# Chapter 1

# Introduction:

# Groucho – A Multifunctional Regulator of Drosophila Development

## Introduction

*Groucho/TLE family corepressors: structure and conservation*

The Groucho/TLE (Gro) family of of corepressors play crucial roles in the interpretation and integration of multiple signaling inputs during development in higher eukaryotes. Groucho, the sole *Drosophila melanogaster* member of this protein family, was first discovered in the context of a slight hypomorphic allele which resulted in the formation of extra supraorbital bristles reminiscent of the bushy eyebrows of Groucho Marx {Lindsley, 1968 #3055}. Subsequent research on Gro in *Drosophila* has served to characterize this factor’s central importance to developmental gene regulation in response to a variety of developmental programs and signaling pathways. As a corepressor, Groucho has no documented ability to bind directly DNA in a sequence-specific manner, instead relying on recruitment to genomic loci through interaction with a diverse array of transcriptional repressors. Groucho is essential to the correct patterning and development of *Drosophila* and is required for viability.

Homologs of Groucho with similar roles in developmental decision making have been identified throughout metazoans {Paroush, 1994 #3090}.

Groucho consists of five domains, two of which are highly conserved throughout higher eukaryotes {Turki-Judeh, 2012 #2385}. A great body of work has arisen documented the diverse contributions of each domain to the overall function and regulation of Groucho. While much of this work has focused on the N and C terminal domains, the central domains of Groucho have been explored for their roles in modulation of Groucho activity through interaction with a number of regulatory kinases and The N-terminal Q (glutamine rich) domain is one of the two conserved domains and is responsible for the formation of tetramers and potentially higher-order oligomers of Gro {Chen, 1998 #267}. Additionally, the Q-domain mediates a subset of interactions with transcriptional repressors, including the Tcf/Lef family of proteins {Brantjes, 2001 #3058}. Assays involving Grg3, a mouse homolog of Gro, on *in vitro* chromatin arrays showed that tetramerization mediated through the Q-domain is not required for recruitment of Gro to chromatin, but is required for subsequent aggregation of chromatinized fragments {Sekiya, 2007 #1658}. However, assays in cell culture revealed that oligomerization-deficient mutants of Gro exhibited similar median peak widths to wild-type Gro {Kaul, 2014 #2204}. The interpretation of this result is somewhat limited by the fact that ChIP-seq data was generated from two *Drosophila* cell lines depleted of endogenous Groucho via RNAi and overexpressing either GFP-tagged wild-type or oligomerization-deficient Groucho. The authors show that levels of endogenous Gro are significantly reduced, but remain detectable by immunoblot. Thus, it remains a possibility that low levels of endogenous Groucho are contributing to peak formation or spreading in both contexts.

Regardless of the role of oligomerization in the definition of the size of Groucho binding domains, the authors do detect large differences between the recruitment patterns of overexpressed wild-type and oligomerization-deficient mutants. Of the approximately 3000 distinct Groucho binding sites identified in the wild-type and mutant samples, 48% are unique to a single condition. Loss of oligomerization potential therefore, while preserving a large fraction of wild-type Groucho binding patterns, does disrupt Groucho association with chromatin in some contexts.

The structure of the Q-domain of TLE1, a human homologue of Gro, was recently solved, revealing the domain to form a dimer of dimers {Chodaparambil, 2014 #3057}. Though this explains the observation that Drosophila Gro forms a tetramer, the current model of association fails to account for the observation of higher-order oligomerization.

The WD-domain is the second conserved domain of Gro and comprises the C-terminus of the protein. The WD-domain consists of a seven-bladed β-propeller domain, and is responsible for the majority of Groucho interactions with DNA-binding repressors {Pickles, 2002 #3060}. The majority of these interactions are mediated through binding of the WD-domain to short peptide motifs {Jennings, 2006 #3059}, the most well-characterized of which are the Engrailed homology domain (Eh1) and WRPW motifs.

The central region of Groucho is divided into three domains, the GP, CcN, and SP domains. The GP domain binds to a histone deacetylase (HDAC1/Rpd3), which is involved with some but not all Groucho-repressive activity {Chen, 1999 #3061}. The CcN domain is involved in Groucho regulation, containing multiple Ck2 and Cdc2 phosphorylation sites {Nuthall, 2002 #3062}. The SP domain contains multiple sites phosphorylated in response to MAPK signaling, resulting in down-regulation of Groucho activity via nuclear export {Hasson, 2005 #3064}. There is evidence that the central regions of Groucho are intrinsically disordered {Turki-Judeh, 2012 #2966}, which has emerged as a common strategy among eukaryotic protein domains to facilitate participation in extensive protein-protein interactions, expose signaling motifs, and/or accept posttranslational modifications {Dunker, 2008 #3091}.

