

RAPID COMMUNICATION

## Identification of two distinct transactivation domains in the pluripotency sustaining factor nanog

GUANG Jin PAN<sup>1</sup>, DUAN QING PEI<sup>1,2,\*</sup>

<sup>1</sup>*Institute of Pharmacology, Department of Biological Sciences and Biotechnology and State Key Laboratory of Biomembrane and Membrane Biotechnology, Tsinghua Institutes of Biomedicine, Tsinghua University, 100084 Beijing, China. E-mail: duanqing@aol.com*

<sup>2</sup>*Department of Pharmacology, University of Minnesota, Minneapolis, MN55455, USA*

### ABSTRACT

Nanog is a newly identified homeodomain gene that functions to sustain the pluripotency of embryonic stem cells. However, the molecular mechanism through which nanog regulates stem cell pluripotency remains unknown. Mouse nanog encodes a polypeptide of 305 residues with a divergent homeodomain similar to those in the NK-2 family. The rest of nanog contains no apparent homology to any known proteins characterized so far. It is hypothesized that nanog encodes a transcription factor that regulates stem cell pluripotency by switching on or off target genes. To test this hypothesis, we constructed fusion proteins between nanog and DNA binding domains of the yeast transcription factor Gal4 and tested the transactivation potentials of these constructs. Our data demonstrate that both regions N- and C- terminal to the homeodomain have transcription activities. Despite the fact that it contains no apparent transactivation motifs, the C-terminal domain is about 7 times as active as the N-terminal one. This unique arrangement of dual transactivators may confer nanog the flexibility and specificity to regulate downstream genes critical for both pluripotency and differentiation of stem cells.

**Key words:** *nanog, pluripotency, stem cells, self renewal, transactivation domain, homeoprotein, homeodomain*

### INTRODUCTION

The pluripotency of embryonic stem cells appears to be under the regulation of multiple transcription factors [1-3]. Oct-4 is recognized as a critical factor in maintaining the self renewing ability of embryonic stem cells, perhaps through the regulation of a plethora of downstream genes both positively and negatively[1-3]. In addition, transcription factors STAT3 has also been implicated in regulating the stemness of stem cells[2, 4]. However, neither STAT3 nor Oct-4 is sufficient to maintain stem cell pluripotency[5, 6]. To this end, one novel homeoprotein named nanog has recently been identified as an intrinsic factor for ES cell self renewal[6-8].

Like Oct-4, the expression of nanog mRNA is restricted to pluripotent stem cells and absent in differentiated ones[7, 8]. Nanog apparently operates in parallel with STAT3 and is sufficient to maintain stem cell pluripotency without gp130/ STAT3 activation[7, 8].

It has been further suggested that nanog is the missing determinant for pluripotency of inner cell mass (ICM) and embryonic stem (ES) cells[7]. While Oct-4 functions to prevent the differentiation of ICM and ES cells into trophectoderm, nanog not only blocks the differentiation into endoderm independently but also actively maintains pluripotency[7]. Currently, the molecular mechanism through which nanog performs these two distinct function remains entirely unknown. Mitsui and colleagues suggest that nanog maintains self renewal of ES cells by transcriptional repression of those genes critical for differentiation[7]. Indeed, either gata4 or gata6 is elevated in nanog null ES cells[7]. Both gata4 or gata6 can trigger extraembryonic endoderm differentiation of ES cells transfected with either factor[9]. A direct link between nanog and gata6 became apparent when a nanog consensus sequence identified through SELEX selection was found in the enhancer region of gata6[7]. Therefore, it is plausible that nanog maintains pluripotency by directly repressing downstream genes such as gata4 or gata6[7]. Yet, little is known about the transactivating

\*Corresponding author: Duan Qing PEI, Ph.D.

E-mail: duanqing@aol.com

Received: Nov-20-2003 Revised: Nov-23-2003 Accepted: Nov-25-2003

or repressing potential of nanog at the present.

Mouse nanog encodes a protein of 305 amino acid residues which can be roughly divided into three regions based on the position of the NK-2 type homeodomain [7, 8]. The 60 residue homeodomain is expected to perform protein-protein interactions and DNA binding as demonstrated for Oct-4 [1]. The region N-terminal to the homeodomain has 95 residues rich in Ser and Thr and acidic residues found in typical transactivators [7, 8, 10]. On the other hand, the C-terminal region of 150 residues have no apparent transactivation motifs, but contains a prominent W-repeat [7, 8]. Therefore, it is plausible that the N-terminal domain of nanog may mediate trans-activation while the C-terminal domain mediates trans-criptional repression. While the repression function of a transcription factor is hard to establish, positive transactivation domains can be identified and characterized by the well-established Gal4 based reporter system [10]. To this end, we have performed fusion experiments between Gal4 DNA binding domain and individual ones from nanog. We report here that both the N- and C-terminal domains have transactivation activities. To our surprise, despite the lack of any known features for transactivators, the C-terminal domain of nanog is a much stronger activator than its N-terminal counterpart.

