devient conséquent. Ceci révèle donc aussi bien sur le fait que le gène Ψγ est le siège d’une transcription germinale non réarrangée (comme tous les gènes constants du locus IgH). On notera d’ailleurs avec intérêt que d’autres gènes « classiques » du locus IgH sont dans le même cas de figure, comme le gène Cε, pour lequel les transcrits non épissés (a priori donc « germinaux ») sont abondants au sein des données RNAseq issus de cellules B mémoires humaines activées, alors que les transcrits fonctionnels VDJ-Cε y sont dans notre expérience indétectables. L’abondance relative des transcrits germinaux CΨγ a clairement l’inconvénient de rendre plus difficile la détection des transcrits fonctionnels VDJ-CΨγ en les « diluant » encore un peu plus parmi tous les transcrits humains Cγ.

Nous nous sommes aussi intéressés à la protéine « IgG5 » qui résulte de la traduction des transcrits fonctionnels VDJ-CΨγ. Nous avons identifié sa signature au sein des données de spectrométrie de masse correspondant à différents fluides biologiques humains (plasma, urine). Nous avons aussi produit des formes recombinantes de cette IgG5 dont les propriétés se sont révélées tout à fait surprenantes. En effet cette IgG partage avec les autres IgG des capacités effectrices d'activation du complément ou de l'ADCC, mais elle montre par contre une faculté unique à s'autocliver en fragments Fc + Fab. La faible quantité d'IgG5 produite peut faire douter de son rôle sous forme circulante, alors que son taux est vraisemblablement inférieur d'au moins 10^5 fois aux autres IgG, soit de l'ordre de 100µg/l (ce qui reste quand même un taux équivalent à celui de l'IgE humaine). Par contre, à proximité d'une cellule ayant « switché » vers Ψγ puis différenciée en plasmocyte, la présence d'une IgG5 auto-clivable en segments Fc+Fab serait attendue comme compétitrice de l'action effectrice des autres IgG, et donc immunosuppressive.

En résumé, nous montrons maintenant qu'une cinquième sous-classe d'IgG, IgG5, est exprimée à un faible niveau chez l'Homme, avec un switch « non canonique » et un répertoire VDJ diversifié, plutôt moins muté que celui des IgG classiques. Son expression résulte d'événements de changement de classe atypiques en aval de la première région régulatrice IgH 3' (3'RR1) au sein du locus IgH. Pour vérifier davantage la fonctionnalité de l'IgG5 correspondante, des versions adaptées des anticorps thérapeutiques rituximab et trastuzumab ont été construites. Sous sa forme complète, cette molécule d'anticorps partage la plupart des propriétés fonctionnelles des quatre premières sous-classes d'IgG, y compris l'activation du complément et la cytotoxicité cellulaire dépendant des anticorps. De plus, il porte une région charnière de séquence inhabituelle et spécifique, lui conférant une capacité unique de clivage spontané en fragments Fab + Fc. L'IgG5 apparaît comme la classe d'anticorps la plus rare produite chez l'Homme et sa contribution aux réponses immunitaires est surtout susceptible d'être régulatrice.

IgG5, the Tenth Member of the Human Antibody Family

Zeinab Dalloul[1], Christelle Oblet[1*], Iman Dalloul[1*], Virginie Pascal[1], François Boyer[1], Alexis Saintamand[1], Brice Laffleur[1], Stéphanie Durand-Panteix[1], Jean-Claude Aldigier[1], Emilie Pinault[3], Yannic Danger[2], Yolla El Makhour[4], Jeanne Cook-Moreau[1] & Michel Cogné[1]

[1]Limoges University, CNRS UMR 7276, INSERM U1262, Control of the B Cell Response & Lymphoproliferation, 2 rue du Dr. Marcland, 87025 Limoges, France.

[2]INSERM U1236, Etablissement Français du Sang Bretagne, 35000 Rennes, France.

[3]Service Commun de Recherche et d'Analyse des Biomolécules de Limoges, Université de Limoges, Limoges, France.

[4]Lebanese University, Immunology Unit MICSU, Faculty of Sciences, Beirut, Lebanon.

[*]These authors equally contributed to this work

Correspondence: cogne@unilim.fr

Around 500 million years, immunoglobulins evolved as master elements of adaptive immunity in jawed vertebrates. Beside ancestral IgM and IgD already present in cartilaginous fishes, mammals developed an extended repertoire of constant regions and effector properties shaping the IgG, IgA and IgE antibody classes. Until now, 9 different human classes and subclasses have been described (IgG1, G2, G3, G4, A1, A2, M, D and E), varying in function after structural changes affecting folding of the molecule, its flexibility, polymerization, complement activation, cytotoxicity, recycling and catabolism. We now show that a fifth IgG sub-class, IgG5, is expressed at low levels in humans, with a diversified VDJ repertoire. Its expression results from atypical class-switching events downstream from the first IgH 3'regulatory region (3'RR1) within the IgH constant gene cluster. In its complete form, this antibody shares some functional properties of the other IgG subclasses, including complement activation and antibody-dependent cell cytotoxicity. By contrast, it harbors an unusual hinge region and a unique ability to spontaneous cleavage into Fc and Fab fragments. This peculiar hinge sequence is conserved in non-human old-world primates. By naturally yielding antigen-binding Fab fragments which lack the main IgG effector properties, its potential contribution to immune responses might then rather be regulatory or tolerogenic. IgG5 thus appears as the rarest antibody class produced in humans. It remains to be explored whether the complete antibody or the resulting cleaved fragments might confer an immunoregulatory role, or whether switching to expression of the γ5 gene merely constitutes a dead end terminating the participation of some activated B-cells in the immune response.

Introduction

In addition to the enormous diversity of their binding domains, immunoglobulins (Ig) have evolved in jawed vertebrates and diversified their functions under the format of different classes [1,2]. Mammalian IgG further diversified with probably at least two different subclasses in early mammals, resulting in four in the mouse as well as in humans [1]. Besides duplication of genes and divergence of their sequences as for other multigene families, gene conversion events also likely maintained some patches of homology between family members. Human IgG is a 150 to 170kDa protein equilibrating in intravascular and extravascular fluids. Due to variegated binding to FcγR receptors, IgG1, IgG2, IgG3 and IgG4 differ in cytophily and hence in ability to mediate antibody-mediated cell cytotoxicity or phagocytosis. The same is true for FcRn binding and IgG recycling determining antibody half-life in body fluids. Efficacy of IgG subclasses for complement recruitment decreases from IgG3 to IgG4 (IgG3 > IgGl > IgG2 > IgG4) [3]. While amino acid composition is 90% similar, the highest level of divergence occurs in the hinge region and the upper CH2 domain [4]. IgG3 exhibits the longest hinge region (62 amino acids), compared to 15 amino acids for IgG1, and 12 for IgG2 or IgG4 [5]. Exon shuffling accounts for this elongated IgG3 hinge region, hence encoded by a variable number of repeats of a prototypic 15 codon-long γ–hinge exon[6]. Origin of the other γ subclasses relates to an ancient partial duplication of the constant gene cluster within the IgH locus [7,8], thus split in two units: an upstream [μ, δ, γ3, γ1, Ψε, α1], followed by a downstream [Ψγ, γ2, γ4, ε and α2] set. Each of these clusters is flanked by a 3' regulatory region (3'RR)[9], and four well-known γ heavy chain genes present in the locus encode the four human IgG subclasses. Little attention has been given to a fifth homologous gene, Ψγ, which was classified as a pseudogene due to the lack of a switch region and any evidence for its expression [10]. However, its global exon-intron organization, open reading frame and polyadenylation sites are completely super-imposable to functional γ genes[11]. We thus wished to re-challenge the pseudogene status of Ψγ using multiple means. We tracked rare rearrangement events at the DNA level and corresponding functional

transcripts in human primary B-cells and we explored the functional properties of the encoded antibody product. Altogether and also based on mass spectrometry analyses of human body fluids, our data demonstrate that the Ψγ gene is expressed in humans, constituting the tenth immunoglobulin subclass and deserving to be renamed Cγ5. IgG5 additionally shows unique properties including an unusual fragile hinge domain, rapidly cleaved upon secretion and spontaneously yielding low molecular mass monovalent Fab fragments.

