Translational Control of Serrate Expression in Drosophila Cells

GEORGIA DELICONSTANTINOS[1], KONSTANTINA KALODIMOU[1,2] and CHRISTOS DELIDAKIS[1,2]

[1]*Institute of Molecular Biology and Biotechnology, F.O.R.T.H., Heraklion, Greece;*
[2]*Department of Biology, University of Crete Heraklion, Heraklion, Greece*

Abstract. *Background/Aim: The DSL proteins, Serrate and Delta, which act as Notch receptor ligands, mediate signalling between adjacent cells, when a ligand-expressing cell binds to Notch on an adjacent receiving cell. Notch is ubiquitously expressed and DSL protein mis-expression can have devastating developmental consequences. Although transcriptional regulation of Delta and Serrate has been amply documented, we examined whether they are also regulated at the level of translation. Materials and Methods: We generated a series of deletions to investigate the initiation codon usage for Serrate using Drosophila S2 cells. Results: Serrate mRNA contains three putative ATG initiation codons spanning a 60-codon region upstream of its signal peptide; we found that each one can act as an initiation codon, however, with a different translational efficiency. Conclusion: Serrate expression is strictly regulated at the translational level.*

The two Drosophila Notch ligands, the DSL proteins Serrate and Delta, are single-pass type I trans-membrane proteins. A ligand-expressing cell binds to Notch on an adjacent cell, mediating an active signal between the two cells (1-3). The highly-conserved cell to cell Notch signalling pathway is essential in multiple developmental processes (4), such as stem cell renewal and maintenance (5) and specification of differentiated cell types (6).

DSL endocytosis is thought to exert a pulling force on a bound Notch receptor, which triggers two consecutive proteolytic cleavages, and as a consequence the Notch intracellular domain (NICD) is released from the plasma membrane (7-9). The NICD translocates into the nucleus and

acts as a transcriptional activator together with a CSL protein (CBF1/RBPjκ/Su(H)/Lag-1). The CSL/NICD complex recruits co-activators, including Mastermind, to drive Notch-dependent gene expression (9-12). The Serrate and Delta ligands have the capacity to act as both agonists and antagonists of the Notch receptor (13-17), depending on whether they are presented to an adjacent cell (trans-activation) or on the Notch-expressing cell (cis-inhibition) (18). Consequently, the exact time, place and levels of DSL protein expression are crucial in mediating the fine-tuning of Notch signalling (19). Serrate consists of an extracellular domain with 14 EGF-like repeats and an N-terminal region containing the conserved MNNL and DSL motifs, necessary for binding to Notch (20, 21), and an intracellular domain containing motifs that promote ubiquitylation by Mindbomb1 and Neuralized (RING domain E3 ubiquitin ligases), which is thought to promote its signalling activity by triggering its endocytosis (2, 22-25).

Serrate has an unusually long 5' UTR of 431 bases (Drosophila average 5' UTR length is 223 bases) (26, 27), containing 3 ATG triplets (Figure 1A). Likewise, *Delta* 5' UTR comprises 679 bp which includes 6 ATG triplets (Figure 1B). 5' UTRs can regulate translation through the action of upstream open reading frames (uORFs) (28). uORFs are regions in the 5' UTR containing an AUG (start) and a stop codon, with no overlap with the coding sequence (CDS) (28). Similarly, regions that start at an AUG prior to the CDS but finish within the CDS are termed overlapping open reading frames (oORFs) and can also act as potent negative regulators of translation (28). uORFS in several systems have been shown to be actively translated by means of high-resolution ribosome footprinting (29). Translational repression by uORFs or oORFs is as ubiquitous as *trans*-acting RNA binding proteins that selectively target defined subsets of mRNAs, such as the miRNA-guided RISC complex (30). Moreover, the extent of negative regulation of translation by uORFs can be as high as that exercised by miRNAs. The best-studied example of uORF repression activity is the stress response gene *GCN4* in yeast, in which uORFs repress translation of the CDS, unless the cell is stressed by amino acid starvation (31).

