



Phosphorylation of NANOG by casein kinase I regulates embryonic stem cell self-renewal

Nicholas P. Mullin¹ , Joby Varghese², Douglas Colby¹, Julia M. Richardson³, Greg M. Findlay² and Ian Chambers¹ 

¹ Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, UK

² Protein Phosphorylation and Ubiquitylation Unit, James Black Centre, School of Life Sciences, Dundee, UK

³ Institute of Quantitative Biology, Biochemistry and Biotechnology, Edinburgh, UK

Correspondence

N. P. Mullin and I. Chambers, Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, Edinburgh EH16 4UU, UK
Tel: 01316519500
E-mail: nick.mullin@ed.ac.uk (N.P.M.)
ichambers@ed.ac.uk (I.C.)

(Received 28 August 2020, revised 13 October 2020, accepted 15 October 2020, available online 18 November 2020)

doi:10.1002/1873-3468.13969

Edited by Ned Mantei

The self-renewal efficiency of mouse embryonic stem cells (ESCs) is determined by the concentration of the transcription factor NANOG. While NANOG binds thousands of sites in chromatin, the regulatory systems that control DNA binding are poorly characterised. Here, we show that NANOG is phosphorylated by casein kinase I, and identify target residues. Phosphomimetic substitutions at phosphorylation sites within the homeodomain (S130 and S131) have site-specific functional effects. Phosphomimetic substitution of S130 abolishes DNA binding by NANOG and eliminates LIF-independent self-renewal. In contrast, phosphomimetic substitution of S131 enhances LIF-independent self-renewal, without influencing DNA binding. Modelling the DNA–homeodomain complex explains the disparate effects of these phosphomimetic substitutions. These results indicate how phosphorylation may influence NANOG homeodomain interactions that underpin ESC self-renewal.

Keywords: casein kinase I; DNA binding; NANOG; phosphorylation; self-renewal

The patterns of gene expression during development and cell fate commitment are controlled by a complex web of sequence-specific transcription factors (TFs). The ability of TFs to regulate gene expression relies on their ability to interact with both DNA and protein partners. Regulation of DNA and partner protein binding, together with modulation of expression levels, enables both temporal control and spatial control of gene expression within a developing organism. One of the best understood mechanisms for regulating TF function is *via* covalent modification of the protein. These modifications include phosphorylation [1], O-GlcNAcylation [2], acetylation [3] and ubiquitination [4]. Examples of regulation of TF function *via* phosphorylation include STAT3, which is translocated into the nucleus upon phosphorylation [5], c-myb, which is

unable to bind to DNA upon phosphorylation of a residue N-terminal to the DNA-binding domain [6] and oestrogen receptor A, which is also blocked from DNA binding upon phosphorylation [7].

In embryonic stem cells (ESCs), the network of TFs that underpin self-renewal is centred on the triumvirate of factors, NANOG, OCT4 and SOX2 that are often referred to as master transcription factors [8–13]. Amongst these TFs, NANOG has the ability to drive LIF-independent ESC self-renewal when overexpressed [8]. In addition to transcriptional interactions between these TFs [14], each has been shown to be modified *in vivo* by covalent modification. For example, various screens have identified phosphorylated residues in NANOG, OCT4 and SOX2 [15–18] and phosphorylation has been reported to modify NANOG protein

Abbreviations

CKI, casein kinase I; TF, transcription factor.

stability [18,19]. Furthermore, OCT4 and SOX2 are regulated by O-GlcNAcylation [20] and SOX2 has been demonstrated to be regulated by acetylation [21].

To investigate how NANOG acts through partner proteins to deliver function, we determined the NANOG interactome [22]. This identified > 130 partner proteins of NANOG, including two kinases, one of which is casein kinase I (CKI). The recent demonstration that inhibition of CKI increases reprogramming efficiency of epiblast stem cells [23] suggests that CKI-mediated phosphorylation is an important regulatory mechanism in pluripotent cells. CKI is a widely expressed enzyme with broad specificity implicated in diverse processes including regulation of circadian rhythms, the cell-cycle, cell adhesion and cytoskeletal structure [24]. CKI also regulates transcription. For example, CKI regulates the activity of the TFs NFAT-1 and HIF by direct phosphorylation [25,26], is a critical component of the β -catenin pathway [27] and regulates p53 activity [28,29]. At sites unprimed by prior phosphorylation, CKI has a preference for a negative charge at a position -3 to the target site and an isoleucine at $+1$ (D/E-x-x-S/T-I) [30,31]. In NANOG, 17 of the 61 serine/threonine residues are potential CKI targets since they either have a negative charge or a phosphorylatable residue at position -3 . Here, we report that CKI phosphorylates NANOG *in vitro*, map sites of phosphorylation and assess the functional effects of phosphorylation at novel sites in the homeodomain on NANOG function in ESCs.

Materials and methods

Preparation and phosphorylation of recombinant NANOG

Recombinant NANOG was prepared as described [32]. Phosphorylation was performed by incubating 2 μ g NANOG with 0.3 U CKI (New England Biolabs, Ipswich, MA, USA, P6030) in the presence of 200 μ M ATP in reaction buffer (50 mM Tris pH7.5, 10 mM MgCl₂, 5 mM DTT, 0.025% NP-40) at 30 °C for 2 h. The CKI inhibitors triamterene and D4476 were used at final concentrations of 50, 10 and 1 μ M. The reaction was stopped by boiling in Laemmli buffer and the samples loaded onto SDS/PAGE.

Phosphosite mapping by mass spectrometry

NANOG phosphorylated by CKI was resolved by SDS/PAGE and stained with colloidal Coomassie. The specific NANOG band was excised, washed in water and shrunk in

acetonitrile. Gel pieces were reswollen in 50 μ L 50 mM Tris/HCl pH 8.0 and reshrunk in acetonitrile. Gel pieces were reduced in 50 mM Tris/HCl pH 8.0, 5 mM DTT at 65 °C for 20 min and alkylated in 20 mM iodoacetamide in 50 mM Tris/HCl pH 8.0 at room temperature for 20 min. Gel pieces were shrunk in acetonitrile, reswollen in 50 mM triethylammonium bicarbonate and reshrunk in acetonitrile. Protein was digested from the gel slice using 5 μ g·mL⁻¹ mass spectrometry grade trypsin in 50 mM triethylammonium bicarbonate overnight at 30 °C. The gel was shrunk by addition of acetonitrile, swollen in 0.1% TFA and peptides extracted twice with acetonitrile. All supernatants were combined and lyophilised in a SpeedVac prior to mass spectrometry analysis.

Mass spectrometric analysis was performed by LC-MS-MS using a linear ion trap-orbitrap hybrid mass spectrometer (Orbitrap-Classic, Thermo, Waltham, MA, USA) equipped with a nanoelectrospray ion source (Thermo) and coupled to a Proxeon EASY-nLC system. Peptides (in 0.1% formic acid, 2% acetonitrile) were injected *via* a 2 cm trap column (Nano Separations, NS-MP-10 BioSphere C18, 5 μ m, 120 Å, 360/100 μ m) onto a Thermo (Part No. 160321) Acclaim PepMap100 reverse phase C₁₈ 3 μ m column (75 μ m \times 15 cm) with a flow of 300 nL·min⁻¹ and eluted with a 45 min linear gradient of 95% solvent A (2% acetonitrile, 0.1% formic acid in H₂O) to 40% solvent B (90% acetonitrile, 0.08% formic acid in H₂O), followed by a rise to 80% B at 48 min. The instrument was operated with the 'lock mass' option to improve the mass accuracy of precursor ions, and data were acquired in the data-dependent mode, automatically switching between MS and MS-MS acquisition. Full scan spectra (m/z 340–1800) were acquired in the orbitrap with resolution $R = 60\,000$ at m/z 400 (after accumulation to an FTMS Full AGC Target; 1 000 000; MSn AGC Target; 100 000). The 5 most intense ions, above a specified minimum signal threshold (20 000), based upon a low-resolution ($R = 15\,000$) preview of the survey scan, were fragmented by collision-induced dissociation and recorded in the linear ion trap (Full AGC Target; 30 000. MSn AGC Target; 5000). Multistage activation was used to provide a pseudo-MS3 scan of any parent ions showing a neutral loss of 48.9885, 32.6570, 24.4942, allowing for 2+, 3+ and 4+ ions, respectively. The resulting pseudo-MS3 scan was automatically combined with the relevant MS2 scan prior to data analysis.

