Defining the heterochromatin localization and repression domains of SALL1

Christian Netzer a,b,1, Stefan K. Bohlander c,1, Markus Hinzke d, Ying Chen c, Jürgen Kohlhase c,*

a Institute of Human Genetics, University of Bonn, Wilhelmstr. 31, 53111 Bonn, Germany
b Institute of Human Genetics, University of Cologne, Kerpener Str. 34, 50931 Cologne, Germany
c Department of Medicine III, University of Munich, and GSF, Clinical Cooperative Group ‘Leukemia’, Marchioninistr. 25, 81377 München, Germany
d Institute of Human Genetics, University of Göttingen, Heinrich-Düker-Weg 12, 37073 Göttingen, Germany
e Institute of Human Genetics and Anthropology, Breisacher Str. 33, 79106 Freiburg i. Br., Germany

Received 20 September 2005; received in revised form 24 November 2005; accepted 5 December 2005
Available online 6 January 2006

Abstract

SALL1 has been identified as one of four human homologues of the Drosophila region-specific homeotic gene spalt (sal), encoding zinc finger proteins of characteristic structure. Mutations of SALL1 on chromosome 16q12.1 cause Townes–Brocks syndrome (TBS, OMIM 107480). We have shown previously that SALL1 acts as a strong transcriptional repressor in mammalian cells when fused to a heterologous DNA-binding domain. Here, we report that SALL1 contains two repression domains, one located at the extreme N-terminus of the protein and the other in the central region. SALL1 fragments with the central repression domain exhibited a punctate nuclear distribution pattern at pericentromeric heterochromatin foci in murine NIH-3T3 cells, suggesting an association between repression and heterochromatin localization. The implications of these findings for the pathogenesis of Townes–Brocks syndrome are discussed.

1. Introduction

SALL1 belongs to the spalt (sal)-like gene family. Spalt-like genes regulate developmental processes in many organisms. Mutations in human SALL1 result in Townes–Brocks syndrome (TBS; OMIM 107480), an autosomal dominantly inherited malformation syndrome [1]. TBS is characterized by anorectal abnormalities (imperforate anus, anal stenosis), abnormalities of the hands (preaxial polydactyly, triphalangeal thumbs) and feet, and deformities of the outer ear, often with preauricular tags. Hearing loss and renal malformations are commonly found, whereas cardiac defects and mental retardation occur with lower frequencies. The penetrance of TBS seems to be complete [2]. Our detection rate of SALL1 mutations among typical TBS patients is about 64% [3].

Mutations in SALL1 have been postulated to cause TBS by haploinsufficiency. However, a mouse model carrying a Sall1- null allele did not exhibit a TBS-like phenotype. The heterozygous knock-out mice were phenotypically normal, whereas homozygous Sall1-deficient mice died perinatally from renal failure caused by severe renal dysplasia or complete renal agenesis, indicating that Sall1 has an essential role in kidney development [4]. Instead, a mouse model carrying a mutant allele designed to mimic a human TBS mutation mostly recapitulated the TBS phenotype [5]. These data support a model for the pathogenesis of TBS in which expression of a truncated SALL1 protein causes abnormal development, either in a dominant-negative or gain-of-function fashion.

First hints towards an understanding of SALL1 function were obtained by demonstrating that a GFP-SALL1 fusion protein localizes to chromocenters and smaller heterochromatin foci in transiently transfected NIH-3T3 cells [6]. Chromocenters consist of clustered pericentromeric heterochromatin and contain telomere sequences. In a yeast two-hybrid screen PIN2, an isoform of telomere-repeat-binding factor 1 (TRF1), was identified as an interaction partner of SALL1. Further SALL1 interacting proteins revealed by this screen were Ubiquitin-conjugating enzyme 2I, the human homologue of
UBC9, and the small Ubiquitin-like modifier-1 (SUMO-1). In an in vitro assay, it was demonstrated that SALL1 is covalently modified by SUMO-1 in the presence of UBA2/AOS1 and UBE2I [7].

