The 5' splice site: phylogenetic evolution and variable geometry of association with U1RNA

Monique Jacob and Hélène Gallinaro

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, 11 rue Humann, 67085 Strasbourg Cédex, France

Received January 3, 1989; Accepted February 17, 1989

ABSTRACT

The 5' splice site sequences of 3294 introns from various organisms (1-672) were analyzed in order to determine the rules governing evolution of this sequence, which may shed light on the mechanism of cleavage at the exon-intron junction. The data indicate that, currently, in all organisms, a common sequence GUAAGU and its derivatives are used as well as an additional sequence and its derivatives, which differ, in metazoa (G/GUGAGU), lower eucaryotes (GUAxGU) and higher plants (AG/GUA). They all partly resemble the prototype sequence AG/GUAAGU whose 8 contiguous nucleotides are complementary to the nucleotides 4-11 of U1RNA, which are perfectly conserved in the course of phylogenetic evolution. Detailed examination of the data shows that U1RNA can recognize different parts of 5' splice sites. As a rule, either prototype nucleotides at position -2 and -1 or at positions 4, 5 or 6 or at positions 3-4 are dispensable provided that the stability of the U1RNA-5' splice site hybrid is conserved. On the basis of frequency of sequences, the optimal size of the hybridizable region is 5-7 nucleotides. Thus, the cleavage at the exon-intron junction seems to imply, first, that the 5' splice site is recognized by U1RNA according to a "variable geometry" program; second, that the precise cleavage site is determined by the conserved sequence of U1RNA since it occurs exactly opposite to the junction between nucleotides C9 and C10 of U1RNA. The variable geometry of the U1RNA-5' splice site association provides flexibility to the system and allows diversification in the course of phylogenetic evolution.

INTRODUCTION

Soon after the discovery that eucaryotic genes are split, it was recognized that the intervening sequences had a GU at their 5' end and a AG at their 3' end (673). For the 5' end, studies of consensus sequences suggested that 3 nucleotides upstream of the GU and 4 nucleotides downstream were involved in 5' splice site recognition (673-674). On the other hand, it was shown that a sequence at the 5' end of U1RNA was complementary to the 5' splice site consensus sequence (675-676). The 5' end U1RNA sequence is not involved in intramolecular base-pairing (677 and Fig. 4) and appears as accessible in both isolated U1RNA (678) and U1snRNP (679).

Further studies showed that the 5' splice site consensus sequence was somewhat different in various organisms (589, 598, 612, 617, 672, 680, 681, 682) and that homologous introns did not have obligatorily the same 5' splice site (examples in 1-672). At the same period, the sequence of U1RNA from a variety of organisms was determined and it was shown that its 5' end was perfectly conserved (683-694). In addition, various experimental approaches indicated that the U1RNP was indeed involved in 5' splice site recognition (for reviews see 695, 696).

Among these data, 2 major facts seemed of interest for the comprehension of the mechanism of 5' splice site recognition and cleavage. First, a same U1RNA sequence is able to recognize a variety of 5' splice sites in pre-messenger RNA. Second, the 5' splice site seems to vary in the course of phylogenetic evolution. To better understand the underlying mechanisms, we analyzed the 5' splice sites of various groups of organisms and we could thus show that the rules governing these variations were different in metazoa, lower eucaryotes and higher plants. The data also permitted a detailed analysis of the U1RNA-5' splice site association. Variable geometry in the association provides a large flexibility of the system while ensuring a precise cleavage at the exon-intron junction.

MATERIAL AND METHODS

The databank (1-672) is made of a personal catalogue of sequences comprising 3294 introns from 761 genes. These genes come from various organisms as detailed in Table 1. The lists of splice site sequences published by many groups is included in the catalogue.

The following conventions were adopted for presentation of sequences. The sequence AG/GUAAGU which is fully complementary to nucleotides 4 to 11 of U1RNA is considered as the "prototype sequence" (Fig. 4). Any nucleotide conforming to this sequence at the splice junction is written in capitals while any nucleotide which does not conform is written in minuscules. For each position, x represents any nucleotide different from the prototype nucleotide (for instance, g, c or u instead of A), y represents a pyrimidine and n any nucleotide, including the prototype nucleotide. The positions are numbered from -3 to +6 and are indicated as follows : C(-3), A(-2), G(-1), G1, U2, A3, A4, G5, U6.

The data were compiled and compared manually as this has proven to be the most efficient method in our particular case.

RESULTS

Composition of the databank and group comparison

The 3294 introns from the databank come from a variety of organisms (Table 1). The number of introns per group varies from 113 (group L) to 1137 (group H). This is due predominantly to the well-known increase of the number of introns in higher organisms relative to lower eucaryotes or invertebrates and accessorially to the more sociological fact that mammalian genes (groups H and M) are the most intensively studied. Even if the same number of genes had been analyzed in mammals and in lower eucaryotes, the existence of many intronless genes in the latter group would still make it difficult to have comparable number of introns.

The difference in the number of introns raises the problem of the validity of group comparison. We considered that a comparison will be biased mostly in the case of overrepresentation of a class of homologous introns from genes derived from a same ancestor. This is not the case, neither in the small nor in the large groups. The genes analyzed code for a variety of products and homologous introns represent only a small percentage of the total. For instance, the globin genes are amongst the most frequently analyzed. Globin introns represent 2,5 %, 2,0 % and 2,2 % of the total introns of groups H, M and V, respectively. Taking into account that even within a group, the sequence of the 5' splice site is quite variable (Table 3), we infer that the conclusions that may be drawn from our study are not significantly biased by the presence of genes of the same origin.

Another problem is raised by the presence of very low frequency splice site sequences (1-2 examples). In such a case, the possibility of sequence errors cannot be dismissed. In the absence of any confirmatory data, we did not draw any conclusions from the existence of such sequences except for low frequency occurrence.

Diversity and evolution of the 6 nucleotide sequences at the 5'end of the intron

The consensus sequence at the exon-intron junction is currently assumed to be made of 9 nucleotides comprising 3 exon nucleotides and 6 intron nucleotides. We first analyzed the intronic part of the splice site as illustrated in Fig. 1.

For each group, we sorted out the sequences differing from the prototype sequence GUAAGU in the three most 5' nucleotides. The first two nucleotides GU are very conserved in agreement with the Chambon's rule (673). There are only 19 examples in which U2 is substituted (0,6 % of total, indicated as x2 in

O R G A N I S M	N U M B E R O F		N U M B E R O F		G R O U P	R E F E R E N C E S
	GENES	INTRONS	GENES	INTRONS		
Man	231	1097			H	1-210,428
Monkeys	19	40	250	1137	Primates	211-218
Mouse	124	512			M	219-337,409
Rat	69	396			Other mammals	338-406,
Hamster	9	42				408,409,412,428,446,454,455
Guinea-Pig	2	5				435, 438
Rabbit	17	68				407,410,411,415-419,430-433,
Beef	15	64				439-441,444,451
Pork	2	11				413,414,420,426,427,429,436,
Other mammals*	13	26	251	1124		437,443,445,447,449,450
						448,453
						421,425,434,442
Chicken	62	393			V	456-469,471-490,497,498,500,
Xenopus	10	83			Other	502,503,505-515,522-524,519,520
Other vertebrates*	8	16	80	492	vertebrates	470,494-496,504,516-518,525
						491-493,499,501,521
Drosophila	42	116			I	529-536,538,540-542,544-551
Bombix Mori	4	13			Invertebrates	557-564,569,570,572,580
Other insects*	4	9				565-567,571
Nematodes	9	21				537,539,555,561
Sea urchin	7	13	66	172		543,552-554,556-560
						526-528,568
Tetrahymena	2	7			L	589,619
Plasmodium	5	6			Lower eucaryotes	594,598,599,609
Acanthamoeba	2	4				582,603
Aspergillus	7	23				587,588,595,604-607,620
Neurospora	5	13				597,602,617,618
Saccharomyces	12	13				583-585,591,592,605,611-616
Dictyostelium	2	3				601,610
Physarum	5	15				581,596,622,623
Other lower eucaryotes*	11	29	51	113		586,590,593,600,608,621
Maize	9	61			P	626,627,638,656,657,670
Barley	2	10			Higher plants	628,671
Tobacco	2	11				629,655
Potato	6	31				647-651
Arabidopsis	4	14				625,632,659,660
Soybean	15	47				624,626,636,639,644-646,667-669
Other beans	2	7				643,665,666,669
Pea	6	17				640-642,654,663
Other higher plants*	17	58	63	256		630,631,633,637,652,653,658,661
						662,672
TOTAL			761	3294		1-672

TABLE 1 : The databank: composition of the groups of organisms. *: only one gene or gene type was studied in each organism. Note that the ratio intron/gene does not indicate the exact number of introns per gene since not all exon-intron junctions were sequenced in each gene.

Fig. 1). The sequences containing the prototype A3 are the most numerous in all 6 groups. The frequency of those containing g3 varies considerably according to the group. It is very low in lower eucaryotes and higher plants. g3 seems to be characteristic of metazoa (Fig. 2) and its usage is particularly amplified in mammals (groups M and H). The substitution of A3 by y3 is a rather infrequent event except in higher plants (Figs 1-2). In this group, y3 is more frequent than g3. Thus, in the course of phylogenetic evolution, metazoa may substitute A3 by g3 whereas higher plants rather substitute it by y3.

Substitutions at positions 4, 5 and/or 6 are quite frequent and more so with A3 than with g3 (Fig. 1). Several characteristics of these substitutions should be underlined : 1) Substitution of U6 is a frequent event in mammals (groups H and M), in particular when the nucleotide at position 3 is the prototype nucleotide A (Fig. 1). Inspection of the data suggests that the substitution of U6 started in early vertebrates and was amplified in mammals (Fig. 2). Remark that U is still the preferred nucleotide at position 6. If this nucleotide would be neutral, the x6/U6 ratio would be 3. 2) The substitution of G5 by x5 is infrequent in metazoa and lower eucaryotes as compared to that of the flanking nucleotides (Figs 1 and 2). 3) The substitution of A4 and of both A4 and U6 is a relatively frequent event in lower eucaryotes (Figs 1 and 2). This gives rise to the consensus sequences of yeast (GUA_AUGU), *Neurospora* (GUA_C^AGU) or *Plasmodium* (GUA_U^AGU) (612, 617, 598). 4) Higher plants are different from all other groups : substitutions of nucleotides 4, 5, 6 occur at approximately the same, relatively high frequency (Figs 1 and 2).

The data show that except for the well-conserved G1U2, the intron sequence is quite variable and that at least some of the evolutionary driven changes are non-random. As will be commented below, one of the factors regulating these changes is related to the interaction of the 5' splice site with U1RNA.

The exon nucleotides of the 5' splice site

It is often assumed that the exon part of the 5' splice site is made of 3 nucleotides (positions -3, -2, -1). The mammalian consensus exon sequence is given as $5'\text{C}_A\text{AG3}'$ and may be complementary to nucleotides 10, 11 and 12 of U1RNA: $3'\text{GUC5}'$ (Fig. 4). Obviously C(-3) is a good candidate for base-pairing to G12 but not a(-3). The distribution of nucleotide -3 in the various groups or in association with the various intron sequences was examined. The average distribution (a : 36 %, g : 17 %, C : 36 %, u : 11 %) was approximately conserved between groups whatever the nucleotides downstream (data not shown). Therefore, we did not consider that the nucleotide at position -3 was

Figure 2. Substitutions of nucleotides at position 3 to 6 in the course of phylogenetic evolution. The data are those of Fig. 1. Rate of substitution: low, white bars ; medium, stripped bars ; high, filled bars.

part of the 5' splice site sequence as far as the association with U1RNA is concerned. Its peculiar and well-conserved distribution rather suggests that it may have another important but yet undetermined function.

The 16 combinations of nucleotides -2 and -1 were classified as follows : α -type : A(-2), G(-1) ; β -type : x(-2), G(-1) ; γ -type : n(-2), x(-1). The distribution of the α , β and γ sequences found upstream of various intron sequences is shown in Figure 3.

The mammalian sequences being similarly distributed (groups H and M Figs 1 and 2) were studied together (Fig. 3A). With the prototype intron sequence, the exon nucleotides are predominantly of the γ -type (bottom left), their distribution being close to a random distribution (bottom right).

