Site-specific Disruption of the Oct4/Sox2 Protein Interaction Reveals Coordinated Mesendodermal Differentiation and the Epithelial-Mesenchymal Transition*

Received for publication, June 28, 2016 Published, JBC Papers in Press, July 1, 2016, DOI 10.1074/jbc.M116.745414

Xiao Pan1, Xiaohui Cang1, Songsong Dan1, Jingchao Li2, Jie Cheng2, Bo Kang2, Xiaotao Duan1, Binghui Shen1, and Ying-Jie Wang1,2

From the 1College of Life Sciences, Zhejiang University, 866 Yuhangtang Road, Hangzhou, Zhejiang 310058, China, the 2State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, First Affiliated Hospital, School of Medicine, Zhejiang University, 79 QingChun Road, Hangzhou, Zhejiang 310003, China, the 3State Key Laboratory of Toxicology and Medical Countermeasures, Beijing Institute of Pharmacology and Toxicology, Beijing 100850, China, and the 4Department of Radiation Biology, City of Hope National Medical Center and Beckman Research Institute, Duarte, California 91010

Although the Oct4/Sox2 complex is crucial for maintaining the pluripotency of stem cells, the molecular basis underlying its regulation during lineage-specific differentiation remains unknown. Here, we revealed that the highly conserved Oct4/Lys-156 is important for maintaining the stability of the Oct4 protein and the intermolecular salt bridge between Oct4/Lys-151 and Sox2/Asp-107 that contributes to the Oct4/Sox2 interaction. Post-translational modifications at Lys-156 and K156N, a somatic mutation detected in bladder cancer patients, both impaired the Lys-151—Asp-107 salt bridge and the Oct4/Sox2 interaction. When produced as a recombinant protein or over-expressed in pluripotent stem cells, Oct4/K156N, with reduced binding to Sox2, significantly down-regulated the stemness genes that are cooperatively controlled by the Oct4/Sox2 complex and specifically up-regulated the mesendodermal genes and the SNAIL family genes that promote the epithelial-mesenchymal transition. Thus, we conclude that Oct4/Lys-156-modulated Oct4/Sox2 interaction coordinately controls the epithelial-mesenchymal transition and mesendoderm specification induced by specific differentiation signals.

Cell fate determination during embryogenesis remains a major focus of developmental biology. Embryonic stem cells (ESCs)3 derived from the inner cell mass of blastocyst-stage embryos and their malignant counterpart embryonal carcinoma cells (ECCs) present in teratocarcinomas can grow in the Petri dish indefinitely while maintaining their pluripotency (i.e. the potential to differentiate into any of the three embryonic germ layers as follows: endoderm, mesoderm, or ectoderm), offering a unique tool for modeling cell fate determination in vitro (1). With such a cultured pluripotent stem cell (PSC) system, it has been well established that the undifferentiated state of PSCs is governed by a network of transcription factors, including Oct4, Sox2, Nanog, Klf5, Esrrb, and Tbx3, which repress differentiation-promoting genes while activating pluripotency genes (2). Among them, Oct4, Sox2, and Nanog are considered to be the master pluripotency factors as each of them is unique and indispensable for pluripotency and self-renewal (3, 4). Remarkably, they can regulate their own or each other’s gene transcription via combinatorial interactions, forming a positive feedback transcriptional regulatory circuit that suppresses differentiation (3, 5). Furthermore, they are at the center of a highly integrated regulatory network composed of many transcriptional and epigenetic regulators (6).

Oct4 (encoded by POLISF1 gene) is a member of the class 5 POU (Pit-Oct-Unc) family of transcription factors, which specifically binds to the canonical octamer motif (with consensus sequence ATGC(A/T)AAAT) at target gene enhancer or promoter regions via the cooperation of two subdomains known as the POUs (for POU-specific, binding to the sequence ATGC) and POUs (for POU homeodomain, binding to the sequence A(T)AAAT) that is connected by a flexible linker (7). Depending on the specific core octamer DNA sequence and its flanking sequence, POUs and POUs domains can position in different orientations relative to each other to allow Oct4 to form monomers, homodimers in PORE motif (ATTTGAAAT/ GGGAAT)- or the More PORE motif (ATGCAATATGCAT)-binding configuration, or to form heterodimers with other transcription factors such as Sox2 (7). Compelling evidence shows that the maintenance of the Oct4/Sox2-centered complex is of paramount importance for pluripotency, and altering the stoichiometry of cellular Oct4/Sox2 will trigger differentiation. For instance, knocking down either Oct4 or Sox2 in mouse ESCs led to their differentiation into trophoectoderm-like cells, and elevating Oct4 or Sox2 protein levels in ESCs induced their
Oct4/Lys-156 Controls Differentiation and EMT
differentiation into either primitive endoderm and mesoderm or most non-endoderm lineages, respectively (6). Furthermore, a contrasting pattern of Oct4 versus Sox2 expression during cell fate determination was reported by Thomson et al. (8), who showed that Oct4 is up-regulated in cells choosing the mesendoderm fate but repressed in cells choosing the neural ectoderm fate, and Sox2 exhibited the opposite expression pattern in both cell fate choices. Thus, it is likely that during gastrulation the developing embryo has differentiation signals that continuously and asymmetrically modulate Oct4 and Sox2 protein levels, altering their binding pattern and binding targets in the genome and leading to cell fate choices. Multiple studies have established that nodal/activin A, BMP, WNT, and FGF pathways are the major signaling pathways that regulate the formation of three germ layers (9). However, it remains unknown how these differentiation signaling pathways initiate the disruption of the Oct4/Sox2 complex and alter the expression of Oct4 and Sox2 during lineage specification.

Biochemical, structural, and cellular analyses have provided substantial mechanistic insights into the Oct4/Sox2 interaction in the context of undifferentiated PSCs. It is established that in ESCs Oct4 and Sox2 can bind cooperatively to two adjacent cis-regulatory elements known as the octamer motif ATGC(A/T)AAT and the SOX-binding motif C(T/A)TTGTT, respectively (10). Recent molecular simulations indicated that Sox2 influences the orientation and dynamics of the DNA-bound configuration of Oct4 (11). Single molecule imaging study on how Oct4 and Sox2 dynamically search for and assemble on their cognate DNA target sites gave rise to a model in which Sox2 engages with the chromatin first and primes the target site for subsequent Oct4 binding in ESCs, and Oct4 in turn helps to stabilize the Oct4/Sox2 complex at composite recognition sites (12). This stepwise-ordered assembly at endogenous chromatin sites indicates that the Oct4/Sox2 heterodimers are unlikely to form prior to binding DNA, consistent with their DNA-dependent interaction reported in live cells (13).

Based on the solved Oct1/Sox2/DNA crystal structures (14, 15) and the Oct4 PORE homodimer structure (16), it can be inferred that the direct interaction between Oct4 and Sox2 is mediated by the POUα α1 helix of Oct4 and the HMG α3 helix of Sox2. Depending on the Oct4/Sox2-binding motif, two models of DNA-bound Oct4/Sox2 were built (7). In one configuration, the octamer motif follows right after the SOX-binding site (Fig. 1A). In this canonical configuration, POUs α1 helix of Oct4 and HMG α3 helix of Sox2 form the Oct4/Sox2 interface, and Fig. 1B highlights the key residues at this interface. Based on our MD simulation results, the most significant electrostatic interaction at the Oct4/Sox2 interface was the salt bridge formed between Oct4/ Lys-151 and Sox2/Asp-107 that is equivalent to the Oct1/Lys- 14/Sox2/Asp-71 (whose amino acid numbering was based on the solved crystal structure) electrostatic interaction inferred from a previous study (15). Because the alterations of electrostatic forces and local structures by multiple post-translational modifications (PTMs) at certain residues are known to critically regulate protein/protein interaction (18), we hypothesized that the key residues mediating the Oct4/Sox2 interaction may be modified during lineage specification, and therefore the same residues may have differential PTMs between PSCs and differentiated cells. To this end, we employed mass spectrometry and an in vitro PTM system (19) to determine multiple PTM profiles of the POUs α1 domain of Oct4 derived from a pluri- 120, Lys-144, Lys-151, Lys-154, and Lys-156) were at the α1 helix that is crucial for forming the Oct4/Sox2 interaction interface (Fig. 1C). In a subsequent expanded experiment involving two PSC lines and two differentiated cell lines, residue Lys-151 was post-translationally modified in two differentiated cell lines but not in two PSCs, whereas residue Lys-156 was only modified in PSCs (Fig. 1D).

