The EMBO Journal vol.4 no.2 pp.393-398, 1985

Nucleotide sequences of five anti-lysozyme monoclonal antibodies

M.J.Darsley and A.R.Rees

Laboratory of Molecular Biophysics, Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

Communicated by D.C.Phillips

The nucleotide sequences of the heavy and light chain immunoglobulin mRNAs derived from five hybridomas (Gloop 1-5) secreting IgGs specific for the loop region of hen egg lysozyme were determined. These monoclonal antibodies recognise three distinct but overlapping epitopes within the loop region. The sequences of two pairs of antibodies with indistinguishable fine specificities were similar in both chains whereas the sequences of antibodies of non-identical specificities were very different. It is proposed that the D-segments expressed in two of the antibodies (Gloop3 and Gloop4) are the products of one, or perhaps two, previously unidentified germ line D-genes. Gloop1 and Gloop2 use a D-segment previously identified in antibodies specific for the hapten 2-phenyloxazolone; however it is recombined in a different reading frame in the anti-lysozyme antibodies, producing a different amino acid sequence.

Key words: anti-protein antibodies/D-segments/frame-shift mutations/immunoglobulin sequences/V-region diversity

Introduction

The murine immune system is estimated to be capable of generating $10^6 - 10^8$ different antibody combining sites (Sigal and Klinman, 1978). The mechanisms by which this diversity is achieved are now quite well understood. The antigen binding site of an antibody is constructed from six hypervariable loops or complementarity determining regions (CDRs). Three CDRs from the heavy chain variable region and three from the light chain variable region are brought together in the tertiary structure to form a contiguous surface. The first two CDRs in both chains are encoded within the germ line V_H and V_L gene segments of which there are thought to be between 100 and 300 for both families. The third CDR in the light chain is encoded in part by the V_L genes and in part by one of four functional J_L segments (in the κ-light chains) which are at separate loci from the V_κ segments in the germ line genome. The third CDR in the heavy chain is generated from the combination of one of another family of gene segments called D (for diversity) and part of one of four J_H segments. The selection and recombination of the members of these five families into a complete antibody is believed to be independent, creating a large library of germ line-encoded combining sites. In addition, the positions of recombination between the segments are imprecise, causing variation of both length and sequence at the junctions. To add to this combinatorial diversity, nucleotides may be introduced in a template-independent manner (N-diversity) during H-chain recombination events (Alt and Baltimore, 1982) and point mutations can accumulate in both functionally recombined immunoglobulin loci during the development of a B-cell lineage. These diversity generating mechanisms are reviewed in Honjo (1983) and Tonegawa (1983). Although the mechanisms by which antibody molecules of different primary structures are generated are reasonably well understood, little is known about the way in which such sequence differences influence the specificity and/or affinity of different antibodies for their respective antigens.

To address this question several laboratories have determined the variable region sequences of a number of immunoglobulins with a common specificity. Families of myeloma proteins and hybridoma antibodies specific for the simple molecules phosphoryl chlorine (PC, Gearhart *et al.*, 1981; Crews *et al.*, 1981), 2-phenyloxazolone (Phe-ox, Kaartinen *et al.*, 1983; Griffiths *et al.*, 1984), *p*-azophenylarsonate (Milner and Capra, 1982) and α(1,6)-linked dextran (Rudikoff *et al.*, 1984) have been sequenced and more recently these studies have been extended to antibodies specific for single protein epitopes (Smith-Gill *et al.*, 1984; A.Caton, in preparation). An alternative and potentially more rewarding approach is to identify antibodies which are similar in sequences but differ in specificity. Although such antibodies are difficult to select for, a small number of cases have been reported. For example, mutation of a single heavy chain amino acid arising in a myeloma line secreting an anti-PC IgA *in vitro* (Rudikoff *et al.*, 1982) resulted in complete loss of binding affinity.

To understand fully the genetic events responsible for the generation of a particular antibody sequence a knowledge of the nucleotide sequence of the mRNA is required in addition to the amino acid sequence of the protein. Precise characterization of the joining events may then be possible and if the appropriate germ line V-region sequence is available the extent of somatic mutation, both expressed and silent, can be assessed (Crews *et al.*, 1981; Bothwell *et al.*, 1981). The development of techniques allowing direct mRNA sequencing (Hamlyn *et al.*, 1978) will allow rapid accumulation of such data.

