Dissecting the roles of the transcriptional coactivator p300 regulating the oncogenic self-renewal gene *NANOG*

by

Conor Harreld Doss

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2014

Doctoral Committee:

Professor Anna K. Mapp, Chair
Assistant Professor Tomasz Cierpicki
Professor Roger Sunahara
Associate Professor Raymond C. Trievel

Table of Contents

List of Tables

List of Figures

Abstract

Dissecting the roles of the transcriptional coactivator p300 regulating the oncogenic self-renewal gene *NANOG*

by

Conor Harreld Doss

Chair: Anna K. Mapp

Eukaryotic gene expression is controlled by the concerted actions and interactions of DNA-regulatory elements, transcriptional activators and associated coactivators. Successful assembly of the proper activator and coactivators at a target gene promoter leads to the stimulation of RNA polymerase activity and the transcription of the gene into messenger RNA. Studying the interactions leading up to transcriptional initiation is challenging due to the relatively weak and promiscuous nature of activator-coactivator interactions. This dissertation documents the identification of coactivator functions that control the expression of the medically relevant target gene *NANOG*.

NANOG is an embryonic transcription factor that confers tumorigenic and self-renewing potential when expressed in human cancer cells, yet its regulation is poorly understood and few methods are currently available to block *NANOG* function. We reasoned that modulating the coactivator(s) that regulate *NANOG* would allow for control over *NANOG* expression. To achieve this control, we identified the histone acetyltransferase (HAT) p300 as a necessary and direct coactivator of *NANOG* expression in a variety of cancer cells.

In order to better manipulate p300, we identified critical p300 domains involved in *NANOG* expression by systematically deleting, mutating, and inhibiting all potential p300 interaction surfaces and catalytic functions. The activator-binding domain CH1 was found to be essential for p300-driven *NANOG* expression, suggesting the CH1 domain may act in localizing p300 to the *NANOG* promoter prior to transcription. Additionally, p300 HAT domain activity was found to be necessary for maintaining high levels of histone acetylation at the *NANOG* promoter and for maintaining *NANOG* expression in cancer cells. These results allow us to propose a model for the activator-coactivator interactions that drive *NANOG* transcription. This model will assist in guiding future drug discovery and chemical probe design by providing the first validated targets capable of downregulating *NANOG* expression in cancer cells.

Chapter 1

Regulatory targets in cancer-initiating cells and the challenges in inhibiting them

1.1. Abstract

Cancer-initiating cells (CICs) represent an emerging and attractive target for therapeutic intervention given their critical roles in tumor development and maintenance. However, studying CICs *in vitro* is challenging due to issues with their identification and stability in isolation. A variety of techniques have been developed to aid CIC research, including the application of small molecule inhibitors and probes to examine their internal functions. Here we introduce the concept of CICs in cancer pathogenesis, as well as review how early small molecule approaches to CICs have generated meaningful insights into CIC biology and show great promise for future efforts to elucidate the nature of this elusive cell type.

1.2. Cancer-initiating cells

Cancer-initiating cells (CICs, also known as cancer stem cells) are a unique subpopulation of cancer cells that possess the ability to generate and maintain tumors (1-4). As a result, even small populations of CICs that remain after conventional treatment may regenerate the tumor following therapy. CICs are often associated with highly drug- and radiation-resistant cancer cell populations,

as well metastatic cells. These observations suggest that traditional anticancer strategies are largely inadequate for the eradication of CICs. Thus, new CIC-targeting therapies are urgently needed to improve the prognosis of cancer patients.

CICs share many features with normal stem cells such as self-renewal potential and the capacity to differentiate. Unfortunately, these similarities also translate into difficulties in studying CICs *in vitro*. First, CICs in isolation are difficult to maintain due to their intrinsic instability and interchangeability (5). Second, CICs are poorly defined in many tumors, leading to little or even conflicting information regarding the actual identity of the CICs (6). Lastly, efforts to manipulate CICs *in vitro* often require complicated gene knockout or knockdown systems that can destroy important signaling or interaction networks that may artificially alter the identity of the CIC.

Because of the difficulties outlined above, developing new tools for studying CICs has proven challenging. In order to fully appreciate the strategies that have been successfully employed against CICs, it is important to first discuss some of the defining characteristics of CICs and how they can be exploited for CIC-targeting inhibitor development.

The CIC hypothesis states, in essence, that cancer cells are hierarchically organized with a CIC at the apex driving tumor growth. (Fig. 1.1A)(1-3) When the

CIC divides, it can produce two general types of daughter cells. First, the CIC can produce a daughter cell that resembles the parental CIC (a process termed

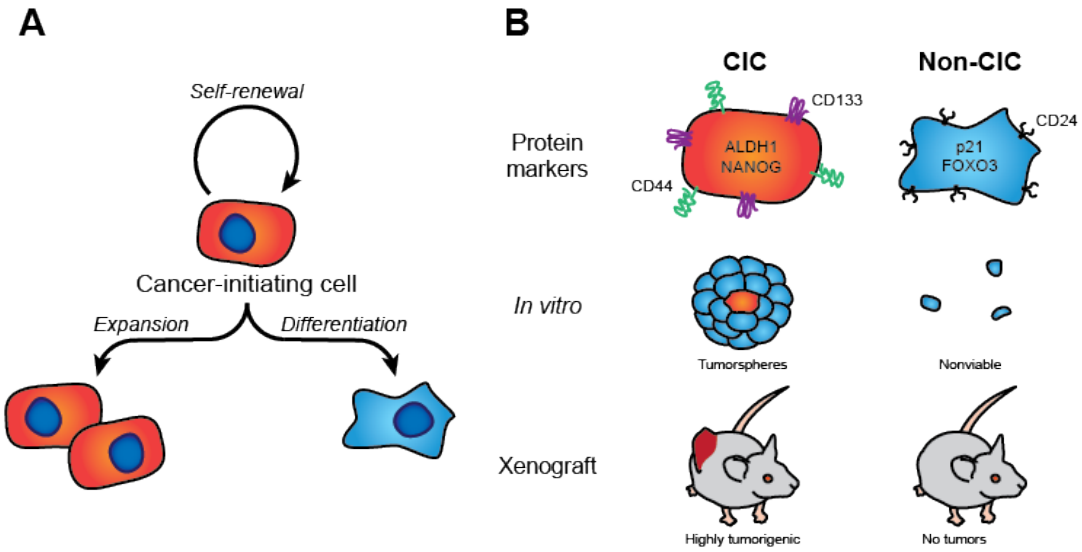

Figure 1.1: Features of cancer-initiating cells (CICs). (A) CICs (red) are defined by their ability to differentiate into the heterogeneous non-CIC cell types found within in a tumor (blue), as well as to undergo self-renewal during division and remain in a multipotent state. Additionally, CICs can undergo symmetrical division to produce two multipotent daughter cells and thereby expand the pool of CICs. (B) CICs and non-CICs can be distinguished based on: i) the differential expression of various protein markers such as CD44 and NANOG; ii) their capacity to form tumorspheres in suspension culture; and iii) their relative ability to found new tumors in immunocompromised mouse xenografts. CICs are highly tumorigenic and readily initiate new tumors while non-CICs cannot.

self-renewal), thereby regenerating or even expanding the pool of CICs present in the tumor. Alternatively, the resulting daughter cell can exit the self-renewing state and differentiate into a specific, fixed identity. These non-CICs represent a "mature" state analogous to the post-mitotic, terminally differentiated cells found in normal tissues. Non-CICs possess little to no tumorigenicity, and typically express known markers of differentiation (Fig. 1.1B). Since only the CICs are able to self-renew, the CICs are exclusively capable of generating tumorspheres *in vitro* or initiating new tumors when injected into immunocompromised mice

(Fig. 1B). Together, the features of self-renewal, infinite growth, and differentiation status distinguish CICs from non-CICs.

These defining features all have a series of underlying regulatory mechanisms within the cell to ensure their sustained activity. The goal of CIC-inhibitor development is to first identify these mechanisms/pathways and then design methods to shut them down. Recently, small molecule tools have proven to be effective for both the discovery and inhibition of essential CIC programs. Here we will review how small molecules have advanced the study of CIC biology before closing with a discussion on how to identify and exploit other unique features of CICs including their transcriptional networks and protein-protein interactions.

1.3. Small molecule CIC inhibitors

In principle, the defining features of self-renewal and multipotency also represent the most effective means to selectively deplete CICs from a mixed cell population. A CIC-targeting inhibitor can alter the cellular identity by any combination of three distinct effects (Fig. 1.2):

1. *Cell death*: An inhibitor is acutely cytotoxic or activates apoptosis in CICs
2. *Growth arrest/senescence*: The small molecule slows or ceases CIC cell cycle progression, thereby halting self-renewal and growth
3. *Differentiation*: Small molecule treatment causes the CICs to change and/or commit to a particular cell lineage, thereby exiting the CIC state

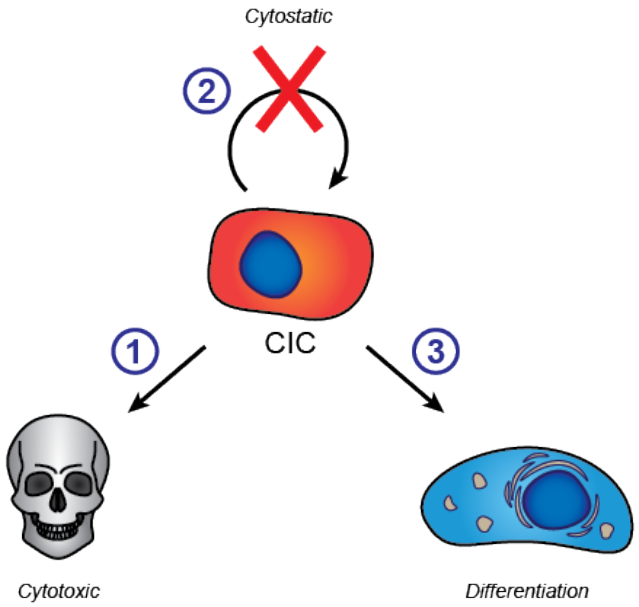

Figure 1.2: Effects of CIC-targeting inhibitors. CICs can be eliminated by inhibitors through one of three possible routes: (1) The inhibitor induces cell death, either by acute cytotoxicity or by inducing apoptosis. (2) The inhibitor halts the cell cycle or self-renewal process, preventing the CICs from expanding and further driving tumor growth. (3) The inhibitor induces differentiate and thereby forces the CICs to exit the self-renewing state.

A number of small molecules have been reported to selectively affect CICs (Table 1.1). Most frequently, these compounds target key signaling enzymes that regulate CIC maintenance such as PI3K/mTOR and γ–secretase that regulate CIC maintenance but, as will be discussed below, these are not the only possible targets. In fact, several of the most successful CIC inhibitors discovered to date act via unknown mechanisms. Identifying the potential mechanisms in these cases can provide highly illuminating insights into CIC biology and thereby inform inhibitor design in the future.

1.3.1 CIC inhibitors with known targets

Conventional CIC inhibitor development follows the reverse chemical genetics approach wherein a cellular protein is first identified as relevant for CIC regulation, and inhibitors are chosen or designed to block the function of that protein. For example, Lonardo et al. (7), found that the Nodal signaling pathway was found to be a driver of self-renewal in pancreatic CICs through a series of knockdown experiments silencing either the Nodal ligand or its receptor, ALK4. Pharmacological inhibition of ALK4 by the compound SB431542 (Fig. 1.3A) dose-dependently suppressed CIC-driven tumorsphere formation and blocked Nodal-enhanced growth. In contrast, the ALK5 antagonist LY2157299 had no effect on tumorsphere formation, indicating that a distinct set of TGFβ receptors may be responsible of pancreatic CIC maintenance.

TABLE 1: CIC-Selective Inhibitors

Biological Pathway	Target	Compound	*In vitro* IC50 (nM)	Effects	References
Notch	γ-secretase	DAPT	~20	Blocks the response of Nestin+ GBM CICs to hypoxia; reduces the fraction of CD44+/CD24- breast CICs in c	41-46
		GSI-18	~50	Depleted CD133+ GBM CIC fraction; reduces neurosphere formation and promotes differentiation *in vitro*	49-51
		RO4929097	~4.0	Reduces melanoma CIC potential	52
		DBZ	1.0-2.0	Suppresses Bcl-2 expression in prostate CICs and restores docetaxel sensitivity when combined with Hh inhi	53
		L685485			54
		MRK-003	~0.5	Irreversibly inhibits CICs mammosphere formation, reduces CIC frequency, and induces differentiation in a m	55-59
TGFβ/SMAD	ALK4/7	SB431542	90-100	Reduces the fraction of CD49f+/CD61+ breast CICs and reversed EMT phenotypes; Dose-dependently inhibi	7, 60
	ALK5	A-83-01	10-15	Reduces the fraction of CD49f+/CD61+ breast CICs and reversed EMT phenotypes	60
		LY2109761	30-40 for TßRI; ~300 fo	Restored an epithelial phenotype to CD44+ breast CICs	61
Hedgehog	SMO	Cyclopamine	40-50	Reduces AML CIC self-renewal *in vitro*; blocks AKT activation in prostate CICs and restores docetaxel sensiti	44,57,62-66
		GDC-0449	~3	Blocks AKT activation in prostate CICs and restores doceetaxel sensitivity in combination with Notch inhibitor	66
		IPI-926	5.0-9.0	Reduces AML CIC self-renewal *in vitro* and *in vivo*	63,67
JAK/STAT	STAT3	S31-201	80,000-90,000	Inhibits expression of the *NANOG* self-renewal gene in hepatocellular CICs; prevents neurosphere growth in	10, 68
		FLLL32	1,000-5,000	Reduced viability and tumorsphere formation in colon CICs	69
		WP1066	~2,500	Inhibits the hypoxia response in Nestin+ GBM CICs	42
		STA-21	10,000-20,000	Prevents neurosphere growth in GBM CICs; depletes GBM CIC marker expression	68
		Stattic	4,300-5,900	Induced apoptosis in GBM CICs	70
	JAK2	NVP-BSK805	~0.5	Reduced breast CIC number in established xenografts	71
RTK/PI3K	PI3K/mTOR	LY294002	500-1,000	Inhibits CIC survival in pancreatic and breast cancer cell lines; restores drug sensitivity in GBM CICs	18, 72, 73
		NVP-BEZ235	5-70	Inhibits CIC survival in pancreatic cell lines; induces GBM CIC differentiation	72, 74, 75
	FKBP12/mTORC1	Rapamycin	~1.0	Inhibits CIC survival in breast cancer cell lines; restores drug sensitivity?; Blocks the hypoxia response in Nes	18, 42, 73
MAPK	MEK1/2	U0126	~60	Synergizes with FTIs to induce CML CIC apoptosis;	74-79
		SL327	150-250	Induces GBM CIC differentiation	74, 75, 77
		PD184352	~17	Inhibits breast CIC self-renewal *in vitro*; synergizes with FTIs to induce CML CIC apoptosis	79, 80
	ERK1/2	PD98059	~2000	Blocks the response of Nestin+ GBM CICs to hypoxia;	18
Histone modifications	HDAC class I	Trichostatin A	1.0-2.0	Reduces ALDH1+, CD133+ GBM neurosphere-derived cells	33
		LBH589	5.0-20	Eliminates CML CICs when combined with imatinib	34
		SAHA			81, 82
		MS-275	5.0-20	Reduces ALDH1+, CD133+ GBM neurosphere-derived cells	33
	LSD1	Tranylcypromine	~2.0	Induces AML CIC differentiation *in vitro*, depletes CIC population *in vivo*	36
Wnt/β-catenin	???	Salinomycin		Reduces CIC fraction in breast and CLL cell lines	5, 12, 15-17
NF-κB	???	Parthenolide		Selectively induces apoptosis and ROS production in AML CICs	83, 84
Dopamine receptors	DR	Thioridazine	*Ki* = 5-10 nM	Induces AML CIC differentiation and granulocytic maturation	25
DNA Repair	CHK1/2	SB218078	~15	Reduces the CD133+ colon CIC fraction; increases cytotoxicity due to DNA damage in lung CICs	85, 86
		AZD7762	5-10	Restores lung CIC sensitivity to chemotherapeutics *in vivo*	85
		DBH	3,000-3,500	Restores GBM CIC sensitivity to ionizing radiation	87
	CHK1	UCN-01	~25	Reduces the CD133+ colon CIC fraction	86
Leukotriene biosynthesi	ALOX5	Zileuton	400-900	Blocks CML CIC maturation into malignant leukemia cells	88
Lipidation	Farnesyltransferase	BMS-214662	1,300-1,400	Induces apoptosis in quiescent CML CICs	79, 89
Nuclear receptors	PPARγ	Ciglitazone		Inhibited GBM CIC growth and modulated expression of stemness genes	91
		15-deoxy-Δ12,14-prostaglandinJ2		Inhibited GBM CIC growth and modulated expression of stemness genes	91
Telomerase	hTERT	imetelstat			92
Wnt/ß-catenin	???	Sulforaphane			93
	???	Curcumin			94, 95
	CXCR1	Repertaxin			96
		Metformin			97, 98
Translation	Ribosome	omacetaxine		Induces apoptosis in primitive CML CICs	99

Abbreviations: FTI (farnesyltransferase inhibitor); CML (chronic myeloid leukemia); CLL (chronic lymphoid leukemia); AML (acute myeloid leukemia); GBM (glioblastoma multiforme); EMT (epithelial-mesenchymal transition)

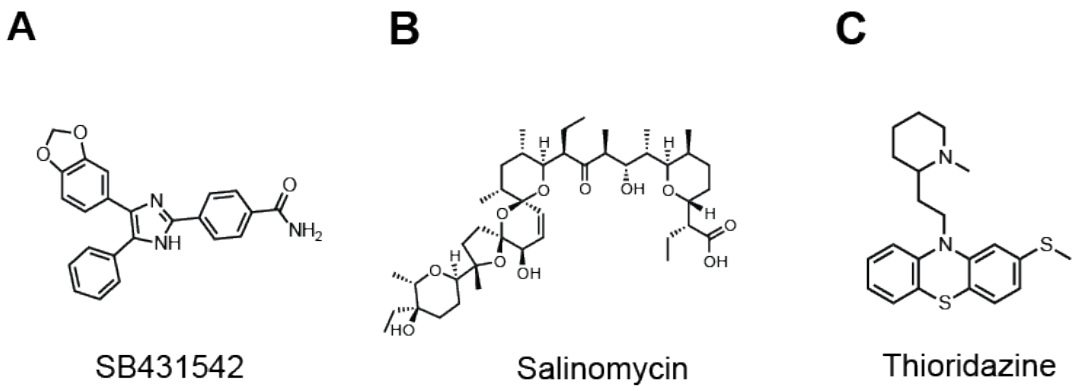

Figure 1.3: Notable CIC-selective small molecules. (A) The TGFβ/ALK receptor inhibitor SB431542. (B) The potassium ionophore salinomycin. (C) The dopamine receptor antagonist thioridazine.

This type of reverse chemical genetics approach has also been applied to other known oncogenes such as STAT3. Constitutive JAK/STAT3 signaling is frequently observed in various cancers and relies upon tyrosine phosphorylation to recruit and activate the transcription factor STAT3 (8). STAT3 recognizes phosphotyrosine residues via its SH2 domain, and numerous small molecules have been developed as competitive SH2 ligands (9). When STAT3 was later identified as a CIC-relevant target, those same SH2 inhibitors demonstrated anti-CIC activity. For instance, the compound S31-201 was shown to deplete CICs in hepatocellular carcinoma cell lines, in part by downregulating the stem cell self-renewal *NANOG* (10). The observation that STAT3 signaling could drive *NANOG* expression in somatic tissues like liver cells suggests that one oncogenic role for STAT3 may be establishing embryonic self-renewal programs in CICs. These cells thus adopt a less mature, more stem-like phenotype and

become freed from extrinsic homeostatic signals, helping to accelerate tumor progression.

These studies highlight how CIC inhibitors can aid in validating other discoveries in CIC biology. However, the study of CICs has also benefitted from the use of so-called "orphan" CIC inhibitors, compounds that as of yet have no clear intracellular targets. Delineating the mechanisms by which these orphan inhibitors affect CICs could reveal novel details of CIC generation and maintenance.

1.3.2 CIC inhibitors without known targets

Forward chemical genetic screens identifying compounds that selectively kill CICs by conventional high-throughput assays have proven technically difficult due to the low frequency, high heterogeneity, and instability of CICs *in vitro*. To overcome these limitations, Gupta *et al*. (5) generated an artificially enriched CIC population by inducing an epithelial-to-mesenchymal transition (EMT) cell phenotype. EMT is an evolutionarily conserved developmental process associated with cellular migration and invasiveness, as well as the acquisition of stem-like properties in cancer cells (11). Stable knockdown of the epithelial marker E-cadherin induced EMT and greatly enriched for the putative $CD44^{+}/CD24^{lo}$ breast CIC population. These induced-EMT cells were stable enough to serve as a screening platform for compounds exhibiting selective cytotoxicity compared to the uninduced parental cells (5). Using a commercially

available library, 32 compounds were identified as hits showing modest selectivity for the transdifferentiated cells, including the natural product salinomycin (Fig. 1.3B; Table 1). Salinomycin treatment alone depleted the $CD44^{+}/CD24^{lo}$ fraction and strongly inhibited tumor engraftment in a mouse xenograft model. Most strikingly, salinomycin also proved effective against established tumors in mice, inducing both cancer cell differentiation and death, suggesting that salinomycin could target and eliminate the CIC subpopulation *in vivo*.

Salinomycin is a highly selective potassium ionophore, suggesting that Na^{+}/K^{+} ion gradients may be finely balanced in CICs. In support of this hypothesis, another polyether potassium ionophore, nigericin, was reported to target breast CICs (5) and possess similar anticancer effects as salinomycin (12, 13).

Salinomycin has been used for decades as a veterinary antibiotic. However, its activity against human CICs was unexpected. Subsequent studies have demonstrated broad anti-CIC activity, targeting the cancer-initiating populations in cancers of the breast (5), lung (14), stomach (15), and bone (16), as well as in chronic myelogenous leukemia (17)(Table 1). In addition to its activity as a cation chelator, salinomycin has also been reported to inhibit the ABC family of multidrug transporters that are frequently upregulated in CICs (17-20). This activity could account for the increased sensitivity of the putative CICs to conventional chemotherapeutics in the presence of salinomycin (15-17).

However such observations do not explain why salinomycin alone displays CIC-selective toxicity.

The general activity of salinomycin against such physiologically diverse tumors may suggest that CICs all possess some common target essential for their self-renewal and survival. Lu *et al*. (12) saw that in primary CLL cells, salinomycin downregulated components of the Wnt/β-catenin signaling pathway, a developmental pathway frequently associated with stem cell and CIC self-renewal (21). Salinomycin has also been show to inhibit GSK3β phosphorylation and β-catenin accumulation in osteosarcoma cells, further supporting the idea that Wnt/β-catenin may be the intracellular target (16). Unfortunately, there are few Wnt pathway inhibitors currently available, with active small molecules only recently being reported (22). Better resolution of salinomycin's mechanism of action will require further experiments to determine whether, in fact, Wnt signaling inhibition is the main activity against CICs, or if there are indirect effects like oxidative or metabolic stress that are not coupled with Wnt downregulation (23).

Most discovery and screening efforts with CICs suffer due to a lack of robust *in vitro* models. Gupta et al. (5) created an artificial EMT assay that allowed them to identify salinomycin as an inhibitor of EMT-dependent processes in CICs. However, CICs can also be defined by other phenotypes such as the expression of undifferentiated markers and resistance to differentiation. Bhatia and

colleagues (24) generated a neoplastic varianthuman ESCs (v-hESC) that recapitulate many features of CICs including high tumorigenicity, enhanced self-renewal, and impaired differentiation potential. These v-hESCs were stable enough *in vitro* to allow for high-throughput screening for compounds capable of inducing differentiation (25).

Unexpectedly, the dopamine receptor (DR) antagonist and antipsychotic drug thioridazine was found to induce v-hESC differentiation (Fig. 1.3C; Table 1). Furthermore, racemic thioridazine reduced the CIC fraction in primary AML cells and increased the expression of mature granulocytic markers, suggesting that thioridazine is capable of inducing differentiation in a therapeutically relevant AML disease model. More importantly, however, thioridazine had no effect on the differentiation of umbilical cord HSCs, indicating a selective effect on the cancer cells. Salinomycin, the previous standard for CIC-targeting small molecules, proved cytotoxic to both HSCs and AML CICs.

The discovery of a CIC-selective DR antagonist prompted an investigation into the possible role of DR signaling in AML CICs. Sachlos *et al*. (25) found that AML blasts frequently expressed all 5 DR subtypes, and that high DR levels overlapped and correlated with the CIC populations. The enantiomers of thioridazine display between 2-10-fold differences in DR subtype selectivity (26), leading to potential questions about the relevant receptor targets. More importantly, the DR subtypes can have opposing signaling effects, with D_1 and D_5

stimulating intracellular cAMP production while D_2, D_3, and D_4 directly repress cAMP production. To clarify the mechanism of thioridazine in CICs, Sachlos *et al*. (25) also tested family-specific DR ligands. The D_1/D_5 antagonist SKF-38,393 blocked AML proliferation, while the D_2 agonist 7OH-DPAT actually promoted AML cell growth. These results suggest that shutting down the DR-mediated production of cAMP may be an effective way of inhibiting AML CIC self-renewal.

As a chemical probe, thioridazine suffers due to a low level of receptor selectivity. In addition to DR, thioridazine and related phenothiazines have been reported to affect various other intracellular targets such as AKT and MALT1 (27, 28). Furthermore, the effective CIC-targeting concentrations were much greater than the typical K_i for DR (1 nM vs. 1 μM; Table 1), possibly indicating a role for additional off-target effects of thioridazine.

From a clinical standpoint, however, these results were encouraging. Thioridazine and other DR antagonists are already FDA-approved and have been prescribed for schizophrenic patients for years, but their effects on CIC populations had never been examined. Further work will be needed to delineate the role of DR in CICs, as well as determining whether gain of DR function is a transformative (early) or self-reinforcing (late) step in tumor development and progression. Remarkably, patients receiving DR antagonist medications have reduced incidences of cancer relative to the general population (29), suggesting

that DR antagonism may well be prophylactic in regards to disease prevention and recurrence.

