FEBS Letters 452 (1999) 386–394 FEBS 22108

Efficient amplification and direct sequencing of mouse variable regions from any immunoglobulin gene family

Thierry Chardès[b,]*, Sylvie Villard[a], Gaelle Ferrières[a], Martine Piechaczyk[c], Martine Cerutti[b], Gérard Devauchelle[b], Bernard Pau[a]

[a]*CNRS UMR 9921, Faculté de Pharmacie, 15 Avenue Charles Flahault, 34060 Montpellier, France*
[b]*INRA-CNRS URA 2209, Laboratoire de Pathologie Comparée, 30380 Saint-Christol-Lez-Alès, France*
[c]*CNRS UPR 1086, CRBM-Laboratoire Infections Rétrovirales et Signalisation Cellulaire, 34060 Montpellier, France*

Received 11 March 1999; received in revised form 29 April 1999

Abstract **We have designed two original sets of oligonucleotide primers hybridizing the relatively conserved motifs within the immunoglobulin signal sequences of each of the 15 heavy chain and 18 kappa light chain gene families. Comparison of these 5′ primers with the immunoglobulin signal sequences referenced in the Kabat database suggests that these oligonucleotide primers should hybridize with 89.4% of the 428 mouse heavy chain signal sequences and with 91.8% of the 320 kappa light chain signal sequences with no mismatch. Following PCR amplification using the designed primers and direct sequencing of the amplified products, we obtained full-length variable sequences belonging to major (VH1, VH2, VH3, Vκ1 and Vκ21) but also small-sized (VH9, VH14, Vκ2, Vκ9A/9B, Vκ12/13, Vκ23 and Vκ33/34) gene families, from nine murine monoclonal antibodies. This strategy could be a powerful tool for antibody sequence assessment whatever the V gene family before humanization of mouse monoclonal antibody or identification of paratope-derived peptides.**

Key words: Immunoglobulin; Signal sequence; V gene; Polymerase chain reaction; Oligonucleotide primer

1. Introduction

Antibody molecules are unique because of their wide range of binding specificities coupled with many different effector functions. Hybridoma cells, mainly of rodent origin, are a pure source of monoclonal antibody (mAb) of desired specificity. The clinical potential of these mAbs is hindered because therapeutic mouse antibodies can elicit an immune response in humans. To overcome this problem, mouse/human chimeric antibodies have been constructed by replacement of all but the murine variable regions [1] or by grafting murine complementarity-determining regions (CDR) with human framework regions (FR) [2]. Furthermore, bioactive paratope-derived peptides (PDP) designed from mouse mAbs are also a powerful tool to obtain minimal antibody-like structure [3,4] showing the lowest human anti-mouse antibody response.

Such humanization of a mouse mAb or identification of PDPs depends on efficient amplification and sequencing of murine variable regions from the mAb under investigation. The polymerase chain reaction (PCR) is generally used with 5′ primers which take advantage of conserved features of framework [5–8] or signal (S) sequences [9–12] within antibody variable regions, in conjunction with a 3′ oligonucleotide probe directed to the constant region. A one-side PCR using a 3′ primer directed to the constant region has also been described as being independent of the sequence of the variable segment [13]. A strategy employing FR primers can lead to the substitution of amino acids in the framework affecting antigen binding [14,15]. Of particular importance is the residue at position 2 in FR1 of the light chain variable region, which indirectly affects the structure of CDR1 [16]. FR sequences can undergo somatic maturation leading to inadequate matching [17], whereas the frequency of possible point mutations in the signal sequence is lower. Priming in the conserved motifs from this immunoglobulin (Ig) signal region is thus a possible powerful way to obtain entire sequences of antibody variable regions. Various sets of signal sequence primers have been described [9–12], but amplification of certain variable regions with these probes systematically fails, mainly due to inadequate matching [12,18–20]. An extensive analysis of the oligonucleotide sequences used [9–12] indicates that they efficiently match Ig signal regions from the major VH gene families, i.e. VH1 (J558), VH2 (Q52), VH3 (36–60) or VH5 (7183), and from the major Vκ gene families, i.e. Vκ1, Vκ4/5, Vκ19/28, Vκ21 or Vκ24/25, but they do not prime signal sequences from minor Ig gene families.

Murine V genes have been classified into 15 VH and 18 Vκ gene families, based upon amino acid and/or nucleotide sequence similarities [21–25]. In an attempt to potentially amplify Ig genes from all V gene families, we have defined two original sets of leader primers which hybridize in the relatively conserved signal sequences of each heavy and light chain gene family. These primers have been routinely used in our laboratory to amplify and directly sequence the full-length variable regions from nine murine mAbs, including domains belonging to seven different Vκ and five different VH gene families. Our strategy allows rapid and accurate sequencing of variable regions from any Ig gene family and should facilitate the design of PDPs or chimeric antibodies of clinical interest.

2. Materials and methods

2.1. Mouse hybridoma cells

The following cell lines were used in the experiments: 27G7 (anti-phosphonamidic hapten: IgG2b/κ), 2C2 (anti-digoxin: IgG1/κ [26]), 1C10 (anti-digoxin: IgG1/κ [27]), Tg6 (anti-human thyroglobulin:

---

*Corresponding author. Fax: (33) 4-67-54-86-10.
E-mail: chardes@pharma.univ-montp1.fr

Abbreviations: mAb, monoclonal antibody; CDR, complementarity-determining region; PDP, paratope-derived peptide; PCR, polymerase chain reaction; FR, framework region; S, signal sequence; Ig, immunoglobulin; VH, variable region of the heavy chain; Vκ, variable region of the kappa light chain

PII: S0014-5793(99)00649-3

IgG1/κ [28]), 8G7B3 (anti-parvalbumin: IgG2a/κ), 1F3 (anti-human antithrombin III: IgG1/κ), 11E12 (anti-human cardiac troponin I: IgG2a/κ [29]), ST40/F142.63 (anti-human CD4: IgG1/κ [30]) and 13B8.2 (anti-human CD4: IgG1/κ [31]). All the mAbs were produced in our laboratory except mAb 13B8.2, which was kindly donated by D. Olive and C. Mawas (INSERM U119, Marseille, France), and mAbs ST40/F142.63 and 11E12, which were provided by D. Carrière and C. Larue, respectively (Sanofi Recherche, Montpellier, France). Hybridoma cells were obtained by fusing splenocytes with P3-X63Ag8.653 myeloma cells except mAb 13B8.2, which was produced by cell fusion of splenocytes with NS1 myeloma cells. Cells were grown in RPMI 1640 medium with 10% fetal calf serum supplemented with 2 mM glutamine, 100 μg/ml streptomycin and 100 IU/ml penicillin.

