

Cell-specific enhancers in the rat exocrine pancreas

(gene expression/transcription/chloramphenicol acetyltransferase)

ANNE M. BOULET*, CHRISTOPHER R. ERWIN, AND WILLIAM J. RUTTER

*Hormone Research Institute and Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143

Contributed by William J. Rutter, January 2, 1986

ABSTRACT Pancreatic amylase, chymotrypsin B, and trypsin I genes are specifically expressed in the exocrine pancreas. The 5'-flanking regions of these genes direct preferential expression of a linked reporter function (chloramphenicol acetyltransferase) in the pancreatic exocrine cell line AR4-2J. The sequences upstream of the amylase and chymotrypsin genes that are required for this cell specific activity possess the characteristics of transcriptional enhancers. We have mapped the regions responsible for enhancer activity by deletion analysis. Modification of specific nucleotide sequences within these regions can alter or eliminate enhancer function. Comparison of the 5'-flanking regions of nine genes expressed in the exocrine pancreas identifies a family of short related sequences. These sequences are located within the enhancer regions that we have mapped and may play a role in the regulation of the expression of pancreatic exocrine-specific genes.

During cellular differentiation, different cell types acquire the ability to stably express characteristic sets of genes. The molecular mechanisms by which this occurs are poorly understood. The results of experiments involving transfection of exogenous genes into appropriate differentiated cells suggest that genes whose expression is restricted to a specific cell type contain *cis*-acting sequences that are required for efficient transcription in those cells (1–7). These elements resemble viral enhancers: they act in conjunction with separate promoter elements and can function in either orientation at a variable distance from the transcription initiation site. Unlike viral enhancers that can act in a variety of cell types, however, the activity of these *cis*-acting sequences appears to be restricted to a particular cell type. We have proposed that differentiated cells contain *trans*-acting proteins (differentiators) that interact with the specific *cis*-acting sequences to stimulate expression of the associated genes (1).

We have used a linked reporter function, the bacterial enzyme chloramphenicol acetyltransferase (CAT) to conveniently distinguish expression of exogenous genes from that of the endogenous cognate gene in transient transfection assays. When recombinants containing the 5'-flanking regions of the insulin and chymotrypsin genes fused to the CAT coding region were transfected into various cells, high level CAT expression was observed only in (insulin-producing) endocrine B cells or (chymotrypsin-producing) pancreatic exocrine cells, respectively (1). In this report, we examine the upstream regions of the set of genes that is expressed specifically in pancreatic exocrine cells. We demonstrate that the amylase, chymotrypsin, and trypsin I genes possess upstream elements that direct preferential expression in the pancreatic exocrine cell line AR4-2J. The elements of the amylase and chymotrypsin genes, which have been mapped in detail, exhibit the properties of cell type-specific enhancers. Further, sequences within the 5'-flanking regions of

seven other pancreatic exocrine genes show homology to sequences within the amylase and chymotrypsin enhancer regions. The results suggest that the expression of specific pancreatic genes is regulated by a common mechanism involving enhancer elements.

METHODS

Plasmid Constructions. The construction of plasmid pTE1, which contains a polylinker located 600 base pairs (bp) upstream from the thymidine kinase (TK) promoter (+55 to -109 with respect to the transcription initiation site) fused to CAT coding sequences, has been described (8). Plasmid pTE0 lacks the TK promoter fragment. The construction of Amy.CAT was as follows: a derivative of pTE0 containing an *Fnu*4HI fragment from the amylase gene (+32 to -1300) inserted at the *Hind*III site by using *Hind*III linkers was digested at a unique *Xba* I site in the polylinker adjacent to the 3' end of the amylase insert. Amylase sequences, including the first AUG, were removed by exonuclease III and S1 nuclease treatment (9), and plasmids were religated after digestion with *Nru* I. Sequence analysis showed that Amy.CAT contains amylase sequences from -17 to -1300 linked to the CAT coding region. Trp I.CAT consists of a 1162-bp *Hgi*AI fragment, -10 to -1172, from the 5'-flanking region of the trypsin I gene (10) inserted at the *Hind*III site of pBR.CAT (1) by using *Hind*III linkers. The construction of Cht.CAT was described (1). This plasmid contains the chymotrypsin 5'-flanking region fragment, -3 to -711, linked to the CAT coding sequences.