1.4. Unaddressed targets in CIC biology

As the studies described above illustrate, small molecule inhibitors have proven successful in selectively blocking crucial functions. However, these studies are also biased towards signaling pathways and catalytic activities that are components of larger, interconnected regulatory networks in CICs (Table 1.1). CICs undoubtedly rely upon multiple internal and external forms of regulation to maintain their identity and self-renewal, and thus there are other potential avenues to explore modulating CIC behavior.

For example, one of the central cell-autonomous regulatory networks in CICs is their unique transcriptional and gene expression profile (30-32). A number of CIC-selective inhibitors target transcriptional coregulatory proteins such as histone deacetylases (33-35) or lysine-specific demethylase 1 (LSD1) (36). These compounds impair CIC growth and self-renewal, however their mechanisms remain unclear. CIC-associated gene networks such as the ESC-like, *NANOG*-driven network (31, 37, 38) are fundamental to CIC identity and therefore represent potential targets for new modulatory strategies. However, these networks operate through protein-protein interactions (PPI) and to date no PPI inhibitors have been pursued as CIC-targeting probes.

The dearth of PPI inhibitors in use for studying CICs originates, in part, due to relatively limited information available regarding the PPIs taking place in CICs. PPIs between transcription factors and coregulatory proteins are often weak (low-micromolar affinities) and mediated via intrinsically disordered domains (39, 40) that hinder structural characterization. Furthermore, transcriptional regulatory proteins can engage in a variety of multivalent and/or cooperative interactions depending on gene context, thereby creating a large number of possible interactions researchers must consider when attempting to probe CIC regulatory networks.

These challenges are not insurmountable however, and we view PPIs in CIC biology as a ripe and currently under-explored field. The work presented in this thesis focuses on our efforts to advance the identification and modulation of PPIs in CIC transcriptional networks, particularly through coactivator proteins essential for driving self-renewal genes in cancer.

1.5. Research goals

One of the goals of this research is to identify essential activators and coactivators of self-renewal genes when little is known about the causes of aberrant gene expression. Specifically, we present our analysis of the coactivator protein p300 and its role in driving expression of the self-renewal gene *NANOG.* We show in Chapter 2 a reasonable method to narrow down a list of candidate activators and coactivators, and demonstrate how the hypotheses

generated can be systematically tested and confirmed. The second challenge in studying multidomain coactivator proteins is dissecting the relevant functions within the protein and producing a clearer picture of how proteins interact with each other and in what order.

A second goal is to generate a molecular-level picture of how activators and coactivators interact and cooperate to drive *NANOG* expression. In Chapter 3 we present a systematic analysis of p300 and show how a combination of mutagenesis and small molecule studies identify two crucial domains for p300-mediated *NANOG* expression. This information allows us to form a more detailed model of how transcriptional complexes are assembled at the *NANOG* promoter. The results and methods present in this thesis thus aid the study of CICs by identifying several new targets for small molecule and potential drug design. However, CIC biology still possesses several technical challenges that require the development of new tools. These challenges and potential ways to overcome them are discussed in Chapter 4.

1.6. References

1. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105-111 (2001).
2. Lobo NA, Shimono Y, Qian D, Clarke MF. The biology of cancer stem cells. *Annu. Rev. Cell Dev. Biol.* **23**, 675-699 (2007).
3. Visvader JE & Lindeman. Cancer stem cells in solid tumors: accumulating evidence and unresolved questions. *Nat. Rev. Cancer* **8**, 755-768 (2008).
4. Dalerba P, Cho RW, Clarke MF. Cancer stem cells: models and concepts. *Annu. Rev. Med.* **58**, 267-284 (2007).
5. Gupta PB, *et al*. Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* **138**, 645-659 (2009).
6. Quintana E, *et al*. Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell* **18**, 510-523 (2010).
7. Lonardo E, *et al*. Nodal/Activin signaling drives self-renewal and tumorigenicity of pancreatic cancer stem cells and provides a target for combined drug therapy. *Cell Stem Cell* **9**, 433-446 (2011).
8. Bowman T, Garcia R, Turkson J, Jove R. STATs in oncogenesis. *Oncogene* **21**, 2474-2488 (2000).
9. Costantino L, Barlocco D. STAT3 as a target for cancer drug discovery. *Curr. Med. Chem.* **15**, 834-843 (2008).
10. Lee TKW, *et al*. $CD24^{+}$ liver tumor-initiating cells drive self-renewal and tumor initiation through STAT3-mediated NANOG regulation. *Cell Stem Cell* **9**, 50-63 (2011).
11. Floor S, van Staveren WCG, Larsimont D, Dumont JE, Maenhaut C. Cancer cells in epithelial-to-mesenchymal transition and tumor-propagating-cancer stem cells: distinct, overlapping or same populations. *Oncogene* **30**, 4609-4621 (2011).
12. Lu D, *et al*. Salinomycin inhibits Wnt signaling and selectively induces apoptosis in chronic lymphocytic leukemia cells. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13253-13257 (2011).
13. Riccioni R, *et al*. The cancer stem cell selective inhibitor salinomycin is a p-glycoprotein inhibitor. *Blood Cells Mol. Dis.* **45**, 86-92 (2010).
14. Wang Y. Effects of salinomycin on cancer stem cell in human lung adenocarcinoma A549 cells. *Med. Chem.* **7**, 106-111 (2011).
15. Zhi QM, *et al*. Salinomycin can effectively kill $ALDH^{high}$ stem-like cells on gastric cancer. *Biomed. Pharmacother.* **65**, 509-515 (2011).
16. Tang QL, *et al*. Salinomycin inhibits osteosarcoma by targeting its tumor stem cells. *Cancer Lett.* **311**, 113-121 (2011).

17. Fuchs D, Daniel V, Sadeghi M, Opelz G, Naujokat C. Salinomycin overcomes ABC transporter-mediated multidrug and apoptosis resistance in human leukemia stem cell-like KG-1a cells. *Biochem. Biophys. Res. Commun.* **394**, 1098-1104 (2010).
18. Bleau AM, *et al*. PTEN/PI3K/Akt pathway regulates the side population phenotype and ABCG2 activity in glioma tumor stem-like cells. *Cell Stem Cell* **4**, 226-235 (2009).
19. Riccioni R, *et al*. The cancer stem cell selective inhibitor salinomycin is a p-glycoprotein inhibitor. *Blood Cells Mol. Dis.* **45**, 86-92 (2010).
20. Dean M. ABC transporters, drug resistance, and cancer stem cells. *J. Mammary Gland Biol. Neoplasia* **14**, 3-9 (2009).
21. Clevers H. Wnt/beta-catenin signaling in development and disease. *Cell* **127**, 469-480 (2006).
22. Dodge ME & Lum L. Drugging the cancer stem cell compartment: lessons learned from the Hedgehog and Wnt signal transduction pathways. *Annu. Rev. Pharmacol. Toxicol.* **51**, 289-310 (2011).
23. Ketola K, *et al*. Salinomycin inhibits prostate cancer growth and migration via induction of oxidative stress. *Br. J. Cancer* **106**, 99-106 (2012).
24. Werbowetski-Ogilvie TE, *et al*. Characterization of human embryonic stem cells with features of neoplastic progression. *Nat. Biotechnol.* **27**, 91-97 (2009).
25. Sachlos E, *et al*. Identification of drugs including a dopamine receptor antagonist that selectively target cancer stem cells. *Cell* **149**, 1284-1297 (2012).
26. Choi S, *et al*. Synthesis, receptor binding and functional studies of mesoridazine stereoisomers. *Bioorg. Med. Chem. Lett.* **14**, 4379-4382 (2004).
27. Choi JH, *et al*. Potential inhibition of PDK1/Akt signaling by phenothiazines suppresses cancer cell proliferation and survival. *Rec. Adv. Clin. Oncol.* **1138**, 393-403 (2008).
28. Nagel D, *et al*. Pharmacologic inhibition of MALT1 protease by phenothiazines as a therapeutic approach for the treatment of aggressive ABC-DLBCL. *Cancer Cell* **22**, 825-837 (2012).
29. Dalton SO, Mellemkaer L, Thomassen L, Mortensen PB, Johansen C. Risk for cancer in a cohort of patients hospitalized for schizophrenia in Denmark, 1969-1993. *Schizophr. Res.* **75**, 315-324 (2005).
30. Ben-Porath I, *et al*. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet.* **40**, 499-507 (2008).
31. Chiou SH, *et al*. Positive correlations of Oct4- and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma. *Clin. Cancer Res*. **14**, 4085-4095 (2008).
32. Widschwendter M, *et al*. Epigenetic stem cell signature in cancer. *Nat. Genet.* **39**, 157-158 (2007).

33. Sun P, *et al*. *DNER*, an epigenetically modulated gene, regulates glioblastoma-derived neurosphere cell differentiation and tumor propagation. *Stem Cells* **2009**, 1473-1486 (2009).
34. Zhang B, *et al*. Effective targeting of quiescent chronic myelogenous leukemia stem cells by histone deacetylase inhibitors in combination with imatinib mesylate. *Cancer Cell* **17**, 427-442 (2010).
35. Nimmanapalli R, *et al*. Cotreatment with the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) enhances imatinib-induced apoptosis of Bcr-Abl-positive human acute leukemia cells. *Blood* **101**, 3236-3239 (2003).
36. Harris WJ, *et al*. The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells. *Cancer Cell* **21**, 473-487 (2012).
37. Loh YH, *et al*. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431-440 (2006).
38. Jeter CR, *et al*. NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. *Oncogene* **30**, 3833-3845 (2011).
39. Demarest SJ, *et al*. Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549-553 (2002).
40. Dyson HJ & Wright PE. Intrinsically unstructured proteins and their functions. *Rev. Mol. Cell Biol.* **6**, 197-208 (2005).
41. Gilbert CA, Daou MC, Moser RP, Ross AH. Gamma-secretase inhibitors enhance temozolomide treatment of human gliomas by inhibiting neurosphere repopulation and xenograft recurrence. *Cancer Res.* **70**, 6870-6879 (2010).
42. Qiang L, *et al*. HIF-1α is critical for hypoxia-mediated maintenance of glioblastoma stem cells by activating Notch signaling pathway. *Cell Death Diff.* **19**, 284-294 (2012).
43. Vathipadiekal V, *et al*. Identification of a potential ovarian cancer stem cell gene expression profile from advanced stage papillary serous ovarian cancer. *PLoS ONE* **7**, e29079 (2012).
44. Wang CYY, *et al*. Hedgehog and Notch signaling regulate self-renewal of undifferentiated pleomorphic sarcomas. *Cancer Res.* **72**, 1013-1022 (2012).
45. McGowan PM, *et al*. Notch1 inhibition alters the CD44(hi)/CD24(lo) population and reduces the formation of brain metastases from breast cancer. *Mol. Cancer Res.* **9**, 834-844 (2011).
46. Wang R, *et al*. Glioblastoma stem-like cells give rise to tumour endothelium. *Nature* **468**, 829-833 (2010).
47. Sullivan JP, *et al*. Aldehyde dehydrogenase activity selects for lung adenocarcinoma stem cells dependent on Notch signaling. *Cancer Res.* **70**, 9937-9948 (2010).

48. Harrison H, *et al*. Regulation of breast cancer stem cell activity by signaling through the Notch4 receptor. *Cancer Res.* **70**, 709-718 (2010).
49. Fan X, *et al*. Notch pathway inhibition depletes stem-like cells and blocks engraftment in embryonal brain tumors. *Cancer Res.* **66**, 7445-7452 (2006).
50. Fan X, *et al*. NOTCH pathway blockade depletes CD133-positive glioblastoma cells and inhibits growth of neurospheres and xenografts. *Stem Cells* **28**, 5-16 (2010).
51. Hu YY, *et al*. Notch signaling contributes to the maintenance of both normal and patient-derived glioma stem cells. *BMC Cancer* **11**, 82 (2011).
52. Huynh C, *et al*. The novel gamma secretase inhibitor RO4929097 reduces the tumor initiating potential of melanoma. *PLoS ONE* **6**, e25264 (2011).
53. Domingo-Domenech J, *et al*. Suppression of acquired docetaxel resistance in prostate cancer through depletion of Notch- and Hedgehog-dependent tumor-initiating cells. *Cancer Cell* **22**, 373-388 (2012).
54. Wang Y. Effects of salinomycin on cancer stem cell in human lung adenocarcinoma A549 cells. *Med. Chem.* **7**, 106-111 (2011).
55. Kondratyev M, *et al*. Gamma-secretase inhibitors target tumor-initiating cells in a mouse model of ERBB2 breast cancer. *Oncogene* **31**, 93-103 (2012).
56. Grudzien P, *et al*. Inhibition of Notch signaling reduces the stem-like population of breast cancer cells and prevents mammosphere formation. *Anticancer Res.* **30**, 3853-3867 (2010).
57. Cao L, *et al*. Sphere-forming cell subpopulations with cancer stem cell properties in human hepatoma cell lines. *BMC Gastroenterol.* **11**, 71 (2011).
58. Cullion K, *et al*. Targeting the Notch1 and mTOR pathways in a mouse T-ALL model. *Blood* **113**, 6172-6181 (2009).
59. Tatarek J, *et al*. Notch1 inhibition targets the leukemia-initiating cells in a *Tal1/Lmo2* mouse model of T-ALL. *Blood* **118**, 1579-1590 (2011).
60. Lo PK, *et al*. CD49f and CD61 identify Her2/neu-induced mammary tumor-initiating cells that are potentially derived from luminal progenitors and maintained by the integrin-TGFβ signaling. *Oncogene* **31**, 2614-2626 (2012).
61. Shipitsin M, *et al*. Molecular definition of breast tumor heterogeneity. *Cancer Cell* **11**, 259-273 (2007).
62. Zhao C, *et al*. Hedgehog signaling is essential for maintenance of cancer stem cells in myeloid leukaemia. *Nature* **458**, 776-779 (2009).
63. Lin TL, *et al*. Self-renewal of acute lymphocytic leukemia cells is limited by the hedgehog pathway inhibitors cyclopamine and IPI-926. *PLoS ONE* **5**, e15262 (2010).

64. Jinushi M, *et al*. Tumor-associated macrophages regulate tumorigenicity and anticancer drug responses of cancer stem/initiating cells. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12425-12430 (2011).
65. Dierks C, *et al*. Expansion of Bcr-Abl-positive leukemic stem cells is dependent on hedgehog pathway activation. *Cancer Cell* **14**, 238-249 (2008).
66. Domingo-Domenech J, *et al*. Suppression of acquired docetaxel resistance in prostate cancer through depletion of Notch- and Hedgehog-dependent tumor-initiating cells. *Cancer Cell* **22**, 373-388 (2012).
67. Feldmann G, *et al*. An orally bioavailable small-molecule inhibitor of Hedgehog signaling inhibits tumor initiation and metastasis in pancreatic cancer (2008).
68. Sherry MM, Reeves A, Wu JLK, Cochran BH. STAT3 is required for proliferation and maintenance of multipotency in glioblastoma stem cells. *Stem Cells* **27**, 2383-2392 (2009).
69. Lin L, *et al*. targeting colon cancer stem cells using a new curcumin analogue, GO-Y030. *Br. J. Cancer* **105**, 212-220 (2011).
70. Wang H, *et al*. Targeting interleukin 6 signaling suppresses glioma stem cell survival and tumor growth. *Stem Cells* **27**, 2393-2404 (2009).
71. Marotta LLC, *et al*. The JAK2/STAT3 signaling pathway is required for growth of CD44(+)CD24(–) stem cell-like breast cancer cells in human tumors. *J. Clin. Invest.* **121**, 2723-2735 (2011).
72. Dubrovska A, *et al*. The role of PTEN/Akt/PI3K signaling in the maintenance and viability of prostate cancer stem-like cell populations. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 268-273 (2009).
73. Zhou J, *et al*. Activation of the PTEN/mTOR/STAT3 pathway in breast cancer stem-like cells is required for viability and maintenance. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 16158-16163 (2007).
74. Sunayama J, *et al*. Crosstalk between the PI3K/mTOR and MEK/ERK pathways involved in the maintenance of self-renewal and tumorigenicity of glioblastoma stem-like cells. *Stem Cells* **28**, 1930-1939 (2010).
75. Sunayama J, *et al*. FoxO3a functions as a key integrator of cellular signals that control glioblastoma stem-like cell differentiation and tumorigenicity. *Stem Cells* **29**, 1327-1337 (2011).
76. Latifi A, *et al*. Cisplatin treatment of primary and metastatic epithelial ovarian carcinomas generates residual cells with mesenchymal stem cell-like profile. *J. Cell. Biochem.* **112**, 2850-2864 (2011).
77. Sato A, *et al*. MEK-ERK signaling dictates DNA-repair gene MGMT expression and temozolomide resistance of stem-like glioblastoma cells via the MDM2-p53 axis. *Stem Cells* **29**, 1942-1951 (2011).
78. Tabu K, *et al*. Analysis of an alternative human CD133 promoter revelas the implication of Ras/ERK pathway in tumor stem-like hallmarks. *Mol. Cancer* **9**, 39 (2010).

79. Pellicano F, *et al*. The MEK inhibitor PD184352 enhances BMS-214662-induced apoptosis in CD34+ CML stem/progenitor cells. *Leukemia* **25**, 1159-1167 (2011).
80. Aceto N, *et al*. Tyrosine phosphatase SHP2 promotes breast cancer progression and maintains tumor-initiating cells via activation of key transcription factors and a positive feedback signaling loop. *Nat. Med.* **18**, 529-537 (2012).
81. Orzan F, *et al*. Enhancer of Zeste 2 (EZH2) is upregulated in malignant gliomas and in glioma stem-like cells. *Neuropathol. Appl. Neurobiol.* **37**, 381-394 (2011).
82. Nalls D, *et al*. Targeting epigenetic regulation of miR-34a for treatment of pancreatic cancer by inhibition of pancreatic cancer stem cells. *PLoS One* **6**, e24099 (2011).
83. Guzman ML, *et al*. The sesquiterpene lactone parthenolide induces apoptosis of human acute myelogenous leukemia stem and progenitor cells. *Blood* **105**, 4163-4169 (2005).
84. Guzman ML, *et al*. An orally bioavailable parthenolide analog selectively eradicates acute myelogenous leukemia stem and progenitor cells. *Blood* **110**, 4427-4435 (2007).
85. Bartucci M, *et al*. Therapeutic targeting of Chk1 in NSCLC stem cells during chemotherapy. *Cell Death Diff.* **19**, 768-778 (2012).
86. Gallmeier E, *et al*. Inhibition of ataxia telangiectasia- and Rad3-related function abrogates the in vitro and in vivo tumorigenicity of human colon cancer cells through depletion of the CD133+ tumor-initiating cell fraction. *Stem Cells* **29**, 418-429 (2011).
87. Bao S, *et al*. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756-760 (2006).
88. Chen Y, *et al*. Inhibitory effects of omacetaxine on leukemic stem cells and BCR-ALB-induced chronic myeloid leukemia and acute lymphoblastic leukemia in mice. *Leukemia* **23**, 1446-1454 (2009).
89. Copland M, *et al*. BMS-214662 potently induced apoptosis of chronic myeloid leukemia stem and progenitor cells and synergizes with tyrosine kinase inhibitors. *Blood* **111**, 2843-2853 (2008).
90. Pellicano F, *et al*. The MEK inhibitor PD184352 enhances BMS-214662-induced apoptosis in CD34+ CML stem/progenitor cells. *Leukemia* **25**, 1159-1167 (2011).
91. Pestereva E, Kanakasabai S, Bright JJ. PPARg agonists regulate the expression of stemness and differentiation genes in brain tumour stem cells. *Brit. J. Cancer* **106**, 1702-1712 (2012).
92. Castelo-Branco P, *et al*. Neural tumor-initiating cells have distinct telomere maintenance and can be safely targeted for telomerase inhibition. *Clin. Cancer Res.* **17**, 111-121 (2011).
93. Li Y, *et al*. Sulforaphane, a dietary component of broccoli/broccoli sprouts, inhibits breast cancer stem cells. *Clin. Cancer Res.* **16**, 2580-2590 (2010).

94. Yu Y, *et al*. Elimination of colon cancer stem-like cells by the combination of curcumin and FOLFOX. *Transl. Oncol.* **2**, 321-328 (2009).
95. Fong D, *et al*. Curcumin inhibits the side population (SP) phenotype of the rat C6 glioma cell line: towards targeting of cancer stem cells with phytochemicals. *Cancer Lett.* **293**, 65-72 (2010).
96. Ginestier C, *et al*. CXCR1 blockade selectively targets human breast cancer stem cells in vitro and in xenografts. *J. Clin. Invest.* **120**, 485-497 (2010).
97. Chen G, *et al*. Metformin inhibits growth of thyroid carcinoma cells, suppresses self-renewal of derived cancer stem cells, and potentiates the effect of therapeutic agents. *J. Clin. Endocrinol. Metab.* **97**, e510-e520 (2012).
98. Jung JW, *et al*. Metformin represses self-renewal of human breast carcinoma stem cells via inhibition of estrogen receptor-mediated OCT4 expression. *PLoS One* **6**, e28068 (2011).
99. Allan EK, Holyoake TL, Craig AR, Jorgensen HG. Omacetazine may have a role in chronic myeloid leukaemia eradication through downregulation of Mcl-1 and induction of apoptosis in stem/progenitor cells. *Leukemia* **25**, 985-994 (2011).

Chapter 2

Investigating the essential activators and coactivators for *NANOG* expression in cancer

2.1. Abstract

Eukaryotic gene transcription relies upon the concerted actions and interactions of activator and coactivator proteins. The multiprotein complexes that regulate transcription can be gene-specific and therefore represent an opportunity for selective modulation of target gene expression. Unfortunately, the regulatory proteins for many genes are unknown. Here we describe the identification of the coactivator p300 as direct and essential regulator of the embryonic transcription factor *NANOG* in human cancer cells. Aberrant *NANOG* expression promotes tumorigenesis and confers self-renewal and drug-resistance, yet no methods are currently available to inhibit *NANOG*. Our results indicate that loss of p300 function attenuates both *NANOG* expression and *NANOG*-dependent phenotypes, suggesting that p300 may be a viable target for inhibiting *NANOG*. We further show that the Kruppel-like factor (KLF) family of transcriptional activators, in particular KLF4, is required for *NANOG* expression, revealing yet another potential target for *NANOG* inhibition. These findings help generate a clearer picture of the activator-coactivator complexes regulating *NANOG* and can be used to design new ways of modulating *NANOG* expression in cancer.

2.2. Transcriptional coactivators and gene expression

Eukaryotic gene activation is regulated through the sequential interactions of a large number of modulatory multiprotein complexes involving both transcriptional activators and coregulatory proteins termed coactivators (1). Transcriptional activators bind to specific cognate sequences in the promoter and enhancer regions of genes and, once localized to the DNA, stimulate transcription by bringing the basal transcriptional machinery to the transcriptional initiation site (2). While some activators are able to directly recruit general factors such as TFIID and RNA polymerase II, most activators rely on coactivators to facilitate gene activation. The resulting multiprotein complexes perform such diverse functions as remodeling the chromatin, stabilizing the transcriptional preinitiation complex, and even conferring temporal- and cell-type selectivity in which genes are expressed in the cell (3-5). As a result, it is imperative to understand the series of protein-protein interactions (PPI) and catalytic functions between transcriptional components in order to fully understand how complex biological programs are regulated.

The need for a detailed understanding of gene regulatory mechanisms becomes especially important when trying to modulate gene expression. A number of human diseases manifest as a result of aberrant gene activation, including cancer, developmental defects, and metabolic disorders (6-9). In the case of cancer, numerous groups have attempted to modulate transcription in order to

treat the disease through the use of coactivator enzymatic inhibitors as well as PPI inhibitors (10-13). However, the design and implementation of these strategies requires knowledge about both identity of the activators/coactivators involved in expression of a target gene and the nature of the PPIs that underlie the regulatory complexes. Unfortunately, this level of information is often lacking for many disease-related genes. We have set out to identify key coactivators and PPIs essential for the expression of one such cancer-associated gene, *NANOG*. We found that the coactivator p300, in particular, is an important and direct regulator of *NANOG* expression in two different carcinoma cell lines, and provide preliminary evidence for how p300 is initially recruited to the *NANOG* promoter. This information provides the foundation for designing new inhibition strategies to downregulate *NANOG* and reverse aggressive, *NANOG*-dependent cancer phenotypes.

2.2.1. The role of NANOG *in cancer*

NANOG is an embryonic transcription factor that has recently been implicated as a driver of several human malignancies, including head and neck squamous cell carcinoma (HNSCC), prostate carcinoma, glioblastoma, and germ cell tumors (14-17). Elevated *NANOG* levels are associated with poor survival in HNSCC patients (17), resistance to standard chemotherapeutics (18), and interestingly, increased features of cancer-initiating cells (CICs) (19, 20). These CICs exhibit complex phenotypes such as self-renewal, sustained multipotency, tumorigenicity, and a gene expression profile similar to that found in embryonic

stem cells (ESCs) (14, 15, 21). Moreover, downregulation of *NANOG* in cancer cells markedly decreases these malignant features, consistent with the idea that *NANOG* is a core factor in a complex CIC transcriptional program. We therefore hypothesized that loss of function of key coactivators that regulate *NANOG* would disrupt both *NANOG* expression and *NANOG*-dependent phenotypes. To test this hypothesis though, we must first examine the known coregulators of *NANOG* in order to narrow down our search for potential targets.

2.2.2. Coregulators of NANOG expression

Interactome studies of NANOG have revealed an extended network of protein-protein interactions that participate in both regulating the pluripotency gene network, as well as *NANOG* itself (22, 23). For example, the Baf (Brg/Brahma-associated factor) ATP-dependent chromatin remodeling complexes are essential for formation of the pluripotent cells of the early embryo. ESCs deficient in the core Brg enzyme maintain the expression of Oct4, Sox2, and Nanog for several divisions but ultimately lose self-renewal capacity and downregulate pluripotency marker expression suggesting that Brg is required to maintain stable expression of Oct4 and Nanog over time (4, 5). Baf complexes in ESCs (esBaf) include Brg, Baf155 and Baf250A. Brg and Baf155 have been reported to bind extensively to ESC regulator genes, including *Oct4*, *Sox2*, *Nanog*, and *Myc*, whereas Brg localizes to *Klf4*. The essential nature of esBaf complexes and their localization within the genome suggest that esBaf are critical

for ESC self-renewal and pluripotency by directly regulating the core transcriptional circuitry (4, 5).

