Groucho And Its Role In The Mouse Eye Development

BSc Thesis

Tomáš HADRAVA

Univerzita Karlova v Praze

Přírodovědecká fakulta

2006

Table of contents

Abbreviations

AES	amino-terminal enhancer of split
AML	acute myeloid leukemia
CcN	central region
DNA	deoxyribonucleic acid
eh-1	engrailed homology 1 motif
G_2/M	gap phase 2/mitosis
GH	here, glycine-histidine
GP	here, glycine-proline
Grg	Groucho mouse homologue
Grg1-S	short form of Groucho1
Grg3b	short form of Groucho3
Gro	Groucho
HDAC	histone deacetylase
Q	here, glutamine
QD	here, glutamine-aspartic acid
So	sine oculis
SP	here, serine-proline
TLE	transducin-like Enhancer of split
WD	here, tryptophane-aspartic acid

Introduction

The original *Drosophila* Groucho gene was identified in 1968. It caused a viable mutation resulting in a specific phenotype – the adult fly had clumps of extra bristles above the eyes, which was reminiscent of the bushy eyebrows of Groucho Marx, a popular comedian at that time (Lindsley and Grell, 1968).

Groucho is a member of the Enhancer of split gene complex. Gene products of this complex are components of the Notch signalling pathway and they were originally recognised as regulators of cell differentiation into epidermal and neural progenitor cells. The well-known phenomenon of lateral inhibition can be found among those signalling pathways incorporating both Notch and Groucho (see "Expression" section for more information).

Another well-known example is the Wnt pathway. Wnt proteins are secreted, locally acting signal molecules that control a lot of developmental processes. They were found in virtually all animals studied (Alberts et al., 2002). B-catenin, an important intermediate in this pathway, on activation migrates to the nucleus, binds to LEF-1/TCF molecule and by doing that acts as a co-activator. B-Catenin needs to displace Groucho though, which, under normal conditions, binds to LEF-1/TCF and acts as a co-repressor. Here, as part of a complex of regulatory proteins, it silences the downstream target genes of the whole Wnt pathway (Alberts et al., 2002).

As further research followed, it has become clear that apart from its involvement in lateral inhibition and its role in the Wnt pathway, Groucho has a lot more roles in development on top of that.

While most genes in the Enhancer of split complex code for basic helix-loop-helix domain-containing transcription factors (Chen and Courey, 2000), the carboxy-terminal region of the Groucho protein 40 displays sequence homology to a particular class of proteins that contain approximately 40 amino acid tandem repeats with characteristically located tryptophan and aspartate residues – the so-called WD-repeats (Neer, 1995). This repetitive motif was originally recognised in the β-subunits of the mammalian heterotrimeric G-proteins, which have a unique role in cell signaling, transducing a range of extracellular signals from the plasma membrane to the inside of the cell (Neer, 1995). G-proteins are transmembrane molecules, and because of the sequence similarity and also because of the role of Groucho in the Notch pathway, it was initially thought that Groucho was a cytoplasmic signal-transducing

molecule. However, subsequent experiments revealed that Groucho is exclusively localised in the nucleus (but see "The Role of Phosphorylation" section for more discussion). This raised the possibility that it functions as a transcription factor. Other evidence for this role came from a yeast two-hybrid screen (Paroush et al., 1994). In these studies, Groucho was found to interact with Hairy and Hairy-related proteins and Groucho was shown to be required for the transcriptional repression of genes that direct *Drosophila* segmentation, neurogenesis and determination of sex. These three processes are known to be regulated by Hairy family repressors. More experiments followed and it became clear that Groucho, although having no intrinsic DNA-binding domain, is recruited to promoters via direct interaction with a variety of sequence specific transcription factors. As it does not bind DNA directly, but is needed for the proper function of DNA-bound repressors, it is termed a co-repressor.

Numerous Groucho-homologous proteins have been found in a variety of eukaryotic organisms, most notably mouse *Mus musculus* and human (Stifani et al., 1992). The human genome encodes at least four members of this family – they are called TLE proteins, meaning transducin-like Enhancer of split (Stifani et al., 1992; Mallo et al.,1993). Gro/TLE family proteins play critical roles in a wide array of cellular and developmental pathways. Through interaction with various transcription factors, these co-repressors can modulate various developmentally important processes, such as, apart from the development of the eye, neural system development and patterning, hematopoiesis, myogenesis, bone formation and intestinal development (see "Other Processes Involving Groucho" section for more information).

A Note to the Reader

Please note that throughout this work, the overall term "Gro/TLE" protein or proteins is used for Groucho molecule as a whole, including homologues as well, as found in human and/or other organisms, with relation to mammals in most of the cases. The term "Groucho" is used when speaking about it in the widest sense, meaning it as a name of the particular gene or its product. When referring to a specific organism and its particular Groucho encoded protein, the fact is clearly stated. TLE refers to human homologue/homologues, Grg specifically to mouse counterparts. The distinction between a gene and a protein form of Groucho is emphasised only when its importance is required for the overall understanding.

Structure

In order to fully understand and appreciate the function of a particular protein, one has to have a look at its molecular structure and contents first.
Grg/TLE consist of five regions, these are (in the amino terminal – carboxy terminal end direction):

- glutamine rich region (Q domain)
- glycine/proline-rich region (GP domain)
- central region (CcN domain)
- serine/proline-rich region (SP domain)
- carboxy-terminal region with multiple tryptophane and aspartate tandem repeats (WD-repeat or WD40 domain, depending on an author) (Gasperowitz and Otto, 2005).

(As is usual, the actual domains are called according to the capital letters denominating the particular amino acids.)

The Q and WD-repeat domains are the most highly conserved, followed by a much less well-conserved central region, or CcN. The reason for this will become apparent after having read about the characteristic features of the particular domains; they are dealt with separately and in an extensive fashion in the following chapters.
On top of that, each subgroup includes the so-called short forms of Groucho – Grg5 in mouse and AES (amino-terminal enhancer of split) in humans. These are composed of Q and GP domains only (Mallo et al., 1993) (see Fig.1 in the "Attachments" section). The main part of the Q domain is highly conserved but the GP domain and carboxy-terminus of Q domain are slightly different from those of the long Grg/TLE proteins (see "The Role of Short Forms" section for more information).

