Floxed exon (Flexon): A flexibly positioned stop cassette for recombinase-mediated conditional gene expression

Justin M. Shaffer* and Iva Greenwald*a,1

*aDepartment of Biological Sciences, Columbia University, New York, NY 10027

Conditional gene expression is a powerful tool for genetic analysis of biological phenomena. In the widely used “lox-stop-lox” approach, insertion of a stop cassette consisting of a series of stop codons and polyadenylation signals flanked by lox sites into the 5’ untranslated region (UTR) of a gene prevents expression until the cassette is excised by tissue-specific expression of Cre recombinase. Although lox-stop-lox and similar approaches using other site-specific recombinases have been successfully used in many experimental systems, this design has certain limitations. Here, we describe the Floxed exon (Flexon) approach, which uses a stop cassette composed of an artificial exon flanked by artificial introns, designed to cause premature termination of translation and nonsense-mediated decay of the mRNA and allowing for flexible placement into a gene. We demonstrate its efficacy in Caenorhabditis elegans by showing that, when promoters that cause weak and/or transient cell-specific expression are used to drive Cre in combination with a gfp(flexon) transgene, strong and sustained expression of green fluorescent protein (GFP) is obtained in specific lineages. We also demonstrate its efficacy in an endogenous gene context: we inserted a flexon into the Argonaute gene rde-1 to abrogate RNA interference (RNAi), and restored RNAi tissue specifically by expression of Cre. Finally, we describe several potential additional applications of the Flexon approach, including more precise control of gene expression using intersectional methods, tissue-specific protein degradation, and generation of genetic mosaics. The Flexon approach should be feasible in any system where a site-specific recombination-based method may be applied.

Author contributions: J.M.S. designed research; J.M.S. performed research; J.M.S. analyzed data; J.M.S. and I.G. wrote the paper; and I.G. supervised research.

Reviewers: J.K., Johns Hopkins University; B.G., Rutgers The State University of New Jersey; and A.R., University of Minnesota Twin Cities.

The authors declare no competing interest.

This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1 To whom correspondence may be addressed. Email: isg4@columbia.edu

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2117451119/-/DCSupplemental

Published January 13, 2022.

Tools that afford spatiotemporal control of gene expression are crucial for studying genes and processes in multicellular organisms. Stop cassettes consist of exogenous sequences that interrupt gene expression and flanking site-specific recombinase sites to allow for tissue-specific excision and restoration of function by expression of the cognate recombinase. We describe a stop cassette called a flexon, composed of an artificial exon flanked by artificial introns that can be flexibly positioned in a gene. We demonstrate its efficacy in Caenorhabditis elegans for lineage-specific control of gene expression and for tissue-specific RNA interference and discuss other potential uses. The Flexon approach should be feasible in any system amenable to site-specific recombination-based methods and applicable to diverse areas including development, neuroscience, and metabolism.

Significance

PNAS 2022 Vol. 119 No. 3 e2117451119

https://doi.org/10.1073/pnas.2117451119  1 of 9
a stop cassette into these sequences could alter expression of the downstream gene; even after the cassette is excised, the lox scar could disrupt important regulatory sequences.

Here, we outline a strategy for a stop cassette that increases insertion site flexibility for both transgenes and endogenous genes, while retaining spatiotemporal control of expression. We call this strategy “Flexon,” for “Floxed exon,” in which we engineer a stop cassette that includes an artificial exon designed to prevent protein translation by causing premature termination in all three reading frames and to trigger nonsense-mediated decay of the mRNA that contains it. A flexon can be inserted into an intron or exon of a gene, providing different options for its placement within a transgene or endogenous locus. We demonstrate the efficacy of the Flexon approach for creating bright lineage markers and for abrogating function of an endogenous Argonaute gene to allow restoration of function for tissue-specific RNA interference (RNAi) and describe other potential applications of Flexon for genetic analysis.

Results

General Design Considerations. As with other stop cassettes, the principle behind the flexon is that it should prevent target gene expression in the absence of tissue-specific recombination-mediated excision. In addition, the flexon is designed to be versatile regarding where it can be inserted in an open reading frame and prevent ectopic gene expression due to spurious initiation downstream of the cassette. Furthermore, although we have used Cre recombinase in this study, the Flexon approach is compatible with any site-specific recombination system. We note that with regard to nomenclature, we use “Flexon” to describe the approach and “flexon” for the cassette used to create a new genotype.

Any flexon contains an artificial exon with one or more stop codons flanked by artificial introns (Fig. 1 A and B). Each flanking intron of the cassette contains a single lox site, and both lox sites are oriented in the same direction to allow for excision of the stop sequence in the exon (23). Due to the flanking introns, the stop sequence is spliced into the transcript if the cassette is inserted into an exon or an intron, allowing for flexible placement within the gene of interest (Fig. L4).

A flexon is designed to block gene expression at the level of translation instead of transcription. The artificial exon contains redundant mechanisms designed to halt translation of the protein: it contains stop codons both in frame and out of frame, relative to the gene of interest, and creates a frameshift that introduces additional premature stop codons in all frames. The flexon used in this study is bounded by lox2272 sites. Specific parameters may be varied as described in the text and SI Appendix.

Design of the flexon Cassette Used in This Study. In the test cases here, we used a flexon with the following sequence elements (Fig. 1B). Design considerations that may apply to other flexon cassettes are addressed in Discussion and in SI Appendix.

1) In-frame exon sequence: The exon contains a three-codon leader to a stop codon that is in frame with the coding region of the gene in which it was inserted (Fig. 1B).

2) Frameshift-generating exon sequence: We used 62 bp of sequence from the neutral 3' UTR from tmb-2 (26) to cause a frameshift of the downstream coding region (Fig. 1B). The total length of the exon sequence is the same as the second exon in a widely used, artificial intron-containing form of the coding region of green fluorescent protein (gfp) (23) to ensure it is of sufficient length to be spliced into the mRNA. The length of the exon was kept as small as possible to ensure efficient Cre-mediated excision, facilitate cloning, and promote homologous repair for endogenous gene insertion.

