THESIS ON NATURAL AND EXACT SCIENCES B173

Identification and Structure-Functional Characterisation of the Gene Transcriptional Repressor Domain of Human Gli Proteins

ROBERT TSANEV

TUT
PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
Faculty of Science
Department of Gene Technology

The dissertation was accepted for the defence of the degree of Doctor of Philosophy in Gene Technology on July 2, 2014

Supervisors: **Priit Kogerman**, PhD
Senior Research Scientist, Faculty of Medicine, University of Tartu, Estonia

Torben Østerlund, PhD
Senior Research Scientist at Proffice Life Science for AstraZeneca, Denmark

Prof. **Lilian Järvekülg**, PhD
Department of Gene Technology, Tallinn University of Technology, Estonia

Opponents: Prof. **Peter Zaphiropoulos**, PhD
Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden

Arnold Kristjuhan, PhD
Senior Research Scientist, Department of Molecular and Cell Biology, University of Tartu, Estonia

Defence of the thesis: August 22, 2014

Declaration:
Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted for any other academic degree.

/Robert Tsanev/

ISSN 1406-4723
ISBN 978-9949-23-658-9 (publication)
ISBN 978-9949-23-659-6 (PDF)

LOODUS- JA TÄPPISTEADUSED B173

Inimese Gli-valkude transkriptsioonilise repressordomeeni määramine ja struktuurilis-funktsionaalne iseloomustus

ROBERT TSANEV

TTÜ
KIRJASTUS

CONTENTS

INTRODUCTION

During embryogenesis, different systems and patterns evolve according to the developmental program of the organism. The spatial arrangement of organ development and embryonic patterns are established by gradients of signalling molecules. These molecules have higher concentration in the place of their synthesis or secretion, diminishing with distance. This drop in concentration creates a signalling gradient. The gradient provides positional information and determines the developmental fate of cells in a multi-cellular embryo. The idea that signalling gradients are responsible for providing positional information was developed by Wolpert already in 1969 and was named the French flag model (Wolpert 1969). Genetic or regulatory disruption of these signals can lead to developmental defects or miscarriage and to cancer development in adult animals.

Sonic hedgehog is a secreted signalling molecule responsible for embryonic pattern formation and organ development. Sonic hedgehog binds to its transmembrane receptor and triggers an intracellular cascade of signalling events. Eventually, this leads to cell proliferation and differentiation. Gli proteins are the effector transcription factors of hedgehog signalling. They regulate the expression of pathway target genes. The repressor forms, Gli-repressors, possess transcriptional repressor activity within their N-terminus. Gli-repressor binds to the target DNA and turns off the gene expression. Disruption of Shh signalling has been associated with various developmental disorders and cancer types.

The gradient of Sonic hedgehog and the opposing Gli-repressor gradient set the patterning in the embryonic neural tube and limb bud. These gradients establish the dorsoventral or anterior-posterior axes in the developing neural tube and limb bud, respectively.

Gli proteins are transcription factors with many build-in functions. There is a Zn-finger DNA binding part and transcriptional activator and repressor functions represented by the domains containing binding motifs for interactions with other molecules. Identification of the Gli protein amino-acid sequence responsible for repressor-activity and its structural and functional characterisation is the scope of this thesis.

ORIGINAL PUBLICATIONS

I. Laht, S., K. Meerits, H. Altroff, H. Faust, **R. Tsaney**, P. Kogerman, L. Jarvekulg, V. Paalme, A. Valkna and S. Timmusk (2008). Generation and characterization of a single-chain Fv antibody against Gli3, a hedgehog signaling pathway transcription factor. Hybridoma (Larchmt) 27(3): 167-74.

II. **Tsanev, R.**, P. Tiigimagi, P. Michelson, M. Metsis, T. Osterlund and P. Kogerman (2009). Identification of the gene transcription repressor domain of Gli3. FEBS Lett 583(1): 224-8.

III. **Tsanev, R.**, K. Vanatalu, J. Jarvet, R. Tanner, K. Laur, P. Tiigimagi, B. B. Kragelund, T. Osterlund and P. Kogerman (2013). The transcriptional repressor domain of Gli3 is intrinsically disordered. PLoS One 8(10): e76972.

ABBREVIATIONS

aa	amino acid
AER	apical ectodermal ridge
CD	circular dichroism
CTD	C-terminal domain
E9.5	embryonic day 9.5 of development
ELISA	enzyme-linked immunosorbent assay
EMSA	electrophoretic mobility shift assay
FGF	fibroblast growth factor
Fp	floor plate
GCPS	Greig cephalopolysyndactyly syndrome
hGli2	human Gli2
HDAC	histone deacetylase
Mab	monoclonal antibody
MoRF	molecule recognition feature
NLS	nuclear localisation signal
NMR	nuclear magnetic resonance
PHS	Pallister-Hall syndrome
PKA	protein kinase A
RD	repressor domain
scFv	single-chain variable fragment
Shh	Sonic Hedgehog
siRNA	small interfering RNA
Smo	Smoothened
Sufu	Suppressor of fused
V_H	heavy chain
V_L	light chain

1. REVIEW OF THE LITERATURE

The developmental role of Gli-repressor molecules is comprehensively demonstrated in development of the neural tube and limb bud. Accordingly, the literature review section begins with description of these embryonic systems. Further, the aspects known so far of Gli repressor formation are discussed. The Gli repressor-domain is intrinsically disordered, therefore, the last section of the literature review describes the properties of this protein group.

1.1. Opposing Hedgehog and Gli-repressor gradients establish dorso-ventral axis of the embryonic neural tube

The neural tube is a vertebrate embryonic precursor for the central nervous system. It stretches the entire length of an embryo and forms brain structures anteriorly and the spinal cord posteriorly. The neural tube appears as an oval structure in embryo cross-sections. At dorsal and ventral poles, structures named the roof plate and floor plate, respectively, are located.

At embryonic day 8.5 of mouse development the central mesodermal cells start to pack tightly and form the notochord that separates from the adjacent somatic mesoderm. At E9.5, the Shh is initially secreted by the notochord, which induces the secondary centre of Shh production from the floor plate (Chang et al. 1994). There is the maximal concentration of Shh that diminishes towards the roof plate. So the Shh gradient is higher ventrally and drops dorsally. This gradient establishes the dorsoventral axis of the developing neural tube (Fig.1). The Shh gradient was revealed by visualisation of Shh protein in this neural tissue (Gritli-Linde et al. 2001; Huang et al. 2007; Chamberlain et al. 2008). The gradient determines the identity of different neurons, so that distinct neurons are generated from the respective neural domain. These domains are established in a correct spatial order along the dorsoventral axis of the neural tube (Fig. 1) (Jessell 2000). *Ex vivo* experiments are in line with the graded distribution of Shh, showing that the neural tissue incubation with recombinant Shh switches the identity of cells towards more ventral cell identities with increased Shh concentrations (Briscoe et al. 2000; Dessaud et al. 2007). The prolonged exposition to Shh also directs the cells to a more ventral identity (Ericson et al. 1996).

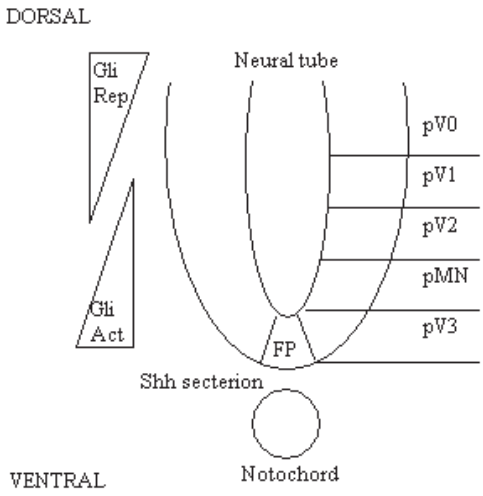

Figure 1. Scheme of a transverse section of the neural tube and notochord. The floor plate (FP) is located at the ventral pole of neural tube. Shh is secreted initially from the notochord, then from the floor plate. The formation of Gli activity gradient is depicted on the left. Distinct subtypes of interneurons (pV0-pV3) and motor neurons (pMN) are generated along the dorsoventral axis shown on the right. Shh is a ventral fate determinant, whereas Gli repressor is a dorsal determinant.

Shh binds to its cellular receptor Patched that in its turn regulates the activity of another transmembrane protein Smoothened (Briscoe and Therond 2013). The molecular cascade downstream of Smo is weakly understood but it has been shown to depend on the primary cilium and intraflagellar transport (Eggenschwiler and Anderson 2007). Ultimately Smo regulates the Gli family of transcription factors. The extracellular Shh activity gradient is converted into an integrated functional gradient of intracellular Gli proteins activity along the dorsoventral axis (Fig. 1). There are three Gli proteins: Gli1, Gli2 and Gli3. Gli1 and Gli2 are transcriptional activators and act together to specify the neural tube two most ventral domains FP and pV3 (Matise et al. 1998; Park et al. 2000; Bai and Joyner 2001). Gli1 null mice have no phenotype, whereas Gli2 null embryos lack FP and most of pV3 domain.

Gli3 is predominantly responsible for the Gli repressor function. Gli transcriptional repressor form is the dorsal fate determinant of progenitor domains in the neural tube (Hui et al. 1994). The Gli repressor formation is inhibited by Shh in ventral compartments of the neural tube. This inhibition is

lost in Shh null mice where, as a result, the Gli repressor is produced uniformly. In these mice only dorsal identities are developed and all ventral progenitors are missing (Chiang et al. 1996; Pierani et al. 1999). Removing of Gli3 in Shh mutant background restores the missing ventral progenitor domains: pV0, pV1, pV2 and pMN (Litingtung and Chiang 2000; Persson et al. 2002; Wijgerde et al. 2002). In the double mutant the progenitor domains are restored but they are abnormally intermingled showing that the role of Shh and Gli repressor gradients is to determine the accurate patterning and positioning.

It is the dynamic molecular dialog between the opposing Shh and Gli repressor gradients that establishes the correct spatial arrangement of neural progenitor domains. The two gradients have an active bilateral communication to refine each other. For instance, the production of Gli repressor in target cells is controlled by Shh (Ribes and Briscoe 2009). The cells are also actively reshaping the Shh gradient by producing Shh promoting or inhibiting membrane-proteins (Goodrich et al. 1996; Allen et al. 2007). Differential response to molecular dialog between the graded Shh signalling and the transcriptional network of the target cell is determined by the regulatory architecture of Shh signalling network and not by differences in sensitivity of gene expression to Shh signalling. In this view, the network can interpret the Shh gradient even when the genes are equally responsive to the signal (Balaskas et al. 2012).

1.2. Opposing Hedgehog and Gli-repressor gradients establish anterior-posterior axis of the developing limb bud

Mouse limbs start to develop at embryonic day 9.5. By that time symmetric swellings appear in the lateral body wall (Wanek et al. 1989). These buddings consist of lateral plate mesenchyme surrounded by ectoderm. Signalling from three organising centres controls the limb bud development (Zeller et al. 2009). These centres are: the apical ectodermal ridge (AER) responsible for the formation of proximodistal (shoulder to fingertips) limb axis; the zone of polarising activity (ZPA) establishing the anterior-posterior (thumb to little finger) limb axis, and the non-AER ectoderm driving the dorsoventral (knuckles to palm) patenting.

Formation of the apical ectodermal ridge is induced by FGF10 expressed from the below lying mesoderm that signals to the surrounding ectoderm (Ohuchi et al. 1997). This signalling centre synthesises factors required for limb bud outgrowth and patterning. Removal of the AER results in truncation of limb development (Summerbell 1974). The earlier the AER was removed, the more proximal the level of limb truncation was observed. The factors needed for the limb outgrowth were identified to be FGF4 and FGF8 (Sun et al. 2002). Their

removal resulted in an absence of limb outgrowth. The proximal signal in limb development was identified to be the retinoic acid (Rosello-Diez et al. 2011).

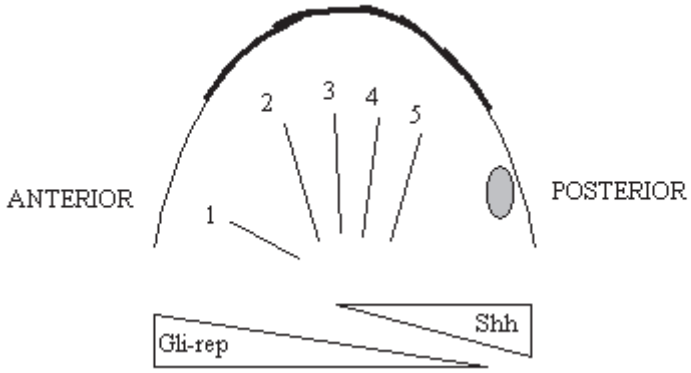

Figure 2. Scheme of a lateral section of the limb bud. The thicker line on the limb bud edge represents the apical ectodermal ridge (AER), the thinner line shows the non-AER ectoderm. The filled ellipse represents the zone of polarising activity. The finger progenitors are shown and numbered beginning from the thumb. Formation of the opposing Gli-repressor and Shh gradients is presented below the limb bud scheme.

Opposing action between Shh produced posteriorly and Gli-repressor expressed anteriorly sets up the anterior-posterior axis of the developing limb. Shh controls the processing of Gli3 to Gli repressor. Shh sets the precise ratio of Gli activator to Gli repressor. This ratio is essential for limb development (Wang et al. 2000). Hox genes activate the expression of Shh (Kmita et al. 2005; Galli et al. 2010). Shh is synthesised from E10.25 to E12.5 from mesodermal cells (Haramis et al. 1995). The cells producing Shh form a signalling centre named the zone of polarising activity (Riddle et al. 1993). Transplanting these cells or Shh soaked beads to the anterior border of a control limb produced a mirror-image duplication of digits (Tickle et al. 1975; Lopez-Martinez et al. 1995). Mice mutant for Shh develop only digit 1 (thumb) in hind limbs and no digits in forelimbs (Kraus et al. 2001), whereas Gli3 null mice limbs develop polydactyly and lack digit identity (Johnson 1967). From these observations it can be concluded that Gli-repressor is to restrict the pattern formation of polydactylous digits, wheras Shh promotes the patterning of multiple digits. However, when removing the both genes, Gli3 and Shh, the resulting mouse limbs were surprisingly identical to the Gli3 null where the Shh gene was present (Litingtung et al. 2002). These observations indicate that the limb has in fact an

intrinsic ability to form digits, even in the absence of Shh. But for development of normal digit number and identity, a balanced counteraction between Shh and Gli repressor is established.

Mutations in the Gli3 gene are associated also with rare human developmental disorders like the Greig syndrome (GCPS) and Pallister-Hall syndrome (PHS) (Biesecker 2006). These syndromes have distinct features, but shared finger/thumb polydactyly. Most of the disorder-causing mutations result in a truncated form of Gli3 protein. It appears that the place of mutation within the Gli3 gene correlates with the arising disorder phenotype (Biesecker 2006). GCPS is caused by mutations localised mostly in the first or last third of the gene. This results in a Gli3 variant with lost DNA-binding domain or with the affected transcriptional activation ability. On the other hand, the PHS causing mutations localises to the central third of Gli3. The Gli3-PHS protein variant is the N-terminal part of Gli3 including the DNA-binding domain. It has been shown that the variant acts as a strong transcriptional repressor (Shin et al. 1999). This indicates that within the truncated protein there is an amino-acid sequence turning the Gli3-PHS construct into a transcriptional repressor.

Shh and Gli3 control the both, proliferative expansion and differentiation of mesenchymal progenitors (Zhu et al. 2008). Gli3, through its interaction with the cell-cycle regulator Cdk6 (Vokes et al. 2008), acts as a regulator of S phase entry for digit progenitors in the anterior hand plate (Lopez-Rios et al. 2012). In addition, Gli3 repressor promotes proliferation exit and BMP-dependent condensation of skeletal elements and chondrogenic differentiation of mesenchymal progenitors. These observations show that Gli3 fulfils a dual role in limb development by regulating both cell-cycle entry and exit to chondrogenic differentiation (Lopez-Rios et al. 2012).

Finally the third axis of the limb, the dorsoventral patterning, is established by the differential gene expression in dorsal and ventral non-AER ectoderm. For instance, the dorsal ectoderm expresses a dorsal fate determining genes like Wint7a (Parr and McMahon 1995), and the ventral ectoderm expresses ventral determinants like Engrailed (Loomis et al. 1996; Logan et al. 1997).

1.3. The formation of Gli repressor is phosphorylation dependent proteasomal processing

Most of the Hedgehog pathway proteins are localised into a cellular compartment named cilium (Haycraft et al. 2005; Rohatgi et al. 2007). This is a small outgrown organelle consisting of a cytoplasm and cytoskeleton covered by a cellular membrane. The Hedgehog receptor Pached1 and co-receptor Smoothened have been detected in cilia as well as Gli proteins and Sufu. It is likely that many of the pathway critical reactions occur in the cilium. Disruption

of this cellular structure results in a severe decrease of Gli repressor level (Haycraft et al. 2005; Cortellino et al. 2009) and failure of full-length Gli to translocate to the nucleus when the pathway is activated (Humke et al. 2010).

In the cytoplasm of Hedgehog unstimulated cells, Gli transcription factors are in a complex with Sufu protein (Pearse et al. 1999; Humke et al. 2010). Sufu binding keeps Gli from translocating to the nucleus and protects Gli from degradation (Kogerman et al. 1999; Humke et al. 2010). In the absence of Sufu, in Sufu knock-out cells, the Gli protein level is drastically reduced and Gli proteins enter the nucleus (Chen et al. 2009; Jia et al. 2009). This results in the start of transcription and ectopic activation of Hedgehog pathway. In wild-type cells Sufu-Gli complex cycles in and out of cilia in a low rate (Tukachinsky et al. 2010). In the absence of Hedgehog signalling, Sufu-Gli complex is recognised by protein kinase A (PKA) on the base of the cilium (Tuson et al. 2011). PKA phosphorylates Gli proteins and that initiates further phosphorylation by GSK3β and casein kinase 1 (CK1) (Pan et al. 2006; Tempe et al. 2006). The phosphorylated full-length proteins Gli3 (~190kDa) and Gli2 (~185kDa) bind to the ubiquitin ligase complex (Skp1/Cullin1/F-box) that targets Gli proteins for cleavage via the ubiquitin-proteasome pathway (Tempe et al. 2006). The proteins are cleaved after the zinc-finger domain to N-terminal repressor fragments, Gli3R ~83kDa and Gli2R ~78kDa (Wang et al. 2000). After the cleavage the Gli repressor form is no longer in a complex with Sufu and translocates to the nucleus to repress transcription of target genes (Humke et al. 2010; Tukachinsky et al. 2010). The processing of Gli proteins is thus dependent on PKA and cilia. However, some Gli processing was detected also in PKA-null cells and in mutants lacking cilia (Huangfu et al. 2003; Liu et al. 2005; Tuson et al. 2011). These findings suggest that there is also a PKA and cilia independent mechanism for Gli repressor formation. The main form of Gli repressor is the Gli3 N-terminal fragment. The processing of Gli2 is less efficient and in the absence of Hedgehog signalling Gli2 is instead degraded (Pan et al. 2009).

Gli1 is not processed and it functions only as a transcriptional activator (Dai et al. 1999; Kaesler et al. 2000). In the presence of Hedgehog, full-length Gli3 and Gli2 are no longer processed and act as transcriptional activators (Wang et al. 2000; Pan et al. 2006). Pathway activation results in a quick accumulation of Sufu-Gli complex in the cilium. This complex cycles within the cilia also at unstimulated conditions, but upon stimulation its concentration rises quickly (Wen et al. 2010). Probably the pathway activation causes a modification of Gli proteins that takes place within the cilium (Wen et al. 2010). That modification is the likely cause of the dissociation of the Sufu-Gli complex. When released from Sufu, Gli proteins translocate into the nucleus (Tukachinsky et al. 2010). There Gli is activated by a nuclear kinase that renders it highly unstable (Humke et al. 2010). Hedgehog-activated full-length Gli transcription factors activate the target genes and are degraded in the nucleus by the nuclear SPOP/Cul3 complex

(Chen et al. 2009; Wen et al. 2010). Hedgehog signalling triggers the degradation of full-length Gli3, but it does not disappear completely. Pathway activation inhibits also the production of Gli repressor form so that it disappears completely but more slowly than the full-length Gli3 protein (Wen et al. 2010).

The primary transcriptional activator of Hedgehog pathway is mostly Gli2 and less so Gli3. They induce the expression of a secondary activator that is a Gli1 transcription factor acting in a positive loop manner (Ding et al. 1998; Dai et al. 1999; Bai and Joyner 2001). To be activated, the newly synthesised Gli1 has to pass through cilia in a complex with Sufu. There, continuing Hedgehog signalling will cause a modification of Gli1. This would ensure that the Hedgehog pathway remains signal dependent (Tukachinsky et al. 2010).

1.4. Intrinsically disordered proteins

This is a large group of proteins, peptides and domains defined on the basis of what they do not have. These proteins are not restricted in a concrete three-dimensional structure, but instead experience less limited spatial freedom, e.g. they are less structurally defined and more flexible. Although these proteins are not totally unstructured, some conformations are more preferable than others (Tompa 2010). Structural disorder is enriched among proteins involved in processes like gene regulation or signal transduction, and depleted in enzymatic and ligand binding reactions (Xie et al. 2007; Tantos et al. 2012). In some instances, the disordered proteins can undergo binding-coupled folding (Dyson and Wright 2005), whereas in others there is a large degree of spatial freedom in the complex (Fuxreiter et al. 2011). In the latter case, the partners contact each other by loosely defined and transient contacts, which do not bring about a well-defined ordered structure even in the complex state (Fuxreiter 2012). The coupling of interaction to folding enables unstructured proteins to enhance specificity in expense of binding affinity (Spolar and Record 1994). Therefore the binding to their functional partner is weak, specific and dynamic. Intrinsically disordered proteins recognise their partners by a number of mechanisms. A predominant one utilises short sequence motifs named molecular recognition features – MoRFs (Mohan et al. 2006). In an unbound state these elements might be structured as well as unstructured and their folding can be initiated by complex formation. Intrinsically disordered proteins are stabilised more by interactions with the partner than from interactions with their own chain (Drobnak et al. 2013). On the contrary, ordered proteins establish more stabilising interactions with their own chain and their fold is self-driven.

As binding can initiate the folding of disordered proteins, such structure induction can be used as a good indicator of interaction. CD-spectroscopy and NMR measurements are often employed in studying the binding of disordered

proteins (Libich and Harauz 2008; Chemes et al. 2012; Dasgupta et al. 2012). On the other hand, since the binding of unstructured proteins to their partner is weak and dynamic, methods like immunoprecipitation and yeast two-hybridisation are not immediately suitable for studying disordered proteins interactions (Tompa 2010). These methods are designed for studying stronger and more stable interactions.

Intrinsically disordered proteins employ bigger capture radius for partner finding, which makes the scanning for the interactor more efficient (Shoemaker et al. 2000). The interaction surface of disordered proteins is bigger compared to a structured protein of a similar size (Gunasekaran et al. 2003). The complex formed by unstructured proteins is held together by weak, mostly hydrophobic contacts separated by linker sequences rather free to change (Meszaros et al. 2007). The contacts between the partners are transient and a loss of individual contact will not disrupt the whole complex. This makes the intrinsically disordered proteins more tolerant to mutations and their complexes more resistant to the loss of the related function. Upon mutation, globular proteins lose their function suddenly, whereas intrinsically disordered proteins lose the function gradually.

Structural disorder provides proteins with the ability to interact with multiple partners by making different contacts from the same interaction surface (Hegyi et al. 2007). The intrinsically disordered proteins and domains are more accessible to proteases. Their shorter life-time ensures quicker reaction to different extracellular signals. These and other advantages make structural disorder suitable for and enrich it in processes like transcription and signal transduction. Also, structural disorder of proteins is more frequent in multicellular organisms (Schlessinger et al. 2011).

As Gli proteins are transcriptional regulators, it is not surprising that parts of these may have features of intrinsically disordered regions. Structural prediction programs will assign structural features to distinct amino-acid sequences within Gli proteins, but the actual degree of order/disorder can be reviled only empirically. The short sequence elements with high structural propensities might well represent the partner recognition motives – MoRFs. In the process of transcriptional regulation, Gli proteins will bind to their partner molecules probably in a transient and loose way. The resulting structure induction in the bound state could be a good indication for complex formation. The pursuit for Gli binding partners will help to comprehend in more detail the architecture of the molecule network used by Hedgehog pathway for the regulation of transcription. The knowledge gained can be used for better diagnosis and treatment of diseases like cancer.

2. AIMS OF THE STUDY

The Hedgehog gradient provides positional information to cells during embryonic development. In organogenesis, this signalling pathway establishes the left-right asymmetry and determines the differentiation fate of the target cells. This is one of the major reasons to explore in more detail the regulation of the pathway. The Hedgehog pathway is active during embryogenesis, but it is mainly quiescent in adults. Inappropriate reactivation, however, contributes to various cancers, thus providing further requirement for research on the pathway inactivation.

To elucidate the transcriptional repression function of the Hedgehog pathway, these main objectives were posed:

1. To generate an anti-Gli3 intrabody.
2. To identify the amino-acid sequence stretch within Gli proteins responsible for the transcriptional repression function.
3. To verify whether the Gli repressor domain binds to Ski co-repressor protein.
4. To test the interaction of possible molecular partners with the Gli3 repressor domain.
5. To analyse the mechanism of Gli3 repressor domain function.
6. To determine the 3D-structure of Gli3 repressor domain.
7. To identify functionally important amino-acid residues within the Gli3 repressor domain sequence.

3. MATERIALS AND METHODS

A detailed description of materials and methods is provided in the publications of this thesis. Briefly, the following methods were used in the present study:

- Cloning and mutagenesis (Publications I, II and III)
- Cell culture and transfection (Publications II and III)
- Luciferase assay (Publication II and III)
- His pull-down and immunoprecipitation (Publication II and III)
- Western blot (Publications II and III)
- Bioinformatic analysis of protein sequence and structure prediction (Publication III)
- Expression and purification of recombinant protein (Publication III)
- MALDI TOF MS (Publication III)
- Electrophoretic mobility shift assay (Publication III)
- Circular dichroism spectroscopy (Publication III)

4. RESULTS

These studies were performed in order to elucidate in more detail the nature of the transcriptional repression function of Gli proteins. As a result of the investigations, a single-chain Fv intrabody against Gli3 was generated (paper I), the minimal transcriptional repressor domain of Gli proteins was defined, and the repression mechanism was shown to be independent of HDACs (paper II). The Gli3RD was described as an intrinsically disordered region, performing autonomously, independent of binding to Ski, DNA and Zn^{2+} (paper III). The residues H141 and H157 were identified as constituting functionally important parts of the domain (paper III).

4.1. Generation and characterization of a single-chain Fv antibody against Gli3, a Hedgehog signaling pathway transcription factor (publication I)

Antibody engineering, i.e. construction of a single-chain variable fragment (scFv) intrabody can overcome the inaccessibility of intracellular antigens like transcription factors. For this purpose, the variable regions of the heavy (V_H) and light (V_L) chain of Gli3 MAb 5E1 were obtained by amplification of hybridoma transcripts. An inert flexible glycine-serine linker connected the isolated V_H and V_L fragments and the resulting construct was cloned into prokaryotic and eukaryotic expression vectors. The bacterially expressed anti-Gli3 scFv displayed a binding capacity more than seven-fold lower than that of the original MAb 5E1, when tested by dose-response ELISA. The eukaryotic His-tagged anti-Gli3 scFv showed diffuse cytoplasmic expression in Gli3 negative cell-lines, Cos-1 and PK15. This diffuse expression changed to discrete perinuclear, when the anti-Gli3 scFv was co-expressed together with Gal4-tagged human Gli3RD in Cos-1 and PK15 cells. Furthermore, the expression pattern of anti-Gli3 scFv was very similar to that of Gal4-tagged human Gli3RD in the tested cells. These results suggest that the intrabody may indeed target cytoplasmic human Gli3. Interestingly, while a weak Gal4-Gli3RD signal was also detected in the nucleus, the scFv showed no nuclear localization.

The expression pattern of anti-Gli3 scFv was not altered when co-expressed with GFP-Gli2 in PK15 cells. Thus, the Gli3 intrabody does not cross-react with Gli2 and is specific for Gli3, despite the high similarity of the recognition epitopes within both proteins. To test the capacity of anti-Gli3 scFv to recognize

mouse Gli3, the anti-Gli3 scFv was over-expressed in the TM4 mouse cell line, containing endogenous Gli3. The anti-Gli3 scFv displayed a diffuse cytoplasmic localisation with a predominant perinuclear staining. This pattern was similar to that when the cells were stained by the parent Mab 5E1, suggesting that the anti-Gli3 scFv indeed recognises the murine Gli3. In some TM4 cells, a dot stained by anti-Gli3 scFv was observed in a location distant from the nucleus, which may indicate a cilia staining of Gli3.

