Groucho proteins: transcriptional corepressors for specific subsets of DNA-binding transcription factors in vertebrates and invertebrates

Alfred L. Fisher and Michael Caudy

Department of Cell Biology and Anatomy, Cornell University Medical College, New York, New York 10021 USA

Extensive analysis during the last 10–15 years has identified many of the mechanisms and factors involved in the activation of eukaryotic gene transcription. Although activation is better studied and more appreciated, a growing body of work has shown that in many circumstances transcriptional repression is as important as activation in the regulation of gene expression (Gray and Levine 1996). Studies of the development of the Drosophila peripheral nervous system (PNS) have revealed that a family of basic helix–loop–helix (bHLH) transcriptional repressors, known as the Hairy-related proteins, plays critical roles during development by repressing target genes at multiple stages of neurogenesis (Fisher and Caudy 1998). Similarly, the early patterning of the Drosophila embryo requires genes encoding transcriptional repressor proteins as well as transcriptional activator proteins, and mutations in either the activators or repressors result in lethal defects in patterning (Carroll 1990; Gray and Levine 1996). In mammals, the expression of certain neuron-specific genes like the type II sodium channel is restricted to neurons by the REST/NRSF protein that transcriptionally represses these genes in the non-neuronal cells in which REST/NRSF is expressed (Chong et al. 1995; Schoenherr and Anderson 1995). In humans, the DAX-1 gene, which is mutated in congenital X-linked adrenal hypoplasia, encodes a transcriptional repressor, and mutations responsible for the disease phenotype map to its repression domain (Lalli et al. 1997; Zazopoulos et al. 1997). This review focuses on a family of transcriptional corepressor proteins, known as Groucho proteins, that are found in flies, mice, humans, frogs, and worms.

Groucho, the founding member of this family, was identified initially as a mutation that affects the development of the Drosophila nervous system, with one allele resulting in thick tufts of sensory bristles over the eyes, resembling the bushy eyebrows of the comedian Groucho Marx. The Groucho proteins serve as non-DNA binding corepressors for specific subsets of DNA-binding transcription factors, including the Hairy-related proteins, Runt domain proteins, Engrailed, and Dorsal, and they are essential for certain aspects of repression by each of these repressors (Paroush et al. 1994; Fisher et al. 1996; Aronson et al. 1997; Dubnicoff et al. 1997; Jimenez et al. 1997). These family members have been shown to be widely expressed both during development and in the adult, in contrast to the more limited expression pattern of their DNA-binding partners (Hartley et al. 1988; Stefan et al. 1992; Miyasaka et al. 1993; Schmidt and Sladek 1993; Choudhury et al. 1997; Pflugrad et al. 1997; Sharief et al. 1997). In addition, both Groucho and the human family member TLE1 have been shown to actively repress transcription when fused to a heterologous DNA-binding domain and directly bound to DNA (Fisher et al. 1996). Hence, the Groucho proteins are recruited to target gene promoters by direct binding to specific DNA-binding repressors, and once recruited, the Groucho proteins repress transcription via a conserved intrinsic repression domain.

Groucho proteins act as corepressors for specific active repressors

Groucho proteins are corepressors that are required for transcriptional repression by several distinct types of active transcriptional repressors. Corepressors are defined in this review as proteins that are required for the repression activity of a specific transcription factor but do not have the ability to bind DNA alone. Hence, such corepressors are recruited to target promoters by protein–protein interactions between a specific DNA-binding partner and the corepressor. For example, the mSin3 protein is a corepressor that is recruited to target promoters by the bHLH leucine-zipper repressor Mad as well as by unliganded nuclear receptor proteins, such as the thyroid hormone receptor (Pazin and Kadonaga 1997). In addition, mSin3 has intrinsic repressor activity and represses target gene promoters when directly bound to DNA by a heterologous DNA-binding domain. In contrast to mSin3, the human Kap-1 protein serves as a corepressor for a specific subfamily of zinc-finger transcription fac-
tors that contain a small domain known as the Krab domain to which Kap-1 binds (Friedman et al. 1996). The contrast between mSin3 and Kap-1 illustrates how some corepressors are relatively promiscuous and interact with several unrelated families of transcription factors, whereas others are more specialized and interact with a limited range of transcription factors. Finally, the Rb protein, which acts as a corepressor for the E2F transcription factor, shows another important property of at least some corepressors, that of regulation. The interaction between E2F and Rb is not constitutive but, instead, is regulated by phosphorylation of the Rb protein during the cell cycle (Sellers and Kaelin 1996). The ability of corepressor–DNA-binding partner interactions to be regulated allows some DNA-binding proteins to act as both repressors and activators, as is the case for E2F, and allows regulation by cell signaling or other pathways to modulate the control of target gene transcription.

