

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE
ENGINEERING AND TECHNOLOGY

**ANALYSIS OF FUNCTIONAL DOMAINS OF NFIB ALTERNATIVE
SPLICING ISOFORMS**



M.Sc. THESIS

Şeyma VARINCA

Department of Molecular Biology-Genetics and Biotechnology

Molecular Biology-Genetics and Biotechnology Programme

AUGUST 2019

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE
ENGINEERING AND TECHNOLOGY

**ANALYSIS OF FUNCTIONAL DOMAINS OF NFIB ALTERNATIVE
SPLICING ISOFORMS**



M.Sc. THESIS

**Şeyma VARINCA
(521171113)**

Department of Molecular Biology-Genetics and Biotechnology

Molecular Biology-Genetics and Biotechnology Programme

Thesis Advisor: Assoc. Prof. Dr. Aslı KUMBASAR

AUGUST 2019

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**NFİB ALTERNATİF KIRPILMA İZOFORMLARININ İŞLEVSEL
BÖLGELERİNİN ANALİZİ**

YÜKSEK LİSANS TEZİ

**Şeyma VARINCA
(521171113)**

Moleküler Biyoloji-Genetik ve Biyoteknoloji Anabilim Dalı

Moleküler Biyoloji-Genetik ve Biyoteknoloji Programı

Tez Danışmanı: Doç. Dr. Aslı KUMBASAR

AĞUSTOS 2019

Şeyma Varınca, a M.Sc. student of ITU Graduate School of Science Engineering and Technology student ID 521171113, successfully defended the thesis/dissertation entitled “ANALYSIS OF FUNCTIONAL DOMAINS OF NFIB ALTERNATIVE SPLICING ISOFORMS”, which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Assoc. Prof. Aslı KUMBASAR**
İstanbul Technical University

Jury Members : **Assoc. Prof. Ceren ÇIRACI**
İstanbul Technical University

Assist. Prof. Necla BİRGÜL
Bogazici University

Date of Submission : 15 August 2019
Date of Defense : 22 August 2019





To my beautiful nieces,



FOREWORD

I would like to express my deepest gratitudes to my supervisor Assoc. Prof. Aslı KUMBASAR for her guidance and instructiveness during both my graduate and undergraduate education.

I'll be appreciated to Sinem SARITAŞ and Betül KUTAY untill last breath of my life. Whenever and wherever we met, suddenly my question marks poured off and they both answer my all questions in detail, always with a smile even at their hard times.. I am also grateful to Dicle MALAYMAR for all her explanations during both my 2 years masters period and previously my undergrad education. I would like to thank Istanbul Technical University Research Fund (ITU BAP Project ID: 42095) for financial support.

I am also appreciated to my jury members, Ceren ÇIRACI and Necla BIRGÜL, for hers understanding during my thesis preparation and for their supports to enhance my thesis.

Of course the biggest part of my thankfulness goes to my perfect family. There is no word to express their companionship..

August 2019

Şeyma VARINCA



TABLE OF CONTENTS

ABBREVIATIONS.....	XIII
SYMBOLS	XV
LIST OF TABLES.....	XVII
LIST OF FIGURES.....	XIX
SUMMARY	XXI
ÖZET	XXIII
1. INTRODUCTION	1
1.1 Nuclear Factor I Family of Transcription Factors	1
1.1.1 Structural and functional domains	1
1.1.2 Four NFI family members: NFIA, NFIB, NFIC, and NFIX	4
1.1.2.1 NFIB.....	5
1.2 Aim of the Study.....	15
2. MATERIALS AND METHODS	17
2.1 Materials.....	17
2.1.1 Enzymes and buffers	17
2.1.2 Commercial kits	17
2.1.3 Devices	18
2.1.4 Cell lines used	18
2.1.5 Growth and freezing media	19
2.1.6 Other requisites	19
2.2 Methods.....	20
2.2.1 pCH_mNfib4 plasmid (pCH_hNFIB1)	20
2.2.2 pCH_MCS control plasmid	22
2.2.3 pCH_GFP plasmid	22
2.2.4 qRT-PCR Experiments.....	23
2.2.5 Dual luciferase reporter assays	23
2.2.6 Electrophoretic mobility shift assays	25
3. RESULTS	29
3.1 Preparation of Expression Constructs	29
3.1.1 pCH_mNfib4 construct	29
3.1.2 pCH_MCS construct	32
3.1.3 pCH_GFP construct	33
3.2 Characterization of NFIB-S in SH-SY5Y, U87MG, T98G and HeLa cell lines	35
3.3 Transcriptional regulation by NFIB isoforms	38
3.3.1 <i>GFAP</i> promoter activation by NFIB isoforms.....	38
3.3.2 <i>p21</i> promoter repression by NFIB isoforms	41
3.4 Binding of NFIB isoforms to <i>GFAP</i> and <i>p21</i> NFI sites.....	44
3.4.1 Binding of NFIB isoforms to NFI sites on <i>GFAP</i>	44
3.4.2 Binding of NFIB isoforms to NFI sites on <i>p21</i>	45
4. DISCUSSION.....	49
REFERENCES	55

APPENDIX.....	63
CIRRICULUM VITAE.....	69



ABBREVIATIONS

Ad	: Adenovirus
cDNA	: Complementary DNA
DBD	: DNA-Binding Domain
DDI	: Domain-domain interactions
HCS	: Hematopoietic Stem Cells
NFI	: Human Nuclear Factor I
NFIB	: Human Nuclear Factor I B
NFIB1	: Human Nuclear Factor I B, transcript variant 1
NFIB3	: Human Nuclear Factor I B, transcript variant 3
NFIB4	: Human Nuclear Factor I B, transcript variant 4
NFIB-S	: Human Nuclear Factor I B- Short isoform
NFIB-L	: Nuclear Factor I B- Long isoform
NFIB-FL	: Nuclear Factor I B- Full Length isoform
Nfi	: Mouse Nuclear Factor I
Nfib	: Mouse Nuclear Factor I B
Nfib3	: Mouse Nuclear Factor I B, transcript variant 3
Nfib4	: Mouse Nuclear Factor I B, transcript variant 4
MK	: Megakaryocyte
MPPs	: Multipotent Progenitor cells
ORF	: Open Reading Frame
PPI	: Protein-protein interactions
PTM	: Post Transcriptional Modifications
TA	: Transcriptional Activation
TFs	: Transcription Factors



SYMBOLS

bp	: Base pair
°C	: Celcius Degree
g	: Grams
kb	: Kilo base pair
L	: Liter
M	: Molar
n	: Mol
nt	: Nucleotide
rpm	: Revolutions per minute
V	: Volt



LIST OF TABLES

	<u>Page</u>
Table 2.1: Enzymes used in this experiment.	17
Table 2.2: Enzyme Buffers used in this experiment.	17
Table 2.3: Commercial kits used in the experiment.	18
Table 2.4: The devices used in the experiment.	18
Table 2.5: Other requisites used in the experiment.	19
Table 2.6: mNfib4 detection primers.	21
Table 2.7: mNfib4 cloning primers.	21
Table 2.8: TOPO Cloning reaction.	21
Table 2.9: UV-Spectrophotometer results for miniprep of Zero Blunt II TOPO_mNfib4 plasmids obtained from the two colonies.	21
Table 2.10: Infusion primers to amplify TOPO_mNfib4.	21
Table 2.11: Cloning enhancer treatment to PCR product.	22
Table 2.12: Infusion reaction components.	22
Table 2.13: pCH_MCS_control oligonucleotide sequences.	22
Table 2.14: pCH_GFP infusion primers.	23
Table 2.15: qRT-PCR primers of all NFIB and NFIB-S transcripts.	23
Table 2.16: Oligonucleotide sequences of GFAP and p21 NFI binding sites used in mobility shift experiments.	25
Table 2.17: Components of DNA-binding reactions in EMSA experiment made with GFAP promoter oligonucleotides.	25
Table 2.18: Components of DNA-binding reactions in EMSA experiment made with p21 promoter oligonucleotides.	26
Table 2.19: Components of DNA-binding reactions in EMSA experiment made with p21 promoter oligonucleotides.	27
Table 3.1: C _t values produced by NFIB and NFIB-S qRT-PCR primers on mRNA extracts of SH-SY5Y, U87MG, T98G, and HeLa cell lines.	38



LIST OF FIGURES

	<u>Page</u>
Figure 1.1: Functional domains of NFI proteins.	2
Figure 1.2: Three conserved cysteines and one histidine residue reside at N-termini of NFI proteins.....	3
Figure 1.3: Ribbon diagram of Smad MH1 domain.....	3
Figure 1.4: Amino acid sequence alignments of NFIX2 and NFIX-X5 isoforms.	8
Figure 1.5: Schematic representation of human NFIB splicing isoforms.	9
Figure 1.6: Schematic representation of mouse NFIB splicing isoforms.....	10
Figure 1.7: Aminoacid sequence alignment of human NFIB alternative splicing isoforms.....	11
Figure 1.8: Aminoacid alignment of human NFIB1 and NFIB3 alternative splicing isoforms.....	12
Figure 1.9: Aminoacid sequence alignment of human NFIB3, human NFIB1, mouse Nfib3 and mouse Nfib4 proteins.....	13
Figure 3.1: mNfib4 3' fragment was amplified from C17.2 cDNA.....	30
Figure 3.2: Amplification of mNfib4 with infusion primers from TOPO_mNfib4 ..	31
Figure 3.3: mNfib4 3' fragment was cloned into pCH expression vector.	32
Figure 3.4: pCH_MCS was generated by ligation.....	33
Figure 3.5: Generation of pCH_GFP.....	34
Figure 3.6: Diagnostic restriction digestion of pCH_GFP and overexpression in HEK293T cells.	34
Figure 3.7: Characterization of NFIB4 expression in SH-SY5Y, U87MG, T98G and HeLa cell lines.....	36
Figure 3.8: Characterization of NFIB4 expression in SH-SY5Y, U87MG, T98G and HeLa cell lines.	37
Figure 3.9: Characterization of NFIB4 expression in SH-SY5Y, U87MG, T98G and HeLa cell lines.....	37
Figure 3.10: <i>GFAP</i> promoter activity in mNfib3, mNfib4 and NFIB-S overexpressing HEK293T cells -I.....	39
Figure 3.11: <i>GFAP</i> promoter activity in mNfib3, mNfib4 and NFIB-S overexpressing HEK293T cells -II.	40
Figure 3.12: <i>GFAP</i> promoter activity in mNfib3, mNfib4 and NFIB-S overexpressing HEK293T cells -III.....	41
Figure 3.13: <i>p21</i> promoter activity in mNfib3, mNfib4, NFIB-S and NFIB-TAD overexpressing HEK293T cells -I.....	42
Figure 3.14: <i>p21</i> promoter activity in mNfib3, mNfib4, NFIB-S and NFIB-TAD overexpressing HEK293T cells -II.	43
Figure 3.15: Gel shift assays showing binding of NFIB isoforms to <i>GFAP</i> NFI sites.	45
Figure 3.16: Gel shift assays showing binding of NFIB isoforms to <i>p21</i> NFI sites-I..	46

Figure 3.17: Gel shift assays showing binding of NFIB isoforms to *p21* NFI sites-II.
..... 47



ANALYSIS OF FUNCTIONAL DOMAINS OF NFIB ALTERNATIVE SPLICING ISOFORMS

SUMMARY

Transcription factors which regulate gene expression in the developing embryo control cellular processes such as proliferation, differentiation and migration. Disruption of the same set of transcription factors is also associated with various types of cancer in adults. Nuclear Factor I (NFI) family of transcription factors, for example, are expressed by stem cells during development and in the adult stem cell niche and drive many developmental processes such as neural stem cell differentiation, lung maturation, and hematopoiesis. In vertebrates there are four NFI genes: *NFIA*, *NFIB*, *NFIC* and *NFIX*; however, each gene produces numerous alternative splicing isoforms that may have different interaction partners and functional characteristics. NFIs carry a highly conserved N-terminal DNA-binding and dimerization domain through which they form homo- or heterodimers and bind to DNA on specific consensus sequences. NFI C-terminal transcription modulation domain is less conserved, and allow NFIs to activate or repress target gene expression. The diverse C-terminal domain may indeed lead to differences in function as demonstrated by NFIX isoforms that vary in the efficiency of target gene activation. Here we set out to investigate functional differences between NFIB isoforms and focused on two mouse *Nfib* alternative splicing isoforms, *Nfib3* and *Nfib4*. While *Nfib4* bears exons 1-10 and 12; *Nfib3* is encoded by exons 1-8 and 12 and 79 residues shorter than *Nfib4*. We also included an unusual human NFIB isoform, NFIB-S. NFIB-S is a truncated protein encoded by exons 5-12; thus lacks the N-terminal DNA-binding and dimerization domain while carrying a long transcription modulation domain. In order to study these proteins, we first generated the mammalian expression constructs that encode these three isoforms. We selected two NFI target genes, *GFAP*, which is induced by NFIs and *p21* which is repressed by NFI proteins and measure promoter driven luciferase activity. While *Nfib3* induced *GFAP* promoter by 1.7-1.8 fold, *Nfib4* activated by 2.5-3 fold in three independent experiments. However, *Nfib4* was consistently expressed at higher levels compared to *Nfib3*, which may account for the difference in *GFAP* activation. *Nfib4* also repressed the *p21* promoter more efficiently, 45-67% compared to 37-44% by *Nfib3*. These differences may be due to differential activation efficacy or expression levels; further experiments are required to determine the underlying mechanisms.

We also tested NFIB-S in parallel experiments. Interestingly, NFIB-S reduced *GFAP* promoter activity compared to pCH mock. When we included GFP in order to control for effects of exogenous protein overexpression, NFIB-S still displayed a modest repressive effect. Neither NFIB-S nor a truncated version of *Nfib3* that lacks the N-terminal domain (NFIB-TAD) repressed the *p21* promoter significantly. Next, we tested binding ability of NFIB-S to *GFAP* and *p21* NFI-sites by electrophoretic mobility shift assays. Once again, neither NFIB-S and NFIB-TAD interacted with the NFI-sites, while full-length NFIB bound to the NFI-motifs. NFIB-S may act by

affecting interactions of full-length NFI proteins with NFI binding partners or NFI sites on genomic DNA. *In vitro* co-incubation of NFIB-S with full-length NFIB prior to DNA binding did not interfere with NFIB binding to NFI-sites on DNA. While the *in vitro* binding and reporter assays may not have been able to detect low affinity interactions or replicate *in vivo* conditions, preliminary data show that NFIB-S does not bind or regulate the selected target genes, *GFAP* or *p21*. Thus, further experiments are required to elucidate mechanisms by which NFIB-S acts to regulate cellular differentiation.



NFIB ALTERNATİF KIRPILMA İZOFORMLARININ İŞLEVSEL BÖLGELERİNİN ANALİZİ

ÖZET

Transkripsiyon faktörleri, hedef genlerinin ekspresyonlarını dolaylı veya direkt olarak aktive etmek veya baskılamak suretiyle düzenleyen proteinlerdir. Bu proteinler tarafından gerçekleştirilen gen regülasyonu, özellikle gelişim sürecinde ve kök hücre farklılaşmasında hayati önem arz etmektedir. Nükleer Faktör I (NFI) transkripsiyon faktörü ailesi, gelişim sürecinde, nöral kök hücre farklılaşmasında, akciğer gelişiminde, megakaryosit özelleşmesinde ve daha birçok önemli hücresel süreçte rol oynarlar. Bu geniş regülasyon alanları, bu transkripsiyon faktörlerinin gelişimin daha erken evrelerinden itibaren ekspresyon kontrolünde rol oynamaya başladıklarını işaret etmektedir. NFI ailesi dört gen (*NFIA*, *NFIB*, *NFIC*, and *NFIX*) kodlanan dört proteinden (*NFIA*, *NFIB*, *NFIC*, and *NFIX*) oluşmaktadır. Bu proteinler özellikle N-ucu bölgelerinde neredeyse tamamen korunmuş olan dimerizasyon ve DNA'ya bağlanma bölgeleri bulundurmaktadır. Homodimer veya heterodimer oluşturmak suretiyle bağlandıkları konsensus sekans olan "TTGGC(N5)GCCAA" dizisine afiniteleri eşittir. C-ucu bölgelerinde ise daha az korunmuş olan ve farklı ekzon kombinasyonları tarafından kodlanan transaktivasyon/baskılama bölgeleri bulunur. Proteine fonksiyonel özelliklerini kazandıran bölge bu C-ucu bölgesidir. Proteine fonksiyonel karakteristiğini kazandıran bölge bu C-ucu bölgesidir. C ucu bölgelerinin en ucunda birçok prolin amino asidinin oluşturduğu bir kısım korunmaktadır ve bu prolince zengin bölgenin histon proteinleriyle etkileşmek suretiyle nükleozom yapısında modifikasyonlara sebep olarak transkripsiyonu regüle ettiği düşünülmektedir. NFI proteinleri hem gelişim sürecinde kök hücre farklılaşmasında hem de yetişkin kök hücrelerinde transkripsiyonel regülasyonda önemli rollere sahiptirler. Merkezi sinir sistemi gelişimi sürecindeki kök hücrelerdeki etkileri, yetişkin beyin kök hücrelerindeki etkilerinden farklıdır. Gelişim evresinde genellikle kök hücre yenilenmesiyle ilişkili genlerin ifadesini bastırırken; hücre özelleşmesi, hücre göçü, akson ve dendrit oluşumu gibi hücre farklılaşması ile ilintili genlerin ifadelerini arttırdıkları saptanmıştır. Yetişkin kök hücrelerinde ise, kök hücre devamlılığı ile ilişkili genlerin ifadesini arttırırken, hücre farklılaşması ile ilişkili genlerin ifadesini baskıladıkları bilinmektedir. Bunlara ek olarak, NFIB kanser ile de ilişkilendirilmektedir. NFIB ifadesi arttırıldığında tümör büyümesinin durdurulduğu gözlemlenmiştir. Ancak, genellikle kaydedilen tümör baskılayıcı rolüne rağmen NFIB'nin onkogenik etkisinin görüldüğü de olmuştur. *NFIB* proteininin insanda dört kanonik izoform mevcuttur* ve bunların ikisi (NFIB1, NFIB2) uzun izoformlar olup 11. ekzon hariç 1-12 ekzonların tümünü bulundururken; bir diğer izoform (NFIB3) daha kısadır ve 9., 10. ve 11. ekzonları barındırmamaktadır. En kısa izoform olan NFIB-S (NFIB4) ise, alternatif bir başlangıç kodonunun kullanılması sonucunda 11. ekzon da dahil olmak üzere 5-12 ekzonları bulundurmaktadır ve 11. ekzonun bulunması sonucu oluşan çerçeve kayması bu izoforma farklı bir N-ucu aminoasit dizisi kazandırmaktadır. Sonuçta oluşan bu NFIB-S izoformunun farklı olan 12.

ekzonu, fare Nfib izoformlarının 12. ekzonlarına benzemekle birlikte, asıl ilgi çekici olan ise NFIB-S izoformunun bu NFI transkripsiyon faktörlerinin tümünde ortak olarak korunmuş olan dimerizasyon ve DNA-bağlanma bölgelerini taşıyan N-ucu bölgesini bulundurmamasıdır. İnsan NFIB izoformlarının taşıdıkları farklı transkripsiyon düzenleme bölgelerinin yol açabileceği farklı fonksiyonların araştırılması, bu proteinlerin gelişim ve nöral hücre farklılaşmasındaki rollerinin anlaşılması ve kanserdeki olası etkilerinin aydınlatılması için önem arz etmektedir. Tüm bunlardan hareketle, bu çalışmada öncelikle ilgili izoformları ifade eden memeli ekspresyon vektörleri hazırlanmış, daha sonra bu proteinlerin HEK293T hücrelerinde ifadesi artırımı suretiyle *in vitro* DNA-bağlanma ve promotor-aktivasyonu/baskılaması deneyleri gerçekleştirilmiştir. Lusiferaz raportör deneylerinde, izoformların aktivasyon kapasitelerinin saptanması için *GFAP* promotoru kullanılırken; izoformların baskılama kapasitelerinin belirlenmesi için *p21* promotoru kullanılmıştır.

Farede bulunan iki Nfib izoformu, Nfib3 ve Nfib4, insandaki NFIB3 ve NFIB4 izoformları ile %99 örtüşmektedir ve bu fare izoformlarının var ise fonksiyonel farklılıklarının, insan NFIB izoformlarının fonksiyonlarını da yansıtacağı düşünülmektedir. Biz bu çalışmada iki fare izoformunun fonksiyonel karşılaştırmak için promotor raportör deneylerinde aktivasyon ve baskılama kapasitelerini belirledik. Ancak, bu deneylerde Nfib4 izoformu, tutarlı olarak Nfib3 izoformundan daha yüksek düzeylerde ifade edilmiş ve her iki izoformun ekspresyon seviyeleri eşitlenememiştir. Bu nedenle, iki izoformun aktivasyon/baskılama kapasitelerinin farklı olup olmaması konusunda kesin bir sonuca ulaşamamıştır.

Diğer yandan, NFIB-S izoformunun olası aktivasyon/baskılama kapasitesinin belirlenmesi amacıyla yapılan promotor raportör deneylerinde doğru kontrol grubunun belirlenmesinin önemi daha da öne çıkmıştır. NFIB-S hem *GFAP* hem de *p21* promotor aktivasyonunda ekzojen protein ifade etmeyen kontrol grubuna göre düşüşe neden olmaktadır. Ancak, aktivasyon seviyelerindeki bu düşüşün, transfeksiyona tabi tutulan hücrelerdeki ekzojen protein üretiminin oluşturduğu strese kaynaklanıyor olabileceğini düşünerek eklediğimiz GFP ifade eden kontrol grubu ile kıyaslandığında, NFIB-S'in promotor aktivitesinde istatistiksel açıdan önemli sayılabilecek bir değişikliğe yol açmadığı düşünülmektedir. Her ne kadar NFIB-S izoformu DNA-bağlanma ve dimerizasyon bölgesini bulundurmasa da, yine de bu izoformun başka bir protein bölgesi aracılığı ile DNA üzerindeki bilinen NFI bağlanma bölgelerine bağlanıp bağlanamayacağı, bu izoformun transkripsiyonel regülasyonda rolü var ise hangi şekilde olabileceğinin aydınlatılması yolunda önemli bir ilk adımdır. Bu amaçla yapılan *in vitro* DNA bağlanma deneyleri, ne NFIB-S'nin ne de NFIB-TAD'ın NFI bağlama motifleri ile etkileşmediği görülmektedir (NFIB-TAD, NFIB'nin N-ucu DNA bağlanma ve dimerizasyon bölgesi çıkartılmak suretiyle oluşturulmuş yapay bir kısaltılmış proteindir). NFIB-S'nin tam boy NFIB ile olası bir etkileşimi sonucunda, NFIB'nin DNA'daki ilgili kısımlara bağlanmasını etkileyebileceği görüşüyle uygulanan ve her iki izoformun birlikte *in vitro* inkübasyonunu müteakip NFIB'nin DNA'ya bağlanmasını inceleyen deneylerde, NFIB-S'nin DNA-bağlanma seviyesine herhangi bir etkisi saptanamamıştır. Bu etki çok düşük miktarda olduğundan veya bu etkileşimlerdeki protein afinitelerinin düşük olmasından dolayı da saptanamamış olabilir. Bu ihtimalin test edilmesi için, birlikte inkübasyon deneylerinde tam boy NFIB miktarına kıyasla daha yüksek oranlarda NFIB-S kullanılabilir. Bir başka yaklaşım ise bu iki proteinin ilgili bir hücre hattında birlikte ifadesi sonucu kromatin immunopresipitasyonu ile elde edilebilecek *in vivo* etkileşimleri ile ilgili bilgilerdir. Bunlara ek olarak, bu proteinlerin işlevlerinin inceleneceği hücre hatlarının seçimi,

evresel faktrlerin transkripsiyonel reglasyonda modler etkisinin yadsınamayacađı geređi dolayısıyla ok mhimdir.

*: Biz bu alıřmaya bařladıđımızda GeneBank verilerine gre 4 kabul edilmiř izoform mevcuttu ancak Mayıs 2019’da yapılan bir gncelleme ile bu sayı 28’e ykseldi. Ancak hala NFIB1 en uzun izoform olma zelliđini korumaktadır. Buna ek olarak, NFIB-S, N-ucu DNA bađlanma blgesini bulundurmayan tek ilgi ekici izoform olmaya devam etmektedir.





1. INTRODUCTION

1.1 Nuclear Factor I Family of Transcription Factors

Stem cell biology has been an attractive field for biologists for decades as all known cells of multicellular organisms arise from a limited set of stem cells. Unlike most differentiated cells, neurons do not divide to proliferate or regenerate once they are committed to their fate, thus, elucidation of stem cell differentiation into neuronal fate is fundamental to the field of neuroscience. During stem cell differentiation or renewal, intrinsic factors, such as transcription factors, play important roles within stem cells as well as in progenitors or later in differentiated cells themselves. Transcription factors bind to specific regulatory sequences to activate or repress the expression of target genes by interacting directly or indirectly with the transcription machinery involving RNA polymerase II (Dynlacht et al., 1991; Tanese et al., 1991; Lin et al., 1991; and Stringer et al., 1990). Transcription factors generally include a DNA-binding domain and a transcriptional activation domain and both of these domains are functional separately (Mitchell and Tjian, 1989; and Ptashne, 1988). Nuclear Factor I (NFI) proteins comprise such a family of transcription factors which act in diverse set of tissues, regulating processes such as neural stem cell differentiation, lung maturation, megakaryocyte differentiation. Moreover, these proteins are found to be misregulated in several types of cancer (Semenova et al., 2016).