A Serrate 5' UTR and coding sequence:

```
-453                                                    CC

-451  CCCCCCCCCC CCCCCCCCCG AGTCGAGCGC CGTGCTTCGA GCGGTGATGA

-401  GCCCCTTTTC TGTCAACGCT AAAGATCTAC AAAACATCAG CGCCTATCAA

-351  GTGGAAGTGT CAAGTGTGAA CAAAACAAAA ACGAGAGAAG CACATACTAA

-301  GGTCCATATA AATAATAAAT AATAATTGTG TGTGATAACA ACATTATCCA

-251  AACAAAACCA AACAAAACGA AGGCAAAGTG GAGAAAATGA TACAGCATCC

-201  AGAGTACGGC CAGTTATTCA GCTATCCAGA GCAAGTGTAG TGTGGCAAAA

-151  TAGAAACAAA CAAAGGCACC AAAATCTGCA TACATGGGCT AATTAAGGCT

-101  GCCCAGCGAA TTTACATTTG TGTGGTGCCA ATCCAGAGTG AATCCGAAAC

- 51  AAACTCCATC TAGATCGCCA ACCAGCATCA CGCTCGCAAA CGCCCCCAGA

001   ATGTACAAAA TGTTTAGGAA ACATTTTCGG CGAAAACCAG CTACGTCGTC
       M   Y   K   M   F   R   K   H   F   R   R   K   P   A   T   S   S

051   GTCGTTGGAG TCAACAATAG AATCAGCAGA CAGCCTGGGA ATGTCCAAGA
           S   L   E   S   T   I   E   S   A   D   S   L   G   M   S   K

101   AGACGGCGAC AAAAAGGCAG CGTCCGAGGC ATCGGGTACC CAAAATCGCG
           K   T   A   T   K   R   Q   R   P   R   H   R   V   P   K   I   A

151   ACCCTGCCAT CGACGATCCG CGATTGTCGA TCATTAAAGT CTGCCTGCAA
           T   L   P   S   T   I   R   D   C   R   S   L   K   S   A   C   N

201   CTTAATTGCT TTAATTTTAA TACTGTTAGT CCATAAGATA TCCGCAGCTG
           L   I   A   L   I   L   I   L   L   V   H   K   I   S   A   A

251   GTAACTTC
           G   N   F
```
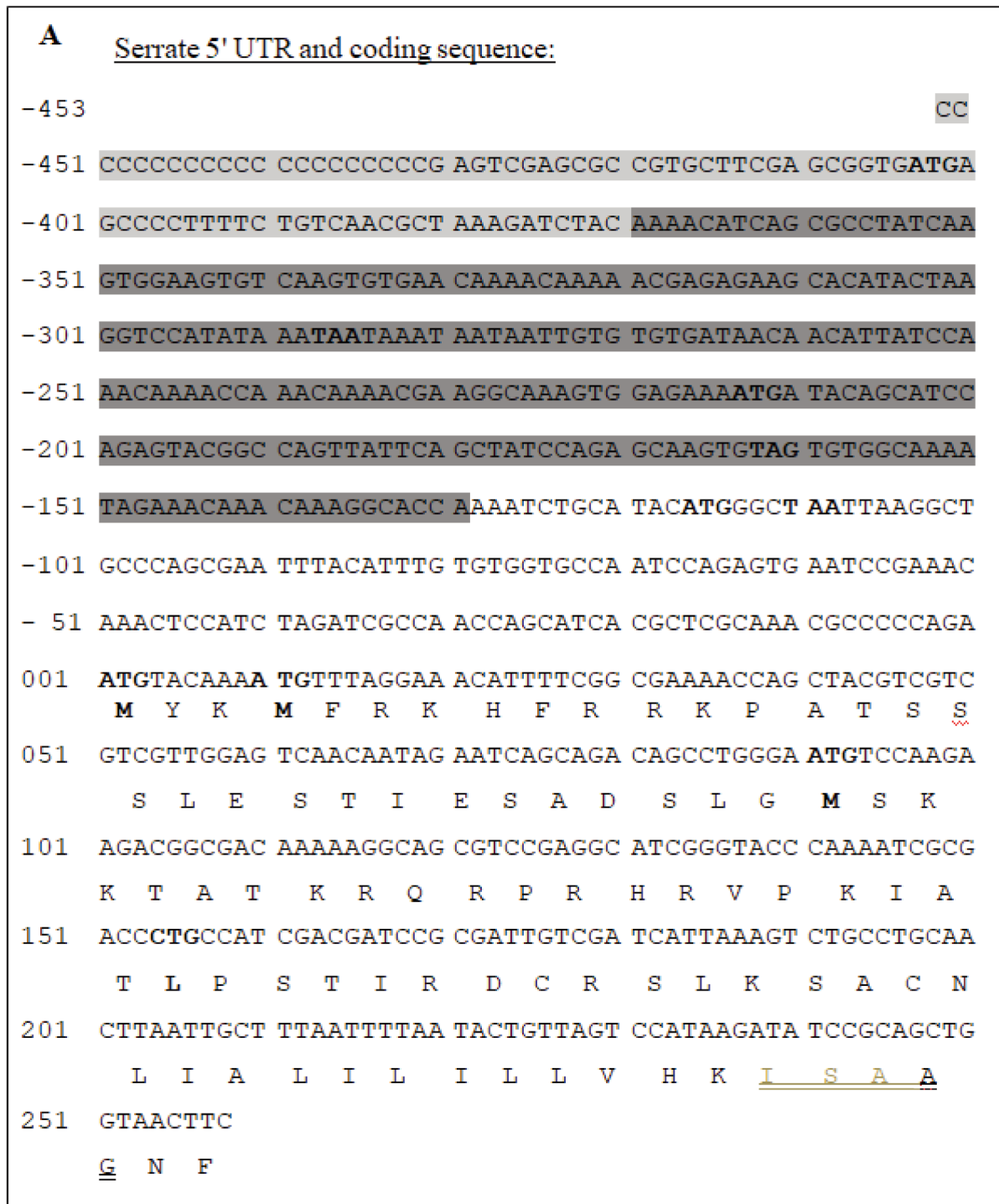
Three in-frame ATGs are located within 93 nt of the 5' end of the *Serrate* CDS. The first two are closely spaced and were originally predicted as the most likely translation initiation codons (32), especially the second, which is located within a sequence context that conforms with the *Drosophila* Kozak consensus sequence *(C/A)AA(A/C)AUG* (33). We asked which ATG is used for initiation of translation. We found that all three 5' proximal ATG initiation codons can

B

Delta5' – UTR:

```
-679                          CGGCATTAC GTTTTTCAAC TCGGCCGCGC

-651 GAATAACTCG AAAAAGGCAG CTCCGCAGCC AAGGCATTTT GAAAAACACA

-601 AGTCCCCGAC TCGAAACGCG ACCAAATATT CGGTGTGTGA CGCGAACTGC

-551 CAATGCAATA GTTCACTTAA GAATTGCAGA TTACCGCGAC TCTGGGCAGT

-501 TCTCATTCGA TATTTGAATG TACCAAAAGA AAAGTGCCAG AACCAGAAAT

-451 CAAAATAAAA GATCTTCTAA CAGAATAACA AGAAGTGTTT CCTCCGAAAG

-401 ATTAAAAAAT CGCGAATGAT TAAGAATCGC GGCACCGTTA GTTCCCTCTC

-351 TCGCTTTTCC CTTTTGCGCT TTTCTGCGTG AAAAGGAGAA GACTCCGCAG

-301 TTCGCGAAAA GAGCTTTTAT TTCGCGACAC TTAAAAATAA ACACAGTCGA

-251 GGAAATTGCA ACAAAAAATG TGGCGCGTGC ACGTGTAGCT TGCAACAACG

-201 ACAACAAGAA AACAACGAAA ATGCTAAAAG CGAATTGCAA AACCATGTGC

-151 GACAAGTAAC CAAAAAAAAA ATTCAAAACA TAAACACAAT AAACAATTTG

-101 AGTAGTTGCC GCACACACAC ACACACACAC ACAGCCCGTG GATTATTACA

- 51 CTAAAAGCGA CACTCAATCC AAAAAATCAG CAACAAAAAC ATCAATAAAC

001 ATGCATTGGA TTAAATGTTT ATTAACAGCA TTCATTTGCT TCACAGTCAT
     M   H   W   I   K   C   L   L   T   A   F   I   C   F   T   V   I

051 CGTGCAGGTTCACAGTTCCGGCAGCTTTGAG
     V   Q   V   H   S   S   G   S   F   E
```

Figure 1. *Nucleotide sequence of the 5' UTR and the first 258 bp of Serrate CDS. (A) The nucleic acid sequence is numbered relative to the first base of the coding region (M1). Within the Serrate 5' UTR, the three putative uORFs are marked with bold initiation codons (at - 405 bp, - 215 bp, - 118 bp) and termination codons. The deletion of the first 61 bp (5' UTR-Δ61) is shown in light grey; the deletion of the next 241 bp (5' UTR-Δ302) is shown in dark grey. The three-putative initiator ATGs of Serrate are in bold, the second ATG codon (M2) is located within a sequence context that perfectly matches the Drosophila Kozak consensus sequence (italics). The potential CTG alternative initiation codon is also in bold (154 nt). The predicted signal peptide is shown in grey. The putative cleavage site within the signal peptide is underlined (double line). Note the artificial 21 nt poly-C stretch included in the beginning of the sequence. (B) The Delta 5' UTR which comprises 679 bp, including 6 uORFs and the first 81 bp of the Delta CDS. Initiation codons and signal peptide cleavage site are marked as in A. The first 5 residues (SGSFE) of mature Delta are included.*

be used for translation initiation with different efficiency. We also entertained the possibility that a CTG (Leu) codon at position 154 (from the most upstream ATG of the CDS) may be the initiation codon, as it has been proposed that CUG can act as an alternative initiation codon in animal cells (34). However, this did not seem to be the case for *Serrate*, since deletion of all three in-frame ATG's abolished translation. We also hypothesized that the ATG triplets in the 5' UTR of *Serrate* – and *Delta* – are putative starts of uORFs and investigated whether these long 5' UTRs negatively regulate translation. We found that removal of a string of 21 cytosines in the *Serrate* 5' UTR, significantly increased the expression of Serrate but removing the uORFs did not have an appreciable effect. Moreover, the *Delta* 5' UTR had a similar repressive effect when used in place of the *Serrate* 5' UTR preceded by a string of 21 cytosines.