The DNA-binding partners for the Groucho proteins all function as active transcriptional repressors in at least some contexts. Active repressors are distinguished from passive repressors by their mechanism of action (Cowell 1994). Passive repressors act by interfering with the transcriptional activator proteins that activate target gene transcription. This repression can occur by various mechanisms, such as competing for DNA sites bound by activator proteins, forming inactive heterodimers with activator proteins, or titrating coactivators required by the activator proteins (Cowell 1994). An example of a passive repressor is the Id protein, which represses transcription by forming non-DNA binding, inactive heterodimers with activator bHLH transcription factors such as MyoD (Benezra et al. 1990). In contrast, active repressors negatively regulate target genes by binding to repressor-specific sites in the target gene and repressing transcription by a distinct intrinsic repression domain (Cowell 1994). By this definition, many known transcriptional repressor proteins are active repressors. Active repression domains are defined operationally as domains that are necessary for repression, and they can be found either in a DNA-binding protein or in a corepressor. Because of the modular nature of proteins, these domains often can also confer repression when fused to a heterologous DNA-binding domain. For example, zinc-finger proteins containing the Krab domain are active repressors because they bind to repressor-specific sites and repress via the recruitment of the Kap-1 corepressor (Friedman et al. 1996). In agreement with this, the Krab domain will repress transcription when fused to a heterologous DNA binding protein and targeted to that protein’s binding sites. On the basis of these defining criteria, the various DNA-binding partners of the Groucho proteins can all act as active transcriptional repressors in at least some contexts (Jaynes and O’Farrell 1991; Jiang et al. 1992; Pan and Courey 1992; Oellers et al. 1994; Ohsako et al. 1994; Van Doren et al. 1994; Aronson et al. 1997). In addition, the Groucho proteins also act as active transcriptional repressors when directly bound to DNA by fusion to a heterologous DNA-binding domain, as discussed further below (Fisher et al. 1996).

Hairy-related proteins

The Hairy-related proteins are a family of bHLH transcription factors that, in Drosophila, are involved in segmentation, neurogenesis, sex determination, and myogenesis (for review, see Paroush et al. 1994; Fisher and Caudy 1998). In vertebrates, family members have been shown to be involved in neurogenesis and somite formation (Kageyama and Nakanishi 1997; Palmeirim et al. 1997). Additionally, members of this family such as the Drosophila Enhancer of split genes and certain mammalian Hairy Enhancer of split (HES) genes are important effectors of the Notch signaling pathway, which controls neuronal cell fate decisions in both vertebrates and Drosophila (Kageyama and Nakanishi 1997; Robey 1997; Fisher and Caudy 1998). The mammalian HES-1 protein is also a functional target for Nerve Growth Factor (NGF) signaling in the PC12 cell line (Ström et al. 1997). In these various pathways, the Hairy-related proteins act genetically as repressors of target genes in vivo.

The Hairy-related proteins are defined by the presence of two specific domains: a proline bHLH domain, which contains a proline at a specific position in the basic region, and a 4-amino-acid WRPW (Trp-Arg-Pro-Trp) domain found at the carboxyl terminus of the protein (Paroush et al. 1994; Fisher et al. 1996). As transcriptional repressors the Hairy-related proteins appear to act both by passive mechanisms, such as directly interacting with activator proteins (Sasai et al. 1992; Kageyama and Nakanishi 1997), and also by active mechanisms involving the binding to repressor specific DNA sites in target genes via the bHLH domain and the recruitment of the Groucho proteins (Sasai et al. 1992; Oellers et al. 1994; Ohsako et al. 1994; Paroush et al. 1994; Van Doren et al. 1994; Fisher et al. 1996; Heitzler et al. 1996; Kageyama and Nakanishi 1997).