In the study described here we have determined the V-region nucleotide sequences of five monoclonal antibodies (Gloop 1-5) specific for three distinct but overlapping epitopes within the loop region of hen egg lysozyme (HEL; Darlsey and Rees, accompanying paper). These sequences confirm the grouping of the five antibodies into three fine specificity groups: each of the two antibodies with indistinguishable epitopes (Gloop1 and Gloop2; Gloop3 and Gloop4) have very similar heavy and light chain sequences, while the sequences of antibodies recognising distinct epitopes are quite different (Gloop1,2 *versus* Gloop3,4 *versus* Gloop5). Comparisons of the sequences of the two pairs of antibodies suggest that each specificity might have arisen from the expression of a single V_κ, J_κ, V_H and J_H combination with differences arising from somatic mutation, although the expression of closely related germ line genes, particularly V_H genes, is also possible.

The D-regions expressed in the Gloop antibodies show considerable differences from published sequences. Although it

is possible that they arose by the somatic mechanisms described above it is likely that at least one new D-segment has been identified. The D-segment expressed in Gloop1 and Gloop2 is probably derived from the same germ line gene as those expressed in the anti-Phe-ox antibodies (Kaartinen *et al.*, 1983); however it is recombined into a different reading frame, resulting in an altered amino acid sequence.

The variable lengths of CDR1 in the light chains and CDR3 in the heavy chains show an inverse correlation with the sizes of the epitopes recognised by the antibodies, demonstrating the importance of the overall structure of the F_V region in determining the specificity for protein antigens.

Results

The sequences of the V_H and V_χ regions of the immunoglobulin mRNAs from the five anti-loop hybridomas and their corresponding amino acid sequences are presented in Figures 1 and 2, respectively. These sequences define three groups of antibodies which are similar in both chains; Gloop1 and Gloop2 form the first, Gloop3 and Gloop4 the second and Gloop5 the third. The sequences of antibodies from different groups are quite different. This grouping of antibodies exactly matches that derived from the specificity studies detailed in the accompanying paper. The nucleotide sequences of the V_χ-encoded regions of Gloop1 and Gloop2 differ in only 4/314 positions (1.3%) and those of Gloop3 and Gloop4 differ in 6/254 positions (2.4%) (Table I). This level of variation could easily be accounted for by somatic mutation operating on single V_χ genes for each of the two pairs of hybridomas: the V_χ segments of antibodies with different fine specificities, however, differ by between 30.7% and 40.9% (Table I). The different lengths of the CDR1 regions in the light chains (Figure 2) show that at least three V_χ genes are expressed.

The V_H sequences of Gloop antibodies with similar specificities are more similar than those with different specificities. The pairs of antibodies with similar specificities have nucleotide sequences differing by 4.8% and 3.6%; the maximum sequence difference (Gloop2 *versus* Gloop5) is 18.2% (Table I). While the pairs of similar specificities may

