Transcriptional Activity and Nuclear Localization of Cabut, the *Drosophila* Ortholog of Vertebrate TGF-β-Inducible Early-Response Gene (TIEG) Proteins

Yaiza Belacortu[1], Ron Weiss[2], Sebastian Kadener[2], Nuria Paricio[1]*

1 Departamento de Genética, Facultad CC Biológicas, Universidad de Valencia, Burjasot, Spain, 2 Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat-Ram, Jerusalem, Israel

Abstract

Background: Cabut (Cbt) is a C_2H_2-class zinc finger transcription factor involved in embryonic dorsal closure, epithelial regeneration and other developmental processes in *Drosophila melanogaster*. Cbt orthologs have been identified in other *Drosophila* species and insects as well as in vertebrates. Indeed, Cbt is the *Drosophila* ortholog of the group of vertebrate proteins encoded by the TGF-ß-inducible early-response genes (TIEGs), which belong to Sp1-like/Krüppel-like family of transcription factors. Several functional domains involved in transcriptional control and subcellular localization have been identified in the vertebrate TIEGs. However, little is known of whether these domains and functions are also conserved in the Cbt protein.

Methodology/Principal Findings: To determine the transcriptional regulatory activity of the *Drosophila* Cbt protein, we performed Gal4-based luciferase assays in S2 cells and showed that Cbt is a transcriptional repressor and able to regulate its own expression. Truncated forms of Cbt were then generated to identify its functional domains. This analysis revealed a sequence similar to the mSin3A-interacting repressor domain found in vertebrate TIEGs, although located in a different part of the Cbt protein. Using β-Galactosidase and eGFP fusion proteins, we also showed that Cbt contains the bipartite nuclear localization signal (NLS) previously identified in TIEG proteins, although it is non-functional in insect cells. Instead, a monopartite NLS, located at the amino terminus of the protein and conserved across insects, is functional in *Drosophila* S2 and *Spodoptera exigua* Sec301 cells. Last but not least, genetic interaction and immunohistochemical assays suggested that Cbt nuclear import is mediated by Importin-α2.

Conclusions/Significance: Our results constitute the first characterization of the molecular mechanisms of Cbt-mediated transcriptional control as well as of Cbt nuclear import, and demonstrate the existence of similarities and differences in both aspects of Cbt function between the insect and the vertebrate TIEG proteins.

Introduction

cabut (*cbt*) encodes a *Drosophila* transcription factor (TF) containing three C_2H_2 zinc finger motifs at the carboxy (C) terminus and a serine-rich (SR) region at the amino (N) terminus [1]. This protein is involved in dorsal closure during *Drosophila* embryogenesis [1], but it is also required for other developmental processes such as the ecdysone response [2], neuroendocrine cell remodeling [3], epithelial regeneration [4], circadian rhythms [5], axon guidance and synaptogenesis [6,7], pole cell formation [8], cell growth [9,10], autophagic cell death [11], cell cycle progression (A.J. Katzaroff and E.A. Bruce, personal communication) and cell proliferation and patterning [12]. Experiments in *Drosophila* embryos and S2 cells have shown that Cbt is a nuclear protein, although it is also present in axons in the central and peripheral nervous systems [13]. Cbt orthologs have been

identified in other *Drosophila* species and insects, including the mosquito (*Anopheles gambiae* and *Aedes aegypti*), red flour beetle (*Tribolium castaneum*), honeybee (*Apis mellifera*) and silkworm (*Bombyx mori*), as well as in other invertebrate organisms such as ascidians, echinoderms and crustaceans [14,15]. The expression patterns of *cbt* transcripts and proteins during embryonic development are highly conserved among Drosophilidae [13,14]. Interestingly, Cbt also presents high similarity to the vertebrate proteins encoded by the TGF-β-inducible early-response genes (TIEGs) [14,16] and has also been named *Drosophila TIEG* (*dTIEG*) [12].

TIEG proteins belong to subgroup III of the Sp1-like/Krüppel-like family of TFs, which contain three highly conserved C-terminal C_2H_2-type zinc finger motifs that mediate binding to GC-rich promoter sequences [17,18,19,20,21]. Their expression is regulated by a plethora of growth factors (e.g., TGF-β superfamily), cytokines (e.g., BMP family and activin A) and hormones (e.g.,

estrogens) (reviewed in [20]). Several proteins of this family have been characterized, including TIEG1 (Krüppel-like factor 10, KLF10) and TIEG3 (KLF11) in humans and mice, and and TIEG2/3 (KLF11) in mice (17,19,20,21, reviewed in [18]). They have been also identified in rat, monkey, pig and zebrafish genomes [14,22]. TIEG proteins are involved in numerous processes, including, among others, proliferation, apoptosis, differentiation, cancer and circadian rhythms [17,23,24,25,26,27,28,29,30]. These proteins can function as either transcriptional repressors [31,32] or activators [32,33,34,35], depending on the cellular context, the promoter to which they bind and the coregulators with which they interact [33,34,35]. Several studies have identified and character-ized functional domains in TIEG proteins. One proline-rich (PR) region and three repression domains (R1, R2 and R3) were identified at the N-terminal region of TIEG proteins [19,31]. Interestingly, a mammalian mSin3A-interacting domain (SID) was identified in the R1 domain of the TIEG3 protein and shown to be essential for TIEG3-mediated transcriptional repression in cell culture [31,32]. This domain interacts with the co-repressor mSin3A, which inhibits transcriptional activation of target genes by histone deacetylation and subsequent remodeling of chromatin structure [36]. Different sequences are found in the R2 and R3 domains [31]. Moreover, the C-terminal end of TIEG3 contains the DNA-binding domain (DBD) and an additional downstream domain, both of which are able to activate transcription in OLI-neu and HeLa cells [31,32]. More recently, other domains involved in transcriptional regulation have been identified in TIEG proteins, including an N-terminal domain in TIEG1 which interacts with the enzyme Jumonji AT-rich interactive domain 1B/lysine-specific demethylase 5B (JARID1B/KDM5B) and mediates transcriptional repression [37], and a C-terminal domain in TIEG2 that interacts with the p300 co-activator to activate expression of the *Pdx1* gene [33]. Regarding their nuclear localization, TIEGs and other KLF proteins, contain a bipartite NLS within the zinc finger domains, that is required for transport to the nucleus [32,38,39]. In general, NLSs consist of either one (monopartite) or two (bipartite) stretches of basic amino acids (usually arginine (R) and lysine (K)) separated by an intervening region of 10–12 residues and recognized by protein carriers called importins [40]. The NLSs frequently overlap with DBDs [41], as occurs in the TIEG3 protein [32]. Because nuclear transport of TFs is essential for cellular function, regulation of TF nuclear availability through NLSs directly affects gene expression, cell growth and proliferation [42].

Cbt is the *Drosophila* ortholog of vertebrate TIEG proteins [14,16] and shares functions with several family members, e.g., rat TIEG1 and murine TIEG3, as it is involved in circadian rhythms as well as cell proliferation and positive regulation of TGF-β signaling [5,12]. Regarding its transcriptional activity, previous results suggested that Cbt may function as an activator of gene expression. We showed that *decapentaplegic* (*dpp*) expression was downregulated at the leading edge of the lateral epidermal sheets during dorsal closure in *cbt* mutant embryos [1]. Cbt also positively regulates the expression of *STAT92E*, *spalt* (*sal*) and *optomotorblind* (*omb*) genes in wing imaginal discs [12]. Although these results suggest that Cbt may activate gene expression in embryos and wing discs, it is not clear whether the transcriptional regulation of these target genes is direct or indirect. In the present study, we performed a functional characterization of the Cbt protein by examining its transcriptional regulatory potential and identifying its functional domains. Gal4-based transcriptional assays in S2 cells demonstrated that Cbt is a transcriptional repressor and contains a SID similar to the one identified in TIEG proteins

[31,32]. We also report that Cbt can downregulate its own expression, probably by directly binding to a sequence located 1 kb upstream of the gene's transcription start site. Finally, we provide evidence that Cbt nuclear localization is mediated by a monopartite NLS located at the N-terminal region of the protein ($_{71}$PNKKPRL$_{77}$), which is conserved in Cbt orthologs from other *Drosophila* species and insects. Genetic interaction assays and immunostaining using *importin-α* mutant strains suggested that the Importin-α2 protein is involved in Cbt nuclear import in *Drosophila*. Together, these results expand our understanding of the mechanisms of Cbt transcriptional regulation and nuclear import, which reveal the biochemical similarities and differences between vertebrate TIEG proteins and Cbt.

Materials and Methods

Plasmid constructs

pIE-β-GalCbt$_{1-428}$, peGFP-Cbt$_{1-428}$ and pIE-Gal4 constructs were generated by in-frame cloning of the *cbt* and *Gal4* coding regions, obtained by PCR amplification using the Pwo Polymerase (Roche diagnostics GmbH, Mannheim, Germany) and the oligos described in Table 1, into the pIE-β-Gal vector without β-Gal stop using the SalI- BamHI sites [43], the peGFP-C3 vector (Clontech Laboratories, Mountain View, CA) using the EcoRI-Asp718I sites and the pIE1-3 vector (Novagen, Madison, WI, USA) using the NotI-Asp718I sites, respectively. Other constructs used in this work were obtained by PCR amplification using pIE-β-GalCbt$_{1-428}$ and peGFPCbt$_{1-428}$ as templates and the oligos described in Table 1, followed by cloning into the pIE-β-Gal, peGFP-C3 and pIE-Gal4 vectors. All of the constructs were confirmed by DNA sequencing. The pG5DE5tkLuc plasmid, which contains the *luciferase* gene under the control of the UAS sequence, was a gift of Dr. Courey (UCLA, Los Angeles, CA, USA). Site-directed mutagenesis was carried out on the pIE-β-GalCbt$_{1-77}$ construct to mutate lysines 73 and 74 (K73 and K74) to asparagine (N). Mutagenesis was performed by GenScript (NJ, USA).

Fly stocks. For genetic interaction assays, a *sev*-Gal4>UAS-Cbt$_{FL}$ line was generated via recombination. The following lines were obtained from the Bloomington Stock Center (http://flystocks.bio.indiana.edu/): y^1 w^*; *P{EP}kapα1^{G4113}*, y^1 w^{67c23}; *P{EPgy2}pen^{EY09095}*, w^{1118}; *P{EP}kapα3C^{G8397}/TM6C, Sb^1* and *71B*-Gal4. The *w;impα2^{D14}/Cyo Actin-GFP* stock [44] was a gift of Dr. Schwarz (Harvard Medical School, Boston, MA, USA). The following *importin* RNAi lines were obtained from the Vienna *Drosophila* RNAi Center (http://stockcenter.vdrc.at/control/main): w^{1118};*P {GD13960}v28920/Cyo*, w^{1118};*P {GD10665}v34265* and w^{1118};*P {GD14213}v36104*. All *Drosophila* strains were maintained at 25°C.

Cell culture and transfection conditions

Drosophila melanogaster Schneider 2 cells (S2) and *Spodoptera exigua* cells (Sec301) were grown at 25°C in Schneider's *Drosophila* Medium with L-Glutamine (Biological Industries, Jerusalem, Israel/Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (FBS, Invitrogen) and 1% penicillin/strepto-mycin (Invitrogen). Chinese hamster ovary (CHO-K1) cells from *Cricetulus griseus* were grown at 37°C and 5% CO_2 in DMEM/F-12 medium (Gibco/Invitrogen) supplemented with 10% FBS and 1% penicillin/streptomycin. For subcellular localization assays, 1×10^6 cells/ml were seeded onto coverslips in 24-well plates. After 24 h, cells were transfected with 0.5–1 μg of DNA for 5 h using 8 μl of Cellfectine reagent (Invitrogen) for insect cells and 6 μl of FugeneHD (Roche) for CHO-K1 cells.

Table 1. PCR-generated constructs and associated primers.

