Published November 15, 1991

Mini-Review

An Analysis of Vertebrate mRNA Sequences: Intimations of Translational Control

Marilyn Kozak

Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Piscataway, NJ 08854

Abstract. Five structural features in mRNAs have been found to contribute to the fidelity and efficiency of initiation by eukaryotic ribosomes. Scrutiny of vertebrate cDNA sequences in light of these criteria reveals a set of transcripts—encoding oncoproteins, growth factors, transcription factors, and other regulatory proteins—that seem designed to be translated poorly. Thus, throttling at the level of translation may be a critical component of gene regulation in vertebrates. An alternative interpretation is that some (perhaps many) cDNAs with encumbered 5′ noncoding sequences represent mRNA precursors, which would imply extensive regulation at a posttranscriptional step that precedes translation.

INITIATION of translation in multicellular eukaryotes is influenced by five aspects of mRNA structure: (*a*) the m7G cap (355); (*b*) the primary sequence or context surrounding the AUG codon (187, 190, 194); (*c*) the position of the AUG codon, i.e., whether it is the first AUG in the message (186); (*d*) leader length (198, 199); and (*e*) secondary structure both upstream (188, 195) and downstream (196) from the AUG codon. Elsewhere (200) I have reviewed the evidence for these five features and explained how they work together to determine the fidelity and efficiency of initiation. A scanning mechanism for initiation can explain many of the effects of cap, context, position, etc. The scanning model (193) in its simplest form postulates that a 40S ribosomal subunit, carrying Met-tRNA$_i^{met}$ and an imperfectly defined set of initiation factors (302), enters at the 5′ end of the mRNA and migrates linearly until it reaches the first AUG codon, whereupon a 60S subunit joins and the first peptide bond is formed. Evidence in support of the model has been adduced previously (62, 193, 197). More recent evidence for scanning includes the apparent queuing of 40S ribosomal subunits on long leader sequences (199) and the stalling of 40S subunits on the 5′ side of a stable hairpin structure introduced between the cap and the AUG codon (195). The possibility of initiation by a mechanism other than scanning has been proposed (158) and is evaluated elsewhere (197).

The trick to identifying elements in 5′ noncoding sequences that can modulate translation was to isolate each feature (200), an approach made possible by the techniques of genetic engineering. For example, by devising a transcript in which the first AUG codon was in an unfavorable context and hence "leaky," we were able to show that downstream secondary structure enhances recognition of the preceding AUG codon, apparently by preventing the 40S ribosomal subunit from scanning too fast or too far (196). The contribution of downstream secondary structure would have been missed had the primary sequence around the first AUG codon been more favorable, and vice versa. Having used the reductionist approach to identify several features that can modulate initiation, I attempt herein to put the story back together by examining the extent to which natural mRNAs conform to the experimentally determined requirements for initiation. A surprising realization is that, although most vertebrate mRNAs have features that ensure the fidelity of initiation (i.e., selection of the correct AUG codon), many do not appear to be designed for efficient translation. This would seem to have important implications for gene regulation.

Leader Sequences on Vertebrate mRNAs: An Overview

In considering the extent to which natural mRNAs conform to the five recognized requirements for initiation, I will focus on mRNAs from vertebrate cells where the rapidly expanding catalogue of sequences provides grist for analysis. mRNAs from animal and plant viruses and yeasts are mentioned only incidentally when they uniquely illustrate a point.

Every cellular mRNA that has been examined is capped (355). Not every mRNA has been examined, of course, but it seems unlikely that uncapped cellular mRNAs will be found inasmuch as the cap is crucial not only for translation (355) but also for mRNA stability and transport (126). Although the uncapped mRNAs from picornaviruses are a vogue topic for discussion (158), it should be remembered that picornaviruses are exceptions. All other animal viruses produce capped mRNAs, even when doing so requires the virus to encode its own capping and modifying enzymes (189).

The requirement for a favorable context around the AUG initiator codon is also met by nearly all mRNAs from higher eukaryotes. The consensus sequence for initiation derived from a compilation of 699 vertebrate mRNAs (191) is GCC$^{A}_{G}$CCAUGG, the same as the experimentally derived optimal sequence (187, 190, 194). While the full consensus sequence is found in only a small number of vertebrate mRNAs, the two positions most critical for function (187) are highly conserved: 97% of vertebrate mRNAs have a purine (usually A) in position −3 and 46% have G in position

Published November 15, 1991

+4 (191). (The A of the AUG codon is designated +1, with positive and negative integers proceeding 3′ and 5′, respectively.) Since context is a principal determinant of the fidelity of initiation, good adherence to the context rules ensures that the full-length protein is the sole translation product from most vertebrate genes. Only six out of 699 mRNAs in the aforementioned survey (191) lacked the preferred nucleotide in both of the key positions flanking the AUG codon. One of the original six entries has since been discounted as a sequencing error (212) but a few more mRNAs have recently been added to the list (154, 180, 311). These rare mRNAs with highly unfavorable initiation sites encode potent regulatory proteins (growth factors, cytokines, etc.), suggesting that a weak context might be an occasional ploy to modulate the yield of proteins that could be harmful if overproduced. Parenthetically, a recent compilation (unpublished results) of 252 plant mRNA sequences reveals that 93% have a purine in position −3, and 74% have G in position +4; thus, in the two most influential positions, plant and animal consensus sequences are the same.

The importance of position of the AUG codon in determining the site of initiation is illustrated by a family of bifunctional genes, described in the next section, and by the fact that the first AUG codon is the unique initiation site in most (perhaps 90% of) vertebrate mRNAs (191). The list of mammalian cDNA sequences that violate the first-AUG-rule is growing, however, and it includes a large number of critical regulatory genes, as documented below. If these AUG-burdened cDNA sequences actually correspond to functional mRNAs, their translation should be compromised. If the AUG-burdened cDNA sequences correspond instead to mRNA precursors or otherwise nonfunctional transcripts (as has been established in some cases), their abundance implies considerable regulation at a level other than translation. Both possibilities are discussed below.

A handful of natural mRNAs, mostly of viral origin, seem to have very efficient 5′ noncoding sequences as documented by leader-shuffling experiments (84, 237, 251, 342, 360). Attempts to pinpoint a motif in any of those sequences that underlies its efficient translation have been notably unsuccessful; that is, virtually every portion of the 5′ noncoding sequence has been mutated or deleted or replaced without impairing translation. The apparent absence of a discrete effector motif, and the fact that the sequences in question are deficient in G residues, fit with the view that a moderately long, unstructured 5′ noncoding sequence may be necessary and sufficient for efficient initiation of translation (199). Indeed, with some experimental constructs, the simple trick of lengthening the 5′ noncoding sequence improves translation by an order of magnitude, creating very efficient in vitro expression vectors (199). (The 5′ sequences were lengthened by reiterating three different synthetic oligonucleotides which were designed simply to preclude secondary structure. It seems unlikely that such arbitrarily designed sequences are recognized by hypothetical "enhancer proteins." Rather, the observed accumulation of extra 40S ribosomal subunits on long 5′ leader sequences (199) may underlie their translational advantage.) The advantage conferred by long, synthetic leader sequences does not hold for most naturally occurring leaders, however: 5′ noncoding sequences hundreds of nucleotides long are not uncommon on vertebrate mRNAs, but their remarkably high GC content implies that they are highly structured; and a structured leader sequence, be it long or short, is a major barrier to translation (188, 195). The frequent presence of such sequences on mRNAs from critical regulatory genes (see below) has notable implications for gene regulation in vertebrates. These considerations do not extend to mRNAs from plants or budding yeasts, which usually have AU-rich, rather than GC-rich, leader sequences.

Bifunctional Genes and Bifunctional mRNAs

The importance of position in determining the functional initiator codon is illustrated by a family of genes that are required to produce two versions of the encoded protein (Fig. 1 *A* and Table I). The general idea is that ribosomes need to initiate translation from the first and second AUG codons in each of these genes. Although the longer mRNA from the model gene in Fig. 1 *A* contains both AUG codons, the presence of a good context around the first AUG codon precludes access to the second. The solution is that the gene produces a second form of mRNA, the 5′ end of which maps between the two AUG codons. In each mRNA ribosomes initiate at the first and only the first AUG codon.

There is a way for two proteins encoded in overlapping open reading frames (ORFs) to be translated from a single mRNA; namely, by introducing a poor context around the first AUG codon. In the bifunctional mRNAs listed in Table

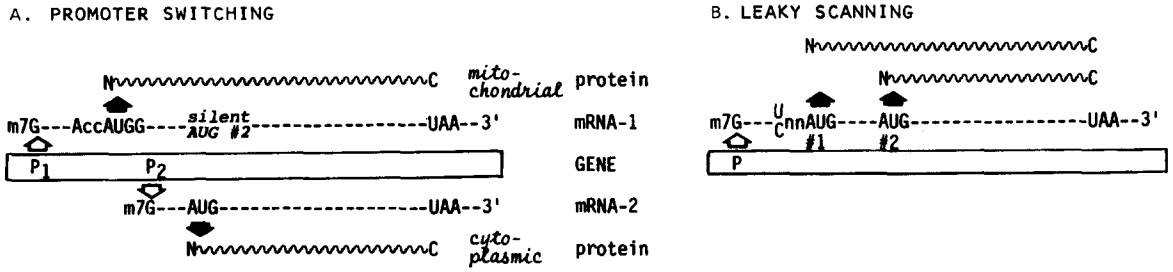

Figure 1. Two mechanisms that enable one gene to produce two versions of the encoded protein. (*A*) When the first AUG codon is in a strong context, as is usually the case, one gene can produce two proteins only by producing two mRNAs, i.e., by initiating one transcript (P_1) upstream from the first AUG codon and initiating a second transcript (P_2) downstream from that AUG. Often the NH_2-terminal amino acid extension targets the long form of the protein to a special intracellular compartment. Examples are given in Table I. (*B*) Leaky scanning permits synthesis of two proteins from one mRNA when the context around the first AUG codon is unfavorable; i.e., when a pyrimidine occurs in position −3, or when there is a G in position −3 and something other than G in position +4. Examples of genes that use leaky scanning are given in Table II.

Published November 15, 1991

Table I. Genes that Produce Two Overlapping Proteins by Initiating Transcription from Two Promoters, as Illustrated in Fig. 1 A

[a] Val-tRNA synthetase (VAS1, yeast) (49)
[a] His-tRNA synthetase (HTS1, yeast) (282)
[a] α-Isopropylmalate synthase (LEU4, yeast) (18)
[a] tRNA dimethyltransferase (TRM1, yeast) (90)
[a] Serine:pyruvate aminotransferase (SPT, rat) (291)
[a] Cyclophilin (N. crassa) (398, 399)
Anion transport protein (band 3, chick) (177)
β1→4-galactosyltransferase (bovine) (234a, 331)
[b] Invertase (SUC2, yeast) (43)
[b] Gelsolin (human) (207)
[b] Surface antigen, Hepatitis B virus (307)
[c] E2 protein, bovine papillomavirus (211)
Family 35 capsid proteins, herpes simplex (228)
[c] Progesterone receptor (forms A & B, human) (167)
Sterol carrier protein (SCP_X & SCP_2, rat liver) (349)
Porphobilinogen deaminase (human) (55)
$G_S\alpha$ protein (human) (156)
Erythroid membrane protein 4.1 (human) (68)

Superscripts refer to ways in which the function of the long isoform differs from that of the shorter protein: (*a*) import into mitochondria, (*b*) secretion, (*c*) control of transcription. The expression of porphobilinogen deaminase, erythroid membrane protein 4.1, and $G_S\alpha$ protein requires alternative splicing as well as promoter switching; they are included because the net effect is activation of an internal AUG initiator codon by making it the first AUG in the mRNA. Not listed are some interesting genes that produce small amounts of 5′ truncated transcripts in extraneous tissues (63, 160, 234, 405, 410). Even when such mRNAs can be shown to direct synthesis of a polypeptide fragment (63), which is almost inevitable if the transcript enters the cytoplasm, the phenomenon might reflect inadvertent expression; what needs to be established is that the NH_2 terminally truncated polypeptide serves a unique function in the ectopic tissue.

II, the first AUG codon deviates from the consensus sequence in either or both of positions −3 and +4. (Three exceptions are discussed in the Table II legend.) The result is "leaky scanning" in which some 40S ribosomes bypass the first AUG codon; initiation occurs from the first and second AUG codons in these mRNAs (Fig. 1 *B*). Curiously, most of these bifunctional mRNAs are of viral origin. Only two cellular mRNAs are listed in Table II, and in neither of those cases has the short protein been shown to mediate a function distinct from the long isoform. Thus, as a practical device for producing two proteins from one gene, cells rely mostly on a transcriptional device (Fig. 1 *A*) while viruses use a translational ploy (Fig. 1 *B*). Leaky scanning does not require virus-induced modifications of the translational machinery, however, inasmuch as the isolated reovirus S1 gene, when expressed in uninfected COS cells, produces the expected two proteins (95). Leaky scanning may also result when the first AUG codon resides close to (within 12 or so nucleotides of) the cap (113, 198), although leakiness due to an unfavorable context is the more common mechanism.

Not included in Table II are a few bifunctional mRNAs (344, 347, 397) in which two proteins are produced from nonoverlapping ORFs. In such cases, reinitiation as well as leaky scanning may provide access to the second AUG codon, and the contributions of the two processes are hard to sort out. As explained elsewhere (192, 200), reinitiation by eukaryotic ribosomes is usually inefficient and seems to occur only when the 5′ proximal ORF is small. Those restrictions probably explain why no bifunctional mRNA has been identified in animal cells or viruses that relies exclusively on reinitiation for expression of the downstream cistron. There are quite a few viral mRNAs that are structurally bicistronic, encoding two full-length proteins in nonoverlapping ORFs, but they are functionally monocistronic, translating only the 5′ proximal ORF (189). In the case of Epstein-Barr virus, a bicistronic mRNA that encodes both the R and EB1 proteins (in that order) does appear capable of translating EB1, albeit inefficiently; however, the virus also produces an abundant transcript that encodes (and translates efficiently) only EB1 (243). The same is true for synthesis

Table II. Genes that Produce Two Proteins from One mRNA by Leaky Scanning, as Illustrated in Fig. 1 B

A. Initiation at 1st and 2nd AUGs generates long and short protein isoforms from the same reading frame:
Simian virus 40 late 19S mRNA - VP2, VP3 (346)
Rotavirus SA11, segment 9 - 37K, 35K, (VP7) (47, 370)
West Nile flavivirus - V2 core proteins (44)
Dengue (type 3) flavivirus - C, C′ (300)
Foot-and-mouth disease virus - p20a, P16 (64)
Hepatitis B virus, human - pre-S, $p24^S$ (306)
Feline leukemia virus - $gPr80^{gag}$, $Pr65^{gag}$ (214)
Rift Valley fever (bunya)virus - M proteins (378)
Cucumber necrosis virus - p21, p20 (325)
Cowpea mosaic virus RNA-M - 105K, 95K (144)
Barley stripe mosaic virus - βb, βb′ (308)
Creatine kinase, chicken brain (363)
N-myc, human tumor cell lines (240)

B. Initiation at 1st and 2nd AUGs in different, overlapping reading frames produces two unrelated proteins:
Sendai (paramyxo)virus - P, C (78)
Reovirus S1 mRNA - σ1, 14K (92, 95)
Bunyavirus s-RNA - N, NS_S (109)
Adenovirus, region E1B - 21K, 55K (27)
Adenovirus, region E3 - 6.7K, gp19K (415)
Human T-Cell Leukemia Virus (HTLV-I)- p27, p40 (279, 359)
Human Immunodeficiency Virus Type I - vpu, env (345)
Potato leafroll luteovirus - CP, 17K (380)
Satellite tobacco mosaic virus - 6.8K, 17.5K (261)
Rotavirus SA11, segment 11 - 28K, 11K (249, 263)
*Turnip yellow mosaic (tymo)virus - 69K OP, 200K RP (170)
*Maize chlorotic mottle virus - p31.6, p50 (290)
Influenza B virus - NB, NA glycoproteins (414)

In several cases the connection between leaky scanning and a suboptimal context around the first AUG codon has been confirmed by mutational analysis (78, 95, 346, 370). Only three cases have been described in which ribosomes initiate at the first and second AUG codons despite a favorable context around the first AUG. The most important of these exceptions is influenza virus B, where the proximity of the second AUG codon to the first AUG seems to allow leaky scanning (414). The other exceptions are barley stripe mosaic virus (308) and cowpea mosaic virus RNA-M (144), where the absence of secondary structure downstream from the first "strong" AUG codon might account for the leakiness. Leaky scanning in those two plant viruses might be inadvertent, inasmuch as the second protein isoform does not contribute to viral infectivity. In contrast, for most of the other viral entries, both proteins produced from the bifunctional mRNA are required for infectivity. The fact that the overlapping arrangement of ORFs is conserved among different members of the paramyxovirus, reovirus, bunyavirus, adenovirus, rotavirus, and tymovirus families constitutes additional evidence that the synthesis of two proteins by leaky scanning is not accidental in those cases. In the case of HIV-1, *vpu* functions more efficiently in promoting the processing of *env* when the two proteins are translated from the same mRNA than when they are expressed experimentally from separate transcripts (M. Martin, K. Strebel and R. Willey, personal communication). An asterisk, preceding some entries in the table, means that only one of the two proteins predicted by the mRNA sequence has been detected so far.

Published November 15, 1991

of the LP and EBNA-2 proteins of Epstein-Barr virus (5). The functionality of an interesting bicistronic transcript for growth/differentiation factor 1 has not yet been established. (Although the only detected transcript for GDF-1 in 14.5-d mouse embryos is a 3-kb bicistronic transcript in which GDF-1 is the downstream cistron, the GDF-1 protein detected by immunohistochemical analysis in 14.5-d embryos might actually have been synthesized a few days earlier, when a 1.4-kb transcript was the predominant form (218).)

The ability to reinitiate to some extent after translating a small 5′ ORF, and the tendency of 40S ribosomes to scan past an AUG codon in a weak context, explain how ribosomes can initiate from an AUG codon that is not first. Nevertheless, the occurrence of upstream AUG codons nearly always reduces the efficiency of initiating from downstream. Thus, mRNA (or, more correctly, cDNA) sequences that are peppered with small upstream ORFs pose a problem.

5′ Noncoding Exons, Introns, and Upstream AUG Codons

The simple question of whether the mRNA from a particular gene has upstream AUG codons cannot always be answered simply. Some of the complexities are due to 5′ noncoding exons and associated phenomena such as alternative splicing, inefficient removal of a 5′ intron, and the presence of alternative promoters. I will first address those complications and then try to assess the frequency and significance of upstream AUG codons in vertebrate mRNAs.

