UNIVERSIDADE DO ALGARVE

DEPARTAMENTO DE CIÊNCIAS
BIOMÉDICAS E MEDICINA

INVESTIGATING THE LOCALIZATION OF AN AVIAN HAIRY HOMOLOG (C-HAIRY1) PROTEIN

Janine Anselmo Cravo

Tese

Mestrado em Ciências Biomédicas

Trabalho efectuado sob a orientação de:

Professora Doutora Isabel Palmeirim e

Doutora Lisa Gonçalves

2013

"Time is nature's way of making sure everything doesn't happen at once."

Henri Bergson

UNIVERSIDADE DO ALGARVE

DEPARTAMENTO DE CIÊNCIAS BIOMÉDICAS E MEDICINA

UNIVERSIDADE DO ALGARVE

DEPARTAMENTO DE CIÊNCIAS

BIOMÉDICAS E MEDICINA

INVESTIGATING THE LOCALIZATION OF AN AVIAN HAIRY HOMOLOG (C-HAIRY1) PROTEIN

Dissertação apresentada à Universidade do Algarve para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Ciências Biomédicas, realizada sob a orientação científica da Professora Doutora Isabel Palmeirim (Departamento de Ciências Biomédicas e Medicina e Centro de Biomedicina Molecular e Estrutural - CBME) e da Doutora Lisa Gonçalves (CBME).

Janine Anselmo Cravo

2013

INVESTIGATING THE LOCALIZATION OF AN AVIAN HAIRY HOMOLOG (C-HAIRY1) PROTEIN

Declaração de autoria de trabalho

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

AGRADECIMENTOS

Após mais uma etapa na minha vida académica, existem contributos de natureza diversa que não podem e nem devem deixar de ser realçados. Por essa razão, desejo expressar os meus sinceros agradecimentos:

À Professora Doutora Isabel Palmeirim, minha orientadora, por me ter concedido a oportunidade de participar neste projecto. Agradeço o apoio, a confiança, e a partilha do saber que tanto me ajudou na realização deste projecto. Acima de tudo, obrigada pela competência científica, pelo acompanhamento do trabalho, pela disponibilidade e generosidade demonstrada.

À Doutora Lisa Gonçalves que vivenciou comigo todas as alegrias e frustrações, e cuja excelente capacidade critica me guiou durante todo o trabalho e sempre me incentivou a ir mais longe.

Ao Tomás quero agradecer toda generosidade e paciência demonstrada ao longo do desenvolvimento do trabalho e que tanto me ajudou com o seu inglês exemplar.

A todo o grupo de Cronobiologia, do qual orgulhosamente fiz parte. A vocês não vos considero colegas mas também amigos. A todos quero agradecer pelo enorme espírito de equipa e por toda a disponibilidade demonstrada.

À Doutora Raquel Andrade por toda a disponibilidade que sempre demonstrou.

À Professora Doutora Solveig, e ao seu grupo por toda a ajuda concedida.

A todos os professores funcionários e colegas do CBME pelo apoio prestado, em especial à Doutora Cláudia Florindo pelos ensinamentos e sugestões no campo da microscopia.

Aos meus amigos que sempre presentes me aconselharam, apoiaram e que nunca me deixaram desistir. A vossa presença tornou a todo o meu percurso mais leve.

A toda a família, com um especial agradecimento aos meus pais, que foram e sempre serão o grande pilar que suporta a minha vida. Obrigado por todo o amor incondicional.

Obrigado a todos.

ABSTRACT

The embryo segmental pattern is first established with the formation of somites. Somites are embryonic segments of vertebrates, periodically formed in a strict temporal precision, which has been believed to be governed by a biological clock, called the 'segmentation clock'. In the mid-70s, the *"Clock and Wavefront Model" was* proposed, predicting the existence of an intracellular clock or oscillator in the presomitic mesoderm (PSM) cells. Several years later, Palmeirim et. al., identified the first molecular evidence of this clock by discovering and characterizing the oscillatory expression of *c-hairy1* in the chicken embryo paraxial PSM, which strikingly matched the period of somite formation (90 min).

This present work aims to characterize the chicken Hairy1 protein, and to study its expression both in early embryo stages and chicken embryonic fibroblasts (CEFs). Bioinformatic tools have predicted the biochemical properties, primary and secondary structure, post transcriptional modifications and subcellular localization of c-Hairy1 protein. By western blot were established the optimal working conditions of the customized monoclonal antibody, as well as the expression of the protein in both chick embryos and CEFs. The protein distribution and its subcellular localization in CEFs was assessed by immunofluorescence.

Results of western blot have demonstrated sensitivity of the antibody, although its specificity for c-Hairy1 protein remains debatable. Even more, both immunofluorescence and bioinformatics analysis showed c-Hariy1 to be localized both in nucleus and cytoplasm. Interestingly, it was also demonstrated that the nucleus:cytoplasm ratio distribution varied between cells.

These work's findings suggest that c-Hairy1 protein holds much investigation potential, and the optimization of the antibody working conditions enables its use for further studies.

Keywords: segmentation clock, chicken embryo, Hes-gene family, c-hairy1, bioinformatics, chicken embryonic fibroblasts

RESUMO

Nos vertebrados, o padrão segmentado do embrião é estabelecido com o aparecimento de unidades metaméricas na mesoderme pre-somítica (MPS), os sómitos, que se formam numa precisão temporal restrita demonstrando um comportamento periódico. Estes originam todas as estruturas segmentadas presentes no animal adulto: vértebras, discos intervertebrais, costelas, a derme das costas e todos os músculos esqueléticos do tronco e membros.

Em meados dos anos 70, foi proposto o modelo "*Clock and Wavefront*" que previa a existência de um relógio intracelular, ou um oscilador em células da MPS, responsável pela regulação da formação dos sómitos, a somitógenese. Vários anos mais tarde, Palmeirim et al., identificou pela primeira vez uma evidência molecular deste relógio, descrevendo e caracterizando uma expressão oscilatória de c-*hairy1* na MPS paraxial de embriões de galinha. Surpreendentemente esta periodicidade correspondia ao período de formação de um par sómitos em galinha (90 min). Vários componentes das vias de sinalização Notch, FGF e Wnt foram, mais tarde, descritos como tendo uma expressão dinâmica semelhante ao de *c-hairy1* na MPS de embriões de galinha, ratinho e peixe zebra. Estas vias de sinalização parecem estar interligadas e constituem uma rede complexa de vias oscilatórias envolvidas na somitogénese.

Os genes da família Hes, à qual *c-hairy1* pertence, são descritos como alvos da via Notch e funcionam como repressores de transcrição que regulam a proliferação e diferenciação celular. Estes participam em diversos processos do desenvolvimento embrionário, funcionando como relógios biológicos, mas também mantendo o estado indiferenciado de células progenitoras. A desregulação da expressão destes genes tem sido ligada a defeitos do desenvolvimento embrionário e oncogénese.

As proteínas da família Hes possuem três domínios conservados, basic Helix-Loop-Helix (bHLH), Orange e WRPW, que lhes conferem características únicas como repressores e osciladores. A região basic é responsável pela ligação da proteína ao DNA, o domínio Helix-Loop-Helix promove a dimerização, o domínio

Orange actua como selecionador de um possível heterodímero ao qual o factor Hes se liga. O motivo WRPW é responsável pela interação de co-repressores e também funciona como sinal de poliubiquitinização.

Alguns membros da família Hes foram descritos como apresentando expressão oscilatória, quer ao nível do RNAm como da proteína. Tal expressão parece ser auto-regulada por um mecanismo de feedback negativo, em que o homodímero da proteína se liga ao promotor do seu gene, reprimindo-o. Por outro lado, o curto tempo de meia-vida do RNAm e da proteína parece também estar na origem deste fenómeno.

Com o objectivo de caracterizar a proteína c-Hairy1, e estudar a sua expressão em estadios iniciais de embriões de galinha e em fibroblastos embrionários de galinha (FEG), foi requerida a uma empresa a produção de um anticorpo monoclonal contra c-Hairy1. Neste contexto, ferramentas bioinformáticas foram usadas para analisar as propriedades bioquímicas, estrutura primária e secundária, modificações pós-transcricionais, bem como a localização subcelular da proteína c-Hairy1. Por western blot foram estabelecidas as condições funcionais ótimas do anticorpo monoclonal, assim como a expressão da proteína em embriões de galinha e em FEGs. A distribuição da proteína e a sua localização subcelular foi determinado por imunofluorescência.

A análise bioinformática da sequência proteica previu c-Hairy1 como sendo uma proteína pequena com características hidrofílicas. A previsão da estrutura secundária mostrou uma grande percentagem de regiões *coil*, apresentando algumas hélices-α e folhas-β localizados nos domínios conservados. Tais características podem conferir a c-Hairy1 uma estrutura globular.

A análise de modificações pós-transcricionais usando dois softwares sugeriu vários possíveis locais de fosforilação distribuídos principalmente nas extremidades da proteína. Foram também sugeridas dez cinases capazes de fosforilar estes locais. Um estudo dos da localização subcelular permitiu a identificação de um sinal de localização nuclear conservado nas sequências proteicas de c-Hairy1 e nos seus homólogos, propondo a localização de c-Hairy1

no núcleo. Este estudo revelou ainda dois potenciais sinais de exportação nuclear, sugerindo a exportação de c-Hairy1 para o citoplasma.

Para a optimização das condições de funcionamento do anticorpo monoclonal estabeleceram as quantidades mais apropriadas de proteína purificada e de extractos proteicos a serem carregadas no gel. A sensibilidade do anticorpo monoclonal foi demonstrada pela detecção da banda de peso molecular esperada de c-Hairy1, tanto nos extractos proteicos de embriões e FEGs. Contudo, outras bandas de peso molecular não esperado foram sistematicamente detectadas, apesar do aumento do tempo de desnaturação das amostras. Estas bandas podem ser justificadas por diferentes graus de fosforilação, formação de dímeros, *splicing* alternativo, isoformas não descritas, e ainda ligação do anticorpo a outras proteínas. Concluiu-se que a especificidade do anticorpo para c-Hairy1 é ainda um tema sujeito a debate. Futuramente, poderá ser realizada uma identificação por espectrometria de massa de forma a garantir que as bandas detectadas correspondem à proteína em estudo.

Os resultados de imunofluorescência, complementados pela previsão bioinformática, mostraram que a proteína c-Hairy1 está localizada tanto no núcleo como citoplasma de FEGs. Estas evidências sugerem um possível *shuttle* de c-Hairy1 entre o núcleo e citoplasma, possivelmente regulado por diferentes níveis de fosforilação da proteína. No entanto, este mecanismo de *shuttle* deverá ser validado em estudos futuros.

Os resultados de imunofluorescência revelaram ainda um sinal de fluorescência mais intenso nas bordas celulares em aproximadamente 40% das FEGs analisadas, indicando uma possível interação entre a c-Hairy1 e outras proteínas do citoesqueleto. Em mitose, c-Hairy1 apresentou uma distribuição homogénea na célula, o que poderá reflectir o papel desempenhado por esta como promotor de proliferação celular.

A análise detalhada da localização de c-Hairy1 em 54 FEGs, demonstrou que o rácio de distribuição núcleo:citoplasma variou entre células. Esta variação pode

sugerir uma distribuição dinâmica da proteína, apoiando a ideia da existência do *shuttle* de c-Hairy1 entre o núcleo e citoplasma.

Este trabalho permitiu estabelecer as condições óptimas de funcionamento do anticorpo contra c-Hairy1, permitindo a realização de novos estudos de caracterização desta proteína a nível da sua expressão e função.

Palavras-chave: Relógio embrionário, embrião de galinha, Hes-gene family, c-Hairy1, bioinformática, Fibroblastos embrionários de galinha.

TABLE OF CONTENTS

ABREVIATIONS LIST

aa	Amino acid
ATM	Ataxia telangiectasia mutated
bHLH	Basic helix-loop-helix
BSA	Bovine serum albumin
CEFs	Chick embryonic Fibroblasts
Cdc2	Cyclin dependent kinase (Cdk1)
CKI	Casein kinase 1
CKII	Casein kinase 2
DAPI	4',6-diamidino-2-phenylindole
DIC	Differential interference contrast microscopy
DMEM	Dulbecco's modified Eagle's medium
DNA	Deoxyribonucleic acid
DNAPK	DNA-dependent protein kinase
EGFR	Epidermal growth factor receptor
FBS	Fetal bovine serum
FGF	Fibroblast growth factor
GRAVY	Grand average of hydropathicity
GSK3	Glycogen Synthase Kinase 3
Hes	Hairy Enhancer of Split
Hrt	Hairy-related transcription factors
INSR	Insulin receptor
Lnfg	Lunatic fringe
Da	Dalton
κ-NN	*k*-Nearest neighbor
NES	Nuclear export signal
NLS	Nuclear localization signals
NCBI	National Center for Biotechnology Information
NICD	Notch intracellular domain
OSCC	Oral squamous cell carcinoma

p38 MAPK	p38 Mitogen-activated protein kinase
PBS	Phosphate buffered saline
PKA	Cyclic AMP dependent protein kinase
PKB	Protein kinase B
PKC	Protein kinase C
PKG	Cyclic GMP-dependent protein kinase
PSM	Presomitic mesoderm
RA	Retinoic acid
RNA	Ribonucleic acid
ROI	Region of interest
RSK	ribosomal s6 kinase
SD	Spondylocostal dysostosis
Src	Proto-oncogene tyrosine-protein kinase
SAP18	Sin3-associated polypeptide, 18 kDa
TLE	Transducin-like enhancer of split

TABLES LIST

FIGURES LIST

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

Life itself is a complex, intertwining and extraordinary process, full of wonder to unveil, and questions to answer. For instance, how does a single cell become a complete individual? The more we discover about the mechanism that regulates life, the more we realize that there are so many points and steps that can go wrong.

Multicellular organisms do not spring forth fully formed. Rather, they arise by a relatively slow process of progressive change. In almost all cases, the development of a multicellular organism begins with a single cell, called the fertilized egg, or zygote, which divides mitotically to produce all the cells of the body (Alberts et al., 2007). The study of animal development from fertilization to birth has traditionally been called embryology, however development does not stop at birth or even at adulthood. In fact, most organisms never stop developing and so the subject responsible for studying embryonic and other development processes is called developmental biology (Gilbert, 2010). Developmental biology is one of the most exciting areas in biology, as it integrates anatomy, physiology, genetics, biochemistry and cellular and molecular biology. It too deals with fascinating biologic events, such as changes in form, structure and function of the organism (Tuan and Lo, 2000). As the embryos go through growth and differentiation, there are three well known dimensions that grant a field for their growth, but the mechanisms underlying the processes of life have to be performed in a synchronicity of events, so time can be considered as the fourth dimension (Andrade et al., 2007). This can lead us to the question, how is temporal control achieved? Biology takes advantage of animal models to answer this and many more questions.

Every model has some characteristics that are more or less advantageous to study a particular question. The models which are most commonly used in development include the nematode *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, the zebrafish *Danio rerio*, the South African clawed toad frog *Xenopus laevis*, the chicken *Gallus gallus*, and the mouse *Mus musculus (Tuan*

and Lo, 2000). Of these, the chicken was the first model be used in developmental studies, and there can be found reports which dates back to ancient egyptians (Stern, 2005).

1.1. CHICKEN EMBRYO AS A MODEL

For a long time the chick embryo (*Gallus gallus*) has been used in different areas in biology, it has a distinguished history as a major model system in developmental biology, and has also contributed with major concepts to immunology, genetics, virology, cancer, and cell biology (Siegel et al., 2006; Stern, 2005)

Chick embryo gives advantages over other embryo models in embryology since the chicken egg is inexpensive, accessible all year, and can be purchased in any specified quantity. Chicken eggs can be incubated to any stage of interest, which facilitates designing an experiment that requires specific stages of development. When the egg is laid, the avian embryo consists of a flat, two-layered blastoderm that lies on the surface of the yolk and, therefore, is readily accessible for experimentation (Gilbert, 2010). Chick development occurs with incubation at 38°C, and the period of incubation is about 21 days, after which the egg will hatch. Experimental design and timely data collection gets easier with such rapid development. Cultures *in ovo* and *ex-ovo* (removed from the shell) can be performed during the period of early development when so much is occurring. Two great advantages of this model are its semitransparency, making viewing of internal tissues possible under the microscope, and their sufficient size that allows several types of micromanipulation at these early stages. Due to its many advantages, there are many detailed experimental studied on the chick embryo, in which its availability adds to the value of chick embryos as a model system for studying development (Tuan and Lo, 2000).

Because the chick embryo forms most of its organs in ways very similarly to those of mammals, it has often served as a surrogate for human embryos (Gilbert, 2010). As so, all the great advances and discoveries in chick have relevance for

other vertebrates, including mammals, and some have brought about dramatic changes in fundamental understanding of development itself (Siegel et al., 2006).

1.2. BRIEF VIEW OF AVIAN EMBRYO DEVELOPMENT STAGES

In embryonic development there are some common stages to all vertebrate species known. The fertilization is the first step and gives rise to the process of development. Immediately following fertilization, a series of extremely rapid mitotic divisions occurs wherein the enormous volume of zygote cytoplasm is divided into numerous smaller cells, by a phenomenon called cleavage. These smaller cells are called blastomeres, and they generally form a sphere known as a blastula, by the end of cleavage. Once the rate of mitotic division has slowed down, the blastomeres go through dramatic movements where they change their positions relative to one another. At this point, the embryo is now in the process of gastrulation and is said to be in gastrula stage. As a result of gastrulation, the embryo ends up with three germ layers: the ectoderm, the endoderm, and the mesoderm. Once the three germ layers are established, the cells interact with one another and rearrange themselves to produce tissues and organs. This process is called organogenesis. Many organs contain cells from more than one germ layer, and it is not unusual for the outside of an organ to be derived from one layer and the inside from another (Gilbert, 2010).

1.3. SOMITOGENESIS AND THE MOLECULAR CLOCK

One of the major tasks of gastrulation is to create a mesodermal layer between the endoderm and the ectoderm. A localized and specialized region of the embryo is established during gastrulation, and has a role of an organizer. In avian it is called Hensen's node and is found at the rostral end of primitive streak, which is the major structural characteristic of avian, reptilian, and mammalian gastrulation (Darnell et al., 1999; Gilbert, 2010). As the primitive streak regresses and the

neural tube folds, beginning to gather at the center of the embryo, somitogenesis is initiated (Gilbert, 2010). Somitogenesis is a periodic and sequential process, wherein each successive bilateral somite pair segregates at a regular time interval from the anterior end of the pre-somitic mesoderm (PSM) as the body axis elongates (Oates et al., 2012).

Somites consist of epithelial spheres of cells and, although these are transient structures, they are of extreme importance for organizing the segmental pattern of vertebrate embryos (Figure 1.1A). These structures are the earliest manifestation of the segmental pattern of the adult vertebrate body and give rise to the vertebrae and ribs, the dorsal dermis, the skeletal muscles of the back, body wall and limbs (Andrade et al., 2007; Gilbert, 2010). Interestingly, the intervals of each somite formation and the total number of somites formed are intrinsic to the species (Described in Table 1.1) (Andrade et al., 2007). This periodic event, somite segmentation, has been believed to be governed by a biological clock, called the 'segmentation clock'. In the mid-70s, the *"Clock and Wavefront Model"* was proposed, and it predicted the existence of an intracellular clock or oscillator in the PSM cells, which temporal periodicity turns into the spatial periodicity of somites (Cooke and Zeeman, 1976). In this model, the wavefront represents the anterior to posterior progression of development of the embryo. Thus, this wavefront governs the maturation of the PSM to become somites. In this sense, a somite unit forms only in the presence of two conditions, the wavefront of maturation must reach a group of cells who are at the appropriate phase of the clock. This model postulates that somite size is regulated by the speed of the wavefront while the rate of somite formation is controlled by the frequency of the oscillator (Gibb et al., 2010).

Table 1.1 - Number of somites and its periodic formation time in several organisms.

Specie	Total number of somite pairs formed	Time intervals
Zebrafish	30 (Stickney et al., 2000)	30 min (Stickney et al., 2000)
Chick	52 (Andrade et al., 2007)	90 min (Andrade et al., 2007)
Mouse	65 (Tam, 1981)	120 min (Bessho and Kageyama, 2003)
Human	33 (Gomez et al., 2008)	8h (Bessho and Kageyama, 2003)

Several years later, molecular evidences supporting both the intrinsic clock and the wavefront were reported (Andrade et al., 2007). For instance, it has been described that the wavefront relies on an intersection of gradients and cross regulatory activities of three transduction pathways, namely FGF (fibroblast growth factor), Wnt and Retinoic acid (RA). Therefore, the determination front is delimited by the intersection of these three gradients, and it denotes the position where the PSM is already committed (Figure 1.1B). As for the molecular oscillator the first evidence came in 1997, when Palmeirim et. al, reported the periodic expression of the avian homolog of the Drosophila segmentation gene *hairy*, in chick embryo PSM. Chick *hairy1* gene was shown to be cyclically expressed in PSM cells, with the same periodicity as somite formation (90 min). This was the first example of the molecular oscillator working in embryos (Bessho and Kageyama, 2003), and it became known as the segmentation clock.

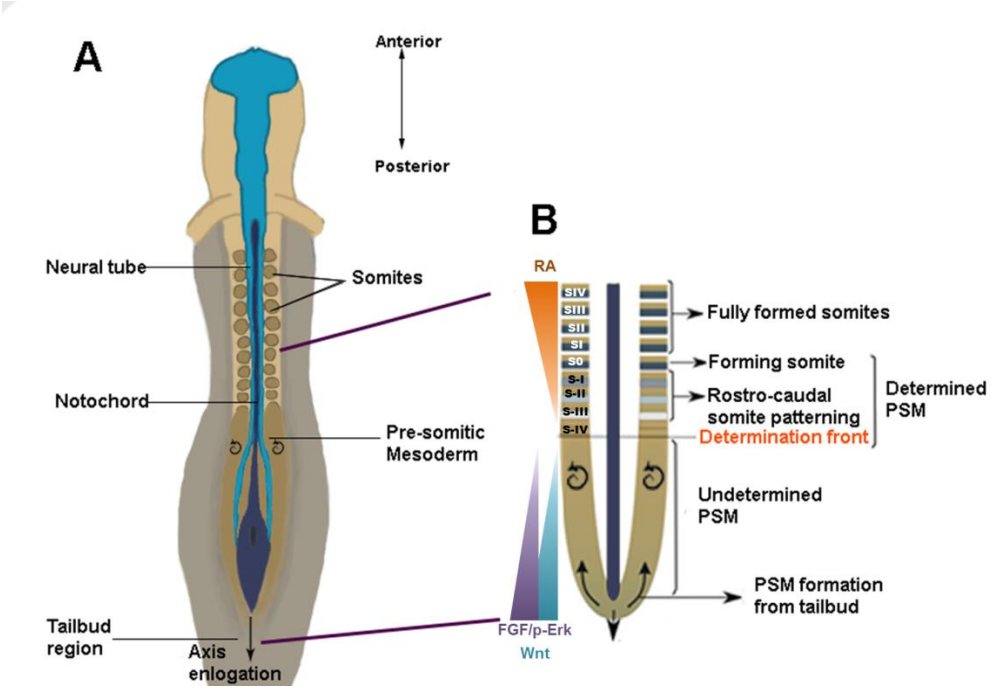


Figure 1.1 - Representation of the somitogenesis, molecular clock and determination front concept. A: Illustration of a HH10 chick embryo, highlighting its major structures. Epithelial Somite pairs are periodically formed from the anterior most tip of the PSM in an AP manner, and due to new entrance of cells from the tailbud region, the embryo elongates posteriorly. B: Enlarged view of posterior embryo highlighting determination front (DF), situated more or less in the two third level of the PSM. This

illustration indicates as well the three gradients by which the determination front is defined, Retinoic Acid (RA) FGF/p-ERK and Wnt. New somites are formed every 90 minutes, having the notion of time and space which are provided respectively by the molecular clock and the determination front. Elongation is made in an Anterior Posterior manner. The forming somite is represented as S0 and the fully formed somites as SI to SIV. Adapted from (Sheeba, 2011).

Chick *hairy1* starts to be expressed rhythmically as early as gastrulation (Jouve et al., 2002), and when the embryo elongates and PSM starts to form, it shows a very particular expression pattern, since it resembles a wave sweeping the PSM in a posterior-to-anterior direction (Figure 1.2). Each cycle of expression goes through 3 different phases. In phase I, *hairy1* appears in the large caudal domain, filling about 70% of the posterior PSM and in the anterior PSM it can be detectable a narrow stripe that marks the position where the future somite will form. In phase II, while the anterior stripe is maintained, the caudal expression of *hairy1* disappears and it seems to shift anteriorly. In phase III, *hairy1* expression shifts to the anterior PSM, corresponding to about one somite-length. Simultaneously, a new caudal domain arises which corresponds to the phase one of the next cycle of expression (Palmeirim et al., 1997).

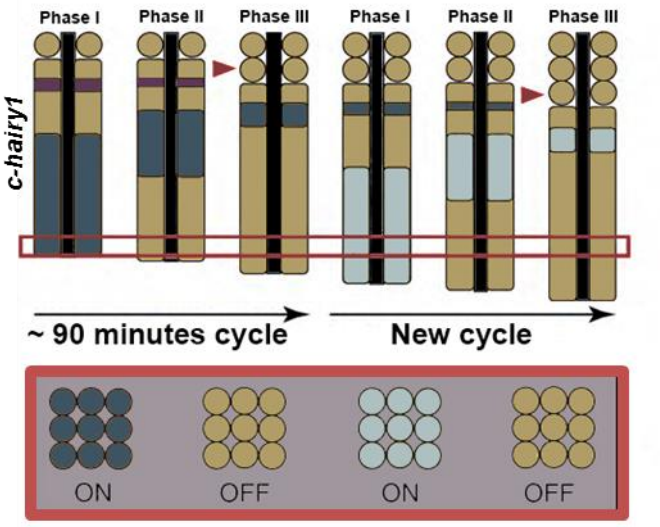


Figure 1.2 - Somitogenesis molecular clock. Illustration of chick PSM, and representation of the three phases, by which is defined a cycle of *c-hairy1* expression, which take place in every 90 minutes. Individual cells over time, turn on and of the gene. This dynamic expression at the level of single cells, by virtue of being synchronized across the PSM, results in kinematic 'waves' of gene expression that 'move' across the PSM. Adapted from (Gibb et al., 2010).

After c-*hairy1* discovery, several more genes were found to be dynamically expressed in the PSM with cycling times equal to the time taken to form one somite (Andrade et al., 2007; Baker et al., 2006; Dequéant et al., 2006; Krol et al., 2011), these genes included not only components of the Notch signalling pathway, but also Wnt and FGF components (Dequéant et al., 2006; Krol et al., 2011). It is now evident that the molecular events underlying somitogenesis are highly conserved among vertebrates, since periodic gene transcription has also been described in other animal models used in Developmental Biology - mouse, zebrafish, frog and medaka (Andrade et al., 2007). Several homologs of the *c-hairy1* have been identified in humans, and other mammals, as well as frogs and zebrafish, these include not only other *hairy* or *Enhancer of split (E(spl)*, but also *hairy/E(spl)* (*hes*)*, hairy/E(spl)-related* (*her*) and *hairy/E(spl)-related with YRPW motif* (*hey*) genes (Table 1.2).

