Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 2385–2388, March 1993
Immunology

Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: Clustering, polarity, and specific hot spots

ALEXANDER G. BETZ, CRISTINA RADA, RICHARD PANNELL, CÉSAR MILSTEIN, AND MICHAEL S. NEUBERGER

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, United Kingdom

Contributed by César Milstein, December 15, 1992

ABSTRACT **We have analyzed somatic hypermutation in mice carrying an immunoglobulin κ transgene in order to discriminate mutations that reflect the intrinsic specificity of the hypermutation mechanism from those highlighted by antigenic selection. We have immunized animals with three different immunogens. With one immunogen, the antigen-specific B cells express a transgenic κ chain, which does not form part of the antibody; the transgene is a passenger free to accumulate unselected mutations. With the other two immunogens, the transgenic κ chain constitutes the light chain of the expressed antibody. A comparison of the transgene mutations obtained under these different circumstances allows us to identify common features that we attribute to the intrinsic specificity of the hypermutation process. In particular, it yields only base substitutions and leads to hot spots occurring in individual positions (e.g., the second base of the Ser-31 codon). The mutations preferentially accumulate around the first complementarity-determining region. The process exhibits specific base substitution preferences with transitions being favored over transversions. We propose that these substitution preferences can be used to discriminate intrinsic from antigen-selected hot spots. We also note that hypermutation distinguishes between the coding and noncoding strands since pyrimidines (particularly thymidines) mutate less frequently than purines.**

Antibodies produced in response to repeated antigen challenges have a higher affinity for the antigen than those elicited initially. This maturation in affinity is largely attributable to a process of hypermutation and selection (1, 2). Thus, nucleotide changes are introduced into the immunoglobulin variable (V) region gene; a small proportion of these changes may improve the affinity of the antibody for the antigen and cells expressing such high-affinity antibodies will be selected (3, 4). Most information about the hypermutation is therefore heavily biased by the effect of antigenic selection; we know little about the intrinsic specificity of the process. In this work, we analyze hypermutation of an immunoglobulin transgene in order to identify these intrinsic features.

MATERIALS AND METHODS

Passenger Sequences. LK3 transgenic mice (5) were immunized i.p. with $KAl(SO_4)_2$-precipitated 4-hydroxy-3-nitrophenacetyl (NP) coupled to chicken serum albumin (200 μg) and 10^9 heat-inactivated *Bordetella pertussis*. Hybridomas were derived 3 days after secondary immunization as described (5). NP-specific hybrids were cloned twice (none reacted with the carrier protein). DNA was directly amplified by PCR (15–25 cycles) from cell lysates using transgene specific primers [VKOXBACK (5) and CGCGGATCCCTTTTCTATCCTGAAGTTCCT] and cloned into M13. At least 12 M13 κ light-chain V region of canonical anti-2-phenyloxazolone antibody ($V_κOx$) subclones were sequenced from each hybridoma, eliminating errors caused by a low frequency (1/3000) of PCR error.

Phycoerythrin (PE) Sequences. LK3 transgenic mice were immunized i.p. with PE (100 μg) and 10^9 heat-inactivated *B. pertussis* and given a booster dose (i.v. 100 μg) 3 days before isolating splenic lymphocytes. IgG-expressing lymphocytes were purified on Ficoll cushions and depleted of IgM-, IgD-, and Thy-1-expressing cells with biotinylated antibodies by use of a streptavidin-based magnetic cell separation system (MACS) sorting (6). PE-binding cells were then purified by flow cytometry and transgene sequences were amplified by PCR (50 cycles and transgene-specific primers as described above) and cloned into M13 for sequencing. Under these conditions, the average PCR error in control samples is 1/1000 to 1/1500 compared to an average mutation in samples from antigen-selected cells of 12.8/100.

Papain Digests and Immunoassays. Antibodies purified from culture supernatants (500 ml) using protein A-Sepharose were digested with 100 μl of immobilized papain (Pierce) in 20 mM sodium phosphate/10 mM EDTA/1.8 mg of cysteine per ml/250 μg of bovine serum albumin per ml, pH 7.0 (final vol, 200 μl) for 5 h at 37°C. Fab fragments were then purified by chromatography over protein A. Immunoassays were performed by incubating intact antibody (or an equivalent molarity of Fab) on NP/bovine serum albumin-coated plastic plates and then developing with biotinylated anti-light-chain antisera and avidin/peroxidase.