MATERIALS AND METHODS

Amplification of Sμ/ψγ junctions and Ion torrent next generation sequencing NGS

Sμ/ψγ junctions were amplified in triplicate from 100 ng DNA of human tonsil DNA by using nested PCR (Phusion HF polymerase; BioLabs) with the following specific primers:
PCR1: Sμ-forward1 5'-CCAGGTAGTGGAGGGTGGTA-3' / ψγ-reverse1 5'-GCGGCTGTGATTGGATTTA-3';
PCR2: Sμ-forward2 5'-CAGGGAACTGGGGTATCAAG-3' / ψγ-reverse2: 5'-CGCCCCTACCATTGAAGTAG -3'.
For global amplification of CSR junctions to Sγ1, γ2, γ3 and γ4, the same Sμ forward primers were combined with the following consensus reverse primers:
PCR1: Sγ consensus-rev1 5'- GGTCACCACGCTGCTGAG -3',
PCR2: Sγ consensus-rev2: 5'- CTTGACCAGGCAGCCCAG-3'.
Each library was prepared using 200 ng PCR2 product. Barcoded libraries with 200-pb read lengths were prepared using Ion Xpress plus Fragment Library Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Each barcoded library was mixed in equal amounts and diluted to 100 pM. Libraries were run on an Ion PI v3 chip on the Ion Proton sequencer (Life Technologies). Data were analyzed using CSReport [12].

Next-generation sequencing for repertoire analysis

Repertoire analysis was carried out on RNA from human tonsils after 5'RACE, PCR and library preparation.

RACE was performed on 500ng RNA using SMARTer RACE 5'/3' kit (Clontech) following the manufacturer's instructions and incubated at 72°C for 3 min and at 42°C for 2 min with a specific human CH2ψγ-reverse downstream primer: 5'-CCAGGGGTTCAGTTGTTGCA-3'.

Libraries were prepared on RACE products using an upstream CH1ψγ-reverse specific biotinylated primer (5'-TTGTGTCACAACATGGGGTT-3'). For PCR, DNA was denatured 30 s at 98°C and then submitted to 30 cycles consisting of 98°C for 10 s, 70°C for 30 s, 72°C for 30 s and 72°C for 5 min. PCR products were purified using MyOne C1 Streptavidin microbeads (Invitrogen). Streptavidin purified PCR products were used to perform another PCR round using the following primers: two forward primers (5'-CTAATACGACTCACTATAGGGC-3' and 5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3') and two CH1ψγ-reverse specific primers: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCGCTGTGCC -3' and 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATCGCCGCTGTGCC-3'.

For γ1, γ2, γ3 and γ4 the following consensus primers were used: Hs Cγ consensus Race: 5'-GTGTTGCTGGGCTTGTGAT-3' and Hs Cγ consensus PCR1: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCGGTTCGGGGAAGTAGTCC-3'.

Amplification was performed with Phusion® High-Fidelity DNA Polymerase (New England Biolabs) according to the following program: DNA was denatured 30 s at 98°C and then submitted to 30 cycles consisting of 98°C for 10 s, 65°C for 30 s and 72°C for 30 s, and 1 cycle at 72°C for 10 min. Adapters were added and the resulting amplicons were sequenced on an Illumina Miseq sequencer.

RNAseq on activated human B-cells

Memory B-cells were purified and activated for 4 days with IL-2. RNA was extracted and libraries were prepared for sequencing on an Illumina NextSeq 500 using 75-bp single-

end reads. After quality control and filtering, reads were aligned on the GRCh38 human genome using STAR 2.4.2a, and blasted for the occurrence of reads containing correctly spliced JH to CH1 and CH1 to Hinge exon.

Production and characterization of recombinant antibodies of the IgG5 subclass

Vectors were constructed which associated the VDJ coding regions of either rituximab (anti-human CD20) or trastuzumab (anti-human HER2) to the Ψγ/γ5 constant coding exons. The corresponding Ig expression vector also included the corresponding κ light chain coding region in order to conserve antigen-recognition. The final complete gene expression vectors G5K-CD20 or G5K-HER2 were transfected either into CHO or HEK293 cells using the ExpiCHO Expression System (Gibco) (maintained in DMEM medium, 10% heat-inactivated fetal bovine serum, 2 mM glutamine, 1 mM sodium pyruvate, 100 U/mL penicillin, 100 mg/mL streptomycin) following the manufacturer's instructions. Anti-CD20 IgG5 secreted in transfected cell supernatants was chromatographically purified using protein A/G Prepacked columns (Pierce). Concentrations of purified antibodies were determined by 280nm UV adsorption.

Western Blots

The integrity of the purified IgG5 antibody preparations was determined by sodium dodecyl sulphate polyacrylamide gel electrophoresis under reducing and non-reducing conditions using 12% StainFree gel (Biorad). After transfer, PVDF-membranes (GE Healthcare) were blocked with 5% non-fat milk, followed by incubation for 1h with goat anti-human IgG antibody (Southern Biotech). After washing with TBS, 0.1% Tween buffer, secondary horseradish peroxidase-conjugated rabbit anti-goat antibody (Dako) was added. After washing, blots were revealed with ECL substrate (Immobilon, Millipore).

Various other antibodies were used for characterizing blotted proteins and their degradation products, including anti-human Igκ chain (Southern Biotech), anti-Ig Fd

fragment (Millipore), anti-Fc fragment (Invitrogen), anti-idiotypes specific for variable domain of either rituximab (R&D Systems) or trastuzumab (R&D Systems).

Flow cytometry

Binding of recombinant antibodies to their cognate antigen was evaluated by cell cytometry using target cells expressing human CD20 (human B lymphoma cell line DHL-4, hCD20-expressing variant "EL4-CD20"), or untransfected CD20neg EL4 cells as negative control. The MCF7 cell line expressing human HER2 was used to test trastuzumab IgG5 binding capacity.

Identical amounts of the anti-CD20 IgG5 preparation or control IgG antibodies (10µg/ml) were incubated with target cells for 30 min. For anti-HER2 IgG5, an excess amount of supernatant was used. Cells were then stained with fluorescein isothiocyanate (FITC)-conjugated secondary antibodies, which were either rabbit anti-human IgG antibodies (Dako), rabbit anti-human κ light chain (Dako), mouse anti-human IgG Fd (Millipore) or anti-idiotypes specific for variable domains of either rituximab (R&D Systems) or trastuzumab (R&D Systems), and then analyzed by flow cytometry.