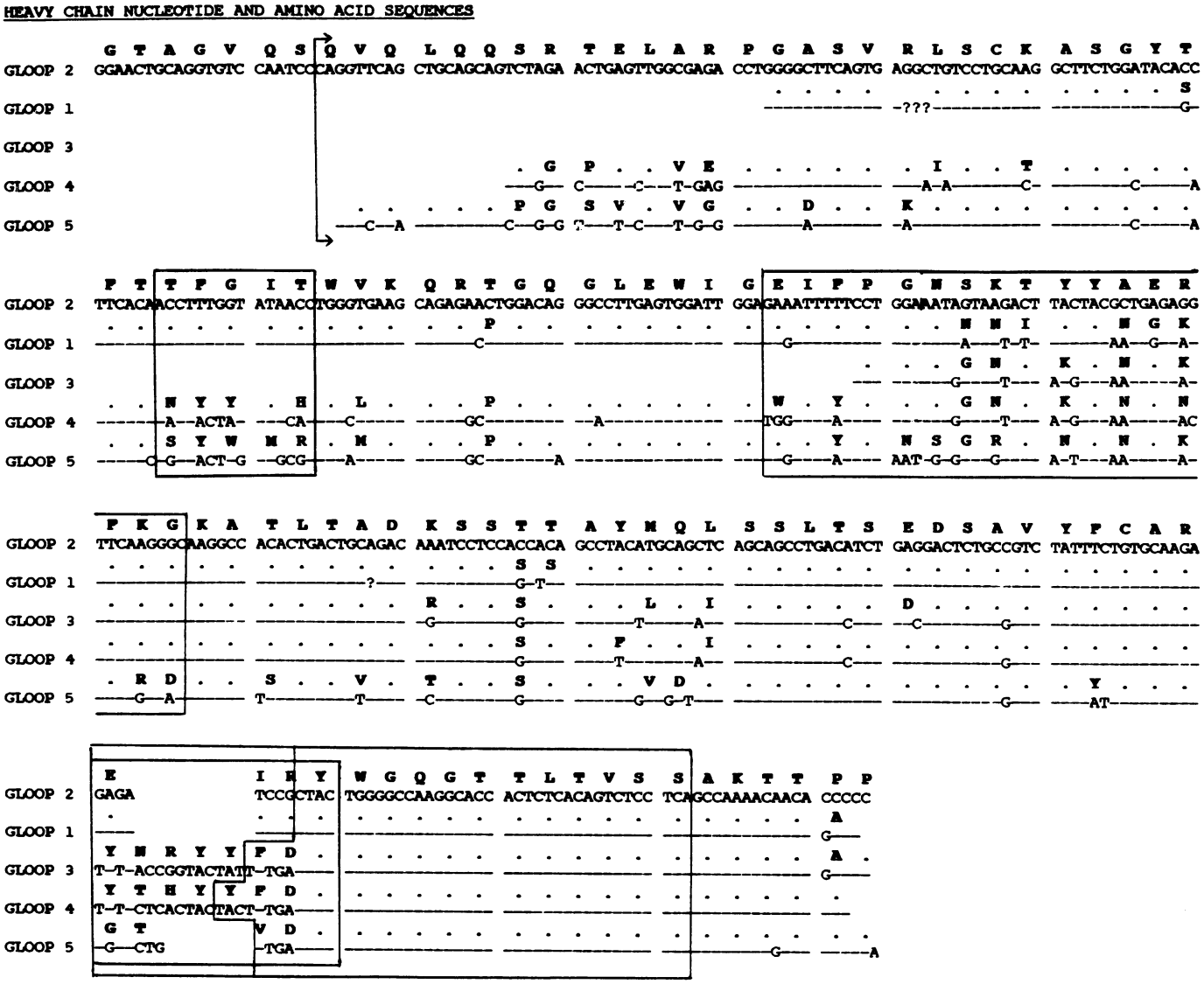

Fig. 1. Sequences are aligned with and compared with that of Gloop2, dashes indicate nucleotide identity, dots indicate amino acid identity with this sequence. Amino acids are represented by the one-letter code of Dayhoff (1969). ↱ indicates the start of the coding sequence. Narrower boxes enclose the three CDRs. Wider boxes enclose the regions encoded by the D- and J-gene segments (but see text). All heavy chains belong to the V_HII subgroup and use the J_H2 segment.

be derived from single V_H genes they may also represent the expression of two closely related V_H genes. It is likely, however, that those antibodies showing dissimilar specificities are encoded by different V_H genes (see Discussion).

In all 10 immunoglobulin chains the J-encoded segments appear to contain no somatic mutations. However, since the precise position of the D/J_H boundary is unclear (for the purpose of this discussion the boundary is assumed to be at the 5′ limit of identity of the determined sequence to the germ line J_H sequence), the 5′-terminal nucleotides in the J_H-segments may have undergone mutation. With this proviso all five heavy chains contain germ line J_H2; Gloop1 and Gloop2 contain germ line $J_\chi5$ and Gloop3, Gloop4 and Gloop5 contain germ-line $J_\chi1$ sequences. These regions contain a total of 397 nucleotides and the absence of a single nucleotide substitution demonstrates a very low mutation rate in these J-segments.

The D-segments identified in the five MABs (Figure 3) do not directly correspond to any of the D-segments sequenced to date (Tonegawa, 1983; Kaartinen *et al.*, 1983). Gloop1 and Gloop2 may use the D-segment identified in anti-Phe-ox MABs (Kaartinen *et al.*, 1983) but with a single nucleotide substitution. Interestingly, it is recombined into a different reading frame (see Discussion). Gloop5 contains a short segment which may be derived from either the D_{Q52} germ line sequence or the $D_{FL16.1}$ (or $D_{FL16.2}$) germ line sequence with one internal point change, plus a GGG sequence at the 5′ end which may be due to N-diversity (Alt and Baltimore, 1982). The D-segments expressed in Gloop3 and Gloop4 are related to the $D_{FL16.1}$ sequence but there are large differences at the 5′ end.

The lengths of the V_χ regions in CDR1 and the V_H regions in CDR3 of immunoglobulins are variable. The lengths of these CDRs are identical in those Gloop antibodies with similar specificities but different in those with different specificities. CDR1 in the light chains of both Gloop1 and Gloop2 is 10 residues long, in Gloop3 and Gloop4 15 residues long and in Gloop5 16 residues long (Figure 2). CDR3 in the heavy chains of both Gloop1 and Gloop2 is four residues long, in Gloop3 and Gloop4 eight residues long and in Gloop5 five