## MATERIALS AND METHODS

### *Effector constructs*

Mouse Nanog cDNA was isolated by RT-PCR using high fidelity polymerase pfx (Invitrogen, CA) from P19 cells. The primers were: accatgagtgtgggtcttctctgtcc for the 5' end and tattcaccgtgtggagtcacagagt for the 3' end. The mNanog cDNAs were inserted into the EcoRV site of modified pCR3.1 [11]. pCR3.1-Gal4DBD was created from a PCR fragment encoding the Gal4 DNA binding domain (residues 1 to 147). An EcoRV site was regenerated in the upstream or downstream of Gal4 DBD for inserting various domains of Nanog. A flag tag was fused N terminally or C terminally to the Gal4 DBD for proteins detection. pCR3.1-ND-Gal4 was created by inserting a PCR fragment encoding Nanog N domain (residue 1-95) to the upstream EcoRV site of pCR3.1-Gal4 DBD. pCR3.1-Gal4-CD and pCR3.1-Gal4-HD were generated by inserting a PCR fragment encoding nanog C domain (residues 156-305) or the nanog homeodomain (residues 96-155) to the downstream EcoRV site of pCR3.1-Gal4 DBD respectively.

### *Cell lysates preparation and western blot analysis*

HEK293T cell lysates were harvested 48h after transfection. Briefly, after washed with PBS, cells were lysed on ice with 250  $\mu$ l of RIPA buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.25% Sodium deoxycholate, 0.1% Nonidet P-40, 0.1% Triton X-100) for

10 min and cleared of debris by centrifugation at 15000 rpm for 15 min at 4°C. Cell lysates were electrophoresed through 10% SDS-PAGE and electro-blotted to PVDF membranes (Millipore). The membranes were blocked in 5% non-fat milk and incubated with anti-flag antibody (1:2000, Sigma, MO). Alkaline phosphatase-conjugated anti-mouse antibody (1:2500, Pierce) was used as 2<sup>nd</sup> antibody and NBT/BCIP as substrates for detection as described [11].

### *Cell culture and reporter assays*

HEK293T cells were cultured in DMEM (Invitrogen, CA) supplemented with 10% FBS (Hyclone, UT) and 100 mg/ml antibiotics (penicillin and streptomycin) as described [11]. For reporter assays, HEK293T cells were seeded in 12-well cell culture plate and transfected by calcium phosphate co-precipitation methods with 0.2  $\mu$ g of p5G-e1b-luciferase and 1  $\mu$ g of each expression vector. Transfection efficiencies were normalized by co-transfection with Renilla (Promega, WI) in each well and DNA concentrations were normalized by using pCR3.1 empty vector. Cells were harvested 36 h after transfection, and luciferase activity was measured by using Dual-luciferase Reporter Assay system (Promega, WI) according to the manufacture's instructions. Each transfection was carried out in duplicate and repeated at least twice.

## RESULTS AND DISCUSSION

To test our hypothesis that nanog contains distinct transcription domains, we operationally divided its open reading frame into three subdomains: the 95 residue N-terminal domain (ND), the 60 residue homeodomain (HD) and the 150 residue C-terminal domain (CD) (Fig 1A). Since the cognate DNA binding sequence for nanog remains largely unknown, we employed the well-established Gal-4 fusion/reporter system to identify any transactivation domains in nanog as described previously [10]. We fused the ND, HD and CD of nanog to gal-4 DNA binding domain as illustrated in Fig 1B. The resulting plasmids were confirmed by sequencing and the corresponding fusion proteins characterized by Western blotting with the anti-FLAG monoclonal antibody M2 (Fig 1C). The transactivation potentials of these fusions were then assessed using a reporter plasmid containing 5 copies of Gal-4 DNA binding sites upstream of the E1b TATA box. As shown in Fig 1D, both Gal4-ND and Gal4-CD plasmids transactivated the reporter efficiently, resulting over 35 or 240 folds of activation over the gal4 construct respectively in HEK293T cells. We also tested both plasmids in HeLa, NIH3T3, and P19 cells and observed transactivation activities for both plasmids (data not shown). The activation potential of ND is consistent with the fact that it is rich in acidic and Ser/Thr residues, hallmarks of transactivation domains [7, 8]. On the other hand, we were surprised to detect