Immunofluorescence

EL4-CD20 cells and HER2 HER2 MCF7 ecells were first washed, resuspended in culture medium, deposited on polysine pre-treated slides (ThermoFisher) and then incubated overnight at 37^0C. Spots of adhered cells were washed twice with phosphate buffered saline (PBS) and fixed with 4% paraformaldehyde for 10 min at room temperature. Spots were washed three times with PBS, 5 min each then blocked with PBS 1% BSA (Bovine Serum Albumin) for 30 min at 37^0C in humidified chamber. Cells were than incubated with either identical amount of anti-CD20 antibodies (10µg/ml) or supernatants containing anti-HER2 antibodies for 1h at room temperature. After three washes, either unlabeled human anti-kappa (Southern Biotech) or anti-Fd (Millipore) antibodies were added and incubated for 1h at room temperature in a humidified chamber. Cells were washed and incubated with secondary anti-goat Alexa 488 antibody for 1h at room

temperature. After three washes, nuclei were stained with Dapi before mounting with Mowiol 4-88 solution (Sigma). Sections were then observed under a confocal microscope.

Cell aggregation in the presence of anti-CD20 antibodies

Cells were plated at $2.5x10^3$ cells/well in flat-bottomed 96-well plates in the presence of 10 µg/mL $IgG5_{Rituximab}$, $IgG1_{Rituximab}$ or control IgG antibodies. Cell adhesion was assessed 24 h later by examination under an optical microscope.

Cell proliferation

An MTS assay (Promega) was used to analyze the effect of anti-CD20 antibodies on B-cell growth *in vitro*. DHL-4 cells were cultured at 37°C in 96-well plates (10^4 cells in 100 µl/well). Proliferation was assessed after 48 h under various concentrations of specific or control monoclonal antibodies. The amount of viable cells was determined (in triplicate) by measuring the absorbance (490 nm) after addition of MTS for 3 h. Cells incubated in medium without antibodies were used as a control. Growth inhibition is expressed as follows: 100 - (Mean $OD_{experimental\ group}$ – Mean OD_{medium}) / (Mean $OD_{control}$ – Mean OD_{medium}) * 100.

Complement-mediated cytotoxicity (CDC)

DHL4 cells ($2x10^5$) were re-suspended in 50% normal human serum or heat inactivated normal human serum with or without anti-CD20 antibodies. Non-CD20 IgG antibody (0.01, 1 and 10µg/ml) was used as a negative control. After incubation for 4 h at 37°C, dead and viable cells were differentially stained by addition of 0.5 mg/mL propidium iodide (PI) and analyzed by flow cytometry. The percentage of specific cytolysis was calculated from cell counts using the following formula: % of specific lysis = (% PI+ experimental- % PI+basal) / (100 - % PI+basal) * 100.

Antibody-dependent cellular cytotoxicity (ADCC)

Blood samples from healthy volunteers were collected in EDTA tubes for preparation of peripheral blood mononuclear cells (PBMCs) by density gradient centrifugation using

Lymphocyte Cell Separation Media (Cedarlane). After washing, human leukocytes containing potential effector cells were adjusted to $10x10^6$ cells/ml in serum free RPMI medium (Lonza). Target DHL4 cells were resuspended at $5x10^5$, labeled with 0.5µM CFSE (Invitrogen) and incubated 20 min at 37^0C. Labeled DHL4 target cells were washed twice and mixed with leukocytes at a 1:20 ratio (Target: Effector). Anti-CD20 IgG5 (0.01, 1 or 10µg/ml) or control antibodies were added and cells were incubated 4h at 37^0C. Dead and viable cells were stained by 0.5 mg/mL propidium iodide (PI) and analyzed by flow cytometry. The percentage of specific cell lysis was calculated from cell counts using the following formula: % of specific lysis = (% PI+ experimental- % PI+basal) / (100 - % PI+basal) * 100.

In-house mass spectrometry

Peptide samples were purified on 1CC 30 mg HLB Cartridge and filtered on a 0.22µm spin column. Resulting peptides were analyzed with a nanoLC 425 in micro-flow mode (Eksigent, Dublin, CA, USA) coupled to a TTOF5600+ mass spectrometer (SCIEX, Framingham, USA) in the Information-Dependent Acquisition mode. The MS/MS data were used to query an in-house database containing the IgG1-5 sequences using the Protein Pilot software (Sciex) with Mascot software (version 2.2, Matrix Science, England). Proteins were validated if at least 2 peptides were identified with a confidence above 99%.

Identification of peptides in mass spectrometry databanks of human samples

Tryptic peptides for the predicted Cγ5 protein were compared to all those from the four classical IgG subclasses (and all their known allelic variants, as deposited in the IMGT database http://www.imgt.org/). were then searched for their occurrence in Deposited mass spectrometry databanks were queried for uniquely specific Cγ5 peptides using the PeptideAtlas algorithm https://db.systemsbiology.net/sbeams/cgi/PeptideAtlas/Search (from the Institute for Systems Biology, Seattle, WA, USA).

RESULTS

Position and structure of Sμ/ψγ junctions in human tonsil B cell DNA.

We designed a nested PCR assay based on an upstream Sμ forward primer and consensus reverse primers specific for any human Cγ gene CH1 exon and thus susceptible to amplify a switching event towards a Cγ gene. After Next generation sequencing (NGS), junctions were analyzed by the CSReport algorithm [12] in order to score γ gene usage. Within the bulk of all CSR junctions detectable in human tonsil cells and besides a majority involving the Cγ1-4 genes, diversified junctions from Sμ to ψγ also occurred (Figure 1A). The frequency of such junctions compared to all Sμ-Sγ junctions was around 3/10000. Ψγ breakpoints were scattered within a large 5' flanking fragment and mostly located between 6 and 8 kb upstream of the Ψγ5 gene CH1 exon but some also occurred in the immediate vicinity of CH1 (Figure 1B).

In comparison to junctions with classical Cγ genes, the mean distances of breaks to AID sites as well as breaks to APOBEC sites were longer for breaks upstream of Ψγ than for those in Sγ1 (Figure 1C). Those breaks involving Ψγ featured a higher level of microhomology, a likely indication that the alternate end joining (A-EJ) pathway is more often involved in junctional repair of Ψγ-CSR double strand breaks, while non-homologous end-joining (NHEJ) dominates in classical junctions towards Cγ1-4 (Figure 1D).

We conclude that, although rare and likely involving a distinct pattern of DNA repair, switch junctions do occur upstream of the human Ψγ gene.

Occurrence and repertoire of mature VDJ-ψγ transcripts

The putative repertoire diversity of functional VDJ-Ψγ transcripts was determined in RNA extracted from human tonsil cells by RACE-PCR and high-throughput repertoire sequencing using Ψγ gene-specific reverse primers, followed by validation of sequence identity of the amplified upstream portion of CH1-Ψγ. Rare, but diversified, normally

spliced and in-frame, ψγ transcripts were found in 5 different individuals. Altogether, we characterized a total of 80 unique functional VDJ regions joined to CΨγ, while in parallel >21854 VDJ-Cγ1-4 junctions were identified with classical Cγ reverse primers. Distribution of the Ψγ repertoire was similar to that of classical Cγ transcripts sequenced from the same individuals with more frequent use of IgHV4 (Figure 2A) and IgHJ4 (Figure 2B).

By contrast, the SHM load was significantly lower, with an overall 4.64% mutation rate for ψγ compared to 6.35% for the other IgGs (Figure 2C). 6.3% of these transcripts were virtually germline (below 1% of mutations) instead of 0.8% for γ1-4 transcripts.

Among all transcripts identified from human memory B-cells by RNAseq, a parallel unbiased search of correctly spliced Ψγ transcripts (correct JH-CH1 junctions, or correctly spliced hinge exons) revealed rare occurrences of mature functional Ψγ transcripts with a ratio to the equivalent mature γ1 transcripts ranging from 1/5923 to 1/ 400000 in different samples analyzed.

