Cell

A DNA Repair Complex Functions as an Oct4/Sox2 Coactivator in Embryonic Stem Cells

Yick W. Fong,[1] Carla Inouye,[1] Teppei Yamaguchi,[1] Claudia Cattoglio,[1] Ivan Grubisic,[2] and Robert Tjian[1,3,*]
[1]Howard Hughes Medical Institute, Department of Molecular and Cell Biology
[2]UC Berkeley-UCSF Graduate Program in Bioengineering
[3]Li Ka Shing Center for Biomedical and Health Sciences
University of California, Berkeley, Berkeley, CA 94720, USA
*Correspondence: jmlim@berkeley.edu
DOI 10.1016/j.cell.2011.08.038

SUMMARY

The transcriptional activators Oct4, Sox2, and Nanog cooperate with a wide array of cofactors to orchestrate an embryonic stem (ES) cell-specific gene expression program that forms the molecular basis of pluripotency. Here, we report using an unbiased in vitro transcription-biochemical complementation assay to discover a multisubunit stem cell coactivator complex (SCC) that is selectively required for the synergistic activation of the *Nanog* gene by Oct4 and Sox2. Purification, identification, and reconstitution of SCC revealed this coactivator to be the trimeric XPC-nucleotide excision repair complex. SCC interacts directly with Oct4 and Sox2 and is recruited to the *Nanog* and *Oct4* promoters as well as a majority of genomic regions that are occupied by Oct4 and Sox2. Depletion of SCC/XPC compromised both pluripotency in ES cells and somatic cell reprogramming of fibroblasts to induced pluripotent stem (iPS) cells. This study identifies a transcriptional coactivator with diversified functions in maintaining ES cell pluripotency and safeguarding genome integrity.

INTRODUCTION

The molecular events leading to the maintenance of pluripotency in embryonic stem (ES) cells and reacquisition of a stem-like state in induced pluripotent stem (iPS) cells during somatic reprogramming represent mechanistically distinct processes that converge on a set of remarkably similar transcriptional events that underpin the pluripotent state. Both ES and iPS cells depend on fundamental transcription frameworks that are governed by a common set of "core" stem cell-specific transcription factors, namely Oct4, Sox2, and Nanog (Jaenisch and Young, 2008). These activators, in turn, collaborate with both ubiquitous and cell type-specific transcription factors to orchestrate complex gene expression programs that confer upon stem cells the unique ability to safeguard stemness while remaining poised to execute a broad range of developmental programs that drive lineage specification (Boyer et al., 2005; Chen et al., 2008; Kim et al., 2008; Marson et al., 2008).

Proper execution of these highly regulated processes by sequence-specific transcription factors often requires the coordinated recruitment of coactivator proteins to their cognate promoters. For example, transcriptional activators direct histone modifiers (e.g., CBP/p300) and chromatin remodelers (e.g., PBAF/BAF) to gene promoters to alter chromatin structure toward a state that is more permissive to transcriptional activation (Näär et al., 2001). Independent of chromatin, a variety of activators recruit other classes of coactivators, such as the multisubunit Mediator, various TBP/TAF complexes, SRC, etc., via direct protein-protein interactions to execute specific transcriptional programs. This class of coactivators often serves as molecular "adaptors" by bridging activators to the general transcription machinery, thereby mediating the synergistic response by these activators (Näär et al., 1999). Interestingly, subunits of Mediator have also been shown to interact with cohesin possibly to promote DNA looping and thereby facilitate long-distance interactions between enhancers and core promoters in vivo (Kagey et al., 2010). Indeed, such coactivators are often multifunctional and can activate transcription through chromatin-dependent as well as independent mechanisms. Further expanding the transcriptional repertoire of coactivator complexes, their protein levels and subunit compositions are frequently modulated in a developmental stage and cell type-specific manner (Roeder, 2005; Taatjes et al., 2004). Additionally, these protein-protein-driven coactivator-activator transactions are often critical nodes in various signal transduction pathways and can serve as molecular "sensors" by integrating cell-intrinsic and -extrinsic cues, thereby coupling gene networks with specific cellular responses to produce complex biological programs of gene expression (Rosenfeld et al., 2006).

Totipotent ES cells employ these same sets of coactivators in conjunction with special activators such as Oct4 and Sox2 to regulate transcription of a large number of genes, including *Nanog*, that form the molecular basis of pluripotency (Gao et al., 2008; Kagey et al., 2010; Kidder et al., 2009; Tutter et al., 2009). The transcription of *Nanog* is exquisitely dependent on

Oct4 and Sox2 (Kuroda et al., 2005; Rodda et al., 2005). However, coexpression of Oct4 and Sox2 failed to robustly activate a *Nanog* promoter reporter construct in differentiated cells like 293 or NIH 3T3 cells, even though Mediator, p300/CBP, and PBAF/BAF complexes remain abundantly expressed and active (Rodda et al., 2005). This led us to speculate that one or more as yet unidentified stem cell-specific cofactors may be required to activate the transcription of *Nanog* and other Oct4/Sox2 target genes in ES cells. Indeed, recent studies of germ cells and differentiated somatic cells revealed that even parts of the general transcriptional machinery may be radically altered in a tissue- or cell-specific context (Goodrich and Tjian, 2010; Müller et al., 2010). Diversification of the transcriptional apparatus may therefore represent a fundamental strategy, particularly in ES cells, to cope with the multidimensional nature of transcription programs that must be precisely tuned to both maintain pluripotency and, at the same time, allow for lineage-specific programs of differentiation (Liu et al., 2011).