RAW files from the OrbiTrap-Classic were analysed by Proteome Discoverer 1.4-SP1 (Thermo) using the phosphoRS [33] node and searched using MASCOT 2.4.1 (www.matrixscience.com) against an in-house database that contained the NANOG Q80Z64.1 sequence. Searches allowed for variable modifications; dioxidation (M), oxidation (M), phospho (ST), phospho (Y) and a fixed modification: carbamidomethyl (C). Peptide mass tolerance was ± 10 ppm, and MS2 fragment mass tolerance was ± 0.6 Da.

Cell transfections and self-renewal assays

For episomal assessment of function of NANOG mutants, DNA was transfected into E14/T cells using Lipofectamine 2000 (Thermo Fisher, 11668027). One million cells were plated per well of a 6-well plate in 2.0 mL Gmem β /10% FCS, LIF 30–60 min before DNA addition. 3 μ g DNA was diluted into 250 μ L of serum-free Gmem β for each transfection. Separately, 3 μ L of Lipofectamine 2000 was diluted into 250 μ L of serum-free Gmem β and incubated at room temperature for 5 min before mixing with the DNA solution and incubation at room temperature for 20 min. The transfection cocktail was then added to the cells and the plates rocked gently before being placed in the incubator. Cells were trypsinised 24 h later, plated at 5×10^4 cells per 9-cm dish in medium supplemented with 2 μ g·mL⁻¹ puromycin and medium replaced every 2 days. Colonies were photographed after 12 days following fixation and alkaline phosphatase staining using a leucocyte alkaline phosphatase kit (Sigma kit 86-R, St. Louis, MO, USA) [34].

For generation of stable lines, plasmids encoding NANOG and NANOG mutants were linearised with ScaI and introduced into 44Cre6 NANOG null cells [35] *via* electroporation. After selection in media containing puromycin, 600 cells were plated in triplicate in 6-well plates plus and minus LIF. Plates were stained for alkaline phosphatase as described above. Antibodies used for determining expression levels are anti-NANOG, A300-397A, Bethyl Laboratories and anti-LAMIN B1, Ab16048, Abcam (Cambridge, MA, USA).

Electrophoretic mobility shift assay

All assays used oligonucleotide labelled at the 5' end of the upper strand with IR700 dye. The DNA probe corresponds to the NANOG-binding site in Tcf3 promoter CTGTTAATGGGAGC [36]. Protein was added to oligonucleotide and buffer to give final concentrations of 10 mM HEPES pH7.9, 10% glycerol, 50 mM KCl, 10 mM NaCl, 0.4 mM EDTA, 2.5 mM DTT, 0.25% Tween 20, 100 mg·mL⁻¹ dI.dc, 1 mg·mL⁻¹ BSA, 10 nM oligonucleotide, 250 nM protein. Reactions were incubated at room temperature for 30 min before analysis on 5% acrylamide gels [37]. Data were collected on a LICOR Odyssey Fc. Recombinant His-tagged homeodomains were expressed in *E. coli* with induction of protein expression at 30 °C for 3 h. Bacterial lysates were prepared in 25 mM Tris pH8.0, 500 mM NaCl, 30 mM imidazole and loaded onto His-select resin (Sigma, P6611). After washing in the same buffer, protein was eluted in 25 mM Tris pH8.0, 500 mM NaCl, 250 mM imidazole. Proteins were dialysed against 25 mM Tris pH8.0, 500 mM NaCl and used directly in EMSAs.

Mutagenesis

Synthetic DNAs (IDT) encoding the mutations were inserted into pPyCAG-NANOG [8] cut with PpuMI and NotI.

Chromatin immunoprecipitation

Ten million E14Tg2a cells were resuspended in 3 mL GMEM/FCS (10%) and cross-linked with 1% formaldehyde (Sigma, P6611) for 10 min at room temperature. The reaction was quenched with 0.125 mM glycine for 5 min at room temperature. Cells were pelleted and washed twice with cold PBS. Cells were resuspended in 300 μ L of swelling buffer (5 mM Pipes pH 8, 85 mM KCl) supplemented with 1 \times protease inhibitor cocktail (Roche) and 0.5% NP-40. The suspension was incubated for 20 min on ice. Nuclei were pelleted and resuspended in 1.5 mL of 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris/HCl pH8, 150 mM NaCl, supplemented with 1 \times protease inhibitor cocktail. Samples were sonicated at 4 °C in a Bioruptor (Diagenode). 20 μ g of sonicated chromatin was used per ChIP. Immunoprecipitation with 5 μ g rabbit anti-NANOG [32] was performed overnight with rotation at 4 °C, in a final volume of 1 mL. Immunocomplexes were purified using Protein-G Dynabeads (Lifetech, Waltham, MA, USA). Washing and elution were as described in Ref. [38]. The sequences of the oligos used for ChIP-Q-PCR are as follows: Xin1: AACCCCTTTTAAGTCCACTGTAAATTCC and TAGAGAGCCAGACAATGCTAAGCC; Xex1b: CATCAGGCTTGGCAGCAA and TTCATCAGCAATGTCATATCAAACA; EE2: TGCTGGTGGTATTCAAC TGC and CGGATCTGTGGAATTCGTG; E7: TAAGG GACACCTCCCTAGCC and CCATACCCCACACACCA TGT; Otx2 A: CAAGAACAAAACCCACGCC and GTGCCAGCCAATGAGTCCTA; Otx2 B: GTGGAAGG GCGTCTAGAAGG and GCAGTCAATGGGCTGAG TCT [39].

Modelling a mouse NANOG homeodomain–DNA structure

NANOG homeodomain sequences from mouse and human were aligned with ClustalOmega [40]. The coordinates of protein chain A from the mouse NANOG homeodomain crystal structure (PDBID: 2VI6) were renumbered and superimposed on chain A of the human NANOG HD cocrystal structure with DNA (PDBID: 4RBO) with SUPERPOSE in the CCP4 software suite [41]. The side chains of serines 130 and 131 were phosphorylated, and the structures were visualised with PYMOL (Schrodinger Inc, New York, NY, USA).