Therefore, several methods concur to demonstrate the existence of functional transcripts originating from the Ψγ gene, thus deserving to be considered as a functional γ5 gene from which a diversified repertoire of VDJ-γ5 transcripts is found.

Analysis of the Ψγ/γ5 coding sequence

Comparing the Ψγ/γ5 gene coding sequence to the γ1-γ4 subclasses shows its strongest similarity with Cγ1 with a noticeable conservation of several features required for IgG function (referenced below with numbering according to Edelman et al[13,14]). Notably, γ5 is identical to γ1, γ2 and γ3 for all C_H2 positions which are considered to be important for recruiting C1q (notably Ala330-Pro331, the substitution of which by serine residues in γ4 suppresses complement activation)[15]. Residues responsible for FcRn binding (*Ile*253 and *His*310 in C_H2, *His*435 in C_H3) are also preserved. Some C_H2 residues binding to FcγR1 are preserved (*Leu* 235 and *His* 268), but not all (*Val/Leu* 234 replaced by *Pro* in Cγ5).

We noticed the conservation of the major IgG N-glycosylation site, known for its role in binding to FcγRIIa, IIb and IIIa (at Asn297, in the D-E turn of the C_H2 β-pleated sheet structure).

The homologous position in the γ5 C_H3 D-E turn carried a unique additional N-glycosylation site (with *Asn*401 instead of *Asp* in other IgG subclasses), at a position identical to the C_H3 N-glycosylation sites of IgM and IgE (Figure 3A).

Besides this additional glycosylation, the major peculiarity of the Ψγ/γ5 coding sequence concerned the hinge region. This region is 17AA-long instead of 15AA for γ1, 12AA for γ2 and γ4, and 62AA for γ3 (one 17 bp motif followed by 3 repeats of 15bp). The γ5 hinge is most closely related to γ1 and to the first hinge motif of γ3, with 12 out of 17 identical residues in both cases. It might thus originate from a common ancestor of these two genes. A phylogenetic tree of the 5 human γ genes is shown in Figure 3B, and additionally shows interspecies conservation of the human Ψγ gene with the orthologous chimpanzee and gorilla Ψγ genes, the coding sequences of which appear as functional as that of humans.

Identification of IgG5 peptide signatures by mass spectrometry

We first analyzed tryptic peptides obtained by digestion of human serum Igs for the presence of IgG5 specific peptides. No specific peak appeared above background level. We then decided to analyze normal human urine samples, which are almost free of high molecular mass protein but are known to be enriched in low molecular mass protein or protein fragments, and notably contain low amounts of Ig fragments. Several IgG5 specific tryptic fragments and mostly mapping to the Fc region C_H3 domain were then detected at low frequencies (Fig 3C). Querying the PeptideAtlas mass spectrometry public databank (http://www.peptideatlas.org/) confirmed that these specific γ5 tryptic peptides have been previously identified in several human data sets corresponding to human samples and most notably human plasma.

Structural characteristics of recombinant IgG5 antibodies

Functional γ5 heavy chains with V regions identical to the therapeutic antibodies rituximab and trastuzumab were cloned into GK or pVitro vectors together with the corresponding κ light chain coding sequence and used for antibody preparation after transfection into CHO or HEK cell lines (Figure 4A). Anti-hCD20 IgG5 specificity was checked by flow cytometry on hCD20 EL4 transfected cells using anti-IgG Fc chain antibody and confirmed specific binding to hCD20 at 10 µg/ml IgG5. The mean fluorescence signal obtained with 10µg/ml IgG5 remained, however, weak and equivalent to that of 0.01 µg/ml of standard IgG1 rituximab antibody (Figure 4B).

The integrity of the purified antibody was assessed by Western blot. Although poorly efficient, only production by stable B-cell transfectants yielded intact IgG5. SDS-PAGE analysis showed the expected full-length 150 KDa assembled IgG in native conditions and normal size 50kDa heavy chain and 25kDa light chain in reducing conditions.

All production conditions usually yielding high concentrations of recombinant IgG by transient transfection of CHO or HEK293 cells generated minute amounts of intact IgG5 together with massive amounts of short sized fragments in non-reducing conditions. These fragments were identified as Fab and Fc fragments by Western blot analysis using anti-human IgG (Fc) and Fd (Figure 4C). Accordingly, SDS-PAGE analysis in reducing conditions demonstrated the presence of low amounts of the 50 KDa full-length γ5 heavy chain, together with abundant additional H chain fragments of shorter sizes and abundant 25 KDa full-length light chains.

Functional characteristics of anti-CD20 IgG5

Direct anti-proliferative effects of anti-hCD20 IgG5 were measured by an MTS assay on the DHL4 lymphoma cell line (Fig 5A). IgG5 inhibited cell growth at 10 µg/ml and to a lesser extent with concentrations of 1 and 0.01 µg/ml in comparison to IgG control antibody.

We next examined whether IgG5 antibody also activated indirect cytotoxic pathways against lymphoma cancer cells, and first evaluated anti-CD20 antibodies for their capacity to trigger CDC against the DHL4 cell line (Figure 5B). IgG5 showed significant CDC at 10 µg/ml (>80% cells killed in the presence of human serum), while this effect was lost when using heat-inactivated serum in a control experiment. An intermediate effect was seen at 1 µg/ml (>50% cytotoxicity) but CDC rapidly disappeared at lower concentrations, in contrast to anti-hCD20 IgG1 which still displayed strong CDC at a 0.01 µg/ml concentrations.

The ability of IgG5 to recruit human effector cells was also evaluated against target DHL4 cells. For this purpose, CFSE-DHL4 labeled cells were incubated with anti-hCD20 IgG5 or IgG1, and compared to an irrelevant control IgG antibody in the presence of human effector cells (Figure 5C). IgG5 only induced minor cell lysis (~8%) at the saturating antibody concentration of 10 µg/ml, whereas IgG1 showed high activity at 10, 1 and 0.01 µg/ml.

DISCUSSION

The human IgH constant gene cluster has long been considered to be made up of nine functional genes and two pseudo-genes, Ψγ and Ψε. While Ψε is clearly a pseudogene lacking a large part of the constant domains, Ψγ surprisingly carries an almost fully functional sequence, except for the fact that it is not preceded by a classical S region, as are the four other Cγ genes and was thus supposedly not expressed in humans [11].

Since we previously studied breaks occurring outside S regions (during atypical class switching events to the Cδ gene and during LSR[16,17]), and since Ψγ is strikingly preceded by the 3'RR1 super-enhancer which is a target for LSR, we decided to explore the occurrence of CSR-like events joining Sµ to the 5' flanking region of Ψγ. We indeed identified such breaks using a consensus strategy anchored on a conserved sequence of human Cγ genes. Among all breaks upstream of a Cγ gene, about 0.1‰ occurred upstream of Ψγ.

It was thus questionable whether such breaks could indeed allow expression of functional mature transcripts joining an in-frame VDJ exon with the Ψγ gene coding sequences. Using both RNAseq data and a RACE-PCR assay designed for specific amplification and validation of a spliced CH1 Ψγ sequence, we were able to characterize a number of such functional transcripts, which deserve consideration as γ5 mature transcripts. Their occurence was considerably less than that of classical mature γ IgH chain transcripts, with a ratio compared to γ1 in the range of a 10^4 to 10^5 lower abundance. The repertoire of these γ5 transcripts was, however, diversified and the VH usage did not significantly differ from that of other VDJ-Cγ transcripts. The level of SHM in γ5 transcripts was slightly lower than in classical VDJ-Cγ transcripts. Compared to the low number of functional VDJ-γ5 transcripts, unspliced or non-functional transcripts were also identified with a relative higher abundance, and indicated that the Ψγ/γ5 gene also undergoes a significant level of germline transcription.

