The EMBO Journal vol.5 no.12 pp.3227–3234, 1986

Isolation of a murine Ly-6 cDNA reveals a new multigene family

Kenneth P.LeClair, Roger G.E.Palfree[2] Patrick M.Flood[1], Ulrich Hammerling[2] and Alfred Bothwell[1]

Department of Biology, and [1]Department of Pathology and Howard Hughes Medical Institute at Yale University Medical School, 310 Cedar Street, New Haven, CT 06510, and [2]Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10028, USA

Communicated by A.Frischauf

The Ly-6 alloantigens have been shown to play a critical role in T lymphocyte activation. To isolate a Ly-6 cDNA, synthetic oligonucleotides, based on the partial amino acid sequence of purified Ly-6E.1 protein, were used to probe a cDNA library. The synthetic oligonucleotides or the isolated cDNA detected a 1.1-kb RNA species. Sequence analysis of the cDNA clone revealed that the Ly-6E.1 protein consists of a 26-amino acid leader followed by a 108-residue, cysteine-rich, core protein with no N-linked glycosylation sites. Southern blot analysis of genomic DNAs revealed multiple bands indicating a family of related genes. Using recombinant inbred and Ly-6 congenic strains of mice, restriction fragment length polymorphisms were demonstrable, and correlated with the *Ly-6* allotype of the DNA donors. This probe will enable further molecular genetic analysis of the role of *Ly-6*-linked proteins in the process of T lymphocyte activation. Isolation of *Ly-6* genomic clones may promote a further understanding of the complex tissue-specific expression patterns characteristic of *Ly-6*-linked genes.

Key words: Ly-6/lymphocyte activation/multigene family/RFLP

Introduction

Serological, biochemical and functional analysis of the murine alloantigen, Ly-6, has suggested a complex genetic system. The *Ly-6* locus was first shown to control the expression of an alloantigen expressed on peripheral T lymphocytes (McKenzie *et al.*, 1977). *Ly-6*-linked determinants have since been found on many other cell types, including B lymphocytes, thymocytes, bone marrow cells and fibroblasts. Two alleles, *Ly-6^a* and *Ly-6^b*, have been defined controlling the 6.1 and 6.2 specificities, respectively. The development of anti-Ly-6 monoclonal antibodies has shown that multiple Ly-6 specificities exist, each with a unique tissue distribution pattern (reviewed by Kimura *et al.*, 1984; Hogarth *et al.*, 1984; Houlden *et al.*, 1986). The consensus view now appears to be that at least five distinct epitopes are encoded by *Ly-6*-linked genes. The first appears primarily on activated T and B cells and is defined with the monoclonal antibodies to Ly-6A.2 and Ly-6E.1 (Kimura *et al.*, 1984). Despite apparent differences in tissue distribution patterns, these antibodies appear to detect allelic forms of the same protein as determined by immunoprecipitation, Western blotting and competition studies using xenogeneic rat monoclonal antibodies (Takei, 1984; Palfree *et al.*, 1986). A second epitope, defined as Ly-6B.2 (Gm2.2), is found exclusively on bone marrow cells and neutrophils (Takei *et al.*, 1980; Kimura *et al.*, 1984; Hibbs *et al.*, 1984). The Ly6C-2 (Ly-28.2) determinant is expressed on subpopulations of peripheral blood lymphocytes, bone marrow cells and neutrophils (Kimura *et al.*, 1984; Hogarth *et al.*, 1984: Hibbs *et al.*, 1984). Ly6D.2 (Ly-27.2) is found on most peripheral lymphocytes (Kimura *et al.*, 1984; Hogarth *et al.*, 1984) while a fifth epitope, ThB, is found on immature thymocytes, B cells and bone marrow cells (Yotoku *et al.*, 1974).

The complexity of the *Ly-6* system has engendered speculation as to whether the various Ly-6 specificities described are the products of multiple structural genes, or whether one gene exists whose protein product undergoes variable post-translational processing (Auchincloss *et al.*, 1981; Kimura *et al.*, 1984). Tentative evidence for the former comes from the presumed naturally occurring recombination between the genes controlling the Ly6B.2 specificity in the NZB mouse strain (Kimura *et al.*, 1984). If multiple genes exist, genetic backcross analyses have shown that they are extremely tightly linked.

The first demonstration of the intimate involvement of *Ly-6*-linked proteins in the process of lymphocyte activation was a study of the consequence of addition of the anti-Ly-6E.1 monoclonal antibody, HD42, to various *in vitro* cultures (Flood *et al.*, 1985). This showed that the anti-Ly-6E.1 monoclonal blocked both the concanavalin A (ConA)-induced and the interleukin-2 dependent proliferative responses of normal T cells, inhibited allogeneic mixed lymphocyte reactions and blocked the primary *in vitro* plaque-forming cell response to sheep erythrocytes. The fact that the addition of anti-Ly-6E.1 antibody showed no effect on the functions of previously activated cells suggested an important role for the *Ly-6*-linked determinant in the process of T cell activation. Subsequently, other monoclonal antibodies were described which reacted with molecules involved in physiological T cell activation processes (Rock *et al.*, 1986). The T cell activation and associated proteins (TAP and TPAa) recognized by these monoclonals were recently shown also to be controlled by genes linked to the *Ly-6* locus (Reiser *et al.*, 1986). Similar transmembrane signalling effects have been noted with monoclonal antibodies specific for Thy-1 and the mouse and human T cell receptor and the associated human T3 complex proteins (Gunter *et al.*, 1984; Kaye and Janeway, 1984; Meuer *et al.*, 1984). Biochemical characterization of Ly6C.2 (Takei, 1982), Ly-6E.1 (Palfree and Hammerling, 1986) and TAP/TAPa (Rock *et al.*, 1986) have shown them to be small proteins with M_r 15 000–18 000, apparently distinct from presently described mouse T cell receptor-associated proteins and Thy-1 (Campbell *et al.*, 1981; Allison and Lanier, 1985; Samelson *et al.*, 1985a; Oettgen *et al.*, 1986). Thus, the *Ly-6* antigens may represent another set of cell surface molecules involved in the control of signal transduction across the plasma membrane.

