Oct transcription factors in development and stem cells: insights and mechanisms

Dean Tantin*

Summary
The POU domain family of transcription factors regulates developmental processes ranging from specification of the early embryo to terminal differentiation. About half of these factors display substantial affinity for an 8 bp DNA site termed the octamer motif, and are hence known as Oct proteins. Oct4 (Pou5f1) is a well-known Oct factor, but there are other Oct proteins with varied and essential roles in development. This Primer outlines our current understanding of Oct proteins and the regulatory mechanisms that govern their role in developmental processes and concludes with the assertion that more investigation into their developmental functions is needed.

Key words: Oct proteins, POU domain, Stem cells

Introduction
A central tenet of developmental biology is the orchestration of hierarchical gene expression patterns by sequence-specific transcription factors. These gene expression patterns allow for the precise translation of genetic information into the specification of distinct cell lineages, which in turn determines the correct formation of morphological structures. Understanding how this process occurs at a molecular level represents a considerable challenge in developmental biology. Oct transcription factors are key developmental regulators capable of coordinating a spectrum of developmental processes, ranging from the establishment of the embryonic pluripotent ‘ground state’ to terminal differentiation. Oct proteins are a subclass of the POU transcription factor family (Box 1) that recognize an 8 bp consensus sequence [ATGC(A/T)AAT] termed the ‘octamer motif’ and variants thereof (Bodner et al., 1988; Ingraham et al., 1988; Kemler et al., 1989). The 5’ ATGC motif associates with the POU-specific domain, while the 3’ half-site associates with the POU homeodomain (Box 1). The octamer motif is found in the regulatory regions of both ubiquitous and tissue-specific target genes. For example, ubiquitously expressed histone and U small nuclear RNA genes contain the motif in their regulatory regions, but so too do the B cell-restricted immunoglobulin heavy and kappa light chain genes (Fletcher et al., 1987; Henderson and Calame, 1995; Ström et al., 1996).

Although the basic octamer motif is the simplest sequence capable of strong association with Oct proteins, combinations of paired and inverted octamer half-sites are known to bind proteins such as Oct1 (Pou2f1), Oct4 (Pou5f1, Oct3) and Brm2 (Oct7, N-Oct3, Pou3f2). The proteins bind these sequences as dimers or higher-order structures (Reményi et al., 2001; Nieto et al., 2007; Tantin et al., 2008). The basis of this recognition lies in the conformational flexibility of the linker domain, a covalent peptide of variable length and sequence that allows the two DNA-binding subdomains to reorganize relative to each other (Box 1). The discovery of many diverse Oct protein binding sites mirrors the pattern for other transcription factors, such as Notch intracellular domain, NRSF/REST and T-box transcription factors, in which variant sites with different half-site spacing have been shown to bind the factor physiologically (Conlon et al., 2001; Cave et al., 2005; Ong et al., 2006; Johnson et al., 2007). The capacity for binding to alternative sequences also provides a mechanism for differential regulation by Oct proteins because binding to some sites but not others can be controlled by upstream signals such as phosphorylation (Nieto et al., 2007; Kang et al., 2009).

With these basics in mind, the aim of this Primer is to provide an overview of the known and likely developmental roles of Oct proteins, in particular their relevance to stem cells and how our knowledge of their mechanisms of action can inform their function. Synthesis of prior literature and new mechanistic developments in the field suggests that the time is right to refocus on the roles of these proteins in development. Doing so might result in a renaissance in our understanding of the developmental roles of these key transcription factors.

Oct proteins: a brief overview
Oct proteins comprise POU domain transcription factor classes II, III and V (Table 1). The distinction of these classes is based on sequence similarity of the DNA-binding domain. The best-known Oct protein is the pluripotency regulator Oct4 (Okamoto et al., 1990; Rosner et al., 1990; Schöler et al., 1990), which together with the related protein Sprm1 (Pou5f2) (Pearse et al., 1997), constitutes class V. Oct4 is expressed in the early mammalian embryo and in the germline. Within the early embryo, Oct4 is expressed in pluripotent cells of the blastocyst inner cell mass (ICM) and epiblast, which will eventually give rise to all the cells of the embryo proper. Oct4-deficient embryos fail to establish pluripotency. Instead of forming an ICM, the whole embryo differentiates into trophoectoderm and fails to develop following implantation (Nichols et al., 1998). Oct4 is also expressed in embryonic stem cells (ESCs) derived from the ICM, and has been shown to induce reprogramming to ESC-like induced pluripotent stem cells (iPSCs), either alone or in combination with other factors (Yamanaka, 2006; Kim et al., 2009; Zhu et al., 2010).

