Title
The structure and regulation of the rat glucocorticoid receptor gene

Author
Jacobson, Michael David

Publication Date
1991

Peer reviewed|Thesis/dissertation

THE STRUCTURE AND REGULATION OF THE RAT GLUCOCORTICOID RECEPTOR GENE

by

MICHAEL DAVID JACOBSON

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

NEUROSCIENCE

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

Dedication

This thesis is dedicated to Ron Harris-Warrick, who
fed my passions for neurobiology and research,
before I knew what I was getting into.

Preface

I came to U.C.S.F. as a neuroscience student, and am leaving as a neuroscientist with training in molecular biology. I took this route because I believe that studying gene expression will help unlock important secrets of how the brain develops and changes: what neurobiologists call the development and plasticity of the nervous system. This is because changes in gene expression underlie the specification of cell phenotype: what makes one cell a neuron and another liver cell, and how those cells respond to their environment.

The glucocorticoid receptor gene serves as a paradigm for studying systems of gene expression in complex eukaryotic systems. An obvious reason is that the receptor is itself a transcription factor, a regulator of other genes. It has also been studied intensively and so we know a lot about how this protein works. The regulation of this receptor is itself an important determinant of cell phenotype. Understanding this regulation is a first step in piecing together the network of regulatory interactions that controls the phenotype of every cell. Thus to understand how expression of the glucocorticoid receptor is regulated, I began by trying to understand its gene.

What this thesis will show is that the receptor gene has a complex structure, and that it appears to be regulated in complex ways. This should not come as a surprise, given the complexity of presumably simpler prokaryotic regulatory networks. Nevertheless, the details of the receptor gene's structure suggest many possibilities for how this gene is regulated. This thesis begins an investigation of these possibilities. As future work deepens our understanding of this gene's regulation and integrates this knowledge with information about the regulation of other regulatory proteins with which

receptor interacts, it will become possible to put together a picture of how complex gene regulation occurs and how cell phenotype is specified.

While working on this thesis I received advice and support from many colleagues and teachers at U.C.S.F., whom I would like to thank and acknowledge here. The neuroscience program was representative of the attitude at U.C.S.F. in allowing me to cross departmental boundaries without presenting any real barriers, and even encouraging me to become a molecular biologist regardless of whether I studied nerve cells or not. My laboratory rotation advisers – Lily and Yuh Nung Jan, Jim Hudspeth, and Mike Stryker – provided me with diverse and stimulating short research projects. In addition, Lou Reichardt gave me encouragement and valuable career advice throughout my graduate studies. I received specific help on my project from William Mobley and his laboratory, who helped with the transcript expression studies (chapter three); Mary Dallman helped me place these studies in their physiological context; Roger Miesfeld and Sandro Rusconi, who were postdoctoral fellows in Keith Yamamoto's lab, helped train me and get me started on the gene cloning and mapping (chapter two); Leslie Taylor helped me with the gene alignments (chapter two); and Stanton Glantz advised me on the proper use of statistics (chapters two and three). My colleagues in Keith Yamamoto's lab gave me more help and support than I can mention here, except to note that the members of the "annex" in which I worked were especially supportive and became real friends. Keith Yamamoto put together this group of excellent scientists and colleagues, and maintained a laboratory that was consistently productive and stimulating. Keith set an example of hard work, dedication, commitment and integrity. He taught me to simplify the complex and to think critically about every aspect of my work; I hope I can live up to his standards throughout my career. I want especially

to thank my friend and companion Madeline Martell for remaining beside me when it really mattered and for keeping me from becoming a complete lab rat. And finally, my parents have supported me regardless of what I did and encouraged my interest in science and my curiosity of how things work, well before my first test tube.

My Ph.D. work was supported by a fellowship from the National Science Foundation, a Neuroscience Program traning grant from the National Institutes of Health, and by grants from the National Science Foundation and National Institutes of Health to Keith Yamamoto.

The Structure and Regulation of the Rat Glucocorticoid Receptor Gene

a Ph.D. Dissertation by Michael David Jacobson;

sponsored by Keith R. Yamamoto

Abstract

The glucocorticoid receptor is up- or down-regulated in a cell-specific manner, yet the mechanisms of this regulation are unknown. To investigate the regulation of the rat glucocorticoid receptor gene, I cloned, sequenced, and mapped this gene and its major transcripts. The glucocorticoid receptor gene has a complex organization, consisting of eight coding exons distributed over at least 125 kilobases of DNA, and containing two polyadenylation sites and a family of at least eight alternative 5' untranslated exons each spliced to the first coding exon. This family of 5' exons is expressed from two or more promoters. I mapped and sequenced the two 5' exons (1A and 1B) and 5' flanking regions associated with the most abundantly expressed transcripts (A and B) in HTC cells. I also mapped the promoter associated with exon 1B. This promoter contains multiple transcript initiation sites, is rich in G and C nucleotides and contains multiple SP-1 consensus sequences and a degenerate CTF consensus sequence, yet lacks a TATA box. Both exons 1A and 1B contain open reading frames (uORFs), and exon 1B contains a mini-cistron 40 amino-acids long.

These complexities at the 5' end suggested possibilities for regulation. To explore the function of multiple promoters, I initiated a survey of transcript expression patterns in the rat using an RNase protection assay. I

found that transcripts A and B were most abundant in non-lymphoid tissues; however, at least one other transcript, detected by this assay but otherwise uncharacterized, was expressed at relatively higher levels in lymphoid cells and appeared to be positively regulated by glucocorticoids in the thymus. These results indicate that glucocorticoid receptors are expressed from multiple promoters, and that these promoters might be differentially regulated in a tissue-specific manner.

To explore the function of the 5' exon family and upstream ORFs, I examined receptors expressed in mouse S49 cells with the mutant phenotype nt^{i}. mRNA from these cells is missing the first coding exon, and instead use upstream initiators from 5' exons to express receptors containing novel amino-terminal sequences. In limited assays, however, I failed to discern a function for these novel sequences.

Table of Contents

List of Tables

List of Figures

Abbreviations

Those abbreviations followed by an asterix are used without definition.

Abbreviation		Name
ACTH		Adrenocorticotropic hormone
AdML		Adenovirus major-late (promoter)
AR		Androgen Receptor
ATP	*	Adenosine 5' triphosphate
cAMP	*	Cyclic AMP
CAT		Chloramphenicol acetyl tranferase
CBG		Corticosterone binding globulin
CoA	*	Coenzyme A
ConA		Conconavalin A
CRF		Corticotropin releasing factor
CTF		CAAT Transcription Factor
Dex		Dexamethasone
DNA	*	Deoxyribonucleic acid
DTT	*	Dithiothreitol
EcR		Ecdysone Receptor
EDTA	*	Ethylenediaminetetraacetate
ER		Estrogen Receptor
ERR		Estrogen-related Receptor
Gal		Gal4, galactose
GLM		General linear modeling
GPDH		Glycerol-phosphate dehydrogenase
GR		Glucocorticoid Receptor
GRE		Glucocorticoid response element
HPA Axis		Hypothalamic-pituitary-adrenal axis
HRE		Hormone response element
Hsp 90	*	Heat-shock protein 90
LDL		Low density lipoprotein
Lex		LexA
LTR		Long-terminal repeat
MMTV		Mouse mammary tumor virus
MR		Mineralocorticoid Receptor
NGF		Nerve growth factor
nt^i		Nuclear transfer increased
ORF		Open reading frame
PCR		Polymerase chain reaction
POMC		Pro-opiomelanocortin
PR		Progesterone Receptor
RACE		Rapid Amplification of cDNA Ends
RAR		Retinoic Acid Receptor

RNA	*	Ribonucleic acid
SDS		Sodium dodecyl sulfate
TR		Thyroid Hormone (T3/T4) Receptor
Tris	*	Tris(hydroxymethyl)aminomethane
UAS		Upstream acivating sequence (of yeast)
UTR		Untranslated region (of RNA)
VDR		Vitamin D3 Receptor

Chapter One: Introduction

Questions and Perspectives

Steroid hormones, derivatives of the biologically essential molecule cholesterol, control developmental and physiological processes in organisms as diverse as fungi and humans. The effects of steroid hormones are mediated by hormone-specific receptor proteins. Upon steroid binding, these proteins associate with specific chromosomal sites and selectively regulate the transcription rate of target genes (Fig. 1.1). Each hormone induces a multitude of different responses in different cells and tissues. These specific hormone-induced changes depend on the cell and its environment, suggesting a combinatorial model of gene regulation involving a network of multiple transcriptional regulatory proteins (Yamamoto, 1985).

This thesis addresses two basic questions. First, how was this steroid-specific gene network generated and modified during evolution? Second, what is the basis for the diversity of tissue-specific responses to steroid hormones? I have approached these questions by focusing on the gene structure and expression of the glucocorticoid receptor.

This receptor has been the subject of intense scrutiny by a number of laboratories, and was the first steroid receptor to be cloned (Hollenberg et al., 1985; Miesfeld et al., 1984, 1986). Its structure and function has served as a paradigm for the analysis of other steroid receptors. Since the glucocorticoid receptor was cloned, more than thirty steroid receptor-like proteins have been cloned and sequenced. These proteins all share a strongly conserved DNA-binding domain with a C_4/C_5 zinc-finger structure (Freedman et al., 1988; Härd et al., 1990b) and belong to a superfamily of ligand-binding zinc-finger proteins, the "nuclear receptor" superfamily (Amero et al., 1991; Evans, 1988; Green and Chambon, 1988).

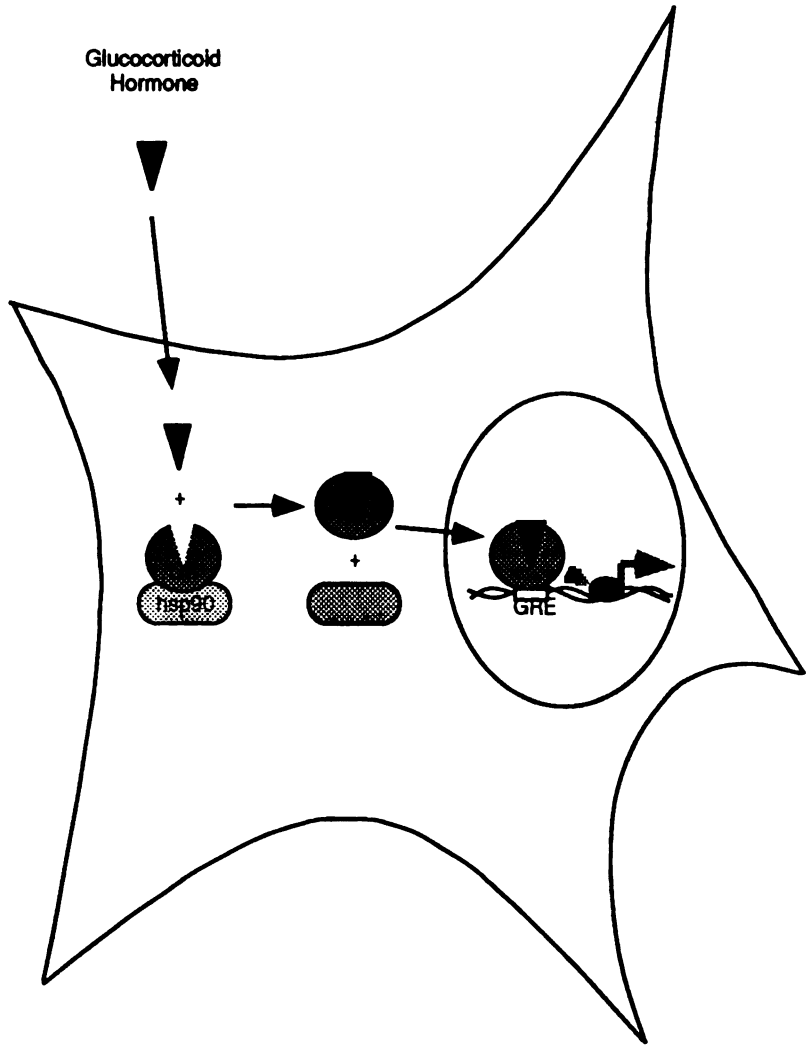


Figure 1.1. Glucocorticoid hormone actions are mediated through a cytoplasmic receptor protein. Serum glucocorticoids (▼) cross the cytoplasmic membrane, either by diffusion or via transport proteins, and bind to glucocorticoid receptor complexes. Hsp90 facilitates hormone binding to the receptor (Bresnick et al., 1989). Upon steroid binding, the complex dissociates, activating steroid-bound receptor (Mendel et al., 1986; Sanchez et al., 1987). Activated receptor translocates to the nucleus, where it binds to specific DNA sequences called glucocorticoid response elements (GREs). DNA-bound receptor regulates the rate of initiation of transcription from nearby promoters. Other steroid hormones appear to behave similarly, regulating the transcription rate of specific genes through the action of steroid-specific receptor proteins.

Gene structures have frequently revealed clues to the function and evolution of the proteins they encode (reviewed in Gilbert, 1987). To investigate the evolution of the glucocorticoid receptor, I determined the intron-exon organization of its gene. Its organization is similar to that of other nuclear receptor superfamily genes, suggesting that these genes descended from a common ancestral ligand-binding zinc-finger protein. I also cloned, sequenced and mapped the 5' ends of the receptor gene and its promoter and began a preliminary investigation of regulation of glucocorticoid receptor expression.

The glucocorticoid receptor gene is organized into eight coding exons distributed over at least 125 kilobases of DNA, contains two polyadenylation sites and at least eight alternative 5' untranslated exons, all transcribed from two or more promoters. The complexity of the receptor gene promoter region suggests both transcriptional and post-transcriptional strategies of regulating glucocorticoid receptor levels.

Evolution and gene structure

Steroids are evolutionarily conserved molecules

All steroids share a common biosynthetic pathway from cholesterol. Cholesterol is synthesized *de novo* from the condensation of five-carbon isoprene units. These terpenes are the building blocks for a large number of biologically important compounds, including cholesterol, retinoids, insect and plant hormones, and respiratory coenzyme Q (Fig. 1.2). Sterols and steroids are themselves highly conserved and nearly ubiquitous biomolecules.

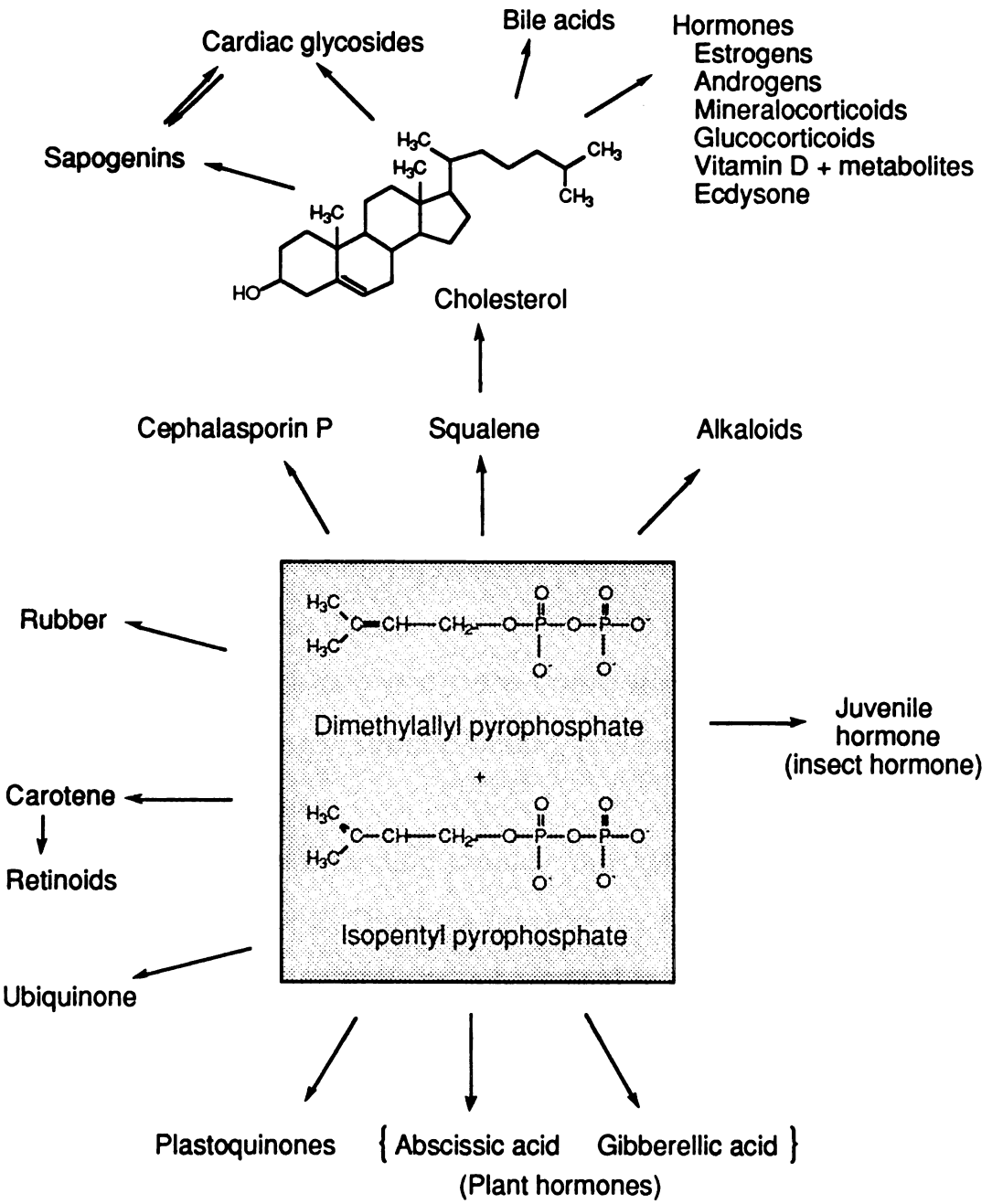


Figure 1.2. Biosynthesis of terpenes. From (Norman and Litwack, 1987). The isoprenoid compounds dimethylallyl phyrophosphate and isopentyl phrophosphate (in box) are synthesized from acetyl CoA, with mevalonic acid as an intermediate. These terpenes are the precursors for a large number of biologically important compounds, including cholesterol (precursor to steroid hormones and vitamin D) and carotene (precursor to vitamin A and other retinoids).

Cholesterol is found in nearly all organisms, including blue-green algae and bacteria (for review, see Sandor and Mehdi, 1979). Most eukaryotic organisms have the enzymatic capability to synthesize cholesterol from acetate. In some organisms that cannot synthesize cholesterol, such as arthropods, cholesterol is an essential dietary nutrient. It is an integral component of cell membranes and necessary for the production of biologically important compounds in both animals and plants (e.g. saponins, glycosids, and phytoecdysones). Steroids and steroid synthetic capacity have been found in both vertebrate and invertebrate species, as well as in plants and fungi (Sandor and Mehdi, 1979). Some prokaryotes are also able to synthesize sterols *de novo* (Nes and McKean, 1977) as well as to metabolize steroids (Berg et al., 1975; Berg et al., 1976; Capek et al., 1966; Charney and Herzog, 1967). Interestingly, specific binding proteins for the steroids estrogen and 11-deoxycortisone have been identified in certain species of yeast, along with endogenous receptor-binding steroids (Burshell et al., 1984; Feldman et al., 1982; Loose et al., 1981). Though a function has not been established for these yeast steroid-binding proteins, a hormonal role for steroids has been established for another fungus species, the water mold *Achlya* (Timberlake and Orr, 1984). This ubiquity in evolution suggests that steroids are ancient signalling molecules. How might the steroids have acquired their roles? And what is the origin of their receptors?

Cholesterol is the only major sterol product in mammalian systems and its structure is conserved despite its wide distribution throughout the biosphere. The biosynthesis of steroids from sterols requires molecular oxygen, a mixed-function oxidase (e.g. cytochrome P_{450}) and electron donors and intermediates in the electron transport chain. These molecules are all ancient biosynthetic components, presumably present in primordial cells

capable of synthesizing or utilizing sterols. Thus steroids may have arisen as metabolic by-products of sterol metabolism. The conservation of steroid systems from vertebrates to bacteria and blue-green algae suggests that these compounds may have had an essential role early in the evolution of unicellular life (Sandor and Mehdi, 1979).

Tomkins (1975) suggested that intercellular communication in metazoans originated in unicellular organisms with the intracellular regulation of metabolic processes. Using cyclic nucleotides as a model, he hypothesized that cAMP was originally formed as a biosynthetic accident, but came to symbolize the metabolic state of the cell. The ability to read this "metabolic code" conferred an adaptive advantage on descendants. A simple form of metabolic regulation may have arisen that allowed important metabolic products (e.g. ATP) to be maintained at relatively constant levels despite changes in environmental conditions. This regulation was achieved by a direct chemical relationship between regulatory effector molecules (e.g. glucose) and their effects (ATP synthesis). Thus, substrates or end products of metabolism could affect their own metabolism. This form of regulation could be positive, as in enzyme induction (e.g. by a substrate), or negative, as in feedback inhibition of enzyme biosynthesis (e.g. by an end product). Since the regulatory molecules were themselves important metabolic intermediates, however, this simple form of regulation could not buffer the intracellular environment from rapid changes in essential nutrients. Since modern organisms display more sophisticated regulatory behavior, Tomkins posited a more complex form of regulation, presumably of later evolutionary origin, whereby the effector molecule was not an essential metabolic intermediate. This molecule could symbolize the metabolic state of the cell without itself being required for the metabolic processes it represented. The relationship

between particular environmental conditions and their corresponding intracellular events could be considered a metabolic code, whereby a specific symbol represents a unique state of the environment. For example in *E. coli*, elevated cyclic AMP levels symbolize glucose starvation, and elevated ppGpp levels symbolize amino-acid deprivation. Each effector molecule (cAMP, ppGpp) elicits a specific "domain" of coordinated responses. For example the "stringent response" to amino acid starvation in *E. coli* is the domain corresponding to elevated ppGpp levels.

Similarly, steroid hormones may have arisen first as metabolic by-products of cholesterol (or more generally, terpene) metabolism. Perhaps they served originally as intracellular symbols in regulating their metabolic precursors. For example, HMG CoA reductase, which catalyzes the limiting step in cholesterol biosynthesis from acetate, is regulated by both cholesterol (simple regulation) and a number of steroids (complex regulation) (Goldstein and Brown, 1990). The stability of steroids, however, would have also made them useful in communication between cells. This metabolic coding of steroids (and other terpene derivatives) in unicellular organisms might have evolved into the metazoan endocrine systems we recognize in modern organisms. Similarly, in unicellular organisms, such signals could have been adapted to roles in intra- and inter-specific communication as pheromones, such as those involved in mating (as in the steroid hormones in the water mould *Achlya*), or in defense systems (as in the plant phytoecdysones, saponins, and cardenolides).

Steroid receptors are molecular adaptors

Steroids are relatively simple molecules and are well-conserved across biological systems. The structural similarities among steroids is partly due to the occurance of only one stereoisomer of cholesterol (out of 256 possible forms) in biological systems. Their simplicity and relative uniformity makes steroids energetically cheap to manufacture, a useful characteristic for a signalling molecule. However, as Tomkins argues (Tomkins, 1975), this simplicity also makes them incapable of mediating alone the diverse and complex response domains observed in modern organisms. These molecular symbols required specific adaptor molecules to "read" the metabolic code, analogous to the role of tRNA molecules for reading the genetic code. Thus, protein receptor molecules for terpene derivatives such as steroids and sterols may have arisen first as proteins involved directly in the metabolism of their precursors. Steroid receptors, for example, may have evolved from enzymes involved in cholesterol metabolism (Yamamoto, 1985). One possibility is that they evolved from transport or carrier proteins. For example, the bacterium *Pseudomonas testosteroni*, upon adaptation to growth on testosterone or certain other steroids, induces a specific steroid binding protein (Watanabe et al., 1973). Thus, steroid receptors could have evolved from transport proteins involved in the uptake of terpene derivatives such as cholesterol. Alternatively, these receptors may have arisen from metabolic enzymes involved in the synthesis and/or breakdown of terpene derivatives. Perhaps a simple feedback control protein arose by the fortuitous combination of a sterol recognition motif (e.g. from a metabolic enzyme) and a protein module capable of binding DNA.

Exon shuffling is a mechanism for the evolution of proteins

Transcription units of eukaryotic genes are organized into discrete segments of coding DNA (exons) separated by segments of non-coding DNA (introns). The introns are removed (spliced) from the nascent mRNA molecule prior to transport of the mature mRNA to the cytoplasm. This organization of genes has suggested a possible mechanism for early protein evolution by the recombination of protein domains (reviewed in Gilbert, 1987).

With this view it is thought that genes were originally assembled by combining exon modules that code for distinct functional elements (Gilbert, 1978; Gilbert, 1987), stably folding peptides (Blake, 1979), or compact modules (Go and Nosaka, 1987). A striking example of exon shuffling is found in the gene coding for low density lipoprotein (LDL) receptor, which appears to be a mosaic of separate exon-encoded domains shared with proteins in different families (Südhof et al., 1985a,b). According to the exon shuffling hypothesis, introns are primitive features retained through evolution, and their positions should reflect the evolutionary history of the genes they interrupt.[1] Studies of the genes for several enzymes indicate that introns predate the divergence of plants and animals (Marchionni and Gilbert, 1986), as well as the endosymbiotic incorporation of chloroplasts and mitochondria (Obaru et al., 1988; Quigley et al., 1988; Shih et al., 1988). The original exons may have been only 15-20 aa long, and coded for distinct functional peptides (Dorit et al., 1990; Gilbert, 1987). The larger and more complex exons of the present day

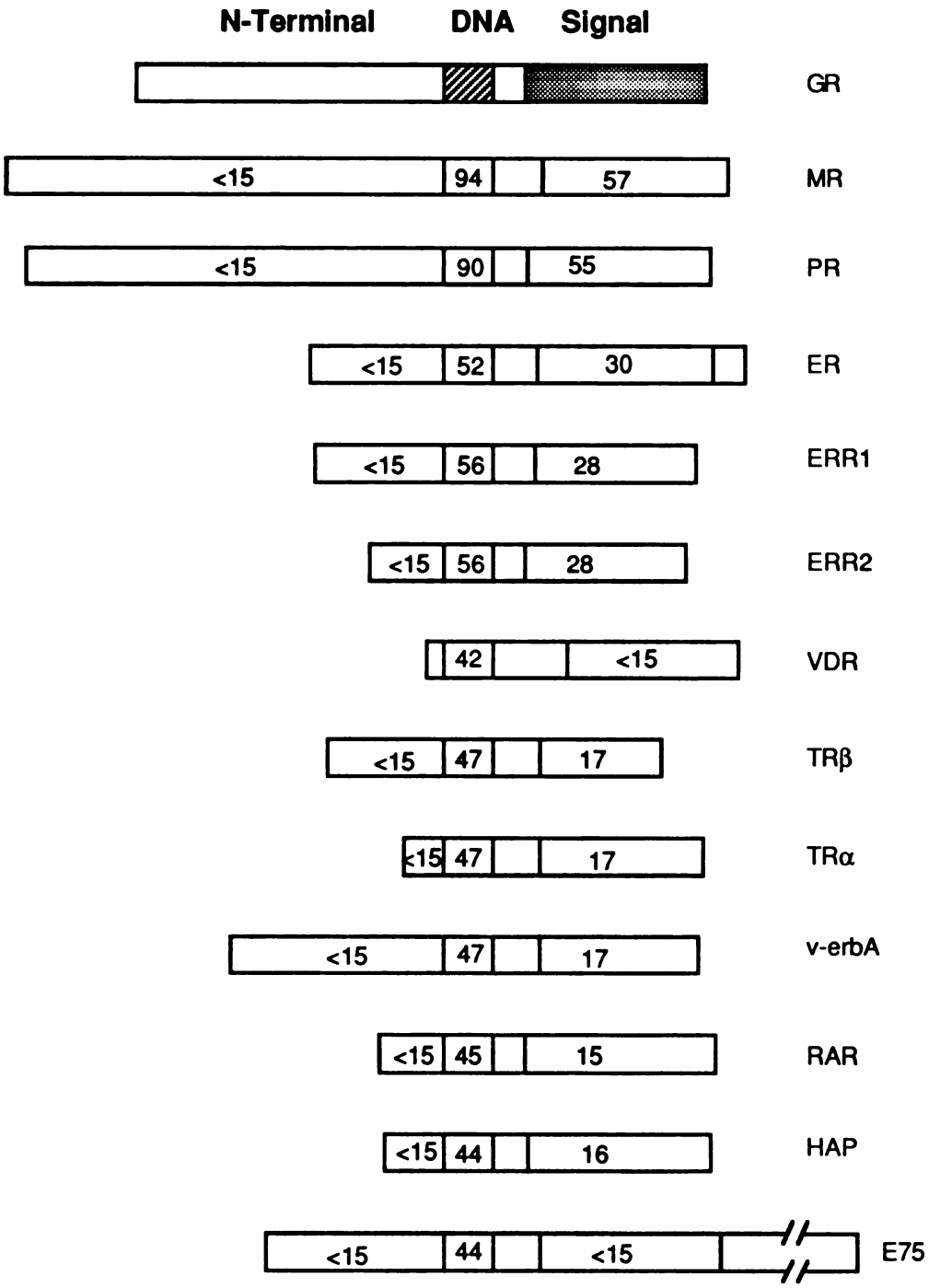
[1] An alternative hypothesis is that introns are a relatively new feature of eukaryotic genomes and were created by the insertion of DNA sequences such as retroposons by bursts of parasitic elements invading early (continuous) eukaryotic coding regions (Cavalier-Smith, 1985; Sharp, 1985; Rogers, 1985; review in Hickey et al., 1986, 1989).

may comprise these ancient functional units, the result of intron sliding and intron loss (Craik et al., 1983; Gilbert, 1987).

Steroid receptors belong to a large gene family

The receptors for glucocorticoids and other steroids belong to a superfamily of ligand-activated transcription factors that also includes receptors for thyroid hormone, retinoic acid, vitamin D3, and a number of receptor-like molecules for which no ligand is known (Evans, 1988; Green and Chambon, 1988)(Fig. 1.3). These nuclear receptor superfamily proteins share a conserved region that corresponds to two C_4/C_5 "zinc fingers"; this region of the steroid, retinoic acid, and thyroid hormone receptors has been shown to bind specific DNA sequences (Beato, 1989). These proteins also share a less well-conserved domain, carboxy-terminal to the zinc-finger region, which binds to receptor-specific ligands.

Complete and partial gene structures have been determined for some members of this family, including other steroid receptors (Arriza et al., 1987; Brinkmann et al., 1989; Huckaby et al., 1987; Hughes et al., 1988; Kuiper et al., 1989; Lazar et al., 1989; Lehmann et al., 1991; Petkovich et al., 1987; Ponglikitmongkol et al., 1988; Ritchie et al., 1989; Segraves and Hogness, 1990; Watson and Milbrandt, 1989). A prominent feature of these genes is that the two zinc fingers are split into separate exons. The position of the intron/exon borders separating the two fingers is conserved among the steroid receptors but differs from the border found in the retinoic acid and thyroid hormone receptors (in which it is also conserved) (Green and Chambon, 1988). This subgrouping by intron/exon position parallels differences between these subfamilies based on protein sequence similarities (Amero et al., 1991) as well



Adapted from: Evans (1988) *Science* **240**:889

Figure 1.3. The glucocorticoid receptor belongs to a large gene family. From Evans, 1988. Sequence similarity (% identity) between human GR and other human nuclear receptor proteins is shown for the zinc-finger (DNA), signal transduction, and N-terminal domains.

as their ligand types (e.g. steroids vs. retinoids), and so may reflect the evolution of nuclear receptors. As I will discuss in chapter two, the other intron/exon borders of nuclear receptors (including glucocorticoid receptor's) also fit this general pattern.

The glucocorticoid receptor is organized into discrete functional domains

The exon shuffling hypothesis predicts that exons encode functional domains of proteins. If the ancestral steroid (or terpenoid) receptor arose from the recombination of exons encoding such modules, then its modern descendants might reflect this initial organization. Biochemical and genetic studies suggest that hormone receptors are organized into discrete functional domains (reviewed in Green and Chambon, 1988). This is illustrated by the functional organization of the glucocorticoid receptor (Fig. 1.4). These domains include regions encoding functions for transcriptional activation (N-terminal domain), DNA-binding (zinc-finger domain), and hormone binding (signal transduction domain). The zinc-finger and signal transduction domains of nuclear receptor superfamily proteins share significant sequence similarity.

The DNA-binding domain is most strongly conserved, and forms a zinc-binding "finger" structure in which each zinc atom is coordinated by four conserved cysteines (Freedman et al., 1988; Härd et al., 1990a; Härd et al., 1990b). Fragments including this domain have been expressed in *E. coli* and have been shown to bind *in vitro* to specific DNA sequences called glucocorticoid response elements (GREs) (Freedman et al., 1989; Rusconi and Yamamoto, 1987); these sequences have been shown to confer glucocorticoid

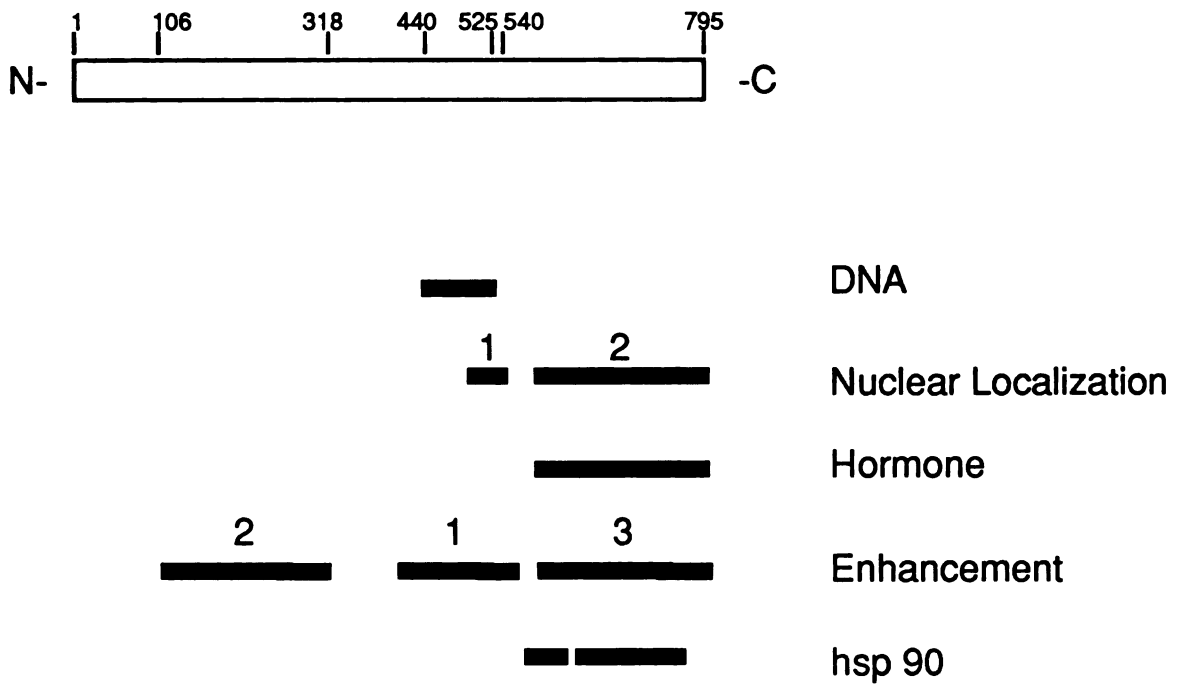


Figure 1.4. Activities within the rat glucocorticoid receptor: functional domains. Domains that have been mapped include: DNA binding (Rusconi and Yamamoto, 1987; Schena et al., 1989), hormone binding (Rusconi and Yamamoto, 1987), nuclear localization (Picard and Yamamoto, 1987), inactivation (Eilers et al., 1989; Godowski et al., 1988a; Hope et al., 1990; Picard et al., 1988; Yamamoto et al., 1988), hsp 90 binding (Howard et al., 1990; Pratt et al., 1988), and enhancement (Godowski et al., 1987; Miesfeld et al., 1987).

responsiveness to nearby promoters (Chandler et al., 1983; DeFranco et al., 1985; Payvar et al., 1983). The carboxy end of this domain (aa 497 - 524) also contains a nuclear localization signal that is reminiscent of the SV40 T-antigen nuclear localization motif (Kalderon et al., 1984; Picard and Yamamoto, 1987).

The signal transduction domain is also conserved (28-57% identity) between GR and other steroid receptors (Evans, 1988). Mutagenesis of parts of this domain reduce the protein's affinity for hormone by ≥1-3 orders of magnitude (Rusconi and Yamamoto, 1987). This and other mutagenesis studies (Danielsen et al., 1987; Garabedian and Yamamoto, 1991; Giguère et al., 1988), as well as hormone cross-linking studies (Carlstedt-Duke et al., 1988; Smith et al., 1988), imply an integrated folded structure bringing separate parts of the protein into close proximity to form the hormone binding pocket. This domain also appears to contain several other functional motifs, including a hormone-dependent nuclear localization signal (Picard and Yamamoto, 1987), a transcriptional enhancement region (Bocquel et al., 1989; Hollenberg and Evans, 1988), a protein inactivation function (Hollenberg et al., 1989; Picard et al., 1988), and the ability to bind to hsp90 (Howard et al., 1990; Picard et al., 1988). Complete deletion of the signal transduction domain in glucocorticoid receptor leaves a protein with constitutive activity; partial deletions are inactive (Godowski et al., 1987; Hollenberg et al., 1989; Rusconi and Yamamoto, 1987). This domain confers hormone regulation when fused to other nuclear proteins, inactivating their function in the absence of hormone, thus acting as a movable and regulable inactivation domain (Hollenberg et al., 1989; Picard et al., 1991, 1990, 1988; Yamamoto et al., 1988). This inactivation function in glucocorticoid receptor may relate to its ability to bind to hsp90 (Picard et al., 1991; Pratt et al., 1988; Yamamoto et al., 1988).

The N-terminal domain is poorly conserved, even between otherwise closely related steroid receptors (e.g. GR and MR). An enhancement function has been mapped to a region in the N-terminal domain, and deletion of the N-terminal domain reduces receptor enhancement activity by 10-20 fold (Miesfeld et al., 1987). The function of this domain may differ between receptor types and appears to depend on cell and DNA contexts. The N-terminal domain is completely absent in some receptors (e.g. vitamin D3 receptor).

Domain swap experiments have shown that the N-terminal, zinc-finger, and signal transduction domains are virtually interchangeable among the various steroid receptors, enabling construction of functional chimeric receptors (Green and Chambon, 1987; Kumar et al., 1987; Picard et al., 1988). Studies of chimeric receptors have been particularly useful in defining the functional activities of the N-terminal and signal transduction domains independent of the steroid receptor's own DNA-binding domain (Godowski et al., 1988a; Hollenberg and Evans, 1988; Webster et al., 1988, 1989).

As the sequences of a large number (> 30) of proteins in this family are now known, it has become possible to deduce some of their evolutionary relationships. Amero et al. (1991) have analyzed the phylogenetic relationships among the sequenced proteins and suggest that they can be grouped into seven subfamilies and five unique proteins. Their results also suggest that the putative ancestral protein contained both a zinc-finger domain and a signal transduction domain, in the same relative positions: in pairwise comparisons among superfamily members, the degree of sequence similarity between zinc finger domains paralleled the level of similarity between signal transduction domains. The evolution of the N-terminal domain was less clear since the sequence of this domain (if present) was not

well conserved between even closely related proteins. These evolutionary relationships may also be reflected in their gene structures.

Evolution of gene networks

As discussed above, an ancestral isoprenoid receptor protein may have arisen as an adaptor, relating an important biological molecule to its own metabolism. This simple chemical relationship may have evolved into a more complex form of regulation in which these molecules lacked any obvious relationship to their own synthesis. These chemical signals may have first represented the metabolic status of the cell. As these molecules became associated with specific environmental conditions (e.g. as a result of the metabolic consequences of these conditions) they became molecular symbols. Since these symbols were not themselves important metabolic intermediates, they were free to acquire signalling roles related more globally to the physiology of the organism.

The diversity of modern isoprene ligand-receptor systems could have arisen by receptor gene duplication and divergence. Occasionally a receptor copy acquired altered specificities for ligands and/or DNA target sites, allowing for a diversity of signal-receptor systems. Such alterations in receptor specificity for ligands and DNA target sequences can be induced in modern steroid receptors by relatively few changes in receptor protein sequence (Danielsen et al., 1989; Garabedian and Yamamoto, 1991; Green et al., 1988; Mader et al., 1989; Umesono and Evans, 1989). As new symbols arose, they acquired new functions appropriate to the needs of the organism. As Tomkins pointed out (Tomkins, 1975), once a chemical symbol interconnects

several important cellular processes, the essential attributes of its response domain would become constrained.

How did a diversity of complex regulatory networks arise and evolve while maintaining their general contexts? Yamamoto suggested that regulatory networks may have evolved primarily by the effects of germ line transpositions and/or gene rearrangements involving long-range regulatory elements (e.g. enhancers) (Yamamoto, 1985). This hypothesis is consistent with the correlation between rates of phenotypic evolution and rates of gene rearrangement (Wilson et al., 1977) and complements Wilson's observation (1977) that the rate of sequence evolution of a given gene is unrelated to its morphological evolution. Sequence changes appear to occur at a relatively constant rate in a given gene, predominantly by substitution events (the "molecular clock"), whereas morphological changes in organisms occur at widely varying rates. For example, frogs, despite a diversity of species, are all morphologically similar. Having arisen more recently, mammalian species are more similar to each other genetically (their sequences are more similar), but include species of widely divergent phenotypes. Therefore Wilson suggested that phenotypic evolution was driven mainly by regulatory changes rather than by sequence alterations.

Receptor regulation

Each steroid appears to regulate the expression of different genes in different target cells. The pattern of regulation by a specific steroid also appears to depend on the action of other hormones and transcription factors. Since steroid hormone action appears to be mediated by a single steroid-specific receptor protein in most cells, it seems at first paradoxical that steroid

hormones mediate such diverse and cell-specific responses. One explanation is that gene expression is controlled by the combinatory action of multiple transcriptional regulatory proteins (Yamamoto, 1985). Thus, the particular repertoire of transcription factors acting on a steroid-responsive gene determines the nature of the response to that steroid. For example, on certain glucocorticoid response elements (GREs), it appears that other cellular transcription factors dictate whether the response to glucocorticoids is positive or negative from that element (Diamond et al., 1990; Jonat et al., 1990; Lucibello et al., 1990; Schüle et al., 1990; Yang-Yen et al., 1990).

Glucocorticoid receptors are nearly ubiquitous and their levels are modulated

The most prominent feature of glucocorticoid receptor expression is that these receptors are present in nearly every cell in the body (Ballard et al., 1974; Kalinyak et al., 1987). This near-ubiquity of receptor expression is consistent with the widespread actions of glucocorticoid hormones. Despite this apparently constitutive pattern of expression, receptor protein and mRNA levels appear to be influenced by a number of extracellular signals and cell states including glucocorticoid hormones (autoregulation) (Cidlowski and Cidlowski, 1981; Dong et al., 1988; Eisen et al., 1988; Kalinyak et al., 1987; Lacroix et al., 1984; McIntyre and Samuels, 1985; Meaney et al., 1985; Okret et al., 1986; Sapolsky and McEwen, 1985; Svec and Rudis, 1981), development (Kalinyak et al., 1989) and aging (Sapolsky et al., 1986), cell type (Kalinyak et al., 1987), extracellular signals (McGinnis and De Vellis, 1981), intracellular cAMP levels (Gruol et al., 1989) and the cell-cycle (Cidlowski and Cidlowski, 1982). Modulation of glucocorticoid receptor expression may in turn affect

responses to glucocorticoids since the level of receptor protein in the cell appears to be limiting for hormone-induced activation (Vanderbilt et al., 1987).

In cases where receptor protein levels have been observed to be particularly high or low *in vivo*, these receptor levels may play a significant physiological role. For example, pro-opiomelanocortin (POMC) is expressed in both the intermediate and anterior lobes of the pituitary, but is processed into different sets of peptide hormones (with different actions) in these two tissues. POMC processing and expression is regulated by glucocorticoids in the anterior, but not the intermediate lobe (Eberwine and Roberts, 1984; Schachter et al., 1982). This difference in glucocorticoid action may be explained by the absence of glucocorticoid receptors in the intermediate lobe (Antakly and Eisen, 1982, 1984; Antakly et al., 1987; Seger et al., 1988). At the other extreme, the hippocampus expresses high levels of receptor protein and mRNA in specific hippocampal neurons (Aronsson et al., 1988; Fuxe et al., 1985). Sapolsky et al. (1986) have suggested that the hippocampus may play a role in downregulating the hypothalamic-pituitary-adrenal (HPA) axis following stress, and that these high receptor levels might be necessary to ensure high sensitivity to glucocorticoids (Sapolsky et al., 1984a). This glucocorticoid sensitivity may also play a role in stress-induced damage to hippocampal neurons *in vivo*. Indeed, aged and/or highly stressed rats appear to be missing hippocampal glucocortiocid receptor-containing neurons, and are impaired in their ability to down-regulate glucocorticoid levels following stress (Sapolsky et al., 1986). These tissue-specific variations in receptor levels and ability to autoregulate receptor might be influenced by transcriptional regulatory signals acting at sites in the glucocorticoid receptor gene.

In chapter three, I investigate the pattern of glucocorticoid receptor mRNA expression by tissue type, glucocorticoid exposure (autoregulation) and postnatal development. These experiments distinguish different classes of receptor mRNA based on differences in their 5' mRNA leaders (and thus, indirectly, their promoters), and suggest the existence of a T-cell specific pattern of receptor mRNA expression and autoregulation.

Mouse T-cell lines have been useful in understanding glucocorticoid receptor function and may reveal clues to receptor regulation

Certain T-cell lines from mouse and human are killed by acute exposure to glucocorticoids. Such cells are said to have a "glucocorticoid-sensitive" phenotype. This cell killing appears to be caused by a specific, programmed response which involves growth arrest, degradation of chromosomal DNA, and ultimately in cell lysis and death (reviews: Bourgeois and Gasson, 1985; Claman, 1972). This programmed response may be similar to the response in target cells to killing by cytotoxic T-lymphocytes (Russell and Dobos, 1980; Russell et al., 1982, 1980), and may reflect certain aspects of T-cell physiology and development *in vivo* (Claman, 1972).

Glucocorticoid induced cell lysis depends on the presence of functional glucocorticoid receptors, and has provided a powerful selection scheme for isolating phenotypic variants defective in various aspects of receptor function. Cells of one such phenotypic class, called nt^i (nuclear transfer increased), contain a truncated form of the glucocorticoid receptor which binds to hormone with wild-type affinity. This receptor also binds to DNA, suggesting that it is defective in activation of transcriptional responses. Previous work (Miesfeld et al., 1988; Northrop et al., 1986) had indicated that

the truncation encompassed the N-terminal domain, and suggested that this receptor truncation was due to an altered mRNA splicing event (Miesfeld et al., 1988). Since alternative splicing could indicate a mechanism of regulating receptor activity, I investigated the nt^i phenotype in detail. These experiments are presented in chapter four, and suggest that an altered splicing pattern in the receptor mRNA could account for the nt^i phenotype. It is not known whether alternate mRNA processing of receptor mRNA occurs *in vivo*.

The signal transduction domain may contain a context-dependent enhancement activity

The N-terminally truncated receptors investigated in chapter four retained some activity, suggesting that the C-terminal portion of the glucocorticoid contained an activation function. An enhancement activity in this domain had been suggested by some studies (Hollenberg and Evans, 1988; Webster et al., 1988), but was not found in others (Godowski et al., 1988a), suggesting that the activation function in this domain might be context-dependent. In chapter five I present an experimental approach and the preliminary results of one experiment that suggest ways to resolve the conflicting observations in the literature.

Summary

In chapter two I present data on the structure of the glucocorticoid receptor gene. I show that the receptor gene is organized into eight coding exons distributed over at least 125 kilobases of DNA, and is transcribed from two or more promoters together expressing a family of alternate 5' exons that

each splice to the first receptor-coding exon. One of these promoters, which I have mapped and characterized, shares features in common with a class of GC-rich, SP-1 responsive TATA-less promoters and appears to be ubiquitously expressed. In chapter three, I investigate the pattern of glucocorticoid receptor mRNA expression by tissue type, glucocorticoid exposure (autoregulation) and postnatal development. These experiments distinguish different classes of receptor mRNA based on differences in their 5' mRNA leaders (and thus, indirectly, their promoters), and suggest the existence of a T-cell specific pattern of receptor mRNA expression and autoregulation. In chapter four I investigate a possible role for the 5' exon family by studying the nt^i class of glucocorticoid-resistant mouse lymphoma cells. I show that these cells use AUGs in their 5' exon family to express receptors containing novel amino-terminal sequences. These experiments suggest that the resistance of these cells to glucocorticoid-induced cell death is caused by the lack of N-terminal domain encoding exon 2 sequences in the nt^i receptor allele; however, in limited assays, I was unable to discern a function for the novel amino-terminal sequences. Finally, in chapter five I describe a preliminary investigation of a potential enhancement activity in the signal transduction domain and suggest ways to study this activity further. Together, these studies show that the glucocorticoid receptor gene is complex and uses multiple strategies for its regulation including the use of alternate 5' noncoding exons, multiple promoters, and upstream open reading frames.

Chapter Two: Glucocorticoid Receptor Gene Structure

Introduction

Most mammalian cells express glucocorticoid receptors (Ballard et al., 1974; Kalinyak et al., 1987); however, receptor protein and mRNA levels appear to be modulated by a number of extracellular signals and cell states including glucocorticoid hormones (autoregulation) (Cidlowski and Cidlowski, 1981; Dong et al., 1988; Lacroix et al., 1984; McIntyre and Samuels, 1985; Meaney et al., 1985; Okret et al., 1986; Svec and Rudis, 1981), development (Kalinyak et al., 1989) and aging (Sapolsky et al., 1986), cell type (Kalinyak et al., 1987), extracellular environment (McGinnis and De Vellis, 1981), cAMP levels (Gruol et al., 1989), and the cell-cycle (Cidlowski and Cidlowski, 1982). Modulation of glucocorticoid receptor expression may be an important mechanism for control of cellular and physiological responses to glucocorticoids since the level of receptor protein in the cell appears to be limiting for hormone-induced activation (Vanderbilt et al., 1987).

Using the cloned glucocorticoid receptor as a probe, I set out to determine the structure of its gene, including its promoter and upstream control regions. I found the glucocorticoid receptor gene to be complex, consisting of eight coding exons, multiple 5' untranslated exons, and alternate promoters. In this chapter I present the genomic cloning and mapping of the rat glucocorticoid receptor gene, including its intron-exon organization, a preliminary characterization of its transcripts, and the mapping of one of its promoters. I discuss the exon structure in the context of the receptor's functional domains and compare the organization of the receptor gene with the known structures of other nuclear receptor

superfamily[2] genes. I also discuss the potential evolutionary relationships among these proteins.

Results

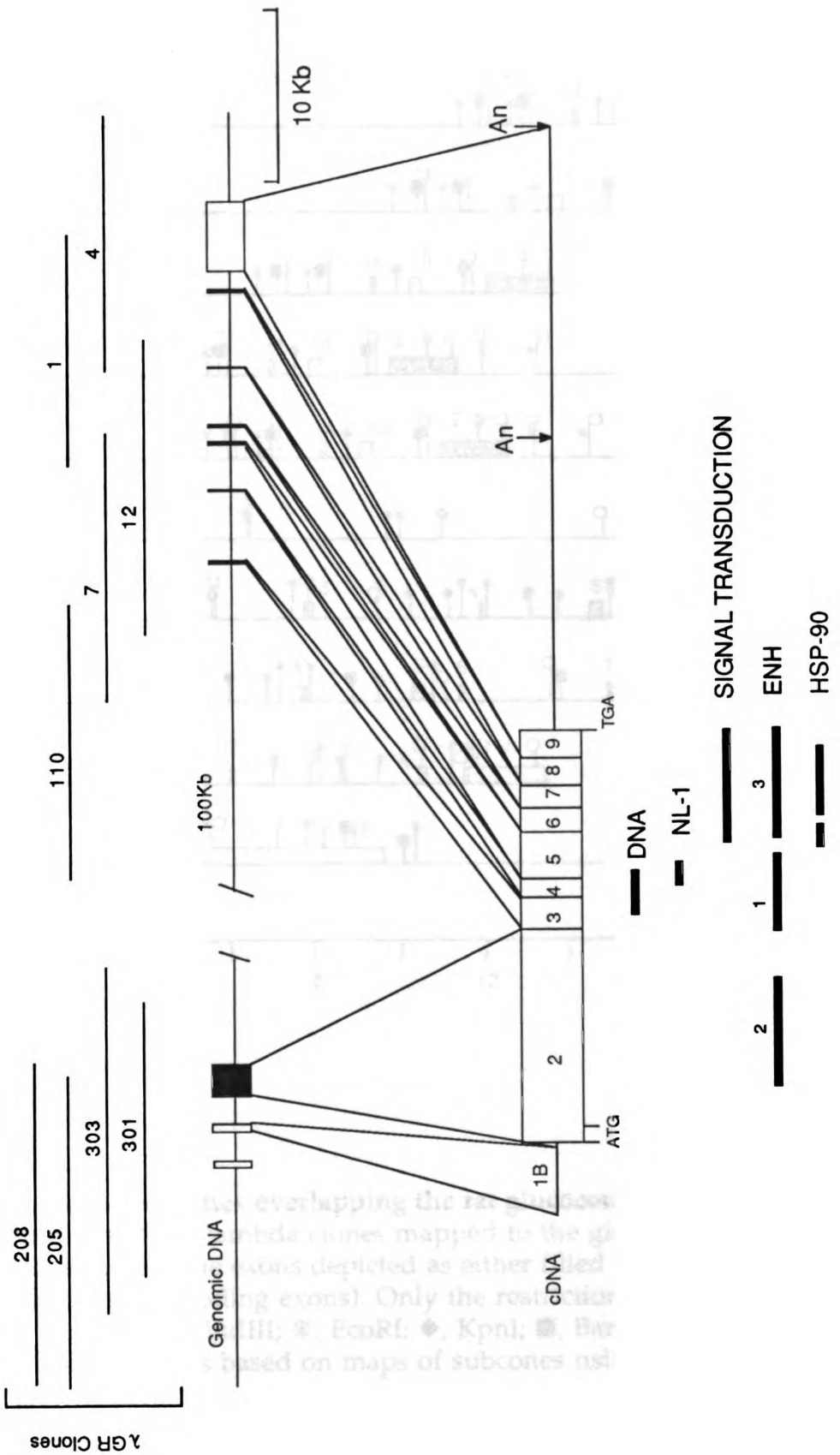
The rat glucocorticoid receptor is encoded by a large gene split into eight coding exons

The genomic organization of the rat glucocorticoid receptor gene is shown in Fig. 2.1. The 3' genomic clones λGR7, 12, 1 and 4 were previously isolated and exons 3 and 9 mapped (Miesfeld et al., 1985). I obtained additional genomic clones (λGR203, 205, 208, 301, 303) by screening a Fisher rat liver genomic library prepared in λEMBL-3B, with nick-translated probes from rat GR cDNA (pRM16, pER41) (Miesfeld et al., 1984, 1986). I isolated an additional λ clone (λGR110) extending into a large intron by screening the library with a probe from the 5' end of λGR7. These overlapping genomic clones were mapped and the exons located by restriction digestion and Southern hybridization to cDNA probes (Fig. 2.2). The precise positions of the intron/exon borders, shown in Fig. 2.3, were determined by sequencing exon-containing genomic subclones, and by comparing these sequences with the published cDNA sequence (Miesfeld et al., 1986).

As shown in Fig. 2.1, the rat GR gene is over 125kb long and is composed of eight coding exons. The amino-terminal half of the receptor is encoded by one large exon (exon 2), while the remainder of the receptor

[2] Nuclear receptor superfamily refers to all homologous zinc-finger proteins and includes receptors for retinoic acid, thyroid hormone, and orphan receptors as well as steroid receptors.

Figure 2.1. Organization of the rat glucocorticoid receptor gene. Lambda clones are shown above, and the cDNA below the map of genomic DNA. Exons are depicted as boxes in the genomic DNA, with the coding region filled. mRNA processing to remove introns is illustrated by lines joining intron/exon boundaries in the genomic map to the splice junctions in the cDNA. The long receptor open reading frame is depicted by a box in the cDNA. Heavy lines under the translated region of the cDNA show the functional regions of receptor that have been mapped: DNA, DNA-binding domain; NL-1, nuclear localization-1; Signal, signal transduction domain; ENH (1, 2, 3), transcriptional enhancement domains; HSP 90, hsp 90 binding regions.


λ GR Clones
208
205
303
301
110
7
12
1
4
Genomic DNA
100Kb
10 Kb
An
An
cDNA
1B
2
3 4 5 6 7 8 9
ATG
TGA
DNA
NL-1
SIGNAL TRANSDUCTION
ENH
HSP-90
1
2
3

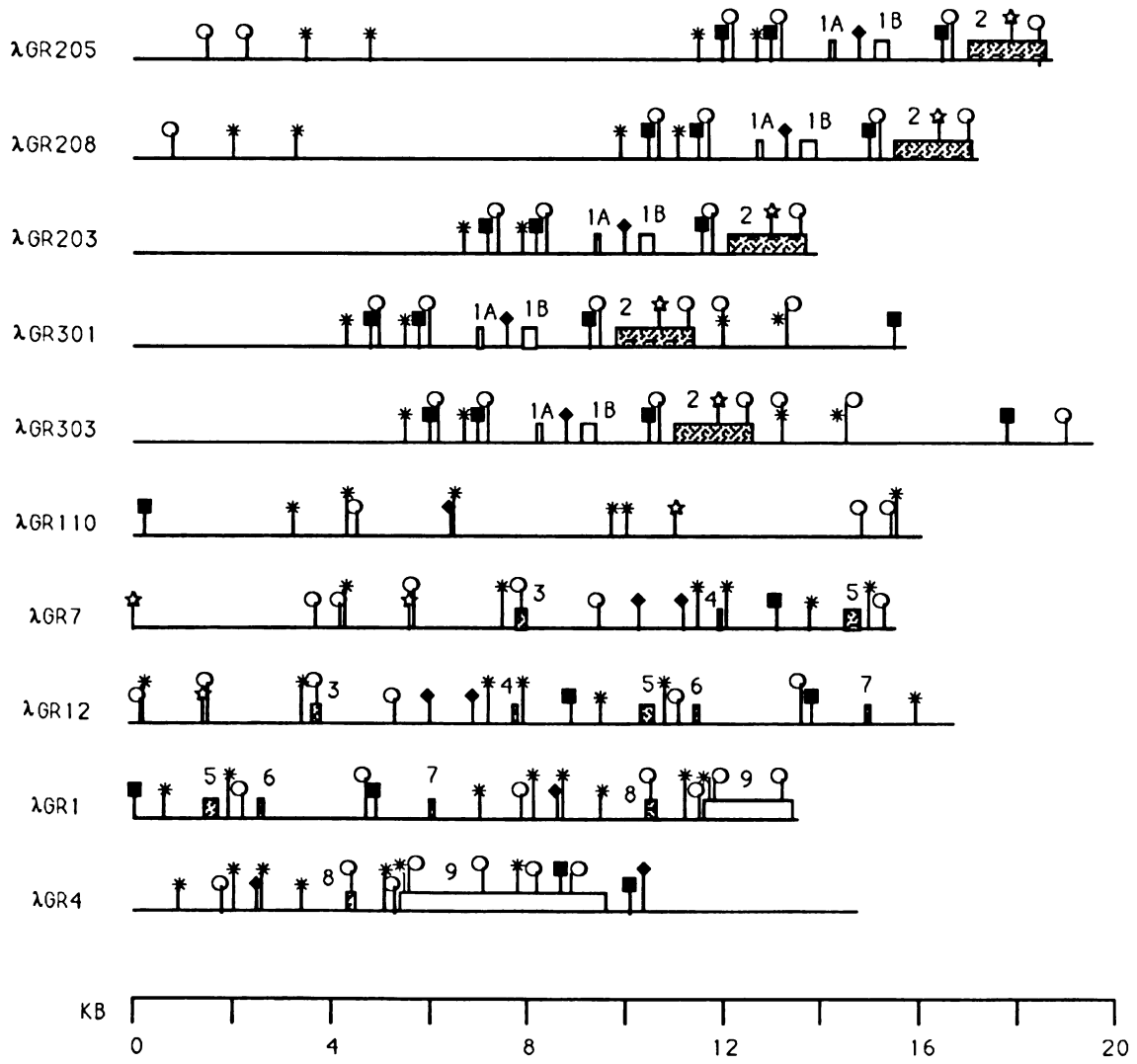


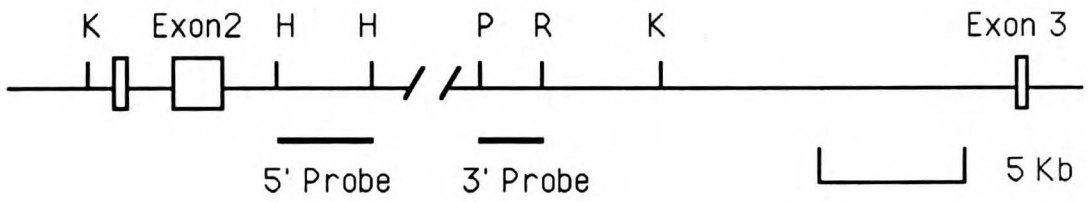
Figure 2.2. Genomic clones overlapping the rat glucocorticoid receptor gene. Restriction maps of the lambda clones mapped to the glucocorticoid receptor gene are shown, with the exons depicted as either filled boxes (coding exons) or unfilled boxes (noncoding exons). Only the restriction sites mapped in all clones are shown: ○, HindIII; ✳, EcoRI; ◆, KpnI; ■, BamHI; ☆, SalI. Detailed restriction mapping was based on maps of subcones using additional enzymes to those shown here.