pTE1 derivatives containing DNA sequences from the amylase and chymotrypsin 5'-flanking regions were constructed as follows: the 182-bp *Sac* I fragment (-93 to -275) from the chymotrypsin upstream region (11) was inserted in both orientations into the *Bgl* II site of the pTE1 polylinker by using *Bam*HI linkers. The 193-bp *Rsa* I fragment (-41 to -235) from the amylase 5'-flanking region (unpublished data) was inserted, in both orientations, at the *Hind*III site of pTE1 by using *Hind*III linkers. The amylase and chymotrypsin sequences were moved to a position directly adjacent to the TK promoter by deletion of the pBR322 sequences (digestion with *Sal* I and *Nru* I, end-repair and ligation). The chymotrypsin -93 to -275 fragment was also inserted, in both orientations at the *Bam*HI site at the 3' end of the CAT/SV40 sequences (12) by using *Bam*HI linkers and a partial digest of pTE1- Δ (*Sal*I-*Nru*I). Deletions 5' from the chymotrypsin enhancer region were generated by partial digestion of Cht.CAT with *Sac* I, treatment with *Bal*31, digestion with *Cla* I, T4 polymerase repair and blunt-end ligation. Deletions from the 3' end of the chymotrypsin -93 to -275 fragment and from the 5' and 3' ends of the amylase -41 to -235 fragment were generated by treatment with exonuclease III

and S1 nuclease (8). To fuse the 55-bp, -170 to -225, chymotrypsin fragment to the TK promoter, a 5' deletion to -225 and a 3' deletion to -170 were recombined at the *Nco* I site at -192.

The 57-bp amylase fragment, -108 to -165, was synthesized in eight overlapping single-stranded segments that were annealed, ligated, and inserted at the *Nru* I site of pTE1 adjacent to the promoter. Blocks of 4 bp were altered by replacing the two short single-stranded segments containing the target nucleotides with two synthetic segments incorporating the desired changes. The complete mutant 57mers were assembled and inserted in pTE1 as described above. In each case, the wild-type sequence was changed to the sequence CAGT. A deletion and duplication of 4 bp (-189 to -192) within the chymotrypsin 5'-flanking region were produced by digestion at the unique *Nco* I site followed by T4 polymerase "fill-in" or S1 nuclease trimming and ligation.

Cell Transfection and Measurement of CAT Activity. Cells were transfected with plasmid DNA purified by two cycles of banding in CsCl gradients by using the calcium phosphate coprecipitation technique (13). Preparation of cell extracts and CAT enzyme assays were as described (1).

RESULTS

Amylase Gene 5'-Flanking Region Directs Preferential Expression in AR4-2J Cells. A transfectable cell line that exhibits characteristics of the differentiated state was required to test the activity of the upstream regions of pancreatic exocrine genes. We screened a number of pancreatic tumor lines and found that the AR4-2J cell line, which was derived from an azaserine-induced tumor of the rat exocrine pancreas (14), displayed a relatively high level of expression of several pancreatic gene products. RNA dot-blot analysis shows that the mRNAs for amylase, chymotrypsin, and trypsin are roughly 10.5, 1.5, and 0.078%, respectively, of the poly(A)⁺ RNA in AR4-2J cells. This represents 42, 18, and 0.9%, respectively, of the amounts present in adult rat pancreas, and 68, 33, and 1.5% of the amounts present in the differentiated embryonic pancreas (15), respectively.

