Gene conversion in human immunoglobulin γ locus shown by unusual location of IgG allotypes

Marie-Paule Lefranc, Ahmed-Noureddine Helal[+], Gerda de Lange[°], Hassan Chaabani[+], Erna van Loghem[°] and Gérard Lefranc

Laboratoire d'Immunogénétique, Université des Sciences et Techniques de Languedoc, F. 34060 Montpellier Cedex, France, [+]Faculté de Pharmacie et de Médecine Dentaire, 5019 Monastir, Tunisia and *[°]Central Laboratory of the Netherlands Red Cross Blood Transfusion Service, POB 9190, 1006 AD Amsterdam, The Netherlands*

Received 15 November 1985

The constant region of the γ1, γ2 and γ3 heavy chains of the human IgG1, IgG2 and IgG3 immunoglobulins carries antigenic determinants or G1m, G2m and G3m allotypes, which are genetic markers of these subclasses. The exceptional presence on γ1 and γ2 chains of Gm allotypes usually located on the CH3 domain of γ3 shows an unexpected clustering of base changes and subsequent identity of short DNA sequences in the CH3 exon of the non-allelic γ1, γ2 and γ3 genes. Such clusters of substitutions are not easily explained on the classical basis of point mutations. A gene conversion, which substituted a segment of the γ1 or γ2 gene with the homologous region of the non-allelic γ3 gene, is more likely. Other examples of possible conversion involving the γ genes are described. The conservation or the restoration of short sequences produced by the conversion events might be related to the biological properties of the constant region of the heavy chains.

Gm allotypic marker *Haplotype* *Gene conversion* *(Human)* *CH gene*

1. INTRODUCTION

In the human immunoglobulin system there are 5 classes or isotypes – IgM, IgD, IgG, IgE and IgA – defined by isotypic antigenic determinants and the physico-chemical and biological activities of their heavy chains (μ, δ, γ, ϵ and α, respectively). Subclasses are also recognised in IgG (IgG1, IgG2, IgG3 and IgG4) and IgA (IgA1 and IgA2) according to the properties of the constant region of the heavy chains (γ1, γ2, γ3 and γ4; α1 and α2, respectively) [1–7]. The CH genes coding for the constant region of the heavy chains are closely linked on chromosome 14 [8,9]. Their order – 5′-μ-δ-γ3-γ1-$\psi\epsilon$1-α1.....γ2-γ4-ϵ-α2-3′ – has been determined from the study of cosmid clones and from the patterns of deletions [10–15]. These class and subclass genes have been generated by duplications of a common ancestor. This was suggested by the homology in amino acid sequences of their products [1–7,16].

The constant region of the γ1, γ2, γ3, α2 and ϵ chains carry antigenic determinants, referred to as G1m, G2m, G3m, A2m and Em allotypes, which are specific of each of the subclasses and represent genetic markers for the respective genes. Of all the immunoglobulin genes, the most polymorphic are the γ3 and then the γ1 genes [17–26]. The Gm, A2m and Em markers, inherited in fixed combinations or haplotypes due to the close linkage between the γ, α and ϵ genes (fig.1), are useful tools for the characterization of populations, analyses of disease associations, forensic medicine and for the study of heavy chain gene evolution. Inheritance of unexpected and unusual sets of allotypes are particularly interesting since they reveal genetic events (point mutations, duplications, deletions, exchanges of CH exons, hybrid

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

genes...) which have occurred at the level of the coding sequences [27–31]. Here we postulate that gene conversion is a major genetic event in the CH multigene family.

2. MATERIALS AND METHODS

Serum samples from Lebanese and Tunisian people were collected and tested for the Gm and Am allotypes by the classical haemagglutination-inhibition method. The specific antisera and the relevant red cell coatings have been described [23,24]. The exceptional Gm phenotypes served as starting point for family investigations to determine unambiguously the unusual haplotypes.

3. RESULTS AND DISCUSSION

3.1. *Occurrence of IgG3 allotypes in IgG1 and IgG2 subclasses*

3.1.1. Converted $\gamma 1$ and $\gamma 2$ genes with $\gamma 3$ sequences

A few years ago, we demonstrated the occurrence of some human Gm allotypes in more than one subclass [24,32]: the Gm (g5 and v) antigenic determinants, usually located on the CH3 domain of $\gamma 3$ chains [33–35] (fig.1), were found on $\gamma 1$ chains encoded by an unusual $G1m^{a,g5,v,z}$ $\gamma 1$ gene. The presence of such an uncommon and stable $\gamma 1$ allele was demonstrated by its regular transmission through 3 generations of a Tunisian family; one child was even found to be homozygous.