The role of Groucho proteins in repression by the Hairy-related proteins has been studied by both biochemical and genetic means. Initially, a yeast two-hybrid screen performed with the Drosophila Hairy protein identified the Drosophila Groucho protein as a specific interacting protein (Paroush et al. 1994). Subsequently, both Groucho and the human TLE1 protein were shown to bind several Hairy-related proteins, and the WRPW motif was shown to be both necessary and sufficient for this interaction (Fig. 1) (Paroush et al. 1994; Fisher et al. 1996; Grbavec and Stifani 1996). The WRPW motif was initially proposed to act as a transcriptional repression domain for the Hairy-related proteins (Ohsako et al. 1994), and assays in cultured cells confirmed that this motif is a repression domain that by itself is sufficient to confer active transcriptional repression when fused to a heterologous DNA-binding domain (Fisher et al. 1996). The interaction between the WRPW motif and Groucho is required for transcriptional repression by these proteins both in Drosophila embryos and cultured cells (Par-
oush et al. 1994; Fisher et al. 1996). Embryos lacking Groucho show defects in segmentation, neurogenesis, and sex determination that are phenotypes consistent with a functional role for Groucho as a corepressor for the Hairy-related proteins shown previously to be involved in these developmental processes (Paroush et al. 1994). This combination of in vitro biochemistry, transcriptional repression assays in cells, and Drosophila genetics indicated that the Groucho proteins are essential corepressors for the Hairy-related proteins (Paroush et al. 1994; Fisher et al. 1996).

Runt domain proteins

The Runt domain family of transcription factors are found in both Drosophila and vertebrates. During Drosophila development, Runt and Lozenge play roles in segmentation, neurogenesis, sex-determination, and eye development (for review, see Duffy and Gergen 1994; Daga et al. 1996). In mammals, family members are essential for bone development and hematopoiesis (Speck and Terryl 1995; Okuda et al. 1996; Ducy et al. 1997; Komori et al. 1997; Otto et al. 1997; Rodan and Harada 1997). In humans, mutation or translocation of the AML1 and CBFA1 genes are a common occurrence in several forms of leukemia and lymphoma in adults and children, or are responsible for the inherited skeletal disorder cleidocranial dysplasia, respectively (Lo Coco et al. 1997; Mundlos et al. 1997). The Runt domain is a distinct DNA-binding domain that mediates both DNA binding and heterodimerization with a non-DNA-binding partner, for example, Brother or Big-Brother in Drosophila or CBFβ in mammals (Speck and Terryl 1995).

Remarkably, at least one isoform of all Runt domain proteins ends with the sequence WRPY, which is very similar to the WRPW motif present in all Hairy-related proteins (Fig. 1; Aronson et al. 1997). The WRPY motif is both necessary and sufficient to mediate protein–protein interactions between the Drosophila Runt protein or the mouse homolog of the AML1 gene, PEBPαB1, and Groucho (Fig. 1; Alifragis et al. 1997; Aronson et al. 1997). Given the large number of Hairy-related proteins and Runt domain proteins currently identified in multiple species, it is quite striking that only Hairy-related proteins have the WRPW motif and only Runt domain proteins have the WRPY motif (Aronson et al. 1997). Why there is a strict division between these two transcription factor families and the four amino acid motifs used to interact with the Groucho proteins is currently unclear. Interestingly, the Runt domain proteins act as both transcriptional activators and repressors, whereas, in contrast, the Hairy-related proteins appear to act only as transcriptional repressors (Aronson et al. 1997). One possible explanation is that Groucho interacts constitutively with the WRPW motif in Hairy-related proteins, whereas the interaction with WRPY-containing Runt domain proteins is regulated.

In Drosophila, the WRPY sequence and Groucho are essential for the repression of specific target genes by the Drosophila Runt protein (Aronson et al. 1997). A Runt protein lacking the WRPY motif is unable to repress transcription in cultured Drosophila cells or to repress the transcription of two specific target genes, hairy and even-skipped, in embryos (Aronson et al. 1997). However, this mutant still represses one in vivo target gene, engrailed, indicating that Runt also acts by an uncharacterized Groucho-independent repression mechanism (Aronson et al. 1997). In addition, the repression of the
target genes hairy and even-skipped in embryos by Runt also depends on the level of Groucho protein in the embryo, which indicates that Groucho is necessary for Runt to repress target genes in vivo (Aronson et al. 1997). Additionally, the mammalian PEBPαB1 and AMIL1b proteins have also been shown to have repressor activity in human HeLa cells, suggesting that Groucho-dependent transcriptional repression is a potential function of all Runt domain proteins with the WRPY motif (Aronson et al. 1997).