Nearly one-fourth of the entries in a recent survey of 699 vertebrate mRNA sequences (191) have turned out to have an intron between the promoter and the start of the major ORF. The high incidence of 5′ introns has theoretical as well as practical consequences. The first intron in a gene sometimes contains sequences that facilitate transcription (26, 42, 60, 73, 150, 174, 176, 185, 269, 280, 287, 303, 368), an effect that sometimes *requires* the intron to be near the 5′ end (42, 287, 303). Some genes that have retained a 5′ intron thereby have the ability to switch promoters, in response to hormonal or tissue-specific inducers, for example, and thus to exchange an inefficiently translated 5′ noncoding exon for one that appears more favorable (9, 54, 77, 106, 289). (The predicted improvement in translation has not yet been verified for all of those genes.) Another kind of regulation takes the form of allowing a gene to be transcribed in an ectopic tissue but preventing its translation by not removing the 5′ intron. The expression of gonadotropin-releasing hormone mRNA in extra-hypothalamic tissues is a striking example (135). Another might be the expression of the tyrosine kinase *fer* gene in testis versus other tissues (101).

From a practical perspective, the frequent presence of 5′ introns (sometimes directly abutting or even interrupting the AUG codon) necessitates caution in picking a probable start site for translation, and caution in scoring upstream AUG triplets. Many claims of mRNAs with AUG-burdened 5′ noncoding sequences (46, 114, 125, 213, 236, 238, 292, 348, 387) have been resolved by finding that the 5′ portion of the cDNA corresponds not to the mature mRNA but to an intron-containing form (36, 135, 225, 235, 284, 316, 388, 407, 408). (see reference 193 for additional examples.) The growing evidence of incomplete (20, 23, 28, 74, 76, 80, 138, 151, 168, 222, 259, 343, 354, 411) or regulated (417) RNA processing in mammalian cells underscores the point that cDNA sequences cannot invariably be equated with functional mRNAs. Some intron-containing transcripts are abundant (70), some enter the cytoplasm (328), and some are even found on polysomes (411). These problems complicate attempts to deduce the real structures of vertebrate mRNAs. While a cDNA sequence that retains an unspliced intron within the coding domain is easily recognizable as a processing intermediate, the presence of unspliced intron(s) in the 5′ noncoding domain is much harder to recognize.

Some genes are transcribed in ectopic tissues from an ill-placed promoter that burdens the 5′ noncoding sequence with AUG codons, thereby impairing translation in that particular tissue. In the tissue that constitutes the major site of expression, however, a different promoter produces a 5′ noncoding sequence that is not so encumbered. Examples include murine complement-B mRNA in hepatic versus extrahepatic tissues (286), rat preproenkephalin mRNA in testis versus brain (172, 175), rat α-crystallin mRNA in brain versus other tissues (157), and rat farnesyl pyrophosphate synthetase mRNA in liver versus testis (390). In the last three cases, the predicted difference in translational efficiency has been verified experimentally. Another ploy involves switching to a shorter, more efficiently translated leader sequence in response to some developmental (61, 425) or environmental cue, such as stimulation with serum (409) or retinoic acid (61) or endotoxin (286), or during T cell maturation (324). Because of promoter switching and/or alternative splicing, many other vertebrate genes produce multiple transcripts that differ near the 5′ end, and failure to detect all pertinent forms has sometimes led to false conclusions. The suggestion of "internal initiation" in the chicken progesterone receptor mRNA (71) is one example of a wrong conclusion that was righted upon discovering other forms of mRNA (166, 167). Detecting alternative transcripts is not always easy! Competition for the primer may cause a minor transcript to be missed (175). Even the major transcript has been missed when the primer was positioned inappropriately (72) or when hybridization conditions were too stringent (338).

Because of the difficulties described above and various other complications in cloning or interpretation (6, 32, 37, 68, 69, 146, 226, 430), the frequency of spurious upstream AUG codons in vertebrate mRNAs is difficult to estimate; but clearly it is not as high as superficial reading of the literature might suggest. When upstream AUG codons do occur, the AUG-burdened leader sequence impairs translation (9a, 12, 105, 157, 247, 272, 277, 281, 326, 390, 409, 412, 424), as expected if initiation occurs by the conventional scanning process.

A partial listing of cDNAs with AUG-burdened leader sequences is given in Table III. It includes many proto-oncogenes as well as genes for transcription factors, a variety of receptor proteins, signal transduction components, and many proteins involved in the immune response. One conclusion might be that mRNAs that encode critical regulatory proteins are intended to be translated poorly. I suspect that conclusion is true for some entries in the table, but some (perhaps many) entries might reflect a different type of regulation. For example, the repeated finding of incompletely spliced transcripts in lymphocytes (20, 107, 389, 411) and recent evidence that undefined posttranscriptional processes improve upon mitogen activation of lymphocytes (67) en-

Published November 15, 1991

Table III. Vertebrate cDNA Sequences that Have Three or More AUG Codons Upstream from the Major Open Reading Frame

Tumor associated (proto-oncogenes, etc.)	Transcription factors and DNA-binding proteins	Receptors for the following ligands
* *abl*, human (21)	NF1-B (TGGCA), chicken (330)	* Acetylcholine, rat (224)
* *bcl-2*, human (401)	NF1-X, hamster (112)	* Angiotensin II, rat (276)
ear-7, human (266)	cAMP response (CRE-BP1) (278)	‡ Atrial natriuretic peptide (110)
* *erb-A*, human (413)	IFN response (IRE-BF1) (421)	D_1 dopamine, mouse (268)
erg, human (320)	PRDII-BF1 (96)	Estrogen, chicken (203)
Evi-1, human (271)	‡ DBP, rat liver (272)	$GABA_A$ α1, mouse (171)
Evi-2, mouse (38)	HOX 2G, human (2)	$GABA_A$ γ2, chicken (115)
* *fgr*, human (122)	* HOX 5.1, human (61)	Gastrin-releasing peptide (13)
‡ *fos-B*, mouse (427)	Hox 2.9, mouse (274)	Glycine, human brain (119)
HCK, human (318)	Hox 3, mouse (31)	Heparin-binding GF (K-sam) (133)
*‡ *int-2*, human (34)	Hox 3.1, mouse (10)	Interferon α, human (402)
* *lck*, mouse (324)	Hox 3.2, mouse (93)	Progesterone, rabbit (262)
* *mos*, mouse (314)	* BTF3 (general) (428)	Prolactin, rabbit (89)
ROS-1, human (22)	OTF-2, human (339)	Retinoic acid (hRAR-α) (30)
*‡ *sis*, (PDGF-2) (98, 319)	TFE, canine (161)	* Retinoic acid (mRAR-β) (426)
* *sno*, human (285)	KUP, human (48)	* Retinoic acid (hRAR-γ) (204)
syn (*slk*), human (352)	polI factor UBF, rat (296)	Serotonin 1c, rat (164)
* T-cell 11p15 (25)	HNF-1β, mouse (256)	Serotonin 5HT-2, rat & CHO (45, 313)
		Substance K, bovine, human (111, 248)
		Substance P, rat (136)
		‡ Thromboxane A_2, human (140)
		‡ Thyroid hormone, rat (277)
		Rev-ErbAα, rat (216)
		Thyroid hormone, rat (394)
		* Thyroid hormone, human (391)

Immune/inflammation mediators	Signal transduction	Growth factors
‡ Interleukin-7, mouse (239, 281)	chick embryo tyr kinase (304)	Keratinocyte GF (99)
IL-1 receptor, mouse (358)	mouse liver tyr kinase (244)	* Insulin-like GF-1 (17, 105)
IL-2 receptor-β, human (132)	FER tyr kinase (130)	Platelet-derived GF-A (328)
IL-3 receptor, mouse (116)	*tyk2* tyr kinase (100)	Epidermal GF (16)
IL-5 receptor, mouse (382)	*elk* tyr kinase (223)	‡ Transforming GF-β3 (9a, 209)
IL-6 receptor, human (420)	ERK3 ser/thr kinase (29)	*See also: proto-oncogenes*
IL-7 receptor, mouse (309)	p58 protein kinase (173)	
G-CSF receptor, mouse (108)	rp-S6 kinase, chick (4)	
C3b receptor (Mac-1α) (315)	protein tyr phosphatase (PTPase) LRP (293)	
CD28, human T cells (217)	PTPase, megakaryocyte (121a)	
CD75, human B cells (365)	Phospholipase C-I, rat (375)	
* Ly-5 (CD45, CALLA) (334)	cAMP phosphodiesterase (230)	
Ttg-1, T cells (252)	insulin receptor substrate-1 (IRS-1, pp185) (376a)	
Surface antigen 114/A10 (87)		
Tyr kinase, leukocyte (19)		
βFcγ receptor II, mouse (143)		
* IgE receptor (high affin) (229)		

Some, perhaps many, of these cDNA sequences are likely to represent mRNA precursors rather than functional mRNAs (see text). The literature contains scattered reports of AUG-burdened cDNA sequences in addition to those listed here. *, The gene produces multiple transcripts with alternative 5′ noncoding sequences; ‡, translation is more efficient with transcripts (natural or derived) that lack the encumbered leader sequence.

courage the idea that many cDNA sequences from immune cells might correspond to precursors rather than to functional mRNAs. The same may be true of AUG-burdened transcripts from transcription factor genes, since some of those transcripts are restricted to the nucleus (61, 70, 275); in other cases, a transcript is detectable but the corresponding protein is not (256). The first report of mitogen-regulated splicing of 5′ introns in vertebrate genes has just been published (417), giving substance to the hypothesis that nontranslatable transcripts may be synthesized and stored for later processing.

Proto-oncogenes, on the other hand, might be genuine candidates for translational modulation via an encumbered leader sequence. Interpretation is complicated by the fact that many proto-oncogenes produce transcripts with alternative 5′ sequences (these are marked by asterisks in Table III), but several observations support the idea that proto-oncogene mRNAs are meant to be translated inefficiently: c-*mos* transcripts are found on very small polysomes (314); some activated oncogenes produce transcripts with simpler 5′ noncoding sequences than the corresponding proto-oncogenes (321, 337); and the experimental expression of many proto-oncogenes improves dramatically upon deleting portions of the leader sequence (12, 50, 247, 319). To propose that proto-oncogene mRNAs might be translated by a mechanism other than scanning (253), inasmuch as their

Published November 15, 1991

AUG-burdened leader sequences seem incompatible with efficient scanning, is to miss the point that these potent proteins probably *have to be translated inefficiently.*

Occurrences and Consequences of Secondary Structure

The catalogue of vertebrate mRNAs with GC-rich (hence highly structured) leader sequences again includes many mRNAs for oncoproteins, growth factors, transcription factors, signal transduction components, and a wide variety of receptor proteins (Table IV). Again, the presence of an encumbered leader sequence suggests that production of these critical regulatory proteins is throttled at the level of translation. The GC-cohort also includes many housekeeping genes, which are generally recognized to be expressed at low levels. While it is easy to show that many of these leader sequences support translation poorly (see below), delineating the cause is not simple. The extraordinarily high GC content (70 to 90%) predicts many alternative base pairings, making it impossible to pinpoint a target for mutagenesis. Consequently, our understanding of how particular base-paired structures affect translation relies heavily on experiments carried out with synthetic transcripts (188, 195, 196) in which discrete stem-and-loop structures have been introduced and their existence documented by genetic techniques.

The best evidence that mRNAs in Table IV are translationally impaired is the dramatic improvement in expression when the GC-rich leader sequences (some of which also contain upstream AUG codons) are truncated experimentally (53, 82, 182, 221, 246, 273, 319, 395). (Discrepancies between mRNA levels and protein accumulation in some stages or conditions of cell growth may be another indication of translational control of transcription factor and other such genes (66, 81, 83, 205, 340, 357, 373); but in most of those cases alternative explanations, such as compartmentalization of the mRNA or accelerated degradation of the protein, have not been ruled out.) Some genes in Table IV actually produce two versions of mRNA, on one of which the leader is shorter and less encumbered than on the other (178, 283, 317, 377). In the few cases where long- and short-leader mRNAs from the same gene have been put to the test, the short-leader transcript nearly always supports translation more efficiently (153, 283, 337, 409). Indeed, the discrepancy in translatability dependent on 5′ leader sequences can be so profound that a minor transcript from certain genes appears to be the major functional mRNA (149, 264, 283). Other genes in the GC-rich cohort produce transcripts with so many different leader sequences (75, 85a, 165, 220, 245, 426) that it is impossible to guess, and no small task to test, their functionality.

Notwithstanding those caveats, the extraordinary number of mRNAs with GC-burdened leader sequences forces the idea that synthesis of critical cellular proteins is probably throttled at the level of translation. Under constitutive conditions, the synthesis of a single molecule of such a protein could conceivably take hours as a 40S subunit slowly maneuvers its way to the downstream AUG codon. If slow initiation of translation is a key to limiting the production of proteins that would be lethal if overproduced, one should not be surprised that such mRNAs are virtually untranslatable in standard in vitro assays in which mRNAs are expected to produce a product in minutes! A compelling rationale for the cumbersome 5′ noncoding sequences on so many regulatory genes is that those transcripts should respond *as a cohort* to shifts in the cell's translational capacity. As for how hypothetical shifts in translational capacity might be accomplished, changes in the extent of phosphorylation of initiation factors and ribosomal proteins have often been remarked (137, 393). With the notable exception of eIF2 (65), however, hard evidence for the functional consequences of phosphorylation remains elusive.

A structured leader sequence may have qualitative as well as quantitative effects on translation. In a small number of vertebrate mRNAs, ribosomes initiate at a non-AUG codon, such as ACG, CUG, or GUG (3, 20, 86, 103, 127, 219, 232a, 336). The list is slightly longer if one counts viral mRNAs (15, 78, 312, 367). It is not valid, however, to count mRNAs in which the use of alternative initiator codons has been documented only in vitro, where inappropriate reaction conditions can activate cryptic sites that would not be used in vivo (194). Initiation at non-AUG codons is usually inefficient and usually occurs in addition to using the first AUG codon. The result is synthesis of an "extra" NH_2 terminally extended version of the protein. (There are only two instances in which a protein derives uniquely from initiation at an upstream non-AUG codon and not, at least in part, from the first in-frame AUG codon. One occurs in cells transfected with *ltk* tyrosine kinase cDNA, in which five out-of-frame AUG triplets occur between the putative CUG initiation site and the first in-frame AUG (20); initiation at the far-upstream CUG codon thus circumvents the problem of getting past out-of-frame AUG codons. As yet, however, initiation at the upstream CUG codon has not been demonstrated with the endogenous *ltk* gene in untransfected cells. The other very intriguing example is the apparently unique use of an AUU codon to initiate translation of the human enhancer factor TEF-1 (416).) All of the vertebrate mRNAs that use a nonstandard initiator codon have GC-rich leader sequences, prompting the speculation that the slow transit of scanning 40S ribosomes across a highly structured 5′ noncoding sequence might be responsible for activating cryptic upstream sites (196). Indeed, initiation at upstream non-AUG codons in synthetic transcripts was considerably enhanced upon introducing secondary structure in an appropriate position 3′ of the cognate initiator codon (196). While the NH_2 terminally extended polypeptides initiated from non-AUG sites in viral and cellular mRNAs occasionally have distinct functions (15, 312, 367) or distributions (3, 40, 232a), it would be simplistic to assume that every instance of initiation from a cryptic upstream site is functionally important. Given the GC-richness of leader sequences on mammalian mRNAs, spurious upstream initiation events may be unavoidable byproducts of the way eukaryotic ribosomes arrive at the AUG codon. In avian cells, the efficiency of initiating at the upstream CUG codon in c-*myc* mRNA is regulable by culture conditions (Stephen Hann, personal communication). This suggests interesting modulation of the translational machinery, but it does not aid the so-far unsuccessful effort to ascribe functional significance to the NH_2 terminally extended form of c-*myc*.

Many vertebrate mRNAs that have highly structured leader sequences also have upstream AUG codons (see the entries marked ‡ in Table IV). This coincidence might be viewed in either of two ways. One rationalization invokes the

Published November 15, 1991

Table IV. Some Vertebrate Genes Predicted to Have Highly Structured 5' Noncoding Sequences

Tumor associated (proto-oncogenes, etc.)		Receptors for the following ligands
‡ *bcl*-3, human (294)	‡ L-*myc*, human (169)	‡ n-acetylcholine, α5 (28)
‡§ *BCR*, human (353)	‡* N-*MYC*, human (366)	‡ n-acetylcholine, β2 (7)
bmk, mouse (145)	*pim*-1, human (323)	‡ α2B-adrenergic, rat (102)
DBL, human (94)	* H-*ras*, human (148)	‡ β_1-adrenergic, rat (356)
‡* *erbA*-1, human THRA (215)	Ki-*ras*, mouse (142)	‡§ β_2-adrenergic, human (91)
erbB (HER1) (124)	*RB*, human (147)	‡ α1A-adrenergic, human (233)
eph, human (139)	*rel*, chicken (128)	‡§ atrial natriuretic peptide (53)
ets-2, human (250)	‡ ret (381)	‡ dopamine D2, rat (267)
fos, human (372)	‡*§ *sis* (PDGF-2) human (98, 319)	‡ *ear*-2 (THR family) (265)
‡ *hck*, mouse (232)	*ski*, human (285)	‡* estrogen, human (118)
‡ HER2 (*neu*) human (384)	‡ *Spi*-1, human (322)	‡ HER3 (EGFR-related) (310)
‡ *hst*, human (423)	* *src*, chicken (85a)	‡ insulin, human (351)
‡ *int*-1, human (406)	‡ *syn*, human (352)	IGF-II, human (270)
‡§* *int*-2, human (34, 120)	GA733 antigen (227)	integrin, chicken (385)
jun, human (8)	GA733-2 antigen (379)	interferon-γ (117)
‡ KS (Kaposi) (79)	*mdm*-1, mouse (361)	‡ interleukin-1 (59)
lyl-1, human (255)	*mdr*-2, mouse (121)	mannose-6-phosphate (231)
lyn, human (419)	*timp*-2, human (369)	N10 (TH receptor family) (333)
myb, mouse (362)	‡ Wilms' tumor (WTI) (39)	nerve growth factor (350)
	(*See also: growth control, receptors*)	poliovirus, human (184)
		‡* progesterone, chicken (162)
		‡* retinoic acid (hRARγ) (204)
		ryanodine, rabbit (383)
		syndecan, human (242)
		thrombin, human (404)
		‡* transferrin, human (341)
		tumor necrosis factor (183)
		(*See also: proto-oncogenes*)

Transcription factors and DNA-binding proteins	Signal transduction	Growth control
RAP30/74 (general) (364)	GAP, bovine (403)	TGF-α, rat (24)
‡ NF-κB, human (258)	G protein G_α11, mouse (374)	* TGF-β1, human (178, 317)
PU.1, mouse B cells (181)	G protein $G_s\alpha$, human (201)	‡ TGF-β1 masking protein (400)
EFI_A CCAAT-BP, rat (301)	G protein $G_i\alpha$, human (14)	TNF-β, human (257)
CCAAT/EBP, mouse (57)	‡ G protein Gα12, mouse (374a)	CSF-1, human (208)
‡ NF-E1b, chicken eryth. (418)	G protein Gα13, mouse (374a)	interleukin 11, human (305)
GATA-1 (NF-E1), human (429)	G protein β_2, human (104)	endothelial GF (hPD-ECGF) (123)
GATA-1 (NF-E1), chick (129)	‡ adenylyl cyclase, bovine (202)	‡ *Egr*-1, mouse (376)
GATA-3, human (141)	protein kinase C (PKCα) (329)	‡ erythropoietin, human (159)
‡ *Krox-24*, mouse (219)	‡ PKC-ζ, rat brain (297)	basic fibroblast GF, human (1)
jun-D, mouse (332)	‡ PKC-L, human (11)	‡§ fibroblast GF-5, mouse (12, 134)
‡ zif/268, serum induced (56)	‡ nPKC, rabbit brain (295)	* IGF-II, human (155)
hepatocyte NF-1, mouse (205)	cAMP-dependent PKCα, hu (241)	IGF-binding prot-1, human (33)
‡ *Cdx-1*, mouse homeobox (88)	cAMP-dependent PKCβ, mu (58)	IGF-binding prot-2, rat (35)
LAP, rat liver (81)	CaM-kinase II, rat (396)	‡* PDGF-A, human (328)
LRF-1, regen. liver (152)	*ltk* tyr kinase, mouse (20)	Schwannoma-derived GF (180)
‡* SCL, human (9)	‡ PTPase (R-PTP-α), mouse (335)	TAPA-1, human (299)
Oct6, mouse embryo (254)	PTPase, human placenta (52)	(*See also: proto-oncogenes*)
‡ Serum response (SRE-BF) (288)	prot. phosphatase 2A, pig (371)	
TFIIS (elongation) (422)	prot. phosphatase 2C, rat (386)	
‡ TCF-1, murine T-cells (298)	phospholipase C-γ, human (41)	
‡ TEF-1, human EBP (416)		
SCIP rat nerve (204a)		

The leader sequences on cDNAs from these genes have a G+C content of 70 to 90%, which would seem to imply extensive secondary structure. Most but not all genes for oncoproteins, receptors, transcription factors and signal transduction components belong to this GC-cohort. Overall, only 19% of the vertebrate mRNAs compiled in reference 191 have GC-rich leader sequences. Of these, ~40% fall into the categories listed in Table IV, although oncogenes, receptors, transcription factors, signal transduction, and growth factor genes constitute only 13% of the total sequences in the compilation. Another ~30% of the GC-burdened leader sequences in reference 191 derive from mRNAs for cytoskeletal and housekeeping proteins. Thus, GC-rich 5'-noncoding sequences are not randomly distributed among vertebrate genes. The frequency of GC-rich mRNAs does appear to be increasing, however, now that technical improvements enable the routine cloning of cDNAs derived from scarce transcripts. *, The gene produces multiple transcripts with alternative 5' noncoding sequences; ‡, the GC-rich leader sequence also contains upstream AUG codon(s); §, translation improves upon deleting portions of the leader sequence.