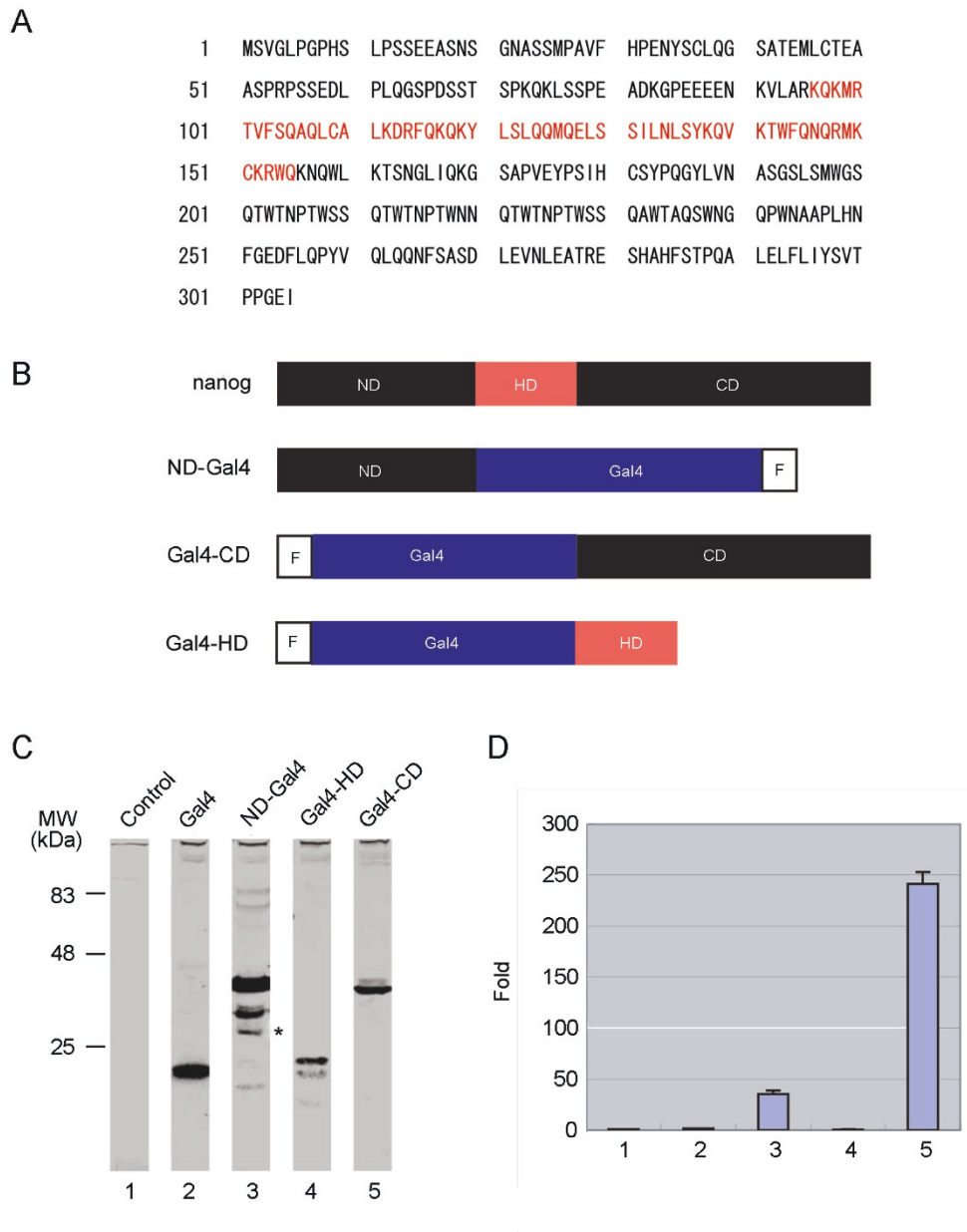


Fig 1: Identification of two distinct transactivation domains in nanog. **(A)** Division of nanog into three distinct domains at the amino acid level. The homeodomain is marked in red. **(B)** Schematic illustrations of fusion constructs between Gal-4 DNA binding domain and the three subdomains of nanog. The FLAG tag is located at the C-terminal end for the ND-Gal4 fusion, but at the N-termini for the Gal4-CD and Gal4-HD. **(C)** Western blotting analysis of fusion constructs. HEK293T cells were transfected with control expression vector (lane 1), Gal4 DNA binding domain alone (lane 2), ND-Gal4 (lane 3), Gal4-CD (lane 4) and Gal4-HD (lane 5) as described in Materials and Methods. Cell lysates were fractionated, blotted and probed with anti-FLAG antibody. The \* at the right of lane 3 depicts the expected position based on its calculated molecular weight. **(D)** Transactivation activities of fusion constructs. The reporter plasmid, p5G-E1b-luciferase (0.2  $\mu$ g), were co-transfected with control vector (1  $\mu$ g, lane 1), pCR3.1-Gal4DBD (1 $\mu$ g, lane 2), pCR3.1-ND-Gal4 (1  $\mu$ g, lane 3), pCR3.1-Gal4-HD (1 $\mu$ g, lane 4) or pCR3.1-Gal4-CD (1 $\mu$ g, lane 5) to HEK293T cells in 12 well plates. The luciferase activity was assayed 36 h after transfection using the Dual-reporter assay system. The results were the average of at least two independent transfection experiments each with duplicates.