One of the most thoroughly characterized *Ly-6*-linked proteins is Ly-6E.1 encoded by the *Ly-6^a* allele and recognized by both the SK70.94 and HD42 monoclonal antibodies (Palfree and Hammerling, 1986). This protein appears as a 17 000 and 18 000 M_r doublet in SDS–polyacrylamide gels, and shows at least five charged species with pI between 4.0 and 5.2, although no N-

```
                    1  2  3  4  5  6  7  8  9  10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
AMINO-TERMINAL
SEQUENCE DATA       ?  GLU GLN TYR GLN cys TYR GLY VAL PRO PHE GLU THR ser ?  PRO ?  ILE THR ?  ?  TYR PRO ASP gly VAL

CODON
POSSIBILITIES    5'    GAR CAR TAY CAR tgy TAY GGX GTX CCX TTY GAR ACX  3'

COMPLEMENTARY
OLIGONUCLEOTIDE  3'    CTC GTC ATG GTC ACG ATA CCG CAC GGT AAA CTC TG   5'
   (35-mer)

COMPLEMENTARY                              A                 A   C
OLIGONUCLEOTIDES                    3'  AT  CCI CAI GGI AA  CT  TG   5'
(Deoxyinosine 20-mers)                     G                 G   T

cDNA-DEDUCED
NUCLEOTIDES      5'CTG GAG TGT TAC CAG TGC TAT GGA GTC CCA TTT GAG ACT TCT TGC CCA TCA ATT ACC TGC CCC TAC CCT GAT GGA GTC 3'

cDNA-DEDUCED
AMINO ACIDS        LEU GLU CYS TYR GLN CYS TYR GLY VAL PRO PHE GLU THR SER CYS PRO SER ILE THR CYS PRO TYR PRO ASP GLY VAL
```

Fig. 1. Ly-6E.1 amino acid sequence and synthetic oligonucleotide constructions. N-terminal sequence analysis of purified Ly-6E.1 protein was performed and enabled assignments of amino acids at positions 2–26. Weak signals allowed only low-confidence (lower case letters) or an ambiguous assignment (?) for certain residues. The second line indicates the possible codon redundancies (R = A or G; Y = C or T; X = A, G, C or T). The two types of oligonucleotides, the 35-mer and the 8-fold mixture of 20-mers, were synthesized with the (−) strand orientation (3′−5′) chosen. The insertion of deoxyinosine at the third position of codons is indicated by an I. The nucleotide and amino acid sequences deduced from the cloned Ly-6E.1 cDNA are shown in the bottom two lines.

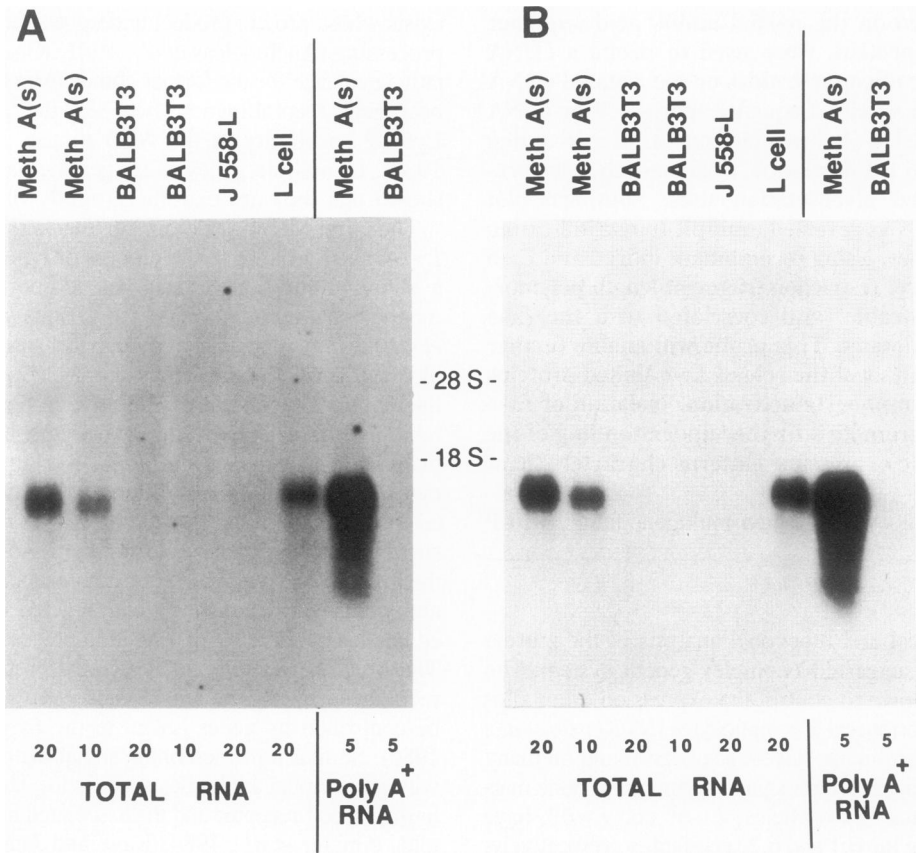

Fig. 2. The synthetic oligonucleotides and the Ly-6E.1 cDNA detect a 1.1-kb RNA species. 20 or 10 μg of total RNA, or 5 μg of poly(A)$^+$ RNA from the cell lines indicated were Northern blotted as described in Materials and methods. The blot in **panel A** was hybridized with the ^{32}P-labeled 35-mer synthetic oligonucleotide, washed and autoradiographed for 4 days. **Panel B** shows a 15 h autoradiograph of the same blot, after stripping and re-probing with nick-translated Ly-6E.1 insert of pKLy6.1-2R.

linked glycosylation is apparent (Palfree and Hammerling, 1986; Palfree *et al.*, 1986). The Ly-6E.1 has been isolated by immunoaffinity chromatography in sufficient quantity for partial amino acid sequence determination. This report presents the N-terminal amino acid sequence of the Ly-6E.1 protein and describes the isolation and characterization of a Ly-6E.1 cDNA clone. Hybridization experiments reveal that the *Ly-6* system does represent a complex, multigene family, as previously suggested by immunogenetic and biochemical analyses.

Results

Construction of synthetic oligonucleotides and cDNA cloning

Amino acid sequence analysis of the amino terminus of purified Ly-6E.1 protein allowed the determination of most of the amino acids from position 2 through position 26 (Figure 1, line 1). At certain residues however, the lack of a strong signal required either a low confidence or an ambiguous amino acid assignment. Since the most reliable amino acid assignment data was for