Engrailed

The Drosophila engrailed gene is a segment polarity gene that has roles in both the embryo and adult (Manak and Scott 1994). Engrailed is a homeobox protein (Fig. 1) that represses the transcription of target genes (Jaynes and O’Farrell 1991; Han and Manley 1993; Smith and Jaynes 1996). Recently, Engrailed has been shown to require Groucho for the repression of specific target genes in vivo (Jimenez et al. 1997). A repression domain from Engrailed located between amino acids 168 and 298 is composed of two subdomains (Jaynes and O’Farrell 1991; Han and Manley 1993; Smith and Jaynes 1996; Jimenez et al. 1997). The first domain, referred to as eh1 (Engrailed homology 1), is conserved in all Engrailed homologs and has been shown to be the primary repression domain in certain assays in transgenic flies (Smith and Jaynes 1996). The eh1 domain directly binds to Groucho (Fig. 1; Jimenez et al. 1997; Tolkunova et al. 1998); thus, the conservation of the eh1 domain in many other homeobox proteins from both Drosophila and vertebrates suggests that a subset of homeodomain transcriptional repressor proteins may show Groucho-dependent repression (Smith and Jaynes 1996; Jimenez et al. 1997). The second domain, referred to as region D, has been identified as the primary repression domain in transiently transfected cells (Han and Manley 1993). Similar to Runt, Engrailed appears to act by both Groucho-dependent and independent mechanisms during development because Engrailed still retains some residual repressor activity in flies when the eh1 domain is deleted, and grouch mutant clones in the developing wing do not exhibit the same phenotype as en engaged mutant clones (de Celis and Ruiz-Gomez 1995; Smith and Jaynes 1996; Jimenez et al. 1997). This may be explained by the presence of both the eh1 domain and region D within Engrailed as each subdomain could have independent repressor activity depending on the target gene being repressed. Hence, the eh1 domain/Groucho-mediated repression may only be required for specific target genes, whereas other targets may require repression mediated by region D.

Dorsal

The Drosophila Dorsal protein also has been shown to interact biochemically with Groucho and to exhibit Groucho-dependent repression in vivo (Dubnicoff et al. 1997). Dorsal is a member of the NF-κB/rel family of transcription factors (Fig. 1) and during development is involved in specifying the dorsoventral axis of the Drosophila embryo (Courey and Huang 1995). In the ventral regions, Dorsal acts to activate the transcription of ventral-specific genes and simultaneously, in the same cells, to inhibit the transcription of dorsal-specific genes (Courey and Huang 1995). Dorsal has been shown to function intrinsically as a transcriptional activator that binds to sites in the promoters of the ventral-specific genes to activate their transcription (Courey and Huang 1995). Repression of the dorsal-specific genes requires the binding of Dorsal to sites within their promoters, and Groucho is also required for Dorsal-mediated repression of these promoters (jiang et al. 1992; Pan and Courey 1992; Courey and Huang 1995; Dubnicoff et al. 1997). Because the interaction between Dorsal and Groucho is mediated by the rel domain that is found in NF-κB and related mammalian transcription factors (Fig. 1; Dubnicoff et al. 1997), it will be important to determine whether Groucho proteins can also interact with these mammalian proteins to similarly repress specific target genes.

In summary, two findings indicate that the Groucho proteins function as active transcriptional corepressors for the various DNA-binding partners described above. First is the genetic requirement for Groucho for in vivo transcriptional repression by the various Drosophila partner proteins (Paroush et al. 1994; Aronson et al. 1997; Dubnicoff et al. 1997; Jimenez et al. 1997). Second is the observation that the Groucho proteins have intrinsic active transcriptional repressor activity when directly bound to target genes by fusion to a heterologous DNA binding protein (Fisher et al. 1996). Together with the biochemical interactions between Groucho proteins and the repression domains of these partner proteins, these two findings clearly indicate that the Groucho proteins are active transcriptional corepressors that are recruited to target genes in vivo by specific subsets of DNA-binding proteins (Paroush et al. 1994; Fisher et al. 1996; Aronson et al. 1997; Dubnicoff et al. 1997; Jimenez et al. 1997).