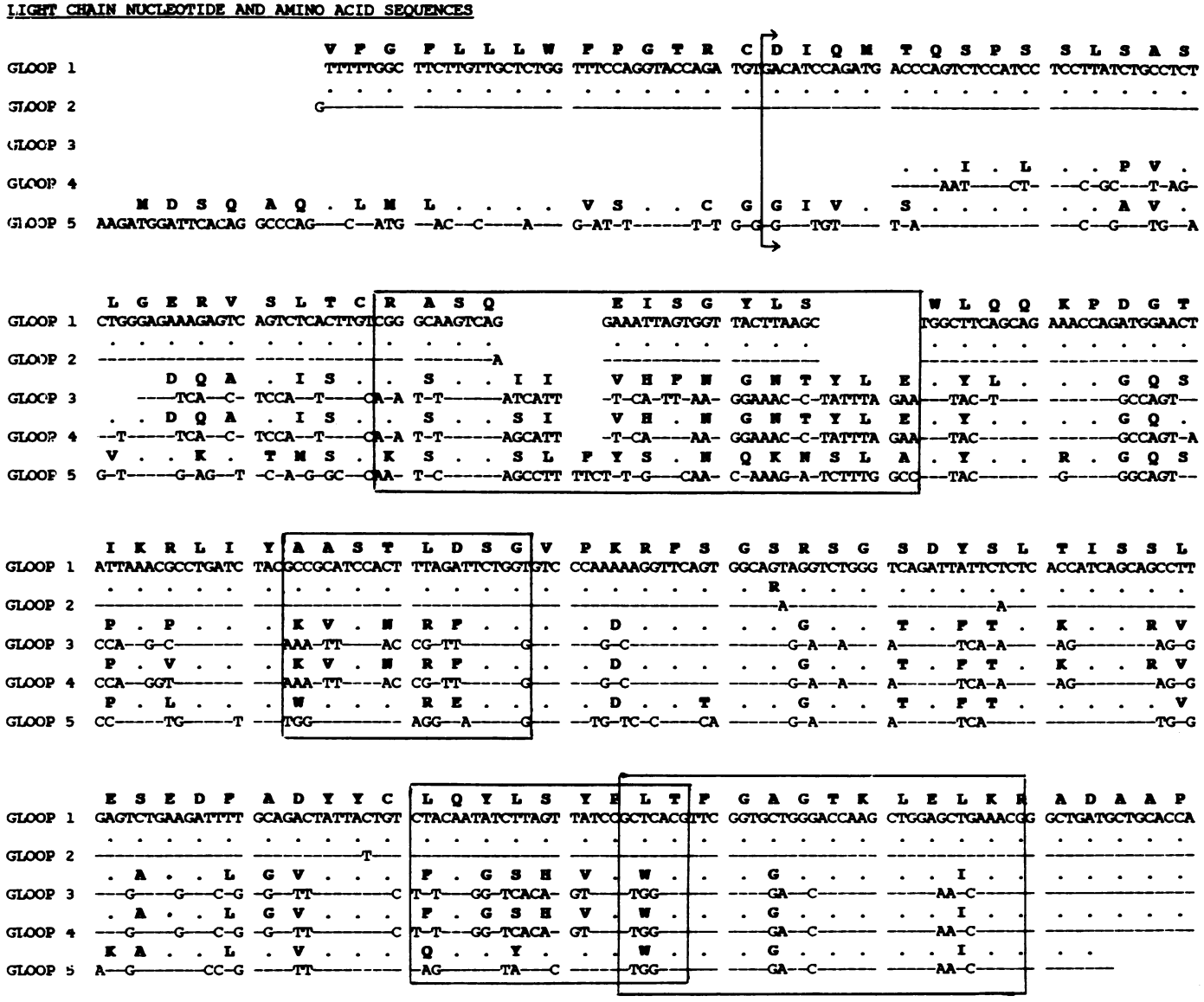

Fig. 2. Sequences are aligned with and compared with that of Gloop1, symbols and boxes are as for Figure 1. Gloop1 and Gloop2 belong to the $V_\chi9$ subgroup and use $J_\chi5$. Gloop3 and Gloop4 belong to the $V_\chi3$ subgroup and use $J_\chi1$. Gloop5 belongs to the $V_\chi8$ subgroup and uses $J_\chi1$.

Table I. Nucleotide and amino acid differences between all pairs of Gloop antibodies

		Nucleotide differences			
		Gloop1	Gloop2	Gloop3	Gloop4	Gloop5
Amino acid differences	G1		12/248 (4.8%) 4/314 (1.2%)	13/138 (9.4%) 104/254 (40.9%)	30/296 (10.1%) 115/288 (39.9%)	40/291 (13.7%) 124/344 (36.0%)
	G2	10/82 (12.2%) 1/105 (0.9%)		14/138 (10.1%) 106/254 (41.7%)	40/276 (14.5%) 117/288 (40.1%)	53/291 (18.2%) 128/345 (37.1%)
	G3	9/46 (19.6%) 48/84 (57.1%)	10/46 (21.7%) 49/84 (58.3%)		5/138 (3.6%) 6/254 (2.4%)	20/138 (14.5%) 79/257 (30.1%)
	G4	18/92 (19.6%) 49/96 (51.0%)	9/92 (9.8%) 50/96 (52.1%)	5/46 (10.9%) 5/84 (6.0%)		42/276 (15.2%) 94/291 (32.3%)
	G5	25/97 (25.7%) 55/115 (47.8%)	32/97 (33.0%) 56/115 (48.7%)	14/46 (30.4%) 35/85 (41.2%)	29/92 (31.5%) 38/97 (39.2%)	