RESULTS AND DISCUSSION

The light chains of many of the mouse antibodies to the hapten 2-phenyloxazolone (phOx) are encoded by V regions composed of a fusion of $V_κOx1$ to the κ light-chain joining region 5 ($J_κ5$) gene segment. High-affinity, mature antibodies usually have mutations in the His-34 and Tyr-36 codons of $V_κOx1$; these mutations improve the affinity of the antibody for the phOx hapten (7). A compilation of sequences (Fig. 1*A*) also reveals other, albeit less striking, hot spots (e.g., all three bases of Ser-31; second base of Ser-77) as well as a preferential accumulation of mutations around complementarity-determining region 1 (CDR1). Analysis of the distribution of silent mutations has indicated that the hypermutation process is not random (7), and it is therefore unclear to what extent the location of observed hot spots is determined by the selecting antigen.

Different Antigens Select Different Patterns of V Gene Mutation but with Some Common Hot Spots. To see how the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: CDR, complementarity-determining region; $J_κ5$, κ light-chain joining region 5; NP, 4-hydroxy-3-nitrophenacetyl; PE, phycoerytherin; phOx, 2-phenyloxazolone; V region, variable region; $V_κOx1$, κ light-chain variable region of canonical anti-phOx antibodies; V_H, heavy-chain V region.