Structural and functional domains of Groucho proteins

The Groucho proteins from Drosophila, Caenorhabditis elegans, Xenopus, rats, mice, and humans all share a similar primary sequence structure consisting of a series of seven highly conserved, carboxyl-terminal WD40 repeats (for discussion of WD40 repeats, see Nee et al. 1994), a highly conserved amino terminus, and a variable region that separates these two domains (Fig. 2A,B) (Hartley et al. 1988; Stifani et al. 1992; Miyasaka et al. 1993; Schmidt and Sladek 1993; Choudhury et al. 1997; Pfliugrad et al. 1997; Sharief et al. 1997). Proteins resembling the Groucho proteins, either consisting of part of the amino terminus but lacking the variable region and WD40 repeats, or consisting of part of the variable region and the WD40 repeats, but lacking the amino terminus, have also been isolated (Fig. 2C; Stifani et al. 1992; Miyasaka et al. 1993; Schmidt and Sladek 1993; Choudhury et al. 1997). At this time we will not classify these as
Groucho proteins as corepressors

Table 1. A summary of known properties of Groucho proteins.

<table>
<thead>
<tr>
<th>Groucho Domain</th>
<th>Interaction with WRWPW and WRPY</th>
<th>Interaction with EH1 Domain</th>
<th>Other Protein-Protein Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcriptional Repression Domain</td>
<td>Amino terminus</td>
<td>Nuclear Localization?</td>
<td>Phosphorylation?</td>
</tr>
<tr>
<td>Dimerization</td>
<td>Variable</td>
<td>WD40 Repeats</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. (A) Schematic diagram of the domain structure of Groucho proteins. Groucho proteins consist of three domains: the amino terminus, a series of seven carboxy-terminal WD40 repeats, and a variable region that separates the other two domains and shows poor sequence homology among family members. Below each domain are the functions that have been assigned to it. (B) Drosophila Groucho, human TLE1, 2, and 3, and C. elegans UNC-37 proteins (solid lines) possess all three of these domains. Given the structural and functional similarities between these proteins, they are considered to be members of the Groucho protein family. (C) In contrast, the human TLE4 and mouse AES proteins lack some of these domains. Additionally, the biological function of these proteins is currently unknown, and they could act as dominant-negative inhibitors of the Groucho proteins. For these reasons, proteins such as these are considered here to be related to Groucho proteins, but not part of the Groucho protein family.

Groucho proteins because their function is not yet known. Although they may act as corepressors for some of the known partners or a distinct set of partners, these proteins could also act as antagonists of the Groucho proteins by titrating an effector or by binding to the targets of repression and acting as dominant-negative inhibitors.

The amino terminus of the Groucho proteins was shown to act as a dimerization and active repression domain (Fig. 2A; Fisher et al. 1996; Pinto and Lobe 1996). This domain also contains potential phosphorylation sites for cdc2 and casein kinase II in close proximity to a nuclear localization sequence (Fig. 2A; Stifani et al. 1992). The amino terminus has been shown to repress both activated and basal transcription when bound directly to DNA and to have repressor activity equivalent to the full-length protein in several assays, suggesting that this domain is the major intrinsic repression domain (Fisher et al. 1996).

The exact mechanism used by the amino terminus to repress transcription is not known, but the recently described interaction between Groucho proteins and histone H3 may provide a mechanism, especially if the histone H3 interaction domain is found to localize to the amino terminus (Palapart et al. 1997). An interaction with histone H3 could lead to the assembly or stabilization of repressive chromatin on target gene promoters and thereby repress transcription. This mechanism is used by the SIR repressors from yeast (Hecht et al. 1995), but for the Groucho proteins the role of the interaction with H3 in transcriptional repression is untested (Palapart et al. 1997). The Groucho proteins may also work by other mechanisms in addition to, or instead of, the assembly of chromatin. Such mechanisms could include interaction with components of the basal transcription complex or the recruitment of other proteins with repressor or enzymatic activity, such as the HDAC1 histone deacteylase protein that is known to interact with other transcriptional repressors (Pazin and Kadonaga 1997). Because the Groucho proteins can repress both activated and basal transcription (Fisher et al. 1996), they are unlikely to act via the quenching (Levine and Manley 1989) of the activation domains of transcriptional activators. Instead, the Groucho proteins appear to be general repressors that can repress both basal transcription and transcription activated by a variety of activator proteins.

