Characterization of human UTF1, a chromatin-associated protein with repressor activity expressed in pluripotent cells

Susanne M. Kooistra, Rajkumar P. Thummer, Bart J.L. Eggen*

Developmental Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands

Received 25 November 2008; received in revised form 27 January 2009; accepted 4 February 2009

Abstract In mice, during early embryonic development UTF1 (undifferentiated embryonic cell transcription factor 1) is expressed in the inner cell mass of blastocysts and in adult animals expression is restricted to the gonads. (Embryonic) Cells expressing UTF1 are generally considered pluripotent, meaning they can differentiate into all cell types of the adult body. In mouse it was shown that UTF1 is tightly associated with chromatin and that it is required for proper differentiation of embryonic carcinoma and embryonic stem cells. In this study we functionally characterized the human UTF1 protein. We show with localization, subnuclear fractionation, and strip-FRAP analyses that human UTF1 is a tightly DNA-associated protein with transcriptional repressor activity. Our data identify human UTF1 as a pluripotency-associated chromatin component with core histone-like characteristics.

Introduction

Embryonic stem (ES) cells are cell lines derived from the inner cell mass of blastocyst embryos. ES cells have the ability to undergo unlimited symmetric cell divisions (self-renewal) and the capacity to differentiate into all fetal and adult cell types (pluripotency). Especially their pluripotent character has generated high expectations that they might be used for cell replacement therapies. The molecular mechanisms regulating ES cell self-renewal are relatively well known, but how the transition from self-renewal to differentiation is controlled is much less understood.

Several factors involved in pluripotency and/or the induction and regulation of differentiation have been identified. In mice, these include external signals like leukemia inhibitory factor (LIF) and bone morphogenetic protein (BMP) (Smith et al., 1988; Williams et al., 1988; Ying et al., 2003). Self-renewal of human ES cells is controlled extrinsically by basic fibroblast growth factor (bFGF) and suppression of BMP signaling (Xu et al., 2005).

In human and mouse ES cells, a core transcriptional regulatory circuit was identified (Boyer et al., 2005; Loh et al., 2006) comprising the transcription factors OCT4, SOX2, and NANOG. OCT4 and NANOG were the first proteins identified that are required for normal embryonic development as well
as maintaining pluripotency in ES cells (Nichols et al., 1998; Chambers et al., 2003; Mitsui et al., 2003). In human ES cells these proteins were found to co-occupy, together with SOX2, the promoter regions of genes involved in maintaining pluripotency. Additionally, they were present on the promoters of many developmentally important transcription factors, thereby contributing to both self-renewal and pluripotency (Boyer et al., 2005).

In addition to this network of transcription factors, also the importance of epigenetic mechanisms has been recognized (reviewed in (Spivakov and Fisher, 2007; Chen and Daley, 2008; Bibikova et al., 2008; Pierssen and van Lohuizen, 2008)). It has become clear that ES cells differ from their committed progeny in abundance of modified histones, polycomb group binding patterns, replication timing, and chromatin accessibility (Azuara et al., 2006; Bernstein et al., 2006; Perry et al., 2004; Hiratani et al., 2004; Mesheror et al., 2006), (reviewed in (Spivakov and Fisher, 2007)).

Mouse ES cell chromatin has been reported to be in general more accessible (Wiblin et al., 2005; Williams et al., 2006; Keohane et al., 1996), or hyperdynamic, a property thought not only relevant for maintaining pluripotency but also essential in the early stages of ES cell differentiation for reshaping the global architecture of the genome (Mesheror et al., 2006).