Recently, we found again the Gm(g5 and v) antigenic determinants at an unusual location but, in this case, they were located on $\gamma 2$ molecules (unpublished).

The unexpected presence of Gm(g5 and v) on these $\gamma 1$ and $\gamma 2$ chains revealed a surprising identity of short DNA sequences, encoding Gm(g5 and v), in the CH3 exon of the non-allelic $\gamma 3$, $\gamma 1$ and $\gamma 2$ genes. Such remarkable clustering of base changes within the $\gamma 1$ and $\gamma 2$ genes appear to be much more frequent than the normal mutation rate. Moreover, they are not easily explained on the classical basis of random mutation unless it is postulated that the mutations were repeated at the same position(s) on 2 different genes. The hypothesis of an unequal intragenic crossover between the homologous $\gamma 1$ and $\gamma 3$ genes (1st example) (fig.2A) or between the $\gamma 2$ and $\gamma 3$ genes (2nd example) may be considered. However, once again, the abnormally increased frequency of such recombination events seems to be unlikely. A better explanation is that, by a gene conversion event, a small region of the CH3 sequence of the $\gamma 1$ or $\gamma 2$ gene was substituted by the homologous region of the non-allelic $\gamma 3$ gene coding for the Gm(g5 and v) allotypes (fig.2B).

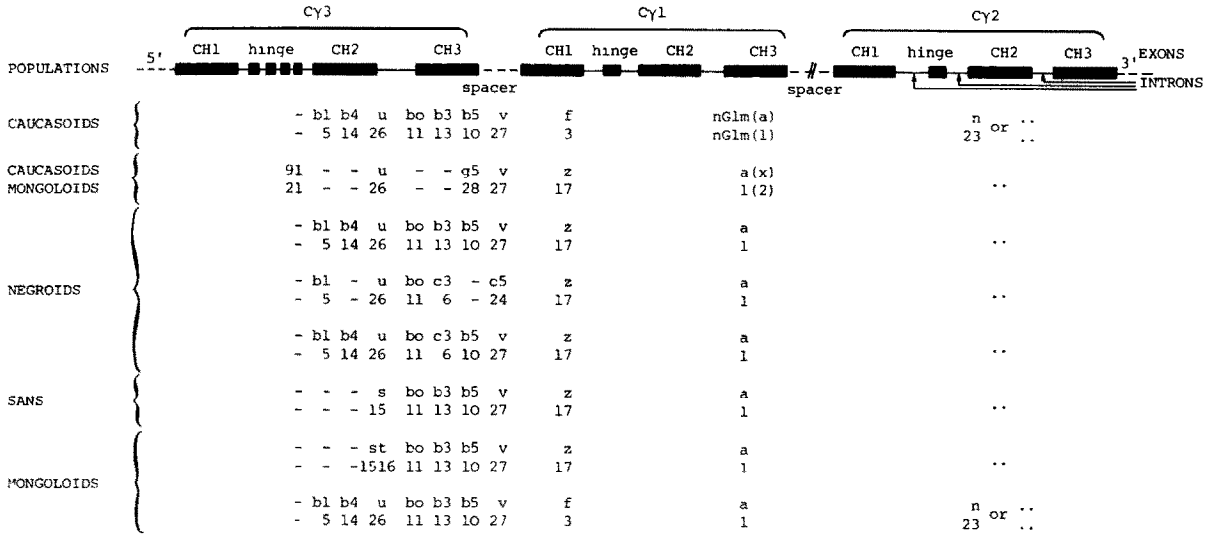

Fig.1. Prevalent haplotypes of the human $\gamma 3$, $\gamma 1$ and $\gamma 2$ heavy chain allotypes in several populations.