Somatic mutations (particularly those cancer-associated driver mutations) identified in many protein kinases are found to correspond to sites and regions that are important for protein function and structure, such as ATP- and substrate-binding sites, conserved residues or domains, and the apolar core of the protein (20). We therefore reasoned that somatic mutation corresponding to amino acid substitutions in Oct4 that are distributed at the Oct4/Sox2 interface may have significant structural/functional relevance in physiological or pathological set-
tings. We searched for two somatic mutation databases and summarized the results in Fig. 2. In the Catalogue of Somatic Mutations in Cancer (COSMIC) database, 13 somatic mutations have been identified so far in the POU5F1 gene from 8651 unique samples, and three of them are located at the POU5 domain. In the Cancer Genome Atlas (TCGA) database, there are 23 somatic mutations of which nine are at the POU5 domain (data not shown). When the two databases were combined...
together, two somatic mutations (K156N and F194L, Fig. 2A, marked in red) overlapped, and only K156N was located at the POUs α1 helix. Sequence alignment showed that Lys-156 is highly conserved across many species (Fig. 2B), implicating its importance for Oct4 function. Taken together, because the highly conserved Oct4/Lys-156 is located close to the Oct4/Sox2 interface (Fig. 1B), differentially modified between undifferentiated and differentiated cells (Fig. 1, C and D), and detected as a K156N somatic mutation in multiple bladder cancer specimens (Fig. 2A), we focused on this particular site in the following studies.

Oct4/Lys-156 Regulates the Oct4/Sox2 Interaction via Multiple Mechanisms—To assess the potential role of Oct4/Lys-156, we determined the impact of substituting the Lys-156 with various mutants (K156N and K156A mutants losing positive charges and K156R mutant retaining positive charge) on Oct4/Sox2 binding in a series of in vitro assays as described previously (21). First, in the electrophoretic mobility shift assays (EMSAs), a biotin-labeled probe of NANOG promoter harboring the no-gap canonical Oct4/Sox2-binding motif was incubated with recombinant Sox2 mixed with either wild type recombinant Oct4 protein (Oct4-WT) or one of the recombinant Oct4 mutant proteins, and the binding of the Oct4/Sox2 heterodimer to the NANOG probe was reflected by a mobility shift of the protein-bound probes versus free probes as resolved by native PAGE. Oct4-WT alone can bind to the NANOG probe, Oct4/K156R had the same level of binding, and either Oct4/K156N or Oct4/K156A had much reduced binding (Fig. 3A, arrow 2). The putative Oct4/Sox2 probe band (Fig. 3A, arrow 1) was strong with the Sox2-Oct4-WT or Sox2-Oct4/K156R combinations, but was significantly reduced when Oct4-WT was replaced with either Oct4/K156N or Oct4/K156A, possibly as a result of reduced binding of the mutant Oct4 proteins to the NANOG probe. EMSA with the UTF1 probe gave a fairly similar result as the NANOG probe (Fig. 3B). However, it is difficult to distinguish by EMSA whether the specific mutations at Lys-156 reduced the direct binding of Oct4 with Sox2 or with the probes.

To circumvent the above caveat and assess the Oct4/Sox2 binding in a more cellular-relevant context, the same NANOG probe was incubated with lysates from NCCIT cells infected with lentiviruses harboring a POU5F1 shRNA (to knock down endogenous Oct4) and one of the FLAG-POU5F1 variants (to functionally replace the endogenous Oct4) (Fig. 3, C and D), followed by pulldown of biotin-labeled probes using streptavidin beads and detection of probe-bound proteins by immunoblotting. Despite the similar total amounts of overexpressed FLAG-Oct4 variant proteins and endogenous Sox2 proteins present in the reaction system and the similar amounts of probe-bound FLAG-Oct4 proteins pulled down by streptavidin beads, there were many reduced probe-bound Sox2 proteins in the K156N or K156A group compared with the WT or K156R group (Fig. 3E), suggesting that the specific Oct4/Lys-156 mutations resulted in a dramatically diminished binding with Sox2 protein besides their reduced binding with the target genes. Accordingly, when similar levels of ectopically expressed FLAG-Oct4 variant proteins were pulled down, there were less Sox2 proteins co-precipitated with K156N or K156A compared with WT or K156R (Fig. 3F). As a functional validation, the transcription of stemness genes that are dually regulated by the Oct4/Sox2 complex was examined, which showed a dramatic decrease in NANOG and UTF1 transcripts and a slight increase in FBX15 transcripts (Fig. 3G). It should be noted that to eliminate the potential effect of the Oct4 variants on their own tran-
FIGURE 3. Oct4/Lys-156 is crucial for regulating the Oct4/Sox2 interaction. A and B, EMSA examining the binding of purified recombinant His-tagged Oct4 (WT and its variants) and Sox2 protein with biotinylated NANOG probe (A) or biotinylated UTF1 probe (B). Arrows 1–3 in A indicate the biotinylated NANOG probe bound by Oct4-Sox2 complex, Oct4 only, or Sox2 only, respectively, and the arrow in B indicates the biotinylated UTF1 probe bound by Oct4-Sox2 complex. C, NCCIT cells were infected with “shRNA/H11001 POU5F1 variants” viruses or with a relatively high titer of “shRNA/H11001 POU5F1 WT” viruses. Cells were harvested and subjected to immunoblotting (IB) with the indicated antibodies. Overexpressed FLAG-Oct4 variants (FLAG-Oct4, upper band) and endogenous Oct4 (lower band) were indicated. D, cells treated as in C were harvested, and the mRNA levels of POU5F1 relative to GAPDH were determined by qRT-PCR. E, biotinylated DNA pulldown assay. Biotinylated NANOG probe was incubated with the NCCIT cell lysates containing normalized equivalent amount of overexpressed FLAG-tagged Oct4 WT or a variant, followed by addition of streptavidin-conjugated agarose beads. The bead-DNA-protein complexes were collected and immunoblotted with indicated antibodies. F, NCCIT cells overexpressing FLAG-Oct4 variants were lysed and immunoprecipitated with anti-FLAG M2 beads. To compare the interaction between Sox2 and Oct4 variants, the loading of immunocomplexes was normalized to make Oct4 approximately the same level. WCL, whole cell lysate. G, knockdown of endogenous Oct4 together with overexpressing POU5F1-K156N in NCCIT cells led to a significant decrease in mRNA levels of the stemness genes governed by the Oct4-Sox2 complex, while overexpressing POU5F1 WT or POU5F1-K156R had no prominent effects. The data were expressed as mean ± S.D. of triplicate measurements from one of three independent experiments, which gave similar results. The difference between the “shRNA + POU5F1-K156N” group and the “shRNA + POU5F1-WT” group was evaluated using the one-way ANOVA and marked if it was statistically significant. ***, p < 0.001. H, NCCIT cells were infected with “shRNA + POU5F1-K156N” viruses or with a relatively lower titer of “shRNA + POU5F1 WT” viruses (as compared with that in C). Cells were harvested and subjected to immunoblotting with the indicated antibodies.
scription, we used the U6 promoter and hPGK promoter (rather than endogenous POLISF1 promoter) to drive the transcription of POLISF1 variants + POLISF1 shRNA that should be independent from the feedback regulation of the expressed Oct4 proteins. To achieve similar protein level as its counterparts, we had to infect cells with higher titers of the "POU5F1-K156N" lentiviruses that were reflected by a significantly higher Oct4/K156N mRNA level (shown in Fig. 3, D and G). The obvious disconnection between the high Oct4/K156N mRNA level and its low protein level (shown in Fig. 3, C and H) is a strong indication that the translated FLAG-Oct4/K156N protein may be degraded by certain mechanisms. Taken together, we concluded from the above experiments that Oct4/Lys-156 critically controls the Oct4/Sox2 interaction, and the transcription of genes dually targeted by the Oct4/Sox2 complex in PSCs.

