



Drosophila Sal and Salr are transcriptional repressors

Jonatan Sánchez, Ana Talamillo, Monika Gonzalez, Luis Sánchez-Pulido,
Silvia Jimenez, Lucia Pirone, James David Sutherland, Rosa Barrio

► To cite this version:

Jonatan Sánchez, Ana Talamillo, Monika Gonzalez, Luis Sánchez-Pulido, Silvia Jimenez, et al..
Drosophila Sal and Salr are transcriptional repressors. Biochemical Journal, 2011, 438 (3), pp.437-445.
10.1042/BJ20110229 . hal-00617326

HAL Id: hal-00617326

<https://hal.science/hal-00617326>

Submitted on 27 Aug 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***Drosophila* Sal and Salr are transcriptional repressors**

Jonatan SÁNCHEZ^{*,†}, Ana TALAMILLO^{*}, Monika GONZÁLEZ^{*}, Luis SÁNCHEZ-PULIDO[‡], Silvia JIMÉNEZ^{*}, Lucia PIRONE^{*}, James D. SUTHERLAND^{*} and Rosa BARRIO^{*,§}

^{*}: CIC bioGUNE, Bizkaia Technology Park, 48160 Derio, Spain.

[†]: Present address: McKnight Brain Institute, University of Florida, FL 32610-0236, USA.

[‡]: MRC Functional Genomics Unit, Department of Physiology, Anatomy and Genetics, University of Oxford, South Parks Road, Oxford OX1 3QX, UK.

Running title: Analysis of transcriptional activity in SALL proteins

(§) Author for correspondence:

Rosa Barrio

Functional Genomics, CIC bioGUNE, Bizkaia Technology Park, 48160 Bilbao, Spain

Telephone: xx 34 944 061 316

Fax: xx 34 944 061 324

email: rbarrio@cicbiogune.es

Keywords: *Drosophila*; SALL; Transcriptional repression; SUMO; HDAC.

Abbreviations used: CFP, cyan fluorescent protein; DBD, DNA Binding Domain; FL, full length; Gln, glutamine; HDAC, Histone-Deacetylase Complex; NAM: Nicotinamide; OS: Okiihiro Syndrome; Sal, Spalt; SALL, Spalt-like; Salr, Spalt related; SUMO, Small Ubiquitin-like Modifier; TBS, Townes-Brooks Syndrome; TSA, Trichostatin A; YFP, yellow fluorescent protein; ZF, zinc fingers pair.

SYNOPSIS

The Spalt-like (SALL) family of zinc finger transcription factors is conserved in metazoans. In *Drosophila* Spalt (Sal) and Spalt-related (Salr) control the expression of genes involved in wing and central nervous system development, including cell adhesion and cytoskeletal proteins. In humans, *SALL* mutations associate with congenital disorders such as the Townes-Brocks and Okhiro syndromes. Human and *Drosophila* SALL proteins are modified by Sumo, which influences their subnuclear localisation. Here, we analyse the transcriptional activity of *Drosophila* Sall proteins in cultured cells. We show that both Sal and Salr act as transcriptional repressors in *Drosophila* cells where they repress transcription through an AT-rich sequence. Furthermore, using the UAS/Gal4 heterologous system, *Drosophila* Sal and Salr repress transcription in human cells. In our experimental conditions, only in the case of Salr the repression activity is dependent on the Histone Deacetylase complex. This complex might interact with the carboxy-terminal zinc fingers of Salr. We describe the differential subcellular localisations of Sal and Salr fragments and identify their repression domains. Surprisingly, both repressors also contain transcription activation domains. In addition, under our experimental conditions sumoylation has differential effects on Sal and Salr repressor activity. Phylogenetic comparison between nematodes, insects and vertebrates identifies conserved peptide sequences that are presumably critical for SALL protein function.

INTRODUCTION

Spalt-like (SALL) proteins are zinc finger transcription factors that are conserved from *C. elegans* to mammals. These proteins are involved in different biological processes such as organogenesis, carcinogenesis and the maintenance of pluripotency in embryonic stem cells [1, 2]. The first *SALL* genes to be described were Spalt (Sal) and Spalt related (Salr) in *Drosophila* [3, 4]. These two genes participate in numerous developmental processes including wing growth and vein formation under the control of the Transforming Growth Factor β pathway via the target genes *iro-C* and *kni-C* in the imaginal discs [5-7]. Sal and Salr are necessary for the determination of neuronal fate in the peripheral nervous system [8, 9] and regulate the expression of cell adhesion and cytoskeletal proteins in the central nervous system and in imaginal discs [10, 11]. However, the direct interaction between Sall proteins and the target DNA sequences has been reported only for Salr, which binds to an AT rich region (AATTATGAAATGCCA) in the promoter of the chorion gene *s15* [4, 12].

In vertebrates, there are four homologous genes (*SALL1-4*). The importance of this family of proteins to human health is linked to mutations that are associated with different inherited diseases such as the Townes-Brocks (TBS) or the Okhiro (OS) syndromes caused by mutations in *SALL1* and *SALL4*, respectively. TBS is an autosomal dominant syndrome characterised by imperforate anus, limb malformations, dysplastic ears and sensorineural hearing loss [13, 14]. OS (also known as Duane-radial ray syndrome) is an autosomal dominant disorder also characterized by limb malformations, together with ocular and renal anomalies [15, 16]. The limb and neuronal abnormalities in those patients are reminiscent of the limb and nervous system defects in *Drosophila sall* mutants, suggesting that SALL

proteins might be involved in the regulation of conserved developmental processes in metazoans [17].