3) Intron sequences: The artificial introns were derived from a similar sequences (Fire Vector Kit, 1995) are commonly used for transgenes in C. elegans because they demonstrate efficient splicing, due to the short sequence length and canonical splice acceptor and donor sequences.

Fig. 1. Design and use of a Floxed exon cassette (flexon). (A) A traditional stop cassette (left) is placed in the 5' UTR of the gene of interest and designed to block gene expression through a transcriptional stop. A flexon stop cassette (right) is designed to abrogate gene expression by a translational stop and frame shift mutations leading to stop codons in other frames, and is also expected to result in nonsense-mediated decay (NMD) of the mRNA. Further differences are discussed in the text. Both traditional and flexon stop cassettes may use strong, ubiquitous promoters to drive expression in specific lineages or tissues after cassette excision by Cre recombinase driven by the promoter of "your favorite gene" (yfg). (B) In this study, the artificial exon of the flexon has an in-frame, three-codon leader sequence that ends with an in-frame stop codon followed by a frameshift-generating sequence that creates additional premature stop codons in all frames. The flexon used in this study is bounded by lox2272 sites. Specific parameters may be varied as described in the text and SI Appendix.

PNAS Shaffer and Greenwald
Flxed exon (flexon): A flexibly positioned stop cassette for recombinase-mediated conditional gene expression
Using the Flexon Approach in a gfp(flexon) Transgene for Strong, Cre-Mediated, Tissue-Specific Expression of GFP. We replaced the first or second intron in codon-optimized gfp sequences typically used in *C. elegans* transgenes (Materials and Methods) to create “gfp(flexon)” sequences (shown schematically in *SI Appendix*, Fig. S1) and used the strong ubiquitous promoter rps-27p (27) to drive expression in all somatic cells. The resulting rps-27p::gfp(flexon) transgenes produce no or minimal visible GFP fluorescence on their own (*SI Appendix*, Fig. S2), indicating little to no ectopic expression in the absence of Cre. We then combined rps-27p::gfp(flexon) with two different Cre drivers that were made using tissue-specific promoters that drive weak and/or transient expression, as described below. The results were remarkable: The expected lineages were brightly marked throughout larval development (Figs. 2 and 3 and *SI Appendix*, Fig. S3).

The first driver we used would be predicted to result in tissue-specific excision of the *flexon* in all somatic gonadal cells. A *C. elegans* L1 larva hatches with two somatic gonad precursor cells, Z1 and Z4, which generate all of the structures of the somatic gonad during postembryonic development (Fig. 2A). During the first phase of gonadogenesis, Z1 and Z4 give rise to 12 cells that form the somatic gonad primordium in the L2 stage, and the somatic gonad blast cells divide in the L3 stage and give rise to many additional cells (28).

The 5′ regulatory region for the *ckb*-3 gene, denoted *ckb*-3p, drives expression in the somatic gonad precursor cells Z1 and Z4 in embryos and L1 larvae, but expression rapidly diminishes as the lineage progresses (29). When *ckb*-3p is used to drive 2xNLS::GFP, nuclear GFP is readily visualized in Z1 and Z4 in the L1 stage, but progressively dims and is essentially undetectable by the time the somatic gonad primordium has formed in the L2 stage (Fig. 2A and B). The combination of a *ckb*-3p::Cre driver that expresses a form of Cre recombinase optimized for *C. elegans* (30) (Materials and Methods) with rps-27p::2xNLS::gfp(flexon) produces sustained, strong, and specific expression of 2xNLS::GFP in the somatic gonad throughout development (Fig. 2B and *SI Appendix*, Fig. S3). Furthermore, when we used a histone tag to stabilize GFP as well as for nuclear localization (rps-27p::gfp(flexon):h2b), fluorescence intensity was sufficient for visualization at lower exposure, and even using a dissecting microscope (Fig. 2 B and C). Excision frequency approached 100% for both gfp(flexon) transgenes; by the time Z1 and Z4 had divided, excision had always been observed.

The second driver we used would be expected to result in tissue-specific excision of the *flexon* in all vulval precursor cells (VPCs). The VPCs are six polarized epithelial cells, named P3.p–P8.p, that are born in the L1 stage and remain quiescent until the L3 stage. At that time, P5.p, P6.p, and P7.p are induced by epidermal growth factor receptor and LIN-12/Notch signaling to adopt a vulval fate, and the descendants of these cells form the vulval primordium in the L4 stage (31–33). The other VPCs, P3.p, P4.p, and P8.p, do not receive spatial patterning signals and divide once to generate daughter cells that fuse with the major hypodermal syncytium (P3.p sometimes fuses directly, without dividing). A composite regulatory region made from the 5′ flanking region and introns of the *lin-31* gene, called *lin-31p* (34–36) (Fig. 3A) displays a dynamic pattern of expression. GFP fluorescence is visible after the cells are born in the late L1 stage and appears uniform in all VPCs during the L2 stage; at the time VPCs commit to their fates in the L3 stage, fluorescence begins to decrease visibly in P5.p, P6.p, and P7.p, and progressively dims and becomes undetectable as their lineages progress (36) (Fig. 3B). Fluorescence in P3.p, P4.p, and P8.p remains stronger than in the other VPCs at the beginning of the L3 stage but becomes undetectable after they divide in the L3 stage. By contrast, when rps-27p::2xNLS::gfp(flexon) was combined with a lin-31p::Cre driver, GFP was strongly, uniformly, and continuously expressed in the VPCs and their descendants (Fig. 3 B and C). As in the somatic gonad, the use of a histone tag to stabilize GFP results in fluorescence visible using a dissecting microscope.

Using the Flexon Approach in an Endogenous Gene: An rde-1(flexon) Allele Enables Tissue-Specific RNAi. In *C. elegans*, RNAi is usually performed by feeding worms with bacteria that express double-stranded RNA for a gene of interest (37). Tissue-specific RNAi has been accomplished by selectively expressing RDE-1/Argonaute in an *rde-1* hypomorphic or null mutant background (38, 39). As with any other transgenic method, rescue experiments are affected by the strength and specificity of available regulatory regions; for example, our laboratory has had difficulty achieving effective tissue-specific RNAi using *ckb*-3p or *lin-31p* to drive RDE-1 expression in an *rde-1* null mutant background. The transgene-based Flexon approach described above would be one way to circumvent this problem, but for proof of concept to show that a flexon works in an endogenous gene context, we created an *rde-1(flexon)* allele that would allow for a physiological level of RDE-1 protein to be restored after tissue-specific, Cre-mediated recombination.

