Clin. exp. Immunol. (1992) **88,** 430–434

A somatically mutated VκIV gene encoding a human rheumatoid factor light chain

A. GAUSE, R. KÜPPERS* & R. MIERAU† *I. Medizinische Klinik, Universitätskliniken des Saarlandes, Homburg/Saar, *I. Medizinische Klinik, Universität Köln D-5000 Köln, and †Rheumaklinik Aachen, Aachen, Germany*

(Accepted for publication 27 January 1992)

SUMMARY

The light chain of an IgA κ rheumatoid factor (RF) produced by a hybridoma derived from a patient with rheumatoid arthritis (RA) has been shown to belong to the VκIV family. This RF light chain has 31 nucleotide differences compared with the single VκIV germline gene reported for the human genome. The patient's VκIV germline gene was sequenced, using the polymerase chain reaction (PCR), and shown to be identical to that previously reported. This demonstrates that the RF light chain is the product of a somatically mutated gene. A comparison with other known VκIV sequences shows that the RF light chain has more replacement mutations than most of the known VκIV light chains.

Keywords rheumatoid arthritis rheumatoid factor somatic mutation immunoglobulin V gene

INTRODUCTION

Classical rheumatoid factors (RF) are IgM autoantibodies with specificity for the Fc-region of IgG. RF may also occur as 'natural' antibodies that are thought to be a normal physiological component of the immune system [reviewed in 1]. In certain diseases, however, RF of different isotypes [reviewed in 2] may be produced which are a hallmark of a pathological tissue damaging state, e.g. in the course of rheumatoid arthritis (RA) or Sjögren's syndrome.

Serological protein sequencing and DNA sequencing studies have shown that cryoglobulins with RF activity preferentially use a restricted set of VκIII light chains encoded by unmutated or minimally mutated germline genes [reviewed in 3–5]. The corresponding heavy chains show a greater diversity of V_H gene usage.

The cross-reacting idiotypes that characterize the light and heavy chains of paraproteins with RF activity have been demonstrated at significantly lower frequencies among RF produced by hybridomas established from synovial B cells of RA patients [6–8]. A greater heterogeneity of RF from RA patients is further shown by their use of different isotypes and lambda light chains in addition to kappa chains [9]. This diversification of RF in RA patients leads to the hypothesis that these human RF are selected by an unknown antigen and have undergone somatic mutations in analogy to the MRL/lpr mouse model, where clonal expansion and accumulation of somatic mutations in antibody V regions could be demonstrated [10].

We were able to characterize a human heterohybridoma derived from B cells of a patient with RA which is producing an IgA RF [11]. The V-region genes encoding the heavy and light chains of this RF were sequenced. The expressed V_HI gene is related to two other V_H genes used by RF [12,13]; however, the corresponding germline gene is not known. In contrast, the light chain is a VκIV gene differing in 31 nucleotides from the published germline sequence. As it is known that the VκIV gene family has only one member [14,15], the nucleotide differences must be due to somatic mutations or to polymorphism of the VκIV gene in the human population. In order to discriminate between these two mechanisms, we sequenced the corresponding germline gene of the patient from whom the IgA RF producing hybridoma was derived.

Correspondence: Angela Gause, I. Medizinische Klinik, Universitätskliniken des Saarlandes, D-6650 Homburg/Saar, Germany.

MATERIALS AND METHODS

DNA preparation

Peripheral blood mononuclear cells were prepared from 20 ml heparinized blood by centrifugation on Ficoll–Hypaque. DNA was prepared from the supernatant of a guanidinium/CsCl centrifugation for RNA isolation [16]. The supernatant fluid was dialysed against TE (10 mM Tris–HCl, pH 8·0, 1 mM EDTA) followed by three extractions with phenol/chloroform/isoamylalcohol (25:24:1), one extraction with chloroform/isoamylalcohol (24:1) and precipitation with ethanol.

