

High-frequency deletional rearrangement of immunoglobulin κ gene segments introduced into a pre-B-cell line

(site-specific recombination/translation initiation)

PETER ENGLER*[†] AND URSULA STORB*

*Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637; and [†]Department of Microbiology and Immunology, University of Washington, Seattle, WA 98195

Communicated by Hewson Swift, March 16, 1987 (received for review January 12, 1987)

ABSTRACT We describe an immunoglobulin gene recombination indicator in which a specific rearrangement via deletion results in the acquisition of a dominant phenotype. The indicator consists of the *Escherichia coli* xanthine/guanine phosphoribosyltransferase (*gpt*) gene, whose translation is prevented by the presence of an upstream initiation codon out of frame with respect to the *gpt* coding sequence. Flanking this barrier initiation codon are the heptamer-spacer-nonamer recognition sequences from a κ chain variable region (V_{κ}) and from a κ chain joining region (J_{κ}). A proper V–J joint results in the deletion of the translational barrier and allows expression of the selectable marker. When tested by transfection into fibroblasts, no rearrangements were detected and the presence of the barrier initiation codon was sufficient to completely abolish *gpt* expression in these cells. Similarly, no rearrangements were detected after transfer of the test gene into myeloma cells. However, when the construct was introduced into the pre-B-cell line 38B9, >80% of the transfected cells showed evidence of a specific rearrangement. These rearrangements were associated with the translation of *gpt*, although no selection for its expression was needed. DNA sequence analysis of six different V–J joints revealed that the rearrangement proceeded with a high degree of accuracy. These results indicate that only very minimal DNA sequences (21 base pairs 5' of the V heptamer and 4 base pairs 3' of its nonamer; <45 base pairs 5' of the J nonamer and 3' of its heptamer) are required for efficient rearrangement and provide formal proof that κ gene segments can rearrange by a deletional mechanism.

As postulated over 20 years ago, the variable (V) and constant regions of immunoglobulin chains are encoded by separate genetic elements (1). During the development of an antibody-producing cell, these elements are rearranged to yield functional immunoglobulin genes. Closely associated with each of these elements [V, D (diversity), and J (joining) in the case of heavy-chain genes; V and J for light-chain genes; see ref. 2 for a review] are highly conserved DNA sequences consisting of a heptamer, a spacer, and a nonamer, which presumably constitute at least a portion of the sites recognized by the enzyme(s) involved in the recombination.

Much of our knowledge of the process of immunoglobulin gene rearrangement has come from comparison of the germ-line genes with their rearranged counterparts. For a more thorough understanding of various aspects of the recombination, a system that actively rearranges introduced substrates is desirable. Recently, two such systems in which recombination substrates are introduced into pre-B-cell lines have been described and have contributed greatly to our understanding of immunoglobulin gene rearrangement. One system depends on a V_{κ} – J_{κ} rearrangement via inversion to flip a selectable marker gene into the proper transcriptional

orientation (3, 4). A number of recombinants have been isolated and sequenced, revealing that authentic immunoglobulin-like rearrangements had occurred. A feature that limits the usefulness of this system is the unknown but generally low rate of rearrangement observed. Another system, which depends on a D–J rearrangement to delete an interposed thymidine kinase gene, has been described (5, 6). Deletion of the thymidine kinase gene, at least under certain conditions, occurs at a remarkably high rate. Sequence analysis of the joints demonstrated that they were similar to normal D–J rearrangements. A similar approach was used to demonstrate that pre-B cells are able to rearrange T-cell receptor gene segments, suggesting that B and T cells use a common recombinase (7). Use of this type of recombination substrate to select for rearrangement is limited to mutant cells deficient in thymidine kinase, which narrows the range of potential recipients. Furthermore, thymidine kinase activity can be lost, albeit at a low frequency, by events other than immunoglobulin-like rearrangements. This may become a serious problem when experiments designed to detect very rare events are undertaken.

