Transcriptional control of the glucocorticoid receptor: CpG islands, epigenetics and more.
Jonathan D. Turner, Simone R. Alt, Lei Cao, Sara Vernocchi, Slavena Trifonova, Nadia Battello, Claude P. Muller

To cite this version:
Jonathan D. Turner, Simone R. Alt, Lei Cao, Sara Vernocchi, Slavena Trifonova, et al. Transcriptional control of the glucocorticoid receptor: CpG islands, epigenetics and more. Biochemical Pharmacology, 2010, 80 (12), pp.1860. 10.1016/j.bcp.2010.06.037. hal-00637161

HAL Id: hal-00637161
https://hal.science/hal-00637161
Submitted on 31 Oct 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
GR Transcription

Transcriptional control of the glucocorticoid receptor: CpG islands, epigenetics and more.

Jonathan D. Turner\textsuperscript{1,2}, Simone R. Alt\textsuperscript{1,2}, Lei Cao\textsuperscript{1,2}, Sara Vernocchi\textsuperscript{1,2}, Slavena Trifonova\textsuperscript{1,2}, Nadia Battello\textsuperscript{1}, Claude P. Muller\textsuperscript{1,2}

\textsuperscript{1}Institute of Immunology, CRP-Santé / Laboratoire National de Santé, 20A rue Auguste Lumière, L-1950, Luxembourg

\textsuperscript{2}Department of Immunology, Graduate School of Psychobiology, University of Trier, D-54290, Germany

Running title: GR transcription

Address for correspondence:
Claude P. Muller, Institute of Immunology, CRP-Santé/Laboratoire National de Santé, 20A rue Auguste Lumière, L-1950 Grand Duchy of Luxembourg.
Tel: +352 490604 Fax: +352 490686
Email: claude.muller@LNS.ETAT.LU
Abstract

The unique variability in the 5’ region of the GR gene, with 9 alternative first exons and 13 splice variants plays a critical role in transcriptional control maintaining homeostasis of the glucocorticoid receptor (GR). This 5’mRNA heterogeneity, common to all species investigated, remains untranslated since the alternative first exons are spliced to exon 2 immediately upstream of the translation initiation codon. These alternative first exons are located either immediately upstream of the coding exons in the CpG island (exons B-H and J), or further upstream (exons 1A and 1I). The mechanisms regulating the differential usage of these first exons in different tissues and individuals, and the role of the 5’UTR in the splicing of the coding exons are still poorly understood. Here we review some of the mechanisms that have so far been identified. Data from our laboratory and others have shown that the multiple first exons represent only a first layer of complexity orchestrated probably by tissue-specific transcription factors. Modulation of alternative first exon activity by epigenetic methylation of their promoters represents a second layer of complexity at least partially controlled by perinatal programming. The alternative promoter usage also appears to affect the 3’ splicing generating the different GR coding variants, GRα, GRβ, and GR-P. Aberrant GR levels are associated with stress-related disorders such as depression, and affect social behaviour, mood, learning and memory. Dissecting how tissue-specific GR levels are regulated, in particular in the brain, is a first step to understand the significance of aberrant GR levels in disease and behaviour.

Keywords: Glucocorticoid receptor; Promoter; Methylation; Epigenetic variability; alternative exon splicing
GR Transcription

1 Introduction

Stress causes an increasing disease burden at least in Western societies. Economic losses due to stress may be as high as 3-4% of the European gross national product. In addition, stress has detrimental effects on behaviour, mood, learning and memory that are difficult to quantify. Glucocorticoids (GCs), the key stress mediators, exert profound effects on a wide range of physiological and developmental processes that are crucial for the adaptation to stress. Psychosocial stress has also been implicated in the development of mental disorders, including schizophrenia, anxiety disorders, and depression.

GCs act via binding to two types of intracellular nuclear receptors, i.e. the glucocorticoid receptor (GR) and the mineralocorticoid receptor (MR). Upon binding of GCs, these receptors translocate to the nucleus where they regulate the activity of specific target genes, in a cell-type specific manner, as transcription factors. This review focuses on the classical GR, NR3C1. Numerous factors have been demonstrated to affect the responsiveness to GC by regulating GR activity, such as GR co-activators and co-repressors [1], GR splice-variants [2-4], and GR isoforms [5-7]. In addition, and perhaps most important for GC responsiveness is the expression level of GR protein [8-11], determined by the mRNA level.

In eukaryotic cells, gene expression is controlled by a variety of mechanisms, both at transcriptional and translational levels, including chromatin condensation, transcription initiation, DNA methylation, alternative RNA splicing, mRNA stability and others. The GR is an ubiquitously expressed nuclear hormone receptor, however, levels of both mRNA and protein vary widely between cell and tissue types. Over the last few years we and others have contributed to the significant progress that has been made to unravel the
GR Transcription

transcriptional mechanisms determining the tissue specific control of GR levels, that will be reviewed here.

Structure of the NR3C1 gene

The human GR gene (OMIM + 138040; NR3C1) is located on chromosome 5q31-q32 [12] and contains 8 translated exons (2-9) and 9 untranslated alternative first exons. We and others have shown that GR levels are under the transcriptional control of a complex 5’ structure of the gene, containing the untranslated first exons important for differential expression of the GR. All of the alternative first exons identified are located in one of the two promoter regions: the proximal or the distal promoter region, located approximately 5kb and 30kb upstream of the translation start site, respectively [13-18]. Alternative first exons 1A, and 1I are under the control of promoters in the distal promoter region, whereas the promoters of exons 1D, 1J, 1E, 1B, 1F, 1C (1C1-3), 1H (Fig. 1A) are located in the proximal promoter region [19, 20]. Exons 1D to 1H are found in an upstream CpG island with a high sequence homology between rats and humans.