Any substitution of nucleotides 4, 5 or 6 completely modifies the picture, the α and β -type exon sequences becoming predominant. The phenomenon is particularly marked when G5 is substituted or when 2-3 substitutions occur (Fig. 3A, lower row). When g replaces A at position 3, the requirement for G(-1)

Figure 1'. Diversity of the 6 nucleotide sequences at the 5' end of the intron. For each group (as defined in Table 1), the number of prototype sequences GUAAGU was taken as 1 and the relative proportions of the other combinations was calculated. The exact number of prototype sequences is indicated at the top of the left bar in each group. The proportion of sequences where U2 was substituted are represented as a bar at the extreme right of each panel (x2). The other sequences were first classified according to nucleotide at position 3 (A3, g3, y3, above the bars). The numbers in parentheses indicate the number of introns indexed in each of these classes. Within each class, the introns were further classified according to substitutions at positions 4, 5 or 6 as shown in abscissa. x indicates the position of the nucleotide substituted and may be any of the 3 non-prototype nucleotides. White bars : no substitution at positions 4, 5, 6. Stripped bars : 1 substitution. Filled bars : 2 substitutions. Dotted bars : 3 substitutions.

Figure 3. The exon nucleotides upstream of various types of intron sequences at the exon-intron junction. The nucleotides at positions -2 and -1 are : AG(α) ; xG(β) ; nx(γ). (see text) and their relative proportion (%) is represented as bars. The number above each block of bars indicates the total number of 5' splice sites in each case. The intron sequences with 2 or 3 substitutions which are relatively infrequent but have similar distribution of α , β , γ sequences were assembled. A random distribution is shown as control at the bottom left of panel A. Panel A : the exon sequence distribution of the 2 groups of mammals are presented together. The distribution of the exon nucleotides is shown for A3, g3 and y3. Panel B : the distribution for the other groups is presented only for A3. (Because of low number of samples, it was senseless for g3 and y3, see Fig. 1). The asterisks indicate the distribution which differ from the mammalian distribution. For group L, the arrows indicate that the 2 splice sites with substitutions of G5 are of β and γ types.

increases (Fig. 3A, middle row, left). However, the intron sequence is still frequently associated with a γ -type exon sequence. This is not so when A3 is substituted by a pyrimidine (upper row, left) ; in this case, there is a strong requirement for AG in the exon. The various intron substitutions apparently provoke the same effect for g3 and y3 as for A3, except that the requirement for AG in the exon becomes more and more stringent.

Figure 4. Association of the 5' splice site of a premessenger RNA intron with the 5' extremity of U1RNA. Only the fragments of interest from U1RNA and premRNA are shown. 8 base-pairs can be formed between U1RNA and the prototype 5' splice site sequence. The arrow indicates the exon-intron junction.

The other groups (V, I, L and P) were studied in the same way (Fig. 3B). Only the sequences containing A3 are presented. The distribution of α , β , and γ sequences upstream of the various intron sequences is very close to that of mammals in groups V and I (with the possible exception of the intron sequence GUAAXU in group V which seem to be less frequently associated with an α -type sequence).

The situation is quite different in lower eucaryotes (group L). The sequence upstream of the prototype intron sequence GUAAGU is predominantly of the γ -type as in the other groups but this is also the case for the sequence GUAXGU (fourth block). Thus, the 2 most frequent intron sequences of group L (Fig. 1 and 2) do not require an AG at the 3' end of the exon. The α and β -type sequences are more frequent when other substitutions occur, but the presence of an AG upstream is not as stringent as in metazoa. The higher plants group has also its own characteristics. On average, the γ -type exon sequence is more frequent than in mammals and the α -type less frequent. This is particularly obvious when U6 is substituted (second block).

In summary, the prototype exon sequence A(-2)G(-1) is primarily found when some prototype nucleotides at positions 3-6 are substituted. This characteristic has important implications for the interaction with U1RNA.

The complementarity of the 5' splice site and of the 5' extremity of U1RNA

The variability of the 5' splice site, which is amply demonstrated in Figures 1-3, contrasts with the constancy of the sequence at the 5' extremity of U1RNA. The nucleotides of U1RNA, which are assumed to base-pair with the 5' splice site (Fig. 4), are the same whatever the origin of the small RNA (683-694). Up to now, no variant U1RNA with substitutions in the 4-12 region was detected in any cell type. As all natural introns have to be excised, a mechanism allowing the recognition of the 5' splice site and the precise

cleavage between exon and intron must exist, whatever the sequence at the junction.

To try to find a denominator common to all splice sites, we examined the extent of complementarity of the various types of 5' splice sites with nucleotide 4-11 of U1RNA. The results are schematically presented in Fig. 5.

The most striking observation is that the zone of complementarity is highly variable. Nucleotides at either the 5' extremity or the 3' extremity of the 5' splice site appear to be dispensable for binding to U1RNA. We conclude that U1RNA may recognize different sites of the exon-intron junction. This allows the large diversity of the splice junctions observed during phylogenetic evolution as well as within a group (Figs 1-3) or an organism (not shown).

As schematically shown in Fig. 5, the most frequent sequences are amongst the most stable. However, there is no direct simple relationship between the number of 5' splice site sequences and the free energy of the hybrid (data not shown). In this respect, it is interesting to remark that the prototype sequence (α A1) with 8 contiguous appairable nucleotides is less frequent than the sequences with 6 or 7 contiguous nucleotides (compare α A1 with α A2-3, β A1-2, γ A1). This suggests that 8 contiguous base-pairs are not necessary for the 5' splice site-U1RNA hybrid formation or may even be less favourable for subsequent reactions. Another remark is that substitutions at positions 3 and/or 4 are permissible, provided that G5 is conserved and that the hybrid comprises at least 4 + 1 or 3 + 2 contiguous base-pairs (compare for instance the sequences 6 and 7 to 5 and 8 in the different panels). Thus, nucleotides 3-4 are dispensable, like nucleotides -1 and -2, or nucleotides 4-5-6, provided that stable hybrids may be formed.

We now adressed the question of the possible signal for determination of the 5' cleavage site. As seen in panels A, B and C (Fig. 5), the only nucleotides common to most existing sequences are G and U at positions 1 and 2. However, U2 is not indispensable for splicing : indeed there is a small but significant number of sequences where U2 is substituted by c2 or a2 (Fig. 1, x2 ; Fig. 5 panel D, E) and such sequences allow normal splicing (697-698).

This leaves us with G1. To the best of our knowledge, there is no natural functional intron where G1 is substituted. However, mutant sequences which allow cleavage between either a c1 or a u1 and a G(-1) have been described (699).

This survey indicates that none of the 8 nucleotides of the splice junction is directly involved in the determination of the cleavage site and strongly

suggests that the signal for cleavage rather resides in the remarkably constant region at the 5' end of U1RNA. We propose that the precise cleavage of the exon-intron junction requires at least two steps :

- 1) The 5' splice site is recognized by U1RNA, not at a well-determined site, but at sites that may vary from an intron to the other. The U1RNA-5' splice site association can thus be qualified as an association with "variable geometry". In all likelihood, in each individual case, the sequence permitting the most stable hybrid between premRNA and U1RNA is chosen. However, the most frequently formed hybrids comprise only 5-7 base-pairs which suggests that there is an optimal rather than a maximal stability.
- 2) Once the hybrid is formed, the premRNA is cleaved between the nucleotides positionned exactly opposite to C9 and C10 of U1RNA, whether they are base-paired to these nucleotides or not.

Usage of the 5' splice sites in the course of phylogenetic evolution

The data of Figs 1-3 indicate that changes of the 5' splice junction sequence occur during phylogenetic evolution. To determine which sequences are preferentially used, we sorted out the most frequent sequence types ($\geq 3,5\%$ of total sequences in each group). These sequences, as well as their frequency, are listed in Table 2. In lower eucaryotes (group L), the most frequent sequences are of the γ -type : **GUAAGU** ($\gamma A1$) and **GUAxGU** ($\gamma A6$). Other relatively frequent sequences seem to be derived from these, primarily by addition of a G at position -1 ($\beta A1$ and $\beta A6$) or by substitution of U6 ($\gamma A7$). In higher plants (group P), there are also 2 preferred sequences and their derivatives : **GUAAGU** ($\gamma A1$ and $\beta A1$, $\gamma A2$) and **AG/GUA** ($\alpha A4$ and $\alpha A3$, $\alpha A8$, $\beta A7$). The 4 groups of metazoa show more diversity. The differences between the groups are quantitative rather than qualitative, a fact which can be attributed to the gradual apparition of substitutions such as those shown in Fig. 2. The 2 major sequences can be considered as **GUAAGU** ($\gamma A1$) and **G/GUgAGU** ($\beta B1$). The derivatives are more diverse and more frequent than in the other groups. Some α or β -type sequences ($\alpha A2$, $\beta A2$, $\alpha B1$, $\beta B1$, $\alpha B2$, $\beta B2$) are relatively more frequent in mammals than in other vertebrates or in invertebrates.

GUAAGU ($\gamma A1$) is common to all groups and is a candidate for an ancestral 5' splice site. At least one additional sequence and its derivatives are used in metazoa, lower eucaryotes or higher plants. This sequence specialization is likely to be accompanied by adaptative changes in the splicing machinery. This point of view is supported by the fact that nuclear extracts from yeast or from higher plants are usually not competent for the excision of

1 5 AGGUAAGU	H	M	V	I	L	P
----- αA1	1.5	2.9	3.3	1.8	•	•
----- βA1	2.5	2.8	3.5	4.5	5.3	4.7
----- γA1	<u>8.2</u>	<u>6.9</u>	<u>13.7</u>	<u>12.7</u>	<u>15.9</u>	<u>12.5</u>
----- αA2	7.3	8.5	5.7	5.0	1.8	1.2
----- βA2	4.9	4.4	4.1	3.2	•	.4
----- γA2	2.3	2.0	2.4	1.4	1.8	3.9
----- αA3	3.0	3.9	5.7	3.2	1.8	6.2
----- αA4	2.9	3.2	3.9	2.7	1.8	6.6
----- - αA8	.8	1.2	2.7	.5	1.8	7.4
----•---- αB1	4.6	3.7	1.4	2.7	•	.4
---•---- βB1	7.3	6.1	5.9	3.6	•	.8
--•---- γB1	<u>5.0</u>	<u>7.8</u>	<u>5.1</u>	<u>4.5</u>	<u>4.4</u>	2.0
----•--- αB2	8.3	7.5	6.3	3.2	•	.4
----•--- βB2	<u>6.3</u>	<u>5.4</u>	<u>3.7</u>	1.4	1.8	2.0
----- --- αA6	1.9	2.5	3.9	4.5	1.8	3.1
----- --- βA6	1.6	2.6	3.3	2.3	8.0	2.3
----- --- γA6	1.1	1.2	.6	.5	<u>15.9</u>	1.6
----- - αA7	5.4	4.3	3.9	1.4	8.0	1.2
----- - βA7	2.2	.7	2.0	.	2.7	5.1
----- - γA7	.4	.5	.2	.9	<u>5.3</u>	2.3

TABLE 2 : The most frequent 5' splice sites in the various groups of organisms. For each group, the frequency of sequence types defined as in Fig. 5 were calculated in per cent of the total number of splice sites in the group. The sequence types $\geq 3.5\%$ in at least one group were sorted out and are underlined in the table (double underlined: frequency $\geq 10\%$). The schematic representation of the splice sites is as in Fig. 5.

metazoan introns (700, 701, 702). In contrast, the HeLa cells nuclear extract which is assumed to be competent for the excision of a large diversity of introns is frequently able to achieve excision of lower eucaryotes and higher plant introns (702, 703, 704).

Rules for changes at the 5' splice site of individual genes

For correct expression of a gene, changes occurring at the 5' splice site are probably submitted to at least 2 constraints. The first is that the coding sequence (exon nucleotides) remains such that a functional protein can be made. The second constraint is that the intron must be precisely excised whatever the exon sequence. The pressure inducing changes in the exon sequences can be assumed to be independent of the splicing mechanism but may have profound repercussions on it. Consider, for example, a transformation

5' SPLICE SITES						5' SPLICE SITES						
NUMBER IN GROUP			TYPE	SEQUENCE						NUMBER IN GROUP		
H	M	V		1	2	3	4	5	6	H	M	V
A	G	G	U	A	A	G	U	A	G	A	G	U
1	2		α A 1							2	1	
2			α A 2						g		1	
1	3		α A 2						c		1	
1	1	1	α A 3						a a		1	
1			α A 3						a c		1	
		1	α A 3						u g			
1			α A 4						g a a			
2	3		α A 6						g			
1			α A 6						u			
9	2		α A 7						g c			
1			α A 7						c c			
1			α A 7						u c			
3			α B 2						g g			
3			α B 2						g c			
		1	α B 6						g g			
			α B 7						g c a			
2		1	α B 7						g c g			
		1	α B 7						g u g			
7	4		α C 6						u g			
1			α C 7						u g c			
		1	β A 6						c g			

5' SPLICE SITES						5' SPLICE SITES						
NUMBER IN GROUP			TYPE	SEQUENCE						NUMBER IN GROUP		
H	M	V		1	2	3	4	5	6	H	M	V
A	G	G	U	A	A	G	U	A	G	A	G	U
2			α A 1									c
		1	α A 3						a a			
		1	α A 4						c c c			
1	1		α A 7						g a			
1			α A 7						u c			
9	2	1	α B 1						g			
		1	α B 2						g		g	
5	2		α B 2						g		c	
1			α B 6						g g			
1			α B 7						g g c			
2			α C 1						c			
		2	α D 2						c c			
6	9		β B 1						g g			
1			β B 1						c g			
		3	β B 2						g g			a
3			β B 6						g g g			
1			γ A 1						a			
2	2		γ B 1						a g			
1			γ B 1						c u g			

TABLE 3 : Evolution of the globin gene 5' splice sites. The two introns of adult, fetal and embryonic genes from man (5-81), monkeys (211-217), mouse (241-245), rat (372), rabbit (415-419), beef (420), goat (421-424) and seal (425) were studied. Panel A: 1st intron; Panel B: 2nd intron. The splice site types are as defined in Fig. 5. Only the non-prototype nucleotides are indicated in the sequence column. The highest number of sequences usually reflects the fact that a specific globin gene was studied in closely related animal species (for instance, γ -globin genes were extensively studied in several primates) rather than a significant evolutionary feature.

which leads to a change from an α to γ -type splice site, which is only poorly complementary to U1RNA (Fig. 5, compare α A4 to γ A4 or α B4 to γ B4 etc.). Compensatory mutations in the intron part of the splice site could produce a functional 5' splice site owing to the variable geometry of the U1RNA-5' splice site association. Thus, multiple combinations of mutations should be possible in the course of gene evolution, within the limits already specified.