1.1.1 Structural and functional domains

Nuclear factor I family proteins were first identified as factors important for adenoviral replication; and it was subsequently found that NFIs regulate the expression of many cellular genes (Gronostajski 2000, Harris et al., 2014). There are four different *NFI* genes in vertebrates: *NFIA*, *NFIB*, *NFIC* and *NFIX*; and NFIs encoded by those genes share a highly conserved N-terminal ~220 amino acid region which is responsible and sufficient for DNA-binding, dimerization, and initiation of adenoviral replication (Mermod et al., 1989; Gounari et al., 1990; Bandyopadhyay and Gronostajski, 1994). Dimerization is essential for DNA binding (Armentero et al., 1994) and they bind to

DNA as homo- or heterodimers on the consensus sequence (TTGGC(N₅)GCCAA) with similar affinity (Goyal et al., 1990, Kruse and Sippel, 1994). While N-terminal domain mediates DNA binding and dimerization (DBD); the C-terminal domain regulates the expression of target genes (Figure 1.1). Also referred to as the ‘transactivation domain,’ this C-terminal domain gives the protein its functional characteristics (Palacios and Peri, 2010). In contrast to the highly conserved N-terminal domain, NFIs their C-terminal domains are more varied (Rupp et al., 1990). Alternative splicing generates many protein isoforms for each gene, leading to further variation at the C-terminal domain (Kruse and Sippel, 1994). Bioinformatic analysis of recent high-throughput sequencing data predicts numerous NFI gene products which generally share conserved characteristics of NFIs and the more prevalent and full-length ones are referred to as canonical NFIs.



Figure 1.1: Functional domains of NFI proteins. Canonical NFI proteins carry the highly conserved N-terminal domain and less conserved C-terminal transactivation and repression domain with its proline-rich terminal region, and the two nuclear localization signals (NLS).

NFI proteins bear four conserved cysteine residues in their N-terminal DNA-binding domain, three of which are required for DNA interactions whereas the remaining one is involved in response to oxidative inactivation (Gronostajski, 2000). Although the structure of NFI proteins has not been determined yet, the N-terminal DBD domain of NFIs carries the MH1 motif that was first identified in another transcription factor, Smad (Sadreyev and Grishin, 2003). The crystal structure of Smad MH1 domain in complex with DNA is available (Shi et al., 1998 and the X-Ray crystal structure data is available on Protein Data Bank: <https://www.rcsb.org/structure/1mhd>, on August 23, 2019). Alignment of the NFI MH1 domain with that of Smads showed that three of the conserved cysteine residues and a histidine constitutes a ‘cys-his box’ which is thought to form a zinc-binding domain (Sadreyev and Grishin, 2003). Indeed, if any one of these three cysteines is mutated, NFI DNA-binding affinity is reduced; while the mutation of the remaining conserved cysteine does not affect the DNA-binding of

NFIs (Bandyopadhyay and Gronojstaski, 1994). Sequence similarity between NFI and Smad DNA binding domains suggests that NFI and Smad MH1 domain evolved from a common ancestor. Previously, MH1 motif was found to be homologous to His-Me endonucleases (Grishin, 2001); thus, NFI DBD domain may have evolved from an ancestral endonuclease which lost its enzymatic activity in time but conserved its DNA-binding ability.

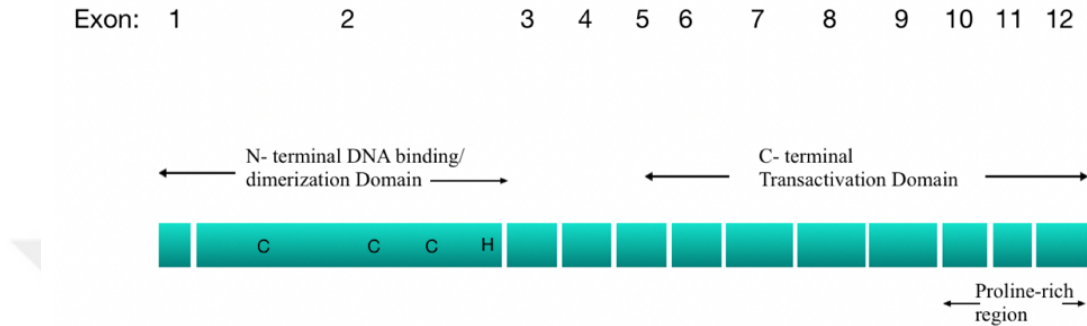


Figure 1.2: Three conserved cysteines and one histidine residue reside at N-termini of NFI proteins. Highly conserved N-terminal domain is responsible for homo or heterodimerization and DNA-binding while less conserved C-terminal domain controls transcriptional activation/repression of target genes. 12 exons encoded by *NFI* genes are represented with green blocks.

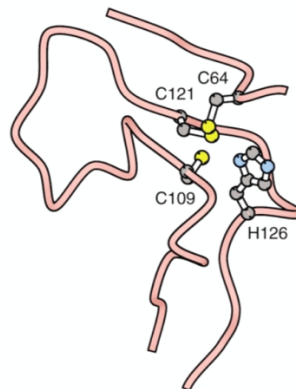


Figure 1.3: Ribbon diagram of Smad MH1 domain. Cys-His box constitutes a Zinc-binding site (Adapted from Sadreyev et al., 2003).

Transcription factors activation domains are categorized according to the amino acid content such as acidic, glutamine-rich, or proline-rich (Attardi and Tjian, 1993; and Mitchell and Tjian, 1989). The proline rich region was first identified in Nuclear Factor I proteins, whose C-terminal transactivation domain comprises about 25% prolines (Mermod et al., 1989). This proline-rich domain may facilitate modification of

nucleosome structure through interacting directly with histone protein, H3 (Gerber et al., 1995).

1.1.2 Four NFI family members: NFIA, NFIB, NFIC, and NFIX

NFIs are expressed in a distinct but overlapping fashion in the developing embryo, and are involved in development of various tissues such as the liver, mammary gland, brain and lung (das Neves et al., 1999)

Absence of each NFI gene gives rise to diverse phenotypes in mice. *Nfia*-null mice display agenesis of the corpus callosum and hydrocephalus (das Neves et al., 1999); while *Nfib*-deficient mice exhibit severe lung hypoplasia and retarded pulmonary development (Gründer et al., 2002) and *Nfix*-deficiency in mice causes brain malformations, including enlargement of the lateral and third brain ventricles and partial agenesis of the corpus callosum, as well as skeletal defects (Driller et al., 2007). On the other hand, *Nfic*-deficient mice show postnatal tooth development defects (Steele-Perkins et al., 2003). Thus, though the NFI DBD domain is highly similar and NFIs bind DNA with similar affinity; each NFI may play distinct roles in development of particular tissues.

Specifically, NFIs regulate stem cells in the embryo and the adult. During central nervous system (CNS) development, NFIs repress genes associated with self-renewal while inducing genes associated with terminal differentiation (Piper et al., 2010; Heng et al., 2014), thus generally promoting cell differentiation. On the other hand, in the adult, NFIs promote stem cell survival and maintenance of melanocyte stem cells (Chang et al., 2013), neural stem cells (NSCs) (Martynoga et al., 2013), and hematopoietic stem cells (Holmfeldt et al., 2013). Interestingly, NFIs were found to occupy enhancer regions associated with NSC quiescence (Martynoga et al., 2013). Loss of function experiments showed that NFIX is required for NSC quiescence *in vitro* and *in vivo*; while NFIA and NFIB are downregulated in quiescent NSCs (Martynoga et al., 2013).

NFIs may regulate gene expression via distinct molecular mechanisms: They could affect recruitment of basal transcriptional machinery (Kim and Roeder, 1994; Xiao et al., 1994) through their C-terminal motifs; they may interact with coactivators (Tanese et al., 1991; Dusserre and Mermoud, 1992) or compete with other activators (Nehls et al., 1991). Alternatively, they can control gene expression by changing the nucleosome

structure through direct interactions with histone proteins (Alevizopoulos et al., 1995) or competing with histone proteins for DNA binding (Ristiniemi and Oikarinen, 1989; Gao et al., 1998; Gronostajski, 2000).

1.1.2.1 NFIB

NFIB is one of the three NFI family members that are expressed in the CNS. NFIB is involved in regulation of astrocyte differentiation from neural stem cells in the hippocampus and the hindbrain (Steele-Perkins et al., 2005; Barry et al., 2008; and Stringer et al., 2016). NFIB also regulates proliferation and epithelial cell differentiation during lung development and differentiation of melanocytes in the hair follicle of stem cell niche (Chang et al., 2013). Indeed, absence of *Nfib* leads to lung hypoplasia and brain development defects: reduced basal pons and hippocampus, corpus callosum agenesis (Hsu et al., 2011; Steele-Perkins et al., 2005). Additionally, *Nfib*^{-/-} mice show a delay in formation of precerebellar nuclei of the hindbrain apparently due to a decrease in neurogenesis (Kumbasar et al., 2009). Interestingly, brain specific conditional *Nfib* knock out mice exhibit increase in cortical size with no other defects in brain structure or connectivity (Schanze et al., 2018). Consistent with these findings, *NFIB* haploinsufficiency is associated with mild intellectual disability and macrocephaly. Moreover, speech delay, muscular hypotonia and attention deficit disorder are also seen in subjects with *NFIB* mutations or deletions including the *NFIB* gene (Schanze et al., 2018). Additionally, SNPs in *NFIB* is also associated with bipolar disorder and schizophrenia (Schanze et al., 2018; Le-Niculescu et al., 2009; Willour et al., 2012; Segurado et al., 2003; Gupta et al., 2015); while some individuals with autism spectrum disorder carry genomic alterations of *NFIB* (Schanze et al., 2018; Prasad et al., 2012; Iossifov et al., 2012; Redin et al., 2012).

NFIB also plays important roles in cancer, in a context dependent manner. Amplification of *NFIB* is observed in breast cancer (Han et al., 2008) and in esophagus squamous cell carcinoma (Yang et al., 2001). Using small cell lung cancer (SCLC) cell lines obtained from both patients and a SCLC mouse model, Dooley et al. (2011) found that NFIB regulates apoptosis, senescence and proliferation and as well as fibroblast growth (Dooley et al., 2011). Higher NFIB levels associated with invasive tumor cell population growth and NFIB was shown to promote metastasis; moreover, in line with data obtained in mice, NFIB overexpression was observed in

approximately all of the tested human metastatic progressive neuroendocrine lung tumors (Semenova et al., 2016). Elsewhere NFIB can act as a tumor suppressor: NFIB is underexpressed in non-small cell lung cancers (Becker-Santos et al., 2016) cutaneous squamous cell carcinoma (Zhou et al., 2014) and osteosarcoma (Mirabello et al., 2015). In classical and mesenchymal glioblastoma (GBM) cell lines obtained from patients, NFIB overexpression resulted in inhibition of cell proliferation and tumorigenesis; furthermore, the inhibition is in direct proportion to NFIB expression levels. However, NFIB exhibited an oncogenic effect in neural GBM (Stringer et al., 2016).

Thus, involvement of NFIB in cancer, in addition to its roles in normal development, makes identification and elucidation of the molecular functions and working mechanisms of NFIB a priority.

Alternative splicing isoforms of NFIB

During mRNA maturation, selective retention or removal of particular exons/introns can generate diverse sets of mature RNA encoded by the same gene (Chow et al., 1977; Early et al., 1980). This process is called alternative splicing. While alternative splicing allows one gene to have many different mature transcripts, resulting proteins may not differ from each other functionally. Nevertheless, inclusion of different domains may change the protein's binding properties, stability, intracellular localization and enzymatic activity (Resch et al., 2004; Xing et al., 2003; Stamm et al., 2005). In humans, exons differ from 50 to 300 nt in length, averaging at 137 nt (Berget et al., 1995). The introns, which are typically about 3400 nt in length (Deutsch and Long, 1999), are removed by the macromolecular spliceosome complex during pre-mRNA synthesis in the nucleus (Kelemen et al., 2013). Five small nuclear RNAs (snRNAs) in addition to at least 170 proteins constitute the spliceosome (Behzadnia et al., 2007). The spliceosome recognizes the sequence elements, 5' and 3' splice sites, and the branch point, on the exons to assemble on the nascent pre-mRNA. U1 snRNP occupies the 5' splice site following the binding of splicing factor 1 (SF1) to the branch point which resulted in U2AF binding to the 3' splice site. This complex is called E (early complex) and the substitution of SF1 by U2 snRNP starts formation of prespliceosomal complex A. The complex exchange and recruit other factors in order to constitute the spliceosomal complex B which yields a transesterification reaction includes removal of introns and joining of the retaining exons (Kelemen et al., 2013).

As eluded to earlier, alternative splicing generates many isoforms for all four NFI proteins (Kruse and Sippel, 1994, Altmann et al., 1994) and the existence of many isoforms has made pinning down the roles of each NFI in transcriptional regulation of development more difficult (das Neves, 1999). Isoforms transcribed from the same gene vary especially in their C-terminal domains and this variation can lead to functional differences. NFIC isoforms with varying C-terminal domain sequences were found to have different transcriptional activity when overexpressed in *Drosophila* cells and yeast (Altmann et al., 1994, Mermod et al., 1989). The difference in their activity arose from the C-terminal domains; thus, the transactivation domain can supply the protein with its functional characteristics. Isoform-specific functions of NFIs has not yet been elucidated, limited data exists on isoform specific expression: expression levels of the five bovine NFIX isoforms vary depending on the type of tissue; with lower expression levels and the isoform diversity in the adult. Specifically, in the adult brain, NFIX3 was not detected, while the other four isoforms were expressed at higher levels. In the embryo, NFIX4 isoform expression was similar to total NFIX expression, suggesting that the main embryonic CNS isoform is NFIX4 (Zhou et al., 2014).

Singh et al. (2011) studied two human NFIX isoforms, a long vs a short isoform (NFIX2; referred to as 'NFIX1' and NFIX-X5; referred to as 'NFIX3' by Singh et al.) generated by different C-terminal exon combinations. The long isoform, NFIX-X5, includes exon 9 which gives rise to a frameshift and a later stop codon at exon 11. The shorter isoform, NFIX2 encodes an earlier stop codon, at exon 10. Thus, NFIX-X5 has a distinct transcriptional activation (TA) domain. NFIX-X5 is highly expressed in astrocytes. However, both total NFIX and NFIX-X5 is downregulated during astrocyte differentiation. Interestingly, NFIX-X5 activates the 1745 bp promoter region of *GFAP*, an NFI target, much more potently than NFIX2 in luciferase reporter assays. Additionally, NFIX-X5 could also activate an artificial construct bearing 10 NFI-binding sequences, at expression levels 1/10 of NFIX2. On the other hand, both NFIX-X5 and NFIX2 were shown to repress the *p21* promoter (2400bp upstream region) with the same efficiency. Hence, NFIX-X5 may promote activation via its distinct TA domain though the repression occurs through other regions found in both isoforms. A fusion protein of the TA domain of NFIX-X5 fused to the Gal4 DNA-binding domain could induce the Gal4 reporter in both astrocytes and HEK293T cells while that of

NFIX2 could not. Moreover, knock-down of NFIX-X5 lead to a 40-50% decrease in expression of important astrocyte markers like *GFAP* and *SPARCL1*, even though NFIX-X5 constituted only 10% of total NFIX. Thus, the NFIX-X5 isoform appears to play an important role in astrocyte differentiation (Singh et al., 2011).

NFIX-2	MYSPLYCLTQDEFHPFIEALLPHVRAFSYTWFNLQARKRKYFKKHEKRMSKDEERAVKDEL	60
NFIX-X5	-----MDEFHPFIEALLPHVRAFSYTWFNLQARKRKYFKKHEKRMSKDEERAVKDEL	52

NFIX-2	LGEKPEIKQKWASRLAKLRDIRPEFREDFVLITITGKKPPCCVLSNPDQKGKIRRIDCL	120
NFIX-X5	LGEKPEIKQKWASRLAKLRDIRPEFREDFVLITITGKKPPCCVLSNPDQKGKIRRIDCL	112

NFIX-2	RQADKVRWRLDLMVILFKGIPLESTDGERLYKSPQCSNPGLCVQPHHIGVTIKELDLILA	180
NFIX-X5	RQADKVRWRLDLMVILFKGIPLESTDGERLYKSPQCSNPGLCVQPHHIGVTIKELDLILA	172

NFIX-2	YFVHTPESGQSDSSNQGDADIKPLPNGHLSFQDCFTSGVWNVTTELVRVSQTPVATASG	240
NFIX-X5	YFVHTPESGQSDSSNQGDADIKPLPNGHLSFQDCFTSGVWNVTTELVRVSQTPVATASG	232

NFIX-2	PNFSLADLESPPSYNNINQVTLGRRSITSPSTSTTKRPKSIDDSEMEPVDDVFYPGTGR	300
NFIX-X5	PNFSLADLESPPSYNNINQVTLGRRSITSPSTSTTKRPKSIDDSEMEPVDDVFYPGTGR	292

NFIX-2	SPAAGSSQSSGWPNDVDAGPASLKKSGKLDFCALSQSSPRMAFTHHPLPVLAVRPG	360
NFIX-X5	SPAAGSSQSSGWPNDVDAGPASLKKSGKLDFCALSQSSPRMAFTHHPLPVLAVRPG	352

NFIX-2	SPRATASALHFPSTSIQQSSPYFTHPTIRYHHHGGDSLKEFVQFVCSGSGQATGQHS	420
NFIX-X5	SPRATASALHFPSTSIQQSSPYFTHPTIRYHHHGGDSLKEFVQFVCSGSGQATGQPN	412

NFIX-2	QRQAPPLPTGLSASDPGTAT-----F-----	441
NFIX-X5	GSGQGKVPGSFLLPPPPVVARPVPLPMPDSKSTSTAPDGAALTPSPSFATTGASSANRF	472
	: * . : * . :	
NFIX-2	-----	441
NFIX-X5	VSIGPRDGNFLNIPQSSQSWFL	494

Figure 1.4: Amino acid sequence alignments of NFIX2 and NFIX-X5 isoforms.

NFIB also produces multiple alternative splice isoforms, with alternating C-terminal transactivation domains. When we started to this study, there were four canonical human *NFIB* isoforms as recognized by GeneBank: NFIB1, NFIB2, NFIB3 and NFIB4 (received from <https://www.ncbi.nlm.nih.gov/gene/4781>; on August 20, 2019). As described below, we were interested in studying the short isoform NFIB4 which lacks a DNA binding and dimerization domain and full-length isoforms that varied in the C-terminal domain, NFIB3 and NFIB1.

However, as of May 2019, the number of canonical human *NFIB* isoforms has been updated to 28. Nevertheless, NFIB1 is still the longest isoform within 28 isoforms of human *NFIB* while mechanisms of action of NFIB4, the unique isoform that lacks the conserved N-terminal domain, still remains to be elucidated.

Human NFIB1, NFIB3, and NFIB4

NFIB1 is the longest human NFIB isoform including the exons 1-10 and 12. NFIB3 is the intermediate-length NFIB isoform which lacks exons 9, 10 and 11. In contrast, NFIB4 (also called NFIB-S as it is truncated) is the shortest human NFIB isoform bearing the exons 5-12 but the use of an alternative start codon in exon 5 and the inclusion of exon 11 shifts the reading frame, generating a different C-terminal amino acid sequence also found in some mouse full-length *Nfib* isoforms. NFIB-S (NFIB4) was first detected in an RNA-seq study (Chen et al., 2014) that focused on differentially expressed alternative splicing isoforms in human blood progenitors during lineage commitment. NFIB-S does not bear the conserved NFI-dimerization domain, and indeed it could not dimerize with NFIC, the full-length NFI previously hypothesized to act in postnatal megakaryopoiesis. However, both NFIC and NFIB-S was required for megakaryocyte differentiation. Because NFIs bind to DNA as homo- or heterodimers, it is not clear whether NFIB-S acts by regulating transcription directly or via other mechanisms.

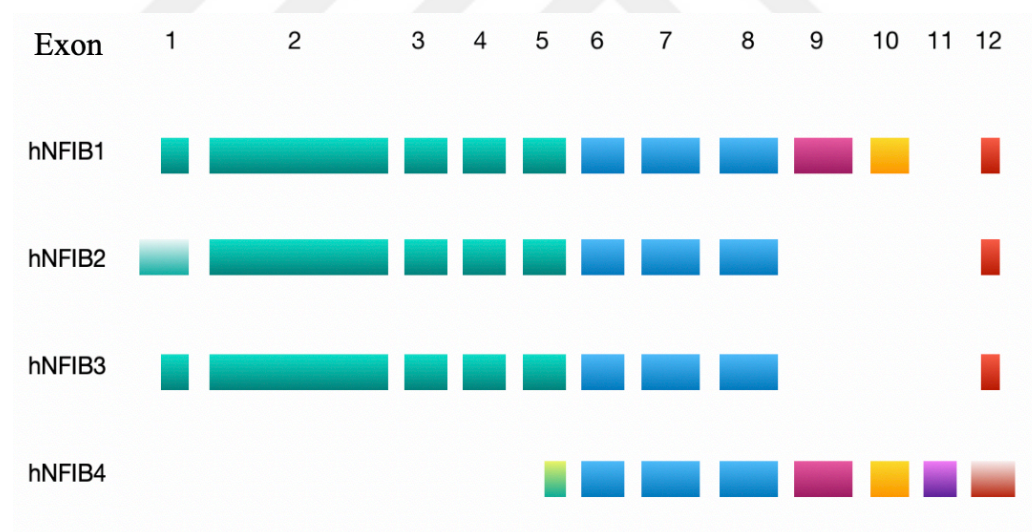


Figure 1.5: Schematic representation of human NFIB splicing isoforms. (from NCBI as of April 2019) Each exon is represented relative to its total amino acid length.

Mouse isoforms *Nfib3* and *Nfib4* are ~99% identical to human NFIB3 and NFIB1 respectively. The transactivation (TA) domain of the ortholog in mouse (mNfib4) differs by 2 amino acids from the long NFIB in human (hNFIB1). Those amino acid substitutions are serine (Ser⁴¹⁸) to glycine (Gly⁴¹⁸) and threonine (Thr⁴⁴⁹) to serine (Ser⁴⁴⁹) from mouse to human. Thus, as the mouse isoforms *Nfib3* (mNfib3) and

mouse Nfib4 (mNfib4) are virtually identical to their human orthologs, hNFIB3 and hNFIB1, we focused on the mouse isoforms as these constructs were readily available.