The table shows: number of differences/length of comparison (percentage difference).
Upper line gives differences between heavy chains.
Lower line gives differences between light chains.

```
Germ line D FL16.1          TTTATTACTACGGTAGTAGCTAC
Germ line D FL16.2          --C----------******----
Germ line D Q52                   -A--T-GGAC

GLOOP 5                           GGG--T--T-AC-

GLOOP 3                         -ATAAC-----C--TTTT-
GLOOP 4                         -ATACT-AC--C--CTTT-

GLOOP 3'                      ---A--**-----C--TTTT-
GLOOP 4'                      ---ACT**-AC--C--CTTT-

GLOOP 3"                       TA--AC*-----C--TTTT-
GLOOP 4"                       TA----*-AC--C--CTTT-

Phe-Ox Concensus               GATCGGGGG

GLOOPS 1 & 2                GA----C-CTAC
```

Fig. 3. All sequences except the last two are aligned with that of $D_{FL16.1}$, dashes indicate identity. The Gloop1/Gloop2 D-sequence is expressed similarly, relative to the consensus Phe-ox D-sequence. Underlined nucleotides are assumed to be derived from J_H segments as they are identical to germ line J_H sequences. The stars represent gaps inserted into the sequences of $D_{FL16.2}$, Gloop3 and Gloop4 which increase the homology with $D_{FL16.1}$. It is postulated that Gloop3 and Gloop4 D-segments were derived from one or two previously unidentified germ line segments which arose from the $D_{FL16.1}$ segment by gene duplication followed by somatic mutation, including the indicated deletions, in the same way that $D_{FL16.2}$ may have been.

```
                    ← VH          D          J →
Phe-ox concensus    Cys Ala Arg   AspArgGly
                    TGT GCA AGA **GATCGGGGG

Gloop1 and Gloop2   --- --- --- GA----C-*** C
                    Cys Ala Arg GluIleArg
```

Fig. 4. The sequence of Gloop1/Gloop2 around the D-region aligned with the consensus sequence observed in anti-Phe-ox hybridomas. Dashes indicate nucleotide identity, stars indicate gaps inserted to maximise homology. The resulting amino acid sequences are shown above and below the respective nucleotide sequences.

residues long (Figure 1). Thus the lengths of these variable loops in each of the antibodies show an inverse correlation with the size of the epitopes recognised by the antibody (see accompanying paper).

Discussion

V- and J-segment usage and somatic mutation

As described in the Results section the sequences of both light and heavy chains of antibodies with different fine specificities are so different that it is likely they arose from different germ line V_χ and V_H genes. Whether the converse is true, that is whether antibodies with the same specificity (Gloops1 and 2; Gloops3 and 4) have been produced from the same V genes, is more difficult to assess. The small number of differences between the light chains of Gloops1 and 2 are compatible with their being derived from the same germ line gene. The same argument holds for Gloops3 and 4. It is impossible, however, to say whether the heavy chains of those antibodies with the same fine specificity are either derived from the same germ line V_H gene by somatic mutation or are the products of closely related V_H genes. A considerably higher level of somatic mutation would have to have occurred in the V_H segments than that occurring in the V_χ, J_χ or J_H segments for the heavy chains of Gloops1 and 2 and Gloops3 and 4 to have been derived from the same V_H germ line genes. This question cannot be resolved by comparison of the extent of sequence differences known to be due to germ line variation or to somatic mutation. Two families of related V_H genes have been previously sequenced, the T15 family (Crews *et al.*, 1981) and the NP^b family (Bothwell *et al.*, 1981). The four T15-like V_H genes differ from one another by between 19 and 42 nucleotides. The seven members of the NP^b family are more closely related to one another, differing by between three and 29 nucleotides. The maximum degree of nucleotide sequence divergence which has been shown to be due to somatic mutation is 10 nucleotides in both the M167 and the S43 V_H regions (Bothwell *et al.*, 1981; Early *et al.*, 1980). Up to eight amino acid substitutions have been shown to be caused by somatic mutation of the Vκ21C gene (McKean *et al.*, 1984). The levels of differences between Gloop1 and Gloop2 and between Gloop3 and Gloop4 heavy chains fall within the observed range of both mechanisms and cannot therefore be attributed to either with certainty. This highlights the difficulties encountered when attempting to decide whether two similar sequences have been produced from related, but

distinct, germ line genes or from a single gene by the process of somatic mutation. The question can only be definitively answered by comparison of the expressed sequences with those of the germ line genes themselves (Kim *et al.*, 1981).