A Note to the Reader

Please note that from now on and throughout this text, when saying "short forms" or "short members", namely Grg5 and AES will be meant, whereas the rest of the Grg and TLE members will be in this context and in a similar way called "long/full forms" or "long members".

Q Domain

The amino terminal Q domain derives its name from the rich glutamine contents, it is a highly conserved domain and it is likely to form coiled-coil structure (Chen et al., 1998). These putative coiled-coils apparently involve a pair of amphipathic α-helical motifs that are found in each Q domain (see Fig.2 in the "Attachments" section). These motifs show certain sequence homology to the leucine zipper motif, identified in many proteins that have coiled-coils in their structure (Chen et al., 1998).

One of the interesting features of both *Drosophila* Groucho and human TLE proteins is the fact that they form large oligomeric structures, both in vitro and in vivo (Chen et al., 1998). Groucho forms a tetramer composed of the same kind of a molecule, therefore referred to as a homotetramer. This homotetramerisation is mediated by the two putative amphipathic α-helices in the Q domain. Likewise, the Q domains in mouse and human have been shown to mediate both homo- and hetero-oligomerisation between Gro/TLE family proteins (Chen et al., 1998; Grbavec et al., 1998) (for information on the function of each domain also see Fig.3 in the "Attachments" section).

While the primary function of the Q domain may be to mediate homo- and hetero-oligomerisation, this domain may have additional functions, too. In some cases, Groucho has been shown to bind other proteins through this domain. This is the case with for example the human PRDI-binding factor, which is a sequence-specific transcription factor required for normal B-cell differentiation as well as the repression of β-interferon and c-myc gene expression (Chen and Courey, 2000). The mechanism of this repression is through recruitment of the human TLE proteins, suggesting that the Q domain may also, at least in some cases, represent a protein interaction motif for interactions between Groucho and DNA-bound repressors.

GP Domain

This domain comes next as we follow from the amino-terminal end to the carboxy-terminal one. It is one of the two poorly conserved, has a rich contents of the amino acids glycine and proline and it aids in transcriptional repression. The GP domain has an ability to recruit histone deacetylases to the DNA template, namely histone deacetylase 1 (HDAC-1) (Chen et

al., 1998; Grbavec et al., 1998), which is clearly connected to the ability of this domain to repress transcription.

CcN region

This central region of Groucho proteins is quite loosely conserved; it is approximately 60 amino acids long. A comparison of this particular region between the one from *Drosophila* Groucho and human TLE proteins shows that they share a cluster of four positively charged amino acids which is separated by a certain distance from serine. The serine residue can be phosphorylated by numerous kinases, most well known participants are casein kinase II and cdc2 kinase. Interestingly, it is not known whether uchon be phosphorylated by casein kinase II and/or cdc2.

In some proteins that have the similar CcN motif, nuclear localisation is controlled by the state of phosphorylation of this particular region. It is a confirmed fact that *Drosophila* Groucho and human TLE proteins are phosphorylated at serine residues in vivo (Husain et al., 1996). It is also known that a region outside the Q and GP domains mediates nuclear localisation of the full-length Groucho proteins (Chen et al., 1998; Chen and Courey, 2000). However, there are additional findings, especially in relation to the localisation of the short forms (see "The Role of Short Forms" section), showing that it cannot be definitely stated that the CcN region serves as and contains a nuclear localisation signal.

SP domain

This very poorly conserved region is rich in serine and proline and it is thought that it plays a role in transcriptional repression. The targets of this domain are not well characterised yet though. (Chen and Courey, 2000).

WD-repeat domain

In Groucho, the WD-repeat region is carboxy-terminal and, along with the Q-domain is the most highly conserved.

Generally speaking, WD-repeats have been found in proteins engaged in a variety of cellular processes, for instance signal transduction, pre-mRNA splicing, cytoskeleton assembly,

vesicular trafficking and, more importantly, transcriptional regulation (Chen and Courey, 2000). As follows from the definition of a repeat, this region is composed of repeating units of roughly the same composition. Each WD repeat unit then consists of a region of a variable length, followed by a conserved core bracketed by glycine-histidine (GH) and tryptophane-aspartate (WD). The length of this core is more or less constant. In symbols, the whole structure of the repeat can therefore be written as {X6-94 ----[GH---X23-41---WD]} (Chen and Courey, 2000), where X stands for any amino acid, GH and WD being defined above.

Most WD-repeat proteins contain four to eight repeats. It is to say that these repeats were originally recognised in β-subunits of heterotrimeric G-proteins in mammals – this was the feature that lead the first investigators into thinking that Groucho might be a transmembrane protein (see "Introduction" section). In G-proteins, this domain adopts a toroidal "β-propeller" structure (Chen and Courey, 2000, Neer et al., 1995) (see Fig.2 in the "Attachments" section). This radially symmetrical structure is made up of seven repeated "blades", each consisting of a four-stranded antiparallel β-sheet. In G-proteins, being transmembrane, these so-called β-blades in a number of seven are positioned next to each other and form a flat ring around a water-solvated channel. The three strands of each β-blade that are internal show very little variation in structure, whereas the outermost strand of each β-blade and also the loop connecting each adjacent blades are relatively more structurally variable. And what is more important, the β-propeller structure is not unique to WD-repeat-containing proteins, but it has been found in even a lot of non-WD-repeat proteins. Even though there is almost no apparent sequence similarity in these cases, the latter proteins adopt almost identical β-propeller folds (Neer and Smith, 1996). Therefore, as a conclusion, WD-repeat proteins, including members of the Gro/TLE family, seem to belong to a larger β-propeller superfamily.

As we have seen, WD-repeat-containing proteins have a large functional diversity. This can be explained by the fact that these domains serve as multifunctional protein-protein interaction sites and they are capable of binding different sets of partner proteins. The particular set of proteins with which any given WD-repeat domain interacts, as well as the other functional domains contained within the polypeptide, determine the functional roles of the protein. As we apply this rule to the protein of interest here, it can be noted that WD-repeat domain in Gro/TLE co-repressor often seem to be involved in making contacts with DNA-bound repressors. To present some examples, the WD-repeat domain of *Drosophila* Groucho is required to direct interactions with the DNA-binding proteins Engrailed and Hairy

(Jimenez et al., 1997). In human, the same region contributes to the interaction with AML1 and AML3 transcription factors (Chen and Courey, 2000).