Polymerase chain reaction (PCR)

One-hundred nanograms DNA were used for a PCR with one oligonucleotide complementary for the 5′ untranslated region (5′CGGGCCGTTTGCATTGTGAACTGAGC 3′) and one for the 3′ spacer and nonamer germline region (5′ GGTTTGTGTTCGAGGCTGAAG 3′). The PCR was per-

formed with 250 nM of each oligonucleotide, 0·2 mM dNTP, 50 mM KCl, 10 mM Tris, pH 8·4, 1·5 mM $MgCl_2$ and 2·5 U Taq polymerase (Cetus). The enzyme was added after a pre-incubation of 5 min at 96°C, followed by 2 min at 55°C and 1 min at 72°C. The following programme consisted of 30 cycles of 30 sec at 95°C, 30 sec at 55°C and 90 sec at 72°C. The synthesis was finished with 5 min of incubation at 72°C.

Cloning and sequencing

The PCR product was controlled by electrophoresis in a 1% agarose gel and prepared by electrophoresis in 2% low melting agarose (NuSieve, Biozym). The excised band was isolated with a geneclean kit (BIO 101) according to the recommendations of the manufacturer. For dG-tailing the purified band was incubated with 140 mM cacodylic acid, 30 mM Tris–HCl, pH 6·8, 1 mM $MnCl_2$, 67 nM dGTP, 0·1 mM DTT with TDT (Pharmacia) for 20 min at 37°C. dG-tailed cDNA was annealed to a *Pst*1 cut and dC tailed pTZ19U vector [17]. *Escherichia coli* strain DH5 alpha was transformed with the annealed plasmid using the procedure of Hanahan [18]. Colonies with plasmids containing inserts of the expected length were expanded and plasmids prepared from about 20 ml cultures. dsDNA sequencing was carried out with Sequenase™ Version 2.0 (USB) with the same primers as used for the PCR.

Calculation of the ratio of replacement to silent mutations

Assuming that point mutations occur randomly over the whole V gene an expected ratio of replacement to silent mutations can be calculated according to the nucleotide composition of the VκIV germline gene. This calculation is based on the assumption that at every position of the codon a point mutation can result in being replaced by three other different nucleotides. One such exchange can either lead to a replacement mutation which results in a different amino acid or a stop codon or can leave the coded amino acid unaffected. For example the ratio of replacement:stop:silent mutations is 7:1:1 for AAG and AAA, 8:0:1 for AAC, AAT and AGC, 6:0:3 for ACX, GCX and CCX and 9:0:0 for ATG, and so on. The ratios of replacement:stop:silent mutions for every codon of the VκIV germline sequence were added and the ratio of the sum of replacement and stop mutations to silent mutations was calculated for the whole gene and for framework and CDR regions separately.

RESULTS

VκIV germline sequence

Figure 1 shows the consensus sequence of eight independent clones of the patient's VκIV germline gene derived from genomic DNA of peripheral blood mononuclear cells. It is identical to the published germline sequence. In more than 3000 bp sequenced there are only six individual nucleotide differences at different positions. There are four exchanges from T to C (nucleotide positions −157, −141, 97, 234) and two exchanges from A to G (nucleotide positions −18, 26) as typical artefacts caused by infidelity of the Taq polymerase [19]. As there are at least seven identical nucleotides at the same position which are derived from independent clones from the same reaction, it can be assumed that the sequence is determined correctly and polymorphism of the patient's VκIV germline gene can be excluded.

