Published in final edited form as:

Int J Biochem Cell Biol. 2010 March ; 42(3): 425–428. doi:10.1016/j.biocel.2009.07.018.

Critical roles for SoxC transcription factors in development and cancer

Alfredo I. Penzo-Méndez*

Department of Cell Biology, Lerner Research Institute, Cleveland Clinic, 9500 Euclid Avenue (NC-10), Cleveland, OH 44195, USA

Abstract

Sox4, Sox11 and Sox12 constitute the group C of Sry-related HMG box proteins. They are co-expressed in embryonic neuronal progenitors and in mesenchymal cells in many developing organs. More closely related to each other than to any other proteins, they nevertheless bind DNA and activate transcription in vitro with different efficiencies. *Sox4*-null embryos and *Sox11*-null newborns die from heart malformations and the latter display widespread defects, while *Sox12*-null mice are viable and do not show obvious malformations. *Sox4* facilitates differentiation of lymphocytes, pancreatic beta cells, osteoblasts and acts in redundancy with *Sox11* to promote neuronal differentiation. *Sox4* and *Sox11* are upregulated in many tumor types in humans, where their roles in cell survival, proliferation, and metastasis remain controversial. Together, these data hint that *Sox4* and *Sox11* regulate cell differentiation, proliferation and survival in multiple essential processes, and suggest that they may act in redundancy to control many more developmental, physiological and pathological processes than currently known.

Keywords

Sox4; Sox11; Sox12

1. Introduction

The group C of Sox transcription factors is constituted in vertebrates by Sox4, Sox11, and Sox12 (formerly known as Sox22). Discovered more than 15 years ago (van der Wetering et al., 1993; Jay et al., 1995, 1997), their molecular properties and functions remain incompletely understood. This is, however, rapidly changing, as impressive evidence has been recently provided by many research groups that they may critically control cell fate and differentiation in major developmental processes, and that their upregulation may be a critical determinant of cancer progression.

2. SoxC molecular features

Sox4, *Sox11* and *Sox12* are single-exon genes. They are highly conserved through vertebrate evolution and closely related to the single *SoxC* gene present in the *Caenorhabditis elegans* worm and the *Drosophila melanogaster* fly (Bowles et al., 2000). The human SOX4, SOX11 and SOX12 proteins have 474, 441 and 315 amino acids, respectively (Fig. 1). They feature two functional domains: an Sry-related HMG box (Sox) DNA-binding domain, located in the N-terminal half of the protein, and a transactivation domain (TAD), located at the C-terminus (Dy et al., 2008).

*Tel.: +1 216 445 2686; fax: +1 216 444 9404. penzoa@ccf.org.

The HMG box is 84% identical across all vertebrate SoxC proteins (Dy et al., 2008). The Sox4 HMG box binds preferentially to the AACAAAG motif in electrophoretic mobility shift assay (EMSA) (van der Wetering et al., 1993), but Sox4, Sox11 and Sox12 also bind and transactivate reporters harboring such distantly related motifs as the GACAATAG and CACAATG sequences present in a bona fide *Tubb3* target gene (Bergsland et al., 2006; Dy et al., 2008; Hoser et al., 2008). This suggests that the SoxC proteins may bind in vivo to target gene sequences that differ quite significantly from their preferred site in vitro. Sox4 binds DNA more efficiently than Sox12 in EMSA, and Sox11 binds very weakly (Wiebe et al., 2003; Dy et al., 2008). Removal of the C-terminus increases the binding efficiency of Sox11 and Sox12, suggesting that the TAD may be inhibitory. Two acidic domains in Sox11 also inhibit DNA binding, perhaps by providing a hinge that facilitates masking of the HMG box by the TAD. The in vivo relevance of this putative auto-regulatory mechanism remains, however, to be demonstrated.