To test this prediction, we examined the two 5' splice sites of the globin genes. These genes which are among the most studied in vertebrates, all have the same general organization with 3 exons separated by 2 introns (Table 3, panel A for IVS1, panel B for IVS2). For both introns, we observe that, in α -type sequences having an A at position 3, many mutations occur at positions 4, 5 and 6 indicating that these nucleotides are dispensable. The mutations at position 5 are not found when g3 or y3 replace A3. In contrast,

when the exon nucleotides AG are substituted, there are only few mutations at positions 4,5,6. The exon changes seem to be compensated by intron changes. Thus, the general rule, showing that U1RNA may recognize various sites at the exon-intron junction, holds true for homologous introns. This conclusion was also supported by examination of homologous introns of other genes, less frequently represented than the globin genes in our databank (results not shown). Furthermore, the data indicate that results such as those shown in Fig. 1-3 are not just due to the addition of various splice sites with a stable sequence in a given gene family but that they are the consequence of substitutions occurring in most genes in the course of their evolution and following strict rules.

Another prediction of the model is that any mutation that considerably reduces the stability of the U1RNA-5' splice site association will prevent cleavage and affect the expression of the gene. Some natural point mutations at the 5' splice site leading to genetic defects have been described (705-717). Table 4 shows that these mutations provoke a change of the sequence type such that the novel sequence is either not present in the databank (mutations at positions 1 and 2) or infrequent. In the latter cases, if some residual splicing occurs, it might be insufficient to allow normal expression of the gene.

Thus, if the variable geometry association of U1RNA and of the 5' splice site permits a large number of sequence changes, there are limits to these changes. Various parameters may probably explain how a particular sequence, which forms a low-stability hybrid with U1RNA, might either be used or be harmful and provoke a defect.

DISCUSSION

The recognition of the 5' splice site by U1RNA is one of the key-reactions involved in the mechanism of splicing. The variety of the 5' splice sites found in the genes compared to the constancy of the U1RNA sequence involved in the recognition requires a mechanism leading to a "variable geometry" association rather than to a rigid association. In this respect, the U1RNA-premRNA 5' splice site interaction resembles the 16SrRNA-procaryotic mRNA ribosome binding site interaction as first described by Shine and Dalgarno (718). In this case, a pyrimidine-rich sequence interacts with a purine-rich sequence and the exact location of the base-paired nucleotides is variable in both partners. Interestingly, the sequence from 16SrRNA

GENE OR PROTEIN INTRON	MUTATED AT POSITION	5' SPLICE SITE SEQUENCE	SEQUENCE TYPE FREQUENCY	REFERENCES
β -globin IVS1	-1	N A G G U u $\dot{g}$ G U M c	α C 6 48 γ C 6 1	705
β -globin IVS1	1	N A G G U u $\dot{g}$ G U M a	α C 6 48 - -	706
Factor IX IVS6	1	N A G G U A c u U M u	α A 8 57 - -	710
Phenylalanine hydroxylase IVS12	1	N c a G U A A G U M a	γ A 1 315 - -	711
β -hexoaminidase α -chain IVS12	1	N u G G U A A G g M c	β A 2 145 - -	712,713, 714
Collagen pro- α 2(I) IVS6	2	N u u G U A u a a M c	γ A 4 14 γ D -	715
Adenovirus E1A 13S IVS	3	N c a G U A A G U M u	γ A 1 315 γ C 1 6	716
β -globin IVS1	5 5	N A G G U u $\dot{g}$ G U M a M c	α C 6 48 α C 8 9 α C 8 9	708,709 706,711
Adenovirus E1A 9S IVS	5-6	N $\dot{g}$ G G U $\dot{g}$ A G g M a U	β B 2 161 β B 5 6	717
β -globin IVS1	6	N A G G U u $\dot{g}$ G U M c	α C 6 48 α C 7 19	706

TABLE 4 : Mutations in the 5' splice site leading to genetic defects. The normal sequence (N) is presented and overlined as in Fig. 5 to schematically show the complementarity to U1 RNA. Only the substituted nucleotides are shown for the mutated sequence (M). The sequence-type data are from Fig. 5. A dash indicates that the sequence was not found in our catalogue of normal exon-intron junctions.

resembles that from U1RNA in that it lacks G and has 1 or 2 CC boxes such as that shown in Figure 5 (718-719).

A variable geometry association has several implications for the mechanism of splicing and its regulation. As already discussed (Fig. 5), the signal for precise cleavage at the exon-intron junction cannot apparently reside in premRNA itself because of the variability of the junction sequence

and may be located in the constant 5' extremity of U1RNA (CC box for instance). The mechanism of 5' cleavage and of intervention of U1RNA remains unknown and would certainly deserve appropriate experimental approaches.

Variable geometry association introduces a large flexibility in the system. Note, for instance, that the U1RNA nucleotides 10-11 or 4-5-6) are not base-paired in many cases (Fig. 5). This may provoke loose local conformation of the splicing complex and allow, for instance, the insertion of factors which could otherwise not reach their target. The fact that hybrids of 5 to 7 contiguous base-pairs (out of 8 possible) are the most frequent may indicate that a few unpaired U1 nucleotides may be favourable for optimal functioning of the system. It is tempting to speculate that the variability of the 5' splice site-U1RNA interaction is a fundamental element for splice site selection in alternative splicing. Different 5' splice sites could be selected under given circumstances because they allow or not the action of some regulation factors. Unfortunately, experimental data are still too few to support this speculation.

The variable geometry association has also a profound impact on the course of phylogenetic evolution. By allowing a large array of mutations at the exon-intron junction, it leads to the selection of 5' splice sites with precise characteristics (GUAxGU in lower eucaryotes or AG/GUA in higher plants or G/GUGAGU in metazoa) and probably indirectly to the adaption of the splicing machinery. This gives birth to novel characteristics such that novel possibilities of regulation are open. Thus, the variable geometry association may participate to speciation and to diversity of eucaryotic cells.

ACKNOWLEDGEMENTS

We are grateful to Dr. B. Waslyk for critical reading of the manuscript and to Dr. Y. Grau for help in the classification of organisms. We also thank the secretarial staff of the laboratory for their patience in the preparation of the manuscript.

This work was supported by grants from CNRS, INSERM and the Association pour la Recherche sur le Cancer.

REFERENCES

1. Noda, M. et al. (1983) *Nature* **305**, 818-823.
2. Hamada, H. et al. (1982) *PNAS USA* **79**, 5901-5905.
3. Nakajima-Iijima, S. et al. (1985) *PNAS USA* **82**, 6133-6137.
4. Ng, S.-Y. et al. (1985) *Mol. Cell. Biol.* **5**, 2720-2732.
5. Erba, H.P. et al. (1988) *Mol. Cell. Biol.* **8**, 1775-1789.
6. Wiginton, D.A. et al. (1986) *Biochemistry* **25**, 8234-8244.
7. Duester, G. et al. (1986) *J. Biol. Chem.* **261**, 2027-2033.
8. Gingeras, T.R. et al. (1982) *J. Biol. Chem.* **13475**-13491.
9. Allard, A. & Wadell, G. (1988) *Virology* **164**, 220-229.
10. Dente, L. et al. (1985) *Nucl. Acids Res.* **13**, 3941-3952.
11. Fujinaka, K. et al. (1984) in 'The Molecular Biology of Adenoviruses 2.' Springer-Verlag. Ed: Doerfler, W. 53-72.
12. van Ormont, H. & Galibert, F. (1984) in 'The Molecular Biology of Adenoviruses 2.' Springer-Verlag. Ed: Doerfler, W. 73-142.
13. Merritt, C.M. & Board, P.G. (1988) *Gene* **66**, 97-106.
14. Buono, P. et al. (1988) *Nucl. Acids Res.* **16**, 4733.
15. Izzo, P. et al. (1988) *Eur. J. Biochem.* **174**, 569-578.
16. Karathanasis, S.K. et al. (1983) *PNAS USA* **80**, 6147-6151.
17. Lackner, K.J. et al. (1985) *Nucl. Acids Res.* **13**, 4597-4608.
18. Knott, T.J. et al. (1985) *Nucl. Acids Res.* **13**, 6387-6398.
19. Shelley, C.S. et al. (1985) *J. Mol. Biol.* **186**, 43-51.
20. Tsao, Y.-K. et al. (1985) *J. Biol. Chem.* **260**, 15222-15231.