A DNA fragment from the amylase 5'-flanking region containing the Goldberg-Hogness box consensus sequence and extending upstream approximately 1300 bp was fused to the CAT gene coding sequences. CAT activity directed by the amylase.CAT recombinant was measured after transfection into AR4-2J cells and rat fibroblasts. In these experiments CAT activity driven by amylase upstream sequences was normalized to the activity directed by the Rous sarcoma virus promoter (RSV.CAT, ref. 16) to control for differences in transfection efficiency. The ratio of CAT activities directed by a herpes simplex virus TK promoter-CAT fusion (pTE1) (17), and RSV.CAT is nearly identical in rat fibroblasts and AR4-2J cells (Table 1). Thus, we conclude that the activity of these promoters is relatively nonspecific with respect to cell type. In contrast to the TK.CAT construct, the amylase.CAT construct displays 75-fold higher relative activity in AR4-2J cells as compared to rat fibroblast cells (Table 1). The magnitude of this selective effect is similar to that observed with the chymotrypsin.CAT recombinant (1).

In similar experiments using a trypsin I.CAT fusion, at least a 15-fold higher level of CAT activity was observed in AR4-2J cells relative to fibroblasts.

The Amylase and Chymotrypsin 5'-Flanking Regions Contain Cell Type-Specific Enhancers. DNA fragments from the 5'-flanking regions of the pancreatic amylase (unpublished data) and chymotrypsin genes (11) were tested for the ability to enhance transcription from the heterologous herpes simplex virus TK promoter in AR4-2J cells. A 182-bp *Sac* I fragment from the chymotrypsin 5'-flanking region (-93 to -275 bp) and a 193-bp *Rsa* I fragment from the amylase

Table 1. Relative activity of 5'-flanking regions

Gene fragment	Relative CAT activity			Cell-type specificity AR4-2J vs. XC
	RAT2	XC	AR4-2J	
RSV (RSV.CAT)	100	100	100	—
TK (pTE1)	5	4	3	0.8
Chymotrypsin (Cht.CAT)	0.4	0.3	33	110
Amylase (Amy.CAT)	0.2	0.2	15	75
Trypsin I (TrpI.CAT)	0.2	0.2	3	15

Relative CAT activity directed by the 5'-flanking regions of pancreatic exocrine genes. The RSV.CAT plasmid consists of 524 bp of the Rous sarcoma virus (RSV) 3' long terminal repeat inserted at the *Hind*III site of pSVO.CAT (16). For each cell type, CAT activity is expressed as a percentage of that directed by the RSV fragment. AR4-2J is a rat pancreatic exocrine cell line, and XC and RAT2 are rat fibroblast cell lines. Cell type specificity represents the ratio of relative CAT activity obtained in AR4-2J cells to that in XC cells.

5'-flanking region (-41 to -235 bp) were inserted upstream from the TK promoter via the polylinker in pTE1.

Plasmids containing the amylase or chymotrypsin fragment fused directly to the TK promoter sequences (110 bp from the TK cap site) elicited a much higher level of CAT activity in AR4-2J cells than that obtained with the parent plasmid pTE1 (Fig. 1). Enhancement of CAT activity is approximately the same for the amylase and chymotrypsin fragments (44- and 41-fold, respectively).

When the distance between the chymotrypsin or amylase fragment and the TK promoter is increased by introduction of 600 bp of pBR322 DNA sequences, the effect on CAT activity is still observed although the levels of enhancement are reduced by factors of 11 (amylase) and 3 (chymotrypsin) (Fig. 1). The chymotrypsin fragment also enhanced expression (4-fold) when placed 1700 bp downstream of the transcription start site at the 3' end of the CAT gene in pTE1 (Fig. 1).

