Development 135, 133-143 (2008) doi:10.1242/dev.014068

A stem-loop structure in the *wingless* transcript defines a consensus motif for apical RNA transport

Gilberto dos Santos, Andrew J. Simmonds* and Henry M. Krause†

Although the subcellular localization of mRNA transcripts is a well-established mechanism for controlling protein localization, the basis for the recognition of mRNA localization elements is only now emerging. For example, although localization elements have been defined for many mRNAs that localize to apical cytoplasm in *Drosophila* embryos, no unifying properties have been identified within these elements. In this study, we identify and characterize an apical localization element in the 3′UTR of the *Drosophila wingless* mRNA. We show that this element, referred to as WLE3, is both necessary and sufficient for apical RNA transport. Full, unrestricted activity, however, requires the presence of one of several downstream potentiating elements. Comparison of WLE3 sequences within the *Drosophila* genus, and their predicted secondary structures, defines a highly conserved stem-loop structure. Despite these high levels of sequence and predicted structure conservation, however, mutagenesis shows significant leeway for both sequence and structure variation in the predicted stem-loop. Importantly, the features that emerge as crucial include an accessible distal helix sequence motif, which is also found in the predicted structures of other apical localization elements.

KEY WORDS: *wingless*, RNA localization, RNA structure, Apical localization element, Stem-loop

INTRODUCTION

The subcellular localization of a mRNA controls the site of protein translation, thereby enriching the protein in areas of particular need and preventing expression in areas where its presence would be useless or damaging. This mechanism for generating cellular asymmetry occurs in most organisms and cell types, and with mRNAs that encode many different kinds of proteins (St Johnston, 2005). During early *Drosophila* development, 70% of expressed mRNAs are subcellularly localized (Lécuyer, 2007). Thus, transcript localization clearly plays a major role in the control of cellular differentiation and development. Transcript localization is achieved by one or more of the following general mechanisms: local synthesis within syncytial tissues (Brenner et al., 1990), local protection from degradation (Bashirullah et al., 1999), diffusion and entrapment (Forrest and Gavis, 2003; Glotzer et al., 1997), or active transport (Bertrand et al., 1998; Betley et al., 2004; Brendza et al., 2000; Cha et al., 2001; Latham et al., 2001; MacDougall et al., 2003; Wilkie and Davis, 2001). Cis-acting elements and trans-acting machineries have been identified for many of these variant processes. However, there is little obvious similarity among elements that mediate similar localization activities, and even less is known about the proteins that recognize and bind them.

In *Drosophila*, most of the actively localized transcripts studied are transported along minus end-directed microtubules via Dynein-based motors (Cha et al., 2001; Hughes et al., 2004; MacDougall et al., 2003; Wilkie and Davis, 2001). This process also requires proteins encoded by the *egalitarian* (*egl*) and *Bicaudal D* (*BicD*) genes (Bullock and Ish-Horowicz, 2001), which are thought to act as adaptors, given that each can interact with the Dynein motor (Hoogenraad et al., 2001; Matanis et al., 2002) and each other (Mach and Lehmann, 1997). Transcripts are then anchored in a manner that requires Dynein, but not Dynein motor activity, Egl or BicD (Delanoue and Davis, 2005).

Apical transcript localization at this and later stages by the Dynein complex is important for the correct function of many developmental control genes. For example, the apical localization of *wingless* (*wg*) mRNA, one of the first transcripts shown to be mediated by Dynein motors (Wilkie and Davis, 2001), ensures the proper processing, secretion and extracellular distribution of the encoded protein (Simmonds et al., 2001). Another example in the segmentation gene category is *hairy* (*h*), for which apical localization and anchoring ensures entry of the translated protein into the nearest nucleus (Bullock et al., 2004).

Little is known about how these Dynein mobilized mRNAs are specifically recognized. For example, despite the relatively large group of mRNAs known to localize apically, only a few discrete elements have been mapped and characterized, and these appear to vary greatly. The apical localization of *bicoid* (*bcd*) transcripts, for example, requires a large 437 nucleotide minimal element, composed of stems III, IV and V of the 3′UTR arranged in a conserved, highly ordered secondary structure (MacDonald, 1990; Snee et al., 2005). In contrast, the 'transport and localization signal' (TLS) of the *K10* [also known as *fs(1)K10* – FlyBase] transcript appears to be a 44 nucleotide stem-loop (Serano and Cohen, 1995), and the '*gurken* (*grk*) localization signal' (GLS) is a 65 nucleotide stem-loop (Bullock and Ish-Horowicz, 2001; Van De Bor et al., 2005). The *h* and *fushi tarazu* (*ftz*) localization elements (HLE and FLE, respectively) each contain two discrete stem-loops, both of which are necessary for activity (Bullock et al., 2003; Snee et al., 2005). Finally, two mapped but uncharacterized apical localization elements in the *wingless* transcript, WLE1 and WLE2, show no obvious similarities to each other or to other apical elements (Simmonds et al., 2001).

This lack of obvious similarities between mapped localization elements suggests that they are either recognized in fundamentally different ways or share cryptic similarities. In yeast, the elucidation of a cryptic transcript localization motif, shared by transcripts colocalized to the bud tip, was achieved by the comparison of localization elements from ten different mRNAs (Jambhekar et al.,

Banting and Best Department of Medical Research, Department of Medical Genetics and Microbiology, University of Toronto, Ontario, M5S 3E1, Canada.

*Present address: Department of Cell Biology, Faculty of Medicine and Dentistry, 5-14 Medical Sciences Building, University of Alberta, Edmonton, Alberta, Canada
†Author for correspondence (e-mail: h.krause@utoronto.ca)

Accepted 27 September 2007

2005; Shepard et al., 2003). Likewise, the identification of a consensus apical localization motif in *Drosophila* may require the identification of more localized mRNAs and their corresponding localization elements.

In this study, we identify and characterize a new apical localization element in the *wingless* mRNA 3′UTR, which we refer to as WLE3. We show that WLE3 is both necessary and sufficient for apical transport in an embryo microinjection assay, although it requires one of several potentiating elements present in the 3′UTR for full activity. A phylogenetic comparison of WLE3 elements predicts a highly conserved stem-loop structure. Further mutagenic analysis, however, shows that much of this conserved sequence is unimportant, so long as a few key residues, base pairs and bulges are maintained. Notably, these essential features are also present in other apical localization elements, defining a consensus motif that is likely to be present in many if not most apically localized transcripts.

MATERIALS AND METHODS

Plasmid construction

Drosophila melanogaster sequences are numbered according to the *wg* cDNA (GenBank accession M17230) starting at the first base of the 3′UTR. Truncations and deletions were generated by standard PCR and assembled in pBluescript SK (Stratagene; GenBank accession X52324) with a 5′ *Xho*I site, a 3′ *Xba*I site, and internal deletions spanned by *Bgl*II sites. WLE3 mutations were generated by cloning synthetic oligonucleotides (spanning nucleotides 518-570) with *Xho*I-*Bgl*II ends upstream of a distal portion of the *wg* 3′UTR (nucleotides 770-1100). All constructs were confirmed by sequencing.

The 2×WLE3 construct was assembled in a modified pBluescript SK vector in which the polylinker was replaced between the *Sac*I and *Kpn*I sites using a synthetic oligonucleotide. The first WLE3 repeat (nucleotides 525-568), amplified as a *Bgl*II-*Bam*HI fragment, was inserted into the *Bgl*II site. A second fully active WLE3 variant (construct 28, see below) was PCR amplified as a *Spe*I-*Xba*I fragment and cloned into the *Spe*I site. The two repeats are separated by a six adenine base spacer, with a downstream *Cla*I site used for plasmid linearization. The 2×WLE3 and flanking polylinker sequence is as follows (WLE3 repeats are in bold and relevant restriction sites underlined): 2×WLE3: GAGCTCAGATCT**TGCTTGCATACTGCTTTGGCCAGGACCAAAACGTATGCGAAGTG**GGATCTAAAAAAAACTAGT**TGCTTGCATACTGCTTTCCCCAGGAGGAAAACGTATGCGAAGTG**TCTAGTAAAAAAATCGATTCTCGAGGGGGGGCCCGGTACC.

UAS-*lacZ* transgene construction

A pre-existing pUAS-*lacZ* P-element vector (Brand and Perrimon, 1993) was modified by the insertion of a polylinker (*Xho*I-*Bgl*II-*Spe*I-*Xba*I) into a unique *Xba*I site 276 nucleotides downstream of the *lacZ* ORF stop codon. The *wg* 3′UTR elements were transferred into this new vector, pUAS-*lacZ* II BS, as *Xho*I-*Xba*I fragments. P-element-mediated genome transformation was performed as described (Robertson et al., 1988; Rubin and Spradling, 1982).