Several other coactivators have been reported to regulate the pluripotency network in ESCs, including Mediator (24), the histone acetyltransferase p300 (25), and surprisingly a nucleotide base-excision repair (NER) complex (26). Using defined *in vitro* transcription assay, Fong et al. (26) screened for cofactors required for OCT4/SOX2-mediated activation from the *Nanog* promoter. Unexpectedly, the XPC-RAD23B-CETN2 NER complex was found to be a necessary coactivator for both *in vitro* and *in vivo* transcription of *Nanog*. Mouse ESCs depleted of XPC showed decreased *Nanog* expression, reduced self-renewal capacity, and increased differentiation, indicating that the NER complex is essential for maintaining *Nanog* expression driven by Oct4/Sox2.

Interestingly, the activators OCT4 and NANOG have also been shown to interact with several corepressor proteins including members of the histone deacetylase NuRD complex (P66β, Sall4 and HDAC2) and polycomb proteins (YY1, Rnf2 and Rybp) (22, 27). These corepressor complexes are generally though to aid in suppressing lineage-specific genes and thereby maintain OCT4 and NANOG expression and self-renewal, however some positive regulatory roles have been reported. For example, the mSin3A-HDAC complex directly stimulates *Nanog* expression in mouse ESCs (27). Knockdown of mSin3A-HDAC components or HDAC inhibitor treatment reduces *Nanog* expression, whereas overexpression of

mSin3A-HDAC subunits increases *Nanog* expression. Interestingly, mSin3A-HDAC occupancy at the *Nanog* promoter does not result in histone deacetylation, indicating that HDAC activity is required for *Nanog* expression but is not histone-mediated.

There appears to be a strong correlation between histone acetylation and the core pluripotency transcriptional network. In particular, the histone acetyltransferase p300 plays a prominent role in regulating pluripotency. Mouse ESCs lacking p300 display abnormal gene marker expression and aberrant differentiation, as well as reduced expression of *Nanog* (28). Forced expression of Nanog rescues these defects, indicating that p300 may be an upstream regulator of *Nanog* expression. In support of this, global-binding studies have found that p300 colocalizes with the core transcription factors Oct4, Sox2, and Nanog at active pluripotency-associated genes (25). These observations led us to test if p300 has an important role in *NANOG* regulation in cancer and cancer-initiating cells.

2.3. The coactivator p300 is essential for *NANOG* expression and function

2.3.1. Loss of p300 attenuates NANOG-dependent phenotypes in HNSCC

The ESC transcriptional program is regulated by a complex network of activators and cofactors yet thus far only a few coactivators have been suggested to

contribute (21, 25). Among these coactivators, the histone acetyltransferase p300 has been reported to regulate self-renewal and cellular identity (25, 28, 29). Given these observations, we hypothesized that p300 may also be an essential coactivator for *NANOG* expression in cancer cells.

NANOG is commonly expressed in head and neck squamous cell carcinomas (HNSCC) and is associated with aggressive phenotypes such as invasiveness and CIC populations (17, 18, 30). To begin testing whether p300 is a necessary coactivator for oncogenic *NANOG* expression, the Pan laboratory (Ohio State University) generated stable p300-deficient clones in the HNSCC model cell line UMSCC74A. UMSCC74A-scramble control cells expressed high levels of NANOG (Fig. 2.1A), consistent with other HNSCC lines (17). However, knockdown of p300 resulted in a strong reduction in NANOG levels (Fig. 2.1A), suggesting p300 is important in sustaining high NANOG levels.

We next examined what impact the loss of p300 would have on *NANOG*-dependent phenotypes. When expressed in human cancers, *NANOG* confers stem cell-like features including greater metastatic potential and self-renewing capacity (20). *In vitro*, these properties can be assessed using wound healing and tumorsphere formation assays, respectively. When the cells lose *NANOG* or stem like-character, they are expected to exhibit reduced wound healing and motility, as well as a reduced ability to self-renew and generate tumorspheres when cultured under nonadherent condtions. When we examined p300-deficient

UMSCC74A cells, we found that the knockdown cells showed greatly reduced motility in wound healing assays, closing ~20% of the wound in the time control cells take

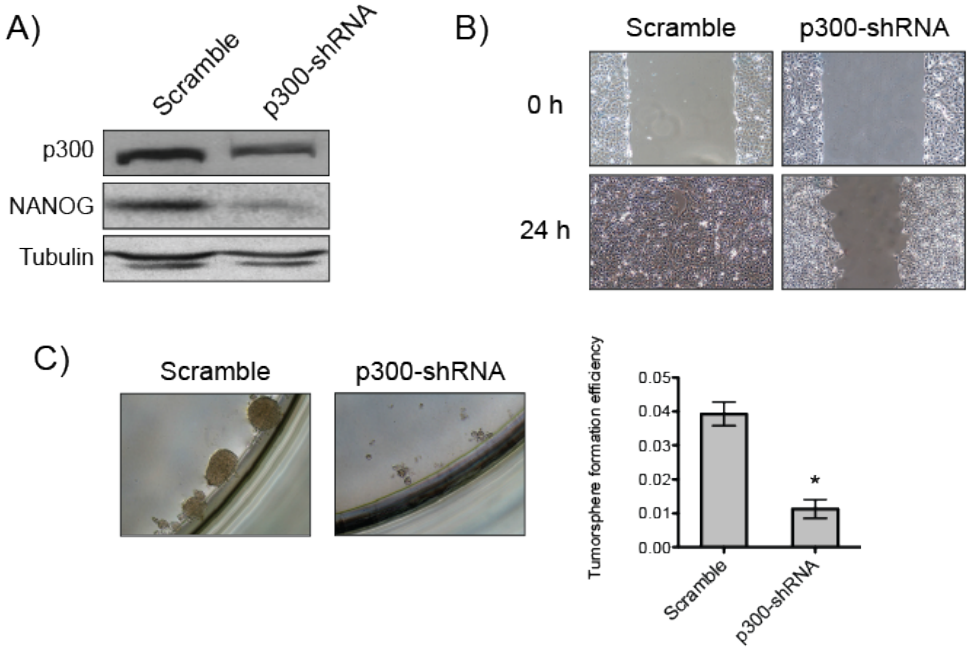

Fig. 2.1. p300 knockdown inhibits NANOG and NANOG-associated phenotypes in HNSCC cells. Stable polyclonal UMSCC74A-scramble and -p300-shRNA cells were generated by transfection and antibiotic selection. (A) Immunoblot showing p300, NANOG, and a-tubulin levels. (B) Wound healing assay. (C) Tumorsphere formation assay. Cells were harvested and seeded on low-attachment plates in a defined, serum-free culture medium for 7 days. Tumorsphere formation efficiency was calculated as the number of tumorspheres (≥50 μm in diameter) formed divided by the original number of cells seeded. Data is presented as mean ± SEM. *p<0.01 (n = 6). Experiments were performed by X. Xie, M. Zhang, and Q. Pan of Ohio State University.

to fully seal reform a monolayer (Fig. 2.1B). p300 knockdown also dramatically inhibited the ability of UMSCC74A cells to form tumorspheres in suspension culture (Fig. 4.2C), suggesting that p300 loss reduced the ability of UMSCC74A cells to self-renew. These results indicated that p300 is important for *NANOG* expression and *NANOG*-driven cancer phenotypes in HNSCC cells.

2.3.2. Loss of p300 downregulates NANOG expression in pluripotent NCCIT cells

We then sought to explore whether p300 is necessary for *NANOG* expression in other cancers besides HNSCC using similar loss-of-function knockdown studies. The teratocarcinoma cell line NCCIT is pluripotent and endogenously expresses

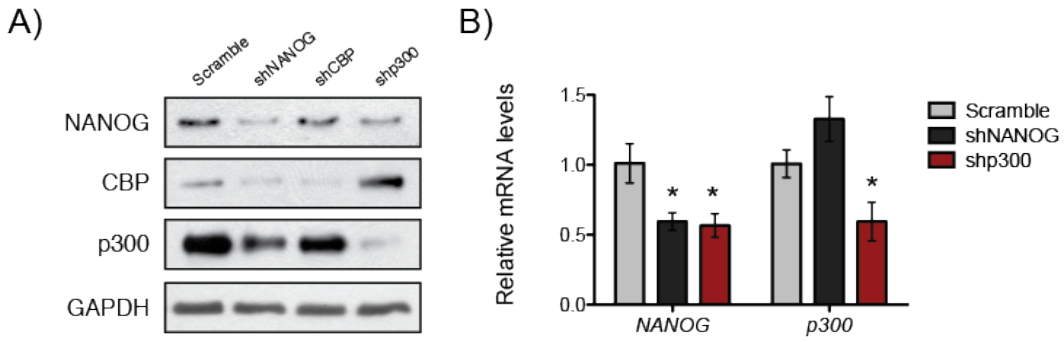

Fig. 2.2. p300 knockdown inhibits NANOG expression in pluripotent NCCIT cells. Cells were transiently transfected with shRNAs and selected with antibiotics for 48 h before harvesting. (A) Immunoblot of NANOG, CBP, p300, and GAPDH levels. (B) mRNA levels of *NANOG* and *p300* 48 h post-transfection. *$p < 0.05$ vs. scramble control (n = 3)(*t*-test).

NANOG and other components of the ESC transcriptional network (31, 32). Transient transfection of NCCIT cells with shRNAs against *NANOG*, *p300*, or the p300-paralog *CBP* all downregulated their targets (Fig. 2.2A). Interestingly, p300-knockdown also inhibited NANOG expression, consistent with the results seen in HNSCC cells. In contrast, loss of the coactivator CBP had no significant impact on NANOG levels, even though CBP and p300 are similar and thought to be largely functionally redundant (9). In support of this, p300-knockdown caused an increase in CBP levels, suggesting that CBP was upregulated to compensate for loss of a redundant protein. However, even elevated CBP levels are

insufficient to maintain NANOG expression, suggesting that CBP and p300 may have differential roles in regulating *NANOG*.

Taken together these results indicate that p300 function is essential for *NANOG* expression in both HNSCC and pluripotent cancer cells, two physiologically and etiologically distinct tumor types. These observations therefore suggest that p300 may be a conserved cofactor for *NANOG* expression and thus a potential target for targeting across the spectrum of *NANOG*-dependent tumors.

2.4. Determination of the genomic origin of *NANOG* in cancer cells

Our results in the loss of p300 function studies suggest that p300 may be an important regulator of *NANOG* expression in cancer cells. However, those experiments did not assess whether p300 functions as a direct or indirect regulator of the *NANOG* gene. p300 could very well be required for expression of an upstream regulator of *NANOG*, and thus knockdown or inhibition of p300 would affect *NANOG* expression through an intermediary gene product. Such a scenario would require identifying the intermediate(s) and likely shift the focus of the experiments away from *NANOG*.

Further complicating the system is the fact that *NANOG* transcripts in several cancer cell lines have been reported to originate from the *NANOGP8* pseudogene located on chromosome 15 (15, 33-36) instead of the *NANOG1* locus associated with embryonic expression.

We reasoned that, if p300 acted as a direct regulator of *NANOG*, it would be localized to the *NANOG* proximal promoter consistent with other direct, p300-dependent genes (37). We therefore set out to determine whether p300 is bound to the *NANOG1* and/or *NANOGP8* promoters in human cancer cell lines.

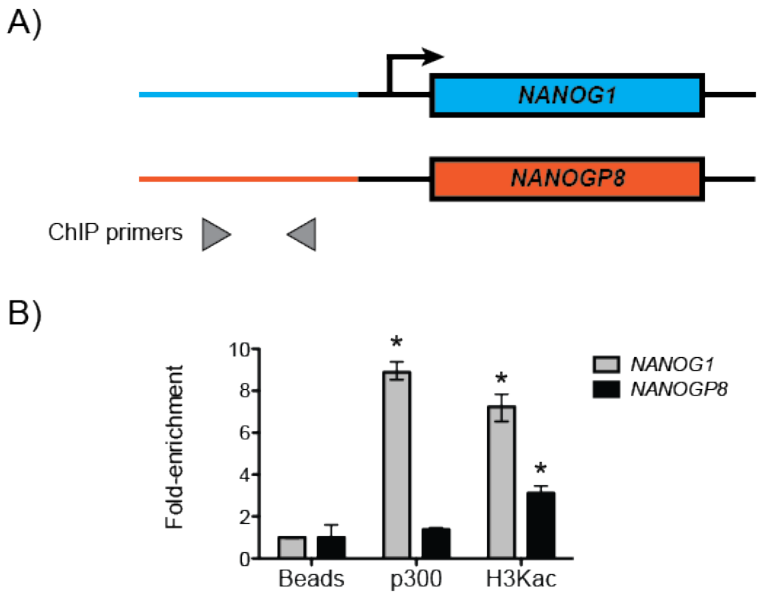

Fig. 2.3. p300 localizes to the *NANOG1* promoter in NCCIT cells. (A) Schematic of the *NANOG1* (blue) and *NANOGP8* (orange) loci. Conserved sequences are indicated in black. ChIP primers were designed to amplify unique but analogous regions of the two proximal promoters. (B) ChIP assay results. Fold-enrichment was calculated by the ΔCt method relative to beads. Data is representative of 3 independent experiments. Error bars indicate the upper and lower limits within measurements. *$p < 0.05$ vs. beads.

Chromatin immunoprecipitation (ChIP) assays were conducted first in the pluripotent embryonal carcinoma NCCIT cells. In this experiment, NCCIT cells were fixed with formaldehyde and chromatin-bound proteins like histone H3 and p300 were immunoprecipitated to pull-down associated sequences of DNA that can be amplified using sequence-specific PCR primers. Given the two potential

sources of *NANOG* mRNA in human cancers, two gene-specific primer sets were designed: one for the *NANOG1* proximal promoter, and one for the analogous 5' region of *NANOGP8* (Fig. 2.3A).

Immunoprecipitation for p300 revealed ~9-fold enrichment for DNA from the *NANOG1* promoter over beads alone (Fig. 2.3B), indicating that the p300 coactivator is directly associating with the *NANOG1* locus and likely engaged with transcriptional machinery driving expression from that gene. In contrast, p300 failed to show significant enrichment over background at the *NANOGP8* promoter. Histone H3 was localized consistently across both loci, showing that chromatin fragments from *NANOGP8* can be recovered and therefore the absence of p300 at *NANOGP8* is not due to differences in PCR primers or amplification efficiency. We also examined the presence of p300-associated histone acetylation, notably histone H3 acetylation (H3Kac), at both loci. H3Kac was observed to be significantly enriched at the *NANOG1* locus relative to the bead control, consistent with acetylation patterns at actively transcribed genes. H3Kac was also observed at the *NANOGP8* locus but at ~50% of the *NANOG1* levels. H3Kac is a common histone modification and its presence at *NANOGP8* may be the result of global acetylation by HAT enzymes other than p300. Thus, even if *NANOGP8* contributes to overall *NANOG* expression in NCCIT cells, p300 does not appear to directly regulate *NANOGP8*. Moreover, loss of p300 function has a significant inhibitory effect *NANOG* expression levels, indicating

that *NANOGP8* has only a minor, if any, contribution to overall *NANOG* expression in NCCIT cells.

Taken together, these data indicate that the *NANOG1* locus is the primary source of *NANOG* mRNA and protein in NCCIT cells, and that the pseudogene *NANOGP8* is not actively transcribed. In support of this, the transcriptional coactivator shows significant localization to the *NANOG1* proximal promoter and no detectable association with the *NANOGP8* upstream region. Furthermore, the *NANOG1* promoter shows high levels of histone H3 acetylation (H3Kac) modification (Fig. 2.3B), a general marker for p300 HAT activity and actively transcribed genes. In contrast, the *NANOGP8* upstream region is relatively depleted of H3Kac modifications, consistent with *NANOGP8* being transcriptionally silent. p300 thus appears to be both a direct and essential regulator of the *NANOG1* promoter, and we elected to focus on how p300 may be recruited to this sequence of DNA.

2.5. *NANOG1* promoter analysis reveals a pair of key KLF motifs

2.5.1. Deletion screening

Significant effort has been invested in studying the transcriptional activators involved in driving *NANOG* expression in embryonic stem cells (see above), yet few activators and even fewer coactivators have been identified in cancer cells. Having identified p300 as a direct regulator of *NANOG*, we next sought to identify

potential activator proteins that could bind p300 and recruit it to the *NANOG* promoter.

The *NANOG1* proximal promoter is loosely defined as spanning from –380 bp to +24 bp from the transcriptional start site. This region contains consensus motifs for many activators and repressors involved in the complicated transcriptional regulation of *NANOG*. We hypothesized that p300 could be recruited to the *NANOG* promoter by binding to one or more of these proteins, however it was not clear what transcription factor(s) would be responsible in cancer cells. Running the promoter sequence through Genomatix MatInspector predicts 115 different transcription factor-binding sites, roughly 1 site every 4 nucleotides (). Several of the predicted motifs match known *NANOG* regulators in ESCs, such as OCT4/SOX2, KLF4, and NANOG itself (Table 2.1). Additionally, several known oncogenes were predicted to bind including GLI, c-Jun, ETS, and SMAD4 (38).

Table 2.1. Genomatix predictions in the *NANOG* proximal promoter

Predicted protein	Start	End	Strand	*P*	
Cell cycle regulators: Cell cycle homology element	15	9	(-)	0.956	gagtTTGAaacca
cAMP-responsive element binding proteins	31	21	(+)	0.997	aactccTGACttcaggtgatc
Human and murine ETS1 factors	32	22	(-)	0.847	ggatcaCCTGaagtcaggagt
Sterol regulatory element binding proteins	32	25	(-)	0.944	ggaTCACctgaagtc
Two-handed zinc finger homeodomain transcription factor	32	26	(-)	0.991	ggatcACCTgaag
Myoblast determining factors	43	35	(-)	0.883	ccgtGGCAggcggatca
HIF-1 ancillary sequence family	46	41	(+)	0.934	tgcCACGgcct
GLI zinc finger family	54	47	(+)	0.864	acggCCTCccaattt
SOX/SRY-sex/testis determinig and related HMG box fact	64	52	(-)	0.888	taatcccagtaAATTgggaggccgt
AT rich interactive domain factor	66	56	(+)	0.926	tcccaaTTTActgggattaca
Abdominal-B type homeodomain transcription factors	64	56	(-)	0.917	taatcccagTAAAttgg
Bicoid-like homeodomain transcription factors	69	61	(-)	0.933	ccctgtAATCccagtaa
Brn POU domain factors	74	65	(+)	0.913	ctgggATTAcaggggtggg
Insulinoma associated factors	74	68	(+)	0.909	ttacaGGGGtggg
Krueppel like transcription factors	78	70	(+)	0.955	ttacagGGGTgggccac
C2H2 zinc finger transcription factors 2	84	73	(-)	0.889	ggcgcggtggcCCACccctgtaa
GLI zinc finger family	77	70	(-)	0.892	tggcCCACccctgta
ZF5 POZ domain zinc finger	90	83	(+)	0.879	cacCGCGcccggcct
RNA polymerase II transcription factor II B	84	81	(+)	1	ccgCGCC
Homolog to deformed epidermal autoregulatory factor-1 fr	99	90	(+)	0.746	cgcCCGGccttttttcttaa
AT-binding transcription factor	104	96	(+)	0.795	ccttttttcttAATTttt
Ccaat/Enhancer Binding Protein	103	96	(-)	0.948	aaaattaaGAAAaag
Homeodomain transcription factors	108	99	(+)	0.866	tttttcTTAAtttttaaaa
snRNA-activating protein complex	108	99	(+)	0.733	tttttCTTAatttttaaaa
NK6 homeobox transcription factors	106	99	(+)	0.917	tttcTTAAttttaa
Paralog hox genes 1-8 from the four hox clusters A, B, C,	110	101	(+)	0.813	tttcttAATTtttaaaaat
NKX homeodomain factors	110	101	(+)	0.957	tttctTAATttttaaaaat
Hepatic Nuclear Factor 1	110	102	(+)	0.848	tCTTAatttttaaaaat
AT-binding transcription factor	111	103	(-)	0.844	tatttttaaaAATTaag
AT rich interactive domain factor	117	107	(-)	0.961	ctttaATATttttaaaaatta
Hepatic Nuclear Factor 1	117	109	(-)	0.806	cTTTAatatttttaaaa
Special AT-rich sequence binding protein	116	109	(-)	0.968	tttAATAtttttaaa
AT rich interactive domain factor	124	114	(+)	0.95	taaaaATATtaaagttttatc
NK6 homeobox transcription factors	119	112	(-)	0.852	aactTTAAtattttt
Vertebrate TATA binding protein factor	123	115	(+)	0.82	aaatattAAAGttttat
Homeodomain transcription factors	125	116	(-)	0.975	ggataaaacttTAATattt
MYT1 C2HC zinc finger protein	124	118	(+)	0.99	ttaAAGTtttatc
Vertebrate caudal related homeodomain protein	130	121	(+)	0.908	ttaaagtTTTAtcccattc
GATA binding factors	128	122	(-)	0.931	atggGATAaaact
C2H2 zinc finger transcription factors 5	131	124	(-)	0.883	ggaatgGGATaaaac
Paralog hox genes 1-8 from the four hox clusters A, B, C,	136	127	(-)	0.959	caacaggAATGggataaaa
Human and murine ETS1 factors	139	129	(-)	0.856	gttcaacAGGAatgggataaa
TEA/ATTS DNA binding domain factors	135	129	(+)	0.944	tccCATTcctgtt
Human and murine ETS1 factors	155	145	(-)	0.889	ttaaatcaGGAAtatggttca
Growth factor independence transcriptional repressor	155	148	(-)	0.864	ttaAATCaggaatat
Vertebrate TATA binding protein factor	159	151	(-)	0.855	actttTAAAtcaggaat
HOX - PBX complexes	160	152	(+)	0.947	ttcctGATTtaaaagtt
Brn POU domain factors	165	156	(+)	0.881	ctGATTtaaaagttggaaa
MYT1 C2HC zinc finger protein	165	159	(+)	0.887	taaAAGTtggaaa
Hypoxia inducible factor, bHLH/PAS protein family	174	166	(+)	0.897	gttggaaaCGTGgtgaa
Nuclear factor of activated T-cells	176	167	(+)	0.875	gttGGAAacgtggtgaacc
NKX homeodomain factors	192	183	(+)	0.844	acctagAAGTatttgttgc
SWI/SNF related nucleophosphoproteins with a RING fing	187	182	(-)	0.987	aaatACTTcta
X-box binding factors	203	194	(-)	0.876	agacaaacccagCAACaaa
Interferon regulatory factors	209	199	(-)	0.694	acctGAAGacaaacccagcaa
Glucocorticoid responsive and related elements	214	205	(+)	0.896	gtttgtcttcagGTTCtgt
MEF3 binding sites	213	207	(+)	0.922	tctTCAGgttctg
X-box binding factors	228	219	(-)	0.867	tagaaaaccgagCAACaga
PAX-2/5/8 binding sites	239	225	(+)	0.805	ctgttGCTCggttttctagttccccacct
Interferon regulatory factors	236	226	(-)	0.866	tgggGAACtagaaaaccgagc
Human and murine ETS1 factors	241	231	(-)	0.978	ctaggtggGGAActagaaaac
C2H2 zinc finger transcription factors 2	243	232	(+)	1	gttttctagttCCCCacctagtc
Myeloid zinc finger 1 factors	237	232	(-)	1	gtGGGGaacta
Serum response element binding factor	250	241	(-)	0.87	taaccCAGActaggtgggg
X-box binding factors	254	245	(+)	0.89	acctagtctggGTTActct
Vertebrate SMAD family of transcription factors	248	243	(+)	0.997	ctaGTCTgggt

Table 2.1. Genomatix predictions in the *NANOG* proximal promoter: continued

Predicted protein	Start	End	Strand	*P*	
X-box binding factors	262	253	(-)	0.912	gtagctgcagaGTAAccca
Octamer binding protein	273	266	(-)	0.958	gtaATGCaaaagtag
Motif composed of binding sites for pluripotency or stem c	278	269	(+)	0.998	tacttttGCATtacaatgg
PAR/bZIP family	277	269	(-)	0.916	cattgTAATgcaaaagt
Octamer binding protein	276	269	(+)	0.816	cttttGCATtacaat
RXR heterodimer binding sites	286	274	(-)	0.794	caccaaGGCCattgtaatgcaaaag
Activator/repressor binding to transcription initiation site	283	273	(-)	0.882	caaggCCATtgtaatgcaaaa
Brn POU domain factors	282	273	(+)	0.944	tttgcATTAcaatggcctt
Estrogen response elements	286	275	(-)	0.884	caccAAGGccattgtaatgcaaa
SOX/SRY-sex/testis determinig and related HMG box fact	291	279	(+)	0.946	gcattACAAtggccttggtgagact
Vertebrate steroidogenic factor	287	280	(-)	0.993	tcacCAAGgccattg
SOX/SRY-sex/testis determinig and related HMG box fact	315	303	(-)	0.883	aattctcagttAATCccgtctacca
Bicoid-like homeodomain transcription factors	310	302	(-)	0.985	tcagtTAATcccgtcta
Hepatic Nuclear Factor 1	315	307	(-)	0.908	aattctcaGTTAatccc
Cellular and viral myb-like transcriptional regulators	315	308	(+)	0.898	gattAACTgagaatt
Core promoter initiator elements	312	307	(-)	0.944	tcTCAGttaat
FAST-1 SMAD interacting proteins	324	316	(-)	0.871	accctTGTGaattctca
CP2-erythrocyte Factor related to drosophila Elf1	333	324	(-)	0.854	tACTGacccacccttgtga
v-ERB and RAR-related orphan receptor alpha	339	328	(+)	0.757	acaagggtggGTCAgtaggggt
MAF and AP1 related factors	338	328	(+)	0.918	caagggTGGGtcagtaggggg
Testis-specific bHLH-Zip transcription factors	331	326	(+)	0.944	gGGTGggtcag
C2H2 zinc finger transcription factors 2	348	337	(-)	0.92	ggcgggcacaCCCCctactgacc
Chorion-specific transcription factors with a GCM DNA bin	342	335	(-)	0.863	cacacCCCCtactga
GLI zinc finger family	342	335	(-)	0.864	cacaCCCCctactga
Krueppel like transcription factors	344	336	(+)	0.967	tcagtaGGGGgtgtgcc
C2H2 zinc finger transcription factors 2	351	340	(-)	0.934	cctggcgggcacaCCCCctactg
Krueppel like transcription factors	346	338	(+)	0.97	agtaggGGGTgtgcccg
GC-Box factors SP1/GC	347	339	(+)	0.934	gtagGGGGtgtgcccgc
Pleomorphic adenoma gene	354	343	(+)	1	taGGGGgtgtgcccgccaggagg
Vertebrate homologues of enhancer of split complex	347	340	(+)	0.929	agggggtGTGCccgc
Retinoblastoma-binding proteins with demethylase activity	343	339	(-)	0.965	GCACacccc
CTCF and BORIS gene family, transcriptional regulators w	365	352	(+)	0.824	tgtgcccgccaggaGGGGtgggtctaa
Selenocysteine tRNA activating factor	364	353	(-)	0.795	tagaCCCAcccctcctggcgggc
Krueppel like transcription factors	360	352	(+)	0.901	ccgccaggagGGGTggg
Krueppel like transcription factors	364	356	(+)	0.954	caggagGGGTgggtcta
C2H2 zinc finger transcription factors 2	370	359	(-)	0.912	tcaccttagacCCACccctcctg
GC-Box factors SP1/GC	365	357	(+)	0.898	aggaggGGTGggtctaa
RXR heterodimer binding sites	374	362	(+)	0.846	ggaggggtgggtctaAGGTgataga
Peroxisome proliferator-activated receptor	373	362	(+)	0.795	gaggggtgggtctaAGGTgatag
Nuclear receptor subfamily 2 factors	376	364	(+)	0.813	aggggtgggtctaAGGTgatagagc
Brn-5 POU domain factors	385	374	(-)	0.818	aTAATgaaggctctatcacctta
C2H2 zinc finger transcription factors 4	378	372	(-)	0.98	agGCTCtatcacc
Paralog hox genes 1-8 from the four hox clusters A, B, C,	389	380	(-)	0.822	atttaTAATgaaggctcta
Brn POU domain factors	393	384	(-)	0.906	ctagaTTTAtaatgaaggc
Brn POU domain factors	396	387	(+)	0.919	ttCATTataaatctagaga
Plant TATA binding protein factor	393	386	(+)	0.909	tcatTATAaatctag
HOX - PBX complexes	395	387	(-)	0.949	ctctagaTTTAtaatga
Vertebrate TATA binding protein factor	396	388	(+)	0.923	cattaTAAAtctagaga
C2H2 zinc finger transcription factors 10	401	394	(+)	0.852	aatCTAGagactcca

Name	Region Deleted	Primer Sequence (5'-3')
del1	–404 to –300 bp	CTCGCTAGCCTCGAG**CGCCCGGCCTTTTTC**
del2	–311 to –205 bp	GATTACAGGGGTGGG**CTAGAAGTATTTGTTG**
del3	–208 to –128 bp	GTTGGAAACGTGGTG**CTGCAGCTACTTTTG**
del4	–121 to –52 bp	GGGTTACTCTGCAGC**CAGTAGGGGGTGTGC**
del5	–49 to –5 bp	CACAAGGGTGGGTCAG**GATCTGGCCTCGGCG**

Table 2.2: Summary of the deletions tested in the NANOG proximal promoter. Mutagenesis primer sequences are listed. Reverse complement primers were used as antisense primers for the reactions. The 5' ORF (upstream from the deletion) are in black. The 3' ORF (immediately downstream from the deletion) are in blue.