2.2. Primer design

Two databases of 491 mouse heavy chain and 356 mouse light chain signal sequences, expanded from the Kabat database ([32], http://immuno.bme.nwu.edu/), were used to design the PCR primers. Each signal sequence was assigned to a Vκ or a VH gene family [23,25], according to the related reference listed in the Kabat database. The signal sequences belonging to the same Ig gene family were aligned and the most conserved region was chosen to define consensus signal sequence primers as depicted on Tables 1 and 2. The antisense primer RevCκ*Sal*I (5′-CGACTAGTCGACTGGTGGGAAGATGGATACAG-3′) contains the reverse complement of codons 113–120 of kappa light chain constant region. The following antisense oligonucleotides were designed to match the heavy chain constant regions: RevCγ_1*Sal*I (5′-CGACAAGTCGACTAGCCCTTGACCAGGCATCC-3′; matching codons 141–147 of Cγ_1), RevCγ_{2a}*Sal*I (5′-CGACAAGTCGACTAACCCTTGACCAGGCATCC-3′; matching codons 141–147 of Cγ_{2a}), RevCγ_{2b}*Sal*I (5′-CGACTAGTCGACCAGGGATCCAGAGTTCCAAG-3′; matching codons 156–166 of Cγ_{2b}) and RevCγ_3*Sal*I (5′-CGACTAGTCGACTAGCCTTTGACAAGGCATCC-3′; matching codons 141–147 of Cγ_3). These primers contained a *Sal*I restriction site (underlined) for further cloning. The 5′-CG-3′ dinucleotide found in this restriction site is extremely rare in mammalian genomes [33] and the frequency of *Sal*I recognition sites in variable domains of mouse antibodies is approximately 0.1–0.2% [34].

2.3. cDNA synthesis and PCR amplification of VH and Vκ genes from mouse hybridomas

Total RNA was extracted from 2×10^8 hybridoma cells using the TRIzol technique (Gibco BRL, Grand Island, NY, USA) as described by the manufacturer. 50 μg of total RNA was annealed at 65°C for 5 min with 10 pM of RevCκ*Sal*I primer for light chain cDNA synthesis or appropriate RevCγ*Sal*I primer for heavy chain cDNA synthesis in a 50 μl final volume of 1×first-strand cDNA synthesis buffer (Gibco BRL) containing 10 μM dNTPs (Gibco BRL) and 1 mM DTT (Gibco BRL). After annealing, the samples were cooled for 20 min at 20°C. 400 U of reverse transcriptase Superscript II (Gibco BRL) and 40 U of RNase inhibitor (Promega, Madison, WI, USA) were then added and the reaction was incubated for 1 h at 42°C. Reverse transcriptase activity was inactivated by incubation at 95°C during 5 min and cDNA was kept at −20°C until use. PCR amplification was carried out in a final volume of 20 μl containing 1 μl of cDNA synthesis reaction, 10 μM dNTPs, 2 μl of 10×PCR buffer (New England Biolabs, Beverly, MA, USA). Fourteen VH and 18 Vκ PCR reactions were set up using each group of family-specific 5′ primers (Tables 1 and 2) and the appropriate 3′-oligonucleotide probe matching the light or heavy chain constant region. Two mixtures, containing all VH or Vκ 5′ primers, were also set up. No PCR reaction was performed with primer specific for the VH13 (3609N) gene family since the only reported member assigned to this family is the non-functional allele of PC3609 [35]. The amount of each primer used was initially 10 pM for non-degenerated oligonucleotide and was increased two- to eight-fold depending on the degeneracy. The reaction mixtures were heated to 94°C for 10 min, then 2 U Vent DNA polymerase (New England Biolabs) was added and 30 cycles of amplification were carried out for 1 min at 94°C, 1 min at 55°C and 2 min at 72°C. After a 10 min extension at 72°C, the PCR products were fractionated through a 1.5% agarose gel (Gibco) and stained with ethidium bromide. Sets of 5′ primers and 3′ primers leading to a 390 bp product for Vκ amplification and a 450 bp product for VH amplification were selected from this family-specific PCR screening. Five replicates using the same selected primers were subjected to a new PCR as described above. The PCR-amplified DNA products were gel purified on a 1.5%, low melting temperature agarose gel (Gibco).

2.4. Direct nucleotide sequencing of the amplified V genes

The gel-purified DNA products were directly sequenced on both strands using the 3′ and 5′ primers selected above and the dideoxy termination kit Thermo Sequenase (Amersham Pharmacia Biotech, Cleveland, OH, USA), as described by the manufacturer except for minor modifications. Briefly, the reaction mixture containing 100 ng DNA amplification product, Thermo Sequenase DNA polymerase, appropriate primer, [α-^{35}S]dATP (ICN Pharmaceuticals, Costa Mesa, CA, USA), dCTP, dGTP and dTTP was subjected to 50 labeling cycles of 15 s at 95°C followed by 30 s at 50°C. Then 3.5 μl of the labeling reaction was transferred onto each tube containing 4 μl of ddATP or ddCTP or ddGTP or ddTTP termination mix. The termination step corresponded to 30 cycles of PCR at 95°C for 15 s, 55°C for 30 s and 72°C for 1 min. The reaction was ended by addition of 4 μl of stop solution. After electrophoresis, gel drying and film exposure, 250–300 bases of V gene sequence were obtained for each PCR product. Sequence comparison and germline gene analysis of VH and Vκ regions used by the nine mAbs were performed using the Kabat database [32].