The recombination indicator described here is designed so that a rearrangement via deletion results in a gain, rather than a loss, of function. This difference should result in a significant reduction of signals in the absence of an authentic immunoglobulin-like rearrangement. The inhibition of selectable marker expression was achieved by placing a translational block just upstream of the xanthine/guanine phosphoribosyltransferase (*gpt*) gene. To allow expression in the event of a specific rearrangement, it was flanked by recognition sequences from V and J such that a rearrangement would delete the block. The rationale behind this approach was provided by the observation that translation of preproinsulin can be eliminated if its initiation codon is duplicated upstream and out of frame with respect to the authentic initiation codon (8). This inhibition is only observed if there are no intervening termination codons in frame with the 5' proximal initiation codon, suggesting that eukaryotic ribosomes can reinitiate at downstream AUG codons. Based on this work, we constructed a rearrangement test gene that yields a translatable message only following a proper V–J joint. We demonstrate that the unrearranged gene is not expressed, but that it can be efficiently rearranged and expressed after introduction into a pre-B-cell line.

MATERIALS AND METHODS

Plasmid Construction. The rearrangement test gene pHRD (Fig. 1) was assembled from two basic units: (i) the *Escherichia coli gpt* coding region under transcriptional control of the mouse metallothionein 1 promoter (P_{MT}) and mouse

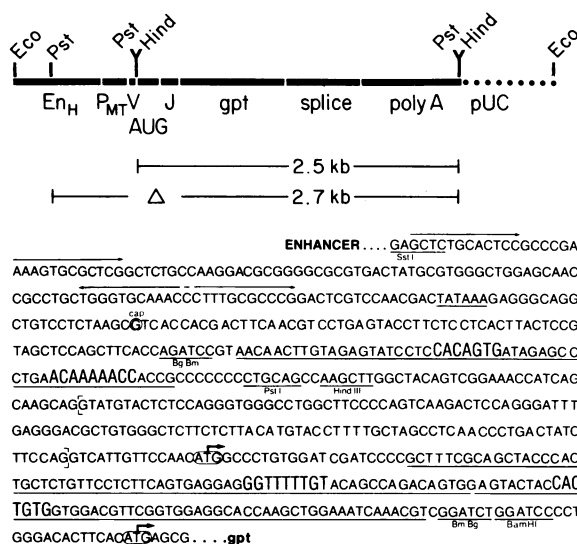


FIG. 1. Partial restriction site map (Upper) and DNA sequence (Lower) of the rearrangement test gene pHRD. Overhead arrows indicate the metal regulatory elements as defined in ref. 9. The "TATA" box is underlined and the 5' end of the mRNA is indicated (cap). Sequences derived from immunoglobulin genes are underlined and the heptamer/nonamer signal sequences are in larger type. Brackets mark an intron in the preproinsulin segment. Selected restriction sites are indicated (Bg Bm denotes a *Bgl* II/*Bam*HI junction) and the preproinsulin and *gpt* initiation codons are highlighted with a circle and an arrow.

immunoglobulin heavy-chain enhancer (En_H), and (ii) an insert consisting of a portion of the rat preproinsulin gene (including its initiation codon) flanked on the 5' side by a V_κ heptamer/nonamer and on the 3' side by a J_κ nonamer/heptamer.

The En_H - P_{MT} -*gpt* unit was itself assembled from several elements. En_H is the 683-base-pair (bp) *Xba* I/*Eco*RI fragment, in reverse orientation, from the J_H - C_μ (constant region μ chain) intron (10). This fragment is linked to P_{MT} by a section of the pUC13 polylinker (11). P_{MT} is the 210-bp *Sst* I/*Bgl* II fragment from pMK (ref. 12; a gift of R. Palmiter, University of Washington, Seattle). The promoter is joined to the *gpt* coding region by an AGATCTGGATCCCC (*Bgl* II/*Bam*HI) sequence derived from a modified pUC polylinker. The *gpt* gene from pSV2*gpt* (13) was altered by deleting to position 186 (i.e., 13 bp 5' of the *gpt* initiation codon; see ref. 14 for numbering scheme) using exonuclease III and S1 nuclease. The remainder of the *gpt* transcription unit, to the *Pst* I site beyond the simian virus 40 (SV40) polyadenylation signal, is identical to pSV2*gpt*. This construct, carried in pUC12, is designated pHMg186 and has been useful for transforming a variety of lymphoid cells to mycophenolic acid resistance.