Class II POU domain proteins include prominent Oct proteins such as the ubiquitously expressed Oct1 (Fletcher et al., 1987; Sturm et al., 1987; Sturm et al., 1988) and Oct2 (Pou2f2), which is expressed in the brain and blood (Landolfi et al., 1986; Staudt et al., 1986; Scheidereit et al., 1987; Scheidereit et al., 1988; Staudt et al., 1988; He et al., 1989). Oct1 and Oct2 are the most closely related members, with more than 85% identity in the DNA-binding domain. Another POU class II protein is Oct11 (Skn-1/Pou2f3), which is expressed predominantly in epidermal keratinocytes and taste receptor cells in the taste buds – two cell types that turn over in adult mammals (Andersen et al., 1997; Matsumoto et al., 2011).
The POU family is a homeodomain subgroup characterized by the presence of a ~60 amino acid homeodomain (POUα) joined by a flexible linker to a second, independently folded ~80 amino acid DNA-binding domain termed the POU-specific (POUβ) domain (Herr et al., 1988). The two domains make separate contacts with the DNA (Klemm et al., 1994), as illustrated (left) for Oct1 (pink) bound to DNA (gray) at the canonical octamer DNA sequence, the 5’ half of which associates with POUγ (s) while the 3’ half associates with POUδ (h); in this case, DNA binding is static. Oct proteins can also form dimers or higher-order structures at variants of the canonical binding site, as depicted (center) for an Oct1 homodimer bound to the alternative MORE (more octamer-related palindromic element) sequence (Remenyi et al., 2001), to which binding is inducible by phosphorylation. The mode of multimer (e.g. Oct4) binding to sites identified in vitro by ChiP remains unknown (right), but might be signal responsive in vivo. The combined 150-160 amino acid DNA-binding domain is the hallmark of POU proteins. The strongest single distinguishing POU domain feature lies in the POUγ recognition helix: all POU proteins contain the sequence RVWFCN, the sole exception being the atypical POU protein HNF1α (Baumhueter et al., 1990). In non-POU proteins, the cysteine residue in this sequence is typically replaced with a glutamine or serine. The POU DNA-binding structure is grossly similar to a complex of Hoxa9 and a cooperative binding partner, Pbx1, bound to DNA (LaRonde-LeBlanc and Wolberger, 2003). It appears that in the course of POU protein evolution this cooperative association has been solidified by fusing the two ancestral DNA-binding domains together.

POU proteins are expressed in patterns that vary from highly tissue-restricted to ubiquitous. Six POU protein subclasses, termed POU I through POU VI, have been defined based on the sequence composition of the DNA-binding domain. Although POU factors recognize a continuum of related sequences, three classes (POU I, IV and VI) do not display high affinity for the standard octamer motif and so are termed non-octamer-binding factors. The other three classes (POU II, III and V) are termed Oct proteins because they display strong affinity for the octamer motif. As a group, POU proteins tend to be broadly expressed early in development and become progressively more restricted as development proceeds. In the adult, they show highly tissue-restricted expression patterns. An exception to this trend is Oct1, which appears to be the only POU protein that is universally expressed.

Mammalian Oct6 (SCIP, Tst-1, Pou3f1), Brn2, Brn1 (Oct8, Pou3f3) and Brn4 (Oct9, Pou3f4) constitute class III of the POU family. Oct6 is expressed in ESCs, the developing brain and in epidermis (Monuki et al., 1989; Suzuki et al., 1990; Andersen et al., 1997). Brn2 is expressed in specific brain substructures including the neuroendocrine hypothalamus and pituitary (Nakai et al., 1995; Schonemann et al., 1995). Interestingly, Brn2 has been shown to induce direct reprogramming to neuronal lineages when delivered together with other neuronal lineage-specific transcription factors (Pang et al., 2011). Brn2 and Oct4 are thus two POU domain transcription factors with potent lineage reprogramming ability.

The developmental roles of Oct proteins are complex and do not correlate in any obvious way with their degree of homology to one another or with the POU protein class in which they are grouped. Furthermore, it is known that some of these factors, such as Oct4, have unique essential functions whereas others have redundant or compensatory functions that can mask important developmental roles. For example, Brn1 and Brn2 together, but not individually, are crucial for the proper migration of cortical neurons and the proper development of the mammalian neocortex (Sugitani et al., 2002).