To better understand the molecular mechanism by which Oct4/Lys-156 regulates the Oct4/Sox2 interaction and how the Oct4/K156N mutant impairs such interaction, we performed 500-ns MD simulations for both wild type Oct4 and Oct4/K156N mutant in the Oct4/Sox2-HOXB1 complex. Our simulation results showed that from 100 ns onward, Lys-151 of the wild type Oct4 formed a stable salt bridge with Asp-107 of Sox2, although in the mutant complex this stable salt bridge did not form (Fig. 4, A and B). To experimentally validate the importance of this specific intermolecular salt bridge in mediating the Oct4/Sox2 interaction, we generated Oct4/K151A and Sox2/D107A recombinant proteins and evaluated their interaction in the presence of the NANOG probe using EMSA. A combination of both mutant proteins or Oct4/K151A alone greatly diminished the Oct4/Sox2 binding. However, Sox2/D107A alone only marginally reduced the binding (Fig. 4, C). This indicates that Oct4/Lys-151 plays a more important role than Sox2/Asp-107 in mediating the Oct4/Sox2 interaction, and additional neighboring residues in Sox2 (e.g. Sox2/Pro-106) may also contribute to the interaction as inferred from the previous study (15).

Because Oct4/Thr-159 appeared to reside at the Oct4/Sox2 binding interface (Fig. 1B), and Oct4/T159P somatic mutation was reported in the TCGA database (Fig. 2A), we also examined the binding of recombinant Oct4/T159P protein to the Sox2 protein in the presence of the NANOG probe using EMSA. Similar to Oct4/K156N, Oct4/T159P also had dramatically reduced binding with Sox2 (Fig. 4D).

Oct4/Lys-156 Regulates the Protein Stability and Subcellular Localization of Oct4 in ECCs—Our MD simulations further showed that in the Oct4/WT-Sox2-HOXB1 complex, from around 90 ns onward, Lys-151 of the wild type Oct4 formed a stable salt bridge with Asp-107 of Sox2, although in the mutant complex this stable salt bridge did not form (Fig. 4, A and B). To experimentally validate the importance of this specific intermolecular salt bridge in mediating the Oct4/Sox2 interaction, we generated Oct4/K151A and Sox2/D107A recombinant proteins and evaluated their interaction in the presence of the NANOG probe using EMSA. A combination of both mutant proteins or Oct4/K151A alone greatly diminished the Oct4/Sox2 binding. However, Sox2/D107A alone only marginally reduced the binding (Fig. 4C). This indicates that Oct4/Lys-151 plays a more important role than Sox2/Asp-107 in mediating the Oct4/Sox2 interaction, and additional neighboring residues in Sox2 (e.g. Sox2/Pro-106) may also contribute to the interaction as inferred from the previous study (15).

Because Oct4/Thr-159 appeared to reside at the Oct4/Sox2 binding interface (Fig. 1B), and Oct4/T159P somatic mutation was reported in the TCGA database (Fig. 2A), we also examined the binding of recombinant Oct4/T159P protein to the Sox2 protein in the presence of the NANOG probe using EMSA. Similar to Oct4/K156N, Oct4/T159P also had dramatically reduced binding with Sox2 (Fig. 4D).
dues turned into a new part of the linker helix. Therefore, the single mutation K156N is likely to alter the structure of the linker region of Oct4, which is known to be crucial for somatic reprogramming (16). The impairment of the Oct4 protein stability by the K156N mutation was indicated by the root mean square fluctuation curves calculated from the last 200 ns of the MD simulation trajectory (Fig. 5A). A similar phenomenon was also observed in the MD trajectories for the Oct4/Sox2–FGF4 complex (Fig. 5B). However, when evaluating the potential effect of the Oct4 Lys-156–Asp-212 salt bridge on the Oct4/Sox2 interaction, Oct4/D212A mutant did not show significantly reduced binding with Sox2 (Fig. 5C), indicating that either Asp-212 plays a less important role than Lys-156 in maintaining the intramolecular Lys-156–Asp-212 salt bridge, and/or besides impairing the Lys-156–Asp-212 salt bridge, Oct4/K156N mutant has other effects that more significantly impact the Oct4/Sox2 interaction (such as impairing the Oct4/Lys-151/Sox2–Asp-107 intermolecular salt bridge as shown above).

We searched the entire UniProt database and found that in all the 74 POU family proteins, the position corresponding to Oct4/Lys-156 is non-exceptionally occupied with positively charged residues (Lys or Arg), and the position corresponding to Oct4/Asp-212 is occupied with negatively charged residues (Asp or Glu) (data not shown), suggesting that the extremely conserved Lys/Arg-156–Asp/Glu-212 salt bridge is likely crucial for the proper structure and function of the POU family proteins.

To experimentally examine the effect of K156N mutant on Oct4 protein stability, we transfected NCCIT cells with FLAG-POU5F1-WT or its variants, treated the cells with cycloheximide (CHX) to block the synthesis of new proteins for up to 12 h, and determined the protein levels of FLAG-Oct4 variants as a function of time by immunoblotting. It was evident that in NCCIT cells, overexpressed Oct4/K156N (Fig. 6A) or Oct4/K156R (Fig. 6B) proteins had a shorter half-life (<2 h) than Oct4/WT proteins (4–6 h), although Oct4/K156R had a similar half-life as Oct4/WT (Fig. 6C). However, in U87 cells, overexpressed Oct4/K156N (Fig. 6D) and Oct4/K156R (Fig. 6E) proteins had similar half-lives that were both longer than that of expressed Oct4/WT. To explore whether Oct4/K156N specifically decreases Oct4 protein stability, Oct4/T159P, another
Oct4/Lys-156 Controls Differentiation and EMT

mutant with similarly reduced binding capacity with Sox2, was tested for its turnover rate in the presence of CHX treatment. Unexpectedly, the Oct4/T159P mutant had a much slower degradation rate than Oct4/WT (Fig. 6F), a phenotype opposite that of Oct4/K156N. These results indicated that in the context of ECCs, the K156N mutation dramatically and uniquely destabilized the Oct4 protein, confirming the results obtained from the above MD simulation.

Based on the previous work from our own (21) and from others (22), endogenous Oct4 proteins or ectopically expressed Oct4/WT proteins degraded mainly via the proteasome-dependent pathway that can be blocked by MG-132, a specific and cell-permeable proteasome inhibitor. However, to our surprise, MG-132 pre-treatment did not completely block the degradation of Oct4/K156N protein (Fig. 7A), implicating its degradation via additional pathways. Because lysosome-dependent and autophagy-dependent degradation pathways represent other major mechanisms for protein degradation (23), we pre-treated the transfected NCCIT cells with chloroquine, which can inhibit both degradation pathways and monitor the degradation kinetics of both WT and K156N proteins. Remarkably, chloroquine treatment prolonged the half-lives of both Oct4/WT and Oct4/K156N protein to 8 and 4 h, respectively (Fig. 7B), indicating that a substantial fraction of Oct4/K156N protein and a minor fraction of Oct4/WT could be subjected to lysosome and/or autophagy-dependent degradation. Interestingly, blocking proteasome-dependent degradation seemed to also block the lysosome-dependent degradation because MG-132 alone almost totally inhibited the degradation of Oct4/WT proteins, and blocking lysosome-dependent degradation of Oct4/WT proteins could in turn reduce their proteasome-dependent degradation (comparing degradation time course by chloroquine treatment in Fig. 7B with that of no treatment in Fig. 6, A and B), implicating certain cross-talk between the two degradation pathways.