INTRON BORDERS

	DONOR	ACCEPTOR	
EXON	----------INTRON----------		EXON
EXON 1A----CTCACA	GTATGTATGCGCTGACCCTCTC....GTCTTTTTTTTTTCTTTTGTAG		TTAATATT
	2.7kb		
EXON 1B----GCTGAG	GTGAGCGGGGGCTGGGCCGAGC....GTCTTTTTTTTTTCTTTTGTAG		TTAATATT
	1.8kb		
EXON 2-----TCA AG	GTAAGTCAGCGCTTTTCTGTTT.....CCCCACAG		C CCT
Ser Ser	100kb		Ser Pro
414 415			415 416
EXON 3-----GAA G	GTAGTGT.....ATTGCCCTTTACTCTCTTGTAG		GA CAG
Glu Gly	4.0kb		Gly Gln
469 470			470 471
EXON 4-----GAA G	GTAATGGAACCTTAAAGGAGCT.....TTCACAG		CT CGA
Glu Ala	2.5kb		Ala Arg
508 509			509 510
EXON 5-----CTA G	GTAGGACATGAAGGACAACAGA....CATACCTTTATTTCCAATCTAG		GC TTC
Leu Gly	1.0kb		Gly Phe
600 601			601 602
EXON 6-----AAT GA	GTAAGTTACATGTGAATCGGGT....TATCTGTATTTGTGTCTCCCAG		G CAG
Asn Gl	3.3kb		u Gln
648 649			649 650
EXON 7-----TCA G	GTTGGTGGATCACCTACTCATT.....TCATATTACCTCTTAG		TT CCT
Ser Val	4.4kb		Val Pro
692 693			693 694
EXON 8-----CAT GAG	GTAAGTACCAGACATAAAGCCA....TCTGCCTCGTCTTCTCTTGCAG		GTG GTT
His Glu	1.1kb		Val Val
744 745			746 747

Figure 2.3. Intron-exon borders. Precise intron/exon borders were determined by sequencing the genomic DNA and comparing the genomic sequence to the published cDNA sequence (Miesfeld et al., 1986). Amino acid numbers correspond to those in the published sequence. Consensus splice acceptor and donor dinucleotides in the intron are underlined.

Figure 2.4. Pulsed-field mapping of intron II. Agar blocks containing DNA from J2.17 cells (10^6 cells per block) were digested overnight with the restriction enzymes KpnI, BssHII, SfiI and SnaBI. Each block was then cut in half and all the half-blocks loaded, along with size markers, on a single 0.8% agarose gel. This gel was then run under a clamped homogeneous electic field (C.H.E.F.) in 0.5xTBE at 16°C at 180 V / 40 sec switch, for 37 h. After staining with EtBr to check for complete digestion, the DNA was transferred to a Zetaprobe® membrane by alkaline transfer (BioRad). The membrane was then cut in two and each half hybridized to either the 5' or the 3' probe.


K
Exon2
H
H
P
R
K
Exon 3
5' Probe
3' Probe
5 Kb

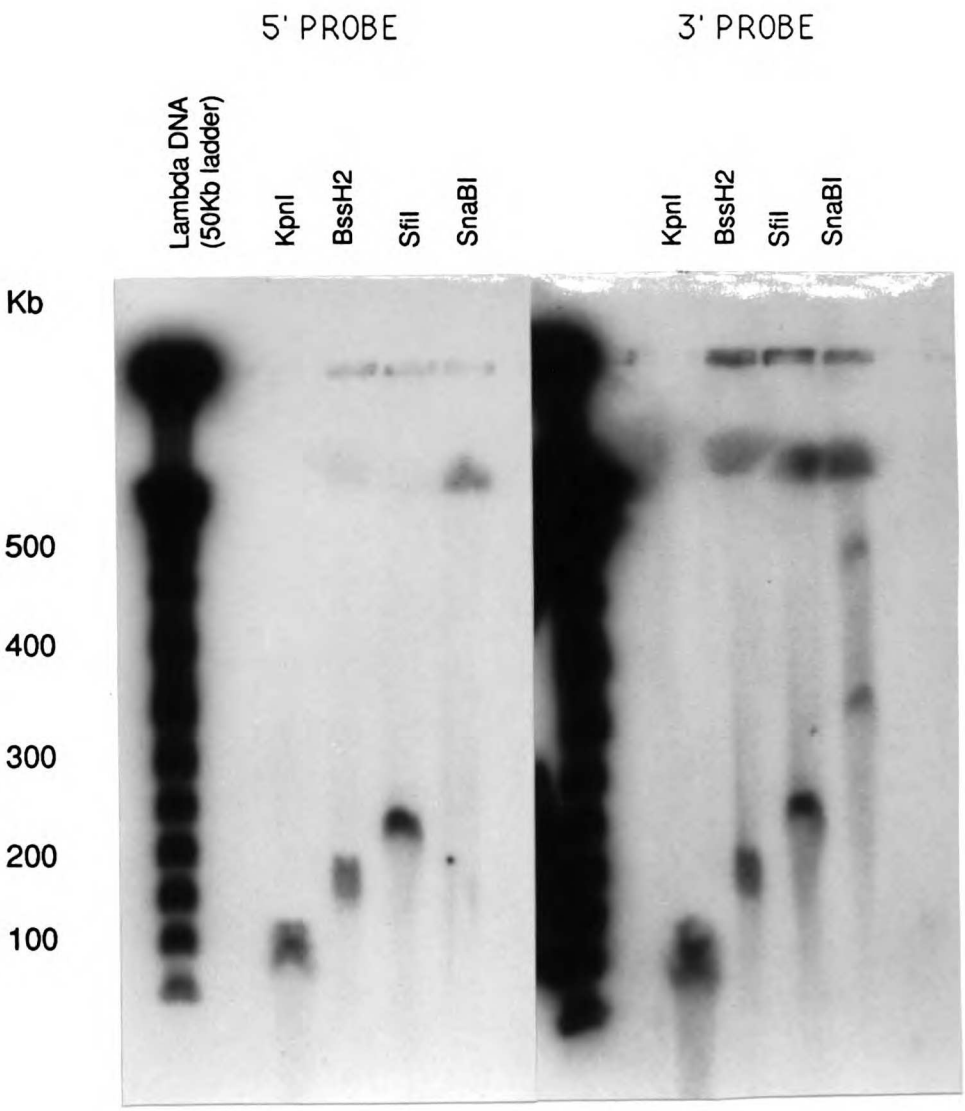

5' PROBE
3' PROBE
Lambda DNA (50Kb ladder)
KpnI
BssH2
SfiI
SnaBI
KpnI
BssH2
SfiI
SnaBI
Kb
500
400
300
200
100

Figure 2.4

coding region is split into seven exons (3-9), with the last exon containing a large 3' UTR.

The clone λGR110 failed to span the intron between exons 2 and 3: though the 3' end of this clone overlapped λGR7, its 5' end did not overlap any of the 5' genomic clones. To estimate this intron's size, I separated restriction digests of rat genomic DNA on pulsed-field gels and probed Southern blots of these gels with labelled genomic fragments flanking the intron. Any band that hybridized to both the 5' and 3' flanking probes would contain both probe sequences as well as all intervening DNA; the size of such a band would provide an upper limit of the intron's size. As shown in Fig. 2.4, the smallest band that hybridized to both the 5' and 3' probes was about 100 kb (*KpnI* digest). Since the *KpnI* sites flanking both probes have been mapped relative to exons 2 and 3 (Fig. 2.2), it was possible to determine that the intron's size (within the resolution of the CHEF gel in Fig. 2.4) is about 100kb.

The exons correspond to functional domains of the rat glucocorticoid receptor

The functions that have been mapped to discrete domains of the glucocorticoid receptor appear to correspond roughly to separate exons. The amino-terminal half of the receptor, which includes activation domain *enh2*, is encoded by one large exon (exon 2), similar to the amino terminal domains of ER (Ponglikitmongkol et al., 1988), AR (Kuiper et al., 1989; Lubahn et al., 1989), PR (Huckaby et al., 1987; Jeltsch et al., 1990), and RARγ (Lehmann et al., 1991). The central domain is composed of two zinc-finger motifs which form an integral globular structure (Härd et al., 1990b) and bind to DNA. Each

finger has been shown by spectrometric analyses to bind a zinc atom (Freedman et al., 1988) and the sequence of this domain is conserved among members of the nuclear receptor superfamily. Each of the two fingers is encoded by a separate exon (3 and 4), as is the case with all fifteen of the genes in the nuclear receptor superfamily analyzed so far[3]. The two fingers are functionally distinct (Green et al., 1988; Härd et al., 1990b; Schena et al., 1989; Schwabe et al., 1990), suggesting that each finger is a separate functional domain (Green and Chambon, 1988).

The fifth exon includes sequences from three functional domains: the 3' end of the DNA-binding domain, which includes the portion of a nuclear-localization signal motif that is reminiscent of the SV40 nuclear localization signal (Kalderon et al., 1984; Picard and Yamamoto, 1987); a proline-rich region poorly conserved among the steroid receptor and TR/RAR sequences (possibly functioning as a "hinge" or spacer separating the zinc finger and signal transduction domains); and the 5' end of the signal transduction domain. Though these three functional regions are encoded by one exon in GR, these regions are split into two separate exons in the genes for thyroid hormone and retinoic acid receptors, NGFIB/N10 and E75B, and into three exons in the hVDR gene, suggesting that GR exon 5 may be the product of intron loss.

The signal transduction domain is encoded by a region spanning the five C-terminal exons (5-9). Several functions reside within this conserved domain. A dimerization function may be located at the 3' end of exon 8, which is similar to a dimerization domain demonstrated in ER (Fawell et al., 1990; Lees et al., 1990). Transcriptional enhancement activity has been

---

[3] rGR, hMR, hER, hAR, ckPR, hVDR, hTRβ, r-rev ERBAα, r-ERBAα, ckERBA, RAR, rN10, mouse NGFIB and Drosophila proteins E75A&B and *egon*; see references in legend to Fig. 2.5.

demonstrated in the carboxy terminal region of ER (Webster et al., 1989) and GR (Hollenberg and Evans, 1988; Webster et al., 1988; see also chapter five). The carboxy terminus also seems to be important for GR interaction with other proteins. Amino acids encoded in exons 5-6 (aa568-616) and 6-9 (aa632-766) appear to be required for GR association with hsp90 (Howard et al., 1990); Holley, pers. comm.). The signal tranduction domain is also important for repression on a composite GRE in the proliferin promoter *in vivo* (M. Diamond, unpubl. data; (Schüle et al., 1990; Yang-Yen et al., 1990), suggesting an interaction with AP-1. More detailed functional mapping as well as determination of the structure of the signal transduction domain will be necessary to determine whether the C-terminal exons correspond to discrete structural and/or functional units.

Similarities in the organization of nuclear receptor superfamily genes are consistent with evolutionary relationships inferred by phylogenetic analyses

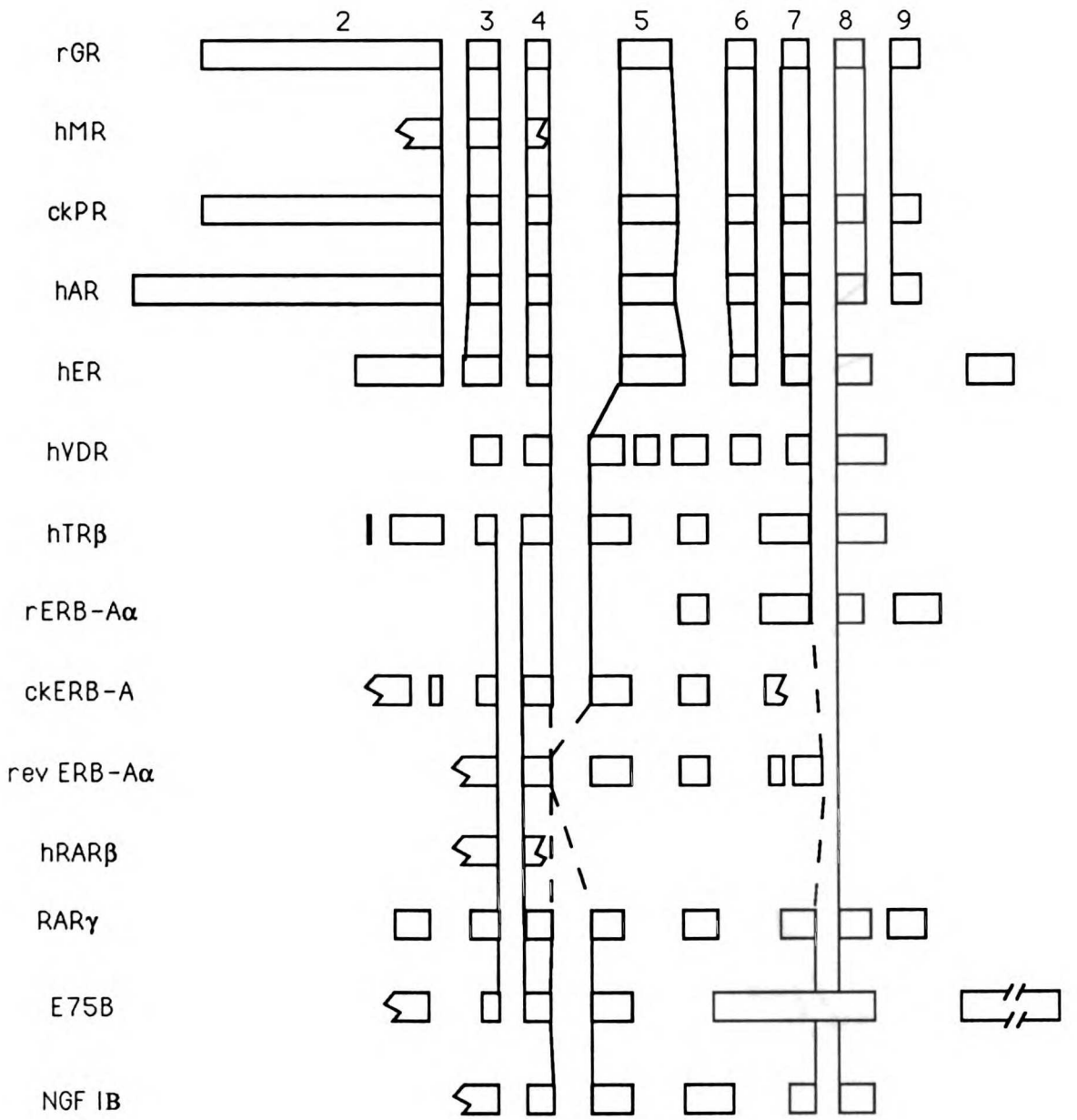
According to the exon shuffling hypothesis of eukaryotic genome evolution, exons should represent discrete protein functional domains. In a family of genes, one might expect similarities in gene structures among family members to parallel the protein sequence similarities, since both are reflections of each gene's evolutionary history. I compared the known gene structures of steroid receptors by aligning the protein sequences by a multiple sequence alignment algorithm (pima.sh; Smith and Smith, 1990) and compared the alignments of the exon/intron boundaries (Fig. 2.5). I noted both the position and phase (which nucleotide position in the amino acid triplet codon) of these juctions. The presence of a common splice juction

between two or more genes (same splice position and phase; indicated by a vertical line joining homologous exon borders in Fig. 2.5) may suggest that these genes share a common ancestor, and that the bordering exons are related. These relationships should parallel the evolutionary patterns inferred from phylogenetic analysis of protein sequence differences.

Amero and her colleagues (Amero et al., 1991) have analyzed the phylogenetic relationships of nuclear receptor superfamily proteins based on comparisons of zinc-finger sequences. Their results, shown in Fig. 2.6, suggest a grouping of these proteins into seven main branches (subfamilies: R, X, D, T, K, E, G) and five unique proteins (nur/77, hTR25, DHR3, E75A, EcR). In this tree, the steroid receptors form distinct subfamilies, as do the retinoic acid and thyroid hormone receptors. In addition, pairwise comparisons between receptor sequences showed that the degree of similarity between signal transduction domain sequences paralleled the similarities between zinc-finger sequences, suggesting that both domains have been retained in their present relationship since the divergence of a common ancestral protein.

The organization of receptor genes shown in Fig. 2.5 is consistent with these phylogenetic relationships. The most strongly conserved introns among all the genes are those between exons 4 and 5 and between 7 and 8 (the exception, rev-erbAα, is unusual since its transcription unit overlaps that of r-erbAα). Other introns appear to be conserved within subfamilies. The phase of the exon 2 - 3 junction is the same among the steroid receptors genes (GR, MR, PR, AR, ER), even though the sequences immediately flanking this junction and in exon 2 are highly divergent among genes in the steroid receptor subfamilies (<15% identity between any pair of genes). The position and phase of the intron between exons 3 and 4 is also conserved within the steroid receptor subfamilies (G, E) and in the TR and RAR subfamilies (T, R),

Figure 2.5. Comparison of genomic structures. The genomic organization of rat glucocorticoid receptor is compared with the organization of other nuclear receptor superfamily genes. Predicted protein sequences were translated from nucleotide sequences in GENBANK and aligned using the pattern-induced multi-(sequence) alignment program (pima.sh; Smith and Smith, 1990). Exons are shown as boxes (the numbers refer to GR exons). Introns that are conserved in both position and phase are connected with vertical lines (dashed lines connecting rev-erbAa indicate that the phase of the indicated introns is not shared). Gene structure references: hAR (Brinkmann et al., 1989; Kuiper et al., 1989; Lubahn et al., 1989), hER (Ponglikitmongkol et al., 1988), ck c-erbA (Zahraoui and Cuny, 1987), E75B (Segraves and Hogness, 1990), rev-erbAα (Lazar et al., 1989), hMR (Arriza et al., 1987), NGF1B (Watson and Milbrandt, 1989), N10 (Ryseck et al., 1989), ckPR (Huckaby et al., 1987), hRAR (Petkovich et al., 1987), hTRβ (Sakurai et al., 1990), hRARγ (Lehmann et al., 1991), hVDR (Hughes et al., 1988; Pike, 1990; Ritchie et al., 1989).


2
3
4
5
6
7
8
9
rGR
hMR
ckPR
hAR
hER
hVDR
hTRβ
rERB-Aα
ckERB-A
rev ERB-Aα
hRARβ
RARγ
E75B
NGF IB

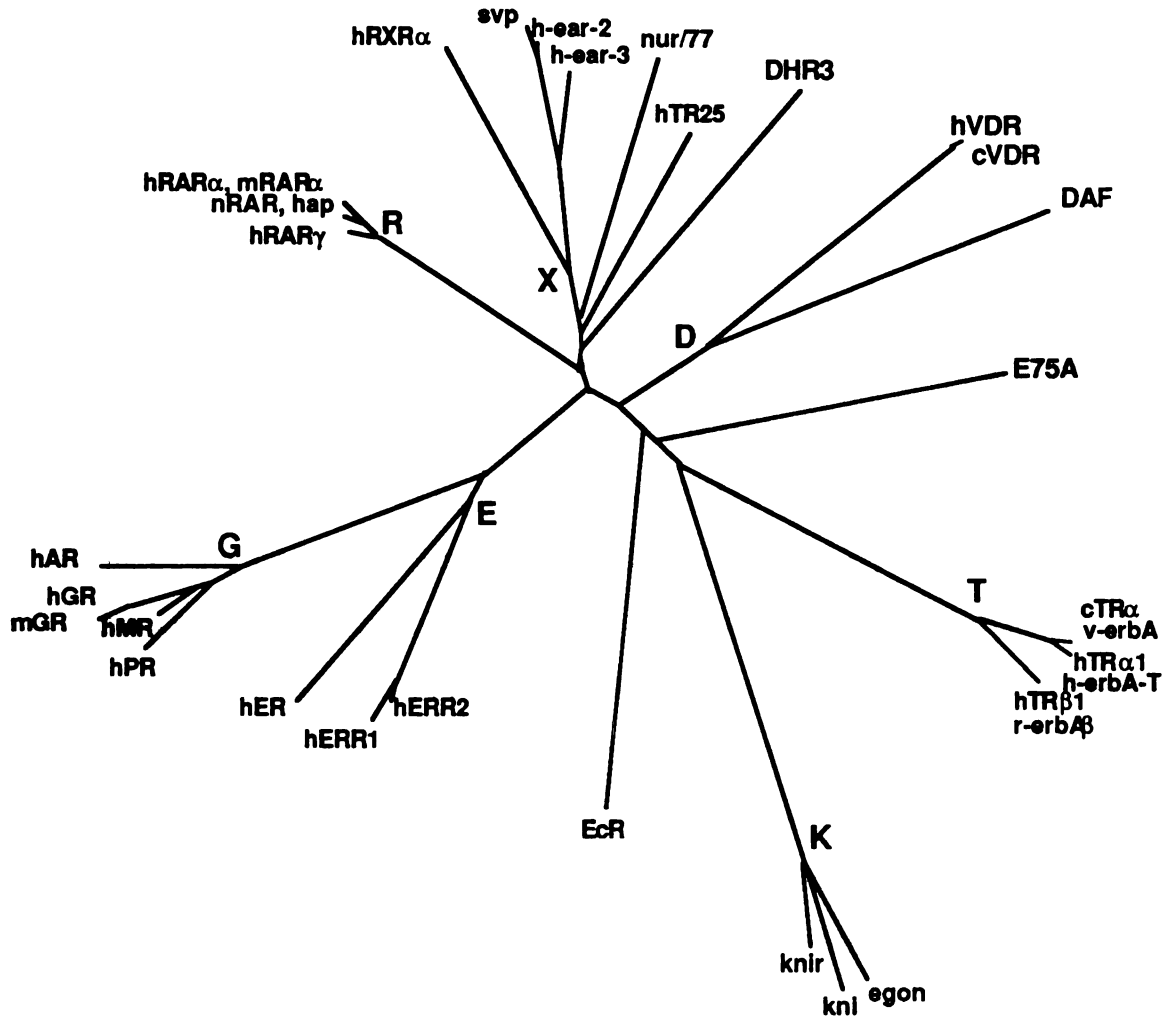


Figure 2.6. Phylogenetic analysis of nuclear receptor family zinc-finger sequences. From Amero et al. (1991).

but both differ from the junction in VDR, NGFIB, E75B and *egon*. The introns separating exons 5 - 9 are all conserved within the steroid receptor subfamilies, with the exception of the last intron in ER, which is in a different phase. These patterns are consistent with the branch clusters in the phylogenetic tree of Amero et al. (1991).

There are multiple 5'-untranslated sequences, all of which splice to the first coding exon

Comparison of the 5' genomic sequence flanking exon 2 with the 5' sequence of the cDNA clone pER41 indicated that there was likely to be one or more additional 5' exons. In order to obtain 5' cDNA sequence to facilitate mapping of the 5' exon(s), I screened a λGT10 cDNA library of J2.17 (an HTC rat hepatoma cell line) mRNA (prepared by Akira Inoue, unpublished), in two successive rounds. I first screened 450,000 plaques with a nick-translated probe from the 5' end of pER41, obtaining 14 positives. I then rescreened these plaques with an oligonucleotide probe to the 5' end of exon 2 (SR01, see appendix two). Seven clones (shown in Fig. 2.7) were isolated that rescreened positive, and these were subcloned into the plasmid pBluescript (KS+), restriction mapped, and their 5' ends sequenced. Of these clones, three (pAR5, 6, 13) contained 5' sequences identical with the 5' end of pER41 plus an additional 40bp at the 5' end (these clones are designated class "A"). The other four clones contained distinct 5' sequences, which I denote classes B (pAR4, 8, 10) and C (pAR1). All seven cDNA clones contain identical exon 2 sequences, suggesting the existence of at least three classes of transcript distinguished only in their 5' exons (exon 1). Interestingly, the clone pAR1 (class C)

contained six nucleotides at the 3' end of exon 1 which were identical to the six 3' nucleotides in the class B clones (see Fig. 2.9).

The 5' exons of transcripts A and B (exons 1A and 1B) were mapped 2.6 and 1.6 kb, respectively, upstream of exon 2 (see Figs. 2.2 and 2.8). As shown in Fig. 2.8, these transcripts could be produced either by initiation from a common upstream promoter and alternate splicing of the intervening exon 1s, or by initiation from multiple promoters. As indicated in Figs. 2.8 and 2.9, some or all of the clones in each class contained additional 5' sequences that were not contiguous with the genomic sequences flanking exons 1A and 1B. The 5' end of clone pAR1 (class C) did not hybridize to λGR205, but did hybridize to a discrete band on a genomic southern blot (data not shown), suggesting that it originated >16 kb upstream of exon 2.

The discovery of three transcript classes differing in their 5' exons suggested that there might be additional transcripts, also differing at their 5' ends. In order to determine whether there were additional 5' exons, I used a PCR-based RACE (rapid amplification of cDNA ends) method (Frohman et al., 1988; Frohman and Martin, 1989) to generate two plasmid libraries enriched in GR 5' cDNA. These libraries were screened by colony hybridization and digestion with SspI (which cleaves at the 5' end of exon 2). The results of this screen are summarized in Fig. 2.10. Out of 470 colonies screened, 296 hybridized to an exon 2 probe. Ninety-six of these were screened for cleavage by SspI (indicating the presence of 5' exon 2 sequences), yielding 47 positive clones, which I sequenced. Of these 47, 31 had DNA that extended more than 2 bp 5' of exon 2[4]. These 5' extending clones fell into seven

[4] Of the fifteen non-5'extending clones, thirteen had just two bases, all AG, 5' of the splice junction. Of the sequences discovered, only B, C, and F have an AG at the exon splice donor

Figure 2.7. rGR cDNA clones. λGT11 glucocorticoid receptor cDNA clones were screened by the method of Wood, et al. (1985) with an oligonucleotide specific to the 5' end of exon 2 (SR01), subcloned into pBluescript at the EcoRI site, and mapped by restriction digestion. The receptor clones are shown aligned with the restriction map of the cDNA published in Miesfeld et al. (1986). The orientation of each insert relative to the vector sequences is denoted by an asterix, indicating the T7, SacI side of the polylinker in pBluescript. Restriction sites: S, SmaI; M, MstII; N, NcoI; Sal, SalI; H, HindIII; A, AvaI; R, EcoRI; X, XbaI.

---

position. Given the frequency distribution of the clones sequenced, most of these truncated clones are probably class B. Thus the frequency distribution of transcript classes based on the frequencies of PCR clones shown in Fig. 2.10 is likely to represent an underestimate of the abundance of class B transcripts.

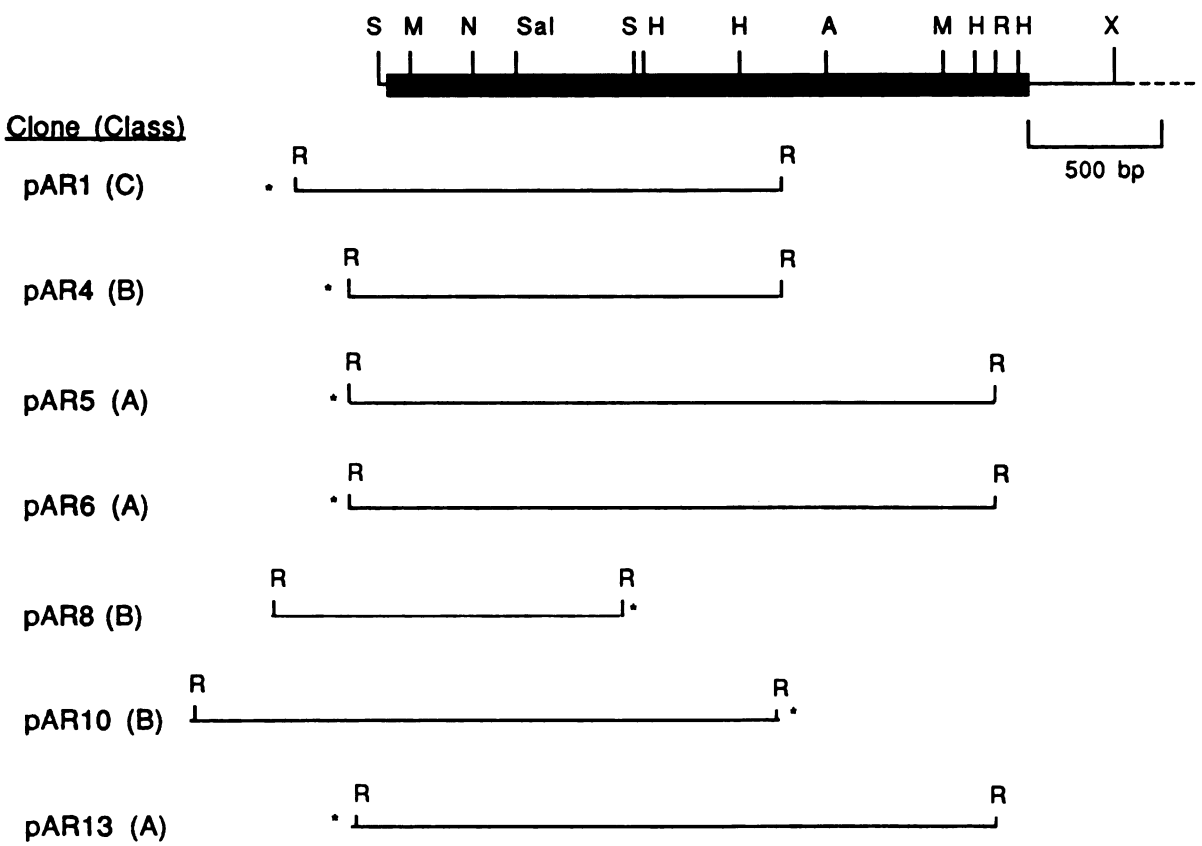

S M N Sal S H H A M H R H X
500 bp
Clone (Class)
pAR1 (C)
pAR4 (B)
pAR5 (A)
pAR6 (A)
pAR8 (B)
pAR10 (B)
pAR13 (A)
R

Figure 2.8. Organization of GR transcripts. The inferred splicing patterns of receptor transcripts are shown aligned with the genomic map. Putative exons shown upstream of exon 1A have not been mapped, and are all different from one another, including the upstream exons in transcripts D-H. Transcripts Bβ, Bγ, A, Aβ, and D-H are based on the PCR clones, and are shown truncated at their 3' ends because they were cloned at their MstII site in exon2. Transcripts Bα, C and Aα are based on the pAR clones (see Fig. 2.6). Transcript B is based on both PCR clones and pAR clones, but cannot be distinguished from either Bα or Bβ because none of the clones extends as far as the putative splice junctions in Bα or Bβ.

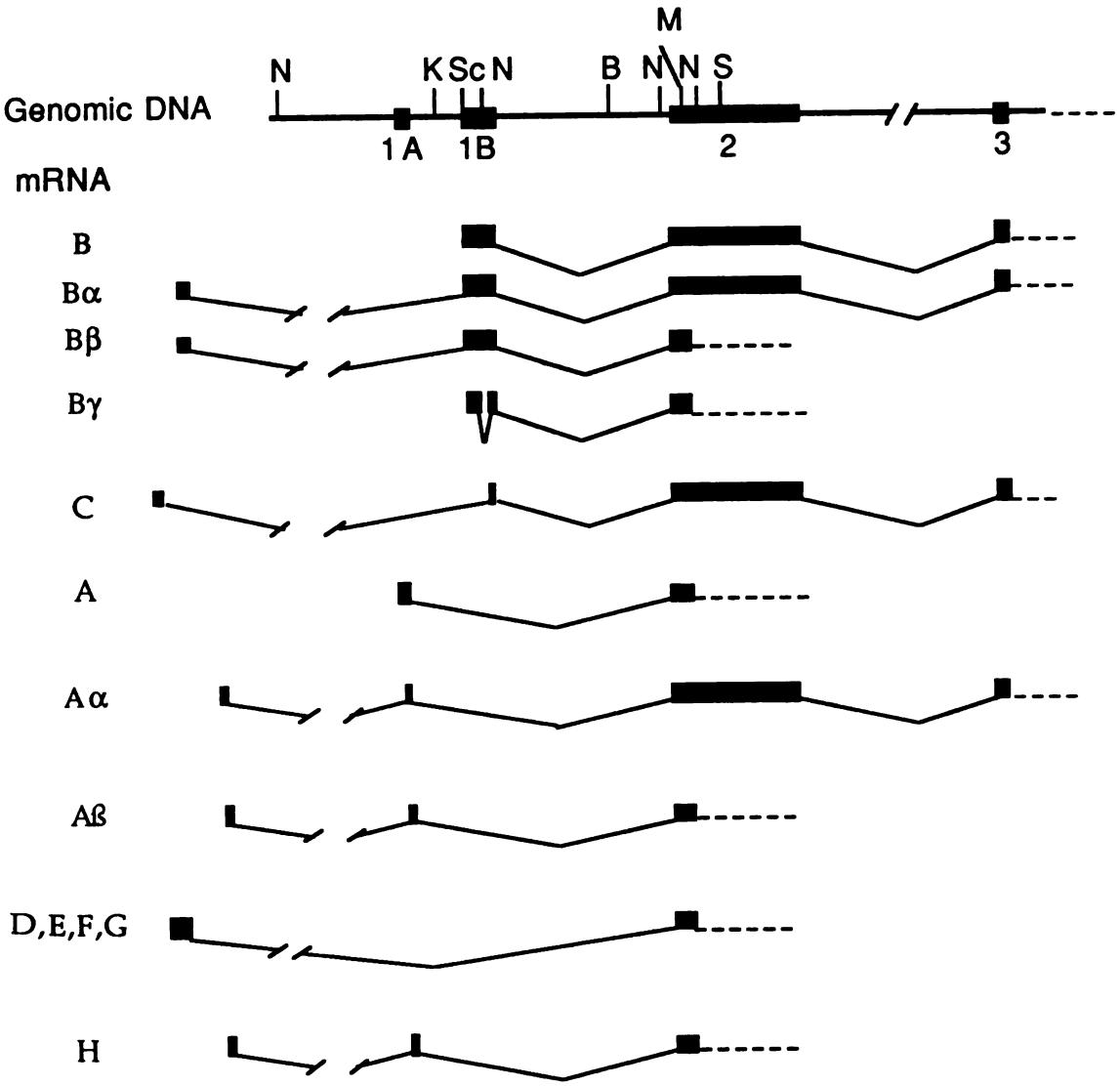

M
N
K Sc N
B N N S
Genomic DNA
1A 1B
2
3
mRNA
B
Bα
Bβ
Bγ
C
A
Aα
Aß
D,E,F,G
H

Figure 2.9. Sequences of GR 5' mRNA leaders. The transcript sequences are based on sequences of cDNA clones. Lowercase letters refer to sequence based on only one cDNA clone; sequences based on multiple clones are represented with uppercase letters. Genomic sequences that have been identified are underlined.

Figure 2.9

GR A

```
        |   10      |   20      |   30      |   40      |   50      |   60
  1 GCTTTGCAAC TTCTCCCGGT TGCGAGCGAG CGCGCGCGCG GCGGCGGCGG CGGCGGCTGC  60
 61 AGACGGGGCC GCCCAGACGC TGCGGGGGTG GGGGACCTGG CGGCACGCGA GTCCCCCCCC 120
121 GGGCTCACA                                                        129
```

GR Aα

```
  1 CCGGGGATCC GTCGACCTGC AGGGGGGGGG GGGGGGGGGA CGCTGCGGGG GTGGGGGACC 60
 61 TGGCGGCACG CGAGTCCCCC CCCGGGCTCA CA                               92
```

GR Aβ

```
  1 ctagagGAGT CCCCCCCCGG GCTCACA                                     27
```

GR B

```
  1 ATTTTGCGAG CTCGAGTCAG TGCCTGGAGC CCGAGTCGCC GCCCGCCGTC GGGGACGGAT  60
 61 TCTAAGTGGG TGGAACAAGA CGCCGCAGCC GGGCGGCGCG GCGCCGGGAC GGGGGACGCG 120
121 CGCGGGAGAC GGGAGCGGCG CGGGGGCCGG CTTGTCAGCC GGGAACGGGT GACTTTCAGC 180
181 GCTAGGGGCT CTCCCCTCCC CCATGGAGAA GAGGGGGCGA CTGTTGACTT CCTTCTCCGT 240
241 GACACGCGCG CCTCCCGCGT CCGCACGCCG ACTTGTTTAT CTGGCTGCGG TGGGAGCCGC 300
301 GAGCGGGCGA GCGCGCGGGT GCTGAG                                       326
```

GR Bα

```
  1 cgcacaaagc ctgtccacta acccagctga ccccacccgg acatttgcct ggctccctgg  60
 61 acggccaggg gtccgggtac tgagtctggg tcctgctcca gaacccccta aacctgccac 120
121 atccaagatc atccttgtca acaagctggg gcaggtattt gtgaagatgg caggagaggg 180
181 tgagcctgtc gtaccccag ttaaagcagc catctctgcc ccccatcatt cccccaacag 240
241 ccccaacctc ctggtactct gcctccagga cccctgctca gtgtgttacc agtggtaggg 300
301 taGAGTCAGT GCCTGGAGCC ....exon 1B...
```

GR Bβ

```
  1 cagATTTTGC ....exon 1B....
```

GR Bγ

```
  1 AAGACGCCGC AGCCGGGCGG CGCG------[deletion]------GCGCCT CCCGCGTCCG  40
 41 CACGCCGACT ....exon 1B....
```

GR C

```
  1 gttannaaga nagatgcaaa aaaccggtaa gagattcctc agctcatcca caaaggactc  60
 61 acgtcatccg tcacatccct gtaccatgga tagacaaagt tttccaacac taattcaagg 120
121 actctgctct tgattctgat gtgatgcttt aagttcctgc cctgacttcc ttgccaataa 180
181 tggactctga gagagatgca tcaaccttgg aagaaacttt caggtctagc cagggctgag 240
```

GR D

```
  1 ggggtaagag gaggcggact acagcagcaa cttactgtcg gtctgcagct tgcctctagg  60
 61 cctgcacaca ccccctcccg ccccgcaagc ttccttaatc acaatttttt ttttaagtgc 120
121 aaagaaaccc ggctctctga gagggttttG CATTCGGCAT GCAACTTCCT CCGAGT     176
```

Figure 2.9 (cont.)

```
GR E
        |   10     |   20     |   30     |   40     |   50     |   60
  1 ggcagcaaac gtcaagattc gggggagggg cctccgcggg gagcttggat gctggcccga  60
 61 anggngtGGA AGGaAGAGGT CAGGAGTTTG G                                 91

GR F

  1 GAgcgGcCGT CTgGRcCgGC GGCTCAGgGC CCGCGGCTCC GGGCTGCGGG CTTGTAGGGT  60
 61 GGATTGGgcT CACAGCCTGC AGCCCAGACT TCGCCCGCCC GGCCTTATCY GCTAGAAGTG 120
121 GgGTGCCGCA GAGaACTCAA CAG

GR G

  1 gggactccac gacgcaccag aaactgacca acgattgtct ccagaaaagt ggcccggcag  60
 61 tctggaacgg tctgtgcgaa cgcgtttgca ggattctagc catgaaatta g          111

GR H

  1 gacgaaccac tgaattccgc attgcagcga tattgtattt ttaaaggtcg ctagctcgat  60
 61 acaataaacg ccacttgacc attcaccaca ttggtgtgca cctccgaagc ttgtcgaccg 120
121 gtacccggga attcgagccg tagaggatcc cgtgcttaca                       160
```

Figure 2.10. Summary of GR 5'RACE library screening. cDNAs made from two independent oligo-dT selected J2.17 RNA preparations were amplified by two rounds of polymerase chain reactions (PCR) by the R.A.C.E. protocol (Frohman and Martin, 1989), cloned into a modified pBluescript plasmid, and transformed into DH5α competent cells. I first screened these two RACE cDNA libraries (J3 and J6) for the presence of exon 2 sequences by colony hybridization with an exon 2 specific probe (5'RdN93 Xba-Sal, random primed). I then screened mini-prep plasmids from 96 positive colonies (48 from each library) for the presence of the SspI restriction site in the insert. Since this site is located at the 5' end of exon 2, plasmids lacking the site would not have upstream (exon 1) sequences. I sequenced all 47 plasmids containing SspI-digestible inserts. I assigned the 5' sequences into classes based on the sequences immediately upstream of exon 2 (A-H, independent sequences; TR, truncated (two or fewer upstream nucleotides); ?, unable to determine). Sequence differences further upstream are assigned to subclasses (e.g., α, β, γ; see Fig. 2.8). Though classes C and H might be considered subclasses of B and A, I assigned them to their own classes since the overlaps with B and A sequences were short (≤ 7 nt).

Figure 2.10

Screen Results

	Library					
	J3		J6		Total	
Measure	No.pos.	% pos.	No.pos.	%pos.	No.pos.	%pos.
Hybe +	123/230	54	173/240	72	296/470	63
SspI +	19/48	40	28/48	58	47/96	49

Sequencing Results

Class	No.	%	%Total	No.	%	%Total	No.	%	%Total
A	2	13		3	20		5	16	
B	11	69		6	40		17	55	
C	0	0		0	0		0	0	
D	1	6		2	13		3	10	
E	1	6		1	7		2	7	
F	1	6		1	7		2	7	
G	0	0		1	7		1	3	
H	0	0		1	7		1	3	
Tot.A-H	16	100	84	15	101	54	31	101	66
TR	3		16	12		43	15		32
?	0		0	1		4	1		2
Total	19		100	28		101	47		100

independent classes (though note that H matched six of seven bases of the 3' end of exon 1A; see Figs. 2.8 and 2.9). A and B were the most abundant (16% and 55%, respectively), and included some clones that extended further 5' than the previously isolated pAR clones.

In addition to classes A and B, five new classes (D-H) were isolated, with D, E and F represented by at least one clone from each library. No clones of class C were found. Thus there appear to be at least eight GR transcripts (A-H) each containing a different 5' mRNA leader encoded by a family of alternate 5' exons, each of which is spliced to exon2. Of these transcripts, A and B appear to be most abundant (based on cloning frequency) in the J2.17 HTC cell line.

The GR gene is transcribed from more than one promoter

The multiple 5' exons discovered by cDNA cloning could have been generated by multiple promoters or by alternate splicing. In order to determine whether the multiple transcripts were generated by more than one promoter, I tested whether exon 1B was associated with its own promoter, downstream of exon 1A.

As Fig. 2.11 illustrates, a 290bp Kpn-Sac fragment immediately upstream of exon 1B displays promoter activity in transient transfection assays in an orientation-dependent manner. The 3' end of this region also contains multiple initiation sites in RNA from J2.17 and XC cells. RNase protection assays (Fig. 2.12) and Northern blots of RNaseH-truncated RNA (Fig. 2.13) indicated the presence of two to six major initiation sites distributed in a ~100 bp region around the SacI site (see Fig. 2.14). This position is

consistent with the location of the 5' ends of the most 5'-extending B cDNA clones (see Figs. 2.9 and 2.15).

The sequence of the upstream region containing promoter B is rich in CpG dinucleotides and has potential promoter elements

The genomic sequence of exons 1A and 1B and their 5' flanking regions is shown in Fig. 2.15. This upstream region is highly rich in G and C nucleotides, with several patches 60 bp or longer with >80% GC content (Fig. 2.17). The sequence of the B-promoter (the Kpn-Sac fragment) is 76% G + C. This upstream region also contains a higher frequency of CpG dinucleotides than is typical of most of the mammalian genome in which CpG dinucleotides are generally underrepresented (Russell et al., 1976; Swartz et al., 1962). Genomic DNA with high CpG content (where the ratio of CpG frequency observed to that expected by chance exceeds 0.6) and a G + C content of more than 50% has been associated with regions of hypomethylation ("HTF Islands"), found at the 5' ends of certain genes, including so-called "housekeeping" genes (Bird, 1986; Gardiner-Garden and Frommer, 1987).

I scanned the genomic sequence shown (Fig. 2.15) for potential transcription factor binding sites commonly found associated with promoters. There was no TATA box, and a weak potential CTF/NF-1 site in the B-promoter region (position -558 relative to ATG1; see Fig. 2.15). The latter sequence matched the CTF/NF-1 core consensus (RCCAR; Jones et al., 1987) perfectly, but did not match the predicted optimal binding site consensus outside this core. However, like the optimal consensus sequence, the -558 GCCAA is located in a palindromic sequence. The center of symmetry of this sequence is shifted -1 bp compared to the center of the palindromic NF-1

Figure 2.11. The 290 bp Kpn-Sac fragment upstream of exon 1B behaves as a promoter when placed upstream of the CAT reporter gene. Various genomic DNA fragments from a 2.2kb region upstream of exon 1B were inserted upstream of the chloramphenicol acetyl transferase (CAT) gene in the promoterless reporter plasmid pUC-OOCO. The fragments and their orientations are indicated by the arrows on the left-hand side of the figure, under a restriction map of the region (N, NcoI; H, HindIII; Sm, SmaI; Sph, SphI; P, PstI; K, KpnI; Sac, SacI). The relative CAT activities of these reporter plasmids transfected into CV1b cells is shown in the bar graph on the right. Equal molar amounts of each reporter were transfected (=6.4μg of NN+ and NN-) along with 0.1-2.0μg p6R.Z as a βgal control and pOTTO as carrier to a total of 6μg. In each experiment, the CAT activity of each transfected plasmid was normalized to βgal activity, and the resulting measure normalized to the experimental mean[5] . The results of 5 independent experiments were then analyzed by one-factor analysis of variance (F-test: $p<.0001$) and differences between means tested for significance by the Student-Newman-Keuls multiple comparisons test (Glantz, 1987). Significance levels are shown for comparisons with the parent reporter pOOCO (†, $p<.025$; *, $p<.01$; **, $p<.001$); error bars are standard errors of the mean, and numbers represent the numbers of experiments in which that plasmid was tested. The means and errors were normalized to NN+ ≡ 100% for graphical display.

[5] Normalizing to the experimental mean is more robust than normalization to a single reference plasmid. In one experiment only 8 of the 13 plasmids were tested, and in another one plasmid was missing. Since the missing vaues did not strongly affect the experimental means, the experiments with missing data were included in the analysis.

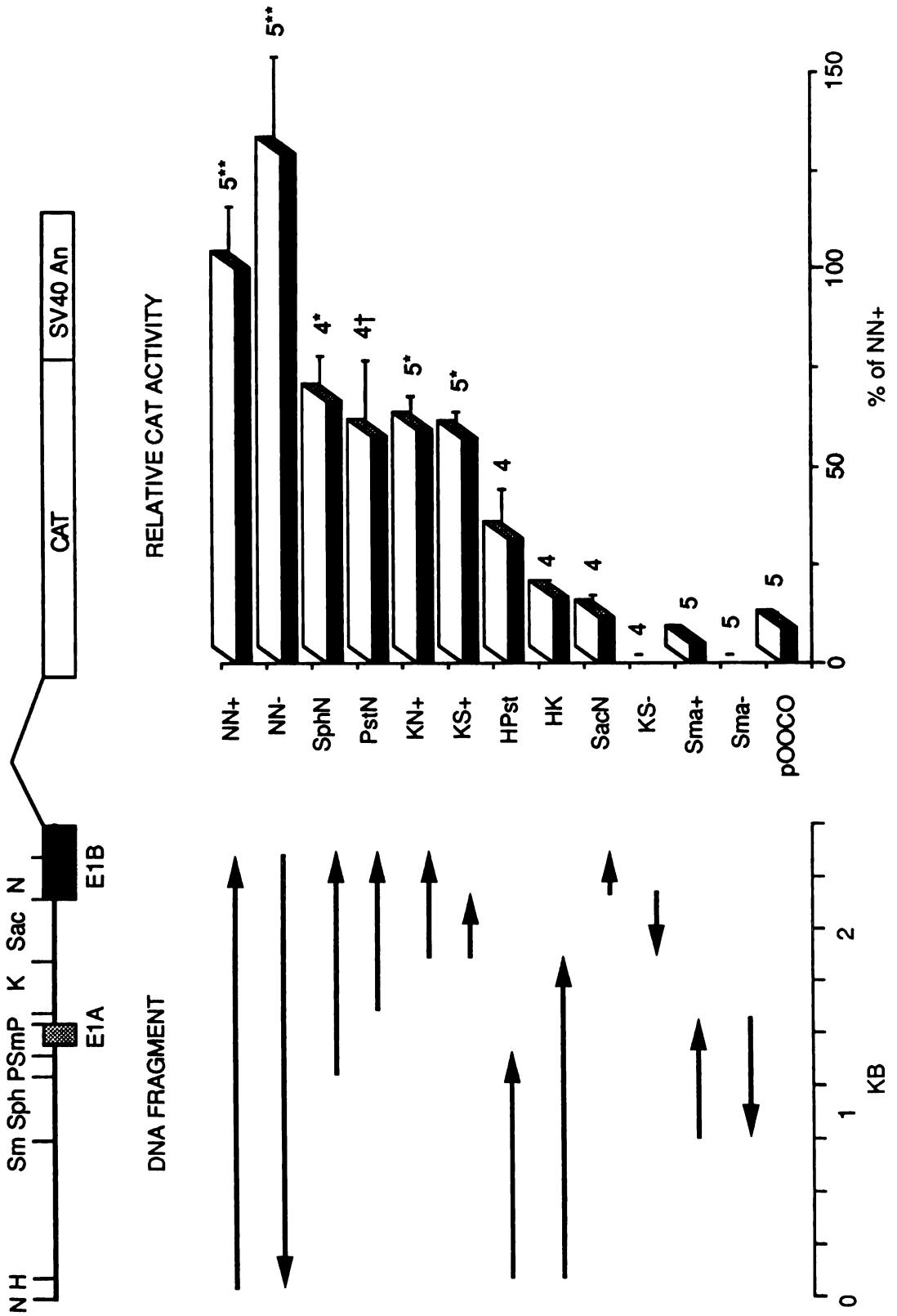

SV40 An
CAT
E1B
E1A
N
Sac
K
P
Sm
Sph
P
Sm
H
N
RELATIVE CAT ACTIVITY
% of NN+
0
50
100
150
5**
5**
4*
4†
5*
5*
4
4
4
4
5
5
5
NN+
NN-
SphN
PstN
KN+
KS+
HPst
HK
SacN
KS-
Sma+
Sma-
pOOCO
DNA FRAGMENT
KB
0
1
2

Figure 2.12. 5' mapping of transcript B by RNAse protection. 2 fmol of the indicated probe was hybridized at 55°C for 12 h with the indicated RNA (0, 50μg carrier tRNA only; S, sense control RNA SMα (0.025 fmol) plus carrier; E, 50μg EDR3 pA+mRNA; X, 20μg XC pA+mRNA; J, 50μg J2.17 pA+mRNA) in 20μl of S1 hybridization buffer. RNase digestion and acrylamide gels were as described in Materials and Methods. Probes are antisense T7 (ENα) or T3 (BSα) transcribed RNA as diagrammed in Fig. 2.14. Undigested probes (B, BSα; E, ENα) were loaded at 0.1 fmol each. Fragment sizes were estimated by reference to a pBR325xHaeIII end-labelled marker (Mkr). The autoradiogram in this figure is from experiment RP31 (see also Fig. 2.14).

Figure 2.12

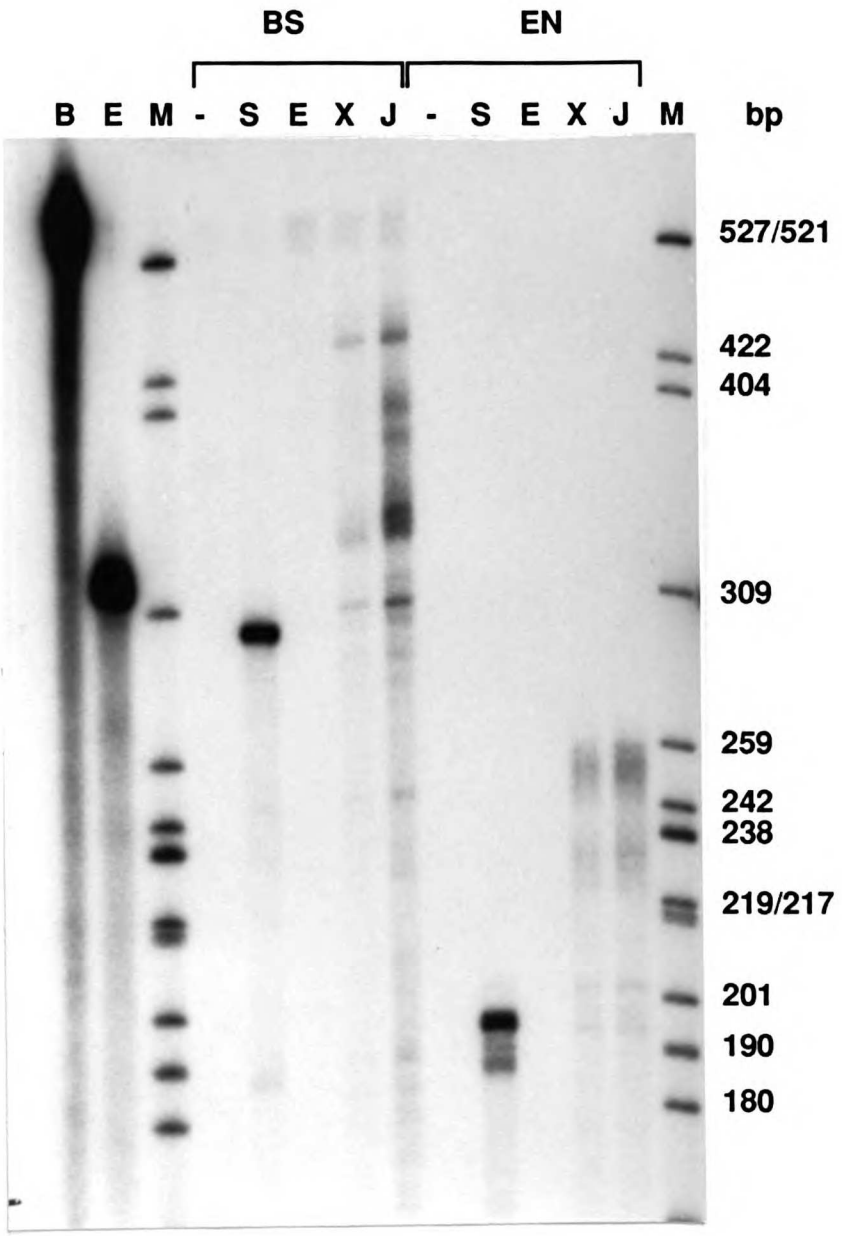

BS
EN
B E M - S E X J - S E X J M bp
527/521
422
404
309
259
242
238
219/217
201
190
180

Figure 2.13. 5' mapping of transcript B by RNAse H/ Northern. Sense control RNA mixture (C, 0.2 fmol each of T3/pAR8 x BamHI, T7/pAR1 x KpnI, T7/pAR5 x HindII, and 15µg tRNA carrier) or 15µg pA+ RNA (E, EDR3; J, J2.17; G, 19G11) were incubated with 300 pmol of oligo MJ04 (5'-CATTGCTTGTGGAG CCTTTCG-3') and then digested with RNase H, as described in Materials and Methods. The expected sizes of the hybridizing (5') portions of the control RNAs were 540nt (pAR8), 522nt (pAR1), and 374nt (pAR5). Nucleic acid was extracted and run on a 3.5% polyacrylamide-Urea gel alongside radiolabelled size markers (M, pBR325xHaeII end-labelled marker) and then transferred by semidry electroblotting (Novablot®, LKB) in 1xTBE at 0.8 mA/ cm^2 for 1 h to a Gene Screen Plus® membrane. The acrylamide gel was dried and exposed to X-ray film along with the wet membrane to monitor transfer efficiency. The labelled marker appeared to transfer completely (no radioactivity was detectable on the acrylamide gel by hand-held monitor). The membrane was then dried and cut into three pieces. Each portion was then hybridized to an antisense RNA probe from either exon 1B, 1A (not shown), or 2, and then washed, as with a Northern blot (see Materials and Methods).