Table 1.2 - Compilation of Hes-family genes and its different transcripts described so far in Human, Mouse, Chicken, *Xenopus leavis* and Zebrafish.

Human	**Mouse**	**Chicken**	***Xenopus leavis***	**Zebrafish**
hes1 (Takebayashi et al., 1994) *hes2* (Nishimura et al., 1998) *hey1, hey 1, hey2 (Steidl et al., 2000)* *hes6.1, hes6.2* (Bae et al., 2000) *hes4.1, hes4.2, hes7.1, hes7.2* (Bessho et al., 2001) *hes3* (Katoh, 2004) *hes5* (Kamakura et al., 2004)	*hes1, hes3* (Sasai et al., 1992) *hes2* (Ishibashi et al., 1993) *hes5* (Sakagami et al., 1994) *hey1, hey2* (Leimeister et al., 1999) *hes6* (Bae et al., 2000) *hes7* (Bessho et al., 2001)	*hairy1a* (Palmeirim et al., 1997) *hairy2* (Jouve et al., 2000) *hey1, hey2* (Leimeister et al., 2000) *hairy1b* (Vasiliauskas et al., 2003) *hes5*(Fior and Henrique, 2005)	*hes4a** (Turner and Weintraub, 1994) *hes1a*** (Dawson et al., 1995) *hey1* (Pichon et al., 2002) *esr10a, hes9.1a**** (Gawantka et al., 1998) *hes1b, hes2, hes3.3***** (Klein et al., 2002) *hes3, hes9.2a (Deblandre et al., 1999)* *hes4b****** (Davis et al., 2001) *hes5.1, hes5.2-b, hes9.1b* (Takada et al., 2005) *hes7.1* (Shinga et al., 2001)	*hey2* (Weinstein et al., 1995) *her1*(Muller et al., 1996) *her5* (Reifers et al., 1998) *her2, her3, her4.2, her6* (Takke et al., 1999) *hey1* (Kudoh et al., 2001) *her7* (Oates and Ho, 2002) *her8a, her9* (Gajewski and Voolstra, 2002) *hes6* (Yoda et al., 2003) *her11 ,her12, her13* (Sieger et al., 2004) *her4.3* (Harden et al., 2006) *her15(Shankaran et al., 2007)*

* - Previously named *hairy2a*; ** - Previously named hairy1; *** - Previously named *esr9* ; **** - Previously named *esr2;***** Previously named *hairy2b*;******

1.4. SIGNALLING PATHWAYS REGULATING CLOCK GENE OSCILLATIONS

In addition to the oscillation of *hes-related* (*hairy/Enhancher of split*) genes, other Notch, Wnt and FGF pathway components have been identified to have a similar dynamic expression in the PSM (Andrade et al., 2007; Gibb et al., 2010). These genes were identified both by classical method (Andrade et al., 2007) and by microarray analysis in mouse, chick and zebrafish (Dequéant et al., 2006; Krol et al., 2011) (Table 1.3).

Table 1.3 - Comprehensive presentation of the PSM oscillatory genes belonging to the Notch, FGF and Wnt signalling pathways in mouse, chick and zebrafish.

	Mouse (82 genes in total)	Chick (182 genes in total)	Zebrafish (24 genes in total)
Notch	*lfng* (Forsberg et al., 1998) *hes1* (Jouve et al., 2000) *hey2 (Leimeister et al., 2000)* *hes7 (Bessho et al., 2001)* *hes5* (Dunwoodie et al., 2002) *nkd1* (Ishikawa et al., 2004) *nrarp* (Dequéant et al., 2006) *hey1, Id1, Efna1* (Krol et al., 2011)	*hairy1* (Palmeirim et al., 1997) *lfng* (McGrew et al., 1998) *hairy2* (Jouve et al., 2000) *hey2 (Leimeister et al., 2000)* *nrarp (Wright et al., 2009)*	*deltaC* (Jiang et al., 2000) *her1* (Holley et al., 2000) *her7* (Oates and Ho, 2002) *nrarp* (Wright et al., 2009) *her15, her2, her4* (Krol et al., 2011)
Wnt	*axin2 (Aulehla et al., 2003)* *dact1* (Suriben et al., 2006) *dkk1, sp5, myc, tnfrsf19, cyr61,shisa2 (Krol et al., 2011)*	*axin2, T, gpr177, rrm2* (Krol et al., 2011)	*tbx16* (Krol et al., 2011)
Fgf	*dusp6, spry2* (Dequéant et al., 2006) *snail1* (Dale et al., 2006) *dusp4 (Niwa et al., 2007)* *spry4 (Hayashi et al., 2009)*	*snail2* (Dale et al., 2006) *raf1, Erk, dusp6* (Krol et al., 2011)	

Dequéant's microarray showed not only a large number of genes cycling in mouse PSM, but also that the FGF and Notch components were activated in parallel and were cycling in phase, while genes belonging to the Wnt pathway exhibit an opposite cycling phase to the Notch/FGF pathway components (Dequéant et al., 2006). These signalling pathways seem to interconnect and constitute a complex oscillating signalling network involved in somitogenesis (Aulehla and Pourquie, 2008).

To date, most of the genes identified as cycling belong to Notch signalling pathway, and among these, it seems that only hairy-related genes have the property of oscillating universally among various species (Bessho and Kageyama, 2003).

Currently, with the knowledge gathered it is possible to observe a great interconnection of these Notch, FGF and Wnt signalling pathways. However, the search for the upstream pacemaker of the segmentation clock remains in open, and is essential to reach a full understanding of the nature and biological relevance of the somitogenesis clock (Andrade, et al 2007).

1.4.1. ROLES OF NOTCH SIGNALLING PATHWAY

Notch signalling is an intercellular communication pathway mediated by the interaction between the Notch receptors (Notch 1-4) with their ligands (Delta or Serrate1 & 2/Jagged1 & 2) by direct cell-cell interaction. When binding occurs, a multiprotein complex containing Presinilin1/ Presinilin2 performs a proteolytic cleavage of the Notch Intracellular Domain (NICD) (Fortini, 2002). NICD is then released from the membrane region and able to translocate into the nucleus, where it converts the DNA-binding transcription factor RBPjk from a repressor to an activator, forming a complex with it, and activates the transcription of its target genes, such as *hes* and *hey* genes (Kageyama et al., 2010; Shimojo et al., 2008).

The mechanism by which cell intrinsic cyclic expression is generated has been deeply analyzed. Strikingly, autoregulatory feedback loops seem to be one of the main mechanisms by which the oscillations of these genes activity are generated, alongside with short lived mRNA and proteins. In fact, several downstream targets

with oscillatory expression are known inhibitors of the pathways that induce their expression. Such simple feedback mechanisms can generate oscillations, provided the presence of a delay in the process (Aulehla and Pourquie, 2008). Dale, et, al (2003) first identified a negative feedback loop acting in the chick PSM, in which *lfng* (Lunatic fringe is a Notch target gene encoding glycosyltransferase that modulates Notch signalling will act, in turn, to inhibit Notch signalling and thus regulate *lfng*'s own expression. In this model, the NICD domain, after translocated to the nucleus, will activate *hairy1*, *hairy2* and *lfng* genes. Lnfg protein modifies Notch, which becomes less sensitive to activation by Delta, yet it's a transient and periodic effect due to the short life and rapid turnover of the Lfng protein. Oscillations are thus generated by alternation between activation of *lfng* expression and repression of Notch by Lfng. The influence of NICD on *lfng* exerted via *hairy* genes will be delayed relative to the influence exerted directly, and this phase shift can, in principle, allows the two influences to act in synergy (Giudicelli and Lewis, 2004).This study proposed that this negative feedback loop involving these genes represents a core component of the avian segmentation clock mechanism (Giudicelli and Lewis, 2004) (Figure 1.3).

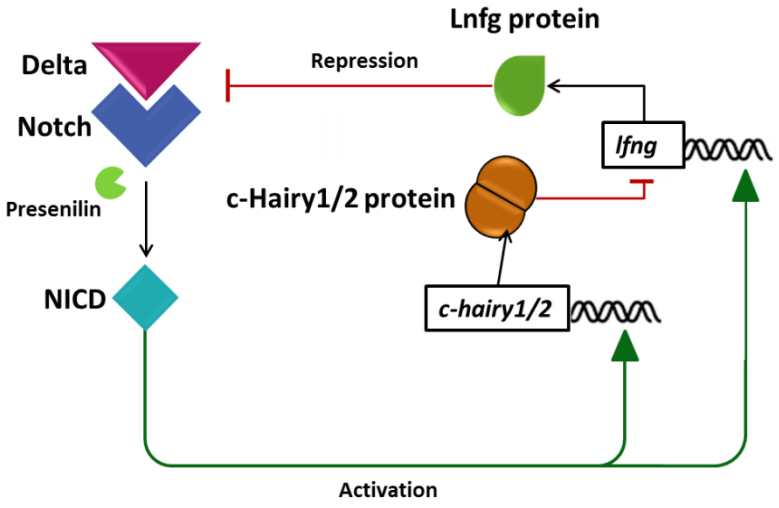


Figure 1.3 – Model of Notch, c-Hairy1/2 and Lnfg feedback loop mechanism. Notch signalling activates cyclic gene transcription of *lfng and c-hairy1/2.* Lnfg protein will in turn modulate Notch, which becomes less sensitive to activation by Delta, inhibiting Notch signalling. The proteins c-Hairy1/2 will inhibit *lfgn* transcription, having an indirect influence on *lfng* via Notch signalling. Adapted from (Giudicelli and Lewis, 2004).

Although generating gene oscillations is an intrinsic property of individual PSM cells and tissue (Masamizu et al., 2006; Palmeirim et al., 1997), several studies demonstrated an out of synchronicity of *hes1* oscillations in dissociated PSM cells, and a necessity of cell-cell contact for synchronized oscillations (Maroto et al., 2005; Masamizu et al., 2006). Hirata, et. al (2002) demonstrated a Delta1 mediated synchronization of *hes1* cycles, which are evident as mouse cell lines exposed to Delta1 expressing cells maintained a synchronous expression of *hes1*, which otherwise presented out of synchrony expression, pointing Notch signalling responsible in synchronizing gene oscillations between neighbouring PSM cells (Jiang et al., 2000). Many other studies in zebrafish either using genetic or chemical inhibition of Notch signalling, or even mathematical modelling, suggested that DeltaC-Notch intercellular interaction synchronize PSM cells oscillations by reducing their internal noise (Dequeant and Pourquie, 2008; Giudicelli et al., 2007; Horikawa et al., 2006; Riedel-Kruse et al., 2007). Notch signalling in zebrafish somitogenesis has been solely assigned to synchronize oscillations of neighbouring cells (Ozbudak and Lewis, 2008). More recently, was observed a delay on the zebrafish segmentation clock by a disruption of Delta-Notch coupling, extending its periodicity and reveled an additional role of Notch signalling in clock period regulation (Herrgen et al., 2010).

1.4.2. ROLES OF FGF SIGNALLING PATHWAY

FGF signalling is another important pathway for the segmentation clock, as it has been demonstrated that it is crutial for the initiation of the expression of *dusp4*, a negative regulator of FGF/MAPK signalling and *hes7*, in the posterior PSM, which is required for their dynamic expression in the anterior PSM (Niwa et al., 2007). FGF signal initiates *hes7* expression as well as oscillation in the posterior PSM, which is propagated and maintained in the anterior PSM by Notch signalling, implementing FGF signalling as the base for Hes7 generated oscillations (Bessho et al., 2003). While the influence of the FGF signalling pathway on the segmentation clock appears to be a controversial topic, an interesting finding shows that FGF components (dusp6, spry2, snail1) does not oscillate in

presenilin1/ presenilin2 double mutant mice lacking NICD, indicating that Notch signalling is necessary for FGF component oscillations (Ferjentsik et al., 2009). Other intriguing study demonstrated that chemical inhibitors of FgfR1 failed to perturb Notch target gene expression in mouse and chick (Delfini et al., 2005; Dubrulle et al., 2001; Gibb et al., 2009; Niederreither et al., 2002), whereas, Wahl et al. (2007) observed a quick disruption of Wnt target gene *axin2* and a slow, but abrogated *lfng* expression in mouse FGFR1 mutants. Despite the constitutive expression of *Dusp4* in hes7 mutants, axin2 still presents dynamic expression, ruling out the possibility of FGF/Dusp4 signalling to be the phasemaker of gene oscillations (Niwa et al., 2007; Hirata et al., 2004).

1.4.3. ROLES OF WNT SIGNALLING PATHWAY

Until recently it was thought that, although Wnt cluster genes oscillate in mouse PSM, there was no indication of these oscillations in chick PSM. However, chick Wnt components that included *axin2, T, gpr177* (commonly known as Wntless) and *rrm2* and zebrafish: *tbx16* (Krol et al., 2011), were identified for the first time in a genome-wide study as being oscillating (Table 1.3). A linkage between Notch signalling and Wnt targets, was disbelieved since it has been already described that, in *hes7,* RBPjk and NICD gain of function mutants, *axin2* retains dynamic expression pattern (Hirata et al., 2004; Feller et al., 2008; Ferjentsik et al., 2009). However, abolishing the expression of *presenilin1, and presenilin2* in mouse PSM, *axin2* oscillations ceased to exist (Ferjentsik et al., 2009) indicating that Notch signalling is operating upstream of Wnt targets. Yet, other results indicate Wnt signalling to be upstream of Notch, since *wnt3a* hypomorphic mouse mutants (*vestigial tail* with less *wnt3a* expression in the tailbud) present downregulated *dll1, notch1* and non-dynamic *lfng* and *hes7* expression in the posterior PSM (Aulehla et al., 2003; Niwa et al., 2007). This cross-talk between Notch and Wnt pathways was considered in Gibb et al. (2009), wherein it was observed a reciprocal regulation of Notch and Wnt pathways upon their components since, the inhibition of any of these pathways affected both

component genes suggesting a mutual interaction between Notch and Wnt signalling in chick as previously observed in mouse PSM (Ishikawa et al., 2004).

1.5. HES GENE FAMILY

Since the initial description of the segmentation clock, multiple other cyclic genes have been shown to oscillate at the mRNA level in the PSM of chick, mouse (*Mus musculus*) and zebrafish (*Danio rerio*) embryos (Gibb et al., 2010). Many of these genes belong to the Hes gene family, in which *c-hairy1* is included, and are described as targets of Notch pathway that act as transcriptional repressors, regulating cell proliferation and differentiation (Kageyama et al., 2007; Vasiliauskas et al., 2003). Hes family genes are implicated in a number of developmental processes such as segmentation of the mesoderm, functioning as biological clocks, measuring time (Palmeirim et al., 1997). The molecular clock has also been showed to be operating during limb development in the chondrogenic precursor cells (Pascoal et al., 2007). Additionally, these repressor genes have been described to play a role in embryogenesis by maintaining the undifferentiated state of progenitor cells, and regulating the binary cell fate decisions (Kageyama et al., 2007).

1.5.1. STRUCTURAL ANALYSIS OF HES FAMILY PROTEINS

Structurally, Hes family proteins contain three conserved domains which endow them with unique features as repressors and oscillators. These are the basic Helix-Loop-Helix (bHLH), Orange and WRPW domains (Kageyama et al., 2007) (Figure 1.4).

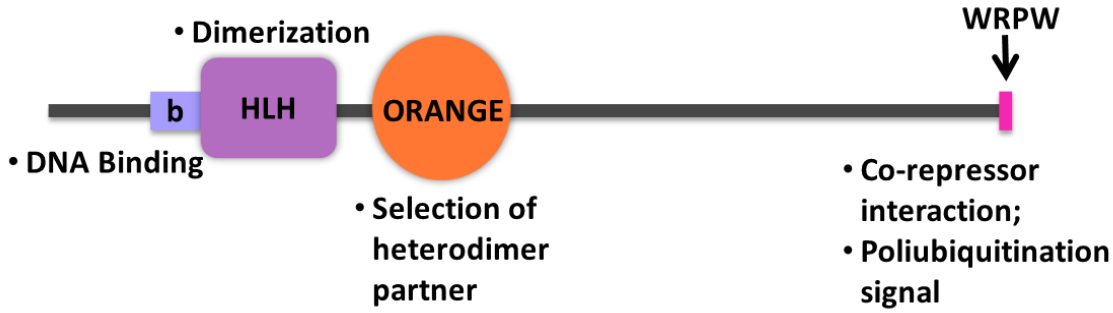


Figure 1.4 – Schematic representation of Hes family proteins conserved domains, and their main functions. Hes family proteins have three conserved domains, the basic Helix-Loop-Heix, Orange and WRPW domains. The basic moitif (b), is responsible for DNA binding, the helix-loop helix (HLH) domain promotes dimerization, the Orange domain acts as the selector of the heterodimer partner which Hes protein will bind, and WRPW is responsible for interaction with Hes co-repressors, and acts a poliubiquitination signal.

The bHLH domain can be distinguished in two regions, the basic region, and the Helix-Loop-Helix domain (Kageyama et al., 2007). The basic region, located at the N-terminal end of the domain consists of approximately 15 amino acids with a high number of basic residues. This motif is present in the majority of proteins containing the HLH domain, facilitating binding to the DNA that contains the canonical 'E box' recognition sequence, CANNTG (Massari and Murre, 2000; Toledo-Ortiz et al., 2003). Contemplating the amino acids within the basic region of the protein, it has been shown that some of them provide recognition of the core consensus site, whereas other residues determine specificity for a given type of hexanucleotide sequence E-box. Hes factors have a highly conserved proline residue located in the middle of basic region, suggesting that this may confer DNA-binding specificity, although the exact role has not yet been established (Dawson et al., 1995; Sasai et al., 1992).

The HLH domain is constituted mainly of hydrophobic residues (Toledo-Ortiz et al., 2003), which comprise two amphipathic α-helices, each 15-20 residues long, separated by a shorter intervening loop that has a more variable length and sequence, which mediates protein dimerization with homo and heterodimers (Norton, 2000; Toledo-Ortiz et al., 2003). Although the role of DNA binding has been assigned to basic region, evidences show residues in the loop and second helix also make contact with DNA (Massari and Murre, 2000).

The bHLH proteins, in animal systems, have been classified into six main groups, in which its classification is based on the evolutionary origin and sequence relatedness as well as the information available on their DNA binding specificities and functional activities (Atchley and Fitch, 1997; Ledent and Vervoort, 2001). Hes family proteins are classified as belonging to group E (previously considered part of group B by Atchley and Fitch [1997]), distinguished from other groups based on the presence of several conserved amino acids flanking the bHLH and the presence of the WRPW peptide (Ledent and Vervoort, 2001). The members belonging to this group are also considered to preferentially bind to N-boxes (CACGGC or CACGAC), having low affinity for E-boxes, as well as possessing a Proline instead of an Arginine residue at a crucial position in the bHLH domain (Fisher and Caudy, 1998)

The orange domain has two amphipathic helices and regulates the selection of bHLH heterodimer partners being particularly important for the efficiency of the interaction (Kageyama et al., 2007; Taelman et al., 2004) This domain is also shown to mediate transcriptional repression, although a co-repressor interacting with this domain is not known yet (Castella et al., 2000) .

Localized at C-terminus is a well-defined simple motif, the WRPW, and it consists of the tetrapeptide Trp-Arg-Pro-Trp. This motif is responsible for binding to the transcriptional corepressor Groucho and its mammalian homologues, the TLE (Transducin-like E(spl)) proteins, and thereby functions as the repressor domain of this family of transcription factors. WRPW is not, however, the only motif required for repression, since functional dissection of the HES proteins revealed other important regions for repression, such as the Orange domain. In *Drosophila* deletion mutants for the gene *Enhancer of split [E(spl)],* it has been demonstrated that both the Orange and the interval between the Orange and the WRPW motif appear to be important to the repression function of the gene (Giebel and Campos-Ortega, 1997). WRPW sequence also acts as a polyubiquitination signal, controlling the half-life of Hes proteins by promotion of proteasome-mediated

degradation, which results in very short half-lives, 22 min for both Hes1 (Hirata,2002), and Hes7 (Hirata, 2004).

1.5.2. HES AND HAIRY-RELATED TRANSCRIPTION FACTORS

Hes factors are similar to but distinct from Hairy-related transcription factors (Hrt), also known as Hesr, Hey, CHF, grl, and HERP. For instance, Hes family members are characterized by a highly conserved proline residue in the basic domain, that contrast a glycine at the comparable position that likely underlies Hrt specificity for E box DNA-binding sites (CACGTG) over the N-box site (CACNAG) favored by Hes family members. In addition, Hrt proteins contain a carboxy-terminal YXPW-TEI/VGAF (Y/T) motif that is alike to but distinct from the WRPW motif of Hes1. This region is necessary for Hes1's recruitment of the co-repressor, Groucho but the Y/T domain of Hrt proteins does not appear to interact with Groucho. (King et al., 2006)

1.5.3. EXPRESSION AND TRANSCRIPTIONAL ACTIVITIES OF HES FACTORS

Hes family oscillation seems to be generated by an auto feedback loop, as it was shown by Hirata, et al. (2002). This study demonstrates that the mouse Hes1 protein directly binds to the N-box in its own promoter repressing the gene transcription. This repression is transient probably due to the short half-life of both *hes1* (22 min) mRNA and Hes1 protein (Figure 1.5) (Hirata et al., 2002). Data suggest that the 3´untranslated region (3´-UTR) of *hes1* gene might be responsible for the short half-life of *hes1* mRNA (Hirata et al., 2002), as it has been previously reported for Xenopus *hairy2* mRNA stability (Davis et al., 2001). Both the protein and mRNA maintain a 2 hour oscillation period, although the protein oscillation is delayed ~15min compared to mRNA, probably due to the time required for protein turnover by the ubiquitin–proteasome pathway (Hirata et al., 2002) (Figure 1.5). When Hes1 protein is constitutively activated either by using proteasome inhibitors or an expression vector, *hes1* transcription is repressed by the persistently high

Hes1 levels. On the other hand, treatment with cyclohexamide, an inhibitor of translation, or over-expression of a dominant-negative form of Hes1 (dnHes1) constitutively upregulates *hes1,* blocking its oscillations. Therefore, *hes1* oscillations requires both *de novo* synthesis and degradation of Hes1 protein and the negative feedback loop, in which Hes1 periodically represses its own transcription, is the central mechanism for the *hes1* oscillations both in cells and in the PSM (Hirata et al., 2002) (Figure 1.5). Furthermore, using mathematical modeling, the authors predict that alterations in synthesis and degradation rates should change the oscillations period.

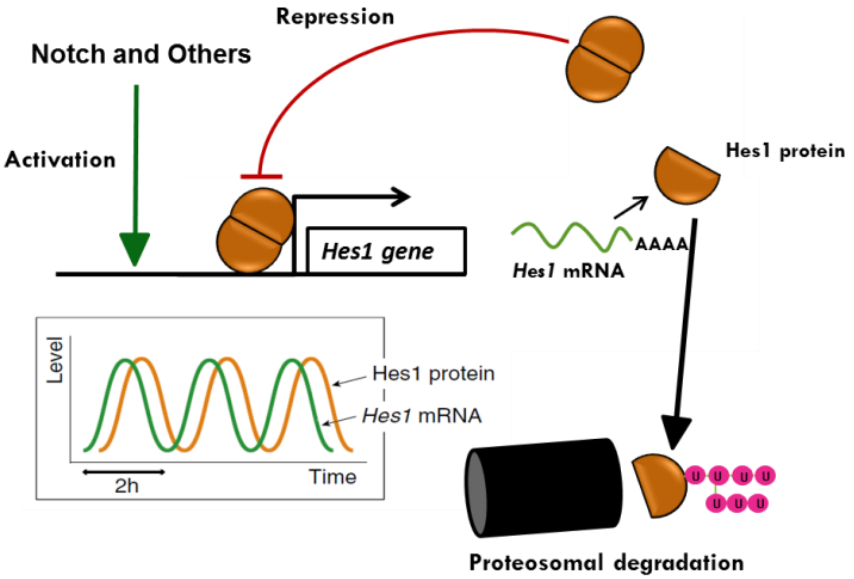


Figure 1.5 – Schematic representation of Hes auto feedback regulation. When *hes1* transcription activated by Notch, for example, Hes1 protein is produced, allowing it to form dimers. The homodimer then will induce the repression of its own gene. Both the proteins and mRNA are degraded rapidly due to their very short half-lives, therefore allowing the start of a new cycle of expression. Adapted from (Kageyama et al., 2007).

Hes factores repression is carried out in two manners, it can repress actively or passively (Figure 1.6). In the active repression, the WRPW domain is involved, and will interact with the co-repressor Transducin-like E(spl) (TLE) genes/Groucho-related gene (Grg) (Kageyama et al., 2007) (Figure 1.6A). Groucho is able to repress transcription by recruiting histone deacetylase, an enzyme responsible for inactivate chromatin (Chen and Courey, 2000). Thus, it is likely that the Hes–Groucho-homolog complex represses transcription by inactivating chromatin.

Studies reveal an interaction between the C-terminal portion of chick hairy1 and Sap18 (Sin3-associated polypeptide, 18 kDa), a component of the Sin3/histone deacetylase (HDAC) transcriptional repressor complex, indicating that in c-hairy1 may mediate gene transcriptional repression by recruiting the Sin3/HDAC complex through a direct interaction with the Sap18 adaptor molecule (Sheeba et al., 2007).

The passive repression is due to the ability of Hes factors to inhibit bHLH activators that bind to the E box, by forming non-DNA-binding heterodimers with them. These activators include Mash1 and E47 that normally activate neuronal-specific gene expression by forming heterodimers and biding to the E box (Figure 1.6C), a process prevented by intervention of Hes1 (Kageyama et al., 2007) (Figure 1.6B).