4.2. Identification of the gene transcription repressor domain of Gli proteins (publication II and unpublished data)

To analyse the Gli3 repressor function in detail, the primary task was to identify the minimal protein region responsible for the repressor activity. To this end, the Gli3-PHS peptide (residues M1-Q691), which is the N-terminally truncated form of full-length Gli3 (1580 residues), was shown to be a strong repressor of Gli1-induced transcription. The repressor effect of PHS was even stronger than that of the full-length Gli3. This was shown using a vector-based luciferase reporter system in HEK293 cells and in a more in-vivo like settings in Shh-Light2 cells. Shh-Light2 cells have a Gli-inducible luciferase reporter construct incorporated into the genome.

The naturally occurring transcriptional repressor forms of Gli proteins are the N-termini of Gli3 and Gli2, whereas Gli1 does not function as transcriptional repressor. By comparing the amino acid sequences of N-termini of Gli proteins a region present in Gli2 and Gli3 but absent in GLI1 was identified. This region was deleted (residues R105-G246) in full-length Gli3 and Gli3-PHS and the resulting constructs were named Gli3ΔRD and Gli3-PHSΔRD, respectively. Removal of this part abolished the repressor function, more clearly seen in Gli3-PHS. The full-length Gli3 behaves as a weaker repressor than Gli3-PHS and the loss of repression in Gli3ΔRD was not so pronounced. On the other hand, the transcription activator function of Gli3 was enhanced in Gli3ΔRD by the removal of the repressor domain (RD).

To assess the minimal repressor domain of Gli3, constructs were made, containing this RD sequence or parts of it in frame with Gal4 DBD. The influence of these constructs to Gal4 binding site containing Luciferase reporter was assessed. There activity was compared to the activity of a mock, only Gal4 DBD containing construct. Thus the minimal RD of Gli3 was mapped to residues G106-E236, in a Gla4 DBD responsive luciferase assay in HEK293 cells. When the RD was shortened beyond these residues, the repression was reduced or lost.

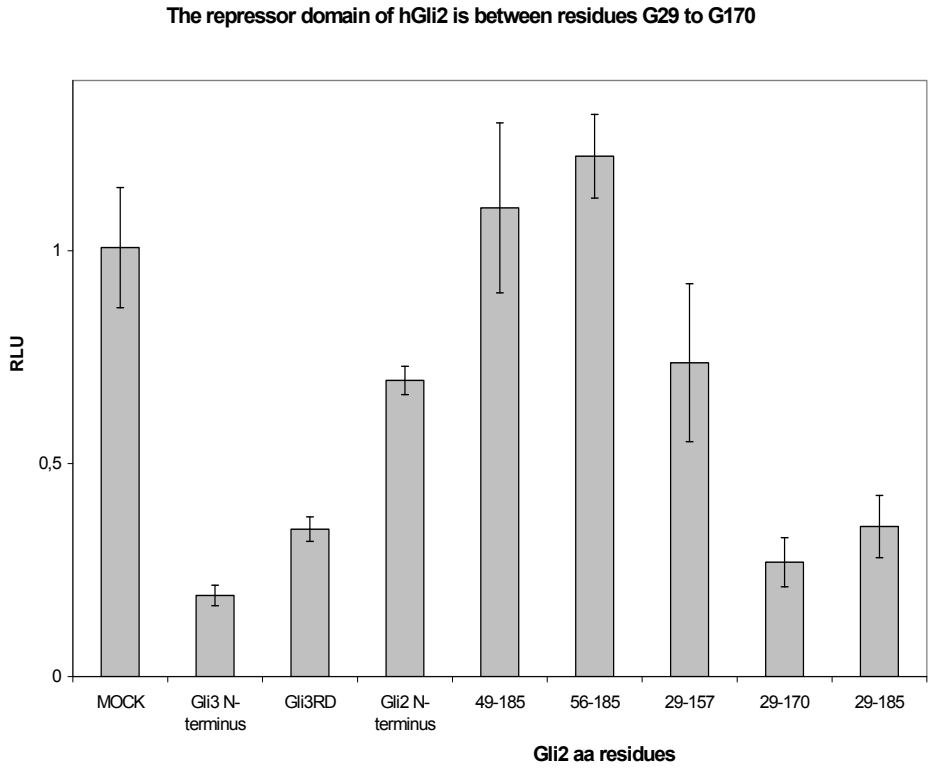

Figure 3. Identification of the minimal repressor domain of human Gli2. Overlapping parts of Gli2 were expressed as Gal4 DBD fusion proteins in HEK293 cells. The repressor function of these constructs was assessed on Gal4 binding site containing Luciferase reporter. The activity of all constructs is relative to an empty, only Gal4 DBD containing vector (mock).

The minimal protein region of Gli2, responsible for the repressor activity of Gli2 was defined analogously to Gli3RD. Constructs containing overlapping parts of Gli2 with different length were fused in frame with Gal4 DBD. There activity was assessed in Luciferase reporter assay in HEK293 cells. Consequently the minimal RD of Gli2 was mapped to residues G29 to G170 (Fig. 3).
The repressor domains of Gli2 and Gli3 display similar strength of activity (Fig. 3). These two protein stretches are also very similar in their amino acid sequence. Still, the N-termini of Gli2 and Gli3 have significantly different activity (Fig. 3). Gli2 N-terminus performs as a weaker repressor because of the intensive degradation of this protein as revealed on a Western blot.

It has been published that HDACs are brought to Gli3 by interaction with Sufu or Ski. So the next question was whether the repressor function of Gli3RD depends on recruitment of HDACs. Thus, the effect of the HDAC inhibitor TSA was tested on the repression of both, Gli3RD and the repressor domain of REST (the latter serves as positive control since it represses transcription by recruiting HDACs). It was determined that the repressor activity of Gli3RD does not depend on recruitment of HDACs.

As a further control, physical interaction between the minimal repressor domain and HDAC was not discernible in pull-down assay.

4.3. The transcriptional repressor domain of Gli3 is intrinsically disordered (publication III)

Here the structural propensities of Gli3RD and the interaction with its potential partners were studied. From the bioinformatic analysis several structured elements were predicted mainly in the C-terminal part of Gli3RD while the N-terminus was shown to be predominantly disordered. Additionally, two Anchor sites for potential partner interactions were identified.

To evaluate the Gli3RD actual structure in native-like conditions, in-cell NMR was performed. This showed Gli3RD to be intrinsically disordered in an intracellular environment. Likewise the purified Gli3RD was intrinsically disordered in a solution environment.

The interaction and structure induction of Gli3RD was tested with Ski, a known partner of Gli3. The NMR spectrum of Gli3RD remained unchanged in the presence of MBP-Ski indicating that the proteins did not interact. Interaction between Gli3RD and Ski was also not observed in a co-immunoprecipitation experiment.

Some transcription factors bind to DNA in sequence unspecific way through their intrinsically disordered regions. So the Gli3RD interaction with DNA was also investigated. The CD spectrum of Gli3RD stayed unchanged in the presence of a 21 base-pair scrambled sequence oligonucleotide or a plasmid DNA, indicating no interaction or at least no structure induction. Then the Gli3RD and DNA interaction was investigated in an EMSA assay. This assay also did not elucidate any binding of DNA to Gli3RD.

The Zn^{2+} binding of Gli3RD was studied in a functional assay by mutating certain histidines, resembling class I or II of Zn^{2+} ligands. Transcriptional repression was preserved in single histidine mutants and even in the double histidine mutant H121/147A, where both classes of Zn^{2+}-ligands should be affected. By this analysis it was verified that Zn^{2+} binding is not involved in the repressor function. The loss of Gli3RD activity was observed in two variants, H141A and H157N. The reason for this might be the altered local structural

propensities, since both mutations significantly increased the predicted occurrence of helicity at their corresponding sites, whereas the non-function affecting mutant H157A leaves the helical propensity unchanged.

5. DISCUSSION

The investigation of the Gli repressor domain led to the development of an anti-Gli3 intrabody. This intrabody can be used for studying Gli3 protein amount, localisation and trafficking. The minimal repressor domain of human Gli proteins was defined to aa residues G29 to G170 in Gli2 and G106 to E236 in Gli3. The repression mechanism was characterised to be independent of HDACs. The in solution and in-cell structure of Gli3 repressor domain was shown to be intrinsically disordered, containing predicted sites for partner interaction. The repressor function of the Gli3RD was demonstrated to be independent from binding to Ski, DNA or Zn^{2+}. Analyses of H141 and H157 mutations identified these residues as being of functional significance or in a functionally important region of the domain.

5.1. Generation of anti-Gli3 intrabody

Intrabodies are antibody-derivates ectopically expressed inside the cell. As such, they can be used for modulating protein expression and trafficking or for protein inactivation. For gene inactivation, intrabodies operate at the protein level, whereas such techniques as mouse knockout or siRNA function at the DNA and RNA level, respectively. It means that an intrabody directed against a given protein's domain could block only a certain biological function, leaving the other protein functions untouched (Li et al. 2007).

The anti-Gli3 intrabody constitutes a tool for studying intracellular localization and trafficking of Gli3. The observed dotty staining of the anti-Gli3 intrabody in TM4 cells suggests a ciliarly concentrated localization of endogenous Gli3 and the usefulness of the anti-Gli3 intrabody. Rinaldi and colleagues report the use of intrabodies for visualisation of lowly expressed oncoprotein, gankyrin, in living cells (Rinaldi et al. 2013). To this aim, they conjugated the intrabody with a green or red fluorescent protein and monitored the fluorescence resonance energy transfer (FRET) on the dual binding of the fluorescent intrabodies to gankyrin. This strategy can be an interesting option also for visualising endogenous Gli3 in living cells. It will be helpful for comparing Gli3 localisation and trafficking in cilia mutants and wild type cells that are exposed to cellular stimuli like Hedgehog or cyclopamine.

In addition, it would also be interesting to know whether the intrabody is able to affect the function of Gli3. Since its epitope is within the repressor domain of Gli3, this might influence the repressor function. But it would be expected that

the intrabody alters the Gli3 repressor function if it was able to translocate to the nucleus. This can be achieved by adding an NLS-signal to the intrabody (which remains to be done). Then, if the intrabody alters the repressor function of Gli3, it will be intriguing to see whether this will have any impact on the activator function of Gli3. In the study of the repressor domain identification, it was seen that removal of the repressor domain increases the activator function of Gli3, and vice versa (Paper 2; Tsanev et al. 2009). It will be challenging to investigate whether and how the activator and repressor domains influence each other's activity.

Intrabodies have been used as a tool for turning off a gene function. Gal-Tanamy et al. have blocked hepatitis C virus replication in infected liver cells by the use of intrabodies against NS3 protease (Gal-Tanamy et al. 2010). Butler and Messer report aggregate binding by intrabody in neurodegenerative disorders (Butler and Messer 2011). They have targeted the aggregates for proteasomal degradation with intrabody bearing a proteasomal PEST signal.

Pallister-Hall syndrome is a developmental human disorder that is associated with an increased level of a Gli3 repressor form (Naruse et al. 2010). Using the anti-Gli3 intrabody to diminish the repression activity by sequestrating Gli3 in the cytoplasm is an attractive strategy. Alternatively, the intrabody could direct Gli3 to the proteasome if it bears a PEST-signal, resulting in Gli3 proteasomal degradation. On the other hand, the intrabody not localising to the nucleus is a useful reagent for visualising Gli3 trafficking without affecting its function as a transcriptional regulator. We have yet to study the influence of the intrabody on the GLI3 activator and repressor functions, but this work has been hampered by the termination of the laboratory.

5.2. Defining the minimal repressor domain of Gli proteins

In this study, the repressor domain of human Gli3 was determined to be between aa residues G106 and E236. The repressor domain of human Gli2 was shown to be between residues G29 to G170. Also, the transcriptional repressor function implemented by this domain was shown to be independent of histone deacetylation.

To localise the repressor domain, first the amino-acid sequences of Gli proteins were compared. There are several conserved regions within N-termini of Gli3, Gli2 and Gli1 (Sasaki et al. 1999). The shared regions are responsible for functions common to all Gli proteins as Sufu or degron N binding (Dunaeva et al. 2003; Huntzicker et al. 2006). Sufu binding and degron N are parts of a negative regulatory mechanism shared among Gli proteins. These mechanisms are leading to degradation of the transcription factor that brings a termination of transcription. Thus the clearance of Gli signal and the resulting transcriptional

repression can be viewed as separate events, possibly combined by more than one mechanism. As this study focused on transcriptional repression, the contribution of other mechanisms could be avoided by defining the minimal amino-acid sequence bearing the repressor function. Previous studies have shown that the N-terminal halves of Gli3 and Gli2 are transcriptional repressors in the Hedgehog pathway and that Gli1 is a transcriptional activator (Sasaki et al. 1999; Lipinski et al. 2006). Based on this, the repressor domain should be present within Gli3 and Gli2 but missing in Gli1. Comparison between the amino-acid sequences of Gli proteins pointed to a conserved region within the N-terminus of Gli2 and Gli3 that was absent in Gli1. Deletion of this sequence abrogated the repressor function. This finding implied that the repressor function is encoded by this region. Shortening the sequence of the GLI3 repressor part in the Gla4 heterologous system defined the minimal transcriptional repressor domain of human Gli3 to be localised between residues G106 and E236.

It has been published that histone deacetylases are recruited to the N-terminus of Gli3 as a result of interaction with Sufu or Ski protein (Cheng and Bishop 2002; Dai et al. 2002). Accordingly, it can be concluded that the repressor activity of Gli3 depends upon histone deacetylation. This assumption was disproved here by showing the retained repressor activity in the presence of a deacetylase inhibitor. From this result it was concluded that Gli3RD uses an HDAC independent mechanism to repress transcription. It is more likely that the HDAC mechanism is part of a general negative regulatory path by which all three Gli proteins are regulated with the Ski binding site being conserved in all three Gli proteins.

The repressor domain deletion within the Gli3 N-terminus resulted, as expected, in loss of repressor activity. However, it was a surprise to see a small transcriptional activation as a result of this deletion. Since this construct misses the C-terminal activation domain, it should have been transcriptionally silent. The deletion of the RD sequence within the context of the full-length Gli3, on the other hand, converted this construct to a better activator. The enhancement of Gli3's activation ability by removing its repressor activity is anticipated. Also, this increase of activator function suggests that the activator and repressor domain might communicate with each other within the context of full-length Gli3.

Signals for transcriptional regulation are gathered up by a molecular complex named the Mediator complex. Mediator is summing up signals coming from transcription factors and other inputs, and acts as a molecular rheostat, modulating the activity of RNA pol II. The function of this complex is to bridge the transcription factors and RNA polymerase II in gene regulation and RNA synthesis (Malik and Roeder 2010). The Mediator complex is interacting with the activator domain of Gli3 (Zhou et al. 2006). Also, it is likely that the Mediator complex receives signals from the repressor domain of Gli3. In this

way, the activation and repression functions of Gli proteins can be translated to a proper level of gene transcription. This might occur if Gli3 repressor domain influences the interaction between mediator and Gli3's activator domain. This can happen by binding of the repressor domain to the activator domain or to the mediator complex itself.

It was reported that the Mediator complex is interacting with Gli3 activator domain through its subunit Med12 (Zhou et al. 2006). However, the communication between Mediator and the repressor domain should proceed through a different subunit, since, in the absence of this subunit the pathway target genes were still repressed in Drosophila model (Janody et al. 2003).

5.3. The repressor domain of Gli3 is intrinsically disordered

To gain more information on the repressor domain's function, its 3D structure was determined and binding sites for partner interaction predicted. Gli3 repressor domain binding to different interaction candidates was tested. The domain's functionally important amino acid residues were identified in a mutational analysis.

The bioinformatical analyses of Gli3 repressor domain predicted it to be mostly disordered containing some short structural elements. The high probability for disorder, however, extends beyond the repressor domain's border to the whole N-terminus of Gli3. Thus the lack of order is not a hallmark of the repressor domain only, but represents the structural state of the Gli3 N-terminus. It is common for proteins involved in transcription that they are entirely disordered or contain large unstructured regions (Tompa et al. 2006; Xie et al. 2007).

Most of the predicted structural stretches of the repressor domain are situated in its C-terminal part. This conditionally divides the domain into two halves: potentially more ordered C-terminus (residues L171–E236) and less ordered N-terminus (residues G106–D170). The balance assessment of hydrophobicity/charge, however, did not support the presence of strongly disordered regions, describing the domain as rather ordered. Thus, these algorithms display alternative results that describe different properties of the domain and characterise the domain's functional features.

To determine the repressor domain's actual degree of order/disorder, it was expressed as a His-tagged recombinant protein in *E. coli* cells and its in-cell NMR spectrum was recorded. In this experimental setup the protein was in a state as close to in-vivo conditions as possible and was characterised to be intrinsically disordered. This is in line with other investigations describing the proteins or their domains involved in transcriptional regulation to be predominantly disordered (Tompa et al. 2006; Xie et al. 2007).

The short sequence stretches within the repressor domain that were predicted to be ordered, probably undergo disorder-to-order transition when in contact with a functional partner of the domain (Mohan et al. 2006; Hinds et al. 2007). The resulting complex is then usually stabilised by intermolecular interactions between the two partners. Therefore, the structure induction is a good indication when studying protein interactions. Although, there are also instances were the proteins stay disordered even in the complex state (Fuxreiter et al. 2011).

A Gli3 binding partner has been published to be the transcriptional co-repressor Ski (Dai et al. 2002). Therefore, the binding to and structure induction within the Gli3 repressor domain was tested firstly together with the Ski protein. Its binding site on Gli3 (residues M1-P397) overlaps potentially with the repressor domain (residues G106-E236). This made Ski a good candidate for testing interaction. The result, however, showed that the two proteins did not interact. Thus the Ski binding site on Gli should be outside the repressor domain. This also confirms our anticipation from the HDAC analyses (Tsanev et al. 2009), since Ski utilises histone deacetylation for transcriptional repression (Nomura et al. 1999) and Gli3 repressor domain does not. Thus, we propose that Ski takes part in a general mechanism for ceasing the overall Gli signalling. This mechanism involves also Sufu and histone deacetylation. It is possible that in this process Ski will bind to Gli through the Sufu protein, and not directly.

Some transcription factors bind DNA in a sequence unspecific manner through their intrinsically disordered region (Liu et al. 2006; Tafvizi et al. 2010). The DNA binding of RNA polymerase II (RNA pol II) is also sequence independent. It is facilitated by its predominantly disordered C-terminal domain (CTD) (Suzuki 1990; Bienkiewicz et al. 2000). This proline rich domain binds DNA by intercalating of the tyrosine aromatic ring into the DNA strands (Suzuki 1990). In the amino-acid sequence of Gli3 repressor domain the spacing between some tyrosines and prolines is the same as in CTD of RNA pol II. Therefore, sequence unspecific DNA binding was the next guess for the repressor domain's interaction. However, there was neither structure induction nor binding as measured by CD spectroscopy and EMSA, respectively. In the EMSA experiments, the repressor domain did not bind DNA, but its DNA-binding tag did. So the repressor domain did not interfere with the DNA binding of its tag. Accordingly the repressor domain does not repress transcription either by masking the target DNA or by preventing the binding of other factors to DNA. From this, it was concluded that the mechanism of transcriptional repression used by this domain is not at the DNA level.

During the sequence analyses it was noticed that within the repressor domain the histidine-residues, H121/H157 and H141/H147, resembled class II and I of Zn-binding motifs, respectively (Karlin and Zhu 1997). To test whether zinc could be an interacting partner of the repressor domain, these histidines were mutated and the resulting effect on the domain's function was examined. The

mutant variants had impaired zinc coordinating potential, but were still functional as long as structural dynamics of the surrounding protein region was preserved. Loss of function and significant alternation of structural propensities was observed in two variants – H141A and H157N, as indicated by the Agadir algorithm. This result highlights the great likelihood of these residues to comprise important functional parts of the domain.

The analyses were performed using the repressor domain expressed alone (fused to a 6×His-tag) and not as part of the full-length Gli3 protein or as a larger part encompassing the repressor domain and the surrounding region. It cannot be ruled out that in the composition of the entire protein, this domain has a diminished spatial freedom due to stabilizing inter-domain connections and is therefore more folded than was observed (Batey and Clarke 2008). However, homologous domains are present in different proteins, meaning that the information about their function and folding is mainly coded within the domain sequence, and only to a minor extent by the surrounding context of the protein. It cannot be ruled out that although the domain appears disordered on its own, it can undergo structural changes upon interaction with a binding partner in the context of the full-length protein only. Nevertheless, as the Gli3RD domain on its own has repressor function, the latter is not likely.

CONCLUSIONS

1. **The anti-Gli3 single chain recombinant intrabody recognises cytoplasmic human and murine Gli3.**
2. **The transcriptional repressor domain of human Gli proteins is located between aa residues G29 to G170 in Gli2 and G106 to E236 in Gli3.**
3. **The mechanism of transcriptional repression of this domain is independent of histone deacetylation.**
4. **The 3D-structure of the transcriptional repressor domain of Gli3 is intrinsically disordered.**
5. **Gli3 repressor domain performs autonomously, independent of binding to Ski, DNA and zinc.**
6. **The residues H141 and H157 are situated within functionally important parts of the domain.**

REFERENCES

Allen, B. L., T. Tenzen and A. P. McMahon (2007). The Hedgehog-binding proteins Gas1 and Cdo cooperate to positively regulate Shh signaling during mouse development. Genes Dev 21(10): 1244-57.

Bai, C. B. and A. L. Joyner (2001). Gli1 can rescue the in vivo function of Gli2. Development 128(24): 5161-72.

Balaskas, N., A. Ribeiro, J. Panovska, E. Dessaud, N. Sasai, K. M. Page, J. Briscoe and V. Ribes (2012). Gene regulatory logic for reading the Sonic Hedgehog signaling gradient in the vertebrate neural tube. Cell 148(1-2): 273-84.

Batey, S. and J. Clarke (2008). The folding pathway of a single domain in a multidomain protein is not affected by its neighbouring domain. J Mol Biol **378**(2): 297-301.

Biesecker, L. G. (2006). What you can learn from one gene: GLI3. J Med Genet **43**(6): 465-9.

Briscoe, J., A. Pierani, T. M. Jessell and J. Ericson (2000). A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. Cell 101(4): 435-45.

Briscoe, J. and P. P. Therond (2013). The mechanisms of Hedgehog signalling and its roles in development and disease. Nat Rev Mol Cell Biol 14(7): 416-29.

Butler, D. C. and A. Messer (2011). Bifunctional anti-huntingtin proteasome-directed intrabodies mediate efficient degradation of mutant huntingtin exon 1 protein fragments. PLoS One **6**(12): e29199.

Chamberlain, C. E., J. Jeong, C. Guo, B. L. Allen and A. P. McMahon (2008). Notochord-derived Shh concentrates in close association with the apically positioned basal body in neural target cells and forms a dynamic gradient during neural patterning. Development 135(6): 1097-106.

Chang, D. T., A. Lopez, D. P. von Kessler, C. Chiang, B. K. Simandl, R. Zhao, M. F. Seldin, J. F. Fallon and P. A. Beachy (1994). Products, genetic linkage and limb patterning activity of a murine hedgehog gene. Development 120(11): 3339-53.

Chemes, L. B., L. G. Alonso, M. G. Noval and G. de Prat-Gay (2012). Circular dichroism techniques for the analysis of intrinsically disordered proteins and domains. Methods Mol Biol 895: 387-404.

Chen, M. H., C. W. Wilson, Y. J. Li, K. K. Law, C. S. Lu, R. Gacayan, X. Zhang, C. C. Hui and P. T. Chuang (2009). Cilium-independent regulation of

Gli protein function by Sufu in Hedgehog signaling is evolutionarily conserved. Genes Dev 23(16): 1910-28.
Cheng, S. Y. and J. M. Bishop (2002). Suppressor of Fused represses Gli-mediated transcription by recruiting the SAP18-mSin3 corepressor complex. Proc Natl Acad Sci U S A **99**(8): 5442-7.
Chiang, C., Y. Litingtung, E. Lee, K. E. Young, J. L. Corden, H. Westphal and P. A. Beachy (1996). Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. Nature 383(6599): 407-13.
Cortellino, S., C. Wang, B. Wang, M. R. Bassi, E. Caretti, D. Champeval, A. Calmont, M. Jarnik, J. Burch, K. S. Zaret, L. Larue and A. Bellacosa (2009). Defective ciliogenesis, embryonic lethality and severe impairment of the Sonic Hedgehog pathway caused by inactivation of the mouse complex A intraflagellar transport gene Ift122/Wdr10, partially overlapping with the DNA repair gene Med1/Mbd4. Dev Biol 325(1): 225-37.
Dai, P., H. Akimaru, Y. Tanaka, T. Maekawa, M. Nakafuku and S. Ishii (1999). Sonic Hedgehog-induced activation of the Gli1 promoter is mediated by GLI3. J Biol Chem 274(12): 8143-52.
Dai, P., T. Shinagawa, T. Nomura, J. Harada, S. C. Kaul, R. Wadhwa, M. M. Khan, H. Akimaru, H. Sasaki, C. Colmenares and S. Ishii (2002). Ski is involved in transcriptional regulation by the repressor and full-length forms of Gli3. Genes Dev **16**(22): 2843-8.
Dasgupta, I., L. Sanglas, J. J. Enghild and I. Lindberg (2012). The neuroendocrine protein 7B2 is intrinsically disordered. Biochemistry 51(38): 7456-64.
Dessaud, E., L. L. Yang, K. Hill, B. Cox, F. Ulloa, A. Ribeiro, A. Mynett, B. G. Novitch and J. Briscoe (2007). Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. Nature 450(7170): 717-20.
Ding, Q., J. Motoyama, S. Gasca, R. Mo, H. Sasaki, J. Rossant and C. C. Hui (1998). Diminished Sonic hedgehog signaling and lack of floor plate differentiation in Gli2 mutant mice. Development 125(14): 2533-43.
Drobnak, I., N. De Jonge, S. Haesaerts, G. Vesnaver, R. Loris and J. Lah (2013). Energetic basis of uncoupling folding from binding for an intrinsically disordered protein. J Am Chem Soc 135(4): 1288-94.
Dunaeva, M., P. Michelson, P. Kogerman and R. Toftgard (2003). Characterization of the physical interaction of Gli proteins with SUFU proteins. J Biol Chem **278**(7): 5116-22.
Dyson, H. J. and P. E. Wright (2005). Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6(3): 197-208.
Ericson, J., S. Morton, A. Kawakami, H. Roelink and T. M. Jessell (1996). Two critical periods of Sonic Hedgehog signaling required for the specification of motor neuron identity. Cell 87(4): 661-73.