20. Wei, C.-F. et al. (1985) *J. Biol. Chem.* **260**, 15211–15221.
21. Paik, Y.-K. et al. (1985) *PNAS USA* **82**, 3445–3449.
22. Carlsson, P. et al. (1986) *Gene* **49**, 29–51.
23. Elshourbagy, N.A. et al. (1987) *J. Biol. Chem.* **262**, 7973–7981.
24. Lauer, S.J. et al. (1988) *J. Biol. Chem.* **263**, 7277–7286.
25. Fojó, S.S. et al. (1987) *FEBS Letters* **213**, 221–226.
26. Nishide, T. et al. (1986) *Gene* **41**, 299–304.
27. Horii, A. et al. (1987) *Gene* **60**, 57–64.
28. Nemer, M. et al. (1984) *Nature* **312**, 654–656.
29. Greenberg, B.D. et al. (1984) *Nature* **312**, 656–658.
30. Hidaka, Y. et al. (1987) *Nucl. Acids Res.* **15**, 9086.
31. Broderick, T.P. et al. (1987) *PNAS USA* **84**, 3349–3353.
32. Freytag, S.O. et al. (1984) *Mol. Cell. Biol.* **4**, 1978–1984.
33. Ovchinnikov, Y.A. et al. (1988) *FEBS Letters* **233**, 87–94.
34. Marazziti, D. et al. (1988) *Biochemistry* **27**, 6529–6534.
35. Wiborg, O. et al. (1984) *PNAS USA* **81**, 1067–1069.
36. Ito, R. et al. (1984) *PNAS USA* **81**, 4662–4666.
37. Takahashi, Y. et al. (1985) *PNAS USA* **82**, 1931–1935.
38. Jonas, V. et al. (1985) *PNAS USA* **82**, 1994–1998.
39. Steenbergh, P.H. et al. (1986) *FEBS Letters* **209**, 97–103.
40. Ferrari, S. et al. (1987) *J. Biol. Chem.* **262**, 8325–8332.
41. Lang, G. et al. (1988) *EMBO J.* **7**, 1675–1682.
42. Stenberg, R.M. et al. (1985) *J. Virol.* **56**, 665–675.
43. Weiss, E.H. et al. (1982) *Nucl. Acids Res.* **10**, 1981–1994.
44. Sangiorgi, F.O. et al. (1985) *Nucl. Acids Res.* **13**, 2207–2225.
45. Barsh, G.S. et al. (1985) *PNAS USA* **82**, 2870–2874.
46. Soininen, R. et al. (1986) *PNAS USA* **83**, 1568–1572.
47. D'Alessio, M. et al. (1988) *Gene* **67**, 105–115.
48. Lei, K.-J. et al. (1985) *J. Biol. Chem.* **260**, 13377–13383.
49. Woo, P. et al. (1985) *J. Biol. Chem.* **260**, 13384–13388.
50. Hogg, D. et al. (1986) *J. Biol. Chem.* **261**, 12420–12427.
51. Shibahara, S. et al. (1983) *EMBO J.* **2**, 775–779.
52. Higashi, Y. et al. (1986) *PNAS USA* **83**, 2841–2845.
53. Kawajiri, K. et al. (1986) *Eur. J. Biochem.* **159**, 219–225.
54. Picado-Leonard, J. & Miller, W.L. (1987) *DNA* **6**, 439–448.
55. Chen, M.-J. et al. (1984) *J. Biol. Chem.* **259**, 3933–3943.
56. Bodescot, M. & Perricaudet, M. (1986) *Nucl. Acids Res.* **14**, 7103–7114.
57. Beisel, C. et al. (1985) *J. Virol.* **54**, 665–674.
58. Shimasaki, S. et al. (1988) *PNAS USA* **85**, 4218–4222.
59. Indik, Z. et al. (1987) *PNAS USA* **84**, 5680–5684.
60. Tal, M. et al. (1987) *Mol. Cell. Biol.* **7**, 2597–2601.
61. Jacobs K. et al. (1985) *Nature* **313**, 806–810.
62. Paoletti, G. et al. (1988) *Nucl. Acids Res.* **16**, 3545–3557.
63. Chung, D.W. et al. (1983) *Biochemistry* **22**, 3244–3250.
64. Rixon, M.W. et al. (1985) *Biochemistry* **24**, 2077–2086.
65. Sweetser, D.A. et al. (1987) *J. Biol. Chem.* **262**, 16060–16071.
66. Costanzo, F. et al. (1986) *Nucl. Acids Res.* **14**, 721–736.
67. Roebroek, A.J.M. et al. (1985) *EMBO J.* **4**, 2897–2903.
68. Nishizawa, M. et al. (1986) *Mol. Cell. Biol.* **6**, 511–517.
69. Gibbs, P.E.M. et al. (1987) *Biochemistry* **26**, 1332–1343.
70. Bishop, D.F. et al. (1988) *PNAS USA* **85**, 3903–3907.
71. Weinstein, L.S. et al. (1988) *FEBS Letters* **232**, 333–340.
72. Itoh, H. et al. (1988) *J. Biol. Chem.* **263**, 6656–6664.
73. White, J.W. & Saunders, G.F. (1986) *Nucl. Acids Res.* **14**, 4719–4730.
74. DeNoto, F.M. et al. (1981) *Nucl. Acids Res.* **9**, 3719–3730.
75. Slightom, J.L. et al. (1980) *Cell* **21**, 627–638.
76. Spritz, R.A. et al. (1980) *Cell* **21**, 639–646.
77. Lawn, R.M. et al. (1980) *Cell* **21**, 647–651.
78. Shen, S.-H. et al. (1981) *Cell* **26**, 191–203.
79. Poncz, M. et al. (1983) *J. Biol. Chem.* **258**, 11599–11609.
80. Michelson, A.M. & Orkins, S.H. (1983) *J. Biol. Chem.* **258**, 15245–15254.
81. Weller, P. et al. (1984) *EMBO J.* **3**, 439–446.
82. Hsu, S.-L. et al. (1988) *Nature* **331**, 94–96.
83. Fiddes, J.C. & Goodman, H.M. (1981) *J. Mol. Appl. Genet.* **1**, 3–18.
84. Policastro, P. et al. (1983) *J. Biol. Chem.* **258**, 11492–11499.
85. Talmadge, K. et al. (1984) *Nature* **307**, 37–40.
86. Mayo, K.E. et al. (1985) *PNAS USA* **82**, 63–67.
87. Spindel, E.R. et al. (1987) *Mol. Endocrinol.* **1**, 224–232.
88. Ishida, K. et al. (1987) *Nucl. Acids Res.* **15**, 10051.
89. Wells, D. et al. (1987) *Nucl. Acids Res.* **15**, 2871–2889.
90. Bensi, G. et al. (1985) *EMBO J.* **4**, 119–126.
91. Hickey, E. et al. (1986) *Nucl. Acids Res.* **14**, 4127–4145.
92. Dworniczak, B. & Mirault, M.-E. (1987) *Nucl. Acids Res.* **15**, 5181–5197.
93. Proia, R.L. & Soravia, E. (1987) *J. Biol. Chem.* **262**, 5677–5681.
94. Proia, R.L. (1988) *PNAS USA* **85**, 1883–1887.
95. Dull, T.J. et al. (1984) *Nature* **310**, 777–781.
96. Rotwein, P. et al. (1986) *J. Biol. Chem.* **261**, 4828–4832.
97. Matthysens, G. & Rabbitts, T.H. (1980) *PNAS USA* **77**, 6561–6565.
98. Bentley, D.L. & Rabbitts, T.H. (1980) *Nature* **288**, 730–733.
99. Rabbitts, T.H. et al. (1981) *Nucl. Acids Res.* **9**, 4509–4524.
100. Ravetch, J.V. et al. (1981) *Cell* **27**, 583–591.
101. Ellison, J. & Hood, L. (1982) *PNAS USA* **79**, 1984–1988.
102. Flanagan, J.G. & Rabbitts, T.H. (1982) *EMBO J.* **1**, 655–660.
103. Ellison, J.W. et al. (1982) *Nucl. Acids Res.* **10**, 4071–4079.
104. Rechavi, G. et al. (1983) *PNAS USA* **80**, 855–859.
105. Jaenichen, H.-R. et al. (1984) *Nucl. Acids Res.* **12**, 5249–5263.
106. Milstein, C.P. et al. (1984) *Nucl. Acids Res.* **12**, 6523–6535.
107. Pech, M. & Zachau, H.G. (1984) *Nucl. Acids Res.* **12**, 9229–9236.
108. Takahashi, N. et al. (1984) *PNAS* **81**, 5194–5198.
109. Klobbeck, H.-G. et al. (1985) *Nucl. Acids Res.* **13**, 4609–4613.
110. Huck, S. et al. (1986) *Nucl. Acids Res.* **14**, 1779–1789.
111. White, M.B. et al. (1985) *Science* **228**, 733–737.
112. Matsuda, F. et al. (1988) *EMBO J.* **7**, 1047–1051.
113. Suter, U. et al. (1987) *Nucl. Acids Res.* **15**, 7295–7308.
114. Yoshikai, Y. et al. (1985) *Nature* **316**, 837–840.
115. Seki, T. et al. (1985) *PNAS USA* **82**, 6657–6661.
116. Van den Elsen, P. et al. (1986) *PNAS USA* **83**, 2944–2948.
117. Tunnacliffe, A. et al. (1986) *EMBO J.* **5**, 1245–1252.
118. Güssow, D. et al. (1987) *J. Immunol.* **139**, 3132–3138.
119. Isobe, M. et al. (1988) *PNAS USA* **85**, 3933–3937.
120. Degraeve, W. et al. (1983) *EMBO J.* **2**, 2349–2353.
121. Fujita, T. et al. (1983) *PNAS USA* **80**, 7437–7441.
122. Furutani, Y. et al. (1986) *Nucl. Acids Res.* **14**, 3167–3179.
123. Clark, B.D. et al. (1986) *Nucl. Acids Res.* **14**, 7897–7914.
124. Campbell, H.D. et al. (1987) *PNAS USA* **84**, 6629–6633.
125. Tanabe, T. et al. (1987) *J. Biol. Chem.* **262**, 16580–16584.
126. Ishida, N. et al. (1985) *Nucl. Acids Res.* **13**, 7579–7589.
127. Leonard, W.J. et al. (1985) *Science* **230**, 633–639.
128. Stewart, A.G. et al. (1986) *FEBS Letters* **206**, 329–334.
129. Van ooyen, A. et al. (1985) *EMBO J.* **4**, 2905–2909.
130. Eckert, R.L. & Green, H. (1986) *Cell* **46**, 583–589.
131. Reich, N. et al. (1987) *PNAS USA* **84**, 6394–6398.
132. Bell, G.I. et al. (1980) *Nature* **284**, 26–32.
133. Tani, T. et al. (1988) *J. Biol. Chem.* **263**, 1231–1239.
134. Schedlich, L.J. et al. (1987) *DNA* **6**, 429–437.
135. Evans, B.A. et al. (1988) *Biochemistry* **27**, 3124–3129.
136. Marchuk, D. et al. (1985) *PNAS USA* **82**, 1609–1613.
137. Johnson, L.D. et al. (1985) *PNAS USA* **82**, 1896–1900.
138. Kitamura, N. et al. (1985) *J. Biol. Chem.* **260**, 8610–8617.
139. Ruppert, J.M. et al. (1988) *Mol. Cell. Biol.* **8**, 3104–3113.
140. Streuli, M. et al. (1987) *J. Exp. Med.* **166**, 1548–1566.
141. Südhof, T.C. et al. (1985) *Science* **228**, 815–822.
142. Gough, N.M. et al. (1988) *PNAS USA* **85**, 2623–2627.
143. Chen, C.-J. et al. (1986) *Cell* **47**, 381–389.
144. Malissen, M. et al. (1982) *PNAS USA* **79**, 893–897.
145. Lee, J.S. et al. (1982) *Nature* **299**, 750–752.
146. Das, H.K. et al. (1983) *PNAS USA* **80**, 3543–3547.
147. Boss, J.M. & Strominger, J.L. (1984) *PNAS USA* **81**, 5199–5203.
148. Soderoy, R. et al. (1984) *EMBO J.* **3**, 879–885.
149. Strachan, T. et al. (1984) *EMBO J.* **3**, 887–894.
150. Servenius, B. et al. (1984) *EMBO J.* **3**, 3209–3214.
151. N'Guyen, C. et al. (1985) *Immunogenetics* **21**, 479–489.
152. Lawrence, S.K. et al. (1985) *Nucl. Acids Res.* **13**, 7515–7528.
153. Kudo, J. et al. (1985) *Nucl. Acids Res.* **13**, 8827–8841.
154. Trowsdale, J. & Kelly, A. (1985) *EMBO J.* **4**, 2231–2237.
155. Boss, J.M. et al. (1985) *Mol. Cell. Biol.* **5**, 2677–2683.
156. Ways, J.P. et al. (1985) *J. Biol. Chem.* **260**, 11924–11933.
157. Gustafsson, K. et al. (1987) *J. Biol. Chem.* **262**, 8778–8786.
158. Geraghty, D.E. et al. (1987) *PNAS USA* **84**, 9145–9149.
159. Repaske, R. et al. (1985) *J. Virol.* **54**, 764–772.
160. Degen, S.J.F. & Davie, E.W. (1987) *Biochemistry* **26**, 6165–6177.
161. Karin, M. & Richards, R.I. (1982) *Nature* **299**, 797–802.