To localize the intracellular compartments where the above degradation occurred, we examined the FLAG-tagged Oct4 variant proteins in NCCIT cells pre-treated with vehicle, MG-132, or chloroquine for 6 h before being fixed and examined by immunofluorescence microscopy for the expressed FLAG-Oct4 variants. Scale bars, 50 μm.
ized in the nucleoplasm. Remarkably, MG-132 or chloroquine pre-treatment resulted in a redistribution of Oct4/WT to the nucleolus or cytoplasm, respectively (Fig. 7C). In comparison, with all treatments, the fluorescence intensity of Oct4/K156N was substantially lower than that of Oct4/WT. Pre-treatment with MG-132 caused a redistribution of Oct4/K156N to both the nucleolus and cytoplasm, although chloroquine pre-treatment mainly promoted its cytoplasmic relocation (Fig. 7C). Collectively, these results suggested that substituting the positively charged Lys-156 to a neutrally charged residue Asn or Ala can dramatically disrupt the Oct4 protein structure, attenuate its interaction with Sox2, and cause a redistribution of Oct4 from the nucleoplasm to the nucleolus or cytoplasm for accelerated degradation.

Oct4/Lys-156 Critically Controls the Mesendodermal Differentiation and the EMT Phenotype of PSCs—To assess the effect of the above site-specific disruption of the Oct4-Sox2 complex on ECC self-renewal and differentiation, we first examined the morphological changes of NCCIT cells infected with the lentiviruses harboring a POLISFI shRNA and one of the FLAG-tagged POLISFI-Lys-156 variants. Although the POLISFI-WT- and POLISFI-K156R-expressing cells had similar cellular morphology as non-transfection cells, elongated and irregular cell shapes were more frequently seen with cells expressing POLISFI-K156N (Fig. 8A), indicative of differentiation. Alkaline phosphatase (ALP) staining assay showed that POLISFI-K156N-expressing cells had the weakest ALP staining, suggestive of the lowest self-renewal capacity (Fig. 8B). Thus, we examined multiple well established lineage-specific markers of the three germ layers by quantitative RT-PCR, and we found that POLISFI-K156N-expressing cells differentiated mainly toward mesendoderm, although the POLISFI-WT- and POLISFI-K156R-expressing cells exhibited no change in the transcription of lineage-specific markers compared with non-transfection cells (Fig. 8, C and D).

As mentioned above, due to decreased protein stability and accelerated degradation, in POLISFI-K156N-expressing cells, the protein levels of overexpressed FLAG-Oct4/K156N were always lower than endogenous Oct4 in non-infected control cells or those of overexpressed FLAG-Oct4/WT counterpart (Fig. 3, C and F), despite that the ratio of Oct4 protein/Sox2 protein was kept at a constant level (Fig. 3F). To make a better comparison, we tried to bring the overexpressed FLAG-Oct4/WT protein level down to the similar level of overexpressed FLAG-Oct4/K156N by infecting cells with less “POLISFI shRNA + POLISFI-WT” lentiviruses (Fig. 3, H versus C). The lineage-specific transcription profiles were fairly similar between Fig. 3H cells (where the total (i.e. overexpressed + endogenous) WT protein level was similar to the total K156N level, Fig. 8D for the corresponding transcriptional profile) and Fig. 3C cells (where the total WT protein level was significantly higher than the total K156N level, Fig. 8C, for the corresponding transcriptional profile), suggesting that the Oct4/K156N mutation rather than the Oct4 protein level is the major determinant for the observed mesendodermal phenotypes.

Since over time a proportion of the POLISFI-K156N-expressing cells exhibited mesenchyme-like phenotype, which were detached from the culture substratum (Fig. 8A), we examined their survival and proliferation status in comparison with the POLISFI-WT-expressing cells. The expression of both POLISFI constructs was adjusted to a similar level by appropriate viral titration so that the results can be more comparable. Compared with the POLISFI-WT-expressing cells, the POLISFI-K156N-expressing cells exhibited slightly reduced survival (Fig. 8E) but significantly dampened proliferation (Fig. 8F). Considering the detached POLISFI-K156N-expressing cells are mostly viable and the detachment could result from the EMT, we next determined the transcription levels of a number of key genes involved in EMT. Although E-cadherin transcripts were not significantly down-regulated, most EMT-promoting genes were up-regulated to varying degrees, with SLUG being the most prominent one (Fig. 9, A and B). Western blotting confirmed the dramatically increased Slug and moderately decreased E-cadherin at the protein level (Fig. 9C). We then searched for the enhancer, promoter, and coding sequences of the SLUG gene and identified a potential canonical Oct4/Sox2-binding motif at its first exon (Fig. 9D), indicating that the transcription of this key EMT gene could be directly regulated by the Oct4/Sox2 complex. The binding of Oct4 protein to the SLUG probe was clearly validated by EMSA (Fig. 9E), but the dual binding of the Oct4/Sox2 appeared to largely depend on the order by which Oct4 or Sox2 was first incubated with the SLUG probe (Fig. 9F). It is interesting to note that the putative Oct4/Sox2 dual binding site at the SLUG gene (+195 to +213 relative to transcription start site) in PSCs differs from the single Sox2-binding site (−679 to −673 relative to transcription start site) reported in differentiated cancer cells (24). Thus, in cells with reduced pluripotency, Oct4 and Sox2 may be dissociated from each other and act independently on their target genes.

To test whether the known mesendodermal differentiation signals or stimuli could trigger the Oct4/Lys-156-regulated Oct4/Sox2 disassociation, we first treated the NCCIT cells with a combination of activin A, BMP4, FGF2, and PI3K inhibitor LY294002 (ABFLY), which is able to drive the differentiation of human ESCs into definitive endodermal cells (25). 16–24 h later, we saw dramatic morphological changes with drug-treated cells (Fig. 10A), where some of the endodermal and mesodermal markers (Fig. 10B) as well as some EMT markers (Fig. 10C) were up-regulated. Then we applied the same ABFLY treatment to NCCIT cells knocked down from the endogenous Oct4 while simultaneously expressing one of the FLAG-tagged POLISFI-Lys-156 variants (Fig. 10D), asking whether any of the Oct4/Lys-156 mutant proteins could functionally overlap with and mimic the ABFLY differentiation stimuli. Unfortunately, most examined lineage-specific genes did not show significant up-regulation even in cells ectopically expressing POLISFI-WT (Fig. 10E), suggesting either the pre-infection treatment or other unknown factors may have interfered with the ABFLY-induced differentiation processes. Nevertheless, a few markers (such as CK18, BRY, and BMP4) did show some increased transcription upon drug treatment. Interestingly, BRY and BMP4, two of the best studied mesoderm markers, were both up-regulated in cells expressing POLISFI-K156N to the same level as those expressing POLISFI-WT, and importantly, the ABFLY treatment did not further enhance their transcription (Fig. 10F).
Oct4/Lys-156 Controls Differentiation and EMT
10E). To better compare the levels of various transcripts in different groups, the ratio values of the transcript levels post-ABFLY treatment versus pre-ABFLY treatment were calculated and plotted in Fig. 10F. Thus, for BRY, the ratio of the POLISF1-WT group was significantly higher than that of the POLISF1-K156N group, which was close to 1 (Fig. 10F). A similar pattern was also seen with BMP4 as well as EOMES, a crucial gene for mesoderm delamination and endoderm specification (25), although the differences were not as significant as BRY. The above results indicated that, in inducing specific mesendoderm genes such as BRY, the effect of the ABFLY treatment could be mimicked by the Oct4/K156N mutant, consistent with the hypothesis that the ABFLY treatment may induce the Oct4/Sox2 disassociation by introducing specific PTMs onto Oct4/Lys-156 to alter the local structure of its side chain, functionally resembling the Oct4/K156N mutant.