Figure 2.13

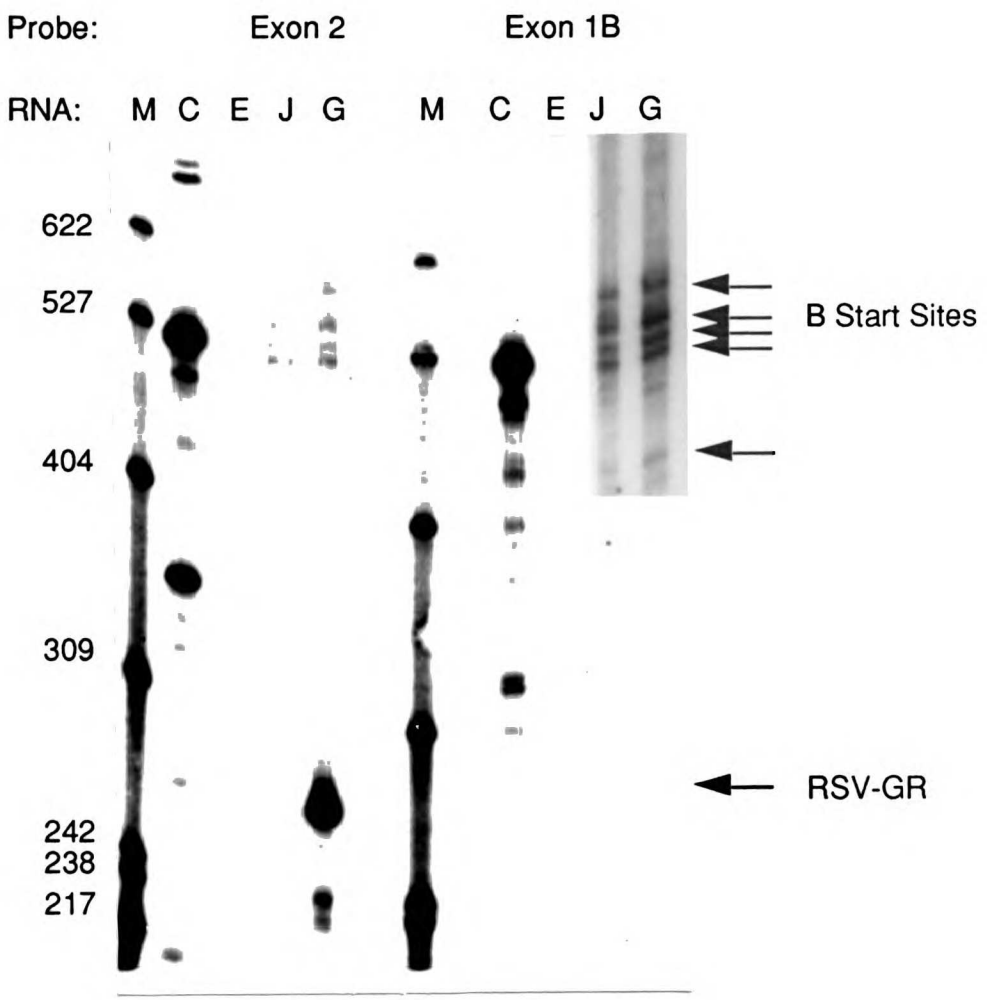
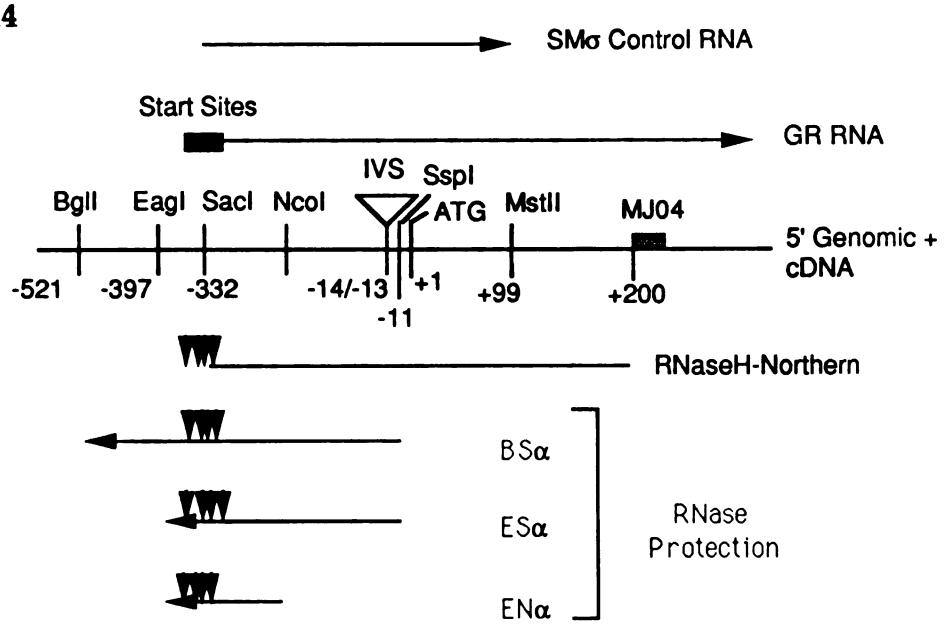

Probe:
Exon 2
Exon 1B
RNA:
M C E J G
M C E J G
622
527
404
309
242
238
217
B Start Sites
RSV-GR

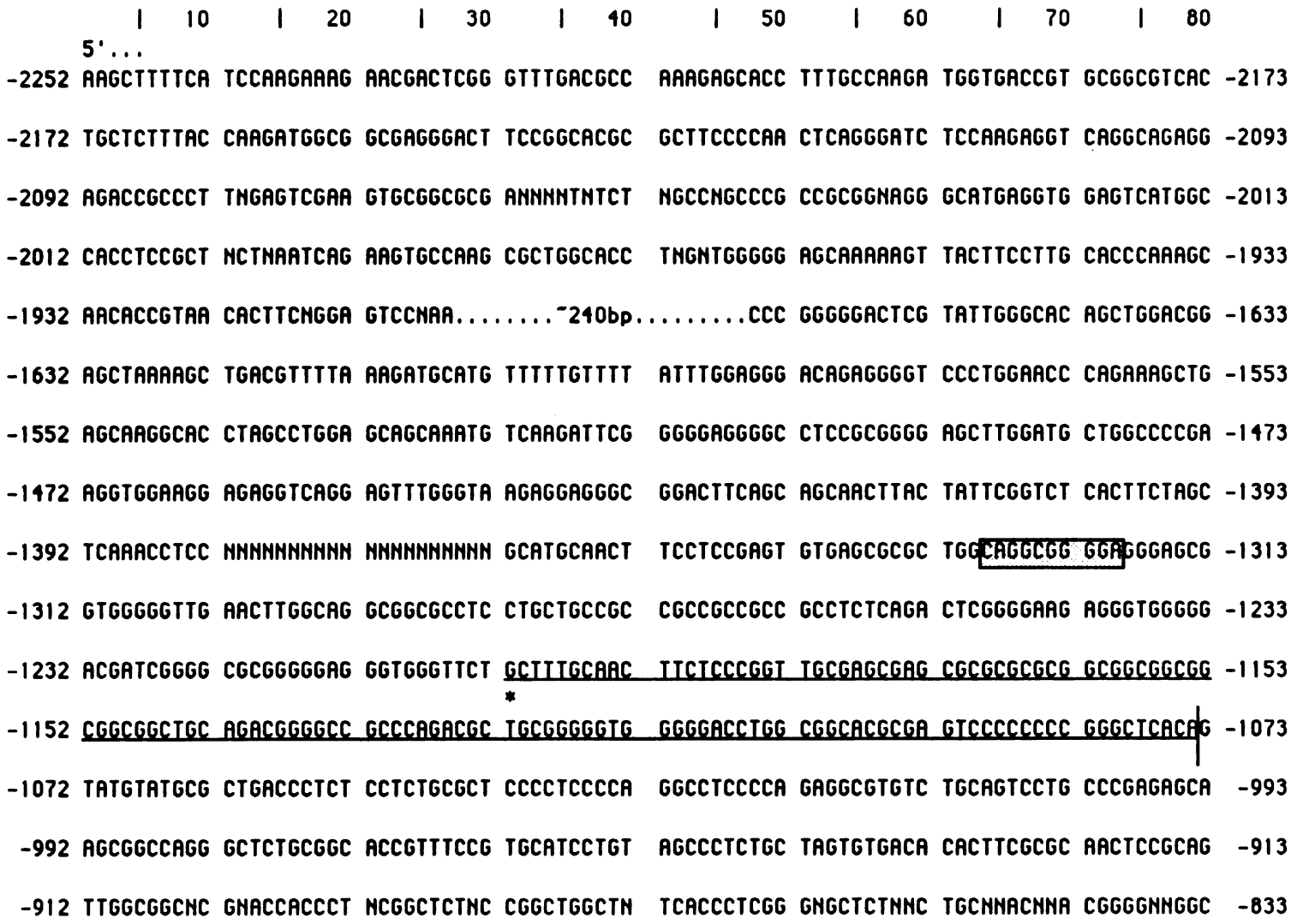
Figure 2.14. Summary of 5' mapping of transcript B. RNase protection and RNaseH-Northern mapping of receptor transcript B reveals the existence of 2-6 start sites located at positions spanning approximately -305 to -461 (where +1is defined as the start of the receptor large open reading frame).

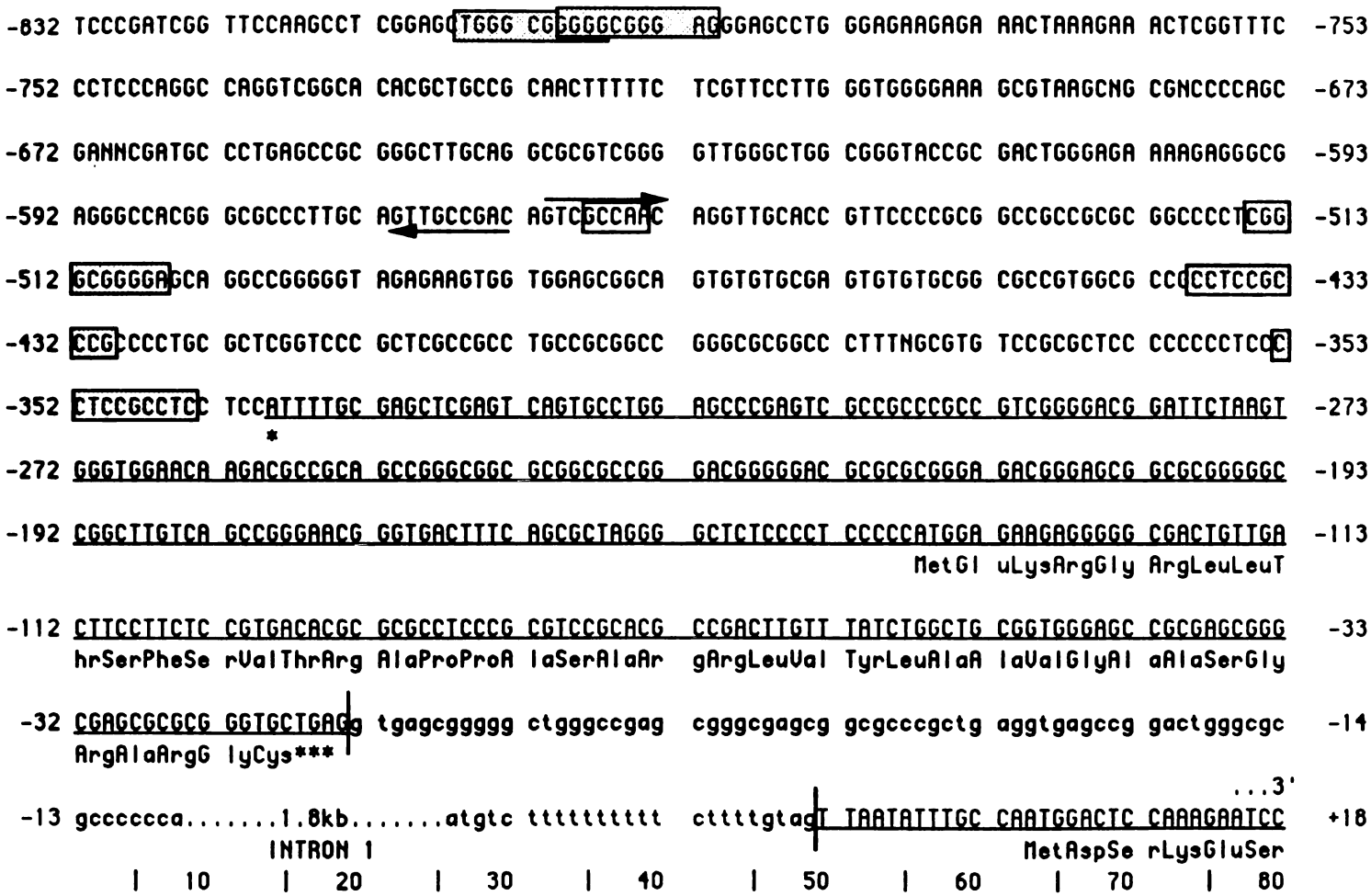
Figure 2.14



Probe/Oligo	Experiment	Fragment Sizes, bp (SMσ)	Est. Start Positions
MJ04	RH5	593, 560, 543, 527	-393, -360, -343, -327
BSα	RP31	453, 403, 381, 339, 331, 307, (298)	-461, -411, -389, -347, -339, -315, (-306)
ESα	RP23	379, 345, 319, 297, (321)	-387, -353, -327, -305, (-329)
ENα	RP23	240, 218, (196)	-369, -347, (-325)
ENα	RP31	256-252, 233, 202, 293, (196)	-385 - -381, -362, -331, -322

Figure 2.15. Nucleotide sequence of the 5' end of the rat glucocorticoid receptor gene. This sequence includes exons 1A and 1B, 5' flanking DNA (including the B-promoter), the ends of intron I, and the beginning of exon 2. The sequencing strategy is shown in Fig. 2.16. Nucleotides are numbered with +1 assigned to the A in the first ATG of receptor, skipping intron I sequence. Intron I sequence is shown in lower case, and exon sequences are underlined. The first bases of the most 5'-extended A and B cDNAs are indicated with a *, and intron/exon junctions with a |. Shaded boxes indicate Sp1-site consensus sequences, and the open box a potential CTF site. Arrows indicate regions of dyad symmetry.


| 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80
5'...
-2252 AAGCTTTTCA TCCAAGAAAG AACGACTCGG GTTTGACGCC AAAGAGCACC TTTGCCAAGA TGGTGACCGT GCGGCGTCAC -2173
-2172 TGCTCTTTAC CAAGATGGCG GCGAGGGACT TCCGGCACGC GCTTCCCCAA CTCAGGGATC TCCAAGAGGT CAGGCAGAGG -2093
-2092 AGACCGCCCT TNGAGTCGAA GTGCGGCGCG ANNNNTNTCT NGCCNGCCCG CCGCGGNAGG GCATGAGGTG GAGTCATGGC -2013
-2012 CACCTCCGCT NCTNAATCAG AAGTGCCAAG CGCTGGCACC TNGNTGGGGG AGCAAAAAGT TACTTCCTTG CACCCAAAGC -1933
-1932 AACACCGTAA CACTTCNGGA GTCCNAA.....~240bp.....CCC GGGGGACTCG TATTGGGCAC AGCTGGACGG -1633
-1632 AGCTAAAAGC TGACGTTTTA AAGATGCATG TTTTTGTTTT ATTTGGAGGG ACAGAGGGGT CCCTGGAACC CAGAAAGCTG -1553
-1552 AGCAAGGCAC CTAGCCTGGA GCAGCAAATG TCAAGATTCG GGGGAGGGGC CTCCGCGGGG AGCTTGGATG CTGGCCCCGA -1473
-1472 AGGTGGAAGG AGAGGTCAGG AGTTTGGGTA AGAGGAGGGC GGACTTCAGC AGCAACTTAC TATTCGGTCT CACTTCTAGC -1393
-1392 TCAAACCTCC NNNNNNNNNN NNNNNNNNNN GCATGCAACT TCCTCCGAGT GTGAGCGCGC TGGCAGGCGG GGAGGGAGCG -1313
-1312 GTGGGGGTTG AACTTGGCAG GCGGCGCCTC CTGCTGCCGC CGCCGCCGCC GCCTCTCAGA CTCGGGGAAG AGGGTGGGGG -1233
-1232 ACGATCGGGG CGCGGGGGAG GGTGGGTTCT GCTTTGCAAC TTCTCCCGGT TGCGAGCGAG CGCGCGCGCG GCGGCGGCGG -1153
-1152 CGGCGGCTGC AGACGGGGCC GCCCAGACGC TGCGGGGGTG GGGGACCTGG CGGCACGCGA GTCCCCCCCC GGGCTCACAG -1073
-1072 TATGTATGCG CTGACCCTCT CCTCTGCGCT CCCCTCCCCA GGCCTCCCCA GAGGCGTGTC TGCAGTCCTG CCCGAGAGCA -993
-992 AGCGGCCAGG GCTCTGCGGC ACCGTTTCCG TGCATCCTGT AGCCCTCTGC TAGTGTGACA CACTTCGCGC AACTCCGCAG -913
-912 TTGGCGGCNC GNACCACCCT NCGGCTCTNC CGGCTGGCTN TCACCCTCGG GNGCTCTNNC TGCNNACNNA CGGGGNNGGC -833


-832 TCCCGATCGG TTCCAAGCCT CGGAGCTGGG CGGGGCGGG AGGGAGCCTG GGAGAAGAGA AACTAAAGAA ACTCGGTTTC -753
-752 CCTCCCAGGC CAGGTCGGCA CACGCTGCCG CAACTTTTTC TCGTTCCTTG GGTGGGGAAA GCGTAAGCNG CGNCCCCAGC -673
-672 GANNCGATGC CCTGAGCCGC GGGCTTGCAG GCGCGTCGGG GTTGGGCTGG CGGGTACCGC GACTGGGAGA AAAGAGGGCG -593
-592 AGGGCCACGG GCGCCCTTGC AGTTGCCGAC AGTCGCCAAC AGGTTGCACC GTTCCCCGCG GCCGCCGCGC GGCCCCTCGG -513
-512 GCGGGGAGCA GGCCGGGGGT AGAGAAGTGG TGGAGCGGCA GTGTGTGCGA GTGTGTGCGG CGCCGTGGCG CCCCTCCGC -433
-432 CCGCCCCTGC GCTCGGTCCC GCTCGCCGCC TGCCGCGGCC GGGCGCGGCC CTTTNGCGTG TCCGCGCTCC CCCCCCTCCC -353
-352 CTCCGCCTCC TCCATTTTGC GAGCTCGAGT CAGTGCCTGG AGCCCGAGTC GCCGCCCGCC GTCGGGGACG GATTCTAAGT -273
-272 GGGTGGAACA AGACGCCGCA GCCGGGCGGC GCGGCGCCGG GACGGGGGAC GCGCGCGGGA GACGGGAGCG GCGCGGGGGC -193
-192 CGGCTTGTCA GCCGGGAACG GGTGACTTTC AGCGCTAGGG GCTCTCCCCT CCCCCATGGA GAAGAGGGGG CGACTGTTGA -113
MetGl uLysArgGly ArgLeuLeuT
-112 CTTCCTTCTC CGTGACACGC GCGCCTCCCG CGTCCGCACG CCGACTTGTT TATCTGGCTG CGGTGGGAGC CGCGAGCGGG -33
hrSerPheSe rValThrArg AlaProProA laSerAlaAr gArgLeuVal TyrLeuAlaA laValGlyAl aAlaSerGly
-32 CGAGCGCGCG GGTGCTGAGg tgagcggggg ctgggccgag cgggcgagcg gcgcccgctg aggtgagccg gactgggcgc -14
ArgAlaArgG lyCys***
...3'
-13 gcccccca.....1.8kb.....atgtc ttttttttt ctttgtagT TAATATTTGC CAATGGACTC CAAAGAATCC +18
INTRON 1
MetAspSe rLysGluSer
| 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80

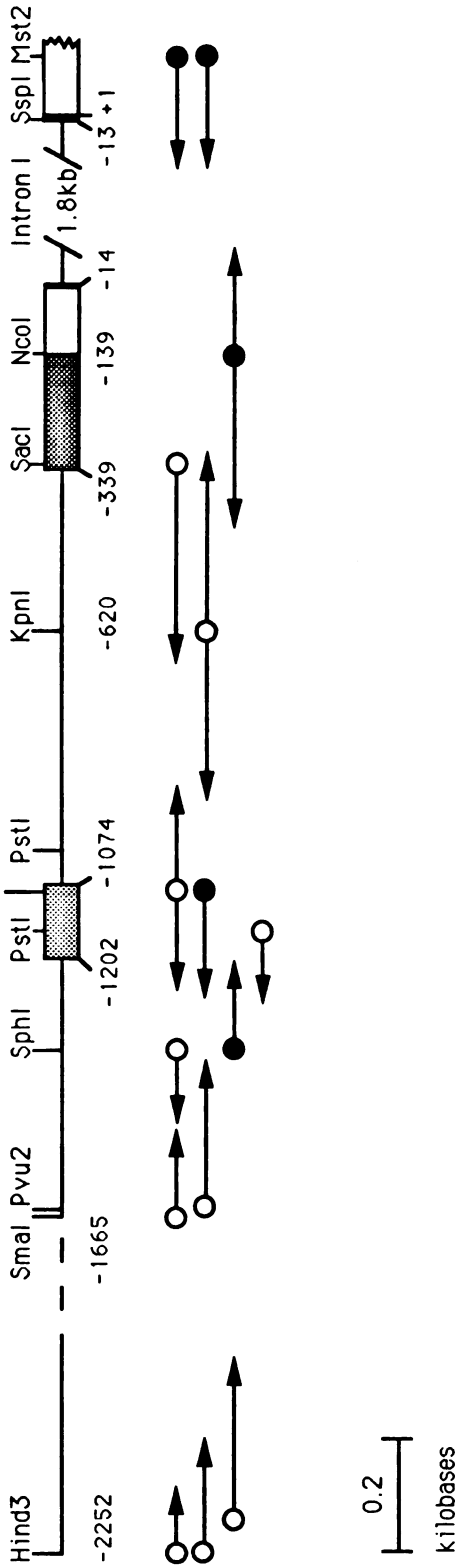
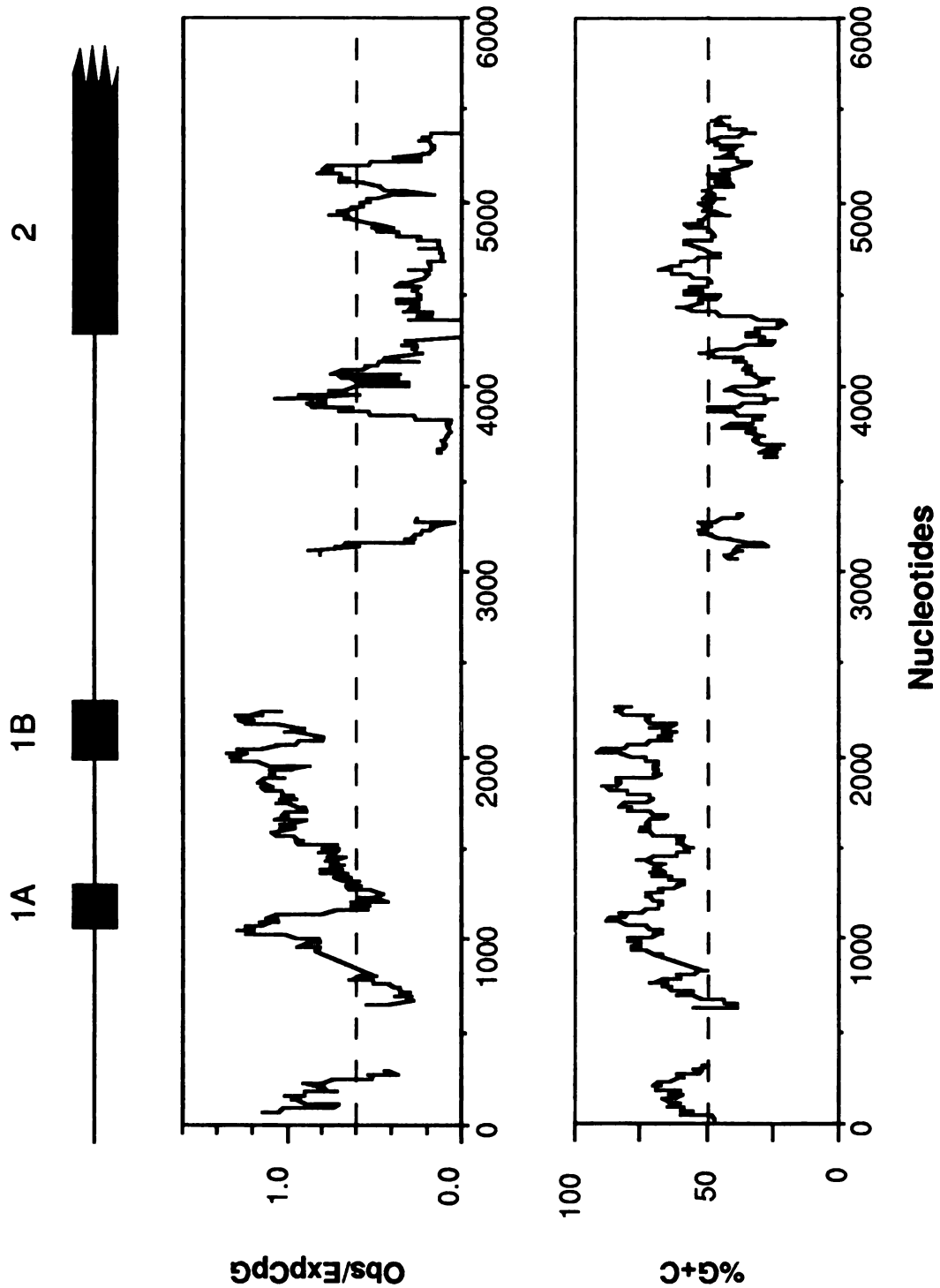


Figure 2.16. Sequencing strategy. Nucleotides are numbered as in Fig. 2.15, with +1 = the start of receptor coding sequence, and skipping intron I. Boxes denote exon sequences, with the unshaded portions indicating coding regions. Arrows indicate sequence runs. Open circles denote the dideoxy method, closed circles the method of Maxam and Gilbert (end labelled DNA).

Figure 2.17. Analysis of the distribution of CpG dinucleotides and G+C content in the 5' flanking region of rat glucocorticoid receptor. Moving averages of %G+C and Observed/Expected CpG (={CpG / (C x G)} x N) were calculated at 5 bp intervals; each point on the graph represents the mean value for a window of size N centered on that point (N = 60 bp for %G+C, N = 120 bp for Obs/Exp CpG). Exons are marked by filled boxes. Sequence with values of %G + C and Obs/Exp CpG both above the dashed lines are considered HTF-island-like regions (Gardiner-Garden and Frommer, 1987).


1A
1B
2
Obs/ExpCpG
1.0
0.0
%G+C
100
50
0
0
1000
2000
3000
4000
5000
6000
Nucleotides

consensus sequence. Six perfect matches were found to the central 8 bp of the 10 bp Sp1 site consensus (KRGGCGKRRY / RYYMCGCCYM); all are potential strong Sp1 binding sites (Briggs et al., 1986; Jones et al., 1986; Kadonaga et al., 1986). Three of these potential Sp1 sites are located in the B-promoter region, each ~80 bp from each other. The B-promoter region also contains a 27 nt stretch of 26/27 pyrimidines immediately upstream of the Sac I site (-366 to -340). Homopyrimidine-homopurine DNA sequences have been found associated with S1 nuclease hypersensitive sites in the 5' flanking regions of β globin genes (Schon et al., 1983) and in the dihydrofolate recuctase promoter (Masters and Attardi, 1985), though their function is not clear.

The 5' leaders contain open reading frames and potential RNA secondary structure

Transcripts B, C, and D all contain upstream initiator codons in their 5' leaders (Fig. 2.18), encoding short peptides (1 - 40 aa). No initiator was found in the transcript A leader, which maintains an open reading frame through the entire length of exon 1A (> 43 aa). Though the sequences of the other transcripts (E-H) cloned by PCR may also contain upstream initiators, these sequences have not been confirmed in more than two independent clones and so are less reliable due to possible mutations introduced by the PCR reaction. 5' leader sequences that have been obtained from the mouse receptor also apper to contain upstream initiators (Fig. 4.5; also in Dieken et al., 1990). One of these mouse 5' leaders is almost identical to the rat A leader except for the substitution in the rat transcript of a C in place of the initiator A in mouse (compare rat A in Fig. 2.15 with mouse 55Rn23 in Fig. 4.5).

Though this degree of sequence similarity (1 mismatch in 53 nucleotides) could indicate conservation of function, it could also simply reflect the recent divergence of rats and mice. To determine whether this degree of similarity might be expected between homologous but non-conserved rat and mouse sequences, I analyzed the number of mismatches between an alignment of rat and mouse receptor 3' untranslated region (3'UTR) sequences. Comparison of the first 208 nucleotides of these sequences (for which published sequence is available in both species (Danielsen et al., 1986; Miesfeld et al., 1986)) revealed 22 mismatches and 2 gaps. For all 50 nucleotide stretches in this region, the range was 1 - 11 mismatches and 0 - 2 gaps. To estimate the probability of finding no more than one mismatch in 50 nucleotides of these sequences, I ran a Mone-Carlo simulation: I shuffled 10 nt sections of the 3' UTR alignment multiple times, choosing five random 10 nt windows, to yield a random frequency distribution of expected 50 nt alignments. This distribution (Fig. 2.19) approximates a normal distribution, with mean = 5.2 and standard deviation = 2.9 (n=150). One mismatch in 53 (0.94 mismatches in 50nt) is within 1.5 standard deviations of the mean of this distribution: assuming a normal distribution, the probability of finding no more than one mismatch in 50 nucleotides is about one in 15. Though these results do not *rule out* functional conservation of the A leader sequences, the observed sequence similarity appears to be within that expected for non-conserved rat and mouse receptor sequences.

The sequence surrounding the ATG codon in B is similar to the receptor long ORF initiator sequence 1, and the initiator C-4 is close to the optimal consensus sequence, suggesting that these ORFs might be translated. Some of these upstream initiators might be expected to affect the translational efficiency of the downstream receptor long ORF (for reviews see Kozak, 1986a,

Figure 2.18. Upstream AUGs in rGR transcripts. Nucleotide positions of receptor open reading frames are given relative to the first initiator AUG in exon 2 (A = +1), and do not include intron sequences. The exon 1/2 splice junction occurs between nucleotides -14 and -13. ORFs C-1 and D-1 both terminate in exon 2; the other 5' ORFs terminate upstream of exon 2.

Transcript Class	Number of ATG Codon	Nucleotide Position of Open Reading Frame		Length of Open Reading Frame (amino acids)	Flanking Sequence Surrounding ATG	Reading Frame Relative to Rat GR
		ATG Codon	Terminator Codon			
A	n.a.	none	none	>43	n.a.	n.a.
B	B-1	-137	-17	40	CCCCCATGG	+1
C	C-1	-58	+33	30	GAGAGATGC	-1
C	C-2	-110	-104	3	ATGTGATGC	+1
C	C-3	-115	-112	1	TTCTGATGT	-1
C	C-4	-168	-123	15	GTACCATGG	same
C	C-5	-240	-123	39	GANAGATGC	same
D	D-1	-31	+33	31	TCGGCATGC	-1
All (exon2)	1	+1	+2386	795	TGCCAATGG	same
All (exon2)	2	+82	+2386	768	GCGTAATGG	same
			Optimal/ Consensus Initiator[a]		CC(A/G)CCATGG	

[a] Kozak, 1986c; Kozak, 1987

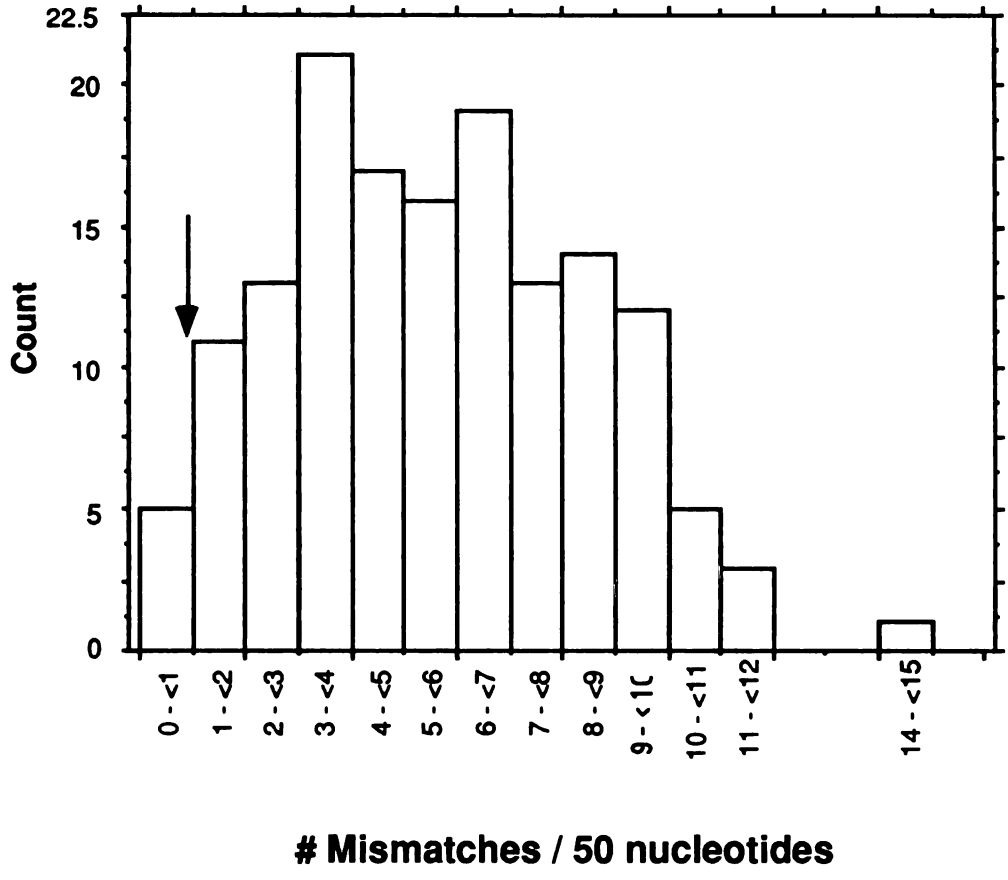


Figure 2.19. Frequency distribution of expected mismatches between rat and mouse receptor untranslated sequence, based on the proximal 3'UTR sequence. The first 208 nucleotides of rat and mouse receptor 3'UTR were aligned and divided into 21 consecutive 10 nt windows. The number of mismatches in each window was counted (the last window had 2 mismatches in 8 nt, but was counted as 2.5 mismatches in 10 nt). Five observations (#mismatches/10nt window) were drawn randomly, without replacement, from this set, to give the number of mismatches/50 nt. Repeating this procedure 150 times gave the frequency distribution shown. Assuming a normal distribution, mean = 5.2, standard deviation = 2.9 (n = 150). The observed mismatch frequency (1/53 = 0.94) between the 3' ends of rat and mouse exons 1A is shown by the vertical arrow.

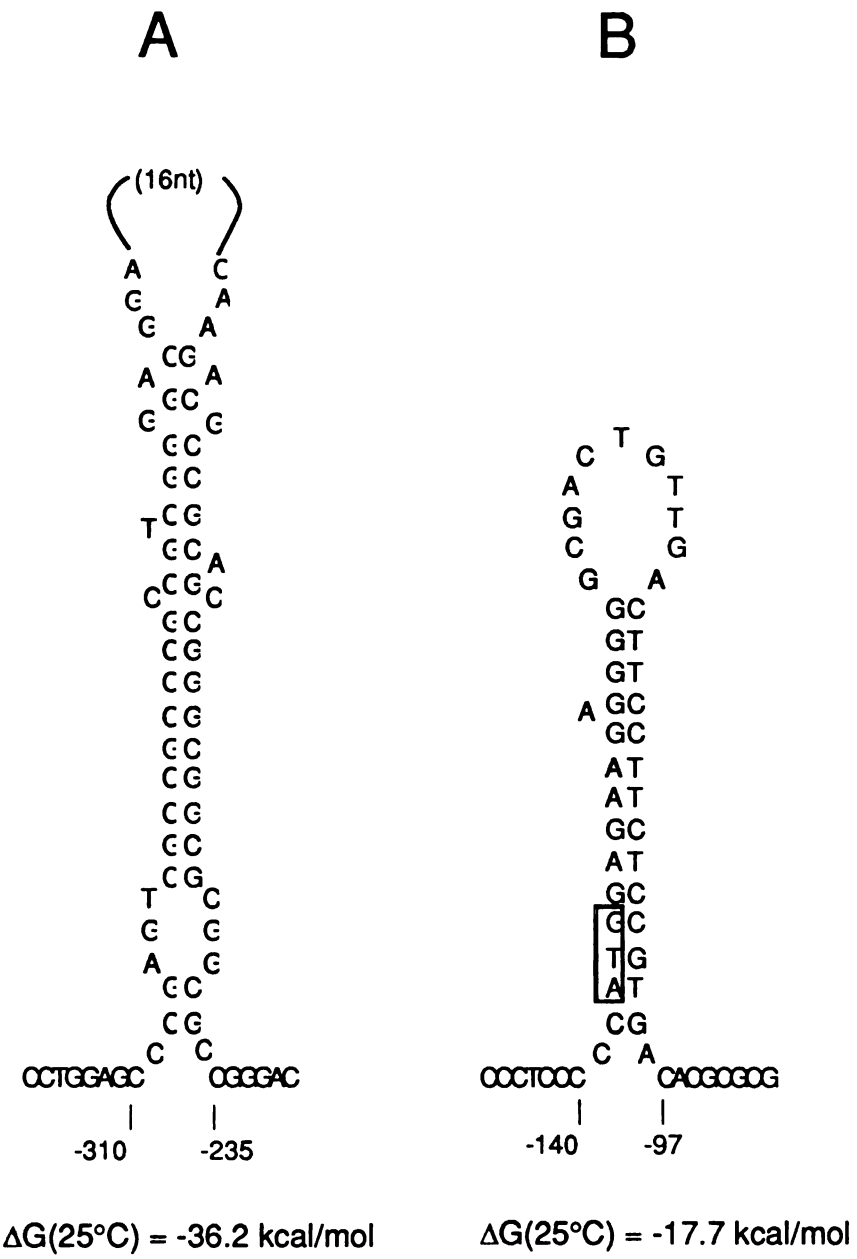


Figure 2.20. Secondary structures in the 5' end of rat glucocorticoid receptor transcripts B mRNA. Two potential hairpin structures (A and B) are shown. These sequences were determined from the cDNA (Fig. 2.9) and genomic DNA (Fig. 2.15). The indicated free energy values for the base-paired regions were calculated according to the method of Salser (1978). Similar values were obtained by the method of Tinoco et al. (1973). Note the ATG initiation codon hairpin B.

1988). Kozak has shown, for example, that the translational efficiency of the preproinsulin gene is inhibited by the presence of an initiator codon in its 5' leader, depending on the sequences surrounding the upstream initiator (Kozak, 1986c) and on the spacing between the upstream cistron and the downstream initiator (Kozak, 1987). The B 5' leader also appears to inhibit downstream receptor translation, as I will show in Chapter Three.

Upstream ORFs might also affect the choice of receptor initiator codon, since some of these ORFs (C-1, D-1) overlap the receptor long ORF at ATG1. If these ORFs were translated, receptor translation might initiate at AUG2 instead of AUG1, or be inhibited altogether. Multiple translation products have been observed with glucocorticoid receptor, consistent with translation from multiple initiaton codons (Miesfeld et al., 1986). The choice of initiation codon has been shown to affect the activity of progesterone receptor in a cell and promoter specific fashion (Kastner et al., 1990a,b). It is unknown, however, whether the choice of initiation codon affects receptor activity.

In addition to an upstream ORF, transcript B contains potential RNA hairpin structures, one of which overlaps the upstream ATG (Fig. 2.20). It is unknown whether these potential hairpin structures are conserved in homologous receptor leader sequences from other species. Since RNA secondary structures are difficult to predict based on free-energy calculations alone, phylogenetic conservation is the strongest test of functional RNA secondary structure. Nevertheless, these potential hairpins indicate strong self-complementarity within this 5' leader sequence and suggest the potential to form stable RNA secondary structures. Since RNA secondary structure has been associated with translational regulation (Aziz and Munro, 1987; Chevrier et al., 1988; Hentze et al., 1987a; Kozak, 1986b, 1989), this 5' leader might affect receptor translation (this possibility is investigated in Chapter

Three). Moderately stable RNA hairpin structures upstream of an initiator codon, for example, appear to inhibit translational efficiency by preventing ribosomal assembly (Kozak, 1989).

Discussion

Steroid receptor sequences and gene structures are highly conserved, as reflected in the clustering of the steroid receptor subfamilies in the phylogenetic tree and in the greater similarity of the steroid receptor gene structures compared to the gene structures of other members of the nuclear receptor superfamily (Fig. 2.5). These data are consistent with the hypothesis that the steroid receptors evolved from a common ancestral gene, which lost several of its introns and diverged from the vitamin D3, thyroid, retinoic acid, and various orphan receptor subfamilies. Positions of the introns support this hypothesis. Splice positions and phases are identical in GR, MR, AR, PR, and ER (with the exception of the last intron of ER and allowing for uncertain alignments in non-conserved regions).

Amino-terminal sequences of the various steroid receptors are completely divergent. Lack of similarity among the N-terminal exons could be due either to genetic drift or to exon shuffling. In either case, sequence divergence may imply a lack of selection pressure. Alternatively, these sequences have been fixed by selection for distinct functions. Though portions of the N-terminal domain may contain distinct receptor-specific functions, this seems unlikely for the entire N-terminal exon since large regions can be deleted without apparent affects on receptor function. That the phase of the splice site at the exon 2 - 3 border is conserved within the steroid receptor subfamilies (G and E) suggests that the ancestor of these proteins also

contained an N-terminal domain, and that this domain has diverged rather than being introduced or exchanged later by exon shuffling.

Eukaryotic gene structures reflect the evolutionary history of the proteins they encode. Structural analyses of many genes and gene families supports the view that exons encode discrete functional or structural modules (reviewed in (Dorit et al., 1990; Gilbert, 1987). The GR gene structure, like the structures of other steroid receptor genes, also appears to preserve correspondences between its exons and functional domains. By reconstructing some of these primitive domains within the steroid receptor subfamily, these correspondences appear even more convincing. For example, GR exon 5 encompasses the non-conserved "hinge" region flanked by portions of the zinc-finger domain (at the 5' end), and the signal transduction domain (at the 3' end). In the thyroid hormone receptors (ck erbA, r erbAα2, rev erbAα, h TRβ), retinoic acid receptor-γ, NGFIB, and E75B genes, there is an intron within the hinge region, separating the homologous portions of the zinc-finger and signal transduction domains in GR exon 5. In the vitamin D3 receptor (VDR) gene, the hinge region is split by two introns, also separating the zinc-finger and signal transduction domain portions homologous to GR exon 5. Interestingly, the intron at the 5' end of exon 5, separating this exon from exon 4, is completely conserved in all the nuclear receptor superfamily genes mapped to date, with the exception of rev ERBAα. This striking degree of conservation parallels the strong sequence conservation in this region, and may reflect the lack of viable splice-site alternatives.

The intron separating exons 3 and 4 is somewhat variable, with four different positions conserved within different subfamilies. As pointed out by (Ponglikitmongkol et al., 1988), the position and phase of this intron is conserved within each of the steroid receptor and thyroid hormone/ retinoic

acid receptor subfamilies. The position and phase of the intron in E75B and egon genes is the same as in the TR and RAR subfamilies. Interestingly, E75A and egon both cluster on the phylogenetic tree most closely with the retinoic acid receptors. In addition, the genes for NGFIB/N10 and for VDR each split the two exons at different positions, consistent with their grouping into separate branches of the phylogenetic tree.

The glucocorticoid receptor gene is large, at least 125 kb long. It is interesting to note that complete transcription of a gene of this size by RNA pol II would be expected to take about 1.5 h, based on a transcript elongation rate of 20-25 bp/sec *in vivo* (Ucker and Yamamoto, 1984). Dexamethasone has been shown to cause a decrease in GR initiation rate (measured by run on assays) without changing GR mRNA half-life (Dong et al., 1988; Rosewicz et al., 1988). Thus the large size of the GR gene would be expected to cause a ≥1.5 h delay before a dex-induced change in the initiation rate would begin to affect steady-state transcript levels. This prediction is consistent with the nuclear run-on data of Dong, et al. (1988), showing no change in transcription rate in rat liver nuclei ~2h after dex treatment, and a maximal decrease after about 4h. This lag would explain the differences observed by Rosewicz, et al. (1988) between rat receptor mRNA half-life (4h) and the half-time for dex-induced down-regulation of steady-state receptor mRNA (6h). Though the lag in receptor mRNA synthesis appears to contribute to the kinetics of receptor mRNA regulation, it is not clear what role this kinetics plays in receptor physiology, or whether the large size of the receptor gene is an important aspect of receptor regulation.

Of the steroid receptor genes that have been mapped, all are longer than 30 kb (chick progesterone receptor (ckPR), ~34 kb; human androgen receptor (hAR), > 77 kb; human estrogen receptor (hER), > 165 kb). Of the

other members of the superfamily that have been mapped, some are long (human vitamin D3 receptor (hVDR), > 37 kb; human thyroid hormone receptor β (hTRβ), > 60 kb; Drosophila E75A, ~ 50 kb), and some are relatively short (rat rev-erbAα, rat erbAα, ~ 5 kb; chick c-erbAα, ~ 8 kb; rat NGFIB/N10, ~ 8 kb). Interestingly, the longer genes are those that are clustered more closely with GR in the phylogenetic tree. This may simply be a reflection of a difference in the gene sizes of early ancestral genes that has propagated to the present gene structures. Nevertheless, it is interesting to speculate on possible functional roles gene size might play. For example, D. Hogness has suggested that large transcription units might serve as developmental clocks (ref. in Laughon et al., 1986). This hypothesis might help explain the observation that the homeotic genes *Antennapedia* and *Bithorax*, which have long transcription units, are expressed after a few hours' delay following the expression of the shorter segmentation genes (which are involved in regulating the homeotic genes).

B transcripts are the highest in relative abundance in HTC cells. The B promoter appears to be similar to the GC-rich promoters of many essential "housekeeping" enzyme or growth-control genes (Hoffman et al., 1987; Ishii et al., 1985a,b; Melton et al., 1984; Reynolds et al., 1985; Valerio et al., 1985) and contains multiple initiation sites, similar to the HMG Co-A reductase promoter (Reynolds et al., 1985). It also shares some similarities with the PR promoters, which are GC-rich, lack a TATA box, and contain upstream ORFs. Inspection of 5' sequences did not reveal a GRE, though this is not surprising since composite GREs do not seem to conform to a consensus sequence like the simple GREs (Sakai et al., 1988). Since there are multiple GR promoters, it is possible that GREs might be located near one of the other promoters, in an intron, or far upstream of the main transcription units. An analysis of a

human GR promoter has been published recently (Zong et al., 1990) that shares may characteristics in common with promoter B (both are GC-rich, lack clearly-defined TATA, CAAT, or GRE sequences, and transcribe 5' leaders containing upstream ORFs), but the sequences of these two promoters do not appear to be similar.

The existence of multiple promoters could reflect the separation of regulatory control regions. If this is the case, one might expect that these GR promoters might be differentially regulated. Different promoters might be expressed in different tissues, developmental stages, or respond differently to glucocorticoids. For example, two rat c-erbA (thyroid hormone receptor) transcripts, β1 and β2, differ at their 5' ends, suggesting they might be expressed from alternate promoters. The β2 transcript is pituitary-specific and autoregulated by thyroid hormone, whereas the β1 form is not (Hodin et al., 1989).

Alternate promoter usage could be functionally important if the promoters produced transcripts with different, independently regulated properties. Different 5' leaders might contain different regulatory elements affecting mRNA stability or translational efficiency. The transcription of many proto-oncogenes and other cellular genes alternates between two promoters: one producing a long, GC-rich leader sequence with upstream out-of-frame AUG codons, while the other produces a shorter, simpler leader sequence (Kozak, 1988). Some of these complicated leader sequences have been shown to impair translation (Marth et al., 1988; Propst et al., 1987; Rao et al., 1988; Ratner et al., 1987). Indeed, evidence from molecular genetic studies supports a role for upstream AUGs and mini-cistrons in the translational control of downstream coding sequences. Engineering upstream cistrons into the preproinsulin mRNA leader, for example, inhibits insulin translation

(Kozak, 1987). A well-studied example of translational control by upstream cistrons in the mRNA leader is the gene for the yeast transcriptional activator GCN4 (Hinnebusch, 1984; Thireos et al., 1984; reviewed in Hinnebusch, 1988). The translational efficiency of this mRNA is modulated in response to amino acid starvation by *trans* acting factors whose sites of action are at four short open reading frames within the 5' leader (Mueller et al., 1987; Mueller and Hinnebusch, 1986; Tzamarias et al., 1986). The peptide products of these short ORFs do not appear to be involved in this regulation. Another example, also from yeast, does involve the peptide product of the uORF in translational regulation. The 5' leader of CPA1 contains a mini-cistron whose peptide product is necessary in *cis* for negative translational regulation of CPA1 by CPA-repressor and arginine (Werner et al., 1987).

RNA secondary structure is also affected by the 5' leader, and could affect translational control. An RNA hairpin, depending on its stability and position relative to the AUG or the 5' cap, has been shown to block translation in eukaryotic cells of the mRNA containing it (Kozak, 1986b, 1989). Translational control by RNA secondary structure, as with upstream AUGs, can act as a site for regulation (Kozak, 1988). In the ferritin receptor transcript, translational control is exerted by a regulatory site created by an RNA hairpin in the 5' leader (Rouault et al., 1988). This site, called an IRE (iron-responsive element), is necessary for regulation of ferritin receptor translation by iron (Aziz and Munro, 1987; Hentze et al., 1987a,b).

At least three of the GR 5'UTRs (A, B and D)[6] appear to contain one or more open reading frames. Transcripts B and D encode upstream AUGs in

[6] The sequences of cDNA clones of transcripts C, E, and G also indicate the presence of 5' AUGs, but these sequences have not been confirmed by analysis of more than two independent PCR clones or by genomic sequences.

their 5'UTRs which could affect translational efficiency. Transcript B in particular contains both a single 40aa open reading frame (Fig. 2.18) and potential RNA hairpin-loop structures (Fig. 2.20), either or both of which might affect GR translational efficiency. By Kozak's scanning model, an upstream AUG in a "good context" would be expected to inhibit translation from downstream AUGs (Kozak, 1987). Because of the short distance between the end of the 5' B leader cistron and the first AUG of the GR ORF (AUG1), it is likely that translation from AUG1 would be inhibited by the B leader (Kozak, 1987). Because AUG B has a weak consensus for translational initiation, a reinitiation mechanism could allow some pre-initiation ribosomal complexes (40S) to skip AUG B and be available to initiate translation from AUG1 (Kozak, 1986a).

Some of the upstream AUGs in other GR 5' leaders might also direct which downstream GR AUG is used. Transcript D, for example, would be expected to produce a slightly smaller GR species translated from AUG2 (aa+28). Indeed, two closely spaced GR bands are often observed on protein gels of both cellular and reticulocyte lysate-produced GR. It is not known whether these two GR species are functionally different. Interestingly, two progesterone receptor forms A and B, one an amino-truncated version of the other, have been observed. These two forms are expressed from distinct mRNAs expressed from alternate promoters and are differentially active on different progestin-responsive promoters. The truncated form A, but not the larger form B, is active on the ovalbumin promoter, whereas both forms enhance transcription from the mouse mammary tumor virsus long-terminal repeat (MMTV LTR; Tora et al., 1988). Finally, one cannot rule out the possibility that 5' ORF-produced peptides themselves might function as

cis-acting translational regulators, similar to the action of the CPA1 leader ORF (Werner et al., 1987).

5' leaders containing short open reading frames have also been found in estrogen receptor (Green et al., 1986; Krust et al., 1986) and progesterone receptor (Kastner et al., 1990b); however, it is not known whether any of the potential RNA secondary structures or the upstream open reading frames in transcripts B or D are conserved in GR 5' leaders from other species. The transcript A 5' leader does have a counterpart in mouse, however. Though mouse and rat are too closely related to infer much by this conservation, the mouse transcript contains an upstream AUG, substituting an A for a C in the rat transcript. It is possible that the rat transcript also contains an upstream AUG, perhaps encoded by an exon further upstream, since there were no termination codons found in the 5' cDNA transcript A sequence.

Together, the complexities of the rat glucocorticoid receptor gene promoters and 5' leader sequenes suggest multiple routes to regulate receptor levels, including alternate promoter usage and translational control. I explore these possibilities in the next chapter.

Chapter Three: Glucocortioid Receptor Gene Expression

Introduction

Evidence from studies of several systems indicates that the magnitude of cellular responses to glucocorticoids is proportional to the number of intracellular glucocorticoid receptors. For instance, early radioligand binding studies showed a direct relationship between nuclear glucocorticoid receptor number and effect (Bloom et al., 1980; Mayo and Palmiter, 1981). The sensitivity of lymphocytes to glucocorticoid-induced cytolysis is also proportional to the number of glucocorticoid receptors (Bloomfield et al., 1981; Bourgeois and Newby, 1979). Vanderbilt et al. (1987) constructed HTC lines expressing from 0.4 to 3.5 times the parental cell line's levels of glucocorticoid receptor and found that the magnitude of several transcriptional responses was directly proportional to the numbers of receptors. For glucocorticoid-induced genes, overall enhancer strength thus appears to be limited by the intracellular concentration of receptor.

Given the direct relationship between receptor levels and response, it is clear that modulation of receptor levels is one way to regulate glucocorticoid-induced responses. In some mouse T-cell lymphoma lines, glucocorticoid-insensitive variants have been isolated which reversibly decrease their levels of glucocorticoid receptors in response to dexamethasone (Danielsen and Stallcup, 1984; Gehring et al., 1982). In C6 glioma cells, for example, ConA reduces the normal induction of glycerol-phosphate dehydrogenase (GPDH) by dexamethasone, coincident with a 90% decrease in receptor binding sites (McGinnis and De Vellis, 1981). Pre-pro opiomelanocortin (pre-POMC) mRNA expression and protein processing are differentially regulated in the intermediate and anterior lobes of the pituitary. POMC expression and processing is regulated by glucocorticoids in the anterior lobe, but not in the

intermediate lobe, resulting in the production and release of different peptides from these two regions. This dramatic difference in glucocorticoid response is explained by a complete lack of receptor immunoreactivity in the intermediate lobe (Antakly and Eisen, 1982, 1984), which appears to be due to a tonic repression by the dopamine innervation of the intermediate lobe (Antakly et al., 1987, 1985; Seger et al., 1988). The distribution of receptor protein and mRNA in the brain is particularly heterogeneous (Aronsson et al., 1988; Fuxe et al., 1987; Gustafsson et al., 1987; Sousa et al., 1988) and correlated, suggesting that much of this variability in receptor expression in brain is controlled at the level of transcription or mRNA stability (Sousa et al., 1988).

The receptors for many neurotransmitters and hormones are regulated by their cognate ligands (Poste and Crooke, 1985). Insulin, for example, decreases the half-life of its receptor, contributing to negative feedback regulation of insulin action (Levy and Olefsky, 1990). Similarly, cholesterol metabolism contains feedback controls: HMG CoA reductase, the limiting enzyme in cholesterol synthesis, is negatively regulated by cholestrol (Goldstein and Brown, 1990), and the mRNA for the low-density lipoprotein (LDL) receptor is feedback regulated by LDL (Russell et al., 1983). The glucocorticoid receptor has been shown to be negatively regulated by glucocorticoids in several systems (Burnstein et al., 1990; Cidlowski and Cidlowski, 1981; Danielsen and Stallcup, 1984; Dong et al., 1988; Gehring et al., 1982; Kalinyak et al., 1987; McIntyre and Samuels, 1985; Okret et al., 1986; Rosewicz et al., 1988; Sapolsky et al., 1984b, 1985; Sapolsky and McEwen, 1985; Schelechte et al., 1982; Shipman et al., 1983; Smith and Shuster, 1984; Svec and Rudis, 1981). Adrenalectomy is associated with an increase in glucocorticoid receptor protein and mRNA (Kalinyak et al., 1987; Meaney et al., 1985; Reul et

al., 1989, 1987; Sapolsky et al., 1984b; Smith and Shuster, 1984), suggesting that receptor levels are under some degree of tonic feedback control by adrenal corticosteroids. Autoregulation of receptor levels is slow (max decrease in 18h or more), involving regulation of both transcription (Dong et al., 1988; Rosewicz et al., 1988) and protein stability (McIntyre and Samuels, 1985; Svec and Rudis, 1981). Though most cells show down-regulation, there may be some tissue specificity. Dexamethasone induced up-regulation has been observed in rat anterior pituitary (Sheppard et al., 1990) and in a human T-cell line (Antakly et al., 1989; Eisen et al., 1988).

Regulation of glucocorticoid receptor levels may play an important role in the regulation of the hypothalamic-pituitary-adrenal (HPA) axis. For instance, Sapolsky et al. (1986) have suggested that hippocampal receptor levels might control the degree of glucocoticoid negative-feedback in turning down HPA axis activity after stress. Glucocorticoids also negatively regulate their own circulating levels by inhibiting the secretion of adrenocorticotrophin (ACTH) by the pituitary (Dallman et al., 1987), and corticotrophin releasing factor (CRF) in the hypothalamus (Kovács et al., 1986).

The HPA system undergoes dramatic anatomical and physiological changes perinatally (reviewed in De Kloet et al., 1988; Milkovic and Milkovic, 1969). Near the end of gestation in the rat, the fetal adrenal cortex is large and there is a large increase in fetal circulating corticosteroids. In sheep this prenatal rise in fetal corticosteroid levels may be important in triggering parturition and induction of fetal pulmonary surfactant (Norman and Litwack, 1987), though the function of this rise is not clear in the rat. At this time prenatally the HPA system is reponsive to negative-feedback inhibition by corticosteroids (). At birth the rat fetal cortex involutes, basal corticosterone

titers drop, and corticosterone binding globulin (CBG) levels decrease; after birth, corticosterone and CBG levels remain low for the first ten days or so, then gradually rise to adult levels (Bartova, 1968; Chatelain et al., 1980; Koch, 1969; Sakly and Koch, 1981; Schoenfield et al., 1980). Though HPA physiology appears to be different in neonates and adults, many of the details are not clear. Much of the literature over the last three decades has suggested that neonates do not respond to stress (the "stress non-responsive period") (Butte et al., 1973; Chatelain and Dupouy, 1981; Corbier and Roffi, 1978; Guillet et al., 1980; Haltmeyer et al., 1966; Jailer, 1950; Levine, 1970; Milkovic and Milkovic, 1969; Schapiro, 1962; Witek-Janusek, 1988); however, recent studies suggest that the concept of a stress non-responsive period may not be true, since neonates can repond to some forms of stress and show sensitivity to corticosteroid-suppression of ACTH levels similar to that seen in adults (Walker et al., 1990, 1986, 1991). Nevertheless it is clear that, early neonatally, the adrenal is quiescent and hyporesponsive (Levine, 1970; Walker et al., 1991). Sapolsky et al. (1986) have suggested that this period of adrenal hyporesponsiveness may serve a protective function, since glucocorticoids have been shown to have deleterious, catabolic effects on the brain (Bohn, 1984; De Kloet et al., 1988).

By the end of the second week, the HPA system matures. CBG and basal corticosterone levels increase, the magnitude of the adrenal response to stress increases, and the circadian rhythm emerges (De Kloet et al., 1988). During early postnatal development, glucocorticoid receptor levels in brain and liver increase (Kalinyak et al., 1989; Meaney et al., 1985; Sakly and Koch, 1981), in parallel with the increase in basal corticosterone titers. Some data suggests that glucocorticoid receptors during the early postnatal period are insensitive

to the down-regulation which occurs in adults (Kalinyak et al., 1989; Meaney et al., 1985).

To investigate the possible function of multiple GR promoters, I initiated a preliminary survey of the distribution of transcript classes in rat tissues. Since the early postnatal insensitivity of rat brain and liver glucocorticoid receptors might be explained by developmental differences in glucocorticoid receptor promoter usage, I used an RNase protection assay to compare brain and liver RNA from neonatal and adult rats that had been treated with dexamethasone or saline control. Since the preliminary survey showed differences in transcript distribution in lymphoid tissues, I also included thymus and spleen RNA in the hormone-treatment study. In all but the lymphoid cells transcripts A and B were the most abundant, and were not significantly affected by hormone treatment. In lymphoid cells, however, there were higher levels of a novel transcript, detected by the RNase protection assay but otherwise uncharacterized, that was positively regulated by dexamethasone in thymus cells. Together these results suggest that there may be at least two independently regulated promoters, one positively autoregulated in T-cells, and another expressed ubiquitously but not positively autoregulated.

Results

Transcripts A and B are the most abundant in J2.17 and XC cell RNA

In order to determine the relative and absolute abundance of the different transcripts in different cells or tissues, I used an RNase-protection assay. Since transcript B appeared to be most abundant in the HTC cell cDNA

libraries studied, I used an antisense probe to transcript B. As shown in Fig. 3.1, this probe (NMα) should yield three protected fragments, α, β, and γ[7] (from transcripts A, B, and C, respectively). Thus, if any transcript other than B (band β) were strongly expressed, the assay would show an increase in the levels of rare transcripts (represented by α, γ, "others") at the expense of B. In addition, since the sequence of B is known, is possible to quantitate the absolute levels of each protected fragment. As an internal control, and to quantitate the level of total receptor mRNA, I also included a probe (AccIα) for a portion of exon 7. Sense RNAs transcribed from cDNA clones (from classes A, B and C) were used as positive controls for the protection pattern of each transcript class. As Fig. 3.2 shows, the assay did indeed discriminate the three transcript classes, A, B and C (bands α, β and γ). Two additional bands, τ and δ, are also apparent in the cellular RNA, but do not match the control fragments. In order to match these additional bands to specific transcript classes, I repeated the experiment using additional positive controls from the PCR clones (classes D-G). As shown in Fig. 3.3, the protected fragments from the control transcripts (A-G) co-migrated with bands α, β, γ, δ, and τ as follows: α = A, D; β = B, γ = C; δ = none; τ = E, F, G (the latter transcripts also showed some protection in band A). Thus, though this assay does not distinguish each individual transcript (except for B), it does distinguish transcript B from the others and is likely to be more sensitive to changes in a single non-B transcript than measurements of B of total transcript levels.

