The Journal of Neuroscience, February 15, 1996, 16(4):1511–1522

Defects in Courtship and Vision Caused by Amino Acid Substitutions in a Putative RNA-Binding Protein Encoded by the *no-on-transient A* (*nonA*) Gene of *Drosophila*

Katherine G. Rendahl,[1] Natasha Gaukhshteyn,[1] David A. Wheeler,[2] Thomas A. Fry,[1] and Jeffrey C. Hall[1]

[1]Department of Biology, Brandeis University, Waltham, Massachusetts 02254, and [2]Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030

The *Drosophila no-on-transient A* (*nonA*) gene is involved in the visual behaviors and courtship song of the fly. The NONA polypeptide contains two copies of the RNA-recognition motif (RRM), a hallmark of proteins involved in RNA binding, and an adjacent conserved charged region. This 311-amino-acid region is found in four other proteins and largely overlaps the *Drosophila*-Behavior/Human Splicing (or DBHS) domain. The newest family member, *Drosophila nAhomo*, was discovered in a database search, and encodes a protein with 80% identity to NONA. In this study, three *nonA* mutations generated by chemical mutagenesis were sequenced and found to fall within the conserved region. Site-directed mutagenesis of the two RRMs, and within a (conserved) charged region located C-terminal to them, was performed to determine the significance of these domains with respect to whole-organismal phenotypes. Behavior and viability were assessed in transformed flies, the genetic background of which lacks the *nonA* locus. Point mutations of amino acid 548 in the charged region confirmed the etiology of the $nonA^{diss}$ courtship-song mutation and showed that a milder substitution at this site produced intermediate singing behavior, although it failed to rescue visual defects. Mutagenesis of the RRM1 domain resulted in effects on viability, vision, and courtship song. However, amino acid substitutions in RNP-II of RRM2 led to near-normal phenotypes, and the *in vivo nonA* mutations located in or near RRM2 caused visual defects only. Thus, we suggest that the first RRM could be important for all functions influenced by *nonA*, whereas the second RRM may be required primarily for normal vision.

Key words: visual behavior; electroretinogram; courtship song; RNA-binding protein; DBHS domain; in vitro *mutagenesis*

no-on-transient A (*nonA*) is a gene required for normal visual behaviors and the wing vibrations produced by males during courtship in *Drosophila melanogaster*. The courtship-song sounds are thought to be important sensory cues involved in species recognition and in sexually stimulating the female (Hall et al., 1980, 1990; Hall, 1994). Jones and Rubin (1990) demonstrated the rescue of the *nonA* visual defect with DNA cloned from the *nonA* region. This genomic fragment was later shown to rescue the courtship-song abnormality caused by $nonA^{diss}$ (Rendahl et al., 1992) and the viability deficit associated with a *nonA*-null deletion (Stanewsky et al., 1993).

nonA encodes a protein containing two RNA-recognition motifs (RRMs) (Besser et al., 1990), a ~80–90 amino acid consensus sequence (Bandziulis et al., 1989; Kenan et al., 1991). Within the RRM are highly conserved RNP-II hexamer and RNP-I octamer motifs, which are thought to contact RNA ligands (for review, see Burd and Dreyfuss, 1994). Recent alignment of NONA and novel sequences resulted in the identification of a conserved charged region adjacent to the RRM domains (Dong et al., 1993).

RNA-binding proteins perform a wide variety of functions including capping, polyadenylation, constitutive and alternative splicing of pre-mRNAs, as well as transport to the cytoplasm, stabilization, and translation of mRNAs. We propose that NONA is involved in some aspect of pre-mRNA processing, perhaps with a primary function in the nervous system. Several RRM-encoding genes with demonstrated splicing functions have been identified in *Drosophila* (for review, see Rio, 1993). A subset of these is hypothesized to be specifically required for mRNA splicing within the nervous system (for review, see Yao et al., 1993).