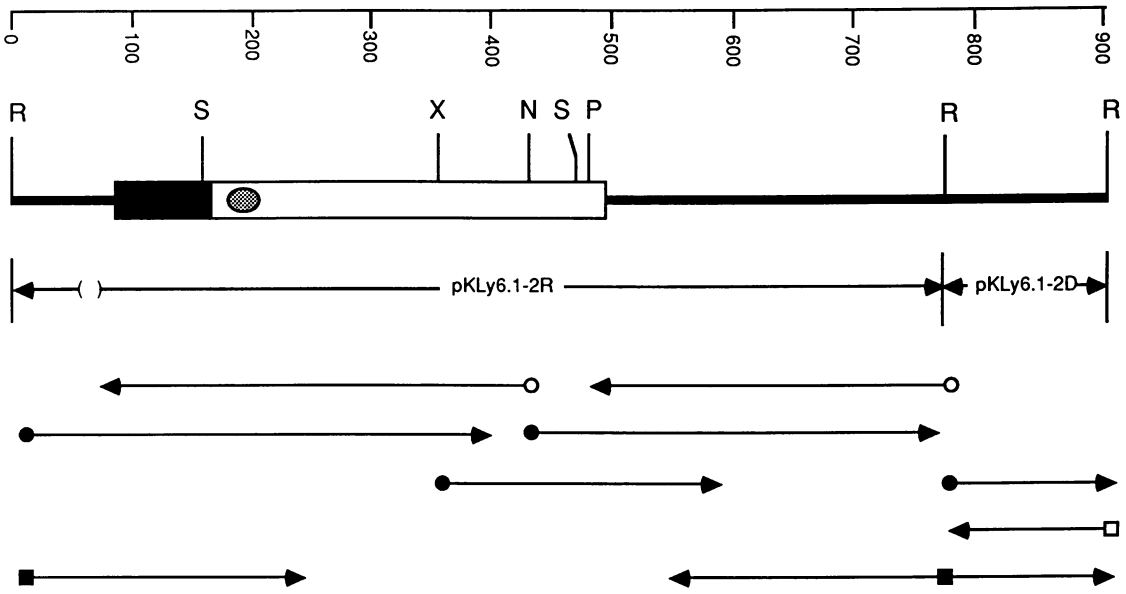

Fig. 3. Restriction enzyme map, structure and sequencing strategy of the Ly-6E.1 cDNA clone. The size scale is given in base pairs and the restriction enzymes designated: *Eco*RI (R), *Sst*I (S), *Xho*II (X), *Nco*I (N), *Pst*I (P). The shaded box represents the leader sequence nucleotides and the open box is the remainder of the coding sequence, with the region of oligonucleotide hybridization shown as the cross-hatched oval. Circles with arrows represent the sites and directions chosen for chemical DNA sequencing (Maxam and Gilbert, 1980) and the squares represent dideoxy sequencing directly from the pUC plasmid (Chen and Seeburg, 1985). Open and closed symbols indicate that the sequence was read from the top strand, or bottom strand, respectively.

residues 2–13, this portion of the protein was chosen for synthetic oligonucleotide construction.

To compensate for codon redundancy, two approaches have been adopted to construct synthetic oligonucleotides from amino acid sequence data. The first method is to construct a long oligonucleotide based on the most probable codon usage (Lathe, 1985). Using this predictive method, we synthesized a complementary, 35 base long oligonucleotide (Figure 1, line 3). An alternative is to synthesize a mixture of shorter oligonucleotides containing every possible sequence (Suggs *et al.*, 1981). The complexity of such a mixture can be reduced by using deoxyinosine at the third position of highly redundant codons (Ohtsuka *et al.*, 1985; Takahashi *et al.*, 1985). Following this strategy, we synthesized the 8-fold degenerate mixture of 20-mers, containing three deoxyinosines (Figure 1, line 4).

The oligonucleotides were designed to correspond to the sequence complementary to the mRNA to permit their use as probes on Northern blots. RNA preparations from various cell lines were resolved by electrophoresis in agarose, transferred to a nylon membrane and probed with ^{32}P-kinased oligonucleotides. The kinased 35-mer (Figure 2A) and the mixture of kinased 20-mers (not shown) produced identical patterns, although the latter gave a higher background. All probes hybridized with a 1.1-kb RNA expressed in MethA and L cells, but not in Ly-6E.1-negative BALB 3T3 fibroblasts or J558-L myeloma cells. As discussed later, the hybridization to RNA from L cells was unexpected as these cells do not express cell surface Ly-6E.1 protein (not shown).

In order to isolate cDNA encoding the Ly-6E.1 protein, poly(A)$^+$ RNA was isolated from MethA fibrosarcoma cells. This line has been shown to express Ly-6E.1 constitutively, and the antigen is indistinguishable from that produced by activated BALB/c lymphocytes when compared by gel electrophoresis (Palfree and Hammerling, 1986). Double-stranded cDNA was prepared and inserted via *Eco*RI linkers into the λgt11 vector, as described in Materials and methods. A portion of this library was plated without amplification and replicate nitrocellulose lifts were taken and probed with the ^{32}P-kinased 20-mers. From the number of phage in the MethA cDNA library which hybridized to the oligonucleotides, we estimate the frequency of the Ly-6.1 message to be of the order of 0.1% of poly(A)$^+$ RNA. This is certainly a conservative estimate, as selection with the oligonucleotides, constructed based on amino-terminal protein data, biases selection toward more full-length cDNAs. Hybridizing phage were selected and subcloned into pUC19. Recombinant plasmids were used to transform DH-5 bacteria and colonies were tested for hybridization to the kinased 20-mers.

Analysis of cDNA clones

The first phage selected for analysis, λKLy6.1-2, contained inserts of ~750 and 250 bp. Only the 750-bp insert (subclone pKLy6.1-2R) hybridized to the labelled oligonucleotides. The DNA sequence of the pKLy6.1-2R cDNA insert was determined (Maxam and Gilbert, 1980) and a region was found to be in excellent agreement with the cDNA nucleotide sequence predicted from amino acid analysis as illustrated in Figure 1. Except for a discrepancy at amino acid 3, (Cys not Gln), the deduced amino acid sequence perfectly corresponds to that obtained by amino-terminal analysis of the purified Ly-6E.1 protein. When the cDNA insert of pKLy6.1-2R was used to re-probe the Northern blot (Figure 2B), the hybridization pattern obtained was identical to that of the oligonucleotide probes but showed a marked increase in sensitivity.

Although pKLy6.1-2R contained the entire coding sequence, it lacked a 3′-terminal poly(A) sequence. Sequencing of the 250-bp insert of subclone pKLy6.1-2D revealed a 109-nucleotide sequence followed by >150 adenosines of the poly(A) tail. A restriction map of the Ly-6E.1 cDNA, together with the sequencing strategies, is presented in Figure 3.

The cDNA insert of pKLy6.1-2R was used as a hybridization probe to determine its relatedness to other recombinant phage (and their subclones) selected by hybridization to the synthetic oligonucleotides. Restriction mapping showed the independent subclone, pKLy6.1-5W, to be slightly longer at the 5′ end. It