The region- or tissue-specific usage of alternative first exons leading to different GR mRNA transcripts [19, 20] (Fig. 1 B) provides a mechanism for the local fine-tuning of GR levels. Since the ATG start codon lies only in the common exon 2, this 5’mRNA heterogeneity remains untranslated, but is important for translational regulation [21].

Alternative mRNA transcript variants are generated by splicing of these alternative first exons to a common acceptor site in the second exon of the GR. Exon 2 contains an in-frame stop codon immediately upstream of the ATG start codon to ensure that this 5’
GR Transcription

heterogeneity remains untranslated, and that the sequence and structure of the GR is not affected.

The GR also has a variable 3’ region. Unlike the 5’ region, the 3’ variability encodes splice variants with different functions. The 3 main 3’ splice variants of the GR are GRα, GRβ, and GR-P (Fig. 1B). GRα and GRβ are generated by two alternatively spliced 3’ exons, 9α and 9β. GR-P lacks both exons 8 and 9 and is translated into a protein with a truncated ligand binding domain (LBD) which is thought to enhance GRα activity. GRα is by far the most active form of the receptor, GRβ is thought to be a dominant negative regulator of the receptor, and little is known about the function of GR-P.

Alternative first exon usage and 3’ splice variants

The recent observation that transcription factors binding to pol II transcribed promoters modulate alternative splicing, supports a physical and functional link between transcription and splicing [22]. Several factors were identified that were critical for the recruitment of a specific set of co-regulators to pol II transcribed gene promoters and the production of a specific splice variants. The splice variant produced depends on the structural organisation of the gene and the nature of the co-regulators involved [23]. A link between transcription initiation sites and the resulting splice variant was suggested since it was shown that promoters controlled alternative splicing also via the regulation of pol II elongation rates or processivity. Slow pol II elongation paired with internal elongation pauses favoured the inclusion of alternative exons governed by an exon skipping mechanism, whereas high elongation rates of pol II, without internal pauses favoured the exclusion of such exons.
GR Transcription

Many eukaryotic genes contain multiple promoters that are alternatively used for the production of different protein isoforms, with important physiological consequences. However, the GR with its variable 5' UTR, and alternative splicing in the 3' coding region is unique. Little is known about the association between the promoter usage and the resulting GR protein isoform. The 5' UTR has tight control over local GR expression levels. There seems to be also a poorly understood statistical link between the 5'UTR and 3' splice variants produced. One of the first studies to address this question showed that exon 1A3, and to a lesser extend 1B and 1C contribute most to the expression of GR-α isoform [24]. By comparing the most abundant exon 1 containing transcripts (1A, 1B 1C) with GRα, GRβ, and GR-P containing transcripts in different tissues and cell lines, Russcher et al. found a correlation between promoter usage and alternative splicing of the GR gene [25]. More specifically they found that the expression of GRα is preferentially regulated by promoter 1C, whereas 1B usage favours the expression of GR-P isoform. No association was found with transcripts including exon 9β or with those transcribed from 1A, suggesting that GRβ splicing may be associated with one of the recently identified exon 1 variants such as 1D to 1F and 1H that were not included in the above study [25].

We also confirmed that in post-mortem brain tissues of patients with major depressive disorder (MDD) altered promoter usage influenced the resulting 3’GR isoform, with a negative correlation between GR-P expression and promoter 1B usage in all brain areas of MDD patients but not in normal control brains. A negative correlation was also found between the 1C promoter usage and GR-P expression in MDD brains. These results suggest that the promoters 1B or 1C do not play a significant role in GR-P expression in MDD, and that they were rather linked to other forms with lower expression [26]. Thus,
GR Transcription

current data suggest a link between the two ends of the mRNA transcript, but there is no
consensus as to the nature of this link.

Transcription factors and transcriptional control within the CpG island

The hGR was initially described as a housekeeping or constitutively expressed gene with
promoters that contain multiple GC boxes and no TATA or TATA-like box [27]. A wide
variety of transcription factors have been identified that bind in the CpG island upstream
of the gene. The description of the transcription factors active within this region is
complicated by their tissue-specific usage. These transcription factors were not assigned
to the different exon 1 promoters since most of this work was performed before our
detailed description of the first exons in this region. The transcription factors so far
identified are summarised (Table 1) and their location within the CpG island shown in
Figure 1C. Initially, 11 DNAse 1 footprints representing unique transcription factor
binding sites were found in the 1C to 1F region of the CpG island (-3259 to -2522 from
the ATG start codon) including, one AP-2 and 5 Sp1 binding sites were identified [28].

It was initially thought that the latter transcription factors played an essential role in the
basal expression of the hGR, although this is now less clear. Further studies identified
one of the footprints in promoter 1C as a binding site for the transcription factor Yin
Yang 1 [15]. YY1, expressed in a wide variety of mammalian cell types, is a zinc-finger
transcription factor that can act as an activator, a repressor, or an initiator of transcription
[17, 29]. The same authors also revealed three other YY1 sites and another Sp1 site,
initially assigned to promoter 1B. The later identification of promoter 1D suggested that
GR Transcription

these YY1 sites are probably associated with this promoter [30]. Similarly, the Sp1 sites  
correspond to a region that was later identified as promoter 1J [20].

Similarly, several transcription factors initially assigned to promoter 1C should be  
reassigned to promoter 1F. AP-1, a transcription complex whose components are encoded  
by c-fos and c-jun proto-oncogenes binds to the AP-1 site within the hGR promoter 1F  
[14, 31]. This same region was also shown to bind Ku70 and Ku 80 in a tissue-specific  
manner [32]. Whilst most of the transcription factors identified upregulate GR  
expression, GRF-1 (glucocorticoid receptor DNA binding factor 1) has been identified as  
a repressor of GR transcription [33, 34]. At the 3’ end of the rat 1γ promoter a NGFI-A  
binding site was identified only 2 bp upstream of the transcription initiation site of this  
exon [35]. Recently, the homologous human NGFI-A binding site, together with  
numerous non-canonical NGFI-A sites were identified in promoter 1F of the hGR [36].