3. Results

3.1. Evaluation of the 5′ primers designed from heavy chain signal sequences

Among the 491 VH signal sequences expanded from the Kabat database under the item ‘mouse AND all.ig.heavy.-chain.signal.sequence’, 64 sequences were eliminated by scanning by eye because (1) they are assigned to another species or (2) they show an incomplete nucleotide sequence (less than 12 bases annotated) or (3) only the amino acid sequence is available. The remaining 427 sequences were assigned to a VH gene family as documented in the related references listed in the Kabat database or in other related references. Furthermore, one VH12-specific signal sequence derived from GenBank was added to our compilation since it was the only sequence reported in both of these databases as corresponding to a signal region from a VH12 hybridoma [36]. In this way, all the selected signal sequences were classified among the 15 heavy chain gene families, except for the VH13 (3609N) family showing only one member reported as a non-functional rearrangement [35]. These subgroups were divided according to amino acid [21] and nucleotide [25,35,37–41] sequence similarity. Alignment of all signal sequences inside each VH gene family allowed the design of 32 non-degenerated or degenerated oligonucleotides (Table 1). These heavy chain S primers (SH) spanned positions −13 and −1 of the signal region.

Primer matching was evaluated by comparison of each designed sequence with the 428 mouse heavy chain signal sequences. As shown in Table 1, 15 VH1 S primers were predicted to hybridize 200 signal sequences assigned to the major VH1 gene family without any mismatch. Groups of two SH primers were designed to potentially hybridize 34 signal sequences from the VH2 family, 28 from VH3, 22 from VH5 and 30 from VH7 with no mismatch. Other families were probed with only one SH oligonucleotide, leading to 69 additional sequences potentially amplified without any mismatch. Taken together, our 32 selected 5′ primers should be able to probe 379 heavy chain signal sequences without any mismatch, indicating that 89.4% of the known murine heavy chain signal sequences from the Kabat database exactly match with the set of 5′-SH oligonucleotides we defined. Finally, the facts that (i) PCR tolerates primer/template DNA mismatches [42] and