Comparison of the mutated light chain with published sequences

Figure 2 compares the deduced amino acid sequence of the IgA RF P61B27 with the germline amino acid sequence and with other published VκIV amino acid sequences. The P61B27 amino acid sequence is one of the sequences with most amino acid exchanges (19/95) of all VκIV light chains known up to now [summarized in 20,21]. Only DA-H and DA-N which are derived from an IgA/IgG paraprotein [22] have relatively more amino acid exchanges (7/25 and 7/32, respectively). For some

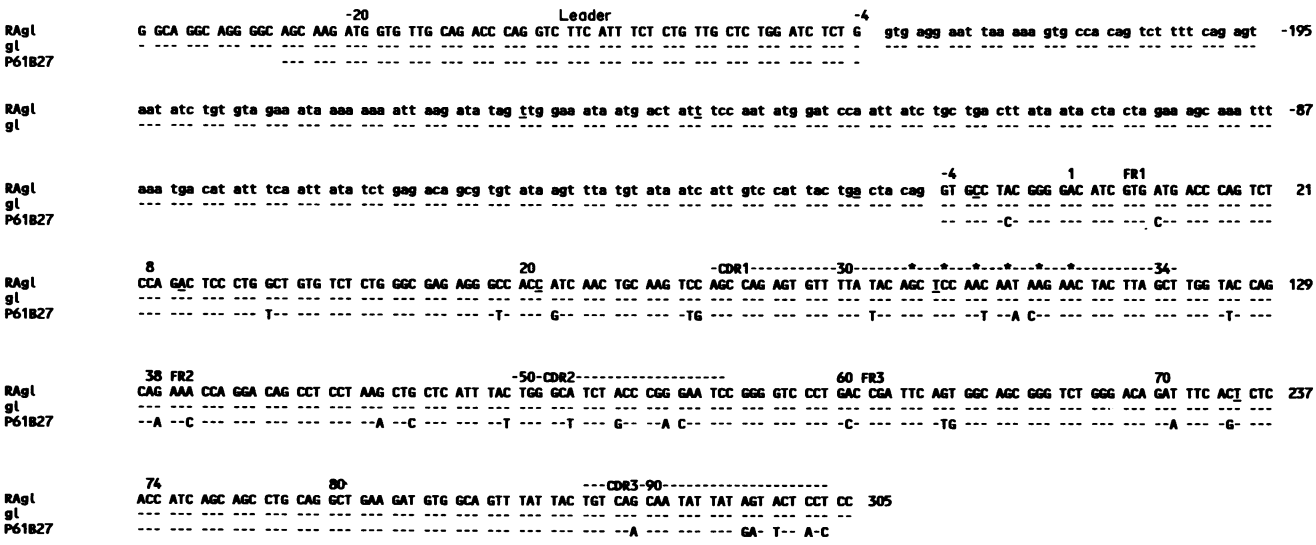

Fig. 1. The VκIV germline sequence of the patient with rheumatoid arthritis (RAg1) shown as consensus sequence of 8 individual clones. In positions −157, −141, 97 and 234 (underlined) the thymidine was substituted by a cytosine in one of eight sequences, in positions −18 and 26 adenosine was substituted by guanidine in one of eight sequences by mis-incorporation through the Taq polymerase. The consensus sequence is compared to the published germline sequence (gl) [14] and to the RF sequence (P61B27) [11]. Numbering is according to Klobeck [14], numbers at the ends of the lines designate nucleotides, numbers above the sequence designate amino acids, the asterisks designate additional amino acids in CDR1 characteristic for the VκIV family. The intron sequence is given in lower-case letters.