There are experiments showing that the structure and function of WD-repeat domains are conserved throughout the Gro/TLE family. To present just one, the WD-repeat domain of human TLE1 can functionally substitute the analogous domain of a Gro-homologue UNC-37 belonging to *Caenorhabditis elegans* in vivo (Chen and Courey, 2000).

Expression

The expression patterns of individual Gro/TLE family proteins during mouse embryonic development are consistent with roles in segmentation, central and peripheral neurogenesis, as well as epithelial differentiation (Koop et al, 1996). Also, TLE proteins are temporarily and spatially co-expressed with the Notch protein as well as other members of the Notch family. This is consistent with possible roles in multiple cell fate decisions governed by this pathway (Husain et al., 1996; Chen and Courey, 2000). Moreover, both Notch and TLE genes are normally expressed in proliferating immature epithelial cells, while their expression begins or decreases in terminally differentiated epithelial cells (Liu et al., 1996). Cells derived from incorrect or incomplete epithelial differentiation exhibit persistent expression of Notch and TLE, which is consistent with the fact that these gene products play a common role in preventing or delaying the differentiation of epithelial cells.

The mammalian Gro/TLE homologues found in any given organism are encoded by different genes that are either located on different chromosomes or can be found clustered in a tandem array on the same chromosome (Mallo et al., 1994). TLE genes were found to be expressed in distinct, non-overlapping patterns during the neural differentiation of mouse embryonic carcinoma cells as well as during mouse embryogenesis. (Koop et al., 1996). Together, these findings imply that individual mammalian Groucho-homologues play non-redundant and differentiated roles in diverse biological processes.

The Role of Short Forms

In addition to encoding full-length Gro/TLE family proteins, mouse and human genomes encode truncated proteins, members of the family that we have come to call "short forms". These are namely:

In mouse:

- Grg5 (Groucho5)
- Grg1-S (a short form of Groucho1)
- Grg3b (a short form of Groucho3)

In human:

- AES (amino-terminal enhancer of split)
- QD (a shortened version of TLE4)

Grg5 and AES are comprised of only Q and GP domains (Mallo et al., 1993). In Grg5 the GP domain differs from GP domains of Grg1 through 4. The other short forms result from alternative splicing of long Grg/TLE mRNA. Grg1-S is a short form of Grg1 and consists of the Q domain and a large part of the GP domain with an additional sequence not translated in the full form of Grg1 (Lepourcelet and Shivdasani, 2002). Grg3b is a short form of Grg3 and consists of the Q and GP domains (Gasperowicz and Otto, 2005). On top of that, a shortened version of TLE4 has been reported, it consists of only Q domain and that is why it was named QD (Gasperowicz and Otto, 2005). While long Grg/TLE proteins function exclusively as co-repressors, the role of short forms is somewhat controversial.

It is believed that Grg5 and AES act as dominant negative form of long Grg/TLE co-repressors. What this basically means is that they can cause the inability to repress transcription when they have a chance of interacting with the Groucho full forms. This was observed with Tcf, when Grg5 de-represses transcriptional activation mediated by Tcf (Gasperowicz and Otto, 2005). Grg5 also reduces Grg4-mediated enhancement of Nkx-dependent repression in vitro (Gasperowicz and Otto, 2005). Likewise, the QD protein inhibits TLE4-Pax5 binding, to name just a few examples. Moreover, mice nullizygous for Grg5 are perfectly viable and show only postnatal growth deficiencies (Mallo et al., 1993).

The phenomenon of de-repression mentioned above is usually explained through the inhibition of the long forms through Q domain binding and there are a couple of reports that are in favour of this concept. In this model, the Q domain of the particular short form binds

other Grg/TLE proteins and by doing this, it inhibits their ability to form functional di/tetramers, or to bind to other transcription factors. In this model, the Q domain binds other Grg/TLE proteins and the lack of a GP domain in case of the QD protein which was shown in Groucho to be essential for interaction with chromatin modifying complexes cause an inability of the multimer to interact with proteins responsible for chromatin architecture and could therefore cause inability to repress transcription (Chen et al., 1999). In Grg5 the GP domain differs from GP domains of Grg1 through 4 and this difference could result in a lack of repressive interaction with chromatin structures yet again (Gasperowicz and Otto, 2005).

Interestingly, the truncated form of Grg1 - Grg1-S - acts as a co-repressor, in a similar fashion as the long forms do. This could be explained by the fact that it contains a GP domain that is almost identical to that of Grg1 and therefore able to mediate the interaction with chromatin components in the same way as the GP domain of Grg1 does.

The situation is a little bit more complicated in the case of AES. Its GP domain differs from that of long Grg/TLE proteins in the same way as the one of Grg5, but it was shown to act as a co-repressor. AES does not interact with histone deacetylases HDAC-1 and HDAC-3 (Gasperowicz and Otto, 2005). Some authors conclude, though, that modifications within the GP domain of this protein cause it to interact with HDACs other than HDAC-1 and HDAC-3 or even with other factors influencing the structure of chromatin, for example with the members of basal transcriptional machinery.

As one can gather from the published work that has to do with the role of the short members of the Grg/TLE family, the exact mechanism of how these influence Grg/TLE-mediated repression and their relevance in this scheme still remains to be resolved.

Repression

A lot of co-repressor complexes, for instance the Sin3 complex, contain histone deacetylases (HDACs) that remove acetyl groups from lysine residues in the amino-terminal tails of core histones (Chen and Courey, 2000). In a similar fashion, co-activator complexes often include histone acetyltransferase activity. By determining the acetylation state of histones, these enzymes presumably modulate local chromatin structure and alter the ability of the general transcription machinery to recognise and transcribe genes. It is known that hyper-acetylated chromatin (that is chromatin with a large portion of acetylated histones) is usually associated with active transcriptional states. On the other hand, hypo-acetylated chromatin (which is just the opposite, chromatin containing relatively few acetylated histones) means that the particular area is under the repressed transcriptional state of activity.