Mutations in RNA-binding proteins have been implicated in *Drosophila* nervous system defects and in human diseases of the nervous system. For example, genetic variations at the *elav* and *cpo* loci of *Drosophila* lead to behavioral or (visual system) physiological defects (Campos et al., 1985; Homyk et al., 1985; Bellen et al., 1992). Human gene products similar to ELAV, called HUD, HUC, and HELEN (Szabo et al., 1991), have been identified, and patients with autoimmune sera generated against these proteins show neural disorders (Graus et al., 1985). In addition, the *FMR1* gene encodes the fragile-X protein containing two motifs common to RNA-binding proteins (Gibson et al., 1993; Siomi et al., 1993). Mutations at this chromosomal site are relatively common and result in mental retardation (Richards and Sutherland, 1992). Thus, elucidating the interactions of NONA and related RNA-binding proteins with their target ligands may have general importance in understanding the development and function of nervous systems in a variety of organisms.

The present molecular studies were initiated to dissect the functional domains of *nonA*, because of the intriguing influence of this gene on a restricted array of behaviors. We investigated whether the NONA open reading frame (ORF) contains distinct

Received Aug. 23, 1995; revised Oct. 27, 1995; accepted Nov. 11, 1995.

This work was supported by National Institutes of Health Grant GM-21473 and the Human Frontier Science Program. We thank Michael J. Palazzolo, Michael Rosbash, Ralf Stanewsky, and Harald Saumweber for discussions, and Donald C. Rio and Ralf Stanewsky for comments on this manuscript.

Correspondence should be addressed to Jeffrey C. Hall, Department of Biology, 235 Bassine Building, Brandeis University, 415 South Street, Waltham, MA 02254-9110.

Dr. Rendahl's present address: Somatix Therapy Corporation, Alameda, CA 94501.

Table 1. Percent identities among NONA and related RRM-containing proteins

	dr-nonA_I	mu-nonO	hu-p54^{nrb}	dr-nAhomo
hu-PSF	44	71	72	42
dr-nonA_I		42	43	80
mu-nonO			99	41
hu-p54^{nrb}				41

These percent identities within the conserved region were derived from the 311-amino-acid region that spans the two RRMs and an adjacent region of relatively high charge density (see legend to Fig. 2*C*). The amino acid sequence numbers for a given protein are shown in Figure 2*A*. NONA I of *Drosophila* refers to the alternative-splice form that contains all the known biological activity encoded by this locus (Rendahl et al., 1992). The two insertion/deletions in the alignment in Figure 2*A* were a gap in the *Drosophila* NONA homolog (dr-nAhomo) at position 103, an Ala insertion in NONA itself (dr-nonA_I), and a Thr insertion in dr-nAhomo) at position 194, relative to the other sequences. The probabilities that the top-ranking identities are similar to NONA by chance, given the length and composition of the query sequence and the size and relative amino acid frequencies of the database, ranged from 6.9×10^{-286} for NONA itself, with a BLAST score of 1414 (cf. Altschul et al., 1990), to 3.0×10^{-78} for the protein from mouse (mu-nonO) that is similar to NONA, with a score of 421.

regions required for wild-type behaviors and viability. Here, we show that three *nonA* alleles contain point mutations within the conserved RRM and charged domains. In addition, novel mutations in these regions were generated by site-directed mutagenesis, and the resulting transgenic flies were assayed for behavioral, physiological, and adult-emergence phenotypes.

MATERIALS AND METHODS

Sequence analysis. The NONA (form I) protein sequence (Genbank accession number M33496) was filtered using the "XNU" program (Claverie and States, 1993) to remove regions of short, repetitive amino acids. The nonredundant nucleic acid database at the National Center for Biotechnology Information (National Library of Medicine, Bethesda, MD) was searched using the TBlastN algorithm (Altshul et al., 1990). The top eight hits in the database matched NONA with a score of 400 or greater. Two of these were self-hits to *nonA* types I and II (cf. Jones and Rubin, 1990). Another (accession number X16850) is a cloning artifact [i.e., is defined as human mRNA for myoblast cell-surface antigen but is now known to be a fragment of polypyrimidine tract-binding protein-associated splicing factor (psf); see below]. Two others, L14459 and U02493, were also duplicates of a *nonA*-related sequence from humans. The sequences with scores greater than 400 were NONA itself of *Drosophila* (accession number M33496), the NONA homolog from this species (nAhomo in Table 1 and Fig. 2, L32750), an RRM-containing protein from human (hu-p54^{nrb}: L14599), a mouse gene similar to NONA (nonO: S64860), and an RNA-splicing factor from human (hu-psf: X70944). These were aligned (see Fig. 2) using the Pattern Induced Multiple Alignment algorithm (Smith and Smith, 1990).