In mammals, SALL1 proteins have been defined as transcriptional repressors. In humans and mice, two different domains mediate SALL1 repression ability: a polyglutamine [poly(Gln)] region located in the N-terminal part of the protein, and the central region where the zinc finger pairs (ZF) 2 and 3 are located [18-23]. In mice, the Histone-Deacetylase Complex (HDAC) mediates repression by the SALL1 N-terminal region via interaction with a 12 aminoacid sequence [20]. The human and murine regions involved in repression localise to sub-nuclear domains that coincide with heterochromatin, binding there to an AT-rich sequence [18-20, 24, 25]. This heterochromatin interaction might represent an additional mechanism of repression beyond HDAC interaction. However, not all the SALL proteins have been identified as transcriptional repressors. In humans, SALL2 activates the transcription of the cyclin-Cdk inhibitor p21, important for G1 checkpoint control [26], and SALL4 activates Bmi-1 through a specific enhancer [27]. In addition, the maintenance of pluripotency in embryonic mouse stem cells is dependent on Sall4 activation of Nanog, Pou5f1, and Sall4 itself [28, 29].

SALL proteins are modified by coupling to Small Ubiquitin-like Modifier 1 (SUMO-1), a protein modification linked to processes including subnuclear localisation and transcriptional activity [30]. In *Drosophila*, Sal and Salr are sumoylated *in vitro* and interact genetically with the sumoylation pathway during wing development. Sumoylation leads to subnuclear localisation changes in Sall proteins, with phenotypic consequences for wing growth and vein formation [7]. In this work, we analyse the functional domains of the *Drosophila* Sall proteins, their localisation and their transcriptional activity in relation to HDAC and to sumoylation. We show that both proteins are transcriptional repressors and that Salr (but not Sal) repression is mediated by HDAC. In addition, sumoylation differentially affects Salr and Sal transcriptional activity. Furthermore, we report that both these repressor proteins also contain transcriptional activation domains. Finally, the evolutionary conservation of these domains among vertebrates and invertebrates is discussed.

EXPERIMENTAL

DNA constructs

NheI-*Apal* or *NotI*-*Apal* fragments containing full-length (FL) *sal* or *salr* were cloned in the *XbaI* or *NotI*-*Apal* sites of pAC5.1-V5-His-A (Invitrogen) to generate *pAc-Sal* or *pAc-Salr*, respectively. *pAc-RLuc* contains the Renilla luciferase gene cloned between the *EcoRV* and *HpaI* sites of pAC5.1-V5-His-A. *pAc-SalrBE-RLuc* contains 5 copies of the sequence AATTATGAAATGCCA [4] in the *BspI* site of *pAc-RLuc*. *pAc-SalrBE*-RLuc* contains 5 copies of the mutated sequence AAccgaattccGCCA cloned in the same site. The *pAC5.1-CFLuc-V5His* vector was purchased from Addgene [31].

The *pECFP-His-Myc-sal*, *pEYFP-His-Fluo-salr*, *pGal4-DNA Binding Domain (DBD)*, *pGal4-sal*, *pGal4-salr*, *pB-CFP-sal-IKDP* and *pB-YFP-salr-IKEA-IKVA* have been described previously [7]. For the constructs *pGal4-S1* to *-S8*, *pGal4-S11*, *pGal4-R1* to *-R8* and *pGal4-R9* to *-R12*, *sal* or *salr* fragments were amplified by PCR using specific oligonucleotides that contained *NheI* and *Apal* restriction sites, cloned into the pST-BlueI vector (Stratagene) and sequenced. The fragments digested with *NheI*-*Apal* were cloned

into the same sites of *pGal4-DBD* or *pGal4-salr*. Fragments S9, S10, and S12 were digested from *pGal4-sal* and cloned into the same sites of *pGal4-DBD* using *KpnI-EcoRV*, *KpnI-EcoRI* or *KpnI-NotI* restriction enzymes, respectively.

pGal4-sal-IKDP was generated by cloning the *EcoRI* fragment from *pB-CFP-sal-IKDP* into *pGal4-sal*. *pGal4-salr-IKEA-IKVA* was generated by cloning the fragment *XbaI-ApaI* from *pB-YFP-salr-IKEA-IKVA* into the same sites of *pGal4-salr*.

Luciferase reporter *pGal4tkLUC* was a gift from Dr. Ericsson [32]. *pRL-SV40* (Promega) was used to normalize the activity in the transcriptional activity assays.

Oligonucleotide sequences are given in the Supplementary Experimental section.

Cell culture and immunocytochemistry

Drosophila S2R+ [33] and Kc167 cells [34] were obtained from the *Drosophila* Genomics Resource Center. Cells were cultured at 25°C in *Drosophila* Schneider's medium (Invitrogen) supplemented with 10% of foetal bovine serum (Gibco) and 1% of penicillin and streptomycin (Gibco). Transfections were done using Effectene (Qiagen) in 96 well plates.

Human embryonic kidney (HEK) 293FT cells (Invitrogen) were maintained at 37°C in DMEM (Gibco) supplemented with 10% of foetal bovine serum, 1% of penicillin-streptomycin and 0.2% of primocin (InvivoGen). For transfections the calcium phosphate method was used in 24 wells plates.

For immunofluorescence, cells were fixed 24 hours after transfection and treated as previously described [7] using anti-Gal4 (1:50; Santacruz) or DAPI (1:2000; Sigma). Confocal images were taken using a Leica TCS-SP2 DM-IRE2 microscope using the 63x objective, resolution 1024x1024 dpi and different zoom values.

Transcriptional activity assays

Drosophila S2R+ cells were seeded in 96 well plates and transfected with 25 ng of *pAC-SalrBE-RLuc* or *pAC-SalrBE*-RLuc*, either alone, or cotransfected with 12,5 ng of *pAc-Sal* or *pAc-Salr* and 12.5 ng of *pAC5.1-CFLuc-V5His* were used for normalization. 293FT cells were seeded in 24 well plates and transfected with 500 ng of the *pGal4-DBD* constructs or 500 ng of *pGal4tkLUC*. 1 ng of *pRL-SV40* was added to each well for normalization. To test the effect of HDAC, 24 hours after transfection cells were treated with either Trichostatin A (TSA; 90 ng/ml, Calbiochem), EX527 (10 µM, Cayman Chemical), Nicotinamide (NAM; 10mM, Sigma) or DMSO (Sigma). Transcriptional activity was measured 24 or 48 hours after transfection using the Dual-Luciferase Reporter Assay System (Promega), for mammalian cells or the Dual-Glo Luciferase Assay System (Promega), for *Drosophila* cells following the manufacturer's instructions. Luminescence was measured in a Microplate Luminometer (Veritas). Results are given as mean + SD, unless otherwise stated. Differences between groups were calculated using Student's t test.