On the basis of studies of other WD40 repeat proteins, the seven WD40 repeats found at the carboxyl terminus of the Groucho proteins (Fig. 2A) are likely to be involved in protein-protein interactions. WD40 repeats generally function as protein-protein interaction domains and have been shown to form a β-propeller structure in which each repeat projects outward radially and is available for interactions with other proteins (Neer et al. 1994; Sondek et al. 1996). For example, the WD40 repeats of the yeast TUP1 protein make direct contacts with the homeodomain protein alpha2 (Komachi et al. 1994; Komachi and Johnson 1997). The WD40 repeats of Groucho appear to be involved in making contact with Engrailed and Hairy as deletion of all of the repeats eliminates the interaction with these proteins (Jimenez et al. 1997). Additionally, the Groucho WD40 repeats were identified in a yeast two-hybrid screen as interacting with the eh1 domain of Engrailed (Tolkunova et al. 1998). However, the repeats alone are not sufficient to mediate a full interaction with either protein, so multiple regions may be involved in the interaction (Jimenez et al. 1997). Further evidence for the importance of the WD40 repeats comes from recent work in C. elegans in which a Groucho protein, UNC-37, was shown to interact genetically with the homeodomain protein UNC-4 to control neuronal development (Pflugrad et al. 1997). Several UNC-37 mutants have been sequenced, and most of the point mutations change specific amino acids within the WD40 repeats (Pflugrad et al. 1997). Interestingly, a chimera containing the amino terminus of UNC-37 and the WD40 repeats of human TLE1 is able to rescue the UNC-37 phenotype, which suggests that the structure and function of the WD40 repeats have been highly conserved among Groucho proteins (Pflugrad et al. 1997).

Relationship of the Groucho proteins to the yeast corepressor TUP1

The Groucho proteins have been compared with the
yeast corepressor TUP1 because of the presence of car
boxyl-terminal WD40 repeats in each protein and their
common function as transcriptional corepressors. How-
ever, it is unclear whether these proteins represent true
homologs as there are major structural and mechanistic
differences between them. First, although the Groucho
proteins and TUP1 both act as corepressors for members
of multiple families of transcription factors, many of the
DNA-binding partners for TUP1 do not directly interact
with TUP1, but instead utilize an accessory protein
known as CYC8 to form a stable complex (Keleher et al.
1992; Tzamarias and Struhl 1995). In contrast to the
Groucho proteins, which directly bind all currently stud-
ied partners, TUP1 directly binds only one DNA-binding
partner, the α2 protein (Komachi et al. 1994; Komachi
and Johnson 1997). Second, both TUP1 and the Groucho
proteins have intrinsic repressor activity, and for TUP1
this activity appears to be mediated both by direct inhi-
bition of the basal complex, perhaps via direct interac-
tion with the RNA polymerase II–associated Srb proteins
(Herschbach et al. 1994; Wahi and Johnson 1995; Carlson
1997; Redd et al. 1997), and by interaction with histones
H3 and H4 (Edmondson et al. 1996). The TUP1 interac-
tion with histones is mediated by a region that coincides
with the identified TUP1 repression domain (Tzamarias
and Struhl 1994; Edmondson et al. 1996). Whereas
Groucho proteins also interact with histone H3 (Pal-
parti et al. 1997), and it is quite possible that direct in-
teractions with histones are important for tran-
scriptional repression by both proteins, no interactions of
Groucho proteins with basal complex factors have yet
been demonstrated. Third, the arrangement of functional
domains and amino acid sequences of the proteins sug-
gest that there may be significant differences. TUP1 does
not conform to the general domain structure of the
Groucho proteins discussed earlier as the immediate
amino terminus of TUP1 is not required for tran-
scriptional repression, but is instead involved in making
contacts with the CYC8 accessory protein (Tzamarias
and Struhl 1994). At the amino acid level, both the repression
domains and WD40 repeats show poor sequence con-
servation, although it is possible that the three dimen-
sional structures will show greater similarity. With these simi-
larities and differences in mind, it may be more accurate
to consider TUP1 and Groucho proteins as analogous
rather than truly homologous.