As a transcription factor, GR also auto-regulates its own CpG island promoters. Several  
glucocorticoid response element (GRE) half-sites, acting in concert with c-Myb, and c-  
Ets protein members have been identified in promoter 1D, 1E, 1F and 1C [37].

The currently known transcription factors provide only an incomplete picture of the  
complex regulatory mechanisms. For instance, little is known about the proximal  
elements in promoters 1B and 1H. Using an in silico phylogenetic footprinting technique  
we were able to find the majority of the experimentally identified transcription factors,  
and predicted a wide variety of factors that are conserved between many species [38].

These are interesting candidate regulators of GR expression that warrant further  
investigations.
GR Transcription

It has not yet been shown whether the transcription factors that bind immediately upstream of exons D, E, F, H, and I, activate the expression of these exons. Only site 13, one of 6 in the region immediately upstream of exon 1F (Fig. 1C), has been shown to activate transcription of the downstream exon. Furthermore, the link between the transcription factors previously identified, or predicted, and the transcription of the new CpG island first exons must be established.

Transcriptional control upstream of the CpG island: Exon 1A and 1I

Whilst the majority of the GR first exons and their promoters are located within the CpG island, exons 1A and 1I map 25 Kb upstream of the CpG island and 32 Kb upstream of the main GR ORF in the distal promoter region [16, 20]. Exon 1A has also been identified in the mouse, and three possible homologues 1, 12, and 13 have been found in the rat [39, 40]. The human promoter 1A generates 3 alternatively spliced transcripts, 1A1, 1A2 and 1A3 [16]. Expression of the 1A transcripts appears to be limited to the immune system in both humans and rodents. The human 1A3 transcript is widely expressed in both acute lymphoblastic leukemia (ALL) cell lines and in children with this malignancy [16, 20, 25, 41-43]. Similarly, exon 1I is used predominantly in T cells, although it is also present in HeLa cells [20].

Promoter 1A is regulated by the GR itself. The human 1A promoter contains a GC-responsive cassette containing a non-canonical GRE adjacent to overlapping binding sites for c-Myb and c-Ets protein family members. In the presence of c-Myb the ligand-bound GR is recruited to the promoter and up-regulates 1A transcripts, while the interaction with c-Ets family members leads to a repression of 1A promoter activity [18, 44]. This
GR Transcription

explains in part the opposite effects observed in different tissues: in most tissues GC
decreases GR expression, but in and certain T cells GCs increase GR expression. The
synergy between c-Ets and the GR has been shown to be responsible for the down
regulation of 1A3 in the B lymphoblastoid cell line IM-9 after GC exposure. In contrast,
the presence of c-Myb in T cells increased the activity of promoter 1A in human cell lines
such as the ALL T cell line CEM-C7. As a result of the higher level of total GR T cells
are more sensitive to GC induced apoptosis, and 1A3- transcripts were shown to be the
most GC responsive of all first exons investigated [16, 24]. Although 1A containing
transcripts correspond to only about 10% of all GR transcripts [42], their contribution to
the tissue-specific response to GC treatment was considered essential [16, 24]. The
human promoter 1A also has a functional binding site for Interferon Regulatory Factors
(IRF-1 and IRF-2), however IFNγ stimulation of CEM-C7 cells did not increase 1A
transcript levels, nor did it alter their susceptibility to GC-mediated apoptosis [45].
Thus the evidence suggests that 1A transcript levels are critically involved in mechanisms
of therapeutic induction of apoptosis in ALL T cells. However, GC resistance in pediatric
T- and B- ALL cells obtained directly from patients did not correlate with either the basal
or the stimulated expression of the 1A-, 1B- or 1C- transcripts. The relation between GC
sensitivity and expression of GR transcripts may be complicated by the overlapping
effects such as multi drug resistance genes that prevent GC concentrations from building
up in the cells. In addition, ALL cells were shown to have an up-regulated GR expression
upon prednisolone treatment regardless of their phenotype or sensitivity to GC-induced
apoptosis [42].
GR Transcription

In mice the presence of the membrane-bound GR was a better correlate of GC-induced apoptosis than the intracellular GR level [46, 47]. Exon 1A was found to be highly expressed in a T lymphoma cell line with elevated levels of membrane-bound GR and enhanced sensitivity to GC-dependent cytotoxicity. 1A transcripts appear to contain all the necessary information for both the synthesis and the subcellular trafficking of the membrane GR, although the exact mechanism remains unknown [41]. It is interesting to hypothesise that the initially presumed link between GC induced apoptosis and 1A transcript levels is indirect, with 1A transcripts producing membrane GR the true correlate of GC sensitivity. Recently a similar membrane GR has been detected in human cells, but its sequence is not known [48, 49]. The membrane GR may eventually explain the discrepancy between the in vivo and in vitro observations in ALL. It is interesting to speculate that, as in the mouse, the different first exons determine the cellular fate of the human GR.

Further sequence analysis of mouse exon 1A found 5 small upstream-ORFs preceding the major ORF of the GR. It has been shown that GR expression from the 1A transcript requires translation of the peptide encoded by uORF-2. The peptide was detectable in both cells with naturally 1A transcripts and cells transfected with 1A. The role of this peptide in the regulation of the membrane GR is still unclear [50].