The TAD is only 66% identical across all vertebrate SoxC proteins, but it is 97% identical amongst Sox4 and Sox11 orthologues and 79% identical amongst Sox12 orthologues (Dy et al., 2008). This 33-residue domain is predicted to form a 20-residue long alpha helix. This helix would be continuous in Sox11, but interrupted by 3 and 7 residues in Sox4 and Sox12, respectively. Consistent with the notion that this configuration may be essential for the strength of the TAD, Sox11 is the most potent and Sox12 the least potent of the 3 proteins in activating transiently transfected reporter genes (Kuhlbrodt et al., 1998; Wiebe et al., 2003; Dy et al., 2008; Hoser et al., 2008). GAL4-SoxC TAD fusion proteins show the same relative efficiencies in activating a GAL4 reporter, a result further supporting the notion that most of the differences in SoxC efficiency are explained by differences in the TAD (Dy et al., 2008).

While the activity of many transcription factors is modulated by post-translational modifications, no evidence has been reported yet that this is also the case for the SoxC proteins. There is evidence, however, that the SoxC proteins are able to cooperate with protein partners. Like the SoxB1 and SoxE proteins, they strongly synergize with the POU domain transcription factors Brn1 and Brn2 to activate an *Fgf4* enhancer that contains adjacent Sox and POU domain binding sites (Kuhlbrodt et al., 1998; Dy et al., 2008). Similarly, expression of *Nestin* in the early neural tube is driven by an enhancer that contains adjacent Sox and POU binding sites. Both sites are required for enhancer activation by Brn2 and either Sox2 or Sox11 in vitro. Sox2 and Sox11 are not co-expressed in the neural tube but each is co-expressed with Brn2 and Nestin, further suggesting that the SoxC proteins may synergize with POU factors in vivo (Tanaka et al., 2004). A recent report indicates that protein interaction may also be used by the SoxC proteins to regulate non-transcriptional targets. Sox4 was found to bind to the transcription factor p53 and thereby to prevent p53 degradation in lung non-small cell carcinoma cells (Pan et al., 2009). Stabilization of p53 involved displacement of the ubiquitin ligase Mdm2. Protein interaction occurred between the Sox4 and p53 DNA-binding domains as well as between the Sox4 TAD and p53 regulatory domain. It is likely that Sox11 and Sox12 are capable of similar interaction, and it is plausible that similar mechanisms allow SoxC proteins to regulate other proteins.

3. SoxC expression

The expression of the SoxC genes has been analyzed in the mouse embryo from mid-organogenesis and found to be widespread and largely overlapping (Hargrave et al., 1997; Sock et al., 2004; Dy et al., 2008; Hoser et al., 2008) (Fig. 2). The highest RNA levels were found in post-mitotic neuron progenitors throughout the neural tube, in dorsal root ganglia, thalamus, retina, and cerebral and cerebellar cortex. The 3 RNAs were also expressed in undifferentiated mesenchyme, with the highest levels in skeletal primordia and genital tubercle. In addition, co-expression occurred in several tissues within developing organs, including

endocardial cushions, lung, gut and pancreas epithelium and mesenchyme, and nephrogenic mesenchyme. In addition, non-overlapping *SoxC* expression was reported in several tissues. *Sox11* and *Sox12*, but not *Sox4*, were found expressed in developing palatal shelves and eyelids, whereas *Sox4* and *Sox12*, but not *Sox11*, were expressed in developing thymus, and only *Sox4* was expressed in hypertrophic chondrocytes (Reppe et al., 2000), teeth bud mesenchyme, and hair follicles (Dy et al., 2008). Expression of the 3 genes becomes more restricted by birth, as cell types reach final differentiation. High *Sox4* and *Sox11* expression levels were detected in adult pancreatic islet cells (Lioubinski et al., 2003), and *Sox4* expression was also detected in adult gonads and thymus, but neither *Sox4* nor *Sox11* was found to be expressed in other adult human or mouse tissues (Jay et al., 1993; van der Wetering et al., 1993). In contrast, *Sox12* was found expressed at a low level in most adult human tissues (Jay et al., 1997). *SoxC* expression studies in other vertebrates suggest that *SoxC* functions have been conserved through evolution. Expression of *Sox4* and *Sox11* in chick embryos was reported to be very similar to that in mice (Uwanogho et al., 1995; Maschhoff et al., 2003). The two *Sox11* genes of the zebrafish have distinct expression patterns, but their combined expression is similar to that in chick and mouse embryos, further suggesting that *Sox11* functions have been conserved after genetic duplication in fish (de Martino et al., 2000).