Quantitation of the protected fragments from the experiment shown in Fig. 3.2 (see table 3.1) shows that transcript B is the most abundant (about 60%), and that total receptor mRNA is about 0.001-0.002% of total cellular

[7] Fragments are labelled with greek letters to distinguish them from transcript classes, since more than one transcript class can yield a single protected fragment.

RNA (0.1-0.2 fmol / 50 μg cellular RNA), or about 1-3 transcripts per cell (assuming 10^{-6} μg RNA per cell). The relative RNA levels represented by the protected fragments ($\beta > \alpha \geq \tau \approx \delta$) agreed well with the frequencies with which the corresponding transcripts were cloned from J2.17 RNA (B > (A + D) > (E + F + G) > C).

The pattern of glucocorticoid receptor transcript expression varies by tissue type

In order to determine whether there might be any tissue-specific expression of the various 5' leader mRNAs, I purified total cellular RNA from nine tissues of an adult male Sprague-Dewly rat and analyzed this RNA by the RNase protection assay described above. As a control I also analyzed RNA from J2.17 and XC cells. The raw data from one RNase protection experiment are shown in Fig. 3.4 and summarized in Fig. 3.5 (this assay was repeated on the same RNA with similar results). Most of the tissues show a pattern of 5'UTR usage similar to that in J2.17 and XC cells, with the exception of the thymus and spleen. In these two tissues both the relative and absolute transcript levels represented by band τ are higher than in the other tissues[8] (about 41% and 35% of the total in thymus and spleen, vs. 11-21% in the other tissues), while the levels of transcript B are lower in the lymphoid tissues (26% and 37% in thymus and spleen vs. 48-63% in the others). Band α varied from 26% - 38% of the total in the tissues surveyed, but didn't appear to show any tissue specificity. Bands γ and δ were faint (the least intense in all the

Figure 3.1. Expected RNAse protection results. Probe NMα is an (α-^{32}P) CTP

[8] Except testes. RNA levels were too low in this tissue to estimate relative transctipt levels reliably.

labelled antisense RNA to transcript B. The three classes of pAR cDNA clones (A, B and C) suggested that NMα would hybridize to three classes of mRNA, yielding three different protected fragments after digestion with RNase. Transcript B would yield nearly full-length protection (except for vector sequences) of the B-specific probe, transcript A would hybridize only to the exon2 portion of the probe, and transcript C would hybridize to the exon 2 portion plus the six nucleotides upstream that match transcript B. Previously uncharacterized transcripts would be either indistinguishable from A, or would yield slightly larger protected fragments corresponding to any upstream contiguity with B (similar to transcript C). By quantifying the results, I would be able to estimate the relative levels of transcripts B, C, A(+others), and any novel transcripts.

Figure 3.1

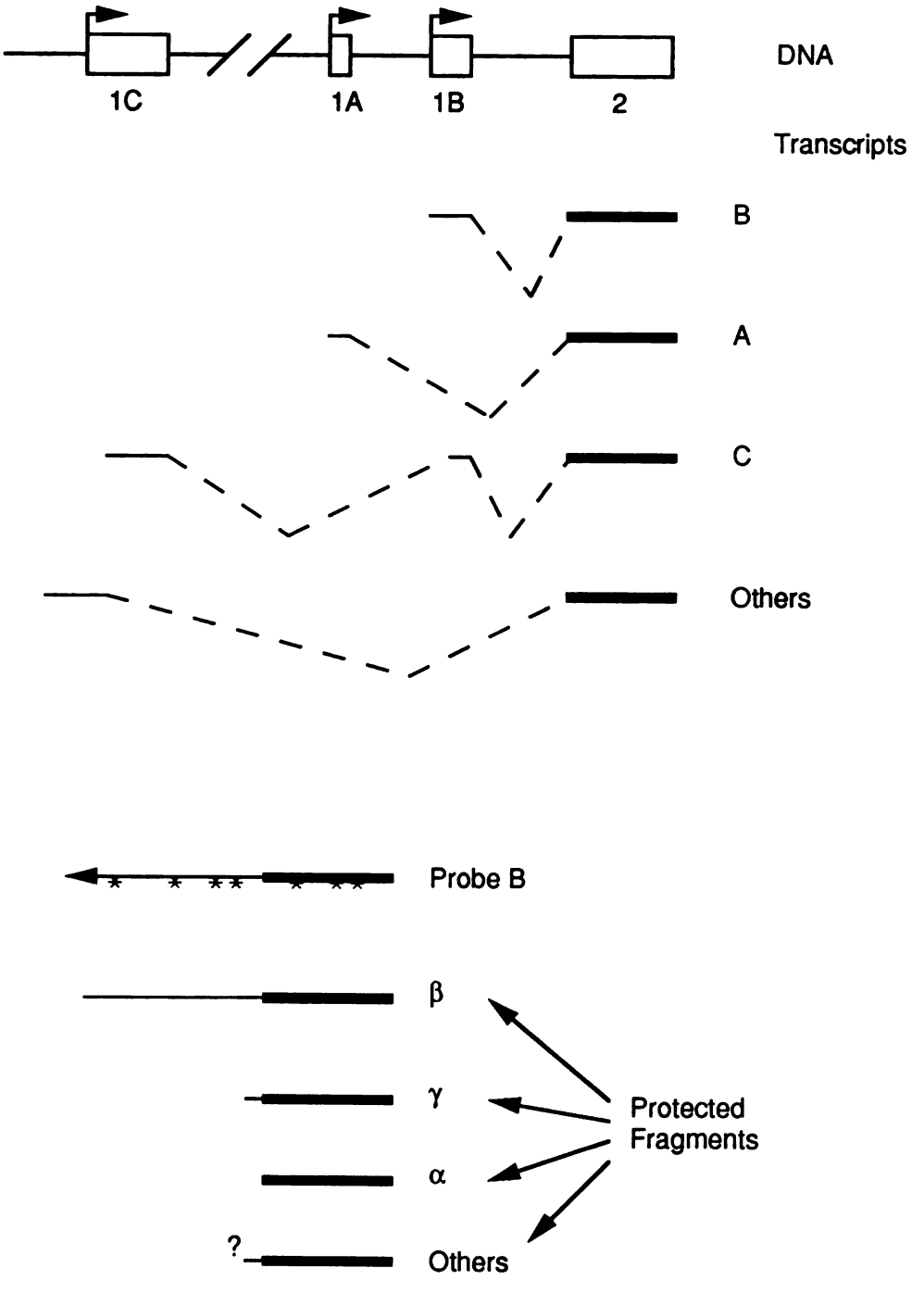

DNA
1C
1A
1B
2
Transcripts
B
A
C
Others
Probe B
β
γ
α
?
Others
Protected
Fragments

Figure 3.2. RNAse protection with probe NMα: cDNA A-C. Fifty micrograms of total cellular RNA (J2.17, XC, EDR3), 4μg poly-A selected RNA plus tRNA carrier (J2.17 pA+), in vitro transcribed RNA from cDNA clones plus tRNA carrier (A - C), or tRNA carrier alone (-) was hybridized to 1 fmol of the uniformly labeled 5' GR RNA probe NMα, and then digested with RNaseA and T1, as described in Materials and Methods. In the lanes indicated with a + at the bottom, the RNA was mixed with in vitro transcribed control RNA from the 3' end of receptor (PMσ) and hybridized to 1 fmol of the exon 7-specific probe AccIα in addition to NMα. The 5', NMα-specific bands are labelled α, β, γ, δ, and τ. There are two AccIα-specific bands, one to the 3' end of receptor mRNA (E7), and one to the control RNA (PMσ). Mkr is HpaI+MspII-digested, end-labelled pBR325 DNA. The probes (left) were loaded .1 fmol each. Full-length probe AccIα is not visible because the figure was cropped, but the probe appeared in the autoradiographic image as a discrete band of the expected size.

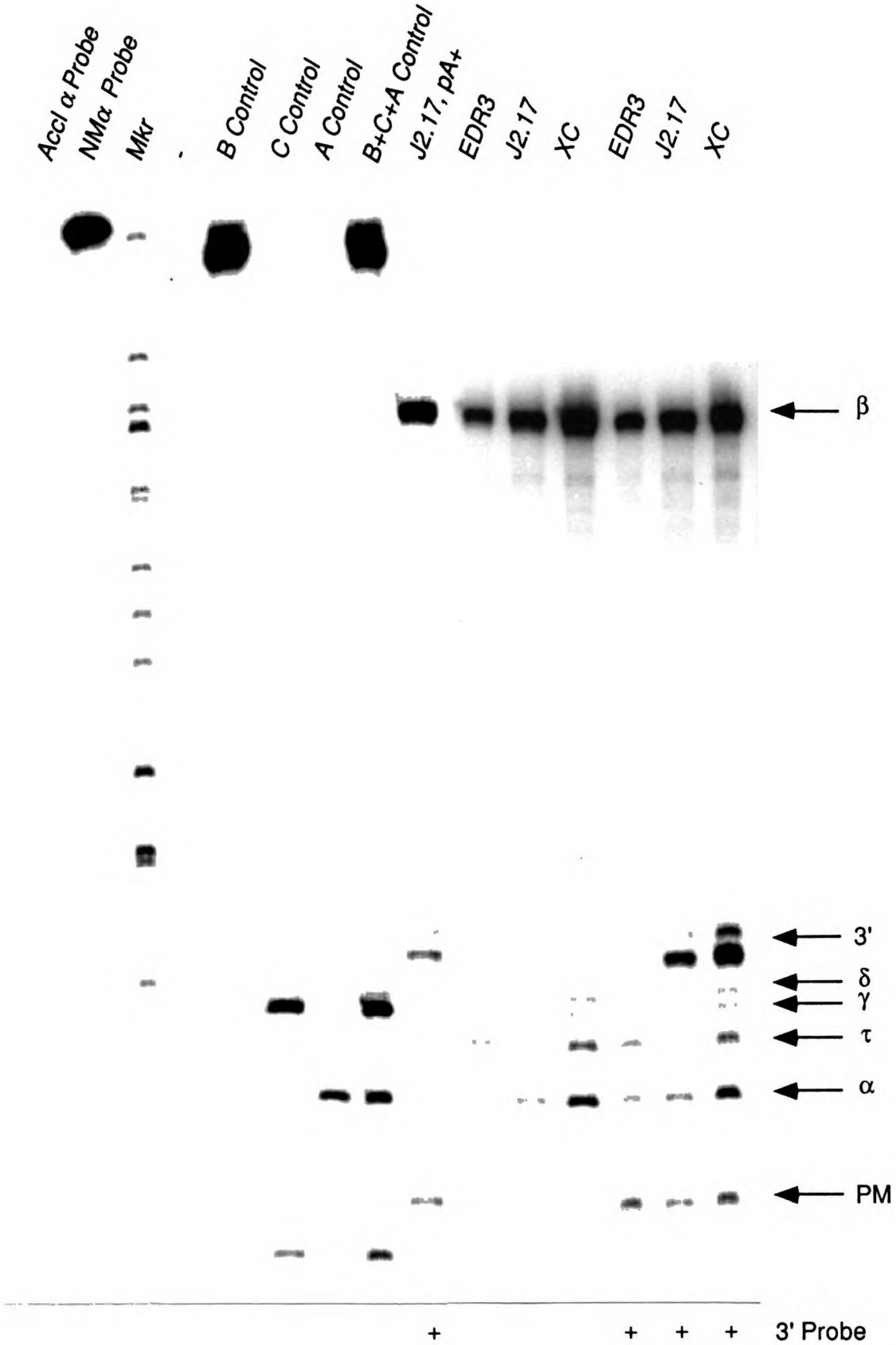

AccI α Probe
NMα Probe
Mkr
-
B Control
C Control
A Control
B+C+A Control
J2.17, pA+
EDR3
J2.17
XC
EDR3
J2.17
XC
β
3'
δ
γ
τ
α
PM
+
+
+
+
3' Probe

Table 3.1. RNA Levels of Receptor Transcript Groups (Quantitation of Fig. 3.2)[9]

Source	Transcript Group (Protected Fragment)				Total GR mRNA	
	α	τ	γ+δ	β	SUM	3'
J2.17[10]	22 (17)	11 (9)	11 (9)	81 (65)	125 (100)	105 (84)
XC	33 (17)	29 (15)	22 (12)	107 (56)	191 (100)	190 (99)
EDR3	15 (16)	15 (16)	11 (12)	53 (57)	94 (100)	27 (28)
J2.17, pA+[11]	12 (17)	4 (6)	nd na	56 (77)	72 (100)	69 (95)

[9]The quantitation method is described in the legend to Fig. 3.5

[10]amol / 50μg total cellular RNA. Numbers in parentheses are the % of total GR mRNA estimated by the sum of levels of the protected fragments.

[11]amol / 4μg poly-A selected RNA. nd, not determined; na, not applicable.

Figure 3.3. RNase protection with probe NMα: cDNA A-G. See legend to Fig. 3.2. A-G are cDNA clones. F3 and F4 are two different cDNA clones of class F that differed by one nucleotide close to the exon 2 junction. Mkr is HpaI+MspII-digested, end-labelled pBR322 DNA.

Figure 3.3

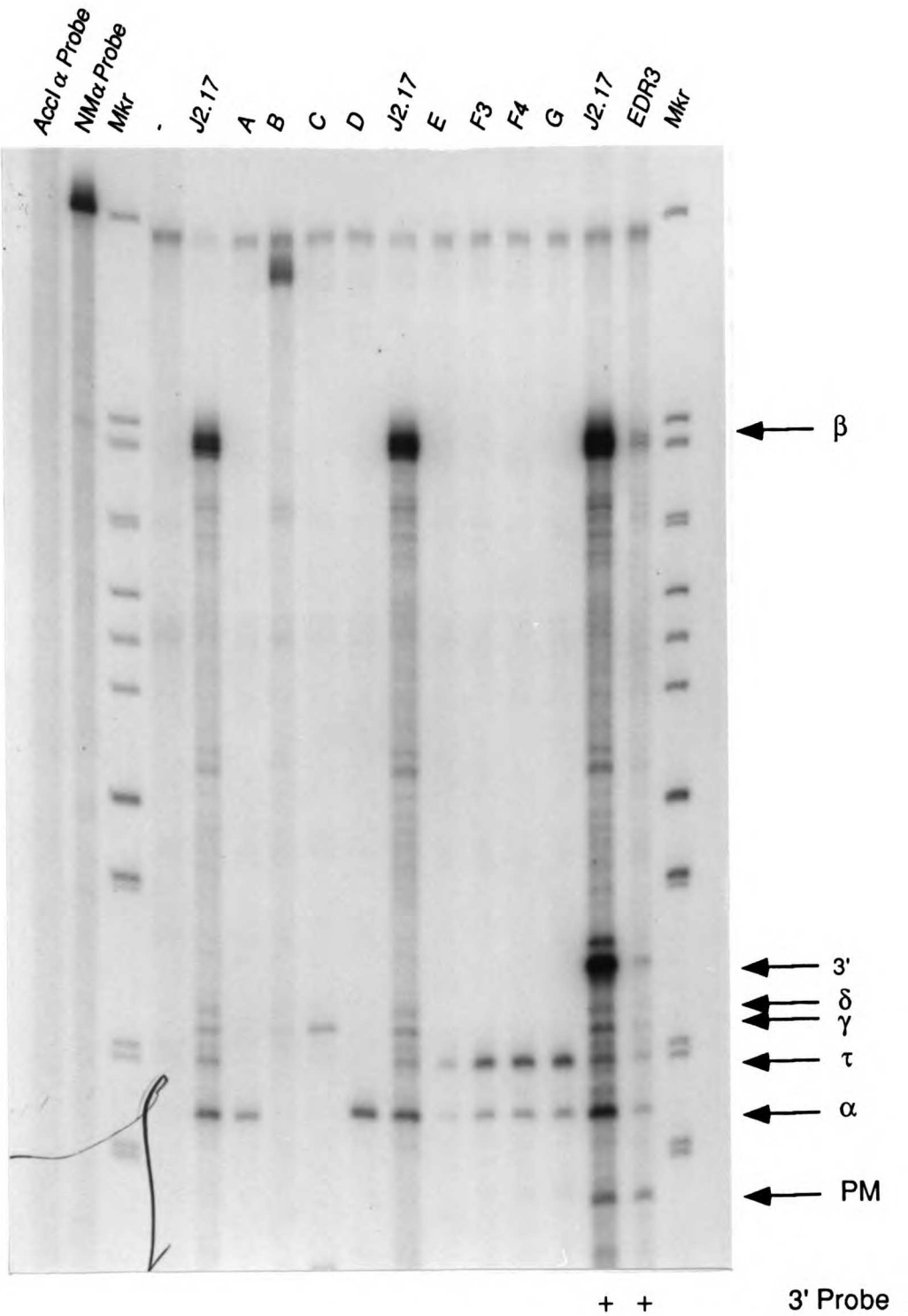

Acc I α Probe
NM α Probe
Mkr
-
J2.17
A
B
C
D
J2.17
E
F3
F4
G
J2.17
EDR3
Mkr
β
3'
δ
γ
τ
α
PM
+ +
3' Probe

Figure 3.4. RNase protection of RNA from rat tissues. Total RNA from the indicated tissues of a male S.D. rat and three cell lines was extracted and 50μg hybridized to 2 fmol each of the probes NMα and AccIα in an RNase protection assay, as described in Figs. 3.1 and 3.2. This figure shows an autoradiogram of the probe fragments separated on a 4.5% acrylamide sequencing gel. See Fig. 3.5 for quantitation.This RNase protection experiment was repeated once with the same RNA and produced similar results.

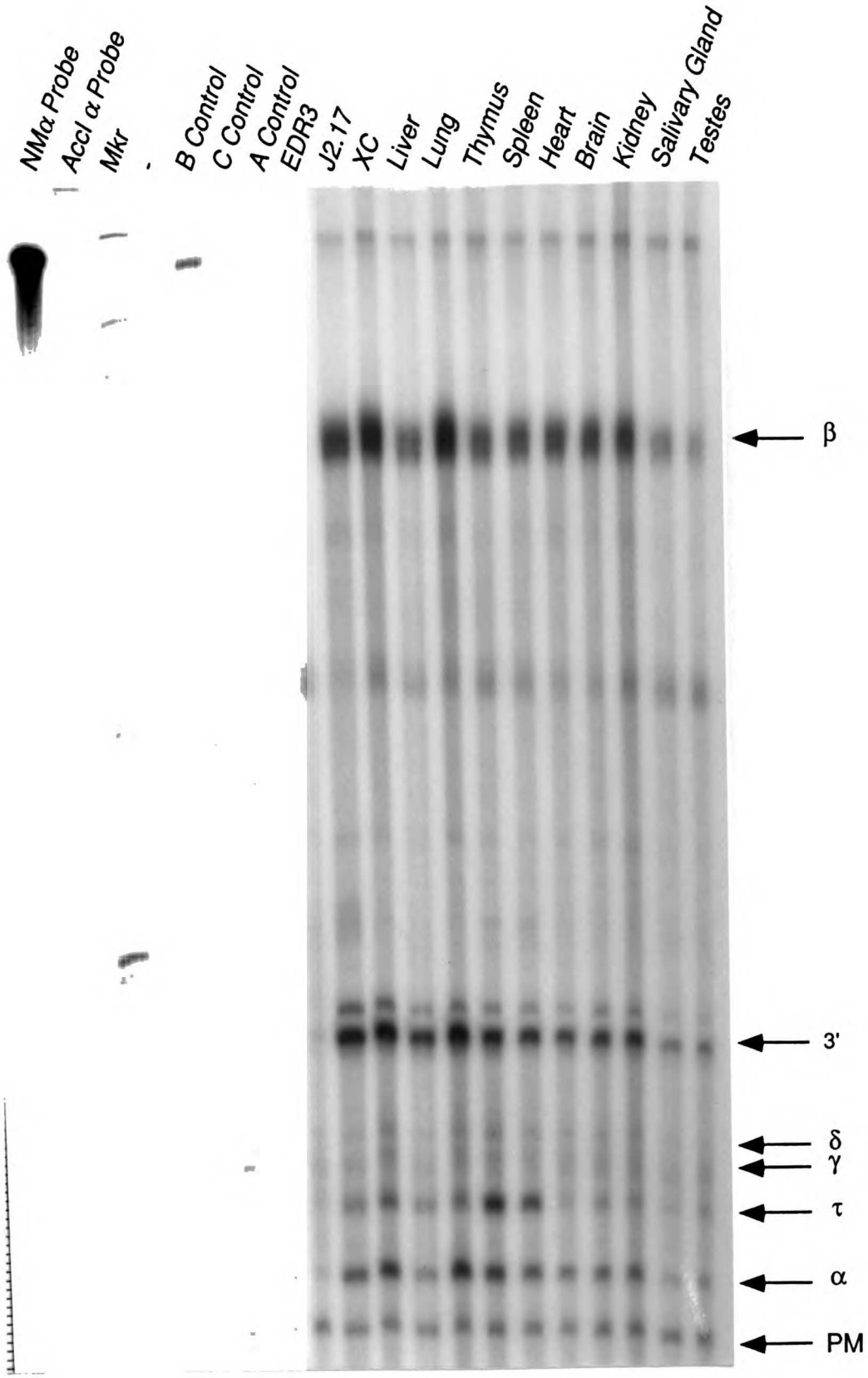
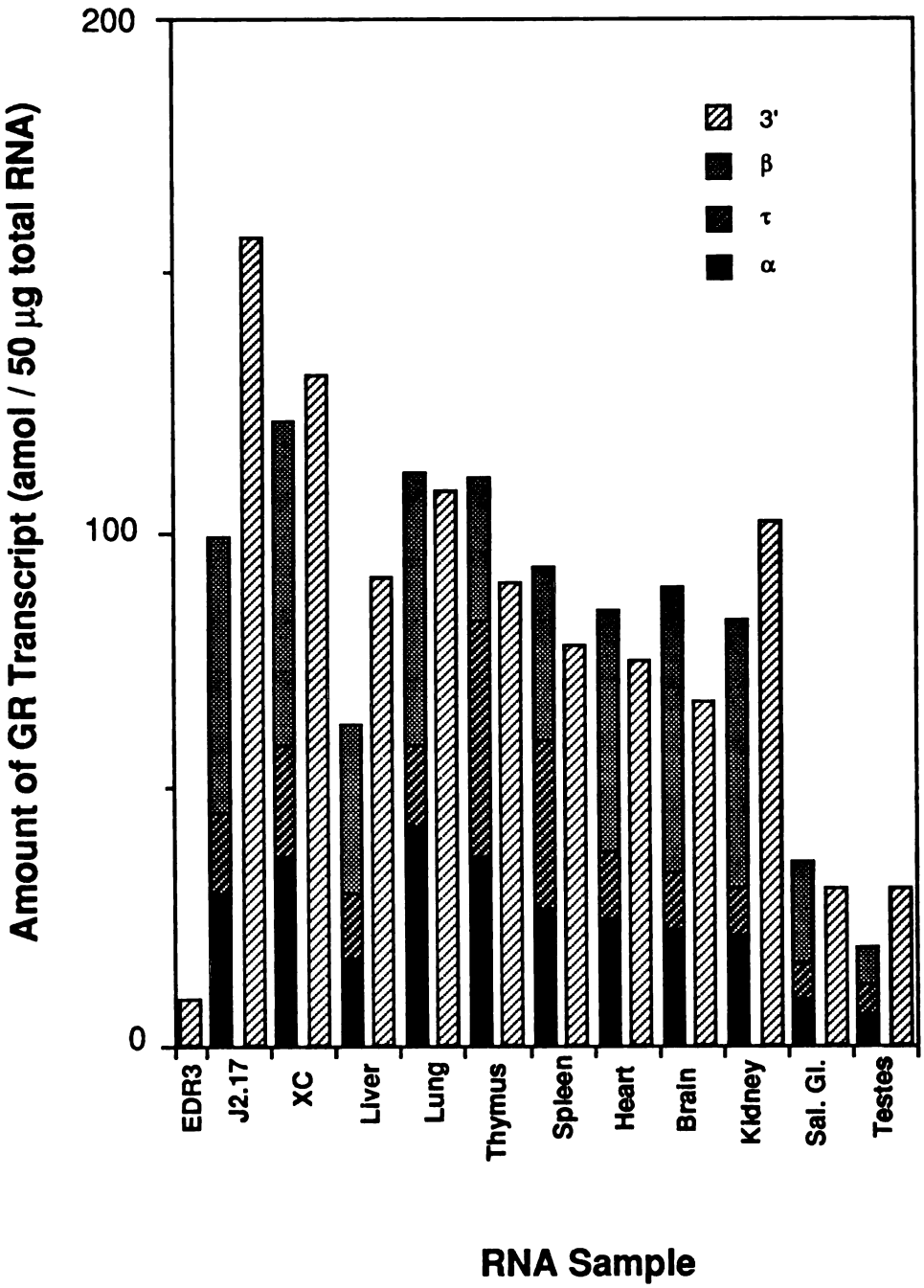

NMα Probe
Accl α Probe
Mkr
B Control
C Control
A Control
EDR3
J2.17
XC
Liver
Lung
Thymus
Spleen
Heart
Brain
Kidney
Salivary Gland
Testes
β
3'
δ
γ
τ
α
PM

Figure 3.5. Quantitation of 5' receptor RNAs from different rat tissues. The bands in the autoradiogram shown in Fig. 3.4 were quantitated by scanning with a laser densitometer. See Materials and Methods for details. Briefly, scanner units (Au x mm) were converted to RNA levels (amol) by normalizing each band intensity to the expected radioactivity of that band (based on its sequence) and the level of the control band PMσ (= 50 amol). The different tissues express various levels of receptor mRNA, but all except thymus and spleen show the same relative levels of 5' protected fragments. The thymus and spleen appear to express higher levels of band τ receptor RNA transcripts, and lower levels of transcript B (band β).

Figure 3.5
5' Transcript Variation in Rat Tissues



lanes) and close to background. These bands could not be quantified reliably, and so are not included in the calculations of total transcript levels in this experiment. Bands α, β, and τ appeared to account for most of the receptor RNA since the total receptor RNA levels estimated by the sum of these bands (α + β + τ) agreed well with the total levels estimated by the 3' band.

The levels and the pattern of receptor mRNA expression in specific tissues change on exposure to dexamethasone

Experiments by others had suggested that autoregulation of glucocorticoid receptor in brain and liver did not appear in rats until 2-3 weeks after birth (Kalinyak et al., 1989). There have also been reports that some lymphoid and myeloma cells respond to glucocorticoids by increasing their receptor mRNA and protein levels (Antakly et al., 1989; Eisen et al., 1988; Gomi et al., 1990), rather than by decreasing receptor mRNA as seems to occur in other cell types investigated. In order to determine whether any of this variation correlates with the mRNA 5'leaders (which might indicate differential promoter activities), I analyzed RNA from neonatal (2d) and adult (28d) rats injected with dexamethasone (1 mg / kg, s.c.) or vehicle (saline) 6 h prior to sacrifice; each sample represents RNA pooled from 3-6 rats (see legend to Table 3.2). This timing was experimentally convenient and in the range (4 - 24 h) for maximal responses to hormone injection reported by Kalinyak et al. (1989). RNA was isolated from liver, brain, thymus and spleen of these rats, and the RNA analyzed by RNase protection as above. The raw data (estimates of RNA levels for each band) are shown in Table 3.2.

I wanted to know if transcript levels varied with respect to tissue, age, or dexamethasone treatment. It was difficult to analyze the absolute transcript

levels reliably with this data set because the variances within data cells differed by over a factor of a million, violating the constant variance assumption (that all the within-cell variances are approximately equal)[12]. There was also variability between measurements of the same sample by as much as two-fold between separate experiments. However, normalizing the data to the sum of the bands in each measurement reduced both the within-cell variances and the spread of the variances, enabling more reliable parametric analysis. Since the means of the sums ($\alpha + \beta + \tau$) were not significantly different ($\alpha < 0.05$) from the total estimated by the 3' band (3' Total), the estimated relative levels reported should be close to the true proportions.

By analyzing relative transcript levels, I could determine how the experimental factors (age, dex, tissue) affected the pattern of transcript expression. Specifically, I could test (1) whether thymus and spleen expressed a different pattern of transcripts from brain and liver (suggesting the existence of tissue-specific promoters); (2) whether dexamethasone affected these patterns (suggesting tissue- and/or promoter-specific regulation); and (3) whether age affected either the tissue-specific pattern or the glucocorticoid regulation of that pattern (suggesting differential regulation of receptor transcripts during postnatal development).

I analyzed the effects of tissue, age, and hormone treatment (DEX) on each set of relative transcript levels (%α, %β, %τ) by analysis of variance (ANOVA) using the general linear models (GLM) procedure (S.A.S. statistical software, release 5.18), as described in Materials and Methods and in the

---

[12] Within cell variance is the standard deviation squared within a data cell, where a cell is the set of observations (transcript levels) for a given set of conditions (age, tissue, hormone). Constant variance is an assumption in analysis of variance.

legend to Fig. 3.6. The results of this analysis are summarized in Fig. 3.6 (A-C).

For all three groups, there was a significant effect by tissue ($p \leq .0001$ for each band). That is, after accounting for the other factors, the relative transcript levels were different in the four tissues. Multiple comparisons (multiple t-tests, controlling the comparison-wide error rate) among tissue groups for all three bands (Fig. 3.7) showed that in thymus and spleen, the relative levels of β were lower, and the relative levels of τ and α were higher than in liver or brain. There were no significant differences between brain and liver for any of the transcript groups. These results confirm the preliminary observation described in the previous section (Fig. 3.5) that the lymphoid tissues expressed a different pattern of receptor transcripts than the brain or liver. Most noticeably, RNA from the thymus contained the highest relative levels of group τ compared to RNA from brain, liver or spleen.

Dexamethasone also appeared to affect the distribution of transcripts among the β and τ bands in a tissue-specific manner. None of the analyses detected a significant hormone effect alone, but detected significant Dex x Tissue effects on both the τ and β bands (Fig. 3.6). This means that hormone did affect relative transcript levels (τ and β), but not the same way in each of the four tissues. In particular, the relative levels of τ increased and β decreased in the thymus of dexamethasone-treated rats compared to control rats (Fig. 3.8). There were no statistically detectable changes in relative transcript levels in brain, liver or spleen.

The analysis of variance detected significant effects of Age and Age x Tissue on the relative levels of τ, but not on either α or β. The main Age effect was a general increase in the proportion of τ in adults compared to neonates ($\alpha < 0.02$). The interaction effect with tissue appeared to be due

Sample	FACTOR			amol RNA				GR/
Group	AGE	DRUG	TISSUE	Total(3')	α	β	τ	actin
DA2	Adult	Dex	Liver	67.97	6.26	66.14	6.31	0.866
DA3	"	"	"	64.17	5.61	53.50	5.81	0.968
DA1	"	"	"	77.34	10.75	40.86	11.21	nd
DA3	"	"	Thymus	128.47	69.99	73.75	127.70	2.396
DA1	"	"	"	94.52	43.04	24.01	63.57	1.812
DA2	"	"	"	157.33	70.71	48.78	96.91	nd
DA2	"	"	Spleen	190.79	67.28	107.49	37.20	nd
DA3	"	"	"	197.24	68.07	105.10	32.80	nd
DA3	"	"	Brain	124.05	22.29	85.30	10.28	0.787
DA1	"	"	"	82.17	12.39	57.55	4.37	0.661
DA2	"	"	"	94.58	24.32	52.21	11.30	nd
CA3	"	Control	Liver	95.94	11.91	62.03	11.43	1.564
CA1	"	"	"	108.42	16.79	81.14	15.88	1.435
CA2	"	"	"	101.70	20.51	59.22	17.21	nd
CA2	"	"	Thymus	124.04	33.39	48.67	39.03	0.704
CA3	"	"	"	118.88	54.41	53.54	66.22	nd
CA1	"	"	Spleen	272.98	82.57	133.46	78.40	nd
CA3	"	"	"	277.35	91.83	139.77	70.72	nd
CA3	"	"	Brain	75.66	13.42	55.80	5.85	0.885
CA2	"	"	"	128.51	25.00	78.58	11.08	1.170
CA1	"	"	"	115.41	31.64	54.42	13.09	nd
DN2	Neonate	Dex	Liver	59.54	6.06	58.37	5.69	0.759
DN1	"	"	"	65.08	12.32	41.52	10.21	1.619
DN1	"	"	Thymus	188.21	52.42	65.61	73.92	0.889
DN2	"	"	"	160.17	58.08	65.56	83.23	0.668
DN1	"	"	Spleen	196.90	66.11	110.82	41.61	nd
DN2	"	"	"	274.90	82.30	160.12	56.27	nd
DN2	"	"	Brain	43.53	11.30	28.65	4.82	0.124
DN1	"	"	"	36.48	7.23	31.60	2.63	0.152
CN2	"	Control	Liver	91.62	14.30	77.45	9.21	0.731
CN1	"	"	"	61.46	15.19	29.93	10.39	1.618
CN1	"	"	Thymus	115.70	41.96	65.95	36.84	0.306
CN2	"	"	"	131.03	35.15	82.74	31.20	0.414
CN1	"	"	Spleen	298.14	81.54	158.46	50.89	nd
CN2	"	"	"	244.76	73.84	147.15	43.07	nd
CN2	"	"	Brain	46.97	10.93	41.39	6.68	0.114
CN1	"	"	"	63.95	17.12	41.24	7.61	0.110

Table 3.2. Estimated rat GR mRNA levels (amol / 50μg RNA) of three transcript species, total GR mRNA, and actin-normalized total GR mRNA levels *in vivo*. RNA levels were estimated by RNase protection, as described in Figs. 3.2 and 3.5. GR/actin was measured by scanning autoradiograms of RNA slot-blots hybridized to a GR or βactin probe. Each sample group represents RNA obtained from 3-4 28d-old or 5-6 2d-old male Sprague-Dewley rats.

Figure 3.6. Three-way analysis of variance, with interactions. Relative transcript levels from data in Table 3.2 were analyzed for bands α (A), β (B), and τ (C). Relative levels for each band (AREL, BREL, TREL) were determined by dividing the each level by the sum (α + β + τ). Analysis of variance was by the general linear models (GLM) procedure using the SAS v 5.18 statistical analysis software. For a detailed description of the method, see (Glantz and Slinker, 1990) and the SAS manual. This procedure is essentially a step-wise least squares linear regression procedure. The data (transcript levels; dependent variable) were fit to the model: transcript level = a*AGE + b*DEX + c*TISSUE + d*AGE*DEX + e *AGE*TISSUE + f*DEX*TISSUE + g*AGE*TISSUE*DEX, where a-g are real coefficients and AGE, DEX and TISSUE are discrete independent variables. The program automatically recodes the independent variables into (DF -1) sets of orthogonal dummy variables. The regression model defined each relative transcript level (AREL, BREL, TREL) as a linear function of AGE (Neonate or Adult), DEX (Control or Dexamethasone), TISSUE (Thymus, Spleen, Liver or Brain), and their interactions (all possible combinations of products). A given independent variable has a significant effect on the regression model if the probability of obtaining an F value larger than the value of F associated with that variable's sums of squares is small[13]. Thus, a large F (probability > F is small) for the entire ANOVA (the top half of Fig. 3.6) signifies that the model accounts for a significant portion of the overall variance in the data. The F values associated

---

[13] $F = MS_{within}/MS_{error}$, where MS = SS/DF. MS, mean square; SS, sum of squares (squared deviations from the mean); DF, degrees of freedom. In the overall analysis of variance, MS_{within} is the variance explained by the model, and MS_{error} is the residual (error) variance. F is thus the amount of variance explained by the model relative to the residual (unexplained) variance. The significance of the value of F is given by the probability of obtaining a larger F, where (PR > F) ≤ .05 is generally considered significant.

with each term in the model (lower half of Fig. 3.6) signify the effect of that variable on the overall variance[14]. For example, in Fig. 3.6A, $F_{ANOVA} = 3.80$ and (PR > F) = 0.003. Thus, there is a significant effect by age, tissue, hormone, and all their interactions, on relative α transcript levels. Most of this effect is from the TISSUE term, since the probability of exceeding the F value associated with its typeIII sums of squares is small (≤0.0001). This analysis (Fig. 3.6A,B,C) detected significant effects by at least one factor for each of the transcript groups (bands) α, β, and τ.

The F value associated with each term in the model represents the effect that term has on the model. In this analysis, since the model explains the variances for α, β and τ, then each term with a low probability {(PR > F) ≤ .05} has a significant effect on the associated relative transcript level. $F = MS_{within}/MS_{error}$, where MS = SS/DF. MS, mean square; SS, sum of squares (squared deviations from the mean); DF, degrees of freedom. In the overall analysis of variance, MS_{within} is the variance explained by the model, and MS_{error} is the residual (error) variance. F is thus the amount of variance explained by the model relative to the residual (unexplained) variance. In the bottom half of each summary, MS_{within} is mean of the type III sums of squares for that variable. The type III SS is the change in the model sums of squares that occurs when that variable is entered last in a stepwise regression.

[14] In this case, MS_{within} is mean of the type III sums of squares for that variable. The type III SS is the change in the model sums of squares that occurs when that variable is entered last in a stepwise regression.

Figure 3.6 A

THREE-WAY ANOVA: TRANSCRIPT α/SUM

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE: AREL

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PR > F	R-SQUARE	C.V.
MODEL	15	1409.6	93.97	3.80	0.0027	0.7306	21.16
ERROR	21	519.8	24.75		ROOT MSE		AREL MEAN
CORRECTED TOTAL	36	1929.4			4.975		23.52

SOURCE	DF	TYPE III SS	F VALUE	PR > F
AGE	1	3.75	0.15	0.7010
DEX	1	12.03	0.49	0.4933
AGE*DEX	1	0.30	0.01	0.9128
TISSUE	3	1126.78	15.17	0.0001 *
AGE*TISSUE	3	56.65	0.76	0.5275
DEX*TISSUE	3	91.98	1.24	0.3207
AGE*DEX*TISSUE	3	5.76	0.08	0.9714

Figure 3.6 B

THREE-WAY ANOVA: TRANSCRIPT β/SUM

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE: BREL

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PR > F	R-SQUARE	C.V.
MODEL	15	9649.2	643.28	10.02	0.0001	0.8775	14.05
ERROR	21	1347.7	64.17		ROOT MSE		TREL MEAN
CORRECTED TOTAL	36	10996.9			8.0109		57.00

SOURCE	DF	TYPE III SS	F VALUE	PR > F
AGE	1	130.27	2.03	0.1689
DEX	1	1.34	0.02	0.8863
AGE*DEX	1	23.78	0.37	0.5493
TISSUE	3	7464.76	38.77	0.0001 *
AGE*TISSUE	3	316.71	1.65	0.2093
DEX*TISSUE	3	724.19	3.76	0.0263 *
AGE*DEX*TISSUE	3	24.86	0.13	0.9417

Figure 3.6 C

THREE-WAY ANOVA: TRANSCRIPT τ/SUM

GENERAL LINEAR MODELS PROCEDURE

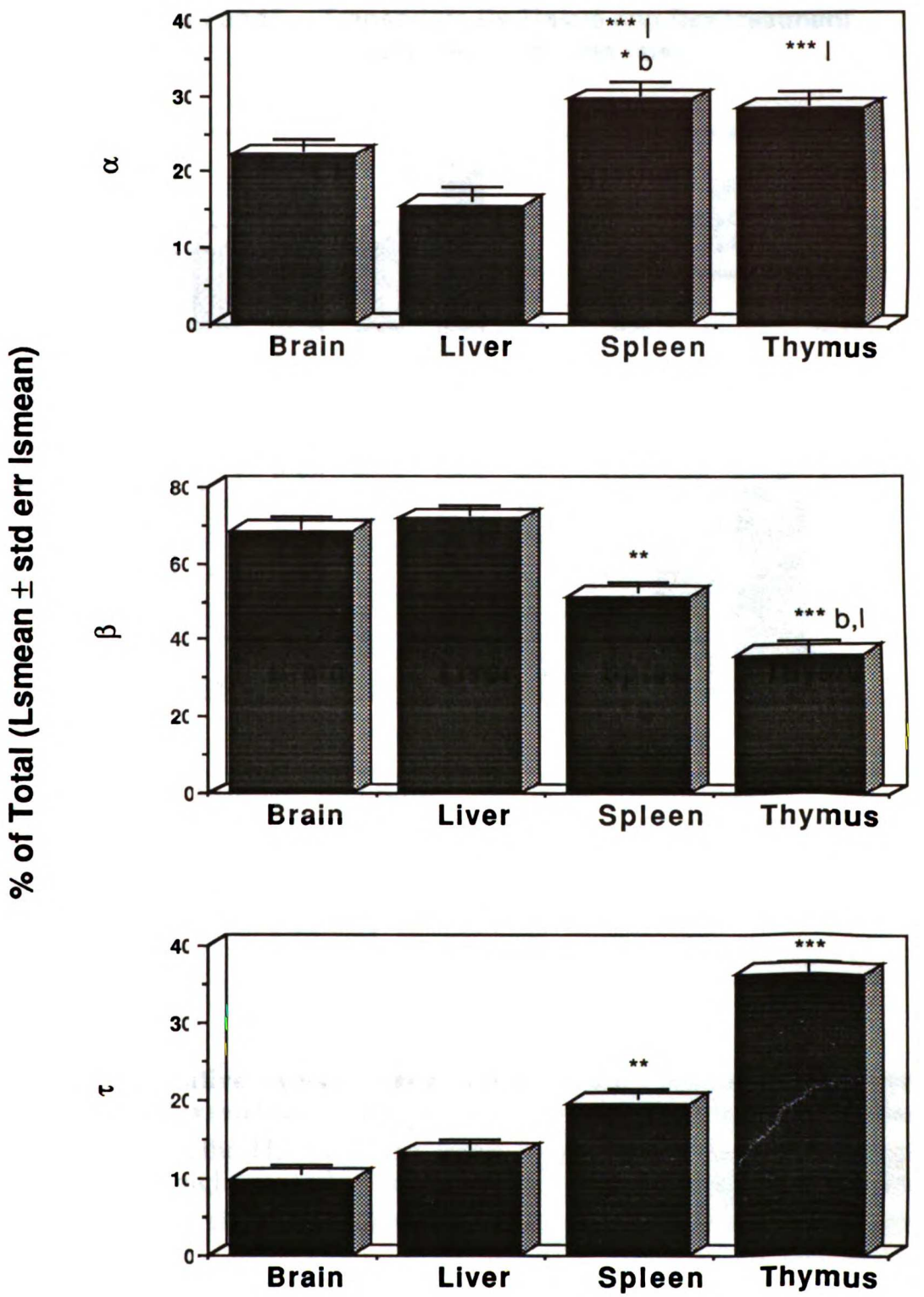
DEPENDENT VARIABLE: TREL

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PR > F	R-SQUARE	C.V.
MODEL	15	5020.6	334.71	28.94	0.0001	0.9539	17.47
ERROR	21	242.8	11.56		ROOT MSE		TREL MEAN
CORRECTED TOTAL	36	5263.5			3.401		19.47

SOURCE	DF	TYPE III SS	F VALUE	PR > F
AGE	1	85.86	7.43	0.0127 *
DEX	1	19.52	1.69	0.2080
AGE*DEX	1	30.12	2.60	0.1215
TISSUE	3	3699.97	106.65	0.0001 *
AGE*TISSUE	3	144.80	4.17	0.0182 *
DEX*TISSUE	3	456.80	13.17	0.0001 *
AGE*DEX*TISSUE	3	42.51	1.23	0.3253

Figure 3.7. Multiple comparisons of relative α, β and τ bands among tissue groups, after accounting for age and dex affects. Least squares means (lsmeans) and standard errors of lsmeans are from the TISSUE terms for each transcript group in the general linear models analysis of variance (Fig. 3.6). Significant differences are based on the probability of a false positive (α_T) for the family of six individual t-test comparisons (Sidak's inequality (); see Materials and Methods). Asterix (*α_T < .05; **α_T < .005; ***α_T < .001) show the largest probability among comparisons with all the other tissues, except where indicated (l = vs. liver, b = vs. brain).

Figure 3.7


% of Total (Lsmean ± std err lsmean)
α
40
30
20
10
0
*** l
* b
*** l
Brain
Liver
Spleen
Thymus
β
80
60
40
20
0
**
*** b,l
Brain
Liver
Spleen
Thymus
τ
40
30
20
10
0
**
***
Brain
Liver
Spleen
Thymus

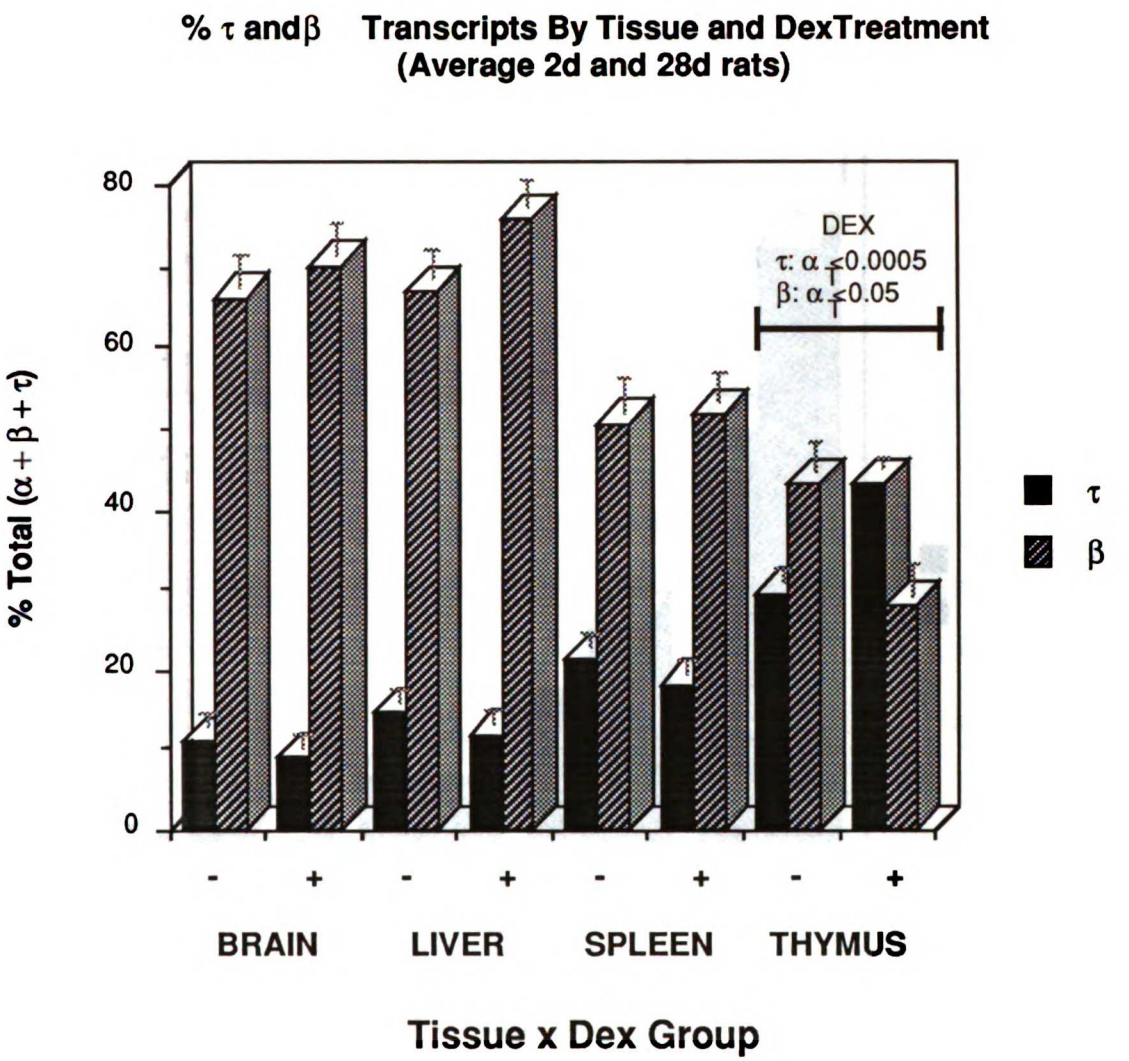


Figure 3.8. Relative levels of band τ increase, and β decrease in the thymus in response to dexamethasone treatment. Least squares means and standard errors are from the TISSUE x DEX terms for the β and τ transcript groups. For the family of eight control vs. dexamethasone pairs within each transcript group, the only significant differences ($\alpha_T < .05$) were in the thymus group.

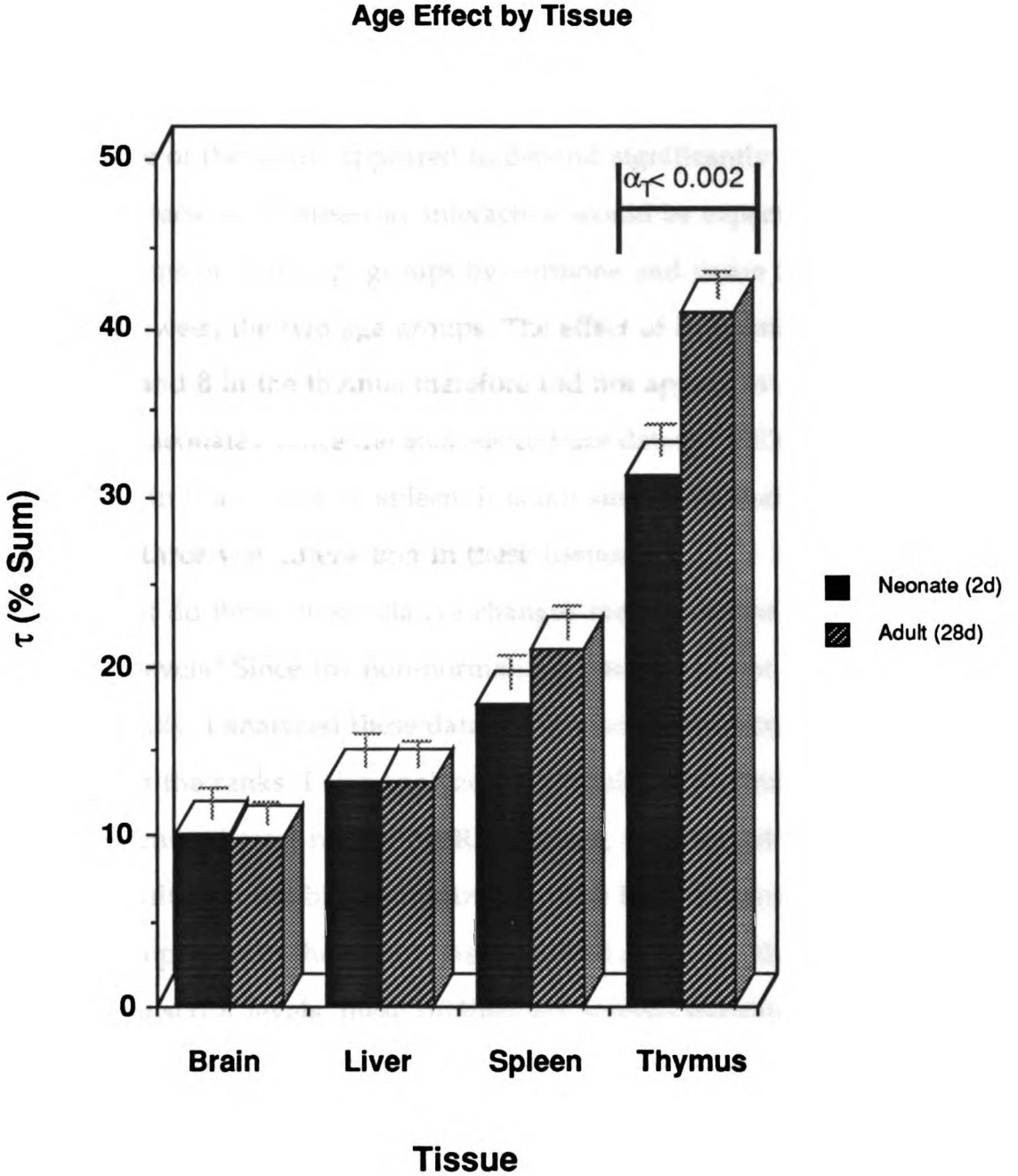


Figure 3.9. Relative levels of band τ are greater in adult than in neonatal rats in the thymus. Least squares means and standard errors are from the TISSUE x AGE term for transcript group τ. For the family of four age (adult vs. neonatal) pairs within each tissue group, the only significant difference (α_T < .05) was in the thymus group.

predominantly to a difference between the adult and neonatal groups in the thymus (Fig. 3.9).

None of the bands appeared to depend significantly on an interaction of all three factors. A three-way interaction would be expected if the changes in proportions of transcript groups by hormone and tissue (DEX x TISSUE) differed between the two age groups. The effect of hormone on the relative levels of τ and β in the thymus therefore did not appear to differ between adults and neonates. Since the analysis did not detect a DEX x TISSUE interaction in brain, liver or spleen, it is not surprising that there was no detectable three-way interaction in these tissues.

What do these these relative changes mean in terms of absolute transcript levels? Since the non-normalized data could not be analyzed parametrically, I analyzed these data using a non-parametric analysis of variance on the ranks. I also analyzed the results of preliminary measurements of total receptor mRNA levels, normalized to βactin mRNA, using quantitative dot-blot hybridization. The latter measurements did not include samples from the spleen. As discussed above for the analyses of relative transcript levels, these preliminary screens indicate the most significant effects that can be detected with the data. However, it is not possible to rule out effects that could not be discerned because they were too small relative to their associated variances and degrees of freedom. Because of the small sample sizes, only the largest effects were significant; more data would likely reveal additional effects in the model.

I was most interested in comparisons within the Dex x Tissue interaction: were there significant differences between +/- hormone pairs within tissue groups? Rank-variance analysis of the dot-blot data of actin-normalized receptor mRNA (Table 3.2) detected an increase in thymus

receptor mRNA ($\alpha_T < .02$[15]), but no significant differences ($\alpha_T \not< .05$) in brain or liver (spleen RNA was not measured). For the RNase protection data (table 3.2), the analysis of ranks detected significant differences in the Dex x Tissue interaction for bands τ ($p = .0015$) and α ($p = .0068$). The specific comparisons indicated that levels of τ were higher in the thymus ($\alpha_T \approx .01$) and that levels of α were lower in the liver ($\alpha_T \approx .0025$) of the hormone-treated groups compared to the controls. No significant three-way interactions (Age x Dex x Tissue) were detected.

The increased total glucocorticoid receptor mRNA levels in the thymus of hormone-treated rats compared to controls is consistent with results of studies of the human T-cell line CEM-C7 showing a similar increase in total receptor mRNA in response to dexamethasone (Antakly et al., 1989; Eisen et al., 1988). Both the absolute and relative τ transcript levels were greater in the thymus dexamethasone groups compared to controls, whereas the relative levels of α and β decreased, suggesting that one or more transcripts in the τ band account for most if not all of the increase in total receptor mRNA levels.

The 5' UTR of transcript B decreases expression of a downstream GR-lacZ fusion protein

Two features of exon B that stand out are the presence of a 40aa-long open reading frame (ORF) entirely within the exon, and the existence of multiple potential RNA hairpin structures, one of which starts at the AUG of the B-ORF (Fig. 2.19). Some 5'-leader sequences with ORFs and with RNA

---

[15] α_T is the probability of a type I (false positive) error for the family of comparisons.

hairpins that have been described in other transcription units have been shown to regulate the translation of their downstream encoded proteins (reviewed in Kozak, 1986a, 1988). In order to determine whether the 5' B-leader was similarly capable of playing a role in the regulation of receptor translation, I made constructs fusing three different 5'mRNA leaders and first 131bp of exon 2 (containing the first two initiation AUGs of GR) to lacZ (Fig. 3.10). I measured expression of the GR-lacZ fusion protein, assayed by βgal activity, in constructs with either the B-leader (B), no B-leader (ΔB), or the tk leader (tk).

In the experiment shown in Fig. 3.11(A) the GR-lacZ plasmids were co-transfected with a CAT reference plasmid, to control for transfection efficiency, in either CV1b or HeLa cells. The CAT-normalized βgal activities were about 3-4 fold lower in the cells transfected with the B-leader construct than with the ΔB-leader construct. The positive control tk-leader construct produced about 2-fold more activity than the ΔB construct. These results suggested that the B-leader decreased the expression of the downstream GR-lacZ fusion protein, possibly by decreasing the translational efficiency of the mRNA. To control for mRNA accumulation, I repeated the above experiment (at one level of transfected plasmid) and measured βgal RNA as well as β-galactosidase activity. As Fig. 3.11(B) shows, the B-leader indeed decreased βgal activity when normalized to βgal RNA levels. The B-leader thus appears to decrease the translational efficiency of receptor.

Discussion

The cloning and mapping results discussed in the previous chapter suggested that receptor is transcribed from two or more promoters, yielding

Figure 3.10. GR-lacZ fusion constructs. Filled boxes represent coding sequences, open arrows initiation codons. The sequences shown were cloned into the expression vectors p65-VAL0 or p6R.Z, as described in Materials and Methods. **(A)** GR(B)-lacZ. **(B)** GR(ΔB)-lacZ. **(C)** Amino acid sequence of fusion. **(D)** tk leader-lacZ (p65-VAL0).