To rapidly identify key regulators, we initially constructed a series of deletions (del1-5) (Table 2.2, Fig. 2.4A) within the promoter sequence to scan for regions essential for *NANOG* promoter activity in a luciferase reporter assay. The transcriptional start site was left intact in all constructs so that transcription was not artificially disrupted by the deletions. In pluripotent NCCIT cells, loss of distal promoter elements had little to no effect on *NANOG* promoter activity. In contrast, the del4 construct that deleted the OCT4/SOX2 consensus sequences displays an ~2-fold reduction in transcriptional activity, consistent with previous reports in ESCs indicating this motif is essential for high levels of promoter activation in pluripotent cells. Surprisingly, the del5 construct exhibited ~10-fold lower activity relative to WT, suggesting that there are regulatory elements in the del5 region with an even greater regulatory effect on *NANOG* expression than OCT4/SOX2.

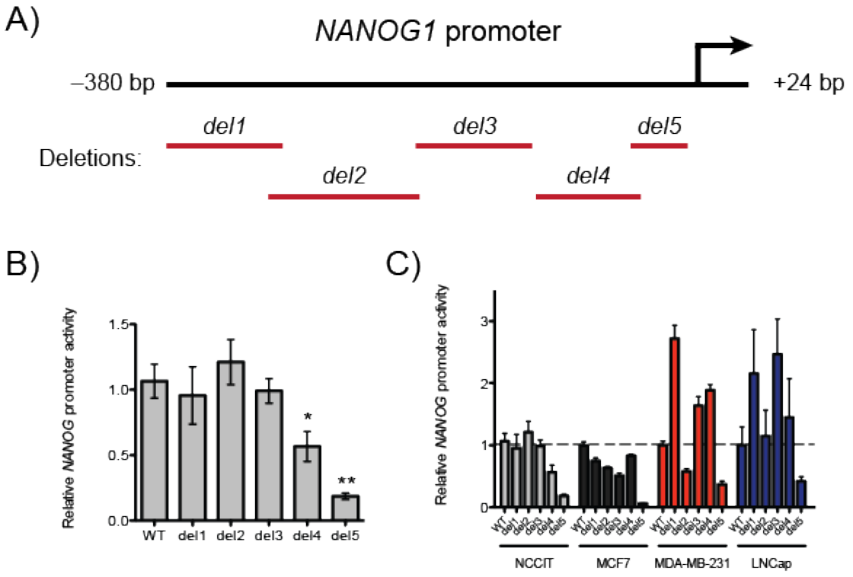

Fig. 2.4. Promoter deletions in NCCIT cells. (A) Schematic of the NANOG1 proximal promoter (black line) and the relative regions that were deleted (red line). (B) Relative promoter construct activity. NCCIT cells were transfected with reporter plasmids driven by various NANOG promoters. After 48 h, cells assayed for luciferase activity according to the manufacturer's protocol. Data is presented as the mean ± SEM relative to the WT control (*$p < 0.05$, t-test). (C) Screen of NANOG promoter deletions in a panel of cancer cell lines. Data is presented as the mean ± SEM relative to the WT control.

Based on these observations in pluripotent NCCIT cells, we then screened the deletion constructs in a small panel of NANOG+ cancer cell lines (Fig. 2.4C). Intriguingly, different cell lines and cancer types display varying promoter element-requirements. For example, NCCIT cells required the del4 element for activity while more differentiated MCF7, MDA-MB-231, and LNCap cells that do not express high levels of OCT4 are not affected by del4 (Fig. 2.4C). Furthermore, del1 had little to no inhibitor effect on promoter activity in NCCIT and *HER2*– MCF7 breast cancer cells, yet greatly increased activity in MDA-MB-231 breast and LNCap prostate cancer cells, suggesting that this region may be bound by a differentially expressed repressor of *NANOG* (Fig. 2.4C). Identification of these repressors could provide valuable details about how

NANOG is regulated in different cancer cell types. However, the present research is focused on activator identification, so we chose to pursue regions required for full activity.

2.5.2. Two GC-rich KLF motifs are broadly required for NANOG expression

In contrast to differentially required regions, the del5 construct displayed greatly reduced transcriptional activity in all cell types tested. The del5 region does not contain traditional TATA or CAAT box motifs associated with general transcriptional machinery, suggesting that the decrease in promoter activity may not be due to nonspecific transcriptional inhibition. MatInspector predicted 2 GC-rich motifs to be bound by Kruppel-like factor (KLFs) (Fig. 2.5A). To test if these motifs are essential for promoter activity, dinucleotide substitutions were introduced to disrupt KLF recognition (39) (Fig. 2.5A). The OCT4 binding site was mutated as well to serve as an additional control (40). In NCCIT cells, mutation of the 5' motif (mutKLF') or the 3' motif (mutKLF'') caused a 50% and 90% reduction in promoter activity, respectively. In contrast, mutation of the OCT4 site only perturbed promoter activity in NCCIT cells and had no impact on activity in more differentiated MCF7 cells. These results indicate that the KLF motifs are essential for *NANOG* promoter activity independent of cellular differentiation status and therefore KLF activators may be a common and pivotal regulator of *NANOG* expression in human cancer cells.

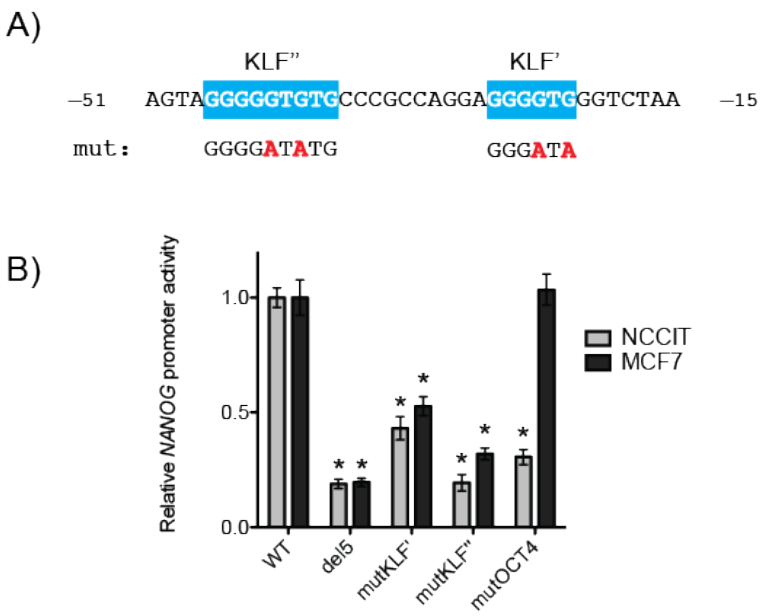

Fig. 2.5. Disruption of the KLF binding motifs inhibits *NANOG* promoter activity. (A) DNA sequence of the del5 region. Two KLF binding sites were predicted by Genomatix (blue). Motif mutations are shown below (red). (B) *NANOG* promoter activity in NCCIT and MCF7 cells. Data is presented as the mean ± SEM relative to the WT control (*p < 0.05, t-test).

2.5.3. Loss of KLF4 downregulates NANOG expression in cancer cells

KLF4 in particular has been documented to be a strong activator of the *NANOG* promoter in ESCs (39). We therefore chose to focus examine whether KLF4 was required for *NANOG* expression in cancer cells. Short-hairpin RNAs were designed to target *KLF4*, *NANOG*, and a scrambled sequence as a negative control. Transfecting NCCIT cells with these shRNAs indicated that robust downregulation could be achieved with the targeted constructs (Fig. 2.6A). Both shKLF4 and shNANOG reduced mRNA levels of their target genes by ~90%, while *OCT4* mRNA levels were not reduced (Fig. 2.6A). Strikingly, *KLF4*-knockdown strongly reduced *NANOG* mRNA levels, suggesting that KLF4 may be important for full *NANOG* expression (Fig. 2.6A).

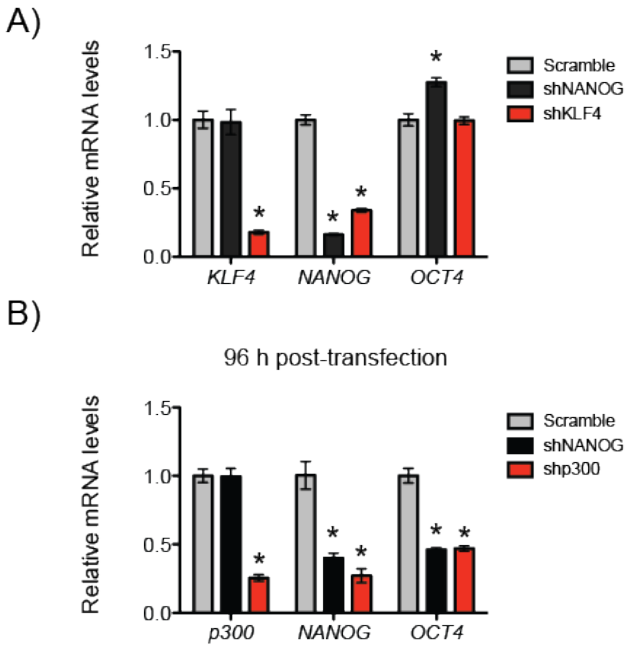

Fig. 2.6. Effects of *KLF4* knockdown on *NANOG* and *OCT4* expression in NCCIT cells. (A) Relative mRNA levels of *KLF4*, *NANOG*, and *OCT4* 72 h post-transfection. (B) Relative mRNA levels 96 h post-transfection. Values were calculated using the –ΔΔCt method normalized against *GAPDH*. Data is representative of 3 independent experiments. *$p < 0.05$ vs. scramble control (*t*-test).

OCT4 may be unperturbed by *KLF4* and *NANOG* knockdown due to an independent regulatory mechanism in NCCIT cancer cells than in ESCs, or, more likely, because residual KLF4 and NANOG protein may persist in the cells following knockdown. When *NANOG*-depleted NCCIT cells were followed for 4+ days after transfection, *OCT4* levels decreased indicating that there is a delay in the responsiveness of *OCT4* expression to *KLF4* and *NANOG* knockdown (Fig. 2.6B).

To further confirm that KLF4 is a direct regulator of *NANOG*, we performed additional ChIP assays to determine KLF4 localization within the genome. We observed that KLF4 was enriched at the *NANOG1* locus in NCCIT cells,

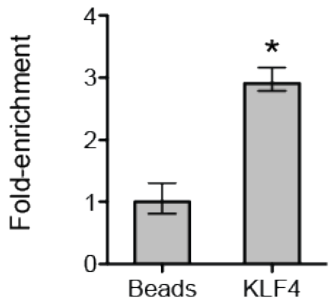

Fig. 2.7. KLF4 localizes to the *NANOG1* promoter in NCCIT cells. Fold-enrichment was calculated by the ΔCt method relative to beads. Data is representative of 3 independent experiments. Error bars indicate the upper and lower limits within measurements.

indicting that KLF4 is a direct and essential positive regulator of *NANOG* expression (Fig. 2.7).

2.6. Conclusions and Future Directions

The data presented here indicate the coactivator p300 is a direct and positive regulator of the *NANOG1* gene in pluripotent human cancer cells. To our knowledge, this is the first work to characterize an essential coactivator for ESC-associated genes in a non-ESC context. This information can be used to design strategies to modulate ESC gene expression profiles as well as guide rational anticancer drug development. Furthermore, these findings reveal significant similarities between p300/NANOG function in ESCs and cancer cells. p300 appears to control *NANOG* expression and function, regardless of cell differentiation status, indicating that at least some of the regulatory mechanisms in place in ESCs retain or regain activity in pathological cell states.

The finding that p300 directly regulates the *NANOG1* locus in pluripotent NCCIT cells stands in contrast to findings from other groups that the *NANOGP8* locus is a major source of *NANOG* expression in cancer cells (15, 35). These differences could partially be attributable to the various cell lines and cancer types examined (embryonal carcinoma and HNSCC here vs. prostate and colon), as well as differences in techniques. Those studies concluded *NANOGP8* was expressed based on sequencing and AlwNI digests of cDNA libraries. However, given the high degree of sequence identity between *NANOG1* and *NANOGP8*, and the relatively error-prone nature of RNA polymerase, the sequencing results may be unreliable. In support of this, Jeter et al. (15) reported multiple independent mutations within sequenced *NANOG* cDNAs, indicating a high rate of nucleotide substitution either in their cultures or in their readout.

We argue that by examining the promoter occupancy of activator and coactivators, we can more concretely identify the active gene loci. Here, we found that both p300 and H3Kac were significantly enriched at the *NANOG1* proximal promoter, whereas little to no detectable levels were found at the analogous *NANOGP8* region (Fig. 2.3B). Protein localization more strongly indicates function and transcriptional activity, and therefore we conclude that *NANOG1* is the predominant source of *NANOG* in the cell lines studied here.

The work presented here also represents the first analysis of *NANOG* promoter requirements in cancer cells. By systematically deleting and mutating proximal

promoter elements, we have identified both differentially required transcription binding sites, as well as a pair of critical KLF consensus motifs (Fig. 2.4-5). Disruption of either KLF' or KLF'' strongly inhibits *NANOG* promoter activity (Fig. 2.5) independent of cancer cell type or differentiation status, suggesting that KLF factors may be a common requirement for *NANOG* expression in cancer. In support of this, knockdown of *KLF4* dramatically reduced *NANOG* mRNA expression in NCCIT cells (Fig. 2.6), confirming that KLF4 is involved in *NANOG* regulation.

KLF4 is a known activator of *NANOG* in ESCs, but had not previously been considered to be a driver of cancer self-renewal. Interestingly, KLF4 has been reported to have both oncogenic and tumor suppressor functions (41). KLF4 is frequently overexpressed in squamous cell carcinomas and breast cancers, and gain of KLF4 appears to represent an early event in tumor development (42, 43). Furthermore, KLF4 also possesses reprogramming capacity and is often including in inducible pluripotent stem cell generation to increase dedifferentiation efficiency (44), suggesting that gain of KLF4 in cancer or precancerous cells could activate *NANOG* expression and thereby establish a self-reinforcing, stem cell-like genetic program that confers both growth and survival advantages.

Intriguingly, the KLF4 activator has also been reported to rely on p300 for full transcriptional activation (45-48). The cooccupancy of both KLF4 and p300 at the *NANOG1* promoter raises the possibility of a physical connection between

KLF4 and p300 that could account for the mutual requirements for both activator and coactivator. Several reports have found that KLF4 is able to associate with p300 (47, 48), suggesting a potential route by which p300 could be recruited to the DNA. Identifying the potential activator-coactivator interactions that lead to p300 recruitment to the *NANOG1* promoter could reveal further details into the regulatory complexes and mechanisms leading to *NANOG* expression, as well as isolate protein-protein interactions to target in small molecule probe development. Such tools could provide new and exciting opportunities for modulating *NANOG* expression in cancer.

2.7. Acknowledgements

I would like to acknowledge the contributions of our collaborators Prof. Quintin Pan (Ohio State University) , Dr. Xiujie Xie, and Dr. Manchao Zang for their work generating the p300-knockdown HNSCC cell lines all data associated with their phenotypes.

2.8. Materials and methods

Cell lines

NCCIT, MCF7, MDA-MB-231, and LNCap cells were purchased from ATCC. UMSCC74B cells were a gift from Dr. Thomas Carey (University of Michigan). NCCIT cells were grown in RPMI media 1640 (Life Technologies, 11875) supplemented with 10% fetal bovine serum (Atlanta Biologicals, S11195) in a 37 ºC incubator with 5% CO_2. All other cell lines were cultured in Dulbecco's

Modified Eagle Medium (DMEM) (Invitrogen, 11965) supplemented with 10% fetal bovine serum.

Chromatin immunoprecipitation assays

ChIP assays were performed on >90% confluent cells in 10-cm plates using the Abcam ChIP (ab-500) protocol.. Briefly, cells were dissociated by trypsin digest and resuspended in media + 10% FBS before counting. 10 x 10^6 cells were then pelleted by centrifugation and then washed with chilled PBS. Cells were then fixed with 1% formaldehyde for 10 min at room temperature before quenching with 1.25 M glycine. The fixed cells were washed with chilled PBS and resuspended in lysis buffer. Samples were then pelleted to collect the fixed chromatin. Mild sonication was carried out to yield sheared chromatin averaging ~500 bp in size. Immunoprecipitations were conducted on the sheared chromatin using anti-histone H3 (Abcam, ab1791), anti-p300 (Santa Cruz, sc-584 X), anti-acetylated histone H3 antibody (Millipore, 06-599), or Protein A beads alone as a negative control. The recovered DNA was used as a template for qPCR reactions using SYBR Green (Applied Biosystems, 4309155) according to the manufacturer's instructions.

Plasmids and cloning

The *NANOG* promoter-driven luciferase plasmid was originally purchased from Addgene (pNANOG-Luc, plasmid 25900). Deletions and point mutations were

then introduced using site-directed mutagenesis. Mutants were validated by DNA sequencing.

Transfections

NCCIT cells were transfected using OptiMEM media (Invitrogen) and Lipofectamine LTX (Invitrogen) according to an optimized protocol. Briefly, DNA•liposome complexes were formed in OptiMEM by sequentially adding DNA, PLUS reagent, and LTX and incubating for 30 min at room temperature. Four volumes of RPMI (10% FBS) was then added to the transfection mix before applying to the cells. The cells were incubated with the transfection mix for 24 h before switching media and grown for an addition 24 h.

Luciferase assays

10,000 cells were plated in 96-well format with 100 µl media + 10% FBS. After overnight incubation, the media was replaced with transfection mix: 100 µL media with 5 ng/ml *Renilla* luciferase plasmid (pRLSV40) and 395 ng/ml firefly luciferase plasmid. After 48 h incubation, luciferase activity is measured on a Berthold luminometer with firefly substrate from the Dual-Luciferase Assay Reporter System (Promega, E1980).

Knockdowns

100,000 NCCIT cells were cotransfected with a puromycin-selectable plasmid (250 ng/ml) and either scrambled or targeted shRNA vectors (1 µg/ml). Cells

were allowed to grow 24 h before selecting for transfected cells with 2 μg/ml puromycin in standard growth media. Cells were incubated for 72 h before harvesting RNA or protein for analysis, and fresh media with puromycin was applied every 24 h.

2.9. References

1. Guarente L. Transcriptional coactivators in yeast and beyond. *Trends Biochem. Sci*. **20**, 517-521 (1995).
2. Ptashne M & Gann A. Transcriptional activation by recruitment. *Nature* **386**, 569-577 (1997).
3. Naar AM, Lemon BD, Tjian R. Transcriptional coactivator complexes. *Annu. Rev. Biochem.* **70**, 475-501 (2001).
4. Ho L, *et al*. An embryonic stem cell chromatin remodeling complex, esBaf, is an essential component of the core pluripotency transcriptional network. *Proc. Natl. Acad. Sci. USA* **106**, 5187-5191 (2009).
5. Ho L, *et al*. An embryonic stem cell chromatin remodeling complex, esBaf, is essential for embryonic stem cell self-renewal and pluripotency. *Proc. Natl. Acad. Sci USA* **106**, 5181-5186 (2009).
6. Wang F, Marshal CB, Ikura M. Transcriptional/epigenetic regulatory CBP/p300 in tumorigenesis: structural and functional versatility. *Cell. Mol. Life Sci.* **70**, 3989-4008 (2013).
7. Farmer SR. Transcriptional control of adipocyte formation. *Cell Metab.* **4**, 263-273 (2006).
8. Scarpulla RC. Transcriptional paradigms in mammalian mitochondrial biogenesis and function. *Physiol. Rev.* **88**, 611-638 (2008).
9. Goodman RH & Smolik S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **14**, 1553-1577 (2000).
10. Santer FR, *et al*. Inhibition of the acetyltransferase p300 and CBP reveals a targetable function for p300 in the survival and invasion pathways of prostate cancer cell lines. *Mol. Cancer Ther.* **10**, 1644-1655 (2011).
11. Sweis RF & Michaelides MR. Recent advances in small-molecule modulation of epigenetic targets: discovery and development of histone methyltransferases and bromodomain inhibitors. *Annu. Reports Med. Chem.* **48**, 185-203 (2013).
12. Kushal S, *et al*. Protein domain mimetics as in vivo modulators of hypoxia-inducible factor signaling. *Proc. Natl. Acad. Sci. USA* **110**, 15602-15607 (2013).
13. Taylor CE, Pan QT, Mapp AK. Synergistic enhancement of the potency and selectivity of small molecule transcriptional inhibitors. *ACS Med. Chem. Lett.* **3**, 30-34 (2012).
14. Zbinden M, *et al*. NANOG regulates glioma stem cells and is essential in vivo acting in a cross-functional network with GLI1 and p53. *EMBO J.* **29**, 2659-2674 (2010).

15. Jeter CR, *et al*. Functional evidence that the self-renewal gene NANOG regulates human tumor development. *Stem Cells* **27**, 993-1005 (2009).
16. Chen C, Wei Y, Hummel M, Hoffmann TK, Gross M, Kaufmann AM, Albers AE. Evidence for epithelial-mesenchymal transition in cancer stem cells of head and neck squamous cell carcinoma. *PLoS ONE* **6**, e16466 (2011).
17. Chiou SH, *et al*. Positive correlations of Oct4- and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma. *Clin. Cancer Res*. **14**, 4085-4095 (2008).
18. Tsai LL, *et al*. Markedly increased Oct4 and Nanog expression correlates with cisplatin resistance in oral squamous cell carcinoma. *J. Oral Pathol. Med*. **40**, 621- 628 (2011).
19. Yu CC, *et al*. MicroRNA let-7a represses chemoresistance and tumourigenicity in head and neck cancer via stem-like properties ablation. *Oral Oncol*. **47**, 202-210 (2011).
20. Jeter CR, *et al*. NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. *Oncogene* **30**, 3833-3845 (2011).
21. Loh YH, *et al*. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431-440 (2006).
22. Wang J, *et al*. A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364-368 (2006).
23. van den Berg DL, *et al*. An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* **6**, 369-381 (2010).
24. Tutter AV, *et al*. Role for Med12 in regulation of Nanog and Nanog target genes. *J. Biol. Chem.* **284**, 3709-3718 (2009).
25. Chen X, *et al*. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106-1117 (2008).
26. Fong YW, *et al*. A DNA repair complex functions as an Oct4/Sox2 coactivator in embryonic stem cells. *Cell* **147**, 120-131 (2011).
27. Baltus GA, *et al*. A positive regulatory role for the mSin3A-HDAC complex in pluripotency through Nanog and Sox2. *J. Biol. Chem.* **284**, 6998-7006 (2009).
28. Zhong XM & Jin Y. Critical roles of coactivator p300 in mouse embryonic stem cell differentiation and Nanog expression. *J. Biol. Chem.* **284**, 9168-9175 (2009).
29. Xie X, *et al*. Phosphorylation of Nanog is essential to regulate Bmi-1 and promote tumorigenesis. *Oncogene* **33**, 2040-2052 (2014).
30. Chen C, Wei Y, Hummel M, Hoffmann TK, Gross M, Kaufmann AM, Albers AE. Evidence for epithelial-mesenchymal transition in cancer stem cells of head and neck squamous cell carcinoma. *PLoS ONE* **6**, e16466 (2011).