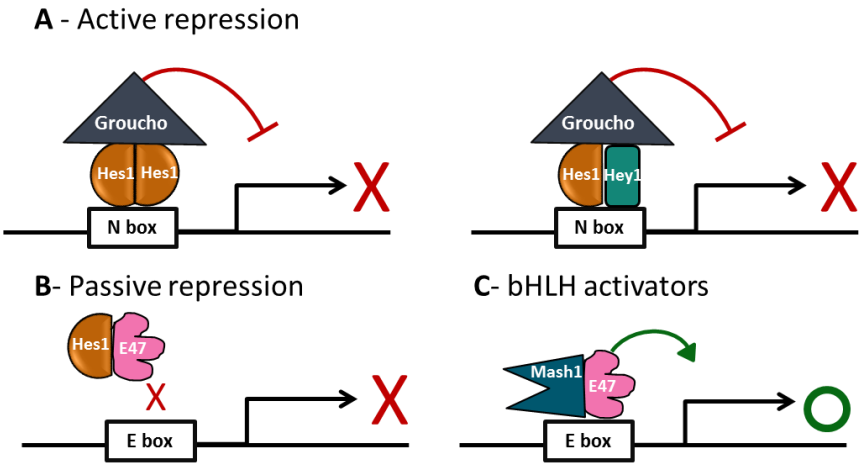


Figure 1.6 - Active and passive repression of Hes1. A – Active repression: Hes factors are able to actively repress transcription, by binding to the N box, in form of homodimers (left panel) or heterodimers with Hey (right panel), interacting with co-repressors, such as Groucho homologs. (**C**) Passive repression: By forming non-DNA binding heterodimers with bHLH activators such as E47, Hes factors are able to passively inhibit transcriptional activation (**D**) Activation: bHLH activators such as Mash1 and E47 form heterodimers that bind to the E box and activate transcription. Adapted from (Kageyama et al., 2007).

1.5.4. HES FACTORS MISREGULATIONS

Misregulations of HES family members have been linked to developmental defects and oncogenesis (Andrade et al., 2007; Davis et al., 2001). Andrade et al. (2007) reviews some development defects associated the segmentation clock, in mutant mouse embryos and in human embryos with congenital malformations in

the axial skeleton. Mouse mutant models can be very useful to investigate the cause of some malformations observed in humans. The most common defect observed in these embryos mutated for segmentation clock genes, when not lethal, are shorter trunks with fused or bifurcated vertebrae and ribs (reviewed in Andrade et al., 2007). Similar segmentation problems can be found in human spondylocostal dysostosis (SD), in which patients exhibit a short trunk due to multiple hemi-vertebrae formation accompanied by rib fusions, bifurcations and deletions and mutations in Notch-regulated genes such as *dll3*, *lfng* and *mesp2* have been observed to induce SD (Giampietro et al., 2009; Turnpenny et al., 2007). Alagille syndrome, a disorder characterized by developmental abnormalities of the liver, heart, eye and skeleton, has been associated with mutations in *jagged1*, a Notch ligand, and *notch2* genes. Mutations of *fgfr1-3* in human have also been shown to result in skeletal disorders, including fusion of the craniofacial sutures and short-limbed dwarfisms (reviewed in Chen and Deng, 2005).

While expanding of the understanding of the mechanisms regulated by Hes-family members the number of observations associating them with oncogenesis also increased. It is, however, controversial whether Hes genes act as a tumor suppressor and/or oncogene, since conflicting observations suggest that these genes may have a dual role depending on the tumor types and the stages of cancer progression (Lee et al., 2012). In the most common of malignant brain tumors, medulloblastomas (MBs), expression of HES1 has been associated with the worse clinical outcome. Down-regulation of Hes1 expression negatively regulates the proliferation rate and anchorage-independent growth of MB cells (Garzia et al., 2009). Studies also suggest that Hes1 may have an oncogenic role in oral squamous cell carcinoma (OSCC), associated with cancer progression and cancer stem cells phenotype in OSCC (Lee et al., 2012). Beyond Hes1, Hes6 adds up to the growing list of Notch signalling pathway components involved in the process of transformed cell growth, being identified as up-regulated in colon, breast, lung ,and renal, primary carcinomas compared to expression in normal tissue (Swearingen et al., 2003).

1.6. C-HAIRY1 PROTEIN AND ANTIBODY PRODUCTION

Since the discovery of the oscillatory expression of *c-hairy1* in the chick embryo PSM, many of the studies have been much focused on the understanding of the gene expression mechanisms rather that in the protein mechanisms (Andrade et al., 2007; Palmeirim et al., 1997). In 2003, Vasiliauskas et al., identified an expression of two products in a RT-PCR amplification of the entire *c-hairy1* open frame from stage 19 chick embryos, which led to the discovery of a c-Hairy1 isoform, c-Hairy1B, since c-Hairy1A have already been described by Palmeirim et al in 1997. The isoforms differed in a insertion of 14 amino acids between the second and third residues belonging to basic region, resulting on a difference in the first two positions of the basic domain, RK (two basic amino acids) in c-Hairy1A, and AE (a hydrophobic and an acid residue) in c-Hairy1B. This change alters the net charge of the c-Hairy1B from negative to positive, which the authors suggest a possibility to change significantly the specificity or weaken the affinity of a transcription factor for DNA. Consequently, the two isoforms of c-Hairy1, by forming homo- or heterodimers, may bind a range of DNA sites. On the other hand, c-hairy1B could antagonize the transcriptional repressor function of c-hairy1A. The study described different functions of these isoforms in regulating the limb bud growth, suggesting the differences were due isoforms different activities rather than to the efficiency.

With the purpose to acquire better understanding about the c-Hairy1 protein, regarding expression, function, and its cellular localization in chicken, our lab ordered the production of an antibody against c-Hairy1 from a commercial company. This task was necessary since there was no antibody available in the market, nor has any other research group developed it before. In order to validate the future results it was crucial to have available the purified protein so it could be used as a positive control, and validate the results obtained. Therefore, a commercial company undertook the task of purifying the protein. The process consisted of the cloning of the c-Hairy1 cDNA fused to a His tag located in the N-

terminal, into the Invitrogen vector pCRII-TOPO vector, and then expressed in *E. Coli*. The protein was purified by His Tag pull-down assay, and removed the His tag from c-Hairy1. The protein was expected to be 90% pure and to have a molecular weight of 30,67 kDa.

At first, a rabbit polyclonal antibody against the c-terminal portion of the protein was produced, and by western blot was verified that the antibody was sensitive for c-Hairy1. In this test, bands with the expected molecular weight were detected in chicken extracts, using the purified protein as a positive control. Controversially, several bands of different molecular weights were also detected in both extract and the purified protein. The most intense bands in the extract corresponded to approximately 32, 49, 63, 72 and 75 kDa, whereas in the purified protein the predominant bands corresponded to 15, 35 and 39 kDa. Results suggested that a low specificity of the polyclonal antibody (Figure 1.7).

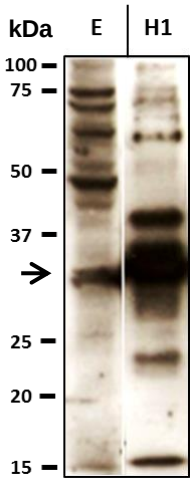


Figure 1.7 - Test of the polyclonal antibody against c-Hairy1. Polyclonal anti-c-Hairy1 was tested in western blot, using as samples chicken embryo extract, E, (15 µg of protein/ lane), and purified c-Hairy1 protein (140 ng of protein/ lane), H1. Several bands were detected, including a band with the predicted c-hairy1 molecular weight, indicated with an arrow.

Regarding these results, a monoclonal antibody against the whole protein was produced in an attempt to increase specificity, as it only has affinity for one epitope (Alberts et al., 2007). Posteriorly serum from several different immunized mice, each specific for one epitope, were sent to be tested by western blot. Results reveal different staining for each serum, in which some presented a stronger staining than others. The serums were selected based on the presence of the predicted molecular weight bands in the samples that were tested, and presented

fewer unspecific bands with unpredicted molecular weights (Figure 1.8). Taking these conditions in considerations, the serums with the reference 8C9G4A10C11 (8C9) and 5C6F11E12 (5C6) were selected, as they detected the predicted band of c-Hairy1 in both embryo extracts and purified protein lanes. Due the fact that 8C9 serum detected fewer unspecific bands than 5C6, it was considered to have more specificity to c-Hairy1 protein. Thus 8C9 was developed as a monoclonal antibody, however the antibody referenced as 5C6 was asked to be kept by the company to possibly validate future studies.

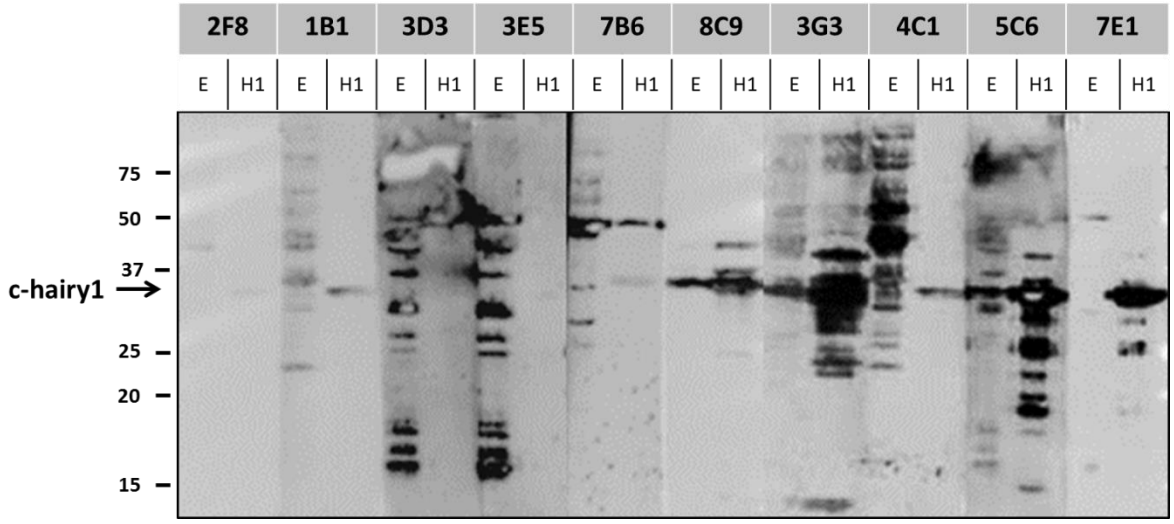


Figure 1.8 - Test for specificity of several hybridomas supernatants against c-Hairy1, in chick embryo extracts and purified protein. Different hybridomas supernatants anti-c-hairy1 were tested in western blot, using as samples chick embryo extract, E, (15 µg each), and purified c-Hairy1 protein, H1 (140 ng each). Each serum reacts against different epitopes, and are distinguished by their own reference: 2F8F6G8 (2F8); 1B11A9F12 (1B1); 3D3A7F7 (3D3); 3E5C6H9 (3E5); 7B6B10H1 (7B6); 8C9G4A10C11 (8C9); 3G3E5G10 (3G3); 4C1G1G10 (4C1); 5C6F11E12 (5C6); 7E10F11E11 (7E1).

1.7. AIMS

The main goal of the present work was to characterize the chick Hairy1 protein expression and study its subcellular localization, first characterized at a bioinformatics level, taking advantage of computational biology tools to study its protein sequence, which was expected to provide a starting point and a solid base to further investigate c-Hairy1 protein expression *in vitro*, correspondent to the second goal of this work.

The specific aims of this work were:

Bioinformatic analysis:

- Study of the protein sequence in terms of primary and secondary structure, also predicting post-translational modifications and subcellular localization signals.

***In vitro* study:**

- Test for the specificity, sensibility and efficiency of the previously produced monoclonal antibody against c-Hairy1, establishing all the optimal parameters for its use on western blot and immunofluorescence.
- Study the expression of c-hairy1 protein in chick embryo and chicken embryonic fibroblasts (CEFs), and its sub-localization in the cells, using the optimized techniques.

CHAPTER 2

MATERIALS AND METHODS

2. MATERIALS AND METHODS

2.1. BIOINFORMATIC APPROACH

2.1.1. SEQUENCES SEARCHES

The c-hairy1 DNA and protein sequences were obtained from the National Center for Biotechnology Information database (http://www.ncbi.nlm.nih.gov/). The chosen sets of sequences were manually verified and a filtration process was proceeded in order to eliminate duplicated sequences and therefore reduce redundancy.

A BLASTP search (Altschul et al., 1997) against the c-Hairy1A protein sequence was performed, obtaining a total of one hundred sequences producing significant alignments and displayed by percentage of maximum identification. For the sequences of interest selection, three conditions were considered: highest similarity; one protein by species; the species must be relevant for developmental biology.

2.1.2. SEQUENCES ALIGNMENT

Sequences were aligned with T-Coffee (http://tcoffee.vital-it.ch/cgi-bin/Tcoffee/tcoffee_cgi/index.cgi?stage1=1&daction=TCOFFEE::Regular) in a regular computational mode, and saved in ClustalW format. ESPript (Easy Sequencing in Postscript, http://espript.ibcp.fr/ESPript/ESPript/), was used to edit these sequence presenting graphical enhancements to the Clustalw outputs.

2.1.3. PHYLOGENETIC ANALYSIS

The phylogenetic analysis was performed by Phylogeny.fr (http://www.phylogeny.fr/), a web server dedicated to reconstructing and analyzing phylogenetic relationships between molecular sequences. Phylogeny.fr runs and connects various bioinformatics programs to reconstruct a robust phylogenetic tree from a set of sequences. This tree has been assigned with genetic distance values, correspondent to the genetic divergence between species. Smaller genetic

distances indicate close genetic relationship whereas large genetic distances indicate a more distant genetic relationship (Dereeper et al., 2008).

The phylogenetic tree was generated by the program One Click, in which the protein sequences were submitted and analyzed in default parameters.

2.1.4. ASSESSMENT OF THE PHYSICO-CHEMICAL PROPERTIES

ExPASy ProtParam, (http://web.expasy.org/protparam/) predicted the physic-chemical properties of c-Hairy1 isoforms. The software calculated several parameters, including the molecular weight, theoretical isoelectric point (pI), amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY). For this study, the parameters considered were the amino acid composition, molecular weight, theoretical pI and GRAVY.

2.1.5. DOMAIN DETERMINATION

To evaluate the conserved domains, or functional sites present on the studied proteins, was used ExPASy Prosite (http://prosite.expasy.org/), in default parameters. Prosite is based on an annotated collection of motif descriptors dedicated to the identification of protein families and domains (Sigrist et al., 2002).

2.1.6. PHOSPHORYLATION SITES AND KINASE PREDICTION

ExPASY's NetPhos, version 2.0 (http://www.cbs.dtu.dk/services/NetPhos/) and the DISPHOS (Disorder-Enhanced Phosphorylation Sites Predictor, version 1.3, http://www.ist.temple.edu/DISPHOS) predicted the possible phosphorylation sites present on c-Hairy1 sequence. Both softwares predicted serine, threonine and tyrosine phosphorylation sites, and confered a score, ranged from 0 to 1, giving the information of the probability of those sites to be phosphorylated. The accuracy of DISPHOS reaches 81.3% +/- 2.2% for Serine, 74.8% +/- 2.5% for Threonine, and 79.0% +/- 2.4% for Tyrosine, whereas NetPhos ranges its sensitivity from 69 to

96% (Blom et al., 1999; Iakoucheva et al., 2004). NetPhos was used with default parameters, and DISPHOS was selected the functional category, regulation.

ExPasy's NetPhosk ((http://www.cbs.dtu.dk/services/NetPhosk/) was used to predict the kinases with affinity for a possible phosphorylation site. Currently NetPhosK covers the following kinases: cyclic AMP dependent protein kinase (PKA), protein kinase C (PKC), cyclic GMP-dependent protein kinase (PKG), casein kinase 2 (CKII), cyclin dependent kinase (cdk1) (Cdc2), Ca2+/calmodulin-dependent protein kinase II (CaM-II), ataxia telangiectasia mutated (ATM), DNA-dependent protein kinase (DNA PK), cyclin-dependent kinase 5 (Cdk5), p38 mitogen-activated protein kinase (p38 MAPK), Glycogen Synthase Kinase 3 (GSK3), casein kinase 1 (CKI), protein kinase B (PKB), ribosomal s6 kinase (RSK), insulin receptor (INSR), epidermal growth factor receptor (EGFR) and proto-oncogene tyrosine-protein kinase (Src) (Blom et al., 2004).

2.1.7. SECONDARY STRUCTURE PREDICTION AND SURFACE ACCESSIBILITY

Merilab's Jufo9D Server predicted the secondary structure from a protein sequence, and has an accuracy of 73.2% for three-state secondary structure prediction (http://www.meilerlab.org/index.php/servers/show?s_id=5). It calculated for each residue a probability for combinations of α-helices, β-strands and coil structures.

2.1.8. PREDICTION OF SUBCELLULAR LOCALIZATION SIGNALS

The prediction of protein localization sites in cells was carried out by NucPred (http://www.sbc.su.se/~maccallr/nucpred/cgi-bin/multi.cgi) combined with PSORT II (http://psort.hgc.jp/form2.html), in order to increase confidence of the results. In the table 2.1 are represented the specificities and sensitivities of the different predicting nuclear localization methods. These programs were tested on human protein sequences, non-predicted to be transmembrane proteins. The performance

of NucPred and PSORT II are roughly equivalent, though presenting different sensitivities (Table 2.1). (Brameier et al., 2007).

Table 2.1 – Specificities and sensitivities for different predictors of subcellular localization signal programs.

Method	Specificity	Sensitivity
NucPred (0.8 threshold)	0.615	0.307
NucPred (0.5 threshold)	0.480	0.626
PSORT II	0.466	0.697
BaCelLo	0.668	0.614

The NucPred results should be interpreted based on a score threshold, in which scores greater than or equal to this threshold are predicted to spend time in the nucleus. Higher thresholds have greater specificity and lower sensitivity, as shown in table 2.2.

Table 2.2 – Specificities and sensitivities correspondent to different NucPred score thresholds.

NucPred score threshold	Specificity	Sensitivity
0.10	0.45	0.88
0.20	0.52	0.83
0.30	0.57	0.77
0.40	0.63	0.69
0.50	0.70	0.62
0.60	0.71	0.53
0.70	0.81	0.44
0.80	0.84	0.32
0.90	0.88	0.21
1.00	1.00	0.02

PSORT II used a *k*-nearest neighbor (*k*-NN) algorithm for assessing the probability of localizing at each candidate sites. It assigned a percentage score corresponding to the probability of the protein to be nucleus localized (Horton and Nakai, 1997).

Nuclear export signals (NES) were predicted using NetNES 1.1 server, which search for leucine-rich nuclear export signals combining neural networks and hidden Markov models, assigning a score correspondent to the probability of being a NES (la Cour et al., 2004).

2.2. PREPARATION OF BIOLOGICAL MATERIAL

2.2.1. EGGS AND EMBRYOS

Fertilized chick (*Gallus gallus*) eggs were purchased from a commercial source, and until incubated, were maintained at 16ºC. Eggs were then placed in incubators at 38ºC with a water container, in order to humidify the environment and stabilize the temperature. The eggs were incubated for a specific time period, so the embryos reached the stages required, according the development table of Hamburger and Hamilton (1992).

2.2.2. CULTURE AND USE OF PRIMARY CHICKEN EMBRYO FIBROBLASTS

Chicken embryonic fibroblasts (CEFs) isolated from 14 day embryos, were maintained at 37 ºC in a humidified 5% CO2 environment and grown as monolayers in Dulbecco's modified Eagle's medium (31966-021, Gibco) supplemented with 10% fetal bovine serum (FBS; 10500, Gibco), 2mM of L-Glutamine (25030, Gibco) and 550 units of penicillin and 5500µg of an antibiotic-mixture, Penicillin Streptomycin, (15140, Gibco). Cells were grown to 75-100% optical confluence before they were detached by incubation with 0.05% trypsin-EDTA (25300, Gibco) for 5 min, at 37ºC. Afterwards cells were centrifuged for 5 min at 1000 rpm and resuspended in medium. Cells were cultured in 75 cm 2 culture flasks containing a final complete growth media volume of 10 mL.

Whenever necessary, the cell number was determined with the help of a haemocytometer.

2.3.MOLECULAR BIOLOGY

2.3.1. PROTEIN EXTRACTION

Cells were harvested, and washed twice in 10 ml of PBS to eliminate FBS residues. Supernatant was discarded and cells were resuspended with lysis buffer (Appendix I). The extract was stored at -20ºC.

Chick embryos of stages HH10-13, were collected and washed twice with PBS, and then added 10 µL/embryo of lysis buffer. Extracts were well homogenized with the help of a vortex, and stored at -20 ºC.

2.3.2. DETERMINATION OF PROTEIN CONCENTRATION

Protein concentration was assessed with the Thermo Scientific NanoDrop 2000c Spectrophotometer, and determined by Bradford assay, using Quick Start Bradford Protein Assay reagent from Biorad. As determined by the reagent manual, 50:1 reagent / sample volume ratio was used to detect a concentration range between 0.125 mg/mL to 1.0 mg/mL, which corresponds to the detection limit of the reagent.

The Bradford assay requires a standard curve before sample proteins can be measured. A standard curve was generated using a serial dilution of BSA (Bovine Serum Albumin) ranging from 0,0675 mg/mL to 1 mg/ml, represented in the figure 2.1 with its values discriminated in the table 2.3.

Table 2.3 – Measurement of absorbance of serial dilutions of BSA.

Measure	mg/mL	Avg Abs.	#1	#2	#3
Standard 1	0,0675	0,009	0,011	0,006	0,010
Standard 2	0,1250	0,019	0,020	0,019	0,018
Standard 3	0,2500	0,038	0,038	0,040	0,037
Standard 4	0,5000	0,067	0,067	0,068	0,067
Standard 5	1,0000	0,112	0,116	0,111	0,110

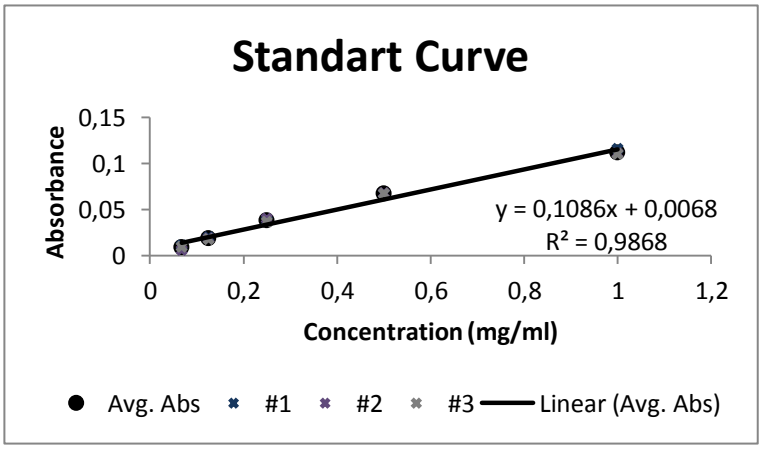


Figure 2.1 – Graphical representation of the standard curve of BSA dilutions. Graph was created using measurements of three replicates of different dilutions of BSA protein, contained in table 2.3. The linear equation was determined using the average absorbance of these values, and corresponds to y=0,1086x+0,0068, with an R^2=0,9868.

Samples were diluted in order to be contained in the standard curve, and three replicates of 1 µL of sample were diluted in 50 µL of Bradford reagent, waiting for 5 minutes, after which, 2 µL of the solution is pipetted onto a measurement pedestal. The final concentration was determined by the mean of the results, and multiplied by the dilution factor.

2.3.3. SDS-PAGE (SODIUM DODECYL SULFATE-POLYACRYLAMIDE GEL ELECTROPHORESIS) AND WESTERN BLOTTING

Samples were prepared to a determined amount of protein, with 10% of loading buffer (Appendix I). Then they were denatured for 10 minutes at 95ºC and loaded onto a 10% polyacrilamide gel (Appendix I). In order to monitoring electrophoretic separation and determine the molecular weight of the bands, a 5 µL of dual color marker was also loaded in the gel (#161-0374, Biorad Precision Plus Protein Dual Color Standards). The electrophoresis was carried out by a Mini-PROTEAN Tetra Electrophoresis System (Bio-Rad) at 15mA/gel for the run in the stacking gel and then raised to 30mA/gel.

After electrophoresis, the proteins on the gel were transferred to a PVDF membrane (88518, Thermo Scientific), sandwiched between three chromatography paper sheets on each side soaked in transfer buffer (Appendix I). The transfer was carried out by a Trans-Blot SD Semi-Dry Transfer Cell (170-3940 Bio-Rad), at 200mA, for 45 min.

Blocking of non-specific binding was achieved by placing the membrane in a10% (w/v) solution of non-fat milk in PBS 0.1% Tween 20 (PBSw).

For protein detection, the membrane was incubated with the primary antibody in 1% non-fat milk in PBSw, overnight, at 4 ºC, with shaking. After rinsing the membrane, three times, with 1% non-fat milk in PBSw, to remove unbound primary antibody, the membrane was incubated with the secondary antibody conjugated with horseradish peroxidase (HRP) (Appendix II) and diluted at a specific concentration in 1% non-fat milk in PBSw. The incubation lasted for 1 hour, at

room temperature, with shaking. After rinsing the membrane, the reactivity of HRP was developed by Biorad Immun-Star WesternC Chemiluminescent kit (more sensitive) (#170-5070, Biorad), or SuperSignal West Pico Chemiluminescent Substrate (34080, Thermo Scientific) (less sensitive), depending on the sensitivity of the chemiluminescent substrate the desired. Digital images of the western blotting were obtained in a ChemiDoc XRS System (Bio-Rad) with Image Lab software. The molecular weights of the bands were determined by the Image Lab software, using the marker as reference of the molecular weight sizing.

2.3.4. IMMUNOFLUORESCENCE

Immunofluorescence uses the specificity of antibodies labeled with fluorochromes to target specific biomolecules within a cell, allowing their visualization in the sample (Oliver and Jamur, 2010). In this work, immunofluorescence was performed using sequentially two antibodies: a primary antibody, that recognizes and binds to the target molecule, and a secondary antibody, which carries the fluorophore and recognizes and binds to the primary antibody.

Cells were grown on 22x22 mm microscope cover glasses (631-1570, VWR) until they reached approximately 70% of confluence. Cells were fixed in methanol/2mM EGTA for 3 minutes at 4º C, hydrated three times with TBS (Appendix II) permeabilized with TBS 0.1% Triton-X100, washed three times with TBS and blocked with 10% (w/v) FBS in TBST (TBS 0,05% Tween 20) for 30 minutes at 4ºC , and incubated with primary antibody in 5% (w/v) FBS in TBST diluted for a specific concentration, for an overnight period at 4ºC. After rinsing three times, 5 minutes, with TBS 1X, cells were incubated with the secondary antibodies (Appendix II) in 5% (w/v) FBS, TBST diluted to a specific concentration for 2 hours, at room temperature. The cover glasses were rinsed, once again, three times with TBST and one time with TBS, prior to be mounted in microscope slides (H868, ROTH) in Glycerol/n-propyl gallate mounting medium (Appendix I) with 1 µg/mL of DAPI (4'-6-diamidino-2-phenylindole). Cells were examined with a magnification 10000x magnification by standard fluorescence microscopy and

differential interference contrast microscopy (DIC). Image stacks with Z-step of 220 nm were acquired by Zeiss Axio Imager Z2 microscope (Zeiss) using Zeiss AxioVision software.

The same procedures described above were performed for negative controls, with the exception of the incubation with the primary antibody.