The insert of the preproinsulin initiation codon flanked by V and J recognition sequences was also assembled from several elements. The V heptamer/nonamer is contained in the 65-bp *Hinc*II/*Pst* I fragment of pES320 (15) cloned into pUC8, thus flanking this piece with *Bam*HI and *Hind*III sites. The *Hind*III site was used to join the V sequences to the preproinsulin 178-bp *Hind*III/*Bam*HI fragment from p255/11 (ref. 16; generously provided by M. Kozak, University of Pittsburgh). The *Bam*HI site on the 3' side of the insulin piece was used for ligation to a pUC derivative carrying the $J_{\kappa 1}$ sequences and was later deleted by filling in using Klenow enzyme. A *Bam*HI linker was placed 40 bp 3' of the $J_{\kappa 1}$ heptamer after treatment with exonuclease III and S1 nuclease, allowing the insert to be excised as a 389-bp *Bam*HI

fragment. This was inserted into the unique *Bgl* II site between P_{MT} and *gpt* of pHMg186 to yield pHRD.

Cell Lines and Transfection. BHK cells (17) were obtained from R. Palmiter. The J558L myeloma (18) was provided by S. Morrison (Columbia University, New York), S194 (S194/5.XXO.BU.1; ref. 19) was from ATTC, and X63-Ag8.653 (20) was from J. Kearney (University of Alabama, Birmingham). The pre-B-cell line 38B9tk⁻ (5) was a gift of F. Alt (Columbia University, New York). All cells were maintained in Dulbecco's modified Eagle's medium or RPMI 1640 medium supplemented with 10% calf serum and antibiotics. The pre-B-cell medium was also supplemented with 50 μ M 2-mercaptoethanol.

BHK cells were transfected by the calcium phosphate method (21). Approximately 10^6 cells were exposed to a precipitate containing 10 μ g of plasmid DNA. Zn induction was performed using 100 μ M ZnSO₄.

Lymphoid cells were transfected by electroporation (22). Five million cells in 0.5 ml of phosphate-buffered saline containing 2 μ g of *Eco*RI-linearized pKOneo [a SV40 promoter-Tn5 neo plasmid from D. Hanahan (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY)] and a 5-fold molar excess of *Eco*RI-linearized pHRD were held in a plastic cuvette with aluminum foil electrodes. After chilling to 0°C, the cells were exposed to a single pulse generated by an ISCO 494 power supply set at 2000 V. Cells were cultured in bulk overnight and then split into microwells and selected for neo expression in G418 (Geneticin, GIBCO; 1 mg/ml). Only wells containing single colonies were chosen for expansion. The colonies were transferred in a vol of 25 μ l after microscopic examination to preclude contamination by other clones.

Assay for *gpt* Expression. Short-term expression of *gpt* was assayed by measuring the extent of conversion of [¹⁴C]xanthine to xanthosine monophosphate using thin-layer chromatography according to the method of Chu and Berg (23).

For selection of stable BHK cell clones expressing *gpt*, transfected cells were grown in medium containing mycophenolic acid (25 μ g/ml) plus xanthine (250 μ g/ml) (13). After 10–14 days, colonies were stained with Giemsa for visualization.

Southern Blotting. DNA was prepared from single G418-resistant 38B9tk⁻ pre-B-cell and myeloma clones 2–3 weeks after transfection. DNA (30 μ g) was cut with *Pst* I, run on 0.8% agarose gels, and transferred to nitrocellulose by standard procedures (24). Filters were hybridized to [³²P]-RNA probes corresponding to the *gpt*, preproinsulin, or P_{MT} sequences contained in pHRD. Labeled probe was removed from nitrocellulose filters by incubation in 100 mM NaOH at 20°C for 5 min before reusing filters for subsequent hybridization.