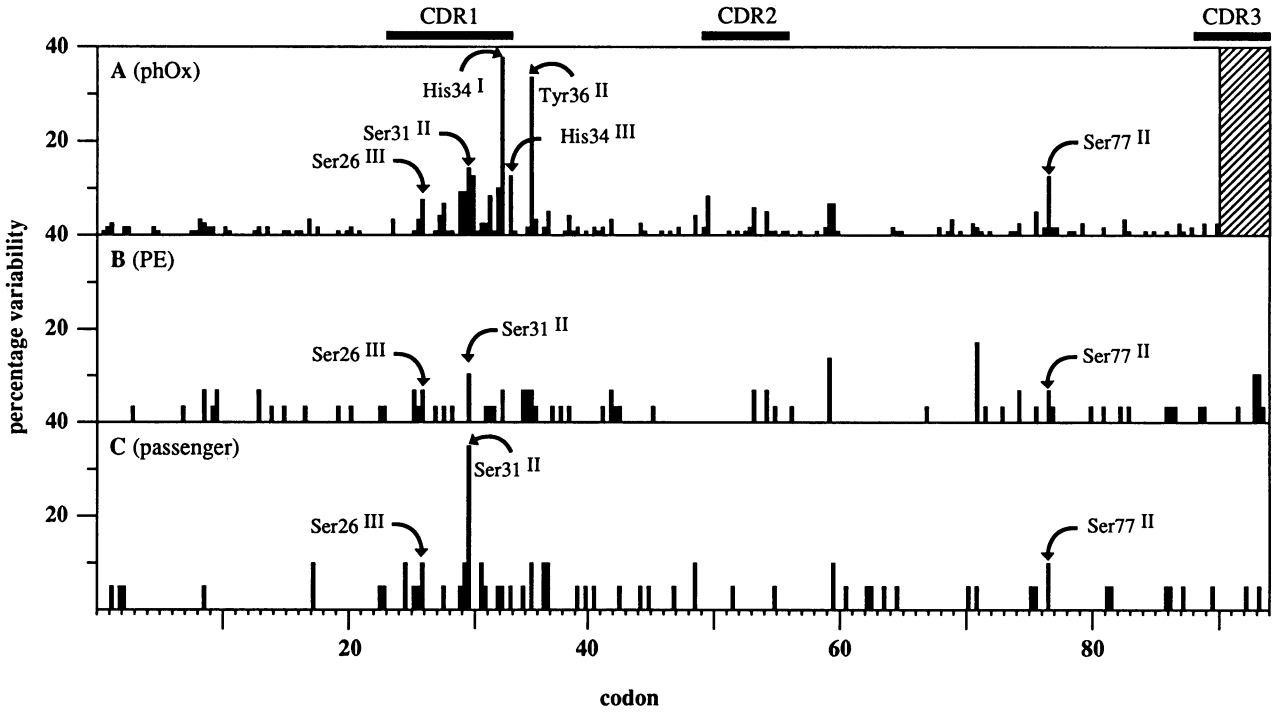

FIG. 1. Variability at each nucleotide position of V_κOx1. Variability is given as percentage of sequences that carry a mutation at each particular residue. Some individual hot spots are marked by their codon designation with the particular position in the codon indicated by a superscript in Roman numerals. (*A*) Variability in endogenous V_κOx1 sequences selected by phOx. Data are compiled from 119 previously published sequences (1, 3, 8) (total of 500 mutations) obtained from anti-phOx hybridomas established from phOx-immunized BALB/c mice. In order to remove irrelevant variation caused by junctional diversity or the use of different J_H segments, the 3′ end of the plot is hatched. (*B*) Variability in V_κOx1 sequences in PE-binding cells selected from PE-immunized LK3 transgenic mice (28 mutated sequences; 89 mutations). (*C*) Variability in passenger V_κOx1 sequences that have been derived from either anti-NP hybridomas or anti-phOx hybridomas (5) established from LK3 transgenic mice in which the transgene does not contribute the hapten binding specificity (19 mutated sequences; 71 mutations).

selecting antigen affects the pattern of mutations obtained, we used a mouse line (LK3) that carries three tandem copies of an immunoglobulin κ transgene in which the mouse V_κOx1–J_κ5 gene is linked to a rat C_κ (constant region of immunoglobulin κ light chain) segment (5). These mice give rise to anti-phOx antibodies with mutational hot spots in the transgene similar to those found in the normal anti-phOx response (5). However, the pattern of hot spots alters when selected by a different antigen. Thus, after hyperimmunizing LK3 transgenic mice with PE, determination of multiple transgene sequences from the PE-binding B lymphocytes (Fig. 1*B*) reveals no evidence for hot spots at His-34 and Tyr-36, although new hot spots appear. At least one of the transgene copies contributes to most of the anti-PE antibody since serological analysis of PE-specific B cells, of PE-specific hybridomas, as well as of PE-binding antiserum antibody all indicate that the anti-PE antibody light chain is mostly derived from the transgene with little or no contribution from endogenously encoded mouse light chains (data not shown). In both PE- and phOx-selected sequences we find an accumulation of mutations around CDR1 as well as some common minor hot spots (e.g., Ser-26, Ser-31, and Ser-77), but the major hot spots depend on the selecting antigen.

Intrinsic Hot Spots Are Revealed by Mutated Transgenes That Have Not Been Selected by Antigen. To examine the response to a third antigen, LK3 mice were immunized with NP. However, we found that although the anti-NP hybridomas expressed the transgene, the transgenic κ chain did not contribute to the anti-NP antibody. All 21 hybridomas analyzed expressed the transgene together with either an endogenous λ chain (14 examples) or an endogenous κ chain (7 examples). Serological analysis of both categories of hybrid-