2.4. DIGITAL IMAGE PROCESSING AND DECONVOLUTION

The images obtained by immunofluorescence were submitted to a process of deconvolution using Huygens Software (Scientific Volume Imaging BV, Hilversum, The Netherlands). Deconvolution consists on an algorithm-based process used to reverse the effects of convolution on recorded data. It uses the imaging properties of the optical system in the form of the point spread function (PSF), which describes the response of an imaging system to a point source or point object, for 'putting back the light where it is coming from'. Deconvolution very effectively removes noise and background and eliminates stray light and blur caused by distortion. The technique yields images of appreciably increased contrast and resolution.

The same deconvolution parameters were used for all the analyzed images, using an absolute background value calculated from the average of automatically estimated background values using 15 images of negative controls (only stained with secondary antibodies). The absolute background value for c-Hairy1 channel was 325, and for DAPI channel was used the value 260.

2.4.1. QUANTIFICATION OF IMAGES

Huygens Software interactive Object Analyzer, built a 3D image using the obtained deconvolved images, and analyzed all the objects present, and obtained the objects statistics, including geometrical data, spatial location, relation with neighbors and/or reference objects and degree of channels overlap, a parameter used to calculate the objects colocalization.

For calculations of cell subcellular compartments volume, such as nucleus and cytoplasm, a region of interest (ROI), was selected manually (Figure 2.2), and the Software calculated the total volume occupied by the selection, and the volume of its individual objects. Cytoplasm volume was calculated subtracting the nucleus to whole cell total volume.

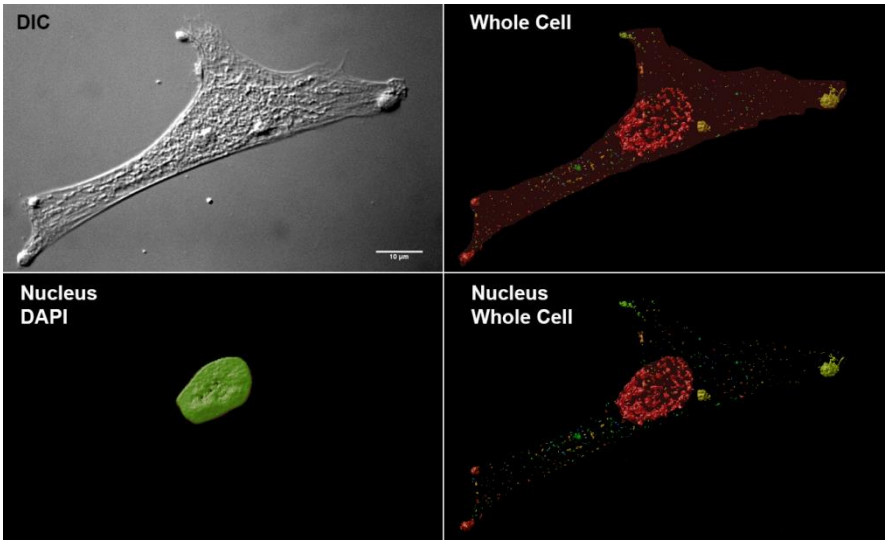


Figure 2.2 – Representation of the selection of a ROI of a whole cell and its nucleus. ROI was determined manually using image of DIC, and the nucleus stained with DAPI as reference for the Whole cell and nucleus ROI selection, respectively.

CHAPTER 3

RESULTS

3. RESULTS

3.1. BIOINFORMATIC ANALISYS OF C-HAIRY1 SEQUENCE

3.1.1. PRIMARY STRUCTURE ANALYSIS

Chick Hairy1A sequence was collected from the NCBI protein database, with the accession number of AAP44728.1. ExPASy's software, Protparam, predicted the physico-chemical properties of the protein, including the molecular weight, theoretical isoelectric point, the grand average of hydropathicity (GRAVY) (Table 3.1) and the amino acid composition (Table 3.2).

The software predicted a molecular weight of approximately 30,7 kDa, and results revealed that c-Hairy1A has a basic isoelectric point, and negative GRAVY values.

Table 3.1 - Protparam predicted molecular weight, theoretical isoelectric point and grand average of hydropathicity for c-Hairy1A.

	c-Hairy1A
Nr of aa	290
Molecular weight (Da)	30666,1
Theoretical pI	9,56
GRAVY	-0,268

In terms of amino acid composition, alanine, proline and serine revealed to be amino acids with greater percentage of abundance (Table 3.2).

Table 3.2 - Amino acid composition and respective percentage of c-HairyA.

Residues	Nr. of residues	Percentage (%)
Alanine (A)	35	12.1
Arginine (R)	16	5.5
Asparagine (N)	10	3.4
Aspartic acid (D)	8	2.8
Cystein (C)	4	1.4
Glutamine (Q)	11	3.8
Glutamic acid (E)	13	4.5
Glycine (G)	21	7.2
Histidine (H)	8	2.8
Isoleucine (I)	9	3.1
Leucine (L)	26	9.0
Lysine (K)	14	4.8
Methionine (M)	8	2.8
Phenylalanine (F)	7	2.4
Proline (P)	33	11.4
Serine (S)	28	9.7
Threonine (T)	15	5.2
Tryptophan (W)	2	0.7
Tyrosine (Y)	4	1.4
Valine (V)	18	6.2

Prosite identified the conserved protein and functional sites of the c-Hairy1 variants using its sequences, detecting the basic, HLH and Orange domains, and also WRPW motif (Table 3.3).

Table 3.3 - Predicted positions of the conserved domains and motifs predicted for c-Hairy1A

Domain	Position	Length	Sequence
Basic	35-48	14	HRKSSKPIMEKRRR
HLH	49-93	58	ARINESLGQLKMLILDALKKDSSRHSKL EKADILEMTVKHLRNLQ
Orange	111-144	34	YRAGFNECMNEVTRFLSTCEGVNADV RARLLGHL
WRPW	287-290	4	WRPW

3.1.2. POST-TRANSLATIONAL MODIFICATIONS

NetPhos and DISPHOS were used, in order to increase the accuracy of the predicted phosphorylation sites. The appendix III contains all the obtained data, which includes all the residues analyzed and the assigned scores.

Both programs detected a total number of 47 phosphorylation sites, although they showed discrepancies on the number of sites that scored higher than 0,5 (27 and 21 for NetPhos and DISPHOS respectively) (Figure 3.1).

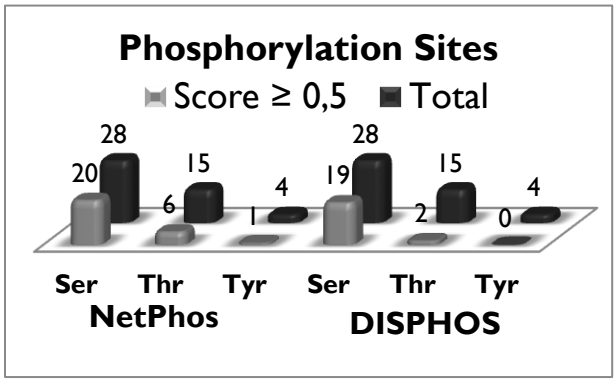


Figure 3.1- Graphical representation of the total phosphorylation sites predicted by NetPhos and DISPHOS and the ones with score higher than 0,5. Data was taken fom the tables presented in apendix C.

Table 3.4 presents the consensus phosphorylation sites in which both softwares assigned scores higher than 0,5, and 17 phosphorylation sites were selected.

Table 3.4 - Selected phosphorylation sites predicted by NetPhos (NP) and DISPHOS (DP).

Pos.	Res.	NP Score	DP Score	Avrg.
13	S	0,951	0,903	0,927
21	S	0,593	0,880	0,737
23	S	0,583	0,926	0,755
25	T	0,679	0,761	0,720
31	S	0,997	0,977	0,987
33	S	0,991	0,988	0,990
38	S	0,997	0,981	0,989
39	S	0,879	0,951	0,915
70	S	0,987	0,884	0,936
71	S	0,727	0,888	0,808
74	S	0,997	0,922	0,960
247	S	0,600	0,742	0,671
251	S	0,779	0,598	0,689
257	S	0,986	0,567	0,777
273	S	0,901	0,546	0,724
277	S	0,979	0,577	0,778
285	S	0,994	0,961	0,978

Kinase specific phosphorylation sites were predicted by NetPhosK and it identified 12 possible kinases with affinity to phosphorylate c-Hairy1A, described in Table 3.5 with their correspondent scores. Correlating the phosphorylation sites determined in previous results shown in Table 3.4 and the predicted kinases, a graphic that constitutes figure 3.2, was constructed showing the distribution of these kinases.

Table 3.5 – Kinases that were predicted to phosphorylate c-Hairy1A.

Site	Kinase	Score
T-5	CKII	0.54
T-11	PKG	0.59
S-23	PKC	0.55
T-25	p38MAPK	0.54
T-25	cdk5	0.65
S-31	PKG	0.56
S-33	RSK	0.54
S-33	PKC	0.64
S-39	RSK	0.52
S-39	PKC	0.54
S-39	PKA	0.67
S-54	DNAPK	0.51
S-54	PKC	0.59
S-54	PKA	0.59
S-70	PKC	0.71
S-71	PKC	0.61
S-74	CKII	0.54
S-127	PKA	0.67
S-127	PKG	0.54
S-197	cdk5	0.50
T-243	CKI	0.56
S-247	PKC	0.63
T-252	PKC	0.61
T-252	cdc2	0.52
T-253	p38MAPK	0.51
S-257	p38MAPK	0.50
S-257	cdk5	0.51
T-263	PKC	0.50
S-273	DNAPK	0.63
S-273	ATM	0.52
S-277	GSK3	0.51
S-285	RSK	0.66
S-285	PKC	0.53
S-285	PKA	0.79

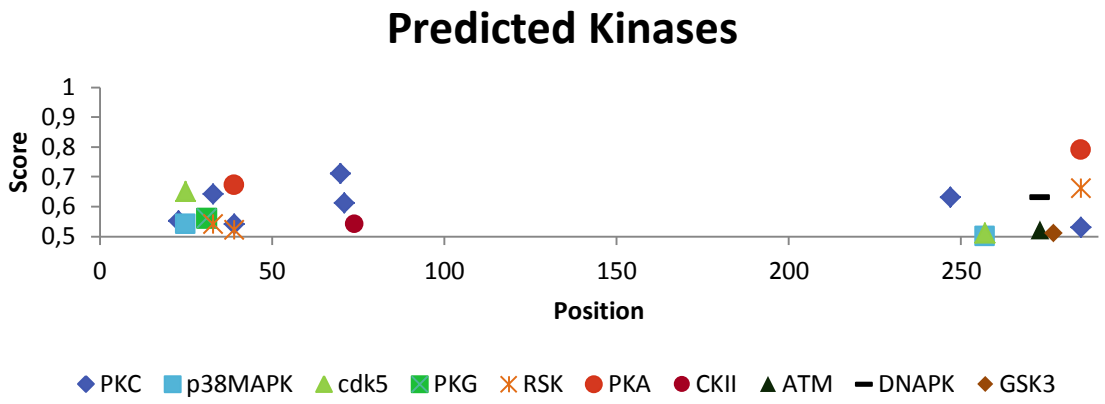


Figure 3.2 – C-Hairy1 specific kinases and the distribution of the predicted phosphorylation sites. This graphical representation was constructed correlating the results obtained in table 3.4 and 3.5, showing the kinases that have affinity to phosphorylate the most probable phosphorylation sites found in c-Hairy1A sequence. PKC- Protein Kinase C, p38 MAPK - p38 Mitogen-Activated Protein Kinase; Cdk5 - Cyclin-dependent kinase 5; PKG - GMP-dependent Protein Kinase, RSK - Ribosomal S6 Kinase; PKA - cyclic AMP dependent Protein Kinase; CKII - Casein Kinase 2; ATM - Ataxia Telangiectasia Mutated; DNAPK DNA-dependent Protein Kinase; GSK3 - Glycogen Synthase Kinase 3.

3.1.3. SECONDARY STRUCTURE ANALYSIS

The Jufo9D Server calculated the secondary structure, predicting three secondary structures types, α-helices, β-strands and coil (turns, loops or random coil), and assigned a probability score for combinations of these structures. NetSurfP calculated the surface accessibility, which determines whether residues are exposed to the surface or buried in the protein. The c-Hairy1 secondary structure and surface accessibility prediction are visual represented in figure 3.3, including also the predicted phosphorylation sites described in Table 3.4. The conserved domains bHLH and orange domains, and WRPW motif, are also highlighted.

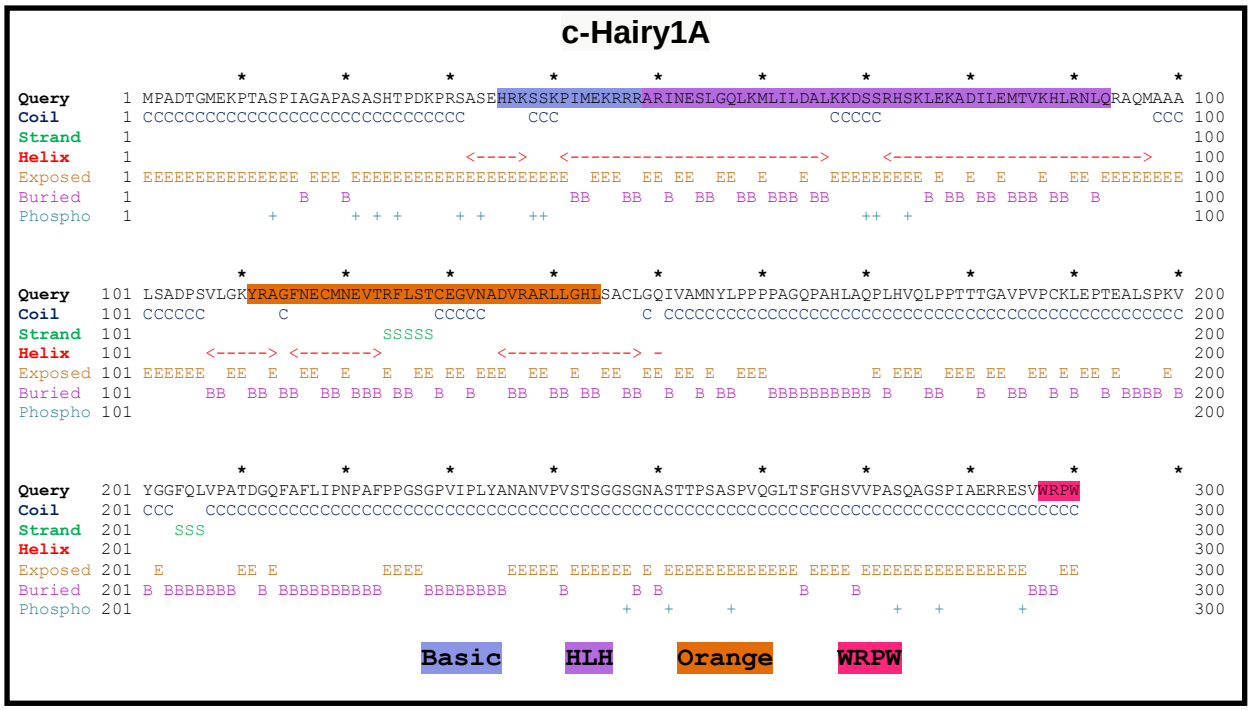


Figure 3.3 - C-Hairy1A secondary structure, surface accessibility and phosphorylation sites. Secondary structure was computed by Jufo9D from c-Hairy1A sequence. From the protein sequence (Query) is calculated three different secondary structures (Helix, Strand Coil), and are represented in the figure the secondary structure that registered highest score. Conserved domains are also highlighted in the protein sequence (Basic motif, HLH, Orange and WRPW). Surface accessibility was predicted in two classes as either exposed or buried. To each residue is associated a secondary structure type and surface accessibility class. The phosphorylation sites are represented in Table 3.4. Asterisks indicate a 10 residue interval.

3.1.4. SUBCELLULAR LOCALIZATION OF C-HAIRY1 AND ITS HOMOLOGS PREDICTION

A search for c-Hairy1 homologs was performed, in order to increase robustness of the subcellular localization signals prediction. Running a protein BLAST against the c-Hairy1A, several homologs were detected as producing significant alignments. Four proteins were selected, which fulfilled the requirements mentioned in the Material and Methods, and are represented in the Table 3.6, by ordained by E-values of the alignment. Sequences of the homologs proteins are described in Appendix III.

Table 3.6 – Selection of c-Hairy1A protein homologs obtained from BLASTP and its respective accession codes, E-value and maximum identity of the alignment.

Order	Accession	Name	Organism	E-value	Max identity
1º	AAP44728.1	c-Hairy1A	*Gallus gallus*	0.0	100%
7º	NP_001082574.1	X-Hes4A	*Xenopus laevis*	5e-147	74%
8º	NP_571948.1	z-Her1	*Danio rerio*	2e-146	75%
41º	NP_005515.1	h-Hes1	*Homo sapiens*	8e-101	59%
52º	NP_032261.1	m-Hes1	*Mus musculus*	1e-98	65%

A phylogenetic tree was generated by Phylogeny.fr (http://www.phylogeny.fr/) and visualy is possible to distiguish two clusters which can be charaterized in non mamalian, and mamalian proteins (Figure 3.4). According to the distance scale given by the software it's observable that the zebrafish Hes1 is the closest analised protein to c-Hairy1, next to the *Xenopus leavis* Hes4A. The mammalian Hes1 proteins belonging to mouse and human presented a greater genetic distance.

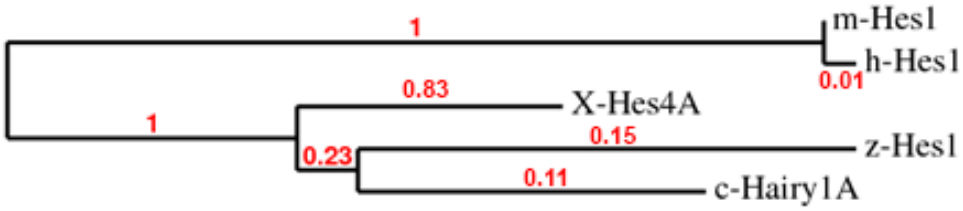


Figure 3.4 – Phylogenetic tree of c-Hairy1A homologs. Distance scale is provided and given as a reference for the comprehension of the tree.

After relating the proteins phylogenetically, a prediction of the proteins localization in the cell was performed, using NucPred and PSORT II programs. NucPred assigned a score correspondent to the probability of the proteins to spend some time in the nucleus, and the results must be analyzed by the aplication of a threshold (see Table 2.2, Materials and Methods).

PSORT II assignes a k-NN value, which represent a probability of the proteins to be localized in different compartiments of the cell, e.g., nucleus, cytoplasm, mithocondria, vesicles of secretory system (VSS) and endoplasmic reticulum (ER).

The results are shown in table 3.7, and according to NucPred results, all the homologs scored higher than 0,5. PSORT II predicted all the homologs to be localized mainly in the nucleus, although the percentage viried. All the analized homologs shared a percentage of 4.3% to be localized either on vesicles of secretory system or endoplasmatic reticulum.

Table 3.7 - **Prediction of subcellular localization of c-Hairy1 homologs.**

Protein	NucPred scores	PSORT II *k*-NN (%)				
		Nuc*	Cyt**	Mit***	VSS****	ER*****
c-Hairy1A	0.56	60.9	17.4	13.0	4.3	4.3
X-Hes4A	0.81	56.5	17.4	17.4	4.3	4.3
z-Hes1	0.59	65.2	17.4	8.7	4.3	4.3
m-Hes1	0.72	60.9	13.0	17.4	4.3	4.3
h-Hes1	0.67	60.9	13.0	17.4	4.3	4.3

* Nucleus; **Cytoplasm; *** Mithocondria; **** Vesicles of Secretory System; *****Endoplasmic Reticulum

Figure 3.5 shows the homolog multiple alignemt provided by NucPred. Each residue was colored according to the probability of being a nuclear signal, and colors range from blue (negative) to red (positive). Observing the results it's possible to notice the presence of a single conserved Nuclear Localization Signal (NLS) (KRRR) located at the end of the basic domain. The predicted phosphorylation sites (Table 3.4) were also indicated in the alignment, in order to observe the conservation of these residues in the homolog sequences.

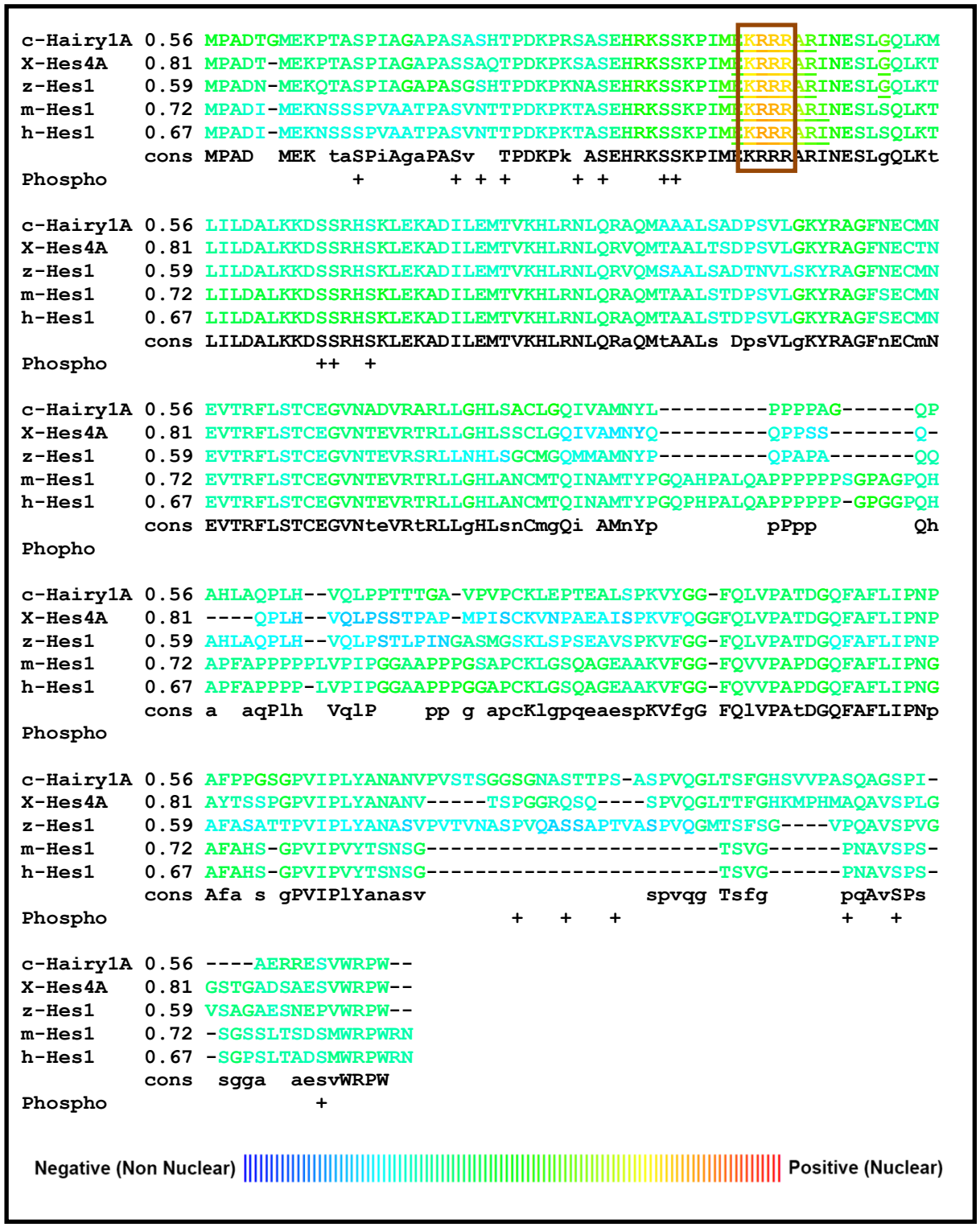


Figure 3.5 –Multiple alignment of c-Hairy1 colored for the presence of NLS, and predicted phosphorylation sites. NucPred performed a ClustalW multiple alignment of the set of homologs, c-Hairy1A, X-Hes4A, z-Hes1, m-Hes1 and h-Hes1. Each residue was colored from blue (negative) to red (positive), indicating the probability to be part of a NLS. In consensus sequence, the residues annotated with uppercase letters represent an

unanimity. In cases of nonunanimity the most common residue is annotated with lowercase letters. The predicted phosphorylation sites (Table 3.4) were also indicated in the alignment.

The prediction of the nuclear localization of c-Hairy1 does not preclude the possibility of it to be export from the nucleus to cytoplasm, and if so, bioinformatics prediction of nuclear export signal (NES) would be rather useful. These are short sequences comprising four spaced hydrophobic residues (denoted Φ^1–Φ^4) and following the consensus Φ1-(x)$_{2-3}$-Φ2-(x)$_{2-3}$-Φ3-x-Φ4, where x is an amino acid preferentially negative charged, polar or small (Kutay and Guttinger, 2005).

NetNES 1.1 searched for nuclear export signals (NES) in the c-Hairy1 homologs sequences, calculating a score for each residue of the probability to be part of a NES. Table 3.8 comprises the two potential NESs showing higher NES activity in the homolog sequences. Both potential NES were highly conserved in the analyzed homologs. NES1 is not a perfect NES consensus, as it has only one amino acid between the second and third hydrophobic residues. The software also registered high NES score in alanine residue (Ø4?) of the mammal homologs sequences, and a lower but still high score in c-Hairy1 sequence. The leucine (Ø5?) next to alanine was also considered to belong to NES since it is a hydrophobic residue. Only the isoleucine (Ø2) reveals a consistent high score in every analyzed sequence.

Table 3.8 – Comparison of NetNES predicted nuclear export sequences (NESs) with consensus NES motifs in c-Hairy1 protein and homologs. In the consensus NES sequence, X indicates any amino acid and Ø indicates a hydrophobic residue, such as leucine, isoleucine, valine or methionine. Letters highlighted in green correspond to critical residues to NES activity (NetNES score>0.5), in orange are highlighted residues that scored between 0.3 and 0.5, and yellow highlighted residues indicate residues with low scores, however represent hydrophobic residues.

	NES1	NES2
Consensus	Ø1XXØ2XØ3XØ4?Ø$_5$?	Ø1XXXØ2XXØ3XØ4
c-Hairy1	58 LKMLILDAL 66	76 LEKADILEMTV 86
X-Hes4A	57 LKTLILDAL 65	75 LEKADILEMTV 75
z-Hes1	57 LKTLILDAL 65	75 LEKADILEMTV 75
m-Hes1	57 LKTLILDAL 65	75 LEKADILEMTV 75
h-Hes1	57 LKTLILDAL 65	75 LEKADILEMTV 75

The results are also represented in the form of a graphic, combining results of neural network and hidden Markov models, resulting on a score of NES calculated by the NetNES server (Figure 3.6).