The *Drosophila* histone deacetylase Rpd3 has recently been identified as a Groucho-interacting protein. Through a range of experiments, it has been shown that Rpd3 and Groucho form a complex in vivo; a notable fact is that they interact directly through the GP domain (see Structure). Out of these experiments there come strong suggestions that histone deacetylation contributes to Groucho-mediated repression. Analysis of the effects of Groucho and Rpd3 mutations on *Drosophila* embryogenesis proves that Groucho and Rpd3 interact in development (Chen et al., 1999).
Based on all the findings mentioned above, the following model for Groucho-mediated transcriptional repression has been proposed. First, Groucho is recruited to the template via a direct interaction with sequence-specific DNA-binding transcription factors. After that, the ability of Groucho to oligomerise together with the favourable interactions between Groucho and the histones results in the building of a polymer of Groucho, which spreads along the template. This template-bound Groucho-polymer then provides a high-affinity interface for the recruitment of key chromatin-modifying factors, including the histone deacetylase Rpd3, given as an example above. These components may subsequently serve to establish and/or maintain a silenced chromatin structure, in fact creating a large transcriptionally silent chromosomal domain (Choi et al., 2005). The interaction between Groucho and histones may be enhanced by histone deacetylation and therefore, it would be possible that histone deacetylation by for example Rpd3 facilitates the recruitment of Groucho to the DNA

template and thus the spread of Groucho along the chromatin fibre, thus even re-enforcing the repressed transcriptional state (see Fig.4 in the "Attachments" section).
This idea of a co-repressor having to polymerise along the template is not entirely new. A precedent of some sort and an older example may be provided by the yeast Sir3 and Sir4 proteins (Chen and Courey, 2000). Sir3 and Sir4 are required for silencing transcription at the chromosomal regions adjacent to telomeres. After recruitment to the DNA, Sir3 and Sir4 are believed to spread along the template into adjacent chromatin regions to induce heterochromatin formation. This inactivates the particular loci and genes near telomeres.

But, to make the described status quo a little bit more complicated, apart from the Rpd3-interacting GP domain, Groucho contains even other repression domains that appear to act in an Rpd3-independent manner. As more research needs to be done on this particular area, one can speculate – these domains may for example function to recruit other general transcription factors and by doing that they interfere with the assembly or activity of the RNA polymerase II machinery.
Similarly, as new research is continually being published, there has only recently been a new set of data suggesting a new mechanism of Groucho-mediated suppression (see "Pax-family" section).

The Role of Phosphorylation

As was mentioned in the Structure section, the Groucho protein has two domains that contain possible sites of phosphorylation – these are the CcN and SP regions which can be phosphorylated by $p34^{cdc2}$ kinase, for instance (Nuthall et al., 2002). This kinase functions as the master regulator of the G_2/M transition and entry into mitosis (Nuthall et al., 2002). Conserved motifs resembling phosphorylation sites for this particular kinase can be found within all Gro/TLE family members. It is not only $p34^{cdc2}$ kinase that can be involved though, these proteins contain evolutionary conserved consensus phosphorylation sites for a number of kinases (Stifani et al., 1992).

During the G_2/M phase of the cell cycle, Gro/TLE proteins become phosphorylated (Nuthall et al., 2002), and this is correlated with a decreased nuclear interaction. More specifically, these phosphorylation events may negatively regulate the ability of the Gro/TLE proteins to interact with the chromatin components and components of the nuclear matrix, thereby playing a negative regulatory role in Gro/TLE functions, as reduced interaction with the nuclear compartment is expected to negatively affect the ability of these proteins to repress transcription (Nuthall et al., 2002).

Finally, the transcription repression ability of Gro/TLE is enhanced by pharmacological inhibition of $p34^{cdc2}$ kinase. Putting these facts together, it is clear that Gro/TLE proteins are phosphorylated as a function of the cell cycle and it is very likely that phosphorylation events do occur during mitosis, which eventually leads to negative regulation of Gro/TLE activity. On the other hand, inhibition of such events in fact enhances Gro/TLE-mediated trancsriptional repression. As with the mechanism of repression itself, some similar cases of negative regulation through phosphorylation can be found (Nuthall et al., 2002). Generally speaking, given the ability of Gro/TLE proteins to form complexes with other DNA-binding proteins, the regulation of the nuclear interaction described above may represent a general mechanism to control the functions of several transcription factors in mitotic cells.

To make matters slightly more complicated, a number of studies exist showing the influence of hyper-phosphorylation, which results in the effects opposite those mentioned above. This process leading to an increase in phosphorylation, consisting of a series of events, is connected with strong association with the nuclear compartment, presumably through the interaction with chromatin.

The two different processes and outcomes described above and the fact that each Gro/TLE protein possesses multiple potential sites of phosphorylation demonstrate just how complex the regulation of Gro/TLE activity can be. It shows the importance of the effort to find out more about these mechanisms for a complete understanding of the action of Gro/TLE co-repressors.

The Involvement in Eye Development

In order to fully understand the importance of Groucho in the mouse eye development and having accepted the fact that this protein functions as a co-repressor - meaning it needs other molecules to interact with - a few other genes and gene products need to be introduced in this chapter. These will be Six-family and Pax-family members.

Six-family - Six3 and Six6

Six3 and Six6 are the only members of the Six-gene family that are expressed in the early stages of the development of visual system (López-Ríos et al., 2003). They are included in the same gene group as the *Drosophila* optix gene, the mouse Six3 gene being originally isolated on the basis of its homology with the *Drosophila* sine oculis (So) gene. Members of the So/Six gene family encode proteins that have a conserved Six-domain and a homeodomain. The Six-domain is involved in both DNA and protein binding, the homeodomain is responsible for DNA binding only (López-Ríos et al., 2003). There are only minor differences between the two domains in Six3 and Six6, they are nearly identical – the amino-terminal portion is longer in Six3 and includes a glycine-rich region of unknown function, this is completely absent in Six6. The carboxy-terminus is the most divergent domain except for the last 15 amino acids that are nearly identical (López-Ríos et al., 2003). Therefore, in spite of their strong homology, Six3 and Six6 behave in a different way in regards to their interaction with other proteins. To be more specific, in the case of Gro/TLE interaction, the Six3/TLE1 complex might be favoured and more effective than its counterpart, Six6/TLE1 complex, because the former is mediated by an additional binding site (Eberhard et al., 2000).
Six3 and Six6 are expressed in the anterior neural plate in an overlapping fashion, a little bit more restricted for Six6. Up to date, six members (Six1-Six6) of this family have been identified in mammals. Through interaction with Groucho proteins, Six3 and Six6 can become strong transcriptional repressors.