Groucho interacts with numerous transcriptional repressors, and through these interactions, is capable of participating in a number of developmental patterning determinations, as well as the reception and interpretation of multiple signaling pathways. *Drosophila* possesses one Groucho gene, which is maternally expressed and deposited into the embryo during oogenesis, ensuring Groucho availability and activity from the very onset of embryonic development {Paroush, 1994 #3090}. In vertebrates, Gro/TLE family proteins play similar roles in development. The genomes of humans and mouse both contain four Gro/TLE family proteins, TLE1-4 and Grg1-4, respectively {Stifani, 1992 #3065} {Mallo, 1993 #3066}. These Gro family members serve non-redundant roles during vertebrate development. An additional, truncated form of TLE, termed AES, is expressed in humans, where it associates with full-length TLE to down-regulate the cellular response to Wnt signaling {Itatani, 2015 #3067}. In humans, Gro/TLE is involved in such processes as organ development, adipogenesis, neurogenesis, hematopoiesis, and osteogenesis {Bajoghli, 2005 #3068} {Villanueva, 2011 #1659} {Javed, 2000 #3070} {Metzger, 2012 #2956}.

In *Drosophila,* Groucho’s roles in signaling pathway response are well documented. The factor plays a role in Ras/MAPK, Notch, Decapentapletic (dpp), and Wingless/Wnt signaling. Groucho activity is down-regulated via the Ras/MAPK pathway in response to signals initiated at EGFR, FGFR, and Torso receptors {Hasson, 2005 #3064}. This relief of Groucho-mediated repression is critical to the cellular response to RTK signaling and is thought to precipitate a cellular memory, where Groucho attenuation is thought to persist after loss of signaling {Cinnamon, 2008 #242} {Helman, 2011 #2938}.

Under the absence of Notch signaling, Groucho represses E(spl) complex genes through interactions with Hairy, which is itself associated with Su(H) {Delidakis, 1991 #3082}. Upon activation of Notch signaling, Notch displaces Hairy binding at Su(H) sites, relieving Groucho repression and initiating expression of E(spl) genes. Groucho then interacts with newly expressed E(spl) to repress a number of proneural genes {Preiss, 1988 #3083} {Wurmbach, 1999 #3084}.

Groucho is also critical to signaling via Decapentaplegic (dpp), a *Drosophila* TGF-β homolog whose diffusion over long distances is essential to patterning during embryogenesis and later during appendage development {Upadhyai, 2013 #2339}. The dpp morphogen is expressed dorsally in the embryo and is critical to the definition of cell-fate along the dorsal-ventral axis {Ferguson, 1992 #3088}. Groucho, through interaction with Dorsal, is essential in repressing ventral expression of *dpp*, meaning that Gro is involved in both the definition and interpretation of dpp signaling{Schwyter, 1995 #3038}*.* In the absence of dppsignaling, Brinker (brk) represses a subset of dpp target genes through two independent repressive mechanisms, one involving dCtBP (a short-range corepressor), and the other involving Gro {Hasson, 2001 #3033}. Upon activation of dpp signaling, Brinker becomes repressed by Schnurri in dorsal regions of the embryo, while continuing to be expressed in ventrolateral regions {Marty, 2000 #3089}.

Finally, Groucho participates in Wingless/Wnt signaling, through interactions with Tcf/Lef family proteins to regulate cell-fate choice {Cavallo, 1998 #3071} {Roose, 1999 #3086}. In unstimulated cells, Groucho assists in repressing Tcf/Lef target genes through interactions with the Q-domain {Clevers, 2006 #3085}. Upon Wnt activation, nuclear beta-catenin (Armadillo) concentration increases, which binds to Tcf, releasing Groucho and leading to gene activation. In this context, Groucho is essential in guarding against spurious activation of Wnt target genes in unstimulated cells {Daniels, 2005 #3087}.

It is primarily through the spatially and temporally precise mediation of gene transcription in response to these extracellular signals that Groucho becomes fundamental to embryonic patterning. Many early embryonic patterning effectors can be divided into dorsal-ventral and anterior-posterior programs, though these processes are complex and heavily interconnected {Jaeger, 2012 #3103}, requiring the coordinated regulation of dozens of transcriptional activators, repressors, and co-regulators {Mannervik, 2014 #2280}. Definition of the dorsal-ventral axis, which is critical to germ layer development, is the specified by the maternally-contributed gradient of nuclear Dorsal along this axis {Roth, 1989 #1112}. On the ventral side of the embryo, high concentrations of nuclear Dorsal initiate transcriptional programs designating the mesoderm {Gonzalez-Crespo, 1993 #3043}. In ventrolateral regions, modest Dorsal concentrations contribute to a neuroectodermal fate {Ip, 1992 #3042}. The strength, spacing, grouping geometry, and distribution of adjacent binding sites modulate Dorsal binding and cofactor recruitment in order to correctly interpret the Dorsal gradient {Zeitlinger, 2007 #3025}. Groucho is critical for repression of a subset of Dorsal-target genes, and is one method by which Dorsal is switched from an activator to a repressor {Dubnicoff, 1997 #2366}.