Materials and Methods

Construction of Serrate expressing vectors. The pMT – 5' UTR – Ser – 3' UTR – V5.His vector was kindly provided by Dr. Sarah Bray (University of Cambridge, Cambridge, UK) (2). In addition to the *Serrate* CDS, the cDNA insert contained the RA transcript 5' UTR comprising 434 bp preceded by a string of 21 cytosines (Figure 1A) and the 3' UTR of *Serrate*. The cDNA had been cloned as an EcoRI-XhoI fragment downstream of a 427 bp fragment containing the metal inducible *MtnA* promoter, which includes 56 bp of the 5' UTR of the *MtnA* gene, as well as 45 bp belonging to the polylinker region of the plasmid. It was confirmed that no ATGs were present in these transcribed sequences.

pMT – Ser – V5.His. This is a deletion of the *Serrate* 3' UTR designed to fuse the vector V5.His epitope in frame with the 3' of the *Serrate* CDS. The last 595 bp of the *Serrate* CDS were amplified using primers: Forward: 5'-CCT CCA TTG TGG AGG TCA AGT TGG AAA-3' and Reverse: 5'-AAG TCT CGA GAA CCA TCA CAG TGG TGG-3', using the pMT – 5' UTR – *Serrate* – 3' UTR– V5.His vector as a template. The 3' UTR was excised by substituting the SfiI – XhoI 1524 bp fragment of the vector with the SfiI – XhoI digested PCR fragment. SfiI cuts within the *Serrate* CDS, whereas XhoI is a vector site abutting the V5-His tag.

pMT – Δ431 – Ser – V5.His. A pUAST–attB–Ser–HA plasmid was kindly provided by Thomas Klein (University of Düsseldorf, Düsseldorf, Germany). The *Serrate* CDS was PCR amplified from this plasmid using primers: Forward: 5'-GGG GGG AAT TCA TGT ACA AAA TGT TTA GGA-3' and Reverse: 5'-CCT CCT CGA GAA CCA TCA CAG TGG TG-3'. The PCR product was digested with EcoRI and XhoI and used to substitute the *Serrate* cDNA in the pMT – Ser – V5.His vector, thus bringing the MtnA promoter (EcoRI site) directly next to the first *Serrate* ATG codon.

Nested deletions of the Serrate 5' UTR. These were made by substituting an EcoRI – AvrII fragment in the *Serrate* 5' UTR of the pMT – Ser – V5.His vector with the following PCR fragments produced from a pMT – Ser – V5.His template. The six PCR products described below were produced using the same reverse primer, which maps 48 bp downstream of the *Serrate* AvrII site.

pMT – Δ61 – Ser – V5.His; 1803 bp of the *Serrate* ORF (comprising 370 bp of the 5' UTR and 1433 bp of the ORF, up to 48 bp downstream from the AvrII site) were amplified using primers: Forward: 5'-GGC GAA TTC AAA ACA TCA GCG-3' and Reverse: 5'-CAG TCG CAA GTG AAG TCG GG-3'. pMT – Δ302 –Ser – V5.His; 1562 bp of the *Serrate* ORF (comprising 129 bp of the 5' UTR and 1433 bp of the ORF) were amplified using primers: Forward: 5'-GGC GAA TTC AAA TCT GCA TAC ATG G-3' and Reverse: 5'-CAG TCG CAA GTG AAG TCG GG-3'. pMT – Δ[431M1] – M2M3LSer –V5.His; 1424 bp of the *Serrate* ORF (comprising the part of the ORF from the second ATG codon, up to 48 bp downstream from the AvrII site) were amplified using primers: Forward: 5'-GGC GAA TTC ATG TTT AGG AAA CAT TTT CG-3' and Reverse: 5'-CAG TCG CAA GTG AAG TCG GG-3'. pMT – Δ[431M2] – M3LSer – V5.His; 1343 bp of the *Serrate* ORF (comprising the part of the ORF from the third ATG codon, up to 48 bp downstream from the AvrII site) were amplified using primers: Forward: 5'-GGC GAA TTC ATG TCC AAG AAG ACG CG-3' and Reverse: 5'-CAG TCG CAA GTG AAG TCG GG-3'. pMT – ΔCC – Ser – V5.His; 1863 bp of the *Serrate* ORF (comprising the 434 bp of the 5' UTR and 1433 bp of the ORF) were amplified using primers: Forward: 5'-GGC GAA TTC AGT CGA GCG CC-3' and Reverse: 5'-CAG TCG CAA GTG AAG TCG GG-3'. pMT – Δ[431M3] – LSer – V5.His; finally, the 1280 bp of the *Serrate* ORF (comprising the part of the ORF from the CTG codon located 165 bp downstream of the first ATG codon, up to 48 bp downstream from the AvrII site) was amplified using primers: Forward: 5'-ATA GAA TTC CTG CCA TCG ACG ATC CG-3' and Reverse: 5'-CAG TCG CAA GTG AAG TCG GG-3'.

PCR product digestion and ligation into vector. The PCR products from the above six reactions were doubly digested with EcoRI and AvrII and then ligated into the pMT – Ser – V5.His vector backbone with complementary sites, *i.e.* digested with the above restriction enzymes and 1809 bp of *Serrate* 5' UTR and part of the CDS excised.

Site directed mutagenesis. We used the QuikChange II Site-Directed Mutagenesis Kit (cat. no: 200522; Agilent, Santa Clara, CA, USA) to generate the constructs shown in Table I.

pMT – Δ431/Dl – Ser – V5.His. The 679 bp of the *Delta* 5' UTR (sequence FBtr0083739) was amplified using primers: Forward: 5'-ATA CCT AGG CGG CAT TAC GTT TTT CAA-3' and Reverse: 5'-GGC GAA TTC GTT TAT TGA TGT TTT TTG TTG C-3'. The PCR product doubly digested with AvrII and EcoRI was ligated into the complementary pMT – Δ431 – Ser – V5.His vector backbone, doubly digested with SpeI and EcoRI (both in the polylinker region downstream of the *MtnA* promoter/ 5' UTR).

All constructs were sequence-verified by Macrogen using the standardized MT Forward primer (Macrogen, Seoul, South Korea) for the sequencing of pMT – V5.His constructs and, furthermore 4 forward and 5 reverse primers were designed for the sequencing of overlapping regions of ~700 to ~1000 bp of the entire *Serrate* 5' UTR and CDS.