Published November 15, 1991

adage that nothing bad can happen to a rotten eggplant: a highly structured 5′ sequence is so inhibitory to translation that the further slight diminution attributable to one or two small upstream ORFs should hardly matter. A more interesting view is that upstream ORFs (initiating at AUG or AUG-cognate codons) might actually be necessary to mitigate the inhibitory effects of a GC-rich leader sequence. The argument here is that *80S* ribosomes engaged in translating the upstream ORFs might be able to penetrate duplex structures that are too stable to be penetrated by scanning 40S ribosomal subunits—an idea which is supported by some evidence from experimental constructs (179, 195). If a smattering of upstream initiator codons indeed facilitates the translation of mRNAs with highly structured leader sequences, they probably provide only a small measure of relief. The major experimental finding is that mRNAs with long, GC-rich leader sequences are translated inefficiently. And a surprising number of vertebrate mRNAs fit that bill.

Coda

The usually favorable context around the AUG start site in vertebrate mRNAs ensures the fidelity of initiation. Because recognition of the AUG codon is a late event in the initiation process, however, a good context should not, and does not (194), affect the ability of one mRNA to outcompete another. Translational efficiency, defined as competitive ability, is probably determined instead by accessibility of the capped 5′ end of the mRNA, since the 5′ end constitutes the apparent entry site for the 40S ribosome/factor complex. Effective competition for the 40S ribosome/factor complex is not sufficient, however; translational efficiency (defined now as actual production of the intended protein) can still be reduced by upstream AUG codons or by base-paired structures that constitute barriers to the scanning 40S ribosome. The particulars, such as how much secondary structure is required to inhibit scanning, and the available evidence are summarized elsewhere (200). If one accepts the general notion that base-paired structures and upstream AUG codons can block ribosome entry and/or scanning, then the encumbered leader sequences described herein pose problems.

Time will tell which of the AUG-burdened cDNA sequences described above represent functional mRNAs and which represent mRNA precursors. The likelihood of the latter explanation increases as the number of upstream AUG codons increases: cDNAs with a dozen or so upstream AUG triplets (45, 85, 140, 151a, 161, 256, 309, 315, 421) almost certainly do not represent translatable transcripts! The importance of using appropriate primers to search for alternative 5′ noncoding sequences cannot be overemphasized. Positioning a primer near the 5′ end of the longest cDNA (85, 260) will nicely pinpoint the start site of the longest transcript, but alternative mRNAs with shorter leader sequences inevitably will be missed unless the primer is positioned close to the AUG initiator codon. The fact that so many AUG-burdened 5′-noncoding sequences have already been traced to retained introns or to other irregularities (documented above and in reference 193) encourages the view that many of the cDNAs in Table III may correspond to nonfunctional transcripts rather than to functional mRNAs. On the other hand, when the correct form of the mRNA is eventually deduced, genes thereby eliminated from Table III often move into Table IV! One way or another, the mRNAs for oncoproteins, transcription factors, growth factors, etc., seem destined to be translated poorly.

For GC-burdened cDNAs, the solution of switching promoters to produce an alternative, less encumbered 5′ noncoding sequence has been documented in only a few cases (178, 283, 317, 377). Because alternative leader sequences can easily be missed, as mentioned above, their frequency might be higher than presently appears. Nevertheless, because of the consistency with which GC-rich leader sequences occur, it seems farfetched to argue that most of those cDNAs (Table IV) derive from nonfunctional transcripts rather than from functional mRNAs. Unlike the tabulation of AUG-burdened leader sequences, which in time tends to be whittled down by corrections, the tabulation of GC-burdened leader sequences keeps growing. It includes mRNAs for many cytoskeletal and housekeeping proteins in addition to the regulatory proteins mentioned in Table IV. Thus, it seems likely that some (probably many) vertebrate mRNAs have enough secondary structure at the 5′ end to throttle translation. The biggest uncertainty may be whether these mRNAs invariably are translated poorly or whether their translation is "derepressed" in response to mitogens, for example, by modifications of the translational machinery or induction of helicases. (Although numerous modifications of the translational machinery correlate with a serum-induced increase in translation, no causal connection has yet been established.) The widespread occurrence of 5′ noncoding sequences that appear unfavorable for translation might be rationalized by the ability of GC motifs to promote transcription. In some vertebrate genes, sequence elements located downstream from the cap site indeed augment the yield of mRNA (51, 97, 131, 131a, 163, 210, 327, 392).

Whatever the explanation, the encumbered leader sequences described herein represent a minority of the vertebrate cDNA sequences that have been analyzed, and they are a distinctly nonrandom set. AUG-burdened/GC-rich leader sequences virtually never occur on mRNAs that encode globins, albumins, caseins, immunoglobulins, histones, or other highly expressed proteins. The fact that genes for growth factors, cytokine receptors, proto-oncogenes, etc., often produce transcripts with encumbered 5′ noncoding sequences suggests extensive regulation of the regulators, at the level of translation and/or RNA processing.

Research in the author's laboratory is supported by the National Institutes of Health (grant GM-33915).

This paper is dedicated to Judith Shakespeare.

Received for publication 8 May 1991 and in revised form 12 July 1991.

References

1. Abraham, J. A., J. L. Whang, A. Tumolo, A. Mergia, J. Friedman, D. Gospodarowicz, and J. C. Fiddes. 1986. Human basic fibroblast growth factor: nucleotide sequence and genomic organization. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2523–2528.
2. Acampora, D., M. D'Esposito, A. Faiella, M. Pannese, E. Migliaccio, F. Morelli, A. Stornaiuolo, V. Nigro, A. Simeone, and E. Boncinelli. 1989. The human HOX gene family. *Nucleic Acids Res.* 17:10385–10402.
3. Acland, P., M. Dixon, G. Peters, and C. Dickson. 1990. Subcellular fate of the *int-2* oncoprotein is determined by choice of initiation codon. *Nature (Lond.).* 343:662–665.
4. Alcorta, D. A., C. Crews, L. Sweet, L. Bankston, S. Jones, and R. L. Erikson. 1989. Sequence and expression of chicken and mouse *rsk*:

homologs of *Xenopus laevis* ribosomal S6 kinase. *Mol. Cell Biol.* 9:3850–3859.
5. Alfieri, C., M. Birkenbach, and E. Kieff. 1991. Early events in Epstein-Barr virus infection of human B lymphocytes. *Virology.* 181:595–608.
6. Amy, C. M., A. Witkowski, J. Naggert, B. Williams, Z. Randhawa, and S. Smith. 1989. Molecular cloning and sequencing of cDNAs encoding the entire rat fatty acid synthase. *Proc. Natl. Acad. Sci. USA.* 86: 3114–3118.
7. Anand, R., and J. Lindstrom. 1990. Nucleotide sequence of the human nicotinic acetylcholine receptor β2 subunit gene. *Nucleic Acids Res.* 18:4272.
8. Angel, P., E. A. Allegretto, S. T. Okino, K. Hattori, W. Boyle, T. Hunter, and M. Karin. 1988. Oncogene *jun* encodes a sequence-specific *trans*-activator similar to AP-1. *Nature (Lond.).* 332:166–171.
9. Aplan, P. D., C. G. Begley, V. Bertness, M. Nussmeier, A. Ezquerra, J. Coligan, and I. R. Kirsch. 1990. The SCL gene is formed from a transcriptionally complex locus. *Mol. Cell. Biol.* 10:6426–6435.
9a. Arrick, B. A., A. L. Lee, R. L. Grendell, and R. Derynck. 1991. Inhibition of translation of transforming growth factor-β3 mRNA by its 5′ untranslated region. *Mol. Cell Biol.* 11:4306–4313.
10. Awgulewitsch, A., C. Bieberich, L. Bogarad, C. Shashikant, and F. H. Ruddle. 1990. Structural analysis of the *Hox-3.1* transcription unit and the *Hox-3.2-Hox-3.1* intergenic region. *Proc. Natl. Acad. Sci. USA.* 87:6428–6432.
11. Bacher, N., Y. Zisman, E. Berent, and E. Livneh. 1991. Isolation and characterization of PKC-L, a new member of the protein kinase C-related gene family specifically expressed in lung, skin and heart. *Mol. Cell Biol.* 11:126–133.
12. Bates, B., J. Hardin, X. Zhan, K. Drickamer, and M. Goldfarb. 1991. Biosynthesis of human fibroblast growth factor-5. *Mol. Cell Biol.* 11:1840–1845.
13. Battey, J. F., J. Way, M. Corjay, H. Shapira, K. Kusano, R. Harkins, J. Wu, T. Slattery, E. Mann, and R. I. Feldman. 1991. Molecular cloning of the bombesin/gastrin-releasing peptide receptor from Swiss 3T3 cells. *Proc. Natl. Acad. Sci. USA.* 88:395–399.
14. Beals, C. R., C. B. Wilson, and R. M. Perlmutter. 1987. A small multigene family encodes G_i signal-transduction proteins. *Proc. Natl. Acad. Sci. USA.* 84:7886–7890.
15. Becerra, S. P., J. A. Rose, M. Hardy, B. M. Baroudy, and C. W. Anderson. 1985. Direct mapping of adeno-associated virus capsid proteins B and C: a possible ACG initiation codon. *Proc. Natl. Acad. Sci. USA.* 82:7919–7923.
16. Bell, G. I., N. Fong, M. Stempien, M. A. Wormsted, D. Caput, L. Ku, M. S. Urdea, L. B. Rall, and R. Sanchez-Pescador. 1986. Human epidermal growth factor precursor: cDNA sequence, expression in vitro and gene organization. *Nucleic Acids Res.* 14:8427–8446.
17. Bell, G. I., M. M. Stempien, N. M. Fong, and S. Seino. 1990. Sequence of a cDNA encoding guinea pig IGF-I. *Nucleic Acids Res.* 18:4275.
18. Beltzer, J. P., S. R. Morris, and G. B. Kohlhaw. 1988. Yeast *LEU4* encodes mitochondrial and nonmitochondrial forms of α-isopropylmalate synthase. *J. Biol. Chem.* 263:368–374.
19. Ben-Neriah, Y., and A. R. Bauskin. 1988. Leukocytes express a novel gene encoding a putative transmembrane protein-kinase devoid of an extracellular domain. *Nature (Lond.).* 333:672–676.
20. Bernards, A., and S. M. de la Monte. 1990. The *ltk* receptor tyrosine kinase is expressed in pre-B lymphocytes and cerebral neurons and uses a non-AUG translational initiator. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:2279–2287.
21. Bernards, A., C. M. Rubin, C. Westbrook, M. Paskind, and D. Baltimore. 1987. The first intron in the human c-*abl* gene is at least 200 kb long and is a target for translocations in chronic myelogenous leukemia. *Mol. Cell Biol.* 7:3231–3236.
22. Birchmeier, C., K. O'Neill, M. Riggs, and M. Wigler. 1990. Characterization of ROS1 cDNA from a human glioblastoma cell line. *Proc. Natl. Acad. Sci. USA.* 87:4799–4803.
23. Birnbaum, M. J. 1989. Identification of a novel gene encoding an insulin-responsive glucose transporter protein. *Cell.* 57:305–315.
24. Blasband, A. J., K. T. Rogers, X. Chen, J. Azizkhan, and D. C. Lee. 1990. Characterization of the rat transforming growth factor alpha gene and identification of promoter sequences. *Mol. Cell Biol.* 10:2111–2121.
25. Boehm, T., J. M. Greenberg, L. Buluwela, I. Lavenir, A. Forster, and T. H. Rabbitts. 1990. An unusual structure of a putative T cell oncogene which allows production of similar proteins from distinct mRNAs. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:857–868.
26. Bornstein, P., J. McKay, J. Morishima, S. Devarayalu, and R. E. Gelinas. 1987. Regulatory elements in the first intron contribute to transcriptional control of the human α1(I) collagen gene. *Proc. Natl. Acad. Sci. USA.* 84:8869–8873.
27. Bos, J. L., L. J. Polder, R. Bernards, P. Schrier, P. J. van den Elsen, A. J. van der Eb, and H. van Ormondt. 1981. The 2.2 kb Elb mRNA of human Ad12 and Ad5 codes for two tumor antigens starting at different AUG triplets. *Cell.* 27:121–131.
28. Boulter, J., A. O'Shea-Greenfield, R. Duvoisin, J. Connolly, E. Wada, A. Jensen, P. Gardner, M. Ballivet, E. S. Deneris, D. McKinnon, S. Heinemann, and J. Patrick. 1990. Three members of the rat neuronal nicotinic acetylcholine receptor gene family form a gene cluster. *J. Biol. Chem.* 265:4472–4482.
29. Boulton, T. G., S. H. Nye, D. Robbins, N. Y. Ip, E. Radziejewska, S. Morgenbesser, R. DePinho, N. Panayotatos, M. Cobb, and G. D. Yancopoulos. 1991. ERKs: a family of protein-serine/threonine kinases that are activated and tyrosine phosphorylated in response to insulin and NGF. *Cell.* 65:663–675.
30. Brand, N. J., M. Petkovich, and P. Chambon. 1990. Characterization of a functional promoter for the human retinoic acid receptor-alpha (hRAR-α). *Nucleic Acids Res.* 18:6799–6806.
31. Breier, G., G. R. Dressler, and P. Gruss. 1988. Primary structure and developmental expression pattern of *Hox* 3.1, a member of the murine *Hox* 3 homeobox gene cluster. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1329–1336.
32. Brenner, M., K. Lampel, Y. Nakatani, J. Mill, C. Banner, K. Mearow, M. Dohadwala, R. Lipsky, and E. Freese. 1990. Characterization of human cDNA and genomic clones for glial fibrillary acidic protein. *Mol. Brain Res.* 7:277–286.
33. Brinkman, A., C. Groffen, D. J. Kortleve, A. G. van Kessel, and S. L. S. Drop. 1988. Isolation and characterization of a cDNA encoding the low molecular weight insulin-like growth factor binding protein (IBP-1). *EMBO (Eur. Mol. Biol. Organ.) J.* 7:2417–2423.
34. Brookes, S., R. Smith, G. Casey, C. Dickson, and G. Peters. 1989. Sequence organization of the human *int-2* gene and its expression in teratocarcinoma cells. 1989. *Oncogene.* 4:429–436.
35. Brown, A. L., and M. M. Rechler. 1990. Cloning of the rat insulin-like growth factor-binding protein-2 gene and identification of a functional promoter lacking a TATA box. *Mol. Endocrinol.* 4:2039–2051.
36. Brown, J. W. S., and G. Feix. 1990. A functional splice site in the 5′ untranslated region of a zein gene. *Nucleic Acids Res.* 18:111–117.
37. Browner, M. F., F. Taroni, E. Sztul, and L. E. Rosenberg. 1991. Sequence analysis, biogenesis, and mitochondrial import of the α-subunit of rat liver propionyl-CoA carboxylase. (Correction) *J. Biol. Chem.* 266:4660.
38. Buchberg, A. M., H. G. Bedigian, N. A. Jenkins, and N. G. Copeland. 1990. *Evi-2*, a common integration site involved in murine myeloid leukemogenesis. *Mol. Cell Biol.* 10:4658–4666.
39. Buckler, A. J., J. Pelletier, D. A. Haber, T. Glaser, and D. E. Housman. 1991. Isolation, characterization, and expression of the murine Wilms' tumor gene (WTI) during kidney development. *Mol. Cell Biol.* 11:1707–1712.
40. Bugler, B., F. Amalric, and H. Prats. 1991. Alternative initiation of translation determines cytoplasmic or nuclear localization of basic fibroblast growth factor. *Mol. Cell Biol.* 11:573–577.
41. Burgess, W. H., C. A. Dionne, J. Kaplow, R. Mudd, R. Friesel, A. Zilberstein, J. Schlessinger, and M. Jaye. 1990. Characterization and cDNA cloning of phospholipase C-γ, a major substrate for heparin-binding growth factor 1 (acidic FGF)-activated tyrosine kinase. *Mol. Cell Biol.* 10:4770–4777.
42. Callis, J., M. Fromm, and V. Walbot. 1987. Introns increase gene expression in cultured maize cells. *Genes and Dev.* 1:1183–1200.
43. Carlson, M., R. Taussig, S. Kustu, and D. Botstein. 1983. The secreted form of invertase in *S. cerevisiae* is synthesized from mRNA encoding a signal sequence. *Mol. Cell Biol.* 3:439–447.
44. Castle, E., T. Nowak, U. Leidner, G. Wengler, and G. Wengler. 1985. Sequence analysis of the viral core protein and the membrane-associated proteins VI and NV2 of the flavivirus West Nile virus. *Virology.* 145:227–236.
45. Chambard, J-C., E. Van Obberghen-Schilling, R. Haslam, V. Vouret, and J. Pouyssegur. 1990. Chinese hamster serotonin (5-HT) type 2 receptor cDNA sequence. *Nucleic Acids Res.* 18:5282.
46. Chan, S. J., B. San Segundo, M. B. McCormick, and D. F. Steiner. 1986. Nucleotide and predicted amino acid sequences of cloned human and mouse preprocathepsin B cDNAs. *Proc. Natl. Acad. Sci. USA.* 83: 7721–7725.
47. Chan, W-K., M. E. Penaranda, S. E. Crawford, and M. K. Estes. 1986. Two glycoproteins are produced from the rotavirus neutralization gene. *Virology.* 151:243–252.
48. Chardin, P., G. Courtois, M-G. Mattei, and S. Gisselbrecht. 1991. The KUP gene, located on human chromosome 14, encodes a protein with two distant zinc fingers. *Nucleic Acids Res.* 19:1431–1436.
49. Chatton, B., P. Walter, J-P. Ebel, F. Lacroute, and F. Fasiolo. 1988. The yeast *VAS1* gene encodes both mitochondrial and cytoplasmic valyl-tRNA synthetases. *J. Biol. Chem.* 263:52–57.
50. Chen, J. H. 1990. Cloning, sequencing, and expression of mouse c-*ets*-1 cDNA in baculovirus expression system. *Oncogene Res.* 5:277–285.
51. Chen, R., H. A. Ingraham, M. N. Treacy, V. R. Albert, L. Wilson, and M. G. Rosenfeld. 1990. Autoregulation of *pit*-1 gene expression mediated by two cis-active promoter elements. *Nature (Lond.).* 346:583–586.
52. Chernoff, J., A. R. Schievella, C. A. Jost, R. L. Erikson, and B. G. Neel. 1990. Cloning of a cDNA for a major human protein-tyrosine-phosphatase. *Proc. Natl. Acad. Sci. USA.* 87:2735–2739.
53. Chinkers, M., D. L. Garbers, M-S. Chang, D. G. Lowe, H. Chin, D. V.