Since SAL and SAL-like proteins contain multiple possibly DNA-binding zinc finger motifs, they are postulated to be transcriptional regulators. Little is known about the direct target genes regulated by SAL and SAL-like proteins, but we could indeed show that SALL1 acts as a strong transcriptional repressor in mammalian cells [6]. For this purpose we linked SALL1 to a GAL4 DNA-binding domain (GAL4-DB) and measured the activity of a luciferase reporter gene containing a GAL4-binding site upstream of its promoter. Using similar assay systems, two groups reported that the homologous genes in mouse (Sall1) and chick (csal1) are also potent transcriptional repressors [8,9]. On the other hand, new results achieved with a Wnt responsive reporter suggest that murine Sall1 activates the canonical Wnt signaling pathway [10]. In concordance with this, an interaction of β-catenin and Sall1 was demonstrated. Furthermore, there is evidence from chick csal1 that both FGF and Wnt signals are required for the regulation of csal1 expression in the limb [11]. In ovarian carcinoma-derived cells and in human ovarian surface epithelial cells, p150 (SALL2) was shown to be a transcriptional activator of the cyclin-Cdk inhibitor p21, a key factor in G1 checkpoint control [12].

In the study presented here, the repression and heterochromatin localization properties of human SALL1 are analyzed in detail.

2. Material and methods

2.1. Plasmids

The cloning of various cDNA fragments of SALL1 into pBluescript and of the full-length SALL1 cDNA into pEGFP-C1 (CLONTECH) as well as the generation of the reporter plasmid pGAL4:tkLUC has been described previously [6,13]. The constructs I–X (see Fig. 1A) were generated by PCR (constructs III, VI, VII, VIII, XIV–XVI) or by restriction enzyme digestion (constructs I, II, IV, V, IX, X) and subcloning into pM1 (CLONTECH), a mammalian vector used for expression of hybrid proteins with the DNA-binding domain of GAL4. The correct insertion was confirmed by sequencing the cloning site and the entire ORF.

2.2. Transcription repression assay

NIH-3T3 cells (1.4 × 10^5 cells/35 mm plate) were transfected with Roti®-Fect (ROTH). Each transfection assay was performed in triplicate and included (i) 0.5 μg of the pM1-SALL1 deletion construct, (ii) 0.5 μg of the pGAL4:tkLUC reporter plasmid. Specific luciferase activity was measured with the Dual-Luciferase System (Promega) and was normalized to the α-tubulin expression level. The luciferase activity of NIH-3T3 cells expressing GAL4DB alone was set to 1. Fold repression was calculated by dividing the normalized luciferase activity of NIH-3T3 cells expressing GAL4DB alone by the activity of the SALL1 fusion protein. Values are plotted as the mean ± standard deviation of triplicate transfections from three independent experiments. They demonstrate that SALL1 contains two strong repression domains: one at the extreme N-terminus and one in the central region.

Fig. 1. (A) Diagram of full-length SALL1 and of the twelve SALL1 deletion mutants used in this study. The ORF of SALL1 is shown as a box. The positions of the double zinc fingers (white) and of the single zinc fingers (black) are depicted as ovals. The SALL1 fragments were expressed as fusion proteins with the GAL4-DNA-binding domain (GAL4DB). The names of the mutants and their number (written in latin capitals) are indicated on the left side. The number of the first and last amino acid (aa) contained in the fragment are put in brackets. The position of the putative N-terminal repression domain is shown as a box, as well as the region of overlap of the three central SALL1 fragments revealing repression properties and heterochromatin association. (B) Percentage of nuclei with punctate heterochromatin localization. (C) Diagram of the repression activity of full-length SALL1 and of the twelve SALL1 deletion mutants. Fold repression was calculated by dividing the normalized luciferase activity of NIH-3T3 cells expressing GAL4DB alone by the activity of the SALL1 fusion protein. Values are plotted as the mean ± standard deviation of triplicate transfections from three independent experiments. They demonstrate that SALL1 contains two strong repression domains: one at the extreme N-terminus and one in the central region.
reporter plasmid and (iii) 0.01 μg of pRL-null (Clontech), expressing the Renilla luciferase. Forty-eight hours after transfection, cells were harvested, lysed in 100 μl of lysis buffer (100 mM KH₂PO₄ pH 7.8, 0.2% (v/v) Triton-X-100, 0.5 mM DTT) and assayed for luciferase and Renilla luciferase activity with an Autolumat LB953 (Berthold, Wildbad, Germany) as described [14]. All assays were repeated three times. The median of the repression activity and the standard deviation were calculated.