162. Schmidt, C.J. et al. (1985) *J. Biol. Chem.* **260**, 7731–7737.
163. Rabbitts, T.H. et al. (1983) *Nature* **306**, 760–765.
164. Stanton, L.W. et al. (1986) *PNAS USA* **83**, 1772–1776.
165. Julien, J.-P. et al. (1987) *Biochim. Biophys. Acta* **909**, 10–20.
166. Ullrich, A. et al. (1983) *Nature* **303**, 821–825.
167. Benesch, P. et al. (1985) *Nucl. Acids Res.* **13**, 1267–1281.
168. DiLella, A.G. et al. (1986) *Biochemistry* **25**, 743–749.
169. Sogawa, K. et al. (1983) *J. Biol. Chem.* **258**, 5306–5311.
170. Jenne, D. & Stanley, K.K. (1987) *Biochemistry* **26**, 6735–6742.
171. Rorsman, F. et al. (1988) *Mol. Cell. Biol.* **8**, 571–577.
172. Michelson, A.M. et al. (1985) *PNAS USA* **82**, 6965–6969.
173. Meeker, T.C. et al. (1987) *Oncogene Res.* **1**, 87–101.
174. Long, G.L. et al. (1984) *Biochemistry* **23**, 4828–4837.
175. Jagd, S. et al. (1985) *FEBS Letters* **193**, 213–216.
176. Stetler, G. et al. (1986) *Nucl. Acids Res.* **14**, 7883–7896.
177. Bao, J.-J. et al. (1987) *Biochemistry* **26**, 7755–7759.
178. Bosma, P.J. et al. (1988) *J. Biol. Chem.* **263**, 9129–9141.
179. Ragg, H. & Preibisch, G. (1988) *J. Biol. Chem.* **263**, 12129–12134.
180. Truong, A.T. et al. (1984) *EMBO J.* **3**, 429–437.
181. Cochet, M. et al. (1982) *Nature* **297**, 335–339.
182. Takahashi, H. et al. (1983) *Nucl. Acids Res.* **11**, 6847–6858.
183. Weiss, M.J. et al. (1988) *J. Biol. Chem.* **263**, 12002–12010.
184. Henthorn, P.S. et al. (1988) *J. Biol. Chem.* **263**, 12011–12019.
185. Knoll, B.J. et al. (1988) *J. Biol. Chem.* **263**, 12020–12027.
186. Kim, H.-S. & Maeda, N. (1986) *J. Biol. Chem.* **261**, 6712–6718.
187. Jeltsch, J.M. et al. (1987) *Nucl. Acids Res.* **15**, 1401–1414.
188. D'Onofrio, C. et al. (1985) *EMBO J.* **4**, 1981–1989.
189. Hardman, J.A. et al. (1984) *DNA* **3**, 457–468.
190. Hobart, P.M. et al. (1984) *PNAS USA* **81**, 5026–5030.
191. Miyazaki, H. et al. (1984) *PNAS USA* **81**, 5999–6003.
192. Nathans, J. & Hogness, D.S. (1984) *PNAS USA* **81**, 4851–4855.
193. Woo, P. et al. (1987) *J. Biol. Chem.* **262**, 15790–15795.
194. Gitschier, J. et al. (1984) *Nature* **312**, 326–330.
195. Foster, D.C. et al. (1985) *PNAS USA* **82**, 4673–4677.
196. Yoshitake, S. et al. (1985) *Biochemistry* **24**, 3736–3750.
197. Cool, D.E. & MacGillivray, R.T.A. (1987) *J. Biol. Chem.* **262**, 13662–13673.
198. Ichinose, A. & Davie, E.W. (1988) *PNAS USA* **85**, 5829–5833.
199. Riley, L.K. et al. (1988) *PNAS USA* **85**, 2489–2493.
200. Derynck, R. et al. (1987) *Nucl. Acids Res.* **15**, 3188–3189.
201. Parma, J. et al. (1987) *J. Mol. Biol.* **196**, 769–779.
202. Ny, T. et al. (1984) *PNAS USA* **81**, 5355–5359.
203. Degen, S.J.F. et al. (1986) *J. Biol. Chem.* **261**, 6972–6985.
204. Lee, M.G.-S. et al. (1984) *Nucl. Acids Res.* **12**, 5823–5836.
205. Hall, J.L. & Cowan, N.J. (1985) *Nucleic Acids Res.* **13**, 207–223.
206. O'Malley, K.L. et al. (1987) *Biochemistry* **26**, 6910–6914.
207. Kobayashi, K. et al. (1988) *J. Biochem.* **103**, 907–912.
208. Baker, R. T. & Board, P.G. (1987) *Nucl. Acids Res.* **15**, 443–463.
209. Riccio, A. et al. (1985) *Nucl. Acids Res.* **13**, 2759–2771.
210. Romana, M. et al. (1987) *Nucl. Acids Res.* **15**, 7343–7356.
211. Scott, A.F. et al. (1984) *Mol. Biol. Evol.* **1**, 371–389.
212. Savatier, P. et al. (1985) *J. Mol. Biol.* **182**, 21–29.
213. Slightom, J.L. et al. (1985) *Mol. Biol. Evol.* **2**, 370–389.
214. Marks, J. et al. (1986) *Nature* **321**, 785–788.
215. Harris, S. et al. (1986) *Mol. Biol. Evol.* **3**, 465–484.
216. Slightom, J.L. et al. (1987) *J. Biol. Chem.* **262**, 7472–7483.
217. Spritz, R.A. & Giebel, L.B. (1988) *Mol. Biol. Evol.* **5**, 21–29.
218. Dobner, P.R. et al. (1987) *Nucl. Acids Res.* **15**, 199–218.
219. Obaru, K. et al. (1988) *J. Mol. Biol.* **200**, 13–22.
220. Garner, I. et al. (1986) *EMBO J.* **5**, 2559–2567.
221. Zhang, K. et al. (1987) *Gene* **57**, 27–36.
222. Ceci, J.D. et al. (1987) *Gene* **59**, 171–182.
223. Williams, S.C. et al. (1986) *Mol. Cell. Biol.* **6**, 3807–3814.
224. Seidman, C.E. et al. (1984) *Science* **226**, 1206–1209.
225. Dush, M.K. et al. (1985) *PNAS USA* **82**, 2731–2735.
226. Kopito, R.R. et al. (1987) *J. Biol. Chem.* **262**, 8035–8040.
227. Negrini, M. et al. (1987) *Cell* **49**, 455–463.
228. Venta, P.J. et al. (1985) *J. Biol. Chem.* **260**, 12130–12135.
229. Barbás, J.A. et al. (1988) *EMBO J.* **7**, 625–632.
230. Keil, G.M. et al. (1987) *J. Virol.* **61**, 1901–1908.
231. Monson, J.M. et al. (1982) *Mol. Cell. Biol.* **2**, 1362–1371.
232. Rossi, P. & de Crombrughe, B. (1987) *PNAS USA* **84**, 5590–5594.
233. Lok, S. et al. (1984) *Nucl. Acids Res.* **12**, 4517–4529.
234. Limbach, K.J. & Wu, R. (1985) *Nucl. Acids Res.* **13**, 617–630.
235. Gonzalez, F.J. et al. (1985) *J. Biol. Chem.* **260**, 5040–5044; 11884–11889.
236. Crouse, G.F. et al. (1982) *J. Biol. Chem.* **257**, 7887–7897.
237. McDonald, J.D. et al. (1986) *Mol. Cell. Biol.* **6**, 842–848.
238. Shoemaker, C.B. & Mitsock, L.D. (1986) *Mol. Cell. Biol.* **6**, 849–858.
239. Phillips, M. et al. (1986) *J. Biol. Chem.* **261**, 10821–10827.
240. Balcarek, J. M. & Cowan, N.J. (1985) *Nucl. Acids Res.* **13**, 5527–5543.
241. Miller, L.H.I. et al. (1978) *PNAS USA* **75**, 6187–6191.
242. Konkel, D.A. et al. (1979) *Cell* **18**, 865–873.
243. Jahn, C.L. et al. (1980) *Cell* **21**, 159–168.
244. Hill, A. et al. (1984) *J. Biol. Chem.* **259**, 3739–3747.
245. Leder, A. et al. (1985) *Mol. Cell. Biol.* **5**, 1025–1033.
246. Wolf, O. et al. (1987) *J. Biol. Chem.* **262**, 16596–16603.
247. Tonegawa, S. et al. (1978) *PNAS USA* **75**, 1485–1489.
248. Bernard, O. et al. (1978) *Cell* **15**, 1133–1144.
249. Max, E.E. et al. (1979) *PNAS USA* **76**, 3450–3454.
250. Sakano, H. et al. (1979) *Nature* **280**, 288–294.
251. Seidman, J.G. et al. (1979) *Nature* **280**, 370–375.
252. Honjo, T. et al. (1979) *Cell* **18**, 559–568.
253. Tucker, P.W. et al. (1979) *Science* **206**, 1303–1306.
254. Yamawaki-Kataoka, Y. et al. (1980) *Nature* **283**, 786–789.
255. Calame, K. et al. (1980) *Nature* **284**, 452–455.
256. Nishioka, Y. & Leder, P. (1980) *J. Biol. Chem.* **255**, 3691–3694.
257. Kawakami, T. et al. (1980) *Nucl. Acids Res.* **8**, 3933–3945.
258. Goldberg, G.I. et al. (1981) *Gene* **15**, 33–42.
259. Sakano, H. et al. (1980) *Nature* **286**, 676–683.
260. Tucker, P.W. et al. (1980) *Science* **209**, 1353–1360.
261. Gough, N.M. & Bernard, O. (1981) *PNAS USA* **78**, 509–513.
262. Ollro, R. et al. (1981) *PNAS USA* **78**, 2442–2446.
263. Max, E.E. et al. (1981) *J. Biol. Chem.* **256**, 5116–5120.
264. Selsing, E. & Storb, U. (1981) *Cell* **25**, 47–58.
265. Kwan, S.-P. et al. (1981) *Cell* **26**, 57–66.
266. Yamawaki-Kataoka, Y. et al. (1981) *Nucl. Acids Res.* **9**, 1365–1381.
267. Ollro, R. et al. (1981) *Nucl. Acids Res.* **9**, 4099–4109.
268. Blomberg, B. & Tonegawa, S. (1982) *PNAS USA* **79**, 530–533.
269. Ollro, R. & Rougeon, F. (1982) *Nature* **296**, 761–763.
270. Kataoka, T. et al. (1982) *J. Biol. Chem.* **257**, 277–285.
271. Yamawaki-Kataoka, Y. et al. (1982) *PNAS USA* **79**, 2623–2627.
272. Early, P. et al. (1982) *Mol. Cell. Biol.* **2**, 829–836.
273. Ishida, N. et al. (1982) *EMBO J.* **1**, 1117–1123.
274. Hawley, R.G. et al. (1982) *PNAS USA* **79**, 7425–7429.
275. Ollro, R. & Rougeon, F. (1983) *Cell* **32**, 515–523.
276. Wu, G.E. et al. (1983) *Cell* **33**, 77–83.
277. Richards, J.E. et al. (1983) *Nature* **306**, 483–487.
278. Parslow, T.G. et al. (1984) *PNAS USA* **81**, 2650–2654.
279. Wels, J.A. et al. (1984) *EMBO J.* **3**, 2041–2046.
280. Giguère, V. et al. (1985) *EMBO J.* **4**, 2017–2024.
281. Ingraham, H.A. & Evans, G.A. (1986) *Mol. Cell. Biol.* **6**, 2923–2931.
282. Winoto, A. et al. (1985) *Nature* **316**, 832–836.
283. Chou, H.S. et al. (1987) *PNAS USA* **84**, 1992–1996.
284. Telford, J.L. et al. (1986) *Nucl. Acids Res.* **14**, 9955–9963.
285. Fuse, A. et al. (1984) *Nucl. Acids Res.* **12**, 9323–9331.
286. Todokoro, K. et al. (1985) *Gene* **39**, 103–107.
287. Miyatake, S. et al. (1985) *PNAS USA* **82**, 316–320.
288. Campbell, H.D. et al. (1985) *Eur. J. Biochem.* **150**, 297–304.
289. Campbell, H.D. et al. (1988) *Eur. J. Biochem.* **174**, 345–352.
290. van Ooyen, A. & Nusse, R. (1984) *Cell* **39**, 233–240.
291. Hug, H. et al. (1988) *Mol. Cell. Biol.* **8**, 3065–3079.
292. Mason, A.J. et al. (1983) *Nature* **303**, 300–307.
293. van Leeuwen, B.H. et al. (1986) *J. Biol. Chem.* **261**, 5529–5535.
294. Stevenson, B.J. et al. (1986) *Nucl. Acids Res.* **14**, 8307–8830.
295. Drinkwater, C.C. & Richards, R.J. (1987) *Nucl. Acids Res.* **15**, 10052.
296. Drinkwater, C.C. et al. (1987) *Biochemistry* **26**, 6750–6756.
297. Krieg, T.M. et al. (1985) *Nature* **260**, 5867–5870.
298. Saga, Y. et al. (1987) *PNAS USA* **84**, 5364–5368.
299. Takeshima, H. et al. (1988) *J. Mol. Biol.* **200**, 1–11.
300. Malissen, M. et al. (1983) *Science* **221**, 750–754.
301. Hyldig-Nielsen, J.J. et al. (1983) *Nucl. Acids Res.* **11**,

