UNIVERSITY OF NOTTINGHAM

From Sequence to Function: Analysis of Sox3 During Early Zebrafish Development

Elisa Marelli, M.Sc.

Thesis Submitted to the Univesity of Nottingham for the Degree of

Master of Research

September 2013

Abstract

Sox proteins are a family of transcription factors characterised by the presence of a conserved HMG box domain that mediates their binding to DNA. Ten groups of Sox proteins have been identified on the basis of their sequence similarities and named A-J. In particular, the SoxB1 subgroup is composed by highly conserved transcription factors that are involved in the differentiation of the cells towards a neural fate and the specification of neural tissue. During the embryonic development of several vertebrate species the first of the SoxB1 proteins to be expressed is Sox3, which is known to act both as a transcriptional activator and a transcriptional repressor at different stages of development. At the present time, little is known about the regulation of the balance between these two functions. Therefore, this study was aimed to identify the regions of Sox3 that are involved in its functioning as a repressor or as an activator. In order to meet this aim a deletion mutagenesis approach was developed to investigate how the deletion of different regions of Sox3 would have changed the protein's function. A specific cloning strategy was designed in order to obtain twelve Sox3 deletion mutants, each carrying a deletion of about 20 amino acids, so that the regions deleted covered most of the protein sequence. The effects caused by the over-expression of each deletion mutant were then tested on zebrafish embryos and compared to the effects of over-expression of the wild type Sox3. Western blot analysis confirmed that microinjection of all the mutants into 1- 4 cell stage embryos, as well as microinjection of wild type *sox3*, resulted in similar levels of protein expression at sphere stage. The embryos microinjected with *sox3* deletion mutants showed different phenotypes at 24 hours post fertilization (hpf), confirming that they affected the functioning of the protein differentially. In order to investigate deeper these functional changes, microinjected embryos were analysed at earlier stage of development. In zebrafish, Sox3 acts as a repressor of the organizer formation at sphere stage. Analysis of the effects of over-expression of

Sox3 deleted mutants on the expression of the organizer marker genes *bozozok* and *goosecoid,* and the comparison with the effects caused by the over-expression of wild type Sox3, led to the identification of protein regions involved in Sox3 transcriptional repressor function. Analysis of the ability of Sox3 deletion mutants to induce the transcription of a *luciferase* reporter gene, compared to the wild type Sox3, allowed the identification of regions of Sox3 involved with its transcriptional activator function. The data obtained allowed us to draw a presumptive functional map of Sox3. The consistency of this map with evidence found in the literature led to the formulation of different hypotheses that would explain the functions associated with the regions identified. These promising data provide a basis for future studies, which will be aimed to the validation of the hypothesis formulated and to the identification of the amino acid residues that are responsible for the functions mediated by the regions identified.

Acknowledgments

A huge thank you to everyone in lab. D13, for your support every time I needed it and for having welcomed me from day one. Thank you Paul for having been such a good teacher and intellectual guide during all the year, and for having inspired my work with your passion for science. Thank you Michelle for your patience when teaching me and for having become such a good friend. Thank you Safiah for having always been so friendly and kind. And thank you Jaime, Magda and Chris for your help every time I needed. The wonderful environment that I found working in D13 was crucial for my work this year.

Thank you to the assessors who will spend time to read my work, because a careful evaluation will make it more valuable and will allow me to learn more.

This thesis is dedicated to my big brother Martino, who I admire and who taught me to follow my dreams doing what I love. This work would have never been possible without the huge support I received from my mum and dad, who I missed so much this year although I felt their love every single day.

Tiziano, a “thank you” is probably not enough to convey the importance of everything you’ve done to support my work and to help me focusing on my studies. Without you I wouldn’t even have been here.

List of Abbreviations

- Aa: Amino acids
- Amp: Ampicillin
- BMP: Bone Morphogenetic Protein
- Boz: Bozozok
- CAP RNA : Capped RNA
- CNS: Central Nervous System
- cSox3: chicken Sox3
- C-terminus: Carboxyl-terminus
- DEPC water: Diethylpyrocarbonate-treated commercial water
- DMSO: Dimethyl Sulfoxide
- Gcs: Goosecoid
- GFP: often used to refer to a group of embryos that were microinjected with GPF capped RNA
- HA tag : Human Influenza Hemagglutinin tag
- HMG-box: High Mobility Group box
- Hpf: Hours post fertilization
- MBT: Mid-Blastula Translation
- Min : Minute/s
- Mu Agar: Solid culture medium for bacterial growth
- Mu Broth: Liquid culture medium for bacterial growth
- Mut/s : Mutant/s
- Ncad: N-cadherin
- Nt: Nucloetides
- N-terminus: amino-terminus
- Ntl: No Tail
- PAF: Paraformaldehyde
- PBST : PBS added with 0,1% Tween20
- PCR: Polymerase Chain Reaction
- RE: Restriction Enzyme
- SDW: Sterile Distilled Water
- Sec : Second/s
- Sqt; Squint
- SSC : Saline-Sodium Citrate Solution
- UN: uninjected (referred to control embryos that were not microninjected)
- Vol. : Volume
- Wt : Wild type

- wtSox3: group of embryos microinjected with wild-type Sox3 (wt Sox3) capped RNA
- YSL: Yolk Syncitial Layer
- zfSox3: zebrafish Sox3

Table of Contents

1. Introduction

1.1. The Zebrafish (*Danio rerio*)

The Zebrafish (scientific name *Danio rerio,* Fig. 1.1) is a bony fish whose natural habitats are the tropical fresh water rivers of northern India, northern Pakistan, Nepal, and Bhutan. Because of its small size and ease of culture, it has become one of the embryologists' favourite and most used model organisms for the study of vertebrate development.

Figure 1.1 The Zebrafish (Credit: ScienceDaily®).

1.1.1. History and Advantages of the use of Zebrafish as a Model Organism

Zebrafish has been largely used as model organism since George Streisinger, a scientist of the University of Oregon, started using it in the 1970s. The reasons why Streisinger, who was also a fish hobbyist, started working with the zebrafish at the University of Oregon was that it was a simpler model than the mouse and easier to manipulate genetically (Streisinger *et al.* 1981). Moreover, the zebrafish is easy to breed and maintain, and it is small enough to easily house large numbers, but also large enough to allow experimental manipulations. Another useful

characteristic of this fish is the rapid development: in the first 24 hours post fertilization (hpf), all major organs form and within 3 days the fish hatches. After three to four months the fish are sexually mature and ready to generate new offspring. A single zebrafish female can lay up to 200 eggs in a week. A great advantage of using the zebrafish as a model for embryonic development is that zebrafish embryogenesis is very similar to the higher vertebrates, including humans. However, unlike mammals, it develops from a fertilised egg to an adult outside the female in a transparent egg. This means that it is possible to observe the developing embryo in its natural environment. Moreover, the embryos themselves are transparent during the first few days of their lives, so researchers can observe the formation of internal organs live inside the living organism (Fig. 1.2). The use of zebrafish as a model also has advantages from a genetic point of view, as many mutations that disrupt embryonic development have now been identified, many of which may serve as models for human diseases (van Heyningen 1997; Zon 1999; Barut and Zon 2000; Dodd *et al.* 2000; Dooley and Zon 2000; Yan and Gu 2013). These mutants will help us to understand the genetic network controlling the development of vertebrates, including humans.

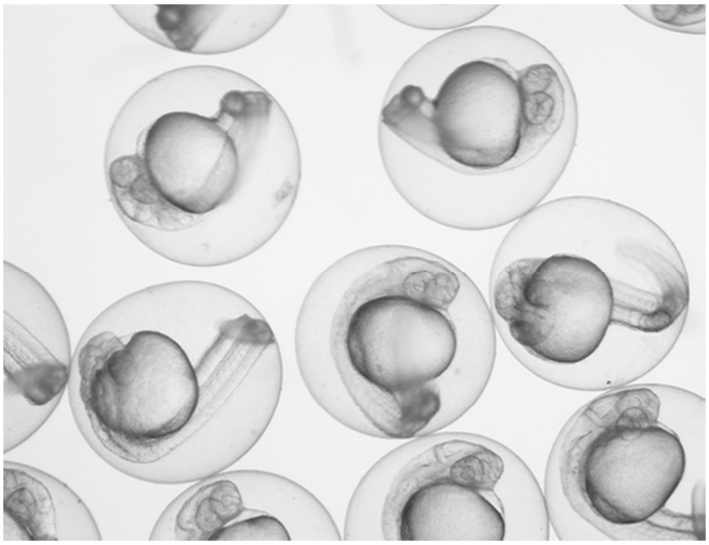


Figure 1.2 Zebrafish embryos 24 hours post fertilization (Credit: Flickr.com, picture by WithoutFins).

1.1.2. Stages of Embryonic Development in Zebrafish

A study published by Kimmel *et al.* in 1995 describes extensive studies on the morphological changes that define the zebrafish body plan during the first three days of embryonic development. Here I describe the developmental changes occurring in first 24 hpf of the zebrafish embryos, adapted from this study (Kimmel *et al.* 1995).

The zebrafish egg consists of a cytoplasm and yolk floating within a prospective chorion. After fertilization, which occurs at the prospective animal pole, the cytoplasm divides from the yolk and becomes the blastodysc, which sits upon the yolk syncytium, and defines the animal-vegetal axis. From this moment every 15 minutes (min) numerous rapid meroblastic cleavage divisions occur synchronously and give rise to a blastoderm, which is formed by cells called blastomeres (Fig. 1.3). The blastomeres perch on top of the yolk as a mound of cells by the 128 cell stage, which is the blastula stage. This continues until approximately 4 hpf. During the blastula stage several important processes take place, such as the mid-blastula transition (MBT), which occurs at the 512-cell stage and defines the beginning of zygotic gene expression (Fig. 1.3). Maternal factors drive early developmental processes before MBT, although it is possible that a few zygotic genes are also transcribed prior to MBT. Moreover, cell cycles lengthen and become asynchronous and complex morphogenetic rearrangements begin (Kane and Kimmel 1993; Kimmel *et al.* 1995). At the interface between the yolk and the blastoderm the multinucleate yolk syncytial layer (YSL) takes shape, as the cells at the margins fall into the yolk. At this point a process called epiboly starts: the morphology of the blastoderm changes and it forms a multi-layered cup which, together with the YSL, thins and spreads radially over the surface of the yolk by complex streaming movements. Epiboly, a process that defines the beginning of the late blastula stage, occurs approximately 4 hpf and continues until the late gastrula stage. A largely used method for indicating developmental stages is the measure of the percentage of the enveloped yolk during epiboly. At 50% epiboly (5 hpf) another process, called

gastrulation, begins (Fig. 1.4B): blastoderm cells at the leading edge/margin converge from more lateral/ventral regions, involute/ingress (internalise) and extend (Kimmel *et al.* 1995). Because this process of convergence occurs at a specific single point, this area becomes thicker and breaks radial symmetry (Fig. 1.4C); this area is now the dorsal part of the embryo (Fig. 1.4 D, E). This thickening become clearly visible by 60% epiboly and is called the shield (Fig. 1.6). It corresponds to the position of the dorsal embryonic organizer (the functional equivalent of the *Xenopus* Spemann-Mangold organizer (see following paragraph for further description). During gastrulation involuting cells converge mediolaterally and extend towards the yolk and then upwards towards the animal pole thereby elongating embryonic tissue anterioposteriorially underneath the overlying blastoderm (Kimmel *et al.* 1995), (Fig. 1.5). Through this process two layers are formed within the proper embryo. The involuting cells form the hypoblast/mesendoderm (the prospective mesoderm and endoderm) which lies underneath the epiplast (the prospective ectoderm). Both gastrulation and epiboly finish by the late gastrula stage (10 hpf), when the yolk is completely surrounded by the three germ layers of cells which are now positioned at their final anteroposterior and dorsoventral locations; ectoderm is the outer layer, mesoderm is the middle layer and endoderm is the internal layer (Fig. 1.4 D). By 24 hpf CNS structures are apparent (Kimmel *et al.* 1995).

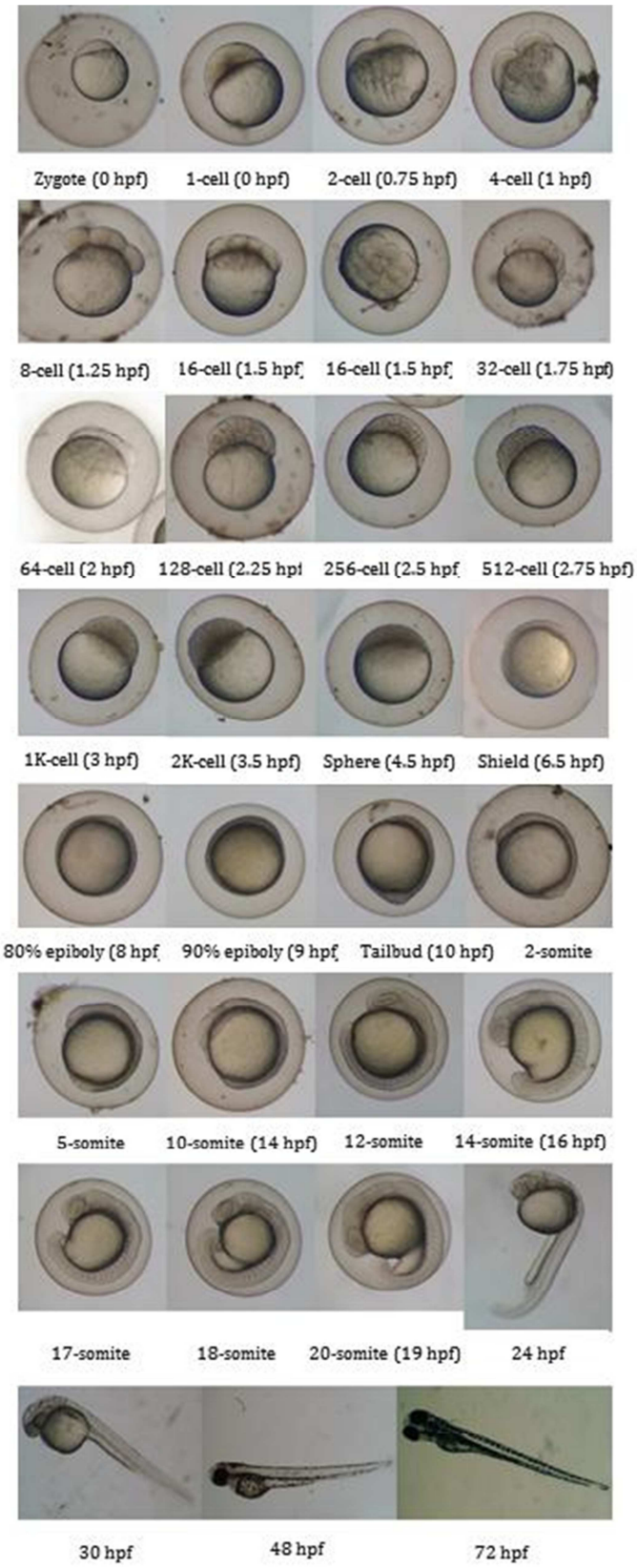


Figure 1.3. Zebrafish embryo developmental stages. The hours post fertilization (hpf) indicated are to be considered approximated as they can vary due to different reasons such as, for example, the temperature of incubation (pictures adapted from Cebra-Thomas, 2004 http://www.swarthmore.edu/NatSci/sgilber1/DB_lab/DB_lab.html).

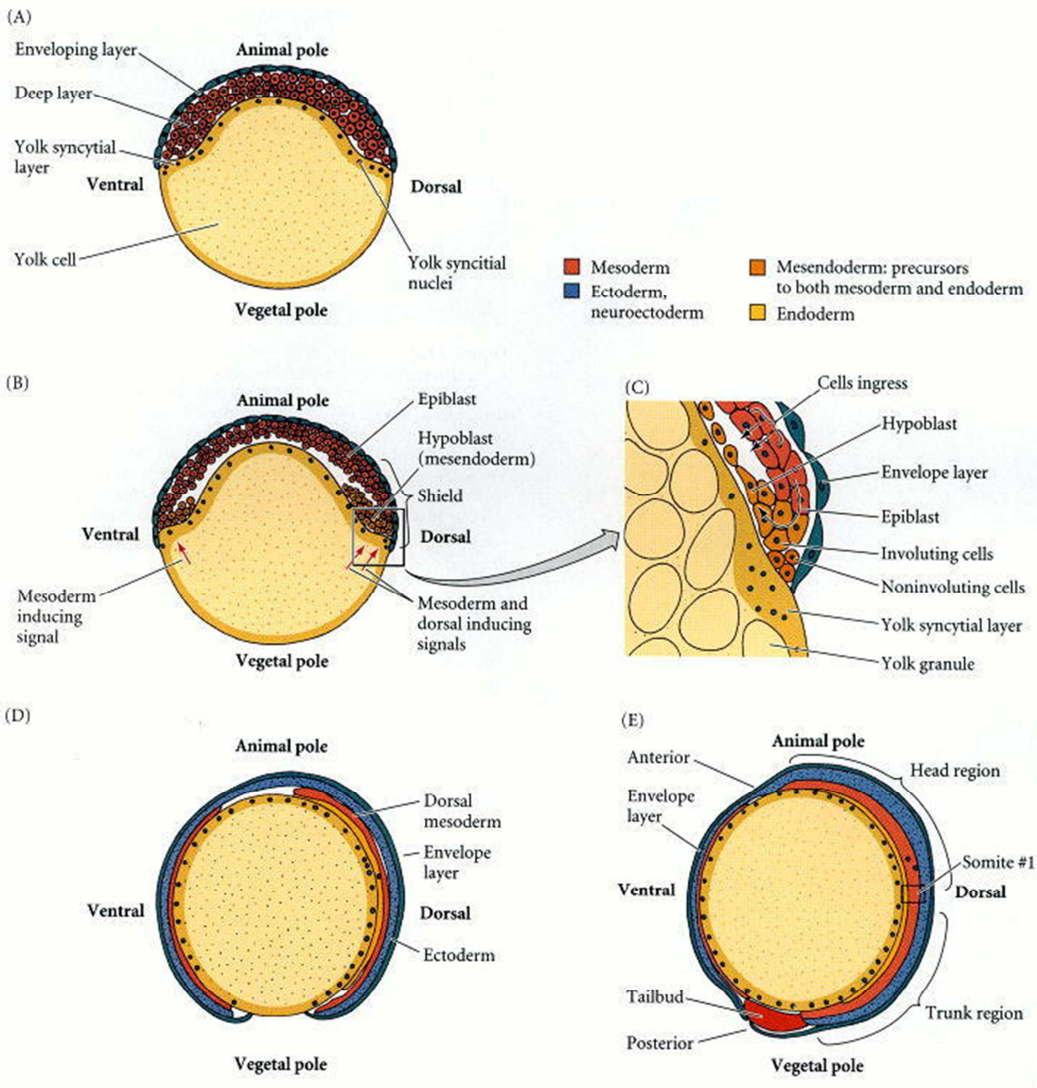


Figure 1.4 Movements of the cells during gastrulation in the zebrafish embryo: blastoderm at 30% epiboly (A); at 50% epiboly the involution of the cells at the margin of the blastoderm create a thickened area were the embryonic shield is loclised (B, C); the embryonic shield corresponds to the region that later becomes dorsal part of the embryo (D, E). (Figure adapted from Gilbert 6th ed.).

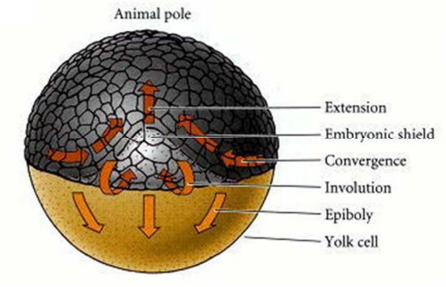


Figure 1.5 Dorsal view of the cell movements that occur during gastrulation in zebrafish embryo. Epiboly moves the blastoderm over the yolk, involution generates the hypoblast. Starting from 50% epiboly convergence and extension movements bring the epiblast and hypoblast (mesendorem) cells towards the dorsal region, forming the organizer. (Figure adapted from Gilbert 6th ed.).

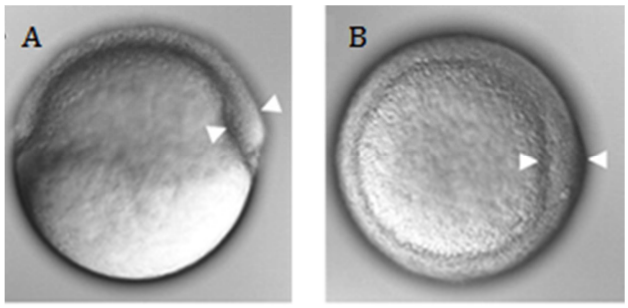


Figure 1.6 At 60% epiboly the thickening corresponding to the embyonic shield becomes visible at the dorsal side of the embryo (white arrows). Lateral view (A) and view from the animal pole (B) of a 60% epiboly embryo (Stemple 2005).

1.1.3. The Role of the Organizer During Development

In zebrafish the formation of the organizer, also known as embryonic shield, is an essential process for the establishment of dorsal-ventral patterning and the induction of neural cell fate. In the following paragraph I briefly describe the molecular mechanisms that regulate this process (Appel 2000; Schier and Talbot 2005).

The firsts understanding of the mechanisms involved in organiser formation came, as in many other biological studies, from the investigation of mutants. Dorsalised (expansion of neural tissue) and ventralised (reduction of neural tissue) mutant phenotypes have been investigated through large scale mutagenic screenings (Driever *et al.* 1996; Hammerschmidt *et al.* 1996; Mullins *et al.* 1996). The mutations that were causing these phenotypes were mapped in genes that are part of the canonical Wnt/β-catenin pathway (Kelly *et al.* 2000) and the Bone Morphogenetic Protein (BMP) pathway (Hammerschmidt *et al.* 1996; Hammerschmidt *et al.* 1999). These pathways are essential for the establishment of the right patterning of the dorsal-ventral axis (Appel 2000; Schier and Talbot 2005).

When the Wnt/β-catenin signalling pathway is activated at the dorsal side of the zygote, β-catenin accumulates in the nucleus, displaces the co-repressor Groucho and partners with members of the TCF/LEF family of transcription factors, thus activating the expression of target transcription factors genes (Fig. 1.7). These genes include *bozozok* (also known as *dharma*) and the nodal-related extracellular signalling molecule *squint*, whose function is crucial for the formation of the presumptive mesoderm. The Wnt/β-catenin signalling pathway overlaps with the Nodal signalling, thus defining a specific region on the dorsal part of the embryo called Nieuwkoop Centre. The Nieukoop Centre, which is located in the area corresponding to the presumptive endoderm (in the vegetal part of the embryo), induces the formation of the embryonic shield. This is a small group of cells located at the dorsal mesoderm edge, and it is functionally equivalent to the Spemann/Mangold organizer found in *Xenopus* (Saúde *et al.* 2000; Niehrs 2004) and to the primitive node found in mouse (Beddington 1994; Shih and Fraser 1996). The zebrafish organizer, as well as the *Xenopus* Spemann/Mangold organizer, can induce the formation of a secondary body axis (including neural tissue and mesoderm) if microsurgically transplanted to the ventral region of another embryo (Shih and Fraser 1996; Driever *et al.* 1997; Spemann and Mangold 2001). A key function of the organizer is to act as a regulatory element determining the dorsal fate of the ectoderm cells by secreting dorsalising signals. These signalling factors are secreted outside of the cells and act in the dorsal region antagonizing BMP such as Bmp2b/7, which are ventralising signalling molecules. Bmp activates the expression of different genes in a concentration-dependent fashion, thus acting as a "morphogen". Some examples of the proteins expressed in the organizer are Chordin, Noggin and Follistatin; they prevent BMP signals from binding DNA and thus induce the cells to become organizer, and then notochord, tissue. Moreover Bozozok, which is induced by the Nieukoop-Centre, represses Bmp2b (Leung *et al.* 2003) as well as the ventralising signals activated by Bmp2b, such as vox/vent/ved (repressors that inhibit Boz and Chd). This results in a de-repression of the organizer genes (such as Chordin)

mediated by Boz. The dorso-ventral regions that give rise to the neural/non-neural ectoderm is defined by the resulting gradient of BMP morphogens. The establishment of the neural/ectodermal fate is described by the "neural default model" (Munoz-Sanjuan and Brivanlou 2002) according to which ectodermal cells are destined to become neural cells unless they are targeted by BMP signals, which is prevented in the dorsal region by mesoderm secreted factors such as Chordin (Fig. 1.8). This means that the organizer induces dorsal fate by secreting dorsalizing factors that block the ventralising BMP proteins, rather than inducing it directly (Linker and Stern 2004; Stern 2005; Stern 2006). In the region where the ventralizing factors are absent, which in other words is the presumptive ectoderm, there is accumulation of transcription factors that induce neural fate (Mizuseki *et al.* 1998; Kudoh *et al.* 2004; Dee *et al.* 2007).

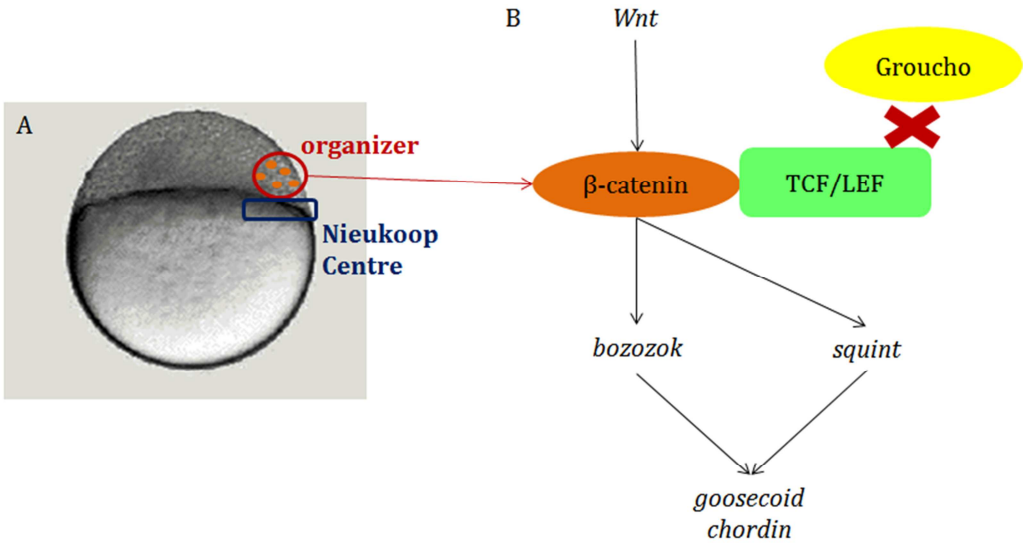


Figure 1.7 Activation of the Wnt/β-catenin pathway in the dorsal side of the zebrafish embryo. Areas of the embryo that correspond to the Nieukoop Centre and to the organizer (A). In the organizer the activation of Wnt signalling and the nuclear localization of β-catenin determine the transcriptional activation of target genes through the displacement of the co-repressor Groucho and the binding to TCF/LEF operated by β-catenin. The activated target genes include *bozozok* and *squint*, which activate *goosecoid* and *chordin* (B).

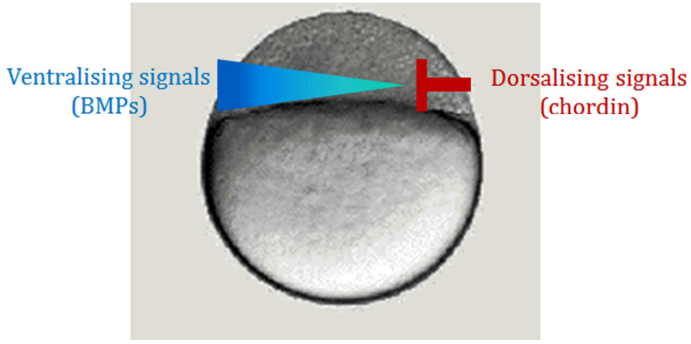


Figure 1.8 The neural default model. According to the "neural default model" ectodermal cells are destined to the neural fate unless they are targeted by ventralising BMP signals, which act in a concentration-dependent fashion. In the dorsal region corresponding to the organizer, dorsalizing signals, such as chordin, block BMP signalling and thus indirectly induce neural fate.

1.2. The Sox Family

1.2.1. The Family History

In the late 1980s many laboratories were investigating the identity of the so called "testis-determining factor" (Tdy or TDF), a gene located on the Y chromosome that was thought to be responsible for the switch in developmental fate occurring in mammals from the default ovarian pathway to that of the testis (Koopman *et al.* 1991; Hacker *et al.* 1995; Collignon *et al.* 1996). The gene was finally identified as the "Sex Determining Region of the Y chromosome" (Sry), which encodes for a transcription factor containing a 79 amino acid DNA-binding domain, the High Mobility Group box (HMG-box), which participates in DNA binding and in some cases also in interactions with partner proteins (Chakravarthy *et al.* 2008). The HMG-box derived its name by the original technique used for sorting proteins closely associated with the DNA: the electrophoresis of denatured proteins through an agarose gel (Lefebvre *et al.* 2007), which showed that the peptides carrying the HMG-box travelled further than other proteins and thus determined the denomination "High Mobility Group". The Sry gene later became the founding member of the family of genes known as the Sox family, initially discovered through homology comparisons between the human and the mouse Sry genes (Gubbay *et al.* 1990; Sinclair *et al.* 1990; Denny *et al.* 1992; Denny *et al.* 1992; Lefebvre *et al.* 2007). On the basis of sequence homology with the Sry HMG-box about thirty other Sry-related genes were later discovered and named "Sry-related HMG-box" (Sox) and numbered chronologically with their discovery, the first to be named were Sox1, Sox2 and Sox3.

1.2.2. The Members of the Family

The nomenclature of the members of the Sox family includes a number indicating the chronological order of discovery and in some cases a prefix indicating the species.

All the genes identified as part of the Sox family have a minimum of 50% of sequence similarity with respect to the HMG-box of Sry. In an extensive evolutionary analysis Bowles *et al.* recognised ten Sox groups (named A-J) based on previous studies as well as full-length protein structure, HMG domain sequence similarity (Fig. 1.9) and structural characteristics such as intron-exon gene organization (Gubbay *et al.* 1990; Wright *et al.* 1993; Pevny and Lovell-Badge 1997; Bowles *et al.* 2000; Schepers *et al.* 2002). Subgroup A contains mammalian proteins, the members of the subgroups B, C, D, E and F are found in a large number of metazoan taxa (Jager *et al.* 2006; Larroux *et al.* 2008) and the subgroups G to J contains members that are specifics for particular lineages (Zhong *et al.* 2011).

Each one of these subgroups is composed of members with a sequence similarity of between 60% and 90% (Bowles *et al.* 2000; Kamachi *et al.* 2000) and while some of them are species-specifics, others can be found in a wide range of organisms (Wegner 1999; Bowles *et al.* 2000; Wegner 2010). About twenty Sox genes were found in the mouse as well as in the human genome (Schepers *et al.* 2002). The majority of vertebrate family subgroups have been found to be represented in invertebrates by a single Sox gene (Wegner 1999; Bowles *et al.* 2000).

Except for the similarity in the HMG-box sequence, the members of distant subgroups have high variability in the rest of their sequences (Bowles *et al.* 2000), while the Sox proteins that belong to the same subgroup often show functional redundancy when they are co-expressed. This can be explained by the presence of conserved structural domains outside the HMG-box (Fig. 1.10), (Bowles *et al.* 2000). Such domains are localised at the C-terminus of the proteins and are implicated in transcriptional regulation.

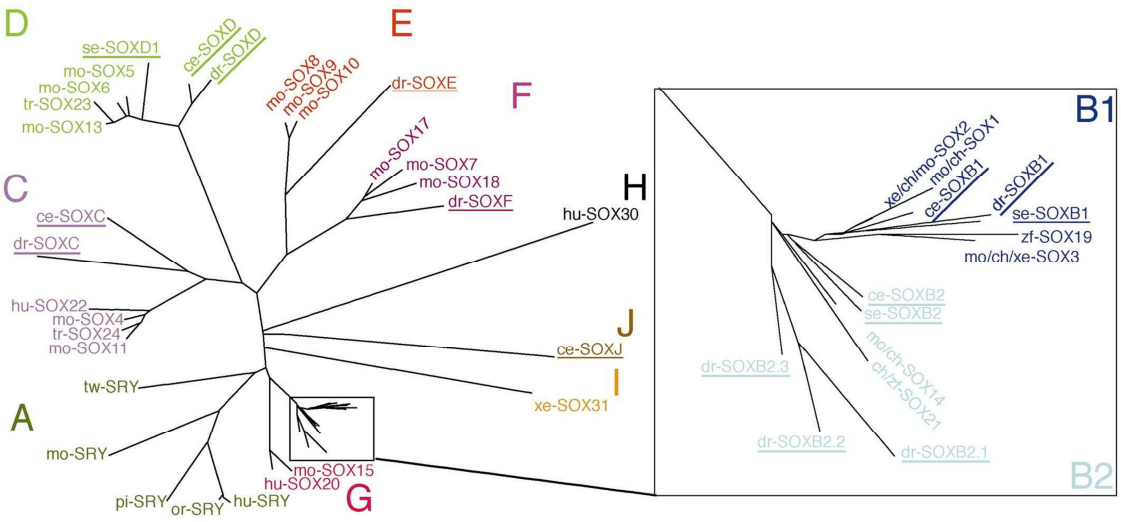


Figure 1.9 Unrooted phylogeny tree of the Sox HMG domain. Branch lengths are representative of the extent of divergence. For groups of presumed mammalian orthologues (other than group A—Sry), only one representative is indicated. Different groups (A-J) are written with different colours. The insert shows group B, which contains the subgroups B1 and B2. Invertebrate sequences are underlined. Abbreviations: al, *Alligator mississippiensis* (alligator); ce, *Caenorhabditis elegans* (nematode); ch, *Gallus gallus* (chicken); dr, *Drosophila melanogaster* (fruit-fly); du or d, *Sminthopsis macroura* (marsupial); fu, *Saccharomyces cerevisiae* (fungi); hu or h, *Homo sapiens* (human); mo or m, *Mus musculus* (mouse); or, *Pongo pygmaeus* (orangutan); pi or p, *Sus scrofa* (pig); ra or r, *Rattus norvegicus* (rat); tw, *Macropus eugenii* (marsupial); sh or s, *Ovis aries* (sheep); tr, *Oncorhynchus mykiss* (rainbow trout); se, *Strongylocentrotus purpuratus* (sea urchin); xe, *Xenopus laevis* (frog); zf, *Danio rerio* (zebrafish). Image from Bowles *et al.* (2000).

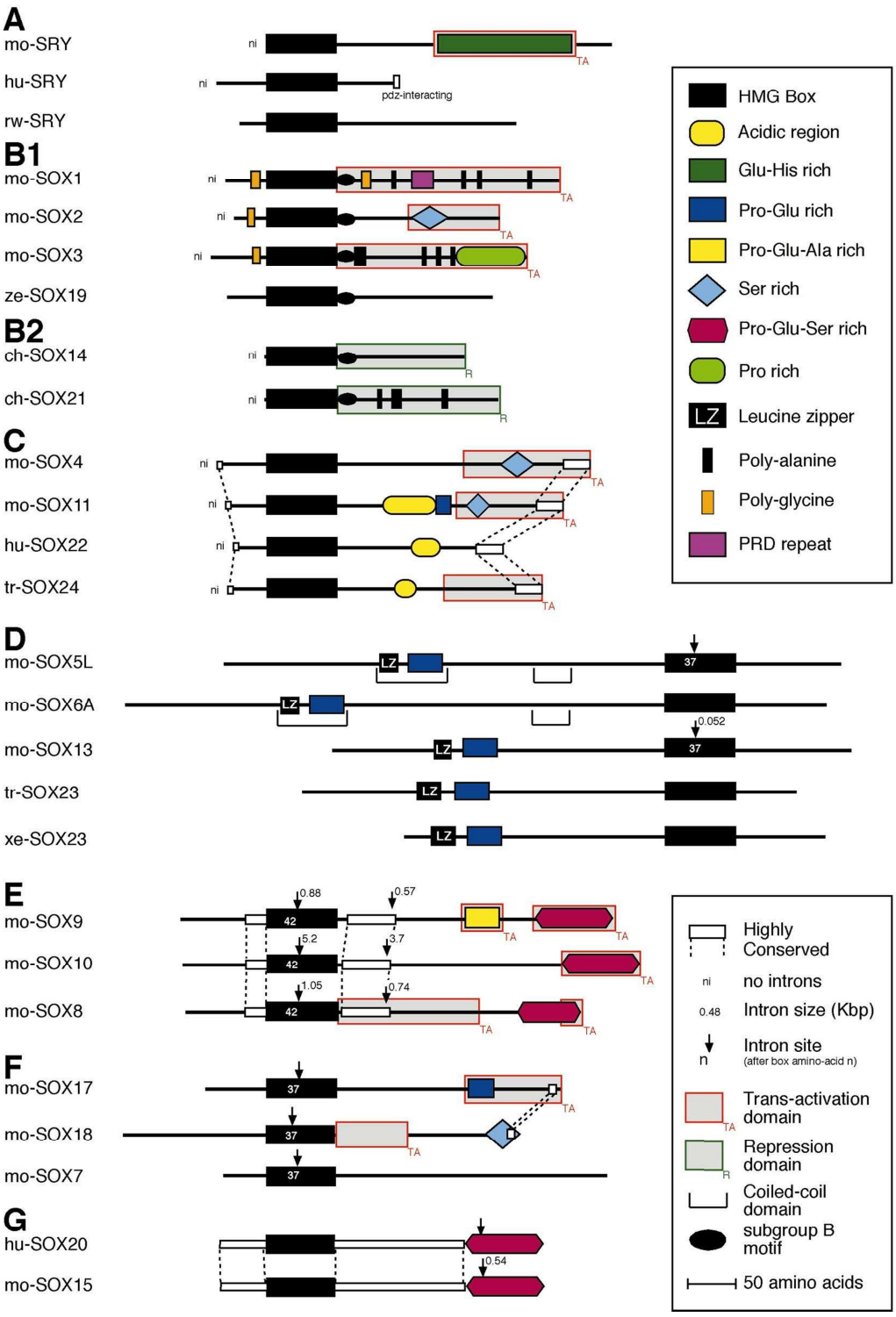


Figure 1.10 Schematic representation of some of the members of different groups (and subgroups) of Sox family proteins that highlights the structural conservation of functional domains and gene organization. Demonstrated and putative structural domains are shown. Picture from Bowles *et al.* (2000).

1.2.3. DNA-binding Property of the Sox Proteins

All the members of the Sox family share high level of similarity and are thought to function as transcription factors (Laudet *et al.* 1993). The Sox proteins contain a single canonical HMG domain composed of 79 amino acids that contains the conserved sequence motif "RPMNAFMVW"; this motif differentiates the Sox proteins from the other members of the HMG superfamily. Due to the presence of the HMG domain, Sox proteins are able to recognise and bind specific sequences of DNA. In particular they bind to the consensus sequence "5'-WWCAAW-3'" (where W=A/T) (Harley *et al.* 1994). However some of these proteins, such as Sox9, showed a higher binding preferences for slightly different consensus sequences (Mertin *et al.* 1999). The secondary structure of the HMG-box is composed by three α-helices and one β-sheet forming a "L" shape which binds to the minor groove of the DNA helix (Fig. 1.11), (van de Wetering and Clevers 1992; Read *et al.* 1994; Weiss 2001; Lefebvre *et al.* 2007). This binding process is a unique feature of the HMG-box as it causes a widening of the minor groove that leads to the bending of the DNA structure (Fig. 1.12), (Ferrari *et al.* 1992; van de Wetering and Clevers 1992; Lefebvre *et al.* 2007). In fact the majority of the DNA-binding proteins bind to the major DNA groove and cause minor changes in its spatial conformation, by contrast the binding of the HMG with the minor groove can bend the DNA with an angle from 30 to 110 degrees depending on experimental conditions (Ferrari *et al.* 1992; Connor *et al.* 1994; Pontiggia *et al.* 1994; Kamachi *et al.* 1999; Kamachi *et al.* 2000; Weiss 2001). Such characteristic can be explained if we consider that the conformation adopted by the DNA might render it more accessible to other proteins. For this reason it has been proposed that Sox factors function as architectural transcription factors that recruit other protein through their binding to the DNA. According to this model the proteins recruited would be other transcription factors, chromatin re-modellers or other regulatory partners that would act co-activating (or co-repressing) target genes (Pontiggia *et al.* 1994).