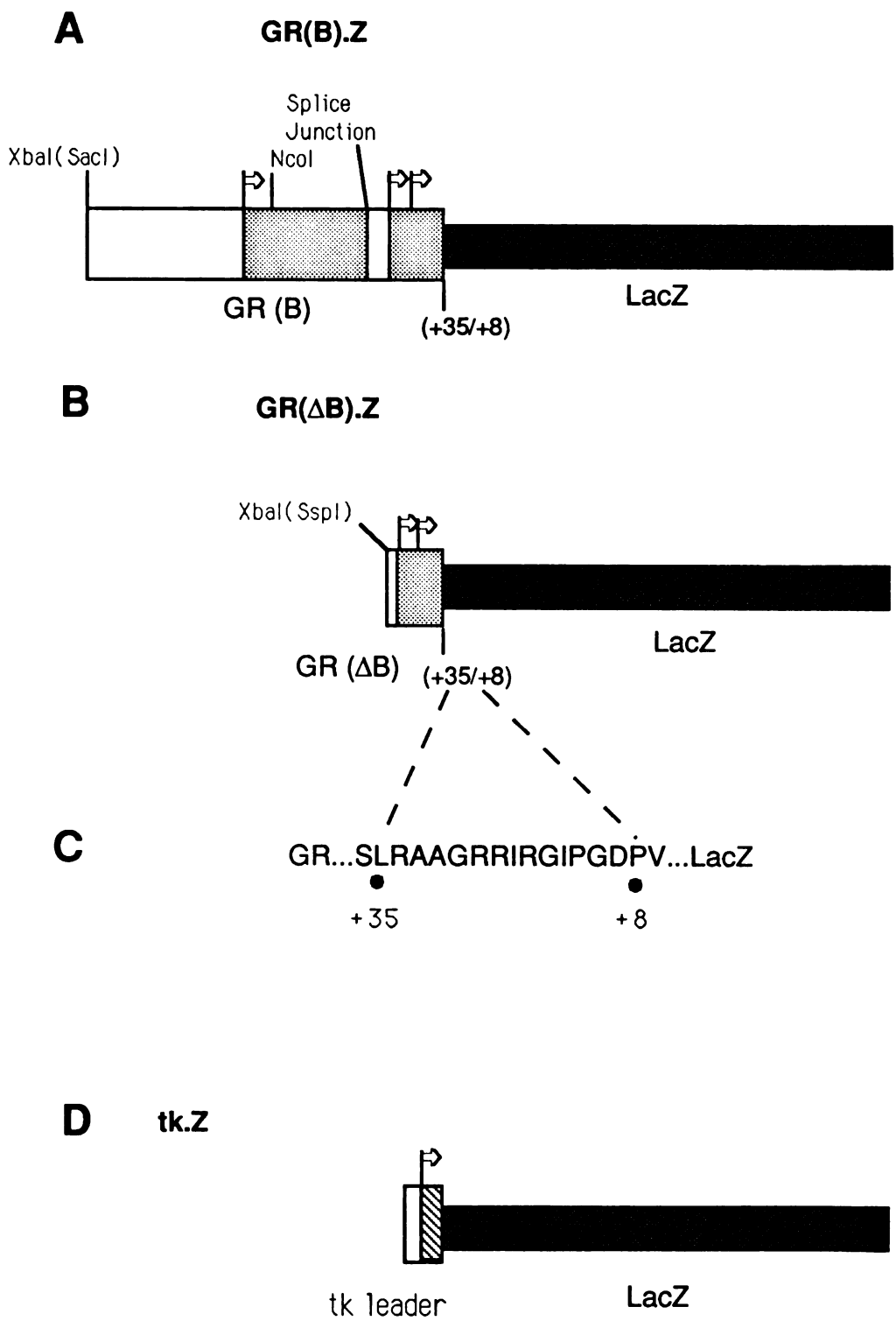
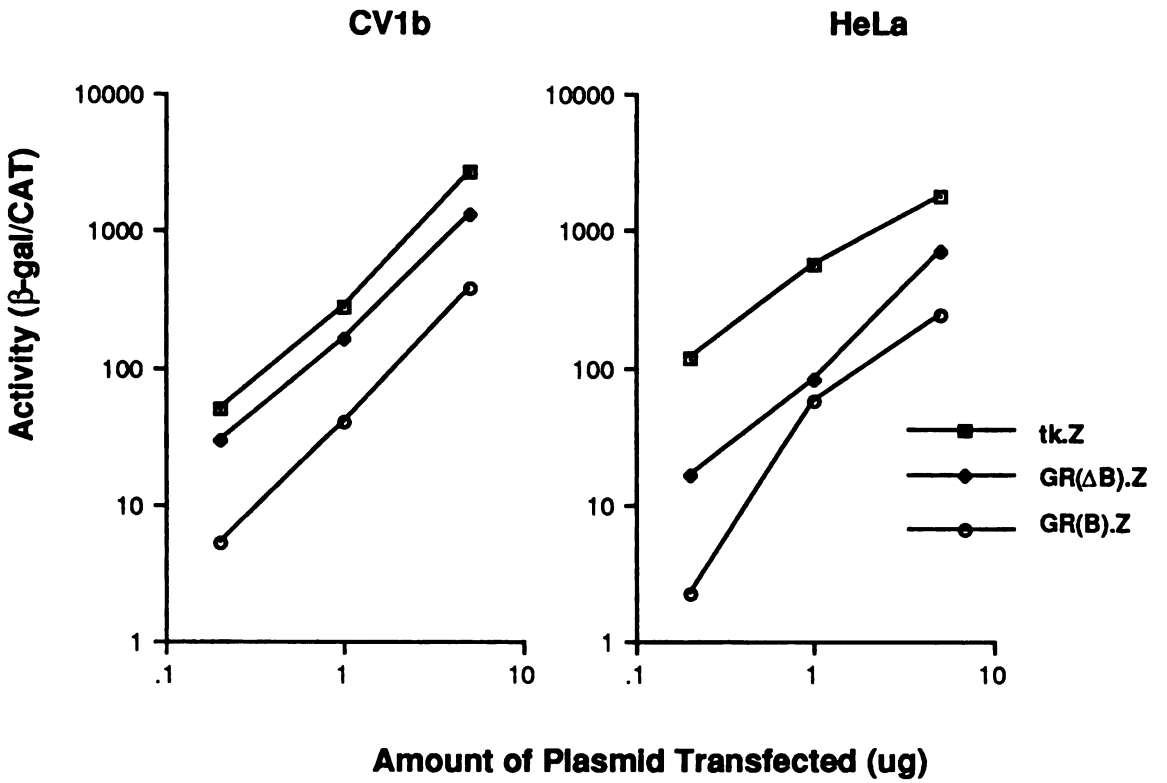

A
GR(B).Z
Splice
Junction
XbaI(SacI)
NcoI
GR (B)
(+35/+8)
LacZ
B
GR(ΔB).Z
XbaI(SspI)
LacZ
GR (ΔB)
(+35/+8)
C
GR...SLRAAGRRIRGIPGDPV...LacZ
+35
+8
D
tk.Z
tk leader
LacZ

Figure 3.11. The B-leader decreases expression of a GR-lacZ fusion construct. **(A)** GR-lacZ fusion expression plasmids (Fig. 3.10) were co-transfected with p6R-CAT. β-galactosidase activity was normalized to chloramphenicol acetyl transferase activity (to control for transfection efficiency). The data points shown represent individual transfections in one experiment. **(B)** GR-lacZ fusion expression plasmids were transfected into CV1b cells in a separate experiment. β-galactosidase activity was normalized to lacZ RNA levels, based on quantitative slot-blots (see Materials and Methods), for each transfection.

Figure 3.11

A



B

Plasmid	βgal (activity/RNA)
mock	0
VA(tk)LO	1656
VA(grB)LO	451
"	207
VA(grΔB)LO	926
"	705

transcripts with at least eight alternate 5' mRNA leaders. Reports in the literature suggested that receptor mRNA was auto-regulated differently in T-cells than in other cells: in most cells and tissues, glucocorticoid hormones down-regulate their receptor RNA, whereas in a human T-cell line it appeared that receptor mRNA was up-regulated. In addition, Kalinyak et al. (1989) reported that dexamethasone down-regulated receptor mRNA from brain and liver in adult rats, but not in neonates. I began a preliminary investigation of whether receptor transcripts are differentially regulated by tissue, glucocorticoids, and postnatal development. I used an RNase protection assay to measure the relative levels of transcript group β (consisting entirely of transcript B) and two other transcript groups α (which includes transcripts A and D) and τ (including transcripts E, F, and G).

A preliminary survey of nine rat tissues and two cell lines showed that most of the tissues (brain, liver, lung, salivary gland, heart, and kidney) along with the rat hepatoma and fibroblast cell lines had similar relative levels of α, β, and τ transcript groups. In these samples, β was the most abundant (about 50 - 75%), followed by α (about 20 - 40%) and τ (about 10 - 20%). This pattern was different in the thymus and spleen RNA, which contained about 2 - 4 fold higher relative levels of τ (35 - 40%) than in other tissues, and correspondingly lower levels of β (25 - 40%). Since band τ represents more than one transcript class, the increased levels of τ in lymphoid cells could be due to the specific expression of one transcript, or to the increased expression of several transcripts. Transcripts E, F and G are candidates for a T-cell specific transcript, but since they were identified in hepatoma cells, there could be a T-cell specific transcript not yet identified. Interestingly, transcripts have been cloned from a mouse T-cell line (chapter three and Dieken et al., 1990)) which could be mouse homologues of such rat lymphoid-specific transcripts. Of

these mouse cDNAs, none matched rat classes B, C, D, E, F, G, or H, and one mouse cDNA clone matched class A. Characterization of cDNAs from rat thymus RNA should help identify T-cell specific transcript(s).

Receptor autoregulation also appears to be different in T-cells than in most other cells. In contrast with the glucocorticoid down-regulation that has been observed in receptor mRNA from most tissues, receptor mRNA from thymus appeared to increase in response to dexamethasone treatment *in vivo*. Glucocorticoid induced increases in receptor protein and mRNA have also been reported in the CEM-C7 human leukemic T-cell line (Antakly et al., 1989; Eisen et al., 1988), in the human myeloma line OPM-2 (Gomi et al., 1990), in the rat pituitary (Sheppard et al., 1990), and transiently in rat HTC cells (but not in liver) at short times after exposure to dexamethasone (Dong et al., 1988; Okret et al., 1986). These examples are in contrast to the large number of observations of receptor down-regulation in a variety of rat and human tissues and cells. In a survey of eight rat tissues, including the spleen (but not the thymus), dexamethasone caused a decrease and adrenalectomy an increase in receptor mRNA levels (Kalinyak et al., 1987). In the only other direct examinations of glucocorticoid regulation of receptor levels specifically in lymphocytes *in vivo*, dexamethasone administration to human volunteers decreased receptor levels in peripheral lymphocytes (Schelechte et al., 1982; Shipman et al., 1983), and cortisol decreased the number of receptor binding sites in human T-lymphocytes maintained in culture (Lacroix et al., 1984). The human lymphoma line IM-9 has also been shown to down-regulate receptor mRNA and protein in response to dexamethasone (Rosewicz et al., 1988).

It is interesting that the cases of receptor up-regulation in lymphocytes cited above all involved T-cells subject to cytolysis or growth inhibition by

glucocorticoids, and the cases of receptor down-regulation in lymphocytes all occurred in glucocorticoid-resistant cells. Rat thymocytes, which consist mainly of cortical immunoincompetent T-cells, are sensitive to glucocorticoids (Claman, 1972), whereas normal human peripheral lymphocytes are resistant (Claman, 1972; Galili, 1983). The rat spleen, which comprises both peripheral B and T lymphocytes, includes both sensitive and resistant cells (Claman, 1972). Receptor levels in glucocorticoid-sensitive T cell lines correlate well with the cell's glucocorticoid sensitivity (growth and viability). These data suggest that receptor up-regulation could be involved in the modulation of glucocorticoid sensitivity in lymphocytes *in vivo*. By creating a positive-feedback loop on glucocorticoid action, receptor up-regulation might ensure that cells do not desensitize or otherwise escape the effects of glucocorticoid exposure. Interestingly, steroid-resistant variants of the sensitive mouse T-cell line W7 have been isolated which reversibly down-regulate receptor in response to dexamethasone (Danielsen and Stallcup, 1984). Thus these resistant variants appear to escape the normal growth-arrest and cell lysis responses of the parent lines because of their negative receptor autoregulation (receptor levels of the parent line W7 do not appear to change in response to hormone). It will be interesting to examine receptor autoregulation in lymphocytes systematically, with particular attention to the sensitivity of the cells to glucocorticoids as well as to other differences in cell phenotype such as developmental stage and subtype.

The dexamethasone-induced increase in receptor mRNA levels in the thymus was specifically detected in the τ transcript group. These changes in τ transcripts were not detected in brain, liver, or spleen, suggesting that a T-cell specific transcript might be up-regulated. No postnatal age-dependence was detected in these thymus-specific changes.

I did not detect any hormone-dependent changes in total transcript levels or in relative levels of α, β or τ in the brain, liver or spleen. Based on data presented by Kalinyak, et al. (1989), I would have expected to observe a decrease in total receptor mRNA levels of about two-fold or less in response to dexamethasone. This magnitude of change was probably too small to detect with the small sample sizes used in this preliminary analysis. This poor sensitivity may also explain why I did not detect differences in hormone-dependent changes between adult and neonatal rats.

Changes in receptor mRNA in response to hormone could be caused by regulation at the level of transcription or mRNA stability. Studies of receptor transcription rate by run-on assays of nuclei from rat liver (Dong et al., 1988) and human IM-9 lymphocytes (Rosewicz et al., 1988) have shown a decrease in receptor transcription rate in response to dexamethasone. Glucocorticoid receptor half-life, however, was unaffected by dexamethasone in hepatoma cells (Dong et al., 1988) or in IM-9 and AR42J cells (Rosewicz et al., 1988). These results suggest that dexamethasone inhibits receptor transcription in these cells without affecting receptor mRNA stability. Nevertheless, since the receptor half-life experiments involved only tissue culture cells, it is still possible that mRNA stability mechanisms could be involved *in vivo*.

If receptor regulates its mRNA by affecting its rate of transcription, the observed tissue- and hormone-dependent differences in relative transcript levels could be due to differential promoter activities. One hypothesis is that there are multiple, negatively autoregulated promoters active in all cells, plus a positively autoregulated T-cell-specific promoter. These promoters could be autoregulated by independent promoter-specific positive and negative glucocortiocid responsive enhancers, or they could be controlled by a single

composite glucocorticoid response element that can confer both responses (Diamond et al., 1990).

Some investigators have suggested that there are one or more glucocorticoid-responsive regulatory elements associated with the receptor cDNA. These elements have been inferred from receptor binding data *in vitro*, and have not yet been shown to behave as enhancers. Okret et al. (1986) reported rGR binding to the 3'UTR of the rat receptor cDNA based on immunoprecipitation and DNase footprinting of rGR-DNA complexes. Though their immunoprecipitation results are suggestive, it is not clear how specifically the receptor binds to the cDNA. The receptor footprints they report are also weak and spread over a large (≥97bp) region. Though weak DNA binding might be expected for a composite GRE (Sakai et al., 1988), these sites in the receptor 3'UTR have not yet been associated with any function. Burnstein et al. (1990) have reported negative autoregulation of transiently transfected hGR. This suggests the presence of a GRE within the receptor coding region since this down-regulation appears to depend on receptor sequences rather than on sequences in the expression vector. It is not clear, however, whether this mRNA regulation was due to transcriptional or to post-transcriptional effects, nor was the putative regulatory region tested for enhancer activity.

The recent discovery of a composite GRE capable of both positively and negatively regulating its cognate promoter (Diamond et al., 1990) provides one possible model to explain the tissue-specific auto-regulation of receptor mRNA. A GRE in the proliferin gene promoter, for example, acts as a negative or a positive regulator depending on the relative levels of c-jun and c-fos in the cell (Diamond et al., 1990), consistent with the existence of AP1-binding sites interdigitated with the GR-binding site (Mordacq and Linzer,

1989). Similarly, the difference in receptor regulation between thymocytes and other cells such as hepatocytes could be in the expression of transcriptional activators that differentially modulate receptor action. Differential expression of a receptor-modulating factor could explain the observed developmental and tissue-specific differences in receptor regulation. Though fos and jun provide a convenient model for modulation of receptor regulation, other transcriptional activators could also fit the paradigm.

One way that differential expression of mRNAs with different 5'ends could be important is in the regulation of receptor translation. The presence of an ORF and potential RNA secondary structure in the B 5'UTR suggested that this UTR might regulate receptor translation. Indeed, GR-lacZ transfection experiments (Fig. 3.11) suggested that the B 5'UTR could inhibit downstream receptor translation. It would be interesting to compare the translational efficiencies of other receptor 5' UTRs with 5'UTR-B. Though transcript B comprises the majority of the receptor RNA present in most of the tissues examined, it is possible that this transcript could contribute much less to the production of receptor protein since it is translationally inefficient. For example, if the B-leader inhibits receptor translation 3-4 fold (Fig. 3.11) and transcript B comprises 70% of receptor transcripts, this transcript might only contribute to the expression of less than 40% of receptor protein.

It is also unknown whether glucocorticoid receptor translation is regulable. An intriguing possibility of translational control of receptor exists in the rat pituitary intermediate lobe, where expression of receptor protein is strongly inhibited by dopamine (Antakly et al., 1987) but receptor mRNA is unaffected (Antakly, personal communication). Unfortunately these cells are difficult to study since they can only be grown in primary culture and are difficult to transfect.

In this chapter I have shown how multiple promoters might affect receptor expression. One mechanism is tissue-specific autoregulation of alternate promoters. My data suggest that there may be a T-cell specific promoter that is positively autoregulated, as well as a ubiquitously expressed promoter that is not positively autoregulated. This is the first demonstration of positive receptor autoregulation in lymphoid cells *in vivo*. In addition, the ubiquitous promoter expresses a 5' mRNA leader that appears capable of inhibiting receptor translation. These results suggest exciting new possibilities for investigating tissue-specific regulation of glucocorticoid receptors.

Chapter Four: The nt^{i} receptor phenotype

Introduction

Glucocorticoid hormones have striking effects on susceptible mouse and human lymphocytes (Claman, 1972), including inhibition of IL-2 secretion (Gillis et al., 1979a,b; Snyder and Unanue, 1982), growth arrest (Harmon et al., 1979), DNA fragmentation (Cohen and Duke, 1984; Distelhorst, 1988; Ucker, 1987; Vedeckis and Bradshaw, 1983; Wyllie, 1980), and cell lysis and death (Claman, 1972; Horibata and Harris, 1970). Steroid-induced cell death is thought to occur through the activation of an endogenous suicide process, perhaps through the induction of cell-lysis specific genes (Gasson and Bourgeois, 1983; Gasson et al., 1983; Harrigan et al., 1989; Yuh and Thompson, 1987). Early attempts to isolate mouse T-cell lines resistant to the cell-killing effects of glucocorticoids yielded resistant variants at high frequency (Sibley and Tomkins, 1974a). Most (~80%) of these clones were unable to bind radiolabeled dexamethasone, and were defined r- (Sibley and Tomkins, 1974b). The remaining mutants were able to bind hormone, but were unable to translocate to the nucleus or bind DNA cellulose (nt-) (Gehring and Tomkins, 1974), or showed increased nuclear transfer and DNA affinity (nt^i) (Gehring and Tomkins, 1974; Yamamoto et al., 1976, 1974), or could not be distinguished from wild-type on the basis of biochemical assays (d-, "deathless") (Sibley and Tomkins, 1974b). These cell lines were derived from the mouse T-cell lymphoma line S49. This parental wild-type line has since been shown to contain one mutant receptor allele (r-) in addition to the wild-type (r+) receptor allele (Danielsen et al., 1986), so these lines are functionally haploid for receptor.

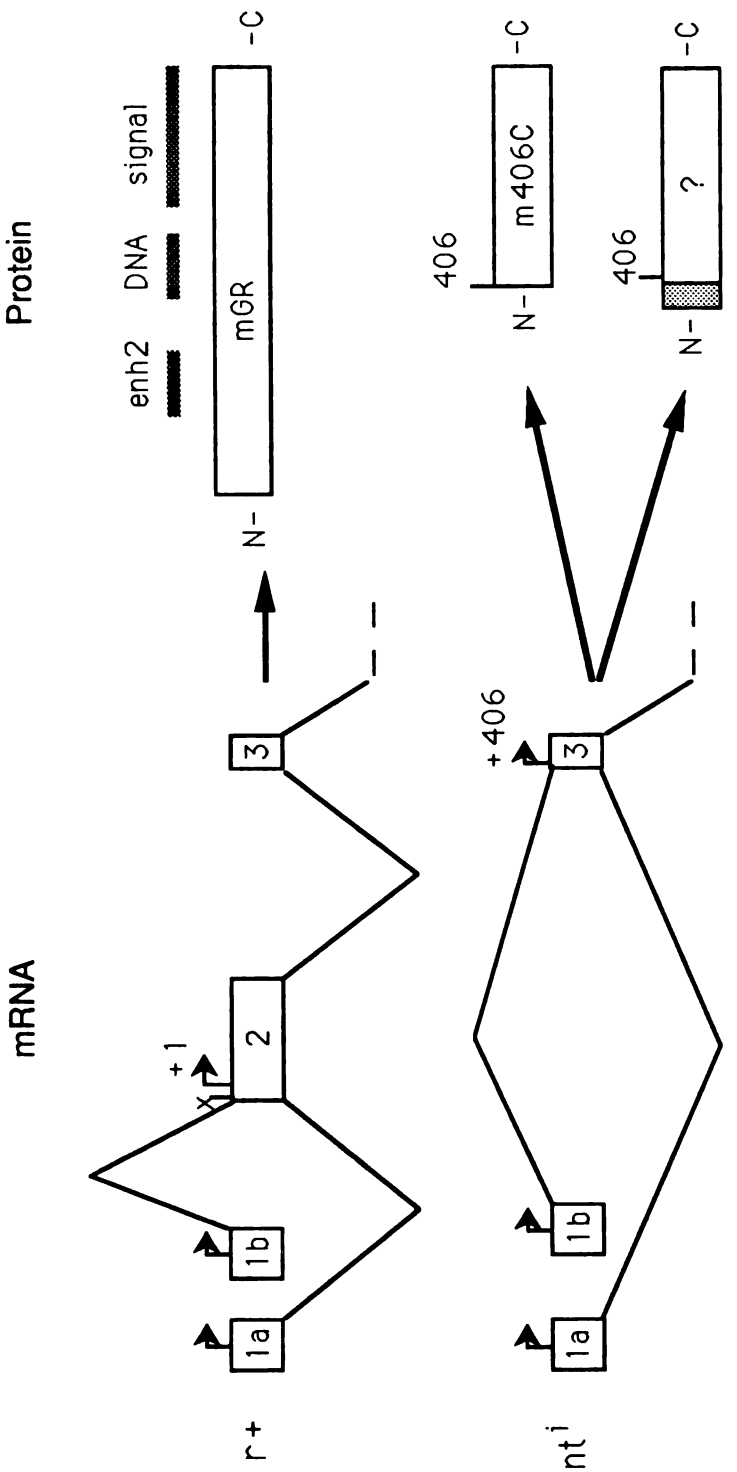
The receptor in nt^i cells was smaller than wild-type (~ 42 kD vs. ~89kD), bound hormone with wild-type affinity, and displayed increased affinity for

nonspecific DNA *in vitro* (Yamamoto et al., 1976, 1974). These alterations in receptor suggested that the nt^i phenotype was due to a receptor mutation. In order to determine the molecular basis for the nt^i phenotype, investigators have cloned and sequenced r- and nt^i alleles from nt^i cell lines (Dieken et al., 1990; Miesfeld et al., 1988).

Mapping of the nt^i mRNA revealed that nt^i was the product of an mRNA species that lacked sequences encoding the N-terminal portion of the receptor (Miesfeld et al., 1984, 1985; Northrop et al., 1986). Comparison of cDNA clones obtained from nt^i cells with the rat genomic structure indicated that nt^i mRNA lacked exon 2 sequences (Miesfeld et al., 1988), possibly reflecting a different mRNA splicing pattern. It is not clear whether this altered mRNA is caused by an alternative splice or by a mutation in the receptor gene. Somatic hybridization studies indicate that nt^i cells do not express a *trans*-acting defect such as a mutation in the splicing machinery, since full-length receptor mRNA and protein is expressed from the wild-type chromosome as well as from the r^- allele (Danielsen et al., 1986; Northrop et al., 1986; Yamamoto et al., 1976). A recent study of receptor genes in nt^i lines did not detect any gross genomic lesions in the line 143R (Dieken et al., 1990).

An in-frame ATG at the beginning of exon 3 (Met 406) could presumably initiate receptor translation in these mRNAs, and would predict a truncated protein of the appropriate size (see Fig. 4.1). Since exon 2 encodes the amino-terminal half of the receptor, these nt^i mRNAs offered an explanation for the decreased size of receptor protein as well as the loss of activity (due to loss of enh 2 sequences). Alternatively, an upstream in-frame ATG in an exon 1 might initiate translation of exon 1-encoded sequences that would fuse to sequences in exon 3 (Dieken et al., 1990; Miesfeld et al., 1988).

Figure 4.1. The nt^i receptor is produced by an mRNA lacking N-terminal sequences. The nt^i receptor is translated from a transcript in which exon 2 sequences are missing, fusing exon 1 family sequences to exon 3 (Miesfeld et al., 1988). In the full-length wild-type receptor (r^+) mRNA, any translational start signals (arrows) in the exon 1 family that are in-frame with receptor would be terminated (x) at the 5' end of exon 2. In the nt^i transcripts, these start codons would also be in-frame with exon 3 receptor sequences, but would not be terminated. Since there is also an initiator codon at the 5' end of exon 3 (aa406), there are two possible sets of translation products: those initiating at aa406 (A), and those initiating at the 5' leader AUGs (B). The latter proteins would contain additional exon 1 encoded sequences (shaded).


Protein
mRNA
enh2
DNA
signal
mGR
-C
N-
406
m406C
?
+1
+406
1a
1b
2
3
r+
nt^i

This predicted protein would contain novel amino-terminal sequences in place of those encoded by exon 2 in the wild-type receptor.

Despite our increasing understanding of the receptor defect in nt^i cells, we still do not know precisely what accounts for the nt^i phenotype. Miesfeld et al. (1988) presented four hypotheses. One explanation is that the receptors in these cells are inactive because they are missing their N-terminal activation sequences. Consistent with this hypothesis is the observation that N-terminal deletions of rat receptor have ~5-10% of the activity of the full-length receptor when transfected both stably and transiently into receptor-deficient HTC and CV1 cells (Miesfeld et al., 1987). This reduction in activity could account for the resistance to dexamethasone since the severity of the cell lysis phenotype appears to depend on receptor dosage in mouse lymphoid cell lines (Bourgeois and Newby, 1979).

Alternatively, if the truncated receptor in nt^i cells contains novel amino-terminal sequences, this portion of the protein might reduce the activity of the remainder of the protein. This hypothesis predicts that removal of these novel amino-terminal sequences would render the protein active. Another hypothesis is that the receptor in nt^i cells is defective specifically in the cell lysis response, perhaps due to the presence of the novel amino-terminal sequences. The amino-terminus of the progesterone receptor for example can distinguish between the ovalbumin and MMTV LTR promoters (Tora et al., 1988), and the activation domain in the amino terminus of the estrogen receptor activates transcription in a cell-specific fashion (Tora et al., 1989). Finally, though it is formally possible that the alterations in receptor are not directly responsible for the nt^i phenotype, studies of wt x nt^i hybrid cell lines indicate that the mutation acts in *cis* with the receptor (Yamamoto et al., 1976).

I present evidence to distinguish among these hypotheses, and conclude that the resistance of nt^i to glucocorticoids is predominantly due to the loss of N-terminally encoded sequences in exon 2. I show that this truncation has reduced the hormone-dependent activity of receptor sufficiently to confer resistance on nt^i cells, as well as to reduce receptor activity more generally. This resistance phenotype did not appear to require exon 1 encoded N-terminal receptor sequences, even though these new peptide sequences appear to be expressed in nt^i cells.

Results

The nt^i phenotype is due to the loss of receptor activity

Experiments with somatic cell hybrids of wild-type and nt^i lines had shown that the nt^i phenotype was "co-dominant" with wild-type (Yamamoto et al., 1976). That is, the killing efficiency of the hybrid was intermediate to the efficient killing of (wt x wt) and the resistance of (nt^i x nt^i). One explanation for this intermediate phenotype was that wild-type receptor concentration was lower in the (wt x nt^i) hybrid because of the larger hybrid cell size (Yamamoto et al., 1976). Alternatively, the receptors from the wt and nt^i alleles might assemble into heterodimers. More importantly, these experiments showed that both the wild-type and nt^i species were produced in the hybrid cell lines, ruling out a *trans*-dominant effect of the nt^i cells on the wild-type receptor allele. Thus it appears most likely that a mutation acting in *cis* with the receptor is responsible for the nt^i phenotype. To confirm this hypothesis, I tested whether I could rescue glucocorticoid sensitivity in an nt^i cell line by transfection with a wild-type receptor.

I infected nt^i cell line S49.143R with a rodent retrovirus (pMV7-mGR) containing a copy of full-length wild-type mouse receptor (mGR) in the expression vector pMV7-MCS (Fig. 4.2). pMV7-MCS is a derivative of the rodent retroviral vector pMV7 (Kirschmeier et al., 1988) which has been used previously to infect S49 cells with high efficiency (Sullivan et al., 1987). I constructed the wild-type receptor by fusing the 5' end of an r^- clone to the 3' end of an nt^i receptor clone and inserting this fusion into the polylinker of pMV7-MCS. Receptor-expressing mRNA is transcribed from the MSV 5' LTR; G418 resistance is expressed from a downstream tk promoter. Using this receptor-containing retroviral plasmid, I produced active retrovirus by the method of Sullivan et al. (1987) and infected 143R cells with this retrovirus. During the same experiment I also infected 7R cells with this and another retroviral construct, pMV7-m406C (discussed below). I plated infected cells in soft agar (Sibley and Tomkins, 1974a) and selected colonies that were G418 resistant.

I selected one mGR-infected 143R clone (143R.GR.1) and analyzed it in detail. Both growth (Fig. 4.3) and cloning efficiency (Table 4.1) showed sensitivity to dexamethasone similar to wild-type and different from the resistance phenotype of 143R. Both the wild-type (S49.A2) and 143R.GR lines failed to grow in the presence of dexamethasone in both assays. Thus the glucocorticoid-sensitivity phenotype in 143R appears to be restored by wild-type receptor, indicating that the glucocorticoid resistance of this nt^i line is due to a mutation affecting glucocorticoid receptor.

Some resistant colonies, however, did appear in one of the soft-agar cloning experiments (expt 2, table 4.1). The resistance of the S49.A2 in expt 2 corresponded to only one resistant colony in the presence of dexamethasone vs. 281 colonies in the absence of hormone: a relative efficiency of $\leq 3.6 \times 10^{-3}$.

Figure 4.2 Retroviral constructs expressing wild-type (mGR) or truncated (m406C) receptor. Retroviral plasmids were constructed by inserting cDNA encoding full-length (mGR) or Met406-initiating (m406C) mouse receptor into the vector pMV7-MCS (a derivative of pMV7 (Kirschmeier et al., 1988) containing the pUC18 multiple cloning site). **(A)** pMV7-m406C was constructed by inserting the 5' Xba-SphI fragment of pBS-55n23ΔSma (a class A nt^i clone with a 5' deletion at the SmaI site) and the 3' SphI-XbaI fragment of pBSr (a wild-type mouse receptor clone constructed by R. Miesfeld by fusing the N-terminal portion of an r^- clone with the C-terminal portion of an nt^i clone; described by Dieken et al. (Dieken et al., 1990)) into the cloning site of pMV7-MCS (cut with XbaI). The first ATG of this receptor clone occurs at Met 406; there are two in-frame termination codons upstream of this initiator in the Bluescript sequence. **(B)** pMV7-mGR was constructed by inserting the EcoRI-XbaI fragment of pBSr into pMV7-MCS at the multiple cloning site. **(C)** pMV7-MCS is a rodent retroviral vector. The pUC18/19 polylinker was inserted into the EcoRI-HindIII sites of pMV7.

Figure 4.2

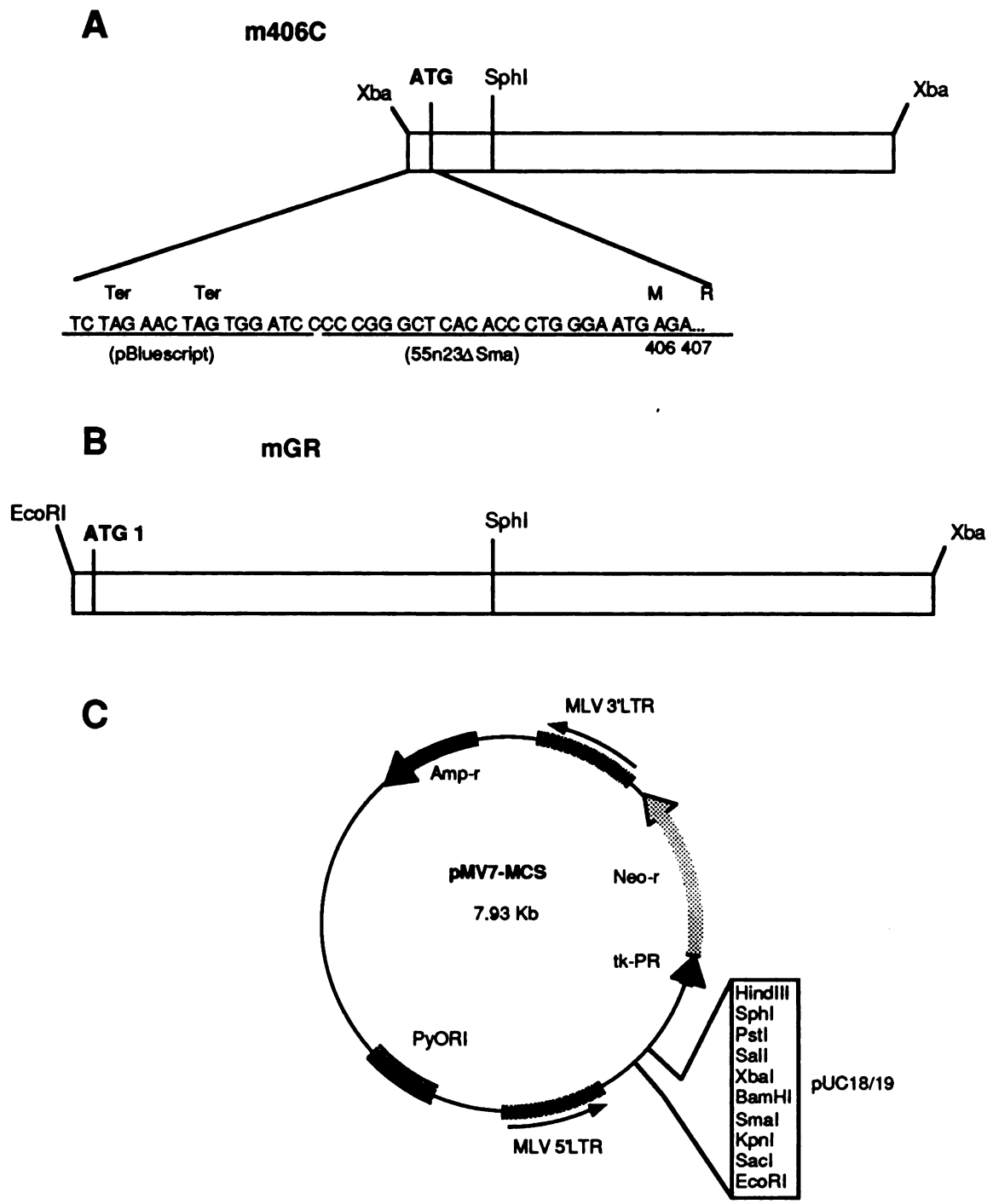
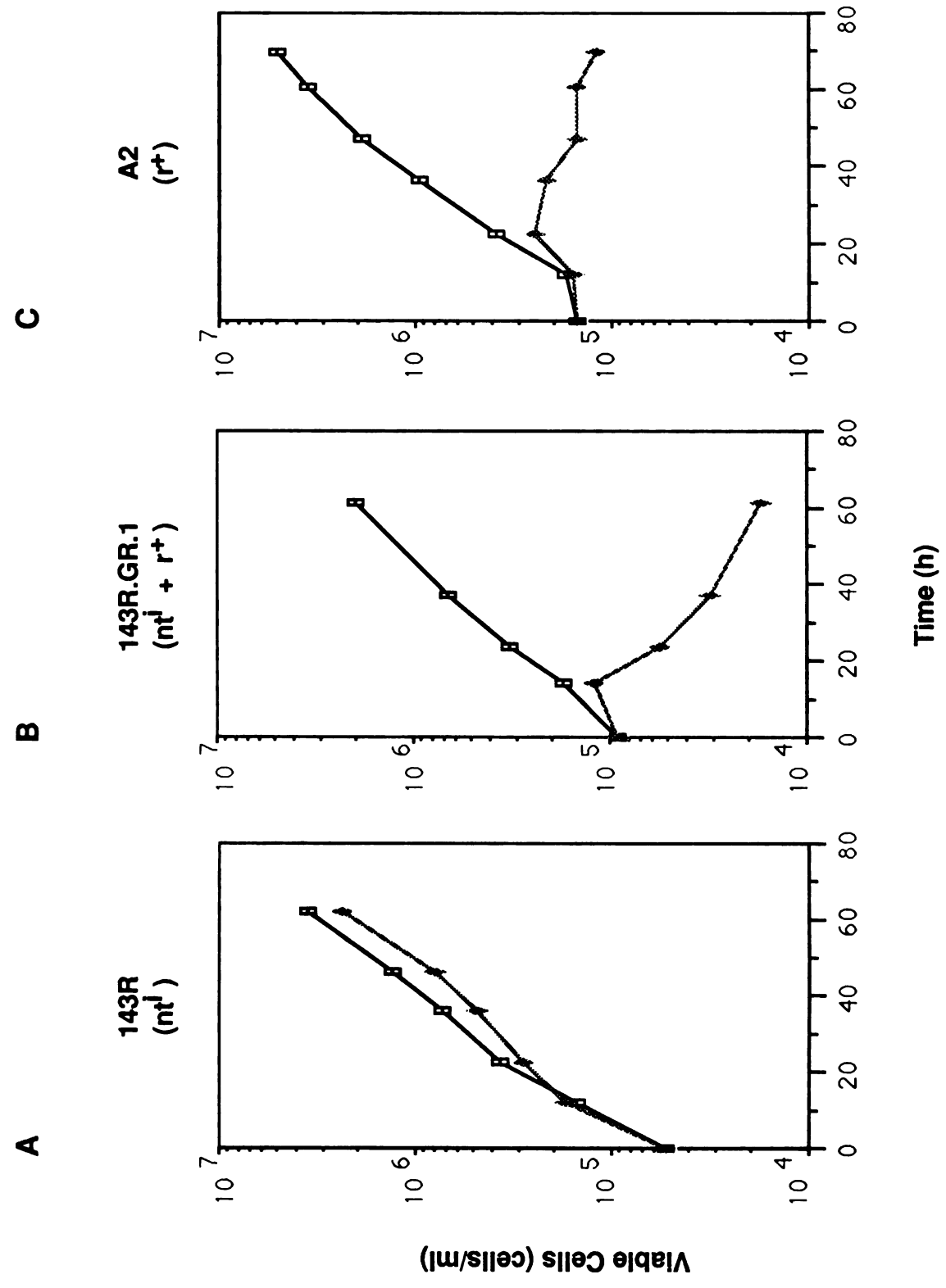

A
m406C
Xba
ATG
SphI
Xba
Ter
Ter
M
R
TC TAG AAC TAG TGG ATC CCC CGG GCT CAC ACC CTG GGA ATG AGA...
(pBluescript)
(55n23Δ Sma)
406 407
B
mGR
EcoRI
ATG 1
SphI
Xba
C
MLV 3'LTR
Amp-r
pMV7-MCS
7.93 Kb
Neo-r
tk-PR
PyORI
MLV 5'LTR
HindIII
SphI
PstI
SalI
XbaI
BamHI
SmaI
KpnI
SacI
EcoRI
pUC18/19

Figure 4.3. Introduction of wild-type receptor into the nt^i line 143R recovers the glucocorticoid-induced cell-killing phenotype. The indicated cell lines were grown either in the presence (gray lines) or absence (black lines) of 1 μM dexamethasone. 143R grows normally in the presence of dexamethasone **(A)**. Expression of wild-type receptor (see Fig. 4.2) in this line renders it dexamethasone-sensitive **(B)**, similar to the wild-type phenotype **(C)**. The densities of viable cells were determined by counting trypan-blue excluding cells on a hemacytometer. The effects of dexamethasone on 143R.GR and S49.A2 were also apparent visually. Healthily growing cells appeared round and refringent, whereas the dex-sensitive lines contained many sick-looking or dead cells (shrivelled or swollen, non-refringent). Counting viable cells by these visual criterea gave the same results as by trypan-blue exclusion.


A
B
C
143R
(nt^i)
143R.GR.1
(nt^i + r^+)
A2
(r^+)
Viable Cells (cells/ml)
Time (h)
10^7
10^6
10^5
10^4
0
20
40
60
80

Cell Line	^{3}H-Dex Bound (receptors/cell)[16]	Relative Hormone Binding (%143R)	Cloning Efficiency in Soft Agar (+dex/control, %) Expt1	Expt2
A2	17,500 ± 4,900(3)[17]	33	0.0	0.36±0.8[18]
143R	53,600 ±20,000(3)	100	28.6±24.2	16.4±13.3
55R	31,600 (1)	59	n.d.	n.d.
143R.GR.1	86,700 ±12,100(2)	162	0.0	2.3±2.1
7R	100 ± 350(3)	0.2	75.0±22.8	114±30
406C.2	229,500 ±60,700(3)	428	62.7±16.0	81.0±16.3
406C.4	136,000 ±38,300(3)	254	80.0±25.7	94.0±10.3
406C.5	130,900 ±24,700(3)	244	0.0	4.4±2.7
406C.6	86,900 ±12,100(3)	162	1.7±1.1	11.0±3.9
406C.7	112,500 ±28,900(3)	210	5.2±2.3	0.67±0.61

Table 4.1. Receptor levels and dex sensitivity of S49 cell lines. Receptor levels were determined by ^{3}H-dexamethasone binding to whole cell extracts (Miesfeld et al., 1986). Cloning efficiency in soft agar was determined by counting the number of colonies visible after 10 - 11 days on five 60 mm plates each with and without 1 μM dexamethasone. The cloning efficiency of cells in the absence of hormone varied from 5 - 55% (for 150 - 500 cells plated per 60 mm dish), with 143R cloning at ~5% and the other cells at >10%.

[16]Estimate based on measured dpm/mg protein and assuming 0.21 mg protein / 10^6 cells (determined for S49.A2).

[17]Mean ± sample standard deviation (n).

[18]Mean ± sample standard deviation (geometric mean of SDs for numerator (n = 5) and denominator (n = 5)).

In expt 1, 0/217 colonies were resistant. Averaging these two experiments gives a relative efficiency with an upper bound of about 2×10^{-3}. This value is roughly intermediate between the relative efficiencies originally reported for the wild-type line S49.1A (about 0.02) and some of its subclones[19] (10^{-6} to 10^{-7}) (Sibley and Tomkins, 1974a), and the relative efficiency reported for wt x wt hybrids (10^{-4}) (Yamamoto et al., 1976).

143R.GR.1 gave 0/103 and 6/266 resistant colonies in expts 1 and 2, respectively (table 4.1). This yields an average[20] relative efficiency of $\sim 1.6 \times 10^{-2}$, fairly close to the efficiency reported for wt x nt^i hybrids (10^{-2}) (Yamamoto et al., 1976). This apparent agreement may simply be coincidental; since the degree of cell-killing in a population seems to correlate with the concentration of wild-type receptors (Bourgeois and Newby, 1979), it may be fortuitous that 143R.GR mimics the hybrid phenotype quantitatively. By this argument, higher levels of wild-type receptor expression would more closely mimic the wild-type phenotype and give higher killing efficiencies. This could be tested directly by measuring full-length receptor concentration in 143R.GR clones expressing different levels of receptor.

The receptor in nt^i cells is impaired in induction of GRE-linked promoters

The receptor in nt^i cells could be specifically impaired in triggering cell-lysis response genes, or it could be generally defective in inducing transcriptional responses of GRE-linked promoters. In order to test whether

---

[19]S49.A2 is a subclone of S49.1A.

[20]The two sets of observations are not significantly different ($p < .05$) from those expected if they were randomly drawn from the same population (data not shown).

other responses to dexamethasone were also reduced in nt^i cells compared to wild-type, I transiently transfected various GRE-containing reporter plasmids into both cell types. Three reporters tested (with AdML, MMTV and β-globin promoters) showed an approximately 20-fold reduction of activity in 143R and 55R compared to S49.A2 (Fig. 4.4; quantitation not shown), even though the nt^i lines contained at least as high receptor levels as wild-type (table 4.1). A tk-promoter-containing reporter (pGTCO) also showed reduced activity in the nt^i lines (data not shown[21]). Thus, the nt^i phenotype appears to be general to GRE-linked promoters, rather than specific to putative cell-lysis response genes. As a more direct positive control, this nt^i phenotype should also be tested by complementation with wild-type receptor (i.e., 143R.GR).

Nt^i mutants contain receptors that lack the first coding exon and carry novel amino terminal peptides

In place of exon 2, nt^i mRNAs contain distinct 5' sequences (presumably encoded by exon 1 family members) fused to the first zinc-finger exon (exon 3). M. Schena and R. Miesfeld had cloned r^- and nt^i receptor cDNAs from 55R and 143R (Miesfeld et al., 1988); unpublished data). M. Schena, R. Miesfeld and I sequenced some of these clones to determine whether they contained upstream initiators. Inspection of these 5' leader sequences revealed two different 5' leader sequences (defined as leaders A and S; see Fig. 4.5), each containing an upstream open reading frame and ATG. In nt^i transcripts these 5' leader-initiated peptides could read through into the

---

[21] The induction in wild-type cells was about 4-fold, based on quantitation by liquid scintillation counting of the acetylated and unacetylated ^{14}C-chloramphenicol on the TLC plates. The activity was too low for the differences to be apparent on the exposed film.

receptor reading frame in exon 3, and encode truncated receptors with novel amino termini in place of the exon 2 encoded N-terminal domain (see Fig. 4.1). In wild-type transcripts, a stop codon in the receptor reading frame at the 5' end of exon 2 would terminate translation products initiated in the 5' leader, prior to initiation of the receptor open reading frame. Recently, (Dieken et al., 1990) have reported the sequences of 21 additional cDNA clones from S49.A2, S49.143R and S49.55R. The published nucleotide and predicted peptide sequences of the most abundant of these (17/21) was identical to corresponding sequences of S-leader-containing clones. Their sequence similarly contained an upstream in-frame ATG encoding a predicted novel amino-terminal peptide fused to exon 3.

In the full-length transcript, these 5′ open reading frames are terminated by a UAA condon at the beginning of exon 2, whereas in the nt^i transcripts these open reading frames are fused in-frame with receptor sequences at exon 3, completely skipping exon 2. Translation of the ca. 40kD nt^i protein could initiate at mouse Met 406 (equivalent to rat Met 418) at the 5′ end of exon 3, or it could express a new short amino terminus by initiating translation from an in-frame ATG encoded by one of the fused 5' leader sequences (Fig. 4.5).

To examine these possibilities, I compared receptor proteins expressed by the nt^i cell line 143R with a mouse receptor protein that initiated at Met406 (m406C; see Fig. 4.2). The 7R.406C lines express a mouse exon2-deleted receptor (m406C) that contains exon 3, but no exon1 encoded initiator, in an r^- background (S49.7R). To visualize these N-terminally truncated proteins I used a polyclonal antiserum raised against an E. coli expressed galK/mGR C-terminus (aa453-783) fusion protein (αGCR3'; courtesy Ray Sweet, Smith Kline Beecham Laboratories).

Figure 4.4. Nt^i cell lines show reduced activity on activation of GRE-linked promoters. Nt^i (143R and 55R) or wild-type (A2) lines were transfected with CAT reporter plasmids containing one or more GREs linked to the indicated promoters: AdML, Adenovirus major late promoter (pMG18C, containing 18 GREs); MMTV, mouse mammary tumor virus promoter (pGMCS); β-globin, rabbit βglobin promoter (pGBCO). A representative experiment is shown. For each +/- dex pair, one ml of cells were electroporated in Bio Rad *(name)* cuvettes (225V, 1300μF) at a density of 1.0×10^7 cells/ml in 1 x HBSP buffer (see Materials and Methods), with 40μg (MG18C) or 100μg (GMCS, GBCO) of plasmid plus sheared salmon sperm DNA to 500μg DNA final. Cells were left (in the cuvettes) at room temerature for 10 - 20 min and then transferred to 25 ml of culture medium (DME H21 containing 10% defined supplemented calf serum, pre-equilibrated in the tissue culture incubator). Each 25 ml suspension was then split equally between two T-25 flasks, to one of which was added dexamethasone to 1 μM final. Cells were harvested 16 h later and protein extracted for CAT assays (see Materials and Methods). Equal quantities of protein (determined by Bio Rad Bradford assay) were assayed within each set of plasmids transfected. After exposing the TLC plate to X-ray film, I cut out acetylated and unacetylated ^{14}C-chloramphenicol spots and counted them in a liquid scintillation counter.

Figure 4.4

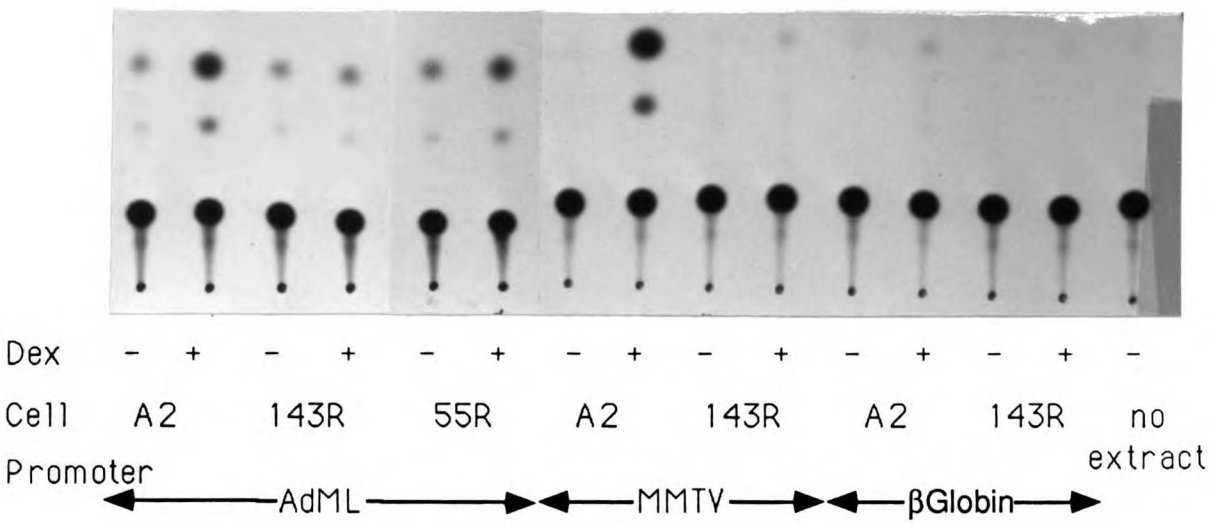

Dex - + - + - + - + - + - + - + - +
Cell A2 143R 55R A2 143R A2 143R no extract
Promoter
AdML
MMTV
βGlobin

Figure 4.5. Two classes of receptor 5' sequences from nt^i cells each contain an ATG in the receptor reading frame. Two different 5' sequences from nt^i lines 55R and 143R each contain an ATG in the receptor reading frame (shown as triplets) within their 5' leader sequences (exon 1). These sequences are named A and S (A because of the similarity with the rat exon 1A sequence; S to avoid confusion with rat transcripts B-H). The exon 2 and 3 potential initiator codons are shown. ATG codons are shown in outline and termination codons are underlined. In wild-type or r^- transcripts, proteins initiated from these 5' ATGs in the receptor reading frame would be terminated in exon 2. In transcripts from the nt^i allele, these upstream initiators would read through into exon 3 and produce an exon 1 encoded peptide fused to the C-terminal portion of receptor. Also note that the S 5' leader contains additional open reading frames. Class A matches the sequence of the rat A transcript, except for a CTG in rat in place of the mouse ATG in the 5' leader.

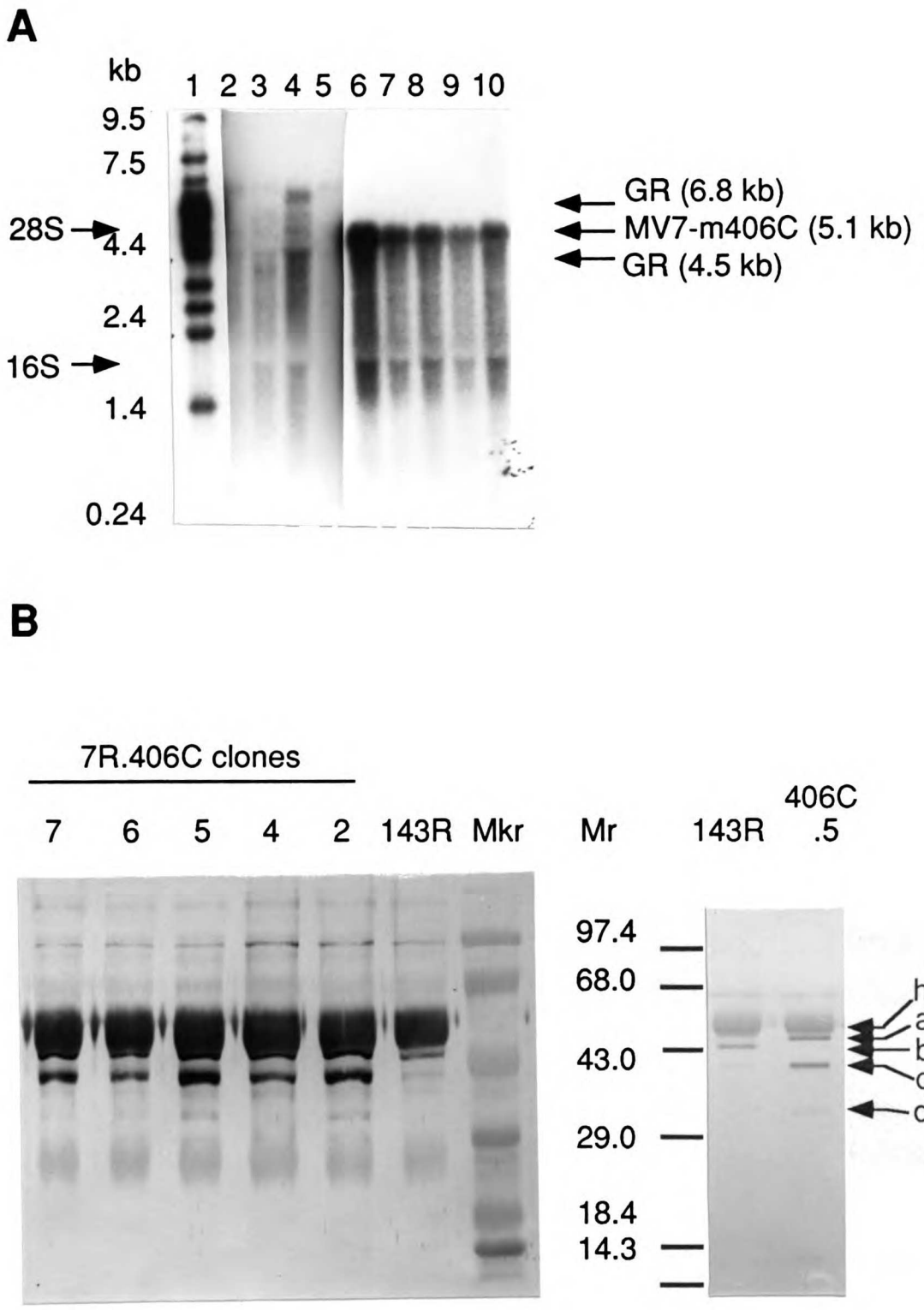
The sequences shown are based on receptor-hybridizing cDNA clones isolated by M. Schena and R. Miesfeld (unpublished data and (Miesfeld et al., 1988)) from a λgt10 library from S49.55R and S49.143R cells (prepared by A. Inoue). Restriction mapping data indicated that most of these clones were from the r^- allele, likely because of the way the library was constructed (M. S. & R. M., unpublished data). Of those clones that were sequenced (M.S, R.M., and my data), four were class S (55r6, 143r1, 143n3, 143n73) and two were class A (55n2, 55n23)[22].

[22]55 and 143 indicate the cell line from which the library was made. r = r^- allele, n = nt^i allele. The sequence of 55n2 was previously published in Miesfeld et al. (1988) as 55.2.

CLASS	5' SEQUENCES	EXON 2	EXON 3
A	....ACG **ATG** CGG GGG TGG GGG ACC TGG CAG CAC GCG AGT CCC CCC CCG GGC TCA CA	T TAA TAT TTG CCA **ATG** GAC ... 1	C CCT GGA **ATG** AGA ... 406
S	.....NA CAA CTN GTA CAG GAC AGT GCA GTG TGG TCC ACA GAT **ATG** AAC TCT GAT AAT CGT GC**A** **TGA** GCT ACT CTG ATC GTA AGA **ATG** GAG AAG AGA GCA GCC CAG CTC CCA CCC TCC TGG GGT TCC CAT CGC AGC CTG ATC ATC TGC AGC CTT CTC AGC CAG GAA GA**T** **G**TT TCA GAT CCT GCT TCG TTA GAG TGT CTG GGA GGA AG		

Figure 4.6. Expression of receptor mRNA and protein in m406C clones. (A) Northern blot of receptor RNA from 7R and 7R.406C clones. Only the ~5.1 kb MV7-GR mRNA is visible on this exposure (3h; lanes 1, 6 - 10); the 6.8 kb receptor band is faintly visible in both 7R and 7R.406C lanes on a much longer exposure (48h; lanes 2 - 5). Total cytoplasmic RNA was extracted by the NP-40 lysis method () and 5μg run on a 0.9% HGT-Agarose/6.6% formaldehyde gel in 1xMOPS buffer. Nucleic acid was transferred to a Zetaprobe (Bio Rad) nylon membrane and hybridized to a ^{32}P-labelled random-primed receptor probe (p0.7RP). Lanes: (1) RNA ladder (BRL), (2) S49.A2, (3) 143R, (4) 143R.GR, (5) 7R, (6) 406C.2, (7) 406C.4, (8) 406C.5, (9) 406C.6, (10) 406C.7. **(B)** Western blot of receptor in 143R and 406C clones. Whole cell extracts were incubated overnight at 4°C in RIPA buffer with the rabbit polyclonal antiserum αGCR3' (courtesy of Ray Sweet, Smith Kline Beecham) at 1:20 dilution. This antiserum recognizes the carboxy-terminal end of mouse glucocorticoid receptor (aa452 - 783). Antibody complexes were then precipitated with protein A sepharose, washed 3-4 times in RIPA buffer followed by one wash in receptor buffer, and precipitates denatured by boiling in SDS sample buffer. Supernatants of this extract were run on a 10% SDS-Polyacrylamide gel, transferred to nitrocellulose, and then incubated with GRC3' primary antibody (1:500) followed by goat anti-rabbit IgG - alkaline phosphatase conjucated secondary antibody (1:3000). Bands were visualized by alkaline phosphatase staining with NBT/BCIP reagent. Labelled bands: (HC) immunoglobulin heavy chain (51.5 kD), (a) "extra" m406C band (47.0 kD), (b) nt^i-specific band (44.0 kD), (c) m406C Met406 (40.5 kD), (d) m406C Met493 (32.5 kD).

Figure 4.6


A
kb
1 2 3 4 5 6 7 8 9 10
9.5
7.5
28S
4.4
2.4
16S
1.4
0.24
GR (6.8 kb)
MV7-m406C (5.1 kb)
GR (4.5 kb)
B
7R.406C clones
7 6 5 4 2 143R Mkr
Mr
143R
406C .5
97.4
68.0
43.0
29.0
18.4
14.3
hc
a
b
c
d

Western blots of SDS gels of 143R and 7R.406C cell extracts are shown in Fig. 4.6 (B). These blots show what appears to be an identical immunoreactive band in the lanes from all m406C clones at a mobility of approximately 40.5 kD (band c), close to the predicted size of the 378aa m406C protein (40.6 kD, assuming 110 D/aa). The most prominent band from 143R (band b), however, is slightly larger, about 44 kD, suggesting that the receptor in 143R contains about 4 kD of additional protein sequence compared to m406C, appproximately what is expected based on the presence of the 5' leader ATGs. In addition to the 44 kD band, 143R also displays a more weakly staining band that comigrates with the major 40.5 kD m406C band, suggesting that 143R also expresses (at lower levels) a protein initiating at Met 406. This expression pattern is consistent with the expected preference of ribosomes to initiate protein synthesis from the first ATG in a "good context", with some leaky scanning that allows some of the ribosomes to initiate from the next AUG downstream (Kozak, 1987). Consistent with this model, m406C lanes also show a comigrating, weakly staining band (d) running below the 40.5 kD m406C band (c). This band runs at the predicted mobility for a protein initiating at the next ATG downstream of Met 406 (i.e., Met493; 32.0 kD assuming 110 D/aa). Thus, it appears that the receptor expressed in an nt^i cell line consists of a population of receptors initiating at different methionines, with the most predominent species initiating upstream of Met 406. Since these upstream ATGs must be located in the 5' leaders, rather than exon 2, the nt^i receptor must contain novel amino-terminal sequences encoded by these 5' leaders.

The 7R.406C lanes also contained a slightly slower migrating band (a) that was not predicted from the sequences. The origin of this band is unclear,

especially since the cloning procedure predicts two upstream in-frame termination codons (Fig. 4.2).