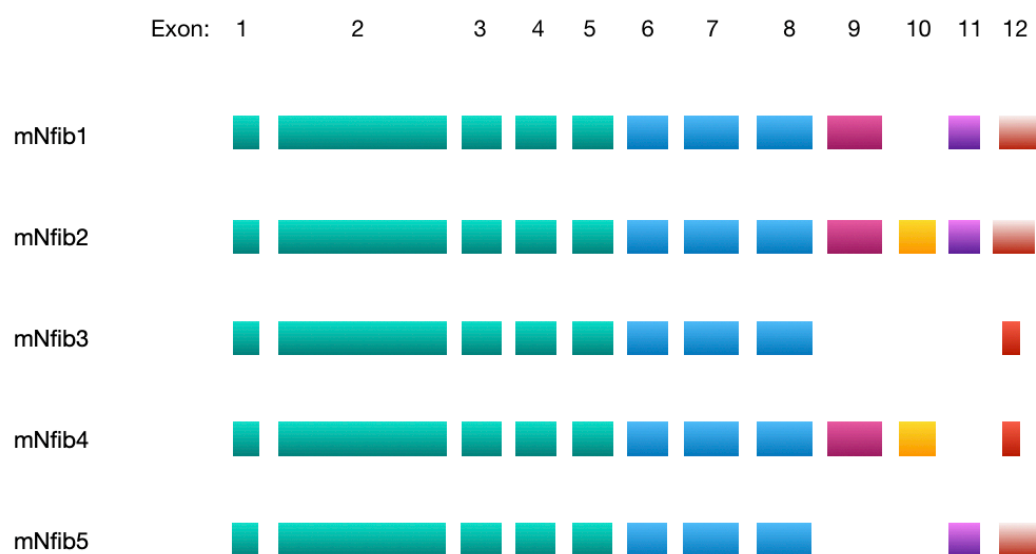


Figure 1.6: Schematic representation of mouse NFIB splicing isoforms. Each exon is represented relative to its total amino acid length.

hNFIB4	-----	0
hNFIB2	MERIPVSVDFVWVCCAVLKCNPGIPKRMSTLCFGFSDEFHPFIEALLPHVRAIAYTWFNL	60
hNFIB1	-----MMYSPICL-TQDEFHPFIEALLPHVRAIAYTWFNL	34
hNFIB3	-----MMYSPICL-TQDEFHPFIEALLPHVRAIAYTWFNL	34
hNFIB4	-----	0
hNFIB2	QARKRKYFKKHEKRMSKDEERAVKDELLSEKPEIKQKWASRLAKLRKDIRQEYREDFVL	120
hNFIB1	QARKRKYFKKHEKRMSKDEERAVKDELLSEKPEIKQKWASRLAKLRKDIRQEYREDFVL	94
hNFIB3	QARKRKYFKKHEKRMSKDEERAVKDELLSEKPEIKQKWASRLAKLRKDIRQEYREDFVL	94
hNFIB4	-----	0
hNFIB2	TVTGKKHPCCVLSNPDQKGKIRRIDCLRQADKVWRDLVMVILFKGIPLESTDGERLMKS	180
hNFIB1	TVTGKKHPCCVLSNPDQKGKIRRIDCLRQADKVWRDLVMVILFKGIPLESTDGERLMKS	154
hNFIB3	TVTGKKHPCCVLSNPDQKGKIRRIDCLRQADKVWRDLVMVILFKGIPLESTDGERLMKS	154
hNFIB4	-----	0
hNFIB2	PHCTNPALCVQPHHITVSVKELDLFLAYYVQEQDSGQSGSPSHNDPAKNPPGYLEDSFVK	240
hNFIB1	PHCTNPALCVQPHHITVSVKELDLFLAYYVQEQDSGQSGSPSHNDPAKNPPGYLEDSFVK	214
hNFIB3	PHCTNPALCVQPHHITVSVKELDLFLAYYVQEQDSGQSGSPSHNDPAKNPPGYLEDSFVK	214
hNFIB4	-----MNSGVNLQRLSLSPPSSKRPKT	22
hNFIB2	SGVFNSELVLRVSRTPITQGTGVNFPIGEIPSQPYHDMNSGVNLQRLSLSPPSSKRPKT	300
hNFIB1	SGVFNSELVLRVSRTPITQGTGVNFPIGEIPSQPYHDMNSGVNLQRLSLSPPSSKRPKT	274
hNFIB3	SGVFNSELVLRVSRTPITQGTGVNFPIGEIPSQPYHDMNSGVNLQRLSLSPPSSKRPKT *****	274
hNFIB4	ISIDENMEPSPTGDFYFSPSSPAAGSRTWHERDQDMSSPTTMKKPEKPLFSSASPQDSSP	82
hNFIB2	ISIDENMEPSPTGDFYFSPSSPAAGSRTWHERDQDMSSPTTMKKPEKPLFSSASPQDSSP	360
hNFIB1	ISIDENMEPSPTGDFYFSPSSPAAGSRTWHERDQDMSSPTTMKKPEKPLFSSASPQDSSP	334
hNFIB3	ISIDENMEPSPTGDFYFSPSSPAAGSRTWHERDQDMSSPTTMKKPEKPLFSSASPQDSSP *****	334
hNFIB4	RLSTFPQHHPGIPGVAHSVISTRTPPPPSPLPFPTQAILPPAPSSYFSHPTIRYPHNL	142
hNFIB2	RLSTFPQHHPGIPGVAHSVISTRTPPPPSPLPFPTQAILPPAPSSYFSHPTIRYPHNL	420
hNFIB1	RLSTFPQHHPGIPGVAHSVISTRTPPPPSPLPFPTQAILPPAPSSYFSHPTIRYPHNL	394
hNFIB3	RLSTFPQHHPGIPGVAHSVISTRTPPPPSPLPFPTQAILPPAPSSYFSHPTIRYPHNL *****	394
hNFIB4	PQDTLKNYVPSYDPSSPQTSQPNGSGQVVGKVPGHFTPVLAPSPHPSAVRPVTLSTMDTK	202
hNFIB2	PQDTLKNYVPSYDPSSPQTSQSWYLG-----	446
hNFIB1	PQDTLKNYVPSYDPSSPQTSQPNGSGQVVGKVPGHFTPVLAPSPHPSAVRPVTLSTMDTK	454
hNFIB3	PQDTLKNYVPSYDPSSPQTSQSWYLG----- ***** *	420
hNFIB4	PITTSTEAYTASGTSQANRYVGLSPRDPSFLHQQLRICDWTMNQNGRHLYPSTSEDTLG	262
hNFIB2	-----	446
hNFIB1	PITTSTEAYTASGTSQANRYVGLSPRDPSFLHQQSWYLG-----	494
hNFIB3	-----	420
hNFIB4	ITWQSPGTWASLVFPQVSNRTPILPANVQNYGLNIIGEPFLQAETSN	309
hNFIB2	-----	446
hNFIB1	-----	494
hNFIB3	-----	420

Figure 1.7: Aminoacid sequence alignment of human NFIB alternative splicing isoforms

hNFIB1	MMYSPICLTQDEFHFFIEALLPHVRAIAYTWFNLQARKRKYFKKHEKRMSKDEERAVKDE	60
hNFIB3	MMYSPICLTQDEFHFFIEALLPHVRAIAYTWFNLQARKRKYFKKHEKRMSKDEERAVKDE *****	60
hNFIB1	LLSEKPEIKQKWASRLAKLRKDIRQEYREDFVLTVTGKKHPCCVLSNPDQKGKIRRIDC	120
hNFIB3	LLSEKPEIKQKWASRLAKLRKDIRQEYREDFVLTVTGKKHPCCVLSNPDQKGKIRRIDC *****	120
hNFIB1	LRQADKVWRDLVMVILFKGIPLESTDGERLMKSPHCTNPALCVQPHHITVSVKELDLFL	180
hNFIB3	LRQADKVWRDLVMVILFKGIPLESTDGERLMKSPHCTNPALCVQPHHITVSVKELDLFL *****	180
hNFIB1	AYYVQEQDSGQSGSPSHNDPAKNPPGYLEDSFVKSGVFNSELVRSRTPITQGTGVNFP	240
hNFIB3	AYYVQEQDSGQSGSPSHNDPAKNPPGYLEDSFVKSGVFNSELVRSRTPITQGTGVNFP *****	240
hNFIB1	IGEIPSQPYHDMNSGVNLQRSLSPPSSKRPKTISIDENMEPSPTGDFYPSPPSPAAGS	300
hNFIB3	IGEIPSQPYHDMNSGVNLQRSLSPPSSKRPKTISIDENMEPSPTGDFYPSPPSPAAGS *****	300
hNFIB1	RTWHERDQDMSSPTTMKKPEKPLFSSASPQDSSPRLSTFPQHHPGIPGVAHSVISTRTP	360
hNFIB3	RTWHERDQDMSSPTTMKKPEKPLFSSASPQDSSPRLSTFPQHHPGIPGVAHSVISTRTP *****	360
hNFIB1	PPPSPLPFPTQAILPPAPSSYFSHPTIRYPPHLNPQDTLKNYVPSYDPSSPQTSQPNGSG	420
hNFIB3	PPPSPLPFPTQAILPPAPSSYFSHPTIRYPPHLNPQDTLKNYVPSYDPSSPQTSQSWYLG ***** *	420
hNFIB1	QVVGKVPGHFTPV LAPSPHPSAVRPVTL SMTDTKPI TTSTEAYTASGTSQANRYVGLSPR	480
hNFIB3	-----	420
hNFIB1	DPSFLHQQSWYLG	494
hNFIB3	-----	420

Figure 1.8: Aminoacid alignment of human NFIB1 and NFIB3 alternative splicing isoforms.

mNfib4	MMYSPICLTQDEFHPFIEALLPHVRAIAYTWFNLQARKRKYFKKHEKRMSKDEERAVKDE	60
hNFIB1	MMYSPICLTQDEFHPFIEALLPHVRAIAYTWFNLQARKRKYFKKHEKRMSKDEERAVKDE	60
hNFIB3	MMYSPICLTQDEFHPFIEALLPHVRAIAYTWFNLQARKRKYFKKHEKRMSKDEERAVKDE	60
mNfib3	MMYSPICLTQDEFHPFIEALLPHVRAIAYTWFNLQARKRKYFKKHEKRMSKDEERAVKDE	60

mNfib4	LLSEKPEIKQKWASRLLAKLRKDIRQEYREDFVLTVTGKKHPCCVLSNPDQKGKIRRIDC	120
hNFIB1	LLSEKPEIKQKWASRLLAKLRKDIRQEYREDFVLTVTGKKHPCCVLSNPDQKGKIRRIDC	120
hNFIB3	LLSEKPEIKQKWASRLLAKLRKDIRQEYREDFVLTVTGKKHPCCVLSNPDQKGKIRRIDC	120
mNfib3	LLSEKPEIKQKWASRLLAKLRKDIRQEYREDFVLTVTGKKHPCCVLSNPDQKGKIRRIDC	120

mNfib4	LRQADKVWRDLVLMVILFKGIPLSTDGERLMKSPHCTNPALCVQPHHITVSVKELDLFL	180
hNFIB1	LRQADKVWRDLVLMVILFKGIPLSTDGERLMKSPHCTNPALCVQPHHITVSVKELDLFL	180
hNFIB3	LRQADKVWRDLVLMVILFKGIPLSTDGERLMKSPHCTNPALCVQPHHITVSVKELDLFL	180
mNfib3	LRQADKVWRDLVLMVILFKGIPLSTDGERLMKSPHCTNPALCVQPHHITVSVKELDLFL	180

mNfib4	AYYVQEQDSGQSGSPSHSDPAKNPPGYLEDSFVKSGVFNVSSELVRSRTPITQGTGVNFP	240
hNFIB1	AYYVQEQDSGQSGSPSHSDPAKNPPGYLEDSFVKSGVFNVSSELVRSRTPITQGTGVNFP	240
hNFIB3	AYYVQEQDSGQSGSPSHSDPAKNPPGYLEDSFVKSGVFNVSSELVRSRTPITQGTGVNFP	240
mNfib3	AYYVQEQDSGQSGSPSHSDPAKNPPGYLEDSFVKSGVFNVSSELVRSRTPITQGTGVNFP	240

mNfib4	IGEIPSQPYHDMNSGVNLQRSLSPPSSKRPKTISIDENMEPSPTGDFYPSPNSPAAGS	300
hNFIB1	IGEIPSQPYHDMNSGVNLQRSLSPPSSKRPKTISIDENMEPSPTGDFYPSPNSPAAGS	300
hNFIB3	IGEIPSQPYHDMNSGVNLQRSLSPPSSKRPKTISIDENMEPSPTGDFYPSPNSPAAGS	300
mNfib3	IGEIPSQPYHDMNSGVNLQRSLSPPSSKRPKTISIDENMEPSPTGDFYPSPNSPAAGS	300

mNfib4	RTWHERDQDMSSPTTMKKPEKPLFSSTSPQDSSPRLSTFPQHHPGIPGVAHSVISTRTP	360
hNFIB1	RTWHERDQDMSSPTTMKKPEKPLFSSTSPQDSSPRLSTFPQHHPGIPGVAHSVISTRTP	360
hNFIB3	RTWHERDQDMSSPTTMKKPEKPLFSSTSPQDSSPRLSTFPQHHPGIPGVAHSVISTRTP	360
mNfib3	RTWHERDQDMSSPTTMKKPEKPLFSSTSPQDSSPRLSTFPQHHPGIPGVAHSVISTRTP	360

mNfib4	PPPSPLPFPTQAILPPAPSSYFSHPTIRYPHLPNPQDTLKNYVPSYDPSSPQTSQPNSSG	420
hNFIB1	PPPSPLPFPTQAILPPAPSSYFSHPTIRYPHLPNPQDTLKNYVPSYDPSSPQTSQPNSSG	420
hNFIB3	PPPSPLPFPTQAILPPAPSSYFSHPTIRYPHLPNPQDTLKNYVPSYDPSSPQTSQSWYLG	420
mNfib3	PPPSPLPFPTQAILPPAPSSYFSHPTIRYPHLPNPQDTLKNYVPSYDPSSPQTSQSWYLG	420

mNfib4	QVVGKVPGHFTPVLAAPSPHSAVRPVTLSMTDTKPIITSTEAYTASGTSQANRYVGLSPR	480
hNFIB1	QVVGKVPGHFTPVLAAPSPHSAVRPVTLSMTDTKPIITSTEAYTASGTSQANRYVGLSPR	480
hNFIB3	-----	420
mNfib3	-----	420
mNfib4	DPSFLHQQSWYLG	494
hNFIB1	DPSFLHQQSWYLG	494
hNFIB3	-----	420
mNfib3	-----	420

Figure 1.9: Aminoacid sequence alignment of human NFIB3, human NFIB1, mouse Nfib3 and mouse Nfib4 proteins. hNFIB3/mNfib3 is the shorter isoform in both organisms compared to the longer isoforms hNFIB1/mNfib4 bearing exons 9 and 10. Aminoacid variations on the TA of the long isoform is indicated in yellow (mouse) to blue (human).

Regulation of *GFAP* by NFIs

Glial fibrillary acidic protein (GFAP) is an intermediate filament found in astrocytes and its expression is regulated by NFIs (das Neves et al., 1999). *NFIA* or *NFIB* disruption results in reduced *GFAP* expression and causes glial-cell related developmental anomalies in mice (das Neves et al., 1999; Shu et al., 2003; Steele-

Perkins et al., 2005). NFIs activate the *GFAP* promoter in differentiating progenitor cells towards astrocytes. In the developing rat brain, NFI expression coincides with the onset of astrocyte differentiation and upregulation of NFIs continues through astrocytogenesis. Elements in the -106/-35 promoter region is sufficient for *GFAP* induction, and NFI binds to regulatory elements within that region to modulate *GFAP* expression (Cebolla and Vallejo, 2006). *GFAP* promoter region includes three NFI-sites (-120/-106; -1585/-1571; and -1633/-1619) that are occupied by NFIs *in vitro* (Gopalan et al., 2006; Brun et al., 2009). When one of the known *NFI* sites on *GFAP* promoter is mutated, the expression of GFAP is reduced (Miura et al., 1990; Krohn et al., 1999). The GFAP-Luc plasmid (pGfa2-luc) used in this study includes all three NFI sites.

Regulation of *p21* by NFIs

p21 is a cyclin-dependent kinase inhibitor and it promotes cell cycle arrest in response to many stimuli by binding to and inhibiting the kinase activity of cyclin dependent kinases and causes growth arrest at certain stages of the cell cycle (Abbas and Dutta, 2009). *p21* regulates many processes like cell proliferation, differentiation, or apoptosis (Gartel et al., 1996; Wu et al., 1996; Ostrovsky and Bengal, 2003). *p21* expression is found to be regulated by NFIs which repress *p21* transcription by interacting with the proximal promoter region *in vivo* (Ouellet et al., 2006). Overexpression of NFIA decrease *p21* mRNA levels; *p21* promoter driven luciferase activity is reduced by NFIA overexpression and increased by knockdown of endogenous NFIA (Lee et al., 2014). The mutagenesis of the *p21* promoter determined NFI sites between -161/-149 upstream region (Ouellet et al., 2006). Additionally, a binding site at -1507/-1493 is also included in the *p21*-Luc (pGL2-*p21*) plasmid used in this study.

1.2 Aim of the Study

NFI proteins regulate stem cells in the developing embryo and in the adult stem cell niche. NFIB, a member of this family, acts in development of the CNS and is implicated in brain cancer, where it may act as a tumor suppressor. Not much is known about regulation of NFIB function, or the functional consequences of the numerous NFIB alternative splicing isoforms. Interestingly, during alternative splicing of NFIX, another NFI family member, inclusion of exons 9 and 10 produces a transcriptionally more potent activator. Here we hypothesized that differences in the C-terminal transactivation domains of NFIB isoforms may also lead to differences in their function and focused on three isoforms. First, we compared mNfib3 with mNfib4, the mouse isoform that includes exons 9 and 10 and thus has a longer C-terminal transactivation domain. We also included NFIB-S, an unusual NFIB isoform that has only been detected in humans and lacks the N-terminal DNA binding and dimerization domain. We tested the functional capabilities of these isoforms with luciferase reporter assays, measuring promoter activation of two NFI target genes, *GFAP* and *p21*. GFAP is an astrocyte terminal differentiation marker and well characterized NFI target induced by NFIs. p21 is a cell cycle-inhibitor and tumor suppressor repressed by NFIs in proliferating cells, including glioblastoma cells, where NFI proteins function. We also carried out *in vitro* DNA-binding experiments to measure interactions of NFIB isoforms with NFI sites on these promoters.



2. MATERIALS AND METHODS

2.1 Materials

2.1.1 Enzymes and buffers

The enzymes used for diagnostic or restriction digestion are shown in the Table 2.1 and enzymes buffers are shown in the Table 2.2.

Table 2.1: Enzymes used in this experiment.

Enzymes	Trademarks
BglII	NEB
EcoRI	NEB
KpnI	NEB
SfiI	Thermofisher Scientific
BsrGI	NEB
NotI	NEB
HindIII	NEB

Table 2.2: Enzyme Buffers used in this experiment.

Enzyme Buffers	Trademarks
NEB 2.1 Buffer	NEB
G Buffer	Thermofisher Scientific
NEB 3.1 Buffer	NEB
NEB 1.1 Buffer	NEB
NEB CutSmart Buffer	NEB

2.1.2 Commercial kits

The commercial kits used in the experiment are indicated in the Table 2.3.

Table 2.3: Commercial kits used in the experiment.

Kit	Commercial Name and Trademark
Transformation Kit	In-Fusion HD Plus kit -Stellar protocol-
Miniprep Kit	NucleoSpin Plasmid Kit-Macherey-Nagel
Infusion kit	In-Fusion Advantage PCR Cloning Kit
TOPO-PCR Cloning kit	Zero Blunt Topo PCR Cloning Kit- invitrogen
Gel and PCR Clean-up Kit	NucleoSpin Gel and PCR Clean-up Kit- Macherey-Nagel
Dual Luciferase Assay Kit	Promega
FuGENE6 Transfection Reagent	Promega
LightShift Chemiluminescent EMSA Kit	ThermoScientific
Pierce Biotin 3'End DNA Labeling Kit	ThermoScientific
Pierce BCA Protein Assay Kit	ThermoScientific

2.1.3 Devices

The devices used in the experiment are indicated in the Table 2.4.

Table 2.4: The devices used in the experiment.

Devices	Trademarks
Autoclave	TOMY SX-700E
Bunsen's Burner	
Centrifuge	VWR MicroStar 17R
Electrophoresis	BIORAD
Heat block	Biometra Ts1 Thermo Shaker
Laminar Hood	BioBeb Faster BH-EN 2003
Micropipette	GILSON, Thermo Scientific
Microwave	ALTUS
PCR device	BIORAD T100 Thermal Cycler
Shaker	Sartorius CERTOMAT SII
UV-light device	Biometra UV Star 312nm

2.1.4 Cell lines used

Human Embryonic Kidney 293T (HEK293T) cell line is used in transient transfection experiments. In order to detect the expression levels of NFIB-S (hNFIB4) in

neuroblastoma and glioblastoma cells, SHSY5Y cell line as a neuroblastoma cell line and U87MG and T98G cell lines as glioma cell lines were used. Moreover, HeLa cells were used to detect the NFIB-S expression due to the fact that its expression in that cell line was detected before (Hein et al., 2015). cDNA from C17.2 immortalised mouse neural progenitor cell line capable of differentiation *in vitro* was also used for pCH_mNfib4 mammalian expression plasmid construct preparation.

2.1.5 Growth and freezing media

HEK293T complete medium

DMEM (high glucose without L-glutamine) with 2 mM L-Glutamine, 10% (v/v) Fetal Bovine Serum (FBS), 500 µg/ml Penicilin/Streptomycin liquid is used as growth medium for HEK293T cells.

HEK293T freezing medium

Freezing medium for HEK293T and SH-SY5Y cell lines composed of 50% FBS, 10% DMSO and 40% DMEM. It was sterilized by filtering through 0.22 µm filter before aliquoting to falcons for later use.

2.1.6 Other requisites

Antibiotics, DNA Markers, LB and Agarose Gel materials used are shown in the Table 2.5.

Table 2.5: Other requisites used in the experiment.

Requisite	Trademark
Ampicillin	ROTH
Agar	LabM
Agarose	SeaKem
Kanamycin	ROTH
DNA Marker	Thermofisher Scientific- Mass Ruler DNA Ladder mix- ready to use
Microtubes	Axygen- MT-060-C
NaCl	EMSURE
PCR tubes	Axygen
Agarose Gel Dye	Safeview
SOC medium	Clonetech
Tryptone	LabM
Yeast powder	LabM

2.2 Methods

Mammalian expression plasmid constructs were prepared for overexpression experiments. Each NFIB isoform contains an NH₂-terminal human influenza virus hemagglutinin (HA) epitope tag for ease of detection. pCH_mNfib3 construct used in these experiments officially called pCHNFIB-B2 which was a gift from Richard Gronostajski (Addgene plasmid # 31405 ; <http://n2t.net/addgene:31405> ; RRID:Addgene_31405). Additionally, newly constructed pCH_MCS and pCH_GFP control plasmids also have the HA-tag. pCH_MCS vector does not express any protein and it was used as a control vector, while pCH_GFP expresses GFP protein and its use was thought as crucial due to that any protein expression within cells may lead to translation-causing stress to the cell and this situation may give unexplainable results.

2.2.1 pCH_mNfib4 plasmid (pCH_hNFIB1)

Mouse mNfib3 isoform was already existed in our laboratory. This isoform resembles to human hNFIB3 isoform which is the intermediate long hNFIB isoform in human. Human intermediate-long hNFIB3 differs from the mouse mNfib3 isoform at 2 aminoacids which has similar chemical characteristics. On the other hand, the other hNFIB isoform which is the longest isoform (hNFIB1) in human has nearly the same aminoacid sequence (only 3 aminoacid difference, but also aminoacids with similar chemical characteristics) with mouse mNfib4. Because we have pCH_mNfib3 and its N-terminal domain will be used for construction of the longest isoform, we decided to clone mouse mNfib4 ORF which was supposed to encode mNfib4 with very similar functions with the hNFIB1 isoform. By digesting the pCH_mNfib3 (hNFIB3) with KpnI at two sites one from the start of the C-terminal domain and the other one from the end of the C-terminal domain, this resulting pCH with N-terminus of the mNfib was used as the backbone for cloning of mNfib4. mNfib4 was detected at C17.2 cDNA by using detection primers shown in Table 2.6. mNfib4 was amplified from C17.2 cDNA using cloning primers shown in the Table 2.7. This C-terminal mNfib4 (hNFIB1) PCR product was introduced into TOPO expression vector; in order to amplify this C-terminus using this TOPO_mNfib4 as a template with infusion primers for use in homologous recombination reaction. TOPO-cloning reaction ingredients are shown in Table 2.8.

Table 2.6: mNfib4 detection primers.

Primer	Sequence
mNfib4_1 (Forward)	5'-CCACTGAAGCCTACACAGCC-3'
mNfib4_2 (Reverse)	5'-GGTACCAGGACTGTTGCTGA-3'

Table 2.7: mNfib4 cloning primers.

Primer	Sequence
mNfib4_1 (Forward)	5'-TGCCAAGAATCCT CCAGGGTA-3'
mNfib4_2 (Reverse)	5'-GGTACCAGGA CTGTTGCTGA-3'

Table 2.8: TOPO Cloning reaction.

Reagent	Amount
PCR product	4 µl
Salt solution	1 µl
pCR-Zero Blunt TOPO II	1 µl
Final volume	6 µl
22°C (RT) incubation for 5 minutes	

2 µl of TOPO_mNfib4 reaction was transformed to 50 µl DH5alpha competent cells. Since TOPO vector bears kanamycin resistance gene, LB plates were prepared as including kanamycin for selection. There were 2 colonies in 4/5 plate and plasmids of those two colonies were isolated by using MN Nucleospin miniprep kit and their concentrations were measured by UV-spectrometry as shown in the Table 2.9. In order to detect the accuracy of the plasmid on gel, EcoRI diagnostic digestion was performed.

Table 2.9: UV-Spectrophotometer results for miniprep of Zero Blunt II TOPO_mNfib4 plasmids obtained from the two colonies.

Sample No.	Concentration(ng/µl)	OD _{260/280}	OD _{260/230}
1	330.9	1.93	2.16
2	325.9	1.9	2.1

Table 2.10: Infusion primers to amplify TOPO_mNfib4.

Primer	Sequence
pCH_mNfib4_F (Forward)	5'-CAAGAATCCTCCAGGGT ACCTCGAGGACAGCTTTGT -3'
pCH_mNfib4_R (Reverse)	5'- ACCAAGCTAGCCCAGATACC AGGACTGTTGCTGATGCAGGA-3'

Table 2.11: Cloning enhancer treatment to PCR product.

Reagent	Volume
PCR Product	5 μ l
Cloning Enhancer	2 μ l
37°C incubation for 15 minutes	
80°C incubation for 15 minutes	

Table 2.12: Infusion reaction components.

Reagent	Sample (+)	Control (-)
5X InFusion HD Enzyme Premix	2 μ l	2 μ l
pCH-vector	2 μ l (50 ng/ μ l)	2 μ l
PCR product (Cloning enhancer-treated)	2 μ l	-
dH ₂ O	4 μ l	6 μ l

2.2.2 pCH_MCS control plasmid

As mammalian expression vector, pCH was used. As a control plasmid, an empty pCH which encodes any protein was prepared which is called ‘‘pCH_MCS’’ in this study. This vector was prepared by digestion of pCH_NFIB3 commercial vector with SfiI and BglII; and then a multiple cloning site including multiple restriction digestion enzyme sites and stop codons were annealed and ligated into it. The sequence of the annealed MCS oligos are written in the Table 2.13. 1 μ M of phosphorylated oligonucleotides were incubated at 95°C for 5 minutes and annealed at 70°C for 10 minutes and then it slowly cooled down to 37°C for 15 minutes and then cooled down to 10°C before storage at -20°C.