Possibly, specific chromatin components are present in ES cells that are responsible for maintaining a chromatin state that allows for self-renewal while maintaining the capability of differentiation. Recently we have shown that in mouse ES cells the undifferentiated embryonic cell transcription factor 1 (UTF1) protein might be such a factor. UTF1 is expressed early during embryonic development in the cells of the ICM and epiblast (Okuda et al., 1998). During development its expression is rapidly down regulated (Okuda et al., 1998), but it is maintained in the primordial germ cells of the developing embryo (Chuva de Sousa Lopes et al., 2005). In adults, expression of UTF1 could be detected in the gonads (Okuda et al., 1998; Chuva de Sousa Lopes et al., 2005; van Bragt et al., 2008; Kristensen et al., 2008).

In mouse, UTF1 is involved in maintaining the proliferation rate and teratoma formation of ES cells (Nishimoto et al., 2005; van den Boom et al., 2007). In human ES cells, UTF1 has also been implied to be important for their proliferation rate (Li et al., 2007). For mouse UTF1 we have shown that it is a tightly chromatin-associated protein with dynamics similar to those of core histones (van den Boom et al., 2007). Although its expression is dispensable for self-renewal, ES cells with reduced UTF1 levels failed to differentiate properly. These data indicate a possible role for UTF1 in the maintenance of a specific epigenetic profile that is required for differentiation of mouse ES cells (van den Boom et al., 2007). This is further supported by a recent observation by Zhao and co-workers who reported that the efficiency of induced pluripotent stem cell (iPS) generation increased approximately 100-fold by the co-expression of UTF1 and siRNAs against p53 with C-MYC, KLF-4, OCT4, and SOX2. Remarkably, UTF1 could replace C-MYC and enhance the efficiency of iPS generation by 10-fold (Zhao et al., 2008).

In this report we describe the characterization of the human UTF1 protein. The subcellular localization and sub-nuclear fractionation of UTF1 in human EC cells, in combination with strip-FRAP, indicated that human UTF1 is a stably chromatin-associated protein with a mobility similar to that of core histones. In luciferase reporter assays human UTF1 displayed transcriptional repressor activity, for which a conserved C-terminal domain is required.

**Results**

**Human UTF1 is tightly associated to chromatin**

From studies on the mouse protein, it has become clear that mUTF1 is a nuclear protein with biochemical characteristics similar to those of core histones (van den Boom et al., 2007). NCCIT cells, a human teratocarcinoma cell line, were used to study the localization of human UTF1 (hUTF1). RT-PCR confirmed expression of the hUTF1 gene in NCCIT cells whereas expression was not detected in differentiated NCCIT cells (8 days of 10 μM retinoic acid) (Fig. 1C). Using immunofluorescence, the localization of hUTF1 was determined in NCCIT cells (Fig. 1A). In these cells, hUTF1 has an inhomogeneous nuclear localization, and it is excluded from the nucleoli. It colocalizes with DNA (visualized by DAPI) in interphase cells, but also during mitosis hUTF1 remains colocalized with chromatin (Fig. 1A). To study the association of hUTF1 with chromatin, hUTF1 was fused to enhanced GFP (GFP-hUTF1) and stably expressed in P19CL6 embryonic carcinoma cells (Habara-Ohkubo, 1996). GFP-hUTF1 in P19CL6 cells showed a localization (Fig. 1B) similar to that of the endogenous protein in NCCIT cells (Fig. 1A). GFP-hUTF1 localization in the nucleus is inhomogeneous (Fig. 1B); it is excluded from the nucleoli and associated with chromatin during mitosis (insets Fig. 1B). This localization of GFP-hUTF1 was also observed in living cells (Fig. 1D).

NCCIT cells were subjected to subnuclear fractionation (Fig. 1E), during which free diffusing nuclear and cytoplasmic proteins (F), weak (D) and strong (AS) DNA-associated proteins, and nuclear matrix proteins (HS and M) are separated. In this assay, the endogenous hUTF1 protein fractionated almost exclusively to the ammonium sulfate fraction which contains tightly DNA-associated proteins like core histone H2A (Fig. 1E). In contrast, chromatin-modifying proteins like HDAC1 and mSin3A primarily fractionated to the free diffusing and weakly DNA-associated protein fractions. Moreover, transcription factors Oct4, A T F -2, and TFIID were detected mainly in the free diffusing protein fraction.