Deletion of the N-terminal domain of receptor appears to reduce specific activity for induction of cell lysis

The receptor in nt^i cells substitutes short amino terminal sequences from members of the exon 1 family for the amino-terminal domain normally translated from exon 2. Are these novel amino-terminal sequences themselves responsible for the nt^i phenotype, or is deletion of exon 2 encoded sequences sufficient? Nt^i receptors may fail to activate GREs because the substituted amino-terminus somehow represses the activity of the rest of the protein, or simply because they are lacking the exon2-encoded activation domain. If the former hypothesis is correct, then a truncated protein lacking this new N-terminus should be active, at least sufficiently to induce cell-lysis. In order to determine whether such a truncated protein has dexamethasone-induced cell-killing activity, I used the 7R.406C stable cell lines discussed in the previous section (see Fig. 4.2). All five lines expressed m406C protein and mRNA of the expected size (Fig. 4.6) and showed approximately 2-4 x greater dex binding than 143R (table 4.1).

All five m406C lines studied were more resistant to dexamethasone than was the wild-type line S49.A2 (see Figs. 4.7 and 4.8 and Table 4.1). However, there was substantial variability in the hormone sensitivity of these lines. Clones 406C.2 and 406C.4 were completely insensitive to dexamethasone, their growth curves indistinguishable from that of 7R. In contrast, clones 406C.5, 406C.6, and 406C.7 were somewhat sensitive to hormone, being more sensitive than 143R or 55R, though not as sensitive as

S49.A2. There did not appear to be any correlation between dosage of m406C receptor and dexamethasone sensitivity (Table 4.1), nor were there any apparent differences in the levels or sizes of m406C protein or mRNA expressed in these clones (Fig. 4.6).

Taken in isolation, the greater sensitivity of clones 406C.5, 6 and 7 compared to the nt^i lines may be explained by their higher expression of m406C protein compared to the nt^i lines (see Table 4.1). The insensitivity of clones 406C.2 and 406C.4, by this argument, might then be explained by a dosage-independent mutation introduced during retroviral infection. Such a mutation would have to be subtle, since m406C mRNAs and proteins were indistinguishable among the clones. An alternative hypothesis to explain these observations would be that there is some underlying variability in the parent 7R lines which is manifest in the phenotype of the m406C expressing clones.

During construction of the 7R.406C line I also made 7R cell lines expressing full-length mouse receptor (7R.GR). If 7R contained a high frequency of non-receptor-linked resistant variants (e.g., akin to the "deathless" phenotype), they should be apparent in the phenotype of some of these 7R.GR lines. As part of the initial screening of stable lines, I tested these lines, along with the 7R.406C, 143R.GR, wild-type and parental lines, for glucocorticoid-sensitivity incubating aliquots of cells with or without dexamethasone in nickel-well plates (data not shown). After 1-3 days I scored the lines as either insensitive, if there were no differences between wells with and without dexamethasone, or sensitive if there was a detectable difference. Sensitive cells appeared shriveled and did not grow well in the presence of 1μM dexamethasone (the medium remained red rather than turning orange, and there were fewer cells than in the control well); differences were

Figure 4.7. Growth-sensitivity of 7R.406C clones to dexamethasone. The indicated cell lines were grown either in the presence (gray lines) or absence (black lines) of 1 μM dexamethasone. All five 7R.406C clones grew in the presence of dexamethasone, despite the variability in growth rates.

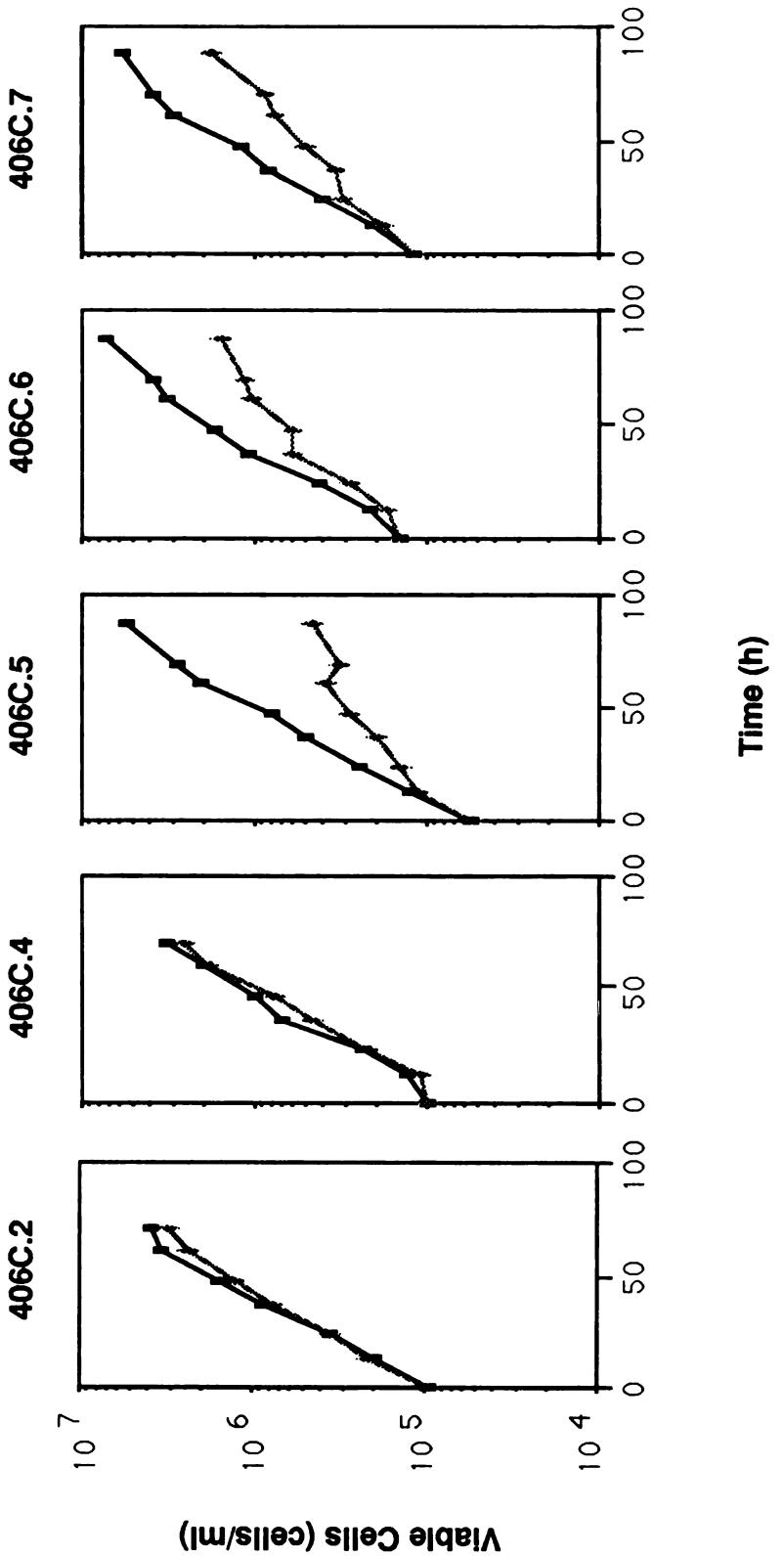
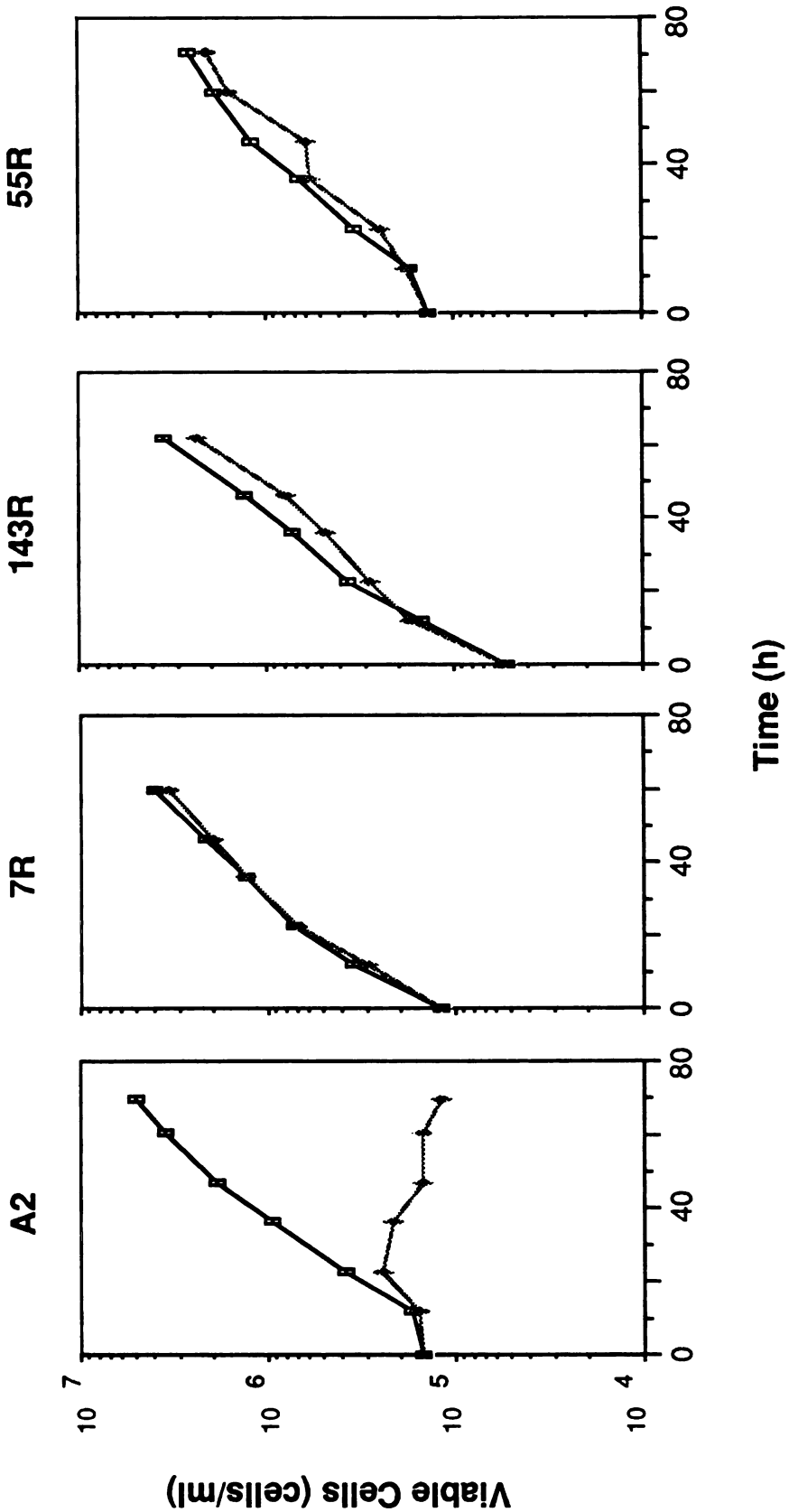

406C.2
406C.4
406C.5
406C.6
406C.7
Viable Cells (cells/ml)
10 7
10 6
10 5
10 4
0
50
100
Time (h)

Figure 4.8. Growth-sensitivity of S49 lines A2 (wild-type), 143R (nt^i), 55R (nt^i), and 7R (r^-) to dexamethasone. The indicated cell lines were grown either in the presence (gray lines) or absence (black lines) of 1 μM dexamethasone.


A2
7R
143R
55R
Viable Cells (cells/ml)
10 7
10 6
10 5
10 4
0
40
80
Time (h)

generally obvious within a two days. This assay gave results similar to the growth curves and soft-agar cloning assays: 7R, 143R, 406C.2, and 406C.4 were all scored as resistant, whereas A2, 406C.5-7, 143R.GR, and all 11 out of 11 tested 7R.GR clones were scored as sensitive. These results suggest that the variability in 7R.406C phenotypes was not due to the unmasking of a dexamethasone-resistant phenotype in part of the 7R population.

Discussion

Receptors from nt^i lines appear to contain additional peptide sequences compared to m406C, supporting the idea that the 5' leaders encode novel amino terminal sequences, fused to exon 3, in place of the exon 2 encoded portion of the receptor. However, the function of these new amino terminal sequences is unclear. It is also unknown whether the splicing pattern of receptor in nt^i cells has a normal biological role or is a phenomenon peculiar to S49 cells.

An intriguing hypothesis is that the receptor splicing pattern found in nt^i cells is revealing an alternate splice pattern that occurrs normally during the development and differentiation of T-cells. For example, suppose a DNA site in the receptor gene necessary for proper splicing of exons 1 and 2 were sensitive to a specific inhibitory signal (e.g. present in specific T-cells) that caused the mRNA splicing machinery to skip exon 2. A constitutive mutation in that site might cause an nt^i-like splicing event and thus confer steroid resistance on those cells. This hypothesis would be supported by the finding of the same splice pattern in other, independently derived T-cell lines, or in T-cells *in vivo*. A simpler question to address is whether the new amino-terminus of receptor is required for the nt^i phenotype (i.e., loss of

steroid sensitivity). If true, it would suggest that the truncated receptor from nt^i cells might have a normal biological function.

The resistance or partial resistance of the 7R.406C clones suggests that deletion of enh2, without the addition of exon1-encoded amino acids, was sufficient to confer glucocorticoid resistance in S49 cells. Thus the amino-terminal amino acids of receptor in nt^i were not essential for the resistance phenotype. The loss of activity of the truncated protein was also not promoter-specific, since transfected CAT reporters driven by four different GRE-linked promoters were about 10-20 fold less active in 143R than in wild-type cells. This result is consistent with experiments showing that amino-terminally truncated rat GR (r407C) was about 5-10% as active as full length rat receptor (rGR) when these receptor genes were stably transfected into receptor-deficient HTC cells (Miesfeld et al., 1987). These experiments do not rule out the possibility that receptor from nt^i cells might be active in other contexts, since this protein contains other enhancement regions such as *enh1*, located in the zinc-finger, and *enh3*, located in the signal transduction domain of receptor. *Enh3* is explored further in the next chapter.

If receptor in nt^i cells is partially active, why are nt^i S49 cells dexamethasone resistant? Though cell lysis is an all-or-none response in an individual cell, glucocorticoid-induced lysis of the cell population appears to occur in a graded, receptor dose-dependent fashion (Bourgeois and Newby, 1977, 1979; Gehring et al., 1982). The nt^i lines 143R and 55R appeared to be slightly more sensitive to dexamethasone than the r^- line 7R in both the growth curve and soft-agar cloning experiments (Fig. 4.8 and Table 4.1), consistent with a partially active receptor in the nt^i lines. The intermediate sensitivity of the m406C lines 7R.406C.5, 6, and 7 can be explained by their higher levels of expression compared to the nt^i lines. The complete resistance

of m406C.2 and 4, however, does not fit this pattern, since these clones also appear to express high levels of receptor. One explanation would be that clones 406C.2 and 4 contain an additional mutation, or that the host cells are resistant (i.e., "deathless") variants. Resistant variants in the 7R population seem unlikely, however, since none of the 7R.GR lines appeared resistant. These lines could also be tested for activity on GRE-containing reporters. A cell-lysis specific variant should show normal glucocorticoid induction of other GRE-containing promoters, whereas cells with a mutation in receptor or a general receptor-interacting factor would be inactive on GRE-linked promoters. Nevertheless, if we consider the glucocorticoid-sensitivity of a cell population as graded and receptor dose-dependent, then nt^i receptor may simply have reduced specific activity in general. Nt^i was indeed partially active or inactive (≤10% activity compared to wild-type) in its ability to induce four GRE-linked promoter-CAT reporters (containing a tk, AdML, MMTV, or βglobin promoter), supporting the hypothesis that nt^i has generally reduced activity.

An alternative hypothesis is that exon 2 sequences, but not enhancement domains elsewhere in the protein (i.e., *enh2* but not *enh1* or *enh3*) interact with factors that regulate cell lysis gene(s). By this hypothesis, steroid resistance in nt^i cells is not related to its generally reduced activity, but to the loss of cell-lysis activating sequences in exon 2. Thus the protein would still be partially active on other promoters (through enhancement activities downstream of exon 2 sequences) but inactive on cell lysis genes. The notion that only exon 2-encoded sequences can activate cell-lysis is unsupported, however, by the observation that overexpression of amino-terminally truncated receptor in S49 r^- cells showed partial sensitivity in some clones (406C.5,6,7). The dexamethasone sensitivity of these clones indicates that the

amino-terminal activation domain is not essential for dexamethasone-induced cell lysis. A more definitive answer will come from identification and analysis of specific cell lysis pathway genes.

The nt^i defect appears to be *cis*-acting since full-length receptor complemented the nt^i phenotype, consistent with the results of somatic cell hybrid studies (Harmon et al., 1985; Yamamoto et al., 1976). Possible *cis*-acting mutations might include a genomic deletion, rearrangement of exon 2 encoding sequences, or sequence changes affecting the splicing signals flanking exon 2. A large DNA rearrangement was found in 55R, but not in 143R (Dieken et al., 1990), suggesting that a genomic mutation might explain the nt^i phenotype of 55R.

The two different receptor 5' leader sequences presented here (mouse A and S), as well as the predicted protein sequence of the most frequently occuring clone (equivalent to S) shown in Dieken et al. (1990), contain in-frame initiator codons. In contrast, the most frequent 5' leaders in rat (Fig. 2.18) and the 5' leader reported in human (Hollenberg et al., 1985; Zong et al., 1990) did not contain in-frame initiators. However, this could reflect tissue-specific differences in transcript expression, as suggested by data presented in chapter three (Figs. 3.4, 3.5, 3.7). In the one transcript class with a homologue found in both mouse and rat, the in-frame initiator is not conserved. The trinucleotide AUG in the mouse A leader is a CUG in the homologous rat A transcript. Since CUG is not an efficient initiator in eukaryotes, the difference between mouse and rat sequences suggests either that this codon does not serve an essential function, that only mouse has maintained a specialized function for this initiator, or that this initiator works specifically in this context. Though the A-leader is well conserved between the known mouse and rat sequences (1 mismatch in 53 nucleotides), this degree of sequence

similarity is also found in portions of the 3' untranslated region of the receptor (see chapter two; Fig. 2.19). Thus it is not clear whether conservation of A-leader sequences between rat and mouse reflects functional conservation of the sequence or simply the phylogenetic proximities of rats and mice.

One hypothesis for the nt^i 5' ATGs is that they serve a function in forming amino terminal sequences in an alternately spliced nt^i -like transcript normally expressed *in vivo*. Though these sequences appear to be used in forming the nt^i receptor, it is not clear whether there is any difference between the activities of a receptor containing these sequences and one in which they are absent (m406C). The heterogeneous hormone sensitivities of the 7R.406C lines leaves the question open. Though an nt^i - like transcript has not been found in RNA from any other source besides S49 nt^i cells, no systematic search for these transcripts has been undertaken. Another explanation for some of the 5' ATGs might be that they play roles as translational regulators in the full-length message, similar to the possible role of the upstream mini-cistron in rat transcript B, as discussed previously. These initiators are clearly used in nt^i cells to translate novel amino-terminal receptor peptides.

There is also an intriguing evolutionary explanation for why the nt^i 5' leaders might contain ATGs or open reading frames in frame with exon 3. Consider the hypothesis that exon 2 is a later addition to the nuclear receptor superfamily, perhaps the result of a single exon insertion event before the divergence of the steroid receptor and other subfamilies from the VDR and krirps-like subfamilies (which lack an amino-terminal domain). By this hypothesis, the ancestral receptor lacked exon 2, so exon 1 served as the first coding exon. Though the amino-terminal coding exons in the steroid receptor subfamilies have diverged considerably from each other in their sequences

and sizes, they are all encoded by a single exon, and all maintain the same splice junction phase between exons 2 and 3. The exon 2-deleted receptor transcript in nt^i could be interpreted as an ancestral form, the upstream ATGs remnants of this ancestral receptor's initiator codons. Their survival in present-day genes would make sense if they were co-opted for regulatory purposes, while others might have been lost or weakend (by mutation of the initiation context). This hypothesis is difficult to test. Lacking the same selection pressures as well as possibly acquiring new regulatory functions, the 5' leaders of the different steroid receptors are likely to have diverged considerably from each other.

Together, the results presented in this chapter demonstrate that AUGs in a family of 5' exons are used to express receptors with novel amino-terminal sequences. In cell lysis assays I was unable to discern any difference between receptors containing these sequences (nt^i receptors) and receptors lacking these sequences (m406C). The resistance of nt^i cells to glucocorticoid-induced cell lysis thus appears to be caused by the loss of sequences encoded by exon 2. More detailed analyses of nt^i receptor expression or activity, however, may be able to determine a function for the upstream initiators used in these receptors.

Chapter Five: An activation function within the signal transduction domain

Introduction

In mouse lymphoma cells, deletion of the amino terminus of the receptor reduces its activity sufficiently to protect the cell from hormone-induced lysis. This observation is consistent with earlier experiments in COS-7 and HTC cells using transient and stable transfections of glucocorticoid receptor deletion mutants (Miesfeld et al., 1987; Rusconi et al., 1987). These experiments had shown that an amino-terminal deletion mutant, in which the DNA and signal transduction domains were left intact, retained 5-10% of the specific activity of full-length receptor. Approximately the same activity was also detected in the DNA-binding domain alone (aa407-556) (Miesfeld et al., 1987). Subsequently, Godowski et al. (1988a) fused a heterologous DNA-binding domain, LexA, to portions of receptor to map its enhancement domains independently of the receptor's own DNA-binding domain. They mapped an enhancement domain in the amino terminus (*enh2*), but found little or no enhancement activity associated with the signal transduction domain. Thus there appeared to be two important enhancement domains, one co-localized with the DNA-binding domain (*enh1*), and another in the amino terminus (*enh2*), but none associated with the signal transduction domain.

In contrast, Webster et al. (1988) and Hollenberg and Evans (1988) both reported significant enhancement activity associated with the C-terminus. Their experiments were similar to those of Godowski et al. (1988a), but they used the Gal4 DNA-binding domain and human glucocorticoid receptor in their fusion constructs instead of LexA and rat receptor. The cell types, reporter constructs (binding sites, promoters), and transfection conditions

were also different in these three experiments. I decided to investigate the discrepancy between these LexGR and GalGR fusion experiments.

Enhancement activity associated with the signal transduction domain would be interesting for several reasons. Its existence would be consistent with the hypothesis (chapter four) that an ancestral steroid receptor lacked an amino-terminal domain. Also, multiple enhancement domains is compatible with the idea that receptor acts differently on different target genes because of distinct specificities of its enhancement domains. Promoter-specific activity of different enhancement domains has been shown in progesterone receptor (Tora et al., 1988). The differences between the activities of different signal transduction domain-containing fusion proteins might also reveal important functional interactions between the DNA-binding and signal transduction domains. To examine these possibilities, I constructed GalGR expression plasmids identical to the LexGR expression plasmids in Godowski et al. (1988a) except for substitution of the Gal4 DNA-binding domain (aa1-74) for the LexA DNA-binding portion (aa1-87). I also constructed a Gal reporter plasmid with convenient restriction sites to enable the exchange of different functional elements (binding site, promoter).

In this chapter I describe preliminary results suggesting that rat receptor - Gal4 fusion constructs containing the signal transduction domain possess enhancement activity. I show that a Gal4-GR C-terminal domain fusion protein is able to induce a reporter containing five Gal-17mer binding sites at least 500-fold in response to dexamethasone, and appears to be as active as a Gal-GR fusion construct containing both the N-terminal and signal transduction domains.

Results

In order to determine whether a Gal4-rat GR fusion protein could activate transcription, and to compare this protein with the Lex-rat GR fusion protein, I constructed GalGR expression and reporter plasmids. The expression plasmids (Fig. 5.1 A,B) are identical to the LexA-GR fusions studied by Godowski et al. (1988a), with the Gal4 DNA-binding domain (aa1-74) substituting for LexA (aa1-87). The reporter plasmid $UAS_G(5x17)$-$\beta_m CO$ (Fig. 5.1C) contains five tandemly repeated Gal binding site 17-mers upstream of the mouse β globin promoter driving the CAT gene. The Lex reporter pXBCO in Godowski et al. (1988a) contained a single Lex operator upstream of a rabbit β globin promoter (Fig. 5.1D).

As shown in Fig. 5.2, GalC and GalGRΔZ displayed similar activities on the Gal UAS reporter when the expression and reporter plasmids were transiently co-transfected in CV1 or HeLa cells. Table 5.1 summarizes these results and compares them to the results of Hollenberg and Evans (1988), Webster et al. (1988) and Godowski et al. (1988a). The data in Fig. 5.2 are consistent with the activities of Gal-GR signal transduction domain fusion proteins (Hollenberg and Evans, 1988; Webster et al., 1988), but are in contrast to the LexA-based plasmids, in which the signal transduction domain fusion tkLxC showed little or no activity in CV-1 cells when co-transfected with pXBCO (Godowski et al., 1988a). Others have shown that Gal(1-74) does not enhance transcription nor stimulate CAT activity in similar UAS-containing reporters in mammalian cells (Webster et al., 1988); however, the recent discovery of a Gal11P mutant that increases transcriptional activity by Gal(1-74) suggests that Gal(1-74) may directly interact with other transcriptional activators in yeast (Himmelfarb et al., 1990). Thus it is still not clear whether

the activity of GalGR fusions is due to direct interactions between the Gal zinc-binding domain and other proteins to contibute to the transcriptional activity of Gal4, or to the presence of an enhancement activity (*enh3*) in the signal transduction domain of glucocorticoid receptor.

DNA-binding domain	Signal transduction domain	Binding site	No. sites	Promoter	Cell	Rel. Activity[23] (ref.)
gal (1-74)	rat (525C)	galUAS-17	5	β glob (m-106)	HeLa	+++ (5.2)
"	"	"	"	"	CV-1	+++ (5.2)
"	hum (r507C[24])	"	1	MTV	CV-1	++ (H)
"	hum (r519C)	"	2	β glob (rb-109)	HeLa	≥+[25] (W)
lex (1-87)	rat (524C)	lex OP	1	β glob (rb-125)	CV-1	+/- (G)

Table 5.1. Comparison of gal- and lex-GR signal transduction domain fusion experiments. References: (5.2) from Fig. 5.2; (H) Hollenberg and Evans (1988); (W) Webster et al. (1988); (G) Godowski et al. (1988a).

[23]Compared to GR-gal-GR or equivalent. +++= 30-100%; ++ = 10-30%; + = <10%; - = no activity.

[24]human amino acid positions are given as their homologous rat positions.

[25]Not compared to full-length gal fusion. > 50% of the activity of full-length GR on the MTV GRE/promoter.

Figure 5.1. Gal- and Lex-GR fusion protein expression plasmids and reporters. **(A)** GalC was constructed by substituting Gal(1-74) (courtesy of Ed Giniger) for tkLex(1-87) in tkLxC (Godowski et al., 1988a). Predicted amino acids at the Lex or Gal / GR junction are indicated; newly created amino acids are in italics. The tkLex constructs contain a tk leader peptide fused to the start of Lex(1-87) to improve translation efficiency (Godowski et al., 1988a). The Gal-GR and Lex-GR fusion constructs were both expressed from the RSV enhancer/promoter in the p6R vector (Godowski et al., 1988a). **(B)** GalGRΔZ was constructed from tkLxGRΔZ (Godowski et al., 1988a). **(C)** pUC-$UAS_{G\ 5x17}B_mCO$ was constructed by inserting a 5x tandemly repeated Gal4 UAS (17mer) (Giniger et al., 1985) and mouse β globin promoter fragment (courtesy L. Stuve) into the CAT promoterless reporter pUC-OOCO. **(D)** pXBCO (Godowski et al., 1988a) contains a single Lex operator (Brent and Ptashne, 1981) inserted into the rabbit β globin promoter.

Figure 5.1

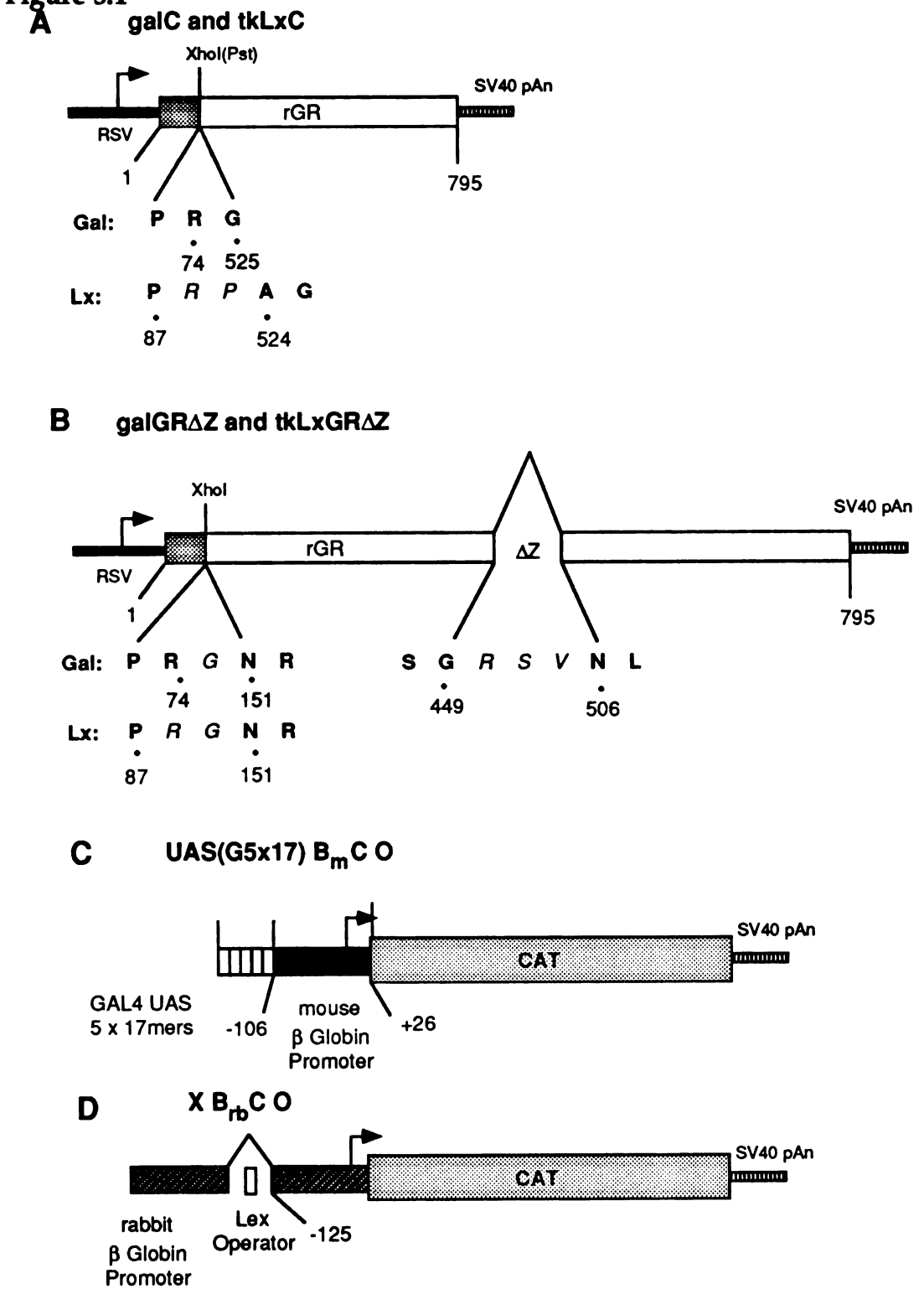
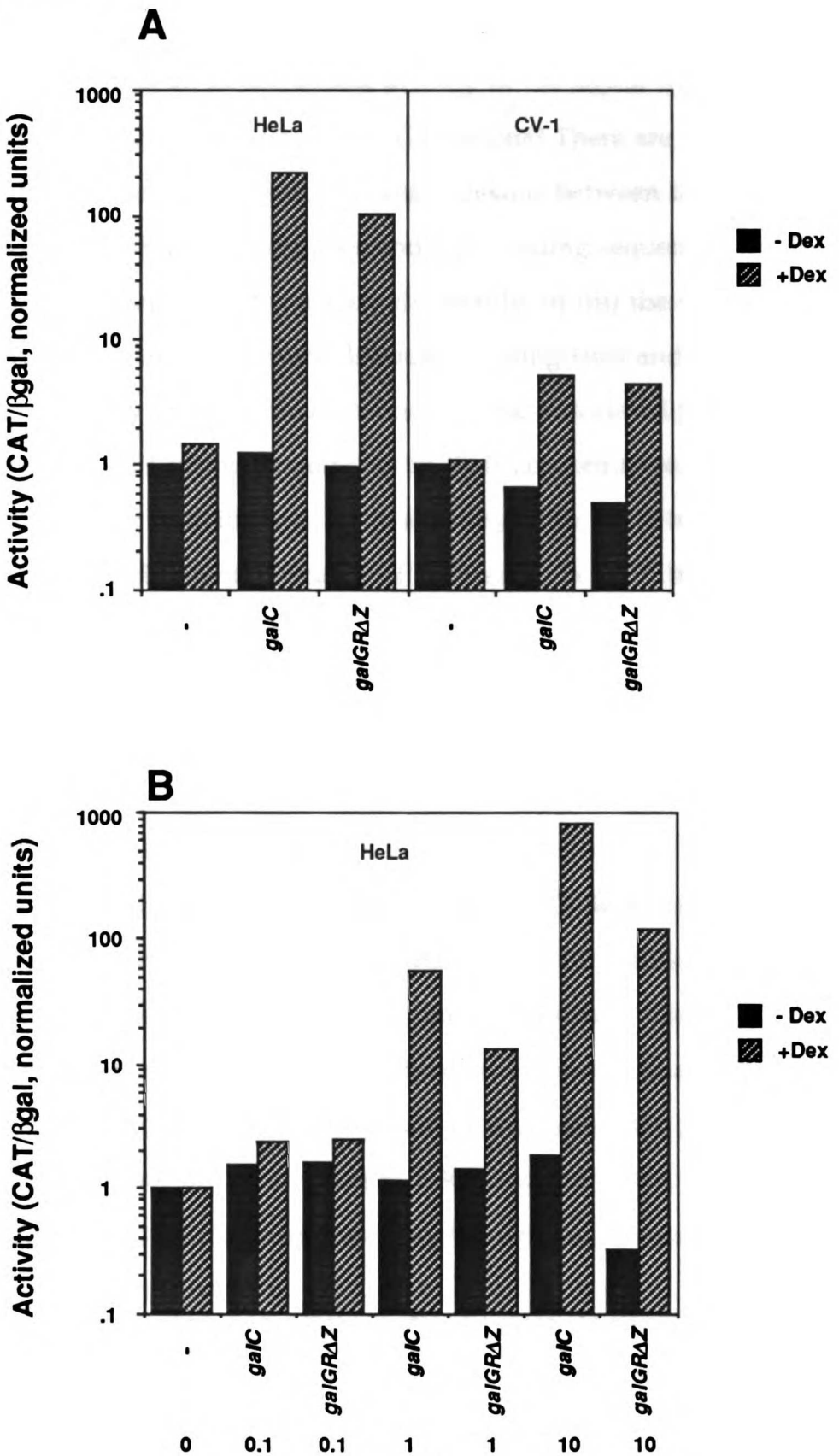

A
galC and tkLxC
XhoI(Pst)
SV40 pAn
rGR
RSV
1
795
Gal: P R G
74 525
Lx: P R P A G
87 524
B
galGRΔZ and tkLxGRΔZ
XhoI
SV40 pAn
rGR
ΔZ
RSV
1
795
Gal: P R G N R S G R S V N L
74 151 449 506
Lx: P R G N R
87 151
C
UAS(G5x17) BmC O
SV40 pAn
CAT
GAL4 UAS
5 x 17mers
-106
mouse
β Globin
Promoter
+26
D
X BrbC O
SV40 pAn
CAT
rabbit
β Globin
Promoter
Lex
Operator
-125

Figure 5.2. Cat activity of Gal-GR fusion proteins in HeLa and CV1 cells. Activities (CAT/βgal) were normalized to activity with no expression plasmid (-) and no hormone. Note the log scale in both (A) and (B). Data in (A) are single observations from one experiment; (B) are averages of duplicate transfections. **(A)** Expression plasmids containing galC or galGRΔZ were cotransfected with the reporter plasmid $pGal_{5x17}\beta_mCO$ at a mass ratio of 5μg expression plasmid : 1 μg reporter plasmid (per 60mm dish) along with a lacZ expression plasmid (- = no expression plasmid). $CaPO_4$ precipitates were split into two dishes of cells, one with and one without dexamethasone (1 μM final). CAT activies were normalized to βgal activity after subtracting background (based on a mock transfection). **(B)** Same experiment as in (A), but in HeLa cells only, and the mass ratio was varied from 0.1:1 to 10:1. Note that the induction ratio (+ hormone / - hormone) increases with the amount of expression plasmid for both galC and galGRΔZ.

Figure 5.2


A
1000
100
10
1
.1
Activity (CAT/βgal, normalized units)
HeLa
CV-1
- Dex
+Dex
-
galC
galGRΔZ
-
galC
galGRΔZ
B
1000
100
10
1
.1
Activity (CAT/βgal, normalized units)
HeLa
- Dex
+Dex
-
galC
galGRΔZ
galC
galGRΔZ
galC
galGRΔZ
0
0.1
0.1
1
1
10
10

Discussion

If there is an enhancement activity in the signal transduction domain, why was it not detected in the Lex-GR fusions? There are three types of hypotheses to explain the differences in results between the receptor fusion experiments: (i) the binding sites and surrounding sequences are different, (ii) the DNA binding domains behave differently, or (iii) the cell types and/or experimental conditions were different. Binding sites and their DNA contexts might differ by virtue of their abilities to bind associated proteins, in their affinities for GR fusion proteins, or in their conformational affects on GR fusion or other proteins (e.g. DNA allostery). The DNA binding domains might behave differently because of differences in DNA binding affinity, interactions with other proteins, or in interactions with other domains of the fusion protein (e.g. that activate transcription). And experimental conditions, such as the ratio of expression to reporter plasmids or the cell lines used, have been shown to affect the activities of different steroid receptor deletion mutants (Bocquel et al., 1989).

My results (Fig. 5.2) and those of others (Table 5.1) suggest that GalC is active in both HeLa and CV1 cells under different transfection conditions and with different fusion constructs (e.g. rat and human). Though I did not test LxC myself, the experimental conditions in Godowski et al. (1988a) were not substantially different from those in my experiments (i.e., 10:1 expression plasmid/reporter in Godowski et al. (1988a) vs. 5:1 in Fig. 5.2A; both were transfected into CV1 cells and used β globin promoter-based reporters[26]). It

[26]Though both experiments used β globin promoter-based reporters (as did Webster et al. (1988)), my experiments used a mouse promoter instead of the rabbit promoter because the mouse promoter was available and was more easily usable as the basis for a shuttle plasmid. Though I obtained results similar to those of Webster et al. (1988) and Hollenberg & Evans (1988), the

thus appears unlikely that the discrepancies between GalC and LxC are due to experimental differences such as the transfection conditions, cell types, receptor sequences (i.e. species or precise fusion sequences), or promoter used. The difference between LxC and GalC activities therefore appears to be caused either by differences in the DNA-binding domains, differences in their binding sites, or both.

The DNA binding sites can be analyzed by swapping or modifying specific elements in the reporter plasmid. The most obvious difference between the Gal UAS reporter used here and the Lex reporter in Godowski et al. (1988a) is that one contains five Gal UAS 17mers and the other contains a single Lex operator. Thus, the sites in these two reporter constructs might differ in their affinities for their respective GR chimeras, and affect the activities observed. Indeed, a different Lex reporter with a Lex operator containing two binding sites has a higher affinity for LexA than a single-site operator (Ebina et al., 1983; Kamens et al., 1990). It is also possible that different enhancement domains behave differently with respect to the multiplicity of binding sites. I constructed the plasmid pUAS(5x17)-βCAT to enable binding site and promoter swapping. Controlled comparisons of binding site multiplicity and type as well as promoter context should be straightforward with this plasmid and its derivatives. One experiment that should be done is to test LxC on a Lex reporter containing multiple binding sites. Similarly, GalC should be tested on a single Gal UAS 17mer reporter. Ideally, these experiments should be run as part of a set of parallel transfections using a series of reporter plasmids identical except for the

---

LexA fusion experiments (Godowski et al., 1988a) should be repeated with a LexOP version of the UAS β_mCAT reporter to make sure that the mouse and rabbit β globin promoters do not behave significantly differently.

multiplicity and type of binding site. These experiments should also include controls for fusion protein expression and DNA-binding.

Another important difference between the Lex and Gal protein fusion systems is that the DNA binding domains are structurally distinct from each other and from the receptor zinc-finger. The LexA DNA-binding domain behaves similar to the helix-turn-helix binding motif of phage repressors (Brent and Ptashne, 1981, 1985), while Gal4 DNA binding region is based on a C_6 binuclear metal ion complex (Pan and Coleman, 1990a,b) and the glucocorticoid receptor DNA-binding domain consists of two zinc fingers with a C_4/C_5 structure (Freedman et al., 1988; Härd et al., 1990a,b). It should not be surprising that the glucocorticoid receptor DNA binding domain is not as independent of the signal transduction domain as a simple modular model of the receptor would imply. Linker-scanning mutants in the receptor DNA binding domain have been isolated that specifically affect receptor action on a positive and not a negative GRE, and vice versa (Godowski et al., 1988b), suggesting that interdomain interactions might be important in enhancement function. By substituting a heterologous DNA binding domain for the receptor DNA binding domain, the activity in *enh3* might be altered depending on the structure of the DNA-binding domain. This modulatory interaction might be influenced also by other context effects such as the binding site, promoter, or cellular environment. It is also possible that Gal(1-74), Lex(1-87), and GR(440-525) each interact in different ways with other proteins and thus directly affect enhancement activity. Indeed, Gal(1-74) and the zinc-binding domains of other yeast C_6-type transcriptional activators may interact with another protein, Gal11, to activate transcription (Himmelfarb et al., 1990; Nishizawa et al., 1990).

One way to analyze *enh3* would be to mutagenize the zinc-finger and signal transduction domains, in both the ΔN-terminal and full-length contexts as well as in heterologous fusions. These experiments might also help characterize possible interactions between *enh3* and the DNA binding domain. The strong hormone induction of UAS(5x17)βCAT by GalC in HeLa cells suggests that this system could be useful in analysis of the signal transduction domain and of the putative activation domain *enh3*. Interpreting these mutagenesis studies is likely to be difficult, however, because of the presence of multiple overlapping functions in both domains. It would therefore be essential to assay for other activities in addition to enhancement (e.g. hormone binding, DNA-binding, nuclear localization). *Enh3* might also be distributed over multiple portions of the C-terminus. Attempts to map the enhancement activity in the hormone-binding domain of the estrogen receptor, for example, failed to localize the activity to a single exon (Webster et al., 1989).

Based on my results and those of others (as presented above), I believe that there is an enhancement activity in the signal transduction domain. I would also predict that LxC is much less active than GalC because the single-site reporter pXBCO is much weaker than the multi-site Gal UAS reporters. The reason for this difference between reporters might be that the single Lex operator has a much lower affinity for LxC than the affinity between the multimerized Gal UAS site and GalC. Alternatively, multiple binding sites might allow synergistic activation by their cognate binding proteins. In any case, by my hypothesis LxC activity should become apparent on induction of multiple Lex site-containing reporters. Indeed, preliminary results indicate that this is the case (D. Pearce, pers. comm.). It is not clear, however, how to interpret the relative activities of full-length and N-terminal deletion

constructs without knowing the relative DNA-binding affinities of these proteins. N-terminally deleted glucocorticoid receptor species, for example, display reduced affinity for GRE sequences when compared to full-length receptor (Payvar and Wrange, 1983).

The existence of enhancement activities associated with the zinc-finger and signal transduction domains of the glucocorticoid receptor is interesting when compared to other nuclear receptors such as ER and VDR, which do not carry amino-terminal transcriptional enhancement domains. It is possible that the ancestral nuclear receptor contained only the zinc-finger and signal transduction domains, and posessed transcriptional enhancement activity. In addition to these evolutionary implications, it will be important to characterize these enhancement domains to determine whether they have cell- and promoter-specific functions. Multiple context-specific enhancement activities may help explain the diverse effects that glucocorticoids have on their target sites.

Chapter Six: Conclusions and Perspectives

Steroid receptor evolution

The receptors for steroid hormones, retinoids, and thyroid hormones are members of a superfamily of zinc-finger containing proteins. Each of these proteins contains a zinc-finger motif consisting of one C_4 followed by one C_5 zinc finger. The zinc-finger domains bind specific DNA sequences called hormone response elements (HREs). These proteins also contain a C-terminal domain similar in sequence to the glucocorticoid receptor's signal transduction domain. This domain binds to receptor-specific ligands and mediates several processes involved in transducing the hormone signal. Some of these proteins (including those in the steroid receptor subfamily) contain additional N-terminal sequences that contribute to receptor enhancement activity, and may have other functions as well. The relative positions of these domains and their organization into discrete exons is well conserved, especially within receptor subfamilies.

The degree of gene structure similarity, defined by the precise intron-exon organization (Fig. 2.5), closely parallels the similarities in protein sequences (Fig. 2.6; Amero et al., 1991). Analysis of nuclear receptor protein sequences indicates that the zinc-finger and C-terminal domains have evolved in parallel (Amero et al., 1991), consistent with the conservation of introns in both domains (Fig. 2.6). In addition, comparison of the zinc-finger domain with sequences in the PIR protein database suggests that the nuclear receptor zinc-finger is evolutionarily distinct from other protein sequences including other types of zinc-binding domains (e.g. C_2H_2, C_6; Amero et al., 1991). Together these observations indicate that these genes have not undergone major rearrangements involving their zinc-finger and C-terminal domains (the evolution of the N-terminal domain is less clear) and that the

nuclear receptors have evolved by a pattern of gene duplication and divergence from an ancestral ligand-binding C_4/C_5-type zinc-finger regulatory protein.

The similarities in overall gene structures and in protein sequences between vertebrate and *Drosophila* nuclear receptor proteins place the emergence of the ancestral receptor to a time before the divergence of vertebrates and arthropods. Indeed, the presence of steroids and their specific binding proteins in fungi (Burshell et al., 1984; Feldman et al., 1982; Loose et al., 1981), including some with a hormonal role (Timberlake and Orr, 1984), suggests that an ancestral steroid hormone receptor was present in early eukaryotes, prior to the divergence of fungi and animals. It will be particularly interesting to determine whether these fungal steroid binding proteins are related to the nuclear receptor superfamily defined in animals. If this superfamily arose from an ancestral family of terpenoid-binding regulatory molecules (Moore, 1990), additional superfamily members should be found in the receptors for other terpenoid signalling molecules.

If this ancestral protein arose by the recombination of zinc-finger and ligand-binding signal transduction motifs, some similarities might be apparent between these motifs and portions of otherwise unrelated or distantly related proteins. For example, the low-density lipoprotein (LDL) receptor is a mosaic of functional domains encoded by separate exons and also found in proteins from other supergene families; one domain, the "EGF repeats", is also present in the epidermal growth factor precursor and in three proteins of the blood clotting system (Südhof et al., 1985a,b), as well as in many other proteins. Such similarities between nuclear receptor superfamily proteins and other proteins may indicate how this ancestral protein arose. For example, if a terpenoid-modulated gene regulator arose from a simple

metabolic regulator, a similar terpene-binding motif might be found in enzymes or carrier proteins involved in isoprenoid metabolism. Similarly, the zinc-finger motif might be similar to other tetrahedrally coordinated metal-binding proteins, or to other DNA-binding proteins. The sequences of the zinc-finger and signal transduction domains, however, do not appear to be similar to non-superfamily sequences present in current databanks, including other types of zinc-finger (Amero et al., 1991). Thus if a combining event happened (e.g., exon shuffling between zinc-finger and ligand binding domains), it happened once, early in the evolution of nuclear receptors.

It is possible that distant relatives have been sequenced but have diverged considerably as they adapted to different functions. Sequence similarities within a shorter (17aa) segment have been noted between steroid-binding regions of steroid receptors and steroidogenic enzymes (Picado-Leonard and Miller, 1988). Structural similarities have also been suggested between the steroid binding domains in the steroid receptor superfamily and members of the serpin superfamily (corticosteroid binding globulin (CBG) and α_1-antitrypsin; Mornon et al., 1989). More meaningful comparisons of shorter (and possibly better conserved) sequence motifs will become possible as functional domains and subdomains are defined more precisely.

Transcription factor multigene families

To date over 100 eukaryotic genes encoding DNA-binding regulatory proteins have been discovered and sequenced. Groups of these proteins appear to share distinct, highly conserved domains involved in DNA-binding (for reviews see (Johnson and McKnight, 1989; Struhl, 1989). Many of these genes are members of large multigene families. The genes for the C_4/C_5 zinc-

finger containing nuclear receptors comprise one such family (reviewed in (Evans, 1988; Green and Chambon, 1988). Genes for other types of zinc-binding proteins appear to be members of other families, and contain distinctly different structural motifs (e.g. the C6 binuclear cluster of gal4 (Pan and Coleman, 1990a,b) and the C_2H_2 class of TFIIIA-like zinc-fingers (Miller et al., 1985; see reviews by Johnson and McKnight, 1989; Mitchell and Tjian, 1989). Other families of DNA-binding regulatory proteins include the families of leucine-zipper-containing proteins (the Jun, Fos, and C/EBP families; see Abate and Curran, 1990; Vogt and Bos, 1990), the large family of homeodomain-containing proteins (including over seventy sequenced members; see reviews by Affolter et al., 1990; Scott et al., 1989), the Helix-Loop-Helix (HLH) family proposed by Baltimore and colleagues (e.g., *MyoD*, *myc*, *daughterless*; Murre et al., 1989a,b), the recently discovered *rel* family (*rel*, NFκB, *dorsal*; Ghosh et al., 1990; Kieran et al., 1990), and a putative *myb* family (Cole, 1990). Other families are likely to emerge as more transcription factors are discovered and cloned. The groupings of these proteins into evolutionary families is based on conserved sequence similarities; however, structural and/or molecular genetic studies of the nuclear receptor, homeodomain, and jun families provide additional support for the hypothesis of a common evolutionary history for these proteins (see reviews cited above).

For example, it has been suggested, based on similarities to prokaryotic and yeast DNA-binding domains, that the homeodomains contain a helix-turn-helix motif which is similarly essential for sequence-specific DNA-binding (Laughon and Scott, 1984; Shephard et al., 1984). This hypothesis has been confirmed by recent structural studies of the Antennapedia (*Antp*) (Quian et al., 1989) and Engrailed (*en*) (Kissinger et al., 1990) homeodomains.

Structural predictions and genetic studies of other homeodomains also support this hypothesis (Desplan et al., 1988; Treisman et al., 1989). Consistent features of homeodomain-containing proteins are considerable conservation of predicted secondary structure of their homeodomains and nuclear localization of the proteins (Scott et al., 1989). A large number of homeobox[27]-containing genes in *Drosophila* (>20) and non-*Drosophila* species (>50) have been sequenced. Phylogenetic analysis by protein parsimony of the *Drosophila* proteins' homeodomains suggests that these proteins have evolved by the duplication and divergence of a common ancestral protein; in addition, proteins closely related on the phylogenetic tree share functional attributes such as related protein domains (other than the homeodomain), expression patterns, genetic loci, or functional roles (Scott et al., 1989).

The characteristics of the homeodomain superfamily are reminiscent of the nuclear receptor superfamily: the family contains many members (>30) which appear to have diverged from a single common ancestor, and subsets of proteins closely related in their DNA-binding homeodomains also share other sequence similarities outside that domain (Scott et al., 1989). The homeodomain also appears to be a functional unit: homeodomains alone are capable of sequence-specific DNA-binding, the homeobox is usually found as a separate exon, homeodomain sequences have been more highly conserved during evolution than the surrounding protein sequences, and functional chimeric proteins can be constructed that exchange homeodomains and surrounding sequences (Scott et al., 1989).

The leucine-zipper motif suggests another superfamily. Originally proposed for the C/EBP protein (Landschulz et al., 1988), the leucine zipper

---

[27]For historical reasons, homeobox refers to the nucleic acid sequence of homeodomains.

domain is responsible for dimerization with other leucine-zipper containing proteins by a parallel coiled-coil interaction (O'Shea et al., 1989). This dimerization appears to be essential for DNA-binding and transcriptional activation by these proteins (for reviews see Curran and Vogt, 1991; Johnson and McKnight, 1989). These proteins also contain a sequence of basic residues adjacent to the leucine-zipper, thought to be directly involved in sequence-specific DNA binding. Many of these basic-leucine-zipper (bZip) proteins share similarities in other protein regions, and so are grouped into families of related genes (e.g. *jun* (*jun, junB, junD,* CREB); *fos* (*fos, fra-1, fosB*). Though each group does not share protein similarities with other groups outside the bZip region, it is possible that the discovery of additional family members with intermediate sequences in other domains may suggest evolutionary links between these families (by comparison, the most distant members of the nuclear receptor superfamily do not appear to share any sequence similarities (<15% identity) outside the zinc-finger domains). Alternatively, these families may be mosaics of protein domains created by exon shuffling events. Nevertheless, the ability of some of these proteins to form functional heterodimers with each other, as well as their similar physiological roles in early reponses to environmental stimuli, suggest a shared evolutionary relationship.

Why are DNA-binding transcriptional regulators organized into large multigene families? As suggested for the nuclear receptors (Amero et al., 1991) and homeodomain-containing proteins (Lewis, 1978; Scott et al., 1989), each gene family may have evolved from a common ancestral gene. Consider the evolution of eukaryotic organisms. The ancestral eukaryotic cell contained a distinct set of regulatory proteins. The number of these transcriptional regulators may have been small (compared to the number of

modern regulators), reflecting the relative simplicity of this single-celled organism. As this organism evolved into more complex multicellular species, its regulatory controls must have become more complex. Some of this increased regulatory complexity may have involved an increase in the number of transcriptional regulators by the duplication and divergence of existing regulatory proteins. Indeed, it is possible to change the DNA-binding specificity (and hence target-gene specificity) of nuclear receptors by changing a few amino acids, apparently without altering the overall structure of the protein. Thus it is easy to imagine how duplicate copies of regulatory proteins might have diverged to accomplish distinct regulatory functions. Modern multigene families of transcriptional regulators may thus be a reflection of the complexity of modern multicellular eukaryotic organisms and the need for more complex regulatory networks involving greater numbers of regulators.

The glucocorticoid receptor gene is complex, suggesting multiple strategies for regulation

The glucocorticoid receptor gene is organized into eight coding exons distributed over at least 125 kilobases of DNA. It contains two polyadenylation sites and at least eight alternative 5' untranslated exons, transcribed from at least two promoters. This organization is reminiscent of another complex regulatory gene, the Antennapedia (*Antp*) gene of *Drosophila melanogaster*. Similar to the receptor gene, the *Antp* locus is long (>100kb), is transcribed from two alternate promoters, and contains two alternate polyadenylation sites (Laughon et al., 1986; Schneuwly et al., 1986; Stroeher et al., 1986). The alternate upstream exons of this gene encode two long 5' mRNA leaders, each

fused to a common downstream exon and the long protein open reading frame. Each of these 5' leaders contains multiple AUGs. Antennapedia promoters P1 and P2 have overlapping, but distinct spatial and temporal patterns of expression in embryos and imaginal disks (Jorgensen and Garber, 1987). Though the function of this promoter arrangement in the *Antp* gene is unclear, there are other examples in which alternate promoters are clearly involved in developmental stage-specific expression (as in the *Drosophila melanogaster* alcohol dehydrogenase gene; Savakis et al., 1986) or in tissue-specific espression (such as the mouse α-amylase promoters; Schibler et al., 1983). The Antennapedia and receptor genes are not unique in their complexities. Many other regulatory genes in Drosophila (Affolter et al., 1990; Scott et al., 1989) and mammals (e.g. jun; Vogt and Bos, 1990) appear to have complex regulatory regions; these most likely reflect the complexities of the controlling regulatory networks.

What is the function of the multiple 5' exons and promoters in the glucocorticoid receptor gene? Do they play a role in receptor regulation? I identified one promoter, promoter B, immediately upstream of exon 1B. Transcript B, the product of this promoter, was expressed in all nine tissues (brain, liver, heart, lung, kidney, testes, salivary gland, thymus, spleen) and two cell lines (hepatocyte and fibroblast-derived) surveyed. It appeared to be the most abundant receptor mRNA species (60 - 75% of the total) in all but the lymphoid tissues (thymus and spleen), and contained multiple initiation sites. The sequence of promoter B was rich in CpG dinucleotides and G + C content, contained multiple consensus sequences for transcription factor SP-1 and a weak consensus sequence for transcription factor CTF, yet contained no TATA box. These features appear to be typical of many eukaryotic promoters, particularly the promoters of essential "housekeeping" and growth-control

genes (Hoffman et al., 1987; Ishii et al., 1985a,b; Melton et al., 1984; Reynolds et al., 1985; Valerio et al., 1985) , suggesting that promoter B contributes to the near ubiquity of receptor expression.

Since this promoter was found downstream of at least one other 5' exon, there must be at least one more distal promoter. What is the function of these other promoters? One possibility is that they are tissue-specific, perhaps mediating tissue-specific regulation of receptor mRNA. It is known, for example, that receptor protein and mRNA expression in the brain is heterogeneous among neuronal types and brain regions (Aronsson et al., 1988; Fuxe et al., 1987; Gustafsson et al., 1987) and may be regulated in distinct ways by developmental and environmental cues (De Kloet et al., 1988; Kalinyak et al., 1989; Meaney et al., 1985; Rosenfeld et al., 1988). There is also evidence that receptor mRNA levels are feedback-regulated by glucocorticoids differently in a T-cell line (where it appears to be up-regulated) than in most other cells and tissues (where down-regulation seems to be the rule). Indeed, I have observed positive autoregulation of thymus receptor mRNA *in vivo*. Though I was unable to confirm negative autoregulation in other tissues, it might be possible to detect small differences between the control and hormone-treated groups with larger sample sizes.

Receptor mRNA expression in lymphoid tissue appears to be novel in several respects. The thymus and spleen contained a greater proportion of non-B transcripts (particularly τ) and less B compared to other tissues. Acute glucocorticoid treatment appeared to accentuate these differences in the thymus, but not in the spleen. Not only did glucocorticoids increase receptor mRNA levels in the thymus, they did so by increasing selectively the proportion of group τ transcripts. These thymus-specific differences were observed in both young (2d) and adult (28d) animals. Thymus tissue consists

almost entirely of immature T-cells, whereas the spleen contains a mixture of B and T cells. Together, these observations suggest that receptor mRNA is up-regulated by a T-cell-specific promoter.

To test this hypothesis, it will be necessary to identify a putative T-cell specific promoter and determine whether it is active in T-cells and responsible for T-cell specific mRNAs. It will also be important to test whether this promoter activity is up-regulated by glucocorticoids, as the observations presented in Chapter Three suggest. Since the thymus tissue from control animals had been exposed to endogenous circulating glucocorticoids, it is possible that the lymphoid pattern of receptor mRNA expression (Figs. 3.5, 3.7) was caused by T-cell-specific glucocorticoid up-regulation of τ transcripts.

This hypothesis could be tested by comparing thymus mRNA from hormone-treated, control, and adrenalectomized animals. It would also be interesting to know whether B transcripts are hormone-regulated in T-cells (as discussed above, the data were insufficient to discern hormone-dependent changes in B levels in any of the tissues). Finally, it is not clear whether these changes in mRNA reflect transcriptional effects or altered mRNA stabilities.

What could be the function of a T-cell specific promoter? Since certain T-cells are killed by glucocorticoids, receptor up-regulation might be a mechanism of reinforcing the response to an initial increase in glucocorticoid levels by lowering the hormone concentration threshold for induction of cell lysis. Indeed, a threshold effect is suggested by the reciprocal relationship between receptor levels and hormone concentration necessary to induce cytolysis in some mouse lymphoid lines (Bourgeois and Newby, 1977, 1979; Danielsen and Stallcup, 1984). Positive autoregulation of the receptor gene might be responsible for prolonging transient stimuli, as has been suggested

for autoregulation of Jun by its product AP-1 (Jun/Fos) (Angel et al., 1988). Positive autoregulation has also been observed in *Drosophila* homeodomain-containing pair-rule genes (Goto et al., 1989; Harding et al., 1989; Hiromi and Gehring, 1987; Ish-Horowicz et al., 1989). This form of regulation may provide a mechanism for maintaining a stable developmental state, similar to the maintenance of the lysogenic state of λ by lambda repressor.