Epigenetic programming of GR promoters

Epigenetic methylation of the 5’-cytosine of a CpG dinucleotide is associated with gene silencing either by inhibition of transcription factor binding (Fig. 2) or by chromatin inactivation [51-53]. For instance, prenatal epigenetic methylation governs genomic
GR Transcription

imprinting and inactivation of one X-chromosome [54]. The epigenetic chromatin status
is sensitive to the host environment. Thus, epigenetic methylation represents a link
between the environment and gene activity. In particular, early life events can have a
long-lasting effect on epigenetic programming [52, 53]. In many instances, minor
changes in GR levels can have a significant impact, for example on feedback regulation
of the HPA axis, where hippocampal or pituitary GR levels determine the HPA axis set-
points and the response to stress.

Experimentally, maternal care such as licking-grooming (LG) and arched-back nursing
(ABN) has been shown to translate into epigenetic methylation of the GR promoter \( l_7 \)
with profound and lasting effects on the stress response of the off-springs. [55]. The
NGFI-A binding site in the GR promoter \( l_7 \) (Fig. 3), homologous to the human 1F, was
highly methylated (>80%) in the offspring of low caring mothers whereas it was rarely
methylated in the offspring of high caring rats [35]. As a result, binding of NGFI-A to the
GR \( l_7 \) promoter was inhibited in the hippocampus of offspring of low caring mothers and
GR \( l_7 \) expression was reduced [56]. Interestingly, these effects were reversed by cross-
fostering indicating a direct effect of maternal care on the epigenome of the offspring
[35]. Infusion of L-methionine reversed these effects on methylation and NGFI-A
binding to the exon \( l_7 \) promoter in the rat brain [57].

In Lewis and Fisher rats that naturally differ in their stress response and hippocampal GR
levels, the \( l_7 \) promoter was shown to be un- or poorly methylated throughout (mostly
below 10% and never exceeding 30%), with no difference between the two strains.

Feeding these rats a methyl-supplemented diet had no significant effect on the GR
promoter \( l_7 \) methylation levels [58].
Using the maternal separation model to change the stress response in rat pups (Table 2), Daniels et al. observed elevated NGFI-A levels and significant behavioural changes. In this model, the 17 promoter, including both CpG sites within the NGFI-A binding site, was uniformly un-methylated even after applying the maternal separation stressor [59].

Epigenetic programming of the GR is not limited to central tissues such as the hippocampus. It has also been proposed that dietary restriction could lead to changes in DNA methyl content, affecting epigenetic programming of the GR promoter both centrally and peripherally [60]. Feeding a protein-restricted diet to pregnant dams lead to a hypomethylation of the major GR promoter 110 and to an increased expression of GR in the liver of these rat pups [61].

These animal models showing changes in GR promoter methylation are not easily transferred to humans. Nevertheless maternal adversities like depression or protein restriction and their effects on the epigenome of offspring have been investigated (Table 2). Oberlander et al. showed that prenatal exposure to maternal depression leads to increased methylation levels of the GR promoter 1F at the NGFI-A binding site in cord blood of newborns [62]. Like in most rat experiments, methylation levels were uniformly low (5-10 %) with small but significant differences between children of depressed and healthy mothers.

Several studies also investigated alterations in the human GR 1F promoter in specific disease populations based on the rat 17 data of Weaver et al. [35]. In neurological disorders such as Parkinson’s disease, Alzheimer’s or dementia no hypermethylation of the 1F promoter and the NGFI-A binding site could be found [63]. We showed that in major depressive disorder there was no methylation of the NGFI-A binding site of the 1F
GR Transcription

promoter in several regions of human post-mortem brains [26]. However, in suicide
victims with a history of child abuse, McGowan et al. found increased methylation
patterns compared to suicide victims without abuse [36]. In this study, methylation of
another putative NGFI-A binding site within promoter 1F resulted in decreased
expression levels of 1F transcripts and overall GR levels. The known NGFI-A binding
site was completely unmethylated in all of the suicide victims. Thus, it is possible that
other transcription factor binding sites are important for the transcriptional regulation of
the GR 1F promoter and that these are more sensitive to epigenetic modifications.

Interestingly, in all four of the above studies [26, 36, 62, 63], levels of methylation were
always very low in comparison to those in the LG-ABN rats [35]. Investigating the
complete GR CpG island, we were able to show highly variable methylation patterns
among different GR promoters in PBMC’s of healthy donors, suggesting that epigenetic
programming may not be restricted to the 1F promoter, but operates throughout the CpG
island [33]. It remains unclear, however, what triggers changes in methylation, and when
are the different tissues most susceptible to epigenetic programming. Despite some
contradictions it seems that levels of methylation are consistently low in the brain, and
somewhat higher and more variable at least in the blood mononuclear cells and liver.

The future - GR post-transcriptional regulation by miRNAs?

MicroRNAs (miRNAs) were discovered in 2001 [64, 65]. In mammalian, miRNAs were
predicted to regulate up to 30% of all genes [66]. So far they have been shown to be
involved in almost every cellular process investigated [67, 68]. It is now recognised that
miRNAs account for about 1% of the human genome and that they play a key role in
many regulatory pathways such as development timing, cell differentiation and apoptosis [69]. MicroRNAs are single-stranded RNA molecules of about 21 nucleotides in length that bind through imperfect base pairing to their target mRNAs, interfering with translational output [70]. miRNAs can either be encoded within an intronic locus [71, 72] or they cluster within intergenic regions [73]. miRNA gene transcription gives a primary miRNA transcript (pri-miRNA) that after subsequent processing (reviewed in [74]) is incorporated into the RNA Induced Silencing Complex (RISC). The seed region, a 6-8 nucleotide long sequence situated at the 5' end of the miRNA, nucleates the binding of miRNAs to their mRNA targets, requiring several contiguous and perfect by matching base pairs [75]. With a few exceptions, miRNA-binding sites in metazoan mRNA lie in the 3'UTR of the target gene and are usually present in multiple copies [76-80]. The exact mechanisms by which miRNAs repress translation remain unclear. However, it is known that miRNAs not only induce mRNA deadenylation and degradation but they also seem to interfere with protein synthesis [81].