In mouse P19CL6 cells, a hUTF1-GFP fusion protein also fractionated to the ammonium sulfate fraction (Fig. 1F). The endogenously expressed mouse UTF1 protein was used as a fractionation control and was exclusively detected in the ammonium sulfate fraction.

**Human UTF1 is a transcriptional repressor**

To study the effect of hUTF1 on transcription we used a GAL4 reporter assay. In this assay hUTF1 is inserted C terminal of the DNA-binding domain of GAL4. The effect on transcription was measured on a luciferase reporter construct containing a constitutively active thymidine Kinase (TK) promoter under control of 5 copies of the GAL4 DNA-binding element, the upstream activating sequence (UAS). A LacZ expression plasmid, pDM2-LacZ, was cotransfected and used as an
internal standard. Repression by mUTF1 and hUTF1 is indicated relative to the luciferase activity in the presence of the GAL4 DNA-binding domain alone. In this assay, mUTF1 fused to GAL4 (m1–339) repressed transcription 8.7 +/- 0.9-fold whereas hUTF1 (h2–341) repressed transcription approximately 6.4 +/- 0.4-fold (Fig. 2).

hUTF1 contains two evolutionarily conserved domains: CD1 which contains high homology to Myb/SANT domains (aa52–167) and CD2 which contains a putative leucine zipper (aa271–334) (Fukushima et al., 1998). To study the role of both these domains in the observed repressor activity, a series of C- and N-terminal deletion mutants was generated (Fig. 2).

A double leucine to proline point mutation in the putative leucine zipper, L293P and L300P, did not affect repressor activity (6.7 +/- 0.9-fold compared to 6.4 +/- 0.4-fold of wild-type hUTF1). Deletion of the CD2 domain (constructs h2–178 and h2–126) resulted in drastically reduced repressor activity (2.2-fold +/- 0.2). When both CD1 and CD2 were deleted (h2–26) the luciferase activity was identical to that of GAL4 alone.

C-terminal deletion of CD1 (constructs h141–341 and h256–341) did not result in decreased repressor activity. These data indicate that hUTF1 can act as a transcriptional repressor and that the CD2 domain is sufficient and required for transcriptional repression. Furthermore, the leucine...
residues at positions 293 and 300 do not seem to be involved in the repressor activity of hUTF1.

**Dynamics of human UTF1 and the role of its conserved domains**

To study the localization and dynamic behavior of hUTF1 and the role of both conserved domains, GFP-NLS, GFP-hUTF1 2–341, GFP-NLS-hUTF1 141–341, and GFP-hUTF1 2–41 L293P L300P fusion constructs (Fig. 3A; the CD1 and CD2 domains are indicated with black and gray boxes, respectively) were generated and stably expressed in P19CL6 cells.

Deletions in the CD1 domain led to cytoplasmic mislocalization of the protein (data not shown). To induce nuclear targeting, we fused amino acids 141 to 341 of hUTF1 to an NLS sequence (PPKKKRKV). To investigate the localization and biochemical properties of the mutants, confocal imaging and subnuclear fractionations were performed (Fig. 3B, Fig. 3C).

GFP-NLS, as a control, was detected in the free diffusing fraction (F) and the full-length GFP-hUTF1 fractionated to the strongly DNA-associated fraction (AS). Confocal imaging of N-terminal deletion mutant GFP-NLS-UTF1 141–341 showed a more dispersed nuclear localization than GFP-hUTF1 2–341 and a fair amount GFP-NLS-UTF1 141–341 localized to the nucleoli (Fig. 3C). Despite this partly nuclear localization, GFP-NLS-hUTF1 141–341 was exclusively detected in the tightly DNA-associated fraction (AS). A double point mutation in the putative leucine zipper in CD2, GFP-hUTF1 2–341 L293P L300P, resulted in a complete shift to the free diffusing protein fraction. This is reflected by a more diffuse localization in the nucleus of interphase cells. However, during cell division GFP-hUTF1 2–341 L293P L300P was detected at mitotic chromosomes (Fig. 3C). In all fractionation experiments, the endogenous mUTF1 protein served as an internal control and localized to the AS fraction (data not shown).