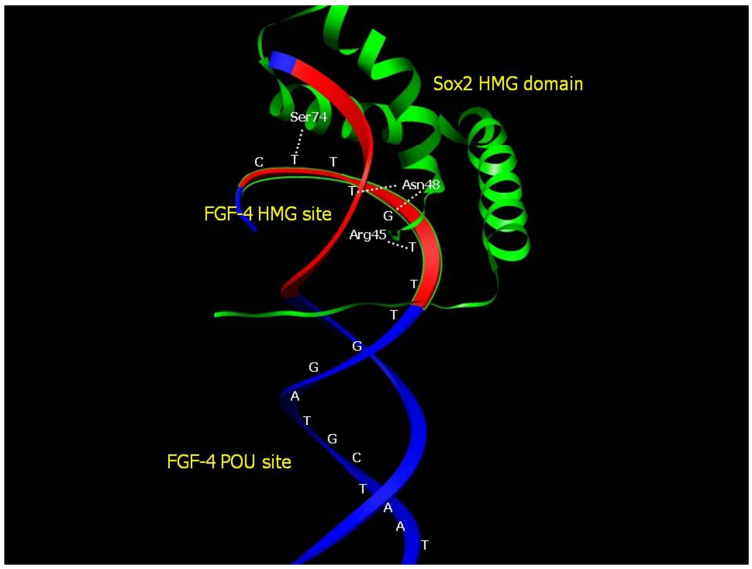


Figure 1.11 Three-dimensional representation of mouse Sox2 HMG domain binding its target sequence on the FGF-4 HMG/POU cassette (indicated in red). The HMG domain is composed by three α-helices and a β-sheet (pictured in green) and binds to the minor groove of the DNA which participate in DNA binding as well as interactions with partner proteins. Hydrogen bonds are indicated as white dotted lines. Picture from Chakravarthy *et al.* (2008).

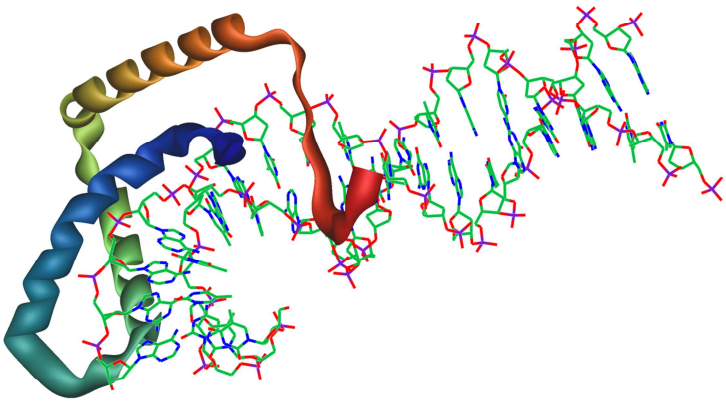


Figure 1.12 Three-dimensional PDB model of mouse Sox2 binding and bending DNA. The presence of three α-helices and a β-sheet give to the domain a characteristic "L" shape. As a result of the binding, the DNA helix result to be dramatically bended of about 90°. Picture from Chakravarthy *et al.* (2008).

1.2.4. Transcriptional Regulation Activity of the Sox Proteins

It has been demonstrated that Sox proteins also function as classical transcription factors regulating the expression levels of target genes (Pevny and Lovell-Badge 1997). They are mostly described as activators of target gene expression. Many of them contain a trans-activation domain at the C-terminus, and this includes the SoxB1 proteins (Kamachi *et al.* 1995; Kamachi 1996; Kamachi *et al.* 1998; Kamachi *et al.* 1999; Chakravarthy *et al.* 2008), SoxC proteins (van de Wetering *et al.* 1993; Chakravarthy *et al.* 2008), SoxE proteins (Bell *et al.* 1997; Ng *et al.* 1997; Kamachi *et al.* 1999; Chakravarthy *et al.* 2008) and SoxF (Chakravarthy *et al.* 2008). If these proteins are deleted at their C-terminus and fused to VP16 (a transactivation domain constitutively active), they function as activators, which is in contrast with the fusion with repressors domains (Koster *et al.* 2000; Bylund *et al.* 2003; Wegner 2010). However some of the Sox factors, for example the SoxB2 group (Uchikawa *et al.* 1999), act as transcriptional repressors.

1.2.5. The Sox Proteins: a Family of Flexible Transcription Factors

All the Sox proteins bind to the DNA by recognizing a specific motif that is present many times throughout the whole genome. This consensus motif is quite short and degenerate (Wegner 2010). Moreover, the specificity with which the HMG-box domains of these different Sox proteins bind and bend the DNA is comparable when they are tested in the same *in vitro* conditions (Kamachi *et al.* 1999; Mertin *et al.* 1999; Kamachi *et al.* 2000). It has been shown that there may be several Sox factors able to bind a certain site *in vitro*, but when tested *in vivo* only one of them is still able to bind that site. For example, although Sox1, Sox2, Sox3 and Sox9 have been shown to have a C-terminal capable of acting as transactivation domain (Ng *et al.* 1997; Kamachi *et al.* 1999; Kamachi *et al.* 2000), they do not bind DNA with sufficient affinity compared to the classical transcription factors

(Lefebvre *et al.* 2007). Therefore, theoretically they would not be able to bind the DNA and act as transcriptional regulators *in vivo* (Kamachi *et al.* 2000). Moreover, given that Sox factors elicit their function in many of different tissues, cell types and developmental stages and that in mammals only twenty of them have been identified in spite of the high number of functions that they have, it is logical to think that they must be strictly and specifically regulated. Given all this, how is it then possible for each of the Sox factor to act in a cell-specific way activating (or in some cases repressing) specific target genes (Kamachi *et al.* 2000)?

It is logical to think that Sox factors have to be necessarily flexible in order to function in many diverse contexts as they were passepartout regulators (Wegner 2010). This can be explained by the interaction of these proteins with co-factors, whose presence would be regulated in a developmental stage- and cell-specific fashion. According to this model the transcriptional regulation activity of the Sox proteins could function only when the proteins and the specific co-factors are co-expressed. Each specific partnering between a given Sox protein and a given co-factor would therefore regulate specific target genes and regulate different processes, as for example cell differentiation, in a cellular, tissue and temporal-specific fashion (Fig. 1.13). Moreover, this could explain why the Sox proteins that *in vitro* have been shown to bind the DNA loosely, on the other hand act as effective transcription factors *in vivo*. Therefore, the Sox proteins contain regions that specifically recruit different co-factors depending on the context (Kamachi *et al.* 2000; Wilson and Koopman 2002). Because of the high homology between the HMG-box domains, it is thought that such motifs would be generally present in the rest of the sequence, which is highly variable. Sox factors that belong to different subgroups and act as trans-activators have low homology in their primary structure, and this could be explained by the necessity of recruiting different partners (Wegner 2010). For example, the C-terminal region of Sox2 recruits p300 and partners with OCT3/OCT4 to activate the enhancer fgf4 (Bernadt *et al.* 2004; Wegner 2010). Some contexts in which Sox

protein action is dependent on the cell-specific presence of co-factors have already been identified. For example, Sox2 (a member of the SoxB1 subgroup) is implicated in lens development, but it can recognise the DC5 target enhancer and regulate this process only when interacting with Pax6 (Kamachi *et al.* 2001; Inoue *et al.* 2007). Through Chromatin Immunoprecipitation (ChIP) techniques it has been demonstrated that Sox2 and Pax6 are not able to bind DC5 enhancer stably when alone, but they do so only when co-expressed, as during normal lens development process (Kamachi *et al.* 1995; Kamachi 1996; Kamachi *et al.* 1998).

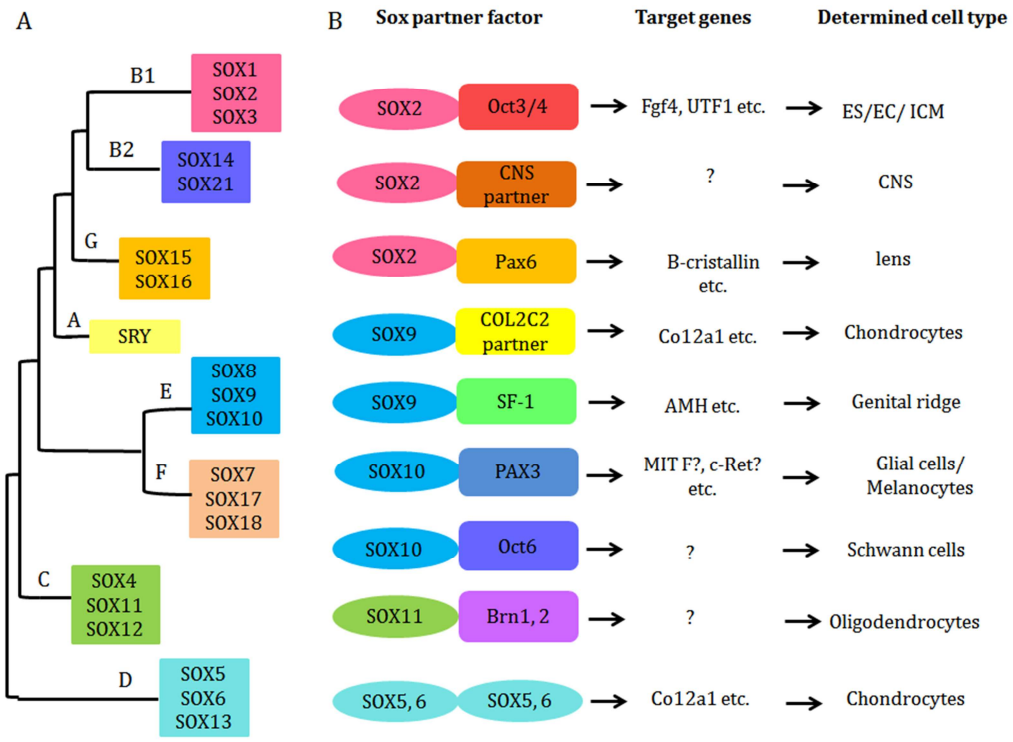


Figure 1.13 (A) Different members of some of the groups (and subgroups) of the Sox family of transcription factors. The name of the groups is indicated in black, the group B contains the two subgroup B1 and B2. (B) Some examples of how Sox proteins partner with specific co-factors in order to determine cell differentiation. The figure represents the interactions between some of the Sox proteins and various co- factors, the target genes selectively activated as a result of the interactions, and the results in terms of cell differentiation. Picture drawn from image on Graduate School of Osaka University website (http://www.fbs.osaka-u.ac.jp/eng/labo/06/).

1.3. The SoxB Group

1.3.1. Classification of the Members of SoxB Group

The SoxB subgroup of transcription factors, which include Sox1, Sox2, Sox3, Sox14 and Sox21, plays a central role in several processes during the embryonic development of vertebrates and insects, such as neurogenesis, gonadogenesis and morphogenesis (Nambu *et al.* 1996; Soriano and Russell 1998; Uchikawa *et al.* 1999; Lefebvre *et al.* 2007; Phochanukul and Russell 2010).

Sox1, *sox2* and *sox3* belong to this group and they were the first *sox* genes to be isolated and characterised as Sry-related (Gubbay *et al.* 1990). In addition to these three genes two others were later classified in the same subgroup because of the high sequence homology of their HMG-box domains: *sox14* (Wright *et al.* 1993) and *sox21* (Rex *et al.* 1997). However, a further subdivision into subgroups SoxB1 and SoxB2 has been proposed on the basis of the full-length sequence alignment and the different roles of these proteins in chick (Uchikawa *et al.* 1999) as well as in other vertebrates (Bowles *et al.* 2000). The homology of the full-length sequences is high in proteins belonging to the same subgroup, but there is no similarity between the sequences of proteins of different subgroups, except for the HMG-box domains and a short proximal C-terminal region. Therefore, the similarity between the sequences outside the HMG-box of members of the different SoxB subgroup (SoxB1 and SoxB2) is low as it is for the members of different Sox groups; for this reason the phylogenetic analysis of the SoxB1 and SoxB2 members have been largely based on the HMG-box sequences (Zhong *et al.* 2011).

In zebrafish the SoxB1 subgroup contains other two members: Sox 19a (Sox19) (Vriz and Lovell-Badge 1995) and Sox19b (Sox31) (Girard *et al.* 2001), which have been identified as orthologues of Sox15 (which belong to the mammalian SoxG group) and are thought to derive from a further genome duplication event (Okuda *et al.* 2006; Okuda *et al.* 2010).

1.3.2. Roles of the Members of SoxB Group

Regarding the functions of SoxB proteins, SoxB1 transcription factors are generally considered to act as activators, while SoxB2 are thought to act as repressors.

SoxB1 proteins contain a C-terminus that act as transcriptional activator. However, it has been shown that Sox14 and Sox21 have a putative repressor C-terminus domain. This evidence is in contrast with the findings on other Sox proteins (Kamachi *et al.* 1995; Kamachi *et al.* 1998; Kamachi *et al.* 1999). In fact, it has been shown that they repress the expression of the δ-crystalline enhancer DC5, while other SoxB1 members activate it in the same experimental conditions (Uchikawa *et al.* 1999).

However, it now seems that SoxB1 and SoxB2 can be either activators or repressors and that their function is highly dependent on the context of when and where they are expressed. This context-dependency could be correlated with the interacting partners of the SoxB factors or with post-transcriptional modifications as, for example, SUOMylation. This is the covalent attachment of a short peptide (the SUMO) to a consensus region of the protein and it is thought to be correlated to this switch between transcriptional activation and repression functions of the SoxB factors (Savare *et al.* 2005; Savare and Girard 2005; Taylor and Labonne 2005; Haldin and LaBonne 2010)

1.4. The SoxB1 Subgroup

Compared with all the other Sox factors, the members of the SoxB1 subgroup show higher sequence conservation (Bowles *et al.* 2000). Moreover, their functions are conserved during evolution as they are all involved in the process of differentiation of the cells towards the neural fate and the specification of the neural tissue. The expression pattern of these proteins during the early stages of embryonic development is also conserved.

Some SoxB1 factors are maternally-derived and they regulate axis formation by modulating the expression of nodal-like proteins. This has been shown to be the case for *Xenopus* Sox3 (Zhang *et al.* 2003; Zhang *et al.* 2004; Zhang and Klymkowsky 2007), of mouse Sox2 (Avilion *et al.* 2003) and Sox3-Sox19b in zebrafish (Okuda *et al.* 2006).

Sox1, Sox2 and Sox3 proteins were found to be present in the dividing neuroepithelium of metazoans embryo during CNS development. It had been demonstrated that their co-expression is crucial for maintaining the proliferative potential of these multipotent progenitor cells by blocking their differentiation. Accordingly, when these neuroepithelial progenitor cells start to differentiate and the post-mitotic neural genes start to be transcribed, the expression of SoxB1 factors begin to be down-regulated (Collignon *et al.* 1996; Pevny and Lovell-Badge 1997; Pevny *et al.* 1998; Uchikawa *et al.* 1999; Wood and Episkopou 1999; Bylund *et al.* 2003; Graham *et al.* 2003; Kan *et al.* 2004; Pevny and Placzek 2005). Moreover, the constitutive overexpression of these proteins in the neural tube of the chicken (Bylund *et al.* 2003; Graham *et al.* 2003) or in the neural plate of *Xenopus* (Rogers *et al.* 2009) inhibits the neural differentiation process and leads to an increase in the number of progenitor cells that causes the expansion of the population of proliferative stem-like cells. When constitutive activator versions of Sox2/3 are over-expressed, they suppress the process of neural differentiation with the same extent of the wt proteins. On the other hand, overexpression of dominant repressor forms of Sox3 has been shown to cause premature exit from the cell cycle and neural differentiation of the cells. These data suggests that the SoxB1 factors inhibit neural differentiation by acting as transcriptional activators instead of repressors (Bylund *et al.* 2003).

As mentioned earlier, SoxB1 proteins can be maternally-inherited by the zygote and their pattern of expression during early stages of embryogenesis is evolutionarily highly conserved. For example Sox3 is maternally-inherited in the *Xenopus* embryo and its regulation of nodal-like protein expression is essential for the axis formation (Zhang *et al.*

2003; Zhang *et al.* 2004; Zhang and Klymkowsky 2007), while in the mouse embryo Sox2 is maternally-derived in a similar way (Avilion *et al.* 2003). On the other hand, some small differences in the spatial and temporal patterns of expression of the SoxB1 factors have been identified. In fact Sox3 is the first of the SoxB1 proteins to be expressed in vertebrates (except for mammals), and while Sox2 starts to be expressed in the neurectoderm at the beginning of gastrulation, Sox1 is expressed only after gastrulation (Rex *et al.* 1997; Mizuseki *et al.* 1998; Wood and Episkopou 1999; Nitta *et al.* 2006).

SoxB1 factors are thought to be functionally redundant during the formation of the CNS because of the high similarity of their primary structure and the overlapping patterns of their expression (Collignon *et al.* 1996; Pevny and Rao 2003). Miyagi, Masui *et al.* showed that the conditional knockout of Sox2 is compensated by the up-regulation of Sox3 expression in mouse (Miyagi *et al.* 2008). Moreover loss-of-function experiments demonstrated that that the knockdown of Sox3 in zebrafish (Dee *et al.* 2008) and in *Xenopus* (Rogers *et al.* 2009) causes only mild defects in the development of the nervous system. The phenotypes that result from these experiments affect only the cells (or the tissues) where that specific SoxB1 factor is usually expressed alone. For example in mouse, mutated Sox1 and Sox3 cause lens fibre and pituitary or cranio-facial defects (Nishiguchi *et al.* 1998; Rizzoti *et al.* 2004; Rizzoti and Lovell-Badge 2007). In humans, different retinal abnormalities, microphthalmia and anophthalmia are due to mutations in the Sox2 gene (Fantes *et al.* 2003; Williamson *et al.* 2006).

1.5. Sox 3

1.5.1. The Importance of Studying Sox3

At the present little is known about the roles of Sox3 in the earliest stages of the embryonic development. Knowing the structure of the protein and how it interacts with its co-factors would help to give a clearer picture. However, its secondary structure is still unknown (except for the HMG domain) and the interacting co-factors are still being identified. Sox3 amino acid sequence, though, does show high similarity with the other SoxB1 proteins, Sox1 and Sox2. As shown in figure 1.14, the physical-chemical properties of the amino acid sequences are conserved among the three SoxB1 factors in zebrafish. This is consistent with the findings that SoxB1 factors have overlapping expression patterns and that these highly homologous proteins have redundant roles in nervous system development (Rogers *et al.* 2009; Zhong *et al.* 2011). Therefore the understanding of the roles and mechanisms of function of Sox3 is very important. In fact, knowing the molecular mechanisms of Sox3 could on one hand clarify the earliest events that lead to the formation of the neural tissue and, on the other help to understand the mechanisms of action of the other SoxB1 factors in the later stages of development. Since Sox3 is the first of the SoxB1 proteins to be expressed in several vertebrate species, and since its presence was detected at very early stages even before the specification of neural and non-neural domains (Rex *et al.* 1997; Dee *et al.* 2007), it is thought to be a key factor for the subsequent development of the neural tissue.

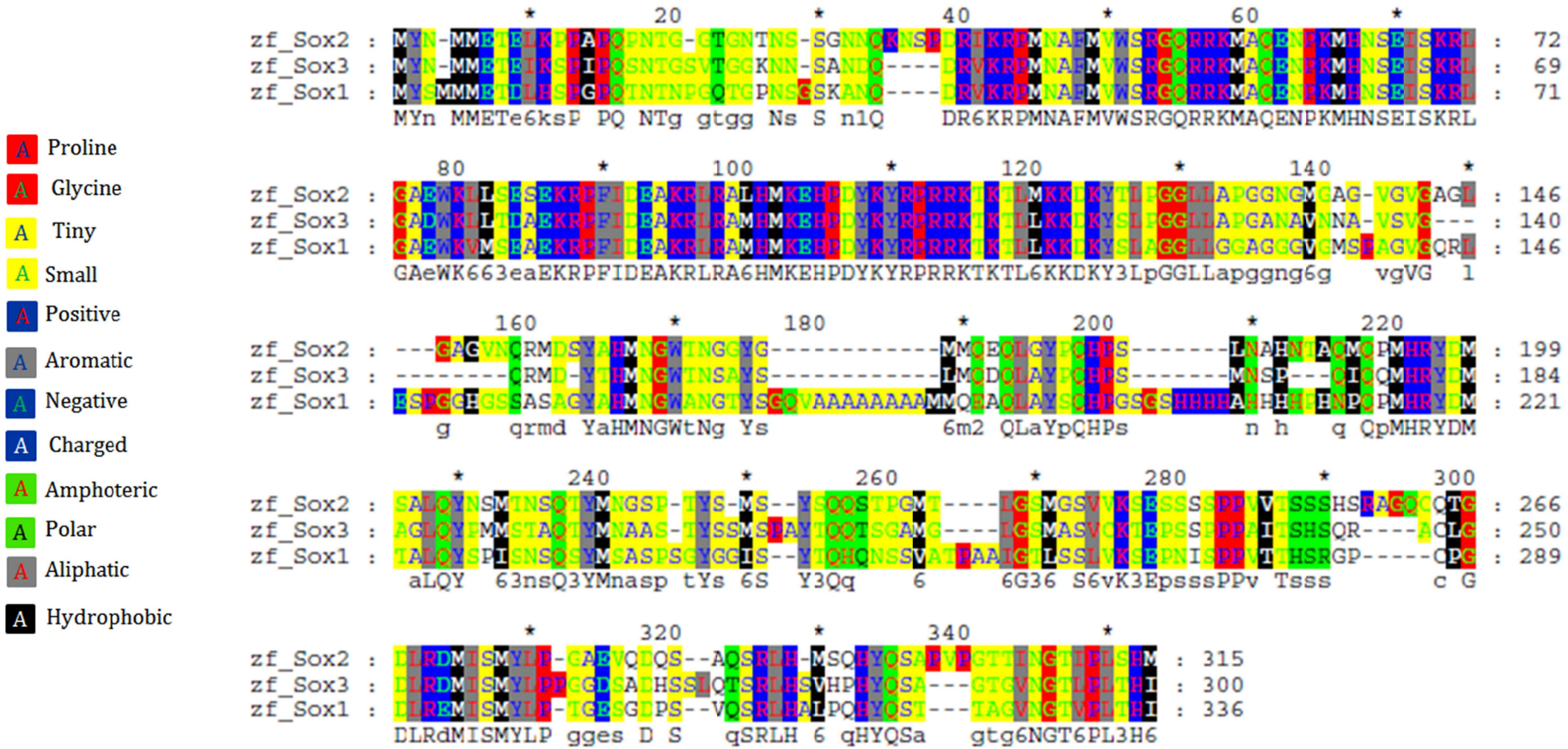


Figure 1.14 Multiple alignment of the amino acid sequences of zebrafish SoxB1 proteins (Sox1, Sox2 and Sox3). The alignment was made using ClustalX2 and shaded, on the basis of amino acid chemical and physical properties, using GenDoc.

1.5.2. Expression and Roles of Sox3 During Zebrafish Development

Sox3 is one of the first and most generally expressed transcription factors during neural development of a huge range of vertebrate species, which include the zebrafish (Okuda *et al.* 2006; Dee *et al.* 2008), *Xenopus* (Penzel *et al.* 1997; Zhang *et al.* 2003; Zhang *et al.* 2004; Zhang and Klymkowsky 2007), chicken (Rex *et al.* 1997) and mouse (Collignon *et al.* 1996; Wood and Episkopou 1999).

In zebrafish, as well as *Xenopus,* Sox3 protein is thought to be maternally-inherited (Zhang *et al.* 2003; Zhang *et al.* 2004) and the mRNA is barely detected at 32-cell stage (Okuda *et al.* 2006). While *sox3* transcript is detectable in the zygote at the beginning of gastrulation and differentiation of the ectoderm towards neural/non-neural fate, *sox1* and *sox2* are not detectable until 30% epiboly (Mizuseki *et al.* 1998; Nitta *et al.* 2006; Okuda *et al.* 2006; Rogers *et al.* 2008).

In particular, *sox3* has been showed to be present at the MBT (Mid-Blastula Transition) and its expression increases throughout the developing epiblast by 30% epiboly (Okuda *et al.* 2006). At the shield stage it is uniformly confined at the presumptive ectoderm (Okuda *et al.* 2006) and at the same time it is lost from the prospective mesodermal and endodermal cells. At the mid-gastrula stage *sox3* expression becomes confined to the neural plate (Koyano *et al.* 1997; Rex *et al.* 1997; Okuda *et al.* 2006; Dee *et al.* 2007; Dee *et al.* 2008). In later development *sox3* expression becomes even more confined and overlaps with the other *soxB1* factors (*sox1-2*) maintaining the stem cells pluripotency state (Bylund *et al.* 2003; Graham *et al.* 2003). The presence of *sox3* throughout the ectoderm by shield stage has inspired the idea that it could play a role in the earliest fate decisions (Dee *et al.* 2008). In the experiments documented by Dee *et al.* (2008) the ectopic over-expression of Sox3 led to the duplication of the central nervous system or to other similar phenotypes characterised by the formation of additional neural tissue

located at the head or the trunk of the developing embryo. This was thought to be due to cell-autonomous effects of the ectopic expression of the protein in the ectoderm, which would have promoted the expansion of the dorsal cells, fated to differentiate into neural tissue during earlier neural induction, towards more ventral regions. A model was proposed according to which Sox3 transcription factor directly activates the expression of *sox2* and *sox31* (both of which are markers for neural fate specification) and at the same time it directly inhibits *gata2* (a non-neural marker) thus re-programming the cells towards a neuro-ectodermal destiny. Therefore, this re-programming process would be the reason why the nervous system expands. Moreover, it has been showed that in zebrafish Sox3 acts both as a transcriptional activator and as a transcriptional repressor (Dee *et al.* 2008; Shih *et al.* 2010; Kuo *et al.* 2013).

1.5.3. Sox3 Acts as Transcriptional Activator During Neurogenesis

It has been demonstrated that in the chicken spinal cord Sox3 acts by activating the expression of target genes downstream of the pro-neural basic Helix-Loop-Helix transcription factors (bHLH family), and maintains the stem-like status of neural progenitor cells (Bylund *et al.* 2003), counteracting proneural proteins (normally directed by bHLH activity). In *Xenopus*, gain-of-function and loss-of-function experiments showed that Sox3 acts during primary neurogenesis by directly activating *sox2* and *geminin*, two early neural genes, in the absence of protein synthesis, and at the same time by indirectly inhibiting the Bmp target *Xvent2*. The resulting phenotypes included an increase of proliferative cells, expansion of the neural plate, delay in the neurogenesis process and, subsequently, in the neural specification (Rogers *et al.* 2009). This same effect was caused by the expression of the constitutive active form of Sox3 (Sox3HMG-VP16) and it is consistent with over-expression experiment conducted in other model organisms (Bylund *et al.* 2003; Graham *et al.* 2003; Dee *et al.* 2008).

In *Xenopus*, Sox3 also has an indirect repressor effect on *Xvent2*, probably by inducing the expression of a repressor; in fact, the expression of either ectopic Sox3 or Sox3HMG-Vp16 led to a similar repression effect, resulting in a decrease of epidermogenesis. This effect was not observed after the expression of the dominant repressor form of Sox3 (Sox3HMG-EnR) (Rogers *et al.* 2009). Together, these data suggest that Sox3 plays a role as a transcriptional activator during neurogenesis. However, as explained in the next paragraph, Sox3 has been shown to act as a transcriptional repressor in earlier stages of development.

1.5.4. Sox3 Acts as Transcriptional Repressor During Organizer Formation

In the *Xenopus* zygote, Sox3 protein and mRNA are maternally-inherited (Zhang *et al.* 2003; Zhang *et al.* 2004) and therefore present before the specification of the three germ layers. This suggested that Sox3 could play a role in the specification of the germ layers (Penzel *et al.* 1997; Rex *et al.* 1997). Accordingly, it has been demonstrated that maternal Sox3 is involved in axis formation by blocking mesoderm differentiation during the specification of the germ layers (Zhang *et al.* 2003; Zhang *et al.* 2004; Zhang and Klymkowsky 2007). It has been shown that the expression of the constitutive active version of Sox3 (Sox3HMG-VP16), knockdown of endogenous Sox3 by morpholinos, or injection of anti-Sox3 antibodies that prevented the protein from binding DNA, all caused up-regulation of *Xnr5* (Zhang *et al.* 2003) and *Xnr6* (Zhang *et al.* 2004), two genes that are regulated by endogenous *VegT* and β-catenin and induce the formation of dorsal mesoderm. The data indicate that Sox3 directly binds sites within the promoter regions of *Xnr5* and *Xnr6* and down-regulates their expression, resulting in a suppression of dorsal axis specification and therefore in ventralised phenotypes. Moreover, it has been demonstrated that this is not related with Sox3 ability to inhibit Wnt signalling by binding β-catenin or by competing with TCF/Lef (which are transcription factors regulated by β-catenin) (Zorn *et al.* 1999; Zorn *et al.* 1999; Zhang *et al.* 2003; Heeg-Truesdell and LaBonne 2006; Sinner *et al.* 2007).

Injection of affinity purified anti-Sox3 antibodies into zebrafish embryos resulted in the up-regulation of the nodal-related protein Cyclops and gastrulation abnormalities, both of which could be rescued by co-injection with mRNA encoding the dominant repressor sox3HMG-*EnR* or the nodal inhibitor *cerS* (Zhang *et al.* 2004). These data suggest that Sox3 acts as transcriptional repressor of nodal signalling in zebrafish as well as in *Xenopus*.

Sox3 plays a crucial role in the formation of the zebrafish organizer by acting as a transcriptional repressor in multiple steps of this process and confining the organizer formation both spatially and temporarily (Shih *et al.* 2010). Ectopic over-expression of Sox3 in the early zebrafish embryos, realised by RNA microinjection, caused down-regulation of *squint* (a mesoderm-derived nodal factor), of the organizer marker *goosecoid* (*gsc*), and of *chordin* (*chd*) and *noggin1* (*nog1*), which are BMP antagonists produced within the organizer. Moreover, in these experiments also the Nieukoop centre marker *bozozok* (*boz*, also known as *dharma*) was down-regulated by Sox3 ectopic over-expression and this is consistent with the observation that in *Xenopus* Sox3 down-regulates the expression of *siamosis*, whose function corresponds to the zebrafish *bozozok* (Zhang *et al.* 2003). This repression activity of Sox3 has been demonstrated to be direct and not caused indirectly by interference of β-catenin. In fact, over-expression of Sox3HMG-EnR had the same effects of the ectopic expression of wt Sox3 on all the analysed organizer markers, but this was found not to be the case of Sox3HMG-VP16. Chromatin Immunoprecipitation PCR analysis (ChIP PCR) demonstrated that Sox3 IP robustly precipitates a *bozozok* promoter fragment located 1.3 kb upstream of the transcription start region (Shih *et al.* 2010). The expression of dominant negative versions of Sox3, which would block Sox3 ability to bind its target regions on the DNA, led to ectopic expression of *chordin*, *goosecoid*, *noggin1* and *squint* into both dorsal and ventral areas of the animal pole in embryos 24 hours post-fertilization. The resulting phenotypes showed vary levels of axis duplications (Shih *et al.* 2010).

Currently, the molecular mechanisms implicated in Sox3 repression function are still not completely solved. It is important to mention that recent studies proposed the implication of SUMOylation in the switch between Sox3 activation and repression functions (Gill 2005; Gill 2005; Savare *et al.* 2005; Savare and Girard 2005; Taylor and Labonne 2005; Fernandez-Lloris *et al.* 2006; Girard and Goossens 2006; Tsuruzoe *et al.* 2006).

1.6. Aim of the Study

The aim of the present study was to gain a deeper insight into the mechanisms that regulate the ability of Sox3 to act both as a transcriptional activator and as a transcriptional repressor in the several developmental processes mentioned above (Paragraph 1.5). Since previous studies focused on specific regions of the protein, I decided to proceed with a more comprehensive approach. Therefore I designed a set of experiments that would allow the functional screening of the whole Sox3 amino acid sequence in order to identify and associate specific regions of the protein with their function/s.

The use of zebrafish as animal model for the experiments was due to its ease to be cultured and bred, and because its embryos develop fast but at the same time are considered to be a good model for studying vertebrate early development. The study was focused on the zebrafish Sox3 with the awareness of the high level of similarity that it shares with its vertebrate homologues both structurally (primary structure) and functionally, in the early stages of embryonic development. Therefore it is highly probable that the protein has conserved functions and roles in molecular regulation in the earliest stages of embryonic development among all the vertebrates.

In particular, attention was focused on one of the earliest functions of Sox3, which is the spatial and temporal regulation of the embryonic organizer formation (see Paragraph 1.5.1 for further description), together with its involvement in neural fate specification.

Because it is known that Sox3 acts as a transcriptional repressor in the context of organizer formation (Shih *et al.* 2010; Kuo *et al.* 2013), the present study aimed to investigate the regulation of the activity of Sox3 both as a transcriptional repressor and a transcriptional activator. Further analyses were performed in order to assess the activator function of Sox3 *in vitro.*

Therefore, the present study has been designed and performed in order to build the basis for subsequent, more specific analyses of the relationship between Sox3 structure and functions as a transcriptional activator and repressor. In particular, I wanted to identify large regions of the protein that contain domains or residues essential for regulating the activation or repression functions of Sox3, in order to obtain a functional map of Sox3 that could be useful for subsequent analysis.

To summarise, the objectives of the present study were:

1. The design of a strategy that could allow the functional screening of zebrafish Sox3 sequence by creating mutant constructs of the protein, based on what is known about its structure (Chapter 3.1: Design of Recombinant Forms of Sox3 for Structural-Functional Analysis);
2. The development and optimization of a cloning strategy that would allowed to obtain all the designed constructs (Chapter 3.2: Development of a Deletion Cloning Strategy);
3. The development and optimization of experimental techniques that allowed ectopic over-expression of wt and mutant Sox3 proteins in the zebrafish embryos at comparable levels (Chapter 3.3: Protein Overexpression in zebrafish Embryo);
4. The comparison between the effects of ectopic over-expression of wt and mutant Sox3 on the development of the Central Nervous System (Chapter 3.4: Effects of the Expression of Recombinant Sox3 on Neural Development);

5. The comparison between the functions of wt and mutant Sox3 proteins in the early stages of development, in particular in the context of embryonic organizer formation (Chapter 3.5: Effects of the Expression of Recombinant Sox3 on Organizer Formation);
6. The comparison between the transcriptional activation function of wt and mutant Sox3 proteins (Chapter 3.6: Transcriptional Activation Function of wt and Deletion Mutant Sox3 proteins).

2. Materials and Methods

2.1. Purification of Plasmid DNA

Bacteria containing pβUT2-zfSox3-HA plasmid were grown overnight at 37˚C on agar plates containing 1% agarose and 100 μg/ml ampicillin (Amp). A single colony was then picked and grown at 37˚C overnight shaking in 5 ml of Mu Broth culture medium (Section 6.1.1) with Amp 100 μg/ml.

2.1.1. Minipreparation of Plasmid DNA

Bacteria were centrifuged at 16000 x *g* for 3 min. Plasmid DNA was purified using a GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich) following manufacturer's instructions. Final elution from the column was performed using DEPC water (Sigma-Aldrich) at 37˚C.

2.1.2. Determination of Quality and Concentration of Purified DNA

DNA concentration was measured with a Nanodrop spectrophotometer (ND-1000). Only minipreparations with 260/280 ≥ 1.8 were considered of good quality and therefore used for subsequent experiments. The quality of purified DNA was then further evaluated by gel electrophoresis using a 1% (w/v) agarose gel.

2.2. PCR Cloning

2.2.1. Polymerase Chain Reactions

Primers were manually designed and purchased online from Sigma-Aldrich. All reagents were mixed in a different area of the laboratory prior to the addition of template DNA. The reagents used were the followings:

Forward Primer (20 µM)	2.5 µl
Reverse Primer (20 µM)	2.5 µl
DMSO	1.5 µl
2X Phusion® Master Mix*	16.5 µl
Diluted template DNA	2.5 ng
DEPC water	up to 50 µl
Total Volume	**50 µl**

*2X Phusion® High-Fidelity PCR MM w/HF Buffer Master Mix and 100% DMSO by New England Biolabs® Inc.

PCR reactions were performed with the following program on a G-Storm machine:

- Heated lid 110 °C;
- Hot Start 98 °C 1min;
- 1st stage: 6 times, 95 °C 5 seconds (sec), 58 °C 10 sec, 72 °C 1min 30 sec;
- 2nd stage: 30 times, 95 °C 5 sec, 62 °C 10 sec, 72 °C 1min 30 sec;
- 3rd stage: 1 time, 72°C 10 sec;
- Storage: 10 °C.

2.2.2. Purification of PCR Products

PCR reactions were run on 1% (w/v) agarose gel. The bands corresponding to PCR product expected size were cut out of the gel and DNA was purified using QIAquick Gel Extraction Kit (50).

Quality and concentration of purified DNA were analysed using a Nanodrop spectrophotometer and by gel electrophoresis.

2.2.3. Phosphorylation and Ligation

After gel extraction PCR products were phosphorylated before ligation. It was performed a negative control lacking of kinase enzyme in order to verify that gel extraction excluded all parental plasmid.

Phosphorylation reactions were performed under the following conditions:

	Control	**Phosphorylation**
DNA template	50 ng	50 ng
DEPC water	up to **45 µl**	up to **43 µl**

- Heating at 70 °C for 10 min;
- Chilling on ice;
- Addition of:

	Control	**Phosphorylation**
T4 Ligase Buffer (10X)*	5 µl	5 µl
T4 Polynucleotide Kinase*	/	20 units (2 µl)
Total Volume	**50 µl**	**50 µl**

*T4 Ligase Buffer and T4 Polynucleotide Kinase by New England Biolabs® Inc.

- Incubation at 37 °C for 30 min;
- Heat inactivation at 65°C for 20 min;
- Chilling on ice;
- Storage at -20 °C.

Ligation reactions of both control and phosphorylated DNA were performed under the following conditions:

DNA template	5 µl
T4 Ligase Buffer	2 µl
T4 Ligase*	400 units
DEPC water	up to 20 µl
Total Volume	**20 µl**

*T4 Ligase Buffer and T4 Ligase by New England Biolabs® Inc.

Reactions were then incubated at 4 °C overnight.

2.2.4. Transformation and Purification of Recombinant DNA

Competent cells (Alpha-Select Bronze or Gold Efficiency by Bioline Reagents Ltd.) were thawed on ice then added with ligation reaction (20 µl) and gently mixed. They were then placed on ice for 30 min, heated at 42 °C for 1 min and placed back on ice for 5 min. Transformation reaction was added with 1 ml SOC medium and incubated at 37 °C for 1 h.