Construct	Oligonucleotides 5′-3′
pIE-β-GalCbt$_{1-428}$	F (SalI):ACGC<u>GTCGAC</u>GGGCGGCGGC**ATG**GACATGGATACCTTGC
	R (BamHI):CGC<u>GGATCCC</u>G**TCA**TCATCGCTGCAGTTGAAG
pIE-β-GalCbt$_{262-428}$	F(SalI):<u>GTCGAC</u>GCGGCCGCCCAGGCGGCGGCCA
	R (BamHI):CGC<u>GGATCCC</u>G**TCA**TCATCGCTGCAGTTGAAG
pIE-β-GalCbt$_{1-141}$	F (SalI):ACGC<u>GTCGAC</u>GGGCGGCGGC**ATG**GACATGGATACCTTGC
	R (BamHI):CGC<u>GGATCCC</u>G**TCA**GTTGACCCGCATGATGAC
pIE-β-GalCbt$_{1-108}$	F (SalI):ACGC<u>GTCGAC</u>GGGCGGCGGC**ATG**GACATGGATACCTTGC
	R (BamHI): <u>CGGATCCC</u>G**TCA**CAGTGGGTGGTACTGAAC
pIE-β-GalCbt$_{1-77}$	F (SalI):ACGC<u>GTCGAC</u>GGGCGGCGGC**ATG**GACATGGATACCTTGC
	R (BamHI): CGC<u>GGATCCC</u>G**TCA**CAAACGGGGTTTCTT
pIE-β-GalCbt$_{1-70}$	F (SalI):ACGC<u>GTCGAC</u>GGGCGGCGGC**ATG**GACATGGATACCTTGC
	R (BamHI): CGC<u>GGATCCC</u>G**TCA**AACCGCTATTTCGGGGAGCCTCGTCTT
pIE-β-GalCbt$_{1-70KKPR}$	F (SalI):ACGC<u>GTCGAC</u>GGGCGGCGGC**ATG**GACATGGATACCTTGC
	R (BamHI):CGC<u>GGATCCC</u>G**TCA**ACGGGGTTTCTTAACCGCTATTCG
pIE-β-Gal-$_{PNKKPRL}$	F (NotI):<u>GCGGCCGC</u>**ATG**AGCGAAAAATACATCG
	R (SalI):CGC<u>GTCGAC</u>**TCA**GTTTGCCCCAAAGAACAACCCTTTTTGACACCAGA
pIE-$_{PNKKPRL}$-β-Gal	F (NotI):<u>GCGGCCGC</u>**ATG**CCCAACAAGAAACCCCGTTTGAGCGAAAAATACATCGTC
	R(BamHI):<u>GGATCC</u>TTATTACGTCGACCCTTTTGACACCAGACCAAC
pIE-β-GalCbt$_{37-77}$	F (SalI):ACG<u>GGTCGAC</u>GAAGGCAAAGCTCAAG
	R (Bam-HI):CGC<u>GGATCCC</u>G**TCA**CAAACGGGGTTTCTT
peGFPCbt$_{1-428}$	F (EcoRI): CC<u>GGAATTCC</u>G**ATG**GACATGGATACCTTGC
	R (EcoRI):TATGAAT**TCA**TGGACATGGATACCTTGC
peGFPCbt$_{1-292}$	F (EcoRI): CC<u>GGAATTCC</u>G**ATG**GACATGGATACCTTGC
	R (KpnI/Asp718I):CGG<u>GGTACCC</u>**TCA**GGGTCGCTCGCCGGTGTG
peGFPCbt$_{1-322}$	F (EcoRI): CC<u>GGAATTCC</u>G**ATG**GACATGGATACCTTGC
	R (KpnI/Asp718I): CG<u>GGTACCC</u>**TCA**TTTCTTTTCACCGGTGTG
PeGFPCbt$_{1-262}$	F (EcoRI): CC<u>GGAATTCC</u>G**ATG**GACATGGATACCTTGC
	R (KpnI/Asp718I):CGG<u>GGTACCC</u>**TCA**GATGCGACTTCTGGTGGC
PeGFPCbt$_{262-428}$	F (KpnI/Asp718I):CC<u>GGAATTCC</u>GACGGCCGCCCAGGCGGCGGCCA
	R (EcoRI):TATGAAT**TCA**TGGACATGGATACCTTGC
pIE-Gal4$_{STOP}$	F(NotI):AATG<u>CGGCCGC</u>AGAC**ATG**AAGCTACTGTCTTCT
	R (BamHI):CGC<u>GGATCC</u>GTACAG**TCA**ACTGTCTTTGACCTT
pIE-Gal4$_{ΔStop}$	F(NotI):AATG<u>CGGCCGC</u>AGAC**ATG**AAGCTACTGTCTTCT
	R (BamHI):CGC<u>GGATCC</u>GTACAGACTGTCTTTGACCTT
pIE-Gal4Cbt$_{1-428}$	F (BamHI): CGC<u>GGATCC</u>GC**ATG**GACATGGATACC
	R (BamHI):CGC<u>GGATCC</u>G**TCA**TCGCTGCACTTGAAGCAG
pIE-Gal4Cbt$_{1-262}$	F(BamHI):CGC<u>GGATCC</u>GC**ATG**GACATGGATACC
	R (BamHI):CGC<u>GGATCC</u>G**TCA**GTAGATGCGACTTCTGGTGG
pIE-Gal4Cbt$_{1-182}$	F (BamHI): CGC<u>GGATCC</u>GC**ATG**GACATGGATACC
	R (BamHI):CGC<u>GGATCC</u>G**TCA**TGGTGTCTCGGGTGTTT
pIE-Gal4Cbt$_{1-165}$	F (BamHI): CGC<u>GGATCC</u>GC**ATG**GACATGGATACC
	R (BamHI):CGC<u>GGATCC</u>GAGGAGG**TCA**TCGCGAGTTCATTTTGAACTTGAGATT
pIE-Gal4Cbt$_{173-428}$	F (BamHI):CGC<u>GGATCC</u>GAGGAGGTCGCGA**ATG**GCCGCCCAGGCGGCGG
	R (BamHI):CGC<u>GGATCC</u>G**TCA**TCGCTGCACTTGAAGCAG
pIE-Gal4Cbt$_{261-347}$	F (BamHI):CGC<u>GGATCC</u>GC**ATG**ATCTACGAGTGCAGT
	R (BamHI): CGC<u>GGATCC</u>GC**TCA**GTCCTTGTTGTGCCG
pIE-Gal4Cbt$_{345-428}$	F (BamHI):CGC<u>GGATCC</u>GC**ATG**AACAAGGACAAGGCG
	R (BamHI):CGC<u>GGATCC</u>G**TCA**TGGACATGGATACCTTGC
pIE-Gal4Cbt$_{261-389}$	F (BamHI):CGC<u>GGATCC</u>GC**ATG**ATCTACGAGTGCAGT
	R (BamHI):CGC<u>GGATCC</u>GC**TCA**AGCGCTGGAGCCCGC

Sequences recognized by restriction enzymes (in parentheses) are underlined. The start and stop codons are in bold. F, forward primer; R, reverse primer.
doi:10.1371/journal.pone.0032004.t001

Repression and UAS/Gal4-based transcriptional assays

For the repression assay, S2 cells were transfected with 0.5 μg of the pMT-Cbt-V5 plasmid, 0.5 μg of the pHStinger-Prom1-2 plasmid (see [13]) and 0.1 μg of the PAC-cherry vector (to normalize for transfection efficiency) in 6-well plates at 70–90% confluence according to supplier's recommendations (Transfection Custom Insect Reagent, Mirus Bio Corp, Madison, WI). Expression of the recombinant protein was induced by incubation in medium containing 0.25/0.75 μM copper sulfate for 48 h. Fluorescence was measured using a TECAN infinit® 200 plate reader. For the UAS/Gal4 assays, S2 cells were co-transfected in 24-well plates with 0.075 μg of each construct, 0.1 μg of the pG5DE5tkLuc vector and 0.1 μg of the PAC-Renilla vector to normalize for transfection efficiency. The transfection protocol was the same as for the repression assay. After 48 h, cells were lysed and incubated with the Dual Luciferase Assay Kit (Promega, Madison, WI, USA) following the manufacturer's instructions. Luciferase activity was measured using a Modulus single tube multimode reader.

Immunochemistry and scanning electron microscopy

Cells were washed with phosphate-buffer saline (PBS) 16–24 h after transfection, fixed in freshly prepared 4% paraformaldehyde/PBS for 20 min, permeabilized with 0.1% Triton-X-100/PBS for 5 min, and blocked with 1% BSA/PBS for 30 min. After incubation with rabbit anti-β-Gal (1:1000, Cappel/ICN Biomedicals, Ohio, USA) and mouse monoclonal anti-Lamin (1:200) (Developmental Studies Hybridoma Bank, Iowa, USA) antibodies for 2 h, cells were washed three times with 0.5% BSA/PBS, incubated with anti-rabbit-FITC (1:200, Calbiochem EMD Biosciences Inc., La Jolla, CA) and anti-mouse-Cy3 (1:200, Invitrogen) antibodies for 1 h, washed three times with PBS, mounted in Vectashield (Vector Laboratory, Peterborough, UK) and examined under a Leica TCS SP1 confocal microscope. Ovaries were dissected as described [45] and stained using the anti-CbtΔZn antibody [13]. Scanning electron microscopy analysis of adult eyes was performed following the critical point dry method using a Hitachi S4100 microscope, as previously described [46].

Western blotting

Cells were collected 16–24 h after transfection, washed with PBS and scraped into 50 μl of sodium dodecyl sulfate-polyacryl-amide gel electrophoresis (SDS-PAGE) loading buffer. After boiling, aliquots were subjected to 10% SDS-PAGE. Subsequently, proteins were electrotransferred to PDVF membranes (Roche) and detected with rabbit anti-β-Gal (1:5000, Cappel) or anti-eGFP (1:2000, Roche) primary antibodies. The secondary antibody was peroxidase-labeled anti-rabbit (1:3000, Calbiochem). The ECL western blotting detecting reagent (Pierce, Rockford, IL, USA) was then used to reveal the chemiluminescence.

Computational analyses

PSORT II (http://psort.nibb.ac.jp/form2.html) and cNLS Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi) software programs were used to predict NLSs and nuclear export signal (NESs). Multiple alignments of protein sequences were performed with the ClustalW algorithm [47]. The automated protein structure homology-modeling server SWISS-MODEL (http://swissmodel.expasy.org/) was used to analyze secondary structure. NetPhos (http://www.cbs.dtu.dk/services/NetPhos/) and NetPhosK (http://www.cbs.dtu.dk/services/NetPhosK/) programs were used to predict putative phosphorylation sites and specific kinase-binding sites. The GENPEPT database (invertebrate and vertebrate sections; GenBank) was searched with the PX(K/R)KX(R/L) string (where X = any or no amino acid) using the Scansite program (http://scansite.mit.edu; Quick matrix method).

Results

Cabut acts as a potent transcriptional repressor and regulates its own transcription

The UAS/Gal4 fusion system [48] was used in *Drosophila* S2 cells to determine the transcriptional repression and/or activation activity of Cbt. Gal4 fusion proteins have little transcriptional background interference in *Drosophila* cells because of their yeast origin. For these experiments, we generated the pIE-Gal4Cbt$_{1-428}$ construct, in which the full-length Cbt protein was fused to the Gal4 DBD. This construct was cotransfected into S2 cells with a reporter construct containing Gal4-binding sites upstream of the firefly *luciferase* gene (the UAS-Luciferase vector). A significant repression of *luciferase* transcription (~6-fold) was observed in this experiment as compared to transfection with the Gal4 DBD alone (Figure 1A). This result indicates that the Cbt full-length protein is able to repress transcription. Interestingly, preliminary results
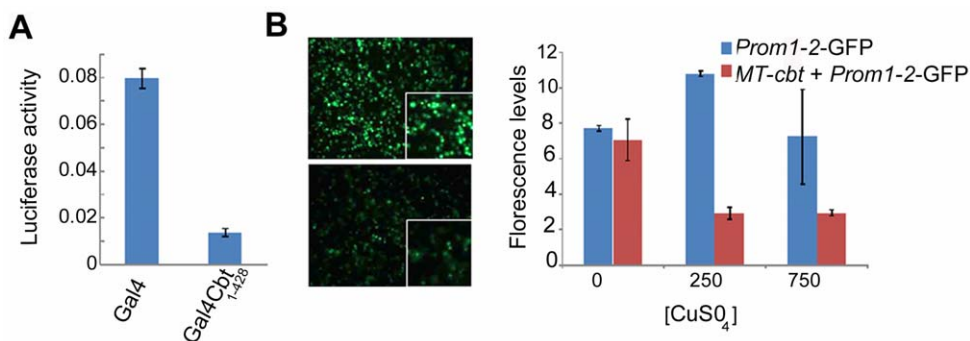

Figure 1. Cabut functions as a transcriptional repressor in S2 cells and regulates its own transcription. (A) Expression of Gal4Cbt$_{1-428}$ led to repression of luciferase activity levels relative to a control Gal4 protein. Luciferase activity was measured 48 h after transfection. pRENILLA was used to normalize for cell number, transfection efficiency, and general effects on transcription (luciferase activity = firefly luciferase/renilla luciferase). (B) S2 cells transfected with *Prom1-2*-GFP as a control or co-transfected with *Prom1-2*-GFP and MT-*cbt*. The MT promoter was induced by exposing the cells to medium containing copper (upper picture, no copper; lower picture, plus copper). Fluorescence levels were reduced following transcriptional induction of Cbt by copper in cells co-transfected with *Prom1-2*-GPF and MT-*cbt* (left panels). Fluorescence was measured 48 h after induction (right panel). pCHERRY was used for normalization (Fluorescence levels = GFP/CHERRY). In A and B, data are presented as the mean ± SD of three replicates.
doi:10.1371/journal.pone.0032004.g001

obtained in chromatin immunoprecipitation assays combined with genomic microarrays (ChIP-on-chip) in Canton-S embryos during dorsal closure (Y.B. and N.P. in collaboration with the modENCODE Project, unpublished results) suggested that Cbt was able to bind to a GC-rich genomic region 1 kb upstream of the *cbt* transcriptional start site (2 L: 479789,480740, Figure S1). Previous electrophoretic mobility shift assays (EMSAs) performed using the Cbt DBD region have demonstrated Cbt binding to GC-rich regions [49]. To confirm this result, we co-transfected *Drosophila* S2 cells with a construct in which the expression of a Cbt-V5 fusion protein was controlled by the methalotionein promoter (*MT*-Cbt) and a plasmid containing that region (named as Prom1-2 in [13]) fused to GFP (*Prom1-2*-GFP). To induce Cbt-V5 expression, the transfected cells were grown in Cu-supplemented medium. Indeed, induction of Cbt expression led to a significant reduction in GFP levels (Figure 1B). Taken together, these results show that Cbt acts as a transcriptional repressor, similar to the mammalian TIEG proteins, at least in S2 cells. Moreover, these experiments also demonstrate that the *cbt* gene autoregulates its own expression via a negative autoregulatory feedback mechanism, as has been shown for several genes involved

in circadian rhythms [50,51,52]. However, this mechanism has not been previously reported for vertebrate TIEG proteins.