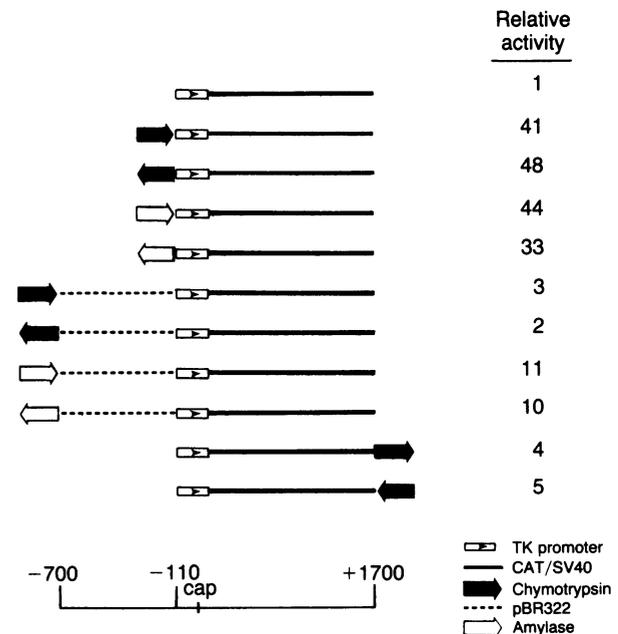


Fig. 1. 5'-Flanking chymotrypsin (-93 to -275) and amylase (-41 to -235) fragments drive specific expression of CAT directed by the TK promoter. The distance separating the amylase and chymotrypsin fragments from the TK promoter cap site is indicated by the scale at the bottom. Fragment orientations are indicated by arrows, and symbols are defined in the legend at the lower right. CAT activities are expressed relative to that of pTE1.

At each location tested, the amylase and chymotrypsin fragments enhanced CAT activity when placed in an inverted, or anti-sense, orientation relative to the direction of transcription from the TK promoter in pTE1. In each case, activity was evident only in AR4-2J cells: the sequences did not enhance expression from the TK promoter in rat XC or RAT2 fibroblasts (data not shown).

When the TK promoter fragment was deleted from the constructs and the amylase and chymotrypsin fragments were juxtaposed with the CAT coding sequences, most of the activity was lost indicating that the TK promoter is required for enhancer function. Deletion of the TK promoter results in the loss of >88% of the activity. The residual CAT activity detected in the absence of the TK promoter may be a result of the recruitment of cryptic promoters located in the amylase and chymotrypsin fragments, a phenomenon that has been observed in other systems (18).

Analysis of the RNA isolated from transfected cells confirms that enhancement of CAT activity is due to an increase in transcription from the TK promoter. RNA transcripts initiating at the TK cap site are nearly undetectable in AR4-2J cells transfected with pTE1. The level of this RNA species is increased dramatically when the amylase or chymotrypsin enhancer is linked to the TK promoter in the transfecting plasmid DNA (data not shown).

Mapping the Chymotrypsin Cell Type-Specific Enhancer Region. The 5' border of the chymotrypsin enhancer was mapped by deleting increasing portions of the chymotrypsin 5'-flanking region from Cht.CAT by using either cleavage by specific restriction enzymes or *Bal31* digestion (1). The constructions were designed such that all 5'-deletion endpoints are flanked by the same pBR322 sequences. A deletion eliminating sequences upstream from -225 has no measurable effect on activity (Fig. 2A). In contrast, deletions to -192 (*Nco* I site) or -184 showed less than 10% of the activity of the -3 to -711 fragment.

To map the 3' border of the enhancer sequence, deletions from the 3' end of the chymotrypsin *Sac* I fragment were generated using exonuclease III and S1 nuclease digestion, and the resulting termini were joined directly to the TK promoter in pTE1 (8). The degree of enhancement of CAT activity decreases progressively as sequences are deleted from the 3' end of the *Sac* I fragment (Fig. 2A). The fragment from -170 to -275 retains only 17% activity but still exhibits 8-fold greater CAT activity than the TK promoter in AR4-2J cells. Further removal of 50 bp from the 5' end, yielding a 55-bp fragment (-170 to -225) (Fig. 2A), does not further decrease the activity. This 55-bp fragment exhibits 27% of the activity of the -93 to -275 fragment, representing at least 10-fold enhancement of CAT activity relative to pTE1 (Fig. 2A). The sequences upstream of -225 do not contribute to enhancer activity in these assays, and there is not any indication that these sequences complement the loss of sequences from the 3' end of the enhancer fragment.

