

POST-TRANSLATIONAL REGULATION OF OCT1/POU2F1
IN THE STRESS RESPONSE AND MITOSIS

by
Jinsuk Kang

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Experimental Pathology

Department of Pathology
The University of Utah
August 2011

Copyright © Jinsuk Kang 2011

All Rights Reserved

ABSTRACT

The Oct1/POU2F1 transcription factor was previously thought to constitutively occupy its cognate DNA binding sites, and to regulate the expression of housekeeping genes. This stereotype led to little attention being paid to Oct1 activity in dynamic cellular responses. In 2005, *Oct1*^{-/-} mouse embryonic fibroblasts were shown to be highly sensitive to oxidative and genotoxic stresses, implicating Oct1 as a stress sensor. However, the mechanism connecting stress with Oct1 transcription regulation was unknown. To identify the mechanism by which Oct1 activity is regulated by post-translational modifications in response to stress exposure, I used affinity purification and mass spectrometry to map specific Oct1 phosphorylation, O-GlcNAc modification and ubiquitination sites.

Serine 385 (Chapter 2): I identified unique mechanisms of Oct1 regulation in response to stress and during mitosis, involving two different Oct1 phosphorylation events. Following genotoxic and oxidative stress, Oct1 binding specificity is mediated by phosphorylation of S385, switching from a monomeric into a dimeric conformation on different binding sequences. I confirmed this mechanism by genome wide ChIPseq. The homologous protein Oct4, a master regulator of embryonic stem cells, uses a similar mode of regulation.

Serine 335 (Chapter 3): I identified an other phosphorylation event (pS335) as a negative regulator of Oct1 binding to DNA. I found that this phosphorylation is induced

in mitotic cells as well as in stressed ones. The phosphorylated Oct1 is enriched in the mitotic spindle poles and midbody. Phospho-Oct1 is also K11-ubiquitinated. Using several different approaches, I showed that Oct1 directly regulates mitosis after being displaced from mitotic chromatin.

Threonine 255 and Serine 728 (Chapter 4): In addition to phosphorylation and ubiquitination, I also identified O-GlcNAc modification of Oct1. Using an Oct1 glycosylation defective mutant, I found that the glycosylated residues of Oct1 regulate its DNA binding and transcriptional activity.

Finally, I confirmed in embryonic stem cells that Oct1 and Oct4 share binding specificity for novel multimeric as well as conventional octamer motifs (**Chapter 5**). Further, multimeric binding motifs recruit Oct1 and Oct4 hetero-complexes suggesting extensive crosstalk between Oct1 and Oct4 in embryonic stem cells.

TABLE OF CONTENTS

ABSTRACT.....	iii
LIST OF TABLES.....	viii
LIST OF FIGURES.....	ix
ACKNOWLEDGMENTS.....	xii
Chapters	
1 INTRODUCTION.....	1
Octamer Transcription Factors.....	1
Cellular Functions of <i>Oct1/POU2F1</i> and <i>Oct4/POU5F1</i>	3
Oct1 Interacting Partners.....	6
DNA Bound Oct Factor Structures.....	9
Transcription Factors in Mitosis.....	15
O-GlcNAc Modification of Transcription Factors.....	16
Plan of the Dissertation.....	19
References.....	20
2 A GENERAL MECHANISM FOR TRANSCRIPTION REGULATION BY OCT1 AND OCT4 IN RESPONSE TO GENOTOXIC AND OXIDATIVE STRESS.....	26
Abstract.....	27
Introduction.....	27
Results.....	28
Discussion.....	36
Materials and Methods.....	38
Acknowledgements.....	39
References.....	39

3	DYNAMIC REGULATION OF OCT1 DURING MITOSIS BY PHOSPHORYLATION AND UBIQUITINATION.....	42
	Abstract.....	42
	Introduction.....	43
	Results.....	45
	Discussion.....	87
	Methods.....	93
	Acknowledgements.....	97
	Author Contributions.....	97
	References.....	98
4	DYNAMIC REGULATION OF OCT1 BY GLUCOSE AND O-GlcNAc MODIFICATION	102
	Abstract.....	102
	Introduction.....	103
	Results.....	104
	Discussion.....	127
	Materials and Methods.....	129
	Acknowledgements.....	131
	References.....	131
5	COMBINATORIAL BINDING OF TRANSCRIPTION FACTORS IN THE PLURIPOTENCY CONTROL REGIONS OF THE GENOME.....	134
	Abstract.....	135
	Introduction.....	135
	Results.....	136
	Discussion.....	141
	Methods.....	142
	Acknowledgements.....	143
	References.....	144
6	CONCLUSION.....	145
	Oct1 as a Stress Sensor.....	148
	Oct1 as a Mitotic Regulator.....	149
	Oct1 as a Potential Glucose Sensor	150
	Oct1 and Oct4 Interaction Network.....	151
	Future Directions: Oct1 and Somatic Stem Cells.....	152
	References.....	153

Appendices

A MAPPING OF CDK7 DEPENDENT OCT1 PHOSPHO-RESIDUE(S).....155

B GENERATION OF iPSCS USING OCT1 AND OCT4 CHIMERAS.....160

C STEM CELLS, STRESS, METABOLISM AND CANCER: A DRAMA
IN TWO OCTS.....166

LIST OF TABLES

Table	Page
1.1 Previous known Oct1 interacting factors.....	7
2.1 Oct1 phosphopeptides identified following exposure of HeLa cells to IR.....	29
6.1 Previous known and newly identified Oct1 interacting factors.....	147
C.1 A subset Oct1 and Oct4 target genes.....	170
C.2 A partial list of Oct1 and Oct4 interaction partners.....	173

LIST OF FIGURES

Figure	Page
1.1. POU domain transcription factors.....	2
1.2. Oct and Non-Oct factors and their expression patterns.....	4
1.3. Structures of the Oct1 DNA binding domain bound to octamer motif.....	11
1.4. Structures of the Oct1 DNA binding domain bound to MORE and PORE motifs....	13
1.5. O-GlcNAc modification.....	17
2.1. Genotoxic and oxidative stress induce Oct1 phosphorylation and dimerization at complex sites.....	28
2.2. An endogenous 2XMORE interacts with four Oct1 molecules in a cooperative stress-induced complex.....	31
2.3. Oct1 functionally regulates <i>Polr2a</i>	32
2.4. Oct1 S385 and S335 mutations modulate DNA selectivity.....	33
2.5. Oct4 complex site binding is IR-inducible.....	35
2.6. ChIPseq identifies inducible Oct1 targets.....	36
3.1. Oct1 ^{pS335} is enriched in M-phase HeLa cells.....	46
3.2. Further characterization of the phospho-specific antibody.....	48
3.3. Mitotic Oct1 ^{pS335} is associated with the spindle pole bodies and midbody.....	50
3.4. Localization of total Oct1 in mitotic HeLa cells.....	52
3.5. Localization of Oct1 phosphorylated at S335 to mitotic A549 lung Adeno- carcinoma cells.....	53

3.6. Comparison of the properties of anti-Oct1 ^{p335} and anti-histone H3 ^{S10} antibodies.....	55
3.7. Modulation of Oct1 levels results in abnormal mitoses in HeLa cells.....	58
3.8. Additional examples of mitotic abnormalities in HeLa cells treated with Oct1-specific siRNAs.....	61
3.9. Phospho-S335 IF staining pattern and cell-cycle phenotype of primary MEFs.....	63
3.10. Effect of WT and S335A Oct1 overexpression on Oct1 ^{pS335} in HeLa cells.....	66
3.11. Ectopic Oct1 expression in HeLa cells increases alpha-tubulin levels, and induces formation of puncta containing alpha-tubulin but lacking DNA.....	68
3.12. Nek6 contributes to mitotic Oct1 phosphorylation at S335.....	69
3.13. Further validation of NEK6.....	72
3.14. Oct1 is present in the spindle matrix and forms a complex with lamin B1 at the midbody in HeLa cells.....	75
3.15. Abnormal localization of Oct1 ^{pS335} by lamin B knockdown in early mitotic cell.....	81
3.16. Dynamic ubiquitination of Oct1 during mitosis.....	83
3.17. APC is responsible for the accumulation of K11-Ub signal at sites of Oct1 phosphorylation upon proteasome inhibition.....	88
3.18. Model for Oct1 localization and modification through the cell cycle.....	91
4.1. The nuclear form of Oct1 is predominantly modified by glycosylation.....	105
4.2. Identification of Oct1 O-linked glycosylated.....	110
4.3. Oct1 supports anchorage-independent growth in p53 deficient MEFs.....	113
4.4. Electrophoretic mobility shift assay comparing WT and DM in binding capacity.....	116
4.5. Glycosylated residues control Oct1 transcription activity.....	119
4.6. Transcription level of Oct1 control target genes.....	121

4.7. Regulation of Oct1 activity by glucose.....	123
4.8. Glycosylated Oct1 residues are critical for Oct1-mediated control of mitochondrial reducing activity but dispensable for the ability of Oct1 to confer sensitivity to long-term glucose withdrawal.....	126
5.1. Resynthesizing 316 pluripotency control regions (PluCRs).....	136
5.2. The map transcriptional complexes on DPPA4, a known target of Oct4.....	137
5.3. Identifying the specificity and clustering tendency of transcription factors.....	138
5.4. Defining the relative specificity for PluCR oligonucleotides for Oct1 and Oct4....	139
5.5. Defining a competition between FoxO1 and Oct4 for DNA binding at the RAB5A and ERMN loci.....	140
5.6. Oct1/Oct4 binding is influence by proximal Sox2 and Nanog binding.....	140
5.7. In vivo validation of predicted Oct1 binding.....	141
6.1. The whole post-translational modifications identified by mass spectrometry in the dissertation works.....	146
A.1. Mapping Oct1 residues phosphorylated by CAK.....	156
A.2. Mutant substrate sequences.....	158
B.1. Cloning Scheme.....	161
B.2. The examples of iPSC like colony.....	163
B.3. Research in progress.....	164
C.1. Dendrograms illustrating the degree of relatedness between different human POU transcription factors.....	168
C.2. POU domain transcription factor expression patterns and inter-relationships.....	169
C.3. Regulation of Oct1 and Oct4.....	172

ACKNOWLEDGMENTS

I would like to thank my academic advisor, Dr. Dean Tantin. His guidance has focused my broad interests toward a deeper insight of gene transcription. I really enjoy scientific discussions with Dr. Tantin: He always encourages me to think logically and corrects my naïve ideas to find better directions. Most of all, he never hesitates to share his time to discuss with me even when he is very busy. In my future career, I will try to follow his ways of thinking and scientific passion. I would like to thank my committee members: Dr. Bradley Cairns, Dr. Tim Formosa, Dr. David Stillman, and Dr. John Weis. Their advice has kept me on the right track during my Ph.D. studies.

I would like to thank my colleagues in the laboratory: Dr. Arvind Shakya, Jessica Maddox, Jared Anderson and other past lab members. We have shared good scientific discussions and enjoyable events.

I would also like to thank all members of Fairbrother laboratory at Brown University. I was able to finish two important chapters (2 and 5) due to their collaborations. Ben Goodman and Dr. Yixian Zheng at the Carnegie Institute provided strong evidence for Oct1 functions in spindle matrix in Chapter 3. Dr. Jaemin Lim and Dr. Lance Wells at the University of Georgia identified the sites of Oct1 O-GlcNAc modifications in Chapter 4. I appreciate their collaborations.

Other collaborators that I would like to thank are Dr. Hiroshi Handa in Tokyo Institute of Technology who provided DNA-linked beads for Oct1 purification, Dr.

Vishiva Dixit from Genentech provided K11-ubiquitin antibody, Dr. Frank G. Whitby from Biochemistry department in the University of Utah, provided expertise in molecular structure, and Dr. Janet Shaw kindly shared lab instruments. I would like to thank them for sharing these valuable resources.

Many Korean friends in the Medical School gave me great support. Whenever I was depressed, they were willing to give me good advice and refreshing parties. I would like to thank them all.

Finally, I would like to thank my family. In Korea, my parents and parents-in-law are sincere supporters. Their cheering encourages me to focus on my studies. My wife and daughter take great care of me by my side. I could not finish my Ph.D. course-work without their dedicated support. Although living in Salt Lake City was not always happy and sometimes difficult because of the economy and culture shock, they were willing to endure those problems.

I love you, all with all my heart.....

Salt Lake City, Utah

May 23, 201

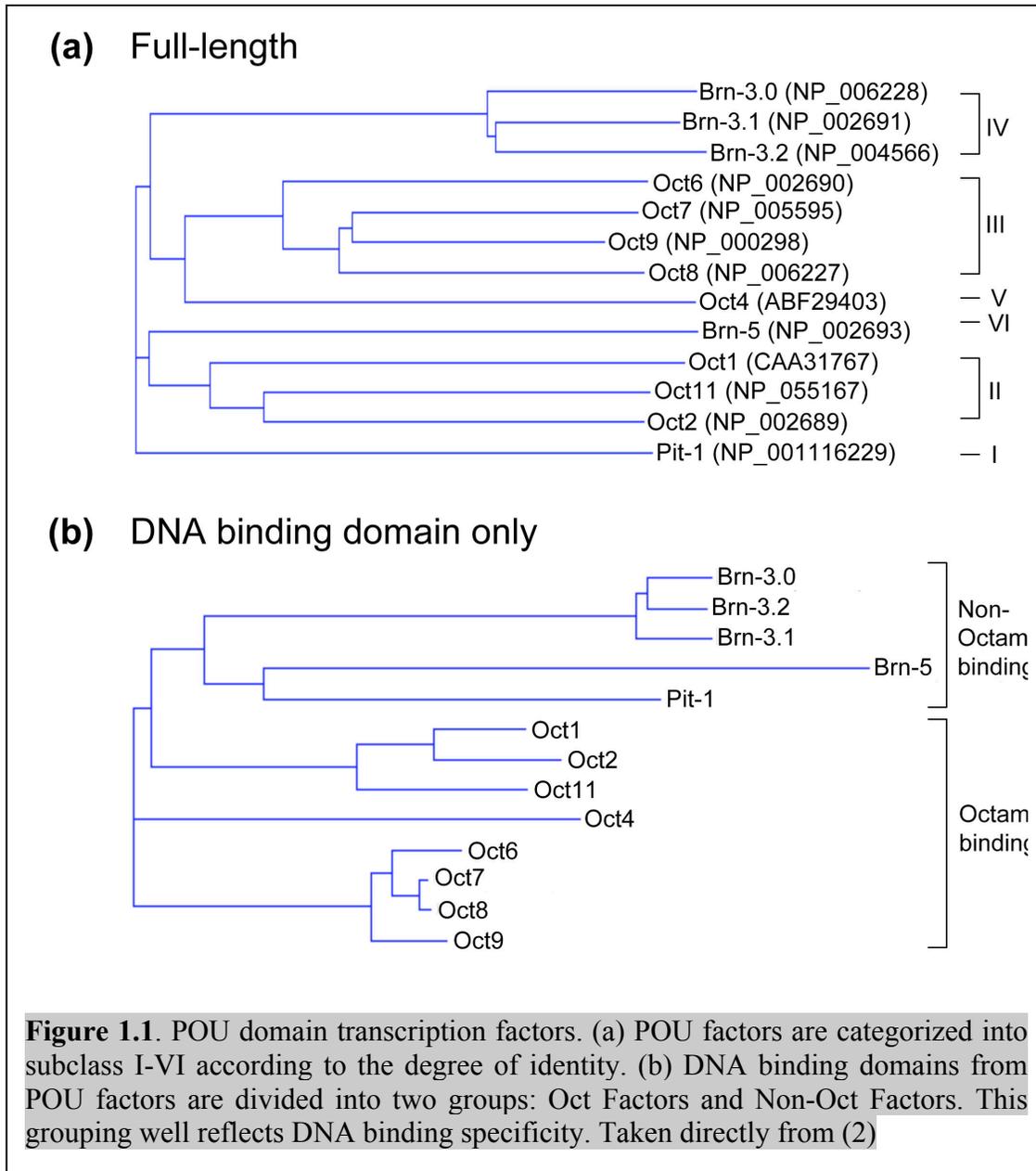
CHAPTER 1

INTRODUCTION

Octamer Transcription Factors

The POU (*Pit1*, *Oct1/2* and *Unc86*) domain transcription factors were first classified in 1988, based on the fact that the four factors share a well-conserved bipartite DNA binding domain: POU specific (POU_S) and POU homeo (POU_H) (3). These subdomains are associated with a linker domain known to be variable in length and amino acid composition. Ensuing studies extensively identified POU family proteins through sequence similarity, as well as expression patterns (4-7). These findings allowed the classification of POU domain factors into subclass I-VI based on sequence identity of POU domains and variable linkers ((8), Figure 1.1a).

Interestingly, POU_S and POU_H subdomains recognize distinct DNA sequences, ‘half site’ (9). The canonical “octamer” binding site consists of two half sites adjacent to each other. The linker domain allows them to bind cooperatively to the combined half sites. The POU domain factors also follow unique classification due to their original identification (10). The factors that bind preferentially to the “octamer” sequence (ATGCAAAT) are further classified as octamer transcription factors (“Oct factors”, Figure 1.1b). *POU1F1* (*Pit1*), *POU4F1* (*Brn3.0*) and *POU6F1* (*Brn5*), classified as non - octamer transcription factors (“Non-oct factors”), are optimized to recognize



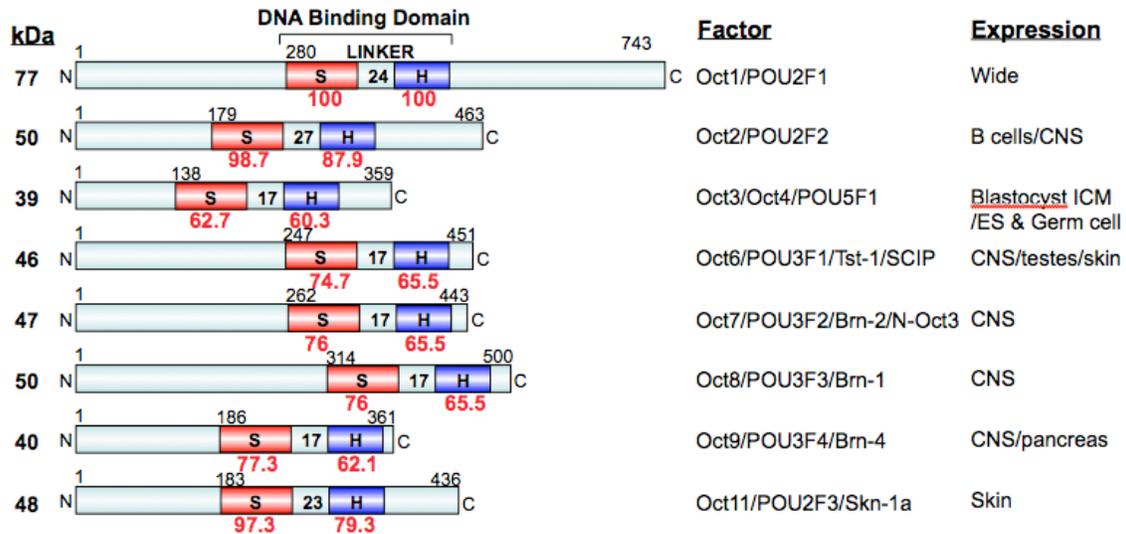
$^A/_T\text{TAT}^T/_c\text{CAT}$, GCATAAATAAT and GCATNN(N)TAAT, respectively (6, 7, 11). This empirical classification turns out to be useful (Figure 1.1 and 1.2). Analyzing POU domains using identity-based dendrograms largely separates them into two groups: Oct and Non-oct factors, and POU domains of Oct factors more closely resemble those of Oct1, the prototype of Oct factors, than those of Non-oct factors. Although the optimal binding sequences are different between POU domain factors, they do not completely lose binding affinity to the others' high affinity sites. Considering some factors are commonly expressed in specific tissues during developmental processes, they are speculated to compete with each other for the binding sites. The level of competition can be determined by affinity to binding sites and amount of expression, which may fine-tune the expression of target genes.

Cellular Functions of *Oct1/POU2F1* and *Oct4/POU5F1*

Oct4/POU5F1

Among the Oct factors, research has mainly dealt with Oct1 and Oct4. Whereas Oct1 seems to be the initial theme of studies, recent attention has been paid on Oct4, a critical pluripotency regulator of stem cells. Oct4 expression was originally identified to be limited to germ cells, the blastocyst inner cell mass and its derived embryonic stem cells (ESCs) suggesting that it regulates early embryonic development (5). The essential role of Oct4 was discovered in the knockout mouse model where its deletion abrogates pluripotency of the inner cell mass (12). Also, when Oct4 expression is doubled, pluripotency is lost suggesting that Oct4 expression is tightly regulated between maintenance of stemness and differentiation (13). Therefore, the complex regulation of

• Octamer Binding POU Transcription Factors



• Non-Octamer Binding POU Transcription Factors

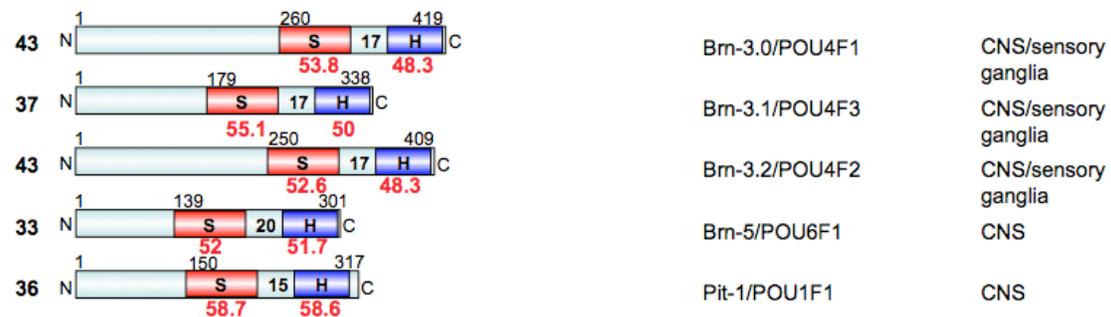


Figure 1.2. Oct and Non-Oct factors and their expression patterns. All POU factors are depicted with conserved DNA binding domains (S:POU_S and H:POU_H) and Linker. kDa represents expected molecular weight. Numbers on light blue box indicate position of amino acids. Red numbers underneath box indicate % identity compared to each domain of Oct1. Black numbers indicate length of linker. Expression patterns are summarized on the right of panel. Taken directly from (2).

Oct4 expression, at multiple levels and by multiple players, is one of the major topics in stem cell biology (reviewed in (2)).

Recently, studies of induced pluripotent stem cells (iPSCs) shed new light on the function of Oct4. Strikingly, only four factors (Oct4, Sox2, Klf4 and c-Myc) are required to induce pluripotent stem cells from fibroblasts. To reduce the potential of malignancy in iPSCs, the field has focused on discovering the factors absolutely necessary for induction (14). Oct4 alone, in combination with small molecule compounds, could generate iPSCs (Reviewed in 15, 16). In comparison to its significance in pluripotency regulation and iPSC biology, little is known about its detailed molecular mechanisms, such as the upstream signal regulator, post-translational modifications and crosstalk with other Oct factors (eg. Oct1). The answers to these questions are imperative to understanding key features in pluripotency at the molecular level.

Oct1/POU2F1

Oct1 was first investigated as a housekeeping regulator because octamer sites are conserved in the regulatory regions of H2B and U1/U2/U6 snRNA where Oct1 binding was confirmed (17, 18). Immunological targets (IL-2 and immunoglobulin) were also a focus for the same reason (19, 20). Surprisingly, the transcription level of these genes is not defective in *Oct1*^{-/-} mouse embryonic fibroblasts (MEFs), although Oct1 knockout mice are embryonic lethal during mid-late gestation (E12.5-E18.5, 21). *Oct1*^{-/-} MEFs are also indistinguishable from wild-type littermates in the aspects of gross morphology and cell division. However, they are hypersensitive to oxidative and genotoxic stress indicating that Oct1 is required for a proper stress response (22). Further, it was shown

that DNA-PK phosphorylates multiple sites on the N-terminus of Oct1 in vitro, and point mutation of all the serine/threonine identified in this study abrogated the ability of Oct1 to promote survival in response to genotoxic stress (23). Research interests on Oct1 are now shifting toward a dynamic stress response activity. The focus now is to identify the molecular mechanism: What kinds of post-translational modifications regulate Oct1 activity responding to stress, and what enzymes are responsible for them? What are the target genes? How does Oct1 regulate transcription?

Oct1 Interacting Partners

Oct1 interacting proteins can be classified by the mechanism by which they regulate transcription (Table 1.1). Oct1 was shown to interact with BRCA1, DNA-PK and PARP-1, well known DNA damage response factors (23-25). DNA-PK phosphorylates N-terminus of Oct1, but further studies need to test whether or not BRCA1 (a ubiquitin ligase) and PARP-1 (an ADP-ribosylase) can post-translationally modify Oct1. Lamin B is a nuclear matrix protein mostly present at the nuclear envelope where it interacts with Oct1. Recently, Lamin B is proposed to spatially sequester Oct1 from target DNA sites (26). Upon stress exposure, Oct1 is released from Lamin B and activates stress response target genes. This interaction network suggests that Oct1 dependent transcription is subjected to multiple upstream regulations in response to stress treatment.

The coactivators not only regulate Oct1 activity under stress conditions. Other coactivators such as OCA-B (OBF-1, Bob.1) and OCA-S are also involved in immune cell development or cell cycle (18, 27). Different from general coactivators, OCA-B

Table 1.1. Previous known Oct1 interacting factors. All factors are categorized by their functional property and by presence or absence of DNA binding activity. * Oct1 interacts with p65 by DNA binding independent manner. Modified from (2).

Functional Category	Factor	DNA binding	Function	References
Regulators/Effectors	BRCA-1	independent	Spindle checkpoint activation Activation of Gadd45a expression	<i>Wang et al., 2004.</i>
	DNA-PK	independent	Oct1 stabilization by phosphorylation	<i>Schild-Poulter et al., 2007.</i>
	PARP-1	independent	Increase in Oct1 DNA binding affinity	<i>Ha et al., 2002.</i>
	LaminB	independent	Oct1 sequestration in nuclear periphery	<i>Malhas et al., 2009.</i>
	OCA-S	independent	Activation of S phase specific H2B expression	<i>Zheng et al., 2003.</i>
	OCA-B	partially independent	Activation of Immunoglobulin expression Activation of PORE-type Oct1 complex	<i>Tomilin et al., 2000.</i>
	SMRT	independent	Repression of Oct-1 transactivation	<i>Kakizawa et al., 2001.</i>
Basal transcription	CAK	independent	Mitosis specific phosphorylation TFIIH recruitment to the preinitiation complex	<i>Inamoto et al., 1997.</i>
	TFIIB	independent	Transcription initiation on TATA-less promoter	<i>Nakshatri et al., 1995.</i>
	TBP	dependent	Interaction with distal activator	<i>Bertolino et al., 2002.</i>
Cooperative TF interactions	SNAPc	dependent	Activation of snRNA transcription	<i>Zhao et al., 2001.</i>
	Sp1	dependent	Activation of snRNA transcription	<i>Lim et al., 2009.</i>
	p65	independent*	Repression of NF-kB transactivation	<i>dela Paz et al., 2007.</i>
	STAT5	dependent	Activation of CyclinD1 expression	<i>Magne et al., 2003.</i>
	Hormone receptors (AR/GR/TR)	dependent or independent	Positive/negative regulation of hormone receptors	<i>Gonzalez et al., 2001, Chandran et al., 1999, Kakizawa et al., 1999.</i>
	Sox2	dependent	Activation of Pax6 expression	<i>Donner et al., 2007.</i>
	C/EBP-b	dependent	Inhibition of immunoglobulin gene transcription	<i>Hatada et al., 2000.</i>

recognizes a specific base pair in the octamer as well as Oct1 through protein-protein interactions (28). The octamer sequence is usually found as two natural variants, 'ATGCAAAT' or 'ATGCTAAT'. It was shown that the fifth 'A' of octamer is critical for OCA-B to associate with the Oct1: DNA complex. Although Oct1 is unable to discriminate between them, OCA-B only uses the former one (ATGCAAAT). This allows Oct factors to differently regulate similar binding sites, increasing the variability of regulation in comparison to limited binding sequence variation. In Oct1 dimeric complexes on alternative binding sequences, OCA-B recognition is through only protein-protein interaction with Oct1 dimer (29). This will be further discussed in a later section. In addition to transcription activation, Oct1 mediates transcription repression by recruitment of a corepressor, silencing mediator for retinoid and thyroid hormone receptors (SMRT), whose repression is achieved by competition with OCA-B for Oct1 binding (30).

Basal transcription factors, TFIIB and TBP as well as their related kinase CAK, interact with Oct1 (31-33). In these contexts, octamer is present near the promoter and Oct1 binding facilitates recruitment of the preinitiation complex, leading to active transcription. Besides of the basal transcription regulation within the proximal promoter, Oct1 also activates transcription together with many other transcription factors by binding to the enhancer. DNA binding of Oct1 with other transcription factors is highly cooperative, which suggests that target genes are dramatically activated only when both factors (Oct1 and partner) reach the threshold in their local concentration. In some cases, two different binding sites for Oct1 and partner are neighboring but, in other cases, they are distant requiring DNA looping for cooperativity. For example, snRNA transcription is

activated by Oct1-SP1 complex in the former way, while by Oct1-SNAPc complex in the latter way (17, 34). The Oct1 and p65 complex is unique in that Oct1 is independent of DNA binding for interaction with p65, which may expand the scope of Oct1 activity beyond DNA binding motifs (35). Oct1 is also well known for regulating hormone receptor dependent transcription (36-38). This issue is an active area of Oct1 research. Another interesting partner is Sox2. Pax6 transcription was shown to be regulated by Oct1-Sox2 complex on the Oct-Sox motifs in developmental processes (39). This finding raises the question of how the core pluripotent regulators (Oct4, Sox2 and Nanog) communicate with Oct1 in ESCs. Core factors are frequently found in the same regulatory regions in genome-wide ChIP analysis (40). Because both Oct1 and Oct4 are co-expressed in ESCs and share DNA binding sites, they may compete with each other on the same sites resulting in two different complexes: Oct1-Sox2-Nanog or Oct4-Sox2-Nanog. It will be interesting to compare the transcriptional outcomes from these two different complexes in terms of regulation of ESC pluripotency.

So far, most studies of Oct1 and its interactor complexes have paid little attention to various epigenetic modifications such as histone modifications and DNA methylation. As a result, Oct1 activities (summarized in Table 1.1) may be only part of the picture, and may need to be revisited with consideration to epigenetic regulation.

DNA Bound Oct Factor Structures

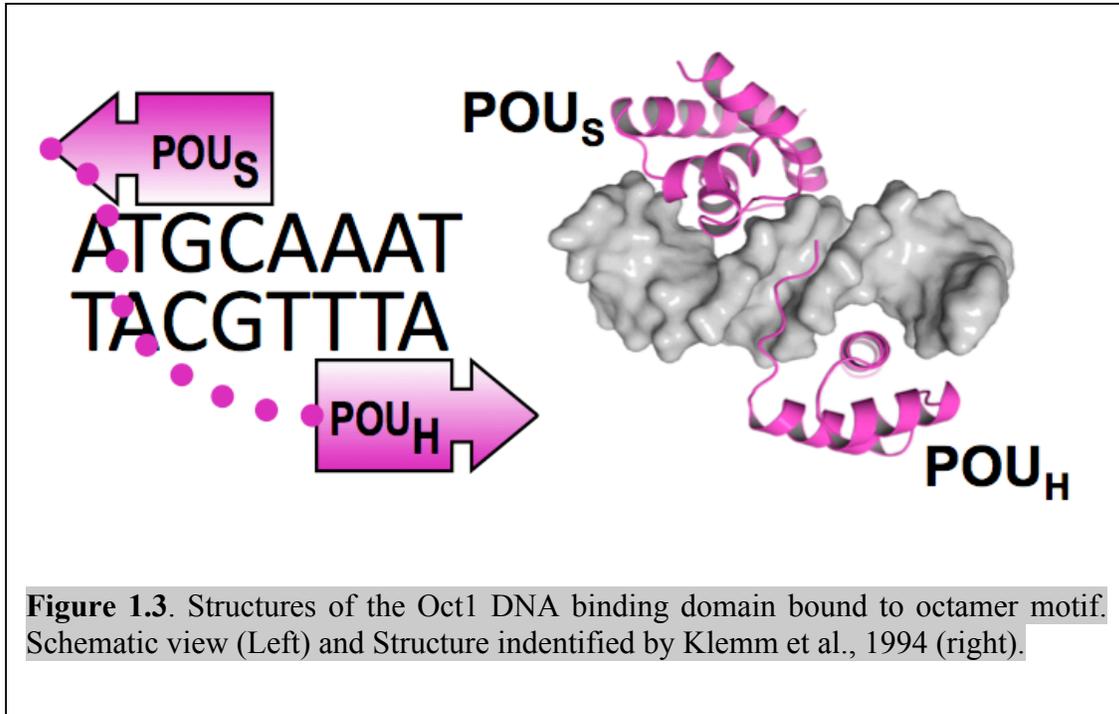
Monomeric binding mode of Oct1

The DNA binding domain of Oct1 was co-crystallized with its corresponding canonical DNA binding sequence unveiling the structure of Oct1 bound to octamer DNA

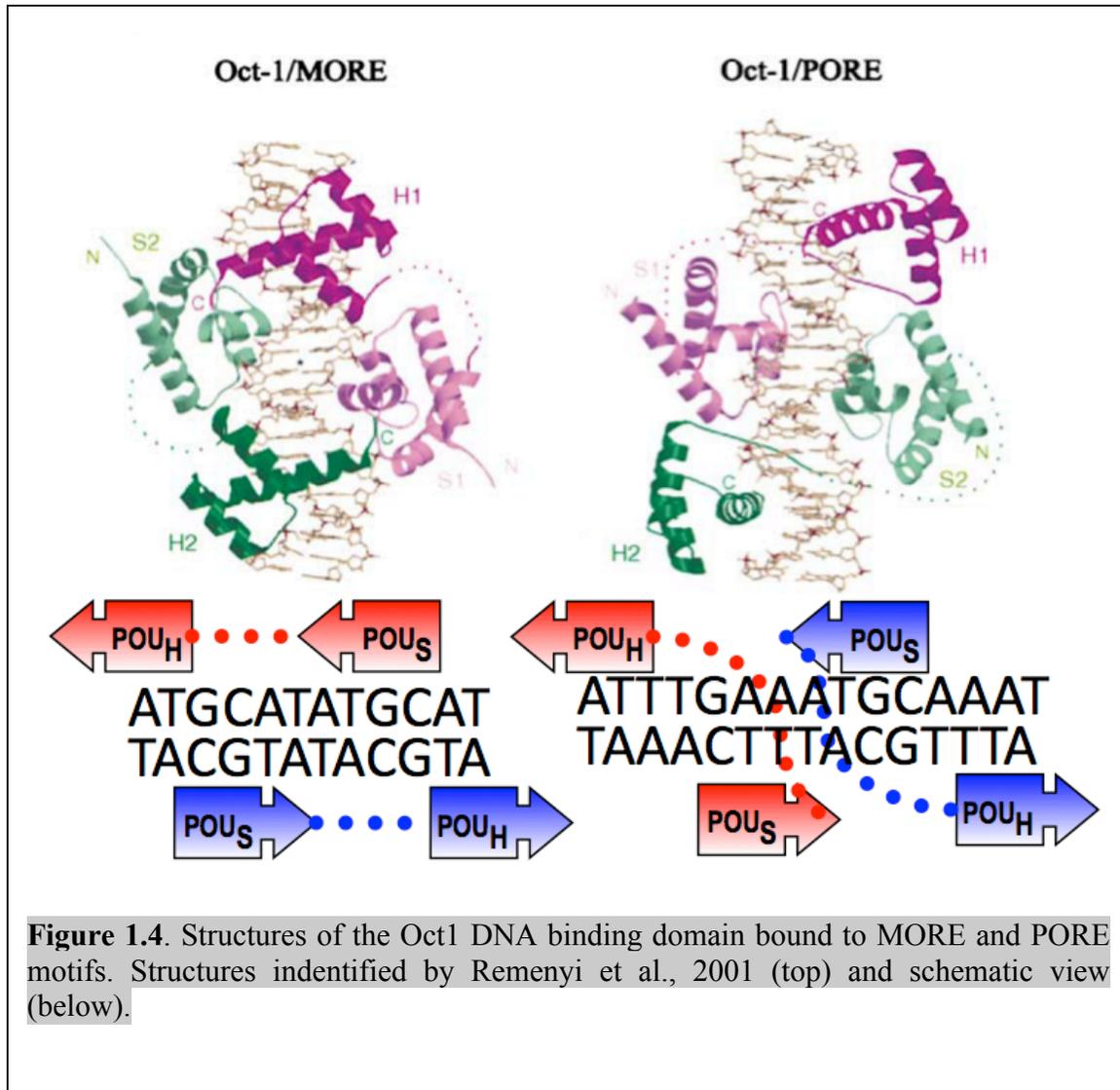
(9, Figure 1.3). Results from this study found that the POU_S domain contains 75 amino acids that build up four α -helices, three of which form a helix-turn-helix (HTH) unit. The 60 amino acid containing POU_H is also composed of three α -helices forming HTH that is quite similar to previously identified homeodomains. The flexible 24-amino acid linker domain is disordered in the crystal. As shown in Figure 1.3, two different domains, both of which recognize DNA using HTH unit, occupy the opposite sides of the DNA double helix. The third α -helix from each domain contacts multiple residues within corresponding major grooves of the octamer sequence (POU_S: ATGC and POU_H: AAAT). The N-terminal arm of POU_H additionally binds in the minor groove.

Dimeric binding mode of Oct1

Oct1 can bind to DNA both as a monomer and as a homo/hetero-dimer. The first identified dimeric binding was on the heptamer-octamer sequences of the immunoglobulin heavy chain promoter: CTCATGAATATGCAAAT (41). However, the way in which this motif regulates expression of immunoglobulin remains unknown. Next, the Palindromic Oct factor Recognition Element (PORE): ATTTGAAATGCAAAT was identified to activate Osteopontin transcription in ES cells (42). This sequence is located at the first intron of *Osteopontin* locus where Oct4/Oct4, Oct4/Oct1 or Oct4/Oct6 dimerization is possible. Interestingly, Oct factors dimerize on both of these DNA motifs, although the Oct dimer configuration of these sites turn out to be quite different (43). Reminyi et al. characterized PORE- and MORE (more PORE)-type Oct1 dimerization using X-ray crystallography. MORE was defined in this study, which is a derivative of heptamer-octamer motif: ATGCATATGCAT. Overall, two Oct1 molecules bind



palindromically to PORE and MORE along the different axis: the MORE half site contains POU_S from one molecule and POU_H from the other one, while the PORE half site contains both domains from a single molecule (Figure 1.4). In detail, POU_H and POU_S bind to AT and ATGC, respectively, on perfect palindromic MORE, while POU_S recognizes two different sequences (ATGC or TTTC) on semipalindromic PORE. The big difference between monomer and dimer configuration is whether or not to form a protein-protein interface between domains: Although two subdomains structurally have no intramolecular interaction on the octamer, two different Oct1 generate protein-protein interfaces through intermolecular interaction between POU_S and POU_H on the PORE and MORE. The interfaces induced from PORE and MORE- type dimerization are also distinguished: the docking of C-terminus of POU_H in loop between 3rd and 4th helix of POU_S in the MORE-type *vs.* the interaction of N-terminus of POU_H with residue in 1st helix of POU_S in the PORE. This structural difference results in selective transcriptional activation. OCA-B, a well-known coactivator of Oct1, favors PORE-type dimerization as an activation target (29). Structurally, the OCA-B binding area is occupied by the C-terminus of the POU_H that is used for protein-protein interface on the MORE. The same space, however, is available in PORE-type dimerization. Therefore, OCA-B introduction into several cell lines shows a huge induction of reporter activity only with the PORE reporter gene. This novel regulation needs to be further validated. For example, it needs to be recapitulated *in vivo* by assessing whether *OCA-B* knockout mice would show defects in Osteopontin expression but not immunoglobulin or other MORE regulated genes. Further studies need to be conducted to identify additional regulatory regions that contain MORE and PORE motifs. In the structure study by Reminyi et al., phospho-



mimetic mutation of Oct1 is shown to choose between the PORE and MORE. Therefore, how post-translational modification of Oct1 regulates octamer-, MORE-, or PORE-type binding needs to be tested, particularly in the context of unique assemblages of cofactors. It will be interesting to identify a specific interacting partner on the MORE induced interface.

Beyond the PORE and MORE

Since a dimeric binding of Oct factor was first identified in 1989, only small numbers of possible target genes have been examined under the PORE and MORE dependent regulation until 2008. It had not been studied whether or not there are unidentified multimeric binding motifs. Therefore, it has been thought that Oct factor dependent transcription is mostly regulated through a monomeric binding on the octamer. This general idea has been challenged by a novel high-throughput study (44). Regular genome-wide ChIP analysis cannot address whether Oct factor binding is monomeric or multimeric. To make it possible, custom tiling DNA oligonucleotide microarrays were designed and synthesized based on genome-wide ChIP analysis for Oct4 occupancy in ESCs, and used for analyzing the binding patterns. Surprisingly, multimeric Oct4 binding is much more pervasive than had been expected, and is formed on a variety of novel multimeric binding motifs that are formed by combinations of the two basic half sites. The other Oct factors are likely to be similar in nature because Oct factors are highly similar in the DNA binding domain. However, as a next step it is imperative to validate that multimeric binding has distinct functions compared to monomeric binding in target gene regulation. Because this novel finding unveils only the tip of the iceberg, it demands us to

expand our scope to identify mechanisms by which “single” or “multiple” Oct factors bind to DNA and regulate transcription.

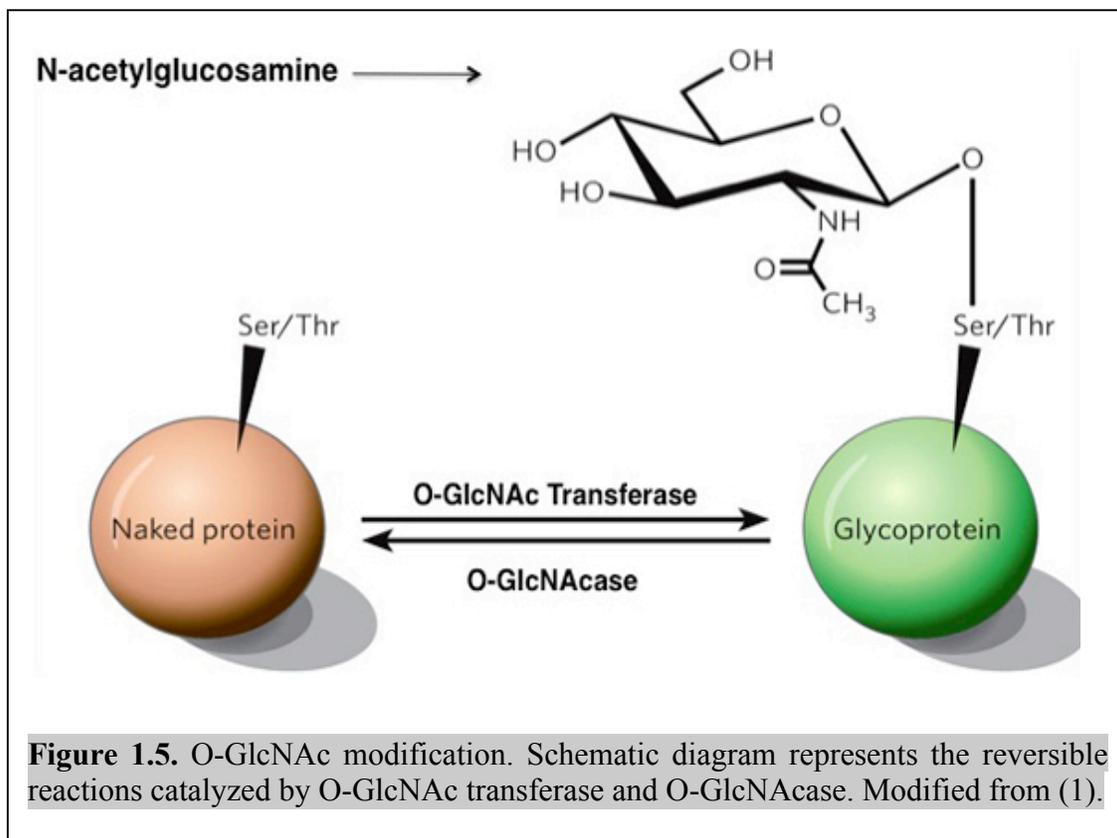
Transcription Factors in Mitosis

Most studies of transcription regulation are focused on interphase cells where transcription actively occurs on euchromatin in the nucleus. When cells enter mitosis, chromosomal DNA is condensed into mitotic chromatin resulting in repression of most transcription (45). However, cyclin B1 transcription is reversely regulated as it is maximized at G2/M but repressed at G1 phase (46). The promoter region of cyclin B1 is protected from mitotic DNA condensation, demonstrated by DNA endonuclease sensitivity, allowing NF-Y transcription factor to bind the promoter for the activation of cyclin B1 transcription. This observation suggests that mitotic DNA condensation is not enough to block the accessibility of transcription factors. Supporting this idea, live image analysis demonstrates that mitotic chromatin is kept accessible to nucleosomes and several transcription factors (47). Nevertheless, because repression of transcription is the hallmark of mitotic chromatin, several different mechanisms are involved in this repression (45). One mechanism is the displacement of transcription factors from mitotic chromatin. C2H2 zinc finger DNA binding domain containing transcription factor such as SP1 and YY1 are excluded from DNA during mitosis after they are phosphorylated in the DNA binding domain (48, 49). Specific phospho-residues have been mapped, and these interfere with DNA binding of those factors by negative charge repulsion against the DNA backbone. Intriguingly, Pit-1, a POU transcription factor family member, undergoes mitotic DNA exclusion via a similar mechanism (50). Its phospho-residue is

positioned in the highly conserved region of the POU_H domain. The matched residue of Oct1 (Ser 385) is also phosphorylated during mitosis (51). Immunofluorescent experiments confirm that Oct1 is clearly excluded from mitotic chromatin (52). The caveat of the previous study is that phospho-Ser 385 did not completely inhibit Oct1 binding on the octamer in vitro and it is unclear whether it inhibited multimeric binding on different motifs (51). Most of all, the central question is why Oct1 needs to be displaced from mitotic chromatin. There are two possibilities. First, during rapid reorganization of chromatin structure in mitosis, transcription factors may not be as tightly controlled as they are in interphase chromatin. Oct1 is a stress response transcription factor that can activate cell-cycle arrest genes (53). Inhibition of Oct1 DNA binding may function as a safety system where normal mitotic cells are protected from unexpected cell-cycle arrest. Second, Lamin B is an Oct1 interactor and a mitotic regulator after breakdown of nuclear envelope (54). Oct1 may move out from DNA to perform one or more unknown mitotic activities with Lamin B during mitosis.

O-GlcNAc Modification of Transcription Factors

Although secretory and cell membrane proteins are subjected to various types of glycosylation moving through ER to golgi, cytosolic and nuclear proteins can be mainly regulated by O-GlcNAc modification. The chemical reaction is summarized with an enzymatic transfer of the N-acetyl glucosamine (GlcNAc) moiety of UDP-GlcNAc onto serines or threonines of proteins (55). This is a reversible reaction in which O-linked b-N-acetylglucosamine transferase (OGT) catalyzes glycosylation but b-D-N-acetylglucosaminidase (O-GlcNAcase) reverts it (Figure 1.5). Although glycosylation has



been poorly studied relative to phosphorylation, recently many transcription factors have been found with O-GlcNAc modification, implicating it as a key post-translational modification that regulates transcription factor activities in response to upstream signals. Two important issues of O-GlcNAc modification in transcription factors have been actively studied. First, many researchers raised the question whether O-GlcNAc modification antagonizes phosphorylation by competing for the same Ser/Thr residues. Studies of p53 and Snail1 tried to identify this competition (56, 57). Both studies identified one glycosylation site each of p53 and Snail1 by mass spectrometry and showed that glycosylation inhibits their own phosphorylation and subsequent ubiquitination pathway, leading to their stabilization. Different from initial speculation, the site of phosphorylation in Snail1 and p53 is not the same as that of glycosylation. Instead, O-GlcNAc modification antagonizes phosphorylation via blockade of the interaction with a kinase. Therefore, the direct antagonism of these modifications needs to be further studied. Second, transcription factors regulated by O-GlcNAc modification were speculated to function as a glucose sensor because glucose is required to form UDP-GlcNAc, the substrate of OGT. NF- κ B and Snail1 were tested for activity and stability under abnormal levels of glucose (57, 58). In hyperglycaemic conditions, levels of O-GlcNAc modification are robustly induced on both proteins. As a result, NF- κ B is released from I κ B and translocates into the nucleus for activating downstream genes. Snail1 is also more activated for transcription via increased stability. In physiological contexts, high glucose in animal models can induce chronic inflammatory genes and cancer metastasis (55, 57, 58).

Comparison between wildtype and *Oct1*^{-/-} MEFs in gene expression profiles

showed that Oct1 is required for the proper regulation of TCA cycle genes (59). Consistent with this result, even under high glucose conditions, *Oct1*^{-/-} MEFs preferentially activate the oxidative phosphorylation pathway using different energy sources rather than the glycolytic pathway. Depending on glucose availability, a group of TCA cycle genes are coordinately regulated in their transcription (60). Taken together, one hypothesis consistent with the data is that Oct1 mediates TCA cycle gene transcription after sensing the level of glucose. In other words, the level of glucose can act as an upstream signal regulator of Oct1. Oct1 is therefore speculated to be regulated by O-GlcNAc modification as shown with p53 and Snail1 whose activities are also regulated by the level of glucose. To test this hypothesis, it is imperative to identify the specific residue(s) of Oct1 modified by O-GlcNAc.

Plan of the Dissertation

As reviewed so far, Oct1 is a bona fide stress response transcription factor. However, it is still unclear how Oct1 regulates downstream target genes in response to genotoxic and oxidative stress. In Chapter 2, I examine the mechanism of Oct1 dependent transcription. Phosphorylation is a known modification to regulate Oct1 activity following genotoxic stress (23). In this study, phosphorylation sites were identified only by an in vitro kinase assay that may be artificial. To resolve this problem, I purified endogenous Oct1 from stress-treated HeLa cells and identified Oct1 modification, using mass spectrometry. As described above, Oct1 binding can be multimeric as well as monomeric. I connected Oct1 modification with multimeric binding regulation. Chapter 3 will discuss how phosphorylated and ubiquitinated Oct1 is displaced from mitotic DNA,

and what kind of function Oct1 performs in mitotic cells. This study will expand Oct1 activity beyond DNA dependent transcription. In Chapter 4, I also identified O-GlcNAc modification in Oct1 using mass spectrometry. To understand the roles of O-GlcNAc modification in Oct1, I generated and assessed mutant Oct1 by comparison with WT Oct1. Finally, I confirmed that Oct1 and Oct4 share binding motifs using ChIP analysis in Chapter 2. The ensuing question is whether or not simultaneously occupied Oct1-Oct4 hetero complex occurs in cells. To address this, I performed sequential ChIPs in Chapter 5. Using our newly established high-throughput analysis, I tried to discriminate Oct1 from Oct4 activity in ESCs. This analysis resulted in the identification of genomic sites that do or do not discriminate between Oct1 and Oct4, based on neighboring elements.

References

1. Hart, G. W., M. P. Housley, and C. Slawson. 2007. Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins. *Nature* 446:1017-1022.
2. Kang, J., A. Shakya, and D. Tantin. 2009. Stem cells, stress, metabolism and cancer: a drama in two Octs. *Trends Biochem Sci* 34:491-499.
3. Herr, W., R. A. Sturm, R. G. Clerc, L. M. Corcoran, D. Baltimore, P. A. Sharp, H. A. Ingraham, M. G. Rosenfeld, M. Finney, G. Ruvkun, and et al. 1988. The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and *Caenorhabditis elegans* unc-86 gene products. *Genes Dev* 2:1513-1516.
4. He, X., M. N. Treacy, D. M. Simmons, H. A. Ingraham, L. W. Swanson, and M. G. Rosenfeld. 1989. Expression of a large family of POU-domain regulatory genes in mammalian brain development. *Nature* 340:35-41.
5. Scholer, H. R., S. Ruppert, N. Suzuki, K. Chowdhury, and P. Gruss. 1990. New type of POU domain in germ line-specific protein Oct-4. *Nature* 344:435-439.
6. Gerrero, M. R., R. J. McEvilly, E. Turner, C. R. Lin, S. O'Connell, K. J. Jenne, M. V. Hobbs, and M. G. Rosenfeld. 1993. Brn-3.0: a POU-domain protein expressed in the sensory, immune, and endocrine systems that functions on elements distinct from known octamer motifs. *Proc Natl Acad Sci U S A* 90:10841-10845.

7. Andersen, B., M. D. Schonemann, R. V. Pearse, 2nd, K. Jenne, J. Sugarman, and M. G. Rosenfeld. 1993. Brn-5 is a divergent POU domain factor highly expressed in layer IV of the neocortex. *J Biol Chem* 268:23390-23398.
8. Ryan, A. K., and M. G. Rosenfeld. 1997. POU domain family values: flexibility, partnerships, and developmental codes. *Genes Dev* 11:1207-1225.
9. Klemm, J. D., M. A. Rould, R. Aurora, W. Herr, and C. O. Pabo. 1994. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* 77:21-32.
10. Sturm, R. A., G. Das, and W. Herr. 1988. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes Dev* 2:1582-1599.
11. Elsholtz, H. P., V. R. Albert, M. N. Treacy, and M. G. Rosenfeld. 1990. A two-base change in a POU factor-binding site switches pituitary-specific to lymphoid-specific gene expression. *Genes Dev* 4:43-51.
12. Nichols, J., B. Zevnik, K. Anastasiadis, H. Niwa, D. Klewe-Nebenius, I. Chambers, H. Scholer, and A. Smith. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95:379-391.
13. Niwa, H., J. Miyazaki, and A. G. Smith. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 24:372-376.
14. Takahashi, K., and S. Yamanaka. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126:663-676.
15. Kang, J., M. Gemberling, M. Nakamura, F. G. Whitby, H. Handa, W. G. Fairbrother, and D. Tantin. 2009. A general mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and oxidative stress. *Genes Dev* 23:208-222.
16. Yuan, X., H. Wan, X. Zhao, S. Zhu, Q. Zhou, and S. Ding. 2011. Brief report: combined chemical treatment enables oct4-induced reprogramming from mouse embryonic fibroblasts. *Stem Cells* 29:549-553.
17. Zhao, X., P. S. Pendergrast, and N. Hernandez. 2001. A positioned nucleosome on the human U6 promoter allows recruitment of SNAPc by the Oct-1 POU domain. *Mol Cell* 7:539-549.

18. Zheng, L., R. G. Roeder, and Y. Luo. 2003. S phase activation of the histone H2B promoter by OCA-S, a coactivator complex that contains GAPDH as a key component. *Cell* 114:255-266.
19. Garrity, P. A., D. Chen, E. V. Rothenberg, and B. J. Wold. 1994. Interleukin-2 transcription is regulated in vivo at the level of coordinated binding of both constitutive and regulated factors. *Mol Cell Biol* 14:2159-2169.
20. Ushmorov, A., O. Ritz, M. Hummel, F. Leithauser, P. Moller, H. Stein, and T. Wirth. 2004. Epigenetic silencing of the immunoglobulin heavy-chain gene in classical Hodgkin lymphoma-derived cell lines contributes to the loss of immunoglobulin expression. *Blood* 104:3326-3334.
21. Wang, V. E., T. Schmidt, J. Chen, P. A. Sharp, and D. Tantin. 2004. Embryonic lethality, decreased erythropoiesis, and defective octamer-dependent promoter activation in Oct-1-deficient mice. *Mol Cell Biol* 24:1022-1032.
22. Tantin, D., C. Schild-Poulter, V. Wang, R. J. Hache, and P. A. Sharp. 2005. The octamer binding transcription factor Oct-1 is a stress sensor. *Cancer Res* 65:10750-10758.
23. Schild-Poulter, C., A. Shih, D. Tantin, N. C. Yarymowich, S. Soubeyrand, P. A. Sharp, and R. J. Hache. 2007. DNA-PK phosphorylation sites on Oct-1 promote cell survival following DNA damage. *Oncogene* 26:3980-3988.
24. Wang, R. H., H. Yu, and C. X. Deng. 2004. A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint. *Proc Natl Acad Sci U S A* 101:17108-17113.
25. Ha, H. C., L. D. Hester, and S. H. Snyder. 2002. Poly(ADP-ribose) polymerase-1 dependence of stress-induced transcription factors and associated gene expression in glia. *Proc Natl Acad Sci U S A* 99:3270-3275.
26. Malhas, A. N., C. F. Lee, and D. J. Vaux. 2009. Lamin B1 controls oxidative stress responses via Oct-1. *J Cell Biol* 184:45-55.
27. Strubin, M., J. W. Newell, and P. Matthias. 1995. OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. *Cell* 80:497-506.
28. Cepek, K. L., D. I. Chasman, and P. A. Sharp. 1996. Sequence-specific DNA binding of the B-cell-specific coactivator OCA-B. *Genes Dev* 10:2079-2088.
29. Tomilin, A., A. Remenyi, K. Lins, H. Bak, S. Leidel, G. Vriend, M. Wilmanns, and H. R. Scholer. 2000. Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. *Cell* 103:853-864.

30. Kakizawa, T., T. Miyamoto, K. Ichikawa, T. Takeda, S. Suzuki, J. Mori, M. Kumagai, K. Yamashita, and K. Hashizume. 2001. Silencing mediator for retinoid and thyroid hormone receptors interacts with octamer transcription factor-1 and acts as a transcriptional repressor. *J Biol Chem* 276:9720-9725.
31. Inamoto, S., N. Segil, Z. Q. Pan, M. Kimura, and R. G. Roeder. 1997. The cyclin-dependent kinase-activating kinase (CAK) assembly factor, MAT1, targets and enhances CAK activity on the POU domains of octamer transcription factors. *J Biol Chem* 272:29852-29858.
32. Nakshatri, H., P. Nakshatri, and R. A. Currie. 1995. Interaction of Oct-1 with TFIIB. Implications for a novel response elicited through the proximal octamer site of the lipoprotein lipase promoter. *J Biol Chem* 270:19613-19623.
33. Bertolino, E., and H. Singh. 2002. POU/TBP cooperativity: a mechanism for enhancer action from a distance. *Mol Cell* 10:397-407.
34. Lim, K., and H. I. Chang. 2009. O-GlcNAc modification of Sp1 inhibits the functional interaction between Sp1 and Oct1. *FEBS Lett* 583:512-520.
35. dela Paz, N. G., S. Simeonidis, C. Leo, D. W. Rose, and T. Collins. 2007. Regulation of NF-kappaB-dependent gene expression by the POU domain transcription factor Oct-1. *J Biol Chem* 282:8424-8434.
36. Gonzalez, M. I., and D. M. Robins. 2001. Oct-1 preferentially interacts with androgen receptor in a DNA-dependent manner that facilitates recruitment of SRC-1. *J Biol Chem* 276:6420-6428.
37. Chandran, U. R., B. S. Warren, C. T. Baumann, G. L. Hager, and D. B. DeFranco. 1999. The glucocorticoid receptor is tethered to DNA-bound Oct-1 at the mouse gonadotropin-releasing hormone distal negative glucocorticoid response element. *J Biol Chem* 274:2372-2378.
38. Kakizawa, T., T. Miyamoto, K. Ichikawa, A. Kaneko, S. Suzuki, M. Hara, T. Nagasawa, T. Takeda, J. Mori, M. Kumagai, and K. Hashizume. 1999. Functional interaction between Oct-1 and retinoid X receptor. *J Biol Chem* 274:19103-19108.
39. Donner, A. L., V. Episkopou, and R. L. Maas. 2007. Sox2 and Pou2f1 interact to control lens and olfactory placode development. *Dev Biol* 303:784-799.
40. Boyer, L. A., T. I. Lee, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner, D. K. Gifford, D. A. Melton, R. Jaenisch, and R. A. Young. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122:947-956.

41. LeBowitz, J. H., R. G. Clerc, M. Brenowitz, and P. A. Sharp. 1989. The Oct-2 protein binds cooperatively to adjacent octamer sites. *Genes Dev* 3:1625-1638.
42. Botquin, V., H. Hess, G. Fuhrmann, C. Anastassiadis, M. K. Gross, G. Vriend, and H. R. Scholer. 1998. New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2. *Genes Dev* 12:2073-2090.
43. Remenyi, A., A. Tomilin, E. Pohl, K. Lins, A. Philippsen, R. Reinbold, H. R. Scholer, and M. Wilmanns. 2001. Differential dimer activities of the transcription factor Oct-1 by DNA-induced interface swapping. *Mol Cell* 8:569-580.
44. Tantin, D., M. Gemberling, C. Callister, and W. G. Fairbrother. 2008. High-throughput biochemical analysis of in vivo location data reveals novel distinct classes of POU5F1(Oct4)/DNA complexes. *Genome Res* 18:631-639.
45. Gottesfeld, J. M., and D. J. Forbes. 1997. Mitotic repression of the transcriptional machinery. *Trends Biochem Sci* 22:197-202.
46. Sciortino, S., A. Gurtner, I. Manni, G. Fontemaggi, A. Dey, A. Sacchi, K. Ozato, and G. Piaggio. 2001. The cyclin B1 gene is actively transcribed during mitosis in HeLa cells. *EMBO Rep* 2:1018-1023.
47. Chen, D., M. Dundr, C. Wang, A. Leung, A. Lamond, T. Misteli, and S. Huang. 2005. Condensed mitotic chromatin is accessible to transcription factors and chromatin structural proteins. *J Cell Biol* 168:41-54.
48. Dovat, S., T. Ronni, D. Russell, R. Ferrini, B. S. Cobb, and S. T. Smale. 2002. A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains. *Genes Dev* 16:2985-2990.
49. Rizkallah, R., and M. M. Hurt. 2009. Regulation of the transcription factor YY1 in mitosis through phosphorylation of its DNA-binding domain. *Mol Biol Cell* 20:4766-4776.
50. Caelles, C., H. Hennemann, and M. Karin. 1995. M-phase-specific phosphorylation of the POU transcription factor GHF-1 by a cell cycle-regulated protein kinase inhibits DNA binding. *Mol Cell Biol* 15:6694-6701.
51. Segil, N., S. B. Roberts, and N. Heintz. 1991. Mitotic phosphorylation of the Oct-1 homeodomain and regulation of Oct-1 DNA binding activity. *Science* 254:1814-1816.
52. Martinez-Balbas, M. A., A. Dey, S. K. Rabindran, K. Ozato, and C. Wu. 1995. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* 83:29-38.