4. Biological functions of the *SoxC* genes

The expression pattern of the *SoxC* genes and the molecular properties of their proteins strongly suggest that these genes could play critical roles in development. *Sox4*-null mice die at embryonic day 14 (Schilham et al., 1996). They show no other obvious defects than severe malformation of the heart outflow tract, which leads to circulatory failure. Endocardial cushions are hypoplastic, resulting in incomplete ventricular septation and partial or total fusion of the proximal aortic and pulmonary trunks. Semilunar valves are also hypoplastic, causing arterial blood backflow. *Sox11*-null mice die immediately after birth, from similar, but less severe, heart outflow tract malformation (Sock et al., 2004). In addition, they display multiple malformations, including microphthalmia with anterior segment dysgenesis (Wurm et al., 2008), open eyelids, cleft palate, cleft lips, hypoplastic lungs, asplenia, omphalocele, undermineralized skull and split lumbar vertebrae. *Sox12*-null mice have no obvious malformations and a normal lifespan and fertility (Hoser et al., 2008). These data have thus started to reveal critical roles for *Sox4* and *Sox11* in multiple major developmental processes. Interestingly, however, none of the single knockouts has led to obvious defects in the CNS and other organs that highly expressed the 3 genes. It is thus tempting to speculate that the 3 genes may have redundant functions, and that these functions will not be revealed until compound mutants are generated.

In vitro inactivation studies have begun to provide support to this proposition. RNAi knockdown of *Sox4* and *Sox11* in chick embryo neural tubes blocked neuronal gene expression while forced expression of *Sox4*, *Sox11* or *Sox12* resulted in neuronal gene upregulation. Functional *SoxC* binding sites were also identified in the regulatory element of the pan-neuronal gene *Tubb3*, suggesting that this gene may be a direct target of the *SoxC* proteins (Bergsland et al., 2006; Hoser et al., 2008). On the other hand, forced expression of *Sox4* resulted in impaired oligodendrocyte differentiation in mice (Pötzner et al., 2007), suggesting that the *SoxC* genes may both promote differentiation of progenitor cells into neurons and prevent differentiation into glia. Similarly, *Sox4* is required for B lymphocyte differentiation, as *Sox4*-null hematopoietic cells grafted into wild-type mice remain blocked at the pro-B cell stage (Schilham et al., 1996). Other studies have suggested that the *SoxC* genes may also be important in specific cell lineages to promote cell proliferation or survival. *Sox4* heterozygous adult mice develop osteopenia, with low osteoblast numbers and low-level expression of osteoblast markers, and *Sox4* knockdown results in impaired proliferation and differentiation of osteoblasts in vitro (Nissen-Meyer et al., 2007). Similarly, *Sox4*-null pancreatic explants

displayed severely reduced numbers of insulin-producing beta cells (Wilson et al., 2005), and knock down of the *Sox4b* in zebrafish embryos resulted in loss of glucagon-producing cells (Mavropoulos et al., 2005). Taken together, these data strongly support the notion that *SoxC* genes regulate differentiation in multiple cell lineages and suggest that additional functions will be identified in the near future.

5. Medical relevance of the *SoxC* genes

While the *SoxC* genes have been implicated in multiple developmental processes in the mouse, no mutations in their human orthologues have yet been associated with any congenital malformation. Many of the defects reported in *Sox4* and *Sox11* mutant mice, such as cleft palate and heart outflow tract malformation, are relatively common birth defects in humans. Many genes have been implicated, but it is tempting to speculate that some cases may arise due to *SOXC* mutations.