- 5055–5071.
302. Saito, H. et al. (1983) PNAS USA **80**, 5520–5524.
303. Arnold, B. et al. (1984) Nucl. Acids Res. **12**, 9473–9487.
304. Sher, B.T. et al. (1985) PNAS USA **82**, 1175–1179.
305. Levi-Strauss, M. et al. (1988) Science **240**, 201–204.
306. Larhammar, D. et al. (1985) J. Biol. Chem. **260**, 14111–14119.
307. Pontarotti, P.A. et al. (1986) PNAS USA **83**, 1782–1786.
308. Schepart, B.S. et al. (1986) J. Immunol. **136**, 3489–3495.
309. Koch, N. et al. (1987) EMBO J. **6**, 1677–1683.
310. Morgan, W.R. & Ward, D.C. (1986) J. Virol. **60**, 1170–1174.
311. Mann, R. & Baltimore, D. (1985) J. Virol. **54**, 401–407.
312. Glanville, N. et al. (1981) **292**, 267–269.
313. Clark, A.J. et al. (1984) EMBO J. **3**, 1045–1052.
314. Lavu, S. & Reddy, E.P. (1986) Nucl. Acids Res. **14**, 5309–5320.
315. DePinho, R.A. et al. (1986) PNAS USA **83**, 1827–1831.
316. Katoh, K. et al. (1988) Nucl. Acids Res. **16**, 3589.
317. Takahashi, N. et al. (1985) Cell **42**, 139–148.
318. Nave, K.-A. et al. (1987) PNAS USA **84**, 5665–5669.
319. Robert, B. et al. (1984) Cell **39**, 129–140.
320. Bourbon, H.-M. et al. (1988) J. Mol. Biol. **200**, 627–638.
321. Lewis, S.A. & Cowan, N.J. (1986) Mol. Cell. Biol. **6**, 1529–1534.
322. Levy, E. et al. (1987) Eur. J. Biochem. **166**, 71–77.
323. Evans, B.A. & Richards, R.I. (1985) EMBO J. **4**, 133–138.
324. Katz, A. & Kahana, C. (1988) J. Biol. Chem. **263**, 7604–7609.
325. Coffino, P. & Chen, E.L. (1988) Nucl. Acids Res. **16**, 2731–2732.
326. Adra, C.N. et al. (1987) Gene **60**, 65–74.
327. Selten, G. et al. (1986) Cell **46**, 603–611.
328. Ahearn, Jr. et al. (1987) J. Biol. Chem. **262**, 10695–10705.
329. Chrivia, J.C. et al. (1988) J. Biol. Chem. **263**, 5739–5744.
330. Ann, D.K. & Carlson, D.M. (1985) J. Biol. Chem. **260**, 15863–15872.
331. Holm, I. et al. (1984) EMBO J. **3**, 557–562.
332. Dudov, K.P. & Perry, R.P. (1984) Cell **37**, 457–468.
333. Wiedemann, L.M. & Perry, R.P. (1984) Mol. Cell. Biol. **4**, 2518–2528.
334. Stearman, R.S. et al. (1986) Nucl. Acids Res. **14**, 797–809.
335. Deng, T. et al. (1986) J. Biol. Chem. **261**, 16000–16005.
336. Semon, D. et al. (1987) Nucl. Acids Res. **15**, 9083–9084.
337. Degen, S.J.F. et al. (1987) Biochemistry **26**, 8270–8279.
338. Osumi, T. et al. (1987) J. Biol. Chem. **262**, 8138–8143.
339. Zakut, R. et al. (1982) **298**, 857–859.
340. Nudel, U. et al. (1983) Nucl. Acids Res. **11**, 1759–1771.
341. Leung, J.O. et al. (1985) J. Biol. Chem. **260**, 12523–12527.
342. Joh, K. et al. (1986) J. Mol. Biol. **190**, 401–410.
343. Boguski, M.S. et al. (1986) J. Biol. Chem. **261**, 6398–6407.
344. Argentin, S. et al. (1985) J. Biol. Chem. **260**, 4568–4571.
345. Ohtake, A. et al. (1988) J. Biol. Chem. **263**, 2245–2249.
346. Yu-Lee, L.-Y. & Rosen, J.M. (1983) J. Biol. Chem. **258**, 10794–10804.
347. Jones, W.K. et al. (1985) J. Biol. Chem. **260**, 7042–7050.
348. Deschenes, R.J. et al. (1985) J. Biol. Chem. **260**, 1280–1286.
349. Amara, S.G. et al. (1984) Mol. Cell. Biol. **4**, 2151–2160.
350. Kohno, K. et al. (1985) J. Biol. Chem. **260**, 4441–4447.
351. Den Dunnen, J.T. et al. (1986) PNAS USA **83**, 2855–2859.
352. Mizukami, Y. et al. (1983) PNAS USA **80**, 3958–3962.
353. Hines, R.N. et al. (1985) Arch. Biochem. Biophys. **237**, 465–476.
354. Sogawa, K. et al. (1985) J. Biol. Chem. **260**, 5026–5032.
355. Suwa, Y. et al. (1985) J. Biol. Chem. **260**, 7980–7984.
356. Morishima, N. et al. (1987) Biochemistry **26**, 8279–8285.
357. Umeno, M. et al. (1988) PNAS USA **263**, 4956–4962.
358. Virbasius, J.V. & Scarpulla R.C. (1988) J. Biol. Chem. **263**, 6791–6796.
359. Perret, C. et al. (1988) Eur. J. Biochem. **172**, 43–51.
360. Ishii, N. et al. (1987) J. Biol. Chem. **262**, 8144–8150.
361. Sakimura, K. et al. (1987) Gene **60**, 103–113.
362. Schwarzbauer, J.E. et al. (1987) EMBO J. **6**, 2573–2580.
363. Crabtree, G.R. et al. (1985) J. Mol. Biol. **185**, 1–19.
364. Morgan, J.G. et al. (1987) Nucl. Acids Res. **15**, 2774–2776.
365. Sweetser, D.A. et al. (1986) J. Biol. Chem. **261**, 5553–5561.
366. Leibold, E.A. & Munro, H.N. (1987) J. Biol. Chem. **262**, 7335–7341.
367. Sargent, T.D. et al. (1981) Mol. Cell. Biol. **1**, 871–883.
368. Esumi, H. et al. (1983) PNAS USA **80**, 95–99.
369. Heinrich, G. et al. (1984) J. Biol. Chem. **259**, 14082–14087.
370. Page, G.S. et al. (1981) Nucl. Acids Res. **9**, 2087–2104.
371. Barta, A. et al. (1981) PNAS USA **78**, 4867–4871.
372. Wong, W.M. et al. (1988) Nucl. Acids Res. **16**, 2342.
373. Jameson, L. et al. (1984) J. Biol. Chem. **259**, 15474–15480.
374. Okuda, A. et al. (1987) J. Biol. Chem. **262**, 3858–3863.
375. Müller, R.M. et al. (1987) J. Biol. Chem. **262**, 6795–6802.
376. Sorger, P.K. & Pelham, H.R.B. (1987) EMBO J. **6**, 993–998.
377. Soares, M.B. et al. (1986) J. Mol. Biol. **192**, 737–752.
378. Ueno, T. et al. (1987) BBRC **148**, 344–349.
379. Steen, M.-L. et al. (1984) J. Mol. Biol. **177**, 19–32.
380. Lomedico, P. et al. (1979) Cell **18**, 545–558.
381. Swift, G.H. et al. (1984) J. Biol. Chem. **259**, 14271–14278.
382. Craik, C.S. et al. (1984) J. Biol. Chem. **259**, 14255–14264.
383. Bell, G.I. et al. (1984) J. Biol. Chem. **259**, 14265–14270.
384. Hijikata, M. et al. (1987) J. Biol. Chem. **262**, 8151–8158.
385. Qasba, P.K. & Safaya, S.K. (1984) Nature **308**, 377–380.
386. Rhodes, C. et al. (1988) J. Biol. Chem. **263**, 6063–6067.
387. Hayashi, K. et al. (1987) Nucl. Acids Res. **15**, 6419–6436.
388. Lemke, G. et al. (1988) Neuron **1**, 73–83.
389. Nudel, U. et al. (1984) Nucl. Acids Res. **12**, 7175–7186.
390. Helfman, D.M. et al. (1986) Mol. Cell. Biol. **6**, 3582–3595.
391. Breitbart, R.E. & Nadal-Ginard, B. (1986) J. Mol. Biol. **188**, 313–324.
392. Strehler, E.E. et al. (1986) J. Mol. Biol. **190**, 291–317.
393. Takiguchi, M. et al. (1987) PNAS USA **84**, 6136–6140.
394. Cognet, M. et al. (1987) J. Mol. Biol. **196**, 11–25.
395. Gubbins, E.J. et al. (1980) J. Biol. Chem. **255**, 8655–8662.
396. Drouin, J. et al. (1985) FEBS Lett. **193**, 54–58.
397. Heinrich, G. et al. (1984) J. Biol. Chem. **259**, 3320–3329.
398. Ruta, M. et al. (1986) Mol. Cell. Biol. **6**, 1706–1710.
399. Hurst, H.C. & Parker, M.G. (1983) EMBO J. **2**, 769–774.
400. Delaey, B. et al. (1987) Nucl. Acids Res. **15**, 1627–1641.
401. Montminy, M.R. et al. (1984) PNAS USA **81**, 3337–3340.
402. Harris, S.E. et al. (1983) PNAS USA **80**, 6460–6464.
403. Musti, A.M. et al. (1986) PNAS USA **83**, 323–327.
404. Lemischka, I. & Sharp, P.A. (1982) Nature **300**, 330–335.
405. Schmale, H. et al. (1984) EMBO J. **3**, 3289–3293.
406. Falany, C.N. et al. (1987) J. Biol. Chem. **262**, 5924–5930.
407. Pan, T.-C. et al. (1987) Eur. J. Biochem. **170**, 99–104.
408. Quax-Jeukens, Y. et al. (1985) PNAS USA **82**, 5819–5823.
409. van den Heuvel, R. et al. (1985) J. Mol. Biol. **185**, 273–284.
410. Zarain-Herzberg, A. et al. (1988) J. Biol. Chem. **263**, 4807–4812.
411. Mitchell, P.J. et al. (1986) Mol. Cell. Biol. **6**, 1926–1935.
412. Nakanishi, T. et al. (1988) J. Biol. Chem. **263**, 6384–6391.
413. Cicila, G. et al. (1985) Biochemistry **24**, 3076–3080.
414. Woychik, R.P. et al. (1982) Nucl. Acids Res. **10**, 7197–7210.
415. Rohrbaugh, M.L. et al. (1985) Mol. Cell. Biol. **5**, 147–160.
416. Hardison, R.C. (1981) J. Biol. Chem. **256**, 11780–11786.
417. Hardison, R.C. (1983) J. Biol. Chem. **258**, 8739–8744.
418. Hardison, R.C. et al. (1979) Cell **18**, 1285–1297.
419. van den Berg, J. et al. (1978) Nature **276**, 37–44.
420. Schimenti, J.C. & Duncan, C.H. (1985) Mol. Biol. Evol. **2**, 505–513.
421. Schon, E.A. et al. (1981) Cell **27**, 359–369.
422. Schon, E.A. et al. (1982) J. Biol. Chem. **257**, 6825–6835.
423. Shapiro, S.G. et al. (1983) J. Mol. Biol. **169**, 31–52.
424. Wernke, S.M. & Lingrel, J.B. (1986) J. Mol. Biol. **192**, 457–471.
425. Blanchetot, A. et al. (1983) Nature **301**, 732–734.
426. Virgin, J.B. et al. (1985) J. Biol. Chem. **260**, 7072–7077.
427. Goodwin, R.G. et al. (1983) Nucl. Acids Res. **11**, 6873–6882.
428. Luskey, K.L. (1987) Mol. Cell. Biol. **7**, 1881–1893.
429. Krawetz, S.A. et al. (1988) J. Biol. Chem. **263**, 321–326.
430. Heidmann, O. & Rougeon, F. (1983) Cell **34**, 767–777.
431. Emorine, L. & Max, E.E. (1983) Nucl. Acids Res. **11**, 8877–8890.
432. Heidmann, O. & Rougeon, F. (1984) Nature **311**, 74–76.
433. McGuire, K.L. et al. (1985) Nucl. Acids Res. **13**, 5611–5628.
434. Kwok, S.C.M. et al. (1983) J. Biol. Chem. **258**, 2357–2363.
435. Chan, S.J. et al. (1984) PNAS USA **81**, 5046–5050.
436. Rieger, M. et al. (1985) EMBO J. **4**, 2261–2267.
437. Bader, B.L. et al. (1986) EMBO J. **5**, 1865–1875.
438. Laird, J.E. et al. (1988) Biochem. J. **254**, 85–94.
439. Marche, P.N. et al. (1985) Immunogenet. **21**, 71–82.