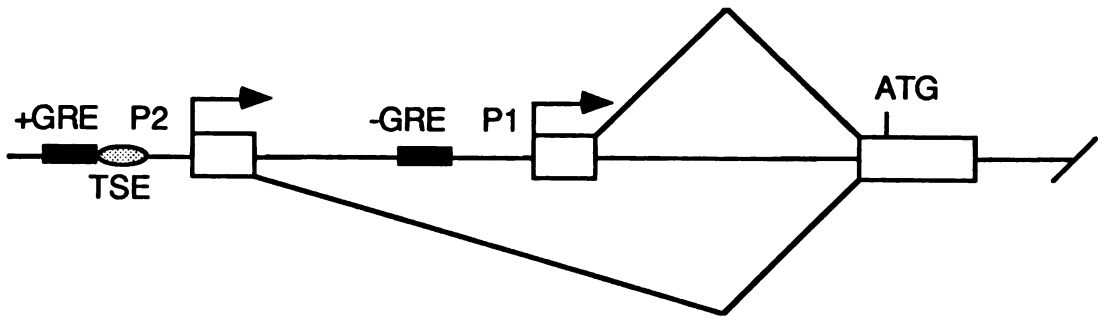
A T-cell-specific promoter might function to ensure that this positive autoregulation occurs only in certain T-cells. This regulation could then function without affecting other promoters (such as B); this speculative scenario is illustrated in Fig. 6.1. In this model, I consider two sets of promoters: a ubiquitously expressed, negatively autoregulated promoter (P1) and a T-cell specific, positively autoregulated promoter (P2). Two potential regulatory schemes are shown. Independent GREs may be associated with each promoter (model A), or one GRE may regulate both promoters (model B). In model B, a composite GRE could confer both positive and negative regulation (). It is unknown whether a single composite GRE can confer positive and negative regulation on two independent promoters simultaneously. Alternatively, P1 might be unaffected by glucocorticoids in T-cells. The GREs in model A could similarly be composite GREs, but the P2 GRE (positive) is enabled only in T-cells when P2 is active, perhaps because of a T-cell specific element (TSE).

To test these models, it will be necessary to determine whether the observed changes in transcript levels are caused by transcriptional or post-transcriptional mechanisms and to identify and characterize promoters P1 and P2. Promoter B is a good candidate for P1; the promoter(s) responsible for transcripts in group τ might behave like P2. Sequences upstream of promoter

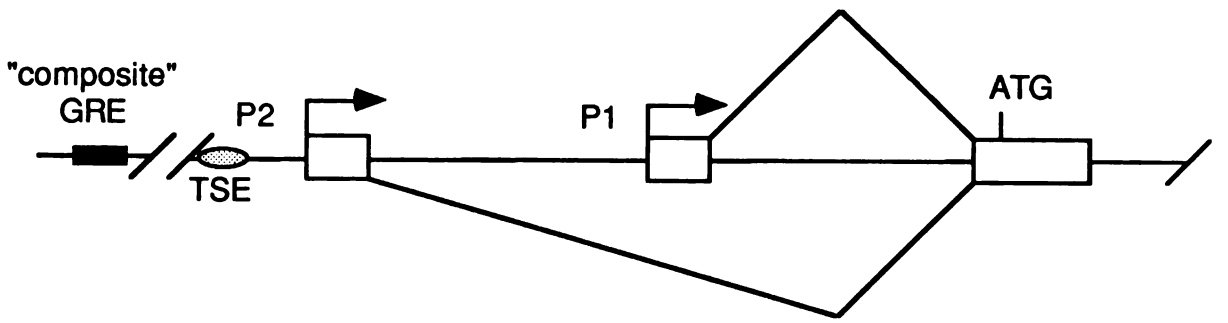
Figure 6.1 Two possible models to explain ubiquitous mRNA expression, general negative autoregulation, and T-cell specific positive autoregulation.(A) Independent GREs. **(B)** Single GRE. **(C)** Predicted expression pattern.

Figure 6.1

A



B



C

	T-Cells		Other Cells	
hormone:	-	+	-	+
P1	++	+ or ++	++	+
P2	-	+++	-	-

B do not reveal any apparent GREs, but this is not surprising since composite GREs do not appear to share a consensus sequence.

Since there are more than two alternate 5' mRNA leaders, there may be more than two promoters. These promoters might have distinct regulatory functions. Several questions are worth investigating more thoroughly. For example, does receptor mRNA expression in specific brain regions or cells correlate with 5' mRNA leader usage? The hippocampus in particular is an area in which receptor mRNA and protein are expressed at high levels relative to other brain areas, are localized to the specific layers, and may play a role in glucocorticoid regulation of the HPA axis. Another region of interest is the pituitary intermediate lobe, where receptor mRNA is expressed, but protein expression appears to be tonically inhibited by dopamine. Is this dopamine regulation mediated by translational control, and if so does this regulation involve a specific receptor 5' leader? 5' leader B appeared to affect receptor translation (chapter three); it would be interesting to know if this translational effect can be regulated by other signalling molecules. These questions should be addressed by first identifying and cloning the 5' leaders expressed in specific tissues and conditions. These 5' leaders could then be used as transcript-specific (and presumably promoter-specific) probes in assays of mRNA levels and transcription rates.

Alternate promoters and translational control elements suggest possible routes to modulate receptor activity by regulation of receptor levels. Another regulatory mechanism is to modulate the activity of the protein directly, by covalent modification or by expression of differently spliced forms. An intriguing idea is that the exon 2-deleted transcripts in nt^i cells are clues to an as yet undiscovered alternative splicing pattern. Since it is not clear what receptor mutation(s) are responsible for the modified receptor message, this is

still a possibility. Such an alternately spliced receptor mRNA has not been found *in vivo* during the course of receptor mRNA analysis (e.g. Northerns); however, an nt^i-like mRNA might be a minor species in a subfraction of cell types during development. (Alternatively, the nt^i phenotype might be peculiar to S49 cells). Nevertheless, it is curious that the most frequently cloned mouse 5' leaders from nt^i cells (S and A) contained in-frame initiator codons. Though the 5'ATG in rat A is not conserved with mouse, it would be useful to clone the rat or human homologue to mRNA S, which appears to be the more abundant mRNA species in nt^i cells, to determine whether the S initiators are conserved. Interestingly, a progesterone mRNA species has been discovered that encodes a truncated receptor missing the amino-terminal domain and first finger; this mRNA is abundantly expressed in progesterone target cells, but its function is unknown (Wei et al., 1990). One way to look for an nt^i-like mRNA species would be to probe mRNA (or cDNA, e.g. with PCR-based methods) from various cell types with an oligonucleotide specific to the exon 1/3 junction. Since S49 is a T-cell line, it would make sense to test subfractions of T-cells from different developmental stages.

Evolution of regulatory networks

Steroid receptors, like other transcriptional regulators, control distinct sets of genes, and are themselves regulated by other transcription factors. These interacting factors form a complex regulatory network that defines the phenotype of the cell by determining which genes are transcribed and to what levels. For example, multiple, cross-regulating genes appear to control phenotype in muscle cells () and early embryogenesis in *Drosophila* (for reviews see Affolter et al., 1990; Ingham, 1988; Scott and Carroll, 1987). These

complex networks presumably evolved from simpler systems involved in essential cellular metabolic functions (Tomkin's "simple regulation"). Indeed, the homeodomain and jun families of regulatory proteins include examples of metabolic regulators in yeast: PHO2 (also known as BAS2), related to the yeast homeodomain-containing developmental regulators Matα2 and Mata1, is involved in the coordination of phosphate regulation of diverse metabolic pathways (Bürglin, 1988; Sengstag and Hinnen, 1987); and GCN4, involved in amino acid metabolism, is a close homologue of mammalian Jun protein (Vogt et al., 1989).

As these networks evolved and became more interconnected, essential interactions may have become fixed. And as complex regulators arose and adapted to new functions, new regulatory interactions in addition to these older, conserved interactions, would have become necessary.

By this argument, the complexity of the receptor gene might reflect its evolutionary history: some features, perhaps adapted to more specialized signalling roles in multicellular eukaryotes, may have been added to regulatory interactions reminiscent of simple metabolic regulators in unicellular organisms (e.g., negative feedback, ubiquitous expression). For example, the nearly ubiquitous expression of receptor, consistent with its essential physiological role in vertebrates, is reminiscent of an essential metabolic regulator. Indeed, promoter B shares many features with promoters of essential "housekeeping" genes. Similarly, negative feedback regulation is typical of many simple metabolic control circuits in prokaryotes; it also makes physiological sense as a long-term homeostatic response to varying hormone levels.

Tissue-specific regulation of nuclear receptors might be viewed as a more recent evolutionary adaptation which evolved under the constraint of

retaining generally ubiquitous expression. The absence of receptor in pituitary intermediate lobe, for example, may have been achieved by adapting a post-transcriptional regulatory mechanism to an ubiquitously expressed mRNA. Similarly, the putative thymus-specific promoter proposed above might be an example of a later addition or modification, perhaps introduced by transposition or gene rearrangement. A transpositional origin for regulatory elements was suggested by Yamamoto (1989), and has found strong support by the recent discovery of an ancient proviral LTR in the 5' region of the mouse sex-limited protein (Slp) gene (Stavenhagen and Robins, 1988). This gene is a close homologue of the complement C4 gene, but unlike C4 is androgen responsive due to the presence of hormone response elements in the proviral LTR. A number of other mouse genes also display androgen responsiveness, not present in related genes or other rodent species (Crank-Tseng and Berger, 1987; Harper et al., 1980; Latimer et al., 1987), suggestive of a recent episode of viral transposition in mouse (Yamamoto, 1989). Such modification in regulatory circuits may be the driving force behind major phenotypic changes underlying morphological evolution (Yamamoto, 1985).

Conclusions

Regulation of the transcriptional regulators, even in the most thoroughly investigated networks such as those controlling embryonic development in *Drosophila*, is still poorly understood. Over the next several years it should become clearer how many of these transcriptional regulators are regulated. The complexity of the glucocorticoid receptor promoter structure suggests that receptor gene regulation is also likely to be complex, reflecting multiple strategies of transcriptional and post-transcriptional

control. I have touched on three possible routes of regulatory control: alternate promoter usage, initiation of translation, and alternate splicing. The data also suggest the intriguing possibility of the existence of at least two independently regulated promoters differing in their patterns of tissue-specific expression and autoregulation. Further investigation will reveal the importance of these and other possible regulatory mechanisms governing receptor expression.

References

1. Abate, C. and T. Curran (1990). Encounters with Fos and Jun on the road to AP-1. *Seminars Cancer Biol.*, **1**:19-26.

2. Affolter, M., A. Schier and W. J. Gehring (1990). Homeodomain proteins and the regulation of gene expression. *Curr. Opin. Cell Biol.*, **2**:485-495.

3. Amero, S. A., R. H. Kretsinger, N. D. Moncrief, K. R. Yamamoto and W. R. Pearson (1991). The origin of nuclear receptor transcription factors: a single precursor distinct from other zinc fingers. *Genes Dev.*, **submitted.**

4. Angel, P., K. Hattori, T. Smeal and M. Karin (1988). The *jun* proto-oncogene is positively autoregulated by its product, Jun/AP-1. *Cell*, **55**:875-885.

5. Antakly, T. and H. J. Eisen (1982). Detection of glucocorticoid receptor by immunocytochemistry. *J. Cell Biol.*, **95**:437a.

6. Antakly, T. and H. J. Eisen (1984). Immunocytochemical localization of glucocorticoid receptor in target cells. *Endocrinology*, **115**:1984-1989.

7. Antakly, T., S. Mercille and J. P. Côté (1987). Tissue-specific dopaminergic regulation of the glucocorticoid receptor in the rat pituitary. *Endocrinology*, **120**:1558-1562.

8. Antakly, T., A. Sasaki, A. S. Liotta, M. Palkovits and D. T. Krieger (1985). Induced expression of the glucocorticoid receptor in the rat intermediate pituitary lobe. *Science*, **229**:277-279.

9. Antakly, T., E. B. Thompson and D. O'Donnell (1989). Demonstration of the intracellular localization and up-regulation of glucocorticoid receptor by *in situ* hybridization and immunocytochemistry. *Cancer Res.*, **49**:2230s-2234s.

10. Aronsson, M., K. Fuxe, Y. Dong, L. F. Agnati, S. Okret and J.-Å. Gustafsson (1988). Localization of glucocorticoid receptor mRNA in the male rat brain by *in situ* hybridization. *Proc. Natl. Acad. Sci. USA*, **85**:9331-9335.

11. Arriza, J. L., C. Weinberger, G. Cerelli, T. M. Glaser, B. L. Handelin, D. E. Housman and R. M. Evans (1987). Cloning of human mineralocorticoid receptor complementary DNA: structural and functional kinship with the glucocorticoid receptor. *Science*, **237**:268-275.

12. Aziz, N. and H. N. Munro (1987). Iron regulates ferritin mRNA translation through a segment of its 5' untranslated region. *Proc. Natl. Acad. Sci. USA*, **84**:8478-8482.

13. Ballard, P. L., J. D. Baxter, S. J. Higgins, G. G. Rosseau and G. M. Tomkins (1974). General presence of glucocorticoid receptors in mammalian tissues. *Endocrinology*, **94**:998-1002.

14. Bartova, A. (1968). Functioning of the hypothalamo-pituitary-adrenal axis system during postnatal development in rats. *Endocrinology*, **10**:235.

15. Beato, M. (1989). Gene regulation by steroid hormones. *Cell*, **56**:335-344.

16. Berg, A., K. Carlström, J.-Å. Gustafsson and M. Ingelman-Sandberg (1975). Demonstration of a cytochrome P-450 dependent steroid 15β-hydroxylase in *Bacillus megaterium. Biochem. Biophys. Res. Commun.*, **66**:1414-1423.

17. Berg, A., J.-Å. Gustafsson, M. Ingelman-Sandberg and K. Carlström (1976). Characterization of a cytochrome P-450 dependent steroid hydroxylase system present in *Bacillus megaterium. J. Biol. Chem.*, **251**:2831-2838.

18. Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature*, **321**:209-213.

19. Blake, C. C. F. (1979). Exons encode protein functional units. *Nature*, **277**:598.

20. Bloom, E., D. T. Metulich, N. C. Lan, S. J. Higgins, S. S. Simons and J. D. Baxter (1980). Nuclear binding of glucocorticoid receptors: correlation between cytosol binding, activation and the biological response. *J. Steroid Biochem.*, **12**:175-184.

21. Bloomfield, C. D., K. A. Smith, B. A. Peterson and A. Munck (1981). Glucocorticoid receptors in adult acute lymphoblastic leukemia. *Cancer Res.*, **41**:4857-4860.

22. Bocquel, M. T., V. Kumar, C. Stricker, P. Chambon and H. Gronemeyer (1989). The contribution of the N- and C-terminal regions of steroid receptors to activation of transcription is both receptor and cell-specific. *Nucleic Acids Res.*, **12**:2581-2595.

23. Bohn, M. C. (1984). Glucocorticoid induced teratologies of the nervous system. In: *Neurobehavioral Teratology* (ed. J. Yanai) Vol. , pp. 365-387. Elsevier,

24. Bourgeois, S. and J. C. Gasson (1985). Genetic and epigenetic bases of glucocorticoid resistance in lymphoid cell lines. In: *Biochemical Actions of Hormones* (ed. Vol. 12, pp. 311-351. Academic Press,

25. Bourgeois, S. and R. F. Newby (1977). Diploid and haploid states of the glucocorticoid receptor gene of mouse lymphoma cell lines. *Cell*, **11**:423-430.

26. Bourgeois, S. and R. F. Newby (1979). Correlation between glucocorticoid receptor and cytolytic response of murine lymphoid cell lines. *Cancer Res.*, **39**:4749-4751.

27. Brent, R. and M. Ptashne (1981). Mechanism of action of the *lexA* gene product. *Proc. Natl. Acad. Sci. U.S.A.*, **78**:4204-4208.

28. Brent, R. and M. Ptashne (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell*, **43**:729-736.

29. Bresnick, E. H., F. C. Dalman, E. R. Sanchez and W. B. Pratt (1989). *J. Biol. Chem.*, **264**:4992-4997.

30. Briggs, M. R., J. T. Kadonaga, S. P. Bell and R. Tjian (1986). Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science*, **234**:47-52.

31. Brinkmann, A. O., P. W. Faber, H. C. J. van Rooij, G. G. J. M. Kuiper, C. Ris, P. Klaassen, J. A. G. M. van der Korput, M. M. Voorhorst, J. H. van Laar, E. Mulder and J. Trapman (1989). The human androgen receptor: domain structure, genomic organization and regulation of expression. *J. Steroid Biochem.*, **34**:307-310.

32. Bürglin (1988). *Cell*, **53**:339-340.

33. Burnstein, K. L., C. M. Jewell and J. A. Cidlowski (1990). Human glucocorticoid receptor cDNA contains sequences sufficient for receptor down-regulation. *J. Biol. Chem.*, **265**:7284-7291.

34. Burshell, A., P. A. Stathis, Y. Do, S. C. Miller and D. Feldman (1984). Characterization of an estrogen-binding protein in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **259**:3450-3456.

35. Butte, C., R. Kakihana, M. L. Farnham and E. P. Noble (1973). The relationship between brain and plasma corticosterone stress in developing rats. *Endocrinology*, **92**:1775-1779.

36. Capek, A., O. Hanc and M. Tara. (1966). Microbial Transformation of Steroids. Academia, Prague.

37. Carlstedt-Duke, J., P. E. Strömstedt, Ö. Wrange, T. Bergman, J.-Å. Gustafsson and H. Jörnvall (1988). Identification of hormone-interacting amino acid residues within the steroid-binding domain of the glucocorticoid

receptor in relation to other steroid hormone receptors. *J. Biol. Chem.*, **263**:6842-6846.

38. Cavalier-Smith, T. (1985). Selfish DNA and the origin of introns. *Nature*, **315**:283-284.

39. Chandler, V. L., B. A. Maler and K. R. Yamamoto (1983). DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter responsive in vivo. *Cell*, **33**:489-499.

40. Charney, W. and H. L. Herzog. (1967). Microbial Transformation of Steroids. A Handbook. Academic Press, New York.

41. Chatelain, A. and J.-P. Dupouy (1981). Activity of the pituitary-adrenal system in rat fetuses subjected to encephalectomy in early or late stages of pregnancy. *Neuroendocrinology*, **33**:148-153.

42. Chatelain, A., J. P. Dupouy and J.-P. Allaume (1980). Fetal-maternal adrenocorticotropin and corticosterone relationships in the rat: effects of maternal adrenalectomy. *Endocrinology*, **106**:1297-1303.

43. Chevrier, D., C. Vézina, J. Bastille, C. Linard, N. Sonenberg and G. Boileau (1988). Higher order structures of the 5'-proximal region decrease the efficiency of translation of the procine pro-opiomelanocortin mRNA. *J. Biol. Chem.*, **263**:902-910.

44. Cidlowski, J. A. and N. B. Cidlowski (1981). Regulation of glucocorticoid receptors by glucocorticoids in cultured HeLa S_3 cells. *Endocrinology*, **109**:1975-1982.

45. Cidlowski, J. A. and N. B. Cidlowski (1982). Glucocorticoid receptors and the cell cycle: evidence that the accumulation of glucocorticoid receptors during the S phase of the cell cycle is dependent on ribonucleic acid and protein synthesis. *Endocrinology*, **110**:1653-1662.

46. Claman, H. N. (1972). Corticosteroids and lymphoid cells. *N. Engl. J. Med.*, **287**:388-397.

47. Cohen, J. J. and R. C. Duke (1984). Glucocorticoid activation of a calcium-dependent endonuclease in thymocyte nuclei leads to cell death. *J. Immunol.*, **132**:38-42.

48. Cole, M. D. (1990). The *myb* and *myc* nuclear oncogenes as transcriptional activators. *Curr. Opin. Cell Biol.*, **2**:502-508.

49. Corbier, P. and J. Roffi (1978). Pituitary adrenocortical response to stress during the first day of postnatal life in the rat. *Biol. Neonate*, **34**:105-107.

50. Craik, C. S., W. J. Rutter and R. Fletterick (1983). Splice junctions: association with variation in protein structure. *Science*, **220**:1125-1129.

51. Crank-Tseng, J. and F. G. Berger (1987). Evolution of steroid-inducible RP2 mRNA expression in the mouse kidney. *Genetics*, **116**:593-599.

52. Curran, T. and P. K. Vogt (1991). Dangerous liaisons: Fos and Jun - oncogenic transcription factors.

53. Dallman, M. F., S. F. Akana, C. S. Cascio, D. N. Darlington, L. Jacobson and N. Levin (1987). Regulation of ACTH secretion: variations on a theme of B. *Rec. Prog. Hormone Res.*, **43**:113-173.

54. Danielsen, M., L. Hinck and G. M. Ringold (1989). Two amino acids within the knuckle of the first zinc finger specify DNA response element activation by the glucocorticoid receptor. *Cell*, **57**:1131-1138.

55. Danielsen, M., J. P. Northrop, J. Jonklaas and G. M. Ringold (1987). Domains of the glucocorticoid receptor involved in specific and nonspecific deoxyribonucleic acid binding, hormone activation, and transcriptional enhancement. *Mol. Endocrinol.*, **1**:816-822.

56. Danielsen, M., J. P. Northrop and G. M. Ringold (1986). The mouse glucocorticoid receptor: mapping of functional domains by cloning, sequencing and expression of wild-type and mutant receptor proteins. *EMBO J.*, **5**:2513-2522.

57. Danielsen, M. and M. R. Stallcup (1984). Down-regulation of glucocorticoid receptors in mouse lymphoma cell variants. *Mol. Cell. Biol.*, **4**:449-453.

58. De Kloet, E. R., P. Rosenfeld, J. A. M. Van Eekelen, W. Sutanto and S. Levine (1988). Stress, glucocorticoids and development. In: *Progress in Brain Research* (ed. G. J. Boer et al.) Vol. 73, pp. 101-120. Elsevier, Amsterdam.

59. DeFranco, D., O. Wrange, J. Merryweather and K. R. Yamamoto (1985). Biological activity of a glucocorticoid regulated enhancer: DNA sequence requirements and interactions with other transcriptional enhancers. In: *Genome Rearrangement: Proceedings of the UCLA Symposium held at Steamboat Springs, Colorado, April 7-13, 1984* (ed. M. Simon and I. Herskowitz) Vol. 20, pp. 305-321. Alan R. Liss, New York.

60. Desplan, C., J. Theis and P. H. O'Farrell (1988). The sequence specificity of homeodomain-DNA interaction. *Cell*, **54**:1081-1090.

61. Diamond, M. I., J. N. Miner, S. K. Yoshinaga and K. R. Yamamoto (1990). Transcription factor interactions: selectors of positive or negative regulaton from a single DNA element. *Science*, **249**:1266-1272.

62. Dieken, E. S., E. U. Meese and R. L. Miesfeld (1990). nti glucocorticoid receptor transcripts lack sequences encoding the amino-terminal transcriptional modulatory domain. *Mol. Cell. Biol.*, **10**:4574-4581.

63. Distelhorst, C. W. (1988). Glucocorticosteroids induce DNA fragmentation in human lymphoid leukemia cells. *Blood*, **72**:1305-1309.

64. Dong, Y., L. Poellinger, J.-Å. Gustafsson and S. Okret (1988). Regulation of glucocorticoid receptor expression: evidence for transcriptional and posttranslational mechanisms. *Mol. Endocrinol.*, **2**:1256-1264.

65. Dorit, R. L., L. Schoenbach and W. Gilbert (1990). How big is the universe of exons? *Science*, **250**:1377-1382.

66. Eberwine, J. H. and J. L. Roberts (1984). Glucocorticoid regulation of proopiomelanocortin gene transcription in the rat pituitary. *J. Biol. Chem.*, **259**:2166-2170.

67. Ebina, Y., Y. Takahara, F. Kishi, A. Nakazawa and C. S. Parker (1983). LexA protein is a repressor of the colicin E1 gene. *J. Biol. Chem.*, **258**:13258-13261.

68. Eilers, M., D. Picard, K. R. Yamamoto and J. M. Bishop (1989). Chimaeras of myc oncoprotein and steroid receptors cause hormone-dependent transformation of cells. *Nature*, **340**:66-68.

69. Eisen, L. P., M. S. Elsasser and J. M. Harmon (1988). Positive regulation of the glucocorticoid receptor in human T-cells sensitive to the cytolytic effects of glucocorticoids. *J. Biol. Chem.*, **263**:12044-12048.

70. Evans, R. M. (1988). The steroid and thyroid hormone receptor superfamily. *Science*, **240**:889-899.

71. Fawell, S. E., J. A. Lees, R. White and M. G. Parker (1990). Characterization and colocalization of steroid binding and dimerization activities in the mouse estrogen receptor. *Cell*, **60**:953-962.

72. Feldman, D., Y. Do, A. Burshell, P. Stathis and D. S. Loose (1982). An estrogen-binding protein and endogenous ligand in *Saccharomyces cerevisiae:* possible hormone receptor system. *Science,* **218:**297-298.

73. Freedman, L. P., B. F. Luisi, Z. R. Korszun, R. Basavappa, P. B. Sigler and K. R. Yamamoto (1988). The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding protein. *Nature,* **334:**543-546.

74. Freedman, L. P., S. K. Yoshinaga, J. N. Vanderbilt and K. R. Yamamoto (1989). In vitro transcription enhancement by purified derivatives of the glucocorticoid receptor. *Science,* **245:**298-301.

75. Frohman, M. A., M. K. Dush and G. R. Martin (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA,* **85:**8998-9002.

76. Frohman, M. A. and G. R. Martin (1989). Rapid amplification of cDNA ends using nested primers. *Technique,* **1:**165-170.

77. Fuxe, K., A. Cintra, A. Härfstrand, L. F. Agnati, M. Kalia, M. Zoli, A.-C. Wikström, S. Okret, M. Aronsson and J.-Å. Gustafsson (1987). Central glucocorticoid receptor immunoreactive neurons: New insights into the endocrine regulation of the brain. *Annals New York Acad. Sci.,* **512:**362-393.

78. Fuxe, K., A.-C. Wikström, S. Okret, L. F. Agnati, A. Härfstrand, Z.-Y. Yu, L. Granholm, M. Zoli, W. Vale and J.-Å. Gustafsson (1985). Mapping of glucocorticoid receptor immunoreactive neurons in the rat tel- and diencephalon using a monoclonal antibody against rat liver glucocorticoid receptor. *Endocrinology,* **117:**1803-1812.

79. Galili, U. (1983). Glucocorticoid induced cytolysis of human normal and malignant lymphocytes. *J. Steroid Biochem.,* **19:**483-490.

80. Garabedian, M. J. and K. R. Yamamoto (1991). A genetic analysis of steroid binding and response by a mammalian glucocorticoid receptor in yeast. *Science,* **submitted:**

81. Gardiner-Garden, M. and M. Frommer (1987). CpG islands in vertebrate genomes. *J. Mol. Biol.,* **196:**261-282.

82. Gasson, J. C. and S. Bourgeois (1983). A new determinant of glucocorticoid sensitivity. *J. Cell Biol.,* **96:**409-415.

83. Gasson, J. C., T. Ryden and S. Bourgeois (1983). Role of de novo DNA methylation in the glucocorticoid resistance of a T-lymphoid cell line. *Nature*, **302**:621-623.

84. Gehring, U. and G. M. Tomkins (1974). A new mechanism for steroid unresponsiveness: Loss of nuclear binding activity of a steroid hormone receptor. *Cell*, **3**:301-306.

85. Gehring, U., J. Ulrich and B. Segnitz (1982). Lymphoma cell variants of decreased glucocorticoid sensitivity. *Mol. Cell. Endocrinol.*, **28**:605-611.

86. Ghosh, S., A. M. Gifford, L. R. Riviere, P. Tempst, G. P. Nolan and D. Baltimore (1990). Cloning of the p50 DNA binding subunit of NF-κB: homology to *rel* and *dorsal*. *Cell*, **62**:1019-1029.

87. Giguère, V., N. Yang, P. Segui and R. M. Evans (1988). Identification of a new class of steroid hormone receptors. *Nature*, **331**:91-94.

88. Gilbert, W. (1978). Why genes in pieces? *Nature*, **271**:501.

89. Gilbert, W. (1987). The exon theory of genes. *Cold Spring Harbor Symp. Quant. Biol.*, **52**:901-905.

90. Gillis, S., G. R. Crabtree and K. A. Smith (1979a). Glucocorticoid-induced inhibition of T cell growth factor production I. The effect on mitogen-induced lymphocyte proliferation. *J. Immunol.*, **123**:1624-1631.

91. Gillis, S., G. R. Crabtree and K. A. Smith (1979b). Glucocorticoid-induced inhibition of T cell growth factor production II. The effect on the in vitro generation of cytolytic T cells. *J. Immunol.*, **123**:1632-1638.

92. Giniger, E., S. M. Varnum and M. Ptashne (1985). Specific binding of GAL4, a positive regulatory protein of yeast. *Cell*, **40**:767-779.

93. Glantz, S. A. (1987). Primer of Biostatistics. 2nd ed. McGraw-Hill, New York.

94. Glantz, S. A. and B. K. Slinker. (1990). Primer of applied regression and analysis of variance. McGraw-Hill, New York.

95. Go, M. and M. Nosaka (1987). Protein architecture and the origin of introns. *Cold Spring Harbor Symp. Quant. Biol.*, **52**:915-924.

96. Godowski, P. J., D. Picard and K. R. Yamamoto (1988a). Signal transduction and transcriptional regulation by glucocorticoid receptor-lexA fusion proteins. *Science*, **241**:812-816.

97. Godowski, P. J., S. Rusconi, R. Miesfeld and K. R. Yamamoto (1987). Glucocorticoid receptor mutants that are constitutive activators of transcriptional enhancement. *Nature*, **325**:365-368.

98. Godowski, P. J., D. D. Sakai and K. R. Yamamoto (1988b). Signal transduction and transcriptional regulation by the glucocorticoid receptor. In: *DNA-Protein Interactions in Transcription* (ed. J. Gralla) Vol. 95, pp. 197-210. Alan R. Liss, Inc., New York.

99. Goldstein, J. L. and M. S. Brown (1990). Regulation of the mevalonate pathway. *Nature*, **343**:425-430.

100. Gomi, M., K. Moriwaki, S. Katagiri, Y. Kurata and E. B. Thompson (1990). Glucocorticoid effects on myeloma cells in culture: correlation of growth inhibition with induction of glucocorticoid receptor messenger RNA. *Cancer Res.*, **50**:1873-1878.

101. Goto, T., P. Macdonald and T. Maniatis (1989). Early and late periodic patterns of *even skipped* expression are controlled by distinct regulatory elements that respond to different spatial cues. *Cell*, **57**:413-422.

102. Green, S. and P. Chambon (1987). Oestradiol induction of a glucocorticoid-responsive gene by a chimeric receptor. *Nature*, **325**:75-78.

103. Green, S. and P. Chambon (1988). Nuclear receptors enhance our understanding of transcription regulation. *Trends Genet.*, **4**:309-313.

104. Green, S., V. Kumar, I. Theulaz, W. Wahli and P. Chambon (1988). The N-terminal DNA-binding 'zinc finger' of the oestrogen and glucocorticoid receptors determines target gene specificity. *EMBO J.*, **7**:3037-3044.

105. 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*, **320**:134-139.

106. Greene, G. L., P. Gilna, M. Waterfield, A. Baker, Y. Hort and J. Shine (1986). Sequence and expression of human estrogen receptor complementary DNA. *Science*, **231**:1150-1154.

107. Gruol, D. J., F. M. Rajah and S. Bourgeois (1989). Cyclic AMP-dependent protein kinase modulation of the glucocorticoid-induced cytolytic response in murine T-lymphoma cells. *Mol. Endocrinol.*, **3**:2119-2127.

108. Guillet, R., M. Saffran and S. M. Michaelson (1980). Pituitary-adrenal response in neonatal rats. *Endocrinology*, **106**:991-994.

109. Gustafsson, J.-Å., J. Carlstedt-Duke, L. Poellinger, S. Okret, A.-C. Wikström, M. Brönnegård, M. Gillner, Y. Dong, K. Fuxe, A. Cintra, A. Härfstrand and L. Agnati (1987). Biochemistry, molecular biology, and physiology of the glucocorticoid receptor. *Endocrine Rev.*, **8**:185-234.

110. Haltmeyer, G. C., V. H. Denenberg, J. Thatcher and M. X. Zarrow (1966). Response of the adrenal cortex of the neonatal rat after subjection to stress. *Nature*, **212**:1371-1372.

111. Härd, T., E. Kellenbach, R. Boelens, R. Kaptein, K. Dahlman, J. Carlstedt-Duke, L. P. Freedman, B. A. Maler, E. I. Hyde, J.-Å. Gustafsson and K. R. Yamamoto (1990a). ^{1}H NMR studies of the glucocorticoid receptor DNA-binding domain: sequential assignments and identification of secondary structure elements. *Biochemistry*, **29**:9015-9023.

112. Härd, T., E. Kellenbach, R. Boelens, B. A. Maler, K. Dahlman, L. P. Freedman, J. Carlstedt-Duke, K. R. Yamamoto, J.-Å. Gustafsson and R. Kaptein (1990b). Solution structure of the glucocorticoid receptor DNA-binding domain. *Science*, **249**:157-160.

113. Harding, K., T. Hoey, R. Warrior and M. Levine (1989). Autoregulatory and gap gene response elements of the *even-skipped* promoter of *Drosophila*. *EMBO J.*, **8**:1205-1212.

114. Harmon, J. M., M. R. Norman, B. J. Fowlkes and E. B. Thompson (1979). Dexamethasone induces irreversible G1 arrest and death of a human lymphoid cell line. *J. Cell Physiol.*, **98**:267-278.

115. Harmon, J. M., E. B. Thompson and K. A. Baione (1985). Analysis of glucocorticoid-resistant human leukemic cells by somatic cell hybridization. *Cancer Res.*, **45**:1587-1593.

116. Harper, G. P., F. L. Pearce and C. A. Vernon (1980). The production and storage of nerve growth factor *in vivo* by tissues of the mouse, rat, guinea pig, hamster and gerbil. *Dev. Biol.*, **77**:391-402.

117. Harrigan, M. T., G. Baughman, N. F. Campbell and S. Bourgeois (1989). Isolation and characterization of glucocorticoid- and cyclic AMP-induced genes in T lymphocytes. *Mol. Cell. Biol.*, **9**:3438-3446.

118. Hentze, M. W., S. W. Caughman, T. A. Rouault, J. G. Barriocanal, A. Dancis, J. B. Harford and R. D. Klausner (1987a). Identification of the iron-responsive element for the translational regulation of human ferritin mRNA. *Science*, **238**:1570-1573.

119. Hentze, M. W., T. A. Rouault, S. W. Caughman, A. Dancis, J. B. Harford and R. D. Klausner (1987b). A cis-acting element is necessary and sufficient for translational regulation of human ferritin expression in response to iron. *Proc. Natl. Acad. Sci. USA*, **84**:6730-6734.

120. Hickey, D. A. and B. F. Benkez (1986). *J. Theor. Biol.*, **121**:283.

121. Hickey, D. A., B. F. Benkez and S. M. Abukashawa (1989). *J. Theor. Biol.*, **137**:41.

122. Himmelfarb, H. J., J. Pearlberg, D. H. Last and M. Ptashne (1990). GAL11P: A yeast mutation that potentiates the effect of weak GAL4-derived activators. *Cell*, **63**:1299-1309.

123. Hinnebusch, A. G. (1984). Evidence for translational regulation of the activator of general amino acid control in yeast. *Proc. Natl. Acad. Sci. USA*, **81**:6442-6446.

124. Hinnebusch, A. G. (1988). Mechanisms of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae*. *Microbiol. Rev.*, **52**:248-273.

125. Hiromi, Y. and W. J. Gehring (1987). Regulation and function of the *Drosophila* segmentation gene *fushi tarazu*. *Cell*, **50**:963-974.

126. Hodin, R. A., M. A. Lazar, B. I. Wintman, D. S. Darling, R. J. Koenig, P. R. Larsin, D. D. Moore and W. W. Chin (1989). Identification of a thyroid hormone receptor that is pituitary specific. *Science*, **244**:76-79.

127. 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.

128. Hollenberg, S. and R. M. Evans (1988). Multiple and cooperative *trans*-activation domains of the human glucocorticoid receptor. *Cell*, **55**:899-906.

129. Hollenberg, S. M., V. Giguère and R. M. Evans (1989). Identification of two regions of the human glucocorticoid receptor hormone binding domain that block activation. *Cancer Res.*, **49**:2292s-2294s.

130. Hollenberg, S. M., C. Weinberger, E. S. Ong, G. Cerelli, A. Oro, R. Lebo, E. B. Thompson, M. G. Rosenfeld and R. M. Evans (1985). Primary structure and expression of a functional human glucocorticoid receptor cDNA. *Nature*, **318**:635-641.

131. Hope, T. J., X. Huang, D. McDonald and T. G. Parslow (1990). Steroid-receptor fusion of the human immunodeficiency virus type 1 Rev transactivator: mapping cryptic functions of the arginine-rich motif. *Proc. Natl. Acad. Sci. USA*, **87**:7787-7791.

132. Horibata, K. and A. W. Harris (1970). Mouse myelomas and lymphomas in culture. *Exptl. Cell Res.*, **60**:61-77.

133. Howard, K. J., S. J. Holley, K. R. Yamamoto and C. W. Distelhorst (1990). Mapping the HSP90 binding region of the glucocorticoid receptor. *J. Biol. Chem.*, **265**:11928-11935.

134. Huckaby, C. S., O. M. Conneely, W. G. Beattie, A. D. W. Dobson, M.-J. Tsai and B. W. O'Malley (1987). Structure of the chromosomal chicken progesterone receptor gene. *Proc. Natl. Acad. Sci. USA*, **84**:8380-8384.

135. Hughes, M. R., P. J. Malloy, D. G. Kieback, R. A. Kesterson, J. W. Pike, D. Feldman and B. W. O'Malley (1988). Point mutations in the human vitamin D receptor gene associated with hypocalcemic rickets. *Science*, **242**:1702-1705.

136. Ingham, P. W. (1988). The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature*, **335**:25-34.

137. Ish-Horowicz, D., S. M. Pinchin, P. W. Ingham and H. G. Gyurkovics (1989). Autocatalytic *ftz* activation and metameric instability induced by ectopic *ftz* expression. *Cell*, **57**:223-232.

138. Ishii, S., G. T. Merlino and I. Pastan (1985a). Promoter region of the human harvey ras proto-oncogene: similarity to the EGF receptor proto-oncogene promoter. *Science*, **230**:1378-1381.

139. Ishii, S., Y.-h. Xu, R. H. Stratton, B. A. Roe, G. T. Merlino and I. Pastan (1985b). Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene. *Proc. Natl. Acad. Sci. U.S.A.*, **82**:4920-4924.

140. Jailer, J. W. (1950). The maturation of the pituitary-adrenal axis in the newborn rat. *Endocrinology*, **46**:420-425.

141. 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.

142. Johnson, P. F. and S. L. McKnight (1989). Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem.*, **58**:799-839.

143. Jonat, C., H. J. Rahmsdorf, K.-K. Park, A. C. B. Cato, S. Gebel, H. Ponta and P. Herrlich (1990). Antitumor promotion and antiinflammation: down-modulation of AP-1 (Fos/Jun) activity by glucocorticoid hormone. *Cell*, **62**:1189-1204.

144. Jones, K. A., J. T. Kadonaga, P. A. Luciw and R. Tjian (1986). Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1. *Science*, **232**:755-759.

145. Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly and R. Tjian (1987). A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell*, **48**:79-89.

146. Jorgensen, E. M. and R. L. Garber (1987). Function and misfunction of the two promoters of the *Drosophila Antennapedia* gene. *Genes Dev.*, **1**:544-555.

147. Kadonaga, J. T., K. A. Jones and R. Tjian (1986). Promoter-specific activation of RNA polymerase II transcription by Sp1. *Trends Biochem. Sci.*, **11**:20-23.

148. Kalderon, D., W. Richardson, A. F. Markham and A. E. Smith (1984). Sequence requirements for nuclear localization of SV40 large T-antigen. *Nature*, **311**:33-38.

149. Kalinyak, J. E., R. I. Dorin, A. R. Hoffman and A. J. Perlman (1987). Tissue-specific regulation of glucocorticoid receptor mRNA by dexamethasone. *J. Biol. Chem.*, **262**:10441-10444.

150. Kalinyak, J. E., C. A. Griffin, R. W. Hamilton, J. G. Bradshaw, A. J. Perlman and A. R. Hoffman (1989). Developmental and hormonal regulation of glucocorticoid receptor messenger RNA in the rat. *J. Clin Invest.*, **84**:1843-1848.

151. Kamens, J., P. Richardson, G. Mosialos, R. Brent and T. Gilmore (1990). Oncogenic transformation by vrel requires an amino-terminal activation domain. *Mol. Cell. Biol.*, **10**:2840-2847.

152. Kastner, P., M.-T. Bocqel, B. Turcotte, J.-M. Garnier, K. B. Horwitz, P. Chambon and H. Gronemeyer (1990a). 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.

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

154. Kieran, M., V. Blank, F. Logeat, J. Vanderkerckhove, F. Lottspeich, O. L. Bail, M. B. Urban, P. Kourlisky, P. A. Baeuerle and A. Israël (1990). The DNA binding subunit of NF-κB is identical to factor KBF1 and homologous to the *rel* oncogene product. *Cell*, **62**:1007-1018.

155. Kirschmeier, P. T., G. M. Housey, M. D. Johnson, A. S. Perkins and I. B. Weinstein (1988). Construction and characterization of a retroviral vector demonstrating efficient expression of cloned cDNA sequences. *DNA*, **7**:219-225.

156. Kissinger, C. R., B. Liu, E. Martin-Blanco, T. B. Kornberg and C. O. Pabo (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell*, **63**:579-590.

157. Koch, B. (1969). Fraction libre de la corticostérone plasmatiquie et réponse hypophyso-surrénalienne au stress durant la période post-natale chez le rat. *Horm. Metab. Res.*, **1**:301-308.

158. Kovács, K., J. A. Kiss and G. B. Makara (1986). Glucocorticoid implants around the hypothalamic paraventricular nucleus prevent the increase of corticotropin-releasing factor and arginine vasopressin immunostaining induced by adrenalectomy. *Neuroendocrinology*, **44**:229-234.

159. Kozak, M. (1986a). Bifunctional messenger RNAs in eukaryotes. *Cell*, **47**:481-483.

160. Kozak, M. (1986b). Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. USA*, **83**:2850-2854.

161. Kozak, M. (1986c). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**:283-292.

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

163. Kozak, M. (1988). A profusion of controls. *J. Cell Biol.*, **107**:1-7.

164. Kozak, M. (1989). Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Molecular and Cellular Biology*, **9**:5134-5142.

165. 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 J.*, **5**:891-897.

166. Kuiper, G. G. J. M., P. W. Faber, H. C. J. van Rooij, J. A. G. M. van der Korput, C. Ris-Stalpers, P. Klassen, J. Trapman and A. O. Brinkmann (1989). Structural organization of the human androgen receptor gene. *J. Mol. Endocr.*, **2**:R1-R4.

167. Kumar, V., S. Green, G. Stack, M. Berry, J.-R. Jin and P. Chambon (1987). Functional domains of the human estrogen receptor. *Cell*, **51**:941-951.

168. Lacroix, A., G. D. Bonnard and M. E. Lippman (1984). Modulation of glucocorticoid receptors by mitogenic stimuli, glucocorticoids and retinoids in normal human cultured T cells. *J. Steroid Biochem.*, **21**:73-80.

169. Landschulz, W. H., P. F. Johnson and S. L. McKnight (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science*, **240**:1759-1764.

170. Latimer, J. J., F. G. Berger and H. Baumann (1987). Developmental expression, cellular localization, and testosterone regulation of α_1-antitrypsin in *Mus caroli* kidney. *J. Biol. Chem.*, **262**:12641-12646.

171. Laughon, A., A. M. Boulet, J. John R. Bermingham, R. A. Laymon and M. P. Scott (1986). Structure of transcripts from the homeotic Antennapedia gene of *Drosophila melanogaster*: two promoters control the major protein-coding region. *Mol. Cell. Biol.*, **6**:4676-4689.

172. Laughon, A. and M. P. Scott (1984). Sequence of a Drosophila segmentation gene: protein structure homology with DNA-binding proteins. *Nature*, **310**:25-31.

173. 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-erbA alpha transcriptional unit. *Mol. Cell Biol.*, **9**:1128-1136.

174. Lees, J. A., S. E. Fawell, R. White and M. G. Parker (1990). A 22-amino-acid peptide restores DNA-binding activity to dimerization-defective mutants of the estrogen receptor. *Mol. Cell. Biol.*, **10**:5529-5531.

175. Lehmann, J. M., B. Hoffmann and M. Pfahl (1991). Genomic organization of the retinoic acid receptor gamma gene. *Nucl. Acids Res.*, **19**:573-578.

176. Levine, S. (1970). The pituitary-adrenal system and the developing brain. *Prog. Brain Res.*, **32**:79-85.

177. Levy, J. R. and J. M. Olefsky (1990). Receptor-mediated internalization and turnover. In: *Insulin* (ed. P. Cuatrecasas and S. Jacobs) Vol. , pp. 237-266. Springer Verlag, Berlin.

178. Lewis, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature*, **276**:565-570.

179. Loose, D. S., D. J. Schurman and D. Feldman (1981). A corticosteroid binding protein and endogenous ligand in C. *albicans* indicating a possible steroid-receptor system. *Nature*, **293**:477-479.

180. Lubahn, D. B., T. R. Brown, J. A. Simental, H. N. Higgs, C. J. Migeon, E. M. Wilson and F. S. French (1989). Sequence of the intron/exon junctions of the coding region of the human androgen receptor gene and identification of a point mutation in a family with complete androgen insensitivity. *Proc. Natl. Acad. Sci. USA*, **86**:9534-9538.

181. Lucibello, F. C., E. P. Slater, K. U. Jooss, M. Beato and R. Müller (1990). Mutual transrepression of Fos and the glucocorticoid receptor: involvement of a functional domain in Fos which is absent in FosB. *EMBO J.*, **9**:2827-2834.

182. Mader, S., V. Kumar, H. de Verneuil and P. Chambon (1989). Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid-responsive element. *Nature*, **338**:271-274.

183. Marchionni, M. and W. Gilbert (1986). The triosephosphate isomerase gene from maize: introns antedate the plant-animal divergence. *Cell*, **46**:133-141.

184. 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*, **332**:171-173.

185. Masters, J. N. and G. Attardi (1985). Discrete human dihydrofolate reductase gene transcripts present in polysomal RNA map with their 5' ends

several hundred nucleotides upstream of the main mRNA start site. *Mol. Cell. Biol.*, **5**:493-500.

186. Mayo, K. E. and R. D. Palmiter (1981). Glucocrticoid regulation of metallothionein-I mRNA synthesis in cultured mouse cells. *J. Biol. Chem.*, **256**:2621-2624.

187. McGinnis, J. F. and J. De Vellis (1981). Cell surface modulation of gene expression in brain cells by down regulation of glucocorticoid receptors. *Proc. Natl. Acad. Sci. USA*, **78**:1288-1292.

188. McIntyre, W. R. and H. H. Samuels (1985). Triamcinolone acetonide regulates glucocorticoid-receptor levels by decreasing the half-life of the activated nuclear-receptor form. *J. Biol. Chem.*, **260**:418-427.

189. Meaney, M. J., R. M. Sapolsky and B. S. McEwen (1985). The development of the glucocorticoid receptor system in the rat limbic brain. I. Ontogeny and autoregulation. *Develop. Brain Res.*, **18**:159-164.

190. Melton, D. W., D. S. Konecki, J. Brennand and C. T. Caskey (1984). Structure, expression, and mutation of the hypoxanthine phosphoribosyltransferase gene. *Proc. Natl. Acad. Sci. U.S.A.*, **81**:2147-2151.

191. Mendel, D. B., J. E. Bodwell, B. Gametchu, R. W. Harrison and A. Munck (1986). *J. Biol. Chem.*, **261**:3758-3763.

192. Miesfeld, R., P. J. Godowski, B. A. Maler and K. R. Yamamoto (1987). Glucocorticoid receptor mutants that define a small region sufficient for enhancer activation. *Science*, **236**:423-427.

193. Miesfeld, R., S. Okret, A.-C. Wikström, Ö. Wrange, J.-Å. Gustafsson and K. R. Yamamoto (1984). Characterization of a steroid hormone receptor gene and mRNA in wild-type and mutant cells. *Nature*, **312**:779-781.

194. Miesfeld, R., S. Rusconi, P. J. Godowski, B. A. Maler, S. Okret, A.-C. Wikström, J.-Å. Gustafsson and K. R. Yamamoto (1986). Genetic complementation of a glucocorticoid receptor deficiency by expression of cloned receptor cDNA. *Cell*, **46**:389-399.

195. Miesfeld, R., S. Rusconi, S. Okret, A.-C. Wikström, J.-Å. Gustafsson and K. R. Yamamoto (1985). Preliminary analyses of the glucocorticoid receptor gene and its expression: a DNA-binding protein essential for hormone-dependent transcriptional enhancement. In: *Sequence Specificity in Transcription and Translation* (ed. R. Calender and L. Gold) Vol. 30, pp. 535-545. Alan R. Liss, New York.

196. Miesfeld, R., D. Sakai, A. Inoue, M. Schena, P. J. Godowski and K. R. Yamamoto (1988). Glucocorticoid receptor sequences that confer positive and negative transcriptional regulation. In: *Steroid Hormone Action: Proceedings of a UCLA Symposium held in Park City, Utah, January 17-23, 1987* (ed. G. Ringold) Vol. 75, pp. 193-200. Alan R. Liss, New York.

197. Milkovic, S. and K. Milkovic (1969). Responsiveness of the pituitary-adrenocortical system during embryonic and early postnatal periods of life. In: *Physiology and Pathology of Adaptation Mechanisms* (ed. E. Bajusz) Vol. 1, pp. 28-47. Pergamon Press, New York.

198. Miller, J., A. D. McLachlan and A. Klug (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.*, **4**:1609-1614.

199. Mitchell, P. J. and R. Tjian (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science*, **245**:371-378.

200. Moore, D. D. (1990). Diversity and unity in the nuclear hormone receptors: a terpinoid receptor superfamily. *The New Biologist*, **2**:100-105.

201. Mordacq, J. C. and D. I. H. Linzer (1989). Co-localization of elements required for phorbol ester stimulation and glucocorticoid repression of proliferin gene expression. *Genes Dev.*, **3**:760-769.

202. Mornon, J. P., V. Bissery, C. Gaboriaud, A. Thomas, T. Ojasoo and J.-P. Taynaud (1989). Hydrophobic cluster analysis (HCA) of the hormone-binding domain of receptor proteins. *J. Steroid Biochem.*, **34**:355-361.

203. Mueller, P. P., S. Harashima and A. G. Hinnebusch (1987). A segment of GCN4 mRNA containing the upstream AUG codons confers translational control upon a heterologous yeast transcript. *Proc. Natl. Acad. Sci. USA*, **84**:2863-2867.

204. Mueller, P. P. and A. G. Hinnebusch (1986). Multiple upstream AUG codons mediate translational control of GCN4. *Cell*, **45**:201-207.

205. Murre, C., P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub and D. Baltimore (1989a). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell*, **58**:537-544.

206. Murre, C., P. Schonleber McCaw and D. Baltimore (1989b). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless, MyoD,* and *myc* proteins. *Cell,* **56**:777-783.

207. Nes, W. and M. L. McKean. (1977). Biochemistry of Steroids and Other Isopentoids. University Park Press, Baltimore.

208. Nishizawa, M., Y. Suzuki, Y. Nogi, K. Matsumoto and T. Fukasawa (1990). Yeast Gal11 protein mediates the transcriptional activation signal of two different transacting factors, Gal4 and general regulatory factor I / repressor / activator site binding protein1 / translation upstream factor. *Proc. Natl. Acad. Sci. USA,* **87**:5373-5377.

209. Norman, A. W. and G. Litwack. (1987). Hormones. Academic Press,

210. Northrop, J. P., M. Danielsen and G. M. Ringold (1986). Analysis of glucocorticoid unresponsive cell variants using a mouse glucocorticoid receptor complementary DNA clone. *J. Biol. Chem.,* **261**:11064-11070.

211. O'Shea, E. K., R. Rutkowski and P. S. Kim (1989). Evidence that the leucine zipper is a coiled coil. *Science,* **243**:538-542.

212. Obaru, K., T. Tsuzuki, C. Setoyama and K. Shimada (1988). Structural organization of the mouse aspartate aminotransferase isoenzyme genes - introns antedate the divergence of cytosolic and mitochondrial isoenzyme genes. *J. Mol. Biol.,* **200**:13-22.

213. Okret, S., L. Poellinger, Y. Dong and J.-Å. Gustafsson (1986). Down-regulation of glucocorticoid receptor mRNA by glucocorticoid hormones and recognition by the receptor of a specific binding sequence within a receptor cDNA clone. *Proc. Natl. Acad. Sci. USA,* **83**:5899-5903.

214. Pan, T. and J. E. Coleman (1990a). The DNA binding domain of GAL4 forms a binuclear metal ion complex. *Biochemistry,* **29**:3023-3029.

215. Pan, T. and J. E. Coleman (1990b). GAL4 transcription factor is not a "zinc finger" but forms a $Zn_2(II)Cys_6$ binuclear cluster. *Proc. Natl. Acad. Sci. USA,* **87**:2077-2081.

216. Payvar, F., D. DeFranco, G. L. Firestone, B. Edgar, Ö. Wrange, S. Okret, J.-Å. Gustafsson and K. R. Yamamoto (1983). Sequence-specific binding of glucocorticoid receptor to MTV DNA at sites within and upstream of the transcribed region. *Cell,* **35**:381-392.

217. Payvar, F. and Ö. Wrange (1983). Relative selectivities and efficiencies of DNA binding by purified intact and protease-cleaved glucocorticoid receptor.

In: *Steroid Hormone Receptors: Structure and Function* (ed. H. Eriksson and J.-Å. Gustafsson) Vol. , pp. 267-284. Elsevier,

218. Petkovich, M., N. J. Brand, A. Krust and P. Chambon (1987). A human retinoic acid receptor which belongs to the family of nuclear receptors. *Nature*, **330**:444-450.

219. Picado-Leonard, J. and W. L. Miller (1988). Homologous sequences in steroidogenic enzymes, steroid receptors and a steroid binding protein suggest a consensus steroid-binding sequence. *Mol. Endocrinol.*, **2**:1145-1150.

220. Picard, D., B. Khursheed, M. J. Garabedian, M. G. Fortin, S. Lindquist and K. R. Yamamoto (1991). Reduced levels of hsp90 compromise steroid receptor action *in vivo*. *Nature*, **348**:166-168.

221. Picard, D., V. Kumar, P. Chambon and K. R. Yamamoto (1990). Signal transduction by steroid hormones: nuclear localization is differentially regulated in estrogen and glucocorticoid receptors. *Cell Regulation*, **1**:291-299.

222. Picard, D., S. J. Salser and K. R. Yamamoto (1988). A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor. *Cell*, **54**:1073-1080.

223. Picard, D. and K. R. Yamamoto (1987). Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. *EMBO J.*, **6**:3333-3340.

224. Pike, J. W. 1990. *Personal Communication.*

225. Ponglikitmongkol, M., S. Green and P. Chambon (1988). Genomic organization of the human oestrogen receptor gene. *EMBO J.*, **7**:3385-3388.

226. Poste, G. and S. T. Crooke. (1985). Mechanisms of receptor regulation.

227. Pratt, W. B., D. J. Dolly, D. V. Pratt, S. M. Hollenberg, V. Giguère, F. M. Cadepond, G. Schweizer-Groyer, M. Catelli, R. M. Evans and E. Baulieu (1988). A region in the steroid binding domain determines formation of the non-DNA-binding, 9S glucocorticoid receptor complex. *J. Biol. Chem.*, **263**:267-273.

228. Propst, F., M. P. Rosenberg, A. Iyer, K. Kaul and G. F. V. 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.

229. Quian, Y. Q., M. Billeter, G. Otting, M. Müller, W. J. Gehring and K. Wüthrich (1989). The structure of the *Antennapedia* homeodomain

determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell*, **59**:573-580.

230. Quigley, F., W. F. Martin and R. Cerff (1988). Intron conservation across the prokaryote-eukaryote boundary: structure of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. *Proc. Natl. Acad. Sci. USA*, **85**:2672-2676.

231. Rao, C. D., M. Pech, K. C. Robbins and S. A. Aaronson (1988). The 5' untranslated sequence of the *c-sis*/PDGF-2 transcript is a potent translational inhibitor. *Mol. Cell. Biol.*, **8**:284-292.

232. Ratner, L., B. Thielan and T. Collins (1987). Sequences of the 5' portion of the human *c-sis* gene: characterization of the transcriptional promoter and regulation of expression of the protein product by 5' untranslated mRNA sequences. *Nucleic Acids Res.*, **15**:6017-6036.

233. Reul, J. M. H. M., P. T. Pearce, J. W. Funder and A. S. Krozowski (1989). Type I and Type II corticosteroid receptor gene expression in the rat: effect of adrenalectomy and dexamethasone administration. *Mol. Endocrinol.*, **3**:1674-1680.

234. Reul, J. M. H. M., F. R. van den Bosch and E. R. de Kloet (1987). Differential response of type I and type II corticosteroid receptors to changes in plasma steroid level and circadian rhythmicity. *Neuroendocrinology*, **45**:407-412.

235. Reynolds, G. A., J. L. Goldstein and M. S. Brown (1985). Multiple mRNAs for 3-hydroxy-3-methylglutaryl coenzyme A reductase determined by multiple transcription initiation sites and intron splicing sites in the 5'-untranslated region. *J. Biol. Chem.*, **260**:10369-10377.

236. Ritchie, H. H., M. R. Hughes, E. T. Thompson, P. J. Malloy, Z. Hochberg, D. Feldman, J. W. Pike and B. W. O'Malley (1989). An ochre mutation in the vitamin D receptor gene causes hereditary 1,25-dihydroxyvitamin D_3-resistant rickets in three families. *Proc. Natl. Acad. Sci. USA*, **86**:9783-9787.

237. Rogers, J. (1985). *Nature*, **315**:458.

238. Rosenfeld, P., J. A. M. Van Eekelen, S. Levine and E. R. De Kloet (1988). Ontogeny of the type 2 glucocorticoid receptor in discrete rat brain regions: an immunocytochemical study. *Dev. Brain Res.*, **42**:119-127.

239. Rosewicz, S., A. R. McDonald, B. A. Maddux, I. D. Goldfine, R. L. Miesfeld and C. D. Logsdon (1988). Mechanism of glucocorticoid receptor down-regulation by glucocorticoids. *J. Biol. Chem.*, **263**:2581-2584.

240. Rouault, T. A., M. W. Hentze, S. W. Caughman, J. B. Harford and R. D. Klausener (1988). Binding of a cytosolic protein to the iron-responsive element of human ferritin messenger RNA. *Science*, **241**:1207-1210.

241. Rusconi, S., R. Miesfeld, P. J. Godowski, J. N. Vanderbilt, B. Maler A. and K. R. Yamamoto (1987). Functional analysis of cloned glucocorticoid receptor sequences. In: *RNA Polymerase and the Regulation of Transcription* (ed. W. S. Reznikoff et al.) Vol. , pp. 257-266. Elsevier, New York.

242. Rusconi, S. and K. Yamamoto (1987). Functional dissection of the hormone and DNA binding activities of the glucocorticoid receptor. *EMBO J.*, **6**:1309-1315.