Furthermore, the clinical relevance of the *SoxC* genes has risen to a high level in recent years as numerous reports have suggested that the *SoxC* genes may contribute to tumor prognosis. *SOX4* and *SOX11* have been shown to be highly expressed in most medulloblastomas, which are brain tumors derived from neural progenitors (Lee et al., 2002). Increased *SOX4* expression is also associated with bladder (Aaboe et al., 2006), prostate (Liu et al., 2006), colon (Andersen et al., 2009), and non-small cell lung tumors (Medina et al., 2009), and high expression of *SOX11* is associated with gliomas (Weigle et al., 2005), non-B cell lymphomas (Wang et al., 2008) and epithelial ovarian tumors (Brennan et al., 2009). The roles of the *SOXC* genes in tumors are not fully understood, as somewhat conflicting data have been reported. While *SOX4* knockdown resulted in apoptosis of ACC3 adenoid cystic carcinoma cells (Pramoonjago et al., 2006), *SOX4* overexpression was shown to promote cell cycle arrest and apoptosis of HCT116 colon carcinoma cells (Pan et al., 2009). The microRNA miR-335 was shown to inhibit metastatic cell invasion and to act at least in part through targeting *Sox4* and its putative target *Tenascin C*, which encodes an extracellular matrix component implicated in cell migration (Tavazoie et al., 2008). In contrast, it was shown that the higher the level of *Sox4* and *Sox11* expression the better the prognosis of medulloblastomas and other tumor types (de Bont et al., 2008). Thus, it is possible that the *SoxC* genes have different effects on tumor cells depending on the context and primary transformation mechanism, but further studies are needed to clarify this issue.

In conclusion, recent studies have started to uncover key functions for the *SoxC* genes as regulators of cell fate, proliferation and survival in major physiological and pathological processes. Our understanding of these functions is, however, still at its beginning, but it is anticipated that increased attention to these genes will bring forward new and exciting advances in the next few years.

References

- Aaboe M, Birkenkamp-Demtroder K, Wiuf C, Sørensen FB, Thykjaer T, Sauter G, et al. *SOX4* expression in bladder carcinoma: clinical aspects and in vitro functional characterization. *Cancer Research* 2006;66(7):3434–42. [PubMed: 16585165]
- Andersen CL, Christensen LL, Thorsen K, Schepeler T, Sørensen FB, Verspaget HW, et al. Dysregulation of the transcription factors *SOX4*, *CBFB* and *SMARCC1* correlates with outcome of colorectal cancer. *British Journal of Cancer* 2009;100:511–23. [PubMed: 19156145]
- Bergsland M, Werme M, Malewicz M, Perlmann T, Muhr J. The establishment of neuronal properties is controlled by *Sox4* and *Sox11*. *Genes & Development* 2006;20:3475–86. [PubMed: 17182872]
- Brennan DJ, Ek S, Doyle E, Drew T, Foley M, Flannelly G, et al. The transcription factor *Sox11* is a prognostic factor for improved recurrence-free survival in epithelial ovarian cancer. *European Journal of Cancer* 2009;45:1510–7. [PubMed: 19272768]