While the abundance of Ψγ/γ5 transcripts relative to γ1 is low, it should be noticed that classical IgG are major products of B-lineage cells which are abundantly distributed in human body fluids and usually reach a concentration of about 10g/liter in plasma. A fifth IgG class expected to be one hundred thousand fold lower than IgG1would still be present at levels around 100µg/liter and might be detectable as a rare component of some body fluids by MS.

Notwithstanding the difficulty of using MS in order to distinguish a rare antibody within an abundant mixture of proteins notably made up of classical and plentiful IgG, we looked for the signature of Ψγ/γ5 tryptic peptides in various types of human samples and data sets. Several of these peptides (more specially those included in the Fc portion of the potential Ψγ/γ5 gene product) were found, notably in human urine samples. Exploration of public databases of MS experiments also identified the occurrence of Ψγ/γ5-encoded peptides in multiple samples.

Because adequate rearrangements were found together with functional transcripts and identification of the corresponding peptides, we thus think that the Ψγ gene can no longer be considered as a pseudo gene but rather as the Cγ5 gene.

To further check functionality of the corresponding IgG5, adapted versions of the therapeutic human IgG1 antibodies rituximab and trastuzumab were constructed.

γ5-rituximab (and γ5-trastuzumab) revealed both conventional IgG functions and unique properties. While assembled Ig were produced by stably transfected plasma cells, transient transfection of CHO and HEK293 cells in conditions providing high-rate production of IgG1 or IgG4 only yielded minute amounts of intact IgG5, together with abundant Fab and Fc fragments. The full-length IgG5 and the Fab fragments maintained antigen-recognition and expression of the idiotype, some level of complement activation and ADCC but to a much lower extent than with IgG1, was also detected. The most surprising property of IgG5 was in fact its unique propensity for auto-cleavage in conditions where IgG1 is perfectly stable such as cell culture supernatants. Even after addition of protease inhibitors, IgG5 indeed rapidly split into Fc and Fab fragments; the latter retained antigen-binding but were of small molecular weight.

Since human body fluids are rich in various proteases, it is thus clear that the *in vivo* half-life of IgG5 in its complete form is likely to be very short, making it most likely a precursor of low-size Ag-binding "mini-bodies". The small size of such expected proteins also likely explains why we more easily detected IgG5 tryptic fragments in the analysis of human urine proteins than from purified plasma Ig. In urine, proteins with a massless than 50KDa can pass freely, in contrast to full-size Ig and fragmented Ig are thus likely to be enriched. By contrast, in serum, micrograms of IgG5 or IgG5 fragments are most likely to be masked by the abundance of large serum proteins and notably of classical IgG representing grams per liter.

Interestingly from a functional point of view, monovalent Fab fragments devoid of effector function might bind antigen and contribute to tolerance. Fc fragments released by IgG proteolysis are indeed known for anti-inflammatory properties[15,18]. Partial nicking

of the IgG hinge by microbial proteases or tumor associated metalloproteases has also been demonstrated to yield monovalent antibodies with suppressive properties[19]. From an evolutionary point of view, it is noteworthy that the Ψγ/γ5 gene is present with a functional coding sequence not only in humans but also in other old-world primates. The available gorilla and chimpanzee sequences notably show perfect preservation of the unique hinge region sequence which joins the Fab and Fc fragments and which appears to promote spontaneous cleavage in humans.

Altogether, IgG5 appears as the tenth subclass of human antibodies. It is clearly the least abundant IgG subclass, but the expected amount indicated by our experiments is close to or lower than the observed concentrations of IgE and IgD (around 100 and 40µg/L, respectively), other rare human Igs but with dramatically important functional effects in both protection and allergy. A next step will be to design specific assays capable of differentiating IgG5 production from that of the other antibodies and determine how IgG5 antibodies are regulated in health and disease. Such explorations will help understand the status of this long hidden member of the human "antibody planetary system". They will help ascertain if IgG5 is either an emerging protein encoded by a nascent gene, or already a protein with a significant specific immune function (likely immunosuppression), or is simply a remnant or phantom, encoded through random gene duplication with progressive accumulation of mutations before finally being inactivated over time.

References

1. Hayashida, H. *et al.* Concerted evolution of the mouse immunoglobulin gamma chain genes. *EMBO J* **3,** 2047–2053 (1984).
2. Hirano, M., Das, S., Guo, P. & Cooper, M. D. The evolution of adaptive immunity in vertebrates. *Adv. Immunol.* **109,** 125–157 (2011).
3. Garred, P., Olsen, J., Mollnes, T. E., Bilde, T. & Glahn, B. E. Biocompatibility of urinary catheters. Effect on complement activation. *Br J Urol* **63,** 367–371 (1989).
4. Schur, P. H. IgG subclasses. A historical perspective. *Monogr Allergy* **23,** 1–11 (1988).
5. Vidarsson, G., Dekkers, G. & Rispens, T. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol* **5,** 520 (2014).
6. Michaelsen, T. E., Frangione, B. & Franklin, E. C. Primary structure of the 'hinge' region of human IgG3. Probable quadruplication of a 15-amino acid residue basic unit. *J. Biol. Chem.* **252,** 883–889 (1977).
7. Ellison, J. & Hood, L. Linkage and sequence homology of two human immunoglobulin gamma heavy chain constant region genes. *Proc. Natl. Acad. Sci. U.S.A.* **79,** 1984–1988 (1982).
8. Flanagan, J. G. & Rabbitts, T. H. Arrangement of human immunoglobulin heavy chain constant region genes implies evolutionary duplication of a segment containing gamma, epsilon and alpha genes. *Nature* **300,** 709–713 (1982).
9. Pinaud, E. *et al.* The IgH locus 3' regulatory region: pulling the strings from behind. *Adv. Immunol.* **110,** 27–70 (2011).

10. Takahashi, N. *et al.* Structure of human immunoglobulin gamma genes: implications for evolution of a gene family. *Cell* **29,** 671–679 (1982).
11. Bensmana, M., Huck, S., Lefranc, G. & Lefranc, M. P. The human immunoglobulin pseudo-gamma IGHGP gene shows no major structural defect. *Nucleic Acids Res* **16,** 3108 (1988).
12. Boyer, F. *et al.* CSReport: A New Computational Tool Designed for Automatic Analysis of Class Switch Recombination Junctions Sequenced by High-Throughput Sequencing. *J. Immunol.* **198,** 4148–4155 (2017).
13. Edelman, G. M. *et al.* The covalent structure of an entire gammaG immunoglobulin molecule. *Proc. Natl. Acad. Sci. U.S.A.* **63,** 78–85 (1969).
14. Bruhns, P. & Jönsson, F. Mouse and human FcR effector functions. *Immunol. Rev.* **268,** 25–51 (2015).
15. Vafa, O. *et al.* An engineered Fc variant of an IgG eliminates all immune effector functions via structural perturbations. *Methods* **65,** 114–126 (2014).
16. Rouaud, P. *et al.* Elucidation of the enigmatic IgD class-switch recombination via germline deletion of the IgH 3' regulatory region. *J. Exp. Med.* **211,** 975–985 (2014).
17. Cogné, M. *et al.* Locus Suicide Recombination actively occurs on the functionally rearranged IgH allele in B-cells from inflamed human lymphoid tissues. *bioRxiv* 430215 (2018). doi:10.1101/430215
18. Blumberg, R. S., Lillicrap, D. & IgG Fc Immune Tolerance Group. Tolerogenic properties of the Fc portion of IgG and its relevance to the treatment and management of hemophilia. *Blood* **131,** 2205–2214 (2018).