Identification of transcriptional repressor domains in the Cbt protein

Previous studies showed that the repressor activity of the human TIEG1, TIEG2 and murine TIEG3 proteins in CHO-K1 and OLI-neu cells is mediated by a Sin3A-interacting domain (SID) (with an EAVEAL consensus sequence), which is required for its interaction with the Sin3A co-repressor and is located at the N-terminal region of the proteins within an α-helix motif [31,32]. Site-mutagenesis analyses revealed that the first alanine (A) residue and the α-helical structure are important for the recognition of this domain by the Sin3A co-repressor [53]. Because Cbt is able to repress transcription in S2 cells, we decided to determine whether the SID is conserved in the *Drosophila* protein. Multiple alignments of Cbt and several vertebrate TFs such as MAD and members of the Sp1 family (TIEG, BTEB), revealed that a similar domain, with equivalent residues, is present in Cbt but in a different region of the protein ($_{168}$AAEVAL$_{173}$ in Figure 2A). Secondary structure analysis of the Cbt protein using the SWISS-MODEL server [54]
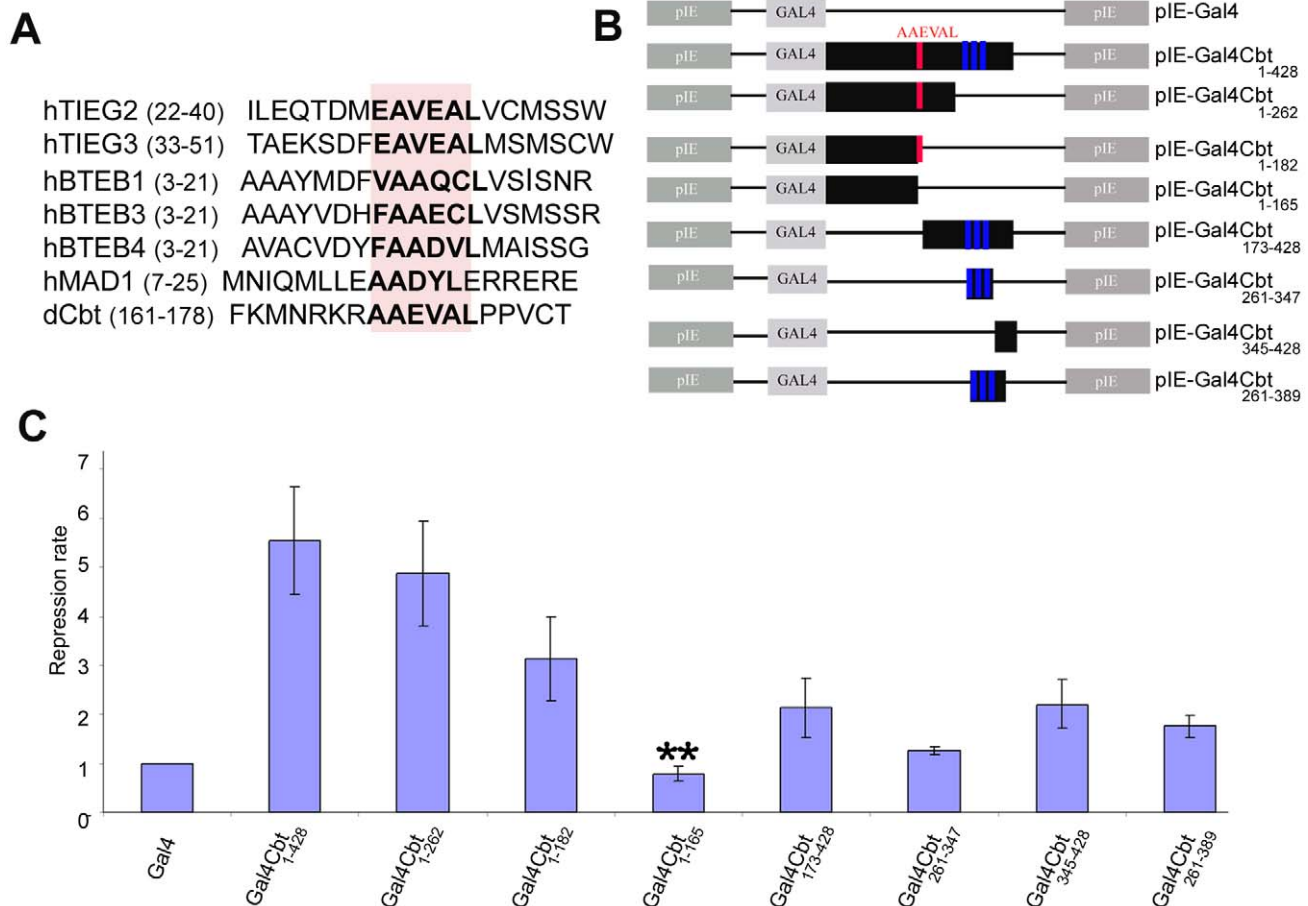
Figure 2. Identification of domains required for Cabut's transcriptional repressor activity. (A) Multiple alignment of the Sin3A interacting domains (SID) from *Drosophila* Cbt and several vertebrate TFs such as the MAD protein and mebers of the Sp1 family (TIEG, BTEB). Note that this domain is highly conserved in sequence (marked by a pink rectangle and residues in bold) but not with respect to location within the protein. (B) Schematic representation of the CbtGal4 fusion constructs transfected into S2 cells to identify transcriptional regulatory domains in the Cbt protein. The AAEVAL sequence is indicated in red and the C_2H_2 zinc fingers in blue. (C) Degree of *luciferase* repression obtained in UAS/Gal4 assay in S2 cells transiently transfected with the constructs shown in (B). Repression rate = luciferase activity of pIE-Gal4/luciferase activity of tested construct. pRENILLA was used for normalization as in Fig. 1. Data are presented as the mean ± SE (n≥5). Note that removal of the AAEVAL sequence completely abolishes the repressor activity of the Cbt protein (asterisks indicate p-value<0.01, t-Student's test).
doi:10.1371/journal.pone.0032004.g002
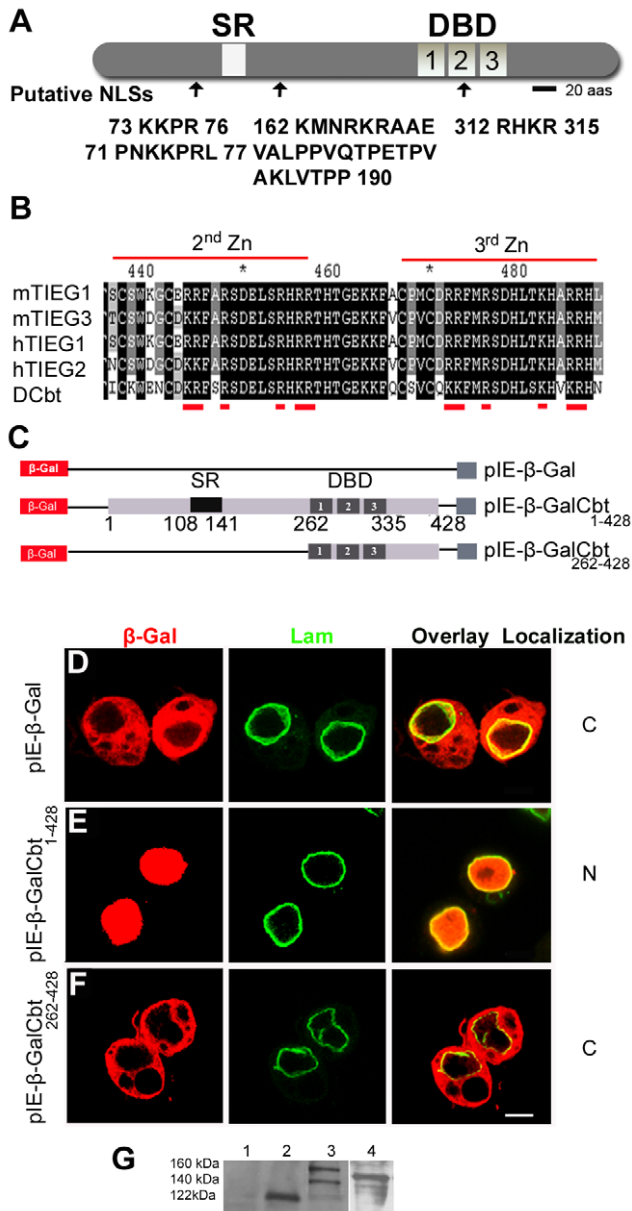
Figure 3. The Cabut zinc finger region is not essential for nuclear localization in S2 cells. (A) Schematic representation of the *Drosophila* Cbt protein, in which the locations of the Ser-rich (SR) region and the DBD are indicated. Sequences, coordinates and locations (arrows) of the predicted NLSs are also indicated. (B) Multiple alignment of the second and third zinc fingers of murine TIEG1 and TIEG3, human TIEG1 and TIEG2 and *Drosophila* Cbt. Red lines below the alignment indicate the amino acids included in the murine and human TIEG NLSs. (C) Schematic representation of the β-GalCbt fusion constructs used to transfect S2 cells. The locations of the SR region and the DBD are indicated by boxes. (D–F) Localization of β-Gal fusion proteins in S2 cells transiently transfected with the constructs shown in (C). Cells were stained with anti-β-Gal (red; first panel) and anti-Lam (green; second panel) to mark nuclear membranes. The overlay panel depicts double staining of cells with both antibodies, and the localization of the fusion proteins is shown in the fourth panel (C, cytoplasmic; N, nuclear). Wild-type β-Gal was located in the cytoplasm (D), but the β-GalCbt fusion protein translocated to the nucleus (E). However, the Cbt zinc finger region alone was not able to translocate β-Gal to the nucleus (F). Scale bar: 10 μm. (G) Western blot of protein extracts from S2 cells transfected with the constructs shown in (C) and stained with anti-β-Gal. (1) Non-transfected cells, (2) empty pIE-β-Gal vector (~122 kDa), (3) pIE-β-GalCbt$_{1-428}$ (~160 kDa) and (4) pIE-β-GalCbt$_{262-428}$ (~140 kDa). Cells transfected with the pIE-β-GalCbt$_{1-428}$

construct presented some degradation that did not affect the subcellular localization of the fusion protein.
doi:10.1371/journal.pone.0032004.g003

confirmed the conservation of an α-helix motif in the Cbt AAEVAL sequence (Figure S2). The presence of this domain could explain the transcriptional repressor activity of Cbt in S2 cells described above. Interestingly, this sequence is also conserved in all *Drosophila* species analyzed (Figure S2), suggesting that Cbt orthologs in these species may present a similar transcriptional activity. To determine whether this sequence is responsible for Cbt's repressor activity and to identify other possible transcriptional repression and/or activation domains in the protein, we generated a collection of Cbt Gal4 DBD fusion proteins (Figure 2B). Plasmids expressing these fusion proteins were cotransfected with the UAS-Luciferase vector (as described in Materials and Methods) into S2 cells, which endogenously express the dSin3A co-repressor [55]. Our results show that fusion proteins containing the AAEVAL sequence (pIE-Gal4Cbt$_{1-428}$, pIE-Gal4Cbt$_{1-262}$, pIE-Gal4Cbt$_{1-182}$) strongly repressed *luciferase* expression (up to 5-fold) in transfected cells (Figure 2C). This repressive effect is completely dependent on that sequence, as cotransfection with the pIE-Gal4Cbt$_{1-165}$ construct fails to repress the UAS-Luciferase reporter (Figure 2C). Thus, this result indicates that the AAEVAL sequence is essential for Cbt-mediated repression in S2 cells. In addition, we found that the C-terminal region of the Cbt protein, which does not contain the SID (pIE-Gal4$_{173-428}$), was able to reduce luciferase expression by ~1.5-fold (Figure 2C), thereby indicating that this region of the protein may contain one or more repression domains close to the DBD. Multiple alignments of the *Drosophila* Cbt orthologs revealed the presence of two regions in the Cbt C terminus that were highly conserved in all proteins analyzed (C4 and C5 in Figure S2). To confirm this possibility and test the transcriptional activity of both the DBD and the C4 and C5 conserved sequences, different constructs were generated in which truncated Cbt proteins containing only the Cbt DBD (pIE-Gal4Cbt$_{261-347}$), the Cbt C-terminal region including the C4 and C5 sequences (pIE-Gal4Cbt$_{345-428}$), and the Cbt DBD plus the C4 sequence (pIE-Gal4Cbt$_{261-389}$) were fused to the Gal4 DBD (Figure 2B) and cotransfected with the UAS-Luciferase vector into *Drosophila* S2 cells. Our results show that the DBD of Cbt by itself is not able to activate *luciferase* expression (Figure 2C). In addition, we found that the C-terminal region of Cbt without the zinc finger domains (compare pIE-Gal4 to Gal4Cbt$_{345-428}$) can significantly repress *luciferase* expression (~1.5-fold) (Figure 2C). Thus, our results show that the conserved sequence $_{379}$LRAIAPA$_{385}$, which we have called REP1/C4 (Figure 2D and Figure S2) and which is not present in the TIEG proteins, appears to have a mild repressive effect on *luciferase* expression (Figure 2C) in the UAS/Gal4 assays and could be in part responsible for the transcriptional repressor activity of the Cbt protein.