**Design of the endogenous rde-1(flexon) allele.** We used CRISPR-Cas9 to replace the small ninth intron of the endogenous locus of *rde-1* with the same *flexon* cassette used in the gfp(flexon) transgenes (Fig. 4A). The resulting allele, *rde-1(ar660)*, is hereafter referred to as *rde-1(flexon)*. The position was chosen to disrupt the catalytic PIWI domain, as the goal was to create a sufficiently strong loss-of-function allele that would greatly reduce or eliminate gene activity even if there was a low level of leakiness in this gene context (Potential Limitations).

**Design of the test of rde-1(flexon) for somatic gonad-specific RNAi.** To create a strain suitable for somatic gonad-specific RNAi, we combined the endogenous *rde-1(flexon)* allele with the *ckb*-3p::Cre driver described above. The transcription factor HLI-1/2/E2A, an essential gene for early embryonic development and for gonadogenesis (40), offered an incisive test case. When L4 larvae are fed bacteria expressing double-stranded RNA for *hlh-2* (“*hlh-2(2-L4-RNAi)*”), all offspring arrest during embryogenesis (41) (Fig. 4B). When embryonic lethality is bypassed by feeding L1 larvae with bacteria expressing double-stranded RNA for *hlh-2* (“*hlh-2(L1-RNAi)*”), the treated larvae become sterile adults that lack or have compromised distal tip cells (DTCs), which serve as the germline stem cell niche (42) (Fig. 4B); sterility can be readily assessed at the dissecting microscope level. *hlh-2* is also required in the DTCs to lead gonad arm outgrowth in adults in the compound microscope. We therefore tested whether *rde-1(flexon)* on its own prevents these deleterious effects of *hlh-2(RNAi)* and whether early lethality is prevented but highly penetrant defects in gonad development result when *rde-1(flexon)* is combined with *ckb*-3p::Cre. We also performed additional supplemental assessments as described below. *rde-1(flexon) strongly reduces rde-1 activity.* We tested whether inserting a *flexon* into the endogenous *rde-1* gene causes a strong loss-of-function phenotype, a necessary prerequisite for creating a system for tissue-specific RNAi. When we performed *hlh-2(L4-RNAi)* by treating strain N2 (wild type) L4 larvae, all of their progeny arrested during embryogenesis (Fig. 4C). By contrast, when we treated *rde-1(flexon) L4* larvae, all of their progeny hatched (Fig. 4C).

Similarly, when we performed *hlh-2(L1-RNAi)* on N2 larvae, we observed highly penetrant sterility, which was not observed when *hlh-2(L1-RNAi)* was performed on *rde-1(flexon)* hermaphrodites.
The Flexon approach enables strong, persistent expression of GFP in all cells of the somatic gonad lineage. (A) ckb-3p drives expression of a high level of GFP specifically in the somatic gonad precursor cells Z1 and Z4 (29), but the level of GFP diminishes rapidly as the lineage progresses (see B, Left column). (B) GFP fluorescence from the transgene arTi433[ckb-3p::2xnls::gfp] (Left column) is dimmer and does not persist as long as GFP fluorescence from arTi435[rps-27p::2xnls::gfp] (flexon) when the flexon excision is mediated by a ckb-3p::Cre driver (arTi237; Middle column; Materials and Methods), as seen in photomicrographs taken at the same exposure time and imaging parameters. Stabilization of GFP using a histone tag (arTi361[rps-27p::gfp(flexon)::h2b]; Right column) results in bright expression visible at a lower exposure time than 2xnls::gfp. (Scale bars, 10 μm.) (C) GFP(flexon)::histone in the presence of ckb-3p::Cre is visible on the dissecting scope at all larval stages. (Magnification, 50x.)

(Tab. 4C). Treated N2 larvae also displayed a highly penetrant lack of gonad arm extension (Fig. 4C), whereas the treated rde-1(flexon) larvae had normal gonad arms, with a single exception of an individual that had one normal and one abnormal gonad arm that extended but did not turn, consistent with reduced hlh-2 activity after the DTCs had formed (42). Similarly, RNAi directed against hlh-12, which encodes the dimerization partner for HLH-2 for gonad arm extension (43), had a low proportion of “escapers”: 1/104 (1%) gonad arms of treated rde-1(flexon) L1 larvae had a DTC with abnormal extension. Because most animals are unaffected, we infer that the rde-1(flexon) allele strongly abrogates rde-1 activity, but the low proportion of escapers suggests that it may not be a true null allele. While there may be some situations in which a low background of residual rde-1 activity may be problematic (39), we demonstrate below that it is eminently feasible to use rde-1(flexon) for tissue-specific RNAi.

We also tested two additional genes that have a cellular focus in tissues other than the somatic gonad: dpy-10, a cuticle collagen (44), and pos-1, an RNA-binding protein required for lineage specification in the early embryo (45). All progeny of treated N2 hermaphrodites displayed the expected phenotype: a Dumpy body shape for dpy-10(RNAi) (84/84) and embryonic lethal progeny for pos-1(RNAi) (716/716 arrested embryos). By contrast, the progeny of treated rde-1(flexon) hermaphrodites were generally unaffected: for dpy-10(RNAi), 0/462 progeny were Dpy, and for pos-1(RNAi), 4/604 arrested embryos were observed (we did not examine these arrested embryos for specific pos-1-associated defects). These results support the inference from hlh-2(RNAi) that insertion of the flexon into rde-1 severely compromises its activity.