Fuxreiter, M., I. Simon and S. Bondos (2011). Dynamic protein-DNA recognition: beyond what can be seen. Trends Biochem Sci 36(8): 415-23.
Fuxreiter, M. (2012). Fuzziness: linking regulation to protein dynamics. Mol Biosyst 8(1): 168-77.
Galli, A., D. Robay, M. Osterwalder, X. Bao, J. D. Benazet, M. Tariq, R. Paro, S. Mackem and R. Zeller (2010). Distinct roles of Hand2 in initiating polarity and posterior Shh expression during the onset of mouse limb bud development. PLoS Genet 6(4): e1000901.
Gal-Tanamy, M., R. Zemel, L. Bachmatov, R. K. Jangra, A. Shapira, R. A. Villanueva, M. Yi, S. M. Lemon, I. Benhar and R. Tur-Kaspa (2010). Inhibition of protease-inhibitor-resistant hepatitis C virus replicons and infectious virus by intracellular intrabodies. Antiviral Res **88**(1): 95-106.
Goodrich, L. V., R. L. Johnson, L. Milenkovic, J. A. McMahon and M. P. Scott (1996). Conservation of the hedgehog/patched signaling pathway from flies to mice: induction of a mouse patched gene by Hedgehog. Genes Dev 10(3): 301-12.
Gritli-Linde, A., P. Lewis, A. P. McMahon and A. Linde (2001). The whereabouts of a morphogen: direct evidence for short- and graded long-range activity of hedgehog signaling peptides. Dev Biol 236(2): 364-86.
Gunasekaran, K., C. J. Tsai, S. Kumar, D. Zanuy and R. Nussinov (2003). Extended disordered proteins: targeting function with less scaffold. Trends Biochem Sci 28(2): 81-5.
Haramis, A. G., J. M. Brown and R. Zeller (1995). The limb deformity mutation disrupts the SHH/FGF-4 feedback loop and regulation of 5' HoxD genes during limb pattern formation. Development 121(12): 4237-45.
Haycraft, C. J., B. Banizs, Y. Aydin-Son, Q. Zhang, E. J. Michaud and B. K. Yoder (2005). Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. PLoS Genet 1(4): e53.
Hegyi, H., E. Schad and P. Tompa (2007). Structural disorder promotes assembly of protein complexes. BMC Struct Biol 7: 65.
Hinds, M. G., C. Smits, R. Fredericks-Short, J. M. Risk, M. Bailey, D. C. Huang and C. L. Day (2007). Bim, Bad and Bmf: intrinsically unstructured BH3-only proteins that undergo a localized conformational change upon binding to prosurvival Bcl-2 targets. Cell Death Differ **14**(1): 128-36.
Huang, X., Y. Litingtung and C. Chiang (2007). Region-specific requirement for cholesterol modification of sonic hedgehog in patterning the telencephalon and spinal cord. Development 134(11): 2095-105.
Huangfu, D., A. Liu, A. S. Rakeman, N. S. Murcia, L. Niswander and K. V. Anderson (2003). Hedgehog signalling in the mouse requires intraflagellar transport proteins. Nature 426(6962): 83-7.
Hui, C. C., D. Slusarski, K. A. Platt, R. Holmgren and A. L. Joyner (1994). Expression of three mouse homologs of the Drosophila segment polarity gene

cubitus interruptus, Gli, Gli-2, and Gli-3, in ectoderm- and mesoderm-derived tissues suggests multiple roles during postimplantation development. Dev Biol 162(2): 402-13.
Humke, E. W., K. V. Dorn, L. Milenkovic, M. P. Scott and R. Rohatgi (2010). The output of Hedgehog signaling is controlled by the dynamic association between Suppressor of Fused and the Gli proteins. Genes Dev 24(7): 670-82.
Huntzicker, E. G., I. S. Estay, H. Zhen, L. A. Lokteva, P. K. Jackson and A. E. Oro (2006). Dual degradation signals control Gli protein stability and tumor formation. Genes Dev **20**(3): 276-81.
Janody, F., Z. Martirosyan, A. Benlali and J. E. Treisman (2003). Two subunits of the Drosophila mediator complex act together to control cell affinity. Development **130**(16): 3691-701.
Jessell, T. M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat Rev Genet 1(1): 20-9.
Jia, J., A. Kolterud, H. Zeng, A. Hoover, S. Teglund, R. Toftgard and A. Liu (2009). Suppressor of Fused inhibits mammalian Hedgehog signaling in the absence of cilia. Dev Biol 330(2): 452-60.
Johnson, D. R. (1967). Extra-toes: anew mutant gene causing multiple abnormalities in the mouse. J Embryol Exp Morphol 17(3): 543-81.
Kaesler, S., B. Luscher and U. Ruther (2000). Transcriptional activity of GLI1 is negatively regulated by protein kinase A. Biol Chem 381(7): 545-51.
Karlin, S. and Z. Y. Zhu (1997). Classification of mononuclear zinc metal sites in protein structures. Proc Natl Acad Sci U S A **94**(26): 14231-6.
Kmita, M., B. Tarchini, J. Zakany, M. Logan, C. J. Tabin and D. Duboule (2005). Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function. Nature 435(7045): 1113-6.
Kogerman, P., T. Grimm, L. Kogerman, D. Krause, A. B. Unden, B. Sandstedt, R. Toftgard and P. G. Zaphiropoulos (1999). Mammalian suppressor-of-fused modulates nuclear-cytoplasmic shuttling of Gli-1. Nat Cell Biol 1(5): 312-9.
Kraus, P., D. Fraidenraich and C. A. Loomis (2001). Some distal limb structures develop in mice lacking Sonic hedgehog signaling. Mech Dev 100(1): 45-58.
Laht, S., K. Meerits, H. Altroff, H. Faust, R. Tsaney, P. Kogerman, L. Jarvekulg, V. Paalme, A. Valkna and S. Timmusk (2008). Generation and characterization of a single-chain Fv antibody against Gli3, a hedgehog signaling pathway transcription factor. Hybridoma (Larchmt) **27**(3): 167-74.
Li, L., C. S. Gondi, D. H. Dinh, W. C. Olivero, M. Gujrati and J. S. Rao (2007). Transfection with anti-p65 intrabody suppresses invasion and angiogenesis in glioma cells by blocking nuclear factor-kappaB transcriptional activity. Clin Cancer Res **13**(7): 2178-90.
Libich, D. S. and G. Harauz (2008). Solution NMR and CD spectroscopy of an intrinsically disordered, peripheral membrane protein: evaluation of aqueous and membrane-mimetic solvent conditions for studying the conformational

adaptability of the 18.5 kDa isoform of myelin basic protein (MBP). Eur Biophys J 37(6): 1015-29.
Lipinski, R. J., J. J. Gipp, J. Zhang, J. D. Doles and W. Bushman (2006). Unique and complimentary activities of the Gli transcription factors in Hedgehog signaling. Exp Cell Res **312**(11): 1925-38.
Litingtung, Y. and C. Chiang (2000). Specification of ventral neuron types is mediated by an antagonistic interaction between Shh and Gli3. Nat Neurosci 3(10): 979-85.
Litingtung, Y., R. D. Dahn, Y. Li, J. F. Fallon and C. Chiang (2002). Shh and Gli3 are dispensable for limb skeleton formation but regulate digit number and identity. Nature 418(6901): 979-83.
Liu, A., B. Wang and L. A. Niswander (2005). Mouse intraflagellar transport proteins regulate both the activator and repressor functions of Gli transcription factors. Development 132(13): 3103-11.
Logan, C., A. Hornbruch, I. Campbell and A. Lumsden (1997). The role of Engrailed in establishing the dorsoventral axis of the chick limb. Development 124(12): 2317-24.
Loomis, C. A., E. Harris, J. Michaud, W. Wurst, M. Hanks and A. L. Joyner (1996). The mouse Engrailed-1 gene and ventral limb patterning. Nature 382(6589): 360-3.
Lopez-Martinez, A., D. T. Chang, C. Chiang, J. A. Porter, M. A. Ros, B. K. Simandl, P. A. Beachy and J. F. Fallon (1995). Limb-patterning activity and restricted posterior localization of the amino-terminal product of Sonic hedgehog cleavage. Curr Biol 5(7): 791-6.
Lopez-Rios, J., D. Speziale, D. Robay, M. Scotti, M. Osterwalder, G. Nusspaumer, A. Galli, G. A. Hollander, M. Kmita and R. Zeller (2012). GLI3 constrains digit number by controlling both progenitor proliferation and BMP-dependent exit to chondrogenesis. Dev Cell 22(4): 837-48.
Malik, S. and R. G. Roeder (2010). The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. Nat Rev Genet **11**(11): 761-72.
Matise, M. P., D. J. Epstein, H. L. Park, K. A. Platt and A. L. Joyner (1998). Gli2 is required for induction of floor plate and adjacent cells, but not most ventral neurons in the mouse central nervous system. Development 125(15): 2759-70.
Meszaros, B., P. Tompa, I. Simon and Z. Dosztanyi (2007). Molecular principles of the interactions of disordered proteins. J Mol Biol 372(2): 549-61.
Mohan, A., C. J. Oldfield, P. Radivojac, V. Vacic, M. S. Cortese, A. K. Dunker and V. N. Uversky (2006). Analysis of molecular recognition features (MoRFs). J Mol Biol 362(5): 1043-59.

Naruse, I., E. Ueta, Y. Sumino, M. Ogawa and S. Ishikiriyama (2010). Birth defects caused by mutations in human GLI3 and mouse Gli3 genes. Congenit Anom (Kyoto) **50**(1): 1-7.

Nomura, T., M. M. Khan, S. C. Kaul, H. D. Dong, R. Wadhwa, C. Colmenares, I. Kohno and S. Ishii (1999). Ski is a component of the histone deacetylase complex required for transcriptional repression by Mad and thyroid hormone receptor. Genes Dev **13**(4): 412-23.

Ohuchi, H., T. Nakagawa, A. Yamamoto, A. Araga, T. Ohata, Y. Ishimaru, H. Yoshioka, T. Kuwana, T. Nohno, M. Yamasaki, N. Itoh and S. Noji (1997). The mesenchymal factor, FGF10, initiates and maintains the outgrowth of the chick limb bud through interaction with FGF8, an apical ectodermal factor. Development 124(11): 2235-44.

Pan, Y., C. B. Bai, A. L. Joyner and B. Wang (2006). Sonic hedgehog signaling regulates Gli2 transcriptional activity by suppressing its processing and degradation. Mol Cell Biol 26(9): 3365-77.

Pan, Y., C. Wang and B. Wang (2009). Phosphorylation of Gli2 by protein kinase A is required for Gli2 processing and degradation and the Sonic Hedgehog-regulated mouse development. Dev Biol 326(1): 177-89.

Park, H. L., C. Bai, K. A. Platt, M. P. Matise, A. Beeghly, C. C. Hui, M. Nakashima and A. L. Joyner (2000). Mouse Gli1 mutants are viable but have defects in SHH signaling in combination with a Gli2 mutation. Development 127(8): 1593-605.

Parr, B. A. and A. P. McMahon (1995). Dorsalizing signal Wnt-7a required for normal polarity of D-V and A-P axes of mouse limb. Nature 374(6520): 350-3.

Pearse, R. V., 2nd, L. S. Collier, M. P. Scott and C. J. Tabin (1999). Vertebrate homologs of Drosophila suppressor of fused interact with the gli family of transcriptional regulators. Dev Biol 212(2): 323-36.

Persson, M., D. Stamataki, P. te Welscher, E. Andersson, J. Bose, U. Ruther, J. Ericson and J. Briscoe (2002). Dorsal-ventral patterning of the spinal cord requires Gli3 transcriptional repressor activity. Genes Dev 16(22): 2865-78.

Pierani, A., S. Brenner-Morton, C. Chiang and T. M. Jessell (1999). A sonic hedgehog-independent, retinoid-activated pathway of neurogenesis in the ventral spinal cord. Cell 97(7): 903-15.

Ribes, V. and J. Briscoe (2009). Establishing and interpreting graded Sonic Hedgehog signaling during vertebrate neural tube patterning: the role of negative feedback. Cold Spring Harb Perspect Biol 1(2): a002014.

Riddle, R. D., R. L. Johnson, E. Laufer and C. Tabin (1993). Sonic hedgehog mediates the polarizing activity of the ZPA. Cell 75(7): 1401-16.

Rinaldi, A. S., G. Freund, D. Desplancq, A. P. Sibler, M. Baltzinger, N. Rochel, Y. Mely, P. Didier and E. Weiss (2013). The use of fluorescent intrabodies to

detect endogenous gankyrin in living cancer cells. Exp Cell Res **319**(6): 838-49.
Rohatgi, R., L. Milenkovic and M. P. Scott (2007). Patched1 regulates hedgehog signaling at the primary cilium. Science 317(5836): 372-6.
Rosello-Diez, A., M. A. Ros and M. Torres (2011). Diffusible signals, not autonomous mechanisms, determine the main proximodistal limb subdivision. Science 332(6033): 1086-8.
Sasaki, H., Y. Nishizaki, C. Hui, M. Nakafuku and H. Kondoh (1999). Regulation of Gli2 and Gli3 activities by an amino-terminal repression domain: implication of Gli2 and Gli3 as primary mediators of Shh signaling. Development **126**(17): 3915-24.
Schlessinger, A., C. Schaefer, E. Vicedo, M. Schmidberger, M. Punta and B. Rost (2011). Protein disorder--a breakthrough invention of evolution?" Curr Opin Struct Biol 21(3): 412-8.
Shin, S. H., P. Kogerman, E. Lindstrom, R. Toftgard and L. G. Biesecker (1999). GLI3 mutations in human disorders mimic Drosophila cubitus interruptus protein functions and localization. Proc Natl Acad Sci U S A **96**(6): 2880-4.
Shoemaker, B. A., J. J. Portman and P. G. Wolynes (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc Natl Acad Sci U S A 97(16): 8868-73.
Spolar, R. S. and M. T. Record, Jr. (1994). Coupling of local folding to site-specific binding of proteins to DNA. Science 263(5148): 777-84.
Summerbell, D. (1974). A quantitative analysis of the effect of excision of the AER from the chick limb-bud. J Embryol Exp Morphol 32(3): 651-60.
Sun, X., F. V. Mariani and G. R. Martin (2002). Functions of FGF signalling from the apical ectodermal ridge in limb development. Nature 418(6897): 501-8.
Tantos, A., K. H. Han and P. Tompa (2012). Intrinsic disorder in cell signaling and gene transcription. Mol Cell Endocrinol 348(2): 457-65.
Tempe, D., M. Casas, S. Karaz, M. F. Blanchet-Tournier and J. P. Concordet (2006). Multisite protein kinase A and glycogen synthase kinase 3beta phosphorylation leads to Gli3 ubiquitination by SCFbetaTrCP. Mol Cell Biol 26(11): 4316-26.
Tickle, C., D. Summerbell and L. Wolpert (1975). Positional signalling and specification of digits in chick limb morphogenesis. Nature 254(5497): 199-202.
Tompa, P., Z. Dosztanyi and I. Simon (2006). Prevalent structural disorder in E. coli and S. cerevisiae proteomes. J Proteome Res **5**(8): 1996-2000.
Tompa, P. (2010). Structure and function of intrinsically disordered proteins. Boca Raton, CRC Press.

Tsanev, R., P. Tiigimagi, P. Michelson, M. Metsis, T. Osterlund and P. Kogerman (2009). Identification of the gene transcription repressor domain of Gli3. FEBS Lett **583**(1): 224-8.

Tsanev, R., K. Vanatalu, J. Jarvet, R. Tanner, K. Laur, P. Tiigimagi, B. B. Kragelund, T. Osterlund and P. Kogerman (2013). The transcriptional repressor domain of Gli3 is intrinsically disordered. PLoS One **8**(10): e76972.

Tukachinsky, H., L. V. Lopez and A. Salic (2010). A mechanism for vertebrate Hedgehog signaling: recruitment to cilia and dissociation of SuFu-Gli protein complexes. J Cell Biol 191(2): 415-28.

Tuson, M., M. He and K. V. Anderson (2011). Protein kinase A acts at the basal body of the primary cilium to prevent Gli2 activation and ventralization of the mouse neural tube. Development 138(22): 4921-30.

Vokes, S. A., H. Ji, W. H. Wong and A. P. McMahon (2008). A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. Genes Dev 22(19): 2651-63.

Wanek, N., K. Muneoka, G. Holler-Dinsmore, R. Burton and S. V. Bryant (1989). A staging system for mouse limb development. J Exp Zool 249(1): 41-9.

Wang, B., J. F. Fallon and P. A. Beachy (2000). Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb. Cell 100(4): 423-34.

Wen, X., C. K. Lai, M. Evangelista, J. A. Hongo, F. J. de Sauvage and S. J. Scales (2010). Kinetics of hedgehog-dependent full-length Gli3 accumulation in primary cilia and subsequent degradation. Mol Cell Biol 30(8): 1910-22.

Wijgerde, M., J. A. McMahon, M. Rule and A. P. McMahon (2002). A direct requirement for Hedgehog signaling for normal specification of all ventral progenitor domains in the presumptive mammalian spinal cord. Genes Dev 16(22): 2849-64.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. J Theor Biol 25(1): 1-47.

Xie, H., S. Vucetic, L. M. Iakoucheva, C. J. Oldfield, A. K. Dunker, V. N. Uversky and Z. Obradovic (2007). Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J Proteome Res 6(5): 1882-98.

Zeller, R., J. Lopez-Rios and A. Zuniga (2009). Vertebrate limb bud development: moving towards integrative analysis of organogenesis. Nat Rev Genet 10(12): 845-58.

Zhou, H., S. Kim, S. Ishii and T. G. Boyer (2006). Mediator modulates Gli3-dependent Sonic hedgehog signaling. Mol Cell Biol **26**(23): 8667-82.

Zhu, J., E. Nakamura, M. T. Nguyen, X. Bao, H. Akiyama and S. Mackem (2008). Uncoupling Sonic hedgehog control of pattern and expansion of the developing limb bud. Dev Cell 14(4): 624-32.

ACKNOWLEDGEMENTS

I am indebted to all my colleagues I have worked with. You gave me strength and inspiration all the way here.

I thank my supervisor Priit Kogerman for building up the nice laboratory we were working in and for giving me the opportunity to work on this topic in his group.

Dear Torben, thank you for supervising me. I am grateful to you for teaching me how to organise my work and express myself in writing. Thank you for advising me on the manuscripts of my articles and the thesis.

Dear Lilian, I am much obliged to you for your advice, assistance and support. I enjoyed the fruitful discussions we had and I am grateful for the work you have done correcting this thesis.

Cordial thanks to my colleague Kalju. Your friendship has been very precious to me.

Dear Rix, thank you for your time and advice on the last manuscript.

Dear Kersti, thank you for your care. You encouraged me in my moments of weakness, and gave me the strength to continue. A great part of what I have done has been done thanks to you. I could not give much back to you though.

Dear Olya, Alla and Indrek, we had some big ideas in our tiny office. Thank you for being there!

I owe much to all the members of our group for the friendly atmosphere in the lab, for practical help provided and theoretical discussions we had: Pirett, Piret, Lagle, Lembi, Taavi, Katrin, Anne, Pille, Illar, Kristi.

Dear Birthe, thank you for your energy and help with ideas, experiments and writing. I liked working at your lab and appreciate your hospitality. My gratitude belongs also to Signe for the technical assistance during my stay at the lab.

I am happy that I had the chance to know and work with Jüri. I am grateful for your time and the patience you showed when explaining to me all the mystic experiments you were doing. Thank you for the comments on the manuscripts.

Dear Viiu, working with you has been a pleasure, thank you.

Dear Sirje, I am much obliged to you for including me in the intrabody project. This made completing this thesis possible.

I express my gratitude to Andres Veske for the support during the last period of my study.

Mu kallid Martin, Peeter ja Kristjan, ma armastan teid.

PUBLICATION I

HYBRIDOMA
Volume 27, Number 3, 2008

DOI: 10.1089/hyb.2008.0004

Generation and Characterization of a Single-Chain Fv Antibody Against Gli3, a Hedgehog Signaling Pathway Transcription Factor

Silja Laht,[1] Kati Meerits,[1] Harri Altroff,[1] Helena Faust,[1] Robert Tsaney,[2] Priit Kogerman,[2] Lilian Järvekülg,[2] Viiu Paalme,[2] Andres Valkna,[1] and Sirje Timmusk[1]

Abstract

Gli3 is a key regulator of development, controlling multiple patterning steps. Here we report the generation of a scFv antibody specific to the repressor domain of human Gli3. We show that this scFv retains the binding capacity of its parent anti-Gli3 monoclonal antibody derived from hybridoma clone 5E1. When expressed in mammalian cells, the anti-Gli3 scFv co-localizes with intracellular Gli3. Immunocytochemical staining of the intrabody in Gli3-positive TM4 cells shows a distinct perinuclear cytoplasmic localization. Such a scFv constitutes a useful tool for studying transcriptional regulation of the hedgehog pathway in mammals and offers a starting point for developing novel Gli-related therapeutic intrabodies.

Introduction

HEDGEHOG SIGNALING IS ONE OF THE MAIN pathways regulating embryogenesis. Vertebrate sonic hedgehog (Shh) signaling is mediated by transcription factors belonging to the Gli family—Gli1, Gli2, and Gli3. In the presence of the Shh ligand, full-length Gli proteins activate transcription of multiple target genes. In the absence of the ligand, Gli proteins are processed by proteases, resulting in inhibition of transcription. Positional information along the anteroposterior Shh gradient is thus translated into distinct intracellular Gli activities. For instance, limb patterning in vertebrates is determined by the counteraction between Shh and transcription factor Gli3.[1] In the anterior region where Shh level is low, protease-cleaved Gli3 acts predominantly as a repressor of Shh signaling, whereas in the posterior region, Shh inhibits Gli3 processing and full-length Gli3 is accumulated. Thus, Shh controls the balance of the activator and repressor forms of Gli3 across a gradient along the anteroposterior axis of the limb bud. These opposing concentration gradients of Shh and Gli3 specify, for example, the number and identity of digits.[2–6] Abnormal Gli3 expression/function causes polydactyly and is associated with a number of human diseases such as Greig cephalopolysyndactyly syndrome,[7–9] Pallister-Hall syndrome,[10] and postaxial polydactyly type A.[1,11] Shin et al.[12] explain the different phenotypes associated with these syndromes by inclusion or loss of functional domains in Gli3 due to distinct mutations. In addition to their role in development, aberrant regulation of Gli proteins has been implicated in many different malignancies including skin cancer[13] and prostate cancer.[14] This makes these proteins desirable targets for therapeutic intervention aimed at controlling Gli-dependent signaling.

The fact that intracellular antigens such as Gli3 are normally not accessible to conventional antibodies places important constraints on their diagnostic and therapeutic utility. This limitation can be overcome by exploiting single-chain variable fragment (scFv) antibodies derived from full-length antibodies by gene engineering. Such scFv retain the binding specificity of the original immunoglobulin and can be expressed in mammalian cells as scFv intrabodies, which are potentially able to alter the folding, subcellular localization, interaction parameters, and/or functional properties of their target antigen *in situ*.

We have recently generated a monoclonal antibody, MAb 5E1,[15] that recognizes the putative repressive motif (residues 106-237) of human Gli3 (Gli3pRM).[16] Here we describe construction and functional characterization of a recombinant scFv based on this antibody. The created anti-Gli3 scFv represents an efficient new tool for investigating the

[1]Celecure AS, Akadeemia tee, Tallinn, Estonia.
[2]Institute of Gene Technology, Tallinn University of Technology, Akadeemia tee, Tallinn, Estonia.

function of Gli3, and may also prove useful for therapeutic applications involving modulation of Gli3 activity.

Materials and Methods

Generation of anti-Gli3 scFv constructs

The V_H and V_L coding regions of anti-Gli3 MAb 5E1[15] were isolated from hybridoma mRNA by nested PCR using the SMART forward primer from the SMART cDNA Synthesis Kit (Clontech, Mountain View, CA) and reverse primers specific to the constant regions of the μ heavy and κ light chains ($C_{\mu 1}$-outer, 5′-TGAAGGA(G/A)ATGGTGCTGGGCAGGAAGT-3′; $C_{\mu 1}$-inner, 5′-CTCGCAGGAGACGAGGGGGAAGACA-3′; C_κ, 5′-CTTGGTCCCCCCTCCGAA-3′). The gel-purified PCR products were cloned into the pGemT-Easy vector (Promega, Madison, WI) and sequenced. Recombinant anti-Gli3 scFv was constructed by joining the obtained V_H and V_L fragments via a synthetic 57-bp linker sequence containing EcoRI and SalI sites (5′-GAATTCTCCTCCGGAGGAGGGGGGTCCGGTGGAGGCGGCTCCGGTGGTTCCGTCGAC-3′), using clone-specific variable domain primers incorporating appropriate restriction sites for domain fusion and subcloning: V_H(F)BamHI-pET, 5′-CTGGG-ATCCGAGGGTTCAGCTGCAGCAG-3′; V_H(F)BamHI pcDNA, 5′CTGGGATCCAGGGTTCAGCTGCAGCAG-3′; V_H(R)EcoRI, 5′-TATGAATTCTGAGGTTCCTTGTCCCCAG-3′; V_L(F)SalI, 5′-TATGTCGACCAAATTGTTCTCACCCAG-3′ V_L(R)HindIII, 5′-TATAAGCTTGGTCCCCCCTCCGAAC-3′; V_L(R)NotI, 5′-TATGCGGCCGCCTACTTGGTCCCCCCTCCGAAC-3′. The resulting scFv5E1 construct was then inserted between the BamHI and HindIII sites of pET40b (Novagen, Gibbstown, NJ) to create pET40-scFv5E1 for bacterial expression, or between the BamHI and NotI sites of pcDNA3.1/His (Invitrogen, Carlsbad, CA) or pcDNA3.1/GST (kindly provided by Lagle Kasak, Tallinn University of Technology, Estonia) to create pcDNA/His-scFv5E1 and pcDNA/GST-scFv5E1 for mammalian expression. Nucleotide sequence of the generated scFv5E1 has been submitted to GenBank (accession no. EU162129).

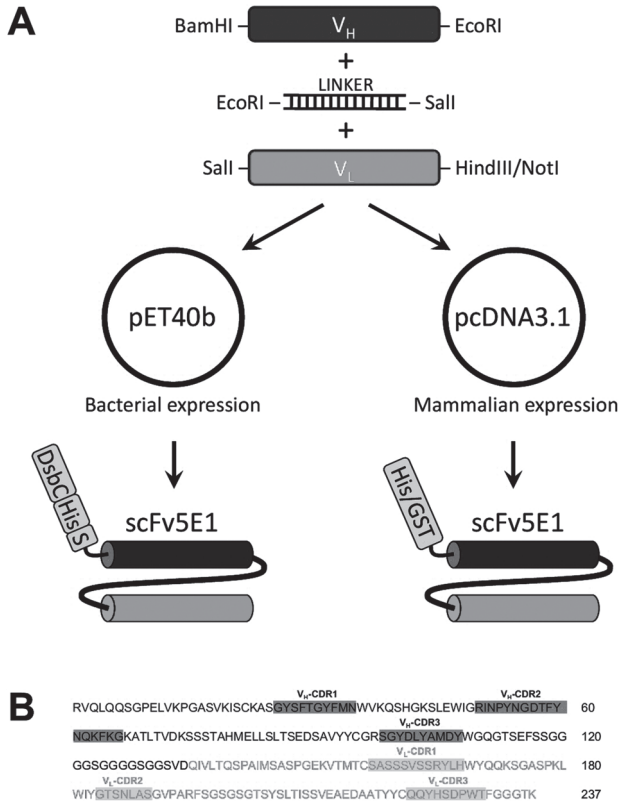

FIG. 1. (**A**) Schematic overview of the construction of anti-Gli3 scFv5E1. Amplified variable regions of the heavy (dark grey) and light (light grey) chain of monoclonal antibody 5E1 were joined by an inert linker and cloned into expression vectors to produce scFv in fusion with indicated N-terminal tags. (**B**) Amino acid sequence of the scFv5E1 construct (V_H, black; V_L, grey), with complementarity-determining regions (CDRs) boxed and marked.

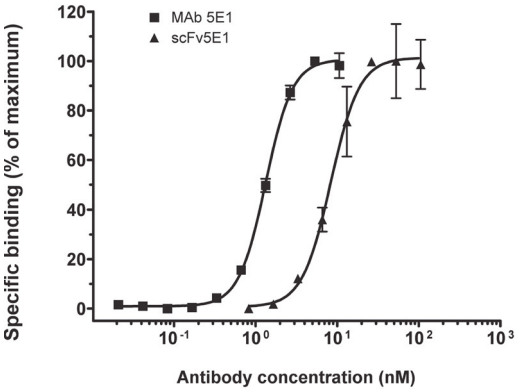

FIG. 2. Solid-phase binding of bacterially expressed DsbC-tagged scFv5E1 (triangles) and MAb 5E1 (squares) to purified recombinant Gli3pRM as measured by ELISA. Results are normalized and expressed as percentages of maximum binding activity.

Plasmids for mammalian expression of Gli3pRM and Gli2

Vector pFA-Gli3pRM for mammalian expression of the repression domain (residues 106-246) of human Gli3 in fusion with an N-terminal Gal4 tag (Gal4-Gli3pRM) was constructed by amplifying the corresponding fragment from the previously described GLI3-FL construct containing full-length Gli3 cDNA[12] and inserting it between the BamHI and XbaI sites of pFA-CMV (Invitrogen). Plasmid pEGFP-Gli2 for mammalian expression of full-length human Gli2 with a green fluorescent protein (GFP) tag was kindly provided by Olga Bragina (Tallinn University of Technology, Estonia).

Bacterial expression and purification of DsbC-scFv5E1 and DsbC

Recombinant scFv5E1 was expressed from pET40b as a DsbC-fusion protein in *Escherichia coli* strain BL21(DE3)pLysS, cultured at 28°C in TB medium supplemented with 40 μg/mL kanamycin and 30 μg/mL chloramphenicol. In addition to the periplasmic chaperone DsbC, the fusion protein contained an N-terminal S-tag and His-tag to facilitate affinity purification. DsbC-scFv5E1 expression was induced by 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 16 h. The cells were disrupted by freezing and sonication and the suspension cleared by centrifugation at 15,000 g for 30 min. DsbC-scFv5E1 was purified in two steps by passing the cell-free lysate over the Talon Superflow resin (BD Biosciences, Franklin Lakes, NJ) and eluting with 250 mM imidazole, followed by size exclusion chromatography on a Superdex 200 column (GE Healthcare, Waukesha, WI) equilibrated with PBS. DsbC was expressed from the pET40b plasmid and purified in an analogous manner. The purity and stability of the obtained proteins were assessed by SDS-polyacrylamide gel electrophoresis and Western blot analysis.