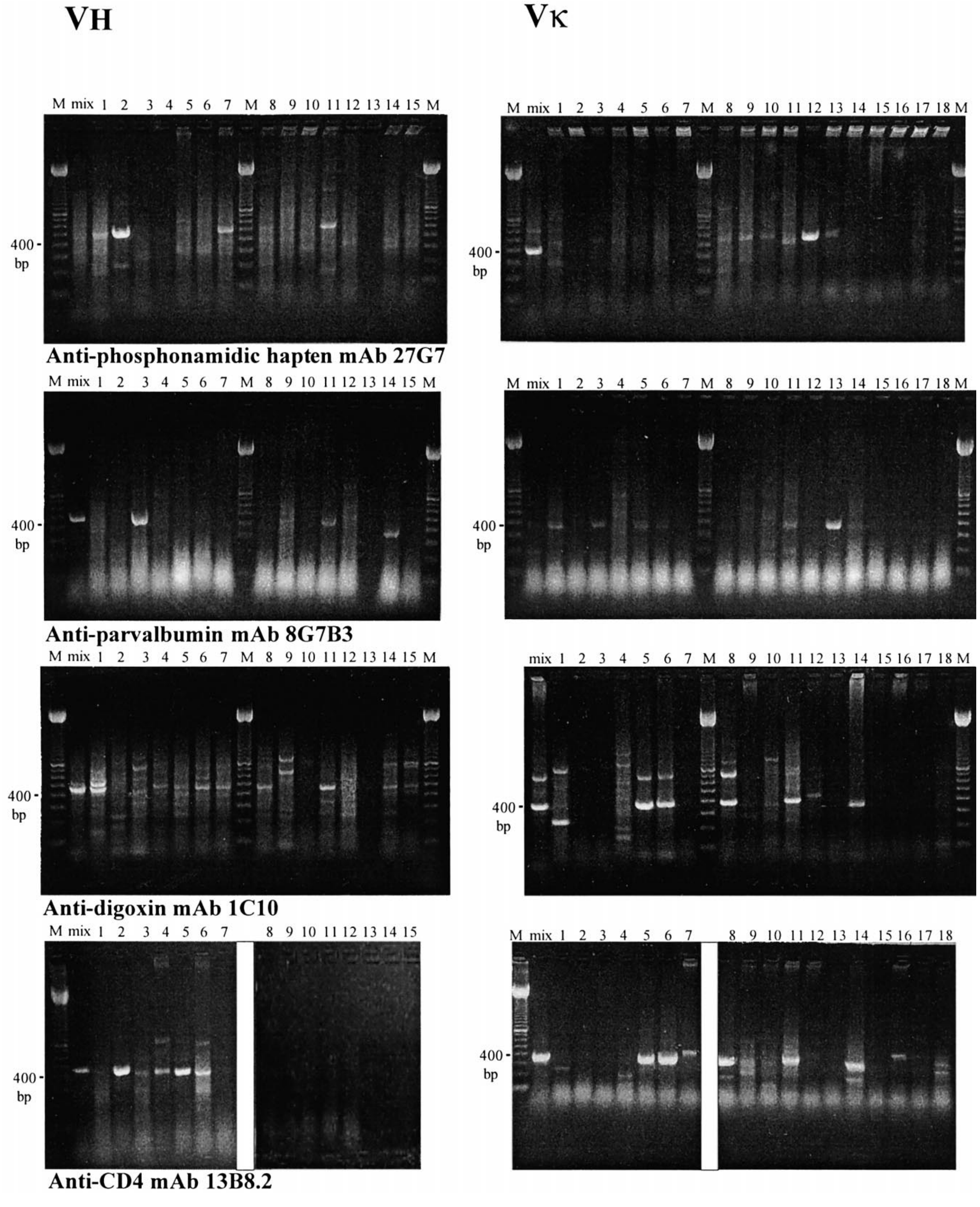

Fig. 1. Immunoglobulin heavy (VH) and kappa light (Vκ) chain PCR products derived from first-strand cDNA of four mouse mAbs. The amplification was performed as described in Section 2 with appropriate constant 3′ primer and sets of signal 5′ primers specific for a given VH (lanes 1–15) or Vκ (lanes 1–18) gene family. The 20 μl PCR reaction was fractionated on a 1.5% agarose gel and the amplification products were stained by ethidium bromide. (VH) The VH1–VH15 family-specific sets of primers were used in lanes 1–15. (Vκ) Lane 1 shows the results of amplification using the Vκ1 set of primers; lane 2, Vκ2 set of primers; lane 3, Vκ4/5; lane 4, Vκ8; lane 5, Vκ9A/9B; lane 6, Vκ10; lane 7, Vκ11; lane 8, Vκ12/13; lane 9, Vκ19/28; lane 10, Vκ20; lane 11, Vκ21; lane 12, Vκ22; lane 13, Vκ23; lane 14, Vκ24/25; lane 15, Vκ32; lane 16, Vκ33/34; lane 17, Vκ31/38C; lane 18, VκRF. In the mixture (mix), PCR products were obtained following amplification of cDNA with a 3′ primer combined with all 5′ signal oligonucleotides. M indicates the 100 bp ladder markers with the 400 bp marker annotated on the left. Note that an inversion occurred for the Vκ amplification of mAb 27G7: lane 11 corresponds to amplification with the Vκ22 primer, whereas lane 12 indicates amplification with the Vκ21 set of primers.

Table 1
Gene family-specific PCR signal primers used to screen hybridoma cDNA for identification of the heavy chain variable region

VH gene family	Kabat subgroup	Selected primers according to the numbering of the Ig signal sequence												SH sequences predicted from the Kabat database				
														Number of mismatches				
		−13	−12	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0	1	2	3	4
VH1 (J558)	IIA, IIB, VA, Misc.							ACT	GCA	GGT	RTC	CAC	TCC	18	2			
								ATA	GCA	GGT	GTC	CAC	TCC	3			1	
								RCT	ACA	GGT	GTC	CAC	TCC	50				
								GCY	ACA	GMT	GTC	CAC	TCC	19	3			
								ACT	GCA	GGT	GTC	CWM	TCC	2				
								RCT	RCA	GGT	GTK	CAC	TCC	3	1			
								GCT	AWM	GGT	GTC	CAC	TCC	2				
								CC	TCA	GGT	GTC	CAC	TCC	2				
								GCT	ACA	GGT	GCT	CAC	TCC	1				
								ACT	GCA	GGT	GTC	CTC	TCT	13	1			
								AYT	GCA	GGT	GTC	CAY	TGC	7	1			
								GCT	AMM	GGT	GTC	CAC	TTC	3				
				CTC	CTG	TCA	KTA	ACT	KCA	GGT				4	1			1
								AAC	TGC	AGG	TGT	CTC	TCT	4				
								RCT	RCA	GGY	GTC	CAC	TCT	69	3			
VH2 (Q52)	IB							CCA	AGC	TGT	ATC	CTT	TCC	2				
								CCA	AGC	TGT	GTC	CTR	TCC	32				
VH3 (36–60)	IA			TG	TTG	ACA	GYC	VTT	CCK	GGT				19		1		
				TG	TTC	ACA	GCC	TTT	CCT	GGT				9				
VH4 (X-24)	IIIB						T	TTA	AAA	GGG	GTC	CAG	TGT	4				
VH5 (7183)	IIID, Misc.					T	AYT	TTA	AAA	RGT	GTC	MAG	TGT	20	8	1		
							GTT	TTA	AAA	GGT	GTC	CTG	TG	2	3	1		
VH6 (J606)	IIIC						CTY	TTA	AAA	GGK	GTC	CAG	WG	10	11			
VH7 (S107)	IIIA						CYT	TTA	MAT	GGT	ATC	CAG	TGT	29	2			
							CTT	TTA	CAT	GGT	TTC	AAG	TGT	1				
VH8 (3609)	IB					YT	GTC	CCT	GCA	TAT	GTC	YT		3				
VH9 (VGam 3–8)	IIA				ATG	GCA	GCW	GCY	CAA	AG				15				
VH10 (MRL-DNA4)	IIID						TT	TAT	CAA	GGT	GTG	CAT	TGT	3	1			
VH11 (CP3)	Misc.						CTT	TTA	AAA	GWT	GTC	CAG	KGT	10				
VH12 (CH-27)[a]	IA				GTG	ACA	GTC	CTT	CCT	GGT	AG			1				
VH13 (3609N)[b]	IIIC																	
VH14 (SM7)	IIB, IIC	C	TTC	CTG	ATG	GCA	GTG	GTT						21		2		
VH15	VA						A	GCT	ACA	GGT	ATC	CAA	TCC	2				

The number of potentially matching SH sequences was determined by comparison of the selected signal primers with the 428 mouse heavy chain signal sequences reported in the Kabat database. All oligonucleotides are listed with their 5′ end to the left. Standard abbreviations are used for mixed sites: R = A or G, Y = T or C, K = T or G, W = A or T, M = C or A, V = A or C or G.

[a] The signal sequence from this family was derived from GenBank (accession number M22439).

[B] No primer specific for the VH13 family has been defined since the only reported member to this family is the non-functional allele of PC3609.

Table 2

Gene family-specific PCR signal primers used to screen hybridoma cDNA for identification of the kappa light chain variable region