Goeddel, and S. Schulz. 1989. A membrane form of guanylate cyclase is an atrial natriuretic peptide receptor. *Nature (Lond.).* 338:78–83.
54. Chiu, I-M., W-P. Wang, and K. Lehtoma. 1990. Alternative splicing generates two forms of mRNA coding for human heparin-binding growth factor 1. *Oncogene.* 5:755–762.
55. Chretien, S., A. Dubart, D. Beaupain, N. Raich, B. Grandchamp, J. Rosa, M. Goossens, and P-H. Romeo. 1988. Alternative transcription and splicing of the human porphobilinogen deaminase gene result either in tissue-specific or in housekeeping expression. *Proc. Natl. Acad. Sci. USA.* 85:6–10.
56. Christy, B. A., L. F. Lau, and D. Nathans. 1988. A gene activated in mouse 3T3 cells by serum growth factors encodes a protein with "zinc finger" sequences. *Proc. Natl. Acad. Sci. USA.* 85:7857–7861.
57. Christy, R. J., K. H. Kaestner, D. E. Geiman, and M. D. Lane. 1991. CCAAT/enhancer binding protein gene promoter: binding of nuclear factors during differentiation of 3T3-L1 preadipocytes. *Proc. Natl. Acad. Sci. USA.* 88:2593–2597.
58. Chrivia, J. C., M. D. Uhler, and G. S. McKnight. 1988. Characterization of genomic clones coding for the Cα and Cβ subunits of mouse cAMP-dependent protein kinase. *J. Biol. Chem.* 263:5739–5744.
59. Chua, A. O., and U. Gubler. 1989. Sequence of the cDNA for the human fibroblast type interleukin-1 receptor. *Nucleic Acids Res.* 17:10114.
60. Chung, S., and R. P. Perry. 1989. Importance of introns for expression of mouse ribosomal protein gene rpL32. 1989. *Mol. Cell Biol.* 9:2075–2082.
61. Cianetti, L., A. Di Cristofaro, V. Zappavigna, L. Bottero, G. Boccoli, U. Testa, G. Russo, E. Boncinelli, and C. Peschle. 1990. Molecular mechanisms underlying the expression of the human HOX-5.1 gene. *Nucleic Acids Res.* 18:4361–4368.
62. Cigan, A. M., L. Feng, and T. F. Donahue. 1988. $tRNA_i^{met}$ functions in directing the scanning ribosome to the start site of translation. *Science (Wash. DC).* 242:93–97.
63. Clark, A. J. L., P. M. Lavender, P. Coates, M. R. Johnson, and L. H. Rees. 1990. In vitro and in vivo analysis of the processing and fate of the peptide products of the short proopiomelanocortin mRNA. *Molec. Endocrinol.* 4:1737–1743.
64. Clarke, B. E., D. V. Sangar, J. N. Burroughs, S. E. Newton, A. R. Carroll, and D. J. Rowlands. 1985. Two initiation sites for foot-and-mouth disease virus polyprotein in vivo. *J. Gen. Virol.* 66:2615–2626.
65. Clemens, M. J., V. M. Pain, S-T. Wong, and E. C. Henshaw. 1982. Phosphorylation inhibits guanine nucleotide exchange on eukaryotic initiation factor 2. *Nature (Lond.).* 296:93–95.
66. Clement, A., J. Campisi, S. R. Farmer, and J. S. Brody. 1990. Constitutive expression of growth-related mRNAs in proliferating and non-proliferating lung epithelial cells in primary culture: evidence for growth-dependent translational control. *Proc. Natl. Acad. Sci. USA.* 87:318–322.
67. Cohen, R. B., T. R. Boal, and B. Safer. 1990. Increased eIF-2α expression in mitogen-activated primary T lymphocytes. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:3831–3837.
68. Conboy, J. G., J. Y. Chan, J. A. Chasis, Y. W. Kan, and N. Mohandas. 1991. Tissue- and development-specific alternative RNA splicing regulates expression of multiple isoforms of erythroid membrane protein 4.1. *J. Biol. Chem.* 266:8273–8280.
69. Conboy, J., Y. W. Kan, S. B. Shohet, and N. Mohandas. 1986. Molecular cloning of protein 4.1, a major structural element of the human erythrocyte membrane skeleton. *Proc. Natl. Acad. Sci. USA.* 83:9512–9516.
70. Condie, B. G., A. H. Brivanlou, and R. M. Harland. 1990. Most of the homeobox-containing Xhox 36 transcripts in early *Xenopus* embryos cannot encode a homeodomain protein. *Mol. Cell Biol.* 10:3376–3385.
71. Conneely, O. M., D. M. Kettelberger, M-J. Tsai, W. T. Schrader, and B. W. O'Malley. 1989. The chicken progesterone receptor A and B isoforms are products of an alternate translation-initiation event. *J. Biol. Chem.* 264:14062–14064.
72. Cortner, J., and P. J. Farnham. 1990. Identification of the serum-responsive transcription initiation site of the zinc finger gene *Krox-20*. *Mol. Cell Biol.* 10:3788–3791.
73. Coulombe, B., A. Ponton, L. Daigneault, B. R. G. Williams, and D. Skup. 1988. Presence of transcription regulatory elements within an intron of the virus-inducible murine TIMP gene. *Mol. Cell Biol.* 8: 3227–3234.
74. Coulter-Karis, D. E., and M. S. Hershfield. 1989. Sequence of full length cDNA for human S-adenosylhomocysteine hydrolase. *Ann. Hum. Genet.* 53:169–175.
75. Crawford, D., K. Hagerty, and B. Beutler. 1989. Multiple splice forms of ribonuclease-inhibitor mRNA differ in the 5′ untranslated region. *Gene.* 85:525–531.
76. Crepin, K. M., M. I. Darville, L. Hue, and G. G. Rousseau. 1989. Characterization of distinct mRNAs coding for putative isozymes of 6-phosphofructo-2-kinase/fructose-2, 6-bisphosphatase. *Eur. J. Biochem.* 183: 433–440.
77. Crumley, G., C. A. Dionne, and M. Jaye. 1990. The gene for human acidic fibroblast growth factor encodes two upstream exons alternatively spliced to the first coding exon. *Biochem. Biophys. Res. Commun.* 171:7–13.
78. Curran, J., and D. Kolakofsky. 1988. Ribosomal initiation from an ACG codon in the Sendai virus P/C mRNA. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:245–251.
79. Delli Bovi, P., A. M. Curatola, F. G. Kern, A. Greco, M. Ittmann, and C. Basilico. 1987. An oncogene isolated by transfection of Kaposi's sarcoma DNA encodes a growth factor that is a member of the FGF family. *Cell.* 50:729–737.
80. DePinho, R. A., K. S. Hatton, A. Tesfaye, G. Yancopoulos, and F. W. Alt. 1987. The human *myc* gene family: structure and activity of L-*myc* and an L-*myc* pseudogene. *Genes & Dev.* 1:1311–1326.
81. Descombes, P., M. Chojkier, S. Lichsteiner, E. Falvey, and U. Schibler. 1990. LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein. *Genes & Dev.* 4:1541–1551.
82. Dixon, M., R. Deed, P. Acland, R. Moore, A. Whyte, G. Peters, and C. Dickson. 1989. Detection and characterization of the fibroblast growth factor-related oncoprotein INT-2. *Mol. Cell Biol.* 9:4896–4902.
83. Dolle, P., J.-L. Castrillo, L. E. Theill, T. Deerinck, M. Ellisman, and M. Karin. 1990. Expression of GHF-1 protein in mouse pituitaries correlates both temporally and spatially with the onset of growth hormone gene activity. *Cell.* 60:809–820.
84. Dolph, P. J., J. Huang, and R. J. Schneider. 1990. Translation by the adenovirus tripartite leader: elements which determine independence from cap-binding protein complex. *J. Virol.* 64:2669–2677.
85. D'Onofrio, M., M. D. Lee, C. M. Starr, M. Miller, and J. A. Hanover. 1991. The gene encoding rat nuclear pore glycoprotein p62 is intronless. *J. Biol. Chem.* 266:11980–11985.
85a. Dorai, T., J. B. Levy, L. Kang, J. S. Brugge, and L.-H. Wang. 1991. Analysis of cDNAs of the proto-oncogene c-*src*. *Mol. Cell Biol.* 11: 4165–4176.
86. Dosaka-Akita, H., R. K. Rosenberg, J. D. Minna, and M. J. Birrer. 1991. A complex pattern of translational initiation and phosphorylation in L-myc proteins. *Oncogene.* 6:371–378.
87. Dougherty, G. J., R. J. Kay, and R. K. Humphries. 1989. Molecular cloning of 114/A10, a cell surface antigen containing highly conserved repeated elements, which is expressed by murine hemopoietic progenitor cells and interleukin-3-dependent cell lines. *J. Biol. Chem.* 264: 6509–6514.
88. Duprey, P., K. Chowdhury, G. R. Dressler, R. Balling, D. Simon, J-L. Guenet, and P. Gruss. 1988. A mouse gene homologous to the *Drosophila* gene *caudal* is expressed in epithelial cells from the embryonic intestine. *Genes & Dev.* 2:1647–1654.
89. Edery, M., C. Jolicoeur, C. Levi-Meyrueis, I. Dusanter-Fourt, B. Petridou, J-M. Boutin, L. Lesueur, P. Kelly, and J. Djiane. 1989. Identification and sequence analysis of a second form of prolectin receptor by molecular cloning of complementary DNA from rabbit mammary gland. *Proc. Natl. Acad. Sci. USA.* 86:2112–2116.
90. Ellis, S. R., A. K. Hopper, and N. C. Martin. 1989. Amino-terminal extension generated from an upstream AUG codon increases the efficiency of mitochondrial import of yeast N^2, N^2-dimethylguanosine-specific tRNA methyl-transferases. *Mol. Cell Biol.* 9:1611–1620.
91. Emorine, L. J., S. Marullo, C. Delavier-Klutchko, S. V. Kaveri, O. Durieu-Trautmann, and A. D. Strosberg. 1987. Structure of the gene for human β_2-adrenergic receptor: expression and promoter characterization. *Proc. Natl. Acad. Sci. USA.* 84:6995–6999.
92. Ernst, H., and A. J. Shatkin. 1985. Reovirus hemagglutinin mRNA codes for two polypeptides in overlapping reading frames. *Proc. Natl. Acad. Sci. USA.* 82:48–52.
93. Erselius, J. R., M. D. Goulding, and P. Gruss. 1990. Structure and expression pattern of the murine *Hox*-3.2 gene. *Development (Camb.).* 110:629–642.
94. Eva, A., G. Vecchio, C. D. Rao, S. R. Tronick, and S. A. Aaronson. 1988. The predicted *DBL* oncogene product defines a distinct class of transforming proteins. *Proc. Natl. Acad. Sci. USA.* 85:2061–2065.
95. Fajardo, J. E., and A. J. Shatkin. 1990. Translation of bicistronic viral mRNA in transfected cells: regulation at the level of elongation. *Proc. Natl. Acad. Sci. USA.* 87:328–332.
96. Fan, C-M., and T. Maniatis. 1990. A DNA-binding protein containing two widely-separated zinc finger motifs that recognize the same DNA sequence. *Genes & Dev.* 4:29–42.
97. Farnham, P. J., and A. L. Means. 1990. Sequences downstream of the transcription initiation site modulate the activity of the murine dihydrofolate reductase promoter. *Mol. Cell Biol.* 10:1390–1398.
98. Fen, Z., and T. O. Daniel. 1991. 5′ Untranslated sequences determine degradative pathway for alternative PDGF B/c-*sis* mRNAs. *Oncogene.* 6:953–959.
99. Finch, P. W., J. S. Rubin, T. Miki, D. Ron, and S. A. Aaronson. 1989. Human KGF is FGF-related with properties of a paracrine effector of epithelial cell growth. *Science (Wash. DC.).* 245:752–755.
100. Firmbach-Kraft, I., M. Byers, T. Shows, R. Dalla-Favera, and J. J. Krolewski. 1990. *tyk2*, prototype of a novel class of non-receptor tyrosine kinase genes. *Oncogene.* 5:1329–1336.
101. Fischman, K., J. C. Edman, G. M. Shackleford, J. A. Turner, W. J. Rutter, and U. Nir. 1990. A murine *fer* testis-specific transcript (*ferT*) encodes a truncated Fer protein. *Mol. Cell Biol.* 10:146–153.
102. Flordellis, C. S., D. E. Handy, M. R. Bresnahan, V. I. Zannis, and H.