Table 2.13: pCH_MCS_control oligonucleotide sequences.

Primer	Sequence
pCH_MCS_F (Forward)	5'-TAACCGGTCTGATGTACACGAGCTC TAGCTTAAGA-3'
pCH_MCS_R (Reverse)	5'-GATCTCTTAAGCTAGAGCTCGTGTACATC AGACCGGTTAGTG -3'

2.2.3 pCH_GFP plasmid

In addition to pCH_MCS which encodes for any protein; GFP-coding pCH construct was also prepared to have a protein-coding control. GFP ORF was amplified from pEGFP-C2 empty vector by pCH_GFP infusion primers (Table 2.14) for homologous recombination reaction. pCH_NFIB2 was cut by SfiI and BglII sequentially, and resulting backbone was extracted from the gel and its concentration was estimated as

Table 2.14: pCH_GFP infusion primers.

Primer	Sequence
pCH_GFP_F (Forward)	5'-GACTATGCTGGGCCTATG GTGAGCAAGGGCGAGG-3'
pCH_GFP_R (Reverse)	5'-AGCTATGCATAGATCTCT AGATCCGGTGGATCCCGG-3'

2.2.4 qRT-PCR Experiments

In order to detect the endogenous expression of NFIB-S in glioblastoma (U87MG and T98G) and neuroblastoma (SHSY5Y) cell lines, quantitative-PCR primers shown in the Table 2.15 were used. Additionally, hNFIB qPCR primers were used to see all hNFIB isoforms' expression, and the expression levels were calculated by normalizing the hNFIB values to GAPDH values.

Table 2.15: qRT-PCR primers of all NFIB and NFIB-S transcripts.

Target Gene	Target Sequence	Binding Sequence
NFIBS	CACAGCCATGAACAGCACAC	4-23
NFIBS	TGGCAGGATCATTGTGGCTT	167-186
NFIBS	GCAATTTAGTTGGGGCGACA	76-95
NFIBS	GGGTTATGGGCGTTCTGGAT	260-279
NFIB	GCACACTGAGGAGGACACAA	18-37
NFIB	GGTCTTTTGCTGCTTGGTGG	378-397

2.2.5 Dual luciferase reporter assays

In order to detect and compare the promoter activation capabilities of canonical isoforms of hNFIB which are hNFIB1, hNFIB3 and hNFIB4; firstly GFAP promoter controlled Firefly luciferase expression vector is used in dual luciferase assay system. HEK293T cells were cultured in 24-well plates as 10×10^4 cells/well 1 day before transfection. Transfections for dual luciferase assays were done by using FuGENE6 Promega transfection reagent with the DNA:Reagent ratio 1:3. Assays were performed with triplicates. Each replicate was transfected with a mixture prepared with 50 μ l DMEM only, 3.6 μ l FuGENE6 reagent, 200 ng GFAP promoter, 200 ng* overexpression plasmid, 5 ng Renilla luciferase, required amount of pGEX**. Firstly DMEM only and FuGENE6 reagent was mixed and are incubated for 5 minutes, and then the remaining DNA constructs were added prior to 15 minutes incubation. Finally, cells were added with mixture and incubated at 37°C for 48 hours at %5 CO₂.

Secondly, in order to detect and compare the promoter repression capabilities of those isoforms, dual luciferase assays were done with p21 promoter controlled Firefly luciferase expression vector. NFIB represses p21 promoter *in vivo* and it could be shown *in vitro*. Experimental procedures were same with the GFAP promoter activation experiments.

* According to the sizes of the mammalian expression vectors or control plasmids, their molar concentrations were equalized. pCH_mNfib3 (pCH_hNFIB3) mole was set to 200 ng, and the remaining pCH_MCS, pCH_GFP, pCH_mNfib4 (pCH_hNFIB1), and pCH_NFIB-TAD constructs' moles were calculated and used in proper amounts.

** In order to equalize the DNA stress on the cells and make certain the 1:3 DNA:FuGENE6 ratio for efficient transfection, the higher amount of DNA added to the mixture was set as the maximum DNA mole and all of the transfection mixtures were filled to this DNA amount by adding pGEX bacterial plasmid. It is supposed not to affect the promoter activation/repression.

At 48th hour, cells were washed with PBS and lysed with the addition of 100 µl 1X Passive Lysis Buffer. Lysates were prepared with incubation of the plates at room temperature by shaking for 15 minutes. For each experiments, plates were freeze and thaw or at least waited for 15 minutes at -20°C for better lysis. Lysates were scraped with the pipette tips and collected to 1.5 ml microcentrifuge tubes and centrifuged at RT for 1-2 minutes at 17.000xg (max speed in our laboratory). 20 µl of supernatants of each replicates were measured in the luciferase assay. Firstly, firefly (*Photinus pyralis*) luciferase luminescence was measured by adding 75 µl Luciferase Assay Reagent II (LAR II) onto 20 µl of the cell lysate. In order to measure the luminescence stemmed from Renilla (*Renilla reniformis*), 75 µl Stop & Glo Reagent was added and the signal was measured. Chemiluminescence was measured using Fluoroskan Ascent FL luminometre (Thermo Electron Corporation) and default settings were used for the PMT voltage (1200) and the integration time was chosen as 10000. Relative firefly luciferase signals were normalized against Renilla luciferase signals to minimize experimental variability caused by differences in transfection efficiency and cell viability and lysis efficiency.

2.2.6 Electrophoretic mobility shift assays

Full length NFIB (NFIB-FL), short NFIB (NFIB-S or hNFIB4) and human NFIB-Trans-Activating-Domain (TAD) were overexpressed in HEK293T cells in order to detect their binding ability to NFI-DNA-binding consensus sequences on GFAP and p21 promoters.

p21 oligonucleotide sequence used in EMSA experiments (Table 2.16) was biotinylated by using Pierce Biotin 3'End DNA Labeling Kit (ThermoScientific) and biotinylation efficiencies of both single strand oligonucleotides were detected. After detection of biotinylation efficiencies, these biotinylated single strand oligonucleotides were annealed to have double stranded bitonylated EMSA oligonucleotides.

Table 2.16: Oligonucleotide sequences of GFAP and p21 NFI binding sites used in mobility shift experiments.

	Oligo strand	Oligo strand
GFAP	Upper strand	5'- CCATAGCTGGGCT GCGGCCCAACCCCA-3'
GFAP	Lower strand	5'-TGGGGTTGGGCCG CAGCCCAGCTATGG-3'
p21	Upper strand	5'-GCCTGCTGGAAGCTCGGCCAG GCTCAGCTGGCTCCGCGCTGG-3'
p21	Lower strand	5'-CCAGCGCGGAGCCAGCTGAG CCTGGCCGAGTTCAGCAGGC-3'

Table 2.17: Components of DNA-binding reactions in EMSA experiment made with GFAP promoter oligonucleotides.

	Control	NFIB-FL OE	NFIB-S OE	NFIB-TAD OE
Ultrapure water	12 µl	10 µl	11.5 µl	11 µl
10X Binding Buffer	2 µl	2 µl	2 µl	2 µl
50% Glycerol	1 µl	1 µl	1 µl	1 µl
100mM MgCl₂	1 µl	1 µl	1 µl	1 µl
20mM EDTA	1 µl	1 µl	1 µl	1 µl
1µg/µl Poly (dI-dC)	1 µl	1 µl	1 µl	1 µl
Extract	-	2 µl	0.5 µl	1 µl
Biotinylated DNA	2 µl	2 µl	2 µl	2 µl

Table 2.18: Components of DNA-binding reactions in EMSA experiment made with p21 promoter oligonucleotides.

	NFIB-FL OE	NFIB-S OE	NFIB-TAD OE	NFIB-FL + NFIB-S OE	NFIB-FL + NFIB-TAD OE
Ultrapure water	10 µl	10 µl	10 µl	10 µl	10 µl
10X Binding Buffer	2 µl	2 µl	2 µl	2 µl	2 µl
50% Glycerol	1 µl	1 µl	1 µl	1 µl	1 µl
100mM MgCl ₂	1 µl	1 µl	1 µl	1 µl	1 µl
20mM EDTA	1 µl	1 µl	1 µl	1 µl	1 µl
1µg/µl Poly (dI-dC)	1 µl	1 µl	1 µl	1 µl	1 µl
NFIB-FL OE extract	1 µl	-	-	1 µl	1 µl
NFIB-S OE extract	-	0.5 µl	-	0.5 µl	-
NFIB-TAD OE extract	-	-	0.5 µl	-	0.5 µl
Untreated HEK293T lysate	1 µl	1.5 µl	1.5 µl	0.5 µl	0.5 µl
Biotinylated DNA	2 µl	2 µl	2 µl	2 µl	2 µl

5% polyacrylamide gel (1.25 ml 40% Acrylamide, 8.75 ml 0.5% TBE Buffer, 200 µl APS, and 10 µl TEMED) was used and pre-run for 50 minutes. *In vitro* binding reaction components are written in the Table 2.17; Table 2.18; and Table 2.19. All the reaction components except biotinylated oligonucleotides were added to the reaction tubes on ice, and after the pre incubation at room teperature for 10 minutes, biotinylated oligonucleotides were also added and incubation was performed at room temperature for 20 minutes. After addition of 5 µl 5X loading buffer to the 20 µl reactions, 20 µl of the mix was loaded into the wells of the pre-run gel and running started at 100 V. Both running and wet transfer was carried out at 0.5% TBE Buffer. Transfer was performed at ~10°C for 40 minutes at 400 mA (voltage was ~66V at all experiments).

Table 2.19: Components of DNA-binding reactions in EMSA experiment made with p21 promoter oligonucleotides.

	Control	NFIB-FL OE	NFIB-S OE	NFIB-FL +NFIB-S OE (1:1)	NFIB-FL + NFIB-S OE	NFIB-N terminus
Ultrapure water	10 µl	10 µl	10 µl	10 µl	10 µl	10 µl
10X Binding Buffer	2 µl	2 µl	2 µl	2 µl	2 µl	2 µl
50% Glycerol	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl
100mM MgCl₂	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl
20mM EDTA	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl
1µg/µl Poly (dI-dC)	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl
NFIB-FL OE extract	-	0.5 µl	0.5 µl	0.5 µl	-	1 µl
NFIB-S OE extract	-	-	0.5 µl	1.5 µl	1.5 µl	-
NFIB-TAD OE extract	-	-	-	-	-	0.5 µl
Untreated HEK293T lysate	2 µl	1.5 µl	1 µl	-	0.5 µl	0.5 µl
NFIB-N terminus	-	-	-	-	-	2 µl
Biotinylated DNA	2µl	2µl	2µl	2µl	2µl	2µl



3. RESULTS

In this study, we set out to explore whether structural differences in NFIB alternative splicing isoforms lead to differences in function. To this end, we focused on two mouse isoforms, mNfib3 and mNfib4, as well as a truncated human NFIB isoform (NFIB-S) that lacks the N-terminal DNA binding and dimerization domain. In order to carry out gain of function studies, first, mammalian expression constructs were generated.

3.1 Preparation of Expression Constructs

We generated several the constructs required for the overexpression experiments, including mock controls. Since the mNfib3 open reading frame (ORF) is already cloned into the pCH expression vector, mNfib4 was also inserted into the same plasmid. For electrophoretic mobility shift assays (EMSA) and luciferase reporter experiments, NFIB isoforms should be ectopically expressed in HEK293T cells, and a mock construct is needed as negative control. An empty vector is transfected to determine background the firefly and renilla luciferase luminescence in the absence of exogenous protein, and to that end, we constructed a control plasmid, pCH_MCS. Preparation of pCH_mNfib4, pCH_MCS and pCH_GFP constructs is described in detail below.

3.1.1 pCH_mNfib4 construct

The ORF of Nfib4 differs from Nfib3 in its 3' end which encodes the C-terminal transactivation domain. In order to prepare pCH_mNfib4, we chose to amplify the 3' end of Nfib4 and insert this fragment into the pCH_NFIB3 plasmid. To this end, pCH_NFIB3 was linearized with KpnI that left the 5' end of the Nfib3 ORF intact in the plasmid. The 3' end of mNfib4 ORF was amplified from C17.2 mouse neural stem cell cDNA (Figure 3.1). This PCR product was first subcloned into the Zero Blunt II TOPO vector, confirmed by sequencing, was amplified out of the Zero Blunt II TOPO construct with infusion primers using homologous ends to the linearized pCH_NFIB3 (Figure 3.2). The amplified fragment was then inserted into pCH_NFIB3 by

homologous recombination. DH5 α bacterial cells were transformed with the recombination reaction, resulting plasmids were purified and then were subjected to diagnostic digestion by KpnI before DNA sequencing (Figure 3.3).

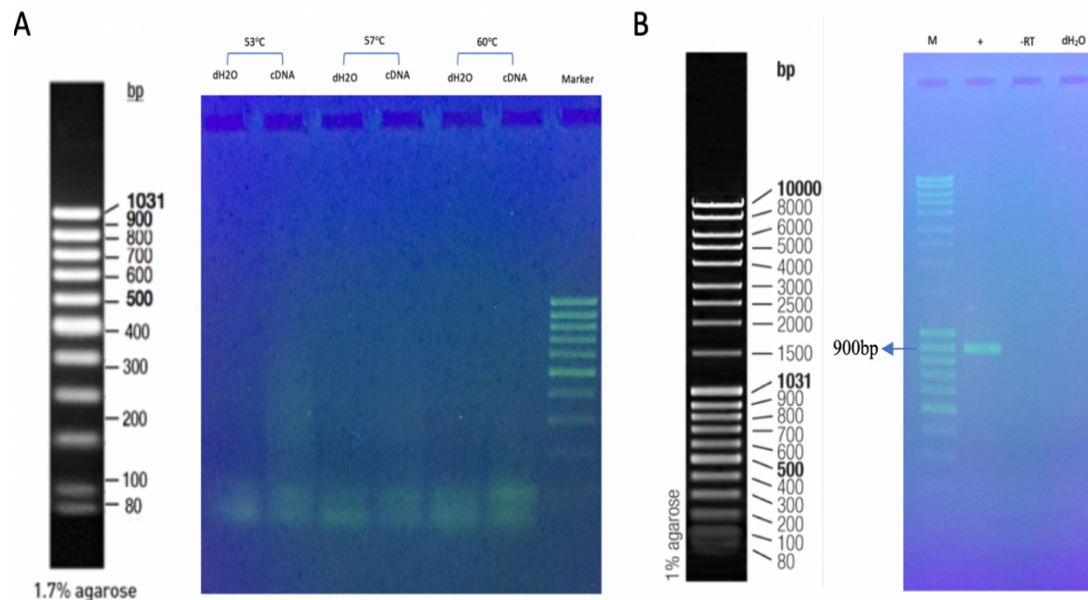


Figure 3.1: mNfib4 3' fragment was amplified from C17.2 cDNA. **A)** In order to determine expression of mNfib4 in C17.2 mouse neural progenitors, we performed gradient PCR using primers that were designed to detect mNfib4. At increasing temperatures, amount of mNfib4 amplicon increased (102 bp bands). **B)** 878bp segment of mNfib4 ORF was amplified C17.2 cDNA and this PCR product was inserted into pCR-BluntII TOPO. In this case, neither the no template controls (water nor -RT) did not produced any PCR products. M: DNA Ladder, +: Reaction of PCR amplification of C17.2 cDNA with mNfib4 cloning primers, -RT: PCR with the control template that was from C17.2 cDNA synthesis reaction which did not added reverse transcriptase enzyme, dh₂O: Control PCR with non-pyrogenic PCR water as a template.

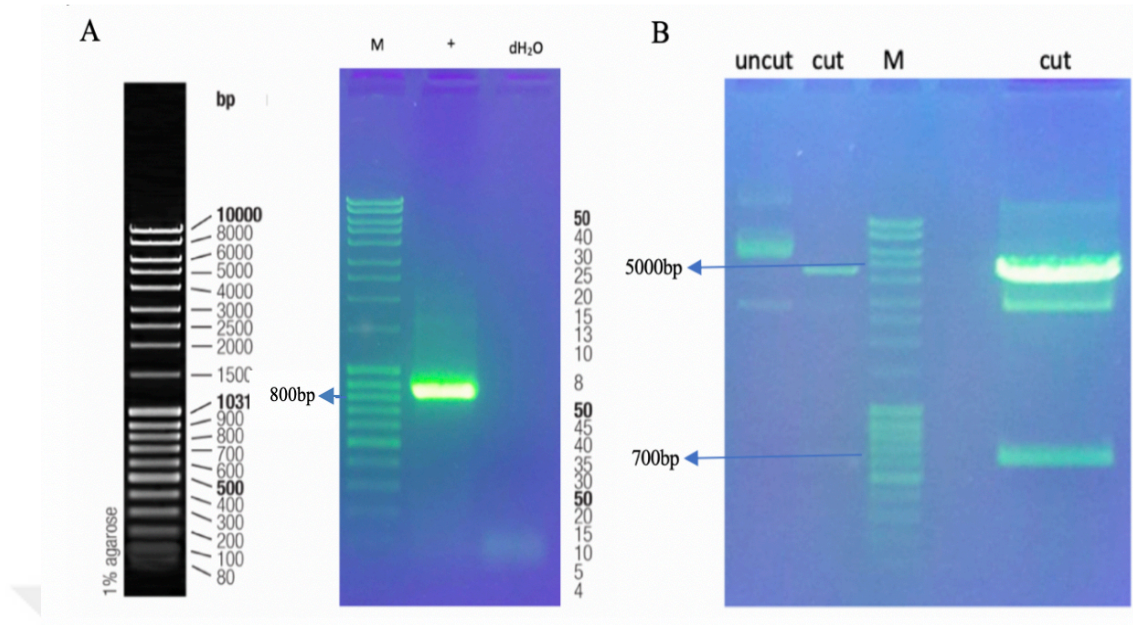


Figure 3.2: Amplification of mNfib4 with infusion primers from TOPO_mNfib4. **A)** After subcloning of C-terminal domain sequence of mNfib4 via TOPO-cloning, this sequence was amplified with infusion primers from TOPO_mNfib4 template to have an infusion PCR product (889 bp) with homologous overhangs to the KpnI-digested pCH backbone. pCH vector is the mammalian expression vector we used in our experiments. M: DNA Ladder, +: Reaction of PCR amplification of C17.2 cDNA with mNfib4 cloning primers, -RT: PCR with the control template that was from C17.2 cDNA synthesis reaction which did not added reverse transcriptase enzyme, dH₂O: Control PCR with non-pyrogenic PCR water as a template. **B)** Digestion of pCH_NFIB plasmid by KpnI. KpnI restriction of pCH_NFIB generated two fragments: 4413 bp backbone encoding the N-terminal domain of mNfib3 and a 633 bp segment that was removed from the plasmid. M: DNA Ladder, uncut: untreated pCH_NFIB plasmid, cut: KpnI-digested pCH_NFIB backbone.

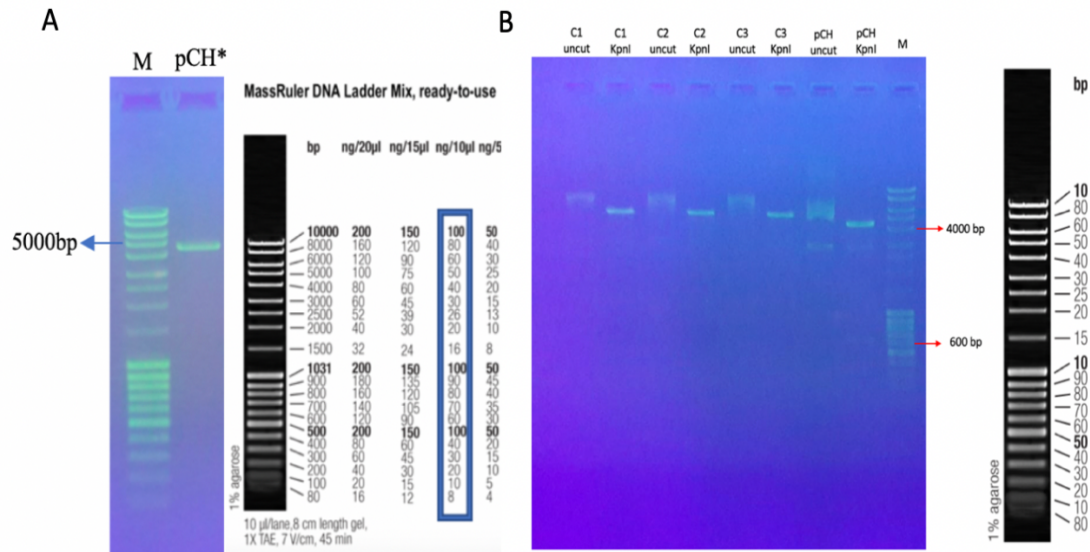


Figure 3.3: mNfib4 3' fragment was cloned into pCH expression vector. **A)** Gel-purified KpnI-digested pCH backbone was analyzed by resolved by agarose gel electrophoresis and its concentration was estimated in order to set up the infusion cloning reaction. Concentration of pCH* (KpnI-digested pCH) was estimated as ~50 ng/µl. M: DNA Ladder, pCH*: KpnI-digested and gel extracted pCH backbone. **B)** pCH_mNfib4 plasmids obtained by recombination were restricted with KpnI and analyzed by agarose gel electrophoresis. All the three plasmids isolated gave the expected restriction products of 5268 bp while parental pCH_mNfib3 gave the restriction products of 4413 and 633 bp. C1-C3 uncut: untreated plasmid isolated from colonies 1-3; C1-C3 KpnI: KpnI-treated plasmid isolated from colonies 1-3; pCH uncut: untreated parental pCH_NFIB plasmid, pCH KpnI: KpnI-treated parental pCH_NFIB plasmid, M: DNA Ladder.

3.1.2 pCH_MCS construct

In order to produce a mock control plasmid, we generated a pCH plasmid with no insert. pCH_NFIB3 that encodes the ORF of mouse Nfib3 had been a generous gift of Richard Gronostajski. Nfib3 ORF was removed by digestion with SfiI at its 5' end and by BglII at its 3' end (Figure 3.4A). Next, a multiple cloning site (MCS) was ligated into the vector to use this vector for further cloning in the future. Resulting pCH_MCS plasmids were cloned in DH5α bacterial cells and purified; the plasmids were subjected to diagnostic digestion by BsrGI and HindIII enzymes before confirmation by DNA sequencing analysis (Figure 3.4B).

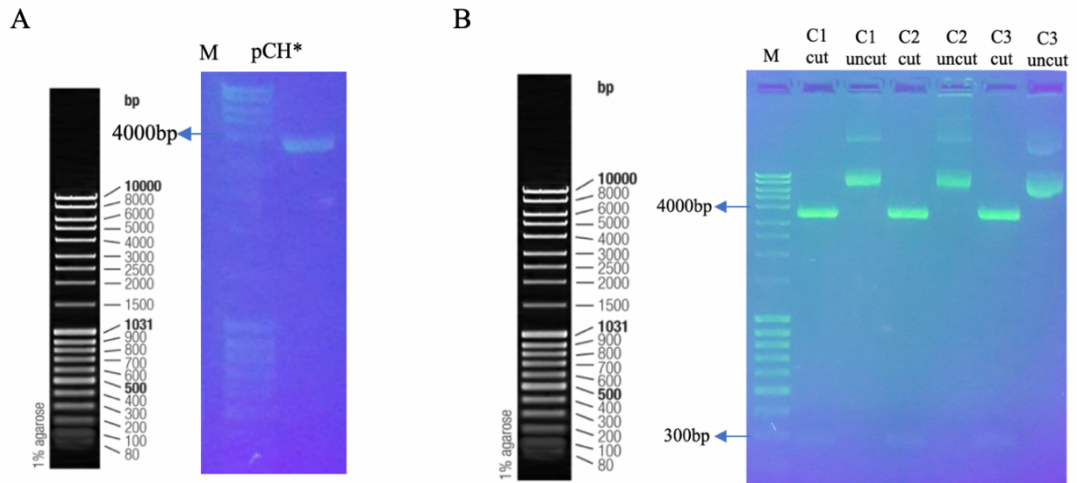


Figure 3.4: pCH_MCS was generated by ligation. **A)** pCH_NFIB was restricted by SfiI and BglII and a 3754 bp backbone was purified by agarose gel electrophoresis. The concentration of this fragment was estimated as ~30 ng/μl. M: DNA Ladder, pCH*: SfiI and BglII-double digested and gel extracted pCH backbone. **B)** pCH_MCS plasmids obtained from the ligation reaction were restricted with BsrGI and HindIII. All the three plasmids isolated gave the expected restriction products of 3516 bp and 280 bp. (BsrGI site was inserted into the ligated multiple cloning site (MCS). M: DNA Ladder, C1-C3 cut: BsrGI and HindIII-treated plasmid isolated from colonies 1-3; C1-C3 uncut: untreated plasmid isolated from colonies 1-3.

3.1.3 pCH_GFP construct

As with the pCH_MCS construct, GFP ORF was cloned pCH_NFIB3 linearized with SfiI and BglII, this time GFP ORF was amplified from pEGFP-C2 vector, and inserted by homologous recombination (Figure 3.5). Resulting pCH_GFP plasmids were cloned in DH5α bacterial cells and purified; then the plasmids were subjected to diagnostic digestion by BsrGI and NotI enzymes before DNA sequencing analysis (Figure 3.6).