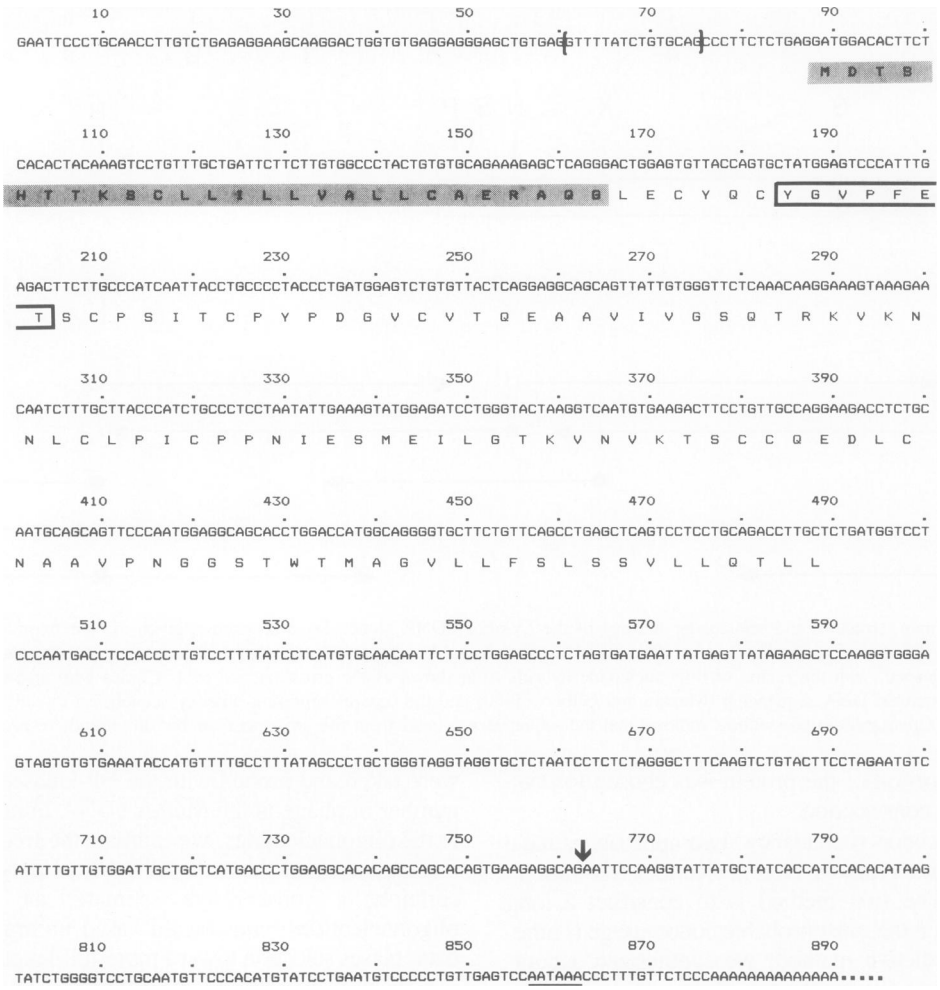

Fig. 4. Ly-6E.1 cDNA nucleotide and amino acid sequence. The complete nucleotide sequence of the Ly-6E.1 cDNA is presented with the deduced amino acids displayed in single letter code below. The amino acids comprising the leader sequence are shown shaded and the location of those used as the basis for synthesis of the 20-mer oligonucleotides are enclosed in the open box. The 15 bases contained in only one of the Ly-6E.1 clones are bracketed. The position of the internal *Eco*RI site is shown by an arrow, and the polyadenylation signal is underlined.

was discovered by DNA sequencing, however, that the pKLy6.1-5W cDNA insert had exactly the same 5′ and 3′ termini as pKLy6.1-2R, and was identical throughout, except for the addition of 15 bp near the 5′ end (position 61−75, shown bracketed in Figure 4). The possible origin and significance of these 15 bp are discussed later. The fact that the two independent subclones pKLy6.1-2R and pKLy6.1-5W have the same 5′ terminus suggests that this is either the actual 5′ end of the Ly-6.1 message, or another natural *Eco*RI site in the cDNA. Since the Ly-6.1 mRNA size is judged to be ~1.1 kb, and we have accounted for at least 1 kb, any additional 5′ sequence would be limited to 0.1 kb. We have looked for such a small cDNA fragment by *Eco*RI digestion and ^{32}P-labelling of purified phage DNA but have no evidence for its existence.

Features of the Ly-6.1 protein

Figure 4 shows the Ly-6E.1 cDNA sequence. The reading frame was determined by the Ly-6E.1 amino-terminal protein sequence. Analysis of the complete cDNA sequence indicated that this was in fact the only large open reading frame possible. The DNA sequence contains 88 nucleotides of 5′-untranslated sequence, 402 nucleotides of coding sequence and 386 bases of 3′-untranslated sequence with the AATAAA polyadenylation signal 15 bp upstream of the start of the poly(A) tail. Protein translation begins at the ATG codon at position 89 and continues for 134 codons until the termination codon TGA is read at nucleotide 491. The initiator methionine is followed by 25 predominantly hydrophobic amino acids which probably constitute the signal sequence characteristic of cell membrane-bound or secreted proteins. The site for signal sequence cleavage is assigned based on the amino-terminal sequence data. This location corresponds to the optimal cleavage position determined by statistical analysis (von Heijne, 1983). Following removal of this signal sequence, the mature Ly-6.1 protein would consist of 108 amino acids with a calculated mol. wt of 11 072 daltons.

The relative decrease in electrophoretic mobility on polyacrylamide gels of Ly-6E.1 protein following reduction suggested a high degree of internal disulfide bonding. Immunoprecipitation experiments using lysates from tunicamycin-treated or untreated cells indicated a lack of asparagine-linked glycosylation of the Ly-6E.1 protein (Palfree and Hammerling, 1986). The presence of 10 cysteines and the absence of [Asn-X-Ser/Thr] sites in the mature protein deduced from the cDNA sequence is consistent with these observations. To visualize the hydrophilic/hydrophobic nature of the Ly-6E.1 protein, hydrophilicity analysis was performed (Hopp and Woods, 1981) with the results shown in Figure 5. This analysis illustrates the hydrophobic core of the signal se-

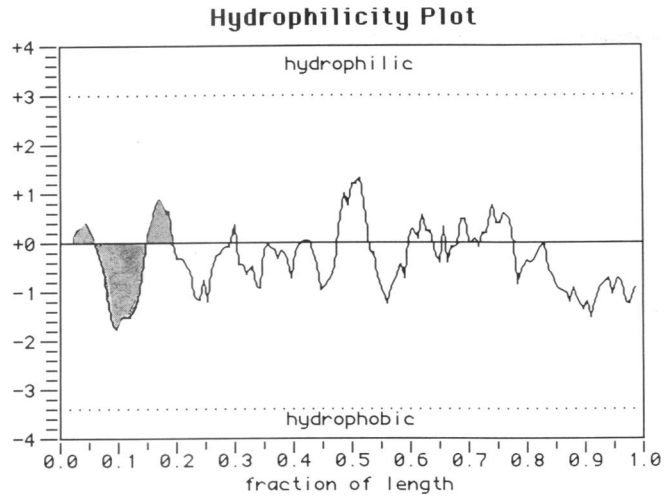

Fig. 5. Hydrophilicity plot of Ly-6E.1 cDNA-deduced amino acids. The hydrophilicity values of amino acids were determined and repetitively averaged across a range of six residues (Hopp and Woods, 1981). The results are plotted as the fraction of the total length of 134 amino acids. The first 26 amino acids represent the leader sequence and are shown shaded.

quence (shaded) as well as the overall hydrophobic nature of the first and last third of the mature protein.

The Ly-6E.1 nucleotide sequence and the deduced amino acid sequence were analyzed for homology to reported sequences in the Genbank and EMBL nucleotide sequence libraries and the NBRF protein sequence database (Devereux *et al.*, 1984; Wilbur and Lipman, 1983; Lipman and Pearson, 1985). No significant homologies at either the protein or nucleic acid level were noted. This indicates that the Ly-6E.1 protein, expressed predominantly in lymphocytes, is a novel protein with no obvious relationship to Thy-1 or the MHC, immunoglobulin, and T cell receptor multigene families (Williams, 1984).