Cloning *wg* 3′UTR sequences from *Drosophila* species

Drosophila and *Zaprionus tuberculatus* fruit flies were obtained from the Tucson *Drosophila* Species Stock Center (Table 1). The strategy in cloning *wg* 3′UTR sequences was to first PCR amplify a specific portion of the *wg* ORF for each species using degenerate primers. The sequenced PCR product was used to design species-specific *wg* primers for 3′ RACE. For degenerate primer PCR, genomic DNA was extracted from adult flies as described previously (Ballinger and Benzer, 1989). Degenerate primers and amplification conditions were the same as those used previously to amplify butterfly *wg* sequences (Brower and DeSalle, 1998). The identity of cloned inserts was confirmed by alignment to other *Wnt1* sequences and they were deposited in GenBank (accession numbers DQ778961-DQ778974).

For 3′RACE, poly(A) RNA was isolated from 100 μl of pupae (exhibiting leg eversion but not eye pigmentation) using the Qiagen RNeasy and Qiagen Oligotex Mini Kits. Total cDNA was synthesized using the avian myeloblastosis virus (AMV) reverse transcriptase based Takara 3′RACE kit according to the manufacturer's directions. The cDNA product was used in an anchor PCR reaction with *wg*-specific 5′ primer at 0.5 μM and universal 3′ primer at 0.05 μM. Anchor PCR products in the range of 1.0 to 1.5 kb were gel purified and used in a nested PCR reaction. Anchor and nested PCR

Table 1. *Drosophila* strains and primers used in cloning *wg* 3′UTRs

Species	Tucson strain		Gene-specific primers
Drosophila ananassae	14024-0371.00	Anchor	5′-CGAAATGGACATCGTCGGGGCCGCAAGCAC-3′
		Nested	5′-CTGGATCCCGGGCGTGGCTAC-3′
Drosophila auraria	14028-0471.00	Anchor	5′-GCCAGGAATGGACGACGCCAGGGCCGCAAG-3′
		Nested	5′-CTGGATCCCGGGCGTGGCTAC-3′
Drosophila ficusphila	14025-0441.01	Anchor	5′-CCGGATCCTGATGTGCTGTGG-3′
		Nested	5′-CTGGATCCGCGTGGCTATCG-3′
Drosophila hydei	15085-1641.00	Anchor	5′-AGGGATCCCCAAGTTTCTGCGAGAAGAATC-3′
		Nested	5′-GAGGATCCATGTGCTGTGGGCGTGGCTATC-3′
Drosophila lucipennis	14023-0331.00	Anchor	5′-CCGGATCCTGATGTGCTGTGG-3′
		Nested	5′-CTGGATCCGCGTGGCTATCG-3′
Drosophila mauritiana	14021-0241.01	Anchor	5′-GAGGATCCGGAGCCTTCGCCGAGCTTCTGC-3′
		Nested	Not applicable
Drosophila prosaltans	14045-0901.07	Anchor	5′-CCGGATCCTGATGTGCTGTGG-3′
		Nested	5′-CTGGATCCGCGTGGCTATCG-3′
Drosophila pseudoobscura	14011-0121.00	Anchor	5′-GGCGGGAGCAAGGTCCAGGAACGGACGTCG-3′
		Nested	5′-CTGGATCCGCGTGGCTATCG-3′
Drosophila robusta	15020-1111.01	Anchor	5′-CCGGATCCTGATGTGCTGTGG-3′
		Nested	5′-CTGGATCCGCGTGGCTATCG-3′
Drosophila simulans	14021-0251.00	Anchor	5′-GAGGATCCAACCCGCACAATCCCGAGCACA-3′
		Nested	Not applicable
Drosophila takahashii	14022-0311.00	Anchor	5′-CCGGATCCCTGCGAGAAGAATCTGCGACAC-3′
		Nested	5′-ACGGATCCGGGCTGATGTGCTGTGGGCGTG-3′

For each species, the Tucson Stock Center strain number and primers used in 3′RACE cloning of the *wg* cDNA are listed.

reactions were performed under standard conditions using an 8:1 Taq:Pfu DNA polymerase mix with the cycling regimen as follows: 94°C for 120 seconds, 1 cycle; 94°C for 45 seconds, 55°C for 45 seconds, 72°C for 90 seconds, 30 cycles; 72°C for 10 minutes, 1 cycle. Gene-specific anchor and nested PCR primers are listed in Table 1. The nested PCR products were cloned and sequenced. These partial *wg* cDNA sequences were deposited in GenBank (accession numbers DQ778975-DQ778988).

Additional *Drosophila wg* 3′UTR sequences were obtained by querying genome projects via the FlyBase Blast Service (http://flybase.bio.indiana.edu/blast/) using the *D. melanogaster* sequence as bait.

Fluorescence in situ hybridization (FISH)

In all cases embryo fixation, DIG-labelled probe synthesis, hybridization and DIG detection were as described previously (Hughes and Krause, 1999), except that for the detection of endogenous *wg* RNA in *Drosophila* species, the anti-DIG antibody was detected by tyramide signal amplification using Alexa Fluor 488 (Molecular Probes) as described previously (Lécuyer et al., 2007). For the detection of *lacZ* reporter transcripts, *patched-Gal4* females were crossed to UAS-*lacZ* males to drive reporter expression in progeny (Brand and Perrimon, 1993). The DIG-labelled probe used was complementary to a *Pvu*II fragment of the *lacZ* ORF. For each construct two transgenic lines were tested. 3′ cDNAs were used to make probes for endogenous *wg* RNA detection.

Phylogenetic WLE3 structure analysis

Drosophila wg 3′UTR sequences were aligned using the ClustalW program (Thompson et al., 1994) as available online (http://www.ebi.ac.uk/clustalw/). The alignment spanning WLE3 (*D. melanogaster* nucleotides 525-570) was then analyzed for conserved secondary structure using ALIFOLD (Hofacker et al., 2002) as available online (http://rna.tbi.univie.ac.at/cgi-bin/alifold.cgi). The secondary structure of each WLE3 sequence was also analyzed using MFOLD 2.3 (Mathews et al., 2004; Mathews et al., 1999) with the temperature set at 25°C.

Fluorescent RNA microinjection into pre-blastoderm embryos

For run-off transcription, 10 μg of template DNA was cut to completion with an appropriate restriction enzyme, followed by phenol:chloroform and chloroform extractions and ethanol precipitation. Template DNA was resuspended in 20 μl RNase-free water. 1.0 μg of purified template DNA was used per transcription reaction with 20 Units of T3 or T7 RNA polymerase, 0.4 mM ATP, CTP, GTP, 0.36 mM UTP and 0.04 mM UTP-Alexa Fluor 488 or UTP-Alexa Fluor 546 (Molecular Probes) and RNase inhibitor. RNA was purified using a G-50 size exclusion RNA spin column (Roche) and precipitated in 1.0 M ammonium acetate, 75% ethanol. The pellet was resuspended in water to a final concentration of 200 ng/μl (confirmed by gel electrophoresis).

Wild-type embryos (OregonR) were collected for 30 minutes and aged 2 hours (25°C). Dechorionated embryos were transferred to a coverslip and covered with halocarbon oil. An Eppendorf 5410C microinjection unit and a Narishige micromanipulator were used to inject about 4 pl RNA/injection. All embryos on a slide were injected within 5 minutes and aged 3 minutes before imaging. For a given slide, the order of embryo image capture followed that of embryo injection. All images were captured within 8 minutes of injection. Localization efficiency for each injection was quantified using Adobe Photoshop 7.0 to measure the ratio of RNA signal intensity apical and basal to the nuclei.

RESULTS

Identification of WLE3

In a previous study that made use of transgenic *wingless* (*wg*) constructs, two regions of the *wg* 3′UTR, referred to as WLE1 and WLE2, were found to be capable of localizing *wg* transcripts to apical cytoplasm (Simmonds et al., 2001). To characterize these elements further, and to test for the existence of additional localization elements, we used a fluorescent RNA microinjection assay described previously by Lall et al. (Lall et al., 1999). Wilkie and Davis (Wilkie and Davis, 2001) have shown that *wg* mRNA