There is now evidence that miRNAs are involved in the regulation of hippocampal GR levels modulating the HPA axis responsiveness [82]. The 3’ UTR of the GR was predicted to contain numerous seed regions recognized by a variety of miRNAs two of which were miR-18 and miR-124a [83]. Both reduced GR protein levels and the overall GR activities measured by target gene expression levels. miR-18 inhibited GR mRNA translation in cultured neuronal cells and was shown to be up-regulated in the paraventricular nucleus of Fischer 344 rats, a strain hypersensitive to stress. Interestingly, it has been known for some time that Fischer 344 rats have low GR levels in the hippocampus compared to Lewis rats. Together, these results suggest a role of miR-18 in
GR Transcription

GR down-regulation, increasing the vulnerability to repeated stress [84]. In contrast to miR-18, which was expressed in numerous tissues, miR-124a expression was restricted to certain brain regions and changed during the stress neonatal hypo-responsive period [85]. Little is known about miRNA targeting 5'UTRs, although target sites for endogenous miRNAs can also be identified in these upstream gene regions. A recent study reported mRNAs to be repressed as efficiently by miRNA binding in the 5' as in the 3' UTR although previously 5' miRNAs were shown to be less effective than those binding in the 3'UTRs [86]. Most surprisingly, by binding to the 5'UTR, miRNA family member miR-10 enhanced the translation of its target mRNA [87].

Nothing is known so far about miRNAs targeting the 5’ UTR of GR mRNAs. Considering the high variability of the GR 5’UTRs and tissue-dependent first exon usage [19, 30], it is attractive to hypothesise that miRNAs may add yet another layer of transcriptional complexity to the GR. Using the online version of PITA (Probability of Interaction by Target Accessibility) [88], we predicted a number of highly significant miRNA binding sites within the GR first exon Figure (4). To determine if and how these miRNAs are involved in GR posttranscriptional regulation remains a rewarding challenge.

Summary

The studies presented in this review demonstrate the multiple layers of complexity involved in the maintenance of the homeostasis of the ubiquitously expressed GR. Data from our laboratory and others have shown that the multiple first exons represent a first layer of complexity playing a particular role in tissue-specific transcriptional regulation.
GR Transcription

The abundance of the alternative exons is modulated by epigenetic methylation of their promoters. Finally, we suggest that the mature mRNA will be susceptible to miRNA effects that are transcript specific. Whilst the amount of the GR is thought to be the main determinant of the GC response, the third layer of complexity, beyond the scope of this review is the interaction of the mature GR with its numerous co-regulators and co-repressor proteins, as well as their interactions with the other steroid receptors upon ligation.

Amongst the many open questions is the effect of the alternative 5′ transcripts on translation. It is likely that the 5′UTR sequence will affect the recruitment of the translational machinery, translation rates, as well as mRNA stability. Together these will further regulate the amount of GR generated, the ultimate determinant of GC sensitivity. Determining how these different mechanisms converge to produce an adequate stress response in the healthy and the sick is one of the next major challenges.

When we published the structure of the GR 5′UTR its complexity was unique. However, over the last few years an unexpected variability in the 5′ UTR of genes throughout the complete genome has unfolded. In 2001 there were only 2 human genes known to have >10 first exons [89], in 2006 this had expanded to at least 43 [90]. Although there is no more recent data, comprehensive analysis of the ENCODE project data suggests that >20% of all human genes have multiple alternative promoters [91]. Data from the Database of Transcription Start sites (DBTSS) even suggests that >50% of all genes have alternative first exons and associated promoters [92]. The discrepancy between these two
GR Transcription

estimates is most probably due to the greater variety of tissue and cell types available in the DBTSS, underscoring tissue specificity of these first exons. Thus, genes with alternative first exons, translated or not, seem to be the rule rather than the exception as initially thought for the GR.
GR Transcription

References:


GR Transcription


GR Transcription


[61] Lillycrop KA, Slater-Jefferies JL, Hanson MA, Godfrey KM, Jackson AA, Burdge GC. Induction of altered epigenetic regulation of the hepatic glucocorticoid receptor in the offspring of rats fed a protein-restricted diet during...
pregnancy suggests that reduced DNA methyltransferase-1 expression is involved in impaired DNA methylation and changes in histone modifications. Br J Nutr 2007;97:1064-73.


[71] Cai X, Hagedorn CH, Cullen BR. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA 2004;10:1957-66.


GR Transcription


[84] Lytle JR, Yario TA, Steitz JA. Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. Proc Natl Acad Sci U S A 2007;104:9667-72.


GR Transcription

Figure 1: Structure of the GR gene (NR3C1; OMIM + 138040; NR3C1; 5q31-q32), the potential mRNA transcripts and the binding sites within the CpG island. Panel A The genomic structure of the GR. □ 5’ untranslated distal exons; ☡ 5’ untranslated CpG island exons; ■ Common exons; ▽ 3’ alternatively spliced exons. Panel B shows the potential mRNA transcripts encoding the three GR isoforms: GRα, GRβ and GR-P. Panel B shows the location of the known transcription factor binding sites. ☻ IRF 1 and IRF2 (position 1); ☺ c-Myb, c-Ets ½ and PU1 (position 4); ☽ Ying Yand 1 (positions 5, 6, 7 and 25); ☞ Glucocorticoid response elements (GRE, positions 2, 3, 8, 21, and 22); ☢ Sp1 binding sites (positions 9, 10, 12, 13, 16, 19, 20, 21, and 24); ☬ NGFIA binding site (position 17); ™ Glucocorticoid response factor-1 (GRF1, position 18); ☢ Ap-1 (position 15); and ☝ Ap-2 (position 23).