440. Sittisombut, N. et al. (1988) *J. Immunol.* **140**, 3237–3243.
441. Tam, Y.C. et al. (1988) *Biochem. Biophys. Res. Com.* **153**, 209–216.
442. Peterson, M.G. et al. (1988) *Eur. J. Biochem.* **174**, 417–424.
443. Hoshimaru, M. & Nakanishi, S. (1987) *J. Biol. Chem.* **262**, 14625–14632.
444. Lee, C.-P. et al. (1987) *J. Biol. Chem.* **262**, 4195–4199.
445. Nakanishi, S. et al. (1980) *Nature* **287**, 752–755.
446. Ann, D.K. et al. (1987) *J. Biol. Chem.* **262**, 3958–3963.
447. Nathans, J. & Hogness, D.S. (1983) *Cell* **34**, 807–814.
448. Haley, J. et al. (1987) *J. Biol. Chem.* **262**, 11940–11946.
449. Irwin, D.M. et al. (1988) *J. Mol. Biol.* **200**, 31–45.
450. Nawa, H. et al. (1984) *Nature* **312**, 729–734.
451. Suske, G. et al. (1983) *Nucl. Acids Res.* **11**, 2257–2271.
452. Bailly, A. et al. (1983) *J. Biol. Chem.* **258**, 10384–10389.
453. Nagamine, Y. et al. (1984) *Nucl. Acids Res.* **12**, 9525–9541.
454. Quax, W. et al. (1983) *Cell* **35**, 215–223.
455. Quax, W. et al. (1985) *Cell* **43**, 327–338.
456. Nef, P. et al. (1984) *PNAS USA* **81**, 7975–7979.
457. Nef, P. et al. (1988) *EMBO J.* **7**, 595–601.
458. Fornwald, J.A. et al. (1982) *Nucl. Acids Res.* **10**, 3861–3876.
459. Kost, T.A. et al. (1983) *Nucl. Acids Res.* **11**, 8287–8301.
460. Chang, K.S. et al. (1985) *Nucl. Acids Res.* **13**, 1223–1237.
461. Bergsma, D.J. et al. (1985) *Mol. Cell. Biol.* **5**, 1151–1162.
462. Eldridge, J. et al. (1985) *Gene* **36**, 55–63.
463. Carroll, S.L. et al. (1986) *J. Biol. Chem.* **261**, 8965–8976.
464. Suminami, Y. et al. (1988) *J. Biochem.* **103**, 611–617.
465. Van het Schip, A.D. et al. (1983) *Nucl. Acids Res.* **11**, 2529–2540.
466. Maguire, D.J. et al. (1986) *Nucl. Acids Res.* **14**, 1379–1391.
467. Yoshihara, C.M. et al. (1987) *Nucl. Acids Res.* **15**, 753–770.
468. Minghetti, P.P. et al. (1988) *Mol. Endocrinology* **2**, 355–367.
469. Wilson, P.W. et al. (1988) *J. Mol. Biol.* **200**, 615–625.
470. Vlasak, R. et al. (1987) *Eur. J. Biochem.* **169**, 53–58.
471. Minvielle, S. et al. (1986) *FEBS Lett.* **203**, 7–10.
472. Simmen, R.C.M. et al. (1985) *J. Biol. Chem.* **260**, 907–912.
473. Yamada, Y. et al. (1980) *Cell* **22**, 887–892.
474. Dickson, L.A. et al. (1981) *J. Biol. Chem.* **256**, 8407–8415.
475. Wozney, J. et al. (1981) *Nature* **294**, 129–135.
476. Tate, V.E. et al. (1983) *Nucl. Acids Res.* **11**, 91–104.
477. Yamada, Y. et al. (1983) *Nucl. Acids Res.* **11**, 2733–2744.
478. Yameda, Y. et al. (1984) *Nature* **310**, 333–337.
479. Aho, S. et al. (1984) *Nucl. Acids Res.* **12**, 6117–6125.
480. Finer, M.H. et al. (1987) *J. Biol. Chem.* **262**, 13323–13332.
481. Yasuda, K. et al. (1984) *EMBO J.* **3**, 1397–1402.
482. Nickerson, J.M. et al. (1986) *J. Biol. Chem.* **261**, 552–557.
483. Thompson, M.A. et al. (1987) *Gene* **56**, 173–184.
484. Zahraoui, A. & Cuny, G. (1987) *Eur. J. Biochem.* **166**, 63–69.
485. Fujiwara, K.T. et al. (1987) *J. Virol.* **61**, 4012–4018.
486. Stone, E.M. et al. (1985) *PNAS USA* **82**, 1628–1632.
487. Dodgson, J.B. & Engel, J.D. (1983) *J. Biol. Chem.* **258**, 4623–4629.
488. Engel, J.D. et al. (1983) *PNAS USA* **80**, 1392–1396.
489. Dodgson, J.B. et al. (1983) *J. Biol. Chem.* **258**, 12685–12692.
490. Dolan, M. et al. (1983) *J. Biol. Chem.* **258**, 3983–3990.
491. Erbil, C. & Niessing, J. (1982) *Gene* **20**, 211–217.
492. Erbil, C. & Niessing, J. (1983) *EMBO J.* **2**, 1339–1343.
493. Erbil, C. & Niessing, J. (1984) *Gene* **32**, 161–170.
494. Patient, R.K. et al. (1983) *J. Biol. Chem.* **258**, 8521–8523.
495. Knöchel, W. et al. (1988) *Nucl. Acids Res.* **16**, 1625.
496. Knöchel, W. et al. (1987) *Nucl. Acids Res.* **15**, 10062.
497. Brush, D. et al. (1985) *Mol. Cell. Biol.* **5**, 1307–1317.
498. Kleinsiek, D.A. et al. (1986) *Nucl. Acids Res.* **14**, 10053–10069.
499. Litman, G.W. et al. (1985) *PNAS USA* **82**, 844–848.
500. Perler, F. et al. (1980) *Cell* **20**, 555–566.
501. Itoh, N. et al. (1988) *J. Biol. Chem.* **263**, 7628–7631.
502. Molloy, P.L. et al. (1982) *Nucl. Acids Res.* **10**, 6007–6021.
503. Jung, A. et al. (1980) *Biochemistry* **19**, 5759–5763.
504. Miyatani, S. et al. (1986) *J. Cell. Biol.* **103**, 1957–1965.
505. Kiss, I. et al. (1987) *PNAS USA* **84**, 6399–6403.
506. Nabeshima, Y.-I. et al. (1984) *Nature* **308**, 333–338.
507. Kropp, K. et al. (1986) *J. Biol. Chem.* **261**, 6613–6618.
508. Breathnach, R. et al. (1978) *PNAS USA* **75**, 4853–4857.
509. Woo, S.L.C. et al. (1981) *Biochemistry* **20**, 6437–6446.
510. Stein, J.P. et al. (1980) *Cell* **21**, 681–687.
511. Heilig, R. et al. (1982) *Nucl. Acids Res.* **10**, 4363–4382.
512. Jeltsch, J.-M. et al. (1987) *Nucl. Acids Res.* **15**, 7643–7645.
513. Hayashi, K. et al. (1988) *BBRC* **152**, 776–782.
514. Lonberg, N. & Gilbert, W. (1985) *Cell* **40**, 81–90.
515. Huckaby, C.S. et al. (1987) *PNAS USA* **84**, 8380–8384.
516. Beccari, E. & Mazzetti, P. (1987) *Nucl. Acids Res.* **15**, 1870–1872.
517. Loreni, F. et al. (1985) *EMBO J.* **4**, 3483–3488.
518. Tso, J.Y. et al. (1986) *Nucl. Acids Res.* **14**, 2187–2200.
519. Sullivan, K.F. et al. (1986) *J. Biol. Chem.* **261**, 13317–13322.
520. Pratt, L.F. & Cleveland, D.W. (1988) *EMBO J.* **7**, 931–940.
521. Smith, D.J. (1988) *Biochem. J.* **249**, 465–472.
522. Bond, U., Schlesinger, M.J. (1986) *Mol. Cell. Biol.* **6**, 4602–4610.
523. Zehner, Z.E. et al. (1987) *J. Biol. Chem.* **262**, 8112–8120.
524. Van het Schip, F.D. et al. (1987) *J. Mol. Biol.* **196**, 245–260.
525. Gerber-Huber, S. et al. (1987) *Nucl. Acids Res.* **15**, 4737–4760.
526. Cooper, A.D. & Crain, Jr. W.R. (1982) *Nucl. Acids Res.* **10**, 4081–4092.
527. Schuler, M.A. et al. (1983) *Mol. Cell. Biol.* **3**, 448–456.
528. Johnson, P.J. et al. (1987) *Nucl. Acids Res.* **15**, 1824–1833.
529. Fyrberg, E.A. et al. (1981) *Cell* **24**, 107–116.
530. Bond, B.J. & Davidson, N. (1986) *Mol. Cell. Biol.* **6**, 2080–2088.
531. Karlik, C.C. et al. (1984) *Cell* **38**, 711–719.
532. Benyajati, C. et al. (1982) *Nucl. Acids Res.* **10**, 7261–7272.
533. Cohn, V.H. & Moore, G.P. (1988) *Mol. Biol. Evol.* **5**, 154–166.
534. Laughon, A. et al. (1986) *Mol. Cell. Biol.* **6**, 4676–4689.
535. Höög, C. et al. (1986) *Nucl. Acids Res.* **14**, 703–719.
536. Eveleth, Jr. et al. (1986) *Nucl. Acids Res.* **14**, 6169–6183.
537. Xanthopoulos, K.G. et al. (1988) *Eur. J. Biochem.* **172**, 371–376.
538. Monson, J.M. et al. (1982) *PNAS USA* **79**, 1761–1765.
539. Kramer, J.M. et al. (1982) *Cell* **30**, 599–606.
540. Eveleth, D.D. et al. (1986) *EMBO J.* **5**, 2663–2672.
541. Morgan, B.A. et al. (1986) *Mol. Cell. Biol.* **5**, 3335–3342.
542. Chen, C.-N. et al. (1987) *Nature* **329**, 721–724.
543. Greenwald, I. (1985) *Cell* **43**, 583–590.
544. Kassis, J.A. et al. (1986) *EMBO J.* **5**, 3583–3589.
545. Martin, C.H. et al. (1988) *J. Mol. Biol.* **201**, 273–287.
546. Cook, J.L. et al. (1988) *J. Biol. Chem.* **263**, 10858–10864.
547. Henikoff, S. et al. (1983) *Cell* **34**, 405–414.
548. Fouts, D. et al. (1988) *Gene* **63**, 261–275.
549. Holmgren, R. et al. (1981) *PNAS USA* **78**, 3775–3778.
550. Hovemann, B. et al. (1986) *Mol. Gen. Genet.* **204**, 334–340.
551. Pauli, D. et al. (1988) *J. Mol. Biol.* **200**, 47–53.
552. Kay, R.J. et al. (1987) *Nucl. Acids Res.* **15**, 3723–3741.
553. Snutch, T.P. et al. (1988) *Gene* **64**, 241–255.
554. Jones, D. et al. (1986) *J. Biol. Chem.* **261**, 12006–12015.
555. DeWet, J.R. et al. (1987) *Mol. Cell. Biol.* **7**, 725–737.
556. Karn, J. et al. (1983) *PNAS USA* **80**, 4253–4257.
557. Wassenberg II, D.R. et al. (1987) *J. Biol. Chem.* **262**, 10741–10747.
558. Basi, G.S. & Storti, R.V. (1986) *J. Biol. Chem.* **261**, 817–827.
559. Hanke, P.D. & Storti, R.V. (1988) *Mol. Cell. Biol.* **8**, 3591–3602.
560. Wharton, K.A. et al. (1985) *Cell* **40**, 55–62.
561. Chisholm, G.E. & Henner, D.J. (1988) *J. Virol.* **62**, 3193–3200.
562. Rosenthal, A. et al. (1987) *EMBO J.* **6**, 433–441.
563. Rafii, F. et al. (1988) *Nucl. Acids Res.* **16**, 4915–4926.
564. Aguadé, M. (1988) *Mol. Biol. Evol.* **5**, 433–441.
565. Tsujimoto, Y. & Suzuki, Y. (1979) *Cell* **18**, 591–600.
566. Okamoto, H. et al. (1982) *J. Biol. Chem.* **257**, 15192–15199.
567. Chevillard, M. et al. (1986) *Nucl. Acids Res.* **14**, 6341–6342.
568. Sucov, H.M. et al. (1987) *Dev. Biol.* **120**, 507–519.
569. Seto, N.O.L. et al. (1987) *Nucl. Acids Res.* **15**, 10601.
570. Frei, E. et al. (1988) *EMBO J.* **7**, 197–204.
571. Sakurai, H. et al. (1988) *J. Biol. Chem.* **263**, 7876–7880.
572. Bowtell, D.D.L. et al. (1988) *Genes & Dev.* **2**, 620–634.
573. Boggs, R.T. et al. (1987) *Cell* **50**, 739–747.
574. Lipshitz, H.D. et al. (1987) *Genes & Dev.* **1**, 307–322.
575. Shyu, A.-B. et al. (1987) *Nucl. Acids Res.* **15**, 10405–10417.
576. Hovemann, B. et al. (1981) *Nucl. Acids Res.* **9**, 4721–4734.
577. Hung, M.-C. & Wensink, P.C. (1981) *Nucl. Acids Res.* **9**, 6407–6419.
578. Hung, M.-C. & Wensink, P.C. (1983) *J. Mol. Biol.* **164**, 481–492.
579. Garabedian, M.J. et al. (1987) *Gene* **55**, 1–8.
580. Zhao, D.-B. et al. (1988) *EMBO J.* **7**, 1115–1119.
581. Nader, W.F. et al. (1986) *Gene* **48**, 133–144.