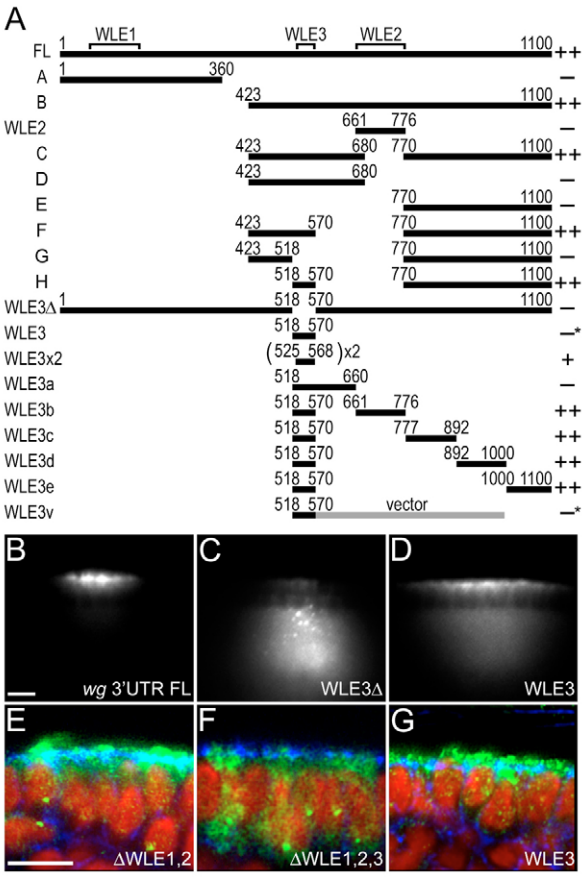

Fig. 1. WLE3 mediates apical localization of *wg* transcripts. (**A**) Fluorescent deletion constructs of *wg* 3′UTR RNA were injected into *D. melanogaster* embryos. Localization efficiency was categorized based on the apical:basal ratio of fluorescent RNA signal as strong (++, ratio ⩾1.5), weak (+, 1.5>ratio⩾1.0) or inactive (–, ratio <1.0). The typical activity of each construct is indicated at the right (* constructs typically inactive but occasionally localized). WLE3 spans nucleotides 518-570, WLE1 57-183, and WLE2 661-776. (**B-D**) Representative images of embryos injected with fluorescently labelled full-length *wg* 3′UTR (B), deleted for WLE3 (C) or WLE3 dimer (D). (**E-G**) Confocal images of stage 11 *patched*-GAL4::UAS-*lacZ* embryos containing different regions of the *wg* 3′UTR fused to *lacZ* ORF: ΔWLE1-ΔWLE2 (E); ΔWLE1-ΔWLE2-ΔWLE3 (F); WLE3 (G). *lacZ* reporter transcripts (green) were detected by FISH. Nuclei are red and membranes blue. Scale bars: 10 μm.

localizes apically in a Dynein-dependent fashion when injected into syncitial stage *Drosophila* embryos. Consistent with their findings, we found that reporter RNA containing the *wg* 3′UTR is localized to apical cytoplasm within 8 minutes after injection, in 97% of embryos (construct FL, Fig. 1B, Table 2). Interestingly, WLE1 alone (construct A) did not localize in this assay, and deleting this region from the 3′UTR (construct B) lowered activity only marginally (Table 2).

The contribution of WLE2 in this system was also difficult to judge. Construct C, in which WLE1 and WLE2 are both removed, does exhibit a small but significant reduction in activity. The minimal WLE2 element, when injected, forms large particles that accumulate below the blastoderm nuclei (Table 2; not shown). As this clumping activity may prevent further passage between the nuclei, we cannot conclude that localizing activity is absent.

Table 2. WLE3 is necessary for apical *wg* RNA localization in injected embryos

Construct	*wg* 3′UTR region	Intensity ratio	Localization (%) ++	+	–	*n*
FL	1-1100	2.80±0.84	**97**	3	0	35
A	1-360 (spans WLE1)	0.80±0.13	0	2	**98***	40
B	423-1100	2.54±0.91	**84**	16	0	32
WLE2	661-776	Not determined	0	0	**100**†	23
C	423-680 + 770-1100	1.37±0.29	33	**56**	11	54
D	423-680	0.81±0.09	0	0	**100**	21
E	770-1100	0.50±0.06	0	0	**100**	15
F	423-570 + 770-1100	1.34±0.46	20	**67**	13	30
G	423-518 + 770-1100	0.54±0.11	0	0	**100**	42
H	518-570 + 770-1100	2.20±0.77	**77**	23	0	56
ΔWLE3	1-518 + 570-1100	0.54±0.12	0	0	**100**†	33
WLE3	518-570	0.96±0.12	0	33	**67**	21
2×WLE3	525-568 dimer	1.40±0.29	32	**66**	2	41
WLE3a	518-570 + 570-660	0.86±0.12	0	9	**91**	76
WLE3b	518-570 + 661-776	2.06±0.47	**86**	14	0	28
WLE3c	518-570 + 777-892	1.50±0.32	42	**58**	0	31
WLE3d	518-570 + 892-1000	2.44±0.57	**100**	0	0	31
WLE3e	518-570 + 1000-1100	1.71±0.27	**78**	22	0	23
WLE3v	518-570 + vector‡	0.98±0.11	0	46	**54**	13

Construct names are shown in the left column, with a description to their right. For each construct, localization efficiencies were measured as the apical:basal ratio of fluorescent RNA signal, with the average ratio and standard deviations shown. The right hand columns indicate the number of injections that resulted in strong (++, ratio≥1.5), weak (+, 1.5>ratio≥1.0) or no (–, ratio<1.0) apical localization (percent). Sample size (*n*) refers to the number of injections. Figures in bold highlight the category with the most prevalent activity level for a given construct.

*Very weak localization was observed in 67% of injections (*n*=9) in post-cellularization embryos. Embryos at this late stage of development were not included in the final tally for this or any other construct.

†In about half of the injections, RNA formed large particles that accumulated at the base of cortical nuclei.

‡The 422-nucleotide *Xba*I-*Afl*II pBluescript SK– sequence was included downstream of the WLE3 fragment.

The remaining localization element that targets construct C apically was mapped to a 53 nucleotide region (nucleotides 518-570), which we designate as WLE3. Deletion of WLE3 in an otherwise full-length *wg* 3′UTR transcript completely abolishes apical accumulation (construct ΔWLE3, Fig. 1C), with transcripts forming large particles reminiscent of the minimal WLE2 construct. Conversely, a dimer of the minimal WLE3 element shows robust apical localization activity (construct 2×WLE3, Fig. 1D). As with other apical minimal localization elements identified by microinjection (Bullock et al., 2003; Snee et al., 2005), a WLE3 monomer shows weak activity on its own, and even as a dimer, does not localize as well as the monomer element in its normal context (Table 2). Thus, WLE3 is necessary for localization activity in this assay, and is sufficient for partial activity.

The incomplete activity of a single WLE3 element indicates a requirement for additional elements, sequences or constraints provided by flanking RNA. Accordingly, regions downstream of WLE3 were scanned in windows of about 100 nucleotides for the ability to confer full activity to WLE3 (Fig. 1A). Interestingly, four of the five regions tested (constructs WLE3b-WLE3e) enabled full activity of the single WLE3 element, despite not giving apical localizing activity on their own (Table 2). A fifth region just downstream of WLE3 (nucleotides 571-659: construct WLE3a) inhibited WLE3 activity, completely abolishing apical accumulation (Table 2). This negative activity is overcome when any of the four potentiating segments is present. Random pieces of pBluescript sequence, placed downstream of WLE3 (e.g. construct WLE3v), did not affect WLE3 activity (Table 2), suggesting something unique and common to the four activity-potentiating regions. Comparison of these four sequences, however, did not reveal any common sequences or structures of note.

To confirm the relevance of WLE3 activity in vivo, transgenic flies carrying UAS-*lacZ* reporters with *wg* 3′UTR sequences (Fig. 1E-G, Table 3) were generated. Fig. 1E shows that a construct containing WLE3 and potentiating sequences (construct ΔWLE1-ΔWLE2) is apically localized. As with the microinjection assay, WLE3 is necessary for this activity, as a triple WLE1/2/3 deletion renders reporter transcripts uniformly localized (Fig. 1F). Somewhat

Table 3. WLE3 mediates apical transcript localization in transgenic reporter constructs

Construct	*wg* 3′UTR region	Cell position	Localization (%)
ΔWLE1-ΔWLE2	423-680 + 770-1100	Dorsal ectoderm	89.4 (47, 3)
		Ventral ectoderm	96.0 (100, 6)*
WLE3	518-570	Dorsal ectoderm	100.0 (103, 7)
		Ventral ectoderm	70.0 (178, 10)†
ΔWLE1-ΔWLE3	423-518 + 570-1100	Ectoderm	0.9 (115, 8)
ΔWLE2-ΔWLE3	1-518 + 570-680 + 770-1100	Ectoderm	0.0 (46, 4)
ΔWLE1-ΔWLE2-ΔWLE3	196-518 + 570-680 + 770-1100	Ectoderm	0.0 (67, 5)

The localization of *patched*-GAL4-driven UAS-*lacZ* reporter transcripts was examined using FISH and confocal microscopy in germ band extended embryos. The reporter transcripts contain a portion of the *wg* 3′UTR between the *lacZ* ORF and *SV40* terminator as indicated. The percentage of cells that displayed apical *lacZ* transcript localization is shown in the right-hand column. Numbers in parentheses indicate the number of cells examined, followed by the number of embryos examined. The probe used is complementary to the *lacZ* ORF.

*Localization was weaker in ventral ectoderm.

†Localization was weak in ventral ectoderm.

surprisingly though, a single copy of the 53 nucleotide WLE3 region, although weak in the ventral ectoderm, is sufficient for strong apical localization in the dorsal ectoderm (Fig. 1G; Table 3). At present, we have no explanation for this dorsal-ventral difference in WLE3 activity, or for the apparent lack of requirement for WLE3 duplication or potentiating sequences when transcribed dorsally. One possibility is that this function can also be fulfilled by sequences in the *SV40* terminator of the UAS-*lacZ* reporter transcript, although only in dorsal ectoderm. Alternatively, unique components of dorsal nuclei/cells may obviate the need for additional elements.