After incubation bacteria were pelleted by centrifugation (7000 x *g* for 3 min), re-suspended in 100 µl Mu Broth medium (Section 6.1.1) and spread onto petri dishes containing a layer of Mu Broth Agar medium (Section 6.1.2) added with 100 µg/ml Amp. Petri dishes were then incubated at 37 °C overnight. The same protocol was applied for transformation with phosphorylated or negative control (without addition of T4 kinase) DNA.

After overnight incubation 4 colonies from each plate were picked and growth in 5 ml mu medium added with 100 μg/ml Amp at 37 °C shaking overnight.

DNA from each colony was purified trough plasmid minipreparation (see Paragraph 2.1.1). Quality and Concentration of purified plasmids were checked both with electrophoresis and Nanodrop measurement (see Paragraph 2.1.2).

2.2.5. Assessment of the Clones' Sequences

The clones obtained were analysed through the following two steps:

1) Firstly, a restriction digestion with *NaeI* was performed to identify the recombinant clones. Both wt Sox3 (pβUT2-zfSox3-HA plasmid) and mutant forms of Sox3 were digested with NaeI (New England Biolabs® Inc.) with the following protocol:

	Wt Sox 3	**Recombinant**
Buffer	2 μl	2 μl
DNA	~0.3 to 0.5 μg	~0.3 to 0.5 μg
Nae I (10.000 units/ml)	1 μl	1 μl
DEPC water	up to 20 μl	up to 20 μl
Total Volume	**20 μl**	**20 μl**

Digestion reactions were incubated at 37 °C for 2 h then analysed.

2) After the identification of the presumptive mutant clones through restriction digestion, sequencing was performed in order to confirm the sequences of each of the selected clones. An amount of 0.5 μg of each predicted mutant was sent to Source Biosciences (http://www.sourcebioscience.com/) for sequencing. Sequencing data were read with FinchTV (http://www.geospiza.com/) and the

sequences were aligned were using Nucleotide BLAST (http://blast.ncbi.nlm.nih.gov/).

2.2.6. Storage of the Clones Obtained

Competent cells containing the recombinant plasmid were grown from a single colony in 5ml Mu Broth medium (Section 6.1.1) for 16 hours at 37 °C. Bacteria were subsequently spin down and pellet was resuspended in 500 µl or 800 µl Mu Broth medium (Section 6.1.1). After mixing an amount of 500 µl or 200 µl of 90% glycerol was added in order to obtain 45% and 18% glycerol stocks. The stocks obtained were stored at -80 °C.

2.3. *In Vitro* Transcription

In vitro transcription was performed on plasmids containing *green fluorescent protein* gene (*pCS2nlsGFP*, see map in Appendix 6.2), wild-type *sox3* (wtSox3, plasmid *pβUT2-zfSox3-HA*, see map in Appendix 6.1) and the deletion mutant constructs of *sox3* (same plasmid as wtSox3) in order to obtain capped RNA for embryo microinjection.

2.3.1. Digestion and Purification of the Template

Template DNA was obtained from glycerol stocks (see Appendix 7.1 and Appendix 7.2 for the maps of the vectors containing wtSox3 and GFP) of from the transformed colonies of competent cells in the case of the mutant Sox3 constructs. Glycerol stocks were transported on ice and bacteria were scraped and added to 5 ml mu, previously added with Amp, using a filtered pipette tip. Cells were grown for 16 hours at 37 °C and then DNA was extracted using GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich).

Plasmids were subsequently digested with *Eco*RI or *Not*I(New England Biolabs® Inc.) according to the following protocol:

Buffer	8 µl
DNA	2,5 µg
*Eco*RI (*Not*I for GFP)	2 µl
DEPC water	up to 80 µl
Total Volume	**80 µl**

Digestion reactions were incubated at 37 °C for 2 h, then 1µl of digest was run on 1% (w/v) agarose gel in order to check digestion.

Phenol-Chloroform extract was performed in order to clean digested DNA template according to the following protocol:

1. Addition of 1 vol. Phenol:Chloroform:Isoamyl Alcohol (purchased from Sigma-Aldich: Phenol:Chloroform:Isoamyl Alcohol 25:24:1 saturated with 10 mM Tris, pH 8.0, 1 mM EDTA);
2. Vortex 3 sec;
3. Centrifuge at 16000 x *g* for 3 min;
4. Transfer of the upper layer into a clean eppendorf tube;
5. Repetition of steps 1,2, 3 and 4;
6. Addition of 1 vol. of Chloroform;
7. Vortex 30 sec;
8. Spin at 16000 x *g* for 3 min;
9. Transfer of the upper layer into a clean eppendorf tube;
10. Repetition of steps 6, 7, 8 and 9;
11. Addition of 0.1 vol. NaAc 3M and 2.5 vol. EtOH;
12. Incubation at -20 °C for 30 min;
13. Spin at 16000 x *g* for 30 min at 4 °C;
14. Rinse of the pellet with ice cold 70% EtOH;
15. Spin at 16000 x *g* for 5 min at 4 °C;
16. Removal of EtOH and air drying of the pellet for 10 min;
17. Resuspension of the pellet in 10 µl RNAse free H_2O;

18. Storage at -20 °C.

After extraction, 1 µl of DNA was run on 1% (w/v) agarose gel in order to check it was not lost during proceedings.

2.3.2. *In Vitro* Transcription and RNA Purification

In vitro transcription reactions were performed using mMessage mMachine® T3 (for obtaining GFP transcript mMessage mMachine® SP6, both by Ambion) transcription kit according to the following protocol:

Linear template DNA	1 µg
2 x NTP / CAP	10 µl
10 x Reaction Buffer	2 µl
Enzyme Mix	2 µl
Total Volume	**20 µl**

The reaction was incubated at 37 °C for 5 h, and then put on ice while running 1 µl of each reaction was tested by gel electrophoresis.

Phenol-Chloroform Extraction was performed to clean capped RNA as follows:

1. Addition of 115 µl Nuclease-Free H_2O (from kit) and 15 µl Ammonium-Acetate STOP Solution;
2. Extraction with 1 vol. Phenol:Chloroform:Isoamyl Alcohol (purchased from Sigma-Aldich: Phenol:Chloroform:Isoamyl Alcohol 25:24:1 saturated with 10 mM Tris, pH 8.0, 1 mM EDTA), vortex 30 sec then spin at 16000 x *g* for 3 min;
3. Extraction with 1 vol. Chloroform, vortex 30" then spin at 16000 x *g* for 3 min;
4. Addition of 1 vol. isopropanol to precipitate RNA;
5. Incubation at -20 °C for 30 min;
6. Spin at 16000 x *g* for 15 min at 4 °C;
7. Removal of supernatant, air drying of the pellet and resuspention in 30 µl Nuclease-Free H_2O;

8. Run 1 µl on 1% (w/v) agarose gel at 150 volt for 25 min and Nanodrop measurement of RNA concentration.

2.3.3. Storage of Capped RNA

A 15 µl aliquot of the capped RNA (CAP RNA) was directly stored at -80 °C. The remaining 15 µl were diluted with Nuclease-Free H_2O to a final concentration of 100 ng/µl and stored in 5 µl aliquots at -80 °C.

2.4. Embryo Injection

2.4.1. Set Up of Zebrafish and Harvesting of the Embryos

Pairs of fishes were set up in the late afternoon in order to allow them to lay fertilized eggs the following morning. Small plastic boxes with mesh bottom were fitted into larger boxes and filled with water. A pair of fish was placed into each of the plastic boxes, divided from each other by a transparent plastic device. The device was removed the following morning and after 20 min the fish were removed as well as the top container. The embryos, which collected from the bottom container, were washed and placed into petri dishes filled with water containing dimethylene blue (2ml of 0.1% methylene blue into 1L of water).

2.4.2. Microinjection of Capped RNA in Zebrafish Embryos

Needles were fabricated by heating and pulling borosilicate glass capillary tubes in a micropipette puller device, and then stored in petri dishes on top of a small stripe of clay. CAP RNA was thawed on ice then used for backloading needles using a microloader pipette. The tip of the needle was broken with steel tweezers, then the needle was inserted into the microinjector. One-two cell stage zebrafish embryos were aligned on a petri dish against a glass microscope slide and then microinjected with 50

pg of CAP RNA. The RNA was microinjected directly into the cytoplasm (Fig. 2.1).

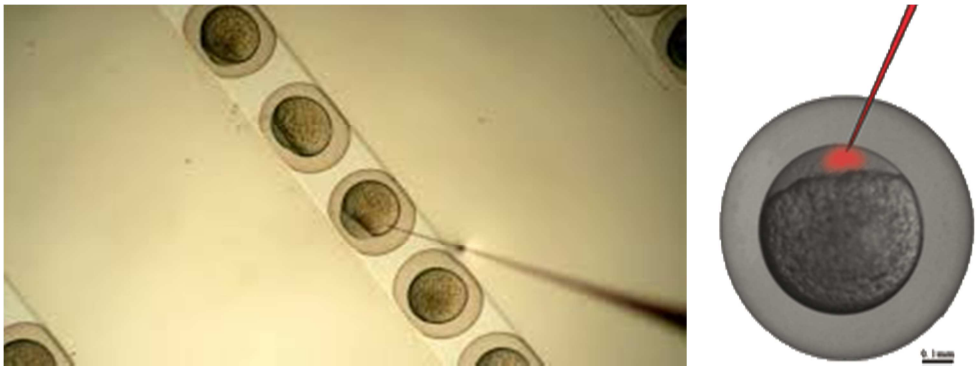


Figure 2.1 RNA microinjection into Zebrafish 1-cell stage embryos (Credits: Grabner Lab website and Dominik Paquet website).

2.4.3. Incubation of the Embryos

Microinjected embryos and their uninjected siblings were incubated all at the same time at 28 ˚C for the required time, which ranged from 4.5 hpf to 24 hpf.

2.5. Western Blot

2.5.1. Preparation of Samples for Western Blot

After microinjection and incubation at 28 ˚C until reaching Sphere Stage, 30 embryos for each sample were collected and the chorions were removed using steel tweezers. Water was removed and replaced with 200 µl of Deyolking Buffer (55 mM NaCl, 1,8 mM KCl, 1,25 mM $NaHCO_3$ in SDW) and embryos were homogenized by pipetting with 200 µl filtered pipette tip for 20 times. Samples were then centrifuged in eppendorf tubes at 124 x *g* for 5 min, supernatant was removed and 30 µl of 2X Laemmli Sample Buffer (Sigma-Aldrich) were added to each sample before mixing and storing them at -80 ˚C.

2.5.2. Preparation of Denaturing Gel and Apparatus for SDS-PAGE

Running gel (12.5% w/v of polyacrylamide) was prepared according to the following recipe, poured between two gel plates in a gel castor and covered with 70% EtOH:

Reagent	**Vol.**
Acryl-Bis 29:1	2.14 ml
Solution B*	1.43 ml
SDW	1.44 ml
APS 10%	75 µl
TEMED	7.5 µl

*Solution B: 1.57 M Tris HCl pH 8,8 and 0.4% SDS.

After removal of EtOH a stacking gel (12.5% w/v of polyacrylamide) was prepared according to the following recipe and poured over polymerized running gel:

Reagent	**Vol.**
Acryl-Bis 29:1	0.65 ml
Solution C**	1.25 ml
SDW	1.73 ml
APS 10%	75 µl
TEMED	7.5 µl

**Solution C: 0.5 M Tris HCl pH 6,8 and 0.4% SDS.

After solidification the gels were transferred into the gel tank, which was then filled with SDS Running Buffer (Section 6.2.1).

2.5.3. Gel Running and Electrophoretic Transfer

Samples were thawed on ice for 10 min, then boiled at 110 ˚C for 10 min on heating block, placed back on ice for 10 min and centrifuged at 124 x *g* for 5 min. Only supernatant was collected and loaded on the SDS gel together with 5 µl of ladder (SeeBlue® Plus2 Pre-Stained Standard, Invitrogen). When performing western blot with recombinant Sox3 forms, half of the supernatant of each sample was loaded into each of two gels. This allowed performing the loading control (α-actin) and detecting Sox3 HA-tagged avoiding superposition of the corresponding bands.

The gels were run at 40 V until all the samples has reached the running gel, then at 90 V until the markers contained into Laemmli Sample Buffer (Sigma-Aldrich) had reached the bottom of the gel. The gel were then removed and a sandwich was made by putting together sponge, filter paper, gel, membrane, filter paper, then sponge again. The sandwich was made by soaking all the components in the same Transfer Buffer (Section 6.2.2), which was later used also to fill the transfer tank.

The transfer was run either at 35 V overnight or 110 V for 1 h RT. The membrane was then removed and washed into 5% Marvel Milk in PBST for at least 1 h at RT, with gentle shaking. The primary antibodies (rabbit anti-HA or anti-actin for loading control, both purchased from abcam®) were diluted 1/4000 in 2% w/v Marvel Milk powder in PBST and used to soak the membrane overnight at 4 ˚C gently shaking. The following morning the membrane was washed in PBST for 3 times 5 min each gently shaking, then it was replaced with secondary antibody (LiCor anti-rabbit, purchased from LI-COR®) diluted 1/4000 in 2% Marvel Milk in PBST and incubated for at least 1 h RT. After washing with PBST for 3 times 5 min each, the membrane was finally scanned using LiCor Scanner.

2.6. Whole Mount *In Situ* Hybridization

2.6.1. RNA Microinjection

RNA microinjection was performed as described in Chapter 2.3 together with each construct GFP capped RNA (50 pg) was injected. After incubation only embryos showing fluorescence under UV light were selected and fixed. This was carried on in order to make sure that microinjection had been effective for all the embryos analysed.

2.6.2. Fixation and Storage of the Embryos for *In Situ* Hybridization

After incubation at 28˚C until the embryos reached the desired developmental stage, only viable (and fluorescent) embryos were collected, washed once with Phosphate-Buffered Saline containing 0.1% Tween20 (PBST) and fixed with 4% Paraformaldehyde (PAF) in Phosphate-Buffered Saline (PBS) for a few minutes shaking at RT, then overnight at 4 ˚C.

After fixation embryos were washed four times with PBST, then the chorions were removed using steel tweezers and washed with an increasing gradient of methanol at RT gently shaking:

Washing	**Time**
100% PBST	5 min
25% MeOH + 75% PBST	5 min
50% MeOH + 50% PBST	5 min
75% MeOH + 25% PBST	5 min
100% MeOH	4x5 min

Embryos were then stored at -20 ˚C for at least overnight.

2.6.3. Preparation of DNA Template for Synthesis of Riboprobes

Cells containing the plasmids with the genes of interest were striped from glycerol stocks on petri dishes containing a layer of Mu Broth Agar medium (Section 6.1.2) with addition of 100 µg/ml Amp, and then grown at 37˚C for 16 h. Single colonies were then transferred into 5 ml Mu Broth (Section 6.1.1) added with Amp 100 µg/ml and grown at 37˚C for 16 h shaking. Plasmid DNA was prepared as described in Section 2.1. The templates were subsequently digested overnight (75 µg DNA in a tot vol. of 40 µl) with the required restriction enzyme, and then gel purified (GenElute™ Plasmid Miniprep Kit by Sigma-Aldrich). The concentration of template DNA was subsequently checked both running 1 µl on gel electrophoresis and with Nanodrop measurement before storage at -20 ˚C.

Sequencing was performed by sending 0.5 µg of each digested template to Source Biosciences.

2.6.4. DIG-labelled Riboprobes Synthesis

The following labelling reaction was set up on ice and then incubated for 3h at 37˚C.

5x Buffer*	4 µl
DTT*	2 µl
10X MIX**	2 µl
RNasin* (40 u/µl)	0.5 µl
Polymerase*	2 µl
DNA template	800 ng
DEPC water	up to 20 µl
Total Volume	**20 µl**

*purchased from Promega, **purchased from Roche.

Probes were cleaned using illustra™ MicroSpin™ G-50 Columns (by GE Healthcare), then checked running 1 µl on 1% (w/v) agarose gel at 150 volt for 25 min and stored at -80˚C after addition of 1 vol. of Hybridization Buffer.

Probes were synthetized for all the following genes:

Gene	Restriction Enzyme	RNA polymerase	Vector
bozozok (*boz*)	*Bam* HI	T7	pCR II
chordin (*chd*)	*Not* I	T7	pCS2
goosecoid (*gsc*)	*Sma* I	T7	pCS2
n-cadherin (*ncad*)	*Xho* I	SP6	unknown
notail (*ntl*)	*Xho* I	T7	unknown
squint (*sqt*)	*Bam* HI	T7	pCS2

2.6.5. In Situ Hybridization

All the buffers necessary for the experiment were previously prepared according to the recipes described in Section 6.3.

Embryos were rehydrated into PBST through a gradient:

Washing	Time
75% MeOH + 25% PBST	5 min
50% MeOH + 50% PBST	5 min
25% MeOH + 75% PBST	5 min
100% PBST	4x5 min

Then they were washed once for 5 min with 50% PBST + 50% Hybridization Buffer and incubated with pre-Hybridisation Buffer for 3h at 67˚C.

The pre-hybridisation buffer was subsequently replaced with 100µl Hybridization Buffer containing a 1:200 dilution of probe and incubated overnight at 67˚C. The Hybridization Buffer containing the probe was then stored at -20 ˚C and reused at least twice.

The following washings were then performed at 67˚C with occasional agitation:

Washing	**Time**
100% Hybridization Buffer	1x10 min
75% Hybridization Buffer + 25% 2x SSC	1x10 min
50% Hybridization Buffer + 50% 2x SSC	1x10 min
25% Hybridization Buffer + 75% 2x SSC	1x10 min
2x SSC	1x10 min
0.2x SSC	4x15 min

The following washings were performed at RT gently shaking:

Washing	**Time**
75% 0.2x SSC + 25% MABTw	5 min
50% 0.2x SSC + 50% MABTw	5 min
25% 0.2x SSC + 75% MABTw	5 min
100% MABTw	5 min

Blocking was performed with 2% Boehringer Blocking Reagent™ in MAB for 1h RT gently shaking.

This was then replaced with antibody (anti-Dig-AP Fab fragments 1:5000 in MAB Blocking buffer), shake for 5 min at RT then incubated at 4˚C overnight.

The following morning the samples were shaken for 1 h at RT, then antibody solution was removed and samples were wahsed 8x15 min with MABTw .

The embryos were then equilibrated with BCL buffer III 3 times 5 min each at RT. This was replaced with developing buffer containing BCL buffer + BM Purple™ 1:1 and incubated at 4 ˚C protected from light for at least overnight or until the colour had developed. The reaction was stopped by washing with PBST 20mM EDTA 3 times 5 min each, then the embryos were fixed in 4%Paf in PBST for 20 min RT. Paraformaldehyde was replaced with PBST by washing 3 times 5 min each and finally the embryos were put into 90% glycerol and stored at 4 ˚C.

2.6.6. Evaluation of Embryos after *In Situ* Hybridisation

After *in situ* hybridization embryos were evaluated by comparison between uninjected controls, embryos injected with wt Sox3 and embryos injected with the different Sox3 mutant forms under a Nikon SMZ15000 microscope. Pictures were taken using a Nikon DS-5M camera and the Nikon ACT-2U 1.40 software. In order to get objective results the evaluation of the stained embryos was always performed with a “blind” approach, which included covering the samples’ labels and mixing them before analysing.

2.7. Luciferase Reporter Assay

2.7.1. Preparation of the Samples

Embryos were co-injected with 50 pg of GFP CAP RNA, 50 pg of wt/mut Sox3 CAP RNA and 5 pg of pGL3-3XSX plasmid (for a map of the plasmid see Appendix 7.4). Subsequently they were incubated at 28 ˚C, collected at Sphere Stage and visualised with an optic microscope under UV light: only embryos that were expressing GFP diffusely and at similar level were selected. For each group about 50 embryos were injected and 30 were selected for the analysis. The chorion was removed from each one of the

selected embryos live using steel tweezers. The embryos were then diluted with Passive Lysis Buffer (5X PLB Promega, 5 µl of 1X PLB in SDW was used for each embryo), vortex and stored at -80 ˚C for up to one week. For each experimental group the experiment was repeated three times, each with samples of 10 embryos.

2.7.2. Luciferase Reporter Assay

The samples, which contained 10 embryos each, were thawed on ice, centrifuged at 16000 x g for 5 min and 35 µl of supernatant was transferred into fresh eppendorf tubes. They were then centrifuged again at 16000 x g for 5 min and 25 µl of supernatant was transferred into white 96-well plate. The reading was made with a GloMax luminometer using the Luciferase Assay System (Promega).

3. Results

3.1. Design of Sox3 Deletion Mutants for Structure-Function Analysis

3.1.1. Introduction

As described in Chapter 1, Sox3 sequence and functions have been highly conserved during evolution. However, at the present time it is still not clear which regions of the protein are involved in its different functions as a regulator of gene expression at different stages of embryonic development. The aim of the present study was therefore to screen through the entire Sox3 protein sequence looking for regions involved in either organizer repressor function or neural induction function. Experiments were designed in order to allow rapid and individual testing of the involvement of different parts of the protein in such mechanisms by testing the effects of their deletions on target genes and comparing them to the effects of wt Sox3.

3.1.2. Development of a Screening Strategy Based on the Evolutionary Conservation of Sox3

The aim of the study was to find a link between Sox3 structure and function in early embryo development. As the three-dimensional structure of Sox3 has not yet been solved, the main idea of the project was based on the assumption that highly conserved positions in proteins are often indicative of structural and/or functional importance (Ashkenazy *et al.* 2010). Multiple alignment of the protein sequence of *Danio rerio* (bony fishes), *Carassius auratus* (bony fishes), *Amphiprion melanopus* (bony fishes), *Gallus gallus* (birds), *Taeniopygia guttata* (birds), *Mus musculus* (rodents), *Xenopus silurana* (frogs), *Xenopus laevis* (frogs), *Pan troglodytes*

(primates) *and Homo sapiens* (primates), showed higher grade of conservation of some regions of the gene with respect to others (Fig. 3.1).

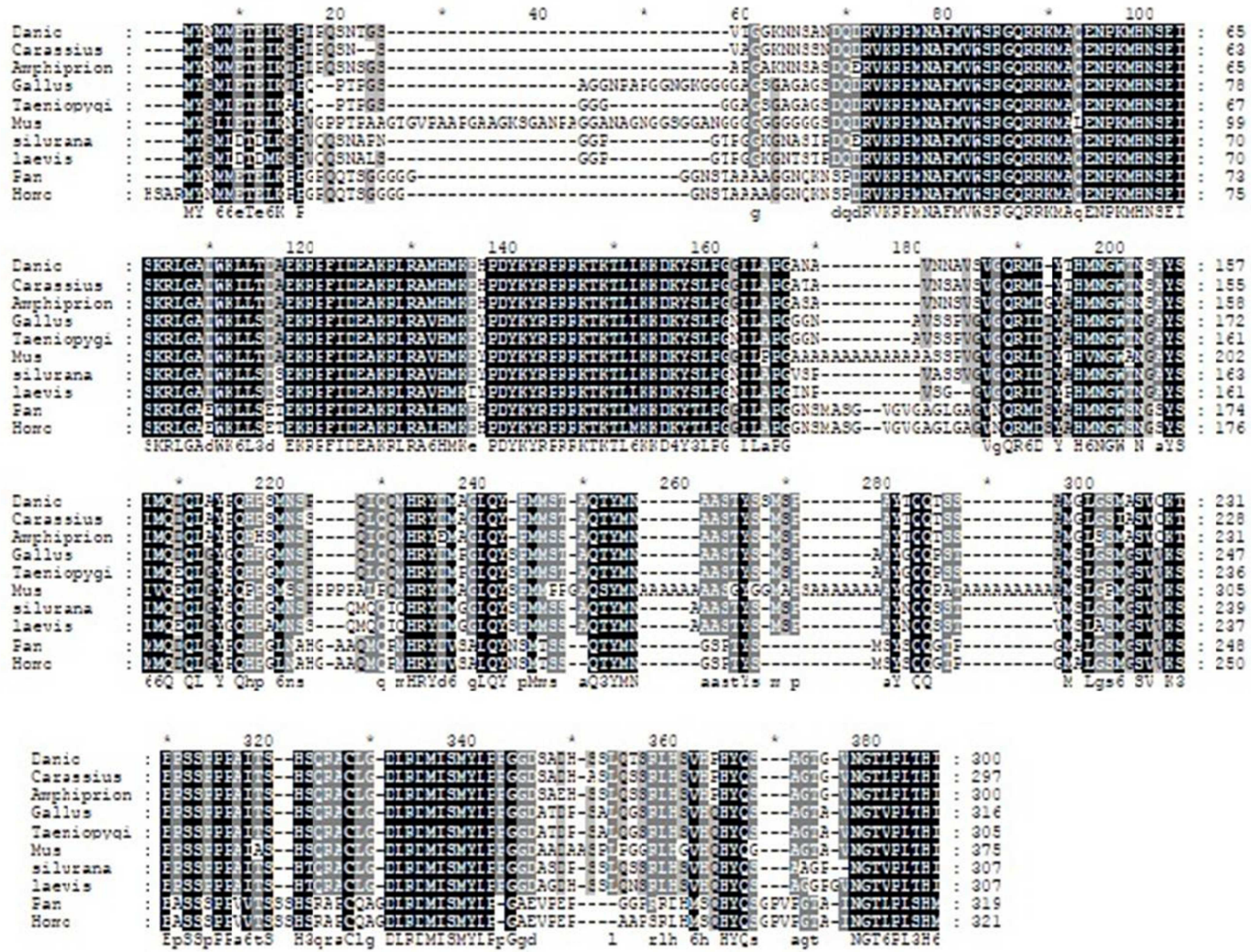


Figure 3.1 Multiple alignment between Sox3 amino acid sequences of *Danio rerio* (Danio*), Carassius auratus* (Carassius), *Amphiprion melanopus* (Amphiprion), *Gallus gallus* (Gallus*), Taeniopygia guttata* (Taeniopygi), *Mus musculus* (Mus), *Xenopus silurana* (silurana), *Xenopus laevis* (laevis), *Pan troglodytes* (Pan) and *Homo sapiens* (Homo). Alignment realized using ClustalX2 and GeneDoc. Black shading represents identical residues, grey shading represents similar residues and white shading represents different residues.

3.1.3. Design of Deletion Mutants of Zebrafish Sox3

The strategy developed for this study was to delete both conserved and non-conserved regions, grouping the amino acids on the basis of conserved/not conserved clusters. Deletions of about 20 amino acids each were designed in order to cover the entire protein sequence but at the same time creating a reasonable number of mutants for a first scan of Sox3. The HMG-box domain was not removed nor modified not to affect Sox3 DNA binding ability. The total number of mutants designed was eleven (Fig. 3.2, Table 3.1 and Appendix 7.4) and they were numbered chronologically on the basis of when they were designed. During cloning experiments a twelfth clone was created and maintained as a negative control during all the following experiments. This clone (Mut12) contains the HMG-box domain and about 39 amino acids of the C-terminal region of the protein but lacks of the central portion; it was expected to lack most of the functions and its over-expression was therefore considered unlikely to affect embryo development.

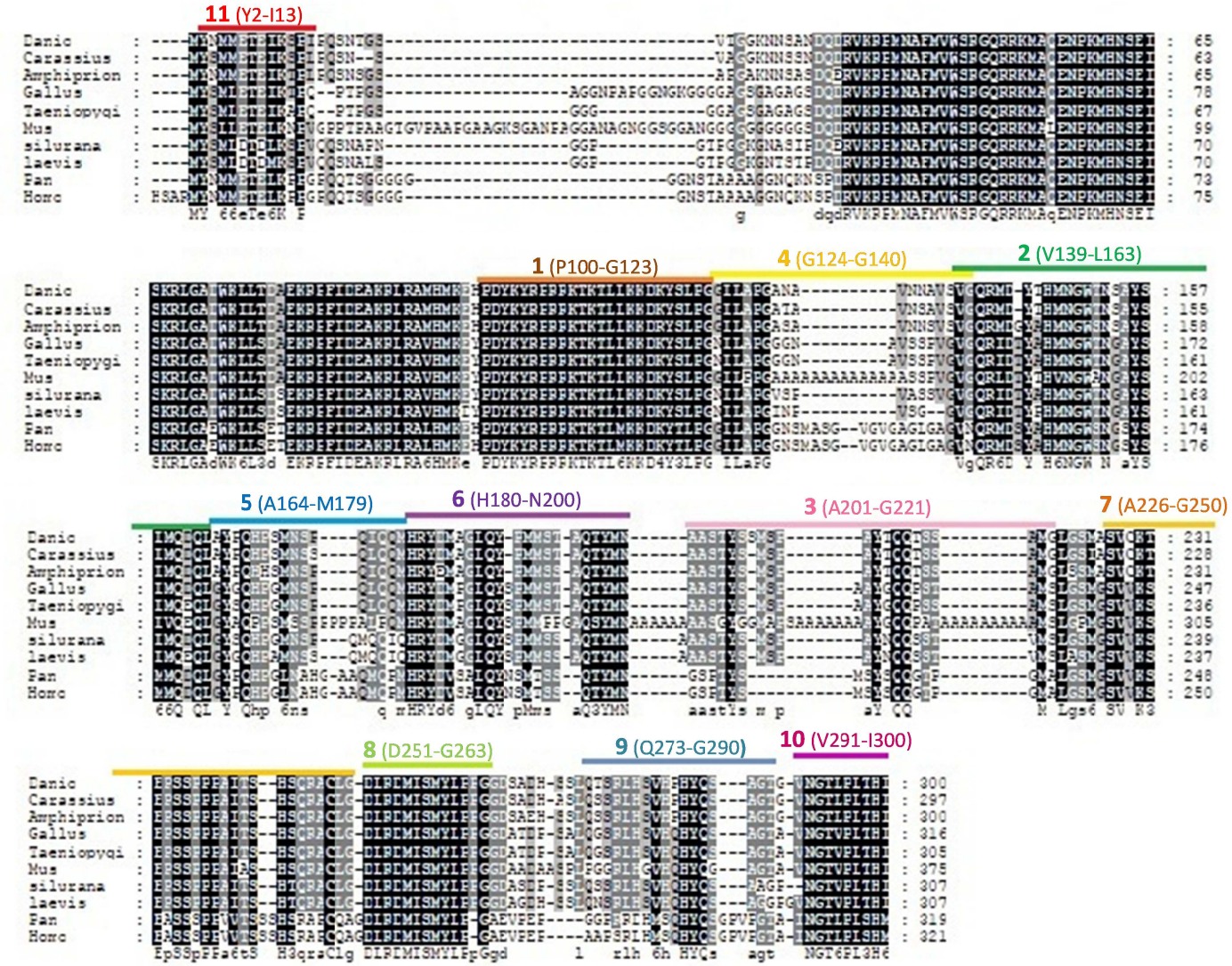


Figure 3.2 The eleven deletion mutants of Sox3 designed on the basis of the sequence conservation during evolution. The regions deleted in each construct are indicated with different colours and numbered accordingly to the names of the mutants. Multiple alignments were realized using ClustalX2 and visualised using GeneDoc. Black shading represents identical residues, grey shading represents similar residues and white shading represents different residues.

Name	Deletion	Level of Conservation
Mut1	P100-G123	Highly conserved region
Mut2	G140-L163	Medium conserved region
Mut3	A201-G221	Poorly conserved region
Mut4	G124-V139	Poorly conserved region
Mut5	A164-M179	Poorly conserved region
Mut6	H180-N200	Medium conserved region
Mut7	A226-G250	Medium conserved region
Mut8	D251-G263	Highly conserved region
Mut9	Q273-G290	Poorly conserved region
Mut10	V291-I300	Highly conserved region
Mut11	Y2-I3	Quite high conserved region
Mut12	P100-P122 + P128-P261	-

Table 3.1 Nature of deleted regions of mutant forms of Sox3 that carry deletion in different regions of the protein.

3.2. Development of a Deletion Cloning Strategy

Zebrafish Sox3 had been previously cloned into the *pβUT2* vector including an HA tag downstream the gene in order to allow immunohistochemical detection of the protein. The same plasmid was used for all the experiments described (Appendix 7.1). The cloning strategy was designed to meet two main objectives: firstly to realise deletions without altering dramatically the remaining part of the protein, and secondly to provide a fast way to screen the clones to identify those containing the mutations before sequencing. The strategy was to replace the deleted region with a restriction site, so that it could be possible to assess the presence of the mutation just performing a restriction digestion followed by gel electrophoresis (Fig. 3.3). The enzyme chosen for these experiments was *Nae*I (restriction site 5'GCCGGC3') as its translation produced Alanine-Glycine, amino acids that were expected to have low impact on the protein secondary structure.

The ability of NaeI to cut *pβUT2-zfSox3-HA* was tested both using NEB cutter V2.0 (http://tools.neb.com/NEBcutter2/index.php) and performing restriction digestion followed by gel electrophoresis. The test showed that the enzyme did not cut the parental plasmid, while it cut the recombinant plasmids that contained the restriction site (Fig. 3.4 and Fig. 3.5).

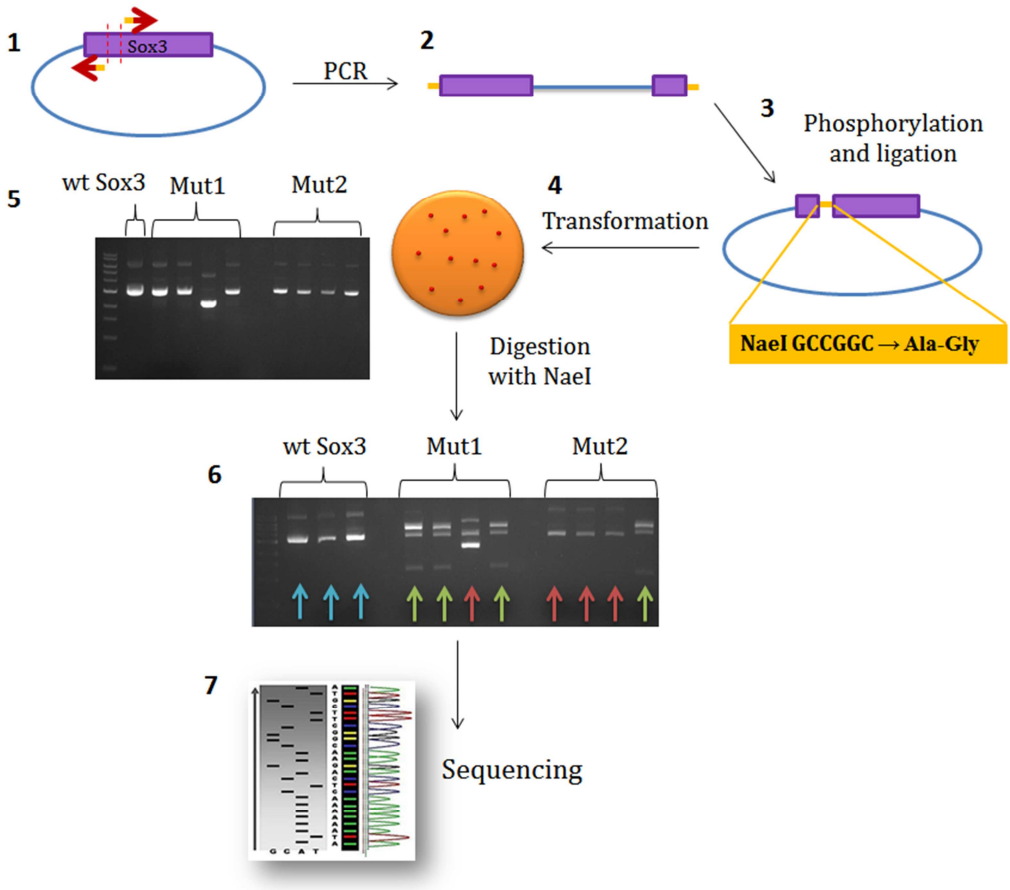


Figure 3.3 Schematic representation of the cloning strategy. For each mutant two primers were designed: a reverse primer located towards the region of the gene corresponding to the N-terminus of the protein, and a forwards primer located towards the region of the gene corresponding to the C-terminus of the protein (1). The PCR products obtained (2) with each pair of primers were phosphorylated and ligated (3), then used for transforming competent cells that were grown on agar plates (4). In order to identify the recombinant colonies, the plasmid DNA was purified (5), digested with NaeI and analysed through gel electrophoresis (6). For each of the eleven deletion mutants one colony was selected and the exact sequence of each recombinant plasmid purified from the selected colonies was subsequently obtained through DNA Sanger sequencing (7).

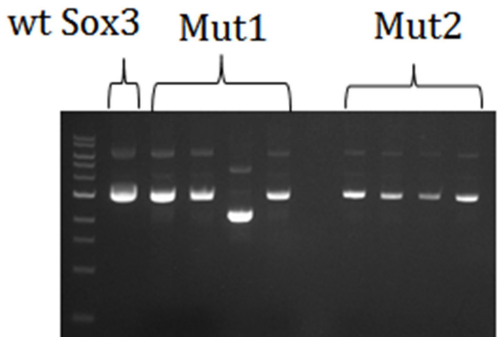


Figure 3.4 Gel represented in point 5 of figure 3.3. The first lane corresponds to 1kb DNA ladder (New England Biolabs®, 0,05 µg), the second lane corresponds to the wt Sox3 plasmid (*pβUT2-zfsox3-HA*), the folllowing four lanes, which are indicated as "Mut1", correspond to plasmids purified from four different colonies of competent cells transformed with *mut1* PCR products. The last four lanes, which are indicated as "Mut2", correspond to plasmids purified from four different colonies of competent cells transformed with *mut2* PCR products.

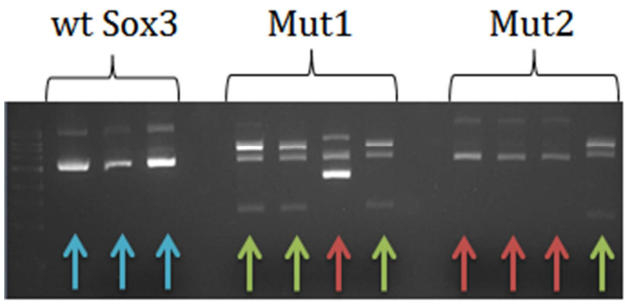


Figure 3.5 Gel represented in point 6 of figure 3.3. Light blue arrows indicate the wt plasmid (*pβUT2-zfsox3-HA*); green arrows indicate plasmids that were successfully digested by NaeI and presumably contain the desired deletion; red arrows indicate plasmids that were not digested by NaeI. The first lane corresponds to 1kb DNA ladder (by New England Biolabs®, 0,05 µg), the second lane to *pβUT2-zfsox3-HA*, the third and the fourth lanes correspond to different concentrations of *pβUT2-zfsox3-HA* after digestion with NaeI (same reaction conditions used for the presumptive mutant plasmids). The four lanes indicated as "Mut1" correspond to the same four plasmids represented in Fig3.4 after digestion with NaeI, the four lanes indicated as "Mut2" correspond to the same four plasmids represented in Fig3.4 after digestion with NaeI. The enzyme proved to partially digest the parental plasmid (this is more evident in the fourth lane, which contains high concentration of digested parental plasmid and where three bands are visible). This explains why the lanes corresponding to digested mutants contain multiple bands.

3.2.1. Primers Design

For each clone a pair of primers of between 19 and 27 nucleotides was designed (Appendix 7.3), so that the primer located towards the 5'-terminus of the gene (region codifying for the N-terminus of the protein) was the Reverse Primer and the primer located towards the 3'-terminus of the gene (region codifying for the C-terminus of the protein) was the Forward Primer (Fig. 3.3). This approach allowed performing a single ligation reaction after PCR. Both the Reverse and the Forward primers had a "GGC" triplet added at their 5'-end in order to obtain the Nae I restriction site inserted in the final transcript.

3.2.2. Results of the Cloning

The first part of this project was to design and develop the cloning strategy described above. This involved the optimization of the protocol, including the testing of different experimental conditions during different steps, such as PCR reaction, ligation and transformation. In particular, the experiments involved the testing of different volumes of reagents and different temperatures for the PCR reaction. Moreover, different ligase enzymes and different steps and times of incubation were tested for the ligation and phosphorylation reactions. Also the transformation protocol was developed through the testing of different conditions, such as the use of different types of competent cells. Once the protocol was optimized, all the designed mutants were effectively obtained through this cloning strategy. Figures 3.6 and 3.7 show, respectively, the nucleotide sequence of wt Sox3 and the sequence of one of the deletion mutant constructs obtained with the cloning strategy described above.

The cloning strategy that was developed allowed the creation of large deletions and to insert a new restriction site by inserting just minor new modifications in the final protein. In fact, the inserted site is translated into Alanine-Glycine, small un-charged amino acids. In addition, the deletions were designed, when possible, in order to include an Ala, or a Gly, or both

at the ends of the deleted regions, so that they would have been replaced by the restriction site (Table 3.2).

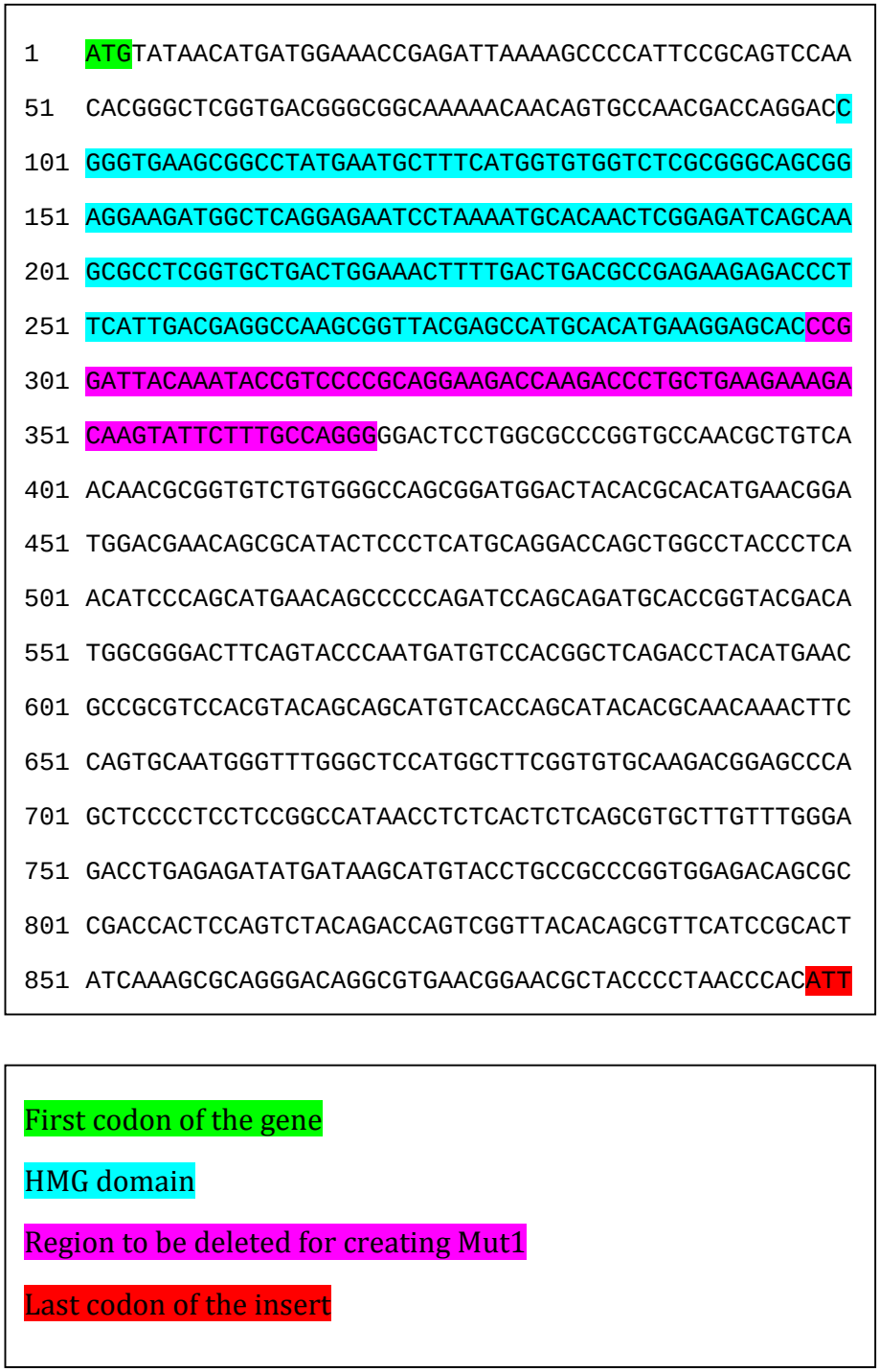


Figure 3.6 Sequence of the wt Sox3 gene inserted in the pβUT2-zfSox3-HA vector. The region highlighted in purple indicates the region that is missing in the Mut1 construct.

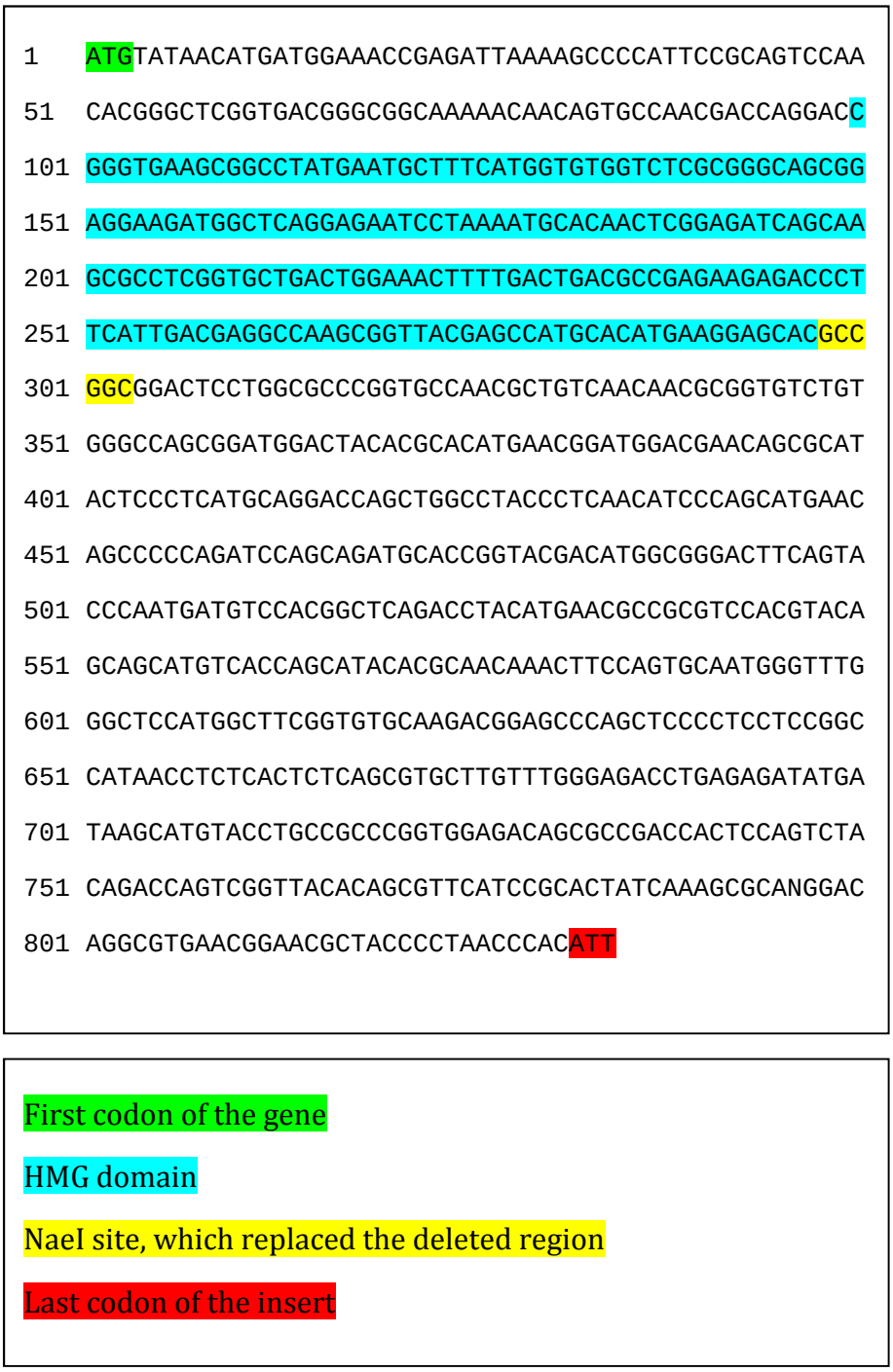


Figure 3.7 Sequence of Mut1 obtained through sequencing. The region of the wt gene that was deleted has been successfully replaced with the restriction site of NaeI (highlighted in yellow), which will be translated in the amino acids Alanine-Glycine.

Name of the construct and deleted region	Length of the deletion (nucleotides)	Residues that replaced the deleted region
Mutant 1 (P100-G123)	24	Ala
Mutant 2 (V139-L163)	25	Ala-Gly
Mutant 3 (A201-G221)	21	-
Mutant 4 (G124-G140)	17	Ala
Mutant 5 (A164-M179)	16	Gly
Mutant 6 (H180-N200)	21	Ala-Gly
Mutant 7 (A226-G250)	25	-
Mutant 8 (D251-G263)	13	Ala
Mutant 9 (Q273-G290)	18	Ala
Mutant 10 (V291-I300)	10	Ala-Gly
Mutant 11 (Y2-I13)	12	Ala-Gly
Mutant 12 (P100-P122 + P128-P261)	157	Ala

Table 3.2 Deletion mutants of Sox3 carrying deletions in different regions of the protein. The table indicates the number of nucleotides deleted from each construct and the amino acids that replaced the deletion.

3.3. Protein Overexpression in Zebrafish Embryos

3.3.1. A Strategy to Induce Ectopic Protein Expression in Early Zebrafish Embryos

In recent years the use of microinjection to study gene function in the zebrafish has become widespread (Rosen *et al.* 2009) . This includes ectopic expression of genes by injecting DNA or RNA into embryos or introduction of blocking molecules, such as RNA encoding truncated proteins or antibodies, in order to alter the activity of endogenous gene products (Paul T. Sharpe and Mason 1999; Dee *et al.* 2008). The method involves microinjection of DNA or RNA molecules directly into the cytoplasm of 1-4 cell stage fertilized zebrafish embryos using a pressure microinjector and micromanipulator, as described in Section 2.4. DNA microinjection results into only a small fraction of the cells within the embryo inheriting the foreign DNA, because of the delay in integration and rapid cell division of the early zebrafish embryos. Hence, the expression of the transgene is highly mosaic and the germ-line transmission of the transgene has low efficiency. In the experiments described below microinjection of *in vitro*-synthetized capped RNA was performed in order to obtain transient (up to 3 days) and widespread ectopic overexpression of Sox3 (or its deleted forms)(Guille 1999). The aim of using this technique instead of DNA microinjection is to obtain rapid translation and more readily diffusion of the construct, avoiding mosaic overexpression which would have perturbed the analysis (Hyatt and Ekker 1999). In the following part of the study, microinjection of capped RNA was used to investigate the functional roles of Sox3 by comparing morphological and molecular changes in embryos overexpressing the wt protein versus mutant forms carrying deletions.

3.3.2. Over-expression of wt Sox3 in Early Zebrafish Embryos

Western blot analysis was performed to assess whether microinjection of capped RNA resulted in ectopic overexpression of protein in embryos at Sphere Stage (about 4 ½ hpf). For assessing overexpression of wt Sox3, 15 embryos were injected with 50 pg of *zfsox3-HA* RNA and dechorionated live at Sphere Stage (about 4 ½ hpf) together with 15 uninjected embryos at the same developmental stage (negative control), followed by SDS PAGE. Western blot was performed to detect the HA tag of the ectopic Sox3-HA (Fig. 3.8 A) and the membrane was subsequently re-blotted to detect endogenous α-actin as loading control (Fig. 3.8 B). The experiment lead to a clear detection of zfSox3-HA as well as α-actin, as demonstrated by the existence of strong bands of about 36 and 38 kDa, respectively, thus verifying successful ectopic expression of the foreign gene at Sphere Stage. HA was not detected in uninjected embryos since it is not endogenously expressed in zebrafish. The other bands that are visible in both samples were probably due to non-specific binding of the antibodies.

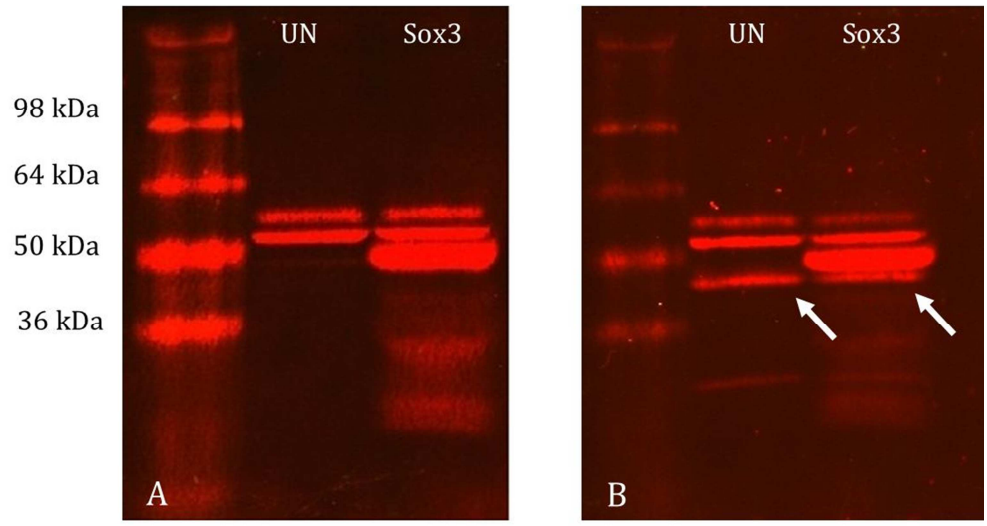


Figure 3.8 Western blot analysis of embryos uninjected (UN) or injected with *wt sox3-HA* (Sox3) at 1-4 cell stage and incubated until sphere stage. Figure A shows detection of HA: the two bands seen in both lanes correspond to non-specific signals, the third strong band in lane 3 corresponds to ectopic Sox3. Figure B shows an additional blotting of the same membrane for detecting α-actin as loading control (white arrows).

3.3.3. Over-expression of Sox3 Deletion Mutants in Early Zebrafish Embryos

Western blot analysis was performed on Sphere Stage embryos microinjected with the Sox3 deletion mutants in order to assess whether the efficiency of over-expression of the protein products was affected by the deletions.

An amount of 45 embryos were injected at 1-4 cell stage with each construct, including *wt sox3-HA* and *mutants 1* to *12* (all containing HA tag fused to the gene), and incubated until they reached Sphere Stage (approximately 4 ½ hpf), while 45 embryos were incubated without any prior injection to be used as negative control (UN). For each of the injection groups 30 embryos were selected, dechorionated live, treated as described in Section 2.4.1, then divided into two half and run on two separate SDS PAGE gels. The gels were then transferred on two membranes, which were incubated respectively with anti-α-actin primary antibodies or anti-HA primary antibodies. This approach was used in order to avoid any superposition between the bands corresponding to α-actin and Sox3, as it could have interfered with the clarity of the blot. The Western Blot for HA showed upper bands that were thought to be non-specific signals, as they were present in the negative control as well as in all the other samples (Fig. 3.10 A) and also in the previous western blot experiment (Fig. 3.8). The third lane of the blot, which correspond to the wtSox3-HA protein, presented another clear large band, which corresponds to the ectopic Sox3. All the lanes corresponding to the deletion mutants contained large bands corresponding to the recombinant proteins but of various sizes. At the same time another Western Blot was performed for each of the samples and incubated with primary antibody specific for α-actin (Fig. 3.10 B). This blot served as a loading control and confirmed that the amounts of embryos analysed were similar in all the samples.

Although the blot of the Sox3 deletion mutants was not perfect (there was insufficient time to repeat the experiment), the variation in the sizes of the bands corresponded well to the sizes of the deletions present in each of the constructs microinjected (Fig. 3.10 A, Table 3.2) and the amount of protein seemed quite similar in all the samples, except Mut12. Even though it is possible that the high intensity of the band corresponding to Mut12 was due to the small size of the construct, a higher expression of such construct was presumed not to be a problem since this construct was a negative control.

Importantly, this analysis proved that the effects of the over-expression of Sox3 mutants showed in the subsequent experiments were not due to the absence, or to the modified expression of the Sox3 deletion constructs.

Figure 3.9 Schematic representations of the deletion constructs that were analysed through western blot. The light blue bars indicate the wt gene and the blue bars represents the HMG domain. For each construct the deleted region is represented as a red bar numbered accordingly to the name of the construct (for example 1 indicates the region missing in Mut1).

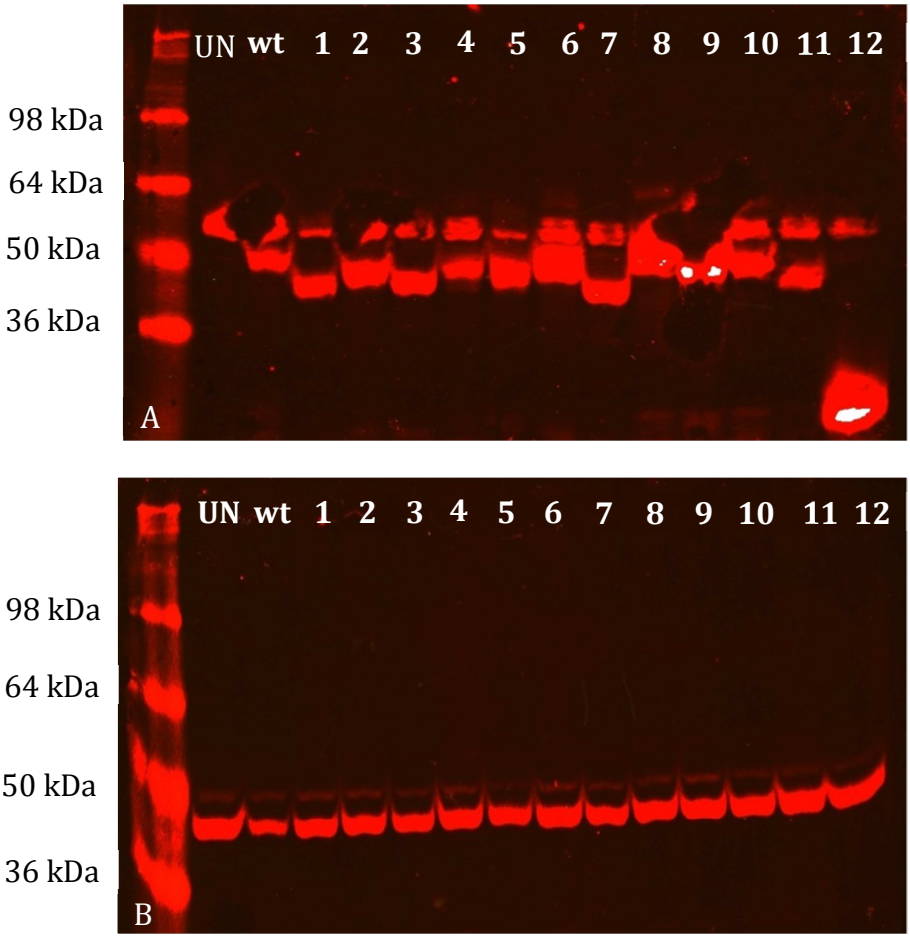


Figure 3.10 Western blot analysis of embryos uninjected (UN) or injected with different constructs at 1-4 cell stage and incubated until sphere stage. The embryos were kept uninjected (UN) or injected with respectively *wt sox3-HA* (wt), *mut1* (1), *mut2* (2), *mut3* (3), *mut4* (4), *mut5* (5), *mut6* (6), *mut7* (7), *mut8* (8), m*ut9* (9), *mut10* (10), *mut11* (11) or *mut12* (12). Figure A: detection of HA, the number of amino acids deleted in each construct is indicated above. Figure B: loading control, detection of endogenous α-actin.

3.4. Effects of the Over-expression of Sox3 Deletion Mutants on Neural Development

3.4.1. Introduction

The ectopic expression of wt Sox3 in zebrafish embryos causes a duplication of the Central Nervous System (CNS) as demonstrated by previous experiments (Dee *et al.* 2008). In order to investigate whether one or more of the regions that were deleted from the protein are involved in this phenomenon, and thus might play a role in Sox3 functioning in the determination of the neural fate, embryos were analysed 24 hpf after microinjection of RNA. In particular, the aim of this experiment was to dissect whether the over-expression of the deletion constructs caused different effects on the development of the nervous system, compared to the effects caused by over-expression of wt Sox3. In order to allow clear visualization of the CNS, the 24 hpf embryos were hybridized with *n-cadherin* riboprobes (*ncad*) by whole mount *in situ* hybridization.

Cadherins are a family of Ca^{2+}-dependent cell adhesion proteins that are essential during several steps of vertebrate embryo development (Takeichi 1988; Halbleib and Nelson 2006). Ncad, in particular, is a cadherin implicated in many aspects of development, including the formation of the neural tube and it is expressed throughout the developing nervous system (Kintner 1992; Klymkowsky *et al.* 2010). The expression of Ncad is quite ubiquitous during early vertebrate development, and only at later stages it becomes restricted to specific regions of the CNS. For this reason in our experiment we decided to use *ncad* as a marker for the visualization of the zebrafish developing CNS.

3.4.2. Ectopic Expression of wt zfSox3 Induces Duplication of the CNS

Firstly, experiments were carried out in order to determine the ideal conditions to see clearly the effects of over-expression of wt Sox3 (wtSox3), compared to uninjected (UN) and *GFP*-injected control embryos (*GFP*). The reason why a control was performed by injecting embryos with only *GFP*, was to confirm that the phenotypes were not affected by the injection procedure itself. Moreover, microinjection of *mut12* deletion construct was also carried out as negative control, since this mutant was expected to have lost Sox3 function.

The experiment was performed by injecting embryos with 50pg of CAP RNA as reported in the literature (Dee *et al.* 2008; Shih *et al.* 2010). The resulting phenotypes (Fig. 3.11) were divided into four categories: normal, when the nervous system was not affected by the experiment (Fig. 3.12 A, B, C), mild CNS duplication, when the nervous tissue showed local duplication or expansion to a range of extents (Fig. 3.12 D, E, F), severe CNS duplication, when the duplication involved most of the length of the CNS (Fig. 3.12 G, H, I), and major disruption of neural tissue, when the embryo appeared disrupted and in most cases very poorly developed and the nervous tissue was very disrupted (Fig. 3.12 L, M, N).

The results of the experiment were the following (Fig. 3.13 and Table 3.3):

1) Both the UN and the *GFP* controls were unaffected by the procedure, as they showed 100% normal embryos;
2) Microinjection of *wt sox3* severely affected the development of the CNS, as 25% of the embryos showed mild CNS duplication, 71% severe CNS duplication and 3% major disruption of the neural tissue;
3) Microinjection of *mut12* did not affect the development of the CNS, as 100% of the embryos had a normal phenotype, suggesting a total loss-of-function of wt Sox3.

Since these results showed a clear difference between the phenotypes of the controls (UN and GFP) and the phenotypes caused by injection of *wt sox3*, all the subsequent experiments were repeated performing injections of 50pg of CAP RNA and whole mount *in situ* hybridization with N-cadherin probes. Moreover, the injection of *mut12* as a control of the loss-of-function of wt Sox3 was also repeated.

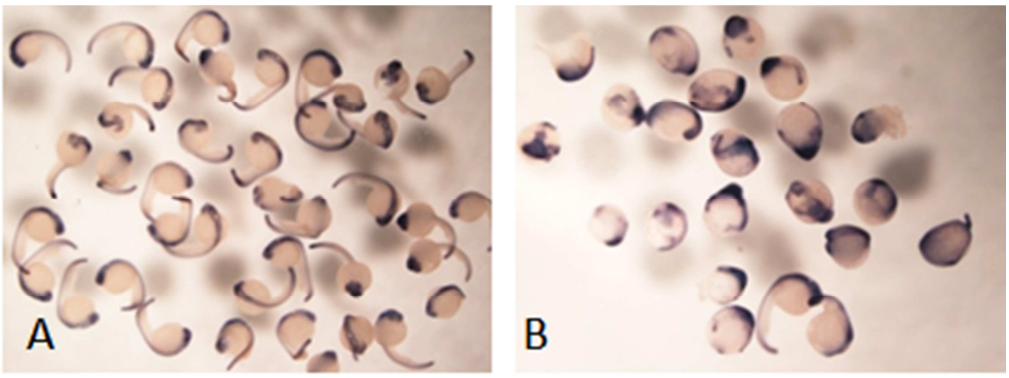


Figure 3.11 Embryos uninjected (A) or injected with *wt sox3* (B), fixed 24 hpf and analysed through whole mount *in situ* hybridization for *ncad* (a neural marker). It is clearly visible how the over-expression of Sox3 affected the development of the nervous system as the neural tissue appeared much more disrupted in the majority of the embryos shown in picture B compared to picture A.

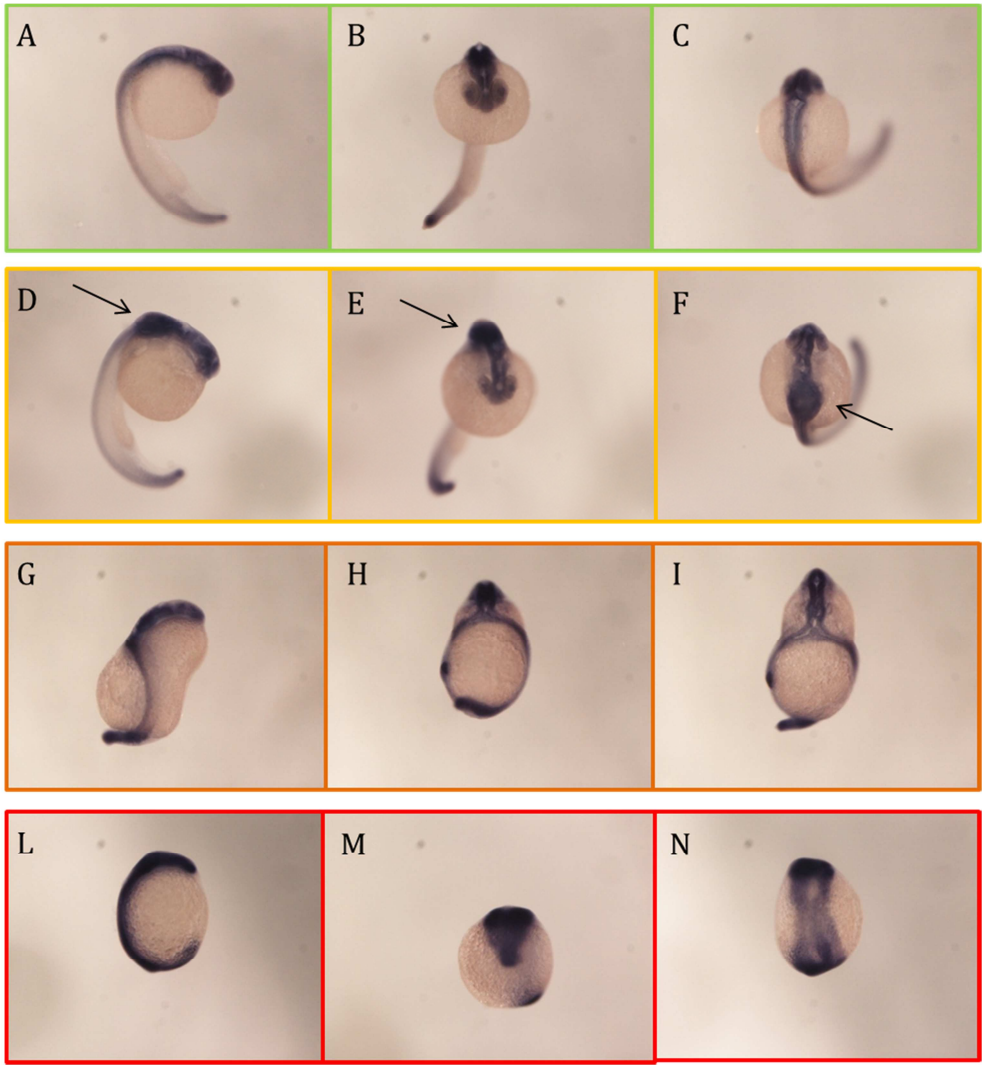


Figure 3.12 Categories of the phenotypes obtained with whole mount *in situ* hybridisation for *ncad* on 24 hpf embryos uninjected or injected with *wt/mut sox3*. The four categories are: normal (A, B, C), mild CNS duplication (black arrows in figures D, E, F), severe CNS duplication (G, H, I) and major disruption of the neural tissue (L, M, N). Each embryo is shown from the lateral side (A, D, G, L), from the frontal side (B, E, M) and from the dorsal side (C, F, I, N). In picture H the embryo is shown from the dorsal side but slightly turned in order to visualize entirely the duplication of the CNS.

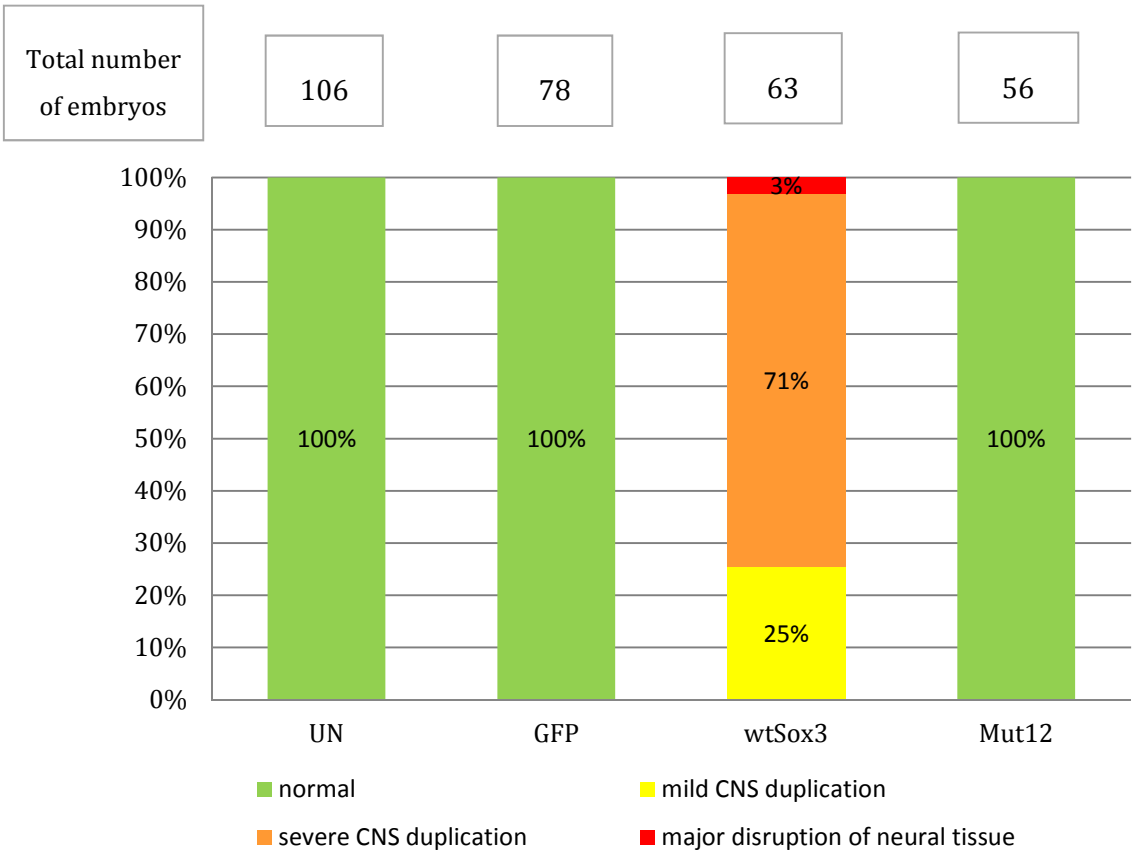


Figure 3.13 Results of whole mount *in situ* hybridisation for *ncad* performed on 24 hpf embryos respectively uninjected (UN), injected with *GFP*, *wt sox3* or *mut12*.

	normal	mild CNS duplication	severe CNS duplication	major disruption of neural tissue	tot
UN	106	0	0	0	**106**
GFP	78	0	0	0	**78**
wt sox3	0	16	45	2	**63**
mut 12	56	0	0	0	**56**

Table 3.3 Numbers of 24 hpf embryos analysed with whole mount *in situ* hybridisation for *ncad*.
The data shown are the raw numbers that correspond to the percentages shown in figure 3.13.

3.4.3. Analysis of Embryos 24 hpf After Microinjection of Sox3 Deletion Mutants

The experiments were carried out in order to analyse whether the over-expression of the Sox3 deletion mutants caused different phenotypes in embryos at 24 hpf compared to those caused by over-expression of wt Sox3 (*wt Sox3)*. Two experimental controls were performed: a group of uninjected embryos (UN) were treated with the same procedures as all the other groups, and a group of embryos were injected with 50pg *GFP* CAP RNA (*GFP*). The effects caused by all the Sox3 deletion mutants were analysed in two separate sets of experiments, as the use of a high number of samples for the same experiment could have caused delays in the experimental procedures (especially in the whole mount *in situ* hybridisation protocol) and therefore could have affected the results. Both the UN and *GFP* control were repeated in the two experiments as microinjection and *in situ* hybridization can show some variability between different experiment and this would have rendered the comparison between wt and mutants Sox3 less reliable. The resulting phenotypes were assigned to four categories (Fig. 3.14A and Fig. 3.14B): normal, when the nervous system was not affected by the experiment; mild CNS duplication, when the nervous tissue showed to be locally duplicated or expanded in a range of extents; severe CNS duplication, when the duplication involved the CNS in all its length; and major CNS disruption, when the embryo appeared disrupted and in most of the cases not developed and the nervous tissue was completely disrupted. Moreover, it has been observed that almost half of the embryos microinjected with *mut10* presented an additional unique phenotype (Fig. 3.15). Because these embryos were characterised by a dramatic disruption of the nervous tissue, they were considered to belong to the "major disruption of the neural tissue" phenotype. However, it is important to note that this phenotype differed from the others in the same category.

In the first set of experiments, uninjected embryos (UN), embryos injected with only *GFP* and embryos injected with *GFP* and *wt sox3* were compared to embryos injected with *GFP* together with *mut2*, *mut3*, *mut4*, *mut5*, *mut6*, *mut7* and *mut12* (Fig. 3.16 ,Table 3.4). The results were as follows:

1) Both UN and *GFP* controls exhibited 100% normal phenotypes embryos.
2) Microinjection of *wt sox3* resulted in only 3% normal phenotypes, while 23% of the embryos presented mild CNS duplication, 55% severe CNS duplication and 20% major CNS disruption.
3) Microinjection of *mut2* seemed to cause slightly less disruption of the CNS, as 21% of the embryos were normal, 67% had mild CNS duplication, 10% severe CNS duplication and 1% major CNS disruption. This suggests mild loss of Sox3 function.
4) The embryos microinjected with *mut3* showed 20% mild CNS duplication, 57% severe CNS duplication and 23% major CNS disruption. These data are comparable to those obtained for the wt and suggest that Mut3 did not present loss of function.
5) Microinjection of *mut4*, *mut5* and *mut6* caused phenotypes very similar to microinjection of *wt sox3*. The total amount of embryos that were affected was almost the same in all these groups compared to *wt sox3*, as they presented respectively only 14%, 4% and 8% of embryos with normal phenotypes (versus the 3% presented in *wt sox3* injected embryos). However, they did show a higher amount of mild phenoypes than *wt sox3* (respectively 79%, 70% and 81% of mild CNS duplication, versus the 23% of *wt sox3*) and a lower amount of severe phenotypes (respectively 3%, 11% and 0% of severe CNS duplication versus the 55% of *wt sox3* and 3%, 15% and 12% of major disruption of neural tissue versus the 20% of *wt sox3*). This suggests mild loss of Sox3 function.
6) Microinjection of *mut7* caused effects comparable to those of *wt sox3*, presenting 2% of embryos with normal phenotypes, 32% with mild CNS duplication, 59% with severe CNS duplication and 7%

with major neural tissue disruption. This suggests no loss of Sox3 function.

7) Finally, as expected, microinjection of *mut12* did not affect the development of the nervous system as its over-expression caused normal phenotype in 100% of the embryos.

In the second set of experiments UN and *GFP*-injection controls were repeated, as well as injections of *mut12* and *wt sox3*, whose effects were compared to the effects of microinjecting *mut1*, *mut8*, *mut9*, *mut10* and *mut11* (Fig. 3.17 ,Table 3.5):

1) Both UN embryos and embryos microinjected with *GFP* showed 100% normal phenotypes;
2) Microinjection of *wt sox3* caused 100% severe CNS duplication phenotypes;
3) Interestingly, embryos microinjected with *mut1* presented mostly normal phenotypes (58%) in addition with 37% of mild CNS duplication, only 3% of severe CNS duplication and no major disruption of neural tissue;
4) Microinjection of *mut8* and *mut9* did cause milder phenotypes as compared to *wt sox3*. These groups presented the majority of embryos with mild CNS duplication (respectively 63% and 75%), only a few with normal phenotypes for *mut8* (9%) and none for *mut9*, in addition to 17% with severe CNS duplication in both *mut8* and *mut9*, and respectively 11% and 8% with major CNS disruption.
5) Microinjection of *mut10* caused mild CNS duplication in 15% of the embryos, severe CNS duplication in 36% of the embryos and major CNS disruption in 48% of the embryos. These last data are indicated in figure 3.17 and in Table 3.3 with a different colour (violet) than the other "major neural tissue disruption" data (red) as although they belonged to the same category (the embryos were, in most of the cases, not developed and the nervous tissue was completely disrupted), their appearance was distinctly different,

resembling the early bud stage, when neural plate forms. This topic is further discussed in Section 4.2.

6) Microinjection of *mut11* resulted in no normal phenotypes, mild CNS duplication in 32% of the embryos, severe CNS duplication in 48% of the embryos and major CNS disruption in 20% of the embryos.
7) Microinjection of *mut12* confirmed the data obtained with the previous experiments, causing 100% normal phenotypes.

The experiments described above showed that microinjection of Sox3 deletion mutants had different effects in 24 hpf embryos compared to microinjection of *wt Sox3*. Therefore, the deletions did modify the function of the protein. The deletion of such large regions of the protein could have disrupted its functionality dramatically, and in this case the injections would have caused the same phenotypes in all the groups of injected embryos. However, the data presented above demonstrated that microinjection of different deletion mutants caused different phenotypes, showing that the deletion approach adopted was effective to identify different functions of Sox3.

However, the analysis of the phenotypes that the embryos presented at 24 hpf was insufficient to determine which were the different effects caused by each deletion, as many other regulatory pathways are involved in the formation of the neural tissue at this stage. Moreover, the functional redundancy showed by the other SoxB1 factors can interfere with the analysis at such a late stage of development (Okuda *et al.* 2006; Miyagi *et al.* 2009; Shih *et al.* 2010). For these reasons, in order to investigate how the deletions changed Sox3 activity, further experiments were carried out at earlier stage of development, when Sox3 is the only SoxB1 factor expressed and it is possible to analyse its function more directly.

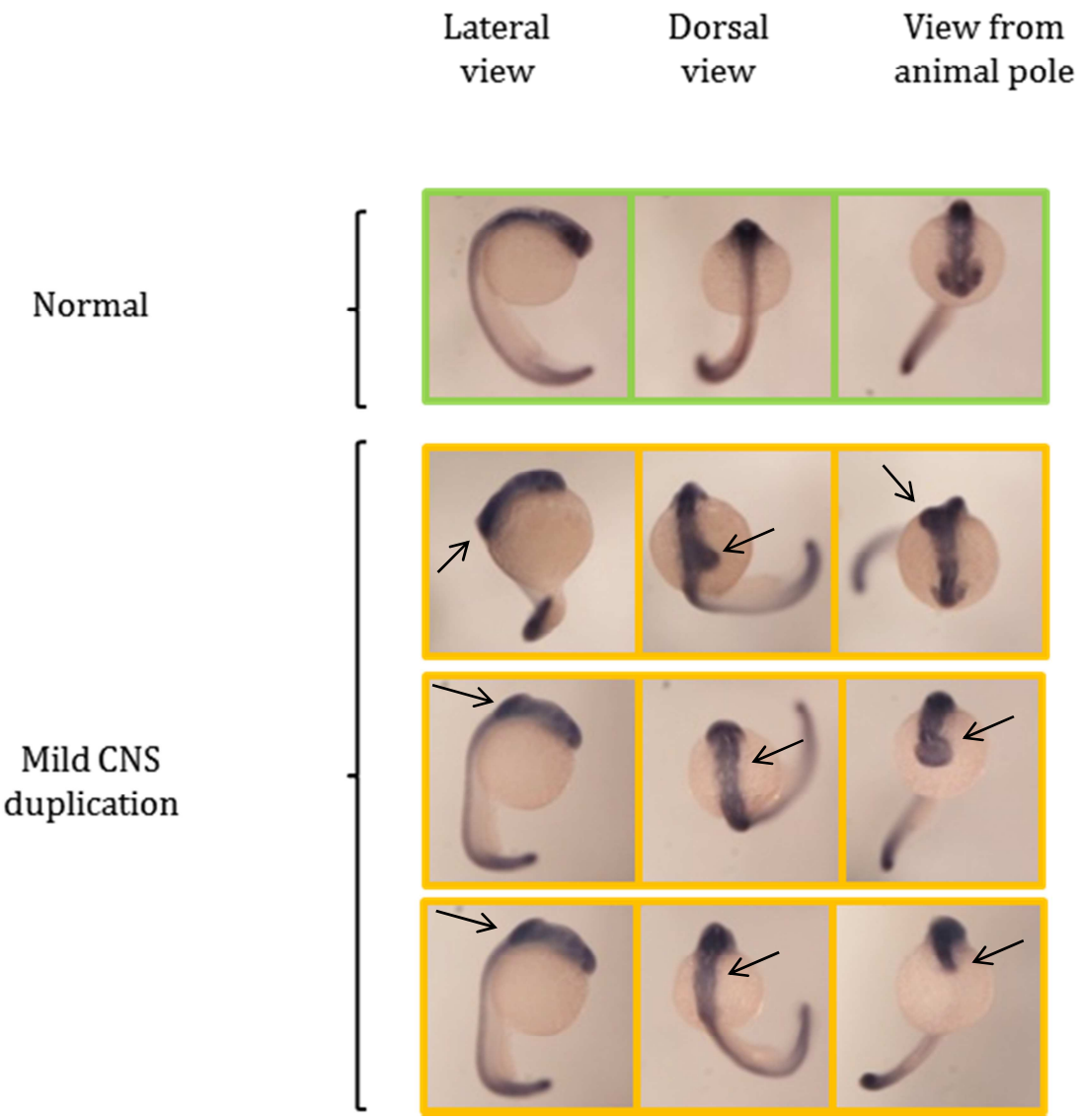


Figure 3.14A "Normal" and "mild CNS duplication" phenotypes obtained through whole mount *in situ* hybridization for *ncad* on 24 hpf embryos uninjected or injected with *GFP* or/and *wt/mut sox3*. Each embryo is shown from the lateral side, from the dorsal side and from the animal pole side. Black arrows indicate localised duplications or abnormalities of the CNS.

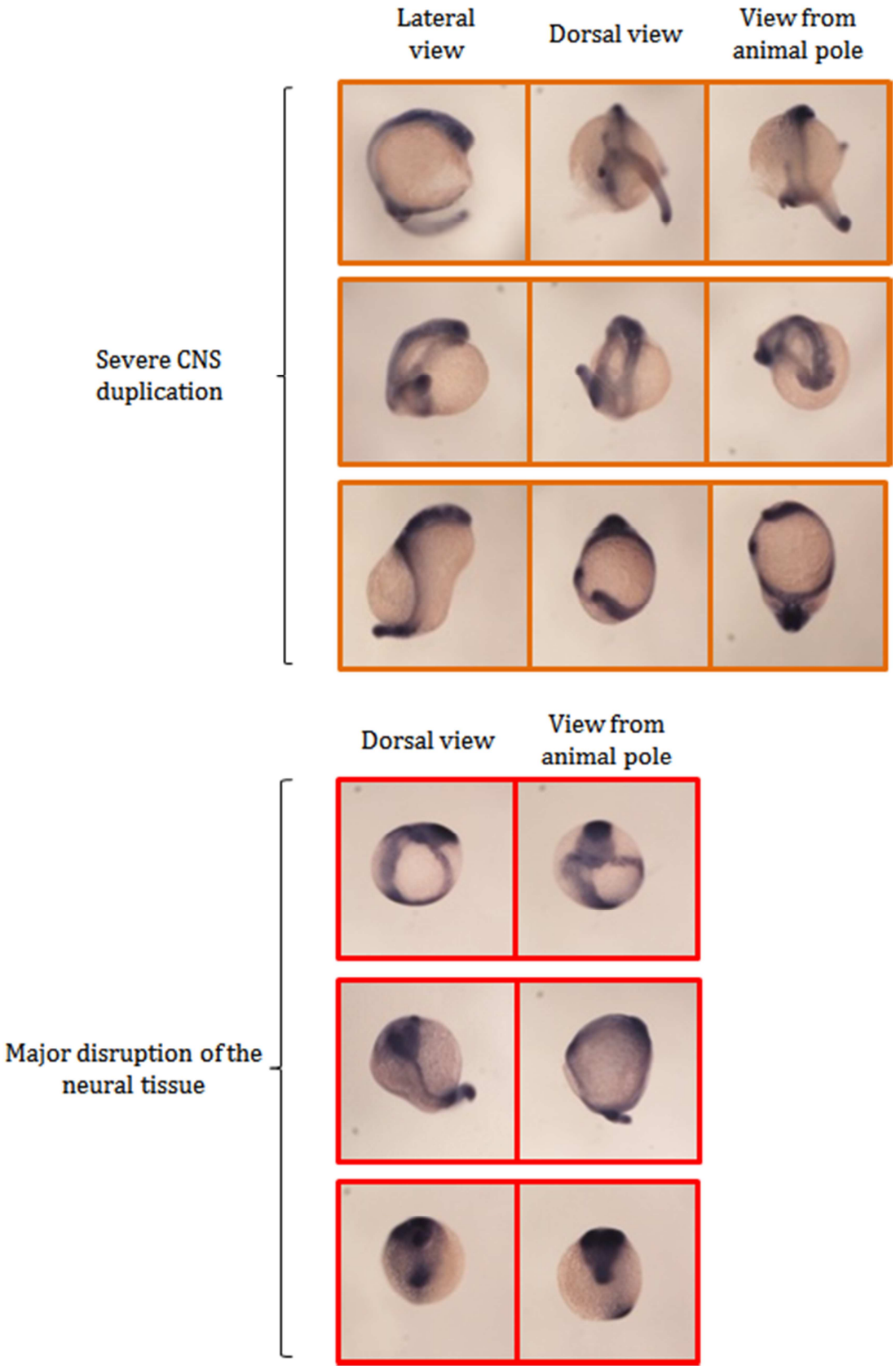


Figure 3.14B "Severe CNS duplication" and "major disruption of the neural tissue" phenotypes obtained through whole mount *in situ* hybridization for *ncad* on 24 hpf embryos uninjected or injected with *GFP* or/and *wt*/*mut sox3*. Each of the embryos with severe CNS duplication is shown is from the lateral side, from the dorsal side and from the animal pole side. Each of the embryos with major disruption of the neural tissue is shown from the dorsal view and from the animal pole.

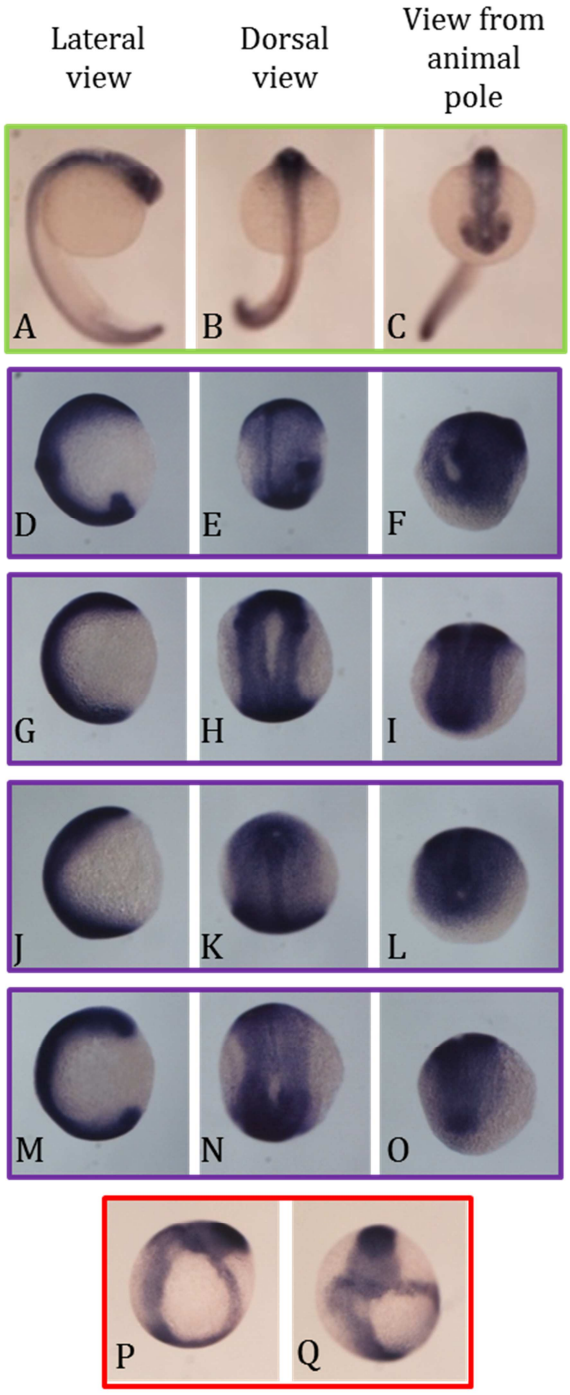


Figure 3.15 Embryos 24 hpf that were microinjected with *mut10* presented a unique phenotype that involved major disruption of the neural tissue but was different from the phenotypes caused by injection of the other constructs. The top panel shows a normal 24 hpf embryo (A, B, C), the following four panels show different embryos that were microinjected with *mut10* (pictures from D to O) and the last panel shows an embryo that was injected with *wt sox3* and was characterised by major disruption of the neural tissue (P, Q).

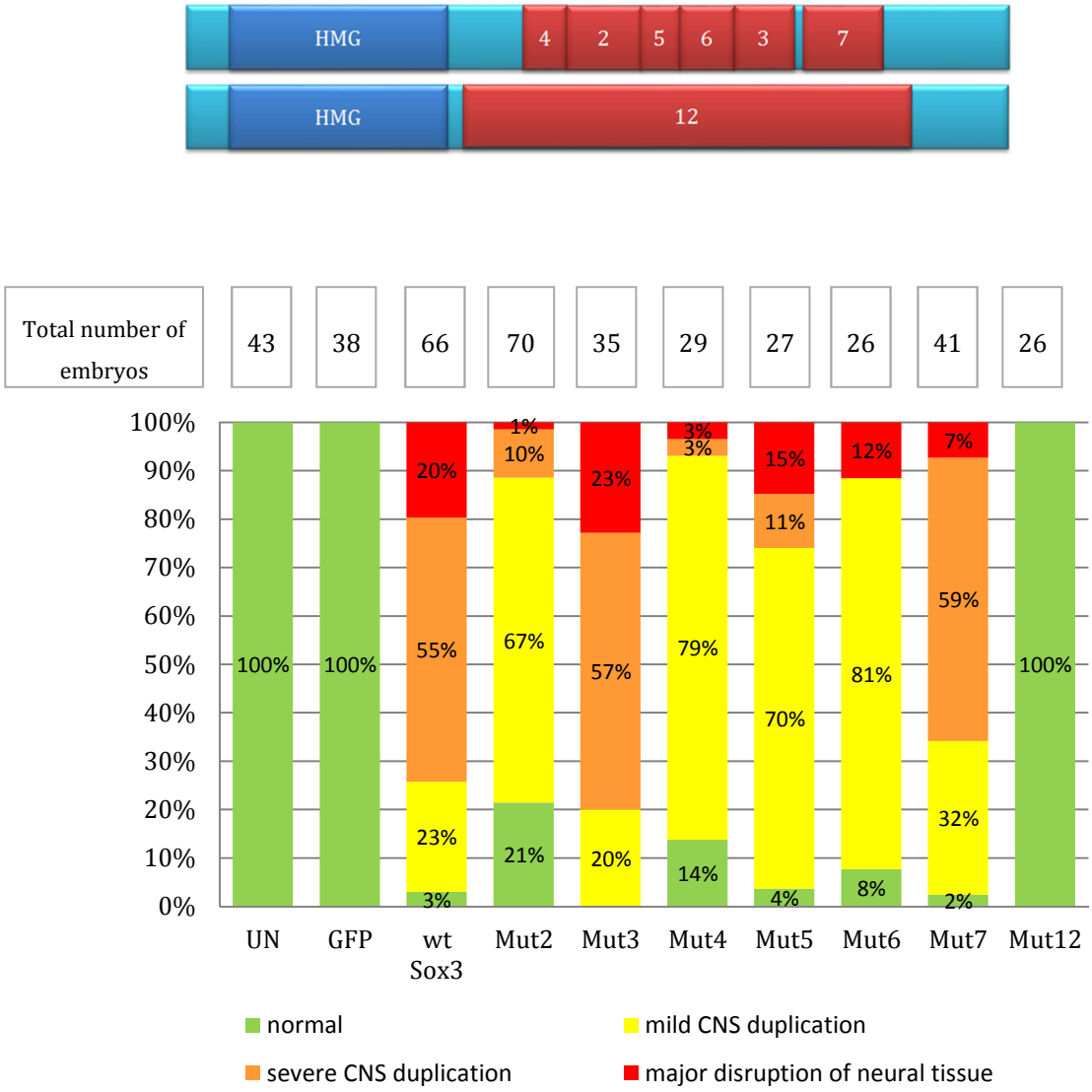


Figure 3.16 The top of the figure shows a schematic representation of the deletion constructs: the regions deleted in each of the mutants is represented as a red bar numbered accordingly to the name of the mutant. The chart shows the results of whole mount *in situ* hybridisation for *ncad* performed on 24 hpf embryos uninjected (UN), injected with *GFP* (*GFP*), co-injected with *GFP* and *wt sox3* or co-injected with *GFP* and *mut2*, *mut3*, *mut4*, *mut5*, *mut6*, *mut7* or *mut12*.

	normal	mild CNS duplication	severe CNS duplication	major disruption of neural tissue	tot
UN	43	0	0	0	**43**
GFP	38	0	0	0	**38**
wtSox3	2	15	36	13	**66**
Mut2	15	47	7	1	**70**
Mut3	0	7	20	8	**35**
Mut4	4	23	1	1	**29**
Mut5	1	19	3	4	**27**
Mut6	2	21	0	3	**26**
Mut7	1	13	24	3	**41**
Mut12	26	0	0	0	**26**

Table 3.4 Numbers of 24 hpf embryos analysed with whole mount *in situ* hybridisation for *ncad* in the second set of experiments. The data shown correspond to the percentages shown in figure 3.16.

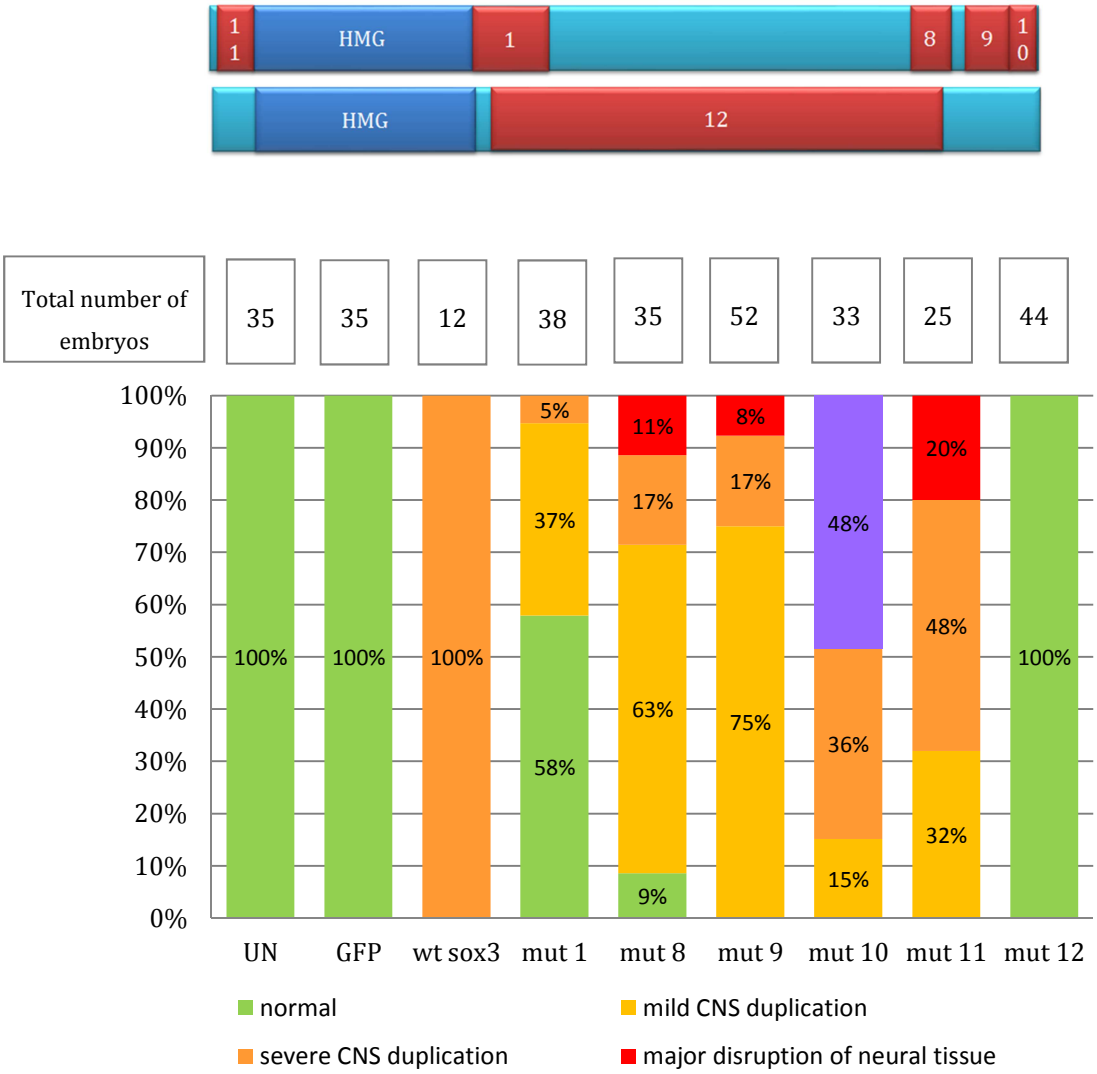


Figure 3.17 The top of the figure shows a schematic representation of the deletion constructs: the regions deleted in each of the mutants is represented as a red bar numbered accordingly to the name of the mutant. The chart shows the results of whole mount *in situ* hybridisation for *ncad* performed on 24 hpf embryos respectively uninjected (UN)or injected with *GFP* (*GFP*)or co-injected with *GFP* and *wt sox3* or *Mut1*, *Mut8*, *Mut9*, *Mut10*, *Mut11*, *Mut12*. The data shown in purple indicates the unique phenotype presented by embryos injected with *mut10* (Fig. 3.15).

	normal	mild CNS duplication	severe CNS duplication	major disruption of neural tissue	tot
UN	35	0	0	0	**35**
GFP	35	0	0	0	**35**
wtSox3	0	0	12	0	**12**
1	22	14	2	0	**38**
8	3	22	6	4	**35**
9	0	39	9	4	**52**
10	0	5	12	16	**33**
11	0	8	12	5	**25**
12	44	0	0	0	**44**

Table 3.5 Numbers of 24 hpf embryos analysed with whole mount *in situ* hybridisation for *ncad* in the first set of experiments. The data shown correspond to the percentages shown in figure 3.17. The data shown in purple indicates the unique phenotype presented by embryos injected with *mut10* (Fig. 3.15).

3.5. Effects of the Over-expression of Sox3 Deletion Mutants on Organizer Formation

3.5.1. Introduction

In order to get deeper insight into the relation between structure and function of Sox3, and in particular to dissect how deletions affected Sox3 functions, the following experiments focused on the early stages of the development, where the regulatory mechanisms are better known and the phenotypes are likely to be a more direct result of the over-expression of the protein. In particular, we analysed the function of Sox3 in the context of the organizer formation, at 4.5 hpf. Therefore, 1-4 cell stage embryos were microinjected with CAP RNA, incubated for 4.5 hpf, then fixed and subsequently analysed with whole mount *in situ* hybridization. *In situ* hybridization was performed in order to detect the expression of the organizer markers *bozozok* and *goosecoid*;

- *bozozok* (also known as *dharma*) is a homeobox gene that in zebrafish is essential for the formation and/or the induction of the Nieuwkoop center and, therefore, for the subsequent formation of the organizer (Ryu *et al.* 2001). In the developing zebrafish embryo, prior to axis formation, there is accumulation of β-catenin in the nuclei situated in that part of the yolk syncytial layer that lies beneath the cells that will later become the organizer (Schneider *et al.* 1996). In the late blastula nuclear localization of β-catenin activates organizer genes such as *bozozok* (*boz*) and *squint*. This homeodomain protein is a transcription factor that works in many different ways: firstly, it can repress *BMP* and *wnt* genes, which promote ventralization, secondly it suppresses the inhibitor of transcription vega1, allowing the activation of the organizer genes, and thirdly it acts with Squint in the activation of *goosecoid* (*gsc*), *noggin* and *dickkopf* (Gritsman *et al.* 2000; Kawahara *et al.* 2000; Solnica-Krezel and Driever 2001).

- *Goosecoid* is a homeobox gene and it is one of the earliest markers expressed during organizer formation (Cho *et al.* 1991; Stachel *et al.* 1993; Toyama *et al.* 1995). *Gsc* blocks the Bone Morphogenetic Proteins (BMPs) and Wnts allowing the formation of the neural ectoderm and the dorsal mesoderm (De Roberts *et al.* 1992; Yasuo and Lemaire 2001).

Ectopic expression of *wt sox3* obtained through capped RNA microinjection has been shown to repress organizer formation, as demonstrated in previous studies from our laboratory (Shih *et al.* 2010; Kuo *et al.* 2013). In the following experiments the effects caused by ectopic over-expression of *sox3* deletion mutants were investigated through the analysis of the expression of *boz* and *gsc in vivo*.

3.5.2. Ectopic Over-expression of wt zfSox3 Represses Organizer Markers

In order to optimise the experimental conditions for the comparison between the effects of ectopic expression of wt Sox3 and its deletion mutants, zebrafish embryos uninjected (UN) or injected with wt Sox3 RNA were incubated until the sphere stage and fixed with 4% paraformaldehyde. The embryos were then analysed through whole mount *in situ* hybridization in order to test the efficiency of newly synthetized DIG-labelled oligo-probes for *bozozok* and *goosecoid*. Moreover, the experiment was aimed to test whether that the experimental conditions allowed the visualisation of a clear signal that would have subsequently be essential for comparing with wtSox3 over-expression. All the uninjected embryos showed a strong purple staining localised in the region of the organizer for both *bozozok* and *goosecoid* probes (Fig. 3.18). The signal was strong enough to see clearly the difference between the uninjected embryos and the embryos injected with wt Sox3 (Fig. 3.19), and presumably to determine whether the deletion mutant constructs would have caused changes in the signal compared to the wt protein. For this reason the subsequent experiments were

performed under the same experimental conditions: microinjection of 50 pg of RNA, incubation of the embryos until sphere stage and whole mount *in situ* hybridization for *boz* or *gsc*.

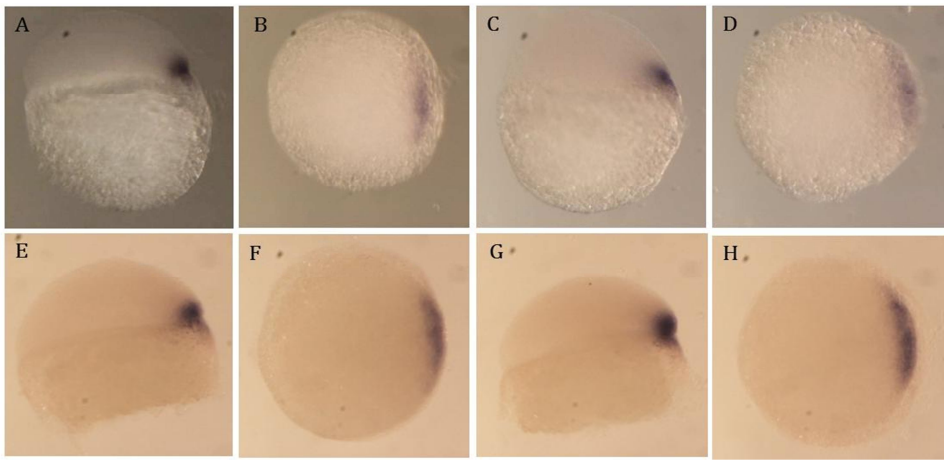


Figure 3.18 Whole mount *in situ* hybridization of sphere stage embryos with oligo-probes for *bozozok* (A, B, C, D) and *goosecoid* (E, F, G, H). Each embryo is shown from the lateral view (A, C, E, G) and from the animal pole view (B, D, F, H). The dark staining developed in the organiser region, which corresponds to the region where *boz* and *gsc* are expressed, is localised and clearly visible.

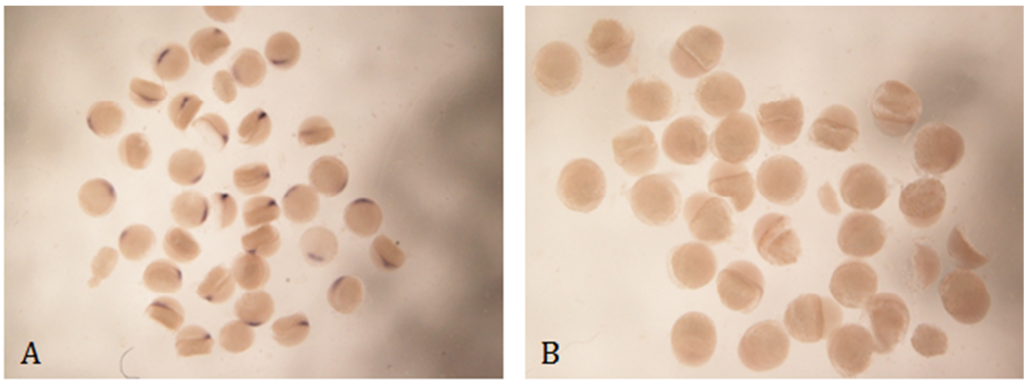


Figure 3.19 Sphere stage embryos uninjected (A) and microinjected with *wt sox3* (B) analysed through whole mount *in situ* hybridization for *boz*. While the majority of the embryos shown in picture A present a strong signal, the embryos shown in picture B present just faint or totally absent signal.

3.5.3. Expression of *bozozok* in Embryos Microinjected with Sox3 Deletion Mutants

In the following experiments zebrafish embryos at 1-4 cell stage were microinjected with RNA encoding GFP, wt Sox3 or each of deletion mutant Sox3 proteins, incubated at 28˚C until they reached Sphere Stage and analysed with whole mount *in situ* hybridization using *bozozok* DIG-labelled riboprobes. The same experiment was performed on uninjected embryos as negative control. The embryos injected with only *GFP* were used as negative control to assess that the injection procedure did not affect the analysis. All the other groups of embryos were co-injected with the desired *sox3* construct together with *GFP*: this allowed verification that the injections were effective by visualising the fluorescence of the expressed GFP under UV light after the incubation of the embryos. As for the analysis of 24 hpf embryos, two separate sets of experiments were performed to analyse all the constructs; each one of the experiments was performed together with the uninjected (UN) and the *GFP* (GFP) controls, as well as the *wt Sox3* (wtSox3).

Normal expression levels of *boz* were expected in uninjected and *GFP*-injected embryos, while lower expression was expected in *wt Sox3*-injected embryos. These expectations were indeed confirmed in both the experiments, where the resulting phenotypes were categorised in three groups: normal expression, when the staining developed was dark purple and clearly visible especially from the lateral view (Fig. 3.20 E, F), partial repression, when the staining was lighter than the normal phenotype from the lateral view and barely visible from any other side of the embryo (Fig. 3.20 C, D), and total repression, where the staining was not detectable from any side of the embryo, which looked completely white (Fig. 3.20 A, B). Moreover, an additional phenotype (ectopic expression of *boz*) was identified only in embryos co-injected with *GFP* and *mut1* (Fig. 3.21).

In the first set of experiments *boz* expression levels found in UN, *GFP*-injected and *wt sox3*-injected embryos were compared to the expression

caused by microinjections of *mut2, mut3, mut4, mut5, mut6, mut7* and *mut12* (Fig. 3.22, Table 3.6). The results were as follows:

1) The data confirmed that ectopic expression of wt Sox3 causes a decrease in *boz* expression levels when compared to the controls, as 98% of UN embryos showed normal expression of *boz* as well as 94% of *GFP* embryos, while only 20% of *wt sox3*-injected embryos showed normal expression and the other 80% showed partial repression (40%) or total repression (40%).
2) The comparison between injection of *wt sox3* and injections of Sox3 deletion mutants showed that the level of repression caused by *mut2*, *mut3* and *mut4* is lower than the *wt sox3* as the percentages of normal embryos are higher, the percentages of partial repression comparable, and the percentages of total repression are lower. However, *mut2*, *mut3* and *mut4* did show a repressive effect on *boz* expression as demonstrated by the comparison with UN and *GFP* controls.
3) Conversely, it appeared that injections of *mut5* and *mut6* caused a substantially milder repressive effect on *boz* expression, showing almost as many normal embryos as the controls, low percentages of partial repression (respectively 8% and 16%) and no total repression.
4) Microinjection of *mut7* resulted in a percentage of normal embryos comparable to *mut2*, *mut3* and *mut4*, which is lower than UN and *GFP*, but higher than *wt sox3*, however there was no embryos showing partial repression and 42% showing total repression.
5) As expected, injection of *mut12* caused a percentage of normal embryos comparable to the UN and *GPF* (88% compared to 98% and 94%) together with a few embryos showing partial repression (13%).

In the second set of experiments the UN and *GFP* controls were repeated, as well as the *mut12* control, which was expected to have lost Sox3 repressor function. In these experiments the levels of expression of *boz*

were compared between the controls, embryos microinjected with *GFP* and *wt sox3* and embryos injected with *GFP* and *mut1*, *mut8*, *mut9*, *mut10*, *mut11* and *mut12* (Fig. 3.23, Table 3.7). The results were as follows:

1) Both UN embryos and embryos injected with *GFP* showed 100% normal phenotypes, while embryos injected with *wt sox3* showed only 11% normal phenotypes and the other 89% of embryos caused partial repression (49%) or total repression (40%). This is consistent with the results of the first experiment.
2) Also the effects caused by injecting *mut12* were consistent with previous data as comparable to the UN and *GFP* controls, showing 97% normal embryos and 3% partial repression.
3) Interestingly, injection of *mut1* caused a unique phenotype in 35% of the embryos. This phenotype was not found in any other sample and was characterised by ectopic expression of *boz* outside the organizer region (Fig. 3.20). Moreover, the remaining 65% of *mut1*-injected embryos showed a normal phenotype a none showed repression of *boz*.
4) Microinjections of *mut8*, *mut9* and *mut10* had milder repressive effects than microinjections of *wt sox3*, showing higher percentages of normal embryos and lower percentages of embryos showing partial and total repression of *boz*.
5) Microinjection of *mut11* caused a lower repression of *boz* compared to *wt Sox3*: 86% of the embryos had a normal phenotype, while only 8% partial repression and 6% total repression.

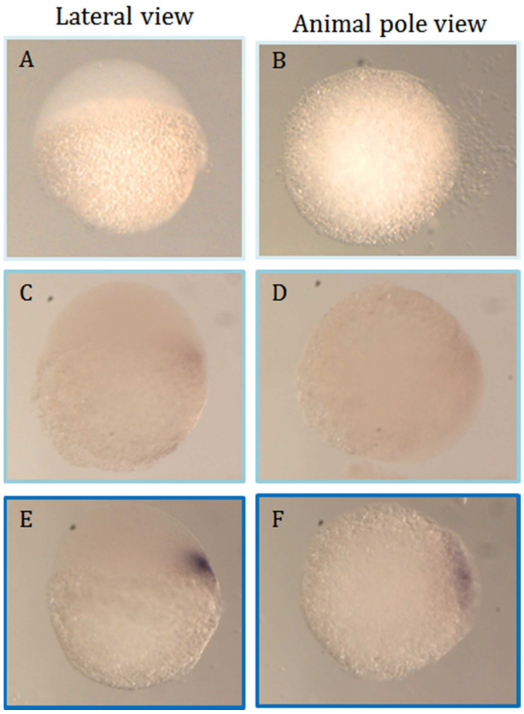


Figure 3.20 Pictures representing the three categories of phenotypes obtained through whole mount *in situ* hybridization for *boz* on sphere stage embryos uninjected or injected with *GFP* or/and *wt*/*mut sox3*. The phenotypes include: total repression of *boz* (A, B), partial repression of *boz* (C, D) and normal expression of *boz* (E, F).

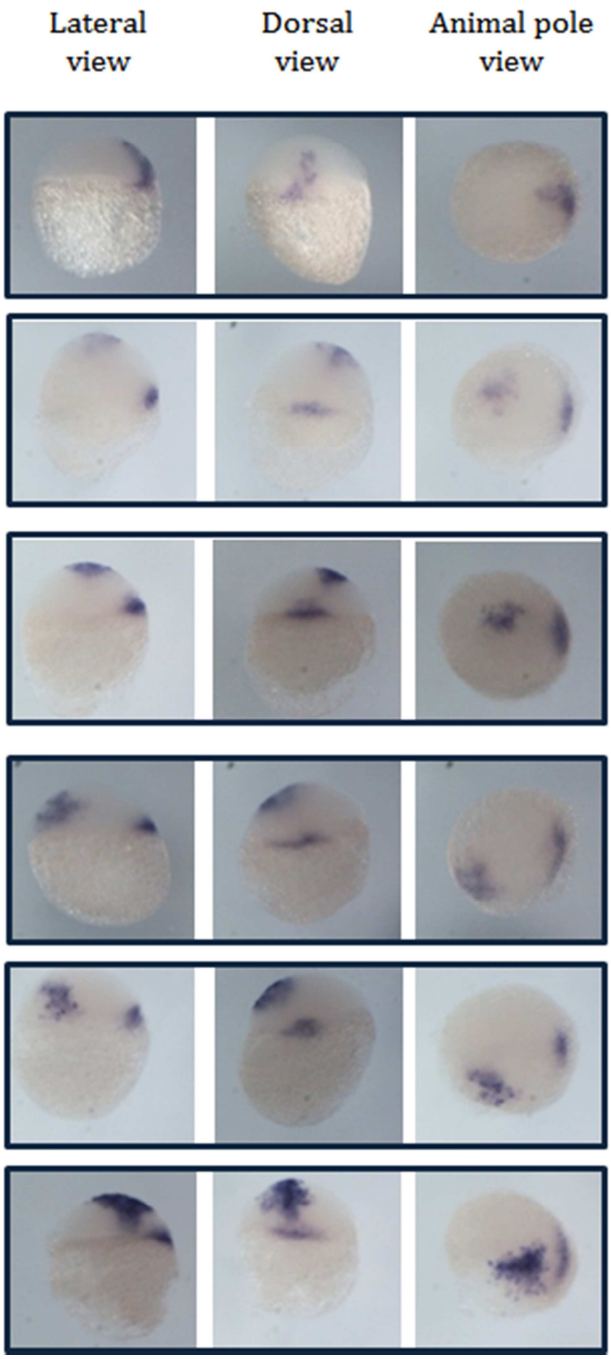


Figure 3.21 Microinjection of *mut1* variable ectopic expression of *bozozok* in sphere stage embryos. Each panel shows the same embryo viewed from different sides. The phenotype of these embryos is characterised by variable ectopic expression of *boz* outside the organizer region.

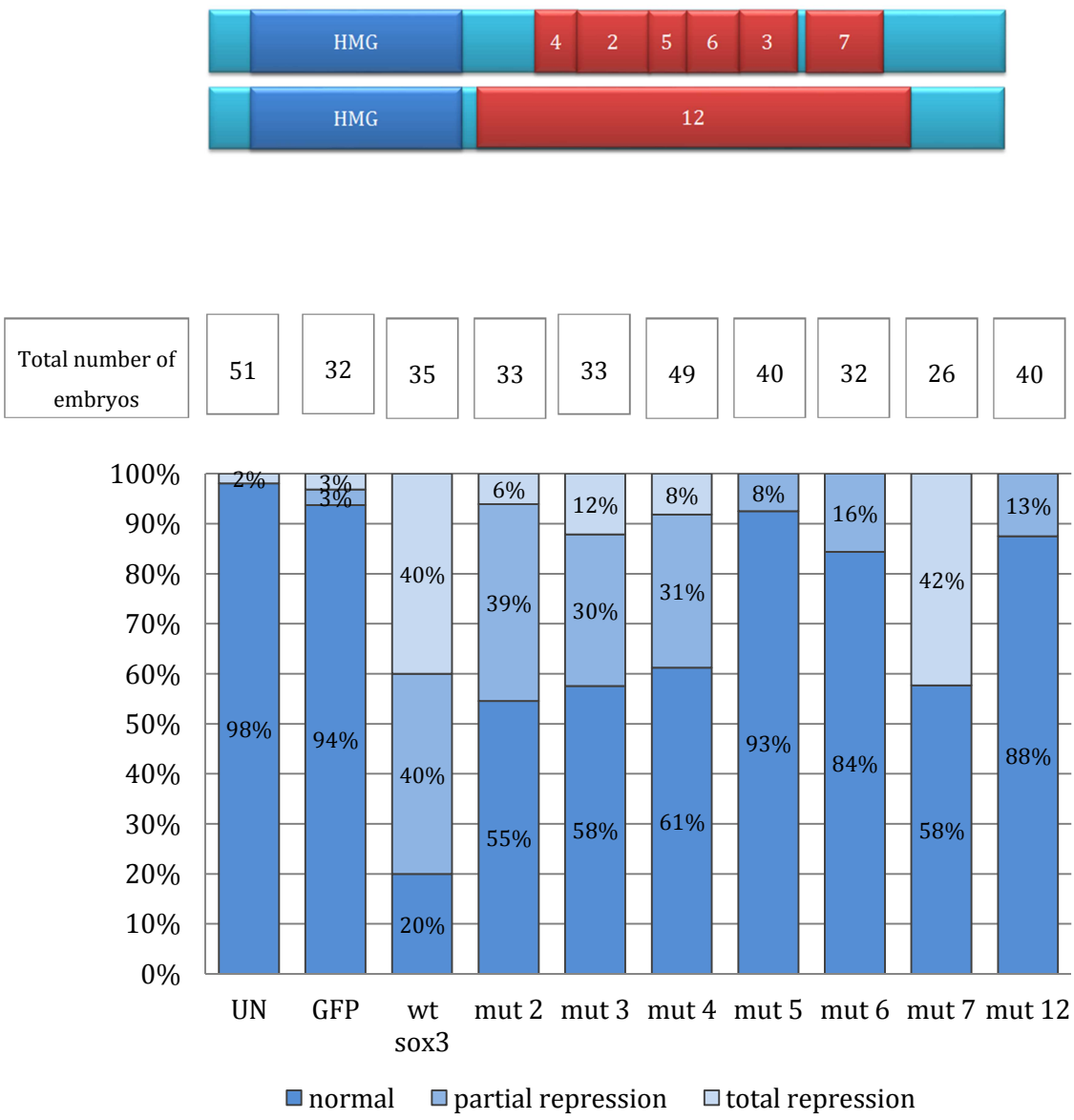


Figure 3.22 The top of the figure shows a schematic representation of the deletion constructs: the regions deleted in each of the mutants is represented as a red bar numbered accordingly to the name of the mutant. The chart shows the results of whole mount *in situ* hybridisation for *boz* performed on Sphere Stage embryos respectively uninjected (UN)or injected with *GFP* (*GFP*)or co-injected with *GFP* and *wt sox3* or *mut2*, *mut3*, *mut4*, *mut5*, *mut6*, *mut7*, *mut12*.

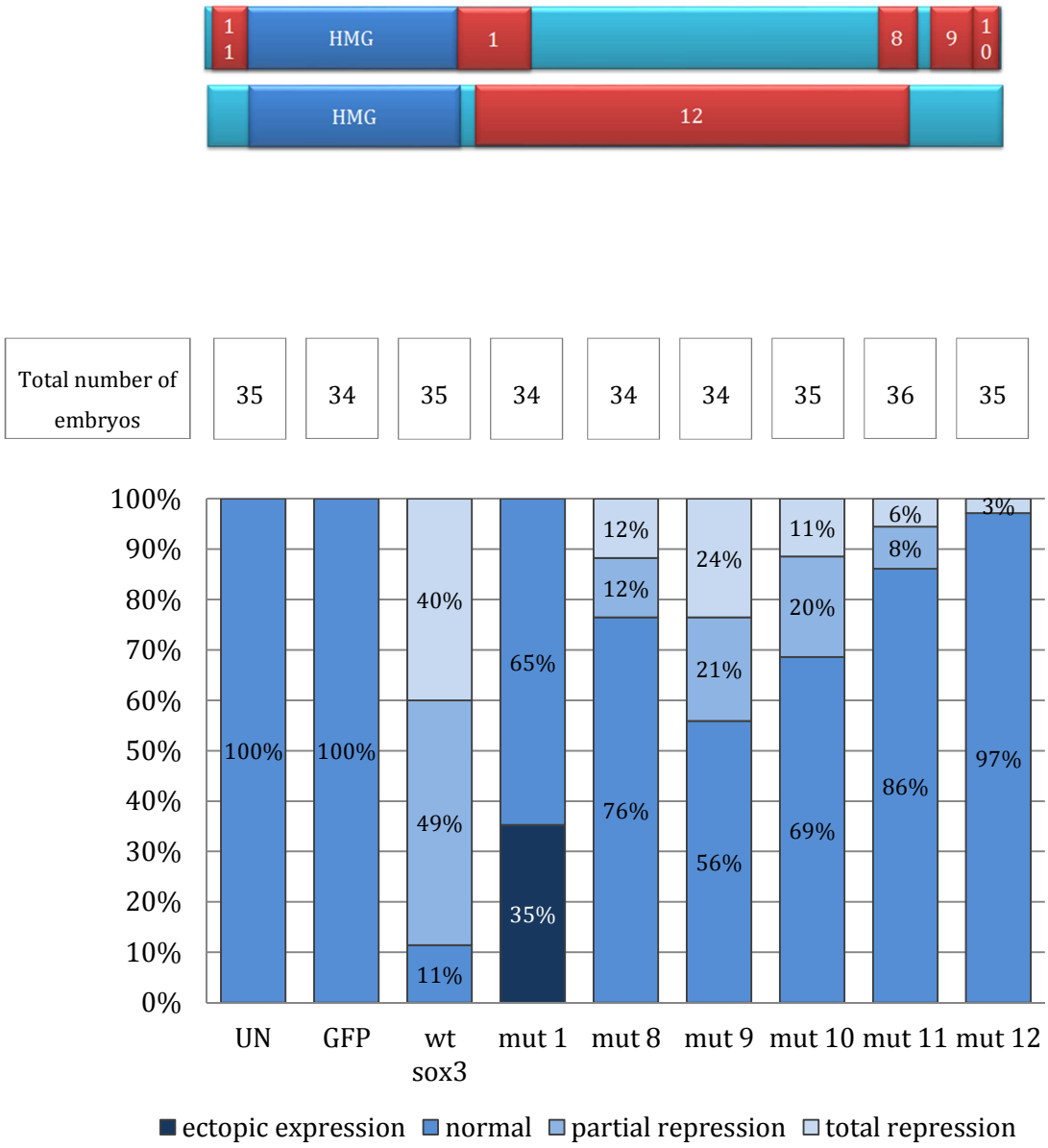


Figure 3.23 The top of the figure shows a schematic representation of the deletion constructs: the regions deleted in each of the mutants is represented as a red bar numbered accordingly to the name of the mutant. The chart shows the results of whole mount *in situ* hybridisation for *boz* performed on sphere stage embryos respectively uninjected (UN)or injected with *GFP* (*GFP*)or co-injected with *GFP* and *wt sox3* or *mut1*, *mut8*, *mut9*, *mut10*, *mut11*, *mut12*.

	normal	partial repression	total repression	total
UN	50	0	1	**51**
GFP	30	1	1	**32**
wtSox3	7	14	14	**35**
Mut2	18	13	2	**33**
Mut3	19	10	4	**33**
Mut4	30	15	4	**49**
Mut5	37	3	0	**40**
Mut6	27	5	0	**32**
Mut7	15	0	11	**26**
Mut12	35	5	0	**40**

Table 3.6 Numbers of Sphere Stage embryos analysed with whole mount *in situ* hybridisation for *boz* in the first set of experiments. The data shown correspond to the percentages shown in figure 3.22.

	ectopic	normal	partial repression	total repression	total
UN	0	35	0	0	**35**
GFP	0	34	0	0	**34**
wtSox3	0	4	17	14	**35**
Mut1	12	22	0	0	**34**
Mut8	0	26	4	4	**34**
Mut9	0	19	7	8	**34**
Mut10	0	24	7	4	**35**
Mut11	0	31	3	2	**36**
Mut12	0	34	0	1	**35**

Table 3.7 Numbers of Sphere Stage embryos analysed with whole mount *in situ* hybridisation for *boz* in the second set of experiments. The data shown correspond to the percentages shown in figure 3.23.

3.5.4. Expression of *goosecoid* in Embryos Microinjected with Sox3 Deletion Mutants

In order to get a deeper insight into the effects of the different Sox3 mutant constructs on the formation of the organizer, the *in situ* hybridization analysis was repeated with the organizer marker *goosecoid* (*gsc*). Also in this case the experiments were performed in two separate sets, in order to test the effects of the over-expression of all the twelve Sox3 deletion constructs compared to the wt Sox3. The phenotypes resulting from the experiments were categorised as normal, when the colour developed by *in situ* hybridization was dark purple and clearly visible in the organizer region (Fig. 2.24 G, H, I) partial repression, when the signal was barely visible (Fig. 2.24 D, E, F) and total repression, when the embryo appeared completely white (Fig. 2.24 A, B, C).

In the first set of experiments uninjected embryos (UN), embryos injected only with *GFP* (*GFP*) and embryos injected with GFP and *wt Sox3* (*wtSox3*) were compared to embryos injected with *GFP* and *mut2*, *mut3*, *mut4*, *mut5*, *mut6*, *mut7* and *mut12* (Fig. 3.25, Table 3.8). The results were as follows:

1) All UN embryos showed the normal phenotype (100%).
2) The 68% of the embryos injected with *GFP* showed normal expression of *gsc* and the remaining 32% partial repression.
3) As expected, over-expression of *wt sox3* resulted in a strong repression of *gsc* as 41% of the embryos showed partial repression phenotype and 59% total repression.
4) Also injections of *mut2* and *mut3* had a comparably strong repressive effect, as only 3% of the embryos were normal after injection of *mut2* a none after injection of *mut3*. Moreover, microinjection of *mut2* and *mut3* caused partial repression in respectively 72% and 43% of embryos and total repression in 25% and 57% of embryos.
5) Microinjection of *mut4* seemed to repress *gsc* slightly less than microinjection of *wt sox3*: 34% of the embryos injected with *mut4*

were found to be normal, but still 34% showed partial repression and 31% total repression.

6) Injection of *mut5*, as well as *mut7*, repressed *gsc* with the same extent of *wt sox3*, having respectively 26% and 39% of partial repression, 74% and 61% of total repression and no normal embryos.
7) On the other hand, injection of *mut6* caused 48% normal phenotype, which seemed to be quite comparable to the 68% found for the *GFP*-injected embryos, 43% of partial repression and only 9% of total repression.
8) Surprisingly, injections of *mut12* did not cause any normal phenotype, but it caused partial repression in 67% of the embryos and total repression in 33% of the embryos.

In the second set of experiments the UN and *GFP* experimental controls were repeated, as well as *mut12* and *wt sox3* for the same reasons previously explained in the context of *bozozok* expression analysis (Chapter 3.5.3). In these experiments the expression on *gsc* was compared between embryos uninjected, or microinjected only with *GFP* or microinjected with *GFP* and *wt sox3,* or *mut1*, *mut8*, *mut9*, *mut10*, *mut11*, *mut12* (Fig. 3.26, Table3.9). The results were as follows:

1) UN and *GFP* showed respectively 88% and 89% of normal phenotype, in addition with 12% and 11% of partial repression;
2) Microinjection of *wt sox3* caused 57% normal phenotype, 17% partial repression and 26% total repression;
3) The effects caused by injection of *mut1* were comparable to the UN and *GFP* controls, as it seemed not to repress *gsc* expression. In fact, 89% of embryos injected with *mut1* presented normal phenotype and 11% partial repression.
4) Injections of *mut8*, *mut9* and *mut10* caused a milder repression of *gsc* compared to *wt sox3*. In fact, they caused higher percentage of normal phenotypes compared to *wt sox3*, even though they still showed repression of *gsc*.

5) Interestingly, injection of *mut11* seemed not to cause repression of *gsc*, as it resulted in a normal phenotype in 94% of the embryos (versus the 57% of *wt sox3* and the 88% and 89% of UN and *GFP*) and partial repression in only the 6% of the embryos, which is comparable to the 12% and 11% found for the UN and *GFP* controls.
6) Injection of *Mut12* did not presented any repressive effect as *gsc* was normally expressed in 100% of the embryos.

It is quite evident that the results obtained with the two set of experiments had some variability. In fact, in the first set of experiments UN embryos showed 100% normal phenotype, while in the second set 88% showed normal phenotype and 12% partial repression (Fig. 3.25 and Fig. 3.26). Also *GFP*-injected embryos seemed to present variability of *gsc* expression between the two sets of experiments, showing 32% of embryos with partial repression phenotype in the first set, while 11% in the second one (Fig. 3.25 and Fig. 3.26). However, although it seemed to be an overlapping between the normal phenotype and the partial repression phenotype, there were no embryos presenting total repression in UN or *GFP* controls, therefore the experiments were considered reliable. The reasons that could explain the variability of *gsc* expression are presented and discussed in Section 4.3.2.

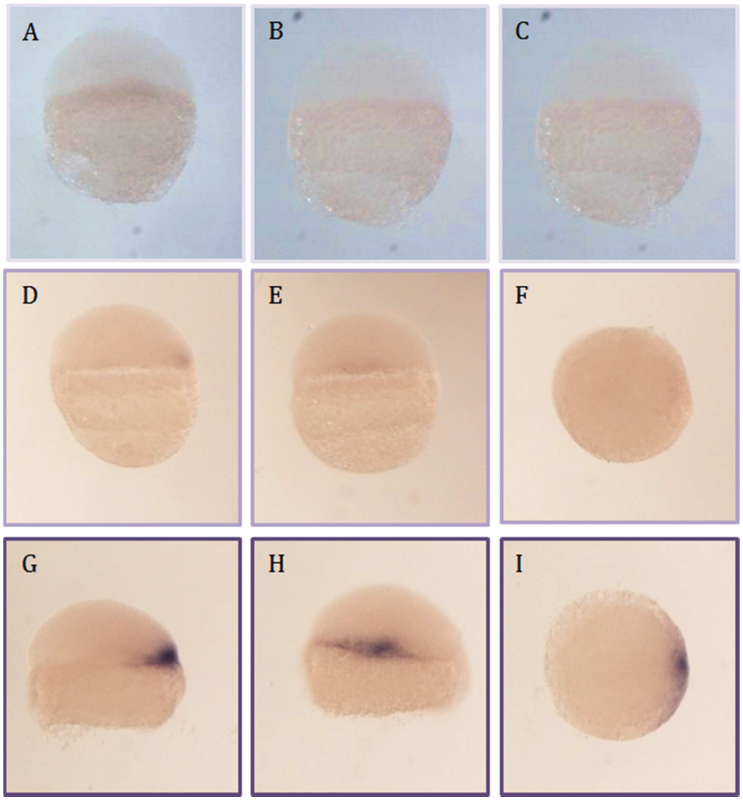


Figure 3.24 Pictures representing the three categories of phenotypes obtained through whole mount *in situ* hybridization for *gsc* on sphere stage embryos uninjected or injected with *GFP* or/and *wt*/*mut sox3*. The phenotypes include: total repression of *gsc* (A, B, C), partial repression of *gsc* (D, E, F) and normal expression of *gsc* (G, H, I).

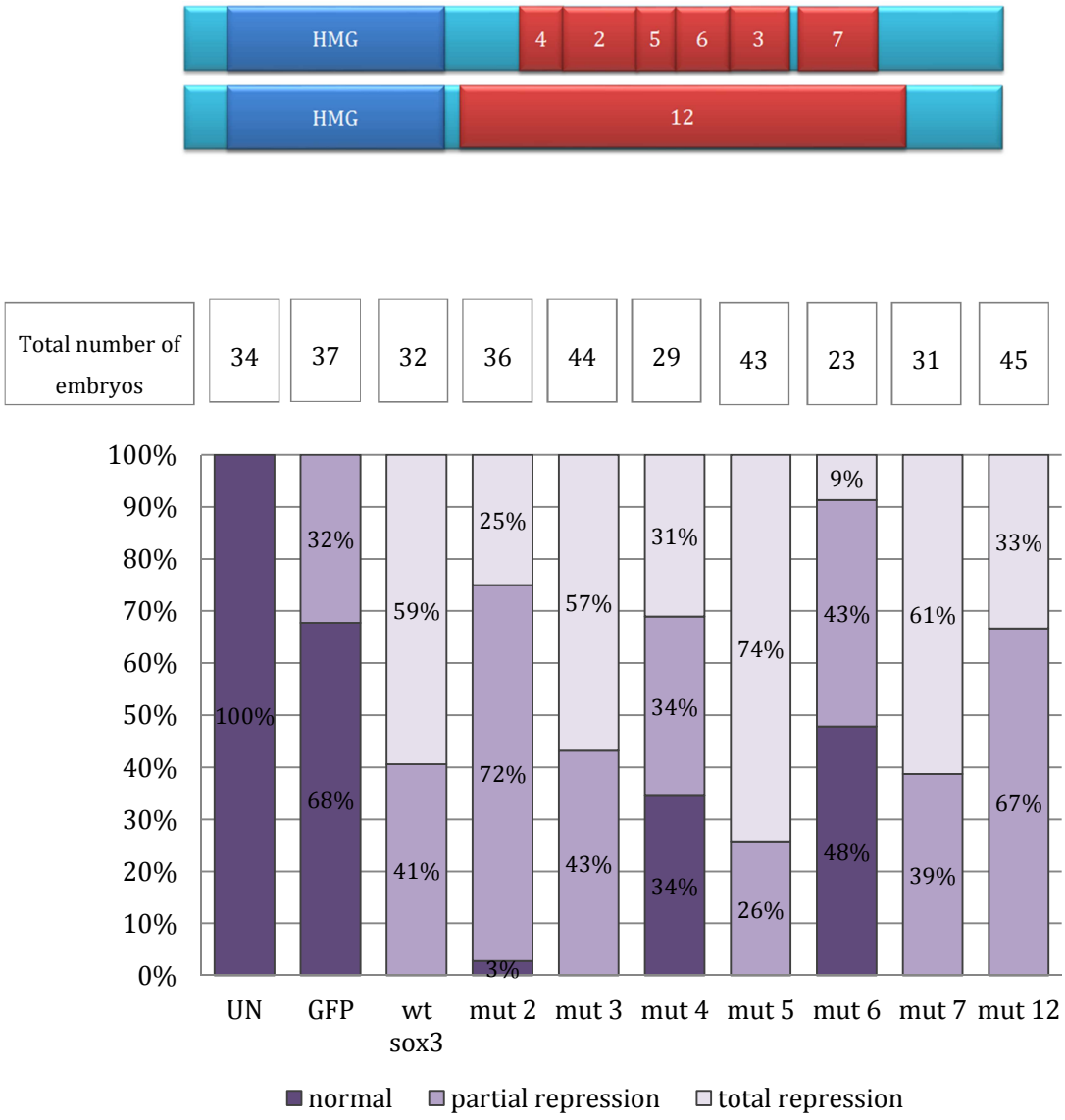


Figure 3.25 The top of the figure shows a schematic representation of the deletion constructs: the regions deleted in each of the mutants is represented as a red bar numbered accordingly to the name of the mutant. The chart shows the results of whole mount *in situ* hybridisation for *gsc* performed on sphere stage embryos respectively uninjected (UN)or injected with *GFP* (*GFP*)or co-injected with *GFP* and *wt sox3* or *mut2*, *mut3*, *mut4*, *mut5*, *mut6*, *mut7*, *mut12*.

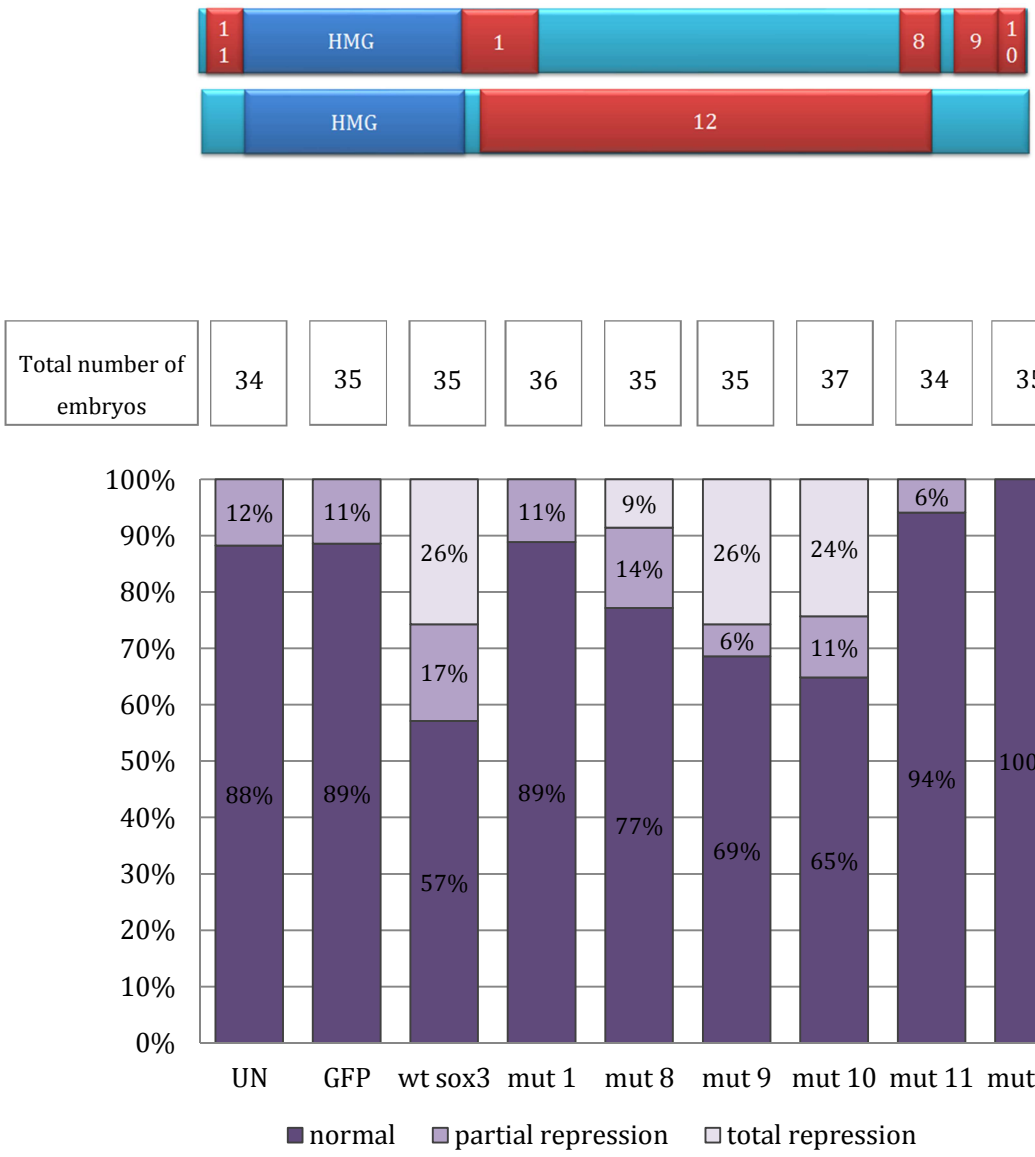


Figure 3.26 The top of the figure shows a schematic representation of the deletion constructs: the regions deleted in each of the mutants is represented as a red bar numbered accordingly to the name of the mutant. The chart shows the results of whole mount *in situ* hybridisation for *gsc* performed on Sphere Stage embryos respectively uninjected (UN)or injected with *GFP* (*GFP*)or co-injected with *GFP* and *wt sox3* or *mut1*, *mut8*, *mut9*, *mut10*, *mut11*, *mut12*.

	normal	partial repression	total repression	total
UN	34	0	0	**34**
GFP	21	10	6	**37**
wtSox3	0	13	19	**32**
Mut2	1	26	9	**36**
Mut3	0	19	25	**44**
Mut4	10	10	9	**29**
Mut5	0	11	32	**43**
Mut6	11	10	2	**23**
Mut7	0	12	19	**31**
Mut12	0	30	15	**45**

Table 3.8 Numbers of sphere stage embryos analysed with whole mount *in situ* hybridisation for *gsc* in the second set of experiments. The data shown correspond to the percentages shown in figure 3.25.

	normal	partial repression	total repression	total
UN	30	4	0	**34**
GFP	31	4	0	**35**
wtSox3	20	6	9	**35**
Mut1	32	4	0	**36**
Mut8	27	5	3	**35**
Mut9	24	2	9	**35**
Mut10	24	4	9	**37**
Mut11	32	2	0	**34**
Mut12	35	0	0	**35**

Table 3.9 Numbers of sphere stage embryos analysed with whole mount *in situ* hybridisation for *gsc* in the first set of experiments. The data shown correspond to the percentages shown in figure 3.26.

3.6. Transcriptional Activation Function of wt and Deletion Mutant Sox3 Proteins

Luciferase reporter assay analysis was performed in order to determine whether the deletions in Sox3 constructs modified the ability of wt Sox3 to activate the transcription of a target gene. Zebrafish 1-4 cell stage embryos were incubated uninjected (UN control), or after co-injection with *sox3* (*wt* or deletion mutants) capped RNA (50pg) and *GFP* capped RNA (50pg) together with a plasmid containing three Sox3 binding sites upstream the Firefly *luciferase* gene (*pGL3-3XSX*, 5 pg, Appendix 7.4). The experiment was also performed on a control group of embryos that were co-injected with *GFP* RNA and *pGL3-3XSX* plasmid, as the endogenous Sox3 was also expected to cause expression of the *luciferase*.

The co-injection of embryos with *GFP* together with the *sox3* constructs allowed the selection of the embryos that showed similar levels of expression of GFP. Microinjection of RNA does not result in the protein being uniformly expressed throughout the entire embryo, it can be quite variable. This can be due to different reasons, including the stage when the embryo is injected and the position of the needle when injecting. In other words, since the manual component of this technique is relevant, it cannot be perfectly reproducible. However, the injection of *GFP* allowed visualisation of the expression pattern of the injected construct under UV light after the incubation of the embryos until sphere stage (Fig. 3.27). Therefore, only embryos that showed similar and diffuse expression of GFP were selected and used for the analysis.

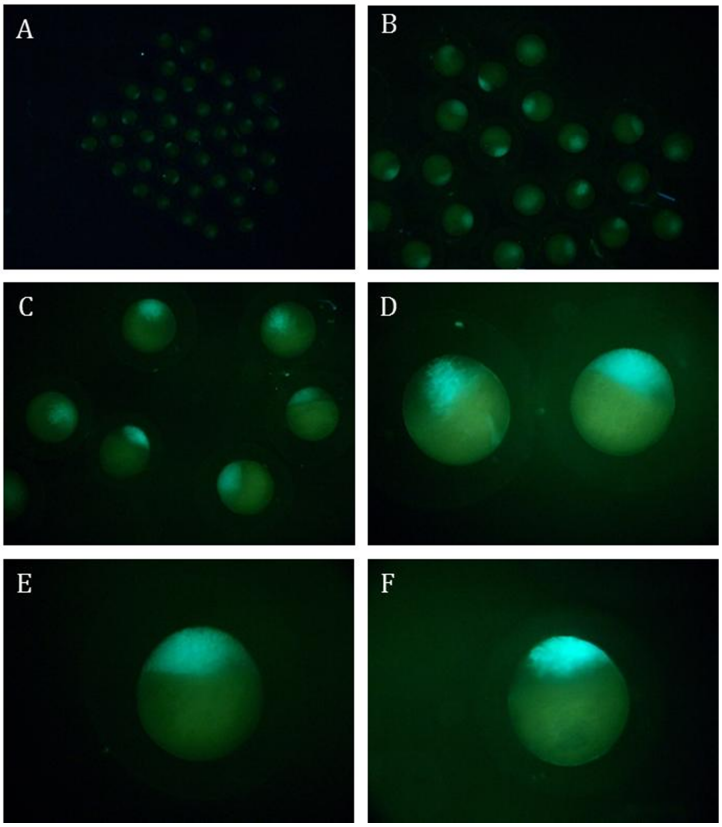


Figure 3.27 Sphere Stage embryos that were previously co-injected at 1-4 cell stage with *GFP* RNA (50 pg), *wt Sox3* RNA (50 pg) and *pGL3-3XSX* plasmid (5 pg) and then incubated for approximately 4 ½ hours. The embryos are shown under UV light to visualise expression of GFP, which can vary due to the microinjection procedure. This is clearly visible in fig.D, where the left embryo shows GFP expression in a restricted area, while the right embryo shows a diffused expression of GFP. Highly expressed GFP appears as blue due to the printing of the image.

The aim of the study was to investigate the activity of Sox3 as a transcriptional activator and repressor. The selection of the constructs to be analysed with the Luficerase Reporter Assay was based on the data obtained with the previous experiments, which showed the repression function of Sox3. For this reason, the experiment was performed on the experimental groups shown in Table 3.10. For each experimental group the experiment was repeated three times on three different groups of embryos in order to normalise the results and check that they had not been affected by technical errors.

The results of the experiments are represented in Fig. 3.28-3.29 and Table 3.11. As expected, the UN control showed a low level of luminescence (the mean of raw readings is 23), which was due to background luminescence. As shown in Fig. 3.28, Fig. 3.29 and Table 3.11, the reporter gene was activated in the GFP co-injected embryos, presumably due to the presence of endogenous Sox3. The reading for wtSox3-injected embryos presented a two-fold increase compared to GFP. Mut1 activated the *luciferase* gene with a two-fold increase compared to wtSox3, while Mut2 did not activate it. Mut5 showed a remarkably strong activation effect, which was approximately tenfold increase compared to wtSox3. The reading for Mut6 is comparable to wtSox3. Mut7 failed to activate the reporter, while Mut10 did show an activation remarkably lower than the GFP control. Mut11 strongly activated the reporter gene compared to wtSox3. Finally, as expected, Mut12 failed to activate the reporter gene, but the reading was even lower than the GFP control.

Name	Microinjected constructs	Reasons
UN	-	Negative control
GFP	• *GFP RNA* • *pGL3-3XSX*	Control for the activation of the reporter gene by endogenous Sox3
wtSox3	• *wt Sox3 RNA* • *GFP RNA* • *pGL3-3XSX*	Necessary for the comparison with the deletion mutants
Mut1	• *mut1 RNA* • *GFP RNA* • *pGL3-3XSX*	Caused ectopic expression of *boz* and did not repress *gsc* expression.
Mut2	• *mut2 RNA* • *GFP RNA* • *pGL3-3XSX*	The deleted region contains highly conserved amino acid patterns and is located between the HMG domain and the region deleted in Mut5 and Mut6.
Mut5	• *mut5 RNA* • *GFP RNA* • *pGL3-3XSX*	Did not caused repression of *boz*.
Mut6	• *mut6 RNA* • *GFP RNA* • *pGL3-3XSX*	Did not caused repression of *boz* and *gsc*.
Mut7	• *mut7 RNA* • *GFP RNA* • *pGL3-3XSX*	The deleted region contains the SUMOylation site (Chapter 1.3.2). The inability to be SUMOylated could affect the activity of Sox3 as transcriptional activator.
Mut10	• *mut10 RNA* • *GFP RNA* • *pGL3-3XSX*	Caused unique and distinctive phenotypes 24 hpf embryos.
Mut11	• *mut11 RNA* • *GFP RNA* • *pGL3-3XSX*	The deleted region is the N-terminal region of the protein, which is thought not to be involved in the activation function.
Mut12	• *mut12 RNA* • *GFP RNA* • *pGL3-3XSX*	Good control as it contains only the N-terminal region and the DNA binding domain; it is expected not to be able to activate the reporter gene.

Table 3.10 Experimental groups chosen for performing Luciferase Reporter Assay, constructs injected in the embryos of each group and reasons why each construct was chosen.

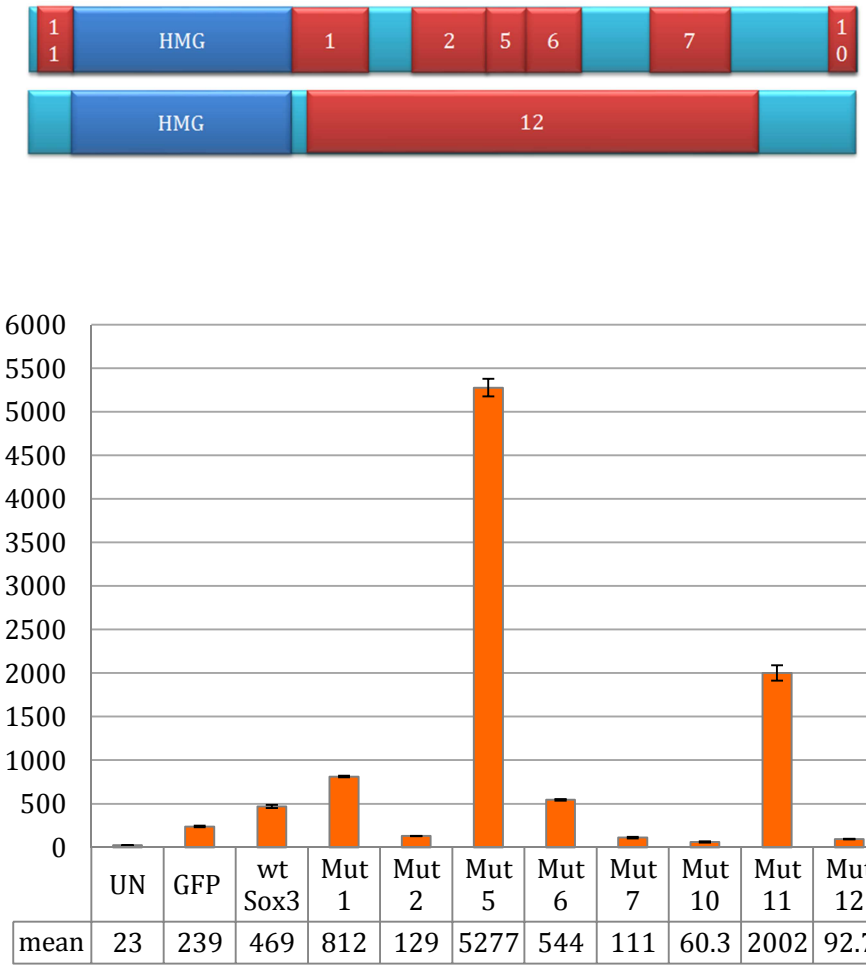
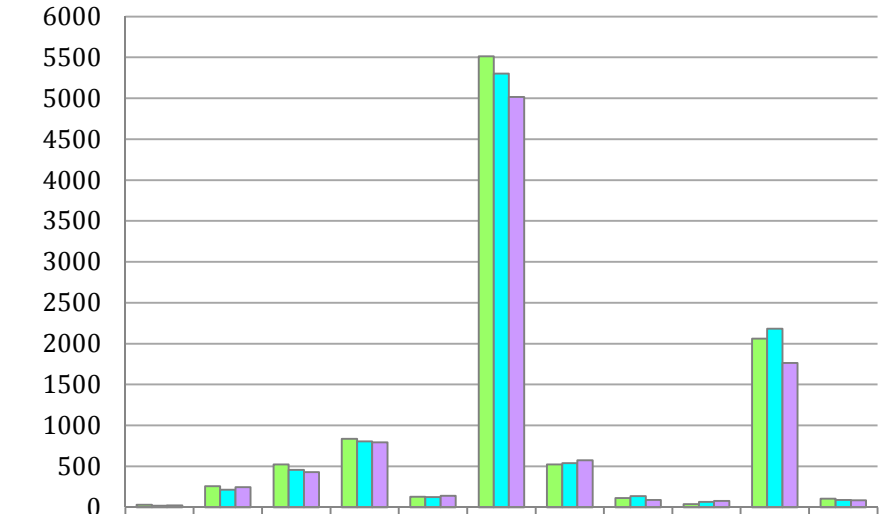


Figure 3.28 The top of the figure shows a schematic representation of the deletion constructs: the regions deleted in each of the mutants is represented as a red bar numbered accordingly to the name of the mutant. The chart represents the mean values of the three readings of the Luciferase Reporter Assay made for each of the experimental groups of embryos. The groups included: uninjected embryos (UN), embryos injected with *GFP* only (GFP), and embryos co-injected with GFP and *wt sox3* or *mut1*, *mut2*, *mut5*, *mut6*, *mut7*, *mut10*, *mut11*, *mut12*. The error bars were built using the standard deviation of the values (Table 3.10).



	UN	GFP	wt Sox3	Mut 1	Mut 2	Mut 5	Mut 6	Mut 7	Mut 10	Mut 11	Mut 12
Series1	30	256	523	838	126	5513	523	112	39	2063	103
Series2	17	213	455	803	123	5302	537	135	64	2181	89
Series3	22	247	430	794	138	5016	573	87	78	1762	86

Figure 3.29 The chart represents the values of the three readings (1, 2, 3) of the Luciferase Reporter Assay made for each of the experimental groups of embryos. The groups included: uninjected embryos (UN), embryos injected with *GFP* only (*GFP*), and embryos co-injected with *GFP* and *wt/mut sox3*.

	1	2	3	mean	st. dev.
UN	30	17	22	**23**	5.35
GFP	256	213	247	**239**	18.52
wtSox3	523	455	430	**469**	39.30
Mut1	838	803	794	**812**	18.98
Mut2	126	123	138	**129**	6.48
Mut5	5513	5302	5016	**5277**	203.67
Mut6	523	537	573	**544**	21.06
Mut7	112	135	87	**111**	19.60
Mut10	39	64	78	**60**	16.13
Mut11	2063	2181	1762	**2002**	176.41
Mut12	103	89	86	**93**	7.41

Table 3.11 Raw data of the Luciferase Reporter Assay made on sphere stage embryos uninjected (UN), or microinjected with *GFP* only (*GFP*), or co-injected with *GFP* and *wt/mut sox3*. The first three columns of data correspond to the reading of three different groups of embryos (1, 2, 3), the fourth column contain the mean of the readings and the last column correspond to the standard deviation (st. dev.).

4. Discussion

4.1. Deletion Cloning and Expression of Mutant Proteins in Zebrafish Allowed Functional Screening of the Entire Sequence of Sox3

The design of the deletions in this study was made entirely on the basis of the pattern of conservation of Sox3 amino acids sequence between different species. This was made without considering any other information known (or presumptive) about the protein, such as post-transcriptional modifications (for example SUMOylation or phosphorylation) or presumptive localisation of functional domains (for example the C-terminal region is thought to contain a trans-activation domain as do the other SoxB1 factors). The reason of this approach was to analyse the entire protein without preconception, so that the results obtained could have confirmed or dispute previous evidence. Furthermore, this approach could have given new clues onto the mechanisms of function of Sox3, thus providing new starting points for subsequent studies that could focus on more restricted regions of the protein.

The design of many different deletions to be realised in the same gene led to problems in the cloning experiments because the efficiency of the cloning proved to be highly variable between different mutants. For this reason the design of a high efficiency cloning strategy was a key step for this study: a strategy that worked for creating all the clones allowed a more reliable comparison of their effects and, on the other hand, also a faster way to obtain all the constructs. In fact, the design of different approaches for different constructs would have added variability in the experiments and would have also required more time for the development and realisation. Moreover, an important feature of the cloning strategy that was developed is that it enabled the generation of large deletions and

the insertion of a diagnostic restriction site by inserting in the final protein just one or two additional amino acids (alanine and/or glycine).

Since the design of the deletion did not take into account any other factor than the amino acids sequence, the obtained constructs could be missing regions necessary for the protein expression or regions that would have radically modified the levels of expression. The results of Western Blot analysis confirmed that the expression of all the mutant proteins was not affected by the mutation at sphere stage, and that the levels of expression were similar for all the constructs, including the wt.

Therefore, the results obtained with the cloning experiments and with the Western Blot built the basis for all the subsequent experiments, which were based on the functional comparison between the different constructs obtained.

4.2. Analysis of the CNS at 24hpf Highlighted Differing Functions of the Sox3 Deletion Mutants

Each of the Sox3 mutant constructs contained a deletion of about 20 amino acids, which is quite a large region considering the full-length protein is only 300 amino acids long. Therefore, such big deletions could have just caused disruption of the three-dimensional integrity of the protein, resulting in complete loss of function. In this case the over-expression of the constructs would not have caused changes in the phenotypes of 24 hpf embryos, or the resulting phenotypes would have all been similar, as the mutant Sox3 proteins would have just been disrupted and not functioning, no matter where the different deletions were localised. However, the results of the experiments conducted on 24 hpf embryos clearly showed that different deletions caused different effects, especially in the cases of Mut1, Mut10 and Mut12.

The results presented in Chapter 3.4 demonstrated that the over-expression of wt Sox3 caused severe duplications of the CNS in the majority of the embryos, while most of the mutants showed milder effects. This means that by deleting the corresponding regions of Sox3, the function of the protein has been changed.

4.3. The Deletion of Different Regions of Sox3 Caused Different Effects on the Protein's Repressor and Activator Functions

The analysis of the effects that microinjection of Sox3 deletion mutants had in 24 hpf embryos was followed by the analysis of their effects at the sphere stage. This analysis was more direct, as it was possible to investigate whether a known function of Sox3 (repression of organizer formation) was changed at a stage when no other SoxB1 proteins are known to be expressed, and therefore no functional redundancy should be present. Over-expression of different Sox3 deletion mutants caused different changes in the expression of the organizer markers compared to the wt protein and allowed us to get a better understanding of how the different deletions affected the repressor function of Sox3.

While the data obtained through the analysis of the expression of *bozozok* were very consistent between the two sets of experiments performed, the results obtained on the transcriptional regulation of *goosecoid* (Section 3.5.4) did show a difference between the two sets of experiments performed. In fact, if comparing the two sets of results, it may seems that the repressive effects were generally stronger in the first set, and milder in the second set, which showed higher percentages of normal phenotypes in all the samples analysed. However, there are different reasons that could explain this phenomenon. Firstly, the two sets of experiments correspond to two different *in situ* hybridisation experiments, which means that the difference can be due to the technical repetition: for example, a change in the temperature during the phases of hybridisation, or longer time in the manipulation of the samples can cause a difference in the hybridization of the probes to the target transcripts. Another factor that can cause variations between different replicas of this technique is the different solutions that are used, and that can be slightly different. However, both the sets of experiments were performed together with two different control groups each (UN and GFP), and the consistency of these controls

allows to trust the results obtained. In fact, in both the sets of experiments the control showed percentages of normal phenotypes higher than the wt Sox3, which means that the repressive effect that Sox3 had on *gsc* expression is visible and consistent in both the sets of experiments, even though it may seem more severe in the first data compared to the second ones. For the reasons just explained it is not correct, however, to compare the results obtained with different experiments without considering the results obtained for the controls: in particular, the GFP control seemed to be more severely affected by injection in the first set of experiments, but it seemed that the decrease in the development of the purple signal is due to the experimental technique rather than a more severe repressive effect on *gcs* expression. Moreover, as discussed below, the data obtained are also consistent with the results on the expression of *bozozok*. In addition, in interpretation of the results obtained by the analysis of the expression of *goosecoid* must be taken in count that the repressive effect that Sox3 has on *gsc* is not as direct has the one that it has on *boz*. In fact, it had been shown that Sox3 represses *gsc* through at least two mechanisms: one is direct repression, the other is mediated by the repression of *boz*, which normally activates *gsc* (Fig. 4.1), (Shih *et al.* 2010). The reason why in this project it was chosen to study the expression of *gsc*, together with *boz* (instead of *squint* for example), was to observe the effects that Sox3 deletion mutants would have caused on target genes located at different steps of the signalling pathway and repressed through different mechanisms by Sox3.

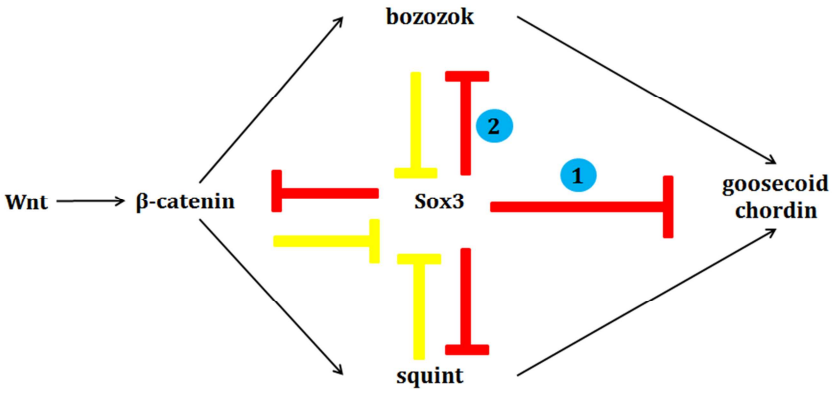


Figure 4.1 Model of the transcriptional repression role of Sox3 during the formation of the organizer. Sox3 represses Gsc both directly (1) and through the repression of Boz (2). Picture adapted from Shih *et al.* (2010).

The luciferase reporter assay allowed the analysis of how the deletions changed the activator function of Sox3. As explained in the Introduction (Chapter 1.6), the main aim of this study was to identify regions of Sox3 that are involved in its ability to function both as a transcriptional repressor and a transcriptional activator with context-dependency. While the results obtained through the analysis of the effects that the different deletion mutants have on the formation of the organizer provided information about the regions of Sox3 that are involved in its repression function, the experiments performed in the Luciferase Assay were meant to provide information about Sox3 activation function. In fact, this assay measures the ability of the protein to activate target genes. However, because Sox3 is both an activator and a repressor, it is uncertain whether an increase in luciferase activity is due to an increase in the activator function of Sox3, or to a decrease in its repressor function. Since all the mutant Sox3 constructs created in this study lacked specific regions of the protein, we can presume that if a mutant caused an increase in the luciferase activity, than the deleted region might contain residues that are important for Sox3 repression function. On the other hand, a decrease in the luciferase activity would indicate that a region important for Sox3 activation function is missing. Alternatively, it is also possible that the decrease in the luciferase activity is due to a dominant-negative interaction of the mutant protein on the endogenous Sox3, which would no longer be able to activate the reporter gene with the same extent.

In the following paragraphs the deletion mutants that generated consistent data in the analysis of 24 hpf embryos, in the analysis of the expression of organizer markers and in the luciferase reporter assay are discussed together with the hypotheses that could explain their effects. The other deletion mutants (Mut2, Mut3, Mut4, Mut8 and Mut9) did generally show milder repressive effects on the organizer marker genes if compared to wt Sox3, but the data were not sufficient to formulate more specific hypothesis for explaining these effects. Only Mut2 was also analysed through Luciferase Reporter Assay and it activated the reporter gene less than the GFP control. This could mean that Mut2 lost the repressor function and also that it acted as a dominant-negative on endogenous Sox3. In order to explain these effects more experiments are needed. Moreover, these data also suggest that it could be interesting to conduct further analysis also on Mut3, Mut4, Mut8 and Mut9.

4.3.1. Mutant 1

Figure 4.2 Schematic representation of *sox3* showing the region deleted in Mut1.

The region missing from Mutant 1 is located imediately C-terminal to the HMG domain (Fig. 4.2). Microinjection of *mut1* affected the development of the CNS in 24 hpf embryos substantially less than microinjection of *wt Sox3*. The over-expression of Mut1 caused a unique pattern of expression of *bozozok*: it was found that this organizer marker was expressed in the region of the organizer but also ectopically in other regions (Fig. 4.3B). This particular phenotype has already been seen in other experiments, where dominant negative forms of Sox3 were over-expressed in zebrafish embryos (Fig. 4.3A), (Shih *et al.* 2010). Shih *et al.* tested the ability of two different dominant negative Sox3 to repress organizer markers. One of the dominant negative (Sox3N40I) contained a point mutation in the HMG domain that prevented the binding to DNA, while the other (Sox3dNLS) contained three point mutations, again in the HMG domain, that prevented the nuclear localisation of the protein. Both the mutants induced ectopic expression of the organizer markers *bozozok*, *squint*, *goosecoid*, *chordin* and *no tail*. Moreover, microinjection of *mut1* did not repress the expression of *gsc*, as it caused phenotypes similar to both UN and GFP controls. Therefore, the data obtained indicates that this mutant present a loss of the repressor function compared to the wt.

In the luciferase reporter assay Mut1 caused a two-fold increase of the luciferase activity compared to wt Sox3. This could be due to the fact that the deletion made in Sox3 did remove a region involved in Sox3 repressor function, thus altering the balance between the activation and repression function of the protein and causing greater activation of the reporter gene.

Two hypotheses that would explain why the over-expression of *mut1* had a milder impact than the wt on the organizer formation are that the deletion affected Sox3 DNA binding properties, or that it affected the

nuclear localisation of the protein. These hypotheses would also be consistent with the activation of the luciferase reporter gene observed after over-expression of Mut1. According to the first hypothesis, a decrease in the ability of binding DNA would allow a lower amount of the over-expressed protein to bind target sequence and activate/repress target genes, therefore mitigating the effect of the over-expression. If, on the other hand, the second hypothesis was true, than it would mean that the amount of protein transported into the nucleus after translation would have decreased (but not completely), and for this reason the mutant Sox3 would have been less efficient in regulating target genes.

Another hypothesis that would explain the ectopic expression of *boz* is that Mut1 lost the ability to compete with the Wnt effector factors Tcf/Lef for binding to β-catenin, and thus it interfered with the Wnt/β-catenin signalling pathway (Fig. 4.4). In fact, in order for β-catenin to bind DNA and activate target genes, the interaction with Tcf/Lef is required; in *Xenopus* it has been shown that Sox3, as well as other Sox proteins, can interfere with this interaction (Zorn *et al.* 1999; Zorn *et al.* 1999). This is consistent with the observation that several other Sox proteins have been shown to be able to interact with β-catenin, sometimes causing repression and other times causing activation of target genes (Zorn *et al.* 1999; Sinner *et al.* 2004). In zebrafish, it has been shown that Sox3 acts as a repressor of the Wnt/β-catenin signalling causing the repression of organizer marker genes and confining the formation of the organizer at the correct time and position (Shih *et al.* 2010). It is also known that the activation of *boz* is induced by β-catenin (Fig. 4.1 and Fig. 4.4), (Schier and Talbot 2005); therefore, it could be possible that the ectopic expression of *boz* caused by the over-expression of Mut1 was due to the fact that the region deleted in Mut1 is involved in the ability of Sox3 to interact with β-catenin. The reason why Mut1 did not cause ectopic expression of *gsc* could be because, as mentioned above, Sox3 does not repress *gsc* as directly as it does repress *boz* (Fig. 4.1). This is consistent with the observation that Mut1 did not repress *gsc* expression.

Zebrafish Sox3 (zfSox3) was found to contain two presumptive SUMOylation sites (Laghari 2010) located respectively in the positions 113-123 and 228-238. SUMOylation consists in the covalent or, in some cases, non-covalent (Merrill *et al.* 2010) attachment of a SUMO polypeptide to a lysine flanked by specific residues on the target protein and it was found to affect different types of proteins, including transcription factors. SUMOylation has been shown to regulate the transcriptional activity of different transcription factors, including Sox proteins (Gill 2005). It is thought that the covalent attachment of SUMO polypeptides to zfSox3 modulates the transcriptional repressor and activator function of this transcription factor. However, at the present time this mechanism is still being studied. The observation that one of the presumptive SUMOylation sites, which is located in the position 113-123, is completely deleted in Mut1 suggests another hypothesis for the function of the deleted region. It is possible that the SUMOylation of this region of the protein is involved in the repressor function of Sox3. That would explain why the deletion caused the loss-of-function in the organizer repression activity and the activation of the luciferase, which could also be caused by a loss of repression function.

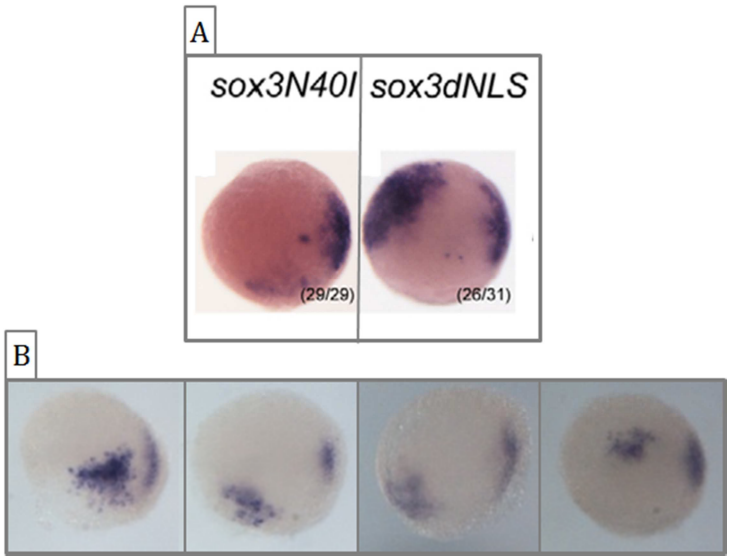


Figure 4.3 Dominant-negative Sox3 constructs induce ectopic expression of *boz* causing the same phenotype as *mut1*. Shih *et al.* showed that microinjection of dominant-negative forms of *sox3* (*sox3N40I* and *sox3dNLS*) causes ectopic expression of *boz* (A), (Shih *et al.* 2010). The same phenotype was found to be caused by microinjection of *mut1* (B). The pictures show different embryos viewed from the animal pole.

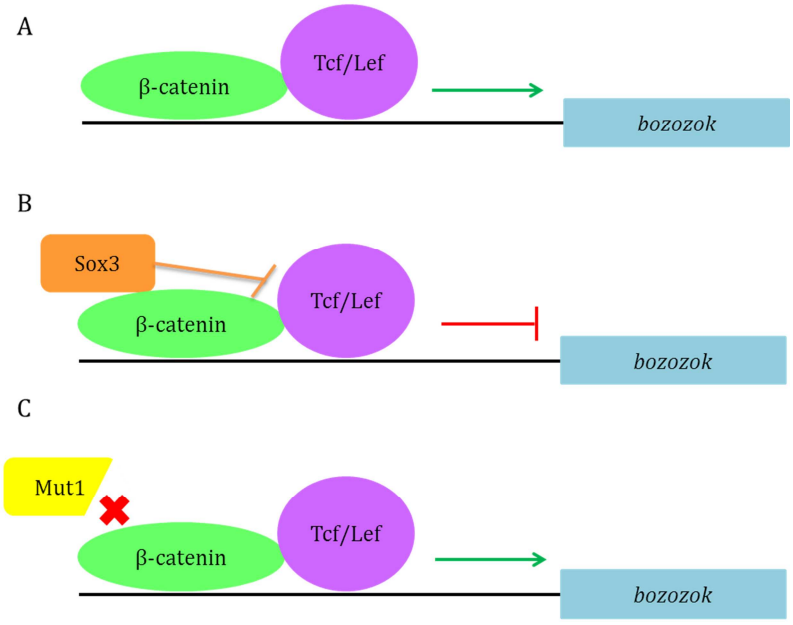


Figure 4.4 One of the hypotheses that would explain ectopic expression of *boz* caused by over-expression of Mut1. B-catenin binds the DNA together with Tcf/Lef to activate *boz* (A); Sox3 can compete with this interaction causing the repression of *boz* (B). If Mut1 lost the ability of interfering with β-catenin-Tcf/Lef interaction, this would explain why *boz* was expressed ectopically (C).

4.3.2. Mutant 5

Figure 4.5 Schematic representation of *sox3* showing the region deleted in Mut5.

Mutant 5 carried deletion of a region of 16 amino acids located approximately in the middle of the protein (A164-M179, Fig. 4.5). The over-expression of Mut5, as well as Mut6, did not cause repression of *boz* compared to wt Sox3. However, microinjection of *mut5*, as well as microinjections of *mut2*, *mut3* and *mut7* seemed not to cause different effects on *gsc* expression compared to the wt. This observation is interesting considering that Mut4, which lacks a region located just before Mut2, and Mut6, which lacks the region between the deletions made in Mut5 and Mut3, and Mut4, did cause significantly milder effects (Fig. 4.6). Therefore, the deletion of regions that are next to each other cause different effects on the organiser repression function. This proved that the deletion strategy developed for the present study is effective to identify regions of the protein that have different functions. The reason why Mut5 did not show the same loss-of-function in the repression of both *boz* and *gsc* is probably because, as mentioned above, *gcs* is not only directly repressed by Sox3, but also through the repression of *boz*.

Figure 4.6 Schematic representation of *sox3* showing the positions of the regions deleted in Mut2, Mut3, Mut5 and Mut7.

Mutant 5 caused a remarkably strong activation of the luciferase compared to wt Sox3. Such dramatic effect could be caused by the lack of a region that mediates the repressive function of Sox3. This is consistent with the findings that Mut5 is unable to repress the expression of *bozozok*. Both experiments support the hypothesis that Mut5 lost the repression function of wt Sox3.

Interestingly, a previous student in the laboratory identified two presumptive Groucho binding sites in the sequence of chicken Sox3 (cSox3) (Laghari 2010). Grouchos are a family of co-repressors that have been proven to interact with Tcf, a Sox-like HMG transcription factor (Brantjes *et al.* 2001; Kuo *et al.* 2013) and it has been suggested that they may interact with Sox2 (Liu 2011). The interaction between Tcf and Grouchos cause repression of target genes, while the binding with β-catenin leads to the release of Tcf from Groucho, thus activating the target genes (Brantjes *et al.* 2001; Kuo *et al.* 2013). It is possible that Grouchos interact with Sox3 and regulate the balance between the activator and the repressor function of Sox3. The presumptive Groucho binding sites that were identified in cSox3 correspond to the amino acids 194-203 and 292-307. In the zebrafish Sox3 (zfSox3) these regions are conserved and corresponding residues are located in the positions 179-188 and 276-291 (Fig. 3.1). The first of these two presumptive binding sites has been deleted in Mut5 and Mut6: Mut5 (A164-M179) was missing the first amino acid of the binding site and Mut6 (H180-N200) was missing the rest of the site. The second presumptive Groucho binding site was deleted in Mut9 (Q273-G290) and in Mut10 (V291-I300).

Therefore, according to this hypothesis, in Mut5 the removal of a region involved in the binding of Sox3 with the co-repressor Groucho caused a loss of the repressor function and changed the balance between the activator and repressor function, causing the failure in the repression of *boz* and the activation of the *luciferase* reporter.

Mut2 led to a reduction in the luciferase activity compared to the GFP control, where the activity was due only to the endogenous Sox3. This

could mean that Mut2 also interfered with the activity of the endogenous Sox3. This observation is interesting considering that the deletion of the region next to the one deleted in Mut2 caused a totally different effect: the dramatic activation showed by Mut5.

4.3.3. Mutant 6

Figure 4.7 Schematic representation of *sox3* showing the region deleted in Mut6.

Mutant 6 lacks of a region of 21 amino acids (H180-N200, Fig. 4.7) which is located approximately in the middle of the protein sequence (Fig. 4.7). The over-expression of Mut6 did not cause repression of the organizer markers compared to wt Sox3 and this could indicate a loss-of-function in the repressor activity. These observations are consistent with the hypothesis that the removal of a part of a Groucho binding site (179-188) changed or prevented the interaction of Mut6 to with Groucho, thus leading to the inability of the protein to repress *boz* and *gsc.* The fact that the phenotypes of the embryos 24 hpf did not differ particularly from the embryos microinjected with wt Sox3 is probably due to the functional redundancy existing between different SoxB1 factors at later stages of development.

However, Mut6, in contrast to Mut5, did not activate luciferase more than wt Sox3. If the above hypothesis that lack of repression is due to loss of interaction with Grouchos, then this cannot be the explanation for the increased luciferase activity caused by Mut5. Only by analysing interaction with Grouchos directly can these possibilities be tested.

4.3.4. Mutant 7

Figure 4.8 Schematic representation of *sox3* showing the region deleted in Mut7.

Mutant 7 contained a deletion of 25 amino acids located towards the C-terminus of the protein (A226-G250, Fig. 4.8). The over-expression of this construct did not show any repressive effect on *boz* compared to *wt sox3.* However, it seemed to have quite a strong effect on *gsc* expression, as it caused a severe phenotype in almost half of the embryos and a normal phenotype in the others, but no mild phenotypes. Moreover, Mut7 activated the *luciferase* reporter gene less than the GFP control, in which the activation of the reporter is thought to be caused by the endogenous Sox3. Therefore, Mut7 seemed to interfere with the transcriptional activator function of the endogenous Sox3. These data suggest that the deletion made in Mut7 caused a loss of the transcriptional activator function of Sox3, and therefore that the deleted region is involved in the functioning of Sox3 as activator.

As mentioned above, zfSox3 was found to contain a presumptive SUMOylation site in the position 228-238, which is comprised in the deletion made in Mut7. In contrast to what was observed for Mut1, in which the other presumptive Sox3 SUMOylation site was deleted, Mut7 seemed to present a loss of activation function compared to wt Sox3. This suggests that SUMOylation of the two different presumptive sites could have opposite effects inducing Sox3 to act as a transcriptional repressor (site deleted in Mut1) or as activator (site deleted in Mut7). However, it is not yet known if SUMOylation occurs at either of these sites.

4.3.5. Mutant 10

Figure 4.9 Schematic representation of *sox3* showing the region deleted in Mut10.

Mutant 10 contained a deletion of 10 amino acids located at the very end of the C-terminal region of the protein (V291-I300, Fig. 4.9). The over-expression of *mut10* caused a unique phenotype in nearly half of the embryos analysed 24 hpf (Fig. 4.8). This phenotype differed from all the others found in the same experiment and presented a distribution of *ncad* transcript that resembled the one expected for embryos at approximately bud stage (Harrington *et al.* 2007; Warga and Kane 2007; Dee *et al.* 2008). Normally, at this stage the dorsal epiblast becomes thicker anteriorly and forms the neural plate. The neural plate is the earliest recognizable dorsal ectodermal primordium of the CNS and forms near the end of gastrulation, at about 9-10 hpf. Therefore it seems that the over-expression of Mut10 affected the development of the CNS in a different way compared to wtSox3. Instead of leading to defects that involve duplication or ectopic expression of neural tissue, Mut10 seemed to block development at an earlier stage, when the first primordium of the nervous system begins to form. In particular, it seems possible that embryos injected with Mut10 presented defects in the extension and conversion cellular movements that occur during gastrulation, and therefore retarded their development before the completion of gastrulation. This hypothesis is consistent with the appearance of these embryos (Fig. 4.10).

Interestingly, at sphere stage embryos over-expressing *mut10* did not show any particular difference in the expression of the organizer markers, compared to the embryos over-expressing *wt sox3*. Therefore, it seems that the mechanisms that were affected by the deletion and that caused the particular phenotype observed 24 hpf are related to processes that occur after the sphere stage.

The reading of the luciferase activity was lower in embryos over-expressing Mut10 than embryos over-expressing wt Sox3. This suggests that the deletion caused a loss in the transcriptional activator function of the protein.

Together, the data obtained by the analysis of the effects caused by the over-expression of Mut10 suggest that this mutant interferes with the development of the CNS at a later stage than the sphere stage, presumably around the bud stage, and that it lost Sox3 activator function. Therefore, the data confirmed the observation that the C-terminal region of Sox3 contains a trans-activation domain (Bowles *et al.* 2000; Xia *et al.* 2000; Sutton *et al.* 2011). As mentioned in the Introduction (Chapter 1.5.3), Sox3 acts in *Xenopus* as a transcriptional activator during the formation of the nervous system inducing the neural progenitor markers *sox2* and *geminin* (Rogers *et al.* 2009). It is possible that the deletion made in Mut10 interfered with the activation of target genes, such as *sox2,* which are normally expressed after the sphere stage in zebrafish and are involved in the specification of the neural tissue.

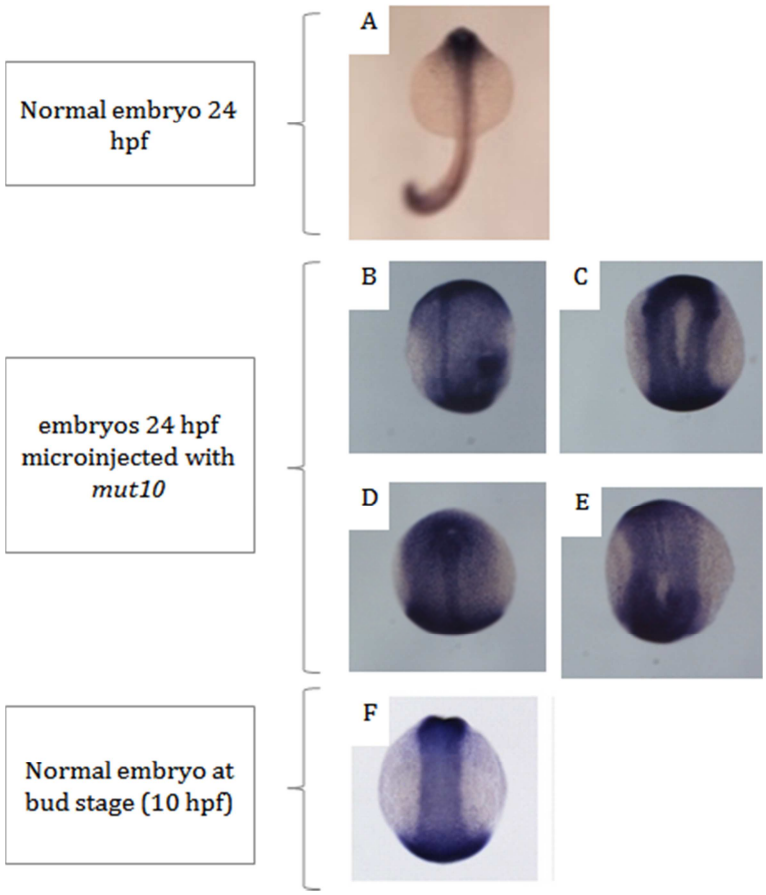


Figure 4.10 Microinjection of *mut10* caused defects in the development of the CNS, which seems not to develop beyond the bud stage. The figure shows the expression of *ncad* in a normal embryo 24 hpf (A), in embryos 24 hpf injected with *mut10* (B, C, D, E) and in a normal embryo at bud stage (F, (Dee *et al.* 2008)). Each picture shows a different embryo viewed from the dorsal side. It is clearly visible that the embryos 24 hpf microinjected with *mut10* presented a phenotype remarkably different to a normal embryo 24 hpf, but similar to the phenotype of an embryo at bud stage.

4.3.6. Mutant 11

Figure 4.11 Schematic representation of *sox3* showing the region deleted in Mut11.

Mutant11 contained a deletion of 11 amino acids that included all the N-terminal region of the protein, except the initial methionine, but not the HMG domain (Fig. 4.11). Interestingly, over-expression of Mut11 did not caused the repression of organizer markers compared to wt Sox3. This result was unexpected because, while it is thought that the C-terminal region of Sox3 contains a trans-activation domain, no particular functions have yet been associated with the short N-terminal end of the protein. The over-expression of Mut11 induced a fivefold increase in the activity of the luciferase and this would suggest that the deleted region is involved in Sox3 repressor function. Together these data suggest the hypothesis that the short N-terminal region could be involved in Sox3 repressor function.

Although the HMG domain is clearly involved in DNA binding, there is also evidence that it can mediate interactions with other proteins (Harley *et al.* 1996; Wilson and Koopman 2002; Zhang *et al.* 2003). Therefore, another hypothesis that would explain the effects caused by the over-expression of Mut11 is that the deleted region is normally involved, together with the HMG box, in the interaction with a co-repressor. The disruption of such interaction would cause the loss of Sox3 repressor function.

4.4. Mapping Different Functions of Zebrafish Sox3

The data obtained with this study allowed drawing a presumptive functional map of Sox3 that is represented Table 4.2.

Mutant	Deletion	Function	Experiments	Hypotheses
Mut11	Y2-I13	Transcriptional repression	• Expression of Boz and Gsc • Luciferase Reporter Assay	• Binding of co-repressors together with the HMG box
Mut1	P100-G123	Transcriptional repression	• Expression of Boz and Gsc • Luciferase Reporter Assay	• DNA binding and/or nuclear localisation • Interaction with Wnt/β-catening signalling • SUMOylation induced repressor function
Mut5	A164-M179	Transcriptional repression	• Expression of Boz • Luciferase Reporter Assay	• Binding with Groucho
Mut6	H180-N200	Transcriptional repression	• Expression of Boz and Gsc	• Binding with Groucho
Mut7	A226-G250	Transcriptional activation	• Luciferase Reporter Assay	• SUMOylation induced activator function
Mut10	V291-I300	Transcriptional activation	• Luciferase Reporter Assay	• Trans-activation

Table 4.2 The region deleted in each of the construct was associated with a presumptive function and one or more hypothesis that would explain these functions. The first column of the table indicates the name of the constructs (Mutant), the second indicates the regions deleted (Deletion), the third the presumptive function of Sox3 associated with the deleted region (Function), the fourth the experiments that suggested the functions and the last column indicates the hypothesis suggested that would explain the presumptive functions. The bar at the left represents the sequence of wt Sox3 (N-terminus upwards and C-terminus downwards). The deletions made in each mutant are indicated with a number that refers to the name of the constructs. The regions with presumptive repressor function are represented in red while the regions with presumptive activator function are represented in green.

4.5. General Discussion and Future Studies

Currently, the mechanisms involved in the modulation of Sox3 function as a transcriptional activator or repressor in early embryogenesis are still to be solved. It is certain that the HMG domain is involved in the binding of the DNA, but there is no certain evidence about the functions mediated by other regions of the protein. The functional screening strategy that was designed and developed in this study allowed the identification and mapping of large regions that are located outside the HMG box and presumably associated with the functioning of Sox3 as an activator or as a repressor. The consistency of the data obtained with the observations found in the literature is encouraging; therefore this study could be continued with the aim of drawing a more complete functional map of Sox3. The first step would be the repetition of the experiments in order to get a statistical confirmation of the data observed. Moreover, other experimental approaches are now being considered in order to meet two main objectives: firstly, the validation, or the rejection, of the hypothesis presented that would explain the function of the regions identified; experiments focused on the study of protein-protein interactions, such as the interaction between Sox3 and the co-repressor Groucho, or the interaction of Sox3 with β-catenin, are now being designed to meet this first objective. Secondly, the aim of further studies will be the identification of the specific residues involved in the functions associated with the regions identified in this study; for this purpose a point mutation approach is being considered.

5. Bibliography

Appel B. (2000). "Zebrafish neural induction and patterning." *Dev Dyn* **219**: 155-168.

Ashkenazy H., Erez E., Martz E., Pupko T. and Ben-Tal N. (2010). "ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids." *Nucleic Acids Res* **38**: W529-533.

Avilion A. A., Nicolis S. K., Pevny L. H., Perez L., Vivian N. and Lovell-Badge R. (2003). "Multipotent cell lineages in early mouse development depend on SOX2 function." *Genes Dev* **17**: 126-140.

Barut B. and Zon L. (2000). "Realizing the potential of zebrafish as a model for human disease." *Physiol. Genomics* **2**: 49-51.

Beddington R. S. (1994). "Induction of a second neural axis by the mouse node." *Development* **120**: 613-620.

Bell D. M., Leung K. K., Wheatley S. C., Ng L. J., Zhou S., Ling K. W., Sham M. H., Koopman P., Tam P. P. and Cheah K. S. (1997). "SOX9 directly regulates the type-II collagen gene." *Nat Genet* **16**: 174-178.

Bernadt C. T., Nowling T. and Rizzino A. (2004). "Transcription factor Sox-2 inhibits co-activator stimulated transcription." *Mol Reprod Dev* **69**: 260-267.

Bowles J., Schepers G. and Koopman P. (2000). "Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators." *Dev Biol* **227**: 239-255.

Brantjes H., Roose J., van De Wetering M. and Clevers H. (2001). "All Tcf HMG box transcription factors interact with Groucho-related co-repressors." *Nucleic Acids Res* **29**: 1410-1419.

Bylund M., Andersson E., Novitch B. G. and Muhr J. (2003). "Vertebrate neurogenesis is counteracted by Sox1-3 activity." *Nat Neurosci* **6**: 1162-1168.

Chakravarthy H., Boer B., Desler M., Mallanna S. K., McKeithan T. W. and Rizzino A. (2008). "Identification of DPPA4 and other genes as putative Sox2:Oct-3/4 target genes using a combination of in silico analysis and transcription-based assays." *J Cell Physiol* **216**: 651-662.

Cho K. W., Blumberg B., Steinbeisser H. and De Robertis E. M. (1991). "Molecular nature of Spemann's organizer: the role of the Xenopus homeobox gene goosecoid." *Cell* **67**: 1111-1120.

Collignon J., Sockanathan S., Hacker A., Cohen-Tannoudji M., Norris D., Rastan S., Stevanovic M., Goodfellow P. N. and Lovell-Badge R. (1996). "A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2." *Development* **122**: 509-520.

Connor F., Cary P. D., Read C. M., Preston N. S., Driscoll P. C., Denny P., Crane-Robinson C. and Ashworth A. (1994). "DNA binding and bending properties of the post-meiotically expressed Sry-related protein Sox-5." *Nucleic Acids Res* **22**: 3339-3346.

De Roberts E. M., Blum M., Niehrs C. and Steinbeisser H. (1992). "Goosecoid and the organizer." *Dev Suppl*: 167-171.

Dee C. T., Gibson A., Rengifo A., Sun S. K., Patient R. K. and Scotting P. J. (2007). "A change in response to Bmp signalling precedes ectodermal fate choice." *Int J Dev Biol* **51**: 79-84.

Dee C. T., Hirst C. S., Shih Y. H., Tripathi V. B., Patient R. K. and Scotting P. J. (2008). "Sox3 regulates both neural fate and differentiation in the zebrafish ectoderm." *Dev Biol* **320**: 289-301.

Denny P., Swift S., Brand N., Dabhade N., Barton P. and Ashworth A. (1992). "A conserved family of genes related to the testis determining gene, SRY." *Nucleic Acids Res* **20**: 2887.

Denny P., Swift S., Connor F. and Ashworth A. (1992). "An SRY-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA-binding protein." *EMBO J* **11**: 3705-3712.

Dodd A., Curtis P., Williams L. and Love D. (2000). "Zebrafish: bridging the gap between development and disease." *Hum. Mol. Genet.* **9**: 2443-2449.

Dooley K. and Zon L. I. (2000). "Zebrafish: a model system for the study of human disease." *Curr Opin Genet Dev* **10**: 252-256.

Driever W., Solnica-Krezel L., Abdelilah S., Meyer D. and Stemple D. (1997). "Genetic analysis of pattern formation in the zebrafish neural plate." *Cold Spring Harb Symp Quant Biol* **62**: 523-534.

Driever W., Solnica-Krezel L., Schier A. F., Neuhauss S. C., Malicki J., Stemple D. L., Stainier D. Y., Zwartkruis F., Abdelilah S., Rangini Z., Belak J. and Boggs C. (1996). "A genetic screen for mutations affecting embryogenesis in zebrafish." *Development* **123**: 37-46.

Fantes J., Ragge N. K., Lynch S. A., McGill N. I., Collin J. R., Howard-Peebles P. N., Hayward C., Vivian A. J., Williamson K., van Heyningen V. and FitzPatrick D. R. (2003). "Mutations in SOX2 cause anophthalmia." *Nat Genet* **33**: 461-463.

Fernandez-Lloris R., Osses N., Jaffray E., Shen L. N., Vaughan O. A., Girwood D., Bartrons R., Rosa J. L., Hay R. T. and Ventura F. (2006).

"Repression of SOX6 transcriptional activity by SUMO modification." *FEBS Lett* **580**: 1215-1221.

Ferrari S., Harley V. R., Pontiggia A., Goodfellow P. N., Lovell-Badge R. and Bianchi M. E. (1992). "SRY, like HMG1, recognizes sharp angles in DNA." *EMBO J* **11**: 4497-4506.

Gill G. (2005). "Something about SUMO inhibits transcription." *Curr Opin Genet Dev* **15**: 536-541.

Gill G. (2005). "SUMO changes Sox for developmental diversity." *Mol Cell* **20**: 495-496.

Girard F., Cremazy F., Berta P. and Renucci A. (2001). "Expression pattern of the Sox31 gene during zebrafish embryonic development." *Mech Dev* **100**: 71-73.

Girard M. and Goossens M. (2006). "Sumoylation of the SOX10 transcription factor regulates its transcriptional activity." *FEBS Lett* **580**: 1635-1641.

Graham V., Khudyakov J., Ellis P. and Pevny L. (2003). "SOX2 functions to maintain neural progenitor identity." *Neuron* **39**: 749-765.

Gritsman K., Talbot W. S. and Schier A. F. (2000). "Nodal signaling patterns the organizer." *Development* **127**: 921-932.

Gubbay, Collignon, Koopman, Capel, Economou, Münsterberg, Vivian, Goodfellow and Lovell-Badge (1990). "A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes." *Nature Reviews Genetics* **346**: 245-250.

Guille M. (1999). *Molecular Methods in Developmental Biology: Xenopus and Zebrafish*. Totowa, New Jersey Humana Press. Inc.

Hacker A., Capel B., Goodfellow P. and Lovell-Badge R. (1995). "Expression of Sry, the mouse sex determining gene." *Development* **121**: 1603-1614.

Halbleib J. M. and Nelson W. J. (2006). "Cadherins in development: cell adhesion, sorting, and tissue morphogenesis." *Genes Dev* **20**: 3199-3214.

Haldin C. E. and LaBonne C. (2010). "SoxE factors as multifunctional neural crest regulatory factors." *Int J Biochem Cell Biol* **42**: 441-444.

Hammerschmidt M., Blader P. and Strahle U. (1999). "Strategies to perturb zebrafish development." *Methods Cell Biol* **59**: 87-115.

Hammerschmidt M., Pelegri F., Mullins M. C., Kane D. A., van Eeden F. J., Granato M., Brand M., Furutani-Seiki M., Haffter P., Heisenberg C. P., Jiang Y. J., Kelsh R. N., Odenthal J., Warga R. M. and Nusslein-Volhard

C. (1996). "dino and mercedes, two genes regulating dorsal development in the zebrafish embryo." *Development* **123**: 95-102.

Harley V. R., Lovell-Badge R. and Goodfellow P. N. (1994). "Definition of a consensus DNA binding site for SRY." *Nucleic Acids Res* **22**: 1500-1501.

Harley V. R., Lovell-Badge R., Goodfellow P. N. and Hextall P. J. (1996). "The HMG box of SRY is a calmodulin binding domain." *FEBS Lett* **391**: 24-28.

Harrington M. J., Hong E., Fasanmi O. and Brewster R. (2007). "Cadherin-mediated adhesion regulates posterior body formation." *BMC Dev Biol* **7**: 130.

Heeg-Truesdell E. and LaBonne C. (2006). "Neural induction in Xenopus requires inhibition of Wnt-beta-catenin signaling." *Dev Biol* **298**: 71-86.

Hyatt T. M. and Ekker S. C. (1999). "Vectors and techniques for ectopic gene expression in zebrafish." *Methods Cell Biol* **59**: 117-126.

Inoue M., Kamachi Y., Matsunami H., Imada K., Uchikawa M. and Kondoh H. (2007). "PAX6 and SOX2-dependent regulation of the Sox2 enhancer N-3 involved in embryonic visual system development." *Genes Cells* **12**: 1049-1061.

Jager M., Queinnec E., Houliston E. and Manuel M. (2006). "Expansion of the SOX gene family predated the emergence of the Bilateria." *Mol Phylogenet Evol* **39**: 468-477.

Kamachi Y. (1996). "Involvement of SOX proteins in activation of crystallin genes and lens development." *EMBO J* **41**: 1113-1123.

Kamachi Y., Cheah K. S. and Kondoh H. (1999). "Mechanism of regulatory target selection by the SOX high-mobility-group domain proteins as revealed by comparison of SOX1/2/3 and SOX9." *Mol Cell Biol* **19**: 107-120.

Kamachi Y., Sockanathan S., Liu Q., Breitman M., Lovell-Badge R. and Kondoh H. (1995). "Involvement of SOX proteins in lens-specific activation of crystallin genes." *EMBO J* **14**: 3510-3519.

Kamachi Y., Uchikawa M., Collignon J., Lovell-Badge R. and Kondoh H. (1998). "Involvement of Sox1, 2 and 3 in the early and subsequent molecular events of lens induction." *Development* **125**: 2521-2532.

Kamachi Y., Uchikawa M. and Kondoh H. (2000). "Pairing SOX off: with partners in the regulation of embryonic development." *Trends Genet* **16**: 182-187.

Kamachi Y., Uchikawa M., Tanouchi A., Sekido R. and Kondoh H. (2001). "Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development." *Genes Dev* **15**: 1272-1286.

Kan L., Israsena N., Zhang Z., Hu M., Zhao L. R., Jalali A., Sahni V. and Kessler J. A. (2004). "Sox1 acts through multiple independent pathways to promote neurogenesis." *Dev Biol* **269**: 580-594.

Kane D. A. and Kimmel C. B. (1993). "The zebrafish midblastula transition." *Development* **119**: 447-456.

Kawahara A., Wilm T., Solnica-Krezel L. and Dawid I. B. (2000). "Antagonistic role of vega1 and bozozok/dharma homeobox genes in organizer formation." *Proc Natl Acad Sci U S A* **97**: 12121-12126.

Kelly, Chin, Leatherman, Kozlowski and Weinberg (2000). "Maternally controlled (beta)-catenin-mediated signaling is required for organizer formation in the zebrafish." *Development* **127**: 3899-3991.

Kimmel C. B., Ballard W. W., Kimmel S. R., Ullmann B. and Schilling T. F. (1995). "Stages of embryonic development of the zebrafish." *Dev Dyn* **203**: 253-310.

Kintner C. (1992). "Molecular bases of early neural development in Xenopus embryos." *Annu Rev Neurosci* **15**: 251-284.

Klymkowsky M. W. (1996). "Intermediate filaments as dynamic structures." *Cancer Metastasis Rev* **15**: 417-428.

Klymkowsky M. W., Rossi C. C. and Artinger K. B. (2010). "Mechanisms driving neural crest induction and migration in the zebrafish and Xenopus laevis." *Cell Adh Migr* **4**: 595-608.

Koopman P., Gubbay J., Vivian N., Goodfellow P. and Lovell-Badge R. (1991). "Male development of chromosomally female mice transgenic for Sry." *Nature* **351**: 117-121.

Koster R. W., Kuhnlein R. P. and Wittbrodt J. (2000). "Ectopic Sox3 activity elicits sensory placode formation." *Mech Dev* **95**: 175-187.

Koyano S., Ito M., Takamatsu N., Takiguchi S. and Shiba T. (1997). "The Xenopus Sox3 gene expressed in oocytes of early stages." *Gene* **188**: 101-107.

Kudoh, Concha, Houart, Dawid and Wilson (2004). "Combinatorial Fgf and Bmp signalling patterns the gastrula ectoderm into prospective neural and epidermal domains." *Development* **131**: 3581-3592.

Kuo C.-L., Lam C. M., Hewitt J. E. and Scotting P. J. (2013). "Formation of the Embryonic Organizer Is Restricted by the Competitive Influences of Fgf Signaling and the SoxB1 Transcription Factors." *PLoS One* **8**.

Laghari Z. A. (2010). *Sox3 as a transcriptional repressor*. Doctor of Philosophy, University of Nottingham.

Larroux C., Luke G. N., Koopman P., Rokhsar D. S., Shimeld S. M. and Degnan B. M. (2008). "Genesis and expansion of metazoan transcription factor gene classes." *Mol Biol Evol* **25**: 980-996.

Laudet V., Stehelin D. and Clevers H. (1993). "Ancestry and diversity of the HMG box superfamily." *Nucleic Acids Res* **21**: 2493-2501.

Lefebvre V., Dumitriu B., Penzo-Mendez A., Han Y. and Pallavi B. (2007). "Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors." *Int J Biochem Cell Biol* **39**: 2195-2214.

Leung T., Bischof J., Soll I., Niessing D., Zhang D., Ma J., Jackle H. and Driever W. (2003). "bozozok directly represses bmp2b transcription and mediates the earliest dorsoventral asymmetry of bmp2b expression in zebrafish." *Development* **130**: 3639-3649.

Linker and Stern (2004). "Neural induction requires BMP inhibition only as a late step, and involves signals other than FGF and Wnt antagonists." *Development* **131**: 5671-5681.

Liu Y.-R. (2011). *Function of Sox2 as a transcriptional repressor*. Doctor of Philosophy, University of Nottingham.

Merrill J. C., Melhuish T. A., Kagey M. H., Yang S. H., Sharrocks A. D. and Wotton D. (2010). "A role for non-covalent SUMO interaction motifs in Pc2/CBX4 E3 activity." *PLoS One* **5**: e8794.

Mertin S., McDowall S. G. and Harley V. R. (1999). "The DNA-binding specificity of SOX9 and other SOX proteins." *Nucleic Acids Res* **27**: 1359-1364.

Miyagi S., Kato H. and Okuda A. (2009). "Role of SoxB1 transcription factors in development." *Cell Mol Life Sci* **66**: 3675-3684.

Miyagi S., Masui S., Niwa H., Saito T., Shimazaki T., Okano H., Nishimoto M., Muramatsu M., Iwama A. and Okuda A. (2008). "Consequence of the loss of Sox2 in the developing brain of the mouse." *FEBS Lett* **582**: 2811-2815.

Mizuseki, Kishi, Matsui, Nakanishi and Sasai (1998). "Xenopus Zic-related-1 and Sox-2, two factors induced by chordin, have distinct activities in the initiation of neural induction." *Development* **125**: 579-587.

Mullins M. C., Hammerschmidt M., Kane D. A., Odenthal J., Brand M., van Eeden F. J., Furutani-Seiki M., Granato M., Haffter P., Heisenberg C. P., Jiang Y. J., Kelsh R. N. and Nusslein-Volhard C. (1996). "Genes establishing dorsoventral pattern formation in the zebrafish embryo: the ventral specifying genes." *Development* **123**: 81-93.

Munoz-Sanjuan I. and Brivanlou A. H. (2002). "Neural induction, the default model and embryonic stem cells." *Nature Reviews Neuroscience* **3**: 271-280.

Nambu J. R., Chen W., Hu S. and Crews S. T. (1996). "The Drosophila melanogaster similar bHLH-PAS gene encodes a protein related to human hypoxia-inducible factor 1 alpha and Drosophila single-minded." *Gene* **172**: 249-254.

Ng L. J., Wheatley S., Muscat G. E., Conway-Campbell J., Bowles J., Wright E., Bell D. M., Tam P. P., Cheah K. S. and Koopman P. (1997). "SOX9 binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse." *Dev Biol* **183**: 108-121.

Niehrs C. (2004). "Regionally specific induction by the Spemann–Mangold organizer." *Nature Reviews Genetics* **5**: 425-434.

Nishiguchi S., Wood H., Kondoh H., Lovell-Badge R. and Episkopou V. (1998). "Sox1 directly regulates the gamma-crystallin genes and is essential for lens development in mice." *Genes Dev* **12**: 776-781.

Nitta K. R., Takahashi S., Haramoto Y., Fukuda M., Onuma Y. and Asashima M. (2006). "Expression of Sox1 during Xenopus early embryogenesis." *Biochem Biophys Res Commun* **351**: 287-293.

Okuda Y., Ogura E., Kondoh H. and Kamachi Y. (2010). "B1 SOX coordinate cell specification with patterning and morphogenesis in the early zebrafish embryo." *PLoS Genet* **6**: e1000936.

Okuda Y., Yoda H., Uchikawa M., Furutani-Seiki M., Takeda H., Kondoh H. and Kamachi Y. (2006). "Comparative genomic and expression analysis of group B1 sox genes in zebrafish indicates their diversification during vertebrate evolution." *Dev Dyn* **235**: 811-825.

Paul T. Sharpe and Mason I. (1999). *Molecular Embryology : Methods and Protocols*. Totowa, New Jersey 07512, Humana Press Inc.

Penzel R., Oschwald R., Chen Y., Tacke L. and Grunz H. (1997). "Characterization and early embryonic expression of a neural specific transcription factor xSOX3 in Xenopus laevis." *Int J Dev Biol* **41**: 667-677.

Pevny L. and Placzek M. (2005). "SOX genes and neural progenitor identity." *Curr Opin Neurobiol* **15**: 7-13.

Pevny L. and Rao M. S. (2003). "The stem-cell menagerie." *Trends Neurosci* **26**: 351-359.

Pevny L. H. and Lovell-Badge R. (1997). "Sox genes find their feet." *Curr Opin Genet Dev* **7**: 338-344.

Pevny L. H., Sockanathan S., Placzek M. and Lovell-Badge R. (1998). "A role for SOX1 in neural determination." *Development* **125**: 1967-1978.

Phochanukul N. and Russell S. (2010). "No backbone but lots of Sox: Invertebrate Sox genes." *Int J Biochem Cell Biol* **42**: 453-464.

Pontiggia A., Rimini R., Harley V. R., Goodfellow P. N., Lovell-Badge R. and Bianchi M. E. (1994). "Sex-reversing mutations affect the architecture of SRY-DNA complexes." *EMBO J* **13**: 6115-6124.

Read C. M., Cary P. D., Preston N. S., Lnenicek-Allen M. and Crane-Robinson C. (1994). "The DNA sequence specificity of HMG boxes lies in the minor wing of the structure." *EMBO J* **13**: 5639-5646.

Rex M., Orme A., Uwanogho D., Tointon K., Wigmore P. M., Sharpe P. T. and Scotting P. J. (1997). "Dynamic expression of chicken Sox2 and Sox3 genes in ectoderm induced to form neural tissue." *Dev Dyn* **209**: 323-332.

Rex M., Uwanogho D. A., Orme A., Scotting P. J. and Sharpe P. T. (1997). "cSox21 exhibits a complex and dynamic pattern of transcription during embryonic development of the chick central nervous system." *Mech Dev* **66**: 39-53.

Rizzoti K., Brunelli S., Carmignac D., Thomas P. Q., Robinson I. C. and Lovell-Badge R. (2004). "SOX3 is required during the formation of the hypothalamo-pituitary axis." *Nat Genet* **36**: 247-255.

Rizzoti K. and Lovell-Badge R. (2007). "SOX3 activity during pharyngeal segmentation is required for craniofacial morphogenesis." *Development* **134**: 3437-3448.

Rogers C. D., Archer T. C., Cunningham D. D., Grammer T. C. and Casey E. M. (2008). "Sox3 expression is maintained by FGF signaling and restricted to the neural plate by Vent proteins in the Xenopus embryo." *Dev Biol* **313**: 307-319.

Rogers C. D., Harafuji N., Archer T., Cunningham D. D. and Casey E. S. (2009). "Xenopus Sox3 activates sox2 and geminin and indirectly represses Xvent2 expression to induce neural progenitor formation at the expense of non-neural ectodermal derivatives." *Mech Dev* **126**: 42-55.

Rosen J. N., Sweeney M. F. and Mably J. D. (2009). "Microinjection of zebrafish embryos to analyze gene function." *J Vis Exp*.

Ryu S. L., Fujii R., Yamanaka Y., Shimizu T., Yabe T., Hirata T., Hibi M. and Hirano T. (2001). "Regulation of dharma/bozozok by the Wnt pathway." *Dev Biol* **231**: 397-409.

Saúde L., Woolley K., Martin P., Driever W. and Stemple D. L. (2000). "Axis-inducing activities and cell fates of the zebrafish organizer." *Development* **127**: 3407-3417.

Savare J., Bonneaud N. and Girard F. (2005). "SUMO represses transcriptional activity of the Drosophila SoxNeuro and human Sox3 central nervous system-specific transcription factors." *Mol Biol Cell* **16**: 2660-2669.

Savare J. and Girard F. (2005). "SUMO modification represses transcriptional activity of Sox proteins." *Med Sci (Paris)* **21**: 917-919.

Schepers G. E., Teasdale R. D. and Koopman P. (2002). "Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families." *Dev Cell* **3**: 167-170.

Schier A. F. and Talbot W. S. (2005). "Molecular genetics of axis formation in zebrafish." *Annu Rev Genet* **39**: 561-613.

Schneider S., Steinbeisser H., Warga R. M. and Hausen P. (1996). "Beta-catenin translocation into nuclei demarcates the dorsalizing centers in frog and fish embryos." *Mech Dev* **57**: 191-198.

Shih J. and Fraser S. E. (1996). "Characterizing the zebrafish organizer: microsurgical analysis at the earlyshield stage." *Development* **122**: 1313-1322.

Shih Y. H., Kuo C. L., Hirst C. S., Dee C. T., Liu Y. R., Laghari Z. A. and Scotting P. J. (2010). "SoxB1 transcription factors restrict organizer gene expression by repressing multiple events downstream of Wnt signalling." *Development* **137**: 2671-2681.

Sinclair A. H., Berta P., Palmer M. S., Hawkins J. R., Griffiths B. L., Smith M. J., Foster J. W., Frischauf A. M., Lovell-Badge R. and Goodfellow P. N. (1990). "A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif." *Nature* **346**: 240-244.

Sinner D., Kordich J. J., Spence J. R., Opoka R., Rankin S., Lin S. C., Jonatan D., Zorn A. M. and Wells J. M. (2007). "Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells." *Mol Cell Biol* **27**: 7802-7815.

Sinner D., Rankin S., Lee M. and Zorn A. M. (2004). "Sox17 and beta-catenin cooperate to regulate the transcription of endodermal genes." *Development* **131**: 3069-3080.

Solnica-Krezel L. and Driever W. (2001). "The role of the homeodomain protein Bozozok in zebrafish axis formation." *Int J Dev Biol* **45**: 299-310.

Soriano N. S. and Russell S. (1998). "The Drosophila SOX-domain protein Dichaete is required for the development of the central nervous system midline." *Development* **125**: 3989-3996.

Spemann H. and Mangold H. (2001). "Induction of embryonic primordia by implantation of organizers from a different species. 1923." *Int J Dev Biol* **45**: 13-38.

Stachel S. E., Grunwald D. J. and Myers P. Z. (1993). "Lithium perturbation and goosecoid expression identify a dorsal specification pathway in the pregastrula zebrafish." *Development* **117**: 1261-1274.

Stemple D. L. (2005). "Structure and function of the notochord: an essential organ for chordate development." *Development* **132**: 2503-2512.

Stern (2006). "Neural induction: 10 years on since the 'default model'." *Curr. Opin. Cell Biol.* **18**: 692-697.

Stern C. D. (2005). "Neural induction: old problem, new findings, yet more questions." *Development* **132**: 2007-2021.

Streisinger G., Walker C., Dower N., Knauber D. and Singer F. (1981). "Production of clones of homozygous diploid zebra fish (Brachydanio rerio)." *Nature* **291**: 293-296.

Sutton E., Hughes J., White S., Sekido R., Tan J., Arboleda V., Rogers N., Knower K., Rowley L., Eyre H., Rizzoti K., McAninch D., Goncalves J., Slee J., Turbitt E., Bruno D., Bengtsson H., Harley V., Vilain E., Sinclair A., Lovell-Badge R. and Thomas P. (2011). "Identification of SOX3 as an XX male sex reversal gene in mice and humans." *J Clin Invest* **121**: 328-341.

Takeichi M. (1988). "The cadherins: cell-cell adhesion molecules controlling animal morphogenesis." *Development* **102**: 639-655.

Taylor K. M. and Labonne C. (2005). "SoxE factors function equivalently during neural crest and inner ear development and their activity is regulated by SUMOylation." *Dev Cell* **9**: 593-603.

Toyama R., O'Connell M. L., Wright C. V., Kuehn M. R. and Dawid I. B. (1995). "Nodal induces ectopic goosecoid and lim1 expression and axis duplication in zebrafish." *Development* **121**: 383-391.

Tsuruzoe S., Ishihara K., Uchimura Y., Watanabe S., Sekita Y., Aoto T., Saitoh H., Yuasa Y., Niwa H., Kawasuji M., Baba H. and Nakao M. (2006). "Inhibition of DNA binding of Sox2 by the SUMO conjugation." *Biochem Biophys Res Commun* **351**: 920-926.

Uchikawa M., Kamachi Y. and Kondoh H. (1999). "Two distinct subgroups of Group B Sox genes for transcriptional activators and repressors: their expression during embryonic organogenesis of the chicken." *Mech Dev* **84**: 103-120.

van de Wetering M. and Clevers H. (1992). "Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson-Crick double helix." *EMBO J* **11**: 3039-3044.

van de Wetering M., Oosterwegel M., van Norren K. and Clevers H. (1993). "Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes." *EMBO J* **12**: 3847-3854.

van Heyningen V. (1997). "Model organisms illuminate human genetics and disease." *Mol Med* **3**: 231-237.

Vriz S. and Lovell-Badge R. (1995). "The zebrafish Zf-Sox 19 protein: a novel member of the Sox family which reveals highly conserved motifs outside of the DNA-binding domain." *Gene* **153**: 275-276.

Warga R. M. and Kane D. A. (2007). "A role for N-cadherin in mesodermal morphogenesis during gastrulation." *Dev Biol* **310**: 211-225.

Wegner M. (1999). "From head to toes: the multiple facets of Sox proteins." *Nucleic Acids Res* **27**: 1409-1420.

Wegner M. (2010). "All purpose Sox: The many roles of Sox proteins in gene expression." *Int J Biochem Cell Biol* **42**: 381-390.

Weiss M. A. (2001). "Floppy SOX: mutual induced fit in hmg (high-mobility group) box-DNA recognition." *Mol Endocrinol* **15**: 353-362.

Williamson K. A., Hever A. M., Rainger J., Rogers R. C., Magee A., Fiedler Z., Keng W. T., Sharkey F. H., McGill N., Hill C. J., Schneider A., Messina M., Turnpenny P. D., Fantes J. A., van Heyningen V. and FitzPatrick D. R. (2006). "Mutations in SOX2 cause anophthalmia-esophageal-genital (AEG) syndrome." *Hum Mol Genet* **15**: 1413-1422.

Wilson M. and Koopman P. (2002). "Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators." *Curr Opin Genet Dev* **12**: 441-446.

Wood H. B. and Episkopou V. (1999). "Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages." *Mech Dev* **86**: 197-201.

Wright E. M., Snopek B. and Koopman P. (1993). "Seven new members of the Sox gene family expressed during mouse development." *Nucleic Acids Res* **21**: 744.

Xia Y., Papalopulu N., Vogt P. K. and Li J. (2000). "The oncogenic potential of the high mobility group box protein Sox3." *Cancer Res* **60**: 6303-6306.

Yan L. F. and Gu A. H. (2013). "[Progress and application of zebrafish in regenerative medicine]." *Yi Chuan* **35**: 856-866.

Yasuo H. and Lemaire P. (2001). "Role of Goosecoid, Xnot and Wnt antagonists in the maintenance of the notochord genetic programme in Xenopus gastrulae." *Development* **128**: 3783-3793.

Zhang C., Basta T., Hernandez-Lagunas L., Simpson P., Stemple D. L., Artinger K. B. and Klymkowsky M. W. (2004). "Repression of nodal expression by maternal B1-type SOXs regulates germ layer formation in Xenopus and zebrafish." *Dev Biol* **273**: 23-37.

Zhang C., Basta T., Jensen E. D. and Klymkowsky M. W. (2003). "The beta-catenin/VegT-regulated early zygotic gene Xnr5 is a direct target of SOX3 regulation." *Development* **130**: 5609-5624.

Zhang C. and Klymkowsky M. W. (2007). "The Sox axis, Nodal signaling, and germ layer specification." *Differentiation* **75**: 536-545.

Zhong L., Wang D., Gan X., Yang T. and He S. (2011). "Parallel expansions of Sox transcription factor group B predating the diversifications of the arthropods and jawed vertebrates." *PLoS One* **6**: e16570.

Zon L. I. (1999). "Zebrafish: a new model for human disease." *Genome Res* **9**: 99-100.

Zorn A. M., Barish G. D., Williams B. O., Lavender P., Klymkowsky M. W. and Varmus H. E. (1999). "Regulation of Wnt signaling by Sox proteins: XSox17 alpha/beta and XSox3 physically interact with beta-catenin." *Mol Cell* **4**: 487-498.

Zorn A. M., Butler K. and Gurdon J. B. (1999). "Anterior endomesoderm specification in Xenopus by Wnt/beta-catenin and TGF-beta signalling pathways." *Dev Biol* **209**: 282-297.

6. Media Recipes

6.1. Buffers for Bacterial Growth

6.1.1. Mu Broth culture medium

Bactotryptone	10g
Yeast Extract	5g
NaCl	10g
NaOH (1M)	2ml
H_2O	up to 1L
Total Volume	**1L**

6.1.2. Mu Broth Agar culture medium

Bactotryptone	10g
Yeast Extract	5g
NaCl	10g
NaOH (1M)	2ml
agar	1%
H_2O	up to 1L
Total Volume	**1L**

6.2. Buffers for Western Blot

6.2.1. SDS Running Buffer

Tris	30.3 g
Glycine	188 g
10% SDS (in SDW)	10 ml
SDW	up to 1L
Total Volume	**1L**

6.2.2. Transfer Buffer

MeOH	200 ml
Tris	3.03
Glycine	14.4
SDW	up to 1L
Total Volume	**1L**

6.3. Buffers for *In Situ* Hybridization

6.3.1. 20XSSC Solution

NaCl	876 g
Tri-Sodium Citrate dehydrate	441 g
SDW	up to 1L
Total Volume	**1L**

The pH of the solution was adjusted to 7.0.

All SSC solutions made from 20xSSC also contain 0.1% Tween20.

6.3.2. Hybridization Buffer

Formamide	250 ml
20x SSC	125 ml
tRNA (50 mg/ml)	0.5 ml
Heparin(100 mg/ml)	25 µl
Citric Acid (1M)	4.6 ml
Tween20 (20%)	2.5 ml
H_2O	112.5 ml
Total Volume	**500 ml**

6.3.3. Prehybridization Buffer

Before preparing the buffer 50 ml Hybridization Buffer were pre-cooled at -20˚C, while waiting for tRNA and Heparin to thaw on ice.

Hybridization Buffer	49.5 ml
tRNA (50 mg/ml)	0.5 ml
Heparin (100 mg/ml)	25 µl
Total Volume	**50 ml**

The buffer was stored at -20˚C.

6.3.4. Maleic Acid Buffer (MAB)

Maleic Acid	0.1 M
NaCl	0.15 M

The pH of the solution was adjusted to 7.5 using NaOH. The solution was then autoclaved.

6.3.5. Blocking Buffer

MAB + Boehringer Blocking Reagent™ 2% w/v, incubated at 80˚C overnight. The buffer was stored at -20˚C.

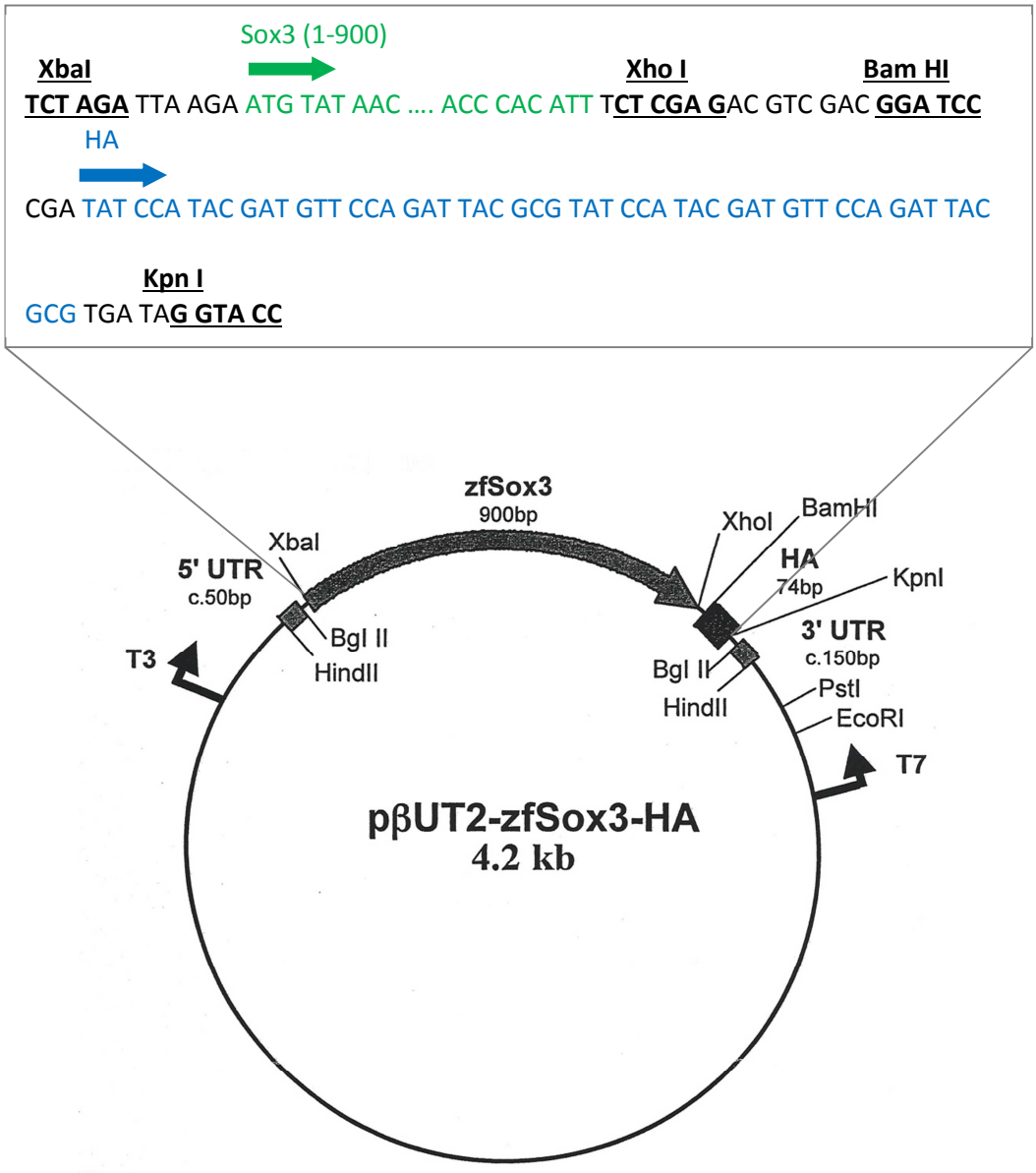
6.3.6. BCL Buffer III

1M Tris-HCl (pH 9.5**)**	5 ml
5M NaCl	1 ml
0.5M $MgCl_2$	5 ml
Tween20 (20%)	0.25 ml
SDW	up to 50 ml
Total Volume	**50 ml**

7.Appendix

7.1. Map of the *pβUT2-zfSox3-HA* Vector

The vector contains *Xenopus* β-globin 5'-UTR (Hind III – Bgl II blunt) and 3'-UTR from PSP64T (Bg I – Hind III) at either ends of Bluescribe's polylinker (*pβUT1*), replaced with Xba-Kpn with synthetic polylinker (*pβUT2*).

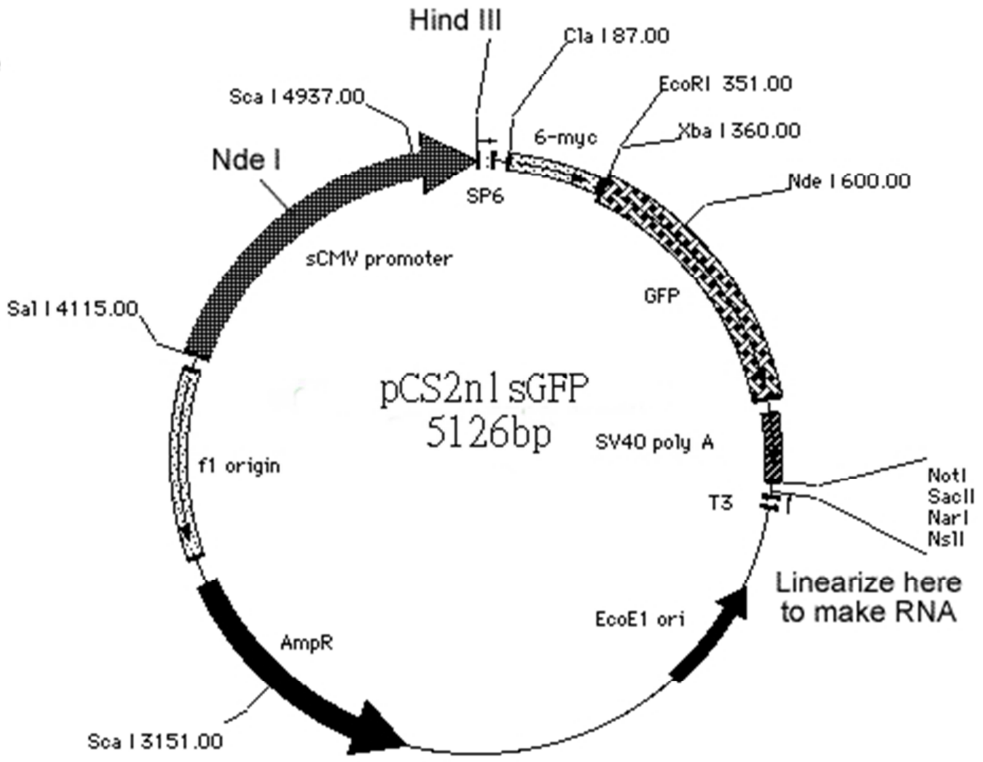


7.2. Map of the *pCS2nls.GFP* Vector

To create *pCS2nlsGFP*, an oligomer encoding a nuclear localisation signal derived from SV40, a large T antigen was inserted into plasmid pCS*mt-SGP (Klymkowsky 1996). Oligonucleotide sequences used were:

- nlsF: aat tcc cca aaa aag aag aga aag gta gaa t
- nlsR: cta gat tct acc ttt ctc ttc ttt ttt ggg g

More information on: pCS2*mt-SGP at http://spot.colorado.edu/~klym/



7.3. Primers Designed and Used for the Cloning of Deletion Mutant Sox3 Constructs

- **Mutant 1 (P100-G123)**

PRIMER **N'R**: ggc gtg ctc ctt cat gtg cat ggc

Length=24 T_m= 65˚C G+C= 63%

PRIMER **C'F**: ggc gga ctc ctg gcg ccc ggt

Length=21 T_m=70˚C G+C=81%

- **Mutant 2 (V139-L163)**

PRIMER **N'R**: ggc cga cac cgc gtt gtt gac agc

Length=24 T_m=67˚C G+C=67%

PRIMER **C'F**: ggc gcc tac cct caa cat ccc agc

Length=24 T_m=67˚C G+C=67%

- **Mutant 3 (A201-G221)**

PRIMER **N'R**: ggc gtt cat gta ggt ctg agc cgt

Length=24 T_m=64˚C G+C=58%

PRIMER **C'F**: ggc ttg ggc tcc atg gct tcg gtg

Length=24 T_m=67˚C G+C=67%

- **Mutant 4 (G124-G140)**

PRIMER **N'R**: ggc ccc tgg caa aga ata ctt gtc

Length= 24 T_m= 62˚C G+C= 54%

PRIMER **C'F**: ggc cag cgg atg gac tac acg cac

Length= 24 T_m=66˚C G+C=67%

- **Mutant 5 (A164-M179)**

PRIMER **N'R**: ggc cag ctg gtc ctg cat gag gga

Length= 24 T_m= 67˚C G+C= 67%

PRIMER **C'F**: ggc cac cgg tac gac atg gcg gga

Length=24 T_m=69˚C G+C=71%

- **Mutant 6 (H180-N200)**

PRIMER **N'R**: ggc cat ctg ctg gat ctg ggg gct

Length=24 T_m= 67˚C G+C= 67%

PRIMER **C'F**: ggc gcc gcg tcc acg tac agc agc

Length=24 T_m=70˚C G+C=75%

- **Mutant 7 (A226-G250)**

PRIMER **N'R**: ggc cat gga gcc caa acc cat

Length=21 T_m= 63˚C G+C=63 %

PRIMER **C'F**: ggc gac ctg aga gat atg ata agc

Length=24 T_m=57˚C G+C=50%

- **Mutant 8 (D251-G263)**

PRIMER **N'R**: ggc tcc caa aca agc acg ctg aga

Length=24 T_m= 64˚C G+C= 58%

PRIMER **C'F**: ggc gga gac agc gcc gac cac tcc

Length=24 T_m=70˚C G+C=75%

- **Mutant 9 (Q273-G290)**

PRIMER **N'R**: ggc tag act gga gtg gtc ggc gct

Length=24 T_m= 67˚C G+C= 67%

PRIMER **C'F**: ggc gtg aac gga acg cta ccc cta

Length=24 T_m=68˚C G+C=71%

- **Mutant 10 (V291-I300)**

PRIMER **N'R**: ggc gcc tgt ccc tgc gct ttg ata

Length=24 T_m= 66˚C G+C= 63%

PRIMER **C'F**: ggc tct cga gac gtc gac gga tcc

Length=24 T_m=65˚C G+C=67%

- **Mutant 11 (Y2-I13)**

PRIMER **N'R**: ggc cat tct taa tct aga gtc gat ctg

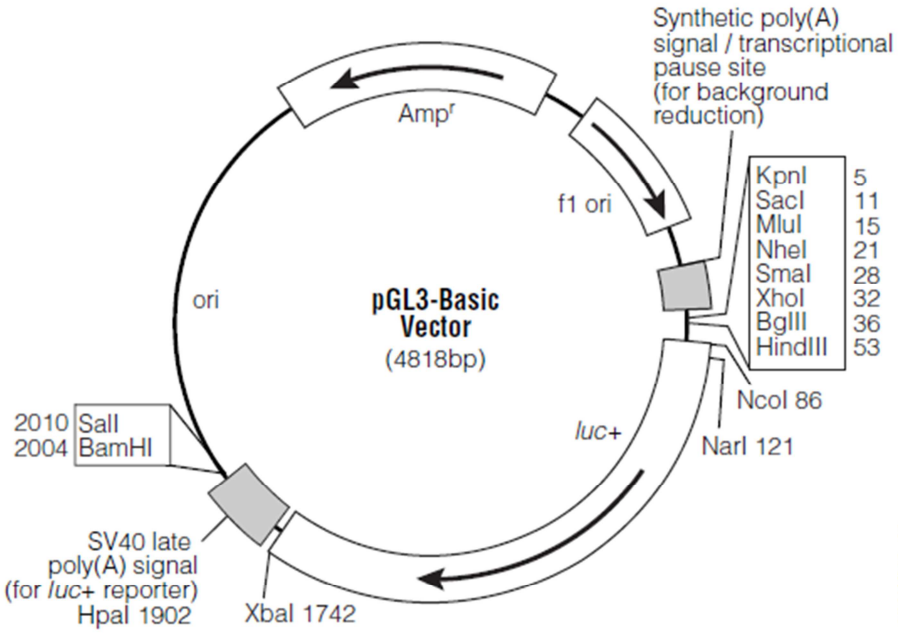
Length=27 T_m= 57˚C G+C= 42%

PRIMER **C'F**: ggc ccg cag tcc aac acg g

Length=19 T_m=65˚C G+C=74%

7.4. Map of the Reporter Plasmid Used for the Luciferase Reporter Assay

The basic *pGL3* vector was modified by insertion of F1 promoter followed by 3 repetitions of Sox3 binding sites and the *firefly luciferase* gene. The plasmid also contains the Ampicillin resistance gene. The insert is shown in the next page.



Insert of the *pGL3-3XSX* plasmid as follows:

TGATNTTCAGCATCTTTACTTTCNCCAGCGTTTCTGGGNGAGCAAAANCAGGAANGCAAAATGCCGC
AAAAAAGGGAATNANGGCGNCACGGAAATGTTGAATACTCATACTCTTCNTTTTTCAATATTATTGAA
GCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAG
GGGTTCCGCGCACATTTCCCCGAAAAGTGCCACNTGACGCGCCCTGTAGCGGCGCATTAAGCGCGGC
GGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTT
TCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGG
GTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTG
GGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCT
TGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGAT
TTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAAC
GCTTACAATTTGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTT
CGCTATTACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGTGGAGGTTTTACTTGCT
TTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTG
TTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTT
CACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATGGTACTGTAACTGAGCTAACATA
ACCCGGGAGGTACCGAGCTCTTACGCGTGCTAGCTCGAGATCCGCGCCTTTGTTCTCCCCAGATCCGC
KpnI SacI MluI NheI XhoI
GCCTTTGTTCTCCCCAGATCCGCGCCTTTGTTCTCCCCAGATCTACTTGGGCATAAAAGGCAGAGCAGG
GCAGCTGCTGNTAAGCTTGGCATTCCGGTACTGTTGGTAAAGCCACCATGGAAGACGCCAAAAACAT
HindIII
AAAGAAAGGCCCGGCGCCATTCTATCCGCTGGAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCT
ATGAAGAGATACGCCCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGGACATCAC
TTACGCTGAGTACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCTGAATACAA
ATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCCGGTGTTGGGCGCGTTATTTA
TCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAATGAACGTGAATTGCTCAACAGTATGGGCATT
TCGCAGCCTACCGTGGTGTTCGTTTCCAAAAAGGGGTTGCAAAAAATTTTGAACGTGCAAAAAAGCT
CCCAATCATCCAAAAAATTATTATCATGGATTCTAAAACGGATTACCAGGGATTTCAGTCGATGTACAC
GTTCGTCACATCTCATCTACCTCCCGGTTTTAATGAATACGATTTTGTGCCAGAGTCCTTCGATAGGGA
CAAGACAATTGCACTGATCATGAACTCCTCTGGATCTACTGGTCTGCCTAAAGGTGTCGCTCTGCCTCA
TAGAACTGCCTGCGTGAGATTCTCGCATGCCAGAGATCCTATTTTTGGCAATCAAATCATTCCGGATAC
TGCGATTTTAAGTGTTGTTCCATTCCATCACGGTTTTGGAATGTTTACTACACTCGGATATTTGATATGT
GGATTTCGAGTCGTCTTAATGTATAGATTTGAAGAAGAGCTGTTTCTGAGGAGCCTTCAGGATTACAA
GATTCAAAGTGCGCTGCTGGTGCCAACCCTATTCTCCTTCTTCGCCAAAAGCACTCTGATTGACAAATA
CGATTTATCTAATTTACACGAAATTGCTTCTGGTGGCGCTCCCCTCTCTAAGGAAGTCGGGGAAGCGG
TTGCCAAGAGGTTCCATCTGCCAGGTATCAGGCAAGGATATGGGCTCACTGAGACTACATCAGCTATT
CTGATT

Ampicillin resistance
F1 promoter
CTTTGTT 3X repeat of Sox3 binding sites
luciferase- firefly