Activation vs. repression of transcription
by Groucho partners

Interestingly, some of the DNA-binding partners for the
Groucho proteins do not always act as transcriptional
repressors, and, in fact, some are better characterized as
activators (Courey and Huang 1995; Speck and Terryl
1995). This is not unprecedented as other transcription
factors that repress transcription with a corepressor also
have been shown to activate transcription, including E2F
and nuclear receptors such as the retinoic acid receptor
(Sellers and Kaelin 1996; Heinzel et al. 1997). Among the
partners for the Groucho proteins, both Dorsal and the
Runt domain proteins are known to directly bind and
activate the transcription of specific target genes (Courey
and Huang 1995; Speck and Terryl 1995). For Dorsal and
possibly the Runt domain proteins, the context of the
target gene promoter appears to be critical for determin-
ing whether activation or repression will occur. Context
here refers to the location and occupancy of DNA-bind-
ing sites for other proteins in the immediate vicinity of
the binding site for a specific transcription factor. The
context-dependent activities of these Groucho partner
proteins suggest that the recruitment of Groucho pro-
teins or their repressor activity might be inhibited in
certain contexts. It also is possible that Groucho proteins
might even function as coactivators in certain situa-
tions.

The Drosophila Dorsal protein has been shown re-
cently to bind directly to Groucho and to exhibit
Groucho-dependent repression in vivo (Dubnicoff et al.
1997). Dorsal appears to repress transcription only when
its binding sites are located near binding sites for DNA-

binding proteins that have been termed corepressors (Fig.
3; Jiang et al. 1993; Kirov et al. 1993; Huang et al. 1995).
Both Dorsal and corepressor binding sites are present in
the promoters of the dorsal-specific genes, and these pro-
teins act together with Groucho to repress the transcrip-
tion of dorsal-specific genes in the ventral regions of the
embryo. In contrast, ventral-specific genes lack corepres-
sor sites and Dorsal activates transcription of these genes
suggesting that Dorsal functions intrinsically as an acti-
vator of transcription (Fig. 3). One possible mechanism
for dorsal-specific gene repression is that the corepres-
sors also directly bind to Groucho and thus stabilize a
Dorsal–Groucho complex on DNA, and/or that the co-
repressors interact with Dorsal to enhance its ability to
bind to Groucho (Fig. 3, double-headed arrows). Because
Groucho and Dorsal can bind to each other in vitro in the
absence of these DNA-binding corepressors (Dubnicoff
1997) it will be important to determine whether the
corepressors can potentiate these binding interactions.

The Runt domain proteins also both activate and re-
press target genes in a context-dependent manner. For
example, the Drosophila Runt protein represses hairy and
even-skipped expression, but activates the expres-
sion of fushi-tarazu and Sex- lethal (Duffy and Gergen
1994; Aronson et al. 1997). The mammalian Runt do-
main proteins have only been characterized as activators
of a large number of target genes in blood cells and bone
(Speck and Terryl 1995; Rodan and Harada 1997), al-
though the recent demonstration of repressor activity for
these proteins may lead to the identification of repressed
targets (Aronson et al. 1997). However, in contrast to
Dorsal, the Runt domain proteins appear to function in-
trinsically as transcriptional repressors, and to act only
as activators in a context-dependent manner (Fig. 4;
Speck and Terryl 1995; Aronson et al. 1997).

A well-studied example of transcriptional activation
by the mammalian Runt domain proteins involves the
TCRα enhancer, which requires for full activation the
simultaneous binding of the transcription factors CREB/
ATF, Ets, LEF-1, and the Runt domain protein AML-1 to
specific sites (Fig. 4; Giese et al. 1995). The context dependence of this enhancer comes from at least two sources: the DNA bend produced by binding of the architectural protein LEF-1 and context-dependent activation domains (CADs) present in AML-1 and LEF-1 (Carlsson et al. 1993; Giese and Grosschedl 1993; Giese et al. 1995; Bruhn et al. 1997). For these two activation domains, transcriptional activation is only observed when the proteins containing them are bound to the TCRκ enhancer at the correct location (Carlsson et al. 1993; Giese and Grosschedl 1993; Giese et al. 1995). In contrast to most activators, no activation by these proteins is seen from artificial promoters with multimerized binding sites for AM 1-1 or LEF-1 (Carlsson et al. 1993; Giese and Grosschedl 1993; Speck and Terryl 1995). Moreover, AML-1 has been shown to bind directly to Groucho and to function intrinsically as a transcriptional repressor when fused to the heterologous GAL4 DNA-binding domain and bound to multimerized DNA-binding sites in an artificial promoter (Fig. 4B; Aronson et al. 1997).