2.3. Indirect fluorescence microscopy

NIH-3T3 cells transiently transfected with pM1-SALL1 deletion constructs were used for indirect immunofluorescence with antibodies against the GAL4-DNA binding domain (Santa Cruz Biotechnology, sc-510). 48 hours after transfection, cells were fixed in ice-cold 0.1% Triton-X and blocked in 10% goat serum. Then, the cells were incubated with the primary antibody, washed and detected with an anti-goat FITC-conjugated secondary antibody. Images were scored and recorded with an inverted epifluorescence microscope (Zeiss Axiovert) and a CCD camera using the appropriate filter settings. Nuclei were counterstained with 1μg/ml DAPI dye (Sigma), and coverslips were mounted using VECTASHIELD antifade (Vector laboratories).

For each construct, 150 GAL4-DB-positive cells on three coverslips (50 on each coverslip) from at least two independent transfections were scored for the pattern of nuclear staining (punctate or diffuse, see text).

3. Results and discussion

A reporter gene assay was used to analyze the ability of human SALL1 to affect transcription. A series of transient transfections of NIH-3T3 cells was performed with full-length SALL1 or fragments of SALL1 fused to the GAL4-DNA-binding domain (see Fig. 1A). The construct pGAL4,tkLUC, which contains the luciferase gene under the control of a thymidine kinase promoter with a GAL4 binding site, served as a reporter plasmid. Our assay system was able to detect both transcriptional activation and repression (data not shown).

We knew from previous experiments [6] that full-length SALL1 (construct I) strongly represses luciferase expression in this system. To map the domains responsible for transcriptional repression, twelve partially overlapping deletion mutants of SALL1 (constructs II–XVI) were used in the reporter gene assay. In these experiments it became apparent that SALL1 contains two repression domains, one located at the extreme N-terminus of the protein and the other one in the central region.

3.1. Repression by the N-terminus of SALL1

The construct expressing the first 183 amino acids of SALL1 (construct II) had a strong repression activity and was capable of repressing transcription about 5-fold (see Fig. 1C). In the final series of experiments depicted here the full-length SALL1 (construct I) exhibited a weaker repression activity than this truncated protein. In more than a dozen experiments performed previously with a slightly different protocol (in which the vector pCMV-β-GAL was co-transfected to normalize transfection efficiency) constructs I and II consistently showed similar repression activities. Removing aa 1–87 of SALL1 from the deletion construct II (resulting in construct III) completely abolished repression, suggesting that the 87 N-terminal amino acids of SALL1 contain a strong repression domain (see Fig. 1).

These results correspond well with published data from murine Sall1: in a similar reporter gene assay, deletion of aa 1–77 of Sall1 reduced the ability to repress transcription to 10%. Deletion of aa 1–129 removed all of the repression activity of Sall1 [8]. The homology between human SALL1 and mouse Sall1 is 89.9% for the entire protein and reaches 97.4% for the first 77 amino acids while dropping to only 64.7% in the region from amino acid 78 to 129. Since the function of SALL1 as a transcriptional repressor has obviously been conserved in evolution, this suggests that the repression domain is located in the highly conserved N-terminus of the protein (amino acids 1–87). This assumption is further supported by the lack of transcriptional repression mediated by construct III expressing aa 87–183 of SALL1 in fusion with the GAL4-DB.

Within the region required for repression, all known vertebrate SAL-like proteins contain a highly conserved single zinc finger domain of the C₂H₂-type [15–24]. Such C₂H₂ zinc finger domains have been postulated to mediate protein–protein interactions rather than DNA binding (reviewed in [25]). For murine Sall1 it has been demonstrated that the first 76 amino acids are required for binding histone deacetylase (HDAC) repression complex proteins (HDAC1, HDAC2, RbAp46/48, MTA1, MTA2). Surprisingly, mutation of the C₂H₂ zinc finger did not affect repression or interaction with the repression complex [5].