ELISA

ELISAs were performed with recombinant human Gli3pRM, expressed from the pET11Gli3pRM plasmid and purified as described previously.[15] MAb 5E1 was purified from the cell culture supernatant of anti-Gli3 hybridoma clone 5E1 as detailed elsewhere.[15] 96-well flat-bottomed Nunc Maxisorp plates (Nunc, Roskilde, Denmark) were coated with Gli3pRM at 50 μg/mL and left overnight at 4°C. The plates were then washed three times with PBS and blocked with 2% bovine serum albumin (BSA) in PBS for 1 h at 37°C. The wells were washed as above and the plates incubated for 1 h at 37°C with MAb 5E1, DsbC-scFv5E1, or DsbC diluted in PBS containing 0.1% BSA. After further washing, either HRP-conjugated goat anti-mouse-IgM antibody (0.2 μg/mL; Santa Cruz Biotechnology, Santa Cruz, CA) or HRP-conjugated polyclonal anti-S-tag antibody (dilution 1:200,000; Abcam, Cambridge, MA) was added in PBS containing 0.1% BSA for the detection of MAb 5E1 or DsbC proteins, respectively. The plates were incubated for 1 h at 25°C, washed as above, and incubated with the Sigma Fast OPD tablet set (Sigma-Aldrich, St. Louis, MO), according to the manufacturer's instructions. The absorbance of the solution was measured at 450 nm. Assays were performed in duplicate. Non-specific binding of the tested proteins to uncoated BSA-blocked wells was measured separately for each

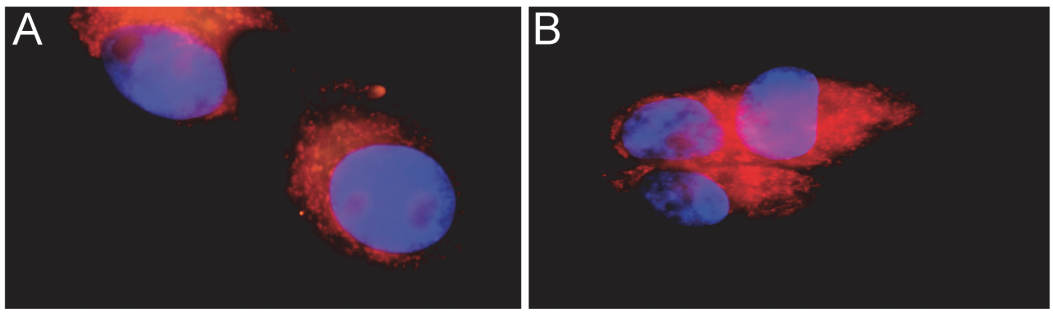

FIG. 3. Expression of His-tagged scFv5E1 in mammalian cells. Anti-His staining (red) reveals granular cytoplasmic expression of the intrabody 24 h after transfection of PK15 (**A**) or Cos-1 (**B**) cells with the expression vector. Nuclei are stained blue.

concentration point and subsequently subtracted from the corresponding values for total binding. Dose-response data from the assays were analyzed by non-linear regression using a sigmoidal curve fit (Prism, GraphPad Software, San Diego, CA). Assay results are expressed as the means ± SEM of three independent experiments.

Cell transfection and immunocytochemistry

Mammalian PK15, Cos-1, or TM4 cells were seeded onto glass cover slips in six-well plates at 5×10^4 cells/well and grown in DMEM supplemented with PEST and 10% FCS for 1 h at 37°C in 5% CO_2. The cells were then (co)transfected with relevant expression vector(s) (pcDNA/His-scFv5E1 or pcDNA/GST-scFv5E1 with or without pFA-Gli3pRM and/or pEGFP-Gli2) using the FuGENE reagent (Roche Diagnostics, Basel, Switzerland), according to the manufacturer's instructions. The cells were incubated for 24 h, washed once with PBS, and fixed with 2% glutaraldehyde, 2% formaldehyde in PBS. Fixed cells were washed again, incubated overnight at 4°C with relevant primary antibody, washed and stained with Texas Red-coupled monoclonal anti-mouse Igκ antibody (dilution 1:2000; Serotech, Oxford, United Kingdom), or FITC-conjugated pig polyclonal anti-rabbit antibody (dilution 1:2000; DAKO, Glostrup, Denmark). Primary antibodies were used at the following dilutions: anti-Gli3 MAb 5E1 hybridoma supernatant, undiluted; anti-His MAb (BD Biosciences), 1:1000; anti-GST MAb (Santa Cruz Biotechnology), 1:2000; and rabbit polyclonal anti-Gal4 antibody (Santa Cruz Biotechnology), 1:4000. Counterstaining of nuclei was performed with Hoechst 33258 dye (Sigma-Aldrich). All fluorescence and phase-contrast microscopy was performed at 1000 × magnification using an Axioplan II imaging fluorescence microscope equipped with appropriate filter sets, an Axiocam charge-coupled device camera, and Axiovision software (Carl Zeiss, Oberkochen, Germany).

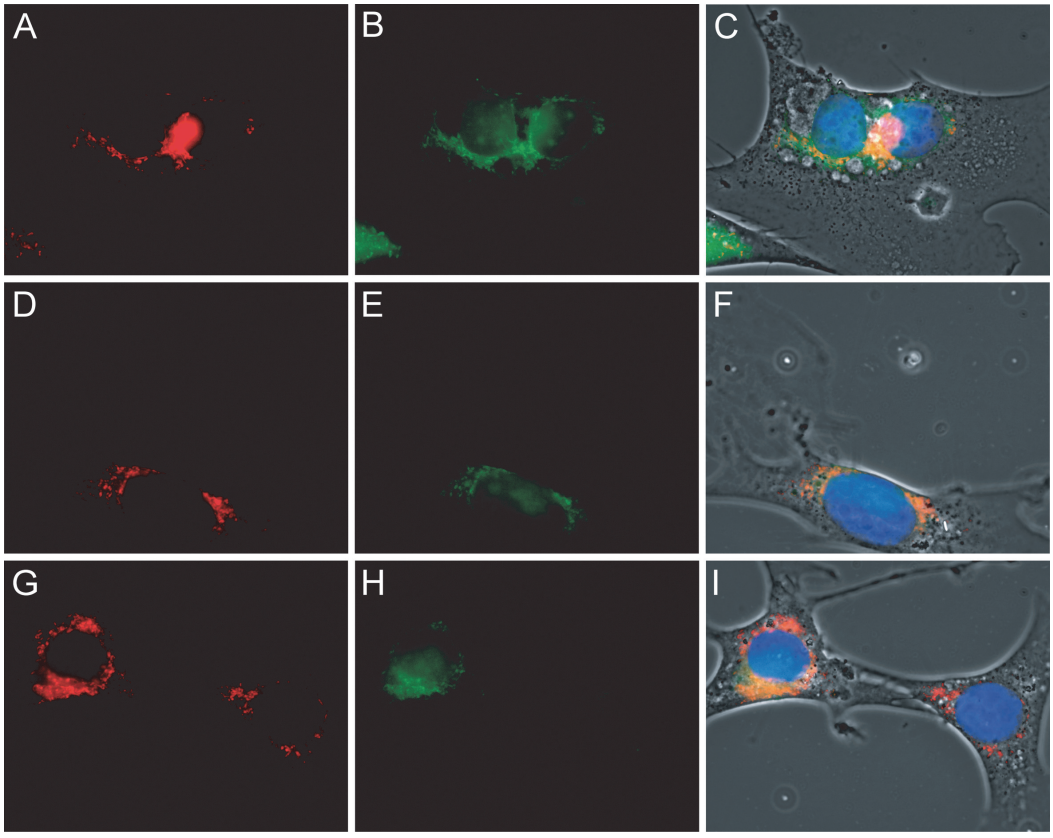

FIG. 4. Co-localization of His-tagged scFv5E1 and recombinant Gli3pRM in PK15 (**A–C**), Cos-1 (**D–F**), and TM4 (**G–I**) cells. A similar perinuclear cytoplasmic distribution is observed for His-tagged scFv5E1 (red) and Gal4-Gli3pRM (green) 24 h after co-transfection with the relevant expression vectors. Overlapping protein localization appears orange on merged phase-contrast images, with nuclei visualized in blue (**C, F, and I**).

Results

Construction and expression of an anti-Gli3 scFv derived from hybridoma clone 5E1

To create expression systems for an anti-Gli3 scFv, we focused our efforts on the recently generated monoclonal antibody 5E1, which is highly specific to human Gli3.[15] The variable regions of the heavy and light chain of MAb 5E1 were amplified from hybridoma transcripts and their sequences determined. The isolated V_H and V_L fragments were connected by an inert flexible glycine-serine linker and the resulting scFv5E1 cloned into prokaryotic and eukaryotic expression vectors (Fig. 1). Bacterial expression and affinity purification of the anti-Gli3 scFv were facilitated by the use of the chaperone and disulfide isomerase DsbC and a histidine tag as fusion partners. This expression system yielded soluble DsbC-scFv5E1 with good stability. Mammalian expression of scFv5E1 with an N-terminal His-tag or GST-tag was achieved by employing pcDNA vectors incorporating the promoter sequence from the immediate early gene of human cytomegalovirus (CMV) for high-level transcription, as well as the polyadenylation signal and transcription termination sequences from the bovine growth hormone gene to enhance RNA stability.

Anti-Gli3 scFv retains the binding affinity of the parent antibody 5E1

In order to compare the affinity of scFv5E1 to that of the parent MAb 5E1, a solid-phase assay was performed using

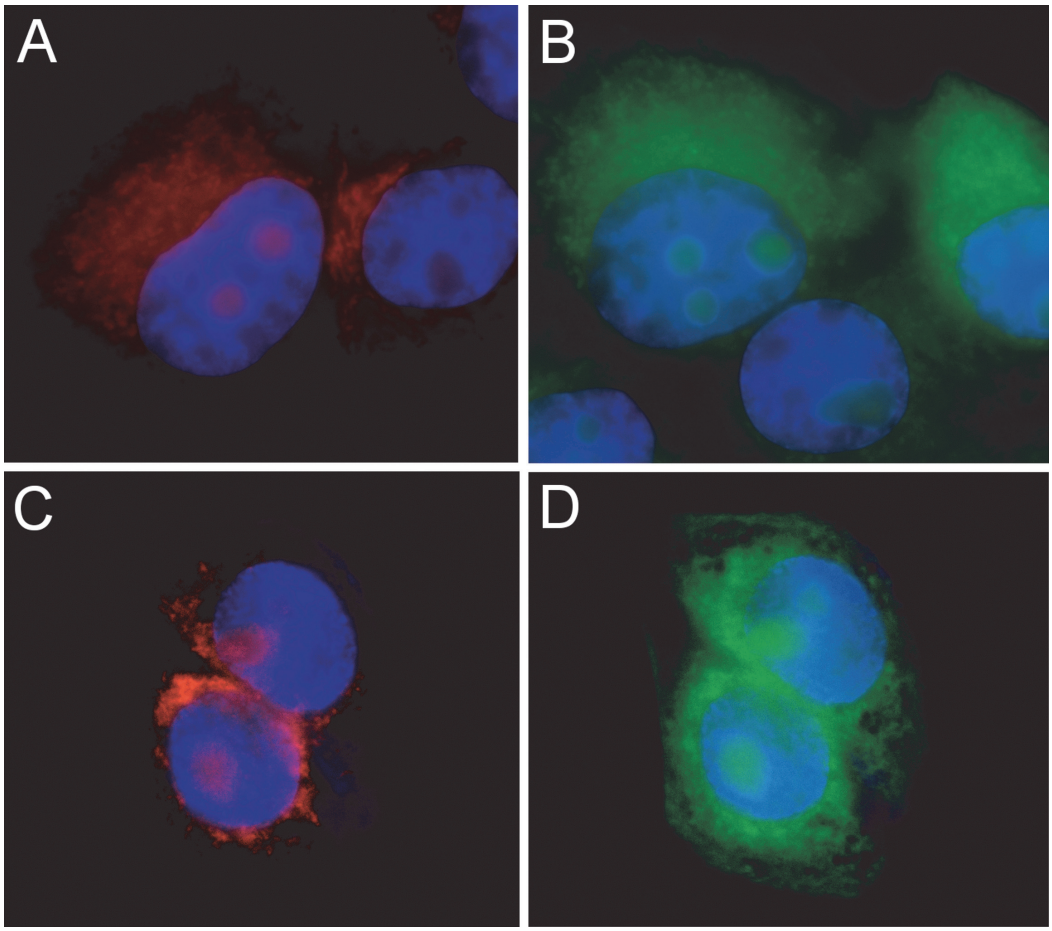

FIG. 5. The anti-Gli3 intrabody does not bind Gli2. Subcellular localization of the anti-Gli3 scFv (red) in PK15 cells co-transfected with vectors encoding His-tagged scFv5E1 and GFP-Gli2 (**A and B**) is distinct from that seen in cells co-expressing scFv5E1, GFP-Gli2, and Gal4-Gli3pRM (**C and D**). Green fluorescence (**B and D**) attests to a robust expression of recombinant GFP-Gli2. Nuclear counterstain appears in blue.

surface-immobilized recombinant human Gli3pRM as a target antigen. Results from the dose-response ELISAs (Fig. 2) show that the bacterially expressed DsbC-scFv5E1 binds to Gli3pRM in a concentration-dependent manner, displaying a binding capacity less than seven-fold lower than that of the original MAb 5E1 (the apparent dissociation constants being 8.4 ± 2.5 nM and 1.3 ± 0.1 nM, respectively). The observed difference in binding presumably reflects reduced avidity of the scFv, stemming from the fact that MAb 5E1 is a pentameric IgM possessing ten identical antigen recognition surfaces, while scFv5E1 retains only one. The affinity of the anti-Gli3 scFv is thus considered to be comparable to that of a single antigen-binding pocket of the original MAb. In addition, parallel testing of DsbC did not yield any specific binding to Gli3pRM, suggesting that this fusion partner is functionally inert and does not interfere with scFv5E1 activity (data not shown).

Anti-Gli3 scFv can be expressed as an intrabody in mammalian cells

To show that it is possible to express the anti-Gli3 scFv in eukaryotic cells as an intrabody, we transfected a pcDNA expression vector encoding His-tagged scFv5E1 into PK15 (porcine kidney epithelial) cells and Cos-1 cells, which do not express endogenous Gli3. Immunostaining of the transfected cells with an anti-His antibody showed that the scFv was stably expressed in the cytoplasm (Fig. 3). The fluorescence pattern suggested that the scFv expression was diffuse in the cytoplasm in both cell lines. No scFv staining was observed in the nucleus or associated with cellular membranes.

Anti-Gli3 scFv co-localizes with intracellular human Gli3

We then expressed His-tagged scFv5E1 together with Gal4-tagged human Gli3pRM in PK15, Cos-1, and TM4 cells in order to study intracellular co-localization of the intrabody and its target antigen (Fig. 4). The anti-Gli3 scFv showed discrete perinuclear cytoplasmic localization in all three cell lines (Fig. 4A, D, and G), very similar to that seen with recombinant Gli3pRM (Fig. 4B, E, and H). Analogous experiments with GST-tagged scFv5E1 yielded the same staining pattern (data not shown), indicating that the nature of the fusion tag and the corresponding detection antibody is unlikely to have significant influence on intrabody distribution in these cell lines. These results suggest that the intrabody may indeed target cytoplasmic Gli3. Furthermore, the co-transfection of Gli3pRM was able to modify the localization of scFv5E1, which was not any more diffuse in the cytoplasm as when the anti-Gli3 scFv was expressed alone (see Fig. 2). Interestingly, while a weak Gli3pRM signal was also detected in the nucleus, the scFv showed no nuclear localization, suggesting that the putative intrabody/Gli3 complex was not able to cross the nuclear membrane.

scFv5E1 is specific to Gli3

Subsequent experiments with human Gli2 further demonstrated that the scFv does not cross-react with another Gli family member. Co-expression of His-tagged anti-Gli3 scFv with a GFP-Gli2 fusion protein in PK15 cells (Fig. 5A and B) showed that while GFP-Gli2 was abundant both in the cytoplasm and in the nucleus, the anti-Gli3 scFv was only dispersed in the cytoplasm in a pattern similar to that observed when the scFv expression vector was transfected alone (see Fig. 2). However, when co-expressed with both GFP-Gli2 and Gal4-Gli3 (Fig. 5C and D), the anti-Gli3 scFv assumed a perinuclear distribution that was clearly distinct from that seen in transfected cells expressing only the scFv and GFP-Gli2.

scFv5E1 also recognizes endogenous murine Gli3

The human Gli3 sequence is highly similar to the mouse homolog (86.5% identity at the amino acid level), particularly within the protein's N-terminal repressive motif (Gli3pRM) recognized by MAb 5E1 (97.7% identity over 132 residues). Earlier studies have shown that the native antibody is indeed able to detect the expression of both human and murine Gli3.[15]

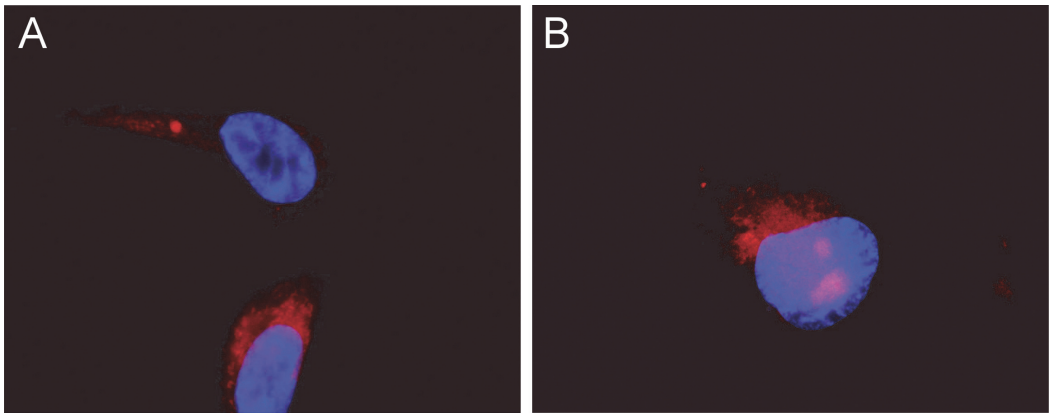

FIG. 6. Intracellular localization of His-tagged scFv5E1 in transfected mouse TM4 cells (**A**) is similar to the endogenous Gli3 pattern revealed by staining with the original MAb 5E1 (**B**), suggesting that the intrabody recognizes endogenous murine Gli3. Blue staining represents nuclei.

To test the capacity of scFv5E1 to recognize the endogenous mouse protein, an expression vector encoding His-tagged anti-Gli3 scFv was transfected into the TM4 cell line derived from mouse Sertoli cells that naturally express Gli3.[15] In contrast to what was observed in Gli3-negative cell lines such as PK15 and Cos-1 (see Fig. 3), scFv5E1 did not show a diffuse cytoplasmic localization but was mainly localized in the perinuclear region (Fig. 6A). This was reminiscent of the staining observed in PK15 or Cos-1 cells expressing exogenous human Gli3 (see Fig. 4). The same pattern was also revealed by staining TM4 cells with the parent MAb 5E1 (Fig. 6B), further suggesting that it indicated the localization of endogenous murine Gli3 in the cytoplasm. In addition, it was confirmed that the intrabody specifically targets Gli3, as no signal was observed when staining was performed on other mouse cell lines not expressing Gli3 (data not shown). In some TM4 cells, a dot stained by scFv5E1 was observed in a location distant from the nucleus, which may indicate the presence of Gli3 in microtubule-based organelles called cilia, a suggestion that is in accordance with recent reports revealing a physical and functional interaction between Gli3 and ciliary proteins.[17,18]

Taken together, these results suggest that the scFv generated against human Gli3 cross-reacts with the mouse protein, and that the affinity of the interaction may allow detection of endogenously expressed Gli3.

Discussion

Over the past decades, there has been growing interest in the use of scFv antibodies aimed at intracellular targets for the purpose of inhibiting or modulating protein expression and trafficking. Intrabodies represent a potent alternative to gene inactivation methods such as antisense oligonucleotides, zinc finger proteins, and RNA interference, which target gene expression at the DNA or mRNA level. Because intrabodies can be directed to specific subcellular compartments and can target precisely defined epitopes on proteins, they provide an interesting possibility to block only one of several functions of an expressed protein at the posttranslational level.[19,20]

The full-length Gli2 and Gli3 proteins are sequestered in the cytoplasm by binding to Fu and SuFu, forming a complex where they can be phosphorylated and processed by proteases (see review[21]). The resulting N-terminal fragments of Gli2 and Gli3 can then translocate to the nucleus and repress transcription. Alternatively, binding of Shh to its receptor Patched initiates a signaling cascade that dissociates the Gli/Fu/SuFu complex in the cytoplasm, allowing full-length Gli proteins to reach the nucleus where they can activate their target genes.

We have generated a single-chain Fv antibody based on the variable regions of monoclonal antibody 5E1[15] directed against the repression domain of human Gli3 (Gli3pRM). To our knowledge, this work represents the first characterization of an intrabody against a human Gli protein. As our results show, the 5E1-derived intrabody is suitable for detection of both human and mouse Gli3 in the cytoplasm. Interestingly, the intrabody staining seems to be restricted to the cytoplasm, suggesting that the reagent may not have access to the nucleus. The generated scFv5E1 constitutes an interesting tool for the study of intracellular Gli3 localization and trafficking. Several studies have shown that in spite of its small size, binding of an intrabody to its target may inhibit the activity of the protein, as demonstrated by pioneering treatments of neurological disorders.[22,23] Further studies are needed to analyze the modalities and consequences of scFv5E1 binding to Gli3, which will provide new insights into Gli3 biology.

Acknowledgments

This work was supported by grants from Enterprise Estonia (EU14872) and the Estonian Science Foundation (ETF5931) to L.J. PK is a Wellcome Trust International Senior Research Fellow. We thank Nora Ausmees and Alla Maloverjan for their technical assistance and Pierre Boudinot for helpful discussions.

References

1. Wang C, Pan Y, and Wang B: A hypermorphic mouse Gli3 allel results in a polydactylous limb phenotype. Dev Dyn 2007;236:769–776.
2. Biesecker LG: What you can learn from one gene: Gli3. J Med Genet 2006;43:465–469.
3. Hui CC, Slusarski D, Platt KA, Holmgren R, and Joyner AL: Expression of three mouse homologs of the Drosophila segment polarity gene cubitus interruptus, Gli, Gli-2, and Gli-3, in ectoderm- and mesoderm-derived tissues suggests multiple roles during postimplantation development. Dev Biol 1994;162:402–413.
4. Pan Y and Wang B: A novel protein-processing domain in Gli2 and Gli3 differentially blocks complete protein degradation by the proteasome. J Biol Chem 2007;282:10846–10852.
5. Pu Y, Huang L, and Prins GS: Sonic hedgehog-patched Gli signaling in the developing rat prostate gland: lobe-specific suppression by neonatal estrogens reduces ductal growth and branching. Dev Biol 2004;273:257–275.
6. Ruiz i Altaba A, Sanchez P, and Dahmane N: Gli and hedgehog in cancer: tumours, embryos and stem cells. Nat Rev Cancer 2002;2:361–372.
7. Kalff-Suske M, Wild A, Topp J, Wessling M, Jacobsen EM, Bornholdt D, Engel H, Heuer H, Aalfs CM, Ausems MG, Barone R, Herzog A, Heutink P, Homfray T, Gillessen-Kaesbach G, Konig R, Kunze J, Meinecke P, Muller D, Rizzo R, Strenge S, Superti-Furga A, and Grzeschik KH: Point mutations throughout the Gli3 gene cause Greig cephalopolysyndactyly syndrome. Hum Mol Genet 1999;8:1769–1777.
8. Vortkamp A, Gessler M, and Grzeschik KH: Gli3 zinc-finger gene interrupted by translocations in Greig syndrome families. Nature 1991;352:539–540.
9. Wild A, Kalff-Suske M, Vortkamp A, Bornholdt D, König R, and Grzeschik KH: Point mutations in human GLI3 cause Greig syndrome. Hum Mol Genet 1997;6:1979–1984.
10. Kang S, Graham JM Jr, Olney AH, and Biesecker LG: Gli3 frameshift mutations cause autosomal dominant Pallister-Hall syndrome. Nat Genet 1997;15:266–268.
11. Radhakrishna U, Bornholdt D, Scott HS, Patel UC, Rossier C, Engel H, Bottani A, Chandal D, Blouin JL, Solanki JV, Grzeschik KH, and Antonarakis SE: The phenotypic spectrum of Gli3 morphopathies includes autosomal dominant preaxial polydactyly type-IV and postaxial polydactyly type-A/B. No phenotype prediction from the position of Gli3 mutations. Am J Hum Genet 1999;65:645–655.
12. Shin SH, Kogerman P, Lindstrom E, Toftgard R, and Biesecker LG: Gli3 mutations in human disorders mimic

Drosophila cubitus interruptus protein functions and localization. Proc Natl Acad Sci USA 1999;96:2880–2884.
13. Nilsson M, Unden AB, Krause D, Malmqwist U, Raza K, Zaphiropoulos PG, and Toftgard R: Induction of basal cell carcinomas and trichoepitheliomas in mice overexpressing Gli-1. Proc Natl Acad Sci USA 2000;97:3438–3443.
14. Karhadkar SS, Bova GS, Abdallah N, Dhara S, Gardner D, Maitra A, Isaacs JT, Berman DM, and Beachy PA: Hedgehog signaling in prostate regeneration, neoplasia and metastasis. Nature 2004;431:707–712.
15. Hunt R, Bragina O, Drews M, Kasak L, Timmusk S, Valkna A, Kogerman P, and Jarvekulg L: Generation and characterization of mouse monoclonal antibody 5E1 against human transcription factor Gli3. Hybridoma 2007;26:131–139.
16. Sasaki H, Nishizaki Y, Hui C, Nakafuku M, and Kondoh H: Regulation of Gli2 and Gli3 activities by an amino-terminal repression domain: implication of Gli2 and Gli3 as primary mediators of Shh signaling. Development 1999;126:3915–3924.
17. Eley L, Yates LM, and Goodship JA: Cilia and disease. Curr Opin Genet Dev 2005;15:308–314.
18. Haycraft CJ, Banizs B, Aydin-Son Y, Zhang Q, Michaud EJ, and Yoder BK: Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. PLoS Genet 2005;1:e53.
19. Li L, Gondi CS, Dinh DH, Olivero WC, Gujrati M, and Rao JS: Transfection with anti-p65 intrabody suppresses invasion and angiogenesis in glioma cells by blocking nuclear factor-kappaB transcriptional activity. Clin Cancer Res 2007;13:2178–2190.
20. Miller TW, and Messer A: Intrabody applications in neurological disorders: progress and future prospects. Mol Ther 2005;12:394–401.
21. Osterlund T, and Kogerman P: Hedgehog signaling: how to get from Smo to Ci and Gli. Trends Cell Biol 2006;16:176–180.
22. Doorbar J, and Griffin H: Intrabody strategies for the treatment of human papillomavirus-associated disease. Expert Opin Biol Ther 2007;7:677–689.
23. Miller TW, Zhou C, Gines S, MacDonald ME, Mazarakis ND, Bates GP, Huston JS, and Messer A: A human single-chain Fv intrabody preferentially targets amino-terminal Huntingtin's fragments in striatal models of Huntington's disease. Neurobiol Dis 2005;19:47–56.

Address reprint requests to:
Harri Altroff
Celecure AS
Akadeemia tee 15
Tallinn 12618
Estonia

E-mail: harri@celecure.com

Received: November 16, 2007
Accepted after revision: February 12, 2008

PUBLICATION II

FEBS Letters 583 (2009) 224–228

FEBS Letters

journal homepage: www.FEBSLetters.org

Identification of the gene transcription repressor domain of Gli3

Robert Tsanev, Piret Tiigimägi, Piret Michelson, Madis Metsis, Torben Østerlund *, Priit Kogerman

Department of Gene Technology, Tallinn University of Technology, Akadeemia tee 15, 12618 Tallinn, Estonia

ARTICLE INFO

Article history:
Received 5 November 2008
Revised 27 November 2008
Accepted 4 December 2008
Available online 10 December 2008

Edited by Lukas Huber

Keywords:
Gli3
Transcription
Domain
Hedgehog
Signaling
Development

ABSTRACT

Gli transcription factors are downstream targets of the Hedgehog signaling pathway. Two of the three Gli proteins harbor gene transcription repressor function in the N-terminal half. We have analyzed the sequences and identified a potential repressor domain in Gli2 and Gli3 and have tested this experimentally. Overexpression studies confirm that the N-terminal parts harbor gene repression activity and we mapped the minimal repressor to residues 106 till 236 in Gli3. Unlike other mechanisms that inhibit Gli induced gene transcription, the repressor domain identified here does not utilize Histone deacetylases (HDACs) to achieve repression, as confirmed by HDAC inhibition studies and pull-down assays. This distinguishes the identified domain from other regulatory parts with negative influence on transcription.