Results

NANOG is phosphorylated by casein kinase I at multiple sites

The presence of CKI in the NANOG interactome [22] combined with the potential of kinases to modulate the function of TFs [5–7] led us to investigate whether

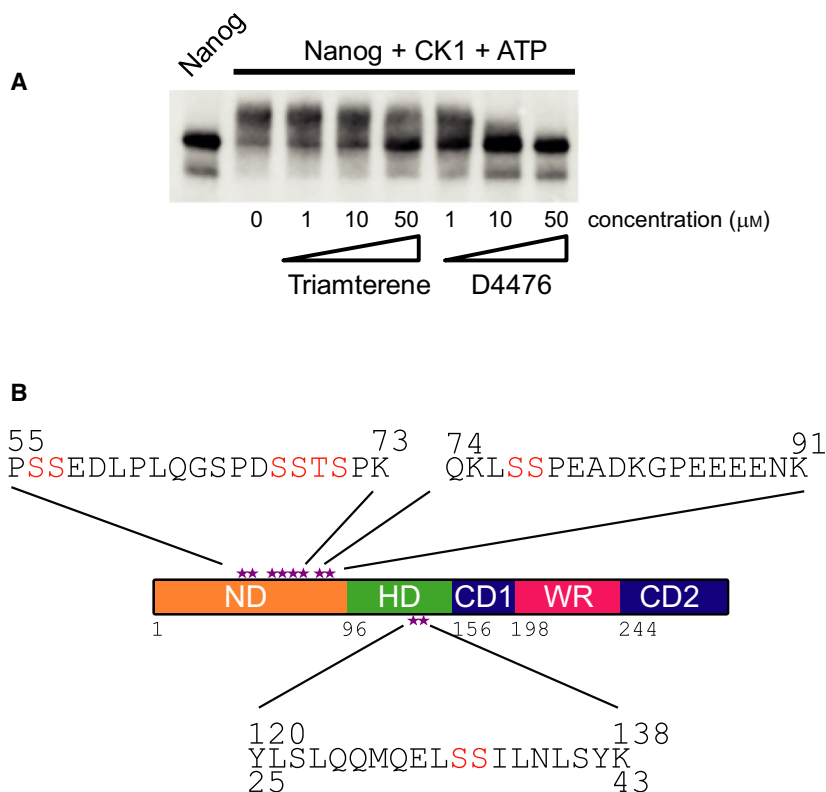


Fig. 1. NANOG is phosphorylated by casein kinase I. (A) Immunoblot of recombinant NANOG incubated with CKI and ATP in 0, 1, 10 or 50 μM triamterene or D4476. (B) Domain diagram of NANOG with phosphorylation sites identified by mass spectrometry marked as stars. ND – N-terminal domain, HD – homeodomain, CD1 – C-terminal domain 1, WR – tryptophan repeat, CD2 – C-terminal domain 2. Sequences of phosphorylated NANOG peptides are shown with potential phosphorylation sites highlighted. For all peptides, NANOG residue numbering is shown above the sequence. For the homeodomain-derived peptide, the canonical homeodomain numbering is shown below the sequence.

CKI could modulate NANOG function by phosphorylation. Recombinant NANOG was therefore incubated with recombinant active CKI and ATP *in vitro*. This resulted in a clear shift in molecular weight of NANOG, which can be indicative of one or more phosphorylation events (Fig. 1A). The shift in mobility was inhibited by inclusion of either of the CKI inhibitors triamterene or D4476, suggesting that retarded electrophoretic mobility is indeed a result of NANOG phosphorylation by CKI. The greater degree of inhibition by D4476 is consistent with the lower IC₅₀ of CKI inhibition by D4476 (0.3 μM) compared to triamterene (6.9 μM) [42,43].

To identify phosphorylation sites, NANOG was incubated with CKI and ATP, subjected to SDS/PAGE and the stained bands analysed by mass spectrometry. Three phosphorylated peptides were detected (Fig. 1B and Table 1), two in the N-terminal domain and one in the homeodomain. The most N-terminal phosphorylated peptide was detected in singly and doubly phosphorylated forms. The single phosphorylation is on either serine 56 or 57, with the mass spectrometry data indicating an equal probability of phosphorylation on each residue. The doubly phosphorylated form is also phosphorylated on either serine 56 or 57. In addition, the doubly phosphorylated

form is likely to be modified on either serine 68 or 69, although there is also a low probability that this second phosphorylation could occur on either threonine 70 or even serine 71 (Table 1). The second phosphorylated N-terminal peptide contains a single modification that is likely to be on serine 77, although the mass spectrometry data indicate a 10% probability that serine 78 could be phosphorylated. These results are consistent with previous work showing that NANOG expressed in 293T cells is phosphorylated on either serine 56 or 57 and either serine 77 or 78 [19]. The third phosphopeptide is derived from the homeodomain. The mass spectrometry data suggest that the modified residue is likely to be serine 131. Notably, S131 is a good match to the consensus CKI sequence (D/E-x-x-S/T-I), since residue 128 is a glutamate and residue 132 is an isoleucine [30,31]. However, the possibility that serine 130 is phosphorylated cannot be discounted. S130 and S131 correspond to residues 35 and 36 in the canonical homeodomain numbering system [44]. Therefore, the phosphorylation sites identified here by mass spectrometry include residues previously reported as phosphorylation sites [19,45,46] as well as novel sites of phosphorylation within the homeodomain; these are highlighted on a domain diagram of NANOG (Fig. 1B).

Table 1. Identification of NANOG sites phosphorylated by CK1 using mass spectrometry.

Peptide sequence ^a	Phos sites	M/Z (Da)	Ion score	phosphoRS probabilities ³³
55-PSEDLPLQGSPDSSTSPK-73	1	1004.9417	59	S56:50.0 S57:50.0
55-PSEDLPLQGSPDSSTSPK-73	2	1044.9257	38	S56:50.0 S57:50.0 S68:47.5 S69:47.5 T70:4.5 S71:0.5
74-QKLSSPEADKGPEEEENK-91	1	698.9794	39	S77:90.0 S78:10.0
120-YLSLQQMQLSSILNLSYK-138	1	780.0541	30	S130:8.9 S131:91.1

^aSequences are numbered according to the primary NANOG sequence.

Mutagenesis of phosphorylation sites within the NANOG homeodomain

The identification of CKI-mediated phosphorylation within the NANOG DNA-binding domain suggests that these modifications may regulate NANOG function. To test for possible roles of phosphorylation, the two potential phosphorylation sites within the homeodomain were mutated simultaneously, either to alanine or to the phosphomimetic residue aspartate. In addition, both residues were individually mutated to aspartate. To determine whether these mutations affected the defining capacity of NANOG to confer LIF-independent self-renewal on transfected ESCs [8], these mutants were cloned into constitutive episomal expression plasmids and transfected into E14/T ESCs to rapidly assess ESC self-renewal (Fig. 2A,B). Examples of colony morphologies containing undifferentiated ES cells, differentiated cells or a mix of both are shown in Fig. 2B. Mutagenesis of both residues to alanine has little effect on the ability of NANOG to confer LIF-independent self-renewal. However, mutagenesis to aspartate either singly or doubly has a marked effect. Substitution of S130 by aspartate (S130D) abolishes the LIF-independent self-renewal function of NANOG. Moreover, in the presence of LIF, S130D acts as a dominant interfering mutant since self-renewal is decreased below the level obtained by transfection of an empty vector control. In contrast, a phosphomimetic at position 131 increases the number of self-renewing colonies formed in either the presence or absence of LIF (Fig. 2B and Fig. S1A,B).

To attempt to verify the findings from E14/T ESCs, and to further investigate the effect of mutating serine 130 and 131, stably expressing ESC lines were generated in a *Nanog*^{-/-} background [35] and the ability of

the mutants to confer LIF-independent self-renewal assayed (Fig. 3). Immunoblot analysis indicated that the stable cells express comparable levels of NANOG (Fig. 3A). Consistent with the clonal self-renewal results from E14/T cells, S131D increases self-renewal efficiency in the presence and absence of LIF whereas S130D decreases self-renewal efficiency (Fig. 3B). Notably, the dominant-negative effect of S130D observed in E14/T ESCs is not observed in a *Nanog*^{-/-} background, consistent with the mutant protein interfering with WT NANOG function. These data suggest that individual phosphorylation at the adjacent sites 130 and 131 has opposite effects on the biological activity of NANOG. Mimicking phosphorylation at 130 abrogates NANOG function, and in wild-type cells interferes with the function of endogenous NANOG. In contrast, mimicking phosphorylation at 131 enhances both the total number and the proportion of cells that undergo efficient self-renewal.