Western blotting

292FT human cells were transfected in 24 well-plates as above. 24 h after transfection cells were collected in 100 µl of Laemmli buffer, protein contents was measured and equivalent amounts were loaded into 4-15% gradient polyacrylamide gels (Biorad). Transfer to PVDF membranes was done during 3 or 4 minutes using the iBlot system (Invitrogen). Anti-Gal4

(Santacruz) antibodies were used, dilution 1:1000. Expected molecular weight and number of aminoacids for each fragment are provided in Supplementary Experimental section.

Sequences comparison and phylogenetic analysis

Sequences search and comparison was done using the Blast tool (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Alignments and unrooted N-J phylogeny tree were done using T-Coffee (<http://www.tcoffee.org>). Databases codes for the sequences used in the analysis are provided in the respective Figure Legends.

RESULTS

Localisation and transcriptional activity of Sall proteins in *Drosophila* cells

We analysed the localisation and transcriptional capacity of Sall proteins in *Drosophila* cell culture. Sall proteins labelled with fluorescent tags were expressed in *Drosophila* Kc167 cells. As in mammalian cells [7], both Sall proteins localised to subnuclear domains (Fig. 1A, B), with the Sal domains being larger in size, and fewer in number, than the Salr domains.

To test the transcriptional capacity of Sall proteins in *Drosophila* cells, we used a Salr downstream region [4]. Salr binds to an AT-rich sequence AATTATGAAATGCCA, which we refer to in this work as SalrBE (Salr binding element). The mutated SalrBE* sequence (AAccgaattccGCCA) blocks Salr binding [4]. Each of these sequences was placed between the strong *actin 5C* promoter and the *TATA* box, upstream of the *Renilla luciferase* (*RLuc*) reporter gene. Plasmids were transfected and luciferase assays were performed in S2R+ *Drosophila* cells. The *pAC-SalrBE*-RLuc* construct gave lower RLuc activity compared to the *pAC-RLuc* or *pAC-SalrBE-RLuc* constructs (data not shown). We found that the RLuc activity from the *pAC-RLuc* construct did not change in presence of Sal or Salr (data not shown). In contrast, both *pAC-Sal* and *pAC-Salr* repressed RLuc activity from *pAC-SalrBE-RLuc* by more than 80% (Fig. 1C). However, the activity of the *pAC-SalrBE*-RLuc* construct, carrying the mutated binding element, was not significantly reduced in the presence of either Sal or Salr (Fig. 1D). These results confirm that the Sal and Salr repressors both act through the AT-rich SalrBE sequence.

Localisation and transcriptional activity of Sall proteins in human cells

Drosophila S2R+ and Kc167 cells are poorly adherent, which makes handling them more difficult. In addition, the small size of these cells is not ideal for determining subcellular localisations. Therefore, localisation and functional assays were made in the human 293FT kidney cell line. *SALL1*, *SALL3* and *SALL4* genes are expressed in the 293FT cells line [35] (www.proteinatlas.org), having SALL proteins an important role in kidney development [36]. To analyse the transcriptional activity and localisation of the different domains of the *Drosophila* Sall proteins, we used a bipartite UAS/Gal4 heterologous transcriptional reporter assay. We tested Gal4-DBD constructs fused to Sal or Salr, and co-transfected into 293FT cells together with a UAS-luciferase reporter. Using an anti-Gal4 antibody, we detected the localisation of Sall proteins in subnuclear domains (Fig. 2A-B), as it was

previously reported for CFP-Sal and YFP-Salr fusion proteins [7], indicating that the Gal4 fusion did not alter the localisation of the proteins.

Luciferase assays showed that both Sall proteins gave strong repression, the activity of Gal4-Sal and Gal4-Salr being around 20% the activity of Gal4-DBD alone (Fig. 2C). These results indicated that *Drosophila* Sall proteins repress transcription in mammalian cells. To test whether the transcriptional activity of Sall proteins was mediated by HDAC, cells transfected with *Gal4-sal* or *Gal4-salr* were treated with the HDAC inhibitor TSA. Only Salr repression activity seems to be mediated by HDAC, the cells treated with TSA showing statistically significant differences from DMSO-treated cells. Cells transfected with Sal did not show statistically significant differences between control and TSA treated cells, suggesting that Sal repression might be mediated by a different mechanism.

TSA inhibits Class I and Class II, the so-called “classical” HDACs. However, the Class III, composed by NAD⁺-dependent enzymes or sirtuins, are not affected by TSA [37, 38]. In order to test whether the repression exerted by Sall proteins is mediated by sirtuins, we used two specific inhibitors: NAM, a general sirtuin inhibitor, and EX527, a specific inhibitor of Sirt1. NAM and EX527 released the repression exerted by Salr in a similar way than TSA. The activities shown by Salr in cells treated with TSA, NAM or EX527 are similar, but show statistically significant differences to non-treated cells ($p < 0.005$). Double treatments, with TSA and NAM, or with TSA and EX527, did not show additive effect (data not shown). In contrast to Salr, these drug treatments gave no significant differences in Sal repression compared to untreated cells. Therefore, these results suggest that in these experimental settings Sal and Salr might repress transcription through different mechanisms.