It is notable that, in our previous study, transgenic constructs carrying an intact WLE1 or WLE2 produced localized transcripts, whereas, in the current set of constructs, the UAS-*lacZ* reporters containing only WLE1 (construct ΔWLE2-ΔWLE3) or WLE2 (ΔWLE1-ΔWLE3) are uniformly localized (Table 3). This may be attributable to differences in construct composition, such as the different ORFs used (*wg* in the previous study; *lacZ* in the current study) or the 3′UTR sequences surrounding each WLE (supported by unpublished data, A.J.S.), which are more extensive in the current study. The inclusion of more extensive 3′UTR sequence has been noted to inhibit the activity of other elements such as *even-skipped* (Davis and Ish-Horowicz, 1991) and *bcd* localization elements (Macdonald and Kerr, 1997; Macdonald and Struhl, 1988).

Evolutionary conservation of WLE3

The comparison of related sequences from different species is very useful for revealing evolutionarily conserved motifs that are critical for activity. This is particularly true for non-coding sequences, which evolve rapidly if non-functional. To identify conserved features of the *wg* 3′UTR important for localization, *wg* cDNAs were cloned from 14 *Drosophila* species that diverged as much as 40 to 60 million years ago (Table 1). Apical localization of *wg* mRNA was confirmed for all of the 13 species tested (Fig. 2A-C; *D. lucipennis* not tested), consistent with previous studies showing apical *wg* localization to be conserved among dipteran insects (Bullock et al., 2004). In addition, all 12 *wg* 3′UTRs tested are active upon injection into *D. melanogaster* embryos (Fig. 2D-F, Table 4), demonstrating that some or all mechanisms of apical localization are conserved.

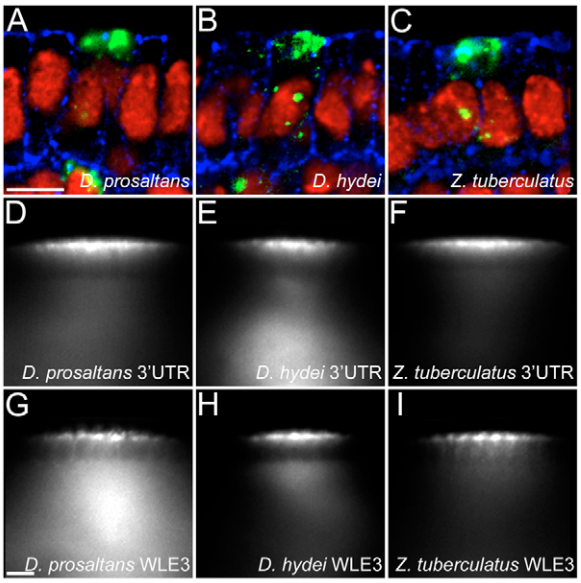

Fig. 2. Apical *wg* localization and WLE3 activity is conserved. (**A-C**) Confocal images of endogenous *wg* RNA (green) detected by FISH in germ band extended embryos of *D. prosaltans* (A), *D. hydei* (B) and *Z. tuberculatus* (C). Nuclei are red and membranes blue. (**D-I**) Representative images of *D. melanogaster* embryos injected with full-length *Drosophila wg* 3′UTR sequences (D-F), or with chimeric sequences in which the *D. melanogaster* WLE3 of construct H was replaced by that of *D. prosaltans* (G), *D. hydei* (H) or *Z. tuberculatus* (I). Scale bars: 10 μm.

An alignment of 21 full-length *wg* 3′UTR sequences (*D. melanogaster* 3′UTR, 14 cloned *wg* 3′UTRs and six *wg* 3′UTR sequences subsequently obtained from *Drosophila* genome projects) show that the three WLEs exhibit varying degrees of conservation. WLE2 shows the least conservation, the sequences being as diverged as other non-conserved regions of the 3′UTR and with no apparent similarities in predicted secondary structure (not shown). By contrast, WLE1 has a high degree of sequence conservation in

Table 4. *Drosophila wg* 3′UTR sequences localize apically in *Drosophila melanogaster*

Species	Intensity ratio	Localization (%) ++	+	–	*n*
Drosophila ananassae	1.58±0.54	45	**52**	3	31
Drosophila auraria	1.58±0.35	**53**	45	2	47
Drosophila ficusphila	1.42±0.32	32	**68**	0	19
Drosophila hydei	1.68±0.32	**73**	27	0	11
Drosophila lucipennis	1.39±0.35	32	**57**	11	75
Drosophila melanogaster	2.80±0.84	**97**	3	0	35
Drosophila prosaltans	1.93±0.53	**78**	20	2	46
Drosophila pseudoobscura	1.79±0.64	**63**	37	0	30
Drosophila takahashii	1.75±0.70	**63**	25	13	16
Drosophila teissieri	2.19±0.85	**79**	14	7	28
Drosophila virilis	1.82±0.71	**66**	28	6	32
Zaprionus tuberculatus	2.04±0.66	**81**	18	1	82
Drosophila hydei WLE3*	1.61±0.56	**43**	43	14	14
Drosophila melanogaster WLE3*	2.20±0.77	**77**	23	0	56
Drosophila prosaltans WLE3*	1.09±0.24	9	**57**	35	23
Zaprionus tuberculatus WLE3*	2.36±0.64	**96**	4	0	24

For each injection, localization efficiency was measured as the apical:basal ratio of fluorescent RNA signal. For each construct, the average apical:basal ratio (±s.d.) is indicated, as well as the number of injections resulting in strong (++, ratio≥1.5), weak (+, 1.5>ratio≥1.0) or no (–, ratio<1.0) apical localization (percent). Sample size (*n*) refers to the number of injections. Bold figures indicate the category with the most prevalent activity level for a given construct.
*The minimal WLE3 sequence of the species indicated was fused to a portion of the *Drosophila melanogaster wg* 3′UTR (nucleotides 770-1100, which alone have no localization activity).

its first half (nucleotides 57-135; ⩾77% for all species), but despite this high sequence similarity, each WLE1 sequence is predicted to adopt one of two alternative secondary structures (not shown).

The most conserved localization element is WLE3, with all sequences being ⩾72% identical to that of *D. melanogaster*, and all sequences predicted to form a similar stem-loop secondary structure (Fig. 3A,B). A consensus secondary structure was determined using the program ALIFOLD (Hofacker et al., 2002). The predicted structure contains distal (base pairs 1-5) and proximal (base pairs 6-14) stems separated by a region with variable extents of base pairing (Fig. 3C). Also noteworthy is a single base bulge between base pairs 9 and 10 of the proximal stem and two invariant residues within the distal loop.

There are two particularly notable features of the ALIFOLD modelled structure. First, this consensus structure is in close agreement with the most thermodynamically favourable structures predicted for each of the *Drosophila* WLE3s (not shown). Although the *D. pseudoobscura* WLE3 is an exception, the ALIFOLD modelled structure is only 6% less stable than the most favourable structure. Second, most substitutions at predicted base pairing positions are compensatory in nature. Of the 48 nucleotide variants that occur at predicted base pairing positions, 43 involve matched substitutions that preserve base pairing, and the five that do not maintain base pairing are at the ends of the proximal stem, with minimal effects on calculated structure stabilities.

This structure was validated further by microinjection of the most diverged WLE3 elements into *D. melanogaster* embryos. WLE3 sequences from *D. prosaltans* (the most diverged sequence within the *Sophophora* subgenus), *D. hydei* (the most diverged sequence within the *Drosophila* genus) and *Z. tuberculatus* (the most diverged of all sequences) all localized apically (Fig. 2G-I, Table 4). Thus, both the predicted WLE3 secondary structure and localization activity are highly conserved.