Discussion

Although the Oct4-Sox2 complex is well established as a master control for cell pluripotency and differentiation, the corresponding interacting sites and regulatory mechanisms remain poorly defined. Structural modeling predicted Asp-166, Ile-158, Gln-155, and Lys-151 in Oct4 (numbered with human Oct4) (12, 14, 15, 17) and Met-102, Ala-99, Arg-98, and Lys-95 in Sox2 (numbered with human Sox2) (11, 12, 15, 17) as critical sites that may mediate the Oct4/Sox2 binding, but the physiological relevance of those sites and the biological consequences of their disruption are unclear. In this study, we identified a physiologically and pathologically relevant site in Oct4 (Lys-156) that critically controls the Oct4/Sox2 interaction. We revealed that Oct4/Lys-156 is important for maintaining the intermolecular salt bridge between Oct4/Lys-151 and Sox2/Asp-107, the only electrostatic interaction identified so far that contributes to the Oct4/Sox2 interaction. Also, Oct4/Lys-156 appears to play a critical role in maintaining the stability of the Oct4 protein possibly by forming a conserved intramolecular salt bridge with Oct4/Asp-212.

Thus far, neither the physiological stimuli/signals nor the molecular mechanisms that modulate Oct4 interactome and its various configurations are clear. We propose that the PTMs at key residues in the Oct4 protein may sense, convey, and integrate such signals (19). Indeed, we identified various PTMs at multiple sites on Oct4, including Lys-156. Notably, Lys-156 is differentially modified between Oct4 derived from PSCs and differentiated cells, possibly reflecting its different biophysicochemical and functional status in those cells. Methylation, acetylation, and ubiquitination at Lys-156 were detected in PSCs but not in differentiated cells, implicating active modifications at this site that may play regulatory roles specifically in PSCs. We envisioned that the differentiation signals may induce the acetylation/methylation of Lys-156, leading to changes in the local structure of the Lys-156 side chain and the disassociation of Oct4 from Sox2. The disassociated Oct4 protein could be deacetylated/demethylated and ubiquitinated at Lys-156 for degradation. Paradoxically, the K156N/K156A mutations accelerated rather than attenuated the degradation of the Oct4 protein, indicating the presence of additional lysine residue(s) that mediate the ubiquitination-dependent degradation of Oct4. Lys-154 that was also identified as an ubiquitination site in NCCIT cells (Fig. 1C) could be one of the candidates, and the K156N/K156A mutations could facilitate the docking of the enzymes required for Lys-154 polypubiquitination. In a recent study, Jin et al. (26) conducted a nonbiased alanine scanning analysis in which each Lys, Arg, Ser, Thr, or Tyr residue in mouse Oct4 protein was individually substituted with Ala, and the protein stability, subcellular localization, transactivation activity, and somatic cell reprogramming capability of each mouse Oct4 mutant were systematically examined. Among all the mutants, K147A, K149A, and R150A (corresponding to K154A, K156A, and R150A in human Oct4, respectively) completely abolished the reprogramming capability of Oct4. Interestingly, the reprogramming defects of these three mutants were accompanied with their dramatically reduced protein stability. Therefore, Lys-156 and Lys-154 may coordinately control the protein stability and functions of Oct4. Accordingly, Thr-159 might be a potential docking site for enzymes catalyzing Lys-154 polypubiquitination as the T159P mutation significantly abrogated the degradation of Oct4 protein (Fig. 6F). In addition, there appears to be a mutually exclusive PTM pattern between Lys-156 and Lys-151; Lys-156 is acetylated/methylated and ubiquitinated in NCCIT but not in U87 cells, although Lys-151 is acetylated/methylated in U87 but not in NCCIT cells (Fig. 1, C and D). Therefore, the PTM-based cross-talk between Lys-156 and its neighboring key residues (such as Lys-154, Lys-151, and Thr-159) in various cellular contexts need to be systematically investigated in future studies. Taken together, we revealed that the alterations of the local structure and the docking properties of Oct4/Lys-156 (either by PTMs or by amino acid substitution) are likely to be the key mechanism that underlies the crucial disassociation of the Oct4-Sox2 complex that leads to the mesendoderm lineage specification. It remains to be seen whether the same mechanism is also responsible for the Oct4/Sox2 disassembly and subsequent Oct4/Sox17 formation during the induction of primitive endoderm in an even earlier embryonic stage of development (27, 28).

One striking feature of coordination is the development between cell fate specification and the EMT-based cell migration. The major regulators of gastrulation, the nodal, BMP,
Wnt, and FGF signaling pathways, can regulate these two interconnected processes in a parallel or sequential manner (29). EMT is often coupled with mesendodermal differentiation (30–32). The first wave of primary EMT after implantation occurs when mesendodermal progenitors are formed from the developing epiblast, prior to the formation of the primitive...
Utilizing the embryoid bodies derived from human ESCs, Chan et al. (33) provided evidence that EMT may occur prior to mesendoderm differentiation. SNAIL and SLUG, the two well known EMT drivers responsible for down-regulating E-cadherin and ZO-1 (34), are implicated as target genes cooperatively suppressed by Oct4 and Sox2 during somatic cell reprogramming (35). Recently, Richter et al. (31) showed that SLUG is up-regulated in human ESCs exposed to BMP4 to initiate EMT and subsequent formation of mesoderm with cardiomyogenic potential. This study confirms this finding and further identifies a potential Oct4/Sox2 dual binding site at the SLUG coding region. It is conceivable that specific differentiation stimuli on the one hand can trigger the disassociation of the Oct4-Sox2 complex from the SLUG (and perhaps SNAIL) gene to launch EMT, and on the other hand can induce their disassociation from the master pluripotency factor genes (such
as NANOG) to allow for mesendodermal differentiation. Therefore, we propose here that during the early developmental stage, the EMT and mesendoderm specification process are likely to be coordinated through the Oct4/Lys-156-modulated disassociation of the Oct4-Sox2 complex that is triggered by specific differentiation signals. Furthermore, because in cells with reduced pluripotency, Oct4 and Sox2 appear to be disassociated from each other and act independently on their target genes (24, 36), it remains to be addressed when and how the Oct4/Sox2 disassociation is achieved and what the role of Oct4/Lys-156 mutation at Oct4/K156N somatic mutation present only in a subpopulation of bladder cancer stem cells or the bulk tumor cells? Can the Oct4/Lys-156 Controls Differentiation and EMT

Experimental Procedures

Cell Culture and Treatment—The sources and culture conditions for NCCIT, U87, H1, and 293T cells were described in detail previously (37). Cells were maintained in a 37 °C incubator with 5% CO2. Where appropriate, cells were treated with 20 μg/ml CHX, 5 μg/ml MG-132, or 200 μM chloroquine (with DMSO as the vehicle) either individually or in combination for the indicated durations. NCCIT cells were induced to differentiate into mesendoderm using DMEM (without FBS) supplemented with 100 ng/ml activin A, 10 ng/ml BMP4, 20 ng/ml FGF2, and 10 μM LY294002 (38).

Antibodies and Reagents—The polyclonal anti-Oct4 antibody (catalog no. sc-9081) was purchased from Santa Cruz Biotechnology. Anti-Sox2 antibody (catalog no. 2478), anti-Slug antibody (catalog no. sc-9081), and phosphatase inhibitors (PhosSTOP, Roche Applied Science catalog no. 04906845001). The whole cell lysates were incubated with 15 μL of anti-FLAG M2 beads (Sigma catalog no. M8823) overnight at 4 °C. Immunocomplexes were washed once with IP buffer before being resolved by SDS-PAGE and immunoblotted with the indicated antibodies.