31. Damjanov V, Bozidar H, Gibas Z. Retinoic acid-induced differentiation of the developmentally pluripotent human germ cell tumor-derived cell line, NCCIT. *Laboratory Investigation* **68**, 220-232 (1993).
32. You JS, *et al*. Depletion of embryonic stem cell signature by histone deacetylase inhibitor in NCCIT cells: involvement in Nanog suppression. *Cancer Res.* **69**, 5716-5725 (2009).
33. Zhang J, *et al*. NANOG modulates stemness in human colorectal cancer. *Oncogene* **32**, 4397-4405 (2013).
34. Zhang JY, *et al*. NANOGP8 is a retrogene expressed in cancers. *FEBS J.* **273**, 1723-1730 (2006).
35. Ishiguro T, *et al*. Differential expression of NANOG1 and NANOGp8 in colon cancer cells. *Biochem. Biophys. Res. Commun.* **418**, 199-204 (2012).
36. Ibrahim EE, *et al*. Embryonic NANOG activity defines colorectal cancer stem cells and modulates through AP1- and TCF-dependent mechanisms. *Stem Cells* **30**, 2076-2087 (2012).
37. Wang Z, *et al*. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019-1031 (2009).
38. Xu RH, *et al*. NANOG is a direct target of TGF beta/activin-mediated SMAD signaling in human ESCs. *Cell Stem Cell* **3**, 196-206 (2008).
39. Chan KKK, *et al*. KLF4 and PBX1 directly regulate NANOG expression in human embryonic stem cells. *Stem Cells* **37**, 2114-2125 (2009).
40. Kuroda T, *et al*. Octamer and Sox elements are required for transcriptional *cis* regulation of *Nanog* gene expression. *Mol. Cell. Biol.* **25**, 2475-2485 (2005).
41. Rowland BD & Peeper DS. KLF4, p21 and context-dependent opposing forces in cancer. *Nat. Rev. Cancer* **6**, 11-23 (2006).
42. Foster KW, *et al*. Oncogene expression cloning by retroviral transduction of adenovirus E1A-immortalized rat kidney RK3E cells: transformation of a host with epithelial features by c-MYC and the zinc finger protein GKLF. *Cell Growth Differ.* **10**, 423-434 (1999).
43. Foster KW, *et al*. Increase of *GKLF* messenger RNA and protein expression during progression of breast cancer. *Cancer Res.* **60**, 6488-6495 (2000).
44. Takahashi K & Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and fibroblast cultures by defined factors. *Cell* **126**, 663-676 (2006).
45. Wang WP, *et al*. The EP300, KMD5A, KDM6A and KDM6B chromatin regulators cooperate with KLF4 in the transcriptional activation of POU5F1. *PLoS ONE* **7**, e52556 (2012).
46. Zhang R, *et al*. Kruppel-like factor 4 interacts with p300 to activate mitiofusin 2 gene expression induced by all-trans retinoic acid in VSMCs. *Act. Pharm. Sin.* **31**, 1293-1302 (2010).

47. Evans PM, *et al*. Kruppel-like factor 4 is acetylated by p300 and regulates gene transcription via modulation of histone acetylation. *J. Biol. Chem*. **282**, 33994-34002 (2007).
48. Geiman DE, *et al*. Transactivation and growth suppression by the gut-enriched Kruppel-like factor (Kruppel-like factor 4) are dependent on acidic amino acid residues and protein-protein interaction. *Nucl. Acid Res.* **28**, 1106-1113 (2000).

Chapter 3

Identification of key p300 functions during *NANOG* transcription

3.1. Abstract

Coactivator-mediated transcription involves a variety of protein-protein interactions (PPIs) and enzymatic activities to regulate gene expression. Multifunctional coactivators such as p300 can fulfill differential roles depending upon the target gene context, and as a result it is often unclear whether a coactivator acts as a scaffolding protein, a catalytic subunit, or a combination of the two. Based on our finding that p300 is an essential and direct regulator of *NANOG* expression in cancer cells, we sought to dissect the relevant PPI and enzymatic contributions of p300 to *NANOG* transcription. Individual domains within p300 were systematically perturbed through the use of dominant-negative constructs, site-directed mutagenesis, and small molecule inhibitors in order to identify the crucial domains participating in *NANOG* transcription. We found that the CH1 domain in particular appears to be essential for overall p300 activity, suggesting a role in initial recruitment of p300 to the *NANOG* promoter. Additionally, we found that inhibition of the histone acetyltransferase (HAT) domain rapidly depletes p300-mediated histone acetylation at the *NANOG* promoter and downregulates *NANOG* expression in pluripotent cancer cells.

Interestingly p300 HAT inhibition does not disrupt p300 localization, suggesting that HAT activity is not necessary for p300 occupancy at the *NANOG* promoter. These results allowed us to propose a model for p300-mediated *NANOG* activation that reveals new potential targets for modulating *NANOG* expression.

3.2. p300 domains and their interactions

Eukaryotic gene activation requires the concerted function of transcription factors and coactivators (1). Transcriptional activators bind cognate sites in the promoters and enhancers of target genes and stimulate transcription by bringing the basal transcriptional machinery, including general factors and RNA Pol II itself, to the transcriptional start site (2, 3). Although some activators directly interact with the general factors such as TFIID, in most cases, additional proteins or multiprotein complexes termed coactivators are required to facilitate this process (4-6). Initially, coactivators were viewed as adaptor protein that connect the sequence-specific transcription factors to the basal transcription machinery, yet it was later discovered that some coactivators have additional functions. In eukaryotic cells, DNA wraps around histone octamers to assemble nucleosomes, which are further packaged into condensed chromatin that is inaccessible for transcription (7). Several classes of transcriptional coactivators were found to possess chromatin-remodeling or modification activity, which opens the chromatin structure to allow effective gene transcription (6).

The coactivator p300 performs both of these functions: bridging of DNA-binding and general transcription factors, and relaxation of chromatin through its intrinsic histone acetyltransferase (HAT) activity (Fig. 3.1A)(8-10). p300 is a large,

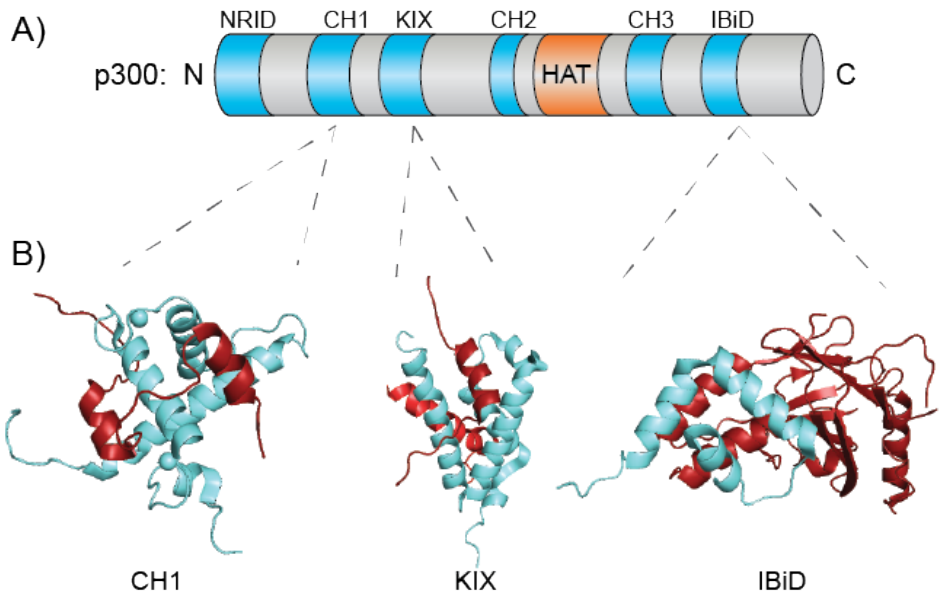

Fig.3.1: The multidomain coactivator p300. (A) Schematic of the domain organization of p300. Domains that engage in protein-protein interaction are shown in blue. The catalytic histone acetyltransferase (HAT) domain is shown in orange. (B) Representative structures of activator-binding domains (blue) complexed with known interaction partners (red). Note the diversity in binding modes. PDBs: CH1-HIF1α (1L3E); KIX-MLL-c-Myb (2AGH); IBiD-IRF3 (1ZOQ).

multifunctional protein with a catalytic HAT domain as well as activator-binding domains that include the Cysteine-Histidine-rich region 1 (CH1)(11, 12), the CREB- and MLL-interacting KIX domain (13), another Cystein-Histidine-rich region (CH3)(12, 14), and the interferon-binding domain (IBiD)(Fig. 4.1A)(15, 16). These domains mediate the protein-protein interactions with DNA-binding transcription factors and the basal transcription machinery, as well as other coactivators such as PCAF/GCN5 (17). The activator-binding domain interactions are frequently with partners that are intrinsically disordered in

isolation and undergo binding-coupled folding. Several structures of p300 domains in complex with transcriptional activation domains (TADs) have been described and shed light on the diverse mechanisms of coactivator recruitment during transcriptional activation (Fig. 3.1B).

The most extensively studied interaction is that between the so-called kinase-inducible domain (pKID) of CREB and the KIX domain of p300 (18). The pKID•KIX complex has been determined by solution nuclear magnetic resonance (NMR) spectroscopy (19). The KIX domain is composed of a three-helix bundle with two additional short 3_{10} helices and is relatively stable on its own compared to the disordered pKID in isolation. pKID binds a largely hydrophobic groove formed by helices H1 and H3 of KIX, with a K_d of ~700 nM. This is ~100-fold tighter than the interaction between unphosphorylated KID peptide and KIX. The NMR structure suggested that the increased binding affinity in pKID stems from an interaction between the phosphate group and Y658 of KIX (19, 20).

The binding mode of pKID to KIX, however, is not representative of all TAD•KIX interactions. For instance, the TAD of the transcription factor c-MYB binds the same site as pKID (KIX H3-H4), but binds as a single helix (21). Furthermore, a second binding site for TADs was identified on the opposite face of KIX involving a hydrophobic groove formed by the H2 and H3 helices. This second site can be occupied by the TAD of the myeloid-lymphoid leukemia (MLL) protein. Remarkably, both c-MYB and MLL TADs can bind to the two KIX sites

simultaneously and cooperatively (22-24). This potential allostery between activator-binding sites in KIX could account for the concerted recruitment of p300 by CREB and the viral protein Tax (25). A variety of other TADs have been reported to bind to one or both sites on KIX, including c-JUN, FOXO3a, and p53 (26-28). The binding of these TADs vary greatly, with differences in binding mode, affinity, and promiscuity– all common features of activator-coactivator PPIs.

Another well-documented activator-binding domain of p300 is the CH1 domain, and the highly similar CH3 domain. CH1 and CH3 are each comprised of four amphiphatic helices that fold around three HCCC-type zinc-binding sites (12, 29, 30). The structures of CH1 and CH3 are similar, however the fourth helix is found in opposite orientations in the two domains and has been invoked as a possible determinant of binding specificity for different TADs (29).

The TADs of some activators are specific for CH1 or CH3, whereas others interact with both as well as other p300 activator-binding domains. For example, the highly promiscuous TAD of p53 binds to both CH1 and CH3 (31, 32), as well as both KIX sites and the IBiD domain (31). In contrast, the CH1 domain recognizes the C-terminal activation domain of HIF1α, which functions in the maintenance of cellular oxygen homeostasis by inducing transcription of adaptive genes under hypoxic conditions (33). The HIF1α TAD, which is intrinsically disordered in isolation, wraps around the CH1 domain and folds into three short

helices linked by extended loop regions. This creates a large surface area for interaction resulting in a tight HIF1α•CH1 interaction (K_d ~ 7 nM) (34). The CBP/p300-interacting transactivator 2 (CITED2) protein negatively regulates HIF1α activity by competing for p300 binding (24, 35-37). CITED2 wraps around CH1 similarly to HIF1α, with a partially overlapping binding site and comparable affinity (K_d ~ 13 nM) (38).

These studies show that p300 can participate in a variety of protein-protein interactions, with diversity observed even within the same domain. Thus, in the absence of known interactions (such as at the *NANOG* promoter), it is difficult to predict which domain(s) may be essential for coactivator function. We therefore set out to identify crucial p300 domains for *NANOG* expression by systematically testing the gain and loss of function for each in order to guide future experiments and potential modulation strategies.

3.3 Identification of essential activator-binding domains

Given the diverse and potentially multivalent types of interactions the p300 activator-binding domains are capable of engaging in, we sought to first identify which domain(s) are required for full p300 activity at the *NANOG* promoter. To do this, we cloned out individual domains from the human p300 sequence for "squelching" experiments (Fig. 3.2A,B). p300 and the related coactivator CBP are expressed at limiting concentrations in the nucleus, and thus transcription of many p300/CBP-target genes is dose-dependent (39-41). Therefore,

overexpressed domains can compete with endogenous full-length p300 for binding to DNA-bound activators and thereby act as dominant-negative constructs to inhibit p300-dependent transcription.

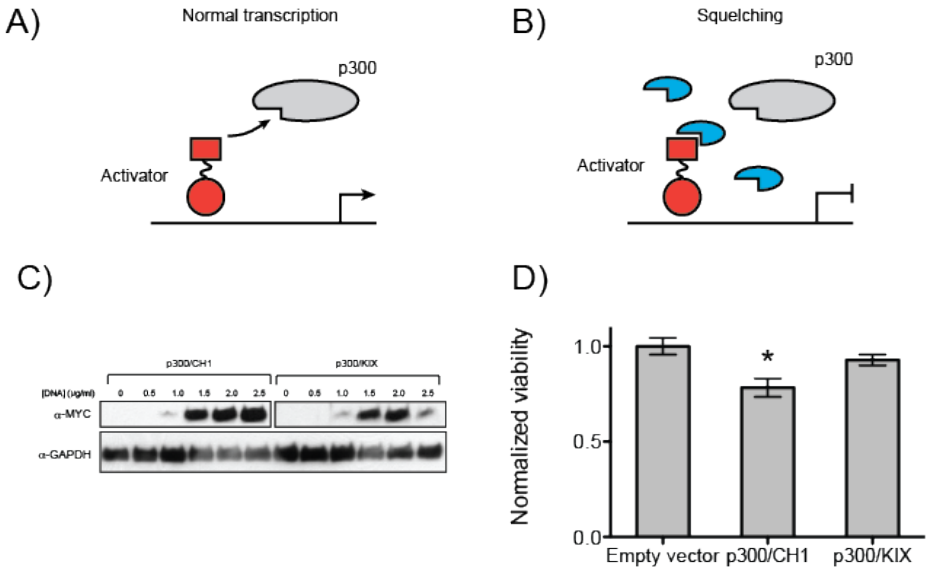

Fig.3.2: Squelching experiments with dominant-negative p300 domains. (A) Normal transcription is mediated by interactions between DNA-bound activators (red) and coactivator proteins like p300 (grey). (B) When dominant-negative domains from p300 are overexpressed (blue), they compete for binding to the activators and block coactivator recruitment and activity, leading to transcriptional inhibition. (C) Expression levels of individual domain constructs in NCCIT cells. (D) Relative viability of NCCIT cells 48 h post-transfection. Viability was assessed by WST-1 assay. Data is represented as mean ± SD. *$p < 0.05$ relative to promoter activity with empty vector (student's t-test).

Immunoblotting confirmed robust expression of the p300/CH1, p300/KIX, and p300/CH3 constructs (Fig. 3.2C). Furthermore, p300/CH1 or p300/KIX overexpression was generally tolerated by the cells, with p300/CH1 eliciting an ~15% decrease in cell viability and KIX having no significant impact on viability (Fig. 3.2D).

To confirm that the expressed domains are functional we tested their ability to suppress the transcriptional activity of two Gal4-fused transcriptional activation domains (TAD): the KIX-dependent TAD MLL and the CH1-dependent TAD HIF1α. As expected, increasing amounts of KIX dose-dependently reduced the activity of Gal4-MLL, but had no effect on Gal4-HIF1α (Fig. 3.3A). In contrast, CH1 dose-dependently inhibited HIF1α-driven transcription without affecting MLL activity. These results confirm that the CH1 and KIX constructs can inhibit transcription from model p300-dependent TADs.

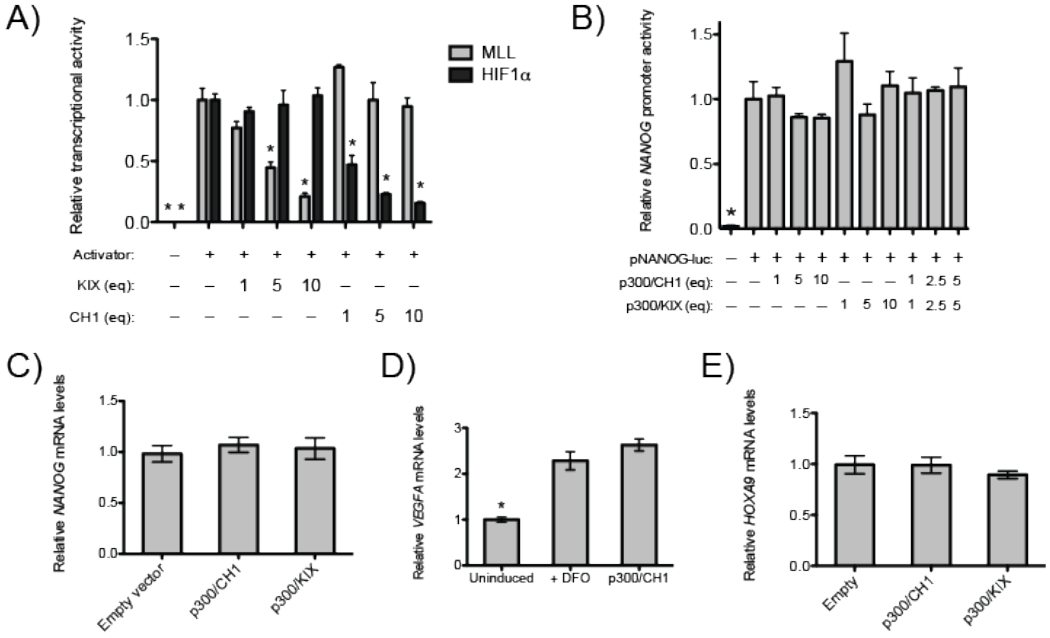

Fig. 3.3. Dominant-negative domain constructs do not inhibit endogenous gene expression in pluripotent NCCIT cells. (A) Gal4-reporter activity in HeLa cells cotransfected with increasing amounts of either the CH1 or KIX constructs. Data is shown as mean ± SD. *$p < 0.05$ vs. activator alone (*t*-test). (B) *NANOG* promoter activity in NCCIT cells cotransfected with CH1, KIX, or both. Data is shown as mean ± SD. *$p < 0.05$ vs. promoter activity alone. (C) *NANOG* mRNA levels in NCCIT cells 48 h post-transfection with CH1, KIX, or empty vector. (D) *VEGFA* mRNA levels in HeLa cells. Cells were treated with DFO for 6 h to induce the hypoxia response. Data is shown as mean ± SD for three independent experiments. *$p < 0.05$ vs. +DFO control. (E) *HOXA9* mRNA levels in HeLa cells. Neither CH1 nor KIX blocked *HOXA9* expression 48 h post-transfection (student's *t*-test).

When stably overexpressed in the head and neck squamous cell carcinoma (HNSCC) cell line UMSCC74A, the CH1 domain inhibited NANOG transcriptional activity by >50% (Fig. 3.4A,B). Furthermore, p300/CH1 decreased the

transcriptional activity of a NANOG-driven reporter gene (Fig. 3.2B), as well as reduced the ability of UMSCC74A cells to form tumorspheres by 80% (Fig 3.4C), suggesting a strong decrease in CIC frequency. To test this, immunocompromised mouse xenograft studies were performed to evaluate the *in vivo* potential of UMSCC74A CICs to generate new tumors. As anticipated, control UMSCC74A cells robustly formed palpable tumors within 60 days (Fig. 3.4D). Remarkably, CH1-overexpressing cells were unable to found tumors even when 1 x 10^6 cells were injected, indicating that the dominant negative form of CH1 effectively blocks CIC potential *in vivo*.

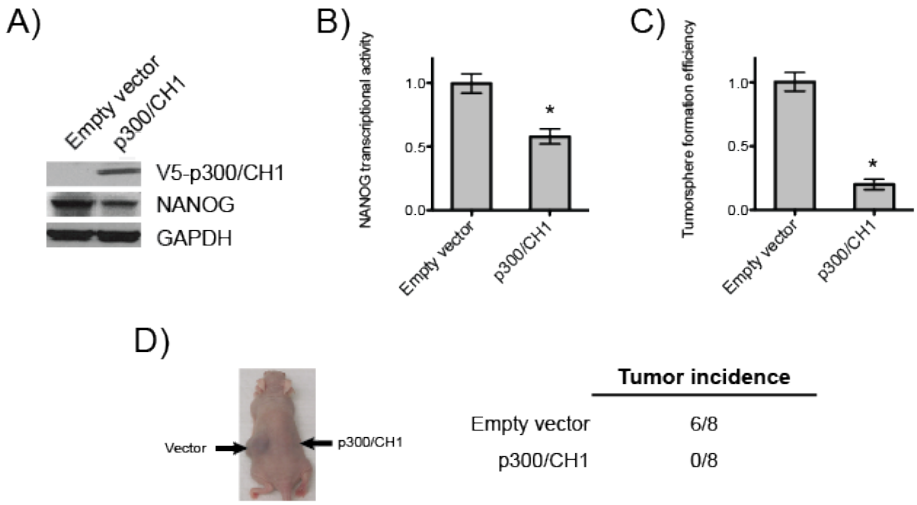

Fig. 3.4. p300/CH1 reduces NANOG transcriptional activity and depletes the CIC population in UM-SCC74A cells. Stable polyclonal UMSCC74A/empty and UMSCC74A/Taz1 cells were generated by transfection and antibiotic selection. (A) Immunoblot for V5-tagged p300/CH1, NANOG, and GAPDH. . (B) NANOG transcriptional activity. NANOG (NANOG TRE-luciferase) transcriptional activity was measured using a luminometer. Data is presented and mean ± SEM. *P<0.01 vs. vector control (n = 3) (C) Cancer initiating cell number. Cells were harvested and seeded on low-attachment plates in a defined, serum-free culture medium for 7 days. Tumorsphere formation efficiency was calculated as the number of tumorspheres (≥50 µm in diameter) formed divided by the original number of cells seeded. Data is presented as mean ± SEM. *P<0.01 (n = 6). (D) In vivo tumor incidence. UMSCC74A-empty or UM-SCC74A-p300/CH1 cells (1x10^6 cells) were suspended in DMEM (1:1 Matrigel) and implanted subcutaneously in the flanks of athymic nude mice. Tumor incidence was monitored for 60 days post-implantation. Experiments were performed by X. Xie, M. Zhang, and Q. Pan of Ohio State University.

When these studies were extended to pluripotent NCCIT cells, neither CH1 nor KIX overexpression were able to inhibit transcription from the *NANOG* proximal

promoter, either alone or in combination (Fig. 3.3B). Furthermore, p300 domain overexpression also failed to inhibit endogenous expression levels of NANOG (Fig. 3.3C), which would suggest that the CH1 and KIX domains are not essential p300 domains for *NANOG* activation. However, further positive controls testing the ability of the dominant-negative domains to inhibit endogenous gene expression also failed. p300/CH1 overexpression failed to inhibit DFO-induced *VEGFA* expression in HeLa cells (Fig. 3.3D), whereas p300/KIX overexpression had no significant effect on MLL-driven *HOXA9* expression (Fig. 3.3E), indicating that the independent domain constructs are not suitable for perturbing expression of endogenous genes in NCCIT cells.

We therefore designed an alternative method for systematically assessing the structural domains of p300 required for *NANOG* expression. We reasoned that, due to the independent folding and function of the activator-binding domains within p300, we could create in-frame deletions within the coactivator for internal loss-of-function studies (Fig. 3.5A). Other groups have used his technique to study the roles of p300 in hematopoiesis (42). A series of internal domain deletions within p300 were constructed (Fig. 3.5A), and their expression was confirmed by immunoblotting (Fig. 3.5B). Due to the high molecular weight of full-length p300, small differences in molecular weight were difficult to resolve by standard SDS-PAGE.

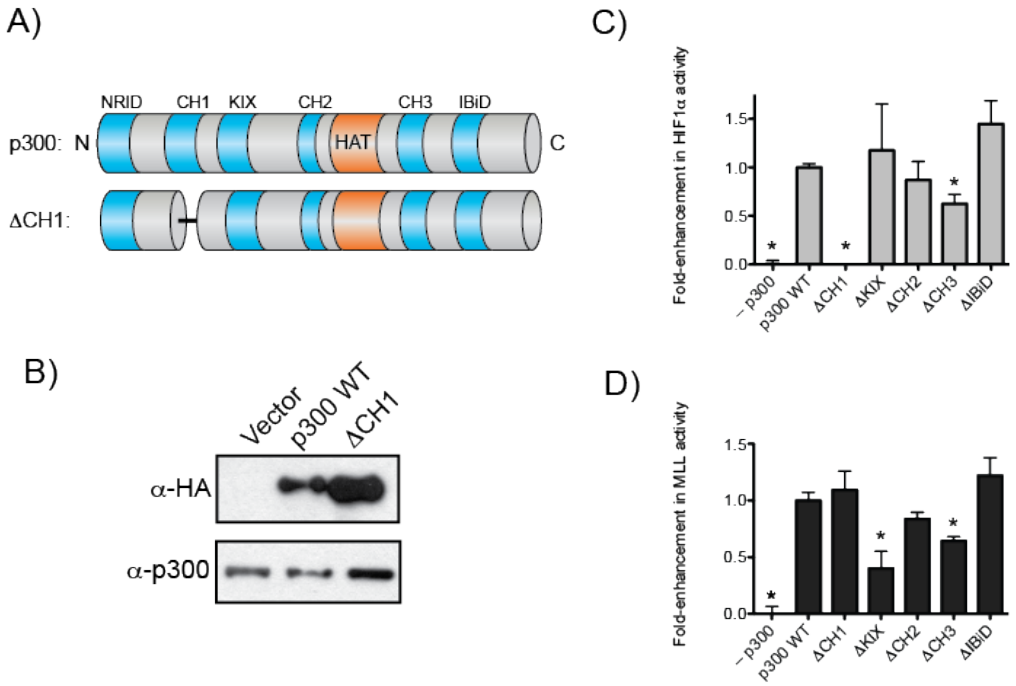

Fig. 3.5. Domain deletion validation. (A) Schematic of full-length p300 and the representative ΔCH1 construct lacking the CH1 domain. Other constructs not shown for brevity. (B) Domain deletion constructs are expressed in HeLa cells. (C and D) Fold-enhancement in Gal4-HIF1a (C) or Gal4-MLL (D) activity in HeLa cells coexpressing WT p300 or various domain deletion constructs. Data is shown as mean ± SD. *$p < 0.05$ vs. WT p300. (student's *t*-test).