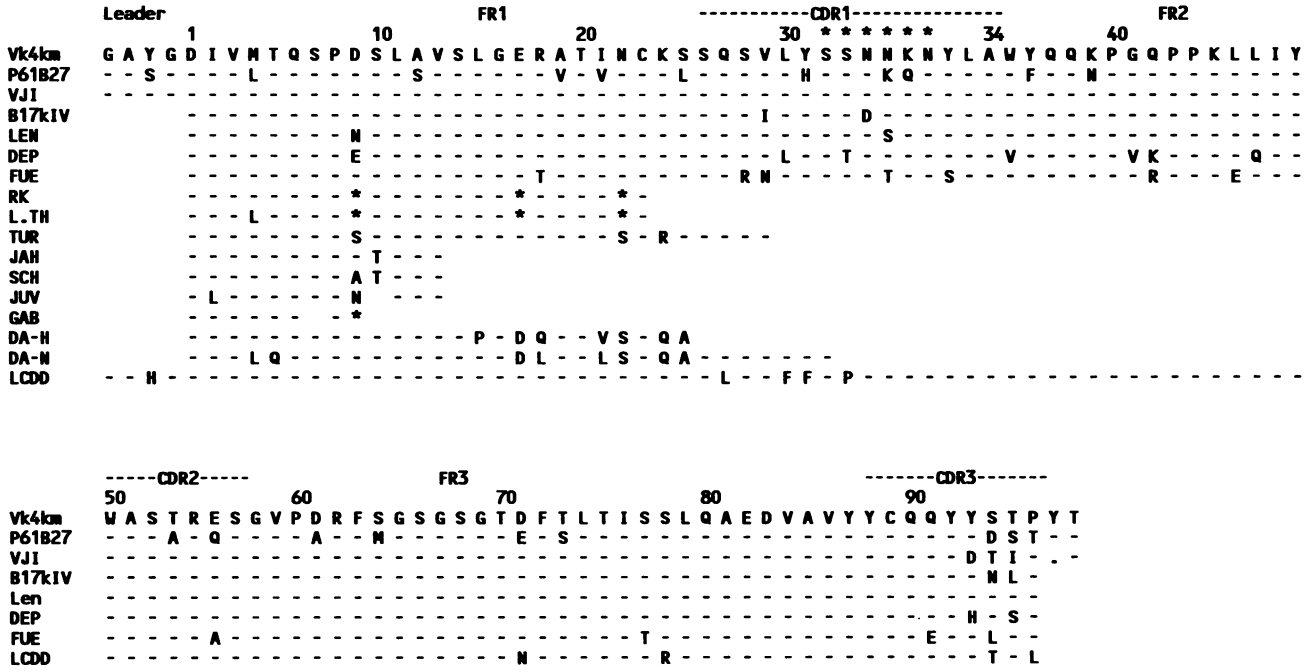

Fig. 2. Comparison of the deduced amino acid sequence of the P61B27 light chain to all other published VκIV sequences [20–24]. Numbering is according to Kabat *et al.* [20], where most protein sequences are summarized.

Table 1. Distribution of nucleotide and amino acid exchanges of the P61B27 light chain over framework (FR) and complementarity determining regions (CDR) compared with the distribution expected for the lengths of these regions

	Nucleotide exchanges	Amino acid exchanges	Replacement/ silent
CDR	16/96	9/32	1·5
FR	15/207	10/69	1·6
CDR expected	10/96	6/32	4·1
FR expected	21/207	13/69	3·2

Replacement/silent gives the ratio of replacement to silent mutations calculated for the P61B27 sequence and the one expected for the VκIV germline gene if point mutations occurred randomly.

proteins autospecificity is reported, e.g. DEP and FUE are IgM paraproteins with anti-myelin activity [23] and RK and L.TH are IgM paraproteins with cold agglutinin specificity [24]. A RF specificity of an antibody with a VκIV light chain has not been reported.

Analysis of somatic mutations of the RF light chain

By absolute numbers the distribution of nucleotide and amino acid exchanges is equally distributed over the framework (FR) and complementarity determining regions (CDR) (Table 1). Relative to the lengths of the regions there are, however, more exchanges in the CDR (32 amino acids with 96 nucleotides) than in the FR regions (69 amino acids with 207 nucleotides), when CDR1 is regarded as running from position 24 to 34 plus the six additional amino acids typical for the VκIV gene, CDR2 from position 50 to 56, and CDR3 from position 88 to 95. For the P61B27 sequence the ratio of replacement to silent mutations is 1·5:1 for all CDR and 1·6:1 for the FR regions; these values are clearly lower than expected if the somatic mutations had occurred randomly. The calculation of the ratio of replacement to silent mutations for the VκIV germline gene produces a ratio of 3·46:1. Calculated for each region it is 2·9:1 for FR1, 4·7:1 for CDR1, 3·8:1 for FR2, 2·7:1 for CDR2, 3·2:1 for FR3 and 5:1 for CDR3.