Differential distribution of zinc finger pairs in Sall proteins across evolution

Our previous results show that, surprisingly, Sal and Salr might mediate transcription through different mechanisms, either HDAC-dependent or independent. This might depend on the differential distribution of the ZFs present in Sall proteins (Fig. 3). Specific ZFs at a given position are more similar to each other in diverse species than to ZFs at other positions. For example, ZF2 from human SALL1 is more related to ZF2 from *Drosophila* Sal than to ZF3, 4 or 5 from human SALL1, indicating that the distribution of finger pairs typical of the SALL proteins occurred before the divergence of insects and mammals. Zinc finger pairs ZF2 and 3 are present in all the *Sall* genes from the insect and mammalian species analysed, the presence of ZF4 or 5 being variable (Figs. 2, S1-S5). Human and murine SALL1 and SALL3 contain ZF2, 3, 4 and 5, while SALL4 does not contain ZF4. In the analysed non-*Drosophilid* insects we found a unique *sall* gene containing ZF2-5, the exception being the mosquito species where a unique *sall* gene lacking ZF5 was found. By contrast, all the *Drosophilid* species analysed contain two *sall* genes in their genomes, one of them encoding a protein equivalent to Sal (ZF2, 3 and 4) and the other to Salr (ZF2, 3, X and 5), suggesting that *sal/salr* duplication occurred after the *Drosophilid* divergence from other diptera. The presence of ZF4 in *Drosophila* Sal, and ZF5 in Salr, might be related to the differential repression mechanisms of these two Sall proteins.

Analysis of Sal transcriptional domains in human cells

Prior to protein domain analysis, we compared the *D. melanogaster* Sal and Salr sequences with that of other insect species to identify conserved peptide sequences (Fig. S2-S5). This

analysis revealed conserved blocks that were used as a guide to generate different fragments of the proteins (Figs. 4A and 5A). These fragments were cloned in frame with Gal4-DBD and a series of transient transfections were performed to analyse their localisation and transcriptional capacity. All the constructs were expressed in the 293FT cells and protein fragments showed the expected molecular weights (Figs. 4C and 5C).

The fragments showed different patterns of localisation, either cytoplasmic, nuclear or both. In the case of Sal (Fig. 4), S1 was mainly localised in the cytoplasm (Fig. 4D), while S2, S7, S9 and S10 were localised in cytoplasm and nucleus (Fig. 4E, J, L, M). Fragment S2, which contains the poly(Gln) sequence, and fragment S7 showed a punctuate pattern in the cytoplasm. All the fragments containing zinc fingers, in addition to fragments S3 and S5, were localised in the nucleus (Fig. 4F-I, K, N, O). The nuclear fragments S3, S4, S5, S6, S8 and S11 showed different distribution than the FL protein. Fragment S12 showed colocalisation with the FL CFP-Sal fusion protein, suggesting that ZF4 is not necessary for Sal subnuclear localisation (Fig. 6A).

We tested the transcriptional activity of the nuclear fragments of Sal (Fig. 4B). Surprisingly, fragments S3, S4, S5, and S11 showed strong transcriptional activation capacity. Although S5 fragment was expressed at lower levels than others (Fig. 4C), this fragment showed the highest activation capacity, having 5 times more activity than Gal4-DBD alone. Interestingly, cells overexpressing S5 were smaller than cells overexpressing other constructs, indicating an interaction of S5 with cell size. Fragments S6, S8, and S12 showed repression capacity similar to the FL protein. This suggests that Sal repression domains might be related to ZF3 and ZF4 and Sal activation is located between the poly(Gln) and ZF3 domains.

Analysis of Salr transcriptional domains in human cells

In the case of Salr, only the regions that contained ZFs were localised in the nucleus (Fig. 5F, H-J, N, O), R3 showing some cytoplasmic location (Fig. 5F), while fragments R4, R8, R9 and R10 were localised mainly in the cytoplasm (Fig. 5G, K-M). Fragments R1 and R2 were localised in nucleus and cytoplasm (Fig. 5D, E). Nuclear fragments exhibited different patterns. Fragments R5, R6 and R7 showed a diffused expression pattern in the nucleus (Fig. 5H-J), R11 was localised in the nuclear periphery (Fig. 5N), and R12 was localised in a similar way to FL Salr (Fig. 5O). Indeed, R12 colocalised with the fusion protein YFP-Salr (Fig. 6B), suggesting that ZF5 is not necessary for Salr localisation in subnuclear domains.

We analysed the transcriptional activity of the nuclear fragments R3, R5, R6, R7, R11 and R12 (Fig. 5B). The ZF2-containing fragment R3 of Salr showed significant transcriptional activation capacity (Fig. 5B), when compared to Gal4-Salr or Gal4-DBD alone. Fragments R5 and R11 showed some repression ability, although less than FL Gal4-Salr, while R6, R7, and R12 did not show statistically differences respect to the full length protein (Fig. 5B). These results suggest that part of Salr repression capacity resides in fragments R6 and R7, which contain ZFX and ZF5, respectively. In addition, R12, which lacks ZF5, showed a similar repression activity than FL Salr, suggesting that, although ZF5 shows repression activity by itself, it is not necessary for the total repression mediated by Salr.

We showed above that Salr might act through HDAC. To test which domain could be responsible of this repression, we analysed the transcriptional activity of the protein fragments in presence of TSA. Our analysis showed that R6, R7 and R12 presented

statistically differences in presence of TSA (Fig. 5B), suggesting that HDAC-dependent repression observed for Salr might be mediated through these fragments. Interestingly, fragments R6 and R7 contain ZFX and ZF5, the two ZFs characteristic of Salr. This suggests that Salr interaction with HDAC might depend on fingers ZFX and ZF5 and might explain the different mechanism observed between Sal and Salr. On the other hand, our results suggest that the lack of ZF5 does not affect the transcriptional repression by Salr, neither its dependence of HDAC, as there are no statistically significant differences between R12 and Salr (Fig. 5B).

Effect of sumoylation on Sall proteins' transcriptional activity

Sumoylation might modify the transcriptional activity of Sall proteins. We analysed the transcriptional capacity of the sumoylation-insensitive mutant forms of Sal and Salr [7]. These forms localise in a similar way than the WT proteins, but they are not modified by sumoylation *in vitro* and they behave different when overexpressed *in vivo*. The repression activity of Sal-IKDP was two-fold higher than that of the WT protein (Fig. 7). Surprisingly, in the case of Salr, the repression of capacity of Salr-IKEA-IKVA was lower than that of the WT protein (Fig. 7). These results suggest that, in our experimental settings, sumoylation modifies the repression capacity of Sal and Salr in opposite ways.