Biotinylated DNA Pulldown Assay—50 μl (50 nm) of biotinylated NANOG probe was incubated with 180 μl (400 μg) NCCIT whole cell lysate in binding buffer (25 mM Tris, pH 7.2, 150 mM NaCl, 1% protease inhibitor mixtures and phosphatase inhibitor mixtures) overnight at 4 °C with gentle rotation. 20 μl of streptavidin-conjugated agarose beads (Pierce catalog no. 20347) pre-washed with binding buffer were added into the DNA-protein complex and further incubated for 1 h at room temperature with gentle rotation. The bead-DNA-protein complexes were sedimented by centrifugation at 1000 × g for 3 min, washed twice with ice-cold binding buffer, and eluted into SDS-PAGE sample loading buffer by heating at 100 °C. Samples were resolved by SDS-PAGE and immunoblotted with the indicated antibodies.

Recombinant Protein Expression and Purification—The His-tagged Oct4 and Sox2 (WT and their variants) constructs were transformed into BL21, and cultures were grown for 12 h at 37 °C in 5 ml of LB medium containing 50 μg/ml kanamycin. The cultures were then diluted into 100 ml of LB supplemented similarly and subsequently until the absorbance (A600) reached about 0.4–0.6, and isopropyl 1-thio-β-D-galactopyranoside (0.5 mM) was added to the culture, which was incubated further at 18 °C for 16 h to induce protein expression. Then the cells were harvested and resuspended in 15 ml of lysis buffer (25 mM Tris, 150 mM NaCl, 20 mM imidazole, 5% (v/v) glycerol, pH 7.4, supplemented with PMSF, lysozyme, and DNase I). After sonication, the lysate was centrifuged at 15,000 rpm for 20 min at 4 °C, and the supernatants were mixed with 1 ml of 50% (w/v) slurry of Ni-NTA beads and incubated at 4 °C on a rotary shaker for 1 h. The mixture was then centrifuged at 1000 rpm for 5 min, and the supernatant was discarded. The beads were then washed successively with lysis buffer containing 50 and 75 mM imidazole and finally eluted with the same buffer containing 500 mM imidazole.
300 mM imidazole. The proteins were aliquoted and stored at −80 °C for later use.

**Immunofluorescence Staining**—To examine subcellular localization of ectopically expressed FLAG-Oct4/WT and FLAG-Oct4/K156N, cells cultured on coverslips were washed twice with warm PBS and fixed with 4% formaldehyde (Sigma catalog no. F8775-25ML) for 20 min at room temperature, washed again three times with PBS, and permeabilized with 0.1% Triton X-100 at room temperature for 10 min. They were then washed three times with PBS and blocked with 3% BSA (Sigma catalog no. A9418) in PBS for 1 h. Thereafter, cells were incubated with specified primary and secondary antibodies for 1 h, respectively, with three washes in between, and counterstained with DAPI.

**Electrophoretic Mobility Shift Assay**—EMSAs were performed with a LightShift chemiluminescent EMSA kit (Thermo, catalog no. 20148) as instructed by the manufacturer. Briefly, the reaction mixture (20 μl) containing 1× binding buffer, 5′-biotin end-labeled dsDNA (50 fmol), and the indicated concentrations of purified recombinant protein were incubated for 30 min at room temperature, and then the reaction was terminated by the addition of 5 μl of gel loading buffer. Samples were separated by 6% native polyacrylamide gel in 0.5× TBE (precooled to −4 °C) and transferred to PVDF membrane in a cold environment. After cross-linking DNA using a UV-light cross-link (120 ml/cm²), the membrane was blocked, washed, and developed using the Chemiluminescent Nucleic Acid Detection Module according to the manufacturer’s instructions.

**Alkaline Phosphatase Staining**—NCCIT cells were plated into 6-well plates and infected with lentiviruses carrying Oct4/WT or its variants. After 3 days, cells were fixed in the citrate/acetate/formaldehyde solution (18 mM citric acid, 9 mM sodium citrate, 12 mM NaCl, pH 3.6, 47% acetone, and 8% formaldehyde) for 30–50 s. ALP staining was carried out with the alkaline phosphatase detection kit (Sigma catalog no. 86R-1KT). Cell colonies were classified into three categories as follows: “undifferentiated” colonies containing only ALP-positive cells; “mixed” colonies containing both ALP-stained and unstained cells; and “differentiated” colonies containing no ALP-stained cells. 150–200 colonies in each group were scored, and the percentages of cells of each category were compared.

**Cell Viability Assay and Proliferation Assay**—NCCIT cells were plated into 12-well plates and infected with different titers of lentiviruses carrying POLISF1-WT or POLISF1-K156N to bring the overexpressed Oct4 variants to approximately the same level. After 3 days, cells were stained with propidium iodide and counted by flow cytometry, and the cell survival rate was determined as the ratio of propidium iodide-negative cells to total cells. To determine the proliferation curve of these cells, NCCIT cells were plated in 6-well plates and infected in the same manner. The total cell numbers were counted using a hemocytometer 1, 3, and 5 days, respectively, after infection.

**Quantitative RT-PCR**—Quantitative real time PCR analysis was performed as described previously (21). Briefly, total mRNA was extracted using RNA plus reagent (TaKaRa catalog no. 9109), and cDNA was synthesized using PrimeScript RT reagent kit with gDNA eraser (TaKaRa catalog no. RR047A) according to the manufacturer’s instructions. The primers used for lineage marker quantitation were described previously (39), and other primers were listed as follows: POU5F1 F, GTGGA-GGAAGCTGAAACA, and R, ATTCCTCCAGGTGGCTCTCA; SOX2 F, CATGGGTCTGGTGTCAAGTC, and R, TCCGGCGCCGGGAGATACA; NANOG F, CGGTTGTTCTGAAGA, and R, TTGGGACGTGTTGAAGAA; FBXO15 F, TGGCTGTGACAGCTATTCCGG, and R, GATAGTAGCCGACCTAATGTGC; UTF1 F, CGGTCC-CACCGGCAGCG, and R, AGGCCGTCCGAGCTTCGTG; EOMES F, ATCTATCCAAAGACCGGAGG, and R, CCGGGTTGTTATTTGTTGAAG; and GAPDH F, GGGGAGCCAAAAAGGTTCATCATCT, and R, GAGGGGCACCATCAGTCTTCT.

qRT-PCR was performed in a MJ Chromo 4 (Bio-Rad DNA engine) by using a reaction mixture with SsoFastTM EvaGreen SuperMix (Bio-Rad catalog no. 172-5202). All the PCR amplifications were performed in triplicate and repeated in three independent experiments. The relative quantities of selected mRNAs were normalized to that of GAPDH.

**In Vitro PTM of Recombinant Oct4 Proteins**—An aliquot (5 μg) of purified recombinant His-Oct4 protein binding to the Ni-NTA beads (100 μl) was incubated with 1 mg of NCCIT or U87 whole cell lysate at 30 °C for 1 h in 1 ml of PMA buffer (50 mM Tris-HCl, pH 7.4, 50 mM KCl, 5 mM MgCl₂, 0.5% Nonidet P-40, 5 mM ATP, 100 μM S-adenosylmethionine, 100 μM acetyl-CoA, 1% EDTA-free protease inhibitor mixtures and phosphatase inhibitor mixtures) (19). The His-Oct4 conjugated beads were sedimented by centrifugation at 800×g for 2 min, washed three times with ice-cold PMA washing buffer (50 mM Tris-HCl, pH 7.4, 50 mM KCl, 5 mM MgCl₂, 0.5% Nonidet P-40), eluted, and denatured in SDS-PAGE sample loading buffer by heating at 100 °C for 5 min. Pooled samples (10–20 μg) were loaded and separated by SDS-PAGE and stained with Coomasie Brilliant Blue R-250. After destaining, the Oct4 bands (with a molecular mass of 45 kDa) were excised and analyzed by mass spectrometry.