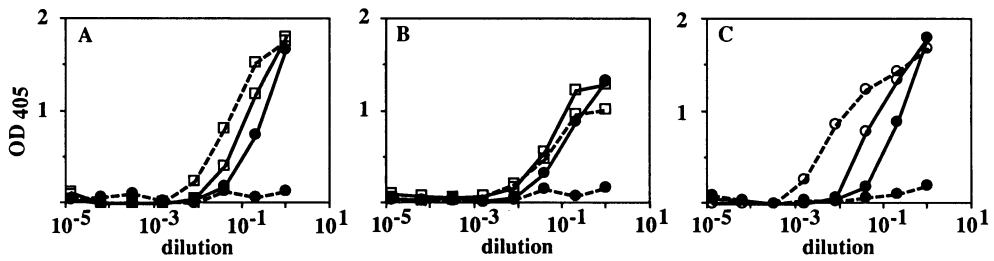

FIG. 2. Demonstration by ELISA that the transgenic κ chain does not contribute to NP-binding activity of the secreted antibody of the anti-NP hybrids. Hybridoma NNT2A5.12.1 (*A*) and hybridoma NNT2A4.24.1 (*B*), both of which express the transgene and mouse λ chains, and hybridoma NNT2A4.18.2 (*C*), which expresses the transgene and mouse κ chains are shown. Dilutions of intact antibody (solid lines) or purified Fab fragment (dashed lines) were incubated with NP/bovine serum albumin-coated plastic plates and the bound light chain was detected with biotinylated anti-light chain antiserum followed by horseradish-coupled streptavidin [□, mouse λ chains; ○, mouse κ chains; ●, rat κ chains (which detects the transgene)]. Results (plotted as absorbance at 405 nm against dilution) from three representative hybridomas are illustrated. Differential ELISA of all 21 transgene-expressing anti-NP hybridomas screened suggested that the transgenic light chain did not contribute to NP specificity; this result was confirmed for six of them with Fab fragments.

Table 1. Nature of base substitutions

	Substitutions in passenger sequences						Substitutions in phOx-selected sequences				
	T	C	A	G	Examples		T	C	A	G	Examples
T	—	0.72	0.14	0.14	7	T	—	0.31	0.24	0.45	42
C	0.8	—	0.2	0	15	C	0.48	—	0.36	0.16	42
A	0.25	0.29	—	0.46	24	A	0.43	0.18	—	0.39	169
G	0.08	0.24	0.68	—	25	G	0.09	0.39	0.52	—	147

Distribution of base substitutions is given as the proportion of the total for each base in the V_κOx1 coding strand with the wild-type base indicated in the vertical and the mutant base indicated in the horizontal. Sequences from which the data are derived are as specified in the legend to Fig. 1.

oma revealed that it was the endogenously encoded light chain that conferred the hapten specificity (Fig. 2). Thus, these NP-specific hybridomas derive from B cells that have escaped allelic exclusion. The transgene behaves as a passenger that can accumulate mutations without affecting the expressed antibody (5). The mutations in these passengers are not biased by selection for improved antigen binding and therefore give information about the intrinsic specificity of the hypermutation process. The sequences (Fig. 1*C*) reveal a dominant hot spot in the second base of the Ser-31 codon and several further less dominant hot spots, as well as preferential accumulation of mutations around CDR1. These features are also manifested in antigen-selected sequences [phOx-specific V_κOx sequences from nontransgenic mice; PE-specific transgene sequences (Fig. 1 *A* and *B*)] and can now be ascribed to the hypermutation itself rather than to selection. The V_κOx Ser-31 hot spot is at the tip of a potential stem–loop structure TGTAA>G<TTACA. Similar structures have been found in other V region genes (9) and could act as entry sites for a repair mechanism favoring local mutation (10).

Intrinsic Base Substitution Preferences of the Hypermutation Mechanism. The mutations in the passenger transgenes are all base substitutions demonstrating that the absence of nucleotide deletions/insertions in hypermutated and antigen-selected antibodies is a consequence of the mutation mechanism and not a result of selection. Passengers also show a marked bias in base substitutions (Table 1) that particularly favors transitions (Table 2). Such a bias has previously been noted in antigen-selected sequences (9, 11–14) and mirrors that reported (15) for substitutions occurring during pseudogene evolution. Regarding transversions, the most striking bias is associated with the third base of His-34 where, responding to selection for mutation to Gln (improving phOx binding), it always (14 examples) uses the C → G rather than the C → A transversion. This does not reflect a general preference among transversions. It is at present unclear whether this result reflects an anomalous base substitution preference or whether it is caused by selection, for example, at the level of differential glutamine codon utilization.

Independent Identification of Intrinsic Hot Spots by Analysis of Base Substitution Preferences. We therefore propose a strategy for identifying intrinsic hot spots within the data base of phOx-selected sequences by comparing the pattern of nucleotide substitutions found at individual positions with the pattern predicted from the intrinsic nucleotide substitution preferences as deduced from the passenger transgenes (Table 3). While extreme disagreements at mutations in His-34 and Tyr-36 confirm that these changes are selected, agreements in Ser-31 (second base) and Ser-77 likely reflect intrinsic hot spots. These latter two positions are also hot spots in all the compilations in Fig. 1. The first and third base of Ser-31 behave quite differently from the second base, showing a strong bias for mutation to arginine, improving phOx binding (16). Thus, among phOx-selected sequences, Ser-31 is highly mutated for two reasons: (*i*) to arginine because it is selected by antigen, (*ii*) to asparagine, threonine, and, less commonly, isoleucine because the second base is an intrinsic hot spot.