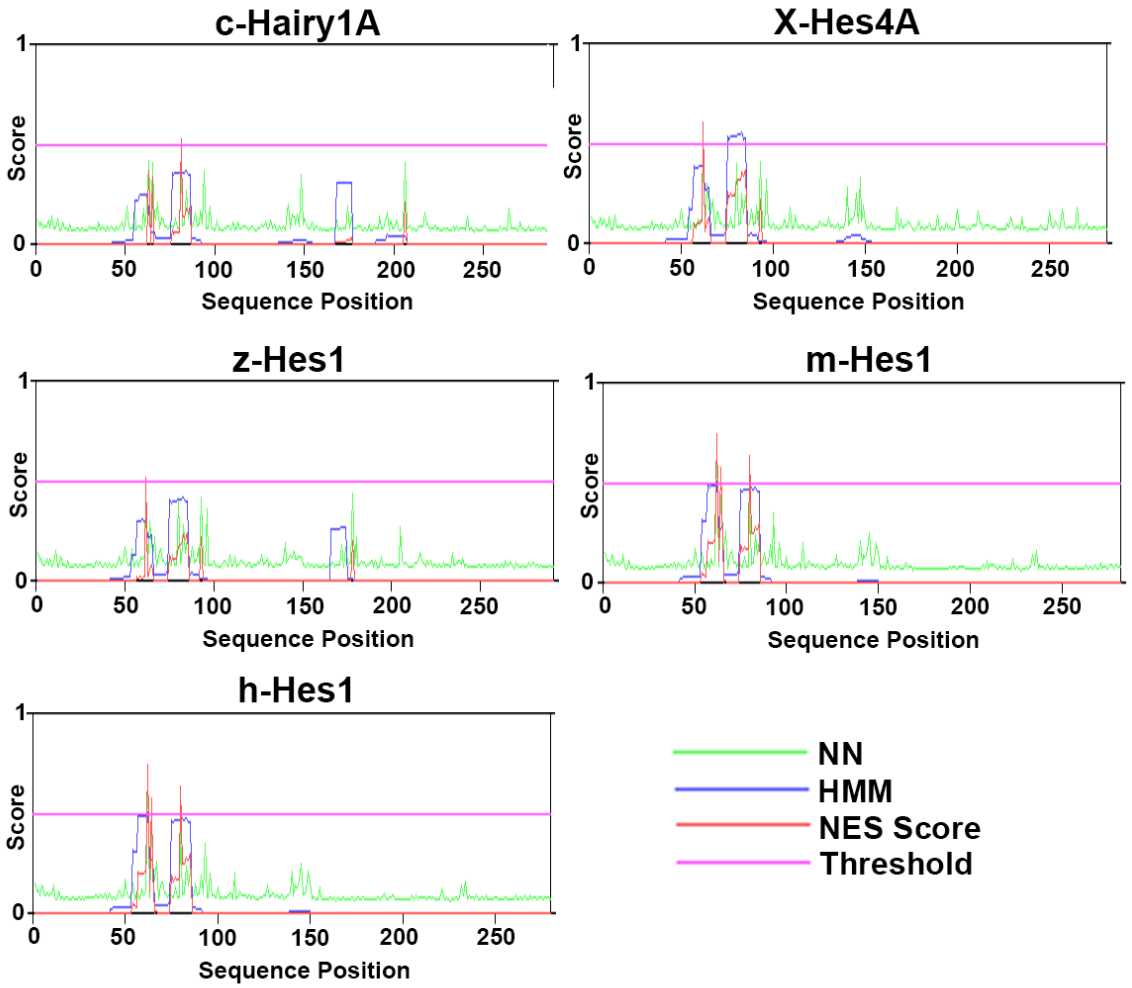


Figure 3.6 - NetNES prediction of the Hairy homologs. All the analyzed sequences presented a region where is evident higher peaks of NES Score, and significant disturbances of hidden Markov models (HMM), between residues 50 and 100. Although the mammalian homologs revealed more accentuated peaks, and crossed the threshold line. NN - Neural Network; HMM – Hidden Markov Model.

3.2. ASSESSING THE C-HAIRY1 ANTIBODY SPECIFICITY AND SENSITIVITY

In order to test the specificity and sensitivity of the purchased monoclonal antibody against c-Hairy1, protein extracts of chicken embryo stage HH10-13, and chicken embryonic fibroblasts (CEFs) were subjected to SDS-PAGE, followed by western blot. The purified c-Hairy1 protein was used as positive control, in order to confirm that the bands detected in the extracts corresponded to the same molecular weights as the purified protein.

Bradford method was used to determine the protein extracts concentration, and the concentration of the chick embryo and CEFs extracts was established as 13,51 mg/ml and 11,59 mg/ml respectively. The purified protein concentration has been already determined by the company that produced it as being 0,14 mg/ml.

As the purified protein was used as positive control of this experience, it was needed to establish the most appropriate loading amount. So, different quantities of protein were loaded in the SDS-PAGE gel. Proceeding with a western blot, it was possible to detect two bands with 35 and 56 kDa, of which molecular weight was estimated by Image Lab software using Marker lane as reference for calculation. Results revealed a consistent increase of intensity of the bands in higher amounts of protein loaded, but the same was not observed relatively to the higher molecular weight band. The optimal loading amount of purified protein was settled as 100 ng per lane, as it revealed to be the lowest amount in which the 35 kDa band was best visible (Figure 3.7).

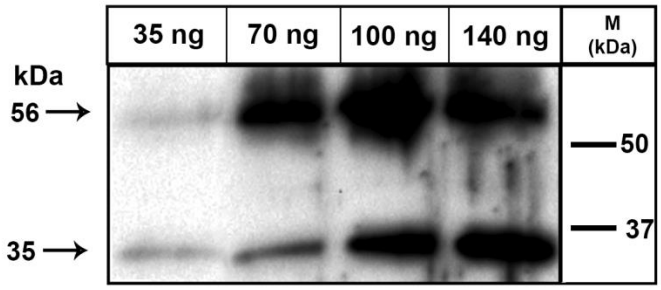


Figure 3.7 - Determination of the most appropriate loading amount of purified protein for c-Hairy1 detection by western blot. Amounts of 35, 70, 100 and 140 ng of protein/lane of c-Hairy1 purified protein were loaded in a SDS-PAGE gel. Western blot for c-Hairy1 detected two bands corresponding to 35 and 56 kDa, and is observable and gradual increase of lower band intensity. For detection of c-Hairy1, several amounts of the protein were loaded in the SDS-PAGE gel, and membrane was blotted with 1:100 of the monoclonal α-c-Hairy1 primary antibody, and 1:1000 of α-mouse secondary antibody. The membrane was exposed with Biorad Immun-Star WesternC Chemiluminescent Kit.

With the loading amount of purified protein established, the conditions were then settled to determine the optimal loading amount of protein extracts. First, and because the protein was expected to be expressed in chick embryos, several amounts of protein extracts were tested by western blot for detection of c-Hairy1. β-tubulin1 detection was used as loading control (Figure 3.8).

Western blot results revealed several bands in the embryo extracts, detecting all the bands found in purified protein (56 and 35 kDa; Figure 3.8), and also other molecular weight bands. In all the amounts tested the 56 kDa band was the most defined band next to the 35 kDa, indicating a strong signal. Faint bands (85, 79, 37 and 33 kDa; Figure 3.8) were also visualized independently of the exposure time (Figure 3.8).

Proportional to the increase of the loaded amounts, an increase of the bands intensity was verifiable. However, the amount of 75 µg of protein per lane revealed to be the lowest loading amount in which all bands in general were better visible, and for that reason was chosen as the most appropriate loading amount (Figure 3.8). The loading control, β-tubulin1, presented a stronger signal in the higher loading amounts, and as it was expected no signal was detected in the purified protein lane (Figure 3.8).

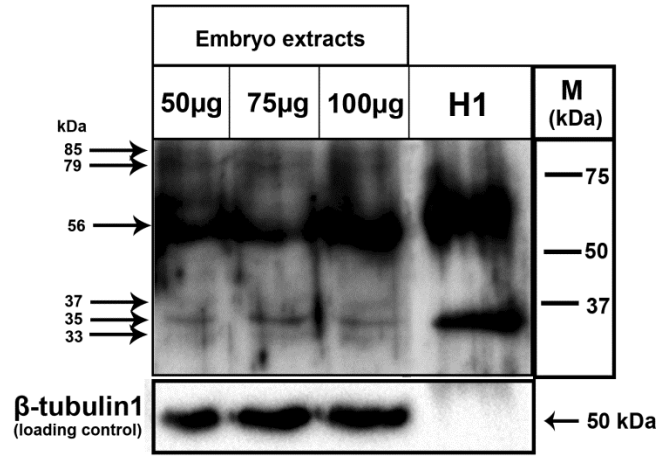


Figure 3.8 - Determination of the most appropriate loading amount of embryo protein extract for c-Hairy1 detection by western blot. Amounts of 50, 75 and 100µg of protein/lane of embryo extracts, and 100 ng of c-Hairy1 purified protein (H1) were loaded in a SDS-PAGE gel and blotted for c-Hairy1 detection. The blotting detected several bands in the embryo extracts lanes, correspondent to a molecular weight of 85, 79, 56,37, 35 and 33 kDa. In the H1 lane were detected 56 and 35 kDa bands. Membrane was blotted with 1:100 of the mouse monoclonal α-c-Hairy1 primary antibody, and 1:1000 of α-mouse secondary antibody, and then exposed with Biorad Immun-Star WesternC Chemiluminescent kit. For detection of β-tubulin1 (loading control), the membrane was stripped and incubated with 1:2000 rabbit α-β-tubulin1 antibody following incubation with 1:4000 α-rabbit secondary, then exposed with Biorad Immun-Star WesternC Chemiluminescent kit. The molecular weights were determined by the marker (M).

A western blot was performed to assess and demonstrate the expression of c-Hairy1 in chick embryo fibroblasts (CEFs), using CEFs protein extracts. But firstly, the appropriate loading amount had to be established, and several amounts of extract, as well as 100 ng of purified c-Hairy1 protein were loaded in the SDS-PAGE gel, and then blotted for c-Hairy1 detection.

Results revealed expression of c-Hairy1 in CEFs, in all the tested loaded amounts, detecting several bands with different molecular weights, including all the bands detected in the purified protein lane (56 and 35 kDa; Figure 3.9)

Proportional to the increase of the loaded amounts, it was verifiable an increase of the bands intensity. However, the loading amount of 75 µg of protein/lane revealed to be the lowest amount in which all bands in general were better visible, and for that reason this value was chosen as the most appropriate loading amount (Figure 3.9).

The loading control, β-tubulin1, presented a stronger signal in the higher loading amounts, and as it was expected no signal was detected in the purified protein lane (Figure 3.9).

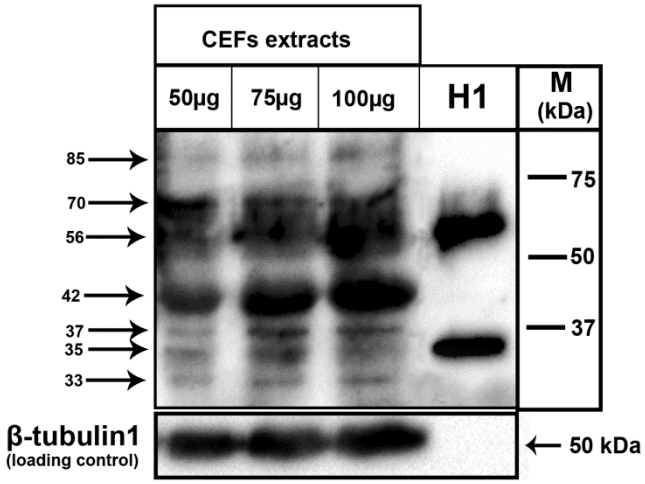


Figure 3.9 - Determination of the most appropriate loading amount of CEFs protein extract for c-Hairy1 detection by western blot. Amounts of 50, 75 and 100µg of protein/ lane of CEFs protein extracts, and 100 ng of c-Hairy1 purified protein (H1) were loaded in a SDS-PAGE gel and blotted for c-Hairy1 detection. The blotting detected several bands in the embryo extracts lanes, correspondent to a molecular weight of 85, 62, 56, 42, 37, 35 and 33 kDa. In the H1 lane were detected 56 and 35 kDa bands. Membrane was blotted with 1:100 of the mouse monoclonal α-c-Hairy1 primary antibody, and 1:1000 of α-mouse secondary antibody, and then exposed with Biorad Immun-Star WesternC Chemiluminescent kit. For detection of β-tubulin1 (loading control), the membrane was stripped and incubated with 1:2000 rabbit α-β-tubulin1 antibody following incubation with 1:4000 α-rabbit secondary, then exposed with Biorad Immun-Star WesternC Chemiluminescent kit. The molecular weights were determined by the marker (M).

It was hypothesized that higher molecular weights bands would could result of dimer formation, and that an increase of the denaturation time would affect the dimer formation, and consequently the higher bands would decrease of intensity. The samples used to be denatured for 10 min and then loaded into the gel, as it was suggested in the SDS-PAGE protocol. However to verify the hypothesis, different denaturation times (5, 10, 15 and 30 min) were tested using 75 µg of protein /lane of embryo protein extracts and 100 ng purified protein, and blotted for c-Hairy1 detection (Figure 3.10).

The membrane was exposed using two chemiluminescent substrates with different sensitivities, Pico Chemiluminescent substrate and Biorad Immun-Star WesternC Chemiluminescent kit.

Results demonstrated that exposing with Immu-Star WesternC Chemiluminescent Kit, a more sensitive chemiluminescent, several bands were detected in all samples in all the denaturation times tested. It could also be observable an intensity decreasing of the higher molecular weight bands.

The exposure with Pico Chemiluminescent substrate, a less sensitive chemiluminescent, showed that in embryo extract samples, the lower molecular weight bands was only clearly detectable in the 5 minutes denaturation time, indicating that higher denaturation times influence not only the higher but also the lower molecular weight bands (Figure 3.10). For that matter the 5 min denaturation time was considered to be the most suitable denaturation time for detection of c-Hairy1.

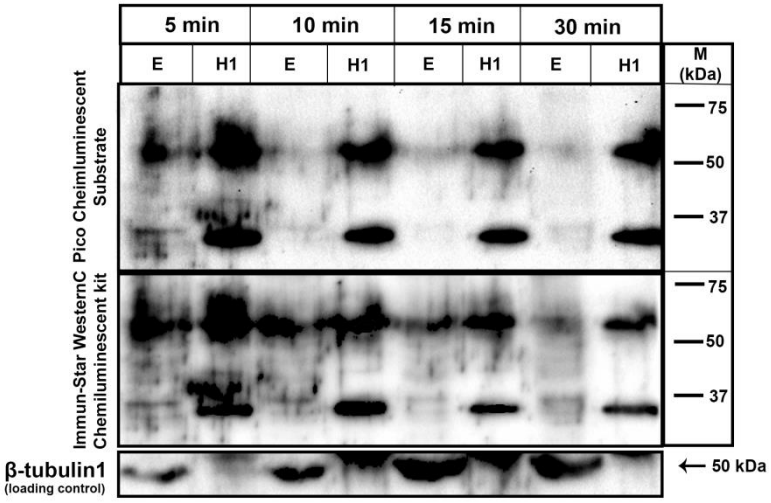


Figure 3.10 - Determination of the most appropriate denaturation time for c-Hairy1 detection by western blot. Samples of embryo protein extracts (75 µg of protein/lane) (E) and c-Hairy1 purified protein (100 ng of protein/lane) (H1) were denatured for 5, 10, 15 and 30 min and loaded in a SDS-PAGE gel, then blotted for detection of c-Hairy1. Exposing the membrane with Imun-Star WestenC Chemiluminescent Kit, were identified several bands in the embryo extract and H1 lanes, in all the tested denaturation times. However, exposing with Pico Chemiluminescent Substrate, the 5 min presented a stronger detection signal than the other tested denaturation times in both embryo and c-Hairy1 purified protein, tough all the H1 always presented a strong detection signal. Membrane was blotted with 1:100 of the mouse monoclonal α-c-Hairy1 primary antibody, and 1:1000 of α-mouse secondary antibody, and then exposed Biorad Immun-Star WesternC

Chemiluminescent kit. For detection of β-tubulin1 (loading control), the membrane was stripped and incubated with 1:2000 rabbit α-β-tubulin1 antibody following incubation with 1:4000 α-rabbit secondary, then exposed with Biorad Immun-Star WesternC Chemiluminescent kit. The molecular weights were determined by the marker (M).

With the most suitable conditions for detection of c-Hairy1 established, a western blot meeting all the conditions was performed, loading 75 µg of embryo (E) and CEFs (C) protein extracts and 100 ng of c-Hairy1 purified protein (H1) denatured for 5 min, and blotted for c-Hairy1 detection. In embryo and CEFs extracts were identified the same bands although presenting intensity differences (Figure 3.11).

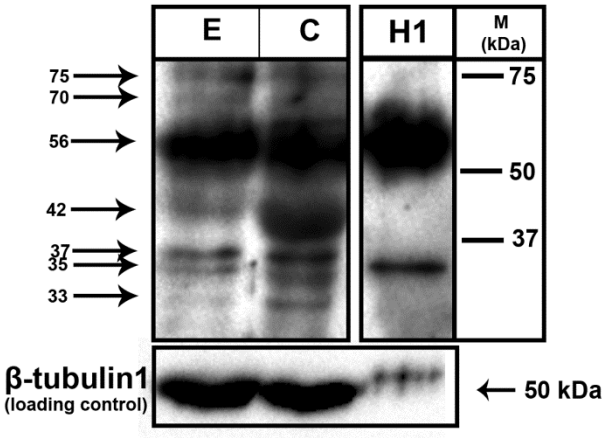


Figure 3.11 - Detection of c-Hairy1 in embryo and chick embryo fibroblasts (CEFs) by western blot. Samples of embryo protein extracts (75 µg of protein/lane) (E), CEFs extracts (75 µg of protein/lane) (C) and c-Hairy1 purified protein (100 ng of protein/lane) (H1) were denatured for 5 min and loaded in a SDS-PAGE gel, then blotted for detection of c-Hairy1. Embryo and CEFs protein extracts lanes presented the same bands, however was noticeable a difference of intensity in the 70, 42 and 33 kDa bands. In the H1 lane were detected 56 and 35 kDa bands Membrane was blotted with 1:100 of the mouse monoclonal α-c-Hairy1 primary antibody, and 1:1000 of α-mouse secondary antibody, and then exposed with Biorad Immun-Star WesternC Chemiluminescent kit. For detection of β-tubulin1 (loading control), the membrane was stripped and incubated with 1:2000 rabbit α-β-tubulin1 antibody following incubation with 1:4000 α-rabbit secondary, then exposed with Biorad Immun-Star WesternC Chemiluminescent kit. The molecular weights were determined by the marker (M).

The negative controls were tested to validate the obtained results, and to confirm that the detected bands were not resultant from unspecific secondary antibody affinity. Results of the negative controls showed that none of the secondary antibody tested showed affinity to any of the samples (Figure 3.12).

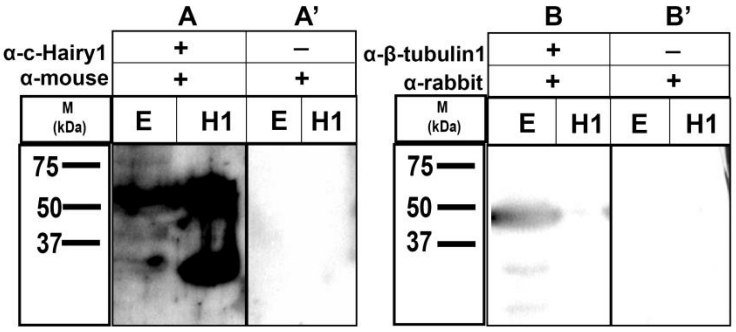


Figure 3.12 – Test for secondary antibodies unspecific staining by western blot. Samples of embryo protein extracts (75 µg of protein/lane) (E) and c-Hairy1 purified protein (100 ng of protein/lane) (H1) were loaded in a SDS-PAGE gel, and when transferred the membrane was split in four parts, and incubated with different antibodies. *(A)* Positive control for c-Hairy1 detection, showing a strong signal of c-Hairy1 detection. Membrane was blotted with 1:100 of the mouse monoclonal α-c-Hairy1 primary antibody, and 1:1000 of α-mouse secondary. *(A')* Negative control for c-Hairy1 detection, showing no detection of c-Hairy1. Membrane was blotted using 1:1000 of α-mouse secondary. *(B)* Positive control for β-tubulin1 detection, presenting a strong signal of detection. Membrane was blotted using 1:2000 of rabbit α-β-tubulin1 antibody and 1:4000 of α-rabbit secondary. *(B')* Negative control for β-tubulin1 detection, showing no sign of detection. Membrane was blotted using 1:4000 of α-rabbit secondary. Membranes were exposed with Biorad Immun-Star WestenC Chemiluminescent kit. The molecular weights were determined by the marker (M).

3.3. SUBCELLULAR DISTRIBUTION OF C-HAIRY1 PROTEIN BY FLUORESCENCE MICROSCOPY

An immunofluorescence using the c-Hairy1 monoclonal antibody was performed, in order to determine and establish the expression pattern of c-Hairy1 in chicken embryonic fibroblasts (CEFs) (Figure 3.13). Two secondary antibodies were tested, the anti-mouse Alexa Fluor 568 and the Alexa Fluor 488. Alexa Fluor 568 was chosen, since produced less background and was not as easily excitable by UV light as Alexa Fluor 488. Testing several dilutions of both primary and secondary antibodies the technique was optimized, and the dilutions 1:50 and 1:250 were established as the highest dilution with the best staining. After acquirement of the immunofluorescence images, it was noticeable that image processing, namely deconvolution using SVI Huygens Software, increased greatly the definition and sharpness of the taken images. The Huygens interactive object

analyzer tool, also allowed a more detailed analysis of the images, converting them in 3D images.

Immunofluorescence results showed c-Hairy1 as being present both in nucleus and cytoplasm, having a granular expression. This expression was possible to observe in the original acquired images and the deconvolved, and the object analyzer tool allowed a better visualization of this protein distribution. β-tubulin1 immunostaining was used as positive control and, as a negative control, an immunofluorescence using only the secondary antibody anti-mouse Alexa Fluor 568 was performed (Figure 3.13).

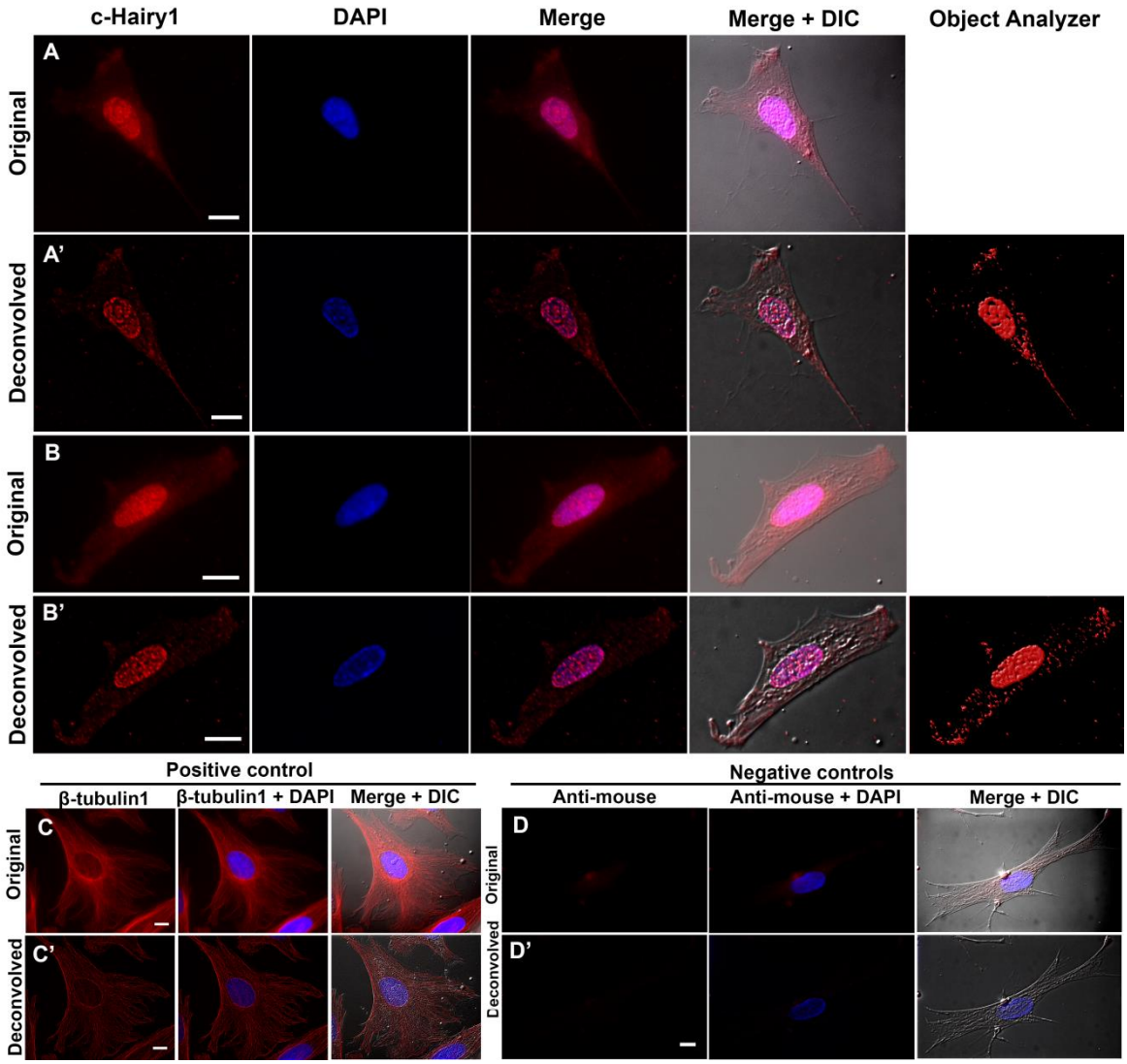


Figure 3.13 - Subcellular distribution of c-Hairy1 protein by fluorescence microscopy of chicken embryonic fibroblasts. Immunodetection of c-Hairy1 reveals the protein to be localized both in nucleus and cytoplasm. This expression pattern is both recognizable in original images (A, B), although is clearer in deconvolved images(A', B'). *(C, C')* Immunostaining for β-tubulin1 was used as a positive control, *(D, D')* Immunostaining for anti-mouse Alexa Fluor 568 secondary antibodies was used as negative control. With SVI's Huygens software was possible to convert the staining in 3D objects, observed in objects analyzer. Immunodetection of endogenous c-Hairy1 in CEFs, using a dilution of 1:50 for c-hairy 1 antibody, and 1:250 for secondary antibody Alexa Fluor 568. Nuclei were stained with DAPI (blue). Fluorescence microscopy (original magnification: 10000x). Images were acquired with the same exposition time. Deconvolution was performed on images to increase the definition and clearance of the signal. Scale bar: 10 μm.

