The EMBO Journal vol.5 no.10 pp.2513–2522, 1986

The mouse glucocorticoid receptor: mapping of functional domains by cloning, sequencing and expression of wild-type and mutant receptor proteins

Mark Danielsen, Jeffrey P.Northrop and Gordon M.Ringold

Department of Pharmacology, Stanford University School of Medicine, Stanford, CA 94305, USA

Communicated by M.Beato

We have isolated mouse glucocorticoid receptor (GR) cDNAs which, when expressed in transfected mammalian cells, produce a fully functional GR protein. Sequence analysis reveals an open reading frame of 2349 bp which could encode a protein of ~86 000 daltons. We have also isolated two receptor cDNAs from mouse S49 nuclear transfer-deficient (nt^-) cells which encode mutant forms of the receptor protein. One cDNA encodes a protein that is unable to bind hormone and represents the endogenous hormone binding deficient receptor recently discovered in S49 cells. The lesion in this receptor is due to a single amino acid substitution (Glu-546 to Gly). The second cDNA from nt^- cells produces a receptor protein that is able to bind hormone but has reduced nuclear binding. This cDNA, therefore, encodes for the S49 nt^- receptor which has been shown to have reduced affinity for DNA. The lesion maps to a single amino acid substitution (Arg-484 to His) located in a highly Cys, Lys, Arg-rich region of the protein previously implicated in DNA binding. Our studies provide unambiguous identification of receptor domains and specific amino acids critical for the hormone and DNA binding properties of this transcriptional regulatory protein. Contained within the first 106 amino acids of the mouse GR is a stretch of nine glutamines with two prolines which are related to the family of transcribed repetitive elements, *opa*, found in *Drosophila melanogaster*. A truncated receptor lacking these 106 amino acids is functionally indistinguishable from the wild-type receptor.

Key words: glucocorticoid receptor/nuclear transfer-deficient cells/receptor proteins

Introduction

Glucocorticoids and other steroid hormones mediate physiological responses in target cells via an interaction with specific, high affinity receptors. Binding of hormone to these receptors results in a structural alteration of the receptor, thereby allowing the steroid–receptor complex to bind tightly to sites within the nucleus (for reviews, see Yamamoto and Alberts, 1976; Ringold, 1985). The association of steroid–receptor complexes with specific DNA sequences in the 5'-flanking regions of target genes such as those encoding the mouse mammary tumor virus (MMTV) (Payvar *et al.*, 1983; Scheidereit *et al.*, 1983), chicken lysozyme (Renkawitz *et al.*, 1984), and human metallothionein-IIa (hMT-IIa) (Karin *et al.*, 1984) appears to be important for transcriptional activation of their respective promoters. However, the detailed mechanisms by which the steroid–receptor complex regulates initiation of gene transcription remain unclear, due largely to our poor understanding of receptor structure and receptor interactions with other components of the transcriptional apparatus.

The isolation and characterization of lymphoma cell variants resistant to the cytolytic effects of glucocorticoid hormones has revealed a wealth of information on the structure and function of the glucocorticoid receptor. Three classes of lymphoma cell variants have been characterized. The r^- (receptorless) variant contains very little and in some cases no specific steroid binding (Sibley and Tomkins, 1974a), while the nt^- (nuclear transfer-deficient) variant contains a receptor which binds hormone but subsequently has reduced affinity for DNA (Gehring and Tomkins, 1974; Yamamoto *et al.*, 1974). The nt^i (nuclear transfer increased) variant receptor has an elevated affinity for nuclei and DNA after binding hormone (Yamamoto *et al.*, 1974).

Recently we and others have shown that, in addition to the well characterized wild-type, nt^i, and nt^- receptors, S49 mouse lymphoma cells contain a receptor which is immunologically reactive but has no hormone binding activity (Northrop *et al.*, 1985; Westphal *et al.*, 1984). Biochemical analysis of the glucocorticoid receptors (GRs) present in these variant cells (Pfahl *et al.*, 1978a; Yamamoto *et al.*, 1976) and studies of partial proteolysis products of these receptors (Gehring and Hotz, 1983; Carlstedt-Duke *et al.*, 1982; Dellweg *et al.*, 1982) has led to the following model for the structure of the GR. The receptor is a 92 000–94 000 dalton polypeptide containing distinct functional domains in which the hormone and DNA binding regions reside on an ~40 000-dalton carboxy-terminal fragment that can be generated by mild chymotrypsin treatment of wild-type receptor. This 40 000-dalton fragment appears to be similar to the truncated form of the wild-type receptor present in the nt^i variant (Dellweg *et al.*, 1982). Further digestion of the 40 000-dalton fragment with trypsin produces a polypeptide of ~25 000 daltons which can bind hormone but not DNA (Wrange *et al.*, 1984; Gehring and Hotz, 1983). The remainder of the molecule harbors the main antigenic sites of the receptor and has been postulated to have a role in modulating DNA binding activity (Carlstedt-Duke *et al.*, 1982; Dellweg *et al.*, 1982). This supposition stems from the following observations: (i) the nt^i receptor, which lacks the amino-terminal ~50 kd, is non-functional in mediating S49 cell killing and has a higher affinity for non-specific DNA and (ii) the ~40 kd chymotryptic fragment of wild-type receptors retains specific DNA binding; however, there appears to be a reduction in its ability to discriminate between specific and non-specific sites (Payvar and Wrange, 1984).