Vκ gene family	Kabat subgroup	Selected primers according to the numbering of Ig signal sequence															Sκ sequences predicted from the Kabat database			
		−15	−14	−13	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	Number of mismatches			
																	0	1	2	3
Vκ1	II						CTG	WTG	TTC	TGG	ATT	CCT	G				37			
											GGT	CAG	ACA	GTC	AGC	AGT	1			
Vκ2	II							GTG	CTC	TGG	ATT	CGG	GAA				9	2		
Vκ4/5	IV, VI	TC	AGC	TTC	YTG	CTA	ATC	AGT	G								65	7		
											CTA	ATC	AGT	GCT	TCA	GGA	1			
Vκ8	I									TGG	GTA	TCT	GGT	RCS	TGT	G	22			
										AAA	TTT	AAA	AGT	ACC	TGT	GGG	1			
Vκ9A/9B	V									G	TTT	CMA	GGT	RCC	AGA	TGT	14	4		2
								G	CTC	TGG	TTY	CCA	GGT	ATC			11			
Vκ10	V									TGT	TTT	CAA	GGT	RCC	AGA	TGT	26			1
Vκ11	V										GTT	GTA	ATG	TCC	AGA	GGA	2	1		
Vκ12/13	V									G	CTT	ACA	GGT	GCC	AGA	TGT	7	4	1	
										TCA	ATT	GTA	GRT	GCC	AGA	TGT	3			
										TT	ACA	GTA	GGT	GTC	AGA	TGT	1			
										CA	GTC	GTA	GTT	GTC	AGA	TGT	1			
											CCT	CCT	TCT	TGG	CCA	AGA	1			
Vκ19/28	V										TTA	TAT	GGA	GCT	GAT	GGG	3			
											GTG	TCT	GGT	GCT	CAT	GGG	8			
									CTS	TGG	TTG	TCT	GGT	GTT	GA		10			
Vκ20	VII										GTC	TCT	GAT	TCT	AGG	GCA	4			
Vκ21	III						TG	CTK	CKC	TGG	GTT	CCA	G				21			
											TT	GCA	GGT	GTT	GAC	GGA	2			
Vκ22	I										GT	AGG	TGC	CTC	GTG	CAC	1			
											TG	TCT	GGT	GCC	TGT	GCA	2			
Vκ23	V									TGG	AYT	YCA	GCC	TCC	AGA		6	2	1	
Vκ24/25	II										WTC	TCT	RGA	GTC	AGT	GGG	7			
									C	TGG	ATC	CCT	GGA	KCY	ACT		11			
Vκ32	V							TG	TTC	TGC	TTT	TTA	GGT	GTG			4			
Vκ33/34	V									GA	ATC	CCA	GGC	ATG	ATA	TGT	4			1
Vκ31/38C	V									G	CTT	CAT	GGT	GCT	CAG	TGT	6			
VκRF	V									G	ATA	TCA	GGT	GCC	CAG	TGT	3			

The number of potentially matching Sκ sequences was determined by comparison of selected signal primers with 319 mouse kappa light chain signal sequences reported in the Kabat database. All oligonucleotides are listed with their 5′ end to the left. Standard abbreviations are used for mixed sites: R = A or G, Y = T or C, K = T or G, W = A or T, M = C or A, S = C or G.

Table 3
Characteristics of the nine murine mAbs

mAb	Isotype	Target antigen	Heavy chain variable region			Light chain variable region		
			V-D-J rearrangement	VH germline [ref.]	CDR3 sequence	V-J rearrangement	Vκ germline [ref.]	CDR3 sequence
27G7	IgG2b/κ	Phospho-hapten	VH2-DSP2.9-JH3	Ox-2 [57]	TGYYGWFAY	Vκ21-Jκ2	Vκ21G [58]	QQSRKVPYT
2C2	IgG1/κ	Digoxin	VH1-DSP2.2-JH3	J558-197 [59]	SGPYDYDEVY	Vκ9A/9B-Jκ1	MOPC41 [60]	LEYARTPWT
1C10	IgG1/κ	Digoxin	VH1-DSP2.2-JH3	J558-197 [59]	SGPYDYDEVY	Vκ9A/9B-Jκ1	MOPC41 [60]	LQYASSPWT
Tg6	IgG1/κ	Thyroglobulin	VH9-DFL16.2-JH2	V-Gam3.8 [38]	RGGFITTALDT	Vκ1-Jκ4	Vκ1-C [61]	FQGSHIPFT
8G7B3	IgG2a/κ	Parvalbumin	VH3-DSP2.3-JH3	36-60 [62]	DRGYASWFAY	Vκ23-Jκ4	L7 [17]	QQSNSWPTT
1F3	IgG1/κ	Anti-thrombin III	VH3-DSP2.7-JH4	36-60 [62]	ARPTMDY	Vκ33/34-Jκ2	Vk34b [45]	QQYWSTPYT
11E12	IgG2a/κ	Troponin I	VH14-?-JH3	H4A3 [63]	GFAY	Vκ2-Jκ1	Undetermined	LQSTHFPRT
ST40	IgG1/κ	CD4	VH9-DSP2.4-JH2	VGK2 [64]	GGVLWSRRGDFDY	Vκ21-Jκ1	Vκ21G [58]	QQSNEDPWT
13B8.2	IgG1/κ	CD4	VH2-DQ52-JH3	Ox-2[57]	NDPGTGFAY	Vκ12/13-Jκ2	k2 [65]	QHHYGNPPT

(ii) 37 VH signal sequences were probed by the defined oligonucleotides with only one mismatch suggest that the true percentage of VH signal sequences potentially amplified is greater than those estimated without mismatch.

3.2. *Evaluation of the 5′ primers designed from the kappa light chain signal sequences*

Three hundred and fifty-six Vκ signal sequences, expanded from the Kabat database, were used to define the Sκ 5′-oligonucleotides. An approach similar to the one described above for SH primers led to the elimination of 36 unexploitable sequences. One signal sequence could not be clearly assigned to a Vκ gene family. As shown in Table 2, 31 consensus Sκ 5′-oligonucleotides were defined to probe the 319 remaining Ig signal sequences from the 18 described Vκ gene families [22,24,43–45]. The designed oligonucleotide hybridized between positions −15 and −1 of the signal region. Sets of one to five primers are needed to match all the Vκ gene families. Among the major light chain families, each selected group of primers should potentially probe, with no mismatch, 38 sequences from the Vκ1 family, 66 sequences from the Vκ4/5 family, 21 from Vκ19/28, 23 from Vκ21 and 18 from Vκ24/25 (Table 2). A group of 154 additional sequences belonging to the 13 other Vκ gene families was further predicted to be potentially matched by the remaining primers. Taken together, 294 kappa light chain signal sequences, corresponding to 91.8% of the referenced Vκ signal sequences in the Kabat database, could theoretically be primed by the selected Sκ 5′-oligonucleotides with no mismatch. Probably more than this calculated percentage of kappa light chain signal sequences can be potentially amplified since the designed Sκ primers hybridized 20 additional sequences with only one mismatch.