**Mass Spectrometric Identification of PTMs in Oct4 Proteins**—Oct4 samples were subjected to overnight digestion with trypsin or chymotrypsin as described by Liu et al. (40). The peptides were extracted with acetonitrile containing 0.1% formic acid and vacuum-dried.

Proteolytic peptides were reconstituted with mobile phase A (2% acetonitrile containing 0.1% formic acid) and then separated on an on-line C18 column (75 μm inner diameter, 360 μm outer diameter, 10 cm, 3 μm of C18). Mobile phase B consisted of 0.1% formic acid in 2% acetonitrile, and mobile phase B was 0.1% formic acid in 84% acetonitrile. A linear gradient from 3 to 100% B over 75 min at a flow rate of 350 nl/min was applied. Mass spectrometry data were collected on an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, San Jose, CA) operated in data-dependent scan mode. Survey scan (m/z 375–1300) was performed at a resolution of 60,000 followed by MS2 scans to fragment the 50 most abundant precursors with collision-induced dissociation. The activation time was set at 30 ms; the isolation width was 1.5 atomic mass units; the normalized activation energy was 35%, and the activation q
Oct4/Lys-156 Controls Differentiation and EMT

was 0.25. Mass spectrometry raw file was searched by Proteome Discovery version 1.3 using Mascot search engine with percolator against the human ref-sequence protein database (updated on 07/04/2013). The mass tolerance was set to be 20 ppm for precursor and 0.5 Da for product ion. Missed cleavages were no more than five for each peptide. Phosphorylation of Ser/Thr and Tyr, acetylation of Lys, monomethylation, dimethylation, and trimethylation of Lys and Arg, and diglycine modification of Lys (ubiquitination) were used as variable modifications. A filter of 90% peptide confidence was applied according to the Peptide Prophet and Protein Prophet parsimony algorithms. Fragment assignment of each modified peptide was subject to manual inspection and validation using the original tandem mass spectra acquired in profile mode using Xcalibur software.

All-atom MD Simulations—Initial coordinates of the wild type (Oct4/WT) and mutated (Oct4/K156N) Oct4-Sox2-DNA ternary complexes were built from the NMR structure of the Oct1-Sox2-HOXB1 complex (PDB code 1O4X) (15), and the missed linker region of Oct4 was modeled base on the structure of the Oct4-PORE complex (PDB code 3L1P) (16). For comparison purposes, the ternary system (Oct4/WT(FGF4), Oct4/K156N(FGF4)) were also built from the crystal structure of the Oct1/Sox2-FGF4 complex (PDB code 1GTO) (14). All MD simulations were carried out using the Gromacs 4.5.3 package (43) with the CHARMM27 force field (44, 45). Each ternary complex was solvated in a cubic box of TIP3P water molecules (41) and 100 mM NaCl in addition to the neutralizing ions, leading to a total of ~90,700 atoms for each system. Periodic boundary conditions were applied, with the particle mesh Ewald method applied to calculate the long range electrostatics. SETTLE and LINCS were applied on the hydrogen-involved covalent bonds in water molecules and in other molecules, respectively. The solvated complex system was first subjected to four rounds of steepest descent minimization, with restraint applied on different atoms of the system as follows: the first round on solute heavy atoms, the second on DNA heavy atoms and the protein main chain, the third round on DNA heavy atoms and protein Ca atoms; and the fourth round without any restraint. Minimization was converged when the maximum force of the four rounds is smaller than 100, 50, 10, 10 (kJ mol⁻¹ nm⁻¹), respectively. In all these steps, the modeled linker region was excluded from any restraint. Minimization procedure was followed by four equilibration steps. The first step is a 50-ps MD simulation with a restraint on solute heavy atoms, and the temperature was slowly increased from 50 to 310 K at constant volume, with a time step of 1 fs. Equilibration continued with another 50-ps MD simulation in the NTP ensemble at constant temperature (310 K) and constant pressure (1 atm) with a time step of 1 fs, and restraint was applied on DNA heavy atoms and protein main chain atoms. In the third and fourth steps, 50- and 100-ps MD simulations were carried out with a time step of 1 and 2 fs, respectively, and in both steps restraint was applied on DNA heavy atoms and protein Ca atoms. In all these steps, the modeled linker region was excluded from any restraint. Equilibration was followed by an unrestrained production MD simulation of up to 500 ns at constant temperature (310 K) and constant pressure (1 atm) with a time step of 2 fs, and the coordinates were saved every 10 ps. Trajectories were viewed with VMD (42).

Statistical Analyses—All quantitative data were presented as mean ± S.E. of three independent experiments. The statistical significance of compared measurements was evaluated with one-way ANOVA and LSD test using SPSS 19.0, and the difference was considered significant and increasing when p < 0.05 (*), p <0.01(**), or p <0.001(***).

Author Contributions—X. P. conducted most biochemical and cellular experiments with critical assistance from S. D., J. L., J. C., and B. K. X. C. performed all molecular dynamics simulations. B. K. and B. S. participated in data analysis and interpretation. X. D. conducted mass spectrometry and data analysis. Y.-J. W. conceived the study and designed and oversaw most biochemical and cellular experiments. The manuscript was written by Y.-J. W. with input from X. P., X. C., and B. S.

Acknowledgments—We are grateful to Drs. Xin-Hua Feng and Jian-Zhong Shao for providing reagents and to Dr. Min-Xin Guan for general support to X. C. We also thank Yuelin Song, Lu Gong, Qi Chen, and Shelong Zhang for technical assistance. The MD simulation was carried out at the National Supercomputer Center in Tianjin, and the calculations were performed on TianHe-1A.

References
1. Yu, J., and Thomson, J. A. (2008) Pluripotent stem cell lines. Genes Dev. 22, 1987–1997.
2. Iovino, N., and Cavalli, G. (2011) Rolling ES cells down the Waddington landscape with Oct4 and Sox2. Cell 145, 815–817.
3. He, S., Nakada, D., and Morrison, S. J. (2009) Mechanisms of stem cell self-renewal. Annu. Rev. Cell Dev. Biol. 25, 377–406.
4. Wang, Z., Oron, E., Nelson, B., Razis, S., and Ivanova, N. (2012) Distinct lineage specification roles for NANOG, OCT4, and SOX2 in hESCs. Cell Stem Cell 10, 440–454.
5. Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenthner, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., and Young, R. A. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122, 947–956.
6. Rizzino, A. (2013) Review—the Sox2-Oct4 connection—critical players in a much larger interdependent network integrated at multiple levels. Stem Cells 31, 1033–1039.
7. Jerabek, S., Merino, F., Schöler, H. R., and Cojocaru, V. (2014) OCT4: dynamic DNA binding pioneers stem cell pluripotency. Biochim. Biophys. Acta 1839, 138–154.
8. Thomson, M., Liu, S. J., Zou, L. N., Smith, Z., Meissner, A., and Ramathan, S. (2011) Pluripotency factors in ESCs regulate differentiation into germ layers. Cell 145, 875–889.
9. Sui, L., Bouwens, L., and Mfopou, J. K. (2013) Signaling pathways during maintenance and DE differentiation of ESCs. Int. J. Dev. Biol. 57, 1–12.
10. Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W., Jiang, J., Loh, Y. H., Yeo, H. C., Yeo, Z. X., Narang, V., Govindarajan, K. R., et al. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133, 1106–1117.
11. Merino, F., Bouvier, B., and Cojocaru, V. (2015) Cooperative DNA recognition modulated by an interplay between protein-protein interactions and DNA-mediated allostery. PLoS Comput. Biol. 11, e1004287.
12. Chen, J., Zhang, Z., Li, L., Chen, B. C., Revyakin, A., Haji, B., Legant, W., Dahan, M., Lionnet, T., Betzig, E., Tjian, R., and Liu, Z. (2014) Single-molecule dynamics of enhanceosome assembly in ESCs–(Oct4&Sox2). Cell 156, 1274–1285.
13. Lam, C. S., Mistri, T. K., Foo, Y. H., Sudhaharan, T., Gan, H. T., Rodda, D.,
Lim, L. H., Chou, C., Robson, P., Wohland, T., and Ahmed, S. (2012) DNA-dependent Oct4/Sox2 interaction and diffusion properties characteristic of the pluripotent cell state revealed by fluorescence spectroscopy. *Biochem. J.* **448**, 21–33