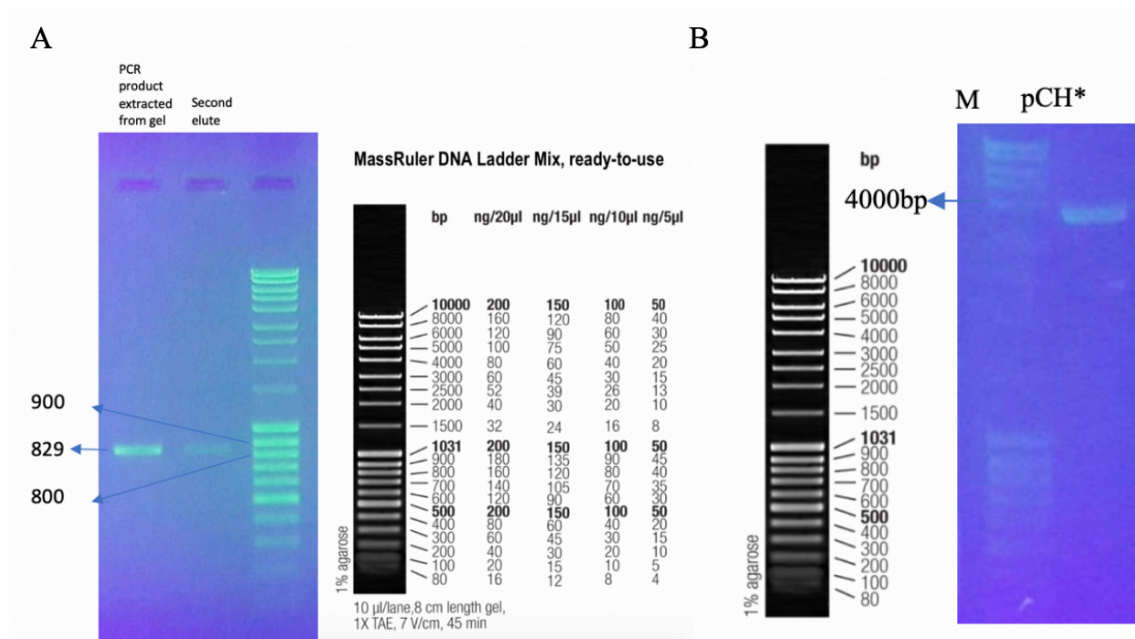


Figure 3.5: Generation of pCH_GFP. **A)** The 829 bp fragment encoding the GFP ORF was amplified from pEGFP-C2 plasmid. **B)** pCH_NFIB was digested with SfiI and BglII and purified by agarose gel electrophoresis. The purified fragment was estimated as ~30 ng/µl. M: DNA Ladder, pCH*: SfiI&BglII double digested pCH_NFIB.

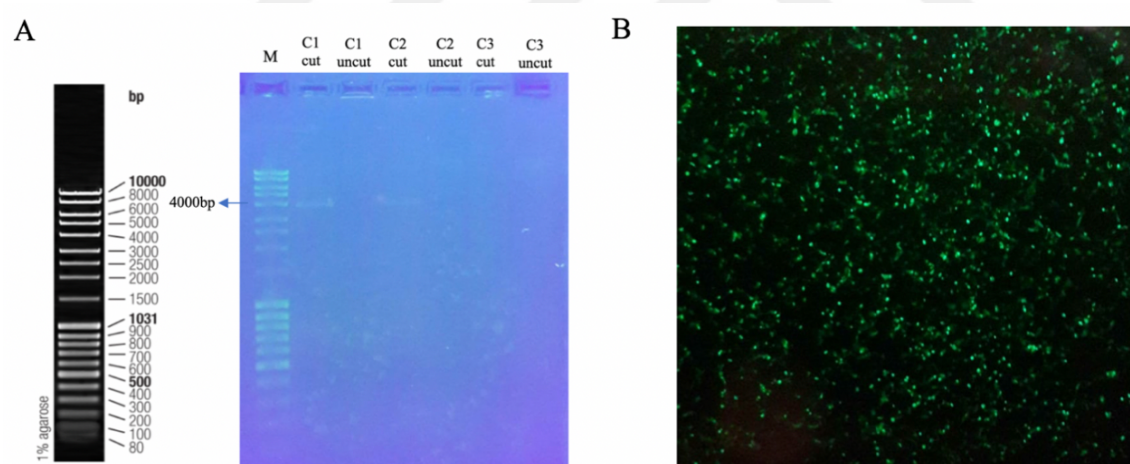


Figure 3.6: Diagnostic restriction digestion of pCH_GFP and overexpression in HEK293T cells. **A)** Plasmids obtained from the recombination reaction were restricted with BsrGI and NotI. Because that the isolated plasmid concentration were very low, low amounts of plasmids were digested and run on gel. Two bands, 3799 bp and 758 bp in sizes, were expected. M: DNA Ladder, C1-C3 cut: BsrGI and NotI-treated plasmid isolated from colonies 1-3; C1-C3 uncut: untreated plasmid isolated from colonies 1-3. **B)** In order to verify GFP expression from these constructs, HEK293T cells were transfected with all three isolated plasmids. All the pCH_GFP constructs were able to express GFP, and the one shown in panel B was chosen for further experiments.

3.2 Characterization of NFIB-S in SH-SY5Y, U87MG, T98G and HeLa cell lines

NFIB-S (human NFIB4) expression has been detected in megakaryocytes before (Chen et al., 2014) but a mouse ortholog has not yet been reported. While exploring NFIB's role in neural stem cell differentiation, expression of different NFIB isoforms in human neural stem cells (hNSCs) was previously analyzed (Cesur, 2015). RT-PCR experiments found that *NFIB-S* is expressed in both progenitor (hNSCs) and SH-SY5Y neuroblastoma cells. To verify these findings and determine *NFIB-S* expression levels in human neural cell lines, qRT-PCR analysis was performed on total mRNA isolated from human glioblastoma and neuroblastoma cells as well as HeLa cells. Since Hein et al. reported NFIB-S protein expression in the HeLa cells in a high throughput proteomics screen, we also included this cell line as positive control. We determined the expression levels of both total *NFIB* and specifically *NFIB-S* in SH-SY5Y, U87MG and T98G as human neural cell lines. NFIB qRT-PCR primers targeting all isoforms and NFIB-S (hNFIB4) isoform-specific qRT-PCR primers were used to detect *NFIB* and *NFIB-S* expression levels. Two sets of NFIB-S qRT-PCR primers were used: primers hNFIB-S_3,4 and hNFIB-S_5,6. Both total *NFIB* and specifically *NFIB-S* expression levels were highest in SH-SY5Y cells followed by U87MG and T98G cell lines. Congruently, among human neural cell lines, SH-SY5Y cells exhibit highest total NFIB expression, followed by U138MG, U87MG, and U251MG cells respectively (received from “<https://www.proteinatlas.org/ENSG00000147862-NFIB/cell>”; on August 16, 2019). Based on Human Protein Atlas data, *NFIB* expression levels in SH-SY5Y cells is approximately 5-fold higher than the *NFIB* expression in U87MG. We found that the first primer set produced C_t values between 26.6-28.6 (26.6 for SH-SY5Y; 28 for U87MG; 28.6 for T98G; and 27.6 for HeLa cells) (Figure 3.7), while the second primer set C_t values ranged between 27.7-28.9 (the C_t value for SH-SY5Y was 28.5; for U87MG was 28.9; for T98G the C_t was 28.7 and for HeLa the C_t value was 27.7) (Figure 3.8). On the other hand, hNFIB primers that detect all four isoforms produced the C_t values of 18.2-20 (18.2 for SH-SY5Y; 20 for U87MG, 20 for T98G; and 18.9 for HeLa cells) (Figure 3.9). Thus, while both NFIB-S qRT-PCR primers gave similar results, primers hNFIB4_3, 4 could detect *NFIB-S* earlier, especially in SH-SY5Y cells. While total *NFIB* measured with hNFIB qRT-PCR primers could be detected at C_t 18.2 in SH-SY5Y cells, *NFIB-S* was detected at C_t 28 (Figure 3.9). Thus there is a 10 cycle difference between the total *NFIB* versus

NFIB-S qPCR products detected in SH-SY5Y cells (Table 3.1). In U87G and T98G glioblastoma cell lines, the cycle difference is approximately 8.5; while the difference is approximately 9 in HeLa cell line. While not quantitative, these data suggest that *NFIB-S* (*hNFIB4*) is present in both neuroblastoma cells and glioma cell lines, even though it does not constitute the majority of *NFIB* mRNA.

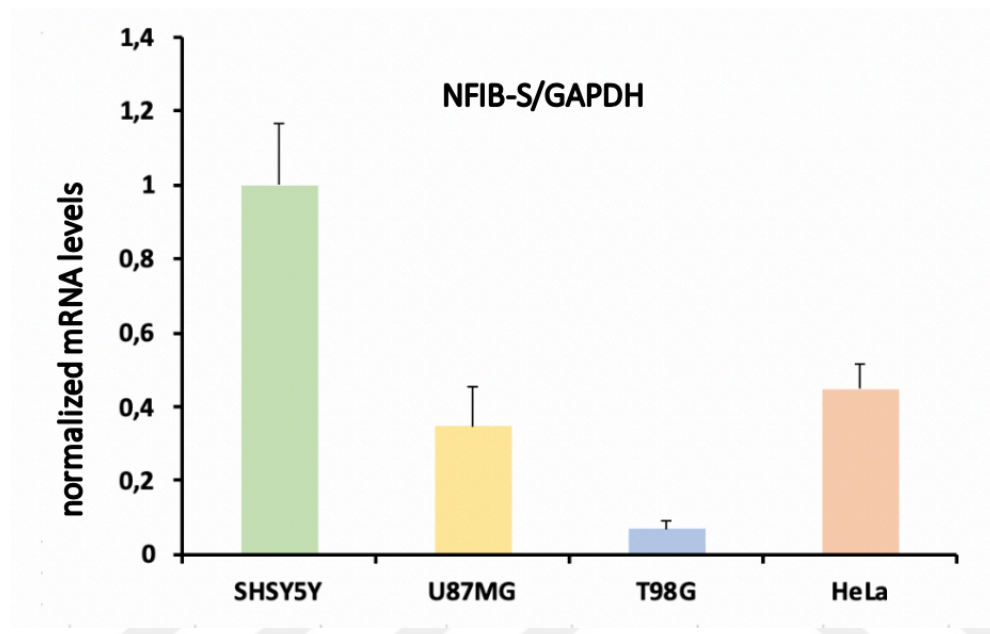


Figure 3.7: Characterization of NFIB4 expression in SH-SY5Y, U87MG, T98G and HeLa cell lines. NFIB-S (NFIB4) expression level is measured by qRT-PCR using primer set hNFIB_3, 4. Gene expression levels were normalized to GAPDH expression. Data show the mean $\pm$ S.D from 2 replicates.

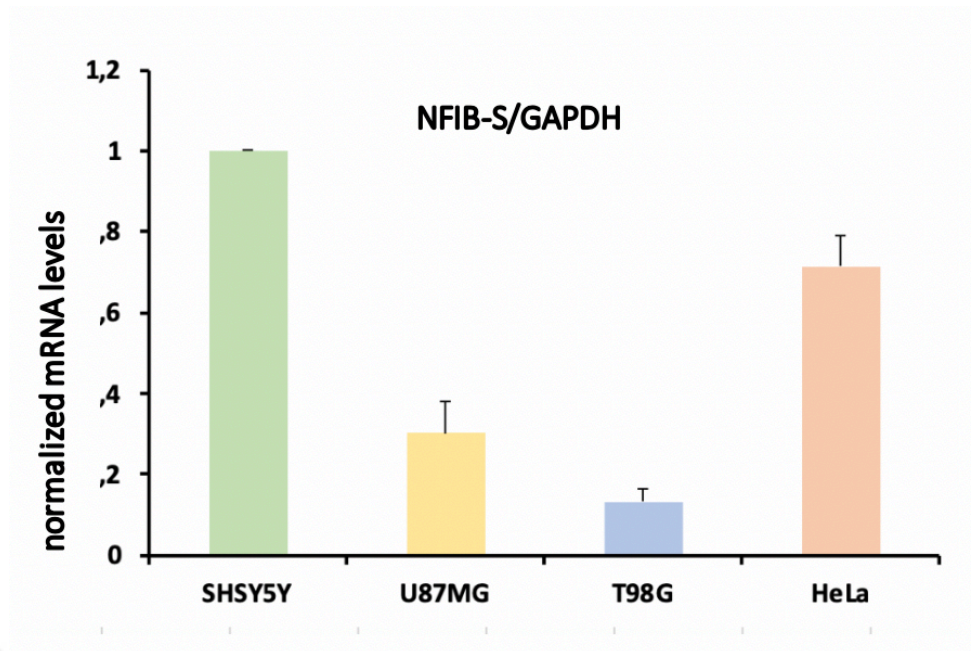


Figure 3.8: Characterization of NFIB4 expression in SH-SY5Y, U87MG, T98G and HeLa cell lines. NFIB-S(NFIB4) expression level is measured by qRT-PCR with hNFIB_5, 6 primer set. Gene expression levels were normalized to GAPDH expression. Data show the mean $\pm$ S.D from 2 replicates.

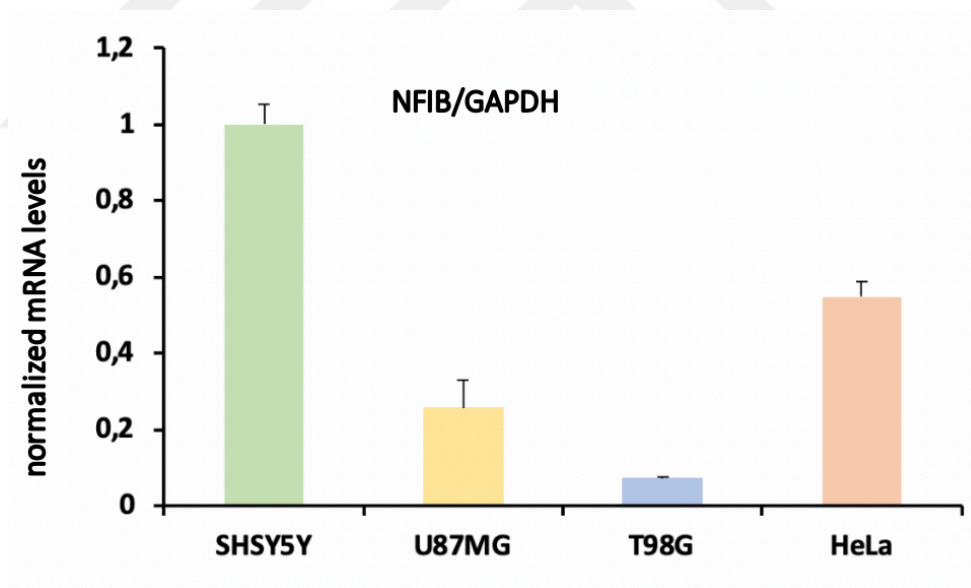


Figure 3.9: Characterization of NFIB4 expression in SH-SY5Y, U87MG, T98G and HeLa cell lines. Total NFIB expression level is measured by qRT-PCR with hNFIB primers. Gene expression levels were normalized to GAPDH expression. Data show the mean $\pm$ S.D from 2 replicates.

Table 3.1: C_t values produced by NFIB and NFIB-S qRT-PCR primers on mRNA extracts of SH-SY5Y, U87MG, T98G, and HeLa cell lines.

Primer set	Target transcript	SH-SY5Y	U87MG	T98G	HeLa
hNFIB-S _{3, 4}	<i>NFIB-S</i>	26.6	28	28.6	27.6
hNFIB-S _{5, 6}	<i>NFIB-S</i>	28.5	28.9	28.7	27.7
hNFIB	Total <i>NFIB</i>	18.2	20	20	18.9

3.3 Transcriptional regulation by NFIB isoforms

Next we examined the promoter activation/repression capabilities of the NFIB isoforms. Luciferase reporter assays were performed and NFIB-TAD which comprises the transactivating C-terminal domain of NFIB (without the DNA-binding domain) was also included in parallel with NFIB-S (NFIB4). In order to measure and compare the promoter activation/repression capacities of two mouse *Nfib* isoforms, *Nfib3* (hNFIB3) and *Nfib4* (hNFIB1), we selected *GFAP* and *p21* promoters, which bear NFI-binding sites. These regulatory regions include the sequences that NFI can bind to and regulate expression of target genes via affecting the nucleosome structure, the recruitment of RNAPolII machinery elements, or the interactions between other transcription factors with regulatory sequences. *GFAP* was used to show activation capacity as this gene is induced by NFIs (Brun et al., 2009) and *p21* was used to study repression capacity because it is repressed by NFIB (Ouellet et al., 2006).

3.3.1 *GFAP* promoter activation by NFIB isoforms

Firefly luciferase luminescence values were normalized to renilla luciferase luminescence in order to compensate errors arising from cell number, transfection efficiency, and lysis efficiency differences of the cell lysates used in measurements. First, when pCH empty vector (pCH_MCS) was used as control, full length NFIB (*Nfib3* or *Nfib4*; which both carry the conserved DNA-binding and dimerization domain) activated the *GFAP* promoter as expected. When the activation capacities of the two mouse isoforms were compared, m*Nfib4* (hNFIB1) was indeed more potent. While m*Nfib3* showed 1.8-fold activation with respect to pCH mock, m*Nfib4* showed

2.6-fold activation. Since mNfib4 has two exons and a longer transactivation domain, it may have a higher promoter activation ability. However, western blot analysis showed mNfib4 to be expressed at higher levels (Figure 3.10). When we repeated this experiment, we observed similar results with 1.7-fold activation by mNfib3 and 3-fold activation by mNfib4 with respect to pCH mock, again with higher mNfib4 expression (Figure 3.11). Higher expression levels may result in higher promoter activation; in order to obtain similar levels of expression, we decreased mNfib4 constructs by half in the next set of experiments (Figure 3.12). We obtained similar results with 1.8-fold activation by mNfib3 and 2.5-fold activation by mNfib4. However, mNfib4 expression levels remained higher than that of mNfib3. This discrepancy may be due to differences in stability of the two isoforms, as mNfib4 contains two additional exons.

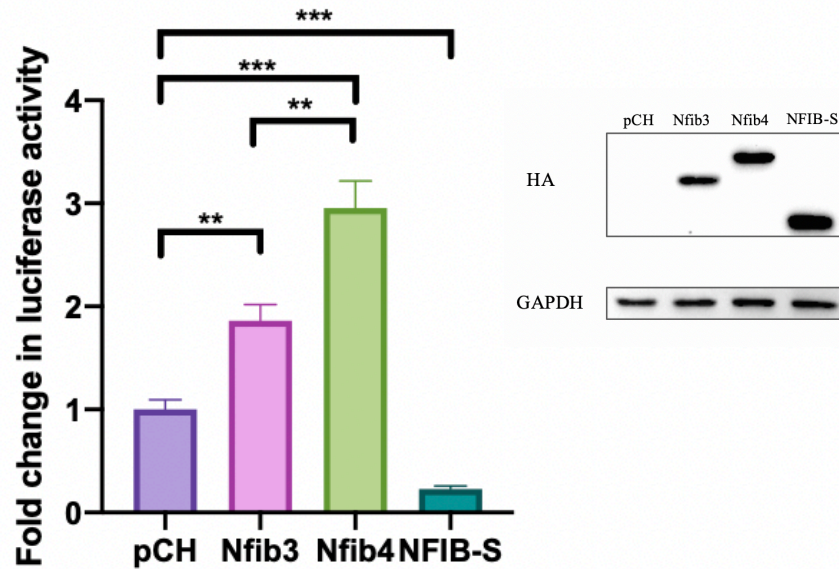


Figure 3.10: *GFAP* promoter activity in mNfib3, mNfib4 and NFIB-S overexpressing HEK293T cells -I. pCH: pCH_MCS which does not express any protein was used as mock. All constructs were encoded in the pCH backbone. 10×10^4 HEK293T cells were co-transfected with pGFA2-luc and pCH_MCS, pCH-mNfib3, pCH-mNfib4, and pCH-NFIB-S expression constructs. The fold increases in Luc activity are relative to the pCH (empty vector) co-transfectants. Luciferase activity was measured in triplicate, and the results were averaged to determine the mean $\pm$ S.D. The Students t test was used for statistical analysis (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ **** $p < 0.0001$). Expression levels of NFI proteins in this experiment was determined by Western Blot (insert). (Nfib3: 49kDA, Nfib4: 56kDA, NFIB-S: 35kDA)

In contrast, the truncated isoform NFIB-S (hNFIB4) appeared to repress *GFAP* with respect to the pCH mock by ~77% (Figure 3.10). Because NFIB-S lacks the conserved

DNA-binding and dimerization domain, this repression was unexpected. In the following experiments, pCH_GFP mock which expresses GFP protein was included as control in case a possible artifactual decrease in *GFAP* activity arises from the burden of translation of overexpressed protein on the cells. In this experiment, overexpression of GFP also lead to decrease in *GFAP* promoter induction, with approximately 55% repression with respect to pCH mock (Figure 3.11). Similarly, NFIB-S expressing samples also showed ~55% repression. When we repeated the experiment, GFP-expressing sample showed ~37% repression with respect to pCH mock while NFIB-S expressing sample showed ~53% repression (Figure 3.12). Thus, NFIB-S expressing cells showed reduced promoter activation with respect to GFP-expressing cells by ~25%. These findings suggest that NFIB-S may repress *GFAP* promoter activity; however further experiments are needed to confirm this result.

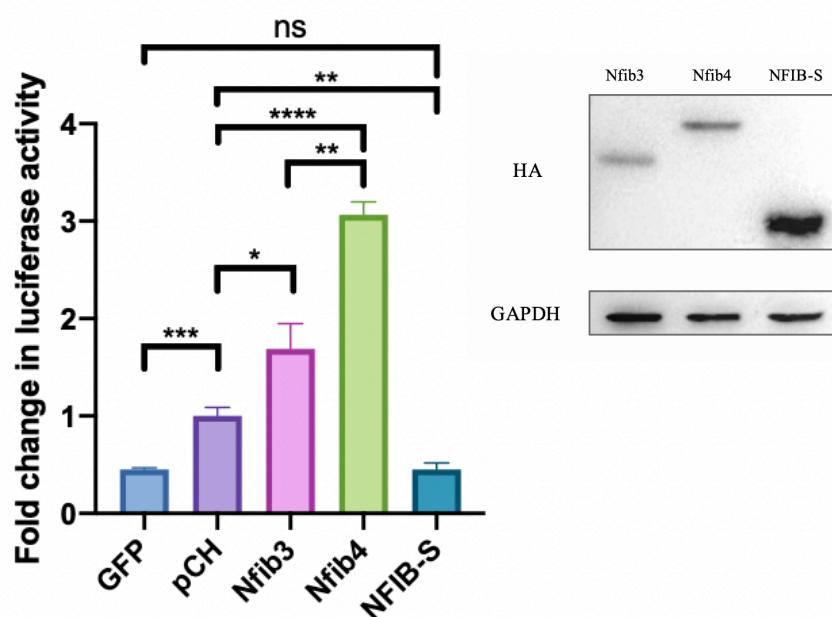


Figure 3.11: *GFAP* promoter activity in mNfib3, mNfib4 and NFIB-S overexpressing HEK293T cells -II. pCH: pCH_MCS which does not express any protein was used as mock. All constructs were encoded in the pCH backbone. 10×10^4 HEK293T cells were co-transfected with pGFA2-luc and pCH_MCS, pEGFP-C2, pCH-mNfib3, pCH-mNfib4, and pCH-NFIB-S expression constructs. The fold increases in Luc activity are relative to the pCH (empty vector) co-transfectants. Luciferase activity was measured in triplicate, and the results were averaged to determine the mean $\pm$ S.D. The Students t test was used for statistical analysis (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ **** $p < 0.0001$). Expression levels of NFI proteins in this experiment was determined by Western Blot (insert). (Nfib3: 49kDA, Nfib4: 56kDA, NFIB-S: 35kDA)

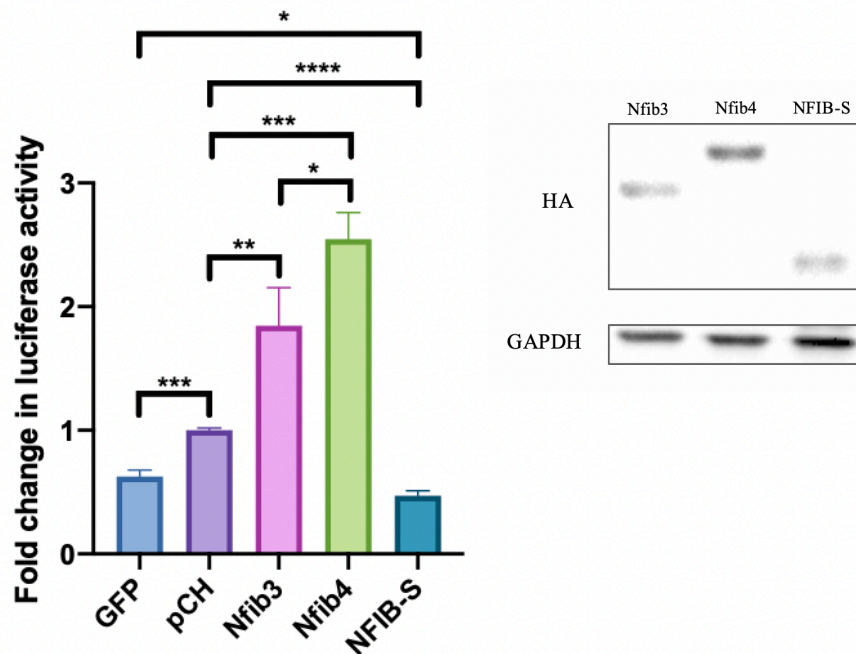


Figure 3.12: *GFAP* promoter activity in mNfib3, mNfib4 and NFIB-S overexpressing HEK293T cells -III. pCH: pCH_MCS which does not express any protein was used as mock. All constructs were encoded in the pCH backbone. 10×10^4 HEK293T cells were co-transfected with pGFA2-luc and pCH_MCS, pCH-GFP, pCH-mNfib3, pCH-mNfib4, and pCH-NFIB-S expression constructs. The fold increases in Luc activity are relative to the pCH (empty vector) co-transfectants. Luciferase activity was measured in triplicate, and the results were averaged to determine the mean $\pm$ S.D. The Students t test was used for statistical analysis (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ **** $p < 0.0001$). Expression levels of NFI proteins in this experiment was determined by Western Blot (insert). (Nfib3: 49kDA, Nfib4: 56kDA, NFIB-S: 35kDA)

3.3.2 *p21* promoter repression by NFIB isoforms

Next we wanted to investigate effects of NFIB isoforms on target genes that are repressed by NFIs. *p21* is one such target gene and the *p21* promoter that is used in these assays contain two NFI binding sites between -161 -149 bp and -1507 -1493 bp. As we described earlier, firefly luciferase luminescence values were normalized against renilla luciferase luminescence values. We also included GFP overexpressing cells as control; indeed, expression of GFP lead to 16% decrease in *p21* promoter activity compared to pCH mock control (Figure 3.13). Full length NFIB isoforms (mNfib3 and mNfib4) repressed the *p21* promoter (by 36% and 45% respectively) compared to the GFP control. When we repeated the experiment, the decrease in *p21* promoter activity was was 44% and 67% in mNfib3 or mNfib4 overexpressing cells

respectively (Figure 3.14). mNfib3 and mNfib4 repressed p21 with similar potency in both two experiments: While mNfib4-expressing samples were able to repress the p21 activation by ~14% (Figure 3.13) and ~39% (Figure 3.14) with respect to mNfib3-expressing samples, mNfib4 expression was higher than the expression of mNfib3 in both experiments (see Figures 3.13 & 3.14, inserts). Preliminary data indicate that mNfib4 may repress the *p21* more efficiently than mNfib3; however further examinations are needed.