Mapping the Amylase Cell Type-Specific Enhancer Region. Progressive deletion of the sequences from the 5' and 3' ends of the -41 to -235 amylase fragment was carried out using exonuclease III and S1 nuclease. While deletion of 80 bp from the 5' end yielded a fragment (-41 to -154) that displayed increased activity, further deletion of 14 bp (-41 to -140 fragment) virtually eliminated enhancer activity (Fig. 2B). Deletion of sequences from the 3' end of the fragment resulted in a sharp loss of enhancer activity. Removal of 36 bp from the 3' end produced a fragment (-77 to -235) that displays 32% the activity of the -41 to -235 fragment. This level of activity is also seen with the -115 to -235 fragment. However, further deletion of 44 bp (-159 to -235 fragment) results in the loss of virtually all activity. Thus, sequences between -115 and -154 appear to be essential for enhancement of expression from the TK promoter in AR4-2J cells.

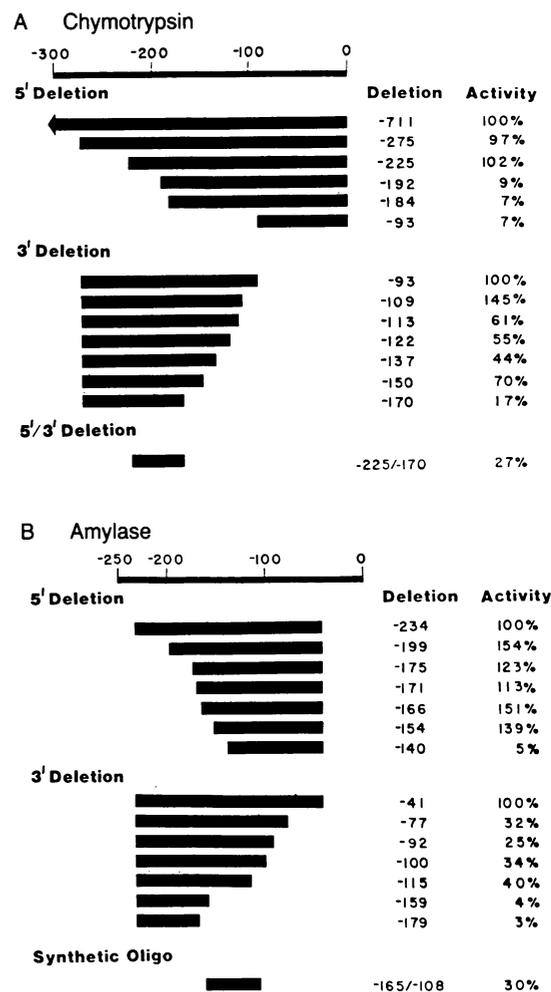


FIG. 2. Effects of progressive 5' and 3' deletions on the ability of the amylase and chymotrypsin enhancer regions to enhance CAT activity directed by the TK promoter. (A) 5' deletions from the chymotrypsin enhancer region when derived from Cht.CAT contain the chymotrypsin promoter. Generation of plasmids with deletions to -275, -192, and -93 was described (1), and activities obtained with these deletions are included here for comparison. Deletions from the 3' end of the chymotrypsin -93 to -275 fragment were generated from the pTE1 derivative. (B) Deletions from the 5' and 3' ends of the amylase -41 to -235 fragment were generated from the pTE1 derivatives. The synthetic oligonucleotide corresponds to the -108 to -165 region of the amylase gene. Activities are expressed as a percentage of that obtained with the largest fragment tested in the appropriate orientation.