Gavras. 1991. Cloning and expression of a rat brain α2B-adrenergic receptor. *Proc. Natl. Acad. Sci. USA.* 88:1019–1023.
103. Florkiewicz, R. Z., and A. Sommer. 1989. Human basic fibroblast growth factor gene encodes four polypeptides: three initiate translation from non-AUG codons. *Proc. Natl. Acad. Sci. USA.* 86:3978–3981.
104. Fong, H. K. W., T. T. Amatruda, B. W. Birren, and M. I. Simon. 1987. Distinct forms of the β subunit of GTP-binding regulatory proteins identified by molecular cloning. *Proc. Natl. Acad. Sci. USA.* 84:3792–3796.
105. Foyt, H. L., D. LeRoith, and C. T. Roberts, Jr. 1991. Differential association of insulin-like growth factor I mRNA variants with polysomes in vivo. *J. Biol. Chem.* 266:7300–7305.
106. Fraser, P., P. Cummings, and P. Curtis. 1989. The mouse carbonic anhydrase I gene contains two tissue-specific promoters. *Mol. Cell Biol.* 9:3308–3313.
107. Froussard, P., P. Chatagner, G. Somme, A. Abadie, W. Greene, J. Theze, and S. Longacre. 1988. p55 IL-2 receptor mRNA precursors in murine T lymphocyte nuclei. *J. Immunol.* 141:1358–1364.
108. Fukunaga, R., E. Ishizaka-Ikeda, Y. Seto, and S. Nagata. 1990. Expression cloning of a receptor for murine granulocyte colony-stimulating factor. *Cell.* 61:341–350.
109. Fuller, F., A. S. Bhown, and D. H. L. Bishop. 1983. Bunyavirus nucleoprotein, N, and a nonstructural protein, NS_S, are coded by overlapping reading frames in the S RNA. *J. Gen. Virol.* 64:1705–1714.
110. Fuller, F., J. G. Porter, A. E. Arfsten, J. Miller, J. Schilling, R. M. Scarborough, J. A. Lewicki, and D. B. Schenk. 1988. Atrial natriuretic peptide clearance receptor. *J. Biol. Chem.* 263:9395–9401.
111. Gerard, N. P., R. L. Eddy, Jr., T. B. Shows, and C. Gerard. 1990. The human neurokinin A (substance K) receptor. *J. Biol. Chem.* 265:20455–20462.
112. Gil, G., J. R. Smith, J. L. Goldstein, C. A. Slaughter, K. Orth, M. S. Brown, and T. F. Osborne. 1988. Multiple genes encode nuclear factor 1-like proteins that bind to the promoter for 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Proc. Natl. Acad. Sci. USA.* 85:8963–8967.
113. Gillman, E. C., L. B. Slusher, N. C. Martin, and A. K. Hopper. 1991. *MOD5* translation initiation sites determine N^6-isopentenyladenosine modification of mitochondrial and cytoplasmic tRNA. *Mol. Cell Biol.* 11:2382–2390.
114. Giorgi, D., J-P. Bernard, S. Rouquier, J. Iovanna, H. Sarles, and J-C. Dagorn. 1989. Secretory pancreatic stone protein messenger RNA. *J. Clin. Invest.* 84:100–106.
115. Glencorse, T. A., A. N. Bateson, and M. G. Darlison. 1990. Sequence of the chicken $GABA_A$ receptor γ2-subunit cDNA. 1990. *Nucleic Acids Res.* 18:7157.
116. Gorman, D. M., N. Itoh, T. Kitamura, J. Schreurs, S. Yonehara, I. Yahara, K-I. Arai, and A. Miyajima. 1990. Cloning and expression of a gene encoding an interleukin 3 receptor-like protein: identification of another member of the cytokine receptor gene family. *Proc. Natl. Acad. Sci. USA.* 87:5459–5463.
117. Gray, P. W., S. Leung, E. H. Fennie, M. A. Farrar, J. T. Pingel, J. Fernandez-Luna, and R. D. Schreiber. 1989. Cloning and expression of the cDNA for the murine interferon γ receptor. *Proc. Natl. Acad. Sci. USA.* 86:8497–8501.
118. Green, S., P. Walter, V. Kumar, A. Krust, J-M. Bornert, P. Argos, and P. Chambon. 1986. Human oestrogen receptor cDNA: sequence, expression and homology to v-*erb-A*. *Nature (Lond.).* 320:134–139.
119. Grenningloh, G., V. Schmieden, P. R. Schofield, P. H. Seeburg, T. Siddique, T. K. Mohandas, C-M. Becker, and H. Betz. 1990. Alpha subunit variants of the human glycine receptor: primary structures, functional expression and chromosomal localization of the corresponding genes. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:771–776.
120. Grinberg, D., J. Thurlow, R. Watson, R. Smith, G. Peters, and C. Dickson. 1991. Transcriptional regulation of the *int*-2 gene in embryonal carcinoma cells. *Cell Growth & Differ.* 2:137–143.
121. Gros, P., M. Raymond, J. Bell, and D. Housman. 1988. Cloning and characterization of a second member of the mouse *mdr* gene family. *Mol. Cell Biol.* 8:2770–2778.
121a. Gu, M., J. D. York, I. Warshawsky, and P. W. Majerus. 1991. Identification, cloning, and expression of a cytosolic megakaryocyte protein-tyrosine-phosphatase with sequence homology to cytoskeletal protein 4.1. *Proc. Natl. Acad. Sci. USA.* 88:5867–5871.
122. Gutkind, J. S., D. C. Link, S. Katamine, P. Lacal, T. Miki, T. J. Ley, and K. C. Robbins. 1991. A novel c-*fgr* exon utilized in Epstein-Barr virus-infected B lymphocytes but not in normal monocytes. *Mol. Cell Biol.* 11:1500–1507.
123. Hagiwara, K., G. Stenman, H. Honda, P. Sahlin, A. Andersson, K. Miyazono, C-H. Heldin, F. Ishikawa, and F. Takaku. 1991. Organization and chromosomal localization of the human platelet-derived endothelial cell growth factor gene. *Mol. Cell Biol.* 11:2125–2132.
124. Haley, J., N. Whittle, P. Bennett, D. Kinchington, A Ullrich, and M. Waterfield. 1987. The human EGF receptor gene: structure of the 110 kb locus and identification of sequences regulating its transcription. *Oncogene Res.* 1:375–396.
125. Halley, C., Y. Courtois, and M. Laurent. 1988. Nucleotide sequence of bovine acidic fibroblast growth factor cDNA. *Nucleic Acids Res.* 16:10913.
126. Hamm, J., and I. W. Mattaj. 1990. Monomethylated cap structures facilitate RNA export from the nucleus. *Cell.* 63:109–118.
127. Hann, S. R., M. W. King, D. L. Bentley, C. W. Anderson, and R. N. Eisenman. 1988. A non-AUG translational initiation in c-*myc* exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas. *Cell.* 52:185–195.
128. Hannink, M., and H. M. Temin. 1990. Structure and autoregulation of the c-*rel* promoter. *Oncogene.* 5:1843–1850.
129. Hannon, R., T. Evans, G. Felsenfeld, and H. Gould. 1991. Structure and promoter activity of the gene for the erythroid transcription factor GATA-1. *Proc. Natl. Acad. Sci. USA.* 88:3004–3008.
130. Hao, Q-L., N. Heisterkamp, and J. Groffen. 1989. Isolation and sequence analysis of a novel human tyrosine kinase gene. *Mol. Cell Biol.* 9:1587–1593.
131. Hariharan, N., D. E. Kelley, and R. P. Perry. 1989. Equipotent mouse ribosomal protein promoters have a similar architecture that includes internal sequence elements. *Genes & Devel.* 3:1789–1800.
131a. Härtig, E., I. F. Loncarević, M. Büscher, P. Herrlich, and H. J. Rahmsdorf. 1991. A new cAMP response element in the transcribed region of the human c-fos gene. *Nucleic Acids Res.* 19:4153–4159.
132. Hatakeyama, M., M. Tsudo, S. Minamoto, T. Kono, T. Doi, T. Miyata, M. Miyasaka, and T. Taniguchi. 1989. Interleukin-2 receptor β chain gene: generation of three receptor forms by cloned human α and β chain cDNAs. *Science (Wash. DC).* 244:551–556.
133. Hattori, Y., H. Odagiri, H. Nakatani, K. Miyagawa, K. Naito, H. Sakamoto, O. Katoh, T. Yoshida, T. Sugimura, and M. Terada. 1990. K-*sam*, an amplified gene in stomach cancer, is a member of the heparin-binding growth factor receptor genes. *Proc. Natl. Acad. Sci. USA.* 87:5983–5987.
134. Haub, O., B. Drucker, and M. Goldfarb. 1990. Expression of the murine fibroblast growth factor 5 gene in the adult central nervous system. *Proc. Natl. Acad. Sci. USA.* 87:8022–8026.
135. Hayflick, J. S., J. P. Adelman, and P. H. Seeburg. 1989. The complete nucleotide sequence of the human gonadotropin-releasing hormone gene. *Nucleic Acids Res.* 17:6403.
136. Hershey, A. D., P. E. Dykema, and J. E. Krause. 1991. Organization, structure, and expression of the gene encoding the rat substance P receptor. *J. Biol. Chem.* 266:4366–4374.
137. Hershey, J. W. B. 1989. Protein phosphorylation controls translation rates. *J. Biol. Chem.* 264:20823–20826.
138. Hillarp, A., and B. Dahlback. 1990. Cloning of cDNA coding for the β chain of human complement component C4b-binding protein: sequence homology with the α chain. *Proc. Natl. Acad. Sci. USA.* 87:1183–1187.
139. Hirai, H., Y. Maru, K. Hagiwara, J. Nishida, and F. Takaku. 1987. A novel putative tyrosine kinase receptor encoded by the *eph* gene. *Science (Wash. DC).* 238:1717–1720.
140. Hirata, M., Y. Hayashi, F. Ushikubi, Y. Yokota, R. Kageyama, S. Nakanishi, and S. Narumiya. 1991. Cloning and expression of cDNA for a human thromboxane A_2 receptor. *Nature (Lond.).* 349:617–620.
141. Ho, I.-C., P. Vorhees, N. Marin, B. K. Oakley, S.-F. Tsai, S. H. Orkin, and J. M. Leiden. 1991. Human GATA-3: a lineage-restricted transcription factor that regulates the expression of the T cell receptor α gene. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:1187–1192.
142. Hoffman, E. K., S. P. Trusko, N. Freeman, and D. L. George. 1987. Structural and functional characterization of the promoter region of the mouse c-Ki-*ras* gene. *Mol. Cell Biol.* 7:2592–2596.
143. Hogarth, P. M., E. Witort, M. D. Hulett, C. Bonnerot, J. Even, W. H. Fridman, and I. F. C. McKenzie. 1991. Structure of the mouse βFcγ receptor II gene. *J. Immunol.* 146:369–376.
144. Holness, C. L., G. P. Lomonossoff, D. Evans, and A. J. Maule. 1989. Identification of the initiation codons for translation of cowpea mosaic virus middle component RNA using site-directed mutagenesis of an infectious cDNA clone. *Virol.* 172:311–320.
145. Holtzman, D. A., W. D. Cook, and A. R. Dunn. 1987. Isolation and sequence of a cDNA corresponding to a *src*-related gene expressed in murine hemopoietic cells. *Proc. Natl. Acad. Sci. USA.* 84:8325–8329.
146. Holzer, K. P., W. Liu, and G. G. Hammes. 1989. Molecular cloning and sequencing of chicken liver fatty acid synthase cDNA. *Proc. Natl. Acad. Sci. USA.* 86:4387–4391.
147. Hong, F. D., H-J. Huang, H. To, L-J. Young, A. Oro, R. Bookstein, E. Lee, and W-H. Lee. 1989. Structure of the human retinoblastoma gene. *Proc. Natl. Acad. Sci. USA.* 86:5502–5506.
148. Honkawa, H., W. Masahashi, S. Hashimoto, and T. Hashimoto-Gotoh. 1987. Identification of the principal promoter sequence of the c-H-*ras* transforming oncogene. *Mol. Cell Biol.* 7:2933–2940.
149. Horiuchi, T., K. Macon, V. J. Kidd, and J. E. Volanakis. 1990. Translational regulation of complement protein C2 expression by differential utilization of the 5'-untranslated region of mRNA. *J. Biol. Chem.* 265: 6521–6524.
150. Horton, W., T. Miyashita, K. Kohno, J. R. Hassell, and Y. Yamada. 1987. Identification of a phenotype-specific enhancer in the first intron of the rat collagen II gene. *Proc. Natl. Acad. Sci. USA.* 84:8864–8868.
151. Howard, T. E., S-Y. Shai, K. G. Langford, B. M. Martin, and K. E. Bern-

stein. 1990. Transcription of testicular angiotensin-converting enzyme (ACE) is initiated within the 12th intron of the somatic ACE gene. *Mol. Cell Biol.* 10:4294–4302.

151a. Hromas, R., S. J. Collins, D. Hickstein, W. Raskind, L. L. Deaven, P. O'Hara, F. S. Hagen, and K. Kaushansky. 1991. A retinoic acid-responsive human zinc finger gene, MZF-1, preferentially expressed in myeloid cells. *J. Biol. Chem.* 266:14183–14187.

152. Hsu, J.-C., T. Laz, K. L. Mohn, and R. Taub. 1991. Identification of LRF-1, a leucine-zipper protein that is rapidly and highly induced in regenerating liver. *Proc. Natl. Acad. Sci. USA.* 88:3511–3515.

153. Hsu, S., D. Cohen, L. S. Kirschner, L. Lothstein, M. Hartstein, and S. B. Horwitz. 1990. Structural analysis of the mouse *mdr*1a (P-glycoprotein) promoter reveals the basis for differential transcript heterogeneity in multidrug-resistant J774.2 cells. *Mol. Cell Biol.* 10:3596–3606.

154. Huang, E., K. Nocka, D. Beier, T-Y. Chu, J. Buck, H-W. Lahm, D. Wellner, P. Leder, and P. Besmer. 1990. The hematopoietic growth factor KL is encoded by the S1 locus and is the ligand of the c-*kit* receptor, the gene product of the W locus. *Cell.* 63:225–233.

155. Irminger, J-C., K. M. Rosen, R. E. Humbel, and L. Villa-Komaroff. 1987. Tissue-specific expression of insulin-like growth factor II mRNAs with distinct 5′ untranslated regions. *Proc. Natl. Acad. Sci. USA.* 84: 6330–6334.

156. Ishikawa, Y., C. Bianchi, B. Nadal-Ginard, and C. J. Homcy. 1990. Alternative promoter and 5′ exon generate a novel Gsα mRNA. *J. Biol. Chem.* 265:8458–8462.

157. Iwaki, A., T. Iwaki, J. E. Goldman, and R. Liem. 1990. Multiple mRNAs of rat brain α-crystallin B chain result from alternative transcriptional initiation. *J. Biol. Chem.* 265:22197–22203.

158. Jackson, R. J. 1988. Picornaviruses break the rules. *Nature (Lond.).* 334: 292–293.

159. Jacobs, K., C. Shoemaker, R. Rudersdorf, S. Neill, R. Kaufman, A. Mufson, J. Seehra, S. Jones, R. Hewick, E. Fritsch, M. Kawakita, T. Shimizu, and T. Miyake. 1985. Isolation and characterization of genomic and cDNA clones of human erythropoietin. *Nature (Lond.).* 313:806–810.

160. Januszeski, M. M., and J. A. Gurr. 1991. Differential translatability in vitro of multiple mRNAs encoding the β subunit of mouse thyrotropin. *Mol. Endocrinol.* In press.

161. Javaux, F., A. Donda, G. Vassart, and D. Christophe. 1991. Cloning and sequence analysis of TFE, a helix-loop-helix transcription factor able to recognize the thyroglobulin gene promoter in vitro. *Nucleic Acids Res.* 19:1121–1127.

162. Jeltsch, J-M., B. Turcotte, J-M. Garnier, T. Lerouge, Z. Krozowski, H. Gronemeyer, and P. Chambon. 1990. Characterization of multiple mRNAs originating from the chicken progesterone receptor gene. *J. Biol Chem.* 265:3967–3974.

163. Johnson, A. C., S. Ishii, Y. Jinno, I. Pastan, and G. T. Merlino. 1988. Epidermal growth factor receptor gene promoter. *J. Biol. Chem.* 263: 5693–5699.

164. Julius, D., A. B. MacDermott, R. Axel, and T. M. Jessell. 1988. Molecular characterization of a functional cDNA encoding the serotonin 1c receptor. *Science (Wash. DC).* 241:558–564.

165. Kastner, Ph., A. Kurst, C. Mendelsohn, J. M. Garnier, A. Zelent, P. Leroy, A. Staub, and P. Chambon. 1990. Murine isoforms of retinoic acid receptor γ with specific patterns of expression. *Proc. Natl. Acad. Sci. USA.* 87:2700–2704.

166. Kastner, P., M-T. Bocquel, B. Turcotte, J-M. Garnier, K. B. Horwitz, P. Chambon, and H. Gronemeyer. 1990. Transient expression of human and chicken progesterone receptors does not support alternative translational initiation from a single mRNA as the mechanism generating two receptor isoforms. *J. Biol. Chem.* 265:12163–12167.

167. Kastner, P., A. Kurst, B. Turcotte, U. Stropp, L. Tora, H. Gronemeyer, and P. Chambon. 1990. Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms A and B. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:1603–1614.

168. Katamine, S., V. Notario, C. D. Rao, T. Miki, M. S. Cheah, S. R. Tronick, and K. C. Robbins. 1988. Primary structure of the human *fgr* proto-oncogene product $p55^{c\text{-}fgr}$. *Mol. Cell Biol.* 8:259–266.

169. Kaye, F., J. Battey, M. Nau, B. Brooks, E. Seifter, J. De Greve, M. Birrer, E. Sausville, and J. Minna. 1988. Structure and expression of the human L-*myc* gene reveal a complex pattern of alternative mRNA processing. *Mol. Cell Biol.* 8:186–195.

170. Keese, P., A. Mackenzie, and A. Gibbs. 1989. Nucleotide sequence of the genome of an Australian isolate of turnip yellow mosaic tymovirus. *Virology.* 172:536–546.

171. Keir, W. J., C. A. Kozak, A. Chakraborti, R. A. Deitrich, and J. M. Sikela. 1991. The cDNA sequence and chromosomal location of the murine $GABA_A$ α1 receptor gene. *Genomics.* 9:390–395.

172. Kew, D., D. F. Jin, F. Kim, T. Laddis, and D. L. Kilpatrick. 1989. Translational status of proenkephalin mRNA in the rat reproductive system. *Mol. Endocrinol.* 3:1191–1196.

173. Kidd, V. J., W. Luo, J. Xiang, F. Tu, J. Easton, S. McCune, and M. L. Snead. 1991. Regulated expression of a cell division control-related protein kinase during development. *Cell Growth & Differ.* 2:85–93.

174. Killen, P. D., P. D. Burbelo, G. R. Martin, and Y. Yamada. 1988. Characterization of the promoter for the α1(IV) collagen gene. DNA sequences within the first intron enhance transcription. *J. Biol. Chem.* 263:12310–12314.

175. Kilpatrick, D. L., S. A. Zinn, M. Fitzgerald, H. Higuchi, S. L. Sabol, and J. Meyerhardt. 1990. Transcription of the rat and mouse proenkephalin genes is initiated at distinct sites in spermatogenic and somatic cells. *Mol. Cell Biol.* 10:3717–3726.

176. Kim, D. W., T. Uetsuki, Y. Kaziro, N. Yamaguchi, and S. Sugano. 1990. Use of the human elongation factor 1α promoter as a versatile and efficient expression system. *Gene.* 91:217–223.

177. Kim, H-R. C., N. Yew, W. Ansorge, H. Voss, C. Schwager, B. Vennstrom, M. Zenke, and J. D. Engel. 1988. Two different mRNAs are transcribed from a single genomic locus encoding the chicken erythrocyte anion transport proteins (Band 3). *Mol. Cell Biol.* 8:4416–4424.

178. Kim, S-J., A. Glick, M. B. Sporn, and A. B. Roberts. 1989. Characterization of the promoter region of the human transforming growth factor-β1 gene. *J. Biol. Chem.* 264:402–408.

179. Kimelman, D., and M. W. Kirschner. 1989. An antisense mRNA directs the covalent modification of the transcript encoding fibroblast growth factor in Xenopus oocytes. *Cell.* 59:687–696.

180. Kimura, H., W. H. Fischer, and D. Schubert. 1990. Structure, expression and function of a schwannoma-derived growth factor. *Nature (Lond.).* 348:257–260.

181. Klemsz, M. J., S. R. McKercher, A. Celada, C. Van Beveren, and R. A. Maki. 1990. The macrophage and B cell-specific transcription factor PU.1 is related to the *ets* oncogene. *Cell.* 61:113–124.

182. Kobilka,, B. K., C. MacGregor, K. Daniel, T. S. Kobilka, M. G. Caron, and R. J. Lefkowitz. 1987. Functional activity and regulation of human β_2-adrenergic receptors expressed in *Xenopus* oocytes. *J. Biol. Chem.* 262:15796–15802.

183. Kohno, T., M. T. Brewer, S. L. Baker, P. E. Schwartz, M. W. King, K. K. Hale, C. H. Squires, R. C. Thompson, and J. L. Vannice. 1990. A second tumor necrosis factor receptor gene product can shed a naturally occurring tumor necrosis factor inhibitor. *Proc. Natl. Acad. Sci. USA.* 87:8331–8335.

184. Koike, S., H. Horie, I. Ise, A. Okitsu, M. Yoshida, N. Iizuka, K. Takeuchi, T. Takegami, and A. Nomoto. 1990. The poliovirus receptor protein is produced both as membrane-bound and secreted forms. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:3217–3224.

185. Konieczny, S. F., and C. P. Emerson, Jr. 1987. Complex regulation of the muscle-specific contractile protein (Troponin I) gene. *Mol. Cell Biol.* 7:3065–3075.

186. Kozak, M. 1983. Translation of insulin-related polypeptides from messenger RNAs with tandemly reiterated copies of the ribosome binding site. *Cell.* 34:971–978.

187. Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell.* 44:283–292.

188. Kozak, M. 1986. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. USA.* 83:2850–2854.

189. Kozak, M. 1986. Regulation of protein synthesis in virus-infected animal cells. *Adv. Virus Res.* 31:229–292.

190. Kozak, M. 1987. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J. Mol. Biol.* 196:947–950.

191. Kozak, M. 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15:8125–8148.

192. Kozak, M. 1987. Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. *Mol. Cell Biol.* 7:3438–3445.

193. Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* 108:229–241.

194. Kozak, M. 1989. Context effects and (inefficient) initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol. Cell Biol.* 9:5073–5080.

195. Kozak, M. 1989. Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol. Cell Biol.* 9:5134–5142.

196. Kozak, M. 1990. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. USA.* 87:8301–8305.

197. Kozak, M. 1991. A consideration of alternative models for initiation of translation in eukaryotes. *BioEssays,* in press.

198. Kozak, M. 1991. A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. *Gene Expression.* 1:111–115.

199. Kozak, M. 1991. Effects of long 5′ leader sequences on initiation by eukaryotic ribosomes in vitro. *Gene Expression.* 1:117–125.

200. Kozak, M. 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* 266: in press.

201. Kozasa, T., H. Itoh, T. Tsukamoto, and Y. Kaziro. 1988. Isolation and characterization of the human Gsα gene. *Proc. Natl. Acad. Sci. USA.* 85:2081–2085.

202. Krupinski, J., F. Coussen, H. A. Bakalyar, W-J. Tang, P. Feinstein, K. Orth, C. Slaughter, R. R. Reed, and A. G. Gilman. 1989. Adenylyl cyclase amino acid sequence: possible channel- or transporter-like structure. *Science (Wash. DC).* 244:1558–1564.