Cell line maintenance and transient transfections. S2-DGRC cells were maintained between $1\text{-}10\times10^6$ cells/ml in M3 insect medium (cat. no: 8398; Sigma-Aldrich, Saint Louis, MO, USA),

Table I. *Constructs created using site-directed mutagenesis.*

Template	Primer sequence	Name of the construct generated Type of conversion
pMT – Ser – V5.His	5'-CCC AAA ATC GCG ACC CTA GGA TCG ACG ATC CGC G-3'	pMT – M1M2M3Ser – V5.His CTG > to CCT (Leu>Pro)
pMT – Δ431 – Ser – V5.His	5'-GGT GAA TTC ATG TAC AAA CCT AGG AGG AAA CAT TTT CGG-3'	pMT – Δ431 – M1M3LSer –V5.His second ATG codon to CCT (Met>Pro) and the subsequent codon TTT to AGG (Phe>Arg)

supplemented with 10% FBS (cat. no: 10270; GIBCO, Thermo Fisher Scientific, Waltham, MA, USA) at 25˚C. For the transfections, we used the calcium phosphate precipitation method. For each sample, 8×10^5 cells were used (1 ml/well/12 well plate). Twenty-four hours later, the medium was supplemented with 0.7 mM $CuSO_4$ for 16 h and the cells were harvested 40 h post-transfection. For each transfection, the following amounts of plasmids were used: 800 ng of one of the *Serrate* expressing vectors described above and 400 ng of Ract – GFP plasmid, a constitutively expressing GFP plasmid for transfection normalization.

Total protein extracts and western blotting. The cells were lysed by freeze-thaw in 100 μl of NP-40 lysis buffer (150 mM NaCl, 50 mM Tris pH8, 1% NP-40, 1 mM PMSF, 1 μg/ml aprotinin, 10 μg/ml leupeptin, 1 μg/ml pepstatin). Samples were centrifuged at 13,000 rpm for 30 min at 4˚C. The protein samples supplemented with Laemmli buffer were incubated at 50˚C for 2 h. SDS-PAGE and transfer of the proteins were performed according to standard protocols. All western blots were performed by loading an equivalent of approximately 110,000 cells per well into an 8% (w/v) SDS polyacrylamide gel.

Antibodies used. 1:10,000 mouse anti-V5 (cat. no: R960-25; Invitrogen, Carlsbad, CA, USA), 1:100,000 rabbit anti-GFP (in-house antibody; Minotech, Heraklion, Crete, Greece), 1:40,000 donkey anti-mouse-HRP (cat. no: 715-035-150; Jackson ImmunoResearch, West Grove, PA, USA) and 1:40,000 donkey anti-rabbit-HRP (cat. no: 711-035-152; Jackson ImmunoResearch).

Sequence comparisons. cDNA sequences of *Serrate* or *Delta* were compared between *D. melanogaster* and *D. virilis*. The "Stretcher" nucleotide pair-wise alignment tool was used (35). For *Serrate* we used the *D. melanogaster* Ser-RA sequence and the *D. virilis* GJ23176-RA orthologue. For *Delta* we used the *D. melanogaster* Dl-RA sequence and the *D. virilis* GJ24543-RA orthologue. The full-length cDNAs were aligned pair-wise in each case, but only their 5' ends are shown in Figure 2.

Results

The M1, M2 and M3 ATGs of the Serrate CDS are interchangeable, efficient initiators of translation. To address the translational regulation of *Serrate* we transiently transfected metal-inducible constructs (with the *Drosophila* metallothionein *MtnA* promoter) containing a *Serrate* CDS incorporating various manipulations of its 5' sequence. The

pMT – V5.His vector (Invitrogen) was used to generate expression constructs that express full-length Serrate C-terminally tagged with the V5 epitope and the protein yielded in the transfected *Drosophila* S2 cells was detected by western blotting using a mouse anti-V5 antibody. All samples were treated identically in terms of transfection, Cu^{2+} induction of the *MtnA* promoter, as well as harvesting and lysis. Nevertheless, to normalize for transfection efficiency and pipetting errors, a Ract-myc-GFP expression construct was co-transfected which constitutively expresses GFP under the *Act5C* promoter (36) (Figure 3, lower panel). In the absence of the entire *Serrate* 5' UTR (Δ431), robust protein levels were obtained (Figure 3, upper panel). Two bands were detected, a major band at 245 kDa and a minor one at 150 kDa. The predicted MW of Serrate is 150 kDa; hence, the higher MW band is likely to result from extensive post-translational modifications (mostly glycosylation). A set of three deletions, as well as two point-mutations, were generated to assess the use of the three ATGs, found within 93 bp at the beginning of the *Serrate* CDS, as putative start codons. We deleted the 5' UTR as well as the first (M1) ATG codon of the pMT – Ser –V5.His expression construct, generating the pMT – Δ[431M1] – M2M3LSer – V5.His expression construct. The latter expression construct yielded robust levels of Serrate, showing that the M2 and M3 methionine codons are sufficient for translation initiation, in the absence of M1 (Figure 3, upper panel). Next, we deleted the upstream region in the CDS up to M3, to construct pMT – Δ[431M2] – M3LSer – V5.His. This deletion, proved M3 to be another initiation codon, albeit not equally efficient as M1 and M2, as reduced protein levels were produced (Figure 3, upper panel). Site-directed mutagenesis was used on the pMT – Δ431 – Ser – V5.His expression construct to convert the M2 ATG codon to CCT, thus generating the pMT – Δ431 – M1M3LSer – V5.His expression construct (comprising only the M1 and M3 ATG codons). This construct produced Serrate protein at levels comparable to the pMT – Δ[431M1] – M2M3LSer – V5.His expression construct, suggesting that the M1 and M3 ATG codons are sufficient to yield high levels of translation, even though the perfect Kozak consensus is not found in their vicinity, as for the M2 ATG codon.