Figure 2: DNA methylation inhibits transcriptional regulation of genes. A complex of transcription factors (TF), the DNA polymerase (Pol) and co-activators (open circles) normally binds to the GR promoter and thus regulates gene expression. Epigenetic modifications can lead to methylation of the transcription factor binding site in the GR promoter (full ovals). Therefore binding to the promoter is inhibited and expression of the GR is reduced. Figure from [26] with permission.

Figure 3: Alignment of the rat GR promoter 17 and the human GR promoter 1F. Solid-lined boxes represent known canonical NGFI-A binding sites either in the rat [35, 58, 59] or the human [38, 26, 62, 63] promoter. The broken-lined box in the human 1F promoter represents a hypothetical non-canonical NGFI-A binding site [36].

Figure 4: The GR first exon 1C is rich in potential microRNA binding sites. MicroRNA binding was predicted based on the probability of interaction by target accessibility (PITA) [88]. Individual binding scores are from the colour scale. Values lower then -11 are normally considered highly relevant.
Table 1. Transcription factor binding sites in the hGR proximal promoter region

<table>
<thead>
<tr>
<th>Promoter</th>
<th>TF a</th>
<th>N°</th>
<th>Sequences</th>
<th>Cell lines / Tissues</th>
<th>Technique b</th>
<th>Location c</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distal Promoter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRF-1</td>
<td></td>
<td></td>
<td>GTAGAGCGGAATCCTTTTACCTGCTGQGG</td>
<td>CEM-C7</td>
<td>FP, EMSA, RG</td>
<td>-34574 -34544</td>
<td>[16, 45]</td>
</tr>
<tr>
<td>IRF-2</td>
<td></td>
<td></td>
<td>GTAGAGCGGAATCCTTTTACCTGCTGQGG</td>
<td>CEM-C7, Jurkat</td>
<td>FP, EMSA, RG</td>
<td>-34574 -34544</td>
<td>[16, 45]</td>
</tr>
<tr>
<td>GR-α</td>
<td></td>
<td></td>
<td>TCTGATACCAAATCACTGGACCTTA</td>
<td>CEM-C7</td>
<td>FP, EMSA</td>
<td>-34490 -34466</td>
<td>[18]</td>
</tr>
<tr>
<td>GR-β</td>
<td></td>
<td></td>
<td>GACCGTCCTATGGCGCAG</td>
<td>CEM-C7, IM-9</td>
<td>FP, EMSA, ChIP</td>
<td>-34436 -34419</td>
<td>[18, 37, 44]</td>
</tr>
<tr>
<td>c-Myb</td>
<td></td>
<td></td>
<td>ATGTGTCACCCGGAAGCACT</td>
<td>CEM-C7</td>
<td>FP, EMSA, ChIP</td>
<td>-34421 -34402</td>
<td>[18, 44]</td>
</tr>
<tr>
<td>Proximal promoter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1D</td>
<td></td>
<td>1</td>
<td>CCAAGATGG</td>
<td>NIH 3T3, Hela</td>
<td>FP, D, E</td>
<td>-4807 -4799</td>
<td>[15]</td>
</tr>
<tr>
<td>mYY1</td>
<td></td>
<td>2</td>
<td>CCAAGATGG</td>
<td>NIH 3T3, Hela</td>
<td>FP, D, E</td>
<td>-4635 -4627</td>
<td>[15]</td>
</tr>
<tr>
<td>pYY1</td>
<td></td>
<td>3</td>
<td>CCAAGATGG</td>
<td>NIH 3T3, Hela</td>
<td>FP, D, E</td>
<td>-4591 -4583</td>
<td>[15]</td>
</tr>
<tr>
<td>GRE</td>
<td></td>
<td>4</td>
<td>GGCTGGCGACGCCTCGGCACTCAGTCTGCTTCAAG</td>