1. Introduction

The morphogens of the Hedgehog (Hh) family are crucial for numerous developmental processes during embryogenesis as well as postnatally. In fact the Hh pathway is one of the four major signaling systems that are controlling the major developmental processes. Gli proteins are Zn-finger transcription factors and are targets as well as mediators of the Hh signaling pathway [1]. Mammals have three *Gli* genes encoding Gli1, Gli2 and Gli3 that are orthologs of the *Drosophila* transcription factor *Cubitus interruptus* (Ci). Ci is mainly a transcriptional activator in its full-length form that dominates in the presence of Hh. In the absence of Hh, a portion of Ci is proteolytically cleaved to produce an N-terminal gene repressor form. In a similar fashion Gli2 and Gli3 can undergo proteolysis to produce a gene repressor form. The full-length forms of Gli2 and Gli3 act as gene activators. A repressor form of Gli1 cannot be generated and Gli1 is considered to be a strong gene activator. The dominating role of Gli2 appears to be gene activation whereas Gli3 often has a gene repression role, mediated by the N-terminal part. In humans, several *GLI3* morphopathies have been described, which can be broadly divided into two classes: Greig's syndrome caused by total loss of GLI3 function and Pallister–Hall syndrome (PHS)/other postaxial polydactylies that are presumed to be caused by abnormally high repressor generation. The first identified mutations causing PHS were found in the *GLI3* gene [2]. Since then several *Gli3* mutations have been identified in the same region (exons 12–14). Both original mutations are single nucleotide deletions that lead to frame shift and premature translational stop [2]. The produced peptide has 691 residues (compared to the 1596 residue full-length protein) but contain alternative residues in the last approximately 20 residues, encoded after the mutations [2]. It was shown that the corresponding peptide Gli3-PHS (residues 1–674) indeed has strong gene repressor activity, which may explain the phenotypes of these patients [3]. Due to its vast impact on cell differentiation and proliferation aberrant Hh signaling is involved in many cancers and several gene members of the pathway are either proto-oncogenes or tumor suppressors [4]. A thorough analysis of the Gli proteins is therefore important in order to understand the associated developmental biology and pathology as well as related carcinogenesis.

To further analyze the repressor function in the PHS part of Gli3 and to identify the specific repressor sequence, we made a series of GLI3 constructs and evaluated their activity in cellular gene regulation assays. This led to the identification of a specific repressor domain in GLI3 also conserved in GLI2 but not in GLI1. The repressor function of this domain is not dependent on histone deacetylases (HDAC) and therefore works through a different mechanism.

Abbreviations: β-gal, β-galactosidase; Ci, *Cubitus interruptus*; DMEM, Dulbecco's modified Eagle's medium; FCS, fetal calf serum; HDAC, histone deacetylases; Hh, Hedgehog; PHS, Pallister–Hall syndrome; TSA, trichostatin-A.

* Corresponding author.
E-mail address: toosathome@yahoo.se (T. Østerlund).

0014-5793/$34.00 © 2008 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2008.12.010

2. Materials and methods

2.1. DNA constructs

Gli1, Gli1(1–407), Gli3, Gli3-PHS, Gli3ΔRD and Gli3-PHSΔRD all of human origin were cloned into pcDNA3.1His expression vector (some of these were described before [3,5]). The 12GliRE-luc and β-galactosidase (β-gal) constructs were described before [5]. The Gli3 repressor domain (residues 106–246) and shorter versions were subcloned into the pFA vector in frame with the DNA Binding Domain (DBD) of yeast Gal4 (Stratagene, La Jolla, CA, USA). As Gal4 reporter construct was used the pMN-Luc plasmid containing a thymidine kinase promoter with five tandem repeats of the yeast GAL4 binding sites that control expression of the firefly luciferase gene. For recruitment of HDAC in gene silencing we employed the C-terminal HDAC dependent repressor domain of the rat REST protein [6] cloned in frame with GAL4 DBD in pFA.

2.2. Cell culture

HEK293 cells were grown and transfected as previously described [7]. The cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS) (PAA Laboratories, Pasching, Austria), streptomycin and penicillin (100 units/ml; Invitrogen, Carlsbad, CA, USA). Cells were grown at 37 °C and 5.0% CO_2 in cell culture incubator. One day before transfection cells were plated into the required growth plates.

Shh-Light2 cells were grown in the same medium as HEK293 supplemented with G 418 (400 μg/ml; Sigma–Aldrich, St. Louis, MO, USA) and Zeocine (100 mg/ml; Invitrogen). At 24 h post transfection the medium was changed to low-serum medium (0.5% of FCS; PAA Laboratories).

2.3. Luciferase assays

Transfections for luciferase assays were performed in 24-well plates. Assessment of Gli1, Gli3 and Gli3-PHS in HEK293 cells was done as previously described [5,7]. Assessment in Shh-L2 cells was performed as described [8], using the incorporated luciferase gene as measurement of gene activation and the co-transfected β-gal as control. Transfections were done with the same amount of total DNA by using empty vector to compensate.

For measurement of the Gli3-RD deletion constructs we transfected HEK293 cells also using the β-gal construct as control. The amount of reporter plasmid (pMN-Luc) used was 300 ng per well and the effector plasmids (pFA Gal4 fusions with RD segments) were 30 ng per well. For normalization we used 100 ng of pCMV-β-gal. As a transfecting agent we used polyethyleneimine (PEI; Sigma–Aldrich) 1 μg per well. DNA and PEI were mixed in 50 μl of DMEM. An additional 150 μl of DMEM was added to the DNA/PEI mixture and then applied to the cells. After 2 h the medium was exchanged for DMEM with 10% FCS. On the following day the medium was changed again and where required, trichostatin-A (TSA) was added at 0.2, 0.5, and 1 μM. Cells were harvested after an additional 24 h. Firefly luciferase and β-gal assays were performed in Ascent FL fluoroskan with the Luciferase Assay Kit (BioTherma, Darlarö, Sweden) and Galacto-Light Plus System (Applied Biosystems, Foster City, CA, USA). In assays measuring HDAC induced gene silencing we used the REST expressing pFA vector as positive control of HDAC recruitment.

2.4. Immunoprecipitation and HDAC enzymatic assay

To test the association of Gli3-RD with HDAC we immunoprecipitated Gli3-RD and measured HDAC enzymatic activity of the precipitate. As a positive control we immunoprecipitated Sin3A that is known to be in a complex with HDAC1 and HDAC2.

We transfected HEK293 cells with Gli3-RD in 15 cm culture dishes. The DNA/PEI complex was prepared as follows: 30 μg of Gli3-RD DNA was mixed with 60 μg of PEI per plate in 500 μl of DMEM. After 10 min of incubation 9 ml of DMEM was added to the DNA/PEI mixture and then applied to the culture dish. After 2 h the medium was exchanged for DMEM with 10% FCS. On the following day, the medium was changed again and cells were lysed in PBS with 1% Triton X-100 (Sigma–Aldrich) after an additional 24 h.

For immunoprecipitation of Gli3-RD 5 μl of anti-Gal4 polyclonal antibodies (Santa Cruz Biotechnology, Santa Cruz, CA, USA) were incubated with 30 μl Protein G agarose (Amersham Biosciences, Bucks, UK). After 1 h 1 ml of cell lysate was added and immunoprecipitation was performed at 4 °C overnight. The immunoprecipitation of Gli3-RD was analyzed by Western blot using Gal4 monoclonal antibody (Santa Cruz Biotechnology). Sin3A was immunoprecipitated and detected as described above using an anti mSin3A antibody (Santa Cruz Biotechnology).

HDAC activity was measured using a fluorescent substrate Fluor de Lys (Biomol, Plymouth Meeting, PA, USA) that contains an acetylated lysine side chain. This substrate was incubated with immunoprecipitated Gli3-RD or mSin3A. If the immunoprecipitate contain HDAC's the substrate is deacetylated and a fluorophore is produced. Assays were performed according to manufacturer's instructions, and measured in Tecan GENios pro microplate spectrofluorometer (Tecan Group, Männedorf, Switzerland) with the Magellan V5.03 system (Tecan Group).

3. Results and discussion

The Gli3-PHS part (residues 1–673) was shown to contain gene repressor activity [3]. Using a two-hybrid screening technology it was shown that almost the same part (residues 1–613) in Gli3 binds to Ski [9]. The Ski binding site on Gli3 was determined to the region from residue 152 to 397 using pull-down assays [9]. Since Ski is known to be part of a gene repressor complex including HDAC, it was suggested that Gli3 exerts its repressor activity through binding of Ski and recruitment of HDAC [9]. However, most of the Ski binding region is conserved between all three Gli proteins, suggesting that Ski binding and HDAC recruitment is part of a general transcription termination signal common to all Gli proteins. Likewise, the SUFU binding site (BS) on Gli proteins (SYGH) is also found in all three Gli proteins [10]. Also SUFU is known to recruit HDAC through recruitment of SAP18 and Sin3A [11] and therefore, SUFU binding may also be regarded as a general mechanism to turn off Gli mediated transcription. Recently, two sites in Gli1 were identified as responsible for protein degradation [12]. One peptide (degron) was in the C-terminal part (D_C) whereas the other was found in the N-terminal part (D_N). In fact the D_N peptide is located very close to the SUFU binding site and is conserved also in Gli2 and Gli3. A previous study identified the peptide 94–280 of Gli2 as a repressor part and removal of a corresponding part in Gli3 (residues 1–344) had strong positive effect on transcription [13]. In contrast, when this part of Gli1 (residues 1–134) was removed there was no effect on transcription as compared to wild type protein [13]. The last approximately 100 residues of this region is conserved between the three Gli proteins, and those are the parts that overlap with the identified Ski binding part, contain D_N and the SYGH peptide. Fig. 1A shows a schematic alignment of the N-terminal parts of mammalian Gli proteins, until the end of the Zn-fingers (corresponding to the PHS-domain), with indications of the respective domains describe above. From this work and a previous paper [3] it is suggested that the gene

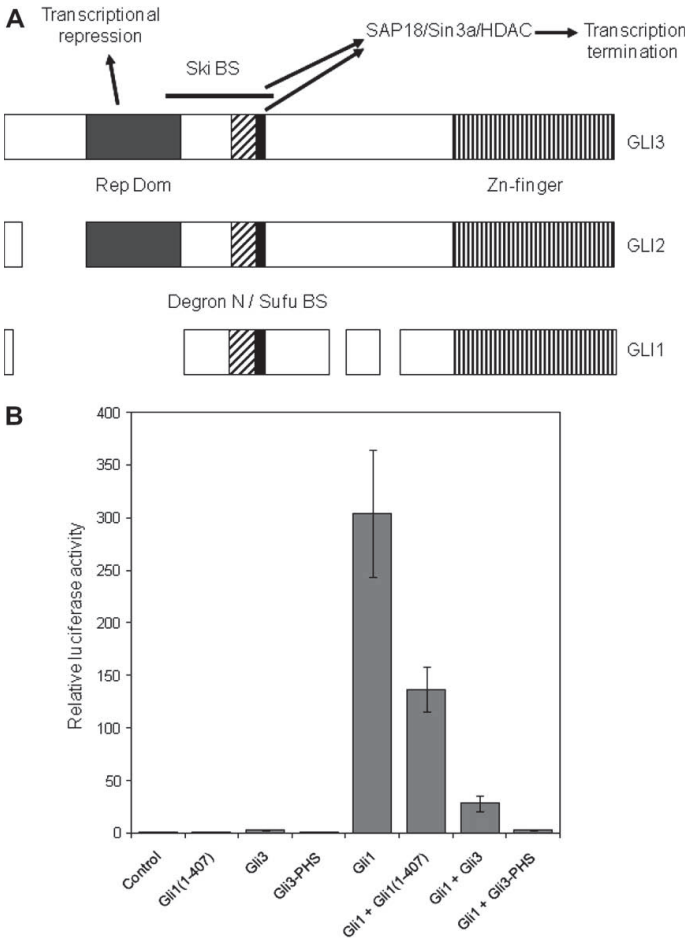

Fig. 1. Sequence analyzes of the N-terminal halves of Gli1, Gli2 and Gli3, as well as analyzes of the transcriptional regulation by Gli1 and Gli3 and the PHS domain of Gli3. (A) Schematic alignment of Gli1, Gli2 and Gli3 in the N-terminal half until the end of the Zn-finger DBD (Zn-finger; lined). The suggested repressor domain (Rep Dom; grey) is only found in Gli2 and Gli3. The sequences for the Sufu binding site (Sufu BS) and $Degron_N$ (Degron N) is found in all three Gli proteins (hatched and black, respectively). Above is a line indicating the presumed Ski binding region (Ski BS). Both Ski and Sufu are likely to recruit HDACs through Sin3A and induce transcriptional termination by this mechanism. (B) HEK293 cells were transfected with a Gli responsive luciferase reporter and Gli1, Gli3 or N-terminal parts of these corresponding to the PHS domain, or combinations to assess the effects of these on transcription. Error bars indicate the standard deviations of triplicate analyzes.

repressor function is localized to the N-terminal (grey) part of Gli2 and Gli3. The physiological significance was shown by regulation of *PTCH1* transcription [3]. We suggest that the other mechanisms with negative gene transcription activity (Ski BS, SUFU BS and D_N peptide) that are common to all Gli proteins, are general means to terminate Gli induced transcription. At least two of these pathways (Ski and SUFU through interaction with SAP18 and Sin3A) probably recruit HDACs to terminate transcription and increase the degradation of the Gli protein.

As stated above it has previously been shown that the Gli3-PHS domain is a repressor of both basal and Gli3 induced transcription [3]. However, since Gli3 only induces a modest transcriptional activation we wanted to test the PHS domain together with the much stronger transcriptional activator Gli1. Therefore, in HEK293 cells Gli1, Gli3 and Gli3-PHS were transfected alone or in combinations together with a Gli-luciferase reporter construct [5]. As shown before [5,7] Gli1 is an effective activator of transcription whereas Gli3 only activates weakly (Fig. 1B). The transcriptional activation of Gli1 is strongly inhibited by co-expressing Gli3-PHS. Expression of full-length Gli3 also leads to repression of Gli1 induced transcription, although significant activity is seen. Gli3 on its own gives much lower activity but it appears that repression of Gli1 is a more pronounced effect (3–4 times Gli3 induction vs. 6 times repression). In other words, not only does the Gli3-PHS have repressor activity on its own [3] it also strongly repress Gli1 induced transcription. Expression of the Gli1 peptide corresponding to Gli3-PHS (residues 1–407) only weakly suppresses Gli1 induced transcription, which is likely to be due to competitive expression and suggest that the N-terminal of Gli1 does not exert any significant repressor function.

In order to analyze the constructs in a more in vivo-like setting, we turned to the Shh-Light2 (Shh-L2) cells that have a Gli-inducible luciferase reporter construct incorporated into the genome. [14]. Transfection of these cells is less efficient and the transcrip-

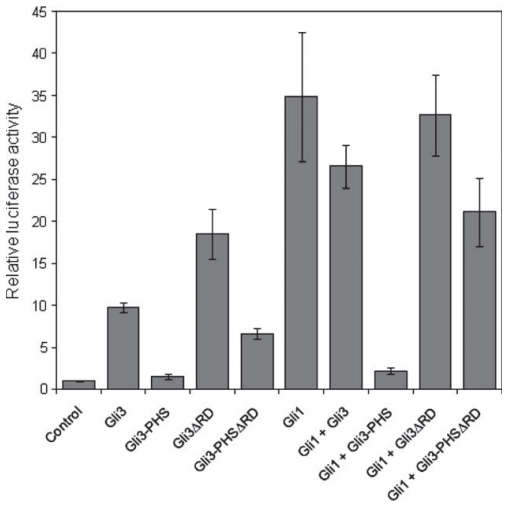

Fig. 2. Analyzes of the repressor function of the Gli3-PHS domain and the repressor domain (RD) in Shh-L2 cells. Shh-L2 cells (with an incorporated Gli responsive luciferase reporter gene) were transfected with Gli1, Gli3, Gli3-PHS, Gli3ΔRD, Gli3-PHSΔRD or combinations of these to assess the effect on transcription alone or on the Gli1 induced transcription. The analyzes were performed at least three times and error bars indicate the standard deviations.

tional induction by Gli1 is much lower than in the HEK293 cells (Fig. 2). The induction posed by Gli3 is also lower but the difference is not as pronounced as in HEK293 cells. This may indicate that in the Shh-L2 cells the transcriptional regulation of the reporter is different from the vector-based one used in HEK293 cells. Alternatively, the differences could reflect differences in the ratios of the expressed Gli peptides. In the Shh-L2 cells Gli3 does not affect Gli1 induced transcription as much as in the HEK293 cells. However, the pattern for both Gli3 and Gli3-PHS induced Gli1 repression is the same as in HEK293 cells, showing that the PHS domain is a strong repressor.

Based on the alignment we made deletion constructs of Gli3 and Gli3-PHS that lack residues 105–246 (Gli3ΔRD and Gli3-PHSΔRD), corresponding to the grey area in Gli2 and Gli3 (Fig. 1A). Removal of this part enhances the gene transcription induction of Gli3 (Fig. 2). This is confirmed in the Gli1/Gli3ΔRD combination, though the effect is small. The effect is much more pronounced when comparing the repression of Gli3-PHS on Gli1 induced transcription to that of Gli3-PHSΔRD (Fig. 2). In the latter case the repression is reduced almost 10 times. This clearly indicates that the particular sequence harbors significant gene transcription repression activity. In fact most (if not all) the repressor function is located in this part of the PHS domain.

We then proceeded to make constructs containing this repressor domain (RD) and parts of this in frame with Gal4 DBD and used the constructs to assess the minimal RD of Gli3. We used HEK293 cells that were also transfected with a Gal4 DBD responsive luciferase reporter. The DBD of the Gli3-RD fusion peptides bind to the reporter plasmid at the GAL4 binding sites. Compared to Gal4 alone (pFA vector) the N-terminus of Gli3 and the Gli3-RD significantly represses gene expression (Fig. 3). However, shorter versions of the Gli3-RD lose the ability to repress transcription and therefore we mapped the minimal RD of GLI3 between residues 106 and 235. When the RD is shortened beyond these residues the repression is reduced or lost.

The suggested Ski binding site on Gli3 has not been exactly mapped (it is in the region from residue 152 to 397 [9]) and may overlap with the identified repressor part. Since Ski and SUFU recruits HDACs to exert their inhibitory role, we wanted to test if the repressor function described here depends on the same mechanism or not. We transfected HEK293 cells with either Gli3-RD or the repressor domain of REST (that serve as positive control since it depend on HDAC to repress transcription) and tested the effect of the HDAC inhibitor TSA as shown in Fig. 4. Again the Gli3-RD suppresses transcription but there is not any effect of TSA up to 1 μM.

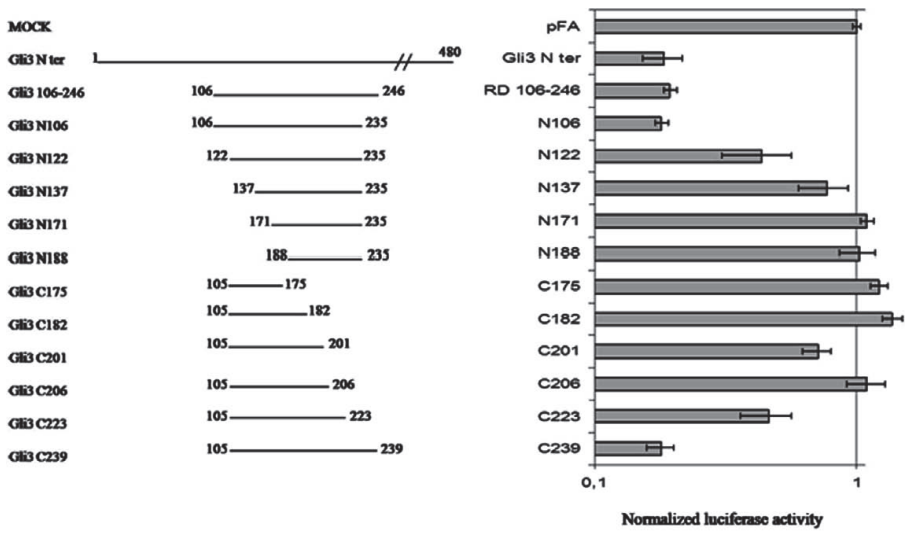

Fig. 3. Determination of the minimal repressor domain of Gli3. The repressor domain (residues 105–246) or parts of this were expressed together with the DBD of Gal4 and assessed for repression of Gal4 induced transcription in HEK293 cells (mock). Also a larger part of the Gli3 N-terminal part was measured since this is known to have significant repressor function (residues 1–480). The analyzes were performed three to five times and error bars show the standard deviations.

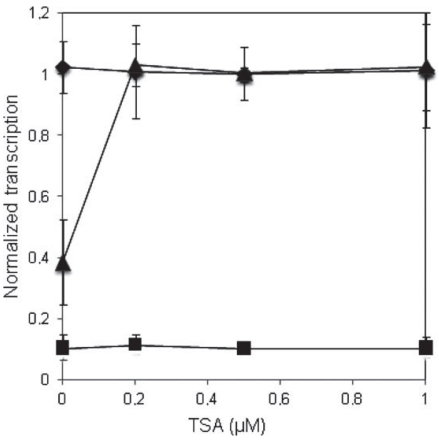

Fig. 4. HDAC recruitment study of the Gli3 repressor domain. HEK293 cells were transfected with Gli3-RD (squares) or the repressor domain of REST (positive control, triangles) and treated with increasing amounts of the HDAC inhibitor TSA. As negative control we used cells transfected with empty vector (diamonds). The analyzes were performed three to six times and the error bars indicate standard deviations.

In contrast the repression by REST is relieved at only 200 nM TSA. This shows that Gli3-RD induced repression is not dependent on HDACs in the way REST is.

As a further control we used pull-down assays to assess if there is any binding of Gli3-RD to HDACs. In cells overexpressing Gli3-RD fused to Gal4 DBD we used anti-Gal4 antibodies to precipitate the fusion protein. HDAC activity was then assessed in the precipitates. Neither Gal4 DBD nor the fusion protein showed any significant HDAC binding (not shown). As a positive control we also immunoprecipitated Sin3A from the HEK293 cells using an mSin3A antibody. In this precipitate there was significant HDAC activity. Sin3A is a transcriptional regulator known to recruit HDACs to induce gene silencing [15,16]. Both Ski and SUFU recruits HDACs through Sin3A. Thus, the repression induced by the Gli3-RD is through a different mechanism than that of Ski and SUFU.

When considering that HDACs are involved in more permanently shutting down a target gene, this also makes good sense. This mechanism is probably more useful when drastic measures are required, e.g. when Hh signaling ceases, or when other signals determines that Gli induced transcription has to end. Therefore, these signals (Ski/SUFU/HDAC) [9,11] are probably common to all the Gli proteins. Also the degradation signal (D_N) is of this general category and leads to removal of the targeted Gli protein and we suggest that the Ski and SUFU signals recruit HDACs to terminate Gli induced transcription at the Gli binding site on DNA (Fig. 1). In contrast our results clearly suggest that the gene transcription repression exerted by the N-termini of Gli2 and Gli3 is mediated by the repressor domain identified here and indicated with grey in Fig. 1A. We also suggest that the Ski binding site is strictly localized in the region common to all three Gli proteins, but a more exact mapping of the Ski binding site requires further experimentation. It remains to be investigated by which mechanism Gli3-RD regulates transcription. Perhaps the domain recruits other proteins than HDACs or interacts (physically or functionally) with the transcriptional machinery. It has been shown that Gli3 interacts with and regulate gene transcription via mediator [17] and perhaps the repressor domain it able to influence this interaction. However, the mediator binding site is localized at the C-terminal part of Gli3 [17] and the Gli3-PHS like repressor that is generated in vivo is not likely to bind mediator, but may exert its repressor function by an independent mechanism. It is clear that the repression is not dependent on HDACs and investigations of the mechanism of repression have been initiated.

Acknowledgements

We thank Lagle Kasak and Kersti Olspert for technical assistance and Dr. Kalju Vanatalu, for fruitful discussions during the study. This study was supported by the Wellcome Trust and Estonian Science Foundation (Grants 5552, 5933 to P.K. and 7279 to T.Ø.).

References

[1] Hooper, J.E. and Scott, M.P. (2005) Communicating with Hedgehogs. Nat. Rev. Mol. Cell Biol. 6, 306–3317.
[2] Kang, S., Graham Jr., J.M., Olney, A.H. and Biesecker, L.G. (1997) GLI3 frameshift mutations cause autosomal dominant Pallister–Hall syndrome. Nat. Genet. 15, 266–268.
[3] Shin, S.H., Kogerman, P., Lindström, E., Toftgård, R. and Biesecker, L.G. (1999) *GLI3* mutations in human disorders mimic *Drosophila Cubitus interruptus* protein functions and localization. Proc. Natl. Acad. Sci. USA 96, 2880–2884.
[4] Rubin, L.L. and de Sauvage, F.J. (2006) Targeting the Hedgehog pathway in cancer. Nat. Rev. Drug Discov. 5, 1026–1033.
[5] Kogerman, P., Grimm, T., Kogerman, L., Krause, D., Undén, A.B., Sandstedt, B., Toftgård, R. and Zaphiropoulos, P.G. (1999) Mammalian suppressor-of-fused modulates nuclear-cytoplasmic shuttling of Gli-1. Nat. Cell Biol. 1, 312–319.
[6] Huang, Y., Myers, S.J. and Dingledine, R. (1999) Transcriptional repression by REST: recruitment of Sin3A and histone deacetylase to neuronal genes. Nat. Neurosci. 2, 867–872.
[7] Østerlund, T., Everman, D.B., Betz, R.C., Mosca, M., Nöthen, M.M., Schwartz, C.E., Zaphiropoulos, P.G. and Toftgård, R. (2004) The *FU* gene and its possible protein isoforms. BMC Genomics 5, 49.
[8] Maloveryan, A., Finta, C., Østerlund, T. and Kogerman, P. (2007) A possible role of mouse Fused (STK36) in Hedgehog signaling and Gli transcription factor regulation. J. Cell Commun. Signal. 1, 165–173.
[9] Dai, P., Shinagawa, T., Nomura, T., Harada, J., Kaul, S.C., Wadhwa, R., Khan, M.M., Akimaru, H., Sasaki, H., Colmenares, C. and Ishii, S. (2002) Ski is involved in transcriptional regulation by the repressor and full-length forms of Gli3. Genes Develop. 16, 2843–2848.
[10] Dunaeva, M., Michelson, P., Kogerman, P. and Toftgard, R. (2003) Characterization of the physical interaction of Gli proteins with SUFU proteins. J. Biol. Chem. 278, 5116–5122.
[11] Cheng, S.Y. and Bishop, J.M. (2002) Suppressor of Fused represses Gli-mediated transcription by recruiting the SAP18-mSin3 corepressor complex. Proc. Natl. Acad. Sci. USA 99, 5442–5447.
[12] Huntzicker, E.G., Estay, I.S., Zhen, H., Lokteva, L.A., Jackson, P.K. and Oro, A.E. (2006) Dual degradation signals control Gli protein stability and tumor formation. Genes Develop. 20, 276–281.
[13] Sasaki, H., Nishizaki, Y., Hui, C., Nakafuku, M. and Kondoh, H. (1999) Regulation of Gli2 and Gli3 activities by an amino-terminal repression domain: implication of Gli2 and Gli3 as primary mediators of Shh signaling. Development 126, 3915–3924.
[14] Taipale, J., Chen, J.K., Cooper, M.K., Wang, B., Mann, R.K., Milenkovic, L., Scott, M.P. and Beachy, P.A. (2000) Effects of oncogenic mutations in smoothened and patched can be reversed by cyclopamine. Nature 406, 1005–1009.
[15] Laherty, C.D., Yang, W.-M., Sun, J.-M., Davie, J.R., Seto, E. and Eisenman, R.N. (1997) Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. Cell 89, 349–356.
[16] Nagy, L., Kao, H.Y., Chakravarti, D., Lin, R.J., Hassig, C.A., Ayer, D.E., Schreiber, S.L. and Evans, R.M. (1997) Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. Cell 89, 373–380.
[17] Zhou, H., Kim, S., Ishii, S. and Boyer, T.G. (2006) Mediator modulates Gli3-dependent Sonic Hedgehog signaling. Mol. Cell Biol. 26, 8667–8682.

PUBLICATION III

OPEN ACCESS Freely available online

PLOS ONE

The Transcriptional Repressor Domain of Gli3 Is Intrinsically Disordered

Robert Tsanev[1]*, Kalju Vanatalu[1], Jüri Jarvet[2], Risto Tanner[2], Kristi Laur[1], Piret Tiigimägi[1], Birthe B. Kragelund[3], Torben Østerlund[4], Priit Kogerman[1]

1 Department of Gene Technology, Tallinn University of Technology, Tallinn, Estonia, **2** National Institute of Chemical Physics and Biophysics, Tallinn, Estonia, **3** Structural Biology and Nuclear Magnetic Resonance Laboratory, University of Copenhagen, Copenhagen, Denmark, **4** Molecular Pharmacology, Zealand Pharma A/S, Glostrup, Denmark

Abstract

The transcription factor Gli3 is acting mainly as a transcriptional repressor in the Sonic hedgehog signal transduction pathway. Gli3 contains a repressor domain in its N-terminus from residue G106 to E236. In this study we have characterized the intracellular structure of the Gli3 repressor domain using a combined bioinformatics and experimental approach. According to our findings the Gli3 repressor domain while being intrinsically disordered contains predicted anchor sites for partner interactions. The obvious interaction partners to test were Ski and DNA; however, with both of these the structure of Gli3 repressor domain remained disordered. To locate residues important for the repressor function we mutated several residues within the Gli3 repressor domain. Two of these, H141A and H157N, targeting predicted helical regions, significantly decreased transcriptional repression and thus identify important functional parts of the domain.