Test of endogenous rde-1(flexon) for tissue-specific RNAi in the somatic gonad. We tested whether restoring rde-1 function specifically to the somatic gonad via excision of the flexon using the ckb-3p::Cre driver would result in highly penetrant, gonad abnormalities. As Cre is expressed in Z1 and Z4, we would expect that RNAi would be restored prior to the birth and specification of the DTCs. Indeed, hlh-2(L1-RNAi) of rde-1(flexon); ckb-3p::Cre resulted in sterility and lack of gonad arms, indicating that RNAi was efficiently restored by excision of the flexon (Fig. 4C). Moreover, all progeny of hlh-2(L4-RNAi) of rde-1(flexon); ckb-3p::Cre hermaphrodites survived (Fig. 4C), indicating that restoration of RNAi was specific to the gonad.

The observed bypass of embryonic lethality by hlh-2(RNAi) is strong evidence that restoration of RNAi by the ckb-3p::Cre driver is limited to the gonad. We corroborated this inference by performing dpy-10 and pos-1 RNAi on the rde-1(flexon); ckb-3p::Cre strain: 555/556 Dpy and 424/433 progeny of pos-1(RNAi) hermaphrodites were viable, indicating that RNAi had not been restored to the hypodermis and germ line.

Tissue-specific restoration of rde-1 activity to body wall muscle. We performed an additional test of tissue-specific RNAi by restoring rde-1 activity to the body wall muscles using hlh-1p::Cre, which drives excision in the MS muscle founder cell lineage (46–48), and performing RNAi against unc-22, which encodes the body wall muscle structural protein Twitchin/Titin (49). unc-22(RNAi) causes a fully penetrant Twitcher phenotype...
when N2 (15/15) or rde-1(ar660);hlh-1p::Cre hermaphrodites (16/16) are treated, but does not cause a phenotype when rde-1(ar660) (0/16) or rde-1(ar660);ckb-3p::Cre (0/13) hermaphrodites are treated. Thus, tissue-specific RNAi was achieved in this additional cell context.

Final comments on rde-1(flexon). We have provided proof of concept that the Flexon approach is applicable to endogenous genes, the main goal of this experiment, by showing that insertion of a flexon strongly reduces rde-1 gene activity and that excision of the flexon after expression of tissue-specific Cre drivers restores gene activity. Our analysis suggests that rde-1(ar660[flexon]) is a strong loss-of-function allele that should facilitate many tissue-specific RNAi applications and be a valuable addition to the tissue-specific RNAi toolkit for C. elegans. It should be straightforward to assess the feasibility of using rde-1(ar660[flexon]) for any specific purpose by performing RNAi in the absence of a Cre driver at the outset of a study to confirm that the associated phenotype is not observable from potential low-level leaky expression at significant penetrance (as indeed would also be necessary for transgene rescue–based approaches).

Discussion
Stop cassettes are a versatile method for achieving conditional gene expression, a powerful approach for genetic analysis of any biological process. Here, we developed and tested a stop cassette we call a flexon, composed of an artificial exon flanked by artificial introns and site-specific recombinase sequences. A flexon can be flexibly positioned in a gene to prevent translation of its protein product until it is excised by tissue-specific recombinase. We have provided a prototype for achieving strong, tissue-specific expression of a desired protein using two transgenes, one using a weak tissue-specific promoter to drive Cre and the other a strong promoter with a gfp(flexon). We also inserted a flexon into the endogenous locus for rde-1, a gene required for RNAi, and showed that gene function was abrogated in the absence of Cre, but restored in specific tissues in the presence of a Cre driver. While it was used here as proof of concept for flexon function in an endogenous gene context, the rde-1(flexon) allele will be useful for tissue-specific RNAi in C. elegans. Here, we first generalize the design consideration for other flexon cassettes and then provide further examples of how the Flexon approach could be adapted to improve the efficiency of commonly used genetic tools and for additional genetic approaches in C. elegans and other experimental systems. Finally, we describe some potential limitations that should be borne in mind when applying the Flexon approach.

Design Considerations for Other flexon Cassettes. In principle, a flexon may be used with any protein-coding gene. In Fig. S4, we show several adjustments that may be made to the flexon cassette we used here and highlight here and in the SI Appendix considerations for designing a flexon. In addition to cassette adjustments, when used in a transgene, different ubiquitous or more restricted promoters may be used to tune the level or tissue in which the flexon-containing construct is expressed after excision of the cassette (Fig. S5B; SI Appendix, Fig. S3).

1) The artificial exon sequence design should be tested in silico by conceptual translation to ensure that it 1) causes a frameshift downstream and 2) introduces stop codons in all three reading frames. The exon design used in this study contains stop codons in two reading frames within the exon; stop codons in the third reading frame were generated in downstream exons by the frameshift. The exon sequence may be modified to include additional out-of-frame stop codons within the exon if necessary.

2) The specific pair of lox sequences selected to flank the exon can be chosen to allow for the possibility to excise multiple flexon cassettes simultaneously or to guard against causing intergenic DNA recombination in the presence of other genes that have loxP, loxN, or lox551 sequences (such as the loxP scar left from the self-excising cassette method of genome engineering in C. elegans) (23). Alternatively, frt or other site-specific recombinase sequences can be used in place of lox sequences for spatiotemporal control of recombination and/or for differential control when combining different flexon-containing transgenes.
3) Insertion of a flexon into an exon can be achieved by appending a splice donor site upstream of the first lox sequence and a splice acceptor sequence downstream of the second lox site. The introns here may be optimized in length or content for high splicing efficiency in alternate applications within C. elegans or for other organisms (SI Appendix).

4) When information about protein functional domains is available, it is preferable to insert the flexon such that expression from potential cryptic, downstream start sites would not result in an active protein. Once a flexon has been designed and the desired insertion obtained, it is important to assess how well gene expression or activity has been abrogated by insertion of the flexon as an important control for tissue-specific excision experiments (Potential Limitations).

Potential Applications Facilitated by Incorporating a flexon into Endogenous Genes or Transgenes. The Flexon approach may be incorporated into many different established strategies for manipulating gene activity or expression. We give some examples here, emphasizing C. elegans but applicable to other experimental systems.