**Oct protein regulation: interaction partners and signal responsiveness**

Recent studies have begun to reveal how Oct proteins respond to signals and partner with other proteins to regulate target gene transcription. Oct proteins can integrate multiple upstream signals, which allows precise regulation of their levels, activity and localization. These changes in Oct protein function ensure correct developmental patterning and adult tissue homeostasis. An example of regulated changes in Oct protein levels comes from mammalian peri-implantation embryos. Here, loss of Oct4 is associated with loss of pluripotency and differentiation of epiblast cells (Nichols et al., 1998). Despite the fact that complete loss of Oct4 abolishes pluripotency, slightly attenuated Oct4 levels have been shown to enhance pluripotency in ESCs (Karwacki-Neisius et al., 2013). Thus, the level of Oct4 protein can exert very fine control over the pluripotent state. Oct protein activity can also be regulated by upstream signals, as seen for example with Oct1 and T lymphocytes. Oct1 has been shown to repress the gene that encodes interleukin 2 (IL2) in naïve T cells but switches to an activating mode upon T cell activation (Shakya et al., 2011). This property may allow a single transcription factor to participate at many levels in multistep developmental processes. Control of Oct protein localization is evidenced by the interactions of both Oct4 and Oct1 with their respective upstream regulators (Table 2). Oct1 can be reversibly tethered to the nuclear periphery through interactions with PIASγ (Pias4) (Tolkunova et al., 2007). Oct1 can similarly reversibly associate with the nuclear periphery through interactions with lamin B1. In this latter case association is known to be regulated by oxidative stress, chronic overgrowth and cellular aging (Imai et al., 1997; Malhas et al., 2009; Kang et al., 2013).
Other Oct proteins have not been afforded this level of scrutiny, but their functions are likely to be regulated at multiple levels as well. Many, possibly all, Oct proteins are subject to regulation through post-translational modifications. These modifications are likely to account for many of the changes in levels, activity and localization described above. The two best-studied Oct proteins in this respect are again Oct4 and Oct1, which are regulated by phosphorylation (Segil et al., 1991; Nieto et al., 2007; Schild-Poulter et al., 2007; Kang et al., 2009; Van Hoof et al., 2009; Kang et al., 2011; Brumbaugh et al., 2012; Lin et al., 2012), O-GlcNAcylation (Webster et al., 2009; Jang et al., 2012; Kang et al., 2013), SUMOylation (Wei et al., 2007; Zhang et al., 2007) and ubiquitylation (Xu et al., 2004; Kang et al., 2011; Kang et al., 2013). These modifications control protein stability, their DNA binding properties, co-factor association and tethering to the nuclear envelope. This array of modifications is unlikely to produce an on/off-type response, but rather a continuum of different outputs.

Developmentally important transcription factors can be regulated via changes in the levels or activity of partner proteins with which they form cooperative (or sometimes antagonistic) DNA-bound complexes. The ability to form these cooperative complexes alters and expands the DNA sequence recognition space available to individual factors. This paradigm applies in the case of Oct proteins, although here the case is more unusual because POU proteins are already in effect cooperative DNA-binding complexes by virtue of their two independently folded DNA-binding subdomains (Box 1). Both Oct1 and Oct4 associate with Sox2 at DNA elements composed of ‘Oct/Sox’ composite binding sites (Ambrosetti et al., 1997). The requirement for Sox2 in ESCs can be eliminated by slightly elevating Oct4 expression, indicating a fine balance of activities necessary to maintain pluripotency (Masui et al., 2007). Recent findings suggest that Oct4 can also bind DNA cooperatively with the related protein Sox17 at slightly different composite elements with shorter spacing between the DNA binding sites. These alternative sites are found in genes that specify endoderm (Aksoy et al., 2013).

Oct4 has been shown to additionally associate with Klf4 (Wei et al., 2009), ESRRB (van den Berg et al., 2008), Nanog (Liang et al., 2008) and Zfp143 (Chen et al., 2008a) at different binding sites. Oct1 has been shown to associate not only with transcription factors such as Sox2, C/EBPβ, AP-1 (Fos), NFAT, NFκB and

<table>
<thead>
<tr>
<th>Table 1. Mammalian POU transcription factors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor (synonyms)</strong></td>
</tr>
<tr>
<td><strong>Octamer-binding POU proteins (Oct proteins)</strong></td>
</tr>
<tr>
<td>Class II</td>
</tr>
<tr>
<td>Oct1 (NF-A1, OTF-1, NF-III)</td>
</tr>
<tr>
<td>Oct2 (NF-A2)</td>
</tr>
<tr>
<td>Oct11 (Skn-1a/i)</td>
</tr>
<tr>
<td>Class III</td>
</tr>
<tr>
<td>Oct6 (SCIP, Tst-1)</td>
</tr>
<tr>
<td>Oct7 (Brn2, N-Oct3)</td>
</tr>
<tr>
<td>Oct8 (Brn1)</td>
</tr>
<tr>
<td>Oct9 (Brn4)</td>
</tr>
<tr>
<td>Class V</td>
</tr>
<tr>
<td>Oct4 (Oct3)</td>
</tr>
<tr>
<td>Sprm1</td>
</tr>
<tr>
<td><strong>Non-octamer-binding POU proteins</strong></td>
</tr>
<tr>
<td>Class I</td>
</tr>
<tr>
<td>Pth1 (GHF-1)</td>
</tr>
<tr>
<td>Class IV</td>
</tr>
<tr>
<td>Brn3a (Brn-3.0)</td>
</tr>
<tr>
<td>Brn3b (Brn-3.2)</td>
</tr>
<tr>
<td>Brn3c (Brn-3.1)</td>
</tr>
<tr>
<td>Class VI</td>
</tr>
<tr>
<td>Brn5</td>
</tr>
<tr>
<td>Atypical</td>
</tr>
<tr>
<td>HNF1a</td>
</tr>
</tbody>
</table>

CNS, central nervous system; PNS, peripheral nervous system.