3.2. Repression by the central domain of SALL1

Our data also show that SALL1 contains a second repression domain: Constructs IV, V, and VI, which do not contain the extreme N-terminus of SALL1, encode for SALL1 fragments that exhibited a strong (3- to 4-fold) ability to repress luciferase expression (see Fig. 1C). The overlap of these SALL1 fragments encompasses aa 434 to aa 690, suggesting that the central repression domain is located within this part of the protein. Surprisingly, a construct expressing this overlapping region (construct XV) had no detectable repression activity, and the same was observed for six additional SALL1 deletion mutants in which C-terminal or N-terminal parts of the repressing SALL1 fragments were systematically removed (encoded by constructs VII–X and XIV–XVI).

A consistent interpretation for this observation would be that the repression properties of the central repression domain depend on a tertiary protein structure which is intact in construct VI (the smallest repressing fragment from the central region) and disturbed by any further deletion. Repression apparently does not depend on the highly conserved second double zinc finger and the associated single zinc finger in construct VI, since construct V is able to repress without them. One might speculate that the first C₂H₂ double zinc finger is necessary for repression, since this zinc finger is present in all three central fragments with repression activity. On the other hand, this double zinc finger would clearly not be sufficient for repression, since regions C-terminal and/or N-terminal of the zinc finger are needed for repression. Possibly two or more sequence elements in the critical region mediate repression, and their effective cooperation depends on the correct conformation of the protein.
Further dissection of the central repression domain may contribute to resolving the issue of structural requirements for the repression function of this region.

A detailed investigation of the central repression domain of murine Sall1 has not been performed so far, but the existence of a second repression domain has been reported: A Sall1 fragment covering aa 436 to 1103 clearly repressed luciferase expression in a comparable reporter gene assay [8]. In contrast to the N-terminal repression domain, this central repression domain was not dependent on HDAC activity.

3.3. Mapping of the heterochromatin localization domain

We used an anti-GAL4-DNA-binding-domain antibody to visualize the intranuclear localization of all SALL1 fragments tested in the reporter gene assay by epifluorescence microscopy. These fusion proteins were transported into the nucleus due to a nuclear localization signal (NLS) contained in the yeast GAL4-DNA-binding domain. Hence, this strategy should enable a mapping of the heterochromatin localization domain in NIH-3T3 cells independent from an intact native NLS. NIH-3T3 mouse fibroblasts contain easily detectable DAPI-bright regions of the nucleus which are known to represent pericentromeric heterochromatin [6].

Only constructs I, IV, V and VI displayed the typical punctate distribution at the chromocenters in 45% (construct I, full-length SALL1), 10% (construct IV), 15% (construct V) and 20% (construct VI) of nuclei (Fig. 2A–D). The other nine deletion mutants showed a diffuse nuclear distribution including construct II, which harbors the N-terminal repression domain (Fig. 2E). It should be noted that due to technical limitations (i.e. indirect immunofluorescence on transiently transfected cells) the percentage of nuclei with punctate pattern is probably an underestimate. When a GFP-SALL1 construct was used nearly 100% of transfected nuclei exhibited a chromocenter associated staining pattern [6].

A consistent explanation for these findings will have to refer to the conformational and/or structural hypothesis discussed above for the central repression domain, i.e. smaller deletion mutants (constructs VII to XVI) may lack the correct tertiary structure and/or additional sequence elements required for heterochromatin binding. Nevertheless our data suggest that there is an association between repression by the central domain and heterochromatin localization, since the three SALL1 deletion mutants with a partially heterochromatic localization were those who also exhibited a detectable repression activity. This suggests that the repression properties of the central SALL1 fragments might depend on heterochromatin localization.