<td>Jurkat, IM-9, CEM-C7</td>
<td>ChIP, E</td>
<td>-4574 -4540</td>
<td>[37]</td>
</tr>
<tr>
<td>1J</td>
<td></td>
<td>5</td>
<td>GCTGGCGCGGGGCTT</td>
<td>NIH 3T3, Hela</td>
<td>FP, E</td>
<td>-4250 -4235</td>
<td>[15]</td>
</tr>
<tr>
<td>1J</td>
<td></td>
<td>6</td>
<td>TTCGCGGCTGG</td>
<td>Jurkat, HepG2, Hela</td>
<td>RG, FP, E</td>
<td>-4011 -3999</td>
<td>[17]</td>
</tr>
<tr>
<td>1E</td>
<td></td>
<td>7</td>
<td>GTGGAAGAGAGGTCGAGGTTTC</td>
<td>Jurkat, IM-9, CEM-C7</td>
<td>ChIP, E</td>
<td>-3962 -3939</td>
<td>[37]</td>
</tr>
<tr>
<td>1B</td>
<td></td>
<td>8</td>
<td>CACATTGGCGGAGGGG</td>
<td>Jurkat, HepG2, Hela</td>
<td>RG, FP, E</td>
<td>-3774 -3757</td>
<td>[17]</td>
</tr>
<tr>
<td>1B</td>
<td></td>
<td>9</td>
<td>TTGACTTGGCACGGGGGCGGCGCC</td>
<td>Jurkat, HepG2, Hela</td>
<td>RG, FP, E</td>
<td>-3750 -3730</td>
<td>[17]</td>
</tr>
<tr>
<td>1F</td>
<td></td>
<td>10</td>
<td>GCACCGTTTTCCGTCGACAAGGCTGCGTGCCGACACT</td>
<td>Jurkat, IM-9, CEM-C7</td>
<td>ChIP, E</td>
<td>-3438 -3393</td>
<td>[37]</td>
</tr>
<tr>
<td>AP-1</td>
<td></td>
<td>11</td>
<td>TGACACA (consensus TGACGTCA)</td>
<td>AT-20, NIH3T3</td>
<td>EMSA</td>
<td>-3401 -3395</td>
<td>[31, 14]</td>
</tr>
<tr>
<td>Sp1</td>
<td></td>
<td>12</td>
<td>TGCGCGCCCGGGGAGA</td>
<td>Hela, NIH3T3, CV1, HepG2</td>
<td>RG, FP, EMSA</td>
<td>-3228 -3213</td>
<td>[28]</td>
</tr>
<tr>
<td>NFI A</td>
<td></td>
<td>13</td>
<td>GGCGCGCGCGGGG</td>
<td>Rat Hippocampi / HEK293</td>
<td>ChIP</td>
<td>-3227 -3216</td>
<td>[35, 36]</td>
</tr>
<tr>
<td>GRF-1</td>
<td></td>
<td>14</td>
<td>GAAGGAGAGGTACGAGAAACGAGGAAATCGGAGGAACCTCGGTTGG</td>
<td>MCF7, CV-1</td>
<td>EMSA</td>
<td>-3215 -3176</td>
<td>[33, 34]</td>
</tr>
<tr>
<td>Sp1</td>
<td></td>
<td>15</td>
<td>TCTAAGCGCGGCCGCGAGA</td>
<td>Hela, NIH3T3, CV1, HepG2</td>
<td>RG, FP, EMSA</td>
<td>-3172 -3153</td>
<td>[28]</td>
</tr>
<tr>
<td>1C</td>
<td></td>
<td>16</td>
<td>GAGTTTGGGCGGCGGGG</td>
<td>Hela, NIH3T3, CV1, HepG2</td>
<td>RG, FP, EMSA</td>
<td>-3107 -3088</td>
<td>[28]</td>
</tr>
<tr>
<td>1C</td>
<td></td>
<td>17</td>
<td>GCGCGCGGGGGGCGGCGGCGGGCGGCC</td>
<td>Hela, NIH3T3, CV1, HepG2</td>
<td>RG, FP, EMSA</td>
<td>-3080 -3061</td>
<td>[28]</td>
</tr>
<tr>
<td>GRE</td>
<td></td>
<td>18</td>
<td>CTGCAGTGGCACAGGCTCACTGCACTGCTGCTCC</td>
<td>Jurkat, IM-9, CEM-C7</td>
<td>ChIP, E</td>
<td>-2971 -2931</td>
<td>[37]</td>
</tr>
<tr>
<td>AP-2</td>
<td></td>
<td>19</td>
<td>CGCGCGCGGGCCCTGCGGCGGGG</td>
<td>Hela, NIH3T3, CV1, HepG2</td>
<td>RG, FP, EMSA</td>
<td>-2923 -2901</td>
<td>[28]</td>
</tr>
<tr>
<td>Sp1</td>
<td></td>
<td>20</td>
<td>CGCGCGCGCGGGCGCGCTCCA</td>
<td>Hela, NIH3T3, CV1, HepG2</td>
<td>RG, FP, EMSA</td>
<td>-2856 -2838</td>
<td>[28]</td>
</tr>
<tr>
<td>IYY1</td>
<td></td>
<td>21</td>
<td>CTCTCCATTGTTT</td>
<td>NIH 3T3, Hela</td>
<td>FP, D, E</td>
<td>-2755 -2743</td>
<td>[28]</td>
</tr>
</tbody>
</table>