The Cys$_2$His$_2$-zinc finger domains of Drosophila Cabut are not essential for its nuclear localization in S2 cells but function as an NLS in mammalian CHO-K1 cells

Proteins larger than 45 kDa in size are unable to pass the nuclear membrane by passive diffusion [56]. Because we recently showed that Cbt is a nuclear protein with a theoretical molecular weight of 48 kDa [13], we assumed that it contains at least one functional NLS, although it could also be transported to the nucleus via an interaction with another NLS-containing protein. Indeed, four potential NLSs were identified in the Cbt protein using PSORT II [57] and cNLSMapper [58] programs (Figure 3A). Two of these

putative NLSs were $_{73}$KKPR$_{76}$ and $_{71}$PNKKPRL$_{77}$, which were predicted by both programs and located in the N-terminal region. The third sequence was a bipartite NLS, $_{162}$KMNRKRAAE-VALPPVQTPETPVAKLVTPP$_{190}$, which yielded the highest score in the cNLSMapper program. The fourth sequence was $_{312}$RHKR$_{315}$, which is located within the second Cbt zinc finger domain and predicted by both programs. A similar sequence (RHRR) was also found in the second zinc finger domain of the murine TIEG3 protein and is included within a functional bipartite NLS that is essential for TIEG3 nuclear import (Figure 3B) [19,32]. Because the zinc finger domains are highly conserved between Cbt and the TIEG proteins (Figure 3B and [14]), we decided to test the role of this putative NLS in Cbt nuclear localization. For doing so, we transiently transfected *Drosophila* S2 cells with the pIE-β-GalCbt$_{262-428}$ construct, in which the C-terminal region of Cbt containing the zinc finger region was fused to the *E. coli* β-Galactosidase (β-Gal, 116 kDa) cytoplasmic protein (Figure 3C, D). The pIE-β-GalCbt$_{1-428}$ construct, in which β-Gal is fused in frame to the full-length Cbt protein (Figure 3C), was used as a control. In these experiments, anti-Lamin (Lam) immunostaining was used to define the nuclear region. Double staining of the transfected cells with anti-β-Gal and anti-Lam antibodies showed that while the β-GalCbt$_{1-428}$ protein was exclusively localized in the nucleus (Figure 3E), β-GalCbt$_{262-428}$ was completely excluded from this cellular compartment, remaining in the cytoplasm (Figure 3F). Western blot analyses of transfected cell extracts were performed to confirm the integrity of the fusion proteins (Figure 3G). These assays showed that while the full-length Cbt protein appears to be partially degraded in S2 cells extracts (without affecting the NLS), the β-GalCbt$_{262-428}$ protein appears to be stable (Figure 3G). This result demonstrates that the putative NLS identified in the Cbt zinc finger region does not play a role in the protein's nuclear localization and suggests that the Cbt NLS is probably located at its N terminus. However, due to the high similarity between the Cbt and TIEG zinc finger domains, where essential basic amino acids are conserved (Figure 3B and [32]), we tested whether this region of the Cbt protein could function as an NLS in mammalian cells. Several expression constructs were generated: truncated Cbt proteins lacking the N-terminal region of Cbt (peGFPCbt$_{262-428}$), the complete zinc finger region (peGFPCbt$_{1-262}$), the third zinc finger (peGFFCbt$_{1-322}$), and both the second and third zinc fingers (peGFPCbt$_{1-292}$) were fused to eGFP (Figure 4A) and used to transiently transfect mammalian CHO-K1 cells. We found that the eGFP protein alone localizes to both the cytoplasm and the nucleus by passive diffusion due to its molecular weight (30 kDa) (Figure 4B). Cells transfected with the peGFPCbt$_{1-428}$ or peGFPCbt$_{262-428}$ construct showed exclusively nuclear eGFP localization (Figure 4C, D). However, cells transfected with deletion constructs affecting the zinc finger domains presented both cytoplasmic and nuclear eGFP signals (Figure 4E–G).Western blot analysis of cell extracts showed no degradation of the fusion proteins (Figure 4H). These results are in agreement with those obtained for the TIEG3 protein in HeLa and OLI-neu cells [32] and indicate that the Cbt bipartite NLS within the second and third zinc fingers is functional in mammalian cells, suggesting that different nuclear import mechanisms for this protein are being used in *Drosophila* and mammalian cells.

The PNKKPRL motif is necessary for Cbt nuclear localization

Our results indicate that a functional NLS is located within the N-terminal region of the Cbt protein. To test whether any of the predicted sequences in that region are required for Cbt nuclear localization (Figure 3A), we transiently transfected S2 cells with several deletion constructs. First, we generated pIE-βGalCbt$_{1-141}$
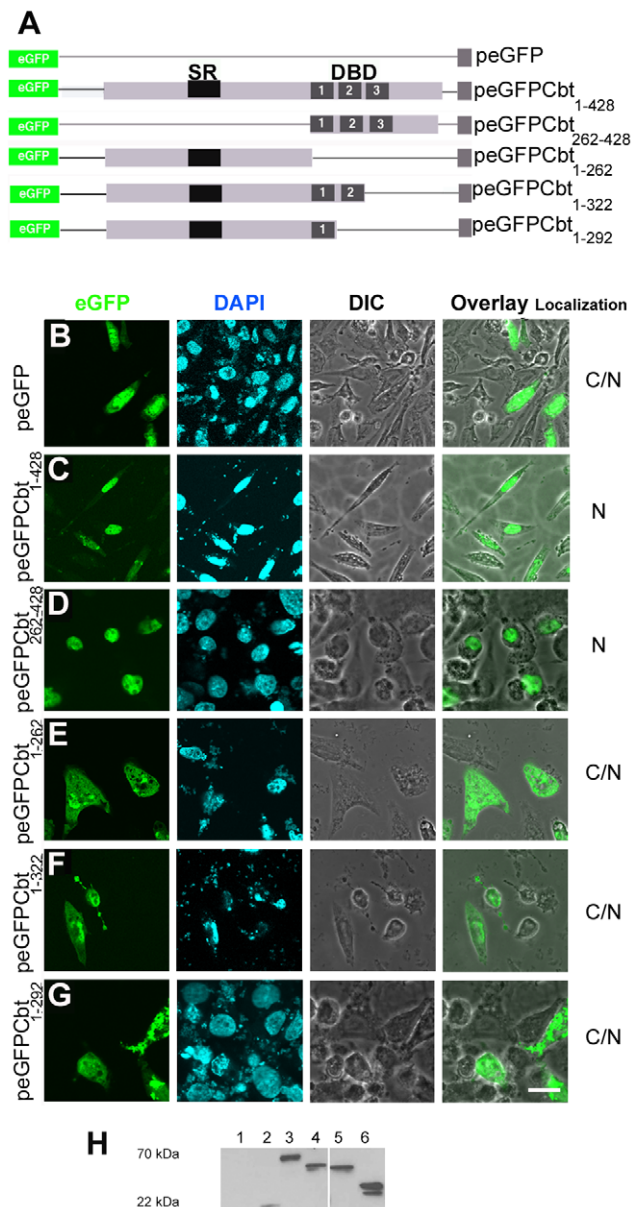
Figure 4. The Cabut zinc finger region is not essential for nuclear localization in CHO-K1 cells. (A) Schematic representation of the eGFPCbt fusion constructs used to transfect CHO-K1 cells. The locations of the SR region (black rectangle) and the zinc fingers (1, 2, 3, grey rectangles) are indicated. (B–G) Localization of eGFP fusion proteins in CHO-K1 cells transiently transfected with the constructs shown in (A). Cells were immunostained with anti-eGFP (green, first panel) and DAPI (blue; second panel). Differential interference contrast (DIC) was used to visualize cell boundaries (third panel). The overlay panel shows DIC and anti-eGFP staining. The localization of the fusion proteins is shown in the fourth panel (C, cytoplasmic; N, nuclear). Wild-type eGFP was located in the cytoplasm and the nucleus (B), but the eGFPCbt fusion protein translocates to the nucleus (C). With the exception of the peGFPCbt$_{262-428}$ fusion protein, which lacks the N-terminal region of the Cbt protein (G), all Cbt deletions affecting the zinc finger region showed cytoplasmic localization (D–F). Scale bar: 10 μm. (H) Western blots of protein extracts from CHO-K1 cells transfected with the constructs shown in (A) and stained with anti-GFP. (1) Non-transfected cells, (2) empty peGFP vector (~27 kDa), (3) peGFPCbt$_{1-428}$ (~70 kDa), (4) peGFPCbt$_{1-322}$ (~60 kDa), (5) peGFPCbt$_{1-292}$ (~60 kDa) and (6) peGFPCbt$_{1-262}$ (~50 kDa).
doi:10.1371/journal.pone.0032004.g004
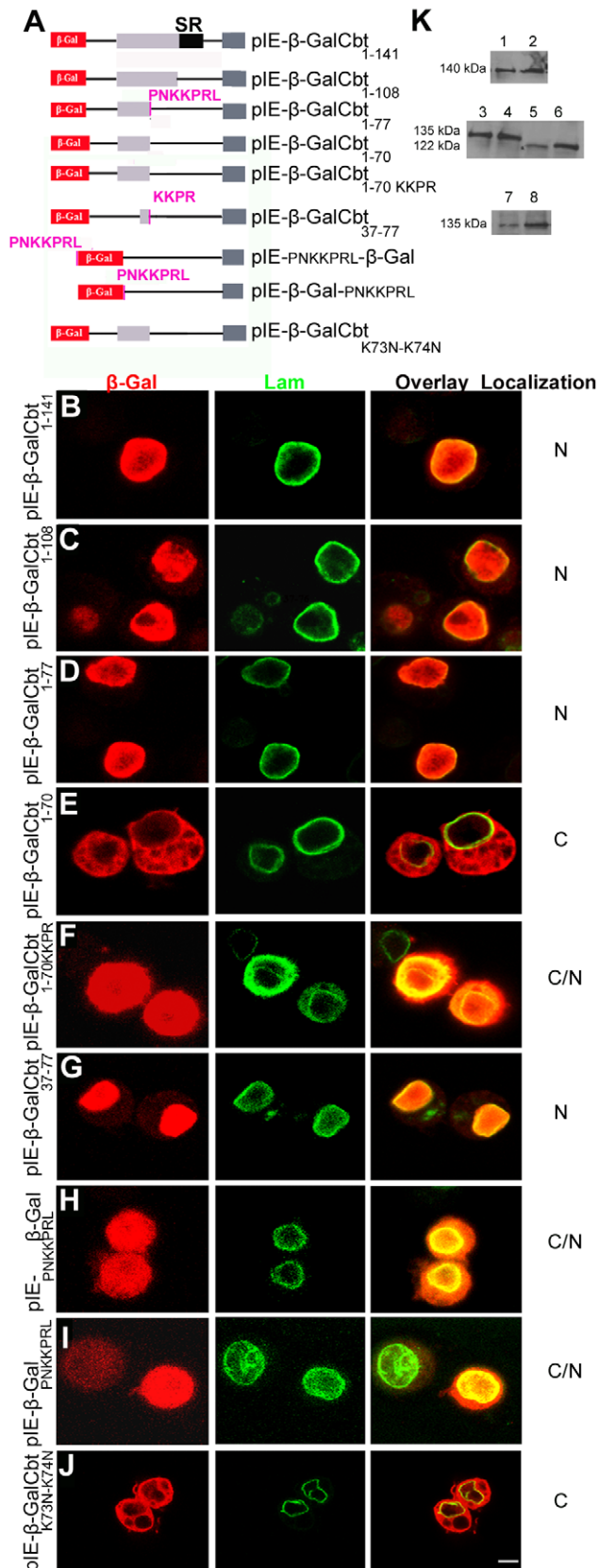
Figure 5. The $_{71}$PNKKPRL$_{77}$ sequence is the NLS of the Cabut protein. (A) Schematic representation of the β-GalCbt fusion constructs transfected into S2 cells to determine whether the $_{162}$KMNRKRAAEVALPPVQTPETPVAKLVTPP$_{190}$ and/or $_{71}$PNKKPRL$_{77}$ sequences are functional NLSs. The Serine-rich (SR) domain is shown in black and the PNKKPRL sequence in fuchsia. (B–J) Localization of β-Gal fusion proteins in S2 cells transiently transfected with the constructs shown in (A). Cells were stained with anti-β-Gal (red; first panel) and anti-Lam (green; second panel) to mark nuclear membranes. The overlay panel depicts double staining of cells with both antibodies, and the localization of the fusion proteins is shown in the fourth panel (C, cytoplasmic; N, nuclear). Note that fusion proteins lacking or containing a mutated $_{71}$PNKKPRL$_{77}$ sequence (E and J) were cytoplasmic. The $_{71}$PNKKPRL$_{77}$ sequence was able to translocate β-Gal to the nucleus when fused to either the N- or the C-terminus of the protein (H and I). Scale bar: 10 μm. (K) Western blots of protein extracts from S2 cells transfected with the constructs shown in (A) and stained with anti-β-Gal. (1) pIE-β-GalCbt$_{1–141}$ (~140 kDa), (2) pIE-β-GalCbt$_{1–108}$ (~140 kDa), (3) and (7) pIE-β-GalCbt$_{1–77}$ (~135 kDa), (4) pIE-β-GalCbt$_{1–70}$ (~135 kDa), (5) pIE-$_{PNKKPRL}$-β-Gal (~120 kDa), (6) pIE-β-Gal-$_{PNKKPRL}$ (~120 kDa) and (8) pIE-β-Gal-Cbt$_{K73N–K74N}$ (~135 kDa).
doi:10.1371/journal.pone.0032004.g005