DISCUSSION

In this work we have shown that Sal and Salr act as transcriptional repressors in two different experimental settings. In *Drosophila* cells, Sal and Salr can repress transcription through an AT-rich sequence. In mammalian cells they act as repressors in a UAS/Gal4 heterologous system. Furthermore, we showed that both proteins contain activation and repression domains and that in our experimental settings they behave differently in relation to HDAC or sumoylation. Despite the limitations inherent to the use of a heterologous system, our results provided valuable information on Sal and Salr transcriptional activity, which is discussed below.

Comparison of insects and vertebrate *Drosophila* proteins

The *Drosophila* Sal and Salr proteins show transcriptional repressor activity in both *Drosophila* and mammalian cells. In *Drosophila* cells, both Sall proteins act through an AT rich sequence present in the upstream regulatory region of the chorion gene *s15*. This is similar to mouse Sall1, which also represses transcription through an AT-rich sequence [25]. In vertebrates, there are two distinct domains of repression: one on the N-terminal part of the proteins, which is not conserved in *Drosophila*, and one formed by the central region ZF3 and ZF2. Sal and Salr ZF3 regions show repressive capacity in a manner similar to that observed in vertebrates. Interestingly, the ZF3 region of Salr that binds to the AT rich sequence is by itself capable of causing an overexpression phenotype similar to that of the FL protein [6], suggesting that it may act in a dominant-negative fashion.

In contrast to Sall1, mouse Sall4 activates transcription through CT or CG rich sequences [28]. Surprisingly, Sal and Salr are also able to activate transcription through a region located between poly(Gln) and ZF3. Repression and activation capacities could reside in different SALL proteins in mammals, while both capacities could exist in the same molecule in *Drosophila*. With all SALL proteins, binding to specific sequences and

cell-specific protein partners could result in different SALL protein conformations, thus exposing either repression or activation domains. Although no direct transcriptional activation of any target gene by *Drosophila* Sall proteins has been reported, this could exist in certain cell types or developmental processes.

Regions S3 in Sal and R3 in Salr possess modest but reproducible activation capacity. These regions contain a box rich in non-polar aminoacids conserved among the insect sequences, suggesting a conserved function (Fig. S3, boxed in orange). Aside from the ZFs and the poly(Gln) region, two other conserved boxes among insects and mammals were identified. One of them is included in the regions S6 and R5, both of them exhibiting transcriptional repression activity (Figs. S1 and S4, boxed in red). The other box is located at the C-terminal end of ZF5, it is rich in charged aminoacids (Fig. S1 and S5, boxed in blue) and it is contained in the R7 fragment. The R7 fragment possesses repression capacity, although it has been shown to be dispensable for the general repression of Salr and also for its subnuclear localisation. Based on their high degree of conservation, these two novel motifs could be potentially relevant for the function of SALL proteins among species.

Sumoylation and the transcriptional activity of Sall proteins

In *Drosophila*, sumoylation increases the transcriptional activity of Vestigial (Vg), Dorsal, and p53 [39-42], while reducing the transcriptional activity of Stat92E, GCMa and SoxN [43]. One conclusion that emerges from our results is that sumoylation seems to affect the transcriptional capacity of Sal and Salr in opposite ways in our experimental settings. In the case of Sal a mutation that prevents sumoylation enhanced its capacity to repress transcription. In contrast, sumoylation-site mutations reduce Salr capacity to mediate transcriptional repression. This has been described as well for some mammalian transcription factors such as Elk-1 or SoxN, where sumoylation increases their repression activity [43-46].

The same "opposite" effect of sumoylation on the transcriptional activity of Sall proteins is in agreement with *in vivo* studies concerning *Drosophila* wing development [7]. We reported that the induction of ectopic wing vein formation by Sal overexpression is partially suppressed in Sal-IKDP mutants, whereas the capacity of Salr to induce veins increases when the sumoylation mutant form is overexpressed. According to the genetic analysis, Sall proteins and Sumo collaborate in the repression of veins in intervein regions. There, the lack of sumoylation of Salr could diminish its repression capacity and cause the generation of ectopic veins. Thus, the overexpression of the sumoylation mutant Salr-IKEA-IKVA promoted the formation of ectopic vein LIII. We hypothesise that sumoylated Salr is responsible for LIII repression in the intervein region based in its transcriptional repression capacity.

It has been proposed that sumoylation has a general role in transcriptional repression, as it is necessary for the interaction of transcription factors with the HDAC components. Such is the case for Elk-1, where mutations in the sumoylation domain block its interaction with HDAC [44-46]. Salr sumoylation sequence IKED is located in the ZF5-containing fragment R7, which we have demonstrated to contain TSA-responsive repressor activity. Therefore we hypothesise that mutation in the IKED sequence would interfere with its putative interaction with HDAC components.

In summary, our results provide insights into the role of *Drosophila* Sall proteins. We showed that both proteins act as transcriptional repressors in *Drosophila* and in

mammalian cells and that both contain repression and activation domains. In addition, we showed that Sal and Salr behave differently in relation to HDAC and that sumoylation affects their transcriptional capacity in opposite ways. Further work will be necessary to understand the role of the conserved boxes among insects and vertebrates and the relation between the sumoylation status of these proteins and their transcriptional capacity.

ACKNOWLEDGEMENTS

We thank R. Hjerpe, J. Ericsson, H. Saumweber, the *Drosophila* Genome Resource Center and Addgene for reagents. We thank C. Pérez technical assistance and A. Carracedo for advise in the sirtuin inhibitory assays. We thank D. Gubb for critical reading of the manuscript.

FUNDING

This work was supported by the Spanish Ministerio de Ciencia e Innovación [BFU2008-01884, RyC-05002168 and CSD2007-008-25120]; the Health Institute CarlosIII [PI070094/]; the Departments of Education [PI2009-16] and Industry of the Basque Government [Ertortek Research Programs 2008-2010]; and the Bizkaia County.