Internal Mutations in the Amylase and Chymotrypsin Enhancer Regions. Comparison of the DNA sequences within the amylase and chymotrypsin enhancer regions led to identification of four noncontiguous regions of homology (Fig. 3). We have altered sequences within these regions of homology. Initially, a 57-bp sequence (-108 to -165) from the amylase 5'-flanking region, containing the sequences required for cell type-specific enhancer function, as predicted by deletion analysis, was synthesized. When inserted directly upstream of the promoter, this fragment displayed about 30% of the enhancer activity of the complete (-41 to -235) amylase fragment. Blocks of four nucleotides were altered by substituting the sequence CAGT within each of the homologous regions at the positions shown in Fig. 3. The mutant 57-bp fragments were tested in both orientations for cell specific enhancer activity. In each case, the 4-bp alteration had a pronounced effect on enhancer activity. Mutations in regions I and IV produced increased enhancer activity relative to wild-type when the fragment was inserted in the

differentiated cells. Although we have not mapped the trypsin regulatory element, we presume that it will show properties similar to those of the chymotrypsin and amylase enhancers. This behavior is also characteristic of the enhancer for the rat insulin gene (8) as well as those in the immunoglobulin genes (4) and the mouse E_B major histocompatibility genes (6). The present experiments do not address the possibility that other regulatory elements may exist, for example, in the promoter region (8).

The -93 to -275 fragment from the chymotrypsin gene and -41 to -235 fragment from the amylase 5'-flanking region are fully competent to drive cell-specific expression from the herpes simplex virus TK promoter in this assay system. The enhancers sequence in the insulin 5'-flanking region lies between -103 and -249 relative to the cap site (8). Thus the control sequences of each of these three pancreatic genes are located in similar positions relative to the respective transcription initiation sites.

A deletion analysis of both the amylase and chymotrypsin enhancers indicates that they contain sequences that are essential for enhancer function and other regions that are not essential but contribute to full enhancer activity. For example, removal of amylase upstream sequences from -41 to -77 reduces activity, while removal of sequences upstream from -115 or downstream from -154 causes a complete loss of activity. The requirement for specific sequences within the enhancer region is demonstrated by the sharp decreases or alterations in activity resulting from some of the internal mutations. A 4-bp deletion within the chymotrypsin enhancer region eliminates activity. A series of 4-bp substitutions in the amylase enhancer region shows variable effects on enhancer function; some virtually eliminate activity, while others have little effect. Another class of modifications changes the symmetry of the enhancer effect. In one case a 4-bp alteration (region I, Fig. 3) eliminates activity of the enhancer in one orientation but does not affect activity in the other orientation. Thus the ability of an enhancer to function bidirectionally appears to be sequence dependent. Enhancers have been reported to be more active in one orientation than the other under certain circumstances (8, 19, 20). Since the wild-type and mutant fragments were tested in an identical sequence environment, the effects observed may be due to changes within the enhancer region itself.

We have reported that the upstream sequences of selectively expressed genes (insulin, somatostatin) are conserved in various species (21, 22). The hypothesis that conserved sequences are involved in specific expression has to some extent been borne out by the functional studies on the insulin gene (1, 8). Sequence analysis of the 5'-flanking region of the set of pancreatic exocrine genes has failed to reveal striking overall sequence relationships although short sequence homologies can be identified (10, 23). We have compared the sequences in the mapped control regions of the amylase and chymotrypsin genes with the 5'-flanking regions of seven other exocrine pancreatic genes and have found the related sequences presented in Table 2. These sequences overlap with the sequences identified by Swift and co-workers (24) and Craik *et al.* (10). The homologous sequences in the amylase and chymotrypsin genes include the similarities pointed out in Fig. 3 and lie within the regions critical for enhancer function. Furthermore, alteration of the nucleotides in the consensus sequence region dramatically reduce enhancer function. The homologous regions of the other exocrine genes are located between -89 and -240 relative to the transcription start sites, a region expected to contain cell-specific control elements by analogy with the insulin, amylase, and chymotrypsin genes. The consensus sequence has not been found in similar regions of other endocrine genes or a number of nonpancreatic genes that have been exam-

ined. The elastase I gene consensus sequence lies within the fragment that directs expression of a linked human growth hormone coding region in pancreatic exocrine cells of transgenic mice (24).