19. Brezski, R. J. *et al.* Tumor-associated and microbial proteases compromise host IgG effector functions by a single cleavage proximal to the hinge. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 17864–17869 (2009).

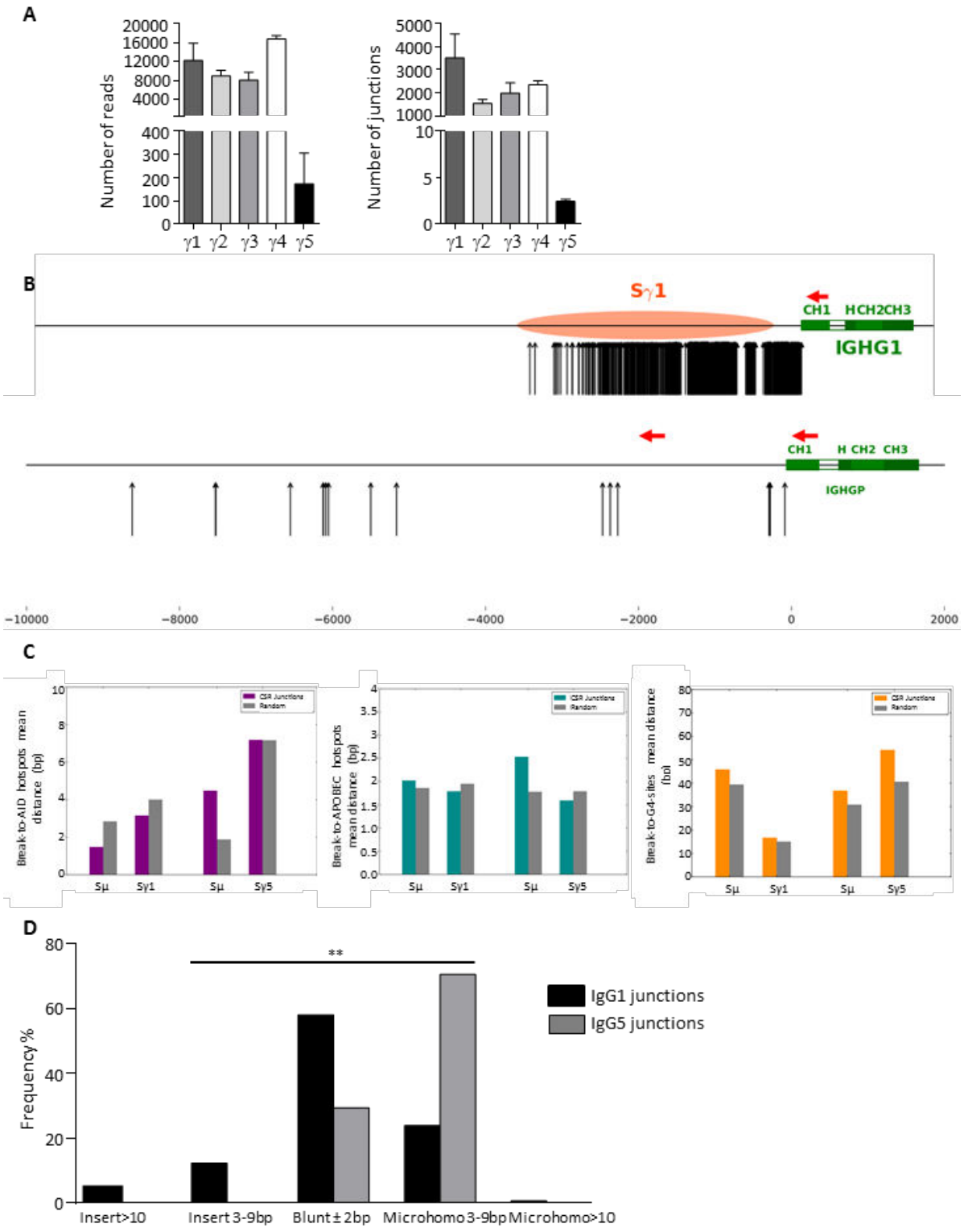

Figure 1: Number, position of breaks and structure of human repaired Sμ-ψγ junctions.

(A) CSR reads (left) and junctions (right) of Sγ regions and 5' flanking regions of ψγ were studied in human tonsil B cells from different individuals. Data represent mean numbers (reads or junctions) ± SEM. (B) Maps reporting breaks observed in sequenced Sμ-Sγ1 compared to Sμ-ψγ junctions. (C) Mean distance to AID target sites (left), APOBEC3 (middle) and G-quadruplex (right) sites. (D) Structure of repaired junctions with regards to mean number of inserted nucleotides and length of microhomology (microhomo) between broken ends for Sμ-Sγ1 and Sμ-ψγ5.

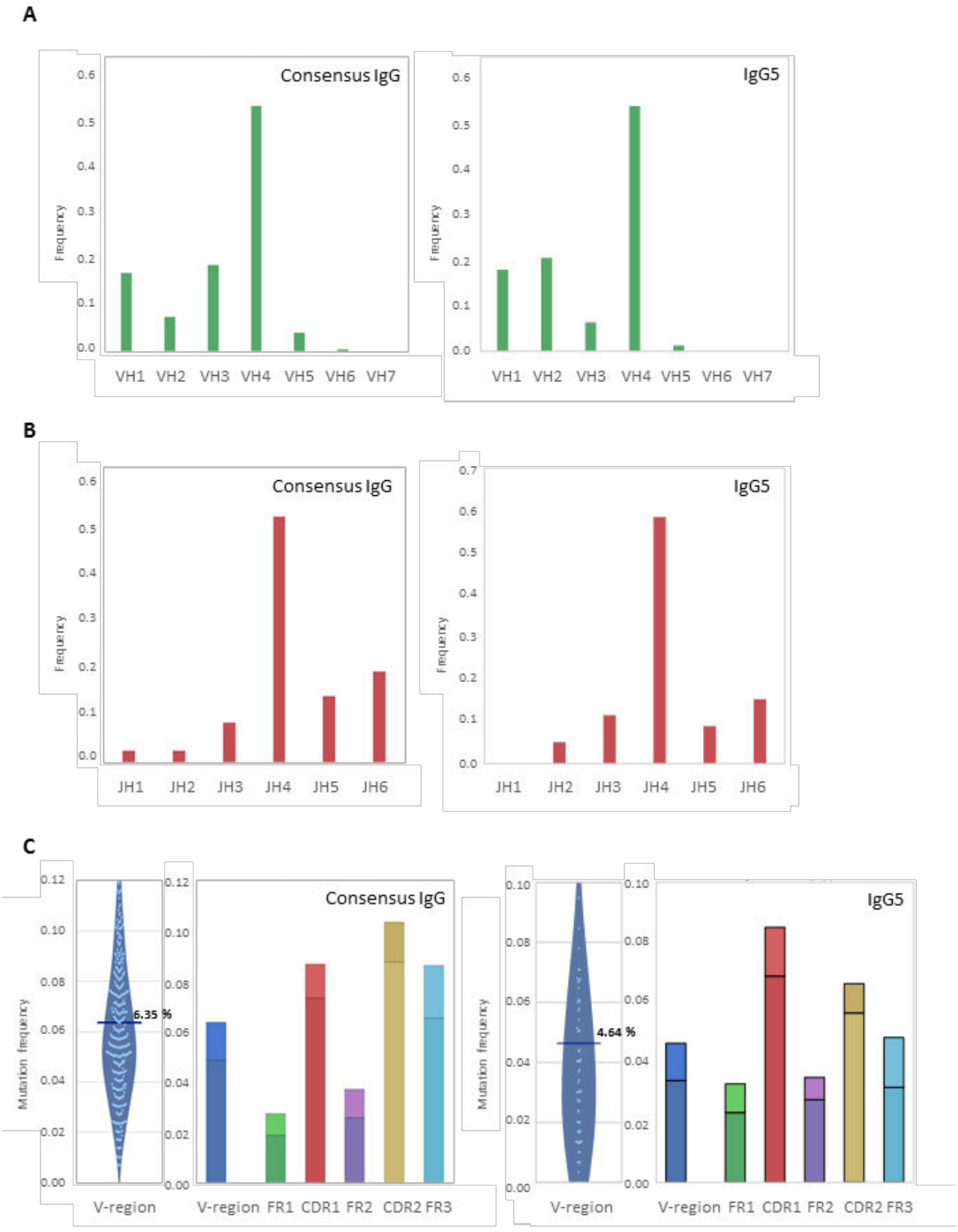

Figure 2: IgG5 VDJ repertoire and mutations.