53. Takahashi, S., S. Saito, N. Ohtani, and T. Sakai. 2001. Involvement of the Oct-1 regulatory element of the gadd45 promoter in the p53-independent response to ultraviolet irradiation. *Cancer Res* 61:1187-1195.
54. Tsai, M. Y., S. Wang, J. M. Heidinger, D. K. Shumaker, S. A. Adam, R. D. Goldman, and Y. Zheng. 2006. A mitotic lamin B matrix induced by RanGTP required for spindle assembly. *Science* 311:1887-1893.
55. Issad, T., and M. Kuo. 2008. O-GlcNAc modification of transcription factors, glucose sensing and glucotoxicity. *Trends Endocrinol Metab* 19:380-389.
56. Yang, W. H., J. E. Kim, H. W. Nam, J. W. Ju, H. S. Kim, Y. S. Kim, and J. W. Cho. 2006. Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. *Nat Cell Biol* 8:1074-1083.
57. Park, S. Y., H. S. Kim, N. H. Kim, S. Ji, S. Y. Cha, J. G. Kang, I. Ota, K. Shimada, N. Konishi, H. W. Nam, S. W. Hong, W. H. Yang, J. Roth, J. I. Yook, and J. W. Cho. 2010. Snail1 is stabilized by O-GlcNAc modification in hyperglycaemic condition. *EMBO J* 29:3787-3796.
58. Yang, W. H., S. Y. Park, H. W. Nam, H. Kim do, J. G. Kang, E. S. Kang, Y. S. Kim, H. C. Lee, K. S. Kim, and J. W. Cho. 2008. NFkappaB activation is associated with its O-GlcNAcylation state under hyperglycemic conditions. *Proc Natl Acad Sci U S A* 105:17345-17350.
59. Shakya, A., R. Cooksey, J. E. Cox, V. Wang, D. A. McClain, and D. Tantin. 2009. Oct1 loss of function induces a coordinate metabolic shift that opposes tumorigenicity. *Nat Cell Biol* 11:320-327.
60. DeRisi, J. L., V. R. Iyer, and P. O. Brown. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686.

CHAPTER 2

A GENERAL MECHANISM FOR TRANSCRIPTION REGULATION BY OCT1 AND OCT4 IN RESPONSE TO GENOTOXIC AND OXIDATIVE STRESS

Jinsuk Kang, Matthew Gemberling, Mitsuhiro Nakamura, Frank G. Whitby, Hiropshi
Handa, William G. Fairbrother, and Dean Tantin

Reprinted from *Genes & Development*, Vol. 23, pp. 208-222

Copyright © 2009, with permission from Cold Spring Harbor Laboratory Press

A general mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and oxidative stress

Jinsuk Kang,¹ Matthew Gemberling,² Mitsuhiro Nakamura,³ Frank G. Whitby,⁴ Hiroshi Handa,³ William G. Fairbrother,² and Dean Tantin^{1,4,5}

¹Department of Pathology, University of Utah School of Medicine, Salt Lake City, Utah 84112, USA; ²Department of Molecular Biology, Cell Biology, and Biochemistry (MCB Department), Brown University, Providence, Rhode Island 02912, USA; ³Integrated Research Institute, Tokyo Institute of Technology, Yokohama 226-8501, Japan; ⁴Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, Utah 84112, USA

Oct1 and Oct4 are homologous transcription factors with similar DNA-binding specificities. Here we show that Oct1 is dynamically phosphorylated in vivo following exposure of cells to oxidative and genotoxic stress. We further show that stress regulates the selectivity of both proteins for specific DNA sequences. Mutation of conserved phosphorylation target DNA-binding domain residues in Oct1, and Oct4 confirms their role in regulating binding selectivity. Using chromatin immunoprecipitation, we show that association of Oct4 and Oct1 with a distinct group of in vivo targets is inducible by stress, and that Oct1 is essential for a normal post-stress transcriptional response. Finally, using an unbiased Oct1 target screen we identify a large number of genes targeted by Oct1 specifically under conditions of stress, and show that several of these inducible Oct1 targets are also inducibly bound by Oct4 in embryonic stem cells following stress exposure.

[Keywords: Oct1; Oct4; stress response]

Supplemental material is available at <http://www.genesdev.org>.

Received June 17, 2008; revised version accepted November 24, 2008.

Oct1 and Oct4 (products of the *Pou2f1* and *Pou5f1* genes) are members of the POU (Pit-1, Oct1/2, Unc-86) domain transcription factor family (Herr et al. 1988; Ryan and Rosenfeld 1997). This family is defined by the presence of a bipartite DNA-binding domain in which two subdomains, covalently connected by a flexible linker, typically recognize DNA through major groove interactions on opposite sides of the helix (Klemm et al. 1994). The classical DNA recognition sequence is known as an octamer motif [5'-ATGCCAAAT-3', hereafter called a "simple" octamer]. However, we demonstrated recently that native binding sites for Oct4 frequently exist in complex paired, overlapping, and nonconsensus configurations (Tantin et al. 2008).

Oct4 is a master regulator of the stem cell state and has recently been shown to be one of three proteins sufficient to reprogram differentiated adult mouse and human cells to the embryonic stem (ES) cell lineage (Okita et al. 2007; Takahashi et al. 2007; Nakagawa et al. 2008). The biological function of Oct1 is more enigmatic. Oct1 is known to interact with regulatory sites in interleukin,

immunoglobulin, and histone genes (Garrity et al. 1994; Zheng et al. 2003; Ushmorov et al. 2004; Murayama et al. 2006). Oct1 also moderately stimulates gene expression reporter constructs linked to target sequences in transient transfection assays (Sive et al. 1986; LeBowitz et al. 1988). However, we showed that Oct1 is nonessential for native *H2B*, *IgH*, and *Igκ* expression (V.E. Wang et al. 2004; V.E.H. Wang et al. 2004). Oct1-deficient cells appear morphologically normal in light microscopy and divide at normal rates. Oct1-deficient mice die in mid-late gestation (embryonic days 12.5–18.5 [E12.5–E18.5]) (V.E.H. Wang et al. 2004).

We determined previously that *Oct1*^{-/-} mouse embryonic fibroblasts (MEFs) are hypersensitive to oxidative and genotoxic stress (Tantin et al. 2005). One explanation for this result is that constitutive products of Oct1-mediated transcription participate in stress response pathways. Support for an alternative hypothesis, namely that Oct1 directly senses cellular stress, comes from the findings that Oct1 interacts with the Ku70 subunit of DNA-dependent protein kinase (DNA-PK) (Schild-Poulter et al. 2001) and is phosphorylated in vitro by DNA-PK at physiologically important serine and threonine residues (Schild-Poulter et al. 2007). Oct1 also interacts with BRCA1 (Fan et al. 2002; R.H. Wang et al.

⁵Corresponding author.

E-MAIL dean.tantin@path.utah.edu; FAX (801) 582-2417.

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.1750709>.

2004), and PARP-1 (Nie et al. 1998), known participants in stress response pathways.

Here, using an affinity purification approach we identify *in vivo* post-translational Oct1 modification events following treatment of HeLa cells with ionizing radiation (IR) or H₂O₂. Comparison with a second Oct1 modification data set indicated that two DNA-binding domain Oct1 phosphorylation events have the potential to modulate Oct1 association with DNA. Using two model sequences, termed PORE (palindromic octamer-related element) and MORE (more PORE) (Remenyi et al. 2001), we show that DNA-binding domain modifications alter the *in vitro* affinity of Oct1 and Oct4 specifically at complex binding sites. We demonstrate stress-induced binding of both Oct1 and Oct4 to physiological targets *in vivo*, show that induced Oct1 binding regulates native gene expression, and expand the repertoire of complex site categories to which Oct1 and Oct4 can bind. Using chromatin immunoprecipitation (ChIP) coupled with deep sequencing (ChIPseq), we identify a large number of constitutive Oct1 targets as well as targets specifically induced in the presence of oxidative stress. These targets frequently contain conserved complex binding sites for Oct1. We demonstrate that Oct4 inducibly binds two of these targets *in vivo* using mouse ES cells and ChIP.

Together, the data show that Oct1 and Oct4 are potent stress response effectors.

Results

Oct1 is phosphorylated in vivo following exposure to genotoxic and oxidative stress

We used a single-step affinity approach to purify Oct1 and other DNA-binding proteins from nuclear extracts (Tantin et al. 2004; V.E.H. Wang et al. 2004). This approach involves ferromagnetic latex nanoparticles ("nanobeads") coupled to multimerized DNA sequence motifs for rapid and sequence-selective enrichment of DNA-binding proteins. Relative to other supports such as sepharose, nanobeads have lower nonspecific binding (due to their nonporous, moderately hydrophilic surface), higher capacity (due to their smaller size), and will remain in suspension until high-speed centrifugation or magnetic separation (Shimizu et al. 2000; Nishio et al. 2008). We used nanobeads coupled to multimerized consensus or mutant simple octamer sequences to purify Oct1 from HeLa cells following IR or H₂O₂ treatment. This affinity purification approach will only isolate Oct1 molecules with modifications that do not inhibit DNA binding. As shown in Figure 1A, an ~100-kDa band was isolated

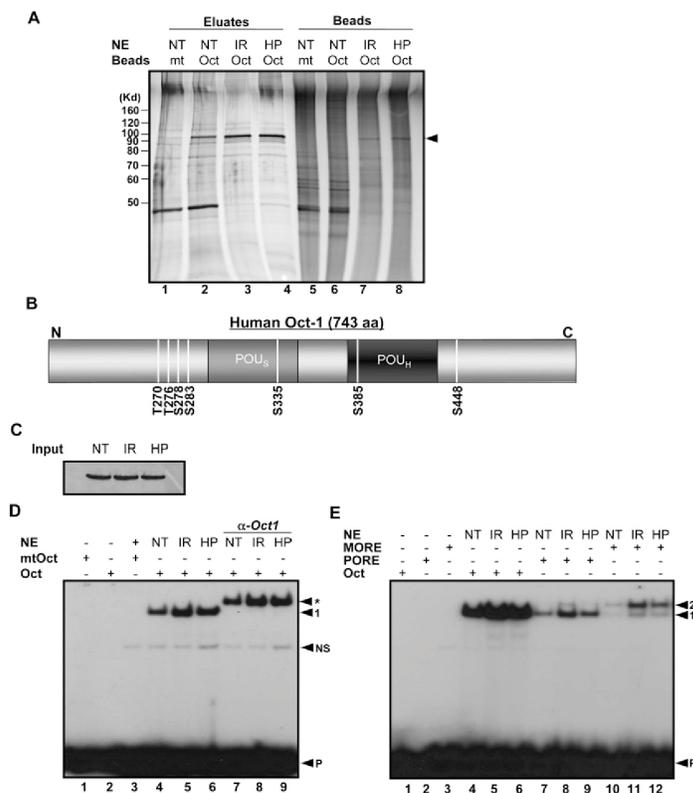


Figure 1. Genotoxic and oxidative stress induce Oct1 phosphorylation and dimerization at complex sites. (A) Oct1 was purified from HeLa nuclear extracts using nanoparticles coupled to multimerized octamer or mutant sequences. HeLa cells were treated with IR or H₂O₂ and incubated for 1 h. The eluted proteins were resolved using SDS-PAGE and silver stained. (NT) No treatment; (HP) H₂O₂. (Arrow) Oct1 band. (B) Identified Oct1 serine and threonine phosphorylation events superimposed on a schematic of the Oct1 amino acid structure. DNA-binding domain modification events identified separately by the Gygi laboratory are also shown. (C) Oct1 Western blot showing protein levels in the input HeLa nuclear extracts. (D) EMSA using a simple octamer probe and extracts from C. Arrows indicate free probe (P), nonspecific band (NS), monomeric Oct1 (1), and Oct1 antibody supershifted band (*). (E) EMSA using radiolabeled simple octamer, PORE, and MORE probes. Arrows indicate monomeric (1) and dimeric (2) occupancy, and free probe (P).

using nanobeads coupled to consensus, but not mutant DNA (Fig. 1A, cf. lane 2 and lane 1). The 100-kDa band was excised, destained, and subjected to tryptic digestion. Mass spectrometry identified peptides consistent with human Oct1. Peptide coverage was good and the sequences were confirmed through sequencing using collision-induced dissociation (CID) in the linear ion trap (data not shown; see the Materials and Methods). The intense band near 40 kDa was identified as β -actin (data not shown).

Using equal amounts of extract from HeLa cells exposed to IR (Fig. 1A, lane 3) or H₂O₂ (Fig. 1A, lane 4), a small increase in bound protein was identified compared with normal HeLa cells (Fig. 1A, lane 2). This increase is consistent with previous reports (Meighan-Mantha et al. 1999; Zhao et al. 2000; B. Wang et al. 2004; Chen and Currie 2005). We excised these bands and subjected them to mass spectroscopy to identify Oct1 modifications. We unambiguously identified Oct1 phosphopeptides, many of which were identified unphosphorylated in the absence of stress (see Table 1). These data show that Oct1 becomes dynamically modified *in vivo* shortly following stress treatment. Although our data were consistent with a large-scale screen for protein phosphorylation events in HeLa cells following IR exposure conducted in the laboratory of Dr. Steven Gygi (S.P. Gygi, unpubl.), two Oct1 DNA-binding domain modifications were identified by the Gygi laboratory that were not observed using our affinity purification approach: phospho-S335 and phospho-S385 (Table 1). Because the Gygi laboratory screen did not rely upon DNA binding to isolate phosphopeptides, we focused on these two modifications, hypothesizing that they regulate DNA binding. The complete panel of *in vivo* phosphorylation events from both studies is shown in Table 1 and in schematic form in Figure 1B.

Alteration in Oct1-binding specificity following stress exposure

The above results demonstrated that Oct1 becomes dynamically modified following stress exposure. We therefore attempted to identify dynamic changes in the properties of the protein that could provide a functional

stress readout. First, we performed Western blotting using Oct1-specific antibodies and extracts from HeLa cells exposed to IR or H₂O₂. Little change was observed (Fig. 1C). We also conducted gel mobility shift experiments using wild-type and mutant simple octamer sequences. A strong octamer sequence-dependent band was identified, but again, little change was observed using extracts from IR or H₂O₂-treated HeLa cells (Fig. 1D, lanes 4–6). There was some increase in DNA-binding activity (Fig. 1D, cf. lanes 5,6 and lane 4), consistent with previous reports (Meighan-Mantha et al. 1999; Zhao et al. 2000; B. Wang et al. 2004; Chen and Currie 2005).

Recently, a PKA-mediated phosphorylation event was identified in the CNS-specific POU domain protein Brn-2/N-Oct-3/Pou3f2 (Nieto et al. 2007). In this case, the modification, at a position homologous to Oct1 S385, was found to alter binding specificity for complex dimeric sites. More recently, we identified native Oct4 genomic binding sites (Tantin et al. 2008). These target sequences were highly enriched for complex multimeric arrangements of half-sites with the capability of binding two or more Oct4 molecules. The potential for these complex arrangements arises from the fact that in the POU domain, the linker connecting the two DNA-binding subdomains is flexible. Because these results showed that native sites for POU transcription factors frequently exist in complex configurations, we tested Oct1 binding to two model complex sequences because structural information was available.

PORE and MORE sequences are both capable of binding Oct proteins as a dimer, but in different configurations (Remenyi et al. 2001). We conducted electrophoretic mobility shift assays (EMSA) using these sequences together with extracts from HeLa cells cultured under standard conditions or exposed to stress. Using a PORE, a somewhat stronger induction of binding by stress was observed compared with simple sequences. This was particularly true of extracts from cells exposed to IR (Fig. 1E, cf. lanes 8,9 and lane 7). Using a MORE, an even stronger induction was observed following stress exposure. This induction was again particularly robust using

Table 1. Oct1 phosphopeptides identified following exposure of HeLa cells to IR (mass error <3 ppm) using LC/MS/MS^a

Peptide ^b	Amino acid position	Modification position	Stress ^c	Missed cleavage	Follow-up?
TIAATPIQLTPQSQST#PK	255–272	T270-P	IR	0	No
RIDT#PSLEEPSDLEEELEQFAK	273–293	T276-P	IR	0	No
RIDTPS#LEEPSDLEEELEQFAK	273–293	S278-P	IR	0	No
RIDT#PS#LEEPSDLEEELEQFAK	273–293	T276-P S278-P	IR	0	No
RIDTPSLEEPS#DLEEELEQFAK	273–293	S283-P	IR	0	No
RIDTPS#LEEPS#DLEEELEQFAK	273–293	S278-P S283-P	IR	0	No
FEALNLS#FK	329–337	S335-P	IR	0	Yes
TS#IETNIR	384–391	S385-P	IR	0	Yes
INPPSSGGTSSS#PIK	437–451	S448-P	IR/HP/HU	0	No

^aModifications identified in this study and in a phosphopeptide screen performed in the Gygi laboratory (S.P. Gygi, unpubl.). S335 and S385 modifications were only identified in the Gygi screen.

^bPhosphorylated amino acid residues are marked with #.

^c[IR] 1200 RAD IR, 1 h; [HP] 1 mM H₂O₂, 1 h; [HU] 1 mM hydroxyurea, 12 h.

IR and was manifested both at the level of overall binding and increased relative dimerization (Fig. 1E, cf. lane 11 and lanes 10,12). The induced binding could be eliminated by shifting the equilibrium to the bound state, as inclusion of higher amounts of extract in the binding reactions increased dimer occupancy and decreased stress induction (data not shown).

Stress-induced association of Oct1 with endogenous complex sites

A MORE has previously been identified in the promoter for the human RNA polymerase II (Pol II) large subunit (*Polr2a*) (Fig. 2A, top panel; Tomilin et al. 2000). We identified a second MORE located immediately upstream. Both MOREs are conserved across multiple mammalian species (Fig. 2A, bottom panel). Using EMSA with HeLa nuclear extract and a probe derived from the *Polr2a* 2XMORE region, we confirmed that this compound element could interact with four Oct1 molecules, and that Oct1 binding could be strongly induced by stress (Fig. 2B). In Figure 2, B and C, the gel was run further to resolve the multiply bound species, and the labeled free probe was consequently run off the gel. In Figure 2B, increasing amounts of normal and IR-treated HeLa cell extract were incubated with either a simple octamer sequence or a 2XMORE probe of identical size. DNA binding was strongly enhanced by stress (Fig. 2B, cf. lanes 7,12,17). We used an Oct1-specific antibody to confirm that the multiply bound species contained Oct1 (Fig. 2B, lanes 8,13,18). In Figure 2C, point mutations were engineered into the *Polr2a* 2XMORE probe to verify the stoichiometry of the protein:DNA complex. In comparison with a simple octamer (Fig. 2C, lanes 1–3), the bands produced by the wild type, and mutant *Polr2a* probes were consistent with four and two bound Oct1 molecules, respectively (Fig. 2C, lanes 4–6,7–12). Complexes of one or three molecules were not observed. Compared with the normal (unstressed) condition (Fig. 2C, lanes 1,4), augmented Oct1 binding was observed using the 2XMORE, but not a simple octamer, following either IR or H₂O₂ treatment (Fig. 2C, cf. lanes 5,6 and lanes 2,3). Quantification of the band intensities from Figure 2C and two replicate experiments showed that the increase in Oct1-shifted DNA in the presence of either stress was progressively higher with the 1XMORE and 2XMORE probes compared with the simple octamer probe, suggesting that the stress-induced 1XMORE and 2XMORE Oct1:DNA complexes were progressively more cooperative (data not shown).

To assess complex stability, we performed kinetic complex dissociation assays using a 200-fold molar excess of cold competitor (Fig. 2D). Complexes were assembled using simple octamer, 1XMORE or 2XMORE probes, and incubated with cold competitor for varying times prior to native gel loading. Band intensities were determined and used to plot exponential decay curves. Complex half-life using a simple octamer was significantly shorter ($T_{1/2}$ = 10.8 sec) compared with the 2XMORE (80.2 sec). Using a 1XMORE, we obtained an intermediate result ($T_{1/2}$ = 13.7 sec). These data show that the 2XMORE complex is

significantly more stable than complexes formed using a simple octamer sequence or a single MORE.

We performed ChIP assays to assess Oct1 binding to the *Polr2a* promoter in vivo. We observed low-level association of Oct1 to *Polr2a* in the absence of stress, but a significant augmentation following H₂O₂ exposure (Fig. 3A). Occupancy was maximal at 60 min and decreased thereafter. In contrast, Oct1 occupancy of the control *histone H2B* promoter (which contains a single simple octamer) was unaffected by H₂O₂. Induced Oct1 *Polr2a* occupancy was also observed following IR treatment; however, in this case the kinetics were different, with *Polr2a* occupancy peaking at 60 min, but subsequently remaining more stable (Fig. 3B). In contrast to the induced binding at these complex sites, simple sites showed either reduced binding or no change. At the *H2B* promoter, Oct1 occupancy was significantly decreased at short time points following IR exposure. These results were reproducible, and we conclude that IR stress affects Oct1 binding to *Polr2a* and *H2B*, but in opposite ways. We also tested Oct1 binding to *Gadd45a*, a known Oct1 target (also containing a simple octamer sequence), finding that Oct1 occupancy was stable. Therefore, the Oct1 transcriptional response to stress is not uniform, but varies depending on the spatial context in which Oct1 associates with the DNA of different targets.

To determine whether Oct1 binding to the 2XMORE is functional, we linked the 2XMORE to an SV-40 promoter luciferase vector and used it in transfections with HeLa cells together with a cotransfected thymidine kinase promoter-linked *Renilla* luciferase internal control (Fig. 3C). A small, but significant increase in luciferase activity was observed using a 2XMORE in either orientation, and activity was decreased by MORE point mutation. Because these experiments used the 2XMORE in the context of a heterologous promoter, we also linked 600 base pairs (bp) of native *Polr2a* promoter sequence to a promoterless luciferase vector. A strong increase was observed relative to the empty vector (Fig. 3D). We engineered a 2XMORE deletion, which significantly decreased activity, although the effect was not complete. We ascribe the residual activity to other bound transcription factors and/or basal promoter activity.

To determine whether Oct1 loss of function effects native *Polr2a* gene expression under stressed conditions, we used real-time RT-PCR with RNA prepared from early-passage primary MEFs derived from littermate wild-type and Oct1-deficient embryos (V.E.H. Wang et al. 2004). The MEFs were either untreated or exposed to H₂O₂ (Fig. 3E, left panel). Three biological replicates were performed for each condition and averages are shown. *Polr2a* is a constitutive gene, and Pol II large subunit levels are thought to be regulated mainly at the protein level in response to stress stimuli. We confirmed this supposition using wild-type cells, in which only small changes in steady-state *Polr2a* mRNA levels were evident following stress (Fig. 3E, dark line). Over repeated experiments, no difference in baseline *Polr2a* expression was observed between wild-type and Oct1-deficient cells (data

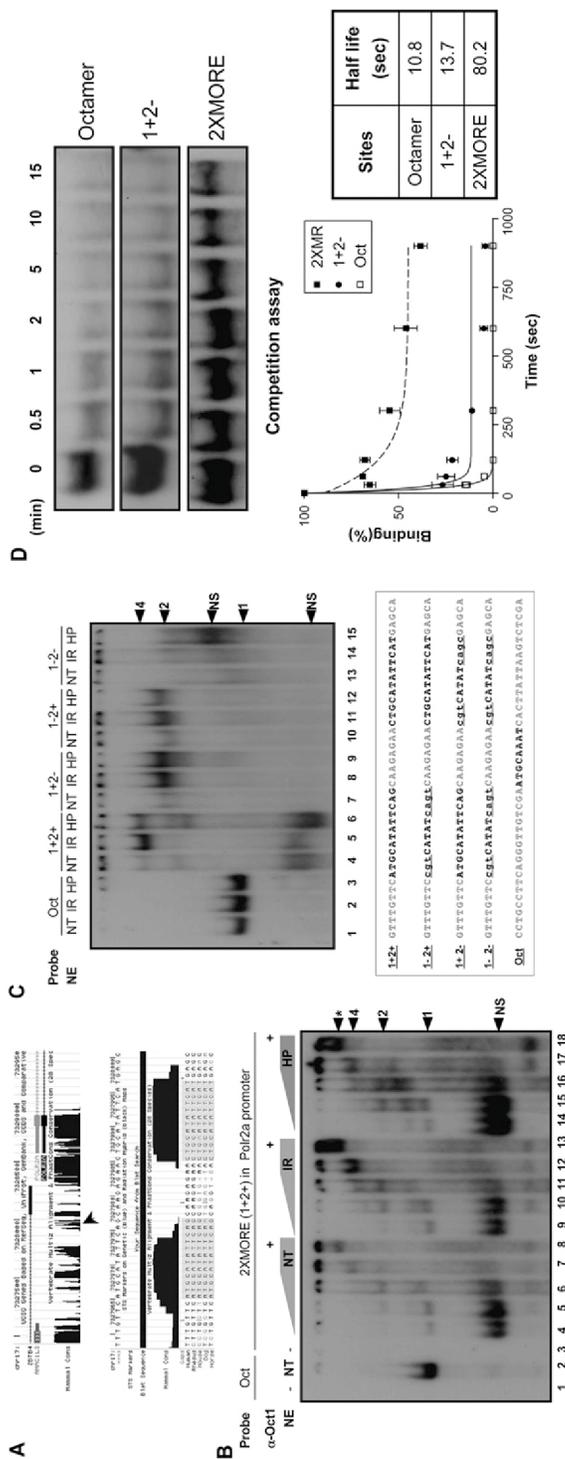


Figure 2. An endogenous 2XMORE interacts with four Oct1 molecules in a cooperative stress-induced complex. (A) UCSC Genome Browser (<http://genome.ucsc.edu>) screenshots of the human *Polr2a* locus. (Top) The region upstream of the TSS and mammalian conservation. Arrow indicates position of the 2XMORE. (Bottom) Higher-resolution screenshot showing the nucleotide sequence. Boxed sequences contain the conserved MOREs. (B) EMSA using probes designed from the sequence in A. Arrows indicate monomeric (1), dimeric (2), and tetrameric (4) occupancy. (*) Antibody supershift. The free probe was run off the gel. (C) EMSA using probes containing point mutations in one or both MOREs. A simple octamer probe of the same length was used as a control for monomeric binding. Arrows indicate monomeric (1), dimeric (2), and tetrameric (4) occupancy. (NS) Nonspecific band. Probe sequences are shown at bottom. (D) Kinetic dissociation assay of Oct1:DNA complexes assembled using simple octamer, MORE or 2XMORE probes. The *top* panels show the cropped bands from the imaged gel mobility shift. The plot shows quantification of average band intensities from three experiments together with decay curves calculated using exponential regression. Half-lives were calculated from these curves.

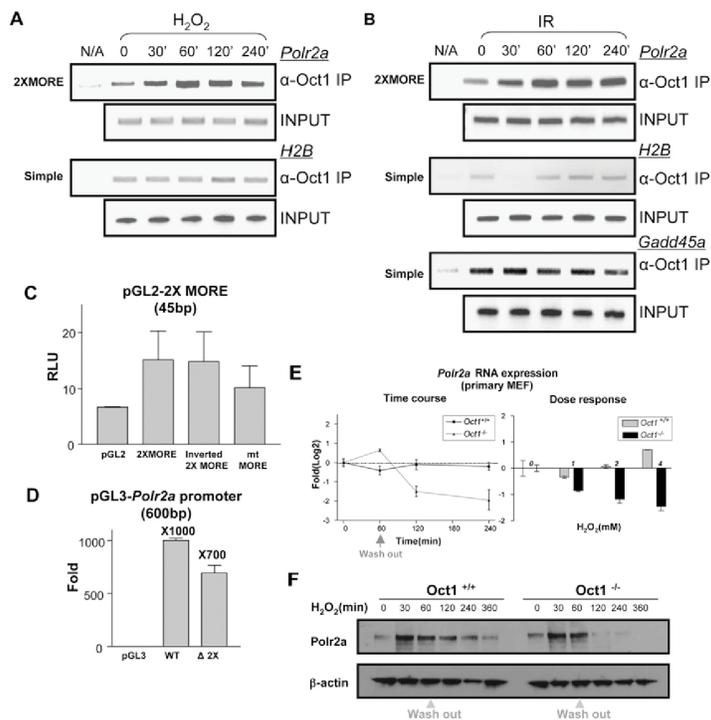


Figure 3. Oct1 functionally regulates *Polr2a*. (A) Time-course ChIP assay of *Polr2a* using HeLa cells treated with H_2O_2 . H2B and input DNA are shown as controls. (B) Similar time-course experiment using IR. Oct1 association with a simple octamer site in *Gadd45a* is additionally shown. (C) Reporter activity of a 2XMORE linked to the SV-40 promoter in a luciferase-based assay. Constructs were transfected into HeLa cells. Experiment was performed in triplicate. Error bars depict standard deviations. (D) Activity of the human *Polr2a* promoter fragment (−600 to +1 relative to the TSS) linked to luciferase. A 45-bp 2XMORE deletion was made in the context of the full sequence. Empty vector was used as a control. (E, left panel) Real-time RT-PCR using intron-spanning mouse *Polr2a* primers and either wild-type or Oct1-deficient MEFs. Message levels are shown following treatment with 2 mM H_2O_2 relative to wild-type fibroblasts under unstressed conditions on a log₂ scale. (Right panel) Dose response experiment. *Polr2a* expression was measured 4 h following exposure to different H_2O_2 doses. (F) Western blot showing time course of Pol II large subunit expression in immortalized MEFs cells treated with H_2O_2 . β-actin is shown as a loading control.

not shown). These results are consistent with our findings that expression of other reported Oct1 targets such as *histone H2B*, *U2 snRNA*, and immunoglobulin are unchanged in the Oct1-deficient and unstressed condition (V.E. Wang et al. 2004; V.E.H. Wang et al. 2004). Strikingly, in the absence of Oct1 a significant depletion in *Polr2a* mRNA levels was observed following stress exposure (gray line). The effect was greatest (approximately four-fold) at 4 h. We also performed a *Polr2a* dose response experiment using a fixed 4-h time point. Increasing doses of H_2O_2 resulted in increased *Polr2a* mRNA repression in the Oct1-deficient condition (Fig. 3E, right panel). These results strongly suggest that Oct1 acts as a functional anti-repressor of *Polr2a* following stress.

We confirmed an effect at the protein level using Western blotting and immortalized MEFs: In the Oct1-deficient condition, Pol II large subunit levels were not detectable several hours after H_2O_2 treatment, although, in this case, a transient and Oct1-independent increase in protein levels was also observed (Fig. 3F).

S335 and S385 control MORE Oct1 dimerization

Oct1 bound to MORE DNA adopts a unique structure in which the two subdomains of each molecule no longer span the helix but instead occupy adjacent major grooves on the same face of the DNA [Remenyi et al. 2001]. We used molecular modeling to identify a possible role for Oct1 DNA-binding domain modifications in stress-induced

MORE dimerization. Figure 4A shows Oct1 occupying MORE DNA, in which a phosphoserine has been modeled at position 385. This phosphoserine lies in the POU-homeodomain and is in close apposition to a lysine residue (K296) in the POU-specific domain of the same molecule (arrows; see inset), suggesting that phosphorylation at this position may structurally preorient Oct1 so as to favor MORE binding.

We used extracts from Oct1-deficient 3T3-immortalized MEFs (V.E.H. Wang et al. 2004) in which Oct1 function was restored using retroviruses encoding human wild-type Oct1, or one of three engineered three point mutations (S385A, S385D, and S385K) to determine the effect on binding in the presence and absence of IR treatment. An alanine substitution at position 385 had little effect (Fig. 4B, lanes 4,5,13,14), suggesting that other modifications may also induce MORE binding (see below). S385D, a potential phosphomimetic mutation, had a deleterious effect on monomer binding (Fig. 4B, lanes 6,7). The same mutation significantly enhanced MORE dimerization (Fig. 4B, lane 16). This mutation also favored MORE binding under unstressed conditions (Fig. 4B, lane 15). In contrast to S385D, the S385K mutation strongly inhibited MORE dimerization, instead favoring the monomeric form (Fig. 4B, lanes 17,18). Band intensity quantification showed that the S385D mutation augmented MORE dimer formation, but only partially decreased dimer stress induction, and that S385K disfavored, but did not completely eliminate MORE dimer stress induction (data not

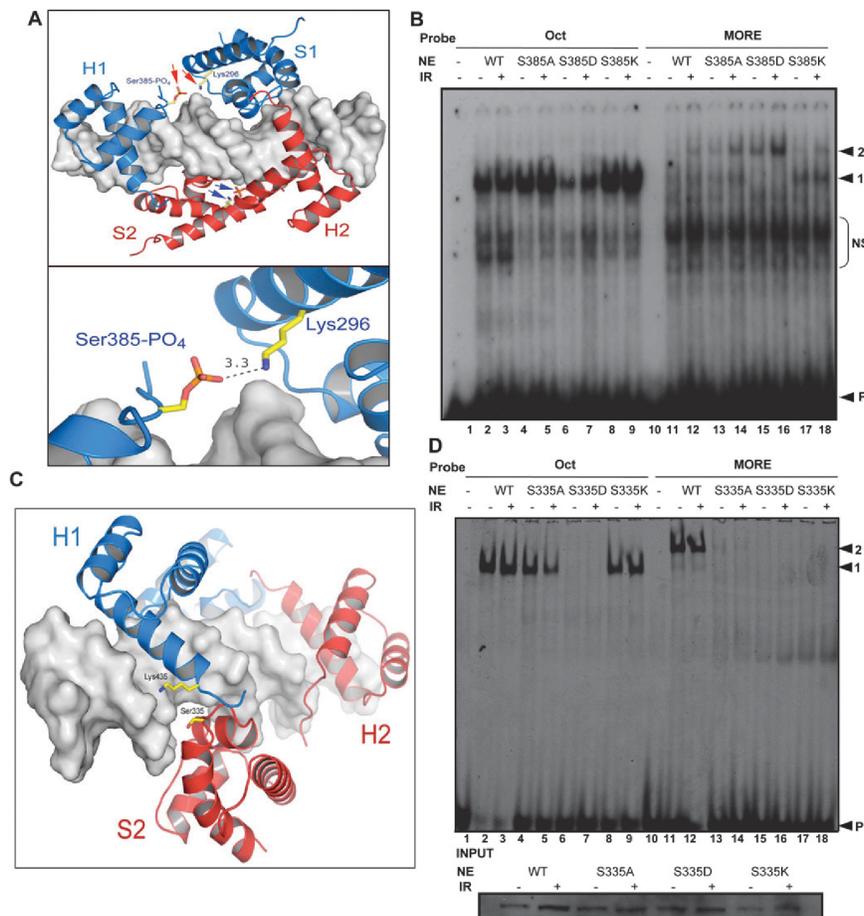


Figure 4. Oct1 S385 and S335 mutations modulate DNA selectivity. (A) Molecular models based on the crystal structure of the Oct1 POU domain dimer bound to MORE DNA (Remenyi et al. 2001). Phospho-Ser 385 and Lys 296 are modeled (red and blue arrows). Bottom image shows predicted Å distance. Images generated using PYMOL (<http://www.pymol.org>). (B) EMSA using nuclear extracts prepared from Oct1-deficient immortalized MEFs infected with retroviruses encoding wild-type, S385A, S385D, or S385K Oct1. Arrows indicate monomeric (1) and dimeric (2) occupancy, and free probe (P). (NS) Nonspecific. (C) MORE DNA-Oct1 dimer structure showing Ser 335 (red). Lys 435 is also shown. (D) EMSA using nuclear extracts prepared from Oct1-deficient immortalized MEFs infected with retroviruses containing either wild-type, S335A, S335D, or S335K Oct1 cDNAs and probes tagged with Cy5. Arrows indicate monomeric (1) and dimeric (2) occupancy, and free probe (P). The gel was scanned using a Typhoon Imaging system (Molecular Dynamics). (Below) Oct1 Western blot of the input extracts.

shown), again suggesting that other modification events can also facilitate dimerization (see below). For these experiments, low titers of Oct1 retrovirus were used, because at higher titers, or higher amounts of extract, dimer occupancy was strongly favored and the differences between the various mutants were obviated. These data are consistent with a model in which phospho-S385 makes and contacts with K296. Mutation of K296 (which is also in close apposition to the DNA backbone) (Fig. 4A) to alanine eliminated MORE binding, including when made as a double mutant with S385D (data not shown).

The other modification identified in the Oct1 DNA-binding domain was phospho-S335. This residue is located in the POU-specific domain and is positioned close to the DNA backbone. Figure 2C shows a different view of the same Oct1 dimer/MORE DNA structure with a phosphoserine modeled at position 335. Although there is a lysine (K435) located across the dimer interface in the POU-homeodomain, the intervening distance is too great to support an interaction without changing the modeled coordinates. We hypothesize that in order to form an optimal interaction, slight rearrangements of these domains take place. We engineered similar point mutations

at this position (S335A, S335D, S335K). The S335 mutants were more robust, such that higher titers of virus could be used to infect cells, or higher amounts of extract could be used in the EMSA and the phenotypes retained (Fig. 3D; see Discussion). Using high viral titers, S335A had minimal effects on simple octamer binding (Fig. 4D, lanes 4,5), but strongly inhibited MORE dimerization (Fig. 4D, lanes 13,14). S335D did not appear to bind DNA under any conditions (Fig. 4D, lanes 7,8,15,16). This serine is closer to the DNA backbone than S385 and it is possible that electrostatic repulsion disfavors binding of the S335D mutant Oct1 to DNA. Interestingly, the homologous Oct4 mutation binds a MORE *in vitro* but is not induced by stress (see below). We conclude that either (1) Oct1 S335D is not a true phosphomimetic or (2) phosphorylation of S335 inhibits DNA binding, but not all Oct1 is phosphorylated at this site, and the unphosphorylated fraction has augmented MORE dimerization activity (possibly because of additional modifications). S335K demonstrated the same selectivity effects as the alanine mutation (Fig. 4D, lanes 9,10,17,18). The mutant proteins were expressed at equivalent levels (Fig. 4D, inset).

We generated a double mutation (S335A/S385D) to determine which was dominant. S335A disrupted MORE DNA binding even in the presence of S385D (data not shown). This finding is consistent with the fact that S385 mutation could be compensated by increased amounts of Oct1 protein, whereas the S335 mutation phenotype was protein level-independent.

Control of Oct4 binding to complex sites by homologous residues

Oct1 Ser 385 and Ser 335 are conserved in Oct2 and Oct4 (Fig. 5A), and across mammalian species (data not shown). The conservation at 335 is interesting because alanine and lysine substitutions at this position do not affect DNA binding to simple octamers, suggesting that MORE dimerization is physiologically significant (see below) and is likely regulated by phosphorylation in multiple Oct proteins, including Oct4. We restored Oct protein function in Oct1-deficient fibroblasts using a retroviral construct encoding mouse Oct4 and used this model system to determine whether stress exposure also regulates Oct4 binding. Fibroblasts were exposed to IR and nuclear extracts prepared after a 1-h incubation. Oct4 expression and loading levels were confirmed by Western blot (Fig. 5B, inset). Compared with Oct1-deficient fibroblasts, in which no octamer-directed DNA-binding activity was observed, Oct4 retroviral expression resulted in binding activity (Fig. 5B, lanes 3,5). As with Oct1, augmented MORE binding was observed in IR-treated fibroblasts compared with untreated cells (Fig. 5B, lanes 7,8). We engineered an Oct4 mutation at a position equivalent to Oct1 S335 (Oct4 S186D), and determined the effect on MORE DNA binding using extracts from cells exposed to methyl methanesulfonate (MMS) or cultured under normal conditions. As with Oct1, mutation of this residue resulted in sequence-selective effects. Unlike Oct1, the aspartic acid mutation did not eliminate DNA binding,

but instead resulted in failure to generate induced MORE dimerization (Supplemental Fig. 1, lanes 7–10). There may therefore be slight differences between the proteins, such that homologous mutations at specific positions can elicit distinct effects *in vitro*.

The above experiments were performed *in vitro* with Oct4 and a synthetic MORE. To verify inducible *in vivo* Oct4 binding to sequences that regulate endogenous genes, we used a previously identified MORE (CTGCA-TATGCAT) from the *Bmp4* regulatory region (Fig. 5C, top panel; Tomilin et al. 2000). This sequence is conserved between humans and rodents (Fig. 5C, bottom panel). *Bmp4* is an Oct4 target (Sharov et al. 2008). Using ChIP, we verified an interaction between Oct4 and this genomic region and showed that Oct4 occupancy was induced by stress (Fig. 5C). The kinetics of induced Oct4 binding were similar to those observed using Oct1, with occupancy peaking at 1 h.

To identify genomic DNA sequences from known Oct4 targets, gene regions that associate with Oct4 in extracts from normal or irradiated mouse ES cells, we generated a pool of tiled oligonucleotides from common human and mouse Oct4 targets (Boyer et al. 2005; Loh et al. 2006) and immunoprecipitation to enrich for Oct4-associated oligonucleotides. Microarray analysis on the purified material identified a number of *in vitro* targets (W.G. Fairbrother, *in prep.*), including a site upstream of the *Taf12* promoter (Fig. 5D). This sequence, which is highly conserved among mammals (Supplemental Fig. 2A), interacts with Oct4 1 h following IR treatment, but not in the untreated condition or at 3 h. We confirmed the ability of both Oct4 and Oct1, which is also expressed in ES cells, to interact with *Taf12* *in vivo* using mouse ES cells and ChIP. Oct1 also binds this region in HeLa cells (Fig. 5E). Because Oct4 binding was inducible at this sequence *in vitro* using extracts from irradiated Oct1-deficient fibroblasts complemented with Oct4 retroviruses (Supplemental Fig. 2B), we performed ChIP using mouse ES cells in a time course following IR treatment. Although slight baseline binding was observed, Oct4 binding was potently induced by stress, with maximal binding at 1 h, decreasing thereafter (Fig. 5E). Oct1 also inducibly binds this sequence with similar kinetics in HeLa cells (Fig. 5E). We used EMSA to show that Oct1 also interacts with the same site *in vitro*, and that multimeric binding could be induced by stress (Supplemental Fig. 2C). The size of the shifted band was consistent with a dimer.

The *Taf12* sequence contains half-sites for both the POU_H and POU_S DNA-binding subdomains, but in a configuration distinct from a MORE or PORE (Supplemental Fig. 2D). We identified highly similar sequences in the mouse *Piwil2* (also known as *Mili*) first intron, and upstream of mouse *Foxo4*. In both cases the sequence was conserved to human (Supplemental Fig. 2C). We also identified similar sequences in conserved intergenic regions between *Zcchc16* and *Lhfp11* and between *Prr16* and *Srfbp1* (data not shown). Using ES cells and ChIP, we confirmed Oct4 binding to *Foxo4* and *Mili* (Supplemental Fig. 2E). We term this sequence a TMFORE (Taf12 Mili Foxo4 octamer-related element) (see Discussion).

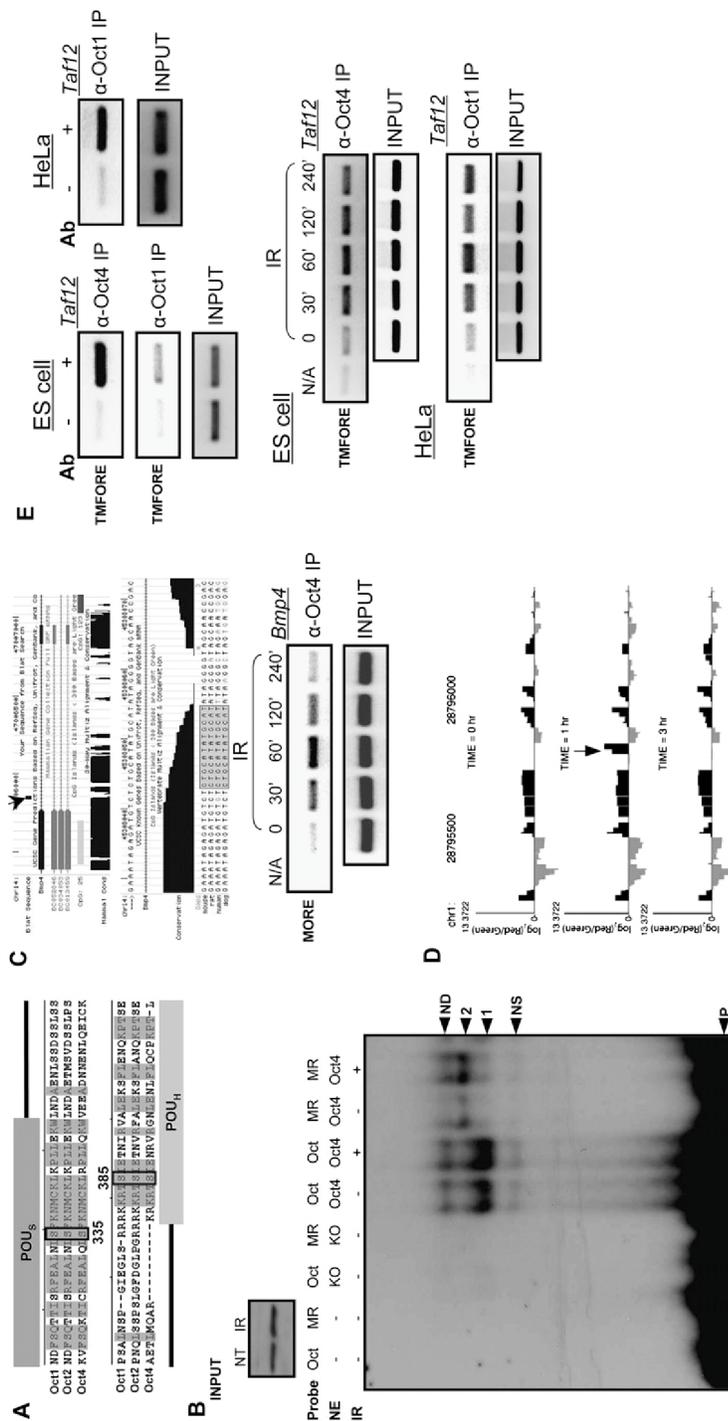


Figure 5. Oct4 complex site binding is IR-inducible. (A) Alignment of DNA-binding domain amino acid sequences of human Oct1, Oct2, and Oct4. Homologous amino acids are highlighted. Oct1 S335 and S385 and the homologous Oct2/Oct4 serines are boxed. (B) EMSA using nuclear extracts from Oct1-deficient MEFs infected with retroviruses encoding murine Oct4 or uninfected controls. (NS) Nonspecific; (ND) not determined. Probes: (Oct) Simple octamer; (MR) MORE. (Inset) Oct4 Western blot using input extracts. (C) (Top panels) Screenshots of mouse *Bmp4* locus intronic region. The MORE is marked with an arrow. Mammalian conservation is also shown. (Below) Higher-resolution image showing the MORE sequence and conservation. (Bottom panel) ChIP assay using mouse ES cells treated with IR and a primer pair spanning the *Bmp4* MORE. (D) Microarray output tracks along the human *Taf12* locus. Log₂ microarray intensities are shown for library oligonucleotides. IR-treated ES extracts were prepared at 0-, 1-, and 3-h time points. (E) (Top panels) Oct1 and Oct4 ChIP assays using mouse ES and human HeLa cells and genomic oligonucleotides encompassing the enriched *Taf12* region. (Bottom panels) ChIP time course showing Oct4 and Oct1 association with the *Taf12* site following IR exposure of mouse ES cells and HeLa cells.

High-throughput identification of Oct1 targets confirms stress-responsive binding

To identify inducible Oct1 targets on a large scale and in an unbiased manner, we used Oct1 ChIP material from normal or H₂O₂-treated HeLa cells together with deep sequencing (ChIPseq). We used 12,190,661 (input DNA), 4,167,664 (normal cells), and 3,107,208 (H₂O₂-treated cells) pipeline quality control-filtered reads with an eland alignment score >13 in the analysis. Targets are provided as Supplemental Table 1 (no treatment) and Supplemental Table 2 (H₂O₂ treatment). In static maps of untreated cells, known Oct1 targets such as genes encoding histones, U snRNAs, and GADD45a were represented with high scores. We also identified differential targets (Supplemental Table 3) using the untreated ChIP material, rather than input DNA, as the control for the H₂O₂ sample (see the Materials and Methods). The top differential target (*Ahcy*) is shown in Figure 6A. The target region mapped to the immediate 5' end of the locus and contained a conserved MORE (Fig. 6B). Additional genes are shown in Supplemental Figure 3 (see Supplemental Table 4 for annotations). *Polr2a* was identified using this analysis, but not as an induced target. We observed some Oct1 occupancy at this locus in the absence of stress (Fig. 3A,B), suggesting that the ChIPseq analysis represents an underestimate of inducible targets.

A number of the most high-scoring differentially occupied targets were previously identified as Oct4 targets in mouse ES cells (Chen et al. 2008). These included *Blcap*, *Tnks1bp1*, *Fbx110*, *Rras*, and *c9orf25* (*2310028H24Rik* in

mouse). In many cases, the Oct1- and Oct4-bound regions overlapped perfectly (Supplemental Table 5). These regions also contained highly conserved MOREs that mapped to the immediate 5' end of their respective targets (see Supplemental Fig. 3 for human and Supplemental Table 5 for mouse positions). A small number of other regions identified using ChIPseq, although not showing strong stress induction, overlapped with known Oct4 target regions. These included *Polr2a*, *Ell*, and *Zmiz2*. Conserved MOREs were also identified within these regions (Supplemental Table 5). An overlapping group of conserved MORE sequences was also identified in the top 500 differential targets that mapped within 250 bp of the nearest TSS (Fig. 6C). We conclude that Oct1 regulates a large number of genes specifically under oxidative stress conditions, a subset of which contain promoter-proximal MOREs.

To both validate the induced Oct1 ChIPseq target results from human HeLa cells, and to determine whether they can be applied to inducible Oct4 target binding in mouse ES cells, we generated primer pairs spanning the conserved equivalent regions of the mouse genome and used them in Oct4 ChIP assays following IR exposure. Oct4 binding was induced by stress at both *Blcap* and *Ell* (Fig. 6D).

Discussion

Here we show that Oct1 is dynamically phosphorylated at serine and threonine residues in vivo in response to

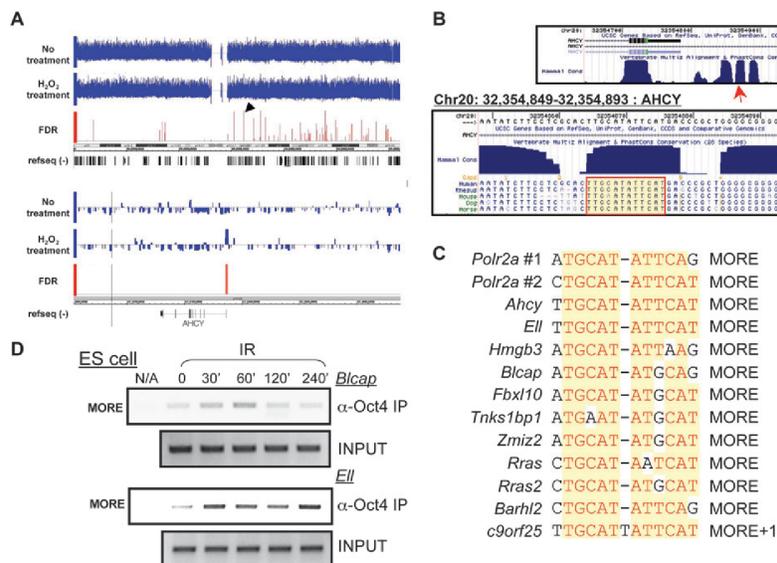


Figure 6. ChIPseq identifies inducible Oct1 targets. (A) Screenshots showing ChIPseq hits covering human chromosome 20. Arrow indicates one significant hit. (Bottom panel) Higher-magnification image showing this hit upstream of *Ahcy*. (B) Screenshots of murine *Ahcy* promoter. Sequence conservation is shown. MORE is highlighted. (C) MORE motifs in Oct1 targets identified by ChIPseq. (D) Oct4 ChIP using primer pairs spanning the homologous mouse region of a subset of inducible human Oct1 targets identified by ChIPseq, together with mouse ES cells treated with IR. Input DNA is shown as a control.

stress stimuli and that phosphorylation at two conserved DNA-binding domain serine residues regulates Oct1 binding to complex DNA configurations. Mutation of a homologous Oct4 residue also supports a role in binding site selectivity. Oct1 binding to specific complex sites (in *Polr2a* and *Taf12*) is inducible by stress *in vivo*, and Oct1 is critical for maintaining normal gene expression at one such site (*Polr2a*) following stress exposure. *In vivo* Oct4 DNA binding to several target loci in ES cells is regulated in the same manner. We also identify a new class of complex site (TMFORE) capable of binding Oct1 and Oct4 as a dimer. We use ChIPseq to confirm inducible target occupancy on a large scale, and to identify new constitutive and inducible Oct1 targets. Finally, using Oct4 ChIP in ES cells, we confirm that Oct4 inducibly interacts with two of these same targets in response to stress stimuli.

We used an affinity purification/mass spectrometry approach to purify Oct1 from HeLa cells under normal and stressed conditions and to identify modifications. Two identified stress-dependent phosphorylation events, T276 and S278, were identified previously as *in vitro* DNA-PK targets (see the supplemental material in Schild-Poulter et al. 2007). Although we did not identify other modifications described in that report, it is notable that there are few substrates for tryptic digestion in the Oct1 N-terminus, where these events were identified. Comparing our data with an unpublished data set generated by the laboratory of Dr. Steven Gygi, we noted two DNA-binding domain modification events (S335 and S385) that were not identified using our affinity approach. We hypothesized that one or both modifications affect DNA binding. *In vitro* studies previously identified the phospho-S385 modification, and it was suggested that this modification inhibited DNA-binding activity during M phase (Segil et al. 1991). Our data suggest that the effect of S385 phosphorylation is more nuanced, acting at the level of selectivity for different classes of target sites. Using S385 and S335 mutations, we confirmed a role for both residues in Oct1 target site selectivity.

The Oct1 S385D phosphomimetic mutation partially inhibits binding to simple octamer motifs in EMSA, consistent with the suggestion that S385 modification could inhibit Oct1 DNA binding (Segil et al. 1991). Because simple octamer binding in extracts from irradiated cells is not inhibited, S385D may not fully mimic serine phosphorylation at this position. Alternatively, only a fraction of Oct1 may be phosphorylated at this position and/or other modifications may simultaneously augment simple octamer binding. In contrast, S385D augmented binding to MORE sites. S385K mutation had the opposite effect: Little change in simple octamer binding was observed, but MORE dimerization was strongly inhibited (Fig. 4B). Our interpretation that phosphorylation of S385 results in an intramolecular stabilization of an orientation that favors dimer binding is consistent with the effect of the S385K, which would create an electrostatic clash with K296 in the MORE structure and disfavor dimerization (Fig. 4A). This in-

terpretation is also consistent with our findings that higher protein concentrations reduce the effects of S385 mutation (data not shown).

Oct1 S335 alanine or lysine substitution caused a major change in the DNA-binding spectrum, strongly inhibiting MORE binding, but having little effect on binding to a simple site (Fig. 4D). Unlike S385, increasing the amount of added protein did not alter the mutant phenotypes, arguing for a more direct role in dimer stabilization. An analogous Oct4 mutation also resulted in specific effects on dimer induction following exposure to the DNA-damaging agent MMS (Supplemental Fig. 1).

We identified a new class of complex recognition sequences that includes sites in *TAF12*, *Mili*, and *Foxo4* (Fig. 5D–F; Supplemental Fig. 2). We term this sequence TMFORE. It is highly likely that there are a number of additional classes of Oct protein-binding sites, each capable of adopting unique protein:DNA configurations, and each capable of responding to different spectrums of Oct protein modification.

Oct protein target genes have functions consistent with a role in the stress response. Examples include Bcl-2 (Heckman et al. 2006), p15^{INK4b} (Hitomi et al. 2007), Cyclin D1 (Magne et al. 2003), GADD45a and GADD45g (Takahashi et al. 2001; Boyer et al. 2005), C-reactive protein (Voleti and Agrawal 2005), iNOS (Kim et al. 1999), PDRG (Luo et al. 2003), and a number of interleukins (Pfeuffer et al. 1994; Duncliffe et al. 1997; Wu et al. 1997; Murayama et al. 2006). Recent studies have also identified a global association between Oct1 recognition sequences and stress response regulation. One study used low pH as a stress agent and identified selective Oct-binding site enrichment in a large group of coordinately up-regulated mRNAs (Duggan et al. 2006). Another study identified sites in >90% of genes, with significant changes in transcription rate following exposure of HeLa cells to camptothecin (Fan et al. 2006). Oct1 also appears to be a mediator of the response to heavy metals (Glahn et al. 2008). These results support our conclusion that Oct1 is a stress response effector. It will now be of importance to re-evaluate previously defined targets to determine whether the sequence composition and Oct protein binding is indicative of simple octamer-, PORE-, MORE-, or TMFORE-type configurations, or less well-defined configurations. It will also be important to define the different kinases and phosphatases that mediate the modification of both proteins and the triggers and signaling cascades involved.

Interestingly, although Oct1 binding to the *Polr2a* locus is induced following stress exposure rather than modulating gene expression in wild-type cells, Oct1 is necessary for maintenance of normal *Polr2a* expression: In the Oct1-deficient condition, message levels are inappropriately repressed. Although it is formally possible that there is a difference in message stability specifically in Oct1-deficient cells and specifically at time points following stress, the most likely interpretation is that Oct1 plays an unusual functional role as a stress-dependent *Polr2a* anti-repressor. In contrast to the effects on expression of the native *Polr2a* gene, stress exposure had little

effect in luciferase-based transient transfection assays (data not shown). It is therefore likely that in response to stress signals, Oct1 relieves repressive activities through a mechanism involving genomic chromatin context not recapitulated in transfections. Our recent results indicate that Oct1 and Oct4 control gene expression, at least in part, through control of a specific negative epigenetic mark, consistent with this hypothesis (J. Kang, A. Shakya, C. Collister, and D. Tartin, unpubl.).

Oct4 also has the capacity to respond to stress signals by selectively altering the affinity for complex binding sites *in vitro*. Further, Oct4 inducibly associates with a number of *in vivo* targets, including *Taf12*, *Bmp4*, *Bicap*, and *Ell*. Oct4 mutation at a serine residue homologous to Oct1 S335 alters binding selectivity for complex sites *in vitro*, suggesting that the stress responsiveness of Oct4 operates through modification of similar sites. The concept that Oct1 and Oct4 share common modes of regulation is not surprising, as the proteins share high identity over a substantial portion of their amino acid sequence, recognize DNA with similar specificity and affinity *in vitro*, and regulate common targets (e.g., histones, *Gadd45*, *osteopontin*, *Taf12*, *Bicap*, and *Ell*) (Botquin et al. 1998; Wang et al. 2000; Boyer et al. 2005; this study).

The most potent and well-characterized stress response pathways involve transcription factors strongly associated with cancer, such as p53 and NF- κ B. Because Oct1 and Oct4 recognize and integrate stress signals to direct changes in gene expression, a logical extension is that deregulation of these proteins may be a causal agent of malignancy. Evidence for this view comes from the findings that malignant T-cell lymphomas arise in transgenic mice overexpressing a fragment of Oct1 under the control of the T-cell-specific Lck promoter (Qin et al. 1994), and that Oct1 is overexpressed in some human breast and intestinal gastric tumors (Jin et al. 1999; Almeida et al. 2005). Oct4 is also strongly associated with hyperplasia and malignancy (Gidekel et al. 2003; Hochedlinger et al. 2005; Covello et al. 2006). One study has demonstrated a significant enrichment in Oct protein target sites in genes overexpressed in a lung cancer model, but not overexpression of Oct1 itself, suggesting that Oct1 activity, rather than levels, may regulate these processes (Reymann and Borlak 2008).

Materials and methods

Oligonucleotides

Sequences for gel-shift probes, site-directed mutagenesis, ChIP, PCR, real-time RT-PCR, and nanoparticle synthesis are presented in Supplemental Table 6.

Plasmids, cloning, and site-directed mutagenesis

pBabe-hOct1 has been described previously (V.E.H. Wang et al. 2004). Mouse Oct4 cDNA was cloned into pBabe by excising the cDNA from pGEX4T2-Oct4 (a gift of Y. Bergman) using EcoRI. The cDNA was inserted into the unique EcoRI site of pBabe. Directionality was confirmed by sequencing. Site-directed mutagenesis was performed using QuickChange (Stratagene).

The 2XMORE from the *Polr2a* gene was inserted into pGL2 (Promega) by synthesizing an oligonucleotide. pGL2 was digested with SmaI, and the oligo was ligated into the digested backbone creating a dead SmaI site. The ligation mixture was recut with SmaI and transformed into competent cells. Correct clones were confirmed by sequencing.

A 600-bp fragment of *Polr2a* was amplified using PCR primers with 5' KpnI and HindIII restriction sites. PCR products were purified, digested with KpnI and HindIII, and ligated into a pGL3 backbone from a plasmid digested with the same enzymes.