D-segment usage

The D-segments observed in the Gloop antibodies are either derived from previously unidentified germ line gene segments or have been generated by a level of somatic mutation much higher than observed in the J-segments. The D-segments from Gloops 1 and 2 contain one substitution relative to the anti-Phe-ox consensus sequence (Kaartinen *et al.*, 1983), suggesting that they were derived from the same germ line segment, and they may possess up to two template-independent nucleotides at their 5′ termini. Especially interesting is the fact that if this attribution for Gloop1 and 2 is correct, rather than their representing a new germ line gene, then the junctional diversity of V_H-D joining has resulted in a frame-shift in the D-segment, compensated for in the D-J_H junction to regain the correct reading frame (Figure 4). This has important implications for the number of heavy chains which can be produced by rearrangements of the germ line pools, effectively trebling the number of available D-segments. To our knowledge this is the first time that such a frame-shift has been observed as a means of producing additional sequence variability in immunoglobulin heavy chains.

The Gloop5 D-segment may derive from either the D_{Q52} segment or from the $D_{FL16.1}$ segment with a single mutation. If it is derived from D_{Q52} or $D_{FL16.1}$ only four or three nucleotides respectively would be encoded by the germ line D-segment, the three 5′ nucleotides being template independent.

The D-sequences expressed in Gloops 3 and 4 are interesting. Their 3′ ends could have derived from $D_{FL16.1}$ with two and three mutations respectively, in which case the six 5′-terminal nucleotides would have arisen from N-diversity but, since the two sequences are identical in four of these six positions, this seems unlikely. Rather, it is more probable that they represent the products of one, or perhaps two, previously unidentified germ line D genes. In Figure 3 we show how such new germ line genes could have been generated from the established $D_{FL16.1}$ sequence by gene duplication followed by deletion of one or two nucleotides plus point mutations. Such a mechanism could have given rise to the $D_{FL16.2}$ germ line gene sequence (Tonegawa, 1983) from the $D_{FL16.1}$ sequence by a duplication event followed by a six nucleotide deletion and a single C for T substitution (Figure 3).

Gloop3 and Gloop4 contain D-segments which differ from each other in five nucleotides, resulting in two amino acid differences in heavy chain CDR3. These two antibodies have identical fine specificities for the loop (although they differ by a factor of 3.5-fold in affinity). This is perhaps surprising as, due to the contribution of the D-segments and double junctional diversity, heavy chain CDR3 is the most variable in sequence and length between antibodies and therefore might be expected to play a large part in producing antibodies of different specificities. This high level of structural diversity is manifested in the Gloop antibodies and yet the different sequences of Gloop3 and Gloop4 in this region do not produce different fine specificities. A recent demonstration that antibodies with differences in the D-encoded region have identical specificities for phosphorylcholine but have lost idiotypic cross-reactivity (Pollok *et al.*, 1984) may suggest that structural diversity of this hypervariable region has functions additional to increasing the diversity of binding specificities.

Size of epitopes and nature of binding sites

The similarities in sequence between Gloop1 and Gloop2 and between Gloop3 and Gloop4 support the conclusion reached in the accompanying paper that these pairs of antibodies recognise very similar epitopes. In addition, the inverse correlation between the lengths of the two variable sized CDRs (L1 and H3) and the proposed extents of their epitopes adds support to these conclusions. Gloop1 and Gloop2 recognise the largest epitope, which is a relatively flat face of the loop surface comprising the side chains of at least eight residues and covering an area of some 16 Å x 15 Å. The corresponding antibody binding sites have the shortest H3 and L1 CDRs which produce an essentially flat surface for interaction with antigen (de la Paz, Darsley, Sutton and Rees, in preparation). In contrast, Gloop3 and Gloop4 which recognise a smaller epitope forming a ridge at the edge of the loop, have longer H3 and L1 CDRs. The extra residues in these two CDRs contribute two raised sides to the binding site with a concave surface between them providing complementarity to the smaller epitope. Gloop5 recognises an epitope intermediate in size between the two epitopes described above; this antibody has a long L1 CDR but a short H3 CDR, producing a binding surface intermediate in size between the other two pairs of antibodies.