Cloning and sequencing of nonA *mutants.* Nearly all procedures were performed according to Sambrook et al. (1989). Genomic DNA containing the five exons of the *nonA^{diss}*, *nonA^{H2}*, and *nonA^{H14}* alleles was amplified by PCR with the following exon-flanking oligonucleotide primers (containing *Eco*RI or *Hin*dIII recognition sequences at their 5′ ends):

exon 1
5′-AAA-GAA-TCC-TTG-TAG-CAA-GCA-CAT-3′
5′-AAA-AAG-CTT-ATT-AAG-TCG-CCG-CGA-3′

exon 2
5′-AAA-GAA-TTC-CTA-AAA-TCA-AAC-TTT-3′
5′-AAA-AAG-CTT-GTA-TGT-TAA-TAA-TTT-3′

exon 3
5′-AAA-GAA-TTC-CTA-AAA-TCA-AAC-TTT-3′
5′-AAA-AAG-CTT-CGA-GCA-TAT-ATA-TTA-3′

exon 4
5′-AAA-GAA-TTC-TCG-ATG-ATC-AGC-TAA-3′
5′-AAA-AAG-CTT-GTA-TGT-TAA-TAA-TTT-3′

exon 5
5′-AAA-GAA-TCT-ACG-GCT-AAT-GCA-ACA-3′
5′-AAA-AAG-CTT-CAA-ATC-TCA-TCC-TTA-3′
5′-GGT-TTC-TCC-GGA-GCT-CCG-3′
5′-AAG-AAC-TTT-ACA-TTC-CTA-3′
5′-GGT-CAA-GGA-TTC-AGA-GGT-CG-3′
5′-GCT-TCC-CTT-CGA-CCG-TGT-CT-3′
5′-TGC-GGA-CCT-CCG-GAC-AAA-CC-3′
5′-GAT-AAT-GGA-GCT-TAC-TTA-TGA-3′

The resulting amplification products were cloned and sequenced by the "dideoxy" method using standard procedures. The following additional primers were designed for sequencing, such that they spanned the clones containing the largest exon (#2):

For ease of identification of base changes in the mutant sequences, reaction mixes containing a particular dideoxynucleoside triphosphate were run side-by-side for all three alleles. After identification of base-pair changes that would lead to amino acid substitutions, the relevant exons were independently amplified, cloned, and sequenced to confirm that the putative mutations were real and not PCR-generated errors.

Site-directed mutagenesis of the 235R11 genomic fragment. Single-stranded template DNA was synthesized in preparation for *in vitro* mutagenesis, using the "*dut- ung-*" method of Kunkel et al. (1987) (see also Sambrook et al., 1989). Two types of mutations were generated in the *Apa*I/*Bam*HI fragment of the 235R11 genomic clone (see Fig. 3). The first class consisted of two point mutations at position 548: R548C and R548K. For the second type, the RRM mutations generated were Y305A N308A in the RNP-II domain of RRM1, K337A F339A in the RNP-I domain of RRM1, S381A N382A in RNP-II of RRM2, and E422A F423A in RNP-I of RRM2. Two primers were used to mutate the Arg residue at position 548 of exon 3, the putative location of the *nonA^{diss}* mutation: R548C, 5′-GCA-GAG-CTG-GCA-AAA-TGA-GAA-GTT-GAC-AAC-GAG-3′; R548K, 5′-GCA-GAG-CTG-GCA-AAA-AAG-GAA-GTT-GAC-AAC-GAG-3′. Additionally, the following primers were synthesized to mutate the four RNP regions of exon 2: RRM1, RNP-II (Y305A N308A): 5′-GGA-CGA-AAT-CGT-CTC-GCT-GTG-GGC-GCC-CTG-ACC-AAT-GAC-ATC-3′; RRM1, RNP-I (K337A F339A): 5′-TTC-TCG-AAC-CTG-GAT-GCG-AAC-GCT-ACA-TTC-CTA-AAG-GTC-3′; RRM2, RNP-II (S381A N382A): 5′-ACC-ATA-TTG-CGG-GTT-GCC-GCT-CTC-ACA-CCG-TTC-GTT-3′; RRM 2, RNP-I (E422A F423A): 5′-GGC-GAG-GGC-ATA-CTT-GCG-GCT-GCC-AAA-AAG-TCA-TCG-3′. All mutants were verified by dideoxy sequencing (Sanger et al., 1977).