Cell culture

Oct1^{-/-} E12.5 MEFs were generated as described (V.E.H. Wang et al. 2004). MEFs and HeLa cells were cultured in Dulbecco's modified Eagle's medium (Hyclone) supplemented with 10% heat-inactivated bovine fetal serum (FBS, Hyclone), 100 U/mL penicillin, 100 μ g/mL streptomycin, 2 mM L-glutamine (Invitrogen), and 50 μ M β -mercaptoethanol (Sigma) at 37°C and 5% CO₂ in a humidified atmosphere. R1.45 ES cells (a gift from M. Capecchi) used 15% FBS and additionally contained 1 mM sodium pyruvate, 0.1 mM nonessential amino acids (Invitrogen), and 1000 U/mL LIF (Chemicon). For IR, cells were exposed to 1200 RAD using an X RAD 320 X-ray irradiator and a hard beam filter (Precision X-Ray). The flux was ~100 RAD/min. H₂O₂ was used at 1 mM except where indicated.

Oct1 purification using latex nanoparticles

HeLa cell spinner cultures were treated with IR or H₂O₂ and incubated for 1 h at 37°C. Cell pellets were resuspended in 10 mM Hepes (pH 7.9), 1.5 mM MgCl₂, 10 mM KCl, 1 mM DTT, protease inhibitors (PIs), and phosphatase inhibitors (PhIs, Roche). Following dounce homogenization, nuclei were collected by centrifugation at 15,000g. Proteins in the nuclear pellet were extracted on ice for 1 h in 20 mM Hepes (pH 7.9), 25% glycerol, 0.42 M NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA, 1 mM DTT, PIs and PhIs. Nuclear extracts were clarified by centrifugation at 15,000g and dialyzed for 4 h at 4°C in 20 mM Hepes (pH 7.9), 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 1 mM DTT, PIs and PhIs. Two milligrams of protein were mixed with 10 μ L of pelleted nanobeads in buffer A (50 mM Tris at pH 8.0, 20% glycerol, 0.5 mM EDTA, 0.1% NP-40, 1 mM DTT, PIs and PhIs) plus 100 mM NaCl. The mixture was incubated for 1 h at 4°C. Nanobeads were washed with buffer A containing 200 mM and 300 mM NaCl, respectively. Bound proteins were eluted with buffer A containing 300 mM NaCl and 30 μ M octamer oligomers for 1 h on ice.

Liquid chromatography-mass spectrometry-mass spectrometry (LC/MS/MS)

Gel slabs were excised from silver-stained polyacrylamide gels and destained using a 1:1 mixture of 30 mM Potassium Ferricyanide/100 mM Sodium Thiosulfate solution and equilibrated in water. Protein gel bands were digested overnight with trypsin (Promega), and reconstituted in 10 μ L of 5% acetonitrile (ACN) with 0.1% formic acid (FA), which was then manually injected (5 μ L) using a nano injector (Valco) onto a self-packed column (100 μ m \times 75 mm, 3- μ m particle size; Waters). An 80-min solvent gradient of 5%–85% A:B (A: 5% ACN/0.1% FA; B: 80% acetonitrile /0.1% FA) was used at 350 nL/min at 5% B for the first 3 min, followed by a linear increase to 55% B in 50 min, and finally maintained at 85% B for 5 min. LC/MS/MS data were acquired using a LTQ-FT hybrid mass spectrometer (ThermoElectron Corp.) equipped with a nanospray ionization source. Peptide

molecular masses were measured by FT-ICR. Peptide sequencing was performed by collision-induced dissociation (CID) in the linear ion trap. Peptides were assigned from MSDB or NCBI protein database searches, using the MASCOT search engine (Matrix Science). All assigned peptides (including phosphopeptides) showed MASCOT scores >20 and had mass errors <3 ppm.

EMSA

EMSA procedures paralleled those described (V.E.H. Wang et al. 2004; Tantin et al. 2008). A monoclonal anti-Oct1 antibody (Upstate Biotechnologies) was used for supershifts. Reactions contained 10 μ g (HeLa cell) or 15 μ g (mouse fibroblast) of nuclear extract.

ChIP/ChIPseq

ChIP was performed as described in Boyd and Farnham (1999) and Shang et al. (2000), except that the ChIP dilution buffer included 190 mM NaCl and Protein G nanobeads (ActivMotif) were used. HeLa or ES cells (1×10^7) were used per individual sample. Oct1 and Oct4 antibodies were purchased from Chemicon and Santa Cruz Biotechnologies, respectively.

For ChIPseq, ChIP material was subjected to high-volume sequencing using a Solexa GAI sequencer (Illumina). Libraries were prepared from ~10 ng of chromatin-precipitated DNA using the Illumina ChIPseq Sample Prep Kit. Adaptor-modified fragments ranging in size from 350 to 450 bp were selected by agarose gel electrophoresis, amplified by PCR, and validated using an Agilent Bioanalyzer. Library samples (2–6 pM) were hybridized to an activated flowcell and amplified. DNA sequence identification was performed on an Illumina Genome Analyzer I using a 26 Cycle Sequencing Kit. Reads were mapped to chromosomal locations in the March 2006 NCBI Build 36.1 human genome using the pipeline's *eland* extended aligner. Further analysis was performed using the USeq analysis package (<http://sourceforge.net/projects/useq>). Briefly, a sliding window (500 bp) binomial *P*-value was used to identify significant regions of enrichment relative to the control. *P*-values were converted to FDRs using Storey's *q*-value method, adjacent windows that exceeded an FDR of 0.01 were joined and ranked to generate a list of putative binding peaks for each condition. For analysis of differential Oct1 occupancy under stress conditions, the no treatment ChIP data was substituted for the input. Raw data are available at the NCBI Short Read archive.

Real-time RT-PCR

RNA was prepared using TRIzol (Invitrogen). Five micrograms of RNA were converted into cDNA using Superscript III and random hexamers (Invitrogen). PCR reactions used SYBR Green (Molecular Probes) to detect signals in proportion to the accumulating target amplicons and intron-spanning primer pairs. PCR results were obtained and analyzed using LightCycler 480 Real-Time PCR System (Roche).

Western blotting

Antibodies used for Western blotting were as follows: mouse anti-Oct1 (Upstate Biotechnologies), mouse anti-Oct4 and β -actin (Santa Cruz Biotechnologies), and rabbit anti-Pol II large subunit (Chemicon).

Acknowledgments

We thank S. Gygi for sharing data prior to publication. We thank D. Stillman for critical reading of the manuscript and M.

Capecchi for ES cells. We also thank D. Stillman and B. Graves for helpful technical advice and A. Shakya for providing reagents. We thank C. Nelson and K. Parsawar for mass spectrometry. We thank B. Dalley and D. Nix for high-throughput sequencing and analysis. This work was supported by a March of Dimes/Basil O'Connor Starter Scholar award, a grant from the American Cancer Society (GMC-115196), to D.T., and by NCI Cancer Center Support Grant P30CA042014.

References

- Almeida, R., Almeida, J., Shoshkes, M., Mendes, N., Mesquita, P., Silva, E., Van Seuning, I., Reis, C.A., Santos-Silva, F., and David, L. 2005. OCT-1 is over-expressed in intestinal metaplasia and intestinal gastric carcinomas and binds to, but does not transactivate, CDX2 in gastric cells. *J. Pathol.* **207**: 396–401.
- Botquin, V., Hess, H., Fuhrmann, G., Anastassiadis, C., Gross, M.K., Vriend, G., and Scholer, H.R. 1998. New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2. *Genes & Dev.* **12**: 2073–2090.
- Boyd, K.E. and Farnham, P.J. 1999. Coexamination of site-specific transcription factor binding and promoter activity in living cells. *Mol. Cell. Biol.* **19**: 8393–8399.
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**: 947–956.
- Chen, Y. and Currie, R.W. 2005. Heat shock treatment suppresses angiotensin II-induced SP-1 and AP-1 and stimulates Oct-1 DNA-binding activity in heart. *Inflamm. Res.* **54**: 338–343.
- Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**: 1106–1117.
- Covello, K.L., Kehler, J., Yu, H., Gordan, J.D., Arsham, A.M., Hu, C.J., Labosky, P.A., Simon, M.C., and Keith, B. 2006. HIF-2 α regulates Oct-4: Effects of hypoxia on stem cell function, embryonic development, and tumor growth. *Genes & Dev.* **20**: 557–570.
- Duggan, S.P., Gallagher, W.M., Fox, E.J., Abdel-Latif, M.M., Reynolds, J.V., and Kelleher, D. 2006. Low pH results in coordinate regulation of gene expression in oesophageal cells. *Carcinogenesis* **27**: 319–327.
- Duncliffe, K.N., Bert, A.G., Vadas, M.A., and Cockerill, P.N. 1997. A T cell-specific enhancer in the interleukin-3 locus is activated cooperatively by Oct and NFAT elements within a DNase I-hypersensitive site. *Immunity* **6**: 175–185.
- Fan, W., Jin, S., Tong, T., Zhao, H., Fan, F., Antinore, M.J., Rajasekaran, B., Wu, M., and Zhan, Q. 2002. BRCA1 regulates GADD45 through its interactions with the OCT-1 and CAAT motifs. *J. Biol. Chem.* **277**: 8061–8067.
- Fan, J., Zhan, M., Shen, J., Martindale, J.L., Yang, X., Kawai, T., and Gorospe, M. 2006. Enmasse nascent transcription analysis to elucidate regulatory transcription factors. *Nucleic Acids Res.* **34**: 1492–1500.
- Garrity, P.A., Chen, D., Rothenberg, E.V., and Wold, B.J. 1994. Interleukin-2 transcription is regulated in vivo at the level of coordinated binding of both constitutive and regulated factors. *Mol. Cell. Biol.* **14**: 2159–2169.
- Gidekel, S., Pizov, G., Bergman, Y., and Pikarsky, E. 2003. Oct-3/4 is a dose-dependent oncogenic fate determinant. *Cancer Cell* **4**: 361–370.

- Glahn, F., Schmidt-Heck, W., Zellmer, S., Guthke, R., Wiese, J., Golka, K., Hergenroder, R., Degen, G.H., Lehmann, T., Hermes, M., et al. 2008. Cadmium, cobalt and lead cause stress response, cell cycle deregulation and increased steroid as well as xenobiotic metabolism in primary normal human bronchial epithelial cells which is coordinated by at least nine transcription factors. *Arch. Toxicol.* **82**: 513–524. doi:10.1007/s00204-008-0331-9.
- Heckman, C.A., Duan, H., Garcia, P.B., and Boxer, L.M. 2006. Oct transcription factors mediate t(14;18) lymphoma cell survival by directly regulating bcl-2 expression. *Oncogene* **25**: 888–898.
- Herr, W., Sturm, R.A., Clerc, R.G., Corcoran, L.M., Baltimore, D., Sharp, P.A., Ingraham, H.A., Rosenfeld, M.G., Finney, M., Ruvkin, G., et al. 1988. The POU domain: A large conserved region in the mammalian pit-1, oct-1, oct-2, and *Caenorhabditis elegans* unc-86 gene products. *Genes & Dev.* **2**: 1513–1516.
- Hitomi, T., Matsuzaki, Y., Yasuda, S., Kawanaka, M., Yogosawa, S., Koyama, M., Tantin, D., and Sakai, T. 2007. Oct-1 is involved in the transcriptional repression of the p15^{INK4b} gene. *FEBS Lett.* **581**: 1087–1092.
- Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. 2005. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell* **121**: 465–477.
- Jin, T., Branch, D.R., Zhang, X., Qi, S., Youngson, B., and Goss, P.E. 1999. Examination of POU homeobox gene expression in human breast cancer cells. *Int. J. Cancer* **81**: 104–112.
- Kim, Y.M., Ko, C.B., Park, Y.P., Kim, Y.J., and Paik, S.G. 1999. Octamer motif is required for the NF- κ B-mediated induction of the inducible nitric oxide synthase gene expression in RAW 264.7 macrophages. *Mol. Cell* **9**: 99–109.
- Klemm, J.D., Rould, M.A., Aurora, R., Herr, W., and Pabo, C.O. 1994. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* **77**: 21–32.
- LeBowitz, J.H., Kobayashi, T., Staudt, L., Baltimore, D., and Sharp, P.A. 1988. Octamer-binding proteins from B or HeLa cells stimulate transcription of the immunoglobulin heavy-chain promoter in vitro. *Genes & Dev.* **2**: 1227–1237.
- Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., et al. 2006. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**: 431–440.
- Luo, X., Huang, Y., and Sheikh, M.S. 2003. Cloning and characterization of a novel gene PDRG that is differentially regulated by p53 and ultraviolet radiation. *Oncogene* **22**: 7247–7257.
- Magne, S., Caron, S., Charon, M., Rouyez, M.C., and Dusanter-Fourt, I. 2003. STAT5 and Oct-1 form a stable complex that modulates cyclin D1 expression. *Mol. Cell. Biol.* **23**: 8934–8945.
- Meighan-Mantha, R.L., Riegel, A.T., Suy, S., Harris, V., Wang, F.H., Lozano, C., Whiteside, T.L., and Kasid, U. 1999. Ionizing radiation stimulates octamer factor DNA binding activity in human carcinoma cells. *Mol. Cell. Biochem.* **199**: 209–215.
- Murayama, A., Sakura, K., Nakama, M., Yasuzawa-Tanaka, K., Fujita, E., Tateishi, Y., Wang, Y., Ushijima, T., Baba, T., Shibuya, K., et al. 2006. A specific CpG site demethylation in the human interleukin 2 gene promoter is an epigenetic memory. *EMBO J.* **25**: 1081–1092.
- Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochizuki, Y., Takizawa, N., and Yamanaka, S. 2008. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* **26**: 101–106.
- Nie, J., Sakamoto, S., Song, D., Qu, Z., Ota, K., and Taniguchi, T. 1998. Interaction of Oct-1 and automodification domain of poly(ADP-ribose) synthetase. *FEBS Lett.* **424**: 27–32.
- Nieto, L., Joseph, G., Stella, A., Henri, P., Burret-Schiltz, O., Monsarrat, B., Clottes, E., and Erard, M. 2007. Differential effects of phosphorylation on DNA binding properties of N Oct-3 are dictated by protein/DNA complex structures. *J. Mol. Biol.* **370**: 687–700.
- Nishio, K., Masaike, Y., Ikeda, M., Narimatsu, H., Gokon, N., Tsubouchi, S., Hatakeyama, M., Sakamoto, S., Hanyu, N., Sandhu, A., et al. 2008. Development of novel magnetic nano-carriers for high-performance affinity purification. *Colloids Surf. B Biointerfaces* **64**: 162–169.
- Okita, K., Ichisaka, T., and Yamanaka, S. 2007. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**: 313–317.
- Pfeuffer, I., Klein-Hessling, S., Heinfing, A., Chuvpilo, S., Escher, C., Brabletz, T., Hentsch, B., Schwarzenbach, H., Matthias, P., and Serfling, E. 1994. Octamer factors exert a dual effect on the IL-2 and IL-4 promoters. *J. Immunol.* **153**: 5572–5585.
- Qin, X.F., Luo, Y., Suh, H., Wayne, J., Misulovin, Z., Roeder, R.G., and Nussenzweig, M.C. 1994. Transformation by homeobox genes can be mediated by selective transcriptional repression. *EMBO J.* **13**: 5967–5976.
- Remenyi, A., Tomilin, A., Pohl, E., Lins, K., Philippsen, A., Reinbold, R., Scholer, H.R., and Wilmanns, M. 2001. Differential dimer activities of the transcription factor Oct-1 by DNA-induced interface swapping. *Mol. Cell* **8**: 569–580.
- Reymann, S. and Borlak, J. 2008. Transcription profiling of lung adenocarcinomas of c-myc-transgenic mice: Identification of the c-myc regulatory gene network. *BMC Syst. Biol.* **2**: 46. doi:10.1186/1752-0509-2-46.
- Ryan, A.K. and Rosenfeld, M.G. 1997. POU domain family values: Flexibility, partnerships, and developmental codes. *Genes & Dev.* **11**: 1207–1225.
- Schild-Poulter, C., Pope, L., Giffin, W., Kochan, J.C., Ngsee, J.K., Traykova-Andonova, M., and Hache, R.J. 2001. The binding of Ku antigen to homeodomain proteins promotes their phosphorylation by DNA-dependent protein kinase. *J. Biol. Chem.* **276**: 16848–16856.
- Schild-Poulter, C., Shih, A., Tantin, D., Yarymowich, N.C., Soubeyrand, S., Sharp, P.A., and Hache, R.J. 2007. DNA-PK phosphorylation sites on Oct-1 promote cell survival following DNA damage. *Oncogene* **26**: 2980–2988.
- Segil, N., Roberts, S.B., and Heintz, N. 1991. Mitotic phosphorylation of the Oct-1 homeodomain and regulation of Oct-1 DNA binding activity. *Science* **254**: 1814–1816.
- Shang, Y., Hu, X., DiRenzo, J., Lazar, M.A., and Brown, M. 2000. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* **103**: 843–852.
- Sharov, A.A., Masui, S., Sharova, L.V., Piao, Y., Aiba, K., Matoba, R., Xin, L., Niwa, H., and Ko, M.S. 2008. Identification of Pou5f1, Sox2, and Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data. *BMC Genomics* **9**: 269. doi:10.1186/2471-2164-9-269.
- Shimizu, N., Sugimoto, K., Tang, J., Nishi, T., Sato, I., Hiramoto, M., Aizawa, S., Hatakeyama, M., Ohba, R., Hatori, H., et al. 2000. High-performance affinity beads for identifying drug receptors. *Nat. Biotechnol.* **18**: 877–881.
- Sive, H.L., Heintz, N., and Roeder, R.G. 1986. Multiple sequence elements are required for maximal in vitro transcription of a human histone H2B gene. *Mol. Cell. Biol.* **6**: 3329–3340.
- Takahashi, S., Saito, S., Ohtani, N., and Sakai, T. 2001. Involvement of the Oct-1 regulatory element of the gadd45

- promoter in the p53-independent response to ultraviolet irradiation. *Cancer Res.* **61**: 1187–1195.
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**: 861–872.
- Tantin, D., Tussie-Luna, M.I., Roy, A.L., and Sharp, P.A. 2004. Regulation of immunoglobulin promoter activity by TFII-I class transcription factors. *J. Biol. Chem.* **279**: 5460–5469.
- Tantin, D., Schild-Poulter, C., Wang, V., Hache, R.J., and Sharp, P.A. 2005. The octamer binding transcription factor Oct-1 is a stress sensor. *Cancer Res.* **65**: 10750–10758.
- Tantin, D., Gemberling, M., Callister, C., and Fairbrother, W. 2008. High-throughput biochemical analysis of in-vivo location data reveals novel distinct classes of POU5F1(Oct4)/DNA complexes. *Genome Res.* **18**: 631–639.
- Tomilin, A., Remenyi, A., Lins, K., Bak, H., Leidel, S., Vriend, G., Wilmanns, M., and Scholer, H.R. 2000. Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. *Cell* **103**: 853–864.
- Ushmorov, A., Ritz, O., Hummel, M., Leithauser, F., Moller, P., Stein, H., and Wirth, T. 2004. Epigenetic silencing of the immunoglobulin heavy-chain gene in classical Hodgkin lymphoma-derived cell lines contributes to the loss of immunoglobulin expression. *Blood* **104**: 3326–3334.
- Voleti, B. and Agrawal, A. 2005. Regulation of basal and induced expression of C-reactive protein through an overlapping element for OCT-1 and NF- κ B on the proximal promoter. *J. Immunol.* **175**: 3386–3390.
- Wang, D., Yamamoto, S., Hijjiya, N., Benveniste, E.N., and Gladson, C.L. 2000. Transcriptional regulation of the human osteopontin promoter: Functional analysis and DNA-protein interactions. *Oncogene* **19**: 5801–5809.
- Wang, B., Ma, L., Tao, X., and Lipsky, P.E. 2004. Triptolide, an active component of the Chinese herbal remedy Tripterygium wilfordii Hook F, inhibits production of nitric oxide by decreasing inducible nitric oxide synthase gene transcription. *Arthritis Rheum.* **50**: 2995–3003.
- Wang, R.H., Yu, H., and Deng, C.X. 2004. A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint. *Proc. Natl. Acad. Sci.* **101**: 17108–17113.
- Wang, V.E., Tantin, D., Chen, J., and Sharp, P.A. 2004. B cell development and immunoglobulin transcription in Oct-1-deficient mice. *Proc. Natl. Acad. Sci.* **101**: 2005–2010.
- Wang, V.E.H., Schmidt, T., Chen, J., Sharp, P.A., and Tantin, D. 2004. Embryonic lethality, decreased erythropoiesis, and defective octamer-dependent promoter activation in Oct-1-deficient mice. *Mol. Cell. Biol.* **24**: 1022–1032.
- Wu, G.D., Lai, E.J., Huang, N., and Wen, X. 1997. Oct-1 and CCAAT/enhancer-binding protein (C/EBP) bind to overlapping elements within the interleukin-8 promoter. The role of Oct-1 as a transcriptional repressor. *J. Biol. Chem.* **272**: 2396–2403.
- Zhao, H., Jin, S., Fan, F., Fan, W., Tong, T., and Zhan, Q. 2000. Activation of the transcription factor Oct-1 in response to DNA damage. *Cancer Res.* **60**: 6276–6280.
- Zheng, L., Roeder, R.G., and Luo, Y. 2003. S phase activation of the histone H2B promoter by OCA-S, a coactivator complex that contains GAPDH as a key component. *Cell* **114**: 255–266.

CHAPTER 3

DYNAMIC REGULATION OF OCT1 DURING MITOSIS BY PHOSPHORYLATION AND UBIQUITINATION

Short title: Oct1 in Mitosis

Jinsuk Kang¹, Ben Goodman², Yixian Zheng², and Dean Tantin^{1,*}

1 Department of Pathology, University of Utah School of Medicine, Salt Lake City, Utah, United States of America, **2** Carnegie Institution of Washington/HHMI, Department of Embryology, Baltimore, Maryland, United States of America

[*Keywords:* Oct1; mitosis; centrosome; spindle pole body; midbody; lamin B; anaphase promoting complex]

* E-mail: Dean Tantin (dean.tantin@path.utah.edu)

Funding: This work was supported by grants from the American Cancer Society (GMC-115196) and National Cancer Institute (1R21CA141009) to D.T.

Financial disclosure: The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Background: Transcription factor Oct1 regulates multiple cellular processes. It is known to be phosphorylated during the cell cycle and by stress. However the upstream kinases and downstream consequences are not well understood. One of these modified

forms, phosphorylated at S335, lacks the ability to bind DNA. Other modification states besides phosphorylation have not been described.

Methodology/Principle Findings: We show that Oct1 is phosphorylated at S335 in the Oct1 DNA binding domain during M-phase by the NIMA-related kinase Nek6. Phospho-Oct1 is also ubiquitinated. Phosphorylation excludes Oct1 from mitotic chromatin. Instead, Oct1^{pS335} concentrates at centrosomes, mitotic spindle poles, kinetochores and the midbody. Oct1 siRNA knockdown diminishes the signal at these locations. Both Oct1 ablation and overexpression result in abnormal mitoses. S335 is important for the overexpression phenotype, implicating this residue in mitotic regulation. Oct1 depletion causes defects in spindle morphogenesis in *Xenopus* egg extracts, establishing a mitosis-specific function of Oct1. Oct1 colocalizes with lamin B1 at the spindle poles and midbody. At the midbody, both proteins are mutually required to correctly localize the other. We show that phospho-Oct1 is modified late in mitosis by non-canonical K11-linked polyubiquitin chains. Ubiquitination requires the anaphase-promoting complex, and we further show that the anaphase-promoting complex large subunit APC1 and Oct1^{pS335} interact.

Conclusions/Significance: These findings reveal mechanistic coupling between Oct1 phosphorylation and ubiquitination during mitotic progression, and a role for Oct1 in mitosis.

Introduction

The Oct1 (POU2F1) transcription factor is a potent regulator of metabolism and tumorigenicity (1). It is widely expressed (2, 3) and interacts with a number of proteins

including poly (ADP-ribose) polymerase-1 (PARP-1), an enzyme that becomes activated by DNA damage and oxidative stress (4), BRCA1, a tumor suppressor protein associated with the DNA damage response (5, 6), and lamin B, a component of the nuclear and spindle matrices (7-9). Oct1 is also a signal integrator that is phosphorylated at multiple residues during the cell cycle and in response to genotoxic and oxidative stress (10, 11). Some of these phosphorylation events alter Oct1 DNA binding selectivity, resulting in altered target gene occupancy (10). Other phosphorylation events have not been carefully studied.

One of the aforementioned phosphorylation events occurs at Ser335 within the DNA binding domain. Ser335 mutation to aspartic acid blocks Oct1 binding to all tested DNA recognition sites (10), and has been associated with mitosis in mass screens (12-14). Little is otherwise known about the function of this modification. Here, we identify a previously unknown role for this form of Oct1. Consistent with the effects of the S335 phosphorylation on Oct1 ability to bind DNA, we find that phosphorylation excludes Oct1 from mitotic chromosomes. Phospho-S335 Oct1 accumulates on centrosomes, spindle pole bodies and kinetochores, with enrichment lost at the anaphase-telophase transition. Late in mitosis the remaining phosphorylated Oct1 is modified by non-canonical K11-linked polyubiquitin chains and colocalizes with lamin B at the midbody. We show that the phosphorylated form of Oct1 interacts with lamin B, and that RNAi knockdown of either Oct1 or lamin B1 in HeLa cells eliminates the midbody localization of the other protein. We implicate the anaphase-promoting complex (APC) in Oct1 ubiquitination. Oct1 RNAi in HeLa cells strongly reduces antibody localization to centrosomes, spindle pole bodies and the midbody, and results in mitotic abnormalities.

Overexpression of wild type Oct1 also disrupts mitoses, resulting in improper chromosome condensation, multinucleated cells and micronuclei. Overexpressed S335A mutant Oct1 does not disrupt mitosis to the same extent, implicating this residue in Oct1 regulation of mitotic functions.

Results

Phosphorylation of Oct1 at serine 335 during mitosis

To study the regulation and function of Oct1 phosphorylation at serine 335 (Oct1^{pS335}), we generated a phospho-specific polyclonal antibody. The peptide sequence used to generate the antibody (EALNLS₃₃₅FKNMC) aligns perfectly to the POU-specific portion of the DNA binding domain of human Oct1, Oct2 and Oct11. This region is less conserved in other human POU domain proteins (not shown). Initial characterization of the antibody using HeLa whole cell extracts and Western blotting indicated the presence of an intense band of high apparent molecular weight (>200 kDa) in cells arrested in mitosis by nocodazole (Figure 3.1A, lane 2, asterisk). A band was also present in mitotic cells corresponding to the expected unmodified molecular weight of ~90 kDa (black arrowhead), as were intermediate forms with apparent molecular weights of ~180 kDa and ~130 kDa (black dot and red arrow). A pan-Oct1 antibody (recognizing the C-terminus) identified the same forms, albeit at different relative intensities, with the two largest forms only found in mitotic cells (Figure 3.1A, lane 4). This result suggests that the four observed species are different forms of Oct1. No augmentation in Western blot signal was observed using HeLa cells arrested in S-phase with hydroxyurea (not shown), suggesting that the effects are specific to mitosis.

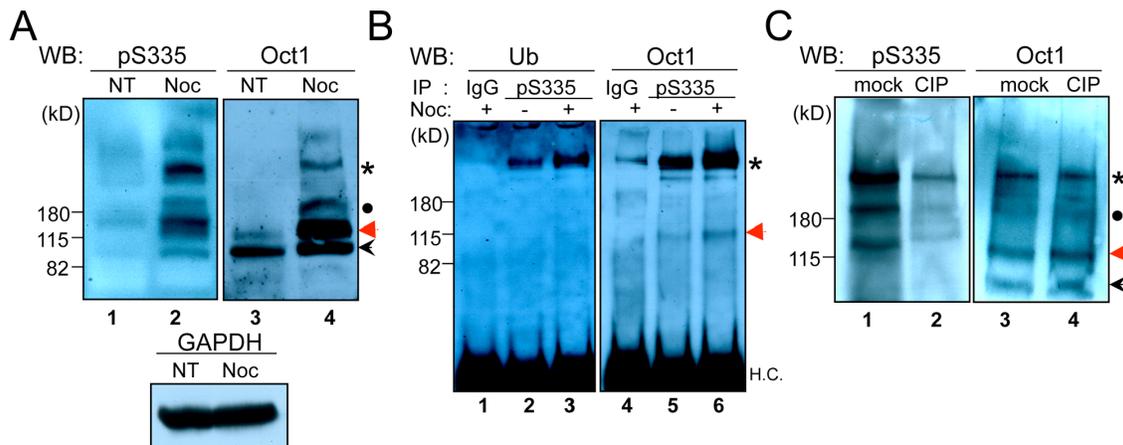


Figure 3.1. Oct1^{pS335} is enriched in M-phase HeLa cells. (A) Whole cell extracts were prepared from normal or nocodazole-arrested HeLa cells. 10% polyacrylamide gels were Western blotted using anti-Oct1^{pS335} or anti-pan-Oct1 (C-terminal) antibodies (Bethyl). Anti-GAPDH is shown as a loading control. (B) HeLa cell whole cell extracts were immunoprecipitated with anti-Oct1^{pS335} antibodies and Western blotted using pan-Ub or pan-Oct1 antibodies. H.C.=immunoglobulin heavy chain. (C) Whole cell extracts from nocodazole-arrested HeLa cells were treated with calf intestinal alkaline phosphatase (CIP), or mock-treated. Western blots using anti-Oct1^{pS335} or anti-pan-Oct1 antibodies are shown.

We hypothesized that one or more of the high molecular weight Oct1^{pS335} bands represented ubiquitinated species. To test this possibility, we immunoprecipitated Oct1 using the phospho-specific antibody and performed Western blots for both ubiquitin (Ub) and total Oct1. A band of the same size corresponding to the high molecular weight form was observed in both cases (Figure 3.1B, lanes 2 and 5, asterisk). The band was further enriched using extracts from nocodazole arrested cells (lanes 3 and 6), indicating the presence of an Oct1 population enriched in mitotic cells that is simultaneously phosphorylated at S335 and ubiquitinated. A similar result was obtained using denaturing conditions, indicating that the phospho-specific antibody is not co-precipitating a ubiquitinated protein of precisely the same apparent molecular weight, but rather recognizes a phosphorylated, ubiquitinated form of Oct1. These experiments also revealed that the high molecular weight form of Oct1 consists of a ladder of bands (Figure 3.2A).

To demonstrate that the form of Oct1 recognized by the antibody is phosphorylated, we treated nocodazole-arrested HeLa cell extracts with calf intestinal alkaline phosphatase (CIP), which resulted in the strong diminution of the high molecular weight form, as well as other Oct1 forms, but had no effect on total Oct1 (Figure 3.1C). No increase in band mobility was observed with the pan-Oct1 antibody following CIP treatment, consistent with the finding that ubiquitination is also present on Oct1^{pS335}. Similar loss of phospho-specific antibody signal with CIP treatment was also observed using indirect immunofluorescence (IF) assays (Figure 3.2B).

We examined phosphorylated Oct1 in HeLa cells using IF and the anti-Oct1^{pS335} antibody. HeLa mitoses were staged using DAPI and anti- α -tubulin. Oct1^{pS335} staining in

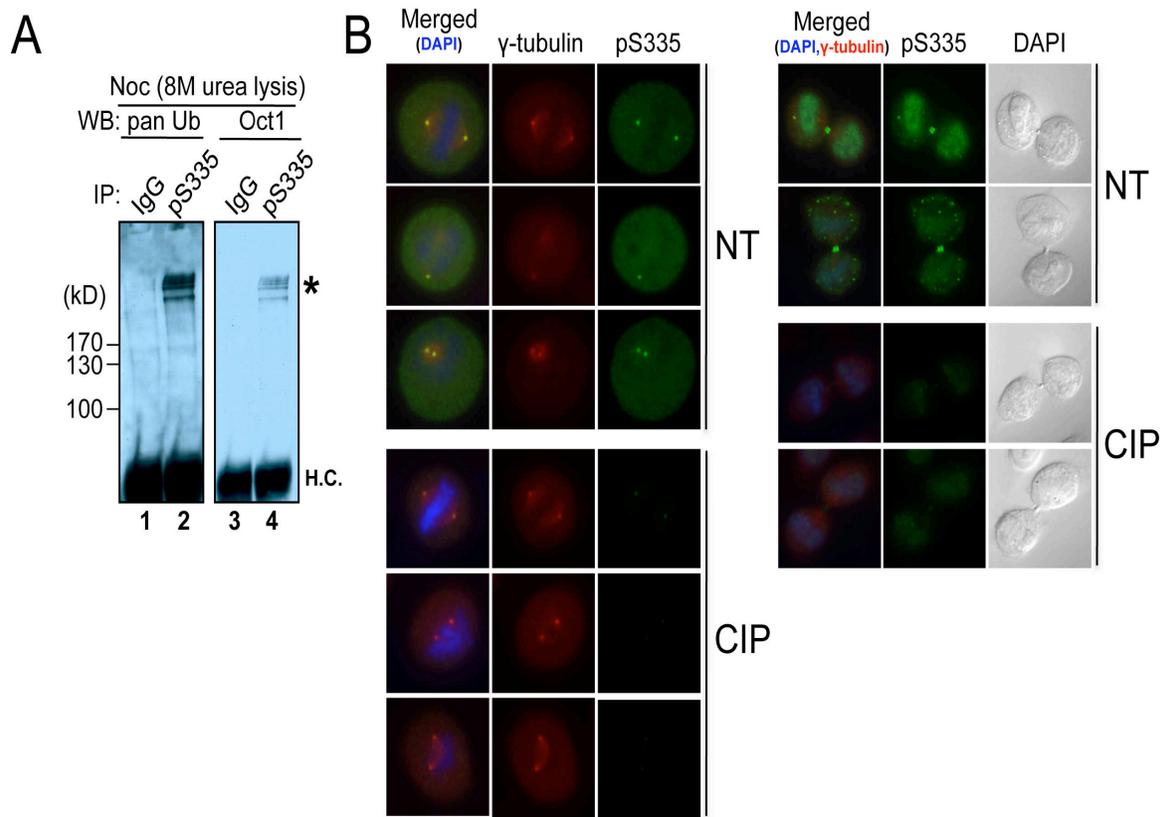


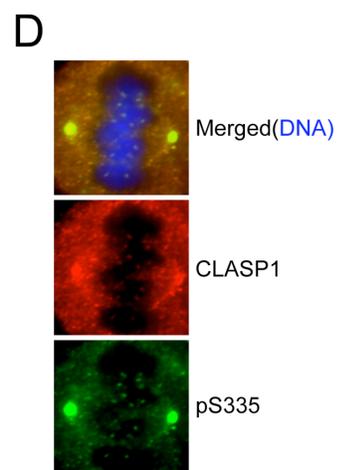
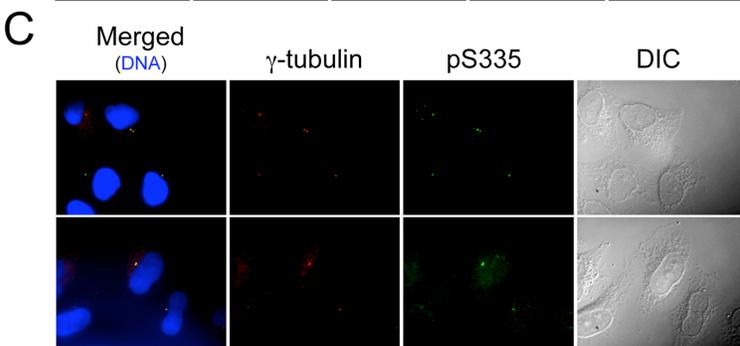
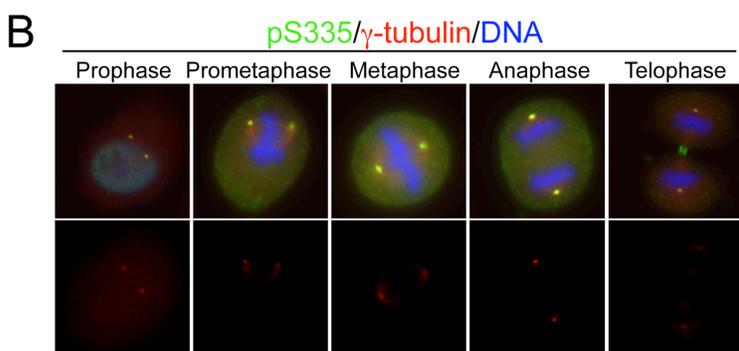
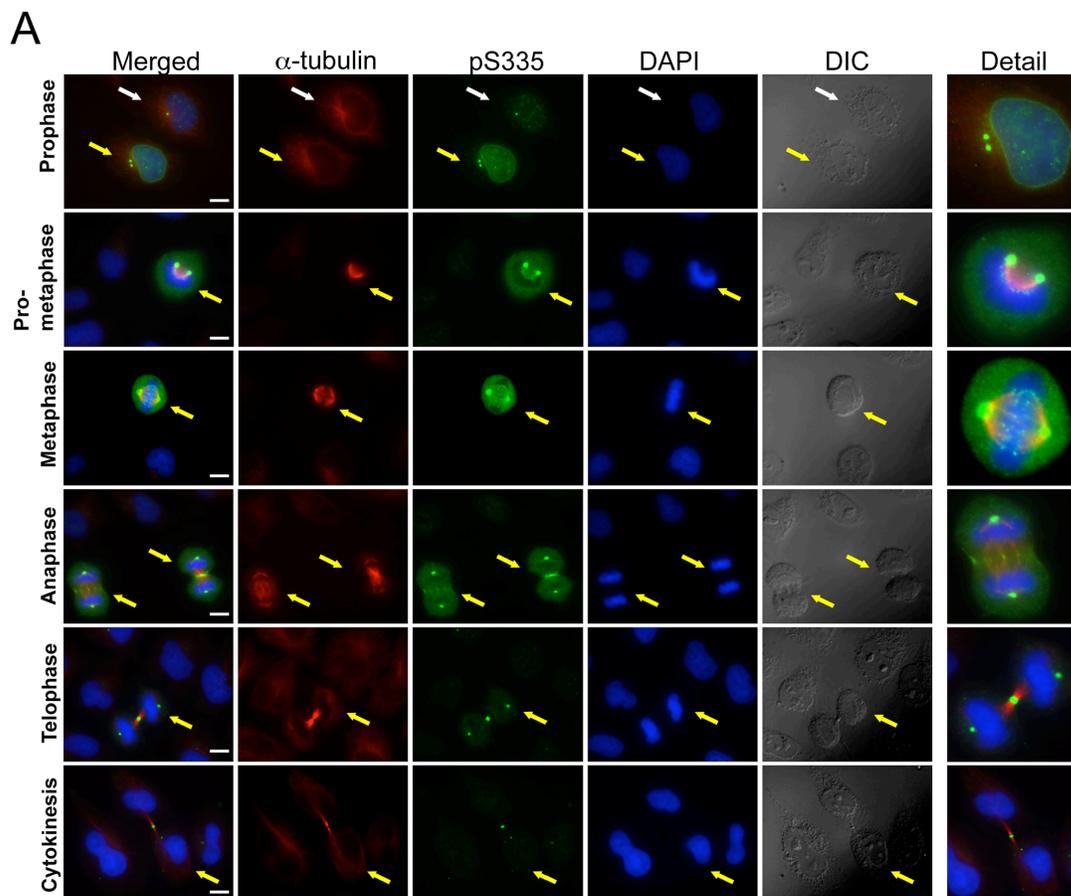
Figure 3.2. Further characterization of the phospho-specific antibody. (A) The phospho-specific antibody immunoprecipitates ubiquitinated Oct1 under denaturing conditions. IP was performed as described in the materials and methods, except cells were lysed in the identical buffer but with 8 M urea. Upon dilution the resulting bead incubation step contained 1.6 M urea in buffer. (B) Immunofluorescence images are shown. Left and right panels show early and late mitosis, respectively. NT=no treatment. METHODS: samples were prepared as described in the materials and methods, except fixed samples were treated with CIP in CSK buffer at 30°C for 30 minutes.

interphase cells (white arrows) was largely confined to individual puncta, the most intense of which correspond to the site of microtubule nucleation. Increased Oct1^{pS335} signal within an intact nuclear envelope was noted in prophase (yellow arrows). Early mitotic cells also showed staining at two puncta suggestive of duplicated centrosomes. Oct1^{pS335} was enriched at the prophase nuclear envelope (see prophase detail). Work of others has associated Oct1 with the nuclear periphery (7, 8, 15). At metaphase, the chromosomes become fully condensed, and Oct1^{pS335} was largely excluded from DNA, except at small puncta consistent with kinetochores. We corroborated spindle pole body and centrosome localization in mitotic and interphase cells using γ -tubulin antibodies (Figure 3.3B and C), and kinetochore localization using CLASP1 antibodies (Figure 3.3D).

Following chromatid separation, Oct1^{pS335} becomes concentrated at the developing midbody (Figure 3.3A). At the transition from late anaphase to telophase, detectable Oct1^{pS335} was greatly diminished with the exception of the midbody, where a concentrated signal was retained throughout cytokinesis. No such change was detected using pan-Oct1 antibodies (Figure 3.4), indicating that the diminution of the phosphorylated form at the anaphase-telophase transition is not the result of changes in total Oct1. IF using pan-Oct1 antibodies showed similar concentrations at spindle pole bodies and the midbody (Figure 3.4). Similar mitotic staining patterns were also obtained using A549 lung adenocarcinoma cells (Figure 3.5), indicating that the pattern is not peculiar to HeLa cells. A comparative analysis of Oct1^{pS335} and the well-established mitotic marker histone H3^{pS10} (16) indicated that all cells that stained for histone H3^{p10} also stained strongly for Oct1^{pS335} (Figure 3.6).

Figure 3.3. Mitotic Oct1^{pS335} is associated with the spindle pole bodies and midbody.

(A) IF images of mitotic HeLa cells are shown. Cells were stained with anti- α -tubulin and anti-Oct1^{pS335} antibodies, and with DAPI. Scale bar: 20 μ m. White arrows show an interphase cell. Yellow arrows show different mitotic stages. (B) Similar images, except γ -tubulin antibodies, were used. Cropped images of individual mitotic cells are shown. (C) Similar images of interphase cells. (D) Detail of a metaphase HeLa cell IF image stained with CLASP-1 and Oct1^{pS335} antibodies.



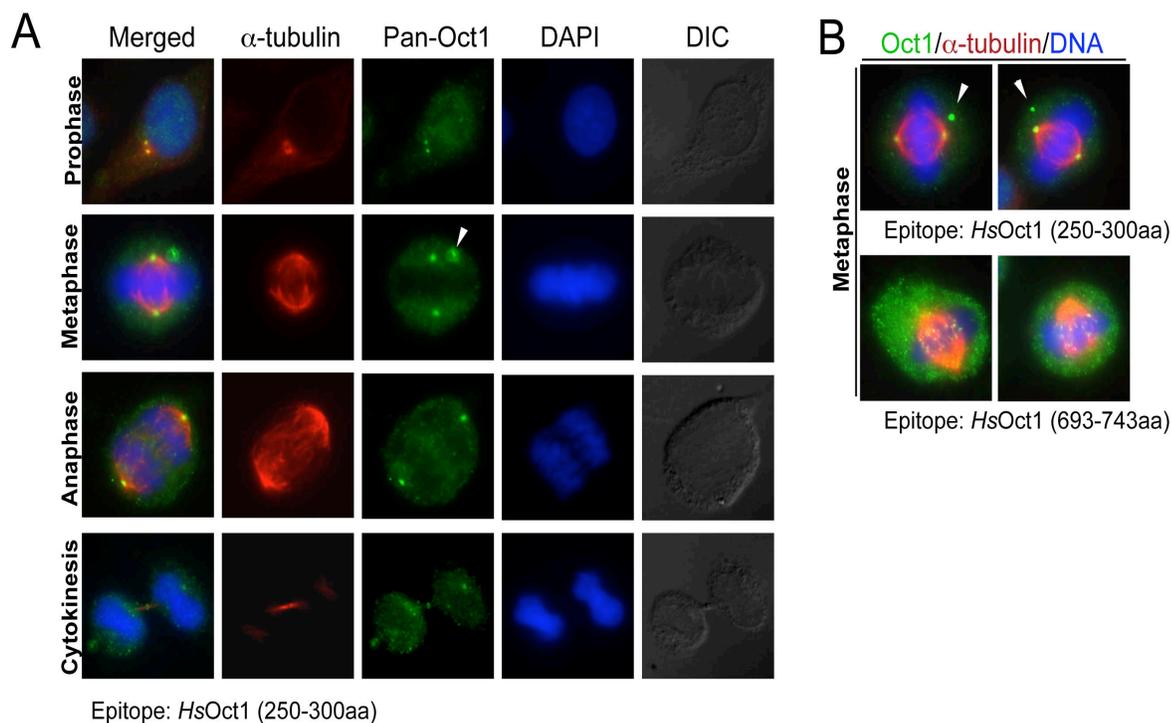


Figure 3.4. Localization of total Oct1 in mitotic HeLa cells. (A) Total Oct1 concentrates at spindle pole bodies and the midbody, and is excluded from mitotic chromosomes. Immunofluorescence images are shown. HeLa cells were fixed and stained with antibodies against α -tubulin and pan-Oct1. Arrow indicates position of additional puncta not recognized by phospho-Oct1 antibodies. (B) Pan-Oct1 antibodies recognizing different epitopes stain different mitotic structures. Cells were fixed and stained as in (A). Arrows indicate additional metaphase puncta of unknown etiology not recognized by phospho-Oct1 antibodies.

Figure 3.5. Localization of Oct1 phosphorylated at S335 to mitotic A549 lung adenocarcinoma cells. (A) Immunofluorescence images are shown. Specific mitotic cell is highlighted with an arrow. The cells were co-stained with antibodies to alpha-tubulin and with DAPI. (B) Similar images using anti-gamma-tubulin antibodies. Protocols were identical to those summarized in the methods section.

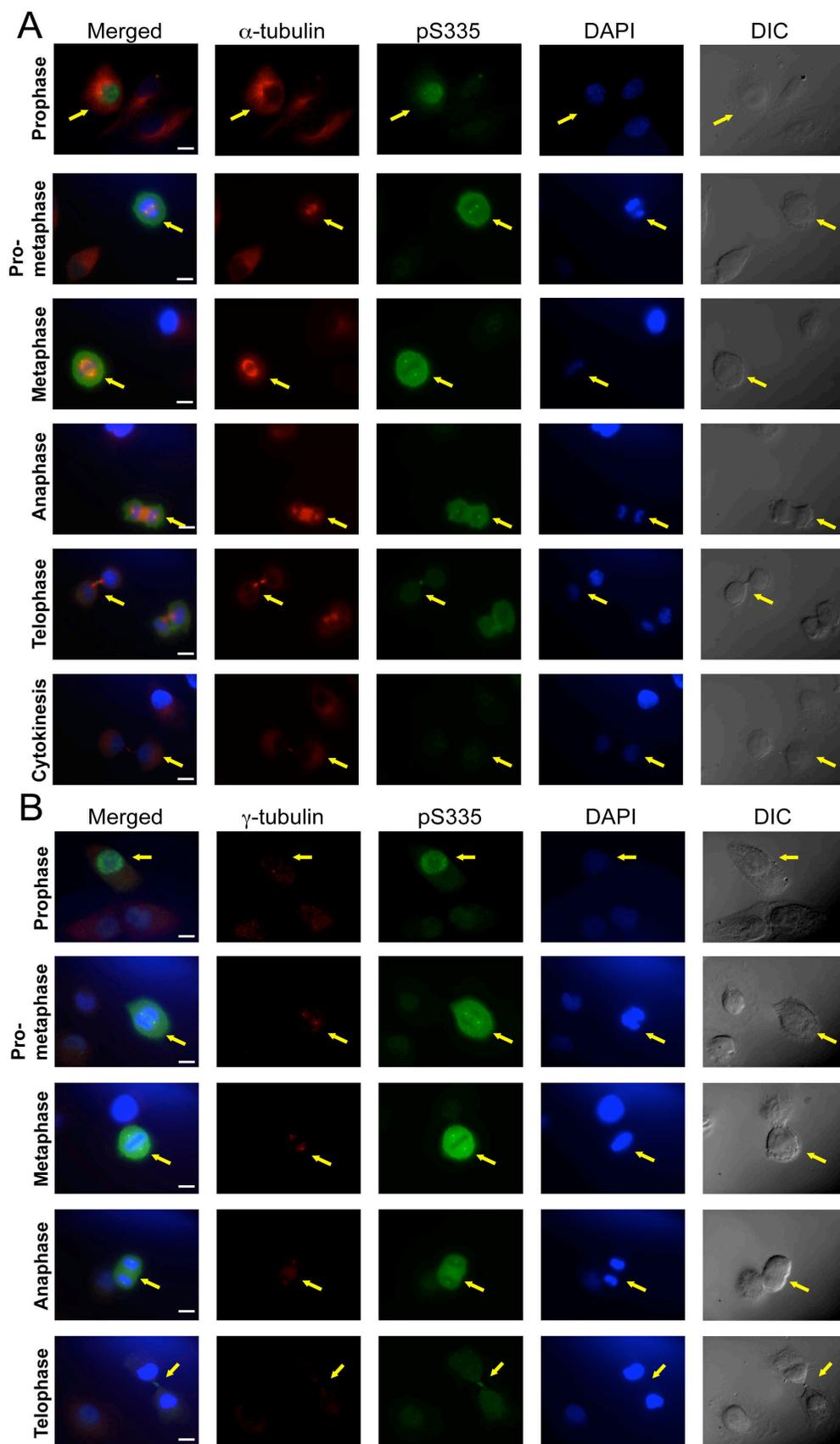
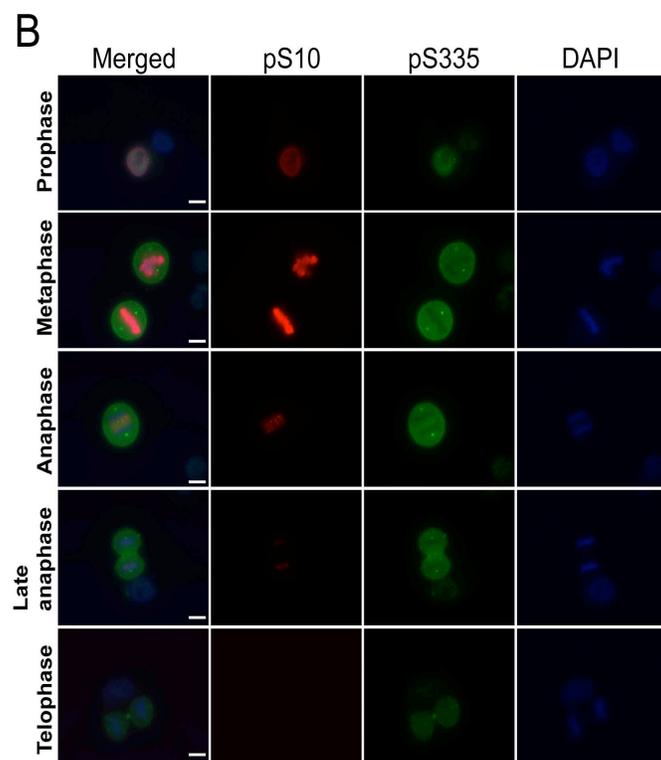
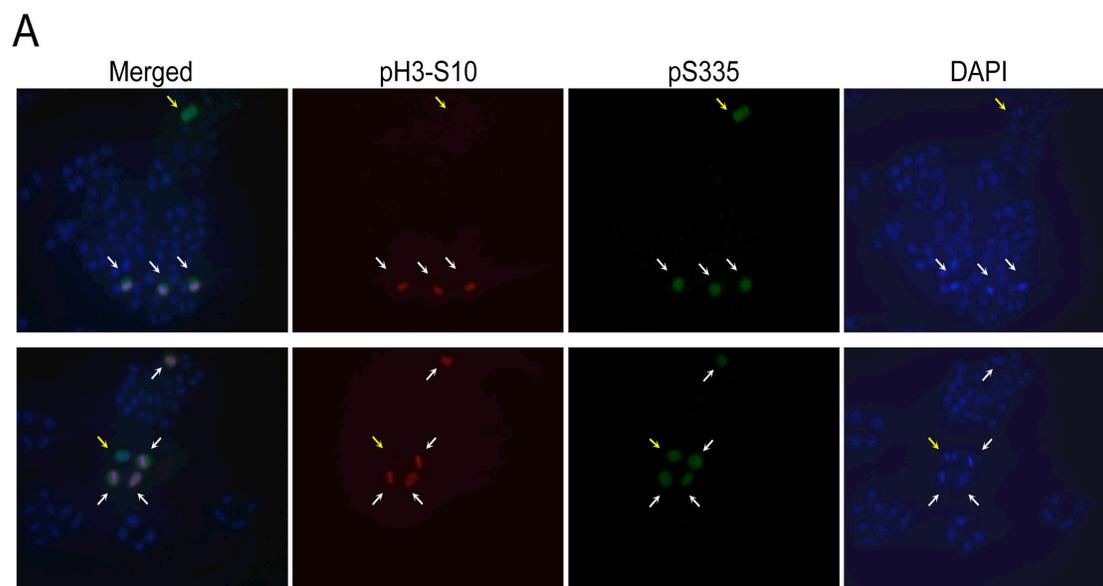


Figure 3.6. Comparison of the properties of anti-Oct1^{p335} and anti-Histone H3^{S10} antibodies. (A) Low-magnification (200X) immunofluorescence images of HeLa cells stained with antibodies against phospho-histone H3 and phospho-Oct1.^{p335} Arrows indicate mitotic cells. White arrows show cells in earlier mitotic stages that co-stain with both antibodies. Yellow arrows show examples of cells staining only with the phospho-335 and not phospho-histone antibody. (B) Immunofluorescence images of mitotic HeLa cells stained with anti-Oct1^{p335} and anti-histone H3^{S10} antibodies. Scale bar: 20 μ M.



Oct1 ablation is associated with abnormal mitoses

We transiently transfected HeLa cells with Oct1 siRNAs, analyzing cells 48 hr post-transfection to determine whether the staining patterns were specifically due to Oct1. Using Oct1-specific but not control siRNAs, we observed mitotic cells lacking or with significantly reduced Oct1^{PS335} spindle pole body/midbody staining (Figure 3.7A). In addition, nearly all (>90%) of the mitotic cells with decreased Oct1 staining lost the normal pattern of α -tubulin staining (Figure 3.7A). We confirmed the effect of transiently transfected Oct1-specific siRNAs by Western blot (Figure 3.7B). We also noted poor/abnormal chromosome segregation and other mitotic irregularities associated with partial and complete Oct1 knockdown (Figure 3.7A and Figure 3.8). To quantify these irregularities, we analyzed 668 total control siRNA and 164 total Oct1-specific siRNA mitoses over three separate experiments. There were fewer mitotic events in the Oct1-specific siRNA condition and more dead/floating cells, suggesting that following acute reductions in Oct1, a higher fraction of mitoses result in apoptosis and/or mitotic catastrophe. We scored approximately 10% of live control mitotic events as abnormal, based on spindle disorganization and chromosomal abnormalities (e.g., incomplete condensation, DNA outside the metaphase plate). Using the same criteria, approximately 60% of the Oct1 knockdown events were abnormal (Figure 3.7C). These findings indicate that Oct1 contributes to accurate mitosis in HeLa cells.

We also examined mitoses in Oct1 deficient primary early-passage fibroblasts. Oct1 protein and activity is undetectable in these cells in Western blotting and gel mobility shift assays using nuclear extracts, although a small amount of residual protein is

Figure 3.7. Modulation of Oct1 levels results in abnormal mitoses in HeLa cells. (A)

IF images are shown of HeLa cells transiently transfected with control or Oct1-specific siRNAs. Arrows indicate positions of disrupted spindle pole/midbody localization.

Formaldehyde fixation was used. **(B)** Western blot showing effect of transfected control and Oct1-specific siRNA on total Oct1 expression in cycling cells. Cells were cultured for 72 hr prior to analysis. **(C)** Quantification of abnormal in HeLa cells treated with control and Oct1-specific siRNAs. Values represent averages from three independent experiments. Error bars depict standard deviations. **(D)** IF images of HeLa cells

transiently transfected with FLAG-Oct1. Transfected cells were incubated for 48 hr prior to formaldehyde fixation and staining with anti- α -tubulin and anti-FLAG antibodies.

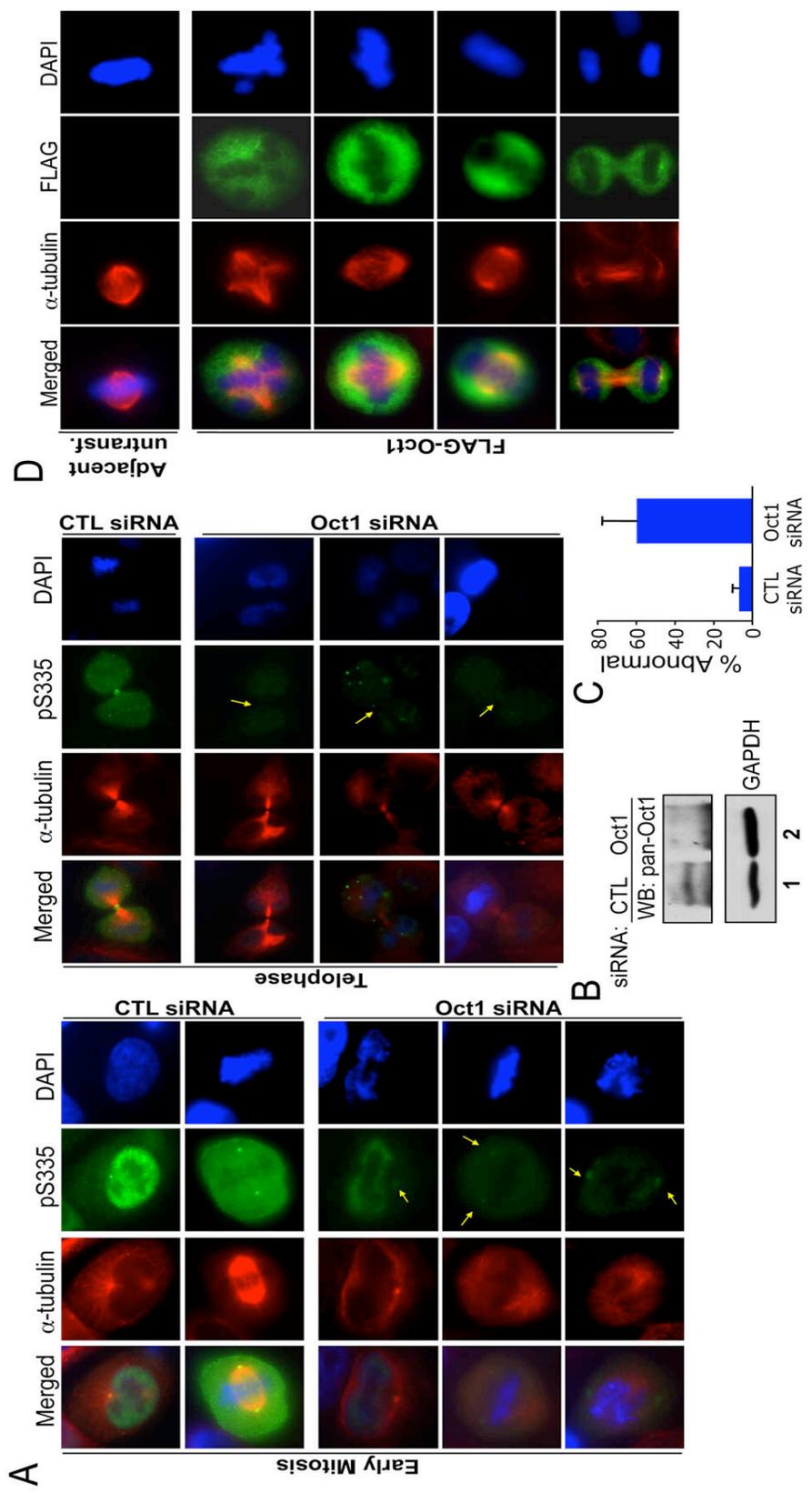
Single mitotic cell images are shown. **(E)** IF images are shown of interphase HeLa cells transiently transfected with FLAG-tagged wild-type Oct1. Cells were stained with DAPI, and anti-FLAG and anti-lamin B (B1+B2) antibodies. Arrows indicate transfected cells. Asterisk indicates an area of specific Oct1 and lamin B co-localization. Example cells showing multinucleated cells and micronuclei are shown. Scale bars indicate 20 μ m. **(F)**

Quantification of the frequency of micronuclei and multinucleated cells in mock transfected, wild-type Oct1 transfected, and S335A transfected interphase HeLa cells. Error bars depict standard deviations. **(G)** Similar experiment as (E) using cells

transiently transfected with an Oct1 S335A point mutant made using site-directed mutagenesis of the human cDNA. Top panels show mitotic HeLa cells stained with anti-FLAG and anti- α -tubulin antibodies. Bottom panels substituted lamin B antibodies.

transiently transfected with an Oct1 S335A point mutant made using site-directed mutagenesis of the human cDNA. Top panels show mitotic HeLa cells stained with anti-FLAG and anti- α -tubulin antibodies. Bottom panels substituted lamin B antibodies.

transiently transfected with an Oct1 S335A point mutant made using site-directed mutagenesis of the human cDNA. Top panels show mitotic HeLa cells stained with anti-FLAG and anti- α -tubulin antibodies. Bottom panels substituted lamin B antibodies.