The published data on the intracellular localization of murine Sall1 fragments are limited and support the complex picture emerging from our observations: A deletion mutant covering aa 1–435 of Sall1 displayed a diffuse nuclear distribution in NIH-3T3 cells, a fragment from aa 1 to 598 (which included the first double zinc finger) showed a punctate heterochromatin localization. A third fragment containing the C-terminal half of Sall1 (aa 599–1324, including double zinc fingers 3, 4 and 5) formed

![Epifluorescence images showing nuclei expressing GAL4DB-SALL1 (A) and -SALL1 deletion mutants (B–E) as well as the corresponding DAPI staining (right pictures).](image-url)
larger intranuclear aggregates [10]. Whereas these data are in concordance with our results, a predominantly punctate nuclear distribution pattern of a Sall1 fragment containing aa 2–435 in COS-1 cells has been reported [8]. Whether these deviating observations reflect differences of NIH-3T3 and COS-1 cells – e.g., with respect to endogenous Sall1 expression and subsequent intranuclear heterodimerization with exogenous Sall1 fragments at heterochromatic sites – is presently unclear. There is no detectable Sall1 expression in NIH-3T3 cells [10], but to our knowledge, this has not been tested for COS-1 cells so far.

3.4. Conclusions

In summary, we have provided strong evidence that human SALL1 is a potent transcriptional repressor and report that it contains two repression domains, one at the extreme N-terminus and one in the central region of the protein. Repression by the central domain was associated with heterochromatin localization.

The presence of multiple independent repression domains has been demonstrated for other transcription factors, e.g. for Brinker, Hesx1, and ZEB [26–28]. In some cases the two domains recruit different co-repressors to mediate transcriptional repression and can function independently or cooperatively. Remarkably, another C2H2 type zinc finger protein, Ikaros, colocalizes to centromeric foci with inactive, developmentally-regulated genes, suggesting that it may contribute to the pericentromeric repositioning and heritable inactivation of these genes [29,30]. Identification of in vivo target genes regulated by SALL1 will be important to investigate whether a similar model can be applied here and to analyze the repression mechanisms used by SALL1 in detail.

The findings presented in this study may have important implications for the pathogenesis of TBS. The majority of SALL1 mutations occur 5′ to the region encoding the first double zinc finger domain [31,32], and all known SALL1 mutations cause premature stop codons. New results [33] document that deletions of the entire SALL1 gene also lead to TBS, confirming that SALL1 haploinsufficiency plays a role in the etiology of the disorder. Alleles resulting from mutations in the 5′ region of exon 2 encode for truncated proteins with strong repressor activity but without the central repression and heterochromatin localization domain. Despite of their potential to act as strong transcriptional repressors, these proteins will probably not localize to the physiological site of action, i.e. the heterochromatic foci. In addition, the truncated proteins are present in both the nucleus and cytoplasm in the TBS mouse model [5] and in the chick as shown by tissue culture experiments [9]. Furthermore, in the latter case, the truncated proteins bind other SAL proteins and move them from the nucleus to the cytoplasm. Mutations further 3′ in the SALL1 gene are thought to result mostly in milder phenotypes than the 5′ mutations [32,34]. If some of those mutations lead to truncated proteins including both repression domains and, more importantly, the heterochromatin localization domain, these proteins could have some residual function which might explain the milder phenotype.

The observation that the human TBS phenotype can result from a SALL1 dosage effect [33] raises the possibility that the critical point in the pathogenesis is the correct dosage of functional SALL1 protein at the heterochromatic foci. A deletion of one allele results in a 50% reduction of this dosage. A 5′ truncating mutation leads to a truncated protein, which does not reach its site of action and in addition probably even removes some full length protein of the normal allele from the nucleus. Therefore, the in most instances more severe phenotype of the 5′ truncating mutations might result from a reduction of the functional protein at the site of action by more than 50%.

However, some riddles still remain to be solved: The interaction between truncated SALL1 and functional SALL1 or other SALL proteins and the relocalization of the functional proteins requires the presence of the evolutionary conserved glutamine-rich region in the amino terminal part of the truncated protein [9]. The identification of the TBS-causing SALL1 mutation c.419delC [3], which would result in a truncated protein lacking the interaction domain, shows that this theory does not fully explain the etiology. Further work on the function of SALL1 and its involvement in various signaling pathways (e.g., Wnt signaling pathway) is required to unravel TBS pathogenesis.

Acknowledgements

The authors would like to thank Wolfgang Engel for his support. This study was funded by the Fritz-Thyssen-Stiftung (grant no. 2000 1071 to S.K.B. and J.K.). C.N. was supported by the DFG (Deutsche Forschungsgemeinschaft) and the BONFOR program of the Medical Faculty, Rheinische Friedrich-Wilhelms-University, Bonn.

References