Figure 2. *An alignment of the nucleotide sequence of the 5' UTR and the first 194 bps of Serrate CDS in D. melanogaster and D. virilis. (A) Putative uORFs are underlined. Their respective start and stop codons are highlighted in grey. The three-putative initiator ATGs of Serrate are in red font and highlighted. The putative CTG alternative initiation codon is also highlighted. (B) The Delta 5' UTR, which comprises 679 bps in D. melanogaster, includes 6 putative uORFs, which are underlined. The D. virilis 5' UTR spans 935 bp and 9 putative uORFs also underlined. Start and stop codons of uORFs are highlighted in grey. Three uORFs are highly conserved between the two species; two of the three are nested in each other. The single initiator ATG of Delta is in red font and highlighted.*
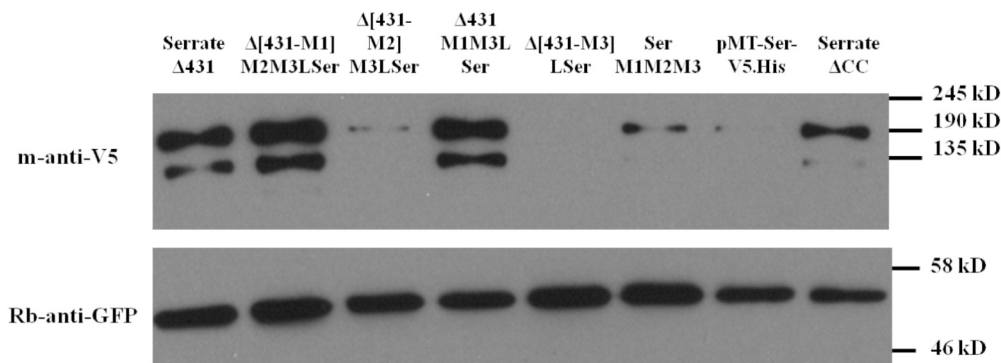
Figure 3. *Western blotting showing the expression of SERRATE starting translation at alternative AUGs. The vectors, pMT – Δ431 – Ser – V5.His, pMT – Δ[431M1] – M2M3LSer –V5.His, pMT – Δ[431M2] – M3LSer – V5.His, pMT – Δ431 – M1M3LSer – V5.His, pMT – Δ[431M3] – LSer – V5.His, pMT – M1M2M3Ser – V5.His, pMT – Ser – V5.His, pMT – ΔCC – Ser – V5. His (upper pane), were co- transfected with the RACT – GFP vector (lower pane) into S2 cells. Serrate fused to V5 epitope, was detected using a mouse anti–V5 antibody. Two bands were detected, ~245 kDa and ~150 kDa. GFP (~55 kDa) was detected using a rabbit anti-GFP antibody.*
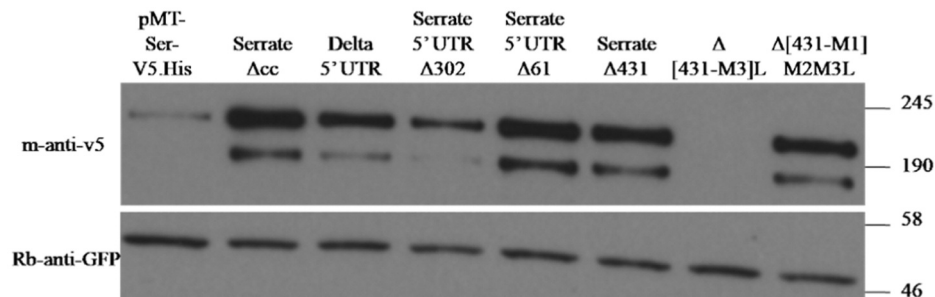
Figure 4. *Western blots showing the expression levels of Serrate with gradual truncation of Serrate 5' UTR, or with the inclusion of Delta 5' UTR. The indicated expression constructs, were co-transfected with Ract – GFP into S2 cells. Serrate fused to V5 epitope, was detected using a mouse anti–V5 antibody. Two bands were detected, ~245 kDa and ~150 kDa. GFP (~55 kDa) was detected using a rabbit anti-GFP antibody. The Δ431 deletion shows higher level of protein expression compared to pMTΔ – Ser – V5.His, where protein expression levels are the lowest.*

Serrate protein bears an unusually long N-terminal extension. An analysis of the Serrate amino acid sequence revealed that the predicted hydrophobic signal peptide is not at the very N-terminus, but located approximately 60 residues into the translated sequence. As predicted by SignalP-4.1 (37), the signal peptide is located 61-84 residues downstream from M1, or at 31-54 residues from M3, (Figure 1A). However, we noticed a CTG (Leu) codon only 9 residues upstream of the putative signal peptide. CTG has been proposed to act as an alternative initiation codon in animal cells (34). To investigate whether this CTG (Leu) might act as an initiation codon for *Serrate*, we deleted the entire 5' UTR and the 5' end of the CDS up to this CTG (Leu) codon. This pMT – Δ[431M3] – LSer – V5.His expression construct yielded no protein (Figure 3 upper panel). Lastly, we used site-directed mutagenesis on the pMT – Ser – V5.His vector (5' UTR and poly-C stretch included)

to modify the CTG codon to CCT, thus generating the pMT – M1M2M3Ser – V5.His (CTG codon destroyed). This expression construct yielded the same protein levels as the original pMT – Ser – V5.His (Figure 3 upper panel). Taken together these data suggest that the CTG (Leu) codon cannot serve as an initiation codon for Serrate. Instead, the Serrate nascent polypeptide chain has an unusually long N-terminal extension (30-60 residues) before the start of its signal peptide, depending on which of the three alternative ATG codons is used for its translation.

The 5' UTRs of Serrate and Delta may attenuate translation. The presence of the *Serrate* 5' UTR along with a 21-cytosine tail (pMT – Ser – V5.His expression construct), which had been incorporated during the cloning procedure, dramatically reduced translational efficiency (Figure 4, cf SerΔ431 with pMT-Ser-V5.His). The reduction was not due to reduced

865

transcription, as the induction of transcription using the *Drosophila* metallothionein *MtnA* promoter (a Cu^{2+} metal-inducible promoter) ensures a standardized method of transcription, therefore equal number of transcripts among samples. We concluded that the 5' UTR contains a translational silencing region. To map the sequences that are responsible for this translational down-regulation, we generated partial-deletions of the *Serrate* 5' UTR in our expression constructs.

Firstly, we removed the 21-cytosine stretch to make pMT – ΔCC – Ser – V5.His. As shown in Figure 4, the removal of the 21-cytosine stretch increased expression robustly. Next, we generated two further deletion constructs; in the pMT – Δ61 – Ser – V5.His, the first 61 bp of the 5' UTR (5' UTR-Δ61) were deleted (including the ATG of the first uORF), then the pMT – Δ302 – Ser – V5.His was generated where 302 bps of the 431 bps of the *Serrate* 5' UTR (including the second uORF) were deleted (5' UTR-Δ302); subsequently, the pMT – Δ431 – Ser – V5.His vector was constructed where the entire 5' UTR (Δ431) (including all three uORFs) was removed. Altogether, the deletion of the entire *Serrate* 5' UTR greatly increased protein expression levels by approximately 10-fold (pMT – Ser – V5.His *vs* Δ431). Substituting the *Delta* 5' UTR upstream of *Serrate*, using the 5'UTR-less construct pMT-Δ431-Ser-V5.His as a vector, caused a reduction in the levels of Serrate expression (Figure 4), suggesting that the *Delta* 5' UTR may also have a negative effect on translation. Our results suggest that the 21-cytosine stretch incorporated during the cloning procedure had the most adverse effect on Serrate translation. In comparison, the presence of uORFs in the *Serrate* and *Delta* 5' UTR had milder effects on reduction of Serrate protein levels. The (Δ[431M1]_M2M3L) expression construct yielded the highest levels of Serrate, showing that the M2 and M3 methionine codons are robust initiators of translation, while the deletion of the three methionine codons abolished protein expression (Δ[431_M3]L).