- Bowles J, Goslik S, Koopman P. Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Developmental Biology* 2000;227:239–55. [PubMed: 11071752]
- de Bont JM, Kros JM, Passier MM, Reddingius RE, Sillevius Smitt PA, Luijckx TM, et al. Differential expression and prognostic significance of SOX genes in pediatric medulloblastoma and ependymoma identified by microarray analysis. *Neuro-Oncology* 2008;10:648–60. [PubMed: 18577562]
- de Martino S, Yan YL, Jowett T, Postlethwait JH, Varga ZM, Ashworth A, et al. Expression of sox11 gene duplicates in zebrafish suggests the reciprocal loss of ancestral gene expression patterns in development. *Developmental Dynamics* 2000;217:279–92. [PubMed: 10741422]
- Dy P, Penzo-Méndez A, Wang H, Pedraza CE, Macklin WB, Lefebvre V. The three SoxC proteins—Sox4, Sox11 and Sox12—exhibit overlapping expression patterns and molecular properties. *Nucleic Acids Research* 2008;36:3101–17. [PubMed: 18403418]
- Hargrave M, Wright E, Kun J, Emery J, Cooper L, Koopman P. Expression of the Sox11 gene in mouse embryos suggests roles in neuronal maturation and epithelio-mesenchymal induction. *Developmental Dynamics* 1997;210:79–86. [PubMed: 9337129]
- Hoser M, Potzner MR, Koch JM, Bösl MR, Wegner M, Sock E. Sox12 deletion in the mouse reveals nonreciprocal redundancy with the related Sox4 and Sox11 transcription factors. *Molecular and Cellular Biology* 2008;28:4675–87. [PubMed: 18505825]
- Jay P, Gozé C, Marsollier C, Taviaux S, Hardelin JP, Koopman P, et al. The human SOX11 gene: cloning, chromosomal assignment and tissue expression. *Genomics* 1995;29:541–5. [PubMed: 8666406]
- Jay P, Sahly I, Gozé C, Taviaux S, Poulat F, Couly G, et al. SOX22 is a new member of the SOX gene family, mainly expressed in human nervous tissue. *Human Molecular Genetics* 1997;6:1069–77. [PubMed: 9215677]
- Kuhlbrodt K, Herbarth B, Sock E, Enderich J, Hermans-Borgmeyer I, Wegner M. Cooperative function of POU proteins and SOX proteins in glial cells. *Journal of Biological Chemistry* 1998;273:16050–7. [PubMed: 9632656]
- Lee CJ, Appleby VJ, Orme AT, Chan WI, Scotting PJ. Differential expression of SOX4 and SOX11 in medulloblastoma. *Journal of Neurooncology* 2002;57:201–14.
- Lioubinski O, Müller M, Wegner M, Sander M. Expression of Sox transcription factors in the developing mouse pancreas. *Developmental Dynamics* 2003;227:402–8. [PubMed: 12815626]
- Liu P, Ramachandran S, Ali Seyed M, Scharer CD, Laycock N, Dalton WB, et al. Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells. *Cancer Research* 2006;66:4011–9. [PubMed: 16618720]
- Mavropoulos A, Devos N, Biemar F, Zecchin E, Argenton F, Edlund H, et al. Sox4b is a key player of pancreatic alpha cell differentiation in zebrafish. *Developmental Biology* 2005;285:211–23. [PubMed: 16055112]
- Maschhoff KL, Anziano PQ, Ward P, Baldwin HS. Conservation of Sox4 gene structure and expression during chicken embryogenesis. *Gene* 2003;320:23–30. [PubMed: 14597385]
- Medina PP, Castillo SD, Blanco S, Sanz-Garcia M, Largo C, Alvarez S, et al. The SRY-HMG box gene, SOX4, is a target of gene amplification at chromosome 6p in lung cancer. *Human Molecular Genetics* 2009;18:1343–433. [PubMed: 19153074]
- Nissen-Meyer LS, Jemtland R, Gautvik VT, Pedersen ME, Paro R, Fortunati D, et al. Osteopenia, decreased bone formation and impaired osteoblast development in Sox4 heterozygous mice. *Journal of Cell Science* 2007;120:2785–95. [PubMed: 17652162]
- Pan X, Zhao J, Zhang WN, Li HY, Mu R, Zhou T, et al. Induction of SOX4 by DNA damage is critical for p53 stabilization and function. *Proceedings of the National Academy of Sciences of the United States of America* 2009;106:3788–93. [PubMed: 19234109]
- Potzner M, Griffel C, Lütjen-Drecoll E, Bösl MR, Wegner M, Sock E. Prolonged Sox4 expression in oligodendrocytes interferes with normal myelination in the central nervous system. *Molecular and Cellular Biology* 2007;27:5316–26. [PubMed: 17515609]
- Pramoonjago P, Baras AS, Moskaluk CA. Knockdown of Sox4 expression by RNAi induces apoptosis in ACC3 cells. *Oncogene* 2006;25:5626–39. [PubMed: 16636670]
- Reppe S, Rian E, Jemtland R, Olstad OK, Gautvik VT, Gautvik KM. Sox-4 messenger RNA is expressed in the embryonic growth plate and regulated via the parathyroid hormone/parathyroid hormone-