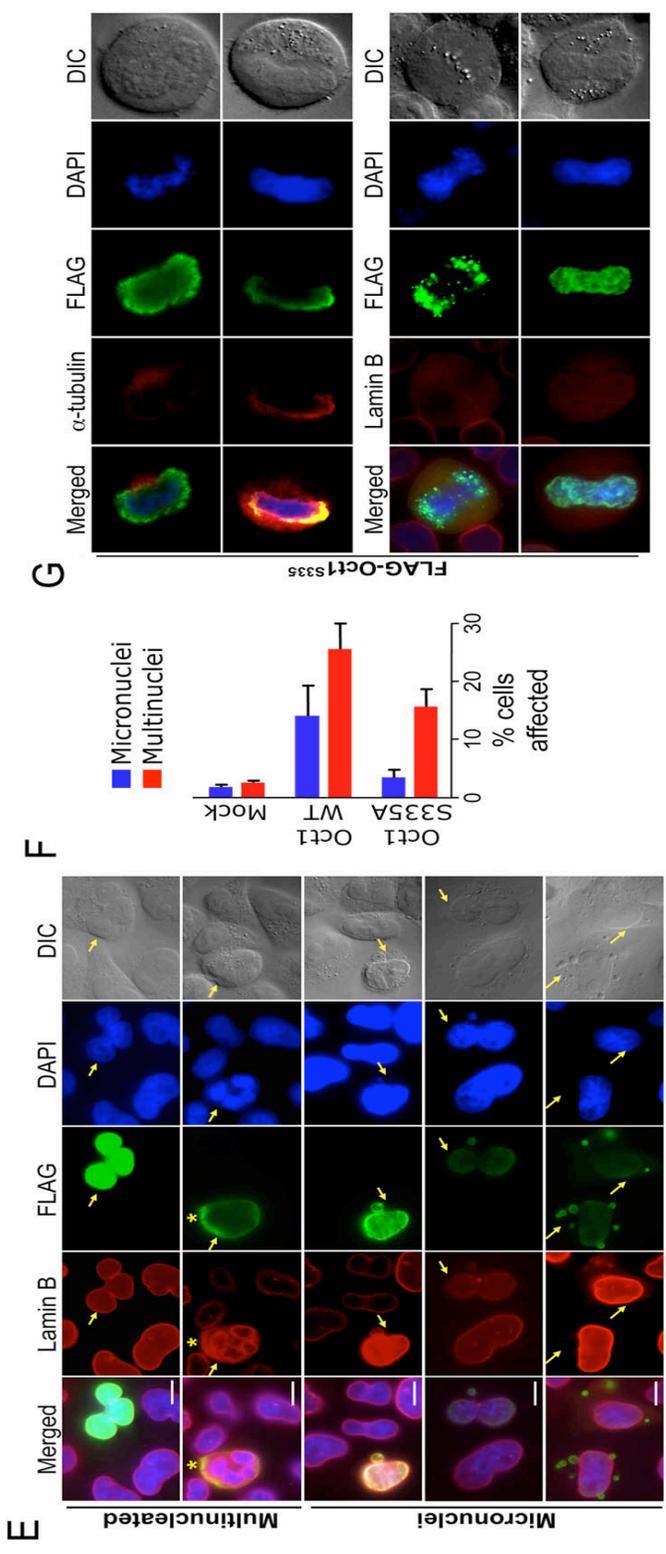


Figure 3.7. Continued

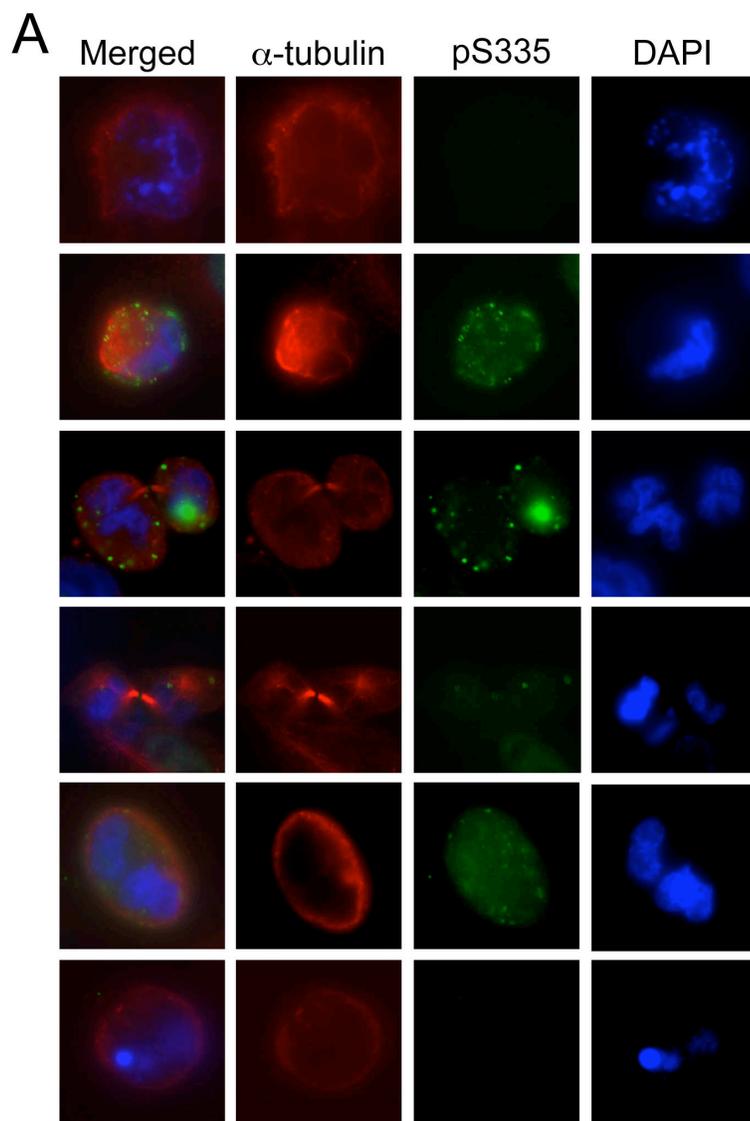


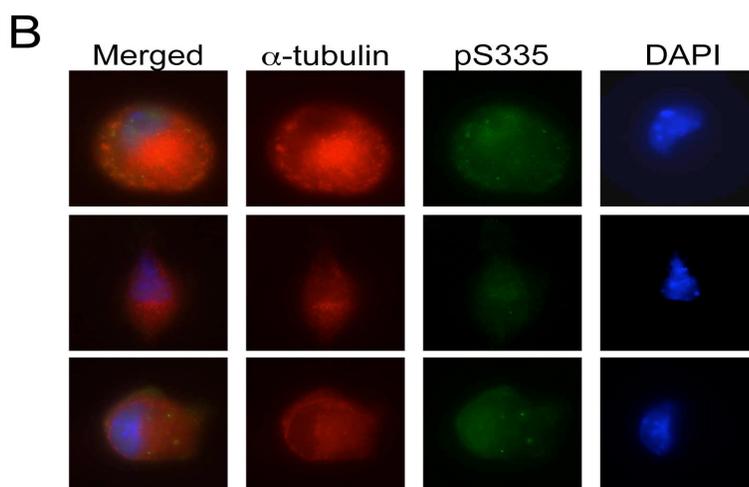
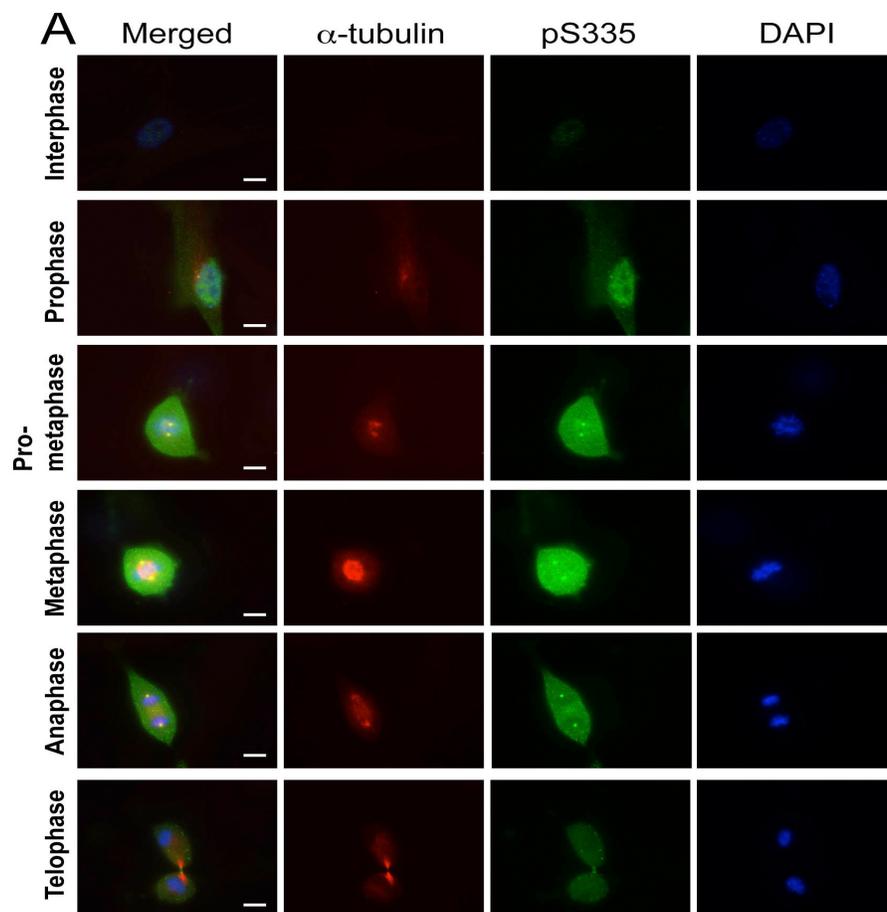
Figure 3.8. Additional examples of mitotic abnormalities in HeLa cells treated with Oct1-specific siRNAs. Reference Fig. 3.7 for normal mitotic and scrambled siRNA controls.

observed upon enrichment by affinity purification (17). No IF signal was detected in Oct1 deficient MEFs using pan-Oct1 antibodies (not shown). Using the phospho-specific antibody, wild-type fibroblasts displayed a similar, albeit less uniform, mitotic staining pattern as HeLa cells. The signal was diminished but not eliminated in *Oct1*^{-/-} fibroblasts (Figure 3.9). These results suggested the presence of a cross-reacting co-expressed POU transcription factor in fibroblasts. Oct1 is the sole detectable octamer DNA binding activity in MEFs (17). We therefore focused on nonoctamer binding POU transcription factors with capacity to cross-react. One murine protein, Pit-1/POU1F1, contains a perfect match to the peptide sequence used to generate the phospho-specific antibody. Western blotting using pan-Pit-1 antibodies indicated that Pit-1 was expressed in murine fibroblasts but not HeLa cells (Figure 3.9). Either due to redundancy with Pit-1 or other differences between HeLa cells and primary murine fibroblasts, we observed more mild evidence of abnormal mitoses in *Oct1*^{-/-} fibroblasts, including occasional abnormal DNA condensation and abnormal spindles. Analysis of DNA content also revealed the presence of aneuploidy, and an increase in cells with >4N DNA in *Oct1*^{-/-} fibroblasts (Figure 3.9).

Overexpression of Oct1 results in abnormal mitoses

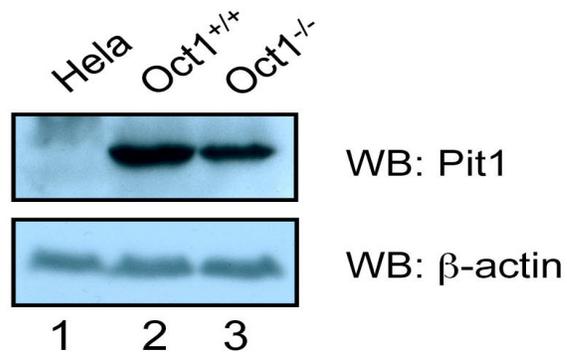
We tested the effect of overexpressed full-length FLAG-tagged wild-type and S335A Oct1 in HeLa cells using transient transfection. Overexpression of wild-type protein resulted in significantly increased Oct1^{pS335} staining, in particular the generation of interphase Oct1^{pS335}-staining puncta, while little change was observed using the S335A mutant (Figure 3.10). In mitotic cells, overexpressed wild-type FLAG-Oct1 was also excluded from mitotic chromatin and resulted in disorganized mitotic microtubules

Figure 3.9. Phospho-S335 IF staining pattern and cell-cycle phenotype of primary MEFs. (A) Immunofluorescence images of mitotic stages from wild-type primary early-passage MEFs. (B) Mitotic examples of *Oct1*^{-/-} MEFs. These examples could not be easily staged due to abnormalities. Note the reduced pS335 staining. (C) Alignment of human Oct1 and mouse Oct1, Oct2 and Pit-1. Alignment was generated using a Clustal W-based algorithm within the Vector NTI software package. Lower panel shows a Western blot of HeLa cells, wild-type and Oct1 deficient fibroblasts. The Pit-1 antibody was obtained from Santa Cruz. (D) Cell cycle profile of primary early passage Oct1 deficient MEFs and wild-type littermate controls. Inset shows expanded view of cells with super-4N DNA content.



C

	(1)	1	15
HsOct1(CAA31767)	(1)	RFEALNLSFKNMCKL	
MmOct1(CAA48423)	(1)	RFEALNLSFKNMCKL	
MmOct2(NP035268)	(1)	RFEALNLSFKNMCKL	
MmPit1(NP035267)	(1)	RFEALNLSFKNMCKL	
Consensus	(1)	RFEALNLSFKNMCKL	



D

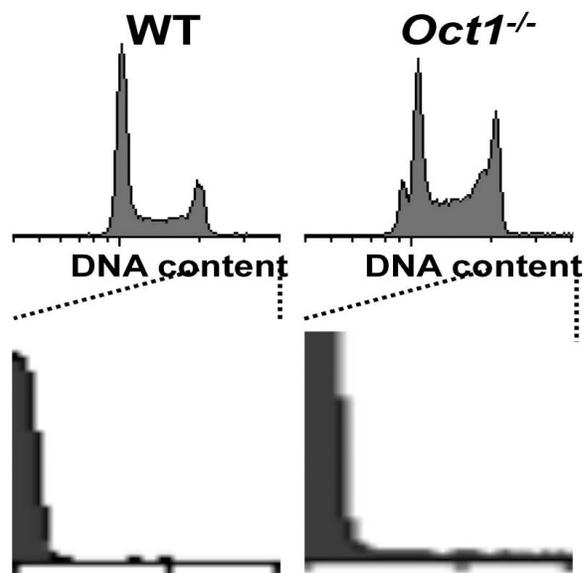


Figure 3.9 Continued

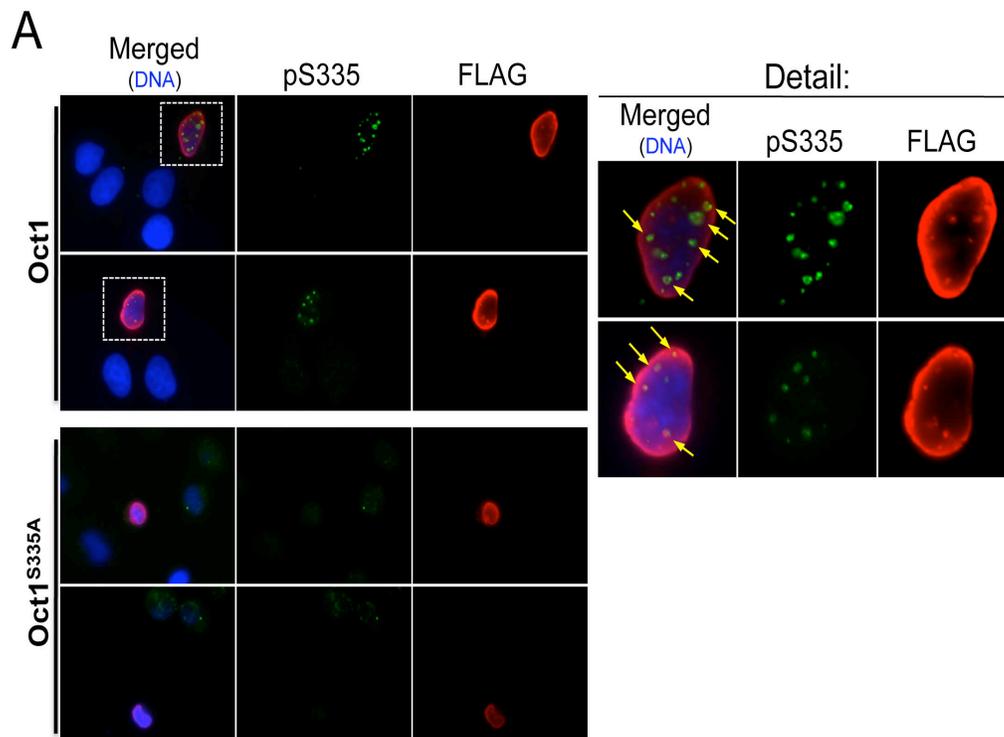


Figure 3.10. Effect of WT and S335A Oct1 overexpression on Oct1^{pS335} in HeLa cells. IF images are shown of interphase and HeLa cells transiently transfected with wild-type or S335A mutant pCG-Oct1.

as compared to adjacent untransfected controls (Figure 3.7D). We also studied interphase cells, using lamin B (B1+B2) antibodies to visualize the nuclear envelope. Adjacent untransfected (FLAG-negative) HeLa cells served as an internal control. Cells in which Oct1 was concentrated in particular areas also displayed lamin B concentrations in the same areas (Figure 3.7E, asterisks), consistent with a described interaction between the two proteins (7-9). Interphase cells overexpressing FLAG-Oct1 displayed increases in multinucleation and micronuclei. The micronuclei contained FLAG-Oct1 (Figure 3.7E, and Figure 3.11). S335A Mutant Oct1 was incapable of inducing micronuclei, and displayed reduced capacity to induce multinucleated cells (Figure 3.7F). Unlike wild-type Oct1, overexpressed S335A mutant Oct1 also could be found at mitotic DNA (Figure 3.7G). This result indicates that S335 is required for exclusion from mitotic chromatin.

Nek6 phosphorylates Oct1 serine335 during mitosis

Computational inspection of Oct1 using the phosphorylation site database PHOSIDA (<http://www.phosida.com/>) identified a consensus Nek6 kinase target site at S335. Nek6 is a NIMA-related kinase required for normal mitosis in HeLa cells (18). We tested the ability of recombinant purified Nek6 to phosphorylate an Oct1 peptide containing S335 fused to recombinant glutathione S-transferase (GST) in vitro. GST-peptide fusions with a mutated target serine residue and a different kinase (Cdk7) were used as controls. Nek6 but not Cdk7 robustly produced a reactive target peptide, but generated no signal using mutant GST-fused peptides (Figure 3.12A). We repeated these experiments using radiolabeled ATP to demonstrate that Nek6 was not phosphorylating the peptide at another position, and to show that full-length recombinant Oct1 was

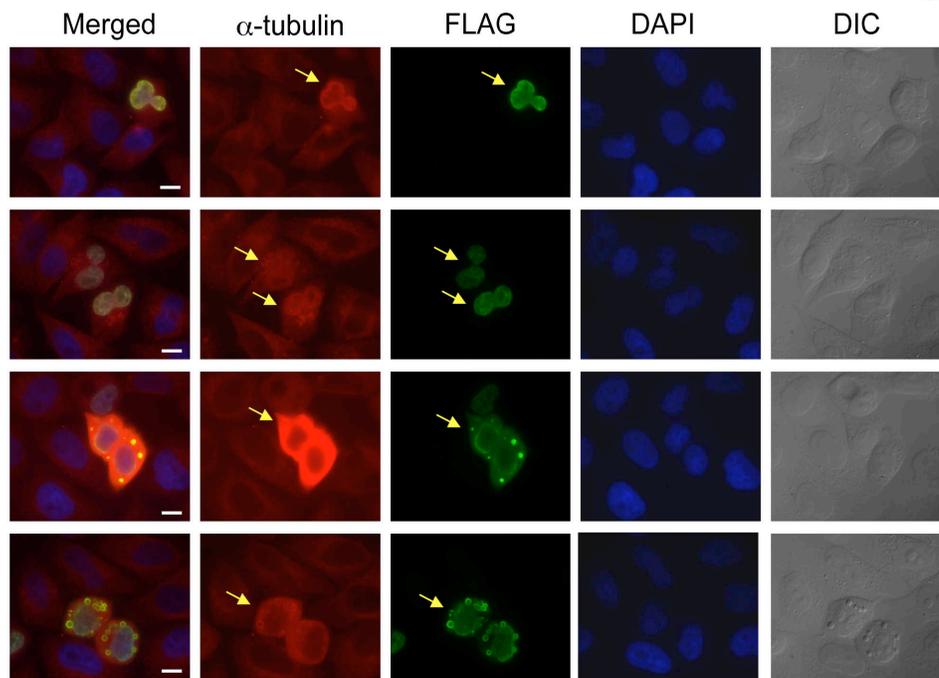
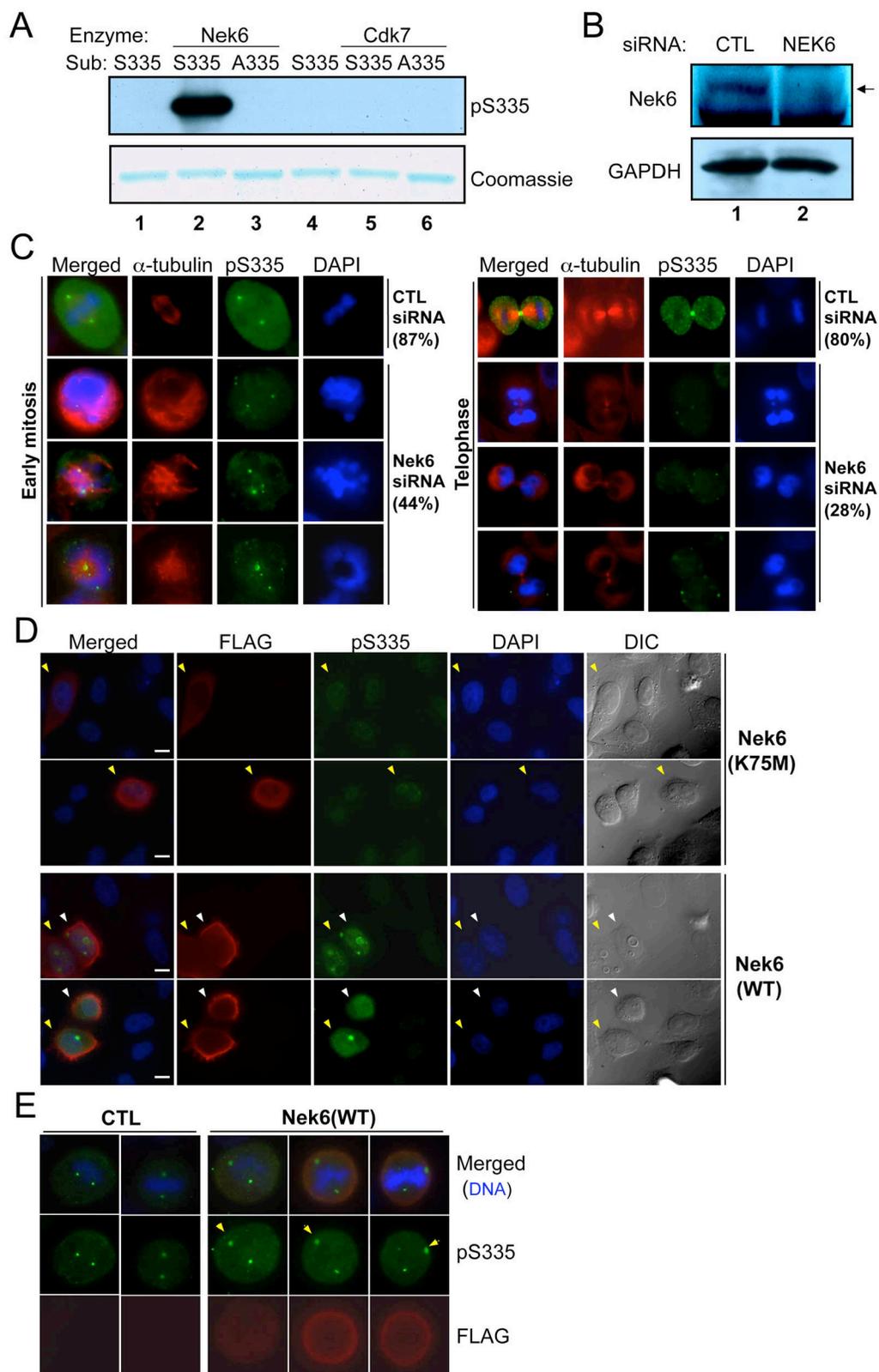


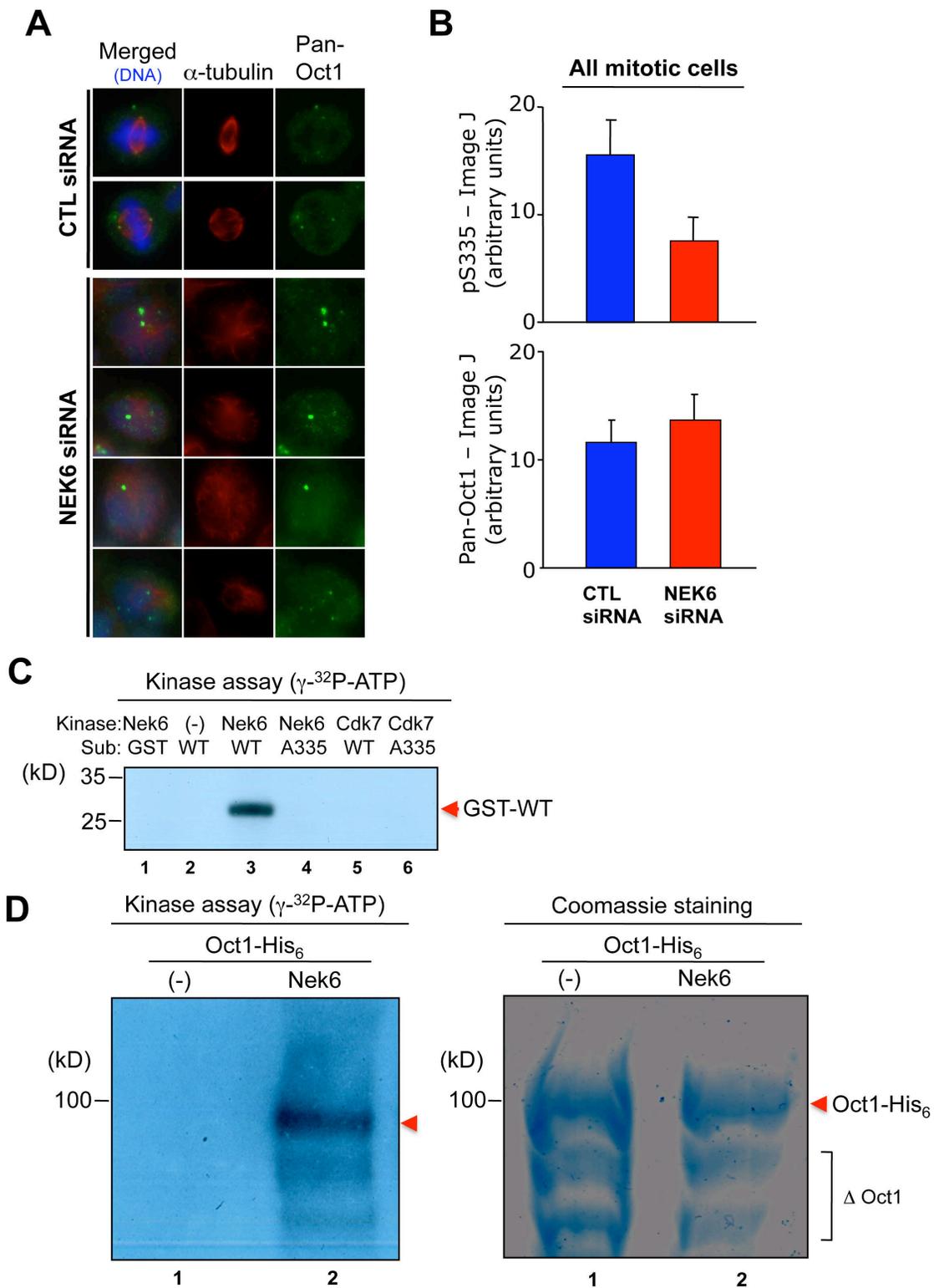
Figure 3.11. Ectopic Oct1 expression in HeLa cells increases alpha-tubulin levels, and induces formation of puncta containing alpha-tubulin but lacking DNA. IF images are shown of interphase HeLa cells transiently transfected as in Figure S7. Fixed cells were stained with DAPI and antibodies against the FLAG epitope and α -tubulin. Scale bars indicate 20 μ M.

Figure 3.12. Nek6 contributes to mitotic Oct1 phosphorylation at S335. (A) In vitro kinase assay using purified recombinant Nek6 or Cdk7, and GST fused to wild-type or mutant Ser335 target peptide sequences. A Coomassie blue-stained SDS-polyacrylamide gel is also shown to confirm presence of the purified peptide. (B) Nek6 knockdown in HeLa cells. A Western blot using anti-Nek6-specific antibodies is shown. Extracts were prepared 72 hr post-transfection. (C) HeLa cells were transfected with scrambled and Nek6-specific siRNAs, incubated for 72 hr, fixed and stained with DAPI, anti- α -tubulin and anti-Oct1^{pS335} antibodies. Examples of early (left) and late (right) mitoses are shown. Early mitotic percentages reflect the number of events showing strong Oct1^{pS335} staining (42/48 in the control vs. 21/47 in the Nek6 specific knockdown). Telophase percentages reflect the number of events showing strong midbody staining (33/41 vs. 11/39). Formaldehyde fixation was used. (D) HeLa cells were transiently transfected with FLAG-tagged wild-type Nek6 or catalytically inactive mutants (K75M), incubated for 24 hr, and prepared as in (B). Examples of interphase cells are shown. Arrows indicated transfected (FLAG-positive) cells. Formaldehyde fixation was used. (E) Mitotic examples. Formaldehyde fixation was used.



phosphorylated by Nek6 (Figure 3.13). The data are consistent with a model in which Nek6 phosphorylates Oct1 during mitosis. To test this model, we transiently transfected Nek6-specific and scrambled control siRNAs into HeLa cells. At 72 hr post-transfection the knockdown was robust (Figure 3.12B). We focused on mitotic events (rounded cells with duplicated centrosomes and partially/fully condensed chromosomes). In early mitosis, Nek6 knockdown reduced the intensity and uniformity of Oct1^{pS335} staining and disrupted the organization of mitotic spindles (Figure 3.12C, left panels). Later in mitosis, Nek6 knockdown also reduced the overall staining intensity, and specifically ablated Oct1^{pS335} detected at the midbody (right panels). We quantified the degree of pan-Oct1 or Oct1^{pS335} signal intensity in all mitotic cells following control or Nek6 siRNA transfection, observing an overall two-fold decrease in Oct1^{pS335} but not pan-Oct1 in Nek6 but not control siRNA (Figure 3.13). We overexpressed FLAG-tagged wild-type or catalytically inactive mutant Nek6 (18) (plasmids a gift of A. Fry) to determine the effect on Oct1 phosphorylation. Overexpression of wild-type but not mutant Nek6 resulted in accumulation of diffuse Oct1^{pS335} and brighter Oct1^{pS335} puncta localizing to interphase centrosomes (Figure 3.12D). Moreover, additional Oct1^{pS335} puncta in interphase, and mis-localization of mitotic spindle poles, were observed. Mitotic HeLa cells also showed increased Oct1^{pS335}, although the baseline expression was higher (Figure 3.12E). These results indicate that Oct1 S335 is a Nek6 target.

Figure 3.13. Further validation of NEK6 (A) Nek6 was knocked down in HeLa cells using siRNAs as in Fig.4C, but using rabbit anti-pan-Oct1 antibodies. (B) Pan-Oct1 and phospho-Oct1 channel signal intensity was analyzed following Nek6 knockdown using ImageJ software (NIH). 6 control and 7 Nek6 siRNA mitotic events were averaged in the case of pan-Oct1, and 5 control and 5 Nek6 siRNA mitotic events were averaged in the case of phospho-Oct1. (C) GST-fused to the substrate peptide was tested as an in vitro kinase target as in Fig. 12A but using γ -³²P-ATP. (D) Full-length, His₆-tagged recombinant Oct1 purified from *E. coli* was used. Right panel shows a Coomassie-stained gel of the same material. “D Oct1” indicates presumptive Oct1 deletion products that are also phosphorylated species. METHODS: Recombinant C-terminal His₆-tag Oct1 was expressed and purified as described (Ström AC, *et al.*, *Nucleic Acids Res.* 1996 Jun 1;24(11):1981-6).



Oct1 is a component of the spindle matrix and participates in a complex with lamin B1 at the midbody

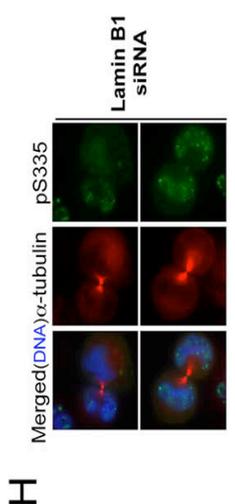
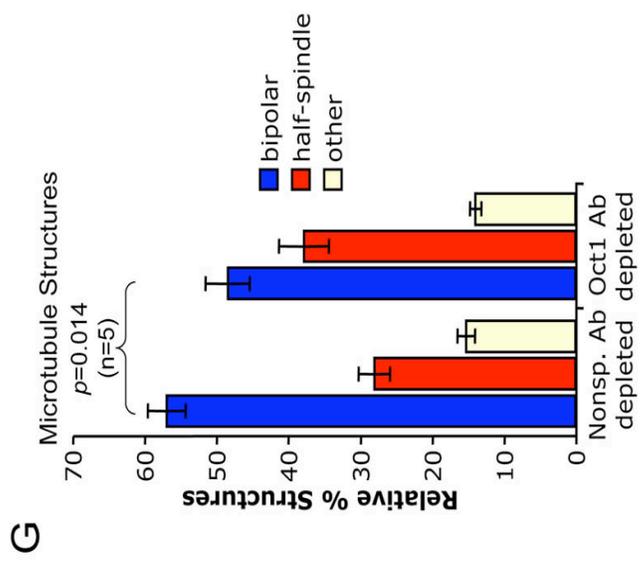
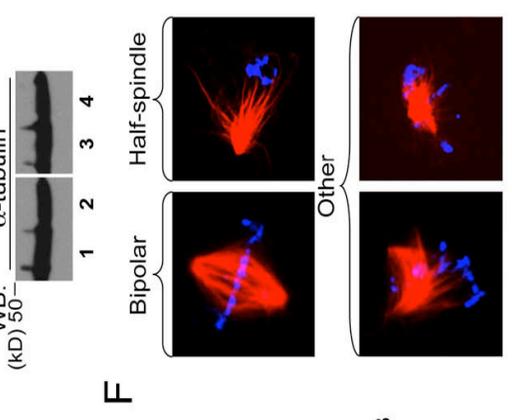
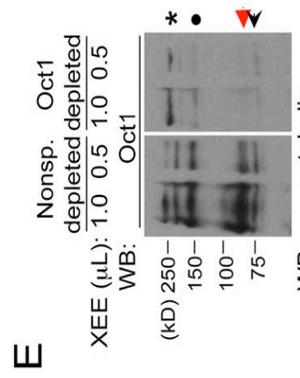
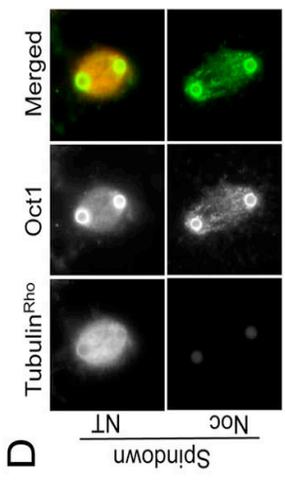
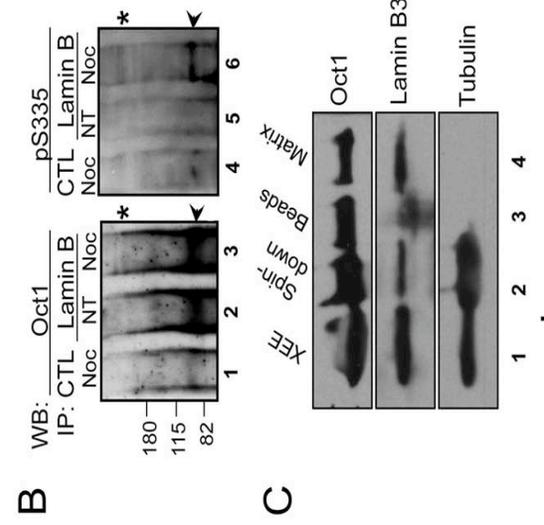
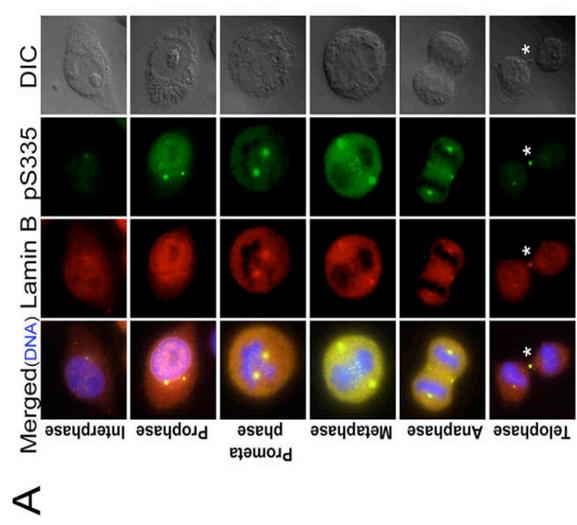
Our results suggested that Oct1 recruitment to mitotic structures is important for normal mitoses. We therefore attempted to identify whether other proteins known to interact with Oct1 recruit it to these structures. Lamin B has been shown to interact with Oct1 and can co-localize with Oct1 at the nuclear envelope (7, 8). Lamin B is also present at a structure known as the spindle matrix, and is required for proper spindle organization (19). We used antibodies against total lamin B1+B2 to observe localization in HeLa mitoses. We observed substantial co-localization between lamin B and Oct1^{pS335}. In particular, the spindle poles and midbody were strongly stained with both proteins (Figure 3.14A, asterisks).

To determine whether phosphorylated Oct1 and lamin B interact, we performed co-immunoprecipitation experiments using whole cell extracts from untreated or nocodazole-arrested HeLa cells. As expected, immunoprecipitation with lamin B antibodies enriched total Oct1 in cycling HeLa cells (Figure 3.14B, lane 2, arrow). Equivalent enrichment was also observed M-phase arrested cells (lane 3), indicating that even after nuclear envelope breakdown the interaction between Oct1 and lamin B is preserved. These data indicate that the known interaction between Oct1 and lamin B1 can be extended to mitosis. Immunoblotting using Oct1^{pS335} antibodies also uncovered an interaction between mitotic phosphorylated Oct1 and lamin B 9 (lane 6, arrow). The high molecular weight form of Oct1 enriched in mitosis interacted only poorly with lamin B (lane 6, asterisk).

Figure 3.14. Oct1 is present in the spindle matrix and forms a complex with lamin B1 at the midbody in HeLa cells. (A) Association of phosphorylated Oct1 with lamin B at the centrosomes and midbody. HeLa cells were fixed and stained with antibodies against lamin B1+B2 and Oct1^{pS335}. Mitotic stage is on the left. Asterisk indicates the midbody structure. (B) Whole cell extracts from cycling HeLa cells and cells arrested in M-phase using nocodazole were immunoprecipitated using mouse anti-lamin B antibodies. Left panel shows a Western blot using pan-Oct1 antibodies. Black arrow shows predicted Oct1 molecular weight. Asterisk shows the high molecular weight form identified in Fig. 3.1. Right panel: the blot was stripped and re-probed using Oct1^{pS335} antibodies. (C) Spindle matrix preparations generated from *Xenopus* oocyte extracts (XEE, lane 1) were Western blotted using pan-Oct1, lamin B3, and α -tubulin antibodies. (D) IF images of bead spindown preparation. Pan-Oct1 antibodies, and rhodamine-conjugated α -tubulin were used. (E) *Xenopus* Oct1 was immunodepleted using magnetic protein A-coupled beads (see methods). Oct1 Western blots are shown of the nonspecific and Oct1-specific depletions. α -tubulin is shown as a loading control. (F) Examples of spindle structures generated using the depleted extracts. Images of structures conforming to the scoring criteria used in (G) are shown. (G) Quantification of spindle structures using non-specific or Oct1-specific depletion. Error bars depict standard error of the mean. (H) HeLa cells were transiently transfected with Lamin B1-specific siRNAs. Cells were incubated for 72 hr, fixed and stained with α -tubulin and pS335 antibodies. Images of cells undergoing abscission are shown. Formaldehyde fixation was used. (I) HeLa cells transfected with control siRNAs, or siRNAs against Oct1 or lamin B1 were fixed and stained with lamin B and Oct1^{pS335} antibodies. IF images of mitotic HeLa cells

undergoing abscission are shown. Detail at right shows isolated midbody structures.

Formaldehyde fixation was used.



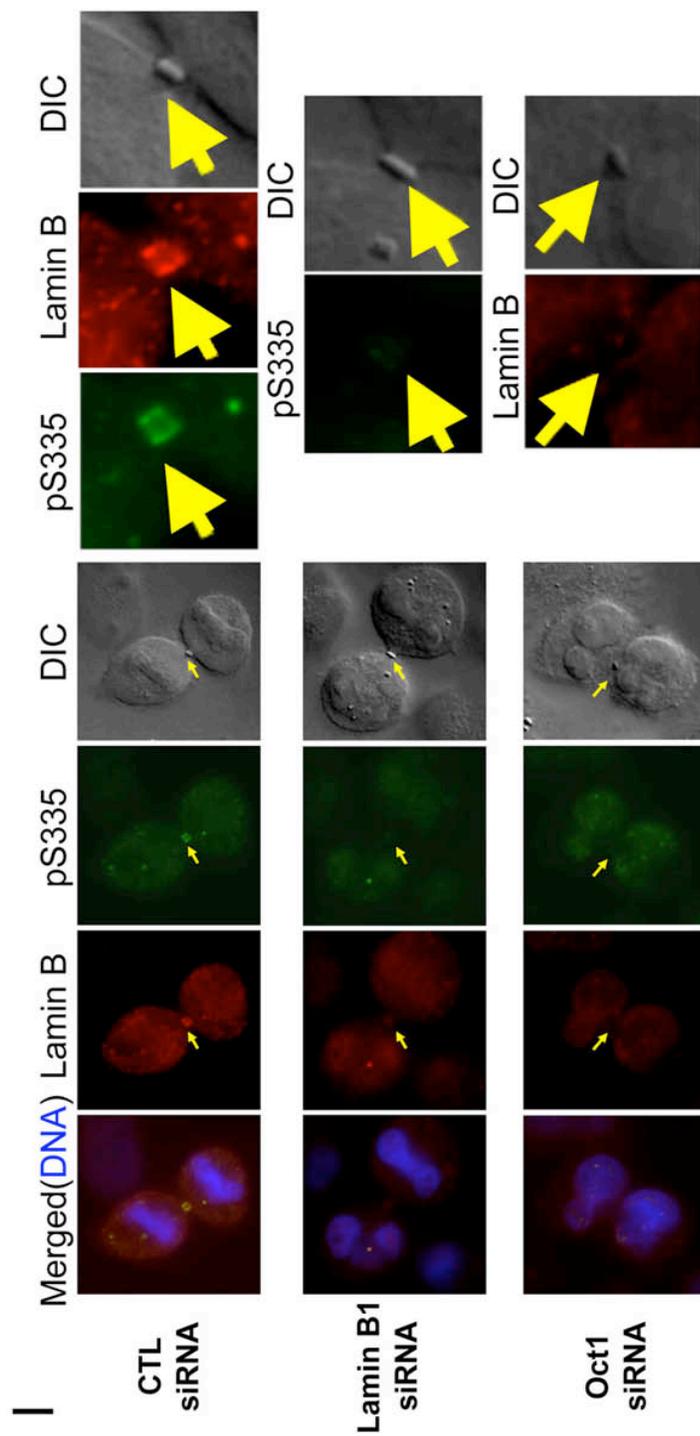


Figure 3.14 Continued

Prior studies in *Xenopus* have shown that lamin B helps to form a network referred to as the spindle matrix during mitosis. This structure associates with the spindle to help maintain spindle pole focus and spindle shape. Mass spectrometry analyses revealed that a number of transcription factors including Oct2, an Oct1 paralog, are present in isolated spindle matrix (19, 20). Spindle matrix components can be isolated from *Xenopus* egg extracts (XEE) using a spindle assembly assay stimulated by magnetic beads coated with the mitotic kinase Aurora A (19, 20). The beads function as potent microtubule nucleating and organizing centers and efficiently organize spindle poles. We retrieved the bead-associated spindles using a magnet (Figure 3.14C, “Spindown”). Buffer containing nocodazole was used to depolymerize spindle microtubules. The beads (Figure 3.14C, lane 3) and their associated spindle matrix (lane 4) were separated from each other. We identified a band corresponding to Oct1 in the spindle, beads, and matrix preparations using pan-Oct1 antibodies. As expected, most lamin B3, the major form of lamin B in XEE, was present in the spindle matrix (Figure 3.14C, lane 4). The presence of Oct1 in the beads and the spindle matrix is consistent with the idea that subsets of Oct1 are associated with the spindle poles and surrounding matrix. We visualized the bead preparations and the associated matrix using fluorescence microscopy. Robust levels of Oct1 associated with the beads themselves as well as the associated matrix (Figure 3.14D).

The above result suggested that loss of *Xenopus* Oct1 could have functional effects in this assay, although there is a distinct possibility of redundancy with the Oct2 protein. To test this prediction, we immunodepleted Oct1 from XEE and reconstituted the assay. Overall Oct1 immunodepletion was estimated at ~85% based on band intensity, although

the antibody showed some variation in the ability to immunodeplete different Oct1 species (Figure 3.14E). Immunodepleted extracts successfully reconstituted mitotic spindle structures. However there was a statistically significant defect in ability to form normal bipolar spindle structures relative to mock-depleted extracts, and a corresponding increase in aberrant monopolar structures (Figure 3.14F-G).

We next silenced lamin B1 using siRNAs and monitored Oct1^{pS335} localization during metaphase and anaphase. Lamin B1 silencing resulted in early mitotic defects such as mitotic spindle disruption and unfocused spindle poles (Figure 3.15), as reported previously (19). Phospho-Oct1 localization to both kinetochores and spindle poles remained intact, suggesting that lamin B1 is not required for Oct1 localization to these two structures. However, interestingly in late mitosis Oct1 midbody localization was abolished (Figure 3.14H). This result suggested that lamin B1 localizes phosphorylated Oct1 to the midbody. To study the interaction of Oct1 and lamin B at the midbody, we knocked down Oct1 and lamin B1 in HeLa cells, and stained with lamin B (B1+B2) and Oct1^{pS335} antibodies. Control-siRNA transfected HeLa cells showed expected concentrations of lamin B and Oct1^{pS335} at the midbody (Figure 3.14I, top panel and detail at right). Due to the presence of lamin B2, total lamin B staining was minimally affected, however lamin B concentration at the midbody was largely eliminated (Figure 3.14I, middle panels), indicating that midbody lamin B consists mostly of lamin B1 in HeLa cells. In the lamin B1 knockdown condition, Oct1^{pS335} midbody staining was also absent in 50% of cells undergoing cytokinesis and significantly depleted in others (detail at right), suggesting that lamin B1 recruits phospho-Oct1. Oct1 knockdown (lower panels)

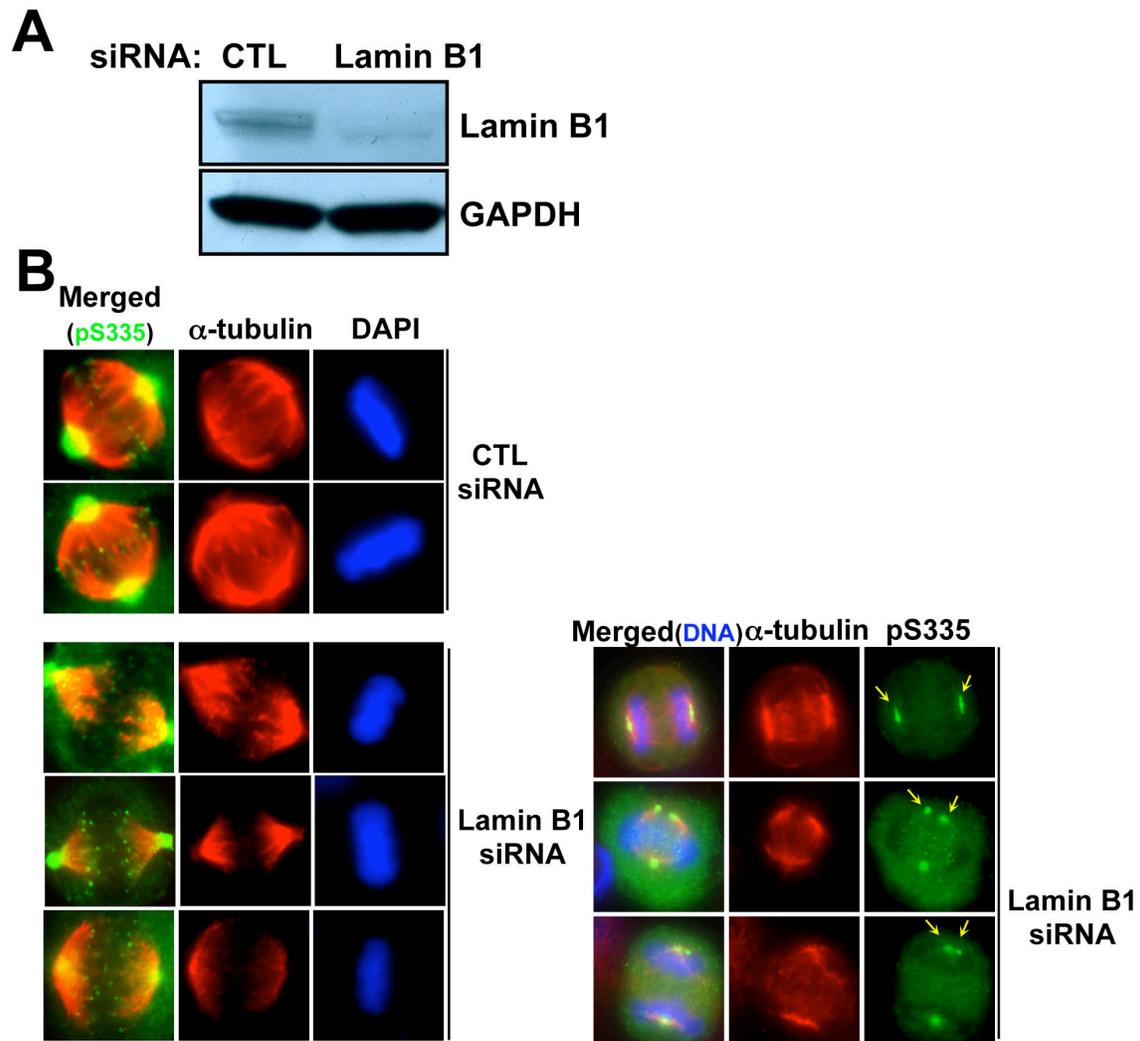


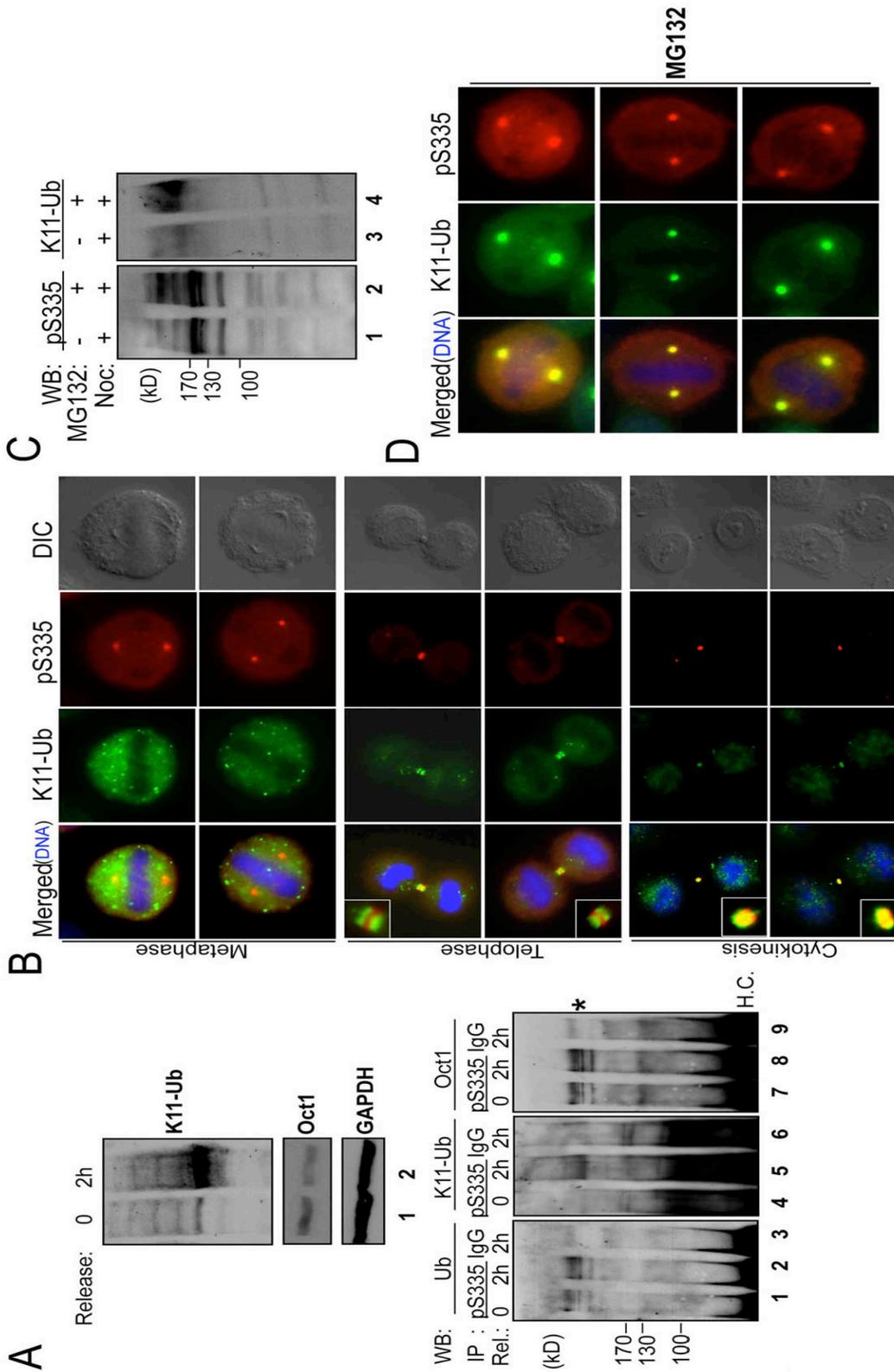
Figure 3.15. Abnormal localization of Oct1^{pS335} by lamin B knockdown in early mitotic cells (A) HeLa cells were transiently transfected with siRNAs against lamin B1 (Santa Cruz), incubated for 72 hr and Western blotted with antibodies against lamin B1 or GAPDH as a loading control. (B) IF images are shown of mitotic HeLa cells transiently transfected with control or lamin B1 siRNAs, and stained with antibodies against lamin B (B1+B2) and Oct1^{pS335} antibodies. Arrows highlight elongated/split spindle poles.

also quantitatively depleted lamin B1+B2 at the midbody. These findings indicate that during cytokinesis lamin B1 and Oct1 are mutually required for midbody localization.

Mitotic Oct1^{pS335} is modified by K11-linked poly-Ub chains associated with the midbody

Our data suggested that there are qualitative differences between Oct1^{pS335} midbody localization and localization to other mitotic structures: midbody association is maintained after most signal is eliminated (Figure 3.3A), and lamin B is required for Oct1^{pS335} midbody localization but not localization to other structures (Figure 3.14F-H). We therefore sought biochemical correlates that may underlie these differences. Proteins modified by noncanonical K11-linked poly-Ub chains are enriched in the midbody (21). We tested whether Oct1 is modified through K11-linked poly-Ub chains (K11-Ub) using a specific antibody (21). Although Oct1 is ubiquitinated in cycling or nocodazole-arrested HeLa cells (Figure 3.1C), immunoprecipitation of phospho-Oct1 from these cells and Western blotting using K11-Ub antibodies produced little evidence of Oct1 K11-linked ubiquitination (not shown). K11 accumulates late in mitosis (21) suggesting that Oct1 may be modified by canonical Ub early in mitosis, but switches to a K11-linked form at later stages. We therefore arrested cells in G1/S with thymidine, released them from the thymidine block and arrested them in mitosis with nocodazole. Following release from nocodazole arrest, K11-Ub linkages were detectable after two hours (Figure 3.16A, upper panels, 21). Under the same conditions, immunoprecipitated Oct1^{pS335} was associated with K11-Ub as assessed by Western blot (Figure 3.16A, lower panels). To

Figure 3.16: Dynamic ubiquitination of Oct1 during mitosis. (A) HeLa cells were blocked in G1 with thymidine, released and arrested in mitosis using 0.1 mM nocodazole as described (21). Top panels show Western blot with K11-Ub and pan-Oct1 antibodies, and GAPDH antibodies as a loading control, from cells arrested with nocodazole or following 2 hr release from nocodazole. For the bottom panels, samples were immunoprecipitated using anti-Oct1^{pS335} antibodies, and probed with antibodies against pan-Ub, K11-Ub and pan-Oct1. H.C.=immunoglobulin heavy chain. (B) IF images of HeLa mitoses stained with Oct1^{pS335} and K11-Ub antibodies. Two metaphase, telophase and late cytokinesis examples are shown. Merged images from the two latter cases also show detail of the midbody structure (inset). (C) HeLa cells were arrested with nocodazole (0.5 mM) for 18 hr, and then incubated with MG132 for a further 6 hr. Whole cell extracts were prepared and subjected to Western blotting using Oct1^{pS335} and K11-Ub antibodies. (D) HeLa cells were treated with MG132, fixed and subjected to IF using K11-Ub and Oct1^{pS335} antibodies. Detail of metaphase cells is shown. (E) Whole cell extracts from HeLa cells arrested as above were immunoprecipitated with anti-Oct1^{pS335} antibodies and Western blotted using pan-Ub, K11-Ub or pan-Oct1 antibodies. H.C.=immunoglobulin heavy chain. (F) Nocodazole-arrested HeLa whole cell extracts were immunoprecipitated using Oct1^{pS335} antibodies and probed using pan-Oct1 or APC1 (AbCam). (G) HeLa cells were transfected with Cdh1 siRNAs. Twenty-four hr-post transfection, cells were treated with 0.5 mM nocodazole. Forty-two hr-post transfection, cells were treated with MG132. Whole cell extracts were prepared after forty-eight hr. (H) siRNA-transfected HeLa cells as in (G) were immunoprecipitated using Oct1^{pS335} antibodies and Western blotted using pan-Ub or K11-Ub antibodies.



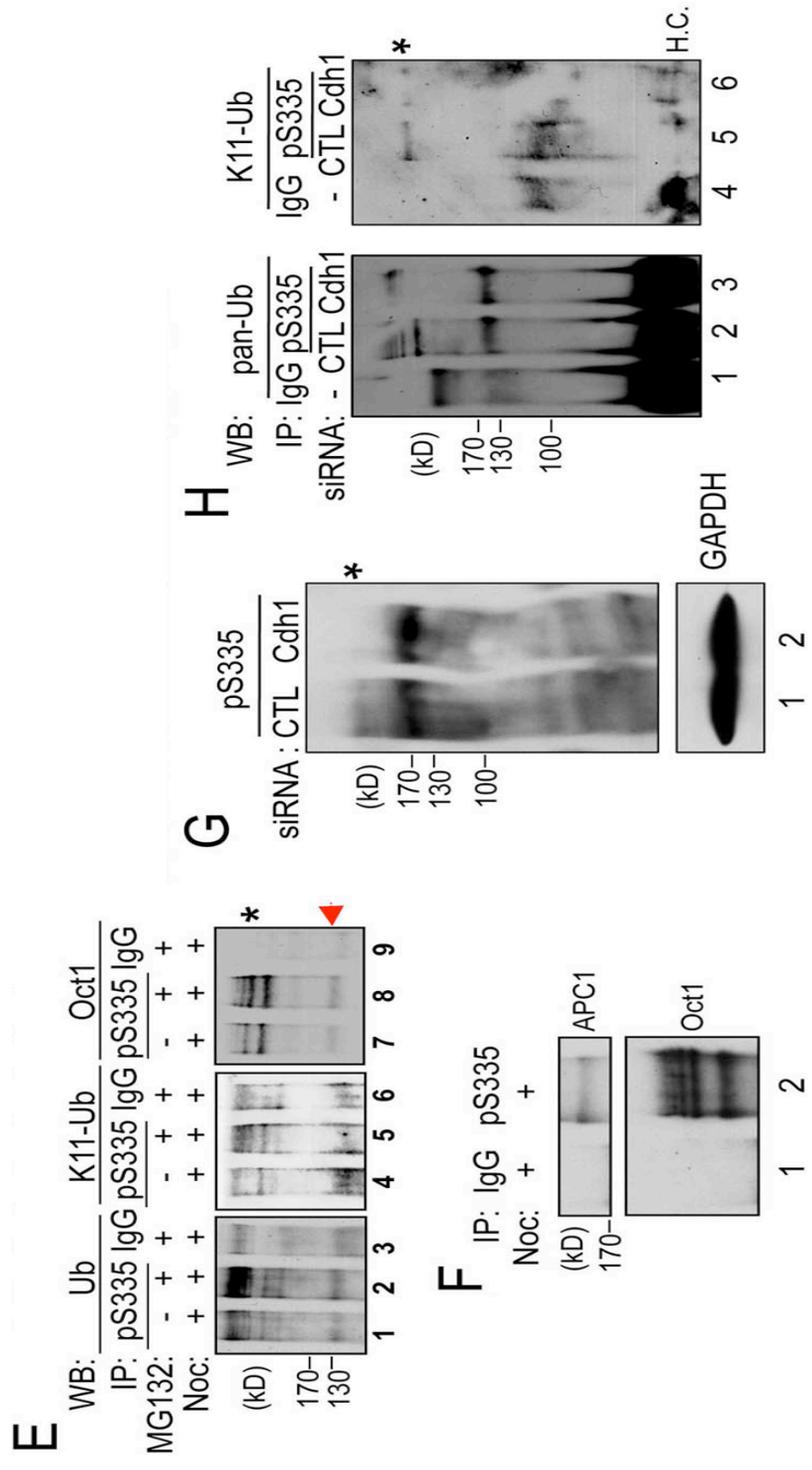


Figure 3.16 Continued

confirm this finding, we co-stained HeLa cells with the Oct1^{pS335} and K11-Ub antibodies and examined late stage mitoses (Figure 3.16B). In metaphase, both K11-Ub and Oct1^{pS335} were excluded from mitotic chromatin. However the two antibody signals displayed little overlap. For example, no K11-Ub was detected in the spindle poles. Telophase cells showed co-localization to the developing midbody, however the two signals remained spatially distinct, with K11-Ub flanking the more centrally localized Oct1^{pS335} (Figure 3.16B, see inset detail at left). In contrast, cells late in cytokinesis showed tight spatial overlap at the midbody.