The amino acid sequence of the human GR has recently been deduced from a full length cDNA clone (Hollenberg *et al.*, 1985). The most notable feature of the sequence is a highly basic and cysteine-containing domain present in the C-terminal half of the protein. This region shows strong homology with the human and chicken estrogen receptor (Greene *et al.*, 1986; Krust *et al.*, 1986) and with the product of the v-*erbA* gene of avian erythroblastosis virus (Weinberger *et al.*, 1985). Based on the fact that this region is highly basic and contains an arrangement of conserved cysteine residues reminiscent of other DNA binding proteins (Brown

```
                                                                                                                    60
GGAAGTTAATATTTGCCAATGGACTCCAAAGAATCCTTAGCTCCCCCTGGTAGAGACGAAGTCCCCAGCAGTTTGCTTGGCCGGGGGAGGGGAAGCGTGATGGACTTGTATAAAACCCTG
......***.........MetAspSerLysGluSerLeuAlaProProGlyArgAspGluValProSerSerLeuLeuGlyArgGlyArgGlySerValMetAspLeuTyrLysThrLeu
                     1                                                                                                  180
AGGGGTGGAGCTACAGTCAAGGTTTCTGCGTCTTCACCCTCAGTGGCTGCTGCTTCTCAGGCAGATTCCAAGCAGCAGAGGATTCTCCTTGATTTTTCAAAAGGCTCAGCAAGCAATGCA
ArgGlyGlyAlaThrValLysValSerAlaSerSerProSerValAlaAlaAlaSerGlnAlaAspSerLysGlnGlnArgIleLeuLeuAspPheSerLysGlySerAlaSerAsnAla
35                                                                                                                  300
CAGCAGCAGCAGCAGCAGCAGCAGCCGCAGCCAGATTTATCCAAAGCCGTTTCACTGTCCATGGGACTGTATATGGGAGAGACCGAAACAAAAGTGATGGGGAATGACTTGGGCTACCCA
GlnGlnGlnGlnGlnGlnGlnGlnProGlnProAspLeuSerLysAlaValSerLeuSerMetGlyLeuTyrMetGlyGluThrGluThrLysValMetGlyAsnAspLeuGlyTyrPro
75                                                                                                                  420
CAGCAGGGCCAGCTTGGCCTCTCCTCTGGGGAAACAGACTTTCGGCTTCTGGAAGAAAGCATTGCAAACCTCAATAGGTCGACCAGCCGTCCAGAGAATCCCAAGAGTTCAACACCTGCA
GlnGlnGlyGlnLeuGlyLeuSerSerGlyGluThrAspPheArgLeuLeuGluGluSerIleAlaAsnLeuAsnArgSerThrSerArgProGluAsnProLysSerSerThrProAla
115                                                                                                                 540
GCTGGGTGTGCTACCCCGACAGAGAAGGAGTTTCCCCAGACTCACTCTGATCCATCTTCAGAACAGCAAAATAGAAAAAGCCAGCCTGGCACCAACGGTGGCAGTGTGAAATTGTATACC
AlaGlyCysAlaThrProThrGluLysGluPheProGlnThrHisSerAspProSerSerGluGlnGlnAsnArgLysSerGlnProGlyThrAsnGlyGlySerValLysLeuTyrThr
155                                                                                                                 660
ACAGACCAAAGCACCTTTGACATCTTGCAGGATTTGGAGTTTTCTGCCGGGTCCCCAGGTAAAGAGACAAACGAGAGTCCTTGGAGGTCAGACCTGTTGATAGATGAAAACTTGCTTTCT
ThrAspGlnSerThrPheAspIleLeuGlnAspLeuGluPheSerAlaGlySerProGlyLysGluThrAsnGluSerProTrpArgSerAspLeuLeuIleAspGluAsnLeuLeuSer
195                                                                                                                 780
CCTTTGGCGGGAGAAGATGATCCATTCCTTCTGGAAGGGGACGTGAATGAGGATTGCAAGCCTCTTATTTTACCGGACACTAAACCTAAAATTCAGGATACTGGAGATACAATCTTATCA
ProLeuAlaGlyGluAspAspProPheLeuLeuGluGlyAspValAsnGluAspCysLysProLeuIleLeuProAspThrLysProLysIleGlnAspThrGlyAspThrIleLeuSer
235                                                                                                                 900
AGCCCCAGCAGTGTGGCACTGCCCCAAGTGAAAACAGAGAAAGATGATTTCATTGAGCTTTGCACCCCTGGGGTAATTAAGCAAGAGAAACTGGGCCCGGTTTATTGCCAGGCAAGCTTT
SerProSerSerValAlaLeuProGlnValLysThrGluLysAspAspPheIleGluLeuCysThrProGlyValIleLysGlnGluLysLeuGlyProValTyrCysGlnAlaSerPhe
275                                                                                                                1020
TCTGGGACAAATATAATTGGGAATAAAATGTCTGCCATTTCTGTTCATGGCGTGAGTACCTCTGGAGGACAGATGTACCACTATGACATGAATACAGCATCCCTTTCTCAGCAGCAGGAT
SerGlyThrAsnIleIleGlyAsnLysMetSerAlaIleSerValHisGlyValSerThrSerGlyGlyGlnMetTyrHisTyrAspMetAsnThrAlaSerLeuSerGlnGlnGlnAsp
315                                                                                                                1140
CAGAAGCCTGTTTTTAATGTCATTCCACCAATTCCTGTTGGTTCTGAAAACTGGAATAGGTGCCAAGGGTCTGGAGAGGACAACCTGACTTCCTTGGGGGCTATGAACTTCGCAGGCCGC
GlnLysProValPheAsnValIleProProIleProValGlySerGluAsnTrpAsnArgCysGlnGlySerGlyGluAspAsnLeuThrSerLeuGlyAlaMetAsnPheAlaGlyArg
355                                                                                                                1260
TCAGTGTTTTCTAATGGATATTCAAGCCCTGGAATGAGACCAGATGTGAGTTCTCCTCCGTCCAGCTCCTCCACAGCAACGGGACCACCTCCCAAACTCTGCCTGGTGTGCTCCGATGAA
SerValPheSerAsnGlyTyrSerSerProGlyMetArgProAspValSerSerProProSerSerSerSerThrAlaThrGlyProProProLysLeuCysLeuValCysSerAspGlu
395                                                                                                                1380
GCTTCGGTATGCCATTATGGGGTGCTGACGTGTGGAAGCTGTAAAGTCTTCTTTAAAAGAGCAGTGGAAGGACAGCACAATTACCTTTGTGCTGGAAGAAATGATTGCATCATTGATAAA
AlaSerValCysHisTyrGlyValLeuThrCysGlySerCysLysValPhePheLysArgAlaValGluGlyGlnHisAsnTyrLeuCysAlaGlyArgAsnAspCysIleIleAspLys
435                                                                                                                1500
ATTCGAAGAAAAAACTGTCCAGCATGCCGCTATCGAAAATGTCTTCAAGCTGGAATGAACCTGGAAGCTCGAAAAACGAAGAAAAAAATTAAAGGAATTCAGCAAGCCACTGCAGGAGTC
IleArgArgLysAsnCysProAlaCysArgTyrArgLysCysLeuGlnAlaGlyMetAsnLeuGluAlaArgLysThrLysLysLysIleLysGlyIleGlnGlnAlaThrAlaGlyVal
475                                                                                                                1620
TCACAAGACACTTCTGAAAACGCTAACAAAACAATAGTTCCTGCCGCGCTGCCACAGCTTACCCCTACCCTGGTGTCACTGCTGGAGGTGATCGAGCCTGAGGTGTTATATGCAGGATAT
SerGlnAspThrSerGluAsnAlaAsnLysThrIleValProAlaAlaLeuProGlnLeuThrProThrLeuValSerLeuLeuGluValIleGluProGluValLeuTyrAlaGlyTyr
515                                                                                                                1740
GACAGCTCTGTTCCAGACTCAGCATGGAGAATTATGACCACGCTCAACATGTTAGGTGGGCGCCAAGTGATTGCCGCAGTGAAATGGGCAAAGGCGATACCAGGATTCAGAAACTTACAC
AspSerSerValProAspSerAlaTrpArgIleMetThrThrLeuAsnMetLeuGlyGlyArgGlnValIleAlaAlaValLysTrpAlaLysAlaIleProGlyPheArgAsnLeuHis
555                                                                                                                1860
CTGGATGACCAAATGACCCTTCTACAGTACTCATGGATGTTTCTCATGGCATTTGCCCTGGGTTGGAGATCATACAGACAAGCAAGTGGAAACCTGCTATGCTTTGCTCCTGATCTGATT
LeuAspAspGlnMetThrLeuLeuGlnTyrSerTrpMetPheLeuMetAlaPheAlaLeuGlyTrpArgSerTyrArgGlnAlaSerGlyAsnLeuLeuCysPheAlaProAspLeuIle
595                                                                                                                1980
ATTAATGAGCAGAGAATGACTCTACCCTGCATGTATGACCAATGTAAACACATGCTGTTTATCTCCACTGAATTACAAAGATTGCAGGTATCCTATGAAGAGTATCTCTGTATGAAAACC
IleAsnGluGlnArgMetThrLeuProCysMetTyrAspGlnCysLysHisMetLeuPheIleSerThrGluLeuGlnArgLeuGlnValSerTyrGluGluTyrLeuCysMetLysThr
635                                                                                                                2100
TTACTGCTTCTCTCCTCAGTTCCTAAGGAAGGTCTGAAGAGCCAAGAGTTATTTGATGAGATTCGAATGACTTATATCAAAGAGCTAGGAAAAGCCATTGTCAAAAGGGAAGGAAACTCC
LeuLeuLeuLeuSerSerValProLysGluGlyLeuLysSerGlnGluLeuPheAspGluIleArgMetThrTyrIleLysGluLeuGlyLysAlaIleValLysArgGluGlyAsnSer
675                                                                                                                2220
AGTCAGAATTGGCAGCGGTTTTATCAACTGACAAAACTTTTGGACTCCATGCATGATGTGGTTGAAAATCTCCTTAGCTACTGCTTCCAAACATTTTTGGATAAGTCCATGAGTATTGAA
SerGlnAsnTrpGlnArgPheTyrGlnLeuThrLysLeuLeuAspSerMetHisAspValValGluAsnLeuLeuSerTyrCysPheGlnThrPheLeuAspLysSerMetSerIleGlu
715                                                                                                                2340
TTCCCAGAGATGTTAGCTGAAATCATCACTAATCAGATACCAAAATACTCAAATGGAAATATCAAAAAGCTTCTGTTTCATCAGAAATGACTGCCTTACTAAGAAAGGCTGCCTTAAAGA
PheProGluMetLeuAlaGluIleIleThrAsnGlnIleProLysTyrSerAsnGlyAsnIleLysLysLeuLeuPheHisGlnLys***.........***............***...
755                                                                                                          783
                                                                                                                   2460
AAGTTGAATTTATAGCTTTTACTGTACAAACTTATCAACTTGTCTTGTAGATGTTTTGTCGTTCTTTTTGTTTGTCTTGTTTGTTTTCTATACGCACTACATGTGGTCTCTAGAGGGCCA
............***

AGACTTGGCAACAGAAGCAGATGAGCCATCACTTTTCAGTGACAGGAAAGCAGAC
```

Fig. 1. Nucleotide sequence of the cDNA encoding mouse GR and the deduced amino acid sequence of the protein. The nucleotides are numbered above the sequence in the 5′ to 3′ direction. The amino acids are numbered on the left hand side of the sequence starting with the initiator methionine. In-frame stop codons are indicated by ***. Met 107 is the first methionine in pΔrec106.

et al., 1985; Rosenberg *et al.*, 1986), one might anticipate that it plays an important role in DNA binding. Supportive evidence for this notion has recently been provided by Miesfeld *et al.* (1985).

We have isolated a full-length mouse GR cDNA clone which, when expressed in mammalian cells, produces a fully functional GR, as evidenced by its ability to induce expression from the MMTV promoter in cells that are ostensibly devoid of GR. In

addition, we have isolated cDNA clones that represent both the hormone binding deficient and DNA binding deficient forms of the GR present in one S49 nt$^-$ cell line. The lesions have been mapped by DNA sequencing as well as functional tests and they indicate directly that amino acids within the Cys/Lys/Arg-rich region are critical for DNA binding. Furthermore, our analysis of a hormone binding mutant indicates that the hormone binding domain may encompass the entire 25-kd C terminus of the GR.