DNA-binding properties of mutants

As S130 and S131 are in the NANOG homeodomain, their phosphorylation may alter the ability of NANOG to bind DNA. To compare the abilities of NANOG mutants to bind DNA, recombinant homeodomains encompassing the D > A mutations were purified for use in electrophoretic mobility shift assays (EMSAs; Fig. 4A). While the S131D homeodomain has a DNA-binding capacity similar to that of the wild-type NANOG homeodomain, EMSAs indicate that the DNA-binding capacity of S130D homeodomain is severely impaired (Fig. 4B).

To investigate the ability of NANOG mutants to bind to chromatin, CHIP was performed using the

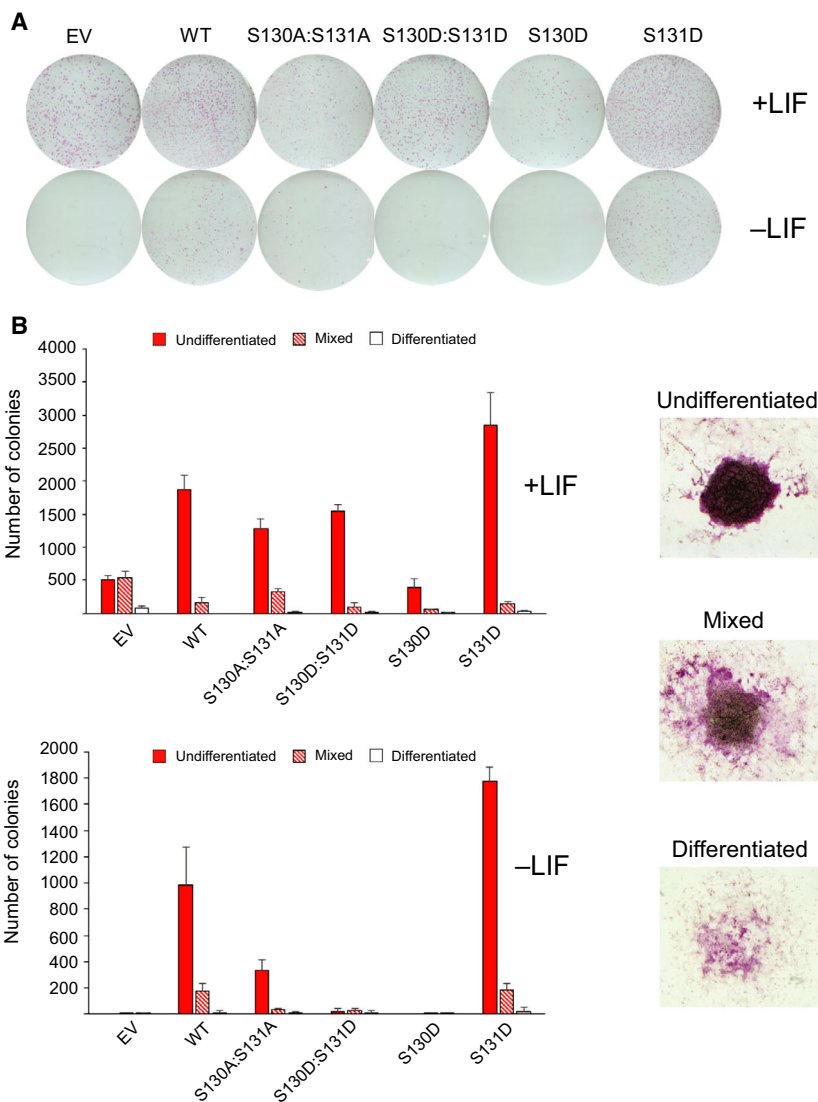


Fig. 2. Self-renewal assays of E14/T cells transfected with NANOG variants. (A) E14/T ESCs were transfected with pPyCAGIP plasmids encoding the indicated NANOG mutants and cultured in media containing puromycin in the presence or absence of LIF (EV, empty vector; WT, wild-type NANOG). After 12 days, plates were stained for alkaline phosphatase activity. Assays were performed in triplicate. (B) Left, Quantitative assessment of colony morphologies from (A) determined in the presence or absence of LIF. Data shown are from a single experiment (data and error bars are the means and standard deviations, respectively, from three technical replicates) and are representative of three independent biological replicates. Biological replicates are shown in Fig. S1. Right, examples of colony morphologies containing undifferentiated ES cells, differentiated cells or a mix of both.

stable lines described above. Mutation of S131 to aspartate has little effect on the ability of NANOG to bind chromatin (Fig. 4C). However, S130D was not localised to chromatin at any of the three loci assayed. This is in agreement with the reduced capacity of S130D to bind DNA *in vitro* (Fig. 4B). Together these data indicate that while phosphorylation of NANOG on S131 has no effect on DNA binding, phosphorylation of S130 may disrupt the binding of NANOG to target sites in chromatin.

Modelling of the NANOG homeodomain–DNA complex

The differing effects observed between mutating S130 and S131 on DNA and chromatin binding are

intriguing, given that the residues are adjacent in the primary NANOG sequence. While no three-dimensional structure is available for mouse NANOG in complex with DNA, the tertiary structure of the mouse homeodomain has been solved [36], as has the structure of the human homeodomain in complex with DNA [47]. An alignment of the mouse and human NANOG homeodomains (Fig. 5A) was used to construct a three-dimensional model of mouse NANOG in complex with DNA (Fig. 5B). The root mean square deviation of superposed mouse and human NANOG HD chains was 0.91 Å over all atoms. Highlighted on the model are S130 and S131, which are present at the C-terminal end of homeodomain α -helix 2. This shows that neither S130 nor S131 are in direct contact with DNA (Fig. 5B). A three-dimensional