To study the dynamics of the interaction of hUTF1 with the chromatin in living cells, fluorescent recovery after photobleaching (FRAP) analysis (Phair et al., 2004) was performed by bleaching fluorescent molecules in a 10 μM strip spanning the nucleus. Subsequent fluorescent recovery in the strip was measured every 20 ms. As a control experiment, FRAP analysis was performed on GFP-NLS (Fig. 3D) showing that the recovery curves of GFP and GFP-NLS are almost identical. Fusion of full-length hUTF1 to GFP results in a highly immobilized protein (slope of curve Fig. 3E), similar to what was previously observed for mouse UTF1 (van den Boom et al., 2007).

To study the role of the CD1 and CD2 domains in the dynamic behavior of hUTF1, strip-FRAP analysis was per-
Human UTF1: a chromatin-associated transcriptional repressor

A. eGFP HA NLS hUTF1

B. GFP-NLS 2-341 NLS-141-341 L293P L300P

C. 2-341 NLS 141-341 L293P L300P

D. Relative Fluorescence vs. Time (s)

E. Relative Fluorescence vs. Time (s)

F. Relative Fluorescence vs. Time (s)

G. Relative Fluorescence vs. Time (s)
formed on GFP-NLS-hUTF1 141–341 and GFP-hUTF1 2–341 L293P L300P, respectively. In the case of the GFP-NLS-hUTF1 141–341 construct, only nuclear (and not nucleolar) localized molecules were bleached. The GFP-NLS-hUTF1 141–341 fusion protein was highly immobilized with dynamics similar to full-length hUTF1 (Fig. 3F), indicating that the CD1 domain is dispensable for long-term immobilization. In contrast, FRAP analysis of GFP-hUTF1 2–341 L293P L300P showed a highly mobile protein (Fig. 3G), indicating that an intact CD2 domain is necessary for stable interaction of hUTF1 with sites of affinity.

Discussion

Recently we have shown that mUTF1 is strongly associated with chromatin in mouse ES cells and that it is capable of transcriptional repression. In addition, both ES and EC cells with severely reduced levels of mUTF1 failed to differentiate properly while their self-renewing capacity was not affected. The histone-like characteristics of mUTF1 and its repressor activity implicate a role in maintaining a specific epigenetic profile required for differentiation either by attracting chromatin-modifying proteins or by compacting chromatin by itself (van den Boom et al., 2007). In humans, UTF1 has an expression pattern similar to that in mice, it is expressed by EC and ES cells, and is rapidly down regulated during differentiation (Fukushima et al., 1998; Phair et al., 2004; Ginis et al., 2004; Carpenter et al., 2004). During embryonic development, UTF1 expression is maintained in PGCs and in spermatogonial stem cells, where it is possibly involved in spermatogonial differentiation (van Bragt et al., 2004; Kristensen et al., 2008).

The aim of this study was to characterize the repressor activity, localization, and dynamic behavior of the human UTF1 protein. Immunofluorescent analysis of hUTF1 shows that it is a nuclear protein that colocalizes with DNA during all stages of the cell cycle, including mitosis. Cotransfection with a luciferase reporter indicates that hUTF1 is capable of transcriptional repression, and both subnuclear fractionation and strip-FRAP analyses show that hUTF1 is a strongly chromatin-associated protein.

Additional analysis of deletion mutants has shown the contribution of the two conserved domains to the biochemical properties of hUTF1. Whereas the CD1 domain appears to be responsible for proper nuclear targeting of the protein, the CD2 domain is involved in the histone-like association to chromatin and the repressor activity of hUTF1.