Citation: Tsanev R, Vanatalu K, Jarvet J, Tanner R, Laur K, et al. (2013) The Transcriptional Repressor Domain of Gli3 Is Intrinsically Disordered. PLoS ONE 8(10): e76972. doi:10.1371/journal.pone.0076972

Editor: Vladimir N. Uversky, University of South Florida College of Medicine, United States of America

Received June 18, 2013; **Accepted** August 26, 2013; **Published** October 17, 2013

Funding: This study was supported by the Wellcome Trust and Estonian Science Foundation (grant 8116 to P.K., 7279 to T.Ø., and 6847 to K.V.) and Estonian Ministry of Education and Research Targeted Financing Theme SF0690034s09 to J.J. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Torben Østerlund is employed by a commercial company, Zealand Pharma A/S. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: robert.tsanev@ttu.ee

Introduction

The expression of human genes is controlled by numerous transcription factors. Depending on the physiological context genes are activated by transcriptional activators or repressed by transcriptional repressors. During the development of organism, the fine-tuning of gene expression depends on an intricate balance between positive and negative regulators. There are three transcription factors (Gli1, Gli2 and Gli3) in the Sonic hedgehog signal transduction pathway [1]. In their central part these proteins contain a conserved DNA binding domain (DBD) consisting of five zinc-fingers. The structure of the DBD from Gli1 has been solved by X-ray analysis [2]. Here, the zinc-fingers 3, 4 and 5 closely contact the DNA with fingers 4 and 5 determining the target DNA sequence (GACCACCCA) [3] of Gli binding. The first finger does not contact DNA and the second finger only partially interacts with DNA. All the three Gli proteins contain a transcriptional activator domain (TAD) at their C-terminus [4,5]. For transcriptional repression Gli proteins utilize at least two mechanisms. The first one, common to all three Gli proteins, is dependent on Sufu and histone deacetylase [6]. The second one is histone deacetylase independent, involving the domain that we previously identified and named the repressor domain (RD) [7]. Here, we investigate the structure and partner interactions of this domain. The RD is only present in the N-terminus of Gli2 and Gli3 but not in Gli1 [4,7]. An alternative, third mechanism of negative transcriptional regulation by Gli3 has been suggested to involve Ski and histone deacetylation, indicating a general mechanism for all Gli proteins [8]. *In vivo*, Gli3 transcriptional repression has been shown in Sufu knockout mice, suggesting Sufu independent repression mechanism [9].

The proteins or protein domains involved in transcriptional regulation often belong to the class of intrinsically disordered proteins (IDP) or regions (IDR) [10,11]. Structural disorder provides advantages in fulfilling the dynamic processes of gene regulation and signal transduction, and is more frequent in multicellular organisms, suggesting complex regulatory mechanisms [12,13].

In some instances the IDR undergoes binding-coupled folding and becomes functional [14], whereas in other cases a larger degree of fuzziness in the complex is possible [15]. Hence, the IDPs recognise and interact with their partners by a number of mechanisms. A predominant one utilizes short sequence elements displaying higher structural propensities in the disordered, unbound state. These sequence elements are known as molecular recognition features or MoRFs [16]. It has been reported that many IDPs contain low-stability structural elements [17,18].

In this work we have studied the biophysical properties of the RD of Gli3 as previously mapped to span residues G106 to E236 [7]. Our results indicate that it constitutes an intrinsically disordered region. In addition, we have investigated its interactions with various potential partners. Sufu is a known partner and negative regulator of all Gli proteins, but the Gli3 interaction site for Sufu ($S_{333}YGH_{337}$ [19]) does not overlap with the RD. Another partner of Gli3, the Ski protein, has also been linked to the repression

function of Gli3 [8]. The interaction site for Ski has been mapped to N-terminus of Gli3 (1–397). This site potentially overlaps with the RD, thus making Ski a possible target for Gli3RD.

Some transcription factors bind DNA through their IDR [20]. Therefore, we investigated whether Gli3RD, as IDR, interacts with DNA. The transcription factor p53 has two DNA-binding domains: a disordered C-terminal domain (CTD) and a structured core domain [21]. The CTD binds to DNA in a sequence-independent manner and slides along DNA. This assists the core domain in finding its consensus site. The DNA binding of RNA polymerase II (RNA pol II) [22] is facilitated by its predominantly disordered CTD [18]. This domain is proline rich and bears some resemblance to the proline rich sequence of Gli3RD. The DNA binding of RNA pol II CTD is not sequence specific but is dependent on the intercalation of the aromatic ring of tyrosine into the DNA strands [22]. There are eight tyrosine residues in the Gli3RD sequence and some of them have the same spacing between tyrosine and proline as in the CTD of RNA pol II. This suggests that if Gli3RD is binding DNA, it might also be sequence-unspecific. It is known that Gli3 binds to its target genes through the Zn-finger domain that recognises a specific DNA sequence-the Gli consensus site. It has not been observed that Gli3 is able to bind any other sequence in addition to its known DNA binding sequence [23]. We reasoned that if the RD was able to bind DNA, it should occur in a sequence-independent manner, otherwise an additional consensus sequence for Gli3 should have been described.

The Gli3RD contains several histidines. Histidines can coordinate Zn^{2+} ions and according to Karlin [24] there are six classes of histidine ligands. Within Gli3RD, positions H121/H157 and H141/H147 resemble class II and I Zn^{2+}-ligands, respectively. For that reason we decided to test whether these histidines coordinate Zn^{2+}. We mutated these histidines and tested the repressor function of the resultant variants.

A useful method for studying protein folding is in-cell NMR where the spectrum is measured directly inside the living cells in physiologically relevant conditions. Measurements are often carried out by expressing the target protein in *E.coli*. We may suppose that the *E.coli* intracellular environment is more native-like even for an eukaryotic protein than the dilute solution conditions. It is known that the conformation of IDPs may be sensitive to molecular crowding of the environment [25,26,27]. Therefore it is appropriate to carry out the studies as in-cell NMR. The $^{15}N,^{1}H$-HSQC NMR spectrum of IDP has low dispersion of signals in the ^{1}H-dimension within a narrow region around 8 ppm. Certain residues give NMR signals which are easily recognizable due to their distinct positioning. These are the cross-peaks of glycines, the mirrored signal from the side chains of asparagines and glutamines and the signal from the side chain of arginines. The prolines are residues that do not give rise to a signal in the NMR spectrum.

We have described here Gli3RD as an IDR. In regard to its function as a transcriptional repressor we aimed to investigate whether Gli3RD binds Ski or DNA. In addition, we also examined the secondary structure induction of Gli3RD. In a functional assay, H141 and H157 were identified as important functional parts of the domain.

Results and Discussion

Gli3RD is predicted to have both order and disorder features, with a mostly disordered N-terminal part

The transcription factor Gli3 contains an RD in its N-terminus that represses gene expression. We previously identified the Gli3RD to reside between residues G106 to E236 [7] as depicted in Fig. 1A. Sequence analyses of Gli3RD revealed higher than average content of histidines, serines and prolines, particularly in the N-terminal region from residues G106 to D170. This makes the sequence being potentially capable of binding to Zn^{2+} ions, post-translationally modified, and extended and rigid. The Gli3RD sequence appears to be mostly disordered when analysed by disorder-predicting algorithm (VL XT, PONDR, [28]), especially its C-terminal region (Fig. 1D, dashed line), while the N-terminal part appears more ordered. The order-predicting programme, Hierarchical Neural Network, HNN method [29], (Fig. 1B, grey line) also reveals Gli3RD to be predominantly disordered, containing several sequence stretches with higher order probability, mostly within the C-terminal part of Gli3RD. These potentially ordered elements are two short extended strands from positions F173 to I176 and S214 to S217 (Fig. 1B, dashed line) and one short α-helix at position N198 to L207 (Fig. 1B, black line). One potentially extended strand stretching from residue G106 to M111 (Fig. 1B, dashed line) was predicted in the N-terminus of RD. The loss of probability of disorder around position P178 (Fig. 1D, dashed line) coincides with the predicted extended strand at positions F173 to I176 (Fig. 1B, dashed line). The next minimum of disorder probability at position I197 (Fig. 1D, dashed line) corresponds to the predicted α-helix at position N198 to L207 (Fig. 1B, black line). Using ANCHOR programme [30] to predict protein binding sites prone to undergo disorder-to-order transitions, we identified two sites in the C-terminal region from residue D170 to I174 and residues P199-T219 (Fig. 1D, black line). The first site (D170-I174) is close to the predicted extended strand (F173-I176) while the second site (P199-T219) overlaps with the sequence prone to form an α-helix (N198-L207), (Fig. 1B and 1E). These sites can probably undergo structuralisation upon binding to a functional partner, while remaining disordered in a free state. When analysed by the method of Uversky [31] that assesses intrinsic disorder based on the ratio between mean hydrophobicity and mean net charge, Gli3RD as a whole was predicted to be an ordered domain. Accordingly, the hydrophobicity/charge balance does not support strongly disordered regions, as the N-terminal part (residues106–170) lies at the order/disorder boundary and C-terminal part (residues171–236) is situated at the order site (Fig. 1C). The hydrophobisity cluster analysis (HCA) performed with the metaserver MeDor [32], (Fig. 1E) reveals the absence of hydrophobic core which, combined with the high proline content will contribute to the extended state of Gli3RD. The hydrophobic residues within RD colocalise with proline residues to form proline- and hydrophobic residues- rich sites connected by short sequences depleted from these residues. While the proline rich sites are rigid, the linkers seem to be more flexible. Based on the HCA, (Fig. 1E) such linker sequences are D138-R145 and T159-S165. One hydrophobic cluster is seen around position 200, corresponding to the predicted α-helix (Fig. 1E).

It has been found that most of the proteins involved in transcriptional regulation are either completely disordered or contain large regions of intrinsic disorder [20]. Consistent with this, Gli3 also contains extensive regions of intrinsic disorder. The whole N-terminus up to the Zn-fingers (residues 1-480), including RD (residues 106-236), shows high probability for disorder (PONDR, HNN; data not shown). Taken together, we conclude that the RD contains properties of both an ordered and disordered protein, suggesting that folding-upon-binding may occur by utilizing the two C-terminal anchor sites.

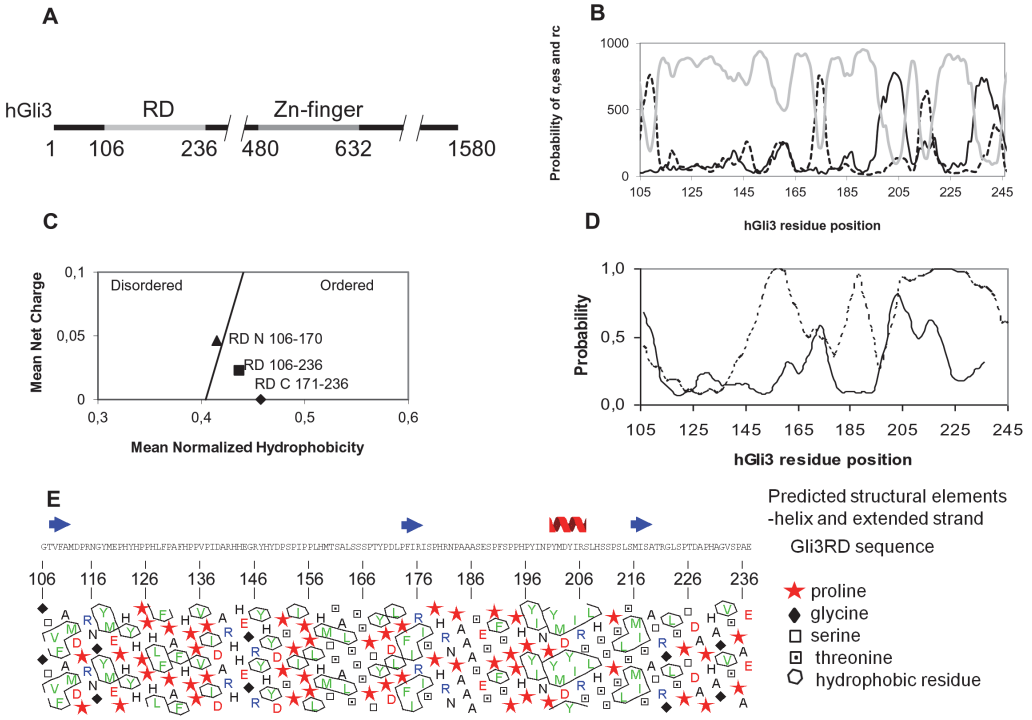

Figure 1. Sequence analysis of human Gli3RD. (A) The relative position of RD within Gli3 from residues G106 to E236. (B) Gli3RD structure prediction with the Hierarchical Neural Network prediction method. The probability for α-helix (α) formation is shown with (-), extended strand (—) and random coil (-). (C) Net charge/hydrophobicity plot of the Gli3RD, as well as of the N- and C-terminus shown separately. The charge/hydrophobicity diagram is divided into two regions by a line (-) corresponding to the equation (R) = 2,743(H)-1,109. Proteins on the left and right side of the diagram are predicted to be disordered and ordered, respectively. (D) Prediction of disorder by PONDR (VL XT algorithm) is shown with a dashed line (—) and the binding regions prediction by ANCHOR is indicated with a line (-). (E) The hydrophobisity cluster analysis combined with secondary structure prediction and amino acid sequence of the human Gli3 repressor domain.
doi:10.1371/journal.pone.0076972.g001

Gli3RD is intrinsically disordered in a native-like environment as determined by in-cell NMR

To determine the structure of Gli3RD, this 15.575-kDa protein domain was His-tagged and expressed in *E. coli*. The NMR spectrum of Gli3RD expressing cells (Fig. 2A) is qualitatively similar to that of the purified protein, measured at pH 7.4 (Fig. 2B). All of the proton peaks lie within a narrow window of 8 ppm that is typical for unfolded peptides. The *E. coli* control cells, not expressing Gli3RD, are shown in Fig. 2C. At physiological pH we detected half of the Gli3RD cross peaks (Fig. 2A) as opposed to the spectrum of the purified protein at pH 5.8 (Fig. 2D). It is possible that a folded fraction bound to a protein partner may have too broad peaks to be seen in the NMR spectrum, or that the missing peaks are not seen due to broadening caused by intra-molecular interaction. However, since Gli3RD is expressed in a heterologous system it most probably has no endogenous binding partner but may interact with the intracellular components unspecifically [33]. The sequence of Gli3RD is proline rich, rendering its structure more rigid and extended thereby making intra-molecular interactions unlikely. Most probably the missing resonances are due to the faster proton exchange occurring at higher pH [34].

In the NMR spectrum of purified Gli3RD we could recognise all five glycines and the side chains of the three asparagines (Fig. 2D), while the signals from the side chains of the seven arginines were visible but overlapping (Fig. 2D). At physiological pH two glycines are detectable (Fig.2B), and they are also visible in the in-cell spectrum (Fig. 2A). This analysis provides supporting evidence that the structure of Gli3RD is disordered in the intracellular environment.

Gli3RD does not bind Ski

It is known that IDRs/IDPs can undergo binding-coupled folding with their functional partners. For that reason we attempted to induce Gli3RD folding using its potential partner Ski, previously reported to interact with the N-terminal part of Gli3 (residues 1–397) that contains the repressor domain (106–236) [8]. To accomplish this, a Ski variant (residues G88 to Y291) was expressed and purified as a maltose-binding (MBP) fusion protein and then added to the ^{15}N-labelled Gli3RD. The NMR spectrum of Gli3RD remained unchanged in the presence of MBP-Ski (data not shown) indicating no major structural changes of Gli3RD in the presence of Ski. We then tested their interaction by co-immunoprecipitation analysis using tagged over-expressed

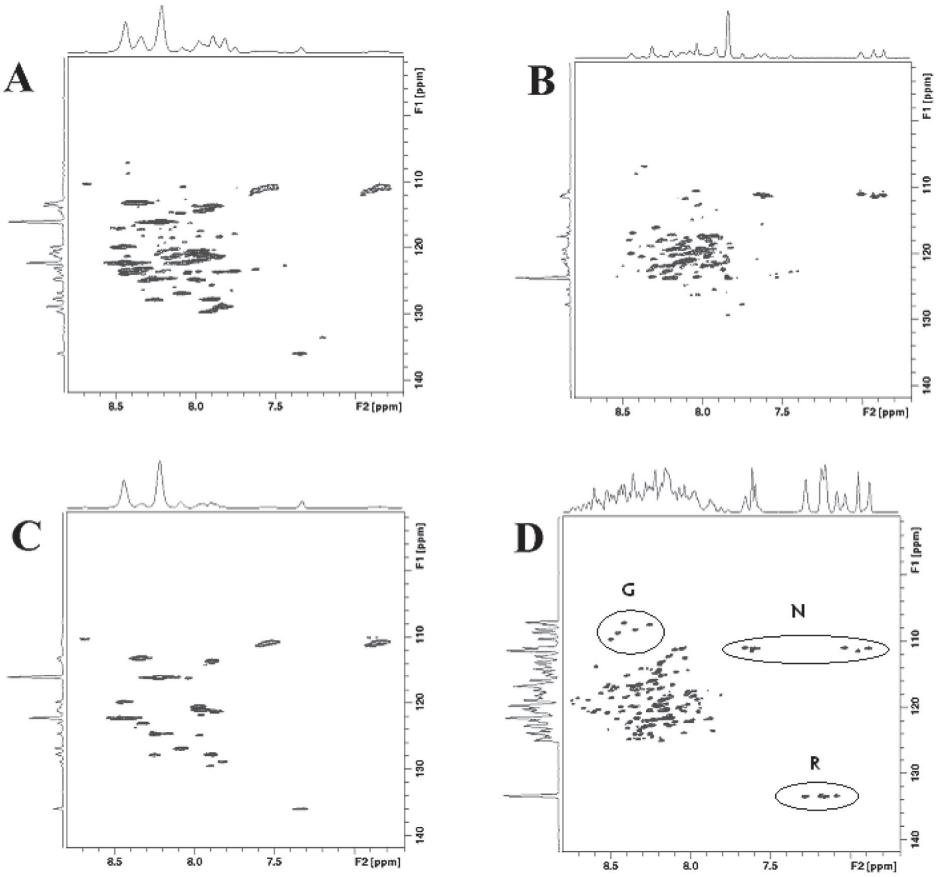

Figure 2. The ^{1}H-^{15}N HSQC NMR spectra. (A) in-cell spectrum of *E.coli* expressing Gli3RD, (B) Gli3RD in solution at pH 7.4, (C) in-cell spectrum of uninduced *E.coli* and (D) Gli3RD in solution at pH 5.8. The cross-peaks from glycines (G), the mirrored cross-peaks from the side chains of asparagines (N) and the signals from the side chain of arginines (R) are denoted in circles.
doi:10.1371/journal.pone.0076972.g002

proteins. However, we did not detect an interaction between Gli3RD and Ski (data not shown). Ski utilizes HDACs to achieve transcriptional repression, suggesting a general mechanism of repression for all three Gli proteins. However, the identified RD does not use HDACs to achieve repression and is found in Gli2 and Gli3 only, but not in Gli1 [7]. Probably the interaction site for Ski is located more upstream from RD, possibly overlapping with the Sufu-binding site ($S_{333}YGH_{337}$). Alternatively, Ski might interact with Gli3 not directly but through Sufu.

Gli3RD does not bind DNA

It has been emphasised that structural disorder is widespread among transcription factors and nucleic acids binding proteins [10,11]. Thus, we subsequently investigated whether GLi3RD interacts with DNA or otherwise. For this we used CD spectroscopy, because this technique has been successfully employed to study the binding of architectural proteins (HMGB1, H1) to DNA [35]. To test if DNA could induce binding-coupled folding of Gli3RD we obtained the CD spectrum of purified recombinant Gli3RD alone or together with a 21 base-pair scrambled sequence DNA oligonucleotide or plasmid DNA to exclude DNA size requirements. After addition of oligonucleotide or plasmid DNA at an stoichiometric ratio of 1:1, no dramatic change in spectral appearance was observed, suggesting that Gli3RD had remained disordered (data not shown). However, although DNA did not induce binding-coupled structure in Gli3RD (CD may not detect low-affinity binding), these two molecules might still interact. We decided to test this possibility by Electrophoretic Mobility Shift Assay (EMSA) and expressed the Gli3RD in HEK293 cells as a Gal4DBD-tagged protein. The cell lysate of HEK293 cells (Fig. 3, lane 2) or a Gli3RD-expressing cells (Fig. 3, lane 3) was incubated with a ^{32}P-labelled 21 base-pairs scrambled DNA oligomers. If Gli3RD was able to bind the labelled oligonucleotide a band corresponding to the DNA-Gli3RD complex should appear. In Fig. 3, comparison of lanes 2 and 3 (with and without Gli3RD, respectively) revealed no additional bands, implying that Gli3RD does not bind DNA. To ascertain that Gli3RD was present in the cell lysate, we used

labelled Gal4BS oligonucleotides binding to the Gal4DBD tag of Gli3RD. The corresponding complex, comprising Gli3RD-Gal4DBD protein / Gal4BS oligo is seen in Fig. 3, lane 5, as a specific band. The specificity of this complex was proved by out-competing with unlabelled Gal4BS oligonucleotide (Fig. 3, lane 7) but not by scrambled oligo (Fig. 3, lane 6). We noticed that Gli3RD did not alter the DNA binding of its tag, Gal4DBD. Therefore, we exclude the mechanism where Gli3RD functions at the DNA level, by preventing the Gli3 Zn-finger to bind to DNA.

The repression activity of Gli3RD is lost in H141A and H157N mutants

Karlin and Zhu describe six classes of Zn-ligands [24]. The positions of H121/H157 and of H141/H147 resemble classes II and I, respectively, of Zn^{2+}-binding ligands. We investigated whether the repression function of Gli3RD is dependent on Zn^{2+}-binding. This was addressed in a functional assay by comparing the activities of wild type to histidine-mutated Gli3RD variants. For such mutational analysis we chose H121 and H157 since they resemble the first and the third histidines from class II site of Zn^{2+}-binding ligands (H_{121}xH and a third $histidine_{157}$ distant in the sequence) [24]. We also selected H141 and H147 because they resemble the third and the second histidines from a class I site ($HExxH_{141}xxGxxH_{147}$). Moreover, the two sites are also positioned in two sequence-stretches predicted by HNN to form low-populated α-helices (and lower than the helix in the C-terminal part). We mutated H121, H141, H147 and H157 to alanines separately and in a double mutant H121/147A, where both classes of Zn^{2+}-ligands were expected to be affected. In addition, the H157 was substituted for asparagines, since asparagine and histidine have similar, but not identical, space requirements and hydrogen bonding capabilities. However, asparagine is unable to coordinate Zn^{2+}-ions. Asparagine was also chosen because, contrary to H157A, its substitution significantly alters the predicted local helicity at that site which might be involved in partner recognition. All proteins were expressed as tagged versions from DNA in HEK293 cells. The constructs encoded Gal4 DBD-fusion proteins and we measured their effect on transcriptional activity of Gal4 binding site-containing luciferase reporter (Fig. 4A).

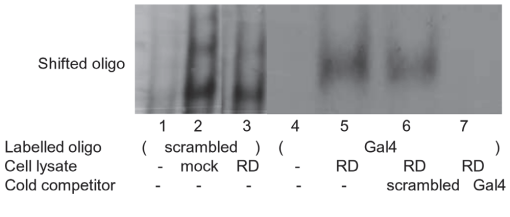

Figure 3. Electrophoretic mobility shift analysis of Gli3RD and DNA binding. Cell lysate from Gli3RD or mock-transfected HEK293 cells incubated with labelled scrambled sequence oligonucleotide is shown in line 3 and line 2, respectively. As a negative control, the labelled scrambled oligonucleotide was loaded alone, without a cell lysate, on line 1. Labelled Gal4 binding site (BS) oligonucleotides were used to confirm the presence of Gal4 DBD tagged Gli3RD. On line 4 the Gal4 BS oligonucleotide is loaded alone. The DNA-protein complex formed by the Gal4 BS oligonucleotide and Gal4 DBD tagged Gli3RD is indicated on line 5. The shifted complex was competed out by unlabelled Gal4 BS oligonucleotide (line 7). Scrambled sequence oligonucleotide did not compete out the complex (line 6).
doi:10.1371/journal.pone.0076972.g003

The mutation of the first histidine (H121) from class II did not affect the repression function. The second histidine mutation (H147) from class I, also left the repression function unaltered. To verify that repression is Zn^{2+}-independent we combined the H121 and H147 mutations in a double mutant H121/147A. In this mutant, both putative Zn^{2+}-ligand classes were affected. Despite targeting both classes I and II, the repression function remained intact. The activity of the double mutant is slightly different but falls within the range of other mutants (Fig. 4A).

Upon mutating the histidines H141 or H147 that resemble the class I Zn-ligands, we observed a significant loss of repression activity in the H141A only. Thus, it is possible that H141 participates in Zn^{2+} coordination together with a more distant histidine. On the one hand, Gli3RD is a proline rich sequence, making its structure extended and therefore any long-range intramolecular interactions more difficult. On the other hand, H141 is at the peak of a low-populated predicted α-helix that reaches a minimum at H147 (Fig 4B). Since the H141A significantly changes the structural ensemble of the above α-helix and correspondingly affects the repressor function, we speculate that it might constitute a protein interaction site or functionally important region with distinct structural dynamics. In the case of H157, the activity of the domain was preserved in H157A variant whereas it was lost in H157N. By analogy with H141A, while the H157N variant increases the predicted helicity by more than 300%, the same does not hold true for the functionally intact H157A variant. Therefore we speculate that the change in the structural ensemble of the predicted α-helix, may respectively either impair or enhance partner recognition via altering a specific MoRF [16]. This remains to be proven when the Gli3RD interaction partner is known. The observed loss of activity in the H141A and H157N variants points that these residues most probably constitute important functional parts of the repressor domain.

Our analyses were performed using the RD expressed alone (fused to a 6xHis-tag) and not as part of the full-length Gli3 protein or as a larger part encompassing the RD and the surrounding region. It cannot be ruled out that in the composition of the entire protein this domain has diminished spatial freedom due to stabilizing inter-domain connections and is therefore more folded than we observed [36]. However, homologous domains are present in different proteins, meaning that the information about their function and folding is mainly coded within the domain sequence, and only to a minor extent by the surrounding context of the protein. It cannot be ruled out that, although the domain appears disordered on its own, it can undergo structural changes upon interaction with a binding partner in the context of the full-length protein only. Nevertheless, as the Gli3RD domain on its own has repressor function (Fig. 4A; [7]), the latter is not likely.

In conclusion, we have described the Gli3RD as an intrinsically disordered domain. The RD does not bind and fold in the presence of neither Ski nor DNA. The H141A and H157N mutants point to important functional parts of the repressor domain, possibly involving α-helical MoRFs. The H157 shows partial functional plasticity, tolerating its substitution for alanine but not for aspargine.

The pursuit for binding partner and complex structure will continue with this and other variants of the Gli3RD. An interesting option would be to study the repressor complex forming on the promoter region of one of the target genes of GLI proteins, i.e. *PTCH* or *GLI1*, as well as the one forming with the DBD itself.

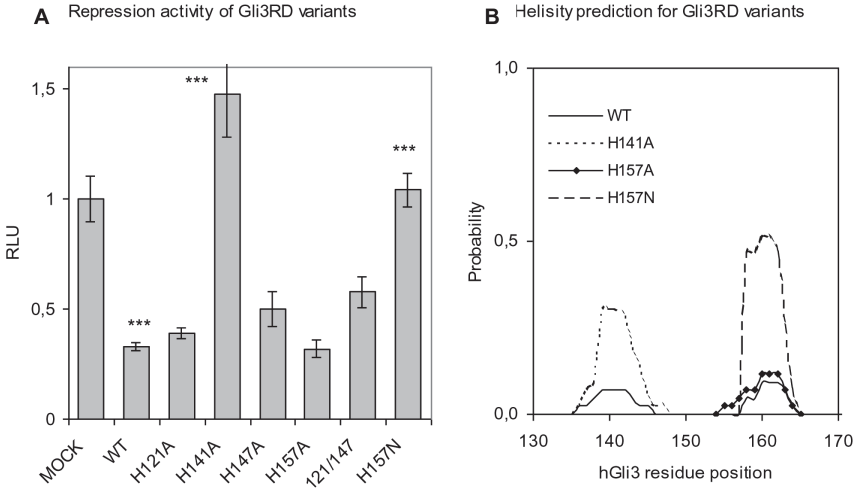

Figure 4. The luciferase reporter activity together with different Gli3RD variants. (A) All Gli3RD constructs were expressed as Gal4 DBD fusions in HEK293 cells and their repressor activity is compared to that of Gal4 DBD alone (mock). The activity of Gla4-Gli3RD is shown as WT. The analyses were performed three to five times and error bars show standard deviations. Statistically significant were WT vs MOCK, WT vs H141A and WT vs H157N, $p<0{,}001$, One-Way ANOVA with Dunnett's Multiple Comparisons Test. (B) The Agadir prediction of helical behaviour for Gli3RD variants.
doi:10.1371/journal.pone.0076972.g004

Materials and Methods

Sequence analyses

Sequence analyses were carried out as described by Uversky [31]. The mean net charge of a protein was determined as the absolute value of the difference between the numbers of positively and negatively charged residues divided by the total number of residues. The mean hydrophobicity was defined as the sum of the normalised Kyte-Doolittle hydrophobicities, divided by the total number of residues. These values were then plotted together with disorder-order boundary, which was defined as (charge) = 2.743 (hydrophobicity) – 1.109. Sequences were also analysed using series neural network predictors as implemented in PONDR VL-XT (pondr.com) [28]. For the secondary structure prediction the Hierarchical Neural Network prediction method was used [29] (http://npsa-devel.ibcp.fr/NPSA/npsa_hnn.html).