Discussion

Alternative initiation codons. Serrate translation uses three alternative ATG initiation codons, found within 93 bp of each other. The three ATG codons are interchangeable, although M3 seems to have lower efficiency (Figure 3). The first two ATG codons are located within a sequence context that conforms better with the *Drosophila* Kozak consensus sequence *(C/A)AA(A/C)AUG* (33), hence, they are expected to produce a robust protein output: M1 is preceded by CAGA (3/4 match), M2 by CAAA (perfect match) and M3 by GGGA (1/4 match). The positioning of a translation initiation codon within a "poor" sequence context results in inefficient ribosomal recognition (and bypassing) a phenomenon termed as "leaky scanning" (29). Since the first

two AUGs of *Serrate* have better Kozak consensus sequences, they are expected to engage the ribosome effectively and make the third AUG less important, although it can still act as an initiation codon when the first two are absent (Figure 3). In this light, it is interesting that all three AUGs are conserved in the distantly related *D. virilis* (Figure 2). In contrast to the three ATGs in the 5' end of the *Serrate* coding sequence, all of which can serve as translation initiation sites, a possible alternative initiation CTG codon (leucine), which is in closer proximity to the signal peptide coding sequence, is incapable of translation initiation, although it is also conserved in *D. virilis* (Figure 3). It thus appears that *Serrate* has a partially conserved 82-88 amino acid N-terminal extension (60 aa in *D. melanogaster*/ 66 aa in *D. virilis* and an identical 22 aa signal peptide), something not seen in *Delta*, which possesses a "classical" 22 aa signal peptide in both *D. melanogaster* and *virilis* located immediately downstream of the initiator methionine. It is unclear why *Serrate* has evolved to contain such a long N-terminal extension which is removed in the mature protein, something also seen in *Drosophila shotgun* (*DE-cadherin*) and *crumbs* genes (38). One possibility is that the cleaved N-terminal peptide may have some novel function in the cell. Alternatively, we suppose that the extended signal peptide has no function at the protein level, but serves at the mRNA level to accommodate the three initiator codons, which may interact with the 5'UTR, in ways yet to be discovered, in order to implement strict translational regulation of S*errate*.

Translational attenuation via 5' UTRs. The 5' UTRs of *Serrate* and *Delta* negatively regulate translation. We were able to document this activity by observing the expression levels of six *Serrate* expressing constructs in *Drosophila* S2 cells after manipulating the 5' UTR. We observed a modest decrease of *Serrate* translation in the presence of the *Delta* and *Serrate* 5' UTRs. These two 5' UTRs*,* which are twice or three times the length of the average 5' UTR in *Drosophila*, respectively, (26) contain three uORFs (*Serrate* 5' UTR) and six uORFs (*Delta* 5' UTR). These are likely to cause the down-regulation of protein expression. However, it cannot be excluded that the 5' UTR folds into a conformation, which may hinder ribosome access, directly or *via* binding of translational repressors. We used the MFOLD algorithm to predict mRNA secondary structure (39) and indeed saw that both the *Serrate* and *Delta* 5' UTRs can fold to form extensive stem-loop structures (data not shown).

The mechanism by which uORFs down-regulate gene expression in eukaryotes is *via* the ribosome, which uses its small 40S ribosomal subunit to bind the mRNA cap and then scan the 5' end of the mRNA until it recognises a start codon. The presence of uORFs (AUG/stop pairs) in the 5'UTR, favours recruitment of scanning ribosomes to these alternative AUGs and as a result reduces the fraction of ribosomes that reach the CDS AUG initiation codon. The

suppressive effect of the 5' UTR on the expression of the downstream CDS is higher in the presence of ORFs overlapping with the CDS (oORFs), as a result of reduced ribosome re-binding efficiency (28); nevertheless, oORFs were not detected in the *Serrate* and *Delta* 5' UTRs. Another way by which uORFs can suppress CDS translation is by ribosome stalling at their termination codons (40). All these effects on productive translation initiation at the actual CDS can secondarily cause mRNA degradation *via* nonsense mediated decay (40). Proteomic analyses have linked predicted uORFs to lower protein levels (41,43). Consistently, our results show that the deletion of the uORFs in the *Serrate* and *Delta* 5' UTRs results in up-regulation of protein expression (Figure 4). Conservation of a uORF is a good indicator of its translation; these are termed conserved peptide uORFs (CPuORFs) (44). Stable peptides encoded by uORFs have been detected by mass spectrometry; however, a functional role has not been determined for these peptides (44). Instead, the translation of the uORFs seems to be responsible for the suppressive effect observed on protein expression, as described above. This is consistent with a lack of CPuORFs and a general lack of conservation of the *Serrate* 5' UTR between *D. melanogaster* and *D. virilis* (Figure 2A). The latter is a distantly related drosophilid (~ 40 Mya diverged from *D. melanogaster*) and contains only a single uORF in its *Serrate* 5' UTR, which is unrelated to any of the three uORFs in the *Serrate* 5'UTR in *D. melanogaster*. Interestingly, the *Delta* 5' UTR of *D. virilis* contains 9 uORFs, three of which are highly conserved with three of the six *D. melanogaster Delta* uORFs (Figure 2B); whether this reflects a role of the encoded peptides or a role of the CPuORF codons' translation at the RNA level is a matter for further investigation.

It is not surprising that *Serrate* and *Delta* are subjected to translational repression. Recent ribosome profiling and bio-informatic studies suggest that one or more uORFs are present in 58.7% of all *Drosophila* 5' UTRs (29) suggesting that uORF–mediated translational regulation is a widespread phenomenon. A well-studied case of translational control in *Drosophila* by a uORF is the silencing of *msl-2* translation, which is necessary for the survival of female flies. The 5' UTR of the *msl-2* gene includes three uORFs which act in conjunction with the female-specific Sex-lethal (Sxl) protein, which binds to a nearby cis-regulatory mRNA element named the B site, to cause silencing of the *msl-2* (29). These uORFs need this nearby RNA-bound protein in order to exert their repressive effect which raises the possibility that uORFs do not simply impose a threshold on translation initiation, but add a heretofore little-studied level of regulation, which involves RNA-binding trans-acting proteins. In the study conducted by Schleich and colleagues (45) in *Drosophila* a translation re-initiation factor was discovered, DENR–MCT-1, which is needed for uORF containing mRNAs, but differs from the well-known GCN4 paradigm, which relies on the GCN2 kinase that phosphorylates the initiation factor eIF2α upon starvation (46). DENR-MCT1 functions in non-stressed cells and affects mRNAs involved in cellular proliferation and tissue growth to promote ribosome re-initiation downstream of a uORF.