The above results are consistent with models in which Oct1 K11-Ub occurs exclusively late in mitosis, or in which Oct1 is continually modified but is rapidly degraded, with degradation slowing or ending at late mitosis. To distinguish these two possibilities, we treated 24 hr nocodazole-arrested HeLa cells with the proteasome inhibitor MG132 during the final 6 hr. Nocodazole-arrested cells showed some evidence of Oct1^{pS335} K11-Ub (Figure 3.16C, lane 1), presumably because the nocodazole arrest was less precise than that in Figure 3.16A using a thymidine block. As expected, MG132 treatment caused total K11-Ub-modified proteins to accumulate (lane 4). In addition, the higher molecular weight forms of phosphorylated Oct1 were increased while the lower molecular weight forms were unaffected (Figure 3.16C, lane 2). These results indicate that Oct1 ubiquitinated species can be enriched by proteasome inhibition. Using IF, we found that mitotic HeLa cells treated with MG132 showed K11-Ub colocalization with Oct1^{pS335} at the spindle poles (Figure 3.16D). Mitotic cells treated with MG132 also showed larger spindle pole puncta (compare Figure 3.16D with Figure 3.16B). We therefore immunoprecipitated Oct1^{pS335} from extracts of nocodazole/MG132-treated

HeLa cells. Treatment with MG132 increased total and K11-specific Oct1^{pS335} ubiquitination (Figure 3.16E, lanes 2 and 5). These results are consistent with a model in which Oct1^{pS335} undergoes cycles of K11 ubiquitination and destruction throughout mitosis, except at later stages when the protein is stabilized and detectable.

One activity known to catalyze K11-linked ubiquitination is the APC (21-23). To test whether Oct1 and APC interact, we performed co-immunoprecipitation using anti-Oct1^{pS335} and extracts from mitotic HeLa cells. Western blotting revealed the presence of not only Oct1 but also the APC large subunit APC1 in the immunoprecipitate (Figure 3.16F). To demonstrate a causal connection, we knocked down *Fzr1*, which encodes the APC component Cdh1, using siRNAs (21). Specific but not control siRNA transfection significantly attenuated the high molecular weight phosphorylated form of Oct1 (Figure 3.16G). Further, Cdh1 knockdown attenuated several ubiquitinated Oct1^{pS335} forms, as measured by Oct1^{pS335} immunoprecipitation followed by Western blotting with pan-Ub or K11-Ub antibodies (Figure 3.16H, lane 3), including a K11-Ub-modified species (lane 6). Lastly, Cdh1 knockdown eliminated the ability of MG132 to redistribute K11-Ub signal to sites of Oct1 phosphorylation, implicating the APC in the deposition of K11-Ub at phosphorylated Oct1 that can be visualized when proteasome degradation is blocked (Figure 3.17).

Discussion

Here we show that Oct1 phosphorylated at position S335 by Nek6 is ubiquitinated and associates with mitotic structures. Oct1^{pS335} is displaced from mitotic chromatin and concentrated at spindle pole bodies and the midbody during mitosis. Interphase cells

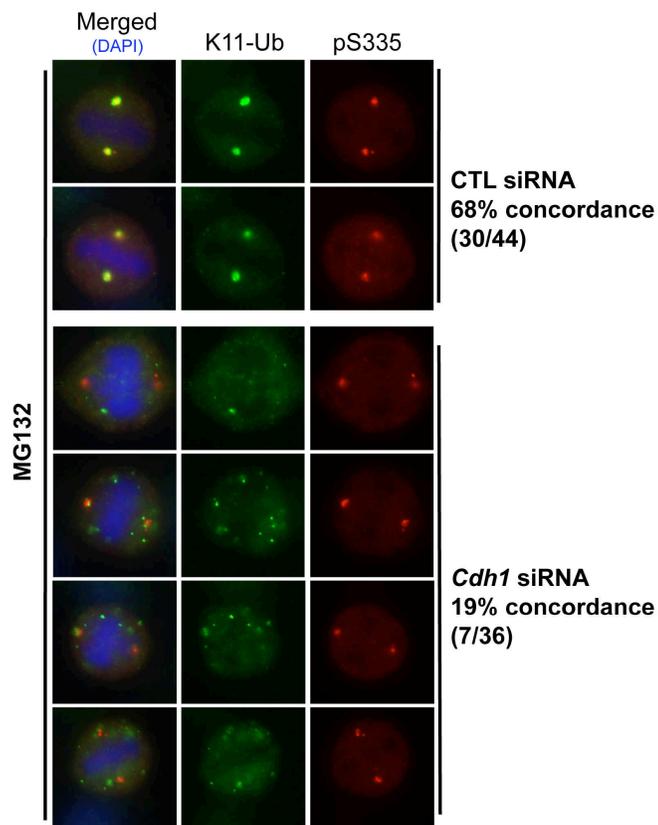


Figure 3.17. APC is responsible for the accumulation of K11-Ub signal at sites of Oct1 phosphorylation upon proteasome inhibition. HeLa cells were transiently transfected with control siRNAs or siRNAs directed against the APC component CDH1 (gene symbol *Fzr1*, Dharmacon), incubated for 48 hr. After 42 hr, cells were treated with MG132 as described in the materials section. Cells were then fixed and processed for IF. For quantification, mitotic events were scored +/- based on concordance between Oct1^{pS335} and K11-Ub staining.

show Oct1^{pS335} staining at centrosomes. The signal detected during mitosis is qualitatively different from that observed during interphase, and cannot be explained simply by increased total Oct1 levels. With the exception of the midbody, Oct1^{pS335} staining is rapidly lost at the anaphase-telophase transition. We identify Nek6 is an upstream Oct1^{pS335} kinase. Previous studies have shown that Nek6 localizes to the centrosomes in interphase cells, and to the spindle poles and midbody during mitosis. Nek6 loss of function also results in mitotic abnormalities and apoptosis in HeLa cells (18). The activities of Oct1-associated proteins are consistent with these findings. For example, the DNA damage sensing factor BRCA1 is known to interact with Oct1 (5, 6) and is a known mitotic regulator that localizes to centrosomes (24-26). PARP-1 also interacts with Oct1 (4) and localizes to centrosomes. PARP-1 is important for centrosomal function, including limiting their duplication (27).

Oct1 phosphorylation has been investigated previously within the context of the cell cycle. Mitotic phosphorylation was described at a different residue, S385 (11). S385 phosphorylation was found to be cell cycle dependent and mediated by PKA. It was also noted that Oct1 purified from M-phase cells did not bind to DNA (28). Later it was shown that Oct1 is excluded from mitotic chromatin (29). Recent screens (12-14) identified enrichment in both Oct1 S335 and S385 phosphorylation during M-phase. We found that phosphorylation at S385 does not block DNA binding but instead alters the Oct1 selectivity for different DNA binding configurations (10). We postulate that S385 phosphorylation correlates with S335 phosphorylation in mitosis, but that it is S335 phosphorylation that causes Oct1 exclusion from mitotic chromosomes. We substantiated

this hypothesis by showing that overexpressed FLAG-tagged Oct1 is excluded from mitotic DNA while Oct1 with a S335A mutation is not.

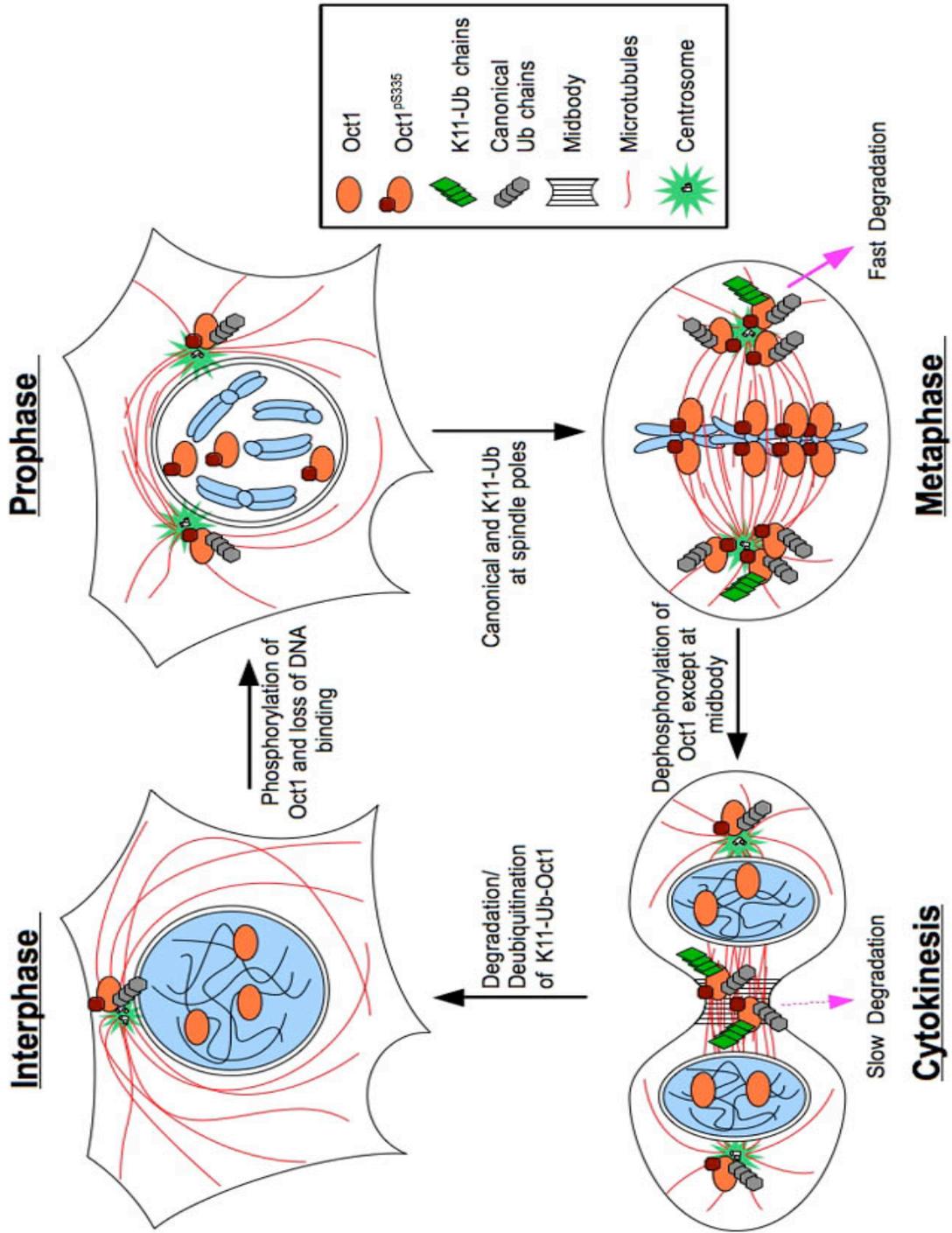
Oct1^{pS335} is also ubiquitinated, including by noncanonical K11-Ub late in mitosis. Ubiquitinated proteins have been previously associated with the spindle pole bodies and midbody (30, 31). Although we found that K11-Ub-modified Oct1 was detectable only in late mitoses in normal HeLa cells, cells treated with the proteasome inhibitor MG132 accumulated K11-Ub at structures to which Oct1 localizes in early mitotic stages. This result suggests a model in which K11-Ub modified Oct1^{pS335} is formed throughout mitosis, but is degraded rapidly prior to telophase, at which time degradation slows or stops. A model for Oct1 phosphorylation and ubiquitination through the cell cycle is shown in Figure 3.18. APC interacts with Oct1 and is required for Oct1 K11 ubiquitination, strongly suggesting that APC is the upstream Oct1 Ub ligase. We observe these interactions and activities outside of anaphase. However it is widely recognized that the APC is active throughout the cell cycle, including in early mitosis (32).

Although a simple model is that Oct1 is phosphorylated at S335 and becomes non-functional, several lines of evidence suggest that Oct1^{pS335} acquires new functions. In *Xenopus* egg extracts, Oct1 co-purifies with the spindle matrix, which helps maintain spindle shape. Oct1 also localizes with lamin B at the midbody. Lamin B1 and Oct1 are mutually required to localize each other to the midbody, suggesting that they form a complex. Aside from the specific localization to mitotic structures, results from both Oct1 loss- and gain-of-function experiments implicate Oct1 as a mitotic regulator, at least in some cell types. For example, the organization of the mitotic spindle is disrupted upon Oct1 siRNA knockdown in HeLa cells. In XEE, Oct1 depletion causes defects in spindle

Figure 3.18: Model for Oct1 localization and modification through the cell cycle.

Oct1 occupies sites in the DNA and regulates gene expression during interphase.

Oct1^{pS335} localizes to centrosomes. Early in mitosis Oct1 is phosphorylated by Nek6 and localizes to spindle pole bodies and kinetochores. Oct1 is also ubiquitinated. Oct1 modified through non-canonical K11-linked Ub chains is rapidly degraded by the proteasome and is not readily detectable unless degradation by the proteasome is inhibited. Late in mitosis the bulk of phosphorylated Oct1 is de-phosphorylated, with the remaining phosphorylated Oct1 concentrated at the midbody. K11-Ub is readily detectable at the midbody, presumably because degradation has slowed or stopped. Following abscission the remaining phosphorylated Oct1 is de-phosphorylated, degraded or relocated to the centrosome.



morphogenesis, implicating Oct1 in mitosis-specific functions. Based on the observed Oct1 localization patterns, it is likely that the mitotic irregularities caused by changes in Oct1 levels are directly associated with mitotic regulation. Furthermore, an intact serine at position 335 is important for the mitotic phenotype of over-expressed Oct1, implicating this residue in mitotic functions.

We found that Oct1 is not required for the completion of mitosis in HeLa cells but rather appears to play a regulatory role. In other cell types, such as murine fibroblasts and A549 cells, the effect of Oct1 is more mild than in HeLa cells. Oct1 deficient embryos survive past gastrulation (17, 33) and Oct1 deficient MEFs proliferate normally in culture (1, 17). Primary Oct1 deficient MEFs undergo oncogenic transformation poorly relative to wild-type controls, but immortalize normally by serial passage (1). In this sense, the role of Oct1 in mitosis may be more akin to BRCA1, which appears to act as a mitotic regulator rather than a core component of the mitotic machinery (24-26). As a second example, lamin-B RNAi results in a delay in prometaphase, following which cells can finish mitosis (19).

Methods

Tissue culture

Oct1 deficient MEFs have been described previously (17). HeLa cells (ATCC) were arrested in M-phase using 0.5mg/ml nocodazole for 18 hr. Thymidine block and release from nocodazole were performed identically to Matsumoto et al. (21). MG132 (Calbiochem) was applied at 5 mM for 6 hr. For experiments using both nocodazole and MG132, cells were treated with nocodazole for 12 hr, following which MG132 was

added at 5 mM and cells were incubated with nocodazole and MG132 for a further 6 hr. Cells were maintained in a humidified environment at 37° C with 5% CO₂. HeLa cells were transiently transfected using polyethyleneimine (PEI, Sigma) and pCG-FLAG-Oct1 (34, 35) or S335 mutant Oct1 generated as described (10).

Antibodies

A commercial rabbit phospho-specific antibody (Bethyl) was raised against the peptide EALNLS₃₃₅FKNMC. The antibody was purified in two steps, first by blocking with unphosphorylated peptide to remove nonphospho-specific antibodies, then by affinity purification using the phosphorylated epitope. Mouse antibodies against α -tubulin, γ -tubulin, CLASP-1, lamin B, and goat anti-Pit-1, were obtained from Santa Cruz. Mouse anti-pS10-H3 was obtained from Cell Signaling, and mouse anti-glyceraldehyde 3-phosphate dehydrogenase (GAPDH) from Chemicon. Human anti-APC1 antibodies were purchased from Abcam. Mouse anti-FLAG and rabbit anti-Nek6 antibodies were from Sigma. For pan-Oct1, a rabbit antibody (Bethyl) was used with the exception of Figure 3.1D, which used a mouse antibody (Millipore). Rabbit anti-Ub antibodies were a gift of M. Rechsteiner. Human anti-K11-Ub antibodies were a gift of V. Dixit.

Spindle matrix preparation

Assembly of spindles from *Xenopus* egg extracts using AuroraA-conjugated beads was conducted as described (19). For Oct1 immunodepletion, rabbit anti-Oct1 (Bethyl) or

nonspecific IgG control (Sigma) antibodies were used similarly to Goodman et al. (36), except that 75 ml of magnetic protein A-Dynabead slurry was used per 100 ml of extract.

Immunofluorescence

Coverslips were coated with Poly-L-lysine (Sigma) for 30 min and placed into 6 well dishes where counted cells were plated. Cells were washed with phosphate buffered saline (PBS) two times prior to fixation. Methanol fixation was performed as described (21). For formaldehyde fixation, cells were incubated in 4% formaldehyde in CSK buffer (100mM NaCl, 300mM Sucrose and 10mM PIPES pH6.8) for 30 min at room temperature (RT), and washed three times with CSK buffer plus protease inhibitor cocktail (PIs, Roche). Permeabilization was achieved by adding CSK buffer (+ 0.5% Triton-X-100 and PIs) for 10 min, followed by 3 washes with PBS-T (PBS + 0.05% Tween-20). IF images using K11-Ub, pan-Ub, and pan-Oct1 antibodies used methanol fixation, while phospho-histone H3 antibodies used formaldehyde. Antibodies against CLASP-1, α and γ -tubulin, lamin B and phospho-S335 worked efficiently with both fixation procedures. Fixed cells were blocked with IF buffer (PBS-T with 1% donkey serum) for 1 hr at RT. Primary and secondary antibodies diluted in the IF buffer and were incubated sequentially. After each incubation, cells were washed with PBS-T for 10 min. Stained coverslips were placed on slides using mounting medium with DAPI (Vector). Images were taken using a Zeiss Axioplan 2 imaging microscope with a 100X oil immersion objective and a numerical aperture of 1.3. Digital fluorescence and DIC images were acquired using an AxioCam MRm monochrome digital camera. Final images were processed, given false color and merged using Photoshop CS3 (Adobe

Systems). All scale bars conform to 20 μ m. Mitotic staging followed established criteria, e.g., Maiato et al. (37) and Pereira et al. (38).

In vitro kinase assay

Two sets of complimentary oligo DNA (S335 and A335) were annealed and ligated into pGEX-4T1 (Promega) using a *Sma*I restriction site. Expressed proteins in BL21 Codon-plus (Stratagene) *Escherichia coli* transformed with each construct were purified using glutathione-sepharose resin (GE Healthcare). Sequences were S335: 5'GCGATTTGAAGCCTTGAACCTCAGCTTTAAGAACATGTGCAAGTGA3', 5'TCACTTGCACATGTTCTTAAAGCTGAGGTTCAAGGCTTCAAATCGC3'; A335: 5'GCGATTTGAAGCCTTGAACCTCGCCTTTAAGAACATGTGCAAGTGA3', 5'TCACTTGCACATGTTCTTAAAGGCGAGGTTCAAGGCTTCAAATCGC3'. The GST-fused peptide (WT, RFEALNLS₃₃₅FKNMCK or S335A) were incubated with recombinant Nek6 and CAK complex (cdk7/cyclin H/MAT1, Millipore). Kinase activities were assayed with the purified substrate according to a manufacturer's protocol.

RNAi

siRNA pools targeting three different regions of Oct1 or Nek6 (Santa Cruz) were mixed with lipofectamine RNAi max (Invitrogen) and transiently transfected according to manufacturer's protocol. Control siRNAs were used in parallel and also purchased from Santa Cruz. Cells were cultured for 3 days prior to analysis. siRNA knockdown of CDH1 (gene symbol *Fzr1*) and control transfections used siRNA pools (Dharmacon). Cells were cultured for 48 hr prior to analysis.

Immunoprecipitation

Cells were lysed with 18mM Hepes pH 7.9, 150mM NaCl, 1mM EDTA, 1mM EGTA, 1% Triton-X-100, protease inhibitors (Roche, PIs), and phosphatase inhibitors (Roche, PhIs). 500 mg of extract was incubated with 4 mg of antibody in IP buffer (50mM Tris pH8.0, 20% glycerol, 0.5mM EDTA, 0.1% NP-40, 1mM DTT, PIs and PhIs) overnight at 4°C. Protein-antibody complexes were precipitated using magnetic beads (Activmotif) and washed three times with IP buffer.

Acknowledgements

We thank R. Chan, K. Ullman, D. Stillman and M. Rechsteiner for critical reading of the manuscript. We thank J. Shaw and members of her laboratory for use of equipment and technical advice, and D. McClain and L. Wells for advice and reagents. A. Fry provided the wild-type and mutant Nek6 mammalian expression plasmids. The pan-Ub antibody was a gift of M. Rechsteiner. The K11-Ub antibody was a gift of V. Dixit. This work was supported by grants from the American Cancer Society (GMC-115196) and National Cancer Institute (1R21CA141009) to D.T.

Author Contributions

Conceived and designed the experiments: JK YZ DT. Performed the experiments: JK BG. Analyzed the data: JK BG. Contributed reagents/materials/analysis tools: YZ. Wrote the paper: JK BG YZ DT.

References

1. Shakya, A., R. Cooksey, J. E. Cox, V. Wang, D. A. McClain, and D. Tantin. 2009. Oct1 loss of function induces a coordinate metabolic shift that opposes tumorigenicity. *Nat Cell Biol* 11:320-327.
2. Ryan, A. K., and M. G. Rosenfeld. 1997. POU domain family values: flexibility, partnerships, and developmental codes. *Genes Dev* 11:1207-1225.
3. Kang, J., A. Shakya, and D. Tantin. 2009. Stem cells, stress, metabolism and cancer: a drama in two Octs. *Trends Biochem Sci* 34:491-499.
4. Nie, J., S. Sakamoto, D. Song, Z. Qu, K. Ota, and T. Taniguchi. 1998. Interaction of Oct-1 and automodification domain of poly(ADP-ribose) synthetase. *FEBS Lett* 424:27-32.
5. Fan, W., S. Jin, T. Tong, H. Zhao, F. Fan, M. J. Antinore, B. Rajasekaran, M. Wu, and Q. Zhan. 2002. BRCA1 regulates GADD45 through its interactions with the OCT-1 and CAAT motifs. *J Biol Chem* 277:8061-8067.
6. Wang, R. H., H. Yu, and C. X. Deng. 2004. A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint. *Proc Natl Acad Sci U S A* 101:17108-17113.
7. Imai, S., S. Nishibayashi, K. Takao, M. Tomifuji, T. Fujino, M. Hasegawa, and T. Takano. 1997. Dissociation of Oct-1 from the nuclear peripheral structure induces the cellular aging-associated collagenase gene expression. *Mol Biol Cell* 8:2407-2419.
8. Malhas, A. N., C. F. Lee, and D. J. Vaux. 2009. Lamin B1 controls oxidative stress responses via Oct-1. *J Cell Biol* 184:45-55.
9. Shakya, A., J. Kang, J. Chumley, M. A. Williams, and D. Tantin. 2011. Oct1 is a switchable, bipotential stabilizer of repressed and inducible transcriptional states. *J Biol Chem* 286:450-459.
10. Kang, J., M. Gemberling, M. Nakamura, F. G. Whitby, H. Handa, W. G. Fairbrother, and D. Tantin. 2009. A general mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and oxidative stress. *Genes Dev* 23:208-222.
11. Segil, N., S. B. Roberts, and N. Heintz. 1991. Mitotic phosphorylation of the Oct-1 homeodomain and regulation of Oct-1 DNA binding activity. *Science* 254:1814-1816.

12. Dephoure, N., C. Zhou, J. Villen, S. A. Beausoleil, C. E. Bakalarski, S. J. Elledge, and S. P. Gygi. 2008. A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci U S A* 105:10762-10767.
13. Daub, H., J. V. Olsen, M. Bairlein, F. Gnad, F. S. Oppermann, R. Korner, Z. Greff, G. Keri, O. Stemmann, and M. Mann. 2008. Kinase-selective enrichment enables quantitative phosphoproteomics of the kinome across the cell cycle. *Mol Cell* 31:438-448.
14. Olsen, J. V., M. Vermeulen, A. Santamaria, C. Kumar, M. L. Miller, L. J. Jensen, F. Gnad, J. Cox, T. S. Jensen, E. A. Nigg, S. Brunak, and M. Mann. 2010. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci Signal* 3:ra3.
15. Guelen, L., L. Pagie, E. Brasset, W. Meuleman, M. B. Faza, W. Talhout, B. H. Eussen, A. de Klein, L. Wessels, W. de Laat, and B. van Steensel. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453:948-951.
16. Crosio, C., G. M. Fimia, R. Loury, M. Kimura, Y. Okano, H. Zhou, S. Sen, C. D. Allis, and P. Sassone-Corsi. 2002. Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. *Mol Cell Biol* 22:874-885.
17. Wang, V. E. H., T. Schmidt, J. Chen, P. A. Sharp, and D. Tantin. 2004. Embryonic Lethality, Decreased Erythropoiesis, and Defective Octamer-Dependent Promoter Activation in Oct-1-Deficient Mice. *Mol Cell Biol* 24:1022-1032.
18. O'Regan, L., and A. M. Fry. 2009. The Nek6 and Nek7 protein kinases are required for robust mitotic spindle formation and cytokinesis. *Mol Cell Biol* 29:3975-3990.
19. Tsai, M. Y., S. Wang, J. M. Heidinger, D. K. Shumaker, S. A. Adam, R. D. Goldman, and Y. Zheng. 2006. A mitotic lamin B matrix induced by RanGTP required for spindle assembly. *Science* 311:1887-1893.
20. Ma, L., M. Y. Tsai, S. Wang, B. Lu, R. Chen, J. R. Iii, X. Zhu, and Y. Zheng. 2009. Requirement for Nudel and dynein for assembly of the lamin B spindle matrix. *Nat Cell Biol* 11:247-256.
21. Matsumoto, M. L., K. E. Wickliffe, K. C. Dong, C. Yu, I. Bosanac, D. Bustos, L. Phu, D. S. Kirkpatrick, S. G. Hymowitz, M. Rape, R. F. Kelley, and V. M. Dixit. 2010. K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. *Mol Cell* 39:477-484.

22. Jin, L., A. Williamson, S. Banerjee, I. Philipp, and M. Rape. 2008. Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. *Cell* 133:653-665.
23. Wu, T., Y. Merbl, Y. Huo, J. L. Gallop, A. Tzur, and M. W. Kirschner. 2010. UBE2S drives elongation of K11-linked ubiquitin chains by the anaphase-promoting complex. *Proc Natl Acad Sci U S A* 107:1355-1360.
24. Joukov, V., A. C. Groen, T. Prokhorova, R. Gerson, E. White, A. Rodriguez, J. C. Walter, and D. M. Livingston. 2006. The BRCA1/BARD1 heterodimer modulates ran-dependent mitotic spindle assembly. *Cell* 127:539-552.
25. Pujana, M. A., J. D. Han, L. M. Starita, K. N. Stevens, M. Tewari, J. S. Ahn, G. Rennert, V. Moreno, T. Kirchhoff, B. Gold, V. Assmann, W. M. Elshamy, J. F. Rual, D. Levine, L. S. Rozek, R. S. Gelman, K. C. Gunsalus, R. A. Greenberg, B. Sobhian, N. Bertin, K. Venkatesan, N. Ayivi-Guedehoussou, X. Sole, P. Hernandez, C. Lazaro, K. L. Nathanson, B. L. Weber, M. E. Cusick, D. E. Hill, K. Offit, D. M. Livingston, S. B. Gruber, J. D. Parvin, and M. Vidal. 2007. Network modeling links breast cancer susceptibility and centrosome dysfunction. *Nat Genet* 39:1338-1349.
26. Parvin, J. D. 2009. The BRCA1-dependent ubiquitin ligase, gamma-tubulin, and centrosomes. *Environ Mol Mutagen* 50:649-653.
27. Kanai, M., W. M. Tong, E. Sugihara, Z. Q. Wang, K. Fukasawa, and M. Miwa. 2003. Involvement of poly(ADP-Ribose) polymerase 1 and poly(ADP-Ribosylation) in regulation of centrosome function. *Mol Cell Biol* 23:2451-2462.
28. Roberts, S. B., N. Segil, and N. Heintz. 1991. Differential phosphorylation of the transcription factor Oct1 during the cell cycle. *Science* 253:1022-1026.
29. Martinez-Balbas, M. A., A. Dey, S. K. Rabindran, K. Ozato, and C. Wu. 1995. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* 83:29-38.
30. Hansen, D. V., J. Y. Hsu, B. K. Kaiser, P. K. Jackson, and A. G. Eldridge. 2002. Control of the centriole and centrosome cycles by ubiquitination enzymes. *Oncogene* 21:6209-6221.
31. Pohl, C., and S. Jentsch. 2008. Final stages of cytokinesis and midbody ring formation are controlled by BRUCE. *Cell* 132:832-845.
32. Harper, J. W., J. L. Burton, and M. J. Solomon. 2002. The anaphase-promoting complex: it's not just for mitosis any more. *Genes Dev* 16:2179-2206.

33. Sebastiano, V., M. Dalvai, L. Gentile, K. Schubart, J. Sutter, G. M. Wu, N. Tapia, D. Esch, J. Y. Ju, K. Hubner, M. J. Bravo, H. R. Scholer, F. Cavaleri, and P. Matthias. 2010. Oct1 regulates trophoblast development during early mouse embryogenesis. *Development* 137:3551-3560.
34. Tanaka, M., J. S. Lai, and W. Herr. 1992. Promoter-selective activation domains in Oct-1 and Oct-2 direct differential activation of an snRNA and mRNA promoter. *Cell* 68:755-767.
35. Tang, H., and P. A. Sharp. 1999. Transcriptional Regulation of the murine 3' Igh enhancer by Oct-2. *Immunity* 11:517-526.
36. Goodman, B., W. Channels, M. Qiu, P. Iglesias, G. Yang, and Y. Zheng. 2010. Lamin B counteracts the kinesin Eg5 to restrain spindle pole separation during spindle assembly. *J Biol Chem* 285:35238-35244.
37. Maiato, H., E. A. Fairley, C. L. Rieder, J. R. Swedlow, C. E. Sunkel, and W. C. Earnshaw. 2003. Human CLASP1 is an outer kinetochore component that regulates spindle microtubule dynamics. *Cell* 113:891-904.
38. Pereira, A. L., A. J. Pereira, A. R. Maia, K. Drabek, C. L. Sayas, P. J. Hergert, M. Lince-Faria, I. Matos, C. Duque, T. Stepanova, C. L. Rieder, W. C. Earnshaw, N. Galjart, and H. Maiato. 2006. Mammalian CLASP1 and CLASP2 cooperate to ensure mitotic fidelity by regulating spindle and kinetochore function. *Mol Biol Cell* 17:4526-4542.

CHAPTER 4

DYNAMIC REGULATION OF OCT1 BY GLUCOSE AND O-GlcNAc MODIFICATION

Jinsuk Kang¹ Jae-Min Lim², Hiroshi Handa³, Lance Wells² and Dean Tantin^{1,§}

¹Department of Pathology, University of Utah School of Medicine, Salt Lake City, Utah 84112, United States

²Complex Carbohydrate Research Center, Department of Biochemistry and Molecular Biology, The University of Georgia, Athens, GA 30602, United States

³Integrated Research Institute, Tokyo Institute of Technology, Yokohama, Japan

[§]Corresponding author (dean.tantin@path.utah.edu)

Running title: Regulation of Oct1 by Glycosylation

Keywords: Oct1, glucose, O-GlcNAc, *Gadd45a*, SUMO, ubiquitin

Abstract

The Oct1 transcription factor is a potent regulator of carbon metabolism and tumorigenicity. Although Oct1 is dynamically regulated by phosphorylation, the presence and importance of other Oct1 modifications is unknown. Here we show that Oct1 is modified by small ubiquitin-like modifier (SUMO) and O-linked N-acetylglucosamine (O-GlcNAc) moieties. We map two glycosylation events at positions T255 and S728 within human Oct1. These glycosylated residues are required for anchorage-independent growth, for a proper transcriptional and cellular, and for the ability of Oct1 to control cell

metabolism. We also, for the first time, identify specific sites of Oct1 ubiquitination. These findings suggest that Oct1 acts as a sensor of glucose levels and cellular adhesion.

Introduction

Oct1 is one of the archetypal POU (Pit-1, Oct1/2, Unc-86) transcription factors. It appears to be expressed in appreciable amounts in all cells and tissues (reviewed in 1), though protein levels are dynamically regulated during development and can vary between cells within a tissue (2, 3). It is related to Oct4, a master regulator of embryonic stem (ES) cell pluripotency, and has similar *in vitro* DNA binding specificity (reviewed in 4). In mouse embryonic fibroblasts (MEFs) and A549 human lung adenocarcinoma cells, loss of Oct1 has little impact on cell growth or viability in culture, but strongly antagonizes oncogenic transformation *in vitro* and tumorigenicity in xenograft assays. This suppression is at least partially metabolic in nature (5), though additional mechanisms are possible.

Although long considered a static transcription factor, the observation that Oct1 promotes resistance to genotoxic and oxidative stress suggested that Oct1 is a stress response effector (6). More recent findings show that Oct1 is dynamically regulated by phosphorylation following stress exposure (7). Some of these phospho-modifications alter Oct1 DNA binding selectivity, resulting in the induced occupancy of a different subset of target genes (7). Others block DNA binding in biochemical assays and result in exclusion from mitotic DNA (8). Proteins reported to interact with Oct1 include DNA-PK, a kinase that becomes activated in response to double-strand DNA breaks (9, 10), and poly (ADP-ribose) polymerase-1 (PARP-1), an enzyme that becomes activated by

DNA breaks and oxidative stress (11), the breast cancer associated gene BRCA1 (12, 13) and lamin B, a component of the nuclear envelope and various mitotic structures (8, 14, 15). At a transcriptional level, Oct1 can both repress and activate target gene expression, even at the same target, in response to different upstream stimuli (16).

During our investigations of Oct1 phosphorylation site mutants, we have identified O-linked glycosylation, SUMOylation and ubiquitination as other post-translational modifications. We map two Oct1 glycosylation sites, at positions T255 and S728, and two ubiquitination sites, at positions K9 and K403. We show that these modifications regulate Oct1 stability and transcriptional activity. We find that Oct1 dampens the induction of *Gadd45a* gene expression in response to acute glucose starvation. Mutation of the Oct1 glycosylation sites to alanine decreases binding to the *Gadd45a* promoter and augments *Gadd45a* induction.

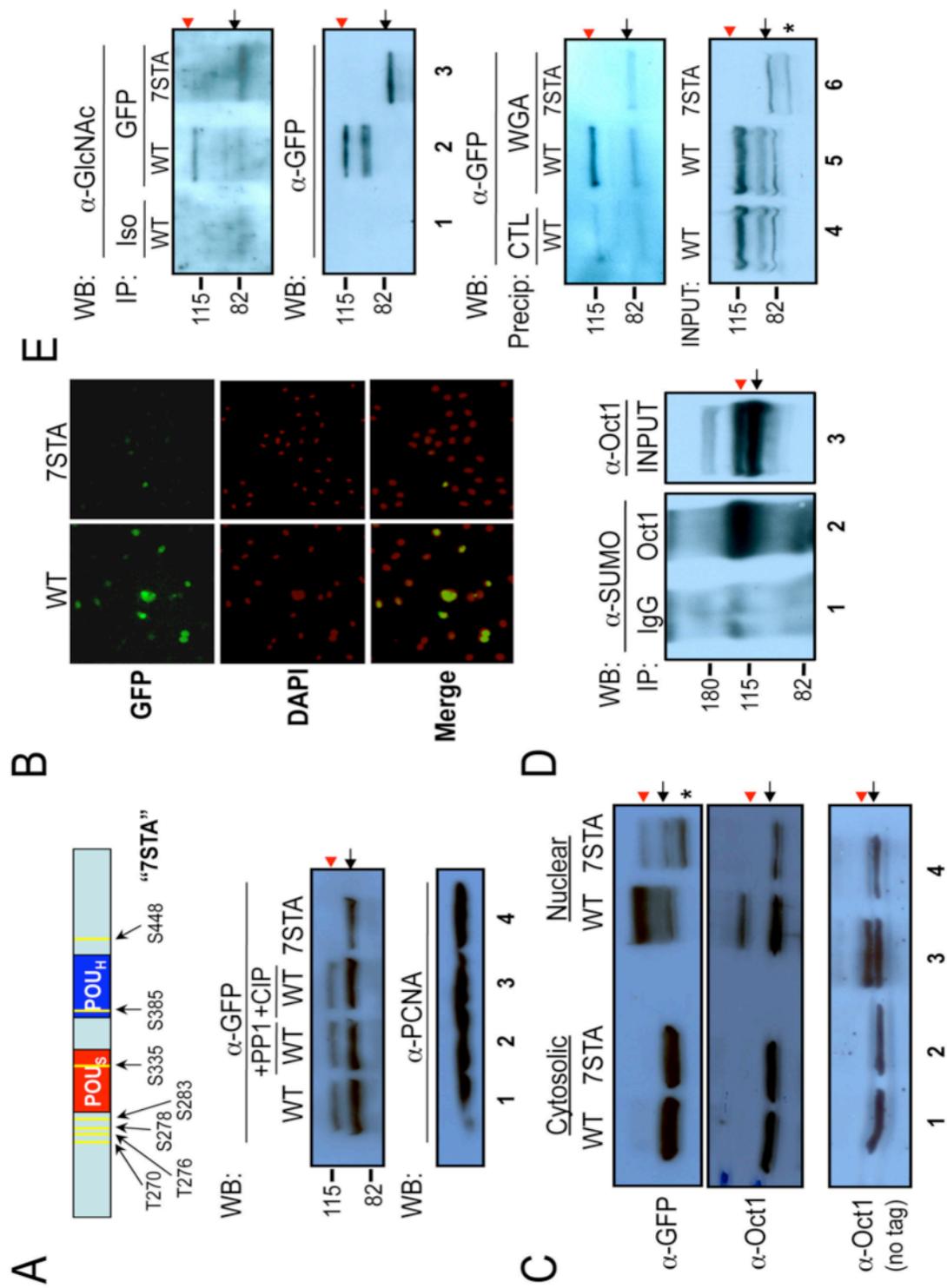
Results

Phosphorylated Oct1 serine/threonine residues control protein glycosylation and stability in the nucleus

Previous work from our laboratory has established that there are no fewer than seven Oct1 Ser/Thr phosphorylation events that occur following exposure to genotoxic and oxidative stress agents (7). The functions of these modifications, the upstream kinases, the fraction modified and the degree of co-occurrence are largely unknown. To simplify the analysis of these residues, we generated a mutant form of Oct1 in which these seven amino acids were changed to alanine (“7STA”, Figure 4.1A, see Materials

Figure 4.1. The nuclear form of Oct1 is predominantly modified by glycosylation. A)

Top panel shows schematic of the Oct1 “7STA” mutation. Bottom panel shows Western blots of Oct1-deficient 3T3-immortalized MEFs transduced with GFP-tagged Oct1 (wild-type and 7STA mutant) probed using monoclonal anti-GST (Sigma) or control anti-PCNA (Santa Cruz) antibodies. Black arrow: predicted molecular weight of GST-Oct1. Red carrot: modified form. PP1=protein phosphatase 1; CIP=calf intestinal alkaline phosphatase. B) Immunofluorescence images of Oct1 deficient 3T3 cells expressing wild-type and 7STA protein. Images were stained with anti-GFP antibodies, counterstained with DAPI, and imaged. C) Material from cytosolic and nuclear fractionation of MEFs was analyzed by Western blotting using monoclonal GFP and rabbit Oct1 (Bethyl) antibodies. Top panels are from cells transduced with a GFP-Oct1 fusion protein. Bottom panel was from cells transduced with untagged protein. D) Left panel: HeLa cell nuclear extracts were immunoprecipitated with anti-Oct1 antibodies or IgG isotype control. Following immunoprecipitation, the products were Western blotted using anti-SUMO antibodies. Right panel: input nuclear extract was Western blotted with anti-Oct1 antibodies. E) HeLa cells were transduced with GFP-Oct1 fusion protein and either immunoprecipitated using anti-GFP antibodies (using rabbit IgG isotype controls) or precipitated using wheat germ agglutinin-agarose (WGA, using glutathione-agarose controls). Western blots were probed using anti-O-GlcNAc or anti-GFP antibodies. IP: immunoprecipitation, precip: precipitation.



and Methods). We then restored Oct1 function in Oct1 deficient 3T3-immortalized MEFs (17) using retroviral expression constructs encoding either wild-type or 7STA Oct1. To more easily monitor Oct1 expression and localization, Oct1 was also tagged with an N-terminal GFP moiety.

Expression of GFP-tagged Oct1 resulted in major and minor bands as determined by Western blotting whole cell extracts with GFP antibodies (Figure 4.1A, lane 1). The major, faster-migrating form (black arrow) corresponded to the predicted molecular weight of GFP-Oct1. The minor, slow-migrating form (red arrow) displayed an apparent molecular weight of ~120 kDa and was interpreted to be a modified species. The Oct1 cDNA encoding the 7STA mutant eliminated the modified form (lane 4). We initially favored the idea that the change in migration was due to *in vivo* phosphorylation events. We therefore treated the wild-type extract with two different protein phosphatases (protein phosphatase 1 and calf intestinal alkaline phosphatase). Surprisingly, no change in migration was observed (lanes 2 and 3), suggesting that the altered migration pattern results from some other modification, although one likely dependent on phosphorylation.

We also noted that Oct1 protein levels in the extracts from cells infected with the 7STA virus were reduced relative to wild-type (Figure 4.1A, compare lane 4 with lane 1). To confirm this finding, we performed indirect immunofluorescence using the transduced Oct1 deficient MEFs and anti-GFP antibodies (Figure 4.1B). Decreased Oct1 protein expression was noted, particularly in the nucleus. We therefore generated nuclear and cytosolic fractions (Figure 4.1C). No difference between cytosolic wild-type and 7STA Oct1 was noted (Figure 4.1C, top panel, lanes 1 and 2). However large changes were identified in the nuclear fraction (lanes 3 and 4). All of the modified form identified in

Figure 4.1A resided in the nuclear fraction, and correspondingly the nuclear pool of Oct1 was mostly modified. This form was largely eliminated in the 7STA mutant. Further, the 7STA form was expressed at lower levels in the nucleus and included degraded species (asterisk). We conclude that one or more of the phosphorylation site residues are required for nuclear Oct1 modification and stability. To confirm that the observed signals were due to tagged Oct1, we stripped the blot and reprobbed with an Oct1 antibody (Figure 4.1C, “a-Oct1”), generating similar results. To confirm that the results were not due to the presence of the GFP tag, we repeated the experiment using untagged retroviral constructs and anti-Oct1 Western blots (“a-Oct1 (untagged)”). Similar results were observed. These findings indicate that one or more of the seven mutated serine/threonine residues regulate nuclear protein stability.

Given that the slow-migrating form of Oct1 identified above did not appear to result directly from phosphorylation, we investigated additional post-translational modifications. We found that in HeLa cells, the high molecular weight but not low molecular weight form of endogenous Oct1 was SUMOylated (Figure 4.1D). We also identified O-linked glycosylation as an additional modification. Nuclear extracts from human HeLa cells expressing GFP-tagged versions of Oct1 (wild-type or 7STA) were subjected to immunoprecipitation with antibodies directed against GFP, or precipitation using wheat germ agglutinin-agarose (WGA, Figure 4.1E). Using HeLa cells infected with wild-type Oct1-encoding virus, immunoprecipitation with anti-GFP antibodies and Western blotting using anti-O-GlcNAc antibodies, we identified a band whose size corresponded to the modified form of Oct1 (Figure 4.1E, lane 2). A fast-migrating glycosylated form was identified using the 7STA mutant (lane 3), indicating that

glycosylation was not eliminated using the 7STA mutant. Similar results were observed by precipitation using WGA (lane 6). These findings indicate that Oct1 is both SUMOylated and O-glycosylated, and suggest that SUMO is the underlying reason for the decreased mobility of wild-type relative to 7STA Oct1.

Mapping of Oct1 glycosylation sites

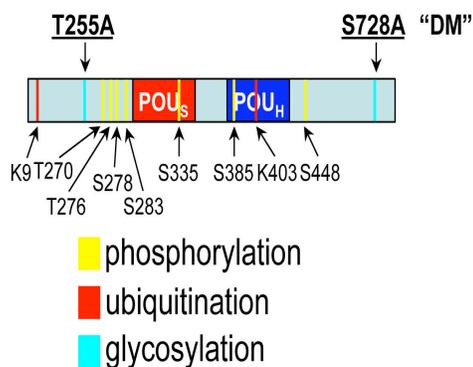
Using latex nanoparticle affinity purification (7), and mass spectrometry methodologies established by Hart and colleagues (18), we definitively mapped two Oct1 glycosylation sites in HeLa cells cultured under normal conditions, at T255 (using trypsin) and S728 (using chymotrypsin) (Figure 4.2A). This analysis also identified multiple peptides consistent with PARP-1 (not shown), which has previously been reported to interact with Oct1 (11), as well as two previously identified Oct1 phosphorylation sites (7), at T276 and S448 (Figure 4.2A). We also identified two ubiquitination sites, at K9 (with >95% confidence), and K403 (with ~75% confidence, Figure 4.2A). We engineered point mutations at the two glycosylation sites (T255A/S728A, “DM”, Figure 4.2B) in the context of the GFP fusion retroviral vector, and expressed this mutant version and wild-type control in Oct1 deficient MEFs. Both constructs were expressed equally (Figure 4.2C, bottom panel). There was a significant, albeit incomplete, decrease in the ability to enrich Oct1 using WGA. We conclude that mutation of T255 and S728 eliminates 66-75% of Oct1 glycosylation.

Figure 4.2. Identification of Oct1 O-linked glycosylated. A) Oct1 (POU2F1) sequence and identified trypsin and chymotrypsin peptides identified by mass spectroscopy. The complete sequence of human Oct1 (NP-002688, UniProtKB/Swiss-Prot: P14859) is shown. Peptides obtained using digestion with trypsin are shown in red. Peptides obtained using chymotrypsin are shown in purple. Glycosylated serine and threonine residues (S_g and T_g) are shown in green. Identified phosphorylation and ubiquitination sites are designated S_p , T_p , and K_{ub} . B) Schematic of the known Oct1 modifications. Double point mutant eliminating the glycosylation sites (“DM”) is show above. C) HeLa cells were transduced with wild-type or DM GFP-Oct1 fusion protein and precipitated using WGA or glutathione-agarose controls. Enrichment was detected by Western blotting for GFP. Input GFP Western blot control is shown beneath. EV: empty vector.

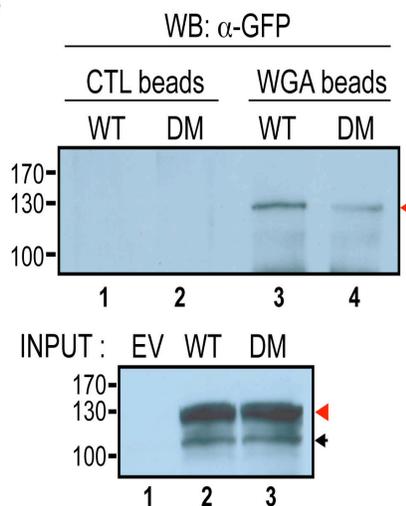
A

1 **MNNPSETSK_{ub}⁹PSMESGDGNTGTQTNGLDFQKQPVPVGGAIStaQAQAFGLHLHQ**
 54 **VQLAGTSLQAAAQSLNVQSKSNEESGDSQQPSQPSQQPSVQAAIPQTQLMLAGGQITGLTLTPAQQLLLQQAQAQA**
 131 **QLLAAAVQQHSASQQHSAAGATISASAATPMTQIPLSQPIQIAQDLQQLQQLQQNLNLQQFVLVHPTTNLQPAQFIISQ**
 211 **TPQGQQGLLQAQNL_{LTQLPQQSQANLLQSQPSITLTSQPATPTR}T_g²⁵⁵IAATPIQTLPOSQST_p²⁷⁰PKRID**
 276 **T_p²⁷⁶PS_p²⁷⁸LEEPS_p²⁸³DLEELEQFAKTFKQRRIKLGFTQGDVGLAMGKLYGNDFSQTT**
 326 **ISRFEALNLS_p³³⁵FKNMCKLKPLLEKWLNDAENLSSDSSLSSPSALNSPGIEGLSRRR**
 381 **KKRTS_p³⁸⁵IETNIRVALEKSFLNQK_{ub}⁴⁰³PTSEEITMIADQLNMEKEVIRVWFCNRRQ**
 433 **KEKRINPPSSGGTSSS_p⁴⁴⁸PIK_{AIFPSPTSL}VATTPSLVTSSAATTLTVSPVLPLTSAAVTNLSVTGTSDTTSNN**
 515 **TATVISTAPPASSAVTSPSLSPSPASASTSEASSASETSTTQTTSTPLSSPLGTSQVMVTASGLQTA AAAALQGAAQLPANASLAAMAAAAG**
 608 **LNPMLMAPSQFAAGGALLSLNPGTLSGALSPALMSNSTLATIQALASGGSLPITSLDATGNLVFANAGGAPNIVTAPLFLNPQNLSSLTSNP**
 690 **VSLVSA AAAASAGNSAPVASLHATSTSAESIQNSLF_{TVAS_g}⁷²⁸ASGAASTTTTASKAQ**

B



C

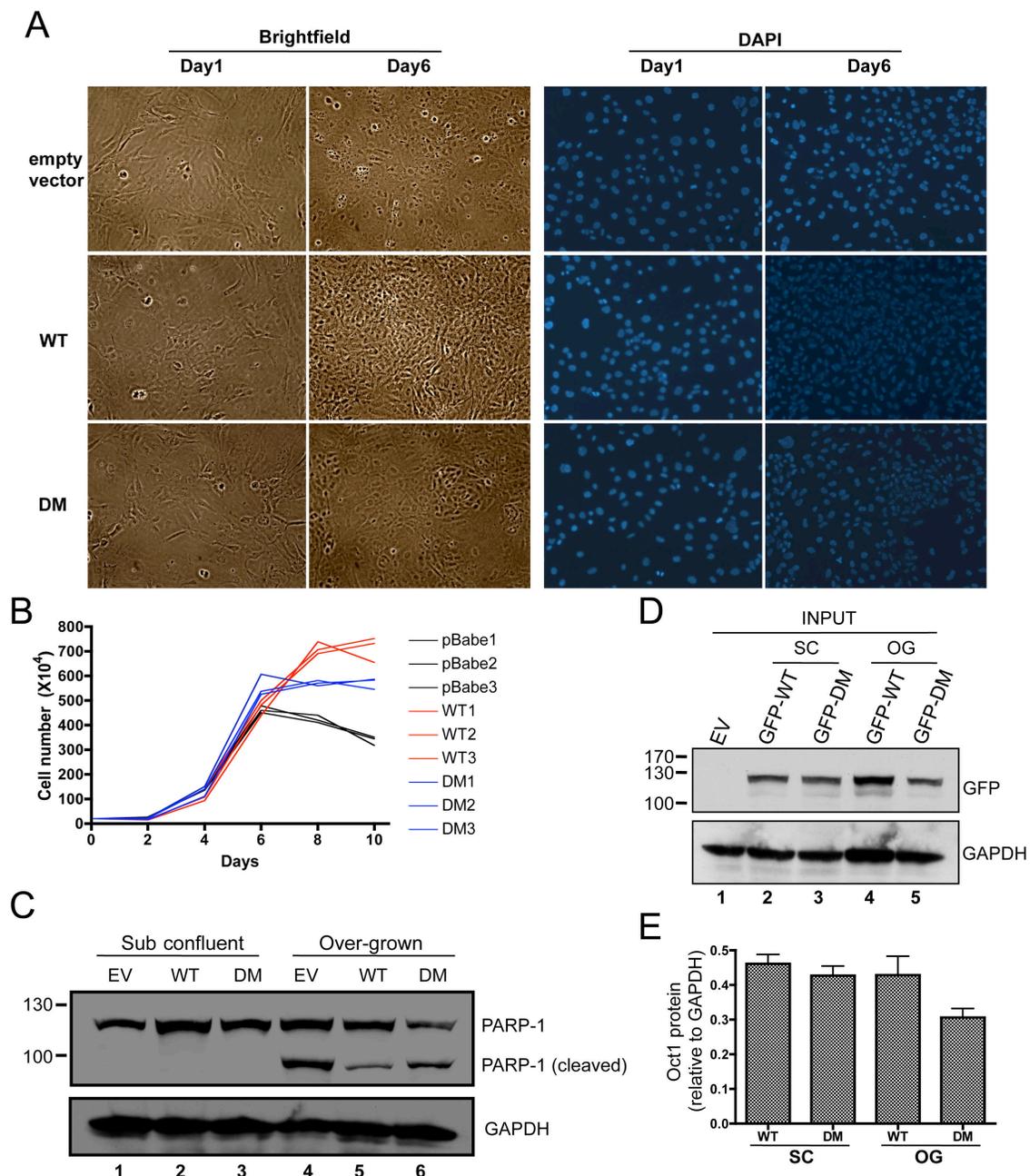


Glycosylated Oct1 serine/threonine residues promote survival in anchorage-independent growth conditions

We tested primary Oct1 deficient MEFs stably transduced with wild-type or DM Oct1 for their effect on cell growth and viability. No gross differences were observed between Oct1 deficient or complemented MEFs under standard culture conditions. However significant differences were manifested following long-term culture (not shown). To study changes resulting from wild-type or mutant GFP-Oct1 in cells that could be propagated indefinitely, we used Oct1 deficient MEFs that were additionally deficient in p53 (5). Primary early passage MEF isolates from *Oct1*^{-/-}, *p53*^{-/-} double-deficient embryos were transduced with retrovirus, selected with puromycin, and tested immediately. As with Oct1 deficient p53 wild-type cells, no gross differences were observed under normal culture conditions (Figure 4.3A, left columns). However, following 6 days of continuous culture we found that p53 deficient MEFs lacking Oct1 remained as a confluent monolayer whereas cells complemented with Oct1 became overconfluent. Cells complemented with DM Oct1 largely failed to grow to overconfluence (Figure 4.3A, right columns). To accurately quantify these data, we counted three independently transduced sets of primary MEFs 2, 4, 6, 8 and 10 days following seeding. *Oct1*^{-/-}, *p53*^{-/-} MEFs transduced with Oct1 reached cell numbers nearly double those of cells complemented with empty vector, while those complemented with DM showed an intermediate phenotype (Figure 4.3B).

The above results could be explained either by a failure to proliferate without attachment to the underlying substratum, and/or by the induction of apoptosis upon loss of adherence (anoikis), in the absence of functional Oct1. We noticed that many dead and

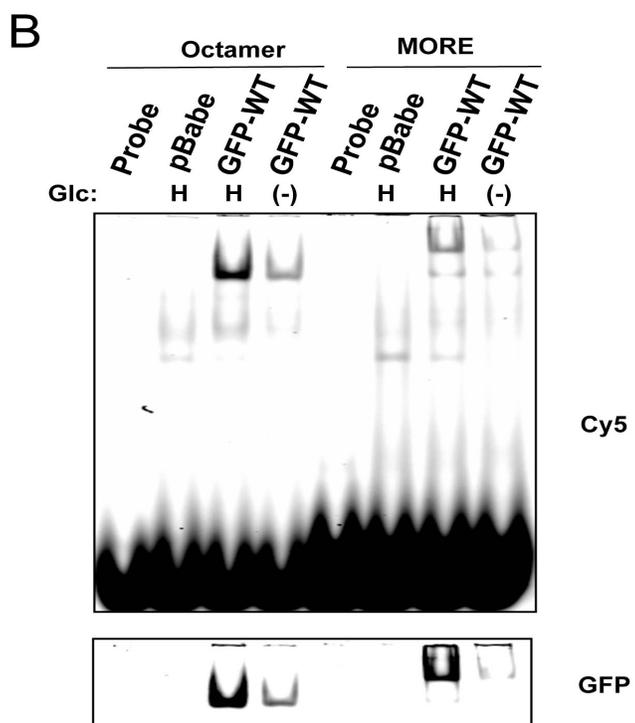
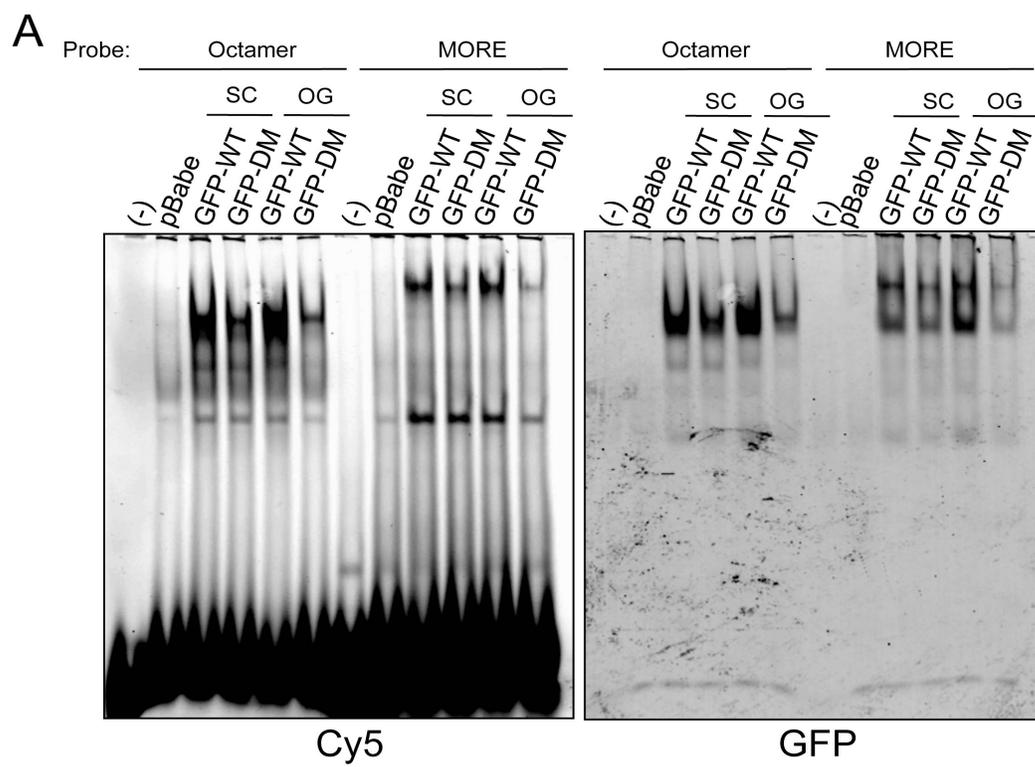
Figure 4.3. Oct1 supports anchorage-independent growth in p53 deficient MEFs. A) Brightfield and DAPI fluorescence images are shown of *p53*^{-/-}, *Oct1*^{-/-} double-deficient MEFs complemented with empty vector, wild-type Oct1, or DM Oct1 with alanine substitutions at the two identified glycosylation sites. 2×10^5 cells were seeded at a density of 2×10^4 cells/cm², and images taken at days 1 and 6. B) 2×10^5 cells were seeded at a density of 4000 cells/cm² and collected every other day for 10 days. Cell numbers are shown for triplicate transduction experiments. C) The induction of apoptosis was measured using whole cell extracts from cells prepared under conditions identical to (B) using antibodies reactive against full-length and cleaved PARP-1 (Cell Signaling). Murine monoclonal GAPDH antibodies (Millipore) were used as a loading control. EV: empty vector. D) Nuclear extracts were prepared from complemented MEFs grown under subconfluent (SC) or over-growth (OG, incubated 6 days longer) conditions, and used in Western blots with anti-GFP antibodies and anti-GAPDH loading controls. E) Quantification from two such experiments. Levels were normalized to GAPDH band intensity using Image J software (NIH).



floating cells were present in the Oct1 deficient condition specifically in over-growth conditions, suggesting that the lack of Oct1 confers a susceptibility to apoptosis in the absence of 2D attachment. To test this possibility, we studied apoptosis in these conditions using a cleaved PARP-1 assay. Cleaved PARP-1 was not observed in *Oct1*^{-/-}, *p53*^{-/-} MEFs in sub-confluent conditions (Figure 4.3C, lanes 1-3). In contrast, cleaved PARP-1 was detectable in over-growth conditions, and was much more prominent in empty vector-transduced cells compared to those transduced with wild-type Oct1 (lanes 4-5). Cells complemented with DM Oct1 displayed an intermediate phenotype (lane 6). These results indicate that Oct1 controls aspects of anchorage-dependent survival, and that glycosylated residues within Oct1 are partially responsible for this phenotype.

To begin to understand the molecular basis of the above findings, we analyzed Oct1 levels and activity in complemented MEFs using Western blots and electrophoretic mobility shift assays (EMSA). We prepared whole cell extracts from *Oct1*^{-/-}, *p53*^{-/-} MEFs transduced with empty vector, or wild-type or mutant GFP-Oct1. Cells were harvested from subconfluent or over-confluent plates incubated 6 days longer. Oct1 levels were assessed using Western blotting with GFP and Oct1 antibodies. Relative to a GAPDH loading control, we observed a significant decrease in DM Oct1 protein in over-confluent cells relative to wild-type Oct1 and relative to cells grown under subconfluent conditions (Figure 4.3D). The effect was relatively small, but consistent across multiple experiments (Figure 4.3E). EMSA confirmed the decreased Oct1 activity present in DM cells grown in over-confluent conditions (Figure 4.4A) as well as conditions of glucose withdrawal (Figure 4.4B). However upon quantification it was apparent that there was no further

Figure 4.4. Electrophoretic mobility shift assay comparing WT and DM in binding capacity. Cy5 labeled octamer and MORE probes are used. A) Nuclear extracts were prepared from MEFs under subconfluent and over-grown conditions. DNA bound GFP-Oct1 and DM were observed by two different channels, Cy5 and GFP. B) MEFs were grown under no glucose condition for 24 hr, and analyzed as experiment (A).