Southern blotting reveals multiple genes and RFLPs

Southern blotting of DNA from the MethA cells and from various inbred mouse strains was performed using the Ly-6E.1 cDNA insert of pKLy6.1-2R as the probe. As shown in Figure 6, this probe hybridized with numerous fragments, suggesting that the *Ly-6* locus contains multiple genes. Preliminary results from genomic cloning of *Ly-6*-related genes support the multigene nature of the *Ly-6* locus. The Ly-6E.1 cDNA has been used to select 47 genomic clones. Restriction mapping and DNA sequence analysis in progress indicate the presence of at least 10 closely related, non-identical genes.

The similarity of the banding pattern of MethA DNA to BALB/c DNA indicates that no gross rearrangement of genes in the *Ly-6* family occurred during the generation or maintenance of this BALB/c-derived fibrosarcoma. Southern blot restriction fragment length polymorphisms (RFLPs) have been defined in studies of other polymorphic, mutigene families (Cami *et al.*, 1981; Wake *et al.*, 1982). The seven (C × B) recombinant inbred strains, and strain C3H and its *Ly-6* congenic partner, C3H.B6-*Ly-6*^{b} (Potter *et al.*, 1980), have proven useful in establishing monoclonal antibodies as recognizing *Ly-6* locus-controlled protein polymorphisms (Flood *et al.*, 1985; Reiser *et al.*, 1986). To determine whether the *Ly-6* genes exhibit any RFLP, we used spleen DNA of these strains for Southern blotting. Figure 6 shows that although many of the hybridizing bands are present in all lanes, others are clearly polymorphic. Except for minor differences, two patterns are discernable. One pattern

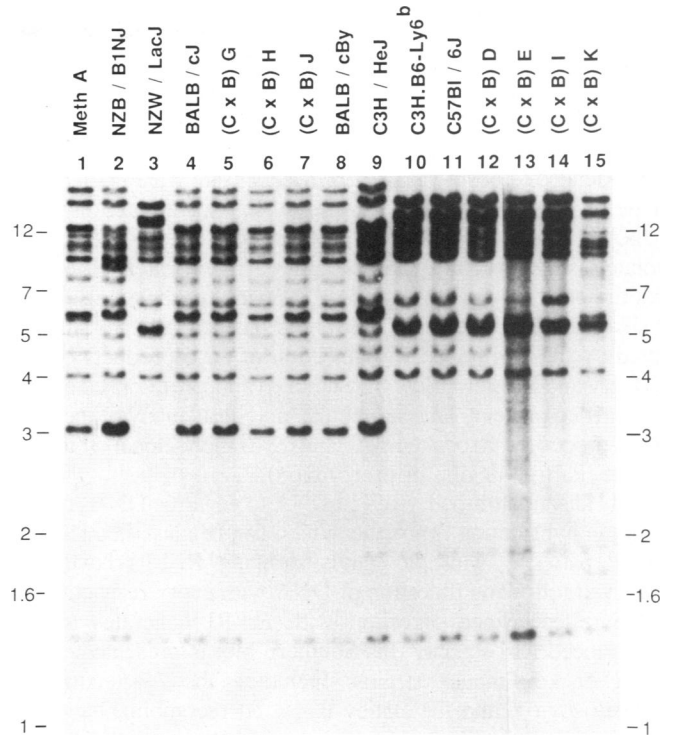

Fig. 6. The Ly-6E.1 cDNA probe identifies multiple bands and RFLPs in murine genomic DNA. 5 μg of DNA from MethA cells, or spleen DNA of the inbred mouse strain indicated, were digested with *Eco*RI and subjected to Southern analysis as described in Materials and methods. To compensate for unequal amounts of DNA loaded, a 3 h autoradiograph is presented for **lanes 2–12**, and one of 15 h shown for **lanes 1** and **13–15**. The autoradiograph shown is from a blot washed with 2 × SSC at 65°C. Approximate sizes of DNA marker fragments, in kilobases, are indicated at either side of the autoradiograph.

is shared by lanes 1, 2, and 4–9, and another is shown by lanes 3, and 10–15. This grouping of strains based on genomic DNA RFLPs exactly correlates with the established *Ly-6* allele. This point is most clearly seen by comparing the hybridization pattern of DNA from the C3H/HeJ (*Ly-6*^{a}) strain in lane 9 with that of DNA from the *Ly-6* congenic C3H.B6-*Ly-6*^{b} (*Ly-6*^{b}) strain shown in lane 10. Under more stringent washing conditions (0.1 × SCC, 65°C), some of the bands faded markedly, yet the polymorphic bands remained relatively unaffected (not shown). A close inspection of the hybridization pattern of NZB/B1NJ DNA (lane 2) reveals a reduction in intensity of the band at ~13 kb and the presence of a novel band at ~9 kb when compared with the patterns of DNA from other *Ly-6*^{a} strains (lanes 1,and 4–9). As mentioned earlier, the NZB mouse strain is the only strain showing a cross-over within the *Ly-6* region. NZB mice express the Ly-6B.2 alloantigen but type as Ly-6.1 for all other *Ly-6* specificities (Kimura *et al.*, 1984). These mice have also been shown to be incapable of expressing the Ly-6C specificity as defined by the xenogeneic 6-C-3 monoclonal antibody (Dumont *et al.*, 1985). Whether either or both of these observations can be linked to the hybridization pattern anomalies reported here is a matter for further study.

Discussion

Using the well-characterized anti-Ly-6E.1 monoclonal antibody SK70.94 (Kimura *et al.*, 1984), Ly-6E.1 protein was immuno-

affinity purified from MethA fibrosarcoma cells (Palfree and Hammerling, 1986). Amino-terminal sequence data obtained from this purified protein formed the basis for construction of synthetic oligonucleotides. The cDNA clone that was isolated from the MethA cDNA library predicted the Ly-6E.1 amino acids found at positions 2−13 with only one error. The correspondence also extended to positions 14−26, residues that were not involved in the construction of the synthetic oligonucleotides used for cDNA selection. This strongly indicates that the cDNA we have isolated is the structural gene for the purified protein. Evidence that the protein is in fact Ly-6E.1 has been presented elsewhere (Palfree and Hammerling, 1986), and has been further substantiated by the highly specific detection of *Ly-6* products on cell surfaces and Western blots by a chicken antiserum raised against the purified protein. This serum also manifests many of the functional properties associated with anti-Ly-6 monoclonal antibodies (R.G.E.Palfree *et al.*, in preparation).