REFERENCES

- 1 de Celis, J. F. and Barrio, R. (2009) Regulation and function of Spalt proteins during animal development. *Int J Dev Biol* 53, 1385-1398
- 2 Sweetman, D. and Munsterberg, A. (2006) The vertebrate *spalt* genes in development and disease. *Dev Biol* 293, 285-293
- 3 Kuhnlein, R. P., Frommer, G., Friedrich, M., Gonzalez-Gaitan, M., Weber, A., Wagner-Bernholz, J. F., Gehring, W. J., Jackle, H. and Schuh, R. (1994) *spalt* encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the *Drosophila* embryo. *Embo J* 13, 168-179
- 4 Barrio, R., Shea, M. J., Carulli, J., Lipkow, K., Gaul, U., Frömmer, G., Schuh, R., Jäckle, H. and Kafatos, F. C. (1996) The *spalt-related* gene of *Drosophila melanogaster* is a member of an ancient family, defined by the adjacent, region-specific homeotic gene *spalt*. *Dev Genes Evol* 206, 315-325
- 5 de Celis, J. F. and Barrio, R. (2000) Function of the *spalt/spalt-related* gene complex in positioning the veins in the *Drosophila* wing. *Mech Dev* 91, 31-41
- 6 de Celis, J. F., Barrio, R. and Kafatos, F. C. (1996) A gene complex acting downstream of dpp in *Drosophila* wing morphogenesis. *Nature* 381, 421-424
- 7 Sánchez, J., Talamillo, A., Lopitz-Otsoa, F., Pérez, C., Hjerpe, R., Sutherland, J. D., Herboso, L., Rodríguez, M. S. and Barrio, R. (2010) Sumoylation modulates the activity of Spalt-like proteins during wing development in *Drosophila*. *J Biol Chem* 285, 25841-25849
- 8 Rusten, T. E., Cantera, R., Urban, J., Technau, G., Kafatos, F. C. and Barrio, R. (2001) Spalt modifies EGFR-mediated induction of chordotonal precursors in the

- embryonic PNS of *Drosophila* promoting the development of oenocytes. Development 128, 711-722
- 9 Elstob, P. R., Brodu, V. and Gould, A. P. (2001) *spalt*-dependent switching between two cell fates that are induced by the *Drosophila* EGF receptor. Development 128, 723-732
 - 10 Cantera, R., Luer, K., Rusten, T. E., Barrio, R., Kafatos, F. C. and Technau, G. M. (2002) Mutations in *spalt* cause a severe but reversible neurodegenerative phenotype in the embryonic central nervous system of *Drosophila melanogaster*. Development 129, 5577-5586
 - 11 Milan, M., Weihe, U., Perez, L. and Cohen, S. M. (2001) The LRR proteins capricious and Tartan mediate cell interactions during DV boundary formation in the *Drosophila* wing. Cell 106, 785-794
 - 12 Shea, M. J., King, D. L., Conboy, M. J., Mariani, B. D. and Kafatos, F. C. (1990) Proteins that bind to *Drosophila* chorion cis-regulatory elements: a new C2H2 zinc finger protein and a C2C2 steroid receptor-like component. Genes Dev 4, 1128-1140
 - 13 Kohlhase, J., Wischermann, A., Reichenbach, H., Froster, U. and Engel, W. (1998) Mutations in the SALL1 putative transcription factor gene cause Townes-Brocks syndrome. Nat Genet 18, 81-83
 - 14 Powell, C. M. and Michaelis, R. C. (1999) Townes-Brocks syndrome. J Med Genet 36, 89-93
 - 15 Al-Baradie, R., Yamada, K., St Hilaire, C., Chan, W. M., Andrews, C., McIntosh, N., Nakano, M., Martonyi, E. J., Raymond, W. R., Okumura, S., Okihiro, M. M. and Engle, E. C. (2002) Duane radial ray syndrome (Okihiro syndrome) maps to 20q13 and results from mutations in SALL4, a new member of the SAL family. Am J Hum Genet 71, 1195-1199
 - 16 Kohlhase, J., Heinrich, M., Schubert, L., Liebers, M., Kispert, A., Laccone, F., Turnpenny, P., Winter, R. M. and Reardon, W. (2002) Okihiro syndrome is caused by SALL4 mutations. Hum Mol Genet 11, 2979-2987
 - 17 Dong, P. D., Todi, S. V., Eberl, D. F. and Boekhoff-Falk, G. (2003) *Drosophila spalt/spalt-related* mutants exhibit Townes-Brocks' syndrome phenotypes. Proc Natl Acad Sci U S A 100, 10293-10298
 - 18 Netzer, C., Bohlander, S. K., Hinzke, M., Chen, Y. and Kohlhase, J. (2006) Defining the heterochromatin localisation and repression domains of SALL1. Biochim Biophys Acta 1762, 386-391
 - 19 Netzer, C., Rieger, L., Brero, A., Zhang, C. D., Hinzke, M., Kohlhase, J. and Bohlander, S. K. (2001) SALL1, the gene mutated in Townes-Brocks syndrome, encodes a transcriptional repressor which interacts with TRF1/PIN2 and localises to pericentromeric heterochromatin. Hum Mol Genet 10, 3017-3024
 - 20 Kiefer, S. M., McDill, B. W., Yang, J. and Rauchman, M. (2002) Murine Sall1 represses transcription by recruiting a histone deacetylase complex. J Biol Chem 277, 14869-14876
 - 21 Kiefer, S. M., Ohlemiller, K. K., Yang, J., McDill, B. W., Kohlhase, J. and Rauchman, M. (2003) Expression of a truncated Sall1 transcriptional repressor is responsible for Townes-Brocks syndrome birth defects. Hum Mol Genet 12, 2221-2227