Structural determinants of WLE3 activity: conserved base pairs

The localization activity of WLE3 in the convenient embryo injection assay makes it possible to further analyze structure-function relationships via targeted mutations. Accordingly, sequence and structural aspects of the WLE3 stem-loop were targeted (nucleotides 518-570; Figs 4, 5, Table 5), and tested in the presence of potentiating sequences (nucleotides 770-1100). All mutant WLE3

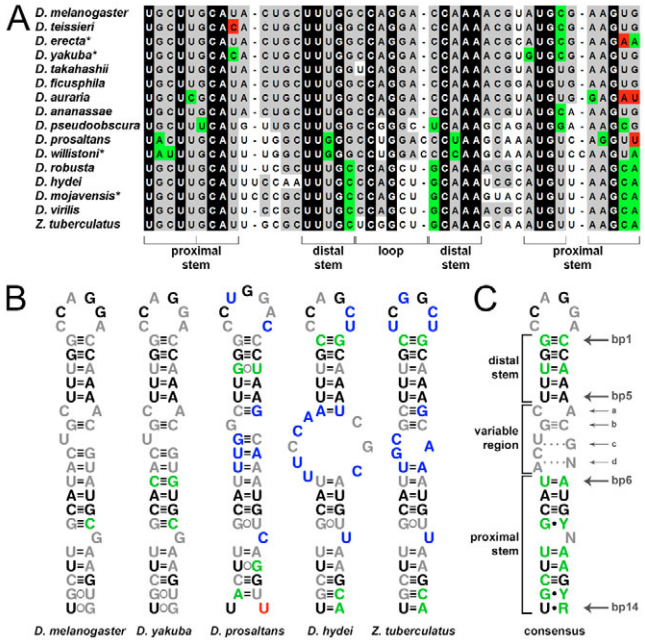

Fig. 3. Predicted WLE3 secondary structure is conserved. (**A**) Alignment of *Drosophila wg* 3′UTR sequences with *D. melanogaster* WLE3 (nucleotides 525-568). Sequences are listed, top to bottom, in order of divergence from *D. melanogaster* (based on full-length 3′UTR pair-wise comparisons). Invariant residues are shaded black, and bases at variable positions similar to the consensus are shaded grey. Compensatory mutations that preserve base-pairing are shaded green and base-pair disrupting mutations red (*sequences obtained from *Drosophila* genome projects). (**B**) Predicted secondary structures for *D. melanogaster*, *D. yakuba*, *D. prosaltans*, *D. hydei* and *Z. tuberculatus* WLE3 sequences. Invariant residues are in black text and consensus residues in grey. Non-consensus residues that preserve or disrupt conserved base pairs are green or red, respectively. Non-consensus residues in the loop and central ‘variable’ region are blue. (**C**) Consensus secondary structure predicted by ALIFOLD. Invariant residues are black, consensus residues grey and positions with conserved base-pairing green.

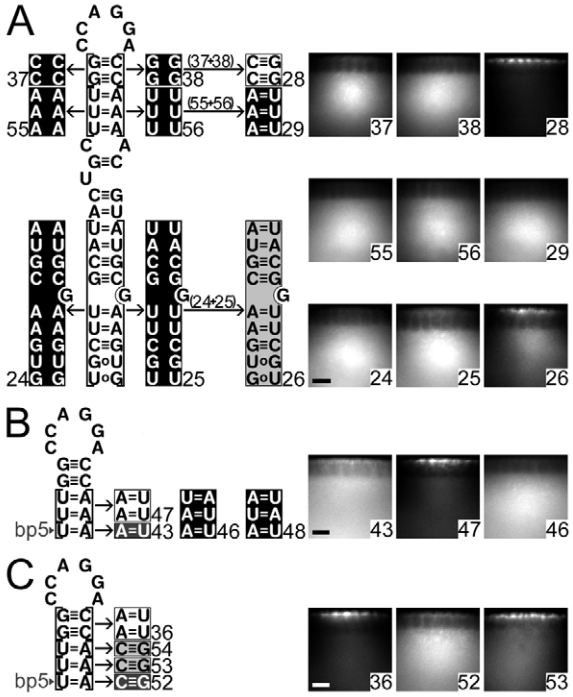

Fig. 4. Importance of sequence and base-pairing at conserved base-pair positions. (**A-C**) Diagrams of *D. melanogaster* WLE3 (A) or just the WLE3 distal stem (B,C). Bases targeted by each mutation are boxed in the wild-type diagram, with arrows pointing to mutant sequences shown to either side (mutant numbers/names are indicated at the bottom). The shading indicates activity: white text on black, no activity (all –); white text on dark grey, weak activity (mostly –); black text on light grey, moderate activity (+); black text on white, full activity (++). To the far right of each diagram are representative images of injected embryos. Scale bars: 10 μm. Base transversions are depicted at the immediate left or right of WLE3, and compensatory mutations further right.

Table 5. Summary of WLE3 mutant analysis

Mutant	Description	Intensity ratio	Localization (%)			*n*
			++	+	–	
H	Wild-type WLE3 construct H (see Fig. 1A)	2.20±0.77	**77**	23	0	56
17	Loop C545U transition	2.09±0.47	**94**	6	0	32
18	Loop G547A transition	2.14±0.46	**96**	4	0	26
20	Bp1-5 disruption (5′ transversion)*	0.58±0.07	0	0	**100**	26
22	Bp1-5 compensatory transversion†	0.53±0.08	0	0	**100**	23
23	Variable region bulge deletion	1.94±0.55	**73**	27	0	22
24	Bp6-14 disruption (5′ transversion)	0.58±0.08	0	0	**100**	41
25	Bp6-14 disruption (3′ transversion)	0.63±0.05	0	0	**100**	34
26	Bp6-14 compensatory transversion (24+25)	1.25±0.21	15	**78**	7	27
27	GAAA tetraloop	2.00±0.71	**77**	18	5	39
28	Bp1-2 compensatory transversion (37+38)	2.03±0.63	**83**	17	0	65
29	Bp3-5 compensatory transversion (55+56)	0.54±0.06	0	0	**100**	21
30	Variable region deletion	0.59±0.08	0	0	**100**	14
31	Variable region bp disruption	0.76±0.14	0	10	**90**	41
36	Bp1-2 G:C→A:U compensatory transition‡	1.78±0.47	**69**	28	3	32
37	Bp1-2 disruption (5′ transversion)	0.57±0.08	0	0	**100**	29
38	Bp1-2 disruption (3′ transversion)	0.55±0.06	0	0	**100**	35
39	Proximal stem bulge deletion	1.35±0.37	30	**59**	11	44
40	No bulges (23+39)	1.23±0.34	17	**60**	23	30
40.5	No bulges, GAAA loop (40+27)	1.39±0.38	27	**60**	13	30
40.6	Bp3-5 swap, no bulges, GAAA loop (40.5+29)	0.57±0.07	0	0	**100**	72
41	Variable region distal bp disruption	0.87±0.21	0	31	**69**	42
42	Variable region proximal bp disruption	1.76±0.50	**75**	25	0	32
43	Bp5 compensatory transversion	0.82±0.23	0	12	**88**	26
44	Bp4 compensatory transversion	1.79±0.47	**74**	20	6	54
45	Bp3 compensatory transversion	1.45±0.34	**46**	42	12	26
46	Bp4-5 compensatory transversion	0.64±0.08	0	0	**100**	13
47	Bp3-4 compensatory transversion	1.53±0.42	**58**	29	13	31
48	Bp3,5 compensatory transversion	0.54±0.08	0	0	**100**	29
49	Variable region bulge CA→AA	0.83±0.15	0	12	**88**	22
50	Variable region bulge CA→CC	1.52±0.35	**50**	45	5	38
51	Variable region bulge CA→AC	1.56±0.20	**53**	47	0	7
52	Bp5 U:A→C:G compensatory transition	0.86±0.19	0	28	**72**	36
53	Bp4 U:A→C:G compensatory transition	1.36±0.33	24	**64**	11	45
54	Bp3 U:A→C:G compensatory transition	1.37±0.37	26	**59**	15	27
55	Bp3-5 disruption (5′ transversion)	0.52±0.06	0	0	**100**	38
56	Bp3-5 disruption (3′ transversion)	0.53±0.07	0	0	**100**	40

Mutant versions of WLE3 (nucleotides 518-570) were placed upstream of potentiating elements (nucleotides 770-1100), fluorescently labelled and injected into *D. melanogaster* embryos. Construct names are shown in the left column, with descriptions to the right. Localization efficiency was measured as the apical:basal ratio of fluorescent RNA signal. Average apical:basal ratios (±s.d.) are indicated, as well as the frequency of strong (++, ratio≥1.5), weak (+, 1.5>ratio≥1.0) or no (–, ratio<1.0) apical localization (percent). Sample size (*n*) refers to the number of injections. Figures in bold indicate the category with the most prevalent activity level for a given construct.
*Transversions limited to the strand upstream of the loop are designated as 5′, those downstream of the loop as 3′.
†Compensatory transversions altered sequence but maintained pairing and base composition (e.g. U:A→A:U).
‡Compensatory transitions maintained pairing but altered sequence and base composition (e.g. U:A→C:G).

sequences were predicted to form WLE3-like structures in this context, with the mutations made producing only anticipated structural changes (not shown).

To begin, two general classes of mutations were made in the regions of predicted base-pairing. The first type alters both the sequence and base-pairing by transversion of bases on one side of the predicted double helix. The second class introduces compensatory transversions to bases on both sides of the helix, altering the sequence on both sides but preserving base-pairing and content (e.g. U:A→A:U). All six of the transversions made, which disrupt both sequence and base-pairing in portions of the predicted proximal and distal stems (constructs 24, 25, 37, 38, 55 and 56, Fig. 4A), completely abolished activity. However, compensatory transversions that reinstate the predicted base pairing and overall base composition revealed different requirements for activity in the three regions tested. For base-pair positions 1 and 2 (distal stem) all compensatory transversions (construct 37+38) restore complete activity. For the proximal stem (construct 24+25) moderate activity is restored. Strikingly though, compensatory transversions of base pair positions 3 to 5 of the distal stem (construct 55+56) yield absolutely no restoration of activity.