<table>
<thead>
<tr>
<th>Table 2. Oct4 and Oct1 upstream protein regulators known to alter function</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activity</strong></td>
</tr>
<tr>
<td>PIASy (Pias4)</td>
</tr>
<tr>
<td>Ubc9 (Ube2i)</td>
</tr>
<tr>
<td>Wwp2</td>
</tr>
<tr>
<td>Lamin B1</td>
</tr>
<tr>
<td>NEK6</td>
</tr>
<tr>
<td>OGT</td>
</tr>
</tbody>
</table>
glucocorticoid receptor (Li-Weber et al., 1998; Préfontaine et al., 1999; Belsham and Mellon, 2000; Hatada et al., 2000; van Heel et al., 2002), but also with itself and with other Oct proteins such as Oct2 and Oct6 to form homodimeric, heterodimeric and higher-order complexes (Verrizier et al., 1992; Reményi et al., 2001). Interestingly, the ability to form these complexes can be controlled by phosphorylation of Oct proteins at specific sites in the DNA-binding domain (Nieto et al., 2007; Kang et al., 2009). Co-incident Oct1 and Oct4 binding has been described for a specific subset of Oct4 targets in ESCs. In some cases, sequential chromatim immunoprecipitation (ChIP) indicates that binding occurs simultaneously at the same DNA site (Ferraris et al., 2011). These binding events might therefore be sites of Oct1/Oct4 heterodimer or higher-order complex formation. Much more work will be required to fully understand the interplay between different Oct proteins at their binding sites in different tissues and developmental stages.

Table 3 shows some of the proteins known to mediate the downstream functions of Oct4 or Oct1. These co-factors can act positively, combining with the Oct protein to activate gene expression, or alternatively they may have inhibitory effects. For example, general transcription factors may play an activating role when the Oct protein-binding site in the DNA is exceptionally close to the core promoter, as has been shown for Oct1 (Nakshatri et al., 1995; Inamoto et al., 1997; Bertolino and Singh, 2002). Paf1, Wdr5 and XPC have all been identified as positive co-factors, interacting with Oct4 to maintain pluripotency in ESCs (Ponnusamy et al., 2009; Ang et al., 2011; Fong et al., 2011). Although the mechanisms are not always clear, these co-factors colocalize with Oct4 at specific target genes and help Oct4 mediate its positive transcriptional functions. By contrast, SETDB1 (ESET), a histone H3K9 methyltransferase, associates with Oct4 to repress target gene expression (Yeap et al., 2009; Yuan et al., 2009). NuRD, a multi-subunit complex with nucleosome remodeling and histone deacetylase activity, is another repressive activity known to associate with Oct4 (Liang et al., 2008). Interestingly, three separate studies have identified proteins that physically interact with Oct4 using mass spectroscopy (Pardo et al., 2010; van den Berg et al., 2010; Ding et al., 2012) but only a small number of the proteins (18/263) overlap between the three studies. It is not clear how these are directly related or indirect Oct4 interaction partners, but among the 18 proteins are subunits of NuRD. Oct1 also interacts robustly with NuRD to repress target genes (Shakya et al., 2011). When Oct1 changes to an activating mode, it associates with KDM3A (Jmjd1a), a histone lysine demethylase that removes inhibitory histone marks (Shakya et al., 2011). Upstream signaling through the MEK/ERK pathway mediates Oct1 switching between NuRD and KDM3A to regulate the expression of target genes such as Il2 and Cdx2 (Shakya et al., 2011).

Despite the identification of numerous Oct-interacting proteins, there are still fundamental unresolved questions concerning the mechanisms that underlie the differential utilization of Oct co-factors and regulators. Most of the work in this area has involved Oct4 or Oct1, and it is not clear which of these mechanisms can be generalized to all or part of the Oct protein family. Crucially, with the exception of the Oct1 example provided above, it is still unclear how the same Oct protein is able to switch between activating and repressive states. Post-translational modifications, as well as the specific complexes formed between the Oct protein and other DNA-binding proteins, may determine the recruited effector protein(s) and thus the transcriptional output.

**Modes of transcriptional regulation and gene poising by Oct proteins**

Oct targets can be placed into three broad categories: (1) positive targets that are actively expressed; (2) positive targets that are transcriptionally silent; and (3) negative (repressed) targets (Fig. 1). For those targets in category 2, the Oct protein or proteins maintain epigenetically ‘poised’ gene expression states: situations in which a target gene is silent but readily inducible upon reception of the appropriate developmental cues.