To confirm that the deleted domains had functional effects, the mutant coactivators were tested for their ability to enhance HIF1α- and MLL-driven reporter gene epxression (Fig. 3.5C). Wild-type (WT) p300 was able to strongly enhance GAL4-HIF1a activity. As anticipated, this enhancement was completely lost when the CH1 domain was deleted (ΔCH1)(Fig. 3.5C). Loss of other domains had little to no inhibitory effect on coactivator-mediated enhancement, except for ΔCH3 which is structurally very similar to the CH1 domain and has been shown to affect HIF1a-mediated transcription (43). In contrast, ΔCH1 had no impact on GAL4-MLL activity (Fig. 3.5D), while ΔKIX showed ~3-fold lower activity relative to WT, indicating that the internal domain deletions selectively perturb coactivator functions. Interestingly, ΔCH3 also displayed a mild (~20%)

reduction in MLL activity enhancement. KIX and CH3 are not structurally similar (Fig. 3.1B), and so far there are no reports of a KIX-CH3 or MLL-CH3 interaction that could potentially account for this observation. The possibility of intraprotein interactions within p300 is intriguing but currently outside the scope of the work presented here.

We next tested the ability of ΔCH1 and ΔKIX to enhance *NANOG*-promoter activity in a luciferase reporter assay. WT p300 increased promoter activity ~4-fold over an empty vector control (Fig. 3.6A), consistent with the proposal that p300 is a limiting coactivator of *NANOG* expression. Strikingly, the ability of ΔCH1 to enhance *NANOG* promoter activity was strongly attenuated (Fig 3.6A), suggesting that the CH1 domain may be a key mediator of p300 function. In contrast, ΔKIX had no effect on promoter activity, indicating that the KIX domain may be largely dispensable (Fig. 3.5A). The loss of other domains such as

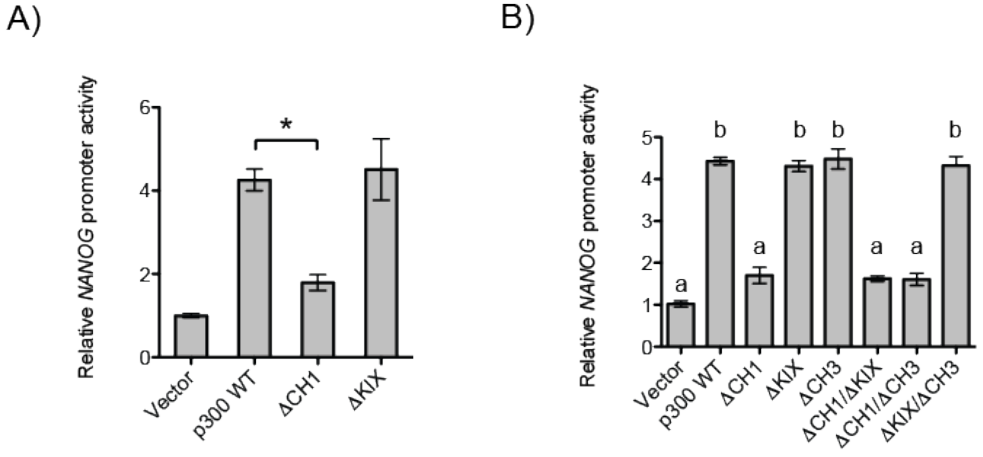

Fig. 3.6. p300 lacking the CH1 domain is unable to enhance transcription from the *NANOG* promoter. (A) *NANOG* promoter activity in NCCIT cells expressing wild-type p300 or ΔCH1 or ΔKIX. Data is shown as mean ± SD (n =6) (B) Compound deletions do not further affect activity. Data is shown as mean ± SD (n = 4). *$p < 0.05$ vs. p300 WT alone (ANOVA and Tukey's post-hoc test).

deletions of ΔCH1/ΔKIX and ΔCH1/ΔCH3 did not exhibit any greater defects than ΔCH1 alone (Fig. 3.6B), suggesting: i) that the CH1 domain is the major PPI interface for p300 during *NANOG* transcription, and ii) that the other domains do not cooperate with CH1 such as in the formation of multivalent interactions. The CH1 domain therefore appears to be of crucial importance for p300 function here.

The CH1 domain itself adopts a compact globular fold around three zinc-binding sites (Fig. 3.7A) (12, 29) and this folded structure serves as a template for intrinsically disordered activators to bind (29). Mutation of the key zinc-coordinating residues to alanine disrupts activator-CH1 interactions and impairs CH1-mediated transcription (44). Some residues appear to be essential for general CH1-domain interactions, while others have effects on specific

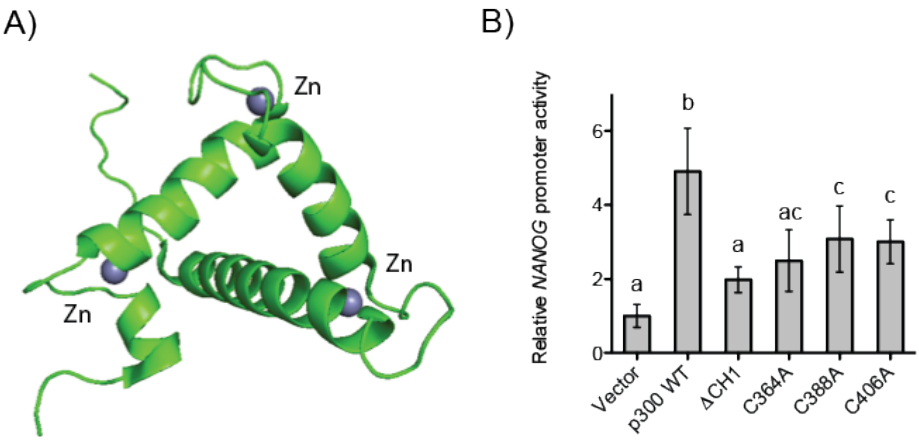

Fig. 3.7. Complete structural integrity of the CH1 domain is required for p300-mediated *NANOG* promoter activity. (A) Solution structure of the p300/CH1 domain (green). The domain folds around three zinc ions (grey) that are located coordinated by distinct cystein/histidine motifs. (B) Mutations that disrupt zinc binding in the CH1 domain attenuate p300-mediated *NANOG* promoter activity. Data is shown as mean ± SD (n = 4). $*p < 0.05$ vs. p300 WT alone (ANOVA and Tukey's post-hoc test).

interactions. For example, C379A in CBP CH1 abrogates binding to HIF1a, NF-κB, and Mrg1, whereas H407A can still bind HIF1a and NF-kB but not Mrg1 (44). Therefore, it appears that the zinc bundles are directly involved in interactions made by some but not all proteins that bind the CH1 domain. We reasoned that, in the absence of a known binding partner at the *NANOG* promoter, we could further explore the role of the CH1 domain by introducing single amino acid substitutions at the zinc-coordinating residues. Three cysteine residues (C364, C388, and C406), representing each of the three zinc-binding sites, were mutated within full-length p300 and tested for their ability to enhance *NANOG* promoter activity as described above.

As shown in Fig. 3.7B, p300 WT consistently enhanced promoter activity 4-5-fold, while ΔCH1 was statistically indistinguishable from basal promoter activity. Interestingly, all three point mutations (C364A, C388A, and C406A) exhibited reduced enhancement potential relative to WT. C364A was not found to be different from ΔCH1 (ANOVA), while C388A and C406A showed milder impairments in activity (~3-fold enhancement). These results further demonstrate that disruption of the CH1 domain, either by deletion or point mutation, greatly reduces the ability of p300 to function as a transcriptional coactivator at the *NANOG* promoter. Furthermore, the fact that all three zinc-binding sites are essential for CH1 function suggests that complete integrity of the CH1 domain is required for full activity. Local misfolding due to improperly coordinated metal ions is sufficient to disrupt CH1 function, likely by disrupting the key interaction(s) CH1 engages in.

Taken together, these data reveal the CH1 domain to be a core contributor to p300 function at the *NANOG* promoter.

3.4. Investigating the role of p300 HAT activity

In addition to the role of the CH1 domain, we also wished to explore the potential catalytic contributions of p300 in *NANOG* expression. p300 contains a HAT domain that acetylates nucleosomal histones (Fig. 3.1A)(8, 9). *In vitro*, p300 acetylates all acetylation sites on histones H2A and H2B, and preferentially acetylates histone H3 at K14 and K18, and histone H4 at K5 and K8 (45). In

addition to the histone N-terminal tails, p300 is also able to acetylate a number of non-histone proteins including transcriptional activators (46-49). The requirement for p300 HAT activity can be gene-dependent (50), so it is possible that the HAT domain may be dispensable for *NANOG* expression. We therefore sought to establish the role of p300 enzymatic activity at the *NANOG* promoter.

A number of HAT inhibitors have been reported for p300 (51-56), but all suffer as chemical probes due to issues with toxicity, solubility, cell permeability, and selectivity. Bowers et al. (57) identified the pyrazolone compound C646 (Fig. 3.8A) as a potent and novel lysine-CoA competitor selective for the p300 HAT domain (K_i ~ 400 nM). More importantly, C646 was shown to suppress histone H3 and H4 acetylation in mouse fibroblasts at low micromolar concentrations

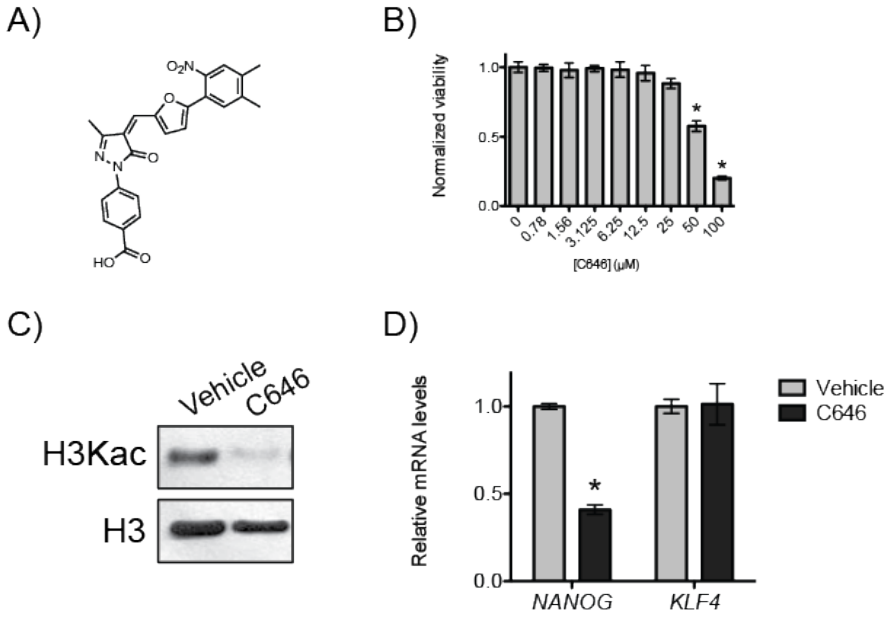

Fig. 3.8. p300 HAT activity is required for *NANOG* expression in NCCIT cells. (A) The p300-selective HAT inhibitor C646. (B) C646 effects on NCCIT viability after 48 h. Data is shown as mean ± SD (n =4). *$p < 0.05$ vs. vehicle (1% DMSO v/v). (C). NCCIT cells show greatly reduced H3Kac levels following 6 h treatment with C646 (40 μM). (D). *NANOG* and *KLF4* mRNA levels in NCCIT cells treated with C646 (20 μM)(black bars) for 6 h. Data is shown as mean ± SD (n = 3). *$p < 0.05$ vs. vehicle control. (student's *t*-test).

without significant cytotoxicity. In contrast, the garcinol-derivative LTK14 has a K_i of ~5 μM with an IC_{50} of ~5-7 μM, indicating a very small index for probing the effects of HAT inhibition (55). We therefore chose to use C646 as a way to explore the role of p300 HAT activity in *NANOG* regulation.

NCCIT cells were first treated with C646 at increasing concentrations to determine the effects on cellular viability (Fig. 3.8B). C646 displayed a GI_{50} of ~50 μM over 48 h and was well tolerated by the cells at lower concentrations, consistent with other reports (57). We chose to test the effects of C646 at both 10 μM and 40 μM initially to try and maximize the effects of p300 inhibition. These concentrations are in line with other reports of cell-based experiments using C646 (57). The large difference between the *in vitro* K_i (400 nM) and cellular assay (low micromolar) concentrations could be due to a variety of effects including cellular permeability, compound stability in cells, and sequestration by other cellular proteins. We therefore sought to verify that C646 could decrease p300 HAT in NANOG-expressing cells. To do this, we tested the ability of C646 to block histone deacetylase inhibitor-driven enrichments in histone H3 acetylation (H3Kac), a known marker for p300 HAT activity. NCCIT cells treated with C646 alone showed significantly reduced levels of H3Kac over basal acetylation levels (Fig. 3.8C), while trichostatin A (TSA) treatment greatly increased H3Kac levels as expected. Impressively, pretreatment with C646 was able to dose-dependently block TSA-induced H3Kac enrichment, indicating that

C646 can potently inhibit p300 HAT activity in NCCIT cells at sub-GI_{50} concentrations and may therefore be useful as a chemical probe.

Next we tested the effects of C646 on *NANOG* expression levels in NCCIT cells. An initial titration experiment indicated that C646 reduced *NANOG* mRNA levels by ~50% at both 20 μM and 40 μM concentrations after 24 h (data not shown). We chose to conduct future experiments at 20 μM to further avoid potential complications from viability effects. We also observed in time course experiments that C646 exhibited greater effects at earlier time points, possibly indicating that the compound is metabolized over time. When C646 was administered to NCCIT cells for 6 h at 20 μM, *NANOG* mRNA levels were still significantly suppressed by ~2-3-fold compared to vehicle treated cells (Fig. 3.8D). Furthermore, C646 had no effects on other ESC-related genes such as *KLF4* (Fig. 3.8D), suggesting that p300 HAT activity is acutely necessary for maintaining *NANOG* transcription, and that *NANOG* mRNA may have a brief half-life in the cell.

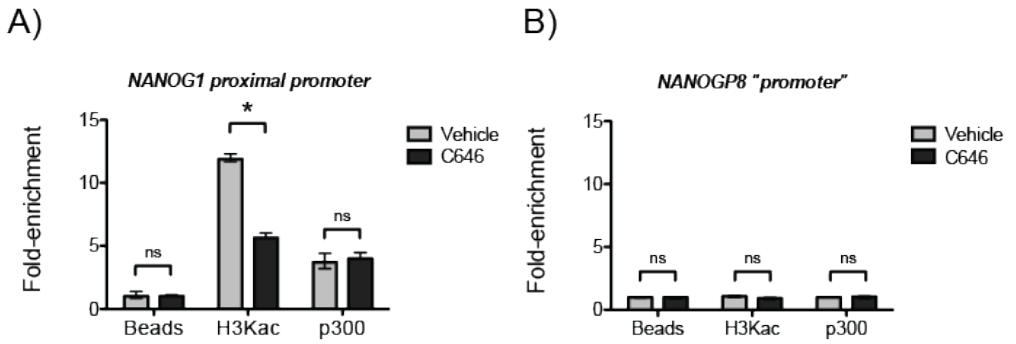

Fig. 3.9. C646 depletes H3Kac at the *NANOG1* proximal promoter. (A and B) Relative enrichment levels of H3Kac and p300 at the *NANOG1* (A) and *NANOGP8* (B) loci in NCCIT cells following 6 h treatment with vehicle or 20 μM C646 (vehicle = 1% DMSO v/v). Data is shown as mean ± SD (n =3). *$p < 0.05$ vs. vehicle control (*t*-test).

We next examined the effect of p300 HAT inhibition on histone acetylation and protein occupancy at the *NANOG* promoter to determine how C646 could be downregulating *NANOG* expression. In chromatin occupancy (ChIP) assays, 6 h treatment with C646 significantly reduced H3Kac levels at the *NANOG* promoter by ~50% (Fig. 3.9A). At the same time, p300 occupancy was not altered, indicating that p300 HAT activity is not required for coactivator recruitment, but is essential in maintaining high levels of histone acetylation. In contrast, when we examined an intragenic region within the *NANOG* gene body, we failed to detect significant levels of either H3Kac or p300 (Fig. 3.9B), consistent with other studies finding that p300 predominantly localizes to the promoter region of target genes (58). Taken together, these results suggest that p300 HAT activity is essential for acetylating histones in the *NANOG* promoter, but that HAT activity is a secondary event to initial p300 recruitment.

3.5. Conclusions and Future Directions

3.5.1.The role of the CH1 domain and its potential interactions

Our results indicate that both the CH1 domain and the catalytic HAT domain of p300 are both required for coactivator function and transcriptional activation of *NANOG* in some human cancer cell lines. Deletion or mutation of the CH1 domain significantly impairs the ability of p300 to drive expression from the *NANOG* promoter (Fig. 3.6-7). Given the role of the CH1 domain in binding DNA-bound activators, these results would suggest that CH1 is likely acting as an initial interface for recruiting p300 to the *NANOG* promoter. The observation that deletion of other activator-binding domains has little to no impact on p300 coactivator function (Fig. 3.5-6) indicates that the CH1 domain may be the major or only domain involved in activator recognition and p300 localization.

The CH1 domain is known to interact with a variety of transcriptional activators, including p53, HIF1a, and ETS1 (43, 59). Intriguingly, the Kruppel-like factor (KLFs) family of activators have also been reported to bind CH1 *in vitro* (60). Furthermore, KLFs have been shown to bind to the CH3 domain of p300 (61) and are functionally dependent upon p300 for transcriptional activity (Evans et al. 2007). Our promoter analysis data indicate that KLF factors, in particular KLF4, are essential for *NANOG* promoter activity (Chapter 2). Thus, we speculate that KLFs may bind the CH1 domain of p300 and thereby recruit the coactivator to the DNA to initiate transcription. It is worth noting that no study thus far has explicitly tested for a KLF4-CH1 interaction. We therefore attempted to detect a KLF4-CH1 domain interaction.

To do so, HA-tagged KLF4 and Myc-tagged CH1 were overexpressed in 293T cells and immunoprecipitated to pull down potential binding partners. Unfortunately, no interaction was detected between the two (Fig. 3.10). Activator-coactivator interactions tend to be modest in affinity, and the KLF1-CH1 interaction has a reported K_d of ~3.0 μM (60). It is possible that a KLF4-CH1 complex could be too weak to survive the lysis and buffer conditions during co-immunoprecipitation, and we therefore cannot rule out the possibility of a KLF4-CH1 interaction. Other techniques such as isothermal calorimetry or fluorescence polarization could more definitively address this question.

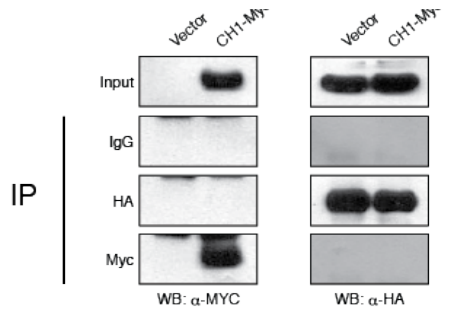

Fig. 3.10. Immunoprecipitation of CH1-Myc and HA-KLF4 in 293T cells. Cells expressing HA-KLF4 were transfected with either Myc-tagged p300/CH1 or an empty vector. 2 days post-transfection, the cells were lysed and probed for potential interactions by immunoprecipitation using HA- or Myc-antibodies.

An important and intriguing observation were the differential responses of HNSCC and NCCIT cells to p300/CH1 domain overexpression. In HNSCC cells, p300/CH1 was sufficient to inhibit NANOG transcriptional activity, tumorsphere formation, and even block xenograft growth (Fig. 3.4D). However, in NCCIT cells, p300/CH1 overexpression had no effect on endogenous gene expression

(Fig. 3.3C). It is currently unclear why these two cell lines display different effects, however there are notable differences in the experiments. For the HNSCC experiments, we generated stable p300/CH1-expressing cell lines, ensuring that all cells in the culture were expressing the domain. This helps to normalize the population by excluding potential contributions from untransfected cels, as well as maintaining p300/CH1 expression over the course of multiple (>7) passages. Thus, the effects observed are relatively long-term compared to the transient transfections done with NCCIT cells. The short-term experiments also underwent a selection step, however this involved the cotransfection of a secondary selectable plasmid. Therefore, we cannot ensure that all surviving cells were cotransfected and thus also expressing p300/CH1. Furthermore, the brief (2-3 day) timecourse of these experiments may mask delays in response due to factors like mRNA/protein turnover rates.

An additional possibility for why p300/CH1 overexpression showed variable effects could be that HNSCC and NCCIT cells in fact have differential requirements for the CH1 domain. For example, HNSCC cells could express relatively lower levels of p300, and thus would be more sensitive to forced expression of the dominant-negative form of CH1. Alternatively, the CH1 domain may be interacting with different partners in HNSCC vs. NCCIT cells, and this could in turn result in a different outcome. What these differences might be, and why they would result in a different effect, are not currently clear and would require further experiments to clarify how broadly the p300/CH1 construct is

expressed in the cellular populations, its lifespan and localization within the cell, and most challengingly, the potential differences in interactions between the cell types. Such insights would provide immense insights into the potential differences in *NANOG* regulatory requirements between cell lines, but are currently outside the scope of the work presented in this thesis. Instead, we chose to focus on how the apparent conserved requirement for the CH1 domain could be utilized for modulating *NANOG* expression.

3.5.2. Small molecule inhibitors of the CH1 domain

The observation that the CH1 domain is crucial for p300 activity at the *NANOG* promoter suggests that small molecule inhibitors of the CH1 domain could be effective at suppressing *NANOG* expression in cancer cells. There are several reported inhibitors of CH1 domain interactions, including the natural product chetomin and related epipolythiodiketopiperazines that disrupt CH1 folding by chelating zinc (62-64), as well as peptidomimetics that compete with activator binding (65). To test the hypothesis that small molecule inhibitors of CH1 could block *NANOG* expression, we treated cells with both chetomin and a hydrogen-bond surrogate peptide designed to mimic HIF-1α (65). Unfortunately, neither compound downregulated *NANOG* expression levels (data not shown), and chetomin showed high cytotoxicity consistent with other reports (63). However, as part of follow up assays, neither compound was able to block DFO-induced *VEGF* expression in positive control experiments, indicating that we were unable to reproduce biological activity with these compounds in our cell models. While

disappointing, these results do not rule out the potential for CH1 domain inhibitors in blocking *NANOG* expression.

There are a number of possible reasons why these small molecules were not active. For the hydrogen-bond surrogate peptide, the lack of activity could likely be explained by poor membrane permeability. Further experiments using a fluorophore-labeled peptide could resolve this question. Additionally, the peptide is designed to resemble the activator HIF1α and bind in the same location on CH1. However, our mutagenesis studies indicate that there is likely a broad interaction interface, and the peptide may either be too small to outcompete such a large interaction, or binding in a noncompetitive site and thus unable to block the interaction. No structures have been reported for a KLF-CH1 interaction, so we do not know at this time whether there is potential overlap between the HIF-1α peptide and KLF binding sites. Future binding assays with HIF-1a and KLF peptides could explore their relative binding modes and determine whether a HIF-1α-inspired peptide could be an effective inhibitor of *NANOG* transcription. Alternatively, the peptidomimetic strategy could be applied to the KLF activation domain as well, however it has yet to be determined if KLF factors adopt an alpha-helical conformation upon binding to the CH1 domain (44).

As for chetomin, the high cytotoxicity seen is unsurprising given its promiscuous zinc-leaching ability. A large number of proteins in the cells rely on zinc as either a catalytic or structural component, and disruption of zinc homeostatsis by

chetomin likely severely stresses the cell. Other analogues like ETP-2 (63) are less toxic and may offer greater selectivity in future experiments. However, ETP-2 still utilizes the same mechanism of action and would likely disrupt similar domains such as the CH3 domain. As a result, chetomin and ETP-2 may not be well suited for use as selective chemical probes of p300/CH1-domain function in *NANOG* expression.

The difficulties encountered with the small molecule studies highlight the need for a broader tool set for probing the p300/CH1 domain. Recently, the small molecules menadione and ethacrynic acid were reported to be sub-micromolar inhibitors of the HIF1α-CH1 interaction (66). These compounds do not appear to chelate metal ions and may therefore inhibit CH1 domain interactions via other mechanisms, suggesting they may be potentially useful in future attempts to probe CH1 function. However, menadione and ethacrynic acid, like chetomin and the hydrogen-bond surrogates, all were designed to inhibit the HIF-1α interaction. Thus, there is still a need for a more diverse catalogue of inhibitors targeting other sites on the CH1 domain. Developing such inhibitors would provide a comprehensive toolkit for probing a variety of CH1 interactions and could be applied to a large number of studies beyond just those described here.

3.5.3. p300 HAT activity is required, but not histone-mediated

In addition to the CH1 domain, we also identified an essential role for p300 HAT activity during *NANOG* transcription. Inhibition using the p300-selective C646

results in significantly decreased expression levels of *NANOG* (Fig. 3.8). This decrease in *NANOG* mRNA coincides with a depletion of H3Kac at the *NANOG* promoter, but not within the gene body of *NANOG*, indicating that p300 HAT activity is confined to the promoter region. Furthermore, C646 does not alter p300 occupancy, suggesting that p300 recruitment is not dependent upon histone recognition by the HAT domain. Taken together, these results would suggest that: i) p300 HAT activity is required at a later step during transcriptional activation, and ii) histone acetylation by p300 is necessary for maintaining high levels of *NANOG* expression.

To explore this latter point, we tested the effects that histone deacetylase inhibitors (HDACIs) have on *NANOG* expression. HDACIs prevent the removal of acetyl groups from histone tails and thereby significantly increase histone acetylation. We reasoned that, if p300-mediated histone acetylation is essential for sustained *NANOG* expression, HDACIs would either enhance or have little effect on *NANOG* levels. Since several HDACIs are FDA-approved as anticancer drugs, such an observation could have significant clinical implications for eliminating cancer-initiating cells during chemotherapy.