203. Krust, A., S. Green, P. Argos, V. Kumar, P. Walter, J-M. Bornert, and P. Chambon. 1986. The chicken oestrogen receptor sequence: homology with v-*erbA* and the human oestrogen and glucocorticoid receptors. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:891–897.

204. Krust, A., P. Kastner, M. Petkovich, A. Zelent, and P. Chambon. 1989. A third human retinoic acid receptor, hRAR-γ. *Proc. Natl. Acad. Sci. USA.* 86:5310–5314.

204a. Kuhn, R., E. S. Monuki, and G. Lemke. 1991. The gene encoding the transcription factor SCIP has features of an expressed retroposon. *Mol. Cell Biol.* 11:4642–4650.

205. Kuo, C. J., P. B. Conley, C-L. Hsieh, U. Francke, and G. R. Crabtree. 1990. Molecular cloning, functional expression, and chromosomal localization of mouse hepatocyte nuclear factor 1. *Proc. Natl. Acad. Sci. USA.* 87:9838–9842.

206. Kuo, C. J., D. B. Mendel, L. P. Hansen, and G. R. Crabtree. 1991. Independent regulation of HNF-1a and HNF-1b by retinoic acid in F9 teratocarcinoma cells. *EMBO (Eur. Mol. Biol. Organ.) J.* In press.

207. Kwiatkowski, D. J., R. Mehl, and H. L. Yin. 1988. Genomic organization and biosynthesis of secreted and cytoplasmic forms of gelsolin. *J. Cell Biol.* 106:375–384.

208. Ladner, M. B., G. A. Martin, J. A. Noble, D. M. Nikoloff, R. Tal., E. S. Kawasaki, and T. J. White. 1987. Human CSF-1: gene structure and alternative splicing of mRNA precursors. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:2693–2698.

209. Lafyatis, R., R. Lechleider, S-J. Kim, S. Jakowlew, A. B. Roberts, and M. B. Sporn. 1990. Structural and functional characterization of the transforming growth factor β3 promoter. *J. Biol. Chem.* 265:19128–19136.

210. Lafyatis, R., R. Lechleider, A. B. Roberts, and M. B. Sporn. 1991. Secretion and transcriptional regulation of transforming growth factor-β3 during myogenesis. *Mol. Cell Biol.* 11:3795–3803.

211. Lambert, P. F., B. Spalholz, and P. M. Howley. 1987. A transcriptional repressor encoded by BPV-1 shares a common carboxy-terminal domain with the E2 transactivator. *Cell.* 50:69–78.

212. Landsman, D., O. W. McBride, N. Soares, M. P. Crippa, T. Srikantha, and M. Bustin. 1989. Chromosomal protein HMG-14. *J. Biol. Chem.* 264:3421–3427.

213. Langridge, P., and G. Feix. 1983. A zein gene of maize is transcribed from two widely separated promoter regions. *Cell.* 34:1015–1022.

214. Laprevotte, I., A. Hampe, C. J. Sherr, and F. Galibert. 1984. Nucleotide sequence of the *gag* gene and *gag-pol* junction of feline leukemia virus. *J. Virol.* 50:884–894.

215. Laudet, V., A. Begue, C. Henry-Duthoit, A. Joubel, P. Martin, D. Stehelin, and S. Saule. 1991. Genomic organization of the human thyroid hormone receptor α (c-*erbA-1*) gene. *Nucleic Acids Res.* 19:1105–1112.

216. Lazar, M. A., R. A. Hodin, D. S. Darling, and W. W. Chin. 1989. A novel member of the thyroid steroid hormone receptor family is encoded by the opposite strand of the rat c-*erbAα* transcriptional unit. *Mol. Cell Biol.* 9:1128–1136.

217. Lee, K. P., C. Taylor, B. Petryniak, L. Turka, C. June, and C. B. Thompson. 1990. The genomic organization of the CD28 gene. *J. Immunol.* 145:344–352.

218. Lee, S.-J. 1991. Expression of growth-differentiation factor 1 in the nervous system: conservation of a bicistronic structure. *Proc. Natl. Acad. Sci. USA.* 88:4250–4254.

219. Lemaire, P., C. Vesque, J. Schmitt, H. Stunnenberg, R. Frank, and P. Charnay. 1990. The serum-inducible mouse gene *Krox-24* encodes a sequence-specific transcriptional activator. *Mol. Cell Biol.* 10:3456–3467.

220. Leung, D. W., S. A. Spencer, G. Cachianes, R. G. Hammonds, C. Collins, W. J. Henzel, R. Barnard, M. J. Waters, and W. I. Wood. 1987. Growth hormone receptor and serum binding protein: purification, cloning and expression. *Nature (Lond.).* 330:537–543.

221. Leung, S., E. Whitelaw, and N. J. Proudfoot. 1989. Transcriptional and translational analysis of the human θ globin gene. *Nucleic Acids Res.* 17:8283–8300.

222. Levitan, E. S., P. Schofield,, D. Burt, L. Rhee, W. Wisden, M. Kohler, N. Fujita, H. Rodriguez, A. Stephenson, M. Darlison, E. Barnard, and P. H. Seeburg. 1988. Structural and functional basis for $GABA_A$ receptor heterogeneity. *Nature (Lond.).* 335:76–79.

223. Lhotak, V., P. Greer, K. Letwin, and T. Pawson. 1991. Characterization of *elk*, a brain-specific receptor tyrosine kinase. *Mol. Cell Biol.* 11:2496–2502.

224. Liao, C-F., A. Themmen, R. Joho, C. Barberis, M. Birnbaumer, and L. Birnbaumer. 1989. Molecular cloning and expression of a fifth muscarinic acetylcholine receptor. *J. Biol. Chem.* 264:7328–7337.

225. Lichtenheld, M. G., and E. R. Podack. 1989. Structure of the human perforin gene. *J. Immunol.* 143:4267–4274.

226. Lightowlers, R., S. Takamiya, R. Wessling, M. Lindorfer, and R. A. Capaldi. 1989. Cloning and sequencing of the cDNA for a 13th different subunit (IHQ) of beef heart cytochrome c oxidase. *J. Biol. Chem.* 264:16858–16861.

227. Linnenbach, A. J., J. Wojcierowski, S. Wu, J. Pyrc, A. Ross, B. Dietzschold, D. Speicher, and H. Koprowski. 1989. Sequence investigation of the major gastrointestinal tumor-associated antigen gene family, GA733. *Proc. Natl. Acad. Sci. USA.* 86:27–31.

228. Liu, F., and B. Roizman. 1991. The promoter, transcriptional unit, and coding sequence of herpes simplex virus 1 family 35 proteins are contained within and in-frame with the U_L26 ORF. *J. Virol.* 65:206–212.

229. Liu, F-T., K. Albrandt, and M. W. Robertson. 1988. cDNA heterogeneity suggests structural variants related to the high-affinity IgE receptor. *Proc. Natl. Acad. Sci. USA.* 85:5639–5643.

230. Livi, G. P., P. Kmetz, M. McHale, L. Cieslinski, G. Sathe, D. Taylor, R. Davis, T. Torphy, and J. M. Balcarek. 1990. Cloning and expression of cDNA for a human low-K_m, rolipram-sensitive cyclic AMP phosphodiesterase. *Mol. Cell Biol.* 10:2678–2686.

231. Lobel, P., N. Dahms, and S. Kornfeld. 1988. Cloning and sequence analysis of the cation-independent mannose 6-phosphate receptor. *J. Biol. Chem.* 263:2563–2570.

232. Lock, P., E. Stanley, D. A. Holtzman, and A. R. Dunn. 1990. Functional analysis and nucleotide sequence of the promoter region of the murine *hck* gene. *Mol. Cell Biol.* 10:4603–4611.

232a. Lock, P., S. Ralph, E. Stanley, I. Boulet, R. Ramsay, and A. R. Dunn. 1991. Two isoforms of murine *hck*, generated by utilization of alternative initiation codons, exhibit different patterns of subcellular localization. *Mol. Cell Biol.* 11:4363–4370.

233. Lomasney, J. W., S. Cotecchia, W. Lorenz, W. Leung, D. Schwinn, T. Yang-Feng, M. Brownstein, R. Lefkowitz, and M. G. Caron. 1991. Molecular cloning and expression of the cDNA for the α1A-adrenergic receptor. *J. Biol. Chem.* 266:6365–6369.

234. Lonberg, N., S. Gettner, E. Lacy, and D. R. Littman. 1988. Mouse brain CD4 transcripts encode only the COOH-terminal half of the protein. *Mol. Cell Biol.* 8:2224–2228.

234a. Lopez, L. C., A. Youakim, S. C. Evans, and B. D. Shur. 1991. Evidence for a molecular distinction between Golgi and cell surface forms of β1,4-galactosyltransferase. *J. Biol. Chem.* 266:15984–15991.

235. Lopez-Casillas, F., and K-H. Kim. 1989. Heterogeneity at the 5′ end of rat-acetyl-coenzyme A carboxylase mRNA. *J. Biol. Chem.* 264:7176–7184.

236. Lowrey, D. M., T. Aebischer, K. Olsen, M. Lichtenheld, F. Rupp, H. Hengartner, and E. R. Podack. 1989. Cloning, analysis, and expression of murine perforin 1 cDNA, a component of cytolytic T-cell granules with homology to complement component C9. *Proc. Natl. Acad. Sci. USA.* 86:247–251.

237. Luckow, V. A., and M. D. Summers. 1989. High level expression of nonfused foreign genes with *Autographa californica* nuclear polyhedrosis virus expression vectors. *Virology.* 170:31–39.

238. Luo, X., K. Park, F. Lopez-Casillas, and K-H. Kim. 1989. Structural features of the acetyl-CoA carboxylase gene: mechanisms for the generation of mRNAs with 5′ end heterogeneity. *Proc. Natl. Acad. Sci. USA.* 86:4042–4046.

239. Lupton, S. D., S. Gimpel, R. Jerzy, L. Brunton, K. Hjerrild, D. Cosman, and R. G. Goodwin. 1990. Characterization of the human and murine IL-7 genes. *J. Immunol.* 144:3592–3601.

240. Makela, T. P., K. Saksela, and K. Alitalo. 1989. Two N-*myc* polypeptides with distinct amino termini encoded by the second and third exons of the gene. *Mol. Cell Biol.* 9:1545–1552.

241. Maldonado, F., and S. K. Hanks. 1988. A cDNA clone encoding human cAMP-dependent protein kinase catalytic subunit Cα. *Nucleic Acids Res.* 16:8189.

242. Mali, M., P. Jaakkola, A-M. Arvilommi, and M. Jalkanen. 1990. Sequence of human syndecan indicates a novel gene family of integral membrane proteoglycans. *J. Biol. Chem.* 265:6884–6889.

243. Manet, E., H. Gruffat, M. C. Trescol-Biemont, N. Moreno, P. Chambard, J. F. Giot, and A. Sergeant. 1989. Epstein-Barr virus bicistronic mRNAs generated by facultative splicing code for two transcriptional *trans*-activators. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:1819–1826.

244. Mano, H., F. Ishikawa, J. Nishida, H. Hirai, and F. Takaku. 1990. A novel protein tyrosine kinase. *tec*, is preferentially expressed in liver. *Oncogene.* 5:1781–1786.

245. Mansour, S. L., and G. R. Martin. 1988. Four classes of mRNA are expressed from the mouse *int*-2 gene, a member of the FGF gene family. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:2035–2041.

246. Manzella, J. M., and P. J. Blackshear. 1990. Regulation of rat ornithine decarboxylase mRNA translation by its 5′ untranslated region. *J. Biol. Chem.* 265:11817–11822.

247. Marth, J. D., R. W. Overell, K. E. Meier, E. G. Krebs, and R. M. Perlmutter. 1988. Translational activation of the *lck* proto-oncogene. *Nature (Lond.).* 332:171–173.

248. Masu, Y., K. Nakayama, H. Tamaki, Y. Harada, M. Kuno, and S. Nakanishi. 1987. cDNA cloning of bovine substance-K receptor through oocyte expression system. *Nature (Lond.).* 329:836–838.

249. Mattion, N. M., D. B. Mitchell, G. W. Both, and M. K. Estes. 1991. Expression of rotavirus proteins encoded by alternative ORFs of genome segment 11. *Virology.* 181:295–304.

250. Mavrothalassitis, G. J., D. K. Watson, and T. S. Papas. 1990. Molecular and functional characterization of the promoter of *ETS2*, the human c-*ets-2* gene. *Proc. Natl. Acad. Sci. USA.* 87:1047–1051.

251. McGarry, T. J., and S. Lindquist. 1985. The preferential translation of Drosophila hsp70 mRNA requires sequences in the untranslated leader. *Cell.* 42:903–911.

252. McGuire, E. A., R. D. Hockett, K. M. Pollock, M. F. Bartholdi, S. J. O'Brien, and S. J. Korsmeyer. 1989. The t(11;14)(p15;q11) in a T-cell acute lymphoblastic leukemia cell line activates multiple transcripts, including *Ttg-1*, a gene encoding a potential zinc finger protein. *Mol. Cell Biol.* 9:2124–2132.
253. Meerovitch, K., J. Pelletier, and N. Sonenberg. 1989. A cellular protein that binds to the 5′ noncoding region of poliovirus RNA: implications for internal translation initiation. *Genes & Devel.* 3:1026–1034.
254. Meijer, D., A. Graus, R. Kraay, A. Langeveld, M. P. Mulder, and G. Grosveld. 1990. The octomer binding factor Oct6: cDNA cloning and expression in early embryonic cells. *Nucleic Acids Res.* 18:7357–7365.
255. Mellentin, J. D., S. D. Smith, and M. L. Cleary. 1989. *lyl*-1, a novel gene altered by chromosomal translocation in T cell leukemia, codes for a protein with a helix-loop-helix DNA binding motif. *Cell.* 58:77–83.
256. Mendel, D. B., L. P. Hansen, M. K. Graves, P. B. Conley, and G. R. Crabtree. 1991. HNF-1α and HNF-1β (vHNF-1) share dimerization and homeodomains but not activation domains. *Genes & Devel.* 5:1042–1056.
257. Messer, G., U. Spengler, M. C. Jung, G. Honold, K. Blomer, G. R. Pape, G. Riethmuller, and E. H. Weiss. 1991. Polymorphic structure of the TNF locus. *J. Exp. Med.* 173:209–219.
258. Meyer, R., E. Hatada, H-P. Hohmann, M. Haiker, C. Bartsch, U. Rothlisberger, H-W. Lahm, E. J. Schlaeger, A. van Loon, and C. Scheidereit. 1991. Cloning of the DNA-binding subunit of human nuclear factor κB. *Proc. Natl. Acad. Sci. USA.* 88:966–970.
259. Miles, J. S., A. W. McLaren, and C. R. Wolf. 1989. Alternative splicing in the human cytochrome P450IIB6 gene generates a high level of aberrant messages. *Nucleic Acids Res.* 17:8241–8255.
260. Miller, R. T., L. Counillon, G. Pages, R. P. Lifton, C. Sardet, and J. Pouyssegur. 1991. Structure of the 5′-flanking regulatory region and gene for the human growth factor-activatable Na/H exchanger. *J. Biol. Chem.* 266:10813–10819.
261. Mirkov, T. E., D. M. Mathews, D. H. Du Plessis, and J. A. Dodds. 1989. Nucleotide sequence and translation of satellite tobacco mosaic virus RNA. *Virology.* 170:139–146.
262. Misrahi, M., H. Loosfelt, M. Atger, C. Meriel, V. Zerah, P. Dessen, and E. Milgrom. 1988. Organization of the entire rabbit progesterone receptor mRNA and of the promoter and 5′ flanking region of the gene. *Nucleic Acids Res.* 16:5459–5472.
263. Mitchell, D. B., and G. W. Both. 1988. Simian rotavirus SA11 segment 11 contains overlapping reading frames. *Nucleic Acids Res.* 16:6244.
264. Mitsuhashi, T., and V. M. Nikodem. 1989. Regulation of expression of the alternative mRNAs of the rat α-thyroid hormone receptor gene. *J. Biol. Chem.* 264:8900–8904.
265. Miyajima, N., Y. Kadowaki, S. Fukushige, S. Shimizu, K. Semba, Y. Yamanashi, K. Matsubara, K. Toyoshima, and T. Yamamoto. 1988. Identification of two novel members of *erbA* superfamily by molecular cloning. *Nucleic Acids Res.* 16:11057–11074.
266. Miyajima, N., R. Horiuchi, Y. Shibuya, S. Fukushige, K. Matsubara, K. Toyoshima, and T. Yamamoto. 1989. Two *erbA* homologs encoding proteins with different T3 binding capacities are transcribed from opposite DNA strands of the same genetic locus. *Cell.* 57:31–39.
267. Monsma, F. J., L. D. McVittie, C. Gerfen, L. Mahan, and D. R. Sibley. 1989. Multiple D_2 dopamine receptors produced by alternative RNA splicing. *Nature (Lond.).* 342:926–929.
268. Monsma, F. J., L. Mahan, L. McVittie, C. Gerfen, and D. R. Sibley. 1990. Molecular cloning and expression of a D_1 dopamine receptor linked to adenylyl cyclase activation. *Proc. Natl. Acad. Sci. USA.* 87: 6723–6727.
269. Moore, D. D., A. Marks, D. Buckley, G. Kapler, F. Payvar, and H. M. Goodman. 1985. The first intron of the human growth hormone gene contains a binding site for glucocorticoid receptor. *Proc. Natl. Acad. Sci. USA.* 82:699–702.
270. Morgan, D. O., J. Edman, D. Standring, V. Fried, M. Smith, R. Roth, and W. J. Rutter. 1987. Insulin-like growth factor II receptor as a multifunctional binding protein. *Nature (Lond.).* 329:301–307.
271. Morishita, K., E. Parganas, E. C. Douglass, and J. N. Ihle. 1990. Unique expression of the human *Evi-1* gene in an endometrial carcinoma cell line: sequence of cDNAs and structure of alternatively spliced transcripts. *Oncogene.* 5:963–971.
272. Mueller, C. R., P. Maire, and U. Schibler. 1990. DBP, a liver-enriched transcriptional activator, is expressed late in ontogeny and its tissue specificity is determined posttranscriptionally. *Cell.* 61:279–291.
273. Muller, A. J., and O. N. Witte. 1989. The 5′ noncoding region of the human leukemia-associated oncogene BCR/ABL is a potent inhibitor of in vitro translation. *Mol. Cell Biol.* 9:5234–5238.
274. Murphy, P., and R. E. Hill. 1991. Expression of the mouse *labial*-like homeobox-containing genes, *Hox 2.9* and *Hox 1.6*, during segmentation of the hindbrain. *Development.* 111:61–74.
275. Murphy, S. P., J. Garbern, W. Odenwald, R. Lazzarini, and E. Linney. 1988. Differential expression of the homeobox gene *Hox-1.3* in F9 embryonal carcinoma cells. *Proc. Natl. Acad. Sci. USA.* 85:5587–5591.
276. Murphy, T. J., R. W. Alexander, K. K. Griendling, M. S. Runge, and K. E. Bernstein. 1991. Isolation of a cDNA encoding the vascular type-1 angiotensin II receptor. *Nature (Lond.).* 351:233–236.
277. Murray, M. B., N. Zilz, N. McCreary, M. MacDonald, and H. C. Towle. 1988. Isolation and characterization of rat cDNA clones for two distinct thyroid hormone receptors. *J. Biol. Chem.* 263:12770–12777.
278. Nagase, T., T. Sudo, T. Maekawa, T. Yoshimura, J. Fujisawa, M. Yoshida, and S. Ishii. 1990. Promoter region of the human CRE-BP1 gene encoding the transcriptional regulator binding to the cAMP response element. *J. Biol. Chem.* 265:17300–17306.
279. Nagashima, K., M. Yoshida, and M. Seiki. 1986. A single species of pX mRNA of human T-cell leukemia virus type I encodes trans-activator $p40^x$ and two other phosphoproteins. *J. Virol.* 60:394–399.
280. Nakano, Y., T. Nishihara, S. Sasayama, T. Miwa, S. Kamada, and T. Kakunaga. 1991. Transcriptional regulatory elements in the 5′ upstream and first intron regions of the human smooth muscle (aortic type) α-actin-encoding gene. *Gene.* 99:285–289.
281. Namen, A. E., S. Lupton, K. Hjerrild, J. Wignall, D. Mochizuki, A. Schmierer, B. Mosley, C. March, D. Urdal, S. Gillis, D. Cosman, and R. G. Goodwin. 1988. Stimulation of B-cell progenitors by cloned murine interleukin-7. *Nature (Lond.).* 333:571–573.
282. Natsoulis, G., F. Hilger, and G. R. Fink. 1986. The *HTS1* gene encodes both the cytoplasmic and mitochondrial histidine tRNA synthetases of S. cerevisiae. *Cell.* 46:235–243.
283. Nielsen, F. C., S. Gammeltoft, and J. Christiansen. 1990. Translational discrimination of mRNAs coding for human insulin-like growth factor II. *J. Biol. Chem.* 265:13431–13434.
284. Noda, C., M. Ohguri, K. Matsuda, T. Nakamura, A. Hasegawa, S. Yagi, and A. Ichihara. 1990. Organization and structure of the 5′ flanking region of the rat serine dehydratase gene. *J. Biochem.* 108:622–628.
285. Nomura, N., S. Sasamoto, S. Ishii, T. Date, M. Matsui, and R. Ishizaki. 1989. Isolation of human cDNA clones of *ski* and the *ski*-related gene, *sno*. *Nucleic Acids Res.* 17:5489–5500.
286. Nonaka, M., N. Ishikawa, J. Passwell, S. Natsuume-Sakai, and H. R. Colten. 1989. Tissue-specific initiation of murine complement factor B mRNA transcription. *J. Immunol.* 142:1377–1382.
287. Nordqvist, K. and G- Akusjarvi. 1990. Adenovirus early region 4 stimulates mRNA accumulation via 5′ introns. *Proc. Natl. Acad. Sci. USA.* 87:9543–9547.
288. Norman, C., M. Runswick, R. Pollock, and R. Treisman. 1988. Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-*fos* serum response element. *Cell.* 55:989–1003.
289. Nudel, U., D. Zuk, P. Einat, E. Zeelon, Z. Levy, S. Neuman, and D. Yaffe. 1989. Duchenne muscular dystrophy gene product is not identical in muscle and brain. *Nature (Lond.).* 337:76–78.
290. Nutter, R. C., K. Scheets, L. C. Panganiban, and S. A. Lommel. 1989. The complete nucleotide sequence of the maize chlorotic mottle virus genome. *Nucleic Acids Res.* 17:3163–3177.
291. Oda, T., T. Funai, and A. Ichiyama. 1990. Generation from a single gene of two mRNAs that encode the mitochondrial and peroxisomal serine:pyruvate aminotransferase of rat liver. *J. Biol. Chem.* 265:7513–7519.
292. Ogawa, H., M. Fujioka, Y. Matsuda, Y. Su, T. Dunn, D. A. Miller, and H. C. Pitot. 1988. Sequence of the rat serine dehydratase gene. *Nucleic Acids Res.* 16:10921–10923.
293. Ohagi, S., M. Nishi, and D. F. Steiner. 1990. Sequence of a cDNA encoding human LRP (leukocyte common antigen-related peptide). *Nucleic Acids Res.* 18:7159.
294. Ohno, H., G. Takimoto, and T. W. McKeithan. 1990. The candidate proto-oncogene *bcl*-3 is related to genes implicated in cell lineage determination and cell cycle control. *Cell.* 60:991–997.
295. Ohno, S., Y. Akita, Y. Konno, S. Imajoh, and K. Suzuki. 1988. A novel phorbol ester receptor protein kinase, nPKC, distantly related to the protein kinase C family. *Cell.* 53:731–741.
296. O'Mahony, D. J., and L. I. Rothblum. 1991. Identification of two forms of the RNA polymerase I transcription factor UBF. *Proc. Natl. Acad. Sci. USA.* 88:3180–3184.
297. Ono, Y., T. Fujii, K. Ogita, U. Kikkawa, K. Igarashi, and Y. Nishizuka. 1989. Protein kinase C ζ subspecies from rat brain: its structure, expression, and properties. *Proc. Natl. Acad. Sci. USA.* 86:3099–3103.
298. Oosterwegel, M., M. van de Wetering, D. Dooijes, L. Klomp, A. Winoto, K. Georgopoulos, F. Meijlink, and H. Clevers. 1991. *J. Exp. Med.* 173:1133–1142.
299. Oren, R., S. Takahashi, C. Doss, R. Levy, and S. Levy. 1990. TAPA-1, the target of an antiproliferative antibody, defines a new family of transmembrane proteins. *Mol. Cell Biol.* 10:4007–4015.
300. Osatomi, K. and H. Sumiyoshi. 1990. Complete nucleotide sequence of dengue type 3 virus genome RNA. *Virology.* 176:643–647.
301. Ozer, J., M. Faber, R. Chalkley, and L. Sealy. 1990. Isolation and characterization of a cDNA clone for the CCAAT transcription factor $EF1_A$ reveals a novel structural motif. *J. Biol. Chem.* 265:22143–22152.
302. Pain, V. M. 1986. Initiation of protein synthesis in mammalian cells. *Biochem. J.* 235:625–637.
303. Palmiter, R. D., E. P. Sandgren, M. R. Avarbock, D. Allen, and R. L. Brinster. 1991. Heterologous introns can enhance expression of transgenes in mice. *Proc. Natl. Acad. Sci. USA.* 88:478–482.
304. Pasquale, E. B. 1990. A distinctive family of embryonic protein-tyrosine kinase receptors. *Proc. Natl. Acad. Sci. USA.* 87:5812–5816.