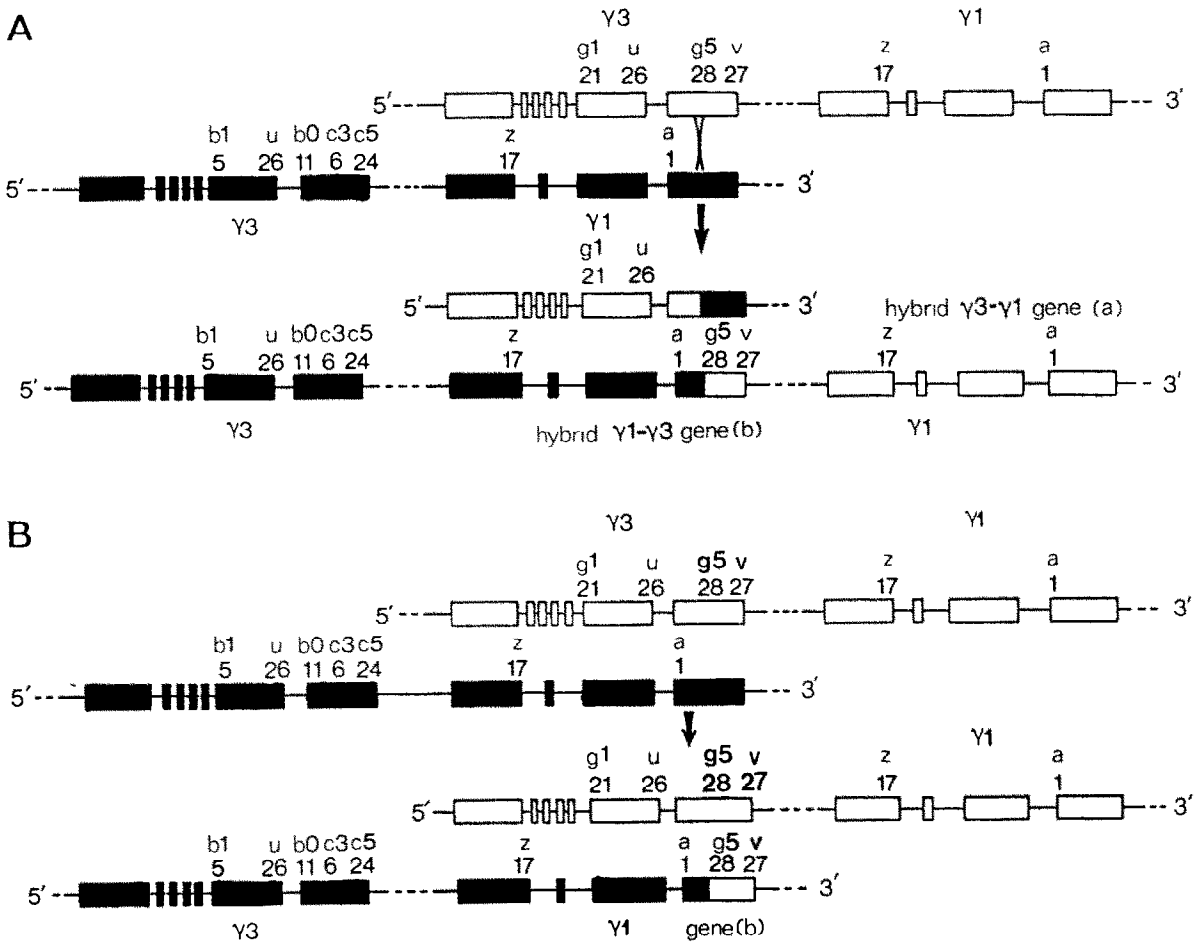

Fig.2. Diagrams of unequal intragenic crossover and of conversion between $\gamma 1$ and $\gamma 3$ genes. (A) Unequal intragenic crossing-over within the CH3 exon. (a) Chromosome carrying a hybrid $\gamma 3$-$\gamma 1$ gene but no normal $\gamma 3$ and $\gamma 1$ genes. (b) Chromosome carrying a hybrid $\gamma 1$-$\gamma 3$ gene and normal $\gamma 3$ and $\gamma 1$ genes. (B) Conversion of a part of the CH3 sequence of the $Gm^{z,a}\gamma 1$ gene by the homologous region of the $Gm^{g1,g5,u,v}\gamma 3$ non-allelic gene.

In the first example (converted $\gamma 1$ gene), the maximum extent of the coding DNA stretch involved in the conversion of the $\gamma 1$ gene may be predicted since the unusual or converted $\gamma 1$ chains still possess the G1m(a) allotype. This marker is related to the presence on usual $\gamma 1$ chains of Asp and Leu residues at positions 356 and 358, respectively [16,36]. Consequently, the converted CH3 sequence in the $\gamma 1$ acceptor gene is downstream of codon 358.

In the second example (converted $\gamma 2$ gene), it is not possible to delimit the 5′-end of the converted region since both the $\gamma 2$ and $\gamma 3$ genes carry the same nG1m(a) isoallotype related to the presence of Glu and Met residues at positions 356 and 358, respectively [16,37,38]. The length of the converted CH3 sequence might be the same as in the first example; alternatively, it may be larger and include the entire CH3 exon.