- 22 Lauberth, S. M., Bilyeu, A. C., Firulli, B. A., Kroll, K. L. and Rauchman, M. (2007) A phosphomimetic mutation in the Sall1 repression motif disrupts recruitment of the nucleosome remodeling and deacetylase complex and repression of Gbx2. *J Biol Chem* 282, 34858-34868
- 23 Lauberth, S. M. and Rauchman, M. (2006) A conserved 12-amino Acid motif in sall1 recruits the nucleosome remodeling and deacetylase corepressor complex. *J Biol Chem* 281, 23922-23931
- 24 Sato, A., Kishida, S., Tanaka, T., Kikuchi, A., Kodama, T., Asashima, M. and Nishinakamura, R. (2004) Sall1, a causative gene for Townes-Brocks syndrome, enhances the canonical Wnt signaling by localising to heterochromatin. *Biochem Biophys Res Commun* 319, 103-113
- 25 Yamashita, K., Sato, A., Asashima, M., Wang, P. C. and Nishinakamura, R. (2007) Mouse homolog of SALL1, a causative gene for Townes-Brocks syndrome, binds to A/T-rich sequences in pericentric heterochromatin via its C-terminal zinc finger domains. *Genes Cells* 12, 171-182
- 26 Li, D., Tian, Y., Ma, Y. and Benjamin, T. (2004) p150(Sal2) is a p53-independent regulator of p21(WAF1/CIP). *Mol Cell Biol* 24, 3885-3893
- 27 Yang, J., Chai, L., Liu, F., Fink, L. M., Lin, P., Silberstein, L. E., Amin, H. M., Ward, D. C. and Ma, Y. (2007) Bmi-1 is a target gene for SALL4 in hematopoietic and leukemic cells. *Proc Natl Acad Sci U S A* 104, 10494-10499
- 28 Wu, Q., Chen, X., Zhang, J., Loh, Y. H., Low, T. Y., Zhang, W., Zhang, W., Sze, S. K., Lim, B. and Ng, H. H. (2006) Sall4 Interacts with Nanog and Co-occupies Nanog Genomic Sites in Embryonic Stem Cells. *J Biol Chem* 281, 24090-24094
- 29 Zhang, J., Tam, W. L., Tong, G. Q., Wu, Q., Chan, H. Y., Soh, B. S., Lou, Y., Yang, J., Ma, Y., Chai, L., Ng, H. H., Lufkin, T., Robson, P. and Lim, B. (2006) Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol* 8, 1114-1123
- 30 Wilkinson, K. A. and Henley, J. M. (2010) Mechanisms, regulation and consequences of protein SUMOylation. *Biochem J* 428, 133-145
- 31 Rehwinkel, J., Behm-Ansmant, I., Gatfield, D. and Izaurralde, E. (2005) A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *Rna* 11, 1640-1647
- 32 Sundqvist, A. and Ericsson, J. (2003) Transcription-dependent degradation controls the stability of the SREBP family of transcription factors. *Proc Natl Acad Sci U S A* 100, 13833-13838
- 33 Yanagawa, S., Lee, J. S. and Ishimoto, A. (1998) Identification and characterization of a novel line of *Drosophila* Schneider S2 cells that respond to wingless signaling. *J Biol Chem* 273, 32353-32359
- 34 Cherbas, P., Cherbas, L., Lee, S. S. and Nakanishi, K. (1988) 26-[125I]iodoponasterone A is a potent ecdysone and a sensitive radioligand for ecdysone receptors. *Proc Natl Acad Sci U S A* 85, 2096-2100
- 35 Shaw, G., Morse, S., Ararat, M. and Graham, F. L. (2002) Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells. *Faseb J* 16, 869-871
- 36 Nishinakamura, R. and Osafune, K. (2006) Essential roles of sall family genes in kidney development. *J Physiol Sci* 56, 131-136

- 37 Frye, R. A. (2000) Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. *Biochem Biophys Res Commun* 273, 793-798
- 38 Finkel, T., Deng, C. X. and Mostoslavsky, R. (2009) Recent progress in the biology and physiology of sirtuins. *Nature* 460, 587-591
- 39 Bhaskar, V., Smith, M. and Courey, A. J. (2002) Conjugation of Smt3 to dorsal may potentiate the *Drosophila* immune response. *Mol Cell Biol* 22, 492-504
- 40 Bhaskar, V., Valentine, S. A. and Courey, A. J. (2000) A functional interaction between dorsal and components of the Smt3 conjugation machinery. *J Biol Chem* 275, 4033-4040
- 41 Mauri, F., McNamee, L. M., Lunardi, A., Chiacchiera, F., Del Sal, G., Brodsky, M. H. and Collavin, L. (2008) Modification of *Drosophila* p53 by SUMO modulates its transactivation and pro-apoptotic functions. *J Biol Chem* 283, 20848-20856
- 42 Takanaka, Y. and Courey, A. J. (2005) SUMO enhances vestigial function during wing morphogenesis. *Mech Dev* 122, 1130-1137
- 43 Savare, J., Bonneaud, N. and Girard, F. (2005) SUMO represses transcriptional activity of the *Drosophila* SoxNeuro and human Sox3 central nervous system-specific transcription factors. *Mol Biol Cell* 16, 2660-2669
- 44 Girdwood, D., Bumpass, D., Vaughan, O. A., Thain, A., Anderson, L. A., Snowden, A. W., Garcia-Wilson, E., Perkins, N. D. and Hay, R. T. (2003) P300 transcriptional repression is mediated by SUMO modification. *Mol Cell* 11, 1043-1054
- 45 Nayak, A., Glockner-Pagel, J., Vaeth, M., Schumann, J. E., Buttmann, M., Bopp, T., Schmitt, E., Serfling, E. and Berberich-Siebelt, F. (2009) SUMOylation of the transcription factor NFATc1 leads to its subnuclear relocalisation and IL2 repression by HDAC. *J Biol Chem* 284, 10935-10946
- 46 Yang, S. H. and Sharrocks, A. D. (2005) PIASx acts as an Elk-1 coactivator by facilitating derepression. *Embo J* 24, 2161-2171

FIGURE LEGENDS

Figure 1. Sal and Salr are transcriptional repressors in *Drosophila* cells.