We have proposed that the activity of *cis*-acting elements in pancreatic genes is mediated by *trans*-acting factors, which we have termed differentiators to suggest a possible role in differentiation. The pancreatic exocrine consensus sequence could represent a recognition site for a single *trans*-acting factor or several closely related factors. The expression profiles of the various specific genes during pancreatic development are remarkably divergent (15), indicating that these genes are differentially regulated. This aspect of pancreatic gene expression is most readily explained by a family of distinct regulatory molecules that interact with the various specific genes. However, differential control could also be explained by a pancreas-specific factor whose activity is regulated during the course of development by chromatin structure or some other factor limiting for expression. Decisive experiments on mechanism will be possible when the *trans*-acting molecules and their genes are identified.

We acknowledge the many constructive suggestions of Dr. M. Walker and our other colleagues in the Rutter laboratory. We thank Drs. C. Gorman and T. Edlund for gifts of plasmids, Dr. C. Craik and J. Barnett for oligonucleotides, J. Turner for technical assistance, and L. Spector for preparation of the manuscript. This work was supported in part by Grants AM21344 and GM28520 from the National Institutes of Health and a grant from the March of Dimes (to W.J.R.). C.R.E. was the recipient of the American Cancer Society Fellowship PF2221.

- Walker, M. D., Edlund, T., Boulet, A. M. & Rutter, W. J. (1983) *Nature (London)* **306**, 557-561.
- Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. (1983) *Cell* **33**, 717-728.
- Banerji, J., Rusconi, S. & Schaffner, W. (1981) *Cell* **27**, 299-308.
- Queen, C. & Baltimore, D. (1983) *Cell* **33**, 741-748.
- Neuberger, M. S. (1983) *EMBO J.* **2**, 1373-1378.
- Gillies, S. D., Folsom, V. & Tonegawa, S. (1984) *Nature (London)* **310**, 594-597.
- Picard, D. & Schaffner, W. (1984) *Nature (London)* **307**, 80-82.
- Edlund, T., Walker, M. D., Barr, P. J. & Rutter, W. J. (1985) *Science* **230**, 912-916.
- Sakonju, S., Bogenhagen, D. F. & Brown, D. D. (1980) *Cell* **19**, 13-25.
- Craik, C. S., Choo, Q.-L., Swift, G. H., Quinto, C., MacDonald, R. J. & Rutter, W. J. (1984) *J. Biol. Chem.* **259**, 14255-14264.
- Bell, G. I., Quinto, C., Quiroga, M., Valenzuela, P., Craik, C. S. & Rutter, W. J. (1984) *J. Biol. Chem.* **259**, 14263-14270.
- Gorman, C. M., Moffat, L. F. & Howard, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044-1051.
- Graham, F. L. & van der Eb, A. V. (1973) *Virology* **52**, 456-467.
- Jessop, N. W. & Hay, R. J. (1980) *In Vitro* **16**, 212 (abstr.).
- Han, J., Rall, L. & Rutter, W. J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 110-114.
- Gorman, C., Merlino, G. T., Willingham, M. C., Pastan, I. & Howard, B. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6777-6781.
- McKnight, S. L. (1980) *Nucleic Acids Res.* **8**, 5949-5964.
- Wasylyk, B., Wasylyk, C., Augereau, P. & Chambon, P. (1983) *Cell* **32**, 503-514.
- Herbomel, P., Bourachot, B. & Yaniv, M. (1984) *Cell* **39**, 653-662.
- Fromm, P. & Berg, P. (1983) *Mol. Cell. Biol.* **3**, 991-999.
- Bell, G. I., Swain, W. F., Pictet, R., Cordell, B., Goodman, H. M. & Rutter, W. J. (1979) *Nature (London)* **282**, 525-527.
- Shen, L.-P. & Rutter, W. J. (1984) *Science* **224**, 168-171.
- Swift, G. H., Craik, C. S., Stary, J., Quinto, C., Lahaie, R. G., Rutter, W. J. & MacDonald, R. J. (1984) *J. Biol. Chem.* **259**, 14271-14278.
- Ornitz, D. M., Palmiter, R. D., Hammer, R. E., Brinster, R. L., Swift, G. H. & MacDonald, R. J. (1985) *Nature (London)* **313**, 600-602.