Following a detailed analysis of immunostained cells, it became evident of c-Hairy1's expression at leading edge of some cells (Figure 3.14). The analysis of 54 cells, revealed leading edge expression in 22 cells, corresponding to an expression in approximately 41% of the analyzed cells (Figure 3.14D).

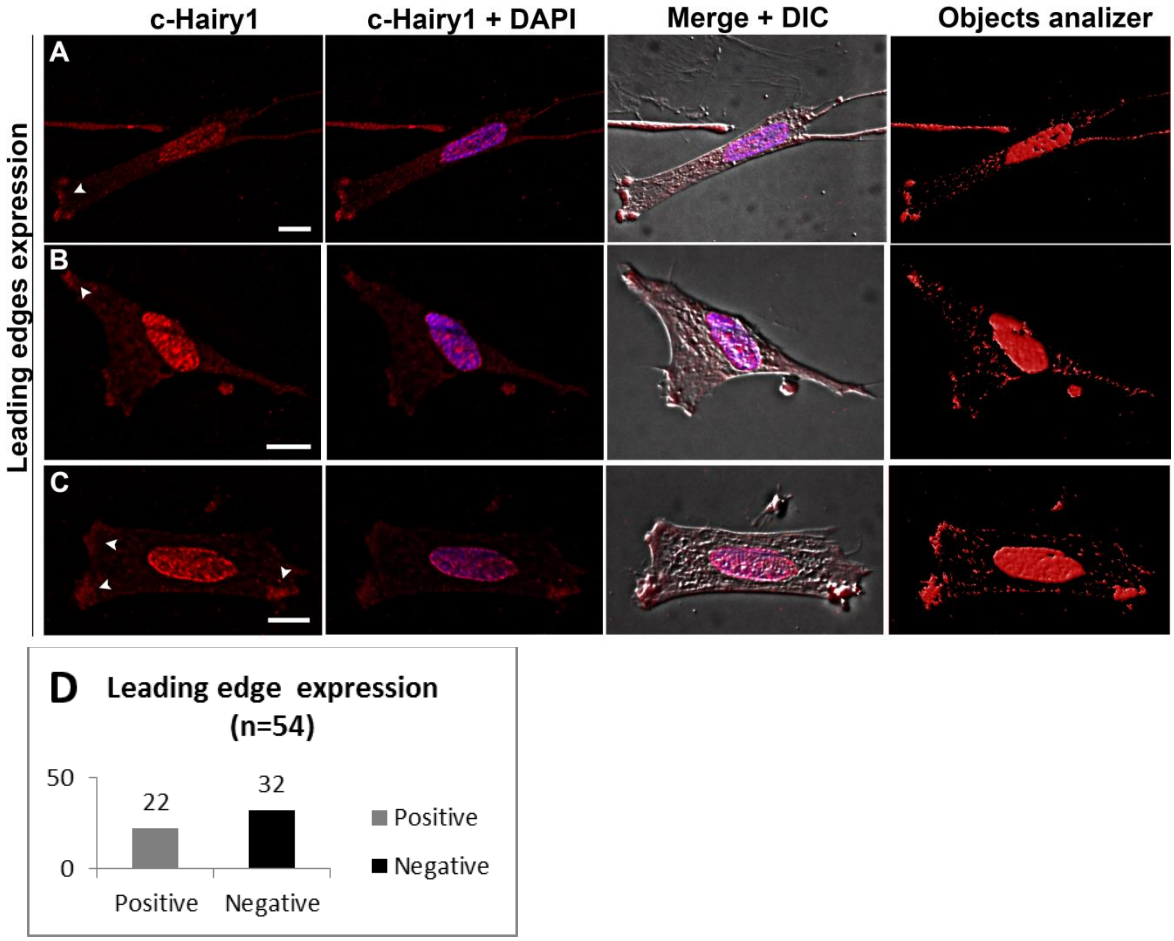


Figure 3.14 - Distribution of c-Hairy1 protein at cell's leading edges by fluorescence microscopy of chicken embryonic fibroblasts. *(A-C)* C-Hairy1 revealed expression at cell's leading edges of CEFs (arrowheads), presenting high levels of fluorescence intensity that stands out from cytoplasmic expression. This expression is noticeable in the 2D images and in the 3D image, converted by SVI's Huygens software. *(D)* An analysis of 54 cells, revealed expression in the leading edges of 22 cells. Immunodetection of endogenous c-Hairy1 in CEFs, using a dilution of 1:50 for c-hairy 1 antibody, and 1:250 for secondary antibody Alexa Fluor 568. Nuclei were stained with DAPI (blue). Fluorescence microscopy (original magnification: 10000x). Images were acquired with the same exposition time. Deconvolution was performed on images to increase the definition and clearance of the signal. Scale bar: 10 µm.

Mitotic cells presented a peculiar expression, as c-Hairy1 presented a much intense fluorescence comparatively to interphase cells, and its expression was homogeneous throughout the cell, except in the presence of DNA (Figure 3.15). In all the mitotic phases c-Hairy1s show the same expression pattern, although in telophase was noticeable a high level of fluorescence located in the contractile ring (Figure 3.15F).

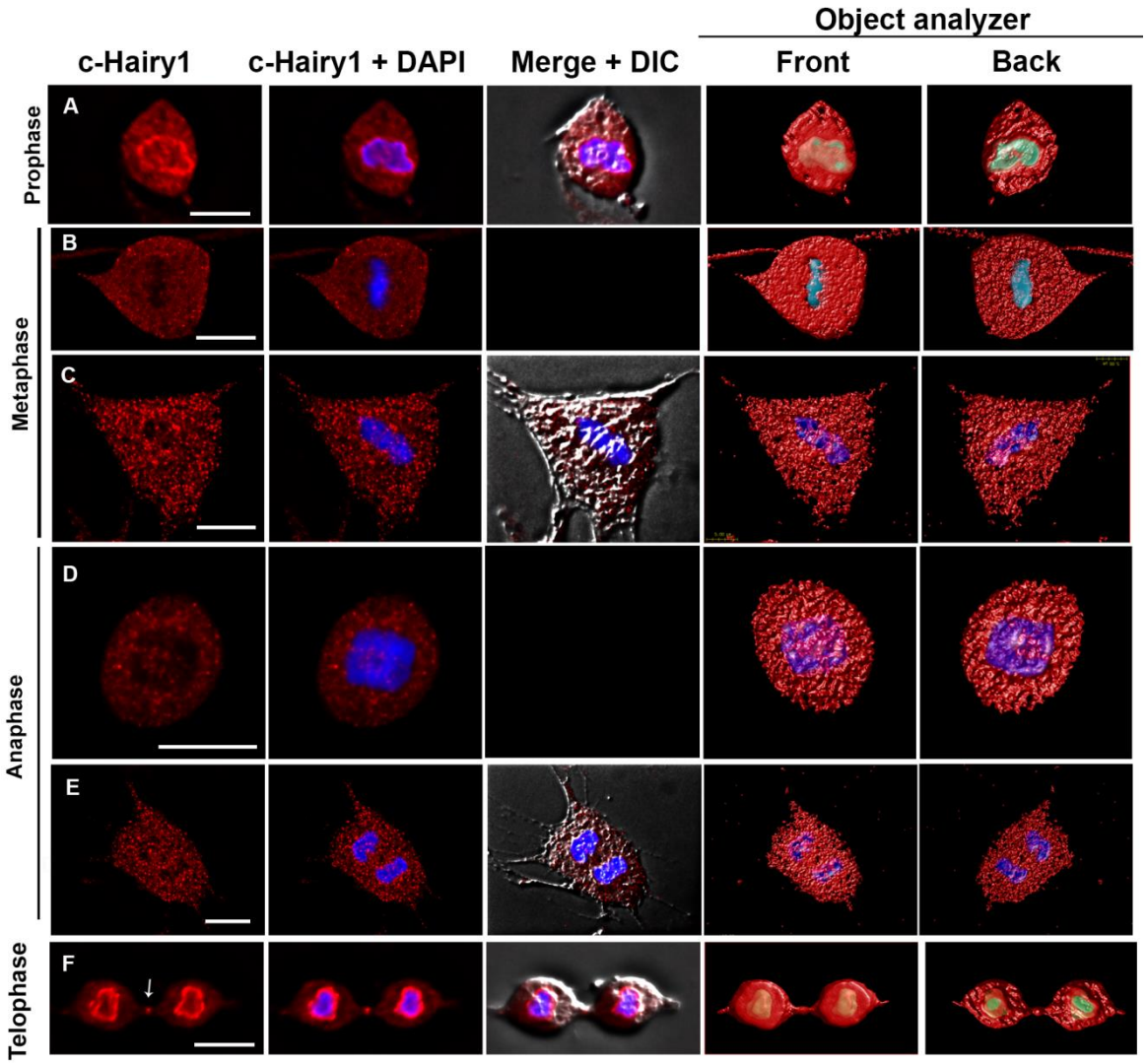


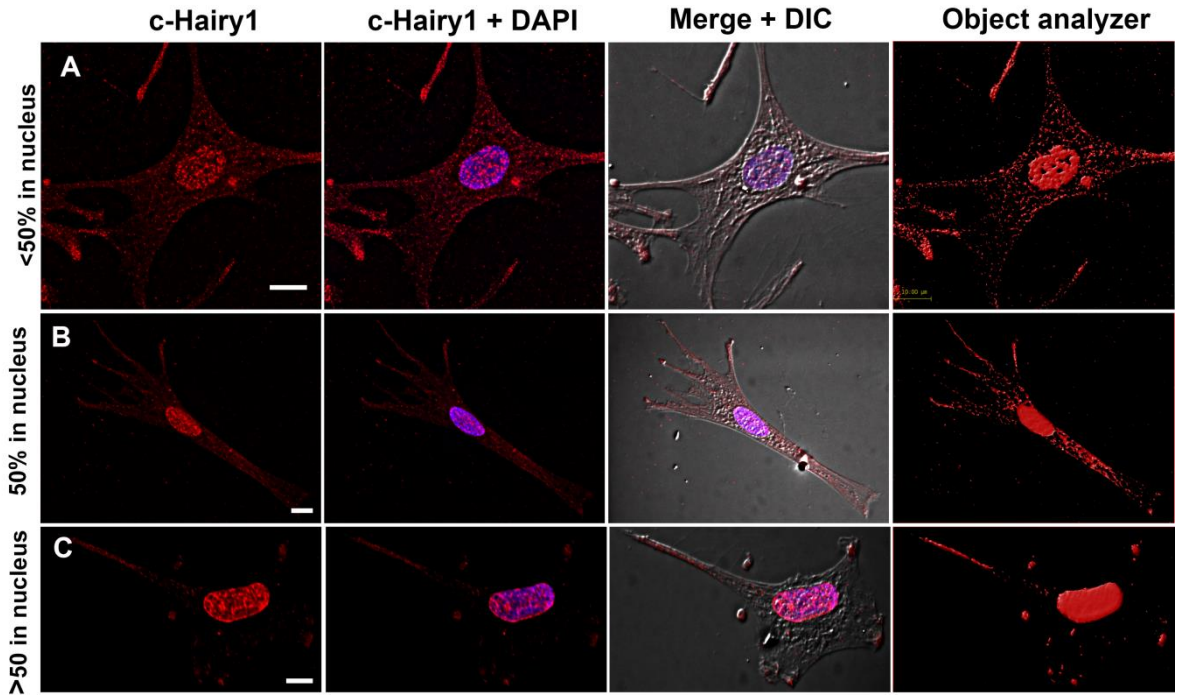
Figure 3.15 – Expression of c-Hairy1 protein in mitotic cells by fluorescence microscopy of chicken embryonic fibroblasts. *(A-F)* C-hairy1 shows an even expression along the entire cell, although absent in DNA localization (Blue, DAPI). This expression is observable in all phases of mitosis. *(B,D)* DIC images were not available. *(F)* In telophase is possible to identify high expression of c-Hairy1 in the contractile ring

(arrow). This expression is noticeable in the 2D images and in the 3D image, converted by SVI's Huygens software. Immunodetection of endogenous c-Hairy1 in CEFs, using a dilution of 1:50 for c-hairy 1 antibody, and 1:250 for secondary antibody Alexa Fluor 568. Nuclei were stained with DAPI (blue). Fluorescence microscopy (original magnification: 10000x). Images were acquired with the same exposition time. Deconvolution was performed on images to increase the definition and clearance of the signal. Scale bar: 10 µm.

As different fluorescence intensities were observable between cells, a study was performed to determine the subcellular distribuition of c-Hairy1, and analyze the differences between cells. The Huygens software interactive object analyzer tool allowed the calculation of several geometrical parameters of the cell, such as the manually outlined cell volume, all the c-Hairy1 objects that were contained in that region of interest (ROI) the volume of the nucleus, using the DAPI staining as reference, and the objects that were contained in the nucleus. The cytoplasm volume was determined, subtracting the volume of the nucleus to the whole cell volume.

In this study, 50 cells were analyzed, and after obtaining the volume data correspondent to the cells, several parameters were calculated, such as the signal density of the c-Hairy1 objects volume in relation to the whole cell volume and the percentages of c-Hairy1 objects contained in nucleus and cytoplasm. These parameteres were then correlated and a statistic study was performed (Figure 3.16A-D).

Results indicate c-Hairy1's density to be variable between 0,05 to 0.31, in a 0 to 1 scale, with an average of 0,15 ± 0,07, and the values variation followed a normal distribution. The percentage of c-Hairy1 in the nucleus ranged from 27,12 to 88,02%, with an average of 64,63% ± 14,62, with an excepction of cells that raged from 80-90% of c-Hairy1 in nucleus (Figure 3.16 E, F).



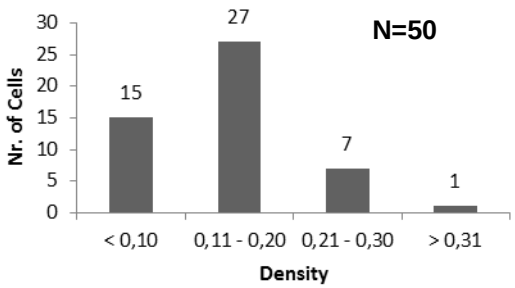
D

		A	B	C
Total Volume (μm^3)	Whole Cell	2953,20	4236,00	6056,80
	Nucleus	270,33	598,85	800,00
	Cytoplasm	2682,87	3637,15	5256,80
c-Hairy1 Volume (μm^3)	Whole Cell	677,09	721,27	1118,19
	Nucleus	183,66	381,35	784,66
	Cytoplasm	493,43	339,92	333,53
	Density*	0,23	0,17	0,18
Percentage (%)	Nucleus**	27,12	52,87	70,17
	Cytoplasm***	72,88	47,12	29,82

*- Whole cell c-Hairy1 volume/ Whole cell total volume.

** - Nucleus c-Hairy1 volume/ Whole cell c-Hairy1 volume.

E



F

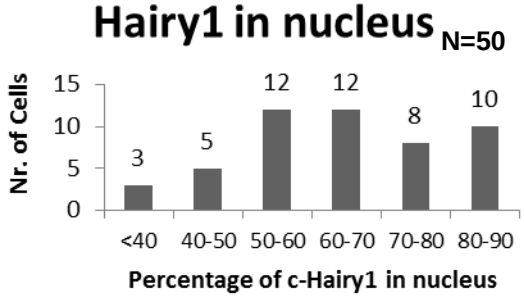


Figure 3.16 – Differential c-Hairy1 distribution in nucleus and cytoplasm. *(A-C)* Hairy1 protein is differently distributed in nucleus and cytoplasm from cell to cell. *(A)* C-Hairy1 is majorly expressed in cytoplasm than in the nucleus. *(B)* C-hairy1 is distributed in the same proportion in nucleus and cytoplasm. *(C)* C-Hairy1 majorly distributed in the nucleus. *(D)* Table that comprises the determined cell volumes and its compartments, the volume occupied by c-Hairy1 objects, density, and percentages. (*E,F)* Statistics study of the variation of c-Hairy1's expression in cells. *(E)* Variation of densities of the c-Hairy1 objects volume in relation to the whole cell volume (n= 50; average= 0,15 ± 0,07; min= 0,05; max=0,36). *(F)* Variation of the percentage of c-Hairy1 objects contained in nucleus (n=50; average=64,63 ± 14,62; min= 27,12; max: 88,02). Results were determined by the analysis of volume occupied by the cell and c-Hairy1 objects, converted by Huygens Professional Objects Analyzer tool. Immunodetection of endogenous c-Hairy1 in CEFs, using a dilution of 1:50 for c-hairy 1 antibody, and 1:250 for secondary antibody Alexa Fluor 568. Nuclei were stained with DAPI (blue). Fluorescence microscopy (original magnification: 10000x). Images were acquired with the same exposition time. Deconvolution was performed on images to increase the definition and clearance of the signal. Scale bar: 10 µm.

CHAPTER 4

DISCUSSION

4. DISCUSSION

4.1. C-HAIRY1 IS LIKELY TO BE A GLOBULAR PROTEIN

Bioinformatics allows the study of biological data in many ways, for instance, it makes possible the identification of genes and their respective proteins, prediction and comparison of protein structures, multiple sequences alignment (MSA), studies of homology of sequences, and many other procedures. Although bioinformatics undeniable potential, the results obtained must be carefully analyzed, since they can lead to misinterpretations and errors. As so, it is important to evaluate the data through different and complementary strategies, bearing in mind that the results obtained are estimated, still lacking biological confirmation (Mount, 2004).

Biochemical parameters prediction indicated c-Hairy1 to be a small protein, as it revealed a molecular weight of approximately 30,6 kDa, (Table 3.1). The small size may facilitate the protein synthesis, degradation and transportation in the cell.

Regarding the theoretical pI, c-Hairy1A is negatively charged at physiological pH (approximately 8 (Sieger et al., 1993)). The negative GRAVY value suggests that c-Hairy1 is a hydrophilic protein.

Amino acid composition analysis demonstrated that alanine, proline and serine presented higher abundance percentage, of approximately 12, 11 and 9,6% respectively. Amino acid composition analysis demonstrated that alanine, proline and serine presented higher abundance percentage, of approximately 12, 11 and 9,6% respectively (Table 3.2). Alanine is one of the most common amino acids occurring in proteins (Fasman, 1989), consistent with the high abundance observed. Proline is often found in protein "turns" (Krieger et al., 2005), suggesting a predisposition of c-Hairy1 to contain a turns in its structure. Serine, alongside with threonine and tyrosine, possesses phosphorylation properties (Blom et al., 1999), increasing the propensity of c-Hairy1 to be phosphorylated.

The domain analysis of the protein sequences (Table 3.3) detected all the described conserved domains belonging to the Hes protein family (Kageyama et al., 2007).

Regarding the secondary structure, the protein showed a mainly coiled structure, presenting some α-helices and a low percentage of β-strands most of them localized within the conserved domains (Figure 3.3). As expected, the HLH domains presented two α-helices separated by a 5 residue coiled region, which can correspond to the loop, reinforcing the accuracy of this prediction (Figure 3.3). The Orange domain was also predicted to have a more organized structure, consisting of α-helices and β-strands, and presenting low coiled structure (Figure 3.3). This structural organization must be conserved as it may confer the domains their activity.

The secondary structure doesn't provide the information of the tertiary structure, since currently it can only be properly determined by experimental techniques, such as X-ray crystallography or nuclear magnetic resonance spectroscopy of proteins (NMR) (Jin and Dunbrack, 2005). However, the secondary structure may present a clue on how structured is the protein, and the high percentage of coiled regions alongside with its hydrophilic properties may confer a globular structure to the protein.

4.2. BIOINFORMATIC RESULTS SUGGEST A HIGH PROPENSITY OF C-HAIRY1 TO BE PHOSPHORYLATED

NetPhos and DISPHOS softwares were used to predict the phosphorylation sites, in order to increase the accuracy and sensitivity of the analysis. Both softwares detected the same number of possible phosphorylation sites, although with different scores assigned to the residues (Figure 3.1). For a more stringent analysis, only the residues that scored higher than 0.5 in both softwares were selected. This restricted to 17 probable phosphorylation sites, where almost half of which (8/17) presented an average score higher than 0.9 (Table 3.4).

Regarding the distribution of these phosphorylation sites, three zones of phosphorylation could be observed at both protein ends, and in the HLH possible loop (Figure 3.3), indicating greater exposure to kinases.

The surface accessibility analysis predicted which residues were exposed to the surface and accessible to kinases or other interactor proteins with c-Hairy1. This way, combining the phosphorylation sites distribution and surface accessibility results could increase the accuracy of the prediction. Results revealed that all the selected phosphorylation sites were predicted as exposed residues (Figure 3.3).

The c-Hairy1 and its homologs sequence alignment revealed that 9 of 17 predicted phosphorylation sites were highly conserved in all the studied proteins, indicating these sites to be crucial for regulation of protein function (Figure 3.5).

A wide variety of kinases were predicted as having affinity to phosphorylate c-Hairy1A (Table 3.5). However, the correlation of the obtained results in Table 3.5 and the most probable phosphorylation sites (Table 3.4), narrowed to 10 the kinases shown to have affinity to the most probable phosphorylation sites (Figure 3.2).

The predicted kinases have been implicated in a variety of functions. Table 4.1 gives an insight on the different functions in which the predicted kinases have been involved.

Table 4.1 - Described cell functions to the predicted kinases.

Kinase	Function	Reference
ATM	• DNA damage signaling; • Cell cycle regulation; • Involved in apoptotic events.	(Lee and Paull, 2007)
cdK5	• Neuronal activity modulation; • Neuronal migration regulation during development; • Neurite outgrowth	(Wang et al., 2006)
CKII	• Cell cycle regulation. • Involved in: • Gene expression, protein synthesis, cell proliferation, apoptosis.	(Pinna and Meggio, 1997) (Gao and Wang, 2006)
DNAPK	• DNA damage signaling; • Cell cycle regulation.	(Moll et al., 1999)
GSK3	• Cell proliferation and migration regulation; • Inflammation; • Immune responses mediation; • Glucose regulation; • Apoptosis regulation.	(Jope et al., 2007)
p38 MAPK	• Immune response mediation; • Inflammation	(Johnson and Lapadat, 2002)
PKA	• Exocytosis regulation and modulation • Cell Cycle Regulation • cAMP signaling in cells	(Seino and Shibasaki, 2005; Stork and Schmitt, 2002)
PKC	• Membrane structure events modulation; • Transcription regulation; • Immune responses mediation; • Cell growth regulation;	(Newton, 1995)
PKG	• Smooth muscle tone regulation; • Platelet activation inhibition; • Neuronal functions modulation.	(Casteel et al., 2010)
RSK	• Cell growth regulation; • Cell cycle regulation; • Nuclear signaling.	(Dufner and Thomas, 1999) (Roux et al., 2003)

Many of the predicted kinases are involved in cell cycle and proliferation regulation, which is consistent with evidences already published role of c-Hairy1A to participate in the maintenance of cell undifferentiation, and promotion of proliferation (Andrade et al., 2007). The cdk5 and PKG are involved in neuronal modulation and development, which agrees with studies that associate Hes family genes with neurogenesis regulation (Kageyama et al., 2007).

The affinity of CKII to phosphorylate c-Hairy1A is interesting, since experimental evidence shows that CKII phosphorylates drosophila Hairy (Kahali et al., 2008). Other studies have reported the involvement of CKII in molecular clocks such as circadian, as it is a well-conserved clock component modulating the stability and subcellular localization of essential clock proteins (Allada and Meissner, 2005; Lu et al., 2011; Mehra et al., 2009). These evidences can provide a hint that c-Hairy1 localization may be modulated by phosphorylation of CKII or other kinases.

It becomes important to note that the fact that this software only makes predictions for some kinases, it doesn't exclude the possibility that c-Hairy1 might be phosphorylated by other kinases.

4.3. C-HAIRY1 IS PREDICTED TO BE LOCALIZED IN BOTH NUCLEUS AND CYTOPLASM

As the main goal of this work was to describe c-Hairy1 localization in the cell, it would be important to verify the presence of some signals in the sequences that could provide a hint on its subcellular localization. Experimental validation of subcellular locations is expensive and time-consuming, and, bioinformatics tools can make fast and accurate predictions (Brameier et al., 2007).

A search for c-Hairy1A homologs was advantageous for the prediction of subcellular localization signals, since a conserved nuclear localization signal (NLS) reinforces the accuracy of the prediction. BLASTp identified several possible homologs, although only the homologs belonging to some developmental animal models were selected. All of the selected proteins belong to Hes protein family (Table 1.2), with maximum identity values raging form 59-95%. A phylogenetic tree provided information on how genetically distant these proteins are. Two protein clusters could be observed, one of non-mammalian, including c-Hairy1A and another of mammalian homologs. Consistent with blast results the mammalian proteins revealed a greater distance to c-Hairy1A (Figure 3.4).

The protein transport to the nucleus is mediated by short binding sites on the protein sequence, called NLSs (Brameier et al., 2007). The task of searching for NLS was performed with NucPred and PSORT II. NucPred results assigned a score higher than 0,5 to all of the studied proteins, which could indicate that these spend at least some time in the nucleus. However, these results must be analyzed based on a threshold representing different specificities and sensitivities of the prediction (See table 2.2, Chapter 2). When considering a threshold of 0.5, all the proteins are considered to spend some time in the nucleus (Table 3.7). Evaluating the multiple alignment of the homologs, a conserved NLS was detected, KRRR, in the c-terminus end of basic region of the proteins (Figure 3.5), supporting the prediction of an NLS on these proteins. PSORT II results revealed that all the homologs had higher probability to be located in the nucleus, than in other cell compartments, with low percentage variability between homologs (Table 3.7).

Overall, all bioinformatics evidences suggest c-Hairy1A and its homologs to be nuclear proteins, and any indication suggesting an opposite idea hasn't been found, consistent with the attributed function to these proteins as transcription factors.

NetNES 1.1 searched for possible nuclear export signals (NES), which consist on a short amino acid sequence of 4 hydrophobic residues, that targets the protein for export from the cell nucleus to the cytoplasm through the nuclear pore complex (la Cour et al., 2004). This transport is most likely lead by biding NES to the exportin CRM1 (Fischer et al., 1995).

The NETNES 1.1 identified two regions with a pronounced NES activity, in which a deep analysis recognized two potential NES localized in the HLH domain, as these sequences have followed the consensus $\Phi1\text{-}(x)_{2-3}\text{-}\Phi2\text{-}(x)_{2-3}\text{-}\Phi3\text{-}x\text{-}\Phi4$, even though they did not correspond perfectly to it. The NES2's hydrophobic residues have not been assigned high scores, with the exception of the isoleucine (Φ2), however this motif followed the consensus sequence, and for that reason, it is more likely to be a potential NES than the NES1.