The gross architecture of the F_v portion of an antibody is thus shown to be very important in determining the specificity for protein epitopes. This is to be expected since we have demonstrated that single epitopes may comprise relatively large areas of the protein surface, requiring that antibody/antigen contact is dependent on correspondingly large areas of the antibody surface. The nature of the antibody interaction with large complex antigens is thus fundamentally different to the interaction with small, essentially rigid, hapten groups. In the latter case the binding energy has to be provided by a small number of contacts with the antibody. The binding sites thus take the form of cavities or grooves allowing contact with a high proportion of the hapten group (e.g., MOPC 315 α-DNP, Padlan *et al.*, 1976; McPc603 α-PC, Segal *et al.*, 1974). The requirement for an accurately constructed binding pocket may account for the restricted response elicited to such haptens in terms of antibody sequence (e.g., Phe-ox, Kaartinen *et al.*, 1983; Griffiths *et al.*, 1984; PC, Crews *et al.*, 1981). The demonstration that the substitution of a single amino acid, not directly involved in hapten binding, in the heavy chain of S107 can abolish the binding of phosphorylcholine (Rudikoff *et al.*, 1982) highlights the low level of tolerance to structural changes exhibited by such binding sites. The specificity of the interaction between an antibody and a protein on the other hand will derive from the summation of a relatively large number of contacts. The other important difference in the two types of interaction is that the surface of a protein is not a rigid entity in the way that a hapten is. The groups responsible for the antibody interaction may exhibit considerable freedom of movement relative to each other. Recent evidence shows that antibodies raised to peptide fragments of proteins show most cross-reactivity with the native structure if they are directed to regions of the protein which are highly mobile (Tainer *et al.*, 1984). This does not necessarily mean, however, that such mobile regions of a protein's surface are more antigenic when the whole molecule

is used as immunogen. An alternative hypothesis is that antigenicity correlates with the relative accessibility of regions of a protein's surface. The work presented here does not address the question of the relative antigenicity of different regions within a protein.

The outcome of these factors (relatively large area of interaction and mobility of interacting groups) is that there may be a number of ways in which the immune system can produce an antibody specific for a particular region of a protein.

In this work we have identified three types of interaction between monoclonal antibodies and the loop region of HEL, although the epitopes recognised by all of the antibodies overlap; The epitope recognised by Gloop5 differs only slightly from that recognised by Gloop1 and Gloop2 and yet the primary structures of the two types of antibody are very different. Those antibodies which recognise indistinguishable epitopes do have similar sequences suggesting that the observed restriction of response to haptenic antigens may have a counterpart in the anti-protein response. This has been suggested previously from the N-terminal protein sequences of two antibodies recognising very similar epitopes of HEL (Smith-Gill *et al.*, 1984). It will be necessary to determine the sequences of a large number of antibodies with identical specificities to address this question more rigorously.

With the availability of the sequences of the V-regions of these five antibodies, combined with detailed definitions of the epitope recognised by each, we are attempting to model their binding sites using molecular graphics. Ultimately we hope to predict the structure of selected Gloop antibody/antigen complexes. Using this information together with the techniques of gene cloning and oligonucleotide-directed mutagenesis it will be possible to construct new anti-protein immunoglobulins with altered affinities or novel specificities.

Materials and methods

Preparation of RNA from hybridomas

Five hybridomas secreting IgG antibodies specific for the loop region of hen egg lysozyme (HEL) were isolated as described previously (Darsley and Rees, accompanying paper). Cells were maintained in logarithmic growth in tissue culture for at least 1 week prior to harvesting by subculturing daily to a density of 2 x 10^5 cells/ml in Dulbecco's Modified Eagles Medium (DMEM, Flow Laboratories) supplemented with 10% foetal calf serum. Alternatively cells were grown as ascites tumours in pristane primed F1 BALB/c x SWR mice.

Cells were pelleted and dissolved directly into 5 ml per 5 x 10^7 cells of lysis buffer (4 M guanidine isothicyanate; 5 mM EDTA; 0.1 M β-mercaptoethanol; 20 mM sodium citrate pH 7.0; 0.5% SDS). RNA was isolated from the lysate by centrifugation through a 5.7 M caesium chloride cushion, extraction of proteins with a 4:1 mixture of chloroform:butan-1-ol and precipitation from 70% EtOH (Chirgwin *et al.*, 1979).

Synthesis of cDNAs complementary to the variable regions of immunoglobulin mRNAs

cDNAs corresponding to both heavy and light chain variable regions were synthesised in a single reaction by the method of A.Caton (in preparation). Total cellular RNA from 10^7 hybridoma cells, or poly(A)-enriched RNA (prepared as described in Maniatis *et al.*, 1982), was denatured by treatment with 30 volumes of dimethyl sulphoxide (DMSO) at 45°C for 20 min, precipitated from 70% EtOH, and annealed with specific oligodeoxynucleotide primers which had been 5′ end-labelled with 50 μCi [γ-^{32}P]ATP (Amersham) and six units of T4 polynucleotide kinase (New England Nuclear).