Results

Cloning the mouse GR

An Okayama and Berg cDNA expression library (Okayama and Berg, 1983) was constructed using polyadenylated RNA from the mouse lymphoma-derived cell line W7.2. This library was screened for mouse GR cDNA clones using as a probe the clone pSG1 which contains 2.2 kb of 3′ non-coding sequence from a rat GR cDNA (Miesfeld *et al.*, 1984). Clone pW28 obtained from this library was 4.2 kb in length and contained the largest open reading frame (ORF) of 2.0 kb. Previous studies have shown the mouse GR to be encoded on a mRNA of ~6.5−7 kb (Miesfeld *et al.*, 1984; Northrop *et al.*, 1986). An overlapping cDNA clone which contained the remainder of the ORF was obtained from a randomly primed cDNA library constructed from W7.2 mRNA and cloned into λgt11 (Young and Davis, 1983). These clones together represnt 4.55 kb of sequence and contain an ORF of 2349 bp which could encode a 783 amino acid polypeptide (Figure 1); an in-frame termination codon is located four codons upstream of the putative initiation codon.

The ORF is flanked on the 3′ side by a large untranslated sequence. Five of 13 clones we have isolated have a 3.85 kb of flanking sequence while the rest have 2.2 kb of 3′-untranslated sequence. The difference in size between these cDNAs correlates well with the size difference in the GR mRNAs one detects in W7.2 cells (5 and 7 kb) (Northrop *et al.*, 1986) and indicates that there are two alternative poly(A) addition sites which can be used. A long 3′-untranslated region and the use of at least two poly(A) addition sites are also features of GR cDNAs from rat (Miesfeld *et al.*, 1985) and human cells (Hollenberg *et al.*, 1985).

Functional expression of the cloned mouse GR

An SV40 expression vector was used to express the cloned GR in mammalian cells (Figure 2, top). pSV2Wrec is a pSV2dhrf (Subramani *et al.*, 1981; Lee *et al.*, 1981) derivative in which the dhfr sequence has been replaced with the 2349 bp of GR coding region, 18 bp of 5′-flanking sequence and ~500 bp of 3′-flanking sequence. Transcription across the GR insert is under the control of the SV40 early promoter and terminates downstream of the SV40 splice site at the SV40 poly(A) addition site. The plasmid contains both pBR and SV40 origins of replication.

MMTV−CAT is a hybrid gene in which the glucocorticoid-inducible promoter of MMTV (Lee *et al.*, 1981) is coupled to the coding region of the bacterial gene chloramphenicol acetyl transferase (CAT) (Alton and Vapnek, 1979). When introduced into cells containing functional GR, the level of CAT expression is under the control of glucocorticoids. Transfection of this plasmid into COS-7 cells (Gluzman, 1981), however, results only in a low basal expression of CAT which cannot be induced by dexamethasone (dex) (Figure 2, bottom) because these cells lack significant quantities of endogenous GR (unpublished observations). Based on these observations we devised an assay to test for the ability of cloned cDNA to produce fully functional GR. COS-7 cells were co-transfected with MMTV−CAT and pSV2-

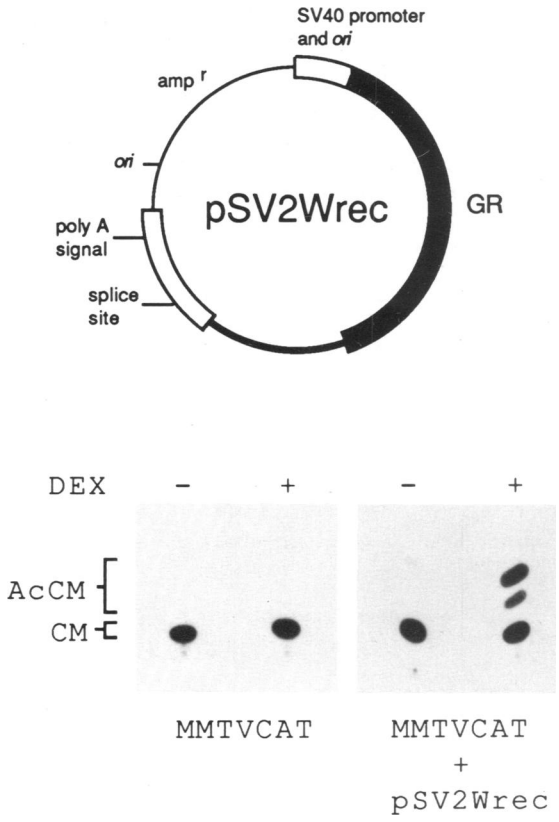

Fig. 2. Functional expression of cloned GR. **(Top)** Structure of the GR expression plasmid pSV2Wrec. The filled-in segment and the thick black line represent coding and 3′ non-coding GR sequences, respectively. The unshaded areas are SV40 sequences and the thin line pBR322-derived sequences. Transcription starts at the SV40 promoter and proceeds clockwise through the GR sequences and terminates at the SV40 poly(A) addition site. **(Bottom)** Expression of CAT in transfected cells. COS-7 cells were transfected with the indicated plasmids and CAT activities were measured in extracts from uninduced (−dex) and induced (+dex) cells. The autoradiogram shows the conversion of [^{14}C]chloramphenicol (CM) to its acetylated products (AcCM).

Wrec and after 2 days the cells were re-fed with medium plus or minus 10^{-6} M dexamethasone. Cells were harvested after 1 day of hormone treatment. CAT activity is induced 14- to 36-fold when pSV2Wrec and MMTV−CAT are co-transfected (Figure 2, bottom), thereby indicating that the mouse GR is indeed capable of producing a fully functional receptor that mediates hormone-regulated transcription from the MMTV long terminal repeat (LTR).

Sequence analysis

The 2349-bp ORF of the mouse GR encodes a polypeptide with a predicted mol. wt of 86 061 which is slightly lower than the 90−94 kd measured using SDS−PAGE. Whether this difference is due to aberrant mobility or whether it represents post-translational modification is not known. The C-terminal half of the molecule is hydrophobic while the N-terminal half of the molecule is hydrophilic. Near the center of the protein is a highly basic region (17/57 Lys or Arg) which overlaps a cysteine-rich region (10/61) (Figure 3, top and bottom). The highly basic nature of this region may indicate that it is the DNA binding domain of the receptor. A homologous Cys,Lys,Arg-rich is also found in the human (Hollenberg *et al.*, 1985) and rat (R.Miesfeld, S.Rusconi and K.R.Yamamoto, personal communication) GRs, in the estrogen receptors from chick (Krust *et al.*, 1986) and human

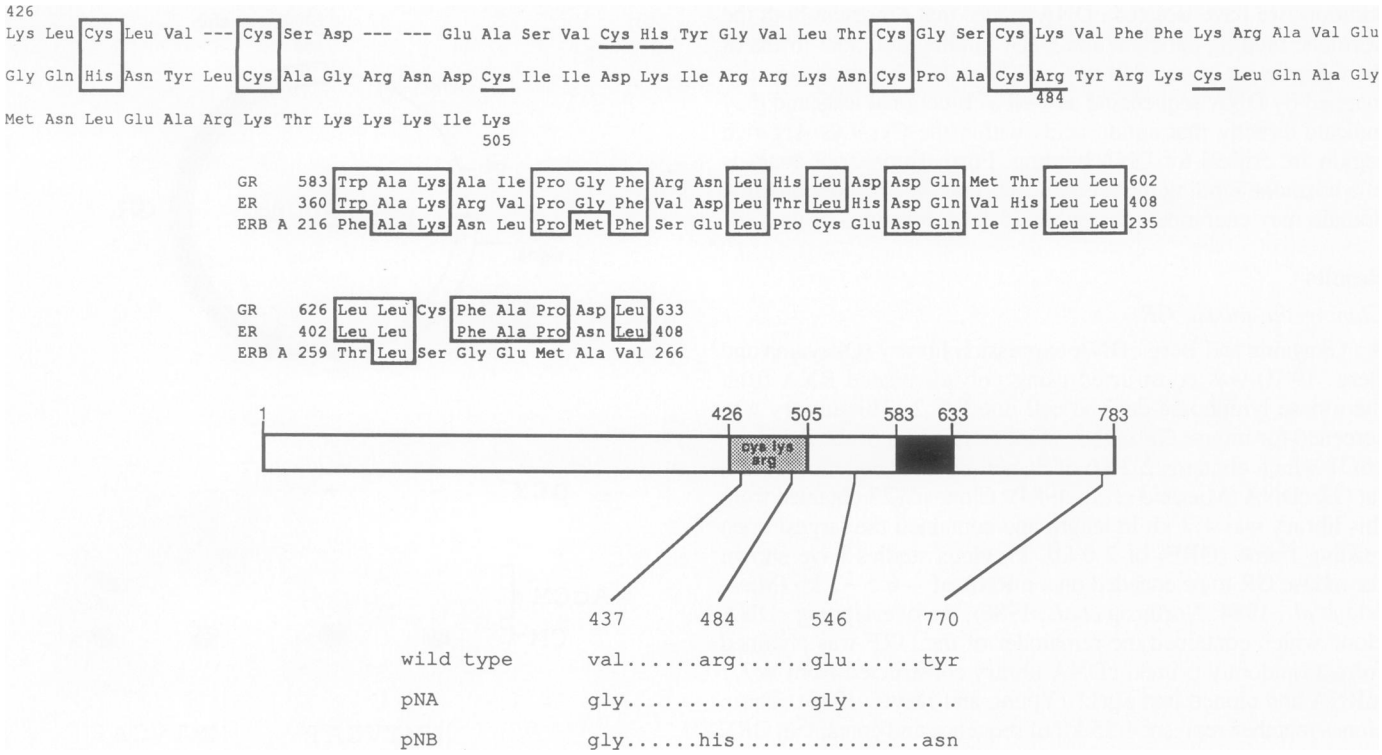

Fig. 3. Sequence analysis of the GR protein. **(Top)** The Cys,Lys,Arg-rich region of the GR. The boxed cysteines have been aligned to form two repeats, with the insertion of three spaces, which show some similarity to the Cys and His repeats found in TFIIIA (see text). The other Cys and His residues are underlined as is Arg-484, which is a His in pSV2NBrec. **(Center)** Homology between mouse GR, human estrogen receptor and v-*erbA* in part of the hormone binding domain of the receptor. Identical residues have been boxed. **(Bottom)** Diagrammatic representation of the GR with the N terminus on the left. The amino acids shown are those that are changed in either pNArec or pNBrec. The positions of the Cys,Lys,Arg-rich region (shaded) and the estrogen homology region (black) are shown.