Hybridization of the Ly-6E.1 cDNA to genomic DNAs reveals multiple hybridizing fragments which can be classified into two RFLP patterns. Multiple bands showing RFLPs have been observed following digestion of DNA with every restriction enzyme tested. Since digestion with *Eco*RI generates a most pronounced RFLP, only this Southern blot is presented. Using DNA of key mouse strains, including the *Ly-6* congenic, C3H.B6-*Ly-6*^{b}, and the Bailey (C × B) recombinant inbreds, a perfect correspondence between the RFLP pattern and the *Ly-6* allele has been demonstrated. This correspondence has been observed for DNA of 25 strains tested as well as for the 12 (B × H) recombinant inbred lines (not shown). The *Ly-6* locus was originally mapped to mouse chromosome 9 (Horton and Hetherington, 1980). Later it was reassigned to chromosome 2 (Meruelo *et al.*, 1982). We have used the Ly-6E.1 cDNA to probe DNA from a panel of (hamster × mouse) somatic cell hybrids. This analysis showed that all the hybridizing fragments are contained on mouse chromosome 15 (K.P.LeClair *et al.*, submitted). This result is consistent with immunogenetic studies which have demonstrated a tight linkage of the Ly-6 genes.

With the exception of L cells, the expression of the Ly-6 mRNA was correlated with cell surface expression of the Ly-6E.1 protein. L cells contain a 1.1-kb mRNA that hybridizes with the Ly-6E.1 cDNA, yet these cells do not express Ly-6E.1 protein on their surface as determined by fluorescence-activated cell sorting (FACS) analysis (not shown). Since L cells, derived from a C3H mouse, are genetically *Ly-6*^{a}, they should be capable of expressing the Ly-6E.1 specificity. The positive hybridization observed could be due to the expression of another member of the *Ly-6* multigene family, a modification of the post-translational processing of the Ly-6E.1 core protein or to a lesion that exists in these cells between *Ly-6*^{a} transcription and the appearance of the Ly-6E.1 epitope at the cell surface.

Inserts from two independent phage were sequenced and found to be identical except for a 15-base sequence present in the 5′-untranslated region of one but not the other. The absence of this sequence could be an artifact due to secondary structure in the templates during cDNA synthesis. Alternatively, the two forms could represent examples of differential splicing. The 15 bases and the two nucleotides just 5′ to them,

AG(GTTTTATCTGTGCAG)

closely resemble the splicing consensus,

AG(GURAG.Y-YYY-CAG)

especially as a 3′ splicing site (Sharp, 1981). Isolation and characterization of the appropriate Ly-6 genomic clone should resolve this matter.

The mol. wt of the Ly-6E.1 protein following removal of the signal sequence is calculated to be 11 072 daltons, somewhat smaller than the 17 kd size estimated from SDS−PAGE analysis. No N-linked glycosylation sites have been found but some O-linked glycosylation would be consistent with the isoelectric focusing pattern of Ly-6.1 protein which shows at least five forms with pIs between 4.0 and 5.2 (Palfree *et al.*, 1986). Further study of O-linked glycosylation of the Ly-6.1 protein is in progress. Analysis of hydrophobicity indicates that the mature protein contains two hydrophobic regions, one near the amino terminus and another at the carboxyl terminus. The last 30 amino acids contain no charged amino acids which have been found to demarcate the end of the transmembrane segment of many integral membrane proteins. However, the presence of the polar amino acids Gln and Thr as part of the end of the protein (-Gln-Thr-Leu-Leu-COOH) might indicate that this short segment extends through the lipid bilayer and is located in the cytoplasm of the cell. Alternatively, the Ly-6E.1 protein could be anchored in the cell membrane by a hydrophobic lipid moiety such as phosphatidylinositol, as has been reported for the Thy-1 glycoprotein (Low and Kincade, 1985; Tse *et al.*, 1985). One class of Thy-1$^-$ mutants has been described which have a biosynthetic defect which also affects the expression of Ly-6 specificities (Horton and Hyman, 1983). The Thy-1 glycoprotein, Ly-6E.1/Ly-6A.2 and TAP are the only structures on murine T cells which have been implicated in antigen-independent proliferative signalling (Gunter *et al.*, 1984; Flood *et al.*, 1985; Rock *et al.*, 1986; Kroczek *et al.*, 1986). Experiments to determine the presence and composition of a lipid-anchoring domain on the Ly-6E.1 protein are in progress.

Phosphorylation of serine residues of membrane-associated proteins has been demonstrated in the activation of T lymphocytes (Samelson *et al.*, 1985b). Functional studies with anti-Ly-6 monoclonal antibodies have shown that *Ly-6*-linked proteins play a critical role in the activation of T lymphocytes. We have noted the presence of four serine (but no tyrosine) residues in the hydrophobic carboxyl-terminal region of the Ly-6E.1 protein, which may be targets of phosphorylation. Experiments to study this are in progress, but we have no evidence yet for phosphorylation of the Ly-6E.1 protein.

Of obvious interest is the relationship between the different members of the *Ly-6* multigene famiy and the known *Ly-6*-controlled antigens. At present, biochemical data support the existence of two distinct categories of *Ly-6*-controlled proteins. One represented by Ly6A.2/Ly6E.1 and TAP, contains proteins tightly constrained by internal disulfide bonds, whereas the other category includes the Ly6C (Ly28) and TAPa proteins which exhibit no internal disulfide bonding (Palfree and Hammerling, 1986; Rock *et al.*, 1986). Since disulfide structures are generally highly conserved among members of gene families it is possible that the genes for Ly6C and TAPa are not members of the family detected by our cDNA probes and that more *Ly-6* complex genes have yet to be identified at the molecular level. In this regard, 47 genomic clones have been isolated from a MethA genomic library by selection with the Ly-6E.1 cDNA. Transfection of these genomic clones may allow the identification and functional analysis of other related members of the *Ly-6* gene family. Characterization of the genomic clones may also allow construction of specific probes for a detailed study of expression in various tissues and at different stages of development.

Materials and methods

Protein sequencing, oligonucleotide synthesis and radiolabeling

Ly−6E.1 protein was isolated from the MethA fibrosarcoma cells by anti-Ly6E.1 (SK70.94) immunoaffinity chromatography followed by Sephacryl S-200 gel filtration as described (Palfree and Hammerling, 1986). Amino-terminal sequencing of the purified protein was performed in the Yale Protein Chemistry Facility using an Applied Biosystems gas phase sequencer. Oligonucleotides were synthesized by the phosphoramidite method on an Applied Biosystems Model 381A DNA synthesizer, and purified on a preparative (12% acrylamide/7 M urea) gel. For 5′-end labelling, equimolar amounts of oligonucleotide and crude [γ-^{32}P]ATP, 7000 Ci/mM (ICN, Irvine, CA) were incubated with >10 units of T4 polynucleotide kinase (Collaborative Research, Lexington, MA) in a buffer containing 70 mM Tris−HCl pH 7.6, 10 mM $MgCl_2$, 0.1 mM KCl, 5 mM dithiothreitol (DTT), 1 mM spermidine, for 30 min at 37°C. EDTA was added to a final concentration of 20 mM to stop the reaction. As measured by trichloroacetic acid precipitation analysis, incorporation of ^{32}P was routinely >80%. Labelled oligonucleotides were used as probes of the Northern blot and the cDNA library with no further purification.