Increasing the Level of Limiting Effectors for Tissue-Specific Protein Degradation. Targeted protein degradation using natural degron/ubiquitin ligase pairs is a powerful method for studying gene function. Several degron/ubiquitin ligase pairs have been successfully used in C. elegans: the TIR1/AID (Auxin-Inducible Degron) system (applied to C. elegans in ref. 50), the ZIF-1/ZF1 system (51), and converting GFP itself into a degron by replacing the interaction domain of an E3 ubiquitin ligase with an anti-GFP nanobody (52). However, the amount of the ubiquitin ligase expressed has been reported to be limiting for degradation (53–55), so we anticipate that the Flexon approach will facilitate such manipulations by enabling stronger tissue-specific expression of the ubiquitin ligase. Furthermore, by using different lox variants, Cre can be used to delete a floxed, degron-tagged gene while simultaneously using a ubiquitin ligase(flexon) to quickly eliminate perduring protein, a strategy that has been demonstrated to be effective when low residual protein levels obscure the null phenotype (2).

Intersectional Approaches to Conditional Gene Expression. One way to achieve specific gene expression is to use two different tissue-specific promoters that have overlapping sites of expression to drive different recombinases, e.g., combining the Cre and Flp recombinases (56–58) (Fig. 5C). The flexible placement of the flexon (Fig. 5D) may facilitate the use of such intersectional strategies. Combinatorial recombinase approaches have been used for marking specific cell types and lineages in mice (reviewed in ref. 59); in C. elegans, these approaches may be used to generate GFP lineage markers that require lower intensity and exposure to visualize, avoiding damage from blue-light overexposure (60, 61).

Engineering Genomic Loci. For engineering genomic loci, the flexibility in placement of a flexon offers several potential advantages over stop cassettes that are placed in a 5’ UTR. First, the Flexon approach facilitates the engineering of conditional alleles for single or multiple isoforms (Fig. 5D). Second, the flexible placement reduces the chance of disrupting uncharacterized but important regulatory sequences in the 5’ UTR by the single recombination site “scar” left after excision of the cassette. Third, in many cases introns are better defined than 5’ UTRs, especially in C. elegans where about 84% of genes are trans-spliced by the addition of a splice-leader sequence to the 5’ end of pre-mRNAs (62), and about 15% of genes are expressed from operons, in which a polycistronic pre-mRNA is processed into separate transcripts by trans-splicing of the downstream gene (63), with the trans-splice

---

**Fig. 5.** Potential adjustments to the flexon design and additional applications. The Flexon system is tunable and may be incorporated into transgenes or endogenous loci to enable or improve genetic tools and approaches for manipulating gene activity as well as to mark lineages. Adjustments to the design for specific purposes (A and B) and some potential applications (C and D) are diagrammed here and described in the text.
acceptor for the downstream gene embedded in a relatively short sequence that also contains a 3′ UTR for the upstream gene. Thus, a flexon can be inserted into a trans-spliced gene, or into either gene of an operon, without potentially compromising regulatory sequences (Fig. 5D).

Finally, we note that manipulating endogenous gene expression using the Flexon approach offers an alternative to transgene-based methods for assessing tissue-specific rescue and creating genetic mosaics that may be advantageous in certain situations. For example, transgenes used for conventional tissue-specific rescue may not be expressed at endogenous levels, and when genes have multiple isoforms, the isoform selected may influence the results. By contrast, insertion of a flexon into an intron of an endogenous gene will allow for all isoforms that share that intron to be expressed under control of its natural regulatory elements after excision using a tissue-specific Cre driver (Fig. 5D).

Potential Limitations. Although the Flexon approach addresses some of the issues that have been observed with traditional lox-stop-lox, it shares other limitations inherent to any stop cassette method. One is that site-specific recombination is irreversible and therefore cannot be used for dynamic control of gene expression by itself, although it may facilitate the implementation of dynamic methods such as AID, as described above. Another limitation is that transiently low-level expression at that is not normally evident when the promoters are identified using fluorescent reporter genes may be more apparent when they are used to drive Cre (46), so tissue-specific promoters used to drive site-specific recombinases may not be as specific as desired. A third limitation is that expression in different cells of a tissue may be uneven shortly after recombination events due to asynchronous cassette excision or when occasionally excision does not occur on both chromosomes in a cell. However, recombination tends to be very efficient given the small size of the flexon, so for many applications, this limitation may not be an issue.

A final limitation shared with other stop cassette approaches is potential “leakiness.” Although we have attempted to address the causes of leakiness in traditional lox-stop-lox while also implementing the advantages of the flexible insertion of the stop cassette that a flexon offers, there is still the potential for leaky expression from a fortuitous downstream translational start site or alternative splicing/exon skipping that excludes the flexon from the final mRNA transcript. Such exon skipping may account for the rare escapers we observed when evaluating rde-1(flexon) for abrogating RNAi, and may depend on the specific gene, the specific tissue, or the specific flexon design itself. For practical purposes, whether leakiness is a problem may be easily addressed by examining the phenotype of novel endogenous flexon alleles. If incomplete penetrance is an issue, repositioning the flexon, adding an additional flexon, or combining the Flexon approach with another conditional method such as AID (2) are possible approaches to reducing such background activity.

One other consideration must be borne in mind when manipulating endogenous genes: Insertion of a flexon will abrogate gene function, so if a lethal or sterile phenotype is anticipated, marked balancer chromosomes (64, 65) or rescuing transgenes (such as C. elegans Strains. C. elegans was grown on 6-cm nematode growth medium plates seeded with Escherichia coli OP50 and maintained at 20 °C. Strain N2 (wild type) (66) and two of the Cre drivers used in this study were previously described: ar1235 [kb-3p::Cre(opt)::rps-27p::gfp(0x2)::rps-27p::2x3′ UTR] X is a single-copy insert transgene made as described in ref. 46 and mapped as part of this study, and ar1235 [nls-1p::Cre(elt):bb-2 3′ UTR] was also described in ref. 46. Cre(elt) refers to the codon-optimized Cre recombinase described previously (30). The generation of other single-copy insertion transgenes and the allele rde-1(ar660(flexon)) is described below. The full genotypes used in this study are listed in SI Appendix, Table S1.