3.3. *PCR amplification of V genes from mouse hybridomas using gene family-specific signal primers*

Fig. 1 shows the PCR amplification products for the VH and Vκ chain regions of mAbs 27G7, 8G7B3, 1C10 and 13B8.2, using a combination of appropriate constant primer and signal primers corresponding to a given VH (lanes 1–15) or Vκ gene family (lanes 1–18). Similar experiments have been successfully performed with five other mAbs (data not shown). Efficient sets of primers should amplify a 450 bp product for VH amplification and a 390 bp product for Vκ amplification. From the cDNA of hybridoma cells secreting the anti-phosphonamidic hapten mAb 27G7 (IgG2b/κ), major bands at the expected size were obtained with the VH2/Rev-$\text{C}\gamma_{2a}$*Sal*I set of primers for heavy chain amplification and with the Vκ21/RevCκ*Sal*I set of primers for Vκ amplification (Fig. 1). In the same manner, VH3- and Vκ23-specific primers in combination with appropriate 3′ primers led to major amplification products corresponding to the expected size for variable regions of anti-parvalbumin mAb 8G7B3 (IgG2a/κ).

A strong PCR amplification of the VH domain of the anti-digoxin mAb 1C10 (IgG1/κ) was obtained using the VH1 set of primers in association with the Rev$\text{C}\gamma_1$*Sal*I oligonucleotide, whereas various sets of Vκ gene family-specific sets of primers (Vκ9A/9B, Vκ10, Vκ12/13) led to a strong 390 bp amplification product corresponding in size to the light chain variable domain. Multiple hybridization was also observed for VH and Vκ amplification of the variable regions from the anti-CD4 mAb 13B8.2. As shown in Fig. 1, a 13B8.2 DNA fragment with the size expected for a VH domain was obtained when each of the different 5′ primer groups (VH2, VH4, VH5 and VH6) was used separately. Similarly, distinct Vκ family-specific oligonucleotides (Vκ9A/9B, Vκ10, Vκ12/13 and Vκ21) led to a 390 bp fragment corresponding in size to a potential light chain domain of 13B8.2 mAb.

Mixtures of all VH or Vκ primers associated with appropriate constant primer-amplified PCR products at the expected size for heavy or light chain variable domain, except for 27G7 VH amplification. DNA from antibodies of various isotypes (IgG1, IgG2a and IgG2b) has been successfully amplified indicating that the isotype specificity does not affect the PCR amplification.

3.4. Gene characterization of variable regions from nine mAbs following direct sequencing of the amplification products

Further analysis of the assembled genes by direct sequencing of the amplification products (Table 3) showed that six of the nine VH genes belong to major VH gene families (VH1, VH2, VH3), whereas the three others are related to minor families (VH9, VH14). Similarly, major (Vκ1, Vκ21) but also small-sized (Vκ2, Vκ9A/9B, Vκ23 and Vκ33/34) gene families are represented among the Vκ genes encoding the nine mAbs. Inside each gene family, the closest germline genes have been identified for all the variable regions sequenced, except for the Vκ domain of the anti-troponin mAb 11E12. DH members from the DSP2, DFL16 and DQ52 families in combination with various JH genes were used for the VDJ rearrangement of heavy chain variable regions. VJ rearrangement of variable kappa light chain regions from the nine mAbs occurred by using various Jκ gene segments. Full-length sequences of the variable domains, as exemplified by the deduced amino acid sequence of the CDR3 region (Table 3), were characterized for the nine mAbs under investigation. The sequences obtained were identical to those previously described for the anti-hapten mAb 27G7 (unpublished result) and for the anti-CD4 mAb ST40 [4] by screening of a cDNA library. Similarly, sequences of the anti-digoxin mAbs 2C2 and 1C10 were identical to those obtained using a PCR method with FR primers [26,27]. The deduced N-terminal amino acid sequence of the anti-troponin mAb 11E12 was also in accordance with the sequence previously obtained following Edman degradation of the heavy and light chain proteins (unpublished data).

To try to explain the multiple hybridization we described above for the DNA amplification of hybridomas 1C10 and 13B8.2, each PCR product obtained with distinct primers was sequenced. DNA amplification using 5′ primers Vκ9A/9B, Vκ10 or Vκ12/13 for 1C10 light chain yielded the same Vκ9A/9B-Jκ1 sequence (Table 3). Only one or two nucleotide differences were noted between certain S oligonucleotides from each group of Vκ primers we defined (Table 2). These strong sequence homologies probably led to cross-matching of these primers with the DNA template. From the 13B8.2 hybridoma cells, the sequencing of the PCR product obtained with VH2 primers yielded a full-length rearranged variable heavy chain region as described in Table 3. In contrast, the sequences resulting from PCR amplification with primers VH5 or VH6 contained a 50 bp deletion at the FR3-CDR3 boundary, which led to a frameshift in the reading frame (data not shown). This abnormal rearrangement corresponds to the P3 myeloma heavy chain pseudogene [46] found in the parental NS1 cells used as the fusion partner. Amplification of light chain cDNA from 13B8.2 hybridoma cells using 5′ primers Vκ9A/9B, Vκ10 or Vκ12/13 yielded an identical sequence completely rearranged as described in Table 3, whereas the Vκ PCR product amplified with Vκ21 primers corresponds to a non-functionally rearranged kappa light chain (data not shown) transcribed in myeloma cell lines, like NS1, derived from the original MOPC21 tumor [47].

4. Discussion

In this study, we designed and demonstrated the usefulness of two original sets of consensus signal primers for the amplification and subsequent direct sequencing of mouse VH and Vκ regions from any V gene family. This strategy can lead to easier and faster humanization of mouse mAb [1] or identification of paratope-derived peptides [3,4]. Although various sets of primers have yet to be described [9–12], 15–35% of V gene amplification fails [48], mainly due to inadequate matching within selected PCR primers. This strong mismatch is probably due to an inappropriate choice of oligonucleotides for amplification of certain V gene families. Particularly, the designed primers [9–12] efficiently probe immunoglobulin signal sequences from major VH or Vκ families showing numerous germline genes but cannot amplify V regions from minor families having a small number of germline genes. For example, major VH gene families, i.e. VH1, VH2 and VH3, represent 40–60% [49], 10–30% and 7%, respectively, of all heavy chain variable regions expressed [50]. Since adult variable region expression is proportional to the number of genes in a given V gene family [51], if one designs primers which only hybridize signal sequences from these major VH families, a large majority of the murine variable heavy chains can potentially be amplified but VH genes from small-sized families would fail to be obtained.