a TF: Transcription Factor
c Locations with respect to the ATG start codon in exon 2.
## Table 2: Methylation analyses of GR promoter regions

<table>
<thead>
<tr>
<th>GR promoter</th>
<th>Species</th>
<th>Model</th>
<th>Method</th>
<th>Tissue</th>
<th>Overall methylation levels</th>
<th>CpG specific comments †</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.7</td>
<td>r. nor.</td>
<td>LG-ABN</td>
<td>colony sequencing</td>
<td>hippocampus</td>
<td>&lt;100%</td>
<td>80-100% at CpG 16-17 in Low LG-ABN</td>
<td>[35]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0-10% at CpG 16-17 in High LG-ABN</td>
<td>[58]</td>
</tr>
<tr>
<td>1.7</td>
<td>r. nor.</td>
<td>met sup</td>
<td>colony sequencing</td>
<td>hippocampus</td>
<td>&lt;15%</td>
<td>at CpG 16-17</td>
<td>[59]</td>
</tr>
<tr>
<td>1.7</td>
<td>r. nor.</td>
<td>maternal sep</td>
<td>direct sequencing</td>
<td>hippocampus</td>
<td>not detected</td>
<td>at CpG 16-17</td>
<td>[61]</td>
</tr>
<tr>
<td>1.10</td>
<td>r. nor.</td>
<td>PR diet</td>
<td>MS-PCR</td>
<td>liver</td>
<td>unknown</td>
<td>relative to control</td>
<td>[61]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PR rats had 30% lower methylation</td>
<td>[61]</td>
</tr>
<tr>
<td>1C</td>
<td>r. nor.</td>
<td>HV</td>
<td>methylation sensitive restriction enzyme PCR assay</td>
<td>umbilical blood</td>
<td>unknown</td>
<td>correlation between relative methylation and DNMT1 expression</td>
<td>[61]</td>
</tr>
<tr>
<td>1D</td>
<td>h. sap.</td>
<td>HV</td>
<td>colony sequencing</td>
<td>PBMCs</td>
<td>&lt;50%</td>
<td>stochastic and unique</td>
<td>[38]</td>
</tr>
<tr>
<td>1E</td>
<td>h. sap.</td>
<td>HV</td>
<td>colony sequencing</td>
<td>PBMCs</td>
<td>&gt;25%</td>
<td>stochastic and unique</td>
<td>[38]</td>
</tr>
<tr>
<td>1F</td>
<td>h. sap.</td>
<td>HV</td>
<td>colony sequencing</td>
<td>PBMCs</td>
<td>&lt;10%</td>
<td>at CpG 37-38</td>
<td>[38]</td>
</tr>
<tr>
<td>1H</td>
<td>h. sap.</td>
<td>HV</td>
<td>colony sequencing</td>
<td>PBMCs</td>
<td>&lt;75%</td>
<td>stochastic and unique</td>
<td>[38]</td>
</tr>
<tr>
<td>1F</td>
<td>h. sap.</td>
<td>Dep</td>
<td>pyrosequencing</td>
<td>cord blood</td>
<td>&lt;7%</td>
<td>&lt;5% at CpG 37-38</td>
<td>[62]</td>
</tr>
<tr>
<td>1F</td>
<td>h. sap.</td>
<td>Dep, Al, Park, Dem</td>
<td>MS-PCR</td>
<td>hippocampus</td>
<td>&lt;10%</td>
<td>only 1 subject had visible methylation all other donors unmethylated</td>
<td>[63]</td>
</tr>
<tr>
<td>1F</td>
<td>h. sap.</td>
<td>Suicide / abuse</td>
<td>colony sequencing</td>
<td>hippocampus</td>
<td>&lt;40%</td>
<td>0% at CpG 37-38</td>
<td>[36]</td>
</tr>
<tr>
<td>1B</td>
<td>h. sap.</td>
<td>MDD</td>
<td>pyrosequencing</td>
<td>hippocampus</td>
<td>0-12%</td>
<td></td>
<td>[26]</td>
</tr>
<tr>
<td>1E</td>
<td>h. sap.</td>
<td>MDD</td>
<td>pyrosequencing</td>
<td>hippocampus</td>
<td>4-9%</td>
<td></td>
<td>[26]</td>
</tr>
<tr>
<td>1F</td>
<td>h. sap.</td>
<td>MDD</td>
<td>pyrosequencing</td>
<td>hippocampus</td>
<td>0-4%</td>
<td>0-2% at CpG 37-38</td>
<td>[26]</td>
</tr>
<tr>
<td>1J</td>
<td>h. sap.</td>
<td>MDD</td>
<td>pyrosequencing</td>
<td>hippocampus</td>
<td>0-7%</td>
<td></td>
<td>[26]</td>
</tr>
</tbody>
</table>

† all CpG numbers as in Figure 3

**Abbreviations:** r. nor, rattus norwegianus; h. sap, homo sapiens; LG-ABN, licking-grooming and arched-back nursing; met sup, methy-supplemented diet; maternal sep, maternal separation; PR, protein-restricted; HV, healthy volunteer; Dep, depression; Al, Alzheimer’s disease; Park, Parkinson’s disease; Dem, dementia; MDD, major depressive disorder; MS-PCR, methylation-specific PCR
Figure 1
Figure 2

GR promoter

Transcription

No transcription

Figure 2
Figure 3

Rat   (1698) GGCCAGGGCTCTGCGGCACCGTTTCCGTGCCATCCT
       GTAGCCCCCTCTG CTAGTGAGACACACTTC1Ge2 GCAACTC C2 GCA-
       CTG-CCCTCTG
       GTAGCCCCTCTG
       CTAGTGTGACACACTTC
       GC
       2
       GCAACTC
       3
       GC

Human   (2494) GGCCGAGACGCTGCGGCACCGTTTCCGTGCAACCC
       GTAGCCCCTCTG
       GTAGCCCCTCTG
       GAAGTG--ACACACTTC
       AC
       12
       GCAACTC
       13
       GCA
       14
       GCC
       15
       GCCC
       16

Rat    (1795) GTTGGC
       4
       GGC
       5
       -GC
       6
       GC
       7
       GGACAC CCCT
       GC
       8
       GGCTCTGCC
       G-
       -------
       -----  
       ----------------------------GCTG G
       9
       GCTG

Human   (2592) GGC
       17
       GGC
       18
       GGC
       19
       GGC
       20
       GC
       21
       GGGCCAC TCAC
       GC
       22
       AGCTCAGCC
       GC
       23
       GGGAGGC
       GCCCC
       24
       GGCTCTGAGCC
       G
       25
       G
       26
       GCC
       27
       G
       28
       G
       29
       G
       30
       G
       31
       G
       32
       GCCAAG GGGAAGGAGGT
       GGGAGGC
       G

Rat    (1857) CTGTCA
       33
       GCTG
       34
       GAGCTCCC
       GAGT
       35
       GGGTCTGGAGCC
       G
       17
       G
       36
       G
       37
       G
       38
       G
       39
       G
       40
       G
       41
       G
       42
       G
       43
       G
       44
       G
       45
       G
       46
       G

Human   (2690) CTGTCA
       36
       GCTG
       37
       GAGCTCCC
       GAGT
       38
       GGGTCTGGAGCC
       G
       39
       G
       40
       G
       41
       G
       42
       G
       43
       G
       44
       G

Rat    (1857) C
       45
       GGAGCTGGGC
       46
       GGGGC
       47
       GGGAGGGAGCC
       G
       48
       G
       49
       G

Human   (2690) C
       50
       GGAGCTGGGC
       51
       GGGGC
       52
       GGAAGGAGGT
       G
       53
       G
       54
       G
       55
       G
       56
       G
       57
       G
       58
       G
       59
       G
       60
       G
       61
       G
       62
       G
       63
       G
       64
       G

hGR exon 1F
-----------------------------AGGT
Glucocorticoid receptor levels are transcriptionally controlled. The variable 5' region produces multiple splice variants coding a single protein, and is regulated by epigenetic methylation and tissue specific transcription factor usage.