Plasmid Rescue and Sequencing. The rearranged plasmids were isolated from total cell DNA by direct plasmid rescue (25). Pre-B-cell transfectant DNA (2 μ g) was cut with *Sst* I or *Xba* I (these enzymes both cut pHRD at a single site between En_H and P_{MT}) and ligated in a vol of 200 μ l. *E. coli* was transformed with this mixture and selected for ampicillin resistance. Sequencing of both plasmid strands was performed by the chain-termination method using synthetic 21-mer primers complementary to P_{MT} and to *gpt* sequences.

RESULTS

The Barrier Initiation Codon Effectively Prevents Expression of *gpt*. To determine the effect of the barrier initiation codon on expression of *gpt*, BHK cells were transfected with pHRD and several related constructs. Fibroblasts were chosen for this test because large numbers of transfectants could be easily generated, allowing a more accurate assessment of efficiency of expression.

Using a short-term enzyme assay, mock-transfected cells did not show detectable activity (Fig. 2 *Upper*). Cells transfected with a SV40 promoter-gpt plasmid (SV-gpt; pSV2gpt, ref. 13) produced a substantial amount of gpt, and this amount was slightly increased when the cells were cultured in Zn^{2+} prior to the assay. The basis for this effect is not understood. The P_{MT} derivative (MT-gpt) also directed synthesis of substantial amounts of gpt but only when the cells were Zn^{2+} induced. This is entirely consistent with previous reports (26), although the magnitude of the induction appears higher than in other systems. As expected for fibroblasts, the addition of an immunoglobulin enhancer (En-MT-gpt) had little effect. When the insert containing the barrier initiation codon was interposed between the promoter and gpt (MT-V-AUG-J-gpt), enzyme expression was eliminated. Again, as expected, the enhancer (En-MT-V-AUG-J-gpt) had no effect.

In addition to studying the behavior of a cell population, it was important to determine whether the translational barrier would inhibit the establishment of stable colonies after selection for expression of gpt. The same experiment was performed, but instead of a transient enzyme assay the cells were selected in mycophenolic acid plus xanthine (Fig. 2 *Lower*). The results of the selection experiment paralleled the transient assay results. Most significantly, no colonies could be detected in populations transfected with plasmids containing the preproinsulin insert even when they were Zn^{2+} induced.

The Recombination Indicator Does Not Rearrange in Fibroblasts or Myeloma Cells. A specific V-J rearrangement should delete the V and J heptamers and the sequences between them, including the preproinsulin sequences, and result in gpt expression. Since no mycophenolic acid-resistant BHK cells were detected, this indicates that rearrangement, if it occurs at all in these cells, is very infrequent.

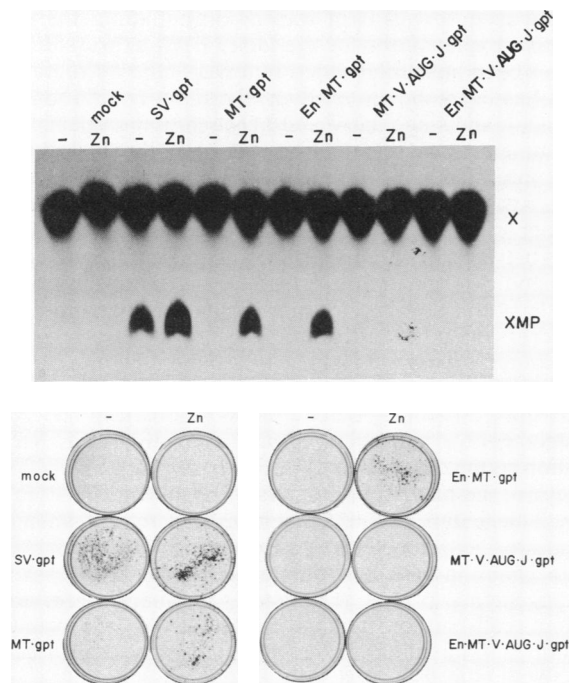


FIG. 2. Detection of gpt enzyme activity in transfected BHK cells by a short-term expression assay (*Upper*) or by selection of stably expressing clones (*Lower*). The general structure of the plasmids is indicated [SV-gpt is pSV2gpt (13), En-MT-V-AUG-J-gpt is pHRD, and other plasmids are all relatives of pHRD]. Zn denotes induction in $ZnSO_4$. The positions of xanthine (X) and xanthosine monophosphate (XMP) are indicated.