Generation of Single-Copy Insertion Transgenes. The plasmids pRHK001 [rps-27p::gfp(flexon)::his-54::unc-54 3′ UTR], pSJ110 [kb-3p::2x3′ UTR] (67) were made in a miniMos vector backbone (pCF910) (67) using Gibson Assembly (NEB) and confirmed by sequencing. pSJ110, pSJ145, and pSJ146 contain a C. elegans codon-optimized GFP sequence tagged with an N-terminal SV40 and C-terminal egl-13 nuclear localization sequences (nis), regulated by the neutral 3′ UTR from unc-54 (68). pRHK001 uses a C. elegans GFP sequence with slightly different codon optimization (23). Plasmids were injected into N2 hermaphrodites. Random, single-copy insertions were obtained and mapped using the standard protocol (67).

Generation of rde-1(ar660), the Endogenous rde-1(flexon) Allele. The template plasmid pSJ155 was made using Gibson cloning to insert the following sequences into a pBluescript vector: The first 500 bp of rde-1 exon 9, a flexon cassette, and the first 500 bp of rde-1 exon 10. The arrangement of the sequences in pSJ155 will effectively replace the native intron 9 of rde-1 with the flexon cassette when the plasmid is used as a repair template. Intron 9 lies between two exons that encode of the catalytic PINI domain of RDE-1.

To generate rde-1(ar660), a correction mix containing pHK001 (50 ng/μL) as the repair template and a crRNA (IDT; tgagttattaaagacagtctcttttagagc-tatgc) was prepared according to an established protocol (69) and injected into the gonads of young adult N2 hermaphrodites. Animals with potential gene editing events were isolated according to the protocol, and animals homozygous for a correct rde-1(ar660) allele were confirmed through PCR genotyping and sequencing. We note that we made another allele by replacing the second intron of the endogenous rde-1 locus with a flexon cassette. In contrast to rde-1(ar660), the allele resulting from this replacement had only a weak effect on rde-1 function. Because the second intron is upstream of the exons for all the known functional domains of RDE-1, we interpret the weak phenotype as an indication that cryptic, in-frame ATG start sites downstream of the flexon allowed for the transcription of a mutant RDE-1 protein that retained significant activity. This concern was incorporated into Design Considerations for Other flexon Cassettes section, point 4, and reiterated in SI Appendix.

Microscopy. C. elegans larvae were imaged for fluorescence on a Zeiss AxioObserver Z1 inverted microscope with a 63×, 1.4 numerical aperture (NA) oil immersion objective equipped with a spinning disk, CSU-X1, a laser bench, and a Photometrics Evolve Electron Multiplying Charge Coupled Device camera. For GFP fluorescence imaging, a 488-nm, 100-mW laser was used for excitation. Larvae were mounted onto 3% agar pads and immobilized with 10 mM levamisole. Z stacks were collected from GS9684, GS9686, GS9691, GS9692, and GS9401 larvae (Figs. 2 and 3) with slices at 500-nm intervals and imaged for GFP fluorescence with the following parameters: 10% laser power, 200-ms exposure (or 50-ms exposure for GS9401), and 400 EM gain. For SI Appendix, Fig. 5G, the scan parameters were used. The animal was positioned so that a 1.4 NA oil immersion objective, and with 1,000-nm intervals between slices. For strains GS9684, GS9692, and GS9401, the number of slices used per larva varied with the size of the gonad to fully capture the nuclei of every somatic gonad cell present. The stage of the animal was determined by the number of somatic gonad cells and somatic gonad morphology. For GS9686 and GS9691, each stack contained 26 slices, which was sufficient to image the full volume of the nuclei of every VPC or VPC descendant. The stage of the animal was determined by the number of VPCs and somatic gonad morphology. Z stacks were collected from GS9401 and GS9407 (SI Appendix, Fig. 5S) with slices at 500-nm intervals and imaged for GFP fluorescence with the following parameters: 25% laser power, 500 EM gain, and the exposure time denoted in the figure. The number of slices used per larva varied with the size of the gonad to fully capture the nuclei of every somatic gonad cell present. The stage of the animal was determined by the number of somatic gonad cells and somatic gonad morphology.

Images in Fig. 2C were taken on a phone camera through the eyepiece of a Zeiss Discovery V.12 Stereoid dissecting microscope, GFP470 filter, Schott ACE 1 fiber optic light source, and X-cite series 120Q fluorescent lamp illuminator. Animals to be imaged were placed on 60-mm plates filled with 1.75% agarose and then immobilized using 10 mM levamisole. Image Quantification. Fluorescence intensity was quantified using Fiji (70, 71). Z stacks were sum projected for all slices containing cells of interest. Nuclei were manually segmented, and the mean green fluorescent background from
RNAi. The following RNAi clones were used: pKM51196, which contains the full-length hhl-2 cDNA (72), and commercially available library clones for unc-22, pos-1, dpy-10, and hhl-12 (73). For feeding RNAi (77), plates were made using NGM with 6 mM isopropyl beta-D-1-thiogalactopyranoside and 100 μM carbobenzoxy-Val-Leu-Leu-72uoro-2-thiol in 70 μL of overnight culture of bacteria expressing a single RNAi clone. Experiments were performed at 25 °C.

To test for embryonic lethality of hhl-2(RNAi), N2, GS9801, and GS9802 individual L4 larvae were placed onto RNAi plates and then removed after 1 d. The number of eggs and surviving larvae on each plate were counted immediately upon removal, and then the number of surviving larvae were counted 1 d later. To test for sterility of dpy-10(RNAi), and then the number of surviving larvae were counted 1 d later. To test for sterility of hhl-2(RNAi), N2, GS9801, and GS9802 embryos were plated in standard bleached nematode control (74) and plated onto unseeded NGM plates. Larvae were allowed to hatch for 1 h, isolated on RNAi plates, and then assessed for fertility by checking for the presence of progeny 4 d after plating (Fig. 4 B and C). To assess gonad arm morphology in hhl-2(RNAi) and hhl-12(RNAi), N2, GS9801, and GS9802 embryos were isolated by bleaching, plated directly onto RNAi plates, and assessed for gonad arm phenotypes 2 d after plating. Assessed phenotypes included gonad arm absence, failure to extend, failure to turn, failure to extend back to the midpoint, and the presence of abnormal bulges. The gonad arms of N2 animals on hhl-12(RNAi) had a very high penetrance of morphological defects (74/48).