582. Nellen, W. & Gallwitz, D. (1982) *J. Mol. Biol.* **159**, 1–18.
583. Gallwitz, D. & Sures, I. (1980) *PNAS USA* **77**, 2546–2550.
584. Ng, R. & Abelson, J. (1980) *PNAS USA* **77**, 3912–3916.
585. Nellen, W. et al. (1981) *J. Mol. Appl. Genet.* **1**, 239–244.
586. Wildeman, A.G. (1988) *Nucl. Acids Res.* **16**, 2553–2564.
587. McKnight, G.L. et al. (1985) *EMBO J.* **4**, 2093–2099.
588. Boel, E. et al. (1984) *EMBO J.* **3**, 1581–1585.
589. Martindale, D.W. & Taylor, F.M. (1988) *Nucl. Acids Res.* **16**, 2189–2201.
590. Ettinger, W.F. et al. (1987) *Biochemistry* **26**, 7883–7892.
591. Uemura, T. et al. (1987) *Nucl. Acids Res.* **15**, 9727–9739.
592. Lin, R.-J. et al. (1985) *J. Biol. Chem.* **260**, 14780–14792.
593. Sundstrom, P. et al. (1987) *Nucl. Acids Res.* **15**, 9997–10006.
594. Simmons, D. et al. (1987) *EMBO J.* **6**, 485–491.
595. Punt, P.J. et al. (1988) *Gene* **69**, 49–57.
596. Wilhelm, M.L. & Wilhelm, F.-X. (1987) *Nucl. Acids Res.* **15**, 5478.
597. Woudt, L.P. et al. (1983) *Nucl. Acids Res.* **11**, 5347–5360.
598. Scherf, A. et al. (1988) *EMBO J.* **7**, 1129–1137.
599. Triglia, T. et al. (1987) *EMBO J.* **6**, 1413–1419.
600. Simmel, T.L. et al. (1988) *Nucl. Acids Res.* **16**, 1219.
601. Kimmel, A.R. & Firtel, R.A. (1980) *Nucl. Acids Res.* **8**, 5599–5610.
602. Munger, K. et al. (1985) *EMBO J.* **4**, 2665–2668.
603. Hammer III, J.A. et al. (1987) *J. Cell Biol.* **105**, 913–925.
604. Clements, J.M. & Roberts, C.F. (1986) *Gene* **44**, 97–105.
605. Simon, M. et al. (1986) *EMBO J.* **5**, 2697–2701.
606. Wilson, L.J. et al. (1988) *Nucl. Acids Res.* **16**, 2339.
607. Oakley, B.R. et al. (1987) *Gene* **61**, 385–399.
608. Cantoral, J.M. et al. (1988) *Nucl. Acids Res.* **16**, 8177.
609. Favaloro, J.M. et al. (1986) *Nucl. Acids Res.* **14**, 8265–8277.
610. Steel, L.F. et al. (1987) *Nucl. Acids Res.* **15**, 10285–10298.
611. Teem, J.L. & Rosbash, M. (1983) *PNAS USA* **80**, 4403–4407.
612. Leer, R.J. et al. (1984) *Nucl. Acids Res.* **12**, 6685–6700.
613. Leer, R.J. et al. (1985) *Nucl. Acids Res.* **13**, 701–709.
614. Leer, R.J. et al. (1985) *Nucl. Acids Res.* **13**, 5027–5039.
615. Dabeva, M.D. & Warner, J.R. (1987) *J. Biol. Chem.* **262**, 16055–16059.
616. Suzuki, K. & Otake, E. (1988) *Nucl. Acids Res.* **16**, 6223.
617. Kreader, C.A. & Heckman, J.E. (1987) *Nucl. Acids Res.* **15**, 9027–9042.
618. Kuiper, M.T.R. et al. (1988) *J. Biol. Chem.* **263**, 2840–2847.
619. Nielsen, H. et al. (1986) *EMBO J.* **5**, 2711–2717.
620. McKnight, G.L. et al. (1986) *Cell* **46**, 143–147.
621. Smith, H.A. et al. (1988) *Gene* **63**, 53–63.
622. Monteiro, M.J. & Cox, R.A. (1987) *J. Mol. Biol.* **193**, 427–438.
623. Werenskiold, A.K. et al. (1988) *Eur. J. Biochem.* **174**, 491–495.
624. Shah, D.M. et al. (1982) *PNAS USA* **79**, 1022–1026.
625. Nairn, C.J. et al. (1988) *Gene* **65**, 247–257.
626. Shah, D.M. et al. (1983) *J. Mol. Appl. Genet.* **2**, 111–126.
627. Dennis, R.F. et al. (1985) *Nucl. Acids Res.* **13**, 727–743.
628. Whittier, E.S. et al. (1987) *Nucl. Acids Res.* **15**, 2515–2535.
629. Boutry, M. & Chua, N.-H. (1985) *EMBO J.* **4**, 2159–2165.
630. Karlín-Neumann, G.A. et al. (1985) *J. Mol. Appl. Genet.* **3**, 45–61.
631. Sommer, H. & Saedler, H. (1986) *Mol. Gen. Genet.* **202**, 429–434.
632. Klee, H.J. et al. (1987) *Mol. Gen. Genet.* **210**, 437–442.
633. Chen, J. & Varner, J.E. (1985) *EMBO J.* **4**, 2145–2151.
634. Brisson, N. & Verma, D.P.S. (1982) *PNAS USA* **79**, 4055–4059.
635. Wiborg, O. et al. (1982) *Nucl. Acids Res.* **10**, 3487–3494.
636. Hyldig-Nielsen, J.J. et al. (1982) *Nucl. Acids Res.* **10**, 689–701.
637. Tischer, E. et al. (1986) *Mol. Gen. Genet.* **203**, 221–229.
638. Rochester, D.E. et al. (1986) *EMBO J.* **5**, 451–458.
639. Czarnecka, E. et al. (1988) *Mol. Cell. Biol.* **8**, 1113–1122.
640. Gatehouse, J.A. et al. (1988) *Biochem. J.* **250**, 15–24.
641. Lycett, G.W. et al. (1984) *Nucl. Acids Res.* **12**, 4493–4505.
642. Bown, D. et al. (1985) *Nucl. Acids Res.* **13**, 4527–4538.
643. Bumlein, H. et al. (1986) *Nucl. Acids Res.* **14**, 2707–2720.
644. Yenofsky, R.L. et al. (1988) *Mol. Gen. Genet.* **211**, 215–222.
645. Mauro, V.P. et al. (1985) *Nucl. Acids Res.* **13**, 239–249.
646. Katínakis, P. & Verma, D.P.S. (1985) *PNAS USA* **82**, 4157–4161.
647. Bevan, M. et al. (1986) *Nucl. Acids Res.* **14**, 4625–4638.
648. Pikaard, C.S. et al. (1986) *Nucl. Acids Res.* **14**, 5564–5566.
649. Rosahl, S. et al. (1986) *Mol. Gen. Genet.* **203**, 214–220.
650. Mignery, G.A. et al. (1988) *Gene* **62**, 27–44.
651. Keil, M. et al. (1986) *Nucl. Acids Res.* **14**, 5641–5650.
652. Fujiyama, K. et al. (1988) *Eur. J. Biochem.* **173**, 681–687.
653. Tumer, N.E. et al. (1986) *Nucl. Acids Res.* **14**, 3325–3342.
654. Coruzzi, G. et al. (1984) *EMBO J.* **3**, 1671–1679.
655. Mazur, B.J. & Chui, C.-F. (1985) *Nucl. Acids Res.* **13**, 2373–2386.
656. Werr, W. et al. (1985) *EMBO J.* **4**, 1373–1380.
657. Marchionni, M. & Gilbert, W. (1986) *Cell* **46**, 133–141.
658. Holdsworth, M.J. et al. (1987) *Nucl. Acids Res.* **15**, 10600.
659. Marks, M.D. et al. (1987) *Plant Mol. Biol.* **10**, 91–104.
660. Oppenheimer, D.G. et al. (1988) *Gene* **63**, 87–102.
661. Takaiwa, F. et al. (1987) *FEBS Lett.* **221**, 43–47.
662. Allen, R.D. et al. (1987) *Mol. Gen. Genet.* **210**, 211–218.
663. Bown, D. et al. (1988) *Biochem. J.* **251**, 717–726.
664. Weschke, W. et al. (1987) *Nucl. Acids Res.* **15**, 10065.
665. Sun, S.M. et al. (1981) *Nature* **289**, 37–41.
666. Slightom, J.L. et al. (1983) *PNAS USA* **80**, 1897–1901.
667. Schuler, M.A. et al. (1982) *Nucl. Acids Res.* **10**, 8225–8243.
668. Marco, Y.A. et al. (1984) *J. Biol. Chem.* **259**, 13436–13441.
669. Doyle, J.J. et al. (1986) *J. Biol. Chem.* **261**, 9228–9238.
670. Klosgen, R.B. et al. (1986) *Mol. Gen. Genet.* **203**, 237–244.
671. Rohde, W. et al. (1988) *Nucl. Acids Res.* **16**, 7185–7186.
672. Brown, J.W.S. (1986) *Nucl. Acids Res.* **14**, 9549–9559.
673. Breathnach, R. & Chambon, P. (1981) *Ann. Rev. Biochem.* **50**, 349–383.
674. Mount, S.M. (1982) *Nucl. Acids Res.* **10**, 459–472.
675. Lerner, M.R. et al. (1980) *Nature* **283**, 220–224.
676. Rogers, J. & Wall, R. (1980) *PNAS USA* **4**, 1877–1879.
677. Brantant, C. et al. (1981) *Nucl. Acids Res.* **9**, 841–858.
678. Lazar, E. et al. (1982) *Nucl. Acids Res.* **10**, 1193–1201.
679. Rinke, J. et al. (1984) *Nucl. Acids Res.* **12**, 4111–4126.
680. Ohshima, Y. & Gotoh, Y. (1987) *J. Mol. Biol.* **195**, 247–259.
681. Shapiro, M.B. & Senapathy, P. (1987) *Nucl. Acids Res.* **15**, 7155–7174.
682. Hanley, B.A. & Schuler, M.A. (1988) *Nucl. Acids Res.* **16**, 7159–7176.
683. Brantant, C. et al. (1980) *Nucl. Acids Res.* **8**, 4143–4154.
684. Lund, E. (1988) *Nucl. Acids Res.* **16**, 5813–5826.
685. Reddy, R. et al. (1974) *J. Biol. Chem.* **249**, 6486–6494.
686. Kato, N. & Harada, F. (1985) *J. Biol. Chem.* **260**, 7775–7782.
687. Mount, S.M. & Steitz, J.A. (1981) *Nucl. Acids Res.* **9**, 6351–6368.
688. Myslinski, E. et al. (1984) *J. Mol. Biol.* **180**, 927–945.
689. Latchman, D.S. (1988) *FEBS Letters* **236**, 5–8.
690. Siliciano, P.G. et al. (1987) *Science* **237**, 1484–1487.
691. Kretzner, L. et al. (1987) *Cell* **50**, 593–602.
692. Van Santen, V.L. & Spritz, R.A. (1987) *PNAS* **84**, 9094–9098.
693. Kiss, T. et al. (1988) *Nucl. Acids Res.* **16**, 2734.
694. Van Santen, V.L. et al. (1988) *Nucl. Acids Res.* **16**, 4176.
695. Green, M.R. (1986) *Ann. Rev. Genet.* **20**, 671–708.
696. Padgett, R.A. et al. (1986) *Ann. Rev. Biochem.* **55**, 1119–1150.
697. King, C.R. & Piatigorski, J. (1983) *Cell* **32**, 707–712.
698. Fischer, H.D. et al. (1984) *PNAS USA* **81**, 2733–2737.
699. Weber, S. & Aebi, M. (1988) *Nucl. Acids Res.* **16**, 471–486.
700. Watts, F. et al. (1983) *EMBO J.* **2**, 2085–2091.
701. Langford, C. et al. (1983) *PNAS USA* **80**, 1496–1500.
702. Van Santen, V.L. & Spritz, R.A. (1987) *Gene* **56**, 253–265.
703. Hartmuth, K. & Barta, A. (1986) *Nucl. Acids Res.* **14**, 7513–7528.
704. Ruskin, B. et al. (1986) *PNAS USA* **83**, 2022–2026.
705. Vidaud, M. et al. (1989) *PNAS USA*, in press.
706. Treisman, R. et al. (1983) *Nature* **302**, 591–596.
707. Lapomroulie, C. et al. (1986) *BBRC* **139**, 709–713.
708. Lapomroulie, C. et al. (1987) *Nucl. Acids Res.* **15**, 8195–8204.
709. Kazarian, H.H. et al. (1984) *EMBO J.* **3**, 593–596.
710. Rees, D.J.G. et al. (1985) *Nature* **316**, 643–645.
711. Dillella, A.G. et al. (1986) *Nature* **322**, 799–803.
712. Myerowitz, R. (1988) *PNAS USA* **85**, 3955–3959.
713. Ohno, K. & Suzuki, K. (1988) *BBRC* **153**, 463–469.
714. Arpaia, E. et al. (1988) *Nature* **333**, 85–87.
715. Weil, D. et al. (1987) *J. Biol. Chem.* **263**, 8561–8564.
716. Montell, C. & Berk, A.J. (1984) *Nucl. Acids Res.* **12**, 3821–3827.
717. Solnick, D. (1981) *Nature* **291**, 508–510.
718. Shine, J. & Dalgarno, L. (1975) *Nature* **254**, 34–38.
719. Steitz, J.A. & Jakes, K. (1975) *PNAS USA* **72**, 4734–4738.