Construction and screening of the MethA cDNA library

Total cellular RNA was isolated from MethA fibrosarcoma cells using the guanidinium thiocyanate technique of Chirgwin *et al.* (1979). Poly(A)$^+$ RNA was selected by two cycles of binding to oligo(dT)−cellulose. With the exception of the addition of actinomycin D (Boehringer-Mannheim, Indianapolis, IN) to the first strand reaction mixture (Watson and Jackson, 1985), the synthesis of double-stranded cDNA was performed exactly as described by Gubler and Hoffman (1983). This method employs RNase-H (Bethesda Research Labs, Gaithersburg, MD), nicotinamide−adenine dinucleotide (B−M), *Escherichia coli* DNA polymerase I and *E. coli* DNA ligase (New England Biolabs, Beverly, MA) in the second strand synthesis reaction. Subsequent steps in processing of the double strand cDNA are discussed in detail by Watson and Jackson (1985), and only the significant modifications and reagent sources are noted below. To protect internal *Eco*RI sites, double-stranded cDNA was incubated with 1 mM S-adenosylmethionine (NEBL), 60 units of *Eco*RI methylase (NEBL), for 30 min at 37°C in a buffer containing 100 mM Tris−HCl pH 8, 1 mM EDTA and 0.4 mg/ml BSA (Reinach and Fischman, 1985). To render the cDNA blunt-ended, it was treated with 9 units of phage T4 DNA polymerase (NEBL) for 30 min at 37°C in a buffer containing 50 mM Tris−HCl pH 8, 6 mM $MgCl_2$, 25 mM KCl, 1 mg/ml BSA and 0.1 mM of each of the deoxynucleotide triphosphates. Kinased, synthetic *Eco*RI linkers, d(GGAATTCC) (NEBL), were ligated onto the cDNA at 4°C for 30 h using 400 units of T4 DNA ligase (NEBL) in 66 mM Tris−HCl pH 7.3, 5 mM $MgCl_2$, 5 mM DTT and 1 mM ATP. Linkered cDNA was heated to 65°C for 10 min to inactivate the ligase and then digested with 300 units of *Eco*RI (NEBL) to remove excess linkers. The cDNA was separated from the excess linkers by fractionation on a Sepharose-4B column. Fractions containing cDNA were collected, concentrated and ligated into *Eco*RI-digested, dephosphorylated λgt11 vector arms (Vector Cloning Systems, San Diego, CA), packaged and titered as described (Huynh *et al.*, 1985; Maniatis *et al.*, 1982).

Recombinant phage were plated without amplification and nitrocellulose lifts were taken and processed (Benton and Davis, 1977). For screening with oligonucleotides, the filters were pre-hybridized for 3 h at 37°C in 6 × SSC (1 × SSC = 0.15 M sodium chloride, 0.015 M sodium citrate pH 7.0), 5 × Denhardt's solution (1 × Denhardt's = 0.02% each of Ficoll 400, polyvinylpyrrolidone and BSA), 0.1% sodium lauryl sarcosine, 1 mM sodium pyrophosphate and 50 μg/ml sonicated, denatured, salmon sperm carrier DNA. The mixture of 20-mers was 5′ end-labeled and ~1 × 10^6 c.p.m./ml (sp. act. 1.1 × 10^8 c.p.m./μg; 3.7 × 10^6 c.p.m./pmol) were added to the filters for hybridization at 37°C for 22 h. Following hybridization, the filters were washed at 37°C with several changes of 6 × SSC, 0.1% sodium lauryl sarcosine, and then dried and exposed to Kodak XAR-5 film at −70°C using an intensifying screen.

Northern and Southern blot analysis

For Northern blot analysis, total RNA or poly(A)$^+$ RNA samples were denatured for 1 h at 50°C in 1 M glyoxal, separated by electrophoresis on a 1.5% agarose gel in 10 mM sodium phosphate pH 7.0 and transferred to a Zetabind nylon filter (AMF Cuno, Meriden, CT). The filter was exposed to short wave u.v. irradiation for 3 min then deglyoxylated by soaking in 20 mM Tris−HCl pH 8.0 at 100°C and allowed to cool to room temperature. The filter was pre-hybridized for 30 min at 65°C in 0.5 M sodium phosphate pH 6.8, 1% BSA, 7% SDS and 1 mM EDTA (Church and Gilbert, 1984). For hybridization, 1 ng of radiolabeled oligonucleotide (see above) was added per ml of hybridization solution, (Wood *et al.*, 1985), and incubated for 26 h at 37°C. The blot was then rinsed at 37°, 48° or 56°C in 6 × SSC, 0.1% SDS. Figure 2A shows the results of probing with the 35-mer oligonucleotide when washed at 56°C.

For re-hybridization, the nylon membrane was soaked three times for 15 min in 0.1 × SSC, 0.1% SDS at 95°C. The blot was autoradiographed to insure complete removal of probe (not shown) then incubated for 6 h at 65°C with the BSA/SDS hybridization buffer described above. The cDNA insert of pKLy6.1-2R was nick-translated (sp. act. 9.6 × 10^8 c.p.m./μg) and 1 × 10^6 c.p.m./ml were added for hybridization at 65°C for 16 h. The blot was washed twice for 10 min at 65°C with 0.5% BSA, 5% SDS, 40 mM sodium phosphate pH 6.8, 1 mM EDTA then four times for 10 min at 65°C with 1% SDS, 40 mM sodium phosphate pH 6.8 and 1 mM EDTA (Church and Gilbert, 1984). The blot was dried and exposed to Kodak XAR-5 film at −70°C with an intensifying screen.

For Southern blotting, DNA was isolated from MethA tissue culture cells as described (Steffen *et al.*, 1979). Spleen DNA from the various mouse strains used was purchased from the DNA Resources of Jackson Laboratories (Bar Harbor, ME). 5 μg of DNA were digested for 6 h at 37°C with 35 units of *Eco*RI, and loaded directly onto a 0.8% agarose gel. Following electrophoresis, the gel was processed and blotted onto nitrocellulose as described (Southern, 1975). After baking, the filter was pre-hybridized for 5 h at 42°C in 50% formamide, 5 × SSC, 100 μg/ml carrier DNA, 2.5 × Denhardt's, 0.2% SDS and 10% dextran sulfate. The 750-bp insert of pKLy6.1-2R was nick-translated to a sp. act. of 1 × 10^9 c.p.m./μg. The probe was denatured by boiling for 5 min and 0.5 × 10^6 c.p.m./ml were added to the filter. After hybridization for 12 h at 42°C, the probe was removed and the blot was washed with several changes of 2 × SSC, 0.1% SDS for 1 h at 65°C, dried, and exposed to Kodak XAR-5 film at −70°C with an intensifying screen for the times indicated in Figure 6.

Acknowledgements

We thank Dr Kenneth Williams and Ms Kathryn Stone of the Yale Protein Chemistry Facility for expert amino-terminal sequencing of the purified Ly-6.1 protein. We are grateful to Ms Patricia Pace for assistance with oligonucleotide syntheses, to Dr William Philbrick for helpful suggestions concerning nick translations and Southern blotting and to Dr Dean Ballard who provided invaluable technical advice throughout many stages of this project. We thank Drs Charles Janeway, Donald Murphy and Adrian Hayday for critical evaluations of this manuscript. This work was supported by grants CA29606 and T32-A1-07019 from the National Institutes of Health. The laboratories of A.B. and P.M.F. are supported by the Howard Hughes Medical Institute.