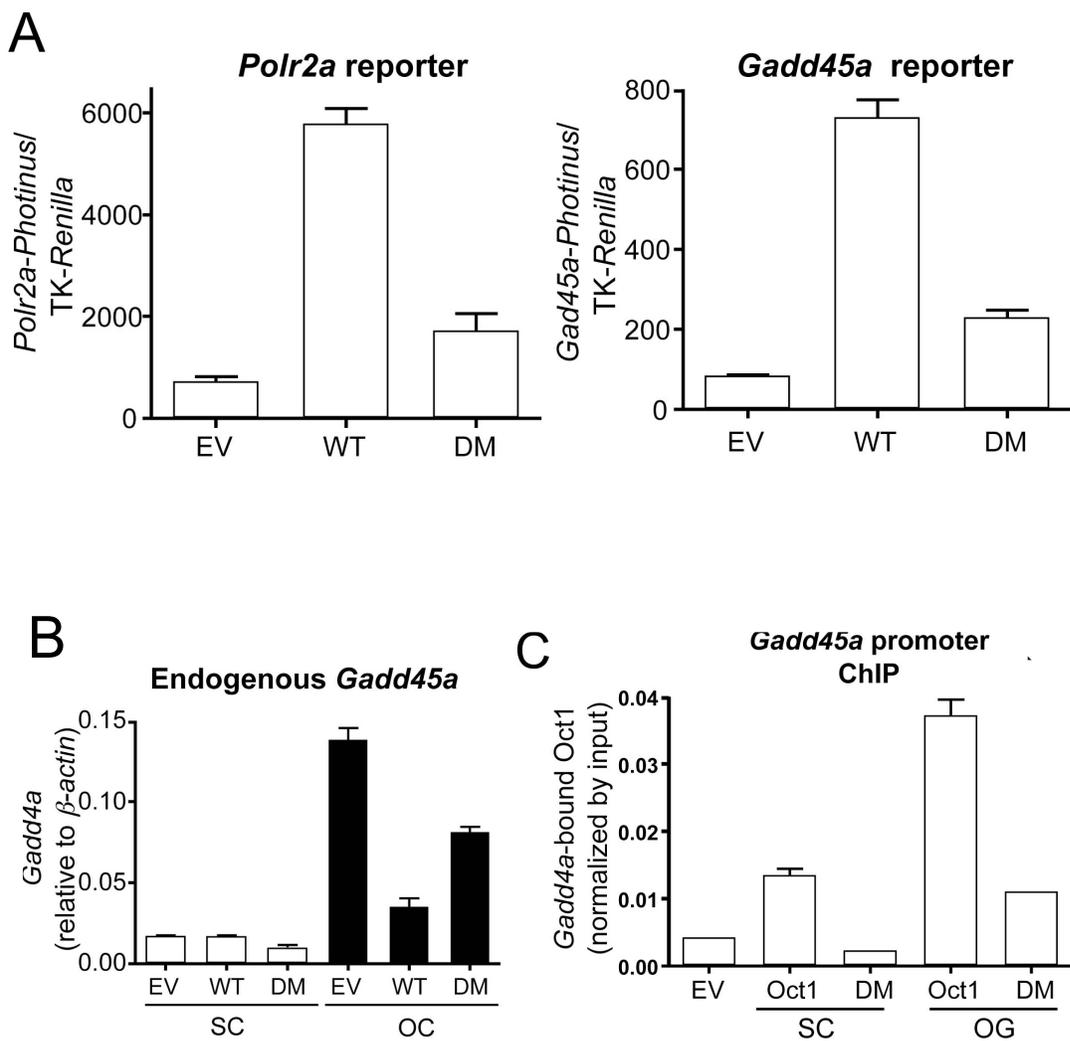


decrease in Oct1 DNA binding activity over the decreased Oct1 levels (not shown).

Glycosylated Oct1 residues control transcriptional activity

We transiently transfected complemented *Oct1^{-/-}, p53^{-/-}* MEFs with firefly luciferase reporter constructs responsive to Oct1, co-transfecting TK-*Renilla* luciferase as an internal control. The *Polr2a* upstream regulatory region contains four Oct1 binding sites, and the *Gadd45a* promoter contains two perfect octamers spaced 19 bp apart. In both cases, we found that the presence of wild-type Oct1 strongly induced transcription activity, while the double mutant was defective (Figure 4.5A). We therefore studied the endogenous *Gadd45a* locus in more detail using complemented *Oct1^{-/-}, p53^{-/-}* MEFs and qRT-PCR. GADD45a is inducible by growth arrest and DNA damage (19). p53 and Oct1 are thought to be the principal mediators of this induction (12, 19, 20). However GADD45a induction during anchorage-independent growth and the role of Oct1 and p53 were not explored. We found that *Gadd45a* was strongly induced in over-growth conditions in a manner that does not require p53 or Oct1 (Figure 4.5B). Further, complementing *Oct1^{-/-}, p53^{-/-}* cells with a retrovirus containing wild-type Oct1 significantly blunted the inductions, indicating that under these conditions, Oct1 acts in a repressive mode at *Gadd45a*. Again, the double glycosylation site mutant showed an intermediate phenotype (Figure 4.5B). A p53 target gene not thought to be regulated by Oct1 (p21/WAF1) and a housekeeping Oct1 target (Ahcy, 16) were not induced under any conditions (with Ahcy remaining active and p21 remaining silent), indicating that the *Gadd45a* gene expression changes were relatively specific (Figure 4.6).

Figure 4.5. Glycosylated residues control Oct1 transcription activity. A) $p53^{-/-}, Oct1^{-/-}$ MEFs were transiently transfected with 1.0 μ g *Polr2a*-TK-pGL2 (7), or *Gadd45a*-pGL3 (see materials and methods), and 100 ng internal control TK-*Renilla* plasmid DNA. Firefly luciferase reporter activity was measured relative to *Renilla* control using a dual luciferase assay (Promega). Experiments were performed in triplicate and error bars depict standard deviation. B) *Gadd54a* mRNA levels were measured in complemented $p53^{-/-}, Oct1^{-/-}$ MEFs relative to b-actin in normal (SC) and overgrowth (OG) conditions using qRT-PCR with intron-spanning primers. Experiments were performed in triplicate and error bars depict standard deviation. C) Quantitative Oct1 ChIP enrichment in $p53^{-/-}, Oct1^{-/-}$ MEFs complemented with wild-type or DM Oct1. Experiments were performed in subconfluent (“SC”) or over-growth (“OG”) conditions cultured for an additional 6 days.



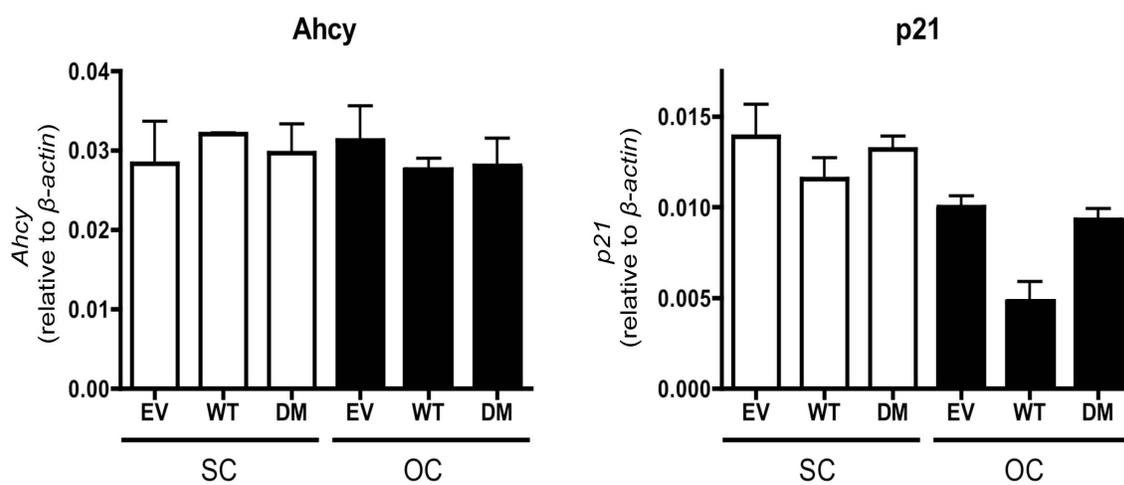
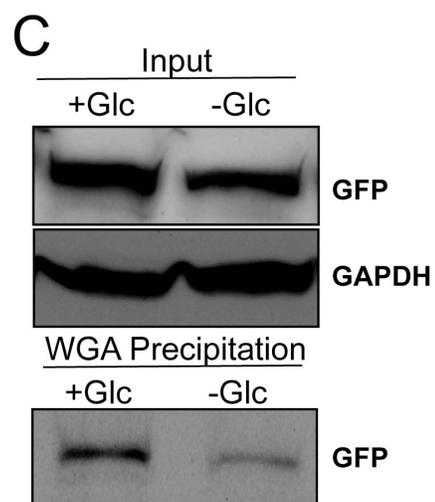
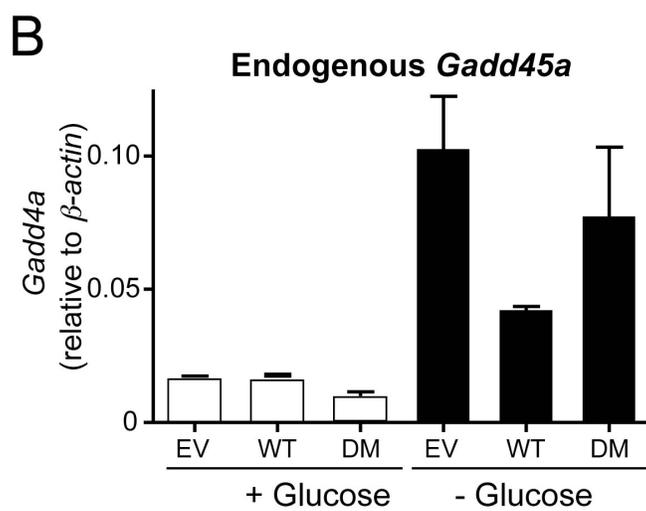
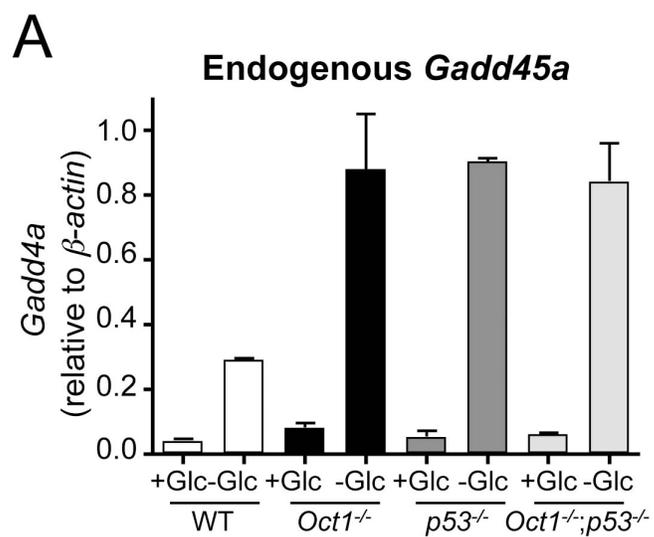


Figure 4.6. Transcription level of Oct1 control target genes. The level of Ahcy and p21 expression is little changed in over-grown MEFs.

We have shown that the double point mutation reduces steady-state protein levels (Figure 4.3D,E). To determine whether or not the decreased amount of protein translated to less bound protein at the *Gadd45a* promoter, we performed quantitative ChIP assays. Bound Oct1 could be detected at *Gadd45a* under normal growth conditions using *Oct1*^{-/-}, *p53*^{-/-} cells complemented with wild-type Oct1. In contrast, no bound Oct1 was detected using cells complemented with DM Oct1 (Figure 4.5C, “SC”). Plates grown for an additional six days (“OG”) displayed a strong induction of Oct1 binding. This binding presumably dampens *Gadd54a* induction. Again, binding was reduced in the DM condition. These results indicate that Oct1 behaves dynamically at *Gadd54a* with binding being inducible by over-growth conditions in a manner controlled by glycosylated Oct1 residues.

Recently, *Gadd45a* was found to be induced by low glucose conditions in a manner promoted by Oct1 (21). We have shown that Oct1 is required for the normal sensitivity of murine fibroblasts to glucose withdrawal (5). To identify transcriptional effects underlying this phenomenon, we tested endogenous *Gadd45a* induction in conditions of low glucose (minus glucose for 24 hr) using uncomplemented primary MEFs prepared from mouse embryos with different genotypes: wild-type, Oct1 deficient, p53 deficient and double deficient. We found that wild-type MEFs inefficiently induced *Gadd45a* under low-glucose conditions (Figure 4.7A). Cells lacking either Oct1 or p53 showed stronger induction, indicating that both proteins can act in a negative capacity at this locus. To test the effect of the glycosylated residues on Oct1 repression of endogenous *Gadd45a* induction in response to glucose withdrawal, we complemented *p53*^{-/-}, *Oct1*^{-/-} double-deficient MEFs with retrovirus encoding wild-type or DM Oct1. In the absence of

Figure 4.7. Regulation of Oct1 activity by glucose. A) Induction of *Gadd45a* was measured as in Figure 4.4B, except wild-type, Oct1 deficient, p53 deficient or double deficient MEFs were used, and the cells were cultured in normal media with 4.5 g/L glucose (“+Glc”) or 0 mM glucose for 24 hr (“-Glc”). Experiments were performed in triplicate and error bars depict standard deviation. B) *p53*^{-/-}, *Oct1*^{-/-} MEFs complemented with wild type or DM Oct1 were cultured in 4.5 g/L glucose or no glucose. Endogenous *Gadd45a* induction was measured using qRT-PCR. Experiments were performed in triplicate and error bars depict standard deviation. EV: empty vector. C) *p53*^{-/-}, *Oct1*^{-/-} MEFs complemented with wild-type Oct1 were cultured under normal conditions or in media lacking glucose for 24 hr. Input material was Western blotted using GFP and GAPDH antibodies to normalize Oct1 levels, following which the same amounts of material were precipitated using WGA-agarose. The amount of recovered protein was detected using GFP antibodies.



either p53 or Oct1, *Gadd45a* was strongly induced, as expected (Figure 4.7B). Wild-type Oct1 blunted the induction, whereas DM Oct1 largely failed to do so. These results indicate that the glycosylated residues are required for full Oct1 transcriptional activity.

Oct1 controls aspects of metabolism following glucose withdrawal

Previous findings from our laboratory indicate that Oct1 controls metabolic parameters in MEFs grown under normal glucose conditions, including a dampening of mitochondrial function. Oct1 deficiency induces a series of metabolic alterations that include decreases in lactate and NADH (relative to NAD^+) and increases in ATP (5). The MTS assay provides a colorimetric readout of cell viability and metabolism through the conversion of a reactive dye to a colored compound that absorbs at 500 nm. The conversion is mediated by mitochondrial reductases (22). Reductase activity, and hence the amount of colored product, is directly proportional to the number of viable cells. However, on a per cell basis the assay can also be used to detect metabolic alterations (23). We plated *p53*^{-/-}, *Oct1*^{-/-} double-deficient MEFs complemented with wild-type or DM Oct1 in normal media and followed MTS activity over 6 days in culture (Figure 4.8A). Wild-type complemented cells displayed significantly more MTS activity as compared to cells transfected with empty vector (“EV”). DM Oct1 failed to complement, behaving identically to the empty vector in terms of ability to augment MTS activity. These findings indicate the metabolic differences identified by Shakya et al. (5) are extendable *p53*^{-/-}, *Oct1*^{-/-} MEFs and to MTS assays. The findings further indicate that

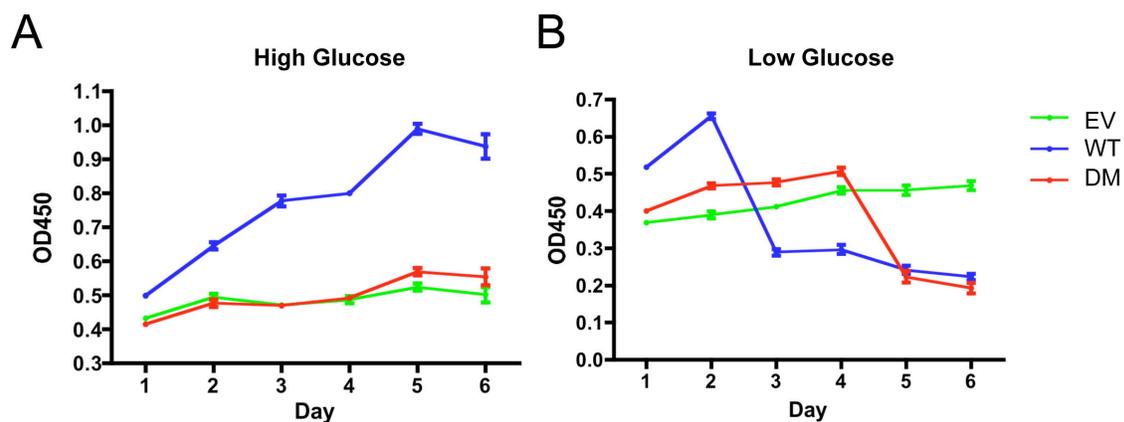


Figure 4.8. Glycosylated Oct1 residues are critical for Oct1-mediated control of mitochondrial reducing activity but dispensable for the ability of Oct1 to confer sensitivity to long-term glucose withdrawal.

A) MTS assay of $p53^{-/-}, Oct1^{-/-}$ MEFs.

MEFs were plated at 2×10^4 cells were plated into each well of a 6 well plate and MTS readings were taken every for 6 days. B) Similar assay except cells were plated in

mediate lacking glucose. EV: empty vector

glycosylated Oct1 residues are critically important for the ability of Oct1 to modulate metabolism.

We and others have shown that Oct1 loss of function also results in resistance to cell death following persistent glucose withdrawal (5, 21). We recapitulated this finding using the MTS assay: in glucose-free media Oct1 complemented cells initially showed activity as before, but after two days lost viability (Figure 4.8B). Interestingly, although cells complemented with DM Oct1 failed to complement the Oct1-mediated metabolic phenotype (DM cells track with empty vector the first few days), they were competent to induce sensitivity to glucose withdrawal (DM cells track with wild-type at late timepoints). These results show that p53 deficient cells lacking Oct1 are similarly resistant to glucose withdrawal, and functionally uncouple enzymatic reduction of the MTS target molecule in mitochondria from sensitivity to glucose withdrawal.

Discussion

Oct4, a transcription factor closely related to Oct1, was recently found to be glycosylated in human ES cells (24), and to interact with O-linked GlcNAc transferase (OGT), the enzyme that catalyzes O-linked glycosylation (25, 26). These findings lead us to test whether Oct1 is also a glycoprotein. We identify two glycosylated Oct1 residues, T255 and S728. The mass spectrometry peptide coverage was relatively low (23%), and mutation of these residues results in only an $\sim 2/3$ diminution of glycosylation, indicating that other residues may also be glycosylated. Nevertheless, we show that T255 and S728 are required for the normal induction of *Gadd45a* to glucose starvation and for survival anchorage-independent conditions. Multiple transcription factors have been found to be

glycosylated, with known examples including p53, NF- κ B, and AP-1. In many of these cases, intramolecular communication between glycosylation and phosphorylation has been reported (27-29), most often taking the form of interference in which the same residue is alternatively phosphorylated or glycosylated. The glycosylation sites identified here are not known to be phosphorylated, and to date we have not identified intramolecular communication between the glycosylation sites and the seven phosphorylated Ser/Thr residues.

We also show that Oct1 is ubiquitinated (at K9 and K403). Oct1 phosphorylation target sites are important for nuclear protein stability, and this dynamic control of protein levels may be mediated by the ubiquitination events, both of which are associated with PEST motifs: K9 with a weak site and K403 with a stronger site (not shown). The third novel Oct1 modification we identify in this study is SUMOylation. SUMO has been found to regulate a large set of molecular processes, including transcriptional regulation, protein stability, stress response, nuclear-cytosolic transport, apoptosis, and cell cycle progression (30). We anticipate that it is the SUMO moiety that causes the shift in molecular weight blocked by mutation of the seven phosphorylated serine/threonine residues (7STA). We found that relative to the cytoplasmic pool, Oct1 in the nucleus was in a predominantly modified form, and that the 7STA mutant results in nuclear protein degradation. Oct1 is known to interact with BRCA1 (12, 13). One hypothesis is that the associated BRCA1 ubiquitin ligase activity is important for the observed changes in Oct1 stability.

The glycosylated residues regulate the Oct1 transcriptional response to glucose withdrawal. Double point mutation to alanine eliminating glycosylation at these two

residues blunted Oct1 transcription potential: either positive potential in transient transfection assays, or repressive potential of the endogenous *Gadd45a* gene locus. Although others have identified an activation function for Oct1 at *Gadd45a* (19, 20), our findings suggest a repressive function under conditions of glucose withdrawal and anchorage-independent growth. Oct1 is known to be capable of repressing and activating the same gene target (16). Following glucose withdrawal, we show that O-GlcNAc moieties on Oct1 are depleted, and that the protein is less active, similar to the point mutant. This model explains why *Gadd45a* is induced following glucose withdrawal, is even more strongly induced in the absence of Oct1, and why the double glycosylation site point mutant partially phenocopies the empty vector control for *Gadd45a* induction in response to glucose withdrawal. We also find that the double mutant fails to complement the Oct1 metabolic function, but interestingly the mutant is fully competent at restoring Oct1-mediated sensitivity to long-term glucose withdrawal. These findings uncover a dynamic and complex regulatory circuitry involving Oct1, glucose levels and target genes such as *Gadd45a*.

Materials and Methods

Cloning: The 7STA mutant was generated by site directed mutagenesis (Stratagene) of the wild-type Oct1 cDNA cloned into the pBabe retroviral vector. The same primer sets for S335A and S385A were used (Kang et al. 2009). The other primer sets are as follows: T270A, T276A, S278A, S283A, S448A (complementary primers are not shown: 5'-TTCCACAGAGCCAGTCAGCACCAAAGCGAATTGAT, 5'-GTCAACACCAAAGCGAATTGATGCTCCCAGCTTGG, 5'-

CGAATTGATGCTCCCGCCTTGGAGGAGCCCAG, 5'-
CTTGGAGGAGCCCGCTGACCTTGAGGAGCT, 5'-
TGGTGGGACCAGCAGCGCACCTATTAAAGCAATT). An ~1 kb human *Gadd45a*
promoter region was amplified using the following restriction site-containing primer set:
5'-GGTACCAAGCTTAGGGCATATCGAGAGCATTTT (*KpnI*), 5'-
GAGCTCGGGCTCCTCCTCCTGTGCCA (*SacI*). This amplicon was ligated into the
pGL3 vector (Promega) via the *KpnI* and *SacI* sites. The *Polr2a*-pGL3 reporter construct
has been described previously (7).

Indirect immunofluorescence: Immunofluorescence assays were performed
identically to Kang et al. (8).

Immunoprecipitation and WGA precipitation: For immunoprecipitation, cells were
lysed with whole cell lysis buffer (18mM Hepes pH 7.9, 150mM NaCl, 1mM EDTA,
1mM EGTA, 1% Triton-X-100, protease inhibitors (Roche, PIs), and phosphatase
inhibitors (Roche, PhIs)). 500ug of extracts were incubated with 4ug of antibody in IP
buffer (50mM Tris pH8.0, 20% glycerol, 0.5mM EDTA, 0.1% NP-40, 1mM DTT, Pis
and PhIs) for overnight (O/N) at 4° C. Using magnetic beads (Activmotif), protein-
antibody complex was precipitated and washed three times with IP buffer.

Immunoprecipitated proteins were analyzed by SDS-PAGE and Western Blotting. For
WGA precipitation, Succinyl WGA Gel was purchased from EY laboratory.

Experimental steps and buffers were used as IP experiment.

MTS assays: were performed using a Kit (Promega) as per the manufacturer's
instructions.

Electrophoretic mobility shift assay: Assays were performed identically to Kang et al. (7).

Acknowledgements

We thank D. McClain for critical reading of the manuscript. We thank J. Shaw for use of laboratory equipment and technical advice, and D. McClain for advice and reagents. This work was funded by the American Cancer Society (GMC-115196, to DT).

References

1. Ryan, A. K., and M. G. Rosenfeld. 1997. POU domain family values: flexibility, partnerships, and developmental codes. *Genes Dev* 11:1207-1225.
2. Donner, A. L., V. Episkopou, and R. L. Maas. 2007. Sox2 and Pou2f1 interact to control lens and olfactory placode development. *Dev Biol* 303:784-799.
3. Maddox, J., S. South, D. Shelton, S. Chidester, A. Shakya, Y. S. DeRose, A. E. Welm, D. A. Jones, B. E. Welm, and D. Tantin. Transcription Factor Oct1 is a Somatic and Cancer Stem Cell Determinant. *Submitted*.
4. Kang, J., A. Shakya, and D. Tantin. 2009. Stem cells, stress, metabolism and cancer: a drama in two Octs. *Trends Biochem Sci* 34:491-499.
5. Shakya, A., R. Cooksey, J. E. Cox, V. Wang, D. A. McClain, and D. Tantin. 2009. Oct1 loss of function induces a coordinate metabolic shift that opposes tumorigenicity. *Nat Cell Biol* 11:320-327.
6. Tantin, D., C. Schild-Poulter, V. Wang, R. J. Hache, and P. A. Sharp. 2005. The octamer binding transcription factor Oct-1 is a stress sensor. *Cancer Res* 65:10750-10758.
7. Kang, J., M. Gemberling, M. Nakamura, F. G. Whitby, H. Handa, W. G. Fairbrother, and D. Tantin. 2009. A general mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and oxidative stress. *Genes Dev* 23:208-222.
8. Kang, J., B. Goodman, Y. Zheng, and D. Tantin. Dynamic Regulation of Oct1 during Mitosis by Phosphorylation and Ubiquitination. *Submitted*.

9. Schild-Poulter, C., L. Pope, W. Giffin, J. C. Kochan, J. K. Ngsee, M. Traykova-Andonova, and R. J. Hache. 2001. The binding of Ku antigen to homeodomain proteins promotes their phosphorylation by DNA-dependent protein kinase. *J Biol Chem* 276:16848-16856.
10. Schild-Poulter, C., A. Shih, D. Tantin, N. C. Yarymowich, S. Soubeyrand, P. A. Sharp, and R. J. Hache. 2007. DNA-PK phosphorylation sites on Oct-1 promote cell survival following DNA damage. *Oncogene* 26:2980-2988.
11. Nie, J., S. Sakamoto, D. Song, Z. Qu, K. Ota, and T. Taniguchi. 1998. Interaction of Oct-1 and automodification domain of poly(ADP-ribose) synthetase. *FEBS Lett* 424:27-32.
12. Fan, W., S. Jin, T. Tong, H. Zhao, F. Fan, M. J. Antinore, B. Rajasekaran, M. Wu, and Q. Zhan. 2002. BRCA1 regulates GADD45 through its interactions with the OCT-1 and CAAT motifs. *J Biol Chem* 277:8061-8067.
13. Wang, R. H., H. Yu, and C. X. Deng. 2004. A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint. *Proc Natl Acad Sci U S A* 101:17108-17113.
14. Imai, S., S. Nishibayashi, K. Takao, M. Tomifuji, T. Fujino, M. Hasegawa, and T. Takano. 1997. Dissociation of Oct-1 from the nuclear peripheral structure induces the cellular aging-associated collagenase gene expression. *Mol Biol Cell* 8:2407-2419.
15. Malhas, A. N., C. F. Lee, and D. J. Vaux. 2009. Lamin B1 controls oxidative stress responses via Oct-1. *J Cell Biol* 184:45-55.
16. Shakya, A., J. Kang, J. Chumley, M. A. Williams, and D. Tantin. 2011. Oct1 is a switchable, bipotential stabilizer of repressed and inducible transcriptional states. *J Biol Chem* 286:450-459.
17. Wang, V. E. H., T. Schmidt, J. Chen, P. A. Sharp, and D. Tantin. 2004. Embryonic Lethality, Decreased Erythropoiesis, and Defective Octamer-Dependent Promoter Activation in Oct-1-Deficient Mice. *Mol Cell Biol* 24:1022-1032.
18. Wang, Z., N. D. Udeshi, M. O'Malley, J. Shabanowitz, D. F. Hunt, and G. W. Hart. 2009. Enrichment and site-mapping of O-Linked N-Acetylglucosamine by a combination of chemical/enzymatic tagging, photochemical cleavage, and electron transfer dissociation (ETD) mass spectrometry. *Mol Cell Proteomics* 19:19.

19. Takahashi, S., S. Saito, N. Ohtani, and T. Sakai. 2001. Involvement of the Oct-1 regulatory element of the gadd45 promoter in the p53-independent response to ultraviolet irradiation. *Cancer Res* 61:1187-1195.
20. Jin, S., F. Fan, W. Fan, H. Zhao, T. Tong, P. Blanck, I. Alomo, B. Rajasekaran, and Q. Zhan. 2001. Transcription factors Oct-1 and NF-YA regulate the p53-independent induction of the GADD45 following DNA damage. *Oncogene* 20:2683-2690.
21. Dalvai, M., K. Schubart, A. Besson, and P. Matthias. 2010. Oct1 is required for mTOR-induced G1 cell cycle arrest via the control of p27(Kip1) expression. *Cell Cycle* 9:3933-3944.
22. Si, F., S. H. Shin, A. Biedermann, and G. M. Ross. 1999. Estimation of PC12 cell numbers with acid phosphatase assay and mitochondrial dehydrogenase assay: dopamine interferes with assay based on tetrazolium. *Exp Brain Res* 124:145-150.
23. O'Neill, J. S., and A. B. Reddy. 2011. Circadian clocks in human red blood cells. *Nature* 469:498-503.
24. Webster, D. M., C. F. Teo, Y. Sun, D. Wloga, S. Gay, K. D. Klonowski, L. Wells, and S. T. Dougan. 2009. O-GlcNAc modifications regulate cell survival and epiboly during zebrafish development. *BMC Dev Biol* 9:28.
25. van den Berg, D. L., T. Snoek, N. P. Mullin, A. Yates, K. Bezstarosti, J. Demmers, I. Chambers, and R. A. Poot. 2010. An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* 6:369-381.
26. Pardo, M., B. Lang, L. Yu, H. Prosser, A. Bradley, M. M. Babu, and J. Choudhary. 2010. An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell Stem Cell* 6:382-395.
27. Issad, T., and M. Kuo. 2008. O-GlcNAc modification of transcription factors, glucose sensing and glucotoxicity. *Trends Endocrinol Metab* 19:380-389.
28. Yang, W. H., J. E. Kim, H. W. Nam, J. W. Ju, H. S. Kim, Y. S. Kim, and J. W. Cho. 2006. Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. *Nat Cell Biol* 8:1074-1083.
29. Park, S. Y., H. S. Kim, N. H. Kim, S. Ji, S. Y. Cha, J. G. Kang, I. Ota, K. Shimada, N. Konishi, H. W. Nam, S. W. Hong, W. H. Yang, J. Roth, J. I. Yook, and J. W. Cho. 2010. Snail1 is stabilized by O-GlcNAc modification in hyperglycaemic condition. *Embo J* 29:3787-3796.
30. Hay, R. T. 2005. SUMO: a history of modification. *Mol Cell* 18:1-12.

CHAPTER 5

**COMBINATORIAL BINDING OF TRANSCRIPTION FACTORS
IN THE PLURIPOTENCY CONTROL REGIONS
OF THE GENOME**

**Luciana Ferraris, Allen P. Stewart, Jinsuk Kang, Alec M. DeSimone, Matthew
Gemberling, Dean Tantin, and William G. Fairbrother**

Reprint from Genome Research, Vol.21, pp. 1-10

Copyright © 2011, with permission from Cold Spring Harbor Laboratory Press

Combinatorial binding of transcription factors in the pluripotency control regions of the genome

Luciana Ferraris,^{1,4} Allan P. Stewart,^{2,4} Jinsuk Kang,^{3,4} Alec M. DeSimone,¹ Matthew Gemberling,¹ Dean Tantin,³ and William G. Fairbrother^{1,2,5}

¹Molecular Biology, Cell Biology, and Biochemistry, Brown University, Providence, Rhode Island 02912, USA; ²Center for Computational Molecular Biology, Brown University, Providence, Rhode Island 02912, USA; ³Department of Pathology, University of Utah School of Medicine, Salt Lake City, Utah 84112, USA

The pluripotency control regions (PluCRs) are defined as genomic regions that are bound by POU5F1, SOX2, and NANOG *in vivo*. We utilized a high-throughput binding assay to record more than 270,000 different DNA/protein binding measurements along incrementally tiled windows of DNA within these PluCRs. This high-resolution binding map is then used to systematically define the context of POU factor binding, and reveals patterns of cooperativity and competition in the pluripotency network. The most prominent pattern is a pervasive binding competition between POU5F1 and the forkhead transcription factors. Like many transcription factors, POU5F1 is co-expressed with a paralog, POU2F1, that shares an apparently identical binding specificity. By analyzing thousands of binding measurements, we discover context effects that discriminate POU2F1 from POU5F1 binding. Proximal NANOG binding promotes POU5F1 binding, whereas nearby SOX2 binding favors POU2F1. We demonstrate by cross-species comparison and by chromatin immunoprecipitation (ChIP) that the contextual sequence determinants learned *in vitro* are sufficient to predict POU2F1 binding *in vivo*.

[Supplemental material is available for this article. The microarray data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo>) under accession no. GSE27535.]

POU5F1 (formerly Oct4), SOX2, and NANOG are generally regarded as the core set of transcription factors necessary for maintaining the pluripotent state in stem cells (Boyer et al. 2005). The loss of POU5F1 or NANOG causes embryonic stem (ES) cells to lose pluripotency and differentiate inappropriately (Hough et al. 2006). Precise POU5F1 levels are also critical to maintaining pluripotency and are at least partially controlled by SOX2, a factor that can heterodimerize to potentially change POU5F1 binding specificity (Niwa et al. 2000; Remenyi et al. 2003). POU5F1 can be regarded as a central pillar in that SOX2 and NANOG interact with POU5F1 but are not known to interact with each other; POU5F1 can be regarded as the hub of this pluripotency network (Zhang et al. 2007).

POU5F1 is part of a larger family of POU factors (formerly known as Oct factors) that share DNA binding specificity and overlapping domains of expression (Phillips and Luisi 2000). The entire family of POU factors was identified on the basis of their ability to bind the octamer sequence ATGCAAAT (Rosales et al. 1987). In addition to POU5F1, the paralog POU2F1 (formerly Oct1) is expressed in ES cells and is capable of heterodimerizing with POU5F1 (Okamoto et al. 1990; Tomilin et al. 2000). This additional complexity raises the question of how distinct biological functions emerge from DNA binding proteins that have similar specificity.

POU binding can be influenced by POU factor-specific interactions with other DNA binding proteins. The human protein reference database reports 36 DNA binding proteins that interact with POU2F1 and 10 that interact with POU5F1, suggesting a large repertoire of complexes that may lend distinct specificities to POU paralogs (Peri et al. 2003). The presence of cofactors can influence

binding and distinguish one POU factor from another. For example, POU2F1 and POU2F2, but not other POU factors, interact with the DNA binding tissue-specific cofactor POU2AF1 (also known as OCA-B and OBF-1) (Gstaiger et al. 1996). POU5F1 is known to dimerize with SOX2 and can be immunoprecipitated with NANOG (Zhang et al. 2007).

Many transcription factors bind DNA via alternate motifs (Badis et al. 2009). Distinct binding specificities between POU factors may lie in their ability to use noncanonical modes of binding octamer-derived motifs on DNA. The octamer motif is a bipartite sequence (i.e., ATGC, AAAT) recognized by two subdomains that together compose the POU domain. As dimers, POU factors are capable of binding a variety of palindromic arrangements of halfsites such as the palindromic octamer recognition element (PORE), ATTGAAATGCAAAT, or the MORE (more PORE), ATGCATATGCAT (Phillips and Luisi 2000; Tantin et al. 2008). The distinct conformations that POU factors form on different classes of binding sites could modulate the access of co-activators (Gstaiger et al. 1996; Kang et al. 2009) and give rise to distinct biological activities. However, there is evidence that the current three classes of binding sites do not describe all the modes through which POU5F1 binds DNA (Tantin et al. 2008).

Here, we present the application of a high-throughput binding assay to map protein/DNA complexes on ~400 kb of transcriptional control regions that regulate pluripotency in ES cells (i.e., regions bound simultaneously by POU5F1, SOX2, and NANOG *in vivo*). We mapped protein occupancy on these DNA targets at a resolution of 10 nucleotides (nt). We discovered that while POU2F1 and POU5F1 share specificity for the octamer, they differ in their ability to bind complex sites. Both POU factors appear to bind their targets through clusters of sites that sometimes stretch for hundreds of nucleotides. Numerous forkhead factor-binding sites co-occur with POU binding sites. We demonstrate that POU5F1 competes with FOXO1 and synergizes with NANOG, while POU2F1 binds synergistically with SOX2. In the process of determining this network of biochemical

⁴These authors contributed equally to this work.

⁵Corresponding author.

E-mail fairbrother@brown.edu.

Article published online before print. Article, supplemental material, and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.115824.110>.

interactions, we also created a rich and detailed map of *cis*-elements on 316 transcription control regions that are important for early development.

Results

POU5F1/SOX2/NANOG modules define 316 pluripotency control regions

We define pluripotency control regions (PluCRs) as genomic regions that bind POU5F1, SOX2 and NANOG *in vivo*. Based on previous chromatin immunoprecipitation (ChIP) data, we identified 316 PluCRs by demanding physically overlapping regions of ChIP enrichment (Fig. 1A; Boyer et al. 2005). ChIP enrichment identified regions that ranged in size from 180–3108 nt (average, 1078 nt). The goal of this study was to map precisely which windows in these regions were bound by the POU factors present in ES cells (POU2F1 and POU5F1) and SOX2 and NANOG. This high-resolution map was then used to discover higher-order patterns and contexts of transcription factor binding in PluCRs. An oligonucleotide pool was constructed by tiling across all 316 PluCRs. (Fig. 1A). Cell extract was added to the oligonucleotide pool, and bound proteins were recovered by immunoprecipitation. The enrichment of each oligonucleotide in the bound fraction relative to the starting pool was quantified by microarray (Tantin et al. 2008; Watkins et al. 2009).

High-throughput binding assay produces a comprehensive map of binding elements on 316 PluCRs

We used this assay to compare DNA/protein complexes in ES cell extract before and after differentiation. The extract was prepared

from the pluripotent J1 ES cells (Li et al. 1992) before and after retinoic acid (RA) differentiation (Bain et al. 1995). Differentiated ES cells adopt a neuronal morphology, stop dividing, and repress POU5F1, SOX2, and NANOG expression (Fig. 1B). We also detected a loss of the repressor of neuronal differentiation REST, which is expressed in ES cells and other non-neural cell types (like HeLa cells) but not in neurons (data not shown). We concluded that our differentiation protocol successfully shifted J1 cells from an ES state to a neuronal-like state.

The oligonucleotide pool was incubated in extract derived from undifferentiated ES cells. POU5F1, POU2F1, SOX2, and NANOG immunoprecipitations were used to separate the oligonucleotide pool into a bound and unbound fraction. The degree of oligonucleotide enrichment in the immunoprecipitate was measured by hybridizing differentially Cy-labeled bound pool and starting pool to a custom oligonucleotide microarray (for detailed methods and comparisons between replicates, see Supplemental Data). A single round of immunoprecipitation achieved a maximal enrichment of 100-fold in our assays. Throughout this article, we refer to the log₁₀ normalized red/green ratios as “enrichment,” and this value is used throughout the article to rank the oligonucleotides by their binding. As these oligonucleotides represent tiled windows of genomic sequence, the enrichment from all microarray probes that overlap each nucleotide position was averaged and rendered on a genomic coordinate system (Fig. 2A). To estimate the selective pressure on these binding elements, the top 1% of oligonucleotides enriched in the immunoprecipitated fractions in these four experiments were examined across 17 vertebrate species. Binding elements of these specific transcription factors were between three- and fourfold more conserved than background levels, suggesting that these elements are functional (Supplemental Table S1).

A well-known target of POU5F1, *DPPA4*, is enriched in the POU5F1 bound fraction

To consider a specific functional element within the more than 316 kb analyzed in this study, we considered a well-characterized target of POU5F1 regulation, the *DPPA4* gene. The *DPPA4* promoter was one of the most enriched signals in the original POU5F1 ChIP experiment in human ES cells, is highly expressed in ES cells, and is regulated by mutationally characterized POU5F1 sites (Boyer et al. 2005; Babaie et al. 2007; Chakravarthy et al. 2008). This example of the *DPPA4* gene illustrates two generalizations about POU5F1 binding that can be made across the 316 PluCRs: (1) Binding of POU5F1 can occur across large regions that can extend hundreds of nucleotides; and (2) the ubiquitously expressed paralog POU2F1 can also bind POU5F1 sites. As POU5F1 possess a nonredundant function in maintaining stemness in ES cells, a competition among POU paralogs could modulate this stemness function within the pluripotency control network.

To verify the results of the array and confirm that POU2F1 was binding *in vivo*, we performed ChIP on the characterized POU5F1 binding site in *DPPA4* (Fig. 2B). While the relative contribution of each POU factor cannot be determined by this assay, both POU2F1 and POU5F1 are enriched at this locus *in vivo* relative to the *beta-ACTIN* control (Fig. 2B). Sequential ChIP (re-ChIP) experiments with a variety of previously identified POU2F1 ligands suggest that POU2F1 and POU5F1 bind together *in vivo*, perhaps by forming heterogeneous complexes similar to those shown *in vitro* (Tomilin et al. 2000). While ChIP is not capable of resolving events at the level of single complexes, the re-ChIP clearly demonstrates that the POU2F1/POU5F1 binding ratio varies with local sequence context.

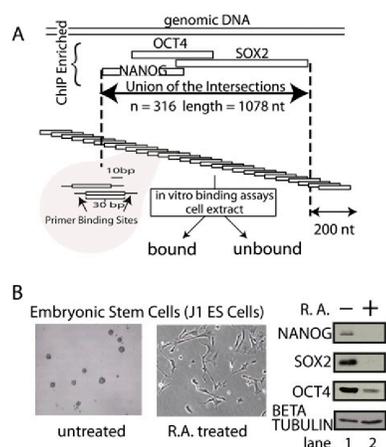


Figure 1. Resynthesizing 316 pluripotency control regions (PluCRs). (A) Regions identified by chromatin immunoprecipitation (ChIP) as binding sites for POU5F1 (OCT4), SOX2, and NANOG were resynthesized as an oligonucleotide pool. The pool tiles through the regions, resulting in a population of 45,793 overlapping oligonucleotides that can be amplified by a universal primer pair. The P32 radiolabeled oligonucleotide pool was incubated in untreated J1 embryonic cell extract and in extract from RA-treated J1 cells and separated into a bound and unbound fraction by two methods. Ligands for unknown cellular DNA binding proteins were identified by their retention in nitrocellulose filter (see below). (B) Morphological changes associated with RA differentiation are recorded in the left panels. Pluripotency markers POU5F1, SOX2, and NANOG were assayed by Western before (lane 1) and after (lane 2) RA treatment.

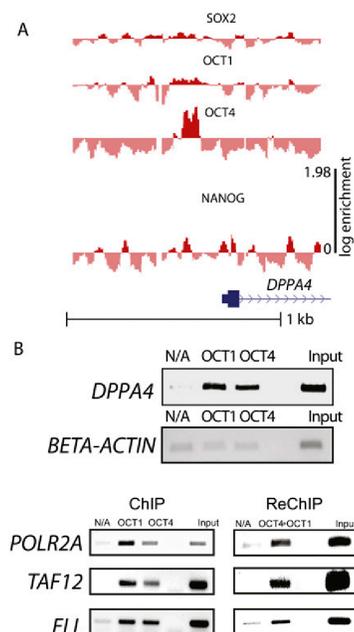


Figure 2. A map of transcriptional complexes on *DPPA4*, a known target of POU5F1 (Oct4). (A) The oligonucleotide pool was subjected to binding assays to identify specific protein DNA complexes by coimmunoprecipitation with SOX2, POU2F1 (Oct1), POU5F1 (Oct4), and NANOG antibodies. Specific oligonucleotide enrichment in the bound fraction was analyzed by two color microarray (bound vs. starting pool). The oligonucleotides were mapped back to genomic coordinates (x-axis), and their average \log_{10} red/green ratio (enrichment) is plotted on the y-axis. UCSC known gene annotation of the *DPPA4* indicates transcription and translation start sites and a portion of the first intron of the *DPPA4* gene. Dotted vertical lines mark the boundaries of complexes identified in the J1 extracted. (B) The enrichment of *DPPA4* and a control region (*BETA-ACTIN*) by chromatin immunoprecipitation with anti-POU2F1 (anti-Oct1), anti-POU5F1 (anti-Oct4), and irrelevant (p53) antibodies measure binding in vivo in J1 ES cells. No antibody and input controls are labeled on the figure. ChIP and sequential ChIP (Re-ChIP) performed on previously identified POU2F1 (Oct1) targets are labeled similarly.

POU5F1 binding is more octamer-based and clustered than POU2F1 binding

To discover signals that influence which paralog joins a POU factor complex, we compared sequence motifs that were discovered de novo in the bound fraction of the POU2F1 immunoprecipitate against those from the POU5F1 immunoprecipitate. As expected, the canonical octamer binding site ATGCAAAT was the dominant motif returned in both the POU2F1 and POU5F1 binding assays (Fig. 3A). The ATGC submotif appeared more prominently in the POU5F1 immunoprecipitate. POU2F1 and POU5F1 also differ in their preference for the previously characterized noncanonical MORE and PORE binding elements. Each of these elements represents palindromic combinations of halfsites (sequences in Fig. 3A) that support POU dimer formation. Although oligonucleotides that contained any octamer-based motif were more enriched in POU5F1 than POU2F1, both POU factors appeared to bind a simple octamer and the PORE element. However, oligonucleotides containing a MORE sequence were differentially enriched for POU5F1

(98th percentile) but not POU2F1 (Fig. 3A; Supplemental Table S5). This result suggests that POU5F1 binds octamer-based oligonucleotides with higher affinity, and MOREs are more “tuned” to POU5F1 than POU2F1. Consistent with this observation, POU dimers bound in a MORE conformation are structurally incompatible with the POU2F1 cofactor POU2AF1. However, when bound in the PORE conformation, these contacts are permitted (Tomilin et al. 2000).

As the flexible linker connecting the two POU subdomains allows some spacing between the ATGC and the AAAT halfsites, the tolerance for intervening nucleotides within the octamer was compared for POU2F1 and POU5F1. Annotating the top 1% of POU2F1 and POU5F1 ligands with various gapped binding models suggests both POU factors tolerate up to 2 nt of spacer (Fig. 3B). However, a greater fraction of high-affinity POU5F1 binding events are found on gapped octamers than perfect octamers.

After determining the primary sequence motifs for POU2F1 and POU5F1, we characterized the transcription factors’ affinity for higher-order motif patterns. In the *DPPA4* gene, POU binding stretches over multiple oligonucleotides encompassing an enriched region of 92 nt (Fig. 2A). An examination of POU binding across all 316 PluCR reveals 42 regions where POU binding seems to span more than a hundred consecutive positions, suggesting a clustering of sites. In order to measure the tendency of sites to cluster, we considered the top 1% of oligonucleotides enriched in the immunoprecipitate to be bona fide POU5F1 “ligands.” We reasoned that an architecture of clustered sites would be evident from examining the oligonucleotides neighboring these ligands. If clustered binding was a prevalent mode of POU binding, the oligonucleotides adjacent to the strong ligands would also be enriched above background in the POU bound fraction. We used this residual enrichment observed in the areas of the peak that lie beyond the dominant “ligand” oligonucleotide to measure a transcription factor’s tendency to bind in cluster (Fig. 3C). For POU5F1 ligands in the PluCRs, we observed a sloping shoulder of enrichment that, on average, accounted for 34% enrichment of the peak and reached background enrichment levels 70 nt away from the ligand. The degree to which binding is clustered is slightly less for POU2F1 and considerably less for NANOG and SOX2 (Fig. 3C; Supplemental Table S2). These findings indicate that POU factor binding extends well beyond the footprint of a traditionally defined transcription factor binding site.

PluCRs contain shared POU2F1/POU5F1 sites, a subpopulation of POU2F1-specific sites and patterns of local associations between POU and other transcription factors

To explore the differences in binding specificity between POU5F1 and POU2F1 at higher resolution, we returned to the output of the high-throughput binding assay. This assay shows that the affinity of all 45,793 oligonucleotides for POU2F1 or POU5F1 can be inferred from their enrichment in the POU2F1 and POU5F1 coimmunoprecipitate. We report a clear positive relationship between the affinity of an oligonucleotide for POU2F1 and its affinity for POU5F1 (Fig. 4A, diagonal trend) and a significant population of oligonucleotides (102 oligos observed, five expected) that are ligands for both factors (Fig. 4A, region 2). This high-resolution comparison of POU factor specificity suggests that the overlapping specificity observed represents shared sites and not two similar classes of sites that are occupied exclusively by POU2F1 or POU5F1. Motif sampling performed on subregions of this distribution indicate that these shared sites contain stronger octamer motifs than found in either comparison alone (Fig. 4A vs. Fig. 3A). While few oligonucleotides recognize only POU5F1 at this threshold, there is a prominent

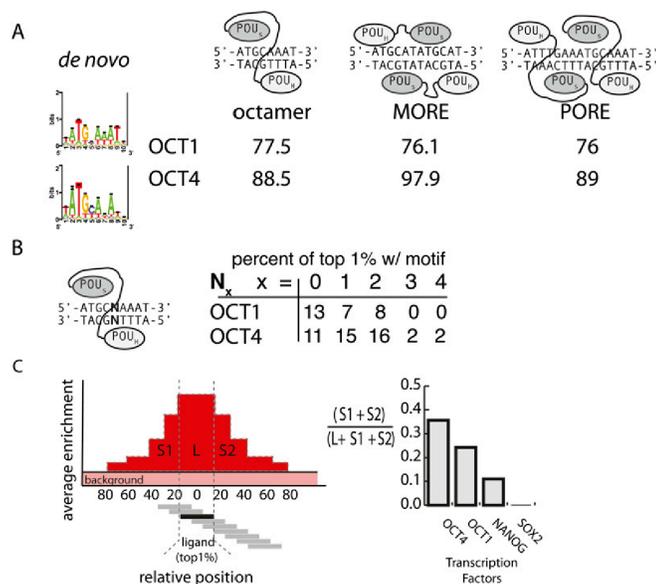


Figure 3. Identifying the specificity and clustering tendency of transcription factors. (A) De novo motif finding was used to identify sequence motifs enriched in the top 1% of oligonucleotides ranked by enrichment in the POU2F1 (OCT1) and POU5F1 (OCT4) immunoprecipitate. The average percentile rank of oligonucleotides matching the three existing POU factor binding models was recorded. Reported conformations of POU factor binding are shown. (B) The distribution of the top 1% of POU2F1 (OCT1) and POU5F1 (OCT4) ligands within gapped octamers is depicted. (C) The tendency of binding events to cluster was analyzed by measuring the average enrichment in the neighborhood of the top 1% of ligands. The percentage of area of the peak that fell outside of the ligand was used to measure a transcription factor's clustering tendency.

subpopulation of oligonucleotides that bound POU2F1 but not POU5F1 (Fig. 4A, region 3 vs. region 1). In contrast to the populations of POU5F1 binding oligonucleotides (region 1 and 2), distinct nonoctamer motifs were found to be enriched in the POU2F1 subpopulation, implying alternate modes of POU2F1 binding or indirect recruitment via factors with alternate DNA binding specificity (Fig. 4B; Supplemental Fig. S3).

To systematically assess the contribution of other transcription factors, we annotated all oligonucleotides with a data set of transcription factor binding motifs and calculated the average POU2F1 and POU5F1 enrichments for the set of oligonucleotides that contain a particular motif. These average enrichments (POU2F1, POU5F1) were plotted as transcription factor binding site "centroids" (Fig. 4B) and summarized the relationship between the TRANSFAC or JASPAR binding models and POU binding. Most centroids clustered at zero enrichment in the immunoprecipitated fraction for both factors, indicating no relationship with POU binding. However, numerous centroids fell along the upper right diagonal, indicating an association with POU factor binding. Similar to the comparison of POU2F1 to POU5F1 oligonucleotide binding enrichment observed in the pool, we observed an overrepresentation of TRANSFAC centroids in the region that represents POU2F1-specific sequences (Fig. 4B, gray areas).

Unsurprisingly, the strongest associations with POU5F1 and POU2F1 binding are found in the numerous highly similar (i.e., overlapping) POU factor binding motifs in the database. However in addition to the many POU factor binding models, certain other transcription factor binding sites were associated with oligonucleotides that bound both POU factors (Fig. 4B, upper right). These factors could

be synergistic cobinders or competitors or coincident with POU factors. Many of these transcription factors (FOXO1, PAX6, VSX2, FOXA2, FOXQ1, and FOXD3) have central roles in development. The strongest non-octamer nonoverlapping motif associated with POU binding is the binding site for FOXO1. FOXO1 is an important regulator of cellular stress associated with aging and has overlapping specificity with POU5F1 and POU2F1 in the pluripotency network. Like POU5F1, FOXO1 is expressed in ES cells, but unlike POU5F1, it persists throughout differentiation (Fig. 5D). While mostly associated with aging and the response to oxidative stress, FOXO1-null mice are characterized by an expanded population of neuronal stem cells, indicating that FOXO1 is capable of restricting multipotency (Brunet et al. 2004; Salih and Brunet 2008; Paik et al. 2009). These genetic data suggest that FOXO1 may function as a POU factor antagonist.

Binding competition between POU5F1 and FOXO1 forms a molecular switch in PluCRs

In order to mechanistically classify the relationship between FOXO1 and POU5F1 as cooperative or competitive, we used recombinant POU5F1 and probes from the *RAB5A* and *ERMN* gene in an *in vitro* binding assay. Both genomic regions contain equally strong matches to the octamer site (ATGCA[A/T]AT, 4C→T in both regions) and two FOXO1 binding sites (TTGTTT) (Fig. 5A). These oligonucleotides map to *in vivo* regions of POU5F1 binding and are located upstream of *ERMN* and *RAB5A* genes. We performed a series of gel mobility shift assays with varying concentrations of recombinant POU5F1 and FOXO1 in order to determine the nature of the binding events. Although each binding site was tuned to bind either FOXO1 (e.g., the *RAB5A* probe) (Fig. 5B, lane 3 vs. lane 2) or POU5F1 (e.g., the *ERMN* probe) (Fig. 5B, lane 2 vs. lane 3), both probes were capable of binding both proteins; however, there was no evidence of an additional species corresponding to both factors binding simultaneously. FOXO1 binding immediately decreased with increasing POU5F1 (Fig. 5B, lanes 4–7). Incubation with antibodies against POU5F1 had a supershifting effect ("←S"; Fig. 5B, lane 8). This antibody also had a blocking effect, thereby reducing the binding of POU5F1 to the probe, which allowed FOXO1 to remain bound (Fig. 5B, lanes 9–12). POU5F1 binds *ERMN* with higher affinity than *RAB5A* (Fig. 5C); however, the quantity of POU5F1-containing species is reduced by increased concentration of FOXO1 (Fig. 5C, decreasing supershift, lanes 4–7). We conclude that POU5F1 and FOXO1 bind competitively to these sequences. As FOXO1 and POU5F1 are expressed together in ES cells, this competition could also be occurring *in vivo* (Fig. 5D). To test whether POU2F1 also competes with FOXO1, we purified recombinant POU2F1 and assayed binding on the *RAB5A* probe. POU2F1 exhibited the same binding preference as POU5F1 on the two substrates tested binding with equal or perhaps slightly lower affinity and exhibiting a greater tendency for dimerization that increased

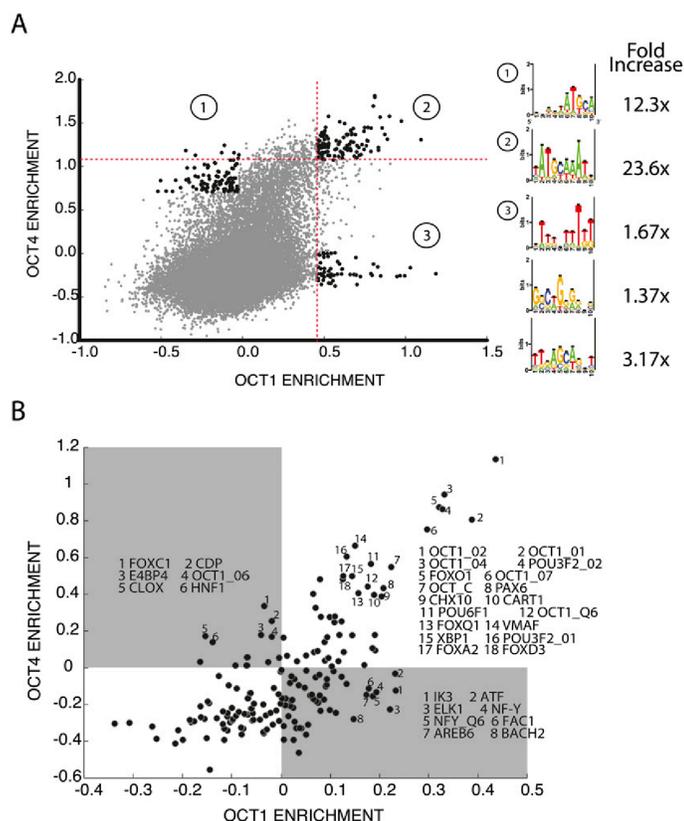


Figure 4. Defining the relative specificity of PluCR oligonucleotides for POU2F1 (Oct1) and POU5F1 (Oct4). (A) The enrichment of all 40,000 oligonucleotides in the bound fraction of POU5F1 (Oct4) and POU2F1 (Oct1) was compared by scatterplot. The top 1% of oligonucleotides defined in each experiment is indicated by vertical (POU2F1) and horizontal (POU5F1) dashed lines. Motifs derived from the doubly enriched and POU specific fraction were identified as in Figure 3. Their enrichment in the bolded data points is printed to the right of the motifs. (B) The association between TRANSFAC/JASPER binding models and POU2F1/POU5F1 binding tendencies. The entire set of vertebrate TRANSFAC/JASPER binding models was used to annotate the oligonucleotide pool. The average POU2F1 (Oct1; x-axis) and POU5F1 (Oct4; y-axis) binding enrichment of the set of oligonucleotides that contain a significant match to a TRANSFAC motif was plotted for each TRANSFAC motif that contained multiple matches.

with POU2F1 concentration (Fig. 5D, lane 2, upper band; data not shown). Unlike POU5F1, there was no detectable decrease in monomeric POU2F1 binding with increasing FOXO1 over several different concentrations (Fig. 5D, lanes 4–14) but rather an increase in multiply bound probe (lane 14).

Proximal binding of NANOG or SOX2 influences POU2F1 versus POU5F1 binding

While there are many more potential cooperative and competitive relationships suggested by this analysis, the most obvious hypothesis for secondary influences on POU5F1 binding is the nearby binding of NANOG or SOX2. To understand the effect of SOX2 and/or NANOG binding on POU specificity, the oligonucleotide pool was ranked by NANOG binding. The nested sets from the top and bottom of this ranked list were then analyzed for POU2F1 and POU5F1 binding. One striking observation is that oligonucleotides

enriched in the bound fraction of both NANOG and SOX2 are also enriched in the bound fraction of POU5F1 and POU2F1, respectively (Fig. 6, diagonal trend toward the upper right quadrant with increasing enrichment in SOX2 and NANOG). While the association between these binding patterns validates previous reports of interactions, the resolution and scale of this binding study afforded the opportunity to search hundreds of biologically relevant binding sites to determine the relative binding locations of POU5F1 to NANOG and SOX2.

Previous studies have demonstrated that SOX2 positioned upstream of the octamer ATGCAAAT can be separated by up to 3 nt and retain protein–protein complexes; however, no structural information has been reported for NANOG (Remenyi et al. 2003; Williams et al. 2004). Comparing the distribution of NANOG sites relative to the octamer motif revealed that high-affinity NANOG sites were significantly more likely (P -value = 0.016) to contain the NANOG recognition element CATT in the forward orientation 3 or 4 nt upstream of the octamer to make the combined motif CATTNN(N)ATGC AAAT (Supplemental Fig. S2; Supplemental Table S3). In comparing NANOG (as a cofactor) versus SOX2, proximal NANOG binding is more strongly associated with POU5F1 binding than POU2F1 binding. Conversely, SOX2 is more highly associated with POU2F1 binding than POU5F1 binding in this assay. This is consistent with reports that while NANOG is known to interact with POU5F1 alone, SOX2 can also interact with POU2F1 (Williams et al. 2004). In addition, the structure of POU5F1/SOX2 contact was determined by homology modeling a complex containing POU2F1's and not POU5F1's POU domain (Remenyi et al. 2003).