Oct4 provides examples of all three classes. Category 1 Oct4 targets in ESCs include active pluripotency genes such as Nanog. Oct4 also positively regulates its own gene (Pou5f1). Nanog, Oct4 and a small number of other proteins form what is termed the ‘core pluripotency network’, which is required to maintain pluripotency. Their expression is maintained not only by Oct4 but also by other proteins encoded in the network such as Nanog and Sox2 (Boyer et al., 2005; Rodda et al., 2005). In addition to the core pluripotency network, a larger number of additional category 1 Oct4 targets encode constitutively expressed proteins such as histones and metabolic enzymes (Chen et al., 2008b). Category 3 Oct4 targets include Cdx2, which is actively repressed by Oct4 and is silent in ESCs (Yeap et al., 2009; Yuan et al., 2009). Finally, hundreds of developmentally poised category 2 Oct4 targets are silent in ESCs but become activated in different developing tissues,

**Table 3. Mediators of downstream Oct protein activity**

<table>
<thead>
<tr>
<th>Activity</th>
<th>Known interaction partners</th>
<th>Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wdr5</td>
<td>Oct4</td>
<td>Transcriptional co-activator</td>
<td>van den Berg et al., 2010; Ang et al., 2011</td>
</tr>
<tr>
<td>XPC</td>
<td>Oct4</td>
<td>Co-activator for Oct4/Sox2 complexes</td>
<td>Fong et al., 2011</td>
</tr>
<tr>
<td>Pafl (PD2)</td>
<td>Oct4</td>
<td>Transcriptional activation</td>
<td>Ponnusamy et al., 2009</td>
</tr>
<tr>
<td>SETDB1 (ESET)</td>
<td>Oct4</td>
<td>Methylation of histone H3K9, represses gene activity</td>
<td>Yeap et al., 2009; Yuan et al., 2009</td>
</tr>
<tr>
<td>NuRD</td>
<td>Oct4 and Oct1</td>
<td>Deacetylase nucleosomes and mediates transcriptional repression</td>
<td>Liang et al., 2008; Shakya et al., 2011</td>
</tr>
<tr>
<td>OCA-B (Pou2af1, OBF-1)</td>
<td>Oct1 and Oct2</td>
<td>B cells and activated T cells, transcriptional activation</td>
<td>Kim et al., 1996; Schubart et al., 1996; Zwilling et al., 1997</td>
</tr>
<tr>
<td>SNAPc</td>
<td>Oct1</td>
<td>Activation of U snRNA transcription</td>
<td>Ford et al., 1998</td>
</tr>
<tr>
<td>OCA-S</td>
<td>Oct1</td>
<td>Activation of S phase-specific transcription</td>
<td>Zheng et al., 2003</td>
</tr>
<tr>
<td>TBP; TFII B and TFIIH</td>
<td>Oct1</td>
<td>General transcription factors, transcriptional activation</td>
<td>Nakshatri et al., 1995; Inamoto et al., 1997; Bertolino and Singh, 2002</td>
</tr>
<tr>
<td>KDM3A (Jmjd1a)</td>
<td>Oct1</td>
<td>Histone H3K9 demethylation and transcriptional derepression</td>
<td>Shakya et al., 2011</td>
</tr>
</tbody>
</table>

Not included are the many transcription factors documented to engage in cooperative interactions with Oct proteins when both associate with proximal binding motifs.
sequence-specific binding site (gray rectangle) and involves the POUH of Oct protein (green) to a target gene depends on the presence of a and POUS domains (light green) within the Oct protein, which are

Fig. 1. Modes of Oct protein transcriptional regulation. The binding of Oct protein (green) to a target gene depends on the presence of a sequence-specific binding site (gray rectangle) and involves the POUH and POUS domains (light green) within the Oct protein, which are connected via a linker sequence (see Box 1). (A) Gene activation. The Oct protein activates the expression of its target genes (green arrows) either directly by promoting the assembly of a transcription complex at the promoter or indirectly by acting through the nucleosomes (pink). In the latter case, the Oct protein interacts with co-factors (blue) to deposit positively acting chromatin modifications (small green circles) or to remove repressive marks (small red circles). ATP-dependent chromatin remodeling may also be employed (not shown). Active gene expression (black arrow on the DNA) is typically associated with promoter DNA that is free of cytosine methylation (white circles). (B) Posing of transcriptionally silent genes. Targets genes of Oct1 and Oct4, and perhaps other Oct proteins, can be transcriptionally silent (red X) but epigenetically poised to rapidly initiate expression upon reception of the correct developmental cues. Present evidence indicates that the Oct protein mechanism in gene posing is positive, removing repressive modifications. The effect of chromatin on gene expression can be differentiated from that in A because a mixture of positively and negatively acting modifications is present (simultaneous green and red arrows). These genes typically lack DNA methylation (white circles). (C) Gene repression. In this case, the Oct protein recruits co-repressor activities that remove activating marks or deposit repressive marks on the chromatin to repress gene transcription (red arrow). These genes are typified by a high degree of promoter DNA methylation (black circles).