(Greene *et al.*, 1986), and in the v-*erbA* oncogene product of the avian erythroblastosis virus (AEV) (Weinberger *et al.*, 1985). The cysteine residues in this region are highly conserved in all these proteins and form two repeats of the general formula [Cys (or His)-X_{2-3}-Cys-X_{13} or $_{15}$-Cys-X_2-Cys]. This motif is highly evocative of the arrangement of cysteines and histidines in the *Xenopus* transcription factor TFIIIA (Brown *et al.*, 1985; Miller *et al.*, 1985) which have the general formula [Cys-X_4-Cys-X_{12}-His-X_3-His]. These repeated residues are thought to bind zinc cations causing the protein to form DNA binding 'fingers'. The residues that might be involved in the formation of such zinc binding domains are indicated in Figure 3, top.

Mouse and human GRs are 89% homolgous at the amino acid level. The amino acid differences are not, however, spread evenly throughout the protein but cluster in the N-terminal half of the protein (Figure 4). This is even more clearly seen when one considers only non-conservative amino acid changes. Only four non-conservative changes are present in the carboxyl half (~43 kd) of the receptor, one of which is a valine to glycine change at position 437 which appears to be polymorphic even among mouse cell lines. The most dramatic difference between these receptors is the presence of seven glutamines and a proline found from position 76 to 83 that are absent in the human receptor; a similar, though longer, stretch of glutamines is found in the rat GR (R.Miesfeld, S.Rusconi and K.R.Yamamoto, personal communication). The highly conserved nature of the C-terminal 25–30 kd of the mouse and human GRs does not extend to the estrogen receptor (ER); however, the GR and ER share limited homology to v-*erbA* in this region. There is, however, one small region of strong homology, found in all of these proteins, within the C-terminal domain that lies between positions 583 and 602 of the mouse GR and the corresponding positions in the other proteins (Figure 3, center and bottom). In addition, the region between positions 626 and 633 of the mouse GR is quite homologous to positions 402–408 of the ER. The possible significance of these regions of homology will be discussed below.

Cloning receptors from the nt⁻ variant

The mouse lymphoma-derived cell lines S49 and WEHI7 undergo cytolysis in response to glucocorticoids. A number of groups have isolated hormone-resistant variants from these two cell lines, all of which have defects either in receptor function or in the amount of receptor in the cell (for reviews, see Yamomoto *et al.*, 1976; Huet-Minkowski *et al.*, 1982). One class of S49-derived variant, nt⁻, is especially interesting as it contains two different mutant receptor proteins. One of these receptor proteins does not bind hormone and has only recently been detected using monoclonal antibodies specific for the GR (Northrop *et al.*, 1985; Westphal *et al.*, 1984). The other receptor protein binds hormone but the hormone receptor complex has reduced affinity for DNA (Gehring and Tomkins, 1974; Yamamoto *et al.*, 1974). To study these mutant proteins in more detail we constructed a cDNA library using polyadenylated RNA from the nt⁻ clone S49.1.A.22R (Sibley and Tomkins, 1974a). Sequencing of two independent clones obtained from this library revealed that a single base change from the wild-type sequence created a *Hae*III restriction site ~700 bp upstream from the 3′ end of the coding region (see below). Hybridization to restriction digest blots revealed that the cDNA library contained a second class of clone lacking this *Hae*III site. This property was used to isolate a clone of this second class. The restriction maps of the two classes were ident-

```
Mouse MDSKESLAPPGRDEVPSSLLGRGRGSVMDLYKTLRGGATVKVSASSPSVAAASQADSKQQRILLDFSKGSASNAQQQQQQ   80
Human ·····T·-··E·N···V·AQF··D···F·····L·V···S····R·L·V··P···V·····----   75

Mouse QQPQPDLSKAVSLSMGLYMGETETKVMGNDLGYPQQGQLGLSSGETDFRLLEESIANLNRSTSRPENPKSSTPAAGCATP  160
Human ----·····F·····IS·····LK·····V·····AST·VS·A·  151

Mouse TEKEFPQTHSDPSSEQQNRKSQPGTNGGSVKLYTTDQSTFDILQDLEFSAGSPGKETNESPWRSDLLIDEN-LLSPLAGE  239
Human ·····K····V·····HL·G·T·····N·····S·····C·····  231

Mouse DDPFLLEGDVNEDCKPLILPDTKPKIQDTGDTILSSPSSVALPQVKTEKDDFIELCTPGVIKQEKLGPVYCQASFSGTNI  319
Human ··S·····NS·····K·N··LV·····N·T·····E·····T··S····P·A··  311

Mouse IGNKMSAISVHGVSTSGGQMYHYDMNTASLSQQQDQKPVFNVIPPIPVGSENWNRCQGSGEDNLTSLGAMNFAGRSVFSN  399
Human ·····I·····D·····TL··P··T····  391

                                                    →
Mouse GYSSPGMRPDVSSPPSSSSTAT-GPPPKLCLVCSDEASVCHYGVLTCGSCKVFFKRAVEGQHNYLCAGRNDCIIDKIRRK  478
Human ·····S·····T·····G·····  471

Mouse NCPACRYRKCLQAGMNLEARKTKKKIKGIQQATAGVSQDTSENA-NKTIVPAALPQLTPTLVSLLEVIEPEVLYAGYDSS  557
Human ·····T····E····PG·····T·····L·····  551

Mouse VPDSAWRIMTTLNMLGGRQVIAAVKWAKAIPGFRNLHLDDQMTLLQYSWMFLMAFALGWRSYRQASGNLLCFAPDLIINE  637
Human ····T·····S·A·····  631

Mouse QRMTLPCMYDQCKHMLFISTELQRLQVSYEEYLCMKTLLLLSSVPKEGLKSQELFDEIRMTYIKELGKAIVKREGNSSQN  717
Human ·····YV·S··H·····D·····  711

Mouse WQRFYQLTKLLDSMHDVVENLLSYCFQTFLDKSMSIEFPEMLAEIITNQIPKYSNGNIKKLLFHQK  783
Human ·····E·····N·····T·····  777
```

Fig. 4. Sequence comparison of mouse and human GRs. Only those residues that differ from moue GR sequence are shown. For the human sequence, identical residues are indicated by dots. Nine spaces have been inserted into the human sequence and three spaces into the mouse sequence to preserve the best alignment. Spaces are represented by dashes. Amino acid differences between the two proteins which are of a non-conservative nature are underlined. The arrow indicates the start of the C-terminal half of the protein which contains both the DNA binding and hormone domains.

ical except for the presence (pNArec) or absence (pNBrec) of this *Hae*III site. As none of these nt^--derived clones were full length, fragments containing the carboxy-terminal halves from each class of clone were substituted for the corresponding wild-type fragments in the pSV2 expression plasmid as described in Materials and methods. The resulting constructions either contain (pSV2NArec) or do not contain (pSV2NBrec) the *Hae*III site.

Characterization of a receptor protein lacking hormone binding

To assess whether the cDNAs from S49 nt^- cells indeed contain mutations that alter receptor function, we co-transfected pSV2-NArec and MMTV−CAT into COS-7 cells and then induced the transfected cells with 10^{-6} M dexamethasone as described above. This cDNA from nt^- cells did not support a dexamethasone response whereas the wild-type cDNA did (Figure 5, top). These data mapped the lesion to within the 2.1-kb *Sal*I−*Xba*I fragment used to construct pSV2NArec. Complete sequence analysis of this 2.1-kb fragment revealed only two single base changes from the wild-type sequence. There is a T to G transversion at residue 1327 which changes a valine to a glycine at position 437 of the protein. The second difference is an A to G transition at residue 1654 creating the *Hae*III site discussed above. This change results in the replacement of glutamic acid by glycine at position 546. Each of these mutations was introduced separately into the wild-type sequence by swapping restriction fragments and tested in the COS-7 cell assay system. The mutant receptor containing the valine to glycine substitution (pSV2NA5′rec) suports a dexamethasone-inducible response; however, substituting a glycine at position 546 for glutamic acid (pSV2NA3′rec) results in the production of a non-functional receptor (Figure 5, top).