and pIE-βGalCbt$_{1–108}$ constructs, encoding β-Gal fused to truncated Cbt proteins lacking either the bipartite NLS, $_{162}$KMNRKRAAEVALPPVQTPETPVAKLVTPP$_{190}$, or that sequence plus the SR domain (Figure 5A). Our results showed that both β-GalCbt fusion proteins were able to translocate to the nucleus (Figure 5B–C), thus indicating that neither the predicted bipartite NLS nor the SR region plays any role in Cbt nuclear localization and that an active NLS sequence is still retained in the truncated proteins. We next wanted to determine whether the overlapping $_{73}$KKPR$_{76}$ and $_{71}$PNKKPRL$_{77}$ sequences could act as functional NLSs. To do so, we generated constructs in which β-Gal was fused to Cbt N termini with (pIE-β-GalCbt$_{1–77}$) or without these sequences (pIE-β-GalCbt$_{1–70}$) (Figure 5A) and used them to transfect S2 cells. Immunostaining revealed that the removal of both sequences completely abolished the nuclear transport of the fusion proteins (Figure 5D, E), thus indicating that a functional NLS is present in this region. We next wanted to determine whether the KKPR residues, which are included within the PNKKPRL sequence and predicted as an NLS by the PSORT II program, were sufficient to target Cbt to the nucleus. Therefore, we generated the pIE-β-GalCbt$_{1–70KKPR}$ construct (Figure 5A), in which only the KKPR sequence is present. Immunostaining of transfected S2 cells showed that the fusion protein was localized in both the nucleus and the cytoplasm (Figure 5F). Because the integrity of the protein was confirmed by western blot analysis (Figure 5K), this observation indicates that the nuclear transport of β-GalCbt$_{1–70KKPR}$ is not efficient and suggests that the additional residues in the long NLS are necessary to increase the efficiency/rate of nuclear transport (compare Figure 5D to Figure 5F). We next wanted to determine whether the PNKKPRL sequence was sufficient to translocate a reporter protein to the nucleus. Therefore, we generated the pIE-β-Gal-$_{PNKKPRL}$ and pIE-$_{PNKKPRL}$-β-Gal constructs, in which the PNKKPRL sequence was fused in frame to either the C- or the N-terminal region of the β-Gal protein (Figure 5A). Immunostaining of transfected S2 cells revealed that the β-Gal protein was transported to the nucleus in both cases (Figure 5H–I), indicating that the PNKKPRL sequence is sufficient for nuclear import. Interestingly, this sequence is very similar to the SV40 large T antigen NLS (PKKKRKV) [59]. To determine which residues within the PNKKPRL sequence are important for Cbt nuclear transport, we performed site-directed mutagenesis of the basic lysine 73 and lysine 74 (K73 and K74) residues to asparagine (N) in the pIE-β-GalCbt$_{1–76}$ construct (designated pIE-β-GalCbt$_{K73N–K74N}$ in Figure 5A). Immunostaining of transfected S2 cells revealed that mutation of both K residues abolished Cbt nuclear transport (Figure 5J), indicating that they are essential for this process. Nuclear transport of the β-

Gal protein fused to the PNKKPRL sequence was not perfectly efficient, as some of the protein remained in the cytoplasm. It is therefore likely that either additional regions of the Cbt protein or other factors must be involved in its nuclear transport. Because no other NLSs in the Cbt protein were predicted by the utilized bioinformatic programs, it is possible that other sequences in the N-terminal region of the protein could be required together with the NLS for Cbt nuclear import. Multiple alignments of the Cbt *Drosophila* orthologs allowed us to identify three highly conserved regions in that part of the protein (C1, C2 and C3 in Figure S2). To determine whether these regions could act cooperatively with the PNKKPRL sequence in Cbt nuclear import, we generated a construct in which the β-Gal protein was fused to a Cbt fragment encompassing residues 37–77 (pIE- β-GalCbt$_{37-77}$) (Figure 5A), which lacked C1 and part of C2 and only contained C3 and the NLS. Immunostaining of transfected S2 cells revealed exclusive nuclear localization of the fusion protein, suggesting that the C3 region, but not the C1 or C2 region, may be required for Cbt nuclear translocation and could act cooperatively with the NLS (Figure 5G). It has been shown that post-translational modifications involving phosphorylation/dephosphorylation of NLSs and adjacent residues represent one of the mechanisms used to regulate nuclear import kinetics (reviewed in [60,61]). It is therefore possible that Cbt nuclear import could be regulated by desphosphorylation/phosphorylation events (see Discussion).

Taken together, these results demonstrate that the PNKKPRL sequence in the Cbt protein is a functional NLS motif in which the central K residues are essential for nuclear transport and the P, N, R and leucine (L) residues are probably required to increase efficiency. We also speculate that the C3 region located upstream of the NLS might regulate this process and thus confirm previous observations regarding the regulatory roles of sequences flanking classical NLSs.

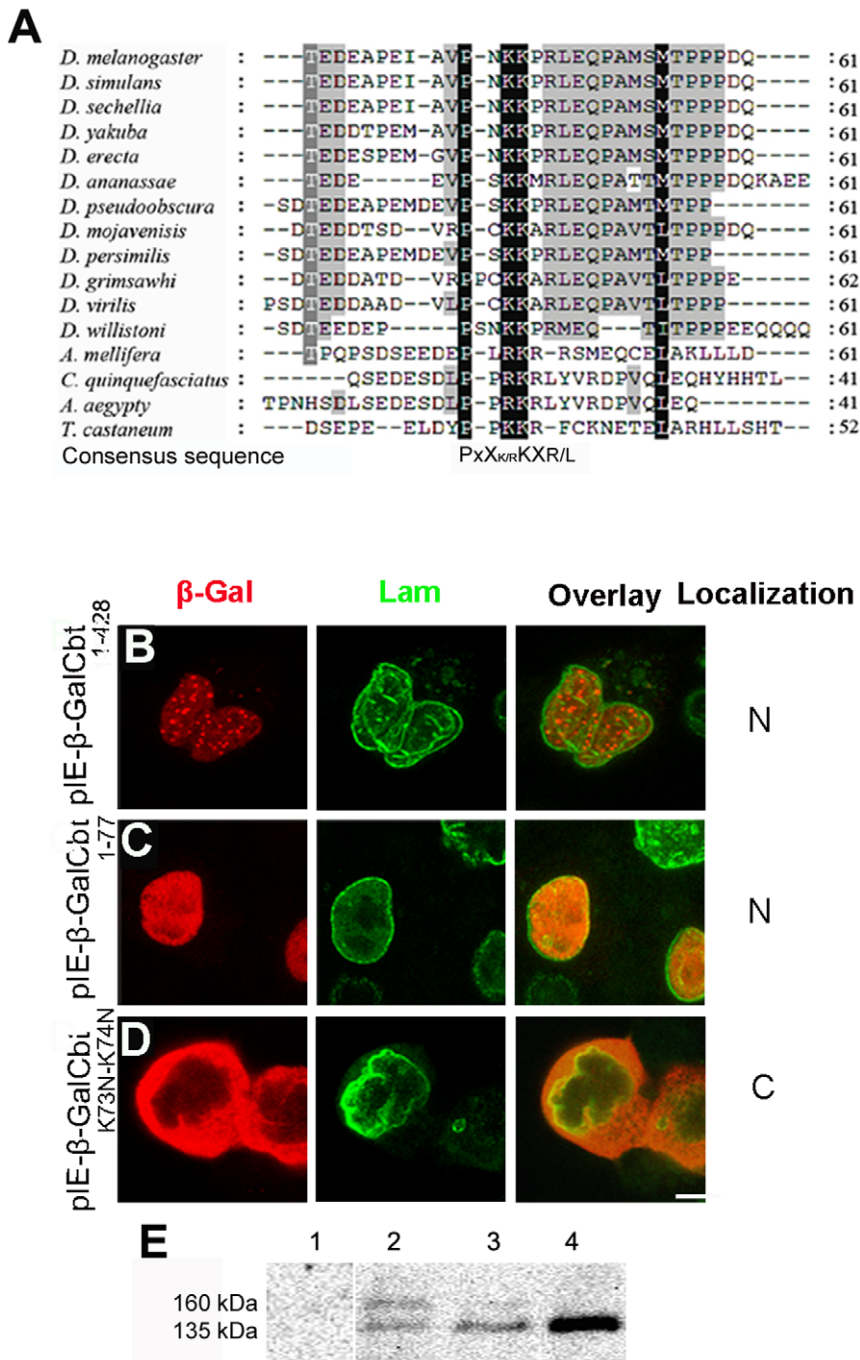## The PNKKPRL NLS is conserved and functional in Cbt insect orthologs

To determine whether the Cbt NLS and other residues important for its nuclear import are conserved in other *Drosophila* species and insects, we performed multiple alignments of the amino acid sequences of their Cbt orthologs. Our results showed that the PNKKPRL sequence was conserved in the twelve *Drosophila* species analyzed (Figure 6A and Figure S2). Consistent with this, we have recently demonstrated the nuclear localization of Cbt proteins in several *Drosophila* species [13]. Moreover, P and basic residues in the PNKKPRL sequence were also found in the Cbt proteins of several other insects, including *Apis mellifera*, *Culex quinquefasciatus*, *Aedes aegyti* and *Tribolium castaneum*, although they present a divergent N-terminal region (Figure 6A). Similar analyses in other Cbt orthologs of *Ciona intestinalis*, *Strongylocentrotus purpuratus*, *Daphnia pulex* and vertebrate TIEGs revealed that the PNKKPRL sequence is not present in these proteins (data not shown). Thus, the consensus NLS of insect Cbt orthologs may be PX(K/R)KX(R/L) (X = any residue). To test whether the Cbt NLS is functional in other insects, we used the pIE-β-GalCbt$_{1-428}$, pIE-β-GalCbt$_{1-77}$ and pIE-β-GalCbt$_{K73N-K74N}$ constructs (Figure 5A) to transiently transfect Sec301 cells from the beet armyworm *Spodoptera exigua*. Immunostaining revealed that both the full-length Cbt protein and a truncated form containing the PNKKPRL NLS but lacking the zinc finger domains and the predicted bipartite NLS were able to translocate the β-GalCbt fusion protein to the nucleus in *S. exigua* cells (Figure 6B–C). However, when the construct encoding β-Gal fused to a Cbt N-terminal region (1–77) with a mutated PNKKPRL sequence (pIE- β-GalCbt$_{K73N-K74N}$) was transfected, the fusion protein was

exclusively localized in the cytoplasm, as in *Drosophila* S2 cells (compare Figure 6D to Figure 5J). Western blot analyses confirmed that the proteins expressed in the transfected cells were of the correct size (Figure 6E). Taken together, these results indicate that the PX(K/R)KX(R/L) consensus motif is evolutionarily conserved and probably functional in insect Cbt orthologs. Its absence from TIEG proteins suggests that it evolved after the divergence between vertebrates and invertebrates.

Further searches in the GENPEPT protein database using the SCANSITE software [62] revealed a total of 57 proteins containing the PX(K/R)KX(R/L) sequence, 32 from invertebrates and 25 from vertebrates (data not shown). Among the identified invertebrate proteins, 53% were nuclear, including the *Drosophila* C$_2$H$_2$-zinc finger TFs Krüppel, Snail, Escargot and Scratch and the *A. gambiae* reverse transcriptase protein (Q868R2), suggesting that this sequence could be involved in the nuclear localization of other proteins, as predicted by the PSORT II program. Interestingly, the PX(K/R)KX(R/L) sequence is conserved and functional in the human BRCA1 (BReast CAncer Type 1) protein, a tumor suppressor protein involved in damaged DNA repair [63]. A $_{606}$PKKNRLRRKS$_{615}$ sequence that is similar to the Cbt NLS is involved in BRCA1 nuclear transport and interacts with the hSRP1α/importin-α2 protein [64].

Importin-α2 is involved in Cbt nuclear import in ovaries

Next, we aimed to identify the molecular mechanism that regulates Cbt nuclear import. The Cbt PNKKPRL motif presents features of classical monopartite NLSs, matching the K(K/R)X(K/R) consensus sequence required for importin-α/importin-β-based nuclear transport. Importin-α binds to NLS-bearing proteins and functions as an adapter to access the importin-β-dependent import pathway [65]. In *Drosophila*, three importin-α proteins have been identified (reviewed in [66]): importin-α1 (impα1, Kapα1 or CG8548), importin-α2 (impα2, Kapα2, Pen or CG4799) and importin-α3 (impα3, Kapα3 and CG9423) [67]. *In vitro* binding studies and nuclear import assays revealed that both NLSs and protein context mediate importin-α specificity for substrate nuclear import [68]. It has been shown that most tissues express all importin-α proteins, which probably perform redundant functions. Indeed, all three importin-α proteins are required for male and female germline development [69,70,71,72,73]. To identify which importin-α is involved in Cbt nuclear transport, we first searched in the BioGrid, DrosID and DpiM interaction databases for reported interactions between importin-α proteins and Cbt [74,75,76]. Because no known interactions were found, other experimental approaches were used to determine possible functional relationships between Cbt and each of these proteins. First, we performed genetic interaction assays using the rough eye phenotype caused by Cbt overexpression with the *sev*-Gal4 driver [1] (Figure 7B). Using this assay, we tested whether dosage reduction of any of the *impα* genes was able to dominantly modify that phenotype. Crosses of a recombinant *sev*-Gal4>UAS-Cbt$_{FL}$ line with *impα1*, *impα2* and *impα3* mutant alleles (*impα1*^{G4113}, *pen*^{EY0909} and *impα3*^{CG397}, respectively) were performed, and the progeny were scored for eye roughness modification. Although no modification was observed with the *impα1* or *impα3* mutant alleles (Figure 7C, E), there was a mild enhancement of the eye phenotype when *impα2* function was reduced (Figure 7D). We validated this interaction using an independent *impα2* allele (*impα2*^{D14}, Figure 7F). This result suggested that Cbt and Impα2 are functionally related. To confirm these results, we used UAS-*impα* RNAi lines to deplete Impα expression in larval salivary glands and then assessed whether Cbt nuclear localization was

Figure 6. The PNKKPRL sequence is conserved and functional in insect Cabut orthologs. (A) Multiple alignment of Cbt proteins from *Drosophila* species and other insects, showing that the PNKKPRL motif is highly conserved. A consensus sequence for the putative NLS is shown below (X = any amino acid). (B–D) Localization of β-Gal fusion proteins in *Spodoptera exigua* Sec301 cells transiently transfected with the constructs shown in (A). Cells were stained with anti-β-Gal (red; first panel) and anti-Lam (green; second panel) to mark nuclear membranes. The overlay panel depicts double staining of cells with both antibodies, and the localization of the fusion proteins is shown in the fourth panel (C, cytoplasmic; N, nuclear). Note that all β-Gal fusion proteins containing the wild-type PNKKPRL sequence were translocated to the nucleus. Scale bar: 10 μm. (E) Western blots of protein extracts from Sec301 cells transfected with constructs shown in (A) and stained with anti-β-Gal. (1) Non-transfected Sec301 cells, (2) pIE-β-Gal-Cbt$_{1-428}$ (~160 kDa), (3) pIE-β-Gal-Cbt$_{1-77}$ (~135 kDa) and (4) pIE-β-Gal-Cbt$_{1-K73N-K74N}$ (~135 kDa).
doi:10.1371/journal.pone.0032004.g006

disrupted. A similar approach was previously applied to demonstrate that the Naked cuticle (Nkd) protein requires Impα3 for nuclear localization [77]. Immunostaining of salivary glands from *71B>impα1^RNAi^*, *71B>impα2^RNAi^* and *71B>impα3^RNAi^* larvae revealed no changes in Cbt nuclear localization (Figure S3A–D).