It is interesting to note that in the tripartite structure (DNA/POU domain/SOX2), the SOX2 protein contacts DNA and the POU_s domain of the POU factor. The POU factor's POU_s subdomain recognizes the ATGC of the octamer through contacts that are presumably stabilized by the interacting SOX2. As a POU factor in complex with SOX2 has the added stability of protein–protein complexes, we hypothesized a lesser requirement for the DNA contacts in the ATGC motif than would be observed in POU2F1 or POU5F1 binding in the absence of SOX2. To test this idea that strong SOX2 binding rescues suboptimal binding of the POU_s domain to DNA, we compared the tolerance of mismatched octamers in POU2F1 ligands that bind SOX2 (top quartile) versus ligands that do not (bottom three quartiles). Within the set of octamer containing POU2F1 ligands, the subset of regions that also bind SOX2 has significantly weaker octamer sequences than regions that do not bind SOX2 (P -value = 0.0007). This result suggests that SOX2 can rescue a deficiency in octamer binding sites in

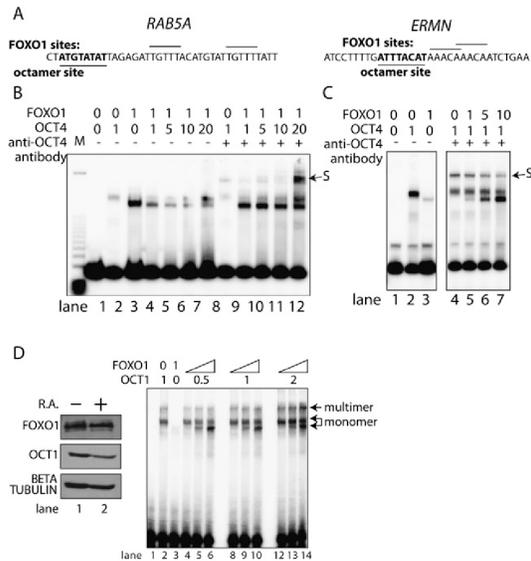


Figure 5. Defining a competition between FOXO1 and POU5F1 (Oct4) for DNA binding at the *RAB5A* and *ERMN* loci. (A) Oligonucleotides originating from the *RAB5A* and *ERMN* upstream control regions that contain annotated octamer and FOXO1 sites were analyzed for FOXO1 or POU5F1 (Oct4) binding. (B) A radiolabeled oligonucleotide was incubated with recombinant POU5F1 (Oct4) or FOXO1 and fractionated by native PAGE. Antibody added prior to loading was used to supershift POU5F1 (Oct4) containing species (marked S). A panel of binding assays containing increasing amounts of POU5F1 (Oct4) was used to assay the ability of POU5F1 (Oct4) to displace FOXO1. (C) An analogous binding assay performed with increasing the amounts of FOXO1 was used to test binding competition at the *ERMN* loci. (D) POU2F1 (Oct1) and FOXO1 protein levels were measured in differentiated and undifferentiated extract. Binding competition between POU2F1 (Oct1) and FOXO1 was tested by gel shift on the *RAB5A* probe. Increasing concentration (1, 5, 10 μ L) of FOXO1 was incubated with a fixed concentration of POU2F1 (Oct1). The ratio of multiply bound to monomeric probe was quantified by phosphorimager.

POU2F1 but not POU5F1 binding. The region most tolerant of mismatches is the ATGC halfsite that is recognized by the POU_s domain. As the SOX2 interaction domain is contained within the POU_s subdomain, this result suggests that the interaction with SOX2 can rescue a loss in canonical protein–DNA contacts with the ATGC halfsite. This phenomenon of protein–protein interactions rescuing suboptimal protein–DNA interactions is only observed in POU2F1/SOX2 complexes and not in POU5F1/SOX2 or POU2F1/NANOG complexes. On the contrary, stronger octamer sites are required for POU2F1 binding in the presence of NANOG binding, suggesting that POU2F1 needs a stronger octamer perhaps to compete with POU5F1 that can be stabilized by NANOG. Consistent with this, octamers in POU5F1 ligands that also bind NANOG are slightly weaker, although this difference did not reach the threshold of statistical significance. This analysis suggests that SOX2 more effectively stabilizes POU2F1 than POU5F1 on DNA and that NANOG is most associated with POU5F1.

Sequence motifs enriched in the POU2F1 subclass of POU ligands are predictive of POU2F1 binding in vivo

The analysis of the POU2F1-specific ligands suggests that there are sequence determinants that allow for the prediction of POU2F1

binding within POU5F1 regulated genes. These sequence determinants are complex and integrate the contribution of a multitude of unknown secondary binding events (perhaps similar to SOX2 and NANOG) with the intrinsic difference in POU binding specificities to types of sites (such as MORE, PORES, gapped octamers). We developed a prediction tool that captures these sequence determinants for the purpose of detecting POU2F1 binding in vivo within POU5F1 ChIP enriched regions, such as the PluCRs (Fig. 7A, prediction scheme). We identified a subset of 5-mers (Supplemental Table S6) enriched in POU2F1 binding sites that were drawn from human POU5F1 ChIP enriched regions. We scored PluCRs for agreement to this word set to return a score that predicted whether an entire PluCR would be predominantly bound by POU2F1 or POU5F1.

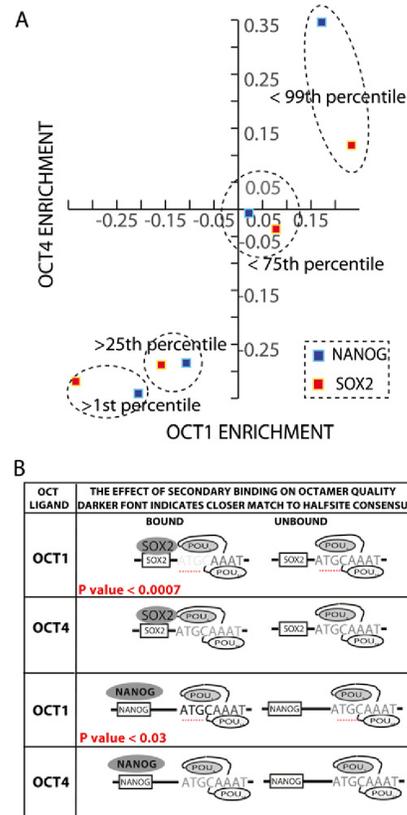


Figure 6. POU2F1(Oct1)/POU5F1(Oct4) binding is influenced by proximal SOX2 and NANOG binding. (A) Subsets of the oligonucleotide pool were created by ranking the pool by SOX2 (red squares) and NANOG (blue squares) enrichment. For each subset, a data point was plotted comparing the average POU2F1 (x-axis) and POU5F1 (y-axis) binding enrichment. (B) The top 1% of POU2F1 (Oct1) and POU5F1 (Oct4) enrichments are denoted as POU2F1 and POU5F1 “ligands” and annotated with an octamer binding model that allowed up to three mismatches. The number (n) and quality of octamer sites were compared for the POU ligands that did or did not bind SOX2. The darker font indicates closer agreement of a halfsite to the octamer consensus in the subset of POU ligands bound by SOX2. The lighter font indicates less agreement. The same analysis was repeated for NANOG. Statistically significant differences in comparisons of half-site strength are listed in red.

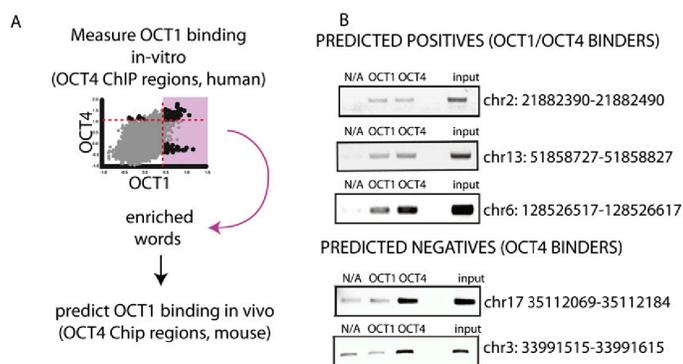


Figure 7. In vivo validation of predicted POU2F1 (Oct1) binding. The prediction algorithm is depicted in A. Significantly enriched words in the top 1% of POU2F1 (Oct1) binding oligonucleotides were added to the predictor with their enrichment score (χ^2 statistic). This score calculated over a window of 100 nt was used to predict POU5F1 ChIP regions in mouse and orthologous regions in human, cow, and dog. (B) High-scoring regions predicting POU2F1 binding were tested by ChIP in mouse. Low-scoring regions predicting background levels of POU2F1 binding were also tested.

Initially, we sought to evaluate the quality of the POU2F1/POU5F1 predictions by particular cross-species comparisons. POU2F1 is found in all vertebrates, whereas POU5F1 is restricted to vertebrates that undergo implantation during embryogenesis. Therefore, we reasoned that if computational prediction correctly distinguishes POU5F1 from POU2F1 binding, then the PluCR targets predicted to only bind POU5F1 should be less conserved than the POU2F1 targets in egg-laying vertebrates. It has recently been demonstrated that sites of POU5F1 binding are rapidly evolving (Kunars et al. 2010). Therefore this comparison was performed on the subset of PluCRs that were retained across three of the four mammalian genomes used in this study. Comparing the conservation of this subset of human PluCR in chicken reveals that orthologs of the predicted POU2F1-bound PluCRs can be found at a higher rate (56%) than orthologs of predicted POU5F1 PluCRs (37%) for cases where the gene is conserved (P -value = 0.18). This observation that POU5F1 predicted targets are under reduced selection in a species lacking POU5F1 suggests that the POU2F1/POU5F1 predictor is accurate; however, the statistical confidence of this claim is modest because of the low numbers of successful double alignments (22 PluCR + gene alignments) and the fact that POU5F1 is only one of many *trans*-acting factors that bind PluCRs.

To directly test the validity of the POU factor predictions, we performed ChIP on predicted POU2F1 and POU5F1 ligands. Positive and negative predictions for POU2F1 binding were selected on the basis of consistently high or low scores in three vertebrate orthologous regions. We found windows that contain words enriched in POU2F1-specific ligands bind POU2F1 in vivo (Fig. 7B, predicted positives), and that the opposite is also true; regions that lack these words are not enriched in POU2F1 ChIP'ed regions and appear to bind POU5F1 exclusively (Fig. 7B, predicted negatives). By incorporating the local context of binding, these findings demonstrate how data derived from in vitro binding results performed in extract can be predictive of transcription factor binding in vivo.

Discussion

We have developed a high-throughput implementation of a classic biochemical binding assay to map 400 Mb of genomic sequence for

the precise sites of complex formation in ES cell extract. Mapping DNA binding at the resolution of 10 nt over thousands of individual sites facilitates the detection of nuanced differences in binding between highly similar paralogs and discovering cooperative and competitive binding events on a genomic or semi-genomic scale. This approach will help solve a common generic problem: There are many families of nucleic acid binding proteins that share similar biological specificity but possess nonredundant functions.

Here we compared binding between POU2F1 and POU5F1, two factors that are expressed in ES cells and were originally discovered (and named) by their affinity for the same octamer sequence, ATGCAAAT. As POU factors are well-studied proteins, many early biochemical characterizations of POU binding performed on synthetic sequences can be revisited with endogenous sequences on a genomic scale. The

ability of POU factors to bind palindromic combinations of halfsites (MORE and PORE motifs) is a point of difference between POU5F1 and POU2F1; POU5F1 has a higher affinity for MORE elements (Fig. 3A). Structural studies demonstrated that the flexible linker connecting the two POU subdomains enables POU binding to tolerate a space of up to 2 nt between halfsites (Tomilin et al. 2000). This analysis finds that the thousands of binding events observed on real sites conform almost perfectly to this limit, but again, there are differences between paralogs. While the single largest class POU2F1 ligand is the perfect octamer, POU5F1 ligands appear to be almost three times more likely to contain a space (i.e., a 1- or 2-nt gap) than to be perfect octamers (Fig. 3B).

Here we report that POU2F1 and POU5F1 both bind the octamer sequence with some important distinctions. POU5F1's predominant mode of interacting with DNA is dominated by full or combinations of half octamer sites. Relative to POU2F1, the more prominent ATGC halfsite is found in POU5F1 ligands, and binding is spread over broader genomic regions. While POU5F1 mostly binds DNA through some combination of octamer halfsites, motif analysis suggests POU2F1 can bind or be recruited through alternate mechanisms (note nonoctamer motifs in Fig. 4A, region 3). The human protein reference database reports 36 interactions between POU2F1 and other DNA binding proteins compared with 10 for POU5F1 (Peri et al. 2003). While there are many potential biases in the interactome data, this study does suggest that POU5F1's primary mode of binding may be more direct than that of POU2F1 in the PluCRs. It is possible that POU2F1 is serving as an adapter and interacting with the transcriptional machinery as heterogeneous complexes with distinct functions. The need to spatially separate these distinct signals may explain the lower tendency of POU2F1 sites to cluster. Indeed, a side population of oligonucleotides seems specific for POU2F1 and not for POU5F1. Scoring the ligands in the oligonucleotide pool for matches to the TRANSFAC motifs suggested patterns of a second transcription factor binding with POU2F1 or POU5F1. Some of these second factors have been implicated in cooperative binding events with a specific POU factor. For example, NF-Y binding motifs are associated with POU2F1 binding across the PluCRs. This trend is supported by reports of NF-Y forming complexes with POU2F1 on the *GADD45* promoter (Fan et al. 2002).

In addition to this interesting class of elements, there are sites that are strongly associated with the binding profile of both POU factors. Not surprisingly, half of the most enriched binding models presented in Figure 4B belong to the multiple TRANSFAC POU position weight matrices. More than a fifth of the remaining transcription factor binding sites belong to forkhead (FOX) transcription factor motifs. We show at multiple loci that POU5F1 and FOXO1 are engaged in a mutually exclusive binding pattern. A similar competition pattern was reported with FOXD3 and POU5F1 on the osteopontin promoter (Guo et al. 2002). However, here we present an analysis that suggests that this relationship exists between the entire family of forkhead and POU factors at numerous regulatory regions within the PluCRs.

Evidence suggests that this antagonistic binding relationship between POU and forkhead transcription factors is mirrored by counteracting roles in maintaining stemness. At an organismal level, studies on null *FOXO* allele mice reveal defects in tissue homeostasis of various adult stem cell populations. The defects are characterized by a small stem cell compartment in the adult and, in some cases, a transient bloom in differentiation at younger stages. FOXO3-deficient animals undergo premature follicle hyperactivation, which depletes oocytes and causes infertility (Castrillon et al. 2003). FOXO-deficient animals develop increased brain size and proliferation of neural stem cells in early life, followed by a greatly reduced stem cell compartment and neurodegeneration in adult life (Paik et al. 2009). This transient increase in a stem cell compartment following *FOXO* depletion argues for an antagonistic function between forkhead factors and factors that maintain a pluripotent state. In the FOXO-deficient background, the hematopoietic stem cell compartment is reduced and marked by elevated levels of oxidative stress (Tothova et al. 2007). This loss of homeostatic control is reversed by antioxidants. FOXO1's role in controlling oxidative stress and in restricting the population of stem cells antagonizes POU factor function, as POU factors have recently been found to increase their binding activity in response to cellular stress (Kang et al. 2009). Taken together, the forkhead and POU factors represent two large transcription factor families that are important in development, have roles in cellular stress response, and bind regulatory regions competitively with overlapping specificities. Furthermore, FOX proteins appear to restrict adult stem cell populations in early development and, in this regard, antagonize the pluripotent function of POU factors.

In addition to competition, the proximal binding of SOX2 and NANOG appears to play a large role in the modulation of POU specificity in the PluCR regions. The association between strong NANOG and POU5F1 binding is compelling and is supported by the finding that NANOG and POU5F1 can be coimmunoprecipitated in vivo (Zhang et al. 2007). The optimal binding location of NANOG lies 3–4 nt apart from the octamer motif, raising the possibility that, at least on these substrates, NANOG may not directly contact POU5F1. SOX2, on the other hand, is associated with both binding events, but strong SOX2 sites are tilted toward POU2F1 but not POU5F1 binding.

The POU/SOX2 interaction surface that forms along the POU₅ domain of POU2F1 ligands appears to have relaxed the stringency of the ATGC halfsite that the POU₅ domain recognizes. Recently, molecular simulation between the POU₅ domain of POU2F1 and SOX2 have kinetically modeled the binding of the POU₅ domain of POU2F1 to its ATGC substrate with and without SOX2 (Lian et al. 2010). Extensive hydrophobic protein/protein interface between POU₅ and HMG (e.g., SOX2) is largely responsible for ensuring a stable ternary complex formation. Using mutations that perturb the protein/protein interface and the POU₅/DNA interface, we conclude that the protein/protein interface effectively prevents the dissocia-

tion of POU₅ from the DNA. This result explains the loss of the requirement of ATGC for POU2F1 when SOX2 is bound proximally. Given this, it is interesting to note that POU5F1 binding motifs observed in this study feature more prominent ATGC motifs in the octamer than POU2F1 (Figs. 3A, 4A).

While the idea that POU2F1 and not POU5F1 binding synergizes more strongly with SOX2 may seem surprising, there is additional evidence that supports this claim. There are examples that these factors interact to coregulate genes outside ES cells. POU2F1 and SOX2 have been shown to function synergistically in the regulation of the *PAX6* and *nestin* genes later in neural development (Donner et al. 2007; Jin et al. 2009). SOX2 is widely expressed and therefore present with POU2F1 in many more tissues than POU5F1, including cell types and cell lines (e.g., HeLa) where POU2F1 is the only POU factor detectable (M Gemberling and L Ferraris, unpubl.). All structural evidence of POU domain/HMG interactions comes from studies on POU2F1. The solution structure of the HMG/POU domain interaction on DNA was performed with POU2F1, not POU5F1 (Williams et al. 2004). Both structures of POU5F1-SOX2 on the *UTF1* and *FGF4* promoter elements were crystallized with POU2F1 and converted to POU5F1 coordinates by homology modeling (Remenyi et al. 2003).

Finally, using this high-throughput binding assay as a means of interrogating the context in which DNA/protein interactions occur, we effectively model DNA/protein interactions in vivo. While chromatin features and epigenetic marks are unlikely to be recapitulated in this assay, we are performing these assays on ChIP'ed regions that are both accessible to and bound by POU, SOX2, and NANOG in vivo. The resolution of ChIPs is on a different scale than the 10-nt resolution achieved in vitro because of the technical limitations of the shearing step during the ChIP protocol. Despite this difficulty, averaging the in vitro-generated prediction over longer windows distinguishes genomic regions that bind POU2F1 from those that bind only POU5F1. This tool can be applied to PluCRs to reevaluate the network and provisionally identify genes that are more likely to be regulated by POU5F1 than POU2F1. High-POU2F1 scoring genes included *HDGF*, *NFIB*, *HES1*, and *TGFA*. The most significant GO term associated with POU2F1-predicted genes was "positive regulation of cell proliferation." The most strongly associated motif was FOXO1, which suggests the possibility that POU2F1 and FOXO1 interact to regulate cell fate (Subramanian et al. 2005). Given the overlapping specificity of POU factors and their distinct binding relationship with FOXO1, it is possible that the exchange of POU factors on these targets during differentiation acts as a switch that triggers FOXO1 binding. Understanding which targets are important to reprogramming a pluripotent state in somatic cells will be of major interest as the therapeutic potential of induced pluripotent stem cells are better evaluated. Likely POU5F1-only genes included *POU5F1* itself, *GAL*, *FOXH1*, *FGFR1*, and *RAB2B*. As POU5F1 is a central player in this triad of factors and POU2F1 and SOX2 are already present in many somatic cells, we intend to focus on predicted POU5F1-specific targets for refining the pluripotency control network in ES cells.

Methods

Cell culture and cell extracts

ES cells were cultured in DMEM + HEPES supplemented with 1 mM glutamine, 1 mM sodium pyruvate, and 1 mM MEM nonessential amino acids (Invitrogen) plus 15% ES cell-qualified heat-inactivated fetal bovine serum (HyClone), 50 mM 2-mercaptoethanol (Sigma),

and leukemia inhibitory factor (LIF/ESGRO; Chemicon). Differentiation of J1 ES cells occurred in the presence of 10^{-7} M (or 100 nM) RA over a 16-d period. Whole-cell extracts were obtained from J1 ES (male) undifferentiated and RA differentiated cells. Cells were pelleted, resuspended, and incubated in extraction buffer (200 mM KCl, 100 mM Tris at pH 8.0, 0.2 mM EDTA, 0.1% Igepal, 10% glycerol, and 1 mM PMSF) for 50 min on ice. Cell debris was pelleted, and extracts were frozen using liquid N_2 and stored at -80°C .

Library design, oligonucleotide synthesis, cloning, and sequencing

A complex pool of 60-mer oligonucleotides was synthesized to contain the union of the intersections of ChIP-ChIP fragments of POU5F1, NANOG, and SOX2, as found in core transcriptional regulatory circuitry in human ES cells (Boyer et al. 2005). The previously described MEGAshift protocol was used to map binding to the PluCRs (Reid et al. 2009). This pool was synthesized as a custom oligonucleotide microarray and liberated from the slide by boiling for 1 h at 99°C . Low-cycle PCR was used to amplify the pool. Electrophoretic analysis was used to ensure log linear amplification as described previously (Reid et al. 2009). Each oligonucleotide was designed as a tiled genomic 30-mer flanked by the common sequences CAGTAGATCTGCCA and ATGGAGTCCAGGTG, which were used as the universal primer binding pair. This oligonucleotide pool was used in coimmunoprecipitation binding studies with antibodies identified below. Nonuniform representations of oligonucleotides are normalized by the two color array approach described in the next section.

Microarrays

We used $8 \times 15\text{k}$ and $2 \times 104\text{k}$ Custom Agilent oligonucleotide microarrays designed complementary to oligonucleotides in each oligonucleotide pool. DNA oligos isolated from four different coimmunoprecipitation reactions (POU5F1, POU2F1, SOX2, NANOG) were amplified with T7 tailed primers and used as templates to produce RNA targets for array analysis. The starting pool was subjected to equal PCR cycles/treatments. RNA targets were produced containing amino-allyl UTP using MEGAshortscript High-Yield Transcription kit (Ambion) after appending a T7 promoter to the oligonucleotides. Amino-Allyl UTPs were then coupled to Cy3 and Cy5 dyes, through a labeling procedure. RNA was pelleted and resuspended in 1 M Na_2CO_3 ; monoreactive dyes were added; and the reaction was allowed to continue for 1 h at room temperature. We used 4 M hydroxylamine to quench the reaction followed by phenol:chloroform and ethanol precipitation to remove remaining free nucleotides. Microarrays were hybridized for 3 h at 50°C using an optimized Agilent gene expression hybridization kit and protocol. Microarrays were scanned at $5 \mu\text{m}$ using a GenePix 4000B scanner and were analyzed using Feature Extraction Software from Agilent. Raw data and UCSC Genome Browser tracks can be downloaded at <http://fairbrother.biomed.brown.edu/data/pluripotent-2010> or through GEO at <http://www.ncbi.nlm.nih.gov/geo/> using accession no. GSE27535. Replicate measurements were averaged, and oligonucleotides were ranked according to enrichment. Oligonucleotides were separated into categories according to enrichment, and a variety of sequence analyses and comparisons were performed on these sets as described in the text. Data analysis is described in detail in the Supplement.

EMSA

Oligonucleotides were prepared for EMSA by end labeling PCR products with $\gamma\text{-}^{32}\text{P}\text{-ATP}$. Samples were prepared in 20 μL

($0.6 \times$ Buffer D, 50 ng/ μL Poly $\text{dI} \cdot \text{dC}$, 1 $\mu\text{g}/\mu\text{L}$ BSA, 1 mM DTT, and 20 ng of probe). Samples were incubated at room temperature for 30 min. Native 4% polyacrylamide gels (29:1 acrylamide:bisacrylamide, 1% glycerol, $0.5 \times$ TBE) were prerun for 1 h at 80 V; samples were loaded and run for 1.75 h at 80 V. A POU5F1 antibody (Santa Cruz Biotechnology) was used for the supershift.

Predictor of POU2F1 binding

Briefly, this tool was trained on the human in vitro binding data described above, and the occurrence of frequency of motifs associated with POU2F1 binding was used to score a test set derived from a mouse POU5F1 ChIP study. To detect words preferentially enriched in the POU2F1 binding subpopulation, we fragmented each 30mer oligo into its constituent k-mer and then compared the frequency of each of these k-mers in the POU2F1 enriched set against the background of the total pool. A power calculation indicates 90% of twofold enrichment events can be detected with 5-mers.

Annotation of POU5F1 ChIP regions for potential POU2F1 binding

We proceeded to use the chi-square scores of 5-mers from our POU2F1 data to determine regions that could be reasonably predicted to interact with POU2F1. The cross-species comparative analysis of POU2F1 word enrichment in known POU5F1-binding regions began by mapping POU5F1 ChIP regions in *Homo sapiens* to *Rattus norvegicus*, *Bos taurus*, and *Pan troglodytes* with the UCSC tool, liftOver. We scanned through all possible 100-mer windows in these four species and annotated regions in which three of the four species had average chi-square scores that were above or below threshold (average chi-square >100). Regions for which at least three of four species had data and in which the average of the window exceeded threshold were used for validation (Fig. 7) and for comparison of conservation in chicken, a species that lacks POU5F1. This final comparison recorded the success rate of identifying the ortholog of the regulatory region (i.e., the PluCR) and its associated gene in chicken.

Western blotting and immunodetection

Following SDS-PAGE separation, proteins were transferred electrophoretically onto a nitrocellulose membrane for 1 h at 30 V following Western blotting. The antibodies used to blot the membrane were anti-POU5F1, anti-SOX2, anti-POU2F1, anti-FOXO1 (Santa Cruz Biotechnology), anti-NANOG (Cosmo Bio), anti-mouse, and goat anti-rabbit HRP-linked (Cell Signaling Technology). The results were visualized by a chemoluminescent reaction using Pierce ECL Substrate Western blot detection (Pierce). For blot development, the membranes were exposed to Kodak Bio Max Light Film.

Acknowledgments

We thank Dr. Lisa Cirillo (Medical College of Wisconsin, Department of Cell Biology, Neurobiology and Anatomy) for the extremely generous gift of FOXO1 proteins and expression constructs, and Richard Freiman and Alex Brodsky for discussions and reagents. The work in the laboratory is partially supported by Award Number R21HG004524 from the National Human Genome Research Institute and Brown University. We also thank the NSF UBM-Group for summer support to A.P.S. (DUE-0734234), a CCMB Scholarship Award (to W.G.F), the Brown University Research Seed Fund Program (to W.G.F), and grants from the American Cancer Society (GMC-115196) and National Cancer Institute (1R21CA141009) to D.T. Finally, we thank the anonymous reviewers for specific suggestions that were incorporated into the manuscript.

References

- Babaie Y, Herwig R, Greber B, Brink TC, Wruck W, Groth D, Lehrach H, Burdon T, Adjaye J. 2007. Analysis of POU5F1-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. *Stem Cells* **25**: 500–510.
- Badis G, Berger MF, Philippakis AA, Talukder S, Gehrke AR, Jaeger SA, Chan ET, Metzler G, Vedenko A, Chen X, et al. 2009. Diversity and complexity in DNA recognition by transcription factors. *Science* **324**: 1720–1723.
- Bain G, Kitchens D, Yao M, Huettner JE, Gottlieb DJ. 1995. Embryonic stem cells express neuronal properties in vitro. *Dev Biol* **168**: 342–357.
- Boyer LA, Lee TI, Cole ME, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, et al. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**: 947–956.
- Brunet A, Sweeney LB, Sturgill JF, Chua KF, Greer PL, Lin Y, Tran H, Ross SE, Mostoslavsky R, Cohen HY, et al. 2004. Stress-dependent regulation of FOXO transcription factors by the SIRT1 deacetylase. *Science* **303**: 2011–2015.
- Castrillon DH, Miao L, Kollipara R, Horner JW, DePinho RA. 2003. Suppression of ovarian follicle activation in mice by the transcription factor Foxo3a. *Science* **301**: 215–218.
- Chakravarthy H, Boer B, Desler M, Mallanna SK, McKeithan TW, Rizzino A. 2008. Identification of DPPA4 and other genes as putative SOX2:POU-3/4 target genes using a combination of in silico analysis and transcription-based assays. *J Cell Physiol* **216**: 651–662.
- Donner AL, Episkopou V, Maas RL. 2007. SOX2 and Pou2f1 interact to control lens and olfactory placode development. *Dev Biol* **303**: 784–799.
- Fan W, Jin S, Tong T, Zhao H, Fan F, Antinore MJ, Rajasekaran B, Wu M, Zhan Q. 2002. BRCA1 regulates GADD45 through its interactions with the POU-1 and CAAT motifs. *J Biol Chem* **277**: 8061–8067.
- Gstaiger M, Georgiev O, van Leeuwen H, van der Vliet P, Schaffner W. 1996. The B cell coactivator Bob1 shows DNA sequence-dependent complex formation with POU-1/POU-2 factors, leading to differential promoter activation. *EMBO J* **15**: 2781–2790.
- Guo Y, Costa R, Ramsey H, Starnes T, Vance G, Robertson K, Kelley M, Reinbold R, Scholer H, Hromas R. 2002. The embryonic stem cell transcription factors POU-4 and FoxD3 interact to regulate endodermal-specific promoter expression. *Proc Natl Acad Sci* **99**: 3663–3667.
- Hough SR, Clements I, Welch PJ, Wiederholt KA. 2006. Differentiation of mouse embryonic stem cells after RNA interference-mediated silencing of POU5F1 and NANOG. *Stem Cells* **24**: 1467–1475.
- Jin Z, Liu L, Bian W, Chen Y, Xu G, Cheng L, Jing N. 2009. Different transcription factors regulate nestin gene expression during P19 cell neural differentiation and central nervous system development. *J Biol Chem* **284**: 8160–8173.
- Kang J, Gemberling M, Nakamura M, Whitby FG, Handa H, Fairbrother WG, Tantin D. 2009. A general mechanism for transcription regulation by POU2F1 and POU5F1 in response to genotoxic and oxidative stress. *Genes Dev* **23**: 208–222.
- Kunarsko G, Chia NY, Jayakani J, Hwang C, Lu X, Chan YS, Ng HH, Bourque G. 2010. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat Genet* **42**: 631–634.
- Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**: 915–926.
- Lian P, Angela Liu L, Shi Y, Bu Y, Wei D. 2010. Tethered-hopping model for protein-DNA binding and unbinding based on SOX2:POU2F1:Hoxb1 ternary complex simulations. *Biophys J* **98**: 1285–1293.
- Niwa H, Miyazaki J, Smith AG. 2000. Quantitative expression of POU-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**: 372–376.
- Okamoto K, Okazawa H, Okuda A, Sakai M, Muramatsu M, Hamada H. 1990. A novel Octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell* **60**: 461–472.
- Paik JH, Ding Z, Narurkar R, Ramkissoon S, Muller F, Kamoun WS, Chae SS, Zheng H, Ying H, Mahoney J, et al. 2009. FoxOs cooperatively regulate diverse pathways governing neural stem cell homeostasis. *Cell Stem Cell* **5**: 540–553.
- Pero S, Navarro JD, Amanchy R, Kristiansen TZ, Jonnalagadda CK, Surendranath V, Niranjan V, Muthusamy B, Gandhi TK, Gronborg M, et al. 2003. Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res* **13**: 2363–2371.
- Phillips K, Luisi B. 2000. The virtuoso of versatility: POU proteins that flex to fit. *J Mol Biol* **302**: 1023–1039.
- Reid DC, Chang BL, Gunderson SI, Alpert L, Thompson WA, Fairbrother WG. 2009. Next-generation SELEX identifies sequence and structural determinants of splicing factor binding in human pre-mRNA sequence. *RNA* **15**: 2385–2397.
- Remenyi A, Lins K, Nissen LJ, Reinbold R, Scholer HR, Wilmanns M. 2003. Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of POU5F1 and SOX2 on two enhancers. *Genes Dev* **17**: 2048–2059.
- Rosales R, Vigneron M, Macchi M, Davidson I, Xiao JH, Chambon P. 1987. In vitro binding of cell-specific and ubiquitous nuclear proteins to the octamer motif of the SV40 enhancer and related motifs present in other promoters and enhancers. *EMBO J* **6**: 3015–3025.
- Salih DA, Brunet A. 2008. FoxO transcription factors in the maintenance of cellular homeostasis during aging. *Curr Opin Cell Biol* **20**: 126–136.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci* **102**: 15545–15550.
- Tantin D, Gemberling M, Callister C, Fairbrother W. 2008. High-throughput biochemical analysis of in vivo location data reveals novel distinct classes of POU5F1(POU5F1)/DNA complexes. *Genome Res* **18**: 631–639.
- Tomilin A, Remenyi A, Lins K, Bak H, Leidel S, Vriend G, Wilmanns M, Scholer HR. 2000. Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. *Cell* **103**: 853–864.
- Tothova Z, Kollipara R, Huntly BJ, Lee BH, Castrillon DH, Cullen DE, McDowell EP, Lazo-Kallanian S, Williams IR, Sears C, et al. 2007. FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress. *Cell* **128**: 325–339.
- Watkins KH, Stewart A, and Fairbrother W. 2009. A rapid high-throughput method for mapping ribonucleoproteins (RNPs) on human pre-mRNA. *J Vis Exp* **34**: 1622. doi: 10.3791/1622.
- Williams DC Jr, Cai M, Clore GM. 2004. Molecular basis for synergistic transcriptional activation by POU2F1 and SOX2 revealed from the solution structure of the 42-kDa POU2F1:SOX2:Hoxb1-DNA ternary transcription factor complex. *J Biol Chem* **279**: 1449–1457.
- Zhang L, Rayner S, Katoku-Kikyo N, Romanova L, Kikyo N. 2007. Successful co-immunoprecipitation of POU5F1 and NANOG using cross-linking. *Biochem Biophys Res Commun* **361**: 611–614.

Received September 24, 2010; accepted in revised form March 8, 2011.

CHAPTER 6

CONCLUSION

Oct1 is a prototypic octamer transcription factor. Although its activity was initially identified as a regulator of housekeeping genes, recently the focus has shifted to its role in regulating stress response genes. Here I established the mechanism by which post-translational modifications regulate Oct1 activities in response to genotoxic and oxidative stress (Figure 6.1, Chapter 2). I unexpectedly found that Oct1 has novel mitotic activity beyond transcriptional regulation through S335-phosphorylation and ubiquitination (Chapter 3). Furthermore, I also showed that O-GlcNAc modifications of Oct1 are involved in Oct1 dependent transcription (Chapter 4). These forms of regulations seem to be similarly applicable to the Oct1 paralogue, Oct4 from my and others' studies (1-3). Besides the similar post-translational modifications, I also showed that Oct1 and Oct4 share multimeric binding motifs in ESCs establishing a form of crosstalk between them (Chapter 5).

Throughout my studies discussed in this dissertation, I also identified novel protein interactions and functions. They can be categorized into Oct1 regulators/effectors. These are updated in Table 6.1.

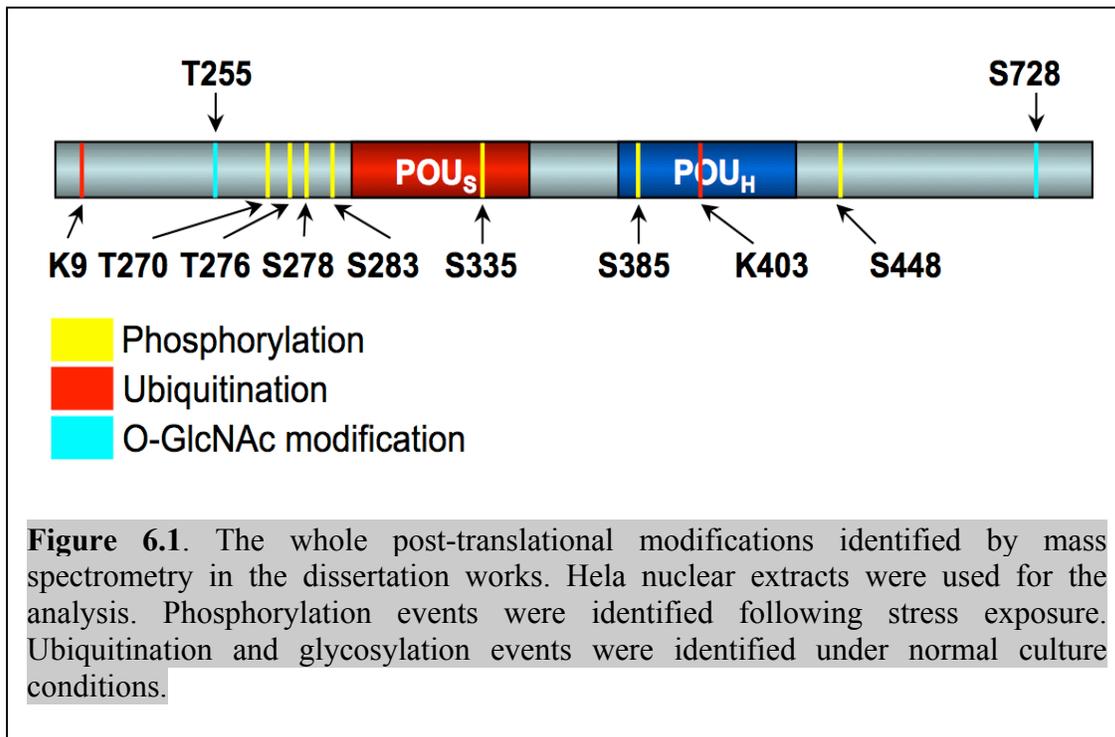


Table 6.1. Previous known and newly identified Oct1 interacting factors. Proteins with bold characters are novel factors and functions. Modified from table1 in chapter 1.

Functional Category	Factor	DNA binding	Function	References
Regulators/Effectors	BRCA-1	independent	Spindle checkpoint activation Activation of Gadd45a expression	<i>Wang et al., 2004.</i>
	DNA-PK	independent	Oct1 stabilization by phosphorylation	<i>Schild-Poulter et al., 2007.</i>
	PARP-1	independent	Increase in Oct1 DNA binding affinity	<i>Ha et al., 2002.</i>
	LaminB	independent	Oct1 sequestration in nuclear periphery Regulates the spindle matrix	<i>Malhas et al., 2009.</i> Kang et al. Chapter 3
	OCA-S	independent	Activation of S phase specific H2B expression	<i>Zheng et al., 2003.</i>
	OCA-B	partially independent	Activation of Immunoglobulin expression Activation of PORE-type Oct1 complex	<i>Tomilin et al., 2000.</i>
	Jmjd1a	independent	Removes H3K4me2 to activate transcription	<i>Shakya et al. 2010.</i>
	SMRT	independent	Repression of Oct-1 transactivation	<i>Kakizawa et al., 2001.</i>
	NuRD	independent	Repression of Oct-1 transactivation	<i>Shakya et al. 2010</i>
	NEK6	independent	Phosphorylation on S335 during mitosis	<i>Kang et al. Chapter 3</i>
	APC1	independent	Ubiquitination of Oct1 during mitosis	<i>Kang et al. Chapter 3</i>
Basal transcription	CAK	independent	Mitosis specific phosphorylation TFIIH recruitment to the preinitiation complex	<i>Inamoto et al., 1997.</i>
	TFIIB	independent	Transcription initiation on TATA-less promoter	<i>Nakshatri et al., 1995.</i>
Cooperative TF interactions	TBP	dependent	Interaction with distal activator	<i>Bertolino et al., 2002.</i>
	SNAPc	dependent	Activation of snRNA transcription	<i>Zhao et al., 2001.</i>
	Sp1	dependent	Activation of snRNA transcription	<i>Lim et al., 2009.</i>
	p65	independent*	Repression of NF-kB transactivation	<i>dela Paz et al., 2007.</i>
	STAT5	dependent	Activation of CyclinD1 expression	<i>Magne et al., 2003.</i>
	Hormone receptors (AR/GR/TR)	dependent or independent	Positive/negative regulation of hormone receptors	<i>Gonzalez et al., 2001,</i> <i>Chandran et al., 1999,</i> <i>Kakizawa et al., 1999.</i>
	Sox2	dependent	Activation of Pax6 expression	<i>Donner et al., 2007.</i>
	C/EBP-b	dependent	Inhibition of immunoglobulin gene transcription	<i>Hatada et al., 2000.</i>

Oct1 as a Stress Sensor

Stress response transcription factors are often highly post-translationally modified following stress exposure. For example, p53 is well studied for dramatic changes in its functions depending on different post-translational modifications (4). Interestingly, several specific phosphorylation and acetylation events restrict p53 DNA binding to the enhancers of apoptotic target genes, although their DNA sequences are not distinct from those of the non-apoptotic targets. Oct1 turns out to adopt similar regulation. Phospho-Ser385 (pS385) converts Oct1 specificity from the octamer to MORE motif (5). Upon stress treatment, Oct1-pS385 is increased resulting in preferred binding to the MORE. However, the octamer binding is constitutive, suggesting that additional Oct1 binding is absent. Genome-wide ChIPseq data showed that inducible MORE-type binding occurs at a large group of specific genes in stressed cells. I also showed that stress-mediated Oct4 binding is also inducible on the MORE and is conserved on the novel MOREs identified by Oct1-ChIPseq indicating that Oct factors can be regulated by the same mechanism.

I also examined *Polr2a* regulation by Oct1. Two consecutive MOREs cooperatively recruit four Oct1 following stress, resulting in stable *Polr2a* expression. Comparing mRNA levels between WT and *Oct1*^{-/-} MEFs, I found that Oct1 acts through an anti-repression mechanism. This mechanism is now elaborated. Stress induces the repressive histone modification, H3K9me2 on the *Polr2a* promoter (6). Oct1 binding to the 2XMORE motif recruits a H3K9me2-specific demethylase, Jmjd1a, and efficiently removes this negative modification.

To better understand Oct1 mediated stress responses, we need to identify substantial sets of inducible Oct1 targets under different conditions and via more

elaborate genome-wide studies. It is also required to define different types of inducible motifs. For example, the MORE is speculated to be just one of a larger set of inducible motifs because the novel TMFORE motif was also shown to be in this category.

Moreover, many binding sites identified by ChIPseq are also inducible motifs that do not match either the octamer or the other known motifs.

Oct1 as a Mitotic Regulator

Previous studies showed that pS385 of Oct1 is responsible for Oct1 displacement from mitotic chromatin (7). However, I found that although pS385 has little negative effect on the octamer binding, it enhances MORE-type binding of Oct1 (5). I instead identified pS335 as the bona-fide negative modification that inhibits both the octamer and MORE-type binding. Using a phospho-specific antibody for pS335, I showed that DNA-released Oct1-pS335 moves to important mitotic structures such as the kinetochore, spindle poles and midbody (Chapter 3). I also showed that pS335 is required for Oct1 displacement from DNA using an Oct1-S335A mutant. In terms of Oct1 mitotic function, Oct1 and Lamin B co-regulate a structure known as the spindle matrix that shapes spindle organization. Further, noncanonical ubiquitination (K11-Ub) is additionally attached on Oct1-pS335 in the spindle poles and midbody. To simplify the findings identified from Chapter 2 and 3, Oct1 has mitotic activity 'OFF' the DNA as well as transcription factor activity 'ON' the DNA.

To further substantiate Oct1 as a mitotic regulator, we need to identify interacting proteins in mitotic structures such as the spindle matrix, spindle poles, and midbody. Also, except for the DNA binding domain, the functions of other regions of Oct1 are not

well defined. A series of truncated Oct1 mutants will provide important clues to identify functional domains in mitotic cells.

Oct1 as a Potential Glucose Sensor

Based on the hypothesis that Oct1 can be regulated by O-GlcNAc modification (details in Chapter 1), I analyzed purified Oct1 by mass spectrometry and mapped two O-GlcNAc modified residues: T255 and S728 (Chapter 4). They are present outside of the DNA binding domain and structure information is not available. Initially, phosphorylation defective mutant Oct1 does not have a slow migrating band with strong glycosylation in Western Blotting suggesting that three different modifications (phosphorylation, glycosylation and sumoylation) may cross talk. This finding should be studied further.

The expression level of T255A/S728A double mutant Oct1 is slightly lower than WT Oct1 suggesting that O-GlcNAc modification partially regulates Oct1 stability. Moreover, both DNA binding affinity and transactivation of reporter genes associated with defined binding sites are much weaker than those of WT Oct1, which is confirmed by Oct1 ChIP findings. Surprisingly, although findings of others have identified an activation function for Oct1 at Gadd45a (8, 9), our findings indicate a repressive function. We already showed that Oct1 can repress and activate the same gene (IL2, 6). In this context, Oct1 as a repressor of Gadd45a recruits a co-repressor complex. NuRD complex recruitment by Oct1 on the Gadd45a promoter is currently being investigated.

Glucose withdrawal significantly reduces the level of O-GlcNAc modification of Oct1 leading to low Oct1 binding on the octamer, consistent with glycosylation defective

mutant Oct1. The data suggest that O-GlcNAc modification of Oct1 is dynamically changed depending on environmental glucose levels. Oct1 regulates TCA cycle genes (10). It is possible that transcription profiles of TCA cycle genes change upon exposure to high and low glucose conditions through O-GlcNAc modification of Oct1. The next efforts will focus on identifying this connection, which will confirm Oct1 as a direct metabolic regulator.

Oct1 and Oct4 Interaction Network

Oct4 is an essential factor in pluripotent ESCs (11). Functionally, it is a part of the core transcriptional regulatory circuitry with Nanog and Sox2 for maintaining self-renewal and pluripotency (12, 13). The three factors frequently co-occupy the regulatory regions (pluripotency control regions) of many target genes. However, Oct1, an Oct4 paralogue, has been ignored in this regulatory network, although substantial amounts of Oct1 are expressed in ESCs (13). Both proteins share similar DNA binding domains and DNA binding specificity (14). Therefore, they are speculated to compete for monomeric binding sites as well as cooperate for multimeric binding sites. Here we performed high-throughput in vitro binding studies and conventional ChIP analysis for Oct1 and Oct4. We found that Oct1 can occupy Oct4 binding sites with similar affinity in the pluripotency control regions. Interestingly, we also identified sequences with preference for Oct4. Oct1 and Oct4 form a hetero-complex on multimeric binding sites indicating that Oct1 is involved in Oct4 dependent transcription. Altogether, I hypothesized that Oct1 is required to maintain developmentally-inducible genes in a poised configuration in ESCs.

Future studies will be dedicated to investigating these phenomena in vivo. For example, our lab has generated Oct1^{-/-} ESCs and can be examined for transcription profiles compared to WT ESCs. Differentially regulated genes will be further studied for Oct1 and Oct4 co-regulated targets.

Future Directions: Oct1 and Somatic Stem Cells

Oct4 expression is highly limited to pluripotent embryonic SCs. This has been controversial because studies reported that multipotent somatic SCs may also express Oct4 (15, 16). However, extensive studies using several different tissue specific Oct4 knockout mice do not identify substantial Oct4 expression and function in somatic SCs (17). The hypothesis consistent with this finding is that somatic SCs use a replacement for Oct4 for at least some functions to maintain “stemness”. This is speculated to be one of Oct factors. Throughout all chapters, we showed that Oct1 and Oct4 share DNA binding specificity and upstream modes of regulation. Oct1 is expressed in all types of tissues, whereas the other Oct factors are expressed in only limited tissues. Therefore, Oct1 is the best candidate to replace Oct4 for the regulation of stemness in somatic SCs, which is consistent with some crucial unpublished findings in our laboratory. Further, asymmetric centrosomal inheritance in mitosis has been reported as a key feature of somatic SCs, which implicates centrosomal Oct1 (Chapter 3) in asymmetric inheritance of centrosomes (18, 19). In the future, our laboratory will attempt to substantiate Oct1 functions in maintaining stemness, which will cover both aspects of transcription regulation and centrosomal function.

References

1. Saxe, J. P., A. Tomilin, H. R. Scholer, K. Plath, and J. Huang. 2009. Post-translational regulation of Oct4 transcriptional activity. *PLoS One* 4:e4467.
2. Xu, H. M., B. Liao, Q. J. Zhang, B. B. Wang, H. Li, X. M. Zhong, H. Z. Sheng, Y. X. Zhao, Y. M. Zhao, and Y. Jin. 2004. Wwp2, an E3 ubiquitin ligase that targets transcription factor Oct-4 for ubiquitination. *J Biol Chem* 279:23495-23503.
3. Webster, D. M., C. F. Teo, Y. Sun, D. Wloga, S. Gay, K. D. Klonowski, L. Wells, and S. T. Dougan. 2009. O-GlcNAc modifications regulate cell survival and epiboly during zebrafish development. *BMC Dev Biol* 9:28.
4. Meek, D., and C. Anderson. 2009. Posttranslational modification of p53: cooperative integrators of function. *Cold Spring Harbor Perspectives in Biology* 1.
5. Kang, J., M. Gemberling, M. Nakamura, F. G. Whitby, H. Handa, W. G. Fairbrother, and D. Tantin. 2009. A general mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and oxidative stress. *Genes Dev* 23:208-222.
6. Shakya, A., J. Kang, J. Chumley, M. A. Williams, and D. Tantin. 2010. Oct1 is a switchable, bipotential stabilizer of repressed and inducible transcriptional states. *J Biol Chem* 286:450-459.
7. Segil, N., S. B. Roberts, and N. Heintz. 1991. Mitotic phosphorylation of the Oct-1 homeodomain and regulation of Oct-1 DNA binding activity. *Science* 254:1814-1816.
8. Takahashi, S., S. Saito, N. Ohtani, and T. Sakai. 2001. Involvement of the Oct-1 regulatory element of the gadd45 promoter in the p53-independent response to ultraviolet irradiation. *Cancer Res* 61:1187-1195.
9. Jin, S., F. Fan, W. Fan, H. Zhao, T. Tong, P. Blanck, I. Alomo, B. Rajasekaran, and Q. Zhan. 2001. Transcription factors Oct-1 and NF-YA regulate the p53-independent induction of the GADD45 following DNA damage. *Oncogene* 20:2683-2690.
10. Shakya, A., R. Cooksey, J. E. Cox, V. Wang, D. A. McClain, and D. Tantin. 2009. Oct1 loss of function induces a coordinate metabolic shift that opposes tumorigenicity. *Nat Cell Biol* 11:320-327.
11. Nichols, J., B. Zevnik, K. Anastassiadis, H. Niwa, D. Klewe-Nebenius, I. Chambers, H. Scholer, and A. Smith. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95:379-391.

12. Boyer, L. A., T. I. Lee, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner, D. K. Gifford, D. A. Melton, R. Jaenisch, and R. A. Young. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122:947-956.
13. Ferraris, L., A. P. Stewart, J. Kang, A. M. Desimone, M. Gemberling, D. Tantin, and W. G. Fairbrother. 2011. Combinatorial binding of transcription factors in the pluripotency control regions of the genome. *Genome Res.*
14. Kang, J., A. Shakya, and D. Tantin. 2009. Stem cells, stress, metabolism and cancer: a drama in two Oct's. *Trends Biochem Sci* 34:491-499.
15. D'Ippolito, G., S. Diabira, G. A. Howard, P. Menei, B. A. Roos, and P. C. Schiller. 2004. Marrow-isolated adult multilineage inducible (MIAMI) cells, a unique population of postnatal young and old human cells with extensive expansion and differentiation potential. *J Cell Sci* 117:2971-2981.
16. Lamoury, F. M., J. Croitoru-Lamoury, and B. J. Brew. 2006. Undifferentiated mouse mesenchymal stem cells spontaneously express neural and stem cell markers Oct-4 and Rex-1. *Cytotherapy* 8:228-242.
17. Lengner, C. J., F. D. Camargo, K. Hochedlinger, G. G. Welstead, S. Zaidi, S. Gokhale, H. R. Scholer, A. Tomilin, and R. Jaenisch. 2007. Oct4 expression is not required for mouse somatic stem cell self-renewal. *Cell Stem Cell* 1:403-415.
18. Wang, X., J. W. Tsai, J. H. Imai, W. N. Lian, R. B. Vallee, and S. H. Shi. 2009. Asymmetric centrosome inheritance maintains neural progenitors in the neocortex. *Nature* 461:947-955.
19. Imai, J. H., X. Wang, and S. H. Shi. 2010. Kaede-centrin1 labeling of mother and daughter centrosomes in mammalian neocortical neural progenitors. *Curr Protoc Stem Cell Biol* Chapter 5:Unit 5A 5.

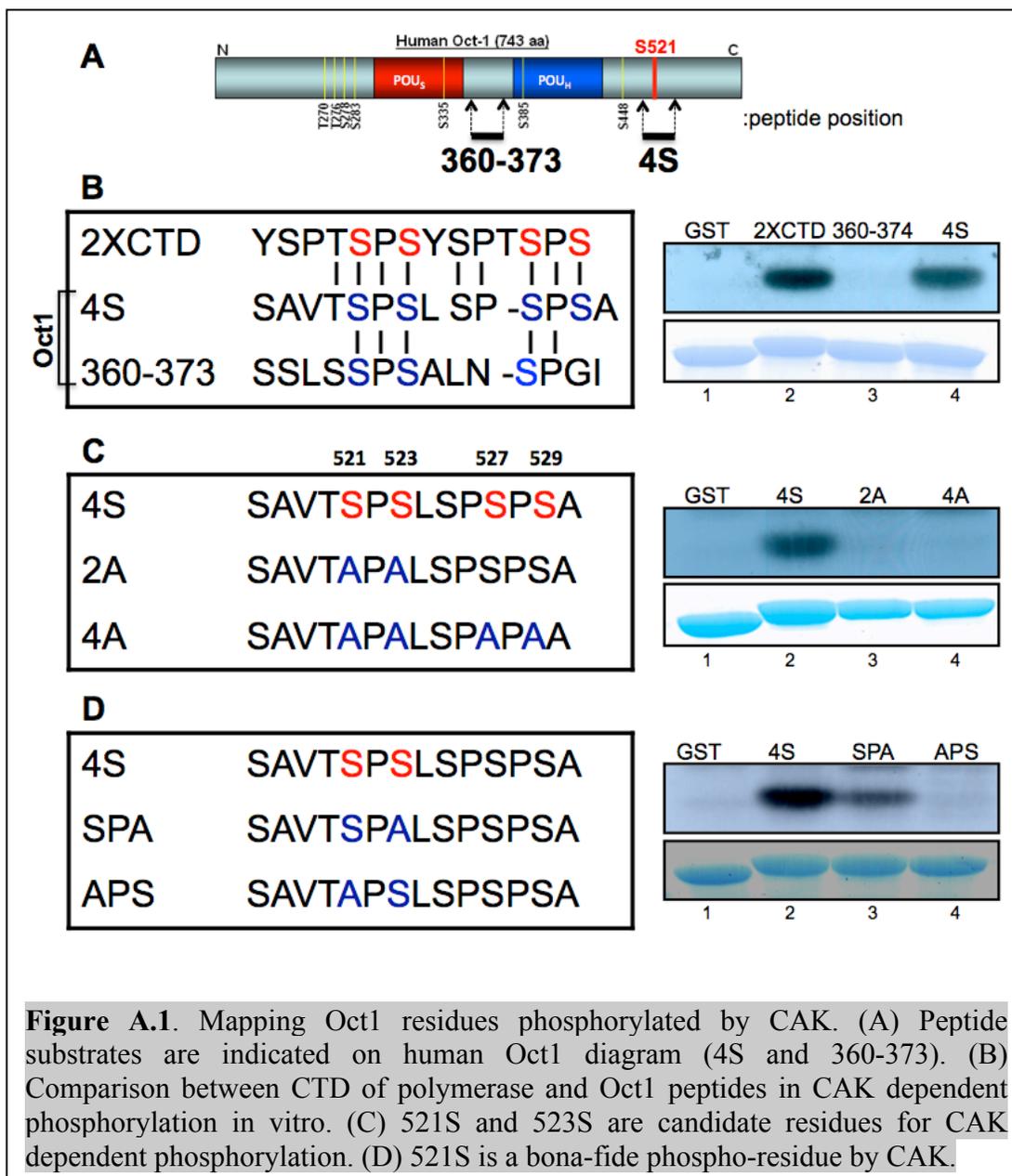
APPENDIX A

MAPPING OF CDK7 DEPENDENT OCT1

PHOSPHO-RESIDUE(S)

The Cyclin-dependent kinase-Activating Kinase (CAK) complex comprises three components: cdk7, cyclin H and MAT1. It interacts with Oct1 through MAT1, and phosphorylates Oct1 (1). However, the specific phosphorylated residue is not known. CAK is a subcomplex of the general transcription complex TFIID (1). Interestingly, TFIID plays important roles in repairing damaged DNA as a nucleotide excision repair complex as well as in initiating transcription as a basal transcription factor complex (2). By analogy, Oct1 can regulate both basal transcription and DNA damage response (3), implicating TFIID as a possible regulator of Oct1 activity.

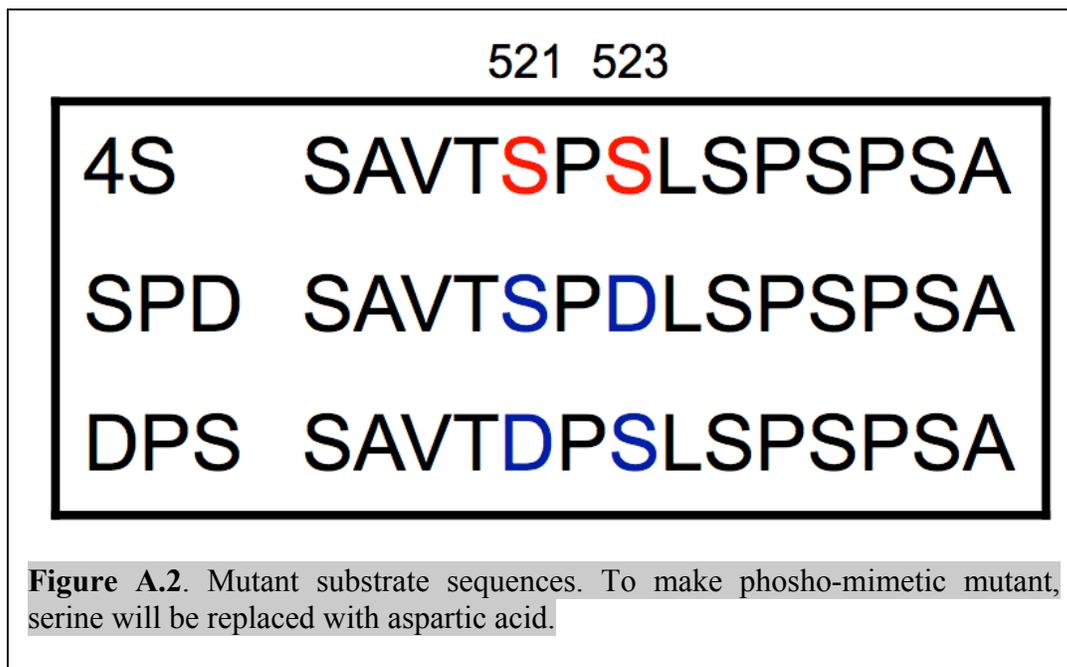
To investigate whether or not TFIID regulates Oct1 activity, here I performed an in vitro kinase assay using CAK, a part of TFIID. RNA polymerase is a well-known substrate for CAK (4). Its CTD region is heavily phosphorylated in repeated motifs (YSPT*S5P*S7: *S is phosphorylated, 5). Using a single CTD motif sequence, I scanned the human Oct1 protein sequence and found two similar sequences covered at positions 360-373 and 517-530 (Figure A.1A). While the former is conserved between limited species, the latter is highly conserved between multiple species including human, mouse and chicken. They resemble two tandem CTD motifs. To test whether CAK



phosphorylates either or both of the two different sequences, I designed and cloned double-stranded oligomers into the pGEX vector system. Using purified GST conjugated substrates, an in vitro CAK assay was conducted (Figure A.1B). As expected, CAK strongly phosphorylates 2XCTD motif but not GST alone (Lanes 1 and 2). In Oct1 derived sequences, 4S (517-530) is the only substrate of CAK (Lanes 3 and 4). To specify the exact residue(s), I introduced alanine mutations into potential serine residues. Both 2A and 4A mutants were not phosphorylated, which narrows down possible residues to S521 and S523 (Figure A.1C, lanes 3 and 4). The fact that SPA (S521/A523) is phosphorylated but not APS (A521/S523) indicates that the S521 is a possible physiological phospho-residue (Figure A.1D, lanes 3 and 4). However, the phosphorylation level of SPA is much lower than that of WT 4S. There are two possibilities. First, S523 could be an interacting site for CAK whose alanine mutation could interfere with the CAK interaction. Second, both S521 and S523 could be phosphorylated just like CTD (YSPTS⁵PS⁷) but S523 phosphorylation could require phospho-S521 as a priming phosphorylation. To test whether both serine residues are phosphorylated by CAK, I designed two different phospho-mimetic mutants (Figure A.2). If S521 is a priming site, DPS (S523) would be phosphorylated and we could conclude that both S521 and S523 are phosphorylated by CAK.

Although I identified phospho-residues in vitro, I should validate them in vivo. A phospho-specific antibody is the best reagent to prove it, but is not available.

Alternatively, I will complement Oct1^{-/-} MEFs with WT, S521A and S521D Oct1 expression constructs using a lentiviral transduction system. Those cell lines will be assessed in Oct1 dependent basal transcription, DNA damage responses and DNA



binding activity. If I find loss or gain of functions from mutant cell lines, I will take this as an evidence favoring the idea that phospho-residue is real and functional. Using more direct approaches, I will confirm in vivo phosphorylation eventl, e.g., phospho-serine antibody blotting after immunoprecipitation of WT and S521A Oct1 and cdk7 knockdown for the same experiments.

References

1. Inamoto, S., N. Segil, Z. Q. Pan, M. Kimura, and R. G. Roeder. 1997. The cyclin-dependent kinase-activating kinase (CAK) assembly factor, MAT1, targets and enhances CAK activity on the POU domains of octamer transcription factors. *J Biol Chem* 272:29852-29858.
2. Drapkin, R., J. T. Reardon, A. Ansari, J. C. Huang, L. Zawel, K. Ahn, A. Sancar, and D. Reinberg. 1994. Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. *Nature* 368:769-772.
3. Kang, J., A. Shakya, and D. Tantin. 2009. Stem cells, stress, metabolism and cancer: a drama in two Octs. *Trends Biochem Sci* 34:491-499.
4. Shiekhattar, R., F. Mermelstein, R. P. Fisher, R. Drapkin, B. Dynlacht, H. C. Wessling, D. O. Morgan, and D. Reinberg. 1995. Cdk-activating kinase complex is a component of human transcription factor TFIIH. *Nature* 374:283-287.
5. Akhtar, M. S., M. Heidemann, J. R. Tietjen, D. W. Zhang, R. D. Chapman, D. Eick, and A. Z. Ansari. 2009. TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Mol Cell* 34:387-393.