mRNA translation can also be regulated by small micro RNAs (miRNAs) that hybridize to mRNA sequences that are frequently located in the 3' UTR (47). Normally, the structural features and regulatory sequences within the mRNA are responsible for its translational outcome (mostly repressive but also activating). The canonical end modifications of mRNA molecules i.e. the cap structure and the poly (A) tail are strong promoters of translation initiation (47). However, the internal ribosome-entry sites (IRESs) which mediate cap-independent translation initiation, uORFs, which normally reduce translation from the main ORF and secondary or tertiary RNA structures, such as hairpins and pseudoknots commonly block initiation (47). Nevertheless, hairpins and pseudoknots can also be part of the IRES elements and, therefore, promote cap-independent translation or can be specific binding sites for regulatory complexes which are crucial determinants of mRNA translation (47). *Serrate* and *Delta* translational repression may require, in addition to their uORFs, putative RNA binding proteins that could recognize their 5' UTR secondary structures. Regardless of the mechanism involved, the characterization of the *Serrate* and *Delta* 5' UTRs as attenuators of translation can be used in the construction of transgenes where high levels of expression may cause undesired effects. Additionally, we found that homopolymeric tails, often used in cDNA cloning, can also strongly inhibit translation and this should also be considered in the design oftransgenes.

Conflicts of Interest

The Authors have no conflicts of interest to declare regarding this study.

Authors' Contributions

Georgia Deliconstantinos and Konstantina Kalodimou performed the experiments. Christos Delidakis designed the study. Georgia Deliconstantinos and Christos Delidakis analyzed the results and wrote the manuscript.

Acknowledgements

References

1 Fleming RJ: Structural conservation of Notch receptors and ligands. Semin Cell Dev Biol *9*: 599-607, 1998. PMID: 9918871. DOI: 10.1006/scdb.1998.0260

2 Glittenberg M, Pitsouli C, Garvey C, Delidakis C and Bray S: Role of conserved intracellular motifs in Serrate signalling, cis-inhibition and endocytosis. EMBO J *25*: 4697-4706, 2006. PMID: 17006545. DOI: 10.1038/sj.emboj.7601337

3 Fleming RJ, Hori K, Sen A, Filloramo GV, Langer JM, Obar RA, Artavanis-Tsakonas S and Maharaj-Best AC: An extracellular region of Serrate is essential for ligand-induced cis-inhibition of Notch signaling. Development *140*: 2039-2049, 2013. PMID: 23571220. DOI: 10.1242/dev.087916

4 Artavanis-Tsakonas S, Rand MD and Lake RJ: Notch signaling: cell fate control and signal integration in development. Science *284*: 770-776, 1999. PMID: 10221902. DOI: 10.1126/science.284.5415.770

5 Yang SA and Deng WM: Serrate/Notch signaling regulates the size of the progenitor cell pool in Drosophila imaginal rings. Genetics *209*: 829-843, 2018. PMID: 29773559. DOI: 10.1534/genetics.118.300963

6 Ferguson GB and Martinez-Agosto JA: Kicking it up a Notch for the best in show: Scalloped leads Yorkie into the haematopoietic arena. Fly (Austin) *8*: 206-217, 2014. PMID: 26151599. DOI: 10.1080/19336934.2015.1055427

7 Meloty-Kapella L, Shergill B, Kuon J, Botvinick E and Weinmaster G: Notch ligand endocytosis generates mechanical pulling force dependent on dynamin, epsins, and actin. Dev Cell *22*: 1299-1312, 2012. PMID: 22658936. DOI: 10.1016/j.devcel.2012.04.005

8 Gordon WR, Zimmerman B, He L, Miles LJ, Huang J, Tiyanont K, McArthur DG, Aster JC, Perrimon N, Loparo JJ and Blacklow SC: Mechanical allostery: evidence for a force requirement in the proteolytic activation of Notch. Dev Cell *33*: 729-36, 2017. PMID: 26051539. DOI: 10.1016/j.devcel.2015.05.004

9 Bray S and Furriols M: Notch pathway: making sense of suppressor of hairless. Curr Biol *11*: R217-221, 2001. PMID: 11301266. DOI: 10.1016/S0960-9822(01)00109-9

10 Petcherski AG and Kimble J: Mastermind is a putative activator for Notch. Curr Biol *10*: R471-473, 2000. PMID: 10898989. DOI: 10.1016/s0960-9822(00)00577-7

11 Bray S: Notch signaling: a simple pathway becomes complex. Nat Rev Mol Cell Biol *7*: 678-689, 2006. PMID: 16921404. DOI: 10.1038/nrm2009

12 Kopan R and Ilagan MXG: The canonical Notch signaling pathway: unfolding the activation mechanism. Cell *137*: 216-233, 2009. PMID: 19379690. DOI: 10.1016/j.cell.2009.03.045

13 de Celis JF and Bray S: Feed-back mechanisms affecting Notch activation at the dorsoventral boundary in the Drosophila wing. Development *124*: 3241-3251, 1997. PMID: 9310319.

14 Klein T, Brennan K and Arias AM: An intrinsic dominant negative activity of Serrate that is modulated during wing development in Drosophila. Dev Biol *189*: 123-134, 1997. PMID: 9281342. DOI: 10.1006/dbio.1997.8564

15 Micchelli CA, Rulifson EJ and Blair SS: The function and regulation of cut expression on the wing margin of Drosophila: Notch, Wingless and a dominant negative role for Delta and Serrate. Development *124*: 1485-1495, 1997. PMID: 9108365.

16 Sakamoto K, Ohara O, Takagi M, Takeda S and Katsube K: Intracellular cell-autonomous association of Notch and its ligands: a novel mechanism of Notch signal modification. Dev Biol *241*: 313-326, 2002. PMID: 11784114. DOI: 10.1006/dbio.2001.0517

17 Li Y and Baker NE: The roles of cis-inactivation by Notch ligands and of neuralized during eye and bristle patterning in Drosophila. BMC Dev Biol *4*: 5-16, 2004. PMID: 15113404. DOI: 10.1186/1471-213X-4-5

18 Pandey A and Jafar-Nejad H: Cell aggregation assays to evaluate the binding of the Drosophila Notch with trans-ligands and its inhibition by cis-ligands. J Vis Exp *131*: e56919-56930, 2018. PMID: 29364239. DOI: 10.3791/56919

19 Palmer WH, Jia D and Deng WM: Cis-interactions between Notch and its ligands block ligand independent Notch activity. eLife *3*: e04415-04426, 2014. PMID: 25486593. DOI: 10.7554/eLife.04415

20 Luca VC, Ge C, Kakuda S, Wu D, Roein-Peikar M, Haltiwanger RS, Zhu C, Ha T and Garcia KC: Notch-Jagged complex structure implicates a catch bond in tuning ligand sensitivity. Science *355*: 1320-1324, 2017. PMID: 28254785. DOI: 10.1126/science.aaf9739