e.g. Hoxa5, Hoxc6, Pax6, Otx2, Gata2 and Pou4f1 (Brn3a) (Bernstein et al., 2006; Chen et al., 2008b). They often encode tissue-specific, lineage-instructive transcription factors. In ESCs, category 2 genes are characterized by a chromatin state free of DNA methylation and containing both activating and repressive histone marks (Bernstein et al., 2006; Meissner et al., 2008). These genes are rapidly induced if differentiating cells encounter the appropriate developmental cues, or become stably repressed if cells proceed down other developmental trajectories. Oct4 is not only expressed in ESCs but also in primordial germ cells and spermatogonial stem cells (Rosner et al., 1990; Kehler et al., 2004). It is possible that Oct4 associates with the same category 2 genes in these cells as well, and thus the activity of Oct4 might be a unifying feature of totipotent/pluripotent stem cells of both the germline and embryo. It is difficult to test mechanistically the poising function of Oct4 in ESCs because of its simultaneous role in maintaining the pluripotency network. Ablating Oct4 in ESCs causes the network to collapse and ESCs to differentiate. Nevertheless, the strong correlation between Oct4 binding to these targets and their poised configuration strongly suggests that Oct4 is executing a poising function. These findings regarding Oct4 target binding to the target is presumably positive, Fig. 1). Simple evaluation of target gene expression levels is inadequate for this task. Instead, the presence of particular co-factors and the status of local chromatin (Fig. 1) at the target need to be determined.

Other Oct protein family members also mediate different target gene expression states. Oct1 has been directly implicated in gene poising at the IL2 locus in resting but previously stimulated CD4 T cells (Shakya et al., 2011). In this study it was found that, within 6 hours of naïve CD4 T cell stimulation, Oct1 switches IL2 from a repressive (category 1) to a posing (category 2) state. Switching was dependent on MAP kinase signals, indicating that Oct proteins can rapidly switch the manner in which they regulate target genes in response to upstream signals. After withdrawal of the initial stimulus and attenuation of IL2 expression, Oct1 blocked the complete repression of this locus, maintaining it in a poised state such that the cells responded more quickly and with greater augmented IL2 expression upon reactivation (Shakya et al., 2011). This ‘anti-repression’ is achieved via an interaction of Oct1 with the chromatin-modifying protein KDM3A, which opposes inhibitory histone methylation of the IL2 promoter.

Mammalian ESCs co-express Oct4 with Oct1 and Oct6 (Okamoto et al., 1990; Rosner et al., 1990; Schöler, 1991). Oct1 co-occupies many of the same poised developmental Oct4 targets but not the active pluripotency Oct4 targets (Ferraris et al., 2011). Oct6 binding to endogenous target genes has not been tested. It is unclear whether Oct1 or Oct6 is also involved in poising developmental genes in ESCs and pluripotent cells of the embryo.

Indirect evidence implicates three other mammalian Oct proteins, Oct2, Brn2 and Brn4, in gene poising. Although Oct2 is increasingly expressed throughout adult B cell development, Oct2-deficient B cells mature normally. Instead, mature B cells manifest multiple defects upon activation, including poor proliferation and poor terminal differentiation to plasma cells (Corcoran et al., 1993; Corcoran and Karvelas, 1994; Emslie et al., 2008). These results are consistent with a gene poising role of Oct2 during B cell development in a manner that might be conceptually and mechanistically similar to Oct1, its closest paralog. Similarly, the POU III Oct protein Brn2 is expressed throughout mammalian neuroendocrine hypothalamic development, but in the null condition defects only manifest during terminal differentiation (Nakai et al., 1995; Schonemann et al., 1995). It is therefore possible that Brn2 is also involved in poising specific targets at the earlier developmental stages, with the phenotype manifesting only
when critical target genes need to be properly induced. In the developing brain, another POU III Oct protein, Brn4, is widely expressed in the ventricular zone, including in striatal stem and progenitor cells. However, depletion of Brn4 specifically reduces the number of differentiated neurons (Shimazaki et al., 1999). These findings are also consistent with a model in which Brn4 poises key targets at earlier developmental stages. These possibilities await formal testing through assessment of the chromatin at target genes and careful measurement of gene induction following manipulation of these activities. In summary, it is possible that many Oct proteins mediate poised chromatin states at subsets of their targets.

### Oct proteins and somatic stem cells

Pluripotent and somatic stem cells are critically dependent on poised gene expression states for proper function. Stem cells also frequently (but not uniformly) show properties such as a glycolytic metabolic signature, relative resistance to oxidative and genotoxic stress, and the ability to divide asymmetrically. Interestingly, both Oct4 and Oct1 regulate target genes involved in core carbon metabolism (Chen et al., 2008b; Shakya et al., 2009). Higher levels of Oct1 promote both stress resistance (Tantin et al., 2005) and a glycolytic metabolic profile associated with cellular transformation (Shakya et al., 2009). The metabolic and stress response phenotypes suggest that Oct1 might regulate functions associated with stem cells. Recent results show that Oct1 is highly expressed in the stem cell compartments of the gastrointestinal and blood systems and that Oct1 promotes multiple phenotypes associated with stem cell function, such as hematopoietic engraftment ability (Maddox et al., 2012). These results strongly suggest that Oct1 regulates stem cell function in a range of tissue-specific contexts.