For unc-22(RNAi), individual N2, GS9801, GS9802, and GS9831 embryos or L1 larvae were placed onto RNAi plates and assessed for twitching at the L4 stage. To determine the effect of RNAi against dpy-10 and pos-1, individual N2, GS9801, and GS9802 L1 larvae were placed onto RNAi plates and removed 2 d later after laying eggs. The progeny were then assessed for phenotypes, dumpy for dpy-10 RNAi and lethality for pos-1 RNAi.

Data Availability. All study data are included in the article and/or **SI Appendix**.

ACKNOWLEDGMENTS. We thank Katherine Luo for generating arT361 and Jessica H. (Henry) Kim for help in generating arT361. We also thank Michael Shen, Michelle Attnner, and Catherine O'Keefe for valuable comments on the manuscript prior to submission; Gary Struhl for valuable comments after revision; and Barth Grant, John Kim, and Ann Rougville for critical feedback. This work was supported by R35GM131746 from the National Institute of General Medical Sciences (to I.J.G.) and F33EY030331 from the National Eye Institute (to J.M.S.). J.M.S. was also supported by Training Grant T32GM00798. Some strains were provided by the Caenorhabditis Genetics Center, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

1. M. Lasko et al., Targeted oncogene activation by site-specific recombination in transgenic mice. Proc. Natl. Acad. Sci. U.S.A. 86, 6322–6326 (1992).
2. A. van der Vaart, M. Godfrey, V. Portegies, S. van de Heuvel, Dose-dependent functions of SWI/SNF BAF in permitting and inhibiting cell proliferation in vivo. Sci. Adv. 6, eaay3822 (2020).
3. E. L. Jackson et al., Analysis of lung tumor initiation and progression using conditional expression of oncogenic Kras. Genes Dev. 15, 3243–3248 (2001).
4. I. H. Maxwell, G. S. Harrison, W. M. Wood, G. M. Wahl, A DNA cassette containing a trimerized SV40 polyadenylation signal which efficiently blocks spurious plasmid-initiated initiation. Biotechniques 7, 276–280 (1989).
5. L. Madison et al., A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. 13, 133–140 (2010).
6. C. J. Potter, B. Tasic, E. V. Ruosler, L. Liang, L. Luo, The Q system: A repressible binary system for transgene expression, lineage tracing, and mosaic analysis. Cell 141, 536–548 (2010).
7. G. Struhl, K. Basler, Organizing activity of wingless protein in Drosophila. Cell 72, 527–540 (1993).
8. K. Basler, G. Struhl, Compartment boundaries and the control of Drosophila limb pattern by hedgehog protein. Nature 368, 208–214 (1994).
9. V. Cardot-Ruffino et al., Generation of a conditional Flpo/FRT mouse model expressing constitutively active Tgfβ in fibroblasts. Sci. Rep. 10, 1–13 (2020).
10. S. Gorman, D. T. Fox, G. M. Wahl, Recombinase-mediated gene activation and site-specific integration in mammalian cells. Science 255, 1351–1355 (1991).
11. L. A. Lyznik, J. C. Mitchell, L. Hirayama, T. K. Hodges, Activity of yeast FLP recombinase in maize and rice protoplasts. Nucleic Acids Res. 21, 969–975 (1993).
12. M. W. Davis, J. J. Morton, D. Carroll, E. M. Jorgensen, Gene activation using FLP recombinase in C. elegans. PLoS Genet. 4, e1000028 (2008).
13. C. Munoz-jimenez et al., An efficient FLP-based toolkit for spatiotemporal control of gene expression in Caenorhabditis elegans. Genetics 206, 1763–1778 (2017).
14. E. Z. Macosko et al., A hub-and-spoke circuit drives pheromone attraction and social integration. Cell 176, 717–729 (2019).
15. K. K. Tamai, K. Nishiwaki, bHLH transcription factors regulate organ morphogenesis and organismal morphology. Curr. Biol. 29, 568–580 (2019).
16. C. de la Cova, I. Greenwald, SEL-10/Vrb7–dependent negative feedback regulation of LIN-45/Barf signaling in C. elegans via a conserved phosphodehydrogenase. Genes Dev. 26, 2524–2532 (2012).
17. K. L. Luo, R. S. Underwood, I. Greenwald, Positive autoregulation of lag-1 and response to LIN-12 activation in cell fate decisions during C. elegans reproductive system development. Development 147, dev19482 (2020).
18. L. Timmons, D. L. Court, A. Fire, Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene 263, 102–112 (2001).
19. H. Qudota et al., Establishment of a tissue-specific RNAi system in C. elegans. Gene 400, 166–173 (2007).
20. J. S. Watts et al., New strains for tissue-specific RNAi studies in Caenorhabditis elegans. G3 10, 4167–4176 (2020).
21. M. Krause et al., A C. elegans p65–dumpy-like bHLH protein marks neuronal but not glial muscle development. Development 124, 2179–2189 (1997).
22. R. S. Kamath, M. Martinez-Campos, P. Zipperlen, A. Frigerio, Effective- ness of specific RNA-mediated interference through ingested double-stranded RNA in Caenorhabditis elegans. Genome Biol. 2, research0002.0001 (2000).
23. K. Karp, I. Greenwald, Multiple roles for the EDaughterless ortholog HLH-2 during C. elegans gonadogenesis. Dev. Biol. 272, 460–469 (2004).
24. K. K. Tamai, K. Nishiwaki, bHLH transcription factors regulate organ morphogenesis via activation of an ADAMTS protease in C. elegans. Dev. Biol. 308, 562–571 (2007).
25. A. D. Levy, J. Yang, J. M. Kramer, Molecular and genetic analyses of the Caenorhabditis elegans dpy-2 and dpy-10 collagen genes: A variety of molecular alterations affect organ morphology. Mol. Biol. Cell 4, 803–817 (1993).
26. H. Tabara, R. J. Hill, C. C. Mello, J. R. Priess, Y. Kohara, pos-1 encodes a cytoplasmic zinc-finger protein essential for germline specification in C. elegans. Development 117, 1–11 (1993).
27. C. C. Tenen, I. Greenwald, Cell non-autonomous function ofdaf-18/PTEN in the somatic gonad coordinates somatic gonad and germline development in C. elegans dauer larvae. Curr. Biol. 29, 1064–1072.e8 (2019).
Conditional gene expression

Floxed exon (Flexon): A flexibly positioned stop cassette for recombinase-mediated conditional gene expression

Shaffer and Greenwald

47. M. Krause, S. W. Harrison, S.-Q. Xu, L. Chen, A. Fire, Elements regulating cell- and stage-specific expression of the C. elegans MyoD family homolog hlh-1. Dev. Biol. 166, 133–148 (1994).