21 Pandey A, Harvey BM, Lopez MF, Ito A, Haltiwanger RS and Jafar-Nejad H: Glycosylation of specific Notch EGF repeats by O-Fut1 and Fringe regulates Notch signaling in Drosophila. Cell Rep *29*: 2054-2066, 2019. PMID: 31722217. DOI: 10.1016/j.celrep.2019.10.027

22 Lai EC, Roegiers F, Qin X, Jan YN and Rubin GM: The ubiquitin ligase Drosophila Mindbomb promotes Notch signaling by regulating the localization and activity of Serrate and Delta. Development *132*: 2319-2332, 2005. PMID: 15829515. DOI: 10.1242/dev.01825

23 Pitsouli C and Delidakis C: The interplay between DSL proteins and ubiquitin ligases in Notch signaling. Development *132*: 4041-4050, 2005. PMID: 16093323. DOI: 10.1242/dev.01979

24 Daskalaki A, Shalaby NA, Kux K, Tsoumpekos G, Tsibidis GD, Muskavitch MA and Delidakis C: Distinct intracellular motifs of Delta mediate its ubiquitylation and activation by Mindbomb1 and Neuralized. J Cell Biol *195*: 1017-1031, 2011. PMID: 22162135. DOI: 10.1083/jcb.201105166

25 Berndt N, Seib E, Kim S, Troost T, Lyga M, Langenbach J, Haensch S, Kalodimou K, Delidakis C and Klein T: Ubiquitylation-independent activation of Notch signalling by Delta. eLife *6*: e27346-27375, 2017. PMID: 28960177. DOI: 10.7554/eLife.27346

26 Chen CH, Lin HY, Pan CL and Chen FC: The plausible reason why the length of 5' untranslated region is unrelated to organismal complexity. BMC Res Notes *4*: 312-324, 2011. PMID: 21871111. DOI: 10.1186/1756-0500-4-312

27 Pesole G, Mignone F, Gissi C, Grillo G, Licciulli F and Liuni S: Structural and functional features of eukaryotic mRNA untranslated regions. Gene *276*: 73-81, 2001. PMID: 11591473. DOI: 10.1016/s0378-1119(01)00674-6

28 Johnstone TG, Bazzini AA and Giraldez AJ: Upstream ORFs are prevalent translational repressors in vertebrates. EMBO J *35*: 706-723, 2016. PMID: 26896445. DOI: 10.15252/embj.201592759

29 Medenbach J, Seiler M and Hentze MW: Translational control via protein-regulated upstream open reading frames. Cell *145*: 902-913, 2011. PMID: 21663794. DOI: 10.1016/j.cell.2011.05.005

30 Sonenberg N and Hinnebusch AG: Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*: 731-745, 2009. PMID: 19239892. DOI: 10.1016/j.cell.2009.01.042

31 Mueller PP and Hinnebusch AG: Multiple upstream AUG codons mediate translational control of GCN4. Cell *45*: 201-207, 1986. PMID: 3516411. DOI: 10.1016/0092-8674(86)90384-3

32 Fleming RJ, Scottdale TN, Diederich RJ and Artavanis-Tsakonas S: The gene Serrate encodes a putative EGF-like transmembrane protein essential for proper ectodermal development in *Drosophila melanogaster*. Genes Dev *4*: 2188-2201, 1990. PMID: 2125287. DOI: 10.1101/gad.4.12a.2188

33 Cavener DR: Comparison of the consensus sequence flanking translational start sites in Drosophila and vertebrates. Nucleic Acids Res *15*: 1353-1361, 1987. PMID: 3822832. DOI: 10.1093/nar/15.4.1353

34 Starck SR, Jiang V, Pavon-Eternod M, Prasad S, McCarthy B, Pan T and Shastri N: Leucine tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. Science *336*: 1719-1723, 2012. PMID: 22745432. DOI: 10.1126/science.1220270

35 Pairwise Sequence Alignment. Available at: https://www.ebi.ac.uk/Tools/psa/emboss_stretcher/nucleotide.html [Last accessed on January 10, 2021]

36 Zarifi I, Kiparaki M, Koumbanakis KA, Giagtzoglou N, Zacharioudaki E, Alexiadis A, Livadaras I and Delidakis C: Essential roles of Da transactivation domains in neurogenesis and in E(spl)-mediated repression. Mol Cell Biol *32*: 4534-4548, 2012. PMID: 22949507. DOI: 10.1128/MCB.00827-12

37 SignalP-5.0 Server. Available at: http://www.cbs.dtu.dk/services/SignalP/http://unafold.rna.albany.edu/?q=mfold [Last accessed on December 30, 2020]

38 Kilic A, Klose S, Dobberstein B, Knust E and Kapp K: The Drosophila Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur J Cell Biol *89*: 449-461, 2010. PMID: 20189678. DOI: 10.1016/j.ejcb.2010.02.001

39 The RNA institute. Available at: http://unafold.rna.albany.edu/?q=mfold [Last accessed on October 30, 2020]

40 Arribere JA and Gilbert WV: Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. Genome Res *23*: 977-987, 2013. PMID: 23580730. DOI: 10.1101/gr.150342.112.

41 Ye Y, Liang Y, Yu Q, Hu L, Li H, Zhang Z and Xu X: Analysis of human upstream open reading frames and impact on gene expression. Hum Genet *134*: 605-612, 2015. PMID: 25800702. DOI: 10.1007/s00439-015-1544-7

42 Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W and Selbach M: Global quantification of mammalian gene expression control. Nature *47*: 337-342, 2011. PMID: 21593866. DOI: 10.1038/nature10098

43 Hayden CA and Bosco G: Comparative genomic analysis of novel conserved peptide upstream open reading frames in Drosophila melanogaster and other dipteran species. BMC Genomics *9*: 61-75, 2008. PMID: 18237443. DOI: 10.1186/1471-2164-9-61

44 Oyama M, Itagaki C, Hata H, Suzuki Y, Izumi T, Natsume T, Isobe T and Sugano S: Analysis of small human proteins reveals the translation of upstream open reading frames of mRNAs. Genome Res *14*: 2048-2052, 2004. PMID: 15489325. DOI: 10.1101/gr.2384604

45 Schleich S, Strassburger K, Janiesch PC, Koledachkina T, Miller KK, Haneke K, Cheng YS, Kuechler K, Stoecklin G, Duncan KE and Teleman AA: DENR-MCT-1 promotes translation reinitiation downstream of uORFs to control tissue growth. Nature *512*: 208-212, 2014. PMID: 25043021. DOI: 10.1038/nature13401

46 Dever TE, Kinzy TG and Pavitt GD: Mechanism and Regulation of Protein Synthesis in *Saccharomyces cerevisiae*. Genetics *203*: 65-107, 2016. PMID: 27183566. DOI: 10.1534/genetics.115.186221

47 Gebauer F and Hentze M: Molecular mechanisms of translational control. Nat Rev Mol Cell Biol *5*: 827-835, 2004. PMID: 15459663. DOI: 10.1038/nrm1488