Apart from Oct1 and Oct4, there is tantalizing but tangential evidence that other Oct proteins regulate the stem cell phenotype. As mentioned above, the mammalian Oct2, Brn2 and Brn4 loss-of-function phenotypes are consistent with a role in gene poising in progenitor/stem cell compartments. Other evidence comes from non-mouse models. A summary of POU proteins in three common non-mammalian model organisms, *Drosophila melanogaster*, *Caenorhabditis elegans* and *Danio rerio*, is shown in Table 4. There are five *Drosophila* POU proteins (Table 4). Two of these, Pdm1 (Nubbin – FlyBase) and Pdm2, are Oct proteins belonging to POU class II, the same group in which mammalian Oct1, Oct2 and Oct11 are found. Pdm1 and Pdm2 are required to maintain self-renewing asymmetric divisions in neuronal progenitor cells and therefore to generate post-mitotic daughter neurons (Bhat and Apsel, 2004). Interestingly, a specifically phosphorylated form of mammalian Oct1 localizes to mitotic structures, such as spindle pole bodies and the midbody, in symmetrically dividing HeLa and fibroblast cells (Kang et al., 2011). The mechanism underlying this function of Pdm1/2, and the potential role of Oct1 modification in asymmetrically dividing cells, have not been tested.

### Table 4. POU domain proteins in three common model organisms

<table>
<thead>
<tr>
<th>Organism (no. of known examples*)</th>
<th>Mammalian class/ortholog</th>
<th>Mutant phenotype</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C. elegans</strong> (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNC-86</td>
<td>POU IV family</td>
<td>Neuroblast differentiation, asymmetric division</td>
<td>Finney et al., 1988</td>
</tr>
<tr>
<td>CEH-6</td>
<td>POU III family</td>
<td>Neuronal, rectal epithelial, excretory cell</td>
<td>Burglin and Rukun, 2001</td>
</tr>
<tr>
<td>CEH-18</td>
<td>POU II family</td>
<td>Oocyte meiotic prophase arrest</td>
<td>Greenstein et al., 1994</td>
</tr>
<tr>
<td><strong>D. melanogaster</strong> (5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pdm1 (Nub)</td>
<td>POU II family</td>
<td>Wing hinge, wing sensory bristles, CNS asymmetric division</td>
<td>Ng et al., 1995; Neumann and Cohen, 1998; Bhat and Apsel, 2004</td>
</tr>
<tr>
<td>Pdm2 (Miti)</td>
<td>POU II family</td>
<td>Neuroblast differentiation, CNS asymmetric division</td>
<td>Yang et al., 1993; Bhat and Apsel, 2004</td>
</tr>
<tr>
<td>Pdm3</td>
<td>POU VI family</td>
<td>Olfactory neuron axon targeting and receptor gene expression</td>
<td>Tichy et al., 2008</td>
</tr>
<tr>
<td>Ct1α (Vvl, Drifter)</td>
<td>POU III family</td>
<td>Trachea, glia migration, wing veins</td>
<td>Certel and Thor, 2004</td>
</tr>
<tr>
<td>Acj6 (I-POU)</td>
<td>POU IV family</td>
<td>Olfactory neuron receptor gene expression</td>
<td>McKenna et al., 1989; Ayer and Carlson, 1991; Clyne et al., 1999</td>
</tr>
<tr>
<td><strong>D. rerio</strong> (12)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pou1f1</td>
<td>Pit1</td>
<td>Untested</td>
<td></td>
</tr>
<tr>
<td>Pou2f1a/b</td>
<td>Oct1</td>
<td>Untested</td>
<td></td>
</tr>
<tr>
<td>Pou3f1</td>
<td>Oct6</td>
<td>Untested</td>
<td></td>
</tr>
<tr>
<td>Pou3f2</td>
<td>Oct7</td>
<td>Untested</td>
<td></td>
</tr>
<tr>
<td>Pou3f3a/b</td>
<td>Oct8 (Brn1)</td>
<td>Untested</td>
<td></td>
</tr>
<tr>
<td>Pou4f1</td>
<td>Brn3a</td>
<td>Untested</td>
<td></td>
</tr>
<tr>
<td>Pou4f2</td>
<td>Brn3b</td>
<td>Untested</td>
<td></td>
</tr>
<tr>
<td>Pou4f3</td>
<td>Brn3c</td>
<td>Untested</td>
<td></td>
</tr>
<tr>
<td>Pou5f1 (Pou2*)</td>
<td>Oct4</td>
<td>Late-phase epiboly microtubule and F-actin defects downstream of defective EGF expression, somites, notochord, endoderm; midbrain and hindbrain specification</td>
<td>Belting et al., 2001; Burgess et al., 2002; Lachnit et al., 2008; Onichtchouk, 2012; Song et al., 2013</td>
</tr>
<tr>
<td>Pou6f1</td>
<td>Brn5</td>
<td>Untested</td>
<td></td>
</tr>
</tbody>
</table>

*Based on the ‘POU domain’ annotation in the UCSC genome browser (http://genome.ucsc.edu/) and/or annotation in the literature. Information on Danio rerio might not be complete.