48. J. I. Murray et al., Automated analysis of embryonic gene expression with cellular resolution in C. elegans. Nat. Methods 5, 703–709 (2008).

49. D. G. Moerman, G. M. Benian, R. J. Barstead, L. A. Schriefer, R. H. Waterston, Identification and intracellular localization of the unc-22 gene product of Caenorhabditis elegans. Genes Dev. 2, 93–105 (1988).

50. L. Zhang, J. D. Ward, Z. Cheng, A. F. Dernburg, The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in C. elegans. Development 142, 4374–4384 (2015).

51. S. T. Armenti, L. L. Lohmer, D. R. Sherwood, J. Nance, Repurposing an endogenous poly-ubiquitination marks AUX/IAA proteins for degradation. PLoS Biol. 19, e3001204 (2021).

52. S. Wang et al., A toolkit for GFP-mediated tissue-specific protein degradation in C. elegans. Development 144, 2694–2701 (2017).

53. U. Aghayeva et al., DAF-16/FoxO and DAF-12/VDR control cellular plasticity both cell-autonomously and via interorgan signaling. PLoS Biol. 19, e3001204 (2021).

54. G. E. Ashley et al., An expanded auxin-inducible degron toolkit for Caenorhabditis elegans. Genetics 217, iyab006 (2021).

55. F. dos Santos Maraschin, J. Memelink, R. Offringa, Auxin-induced, SCF(TIR1)-mediated trans-splicing. Genes Dev. 24, 141–148 (2010).

56. S. Iwata, S. Yoshina, Y. Suehiro, S. Hori, S. Mitani, Engineering new balancer chromosome somes in C. elegans. Plant J. 109, 93–105 (1998).

57. K. Liu, H. Jin, B. Zhou, Genetic lineage tracing with multiple DNA recombinases: A user's guide for conducting more precise cell fate mapping studies. J. Biol. Chem. 295, 6413–6424 (2020).

58. K. Anastassiadi, S. Glaser, A. Kranz, K. Bernhardt, A. F. Stewart, "A practical summary of site-specific recombination, conditional mutagenesis, and tamoxifen induction of CreERT2" in Methods in Enzymology, P. M. Wassarman, P. M. Soriano, Eds. (Academic Press, 2010), vol. 477, pp. 109–123.

59. S. M. Dymecki, R. S. Ray, J. C. Kim, “Mapping cell fate and function using recombinase-based intersectional strategies” in Methods in Enzymology, P. M. Wassarman, P. M. Soriano, Eds. (Academic Press, 2010), vol. 477, pp. 183–213.

60. S. L. Edwards et al., A novel molecular solution for ultraviolet light detection in Caenorhabditis elegans. PLoS Biol. 6, e118 (2008).

61. A. Ward, J. Liu, Z. Feng, X. Z. S. Xu, Light-sensitive neurons and channels mediate phototaxis in C. elegans. Nat. Neurosci. 11, 916–922 (2008).

62. N. J. Tourasse, J. R. M. Millet, D. Dupuy, Quantitative RNA-seq meta-analysis of alternative exon usage in C. elegans. Genome Res. 27, 2120–2128 (2017).

63. M. A. Allen, L. W. Hillier, R. H. Waterston, T. Blumenthal, A global analysis of C. elegans trans-splicing. Genome Res. 21, 255–264 (2011).

64. K. Dejima et al., Auxin-induced, SCF(TIR1)-mediated trans-splicing. Genes Dev. 24, 141–148 (2010).

65. K. Dejima et al., Auxin-induced, SCF(TIR1)-mediated trans-splicing. Genes Dev. 24, 141–148 (2010).

66. S. Brenner, The genetics of C. elegans. Genetics 77, 71–94 (1974).

67. C. Frøkjær-Jensen et al., Random and targeted transgene insertion in Caenorhabditis elegans using a modified Mos1 transposon. Nat. Methods 11, 529–534 (2014).

68. R. Hunt-Newbury et al., High-throughput in vivo analysis of gene expression in Caenorhabditis elegans. PLoS Biol. 5, e237 (2007).

69. K. S. Ghanta, C. C. Mello, Melting dsDNA donor molecules greatly improves precision genome editing in Caenorhabditis elegans. Genetics 216, 643–650 (2020).

70. M. Linkert et al., Metadata matters: Access to image data in the real world. J. Cell Biol. 189, 777–782 (2010).

71. J. Schindelin et al., Fiji: An open-source platform for biological-image analysis. Nat. Methods 9, 676–682 (2012).

72. X. Karp, I. Greenwald, Post-transcriptional regulation of the E/Daughterless ortholog HLH-2, negative feedback, and birth order bias during the AC/VU decision in C. elegans. Develop. Biol. 105, A105 (1988).

73. K. Liu, H. Jin, B. Zhou, Genetic lineage tracing with multiple DNA recombinases: A user's guide for conducting more precise cell fate mapping studies. J. Biol. Chem. 295, 6413–6424 (2020).

74. K. Anastassiadi, S. Glaser, A. Kranz, K. Bernhardt, A. F. Stewart, "A practical summary of site-specific recombination, conditional mutagenesis, and tamoxifen induction of CreERT2" in Methods in Enzymology, P. M. Wassarman, P. M. Soriano, Eds. (Academic Press, 2010), vol. 477, pp. 109–123.