(A, B) Confocal micrographs showing in green the localisation of CFP-Sal (A) or YFP-Salr (B) in subnuclear domains of *Drosophila* Kc167 cells. DAPI nuclear marker is shown in magenta. Grey pictures show single green channels (A', B'). Bars indicate 4 μ m. (C, D) Graphic representation of Renilla luciferase activity of *pAC-SalrBE-RLuc* (C) or *pAC-SalrBE*-RLuc* (D) in absence (-) or presence of vectors expressing Sal or Salr. Values are the average of three independent experiments performed in quadruplicates. Error bars indicate standard deviation. Asterisks indicate statistically significant differences ($p \leq 0.005$).

Figure 2. *Drosophila* Sal and Salr are transcriptional repressors in human cells.

(A, B) Confocal micrographs showing in green the localisation of Gal4-Sal (A) or Gal4-Salr (B) in subnuclear domains of human 293FT cells. DAPI nuclear marker is shown in magenta. Grey pictures show single green channels (A', B'). Bars indicate 4 μ m. (C) Graphic representation of firefly luciferase activity of *pGal4tkLUC* alone (Gal4), Gal4-DBD fused to FL Sal (Gal4-Sal) or FL Salr (Gal4-Salr), either in presence of DMSO (black bars), TSA (dark grey bars), NAM (white bars) or EX527 (light grey bars). Values are the average of three independent transfections performed in triplicates. Error bars indicate standard deviation. Asterisks indicate statistically significant differences between Gal4-DBD and Gal4-Sal or Gal4-Salr (black asterisks, $p \leq 10^{-6}$) or between Gal4-Salr with or without inhibitors treatments (white asterisks, $p \leq 0.005$).

Figure 3. Zinc fingers distribution in Sall proteins is conserved throughout evolution.

Upper part, schematic representation of the zinc fingers and poly(Gln) domains in SALL proteins of the indicated species belonging to the *phyla* specified to the left. Bottom part, unrooted N-J tree analysis of the zinc finger domains of vertebrate and invertebrate SALL proteins, shown in Supplementary Fig. S1. Species analysed and accession numbers of the sequences were: Bruma, *Brugia malayi*, A8P9M2; Cael, *Caenorhabditis elegans*, Q17396; Drome, *Drosophila melanogaster*, P39770 (Sal) and P91639 (Salr); Hs, *Homo sapiens*, Q9NSC2 (SALL1), Q9Y467 (SALL2), Q9BXA9 (SALL3) and Q9UJQ4 (SALL4); Pipra, *Pristionchus pacificus*, Q27ZK0; Sako, *Saccoglossus kowalevskii*, XP_002738555; Stpur, *Strongylocentrotus purpuratus*, XP_781376.

Figure 4. Analysis of Sal repression and activation domains in human cells.

(A) Schematic representation of the S1-S12 fragments used in the analysis. Ovals represent zinc fingers and grid square represents the poly(Gln) domain. (B) Graphic representation of firefly luciferase activity of *pGal4tkLUC* alone (Gal4), Gal4 fused to FL Sal (FL-Sal) or Gal4 fused Sal fragments (S3-S6, S8, S11, S12). Values are the average of three independent transfections performed in triplicates. Error bars indicate standard deviation. Asterisks indicate statistically significant differences respect to FL-Sal ($p \leq 10^{-5}$). (C) Western blot analysis using anti-Gal4 antibodies showing the expression of the constructs

used in the luciferase assays. Molecular weight marker in kilodaltons is indicated to the right of each panel. Expected size for each construct is indicated in the Supplementary Experimental section. **(D-O)** Confocal micrographs showing in green the localisation of the indicated Sal fragments in 293FT cells. DAPI nuclear marker is shown in magenta. **(D'-O')** Black and white panels show single green channels. Bar indicates 4 μm . All pictures were taken at the same magnification.

Figure 5. Analysis of Salr repression and activation domains in human cells.

(A) Schematic representation of the R1-R12 fragments used in the analysis. Ovals represent zinc fingers and grid square represents the poly(Gln) domain. **(B)** Graphic representation of firefly luciferase activity of *pGal4tkLUC* alone (Gal4), Gal4 fused to FL Salr (FL-Salr) or Gal4 fused Salr fragments (R3, R5-R7, R11, R12), either in presence of DMSO (black bars) or TSA (grey bars). Values are the average of three independent transfections performed in triplicates. Error bars indicate standard deviation. Black asterisks indicate statistically significant differences respect to FL-Salr ($p \leq 10^{-4}$). Empty asterisks indicate statistically significant differences of each fragment in presence or absence of TSA ($p \leq 10^{-3}$). **(C)** Western blot analysis using anti-Gal4 antibodies showing the expression of the constructs used in the luciferase assays. Molecular weight marker in kilodaltons is indicated to the right of each panel. Expected size for each construct is shown in the Supplementary Experimental section. **(D-O)** Confocal micrographs showing in green the localisation of the indicated Salr fragments in 293FT cells. DAPI nuclear marker is shown in magenta. **(D'-O')** Black and white panels show single green channels. Bar indicates 4 μm . All pictures were taken at the same magnification.

Figure 6. S12 and R12 fragments colocalise with the full-length proteins.

(A, B) Confocal micrographs showing colocalisation of CFP-Sal **(A)** or YFP-Salr **(B)** shown in green, with the S12 or R12 fragments shown in magenta, respectively. Black and white panels show single green **(A', B')** or magenta pictures **(A'', B'')**. Bar indicates 4 μm . Pictures were taken at the same magnification.

Figure 7. Sumoylation influences the transcriptional activity of *Drosophila* Sal proteins in human cells.

Graphic representation of firefly luciferase activity of *pGal4tkLUC* alone (Gal4), Gal4 fused to FL Sal (Gal4-Sal), sumoylation-insensitive Sal mutant (Gal4-Sal-IKDP), FL Salr (Gal4-Salr) or sumoylation-insensitive Salr mutant (Gal4-Salr-IKEA-IKVA). Values are the average of triplicate transfections. Error bars indicate standard deviation. Asterisks indicate statistically significant differences ($p < 0.005$).