Several experiments using transgenic mice show the importance of these two genes in eye formation (Kobayashi et al., 2001; Gasperowicz and Otto, 2005). While gain-of-function studies show the capability of Six3 and Six6 to control eye field growth and to enhance it and

in most cases increase the size of it, loss-of-function analysis and their specific expression pattern suggest that their function may have separated and diversified.

Groucho family members were shown to be co-expressed, interact with and modulate the activity of Six3/Six6 transcription factors in the developing eye. The expression patterns of Grg4 and Grg5 in mouse embryos are similar to the one of Six3, namely the expression of Grg5 overlaps that of Six3; all are expressed in this way in a developing optic vesicle (Zhu et al., 2002). Grg3 has a similar pattern of expression as Six3, this co-localisation takes place in the neuroblastic layer of retina in the embryonic day 16.5 (Zhu et al., 2002). Grg4 and Grg5 do interact physically with Six3 and Six6; in humans, TLE1 and AES interact with SIX3 and SIX6 (being the human homologues of Six3 and Six6 respectively (Zhu et al., 2002; López-Ríos et al., 2003). All of these interactions take place through the Q domain of Gro/TLE proteins and eh1-like motif that can be found in the highly conserved Six-domain of the Six3/Six6 proteins (Zhu et al., 2002; López-Ríos et al., 2003). On top of that, in humans, SIX3 interacts with TLE proteins via the WD domain (López-Ríos et al., 2003).

The interaction with Gro/TLE proteins is important for Six3-mediated repression in vitro and in vivo (Zhu et al., 2002; López-Ríos et al., 2003). Interaction of Six3 with Groucho proteins was shown to be relevant for the formation and differentiation of photoreceptors in the developing rat (*Rattus norvegicus)* retina (Zhu et al., 2002; López-Ríos et al., 2003). SIX3 and SIX6 with TLE1 can in their own expand the eye field in medaka fish *Oryzias latipes* (López-Ríos et al., 2003). Conversely, AES alone decreases the eye size and abrogates the phenotypic consequences of SIX3/SIX6 over-expression (López-Ríos et al., 2003).

The proposed models for TLE/AES modulation of Six3/Six6 transcriptional activities are quite worth-looking at.

1) Six3 and Six6 could bind to distinct DNA binding sites and followingly, their interaction with TLE1 or AES would lead to transcriptional repression or activation, respectively.
2) Six3 and Six6 could bind to a larger transcriptional repression complex, the sometimes so called “repressosome” (Courey and Jia, 2001). Gro/TLE proteins and possibly additional factors could be part of that complex. AES presence as part of the complex would then provide a mechanism of de-repression. In support of this model, Six3 can contact other nuclear factors that are involved in chromatin remodelling.
3) The Six3/TLE1 complex could act as a single transcriptional repression unit, which could be regulated by a dominant negative complex of Six6/AES. This model presents

a specific function for each of the Six proteins in question and fits in well with the current data on expression, gain- and loss-of-function of the two molecules (López-Ríos et al., 2003), while taking into consideration the differential interactions of TLE and AES with Six3 and Six6.

It is important to note though, that for other Six proteins involved in eye development, besides Six3 and Six6, no interaction with Gro/TLE family proteins has been reported, or is thought to be weak and therefore not easily detected. This applies despite the similarities in the eh-1-like motif identified in the Six-family members (Zhu et al., 2002).

All of these facts point to a crucial role of Gro/TLE co-repressors in a developing visual system. Also, the reports about the involvement of Gro/TLE-mediated Six3/Six6 regulation in eye development in other vertebrates, such as chick *Gallus gallus* (lens morphogenesis and crystallin regulation), medaka fish *Oryzias latipes* (retina development) and zebrafish *Danio rerio* (eye and forebrain formation) suggest the existence of similar mechanisms in mammals (Zhu et al., 2002; López-Ríos et al., 2003).

Pax-family - Pax2

The Pax proteins form a family of transcription factors containing the so-called paired domain. They play important roles in early development in a range of organisms, from *Drosophila* through mouse to human. Based on sequence similarities, the Pax genes can be divided into four subfamilies (Noll, 1993). The one that is of interest here consists of a single member found in *Drosophila* (dPax258) and three genes found in mammals (Pax2, Pax5 and Pax8) (Eberhard et al., 2000).

Members of this subgroup encode DNA-binding proteins that have both repression and activation domains. The Pax2/5/8 activation domain is located at the carboxy-terminus and is rich in serine, threonine and proline residues (Cai et al., 2003). The ability to activate transcription is enhanced by phosphorylation of its activation domain by the c-Jun amino-terminal kinase (Cai et al., 2003).

All of the Pax genes and their products are essential developmental regulators in a wide variety of tissues. In mouse (as well as humans), the Pax2 gene is needed for the

differentiation of epithelial cells in the urogenital tract, hindbrain and inner ear patterning and, more importantly, the eye development.

Expression of Grg1, Grg3 and Grg4 to a certain extent overlaps with that of Pax2/5/8 family in the neural tube. There are some studies showing that in the case of Pax genes activity, yet another example of Groucho action can be deciphered. As is becoming clear, it is not only the deacetylation of histones through the potential interaction with HDACs that brings about Groucho-mediated repression as such, but the specific inhibition of phosphorylation of the activation domain of a DNA-bound protein that causes the suppression, in this case it is Pax2 activation domain and also the protein itself bound to DNA. This has been found out to apply for Grg4, namely, which can completely suppress Pax2-dependent activation of a reporter gene, even when c-Jun amino-terminal kinase is present. This ability of Grg4 to block Pax2 phosphorylation requires binding of Pax2 to a DNA target sequence (Cai et al., 2003).