3.2. *Absence of some IgG3 allotypes*

3.2.1. Converted $\gamma 3$ genes with non-polymorphic $\gamma 1$, $\gamma 2$ or $\gamma 4$ sequences

In the cases described above a short DNA region has been transferred in the direction 5′–3′, from the $\gamma 3$ donor gene to the $\gamma 1$ and $\gamma 2$ acceptor genes to its right (fig.1).

We must consider also the conversion of

polymorphic $\gamma 3$ genes by non-polymorphic $\gamma 1$, $\gamma 2$ and $\gamma 4$ sequences located at their 3′-terminal side. We have described previously, in 3 unrelated Tunisian families [24], unusual $\gamma 3$ chains, phenotypically G3m(g1 and u), lacking the G3m(g5 and v) allotypes usually present on the CH3 domain of the G3m(g1,g5,u,v)$\gamma 3$ chains (fig.1). Several other examples of such $\gamma 3$ chains have been reported [34]. The absence of the G3m(g5 and v) allotypes on these $\gamma 3$ chains might be due to the conversion of the $\gamma 3$ CH3 coding region by the homologous sequences of either $\gamma 1$, $\gamma 2$ or $\gamma 4$ subclass gene which lack these allotypes.

It is particularly intriguing that they are the same Gm(g5 and v) antigenic determinants which were unexpectedly found on the aforementioned converted $\gamma 1$ and $\gamma 2$ chains. It seems likely that specific sequences favour the conversion within the CH3 coding (and flanking non-coding) sequences of the $G3m^{g1,g5,u,v}\gamma 3$ allele and that gene conversion occurs in both 5′-3′ and 3′-5′ directions, the CH3 exon of this $\gamma 3$ gene being either the donor or the acceptor of the sequence.

3.3. *Inter-allelic conversions*

3.3.1. Conversion between $\gamma 3$ allelic genes

Gene conversion between $\gamma 3$ allelic genes may be responsible for situations already reported. An uncommon $Gm^{b1,b4,g5,u,v}\gamma 3$ gene has been found in a Tunisian family and transmitted through 3 generations [24]. This unusual gene may be explained by a gene conversion between $G3m^{b0,b1,b3,b4,b5,u,v}$ and $G3m^{g1,g5,u,v}$, the most common $\gamma 3$ alleles among the Caucasoids, the CH3 exon of the first allele being converted by the homologous sequences of the second allele (fig.3A). The converted sequence, encoding Gm(g5 and v), might be, again, the same as that described earlier.

3.3.2. Conversion between $\gamma 1$ allelic genes

An example of possible gene conversion between $\gamma 1$ allelic genes is represented by an unusual and