DISCUSSION

Our previous work had revealed 31 nucleotide exchanges in the VκIV light chain gene of an IgA RF hybridoma, compared with the known human VκIV germline gene [11]. However, in order to exclude polymorphism in the human population as the underlying cause of the sequence differences, it was necessary to sequence the corresponding germline gene of the patient from whom the hybridoma had been derived. For the VκIV gene this could be accomplished using PCR, since there is only one human germline gene of the VκIV family [14,15].

Although there are some RF light and heavy chains reported which differ from known germline genes, there is so far no direct comparison of a gene encoding a RF with the germline gene of the same individual [12,25–27]. While we present the sequence of a single chain of one RF, our findings are of importance to the recent discussion on the biological significance of RF, for the following reasons. Most of the RF sequenced so far are encoded by V genes of the larger V gene families (VκIII or V_HI). In these cases it cannot be determined whether sequence differences are derived from somatic mutations, from variability of V-gene family members or from polymorphism in the human population, as the corresponding germline genes from the same patient are not known. Although sequencing of germline genes using

PCR is easy, the number of clones that must be sequenced for statistical reasons is increasing with increasing size of the V-gene family. As the exact number of family members is not known for the large human V_H-gene families (e.g. V_HI and V_HIII), this approach becomes practically impossible.

Recent publications concerning the sequences of V genes encoding RF from patients with RA also present strong evidence that some RF from patients with RA differ from those from patients with lymphoproliferative disorders and that somatic mutation is involved in the diversification [12,25–27]. These investigations extend and support earlier serologic studies of fine specificities and cross-reacting idiotypes of hybridomas and serum immunoglobulins of patients with RA and healthy controls [6–9,28,29].

The finding of unmutated germline V-region genes in a RF producing hybridoma of a patient with RA [30] does not contradict the above findings as the somatic diversification starts from unmutated germline genes so that these genes and RF with the typical cross-reacting idiotypes should also be found.

Investigations in the mouse [reviewed in 31] revealed that the process of somatic mutation is driven by antigen selection and accompanied by the isotype class switch during the process of clonal expansion. Therefore, the fact that the RF we analysed is an IgA isotype fits well with the finding of a somatically mutated V region.

Although the distribution of replacement and silent mutations are not comparable, the number of somatic mutations in the P61B27RF light chain (31 nucleotide exchanges and 19 amino acid exchanges) is comparatively high and in the range of that of secondary response antibodies analysed in the mouse. The distribution of the mutations shows a predominance of nucleotide and amino acid exchanges in the CDR when the length of the regions is taken into account (Table 1). The ratio of replacement to silent mutations of 1·6:1 is, however, lower than expected if somatic mutations were introduced randomly. A low rate of replacement to silent mutations argues for the conservation of certain amino acids by selection through an unknown mechanism.

By analogy with anti-DNA antibodies of IgG isotype in patients with systemic lupus erythematosus [32], these data suggest that the disease process of RA leads to somatic mutation of genes encoding RF and class switch to IgA and IgG isotypes occurs with the production of affinity maturated RF that accelerate the disease process, e.g. by immune complex formation, deposition of immune complexes and exacerbation of inflammation. In a mouse model of RA, the MRL/lpr mouse, oligoclonal expansion of somatically mutated RFs analogous to a secondary antigen-driven response was demonstrated [1]. Biochemical studies on immunoglobulins in synovial fluid and serum already indicate a quantitative predominance of certain clones [33,34].

Further confirmation of an antigen-driven immune reaction in RA requires studies on more RF from RA patients and the demonstration of clonal relation and expansion of RF-producing cells in humans.