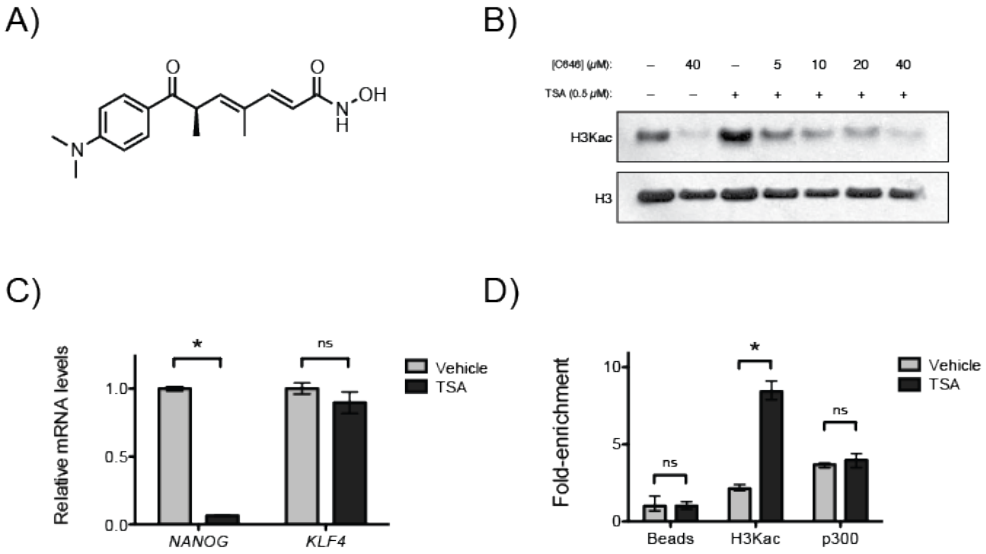

Fig. 3.11. Histone deacetylase inhibitors downregulate *NANOG* expression while retaining high acetylation levels. (A) The HDACI trichostatin A (TSA). (B) Histone H3 acetylation levels in NCCIT cells treated with 0.5 μM TSA and varying concentrations of C646 after 6 h. (C) Relative mRNA levels of *NANOG* and *KLF4* in NCCIT cells after treatment with 0.1 μM TSA for 6 h. (D) Relative enrichment levels of H3Kac and p300 at the *NANOG* promoter in NCCIT cells after 6 h treatment with 0.1 μM TSA. Data ishown as mean ± SD (n = 3). *$p < 0.05$ vs. vehicle control (1% DMSO v/v).

When the HDACI trichostatin A (TSA)(Fig. 3.11A) was applied to NCCIT cells though, *NANOG* mRNA levels were significantly reduced after just 6 h (Fig. 3.11B). Furthermore, TSA strongly enriched for H3Kac levels both globally (Fig. 3.11B-C), and at the *NANOG* promoter (Fig. 3.11D). Similar results were observed with the HDACIs SAHA and the class II-selective PD106 (data not shown). These results would suggest that TSA inhibits *NANOG* expression even though histone acetylation levels increase, in contrast to our initial hypothesis. One possible explanation could be that histone hyperacetylation disrupts expression at the *NANOG* locus, such as through excessive chromatin relaxation and spurious transcriptional initiation. Alternatively, if histone acetylation is not a key determinant in *NANOG* transcription but p300 HAT activity is required, then p300 may be acetylating a non-histone substrate and that acetylated protein is

required for maintaining *NANOG* mRNA levels. The possibility of a non-histone substrate was intriguing and we chose to first explore this potential mechanism.

p300 has a number of non-histone substrates, including transcriptional activators. Interestingly, p300 is known to acetylate KLF4 at K225 and K229, and p300-mediated acetylation is essential for KLF4 transcriptional potential (48). Given our earlier data indicating KLF4 is an essential activator of *NANOG* (Chapter 2), we sought to determine if p300-mediated KLF4 acetylation was required for *NANOG* expression. However, when the acetylation sites were mutated to arginine (K225R and K229R) no significant difference in activation potential was observed (Fig. 3.12A,B), suggesting that KLF4 may not be the necessary substrate of p300 HAT activity. Identifying this protein would likely require unbiased mass spectrometry screening to look for changes in acetylated proteins following C646 treatment (67).

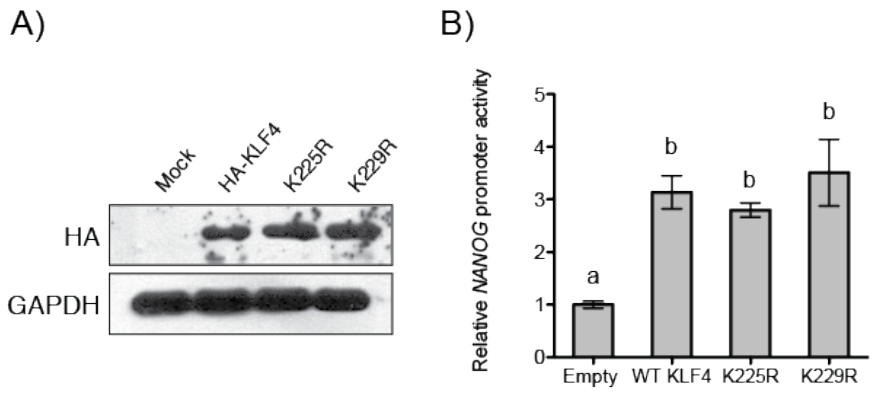

Fig. 3.12. Acetylation is not required for KLF4-mediated *NANOG* promoter activity enhancement. (A) WT KLF4 and K225R and K229R expressed in NCCIT cells. (B) *NANOG* promoter activity in NCCIT cells. Data is shown as mean ± S (n = 4). Statistical groups were determined by ANOVA followed by Tukey's post-hoc test.

3.5.4. A broader model of NANOG transcriptional activation

Take together, our results provide the first systematic interrogation of p300 functions for *NANOG* expression in cancer cells. When combined with our promoter analysis and activator identification data, we can begin to assemble a more comprehensive model for how transcriptional initiation occurs at the *NANOG* promoter (Fig. 3.13). Based on our results, we propose that the transcriptional activator KLF4, or related KLF family members, bind to the proximal promoter region just upstream of the transcriptional start site. Once localized to the DNA, KLF4 activates transcription of *NANOG*. KLF4, like many activators, may recruit coactivators like p300. p300 is localized to the *NANOG* promoter through an as yet unknown series of interactions, but one possible explanation would be via CH1 domain interactions. These interactions are not well characterized yet, but it would appear to be dependent upon the structural integrity of the entire CH1 domain, and thus the proteins may form extensive contacts with one another.

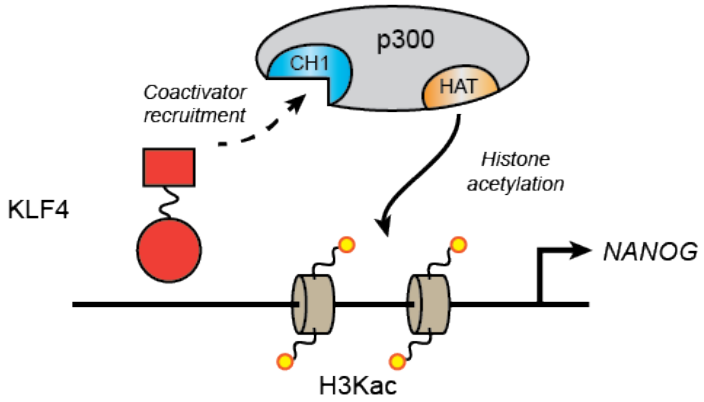

Fig. 3.13: Proposed model of transcriptional activation during *NANOG* expression. First, the activator KLF4 (red) binds the NANOG promoter and initiates recruitment of coregulatory proteins. One such coregulator is the coactivator p300 coactivator (grey). The CH1 domain of p300 (blue) appears to be essential for p300-mediated promoter activity. The CH1 domain may act as a docking interface for p300 recruitment. Once localized to the *NANOG* promoter, p300's HAT domain acetylates local histones (yellow circles) and promotes high levels of transcription.

Once localized to the *NANOG* promoter, the p300 HAT domain will then acetylate histone N-terminal tails. Histone acetylation is associated with a more relaxed chromatin structure and may thereby facilitate transcription. However, histone acetylation by p300 is not the sole factor since *NANOG* can be downregulated even when H3Kac levels are high. Therefore, p300 may also be acetylating some other protein(s) that are also essential for full transcriptional activity, but for now remain unidentified. These unknown proteins must operate after p300 recruitment since p300 promoter occupancy is not sufficient to stimulate transcription. It may be that p300 HAT activity is required to stabilize a larger "enhanceosome"-like transcriptional complex (68). p300 is known to participate in large multiprotein complexes (69-70), and it would be further

illuminating to identify other transcriptional cofactors involved in *NANOG* regulation. Already, the results shown here reveal two potential opportunities for modulating *NANOG* levels just by targeting the p300 coactivator: i) blocking CH1 domain interactions and ii) inhibiting p300 HAT activity. Such information could lead to the design of more selective combinatorial strategies to inhibit *NANOG* expression in cancer, or modulate its expression in cellular reprogramming efforts and regenerative medicine.

3.6. Acknowledgements

I would like to acknowledge the contributions of our collaborators Prof. Quintin Pan (Ohio State University), Dr. Xiujie Xie, and Dr. Manchao Zang for their work generating the p300/CH1-overexpressing HNSCC cell lines as well as the *in vitro* and *in vivo* phenotypic assays using those lines.

3.7. Materials and Methods

Cell lines

NCCIT, HeLa, and 293T cells were purchased from ATCC. NCCIT cells were grown in RPMI media 1640 (Life Technologies, 11875) supplemented with 10% fetal bovine serum (Atlanta Biologicals, S11195) in a 37 ºC incubator with 5% CO_2. All other cell lines were cultured in Dulbecco's Modified Eage Medium (DMEM) (Invitrogen, 11965) supplemented with 10% fetal bovine serum.

For C646 and TSA treatments, compounds were dissolved in DMSO and administered to the cells at varying concentrations. C646 appeared to have a short half-life in the cells as indicated by a cell-dependent loss of color. Therefore, fresh compound was applied every 24 h for longer time courses. Changes in cellular viability were measured using the WST-1 assay and data was normalized against the DMSO vehicle control.

For gene expression experiments, cellular RNA was harvested using the RNeasy kit (Qiagen). 1 µg RNA was used as a template for reverse transcriptase (iScript cDNA Synthesis Kit, Bio-Rad) following the manufacturer's instructions. 1 µl cDNA was used for quantitative PCR reactions using Taqman probes for *NANOG*, *KLF4*, and *GAPDH* as an internal control. Data was normalized using the $\Delta\Delta$Ct method against *GAPDH* levels.

Cell proliferation was assessed using the MTT reagent (Roche Molecular Biochemicals, Nutley, NJ, USA) to detect metabolic active cells. Absorbance was measured at 570nm in the Spectra Max 190 ELISA reader (Molecular Devices, Sunnyvale, CA, USA) after overnight incubation. For clonogenic survival, 300 cells per well were plated in complete growth media and allowed to grow until visible colonies were formed (14 days). Cell colonies were fixed with 10% cold methanol, stained with 0.1% crystal violet in 20% methanol, washed and air dried. The soft agar assay was performed in 35mm plates containing two layers of Agar (Invitrogen). The bottom layer consisted of 0.6% agar in 1.5 ml of DMEM

with 10% FBS. Cells were dissociated and placed (5×10^3/well) in the top layer containing 0.3% agar in the same medium as the bottom. Cells were cultured for three weeks and colonies were photographed under a microscope and measured NIS-Elements software (Nikon, Melville, NY, USA). Colonies with diameters larger than 80 μm were counted.

Tumorsphere formation

Cells were collected and seeded in a serum-free defined medium consisting of keratinocyte serum-free medium supplemented with epidermal growth factor, basic fibroblast growth factor, insulin and hydrocortisone in low-attachment plates (Corning Incorporated, Corning, NY, USA) for tumorspheres. Tumorsphere-formation efficiency was calculated as the number of tumorspheres (≥50 μm in diameter) formed in 7 days divided by the original number of cells seeded. Tumorsphere diameter was measured using NIS-Elements software.

Tumor incidence and growth in athymic nude mice

UMSCC47/CH1 cells were suspended in 50:50 Dulbecco's modified Eagle's Medium : Matrigel and implanted subcutaneously into the left and right flanks of 6-week-old athymic nude mice (eight mice/group), respectively. After 3 weeks, tumors were measured once a week using a digital caliper, and tumor volumes were calculated using the formula d1 x d2 x d3 x 0.5236, where 'd' represents the three orthogonal diameters. Tumor growth and incidence were monitored for 49 days following tumor cell implantation.

Chromatin immunoprecipitation assays

ChIP assays were performed on >90% confluent cells in 10-cm plates using the Abcam ChIP (ab-500) protocol.. Briefly, cells were dissociated by trypsin digest and resuspended in media + 10% FBS before counting. 10 x 10^6 cells were then pelleted by centrifugation and then washed with chilled PBS. Cells were then fixed with 1% formaldehyde for 10 min at room temperature before quenching with 1.25 M glycine. The fixed cells were washed with chilled PBS and resuspended in lysis buffer. Samples were then pelleted to collect the fixed chromatin. Mild sonication was carried out to yield sheared chromatin averaging ~500 bp in size. Immunoprecipitations were conducted on the sheared chromatin using anti-histone H3 (Abcam, ab1791), anti-p300 (Santa Cruz, sc-584 X), anti-acetylated histone H3 antibody (Millipore, 06-599), or Protein A beads alone as a negative control. The recovered DNA was used as a template for qPCR reactions using SYBR Green (Applied Biosystems, 4309155) according to the manufacturer's instructions.

Plasmids and cloning

The *p300* and KLF4 plasmids were originally purchased from Addgene (pCMVb-p300, plasmid 10717; pcDNA3.1-HA-KLF4 FL, plasmid 34593). Deletions and point mutations were then introduced using site-directed mutagenesis. Mutants were validated by DNA sequencing.

Transfections

NCCIT cells were transfected using OptiMEM media (Invitrogen) and Lipofectamine LTX (Invitrogen) according to an optimized protocol. Briefly, DNA•liposome complexes were formed in OptiMEM by sequentially adding DNA, PLUS reagent, and LTX and incubating for 30 min at room temperature. Four volumes of RPMI (10% FBS) was then added to the transfection mix before applying to the cells. The cells were incubated with the transfection mix for 24 h before switching media and grown for an addition 24 h.

Luciferase assays

10,000 cells were plated in 96-well format with 100 µl media + 10% FBS. After overnight incubation, the media was replaced with transfection mix: 100 µL media with 5 ng/ml *Renilla* luciferase plasmid (pRLSV40) and 395 ng/ml firefly luciferase plasmid. After 48 h incubation, luciferase activity is measured on a Berthold luminometer with firefly substrate from the Dual-Luciferase Assay Reporter System (Promega, E1980).

Co-immunoprecipitation assays

293T cells were transiently transfected by lipofectamine with HA-KLF4 and p300/CH1-Myc or empty vector for 3 hr according to the manufacturer's protocol. After transfection, cells were incubated overnight in DMEM (+10% FBS). Cells were then washed with media followed by three washes with 10 ml ice cold PBS and collected by scraping and centrifugation. A modified RIPA buffer (1% NP-40,

0.1% SDS, 0.5% Na deoxycholate, and protease inhibitor cocktail) at a volume of 400 µl/10 cm dish was then used to lyse the cells. The cellular extracts were divided into fractions and aliquots were taken to serve as input controls. Further aliquots were taken and immunoprecipitated using anti-Myc (Santa Cruz, sc-40), anti-hemagglutinin (sc-7392), and anti-IgG (sc-2025, sc-2027) antibodies. Sixty microliters of a 50% protein G-sepharose solution were then added and the incubation continued at 4 ºC for 1 h. The beads were collected by centrifugation and washed three times with 1 ml modified RIPA buffer before being resuspended in sample buffer for electrophoresis.

3.8. References

1. Guarente L. Transcriptional coactivators in yeast and beyond. *Trends Biochem. Sci.* **20**, 517-521 (1995).
2. Ptashne M & Gann A. Transcriptional activation by recruitment. *Nature* **386**, 569-577 (1997).
3. Thomas MC & Chiang CM. The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41**, 105-175 (2006).
4. Dynlacht BD, Hoey T, Tjian R. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66**, 563-576 (1991).
5. Kim YJ, *et al.* A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**, 599-608 (1994).
6. Naar AM, Lemon BD, Tjian R. Transcriptional coactivator complexes. *Annu. Rev. Biochem.* **70**, 475-501 (2001).
7. Bednar J, *et al*. Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14173-14178 (1998).
8. Bannister AJ & Kouzarides T. The CBP co-activator is a histone acetyltransferase. *Nature* **384**, 641-643 (1996).
9. Ogryzko VV, *et al.* The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953-959 (1996).
10. Shiama N. The p300/CBP family: integrating signals with transcription factors and chromatin. *Trends Cell Biol.* **7**, 230-236 (1997).
11. Bhattacharya S, *et al*. Functional role of p35srj, a novel p300/CBP binding protein, during transactivation by HIF-1. *Genes Dev.* **13**, 64-75 (1999).
12. De Guzman RN, *et al*. Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP. *J. Mol. Biol.* **303**, 243-253 (2000).
13. Parker D, *et al*. Phosphorylation of CREB at Ser-133 induces complex formation with CREB-binding protein via a direct mechanism. *Mol. Cell Biol.* **15**, 694-703 (1996).
14. Albanese C, *et al*. Activation of the cyclin D1 gene by the E1A-associated protein p300 through AP-1 inhibits cellular apoptosis. *J. Biol. Chem.* **274**, 34186-34195 (1999).
15. Lin CH, *et al*. A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies. *Mol. Cell* **8**, 581-590 (2001).
16. Demarest SJ, *et al*. Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549-553 (2002).

17. Nagy Z & Tora L. Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26**, 5341-5357 (2007).
18. Chrivia JC, *et al*. Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365**, 855-859 (1993).
19. Radhakrishnan I, *et al*. Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell* **91**, 741-752 (1997).
20. Sugase K, *et al*. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021-1025 (2007).
21. Zor T, *et al.* Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. *J. Mol. Biol.* **337**, 521-534 (2004).
22. Zor T, *et al*. Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators. *J. Biol. Chem.* **277**, 42241-42248 (2002).
23. Goto NK, *et al*. Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. *J. Biol. Chem.* **277**, 43168-43174 (2002).
24. De Guzman RN, *et al.* Structural basis for cooperative transcription factor binding to the CBP coactivator *J. Mol. Biol.* **355**, 1005-1013 (2006).
25. Vendel AC, *et al*. KIX-mediated assembly of the CBP-CREB-HTLV-1 Tax coactivator-activator complex. *Biochem.* **42**, 12481-12487 (2003).
26. Lee CW, *et al*. Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP. *Biochem.* **48**, 2115-2124 (2009).
27. Wang F, *et al*. Structures of KIX domain of CBP in complex with two FOXO3a transactivation domains reveal promiscuity and plasticity in coactivator recruitment. *Proc. Natl. Acad. Sci. USA* **109**, 6078-6083 (2012).
28. Campbell KM & Lumb KJ. Structurall distinct modes of recognition of the KIX domain by Jun and CREB. *Biochemistry* **41**, 13956-13964 (2002).
29. De Guzman RN, *et al.* CBP/p300 TAZ domain forms a structured scaffold for ligand binding. *Biochemistry* **44**, 490-497 (2005).
30. Miller M, *et al*. Structure of the Taz2 domain of p300: insights into ligand binding. *Acta. Crystallogr. D Biol. Crystallogr.* **65**, 1301-1308 (2009).
31. Teufel DP, et al. Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53. *Proc. Natl. Acad. Sci. USA* **104**, 7009-7014 (2007).
32. Ferreon JC, *et al*. Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2. *Proc. Natl. Acad. Sci. USA* **106**, 6591-6596 (2009).
33. Ruas JL, *et al*. Role of CBP in regulating HIF-1-mediated activation of transcription. *J. Cell Sci.* **118**, 301-311 (2005).
34. Dames SA, *et al*. Structural basis for Hif-1 alpha/CBP recognition in the cellular hypoxic response. *Proc. Natl. Acad. Sci. USA* **99**, 5271-5276 (2002).

35. Yin Z, *et al.* The essential role of Cited2, a negative regulator for HIF-1alpha, in heart development and neurulation. *Proc. Natl. Acad. Sci. USA* **99**, 10488-19493 (2002).
36. Freedman SJ, *et al*. Structural basis for negative regulation of hypoxia-inducible factor-1alpha by CITED2. *Nat. Struct. Biol.* **10**, 504-512 (2003).
37. De Guzman RN, *et al.* Structural basis for cooperative transcription factor binding to the CBP coactivator *J. Mol. Biol.* **355**, 1005-1013 (2006).
38. De Guzman RN, *et al.* Interaction of the TAZ1 domain of the CREB-binding protein with the activation domain of CITED2: regulation by competition between intrinsically unstructured ligand for non-identical binding sites. *J. Biol. Chem.* **279**, 3042-3049 (2004).
39. Yao TP, *et al*. Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93**, 361-372 (1998).
40. Pasqualucci L, *et al*. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* **471**, 189-195 (2011).
41. Nettles KW, *et al*. CBP is a dosage-dependent regulator of nuclear factor-kappa B suppression by the estrogen receptor. *Mol. Endocrinol.*
42. Kimbrel EA, *et al*. Systematic in vivo structure-function analysis of p300 in hematopoiesis. *Blood* **114**, 4804-4812 (2009).
43. Ruas JL, *et al*. Complex regulation of the transactivation function of hypoxia-inducible factor alpha by direct interaction with two distinct domains of the CREB-binding protein/p300. *J. Biol. Chem.* **285**, 2601-2609 (2010).
44. Newton AL, *et al.* The transactivation domain within cysteine-histidine-rich region 1 of CBP comprises two novel zinc-binding modules. *J. Biol. Chem*. **275**, 15128-15134 (2000).
45. Schiltz RL, *et al.* Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates. *J. Biol. Chem.* **274**, 1189-1192 (1999).
46. Gu W, Shi XL, Roeder RG. Synergistic activation of transcription by CBP and p53. *Nature* **387**, 819-823 (1997).
47. Ma K, Chan JK, Zhu G, Wu Z. Myocyte enhancer factor 2 acetylation by p300 enhances its DNA binding activity, and myogenic differentiation. *Mol. Cell Biol.* **25**, 3575-3582 (2005).
48. Evans PM, *et al.* Kruppel-like factor 4 is acetylated by p300 and regulates gene transcription via modulation of histone acetylation. *J. Biol. Chem*. **47**, 33994-34002 (2007).
49. Baltus GA, *et al*. Acetylation of Sox2 induces its nuclear export. *Stem Cells* **27**, 2175-2184 (2009).
50. Ma K, Chan JK, Zhu G, Wu Z. Myocyte enhancer factor 2 acetylation by p300 enhances its DNA binding activity, and myogenic differentiation. *Mol. Cell Biol.* **25**, 3575-3582 (2005).
51. Balasubramanyam K, *et al*. Curcumin, a novel p300/CREB-binding protein-specific inhibitor of acetyltransferase, represses the acetylation of

histone/nonhistone proteins and histone acetyltransferase-dependent chromatin transcription. *J. Biol. Chem.* **279**, 51163-51171 (2004).
52. Balasubramanyam K, *et al*. Polyisoprenylated benzhophenone, garcinal, a natural histone acetyltransferase inhibitor, represses chromatin transcription and alters global gene expression. *J. Biol. Chem*. **279**, 33716-33726 (2004).
53. Lau OD, *et al.* HATs off: selective synthetic inhibitors of the histone acetyltransferases p300 and PCAF. *Mol. Cell* **5**, 589-595 (2000).
54. Sagar V, *et al.* Bisubstrate analogue structure-activity relationships for p300 histone acetyltransferase inhibitors. *Bioorg. Med. Chem.* **12**, 3383-3390 (2004).
55. Mantelingu K, *et al*. Specific inhibition of p300-HAT alters global gene expression and represses HIV replication. *Chem. Biol.* **14**, 645-657 (2007).
56. Dekker FJ & Jaisma HJ. Histone acetyltransferases as emerging drug targets. *Drug Discovery Today* **14**, 942-948 (2009).
57. Bowers EM, *et al*. Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. *Chem. Biol.* **17**, 471-482 (2010).
58. Wang Z, *et al.* Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019- 1031 (2009).
59. Goodman RH & Smolik S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **14**, 1553-1557 (2000).
60. Mas C, *et al*. Structural and functional characterization of an atypical activation domain in erythroid Kruppel-like factor (EKLF). *Proc. Natl. Acad. Sci. USA* **108**, 10484-10489 (2011).
61. Geiman DE, *et al.* Transactivation and growth suppression by the gut-enriched Kruppel-like factor (Kruppel-like factor 4) are dependent on acidic amino acid residues and protein-protein interaction. *Nucleic Acids Res.* **28**, 1106-1113 (2000).
62. Kung AL, *et al.* Small molecule blockade of transcriptional coactivation of the hypoxia-inducible factor pathway. *Cancer Cell* **6**, 33-43 (2004).
63. Kushal S, *et al*. Inhibition of hypoxia-inducible transcription factor with designed epipolythiodiketopiperazine. *Biopolymers* **95**, 8-16 (2011).
64. Cook KM, *et al*. Epidithiodiketopiperazines block the interaction between hypoxia-inducible factor-1 alpha (HIF-1alpha) and p300 by a zinc ejection mechanism. *J. Biol. Chem.* **284**, 26831-26838 (2009).
65. Henchey LK, *et al*. Inhibition of hypoxia inducible factor 1-transcription coactivator interaction by a hydrogen bond surrogate alpha-helix. *J. Am. Chem. Soc.* **132**, 941-943 (2010).
66. Na YR, *et al*. Menadione and ethacrynic acid inhibit the hypoxia-inducible (HIF) pathway by disrupting HIF-1alpha interaction with p300. *Biochem. Biophys. Res. Commun.* **434**, 879-884 (2013).

67. Yang YY, *et al.* Identification of lysine acetyltransferase p300 substrates using 4-pentynoyl-coenzyme A and bioorthogonal proteomics. *Bioorg. Med. Chem. Lett.* **21**, 4976-4979 (2011).
68. Chambers I & Tomlinson SR. The transcriptional foundation of pluripotency. *Development* **136**, 2311-2322 (2009).
69. Kamei Y, *et al*. A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* **85**, 403-414 (1996).
70. Chen HW, *et al.* Nuclear receptor coactivato ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell* **90**, 569-580 (1997).