VDJ repertoire was studied in tonsil B cells after NGS sequencing of RACE products generated using either Cγ consensus (left) or Cγ5 (right) specific reverse primers. (A) Usage of VH and JH gene segments is shown. (B) Mutation frequencies in the full V region and in the various subdomains (FR1, CDR1, FR2, CDR2 and FR3) are shown.

A

```
CH1
IgG1 RSTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV
IgG2 A----------C-R---ES--------------------------------------------------T--NF-----T---D----------T-
IgG3 A----------C-R-------------------------------------------------------T----------R-
IgG4 A----------C-R---ES--------------------------------------------------K--T---D----------R-
IgG5 A----------V---R-V-E---------------------RS-------------------------T---D----------T-

HINGE
IgG1 EPKSC..DKTHTCPPCP
IgG2 -R-C-VE.....-----
IgG3 -L-TPLG-T-----R--EPKSCDTPPPCPRCPEPKSCDTPPPCPRCPEPKSCDTPPPCPRCP
IgG4 -S-YGPP.....--S--
IgG5 ---TPCC-T-------A

                FcGRI binding      FcRn binding   FcRn binding   N-glycosylation      FcRn binding
CH2
IgG1 APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
IgG2 --PVA.----------Q-------M-------------F---F--------V----------G----------T-
IgG3 ----------------Q-K-----------------------F----------------------------------T-
IgG4 ---F----------Q-----Q---------------------F--------------------------G--SS---------
IgG5 TT-P-----------------------------------------------W--------H-------V-N-----R---------G----------T-

                                                N-glycosylation                FcRn binding
CH3
IgG1 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK
IgG2 ----------E-M---------------------------------M-------------------------------------------------
IgG3 ----------E-M---------------------S-------N----M-----------------------I----------RF----------
IgG4 ----------QE-M---------------------------------------------R---------E----------L--
IgG5 ----------QK.M-----T----------T-----------------M---N-----------------G----------

M1M2
IgG1 ELQLEESCAEAQDGELDGLWTTITIFITLFLLSVCYSATITFFKVKWIFSSVVDLKQTIIPDYRNMIGQGA
IgG2 ----------------------------------------------------V-------R---
IgG3 ----------------------------------V-------------------------------
IgG4 ----------------------------------V-------------------V-------R---
IgG5 ----------------------------------V-------------------V----------
```

B

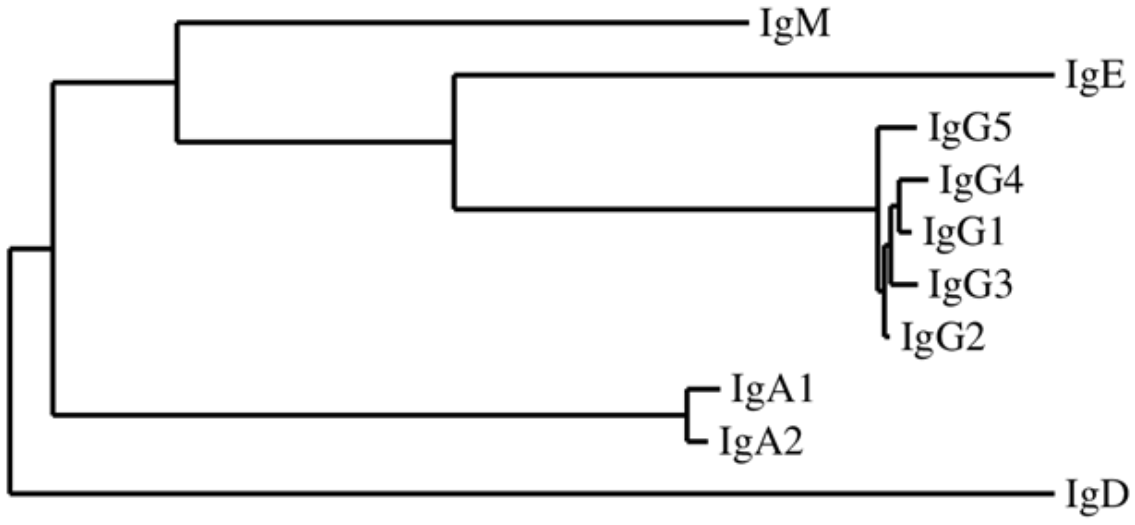

Figure 3: Ψγ coding sequence.

(A) Comparison of the ψγ/γ5 gene coding sequence to γ1-γ4 subclasses. N-glycosylation sites are boxed. Specific γ5 peptides identified by MS are underlined in grey. (B) Phylogenetic tree.