However, only the *impα3^RNAi^* line has been shown to reduce Impα3 immunoreactivity [77].

Previous work has shown that *impα2* is expressed during embryonic development in the male and female germline [69,70,78], the nervous system [79] and larval muscles [44] and
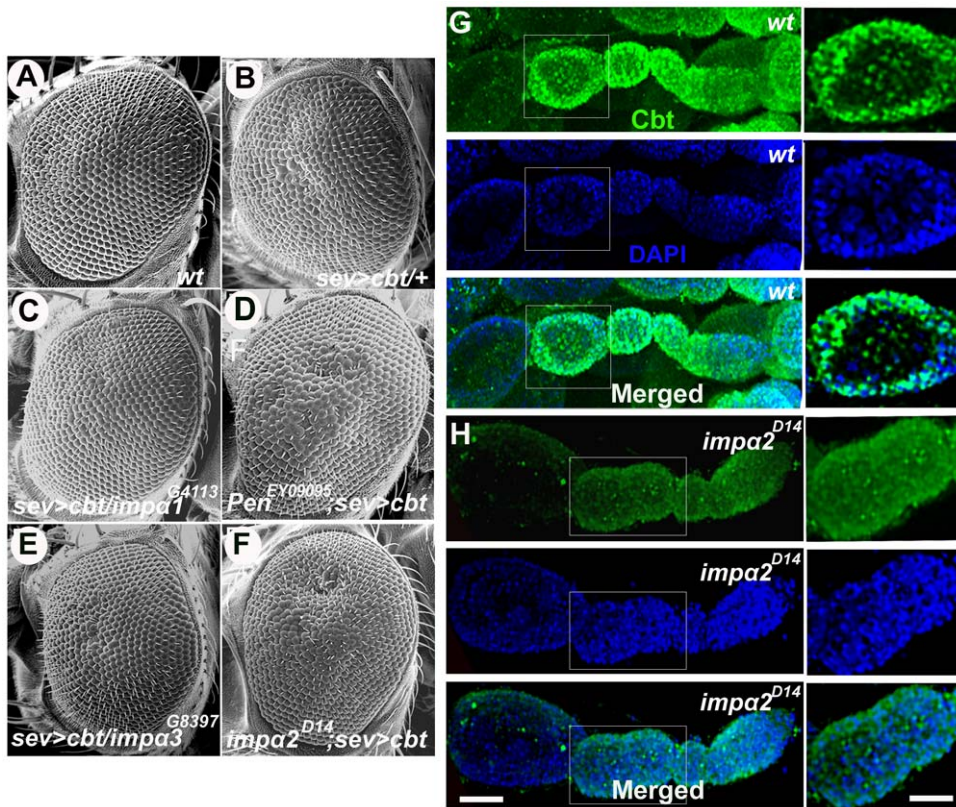
Figure 7. Importin-α2 is required for Cabut nuclear import in the *Drosophila* ovary. (A–F) Scanning electron micrographs of adult eyes of the indicated genotypes. Note that eye roughness in Cbt-overexpressing flies was dominantly enhanced when the *impα2* dosage was reduced (compare B to D and F), but was not modified in the presence of *impα1* or *impα3* mutant alleles. (G–H) Representative confocal images of ovarian follicle cells stained with an anti-Cbt antibody (green) and DAPI (blue). High-magnification images of ovarioles are shown on the right. Note that whereas the Cbt protein is detected in follicle cell nuclei (G) from wild-type flies, its nuclear localization is reduced in *impα2^{D14}* mutant ovaries. Scale bars: 10 μm and 8 μm for regular and high-magnification images, respectively.
doi:10.1371/journal.pone.0032004.g007

may be involved in cell proliferation and cell cycle progression [80,81]. However, it is not required for general development or cell viability because *impα2* mutants survive to adulthood, possibly due to functional redundancy with other importins. Therefore, to determine whether Impα2 is required for nuclear Cbt import, we analyzed Cbt expression in the larval central nervous system (CNS) and in ovaries of wild-type and *impα2^{D14}* mutants. Previously, we and others described that Cbt was a maternal factor that was expressed in the CNS during embryonic development [8,13]. Our results show that Cbt nuclear localization is not altered in the larval brain hemispheres of *impα2* mutants (Figure S3E–F). In the ovary, Cbt is normally localized to the nuclei of follicular cells in the germarium. We observed that the nuclear localization of Cbt in these cells was significantly reduced in *impα2* mutants (Figure 7G–H). These results indicate that Impα2 may be involved in Cbt nuclear import in ovaries but not the larval CNS and suggest that no functional redundancy exists between Impα2 and the other importins during oogenesis, as previously proposed [70]. However, our results can not exclude the possibility that other Impα proteins may be required for Cbt nuclear localization in other tissues. This is supported by the fact that Cbt has a complex expression pattern during *Drosophila* embryogenesis [13].

Discussion

Cbt is an evolutionarily conserved C₂H₂ zinc finger TF involved in the regulation of different developmental processes in *Drosophila*

[1,2,3,4,5,6,7,8,9,10,12,13]. Indeed, it is the *Drosophila* ortholog of the vertebrate TIEG proteins, which belong to the Sp1-KLF family of TFs) [14,16]. However, little is known about the molecular mechanisms that regulate Cbt function. To fully characterize the function of a TF, it is important to identify at least three domains in its sequence: the DBD, the NLS and the transcriptional regulatory domain(s). The aims of the present study were to determine its role in transcriptional regulation as well as characterize in detail the molecular mechanisms of Cbt nuclear import by identifying the relevant functional domains. Our experiments were also designed to test the functional conservation between Cbt and its vertebrate orthologs, based on previous results obtained from functional studies of the TIEG proteins. Although it is generally accepted that evolutionarily conserved sequences will perform the same molecular function, this is not always true, and evidence for functional conservation must come from functional studies and not from sequence similarity analyses.

Transcriptional repressor/activator activity of Cbt

In the present study we have demonstrated for the first time that Cbt acts as a transcriptional repressor in *Drosophila* S2 cells. However, previous studies suggested that it might function as a transcriptional activator. We first showed that *dpp* expression was downregulated at the leading edge of the lateral epidermal sheets in *cbt* mutant embryos during dorsal closure [1]. Furthermore, it has been recently shown that Cbt acts as positive regulator of TGF-β signaling during wing imaginal disc development [12].

This latter study suggested that the function of Cbt as a transcriptional activator was consistent with the fact that repressor domains identified in the TIEG proteins (R1, R2 and R3 domains) were not conserved in Cbt [12]. However, our results clearly show that Cbt acts as a transcriptional repressor in S2 cells and that this activity is mediated in part by a conserved SID located in its N-terminal region. This domain was found within the R1 region of the TIEG proteins and was shown to be essential for human and murine TIEGs-mediated transcriptional repression in cell culture [31,32]. It is interesting to note that the putative SID in the Cbt protein is also located in an α-helix but in a different part of the protein. In addition, we found that the REP1 sequence located in the C-terminal region of Cbt could also account for its transcriptional repressor activity. Supporting this observation, the REP1 sequence contains charged, hydrophobic A and P residues, as has been shown for other transcriptional repressor domains [82,83]. However, no information about similar sequences in TIEGs or other repressor proteins has been found in the literature. Interestingly, the C-terminal region of murine TIEG3 protein presents transcriptional activator activity, although the domain(s) responsible of that function has not been characterized yet [32]. Transcriptional repression is crucial for the regulation of gene expression and morphogenesis. Hairy-related proteins play critical roles during development by repressing target genes at multiple stages of neurogenesis [84]. Similarly, early patterning of the *Drosophila* embryo requires multiple genes encoding transcriptional repressor proteins [83,84,85].

TIEG proteins were originally described as transcriptional repressors, but several studies have demonstrated that these proteins and other Sp1/KLF family members can be repressors or activators depending on the cellular and binding site context [31,32,33,34,35]. Indeed, it has been shown that phosphorylation of S/T residues adjacent to the SID may disrupt the mSin3A interaction, thus inhibiting TIEG2 repressor activity [86]. Besides, the KLF13 protein presents a SID overlapping with an activation domain, and its activator/repressor activity depends on the acetylation state of its DBD or its target promoters [87]. Although we do not know whether this double function is conserved in the *Drosophila* Cbt protein, it is also interesting to consider that the SR region is conserved in most of the *Drosophila* Cbt orthologs (Figure S2). SR domains have been shown to be involved in transactivation, e.g., in the v-Rel protein [88]. Maybe phosphorylations in S/T residues within the SR domain could be involved in the regulation of the transcriptional activity of Cbt. Moreover, it is important to note that despite the high similarity in the zinc finger domain between Cbt and the TIEG proteins, our results show that this region is not able to activate transcription as it does in the murine TIEG3 protein [31,32]. Further experiments are necessary to determine whether Cbt can act as a transcriptional activator in *Drosophila*.

Cbt regulates its own expression: negative feedback?

We also show that Cbt is able to recognize its own promoter and negatively regulate its own expression in S2 cells. This report is the first of a direct Cbt target. These data are supported by ChIP-on-chip assay results in which Cbt was found to bind to its promoter region (Figure S1). Negative feedback loops are used to regulate the levels of signaling molecules and contribute to signal homeostasis. In many cases, the molecular component that executes the feedback-mediated inhibition is transcriptionally targeted by the pathway that it regulates. This mechanism ensures an interdependence of signaling activity and feedback regulation and is often viewed as an inherent means of downregulating signaling pathways during development [89,90]. One interesting

observation that is consistent with the existence of Cbt negative feedback is that *cbt* overexpression in different tissues during embryonic dorsal closure, where it acts downstream of the JNK pathway, does not cause embryonic lethality [1] (Y.B and N.P, unpublished results), as occurs when other components of that pathway are overexpressed [91]. Cbt likely acts to negatively regulate its own expression and modulate JNK signaling levels. This negative feedback is also consistent with a role for Cbt in regulating circadian rhythms [5], as most transcriptional circadian regulators have a strong transcriptional effect (often direct) on their own synthesis in both mammals and *Drosophila* [92,93,94]. Interestingly, several TFs of the KLF family can regulate their own expression as well as the expression of other family members. KLF4, for example, can activate its own expression in the intestinal epithelium, while KLF5 represses KLF4 expression through competitive interaction with the same cis-element [95]. Currently, no evidence of such an autoregulatory mechanism has been demonstrated in vertebrate TIEG proteins. Several studies are being performed to identify other direct Cbt target genes and to further analyze how they are transcriptionally regulated by this protein.

Molecular mechanism of Cbt nuclear import

NLSs of proteins belonging to the KLF family of TFs are diverse in different organisms although most are either within or nearby their DBDs [32,38,39,96]. Indeed, the murine TIEG3 protein contains a bipartite NLS within the zinc finger region conserved in other TIEG proteins in mice and humans [31,32]. In this work, we demonstrate that Cbt nuclear import is mediated by a monopartite NLS (PNKKPRL) located within the N-terminal region of the protein. This NLS is not conserved in vertebrate KLF family members but is present in insect Cbt orthologs as well as in other unrelated proteins from invertebrates and vertebrates. We also demonstrate that this NLS is functional in *S. exigua* Sec301 cells, suggesting that it is probably functional in all insect Cbt orthologs. Interestingly, Cbt also contains a second NLS, which resembles that described in the TIEG3 protein [31,32]. Importantly, experiments investigating Cbt localization in hamster CHO-K1, *Drosophila* S2 and *S. exigua* Sec301 cells confirmed that this second NLS is functional in mammalian but not insect cells. This finding clearly demonstrates that protein sequence conservation among different species does not always indicate functional conservation and indicates that additional factors such as cellular context are also important and must be considered when ascribing molecular functions to certain sequences. A similar situation has been found in the *Aspergillus nidulans* HapB protein, a subunit of CCAAT-binding factor [97]. A likely explanation for the results obtained in our studies is that in insects, the sequence at the N terminus of the Cbt protein is recognized as an NLS by the importin-α/importin-β-based nuclear transport machinery, whereas that contained in the second and third finger domains is not [65].