Other Processes Involving Groucho

Groucho not only plays its role as a co-repressor in eye development, but it can be seen as a participant in numerous developmental settings as well. In order to give the reader an idea of how complex and varied the interactions of Groucho can be, a list of other processes apart from the eye development is given, although having omitted the details, as a more in-depth coverage of these topics is not the subject of this work. As "Grg" indicates, all of them are mouse homologues (Gasperowicz and Otto, 2005).

Grg3, Grg3b, Grg5	Osteogenesis
Grg1, Grg3, Grg4, Grg5	Hematopoiesis
Grg1	Pituitary gland development
Grg4	B-cell development
Grg4	Neural tube patterning

Possible Approaches In the Future

As we have seen, Groucho as such has been known to science for a certain period of time. Studies on the structure of Groucho and the presence of all the various homologues found in different organisms (namely the mouse) date back to 1993. Since then, there has been a great amount of work carried out on this subject. However, there are areas that are still not very well covered and in need of some attention. Among those, I would include the following:

- The overall expression patterns of the different forms of Groucho, especially in mouse and with special relation to the eye, which is of interest here.
- With relation to the previous point, the effort to establish if there is redundancy or specificity, in other words if all the different forms and splicing variants of Groucho have their specific function or if these can be sometimes interchanged.
- The exact mechanisms of Groucho-mediated repression, which is still not entirely clear – there are more repression domains present than was originally thought and there is a new emerging possible mode, as described in the case of Pax2
- The role and action of the Groucho short forms, which seem to be vital in development, with their ability to somehow oppose and counteract the action of the long forms.

Thus, further studies of the Gro/TLE family of proteins will not only provide insights into mechanisms of repression but will also show the importance of transcriptional repression in diverse biological processes, generally speaking, also showing the significance of basic research, as opposed to applied research.

What I would like to do personally is to, at first, look at the expression patterns of a chosen array of Groucho in mouse eye, try to establish the overall expression using histological immuno-staining. I think that, despite the fact that Groucho acts as a co-repressor, it needs to be looked at on its own. There is a huge amount of work and articles carried out on Groucho, but in majority, they all use Groucho as something that works with the particular author´s transcription factor of interest. I would really like to have Groucho as the molecule of my primary attention and my subject of choice.

Followingly, I would like to use appropriate mouse strains/transgenes for further studies, for instance and particularly BAT-gal transgenic mice, which seem to be best suited for this purpose.

Acknowledgments

I would like to thank Zbyněk Kozmík for his suggestions and help, Lenka Libusová, for her support and hot water for tea, my parents for their understanding (the fact that they have not seen me for a while whatsoever) and dearest Miss P. (“Hi!”) for being just who she is, really really.

References

Alberts,B., Johnson,A., Lewis,J., Raff,M., Roberts,K., Walter,P. (2002). Molecular Biology of The Cell, 4th edition. Garland Science, New York.

Cai,Y., Brophy,P.D., Levitan,I., Stifani,S., Dressler,G.R. (2003). Groucho suppreses Pax2 transactivation by inhibition of JNK-mediated phosphorylation. The EMBO Journal 22, 5522-5529.

Chen,G., Nguyen,P.H., Courey,A.J. (1998). A role for Groucho tetramerization in transcriptional repression. Mol. Cell. Biol. 18, 7259-7268.

Chen,G., Fernandez,J., Mische,S., Courey,A.J. (1999). A functional interaction between the histone deacetylase Rpd3 and the corepressor Groucho in *Drosophila* development. Genes and Development 13, 2218-2230.

Chen,G., Courey,A.J. (2000). Groucho/TLE family proteins and transcriptional repression. Gene 249, 1-16.

Choi,C.Y., Kim,Y.H., Kim,Y.O., Park,S.J., Kim,E.A., Riemenschneider,M., Gajewski,K., Schulz,R.A., Kim,Y. (2005). Phosphorylation by the DHIPK2 Protein Kinase Modulates the Corepressor Activity of Groucho. J. Biol. Chem., 280 (22), 21427-21436.

Courey,A.J., Jia,S. (2001). Transcriptional repression: the long and the short of it. Genes Dev. 15, 2786-2796.

Eberhard,D., Jiménez,G., Heavey,B., Busslinger,M. (2000). Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. The EMBO Journal 19, 2292-2303.

Gasperowitz,M., Otto,F. (2005). Mammalian Groucho Homologs: Redundancy or Specificity? J.of Cellular Bicohemistry 95, 670-687.

Goldstein,R.E., Cook,O., Dinur,T., Pisante,A., Karandikar,U.C., Bidwai,A., Paroush,Z. (2005). An eh1-like motif in odd-skipped mediates recruitment of Groucho and repression in vivo. Mol.Cell.Biol. 25(24), 10711-10720.

Grbavec,D., Lo,R., Liu,Y., Stifani,S. (1998). Transducin-like Enhancer of split 2, a mammalian homologue of *Drosophila* Groucho, acts as a transcriptional repressor, interacts with Hairy/Enhancer of split proteins and is expressed during neuronal development. Eur. J. Biochem. 258, 339-349.

Husain,J., Lo, R. Grbavec,D., Stifani,S. (1996). Affinity for the nuclear compartment and expression during cell differentiation implicate phosphorylated Groucho/TLE1 forms of higher molecular mass in nuclear functions. Biochem. J. 317, 523-531.

Jimenez,G., Paroush,Z., Ish-Horowicz,D. (1997). Groucho acts as a corepressor for a subset of negative regulators, including Hairy and Engrailed. Genes Dev. 11, 3072-3081.

Kobayashi,M., Nishikawa,K., Suzuki,T., Yamamoto, M. (2001). The homeobox protein Six3 interacts with the Groucho corepressor and acts as a transcriptional repressor in eye and forebrain formation. Dev.Biology 232, 315-326.

Koop,K.E., MacDonald,L.M., Lobe,C.G. (1996). Transcripts of Grg4, a murine groucho-related gene, are detected in adjacent tissues to other murine neurogenic gene homologues during embryonic development. Mech. Dev. 59, 73-87.