In these results, the potential NESs located region can be observed in every sequence, where higher scores were registered, highlighted from the rest of the sequence (Figure 3.6). The mammalian homologs, however, in comparison with the non-mammalian ones presented highest peaks and scores in more than one residue, which may indicate a higher NES activity on mammalian homologs (Figure 3.6).

4.4. THE MONOCLONAL ANTIBODY RECOGNIZES C-HAIRY1 PROTEIN

The determination of the most adequate loading amount of purified protein was a crucial step in order to test the sensitivity of the antibody. The antibody was not only capable to detect c-Hairy1 but also to detect it in 35 ng of protein extract.

Western blot results revealed the presence of two persistent bands in the purified protein, one with a higher molecular weight of 56 kDa, and another with 35 kDa. The 35 kDa band was considered to represent c-Hairy1A protein, since the purified protein was produced from the c-Hairy1A sequence, and for that reason it would serve as guide for its identification in protein extracts (Figure 3.7). The molecular weight difference between the bioinformatically predicted 30.7 kDa and the experimental determined 35 kDa, may be due to a weaker capacity of migration of the protein in the gel. Also the marker's buffer may provide the marker an different migration dynamics than the samples buffer.

The protein was purified by cloning c-Hairy1A c-DNA fused to a His-tag, and the company guaranteed a purification degree of minimum 90%. This suggests that the higher band would not be result of other unspecific proteins and may be derived from c-Hairy1 protein interactions, for instance dimer formation. Data from literature also supports the dimer formation hypothesis, as c-Hairy1 has a great propensity to form dimers (Sheeba et al., 2007), The dimer binding may be so resistant that even with strong denaturing conditions it fails to totally eliminate the interactions between the proteins. This hypothesis could explain why the gradual increasing of the amount of purified protein did not represent a gradual increase of

the 56 kDa band, since the ability to form dimers may be relative and dependent on the denaturing conditions.

Theoretically, a dimer would be expected to be twice the molecular weight of a single protein, though a band with such molecular weight was never found. However, the denaturing conditions may be causing different conformations of the homodimers, which may result in different folding degrees, explaining the smear and blurry band observed most of the times. Also dimers could present a conformation that migrates easily on the gel and would confer the 56 kDa band. On the contrary, the 35 kDa appeared always much sharper and defined than the 56 kDa band, indicating a stable migration of this denatured conformation of the protein in the gel.

The most suitable loading amount of purified c-Hairy1 was considered to be 100 ng of protein/lane, as results revealed it to be the lowest amount in which the 35 kDa band was better visualized.

Having demonstrated that the antibody showed affinity to the purified protein, the assessment process continued with the detection of c-Hairy1 in protein extracts, using the purified protein as a positive control. Despite the blur and lack of definition, several bands were identified in the embryo protein extract lanes, two of them also detected in the purified protein lane, indicating a sensitivity of the antibody to the embryo c-Hairy1 (Figure 3.8).

The need of high amount of extracts used for the antibody proper function, can suggest two conditions, the antibody may have low sensitivity for detection of c-Hairy1 in the extracts or the protein may be present at low concentrations in the embryo extracts, however data not shown revealed the presence of the 56 kDa band in lower loading amounts of embryo protein extracts. In order to spare the most protein extract, and because 75 µg of protein/lane showed the best definition of the bands, it was considered as the most appropriate loading amount of embryo protein extract (Figure 3.8).

Although the expression of c-Hairy1 in chicken embryonic fibroblasts had never been published, an oscillatory expression of Hes1 (c-Hairy1 homolog) in

mouse embryonic fibroblasts has been demonstrated (Hirata et al., 2002). For that matter, since the antibody was already proven to recognize c-Hairy1 in the purified protein and embryo extracts by western blot, the conditions were settled to assess the expression of c-Hairy1 in CEFs extracts. The best loading amount had to be established, and results revealed that in amounts superior to 50 µg of protein/lane it was possible to detect bands of c-Hairy1. The 75 µg of protein/lane was the amount in which the best definition of the bands was observable. Similar to embryo extracts, the CEF extracts blotting for c-Hairy1 detected several bands with different signal intensities (Figure 3.9).

Also, several denaturation times were tested in order to investigate whether the high molecular weight bands were resultant of dimer formation. Despite of the blurry results, it was noticeable that increasing the denaturation time did not completely eliminate the higher bands of 56, 79 and 92 kDa, however it decreased these bands intensities. The lower bands of 37, 35 and 33 kDa were still observable in the samples denatured for 30 min, even though they didn't present much definition (Figure 3.10). Analyzing the purified protein signal, it is observable that the increase of the denaturation time lowered the protein detection signal, however becoming more delineated. The exposure of the membrane with Pico chemiluminescent, less sensitive than the Immun-Star WesternC kit, revealed that the c-Hairy1 bands in the embryo extracts were best detectable at 5 min, suggesting that higher denaturation time may be degrading the proteins, and for that reason it was chosen as the most suitable denaturation time (Figure 3.10).

4.5. C-HAIRY1 IS LIKELY TO BE EXPRESSED DIFFERENTLY IN CHICK EMBRYOS AND CEFS

After establishing the best SDS-PAGE and western blot conditions analysis of c-Hairy1 expression in embryo and CEF extracts, detected the protein in both extracts, with the same identified bands presenting different intensities (Figure 3.11). CEFs presented relatively stronger signal of the 42 and 33 kDa bands than the embryo, indicating a higher expression of these protein forms (Figure 3.11).

These differences may be caused by the fact that CEFs were isolated from a much later embryonic stage (14 days) than the chick embryos (10-20 somites), being already differentiated, unlike chick embryo cells which still present a high rate of proliferation and undifferentiated cells (Gilbert, 2010).

Three close molecular weight bands with 33, 35 and 37 kDa were identified in the protein extracts lanes, values approximated to the predicted c-Hairy1 molecular weight of 30.7 kDa. The 35 kDa band identified in the purified c-Hairy1 protein lane was considered to represent the isoform c-Hairy1A. This band was also clearly identifiable in the protein extracts, indicating the expression of this isoform in chick embryo end CEFs. As previously described in the introduction, c-Hairy1 has two known isoforms, c-Hairy1A and B (Vasiliauskas et al., 2003) , of which the last has an insertion of 14 amino acids which results in an estimated difference of 1,6 kDa. Taking these results into account, the 37 kDa may be considered to represent c-Hairy1B isoform, while the 33 kDa band remains unidentified, not excluding the possibility of the existence of another c-Hairy1 isoform. As bioinformatic results predicted c-Hairy1 to have high propensity to be phosphorylated, it could also be suggested that different phosphorylation degrees may explain these close molecular weight bands.

The other detected bands of 75, 70 56, 42 (Figure 3.11) may be due to dimer formation, alternative splicing, unpublished isoforms or affinity of the antibody to other proteins. Post-translational modifications, such as ubiquitination may be causing the detection these bands, as it is known that c-Hairy1 has to be ubiquitinated in order to be degraded by proteasome (Kageyama et al., 2007). Ubiquitin is a small protein with a molecular weight of 8,5 kDa, that covalently marks proteins for degradation (Hochstrasser, 2009). This bond may be so resistant that the strong denaturing SDS-PAGE conditions cannot break them, increasing the c-Hairy1 protein molecular weight.

Being a monoclonal antibody greatly improves the antibody specificity, as all the proteins that were detected contained the same epitope against what the antibody was produced (Alberts et al., 2007). On the other hand, most of the

probable proteins the antibody would have affinity to would be member of the Hes family proteins, such as Hes1-B-like, or even Hes5, however data from NCBI (HES-1-B-like [Gallus gallus], 213 aa, NCBI reference sequence: XP_003641836.1; HES-5 [Gallus gallus], 157 aa, NCBI reference sequence: NP_001012713.1) suggest these protein to be smaller than c-Hairy1 (390 aa), which doesn't explain the presence of the higher molecular weight bands.

The negative controls were tested in order to exclude the possibility that the unspecific affinity of secondary antibodies may be responsible for the detection of the unpredicted bands. As it was confirmed that the bands in question were not detected by the α-mouse negative controls, the results suggest that the unpredicted bands were originated only by affinity of the primary monoclonal antibody against c-Hairy1 (Figure 3.12). The α-rabbit negative control also showed no signal of affinity to any protein contained in the embryo extract or the purified c-Hairy1 protein, and, as expected, the β-tubulin1 labeling was detected only by the primary antibody conjugated with the secondary (Figure 3.12).

One way to prove whether the bands correspond to other forms of c-Hairy1 would be by identification of the bands by mass spectrometry, and confirmation of the c-Hairy1 identity. This analysis could only make sense sending the purified proteins bands for examination, since the process consists in removing the bands of interest directly from gel for analysis, and the protein extract lanes contain several proteins. Another validation of the specificity of the antibody would consist on using a morpholino for c-Hairy1 in CEFs, making protein extracts and blotting for c-Hairy1. In this case, the antibody would only detect unspecific bands. Overexpression of c-Hairy1 in CEFs would also help to verify the blotting results, as differences in c-Hairy1 expression would increase the detection signal of the specific bands.

Even though the specificity has not been totally clarified, the identification of the bands of interest (35 and 37 kDa) by western blot, and evidenced sensitivity of the monoclonal antibody to the c-Hairy1 protein, opens a door of possibilities for new studies on this protein. For instance, immunohistochemistry of c-Hairy1

protein in several tissues, studies on the oscillatory expression of the protein, functional studies, immunocytochemistry, or even studies on the subcellular localization of the protein using fractioned CEFs protein extracts.

4.6. IMMUNOFLUORESCENCE RESULTS INDICATE A DIFFERENT C-HAIRY1 DISTRIBUITION IN NUCLEUS AND CYTOPLASM

As the monoclonal antibody against c-Hairy1 had already been tested and proved to be working, it allowed the possibility to perform an immunofluorescence in order to determine the expression pattern of c-Hairy1 in cells. CEFs were chosen to perform this task, since c-Hairy1 expression has already been shown by western blot, and also they are relatively easy to isolate and maintain in culture.

Results from immunofluorescence indicated a high expression of c-Hairy1 in the nucleus, and some expression in the cytoplasm, with variable levels of intensity from cell to cell (Figures 3.13 and 3.16). These results were complemented with positive and negative controls, in which the positive control presented an expected staining of the cell cytoskeleton, and negative control revealed no specific expression. These results could be observed both in original and deconvolved images, however the deconvolution treatment to the images decreased the noisy signals, increasing the definition and clarity of the staining, allowing a better analysis of the results (Figure 3.13).

C-Hairy1 being described as a transcription factor would not be predictable to be localized in cytoplasm, consistent with the bioinformatic prediction of an NLS which indicated a high probability of c-Hairy1 to be a nuclear protein. Two potential NES motif have also been predicted in the c-Hairy1 sequence, raising the question of whether c-Hairy1 shuttles from nucleus to cytoplasm. In this case transportation proteins responsible for carrying c-Hairy1 would still need to be identified, being CMR1 a possible candidate, as the nucleus to cytoplasm transport is most likely lead by biding NES to the exportin CRM1. The shuttle may be regulated by the protein phosphorylation state and the high number of probable phosphorylation

sites may be involved in c-Hairy1 transportation. In order to validate this shuttle, is would be plausible to construct a vector with a Green Fluorescence Protein (GFP) fused to a protein extremity and evaluate in vivo the shuttle of c-Hairy from nucleus to cytoplasm.

A possible experimental validation of this shuttle would consist on the inhibition of CRM1 with Leptomycin B, and in this way verifying the c-Hairy1 exportation and also if CRM1is indeed exporting c-Hairy1 from the nucleus.

Results from an unpublished Bimolecular Fluorescence Complementation (BiFCo) performed in Isabel Palmeirim's lab have identified the c-Hairy1 homodimer formation in Hella cells cytoplasm (Palmeirim et al., unpublished). Two different cloned vectors containing c-Hairy1 cDNA, each of them possessing the N and C-terminus portions of the Yellow Fluorescent protein (YFP) were transfected. Since the YFP protein portion have no affinity to each other, the interaction of the YFP portions must have been driven by c-Hairy1 dimerization (Kerppola, 2006). This data is consistent with the found c-Hairy1 cytoplasm localization and reinforces the idea of a biological function of c-Hairy1 in this cell compartment, since the dimer is a functional active conformation of c-Hairy1 (Sheeba et al., 2007)

Nevertheless, the presence of c-Hairy1 in cytoplasm may be also justified by proteosomal degradation of protein, as it has to migrate to the cytoplasm in order to be degraded.

The detailed analysis of c-Hairy1's expression indicated it to be highly expressed at cells leading edges, with a registered occurrence in 41% of the analyzed cells (Figure 3.14). The leading edges structures are highly dynamic, based on actin filaments and are implied on cell motility, propelling it or functioning as anchors (Alberts et al., 2007). This peculiar localization of c-Hairy1 is curious, as it may give a clue of the interaction of c-Hairy1 with cytoskeleton proteins. Other hints have been provided by results of a non-published yeast-two-hybrid screening assay that revealed interaction of the c-Hairy1 c-terminus portion with some cytoskeleton protein, such as Stathmin1, Dynactin6, Tubulin beta chain (TUBB)

(Andrade, et al unpublished). In the future, an immuno co-localization between c-Hairy1 and these proteins would be useful to better understand its interactions.

In mitotic cells c-Hairy1 is homogeneously distributed throughout the cell (Figure 3.15). The high expression of c-Hairy1 in mitotic cells revealed to be an interesting finding, which could correlate with the described function of c-Hairy1 as a repressive transcription factor that prevents cellular differentiation, and promotes cellular proliferation. As so, high levels of c-Hairy1 may be necessary for mitosis and its proper occurrence. The actin and myosin contractile ring expression would reinforce even more the interaction between the c-Hairy1 and cytoskeleton proteins.

Intriguingly, none of the analyzed cells were absent of c-Hairy1's expression, although varied levels of c-Hairy1 signal density between cells were noticeable, as well as the percentage of c-Hairy1 localized in nucleus *versus* cytoplasm. Relatively to the density variation, values ranged from 0,05 to 0,31 with an average of 0.15 ± 0,07, which can be considered as the lowest and highest levels of expression of this protein.

The study revealed that 84% of the cells present more than 50% of c-Hairy1 in the nucleus, which is consistent with the previous predictions to be localized mainly in the nucleus. A direct correlation between a great percentage in nucleus and the density of c-Hairy1's expression is not assumable, since the nucleus only occupies a small fraction of the whole cell. In this case, a high percentage in the nucleus may be inversely proportional to the signal density of c-Hairy1 in the whole cell.

The analysis of c-Hairy1 density variation from cell to cell followed a normal distribution, as it would be expected in a cyclically expressed protein (Figure 3.16E). For that matter, the minimum levels of c-Hairy1 encountered may be correspondent to be the initiation or the ending of an expression cycle and the 0.36 correspondent to the highest peak on the expression cycle. The variation of the percentage of c-Hairy1 in nucleus also followed a normal distribution with the exception of the point that ranged from 80-90% (Figure 4F). These evidences

indicate a dynamic distribution of c-Hairy1 in cell, supporting the suggestion of the protein shuttle between the nucleus and cytoplasm,

CHAPTER 5

CONCLUSION

5. CONCLUSION

The central aim of this work was to study the c-Hairy1 protein in order to increase the understanding of its working mechanism in the cell. The bioinformatic prediction of c-Hairy1 protein regarding its biochemical properties, secondary structure and even the subcellular localization contributed to the understanding and supporting of the experimental results.

The optimal working conditions of the monoclonal antibody, as well as the most adequate sample preparation criteria were established. The detection of the expected protein bands in the purified protein, and the tested protein extracts, indicated a sensitivity of the antibody and confirmed the expression of c-Hairy1 in chick embryo and chicken embryonic fibroblasts. Other bands with different molecular weight were also systematically detected, even in severe denaturation conditions. For that reason, more studies may be performed to prove the specificity of the antibody.

The subcellular localization of c-Hairy1, predicted by the bioinformatics analysis, was confirmed by the immunofluorescence results that revealed c-Hairy1 in both nucleus and cytoplasm. Moreover, a detailed analysis of c-Hairy1 localization showed that the nucleus:cytoplasm ratio distribution varied between cells. However, a biological function of the protein in the cytoplasm is far from being completely understood. In the future, and taking advantages of the monoclonal antibody, several studies may be performed in order to address this question. Further studies will be also needed to unravel an oscillatory property of the protein.

Overall, the findings herein presented uncover new aspects of the c-Hairy1 protein and will hopefully open several possibilities for further studies relevant to the line of work that has been developed in Palmeirim's lab.

CHAPTER 6

REFERENCES

6. REFERENCES

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2007). Molecular Biology of the Cell. New York: Garland Science.

Allada, R. and Meissner, R. A. (2005). Casein kinase 2, circadian clocks, and the flight from mutagenic light. *Mol Cell Biochem* **274**, 141-9.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-402.

Andrade, R. P., Palmeirim, I. and Bajanca, F. (2007). Molecular clocks underlying vertebrate embryo segmentation: A 10-year-old hairy-go-round. *Birth Defects Res C Embryo Today* **81**, 65-83.

Atchley, W. R. and Fitch, W. M. (1997). A natural classification of the basic helix-loop-helix class of transcription factors. *Proc Natl Acad Sci U S A* **94**, 5172-6.

Aulehla, A. and Pourquie, O. (2008). Oscillating signaling pathways during embryonic development. *Curr Opin Cell Biol* **20**, 632-7.

Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B. and Herrmann, B. G. (2003). Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Dev Cell* **4**, 395-406.

Bae, S., Bessho, Y., Hojo, M. and Kageyama, R. (2000). The bHLH gene Hes6, an inhibitor of Hes1, promotes neuronal differentiation. *Development* **127**, 2933-43.

Baker, R. E., Schnell, S. and Maini, P. K. (2006). A clock and wavefront mechanism for somite formation. *Dev Biol* **293**, 116-26.

Bessho, Y. and Kageyama, R. (2003). Oscillations, clocks and segmentation. *Curr Opin Genet Dev* **13**, 379-84.

Bessho, Y., Miyoshi, G., Sakata, R. and Kageyama, R. (2001). Hes7: a bHLH-type repressor gene regulated by Notch and expressed in the presomitic mesoderm. *Genes Cells* **6**, 175-85.

Blom, N., Gammeltoft, S. and Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* **294**, 1351-62.

Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S. and Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* **4**, 1633-49.

Brameier, M., Krings, A. and MacCallum, R. M. (2007). NucPred--predicting nuclear localization of proteins. *Bioinformatics* **23**, 1159-60.

Casteel, D. E., Smith-Nguyen, E. V., Sankaran, B., Roh, S. H., Pilz, R. B. and Kim, C. (2010). A crystal structure of the cyclic GMP-dependent protein kinase I{beta} dimerization/docking domain reveals molecular details of isoform-specific anchoring. *J Biol Chem* **285**, 32684-8.

Castella, P., Sawai, S., Nakao, K., Wagner, J. A. and Caudy, M. (2000). HES-1 repression of differentiation and proliferation in PC12 cells: role for the helix 3-helix 4 domain in transcription repression. *Mol Cell Biol* **20**, 6170-83.

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

Cooke, J. and Zeeman, E. C. (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *Journal of Theoretical Biology* **58**, 22.
Dale, J. K., Malapert, P., Chal, J., Vilhais-Neto, G., Maroto, M., Johnson, T., Jayasinghe, S., Trainor, P., Herrmann, B. and Pourquie, O. (2006). Oscillations of the snail genes in the presomitic mesoderm coordinate segmental patterning and morphogenesis in vertebrate somitogenesis. *Dev Cell* **10**, 355-66.
Darnell, D. K., Stark, M. R. and Schoenwolf, G. C. (1999). Timing and cell interactions underlying neural induction in the chick embryo. *Development* **126**, 2505-14.
Davis, R. L., Turner, D. L., Evans, L. M. and Kirschner, M. W. (2001). Molecular targets of vertebrate segmentation: two mechanisms control segmental expression of Xenopus hairy2 during somite formation. *Dev Cell* **1**, 553-65.
Dawson, S. R., Turner, D. L., Weintraub, H. and Parkhurst, S. M. (1995). Specificity for the hairy/enhancer of split basic helix-loop-helix (bHLH) proteins maps outside the bHLH domain and suggests two separable modes of transcriptional repression. *Mol Cell Biol* **15**, 6923-31.
Deblandre, G. A., Wettstein, D. A., Koyano-Nakagawa, N. and Kintner, C. (1999). A two-step mechanism generates the spacing pattern of the ciliated cells in the skin of Xenopus embryos. *Development* **126**, 4715-28.
Delfini, M. C., Dubrulle, J., Malapert, P., Chal, J. and Pourquie, O. (2005). Control of the segmentation process by graded MAPK/ERK activation in the chick embryo. *Proc Natl Acad Sci U S A* **102**, 11343-8.
Dequéant, M. L., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A. and Pourquie, O. (2006). A complex oscillating network of signaling genes underlies the mouse segmentation clock. *Science* **314**, 1595-8.
Dequeant, M. L. and Pourquie, O. (2008). Segmental patterning of the vertebrate embryonic axis. *Nat Rev Genet* **9**, 370-82.
Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M. et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**, W465-9.
Dubrulle, J., McGrew, M. J. and Pourquie, O. (2001). FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* **106**, 219-32.
Dufner, A. and Thomas, G. (1999). Ribosomal S6 kinase signaling and the control of translation. *Exp Cell Res* **253**, 100-9.
Dunwoodie, S. L., Clements, M., Sparrow, D. B., Sa, X., Conlon, R. A. and Beddington, R. S. (2002). Axial skeletal defects caused by mutation in the spondylocostal dysplasia/pudgy gene Dll3 are associated with disruption of the segmentation clock within the presomitic mesoderm. *Development* **129**, 1795-806.
Fasman, G. D. (1989). Prediction of Protein Structure and the Principles of Protein Conformation, (ed. New York: Springer.
Fior, R. and Henrique, D. (2005). A novel hes5/hes6 circuitry of negative regulation controls Notch activity during neurogenesis. *Dev Biol* **281**, 318-33.

Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W. and Luhrmann, R. (1995). The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**, 475-83.
Fisher, A. and Caudy, M. (1998). The function of hairy-related bHLH repressor proteins in cell fate decisions. *Bioessays* **20**, 298-306.
Forsberg, H., Crozet, F. and Brown, N. A. (1998). Waves of mouse Lunatic fringe expression, in four-hour cycles at two-hour intervals, precede somite boundary formation. *Curr Biol* **8**, 1027-30.
Fortini, M. E. (2002). Gamma-secretase-mediated proteolysis in cell-surface-receptor signalling. *Nat Rev Mol Cell Biol* **3**, 673-84.
Gajewski, M. and Voolstra, C. (2002). Comparative analysis of somitogenesis related genes of the hairy/Enhancer of split class in Fugu and zebrafish. *BMC Genomics* **3**, 21.
Gao, Y. and Wang, H. Y. (2006). Casein kinase 2 Is activated and essential for Wnt/beta-catenin signaling. *J Biol Chem* **281**, 18394-400.
Garzia, L., Andolfo, I., Cusanelli, E., Marino, N., Petrosino, G., De Martino, D., Esposito, V., Galeone, A., Navas, L., Esposito, S. et al. (2009). MicroRNA-199b-5p impairs cancer stem cells through negative regulation of HES1 in medulloblastoma. *PLoS One* **4**, e4998.
Gawantka, V., Pollet, N., Delius, H., Vingron, M., Pfister, R., Nitsch, R., Blumenstock, C. and Niehrs, C. (1998). Gene expression screening in Xenopus identifies molecular pathways, predicts gene function and provides a global view of embryonic patterning. *Mech Dev* **77**, 95-141.
Giampietro, P. F., Dunwoodie, S. L., Kusumi, K., Pourquie, O., Tassy, O., Offiah, A. C., Cornier, A. S., Alman, B. A., Blank, R. D., Raggio, C. L. et al. (2009). Progress in the understanding of the genetic etiology of vertebral segmentation disorders in humans. *Ann N Y Acad Sci* **1151**, 38-67.
Gibb, S., Maroto, M. and Dale, J. K. (2010). The segmentation clock mechanism moves up a notch. *Trends in Cell Biology* **20**, 593-600.
Gibb, S., Zagorska, A., Melton, K., Tenin, G., Vacca, I., Trainor, P., Maroto, M. and Dale, J. K. (2009). Interfering with Wnt signalling alters the periodicity of the segmentation clock. *Dev Biol* **330**, 21-31.
Giebel, B. and Campos-Ortega, J. A. (1997). Functional dissection of the Drosophila enhancer of split protein, a suppressor of neurogenesis. *Proc Natl Acad Sci U S A* **94**, 6250-4.
Gilbert, S. F. (2010). Developmental Biology, (ed. Sunderland, MA, U.S.A: Sinauer Associates Inc.
Giudicelli, F. and Lewis, J. (2004). The vertebrate segmentation clock. *Curr Opin Genet Dev* **14**, 407-14.
Giudicelli, F., Ozbudak, E. M., Wright, G. J. and Lewis, J. (2007). Setting the tempo in development: an investigation of the zebrafish somite clock mechanism. *PLoS Biol* **5**, e150.
Gomez, C., Ozbudak, E. M., Wunderlich, J., Baumann, D., Lewis, J. and Pourquie, O. (2008). Control of segment number in vertebrate embryos. *Nature* **454**, 335-9.

Harden, M. V., Newton, L. A., Lloyd, R. C. and Whitlock, K. E. (2006). Olfactory imprinting is correlated with changes in gene expression in the olfactory epithelia of the zebrafish. *J Neurobiol* **66**, 1452-66.
Hayashi, S., Shimoda, T., Nakajima, M., Tsukada, Y., Sakumura, Y., Dale, J. K., Maroto, M., Kohno, K., Matsui, T. and Bessho, Y. (2009). Sprouty4, an FGF inhibitor, displays cyclic gene expression under the control of the notch segmentation clock in the mouse PSM. *PLoS One* **4**, e5603.
Herrgen, L., Ares, S., Morelli, L. G., Schroter, C., Julicher, F. and Oates, A. C. (2010). Intercellular coupling regulates the period of the segmentation clock. *Curr Biol* **20**, 1244-53.
Hirata, H., Yoshiura, S., Ohtsuka, T., Bessho, Y., Harada, T., Yoshikawa, K. and Kageyama, R. (2002). Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop. *Science* **298**, 840-3.
Hochstrasser, M. (2009). Origin and function of ubiquitin-like proteins. *Nature* **458**, 422-9.
Holley, S. A., Geisler, R. and Nusslein-Volhard, C. (2000). Control of her1 expression during zebrafish somitogenesis by a delta-dependent oscillator and an independent wave-front activity. *Genes Dev* **14**, 1678-90.
Horikawa, K., Ishimatsu, K., Yoshimoto, E., Kondo, S. and Takeda, H. (2006). Noise-resistant and synchronized oscillation of the segmentation clock. *Nature* **441**, 719-23.
Horton, P. and Nakai, K. (1997). Better prediction of protein cellular localization sites with the k nearest neighbors classifier. *Proc Int Conf Intell Syst Mol Biol* **5**, 147-52.
Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z. and Dunker, A. K. (2004). The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* **32**, 1037-49.
Ishibashi, M., Sasai, Y., Nakanishi, S. and Kageyama, R. (1993). Molecular characterization of HES-2, a mammalian helix-loop-helix factor structurally related to Drosophila hairy and Enhancer of split. *Eur J Biochem* **215**, 645-52.
Ishikawa, A., Kitajima, S., Takahashi, Y., Kokubo, H., Kanno, J., Inoue, T. and Saga, Y. (2004). Mouse Nkd1, a Wnt antagonist, exhibits oscillatory gene expression in the PSM under the control of Notch signaling. *Mech Dev* **121**, 1443-53.
Jiang, Y. J., Aerne, B. L., Smithers, L., Haddon, C., Ish-Horowicz, D. and Lewis, J. (2000). Notch signalling and the synchronization of the somite segmentation clock. *Nature* **408**, 475-9.
Jin, Y. and Dunbrack, R. L., Jr. (2005). Assessment of disorder predictions in CASP6. *Proteins* **61 Suppl 7**, 167-75.
Johnson, G. L. and Lapadat, R. (2002). Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. *Science* **298**, 1911-2.
Jope, R. S., Yuskaitis, C. J. and Beurel, E. (2007). Glycogen synthase kinase-3 (GSK3): inflammation, diseases, and therapeutics. *Neurochem Res* **32**, 577-95.