The test gene pHRD was also introduced into the myelomas J558L, S194, and X63-Ag8.653 by coelectroporation with an unlinked SVneo plasmid. After selection in G418, 8 J558L, 12 S194, and 15 X63-Ag8.653 clones were analyzed by Southern blotting (data not shown). There was no evidence of a specific rearrangement in any of the 35 pHRD-containing myelomas.

The Recombination Indicator Rearranges Very Frequently When Introduced into a Pre-B-Cell Line. The rearrangement test gene pHRD was introduced into the pre-B-cell line 38B9tk⁻ by an unlinked cotransfection with pKOneo. The cells were selected only for neo expression by growth in G418. Of 36 G418-resistant colonies derived from five independent electroporations, 23 had also incorporated pHRD as determined by hybridization using a gpt probe. No selection for gpt expression was used nor was the set of gpt-positive transfectants culled in any way prior to analysis.

DNA from these clones was cut with *Pst* I, blotted onto nitrocellulose, and hybridized with a gpt probe (Fig. 3 *Top*).

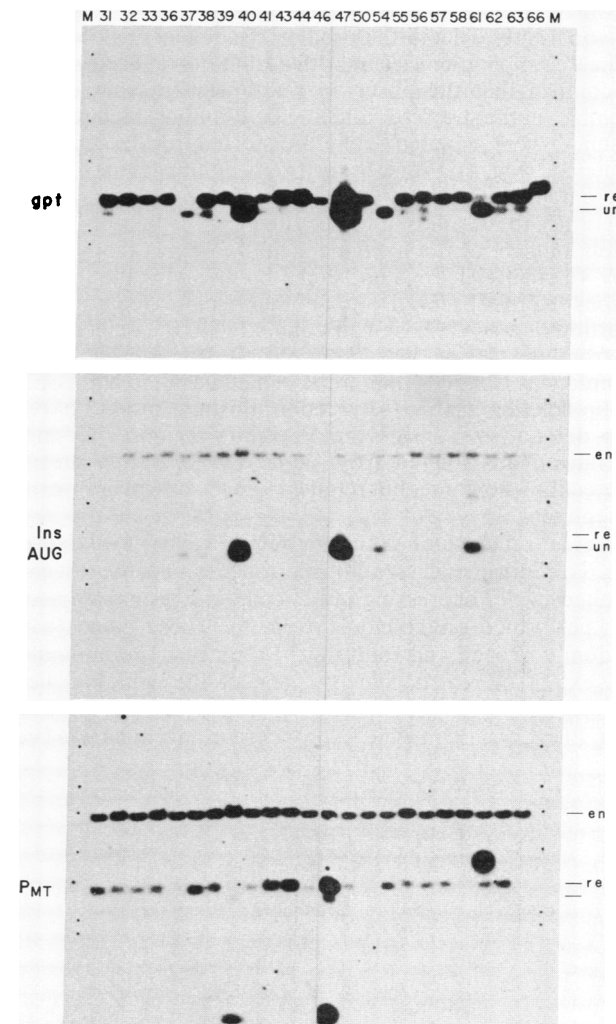


FIG. 3. Southern blot analysis of pHRD rearrangements in transfected pre-B cells. The *Pst* I-cleaved DNA, after transfer to nitrocellulose, was hybridized with ³²P-labeled gpt (*Top*), preproinsulin (Ins AUG) (*Middle*), or P_{MT} (*Bottom*) sequences. The positions of the *Pst* I fragments corresponding to the rearranged (re) or unrearranged (un) plasmid are indicated, as are the positions of the endogenous (en) preproinsulin or metallothionein genes. The P_{MT} -containing fragment of the unrearranged pHRD has run off the gel. Marker (M) lanes contained *Hind*III-cleaved λ DNA: 23.1, 9.4, 6.7, 4.4, 2.3, and 2.0 kb. Numbers above lanes correspond to clone numbers.