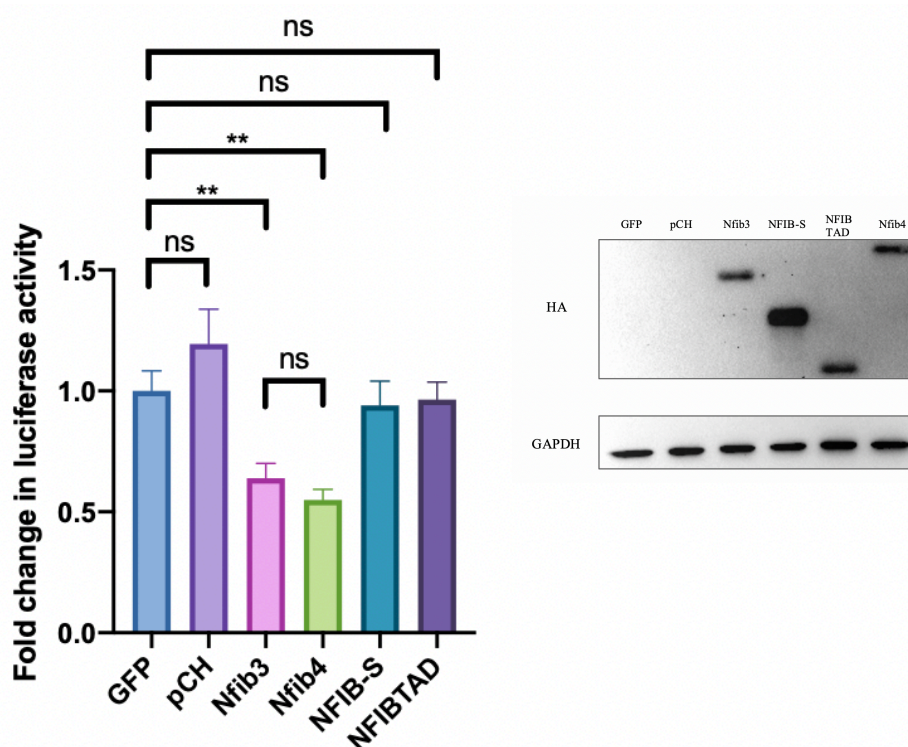


Figure 3.13: *p21* promoter activity in mNfib3, mNfib4, NFIB-S and NFIB-TAD overexpressing HEK293T cells -I. pCH: pCH_MCS which does not express any protein was used as mock. All constructs were encoded in the pCH backbone. 10×10^4 HEK293T cells were co-transfected with pGL2-p21-luc and pCH_MCS, pEGFP-C2, pCH-mNfib3, pCH-mNfib4, pCH-NFIB-S, and pCH-NFIB-TAD expression constructs. The fold increases in Luc activity are relative to the pCH (empty vector) co-transfectants. Luciferase activity was measured in triplicate, and the results were averaged to determine the mean $\pm$ S.D. The Students t test was used for statistical analysis (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ **** $p < 0.0001$). Expression levels of NFI proteins in this experiment was determined by Western Blot (insert). (Nfib3: 49kDA, Nfib4: 56kDA, NFIB-S: 35kDA)

The truncated NFIB isoform, NFIB-S repressed the p21 promoter activity by ~6% compared to the GFP control (Figure 3.13). NFIB-TAD, which is an artificial truncated version of mNfib3, lacks the N terminal DNA-binding and dimerization domain. Specifically, NFIB-TAD is shorter than NFIB-S, which has a different C-terminus end

due to the inclusion of exons 9, 10 and 11. NFIB-TAD overexpressing cells showed a decrease in *p21* promoter activity like those expressing NFIB-S. This result was not reproduced, however, there was a ~17% increase in *p21* promoter activity in NFIB-S overexpressing cells compared to the GFP control, while NFIB-TAD repressed *p21* promoter activity by ~15% compared to the GFP control (Figure 3.14). In this case, expression of NFIB-S was quite low (see Figure 3.14, insert). Preliminary data indicate that NFIB-S does not have an effect of *p21* promoter activation, however, these experiments need to be repeated to draw definitive conclusions.

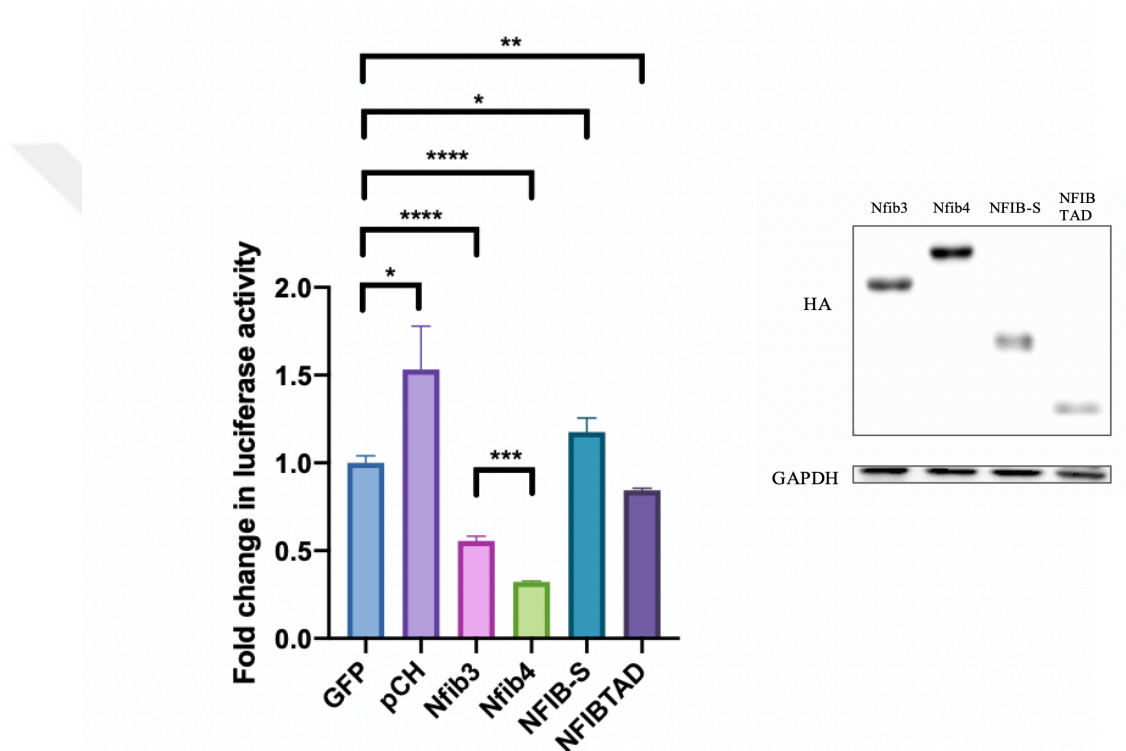


Figure 3.14: *p21* promoter activity in mNfib3, mNfib4, NFIB-S and NFIB-TAD overexpressing HEK293T cells -II. pCH: pCH_MCS which does not express any protein was used as mock. All constructs were encoded in the pCH backbone. 10×10^4 HEK293T cells were co-transfected with pGL2-*p21*-luc and pCH_MCS, pCH-GFP, pCH-mNfib3, pCH-mNfib4, pCH-NFIB-S, and pCH-NFIB-TAD expression constructs. The fold increases in Luc activity are relative to the pCH (empty vector) co-transfectants. Luciferase activity was measured in triplicate, and the results were averaged to determine the mean $\pm$ S.D. The Students t test was used for statistical analysis (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ **** $p < 0.0001$). Expression levels of NFI proteins in this experiment was determined by Western Blot (insert). (Nfib3: 49kDA, Nfib4: 56kDA, NFIB-S: 35kDA)

3.4 Binding of NFIB isoforms to *GFAP* and *p21* NFI sites

In both *GFAP* and *p21* promoter regulation assays, expression of NFIB-S appeared to repress the promoter of interest. Overexpression of NFIB-S decreased *GFAP* promoter activity by approximately 25%. On the other hand, we observed a modest effect on the *p21* promoter, of 6%. In order to find out if NFIB-S affected NFIB binding to the NFI motifs, we carried out *in vitro* DNA-binding assays. NFIB-S does not bear the conserved N-terminal DNA-binding and dimerization domain, nevertheless we examined whether it could interact with the consensus NFI-binding sequence on *GFAP* or *p21* promoters via another region on the protein.

3.4.1 Binding of NFIB isoforms to NFI sites on *GFAP*

Full length NFIB (NFIB-FL) bears the N terminal DNA-binding and dimerization domain which allows it to bind the consensus DNA-binding sequence on the *GFAP* promoter region and regulate *GFAP* expression (Brun et al., 2009). As truncated NFIB (NFIB-S) lacks this domain, we wondered whether it can interact with the NFI consensus site on *GFAP* to affect the activation of this promoter. In order to explore this possibility, a gel shift experiment was performed by using one of the NFIB full length isoforms, mNfib3 (will be referred to as NFIB-FL) as positive control and NFIB-TAD and untransfected HEK293T lysates as negative controls. NFIB-FL, NFIB-S, and NFIB-TAD were overexpressed in HEK293T cells and the lysates were analyzed by Western blot for an approximate quantification (Figure 3.15, insert). NFIB overexpressing lysates were diluted to approximately similar amounts using untreated HEK293T lysate. While NFIB-FL could bind to biotinylated double stranded *GFAP* oligonucleotides, neither NFIB-S (NFIB4) nor NFIB-TAD produced a shift, as expected (Figure 3.15).

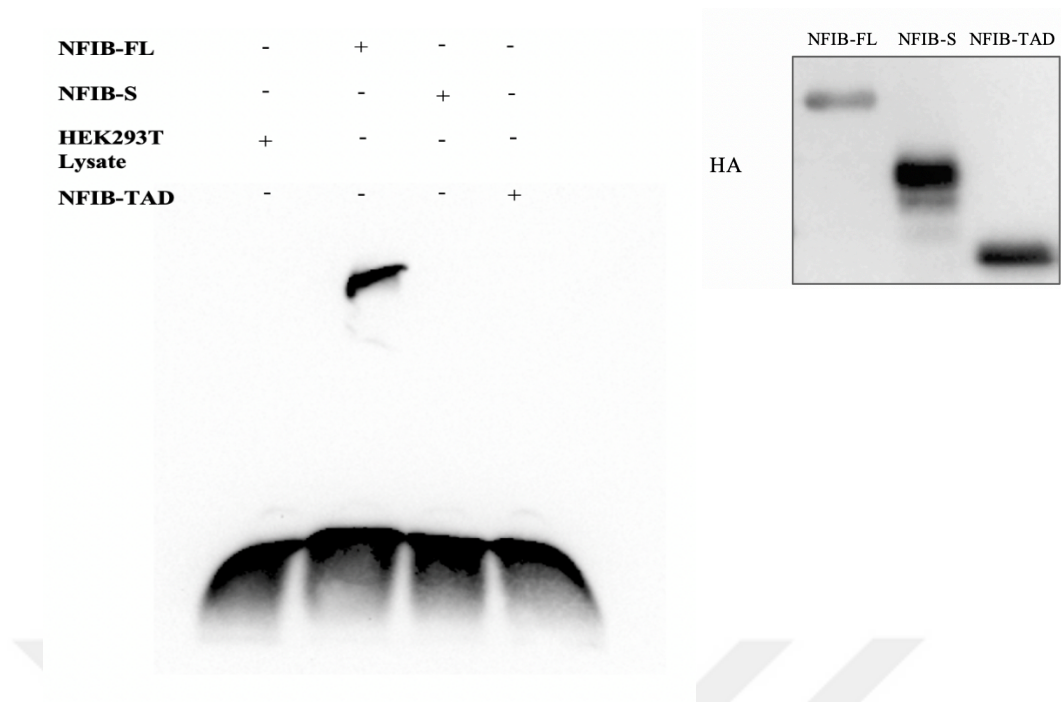


Figure 3.15: Gel shift assays showing binding of NFIB isoforms to *GFAP* NFI sites. NFIB-FL: Full length NFIB; NFIB-S: Short NFIB or NFIB4; NFIB-TAD: Transactivation domain of NFIB, truncated, HEK293T lysate: untransfected HEK293T cell lysate. Gel shift experiments were carried out with biotinylated *GFAP* double stranded oligonucleotides and HEK293T cell lysates. DNA binding reactions were electrophoresed in a 5% polyacrylamide gel in 0.5% TBE buffer to separate unbound DNA and DNA-protein complexes. Expression levels of NFI proteins in this experiment was determined by Western Blot (insert). (NFIB-FL: 49kDA, NFIB-S: 35kDA, NFIBTAD: 27kDA)

3.4.2 Binding of NFIB isoforms to NFI sites on *p21*

Next, we tested whether NFIB-S can bind to the NFI motifs on the *p21* promoter region *in vitro* and to that end, biotinylated double stranded *p21* oligonucleotides were incubated with NFIB isoforms. Nucleotide sequence corresponds to upper strand of *p21* oligonucleotide used in these assay corresponds to the -167/-127 region of *p21* including the NFI site, 5'-TGGAAGCTCGGCCA-3'. This sequence was shown to be shifted when incubated with NFIB before (Ouellet et al., 2006). While NFIB could bind to biotinylated double stranded *p21* oligonucleotides, neither NFIB-S nor NFIB-TAD produced a shift in *p21* oligonucleotide mobility (Figure 3.16). Since NFIB-S lacks a DNA-binding and dimerization domain, if it has transcriptional regulatory roles, it is likely to interfere with NFIB-FL (or other NFIs) binding partners or it may somehow reduce NFIB-FL repression of target genes by blocking NFIB-FL's repressive actions. To address these possibilities, we determined whether NFIB-S can reduce the NFIB-FL binding to target DNA *in vitro*: when NFIB-FL and NFIB-S were

incubated *in vitro*, and then allowed to bind the p21 oligonucleotides, the density of the shifted band decreased compared to that of NFIB-FL alone (Figure 3.16). The assay was repeated, to verify this result. NFIB-N (truncated N-terminal domain of NFIB) was also included as positive control. Additionally, varying amounts of the two isoforms were used in co-incubation. Two NFIB-FL:NFIB-S ratios (1:1 and 1:3) were used in *in vitro* co-incubation experiments (Figure 3.17). Nevertheless, there was no significant difference between the densities of the shifted bands.

Taken together, these data indicate that NFIB-S does not interact with NFI sites with a significant affinity. Furthermore, under these *in vitro* conditions, NFIB-S does not seem to affect NFIB-FL's interaction with NFI sites. Further experiments are needed to explore if NFIB-S does affect NFI function *in vivo*.

NFIB-FL	+	+	-	-	+
NFIB-S	-	+	-	+	-
HEK293T Lysate	-	-	+	+	+
NFIB-TAD	+	-	+	-	-

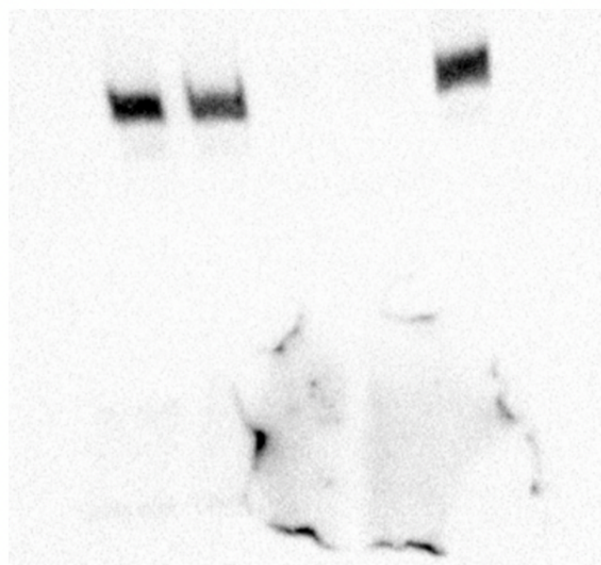


Figure 3.16: Gel shift assays showing binding of NFIB isoforms to p21 NFI sites-I.

NFIB-FL: Full length NFIB; NFIB-S: Short NFIB or NFIB4; NFIB-TAD: Transactivation domain of NFIB, truncated, HEK293T lysate: untransfected HEK293T cell lysate. Gel shift experiments were carried out with biotinylated p21 double stranded oligonucleotides and HEK293T cell lysates. DNA binding reactions were electrophoresed in a 5% polyacrylamide gel in 0.5% TBE buffer to separate unbound DNA and DNA-protein complexes. For co-incubation reactions, pre-

incubation was achieved at room temperature for 10 minutes followed by 20 minutes incubation for DNA-binding reaction.

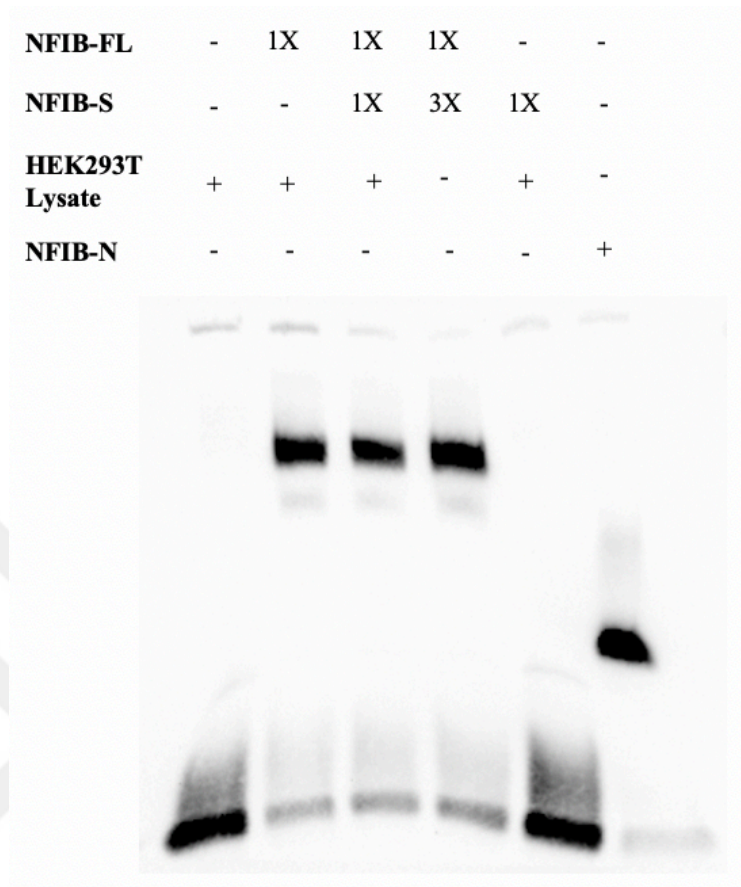


Figure 3.17: Gel shift assays showing binding of NFIB isoforms to *p21* NFI sites-II. NFIB-FL: Full length NFIB; NFIB-S: Short NFIB or NFIB4; NFIB-TAD: Transactivation domain of NFIB, truncated, HEK293T lysate: untransfected HEK293T cell lysate. Gel shift experiments were carried out with biotinylated *p21* double stranded oligonucleotides and HEK293T cell lysates. DNA binding reactions were electrophoresed in a 5% polyacrylamide gel in 0.5% TBE buffer to separate unbound DNA and DNA-protein complexes. For co-incubation reactions, pre-incubation was achieved at room temperature for 10 minutes followed by 20 minutes incubation for DNA-binding reaction.



4. DISCUSSION

NFIB belongs to the NFI family of transcription factors that regulates diverse cellular events such as neural stem cell differentiation during CNS development, proliferation and differentiation during lung maturation or melanocyte stem cell quiescence in the hair follicle niche. NFIB can also act as a tumor suppressor in brain cancer while acting as an oncogene in others, eg. small cell lung cancer, melanoma (Semenova et al., 2016). Molecular mechanisms that regulate NFIB function have not yet been elucidated. Moreover, while NFIB has many alternative splicing isoforms, the potential functional differences among those isoforms is unknown. NFIB isoforms may indeed function differently due to the inclusion of different exon combinations arising from the alternative splicing. Hypothesizing that exons 9 and 10 may lead to functional differences between NFIB isoforms, we tested whether mNfib3 and mNfib4 differed in their activation of two NFI target genes, *GFAP* and *p21*. While mNfib4 induced *GFAP* promoter activity by 2.5-3 fold, mNfib3 activated the promoter by 1.7-1.8 fold in three independent experiments. Thus, mNfib4 appeared to induce *GFAP* promoter activity more efficiently than mNfib3. *p21* promoter activity was reduced in the presence of mNfib4 by 45-67%, while mNfib3 overexpression decreased the activity by 37-44%. Hence, mNfib4 repressed *p21* promoter activity more potently than mNfib3. We also included an unusual NFIB isoform, NFIB-S (NFIB4) and characterized its expression, DNA binding and transcription regulation capacity, using the same target genes. NFIB-S is expressed in both neuroblastoma (SH-SY5Y) and glioma (U87MG and T98G) cell lines; however, the highest expression was seen in SH-SY5Y cells. Interestingly, *GFAP* promoter activity decreased in NFIB-S expressing cells compared to pCH mock, with a modest decrease compared to additional GFP control. On the other hand, NFIB-S had no effect on *p21* promoter activity compared to the GFP control. NFIB-S does not carry a N-terminal DBD domain, nevertheless, we tested whether it could bind to consensus NFI sites *in vitro* or interfere with binding of full length NFIB. Mobility shift assays indicated that biotinylated double stranded oligonucleotides corresponding to NFI sites on the *GFAP*

or *p21* promoter were not bound by NFIB-S. Furthermore, co-incubation of NFIB-S with full length NFIB did not alter NFIB binding to the DNA.

Interestingly, in NFIX-X5, exons 9 -11 encode a longer C-terminal transactivation domain that provides this NFIX isoform with important functional capabilities (Singh et al., 2011). Singh et al. found that the long NFIX isoform (NFIX-X5) with the distinctive transactivation domain activates *GFAP* promoter in astrocytes and glioblastoma more effectively than the short isoform (NFIX2). However, these isoforms did not differ in repression of the *p21* promoter. This result suggests that activation and repression occurs through distinct mechanisms, where the distinct transactivation domain of the long isoform is involved in activation of the *GFAP* promoter. Similarly, for NFIB, C-terminal domain of mNfib4 bears the exons 9 and 10, while the isoform used in most gain of function experiments, mNfib3, lacks these exons. These mouse isoforms correspond to NFIB1 and NFIB3 in humans. Since the mouse and human orthologs are ~99% identical (98.9% mouse Nfib4 with human NFIB1 and 99.3% mouse Nfib3 with human NFIB3), investigation of the functional significance of domains encoded by exons 9 and 10 would be relevant to the human orthologs, NFIB1 and NFIB3, as well. The inclusion of exons 9 and 10 also provides the long isoform, mNfib4, with a distinct transactivation domain which could lead to distinct functions. We wanted to find out if in fact, the long isoform differs functionally. While the long isoform, mNfib4, showed higher activation of *GFAP* promoter in all three independent experiments, in fact the higher activation potency of Nfib4 correlated with its expression levels. We tried to adjust expression levels by decreasing the amount of transfected plasmid, however mNfib4 levels remained higher than that of mNfib3. Indeed, promoter activation levels could be similar if both proteins were expressed at similar levels. When we compare the repression potencies of the two isoforms on *p21* promoter, though mNfib4 reduced the activity more potently than mNfib3; the higher expression levels of mNfib4 still correlated with its higher repression potency.