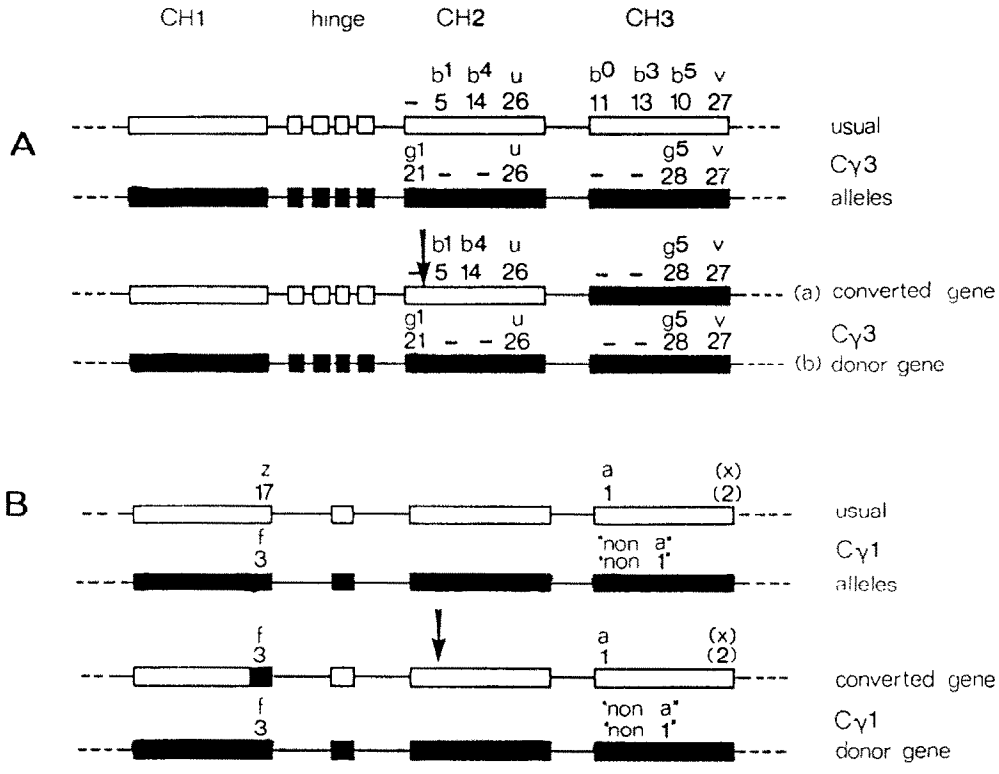

Fig.3. Postulated conversion of allelic γ genes. (A) Conversion of the CH3 exon of the usual $\gamma 3$ $G3m^{b0,b1,b3,b4,b5,u,v}$ allele. (B) Conversion of the 3′-side of the CH1 exon of the usual $\gamma 1$ $G1m^{z,a}$ allele.

unexpected $\gamma 1$ gene inherited by only 1 of the 10 children in a Lebanese family [39] but absent from the genotypes of the parents. The phenotype of the child could not be explained by various combinations of maternal and paternal haplotypes. To explain the unexpected presence of the Gm(f) antigenic determinant we must postulate either a recombination or a gene conversion during the maternal or paternal meiosis. A single equal or unequal crossover, the most likely of all the recombination events between the $\gamma 3$ and $\gamma 1$ genes, must be excluded, at least during the maternal meiosis, for the following reasons: the $G1m^{f}\gamma 1$ allele and the $A2m^{1}\alpha 2$ gene, both these located at the right of the $G3m^{b0,b1,b3,b4,b5,u,v}\gamma 3$ gene [10], should have been transmitted together and the child should be homozygous $A2m^{1}/A2m^{1}$ (fig.4A and D), whereas she is heterozygous $A2m^{1}/A2m^{2}$. Due to the fact