To define the sequence requirements of the distal stem more precisely, compensatory transversions were tested, alone or in combination, at base pair positions 3 to 5 (Fig. 4B). The presence of a single U:A→A:U base-pair switch at position 5 (construct 43) results in a strong loss of activity, and is enhanced to a complete loss of activity by additional U:A→A:U base-pair changes at positions 3 or 4 (constructs 46 and 48). The latter transversions, on their own, have no discernable effect on activity (constructs 44, 45 and 47). The sequence at base pair position 5 is therefore critical for WLE3 activity, and sensitive to sequence changes at base pair positions 3 and 4 of the distal stem.

The sequence requirements of the distal stem were probed further with compensatory transition mutations, which also maintain base pairing but alter base composition (e.g. U:A→C:G). Consistent with the results above, mutation of the U:A base pair at position 5 to a C:G (construct 52) reduces almost all activity, similar to the A:U

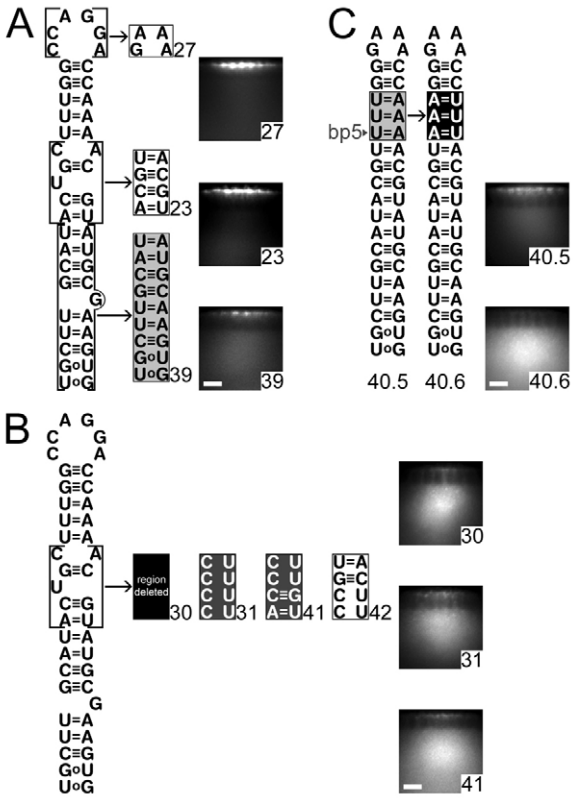

Fig. 5. Importance of the loop sequence, bulges and the variable region. (**A-C**) Diagrams of *D. melanogaster* WLE3 (A,B) or the minimal WLE3 mutant 40.5 (C). Mutant base changes and activities are depicted as in Fig. 4. Representative images of injected embryos are on the right. Scale bars: 10 μm.

substitution at the same position (Fig. 4C). Compensatory transition mutations at base-pair positions 3 or 4 (constructs 53 and 54) cause moderate reductions in activity, and mutations at base-pair positions 1 and 2 (construct 36) have almost no effect. To summarize, primary sequence is critical at position 5 (U:A required), and base-pair composition is moderately important at positions 3 and 4 (U:A or A:U permitted). Positions 1 and 2 of the distal stem require base-pairing, but sequence appears to be relatively unimportant. This is in contrast to the proximal stem, where secondary structure is the major determinant of activity with only a modest role for primary sequence.

Structural determinants of WLE3 activity: loop sequence and bulges

Exposed bases in loops and bulges can create unique structural features critical to the specificity of RNA-protein interactions (Hermann and Patel, 2000; Moras and Poterszman, 1996). Such features are prominent in the predicted WLE3 structure, and their importance in apical localization was tested. The terminal loop contains two invariant residues, C545 and G547, each of which was altered by a transition mutation (constructs 17 and 18). The entire WLE3 loop was also replaced with a GAAA tetraloop (construct 27), which forms a very stable structure at the end of a helix (Jucker et al., 1996). Surprisingly, the localization activity of all three mutants is normal (Fig. 5A, Table 5). Thus, despite the evolutionary conservation of loop size and invariant residues, these properties are inconsequential in this assay.

The WLE3 consensus contains a number of unpaired bulges that, although variable in terms of sequence, are conserved in position, suggesting a functional role. These include a single base bulge of variable identity within the proximal stem below base pairs 9 or 10 (Fig. 3C) and multiple bulges in the variable central part of the stem loop that vary in size, position and identity. These bulges were deleted alone or in combination to test their importance in localization (Fig. 5A). Somewhat surprisingly, conversion of the largely unpaired variable region to a bulge-free helix has no effect on localization (construct 23). However, deletion of the single-nucleotide proximal stem bulge does reduce the efficiency of localization (construct 39). When these two mutations are combined to create a bulge-free stem (construct 40), localization activity is similar to that of construct 39, indicating that its loss of activity is due mainly to removal of the proximal stem bulge.

Although removing bulges in the variable region, as in construct 23, had no effect on activity, removal of this region altogether (construct 30) abolishes all activity (Fig. 5B). Similarly, replacement with a large internal loop also strongly reduces activity (construct 31). This indicates that this region must be present and must include some base-pairing. This was somewhat surprising since *D. hydei* WLE3, which is active, is predicted to have a large internal loop. However, a potential C:G base pair within the *D. hydei* loop may stabilize the region sufficiently.

Smaller internal loops were also introduced into the variable region to further define the base pairing requirements. Although a small internal loop near the proximal stem has no effect on activity (construct 42), a small loop near the distal stem (construct 41) reduces activity by the same extent as the large internal loop (Fig. 5B). The observation that most *Drosophila* WLE3 elements contain a G:C base pair two positions below the critical distal helix U:A base-pair (position 5) suggests that a G:C base pair at this position (position b, Fig. 3C) is important.

To test whether the loops and bulges may have weak additive or redundant effects on WLE3 activity, a minimal WLE3 element free of bulges and containing a terminal GAAA tetraloop was synthesized (construct 40.5). This minimal element retains moderate localization activity (Fig. 5C). Notably, its activity is similar to that of the 'proximal stem' bulge deletion (construct 39, Fig. 5A, Table 5), suggesting that the weakened activity of this construct may again be primarily attributed to deletion of the 'proximal stem' bulge. Notably, the distal stem sequence remains a critical determinant in the context of this minimal WLE3 element, as compensatory transversions (U:A→A:U) at positions 3-5 abolish activity (construct 40.6, Fig. 5C). Thus, it appears that the localization apparatus can make sequence discriminations in the absence of nearby bulges. This is surprising, as the compact nature of double-stranded RNA, lacking bulges that distort the helix, generally precludes access of RNA binding proteins to the base sequence (Saenger, 1984; Seeman et al., 1976).

DISCUSSION

WLE3; a new apical localization motif

This study identifies a new apical localization element, WLE3, which functions in both a microinjection assay and in transgenic reporters. The sequence and predicted secondary structure of this element are highly conserved among *Drosophila* species, and were deemed necessary for WLE3 function by mutational analyses. Interestingly, full WLE3 activity is dependent on appropriate downstream sequences, and can also be blocked by downstream sequences. This dependence on context suggests that the assessment of other localization elements should also take context into account.

For example, activities of the *h* SL1 (stem loop 1) dimer and *K10* TLS in the microinjection assay were tested in the context of the 806 nucleotide *h* 3′UTR and a 2280 nucleotide *stg* reporter transcript, respectively (Bullock and Ish-Horowicz, 2001; Bullock et al., 2003). Importantly, neither the *h* SL1 nor the *K10* TLS has been tested in the absence of any flanking sequence whatsoever. Given the presence of multiple potentiating elements within the *wg* 3′UTR, we suggest that similar potentiators may also be present in other reporter constructs. These may recruit relatively general factors, used for other mRNA functions, that also contribute to the assembly or stabilization of fully functional localization signals. Alternatively, they may be required to block the effects of negatively acting structures or factors. Such interactions could provide a means of coupling processes such as translation and stability modulation to localization. A direct comparison of the activities of the HLE-SL1, *K10* TLS and WLE3 within identical contexts will be required to resolve whether the requirement for potentiating elements is exclusive to WLE3.

Determinants of WLE3 activity

Our mutagenesis data implicate several aspects of the WLE3 predicted secondary structure in its activity. Most important is the U:A base pair at position 5 of the distal helix. The two U:A base pairs at positions 4 and 3 above this are also important, but can be changed to A:U with little perceived effect. Also important are the G:C base pairs that flank the U:A tract of the distal helix. In the proximal helix, base pairing is essential, and sequence also plays a minor role. Our results also implicate the proximal stem, single nucleotide bulge as important for robust localization activity, although its sequence identity may be unimportant. This bulge may provide access to the proximal stem sequence. Alternatively, it may itself constitute a unique backbone geometry for protein interactions, or may present the proximal and distal stems at the correct angle for protein interactions (Hermann and Patel, 2000). The central variable region must be present, but its sequence and bulges appear to be unimportant in this assay. The prevalence of these central region bulges among *Drosophila* WLE3 sequences may reflect a role not detected in this assay – to prevent RNAi-based degradation, for example.