Experiments using β-GalCbt fusion proteins also demonstrated the relevance of the central K residues (K73 and K74) in Cbt nuclear import as well as the existence of putatively critical residues, such as the flanking P, N, P, R and L residues and maybe residues within the C3 region (Figures 5 and S2). Nuclear transport can be regulated at multiple levels, via a diverse range of mechanisms that include (1) accessibility or masking of the NLS and availability of import factors; (2) existence of cytoplasmic or nucleoplasmic retention signals; (3) regulation of the NLS affinity for its import receptor, e.g., by phosphorylation; (4) regulation of nuclear pore complex permeability; and (5) possible regulation of cargo affinity to the hydrophobic central channel. Among these mechanisms, post-translational modification of proteins through

phosphorylation/dephosphorylation is the best understood mechanism regulating nuclear transport (reviewed in [42]). To analyze this possibility, we used the NetPhos software [98] to perform *in silico* predictions of putative S, T and tyrosine (Y) phosphorylation sites around the Cbt NLS and to identify the kinases putatively responsible for the predicted phosphorylation events. These analyses showed that although the $_{71}$PNKKPRL$_{77}$ sequence is not probably modified, the conserved S59 and T61 residues could be phosphorylated by Casein Kinase II (CKII) and the T85 amino acid could be targeted by the p38 Mitogen-activated protein kinase (p38 MAPK) (Figures S2 and S4). Because the β-GalCbt$_{1-76}$ fusion protein, which lacks S83 and T85, can be translocated to the nucleus (Figure 5D), the S59 and T61 residues are better candidates for phosphorylation sites that affect Cbt nuclear transport. Interestingly, we found that the S59 and T61 residues near the PNKKPRL sequence were conserved in the twelve *Drosophila* species analyzed (Figure S2). This finding could explain why the β-Gal protein is not efficiently transported to the nucleus when fused to the PNKKPRL sequence (Figure 5H–I). Similar results have been reported for proteins fused to the SV40 NLS and in the KLF8 protein [99,100]. Phosphorylation of residues flanking the NLS can affect nuclear trafficking in different ways, e.g., enhancing the binding affinity of importins to cargo or enhancing the docking of cargo to the nuclear pore complex as well as causing conformational changes that expose the NLS to the protein surface (reviewed in [61]). Additional experiments will be necessary to confirm that Cbt nuclear translocation is regulated by the phosphorylation/desphosphorylation of these residues.

Regarding Cbt, previous studies revealed that this protein is expressed as a Pumilio target in central and peripheral nervous system axons [13] and probably synapses [101] of the *Drosophila* embryo. Although we have not yet identified a nuclear export signal (NES) in the Cbt sequence (data not shown), these findings suggest that Cbt nuclear import might be tightly regulated either by post-translational modifications (such as phosphorylation of S59 and T61 residues) or conformational changes that prevent NLS recognition by importins. Cbt has also been reported to be involved in circadian rhythms [5], a process in which the control of nuclear trafficking has been demonstrated [102]. Indeed, several clock proteins contain NLSs that facilitate their cellular trafficking [20,103]. This control of trafficking is key for the generation and maintenance of robust and coherent circadian rhythms. Although the mechanisms of nuclear transport regulation mentioned above have been previously described for other proteins (reviewed in [104]), further analyses will be required to confirm whether they also regulate Cbt nuclear trafficking

Finally, genetic interaction assays and Cbt immunostaining in *impα* mutants suggest that the Impα2 protein may interact with Cbt, probably recognizing the PNKKPRL sequence, and seems to be required to transport Cbt to the nucleus in several tissues, including the *Drosophila* ovary, where a reduction of Cbt nuclear localization was observed in *impα2* mutant flies. In support of this hypothesis, the human BRCA protein contains an NLS similar to the one detected in Cbt; this NLS is recognized by the Impα2 protein and is responsible for the protein's nuclear localization [64]. These data suggest that the nuclear transport mechanism mediated by the PX(K/R)KX(R/L) sequence may be conserved between vertebrates and invertebrates. *Drosophila* Impα2 was recently shown to be involved in Frizzled regulation in muscle and in the central and peripheral nervous systems of embryos and larvae [44,79]. However, we do not exclude the possibility that other Impα proteins (α1 or α3) may be involved in Cbt nuclear transport in other tissues because Cbt presents a ubiquitous

expression pattern at embryonic and larval stages [13]. Co-immunoprecipitation assays will be necessary to determine whether other importin proteins interact with Cbt in different contexts.

Supporting Information

Figure S1 Cabut binds to its own promoter region. (A) Integrate Genome Browser (IGB) [105] overview of the *cabut* genomic region on chromosome 2L identified in the ChIP-on-chip analysis. From top to bottom: signal represents the $\log_2$ of normalized ratios of IP/input, p-value is on a $-10\log_{10}$ scale (based on Wilcoxon test), coordinates of the genomic fragment and structure of the *cabut* gene. (B) Nucleotide sequence of the *cabut* Prom1-2 region. The GC-rich sequences that might be recognized by the Cbt protein are boxed.
(TIF)

Figure S2 Multiple alignment of Cabut orthologs from twelve *Drosophila* species, showing conserved domains and secondary structure. C1, C2, C3, C4/Rep1 and C5 boxes contain conserved regions in the N and C termini of *Drosophila* Cbt proteins; the NLS box indicates the position of the sequence required for Cbt nuclear localization in *D. melanogaster*. The AAEVAL and Rep1 boxes mark the transcriptional repressor motifs identified in this work. The positions of the serine-rich (SR) region and the three zinc fingers (Zn1, Zn2 and Zn3) are also indicated. Black diamonds mark the position of S and T residues that are potentially phosphorylatable (according to NetPhos program) and putatively required for regulation of Cbt nuclear import. Red arrows and yellow helixes indicate the positions of predicted β-sheet motifs and α-helices, respectively. The secondary structure topology was obtained using the SWISS-MODEL program.
(TIF)

Figure S3 Immunohistochemical detection of the Cabut protein in *importin-α* mutants. (A–C) Immunostaining of salivary glands from (A) wild-type, (B) *71B>impα3^{RNAi}*, (C) *71B>impα1^{RNAi}*, and (D) *71B>impα2^{RNAi}* third instar larvae with anti-Cbt (red) and anti-Lamin (green) antibodies. Note that Cbt was still detected in the nuclei of salivary gland cells (arrow) and fat cells (arrowhead) in all genotypes. Scale bar: 15 µm. (E–F) Immunostaining of brain hemispheres from (E) wild-type and (F) *impα2^{D14}* mutant larvae with an anti-Cbt (green) antibody. Note that Cbt nuclear localization was not reduced in brains of *impoα2* mutants. Scale bar: 10 µm.
(TIF)

Figure S4 *in silico* prediction of putative phosphorylatable residues in the Cabut sequence. The positions of S and T residues predicted to be susceptible to phosphorylation by different kinases (indicated in pink: CK II and p38-MAPK) and adjacent to the PNKKPRL sequence (whose position is marked by a red arrow) are shown. Putative phosphorylatable residues and responsible kinases were determined using the NetPhos and NetPhosK programs, respectively.
(TIF)

Acknowledgments

Center and the Vienna Drosophila RNAi Center for fly stocks and to ENCODE project members at UCSC for the ChIP-on-chip assay

Author Contributions

References

1. Munoz-Descalzo S, Terol J, Paricio N (2005) Cabut, a C2H2 zinc finger transcription factor, is required during Drosophila dorsal closure downstream of JNK signaling. Dev Biol 287: 168–179.

2. Beckstead RB, Lam G, Thummel CS (2005) The genomic response to 20-hydroxyecdysone at the onset of Drosophila metamorphosis. Genome Biol 6: R99.

3. Zhao T, Gu T, Rice HC, McAdams KL, Roark KM, et al. (2008) A Drosophila gain-of-function screen for candidate genes involved in steroid-dependent neuroendocrine cell remodeling. Genetics 178: 883–901.

4. Blanco E, Ruiz-Romero M, Beltran S, Bosch M, Punset A, et al. (2010) Gene expression following induction of regeneration in Drosophila wing imaginal discs. Expression profile of regenerating wing discs. BMC Dev Biol 10: 94.

5. Kadener S, Stoleru D, McDonald M, Nawathean P, Rosbash M (2007) Clockwork Orange is a transcriptional repressor and a new Drosophila circadian pacemaker component. Genes Dev 21: 1675–1686.

6. Kraut R, Menon K, Zinn K (2001) A gain-of-function screen for genes controlling motor axon guidance and synaptogenesis in Drosophila. Curr Biol 11: 417–430.

7. Mindorff EN, O'Keefe DD, Labbe A, Yang JP, Ou Y, et al. (2007) A gain-of-function screen for genes that influence axon guidance identifies the NF-kappaB protein dorsal and reveals a requirement for the kinase Pelle in Drosophila photoreceptor axon targeting. Genetics 176: 2247–2263.

8. Yatsu J, Hayashi M, Mukai M, Arita K, Shigenobu S, et al. (2008) Maternal RNAs encoding transcription factors for germline-specific gene expression in Drosophila embryos. Int J Dev Biol 52: 913–923.

9. Bulow MH, Aebersold R, Pankratz MJ, Junger MA (2010) The Drosophila FoxA ortholog Fork head regulates growth and gene expression downstream of Target of rapamycin. PLoS One 5: e15171.

10. Guertin DA, Guntur KV, Bell GW, Thoreen CC, Sabatini DM (2006) Functional genomics identifies TOR-regulated genes that control growth and division. Curr Biol 16: 958–970.

11. Gorski SM, Chittaranjan S, Pleasance ED, Freeman JD, Anderson CL, et al. (2003) A SAGE approach to discovery of genes involved in autophagic cell death. Curr Biol 13: 358–363.

12. Rodriguez I (2011) Drosophila TIEG is a modulator of different signalling pathways involved in wing patterning and cell proliferation. PLoS One 6: e18418.

13. Belacortu Y, Weiss R, Kadener S, Paricio N (2011) Expression of Drosophila Cabut during early embryogenesis, dorsal closure and nervous system development. Gene Expr Patterns 11: 190–201.

14. Munoz-Descalzo S, Belacortu Y, Paricio N (2007) Identification and analysis of cabut orthologs in invertebrates and vertebrates. Dev Genes Evol 217: 289–298.

15. Seetharam A, Bai Y, Stuart GW (2010) A survey of well conserved families of C2H2 zinc-finger genes in Daphnia. BMC Genomics 11: 276.

16. Suske G, Bruford E, Philipsen S (2005) Mammalian SP/KLF transcription factors: bring in the family. Genomics 85: 551–556.

17. Cook T, Gebelein B, Mesa K, Mladek A, Urrutia R (1998) Molecular cloning and characterization of TIEG2 reveals a new subfamily of transforming growth factor-beta-inducible Sp1-like zinc finger-encoding genes involved in the regulation of cell growth. J Biol Chem 273: 25929–25936.

18. Kaczynski J, Cook T, Urrutia R (2003) Sp1- and Kruppel-like transcription factors. Genome Biol 4: 206.

19. Wang Z, Peters B, Klussmann S, Bender H, Herb A, et al. (2004) Gene structure and evolution of Tieg3, a new member of the Tieg family of proteins. Gene 325: 25–34.

20. Spittau B, Krieglstein K (2011) Klf10 and Klf11 as mediators of TGF-beta superfamily signaling. Cell Tissue Res;DOI 10.1007/s00441-011-1186-6.

21. Subramaniam M, Harris SA, Oursler MJ, Rasmussen K, Riggs BL, et al. (1995) Identification of a novel TGF-beta-regulated gene encoding a putative zinc finger protein in human osteoblasts. Nucleic Acids Res 23: 4907–4912.

22. Chen Z, Lei T, Chen X, Zhang J, Yu A, et al. (2010) Porcine KLF gene family: Structure, mapping, and phylogenetic analysis. Genomics 95: 111–119.

23. Cook T, Urrutia R (2000) TIEG proteins join the Smads as TGF-beta-regulated transcription factors that control pancreatic cell growth. Am J Physiol Gastrointest Liver Physiol 278: G513–521.

24. Subramaniam M, Gorny G, Johnsen SA, Monroe DG, Evans GL, et al. (2005) TIEG1 null mouse-derived osteoblasts are defective in mineralization and in support of osteoclast differentiation in vitro. Mol Cell Biol 25: 1191–1199.

25. Bender H, Wang Z, Schuster N, Krieglstein K (2004) TIEG1 facilitates transforming growth factor-beta-mediated apoptosis in the oligodendroglial cell line OLI-neu. J Neurosci Res 75: 344–352.

26. Chalaux E, Lopez-Rovira T, Rosa JL, Pons G, Boxer LM, et al. (1999) A zinc-finger transcription factor induced by TGF-beta promotes apoptotic cell death in epithelial Mv1Lu cells. FEBS Lett 457: 478–482.

27. Gohla G, Krieglstein K, Spittau B (2008) Tieg3/Klf11 induces apoptosis in OLI-neu cells and enhances the TGF-beta signaling pathway by transcriptional repression of Smad7. J Cell Biochem 104: 850–861.

28. Hirota T, Kon N, Itagaki T, Hoshina N, Okano T, et al. (2010) Transcriptional repressor TIEG1 regulates Bmal1 gene through GC box and controls circadian clockwork. Genes Cells 15: 111–121.

29. Subramaniam M, Hawse JR, Johnsen SA, Spelsberg TC (2007) Role of TIEG1 in biological processes and disease states. J Cell Biochem 102: 539–548.

30. Tachibana I, Imoto M, Adjei PN, Gores GJ, Subramaniam M, et al. (1997) Overexpression of the TGFbeta-regulated zinc finger encoding gene, TIEG, induces apoptosis in pancreatic epithelial cells. J Clin Invest 99: 2365–2374.