DNA constructs

For production of recombinant protein the cDNA for human Gli3RD (residues 106 to 236) was cloned into pET11C (Novagen, USA). An N-terminal 6xHis-tag was designed into the primer followed immediately by the RD sequence. The cDNA for human Ski (residues 88 to 291) was cloned into pMAL vector. For expression of Gli3RD in mammalian cells, we used the pFA Gli3RD construct described in [7] where the Gli3RD cDNA was in frame with the DBD of yeast Gal4 (Stratagene, La Jolla, CA). We used the pMN-Luc plasmid containing a thymidine kinase promoter with five tandem repeats of the yeast GAL4 binding sites that control expression of the firefly luciferase gene as a Gal4 reporter construct. The Gli3RD mutants were cloned in pFA vector produced as described in QuikChange protocol (Stratagene).

Protein expression and purification

Expression of recombinant 6xHis-Gli3RD (Gli3RD) was carried out in *E. coli* strain BL21. For ^{15}N labelling the cells were grown in minimal medium supplemented with glucose and $^{15}NH_4Cl$ in a volume of 1 litre in fermentor and expressed for 5 hours by induction with 0.5 mM IPTG at 30°C. Cells were harvested by centrifugation. The pellet was resuspended in 40 ml of ice-cold lysis buffer (100 mM NaH_2PO_4; 10 mM Tris·Cl; 8 M urea pH 8). The cells were lysed by 3 cycles of freeze-and-thaw in liquid nitrogen and cold-water bath. Then, the sample was subjected to sonication with 5 s ON and 15 s OFF cycles for 90 s on ice water bath. The cell lysate was cleared by centrifugation for 15 min, 4°C, 10 000 rpm. The supernatant was transferred to a new vial and 2 ml of lysis buffer-washed Ni-resin (Qiagen, Hilden, Germany) was added. The binding was carried out at 4°C for 20 min. The Gli3RD bound resins were collected by centrifugation for 10 min, 4°C, 6 000 rpm. The resin was washed 2 times with 10 ml of washing buffer (50 mM NaH_2PO_4; 300 mM NaCl; 20 mM imidazole; pH 8) and collected by centrifugation for 10 min, 4°C, 6 000 rpm. The Gli3RD was eluted with 2 ml of elution buffer (50 mM NaH_2PO_4; 300 mM NaCl; pH 4,5) and then further purified with C_{18} reversed-phase chromatography (described in Methods S1) and lyophilized until used. The reversed-phase chromatograms are shown in Figure S1 and Figure S3, the mass spectrum is shown in Figure S2. The final yield of purified Gli3RD was 4 mg per 1 g of biomass, the SDS-PAGE shown in Figure S4.

The soluble Ski was expressed in BL21 cells that have been transformed with pMal plasmid (New England Biolabs, Ipswich, MA). Ski was tagged with MBP and GFP to facilitate expression and purification process. The expression of the recombinant protein was indused with 0.5 mM IPTG for 16 h at 20°C in LB medium. The centrifuged cells were resuspended in 40 ml of ice-cold column buffer (20 mM Tris-HCl; 200 mM NaCl; 1 mM EDTA pH 7.4). Cells were lysed with lysozyme 1 mg/ml for 30 min on ice/water bath, followed by freeze-and-thaw and sonication as for Gli3RD. The cleared lysate was incubated with 1 ml of column buffer washed amylose resin (New England

Biolabs). The binding was carried out at 4°C for 20 min. Ski-bound resin was washed 2 times with 10 ml of column buffer and eluted with 1 ml column buffer containing 10 mM maltose. The eluate was kept at 4°C until further use and concentrated to 0.05 mM on Amicon Ultra-0.5 mL Centrifugal Filters (Millipore, Billerica, MA, USA).

Immunoprecipitation

To test the association of Gli3RD protein and Ski protein we immunoprecipitated one protein and detected whether the other one was co-precipitated. Both proteins were over-expressed as tagged proteins in 293HEK cells. The cells were transfected in 15 cm culture dishes. For transfection the DNA/PEI complex was prepared as follows: 30 µg of Gli3RD or Ski encoding DNA was mixed with 60 µg of PEI per plate in 500 µl of DMEM. After 10 min of incubation 9 ml of DMEM was added to the DNA/PEI mixture and then applied to the culture dish. After 2 h the medium was exchanged for DMEM supplemented with 10% FCS. On the following day, the medium was changed again and after an additional 24 h cells were lysed in 1 ml PBS with 1% Triton X-100 (Sigma–Aldrich). The lysate was cleared by centrifugation at 13 000 rpm for 15 min at 4°C.

For immunoprecipitation of Gli3RD, 5 µl of anti-Gal4 polyclonal antibodies (Santa Cruz Biotechnology, Santa Cruz, CA, USA) or for Ski, 5 µl of anti-GluGlu polyclonal antibodies (Abcam, Cambridge,UK) were incubated with 30 µl Protein G agarose (Amersham Biosciences, Bucks, UK) for 1 h at 4°C. After that 1 ml of cell lysate was added and immunoprecipitation was performed at 4°C overnight and analyzed by Western blot using Gal4 monoclonal antibodies (Santa Cruz Biotechnology), to detect Gli3RD or GluGlu monoclonal antibodies (Abcam, Cambridge,UK) to detect Ski.

Electrophoretic Mobility Shift Assay

To test the association of Gli3RD with DNA we incubated cell lysate of HEK293 cells expressing Gli3RD and oligonucleotides containing either Gal4 binding site (Gal4BS) or scrambled sequence. The transfected cells were washed with PBS and cells were collected from the 10 mm plate in 0.5 ml of whole cell extraction buffer (20 mM Hepes-KOH, pH 7.9, 400 mM KCl, 1 mM EDTA, 10% (v/v) Glycerol, with freshly added 10 mM DTT, 1 mM PMSF and 1× Complete protease inhibitor cocktail, Roche Diagnostics GmbH, Mannheim, Germany). The cells were lysed in 3 cycles of freezing and thawing in liquid nitrogen and ice bath. The lysate was cleared by 15 min centrifugation, 4°C, 14 000 rpm. The supernatant was aliquoted and kept in −80°C until needed. The untransfected cells used as a negative control were treated likewise.

The two strands of Gal4BS or scrambled oligonucleotides were annealed and labelled with Klenow fragment of DNA polymerase (Bioron, HeidelbergGermany). The labelling reaction was set as follows: 10.5 µl H_2O, 1.5 µl Reaction buffer (Bioron), 1 µl oligonucleotides (from 4 µM annealed stock), 1 µl [α-^{32}P] dCTP (GE Healthcare, UK) and 1 µl Klenow. The reaction was incubated for 30 min at 37°C. Then, 2 µl dCTP (10 mM stock) was added and incubated for an additional 10 min at 37°C. For purification of labelled oligonucleotides from unincorporated label 80 µl Tris-EDTA (10 mM Tris-HCl, pH 7.5; 1 mM EDTA) was added to the probe and the total of 100 µl was run over a NAP-5 Sephadex G25 column (GE Healthcare, London, UK). The column was washed with 400 µL Tris-EDTA. Additional 0.5 ml of Tris-EDTA was added to elute the labelled oligonucleotides.

Band shift binding reactions were assembled by adding 10 µl binding buffer (2x stock: 100 mM HEPES, pH 7.4; 100 mM KCl; 10 mM $MgCl_2$; 20 µM $ZnSO_4$; 2 mM DTT; 40% glycerol) [2], 1 µl cell lysate, 2 µl labelled Gal4 BS or scrambled oligonucleotides and H_2O to the total volume of 20 µL. While using Gal4BS oligonucleotides 1 µl of poly-(dI-dC) (1 µg/µL, Sigma) was added to eliminate unspecific binding to DNA. To test the specificity of the binding complex 2 µl (from 4 µM stock) of unlabelled Gal4BS oligonucleotides or unlabelled scrambled oligonucleotides were added, and preincubated for 15 min at room temperature. Then all the band shift binding reactions were incubated at room temperature for 15 min and loaded on 5% nondenaturing polyacrylamide gel in 1x TBE buffer. Before loading, the gel was pre run for 1.5 h, 100 V at 4°C. After loading, the gel was run for 2 h, 200 V at 4°C. and then dried on a Whatman paper (Whatman Ltd, Maidstone, UK) and visualised on a Roentgen film (Agfa HealthCare,Mortsel, Belgium). The gel and the film were assembled in cassette (Kodak, New York, NY USA) and exposed at −70°C for 24 hours.

Cell culture

HEK293 cells were grown and transfected as previously described [7]. The cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% Fetal Calf Serum FCS (PAA Laboratories, Pasching, Austria), streptomycin and penicillin (100 units/ml; Invitrogen, Carlsbad, CA). Cells were grown at 37°C and 5.0% CO_2 in cell culture incubator. One day before transfection the cells were splitted into the required plates.

Luciferase assays

A DNA plasmid encoding luciferase gene with upstream Gal4 binding sites was used as a reporter. The effector plasmid contained Gli3RD DNA, cloned in frame with Gal4 DNA binding domain (DBD). The plasmids were transfected into mammalian cells. The effector plasmid expressed a fusion protein consisting of Gal4 DBD and Gli3RD. This fusion protein bind to the Gla4 binding sites of the reporter plasmid through the Gal4 DBD and alters the expression of the liciferase gene. We compare the activity of the reporter in the presence and absence of Gla4-Gl3RD. Transfections for luciferase assays were performed in 24-well plates. The Gli3RD WT and mutated constructs were transfected in HEK293 cells. For transfection efficiency control the β-gal construct was used. The amount of reporter plasmid (pMN-Luc) and the effector plasmids (pFA Gal4 fusions expressing RD or RD mutants) used were 300 ng and 100 ng per well, respectivly. For normalization we used 50 ng of pCMV-β-gal. As a transfecting agent we used Polyethyleneimine (PEI; Sigma-Aldrich, St. Louise, USA) 1 µg per well. DNA and PEI were mixed in 50 µl of DMEM. An additional 150 µl of DMEM was added to the DNA/PEI mixture and then applied to the cells. After 2 hours of incubation the medium was exchanged by DMEM with 10% FCS. On the following day the medium was changed again. Cells were harvested after an additional 24 hours of incubation. Firefly luciferase and β-gal assays were performed in Ascent FL fluoroscan with the Luciferase Assay Kit (BioTherma, Dalarö, Sweden) and Galacto-Light Plus System (Applied Biosystems, Foster City, CA). One-Way ANOVA with Dunnett's Muliple Comparisons Test (GraphPad) was used to determine the statistical significance of differential findings between experimental groups.

CD and NMR analyses of Gli3RD

The in-cell NMR measurements were carried out with cells prepared as described by Serber [37]. The recombinant *E.coli* was grown to OD_{600}~1 in LB medium, then the cells were pelleted by

centrifugation (3750 g for 3 min) and transferred to the expression medium (M9) containing $^{15}NH_4Cl$ and glycerol. After induction with 0.5 mM IPTG for 7 h at 30°C the cells were pelleted (20 min at 800 g), transferred to PBS buffer containing 10% D_2O and kept on ice until the NMR measurements.

For NMR measurements the lyophilised Gli3RD was resuspended in 50 mM phosphate buffer pH 5.8 or PBS containing 10% D_2O, to a final concentration of 0.13 mM. NMR spectra were acquired using a Bruker Avance III spectrometer operating at 800 MHz proton resonance frequency. ^{1}H-^{15}N HSQC spectra were recorded using 79 ms acquisition time in the ^{1}H dimension and 51 ms acquisition time in the ^{15}N dimension. 4 repetitions were averaged for 256 increments in the indirect dimension. Spectral widths were 8000 Hz for ^{1}H and 2500 Hz for the ^{15}N dimension. All NMR data processing was performed using TopSpin software (Bruker, Germany).

The CD measurements of DNA binding to Gli3RD were assembled in binding buffer (20 mM HEPES pH 7.9, 80 mM KCl and 4 mM $MgCl_2$) by adding equimolar concentrations of lyophilised protein and 21 bp scrambled sequence synthetic oligonucleotides or a plasmid DNA to exclude DNA size requirements. The concentration of the protein was measured spectrophotometrically on ND-1000 (NanoDrop Technologies, Wilmington, DE, USA). CD spectra were obtained with a Jasco J-720 spectropolarimeter (Jasco, Easton, MO) and the temperature was controlled with a PTC-343 temperature controller. A quartz cell with 2 mm optical path was used. The spectral range was 190 – 250 nm with a resolution of 0.2 nm and a bandwidth of 2 nm. A scan speed of 50 nm/min with 2 s response time was employed. The buffer-background spectrum was subtracted.

Supporting Information

Figure S1 The UV chromatogram of the sample Gli3RD, 280 nm. The peak RT 16,52 min was identified as protein with MW 15570.9 Da on the basis of the ESI-MS spectrum deconvulated with the MagTrans software.
(TIF)

Figure S2 The ESI-MS spectrum of Gli3RD.
(PDF)

Figure S3 Overlaid chromatograms of the sample Gli3RD. Black - UV 280 nm, red - TIC m/z 300 – 2000 Da.
(TIF)

Figure S4 The SDS-PAGE of Gli3RD. The gel was made to 12% acrylamide, bis-acrylamide 29:1.
(JPG)

Methods S1 HPLC purification.
(DOC)

Acknowledgments

We thank Signe A. Sjørup and Kersti Olspert for excellent technical assistance and Richard Tamme, Leonora Brunie and Katrin Tomson for fruitful discussions during the study.

Author Contributions

Conceived and designed the experiments: R. Tsanev KV JJ KL PT BK TØ PK. Performed the experiments: R. Tsanev KV JJ R. Tanner KL PT. Analyzed the data: R. Tsanev KV JJ KL PT BK TØ PK. Contributed reagents/materials/analysis tools: KV JJ R. Tanner BK TØ PK. Wrote the paper: R. Tsanev KV JJ R. Tanner KL BK TØ.

References

1. Briscoe J (2009) Making a grade: Sonic Hedgehog signalling and the control of neural cell fate. Embo J 28: 457–465.
2. Pavletich NP, Pabo CO (1993) Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers. Science 261: 1701–1707.
3. Kinzler KW, Vogelstein B (1990) The GLI gene encodes a nuclear protein which binds specific sequences in the human genome. Mol Cell Biol 10: 634–642.
4. Sasaki H, Nishizaki Y, Hui C, Nakafuku M, Kondoh H (1999) Regulation of Gli2 and Gli3 activities by an amino-terminal repression domain: implication of Gli2 and Gli3 as primary mediators of Shh signaling. Development 126: 3915–3924.
5. Kalff-Suske M, Wild A, Topp J, Wessling M, Jacobsen EM, et al. (1999) Point mutations throughout the GLI3 gene cause Greig cephalopolysyndactyly syndrome. Hum Mol Genet 8: 1769–1777.
6. Cheng SY, Bishop JM (2002) Suppressor of Fused represses Gli-mediated transcription by recruiting the SAP18-mSin3 corepressor complex. Proc Natl Acad Sci U S A 99: 5442–5447.
7. Tsanev R, Tiigimagi P, Michelson P, Metsis M, Osterlund T, et al. (2009) Identification of the gene transcription repressor domain of Gli3. FEBS Lett 583: 224–228.
8. Dai P, Shinagawa T, Nomura T, Harada J, Kaul SC, et al. (2002) Ski is involved in transcriptional regulation by the repressor and full-length forms of Gli3. Genes Dev 16: 2843–2848.
9. Wang C, Pan Y, Wang B (2010) Suppressor of fused and Spop regulate the stability, processing and function of Gli2 and Gli3 full-length activators but not their repressors. Development 137: 2001–2009.
10. Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, et al. (2007) Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J Proteome Res 6: 1882–1898.
11. Tompa P, Dosztanyi Z, Simon I (2006) Prevalent structural disorder in E. coli and S. cerevisiae proteomes. J Proteome Res 5: 1996–2000.
12. Romero PR, Zaidi S, Fang YY, Uversky VN, Radivojac P, et al. (2006) Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. Proc Natl Acad Sci U S A 103: 8390–8395.
13. Schlessinger A, Schaefer C, Vicedo E, Schmidberger M, Punta M, et al. (2011) Protein disorder—a breakthrough invention of evolution? Curr Opin Struct Biol 21: 412–418.
14. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6: 197–208.
15. Fuxreiter M, Simon I, Bondos S (2011) Dynamic protein-DNA recognition: beyond what can be seen. Trends Biochem Sci 36: 415–423.
16. Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, et al. (2006) Analysis of molecular recognition features (MoRFs). J Mol Biol 362: 1043–1059.
17. Hinds MG, Smits C, Fredericks-Short R, Risk JM, Bailey M, et al. (2007) Bim, Bad and Bmf: intrinsically unstructured BH3-only proteins that undergo a localized conformational change upon binding to prosurvival Bcl-2 targets. Cell Death Differ 14: 128–136.
18. Bienkiewicz EA, Moon Woody A, Woody RW (2000) Conformation of the RNA polymerase II C-terminal domain: circular dichroism of long and short fragments. J Mol Biol 297: 119–133.
19. Dunaeva M, Michelson P, Kogerman P, Toftgard R (2003) Characterization of the physical interaction of Gli proteins with SUFU proteins. J Biol Chem 278: 5116–5122.
20. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, et al. (2006) Intrinsic disorder in transcription factors. Biochemistry 45: 6873–6888.
21. Tafvizi A, Huang F, Fersht AR, Mirny LA, van Oijen AM (2010) A single-molecule characterization of p53 search on DNA. Proc Natl Acad Sci U S A 108: 563–568.
22. Suzuki M (1990) The heptad repeat in the largest subunit of RNA polymerase II binds by intercalating into DNA. Nature 344: 562–565.
23. Vokes SA, Ji H, Wong WH, McMahon AP (2008) A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. Genes Dev 22: 2651–2663.
24. Karlin S, Zhu ZY (1997) Classification of mononuclear zinc metal sites in protein structures. Proc Natl Acad Sci U S A 94: 14231–14236.
25. Dedmon MM, Patel CN, Young GB, Pielak GJ (2002) FlgM gains structure in living cells. Proc Natl Acad Sci U S A 99: 12681–12684.
26. McNulty BC, Young GB, Pielak GJ (2006) Macromolecular crowding in the Escherichia coli periplasm maintains alpha-synuclein disorder. J Mol Biol 355: 893–897.
27. Johansen D, Jeffries CM, Hammouda B, Trewhella J, Goldenberg DP (2011) Effects of macromolecular crowding on an intrinsically disordered protein characterized by small-angle neutron scattering with contrast matching. Biophys J 100: 1120–1128.
28. Li X, Romero P, Rani M, Dunker AK, Obradovic Z (1999) Predicting Protein Disorder for N-, C-, and Internal Regions. Genome Inform Ser Workshop Genome Inform 10: 30–40.

29. Guermeur Y, Geourjon C, Gallinari P, Deleage G (1999) Improved performance in protein secondary structure prediction by inhomogeneous score combination. Bioinformatics 15: 413–421.
30. Dosztanyi Z, Meszaros B, Simon I (2009) ANCHOR: web server for predicting protein binding regions in disordered proteins. Bioinformatics 25: 2745–2746.
31. Uversky VN, Gillespie JR, Fink AL (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins 41: 415–427.
32. Lieutaud P, Canard B, Longhi S (2008) MeDor: a metaserver for predicting protein disorder. BMC Genomics 9 Suppl 2: S25.
33. Crowley PB, Chow E, Papkovskaia T (2011) Protein interactions in the Escherichia coli cytosol: an impediment to in-cell NMR spectroscopy. Chembiochem 12: 1043–1048.
34. Zhang YZ, Paterson Y, Roder H (1995) Rapid amide proton exchange rates in peptides and proteins measured by solvent quenching and two-dimensional NMR. Protein Sci 4: 804–814.
35. Polyanichko AM, Chikhirzhina EV, Andrushchenko VV, Vorob'ev VI, Wieser H (2006) The effect of manganese(II) on the structure of DNA/HMGB1/H1 complexes: electronic and vibrational circular dichroism studies. Biopolymers 83: 182–192.
36. Batey S, Clarke J (2008) The folding pathway of a single domain in a multidomain protein is not affected by its neighbouring domain. J Mol Biol 378: 297–301.
37. Serber Z, Corsini L, Durst F, Dotsch V (2005) In-cell NMR spectroscopy. Methods Enzymol 394: 17–41.

SUMMARY

Gli proteins are the effector transcription factors of Hedgehog signalling. There are three Gli proteins – Gli1, Gli2 and Gli3 – that control the expression of the pathway target genes. Gli proteins regulate cell proliferation and specification during development of different embryonic systems and organs. In adults, reactivation of Gli transcription factors is associated with various types of cancer. Thus, studying the transcriptional regulation of this pathway will help to understand in more detail the complex processes of normal and aberrant organism development and also cancer formation.

Gli transcription factors have many cellular functions including DNA-binding, transcriptional activation and repression, localisation signals and more. The domain responsible for the transcriptional repressor activity of Gli proteins was defined in this thesis.

The intracellular trafficking of Gli3 can be visualised with the recombinant anti-Gli3 intrabody generated in this study. In disease conditions, the transcriptional response of Hedgehog pathway could be guided with the help of this intrabody.

Here, the repressor domain of human Gli proteins was localised to aa residues G29 to G170 of Gli2 and G106 to E236 of Gli3. This domain is conserved between Gli3 and Gli2, and is absent in Gli1. Deletion of this region within the context of full-length Gli3 converted the construct to a better transcriptional activator. This result suggests that the repressor domain of Gli3 is communicating with and influencing the activity of the activator domain. The mechanism of repression used by this domain was shown to be independent of histone deacetylation and not to involve binding to the Ski co-repressor protein. The site on Gli3 responsible for the interaction with Ski should be, thus, outside the repressor domain. The Ski binding site could be close to or overlapping with the Sufu binding site, which is conserved within all Gli proteins. Since both Ski and Sufu way of repression depend on histone deacetylation, they can be a part of a more general pathway for Gli regulation common to all Gli proteins. Thus, Ski binding will result in removal of all Gli proteins and the signal ceasing.

The repressor domain of Gli3 was characterised to be intrinsically disordered. The disorder, however, was predicted to extend almost to the whole N-terminus of Gli3. It is common for proteins involved in transcriptional regulation to contain large unfolded regions. Structure prediction programs display short regions with higher structural propensities that can represent recognition elements for partner interaction – MoRFs. In a free, unbound state, these sequence elements of Gli3 repressor domain were shown to be disordered. Their folding will be induced and stabilised by intermolecular connections between the

binding partners. The interactions between intrinsically disordered proteins are usually weak and transient. Therefore, folding can be a good indication for complex formation when studying Gli3 binding partners. Still, there are also instances of fuzziness even in the bound, complex state.

Gli3 repressor domain functions in the cell nucleus without binding to DNA itself. The mechanism of transcriptional repression is not at DNA level by masking DNA or by interfering with the binding of transcriptional activators to DNA. Instead, the Gli3 repressor domain presumably functions at the protein level, very likely by directly binding to or influencing the function of the Mediator complex. The Mediator complex is summing up signals coming from transcription factors to modulate the activity of RNA polymerase II to a proper level of gene transcription.

Certain histidines within the amino-acid sequence of Gli3 repressor domain resemble zinc binding ligands. Mutation of these histidines provided evidence that zinc binding is not connected with the repression function of this domain. Instead, amino-acid residues H141 and H157 were found to constitute functionally important parts of Gli3 repressor domain. Changing the structural dynamics of the corresponding region by mutating these amino-acid residues, abolished the repressor function.

Conclusively, the results presented in this thesis help to understand the process of transcriptional repression of Gli proteins. They are also the foundation for further studies into Gli protein function and gene regulation. It will be important to further characterize the repressor function in a cellular context and identify the interactions this domain has. An important deduction of the studies has been the separation of Ski and Sufu induced negative regulation of Gli proteins (probably involving HDACs) from the more acute repressor function defined by the studied domain.

KOKKUVÕTE

Gli valgud on efektortranskriptsiooni faktorid hedgehog-signaalülekande rajas. Sellesse perekonda kuulub kolm valku: Gli1, Gli2 ja Gli3, mis kontrollivad raja sihtmärkgeenide ekspressiooni. Organismi embrüonaalse arengu käigus osalevad Gli valgud rakkude jagunemist ja spetsialiseerumist reguleerides. Mitmete vähihaiguste puhul on ilmnenud Gli transkriptsioonifaktorite taasaktiveerumine täiskasvanud organismis. Sellega seoses aitaks hedgehog-raja transkriptsioonilise regulatsiooni uurimine paremini mõista organismi arengu keerukaid protsesse nii normaalse arengu kui ka patoloogia ja vähi puhul.

Gli valkudel on palju rakulisi funktsioone, nagu DNA sidumine, geenide aktivatsioon või selle mahasurumine, erinevad seondumis- ning lokaliseerimissignaalid jms. Käesolevas töös defineeriti Gli valkude piirkond, domeen, mis vastutab transkriptsioonilise repressiooni eest.

Gli3 valgu rakusisest liikumist on võimalik visualiseerida käesolevas töös väljatöötatud Gli3-vastase intrakeha abil. Haigusseisundi puhul oleks võimalik suunata selle raja transkriptsioonilist aktiivsust, kasutades nimetatud intrakeha.

Töö käigus lokaliseeriti inimese Gli2 transkriptsiooniline repressordomeen aminohappe jääkide vahemikku G29 kuni G170 ja Gli3 repressordomeen vahemikku G106 kuni E236. See valgupiirkond on konserveerunud Gli3 ja Gli2 vahel ning puudub Gli1's. Selle domeeni eemaldamine täispika Gli3-valgu koosseisust muutis konstrukti paremaks aktivaatoriks. Siit lähtuvalt toimub Gli3 aktivaator- ja repressordomeenide vahel kommunikatsioon, mille tulemusena repressordomeen mõjutab aktivaatordomeeni funktsioneerimist. Selles töös näidati, et transkriptsioonilise repressiooni mehhanism on sõltumatu histoonide deatsetüleerimisest ega vaja Ski repressorvalgu kaasamist. Seega peaks Ski seondumise koht Gli3'ga olema väljaspool repressordomeeni piire. Võimalik, et Ski seondumiskoht on Sufu seondumiskoha läheduses või isegi sellega kattuv. Kuna mõlemad valgud, nii Ski kui ka Sufu, kaasavad histoondeatsetülaase transkriptsiooni mahasurumiseks, siis on tõenäoline, et nad on osa ühtsest regulatsioonimehhanismist. See mehhanism kehtib ilmselt kõikide Gli valkude puhul, kuna Sufu'ga seonduvad kõik Gli valgud. Nii kutsuks Ski seondumine Gli valkudega esile nende eemaldamise ja sellega kaasneva signaalikatkestuse.

Töös iseloomustati Gli3 repressordomeeni kui sisemiselt korrastamatut valgupiirkonda. Korrastamatust ennustati aga peaaegu kogu Gli3 N-terminusele. Transkriptsioonis osalevaid valke iseloomustab täielik korrastamatus või suurte korrastamatute regioonide olemasolu. Kõrgenenud korrastatusega lühikeste regioonide esinemine repressordomeeni järjestuses tuvastati struktuuriennustusprogrammide abil. Nimetatud lühikesed regioonid on suure tõenäosusega

partneri äratundmise elemendid – MoRF'id. Näidati, et vabas, sidumata olekus on need järjestuselemendid korrastamata. Nende pakkimist ajendatakse ja stabiliseeritakse molekulidevaheliste sidemete kaudu, mis tekivad kompleksis olevate partnerite vahel. Korrastamata valkude vahelised sidemed on tavaliselt nõrgad ja kiiresti mööduvad. Seega võib pakkimine olla heaks indikaatoriks kompleksi moodustumise tuvastamisel Gli3 repressordomeeni partnereid otsides. Siiski on ka näiteid valgu korrastamatuse säilimisest isegi seotud olekus, kompleksis.

Gli3 täidab oma funktsiooni raku tuumas ise DNA külge otseselt seondumata. Seega ei ole transkriptsioonilise repressiooni mehhanism DNA tasemel, näiteks maskeerides DNA'd või takistades aktivaatorvalkudel DNA'ga seondumast. Selle asemel on Gli3 repressordomeeni toimemehhanism tõenäoliselt valgu tasemel. Väga võimalik, et see toimub otseselt seondudes või kaudselt mediaatorkompleksi aktiivsust mõjutades. Mediaatorkompleks summeerib transkriptsioonifaktoritelt tulevaid signaale ning mõjutab vastavalt RNA polümeraas II aktiivsust saavutamaks vajalikku geeni ekspressioonitaset.