Lepourcelet,M., Shivdasani,R.A. (2002). Characterization of a novel mammalian Groucho isoform and its role in transcriptional regulation. J. Biol. Chem. 277, 47732-47741.

Lindsley,E.B., Grell,E.H. (1968). Genetic variation of *Drosophila melanogaster* – Internet source.

Liu,Y., Dehni,G., Purcell,K.J., Sokolow,J., Carcangiu,M.L., Artavis-Tsakonas,S., Stifani,S. (1996). Epithelial expression and chromosomal location of human TLE genes, implications of notch signalling and neoplasia. Genomics 31, 58-64.

López-Ríos,J., Tessmar,K., Loosli,F., Wittbrodt,J., Bovolenta,P. (2003). Six3 and Six6 activity is modulated by members of the groucho family. Development 130, 185-195.

Mallo,M., Franco del Amo,F., Gridley,T. (1993). Cloning and developmental expression of Grg, a mouse gene related to the groucho transcript of the *Drosophila* Enhancer of split complex. Mech. Dev. 42, 67-76.

Mallo,M., Steingrimsson,E., Copeland,N.G., Jenkins,N.A., Gridley,T. (1994). Genomic oragnization, alternative polyadenylation and chromosomal localization of Grg, a mouse gene related to the groucho transcript of the *Drosophila* Enhancer of split complex. Genomics 21, 194-202.

Neer,E.J. (1995), Heterotrimeric G proteins, organizers of transmembrane signals. Cell 84, 175-178.

Neer,E.J., Smith,T.F. (1996). G protein heterodimers, new structures propel new questions. Cell 84, 175-178.

Noll,M. (1993). Evolution and role of Pax genes. Curr. Opin. Genet. Dev. 3, 595-605.

Nuthall,H.N., Joachim,K., Palaparti,A., Stifani,S. (2002). A role for cell cycle-regulated phosphorylation in Groucho-mediated transcriptional repression. J. Biol. Chem. 277(52), 51049-51057.

Paroush,Z., Finley Jr,R., Kidd,T., Wainwright,S.M., Ingham,P.W., Brent,R., Ish-Horowitz,D. (1994). Groucho is required for *Drosophila* neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. Cell 79, 805-815.

Song,H., Passon,P., Paroush,Z., Courey,A.J. (2004). Groucho oligomerization is required for repression in vivo. Mol.Cell.Bio. 24(10), 4341-4350.

Stifani,S., Blaumueller,C.M., Redhead,N.J., Hill,R.E., Artavanis-Tsakonas, S. (1992). Human homologs of a *Drosophila* Enhancer of split gene product define a novel family of nuclear proteins. Nat. Genet. 2(4), 119-127.

Zhu,C., Dyer,M.A., Uchikawa,M., Kondoh,H., Lagutin,O.V., Oliver,G. (2002). Six3-mediated auto repression and eye development requires its interaction with members of the Groucho-related family of co-repressors. Development 129, 2835-2849.