ACKNOWLEDGMENTS

This work was supported by the Deutsche-Forschungsgemeinschaft through Di 184.

REFERENCES

1 Siminovitch KA, Chen P. The biologic significance of human natural autoimmune responses: relationship to the germline, early immune and malignant B cell variable gene repertoire. Int Rev Immunol 1990; **5**:265–77.

2 Hay FC, Meinir PD, Bond A, Soltys AJ. Rheumatoid factors and complex formation. The role of light-chain framework sequences and glycosylation. Clin Orthop 1991; **265**:54–62.

3 Carson DA, Chen PP, Kipps TJ, *et al.* Regulation of rheumatoid factor synthesis. Clin Exp Rheumatol 1989; **7/S-3**:69–73.

4 Chen PP, Silverman GJ, Liu MF, Carson DA. Idiotypic and molecular characterization of human rheumatoid factors. Chem Immunol 1990; **48**:63–81.

5 Carson DA, Chen PP, Kipps TJ. New roles for rheumatoid factor. J Clin Invest 1991; **87**:379–83.

6 Thompson KM, Randen I, Natvig JB, *et al.* Human monoclonal rheumatoid factors derived from the polyclonal repertoire of rheumatoid synovial tissue: incidence of cross-reactive idiotopes and expression of V_H and Vκ subgroups. Eur J Immunol 1990; **20**:863–8.

7 Koopman WJ, Schrohenloher RE, Carson DA. Dissociation of expression of two rheumatoid factor cross-reactive κ L chain idiotopes in rheumatoid arthritis. J Immunol 1990; **144**:3468–72.

8 Schrohenloher RE, Accavitti MA, Bhown AS, Koopman WJ. Monoclonal antibody 6B6.6 defines a cross-reactive kappa light chain idiotope on human monoclonal and polyclonal rheumatoid factors. Arthritis Rheum 1990; **33**:187–98.

9 Natvig JB, Randen I, Thompson K, *et al.* Probing of the rheumatoid factor (RF) V gene repertoire in rheumatoid arthritis (RA) by hybridoma clones. Clin Exp Rheumatol 1990; **8** (Suppl 5):75–80.

10 Shlomchik MJ, Marshak-Rothstein A, Wolfowicz CB, *et al.* The role of clonal selection and somatic mutation in autoimmunity. Nature 1987; **328**:805–11.

11 Mierau R, Gause A, Küppers R, *et al.* A human monoclonal IgA rheumatoid factor using the VκIV light chain gene. Rheumatol Int 1992; **12**:23–31.

12 Pascual V, Randen I, Thompson K, *et al.* The complete nucleotide sequences of the heavy chain variable regions of six monospecific rheumatoid factor derived from Epstein–Barr virus-transformed B cells isolated from the synovial tissue of patients with rheumatoid arthritis. Further evidence that some autoantibodies are unmutated copies of germ line genes. J Clin Invest 1990; **86**:1320–8.

13 Sanz I, Casali P, Thomas JW, *et al.* Nucleotide sequences of eight human natural autoantibody VH regions reveals apparent restricted use of VH families. J Immunol 1989; **142**:4054–61.

14 Klobeck HG, Bornkamm GW, Combriato G, *et al.* Subgroup IV of human immunoglobulin κ light chains is encoded by a single germline gene. Nucl Acids Res 1985; **13**:6515–29.

15 Marsh P, Mills F, Gould H. Detection of a unique human VκIV germline gene by a cloned cDNA probe. Nucl Acids Res 1985; **13**:6531–44.

16 Maniatis T, Fritsch E, Sambrook J. Molecular cloning. New York: CSH Cold Spring Harbour, 1982:187,382.

17 Mead D, Szczesna-Skorupa E, Kemper B. Single-stranded DNA 'blue' T7 promotor plasmids: a versatile tandem promotor system for cloning and protein engineering. Protein Eng 1986; **1**:67–74.