References

Allison,J.P. and Lanier,L. (1985) *Nature*, **314**, 107−109.
Auchincloss,H., Ozato,K. and Sachs,D. (1981) *J. Immunol.*, **127**, 1839−1843.
Benton,W.D. and Davis,R.W. (1977) *Science*, **196**, 180−182.
Cami,B., Bregegere,F., Abastado,J.P. and Kourilsky,P. (1981) *Nature*, **291**, 673−675.
Campbell,D.G., Gagnon,J., Reid,K.B.M. and Williams,A.F. (1981) *Biochem. J.*, **195**, 15−30.
Chen,E.Y. and Seeburg,P.H. (1985) *DNA*, **4**, 165−170.
Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294−5299.
Church,G.M. and Gilbert,W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1991−1995.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387−395.
Dumont,F.J., Coker,L.Z., Habbersett,R.C. and Treffinger,J.A. (1985) *J. Immunol.*, **134**, 2357−2365.
Flood,P.M., Murphy,D.B., Horowitz,M., LeClair,K.P., Smith,F., Stockert,E., Palladino,M.A. and DeLeo,A.B. (1985) *J. Immunol.*, **135**, 63−72.
Gubler,U. and Hoffman,B.J. (1983) *Gene*, **25**, 263−269.
Gunter,K.C., Malek,T.R. and Shevach,E.M. (1984) *J. Exp. Med*, **159**, 716−730.
Hibbs,M.L., Hogarth,P.M., Scott,B.M., Harris,R.A. and McKenzie,I.F.C. (1984) *J. Immunol.*, **133**, 2619−2623.
Hogarth,P.M., Houlden,B.A., Latham,S.E., Sutton,V.R. and McKenzie,I.F.C. (1984) *Immunogenetics*, **20**, 57−69.
Hopp,T.P. and Woods,K.R. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 3824−3828.
Horton,M.A. and Hetherington,C.M. (1980) *Immunogenetics*, **11**, 521−525
Horton,M.A. and Hyman,R. (1983) *Immunogenetics*, **17**, 261−270.
Houlden,B.A., Hogarth,P.M. and McKenzie,I.F.C. (1986) *Immunogenetics*, **23**, 226−232.
Huynh,T., Young,R.A. and Davis,R.W. (1985) In Glover,D.M. (ed.), *DNA Cloning, A Practical Approach*. IRL Press, Oxford and Washington DC, Vol. 1, pp. 49−78.
Kaye,J. and Janeway,C.A. (1984) *J. Exp. Med.*, **159**, 1397−1412.
Kimura,S., Tada,N., Liu-Lam,Y. and Hammerling,U. (1984) *Immunogenetics*, **20**, 47−56.
Kroczek,R.A., Gunter,K.C., Germain,R.N. and Shevach,E.M. (1986) *Nature*, **322**, 181−184.
Lathe,R. (1985) *J. Mol. Biol.*, **123**, 1−12.
Lipman,D.J. and Pearson,W.R. (1985) *Science*, **227**, 1435−1441.
Low,M.G. and Kincade,P.W. (1985) *Nature*, **318**, 62−64.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, NY.
Maxam,A. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.
McKenzie,I.F.C., Cherry,M. and Snell,G.D. (1977) *Immunogenetics*, **5**, 25–32.
Meuer,S.C., Hussey,R.E., Cantrell,D.A., Hodgdon,J.C., Schlossman,S.F., Smith,K.A. and Reinherz,E.L. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1509–1513.
Meruelo,D., Offer,M. and Rossomando,A. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7460–7464.
Oettgen,H.C., Pettey,C.L., Maloy,W.L. and Terhorst,C. (1986) *Nature*, **320**, 272–275.
Ohtsuka,E., Matsuki,S., Ikehara,M., Takahashi,Y. and Matsubara,K. (1985) *J. Biol. Chem.*, **260**, 2605–2608.
Palfree,R.G.E. and Hammerling,U. (1986) *J. Immunol.*, **136**, 594–600.
Palfree,R.G.E., Dumont,F.J. and Hammerling,U. (1986) *Immunogenetics*, **23**, 197–207.
Potter,T.A., McKenzie,I.F.C., Morgan,G.M. and Cherry,M. (1980) *J. Immunol.*, **125**, 541–545.
Reinach,F. and Fischman,D. (1985) *J. Mol. Biol.*, **181**, 411–422.
Reiser,H., Yeh,E.T., Gramm,C.F., Benacerraf,B. and Rock,K.L. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2954–2958.
Rock,K.L., Yeh,E.T., Gramm,C., Haber,S.I., Reiser,H. and Benacerraf,B. (1986) *J. Exp. Med.*, **163**, 315–333.
Samelson,L.E., Harford,J.B. and Klausner,R.D. (1985a) *Cell*, **43**, 223–231.
Samelson,L.E., Harford,J., Schwartz,R.H. and Klausner,R.D. (1985b) *Proc. Natl. Acad. Sci. USA*, **82**, 1969–1973.
Sharp,P. (1981) *Cell*, **23**, 643–646.
Southern,E. (1975) *J. Mol. Biol.*, **98**, 503–517.
Steffen,D., Bird,S., Rowe,W.P. and Weinberg,R.A. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4554–4558.
Suggs,S.V., Wallace,R.B., Hirose,T., Kawashima,E.H. and Itakura,K. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 6613–6617.
Takahashi,Y., Kato,K., Hayashizaki,Y., Wakabayashi,T., Ohtsuka,E., Matsuki,S., Ikehara,M. and Matasubara,K. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1931–1935.
Takei,F., Galfre,G., Anderson,T., Lennox,E.S. and Milstein,C. (1980) *Eur. J. Immunol.*, **10**, 241–246.
Takei,F. (1982) *Immunogenetics*, **16**, 201–208.
Takei,F. (1984) *J. Immunol.*, **133**, 345–350.
Tse,A.G.D., Barclay,A.N., Watts,A. and Williams,A.F. (1985) *Science*, **230**, 1003–1008.
von Heijne,G. (1983) *Eur. J. Biochem.*, **133**, 17–21.
Wake,C.T., Long,E.O. and Mach,B. (1982) *Nature*, **300**, 372–374.
Watson,C.J. and Jackson,J.F. (1985) In Glover,D.M. (ed.), *DNA Cloning. A Practical Approach*. IRL Press, Oxford and Washington DC, Vol. 1, pp. 79–88.
Wilbur,W.J. and Lipman,D.J. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 726–730.
Williams,A. (1984) *Nature*, **308**, 12–13.
Wood,W.I., Gitschier,J., Lasky,L.A. and Lawn,R.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1585–1588.
Yotoku,M., Grossberg,A.L. and Pressman,D. (1974) *J. Immunol.*, **112**, 1774–1781.

Received on 6 August 1986