As previously described, S49.14.22R (nt^-) cells contain two mutant receptors, one binds hormone but not DNA while the other fails to bind hormone. To determine whether pSV2NArec encodes the hormone binding or DNA binding mutant, a dexamethasone binding assay was performed on pSV2NA3′rec-transfected COS-7 cells (Figure 5, center). The results clearly indicate that the glutamic acid to glycine change at position 546 eliminates the receptor's hormone binding activity. As expected, the valine to glycine change (pSV2NA5′rec) has no effect on the receptor's ability to bind hormone. To verify that receptor message was being produced in the COS-7 cells, RNA blot analysis of the transfected cells was performed (Figure 6). The results show that in each case an equal amount of receptor message was produced from the transfected plasmids. Pre-treatment of RNA samples with RNase before analysis completely eliminated the signals (data not shown) indicating that contaminating plasmid was not the source of hybridization. Although the COS-7 cells produce large quantities of receptor message from the introduced cDNA, we can detect only very small amounts of receptor protein immunologically (not shown). Whether this is due to poor translation efficiency or rapid receptor degradation is not known; however, as we have shown, the receptor that is produced is sufficient to mediate hormonal induction of CAT activity from the MMTV LTR.

Characterization of a receptor protein lacking DNA binding (nt^-)

As with pSV2NArec, the expression vector pSV2NBrec did not produce functional receptor protein in transfected COS-7 cells (Figure 5, top), although recombinant GR mRNA (Figure 6) was detected in these cells at levels similar to pSV2Wrec-transfected cells. Sequence analysis revealed three single-base substitutions in the pNBrec DNA that give rise to three amino acid changes relative to W7.2 wild-type GR. The substitution of glycine for valine at position 437 found in pSV2NArec is also present in pNBrec. In addition arginine-484 is changed to histidine and

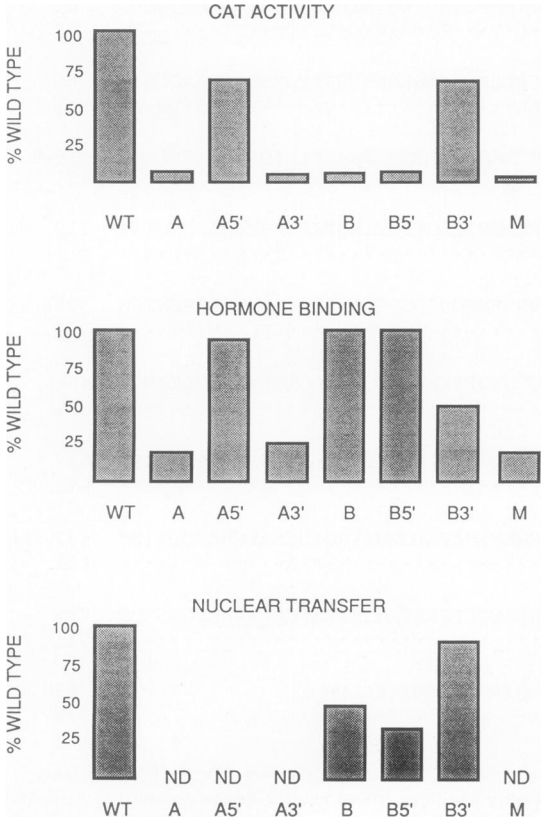

Fig. 5. Properties of the wild-type and mutant forms of the GR. **(Top)** CAT activity in transfected COS-7 cells. COS-7 cells were transfected with either MMTV−CAT alone (M) or with MMTV−CAT and the indicated plasmid and CAT activities were measured in extracts from uninduced (−dex) and induced (+dex) cells. The graph shows the ratio of induced to uninduced CAT activity expressed as a percentage of the wild-type value. WT, pSV2Wrec; A, pSV2NArec; A5′, pSV2NA5′rec; A3′, pSV2NA3′rec; B, pSV2NBrec; B5′, pSV2NB5′rec; B3′, pSV2NB3′rec. **(Center)** Hormone binding in transfected COS-7 cells. COS-7 cells were transfected with the indicated plasmid and a whole cell hormone binding assay performed. The graph shows the binding activity expressed as a percentage of that obtained with wild-type receptor. **(Bottom)** Association of hormone−receptor complexes with the nucleus. COS-7 cells were transfected with the indicated plasmid and a whole cell binding assay performed. Association of receptor−hormone complexes with the nucleus was determined as described in Materials and methods. The values shown are expressed relative to the percentage of [^{3}H]dex−receptor complexes found in the nucleus of COS-7 cells transfected with wild-type receptor. (ND) not determined.

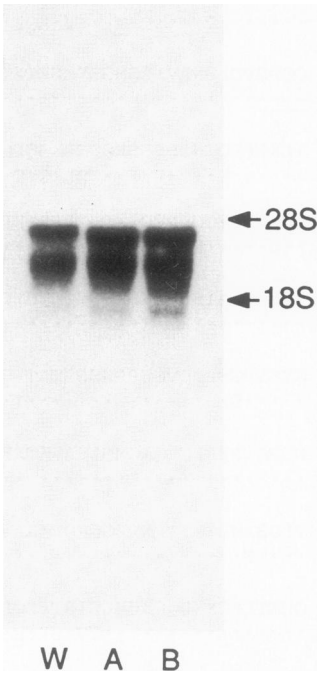

Fig. 6. Analysis of GR RNA produced in transfected COS-7 cells. COS-7 cells were transfected with the indicated plasmid and RNA extracted from the cells 2 days later. 5 μg of RNA was run per lane, transferred to nitrocellulose and probed with a GR-specific probe. Lanes are **W**, pSV2Wrec; **A**, pSV2NA3′rec; **B**, pSV2NBrec. The positions of 28S and 18S rRNA bands are shown.

tyrosine-770 is changed to asparagine. We have shown with pSV2NA5′rec that the valine to glycine change (not present in pSV2NBrec) does not affect receptor function. We therefore separately introduced each of the two amino acid substitutions found in pSV2NBrec into the wild-type sequence and assayed sequence function in COS-7 cells.

Replacement of arginine-484 with histidine (pSV2NB5′rec) results in a functionally inactive receptor protein. As one would predict, however, this variant receptor is able to bind hormone (Figure 5, center). In addition, this receptor has reduced ability to translocate to the nucleus upon hormone treatment of transfected COS-7 cells (Figure 5, bottom). This single amino acid change at position 484 therefore leads to a receptor phenotype similar to that originally described for nt$^-$ cells (Sibley and Tomkins, 1974a) and presumably accounts for its reduced DNA binding ability. Replacement of tyrosine-770 with asparagine (pSV2NB3′rec) did not significantly affect the ability of the protein to mediate hormone induction of CAT activity (Figure 5, top), however Spindler-Barth and Gehring (1982) have reported that the nt$^-$ (i.e. pNBrec) harbors a second defect that leads to reduced hormone binding affinity. Consistent with this interpretation, transfection of COS cells with the Asn-770 mutant leads to production of a receptor that requires a 3- to 4-fold higher concentration of dexamethasone, as compared with wild-type, to elicit half-maximal induction of MMTV-CAT (observations of the authors).

An N-terminal truncated receptor is functional

We have taken advantage of the fact that all of the cDNA clones generated in the pCD vector terminated prior to reaching the 5′ end of the GR mRNA to construct an expression vector (pΔrec106; Figure 7, top) missing the first 295 bp of coding sequence (see Materials and methods). As seen in Figure 1, the first in-frame methionine that might be used as a translation initiation site in this truncated GR cDNA occurs at position 107; the next methionine is not found until position 324. Co-transfection of this plasmid into COS-7 cells with MMTV−CAT leads to the production of a fully functional receptor capable of eliciting dexamethasone-inducible CAT activity comparable with that seen with wild-type receptor (Figure 7, bottom). Immunoblotting experiments indicate that a receptor protein with greater mobility in SDS gels is produced in pΔrec106 transfected COS-7 cells (data not shown). Thus, it appears that this highly variable region of the receptor is not essential for function. Apparent in the CAT assay of uninduced COS-7 cells co-transfected with MMTV−CAT and pΔrec106 (Figure 7, bottom), as well as with MMTV−CAT and pSV2Wrec (Figure 2, bottom) is a high basal level of CAT activity relative to cells transfected with MMTV−CAT alone. This consistent phenomenon in our assays could be due to replication of the MMTV−CAT plasmid in the presence of