Chapter 4

Expanding the toolkit for targeting cancer-initiating cells through chemical biology

4.1. Abstract

Cancer-initiating cells (CICs) are an unstable and poorly understood subpopulation of cancer cells that also represent an incredible opportunity for new therapeutic developments. New technologies and tools are necessary to overcome some of the inherent challenges associated with studying CICs *in vitro* in order to deliver on the promise of CIC-targeting therapies. Here we discuss important areas of CIC biology that are in need of new tools and describe ways chemical tools could address those needs.

4.2. Challenges involved in studying cancer-initiating cells

The model of cancer-initiating cells (CICs) in cancer biology has generated a tremendous amount of new theories into how tumors originate, respond to therapeutic strategies, and how to effectively combat the disease (1-3). These theories will require rigorous testing in order to usher in the next wave of medical advances. However, CICs possess multiple challenges for their study. First, CICs are frequently found in heterogeneous cell populations (1, 4) and therefore

require methods to identify and isolate CICs from non-CICs (5). Second, CICs are multipotent by definition and are capable of spontaneous differentiation. This intrinsic instability makes CICs difficult to control and maintain. Lastly, very little is known about how CICs are controlled in terms of both extra- and intracellular regulatory programs. This information is essential for understanding the underpinning biology of CICs, as well as directing future drug discovery strategies.

In order to address the challenges outlined above, CIC biology requires new tools for probing the inner workings of a rare, unstable cell type. Chemical approaches and small molecules have provided unique ways of observing and manipulating biological systems, and as such are poised to provide the necessary tools for studying CICs. This chapter identifies areas in CIC biology that would greatly benefit from new techniques, as well as possible strategies for providing such tools.

4.3. Protein-protein interactions in CIC-regulatory networks

4.3.1. Strategies for examining the global interaction networks

The unique transcriptional and signaling networks that define CICs are among the most intriguing and sought after targets in CIC biology. Such interaction networks feature prominently in nearly every biological process, but remain difficult to study due to their complexity. For example, transcriptional networks regulating self-renewal can involve dozens of proteins driving the expression, or

suppression, of hundreds of genes (6-8). To further complicate the network, most of these proteins can engage in multiple protein-protein interactions (PPIs). As a result, identifying the key interactions can be challenging due to the sheer complexity of the networks.

Other stem cell model systems have tackled questions about protein-protein and transcriptional networks through the use of affinity-based capture methods in combination with mass spectrometry to identify interacting proteins. For instance, Wang *et al*. (8) described an interaction network for the transcriptional activators Oct4 and Nanog in mouse embryonic stem cells (ESCs). To do this, the ESCs were made to express low levels of exogenous *Oct4* and *Nanog* that bore an N-terminal peptide substrate that could be biotinylated *in cellulo*. The biotinylated activators could then be purified along with proteins that likely interact with Oct4 and Nanong under physiological conditions. Mass spectrometry-based identification of the associated proteins generated a map of PPIs for Oct4 and Nanog in ESCs, and the network overall was found to be significantly enriched in proteins that regulate ESC identity. Moreover, the network members also appeared to be colocalize at ESC-specific genes in global chromatin immunoprecipitation (ChIP) studies (7, 9), suggesting the Oct4/Nanog-PPI network is functionally cooperating to regulate ESC pluripotency and self-renewal.

CICs can express many of the same proteins as ESCs (e.g. *NANOG*, *OCT4*) (10-12), yet it is unclear whether the same PPIs occur. While CICs and stem cells may share features like self-renewal, the differences in cellular context can translate into very different sets of expressed proteins and regulatory signals that can greatly alter the types of interactions NANOG and OCT4 engage in. For example, NANOG has been reported to be phosphorylated by protein kinase Cε (PKCε) (13, 14). PKCε is frequently mutated or overexpressed in cancer leading to elevated kinase activity (15, 16) and these altered phosphorylation levels could easily affect NANOG interactions (17). Therefore, due to concerns about the potential differences between cell types, global PPI and ChIP studies should be revisited in CICs to determine just how much overlap still exists in the PPI partners for proteins of interest in CIC and other stem cell models.

4.3.2. Strategies for examining local interactions

While unbiased and informative, global studies like those described above (Chapter 4.3.1) suffer from lack of information about the specific PPIs. Global, low-resolution pull-down experiments only identify the presence of an interaction between proteins; these interactions could be direct or indirect, mono- or multivalent, low- or high-affinity, etc. These sorts of details need to be determined in order to design efficient probes of PPIs in CICs, especially for multifunctional proteins like p300.

The respective functions of p300 have been studied in other stem cell systems using mutagenesis strategies (18). Specifically, mouse hematopoietic stem cells (HSCs) expressing an allelic series of p300 domain mutants revealed that the CH1 and KIX domains are necessary for HSC engraftment in the bone marrow niche and for proper hematopoietic differentiation (18). Deletion of other domains, or mutational inactivation of the histone acetyltransferase (HAT) domain had little effect on hematopoiesis.

Similarly, the work presented in this thesis utilized a series of complementary mutational and small molecule antagonists to parse out the relevant functions of p300 during *NANOG* expression in cancer cells (Chapter 2.3-4, 3.3-4). Our results indicate that the CH1 and HAT domains are both essential for *NANOG* expression and therefore may be strong candidates for future inhibition strategies. Importantly, p300 can be viewed as a model for other multidomain proteins. Therefore, the strategies described here to dissect multifunctional proteins can be applied to other targets in CIC biology.

4.4. CIC cultivation and external interactions

4.4.1. Engineering microenvironments for CICs *in vitro*

One of the main challenges in the study and clinical application of CICs is that although these cells are multipotent and highly self-renewing *in vivo*, they are difficult to maintain and expand *in vitro* in homogeneous populations. This can

complicate experiments since any change or effect must be evaluated within a mixed sample of cells. Several groups have circumvented this hurdle by generating their own *in vitro* model for CICs. Gupta *et al*. (4) induced an artificial epithelial-mesenchymal transition (EMT) state to create a more CIC-like cell state. Similarly, Sachlos *et al.* (19) used a tumorigenic variant of normal stem cells to form a more stable, uniform cell population. These studies have the important caveat that the resulting cultures only resemble CICs; they are not, in fact, CICs identified from tumors. However, the results from such artificial systems indicate that stable CIC cultures could greatly improve efforts to study CICs and aid drug discovery. The challenge for the field going forward, therefore, is developing methods for isolating, expanding, and cultivating stable CIC populations *in vitro*.

One of the most promising techniques for *in vitro* CIC cultivation is through the use of defined culture conditions. Indeed, defined soluble factors such as bFGF, EGF, and insulin are already in use for standard tumorsphere formation assays that measure clonogenicity and self-renewal *in vitro*. However, soluble signals are only part of the total inputs a CIC receives from the external environment that support CIC maintenance. Another, perhaps more important input comes from the extracellular matrix (ECM). The ECM consists of a complex mixture of proteins, proteoglycans, and polysaccharides that provide adhesive interactions essential for stem cell- and CIC-maintenance (20, 21). The ECM helps to define a supportive microenvironment and this microenvironment can, in theory, be

recreated *in vitro* provided key factors are identified. Most substrata currently used in cancer cell line and CIC cultures are comprised of mixtures of extracellular matrix proteins derived from animal sources, such as collagen or Matrigel (22, 23). As a result, these products often suffer from variability between batches and present issues with reproducibility. Additionally, the complex mixture of proteins and signaling factors present in these substrates make it more difficult to understand and control the cell-substrate interactions. The key to engineering a suitable CIC niche *in vitro* therefore resides in first identifying the components that precisely control CIC identity.

To the best of our knowledge, no groups have yet defined an *in vitro* niche for CICs, but recent advances in artificial microenvironments for other stem cells can provide a template for future studies. For example, Villa-Diaz *et al*. (24) reported that acrylate polymers could support ES cell cultures in defined media, opening up possibilities for future experiments to examine physicochemical properties necessary for maintaining a stem cell niche including ionic charge and mechanical stiffness (25). The relative ease with which these polymeric substrates can be synthesized and modified would not only facilitate the study of ES cell-niche factors, but could also be transferred to additional cell types including CICs.

While polymer-based substrates can support stem cell cultivation, they may not reveal new binding interactions or requirements supplied for the niche due to lack

of identified receptors for most polymers. In order to explore interactions that better reflect the endogenous ECM-stem cell interactions, Kiessling and colleagues (26, 27) selectively cultivated pluripotent embryonic stem (ES) cells on defined peptide substrate monolayers. The authors chose to synthesize an array of laminin-derived peptides presented on a gold substratum by coupling the variable peptide components to a self-assembling alkanethiol moiety (26). These monolayers could be spatially patterned to facilitate screening for peptides capable of supporting ES cell adherence and growth in the absence of MEF feeder layers used in typical ES cell culture. The peptides identified in this screen were able to maintain ES cell pluripotency over the course of the screen (5-7 d) and, impressively, were comparable to Matrigel in terms of supporting ES cell proliferation. Furthermore, the active peptides help to identify critical domains and regions within the large laminin ECM proteins that mediate ES cell maintenance. In addition to laminin-derived peptides, the Kiessling group demonstrated that heparin-binding peptides were also capable of supporting long-term *in vitro* ES cell cultivation (27).

Taken together, these studies provide proof-of-principle that multipotent stem cells can be maintained and manipulated *in vitro* using defined microenvironments. ES cells have been the most popular objects of study with these experiments, but similar artificial niches have been described for a variety of adult stem cell types (reviewed in ref. 28). As a result, it is reasonable to predict that the strategies employed in studying stem cell microenvironments

could be applied to CICs, both to aid cultivation and to generate new insights into their behavior *in vivo*.

4.4.2. Screening for substrates for CICs *in vitro*

The development of a supportive CIC-substrate, such as through the use of defined peptide substrates similar to the strategies described above, would first require screening for matrix molecules (e.g. polymer, peptide, hydrogel, etc.) suitable for maintaining CICs. Fortunately, the synthesis of these molecules is fairly straightforward and amenable to modification and diversification, allowing potentially large libraries of candidate molecules to be screened. Such screens would require extensive controls in place to ensure that the cells growing on the "hit" substrates are: i) identified or putative CICs, ii) viable and propagating, and iii) multipotent. This is important since the primary goal of the screen would be to distinguish between substrates that support CICs as opposed to allowed nonspecific attachment of more differentiated cancer cells. CICs have a variety of markers and/or assays established for measuring these qualities (e.g. ALDH1-staining, cell-surface markers, etc.) that can be incorporated into screens as potential readouts.

A reasonable outline for a CIC-substrate screen, therefore, would begin first with the synthesis of a library of potential substrates followed by immobilizing or printing the library onto a high throughput-compatible platform such as glass slides or multiwall dishes (Fig. 4.1.A-B). The substrates could then be incubated

with dissociated CICs that were isolated through techniques such as CIC enrichment assays (e.g. tumorspheres, fluorescence-activated cell sorting)(Fig. 4.1.B). A simple counterscreen would be to run a parallel plate using the non-

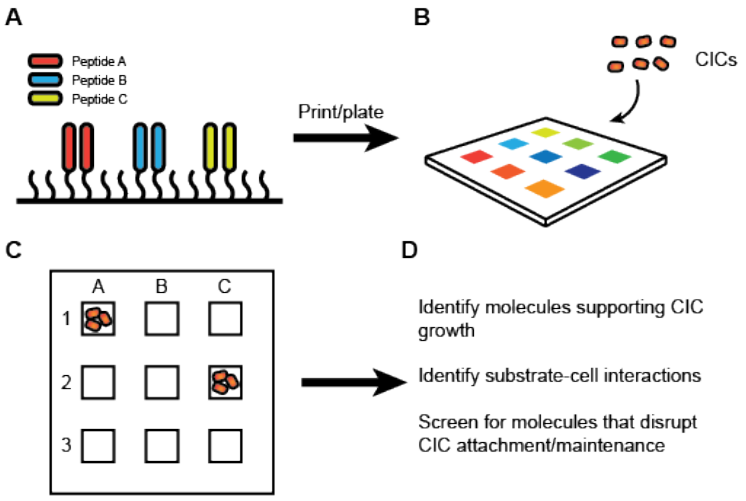

Fig. 4.1. Outline for screening extracellular matrix molecules that support CIC growth. (A) Molecules of interest can be synthesized and immobilized on a solid surface. (B) The library of molecules are then spatially arranged to create a multiplex screening platform, either in wells or across a slide. The library is then exposed to supsended CICs. (C) Molecules that support CIC attachment and growth will retain cells following washing and controls to confirm the propagating cells are, in fact, CICs. (D) Once hit molecules are identified, new experiments can begin including: i) further rounds of modification and optimization on the inital leads; ii) discovery of unique CIC-niche interactions; and iii) developing probes and tools to disrupt CIC-niche interactions.

CIC populations to compare for selective attachment. After several days, the plates could be inspected for potential expansion and staining for CIC markers such as CD133 (Fig. 4.1.C-D). Candidate substrates could then be identified and subjected to further rounds of testing and modification to improve the efficiency of CIC attachment and propagation.

Identifying reliable and robust substrates for *in vitro* cultivation would promote CIC research across the board by increasing the accessibility to a challenging

cell type to study in isolation. Additionally, knowing what substrates foster CIC growth can reveal new information into CIC biology through the identification of cell surface receptors and signaling factors responsible for regulating CIC self-renewal and identity (25).

4.4.3. Substrate platforms in probe and drug discovery screens

A robust *in vitro* platform for studying the CIC microenvironment could also open up exciting new possibilities for small molecule discovery. To date, that majority of small molecules described as CIC inhibitors target intracellular signaling pathways associated with CIC biology (Table 1.1). However, if essential ECM-CIC interactions are identified then additional CIC-targeting strategies can be pursued, such as through the use of antibodies to mask potential niche sites or the development of small molecule ECM-receptor inhibitors to antagonize critical extracellular signals. These molecules could be used to inhibit ongoing niche signals and thereby induce CIC differentiation and exhaustion, or act as a blockade against CIC engraftment during metastasis. Consequently, these alternative strategies could strongly complement current CIC probe and drug discovery methods.

4.5. CIC identification

Since CICs are often a minor subpopulation within the heterogeneous population of cells present in a tumor, a major challenge to CIC research remains the ability to identify and separate CICs from non-CICs. CIC identification relies upon two

categories of defining properties: CIC-specific phenotypes and CIC-specific markers. CIC-specific phenotypes such as self-renewal and multipotency are assay-dependent and rely on distinguishing features like CIC self-renewal in tumorsphere assays. These assays do not always allow for prospective identification of CICs, but benefit from working with viable cells.

On the other hand, CIC-specific markers such as uniquely expressed proteins permit prospective CIC identification, but may compromise viability. Conjugated antibodies are the most frequently used method for labeling CICs thanks to their sensitivity and potential selectivity. However, unless the CICs are permeabilized, antibodies are unable to penetrate the cell membrane and label intracellular markers. Therefore, antibody-based identification of viable CICs is restricted to cell-surface proteins such as CD44 and CD133 (3). Cell-surface markers are often only correlated with CICs and single labels are often not sufficient to accurately identify the CIC subpopulation (29, 30). Thus, additional markers are desirable for high-confidence CIC identification.

Intracellular markers can provide a more reliable indicator of CIC-associated processes. For example, transcription factors such as OCT4 and NANOG regulate self-renewal in CICs and their presence is strongly associated with CIC identity (31). Enzymatic markers such as aldehyde dehydrogenase 1 (ALDH1) are also valuable intracellular markers of CICs in a variety of cancer types (3, 32, 33). ALDH1 is believed to act in CICs as a cytoprotective enzyme that detoxifies

harmful agents and thereby helps to maintain CIC identity in the presence of stress- or differentiating-inducing signals. ALDH1 activity in particular can be determined within living cells and is often used to sort viable CICs in the absence of well-defined cell surface markers (3). ALDH1 enzymatic activity can be measured using the substrate BODIPY-aminoacetaldehyde that is converted into a negatively charged, fluorescent dye and retained in cells expressing ALDH1. This assay can be used in heterogeneous cell populations, and as a result is a very popular method for quantifying CIC frequencies by flow cytometry as well as for CIC isolation by fluorescence-activated cell sorting (FACS).

Unfortunately, there are relatively few enzymatic markers available for CIC-labeling besides ALDH1-based assays. ALDH1 is not a universal marker for CICs, so there is a pressing need for more activity-based assays to label CICs. Other enzymatic activities linked to CICs include telomerase, the ATP-binding cassette membrane transporter ABCG2, and BMI-1 (34, 35). ABCG2 is a broad specificity drug transporter and its activity can be interrogated through the use of dye-exclusion assays (36). Dye-exlcusion assays, like ALHD1-assays, allow CICs to be quantified and sorted, but the potential overlap in transporter functions and low selectivity in substrates cause these assays to be less stringent measurements of CIC properties. Other enzymes, such as BMI-1, lack robust *in cellulo* assays in order to measure activity in viable cells. The development of new assays and labeling methods for identifying and tracking

CICs in mixed cell populations would greatly benefit the study of CICs in general by expanding the tools and targets available to researchers.

4.6. Conclusions and Future Directions

The discovery of CICs as the driving source of many tumors has spurred tremendous interest in determining the intra- and extracellular factors that control their behavior. Studying these factors has been challenging due to issues with CIC frequency and instability, as well as more general challenges with studying complex protein and gene interaction networks. The works presented in this thesis present our efforts towards studying how protein-protein interactions involved in transcriptional regulation can be interrogated and reveal new targets for future CIC modulation (Chapters 2, 3). Specifically, we showed how multidomain transcriptional coactivators can be identified and manipulated to downregulate CIC-associated genes in the absence of information about upstream signaling pathways or interaction partners (Chapter 2). We demonstrated that the coactivator protein p300 is an essential factor involved in *NANOG* expression in undifferentiated cancer cells (Chapter 2.3). The regulation of *NANOG*, like many other CIC genes, is currently not well understood in pathological contexts. As a result, our functional analysis of the *NANOG* promoter and its associated activators and coactivators can serve as an example of how further CIC genes can be explored (Chapter 2.5).

We have also described how individual functions within a multidomain protein like p300 can be tested through a series of complementary techniques including domain loss- and gain-of-function (Chapter 3.3), mutational analysis (Chapter 3.3), and small molecule probes (Chapter 3.4, 3.5.2). These studies allowed us to generate a working model for how *NANOG* is regulated in CICs and other undifferentiated cancer cells (Chapter 3.6.4). Based on this model, we can now propose new targets for inhibiting *NANOG* expression, including protein-protein interaction inhibitors for the CH1 domain, and p300-selective histone acetyltransferase (HAT) inhibitors. Similar strategies can be applied to other proteins of interest and thereby simplifying some of the challenges associated with complex protein-protein interaction networks (Chapter 4.3). Overall, the work described in this thesis reveals new potential targets for modulating *NANOG* expression in cancer cells. Importantly, this work is novel because it is the first effort at modulating *NANOG* expression by targeting the activators and coactivators involved in its expression. The NANOG protein itself is difficult to directly target pharmacologically due to poor structural characterization and limited information about its binding interactions. We reasoned that we could potentially disrupt *NANOG* expression and function by targeting the proteins that direct *NANOG* transcription, and demonstrate that the coactivator p300 and its CH1 and HAT domains are potential targets for downregulating *NANOG.* This information can be used to guide the development of new drug and therapeutic strategies for *NANOG*-expressing tumors.

Our results using the HAT inhibitor C646 and the histone deacetylase (HDAC) inhibitor trichostatin A (TSA) were unexpected and suggest a more complex series of interactions between histone modifying proteins, their non-histone substrates, and the impact on *NANOG* expression. The observation that C646 inhibits *NANOG* expression while decreasing histone H3 acetylation (H3Kac) at the *NANOG* promoter, while TSA inhibited *NANOG* while increasing H3Kac could indicate a non-histone substrate for one, or both, HATs and HDACs that helps modulate *NANOG* expression. We initially tested the possibility that the activator KLF4, which is a known p300 substrate, could be one such protein. However, mutation of the p300 acetylation sites within KLF4 had no significant impact on KLF4-driven *NANOG* promoter activity (Fig. 3.12), suggesting that some other protein may be the target. The p300 HAT has a large number of reported substrates (6), and it is difficult to predict which known substrates may be important for *NANOG* regulation. Additionally, there may be a number of as yet unidentified p300 HAT substrates that also help regulate *NANOG* expression. One potential future direction for investigating the relative substrates and roles of the p300 HAT and HDACs in *NANOG* expression would be through the use of peptide sequencing in C646- and TSA-treated cancer cells to identify differentially acetylated proteins. Such as study could generate a list of candidate proteins for further testing. Additionally, peptide sequencing techniques would allow for a closer inspection of the specific histone modifications associated with basal *NANOG* promoter activity, as well as those following C646 and TSA treatment. Our data so far has only examined general

H3 lysine acetylation. Specific histone modifications can carry dramatically different effects on transcription, and consequently having a detailed picture of how histone modifications change could reveal a more comprehensive picture of how p300 is specifically working at the *NANOG* promoter.

A number of challenges still remain more broadly when studying CICs, including issues with their identification, isolation, and cultivation. Several promising technologies have emerged for addressing these issues in other challenging cell types (Chapter 4.4, 4.5), and tailoring technologies to CICs will undoubtedly generate profound advances in the ways cancer is understood, diagnosed, and treated in the future. This thesis highlights just some of the many ways chemical biologists, with their unique perspectives, can contribute to the field of stem cell biology that is always need of better tools.

4.7. References

1. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105-111 (2001).
2. Lobo NA, Shimono Y, Qian D, Clarke MF. The biology of cancer stem cells. *Annu. Rev. Cell Dev. Biol.* **23**, 675-699 (2007).
3. Visvader JE & Lindeman. Cancer stem cells in solid tumors: accumulating evidence and unresolved questions. *Nat. Rev. Cancer* **8**, 755-768 (2008).
4. Gupta PB, *et al*. Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* **138**, 645-659 (2009).
5. Xia T, *et al*. Molecular imaging in tracking tumor stem-like cells. *J. Biomed. Biotechnol.* 420364 (2012).
6. Goodman RH & Smolik S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **14**, 1553-1577 (2000).
7. Boyer LA, *et al*. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-956 (2005).
8. Wang J, *et al*. A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364-368 (2006).
9. Loh YH, *et al*. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431-440 (2006).
10. Chiou SH, *et al*. Positive correlations of Oct4- and Nanog in oral cancer stem-like cells and high-grade squamous cell carcinoma. *Clin. Cancer Res.* **14**, 4085-4095 (2008).
11. Jeter CR, *et al*. Functional evidence that the self-renewal gene NANOG regulates human tumor development. *Stem Cells* **27**, 993-1005 (2009).
12. Jeter CR, *et al*. NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. *Oncogene* **30**, 3833-3845 (2011).
13. Bourguignon LY, *et al*. Hyaluronan-CD44 interaction with protein kinase Cε promotes oncogenic signaling by the stem cell marker Nanog and the production of microRNA-21, leading to down-regulation of the tumor suppressor protein PDCD4, anti-apoptosis, and chemotherapy resistance in breast tumor cells. *J. Biol. Chem.* **284**, 26533-26546 (2009).
14. Xie X, *et al*. Phosphorylation of Nanog is essential to regulate Bmi-1 and promote tumorigenesis. *Oncogene* **33**, 2040-2052 (2014).
15. Blobe GC, Obeid, Hannun YA. Regulation of protein kinase C and role in cancer biology. *Cancer Metastasis Rev.* **13**, 411-431 (1994).

16. Bertolotto C, *et al*. Protein kinase C theta and epsilon promote T-cell survival by a rsk-dependent phosphorylation and inactivation of BAD. *Biol. Chem.* **275**, 37246-37250 (2000).
17. Moretto-Zita M, *et al*. Phosphorylation stabilizes Nanog by promoting its interaction with Pin1. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13312-13317 (2010).
18. Kimbrel EA, *et al*. Systematic in vivo structure-function analysis of p300 in hematopoiesis. *Blood* **114**, 4804-4812 (2009).
19. Sachlos E, *et al*. Identification of drugs including a dopamine receptor antagonist that selectively target cancer stem cells. *Cell* **149**, 1284-1297 (2012).
20. Marthiens V, *et al*. Adhesion molecules in the stem cell niche-more than just staying in shape? *J. Cell Sci.* **123**, 1613-1622 (2010).
21. Brafman DA. Constructing stem cell microenvironments using bioengineering approaches. *Physiol. Genomics* **45**, 1123-1135 (2013).
22. Ludwig TE, *et al*. Derivation of human embryonic stem cells in defined conditions. *Nat. Biotechnol.* **24**, 185-187 (2006).
23. Hughes CS, Postovit LM, Lajoie GA. Matrigel: a complex protein mixture required for optimal growth of cell culture. *Proteomics* **10**, 1886-1890 (2010).
24. Villa-Diaz LG, *et al*. Synthetic polymer coatings for long-term growth of human embryonic stem cells. *Nat. Biotechnol.* **28**, 581-583 (2010).
25. Musah S, *et al*. Glycosaminoglycan-binding hydrogels enable mechanical control of human pluripotent stem cell self-renewal. *ACS Nano*. **6**, 10168-10177 (2012).
26. Derda R, *et al.* Defined substrates for human embryonic stem cell growth identified from surface arrays. *ACS Chem. Biol.* **2**, 347-355 (2007).
27. Klim JR, *et al*. A defined glycosaminoglycan-binding substratum for human pluripotent stem cells. *Nat. Methods* **7**, 989-994 (2010).
28. Barrilleaux B, *et al.* Ex vivo engineering of living tissues with adult stem cells. *Tissue Engineering* **12**, 3007-3019 (2006).
29. Al-Hajj M, *et al*. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3983-3988 (2003).
30. Quintana E, *et al*. Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell* **18**, 510-523 (2010).
31. Chiou SH, *et al*. Positive correlations of Oct4- and Nanog in oral cancer stem-like cells and high-grade squamous cell carcinoma. *Clin. Cancer Res.* **14**, 4085-4095 (2008).
32. Ma S, *et al*. Aldehyde dehydrogenase discriminates the CD133 liver cancer stem cell populations. *Mol. Cancer Res.* **6**, 1146-1153 (2008).
33. Croker AK, *et al*. High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability. *J. Cell. Mol. Med.* **13**, 2236-2252 (2009).
34. Gires O. Lessong from common markers of tumor-initiating cells in solid cancers. *Cell. Mol. Life Sci.* **68**, 4009-4022 (2011).

35. Prince MEP & Ailles LE. Cancer stem cells in head and neck squamous cell cancer. *J. Clin. Oncol.* **26**, 2871-2875 (2008).
36. Sharom FJ. ABC multidrug transporters: structure, function and role in chemoresistance. *Pharmacogenomics* **9**, 105-127 (2008).