robust activation function in CD given its lack of any discernable transactivation motifs. The fact that CD is almost seven times as active as ND suggests a dominant role for CD in mediating activation of downstream gene.

The data presented here establish for the first time that nanog possesses two independent transactivation domains. Oct-4, another homeodomain protein implicated in stem cell pluripotency [1, 2, 12, 13], also contains two separate transactivation domains with no apparent similarity to the ones we defined in nanog. The divergent structures associated with these activation domains may confer the specificity of regulation for the target genes important for stem cell pluripotency. The ND of nanog is rich in Ser and Thr, suggesting that it may be regulated by phosphorylation. Indeed, western blotting analysis of ND-Gal4 indicates that its apparent molecular weight is 5 kDa larger than predicted, perhaps, as a result of phosphorylation at some of the Ser and Thr residues. More experiments are needed to map the sites of phosphorylation and identify the kinases responsible for the phosphorylations. On the other hand, the CD of nanog is entirely novel without any homology or similarity to other transactivation domains. Taken together, our data presented in this communication definitively prove that nanog contains two distinct transactivation domains with novel structural features. These findings may lead to the eventual elucidation of nanog's role in sustaining stem cell pluripotency at the molecular levels.

#### ACKNOWLEDGEMENTS

This work was supported in part by the Tsinghua University BaiRen Scholar Program, NSFC 30270287, and the 973 Project--2001CB5101 from The Ministry of Science and Technology of China. DP is a CheungKong Scholar of the Ke Shing LI Foundation. The authors wish to acknowledge the kind support of Professors Nan Ming ZHAO and Zhijie CHANG at Tsinghua Univer-

and the generous assistance from members of the Pei laboratory during this study. The authors would like to thank Dr. Sudhof of HHMI, UT southwestern, Dallas, TX for the generous gift of p5G-Elb-Luc.

#### REFERENCE

- 1 Pan GJ, Chang ZY, Scholer HR, Pei D. Stem cell pluripotency and transcription factor Oct4. *Cell Res* 2002; **12**:321-9.
- 2 Pesce M, Scholer HR. Oct-4: gatekeeper in the beginnings of mammalian development. *Stem Cells* 2001; **19**:271-8.
- 3 Constantinescu S. Stemness, fusion and renewal of hematopoietic and embryonic stem cells. *J Cell Mol Med* 2003; **7**:103-12.
- 4 Guo Y, Costa R, Ramsey H, Starnes T, Vance G, Robertson K, Kelley M, Reinbold R, Scholer H, Hromas R. The embryonic stem cell transcription factors Oct-4 and FoxD3 interact to regulate endo dermal-specific promoter expression. *Proc Natl Acad Sci USA* **99**: 3663-7.
- 5 Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 2000; **24**:372-6.
- 6 Cavaleri F, Scholer HR. Nanog: a new recruit to the embryonic stem cell orchestra. *Cell* 2002; **113**:551-2.
- 7 Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* 2003; **113**:631-42.
- 8 Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* 2003; **113**:643-55.
- 9 Fujikura J, Yamato E, Yonemura S, Hosoda K, Masui S, Nakao K, Miyazaki Ji J, Niwa H. Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev* 2002; **16**:784-9.
- 10 Pei DQ, Shih CH. An "attenuator domain" is sandwiched by two distinct transactivation domains in the transcription factor C/EBP. *Mol Cell Biol* 1991; **11**:1480-7.
- 11 Jiang A, Pei D. Distinct roles of catalytic and pexin-like domains in membrane-type matrix metalloproteinase (MMP)-mediated pro-MMP-2 activation and collagenolysis. *J Biol Chem* 2003; **278**:38765-71.
- 12 Niwa H, Burdon T, Chambers I, Smith A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* 1998; **12**:2048-60.
- 13 Pesce M, Scholer HR. Oct-4: control of totipotency and germline determination. *Mol Reprod Dev* 2000; **55**:452-7.