31. Cook T, Gebelein B, Belal M, Mesa K, Urrutia R (1999) Three conserved transcriptional repressor domains are a defining feature of the TIEG subfamily of Sp1-like zinc finger proteins. J Biol Chem 274: 29500–29504.

32. Spittau B, Wang Z, Boinska D, Krieglstein K (2007) Functional domains of the TGF-beta-inducible transcription factor Tieg3 and detection of two putative nuclear localization signals within the zinc finger DNA-binding domain. J Cell Biochem 101: 712–722.

33. Fernandez-Zapico ME, van Velkinburgh JC, Gutierrez-Aguilar R, Neve B, Froguel P, et al. (2009) MODY7 gene, KLF11, is a novel p300-dependent regulator of Pdx-1 (MODY4) transcription in pancreatic islet beta cells. J Biol Chem 284: 36482–36490.

34. Neve B, Fernandez-Zapico ME, Ashkenazi-Katalan V, Dina C, Hamid YH, et al. (2005) Role of transcription factor KLF11 and its diabetes-associated gene variants in pancreatic beta cell function. Proc Natl Acad Sci U S A 102: 4807–4812.

35. Noti JD, Johnson AK, Dillon JD (2004) The zinc finger transcription factor transforming growth factor beta-inducible early gene-1 confers myeloid-specific activation of the leukocyte integrin CD11d promoter. J Biol Chem 279: 26948–26958.

36. Zhang JS, Moncrieffe MC, Kaczynski J, Ellenrieder V, Prendergast FG, et al. (2001) A conserved alpha-helical motif mediates the interaction of Sp1-like transcriptional repressors with the corepressor mSin3A. Mol Cell Biol 21: 5041–5049.

37. Kim J, Shin S, Subramaniam M, Bruinsma E, Kim TD, et al. (2010) Histone demethylase JARID1B/KDM5B is a corepressor of TIEG1/KLF10. Biochem Biophys Res Commun 401: 412–416.

38. Pandya K, Townes TM (2002) Basic residues within the Kruppel zinc finger DNA binding domains are the critical nuclear localization determinants of EKLF/KLF-1. J Biol Chem 277: 16304–16312.

39. Quadrini KJ, Bieker JJ (2002) Kruppel-like zinc fingers bind to nuclear import proteins and are required for efficient nuclear localization of erythroid Kruppel-like factor. J Biol Chem 277: 32243–32252.

40. Boulikas T (1993) Nuclear localization signals (NLS). Crit Rev Eukaryot Gene Expr 3: 193–227.

41. Cokol M, Nair R, Rost B (2000) Finding nuclear localization signals. EMBO Rep 1: 411–415.

42. Poon IK, Jans DA (2005) Regulation of nuclear transport: central role in development and transformation? Traffic 6: 173–186.

43. Fernandez-Costa JM, Artero R (2010) A conserved motif controls nuclear localization of Drosophila Muscleblind. Mol Cells 30: 65–70.

44. Mosca TJ, Schwarz TL (2010) Drosophila Importin-alpha2 is involved in synapse, axon and muscle development. PLoS One 5: e15223.

45. Wong LC, Schedl P (2006) Dissection of Drosophila ovaries. J Vis Exp. pp 52.

46. Munoz-Soriano V, Belacortu Y, Durupt FC, Munoz-Descalzo S, Paricio N (2011) Mtl interacts with members of Egfr signaling and cell adhesion genes in the Drosophila eye. Fly (Austin) 5: 88–101.

47. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680.

48. Brand AH, Perrimon N (1993) Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118: 401–415.

49. Brown JL, Grau DJ, DeVido SK, Kassis JA (2005) An Sp1/KLF binding site is important for the activity of a Polycomb group response element from the Drosophila engrailed gene. Nucleic Acids Res 33: 5181–5189.

50. Lowrey PL, Takahashi JS (2004) Mammalian circadian biology: elucidating genome-wide levels of temporal organization. Annu Rev Genomics Hum Genet 5: 407–441.

51. Reppert SM, Weaver DR (2002) Coordination of circadian timing in mammals. Nature 418: 935–941.

52. Wager-Smith K, Kay SA (2000) Circadian rhythm genetics: from flies to mice to humans. Nat Genet 26: 23–27.
53. Pang YP, Kumar GA, Zhang JS, Urrutia R (2003) Differential binding of Sin3 interacting repressor domains to the PAH2 domain of Sin3A. FEBS Lett 31: 108–12.
54. Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res 31: 3381–3385.
55. Spain MM, Caruso JA, Swaminathan A, Pile LA (2010) Drosophila SIN3 isoforms interact with distinct proteins and have unique biological functions. J Biol Chem 285: 27457–27467.
56. Miller M, Park MK, Hanover JA (1991) Nuclear pore complex: structure, function, and regulation. Physiol Rev 71: 909–949.
57. Nakai K, Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci 24: 34–36.
58. Kosugi S, Hasebe M, Tomita M, Yanagawa H (2009) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proc Natl Acad Sci U S A 106: 10171–10176.
59. Kalderon D, Roberts BL, Richardson WD, Smith AE (1984) A short amino acid sequence able to specify nuclear location. Cell 39: 499–509.
60. Boulikas T (1994) Putative nuclear localization signals (NLS) in protein transcription factors. J Cell Biochem 55: 32–58.
61. Nardozzi JD, Lott K, Cingolani G (2010) Phosphorylation meets nuclear import: a review. Cell Commun Signal 8: 32.
62. Obenauer JC, Cantley LC, Yaffe MB (2003) Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. Nucleic Acids Res 31: 3635–3641.
63. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, et al. (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266: 66–71.
64. Chen CF, Li S, Chen Y, Chen PL, Sharp ZD, et al. (1996) The nuclear localization sequences of the BRCA1 protein interact with the importin-alpha subunit of the nuclear transport signal receptor. J Biol Chem 271: 32863–32868.
65. Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, et al. (2007) Classical nuclear localization signals: definition, function, and interaction with importin alpha. J Biol Chem 282: 5101–5105.
66. Mason DA, Stage DE, Goldfarb DS (2009) Evolution of the metazoan-specific importin alpha gene family. J Mol Evol 68: 351–365.
67. Dockendorff TC, Tang Z, Jongens TA (1999) Cloning of karyopherin-alpha3 from Drosophila through its interaction with the nuclear localization sequence of germ cell-less protein. Biol Chem 380: 1263–1272.
68. Friedrich B, Quensel C, Sommer T, Hartmann E, Kohler M (2006) Nuclear localization signal and protein context both mediate importin alpha specificity of nuclear import substrates. Mol Cell Biol 26: 8697–8709.
69. Giarre M, Torok I, Schmitt R, Gorjanacz M, Kiss I, et al. (2002) Patterns of importin-alpha expression during Drosophila spermatogenesis. J Struct Biol 140: 279–290.
70. Mason DA, Fleming RJ, Goldfarb DS (2002) Drosophila melanogaster importin alpha1 and alpha3 can replace importin alpha2 during spermatogenesis but not oogenesis. Genetics 161: 157–170.
71. Mason DA, Mathe E, Fleming RJ, Goldfarb DS (2003) The Drosophila melanogaster importin alpha3 locus encodes an essential gene required for the development of both larval and adult tissues. Genetics 165: 1943–1958.
72. Mathe E, Bates H, Huikeshoven H, Deak P, Glover DM, et al. (2000) Importin-alpha3 is required at multiple stages of Drosophila development and has a role in the completion of oogenesis. Dev Biol 223: 307–322.
73. Ratan R, Mason DA, Sinnot B, Goldfarb DS, Fleming RJ (2008) Drosophila importin alpha1 performs paralog-specific functions essential for gametogenesis. Genetics 178: 839–850.
74. Stark C, Breitkreutz BJ, Chatr-Aryamontri A, Boucher L, Oughtred R, et al. (2011) The BioGRID Interaction Database: 2011 update. Nucleic Acids Res 39: D698–704.
75. Murali T, Pacifico S, Yu J, Guest S, Roberts GG, 3rd, et al. (2011) DroID 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for Drosophila. Nucleic Acids Res 39: D736–43.
76. Yu J, Pacifico S, Liu G, Finley RL, Jr. (2008) DroID: the Drosophila Interactions Database, a comprehensive resource for annotated gene and protein interactions. BMC Genomics 9: 461.
77. Chan CC, Zhang S, Rousset R, Wharton KA, Jr. (2008) Drosophila Naked cuticle (Nkd) engages the nuclear import adaptor Importin-alpha3 to antagonize Wnt/beta-catenin signaling. Dev Biol 318: 17–28.
78. Gorjanacz M, Adam G, Torok I, Mechler BM, Szlanka T, et al. (2002) Importin-alpha 2 is critically required for the assembly of ring canals during Drosophila oogenesis. Dev Biol 251: 271–282.
79. Mosca TJ, Schwarz TL (2010) The nuclear import of Frizzled2-C by Importins-beta11 and alpha2 promotes postsynaptic development. Nat Neurosci 13: 935–943.
80. Kussel P, Frasch M (1995) Pendulin, a Drosophila protein with cell cycle-dependent nuclear localization, is required for normal cell proliferation. J Cell Biol 129: 1491–1507.
81. Torok I, Strand D, Schmitt R, Tick G, Torok T, et al. (1995) The overgrown hematopoietic organs-31 tumor suppressor gene of Drosophila encodes an Importin-like protein accumulating in the nucleus at the onset of mitosis. J Cell Biol 129: 1473–1489.
82. Hanna-Rose W, Hansen U (1996) Active repression mechanisms of eukaryotic transcription repressors. Trends Genet 12: 229–234.
83. Gray S, Levine M (1996) Transcriptional repression in development. Curr Opin Cell Biol 8: 358–364.
84. Fisher AL, Caudy M (1998) Groucho proteins: transcriptional corepressors for specific subsets of DNA-binding transcription factors in vertebrates and invertebrates. Genes Dev 12: 1931–1940.
85. Carroll SB (1990) Zebra patterns in fly embryos: activation of stripes or repression of interstripes? Cell 60: 9–16.
86. Ellenrieder V, Zhang JS, Kaczynski J, Urrutia R (2002) Signaling disrupts mSin3A binding to the Mad1-like Sin3-interacting domain of TIEG2, an Sp1-like repressor. EMBO J 21: 2451–2460.
87. Song CZ, Keller K, Chen Y, Stamatoyannopoulos G (2003) Functional interplay between CBP and PCAF in acetylation and regulation of transcription factor KLF13 activity. J Mol Biol 329: 207–215.
88. Chen C, Agnès F, Gélinas C (1999) Mapping of a serine-rich domain essential for the transcriptional, antiapoptotic, and transforming activities of the v-Rel oncoprotein. Mol Cell Biol 19: 307–16.
89. Perrimon N, McMahon AP (1999) Negative feedback mechanisms and their roles during pattern formation. Cell 97: 13–16.
90. Freeman M (2000) Feedback control of intercellular signalling in development. Nature 408: 313–319.
91. Harden N (2002) Signaling pathways directing the movement and fusion of epithelial sheets: lessons from dorsal closure in Drosophila. Differentiation 70: 181–203.
92. Hardin PE, Hall JC, Rosbash M (1990) Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels. Nature 343: 536–540.
93. Allada R, Chung BY (2010) Circadian organization of behavior and physiology in Drosophila. Annu Rev Physiol 72: 605–624.
94. Hogenesch JB, Ueda HR (2011) Understanding systems-level properties: timely stories from the study of clocks. Nat Rev Genet 12: 407–416.
95. Dang DT, Zhao W, Mahatan CS, Geiman DE, Yang VW (2002) Opposing effects of Kruppel-like factor 4 (gut-enriched Kruppel-like factor) and Kruppel-like factor 5 (intestinal-enriched Kruppel-like factor) on the promoter of the Kruppel-like factor 4 gene. Nucleic Acids Res 30: 2736–2741.
96. Rodríguez E, Martignetti JA (2009) The Krüppel traffic report: cooperative signals direct KLF8 nuclear transport. Cell Res 19: 1041–3.
97. Tuncher A, Sprote P, Gehrke A, Brakhage AA (2005) The CCAAT-binding complex of eukaryotes: evolution of a second NLS in the HapB subunit of the filamentous fungus Aspergillus nidulans despite functional conservation at the molecular level between yeast, A.nidulans and human. J Mol Biol 352: 517–533.
98. Blom N, Gammeltoft S, Brunak S (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J Mol Biol 294: 1351–1362.
99. Rihs HP, Peters R (1989) Nuclear transport kinetics depend on phosphorylation-site-containing sequences flanking the karyophilic signal of the Simian virus 40 T-antigen. EMBO J 8: 1479–1484.
100. Mehta TS, Lu H, Wang X, Urvalek AM, Nguyen KH, et al. (2009) A unique sequence in the N-terminal regulatory region controls the nuclear localization of KLF8 by cooperating with the C-terminal zinc-fingers. Cell Res 19: 1098–109.
101. Chen G, Li W, Zhang QS, Regulski M, Sinha N, et al. (2008) Identification of synaptic targets of Drosophila pumilio. PLoS Comput Biol 4: e1000026.
102. Tamanini F, Yagita K, Okamura H, van der Horst GT (2005) Nucleocytoplasmic shuttling of clock proteins. Methods Enzymol 393: 418–435.
103. Miyazaki K, Mesaki M, Ishida N (2001) Nuclear entry mechanism of rat PER2 (rPER2): role of rPER2 in nuclear localization of CRY protein. Mol Cell Biol 21: 6651–6659.
104. Jans DA, Hubner S (1996) Regulation of protein transport to the nucleus: central role of phosphorylation. Physiol Rev 76: 651–685.
105. Nicol JW, Helt GA, Blanchard SG, Jr., Raja A, Loraine AE (2009) The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics 25: 2730–2731.