- related protein receptor in osteoblast-like cells. *Journal of Bone Mineral Research* 2000;15:2402–12.
- Schilham MW, Oosterwegel MA, Moerer P, Ya J, de Boer PA, van de Wetering M, et al. Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. *Nature* 1996;380:711–4. [PubMed: 8614465]
- Sock E, Rettig SD, Enderich J, Bösl MR, Tamm ER, Wegner M. Gene targeting reveals a widespread role for the high-mobility-group transcription factor Sox11 in tissue remodeling. *Molecular and Cellular Biology* 2004;24:6635–44. [PubMed: 15254231]
- Tanaka S, Kamachi Y, Tanouchi A, Hamada H, Jing N, Kondoh H. Interplay of SOX and POU factors in regulation of the *Nestin* gene in neural primordial cells. *Molecular and Cellular Biology* 2004;24:8834–46. [PubMed: 15456859]
- Tavazoie SF, Alarcón C, Oskarsson T, Padua D, Wang Q, Bos PD, et al. Endogenous human microRNAs that suppress breast cancer metastasis. *Nature* 2008;451:147–233. [PubMed: 18185580]
- Uwanogho D, Rex M, Cartwright EJ, Pearl G, Healy C, Scotting PJ, et al. Embryonic expression of the chicken Sox2, Sox3 and Sox11 genes suggests an interactive role in neuronal development. *Mechanisms of Development* 1995;49:23–36. [PubMed: 7748786]
- van de Wetering M, Oosterwegel M, van Norren K, Clevers H. Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. *EMBO Journal* 1993;12:3847–54. [PubMed: 8404853]
- Wang X, Asplund AC, Porwit A, Flygare J, Smith CI, Christensson B, et al. The subcellular Sox11 distribution pattern identifies subsets of mantle cell lymphoma: correlation to overall survival. *British Journal of Haematology* 2008;143:248–333. [PubMed: 18729857]
- Weigle B, Ebner R, Temme A, Schwind S, Schmitz M, Kiessling A, et al. Highly specific overexpression of the transcription factor SOX11 in human malignant gliomas. *Oncology Reports* 2005;13:139–44. [PubMed: 15583815]
- Wiebe MS, Nowling TK, Rizzino A. Identification of novel domains within Sox-2 and Sox-11 involved in autoinhibition of DNA binding and partnership specificity. *Journal of Biological Chemistry* 2003;278:17901–11. [PubMed: 12637543]
- Wilson ME, Yang KY, Kalousova A, Lau J, Kosaka Y, Lynn FC, et al. The HMG box transcription factor Sox4 contributes to the development of the endocrine pancreas. *Diabetes* 2005;54:3402–9. [PubMed: 16306355]
- Wurm A, Sock E, Fuchshofer R, Wegner M, Tamm ER. Anterior segment dysgenesis in the eyes of mice deficient for the high-mobility-group transcription factor Sox11. *Experimental Eye Research* 2008;86:895–907. [PubMed: 18423449]

Fig. 1. Conserved structural features of human SoxC proteins. HMG, high-mobility-group box DNA-binding domain; TAD, transactivation domain. Amino acid positions at the beginning and ending of each domain are shown.

Fig. 2.

Widespread, overlapping expression of SoxC genes during embryonic development. In situ hybridization analysis of SoxC gene expression in 11.5-day-old mouse embryos. Signal from hybridized RNA is shown in red. Cell nuclei are counterstained with DAPI. f, forebrain; m, midbrain; s, spinal cord; c, craniofacial mesenchyme; v, vertebrae primordia; h, heart; li, liver; g, gut; t, genital tubercle. Reproduced with permission from Dy et al. (2008).