During the last 10 years, many signal sequences from immunoglobulin variable regions have become available leading to an enriched and more accurate database [32]. Since V genes have been classified into 15 VH and 18 Vκ families depending on amino acid and/or nucleotide similarities [23,25], we can assume that grouping signal sequences according to this classification can lead to the design of primers representative of all the gene families. We deliberately chose to design a wide range of PCR primers with no or limited degeneracy and to adapt the primer concentration according to this degeneracy. In this way, a similar efficient concentration for each primer was maintained and thus the chances of amplifying certain variable regions were not diminished. Using our sets of signal primers, we have been able to obtain the complete V region sequences belonging to major (e.g. VH1 or Vκ21) but also minor gene families (e.g. VH14 or Vκ23), from nine mAbs of different isotypes. These antibodies showed different antigenic specificities; notably they recognized a wide range of molecules from hapten to large protein. The deduced full-length amino acid sequence of each mAb is available, and thus any residue alteration that could influence CDR conformation and hence antigen binding was avoided [14–16]. Furthermore, the dissection of V gene amplifications for a given hybridoma by using multiple family-specific sets of signal primers allows the discrimination between the productive VDJ or VJ rearrangement and aberrant transcripts. By direct sequencing of PCR products, we were able to rapidly identify pseudogenes belonging to the fusion partner [46,47], described by us and others as contaminating the PCR amplification results [52–54], and non-productive rearrangement with a reading frameshift, as exemplified by the 11E12 mAb (unpublished data). With regard to our PCR strategy, other successful methods such as anchored PCR [55] or inverted PCR [54] are more time consuming, require extra steps, demand the synthesis of large PCR products, which may be difficult to sequence directly and lead to reduced cloning efficiency.

Our approach makes it possible to efficiently amplify and directly sequence in less than 2 weeks full-length immunoglobulin variable domains whatever the V gene family. These family-specific signal primers should be helpful in studying V gene selection in B-cell development or in various B cell disorders and in characterizing hybridomas for further chimerization. In addition, the amplified V genes could easily be inserted into our baculovirus expression cassettes for the production of either the entire immunoglobulin molecule or only the Fab fragment [56] or for the production of single chain Fv [27].

Acknowledgements: The authors thank Dr. C. Granier for advice and critical comments and Dr. S.L. Salhi for the editorial revision of the manuscript. We also gratefully acknowledge Dr. D. Carrière, Dr. C. Larue, Dr. D. Olive and Dr. C. Mawas for providing hybridoma cells. We thank G. Gadea and F. Pascual for excellent technical assistance. This work was supported by institutional funds from the Centre National de la Recherche Scientifique (CNRS) and grants from ELF and from the Biotechnology Program of the Ministère de l'Enseignement Supérieur et de la Recherche (MESR).