The oligonucleotide primers were the kind gift of Dr A.Caton. They were a 15-mer complementary to a sequence 25 – 39 nucleotides from the 5′ end of the C_x segment: 5′-GATGGTGGGAAGATG-3′ and a 17-mer complementary to a sequence 22 – 38 nucleotides from the 5′ end of the $C_{\gamma 1}$, $C_{\gamma 2a}$ and $C_{\gamma 2b}$ segments: 5′-GGGGCCAGTGGATAGAC-3′.

Thirty units of TMV reverse transcriptase (Boehringer or Life Sciences) was used to synthesise cDNAs, extending from these primers towards the 5′ end of the mRNAs, in the presence of 3 mM dATP, dGTP, dCTP and dTTP at pH 8.2 and 42°C for 2 h. Full length cDNA copies were purified by preparative electrophoresis on 40 cm 5% polyacrylamide gels (bis:acrylamide 1:20) run at 40 W for 3 h. The separated bands containing heavy and light chain cDNAs were cut out from the gel and eluted by incubation overnight in 2 M ammonium acetate at 37°C.

Sequencing of cDNA

The cDNAs were sequenced by the method of Maxam and Gilbert (1980). Four base-specific modifications were employed to generate tracks corresponding to G; A+G; C; C+T. Four sequencing gels were run and autoradiographed using pre-flashed films and intensifying screens for 2 – 3 weeks allowing reading of up to 400 bases.

Acknowledgements

We thank Dr A.Caton for the oligonucleotide primers and for generous assistance with the mRNA sequencing methodology. This work was supported by the Medical Research Council. M.J.D. was the recipient of a M.R.C. Research Studentship.

References

Alt,F. and Baltimore,D. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4118-4122.
Bothwell,A.L.M., Paskind,M., Reth,M., Imanishi-Kari,T., Rajewski,K. and Baltimore,D. (1981) *Cell*, **24**, 625-637.
Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry (Wash.)*, **18**, 5294-5299.
Crews,S., Griffin,J., Huang,H., Calame,K. and Hood,L. (1981) *Cell*, **25**, 59-66.
Dayhoff,M.O., ed. (1969) *Atlas of Protein Sequence and Structure*, published by National Biomedical Research Foundation, Silver Springs.
Early,P., Huang,H., Davis,M., Calame,K. and Hood,L. (1980) *Cell*, **19**, 981-992.
Gearhart,P.J., Johnson,N.D., Douglas,R. and Hood,L. (1981) *Nature*, **291**, 29-34.
Griffiths,G.M., Berek,C., Kaartinen,M. and Milstein,C. (1984) *Nature*, **312**, 271-275.
Hamlyn,P.H., Brownlee,G.G., Cheng,C.C., Gait,M.J. and Milstein,C. (1978) *Cell*, **15**, 1067-1075.
Honjo,T. (1983) *Annu. Rev. Immunol.*, **1**, 499-528.
Kaartinen,M., Griffiths,G.M., Markham,A.F. and Milstein,C. (1983) *Nature*, **304**, 320-324.
Kim,S., Davis,M., Sinn,E., Patten,P. and Hood,L. (1981) *Cell*, **27**, 573-581.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning, A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
McKean,D., Huppi,K., Bell,M., Staudt,L., Gerhard,W. and Weigert,M. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3180-3184.
Milner,E.C.B. and Capra,J.D. (1982) *J. Immunol.*, **129**, 193-199.
Padlan,E.A., Davies,D.R., Pecht,I., Givol,D. and Wright,C. (1976) *Cold Spring Harbor Symp. Quant. Biol.*, **41**, 627-637.
Pollock,B.A., Kearney,J.F., Vakil,M. and Perry,R.P. (1984) *Nature*, **311**, 376-379.
Rudikoff,S.,Giusti,A.M., Cook,W.D. and Scharff,M.D. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 1979-1983.
Rudikoff,S., Pawlitta,M., Pumphrey,J. and Heller,M. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2162-2166.
Segal,D., Padlan,E.A., Cohen,G., Rudikoff,S., Potter,M. and Davies,D.R. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 4298-4302.
Sigal,N. and Klinman,N. (1978) *Adv. Immunol.*, **26**, 255-337.
Smith-Gill,S.J., Mainhart,C.J., Lavoie,T.B., Rudikoff,S. and Potter,M. (1984) *J. Immunol.*, **132**, 963-967.
Tainer,J.A., Getzoff,E.D., Alexander,H., Houghten,R.A., Olsen,A.J., Lerner,R.J. and Hendrickson,W.A. (1984) *Nature*, **312**, 127-134.
Tonegawa,S. (1983) *Nature*, **302**, 575-581.

Received on 5 December 1984