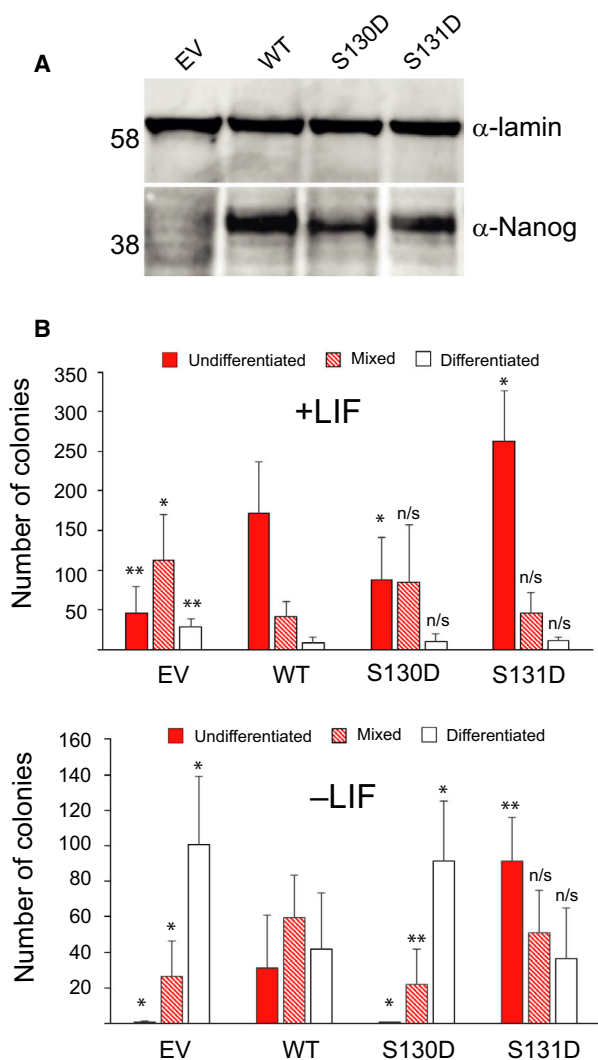


Fig. 3. Self-renewal assays of *NANOG* null ESCs rescued with *NANOG* variants. (A) *NANOG*^{-/-} ESCs (44cre6) were electroporated with linearised pPyCAGIP plasmids encoding the indicated *NANOG* mutants and selected in puromycin LIF (EV, empty vector; WT, wild-type *NANOG*). After expansion, cell extracts were assayed for expression level by immunoblotting. (B) Quantitative assessment of colony morphologies of stably expressing lines from (A) plated at clonal density in the presence or absence of LIF and stained for alkaline phosphatase activity after 6 days. Data and error bars are the means and standard deviations, respectively, of three independent experiments. Significance of variation from WT **P* < 0.1, ***P* < 0.05 and n/s – not significant (Student's *t*-test).

model of mouse *NANOG* in complex with DNA was also prepared after addition of a phosphate group to either S130 or S131 (Fig. 5C). Mutation of S130 to aspartate creates steric clashes with Tyr 137 and Val 140 at the N terminus of α -helix 3 (or in an alternative rotamer conformation with backbone atoms of Leu

135 and Ser 136; Fig. 5D,E). The phosphorylation of Ser 130 is therefore likely to lead to a change in the position or orientation of at least the N-terminal end of α -helix 3, which may interfere with DNA binding. In particular, the critical contacts made with the DNA backbone by Tyr 137 at the N-terminal end of this helix are likely to be disrupted. In contrast to S130, the side chain of S131 faces away from helix 3, making it unlikely that phosphorylation of this residue would directly influence DNA binding. This is consistent with the EMSA and ChIP results showing that S131D retains the ability to bind DNA and to localise to *NANOG* target sites in chromatin (Fig. 4B,C).

Discussion

The *NANOG* interactome contains more than 130 proteins of which only two are kinases (casein kinases I and II) [22]. A role for CKI in pluripotent cells was recently reported, with the demonstration that CKI inhibition improved the efficiency of reprogramming of epiblast stem cells to ESCs [23,42]. To further investigate CKI function in ESCs, the ability of CKI to directly phosphorylate *NANOG*, one of the central members of the pluripotency gene regulatory network, was investigated. This showed that *NANOG* is a CKI substrate in vitro. Of the sites we identify as phosphorylated by CKI, residues 70 and 131 are previously unreported phosphorylation sites that each have a negatively charged residue at position -3 that fits the D/E-x-x-S/T consensus for CKI phosphorylation [30,31]. Of the remaining sites, residues 68 and 71 have serines at the -3 position which could act as priming sites in vivo. Consistent with this possibility, serine 65 has been reported to be phosphorylated in cells [19,45]. The probability of serine 71 being phosphorylated by CKI is low (Table 1), and furthermore, there has been no in vivo identification of a potential priming phosphorylation on serine 68. We also identified additional CKI sites that have been reported as sites of phosphorylation. Specifically, phosphorylation of mouse *NANOG* at S56/S57 has been demonstrated in ESCs and in 293T cells [19,45]. In addition, protein kinase A can phosphorylate recombinant human *NANOG* on the residue equivalent to S78 in mouse *NANOG* [46]. However, these latter sites do not match the consensus for CKI phosphorylation. Moreover, functional roles have not been assigned to any of the above phosphorylated residues. Interestingly, previous studies have not reported phosphorylation within the *NANOG* homeodomain. Our work identifies a novel phosphorylation site(s) within the homeodomain which could potentially occur on S130 and/or S131 (corresponding

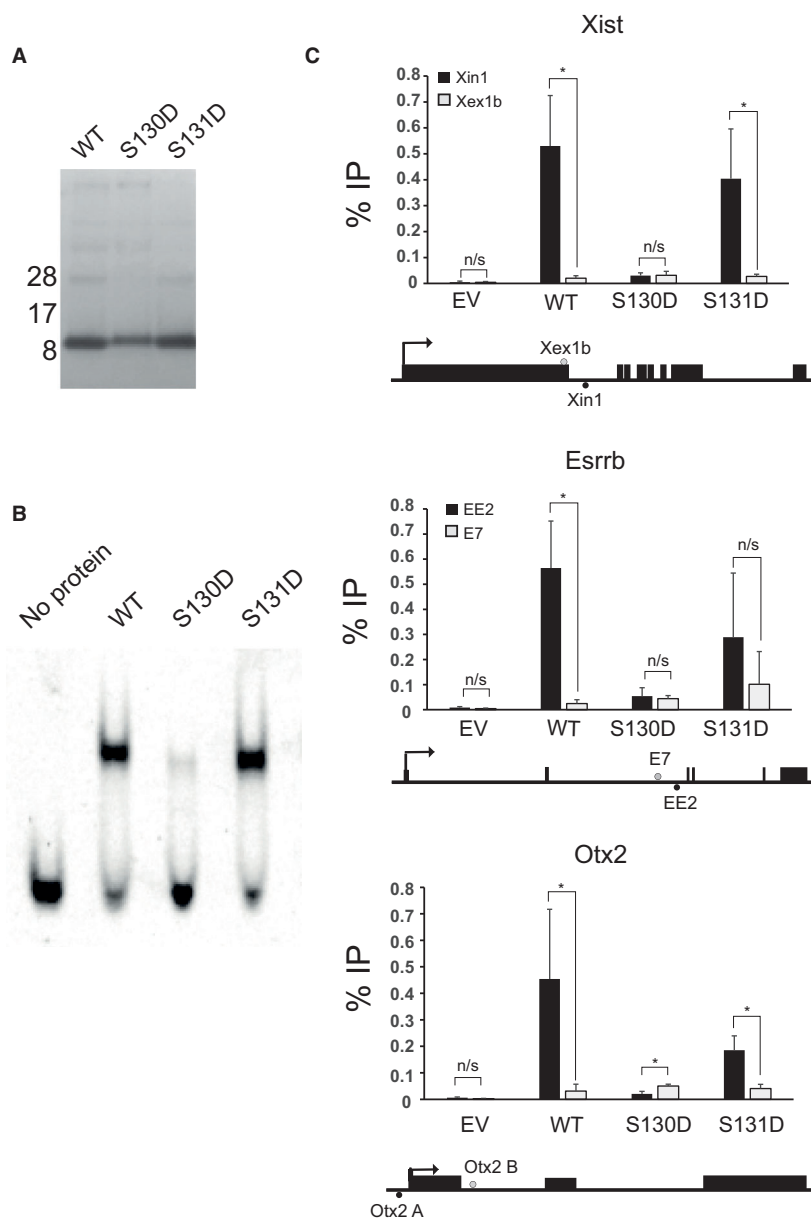


Fig. 4. DNA binding by NANOG mutants. (A) SDS/PAGE of purified homeodomains expressed in *E. coli*. (B) EMSAs with recombinant homeodomains performed using a probe corresponding to the NANOG-binding site within the *Tcf3* locus [36]. (C) ChIP-PCR of NANOG at known binding sites. ChIP was performed with NANOG antibody using the stable populations described in Fig. 3. Three known NANOG-binding sites within *Xist*, *Esrrb* and *Otx2* were assessed; regions within the same genes which do not bind NANOG provided negative controls. The positive binding regions (Xin1, EE2 and Otx2 A) are diagrammed below each gene and negative control regions (Xex1b, E7 and Otx2 B) above each gene. Data are mean of three replicates. Error bars are standard deviation. * $P < 0.05$, n/s – not significant.

to residues 35 and 36 in the canonical homeodomain numbering system [44]).