Groucho is also crucial to specification of the anterior-posterior axis. Early in development, the transcriptional groundwork is laid for the segmentation of the adult fly via multiple gradients, beginning with Bicoid and Nanos, which specify the expression of multiple gap genes, which in turn give rise to a striped pattern of pair rule genes, which then specify the expression of multiple segment polarity genes {Levine, 2008 #3104}. Groucho is critical to the viable specification of this axis through multiple direct interactions with transcription factors, including engrailed, a segment polarity gene.

In addition to dorsal-ventral and segmentation pattering, Groucho is involved in terminal patterning of the embryo through interaction with Capicua {Ajuria, 2011 #2947}, a process regulated by Ras/MAPK signaling {Chen, 2009 #3073} {Paroush, 1997 #3074}. Groucho activity is not limited to the embryo, as it participates in patterning of the imaginal discs during larval morphogenesis through the control of dpp signaling {Winkler, 2010 #2964}.

While a great deal is known about the developmental participation and interactors of Gro, the mechanism (or more likely, mechanisms) by which Gro achieves repression has remained elusive. Multiple models have been proposed to explain Groucho’s ability to fully and reversibly initiate and maintain short- and long-range repression, yet a full picture, able to account for all observations of Groucho behavior, has yet to emerge.

Much of the speculation surround Groucho activity centers on to what degree it forms oligomeric structures *in vivo*, how these structures interact with chromatin, and what relevance these structures have on its repressive abilities. Early evidence showed Groucho tetramerizes *in vitro* via the Q-domain {Chen, 1998 #267} {Song, 2004 #1161}. Groucho has been found to be associated with chromatin over 2 kb away from its recruitment site, leading to the hypothesis that Groucho spreads from its recruitment site, analogous to the spreading activity of Sir family corepressors {Pirrotta, 2005 #3106}. Experiments on a mouse Gro homolog showed that while tetramerization was not required for recruitment to chromatin, it is necessary for the aggregation of independent nucleosomal arrays in vitro {Sekiya, 2007 #1658}. Monomeric forms of the protein successfully bind to and increase the density of dinucleosomes *in vitro* {Sekiya, 2007 #1658}*.*  *In vivo,* the loss of tetramerization is not viable but does not entirely abolish Gro-mediated repression {Jennings, 2007 #2990}. More recent evidence in cell culture has shown that Gro binds in discrete peaks, though long stretches of binding do occur {Kaul, 2014 #2204}.

Gro preferentially associates with histone tails and can do so without the involvement of additional DNA-binding interacting factors {Flores-Saaib, 2000 #656}. Additionally, Gro associates with a histone deacetylase, HDAC1/Rpd3, {Chen, 1999 #3105}, however this association accounts for some but not all of Groucho’s repressive ability *in vivo* {Winkler, 2010 #2964}, indicating that other mechanisms, independent of both oligomerization and HDAC association are likely involved in Groucho-mediated repression. **Figure 1-1. Groucho interacts with numerous transcriptional co-regulators.** Groucho interacting partners are involved in diverse and interconnected aspects of *Drosophila* development, including anterior-posterior and dorsal-ventral patterning and the interpretation of extra-cellular signaling. These interactions are largely mediated through interactions between short peptide motifs and the Gro WD-domain. Some factors, notably Dorsal, require the participation of additional factors to facilitate Gro recruitment.

**Fig. 1-1**

|  |  |  |
| --- | --- | --- |
| **Interacting Protein** | **Biological Role** | **Citation** |
| Capicua | RTK signaling; embryonic terminal gene expression | {Jimenez, 2000 #3093} |
| Huckebein | Embryonic terminal gene expression | {Goldstein, 1999 #3094} |
| Hairy | Segmentation/ Anterior-posterior patterning | {Paroush, 1994 #3090} |
| Runt | Segmentation/ Anterior-posterior patterning | {Aronson, 1997 #3095} |
| Even-skipped | Segmentation/ Anterior-posterior patterning | {Kobayashi, 2001 #3076} |
| Odd-skipped | Segmentation/ Anterior-posterior patterning | {Goldstein, 2005 #3096} |
| Sloppy-paired 1 | Segmentation/ Anterior-posterior patterning | {Andrioli, 2004 #3097} |
| Engrailed | Segmentation/ Anterior-posterior patterning | {Jimenez, 1997 #3075} |
| Knirps | Segmentation/ Anterior-posterior patterning | {Payankaulam, 2009 #2955} |
| Goosecoid | Segmentation/ Anterior-posterior patterning | {Jimenez, 1999 #3092} |
| Dorsal | Dorsal-ventral patterning | {Dubnicoff, 1997 #2366} |
| Brinker | Dorsal-ventral patterning | {Zhang, 2001 #3099} |
| Ind | Dorsal-ventral patterning | {Von Ohlen, 2007 #3101} |
| Vnd | Dorsal-ventral patterning | {Cowden, 2003 #3102} |
| Su(H) | Notch signaling | {Barolo, 2002 #3072} |

## References