Reményi, A., Lins, K., Nissen, L. J., Reinbold, R., Schüler, H. R., and Wilmanns, M. (2003) Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev.* **17**, 2048–2059

Williams, D. C., Jr., Cai, M., and Clore, G. M. (2004) Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1-Sox2-Hoxb1-DNA ternary transcription factor complex. *J. Biol. Chem.* **279**, 1449–1457

Esch, D., Vahokoski, J., Groves, M. R., Poggenburg, V., Coccojara, V., Vom Bruch, H., Han, D., Drewler, H. C., Araújo-Bravo, M. J., Ng, C. K., Jauch, R., Wilmanns, M., and Schüler, H. R. (2013) A unique Oct4 interface is crucial for reprogramming to pluripotency. *Nat. Cell Biol.* **15**, 295–301

Merino, F., Ng, C. K., Veerapandian, V., Schöler, H. R., and Jauch, R. (2007) The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Curr. Opin. Genet. Dev.* **17**, 1274–1286

Hunter, T. (2007) The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Nat. Rev. Mol. Cell Biol.* **8**, 779–793

Dan, S., Kang, B., Duan, X., and Wang, Y. J. (2015) A cell-free system toward deciphering the post-translational modification barcodes of Oct4 in different cellular contexts. *Biochem Biophys. Res. Commun.* **456**, 714–720

Izarzugaza, J. M., Redfern, O. C., Orengo, C. A., and Valencia, A. (2009) Cancer-associated mutations are preferentially distributed in protein kinase functional sites. *Proteins* **77**, 892–903

Lin, Y., Yang, Y., Li, W., Chen, Q., Li, J., Pan, X., Zhou, L., Liu, C., Chen, C., He, J., Cao, H., Yao, H., Zheng, L., Xu, X., Xia, Z., et al. (2012) Reciprocal regulation of Akt and Oct4 promotes the self-renewal and survival of embryonal carcinoma cells. *Mol. Cell* **48**, 627–640

Yu, H., Wang, W., Li, C., Yu, H., Yang, A., Wang, B., and Jin, Y. (2009) WWP2 promotes degradation of transcription factor OCT4 in human embryonic stem cells. *Cell Res.* **19**, 561–571

Cecconi, F., and Levine, B. (2008) The role of autophagy in mammalian development: cell makeover rather than cell death. *Dev. Cell* **15**, 344–357

Sun, C., Sun, L., Li, Y., Kang, X., Zhang, S., and Liu, Y. (2013) Sox2 expression predicts poor survival of hepatocellular carcinoma patients and it promotes liver cancer cell invasion by activating Slugh. *Mol. Oncol.* **30**, 503

Teo, A. K., Arnold, S. I., Trotter, M. W., Brown, S., Ang, L. T., Chng, Z., Robertson, E. J., Dunn, N. R., and Vallier, L. (2011) Pluripotency factors regulate definitive endoderm specification through esomesdermin. *Genes Dev.* **25**, 238–250

Jin, W., Wang, L., Zhu, F., Tan, W., Lin, W., Chen, D., Sun, Q., and Xia, Z. (2016) Critical POU domain residues confer Oct4 uniqueness in somatic cell reprogramming. *Sci. Rep.* **6**, 20818

Aksoy, I., Jauch, R., Chen, J., Dyla, M., Divakar, U., Bogu, G. K., Teo, R., Leng Ng, C. K., Herath, W., Lili, S., Hutchins, A. P., Robson, P., Kolakar, P. R., and Stanton, L. W. (2013) Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. *EMBO J.* **32**, 938–953

Frum, T., Halbisen, M. A., Wang, C., Amir, H., Robson, P., and Ralston, A. (2013) Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Dev. Cell* **25**, 610–622

Heisenberg, C. P., and Solnica-Krezel, L. (2008) Back and forth between cell fate specification and movement during vertebrate gastrulation. *Curr. Opin. Genet. Dev.* **18**, 311–316

Slawny, N., and O’Shea, K. S. (2013) Geminin promotes an epithelial-to-mesenchymal transition in an embryonic stem cell model of gastrulation. *Stem Cells Dev.* **22**, 1177–1189

Richter, A., Valdimarsdottir, L., Hrafnkelsdottir, H. E., Runarsson, I. F., Omarsdottir, A. R., Ward-van Oostwaard, D., Mummery, C., and Valdimarsdottir, G. (2014) BMP4 promotes EMT and mesodermal commitment in human embryonic stem cells via SLUG and MSX2. *Stem Cells* **32**, 636–648

Eastham, A. M., Spencer, H., Soncin, F., Ritson, S., Merry, C. L., Stern, P. L., and Ward, C. M. (2007) Epithelial-mesenchymal transition events during human embryonic stem cell differentiation. *Cancer Res.* **67**, 11254–11262

Chan, D. N., Azghadi, S. F., Feng, J., and Lowry, W. E. (2012) PTK7 marks the first human developmental EMT in vitro. *PLoS ONE* **7**, e50432

Peinado, H., Olmeda, D., and Cano, A. (2007) Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat. Rev. Cancer* **7**, 415–428

Shu, X., and Pei, D. (2014) The function and regulation of mesenchymal-to-epithelial transition in somatic cell reprogramming. *Curr. Opin. Genet. Dev.* **26**, 32–37

Herreros-Villanueva, M., Zhang, J. S., Koening, A., Abel, E. V., Smyrk, T. C., Bamlet, W. R., de Narvajas, A. A., Gomez, T. S., Simeone, D. M., Bujanda, L., and Billadeau, D. D. (2013) SOX2 promotes de differentiation and impacts stem cell-like features to pancreatic cancer cells. *Oncogenesis* **2**, e61

Cheng, J., Li, W., Kang, B., Zhou, Y., Song, J., Dan, S., Yang, Y., Zhang, X., Li, J., Yin, S., Cao, H., Yao, H., Zhu, C., Yi, W., Zhao, Q., et al. (2015) Tryptophan derivatives regulate the transcription of Oct4 in stem-like cancer cells. *Nat. Commun.* **6**, 7209

Teboul, T., Hannan, N. R., Corbineau, S., Martinez, A., Martinet, C., Branchereau, S., Mainot, S., Strick-Marchand, H., Pedersen, R., Di Santo, J., Weber, A., and Vallier, L. (2010) Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. *Hepatology* **51**, 1754–1765

You, J. S., Kang, I. K., Seo, D. W., Park, J. H., Park, J. W., Lee, I. C., Jeon, Y. J., Cho, E. J., and Han, J. W. (2009) Depletion of embryonic stem cell signature by histone deacetylase inhibitor in NCCIT cells: involvement of Nanog suppression. *Cancer Res.* **69**, 5716–5725

Liu, Q., Ding, C., Liu, W., Song, I., Liu, M., Qi, L., Fu, T., Malovannaya, A., Wang, Y., Qin, J., and Zhen, B. (2013) In-depth proteomic characterization of endogenous nuclear receptors in mouse liver. *Mol. Cell Proteomics* **12**, 473–484

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulati ng liquid water. *J. Chem. Phys.* **79**, 926–935

Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38

Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616

Lindahl, E., Bjellmar, P., Larsson, P., Cuenod, M. A., and Hess, B. (2010) Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models. *J. Chem. Theory Comput.* **6**, 459–466