Intriguingly, phosphorylation of residues within other homeodomains has functional consequences. For example, phosphorylation of PRH at a serine immediately before the start of helix 3 of the homeodomain (residue 41 of the HD numbering scheme) decreases DNA binding [48]. Structural modelling suggests this is due to a disruption of the side chain interactions that mediate DNA binding [49]. This is interesting relative to our study, since the phosphate group added to residue 41 of the PRH HD is in a 3D space close to that occupied by introduction of a phosphate group to

homeodomain residue 35 (S130) of NANOG. Of further note, mimicking phosphorylation of NANOG HD residue 35 reduces DNA binding and abolishes the biological activity of NANOG, echoing the effect of mimicking phosphorylation at PRH HD residue 41 on abolition of activity *in vivo* [48]. In contrast, phosphorylation within homeodomains can also increase DNA affinity. This is the case for Nkx2.5, where phosphorylation of a residue between the first and second helices of the homeodomain increases affinity for DNA and is also important for transcriptional activity [50]. Phosphorylation of homeodomains can also have functional consequences without affecting DNA

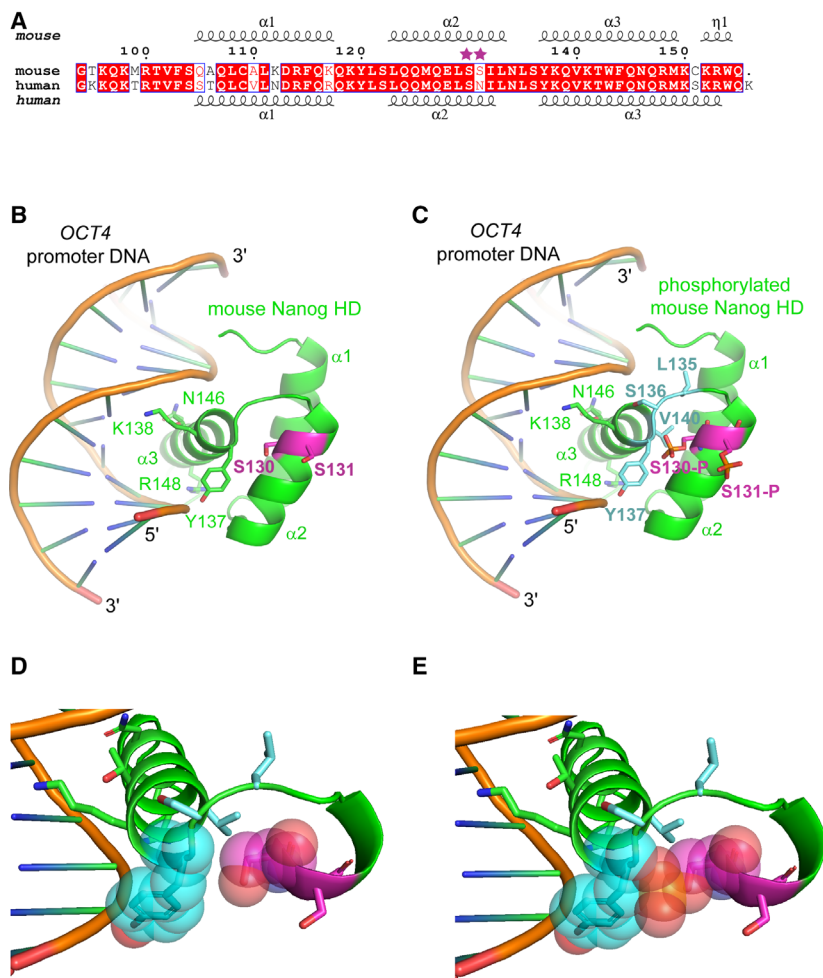


Fig. 5. Model of mouse NANOG homeodomain in complex with DNA. (A) Alignment of the homeodomain (HD) sequences of mouse (amino acids 94–155) and human NANOG (93–155). White letters on red background indicate conserved amino acids; red letters on white background indicate similar amino acids. Two potential phosphorylation sites, S130 and S131, are highlighted with stars. Secondary structure elements of the mouse NANOG HD crystal structure (PDBID:2VI6) and the human NANOG HD cocrystal structure with DNA (PDBID:4RBO) are shown above and below the alignment, respectively. (B) Structural model of mouse NANOG HD bound to OCT4 promoter DNA, based on the human NANOG HD:DNA cocrystal structure. Four amino acids in α -helix 3 that contact DNA (Y137, K138, N146, R148) are shown as sticks, along with S130 and S131 (purple). (C) Structural model of mouse NANOG HD, phosphorylated at S130 (S130-P) and S131 (S131-P), bound to OCT4 promoter DNA. Phosphorylation of S130 introduces steric clashes with the amino acid residues shown in blue (L135, S136, Y137, V140) sticks. (D, E) Expanded view of region around S130. Atoms of Y137 and Ser130/pSer130 are shown as spheres representing van der Waal radii. Residues in blue as per C.

binding. For example, Ftz can be phosphorylated by protein kinase A on a threonine in the first helix of the homeodomain. Substitution of threonine with alanine does not affect DNA binding but fails to rescue the segmentation defect in *ftz* null drosophila embryos [51].

Mass spectrometry identified a singly phosphorylated residue in the NANOG peptide covering serines 130 and 131. While it is possible that NANOG could either be phosphorylated on 130 or 131, probability analysis of the mass spectrometry data suggests that phosphorylation of residue 131 is the more likely. Moreover, it is notable that phosphorylation at S131 is consistent with studies of CKI specificity, since the residue at -3 is a glutamate and the residue at $+1$ is an isoleucine [30,31]. Nevertheless, to obtain comprehensive insight into the potential functional consequences of phosphorylation of either of these residues, we mimicked phosphorylation by mutagenesis of serines to aspartate. Intriguingly, aspartate substitution of

either S130 or S131 of NANOG had functionally distinct effects. Similar to the effect of phosphorylation of S41 within the PRH homeodomain [48], aspartate substitution of S130 reduces the capacity of NANOG to bind DNA. Consistent with this loss of DNA-binding activity, S130D showed a loss of ESC self-renewal capacity. Moreover, the modelled structure of a NANOG–DNA complex shows that phosphorylation of S130 would cause steric hindrance to the spatial disposition of residues L135, S136, Y137 and V140 (HD residues 40, 41, 42 and 45). Notably, this would include interference with the critical contacts between the DNA backbone and Y137. Alanine substitution of Y137 decreases DNA binding activity [47] and impairs the reprogramming activity of mouse NANOG [52]. These contact disruptions differ in detail to those seen in PRH, where phosphorylation of S41 disrupts a hydrogen bonding network centred on Q44.