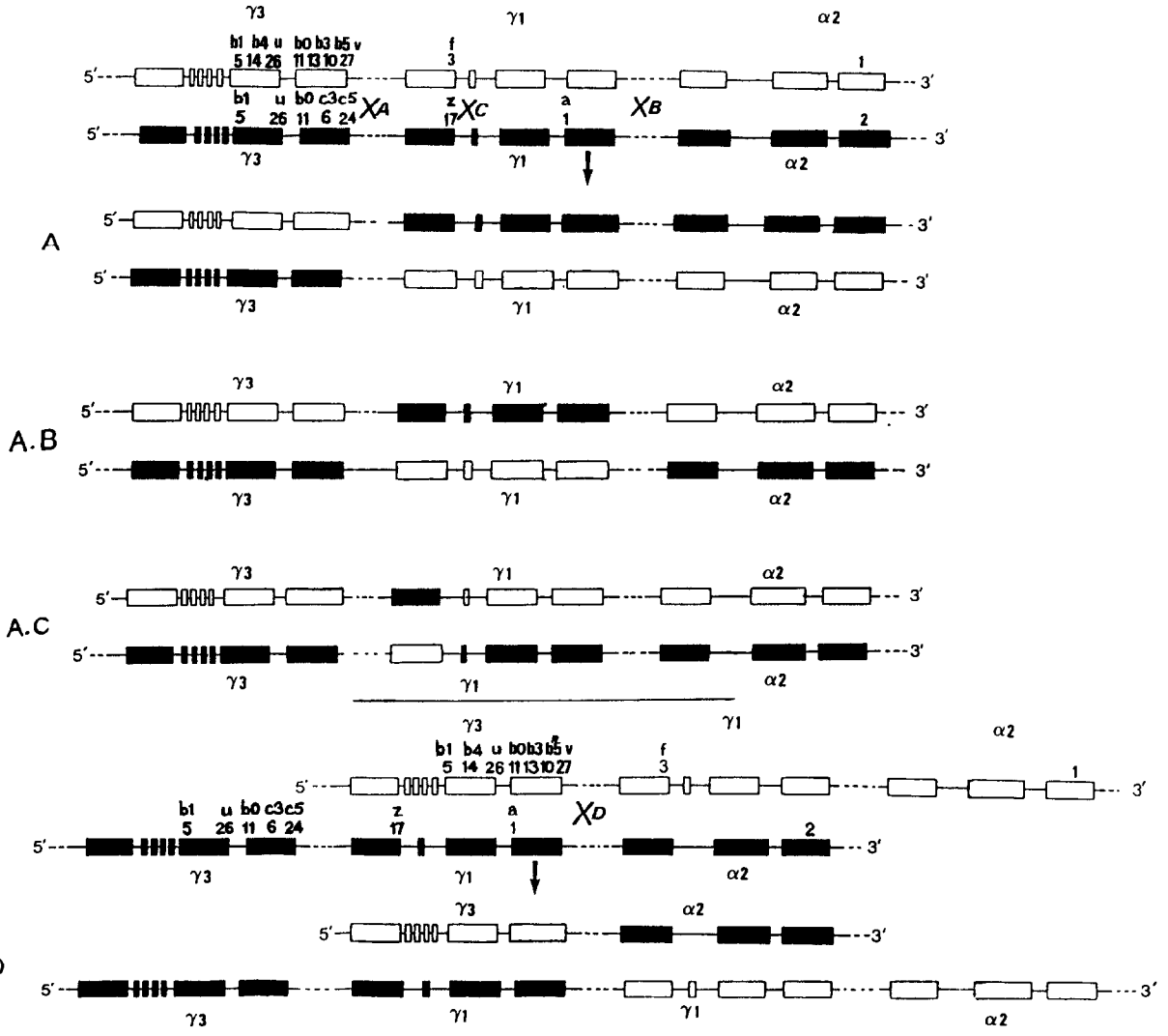

Fig.4. Diagrams of postulated single and double-, equal and unequal-, inter- and intragenic crossovers during the maternal meiosis unable to explain the unexpected presence of Gm(f). The first (A) and the last (D) recombinant haplotypes were not observed, the others are unlikely. (A) Single equal intergenic crossover between the $\gamma 3$ and $\gamma 1$ genes. (A.B) Double equal intergenic crossover between, on the one hand, the $\gamma 3$ and $\gamma 1$ genes and, on the other, the $\gamma 1$ and $\alpha 2$ genes. (A.C) Double equal crossover, the first one between the $\gamma 3$ and $\gamma 1$ genes and the second one within the $\gamma 1$ gene. (D) Single unequal intergenic crossover.

that the $\alpha 2$ alleles have not been affected by the postulated recombination, an unlikely double intergenic crossover is required (fig.4A,B). Furthermore, if the unexpected Gm(f) antigenic determinant is supposed to be located on $\gamma 1$ chains phenotypically G1m(a,f), a double equal intragenic crossover must even be postulated which is far more unlikely (fig.4A,C). In contrast, the presence of Gm(f) on G1m(a and f)$\gamma 1$ chains may be readily explained by a conversion of the 3'-end of the CH1 exon of the $\gamma 1$ $G1m^{z,a}$ allele by the sequence encoding G1m(f) from the $G1m^{f}\gamma 1$ allele during the maternal meiosis (fig.3B).

4. CONCLUSION

In the examples described here and in the case of the Gm(c3) epitope found on $\gamma 2$ chains instead of $\gamma 3$ [32], the converted amino acid sequences have been mainly recognised due to the numerous and mutually exclusive allotypes of the polymorphic $\gamma 3$ chains. Such uncommon chains are more often unmasked in consanguineous populations because of the increased frequency of individuals homozygous for the corresponding unusual alleles. Other gene-conversion events within the human immunoglobulin CH genes may occur, as demonstrated between the $\alpha 1$ and $\alpha 2$ genes [40] but would be masked at the heterozygous state or because of absence either of antigenic determinants or of antibodies able to detect them. Monoclonal antibodies should allow us to find more converted chains, by detecting Gm allotypes unrecognized to date by the conventional anti-allotypes antibodies [41].

Two major points arise from these postulated gene conversions within the human immunoglobulin CH genes. Firstly, to account for the frequency of clusters of substitutions much higher than the normal mutation rate and for the polarity effect favouring the conversion of one short sequence more than an adjacent one, the gene conversion model must assume that the DNA cutting and joining processes occur at specific sites. Some of the examples described here, involving the same DNA region of a $\gamma 3$ gene, tend to confirm the existence of such hot-spots. Some nucleotide sequences, which seem to be present in regions more susceptible to breakage and which can form Z-DNA have been proposed as sites of gene conversion and recombination events (for review see [42]).