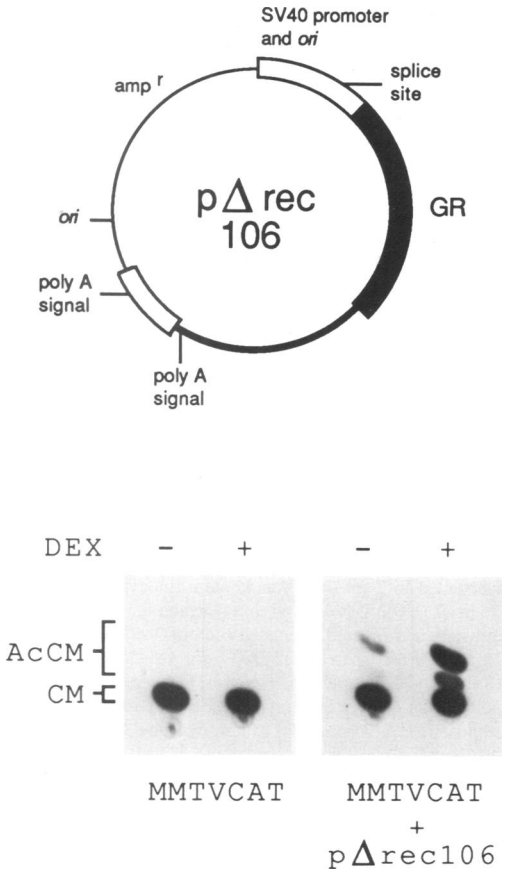

Fig. 7. Expression of a truncated GR. **(Top)** Structure of the truncated GR expression plasmid pΔrec106. The filled-in segment and the thick black line represent coding and 3′ non-coding GR sequences, respectively. The unshaded areas are SV40 sequences and the thin line pBR322-derived sequences. The GR insert lacks the first 295 bp of the coding region. The first Met is at position 107. **(Bottom)** Expression of CAT in transfected cells. COS-7 cells were transfected with the indicated plasmids and CAT activities were measured in extracts from uninduced (−dex) and induced (+dex) cells. The autoradiogram shows the conversion of [^{14}C]chloramphenicol (CM) to its acetylated products (AcCM).

the SV2 receptor plasmids through recombination. Alternatively, this CAT activity could result from receptor mediated induction of the MMTV promoter through small amounts of endogenous glucocorticoids in the culture medium. This second possibility is unlikely as the high basal CAT activity is seen with SV2 receptor plasmids which code for non-functional receptor proteins. Co-transfection of COS-7 cells with MMTV−CAT and an SV2 plasmid not containing receptor cDNA may help resolve this issue.

Discussion

Biochemical analysis has suggested that the formation of a steroid−receptor complex brings about a conformational change in the receptor which allows the complex to bind to specific DNA sequences. The GR is thought to consist of three domains, a DNA binding, a hormone binding and a specificity conferring or modulatory domain (Carlstedt-Duke *et al.*, 1982). The C-terminal half of the molecule contains both the DNA binding and hormone binding domains (Miesfeld *et al.*, 1985; Northrop *et al.*, 1986). S49 nt^{i} cells produced a 40-kd protein which retains both hormone and DNA binding activities (Yamamoto *et al.*, 1974); we and others have shown that this is due to the production by these cells of a GR mRNA which lacks approximately the first 1000 bases of 5′ coding sequence (Northrop *et al.*, 1986; Miesfeld *et al.*, 1985). A protein of similar size and properties can be produced by mild chymotryptic digestion of the wild-type receptor (Dellweg *et al.*, 1982). The 40-kd fragment from S49 nt^{i} cells or generated by proteolytic digestion of wild-type receptor appears to bind to non-specific DNA more tightly than wild-type receptor (Yamamoto *et al.*, 1974; Gehring and Hotz, 1983). For this reason the N-terminal domain of the receptor has been thought to have, among other functions, the ability to confer DNA binding specificity on the rest of the molecule. Analysis of the receptor sequence suggests an alternative possibility. The full size GR is an acidic protein while a truncated 40-kd C-terminal protein, which contains the DNA binding domain, would be quite basic. It is therefore conceivable that all the information required for hormone binding and specific DNA binding resides within the 40-kd fragment and that its increased binding to non-specific DNA is simply due to its highly basic charge. This could be tested by substituting varying extents of acidic or basic fragments onto the amino terminus of the 40-kd C-terminal domain and testing the resultant proteins' DNA binding properties.

If the N-terminal domain encodes determinants required for function then they do not reside within the first 106 amino acids as deletion of these amino acids in pΔrec106 results in a perfectly functional receptor. Although these amino acids are not essential for receptor activity they may be involved in subtle modulatin of receptor function. A relatively small effect on receptor turnover, for instance, may have an effect *in vivo* but would not be detected in our assays. Contained within these first 106 amino acids is a stretch of nine glutamines encoded by the codon CAG, and two prolines. This transcribed repetitive element, first found in *Drosophila melanogaster*, and called *opa* (Wharton and Yedvobnick, 1985) does not have any apparent function. The stretch of glutamines is part of a much larger region, extending to amino acid 275, which bears the most sequence variation when mouse and human receptor sequences are compared. Furthermore, it is the N-terminal domain that is most variable in length and sequence among steroid receptors. Whereas the GR contains ~425 amino acids upstream of the first *cys* in the DNA binding domain, the ER contains only 180 amino acids. Analysis of the overall size and structure of other steroid receptors would indicate that the androgen receptor (~120 kd) could have as many as 700 and the vitamin D receptor (~55 kd) as few as 100 amino acids upstream of the DNA binding domain (Vedeckis, 1984). Since it seems unlikely that each receptor would interact with different components of the basic transcription machinery it is tempting to speculate that most or all of the transcription regulatory functions of the steroid receptors could lie very near the DNA binding domain. Additional mutation analysis combined with the rapid functional assay described herein should help to resolve this critical issue.

A number of steroid receptors including the human and rat glucocorticoid, human estrogen, and chicken estrogen as well as v-*erbA* have highly homologous regions rich in cysteine, lysine, and arginine. The highly basic nature of this region and its possible involvement in zinc binding in a similar manner to TFIIIA, indicate that it may be a DNA binding domain. We have shown that a single amino acid change (Arg-484 to His) within this region completely abolishes receptor function apparently by altering the nuclear association of this protein. It is possible that arginine-484 interacts directly with DNA in a manner that histid-

ine cannot, or perhaps a basic residue is required at this position. Another more intriguing possibility is that a histidine at position 484 creates an alternative zinc binding site. His-484 might pair with Cys-480 or Cys-488 forming a structure which is unable to interact correctly with the DNA (Figure 3a). This possibility is amenable to experimental testing by substitution of alternative amino acids at position 484.

Tryptic digests of the truncated 40-kd receptor produce a fragment of ~25 kd which can bind hormone but not DNA (Sherman *et al.*, 1983; Wrange *et al.*, 1984). The mapping of the DNA binding domain to the N-terminal region of the truncated 40-kd receptor places the hormone binding domain within the C-terminal 25 kd of the receptor. Furthermore, replacement of the C-terminal 50 amino acids of the human GR with 15 amino acids derived from an alternatively spliced mRNA ablates hormone binding *in vitro* (Hollenberg *et al.*, 1985). Replacement of glutamine-546 with glycine (pSV2NA3′rec) totally inactives hormone binding. These data taken together indicate that most or all 25 kd are required for the formation of a functional hormone binding domain. The binding of hormone to the receptor is thought to bring about a conformational change which unmasks or creates the specific DNA binding characteristics of the receptor. In addition, the nt$^-$ receptor has been shown to bind hormone less well than wild-type receptor due to a 3-fold increase in the K_d for dexamethasone (Spindler-Barth and Gehring, 1982). Preliminary evidence suggests that this may be due to the tyrosine to asparagine change at position 770; this position may therefore define a site which is involved in hormone binding but not in the subsequent conformational change.

The hormone binding domains of the ER and GR are only 30% homologous at the amino acid level. The amino acid domains on the other hand are highly conserved (62%) between the two types of receptor. It is generally assumed that the hormone binding domain has to interact with the DNA binding domain in order to bring about the ligand-induced conformational change that exposes the DNA binding site on the receptor. Such an interaction might be expected to involve the more conserved sequences in the hormone binding domain. There is in fact a small region of striking homology within the C-terminal domain of the ER and GR (Figure 3b) which could be involved in such an interaction. One could envision that these 20–30 amino acids form a loop that physically masks all or part of the DNA binding domain. Hormone binding would cause an unmasking of the DNA binding site by altering the conformation of the loop.

An alternative and perhaps more attractive possibility is that masking/unmasking of the DNA binding site is mediated by another receptor-associated protein. The protein might interact with the conserved region of the hormone binding domain and with the DNA binding domain simultaneously in such a way as to prevent DNA binding. The role for the ligand would be to bring about the dissociation of this protein from receptor thereby exposing the DNA binding domain. There is in fact a candidate for such a receptor-associated protein. Most, if not all, steroid receptors are complexed in the unactivated state to a 90-kd heat shock protein (Catelli *et al.*, 1985). Hormone binding causes a dissociation of receptor from the heat shock protein with simultaneous transformation of the 4S DNA binding form of the receptor (Sanchez *et al.*, 1985). If this small region of homology shared by the GR, ER and v-*erbA* protein is involved in receptor activation by either of the possibilities mentioned above, alterations in the region might be expected to produce a receptor that could not associate with the 90-kd heat shock protein and would thus be transcriptionally active in the absence of inducer.