305. Paul, S. R., F. Bennett, J. Calvetti, K. Kelleher, C. Wood, R. O'Hara, A. Leary, B. Sibley, S. Clark, D. Williams, and Y-C. Yang. 1990. Molecular cloning of a cDNA encoding interleukin 11, a stromal cell-derived lymphopoietic and hematopoietic cytokine. *Proc. Natl. Acad. Sci. USA.* 87:7512–7516.
306. Persing, D. H., H. E. Varmus, and D. Ganem. 1985. A frameshift mutation in the pre-S region of the human HBV genome allows production of surface antigen particles but eliminates binding to polymerized albumin. *Proc. Natl. Acad. Sci. USA.* 82:3440–3444.
307. Persing, D. H., H. E. Varmus, and D. Ganem. 1986. Inhibition of secretion of hepatitis B surface antigen by a related presurface polypeptide. *Science (Wash. DC).* 234:1388–1391.
308. Petty, I. T. D., and A. O. Jackson. 1990. Two forms of the major barley stripe mosaic virus nonstructural protein are synthesized in vivo from alternative initiation codons. *Virology.* 177:829–832.
309. Pleiman, C. M., S. D. Gimpel, L. S. Park, H. Harada, T. Taniguchi, and S. F. Ziegler. 1991. Organization of the murine and human interleukin-7 receptor genes. *Mol. Cell Biol.* 11:3052–3059.
310. Plowman, G. D., G. S. Whitney, M. G. Neubauer, J. M. Green, V. McDonald, G. J. Todaro, and M. Shoyab. 1990. Molecular cloning and expression of an additional epidermal growth factor receptor-related gene. *Proc. Natl. Acad. Sci. USA.* 87:4905–4909.
311. Plowman, G. D., J. M. Green, V. McDonald, M. Neubauer, C. Disteche, G. J. Todaro, and M. Shoyab. 1990. The amphiregulin gene encodes a novel EGF-related protein with tumor-inhibitory activity. *Mol. Cell Biol.* 10:1969–1981.
312. Prats, A.-C., G. DeBilly, P. Wang, and J.-L. Darlix. 1989. CUG initiation codon used for the synthesis of a cell surface antigen coded by the murine leukemia virus. *J. Mol. Biol.* 205:363–372.
313. Pritchett, D. B., A. Bach, M. Wozny, O. Taleb, R. Dal Toso, J. C. Shih, and P. H. Seeburg. 1988. Structure and functional expression of cloned rat serotonin 5HT-2 receptor. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:4135–4140.
314. Propst, F., M. Rosenberg, A. Iyer, K. Kaul, and G. F. Vande Woude. 1987. c-*mos* proto-oncogene RNA transcripts in mouse tissues: structural features, developmental regulation, and localization in specific cell types. *Mol. Cell Biol.* 7:1629–1637.
315. Pytela, R. 1988. Amino acid sequence of the murine Mac-1 α chain reveals homology with the integrin family and an additional domain related to von Willebrand factor. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1371–1378.
316. Qian, F., A. Frankfater, S. J. Chan, and D. F. Steiner. 1991. The structure of the mouse cathepsin B gene and its putative promoter. *DNA Cell Biol.* 10:159–168.
317. Qian, S. W., P. Kondaiah, W. Casscells, A. B. Roberts, and M. B. Sporn. 1991. A second messenger RNA species of transforming growth factor β1 in infarcted rat heart. *Cell Regulation.* 2:241–249.
318. Quintrell, N., R. Lebo, H. Varmus, J. M. Bishop, M. J. Pettenati, M. M. Le Beau, M. O. Diaz, and J. D. Rowley. 1987. Identification of a human gene (*HCK*) that encodes a protein-tyrosine kinase and is expressed in hemopoietic cells. *Mol. Cell Biol.* 7:2267–2275.
319. Rao, C. D., M. Pech, K. C. Robbins, and S. A. Aaronson. 1988. The 5′ untranslated sequence of the c-*sis*/platelet-derived growth factor 2 transcript is a potent translational inhibitor. *Mol. Cell Biol.* 8:284–292.
320. Rao, V. N., T. S. Papas, and E. S. P. Reddy. 1987. *Erg*, a human *ets*-related gene on chromosome 21: alternative splicing, polyadenylation, and translation. *Science (Wash. DC).* 237:635–639.
321. Ray, D., P. Meneceur, A. Tavitian, and J. Robert-Lezenes. 1987. Presence of a c-*myc* transcript initiated in intron 1 in Friend erythroleukemia cells and in other murine cell types with no evidence of c-*myc* gene rearrangement. *Mol. Cell Biol.* 7:940–945.
322. Ray, D., S. Culine, A. Tavitian, and F. Moreau-Gachelin. 1990. The human homologue of the putative proto-oncogene Spi-1: characterization and expression in tumors. *Oncogene.* 5:663–668.
323. Reeves, R., G. A. Spies, M. Kiefer, P. Barr, and M. Power. 1990. Primary structure of the putative human oncogene, *pim*-1. *Gene.* 90:303–307.
324. Reynolds, P. J., J. Lesley, J. Trotter, R. Schulte, R. Hyman, and B. M. Sefton. 1990. Changes in the relative abundance of type I and type II *lck* mRNA transcripts suggest differential promoter usage during T-cell development. *Mol. Cell Biol.* 10:4266–4270.
325. Rochon, D. M., and J. C. Johnston. 1991. Infectious transcripts from cloned cucumber necrosis virus cDNA: evidence for a bifunctional subgenomic mRNA. *Virology.* 181:656–665.
326. Romisch, K., J. Webb, J. Herz, S. Prehn, R. Frank, M. Vingron, and B. Dobberstein. 1989. Homology of 54K protein of signal-recognition particle, docking protein and two E. coli proteins with putative GTP-binding domains. *Nature (Lond.).* 340:478–482.
327. Ronen, D., V. Rotter, and D. Reisman. 1991. Expression from the murine p53 promoter is mediated by factor binding to a downstream helix-loop-helix recognition motif. *Proc. Natl. Acad. Sci. USA.* 88:4128–4132.
328. Rorsman, F., M. Bywater, T. J. Knott, J. Scott, and C. Betsholtz. 1988. Structural characterization of the human PDGF A-chain cDNA and gene: alternative exon usage predicts two different precursor proteins. *Mol. Cell Biol.* 8:571–577.
329. Rose-John, S., A. Dietrich, and F. Marks. 1988. Molecular cloning of mouse protein kinase C (PKC) cDNA from Swiss 3T3 fibroblasts. *Gene.* 74:465–471.
330. Rupp, R. A. W., U. Kruse, G. Multhaup, U. Gobel, K. Beyreuther, and A. E. Sippel. 1990. Chicken NF1/TGGCA proteins are encoded by at least three independent genes: NF1-A, NF1-B and NF1-C with homologues in mammalian genomes. *Nucleic Acids Res.* 18:2607–2616.
331. Russo, R. N., N. L. Shaper, and J. H. Shaper. 1990. Bovine β1-4 galactosyltransferase: two sets of mRNA transcripts encode two forms of the protein with different amino-terminal domains. *J. Biol. Chem.* 265: 3324–3331.
332. Ryder, K., A. Lanahan, E. Perez-Albuerne, and D. Nathans. 1989. *Jun*-D: a third member of the *Jun* gene family. *Proc. Natl. Acad. Sci. USA.* 86:1500–1503.
333. Ryseck, R-P., H. Macdonald-Bravo, M-G. Mattei, S. Ruppert, and R. Bravo. 1989. Structure, mapping and expression of a growth factor inducible gene encoding a putative nuclear hormonal binding receptor. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:3327–3335.
334. Saga, Y., J-S. Tung, F-W. Shen, T. C. Pancoast, and E. A. Boyse. 1988. Organization of the *Ly-5* gene. *Mol. Cell Biol.* 8:4889–4895.
335. Sap, J., P. D'Eustachio, D. Givol, and J. Schlessinger. 1990. Cloning and expression of a widely expressed receptor tyrosine phosphatase. *Proc. Natl. Acad. Sci. USA.* 87:6112–6116.
336. Saris, C. J. M., J. Domen, and A. Berns. 1991. The *pim*-1 oncogene encodes two related protein-serine/threonine kinases by alternative initiation at AUG and CUG. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:655–664.
337. Sartor, O., F. S. Gregory, N. S. Templeton, S. Pawar, R. M. Perlmutter, and N. Rosen. 1989. Selective expression of alternative *lck* mRNAs in human malignant cell lines. *Mol. Cell Biol.* 9:2983–2988.
338. Sazer, S., and R. T. Schimke. 1986. A re-examination of the 5′ termini of mouse dihydrofolate reductase RNA. *J. Biol. Chem.* 261:4685–4690.
339. Scheidereit, C., J. A. Cromlish, T. Gerster, K. Kawahami, C-G. Balmaceda, R. A. Currie, and R. G. Roeder. 1988. A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homeobox protein. *Nature (Lond.).* 336:551–557.
340. Schindler, R., B. Clark, and C. A. Dinarello. 1990. Dissociation between interleukin-1β mRNA and protein synthesis in human peripheral blood mononuclear cells. *J. Biol. Chem.* 265:10232–10237.
341. Schneider, C., M. J. Owen, D. Banville, and J. G. Williams. 1984. Primary structure of human transferrin receptor deduced from the mRNA sequence. *Nature (Lond.).* 311:675–678.
342. Schoffl, F., M. Rieping, G. Baumann, M. Bevan, and S. Angermuller. 1989. The function of plant heat shock promoter elements in the regulated expression of chimaeric genes in transgenic tobacco. *Mol. Gen. Genet.* 217:246–253.
343. Schulz, P., R. Stucka, H. Feldmann, G. Combriato, H-G. Klobeck, and F. Fittler. 1988. Sequence of a cDNA clone encompassing the complete mature human prostate-specific antigen (PSA) and an unspliced leader sequence. *Nucleic Acids Res.* 16:6226.
344. Schwartz, S., B. Felber, D. Benko, E-M. Fenyo, and G. N. Pavlakis. 1990. Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type I. *J. Virol.* 64:2519–2529.
345. Schwartz, S., B. Felber, E-M. Fenyo, and G. N. Pavlakis. 1990. Env and vpu proteins of HIV type 1 are produced from multiple bicistronic mRNAs. *J. Virol.* 64:5448–5456.
346. Sedman, S. A., and J. E. Mertz. 1988. Mechanisms of synthesis of virion proteins from the functionally bigenic late mRNAs of simian virus 40. *J. Virol.* 62:954–961.
347. Sedman, S. A., P. J. Good, and J. E. Mertz. 1989. Leader-encoded ORFs modulate both the absolute and relative rates of synthesis of the virion proteins of SV40. *J. Virol.* 63:3884–3893.
348. Seeburg, P. H., and J. P. Adelman. 1984. Characterization of cDNA for precursor of human luteinizing hormone releasing hormone. *Nature (Lond.).* 311:666–668.
349. Seedorf, U., and G. Assmann. 1991. Cloning, expression, and nucleotide sequence of rat liver sterol carrier protein 2 cDNAs. *J. Biol. Chem.* 266:630–636.
350. Sehgal, A., N. Patil, and M. Chao. 1988. A constitutive promoter directs expression of the nerve growth factor receptor gene. *Mol. Cell Biol.* 8:3160–3167.
351. Seino, S., M. Seino, S. Nishi, and G. I. Bell. 1989. Structure of the human insulin receptor gene and characterization of its promoter. *Proc. Natl. Acad. Sci. USA.* 86:114–118.
352. Semba, K., M. Nishizawa, N. Miyajima, M. Yoshida, J. Sukegawa, Y. Yamanashi, M. Sasaki, T. Yamamoto, and K. Toyoshima. 1986. *Yes*-related protooncogene, *syn*, belongs to the protein tyrosine kinase family. *Proc. Natl. Acad. Sci. USA.* 83:5459–5463.
353. Shah, N. P., O. N. Witte, and C. T. Denny. 1991. Characterization of the BCR promoter in Philadelphia chromosome-positive and negative cell lines. *Mol. Cell Biol.* 11:1854–1860.
354. Sharpe, P. T., J. R. Miller, E. P. Evans, M. D. Burtenshaw, and S. J. Gaunt. 1988. Isolation and expression of a new mouse homeobox gene. *Development.* 102:397–407.
355. Shatkin, A. J. 1976. Capping of eucaryotic mRNAs. *Cell.* 9:645–653.
356. Shimomura, H., and A. Terada. 1990. Primary structure of the rat beta-1