Secondly, the gene conversion explains the maintenance of homology between non-allelic genes of highly polymorphic [43–47] and of less polymorphic [40,48–50] multigene families. The converted sequences in the IgCH region genes could be related to the structural and biological properties of the constant region of the heavy chains and should be, therefore, located at positions strongly selected for survival during the evolution. For example, these sequences might be those involved in the binding to the poly-Ig receptor, to the various Fc receptors and to the receptors for transepithelial transport of IgG across the placenta as well as in the binding to the first component of Complement. They might be also involved in the allotypic recognition signals on the CH region of the B cell-surface immunoglobulins for the T and B lymphocyte cooperation necessary to the B cell antigen-driven differentiation.

REFERENCES

[1] Cohen, S. and Porter, R.R. (1964) Adv. Immunol. 4, 287–349.
[2] Feinstein, D. and Franklin, E.C. (1966) Nature 212, 1496–1498.
[3] Grey, H.M. and Kunkel, H.G. (1967) Biochemistry 6, 2325–2334.
[4] Frangione, B., Milstein, C. and Pink, J.R.L. (1969) Nature 221, 145–149.
[5] Van Loghem, E. (1971) Ann. NY Acad. Sci. 190, 136–149.
[6] Natvig, J.B. and Kunkel, H.G. (1973) Adv. Immunol. 16, 1–59.
[7] Wang, A.C. and Fudenberg, H.H. (1974) J. Immunogenet. 1, 3–31.
[8] Croce, C.M., Shander, M., Martinis, J., Cicurel, L., D'Amona, G.G., Dolby, T.W. and Koprowski, P. (1979) Proc. Natl. Acad. Sci. USA 76, 3416–3419.
[9] Hobart, M.J., Rabbitts, T.H., Goodfellow, P.N., Solomon, E., Chambers, S., Spurr, N. and Povey, S. (1981) Ann. Hum. Genet. 45, 331–335.
[10] Flanagan, J.G. and Rabbitts, T.H. (1982) Nature 300, 709–713.
[11] Lefranc, M.P., Lefranc, G. and Rabbitts, T.H. (1982) Nature 300, 760–762.
[12] Lefranc, G., Chaabani, H., Van Loghem, E., Lefranc, M.P., De Lange, G. and Helal, A.N. (1983) Eur. J. Immunol. 13, 240–244.