243. Russell, D. W., T. Yamamoto, W. J. Schneider, C. J. Slaughter, M. S. Brown and J. L. Goldstein (1983). cDNA cloning of the bovine low density lipoprotein receptor: Feedback regulation of a receptor mRNA. *Proc. Natl. Acad. Sci. USA*, **80**:7501-7505.

244. Russell, G. J., P. M. B. Walker, R. A. Elton and J. H. Subak-Sharpe (1976). Doublet frequency analysis of fractionated vertebrate nuclear DNA. *J. Mol. Biol.*, **108**:1-23.

245. Russell, J. H. and C. B. Dobos (1980). *J. Immun.*, **125**:1256-1261.

246. Russell, J. H., V. Masakowski, T. Rucinsky and G. Phillips (1982). *J. Immun.*, **128**:2087-2094.

247. Russell, J. H., V. R. Masakowski and C. B. Dobos (1980). *J. Immun.*, **124**:1100-1105.

248. 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 J.*, **8**:3327-3335.

249. Sakai, D. D., S. Helms, J. Carlstedt-Duke, J.-Å. Gustafsson, F. M. Rottman and K. R. Yamamoto (1988). Hormone-mediated repression: a negative glucocorticoid response element from the bovine prolactin gene. *Genes Dev.*, **2**:1144-1154.

250. Sakly, M. and B. Koch (1981). Ontogenesis of glucocorticoid receptors in anterior pituitary gland: transient dissociation among cytoplasmic receptor density, nuclear uptake, and regulation of corticotropic activity. *Endocrinology*, **108**:591-596.

251. Sakurai, A., A. Nakai and L. J. DeGroot (1990). Structural analysis of human thyroid hormone receptor β gene. *Mol. Cell. Endocrinol.*, **71**:83-91.

252. Salser, W. (1978). Globin mRNA sequences: analysis of base pairing and evolutionary implications. *Cold Spring Harbor Symp. Quant. Biol.*, **42**:985-1002.

253. Sanchez, E. R., S. Meshinchi, W. Tienrungroj, M. J. Schlesinger, D. O. Toft and W. B. Pratt (1987). *J. Biol. Chem.*, **262**:6986-6991.

254. Sandor, T. and A. Z. Mehdi (1979). Steroids and evolution. In: *Hormones and evolution* (ed. E. J. W. Barrington) Vol. 1, pp. 1-72. Academic Press, London.

255. Sapolsky, R. M., L. C. Krey and B. S. McEwen (1984a). Glucocorticoid-sensitive hippocampal neurons are involved in terminating the adrenocortical stress response. *Proc. Natl. Acad. Sci. USA*, **81**:6174-6177.

256. Sapolsky, R. M., L. C. Krey and B. S. McEwen (1984b). Stress down-regulates corticosterone receptors in a site-specific manner in the brain. *Endocrinology*, **114**:287-292.

257. Sapolsky, R. M., L. C. Krey and B. S. McEwen (1986). The neuroendocrinology of stress and aging: the glucocorticoid cascade hypothesis. *Endocrine Reviews*, **7**:284-301.

258. Sapolsky, R. M. and B. S. McEwen (1985). Down-regulation of neural corticosterone receptors by corticosterone and dexamethasone. *Brain Res.*, **339**:161-165.

259. Sapolsky, R. M., M. J. Meaney and B. S. McEwen (1985). The development of the glucocorticoid receptor system in the rat limbic brain. III. Negative-feedback regulation. *Develop. Brain Res.*, **18**:169-173.

260. Savakis, C., M. Ashburner and J. H. Willis (1986). The expression of the gene coding for alcohol dehydrogenase during the development of *Drosophila melanogaster*. *Dev. Biol.*, **114**:194-207.

261. Schachter, B. S., L. K. Johnson, J. D. Baxter and J. L. Roberts (1982). Differential regulation by glucocorticoids of proopiomelanocortin in ribonucleic acid levels in the anterior and intermediate lobes of the rat pituitary. *Endocrinology*, **110**:1442-1444.

262. Schapiro, S. (1962). Pituitary ACTH and compensatory adrenal hypertrophy in stress-non-responsive infant rats. *Endocrinology*, **71**:986-989.

263. Schelechte, J. A., B. H. Ginsberg and B. M. Sherman (1982). Regulation of the glucocorticoid receptor in human lymphocytes. *J. Steroid Biochem.*, **16**:69-74.

264. Schena, M., L. P. Freedman and K. R. Yamamoto (1989). Mutations in the glucocorticoid receptor zinc finger region that distinguish interdigitated DNA binding and transcriptional enhancement activities. *Genes Dev.*, **3**:1590-1601.

265. Schibler, U., O. Hagenbüchle, P. Wellauer and A. C. Pittet (1983). Two promoters of different strengths control the transcription of the mouse alpha-amylase gene *Amy-1^a* in the parotid gland and the liver. *Cell*, **33**:501-508.

266. Schneuwly, S., A. Kuroiwa, P. Baumgartner and W. J. Gehring (1986). Structural organization and sequence of the homeotic gene *Antennapedia* of *Drosophila melanogaster*. *EMBO J.*, **5**:733-739.

267. Schoenfield, N. M., J. H. Leathen and J. Rabii (1980). Maturation of adrenal stress response in the rat. *Neuroendocrinology*, **31**:101-105.

268. Schon, E., T. Evans, J. Welsh and A. Efstratiadis (1983). Conformation of promoter DNA: fine mapping of S1-hypersensitive sites. *Cell*, **35**:837-848.

269. Schüle, R., P. Rangarajan, S. Kliewer, L. J. Ransome, J. Bolado, N. Yang, I. M. Verma and R. M. Evans (1990). Functional antagonism between oncoprotein c-jun and the glucocorticoid receptor. *Cell*, **62**:1217-1226.

270. Schwabe, J. W. R., D. Neuhaus and D. Rhodes (1990). Solution structure of the DNA-binding domain of the oestrogen receptor. *Nature*, **348**:458-461.

271. Scott, M. P. and S. B. Carroll (1987). The segmentation and homeotic gene network in early Drosophila development. *Cell*, **51**:689-698.

272. Scott, M. P., J. W. Tamkun and I. George W. Hartzell (1989). The structure and function of the homeodomain. *Biochim. Biophys. Acta*, **989**:25-48.

273. Seger, M. A., J. A. M. van Eekelen, J. A. Kiss, J. P. H. Burbach and E. R. de Kloet (1988). Stimulation of pro-opiomelanocortin gene expression by glucocorticoids in the denervated rat intermediate pituitary gland. *Neuroendocrinology*, **47**:350-357.

274. Segraves, W. A. and D. S. Hogness (1990). The *E75* ecdysone-inducible gene responsible for the 75B early puff in *Drosophila* encodes two new members of the steroid receptor superfamily. *Genes Dev.*, **4**:204-219.

275. Sengstag and Hinnen (1987). *Nucleic Acids Res.*, **15**:233-246.

276. Sharp, P. A. (1985). On the origin of RNA splicing and introns. *Cell*, **42**:397-400.

277. Shephard, J. C. W., W. McGinnis, A. E. Carrasco, E. M. DeRobertis and W. J. Gehring (1984). Fly and frog homeodomains show homologies with yeast mating type regulatory proteins. *Nature*, **310**:70-71.

278. Sheppard, K. E., J. L. Roberts and M. Blum (1990). Differential regulation of type II corticosteroid receptor messenger ribonucleic acid expression in the rat anterior pituitary and hippocampus. *Endocrinology*, **127**:431-439.

279. Shih, M.-C., P. Heinrich and H. M. Goodman (1988). Intron existence predated the divergence of eukaryotes and prokaryotes. *Science*, **242**:1164-1166.

280. Shipman, G. F., C. D. Bloomfield, K. J. Gajl-Peczalska, A. U. Munck and K. A. Smith (1983). Glucocorticoids and lymphocytes. III. Effects of glucocorticoid administration on lymphocyte glucocorticoid receptors. *Blood*, **61**:1086-1090.

281. Sibley, C. H. and G. M. Tomkins (1974a). Isolation of lymphoma cell variants resistant to killing by glucocorticoids. *Cell*, **2**:213-220.

282. Sibley, C. H. and G. M. Tomkins (1974b). Mechanisms of steroid resistance. *Cell*, **2**:221-227.

283. Smith, K. and S. Shuster (1984). Effect of adrenalectomy and steroid treatment on rat skin cytosol glucocorticoid receptor. *J. Endocrinol.*, **102**:161-165.

284. Smith, L. I., J. E. Bodwell, D. B. Mendel, T. Ciardelli, W. G. North and A. Munck (1988). Identification of cysteine-644 as the covalent site of attachment of dexamethasone 21-mesylate to murine glucocorticoid receptors in WEHI-7 cells. *Biochemistry*, **27**:3747-3753.

285. Smith, R. F. and T. F. Smith (1990). Molecular Biology Computer Resource, Galleria Level 1. Dana-Farber Cancer Institute and School of Public Health, Harvard University, 44 Binney St., Boston, MA 02115 USA (617)732-3746. ARPANET rsmith@mbcrr.harvard.edu; BITNET rsmith%mbcrr@husc6.bitnet.

286. Snyder, D. S. and E. R. Unanue (1982). Corticosteroids inhibit murine macrophage Ia expression and interleukin 1 production. *J. Immunol.*, **129**:1803-1805.

287. Sousa, R. J., N. H. Tannery and E. M. Lafer (1988). In situ hybridization mapping of glucocorticoid receptor messenger ribonucleic acid in rat brain. *Mol. Endocrinol.*, **3**:481-494.

288. Stavenhagen, J. B. and D. M. Robins (1988). An ancient provirus has imposed androgen regulation on the adjacent mouse sex-limited protein gene. *Cell*, **55**:247-254.

289. Stroeher, V. L., E. M. Jorgensen and R. L. Garber (1986). Multiple transcripts from the Antennapedia gene of *Drosophila melanogaster*. *Mol. Cell. Biol.*, **6**:4667-4675.

290. Struhl, K. (1989). Helix-turn-helix, zinc-finger, and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins. *Trends Biochem.*, **14**:137-140.

291. Südhof, T. C., J. L. Goldstein, M. S. Brown and D. W. Russell (1985a). The LDL receptor gene: a mosaic of exons shared with different proteins. *Science*, **228**:815-822.

292. Südhof, T. C., D. W. Russell, J. L. Goldstein, M. S. Brown, R. Sanchez-Pescador and G. I. Bell (1985b). Cassette of eight exons shared by genes for LDL receptor and EGF precursor. *Science*, **228**:893-895.

293. Sullivan, K. A., R. T. Miller, S. B. Masters, B. Beiderman, W. Heideman and H. R. Bourne (1987). Identification of receptor contact site involved in receptor-G protein coupling. *Nature*, **330**:758-760.

294. Svec, F. and M. Rudis (1981). Glucocorticoids regulate the glucocorticoid receptor in the AtT-20 cell. *J. Biol. Chem.*, **256**:5984-5987.

295. Swartz, M. N., T. A. Trautner and A. Kornberg (1962). Enzymatic synthesis of deoxyribonucleic acid. XI. Further studies on nearest neighbor base sequences in deoxyribonucleic acids. *J. Biol. Chem.*, **237**:1961-1967.

296. Thireos, G., M. D. Penn and H. Greer (1984). 5' untranslated sequences are required for the translational control of a yeast regulatory gene. *Proc. Natl. Acad. Sci. USA*, **81**:5096-5100.

297. Timberlake, W. E. and W. C. Orr (1984). Steroid hormone regulation of sexual reproduction in achlya. In: *Hormone Action* (ed. R. F. Goldberger and K. R. Yamamoto) Vol. 3B, pp. 255-283. Plenum Press, New York.

298. Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers and J. Gralla (1973). Improved estimation of secondary structure in ribonucleic acids. *Nature New Biol.*, **246**:40-41.

299. Tomkins, G. M. (1975). The metabolic code. *Science*, **189**:760-763.

300. Tora, L., H. Gronemeyer, B. Turcotte, M.-P. Gaub and P. Chambon (1988). The N-terminal region of the chicken progesterone receptor specifies target gene activation. *Nature*, **333**:185-188.

301. Tora, L., J. White, C. Brou, D. Tasset, N. Webster, E. Scheer and P. Chambon (1989). The human estrogen receptor has two independent nonacidic transcriptional activation functions. *Cell*, **59**:477-487.

302. Treisman, J., P. Gonczy, M. Vashishtha, E. Harris and C. Desplan (1989). A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell*, **59**:553-562.

303. Tzamarias, D., D. Alexandraki and G. Thireos (1986). Multiple *cis*-acting elements modulate the translational efficiency of GCN4 mRNA in yeast. *Proc. Natl. Acad. Sci. USA*, **83**:4849-4853.

304. Ucker, D. S. (1987). Cytotoxic T lymphocytes and glucocorticoids activate an endogenous suicide process in target cells. *Nature*, **327**:62-64.

305. Ucker, D. S. and K. R. Yamamoto (1984). Early events in the stimulation of mammary tumor virus RNA synthesis by glucocorticoids. *J. Biol. Chem.*, **259**:7416-7420.

306. Umesono, K. and R. M. Evans (1989). Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell*, **57**:1139-1146.

307. Valerio, D., M. G. C. Duyvesteyn, B. M. M. Dekker, G. Weeda, T. M. Berkvens, L. van der Voorn, H. van Ormondt and A. J. van der Eb (1985). Adenosine deaminase: characterization and expression of a gene with a remarkable promoter. *EMBO J.*, **4**:437-443.

308. Vanderbilt, J. N., R. Miesfeld, B. A. Maler and K. R. Yamamoto (1987). Intracellular receptor concentration limits glucocorticoid-dependent enhancer activity. *Mol. Endocrinol.*, **1**:68-74.

309. Vedeckis, W. V. and H. D. Bradshaw Jr. (1983). DNA fragmentation in S49 lymphoma cells killed with glucocorticoids and other agents. *Mol. Cell. Endocrinol.*, **30**:215-227.

310. Vogt, P. K. and T. J. Bos (1990). *jun*: oncogene and transcription factor. *Adv. Cancer Res.*, **55**:1-35.

311. Vogt, P. K., T. J. Bos and R. F. Doolittle (1989). Homology between the DNA-binding domain of the GCN4 regulatory protein of yeast and the carboxyl-terminal region of a protein coded for by the oncogene jun. *Proc. Natl. Acad. Sci USA*, **84**:3316-3319.

312. Walker, C.-D., S. F. Akana, C. S. Cascio and M. F. Dallman (1990). Adrenalectomy in the neonate: adult-like adrenocortical system responses to both removal and replacement of corticosterone. *Endocrinology*, **127**:832-842.

313. Walker, C.-D., R. M. Sapolsky, M. J. Meaney, W. W. Vale and C. L. Rivier (1986). Increased pituitary sensitivity to glucocorticoid feedback during the stress nonresponsive period in the neonatal rat. *Endocrinology*, **119**:1816-1821.

314. Walker, C.-D., K. A. Scribner, C. S. Cascio and M. F. Dallman (1991). The pituitary-adrenocortical system of neonatal rats is responsive to stress throughout development in a time-dependent and stressor-specific fashion. *Endocrinology*, **128**:1385-1395.

315. Watanabe, M., K. Phillips and H. Watanabe (1973). Induction of steroid binding activity in *Pseudomonas testosteroni*. *J. Steroid Biochem.*, **4**:623-632.

316. Watson, M. A. and J. Milbrandt (1989). The NGFI-B gene, a transcriptionally inducible member of the steroid receptor gene superfamily: genomic structure and expression in rat brain after seizure induction. *Mol. Cell. Biol.*, **9**:4213-4219.

317. Webster, N. J. G., S. Green, J. R. Jin and P. Chambon (1988). The hormone-binding domains of the estrogen and glucocorticoid recptors contain an inducible transcription activation function. *Cell*, **54**:199-207.

318. Webster, N. J. G., S. Green, D. Tasset, M. Ponglikitmongkol and P. Chambon (1989). The transcriptional activation function located in the hormone-binding domain of the human oestrogen receptor is not encoded in a single exon. *EMBO J.*, **8**:1441-1446.

319. 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.

320. Werner, M., A. Feller, F. Messenguy and A. Piérard (1987). The leader peptide of yeast gene *CPA1* is essential for the translational repression of its expression. *Cell*, **49**:805-813.

321. Wilson, A. C., S. S. Carlson and T. J. White (1977). Biochemical evolution. *Ann. Rev. Biochem.*, **46**:573-639.

322. Witek-Janusek, L. (1988). Pituitary-adrenal response to bacterial endotoxin in developing rats. *Am. J. Physiol.*, **255**:E525-E530.

323. Wood, W. I., J. Gitschier, L. A. Lasky, and R. M. Lawn (1985). Base composition-independent hybridization in tetramethylammonium chloride: A method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. U.S.A.*, **82**:1585-1588.

324. Wyllie, A. H. (1980). Glucocorticoid-induced thymocyte apoptosis is associated with endogenous endonuclease activation. *Nature*, **284**:555-556.

325. Yamamoto, K. R. (1985). Steroid receptor regulated transcription of specific genes and gene networks. *Ann. Rev. Genet.*, **19**:209-252.

326. Yamamoto, K. R. (1989). A conceptual view of transcriptional regulation. *Amer. Zool.*, **29**:537-547.

327. Yamamoto, K. R., U. Gehring, M. R. Stampfer and C. H. Sibley (1976). Genetic approaches to steroid hormone action. *Rec. Prog. Horm. Res.*, **32**:3-32.

328. Yamamoto, K. R., P. J. Godowski and D. Picard (1988). Ligand-regulated nonspecific inactivation of receptor function: a versatile mechanism for signal transduction. *Cold Spring Harbor Symp. Quant. Biol.*, **53**:803-811.

329. Yamamoto, K. R., M. R. Stampfer and G. M. Tomkins (1974). Receptors from glucocorticoid-sensitive lymphoma cells and two classes of insensitive clones: physical and DNA-binding properties. *Proc. Natl. Acad. Sci. USA*, **71**:3901-3905.

330. Yang-Yen, H.-F., J.-C. Chambard, Y.-L. Sun, T. Smeal, T. J. Schmidt, J. Drouin and M. Karin (1990). Transcriptional interference between c-Jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein-protein interaction. *Cell*, **62**:1205-1215.

331. Yuh, Y. and E. B. Thompson (1987). Complementation between glucocorticoid receptor and lymphocytolysis in somatic cell hybrids of two glucocorticoid-resistant human leukemic clonal cell lines. *Somatic Cell Mol. Genet.*, **13**:33-46.

332. Zahraoui, A. and G. Cuny (1987). Nucleotide sequence of the chicken proto-oncogene c-*erbA* corresponding to domain 1 of v-*erbA*. *Eur. J. Biochem.*, **166**:63-69.

333. Zong, J., J. Ashraf and E. B. Thompson (1990). The promoter and first, untranslated exon of the human glucocorticoid receptor gene are GC rich but lack consensus glucocorticoid receptor element sites. *Mol. Cell. Biol.*, **10**:5580-5585.

Appendix One: Materials and Methods

Cell Culture and Transfection

Monkey CV-1 cells and human HeLa cells were grown in Dulbecco's modified Eagle's medium (GIBCO) supplemented with 5% fetal calf serum (Hyclone). Rat J2.17, XC, and 19G.11 cells were supplemented with 10% defined-supplemented calf serum (Hyclone). Mouse S49 cells were supplemented with 10% horse serum (Hyclone) and 4.5 g/l glucose. Rat EDR3 cells were grown in

Cells were transfected with 6 μg DNA on 60 mm dishes or 17 μg on 100 mm dishes by the calcium phosphate coprecipitation technique (Graham and van der Eb, 1973; Wigler et al., 1978) with the modifications of (Picard and Yamamoto, 1987): after incubating the cells with the coprecipitate for 15-18 h, cultures were washed with 5 ml TBS (25 mM Tris-HCl, 137 mM NaCl, 5 mM KCl, 0.7 mM CaCl2, 0.5 mM MgCl2, 0.6 mM Na2HPO4, pH 7.4; (Kimura and Dulbecco, 1972) to dissolve the precipitate, and then were incubated with fresh medium. Where indicated, dexamethasone was added concomitantly with the DNA and with the fresh medium. S49 cells were transiently transfected by electroporation or stably transfected by retroviral-mediated DNA transfer, as described below.

CAT Assays

Following transfection, cells were incubated for an additional 36 hours (total time 48 hours), harvested, and extracts prepared by three freeze-thaw cycles (-70°C, 37°C or 68°C) and centrifugation at 15,000g. Samples that were not to be assayed for β-galactosidase activity were thawed at 68°C. Samples to be assayed for β-galactosidase activity were thawed at 37°C and heat-treated portions (5 minutes, 68°C) subjected to either a chromatographic (Gorman et al., 1982) or nonchromatographic CAT assay (Sleigh, 1986); a second portion

was subjected to a β-galactosidase assay (Stuart et al., 1984). The amount of protein and/or incubation times were adjusted to keep the CAT assays within a linear range (<30% conversion).

Cloning, Library Screening, and Analysis of RNA and DNA

Except as indicated in the figure legends or text, standard recombinant DNA techniques (Ausubel et al., 1989; Sambrook et al., 1989) were used.

Electroporation

Mouse S49 cell lines were transfected by electroporation, using a device similar to that described by (Chu et al., 1987) except that the capacitance was 1300 μF instead of 530 μF. Cells were washed once in HBS (21 mM HEPES (pH 7.05), 137 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4, 6 mM glucose) and resuspended to 1.25×10^7 cells/ml in HBS. Cells (0.8 ml) and DNA (0.5 mg total plasmid plus salmon-sperm DNA carrier in 0.2 ml HBS) were mixed and electroporated in disposable electroporation cuvettes (0.4 cm, Bio Rad) at 225 V, 1300 μF. These settings were found to give maximal transfection efficiencies (CAT activity / sample). Following a 10 - 20 min incubation at room temp, electroporated cells were transferred to T-25 flasks containing medium (DME-H21 + 10% defined-supplemented calf serum) pre-equilibrated at 37 °C.

Rapid Amplification of cDNA Ends (R.A.C.E.)

The second strand synthesis and amplification steps described below are based on the RACE procedure with nested primers developed by Mike Frohman (Frohman et al., 1988; Frohman and Martin, 1989), except that the second strand synthesis and amplification steps have been separated.

cDNA synthesis. 1.5 µg poly-A selected RNA in 9 µl H_2O was denatured with 1 µl 0.1M methyl-mercuric hydroxide for 5 min at room temperature. 10 µl 0.14M β-mercaptoethanol + 40 u RNaseIN (Promega) were added and the mixture incubated at room temp. for another 5 min. cDNA was then synthesized by adding 10 µl 5XM-MLV RT reaction buffer (BRL), 5 µl acetylated BSA (1 mg/ml, Promega), 10 µl dNTP mix (2.5 mM each dNTP), 0.5 µl (α-^{32}P)dGTP (3,000 Ci/mmol, Amersham) as a tracer, 2 pmol primer MJ04, and (lastly) 3 µl M-MLV Reverse Transcriptase (200 u/µl, BRL) in a final volume of 50 µl. After 1 min at room temp., the reaction mixture was incubated for 2 h at 37-39°C. The reaction was stopped and RNA hydrolyzed by adding 4 µl 0.5M EDTA (pH 8.0) and 2.5µl 7.5 M NaOH/0.25 M EDTA and incubating at 70°C for 30 min. The reaction was neutralized with 7.5 µl 2.5M HCl and 3 µl 1M Tris-HCl (pH 8.0), then desalted and primer removed by centrifuging three times (1000 x g, 20 min.) through a Centricon-100 filter with 2 ml TE. Before the last spin, the concentrate was resuspended in 0.15 X TE; afterwards the concentrate was removed from the filter, vacuum dried, and resuspended in 20 µl H_2O.

Tailing. The antisense cDNA strand (20 µl) was tailed with $(dA)_n$ by adding 6 µl 5 X Tailing Buffer, 1 µl 6mM dATP, and 1 µl TdT (15u, BRL), and incubated for 25 min at 37°C followed by 15 min at 68°C. The tailing reaction was desalted by two centrifugations in a Centricon-100 (as above, with the last spin in dilute TE) and then dried and resuspended in 20 µl H_2O.

Second Strand Synthesis. The sense cDNA strand was synthesized by mixing the tailed cDNA (in 20 µl H_2O in a 0.5 ml tube) with 2 µl RoRi(dT)17 primer (25 pmol), 3 µl 25 mM dNTP mix, and 25 µl 2 X Taq Buffer (33.2 mM $(NH_4)_2SO_4$, 134 mM Tris-HCl (pH 8.8 at 25 °C), 13.4 mM $MgCl_2$, 20 mM 2-mercaptoethanol, 340 µg/ml BSA). The reaction mix was heated in a thermal

cycler to 95°C for 5 min, then incubated at 75°C. Taq Polymerase (Perkin-Elmer/Cetus) was then added (0.5 µl, 5 U/µl), the condensate from the lid transferred to the tube, and the reaction mixed and overlayed with mineral oil (pre-equilibrated to 75°C). The mixture was then incubated for one cycle of denaturation (95°, 1 min), annealing (55°, 2 min) and polymerization (72°, 40 min), extracted with chloroform, and diluted to 0.5 ml in TE.

Amplification. cDNA was amplified in two rounds with a nested set of primers. The first round (outer nest) primers were Ro (5' end of primer sequence RoRi(dT)17) and GSP1 (gene-specific primer 1, to GR sequence upstream of RT primer MJ04). The second round (inner nest) primers were Ri and GSP2 (upstream of GSP1). For the first round, 2 µl cDNA was mixed with 1 µl GSP1 (25 pmol), 1 µl Ro (25 pmol), 12.5 µl 4 X Taq Buffer, 3 µl 25 mM dNTP mix, and 30.5 µl H_2O (for 50 µl final volume) in a 0.5 ml tube. The reaction mix was heated in a thermal cycler to 95°C for 5 min then cooled to 75°C. Taq polymerase (2.5U) was added, the mixture overlayed with pre-heated mineral oil, and the cDNA amplified in 30 cycles of 94°C, 40s; 55°C, 60s; 72°C, 90s. The second round was run exactly as the first, using 1 µl of a 1:20 dilution of the first round products instead of cDNA, and using primers Ri and GSP2.

In addition to the amplification reactions described above, control reactions without DNA or with only one primer were run in parallel. Products of control and experimental reactions from both first and second rounds were run on a 1.5% agarose gel, transferred to Zetaprobe (BioRad) nylon filters, and probed with a receptor-specific probe to verify that receptor sequences were amplified.

Cloning. Second round amplification products were extracted with chloroform, desalted by Centricon-100 centrifugation, and digested with

restriction enzymes *Oxa*NI (isoschisomer of *Mst*II) and *Cla*I (there is an *Oxa*NI site 120bp downstream of the 5' end of exon 2, and a *Cla*I site in the Ri primer sequence). The digestion products were then ligated into *Oxa*NI-*Cla*I digested plasmid vector pBS(MstII) (see appendix 2).

RNase Protection

Antisense RNA probes. DNA templates were prepared by digesting GR DNA-containing Bluescript-based plasmids with a restriction enzyme to cut at the downstream side of the desired antisense probe. ^{32}P-labelled antisense RNA was synthesized by mixing 1.0 µl of template DNA (1 mg/ml) with (in the following order): 0.9-1.4µl distilled water, 0.7 µl 10 mM NTP mix (10 mM each UTP, GTP, ATP), 0.4 µl 1 M DTT, 6.1 µl pre-diluted (α^{32}P) CTP {5.0 µl hot nucleotide (800 Ci/mmol, 40 mCi/ml) plus 1.1 µl cold nucleotide (0.5 mM) to yield 250 Ci/mmol, 0.13 mM}, 0.7 µl RNaseIN, 2.7 µl 5xT7/T3 buffer (), and 0.5-1.0µl T3 or T7 RNA polymerase (25-100u), all to a final volume of 13.5 µl. This mixture was incubated at 37°C for 1 hour. Template DNA was digested by adding 4 µg RNase-free DNase I (Worthington) in 36 µl DNAse I buffer (50mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$) and incubating at 37°C for 15 min. 2 µl EDTA (0.5M) and 20 µg tRNA carrier were added, labelled RNA extracted with phenol/chloroform, and RNA separated from unincorporated nucleotide by one NaOAc-ethanol precipitation followed by one ammonium acetate-ethanol precipitation. Labelled antisense RNA was resuspended in RNase-free distilled water and radioactive concentration determined by scintillation counting (in scintillant). From the radioactive concentration of this mixture, the specific activity of radioactive nucleotide in the transcription reaction, and the sequence of the antisense RNA, it was then possible to calculate the concentration of transcripts as follows:

$$\text{Transcript Concentration} = 0.45 \times \left\{\frac{R}{S \times N}\right\} \text{ fmol}/\mu\text{l},$$

where R = radioactive concentration, dpm/μl

S = specific activity of radioactive nucleotide, Ci/mmol

N = number of hot nucleotides per transcript.

RNase Protection. One to two fmol of fresh 32-P-labelled antisense GR mRNA probe (250 Ci/mmol labelled nucleotide) was mixed with fifty micrograms of total RNA (or control sense RNA plus tRNA carrier) and dried *in vacuuo*. Dried RNA was resuspended in 8μl deionized formamide, mixed with 2μl 5xhybridization buffer (200mM Na_2•PIPES (pH 6.4), 2M NaCl, 10mM EDTA), heated to 90°C for 5 min., and incubated at 50°C overnight (6-16h). Hybridized RNA was then digested by adding RNase A (3 μg/ml) and RNase T1 (.15 U/ml) in 0.3 ml digestion buffer (0.3 ml NaOAc (pH 7.0), 5 mM EDTA) and incubating at 30°C for 1 hour. Digestion was stopped by transferring the reactions to fresh tubes containing 5 μl Proteinase K (10 mg/ml), 10 μl 10% SDS, and 5 μg tRNA and incubating at 50°C for 15 min. RNA was extracted with phenol/chloroform, precipitated twice in ethanol (the second extraction in ammonium acetate/ethanol), washed with 100% ethanol, dried, resuspended in formamide loading buffer, and electrophoresed on a 4.5% acrylamide gel containing 1XTBE and 8M urea.

RNaseH/Northern Transcript Mapping

Transcription initiation sites were determined by hybridizing mRNA to a synthetic oligonucleotide complementary to a portion of the 5' exon2 region of GR mRNA. This hybrid was treated with RNase H (which specifically digests RNA in the region of DNA•RNA hybridization), yielding two GR mRNA fragments: a 5' fragment extending from the transcriptional

initiation sites to the beginning of the region of hybridization and a 3' fragment extending from the end of the hybrid to the 3' termini of GR mRNA (Fig. A1.2). The sites of GR mRNA transcriptional initiation from promoter B were deduced from the size of the 5' GR mRNA fragment, the location of the region of DNA•RNA hybridization, and the rat GR gene DNA sequence (Fig. 2.15). Fifteen micrograms of poly-A+ selected RNA was mixed with 300 pmol of MJ04 (a synthetic oligodeoxynucleotide 20 nucleocleotides in length and complementary to bases 200 - 220 of the rat GR gene; see appendix 2). This mixture was dried *in vacuuo*, resuspended into 18µl of annealing buffer (125mM KCl, 10mM Tris-HCl (pH 8.0), 0.1mM EDTA, 10 U RNaseIN® (Promega) per tube), heated to 70°C for 5 min, and then incubated at 37°C for 30 min. RNase H (Boehringer Mannheim Biochemicals) was added in 12 µl of buffer (130 U/ml RNaseH, 70mM Tris-HCl (pH 8.0), 25mM $MgCl_2$, 2.5 mM DTT, 2.2 U RNaseIN® per µl), and the samples incubated at 37°C for 30 min. The reaction was stopped by adding 2 µl 0.5M EDTA, 1.5 µl 10% SDS, and digesting with 30 µg proteinase K at 37°C for 15 min. RNA was extracted with phenol/chloroform, precipitated in ethanol, resuspended in formamide, electrophoresed on a 3.5% acrylamide 1XTBE sequencing gel containing 8M urea, and electroblotted (0.8mA/cm^2, 1h, 1X TBE buffer) onto GeneScreen Plus (DuPont). Radioactive size markers (pBr325 x HaeIII) were run on the gel alongside the sample RNAs. Filters were probed with ^{32}P-labelled (1500 Ci/mmol) antisense RNA complementary to exon 1B (from p5'RdN93) or exon 2 (from pE1B(SS)). Filter hybridization was at 50°C in 1.5X SSPE / 50% Formamide / 1%SDS / 1X Denhart's / 0.2 mg tRNA per ml for 24 hours. Filters were washed once in 2X SSPE / 0.1%SDS (RT) followed by three 30 min washes at 50°C in 0.1XSSPE / 0.1%SDS. Filters were then washed in 1X TNE (10mM Tris, pH 7.5/150mM NaCl/10mM EDTA) to remove SDS, incubated in

1X TNE + 10 μg/ml RNase A at 37°C for 15 min, and then rinsed in 40 mM $NaPO_4$ (pH 7) / 1%SDS to remove RNase A. Washed filters were exposed to Kodak X-AR film with an intensifying screen at -70°C.

References

1. Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, eds. (1989). Current Protocols in Molecular Biology. Wiley & Sons, New York.

2. Chu, Hayakawa and Berg (1987). *Nucl. Acids Res.*, **15**:1311-1326.

3. Frohman, M. A., M. K. Dush and G. R. Martin (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA*, **85**:8998-9002.

4. Frohman, M. A. and G. R. Martin (1989). Rapid amplification of cDNA ends using nested primers. *Technique*, **1**:165-170.

5. Gorman, C. M., L. F. Moffat and B. H. Howard (1982). Recombinant genomes which express chlorapmphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.*, **2**:1044-1051.

6. Graham, F. L. and A. J. van der Eb (1973). *Virology*, **52**:456-467.

7. Kimura, G. and R. Dulbecco (1972). *Virology*, **49**:394-403.

8. Picard, D. and K. R. Yamamoto (1987). Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. *EMBO J.*, **6**:3333-3340.

9. Sambrook, J., E. F. Fritsch, and T. Maniatis (1989). Molecular Cloning: A Laboratory Manual (2nd ed.). Cold Spring Harbor Laboratory Press.

10. Sleigh, M. J. (1986). *Anal. Biochem.*, **156**:251.

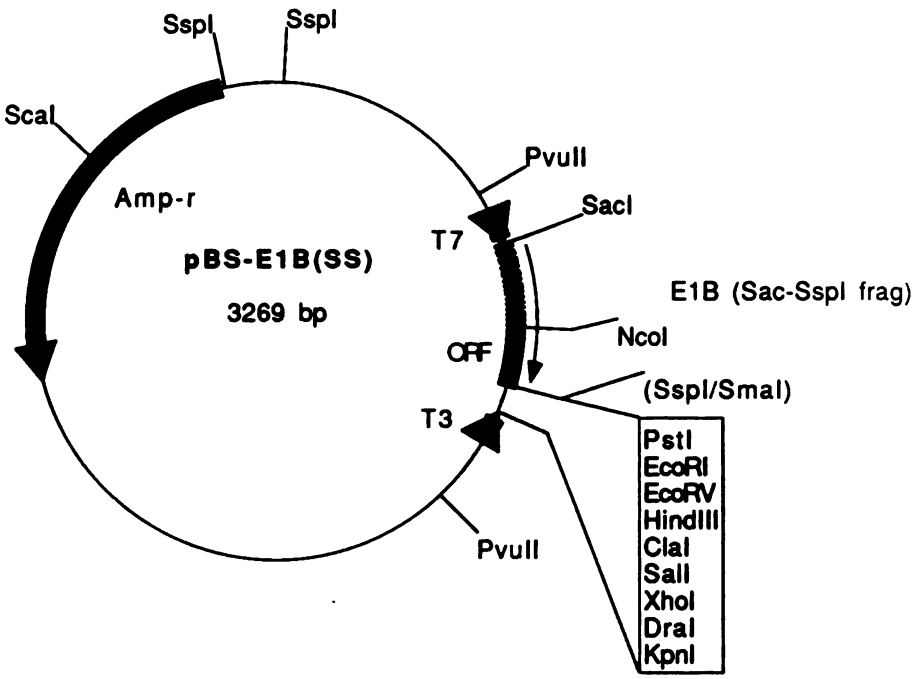
11. Stuart, G. W., P. F. Searle, H. Y. Chen, R. L. Brinster and R. D. Palmiter (1984). *Proc. Natl. Acad. Sci. U.S.A.*, **81**:7318.

12. Wigler, M., A. Pellicer, S. Silverstein and R. Axel (1978). *Cell*, **14**:725-731.

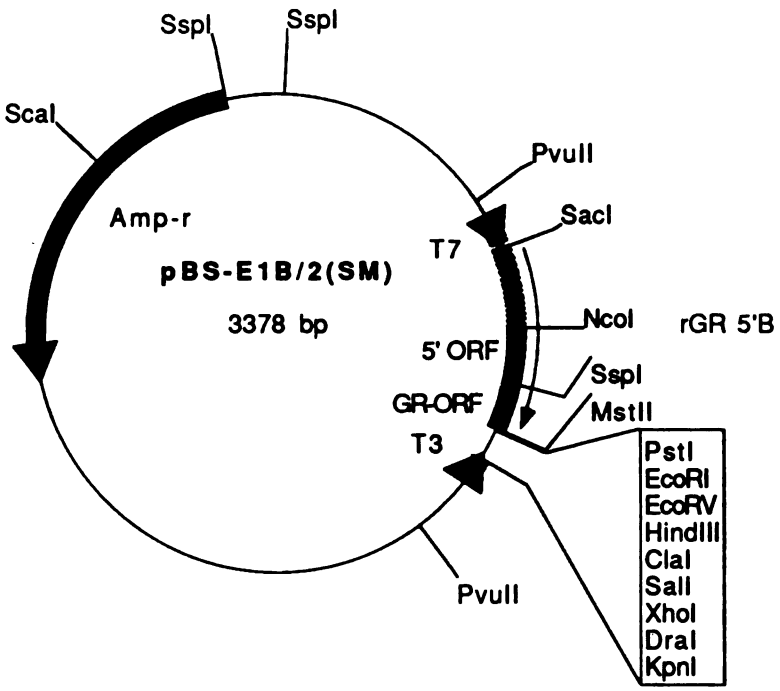
Appendix Two: Plasmids and Oligonucleotides

Oligonucleotides

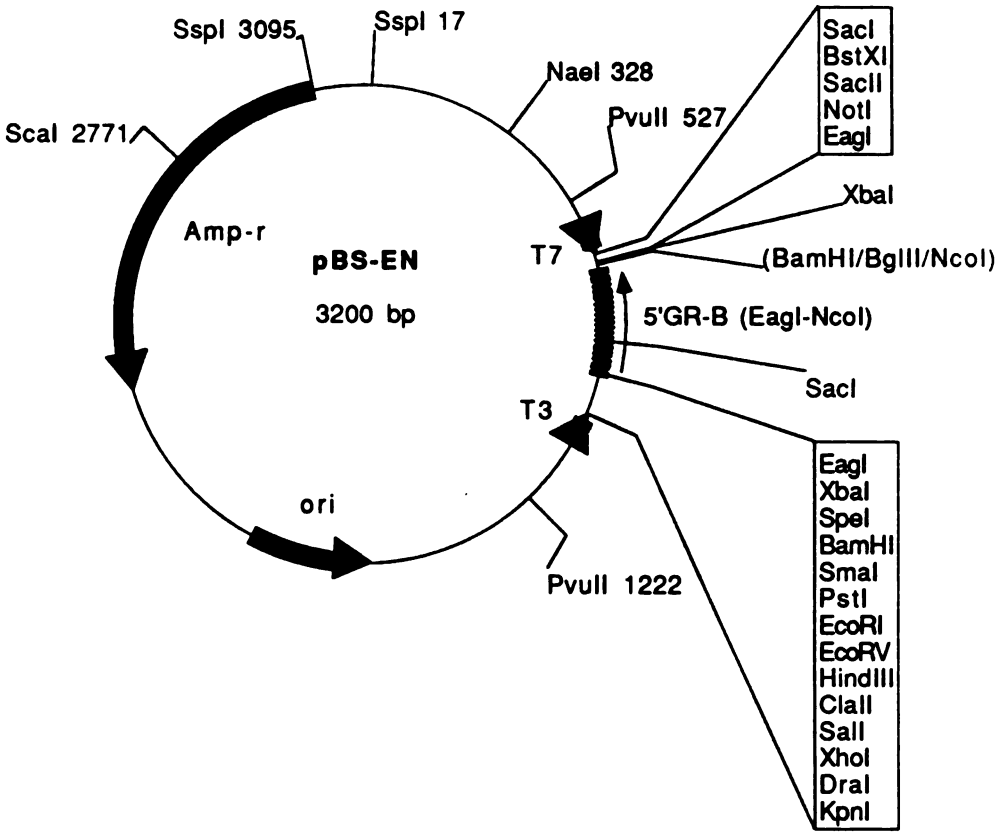
Name(s)	Length	Sequence
MJ04	20	5'- CATTGCTTGTGGAGCCTTTC
MJ07, GSP1	17	5'- GAATCTGCCTGAGAAGC
MJ08, GSP2	17	5'- GAAACCTTGACTGTAGC
Ro	17	5'- AAGGATCCGTCGACATC
Ri	17	5'- GACATCGATAATACGAC
RoRi(dT)17	57	5'- AAGGATCCGTCGACATCGAT AATACGACTCACTATAGGGA TTTTTTTTTTTTTTTTT
SR01	21	5'- GGATTCTTTGGAGTCCATTGG



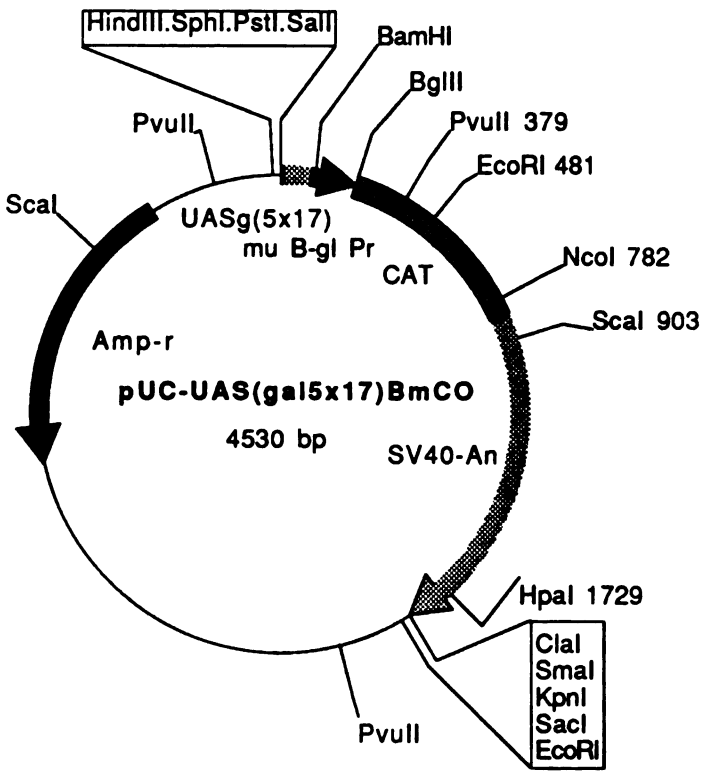
Plasmid name: pBS-E1B(SS)
Plasmid size: 3269 bp
Constructed by: Michael Jacobson
Construction date: 1/20/88
Comments/References: SacI-SspI subclone of pE1B/2 (SM): SacI-Nco from genomic DNA, NcoI-SspI from cDNA (pAR10). pBluescript backbone.



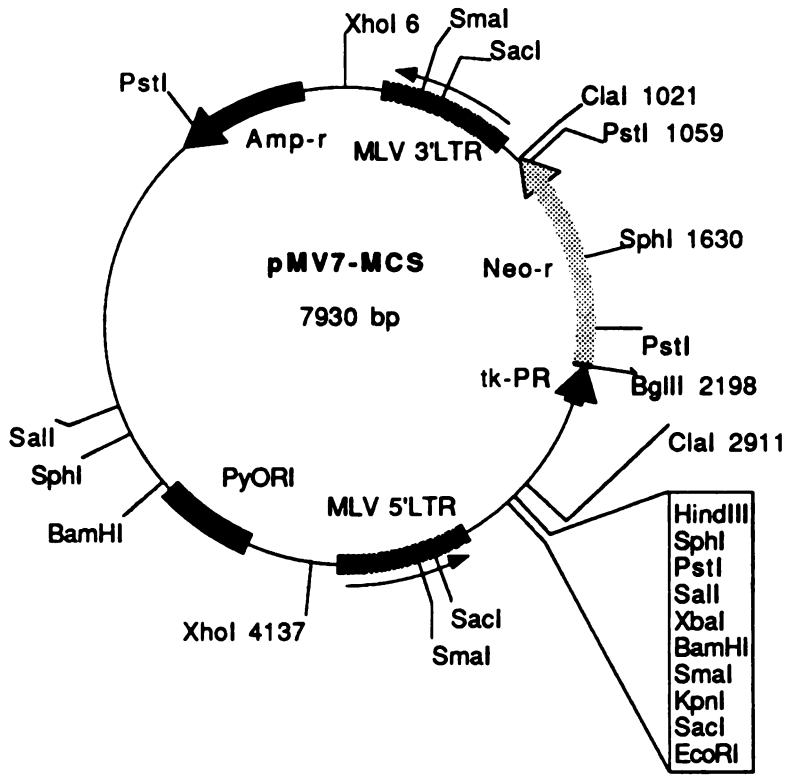
Plasmid name: pBS-E1B/2(SM)
Plasmid size: 3378 bp
Constructed by: Michael Jacobson
Construction date: 7/22/87
Comments/References: SacI-Nco from genomic DNA, NcoI-MstII from cDNA (pAR10); pBluescript backbone. MstII(filled in)-SmaI fusion reconstituted MstII site; reading frame at 3' end = AGC CTG AGG GCT GCA G...



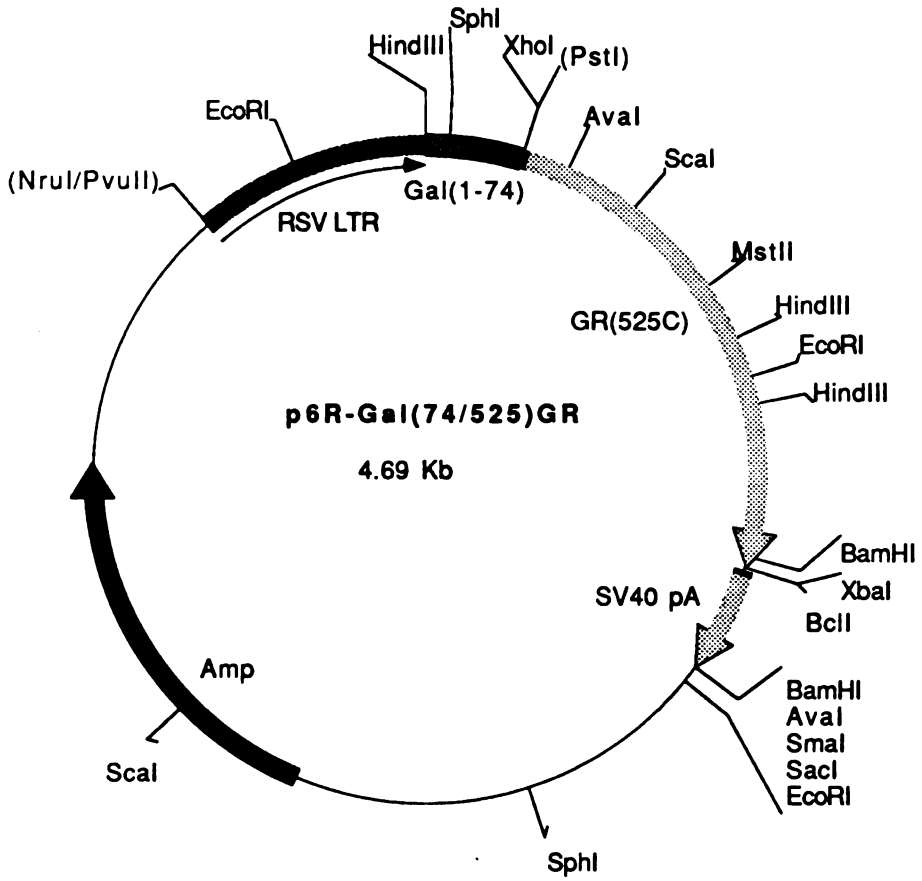
Plasmid name: pBS-EN
Plasmid size: 3200 bp
Constructed by: Michael Jacobson
Construction date: 4/89
Comments/References: EagI-NcoI 5' GR Exon 1B + 5'flanking genomic fragment. Constructed by genomic/cDNA fusion. Linearize with BamHI for antisense T7 RNA.



Plasmid name: pUC-UAS(gal5x17)BmCO
Plasmid size: 4530 bp
Constructed by: Michael Jacobson
Construction date: 11/89
Comments/References: Mouse beta-globin promoter (-106 to +26) from pBetaMT-F7 (L. Stuve, 3/89). UASgal (5x17mer) from p5UAS/pUC18. pUC18 backbone; CAT-SV40 from pUC-OOCO.

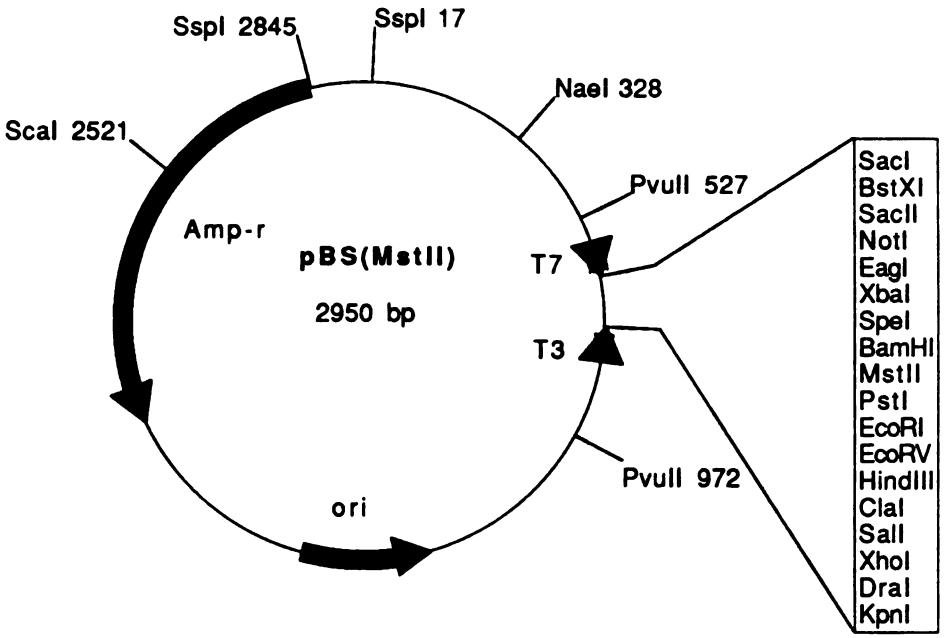


Plasmid name: pMV7-MCS
Plasmid size: 7930 bp
Constructed by: Mike Jacobson
Construction date: 11/14/88
Comments/References: pMV7 with pUC18 polylinker. Unique sites: HindIII, XbaI, KpnI, EcoRI. pMV7 ref.: Kirschmeier, et al. (1988) DNA 7: 219-225.

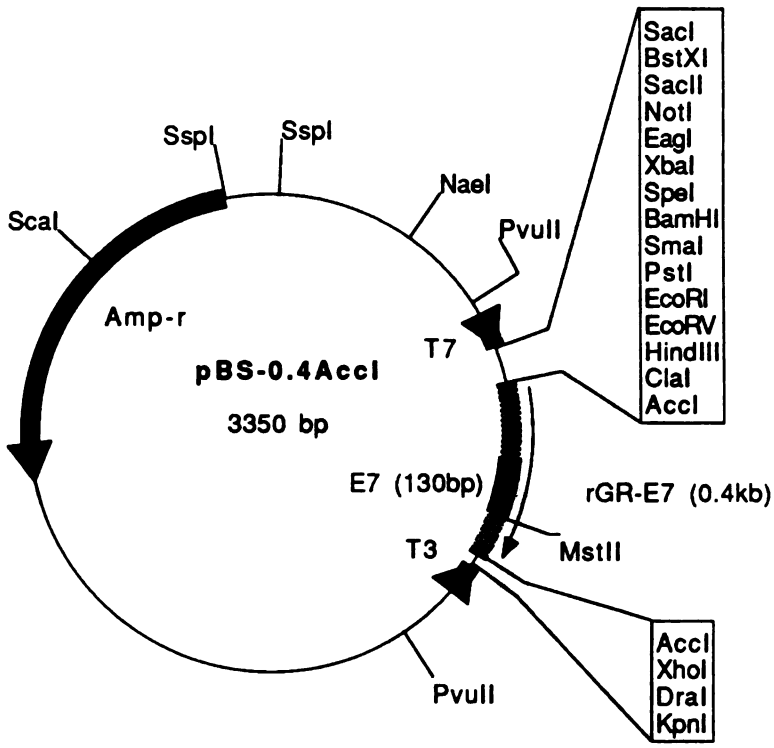


Plasmid name: p6R-Gal(74/525)GR
Plasmid size: 4.69 kb
Constructed by: Mike Jacobson
Construction date: ?
Comments/References: Mammalian RSV expression plasmid for Gal-GR fusion protein, containing Gal (1-74) fused to rGR (525C). In-frame fusion at XhoI/PstI sites:

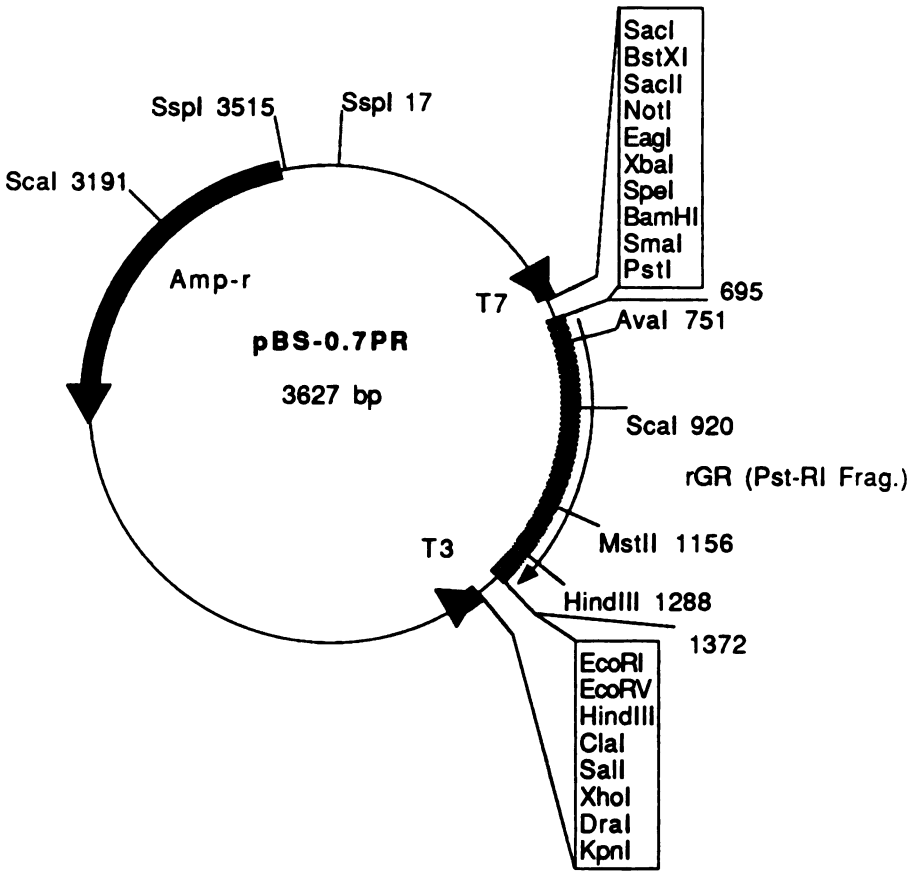
Gal . . . Pro Arg / Gly . . . rGR .
. . . CCT CGA GGA



Plasmid name: pBS(MstII)
Plasmid size: 2950 bp
Constructed by: Michael Jacobson
Construction date: 11/89
Comments/References: Modification of pBluescript + (Stratagene): MstII site at Sma site in polylinker. Constructed by ligating MstII-Sca (filled-in) from pBS-E1B/2 to Sca-Sma from pBS. Predicted sequence at MstII site: 5'-GGATCCCCCTGAGGGCTGC-3'



Plasmid name: pBS-0.4AccI
Plasmid size: 3350 bp
Constructed by: Michael Jacobson
Construction date: 8/10/87
Comments/References: 0.4 kb AccI fragment from rat genomic clone (lambda GR1) containing exon 7; subcloned from p1BK.



Plasmid name: pBS-0.7PR
Plasmid size: 3627 bp
Constructed by: Michael Jacobson
Construction date: 5/88
Comments/References: p0.7PR (Miesfeld et al. (1985) In: Sequence Specificity in Transcription and Translation. 30: 535-545) Pst-RI fragment recloned into pBluescript. Linearize with BamHI or XbaI for T3 antisense RNA; EcoRI for T7 sense RNA.

Constructed by: Mike Jacobson

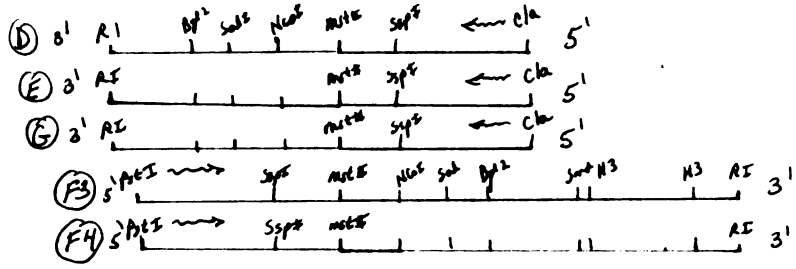
DATE: 10/29/90

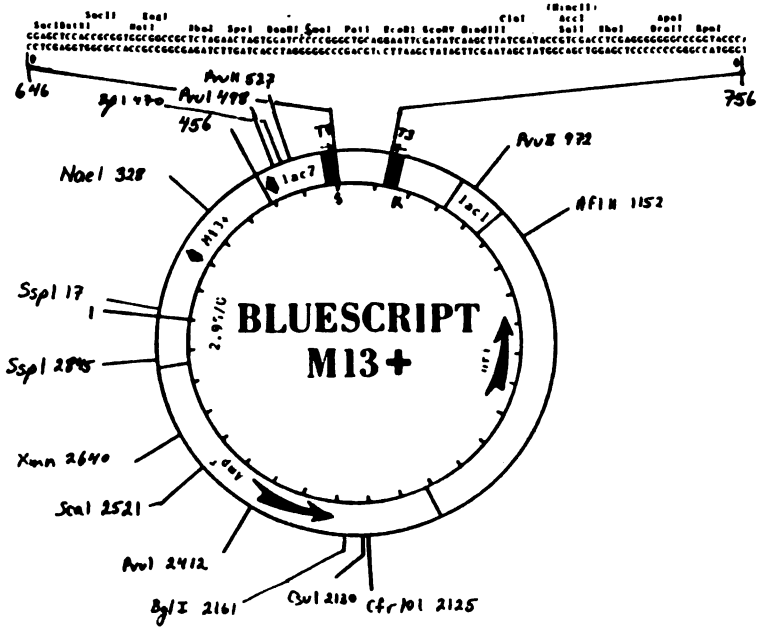
PLASMID NAME: pGR. D, E, F3, F4, G

VECTOR: pBLUESCRIPT (+)

INSERTS: 5' GR sequence from PCR clones from 5' end to SspI site; downstream sequence from E182 (SspE-mstII) and pAR8 (D,E,G) or pAR1 (F3,F4).

RELEVANT FUNCTIONAL ELEMENTS:
(promoters etc.)





DRUG RESISTANCE: AMP

COMMENTS:
(unique sites, deletions, yield, etc.)

REFERENCE:

SOURCE: