1 Introduction

Opitz BBB/G syndrome (OS, MIM 300000 and MIM 145410) is a congenital disorder characterised by malformations of the ventral midline. BBB syndrome and G syndrome (Borden, 2000; Opitz, 1969b; Opitz, 1969a), formerly described as two different disorders, were merged into Opitz BBB/G syndrome subsequent to description of patients with a mixture of symptoms (Cordero and Holmes, 1978). The two cardinal phenotypic manifestations of OS are hypertelorism/telecanthus (abnormally widely spaced/increased innercanthal distance) and hypospadias (urethral cleft in male patients). However, they are often found to have in addition other typical features such as cleft lip and/or palate, laryngotracheoesophageal (LTE) anomalies, congenital heart defects, imperforate anus, agenesis of the corpus callosum, ear abnormalities, and mental retardation (Cox et al., 2000; De Falco et al., 2003; So et al., 2005).

OS is genetically heterogeneous with an autosomal and an X-linked locus. The two forms are clinically indistinguishable (Robin et al., 1996). In 1997, Quaderi et al identified by positional cloning the MID1 gene in Xp22.3; MID1 harbours mutations in about 68% of patients with X-linked OS. In contrast, while the autosomal locus has been mapped by linkage analysis to chromosome 22, the responsible gene remains unknown (Robin et al., 1995).

1.1 Ventral midline development

The development of the ventral midline of the human body is a highly complicated multistep process during embryogenesis which involves formation, development, and correct positioning of structures in the midline of the body (e.g. heart, oesophagus, trachea or craniofacial structures). Many of the molecular mechanisms involved in the development of the ventral midline are unidentified. However, neural crest cell (NCC) migration and ephithelial-mesenchymal transition (EMT) are thought to be fundamental processes during these stages of embryonic development (Schweiger and Schneider, 2003).

Due to their importance during embryogenesis, NCCs have been considered the fourth germ layer, in addition to the three traditional germ layers, ectoderm, mesoderm and endoderm (Hall, 2000). Moreover, some scientists consider that migrating embryonic NCCs are the best evolutionary acquisition from vertebrates (Thorogood, 1989).

NCCs derive from the ectoderm and originate during neurulation at the dorsal most region of the neural tube between the neural and non-neural ectoderm (Figure 1.1; Gilbert, 2003). NCCs delaminate from the neural folds and migrate to different places, where they generate a prodigious variety of cells, such as neurons, pigment, smooth muscle, connective tissue, or bone, among others. Although the mechanisms involved in NCC formation and migration are highly conserved among vertebrates, the precise moment in which NCCs start their journey differs between species. For instance, in mammals NCCs are induced and

migrate concomitantly with neural plate folding. In contrast, in frogs and birds NCCs start migrating only after neural tube closure (Jones and Trainor, 2005).

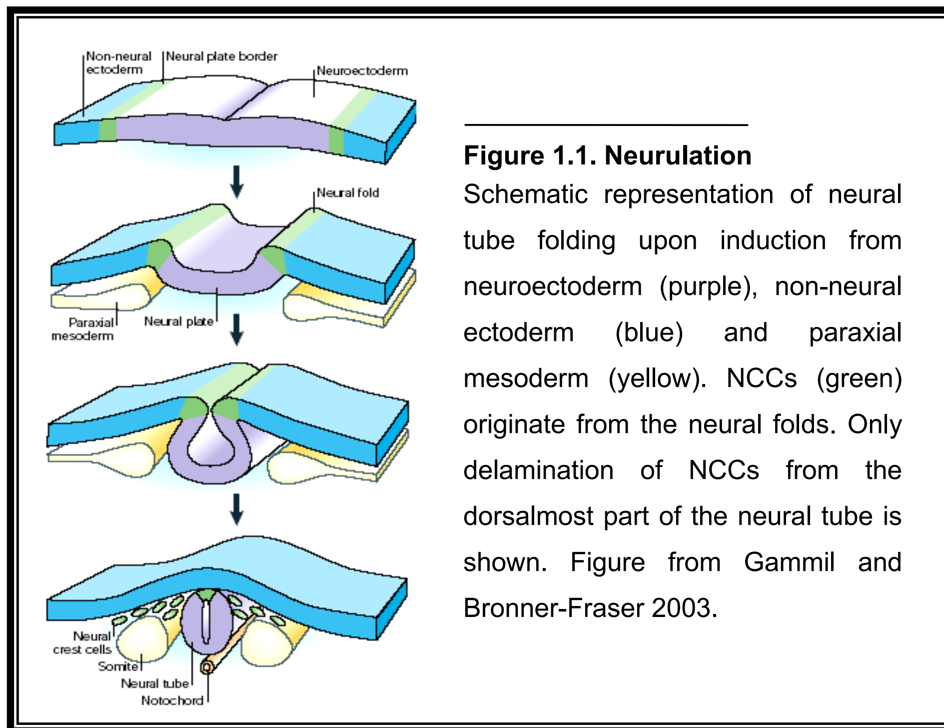

Figure 1.1. Neurulation
Schematic representation of neural tube folding upon induction from neuroectoderm (purple), non-neural ectoderm (blue) and paraxial mesoderm (yellow). NCCs (green) originate from the neural folds. Only delamination of NCCs from the dorsalmost part of the neural tube is shown. Figure from Gammil and Bronner-Fraser 2003.

NCCs are classified into four subpopulations that give rise to specific cell and tissue types (Gilbert, 2003; Jones and Trainor, 2005):

- **Trunk NCCs** differentiate into neurons of the glia and peripheral nervous system, part of the enteric nervous system, and melanocytes.
- **Cranial NCCs** give rise to cartilage, bone, cranial neurons, glia, and connective tissue of the face and, in addition, they produce pigment cells, and sensory and parasympathetic ganglia.
- **Cardiac NCCs** participate in the formation of the heart and, additionally, generate pigment, skeletal, and neuronal cells
- **Vagal NCCs** form nearly the entire gut, and the majority of neurons and glia of the enteric system.

Each population occupies a different region along the anterior-posterior axis of the neural tube, from which they migrate along unique pathways, that are defined by a Homeobox gene (Hox) gradient (Le Douarin et al., 2004; Lumsden et al., 1991; Selleck and Bronner-Fraser, 1995; Trainor and Krumlauf, 2001).

As previously mentioned, NCCs undergo different processes common to all kinds of NCCs and to all species. After induction of neural plate folding, some previously indistinguishable neuroepithelial cells are induced by signals coming from the presumptive epidermis and neural tube (Mancilla and Mayor, 1996; Selleck and Bronner-Fraser, 1995).

Induced cells then enter into EMT to finally delaminate and migrate (Duband et al., 1995; Knecht and Bronner-Fraser, 2002; Thiery, 2003). During all of these processes, contact-mediated interactions between neural and non-neural tissue take place. All these events take place under a very complex temporospacial collaboration of many different morphogens, such as Wnt, transforming growth factor β family (TGFβ) (including bone morphogenic protein, BMP), Hedgehog (HH) family, fibroblast growth factor (FGF) family, Notch, and even retinoic acid and its receptors play a role here (extensively reviewed in Gammill and Bronner-Fraser, 2003; Jones and Trainor, 2005; Knecht and Bronner-Fraser, 2002). All of these pathways participate repeatedly during the life of NCCs; NCC fate is not a sequence of events, but a combination of events and intercalation of pathways functioning concomitantly.

Nevertheless, BMP, Wnt, and FGF signalling by activating transcription factors such as Slug or FoxD3 seem to control the early steps in the induction of NCCs (Barembaum and Bronner-Fraser, 2005; Huang and Saint-Jeannet, 2004; Nieto et al., 1994). Slug is one of the best-characterised NCC induction markers; it belongs to the Snail/Slug family, and is closely related to Snail. Slug and Snail are both expressed in NCCs in a vertebrate-specific manner (Barrallo-Gimeno and Nieto, 2005; Locascio et al., 2002; Nieto et al., 1994).

After induction, NCCs undergo EMT, process during which epithelial cells lose their polarity and adhesion properties by down-regulation of cell adhesion molecules such as E-cadherin (Pla et al., 2001), transform into mesenchymal cells, and acquire migratory properties. During EMT, an important reorganization of cytoskeletal structures and increase in expression of mesenchymal genes take place (Hay, 2005; Savagner, 2001). Important key players in EMT include again Snail/Slug.

Premigratory NCCs also respond to signals coming from the neural plate and non-neural ectoderm that assure their survival. Induction of Foxd3, Sox9 and Sox10, Zic genes, Pax genes, Notch, Delta, Msx genes, Rhob and Twist is important here (Gammill and Bronner-Fraser, 2003; Jones and Trainor, 2005; Steventon et al., 2005). Adjacent territories also have a strong influence on the delamination of NCCs. Thus, although other pathways like Wnt and FGF also participate (Morales et al., 2005), correct balance between BMP and its antagonist Noggin has been shown to be decisive for correct delamination (Sela-Donenfeld and Kalcheim, 1999).

1.1.1 Craniofacial development

Craniofacial development is one of the most complicated processes during development, which involves the formation of a large number of different tissues and the correct emplacement of the structures of the skull and face (Cohen, 2002; Wilkie and Morriss-Kay, 2001). It begins with the formation of five paired prefacial structures known as branchial or pharyngeal arches, at the ventral surface of the embryo. Next, invading NCCs (Figure 1.2C) in combination with mesodermal cells form the facial prominences (maxilomandibular, lateral and frontonasal), which grow, fuse and expand into more mature craniofacial structures, including the skull (Figure 1.2A). From the first pair of pharyngeal arches emerge the maxillomandibular prominences (Figure 1.2A,B, in green), which give rise to the lateral middle and lower face, the sides of the lips and the lower jaw (only the mandibular prominence) (Cox, 2004; Tapadia et al., 2005). The frontonasal process (Figure 1.2A,B, in pink) contributes to the forehead, the

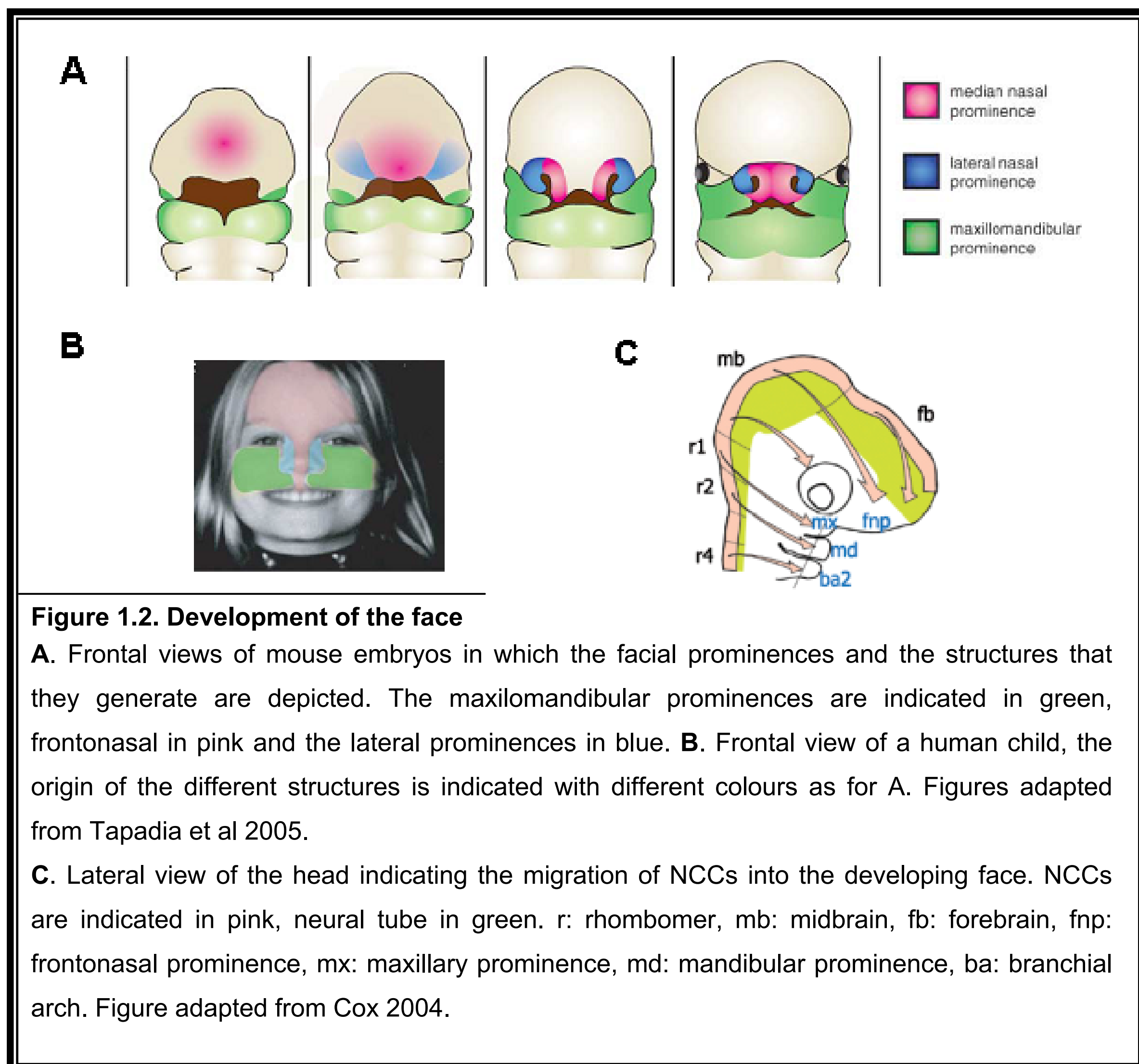

Figure 1.2. Development of the face

A. Frontal views of mouse embryos in which the facial prominences and the structures that they generate are depicted. The maxilomandibular prominences are indicated in green, frontonasal in pink and the lateral prominences in blue. **B**. Frontal view of a human child, the origin of the different structures is indicated with different colours as for A. Figures adapted from Tapadia et al 2005.

C. Lateral view of the head indicating the migration of NCCs into the developing face. NCCs are indicated in pink, neural tube in green. r: rhombomer, mb: midbrain, fb: forebrain, fnp: frontonasal prominence, mx: maxillary prominence, md: mandibular prominence, ba: branchial arch. Figure adapted from Cox 2004.

middle of the nose, and, in combination with the maxillary process, to the formation of the midlip and primary palate. Defective coordination of growth and fusion (governed by NCC invasion, EMT and programmed cell death (PCD)) of the prominences can result in cleft lip and palate (Cox, 2004; Kang and Svoboda, 2005; Kang and Massague, 2004; Schweiger and Schneider, 2003). The lateral prominences (Figure 1.2B, in blue) give rise to the lateral nasal structures.

Invading NCCs arising from different areas along the anterior-posterior neural axis (Figure 1.2C, Kontges and Lumsden, 1996) play an important role during craniofacial development. Positional identity along the anterior-posterior axis is obtained through a Hox gene gradient that defines each rhombomer (the parts into which the hindbrain is divided) (Krumlauf, 1993; Santagati and Rijli, 2003). A complex and cautiously regulated coordination of signalling pathways coming from both epithelia and NCCs are of high importance for a successful completion of craniofacial structures (Helms et al., 2005; Noden and Trainor, 2005; Tapadia et al., 2005).

1.1.2 Heart development

At embryonic day 21, the heart is a single chambered tube formed by fusion of two simple epithelial tubes (Figure 1.3A). By day 28, cardiac looping occurs and the presumptive atria are placed anterior to the presumptive ventricles. Upon myocardial signals, cells from the adjacent endocardium detach and form the endocardial cushions, which divide the primitive tube into the atrioventricular channels. Concomitantly, the primitive atrium becomes divided by two septa that grow towards the endocardial cushion. Up to the eighth week of gestation, there is a continuous process of looping and remodelling until the septa close to form the four definitive chambers of the heart, at which point the circulatory and pulmonary systems become independent, with the pulmonary artery connected to the right ventricle and the aorta to the left ventricle (Figure 1.3A, Gilbert, 2003; Schweiger and Schneider, 2003).

Cardiac NCCs, generated from the first to third somites, give rise to the smooth muscle of the aortic arches. In addition, they migrate through the third and more caudal branchial arches and invade the aorticopulmonary septum and endocardial cushions collaborating to achieve correct articopulmonary segmentation (Kirby et al., 1983; Waldo et al., 1998). Similarly to craniofacial development, induced cells producing the endocardial cushion (see above) undergo EMT (Runyan and Markwald, 1983), probably in parallel to PCD (Figure 1.3B, Keyes and Sanders, 2002). NCC invaded regions are commonly affected in patients with complex congenital disorders, such as OS.

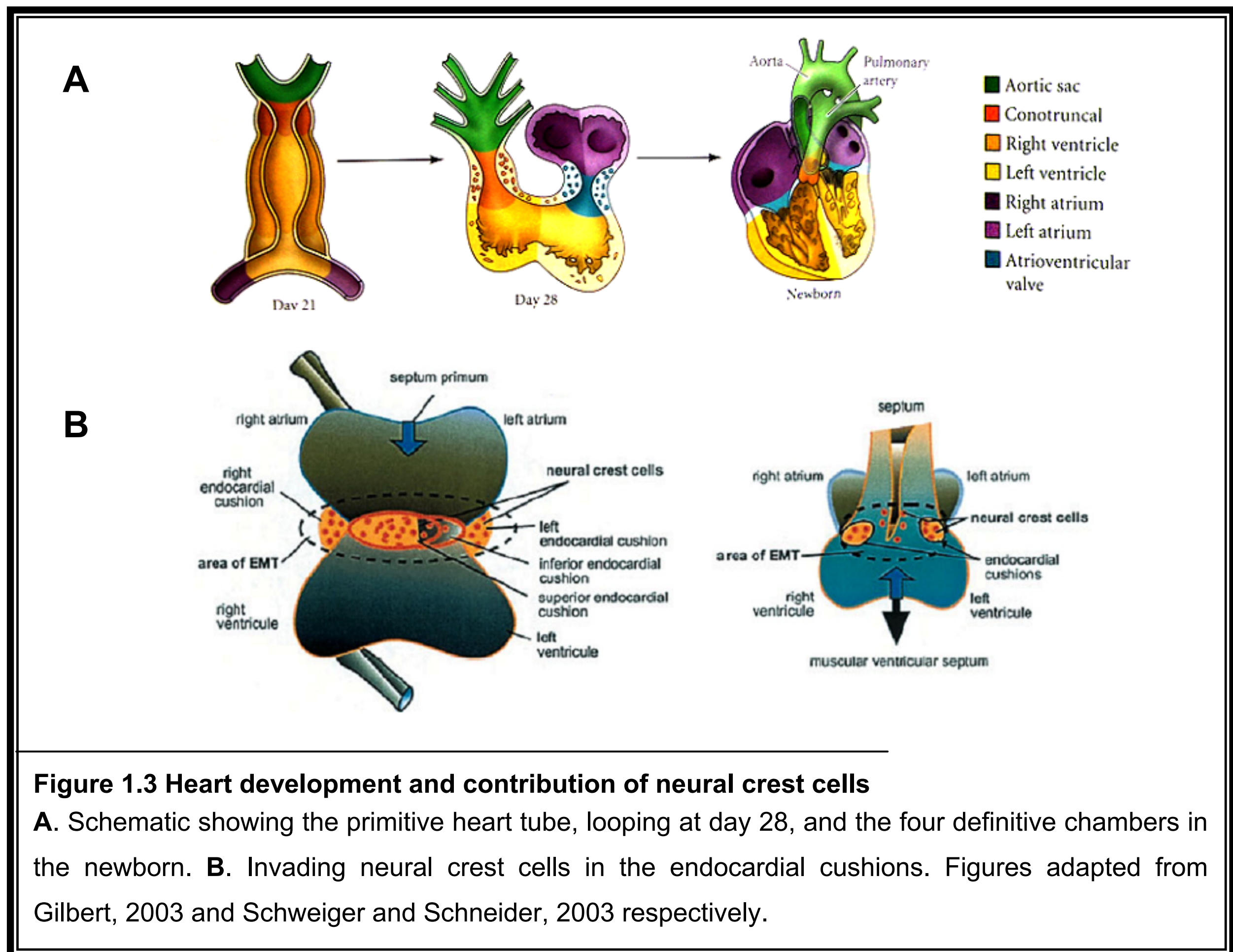

Figure 1.3 Heart development and contribution of neural crest cells
A. Schematic showing the primitive heart tube, looping at day 28, and the four definitive chambers in the newborn. **B**. Invading neural crest cells in the endocardial cushions. Figures adapted from Gilbert, 2003 and Schweiger and Schneider, 2003 respectively.

1.1.3 Ephrins as NCC pathfinders

NCCs migrate in well-separated streams during development until they reach their site of differentiation. In addition to the earlier mentioned Hox genes, ephrin receptors (Eph receptors) and ephrins have been shown to be essential in delineating migratory routes of NCCs. Obstruction of single Eph receptors can cause a mixture of different streams, with severe developmental consequences (Helbling et al., 1998; Robinson et al., 1997; Smith et al., 1997).

Eph receptors, the largest family of membrane-bound receptor tyrosine kinases, interact with ephrin ligands, which are also membrane-bound proteins. Two classes of Eph receptors, A and B, have been described, distinguished by their predominant binding affinity to ephrin-A ligands (anchored to the membrane through glycosylphosphatidylinositol anchor, GPI anchor) or to ephrin-B ligands (transmembrane), respectively. On the cellular level, the interaction of Eph receptors with their ligands principally influences cytoskeletal organization and cell adhesion, and, consequently, cell migration (Pasquale, 2005). Furthermore, many developmental processes, such as morphogenesis, tissue border formation, vascularization, axon guidance or synaptic plasticity, mainly depend on correct Eph receptor-ephrin signalling (Davy and Soriano, 2005; Holder and Klein, 1999; Murai and Pasquale, 2004; Palmer and

Klein, 2003). Since both Eph receptors and ephrins are membrane-bound proteins, cell-to-cell contact is required for them to interact. They interact in a “kiss and tell” fashion, thereby inducing forward and reverse signals that result in attraction and adhesion or repulsion of participating cells (Gauthier and Robbins, 2003; Kullander and Klein, 2002). For instance, ephrin-B1 can either attract or repel EphB-expressing NCCs (Santiago and Erickson, 2002) and, in this way, the flow of NCC streams is regulated.

Eph receptors and ephrins are powerful patterning molecules. Interestingly, ephrin defects have more drastic effects in some cells than in others. Differential expression of ephrin genes in cells that should be the same makes them behave differently, and in the absence of ephrin expression, all the cells remain the same and their functions can be accomplished by redundant molecules (Compagni et al., 2003; Davy et al., 2004). A clear example is craniofrontonasal syndrome, where the loss of expression of ephrin-B1 (*EFNB1*), leads specifically to defects of the ventral midline, while other structures, where EFNB1 differential expression is probably not needed, remain unaffected (Twigg et al., 2004; Wieland et al., 2004).

1.1.4 Midline development and disease

The development of the craniofacial structures is a very complicated process that is found altered in a multitude of congenital disorders (Wilkie and Morriss-Kay, 2001). Additionally, there is a high incidence (1 in 700 to 1 in 1000) of non-syndromic cleft lip and palate that involves a complex inheritance pattern (Prescott et al., 2001; Schutte and Murray, 1999). Interestingly, OS is one of the few known causes of monogenic cleft lip and palate.

Mutations of diverse genes participating in the establishment of the ventral midline result in overlapping phenotypes, demonstrating the complexity of this process and the concomitant participation of many pathways. An extensively studied example is holoprosencephaly (HPE) (Roessler and Muenke, 2001), which is characterised by defective forebrain development normally accompanied by cleft lip, cyclopia and overlying proboscis. Alterations in several members of the Sonic Hedgehog (Shh) pathway have been shown to cause HPE (Roessler and Muenke, 2003). In addition, mutations in GLI3, a target gene of Shh signalling, cause Greig syndrome (OMIM 175700), which is characterised by hypertelorism and polydactyly.

It has been suggested that defects in NCC migration and epithelial EMT lead to malformations in the ventral midline (Schweiger and Schneider, 2003). DiGeorge Syndrome (DGS; OMIM 188400), also referred as to del22q11, is characterised by aortic arch defects, contruncal heart defects, and pharyngeal and craniofacial anomalies, all structures derived from of NCCs. Since loss of *Tbx1* expression in mouse has been shown to generate a phenotype resembling DGS, *TBX1* has been proposed to be the key gene in DGS (Jerome and Papaioannou, 2001; Merscher et al., 2001). Loss of *Tbx1* expression has also been

shown to severely affect neural crest and cranial nerve migratory pathways (Moraes et al., 2005; Vitelli et al., 2002)

Mowat-Wilson syndrome (OMIM 235730) is an autosomal dominant complex disorder caused by mutations in the *ZFHX1B* gene. The phenotype of Mowat-Wilson patients is characterised by craniofacial malformations, Hirschsprung disease, mental retardation, congenital heart disease, hypospadias and agenesis of the corpus callosum; all features are suggestive of a neurocristopahty affecting cranial, cardiac and vagal NCCs (Mowat et al., 2003).

In the previously mentioned craniofrontonasal syndrome (OMIM 304110), which is characterised by hypertelorism and broad nasal bridge, mutations in the *EFNB1* gene also result in improper NCC migration (Twigg et al., 2004; Wieland et al., 2004). Moreover, defective NCC migration and EMT have also been proposed to be involved in the pathogenesis of OS (Cox, 2004; Richman et al., 2002; Short and Cox, 2006).

1.2 X-linked Opitz BBB/G syndrome (OS) and MID1

MID1 covers a genomic region of about 300 kb, from which nine exons are constitutively expressed to give an open reading frame (ORF) of approximately 2 kb (Figure 1.4). While the existence of 3´and 5´UTR alternative MID1 transcripts has been known for some time (Landry and Mager, 2002; Winter et al., 2004), 14 novel alternative exons were recently identified in the human MID1 gene (Figure 1.4, Landry and Mager, 2002; Winter et al., 2004).

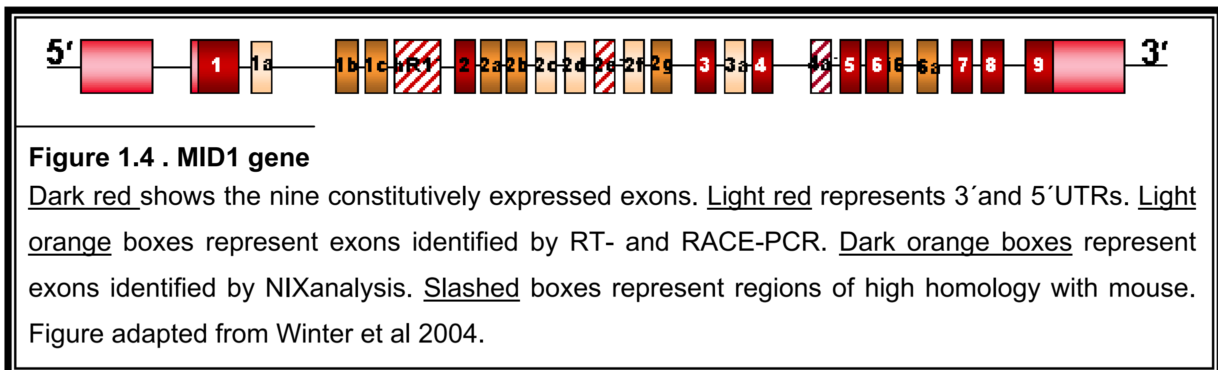

Figure 1.4 . MID1 gene
Dark red shows the nine constitutively expressed exons. Light red represents 3´and 5´UTRs. Light orange boxes represent exons identified by RT- and RACE-PCR. Dark orange boxes represent exons identified by NIXanalysis. Slashed boxes represent regions of high homology with mouse. Figure adapted from Winter et al 2004.

Most of the mutations identified in OS patients are located at the 3′ end of the ORF, hence affecting the C-terminus of MID1. With the detection of MID1 in 68% of X-linked OS patients, the majority of the identified mutations are nonsense and frameshift mutations. However, missense mutations, in-frame deletions and in-frame insertions have also been found (Figure 1.5, Schweiger and Schneider, 2003).

The ubiquitously expressed *MID1*-gene encodes a 667 amino acid phosphoprotein, MID1 (Figure 1.5), with a molecular weight of approximately 72 kDa. MID1, also known as TRIM18, belongs to the RBCC protein family, characterised by an N-terminal tripartite motif

(TRIM) and an RFP-like domain at the C-terminus. The RBCC domain comprises a RING finger, two Bboxes and a coiled-coil domain (Schweiger and Schneider, 2003).

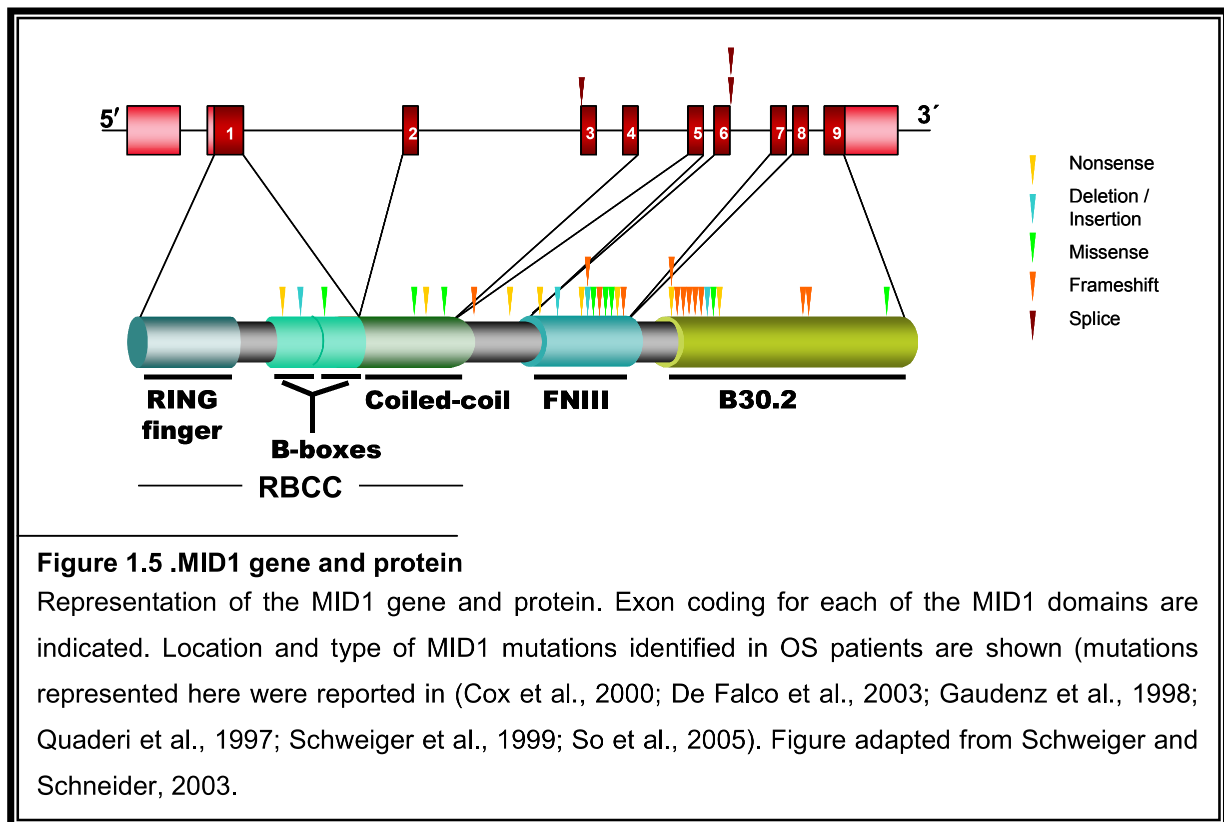

Figure 1.5 .MID1 gene and protein
Representation of the MID1 gene and protein. Exon coding for each of the MID1 domains are indicated. Location and type of MID1 mutations identified in OS patients are shown (mutations represented here were reported in (Cox et al., 2000; De Falco et al., 2003; Gaudenz et al., 1998; Quaderi et al., 1997; Schweiger et al., 1999; So et al., 2005). Figure adapted from Schweiger and Schneider, 2003.

1.2.1 The RBCC protein MID1

The first RBCC domain was identified in Xenopus nuclear factor 7 (XNF7), a protein that participates in the dorso-ventral patterning of Xenopus during development (El-Hodiri et al., 1997). The list of identified RBCC proteins has since then grown rapidly and today encompasses 68 members (Nisole et al., 2005; Reymond et al., 2001).

RBCC proteins participate in various important cellular processes such as development, cell growth, and proliferation. In addition, some of them have antiviral properties by targeting retroviruses (Nisole et al., 2005). However, they are recently gaining attention due to the increasing number of proteins in this family involved in human disease. Well-studied examples include the proto-oncogenic proteins RFP (ret finger protein) (Hasegawa et al., 1996), PML (promyelocytic leukaemia) and TIF1-α, which gain oncogenic properties when fused to RET, retionic acid receptor alpha (RARα), or Braf (Borden et al., 1995a; Grignani et al., 1994; Jensen et al., 2001; Klugbauer and Rabes, 1999; Le Douarin et al., 1995). PYRIN/MARESNOSTRIN, which is mutated in familial Mediterranean fever (FMF) (consortium, 1997), MUL, mutated in Mulibrey nanism (Avela et al., 2000), and MID1, mutated in OS, are some examples of RBCC proteins involved in monogenic disorders.

So far, the only more general function of RBCC proteins that could be defined is ubiquitin ligase activity mediated by the RING finger domain (Meroni and Diez-Roux, 2005). Additionally, some members of this family have been shown to orchestrate the formation of large macromolecular complexes (Borden, 1998; Saurin et al., 1996; Torok and Etkin, 2001).

The RING finger domain of MID1 contains a cysteine rich metal binding domain corresponding to the RING finger consensus sequence (Figure 1.6A). This motif binds two zinc atoms through eight potential metal ligands to form a unique “cross brace” motif (Figure 1.6B, Borden, 2000).

Two Bboxes follow the RING domain in MID1 after a linker of forty-five amino acids, in agreement with the conserved relative position of different domains inside the RBCC motif. Bboxes, like RING fingers, are cysteine- and histidine-rich zinc binding motifs. On account of slightly differing consensus sequences, two different types of Bboxes have been described, Bbox type 1 (Bbox1) and Bbox type 2 (Bbox2) (Figure 1.6C, Torok and Etkin, 2001). The

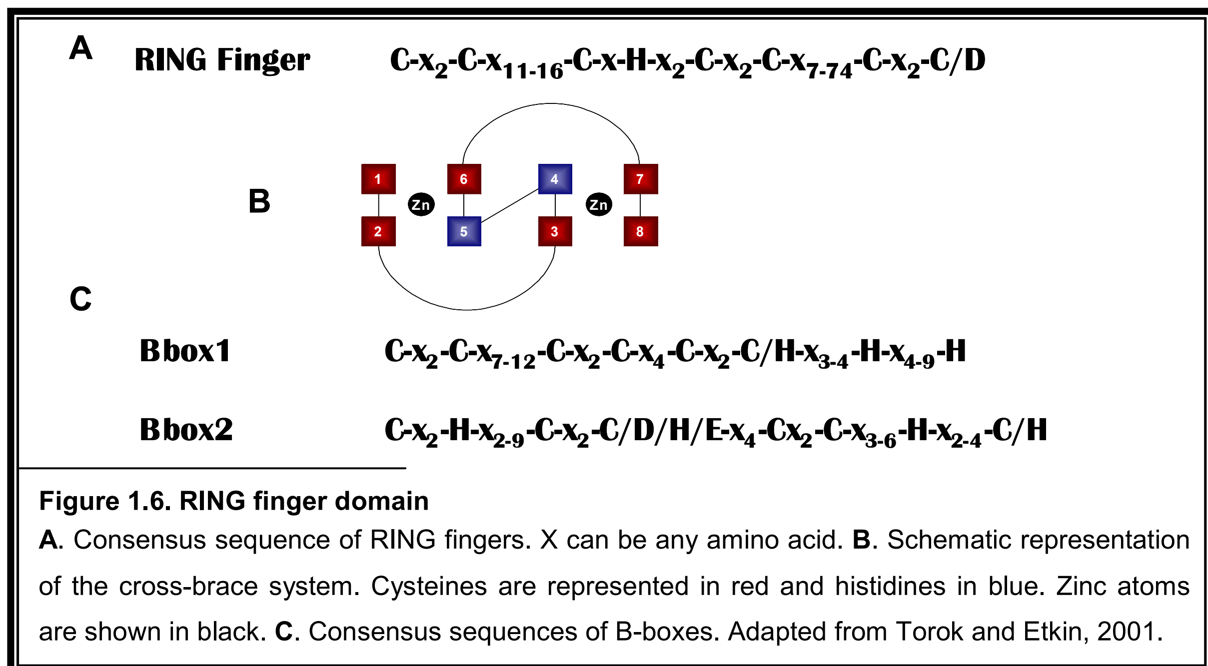

Figure 1.6. RING finger domain
A. Consensus sequence of RING fingers. X can be any amino acid. **B**. Schematic representation of the cross-brace system. Cysteines are represented in red and histidines in blue. Zinc atoms are shown in black. **C**. Consensus sequences of B-boxes. Adapted from Torok and Etkin, 2001.

RBCC motif contains one or two Bboxes; when both are present, as in MID1, Bbox1 always precedes Bbox2, but if only one is present, it is always Bbox2 (Reymond et al., 2001). Bboxes and Bboxes2 include seven potential metal binding ligands, while Bboxes1 include eight potential metal binding ligands. ^{1}H NMR studies performed on XNF7 Bbox have shown that they bind a single zinc atom in a tetrahedral disposition (Borden et al., 1995b). Curiously, Bboxes only exist in the RBCC protein family, suggesting that they fulfil an essential function of RBCC proteins. Thus far, however, no specific function apart from the mediation protein-protein interactions has been identified (Reymond et al., 2001).

Slightly overlapping and C-terminally situated to Bbox2 is the coil-coiled domain of MID1. Coiled-coil domains are bundles of α-helices, twisted around one another into a super coil or

super helix, which participate in protein-protein interactions and, particularly, in dimerisation processes (Burkhard et al., 2001). Many RBCC proteins homo- and/or heterodimerise through the coiled-coil domain, (e.g. MID1, MID2, RFP, TIF) (Reymond et al., 2001). For MID1, apart from forming homodimers, it can also heterodimerise with MID2, a well-known homologue with the same subcellular localisation (Cainarca et al., 1999). Despite having different expression patterns during development, MID1 and MID2 have 89% amino acid identity and perform overlapping functions (Granata et al., 2005; Short et al., 2002)

As mentioned previously, members of the RBCC protein family have been shown to participate in the formation of macromolecular complexes. PML is one of the most well-studied members; it forms parts of a multiprotein complex called PML nuclear bodies. The RING finger and Bboxes have been found to be necessary for nuclear bodies formation (Jensen et al., 2001). The RFP protein and TIF1-α, two other members of the family, have also been shown to associate with PML nuclear bodies and, additionally, the RFP protein has also been shown to be involved in the formation of homomultiprotein complexes through its Bboxes and coiled-coil domain (Cao et al., 1997).

In the case of MID1, Cainarca et al demonstrated in 1999 that MID1 is part of a large macromolecular complex. Interestingly, mutations in the coiled-coil domain do no disrupt the complex, but produce even larger aggregates. However, most of the interaction partners of the complex remain unknown, and no other members apart from MID1 itself, MID2, α4 (Liu et al., 2001; Trockenbacher et al., 2001), and Mig12 (Berti et al., 2004) have been identified.

An RFP-like domain, also known B30.2 domain, is located at the very C-terminal end of MID1. We have shown previously that MID1 interacts with microtubules through its C-terminal end, as evidenced by the detachment of MID1 from microtubules and subsequent formation of cytosolic aggregates when this domain is mutated (Schweiger et al., 1999). B30.2 domains are present in three different families of proteins: RBCC proteins, BTNs-receptor glycoproteins of the immunoglobulin superfamily and stonutoxin (STNX)- protein secreted with the venom expectorated by *Synanceia horrida* (Henry et al., 1998). As yet, attempts at elucidating a common function for B30.2 domains have been unsuccessful. However, it has been suggested that these domains, as in MID1, fulfil important roles in providing the correct subcellular location of RBCC proteins (Reymond et al., 2001).

Between the RBCC motif and the B30.2 domain, an FNIII domain has been defined in the MID1 protein (Cox et al., 2000). FNIII domains are found in a huge variety of proteins such as proteins of the extracelullar matrix, cell adhesion molecules, enzymes, or muscle proteins. In spite of being highly conserved domains, there is no common function attributable to all proteins in which the domain is found. FNIII domains have a β-sandwich fold closely related to the immunoglobulin fold, and contain several highly conserved prolines. Varying numbers of repeats of the motif are usually found in the different protein families. Last year, it was

proposed that the FNIII domain acts as an elbow linking the various functional sites of the MID1, and anchors the protein to the cytoplasmic region where it is needed, namely at the microtubules (Mnayer et al., 2005). However, these observations require further experimental proof.

1.2.2 The MID1 function

Apart from a microtubule stabilizing role (Schweiger et al., 1999), we could show previously that MID1 interacts with the α4 protein through its Bbox1 and that mutations in the coiled-coil domain or the C-terminus do not affect this interaction (Liu et al., 2001; Trockenbacher et al., 2001). α4, encoded by the *IGBP1* gene, binds and negatively regulates the catalytic subunit of phosphatase 2A (PP2Ac) (Figure 1.7, Chen et al., 1998; Murata et al., 1997). This mechanism was first observed in yeast, where TAP42 (yeast homologue of α4) regulates PP2Ac upon Target Of Rapamycin (TOR) signalling (see section 1.3, Nanahoshi et al., 1999; Raught et al., 2001). α4 is a cytosolic protein with a diffuse distribution, but in association with MID1, it is recruited to microtubules. Reciprocally, an excess of α4 can detach MID1 from microtubules and bring it to the cytosol, indicating that both proteins interact and collaborate in determining their subcellular localisation. The interaction between MID1 and α4 leads to a spatial proximity of the ubiqutin ligase activity of MID1, mediated by its RING-finger domain, and microtubule-associated PP2Ac, which results in a ubiquitin-specific modification and degradation of the phosphatase (Figure 1.7). C-terminally mutated MID1 can no longer bind the microtubules and, therefore, the interaction of the MID1-α4 complex with the pool of microtubule-associated PP2Ac can not take place. As a result, PP2Ac can not be degraded and accumulates at the microtubules, resulting in the hypophosphorylation of downstream regulated microtubule-associated proteins (MAPs) (Schweiger and Schneider, 2003; Trockenbacher et al., 2001).

Interestingly, the phosphorylation status of MAPs targeted by PP2A has been shown to be involved in a variety of developmental disorders, such as Miller-Dieker syndrome, X-linked double cortex syndrome (Avila et al., 1994; Gong et al., 2000). In addition, a similar pathomechanism has been observed in Alzheimer´s disease (Trojanowski and Lee, 1995), where hyperphosphorylation of Tau protein is involved (Lim and Lu, 2005).

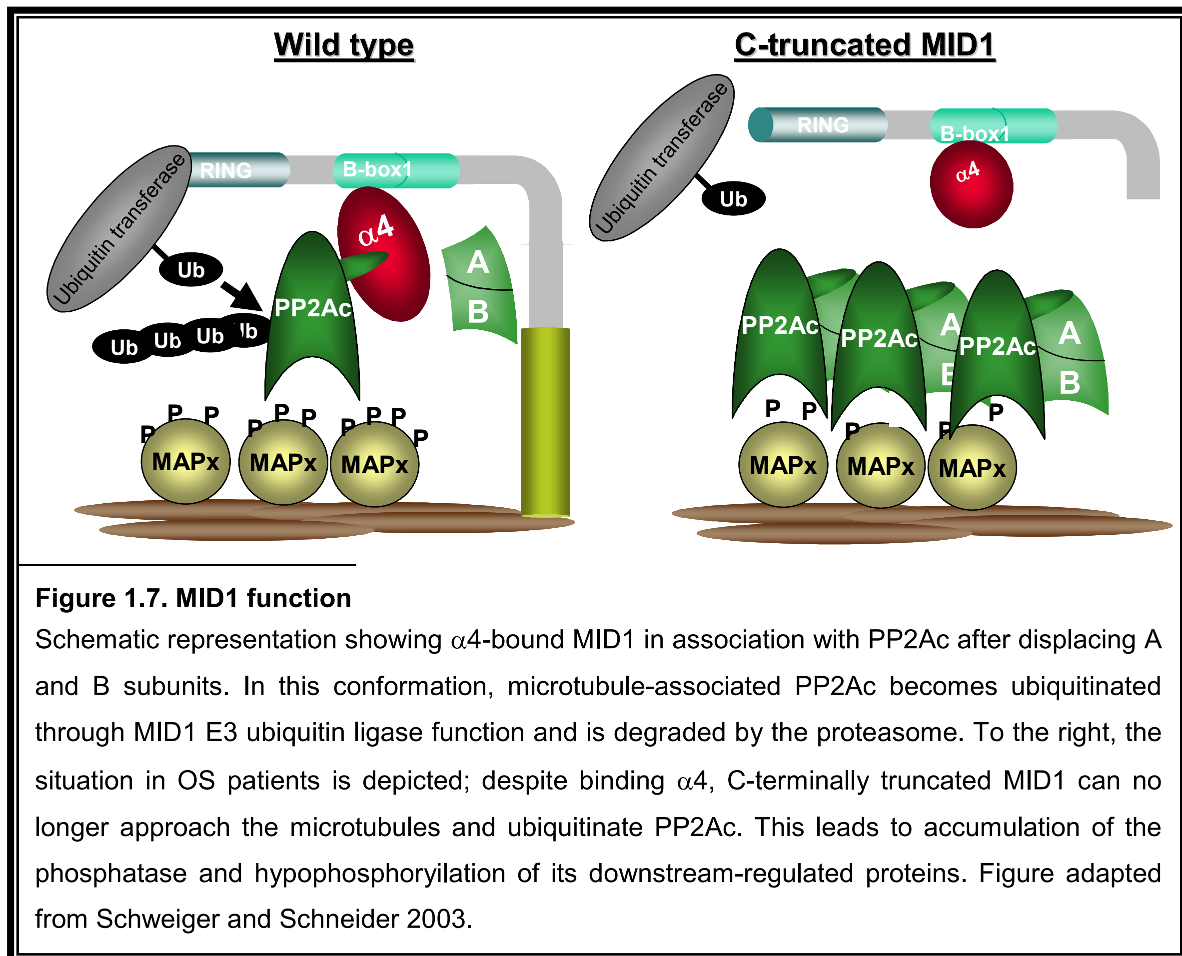

Figure 1.7. MID1 function

Schematic representation showing α4-bound MID1 in association with PP2Ac after displacing A and B subunits. In this conformation, microtubule-associated PP2Ac becomes ubiquitinated through MID1 E3 ubiquitin ligase function and is degraded by the proteasome. To the right, the situation in OS patients is depicted; despite binding α4, C-terminally truncated MID1 can no longer approach the microtubules and ubiquitinate PP2Ac. This leads to accumulation of the phosphatase and hypophosphoryilation of its downstream-regulated proteins. Figure adapted from Schweiger and Schneider 2003.

1.2.3 Regulation of MID1 function

Although MID1 is ubiquitously expressed, OS phenotype specifically affects the ventral midline, suggesting that MID1 function is carefully regulated in a tissue-dependent manner (Schweiger and Schneider, 2003). Moreover, a wide phenotypic variability among OS patients has been described (Cox et al., 2000; So et al., 2005) pointing at modifying factors that participate to regulate MID1 function. Three different mechanisms have been proposed for the regulation of MID1 function:

A - Splicing mechanisms. *MID1* produces a multitude of alternatively spliced transcripts that are differentially expressed in diverse tissues (Winter et al., 2004). The splice variants lead to loss-of-function through several mechanisms conserved in human, mouse and fugu. Some transcripts produce C-terminally truncated MID1 that binds more strongly to α4 and have a dominant negative effect over MID1. Others contain premature stop codons that, in contrast to the previously described transcripts, lead to non-sense mediated decay (NMD). Interestingly, a splice variant lacking the Bbox2 domain has been shown to bind with higher affinity to α4, hinting at an important role of this domain in regulating the interaction of MID1 with the PP2A complex.

B - Protein interaction. MID1 is a member of a multiprotein complex. Thus, the regulation of other members of the complex implies the regulation of MID1. Recently, it has been shown that α4 is expressed during development in tissues affected by OS, suggesting that the tissue-specific expression of α4 could contribute to regulate MID1 function (Everett and Brautigan, 2002). Similar observations have been made for Mig12 (Berti et al., 2004).

It is also known that MID2 is functionally equivalent to MID1. It has been proposed that in some tissues, MID2 substitutes for MID1 to a greater extent than in others and, consequently, only those tissues with less MID2 activity would suffer from MID1 loss (Buchner et al., 1999; Granata et al., 2005; Short et al., 2002).

C - Phosphorylation status. Having been shown to be a phosphoprotein, it has been proposed that MID1 phosphorylation status (or other post-transcriptional modifications) regulates its association with, and subsequent stabilization of, microtubules. In 2001, Liu et al showed that MAP kinase activity is required to maintain MID1 association to microtubules. Unfortunately, most of the mechanisms underlying MID1 posttranslational modification and regulation remain unknown.

1.3 mTOR pathway

In 1991, a screening of mutations conferring resistance to rapamycin in *S.cerevisae* led to identify TOR1, TOR2 (Target of Rapamcycin) and FPR1 kinases (Heitman et al., 1991). Rapamycin, a potent cell proliferation and growth blocker, and immunosuppressant, forms a complex with FKBP12 (FK506-binding protein of 12kDa; coded by FPR1) that binds and inhibits TOR. This mode of action is well-conserved from yeast to mammals. TOR is an evolutionary conserved serine/threonine kinase, member of the phosphatidilinositol kinase-related protein kinase complex (PIKK). While the majority of lower eukaryotes have two Tor proteins, in higher eukaryotes only one orthologue is found (Jacinto and Hall, 2003).

mTOR, the mammalian TOR orthologue, is a large protein that forms large macromolecular complexes through its multiple protein-protein binding domains. So far, only three interaction partners of mTOR have been identified, RAPTOR (regulatory associated protein of TOR), GβL (G-protein β-subunit-like protein) and RICTOR (rapamycin-insensitive companion of TOR) (Hara et al., 2002; Kim et al., 2002; Kim et al., 2003; Sarbassov et al., 2004). The RICTOR complex, like the Tor2 complex in yeast, specifically regulates the cytoskeleton, independent of rapamycin-sensitive pathways (Jacinto et al., 2004).

As a key regulator of cell proliferation and growth, mTOR senses and integrates nutrients and growth factor signals to regulate protein translation and ribosome biogenesis (reviewed in Fingar and Blenis, 2004). Although the best-characterised downstream targets of mTOR are S6K1 and 4E-BP1 (Figure 1.8, Gingras et al., 2001), other upstream and downstream factors of mTOR have been extensively studied and reviewed (Hay and

Sonenberg, 2004). A complex pattern of signals regulates mTOR (Figure 1.8), probably because protein synthesis requires large amounts of energy, and the cell has to make sure that the conditions are adequate to proceed. mTOR activation is controlled by levels of nutrients (branched amino acids and energy) and growth factors, which activate PI3K signalling and regulate mTOR through a sequential pathway known as PI3K/Akt/TSC/Rheb pathway. Interestingly, PI3K can phosphorylate 4E-BP1 and S6K1 and, thus, control cell proliferation independently from mTOR (reviewed in Fingar and Blenis, 2004).

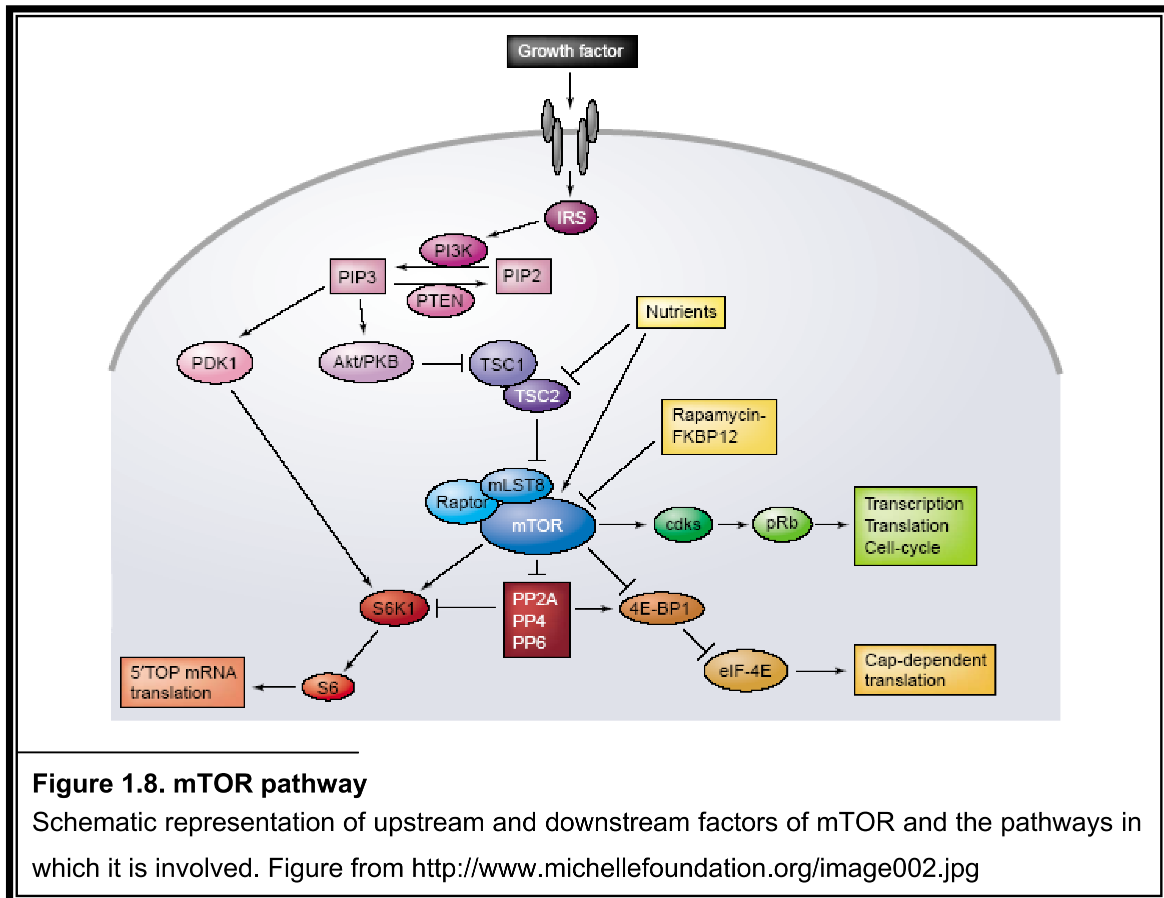

Figure 1.8. mTOR pathway
Schematic representation of upstream and downstream factors of mTOR and the pathways in which it is involved. Figure from http://www.michellefoundation.org/image002.jpg

Activated mTOR, forming a complex with RAPTOR and GβL, and in collaboration with the PIK3 pathway, phosphorylates S6K at several residues. Once phosphorylated, S6K phosphorylates the 40S ribosomal protein S6 (rpS6). This mechanism was formerly believed to regulate 5´TOP mRNA translation, but recently, this has been shown to be unlikely (Tang et al., 2001a). However, it is clear that mTOR is involved in the induction of 5´TOP marked mRNA translation (Pende et al., 2004), and since many of these RNAs encode ribosomal proteins and elongation factors, the translational capacity of the cell as a total increases.

4E-BP1, the second well-known target of mTOR, is a repressor of the translation initiation factor eIF4E. When activated, eIF4E recognizes the cap structure (m7GpppN) at the 5´end of some mRNAs and initiates cap dependent translation by interacting with eIF4G, a scaffolding protein that recruits the translation machinery to the 5´end of mRNA (Gingras et al.,

2001). Under mTOR activation and PIK3 signalling, 4E-BP1 is phosphorylated and disassembles from eIF4E, thereby allowing the initiation of cap-dependent translation.

In addition, TOR regulates phosphatase activity by promoting the binding of TAP42 (type 2A associated protein) to the catalytic subunit of PP2A and to SIT4 (type 2A related phosphatase) in a rapamycin-sensitive manner. This results in the inhibition of PP2A activity (Di Como and Arndt, 1996; Jiang and Broach, 1999). In mammals, a similar mechanism mediated by the TAP42 homologue α4 has been proposed (Chen et al., 1998; Schweiger and Schneider, 2003). Although it has been shown that mTOR regulates phosphatase activity and that α4 negatively regulates PP2A activity, the contribution of rapamycin to the α4-PP2A complex remains controversial (Dennis et al., 1999; Peterson et al., 1999).

1.4 Aim of the study

Upon binding the linker protein α4, the microtubule-associated RBCC protein MID1 targets the microtubule-associated PP2Ac towards ubiquitin specific modification and degradation. In OS patients, which are characterised by ventral midline malformations, MID1 harbours mutations in the C-terminal end, resulting in the loss of microtubule-association and formation of cytoplasmic clumps in the cell. Despite preserving its association with α4, C-terminally mutated MID1 can not approach the vicinity of PP2Ac and therefore, the ubiquitination and degradation of microtubule-associated PP2Ac becomes disrupted, leading to hypophosporylation of its downstream targets. Recently, several OS patients have been found to hold mutations in the Bboxes in the N-terminal end of MID1. Although Bboxes are known to participate in protein-protein interactions, no common functional roles could be attributed to them thus far. For the MID1 protein, it has been shown that the first of the two Bboxes interacts with α4, and a regulatory role of this interaction has been suggested for the second Bbox. Consequently, any disruption of these domains interfering with α4-binding would be expected to lead to accumulation of PP2Ac, thus causing OS. During this thesis, a detailed study of the influence of Bbox mutations on the MID1-α4 and MID1-microtubules interactions by immunofluorescence, immunoprecipitation and yeast two-hybrid experiments was performed to further analyse the functions of the two Bboxes and thus, pave the way for a novel pathomechanism responsible for OS.

Not only the Bboxes, but also, similar to other RBCC proteins, several other domains of MID1 have been proposed to participate in protein-protein interactions, thus forming a multiprotein complex consisting of as-yet-unknown protein components. Proteins belonging to the MID1 complex are potentially important players in the development of the ventral midline and the pathogenesis of OS. Consequently, another main focus of this study was to use affinity chromatography and mass spectrometry to biochemically purify the MID1 complex and identify associated proteins. Subsequently, the MID1 complex should be characterised.