[13] Lefranc, M.P., Lefranc, G., De Lange, G., Out, T.A., Van den Broek, P.J., Van Nieuwkoop, J., Radl, J., Helal, A.N., Chaabani, H., Van Loghem, E. and Rabbitts, T.H. (1983) Mol. Biol. Med. 1, 207–217.
[14] Bech-Hansen, N.T., Linsley, P.S. and Cox, D.W. (1983) Proc. Natl. Acad. Sci. USA 80, 6952–6956.
[15] Migone, N., Oliviero, S., De Lange, G., Delacroix, D.L., Boschis, D., Altruda, F., Silengo, L., Demarchi, M. and Carbonara, A.O. (1984) Proc. Natl. Acad. Sci. USA 81, 5811–5815.
[16] Kabat, E.A., Wu, T.T., Bilofsky, H., Reid-Miller, M. and Perry, H. (1983) Sequences of Proteins of Immunological Interest (Edited by US Department of Health and Human Services) p.275, Public Health Service NIH.
[17] Grubb, R. and Laurell, A.B. (1956) Acta Path. Microbiol. Scand. 39, 390–398.
[18] Steinberg, A.G. (1969) Ann. Rev. Genet. 3, 25–52.
[19] Van Loghem, E., Natvig, J.B. and Matsumoto, H. (1970) Ann. Hum. Genet. 33, 351–359.
[20] Steinberg, A.G. (1973) Isr. J. Med. Sci. 9, 1249–1256.
[21] Schanfield, M.S. and Fudenberg, H.H. (1975) in: Biosocial Inter-relations in Population Adaption (Mouton, ed.) p.105, Amsterdam.
[22] WHO (1976) WHO Meeting on Human Immunoglobulin Allotypic Markers, J. Immunogenet. 3, 357–362.
[23] Lefranc, G., Rivat, L., Serre, J.L., Lalouel, J.M., Pison, G., Loiselet, J., Ropartz, C., De Lange, G. and Van Loghem, E. (1978) Hum. Genet. 41, 197–209.
[24] Lefranc, G., De Lange, G., Rivat, L., Langaney, A., Lefranc, M.P., Ellouze, F., Sfar, G., Sfar, M. and Van Loghem, E. (1979) Hum. Genet. 50, 199–211.
[25] Van Loghem, E. and Biewenga, J. (1983) Mol. Immunol. 20, 1001–1007.
[26] Van Loghem, E., Aalberse, R.C. and Matsumoto, H. (1984) Vox Sang. 46, 195–206.
[27] Natvig, J.G. and Kunkel, H.G. (1969) 17th Coll. of Protides of the Biological Fluids, pp.141–145, Pergamon, New York.
[28] Lefranc, G., Rivat, L., Rivat, C., Loiselet, J. and Ropartz, C. (1976) Am. J. Hum. Genet. 28, 51–61.
[20] Lefranc, G., Dumitresco, S.M., Salier, J.P., Rivat, L., De Lange, G., Van Loghem, E. and Loiselet, J. (1979) J. Immunogenet. 6, 215–221.
[30] Van Loghem, E., Sukernik, R.A., Osipova, L.P., Zegers, B.J.M., Matsumoto, H., De Lange, G. and Lefranc, G. (1980) J. Immunogenet. 7, 285–299.
[31] Lefranc, G., Lefranc, M.P., Helal, A.N., Boukef, K., Chaabani, H., Sfar Gandoura, M. and Van Loghem, E. (1982) J. Immunogenet. 9, 1–9.
[32] Van Loghem, E., De Lange, G., Van Leeuwen, A.M., Van Eede, P.H., Nijenhis, L.E., Lefranc, M.P. and Lefranc, G. (1982) Vox Sang. 43, 301–309.
[33] Schanfield, M.S. and Fudenberg, H.H. (1974) Vox Sang. 26, 133–140.
[34] Van Loghem, E., Blanc, M. and De Lange, G. (1977) J. Immunogenet. 4, 371–383.
[35] Rivat, L., Rivat, C., Cook, C.E. and Steinberg, A.G. (1978) Ann. Immunol. Inst. Pasteur 129C, 33–45.
[36] Edelman, G.M., Cunningham, B.A., Gall, W.E., Gottlieb, P.D., Rutishauser, U. and Waxdal, M.J. (1969) Proc. Natl. Acad. Sci. USA 63, 78–85.
[37] Ellison, J. and Hood, L. (1982) Proc. Natl. Acad. Sci. USA 79, 1984–1988.
[38] Pink, J.R.L., Buttery, S.H., De Vries, G.M. and Milstein, C. (1970) Biochem. J. 117, 33–47.
[39] Lefranc, G., Rivat, L., Salier, J.P., Van Loghem, E., Aydenian, H., Zalzal, P., Chakhachiro, L., Loiselet, J. and Ropartz, C. (1977) Am. J. Hum. Genet. 29, 523–536.
[40] Flanagan, J.G., Lefranc, M.P. and Rabbitts, T.H. (1984) Cell 36, 681–688.
[41] Zelaschi, D., Newby, C., Parsons, M., Van West, B., Cavalli-Sforza, L.L., Herzenberg, L.A. and Herzenberg, L.A. (1983) Proc. Natl. Acad. Sci. USA 80, 3762–3766.
[42] Rogers, J. (1983) Nature 305, 101–102.
[43] Weiss, E., Golden, L., Zakut, R., Mellor, A., Fahrner, K., Kvist, S. and Flavell, R.A. (1983) EMBO J. 2, 453–462.
[44] Mellor, A.L., Weiss, E.H., Ramachandran, K. and Flavell, R.A. (1983) Nature 306, 792–795.
[45] Clarke, S.H., Claflin, J.L. and Rudikoff, S. (1982) Proc. Natl. Acad. Sci. USA 79, 3280–3284.
[46] Bentley, D.L. and Rabbitts, T.H. (1983) Cell 32, 181–189.
[47] Krawinkel, U., Zoebelein, G., Bruggemann, M., Radbruch, A. and Rajewsky, K. (1983) Proc. Natl. Acad. Sci. USA 80, 4997–5001.
[48] Slightom, J.L., Blechl, A.E. and Smithies, O. (1980) Cell 21, 627–638.
[49] Schreier, P.H., Bothwell, A.L.M., Mueller-Hill, B. and Baltimore, D. (1981) Proc. Natl. Acad. Sci. USA 78, 4495–4499.
[50] Ollo, R. and Rougeon, F. (1983) Cell 32, 515–523.