*Pou2, which refers to a class V Oct protein, and class II factors such as Pou2f1 are not to be confused.
Box 2. Zebrafish Oct proteins

The zebrafish genome encodes at least ten non-redundant POU domain-containing factors, five of which are clear Oct protein orthologs (Table 4). The roles of most of these are poorly characterized, although Pou5f1 (Pou2) is the best studied (Belting et al., 2001; Burgess et al., 2002; Lachnit et al., 2008; Onichtchouk, 2012; Song et al., 2013). Historically, the term Pou2 has been applied to numerous POU class V Oct4 orthologs in different vertebrates, including zebrafish. These Oct4-like class V Pou2 factors should not be confused with POU class II transcription factors. Although zebrafish Pou5f1 and mammalian Oct4 are related to each other and both knockout results in developmental defects, their expression patterns are different and so the two proteins cannot be directly equated. Fish deficient in maternal zygotic Pou5f1 display pleiotropic phenotypes around the time of gastrulation. These include overall dorsralization of the embryo, defective Egf expression leading to cleftskelatal defects and delay of epiboly, and deficient endoderm formation (Burgess et al., 2002; Lachnit et al., 2008; Onichtchouk, 2012; Song et al., 2013). In addition, in the developing brain Pou5f1 ablation results in severe defects in the midbrain and hindbrain primordia (Belting et al., 2001; Burgess et al., 2002). This complex phenotype suggests that the single zebrafish Pou5f1 activity executes the functions of multiple mammalian paralogs, some of which are unlikely to be specific to stem cells. Among vertebrates, zebrafish Pou5f1 might be somewhat of an outlier in this regard. Whereas zebrafish Pou5f1 cannot reprogram mammalian fibroblasts to pluripotency, medaka and axotolx Pou5f1 can (Tapia et al., 2012). These findings suggest that this function is ancestral and has been lost in zebrafish.

studies of Oct proteins from a number of other organisms (e.g., Xenopus, fugu, medaka, axotolx, opossum). Most of this work involves comparative analysis of Oct4 orthologs and is beyond the scope of this Primer. Work on other Oct proteins in these organisms largely awaits investigation.

Not all Oct proteins are likely to promote stem cell phenotypes. Some Oct proteins could antagonize stem cell phenotypes by opposing other Oct proteins, for example by competitive binding and/or through differential transcriptional regulatory mechanisms. One interesting example comes from mammalian brain development and POU class III Oct proteins. In neural progenitor cells of the forebrain and midbrain the members of this class – Oct6, Brn2, Brn1 and Brn4 – all associate with an Oct4 upstream enhancer. By contrast, an antagonistic transcription factor, Gbx2, associates and represses this locus in hindbrain (Inoue et al., 2012). These results reveal a complex interplay between multiple Oct proteins and other factors at a single enhancer binding site. Also, different isoforms expressed from the same gene locus can oppose other isoforms; for example, different Oct2 isoforms appear to play opposing roles in the neuronal differentiation of ESCs (Theodorou et al., 2009).

Oct proteins in development and stem cells: an interim conclusion

Oct protein research should be spurred on by the impressive findings of the last 10 years regarding Oct4. Nevertheless, many unanswered questions remain regarding their mechanisms of action and developmental functions. Importantly, the great majority of zebrafish Oct proteins remain completely uncharacterized. These issues are best addressed by leveraging new molecular tools that enable a better understanding of the target genes and regulatory mechanisms utilized by Oct proteins. Powerful new methods for ablating specific genes in zebrafish can be applied to manipulate Oct proteins. For many of the mammalian family members, conditional knockout alleles have not been constructed. The ability to ablate the activity of different factors in a given cell type and in a temporally controlled fashion will help to further elucidate the developmental roles of these important transcription factors.

Based on the prior work in this field, the results of future investigations are likely to be complex but also enlightening. One specific area of focus will be gene poising by Oct proteins and how this mechanism enables their different roles in pluripotent stem cell, somatic stem cell and progenitor cell compartments. The fact that Oct proteins are implicated in gene poising could complicate the interpretation of their function during development. A specific expression or DNA binding pattern may only partially indicate function, as the protein might be poising critical target genes that become activated at later developmental stages. Therefore, investigations of developmental phenotypes associated with Oct proteins in different model organisms and developmental paradigms must go hand-in-hand with advances in understanding the mechanisms by which these proteins function.

Acknowledgements

We thank R. Dorsky, D. Stillman and members of the D.T. laboratory for their critical reading of the manuscript. We apologize to authors whose work was omitted from this article due to space limitations.

Funding

Work in the D.T. laboratory was supported by the National Institutes of Health and by an award from the Concern Foundation. Deposited in PMC for release after 12 months.

Competing interests statement

The author declares no competing financial interests.

References