The differential expression of mNfib3 and mNfib4 may arise from differences in their stability. The special transactivation domain of mNfib4 at its end may give rise to a more stable mRNA/protein or more mNfib4 might be translated into protein. The difference of mNfib4 from mNfib3 is additional 79 amino acids encoded by exons 9 and 10. There are additional 10 prolines in mNfib4 and these proline residues may

provide the protein with additional functions or regulatory features since the proline-rich region in the NFI C-terminal transactivation domain can directly interact with H3 histone protein and/or modify nucleosome structure affecting gene expression (Gerber et al., 1995). Singh et al. (2011) reported that NFIX2 and NFIX-X5 isoforms differ at their C-terminus in 84 amino acids, 17 of which are proline residues. Further enrichment of proline residues were seen as a possible reason for the long isoform to have enhanced *GFAP* activation capacity which may be due to more efficient modification of the nucleosome through histone interactions.

In order to determine whether mNfib4 is more potent than mNfib3 in *GFAP* promoter activation or *p21* promoter repression, promoter activity should be measured in cells overexpressing the isoforms at equal levels. Even though our attempt for matching the isoform expression levels was not successful, additional approaches can be taken to achieve this. For example, cells could be transfected with serial dilutions of expression constructs to achieve similar expression levels. An alternative approach is to measure induction of endogenous *GFAP* expression in a neural cell line, such as glioma cells, by qRT-PCR where overexpression levels can also be quantitated and normalized to determine fold change in *GFAP* mRNA levels. Another important point is the choice of the appropriate cell line for functional assays since the protein composition of the cell is inevitably important for transcriptional regulation. If the distinct transactivation domain of mNfib4 functions via interactions with astrocyte-specific factors for example, then astrocytes may be the correct cell type to study mNfib4. Yet another complementary approach is construction of a fusion protein bearing the distinct transactivation domain of the mNfib4 fused to an independent DNA binding domain (Singh et al., 2011). Thus, only the effect derived from the particular mNfib4 transactivation domain could be determined.

Another unusual NFIB splicing isoform (NFIB-S), which lacks the N-terminal DNA binding and dimerization domain, exists only in humans, the third NFIB isoform that we selected for further study. First discovered in an RNA-seq study on human blood progenitors, NFIB-S was shown to direct megakaryocyte differentiation when co-expressed with NFIC (Chen et al., 2014). Importantly, when full-length NFIB was co-expressed with NFIC, megakaryocyte differentiation did not result in blood progenitor population. Thus, during megakaryopoiesis, NFIB-S has a unique function in megakaryocyte differentiation which full length NFIB could not carry out.

When we started this study, there were four annotated human NFIB alternative splice variants on the NCBI database. Of these isoforms, two were nearly identical and showed ~99% homology with mNfib3 (translation start site differs in NFIB2 and NFIB3 and as a result there are 25 additional amino acids at the very N-terminus of the NFIB2, but this region is not thought to render the protein any function since it is outside the DBD domain and the transactivation domain). NFIB1 is the longest human isoform and exhibits ~99% homology with mNfib4. The final isoform, NFIB4 (or NFIB-S as referred to in this study) does not carry the conserved N-terminal DBD domain and has a different C-terminal end due to an alternative translation start site within exon 5 and the frameshift arising from the inclusion of exon 11. However, as of May 2019, the number of accepted NFIB isoforms has been updated to as many as 28. These include 21 full length isoforms, 5 isoforms with partial lack in N-terminal DBD, 1 truncated isoform with only N-terminal DBD without a C-terminal TAD which is probably a dominant negative isoform, and finally NFIB-S as a truncated isoform lacking in the whole N terminal DBD. Nevertheless, regardless of the diversity of NFIB alternative splicing isoforms, NFIB-S is a unique, unusual isoform missing the conserved N-terminal DNA-binding and dimerization domain. This brings the question, how does a transcription factor function without a DNA-binding domain?

When we tested the ability of NFIB-S to activate the *GFAP* promoter, we were surprised to find that NFIB-S repressed *GFAP* by 77% compared to pCH mock control. In the following experiments, pCH_GFP was also used as a control since exogenous protein translation may negatively affect the cells. NFIB-S exhibited a modest decrease in *GFAP* promoter activation by ~25% compared to the GFP control. In contrast, NFIB-S and GFP reduced *p21* promoter activity equally; thus, NFIB-S had no effect on *p21* activity compared to GFP. Thus, while the choice of the correct control is the key prerequisite to interpret the data, further experiments are needed to obtain reproducible and reliable results.

In vitro binding assays showed that NFIB-S does not interact with NFI-binding sites on the *GFAP* and *p21* promoters while full length NFIB was able to shift the biotinylated bands. One possibility is that while NFIB-S does not bind DNA itself, it can affect binding of NFI proteins to the NFI sites. Co-incubation experiments where full length NFIB was incubated with truncated NFIB constructs, NFIB-S and NFIB-TAD, showed no difference in binding. However, these experiments may be not

sufficiently sensitive to detect low affinity interactions that may occur *in vivo*. For *in vivo* investigation of NFIB-S effect on full length NFIB binding to the consensus sequences, the isoforms can be co-transfected in cells, instead of co-incubation *in vitro*. If NFIB-S indirectly interact with NFIB, for example through adapter proteins, and then these interaction could occur *in vivo* but disrupted by lysis of the proteins before *in vitro* binding experiments. Co-transfected cells overexpressing both isoforms could be included in luciferase reporter assays also, in order to detect a possible change in promoter activity in co-transfected cells, compared to the cells that overexpresses NFIB full length. As before, the choice of correct cell line is critical since transcription factors regulate target gene expression through interacting with other factors. The molecular composition of the cell could determine the function of isoforms. NFIB-S may interact with the local cofactors needed for NFIB binding to DNA, and a possible competition for those cofactors could affect DNA binding of full length NFIB. Alternatively, chromatin immunoprecipitation assays may reveal the changes in DNA binding of NFIB *in vivo*. NFIB-S could be overexpressed in a cell line with endogenous NFIB expression, such as SH-SY5Y cells, and then by immunoprecipitating NFIB, bound NFI-sites could be quantified in control and NFIB-S overexpressing cells. If NFIB-S interferes with binding of full length NFIB to DNA, the amplified NFI site DNA would be lower in NFIB-S expressing cells than that of untreated cells.

It is important to emphasize that NFIB-S was found to be important in megakaryocyte differentiation as explained before (Chen et al., 2014). We were able to detect its expression in neuroblastoma (SH-SY5Y) and glioma cell lines (U87MG and T98G); however, it may not play role in transcriptional regulation of *GFAP* or *p21* in these cells as we expected. Instead, transcriptional profiling would be more beneficial to reveal the effect of NFIB-S on mRNA composition of the cells which already expresses NFIB-S endogenously. NFIB-S upregulation by overexpression in these cells would exhibit the changes in transcript level of differentiation markers. Moreover, NFIB-S could be downregulated by RNAi knock down in order to detect the differentiation marker gene transcripts. Thus, if NFIB-S play roles in transcriptional regulation, transcriptional profiling of cells expressing or lacking NFIB-S would be informative.

Finally, throughout this study we observed that mNfib4 and NFIB-S, which both bear exons 9 and 10, were expressed at higher levels with respect to mNfib3 which lacks both exons. Moreover, NFIB-S, carries exon 11 as well, was expressed at the highest levels, followed by mNfib4 and then mNfib3. These exons may affect the stability of the transcripts at the mRNA level. Moreover, the structure of the C-terminal domain of each isoform may be control the post-translational modification of the protein. Post-translational modifications may affect protein stability, which would indirectly regulate transcriptional output from these isoforms.

Although most of the human multi-exon genes are estimated to be alternatively spliced (Mollet et al., 2010); the number of isoforms encountered at protein level is not as much as it is expected based on the RNA-seq studies. Instead, it is claimed that the isoforms with smaller changes in structural and functional domains would have more opportunity to be functionally important. The isoforms with huge differences in structural and functional domains seem less likely to have functional importance for the cell (Tress et al., 2017). Therefore, since NFIB-S lacks very conserved N-terminal DBD domain, it may not be included in transcriptional regulation directly. However, due to the inclusion of transactivation domain which is important in establishing protein protein interaction (PPI) networks with other cofactors, NFIB-S still has the potential to affect transcription of target genes indirectly.

To conclude, NFIB is an important transcription factor that it regulates stem cell differentiation or maintenance both during development and in the adult; moreover, its misregulation is seen in various types of cancer inclufind brain and lung tumors. Additionally, in humans, mutations in even one copy of *NFIB* gene results in CNS phenotypes, such as intellectual disability. Thus, elucidation of molecular mechanisms that regulate NFIB function is critical to understand mechanisms of brain development and pathologies.

REFERENCES

- Abbas, T., & Dutta, A.** (2009). p21 in cancer: intricate networks and multiple activities. *Nature Reviews Cancer*, 9(6), 400.
- Alevizopoulos, A., Dusserre, Y., Tsai-Pflugfelder, M., Von Der Weid, T., Wahli, W., & Mermoud, N.** (1995). A proline-rich TGF-beta-responsive transcriptional activator interacts with histone H3. *Genes & development*, 9(24), 3051-3066.
- Altmann, H., Wendler, W., & Winnacker, E. L.** (1994). Transcriptional activation by CTF proteins is mediated by a bipartite low-proline domain. *Proceedings of the National Academy of Sciences*, 91(9), 3901-3905.
- Armentero, M. T., Horwitz, M., & Mermoud, N.** (1994). Targeting of DNA polymerase to the adenovirus origin of DNA replication by interaction with nuclear factor I. *Proceedings of the National Academy of Sciences*, 91(24), 11537-11541.
- Attardi, L. D., & Tjian, R.** (1993). Drosophila tissue-specific transcription factor NTF-1 contains a novel isoleucine-rich activation motif. *Genes & Development*, 7(7b), 1341-1353.
- Bandyopadhyay, S. & Gronostajski, R. M.** (1994). Identification of a conserved oxidation-sensitive cysteine residue in the NFI family of DNA-binding proteins. *J. Biol. Chem.* 269, 29949–29955.
- Bandyopadhyay, S., & Gronostajski, R. M.** (1994). Identification of a conserved oxidation-sensitive cysteine residue in the NFI family of DNA-binding proteins. *Journal of Biological Chemistry*, 269(47), 29949-29955.
- Barry, G., Piper, M., Lindwall, C., Moldrich, R., Mason, S., Little, E., Sarkar, A., Tole, S., Gronostajski, R. M., & Richards, L. J.** (2008). Specific glial populations regulate hippocampal morphogenesis. *Journal of Neuroscience*, 28, 12328-12340.
- Barry, G., Piper, M., Lindwall, C., Moldrich, R., Mason, S., Little, E., ... & Richards, L. J.** (2008). Specific glial populations regulate hippocampal morphogenesis. *Journal of Neuroscience*, 28(47), 12328-12340.
- Behzadnia, N., Golas, M. M., Hartmuth, K., Sander, B., Kastner, B., Deckert, J., ... & Luhrmann, R.** (2007). Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. *The EMBO journal*, 26(6), 1737-1748.
- Berget, S. M.** (1995). Exon recognition in vertebrate splicing. *Journal of biological Chemistry*, 270(6), 2411-2414.

- Chang, C. Y., Pasolli, H. A., Giannopoulou, E. G., Guasch, G., Gronostajski, R. M., Elemento, O., & Fuchs, E.** (2013). NFIB is a governor of epithelial–melanocyte stem cell behaviour in a shared niche. *Nature*, 495(7439), 98.
- Chaudhry, A. Z., Vitullo, A. D., & Gronostajski, R. M.** (1998). Nuclear factor I (NFI) isoforms differentially activate simple versus complex NFI-responsive promoters. *Journal of Biological Chemistry*, 273(29), 18538-18546.
- Chow, L. T., Gelinas, R. E., Broker, T. R., & Roberts, R. J.** (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*, 12(1), 1-8.
- das Neves, L. et al.** (1999). Disruption of the murine nuclear factor I-A gene (Nfia) results in perinatal lethality, hydrocephalus, and agenesis of the corpus callosum, *PNAS*, 96: 11946–11951.
- Dooley, A.L., Winslow, M.M., Chiang, D.Y., Banerji, S., Stransky, N., Dayton, T.L., Snyder, E.L., Senna, S., Whittaker, C.A., Bronson, R.T., et al.** (2011). Nuclear factor I/B is an oncogene in small cell lung cancer, *Genes Dev.*, 25, 1470–1475.
- Dusserre, Y., & Mermoud, N.** (1992). Purified cofactors and histone H1 mediate transcriptional regulation by CTF/NF-I. *Molecular and cellular biology*, 12(11), 5228-5237.
- Dynlacht, B. D., Hoey, T., & Tjian, R.** (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell*, 66(3), 563-576.
- Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R., & Hood, L.** (1980). Two mRNAs can be produced from a single immunoglobulin μ gene by alternative RNA processing pathways. *Cell*, 20(2), 313-319.
- Gao, B., Jaffe, H., & Kunos, G.** (1998). Histone H1 isoforms purified from rat liver bind nonspecifically to the nuclear factor 1 recognition sequence and serve as generalized transcriptional repressors. *Molecular and cellular biochemistry*, 178(1-2), 187-196.
- Gartel, A. L., Serfas, M. S., & Tyner, A. L.** (1996). p21—negative regulator of the cell cycle. *Proceedings of the Society for Experimental Biology and Medicine*, 213(2), 138-149.
- Gobius, I., Morcom, L., Suárez, R., Bunt, J., Bukshpun, P., Reardon, W., ... & Richards, L. J.** (2016). Astroglial-mediated remodeling of the interhemispheric midline is required for the formation of the corpus callosum. *Cell reports*, 17(3), 735-747.
- Gopalan, S. M., Wilczynska, K. M., Konik, B. S., Bryan, L., & Kordula, T.** (2006). Nuclear factor-1-X regulates astrocyte-specific expression of the α 1-antichymotrypsin and glial fibrillary acidic protein genes. *Journal of Biological Chemistry*, 281(19), 13126-13133.
- Gounari, F., De Francesco, R., Schmitt, J., Van der Vliet, P., Cortese, R., & Stunnenberg, H.** (1990). Amino-terminal domain of NF1 binds to

DNA as a dimer and activates adenovirus DNA replication. *The EMBO Journal*, 9(2), 559-566.

- Goyal, N., Knox, J., & Gronostajski, R. M.** (1990). Analysis of multiple forms of nuclear factor I in human and murine cell lines. *Molecular and cellular biology*, 10(3), 1041-1048.
- Grishin, N. V.** (2001). Mh1 domain of Smad is a degraded homing endonuclease. *Journal of molecular biology*, 307(1), 31-37.
- Gronostajski, R. M.** (2000). Roles of NFI/CTF gene family in transcription and development. *Gene*, 249: 31-45.
- Gupta, C. N., Chen, J., Liu, J., Damaraju, E., Wright, C., Perrone-Bizzozero, N. I., ... & Calhoun, V. D.** (2015). Genetic markers of white matter integrity in schizophrenia revealed by parallel ICA. *Frontiers in human neuroscience*, 9, 100.
- Harris, L., Genovesi, L. A., Gronostajski, R. M., Wainwright, B. J., & Piper, M.** (2014). Nuclear factor one transcription factors: divergent functions in developmental versus adult stem cell populations. *Developmental Dynamics*. doi: 10.1002/DVDY.24182.
- Hsu, Y.C., Osinski, J., Campbell, C.E., Litwack, E.D., Wang, D., Liu, S., Bachurski, C.J., & Gronostajski, R.M.** (2011). Mesenchymal nuclear factor I B regulates cell proliferation and epithelial differentiation during lung maturation, *Dev. Biol.*, 354, 242–252.
- Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., ... & Kendall, J.** (2012). De novo gene disruptions in children on the autistic spectrum. *Neuron*, 74(2), 285-299.
- Kim, T. K., & Roeder, R. G.** (1994). Proline-rich activator CTF1 targets the TFIIB assembly step during transcriptional activation. *Proceedings of the National Academy of Sciences*, 91(10), 4170-4174.
- Krohn, K., Rozovsky, I., Wals, P., Teter, B., Anderson, C. P., & Finch, C. E.** (1999). Glial fibrillary acidic protein transcription responses to transforming growth factor- β 1 and interleukin-1 β are mediated by a nuclear factor-1-like site in the near-upstream promoter. *Journal of neurochemistry*, 72(4), 1353-1361.
- Kruse, U., & Sippel, A. E.** (1994). The genes for transcription factor nuclear factor I give rise to corresponding splice variants between vertebrate species. *Journal of molecular biology*, 238(5), 860-865.
- Kruse, U., & Sippel, A. E.** (1994). Transcription factor nuclear factor I proteins form stable homo- and heterodimers. *FEBS letters*, 348(1), 46-50.
- Kruse, U., F. Qian, and A. E. Sippel.** (1991). Identification of a fourth nuclear factor I gene in chicken by cDNA cloning: NFI-X, *Nucleic Acids Res.* 19: 6641.
- Kumbasar, A., Plachez, C., Gronostajski, R. M., Richards, L. J., and Litwack, E. L.** (2009). Absence of the transcription factor Nfib delays the formation of the basilar pontine and other mossy fiber nuclei, *The Journal of Comparative Neurology*, 513: 98–112.

- Le-Niculescu, H., Patel, S. D., Bhat, M., Kuczenski, R., Faraone, S. V., Tsuang, M. T., ... & Niculescu Iii, A. B.** (2009). Convergent functional genomics of genome-wide association data for bipolar disorder: Comprehensive identification of candidate genes, pathways and mechanisms. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 150(2), 155-181.
- Lin, Y. S., & Green, M. R.** (1991). Mechanism of action of an acidic transcriptional activator in vitro. *Cell*, 64(5), 971-981.
- Lundqvist, J., EL Andaloussi-Lilja, J., Svensso, C., Dorfh, H. G., & Forsby, A.** (2013). Optimisation of culture conditions for differentiation of C17.2 neural stem cells to be used for in vitro toxicity tests, *Toxicology in vitro*, 27: 1565-1569.
- Mermod, N. O., Neil, E., Kelly, T., & Tjian, R.** (1989). The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain, *Cell*, 58: 741-753.
- Mermod, N., O'Neill, E. A., Kelly, T. J., & Tjian, R.** (1989). The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. *Cell*, 58(4), 741-753.
- Michael, D., & Manyuan, L.** (1999). Intron—exon structures of eukaryotic model organisms. *Nucleic acids research*, 27(15), 3219-3228.
- Mitchell, P. J., & Tjian, R.** (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science*, 245(4916), 371-378.
- Miura, M., Tamura, T. A., & Mikoshiba, K.** (1990). Cell-specific expression of the mouse glial fibrillary acidic protein gene: identification of the cis- and trans-acting promoter elements for astrocyte-specific expression. *Journal of neurochemistry*, 55(4), 1180-1188.
- Mollet, I. G., Ben-Dov, C., Felício-Silva, D., Grosso, A. R., Eleuterio, P., Alves, R., ... & Carmo-Fonseca, M.** (2010). Unconstrained mining of transcript data reveals increased alternative splicing complexity in the human transcriptome. *Nucleic acids research*, 38(14), 4740-4754.
- Nehls, M. C., Rippe, R. A., Veloz, L., & Brenner, D. A.** (1991). Transcription factors nuclear factor I and Sp1 interact with the murine collagen alpha 1 (I) promoter. *Molecular and cellular biology*, 11(8), 4065-4073.
- Nishimura M, Isaka F, Ishibashi M, Tomita K, Tsuda H, Nakanishi S, & Kageyama R.** (1998). Structure, chromosomal locus, and promoter of mouse Hes2 gene, a homologue of Drosophila hairy and Enhancer of split. *Genomics*. 49:69-75.
- Orford KW, Scadden DT.** (2008). Deconstructing stem cell self-renewal: Genetic insights into cell-cycle regulation. *Nat Rev Genet* 9: 115-128.
- Ostrovsky, O., & Bengal, E.** (2003). The mitogen-activated protein kinase cascade promotes myoblast cell survival by stabilizing the cyclin-dependent kinase inhibitor, p21WAF1 protein. *Journal of Biological Chemistry*, 278(23), 21221-21231.

- Palacios D. & Puri P.L.** (2010). Switch NFIB developmental myogenesis. *Developmental Cell*, 18, 340–1.
- Piper, M., Barry, G., Harvey, T. J., McLeay, R., Smith, A. G., Harris, L., ... & Gronostajski, R. M.** (2014). NFIB-mediated repression of the epigenetic factor Ezh2 regulates cortical development. *Journal of Neuroscience*, 34(8), 2921-2930.
- Piper, M., Moldrich, R. X., Lindwall, C., Little, E., Barry, G., Mason, S., ... & Richards, L. J.** (2009). Multiple non-cell-autonomous defects underlie neocortical callosal dysgenesis in Nfib-deficient mice. *Neural development*, 4(1), 43.
- Prasad, A., Merico, D., Thiruvahindrapuram, B., Wei, J., Lionel, A. C., Sato, D., ... & Fernandez, B. A.** (2012). A discovery resource of rare copy number variations in individuals with autism spectrum disorder. *G3: Genes, Genomes, Genetics*, 2(12), 1665-1685.
- Ptashne, M.** (1988). How eukaryotic transcriptional activators work. *Nature*, 335(6192), 683.
- Redin, C., Brand, H., Collins, R. L., Kammin, T., Mitchell, E., Hodge, J. C., ... & Abdul-Rahman, O. A.** (2017). The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. *Nature genetics*, 49(1), 36.
- Resch, A., Xing, Y., Modrek, B., Gorlick, M., Riley, R., & Lee, C.** (2004). Assessing the impact of alternative splicing on domain interactions in the human proteome. *Journal of proteome research*, 3(1), 76-83.
- Ristiniemi, J., & Oikarinen, J.** (1989). Histone H1 binds to the putative nuclear factor I recognition sequence in the mouse alpha 2 (I) collagen promoter. *Journal of Biological Chemistry*, 264(4), 2164-2174.
- Roulet, E., Armentero, M. T., Krey, G., Corthesy, B., Dreyer, C., Mermoud, N., & Wahli, W.** (1995). Regulation of the DNA-binding and transcriptional activities of *Xenopus laevis* NFI-X by a novel C-terminal domain, *Mol. Cell. Biol.*, 15: 5552–5562.
- Rupp, R. A. W., U. Kruse, G. Multhaup, U. Goebel, K. Beyreuther, & A. E. Sippel.** (1990). Chicken NFI/TGGCA proteins are encoded by at least three independent genes: NFI-A, NFI-B and NFI-C with homolog in mammalian genomes. *Nucleic Acids Res.* 18: 2607–2616.
- Segurado, R., Detera-Wadleigh, S. D., Levinson, D. F., Lewis, C. M., Gill, M., Nurnberger Jr, J. I., ... & Ekholm, J.** (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. *The American Journal of Human Genetics*, 73(1), 49-62.
- Semenova, E. A., et al.** (2016). Transcription factor NFIB is a driver of small cell lung cancer progression in mice and marks metastatic disease in patients, *Cell Reports*, 16: 631–643.
- Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., ... & Soreq, H.** (2005). Function of alternative splicing. *Gene*, 344, 1-20.
- Steele-Perkins, G., Butz, K. G., Lyons, G. E., Zeichner-David, M., Kim, H. J., Cho, M. I., & Gronostajski, R. M.** (2003). Essential role for NFI-

C/CTF transcription-replication factor in tooth root development. *Molecular and cellular biology*, 23(3), 1075-1084.

Steele-Perkins, G., et al. (2004). The transcription factor gene Nfib is essential for both lung maturation and brain development, *Molecular and Cellular Biology*, 25: 685–698. doi:10.1128/MCB.25.2.685–698.2005.

Steele-Perkins, G., Plachez, C., Butz, K. G., Yang, G., Bachurski, C. J., Kinsman, S. L., ... & Gronostajski, R. M. (2005). The transcription factor gene Nfib is essential for both lung maturation and brain development. *Molecular and cellular biology*, 25(2), 685-698.

Steele-Perkins, G., Plachez, C., Butz, K.G., Yang, G., Bachurski, C.J., Kinsman, S.L., Litwack, E.D., Richards, L.J., & Gronostajski, R.M. (2005). The transcription factor gene Nfib is essential for both lung maturation and brain development, *Mol. Cell. Biol.*, 25, 685–698.

Suh H, Consiglio A, Ray J, Sawai T, D'Amour KA, Gage FH. (2007). In vivo fate analysis reveals the multipotent and self-renewal capacities of Sox2⁺ neural stem cells in the adult hippocampus. *Cell Stem Cell* 1:515–528.

Tanese, N., Pugh, B. F., & Tjian, R. (1991). Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. *Genes & Development*, 5(12a), 2212-2224.