sion. Certain clones (e.g., nos. 32 and 33) show complete or near 100% rearrangement, whereas others (e.g., no. 31) have evidence of only fractional rearrangement. Some of the lines with such partial rearrangement have been shown to have only a single integrated copy (results not shown). Although we cannot absolutely rule out the possibility of mixed clones, these data suggest that, at least in certain cases, integration involved an unrearranged copy and that rearrangement is an ongoing process. In the clones showing 100% rearrangement, it is not possible to determine whether the plasmid rearranged prior to integration or recombination occurred after insertion into the chromosome.

In another system, high-frequency rearrangement was also observed (5, 6). However, selection for expression of an adjacent thymidine kinase gene was required. In our experiments, in contrast, no selection for expression was necessary, yet rearrangement was very frequent. It is not clear if this is a function of the promoter, the enhancer, or a combination of the two. The modular design of this construct should allow the relative contribution of these various elements to be determined.

The DNA sequences of the six V-J junctions reveal that the rearrangement was quite precise. In two cases, not a single base pair was lost; in the others, a net deletion of 2 or 3 bp was observed. One of those joints (clone 47, copy 3) had an insertion of a single base pair at the point of recombination. Additions of noncoded nucleotides (N regions) are common in heavy-chain genes (31) but are not generally seen in light-chain joints. The two joints isolated from low copy number transfectants are unusual. In one (clone 38), it appears that a deletion occurred 2 bp away from the point of recombination; perhaps this results from error-prone repair during rearrangement. In another (clone 63), it seems that the cleavage occurred 1 bp inside the V heptamer. However, both of these joints could also be explained as N-region addition. It is interesting to note that, using the same pre-B-cell line, heavy-chain or T-cell receptor gene recombination substrates sustained longer deletions of 5–10 bp (5–7), while κ gene segments introduced into a different line were rearranged with shorter deletions (3, 4). These transfection results parallel the situation with endogenous heavy- and light-chain genes (32, 33). This suggests that some subtle feature of the substrate's structure might influence the precision of the joining reaction.

The data provide formal proof that κ genes can rearrange by deletion. It had previously been shown both for transfected substrates (3, 4) and for endogenous genes (34, 35) that κ segments can rearrange by inversion. *H* genes, on the other hand, appear generally to rearrange by deletion (36), although one case of inversion has been described (37). The latter occurred in the 38B9tk⁻ cell line, which was used in our experiments of κ rearrangement by deletion. Thus, all the data taken together support the idea (34) that whether rearrangement occurs by a deletional or an inversional mechanism simply depends on the orientation of the joining DNA segments.

Only minimal sequences are required for this process: 21 bp 5' of the V heptamer and 4 bp 3' of its nonamer; <45 bp 5' of the J nonamer and 3' of its heptamer were present in our construct. In contrast, in previous transfection experiments several kilobases of DNA associated with V or D and J were included in the rearrangement constructs (3–7). Our findings very strongly support the original idea that the heptamer-spacer-nonamer are the essential recognition sequences for rearrangement (38, 39). It remains to be seen, for example by studying transgenic mice, whether these minimal sequences allow regulated rearrangement in the whole animal or whether additional target sequences for proper control of rearrangement may be required.

We are grateful to our colleagues, who are mentioned in the text, for freely sharing cell lines and plasmids. This work was supported by National Institute of Health Grants AI 24780 and HD 23089. P.E. was supported by Predoctoral Training Grants GM 07270 and CA 09537.