The human *Nanog* promoter contains a prototypic composite oct-sox *cis*-acting regulatory element located immediately upstream of the transcription start site that is conserved across several mammalian species (Kuroda et al., 2005; Rodda et al., 2005). A *Nanog* promoter-GFP reporter construct containing a DNA fragment encompassing this promoter-proximal oct-sox element is sufficient to recapitulate the robust expression pattern of endogenous Nanog in ES cells in an Oct4-, Sox2-dependent manner (Kuroda et al., 2005; Rodda et al., 2005). Unbiased genome-wide motif searching analyses of Oct4 in both mouse and human ES cells identified an oct-sox composite consensus sequence element, confirming that Oct4 likely orchestrates an ES-specific gene expression program primarily through cooperation with Sox2 (Chen et al., 2008; Loh et al., 2006). Because the oct-sox *cis*-control element in the *Nanog* promoter represents a common configuration that is present in the promoters of many other Oct4- and Sox2-activated genes in ES cells, the well-characterized *Nanog* proximal promoter provided us with a useful model template for identifying uncharacterized transcriptional cofactors required for Oct4- and Sox2-directed activation. Therefore, we took advantage of a fully reconstituted in vitro transcription system in which one can unambiguously and systematically test and identify transcriptional cofactors that may be directly required to potentiate Oct4- and Sox2-dependent gene activation of *Nanog*. Here, we report the biochemical purification and identification of a multisubunit stem cell coactivator (SCC) that is required for the synergistic activation of *Nanog* by Oct4 and Sox2 in vitro. After extensive biochemical characterization, we surprisingly found that SCC is none other than the XPC-RAD23B-CETN2 (XPC) nucleotide excision repair (NER) complex. SCC/XPC interacts directly with Oct4 and Sox2 and co-occupies a majority of Oct4 and Sox2 targets genome-wide in mouse ES cells. Importantly, SCC/XPC is required for stem cell self-renewal and efficient somatic cell reprogramming. Thus, our findings unmask an unanticipated selective coactivator role of an NER complex in transcription in the context of ES cells and may provide a previously unknown molecular link that couples stem cell-specific transcription to DNA damage response with potential implications for enhanced ES cell genome stability.

RESULTS

Detection of an Oct4- and Sox2-Dependent Coactivator Activity in EC and ES Cells

Having chosen the *Nanog* promoter as our model template, we next set out to develop an in vitro reconstituted transcription assay that could recapitulate the Oct4- and Sox2-dependent transactivation at the *Nanog* promoter observed in vivo. To enhance the sensitivity of the assay, we inserted four copies of the *Nanog* oct-sox-binding sites immediately upstream of the native oct-sox element found in the human *Nanog* promoter. Our basal in vitro transcription assay consisted of purified recombinant TFIIA, -B, -E and -F together with immunoaffinity-purified native RNA polymerase II, TFIID, and TFIIH (Figure S1A available online). When purified Oct4 and Sox2 were added to this reconstituted transcription system, only a very weak activation of the *Nanog* promoter was detected (Figure 1A, lanes 1 and 2). As a control, we could show that the same complement of general transcription factors (GTFs) was able to support strong Sp1-dependent activation from a GC box-containing "generic" transcription template (G3BCAT) (Figure 1A, lanes 5 and 6). This initial result suggested that efficient activation of *Nanog* by Oct4 and Sox2 may require additional cofactors to potentiate a full activator-dependent response.

We reasoned that such a putative coactivator ought to be selectively active in pluripotent cell types that express Nanog under the control of Oct4 and Sox2. For example, NTERA-2 (NT2) is a pluripotent human embryonal carcinoma (EC) cell line that expresses Oct4, Sox2, and Nanog and shares with ES cells core molecular mechanisms that govern self-renewal (Pal and Ravindran, 2006). Detailed expression profiling of NT2 and bona fide human ES cell lines revealed many similarities, including robust expression of Nanog (Schwartz et al., 2005; Sperger et al., 2003). However, unlike human ES cells, NT2 cell culture can be more readily scaled up, a prerequisite to generating sufficient quantities of starting materials for the biochemical purification of putative Oct4/Sox2 coactivators. We therefore chose extracts derived from NT2 cells as our starting material in our efforts to develop a "biochemical complementation" assay to hunt for pluripotent stem cell-selective cofactors.

We first fractionated NT2 nuclear extracts by conventional phosphocellulose ion exchange chromatography. Next, we supplemented our "basal" reconstituted transcription reactions with various salt-eluted fractions from the phosphocellulose column to see whether there was any activity that could restore Oct4/Sox2-dependent activation of our *Nanog* promoter. This strategy allowed us to unmask an activity in the high salt phosphocellulose fraction (P1M) prepared from NT2 nuclear extracts (but not HeLa extracts) (Figure S1B) that strongly potentiated transcription of the *Nanog* promoter in an Oct4- and Sox2-dependent manner using either a naked (Figure 1A, lanes 3 and 4) or a *Nanog* chromatin template assembled with a crude *Drosophila* cytosolic extract (data not shown). This new cofactor activity is selectively required for transcription of *Nanog*, as it had no effect on either basal- or Sp1-activated transcription from a control G3BCAT template (Figure 1A, lanes 5–8). Importantly, this P1M fraction also stimulated the Oct4/Sox2-dependent transcription from a native *Nanog* promoter template (Figure 1B),