adrenergic receptor gene. *Nucleic Acids Res.* 18:4591.
357. Simmons, D. M., J. Voss, H. Ingraham, J. Holloway, R. Broide, M. Rosenfeld, and L. W. Swanson. 1990. Pituitary cell phenotypes involve cell-specific Pit-1 mRNA translation and synergistic interactions with other classes of transcription factors. *Genes Dev.* 4:695–711.
358. Sims, J. E., C. March, D. Cosman, M. Widmer, H. R. MacDonald, C. McMahon, C. Grubin, J. Wignall, J. Jackson, S. Call, D. Friend, A. Alpert, S. Gillis, D. Urdal, and S. K. Dower. 1988. cDNA expression cloning of the IL-1 receptor, a member of the immunoglobulin superfamily. *Science (Wash. DC).* 241:585–588.
359. Siomi, H., H. Shida, S. H. Nam, T. Nosaka, M. Maki, and M. Hatanaka. 1988. Sequence requirements for nucleolar localization of HTLV Type I pX protein, which regulates viral RNA processing. *Cell.* 55:197–209.
360. Sleat, D. E., R. Hull,, P. C. Turner, and T. M. A. Wilson. 1988. Studies on the mechanism of translational enhancement by the 5'-leader sequence of tobacco mosaic virus RNA. *Eur. J. Biochem.* 175:75–86.
361. Snyder, L. C., S. P. Trusko, N. Freeman, J. R. Eshleman, S. S. Fakharzadeh, and D. L. George. 1988. A gene amplified in a transformed mouse cell line undergoes complex transcriptional processing and encodes a nuclear protein. *J. Biol. Chem.* 263:17150–17158.
362. Sobieszczuk, P. W., T. J. Gonda, and A. R. Dunn. 1989. Structure and biological activity of the transcriptional initiation sequences of the murine c-*myb* oncogene. *Nucleic Acids Res.* 17:9593–9611.
363. Soldati, T., B. W. Schafer, and J-C. Perriard. 1990. Alternative ribosomal initiation gives rise to chicken brain-type creatinine kinase isoproteins with heterogeneous amino termini. *J. Biol. Chem.* 265: 4498–4506.
364. Sopta, M., Z. F. Burton, and J. Greenblatt. 1989. Structure and associated DNA-helicase activity of a general transcription initiation factor that binds to RNA polymerase II. *Nature (Lond.).* 341:410–414.
365. Stamenkovic, I., H. C. Asheim, A. Deggerdal, H. Blomhoff, E. Smeland, and S. Funderud. 1990. The B cell antigen CD75 is a cell surface sialyltransferase. *J. Exp. Med.* 172:641–643.
366. Stanton, L. W., and J. M. Bishop. 1987. Alternative processing of RNA transcribed from *NMYC. Mol. Cell Biol.* 7:4266–4272.
367. Stephens, R. M., D. Derse, and N. R. Rice. 1990. Cloning and characterization of cDNAs encoding equine infectious anemia virus tat and putative rev proteins. *J. Virol.* 64:3716–3725.
368. Sternberg, E. A., G. Spizz, W. M. Perry, D. Vizard, T. Weil, and E. N. Olson. 1988. Identification of upstream and intragenic regulatory elements that confer cell-type-restricted and differentiation-specific expression on the muscle creatine kinase gene. *Mol. Cell Biol.* 8:2896–2909.
369. Stetler-Stevenson, W. G., P. D. Brown, M. Onisto, A. Levy, and L. A. Liotta. 1990. Tissue inhibitor of metalloproteinases-2 (TIMP-2) mRNA expression in tumor cell lines and human tumor tissues. *J. Biol. Chem.* 265:13933–13938.
370. Stirzaker, S. C., P. L. Whitfeld, D. Christie, A. Bellamy, and G. W. Both. 1987. Processing of rotavirus glycoprotein VP7: implications for the retention of the protein in the endoplasmic reticulum. *J. Cell Biol.* 105:2897–2903.
371. Stone, S. R., J. Hofsteenge, and B. A. Hemmings. 1987. Molecular cloning of cDNAs encoding two isoforms of the catalytic subunit of protein phosphatase 2A. *Biochemistry.* 26:7215–7220.
372. Straaten, F. V., R. Muller, T. Curran, C. Van Beveren, and I. M. Verma. 1983. Complete nucleotide sequence of a human c-*onc* gene: deduced amino acid sequence of the human c-*fos* protein. *Proc. Natl. Acad. Sci. USA.* 80:3183–3187.
373. Strait, K. A., H. L. Schwartz, A. Perez-Castillo, and J. H. Oppenheimer. 1990. Relationship of c-erbA mRNA content to tissue triiodothyronine nuclear binding capacity and function in developing and adult rats. *J. Biol. Chem.* 265:10514–10521.
374. Strathmann, M., and M. I. Simon. 1990. G protein diversity: a distinct class of α subunits is present in vertebrates and invertebrates. *Proc. Natl. Acad. Sci. USA.* 87:9113–9117.
374a. Strathmann, M. P., and M. I. Simon. 1991. Gα12 and Gα13 subunits define a fourth class of G protein α subunits. *Proc. Natl. Acad. Sci. USA.* 88:5582–5586.
375. Suh, P.-G., S. H. Ryu, K. Moon, H. W. Suh, and S. G. Rhee. 1988. Cloning and sequence of multiple forms of phospholipase C. *Cell.* 54:161–169.
376. Sukhatme, V. P., X. Cao, L. Chang, C-H. Tsai-Morris, D. Stamenkovich, P. Ferreira, D. Cohen, S. Edwards, T. Shows, T. Curran, M. LeBeau, and E. D. Adamson. 1988. A zinc finger-encoding gene coregulated with c-*fos* during growth and differentiation, and after cellular depolarization. *Cell.* 53:37–43.
376a. Sun, X. J., P. Rothenberg, C. R. Kahn, J. Backer, E. Araki, P. Wilden, D. Cahill, B. Goldstein, and M. F. White. 1991. Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein. *Nature (Lond.).* 352:73–77.
377. Suva, L. J., K. Mather, M. Gillespie, G. Webb, K. Ng, G. Winslow, W. I. Wood, T. J. Martin, and P. J. Hudson. 1989. Structure of the 5' flanking region of the gene encoding human parathyroid-hormone-related protein (PTHrP). *Gene.* 77:95–105.
378. Suzich, J. A., L. T. Kakach, and M. S. Collett. 1990. Expression strategy of a phlebovirus: biogenesis of proteins from the Rift Valley fever virus M. segment. *J. Virol.* 64:1549–1555.
379. Szala, S., M. Froehlich, M. Scollon, Y. Kasai, Z. Steplewski, H. Koprowski, and A. J. Linnenbach. 1990. Molecular cloning of cDNA for the carcinoma-associated antigen GA733-2. *Proc. Natl. Acad. Sci. USA.* 87:3542–3546.
380. Tacke, E., D. Prufer, F. Salamini, and W. Rohde. 1990. Characterization of a potato leafroll luteovirus subgenomic RNA: differential expression by internal translation initiation and UAG suppression. *J. Gen. Virol.* 71:2265–2272.
381. Takahashi, M., Y. Buma, and H. Hiai. 1989. Isolation of *ret* protooncogene cDNA with an amino-terminal signal sequence. *Oncogene.* 4:805–806.
382. Takaki, S., A. Tominaga, Y. Hitoshi, S. Mita, E. Sonoda, N. Yamaguchi, and K. Takatsu. 1990. Molecular cloning and expression of the murine interleukin-5 receptor. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:4367–4374.
383. Takeshima, H., S. Nishimura, T. Matsumoto, H. Ishida, K. Kangawa, N. Minamino, H. Matsuo, M. Ueda, M. Hanaoka, T. Hirose, and S. Numa. 1989. Primary structure and expression from complementary DNA of skeletal muscle ryanodine receptor. *Nature (Lond.).* 339:439–445.
384. Tal, M., C. R. King, M. H. Kraus, A. Ullrich, J. Schlessinger, and D. Givol. 1987. Human HER2 (*neu*) promoter: evidence for multiple mechanisms for transcriptional initiation. *Mol. Cell Biol.* 7:2597–2601.
385. Tamkun, J. W., D. DeSimone, D. Fonda, R. Patel, C. Buck, A. Horwitz, and R. O. Hynes. 1986. Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell.* 46:271–282.
386. Tamura, S., K. Lynch, J. Larner, J. Fox, A. Yasui, K. Kikuchi, Y. Suzuki, and S. Tsuiki. 1989. Molecular cloning of rat type 2C (IA) protein phosphatase mRNA. *Proc. Natl. Acad. Sci. USA.* 86:1796–1800.
387. Tedder, T. F., M. Streuli, S. Schlossman, and H. Saito. 1988. Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen of human B lymphocytes. *Proc. Natl. Acad. Sci. USA.* 85:208–212.
388. Tedder, T. F., G. Klejman, S. Schlossman, and H. Saito. 1989. Structure of the gene encoding the human B lymphocyte differentiation antigen CD20 (B1). *J. Immunology.* 142:2560–2568.
389. Tedder, T. F., and C. M. Isaacs. 1989. Isolation of cDNAs encoding the CD19 antigen of human and mouse B lymphocytes. *J. Immunology.* 143:712–717.
390. Teruya, J. H., S. Kutsunai, D. Spear, P. Edwards, and C. F. Clarke. 1990. Testis-specific transcription initiation sites of rat farnesyl pyrophosphate synthetase mRNA. *Mol. Cell Biol.* 10:2315–2326.
391. de The, H., A. Marchio, P. Tiollais, and A. Dejean. 1987. A novel steroid thyroid hormone receptor-related gene inappropriately expressed in human hepatocellular carcinoma. *Nature (Lond.).* 330:667–670.
392. Theill, L. E., O. Wiborg, and J. Vuust. 1987. Cell-specific expression of the human gastrin gene: evidence for a control element located downstream of the TATA box. *Mol. Cell Biol.* 7:4329–4336.
393. Thomas, G., J. Martin-Perez, M. Siegmann, and A. M. Otto. 1982. The effect of serum, EGF, PGF2α and insulin on S6 phosphorylation and the initiation of protein and DNA synthesis. *Cell.* 30:235–242.
394. Thompson, C. C., C. Weinberger, R. Lebo, and R. M. Evans. 1987. Identification of a novel thyroid hormone receptor expressed in the mammalian central nervous system. *Science (Wash. DC).* 237:1610–1614.
395. Timmons, M. S., and O. N. Witte. 1989. Structural characterization of the BCR gene product. *Oncogene.* 4:559–567.
396. Tobimatsu, T., and H. Fujisawa. 1989. Tissue-specific expression of four types of rat calmodulin-dependent protein kinase II mRNAs. *J. Biol. Chem.* 264:17907–17912.
397. Tollefson, A. E., P. Krajcsi, M. Pursley, L. Gooding, and W. Wold. 1990. A 14,500 MW protein is coded by region E3 of group C human adenoviruses. *Virology.* 175:19–29.
398. Tropschug, M. 1990. Nucleotide sequence of the gene coding for cyclophilin/peptidyl-prolyl cis-trans isomerase of *Neurospora crassa. Nuc. Acids Res.* 18:190.
399. Tropschug, M., D. W. Nicholson, F. Hartl, H. Kohler, N. Pfanner, E. Wachter, and W. Neupert. 1988. Cyclosporin A-binding protein (cyclophilin) of *Neurospora crassa. J. Biol. Chem.* 263:14433–14440.
400. Tsuji, T., F. Okada, K. Yamaguchi, and T. Nakamura. 1990. Molecular cloning of the large subunit of transforming growth factor type β masking protein and expression of the mRNA in various rat tissues. *Proc. Natl. Acad. Sci. USA.* 87:8835–8839.
401. Tsujimoto, Y., and C. M. Croce. 1986. Analysis of the structure, transcripts, and protein products of *bcl-2*, the gene involved in human follicular lymphoma. *Proc. Natl. Acad. Sci. USA.* 83:5214–5218.
402. Uze, G., G. Lutfalla, and I. Gresser. 1990. Genetic transfer of a functional human interferon α receptor into mouse cells: cloning and expression of its cDNA. *Cell.* 60:225–234.
403. Vogel, U. S., R. Dixon, M. Schaber, R. Diehl, M. Marshall, E. M. Scolnick, I. S. Sigal, and J. B. Gibbs. 1988. Cloning of bovine GAP and its interaction with oncogenic *ras* p21. *Nature (Lond.).* 335:90–93.
404. Vu, T.-K., D. T. Hung, V. I. Wheaton, and S. R. Coughlin. 1991. Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell.* 64:1057–1068.
405. Vu, T. H., G. R. Martin, P. Lee, D. Mark, A. Wang, and L. T. Williams.

1989. Developmentally regulated use of alternative promoters creates a novel PDGF receptor transcript in mouse teratocarcinoma and embryonic stem cells. *Mol. Cell Biol.* 9:4563–4567.

406. Wainwright, B. J., P. Scambler, P. Stanier, E. Watson, G. Bell, C. Wicking, X. Estivill, M. Courtney, A. Boue, P. Pedersen, R. Williamson, and M. Farrall. 1988. Isolation of a human gene with protein sequence similarity to human and murine int-1 and the *Drosophila* segment polarity mutant *wingless*. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1743–1748.
407. Wang, W-P., K. Lehtoma, M. L. Varban, I. Krishnan, and I-M. Chiu. 1989. Cloning of the gene coding for human class 1 heparin-binding growth factor and its expression in fetal tissues. *Mol. Cell Biol.* 9:2387–2395.
408. Watanabe, T., H. Yonehura, K. Terazono, H. Yamamoto, and H. Okamoto. 1990. Complete nucleotide sequence of human *reg* gene and its expression in normal and tumoral tissues. *J. Biol. Chem.* 265:7432–7439.
409. Waterhouse, P., R. Khokha, and D. T. Denhardt. 1990. Modulation of translation by the 5′ leader sequence of the mRNA encoding murine tissue inhibitor of metalloproteinases. *J. Biol. Chem.* 265:5585–5589.
410. Wei, L. L., C. Gonzalez-Aller, W. M. Wood, L. A. Miller, and K. B. Horwitz. 1990. 5′-Heterogeneity in human progesterone receptor transcripts predicts a new amino-terminal truncated "C" receptor and unique A-receptor messages. *Mol. Endocrinol.* 4:1833–1840.
411. Weil, D., S. Brosset, and F. Dautry. 1990. RNA processing is a limiting step for murine tumor necrosis factor β expression in response to interleukin-2. *Mol. Cell Biol.* 10:5865–5875.
412. Weinberg, A. D., and S. L. Swain. 1990. IL-2 receptor (Tac antigen) protein expression is down-regulated by the 5′ untranslated region of the mRNA. *J. Immunol.* 144:4712–4720.
413. Weinberger, C., C. C. Thompson, E. S. Ong, R. Lebo, D. J. Gruol, and R. M. Evans. 1986. The c-*erb-A* gene encodes a thyroid hormone receptor. *Nature (Lond.).* 324:641–646.
414. Williams, M. A., and R. A. Lamb. 1989. Effect of mutations and deletions in a bicistronic mRNA on the synthesis of influenza B virus NB and NA glycoproteins. *J. Virol.* 63:28–35.
415. Wilson-Rawls, J., S. Saha,, P. Krajcsi, A. E. Tollefson, L. R. Gooding, and W. Wold. 1990. A 6700 MW membrane protein is encoded by region E3 of adenovirus type 2. *Virology.* 178:204–212.
416. Xiao, J. H., I. Davidson, H. Matthes, J.-M. Garnier, and P. Chambon. 1991. Cloning, expression, and transcriptional properties of the human enhancer factor TEF-1. *Cell.* 65:551–568.
417. Xie, W., J. G. Chipman, D. L. Robertson, R. L. Erikson, and D. L. Simmons. 1991. Expression of a mitogen-responsive gene encoding prostaglandin synthase is regulated by mRNA splicing. *Proc. Natl. Acad. Sci. USA.* 88:2692–2696.
418. Yamamoto, M., L. J. Ko, M. W. Leonard, H. Beug, S. H. Orkin, and J. D. Engel. 1990. Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. *Genes Dev.* 4:1650–1662.
419. Yamanashi, Y., S-I. Fukushige, K. Semba, J. Sukegawa, N. Miyajima, K-I. Matsubara, T. Yamamoto, and K. Toyoshima. 1987. The *yes*-related cellular gene *lyn* encodes a possible tyrosine kinase similar to $p56^{lck}$. *Mol. Cell Biol.* 7:237–243.
420. Yamasaki, K., T. Taga, Y. Hirata, H. Yawata, Y. Kawanishi, B. Seed, T. Taniguchi, T. Hirano, and T. Kishimoto. 1988. Cloning and expression of the human interleukin-6 (BSF-2/IFNβ2) receptor. *Science (Wash. DC).* 241:825–828.
421. Yan, C., and I. Tamm. 1991. Molecular cloning and characterization of interferon α/β response element binding factors of the murine (2′-5′) oligoadenylate synthetase ME-12 gene. *Proc. Natl. Acad. Sci. USA.* 88:144–148.
422. Yoo, O., H. Yoon, K. Baek, C. Jeon, K. Miyamoto, A. Ueno, and K. Agarwal. 1991. Cloning, expression and characterization of the human transcription elongation factor, TFIIS. *Nucleic Acids Res.* 19:1073–1079.
423. Yoshida, T., K. Miyagawa, H. Odagiri, H. Sakamoto, P. Little, M. Terada, and T. Sugimura. 1987. Genomic sequence of *hst*, a transforming gene encoding a protein homologous to fibroblast growth factors and the *int-2*-encoded protein. *Proc. Natl. Acad. Sci. USA.* 84:7305–7309.
424. Young, J. C., M. L. Gishizky, and O. N. Witte. 1991. Hyperexpression of interleukin-7 is not necessary or sufficient for transformation of a pre-B lymphoid cell line. *Mol. Cell Biol.* 11:854–863.
425. Zakany, J., C. K. Tuggle, M. Patel, and M. C. Nguyen-Huu. 1988. Spatial regulation of homeobox gene fusions in the embryonic central nervous system of transgenic mice. *Neuron.* 1:679–691.
426. Zelent, A., C. Mendelsohn, P. Kastner, A. Krust, J-M. Garnier, F. Ruffenach, P. Leroy, and P. Chambon. 1991. Differentially expressed isoforms of the mouse retinoic acid receptor β are generated by usage of two promoters and alternative splicing. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:71–81.
427. Zerial, M., L. Toschi, R-P. Ryseck, M. Schuermann, R. Muller, and R. Bravo. 1989. The product of a novel growth factor activated gene, *fos B*, interacts with JUN proteins enhancing their DNA binding activity. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:805–813.
428. Zheng, X. M., D. Black, P. Chambon, and J. M. Egly. 1990. Sequencing and expression of complementary DNA for the general transcription factor BTF3. *Nature (Lond.).* 344:556–559.
429. Zon, I. I., S-F. Tsai, S. Burgess, P. Matsudaira, G. Bruns, and S. H. Orkin. 1990. The major human erythroid DNA-binding protein (GF-1): primary sequence and localization of the gene to the X chromosome. *Proc. Natl. Acad. Sci. USA.* 87:668–672.
430. Zot, A. S., and D. M. Fambrough. 1990. Structure of a gene for a lysosomal membrane glycoprotein (LEP100). *J. Biol. Chem.* 265:20988–20995.