- Dreyer, W. J. & Bennett, J. C. (1965) *Proc. Natl. Acad. Sci. USA* **54**, 864–869.
- Tonegawa, S. (1983) *Nature (London)* **302**, 875–881.
- Lewis, S., Gifford, A. & Baltimore, D. (1984) *Nature (London)* **308**, 425–428.
- Lewis, S., Gifford, A. & Baltimore, D. (1985) *Science* **228**, 677–685.
- Blackwell, T. K. & Alt, F. W. (1984) *Cell* **37**, 105–112.
- Blackwell, T. K., Moore, M. W., Yancopoulos, G. D., Suh, H., Lutzker, S., Selsing, E. & Alt, F. W. (1986) *Nature (London)* **324**, 585–589.
- Yancopoulos, G. D., Blackwell, T., Suh, H., Hood, L. & Alt, F. (1986) *Cell* **44**, 251–259.
- Kozak, M. (1984) *Nucleic Acids Res.* **12**, 3873–3893.
- Stuart, G. W., Searle, P. F., Chen, H. Y., Brinster, R. L. & Palmiter, R. D. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7318–7322.
- Gillies, S. D., Morrison, S., Oi, V. & Tonegawa, S. (1983) *Cell* **33**, 717–728.
- Norrander, J., Kempe, T. & Messing, J. (1983) *Gene* **26**, 101–106.
- Brinster, R. L., Chen, H. Y., Warren, R., Sarthy, A. & Palmiter, R. D. (1982) *Nature (London)* **296**, 39–42.
- Mulligan, R. C. & Berg, P. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2072–2076.
- Mulligan, R. C. & Berg, P. (1981) *Mol. Cell. Biol.* **1**, 449–459.
- Selsing, E. & Storb, U. (1981) *Cell* **25**, 47–58.
- Kozak, M. (1984) *Nature (London)* **308**, 241–246.
- Waechter, D. E. & Baserga, R. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 1106–1110.
- Oi, V. T., Morrison, S. L., Herzenberg, L. A. & Berg, P. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 825–829.
- Hyman, R., Ralph, P. & Sarkar, S. (1972) *J. Natl. Cancer Inst.* **48**, 173–184.
- Kearney, J. F., Radbruch, A., Liesegang, B. & Rajewsky, K. (1979) *J. Immunol.* **123**, 1548–1550.
- Luthman, H. & Magnusson, G. (1983) *Nucleic Acids Res.* **11**, 1295–1308.
- Tonegawa, F., Hayday, A. & Keating, A. (1986) *Mol. Cell. Biol.* **6**, 703–706.
- Chu, G. & Berg, P. (1985) *Nucleic Acids Res.* **13**, 2921–2930.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- Hanahan, D. (1985) in *DNA Cloning: A Practical Approach*, ed. Glover, D. M. (IRL, Oxford), Vol. 1, pp. 109–135.
- Durnam, D. M. & Palmiter, R. D. (1981) *J. Biol. Chem.* **256**, 5712–5716.
- Kozak, M. (1986) *Cell* **44**, 283–292.
- Kozak, M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2850–2854.
- Alt, F., Enea, V., Bothwell, A. & Baltimore, D. (1980) *Cell* **21**, 1–12.
- Ritchie, K. A., Brinster, R. L. & Storb, U. (1984) *Nature (London)* **312**, 517–520.
- Desiderio, S. V., Yancopoulos, G. D., Paskind, M., Thomas, E., Boss, M. A., Landau, N., Alt, F. W. & Baltimore, D. (1984) *Nature (London)* **311**, 752–755.
- Kurosawa, Y. & Tonegawa, S. (1982) *J. Exp. Med.* **155**, 201–218.
- Alt, F. W., Blackwell, T. K., DePinho, R. A., Reth, M. G. & Yancopoulos, G. D. (1986) *Immunol. Rev.* **89**, 5–30.
- Lewis, S., Rosenberg, N., Alt, F. & Baltimore, D. (1982) *Cell* **30**, 807–816.
- Fedderson, R. M. & Van Ness, B. G. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4793–4797.
- Alt, F. W., Yancopoulos, G. D., Blackwell, T. K., Wood, C., Thomas, E., Boss, M., Coffman, R., Rosenberg, N., Tonegawa, S. & Baltimore, D. (1984) *EMBO J.* **3**, 1209–1219.
- Alt, F. W. & Baltimore, D. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4118–4122.
- Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. (1979) *Nature (London)* **280**, 288–294.
- Max, E. E., Seidman, J. G. & Leder, P. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3450–3454.