Conspicuously, the WLE3 loop sequence, with two invariant residues, is not required for localization. This was unexpected, as loop residues are common recognition sites for RNA binding proteins (Aviv et al., 2006; Cilley and Williamson, 2003; Stefl et al., 2006; Wu et al., 2004; Zanier et al., 2002). It cannot be ruled out that localization activity imposes constraints on the loop sequence that might have been revealed by further mutagenesis and analyses. Alternatively, the loop sequence may help to coordinate or discourage interactions with other RNA processing pathways.

Stem sequence recognition by the localization machinery

The sequence requirements for the WLE3 distal stem are somewhat surprising given that the major groove within RNA stems, where sequence recognition occurs, generally requires stem distortions such as bulges and internal loops to access the sequence information (Battiste et al., 1996; Hermann and Patel, 2000; Moras and Poterszman, 1996; Weeks and Crothers, 1993). However, structure prediction, conservation and mutagenesis all indicate an uninterrupted double helix. Hence, we speculate that recognition of the distal stem sequence requires local distortion. This might be achieved by an RNA helicase-type factor, similar to Vasa for example, which bends double-stranded RNA and forces local unwinding (Sengoku et al., 2006). This could also explain the requirement for weak A:U/U:A base pairs in this region. Notably, many RNA helicases have been implicated in RNA localization, although their specific roles remain unclear (Irion and Leptin, 1999; Palacios et al., 2004; Tinker et al., 1998). It is curious though, that the A/U-rich portion of the distal helix is flanked by conserved C:G/G:C base pairs, suggesting that these may be required to discourage full unwinding of the region. Alternatively, the distal helix of WLE3 may adopt an atypical helical conformation having recognizable backbone distortions or a directly accessible major groove as is found in double stranded DNA.

Shared features of apical localization signals in *Drosophila*

The wide array of transcripts known to localize apically in the embryo injection assay exhibit a common requirement for microtubules, Dynein, BicD and Egl (Bullock and Ish-Horowicz, 2001; Wilkie and Davis, 2001), suggesting that these similarities may extend to the apical localization elements themselves. Fig. 6 shows the predicted structures for each of these elements. In all cases where predicted structures have been analyzed by mutagenesis (*wg* WLE3, *h* HLE-SL1 and *K10* TLS), the double-stranded stem regions are indispensable (Bullock and Ish-Horowicz, 2001; Bullock et al., 2003; Cohen et al., 2005; Macdonald and Kerr, 1998; Serano and Cohen, 1995). Each of these stem-loops also contains a distal U:A-rich region that is bracketed by regions of increased stability. In the *wg*, *h* and *ftz* elements, this ‘bracket’ consists of strong

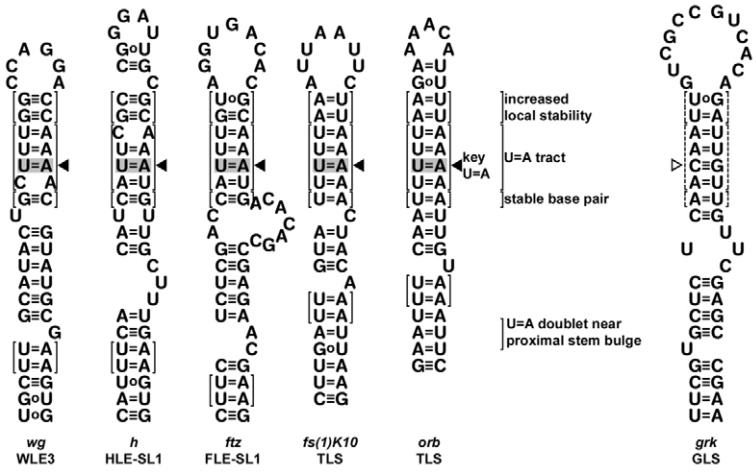

Fig. 6. Recognition of apical localization elements. Shared features of elements that mediate apical localization in the embryo injection assay. Gene and element names are indicated below each structure. Features shared between the *wg*, *h*, *ftz*, *K10*, and *orb* elements are bracketed. The conserved fifth U:A base pair is shaded in light grey. These shared features do not extend to the *grk* GLS.

G:C/C:G base pairs. In the *K10* and *orb* elements, similar increases in local stability may be effected by longer stems or favourable stacking interactions between U:A and adjacent A:U base pairs. Most notable though is the consistent presence of a U:A base pair in the vicinity of the essential fifth base pair position of WLE3. In the set of known elements, this U:A base pair is the third base pair of the U:A tract and is essential for the full activity of both the *K10* TLS (mutant rev5) and the *h* HLE-SL1 (mutant g15). The parallels between these motifs may also extend to their proximal stems, each of which is predicted to possess a bulge in the vicinity of two U:A base pairs.

Taking all of these observations together, it is possible to speculate on the existence of a stripped-down consensus motif common to many or most Dynein-mediated apical localization elements. This motif contains a stem-loop averaging ~16 base pairs in length with a distal U:A-rich tract bracketed by regions of increased local stability. The third U:A base pair in this tract, five positions removed from the terminal loop, is critical to activity. Additionally, the proximal stem contains a bulge with two U:A base pairs in close proximity (Fig. 6). Furthermore, we predict that recognition of this WLE3-like consensus element requires opening or distortion of the distal stem U:A-rich region to allow sequence recognition, and for full activity, recognition of the proximal helix bulge. Notably, this motif does not fit the *grk* GLS, which in the oocyte directs dorsoanterior Dynein-dependent localization. However, although Dynein-mediated, this movement is distinct from the anterior Dynein-dependent localization mediated by the *K10*, *orb* and *ftz* localization elements.

The identification and mutagenic analysis of additional Dynein-mediated localization elements should allow for further testing and refinement of the apical element consensus and its mode of action. In turn, this should aid in the identification of other localization elements in the large number of localized transcripts that are yet to be characterized, and will contribute to an understanding of the interactions between localization signals and corresponding transport and anchoring complexes.

We thank Craig Smibert, Rick Collins and Howard Lipshitz for comments on the manuscript. We also thank Chun Hu for establishing transgenic lines and Neela Parthasarathy for insights on common aspects of apical localization elements. G.d.-S. was supported by a Medical Research Council of Canada Studentship, H.M.K. by a Canadian Institutes of Health Research Senior Scientist award and the project by a grant provided by the National Cancer Institute of Canada with funds from the Cancer Research Society of Canada.

References

Aviv, T., Lin, Z., Ben-Ari, G., Smibert, C. A. and Sicheri, F. (2006). Sequence-specific recognition of RNA hairpins by the SAM domain of Vts1p. *Nat. Struct. Mol. Biol.* **13**, 168-176.

Ballinger, D. G. and Benzer, S. (1989). Targeted gene mutations in Drosophila. *Proc. Natl. Acad. Sci. USA* **86**, 9402-9406.

Bashirullah, A., Halsell, S. R., Cooperstock, R. L., Kloc, M., Karaiskakis, A., Fisher, W. W., Fu, W., Hamilton, J. K., Etkin, L. D. and Lipshitz, H. D. (1999). Joint action of two RNA degradation pathways controls the timing of maternal transcript elimination at the midblastula transition in Drosophila melanogaster. *EMBO J.* **18**, 2610-2620.

Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Muhandiram, D. R., Kay, L. E., Frankel, A. D. and Williamson, J. R. (1996). Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **273**, 1547-1551.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H. and Long, R. M. (1998). Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* **2**, 437-445.

Betley, J. N., Heinrich, B., Vernos, I., Sardet, C., Prodon, F. and Deshler, J. O. (2004). Kinesin II mediates Vg1 mRNA transport in Xenopus oocytes. *Curr. Biol.* **14**, 219-224.

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-415.

Brendza, R. P., Serbus, L. R., Duffy, J. B. and Saxton, W. M. (2000). A function for kinesin I in the posterior transport of oskar mRNA and Staufen protein. *Science* **289**, 2120-2122.

Brenner, H. R., Witzemann, V. and Sakmann, B. (1990). Imprinting of acetylcholine receptor messenger RNA accumulation in mammalian neuromuscular synapses. *Nature* **344**, 544-547.

Brower, A. V. and DeSalle, R. (1998). Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference. *Insect Mol. Biol.* **7**, 73-82.

Bullock, S. L. and Ish-Horowicz, D. (2001). Conserved signals and machinery for RNA transport in Drosophila oogenesis and embryogenesis. *Nature* **414**, 611-616.

Bullock, S. L., Zicha, D. and Ish-Horowicz, D. (2003). The Drosophila hairy RNA localization signal modulates the kinetics of cytoplasmic mRNA transport. *EMBO J.* **22**, 2484-2494.

Bullock, S. L., Stauber, M., Prell, A., Hughes, J. R., Ish-Horowicz, D. and Schmidt-Ott, U. (2004). Differential cytoplasmic mRNA localisation adjusts pair-rule transcription factor activity to cytoarchitecture in dipteran evolution. *Development* **131**, 4251-4261.