APPENDIX B

GENERATION OF iPSCs USING OCT1 AND

OCT4 CHIMERAS

The transduction of four factors (Oct4, c-Myc, Klf4 and Sox2) reprograms differentiated fibroblasts into induced pluripotent stem cells (iPSCs, 1). In many cases, the three factors, except for Oct4, can be replaced by family members (2). This finding is surprising if it is considered that Oct1 and Oct4 share DNA binding specificity and modes of regulation, as shown throughout my dissertation. Therefore, this Oct4 specificity raises interesting questions of what differences between Oct1 and Oct4 result in distinct outcomes: Oct4 but not Oct1 induces pluripotency in the context of the three factors. We hypothesized that regions outside of DNA binding domain determine whether pluripotency can be induced. Oct1 and Oct4 have functionally poorly defined N- and C-terminal domains. Those domains of Oct1 are much longer than those of Oct4. The first possibility is that longer N- or C- domains of Oct1 have a negative effect on induction of pluripotency. To test this, we generated a series of Oct1 deletion mutants: Δ N-Oct1, Δ C-Oct1 and Δ NC-Oct1 (Figure B.1). The second possibility is that N- or C-terminus of Oct4 may be required to interact with crucial effectors to induce pluripotency. Therefore, we replaced N-, C- or N&C- termini of Oct1 with those of Oct4. Lastly, although DNA

No	Name	Sequence
1	Bam.dNoct1.for	GGATCCATGCAGCTCCTGGCCGCTGCA
2	Sal.dNOct1.rev	GTCGACTCACTGTGCCTTGGAGGC
3	Bam.dCOct1.for	GGATCCATGAATAATCCATCAGAAACC
4	Bam.dNCOct1.for	GGATCCATGGCCCAGCTCCTGGCCGCT
5	Sal.dNCOct1.rev	GTCGACTTACGTCGAGTCTGTAGTGCC
6	SnaB1.Oct1DBD.for	TACGTAGAGGAGCCCAGTGACCTT
7	Hpa1.Oct1DBD.rev	GTAACTCTTTTCTCCTTCTGGCG
8	SnaB1.Oct4DBD.for	TACGTAGACATGAAAGCCCTGCAG
9	Hpa1.Oct4DBD.rev	GTAACTCTTTTGCCCTTCTGGCG
10	Oct1N.for	ATGAATAATCCATCAGAAAC
11	Oct1N.rev	CAAGCTGGGAGTGTCAAT
12	Oct1C.for	ATCAACCCGCCAGCAGT
13	Oct1C.rev	TCACTGTGCCTTGGAGGC
14	Oct4N.for	ATGGCTGGACACCTGGCTT
15	Oct4N.rev	CTGGGACTCCTCGGGAGT
16	Oct4C.for	TCAAGTATTGAGTATTCCCA
17	Oct4C.rev	TCAGTTTGAATGCATGGG

	Construct	Primers	Cloning sites
	Δ N-Oct1	1/2	BamHI/Sall
	Δ C-Oct1	3/5	BamHI/Sall
	Δ NC-Oct1	4/5	BamHI/Sall
	Oct1:Oct4N	14/15:6/13	SnaB1
	Oct1:Oct4C	10/7:16/17	Hpa1
	Oct1:Oct4NC	14/15:6/7:16/17	SnaB1/Hpa1
	Oct4:Oct1NC	10/11:8/9:12/13	SnaB1/Hpa1

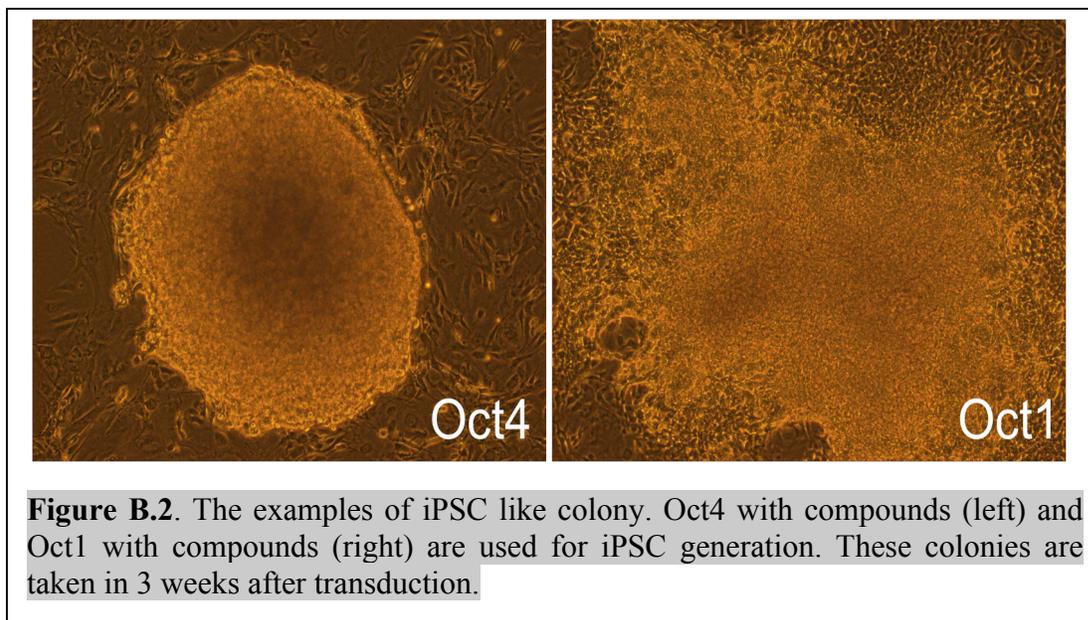
Figure B.1. Cloning scheme. Upper table summarizes all primers used for all cloning procedure. Lower diagram summarizes chimeric constructs and corresponding primer sets and restriction enzymes. Blue and yellow box is derived from MmOct1 and MmOct4. S and H represent POU specific and homeo domain. These constructs will be inserted into pBABE vector using either BamHI/Sall or SnaBI site.

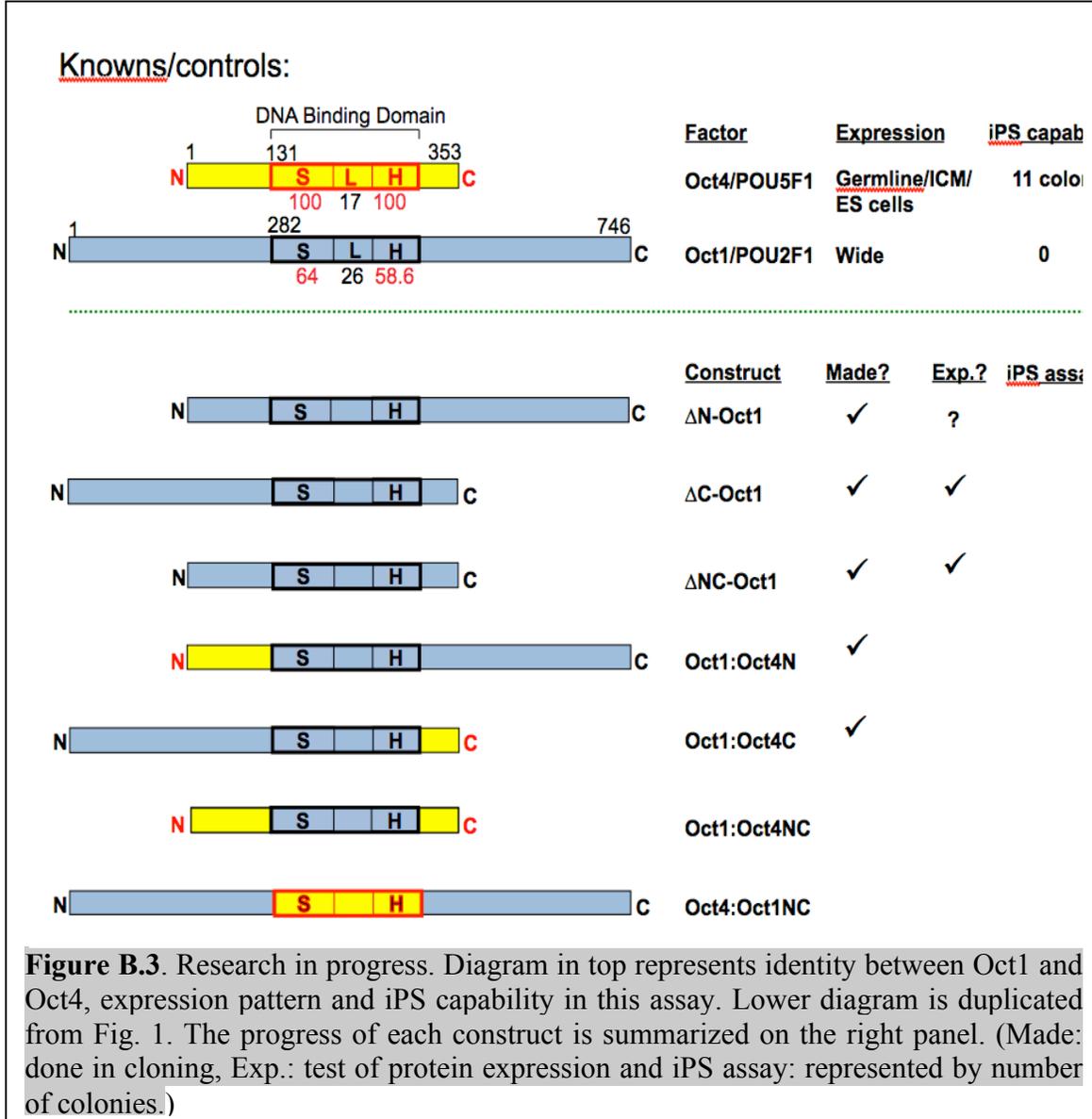
binding domain of Oct1 and Oct4 is highly similar, they may discriminate between some of the crucial pluripotency-related targets. Consistent with this idea, we identified Oct4 preferential binding sequences (3). In this case, N- and C- termini of Oct4 will be replaced by those of Oct1. To maintain the original sequence of each protein, we designed multiple primers for the same domain.

Recent protocol allows us to generate iPSCs using only Oct4 with several small molecule inhibitors (4). The previous protocols using 3 or 4 factors may increase complexity of data interpretation, considering multiple interactions between factors. The novel protocol makes it possible to directly compare Oct4 with Oct1-Oct4 chimeras in induction of pluripotency, which more corresponds to our purposes.

Following the protocol developed by Yuan et al., we transduced the early passage of MEFs using Oct4, Oct1 and ΔC Oct1. Beginning 2 days later, we regularly changed the ES culture media and observed transduced MEFs. In three weeks, iPSC-like colonies were formed and counted. As shown in Figure B.2, colonies from Oct4 expressing MEFs have a clear borderline like ESCs, but those from Oct1 and ΔC -Oct1 fail. Research in progress is summarized in Figure B.3.

Until now, because studies have focused on Oct4 transcription targets, little is known about Oct4 activities at the molecular level. The proposed experiments will provide directions for studying molecular mechanisms by which Oct4 regulates pluripotency. For example, if we identify a critical domain of Oct4 for pluripotency induction, we will make attempts to identify interacting proteins using that domain. The interaction network through this domain may represent an Oct4-specific pluripotency





activity. Or, if we find that chimera of the Oct4 DNA binding domain on the Oct1 N- and C- terminal domains has similar potential in pluripotency compared to WT Oct4, we will conclude that Oct4 preferential targets are critical for pluripotency. This research is currently ongoing.

References

1. Takahashi, K., and S. Yamanaka. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126:663-676.
2. Nakagawa, M., M. Koyanagi, K. Tanabe, K. Takahashi, T. Ichisaka, T. Aoi, K. Okita, Y. Mochiduki, N. Takizawa, and S. Yamanaka. 2008. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* 26:101-106.
3. Ferraris, L., A. P. Stewart, J. Kang, A. M. Desimone, M. Gemberling, D. Tantin, and W. G. Fairbrother. 2011. Combinatorial binding of transcription factors in the pluripotency control regions of the genome. *Genome Res.*
4. Yuan, X., H. Wan, X. Zhao, S. Zhu, Q. Zhou, and S. Ding. 2011. Brief report: combined chemical treatment enables oct4-induced reprogramming from mouse embryonic fibroblasts. *Stem Cells* 29:549-553.

APPENDIX C

STEM CELLS, STRESS, METABOLISM AND CANCER:

A DRAMA IN TWO ACTS

Jinsuk Kang, Arvind Shukya and Dean Tantin

Reprinted from Trends in Biochemical Sciences, Vol. 34, pp.491-499

Copyright © 2009, with permission from Elsevier

Stem cells, stress, metabolism and cancer: a drama in two *Octs*

Jinsuk Kang, Arvind Shakya and Dean Tantin

Department of Pathology, University of Utah School of Medicine, Salt Lake City, Utah 84112, USA

It is a classic story of two related transcription factors. Oct4 is a potent regulator of pluripotency during early mammalian embryonic development, and is notable for its ability to convert adult somatic cells to pluripotency. The widely expressed Oct1 protein shares significant homology with Oct4, binds to the same sequences, regulates common target genes, and shares common modes of upstream regulation, including the ability to respond to cellular stress. Both proteins are also associated with malignancy, yet Oct1 cannot substitute for Oct4 in the generation of pluripotency. The molecular underpinnings of these phenomena are emerging, as are the consequences for adult stem cells and cancer, and thereby hangs a tale.

Setting the stage: Oct transcription factors in development, pluripotency and disease

Oct proteins [1] are odd ducks. Defined by their ability to interact with a DNA sequence known as the 'octamer motif', an important aspect of Oct protein biology relies upon interactions with alternative sequences [2–4]. Although implicated in biological processes as vital and diverse as housekeeping gene expression [5], early embryonic development [6], the development of complex organ systems [7], and the immune/inflammatory response [8], there are nevertheless profound unanswered questions about their biology. One of these proteins (Oct4, also known as POU5F1) currently receives far more attention than all other family members combined owing to its role in embryonic stem (ES) cell pluripotency (ability of ES cells to differentiate into all cells and tissues of the organism), as well as its thus far uniquely essential role in generating induced pluripotent stem (iPS) cells [9]. However, based on sequence composition of full-length human proteins, Oct4 appears to be a typical member of this group (Figure 1A). Despite the high level of attention, little is known about how Oct4 and other Oct proteins regulate target gene transcription. Another member of this family, Oct1, has been reported to regulate multiple housekeeping genes *in vivo*, yet *Oct1*^{-/-} cells are viable but stress-sensitive. Here, we review recent developments relating to Oct1 and Oct4 that resolve some of these paradoxes. Other questions awaiting explanation reveal current boundaries to our understanding and provide avenues for future research.

Oct proteins constitute a sizable portion of the metazoan-specific transcription factor family containing an ~150 amino acid POU (Pit-1, Oct1 and /2, UNC-86) domain

(Figure 2). POU proteins have been divided into six classes (POU1 to POU6) based on DNA binding domain homology [10], Figure 1B). This method also reveals the Oct4 DNA binding domain as an outlier, and partitions these proteins by their ability to bind the classical octamer sequence (5'ATGCAAAT3'): classes 2 (Oct1, Oct2, Skn-1a), 3 (Tst-1, Brn-1, Brn-2, Brn-4) and 5 (Oct4) have high affinity for the octamer motif and are classified as Oct proteins. Classes 1 (Pit-1), 4 (Brn-3.0, Brn-3.1, Brn-3.2) and 6 (Brn-5), though structurally similar, do not recognize this sequence with appreciable specificity. DNA recognition by POU proteins is mediated by two covalently linked DNA binding sub-domains (known as the POU-specific domain or POU_S and POU-homeodomain or POU_H) which independently interact with specific sites in the helix [11]. Notably, Oct proteins can also recognize fundamentally different (non-octamer) DNA sequences [12]. The basis of this recognition lies in the conformational flexibility of the linker domain, which allows the two DNA binding sub-domains to spatially reorganize themselves relative to each other.

Oct1 is widely expressed in adult and embryonic tissues (Figure 2). Oct1-deficient mice die during embryogenesis over a relatively wide developmental window (E12.5–E18.5), and Oct1 is surprisingly dispensable for the viability and growth of cells [13]. However, Oct1-deficient cells display striking phenotypes, including sensitivity to ionizing radiation and resistance to glucose withdrawal [14,15]. Oct1 is a transcription factor, so these phenotypes are almost certainly transcriptional.

Oct4 expression is much more restricted and largely limited to the germline and early embryonic development, where it is expressed in the inner cell mass of the blastocyst (Figure 2). Oct4 is also expressed in ES cells, which are derived from the inner cell mass. Oct4 is necessary for the maintenance of the ES cell phenotype and for progression past the blastocyst stage [6]. In a major breakthrough in regenerative medicine, combinations of transcription factors and in some instances small molecules have been shown to convert somatic cells of multiple different lineages to a pluripotent phenotype apparently identical to ES cells. These reprogrammed cells are termed iPS cells (for review see [16]). To date, all combinations of factors and small molecules used to generate iPS cells include Oct4 as an obligate component (Box 1). These findings suggest a unique feature of the activity or potency of Oct4 that creates a requirement for this transcription factor. For example, SRY (sex determining region Y)-box 1 (Sox1) and Sox3 can replace Sox2, and L-Myc and N-Myc can replace c-Myc in the generation of iPS cells. By

Corresponding author: Tantin, D. (dean.tantin@path.utah.edu).

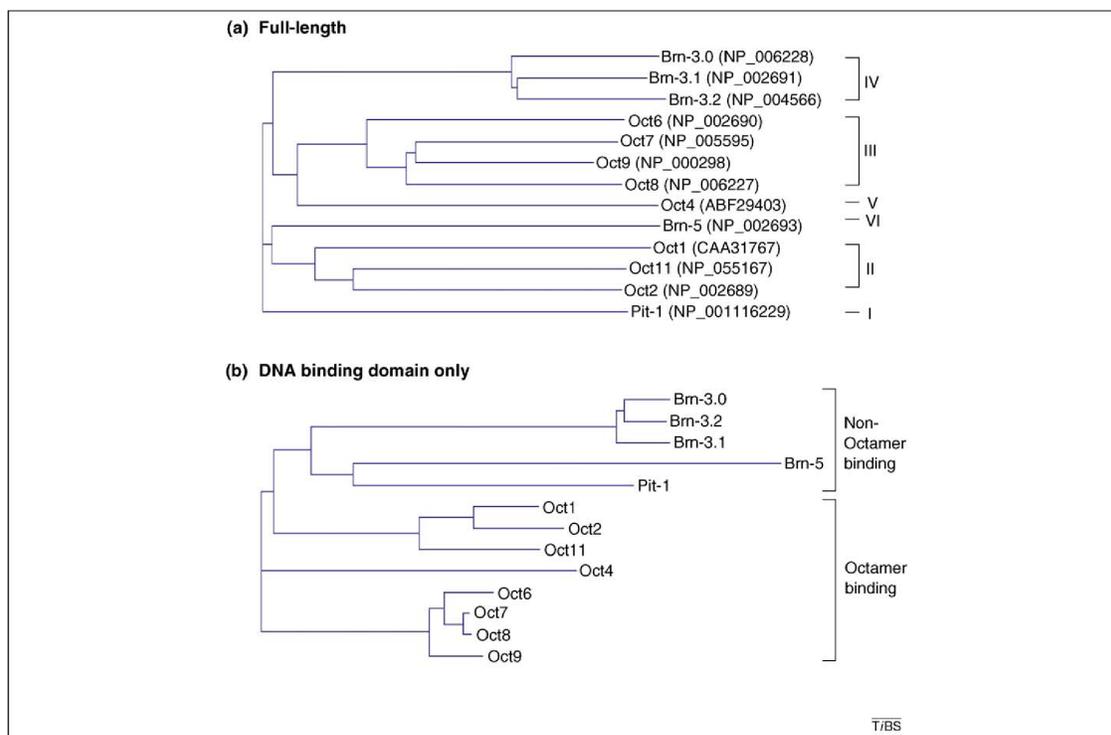


Figure 1. Dendrograms illustrating the degree of relatedness between different human POU transcription factors. (a) Dendrogram generated using full-length proteins. (b) Dendrogram generated using only DNA binding domains. Oct4 becomes more of an outlier when analyzing the DNA binding domains alone. Trees were generated based on sequence distance using guide tree calculation (neighbor joining) algorithm within the Vector NTI software package.

contrast, Oct1 cannot substitute for Oct4 in iPS cell generation [17]. Although it is likely that future reprogramming protocols will use small molecules and particular somatic cell types to minimize the number of exogenous factors, these findings suggest that Oct4 likely will remain a requirement in many instances.

Target genes

Classification of Oct1 and Oct4 targets is an imperfect exercise. Some gene targets can fall into multiple categories, and others might not be classified into any category. Yet there are important lessons to be learned by highlighting a small number of prototypic targets, and demarcating targets shared or exclusive to Oct1 and Oct4. A sizable number of targets for both proteins were identified using traditional molecular biological techniques. More recently, however, far larger numbers have been identified using high-throughput chromatin immunoprecipitation (ChIP)-microarray or ChIPseq methods [2,23–25]. The total number of functional Oct1 and Oct4 targets is not known, but a lower conservative estimate can be put at hundreds of different loci for both proteins. This type of analysis is made particularly difficult because stress signals dynamically regulate the DNA binding specificity of these proteins. Although a careful analysis of common and distinct Oct1 and Oct4 targets under different conditions has not been undertaken, comparing the datasets in two

recent ChIPseq studies [2,26] identifies approximately 150 genes in common. We present a subset of common and unique targets, with an emphasis on genes for which functional data exist (Table 1).

Unlike traditionally identified target genes, target regions identified using an unbiased approach frequently lack identifiable octamer motifs. In these instances, Oct1 and Oct4 appear to recognize DNA through paired, overlapping and other higher-order arrangements of octamer motif half-sites with unusual orientations and spacing [2,23].

Housekeeping genes

Housekeeping genes comprise some of the first identified (and best described) Oct1 and Oct4 targets. For example, genes encoding components of the core transcription machinery such as TBP-associated factor 12 (*Taf12*), elongation factor RNA polymerase II (*Ell*) and polymerase (RNA) II (DNA-directed) polypeptide A (*Polr2a*) contain conserved Oct protein binding sites occupied by Oct1 and Oct4 in somatic and ES cells, respectively (Table 1). In the one case where it has been studied (*Taf12*), occupancy by Oct1 in ES cells is also observed [2]. The wide expression pattern of these targets might seem inconsistent with the known role of Oct4 as a specific regulator of stem-cell pluripotency, and in fact these genes are not frequently highlighted. Nevertheless, Oct4 is clearly associated with

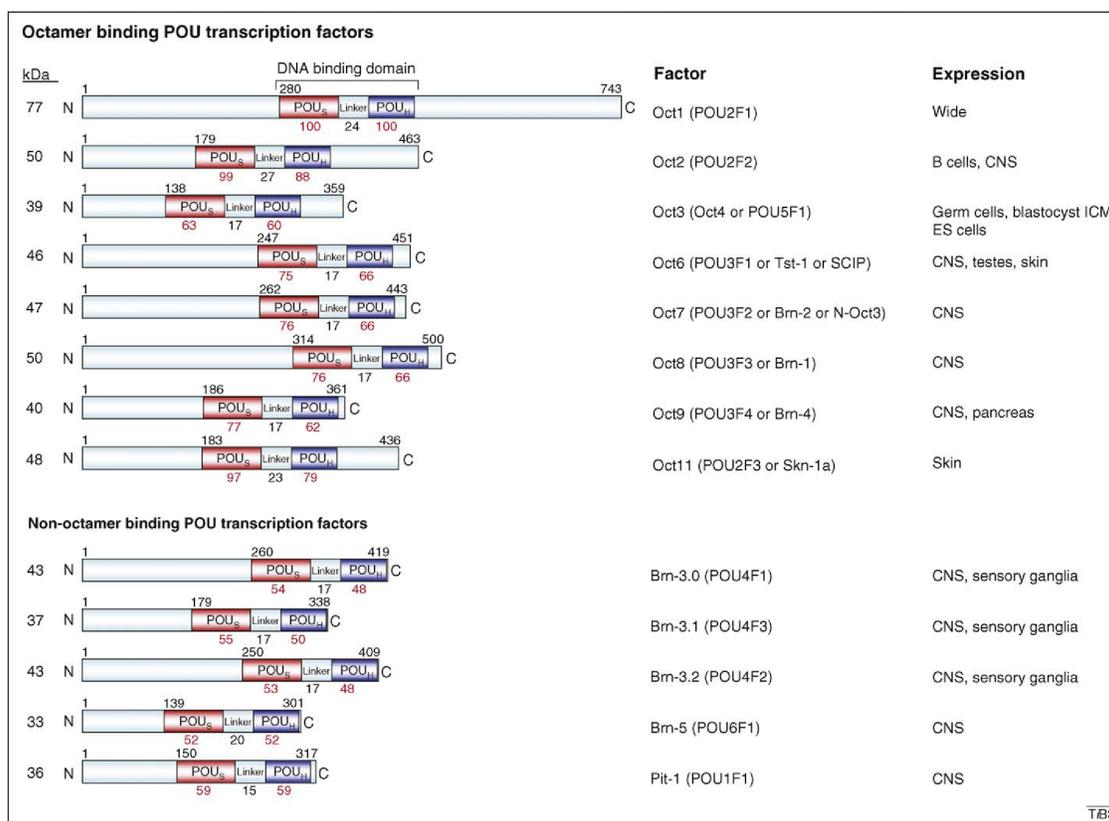


Figure 2. POU domain transcription factor expression patterns and inter-relationships. Schematics of the most common isoforms of the human POU family members are shown. The upper group constitutes the Oct subfamily. The lower group is composed of POU domain proteins that lack specificity for the octamer motif. The predicted protein molecular weight in kDa is shown at the left of each protein. The tissues in which prominent expression is observed are shown on the right. The numbers above each protein show the amino acid position at the start of the DNA binding domain and at the C terminus. Beneath each protein is the percentage identity to Oct1 for the equivalent DNA binding subdomain (POU_S and POU_N), and the number of amino acids in the linker region between the two subdomains. CNS: central nervous system. ICM: inner cell mass.

multiple housekeeping gene loci in ES cells [23–25]. The static nature of these targets has led to the interpretation that Oct1 and Oct4 are static transcription factors, although several lines of evidence have helped to dismantle this idea. In the absence of Oct1 (or any other octamer motif-binding activity) there is little change in the expression of multiple housekeeping targets such as histone H2B, *U2* and *U6 snRNA*, and *Polr2a* [2,13]. Instead of affecting cell viability, loss of Oct1 confers hypersensitivity to genotoxic and oxidative stress agents such as ionizing

radiation, doxorubicin and hydrogen peroxide [15]. Oct1 responds to these stress agents by inducibly binding to a subset of these targets, including *Polr2a* and *Taf12*, through mechanisms that will be summarized below. The traditional prediction from these results is that the expression of these targets should be modulated (up or down) in wild-type cells but not in Oct1-deficient cells after stress exposure. Interestingly, this prediction is incorrect: wild-type cells show stable *Polr2a* expression after stress exposure whereas Oct1-deficient cells display

Box 1. iPS cell induction requires Oct4

iPS cells have been generated from somatic cell types corresponding to all three germ layers (endoderm, ectoderm, mesoderm) using various approaches involving small molecules and transduced factors [10]. These reprogrammed cells, which appear to be indistinguishable from ES cells, can differentiate into all three germ layers and can form teratomas *in vivo*. Additionally, murine iPS cells can generate complete viable and fertile animals just like their ES cell equivalents. Despite the breadth of different protocols used to generate iPS cells, and the differing efficiencies depending on the protocols employed, Oct4 is a common requirement (Table I).

Table I. Known iPS cell inducers

Factors	Compounds	Reference
Oct4 Sox2 Klf c-Myc	None	[9]
Oct4 Sox2 Klf4	None	[17,18]
Oct4 Sox2 Nanog LIN28	None	[19]
Oct4 Sox2	valproic acid	[20]
Oct4 Klf4	BIX-01294, BayK8644	[21]
Oct4 Sox2 Esrrb	None	[22]

Table 1. A subset of Oct1 and Oct4 target genes.

Cellular Process	Target gene	Oct1 target?	Oct4 target?	Refs
Housekeeping	<i>H2B</i>	Yes	Yes	[24,27]
	<i>Polr2a</i>	Yes	Yes	[2,23,28]
	<i>Taf12</i>	Yes	Yes	[2,24]
	<i>E11</i>	Yes	Yes	[2,23]
Developmental Regulators	<i>Pax6</i>	Yes	Yes	[24,29]
	<i>Otx2</i>	Not determined ^a	Yes	[23]
	<i>Pou4f1</i>	Not determined	Yes	[23]
	<i>Hoxb1</i>	Yes	Yes	[30]
Immune/Immunomodulatory/Inflammatory	<i>IgH</i> and <i>Igk</i> variable promoters	Yes	Not detected ^b	[8]
	<i>Il-2</i>	Yes	Not detected	[31,32]
	<i>Il-3</i>	Yes	Yes	[23,33]
	<i>Il-4</i>	Yes	Yes	[23,31]
	<i>Il-5</i>	Yes	Yes	[23,33]
Metabolic	<i>Pcx</i>	Yes	Yes	[14,23]
	<i>Hk1</i> ^c	Yes	Yes	[2,23]
	<i>Ndufa3</i> ^d	Yes	Yes	[2,23]
	<i>Atp5d</i>	Not determined	Yes	[23]
	<i>Atp5g2</i>	Not determined	Yes	[25]
Embryonic stem cell identity	<i>Pou5f1</i>	Not determined	Yes	[34]
	<i>Sox2</i>	Not determined	Yes	[34]
	<i>Nanog</i>	Not determined	Yes	[35,36]
	<i>Fgf4</i> ^e	Not determined	Yes	[37]
	<i>Spp1 (Opn)</i> ^f	Yes	Yes	[38,39]
Other common target examples	<i>Gadd45a</i> ^g	Yes	Yes	[24,40,41]

^aNo published study definitively identifies this gene as a functional Oct1 target.

^bNo high-throughput Oct4 target gene identification study identifies this gene as an *in vivo* target.

^cEncodes Hexokinase 1.

^dEncodes NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3.

^eEncodes Fibroblast growth factor-4.

^fEncodes Osteopontin.

^gEncodes Growth arrest, DNA damage-alpha.

marked inhibition [2]. These results indicate that Oct1 carries out a stress signal-dependent anti-repressive function at this gene. The molecular mechanism underlying this observation is unclear, but might involve the ability of Oct1 to block deposition of negative epigenetic marks. Further experiments are needed to justify or invalidate this model.

Developmental regulators

Another group of target genes encode lineage-specifying master regulators of patterning and organogenesis; these targets are tissue-specific and often encode transcription factors. This group includes paired box 6 (*Pax6*), orthodenticle homeobox 2 (*Otx2*), *Pou4f1*, and genes encoding Hox family proteins (Table 1). These genes are Oct4 targets, but are not expressed in pluripotent tissues or ES cells. Instead, Oct4 helps to maintain these genes in an inactive but 'poised' state such that they can be rapidly induced in response to the appropriate developmental cues, or stably repressed if an alternative developmental lineage is specified. This situation is consistent with an anti-repressive (rather than activating) role for Oct4, extending the analogy with Oct1. The degree to which Oct1 co-occupies these targets in ES cells has not been determined, nor is it known if Oct1 targets a substantial number of these genes in developing embryos and specific adult tissues. However, anecdotal evidence suggests that this might be the case. *Pax6*, for example, is also a functional Oct1 target [29]. Oct1 and Oct4 appear to utilize cooperative transcription factor interactions to regulate many of these targets. For example, Oct1 and Oct4 form cooperative complexes with Sox2 at the *Pax6* and *Hoxb1* enhancers, although it is not

clear if Oct1 and Oct4 recognize the same sites within the *Pax6* enhancer [23,29,30].

Immunological targets

Immune, immunomodulatory and inflammatory genes comprise a large group of Oct1 targets. Regulatory regions in the *IgH* and *Igk* loci, and at the interleukin-2 promoter constitute some of the earliest identified Oct1 binding sites [8]. Although Oct4 expression and the expression of these targets do not significantly overlap, high-throughput data suggests that some of these targets might behave similarly to the lineage-specifying developmental regulators described above because they are bound by Oct4, silent with respect to transcription, and poised for expression in ES cells. Interleukin-3 and interleukin-4, for example, are known Oct1 targets that also appear to be bound by Oct4 in ES cells [23].

Metabolic targets

Several common Oct4 and Oct1 targets are associated with metabolism. Much like housekeeping genes, these targets are not frequently highlighted as examples, but they constitute a major class. Examples include pyruvate carboxylase (*Pcx*) and genes encoding multiple mitochondrial ATP synthases, including *Atp5d* (Table 1). The products of these genes mediate the decision between oxidative and anaerobic glycolytic metabolism which is implicated in carcinogenesis and stem-cell identity. In Oct1-deficient cells, carbon utilization is less fermentative and more oxidative in nature [14]. These metabolic changes oppose those frequently encountered in tumor cells, suggesting that loss of Oct1 might have cancer-protective effects.

Consistent with this prediction, loss of Oct1 reduces oncogenic transformation in culture and cancer incidence in mouse models. Transformation by serial passage is unaffected [14]. Although Oct4 is not known to directly govern a similar metabolic switch, the high glycolytic rate in ES cells points to a parallel function [42]. A study by Plath and colleagues stratified genes induced by an iPS cell induction cocktail (Oct4, Sox2, Klf4 and c-Myc, **Box 1**) into those induced rapidly and relatively efficiently, and those induced slowly and inefficiently [25]. Metabolic targets tend to fall into the former group, suggesting that they might not be rate-limiting targets in the generation of pluripotency.

Genes associated with pluripotency

Oct4 regulates the circuitry governing ES cell pluripotency, in particular the genes encoding Oct4, Sox2 and Nanog (reviewed in [43]). These transcription factors mutually enforce each other's expression in a self-sustaining network that helps to maintain a pluripotent state. This core circuitry has been refined and extended in a recent study [44]. Unlike metabolic genes, these targets are induced slowly and inefficiently by iPS cocktails [25]. These genes could represent largely unique Oct4 targets because no evidence exists for their occupancy by Oct1. Therefore, Oct1 may help mediate the early, relatively efficient stages of iPS generation, but not the late and inefficient ones. The molecular explanations for these differences in target profile are unknown, but might involve unique modifications or interacting partners.

Regulation of Oct factors

Regulation of Oct4 expression has been a topic of continued investigation (Figure 3a). In addition to Oct4 positively regulating its own synthesis, evidence implicates hypoxia inducible factor-2 (HIF-2), estrogen-related receptor beta (Esrrb) and Sal-like 4 (Sall4) as potential positive regulators [45–47]. The Pol II-associated factor 1 complex (Paf1C) is also thought to positively regulate *Pou5f1* expression through its ability to maintain histone H3K4 trimethylation in the promoter and coding region [48]. Factors implicated in *Pou5f1* repression during ES cell differentiation include Cdk2 associated protein 1 (Cdk2ap1) and germ cell nuclear factor (GCNF) [49,50]. The orphan nuclear receptors steroidogenic factor-1 and Tr2 also appear to negatively modulate *Pou5f1* expression [51,52]. Epigenetic regulators, most notably G9a, also repress Oct4 expression during ES differentiation [53,54]. After generation and processing of a *Pou5f1* transcript, the mature mRNA can be repressed by binding of the microRNA miR470 to target sites within the coding region [55]. The large number of currently delineated factors involved in controlling Oct4 expression suggests that its levels must be tightly regulated. Indeed, finely tuned Oct4 levels are crucial as elevated or reduced expression leads to loss of pluripotency [56]. Therefore, induction and repression of Oct4 during early mammalian embryogenesis and in the germline are likely to be heavily regulated, and additional layers of regulation might emerge.

Relatively less is known about how *Oct1* (*Pou2f1*) expression is regulated, although some evidence points

to a role for Oct1 in regulating its own synthesis [57]. Although Oct1 expression has been thought to be uniform, work in the developing eye shows that at a fine level its expression is variegated [29]. Moreover, preliminary studies from our research team indicate that Oct1 expression in epithelial tissues is non-uniform. Therefore, key insights remain to be uncovered about the regulation of expression of Oct1 and Oct4.

Until recently, comparatively little attention has been given to regulation of these proteins at the post-translational level. It is now evident that the activities of Oct1 and Oct4 are regulated at multiple levels. For example, Oct4 is sumoylated and ubiquitylated (Figure 3a), stabilizing and destabilizing the protein respectively [58,59]; however, neither Oct1 ubiquitylation nor sumoylation have been studied. Particularly in the case of Oct1, modification by phosphorylation has been more fully characterized.

Oct1 and Oct4 are modified by phosphorylation at multiple positions (Figure 3b). The kinases responsible for these phosphorylation events have not been fully described. DNA-dependent protein kinase (DNA-PK), which is activated after DNA damage, is a known kinase *in vitro* [60], as is protein kinase A, which phosphorylates mouse Oct4 S229 [4]. DNA-PK phosphorylates Oct1 at physiologically important serines and threonines in the N-terminus [60]. Oct1 is also phosphorylated at multiple sites around the DNA binding domain in response to stimuli, including oxidative stress [2]. Mounting evidence indicates that upstream signaling events can modulate the DNA binding selectivity of Oct1 and Oct4 through DNA binding domain phosphorylation events that favor or disfavor particular orientations of the DNA binding subdomains relative to one another. Two human Oct1 phosphoserines (S335 and S385) are known to alter its sequence selectivity, in particular for non-consensus 'complex sites'. Many of the phosphorylation target residues are conserved in Oct4, with some of these sites undergoing phosphorylation (Figure 3b, [4]). An amino acid substitution at one of these sites modulates Oct4 DNA binding selectivity similarly to an analogous Oct1 mutant [2]. The 'complex' recognition sequences described above generate Oct1–DNA and Oct4–DNA structures distinct from classical octamer sequences. In these structures, the two DNA binding sub-domains are rearranged relative to one another (Figure 3c). These different structures are favored or disfavored by different phosphorylation events, allowing the sequence specificity of these factors to be regulated through a phosphorylation 'code' that might control target gene expression patterns.

Oct factor interaction partners

Oct1 and Oct4 interact with many proteins, more than can be effectively summarized here. Instead we will focus on a subset of prototypic interaction partners (Table 2). Oct1 interacts with several factors implicated in the response to genotoxic stress, including breast cancer 1, early onset (BRCA1) and poly (ADP-ribose) polymerase 1 (PARP-1) [61–63]. Oct1 also interacts with lamin B, and associates with the nuclear periphery in a manner that could be controlled by oxidative stress [65]. A recent study of lamin-associated DNA regions identified a strong enrichment in Oct1 binding sites, thus lending support to the idea

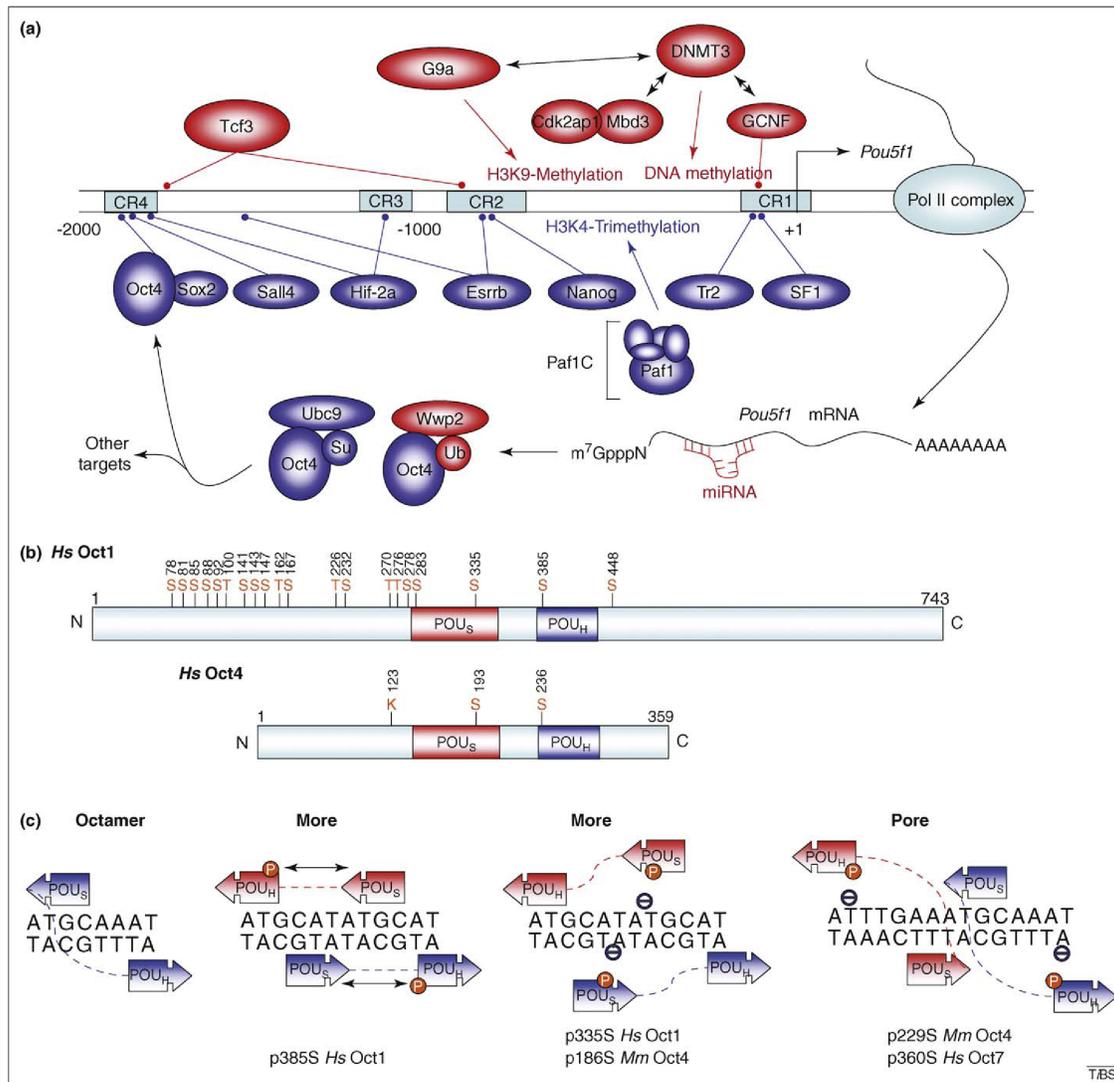


Figure 3. Regulation of Oct1 and Oct4. Various mechanisms are in place to regulate the activity of Oct1 and Oct4. (a) Oct4 transcription, translation, and feedback positively regulate its own synthesis. Points of regulation by different activities are shown by colored lines. Factors in blue (such as Oct4 itself) carry out positive regulatory activities, including H3K4 methylation; those in red are thought to act in a negative capacity (e.g. DNA and H3K9 methylation). Black arrows show known protein–protein interactions. CR1–4 are *Pou5f1* promoter and enhancer regions with a high degree of sequence conservation. (b) Known Oct1 and Oct4 modifications. Amino acid positions of the human proteins are shown. Serines (S) and threonines (T) undergo site-specific phosphorylation events; lysine (K) can be sumoylated or ubiquitylated. (c) Specific examples of how Oct1 and Oct4 protein–DNA complexes can be favored and disfavored by specific phosphorylation events (shown in orange). Dotted lines show the linker domain connecting the two sub-domains. Putative electrostatic interactions mediated by phosphorylation are shown by arrows. Putative repulsive events that disfavor binding are shown with a “e”. The depicted non-canonical sequences are termed PORE (Palindromic Octamer Related Element) and a MORE (More PORE). *Mm*: *Mus musculus*; *Hs*: *Homo sapiens*.

that Oct1 can tether DNA to the nuclear periphery [72]. Two Oct1 co-activators have been described. One is a widely expressed multi-subunit complex known as OCA-S [27] that contains two glycolytic enzymes, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and lactate dehydrogenase (LDH), yet resides in the nucleus. The complex also contains the DNA repair protein uracil-DNA glycosylase (UNG; also called UDG), two chaperones

(Hsp70 and STIP1; stress-induced-phosphoprotein 1) and two proteins of poorly-defined function (nm23-H1 and nm23-H2). The other co-activator, OCA-B (also known as OBF-1 or Bob1), is a single polypeptide. OCA-B is expressed predominantly in the B lymphocyte lineage, and interacts with Oct1 or the related protein Oct2, but not Oct4, Oct6, or Pit-1 [73]. Interestingly, certain complex site DNA binding configurations allow OCA-B binding

Table 2. A partial list of Oct1 and Oct4 interaction partners.

Category	Protein	Function	Refs
Oct1 interacting proteins			
Regulators/effectors	BRCA1	Checkpoint control; Activation of <i>Gadd45a</i> expression	[61,62]
	PARP-1	Alters DNA binding affinity	[63]
	DNA-PK (Ku70)	Oct1 phosphorylation	[64]
	LaminB	Sequestration to the nuclear periphery	[65]
	OCA-S	Regulation of H2B expression	[27]
	OCA-B (POU2AF1)	Activation of B cell-specific genes	[8]
Basal transcription	TBP	Transcription initiation	[66]
	TFIIB	Transcription initiation	[67]
	TFIIH (MAT1)	Transcription initiation/RNA Pol II phosphorylation	[68]
Oct4 interacting proteins			
Regulators/effectors	EWS	Transactivation	[69]
	PIASy	Sequestration in nuclear periphery and inhibition of activity	[70]
	Ubc9	Sumoylation and stabilization	[71]
	Wwp2	Ubiquitination and degradation	[58]

whereas others do not [28]. Oct1 and Oct4 possess similar DNA binding specificity, so differential recruitment of these and possibly other undiscovered co-activators might provide a basis for some of their biological differences.

Another group of Oct1-interacting proteins are involved in basal transcription. Oct1 can form cooperative complexes with TATA box binding protein (TBP) on DNA containing TATA boxes and octamer motifs [66]. Moreover, the general transcription factor TFIIB as well as ménage à trois 1 (MNAT1), a component of the Cdk activating kinase (CAK) TFIIH sub-complex, also interact with Oct1 [67,68]. These interactions might partially mediate transcriptional activity, particularly in situations in which the Oct1 binding site is close to the core promoter.

Known Oct4 interaction partners are fewer in number and include Ewing's sarcoma protein (EWS), which might function as a transcriptional coactivator [69]. Another interacting partner, the protein inhibitor of activated STAT (PIASy), was isolated in a yeast two-hybrid screen and appears to sequester Oct4 to the nuclear periphery [70]. Oct4 also cooperates with several sequence-specific transcriptional regulatory proteins, including Zfp143, to regulate target gene expression [74]. Additionally, Oct1 interacts cooperatively with many sequence-specific DNA binding proteins at compound DNA binding sites, including Sox2 and CCAAT-enhancer binding protein beta (C/EBP β) [29,75].

Roles for Oct proteins in somatic stem-cell function and cancer

The high degree of homology (Figure 2), similar binding specificity, common targets (Table 1) and common modes of upstream regulation (Figure 3b,c) suggest that these two proteins might have parallel functions, or that one protein carries out a subset of the other's functions in some instances. The known role of Oct4 as a master regulator of stem-cell identity therefore suggests that Oct1 or other Oct proteins might carry out similar functions in somatic cells, and/or that Oct1 or other Oct proteins assume Oct4-like functions in malignancy. However, the properties of pluripotent and somatic stem cells are not identical. One possibility is that in those aspects where the biology of pluripotent and somatic stem cells diverge, other regulators could have prominent roles, whereas Oct1 and Oct4

carry out key regulatory functions where the biology is similar.

Several studies indirectly suggest that Oct proteins other than Oct4 function as regulators of 'stemness'. A pluripotency gene expression signature has been identified in aggressive human breast carcinomas without concomitant Oct4 expression. It was suggested that Oct4 paralogs assume Oct4-like functions in these tissues [76]. Similarly, a recent study identified an embryonic stem-cell gene expression signature in myeloid leukemia stem cells associated with myeloid/lymphoid or mixed-lineage leukemia (*MLL*), again without Oct4 expression [77]. Oct1 is associated with a metabolic pattern characterized by low mitochondrial function and high rates of glycolysis (see above); similar patterns are observed in ES cells, somatic stem cells and cancer stem cells [42,78]. Loss of Oct1 function inhibits transformation and tumorigenicity, apparently through its ability to modulate metabolism [14]. Although it remains an open question whether Oct proteins other than Oct4 regulate stem-cell phenotypes, these observations suggest that targeting Oct1 (or the pathways in which it operates) are viable avenues for cancer research.

Oct4 is expressed in the germline and its expression is clearly associated with germ-cell malignancy [79]. Several studies have also identified Oct4 expression in adult somatic stem cells and soma-derived malignancy [80–82]. Although cross-reactivity with other Oct proteins might have resulted in mis-interpretation of data in many of these cases [83], it is hard to ignore the bulk of these data. Further study of this issue should resolve whether Oct4 or other related Oct proteins are involved in the etiology of somatic malignancy.

Curtain call: the future

A striking dichotomy remains between the important roles of Oct transcription factors, e.g. as regulators of stem-cell identity, key components in the generation of pluripotent cells from adult somatic cells, and as mediators of transformation and tumorigenesis, and our understanding of their basic biology. Which pathways control their phosphorylation, sumoylation, ubiquitylation, and perhaps other modifications? How do Oct proteins modulate target gene expression? Are there unknown mechanisms which

regulate the expression, activity and stability of these proteins? What is the biological role of the known Oct1 and Oct4 interacting proteins, and will others be identified? Which Oct proteins have important roles in human tumor etiology and through which pathways? What are the molecular underpinnings of the apparently unique ability of Oct4 to generate iPS cells? By addressing such questions, we should gain a clearer understanding of these interesting and important transcription factors.

Acknowledgements

Our apologies to those many research laboratories whose work we were unable to cite due to space restrictions. We thank D. Stillman and W. Fairbrother for critical reading of the manuscript, and the anonymous reviewers for helpful insights. Our work was supported by the Department of Pathology at the University of Utah, a seed grant from the University of Utah, and awards from the March of Dimes (D.T.) and American Cancer Society (GMC-115196, to D.T.).

References

- Herr, W. *et al.* (1988) The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and *Caenorhabditis elegans* unc-86 gene products. *Genes Dev.* 2, 1513–1516
- Kang, J. *et al.* (2009) A general mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and oxidative stress. *Genes Dev.* 23, 208–222
- Nieto, L. *et al.* (2007) Differential effects of phosphorylation on DNA binding properties of N Oct-3 are dictated by protein/DNA complex structures. *J Mol Biol* 370, 687–700
- Saxe, J.P. *et al.* (2009) Post-translational regulation of Oct4 transcriptional activity. *PLoS ONE* 4, e4467
- Sive, H.L. *et al.* (1986) Multiple sequence elements are required for maximal *in vitro* transcription of a human histone H2B gene. *Mol. Cell Biol.* 6, 3329–3340
- Nichols, J. *et al.* (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95, 379–391
- Latchman, D.S. (1999) POU family transcription factors in the nervous system. *J Cell Physiol* 179, 126–133
- Matthias, P. (1998) Lymphoid-specific transcription mediated by the conserved octamer site: who is doing what? *Semin. Immunol.* 10, 155–163
- Okita, K. *et al.* (2007) Generation of germline-competent induced pluripotent stem cells. *Nature* 448, 313–317
- Ryan, A.K. and Rosenfeld, M.G. (1997) POU domain family values: flexibility, partnerships, and developmental codes. *Genes Dev* 11, 1207–1225
- Phillips, K. and Luisi, B. (2000) The virtuoso of versatility: POU proteins that flex to fit. *J Mol Biol* 302, 1023–1039
- Remenyi, A. *et al.* (2001) Differential dimer activities of the transcription factor Oct-1 by DNA-induced interface swapping. *Mol Cell* 8, 569–580
- Wang, V.E.H. *et al.* (2004) Embryonic lethality, decreased erythropoiesis, and defective octamer-dependent promoter activation in Oct-1-deficient mice. *Mol Cell Biol* 24, 1022–1032
- Shakya, A. *et al.* (2009) Oct1 loss of function induces a coordinate metabolic shift that opposes tumorigenicity. *Nat Cell Biol* 11, 320–327
- Tantin, D. *et al.* (2005) The octamer binding transcription factor Oct-1 is a stress sensor. *Cancer Res* 65, 10750–10758
- Hochedlinger, K. and Plath, K. (2009) Epigenetic reprogramming and induced pluripotency. *Development* 136, 509–523
- Nakagawa, M. *et al.* (2008) Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* 26, 101–106
- Wernig, M. *et al.* (2008) c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* 2, 10–12
- Yu, J. *et al.* (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917–1920
- Huangfu, D. *et al.* (2008) Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol* 26, 1269–1275
- Shi, Y. *et al.* (2008) Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 3, 568–574
- Feng, B. *et al.* (2009) Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat Cell Biol* 11, 197–203
- Chen, X. *et al.* (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133, 1106–1117
- Boyer, L.A. *et al.* (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947–956
- Sridharan, R. *et al.* (2009) Role of the murine reprogramming factors in the induction of pluripotency. *Cell* 136, 364–377
- Tantin, D. *et al.* (2008) High-throughput biochemical analysis of *in-vivo* location data reveals novel distinct classes of POU5F1(Oct4)/DNA complexes. *Genome Res* 18, 631–639
- Zheng, L. *et al.* (2003) S phase activation of the histone H2B promoter by OCA-S, a coactivator complex that contains GAPDH as a key component. *Cell* 114, 255–266
- Tomilin, A. *et al.* (2000) Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. *Cell* 103, 853–864
- Donner, A.L. *et al.* (2007) Sox2 and Pou2f1 interact to control lens and olfactory placode development. *Dev Biol* 303, 784–799
- Di Rocco, G. *et al.* (2001) The recruitment of SOX/OCT complexes and the differential activity of HOXA1 and HOXB1 modulate the Hoxb1 auto-regulatory enhancer function. *J Biol Chem* 276, 20506–20515
- Pfeuffer, I. *et al.* (1994) Octamer factors exert a dual effect on the IL-2 and IL-4 promoters. *J Immunol* 153, 5572–5585
- Murayama, A. *et al.* (2006) A specific CpG site demethylation in the human interleukin 2 gene promoter is an epigenetic memory. *EMBO J* 25, 1081–1092
- Kaushansky, K. *et al.* (1994) Coordinate regulation of multiple human lymphokine genes by Oct-1 and potentially novel 45 and 43 kDa polypeptides. *J Immunol* 152, 1812–1820
- Chew, J.L. *et al.* (2005) Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* 25, 6031–6046
- Rodda, D.J. *et al.* (2005) Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem* 280, 24731–24737
- Wu da, Y. and Yao, Z. (2005) Isolation and characterization of the murine Nanog gene promoter. *Cell Res* 15, 317–324
- Ambrosetti, D.C. *et al.* (2000) Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. *J Biol Chem* 275, 23387–23397
- Wang, D. *et al.* (2000) Transcriptional regulation of the human osteopontin promoter: functional analysis and DNA-protein interactions. *Oncogene* 19, 5801–5809
- Botquin, V. *et al.* (1998) New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2. *Genes Dev* 12, 2073–2090
- Jin, S. *et al.* (2001) Transcription factors Oct-1 and NF-YA regulate the p53-independent induction of the GADD45 following DNA damage. *Oncogene* 20, 2683–2690
- Takahashi, S. *et al.* (2001) Involvement of the Oct-1 regulatory element of the gadd45 promoter in the p53-independent response to ultraviolet irradiation. *Cancer Res* 61, 1187–1195
- Kondoh, H. *et al.* (2007) A high glycolytic flux supports the proliferative potential of murine embryonic stem cells. *Antioxid Redox Signal* 9, 293–299
- Jaenisch, R. and Young, R. (2008) Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* 132, 567–582
- Kim, J. *et al.* (2008) An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 132, 1049–1061
- Covello, K.L. *et al.* (2006) HIF-2[alpha] regulates Oct-4: effects of hypoxia on stem cell function, embryonic development, and tumor growth. *Genes Dev* 20, 557–570
- Zhang, X. *et al.* (2008) Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells. *J Biol Chem* 283, 35825–35833
- Zhang, J. *et al.* (2006) Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol* 8, 1114–1123

- 48 Ding, L. *et al.* (2009) A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. *Cell Stem Cell* 4, 403–415
- 49 Deshpande, A.M. *et al.* (2009) Cdk2ap1 is required for epigenetic silencing of Oct4 during murine embryonic stem cell differentiation. *J Biol Chem* 284, 6043–6047
- 50 Sato, N. *et al.* (2006) The orphan nuclear receptor GCNF recruits DNA methyltransferase for Oct-3/4 silencing. *Biochem Biophys Res Commun* 344, 845–851
- 51 Yang, H.M. *et al.* (2007) Transcriptional regulation of human Oct4 by steroidogenic factor-1. *J Cell Biochem* 101, 1198–1209
- 52 Park, S.W. *et al.* (2007) SUMOylation of Tr2 orphan receptor involves Pml and fine-tunes Oct4 expression in stem cells. *Nat Struct Mol Biol* 14, 68–75
- 53 Epsztejn-Litman, S. *et al.* (2008) De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol* 15, 1176–1183
- 54 Tachibana, M. *et al.* (2008) G9a/ GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *Embo J* 27, 2681–2690
- 55 Tay, Y. *et al.* (2008) MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* 455, 1124–1128
- 56 Niwa, H. *et al.* (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 24, 372–376
- 57 Pankratova, E. *et al.* (2006) Autoregulation of Oct-1 gene expression is mediated by two octa-sites in alternative promoter. *Biochimie* 88, 1323–1329
- 58 Xu, H.M. *et al.* (2004) Wwp2, an E3 ubiquitin ligase that targets transcription factor Oct-4 for ubiquitination. *J Biol Chem* 279, 23495–23503
- 59 Wei, F. *et al.* (2007) Sumoylation of Oct4 enhances its stability, DNA binding, and transactivation. *J Biol Chem* 282, 21551–21560
- 60 Schild-Poulter, C. *et al.* (2007) DNA-PK phosphorylation sites on Oct-1 promote cell survival following DNA damage. *Oncogene* 26, 2980–2988
- 61 Fan, W. *et al.* (2002) BRCA1 regulates GADD45 through its interactions with the OCT-1 and CAAT motifs. *J Biol Chem* 277, 8061–8067
- 62 Wang, R.H. *et al.* (2004) A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint. *Proc Natl Acad Sci U S A* 101, 17108–17113
- 63 Nie, J. *et al.* (1998) Interaction of Oct-1 and automodification domain of poly(ADP-ribose) synthetase. *FEBS Lett* 424, 27–32
- 64 Schild-Poulter, C. *et al.* (2001) The binding of Ku antigen to homeodomain proteins promotes their phosphorylation by DNA-dependent protein kinase. *J Biol Chem* 276, 16848–16856
- 65 Malhas, A.N. *et al.* (2009) Lamin B1 controls oxidative stress responses via Oct-1. *J Cell Biol* 184, 45–55
- 66 Bertolino, E. and Singh, H. (2002) POU/TBP cooperativity: a mechanism for enhancer action from a distance. *Mol Cell* 10, 397–407
- 67 Nakshatri, H. *et al.* (1995) Interaction of Oct-1 with TFIIIB. Implications for a novel response elicited through the proximal octamer site of the lipoprotein lipase promoter. *J Biol Chem* 270, 19613–19623
- 68 Inamoto, S. *et al.* (1997) The cyclin-dependent kinase-activating kinase (CAK) assembly factor, MAT1, targets and enhances CAK activity on the POU domains of octamer transcription factors. *J Biol Chem* 272, 29852–29858
- 69 Lee, J. *et al.* (2005) Stimulation of Oct-4 activity by Ewing's sarcoma protein. *Stem Cells* 23, 738–751
- 70 Tolkunova, E. *et al.* (2007) PIAS proteins as repressors of Oct4 function. *J Mol Biol* 374, 1200–1212
- 71 Zhang, Z. *et al.* (2007) Post-translational modification of POU domain transcription factor Oct-4 by SUMO-1. *FASEB J* 21, 3042–3051
- 72 Guelen, L. *et al.* (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453, 948–951
- 73 Sauter, P. and Matthias, P. (1998) Coactivator OBF-1 makes selective contacts with both the POU-specific domain and the POU homeodomain and acts as a molecular clamp on DNA. *Mol Cell Biol* 18, 7397–7409
- 74 Chen, X. *et al.* (2008) Zfp143 regulates Nanog through modulation of Oct4 binding. *Stem Cells* 26, 2759–2767
- 75 Hatada, E.N. *et al.* (2000) Interaction and functional interference of C/EBPbeta with octamer factors in immunoglobulin gene transcription. *Eur J Immunol* 30, 174–184
- 76 Ben-Porath, I. *et al.* (2008) An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* 40, 499–507
- 77 Somervaille, T.C. *et al.* (2009) Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells. *Cell Stem Cell* 4, 129–140
- 78 Diehn, M. *et al.* (2009) Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature* 4, 4
- 79 Cheng, L. *et al.* (2007) OCT4: biological functions and clinical applications as a marker of germ cell neoplasia. *J Pathol* 211, 1–9
- 80 Atlasi, Y. *et al.* (2007) OCT-4, an embryonic stem cell marker, is highly expressed in bladder cancer. *Int J Cancer* 120, 1598–1602
- 81 Tai, M.H. *et al.* (2005) Oct4 expression in adult human stem cells: evidence in support of the stem cell theory of carcinogenesis. *Carcinogenesis* 26, 495–502
- 82 Zangrossi, S. *et al.* (2007) Oct-4 expression in adult human differentiated cells challenges its role as a pure stem cell marker. *Stem Cells* 25, 1675–1680
- 83 Lengner, C.J. *et al.* (2008) The pluripotency regulator Oct4: a role in somatic stem cells? *Cell Cycle* 7, 725–728