18 Hanahan D. Techniques for transformation of *E. coli*. In: DNA cloning. Oxford: IRL press, 1985:109.

19 Weiss U, Rajewsky K. The repertoire of somatic antibody mutants accumulating in the memory compartment after primary immunization is restricted through affinity maturation and mirrors that expressed in the secondary response. J Exp Med 1990; **172**:1681–9.

20 Kabat EA, WU TT, Reid-Miller M, *et al.* Human kappa light chains subgroup IV. In: Sequences of proteins of immunological interest. Bethesda, MD: NIH Publications, 1987:60–62.

21 Cogné M, Preud'homme JL, Bauwens M, *et al.* Structure of a monoclonal kappa chain of the V kappa IV subgroup in the kidney

and plasma cells in light chain deposition disease. J Clin Invest. 1991; **87**(6):2186–90.

22 Bouvet JP, Liancopoulos P, Pillot J, *et al.* Three M components IgAlambda, IgGkn, IgGkh in one patient. J Immunol 1982; **129**:1519–24.

23 Mihaesco E, Ayadi H, Cogny N, *et al.* Multiple mutations in the variable region of the κ light chains of three monoclonal human IgM with anti myelin-associated glycoprotein activity. J Biol Chem 1989; **264**:21481–5.

24 Wang AC, Fudenberg HH, Wells JV, Roelcke D. A new subgroup of the kappa chain variable region associated with anti-PR cold agglutinins. Nature 1973; **243**:126–7.

25 Victor KD, Randen I, Thompson K, *et al.* Rheumatoid factors isolated from patients with autoimmune disorders are derived from germline genes distinct from those encoding the Wa, Po, and Bla cross-reactive idiotypes. J Clin Invest 1991; **87**:1603–13.

26 Blaison G, Kuntz JL, Pasquali JL. Molecular analysis of V.III variable regions of polyclonal rheumatoid factors during rheumatoid arthritis. Eur J Immunol 1991; **21**:1221–7.

27 Harindranath N, Goldfarb IS, Ikematsu H, *et al.* Complete sequence of the genes encoding the VH and VL regions of low- and high-affinity monoclonal IgM and IgA1 rheumatoid factors produced by CD5+ B cells from a rheumatoid arthritis patient. Int Immunol 1991; **2**:865–75.

28 Kouri T, Crowley J, Aho K, *et al.* Occurrence of two germline-related rheumatoid factor idiotypes in rheumatoid arthritis and in non-rheumatoid seropositive individuals. Clin Exp Immunol 1990; **82**:250–6.

29 Shokri F, Mageed RA, Tunn E, *et al.* Qualitative and quantitative expression of VHI associated cross reactive idiotopes within IgM rheumatoid from patients with early synovitis. Ann Rheum Dis 1990; **49**:150–4.

30 Ezaki I, Kanda H, Sakai K, *et al.* Restricted diversity of the variable region nucleotide sequences of the heavy and light chains of a human rheumatoid factor. Arthritis Rheum 1991; **34**:343–50.

31 Allen D, Cumano A, Dildrop R, *et al.* Timing, genetic requirements and functional consequences of somatic hypermutation during B-cell development. Immunol Rev 1987; **96**:5–22.

32 Van Es JH, Gmelig Meyling FHJ, van de Akker WRM, *et al.* Somatic mutations in the variable regions of a human anti-double stranded DNA autoantibody suggests a role for antibody in the induction of systemic lupus erythematosus. J Exp Med 1991; **173**:461–70.

33 Carpenter AB, Huczko E, Eisenbeis CH, Kelly RH. Evidence for locally synthesized and clonally restricted immunoglobulin in the synovial fluid from rheumatoid arthritis patients. Clin Chim Acta 1990; **193**:1–12.

34 Bouvet JP, Xin WY, Pillot J. Restricted heterogeneity of polyclonal rheumatoid factors. Arthritis Rheum 1987; **30**:998–1005.