Materials and methods

Cells and media

W7.2 is a glucocorticoid-sensitive subclone of the mouse T-lymphosarcoma cell line WEHI7 (Harris *et al.*, 1973). S49.1.A.22R (nt$^-$), selected by Sibley and Tomkins (1974a,b), is a glucocorticoid-resistant variant of the mouse lymphoma line S49.1T.B4.1A (Yamamoto *et al.*, 1974). The cell line, COS-7, is a derivative of the simian kidney cell line CV-1 transformed by a mutant of SV40 (Gluzman, 1981). COS-7 cells produce SV40 T antigen which allows plasmids containing an SV40 origin of replication to replicate to high copy numbers. COS-7 cells were grown as monolayers in Dulbecco's modified Eagle's medium (DMEM, Irvine Scientific) supplement with 10% fetal bovine serum (FBS, Irvine Scientific) in a 37°C humidified atmosphere containing 10% CO_2. Lymphoma cells were grown in suspension in RPMI 1640 medium (Irvine Scientific) supplemented with 10% FBS and 1 mM sodium pyruvate (Gibco) in a 5% CO_2 atmosphere at 37°C. Restriction enzymes, T4 DNA ligase, T4 DNA polymerase, *Escherichia coli* DNA polymerase I, and agarose were obtained from Bethesda Research Laboratories. Avian myeloblatosis virus reverse transcriptase was from Seikagaku. Calf thymus terminal deoxynucleotidyl transferase, *E. coli* DNA ligase, and *E. coli* RNase H were from PL Biochemicals. Riboprobe (SP6) polymerase and RNasin were from Promega Biotech. All other routine chemicals were from Sigma or Baker. [α-^{32}P]GTP (600 Ci/mmol) was purchased from Amersham.

Isolation and blotting of RNA

Total cellular RNA was isolated by the method of Chirgwin *et al.* (1979). Briefly, cells were suspended in 2 ml 4 M guanidinium thiocyanate (Kodak), 25 mM sodium citrate pH 7.0, 0.2% N-lauroylsarcosine and 0.1 M β-mercaptoethanol. The resulting lysate was layered over a cushion of 3 ml of 5.7 M CsCl (KBI), 0.1 M EDTA pH 7.0 and spun at 115 000 *g* for 18–20 h. The pelleted RNA was dissolved in 10 mM Tris pH 7.5, 1 mM EDTA and its concentration determined by reading absorbance at 260 nm. Polyadenylated RNA was selected using oligo(dT)-cellulose (PL Biochemicals) as described by Aviv and Leder (1972). RNA was separated on agarose gels containing 2.2 M formaldehyde (Lehrach *et al.*, 1977) in a buffer of 20 mM MOPS (3-[N-morpholino]propanesulfonic acid) pH 7.0, 5 mM sodium acetate, and 1 mM EDTA. Gels were then treated with 50 mM NaOH, 10 mM NaCl for 30 min, neutralized for 1 h in 0.1 M Tris pH 7.5, and then transferred to nitrocellulose (Schleicher and Schuell) in 20 × SSPE (1 × SSPE is 180 mM NaCl, 10 mM sodium phosphate pH 7.7, and 1 mM EDTA). Blots were pre-hybridized in H-buffer [50% formamide, 5 × SSPE, 0.1% SDS, 2 × Denhardt's solution (1 × Denhardt's solution is 0.02% w/v each of bovine serum albumin, polyvinylporrolidone and Ficoll) 0.5 mg/ml yeast RNA, 0.1 mg/ml denatured and sheared salmon sperm DNA] at 42°C for 16 h. Blots were hybridized in H-buffer containing 1 × 10^6 d.p.m./ml of radioactive probe for 24 h at 60°C. ^{32}P-Labeled probes were made using the Riboprobe system (Melton *et al.*, 1984). Blots were washed at 65°C for 2 × 30 min in 0.2 × SSPE, 0.1% SDS and autoradiographed.

Construction and screening of cDNA libraries

cDNA libraries were constructed in the pCD vector system using 4.5 μg of polyadenylated RNA from W7.2 and from nt$^-$ cells, essentially as described (Okayama and Berg, 1983), except that *E. coli* strain MC 1061 (Casadaban and Cohen, 1980) was used as the host. The resulting libraries contained 4.1 × 10^5 (W7.2) and 3.8 × 10^5 (nt$^-$) independent recombinants. A randomly primed cDNA library in λgt11 (Young and Davis, 1983), constructed using polyadenylated mRNA from W7.2 cells, was kindly provided to us by Dr Tanya Schulz. The λgt11 cDNA library from W7.2 cells contained 1 × 10^7 independent recombinants.

Screening of the λgt11 library was accomplished by a modification of the method of Benton and Davis (1977). *E. coli* strain LE392 (Young and Davis, 1983) was infected with phage and plated at an initial density of 10 000 plaques per 100 mm dish (Falcon). Nitrocellulose filter disks (BA85, 0.45 μM × 82 mm, Schleicher and Schuell) with transferred phage were denatured (0.5 N NaOH, 1.5 M NaCl for 5 min), neutralized [0.5 M Tris (pH 7.4), 3 M NaCl for 5 min] and washed (2 × SSPE for 5 min) before baking. The filters were pre-hybridized as described under blotting of RNA except that hybridizations were done at 42°C and washing at 50°C. Screening of pCD libraries was performed by plating bacteria directly on agar plates containing 50 μg/ml ampicillin at a density of 10 000 per 100 mm dish. After 9 h of growth at 37°C colonies were just visible. Colonies were lifted onto nitrocellulose filter disks which were then transferred onto agar plates containing 170 μg/ml chloramphenicol and incubated at 37°C for 16 h. The filters were then denatured and further processed exactly as described above for the λgt11 library.

Construction of plasmids

The vector pΔrec106 was constructed using two separate pCD clones; one, pN10 isolated from the nt$^-$ pCD library, has been previously described (Northrop *et al.*, 1986). pN10 contains 5.9 kb of cDNA with 3.85 kb of 3′-untranslated sequence. The second, pW28, was isolated from the W7.2 pCD library and contains ~4.2 kb of cDNA with 2.2 kb of 3′-untranslated sequence. Therefore, pN10

contains ~50 bases more coding region than does pW28. To form pΔrec106 a 700-bp *Sal*I fragment containing the SV40 early region promoter and 16S late mRNA splice junction and the 5′ end of the cDNA was removed from pW28 and the analogous 750-bp fragment from pN10 ligated into pW28.

The expression vector pSV2Wrec was constructed from portions of pW28, clone 21 isolated from the W7.2 λgt11 library which contains a 1.52-kb cDNA insert, and pSV2dhfr (Lee *et al.*, 1981). A 975-bp *Eco*RI to *Apa*I fragment from clone 21 was ligated to a 1.9-kb *Apa*I to *Pvu*II fragment from pW28 containing ~500 bp of 3′-untranslated sequences. The resulting fragment containing the whole open reading frame was then inserted into pSV2dhfr in place of the dhfr cDNA. Plasmid pSV2NArec was constructed by replacing the 2.1-kb *Sal*I to *Xba*I cDNA fragment from pSV2Wrec with the same fragment from pN10. pSV2NArec as well as all other expression vectors described below are missing a 200-bp *Xba*I cDNA fragment and therefore have only 300 bp of 3′-untranslated sequence. Vectors pSV2NA5′rec and pSV2NA3′rec were constructed using the unique *Sph*I site which lies between the two amino acid changes in pSV2NArec. A 1.0-kb *Sph*I to *Xba*I fragment was removed from pSV2Wrec and replaced with the same fragment from pSV2NArec to form pSV2NA3′rec. pSV2NA5′rec was constructed by replacing the 500-bp *Apa*I to *Sph*I fragment from pSV2Wrec with the same fragment from pSV2NArec. pSV2NBrec was constructed by removing the 1.0-kb *Sph*I to *Xba*I fragment from pSV2Wrec and replacing it with the same fragment from the nt$^-$ pCD clone pNB1. This vector contains the Arg-484 to His and Tyr-770 to Asn but not the Val-437 to Gly amino acid changes. Switching the 900-bp *Pst*I to *Xba*I fragment from pSV2NBrec to pSV2Wrec was used to create pSV2NB3′rec; while the opposite switch (pSV2Wrec to pSV2NBrec) created pSV2NB5′rec.

The plasmid pMMTVCAT was constructed by inserting a 2.2-kb genomic fragment containing the MMTV LTR and 800 bp of 5′-flanking sequences cloned from the MMTV-infected HTC cell line JZ.2 (Grove and Ringold, 1981) into the unique *Hin*dII site of pSVOCAT (Gorman *et al.*, 1982).

Sequencing was performed using the dideoxy technique (Sanger *et al.*, 1977) using M13mp11 and M13mp18 (reviewed by Yanisch-Perron *et al.*, 1985) as vectors and the M13 15-bp or 17-bp universal primers (PL Biochemicals).

Transfections

Plasmid DNAs were purified by the method of Birnboim and Doly (1979) with subsequent removal of high mol. wt RNA by precipitation with 2.5 M LiCl, and finally banding on CsCl gradients.