Table 2. Percentage transitions

	Found	Expected*
Passengers	64	33
phOx selected		
Total	44	33
Excluding His-34 and Tyr-36 hot spots	55	33
Silent mutations	79	48†

*Expected for random mutation.
†Corrected for codon composition of the V_κOx1 gene.

The Mutational Preferences of the Hypermutation Mechanism Suggest Polarity. There is a strong intrinsic bias against mutations occurring in T. This is apparent not only in the passenger sequences (Table 1) but also from the phOx-selected sequences (despite the selected Ser-31 → Arg). In addition, a bias against C is revealed by the passengers but only becomes apparent in the phOx-selected sequences after the overwhelming effect of the selected His-34 mutations (58 of the 142 sequences) is removed. These biases do not reflect nonrandom base composition of V_κOx1. Similar biases have already been noted among antigen-selected mutations in the heavy-chain V regions of antibodies to *p*-azophenylarsonate (14), and we have also noted them among published mutations in other antigen-selected V_H sequences [10% T and 13% C of the total of 227 mutations (17) for V_H186-2; 7% T and 11% C of the total of 71 mutations (1) in V_κOx1]. Our evidence that the bias against pyrimidines is intrinsic and cannot be attributed to antigenic selection shows that the hypermutation is targeted to one strand. This has been previously proposed from analyses of mutations in antigen-selected sequences (14, 18). Strand polarity together with a likely association of hypermutation with transcription (19) is reminiscent of a form of DNA repair described in mammalian cells (20, 21) and is consistent with models in which hypermutation is achieved by error-prone repair following specific endonucleolytic cleavage (22).

CONCLUSION

By comparing the pattern of somatic hypermutation of a transgene that has either been selected by different antigens or that was expressed as a passenger, we have deduced certain intrinsic characteristics of the mutational mechanism. It exhibits specific hot spots and only yields base substitutions; these substitutions show marked preferences with a bias to transitions and indicate a polarity to the hypermutation mechanism.

The intrinsic specificity of the hypermutation process leads to preferential clustering of mutations around CDR1 with a major hot spot in the second base of Ser-31. This conclusion was obtained by using a V_κOx1 transgene but analysis (unpublished observations) of the nucleotide substitution preferences of mutations reported in two V_H genes (1, 17) suggests that they also harbor intrinsic hot spots (i.e., not due to selection). Thus, the concentration of mutations around CDR1 commonly observed during antibody affinity matura-

Table 3. Analysis of nucleotide substitutions at individual hot spots

Selected hot spots				Unselected hot spots			
Position	Substitution	Found	Predicted	Position	Substitution	Found	Predicted
phOx-selected sequences				Passenger sequences			
Ser-31,I (A)	T Cys	0	3	Ser-31,II (G)	T Ile	1	0.5
(11 examples)	C Arg	9	3	(7 examples)	C Thr	1	1.5
	G Gly	2	5		A Asn	5	5
Ser-31,III (T)	C —	0	11	phOx-selected sequences			
(15 examples)	G Arg	12	2	Ser-26,III (C)	T —	9	7
	A Arg	3	2	(10 examples)	G Arg	0	1.5
His-34,I (C)	T Tyr	0	36		A Arg	1	1.5
(45 examples)	G Asp	0	0	Ser-31,II (G)	T Ile	1	1
	A Asn	45	9	(17 examples)	C Thr	7	4
His-34,III (C)	T —	1*	12		A Asn	9	12
(15 examples)	G Gln	14	0	Ser-77,II (G)	T Ile	1	1
	A Gln	0	3	(15 examples)	C Thr	7	4
Tyr-36,II (A)	T Phe	39	10		A Asn	7	10
(40 examples)	C His	0	12				
	G Cys	1	18				

Position is identified as the base position in the mutated codon; thus Ser-31,I (A) is the first base in the Ser-31 codon of the V_κOx1 gene and is an adenine in the unmutated gene. Predicted number of nucleotide substitutions was calculated based on intrinsic nucleotide substitution preference deduced from passenger transgenes (see Table 1).