References

[1] Poul, M.A., Cerutti, M., Chaabihi, H., Devauchelle, G., Kaczorek, M. and Lefranc, M.P. (1995) Immunotechnology 1, 189–196.
[2] Jones, P.T., Dear, P.H., Foote, J., Neuberger, M.S. and Winter, G. (1986) Nature 321, 522–525.
[3] Laune, D. et al. (1997) J. Biol. Chem. 272, 30937–30944.
[4] Monnet, C. et al. (1999) J. Biol. Chem. 274, 3789–3796.
[5] Kettleborough, C.A., Saldanha, J., Ansell, K.H. and Bendig, M.M. (1993) Eur. J. Immunol. 23, 206–211.
[6] Orlandi, R., Gussow, D.H., Jones, P.T. and Winter, G. (1989) Proc. Natl. Acad. Sci. USA 86, 3833–3837.
[7] LeBoeuf, R.D., Galin, F.S., Hollinger, S.K., Peiper, S.C. and Blalock, J.E. (1989) Gene 82, 371–377.
[8] Chaudhary, V.K., Batra, J.K., Gallo, M.G., Willingham, M.C., FitzGerald, D.J. and Pastan, I. (1990) Proc. Natl. Acad. Sci. USA 87, 1066–1070.
[9] Coloma, M.J., Larrick, J.W., Ayala, M. and Gavilondo-Cowley, J.V. (1991) Biotechniques 11, 152–154.
[10] Coloma, M.J., Hastings, A., Wims, L.A. and Morrison, S.L. (1992) J. Immunol. Methods 152, 89–104.
[11] Gavilondo-Cowley, J.V., Coloma, M.J., Vazquez, J., Ayala, M., Macias, A., Fry, K.E. and Larrick, J.W. (1990) Hybridoma 9, 407–417.
[12] Jones, S.T. and Bendig, M.M. (1991) Biotechnology (NY) 9, 88–89.
[13] Heinrichs, A., Milstein, C. and Gherardi, E. (1995) J. Immunol. Methods 178, 241–251.
[14] Panka, D.J., Mudgett-Hunter, M., Parks, D.R., Peterson, L.L., Herzenberg, L.A., Haber, E. and Margolies, M.N. (1988) Proc. Natl. Acad. Sci. USA 85, 3080–3084.
[15] Chien, N.C., Roberts, V.A., Giusti, A.M., Scharff, M.D. and Getzoff, E.D. (1989) Proc. Natl. Acad. Sci. USA 86, 5532–5536.
[16] Chothia, C. et al. (1989) Nature 342, 877–883.
[17] Pech, M., Hochtl, J., Schnell, H. and Zachau, H.G. (1981) Nature 291, 668–670.
[18] Owens, R.J. and Young, R.J. (1994) J. Immunol. Methods 168, 149–165.
[19] Ruberti, F., Cattaneo, A. and Bradbury, A. (1994) J. Immunol. Methods 173, 33–39.
[20] Lan, F., Umeda, M. and Inoue, K. (1996) J. Immunol. Methods 195, 27–32.
[21] Dildrop, R. (1984) Immunol. Today 5, 85–89.
[22] Potter, M., Newell, J.B., Rudikoff, S. and Haber, E. (1982) Mol. Immunol. 19, 1619–1630.
[23] Kofler, R., Geley, S., Kofler, H. and Helmberg, A. (1992) Immunol. Rev. 128, 5–21.
[24] Strohal, R., Helmberg, A., Kroemer, G. and Kofler, R. (1989) Immunogenetics 30, 475–493.
[25] Mainville, C.A., Sheehan, K.M., Klaman, L.D., Giorgetti, C.A., Press, J.L. and Brodeur, P.H. (1996) J. Immunol. 156, 1038–1046.
[26] Navarro-Teulon, I., Peraldi-Roux, S., Bernardi, T., Marin, M., Piechaczyk, M., Shire, D., Pau, B. and Biard-Piechaczyk, M. (1995) Immunotechnology 1, 41–52.
[27] Lemeulle, C. et al. (1998) FEBS Lett. 423, 159–166.
[28] Piechaczyk, M., Chardes, T., Cot, M.C., Pau, B. and Bastide, J.M. (1985) Hybridoma 4, 361–367.
[29] Larue, C., Defacque-Lacquement, H., Calzolari, C., Le Nguyen, D. and Pau, B. (1992) Mol. Immunol. 29, 271–278.
[30] Carriere, D. et al. (1994) Clin. Chem. 40, 30–37.
[31] Corbeau, P. et al. (1993) J. Immunol. 150, 290–301.
[32] Johnson, G., Kabat, E. and Wu, T. (1996) in: Weir's Handbook of Experimental Immunology (Herzenberg, L., Weir, W., Herzenberg, L. and Blackwell, C., Eds.), pp. 6.1–6.21, Blackwell Science, Cambridge, MA.
[33] McClelland, M. and Nelson, M. (1987) Gene Amplif. Anal. 5, 257–282.
[34] Ge, L., Knappick, A., Pack, P., Freund, C. and Plücktun, A. (1995) in: Antibody Engineering (Borrebaeck, C., Ed.), pp. 229–267, Oxford University Press, New York.
[35] Tutter, A., Brodeur, P., Shlomchik, M. and Riblet, R. (1991) J. Immunol. 147, 3215–3223.
[36] Hardy, R.R., Carmack, C.E., Shinton, S.A., Riblet, R.J. and Hayakawa, K. (1989) J. Immunol. 142, 3643–3651.
[37] Brodeur, P.H. and Riblet, R. (1984) Eur. J. Immunol. 14, 922–930.
[38] Winter, E., Radbruch, A. and Krawinkel, U. (1985) EMBO J. 4, 2861–2867.
[39] Kofler, R. (1988) J. Immunol. 140, 4031–4034.
[40] Reininger, L., Kaushik, A., Izui, S. and Jaton, J.C. (1988) Eur. J. Immunol. 18, 1521–1526.
[41] Pennell, C.A., Sheehan, K.M., Brodeur, P.H. and Clarke, S.H. (1989) Eur. J. Immunol. 19, 2115–2121.
[42] Larrick, J.W., Coloma, M.J., del Valle, J., Fernandez, M.E., Fry, K.E. and Gavilondo-Cowley, J.V. (1990) Scand. J. Immunol. 32, 121–128.
[43] Shefner, R., Mayer, R., Kaushik, A., D'Eustachio, P., Bona, C. and Diamond, B. (1990) J. Immunol. 145, 1609–1614.
[44] D'Hoostelaere, L.A. and Klinman, D. (1990) J. Immunol. 145, 2706–2712.
[45] Valiante, N.M. and Caton, A.J. (1990) Immunogenetics 32, 345–350.
[46] Thammana, P. (1994) Mol. Immunol. 31, 77–78.
[47] Strohal, R., Kroemer, G., Wick, G. and Kofler, R. (1987) Nucleic Acids Res. 15, 2771.
[48] Linke, B., Bolz, I., Pott, C., Hiddemann, W. and Kneba, M. (1995) Leukemia 9, 2133–2137.
[49] Freitas, A.A., Andrade, L., Lembezat, M.P. and Coutinho, A. (1990) Int. Immunol. 2, 15–23.
[50] Delassus, S., Gey, A., Darche, S., Cumano, A., Roth, C. and Kourilsky, P. (1995) J. Immunol. Methods 184, 219–229.
[51] Dildrop, R., Krawinkel, U., Winter, E. and Rajewsky, K. (1985) Eur. J. Immunol. 15, 1154–1156.
[52] Berdoz, J., Monath, T.P. and Kraehenbuhl, J.P. (1995) PCR Methods Appl. 4, 256–264.
[53] Kleymann, G., Ostermeier, C., Ludwig, B., Skerra, A. and Michel, H. (1995) Biotechnology (NY) 13, 155–160.
[54] Kaluza, B., Betzl, G., Shao, H., Diamantstein, T. and Weidle, U.H. (1992) Gene 122, 321–328.
[55] Loh, E.Y., Elliott, J.F., Cwirla, S., Lanier, L.L. and Davis, M.M. (1989) Science 243, 217–220.
[56] Poul, M.A. et al. (1995) Eur. J. Immunol. 25, 2005–2009.

[57] Kaartinen, M., Griffiths, G.M., Markham, A.F. and Milstein, C. (1983) Nature 304, 320–324.

[58] Alanen, A. and Weiss, S. (1989) Eur. J. Immunol. 19, 1961–1963.

[59] Sheehan, K.M., Mainville, C.A., Willert, S. and Brodeur, P.H. (1993) J. Immunol. 151, 5364–5375.

[60] Seidman, J.G., Max, E.E. and Leder, P. (1979) Nature 280, 370–375.

[61] Kim, H.N.G., Lavigueur, A., Ricard, L., Boivrette, M., Maclean, S., Cloutier, D. and Gibson, D.M. (1989) J. Immunol. 143, 638–648.

[62] Near, R.I., Juszczak, E.C., Huang, S.Y., Sicari, S.A., Margolies, M.N. and Gefter, M.L. (1984) Proc. Natl. Acad. Sci. USA 81, 2167–2171.

[63] Schiff, C., Milili, M., Hue, I., Rudikoff, S. and Fougereau, M. (1986) J. Exp. Med. 163, 573–587.

[64] Press, J.L. and Giorgetti, C.A. (1993) J. Immunol. 151, 1998–2013.

[65] Tillman, D.M., Jou, N.T., Hill, R.J. and Marion, T.N. (1992) J. Exp. Med. 176, 761–779.