In contrast, aspartate substitution of S131 increased the ESC self-renewal capacity of NANOG without

affecting DNA binding. The results from self-renewal assays demonstrate that a phosphomimetic at position S131 has a large effect on NANOG function, increasing its ability to support LIF-independent self-renewal several fold. It will be interesting in future to see whether S131D also enhances reprogramming to an ES cell state [52–54]. In the model of the NANOG–DNA complex, the side chain of serine 131 points away from the DNA. This is consistent with the lack of an effect on DNA binding for S131D as assayed by EMSA. Therefore, instead of mediating protein–DNA interactions S131 may mediate interaction with NANOG partner proteins. Although NANOG is known to interact with > 130 partner proteins [22,55,56], the residues within NANOG that mediate these interactions have been identified in only a minority of cases [22,57]. The fact that modification of S131 results in an increase in activity that is independent of DNA binding suggests that reversible modification of the NANOG homeodomain may be used as a switch to regulate interactions between NANOG and partner proteins and to thereby modulate ESC self-renewal efficiency. Future work should address this issue and the alternative possibility that S131D affects protein stability as is the case for the SOX17 gain-of-function mutant SOX17 V118M [58].

Acknowledgements

Research in IC's lab is supported by the Medical Research Council and the Biotechnology and Biological Sciences Research Council of the UK. Work in GF's laboratory is supported by a Medical Research Council New Investigator Research Grant and a Wellcome Trust/Royal Society Sir Henry Dale Fellowship. JV is supported by the Medical Research Council Protein Phosphorylation and Ubiquitylation Unit core grant.

Author contributions

IC and NPM conceived the project, designed the experiments and wrote the manuscript. NPM performed phosphorylation assays, constructed mutants and performed EMSAs and ChIP assays. GMF and JV performed phosphorylation analysis. JMR modelled the NANOG–DNA structure. DC performed cell culture.

References

- Whitmarsh AJ and Davis RJ (2000) Regulation of transcription factor function by phosphorylation. *Cell Mol Life Sci* **57**, 1172–1183.

- Ozcan S, Andrali SS and Cantrell JE (2010) Modulation of transcription factor function by O-GlcNAc modification. *Biochim Biophys Acta* **1799**, 353–364.
- Das C and Kundu TK (2005) Transcriptional regulation by the acetylation of nonhistone proteins in humans – a new target for therapeutics. *IUBMB Life* **57**, 137–149.
- Hammond-Martel I, Yu H and el Affar B (2012) Roles of ubiquitin signaling in transcription regulation. *Cell Signal* **24**, 410–421.
- Levy DE and Darnell JE Jr (2002) Stats: transcriptional control and biological impact. *Nat Rev Mol Cell Biol* **3**, 651–662.
- Luscher B, Christenson E, Litchfield DW, Krebs EG and Eisenman RN (1990) Myb DNA binding inhibited by phosphorylation at a site deleted during oncogenic activation. *Nature* **344**, 517–522.
- Chen D, Pace PE, Coombes RC and Ali S (1999) Phosphorylation of human estrogen receptor alpha by protein kinase A regulates dimerization. *Mol Cell Biol* **19**, 1002–1015.
- Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S and Smith A (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643–655.
- Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M and Yamanaka S (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631–642.
- Niwa H, Miyazaki J and Smith AG (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**, 372–376.
- Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov AA *et al.* (2007) Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* **9**, 625–635.
- Mullin NP and Chambers I (2012) The function of Nanog in pluripotency. In *Nuclear Reprogramming and Stem Cells. Stem Cell Biology and Regenerative Medicine* (Ainscough J, Yamanaka S and Tada T, eds), pp. 99–112. Springer, Totowa, NJ.
- Wong FCK, Chambers I and Mullin NP (2016) Sox2-dependent regulation of pluripotent stem cells. In *Sox2: Biology and Role in Development and Disease* (Kondoh H and Lovell-Badge R, eds), pp. 163–185. Elsevier, London, UK.
- Navarro P, Festuccia N, Colby D, Gagliardi A, Mullin NP, Zhang W, Karwacki-Neisius V, Osorno R, Kelly D, Robertson M *et al.* (2012) OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. *EMBO J* **31**, 4547–4562.

- 15 Li QR, Xing XB, Chen TT, Li RX, Dai J, Sheng QH, Xin S-M, Zhu L-L, Jin Y, Pei G *et al.* (2011) Large scale phosphoproteome profiles comprehensive features of mouse embryonic stem cells. *Mol Cell Proteomics* **10**, M110 001750.
- 16 Brumbaugh J, Hou Z, Russell JD, Howden SE, Yu P, Ledvina AR, Coon JJ and Thomson JA (2012) Phosphorylation regulates human OCT4. *Proc Natl Acad Sci USA* **109**, 7162–7168.
- 17 Jeong CH, Cho YY, Kim MO, Kim SH, Cho EJ, Lee SY, Yao K, Keum YS, Bode AM and Dong Z (2010) Phosphorylation of Sox2 cooperates in reprogramming to pluripotent stem cells. *Stem Cells* **28**, 2141–2150.
- 18 Kim SH, Kim MO, Cho YY, Yao K, Kim DJ, Jeong CH, Yu DH, Bae KB, Cho EJ, Jung SK *et al.* (2014) ERK1 phosphorylates Nanog to regulate protein stability and stem cell self-renewal. *Stem Cell Res* **13**, 1–11.
- 19 Moretto-Zita M, Jin H, Shen Z, Zhao T, Briggs SP and Xu Y (2010) Phosphorylation stabilizes Nanog by promoting its interaction with Pin1. *Proc Natl Acad Sci USA* **107**, 13312–13317.
- 20 Jang H, Kim TW, Yoon S, Choi SY, Kang TW, Kim SY, Kwon Y-W, Cho E-J and Youn H-D (2012) O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. *Cell Stem Cell* **11**, 62–74.
- 21 Baltus GA, Kowalski MP, Zhai H, Tutter AV, Quinn D, Wall D and Kadam S (2009) Acetylation of Sox2 induces its nuclear export in embryonic stem cells. *Stem Cells* **27**, 2175–2184.
- 22 Gagliardi A, Mullin NP, Ying Tan Z, Colby D, Kousa AI, Halbritter F, Weiss JT, Felker A, Bezstarosti K, Favaro R *et al.* (2013) A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell self-renewal. *EMBO J* **32**, 2231–2247.
- 23 Illich DJ, Zhang M, Ursu A, Osorno R, Kim KP, Yoon J, Araúzo-Bravo MJ, Wu G, Esch D, Sabour D *et al.* (2016) Distinct signaling requirements for the establishment of ESC pluripotency in late-stage EpiSCs. *Cell Rep* **15**, 787–800.
- 24 Knippschild U, Gocht A, Wolff S, Huber N, Lohler J and Stoter M (2005) The casein kinase I family: participation in multiple cellular processes in eukaryotes. *Cell Sig* **17**, 675–689.
- 25 Kalousi A, Mylonis I, Politou AS, Chachami G, Paraskeva E and Simos G (2010) Casein kinase I regulates human hypoxia-inducible factor HIF-1. *J Cell Sci* **123**(Pt 17), 2976–2986.
- 26 Okamura H, Garcia-Rodriguez C, Martinson H, Qin J, Virshup DM and Rao A (2004) A conserved docking motif for CK1 binding controls the nuclear localization of NFAT1. *Mol Cell Biol* **24**, 4184–4195.
- 27 Cruciat CM (2014) Casein kinase I and Wnt/beta-catenin signaling. *Curr Opin Cell Biol* **31**, 46–55.
- 28 Chen L, Li C, Pan Y and Chen J (2005) Regulation of p53-MDMX interaction by casein kinase I alpha. *Mol Cell Biol* **25**, 6509–6520.
- 29 Knippschild U, Milne DM, Campbell LE, DeMaggio AJ, Christenson E, Hoekstra MF and Meek DW (1997) p53 is phosphorylated in vitro and in vivo by the delta and epsilon isoforms of casein kinase I and enhances the level of casein kinase I delta in response to topoisomerase-directed drugs. *Oncogene* **15**, 1727–1736.
- 30 Songyang Z, Lu KP, Kwon YT, Tsai LH, Filhol O, Cochet C, Brickey DA, Soderling TR, Bartleson C, Graves DJ *et al.* (1996) A structural basis for substrate specificities of protein Ser/Thr kinases: primary sequence preference of casein kinases I and II, NIMA, phosphorylase kinase, calmodulin-dependent kinase II, CDK5, and Erk1. *Mol Cell Biol* **16**, 6486–6493.
- 31 Meggio F, Donella-Deana A and Pinna LA (1979) Studies on the structural requirements of a microsomal cAMP-independent protein kinase. *FEBS Lett* **106**, 76–80.
- 32 Mullin NP, Yates A, Rowe AJ, Nijmeijer B, Colby D, Barlow PN, Walkinshaw MD and Chambers I (2008) The pluripotency rheostat Nanog functions as a dimer. *Biochem J* **411**, 227–231.
- 33 Taus T, Kocher T, Pichler P, Paschke C, Schmidt A, Henrich C and Mechtler K (2011) Universal and confident phosphorylation site localization using phosphoRS. *J Proteome Res* **10**, 5354–5362.
- 34 Robertson M, Stenhouse F, Colby D, Marland JR, Nichols J, Tweedie S and Chambers I (2006) Nanog retrotransposed genes with functionally conserved open reading frames. *Mamm Genome* **17**, 732–743.
- 35 Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L and Smith A (2007) Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230–1234.
- 36 Jauch R, Ng CK, Saikatendu KS, Stevens RC and Kolatkar PR (2008) Crystal structure and DNA binding of the homeodomain of the stem cell transcription factor Nanog. *J Mol Biol* **376**, 758–770.
- 37 Novo CL, Tang C, Ahmed K, Djuric U, Fussner E, Mullin NP, Morgan NP, Hayre J, Sieners AR, Elderkin S *et al.* (2016) The pluripotency factor Nanog regulates pericentromeric heterochromatin organization in mouse embryonic stem cells. *Genes Dev* **30**, 1101–1115.
- 38 Navarro P, Moffat M, Mullin NP and Chambers I (2011) The X-inactivation trans-activator Rnf12 is negatively regulated by pluripotency factors in embryonic stem cells. *Hum Genet* **130**, 255–264.
- 39 Karwacki-Neisius V, Goke J, Osorno R, Halbritter F, Ng JH, Weisse AY, Wong FCK, Gagliardi A, Mullin NP, Festuccia N *et al.* (2013) Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. *Cell Stem Cell* **12**, 531–545.