*The one example of this third base C-to-T transition at His-34 among phOx-selected sequences is accompanied by a C-to-A transversion at the first position of this codon, resulting in a change to Asn.

tion (3, 23) is probably not due to antigen selection alone but also to the concentration of favored mutation sites resulting from germ-line evolutionary selection. It is interesting to contrast this CDR1 clustering with the role of V–(D)–J joining (D, diversity) in generating CDR3 diversity and the fact that CDR2 appears to be the region of greatest diversity among germ-line V region segments (24, 25). Study of hypermutation in modified transgenes should allow one to ascertain whether the CDR1 clustering of hypermutation and the intrinsic hot spots reflect structures necessary for recruiting the relevant enzymes.

We thank Melanie Sharpe for advice, David Gilmore for cell sorting, Roger Staden for advice on data handling, and Andrea Löhndorf for help with sequencing. A.G.B. and C.R. were supported by fellowships from Boehringer Ingelheim Fonds and the Spanish Ministry of Science and Education, respectively.

1. Griffiths, G. M., Berek, C., Kaartinen, M. & Milstein, C. (1984) *Nature (London)* **312,** 271–275.
2. Möller, G., ed. (1987) *Immunological Reviews* (Munskgaard, Copenhagen), Vol. 96.
3. Berek, C., Jarvis, J. M. & Milstein, C. (1987) *Eur. J. Immunol.* **17,** 1121–1129.
4. Allen, D., Simon, T., Sablitzky, F., Rajewsky, K. & Cumano, A. (1988) *EMBO J.* **7,** 1995–2001.
5. Sharpe, M. J., Milstein, C., Jarvis, J. M. & Neuberger, M. S. (1991) *EMBO J.* **10,** 2139–2145.
6. Schittek, B. & Rajewsky, K. (1990) *Nature (London)* **346,** 749–751.
7. Berek, C. & Milstein, C. (1987) *Immunol. Rev.* **96,** 23–41.
8. Rada, C., Gupta, S. K., Gherardi, E. & Milstein, C. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 5508–5512.
9. Levy, S., Mendel, E., Kon, S., Avnur, Z. & Levy, R. (1988) *J. Exp. Med.* **168,** 475–489.
10. Bhattacharyya, A., Murchie, A. I. H., von Kitzing, E., Diekmann, S., Kemper, B. & Lilley, D. M. J. (1991) *J. Mol. Biol.* **221,** 1191–1207.
11. Golding, G. B., Gearhart, P. J. & Glickman, B. W. (1987) *Genetics* **115,** 169–176.
12. Both, G. W., Taylor, L., Pollard, J. W. & Steele, E. J. (1990) *Mol. Cell. Biol.* **10,** 5187–5196.
13. Rogerson, B., Hackett, J., Peters, A., Haasch, D. & Storb, U. (1991) *EMBO J.* **10,** 4331–4341.
14. Manser, T. (1990) in *Somatic Hypermutation in V-Regions*, ed. Steele, E. J. (CRC, Boca Raton, FL).
15. Li, W., Wu, C. & Luo, C. (1984) *J. Mol. Evol.* **21,** 58–71.
16. Dreher, M. L., Gherardi, E., Skerra, A. & Milstein, C. (1991) *J. Immunol. Methods* **139,** 197–205.
17. Weiss, U., Zoebelein, R. & Rajewsky, K. (1992) *Eur. J. Immunol.* **22,** 511–517.
18. Lebecque, S. G. & Gearhart, P. J. (1990) *J. Exp. Med.* **172,** 1717–1727.
19. Roes, J., Hüppi, K., Rajewsky, K. & Sablitzky, F. (1989) *J. Immunol.* **142,** 1022–1026.
20. Mullender, L. H., Vireling, H. & Venema, J. (1991) *Mutat. Res.* **250,** 223–228.
21. Hanawalt, P. C. (1991) *Mutat. Res.* **247,** 203–211.
22. Brenner, S. & Milstein, C. (1966) *Nature (London)* **211,** 242–243.
23. Bahler, D. W. & Levy, R. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 6770–6774.
24. Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. & Winter, G. (1992) *J. Mol. Biol.* **227,** 776–798.
25. Milstein, C., Even, J., Jarvis, J. M., Gonzalez-Fernandez, A. & Gherardi, E. (1992) *Eur. J. Immunol.* **22,** 1627–1634.