COS-7 cells grown on 100 mm dishes were transfected by a modification of the method of Sompayrac and Danna (1981). Nearly confluent dishes of cells were washed twice with phosphate-buffered saline (PBS) and then exposed to 5 μg of receptor containing vector DNA and/or 20 μg of pMMTVCAT in 3 ml DMEM (without serum) containing 400 μg/ml DEAE-dextran (Pharmacia) per plate for 3 h. Plates were then washed twice with PBS and fresh medium (with serum) was added. After 2–3 days of growth, cells were washed with PBS and harvested for either CAT assays, binding assays (see below), or RNA isolation (see above). Inductions of CAT activity were typically performed by adding dexamethasone to 10^{-6} M to the culture medium 16–20 h before harvesting cells. Assays of CAT were carried out essentially as described by Gorman *et al.* (1982). Autoradiographs of silica gel thin-layer plates were used to cut out separated acetylated and non-acetylated chloramphenicol for counting in aqueous counting scintillant (Amersham).

Glucocorticoid binding and nuclear translocation were measured using the [^{3}H]dexamethasone whole cell binding assay (Pfahl *et al.*, 1978b) except that a single saturating concentration of [^{3}H]dexamethasone (10^{-7} M) was used. Nonspecific binding was determined using a 500-fold excess of unlabelled dexamethsone in a parallel incubation.

Acknowledgements

We thank Tanya Schulz for constructing the λgt11 cDNA libraries, Keith Yamamoto for the nt$^-$ cell line, S49.1A.22R, and Roger Miesfeld, Keith Yamamoto and Sandro Rusconi for the clone pSG1 and for communicating results prior to publication. We also thank Karen Benight, Karen Bird and Nancy Schonher for their expert preparation of the manuscript. This work was supported by NIH grant GM25821 to G.M.R., and NIH predoctoral training grant GM07149. M.D. is a Leukemia Society of America Special Fellow. G.M.R. holds an Established Investigator Award from the American Heart Association.

References

Alton,N. and Vapnek,D. (1979) *Nature*, **282**, 864–869.
Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408–1412.
Benton,W.D. and Davis,R.W. (1977) *Science*, **196**, 180–182.
Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513–1523.
Brown,R.S., Sander,C. and Argos,P. (1985) *FEBS Lett.*, **186**, 271–274.
Carlstedt-Duke,J., Okret,S., Wrange,O. and Gustafsson,J.-A. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4260–4264.
Casadaban,M.J. and Cohen,S.N. (1980) *J. Mol. Biol.*, **138**, 179–207.
Catelli,M.G., Binart,N., Jung-Testas,I., Renoir,J.M., Baulieu,E.E., Feramisco, J.R. and Welch,W.J. (1985) *EMBO J.*, **4**, 3131–3135.
Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W. (1979) *Biochemistry*, **18**, 5294–5299.
Dellweg,H.-G., Hotz,A., Mugele,K. and Gehring,U. (1982) *EMBO J.*, **1**, 285–289.
Gehring,U. and Hotz,A. (1983) *Biochemistry*, **22**, 4013–4018.
Gehring,U. and Tomkins,G.M. (1974) *Cell*, **3**, 301–306.
Gluzman,Y. (1981) *Cell*, **23**, 175–182.
Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
Greene,G.L., Gilna,P., Waterfield,M., Baker,A., Hort,Y. and Shine,J. (1986) *Science*, **231**, 1150–1154.
Grove,J.R. and Ringold,G.M. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4349–4353.
Harris,A.W., Bankhurst,A.D., Mason,S. and Warner,N.L. (1973) *J. Immunol.*, **110**, 431–438.
Hollenberg,S.M., Weinberger,C., Ong,E.S., Cerelli,G., Oro,A., Lebo,R., Thompson,E.B., Rosenfeld,M.G. and Evans,R.M. (1984) *Nature*, **318**, 635–641.
Huet-Minkowski,M., Gasson,J.C. and Bourgeois,S. (1982) In Bruchovsky,N. and Goldie,J.H. (eds), *Drug and Hormone Resistance in Neoplasia.* CRC Press, Boca Raton, FL, Vol. 1, pp. 79–94.
Karin,M., Haslinger,A., Holtgreve,H., Cathala,G., Slater,E. and Baxter,J.D. (1984) *Cell*, **36**, 371–379.
Krust,A., Green,S., Argos,P., Kumar,V., Walter,P., Bornert,J.-M. and Chambon,P. (1986) *EMBO J.*, **5**, in press.
Lee,F., Mulligan,R., Berg,P. and Ringold,G. (1981) *Nature*, **294**, 228–232.
Lehrach,H., Diamond,D., Wozney,J.M. and Boedtker,H. (1977) *Biochemistry*, **16**, 4743–4751.
Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green, M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
Miesfeld,R., Okret,S., Wikstrom,A.-C., Wrange,O., Gustafsson,J.-A. and Yamamoto,K.R. (1984) *Nature*, **312**, 779–781.
Miesfeld,R., Rusconi,S. and Yamamoto,K.R. (1985) In Calendar,R. and Gold,L. (eds), *Sequence Specificity in Transcription and Translation.* Alan R. Liss, Inc., NY, Vol. 30, pp. 535–545.
Miller,J., McLachlan,A.D. and Klug,A. (1985) *EMBO J.*, **4**, 1609–1614.
Northrop,J.P., Gametchu,B., Harrison,R.W. and Ringold,G.M. (1985) *J. Biol. Chem.*, **260**, 6398–6403.
Northrop,J.P., Danielsen,M. and Ringold,G.M. (1986) *J. Biol. Chem.*, in press.
Okayama,H. and Berg,P. (1983) *Mol. Cell. Biol.*, **3**, 280–298.
Payvar,F. and Wrange,O. (1984) In Eriksson,H., Gustafsson,J.-A. and Hogberg, B. (eds), *Steroid Hormone Receptors: Structure and Function.* Elsevier/North-Holland Biomedical Press, pp. 267–282.
Payvar,F., DeFranco,D., Firestone,G.L., Edgar,B., Wrange,O., Okret,S., Gustafsson,J.-A. and Yamamoto,K.R. (1983) *Cell*, **35**, 381–392.
Pfahl,M., Kelleher,R.J., Jr. and Bourgeois,S. (1978a) *Mol. Cell. Endocrinol.*, **10**, 193–207.
Pfahl,M., Sandros,T. and Bourgeois,S. (1978b) *Mol. Cell. Endocrinol.*, **10**, 175–191.
Renkawitz,R., Schutz,G., von der Ahe,D. and Beato,M. (1984) *Cell*, **37**, 503–510.
Ringold,G.M. (1985) *Annu. Rev. Pharmacol. Toxicol.*, **25**, 529–566.
Rosenberg,U.B., Schroder,C., Preiss,A., Kienlin,A., Cote,S., Reide,I. and Jackle, H. (1986) *Nature*, **319**, 36–339.
Sanchez,E.R., Toft,D.O., Schlesinger,M.J. and Pratt,W.B. (1985) *J. Biol. Chem.*, **260**, 12398–12401.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Scheidereit,C., Geisse,S., Westphal,H. and Beato,M. (1983) *Nature*, **304**, 749–752.
Sherman,M.R., Moran,M.C., Tuazon,F.B. and Stevens,Y.-W. (1983) *J. Biol. Chem.*, **258**, 10366–10377.
Sibley,C.H. and Tomkins,G.M. (1974a) *Cell*, **2**, 221–227.
Sibley,C.H. and Tomkins,G.M. (1974b) *Cell*, **2**, 213–220.
Sompayrac,L.M. and Danna,K.J. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 7575–7578.
Spindler-Barth,M. and Gehring,U. (1982) *FEBS Lett.*, **138**, 91–94.
Subramani,S., Mulligan,R. and Berg,P. (1981) *Mol. Cell. Biol.*, **1**, 854–864.
Vedeckis,W.V. (1985) In Hollander,V.P. (ed.), *Hormonally Responsive Tumors.* Academic Press, NY, pp. 3–61.
Westphal,H.M., Mugele,K., Beato,M. and Gehring,U. (1984) *EMBO J.*, **3**, 1493–1498.
Weinberger,C., Hollenberg,S.M., Rosenfeld,M.G. and Evans,R.M. (1985) *Nature*, **318**, 670–672.

Wharton,K.A. and Yedvobnick,B. (1985) *Cell*, **40**, 55–62.
Wrange,O., Okret,S., Radojcic,M., Carlstedt-Duke,J. and Gustafsson,J.-A. (1984) *J. Biol. Chem.*, **259**, 4534–4541.
Yamamoto,K.R. and Alberts,B.M. (1976) *Annu. Rev. Biochem.*, **45**, 721–746.
Yamamoto,K.R., Stampfer,M.R. and Tomkins,G.M. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 3901–3905.
Yamamoto,K.R., Gehring,U., Stampfer,M.R. and Sibley,C.H. (1976) *Recent Prog. Horm. Res.*, **32**, 3–32.
Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.
Young,R.A. and Davis,R.W. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1194–1198.

Received on 18 July 1986