Cha, B. J., Koppetsch, B. S. and Theurkauf, W. E. (2001). In vivo analysis of Drosophila bicoid mRNA localization reveals a novel microtubule-dependent axis specification pathway. *Cell* **106**, 35-46.

Cilley, C. D. and Williamson, J. R. (2003). Structural mimicry in the phage phi21 N peptide-boxB RNA complex. *RNA* **9**, 663-676.

Cohen, R. S., Zhang, S. and Dollar, G. L. (2005). The positional, structural, and sequence requirements of the Drosophila TLS RNA localization element. *RNA* **11**, 1017-1029.

Davis, I. and Ish-Horowicz, D. (1991). Apical localization of pair-rule transcripts requires 3′ sequences and limits protein diffusion in the Drosophila blastoderm embryo. *Cell* **67**, 927-940.

Delanoue, R. and Davis, I. (2005). Dynein anchors its mRNA cargo after apical transport in the Drosophila blastoderm embryo. *Cell* **122**, 97-106.

Forrest, K. M. and Gavis, E. R. (2003). Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in Drosophila. *Curr. Biol.* **13**, 1159-1168.

Glotzer, J. B., Saffrich, R., Glotzer, M. and Ephrussi, A. (1997). Cytoplasmic flows localize injected oskar RNA in Drosophila oocytes. *Curr. Biol.* **7**, 326-337.

Hermann, T. and Patel, D. J. (2000). RNA bulges as architectural and recognition motifs. *Structure* **8**, R47-R54.

Hofacker, I. L., Fekete, M. and Stadler, P. F. (2002). Secondary structure prediction for aligned RNA sequences. *J. Mol. Biol.* **319**, 1059-1066.

Hoogenraad, C. C., Akhmanova, A., Howell, S. A., Dortland, B. R., De Zeeuw, C. I., Willemsen, R., Visser, P., Grosveld, F. and Galjart, N. (2001). Mammalian Golgi-associated Bicaudal-D2 functions in the dynein-dynactin pathway by interacting with these complexes. *EMBO J.* **20**, 4041-4054.

Hughes, J. R., Bullock, S. L. and Ish-Horowicz, D. (2004). Inscuteable mRNA localization is dynein-dependent and regulates apicobasal polarity and spindle length in Drosophila neuroblasts. *Curr. Biol.* **14**, 1950-1956.

Hughes, S. C. and Krause, H. M. (1999). Single and double FISH protocols for Drosophila. *Methods Mol. Biol.* **122**, 93-101.

Irion, U. and Leptin, M. (1999). Developmental and cell biological functions of the Drosophila DEAD-box protein abstrakt. *Curr. Biol.* **9**, 1373-1381.

Jambhekar, A., McDermott, K., Sorber, K., Shepard, K. A., Vale, R. D., Takizawa, P. A. and DeRisi, J. L. (2005). Unbiased selection of localization elements reveals cis-acting determinants of mRNA bud localization in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA* **102**, 18005-18010.

Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H. and Pardi, A. (1996). A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J. Mol. Biol.* **264**, 968-980.

Lall, S., Francis-Lang, H., Flament, A., Norvell, A., Schüpbach, T. and Ish-Horowicz, D. (1999). Squid hnRNP protein promotes apical cytoplasmic transport and localization of Drosophila pair-rule transcripts. *Cell* **98**, 171-180.

Latham, V. M., Yu, E. H., Tullio, A. N., Adelstein, R. S. and Singer, R. H. (2001). A Rho-dependent signaling pathway operating through myosin localizes beta-actin mRNA in fibroblasts. *Curr. Biol.* **11**, 1010-1016.

Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T. R., Tomancak, P. and Krause, H. M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**, 174-187.

MacDonald, P. M. (1990). bicoid mRNA localization signal: phylogenetic conservation of function and RNA secondary structure. *Development* **110**, 161-171.

Macdonald, P. M. and Struhl, G. (1988). cis-acting sequences responsible for anterior localization of bicoid mRNA in Drosophila embryos. *Nature* **336**, 595-598.

Macdonald, P. M. and Kerr, K. (1997). Redundant RNA recognition events in bicoid mRNA localization. *RNA* **3**, 1413-1420.

Macdonald, P. M. and Kerr, K. (1998). Mutational analysis of an RNA recognition element that mediates localization of bicoid mRNA. *Mol. Cell. Biol* **18**, 3788-3795.

MacDougall, N., Clark, A., MacDougall, E. and Davis, I. (2003). Drosophila gurken (TGFalpha) mRNA localizes as particles that move within the oocyte in two dynein-dependent steps. *Dev. Cell* **4**, 307-319.
Mach, J. M. and Lehmann, R. (1997). An Egalitarian-BicaudalD complex is essential for oocyte specification and axis determination in Drosophila. *Genes Dev.* **11**, 423-435.
Matanis, T., Akhmanova, A., Wulf, P., Del Nery, E., Weide, T., Stepanova, T., Galjart, N., Grosveld, F., Goud, B., De Zeeuw, C. I. et al. (2002). Bicaudal-D regulates COPI-independent Golgi-ER transport by recruiting the dynein-dynactin motor complex. *Nat. Cell Biol.* **4**, 986-992.
Mathews, D. H., Sabina, J., Zuker, M. and Turner, D. H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911-940.
Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M. and Turner, D. H. (2004). Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. USA* **101**, 7287-7292.
Moras, D. and Poterszman, A. (1996). Getting into the major groove. Protein-RNA interactions. *Curr. Biol.* **6**, 530-532.
Palacios, I. M., Gatfield, D., St Johnston, D. and Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature* **427**, 753-757.
Robertson, H. M., Preston, C. R., Phillis, R. W., Johnson-Schlitz, D. M., Benz, W. K. and Engels, W. R. (1988). A stable genomic source of P element transposase in Drosophila melanogaster. *Genetics* **118**, 461-470.
Rubin, G. M. and Spradling, A. C. (1982). Genetic transformation of Drosophila with transposable element vectors. *Science* **218**, 348-353.
Saenger, W. (1984). *Principles of Nucleic Acids, Advanced Texts in Chemistry.* New York: Springer-Verlag.
Seeman, N. C., Rosenberg, J. M. and Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. USA* **73**, 804-808.
Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S. and Yokoyama, S. (2006). Structural basis for RNA unwinding by the DEAD-box protein Drosophila vasa. *Cell* **125**, 287-300.
Serano, T. L. and Cohen, R. S. (1995). A small predicted stem-loop structure mediates oocyte localization of *Drosophila K10* mRNA. *Development* **121**, 3809-3818.
Shepard, K. A., Gerber, A. P., Jambhekar, A., Takizawa, P. A., Brown, P. O., Herschlag, D., DeRisi, J. L. and Vale, R. D. (2003). Widespread cytoplasmic mRNA transport in yeast: identification of 22 bud-localized transcripts using DNA microarray analysis. *Proc. Natl. Acad. Sci. USA* **100**, 11429-11434.
Simmonds, A. J., dos Santos, G., Livne-Bar, I. and Krause, H. M. (2001). Apical localization of wingless transcripts is required for wingless signaling. *Cell* **105**, 197-207.
Snee, M. J., Arn, E. A., Bullock, S. L. and Macdonald, P. M. (2005). Recognition of the bcd mRNA localization signal in Drosophila embryos and ovaries. *Mol. Cell. Biol.* **25**, 1501-1510.
Stefl, R., Xu, M., Skrisovska, L., Emeson, R. B. and Allain, F. H. (2006). Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. *Structure* **14**, 345-355.
St Johnston, D. (2005). Moving messages: the intracellular localization of mRNAs. *Nat. Rev. Mol. Cell Biol.* **6**, 363-375.
Thompson, J. D., Higgins, D. G. and Gibson, T. J. (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.
Tinker, R., Silver, D. and Montell, D. J. (1998). Requirement for the vasa RNA helicase in gurken mRNA localization. *Dev. Biol.* **199**, 1-10.
Van De Bor, V., Hartswood, E., Jones, C., Finnegan, D. and Davis, I. (2005). gurken and the I factor retrotransposon RNAs share common localization signals and machinery. *Dev. Cell* **9**, 51-62.
Weeks, K. M. and Crothers, D. M. (1993). Major groove accessibility of RNA. *Science* **261**, 1574-1577.
Wilkie, G. S. and Davis, I. (2001). Drosophila wingless and pair-rule transcripts localize apically by dynein-mediated transport of RNA particles. *Cell* **105**, 209-219.
Wu, H., Henras, A., Chanfreau, G. and Feigon, J. (2004). Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc. Natl. Acad. Sci. USA* **101**, 8307-8312.
Zanier, K., Luyten, I., Crombie, C., Muller, B., Schumperli, D., Linge, J. P., Nilges, M. and Sattler, M. (2002). Structure of the histone mRNA hairpin required for cell cycle regulation of histone gene expression. *RNA* **8**, 29-46.