The context-dependent activation domains in LEF-1 and AML1 bind a coactivator known as ALY that appears to stabilize DNA binding and facilitate transcriptional activation (Fig. 4; Bruhn et al. 1997). Additionally, AML1 mutants lacking the WRPY motif that mediates interaction with the Groucho proteins still act as transcriptional activators, strongly suggesting that the Groucho proteins do not play a role in activation (Kurokawa et al. 1996). It will be interesting to determine whether, in contrast to Dorsal, the Runt domain proteins could intrinsically be repressors that are converted to activators through interactions with both coactivators and other transcription factors (Fig. 4). These interactions could either mask the WRPY motif, sterically block access of Groucho proteins to the Runt domain protein, or stabilize a conformation of the Runt domain protein that is unable to bind to Groucho proteins.

Thus, the Dorsal protein appears to be an intrinsic activator that can only repress transcription in the context of adjacent binding by specific cofactors. In contrast, the Runt domain protein AML-1 appears to be an intrinsic repressor that can only activate transcription in specific promoter contexts, perhaps in which the interaction with a Groucho protein is blocked by context-specific binding of cofactors. Additional studies will be needed to determine whether these are general properties for the various Runt domain and Rel domain proteins.

**Future directions**

Whereas significant progress has been made regarding
Groucho proteins and how they function, there are still many important questions that remain unanswered. First, the total number of families of DNA-binding partners for Groucho proteins is not known, and it is likely that there are more partners awaiting discovery. Second, the interaction of Groucho with transcription factors that both activate and repress transcription suggests that there must be regulation of the Groucho proteins at the level of either interaction with DNA-binding partners or their intrinsic repressor activity. A related question is whether Groucho proteins are regulated in more global ways, such as by cell signaling. There is preliminary evidence that cell signaling can regulate the Groucho proteins, as the phosphorylation status of several TLE proteins changes during the induction of neuronal differentiation of P19 cells by retinoic acid (Husain et al. 1996). Additionally, the Torso receptor tyrosine kinase inhibits the Groucho-dependent repression of tailless and hunchback expression in the anterior and posterior terminal regions of the Drosophila embryo. This regulation could occur either by inhibition of Groucho function or by inhibition of a yet to be characterized DNA-binding partner (Paroush et al. 1997). Third, an important mechanistic question is how the Groucho proteins repress both activated and basal transcription. The recent finding that Groucho proteins can interact with histone H3 may provide a mechanism, although by analogy with TUP1 there may be multiple mechanisms involved in repression, including direct interactions with the basal transcriptional machinery.

Acknowledgments

We thank Drs. Claude Desplan, James Jaynes, Michael Levine, and Anthony M.C. Brown for helpful comments and discussion about this review. A.F. was supported in part by the Tri-Institutional M.D.–Ph.D. program [National Institutes of Health (NIH) Medical Science Training Program (MSTP) grant GM07739] and funds from the Surdna Foundation. M.C. was supported by NIH grant R01 NS28652, the Alfred P. Sloan Foundation, the Pew Scholars in Biomedical Sciences Program, and the Cornell Scholars Program.

References


———. 1995. Distinct TPR motifs of Cyc8 are involved in recruiting the Cyc8-Tup1 corepressor complex to differentially regulated promoters. Genes & Dev. 9: 821–831.

Groucho proteins: transcriptional corepressors for specific subsets of DNA-binding transcription factors in vertebrates and invertebrates

Alfred L. Fisher and Michael Caudy

*Genes Dev.* 1998, 12:
Access the most recent version at doi:10.1101/gad.12.13.1931

**References**
This article cites 79 articles, 32 of which can be accessed free at:
http://genesdev.cshlp.org/content/12/13/1931.full.html#ref-list-1

**License**

**Email Alerting Service**
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.