- 40 Robert X and Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**(Web Server issue), W320–W324.
- 41 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**(Pt 4), 235–242.
- 42 Ursu A, Illich DJ, Takemoto Y, Porfetye AT, Zhang M, Brockmeyer A, Janning P, Watanabe N, Osada H, Vetter IR *et al.* (2016) Epiblastin A induces reprogramming of epiblast stem cells into embryonic stem cells by inhibition of casein kinase I. *Cell Chem Biol* **23**, 494–507.
- 43 Rena G, Bain J, Elliott M and Cohen P (2004) D4476, a cell-permeant inhibitor of CK1, suppresses the site-specific phosphorylation and nuclear exclusion of FOXO1a. *EMBO Rep* **5**, 60–65.
- 44 Gehring WJ, Affolter M and Burglin T (1994) Homeodomain proteins. *Annu Rev Biochem* **63**, 487–526.
- 45 Saunders A, Li D, Faiola F, Huang X, Fidalgo M, Guallar D, Ding J, Yang F, Xu Y, Zhou H *et al.* (2017) Context-dependent functions of NANOG phosphorylation in pluripotency and reprogramming. *Stem Cell Rep* **8**, 1115–1123.
- 46 Xie X, Piao L, Cavey GS, Old M, Teknos TN, Mapp AK and Pan Q (2014) Phosphorylation of Nanog is essential to regulate Bmi1 and promote tumorigenesis. *Oncogene* **33**, 2040–2052.
- 47 Hayashi Y, Caboni L, Das D, Yumoto F, Clayton T, Deller MC, Nguyen P, Farr CL, Chiu H-J, Miller MD *et al.* (2015) Structure-based discovery of NANOG variant with enhanced properties to promote self-renewal and reprogramming of pluripotent stem cells. *Proc Natl Acad Sci USA* **112**, 4666–4671.
- 48 Soufi A, Noy P, Buckle M, Sawasdichai A, Gaston K and Jayaraman PS (2009) CK2 phosphorylation of the PRH/Hex homeodomain functions as a reversible switch for DNA binding. *Nucleic Acids Res* **37**, 3288–3300.
- 49 Soufi A, Smith C, Clarke AR, Gaston K and Jayaraman PS (2006) Oligomerisation of the developmental regulator proline rich homeodomain (PRH/Hex) is mediated by a novel proline-rich dimerisation domain. *J Mol Biol* **358**, 943–962.
- 50 Kasahara H and Izumo S (1999) Identification of the in vivo casein kinase II phosphorylation site within the homeodomain of the cardiac tissue-specifying homeobox gene product Csx/Nkx2.5. *Mol Cell Biol* **19**, 526–536.
- 51 Dong J, Hung LH, Strome R and Krause HM (1998) A phosphorylation site in the ftz homeodomain is required for activity. *EMBO J* **17**, 2308–2318.
- 52 Theunissen TW, Costa Y, Radzisheuskaya A, van Oosten AL, Laval F, Pain B, Castro LFC and Silva JCR (2011) Reprogramming capacity of Nanog is functionally conserved in vertebrates and resides in a unique homeodomain. *Development* **138**, 4853–4865.
- 53 Wang B, Wu L, Li D, Liu Y, Guo J, Li C, Yao Y, Wang Y, Zhao G, Wang X *et al.* (2019) Induction of pluripotent stem cells from mouse embryonic fibroblasts by Jdp2-Jhdm1b-Mkk6-Glis1-Nanog-Essrb-Sall4. *Cell Rep* **27**, 3473–3485.e5.
- 54 Festuccia N, Osorno R, Halbritter F, Karwacki-Neisius V, Navarro P, Colby D, Wong F, Yates A, Tomlinson SR and Chambers I (2012) Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell* **11**, 477–490.
- 55 Liang J, Wan M, Zhang Y, Gu P, Xin H, Jung SY, Qin J, Wong J, Cooney AJ, Liu D *et al.* (2008) Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nat Cell Biol* **10**, 731–739.
- 56 Costa Y, Ding J, Theunissen TW, Faiola F, Hore TA, Shliha PV, Fidalgo M, Saunders A, Lawrence M, Dietmann S *et al.* (2013) NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* **495**, 370–374.
- 57 Mullin NP, Gagliardi A, Khoa LT, Colby D, Hall-Ponsole E, Rowe AJ and Chambers I (2016) Distinct contributions of tryptophan residues within the dimerization domain to Nanog function. *J Mol Biol* **429**, 1544–1553.
- 58 Srivastava Y, Tan DS, Malik V, Weng M, Javed A, Cojocar V, Wu G, Veerapandian V, Cheung LWT and Jauch R (2020) Cancer-associated missense mutations enhance the pluripotency reprogramming activity of OCT4 and SOX17. *FEBS J* **287**, 122–144.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig S1. Replicate self-renewal assays of E14/T cells transfected with NANOG variants.