Attachments

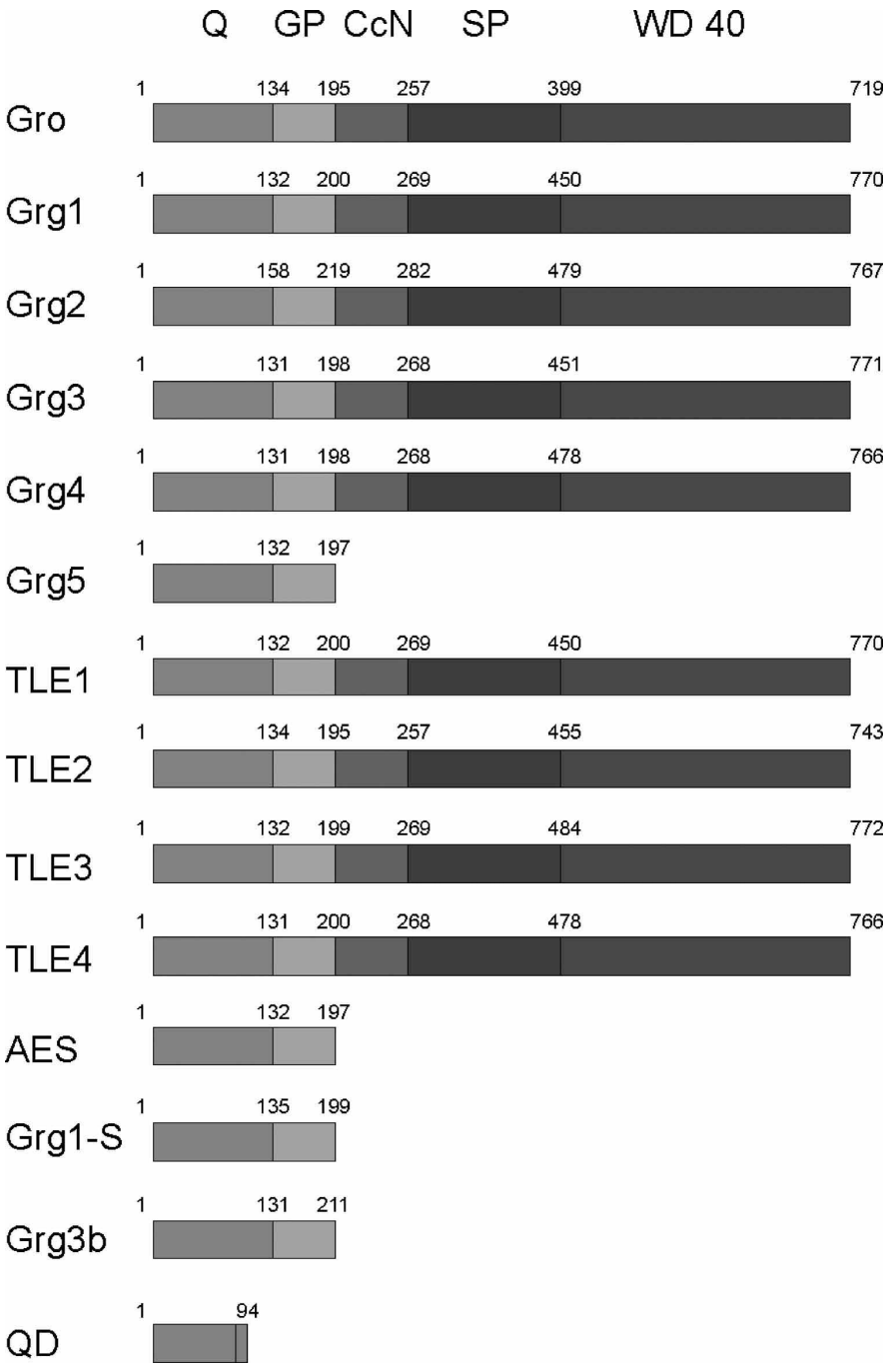

Fig.1

Schematic overview of the structure of Gro/TLE protein family members. Gro – a *Drosophila melanogaster* member, Grg – forms found in mouse *Mus musculus,* TLE – human homologues. Grg5, Grg1-S, Grg3b, AES, QD – short splice variants of Grg/TLE proteins. For domain names and other details see text. Numbers stand for amino acids (taken from Gasperowicz and Otto, 2005).

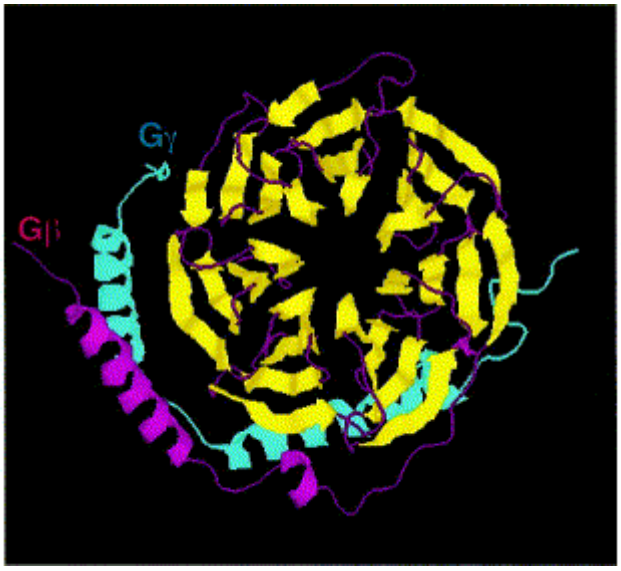

Fig.2

Ribbon diagram of the crystal structure of seven WD-repeats adopting a toroidal „β-propeller" structure. Letters indicate a $G_{\beta\gamma}$ heterodimer. (taken from Chen and Courey, 2000).

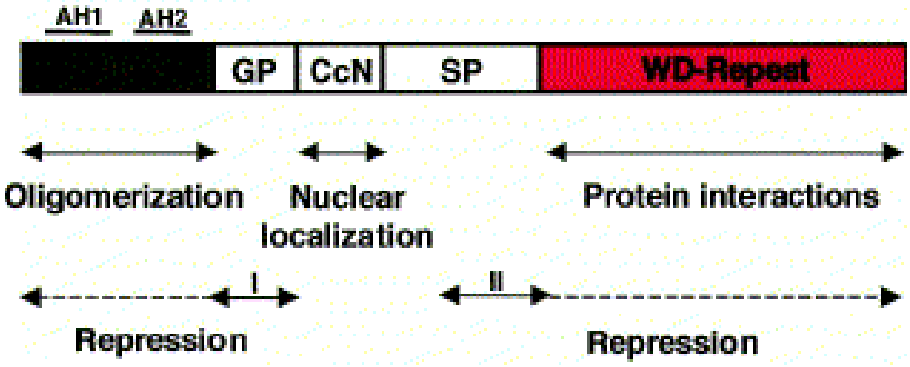

Fig.3

Structural representation of Gro/TLE domains with regards to function. Conserved regions are in colour, more flexible regions depicted by open boxes. A pair of putative amphipathic α-helical motifs is indicated by two lines with labels at the top of the Q-domain. For domain names and other details see text. (taken from Chen and Courey, 2000).

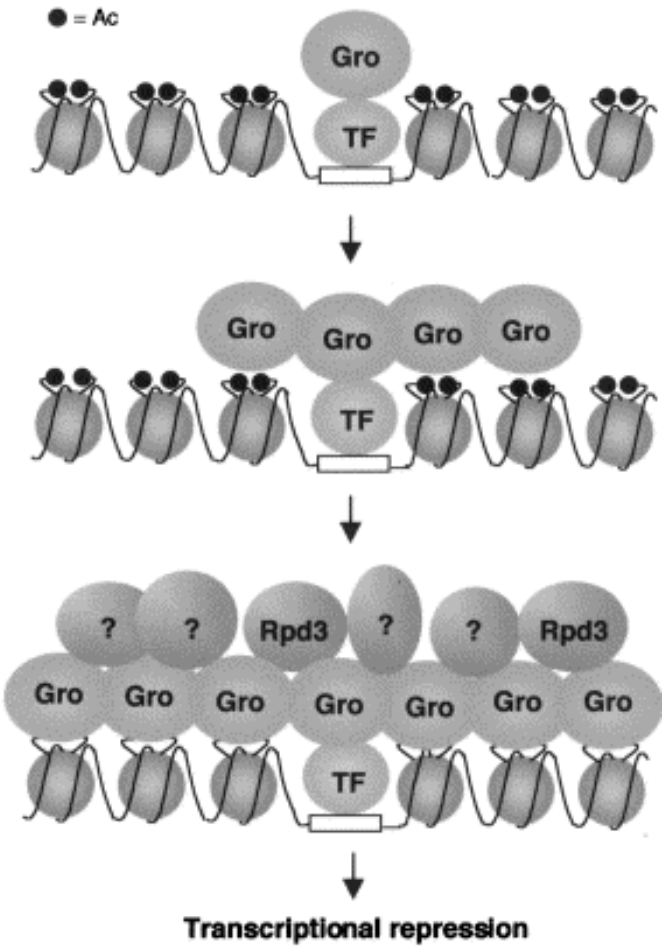

Fig.4

Possible model of Groucho-mediated transcriptional repression. Ac – acetylated lysine residues of histones. Gro – Groucho molecules. TF – A general DNA-binding transcription factor (taken from Chen and Courey, 2000).