Jouve, C., Iimura, T. and Pourquie, O. (2002). Onset of the segmentation clock in the chick embryo: evidence for oscillations in the somite precursors in the primitive streak. *Development* **129**, 1107-17.
Jouve, C., Palmeirim, I., Henrique, D., Beckers, J., Gossler, A., Ish-Horowicz, D. and Pourquie, O. (2000). Notch signalling is required for cyclic expression of the hairy-like gene HES1 in the presomitic mesoderm. *Development* **127**, 1421-9.
Kageyama, R., Niwa, Y., Shimojo, H., Kobayashi, T. and Ohtsuka, T. (2010). Ultradian oscillations in Notch signaling regulate dynamic biological events. *Curr Top Dev Biol* **92**, 311-31.
Kageyama, R., Ohtsuka, T. and Kobayashi, T. (2007). The Hes gene family: repressors and oscillators that orchestrate embryogenesis. *Development* **134**, 1243-51.
Kahali, B., Trott, R., Paroush, Z. e., Allada, R., Bishop, C. P. and Bidwai, A. P. (2008). Drosophila CK2 phosphorylates Hairy and regulates its activity in vivo. *Biochemical and Biophysical Research Communications* **373**, 637-642.
Kamakura, S., Oishi, K., Yoshimatsu, T., Nakafuku, M., Masuyama, N. and Gotoh, Y. (2004). Hes binding to STAT3 mediates crosstalk between Notch and JAK-STAT signalling. *Nat Cell Biol* **6**, 547-54.
Katoh, M. (2004). Identification and characterization of human HES2, HES3, and HES5 genes in silico. *Int J Oncol* **25**, 529-34.
Kerppola, T. K. (2006). Visualization of molecular interactions by fluorescence complementation. *Nat Rev Mol Cell Biol* **7**, 449-56.
King, I. N., Kathiriya, I. S., Murakami, M., Nakagawa, M., Gardner, K. A., Srivastava, D. and Nakagawa, O. (2006). Hrt and Hes negatively regulate Notch signaling through interactions with RBP-Jkappa. *Biochem Biophys Res Commun* **345**, 446-52.
Klein, S. L., Strausberg, R. L., Wagner, L., Pontius, J., Clifton, S. W. and Richardson, P. (2002). Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative. *Dev Dyn* **225**, 384-91.
Krieger, F., Moglich, A. and Kiefhaber, T. (2005). Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *J Am Chem Soc* **127**, 3346-52.
Krol, A. J., Roellig, D., Dequeant, M. L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A. C. and Pourquie, O. (2011). Evolutionary plasticity of segmentation clock networks. *Development* **138**, 2783-92.
Kudoh, T., Tsang, M., Hukriede, N. A., Chen, X., Dedekian, M., Clarke, C. J., Kiang, A., Schultz, S., Epstein, J. A., Toyama, R. et al. (2001). A gene expression screen in zebrafish embryogenesis. *Genome Res* **11**, 1979-87.
Kutay, U. and Guttinger, S. (2005). Leucine-rich nuclear-export signals: born to be weak. *Trends Cell Biol* **15**, 121-4.
la Cour, T., Kiemer, L., Molgaard, A., Gupta, R., Skriver, K. and Brunak, S. (2004). Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng Des Sel* **17**, 527-36.
Ledent, V. and Vervoort, M. (2001). The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. *Genome Res* **11**, 754-70.

Lee, J. H. and Paull, T. T. (2007). Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. *Oncogene* **26**, 7741-8.
Lee, S. H., Hong, H. S., Liu, Z. X., Kim, R. H., Kang, M. K., Park, N. H. and Shin, K. H. (2012). TNFalpha enhances cancer stem cell-like phenotype via Notch-Hes1 activation in oral squamous cell carcinoma cells. *Biochem Biophys Res Commun* **424**, 58-64.
Leimeister, C., Dale, K., Fischer, A., Klamt, B., Hrabe de Angelis, M., Radtke, F., McGrew, M. J., Pourquie, O. and Gessler, M. (2000). Oscillating expression of c-Hey2 in the presomitic mesoderm suggests that the segmentation clock may use combinatorial signaling through multiple interacting bHLH factors. *Dev Biol* **227**, 91-103.
Leimeister, C., Externbrink, A., Klamt, B. and Gessler, M. (1999). Hey genes: a novel subfamily of hairy- and Enhancer of split related genes specifically expressed during mouse embryogenesis. *Mech Dev* **85**, 173-7.
Lu, S. X., Liu, H., Knowles, S. M., Li, J., Ma, L., Tobin, E. M. and Lin, C. (2011). A role for protein kinase casein kinase2 alpha-subunits in the Arabidopsis circadian clock. *Plant Physiol* **157**, 1537-45.
Maroto, M., Dale, J. K., Dequeant, M. L., Petit, A. C. and Pourquie, O. (2005). Synchronised cycling gene oscillations in presomitic mesoderm cells require cell-cell contact. *Int J Dev Biol* **49**, 309-15.
Masamizu, Y., Ohtsuka, T., Takashima, Y., Nagahara, H., Takenaka, Y., Yoshikawa, K., Okamura, H. and Kageyama, R. (2006). Real-time imaging of the somite segmentation clock: revelation of unstable oscillators in the individual presomitic mesoderm cells. *Proc Natl Acad Sci U S A* **103**, 1313-8.
Massari, M. E. and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol* **20**, 429-40.
McGrew, M. J., Dale, J. K., Fraboulet, S. and Pourquie, O. (1998). The lunatic fringe gene is a target of the molecular clock linked to somite segmentation in avian embryos. *Curr Biol* **8**, 979-82.
Mehra, A., Shi, M., Baker, C. L., Colot, H. V., Loros, J. J. and Dunlap, J. C. (2009). A role for casein kinase 2 in the mechanism underlying circadian temperature compensation. *Cell* **137**, 749-60.
Moll, U., Lau, R., Sypes, M. A., Gupta, M. M. and Anderson, C. W. (1999). DNA-PK, the DNA-activated protein kinase, is differentially expressed in normal and malignant human tissues. *Oncogene* **18**, 3114-26.
Muller, M., v Weizsacker, E. and Campos-Ortega, J. A. (1996). Expression domains of a zebrafish homologue of the Drosophila pair-rule gene hairy correspond to primordia of alternating somites. *Development* **122**, 2071-8.
Newton, A. C. (1995). Protein kinase C: structure, function, and regulation. *J Biol Chem* **270**, 28495-8.
Niederreither, K., Vermot, J., Schuhbaur, B., Chambon, P. and Dolle, P. (2002). Embryonic retinoic acid synthesis is required for forelimb growth and anteroposterior patterning in the mouse. *Development* **129**, 3563-74.
Nishimura, M., Isaka, F., Ishibashi, M., Tomita, K., Tsuda, H., Nakanishi, S. and 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.

Niwa, Y., Masamizu, Y., Liu, T., Nakayama, R., Deng, C. X. and Kageyama, R. (2007). The initiation and propagation of Hes7 oscillation are cooperatively regulated by Fgf and notch signaling in the somite segmentation clock. *Dev Cell* **13**, 298-304.

Norton, J. D. (2000). ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. *J Cell Sci* **113 (Pt 22)**, 3897-905.

Oates, A. C. and Ho, R. K. (2002). Hairy/E(spl)-related (Her) genes are central components of the segmentation oscillator and display redundancy with the Delta/Notch signaling pathway in the formation of anterior segmental boundaries in the zebrafish. *Development* **129**, 2929-46.

Oates, A. C., Morelli, L. G. and Ares, S. (2012). Patterning embryos with oscillations: structure, function and dynamics of the vertebrate segmentation clock. *Development* **139**, 625-39.

Oliver, C. and Jamur, M. C. (2010). Overview of antibodies for immunochemistry. *Methods Mol Biol* **588**, 3-9.

Ozbudak, E. M. and Lewis, J. (2008). Notch signalling synchronizes the zebrafish segmentation clock but is not needed to create somite boundaries. *PLoS Genet* **4**, e15.

Palmeirim, I., Henrique, D., Ish-Horowicz, D. and Pourquie, O. (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* **91**, 639-48.

Pascoal, S., Carvalho, C. R., Rodriguez-Leon, J., Delfini, M. C., Duprez, D., Thorsteinsdottir, S. and Palmeirim, I. (2007). A molecular clock operates during chick autopod proximal-distal outgrowth. *J Mol Biol* **368**, 303-9.

Pichon, B., Taelman, V., Kricha, S., Christophe, D. and Bellefroid, E. J. (2002). XHRT-1, a hairy and Enhancer of split related gene with expression in floor plate and hypochord during early Xenopus embryogenesis. *Dev Genes Evol* **212**, 491-5.

Pinna, L. A. and Meggio, F. (1997). Protein kinase CK2 ("casein kinase-2") and its implication in cell division and proliferation. *Prog Cell Cycle Res* **3**, 77-97.

Reifers, F., Bohli, H., Walsh, E. C., Crossley, P. H., Stainier, D. Y. and Brand, M. (1998). Fgf8 is mutated in zebrafish acerebellar (ace) mutants and is required for maintenance of midbrain-hindbrain boundary development and somitogenesis. *Development* **125**, 2381-95.

Riedel-Kruse, I. H., Muller, C. and Oates, A. C. (2007). Synchrony dynamics during initiation, failure, and rescue of the segmentation clock. *Science* **317**, 1911-5.

Roux, P. P., Richards, S. A. and Blenis, J. (2003). Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity. *Mol Cell Biol* **23**, 4796-804.

Sakagami, T., Sakurada, K., Sakai, Y., Watanabe, T., Nakanishi, S. and Kageyama, R. (1994). Structure and chromosomal locus of the mouse gene encoding a cerebellar Purkinje cell-specific helix-loop-helix factor Hes-3. *Biochem Biophys Res Commun* **203**, 594-601.

Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R. and Nakanishi, S. (1992). Two mammalian helix-loop-helix factors structurally related to Drosophila hairy and Enhancer of split. *Genes Dev* **6**, 2620-34.
Seino, S. and Shibasaki, T. (2005). PKA-dependent and PKA-independent pathways for cAMP-regulated exocytosis. *Physiol Rev* **85**, 1303-42.
Shankaran, S. S., Sieger, D., Schroter, C., Czepe, C., Pauly, M. C., Laplante, M. A., Becker, T. S., Oates, A. C. and Gajewski, M. (2007). Completing the set of h/E(spl) cyclic genes in zebrafish: her12 and her15 reveal novel modes of expression and contribute to the segmentation clock. *Dev Biol* **304**, 615-32.
Sheeba, C. J., Palmeirim, I. and Andrade, R. P. (2007). Chick Hairy1 protein interacts with Sap18, a component of the Sin3/HDAC transcriptional repressor complex. *BMC Dev Biol* **7**, 83.
Shimojo, H., Ohtsuka, T. and Kageyama, R. (2008). Oscillations in Notch Signaling Regulate Maintenance of Neural Progenitors. *Neuron* **58**, 52-64.
Shinga, J., Itoh, M., Shiokawa, K., Taira, S. and Taira, M. (2001). Early patterning of the prospective midbrain-hindbrain boundary by the HES-related gene XHR1 in Xenopus embryos. *Mech Dev* **109**, 225-39.
Sieger, D., Tautz, D. and Gajewski, M. (2004). her11 is involved in the somitogenesis clock in zebrafish. *Dev Genes Evol* **214**, 393-406.
Sieger, U., Reinhardt, C. and Baumann, R. (1993). Control of cell pH in immature primitive red cells from chick embryo. *Comp Biochem Physiol Comp Physiol* **104**, 765-70.
Sigrist, C. J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A. and Bucher, P. (2002). PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform* **3**, 265-74.
Steidl, C., Leimeister, C., Klamt, B., Maier, M., Nanda, I., Dixon, M., Clarke, R., Schmid, M. and Gessler, M. (2000). Characterization of the human and mouse HEY1, HEY2, and HEYL genes: cloning, mapping, and mutation screening of a new bHLH gene family. *Genomics* **66**, 195-203.
Stickney, H. L., Barresi, M. J. and Devoto, S. H. (2000). Somite development in zebrafish. *Dev Dyn* **219**, 287-303.
Stork, P. J. and Schmitt, J. M. (2002). Crosstalk between cAMP and MAP kinase signaling in the regulation of cell proliferation. *Trends Cell Biol* **12**, 258-66.
Suriben, R., Fisher, D. A. and Cheyette, B. N. (2006). Dact1 presomitic mesoderm expression oscillates in phase with Axin2 in the somitogenesis clock of mice. *Dev Dyn* **235**, 3177-83.
Swearingen, M. L., Sun, D., Bourner, M. and Weinstein, E. J. (2003). Detection of differentially expressed HES-6 gene in metastatic colon carcinoma by combination of suppression subtractive hybridization and cDNA library array. *Cancer Lett* **198**, 229-39.
Taelman, V., Van Wayenbergh, R., Solter, M., Pichon, B., Pieler, T., Christophe, D. and Bellefroid, E. J. (2004). Sequences downstream of the bHLH domain of the Xenopus hairy-related transcription factor-1 act as an extended dimerization domain that contributes to the selection of the partners. *Dev Biol* **276**, 47-63.

Takada, H., Hattori, D., Kitayama, A., Ueno, N. and Taira, M. (2005). Identification of target genes for the Xenopus Hes-related protein XHR1, a prepattern factor specifying the midbrain-hindbrain boundary. *Dev Biol* **283**, 253-67.

Takebayashi, K., Sasai, Y., Sakai, Y., Watanabe, T., Nakanishi, S. and Kageyama, R. (1994). Structure, chromosomal locus, and promoter analysis of the gene encoding the mouse helix-loop-helix factor HES-1. Negative autoregulation through the multiple N box elements. *J Biol Chem* **269**, 5150-6.

Takke, C., Dornseifer, P., v Weizsacker, E. and Campos-Ortega, J. A. (1999). her4, a zebrafish homologue of the Drosophila neurogenic gene E(spl), is a target of NOTCH signalling. *Development* **126**, 1811-21.

Tam, P. P. (1981). The control of somitogenesis in mouse embryos. *J Embryol Exp Morphol* **65 Suppl**, 103-28.

Toledo-Ortiz, G., Huq, E. and Quail, P. H. (2003). The Arabidopsis basic/helix-loop-helix transcription factor family. *Plant Cell* **15**, 1749-70.

Tuan, R. S. and Lo, C. W. (2000). Developmental biology protocols. Overview I. *Methods Mol Biol* **135**, 3-5.

Turner, D. L. and Weintraub, H. (1994). Expression of achaete-scute homolog 3 in Xenopus embryos converts ectodermal cells to a neural fate. *Genes Dev* **8**, 1434-47.

Turnpenny, P. D., Alman, B., Cornier, A. S., Giampietro, P. F., Offiah, A., Tassy, O., Pourquie, O., Kusumi, K. and Dunwoodie, S. (2007). Abnormal vertebral segmentation and the notch signaling pathway in man. *Dev Dyn* **236**, 1456-74.

Vasiliauskas, D., Laufer, E. and Stern, C. D. (2003). A role for hairy1 in regulating chick limb bud growth. *Dev Biol* **262**, 94-106.

Wang, C. X., Song, J. H., Song, D. K., Yong, V. W., Shuaib, A. and Hao, C. (2006). Cyclin-dependent kinase-5 prevents neuronal apoptosis through ERK-mediated upregulation of Bcl-2. *Cell Death Differ* **13**, 1203-12.

Weinstein, B. M., Stemple, D. L., Driever, W. and Fishman, M. C. (1995). Gridlock, a localized heritable vascular patterning defect in the zebrafish. *Nat Med* **1**, 1143-7.

Wright, D., Ferjentsik, Z., Chong, S. W., Qiu, X., Jiang, Y. J., Malapert, P., Pourquie, O., Van Hateren, N., Wilson, S. A., Franco, C. et al. (2009). Cyclic Nrarp mRNA expression is regulated by the somitic oscillator but Nrarp protein levels do not oscillate. *Dev Dyn* **238**, 3043-55.

Yoda, H., Momoi, A., Esguerra, C. V., Meyer, D., Driever, W., Kondoh, H. and Furutani-Seiki, M. (2003). An expression pattern screen for genes involved in the induction of the posterior nervous system of zebrafish. *Differentiation* **71**, 152-62.

APPENDICES

APPENDIX I

Solutions recipes

Lysis Buffer (1 mL)	200 µL 100mM Hepes Buffer 250 µL 200 mMGlicerophosphate 200 µL 10mM EGTA 100 µL Glicerol 10 µL 0.01% Triton-X100 10 µL 100 mM Sodium Vanadate 10 µL Protease Inibitor Cocktail 220 µL H2O
Loading Buffer 10x	Tris-CL (0.5 M, pH 6.8) SDS 20% (w/v) Bromophenol Blue 0.2% β-Mercaptoethanol 7% Glicerol 40%
4% acrylamide stacking gel (1 mL)	680 µL H20 170 µL 30% Acrylamide mix 130 µL Tris-Cl (1.0 M, pH 6.8) 10 µL SDS (10% w/v) 10 µL Ammonium Persulfate 1 µL TEMED
10% acrylamide gel (5 mL)	1.9 mL H20 1.7 mL 30% Acrylamide mix 1.3 mL Tris-Cl (1.5 M, pH 8.8) 50 µL SDS (10% w/v) 50 µL Ammonium Persulfate 2 µL TEMED
Running Buffer (1L)	14,4g Glicine 3,0g Tris-Base 10 mL SDS (10% w/v) Dilute to 1L of distilled water
Transfer Buffer (1L)	14,4g Glicine 3,0g Tris-Base Add 200 mL of Methanol Dillute to 1 L of distilled water
TBS	50 mM Tris.HCl, pH 7.4 150 mM NaCl.
Glycerol/n-propyl gallate mounting medium	9 mL Glycerol 1mL Tris-Cl (pH 8.0) 0.05 g *n*-Propyl gallate

APPENDIX II

Antibodies list:

Primary antibodies

Antibody	Host	Clone	Concentration	Expected protein size	Reference
α-cHairy1	Mouse	Monoclonal	2,15 mg/mL	≃ 31 kDa	——
α-β-tubulin1	Rabbit	Polyclonal	1 mg/mL	50 kDa	Abcam ab6046

Secondary antibodies

Reactivity	Host	Label	Concentration	Reference
Mouse	Sheep	Horseradish Peroxidase	2 mg/ml	Abcam ab6808
Rabbit	Goat	Horseradish Peroxidase	2 mg/ml	Abcam ab6721
Mouse	Goat	Alexa Fluor 568	2 mg/ml	Molecular Probes A-11019

APPENDIX III

Protein sequences

Protein	RefSeq	Sequence
c-Hairy1A	AAP44728.1	MPADTGMEKPTASPIAGAPASASHTPDKPRSASEHRKSSKPIM EKRRRARINESLGQLKMLILDALKKDSSRHSKLEKADILEMTV KHLRNLQRAQMAAALSADPSVLGKYRAGFNECMNEVTRFLSTC EGVNADVRARLLGHLSACLGQIVAMNYLPPPPAGQPAHLAQPL HVQLPPTTTGAVPVPCKLEPTEALSPKVYGGFQLVPATDGQFA FLIPNPAFPPGSGPVIPLYANANVPVSTSGGSGNASTTPSASP VQGLTSFGHSVVPASQAGSPIAERRESVWRPW
c-Hairy1B	NP_989803.1	MPADTGMEKPTASPIAGAPASASHTPDKPRSASEHRKVNGSGW RRARGRAESSKPIMEKRRRARINESLGQLKMLILDALKKDSSR HSKLEKADILEMTVKHLRNLQRAQMAAALSADPSVLGKYRAGF NECMNEVTRFLSTCEGVNADVRARLLGHLSACLGQIVAMNYLP

		PPPAGQPAHLAQPLHVQLPPTTTGAVPVPCKLEPTEALSPKVY GGFQLVPATDGQFAFLIPNPAFPPGSGPVIPLYANANVPVSTS GGSGNASTTPSASPVQGLTSFGHSVVPASQAGSPIAERRESVW RPW
X-Hes4A	NP_001082574.1	MPADTMEKPTASPIAGAPASSAQTPDKPKSASEHRKSSKPIME KRRRARINESLGQLKTLILDALKKDSSRHSKLEKADILEMTVK HLRNLQRVQMTAALTSDPSVLGKYRAGFNECTNEVTRFLSTCE GVNTEVRTRLLGHLSSCLGQIVAMNYQQPPSSQQPLHVQLPSS TPAPMPISCKVNPAEAISPKVFQGGFQLVPATDGQFAFLIPNP AYTSSPGPVIPLYANANVTSPGGRQSQSPVQGLTTFGHKMPHM AQAVSPLGGSTGADSAESVWRPW
z-Hes1	NP_571948.1	MPADNMEKQTASPIAGAPASGSHTPDKPKNASEHRKSSKPIME KRRRARINESLGQLKTLILDALKKDSSRHSKLEKADILEMTVK HLRNLQRVQMSAALSADTNVLSKYRAGFNECMNEVTRFLSTCE GVNTEVRSRLLNHLSGCMGQMMAMNYPQPAPAQQAHLAQPLHV QLPSTLPINGASMGSKLSPSEAVSPKVFGGFQLVPATDGQFAF LIPNPAFASATTPVIPLYANASVPVTVNASPVQASSAPTVASP VQGMTSFSGVPQAVSPVGVSAGAESNEPVWRPW
m-Hes1	NP_032261.1	MPADIMEKNSSSPVAATPASVNTTPDKPKTASEHRKSSKPIME KRRRARINESLSQLKTLILDALKKDSSRHSKLEKADILEMTVK HLRNLQRAQMTAALSTDPSVLGKYRAGFSECMNEVTRFLSTCE GVNTEVRTRLLGHLANCMTQINAMTYPGQAHPALQAPPPPPPS GPAGPQHAPFAPPPPPLVPIPGGAAPPPGSAPCKLGSQAGEAA KVFGGFQVVPAPDGQFAFLIPNGAFAHSGPVIPVYTSNSGTSV GPNAVSPSSGSSLTSDSMWRPWRN
h-Hes1	NP_005515.1	MPADIMEKNSSSPVAATPASVNTTPDKPKTASEHRKSSKPIME KRRRARINESLSQLKTLILDALKKDSSRHSKLEKADILEMTVK HLRNLQRAQMTAALSTDPSVLGKYRAGFSECMNEVTRFLSTCE GVNTEVRTRLLGHLANCMTQINAMTYPGQPHPALQAPPPPPPG PGGPQHAPFAPPPPLVPIPGGAAPPPGGAPCKLGSQAGEAAKV FGGFQVVPAPDGQFAFLIPNGAFAHSGPVIPVYTSNSGTSVGP NAVSPSSGPSLTADSMWRPWRN

C-hairy1 Isoform 1 predicted phosphorylation sites.

Position	Residue	Sequence	NetPhos Score	DIPHOS Score
5	T	MPAD**T**GMEK	0.917	0.377
11	T	MEKP**T**ASPI	0.024	0.597
13	S	KPTA**S**PIAG	0.951	0.903
21	S	GAPA**S**ASHT	0.593	0.880
23	S	PASA**S**HTPD	0.583	0.926
25	T	SASH**T**PDKP	0.679	0.761
31	S	DKPR**S**ASEH	0.997	0.977
33	S	PRSA**S**EHRK	0.991	0.988
38	S	EHRK**S**SKPI	0.997	0.981
39	S	HRKS**S**KPIM	0.879	0.951
54	S	RINE**S**LGQL	0.186	0.479
70	S	LKKD**S**SRHS	0.987	0.884
71	S	KKDS**S**RHSK	0.727	0.888

74	S	SSRH**S**KLEK	0.997	0.922
85	T	ILEM**T**VKHL	0.738	0.059
102	S	AAAL**S**ADPS	0.405	0.513
106	S	SADP**S**VLGK	0.030	0.404
111	Y	VLGK**Y**RAGF	0.008	0.038
123	T	MNEV**T**RFLS	0.039	0.058
127	S	TRFL**S**TCEG	0.990	0.303
128	T	RFLS**T**CEGV	0.259	0.026
145	S	LGHL**S**ACLG	0.003	0.068
156	Y	VAMN**Y**LPPP	0.020	0.108
179	T	QLPP**T**TTGA	0.043	0.083
180	T	LPPT**T**TGAV	0.489	0.115
181	T	PPTT**T**GAVP	0.718	0.097
193	T	KLEP**T**EALS	0.048	0.108
197	S	TEAL**S**PKVY	0.992	0.334
201	Y	SPKV**Y**GGFQ	0.047	0.289
210	T	LVPA**T**DGQF	0.022	0.088
227	S	FPPG**S**GPVI	0.021	0.527
234	Y	VIPL**Y**ANAN	0.658	0.379
242	S	NVPV**S**TSGG	0.981	0.486
243	T	VPVS**T**SGGS	0.654	0.143
244	S	PVST**S**GGSG	0.120	0.447
247	S	TSGG**S**GNAS	0.600	0.742
251	S	SGNA**S**TTPS	0.779	0.598
252	T	GNAS**T**TPSA	0.096	0.177
253	T	NAST**T**PSAS	0.821	0.139
255	S	STTP**S**ASPV	0.020	0.620
257	S	TPSA**S**PVQG	0.986	0.567
263	T	VQGL**T**SFGH	0.104	0.124
264	S	QGLT**S**FGHS	0.181	0.428
268	S	SFGH**S**VVPA	0.957	0.375
273	S	VVPA**S**QAGS	0.901	0.546
277	S	SQAG**S**PIAE	0.979	0.577
285	S	ERRE**S**VWRP	0.994	0.961