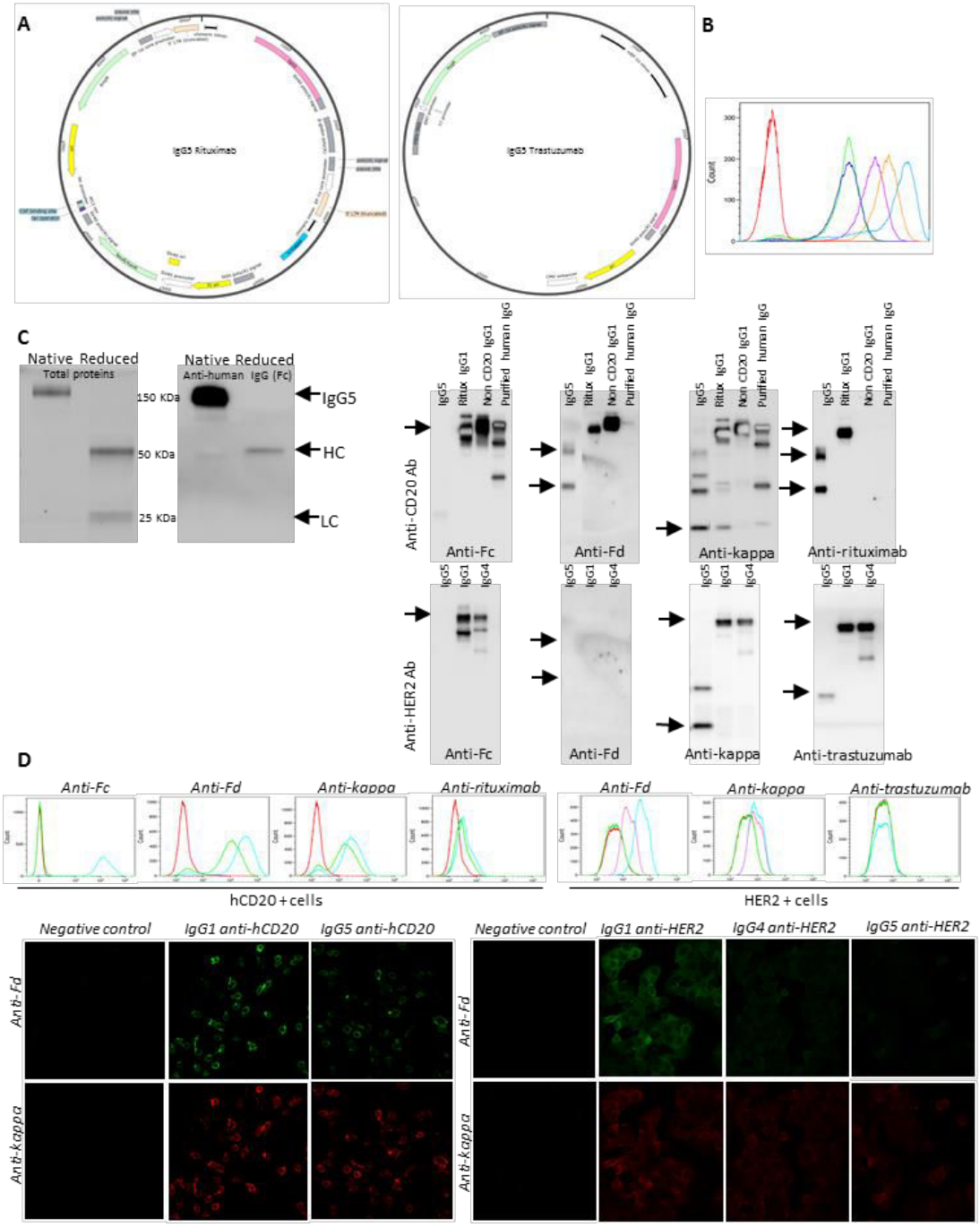

Figure 4: Production of recombinant specific IgG5 mAbs.

(A) Maps of vectors encoding IgG5 versions of rituximab (left) and trastuzumab (right). (B) Binding characteristics of anti-CD20 antibody measured by flow cytometry on hCD20-expressing variant (EL4-CD20). Red curves represent negative control, green curves represent 10µg/ml IgG5 anti-hCD20. Other curves curves represent IgG1 anti-hCD20 at 10µg/ml (light blue), 0.1 µg/ml (orange), 0.05µg/ml (purple) and 0.01µg/ml (dark blue). (C) Western blot analysis of recombinant IgG5 Abs produced either in stably transfected plasma cells (left) or in transiently transfected Expi-CHO cells (right). Fc, Fd, rituximab and kappa regions from IgG5, IgG1 and control antibodies were analyzed. (D) Analysis of specific binding to target cells (hCD20+ DHL4 cells, or HER2+ MCF7 cells) for spontaneously released Ig fragments upon *in vitro* IgG5 production. Red curves represent negative controls, light blue curves represent IgG1 anti-hCD20 (10µg/ml) or anti-HER2, purple curves represent IgG4 anti-HER2, green curves represent IgG5 anti-hCD20 (10µg/ml) or anti-HER2.

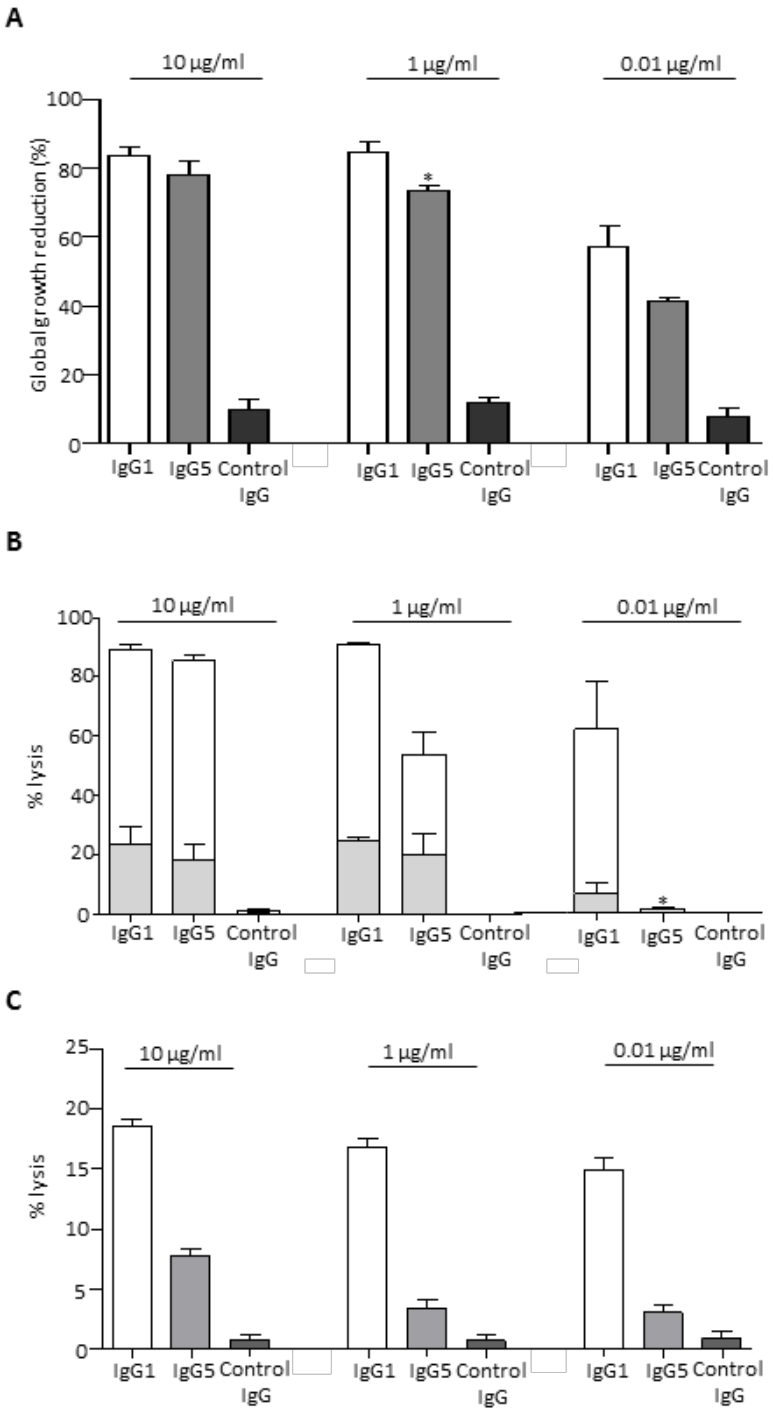

Figure 5: Functional evaluation of anti-CD20 IgG5 mAb.

Evaluations were carried out with IgG5 mAb produced *in vitro* in the standard Expi-CHO system, and thus included a low amount of full-size IgG5 together with a large amount of cleaved IgG5. (A) Direct effects of anti-CD20 mAb preparations (10 µg/mL, 1 µg/mL or 0.01 µg/mL) were measured by MTS assay on DHL4 cells. (B) Anti-CD20 mediated complement-dependent cytotoxicity evaluated against human DHL4 target cells incubated with various mAb concentrations (10 µg/mL, 1 µg/mL or 0.01 µg/mL) in the presence of either active (open bars) or heat-inactivated human serum (grey bars). Late apoptotic cells were evaluated by cell cytometry. (C) ADCC induced by anti-CD20 antibodies in DHL4 cells (10 µg/mL, 1 µg/mL or 0.01 µg/mL).