In mouse UTF1, the CD1 domain is also required for nuclear localization and proper targeting to sites of affinity and the CD2 domain for long-term immobilization. The localization of the domains with repressor activity in the hUTF1 protein differs from mUTF1. In hUTF1, repressor activity can almost completely be abolished by deletion of the CD2 domain whereas in mUTF1 each of the two conserved domains is responsible for approximately half of the repressor activity.

In earlier studies it was reported that mUTF1 represses TATA-containing promoter constructs. This mUTF1 repressor activity was dependent on its CD2 domain, as deletion of the 42 carboxy-terminal amino acids resulted in a complete loss of repressor activity and even potentiated reporter activity (Fukushima et al., 1999). Similar studies using the human UTF1 protein yielded different results. Fukushima and co-workers reported that hUTF1 can interact with the ATF-2 protein and activate transcription. This coactivator activity is dependent on intact CD1 and CD2 domains as deletion of either one resulted in a loss of coactivator activity (Fukushima et al., 1998). Summarizing, these and our observations show that there are differences in function between mouse and human UTF1 CD2 domains, which in view of their ~87% sequence identity was unexpected.

Interestingly, although mutation of leucines at positions 293 and 300 into prolines in hUTF1 did not result in decreased repressor activity, it did result in a completely different dynamic behavior. This suggests that the leucine zipper in CD2 is required for immobilization of hUTF1 where repressor activity depends on a different domain within CD2.

The fact that the biochemical and histone-like properties are conserved between both mouse and human UTF1 indicates that UTF1 is a chromatin component of mammalian embryonic stem cells. Presumably, the presence of UTF1 on the chromatin in stem cells is involved in creating or maintaining an ES cell-specific chromatin structure that is required for pluripotency.

Materials and methods

Plasmids

pSG424 (GAL4-DNA-binding domain) plasmids
pSG424-mUTF1 1–339 was described previously (van den Boom et al., 2007). pCEP4-FLAG-hUTF1 2–341 and pCEP4-FLAG-hUTF1 2–341 L293P L300P were provided by A. Okuda. pSG424-hUTF1 2–341 L293P L300P was generated by ligating a KpnI (Klenow) and XbaI fragment from pcDNA3-flag-hUTF1 2–341 L293P L300P into pSG424 digested with SmaI and XbaI.

pSG424-hUTF1 2–341 was generated by AcclIII-BamHI digestion of pCEP4-hUTF1 and ligation into pSG424-hUTF1 2–341 L243P L300P that was digested with AcclIII and BamHI.

pSG424-hUTF1 2–178 was generated by ligating the EcoRI-Nael fragment of pSG424-hUTF1 2–341 into pSG424-hUTF1 2–341 digested with XbaI and EcoRI.

pSG424-hUTF1 2–126 was generated by digestion of pSG424-hUTF1 2–341 with BamHI (Klenow)-NruI and subsequent ligation. pSG424-hUTF1 2–26 was generated by digestion of pSG424-hUTF1 2–431 with BamHI (T4 polymerase) and SacII (T4 polymerase) followed by ligation. pSG424-hUTF1 141–341 was generated by ligation of the EcoRI-AccIII (Klenow) fragment from pSG424-hUTF1 2–341 into pSG424-hUTF1 2–341 digested with EcoRI (Klenow). pSG424-hUTF1 256–341 was generated by digestion of pSG424-hUTF1 2–341 with BbsI (Klenow) and XbaI followed by ligation in pSG424-hUTF1 2–341 that was digested with EcoRI (Klenow) and XbaI.

peGFP plasmids

peGFP-HA-hUTF1 2–341 L293P L300P was generated by ligation of a SalI fragment from pcDNA3-HA-hUTF1 2–341 L293P L300P into SalI-digested peGFP-HA-hUTF1 2–341 L293P L300P digested with AccIII and BamHI to generate pSG424-hUTF1 2–341 L293P L300P that was digested with AccIII and BamHI.

peGFP-HA-hUTF1 2–26 was generated by digestion of pSG424-hUTF1 2–431 with BamHI (T4 polymerase) and SacII (T4 polymerase) followed by ligation. pSG424-hUTF1 141–341 was generated by ligation of the EcoRI-AccIII (Klenow) fragment from pSG424-hUTF1 2–341 into pSG424-hUTF1 2–341 digested with EcoRI (Klenow). pSG424-hUTF1 256–341 was generated by digestion of pSG424-hUTF1 2–341 with BbsI (Klenow) and XbaI followed by ligation in pSG424-hUTF1 2–341 that was digested with EcoRI (Klenow) and XbaI.
gesting peGFP-HA-hUTF1 2–341 with BbsI (Klenow) and KpnI (Klenow) followed by self-ligation. peGFP-HA-NLS was generated by PCR on peGFP-C1 using the following primers: F, GTT TGC TGA CCT TGA; R, ATA GCC GGC GATATC TAA CCT. The PCR product was ligated into pBluescript II SK+ digested with EcoRV, resulting in pBluescript II SK+–GFP-HA-NLS. Subsequently the NheI-Nael fragment from pBluescript II SK+–GFP-HA-NLS was ligated into peGFP-C1 digested with NheI and Small. peGFP-HA-NLS-hUTF1 141–341 was generated by digesting pBluescript II SK+–GFP-HA-NLS with Nhel and NgoMIV followed by ligation into peGFP-HA-hUTF1 2–341 digested with Nhel and Kpn2I.

Cell culture

NCCIT cells were grown in RPMI 1640 medium containing 10% FBS (PAA), 100 U/ml penicillin, 100 μg/ml streptomycin (Invitrogen). NCCIT cells were differentiated with 10 μM retinoic acid. HepG2 and P19C6 cells were cultured as described in (van den Boom et al., 2007).

Reporter assays

Luciferase reporter assays were performed as described previously (van den Boom et al., 2007).

Immunofluorescence and microscopy

UTF1 was detected in NCCIT cells with a monoclonal anti-UTF1 antibody, clone 5G10.2 (MAB 4337, Millipore). Goat anti-mouse IgG Alexa Fluor 488 (Molecular Probes) was used for visualization. For technical details on procedures and microscopy see (van den Boom et al., 2007).

RT-PCR

Total RNA was extracted using TRIzol reagent (Invitrogen). One microgram of RNA was treated with RNase-free DNaseI (Invitrogen). NCCIT cells were differentiated with 10 μg/ml streptomycin (Gibco, Grand Island, NY). Cells were grown in RPMI 1640 medium containing 10% FBS (PAA), 100 U/ml penicillin, 100 μg/ml streptomycin (Invitrogen). NCCIT cells were differentiated with 10 μM retinoic acid. HepG2 and P19C6 cells were cultured as described in (van den Boom et al., 2007). Secondary antibodies used are donkey anti-goat IgG-HRP (Santa Cruz Biotechnology, Inc.), donkey anti-rabbit IgG-HRP (GE healthcare), and goat anti-mouse IgG-HRP (Santa Cruz Biotechnology, Inc.).

Strip-FRAP

Strip-FRAP experiments were performed as described previously (van den Boom et al., 2007).

Acknowledgments

We thank Dr. A. Okuda (Division of Developmental Biology, Saitama Medical School, Saitama, Japan) for supplying human UTF1 constructs, Dr. L.H.J. Looijenga (Dept. of Pathology, Erasmus MC, Rotterdam, the Netherlands) for NCCIT cells, L. Drenth-Diephuis for technical assistance, and Dr. V. van den Boom for assistance with the Strip-FRAP experiments. This work was supported by the Groningen Biomolecular Sciences and Biotechnology Institute.

References