Tanese, N., Pugh, B. F., & Tjian, R. (1991). Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. *Genes & Development*, 5(12a), 2212-2224.

Tress, M. L., Abascal, F., & Valencia, A. (2017). Alternative splicing may not be the key to proteome complexity. *Trends in biochemical sciences*, 42(2), 98-110.

Willour, V. L., Seifuddin, F., Mahon, P. B., Jancic, D., Pirooznia, M., Steele, J., ... & Perlis, R. H. (2012). A genome-wide association study of attempted suicide. *Molecular psychiatry*, 17(4), 433.

Wu, H., Wade, M., Krall, L., Grisham, J., Xiong, Y., & Van Dyke, T. (1996). Targeted in vivo expression of the cyclin-dependent kinase inhibitor p21 halts hepatocyte cell-cycle progression, postnatal liver development and regeneration. *Genes & development*, 10(3), 245-260.

Wu, H., Wade, M., Krall, L., Grisham, J., Xiong, Y., & Van Dyke, T. (1996). Targeted in vivo expression of the cyclin-dependent kinase inhibitor p21 halts hepatocyte cell-cycle progression, postnatal liver development and regeneration. *Genes & development*, 10(3), 245-260.

Xiao, H., Lis, J. T., Xiao, H., Greenblatt, J., & Friesen, J. D. (1994). The upstream activator CTF/NF1 and RNA polymerase II share a common element involved in transcriptional activation. *Nucleic acids research*, 22(11), 1966-1973.

Xing, Y., Xu, Q., & Lee, C. (2003). Widespread production of novel soluble protein isoforms by alternative splicing removal of transmembrane anchoring domains. *FEBS letters*, 555(3), 572-578.

Yang, Z.Q., Imoto, I., Pimkhaokham, A., Shimada, Y., Sasaki, K., Oka, M., & Inazawa, J. (2001). A novel amplicon at 9p23 - 24 in squamous cell

carcinoma of the esophagus that lies proximal to GASC1 and harbors NFIB, *Jpn. J. Cancer Res.*, 92, 423–428.

Zhou, Y., Cai, H., Xu, Y., Sun, J., Lan, X., Lei, C., & Chen, H. (2014). Novel isoforms of the bovine Nuclear factor I/X (CCAAT-binding transcription factor) transcript products and their diverse expression profiles, *Animal Genetics*, 45, 581-584. doi: 10.1111/age.12177.





APPENDIX

TOPO_mNfib4 construct supposed to be

Sequence ID: Query_181677 Length: 861 Number of Matches: 1

Range 1: 1 to 861 [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1591 bits(861)	0.0	861/861(100%)	0/861(0%)	Plus/Minus
Query 103	GGTACCAGGACTGTTGCTGATGCAGGAAGGATGGGTCTCTTGGGCTTAGTCCCACATATC	162		
Sbjct 861	GGTACCAGGACTGTTGCTGATGCAGGAAGGATGGGTCTCTTGGGCTTAGTCCCACATATC	802		
Query 163	GATTGGCTTGAGATGTGCTGAGGCTGTGTAGGCTTCAGTGGATGTAGTGATGGGTTTAG	222		
Sbjct 801	GATTGGCTTGAGATGTGCTGAGGCTGTGTAGGCTTCAGTGGATGTAGTGATGGGTTTAG	742		
Query 223	TATCTGTCATGGTCAGGGTCACAGGTCGCACTGCACTGGGATGGGGAGAGGGTGCCAAGA	282		
Sbjct 741	TATCTGTCATGGTCAGGGTCACAGGTCGCACTGCACTGGGATGGGGAGAGGGTGCCAAGA	682		
Query 283	CAGGTGTGAAATGGCCAGGCACCTTCCCTACTACTTGACCACTGCTGTTAGGCTGGCTCG	342		
Sbjct 681	CAGGTGTGAAATGGCCAGGCACCTTCCCTACTACTTGACCACTGCTGTTAGGCTGGCTCG	622		
Query 343	TTTGAGGACTGGATGGGTGATAGGAAGGTACGTAGTTCTTCAGAGTATCCTGAGGATTCA	402		
Sbjct 621	TTTGAGGACTGGATGGGTGATAGGAAGGTACGTAGTTCTTCAGAGTATCCTGAGGATTCA	562		
Query 403	GGTGAGGAGGATATCTGATTGTTGGATGAGAGAAGTAGCTGGAAGGTGCCGAGGAAGGA	462		
Sbjct 561	GGTGAGGAGGATATCTGATTGTTGGATGAGAGAAGTAGCTGGAAGGTGCCGAGGAAGGA	502		
Query 463	TAGCTTGCCTCGGAAATGGCAACGGTGAGGGCGGAGGTGGAGTTGAGATTGAGATGACAC	522		
Sbjct 501	TAGCTTGCCTCGGAAATGGCAACGGTGAGGGCGGAGGTGGAGTTGAGATTGAGATGACAC	442		
Query 523	TGTGCGCGACTCCAGGTATTCGGGATGATGGTGCTGGGGGAAAGTGCTCAATCTTGGGG	582		
Sbjct 441	TGTGCGCGACTCCAGGTATTCGGGATGATGGTGCTGGGGGAAAGTGCTCAATCTTGGGG	382		
Query 583	AAGAATCCTGTGGAGATGTAGAGCTAAACAGTGGCTTCTCAGGCTTCTTCATTGTAGTTG	642		
Sbjct 381	AAGAATCCTGTGGAGATGTAGAGCTAAACAGTGGCTTCTCAGGCTTCTTCATTGTAGTTG	322		
Query 643	GAGAAGACATATCTTGATCTCGTTTCATGCCATGTTTCGACTTCCAGCAGCTGGTGAATTTG	702		
Sbjct 321	GAGAAGACATATCTTGATCTCGTTTCATGCCATGTTTCGACTTCCAGCAGCTGGTGAATTTG	262		
Query 703	GAGAGGGGTAAAAGTCTCCTGTAGGACTTGGCTCCATATTTTCATCTATAGATATAGTTT	762		
Sbjct 261	GAGAGGGGTAAAAGTCTCCTGTAGGACTTGGCTCCATATTTTCATCTATAGATATAGTTT	202		
Query 763	TGGGTCTTTTGCTGCTCGGTGGAGAAGACAGCGACCTCTGCAGGTTACACACAGAGTTCA	822		
Sbjct 201	TGGGTCTTTTGCTGCTCGGTGGAGAAGACAGCGACCTCTGCAGGTTACACACAGAGTTCA	142		
Query 823	TGTCATGATAGTATGGTTGGCTGGGAATTTCTCCGATTGGGAAGTTGACTCCAGTTCCCT	882		
Sbjct 141	TGTCATGATAGTATGGTTGGCTGGGAATTTCTCCGATTGGGAAGTTGACTCCAGTTCCCT	82		
Query 883	GGGTTATGGGTGTTCTGGATACTCTCACAAGCTCTGATACATTGAAGACTCCGGATTTTA	942		
Sbjct 81	GGGTTATGGGTGTTCTGGATACTCTCACAAGCTCTGATACATTGAAGACTCCGGATTTTA	22		
Query 943	CAAAGCTGTCCTCGAGGTACC	963		
Sbjct 21	CAAAGCTGTCCTCGAGGTACC	1		

TOPO_mNfib4 sequencing result aligned with the mNfib4-C termini being cloned,
m13r reverse primers

Score	Expect	Identities	Gaps	Strand
2032 bits(1100)	0.0	1239/1301(95%)	30/1301(2%)	Plus/Minus
Query 71	CTAGCCCAGATACCAGGACTGTTGCTGATGCAGGAAGGATGGGTCTCTTGGGCTTAGTCC	130		
Sbjct 1485	CTAGCCCAGATACCAGGACTGTTGCTGATGCAGGAAGGATGGGTCTCTTGGGCTTAGTCC	1426		
Query 131	CACATATCGATTGGCTTGAGATGTGCCTGAGGCTGTGTAGGCTTCAGTGGATGTAGTGAT	190		
Sbjct 1425	CACATATCGATTGGCTTGAGATGTGCCTGAGGCTGTGTAGGCTTCAGTGGATGTAGTGAT	1366		
Query 191	GGGTTTAGTATCTGTTCATGGTCAGGCTCACAGGTCGCACTGCACTGGGATGGGGAGAGGG	250		
Sbjct 1365	GGGTTTAGTATCTGTTCATGGTCAGGCTCACAGGTCGCACTGCACTGGGATGGGGAGAGGG	1306		
Query 251	TGCCAAGACAGGTGTGAAATGGCCAGGCACCTTTCCTACTACTTGACCACTGCTGTTAGG	310		
Sbjct 1305	TGCCAAGACAGGTGTGAAATGGCCAGGCACCTTTCCTACTACTTGACCACTGCTGTTAGG	1246		
Query 311	CTGGCTCGTTTGAGGACTGGATGGGTCATAAGAAGGTACGTAGTTCTTCAGAGTATCCTG	370		
Sbjct 1245	CTGGCTCGTTTGAGGACTGGATGGGTCATAAGAAGGTACGTAGTTCTTCAGAGTATCCTG	1186		
Query 371	AGGATTCAGGTGAGGAGGATATCTGATTGTTGGATGAGAGAAGTAGCTGGAAGGTGCCGG	430		
Sbjct 1185	AGGATTCAGGTGAGGAGGATATCTGATTGTTGGATGAGAGAAGTAGCTGGAAGGTGCCGG	1126		
Query 431	AGGAAGGATAGCTTGCCTCGGAAATGGCAACGGTGAGGGCGGAGGTGGAGTTCGAGTTGA	490		
Sbjct 1125	AGGAAGGATAGCTTGCCTCGGAAATGGCAACGGTGAGGGCGGAGGTGGAGTTCGAGTTGA	1066		
Query 491	GATGACACTGTGCGCGACTCCAGGTATTCCGGGATGATGGTGCTGGGGGAAAGTGCTCAA	550		
Sbjct 1065	GATGACACTGTGCGCGACTCCAGGTATTCCGGGATGATGGTGCTGGGGGAAAGTGCTCAA	1006		
Query 551	TCTTGGGGGAAAGATCCTGTGGAGATGTAGAGCTAAACAGTGGCTTCTCAGGCTTCTTCAT	610		
Sbjct 1005	TCTTGGGGGAAAGATCCTGTGGAGATGTAGAGCTAAACAGTGGCTTCTCAGGCTTCTTCAT	946		
Query 611	TGTAGTTGGAGAAGACATATCTTGATCTCGTTTCATGCCATGTTTGGACTTCCAGCAGCTGG	670		
Sbjct 945	TGTAGTTGGAGAAGACATATCTTGATCTCGTTTCATGCCATGTTTGGACTTCCAGCAGCTGG	886		
Query 671	TGAATTTGGAGAGGGGTAAAGTCTCCTGTAGGACTTGGCTCCATATTTTCATCTATAGA	730		
Sbjct 885	TGAATTTGGAGAGGGGTAAAGTCTCCTGTAGGACTTGGCTCCATATTTTCATCTATAGA	826		
Query 731	TATAGTTTGGGCTCTTTGCTGCTCGGTGGAGAAGACAGCGACCTCTGCAGGTTACACCC	790		
Sbjct 825	TATAGTTTGGGCTCTTTGCTGCTCGGTGGAGAAGACAGCGACCTCTGCAGGTTACACCC	766		
Query 791	AGAGTTCATGTCATGATAGTATGGTTGGCTGGGAATTTCTCCGATTGGGAAGTTGACTCC	850		
Sbjct 765	AGAGTTCATGTCATGATAGTATGGTTGGCTGGGAATTTCTCCGATTGGGAAGTTGACTCC	706		
Query 851	AGTTCCTCGGGTTATGGGTGTTCTGGATACTCTCACAAGCTCTGATACATTGAAGACTCC	910		
Sbjct 705	AGTTCCTCGGGTTATGGGTGTTCTGGATACTCTCACAAGCTCTGATACATTGAAGACTCC	646		
Query 911	GGATTTTACAAAGCTGTCTCGAGGTACCTGGGAGGATTCTTGGCAGGATCACTGTGGC	970		
Sbjct 645	GGATTTTACAAAGCTGTCTCGAGGTACCTGGGAGGATTCTTGGCAGGATCACTGTGGC	587		
Query 971	TTGGACTTCCGATTGTCCAGAATCTTGCTCCTGCACGTAGTATGCCAAAAACAAGTCA-	1029		
Sbjct 586	TTGGACTTCCGATTGTCCAGAATCTTGCTCCTGCACGTAGTATGCCAAAAACAAGTCAA	527		
Query 1030	GCTCCT-A-CTGATACCGGTGGATGTGGTGTGGCTGAACACAAGGGGCTGGGTT-GTGCA	1086		
Sbjct 526	GCTCCTTAACGTGATACTG-T-GATGTGGTGTGGCTGAACACAAGGTGCTGGGTTGTGCA	469		
Query 1087	GTGCGGGGACTTCATGAGTCGCTCTCCATCCGTACTCTCCAAAGGGATGCCTTTGAACAG	1146		
Sbjct 468	GTGCGGGGACTTCATGAGTCGCTCTCCATCCGTACTCTCCAAAGGGATGCCTTTGAACAG	409		
Query 1147	GATC-CCATGACTAGATCCGAACGCC-GACTTTGTCTGCC-GTGGCAGG-AGTCGATC-T	1201		
Sbjct 408	GATCACCATGACTAGATCCAGACGCCAGACTTTGTCTGCCTGTGCGAGGCAGTCGATCCT	349		
Query 1202	CCTA-T-TTAACTTTTGGGTGGGGATGGGAGA-GAAAAAACACGGGGGTTCTTGCCA-	1257		
Sbjct 348	CCTAATCTTACCCTTCT-GGTCTGGATTGGATAAGACACAGCACGGGTGCTTCTTGCCAG	290		

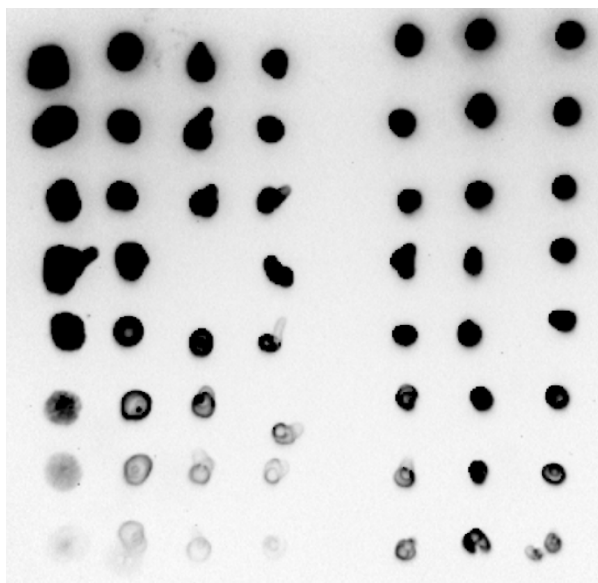
Alignment of pCH_mNfib4 infusion sequencing results with pCH_mNfib4 construct supposed to be, EBV-R reverse primers

Score	Expect	Identities	Gaps	Strand
2146 bits(1162)	0.0	1232/1263(98%)	16/1263(1%)	Plus/Plus
Query 4	GTGGGA-GT-T-TATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCA	60		
Sbjct 510	GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCA	569		
Query 61	TCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGGACTCTAGA	120		
Sbjct 570	TCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGGACTCTAGA	629		
Query 121	GGATCCGGTACTCGAGGAACTGAAAAACCAGAAAGTTAACTGGTAAGTTTAGTCTTTTTG	180		
Sbjct 630	GGATCCGGTACTCGAGGAACTGAAAAACCAGAAAGTTAACTGGTAAGTTTAGTCTTTTTG	689		
Query 181	TCTTTTATTTTCAGGTCCCGGATCCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGG	240		
Sbjct 690	TCTTTTATTTTCAGGTCCCGGATCCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGG	749		
Query 241	ATGTTGCCTTTACTTCTAGGCCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAA	300		
Sbjct 750	ATGTTGCCTTTACTTCTAGGCCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAA	809		
Query 301	TTGTACCCGCGGCCGCCAGCCATGTATCCTTACGATGTACCAGACTATGCTGGGCCTCAC	360		
Sbjct 810	TTGTACCCGCGGCCGCCAGCCATGTATCCTTACGATGTACCAGACTATGCTGGGCCTCAC	869		
Query 361	GGGCCGATGATGTATTCTCCCATCTGTCTCACTCAGGATGAATTCACCCATTTATTGAG	420		
Sbjct 870	GGGCCGATGATGTATTCTCCCATCTGTCTCACTCAGGATGAATTCACCCATTTATTGAG	929		
Query 421	GCACTTCTTCCTCACGTCCGTGCAATTGCCTATACTTGGTTCAACCTGCAGGCTCGAAAA	480		
Sbjct 930	GCACTTCTTCCTCACGTCCGTGCAATTGCCTATACTTGGTTCAACCTGCAGGCTCGAAAA	989		
Query 481	CGCAAGTACTTTAAAAAGCATGAGAAACGAATGTCGAAGGATGAAGAAAGGGCAGTCAA	540		
Sbjct 990	CGCAAGTACTTTAAAAAGCATGAGAAACGAATGTCGAAGGATGAAGAAAGGGCAGTCAA	1049		
Query 541	GACGAGCTGCTCAGTGAGAAGCCCGAAATCAAGCAGAAGTGGGCATCCAGGCTCCTGGCC	600		
Sbjct 1050	GACGAGCTGCTCAGTGAGAAGCCCGAAATCAAGCAGAAGTGGGCATCCAGGCTCCTGGCC	1109		
Query 601	AAACTGCGCAAAGATATCCGCCAGGAGTACCGGGAGGACTTTGTGCTTACCGTGACTGGC	660		
Sbjct 1110	AAACTGCGCAAAGATATCCGCCAGGAGTACCGGGAGGACTTTGTGCTTACCGTGACTGGC	1169		
Query 661	AAGAAGCACCCGTGCTGTGTCTTATCCAATCCAGACCAGAAGGGTAAGATTAGGAGGATC	720		
Sbjct 1170	AAGAAGCACCCGTGCTGTGTCTTATCCAATCCAGACCAGAAGGGTAAGATTAGGAGGATC	1229		
Query 721	GACTGCCTGCGACAGGCAGACAAAGTCTGGCGTCTGGATCTAGTCATGGTGATCCTGTTC	780		
Sbjct 1230	GACTGCCTGCGACAGGCAGACAAAGTCTGGCGTCTGGATCTAGTCATGGTGATCCTGTTC	1289		
Query 781	AAAGGCATCCCTTTGGAGAGTACGGATGGAGAGCGACTCATGAAGTCCCCGCATGCACA	840		
Sbjct 1290	AAAGGCATCCCTTTGGAGAGTACGGATGGAGAGCGACTCATGAAGTCCCCGCATGCACA	1349		
Query 841	AACCCAGCACTTTGTGTTTCAGCCACACCACATCACAGTATCAGTTAAGGAGCTTGACTTG	900		
Sbjct 1350	AACCCAGCACTTTGTGTTTCAGCCACACCACATCACAGTATCAGTTAAGGAGCTTGACTTG	1409		
Query 901	TTTTTGGCATACTACGTGCAGGAGCAAGATTCTGGACAATCAGGAAGTCCAAGCCACAGT	960		
Sbjct 1410	TTTTTGGCATACTACGTGCAGGAGCAAGATTCTGGACAATCAGGAAGTCCAAGCCACAGT	1469		
Query 961	GATCCTGCCAAGAATCCTCCAGGGTACCTCGAGGACAGCTTTGTAAAAATCCGGAGTCTTC	1020		
Sbjct 1470	GATCCTGCCAAGAATCCTCCAGGGTACCTCGAGGACAGCTTTGTAAAAATCCGGAGTCTTC	1529		
Query 1021	AATGTATCAGAGCTTGTGAGAGTATCCAGAACACCCATAACCCAGGGAACCTGGAGTC-AC	1079		
Sbjct 1530	AATGTATCAGAGCTTGTGAGAGTATCCAGAACACCCATAACCCAGGGAACCTGGAGTCAAC	1589		
Query 1080	TTCCCAATCGGAGAAAT-CCCAGCC-ACCATACTATCATGACAGGAACCTGGTGTGAAC	1137		
Sbjct 1590	TTCCCAATCGGAGAAATTTCCAGCCAACCATACTATCATGACATGAACTCTGGTGTGAAC	1649		
Query 1138	TGGCAGAGGTCGCTGTCTTCTCC-CCGAGCAGCAAA-GACCCAA--CTTTATCTTTTAAT	1193		
Sbjct 1650	CTGCAGAGGTCGCTGTCTTCTCCACCGAGCAGCAAAAGACCCAAAACCTATATCTATAGAT	1709		

pCH_mNfib4 construct sequencing with forward CMV-40 primers

Query	5	GTGGG-GGGCT-TAT-AGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCA	61
Sbjct	510	GTGGGAGGTCATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCA	569
Query	62	TCCACGCTGTTTGTGACCTCCATAGAAGACACCGGGACCGATCCAGCTCCGGACTCTAGA	121
Sbjct	570	TCCACGCTGTTTGTGACCTCCATAGAAGACACCGGGACCGATCCAGCTCCGGACTCTAGA	629
Query	122	GGATCCGGTACTCGAGGAAGTAAAAACAGAAAGTTAACTGGTAAGTTTAGTCTTTTTG	181
Sbjct	630	GGATCCGGTACTCGAGGAAGTAAAAACAGAAAGTTAACTGGTAAGTTTAGTCTTTTTG	689
Query	182	TCTTTTATTTTCAAGTCCCGGATCCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGG	241
Sbjct	690	TCTTTTATTTTCAAGTCCCGGATCCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGG	749
Query	242	ATGTTGCTCTTACTTCTAGGCCTGTACGGAAGTGTACTTCTGCTCTAAAAGCTGCGGAA	301
Sbjct	750	ATGTTGCTCTTACTTCTAGGCCTGTACGGAAGTGTACTTCTGCTCTAAAAGCTGCGGAA	809
Query	302	TTGTACCCGCGGCCGCGCAGCCATGTATCCTTACGATGTACCAGACTATGCTGGGCCTCAC	361
Sbjct	810	TTGTACCCGCGGCCGCGCAGCCATGTATCCTTACGATGTACCAGACTATGCTGGGCCTCAC	869
Query	362	TAACCGGTCTGATGTACACGAGCTCTAGCTTAAGAGATCTATGCATAGCTAGCTAGGGCC	421
Sbjct	870	TAACCGGTCTGATGTACACGAGCTCTAGCTTAAGAGATCTATGCATAGCTAGCTAGGGCC	929
Query	422	GCGGGGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGC	481
Sbjct	930	GCGGGGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGC	989
Query	482	AGTGaaaaaaTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA	541
Sbjct	990	AGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA	1049
Query	542	TAAGCTGCAATAAACAAAGTTAACAAACAATTCATTCATTTTATGTTTCAGGTTTCAGG	601
Sbjct	1050	TAAGCTGCAATAAACAAAGTTAACAAACAATTCATTCATTTTATGTTTCAGGTTTCAGG	1109
Query	602	GGGAGGTGTGGGAGGTTTTTTCGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGG	661
Sbjct	1110	GGGAGGTGTGGGAGGTTTTTTCGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGG	1169
Query	662	CGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAAATCCACACA	721
Sbjct	1170	CGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAAATCCACACA	1229
Query	722	ACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCA	781
Sbjct	1230	ACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCA	1289
Query	782	CATTAATGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGC	841
Sbjct	1290	CATTAATGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGC	1349
Query	842	ATTAATGAATCGGCCAACGCGGGGAGAGGCGGTTTGCCTATTGGGCGCTCTTCCGCTT	901
Sbjct	1350	ATTAATGAATCGGCCAACGCGGGGAGAGGCGGTTTGCCTATTGGGCGCTCTTCCGCTT	1409
Query	902	CCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGCAGCGGTATCAGCTCACT	961
Sbjct	1410	CCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGCAGCGGTATCAGCTCACT	1469
Query	962	CAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAG	1021
Sbjct	1470	CAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAG	1529
Query	1022	CAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATA	1081
Sbjct	1530	CAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATA	1589
Query	1082	GGCTCCGCCCCCTGACGAGCATCAC-AAAATCGACGCTCAAGTCAGAGGTGGCGAAACC	1140
Sbjct	1590	GGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC	1649
Query	1141	CGACGGGACTATAAGATACAGGCGTTTTCCCCCTGGAAACTCCCTCGGG-GCCTTCCTG	1199
Sbjct	1650	CGACAGGACTATAAGATACAGGCGTTTTCCCCCTGGAAGTCCCTCGTGCCTCTCCTG	1709
Query	1200	TT-CGAACCTGGCGCTTAC-GGAAACTGT-CGCCTTTTTCCCTTCGGGAAG	1248
Sbjct	1710	TTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAG	1761

pCH_MCS control plasmid construct sequenced with CMV-F. 362-397 nucleotides belongs to the multiple cloning site inserted to this control plasmid.



Detection of bitonylation efficiencies of p21 EMSA single strand oligonucleotides.



PHOTO

CURRICULUM VITAE

Name Surname : Şeyma VARINCA

Place and Date of Birth : ISTANBUL 24.05.1994

E-Mail : seymavarinca@gmail.com

EDUCATION :

- **B.Sc.** : 2017, University, Faculty of Science and Letters, Dept.
of Molecular Biology and Genetics