Teatud histidiinijääkide paiknemine Gli3 repressordomeeni järjestuses meenutab tugevalt tsinki siduvaid ligande. Nende jääkide muteerimine tõestas, et repressordomeeni talitlemiseks ei ole vaja tsinki siduda. Sellega seoses tuvastati, et aminohappe jäägid H141 ja H157 paiknevad funktsionaalselt olulistes piirkondades repressordomeeni järjestuses. Repressori funktsioon kadus, kui vastava regiooni struktuurne dünaamika muutus seoses nimetatud jääkide muteerimisega.

Kokkuvõtvalt võib öelda, et siin esitatud tulemused aitavad paremini mõista Gli valkude transkriptsioonilise repressiooni protsesse. Need tulemused on ka aluseks edaspidistele Gli valkude toimimise ja geeniregulatsiooni uuringutele. Tulevikus oleks oluline veel detailsemalt iseloomustada repressori funktsiooni rakusiseses kontekstis ning tuvastada repressordomeeni koostoime partnerid. Oluline järeldus, mida saab teha antud tulemuste põhjal, on vajadus lahutada Ski ja Sufu ajendatud Gli-valkude negatiivne regulatsioon uuritava domeeniga määratud intensiivsemast repressorfunktsioonist.

CURRICULUM VITAE

Personal data

Name:	Robert Tsanev
Data and place of birth:	16.05.1973, Sofia, Bulgaria

Contact information

Address:	Department of Gene Technology, Tallinn University of Technology (TUT), Akadeemia tee 15, 12618
Phone:	+3726204430
e-mail:	robert.tsanev@ttu.ee

Education

2000 – 2014	PhD student, TUT, Department of Gene Technology
1991 – 1997	MSc student, Sofia University, Faculty of Biology, Department of Biotechnology

Employment

2012 – 2014	lecturer, TUT, Department of Gene Technology, Chair of Molecular Diagnostics
2010 – 2013	researcher, TUT, Department of Gene Technology
2009 – 2013	researcher, Competence Center for Cancer Research
2005 – 2009	research assistant, TUT, Department of Gene Technology
2002 – 2005	researcher, National Institute of Chemical Physics and Biophysics

Supervised theses

Lea Pallon, 2010, BSc, “Method development for production of recombinant peptides motilin and galanin”.

Valentina Božok, 2012, BSc, “Constructing and testing of expression vectors for galanin peptide production”.

Viktor Bolkhin, 2012, BSc, “Testing the impact of TPD signal sequence to the function of Gli3 repressor domain”.

Margot Lidemann, 2014, BSc, “Condition optimisation for studying the activity of Gli3 repressor domain in luciferase reporter assay”.

Publications

1. Tsanev, R., K. Vanatalu, J. Jarvet, R. Tanner, K. Laur, P. Tiigimagi, B. B. Kragelund, T. Osterlund and P. Kogerman (2013). The transcriptional repressor domain of Gli3 is intrinsically disordered. PLoS One 8(10): e76972.

2. Tsanev, R., P. Tiigimagi, P. Michelson, M. Metsis, T. Osterlund and P. Kogerman (2009). Identification of the gene transcription repressor domain of Gli3. FEBS Lett 583(1): 224-8.

3. Laht, S., K. Meerits, H. Altroff, H. Faust, R. Tsaney, P. Kogerman, L. Jarvekulg, V. Paalme, A. Valkna and S. Timmusk (2008). Generation and characterization of a single-chain Fv antibody against Gli3, a hedgehog signaling pathway transcription factor. Hybridoma (Larchmt) 27(3): 167-74.

4. Vanatalu, K., Tsanev, R. (2008). Automatic fed-batch cultivation for economic in vivo isotopic labelling. Journal of Biotechnology 136/S: S512–S513.

ELULOOKIRJELDUS

Isikuandmed

Nimi:	Robert Tsanev
Sünniaeg ja -koht:	16.05.1973, Sofia, Bulgaaria

Kontaktandmed

Aadress:	Geenitehnoloogia instituut, Tallinna Tehnikaülikool, Akadeemia tee15, 12618 Tallinn
Telefon:	+3726204430
e-post:	robert.tsanev@ttu.ee

Hariduskäik

2000 – 2014	doktorant, TTÜ, geenitehnoloogia instituut
1991 – 1997	MSc, Sofia Ülikool, bioloogia instituut, biotehnoloogia õppetool

Teenistuskäik

2012 – 2014	lektor, TTÜ, geenitehnoloogia instituut, molekulaardiagnostika õppetool
2010 – 2013	teadur, TTÜ, geenitehnoloogia instituut
2009 – 2013	teadur, Vähiuuringute Tehnoloogia Arenduskeskus AS
2005 – 2009	insener, TTÜ, geenitehnoloogia instituut
2002 – 2005	teadur, Keemilise ja Bioloogilise Füüsika Instituut, Tallinn

Juhendatud lõputööd

Lea Pallon, 2010, bakalaureusekraad, “Rekombinantse galaniini ja motiliini tootmise metoodika väljatöötamine”.

Valentina Božok, 2012, bakalaureusekraad, “Ekspressiooni vektorite konstrueerimine ja testimine galaniini peptiidi tootmiseks”.

Viktor Bolkhin, 2012, bakalaureusekraad, “PTD signaaljärjestuse mõju testimine Gli3 repressordomeeni funktsioonile”.

Margot Lidemann, 2014, bakalaureusekraad, “Katsetingimuste optimeerimine Gli3 repressordomeeni aktiivsuse uurimiseks lutsiferaasi reportertesti meetodil”.

Publikatsioonid

1. Tsanev, R., K. Vanatalu, J. Jarvet, R. Tanner, K. Laur, P. Tiigimagi, B. B. Kragelund, T. Osterlund and P. Kogerman (2013). The transcriptional repressor domain of Gli3 is intrinsically disordered. PLoS One 8(10): e76972.

2. Tsanev, R., P. Tiigimagi, P. Michelson, M. Metsis, T. Osterlund and P. Kogerman (2009). Identification of the gene transcription repressor domain of Gli3. FEBS Lett 583(1): 224-8.

3. Laht, S., K. Meerits, H. Altroff, H. Faust, R. Tsaney, P. Kogerman, L. Jarvekulg, V. Paalme, A. Valkna and S. Timmusk (2008). Generation and characterization of a single-chain Fv antibody against Gli3, a hedgehog signaling pathway transcription factor. Hybridoma (Larchmt) 27(3): 167-74.

4. Vanatalu, K., Tsanev, R. (2008). Automatic fed-batch cultivation for economic in vivo isotopic labelling. Journal of Biotechnology 136/S: S512 - S513.

DISSERTATIONS DEFENDED AT TALLINN UNIVERSITY OF TECHNOLOGY ON *NATURAL AND EXACT SCIENCES*

1. **Olav Kongas**. Nonlinear Dynamics in Modeling Cardiac Arrhytmias. 1998.
2. **Kalju Vanatalu**. Optimization of Processes of Microbial Biosynthesis of Isotopically Labeled Biomolecules and Their Complexes. 1999.
3. **Ahto Buldas**. An Algebraic Approach to the Structure of Graphs. 1999.
4. **Monika Drews**. A Metabolic Study of Insect Cells in Batch and Continuous Culture: Application of Chemostat and Turbidostat to the Production of Recombinant Proteins. 1999.
5. **Eola Valdre**. Endothelial-Specific Regulation of Vessel Formation: Role of Receptor Tyrosine Kinases. 2000.
6. **Kalju Lott**. Doping and Defect Thermodynamic Equilibrium in ZnS. 2000.
7. **Reet Koljak**. Novel Fatty Acid Dioxygenases from the Corals *Plexaura homomalla* and *Gersemia fruticosa.* 2001.
8. **Anne Paju**. Asymmetric oxidation of Prochiral and Racemic Ketones by Using Sharpless Catalyst. 2001.
9. **Marko Vendelin**. Cardiac Mechanoenergetics *in silico*. 2001.
10. **Pearu Peterson**. Multi-Soliton Interactions and the Inverse Problem of Wave Crest. 2001.
11. **Anne Menert**. Microcalorimetry of Anaerobic Digestion. 2001.
12. **Toomas Tiivel**. The Role of the Mitochondrial Outer Membrane in *in vivo* Regulation of Respiration in Normal Heart and Skeletal Muscle Cell. 2002.
13. **Olle Hints**. Ordovician Scolecodonts of Estonia and Neighbouring Areas: Taxonomy, Distribution, Palaeoecology, and Application. 2002.
14. **Jaak Nõlvak**. Chitinozoan Biostratigrapy in the Ordovician of Baltoscandia. 2002.
15. **Liivi Kluge**. On Algebraic Structure of Pre-Operad. 2002.
16. **Jaanus Lass**. Biosignal Interpretation: Study of Cardiac Arrhytmias and Electromagnetic Field Effects on Human Nervous System. 2002.
17. **Janek Peterson**. Synthesis, Structural Characterization and Modification of PAMAM Dendrimers. 2002.
18. **Merike Vaher**. Room Temperature Ionic Liquids as Background Electrolyte Additives in Capillary Electrophoresis. 2002.
19. **Valdek Mikli**. Electron Microscopy and Image Analysis Study of Powdered Hardmetal Materials and Optoelectronic Thin Films. 2003.
20. **Mart Viljus**. The Microstructure and Properties of Fine-Grained Cermets. 2003.
21. **Signe Kask**. Identification and Characterization of Dairy-Related *Lactobacillus.* 2003
22. **Tiiu-Mai Laht**. Influence of Microstructure of the Curd on Enzymatic and Microbiological Processes in Swiss-Type Cheese. 2003.

23. **Anne Kuusksalu**. 2–5A Synthetase in the Marine Sponge *Geodia cydonium*. 2003.
24. **Sergei Bereznev**. Solar Cells Based on Polycristalline Copper-Indium Chalcogenides and Conductive Polymers. 2003.
25. **Kadri Kriis**. Asymmetric Synthesis of C_2-Symmetric Bimorpholines and Their Application as Chiral Ligands in the Transfer Hydrogenation of Aromatic Ketones. 2004.
26. **Jekaterina Reut**. Polypyrrole Coatings on Conducting and Insulating Substracts. 2004.
27. **Sven Nõmm**. Realization and Identification of Discrete-Time Nonlinear Systems. 2004.
28. **Olga Kijatkina**. Deposition of Copper Indium Disulphide Films by Chemical Spray Pyrolysis. 2004.
29. **Gert Tamberg**. On Sampling Operators Defined by Rogosinski, Hann and Blackman Windows. 2004.
30. **Monika Übner**. Interaction of Humic Substances with Metal Cations. 2004.
31. **Kaarel Adamberg**. Growth Characteristics of Non-Starter Lactic Acid Bacteria from Cheese. 2004.
32. **Imre Vallikivi**. Lipase-Catalysed Reactions of Prostaglandins. 2004.
33. **Merike Peld**. Substituted Apatites as Sorbents for Heavy Metals. 2005.
34. **Vitali Syritski**. Study of Synthesis and Redox Switching of Polypyrrole and Poly(3,4-ethylenedioxythiophene) by Using *in-situ* Techniques. 2004.
35. **Lee Põllumaa**. Evaluation of Ecotoxicological Effects Related to Oil Shale Industry. 2004.
36. **Riina Aav**. Synthesis of 9,11-Secosterols Intermediates. 2005.
37. **Andres Braunbrück**. Wave Interaction in Weakly Inhomogeneous Materials. 2005.
38. **Robert Kitt**. Generalised Scale-Invariance in Financial Time Series. 2005.
39. **Juss Pavelson**. Mesoscale Physical Processes and the Related Impact on the Summer Nutrient Fields and Phytoplankton Blooms in the Western Gulf of Finland. 2005.
40. **Olari Ilison**. Solitons and Solitary Waves in Media with Higher Order Dispersive and Nonlinear Effects. 2005.
41. **Maksim Säkki**. Intermittency and Long-Range Structurization of Heart Rate. 2005.
42. **Enli Kiipli**. Modelling Seawater Chemistry of the East Baltic Basin in the Late Ordovician–Early Silurian. 2005.
43. **Igor Golovtsov**. Modification of Conductive Properties and Processability of Polyparaphenylene, Polypyrrole and polyaniline. 2005.
44. **Katrin Laos**. Interaction Between Furcellaran and the Globular Proteins (Bovine Serum Albumin β-Lactoglobulin). 2005.

45. **Arvo Mere**. Structural and Electrical Properties of Spray Deposited Copper Indium Disulphide Films for Solar Cells. 2006.
46. **Sille Ehala**. Development and Application of Various On- and Off-Line Analytical Methods for the Analysis of Bioactive Compounds. 2006.
47. **Maria Kulp**. Capillary Electrophoretic Monitoring of Biochemical Reaction Kinetics. 2006.
48. **Anu Aaspõllu.** Proteinases from *Vipera lebetina* Snake Venom Affecting Hemostasis. 2006.
49. **Lyudmila Chekulayeva**. Photosensitized Inactivation of Tumor Cells by Porphyrins and Chlorins. 2006.
50. **Merle Uudsemaa**. Quantum-Chemical Modeling of Solvated First Row Transition Metal Ions. 2006.
51. **Tagli Pitsi**. Nutrition Situation of Pre-School Children in Estonia from 1995 to 2004. 2006.
52. **Angela Ivask**. Luminescent Recombinant Sensor Bacteria for the Analysis of Bioavailable Heavy Metals. 2006.
53. **Tiina Lõugas**. Study on Physico-Chemical Properties and Some Bioactive Compounds of Sea Buckthorn (*Hippophae rhamnoides* L.). 2006.
54. **Kaja Kasemets**. Effect of Changing Environmental Conditions on the Fermentative Growth of *Saccharomyces cerevisae* S288C: Auxo-accelerostat Study. 2006.
55. **Ildar Nisamedtinov**. Application of ^{13}C and Fluorescence Labeling in Metabolic Studies of *Saccharomyces* spp. 2006.
56. **Alar Leibak**. On Additive Generalisation of Voronoï's Theory of Perfect Forms over Algebraic Number Fields. 2006.
57. **Andri Jagomägi**. Photoluminescence of Chalcopyrite Tellurides. 2006.
58. **Tõnu Martma**. Application of Carbon Isotopes to the Study of the Ordovician and Silurian of the Baltic. 2006.
59. **Marit Kauk**. Chemical Composition of $CuInSe_2$ Monograin Powders for Solar Cell Application. 2006.
60. **Julia Kois**. Electrochemical Deposition of $CuInSe_2$ Thin Films for Photovoltaic Applications. 2006.
61. **Ilona Oja Açik**. Sol-Gel Deposition of Titanium Dioxide Films. 2007.
62. **Tiia Anmann**. Integrated and Organized Cellular Bioenergetic Systems in Heart and Brain. 2007.
63. **Katrin Trummal**. Purification, Characterization and Specificity Studies of Metalloproteinases from *Vipera lebetina* Snake Venom. 2007.
64. **Gennadi Lessin**. Biochemical Definition of Coastal Zone Using Numerical Modeling and Measurement Data. 2007.
65. **Enno Pais**. Inverse problems to determine non-homogeneous degenerate memory kernels in heat flow. 2007.

66. **Maria Borissova**. Capillary Electrophoresis on Alkylimidazolium Salts. 2007.
67. **Karin Valmsen**. Prostaglandin Synthesis in the Coral *Plexaura homomalla*: Control of Prostaglandin Stereochemistry at Carbon 15 by Cyclooxygenases. 2007.
68. **Kristjan Piirimäe**. Long-Term Changes of Nutrient Fluxes in the Drainage Basin of the Gulf of Finland – Application of the PolFlow Model. 2007.
69. **Tatjana Dedova**. Chemical Spray Pyrolysis Deposition of Zinc Sulfide Thin Films and Zinc Oxide Nanostructured Layers. 2007.
70. **Katrin Tomson**. Production of Labelled Recombinant Proteins in Fed-Batch Systems in *Escherichia coli*. 2007.
71. **Cecilia Sarmiento**. Suppressors of RNA Silencing in Plants. 2008.
72. **Vilja Mardla**. Inhibition of Platelet Aggregation with Combination of Antiplatelet Agents. 2008.
73. **Maie Bachmann**. Effect of Modulated Microwave Radiation on Human Resting Electroencephalographic Signal. 2008.
74. **Dan Hüvonen**. Terahertz Spectroscopy of Low-Dimensional Spin Systems. 2008.
75. **Ly Villo**. Stereoselective Chemoenzymatic Synthesis of Deoxy Sugar Esters Involving *Candida antarctica* Lipase B. 2008.
76. **Johan Anton**. Technology of Integrated Photoelasticity for Residual Stress Measurement in Glass Articles of Axisymmetric Shape. 2008.
77. **Olga Volobujeva**. SEM Study of Selenization of Different Thin Metallic Films. 2008.
78. **Artur Jõgi**. Synthesis of 4'-Substituted 2,3'-dideoxynucleoside Analogues. 2008.
79. **Mario Kadastik**. Doubly Charged Higgs Boson Decays and Implications on Neutrino Physics. 2008.
80. **Fernando Pérez-Caballero**. Carbon Aerogels from 5-Methylresorcinol-Formaldehyde Gels. 2008.
81. **Sirje Vaask**. The Comparability, Reproducibility and Validity of Estonian Food Consumption Surveys. 2008.
82. **Anna Menaker**. Electrosynthesized Conducting Polymers, Polypyrrole and Poly(3,4-ethylenedioxythiophene), for Molecular Imprinting. 2009.
83. **Lauri Ilison**. Solitons and Solitary Waves in Hierarchical Korteweg-de Vries Type Systems. 2009.
84. **Kaia Ernits**. Study of In_2S_3 and ZnS Thin Films Deposited by Ultrasonic Spray Pyrolysis and Chemical Deposition. 2009.
85. **Veljo Sinivee**. Portable Spectrometer for Ionizing Radiation "Gammamapper". 2009.
86. **Jüri Virkepu**. On Lagrange Formalism for Lie Theory and Operadic Harmonic Oscillator in Low Dimensions. 2009.

87. **Marko Piirsoo**. Deciphering Molecular Basis of Schwann Cell Development. 2009.
88. **Kati Helmja**. Determination of Phenolic Compounds and Their Antioxidative Capability in Plant Extracts. 2010.
89. **Merike Sõmera**. Sobemoviruses: Genomic Organization, Potential for Recombination and Necessity of P1 in Systemic Infection. 2010.
90. **Kristjan Laes**. Preparation and Impedance Spectroscopy of Hybrid Structures Based on $CuIn_3Se_5$ Photoabsorber. 2010.
91. **Kristin Lippur**. Asymmetric Synthesis of 2,2'-Bimorpholine and its 5,5'-Substituted Derivatives. 2010.
92. **Merike Luman**. Dialysis Dose and Nutrition Assessment by an Optical Method. 2010.
93. **Mihhail Berezovski**. Numerical Simulation of Wave Propagation in Heterogeneous and Microstructured Materials. 2010.
94. **Tamara Aid-Pavlidis**. Structure and Regulation of BDNF Gene. 2010.
95. **Olga Bragina**. The Role of Sonic Hedgehog Pathway in Neuro- and Tumorigenesis. 2010.
96. **Merle Randrüüt**. Wave Propagation in Microstructured Solids: Solitary and Periodic Waves. 2010.
97. **Marju Laars**. Asymmetric Organocatalytic Michael and Aldol Reactions Mediated by Cyclic Amines. 2010.
98. **Maarja Grossberg**. Optical Properties of Multinary Semiconductor Compounds for Photovoltaic Applications. 2010.
99. **Alla Maloverjan**. Vertebrate Homologues of Drosophila Fused Kinase and Their Role in Sonic Hedgehog Signalling Pathway. 2010.
100. **Priit Pruunsild**. Neuronal Activity-Dependent Transcription Factors and Regulation of Human *BDNF* Gene. 2010.
101. **Tatjana Knjazeva**. New Approaches in Capillary Electrophoresis for Separation and Study of Proteins. 2011.
102. **Atanas Katerski**. Chemical Composition of Sprayed Copper Indium Disulfide Films for Nanostructured Solar Cells. 2011.
103. **Kristi Timmo.** Formation of Properties of $CuInSe_2$ and $Cu_2ZnSn(S,Se)_4$ Monograin Powders Synthesized in Molten KI. 2011.
104. **Kert Tamm**. Wave Propagation and Interaction in Mindlin-Type Microstructured Solids: Numerical Simulation. 2011.
105. **Adrian Popp**. Ordovician Proetid Trilobites in Baltoscandia and Germany. 2011.
106. **Ove Pärn**. Sea Ice Deformation Events in the Gulf of Finland and This Impact on Shipping. 2011.
107. **Germo Väli**. Numerical Experiments on Matter Transport in the Baltic Sea. 2011.

108. **Andrus Seiman**. Point-of-Care Analyser Based on Capillary Electrophoresis. 2011.
109. **Olga Katargina**. Tick-Borne Pathogens Circulating in Estonia (Tick-Borne Encephalitis Virus, *Anaplasma phagocytophilum*, *Babesia* Species): Their Prevalence and Genetic Characterization. 2011.
110. **Ingrid Sumeri**. The Study of Probiotic Bacteria in Human Gastrointestinal Tract Simulator. 2011.
111. **Kairit Zovo**. Functional Characterization of Cellular Copper Proteome. 2011.
112. **Natalja Makarytsheva**. Analysis of Organic Species in Sediments and Soil by High Performance Separation Methods. 2011.
113. **Monika Mortimer**. Evaluation of the Biological Effects of Engineered Nanoparticles on Unicellular Pro- and Eukaryotic Organisms. 2011.
114. **Kersti Tepp**. Molecular System Bioenergetics of Cardiac Cells: Quantitative Analysis of Structure-Function Relationship. 2011.
115. **Anna-Liisa Peikolainen**. Organic Aerogels Based on 5-Methylresorcinol. 2011.
116. **Leeli Amon**. Palaeoecological Reconstruction of Late-Glacial Vegetation Dynamics in Eastern Baltic Area: A View Based on Plant Macrofossil Analysis. 2011.
117. **Tanel Peets**. Dispersion Analysis of Wave Motion in Microstructured Solids. 2011.
118. **Liina Kaupmees**. Selenization of Molybdenum as Contact Material in Solar Cells. 2011.
119. **Allan Olspert**. Properties of VPg and Coat Protein of Sobemoviruses. 2011.
120. **Kadri Koppel**. Food Category Appraisal Using Sensory Methods. 2011.
121. **Jelena Gorbatšova**. Development of Methods for CE Analysis of Plant Phenolics and Vitamins. 2011.
122. **Karin Viipsi**. Impact of EDTA and Humic Substances on the Removal of Cd and Zn from Aqueous Solutions by Apatite. 2012.
123. **David Schryer**. Metabolic Flux Analysis of Compartmentalized Systems Using Dynamic Isotopologue Modeling. 2012.
124. **Ardo Illaste**. Analysis of Molecular Movements in Cardiac Myocytes. 2012.
125. **Indrek Reile**. 3-Alkylcyclopentane-1,2-Diones in Asymmetric Oxidation and Alkylation Reactions. 2012.
126. **Tatjana Tamberg**. Some Classes of Finite 2-Groups and Their Endomorphism Semigroups. 2012.
127. **Taavi Liblik**. Variability of Thermohaline Structure in the Gulf of Finland in Summer. 2012.

128. **Priidik Lagemaa**. Operational Forecasting in Estonian Marine Waters. 2012.
129. **Andrei Errapart**. Photoelastic Tomography in Linear and Non-linear Approximation. 2012.
130. **Külliki Krabbi**. Biochemical Diagnosis of Classical Galactosemia and Mucopolysaccharidoses in Estonia. 2012.
131. **Kristel Kaseleht**. Identification of Aroma Compounds in Food using SPME-GC/MS and GC-Olfactometry. 2012.
132. **Kristel Kodar**. Immunoglobulin G Glycosylation Profiling in Patients with Gastric Cancer. 2012.
133. **Kai Rosin**. Solar Radiation and Wind as Agents of the Formation of the Radiation Regime in Water Bodies. 2012.
134. **Ann Tiiman**. Interactions of Alzheimer's Amyloid-Beta Peptides with Zn(II) and Cu(II) Ions. 2012.
135. **Olga Gavrilova**. Application and Elaboration of Accounting Approaches for Sustainable Development. 2012.
136. **Olesja Bondarenko**. Development of Bacterial Biosensors and Human Stem Cell-Based *In Vitro* Assays for the Toxicological Profiling of Synthetic Nanoparticles. 2012.
137. **Katri Muska**. Study of Composition and Thermal Treatments of Quaternary Compounds for Monograin Layer Solar Cells. 2012.
138. **Ranno Nahku**. Validation of Critical Factors for the Quantitative Characterization of Bacterial Physiology in Accelerostat Cultures. 2012.
139. **Petri-Jaan Lahtvee**. Quantitative Omics-level Analysis of Growth Rate Dependent Energy Metabolism in *Lactococcus lactis*. 2012.
140. **Kerti Orumets**. Molecular Mechanisms Controlling Intracellular Glutathione Levels in Baker's Yeast *Saccharomyces cerevisiae* and its Random Mutagenized Gluthatione Over-Accumulating Isolate. 2012.
141. **Loreida Timberg**. Spice-Cured Sprats Ripening, Sensory Parameters Development, and Quality Indicators. 2012.
142. **Anna Mihhalevski**. Rye Sourdough Fermentation and Bread Stability. 2012.
143. **Liisa Arike**. Quantitative Proteomics of *Escherichia coli*: From Relative to Absolute Scale. 2012.
144. **Kairi Otto**. Deposition of In_2S_3 Thin Films by Chemical Spray Pyrolysis. 2012.
145. **Mari Sepp**. Functions of the Basic Helix-Loop-Helix Transcription Factor TCF4 in Health and Disease. 2012.
146. **Anna Suhhova**. Detection of the Effect of Weak Stressors on Human Resting Electroencephalographic Signal. 2012.
147. **Aram Kazarjan**. Development and Production of Extruded Food and Feed Products Containing Probiotic Microorganisms. 2012.

148. **Rivo Uiboupin**. Application of Remote Sensing Methods for the Investigation of Spatio-Temporal Variability of Sea Surface Temperature and Chlorophyll Fields in the Gulf of Finland. 2013.
149. **Tiina Kriščiunaite**. A Study of Milk Coagulability. 2013.
150. **Tuuli Levandi**. Comparative Study of Cereal Varieties by Analytical Separation Methods and Chemometrics. 2013.
151. **Natalja Kabanova**. Development of a Microcalorimetric Method for the Study of Fermentation Processes. 2013.
152. **Himani Khanduri**. Magnetic Properties of Functional Oxides. 2013.
153. **Julia Smirnova**. Investigation of Properties and Reaction Mechanisms of Redox-Active Proteins by ESI MS. 2013.
154. **Mervi Sepp**. Estimation of Diffusion Restrictions in Cardiomyocytes Using Kinetic Measurements. 2013.
155. **Kersti Jääger**. Differentiation and Heterogeneity of Mesenchymal Stem Cells. 2013.
156. **Victor Alari**. Multi-Scale Wind Wave Modeling in the Baltic Sea. 2013.
157. **Taavi Päll**. Studies of CD44 Hyaluronan Binding Domain as Novel Angiogenesis Inhibitor. 2013.
158. **Allan Niidu**. Synthesis of Cyclopentane and Tetrahydrofuran Derivatives. 2013.
159. **Julia Geller**. Detection and Genetic Characterization of *Borrelia* Species Circulating in Tick Population in Estonia. 2013.
160. **Irina Stulova**. The Effects of Milk Composition and Treatment on the Growth of Lactic Acid Bacteria. 2013.
161. **Jana Holmar**. Optical Method for Uric Acid Removal Assessment During Dialysis. 2013.
162. **Kerti Ausmees**. Synthesis of Heterobicyclo[3.2.0]heptane Derivatives *via* Multicomponent Cascade Reaction. 2013.
163. **Minna Varikmaa**. Structural and Functional Studies of Mitochondrial Respiration Regulation in Muscle Cells. 2013.
164. **Indrek Koppel**. Transcriptional Mechanisms of BDNF Gene Regulation. 2014.
165. **Kristjan Pilt**. Optical Pulse Wave Signal Analysis for Determination of Early Arterial Ageing in Diabetic Patients. 2014.
166. **Andres Anier**. Estimation of the Complexity of the Electroencephalogram for Brain Monitoring in Intensive Care. 2014.
167. **Toivo Kallaste**. Pyroclastic Sanidine in the Lower Palaeozoic Bentonites – A Tool for Regional Geological Correlations. 2014.
168. **Erki Kärber**. Properties of ZnO-nanorod/In_2S_3/$CuInS_2$ Solar Cell and the Constituent Layers Deposited by Chemical Spray Method. 2014.
169. **Julia Lehner**. Formation of Cu_2ZnSnS_4 and $Cu_2ZnSnSe_4$ by Chalcogenisation of Electrochemically Deposited Precursor Layers. 2014.