Generation of constructs for in vitro *mutagenesis-related transformation.* The mutated *Apa*I–*Bam*HI fragments were then cloned into the wild-type 11 kb genomic fragment 235R11 in the shuttle vector pHSΔB, a derivative of the plasmid pHSS7 (Seifert et al., 1986) (F. Heffron, unpublished data), which contains a modified polylinker deleted of a *Bam*HI site (Jones and Rubin, 1990). The mutant fragments were then excised from the vector with *Not*I, a unique restriction site in the polylinker, and ligated to the appropriate vector cut with *Not*I. In the case of some of the RNP mutations, the fragments were cloned into the pDM30 vector marked with ry$^+$ (obtained from L. Gavis, Whitehead Institute, Cambridge, MA) (cf. Mismer and Rubin, 1987). A majority of the mutated fragments were cloned into the *Drosophila*-transformation vector (CaSpeR-4) marked with w^+ or "mini-*white*" (Pirrotta, 1988). The sequences of the mutated regions were reconfirmed at this step, and the overall integrity of the 235R11 fragment was checked by diagnostic restriction digests.

Transformations of Drosophila *embryos.* P-element-mediated transformation by microinjection of DNA into embryos was performed according to standard methods (Rubin and Spradling, 1982). For the generation of some transgenics, embryos were also injected with transposase-encoding helper plasmid: *ry^{506}* embryos were injected with *nonA* fragments cloned into the pDM30 vector (Mismer and Rubin, 1987), or *Df(1)w* embryos with the constructs made using pCaSpeR-4 (cf. Pirrotta, 1988). For other transformations, the P-element transposase-producing stocks P[(*ry*$^+$)Δ2–3 99B *ry^{506}*] or *y w*; P[(*ry*$^+$)Δ2–3 99B *ry^{506}*] (Robertson et al., 1988) were the sources of embryos to be injected.

Genetics. All genetic variations are listed in Lindsley and Zimm (1992) unless noted otherwise. Surviving G_0 embryos were crossed to the relevant host strain flies (*ry^{506}*, *Df(1)w*, or *y w*), and the G_1 offspring exhibiting transformed eye colors had their chromosomal insert positions determined by standard *Drosophila* crosses; these included "balancing" a given insert (in almost every transgenic strain; compare Table 2) with a dominantly marked, multiply inverted autosome. When only a single line of a given construct type was generated, the P-element was mobilized using the transposase-producing Δ2–3 line (cf. Robertson et al., 1988). For behavioral analysis, virgin females heterozygous for the *nonA*-null translocation/deletion (Fig. 1) were crossed to balanced transformant males, producing *nonA*-null (*nonA*$^-$) males carrying the transgene.

A minimum of two transgenic lines of each mutant type was generated. Southern analysis was performed on transformed adults to verify the

Table 2. Viability of *nonA* point mutation-containing transgenics

Transformant genotype	Viability in *T(1;4)9e2-10Y* genetic background (%)	Viability in *1(1)i19e²/Y* genetic background (%)
Controls		
235R11		
line 1	81 ($n > 200$)	112 ($n > 200$)
line 2	96 ($n > 200$)	100 ($n > 200$)
211S12		
line 1	16 ($n > 200$)	100 ($n > 200$)
line 2	25 ($n > 200$)	129 ($n > 200$)
diss-site mutations		
R548K		
line 1	55 ($n = 115$)	79 ($n = 110$)
line 2	66 ($n = 106$)	109 ($n = 157$)
R548C		
line 1	32 ($n = 114$)	86 ($n = 146$)
line 2	ND	ND
RRM mutations		
Y305A N308A		
line 1	15 ($n = 199$)	96 ($n = 124$)
line 2	44 ($n = 174$)	113 ($n = 120$)
K337A F339A		
line 1	21 ($n = 118$)	81 ($n = 126$)
line 2	11 ($n = 128$)	113 ($n = 115$)
S381A N382A		
line 1	102 ($n = 116$)	87 ($n = 137$)
line 2	88 ($n = 152$)	82 ($n = 138$)
E422A F423A		
line 1	15 ($n = 203$)	78 ($n = 180$)
line 2	10 ($n = 174$)	110 ($n = 160$)

Viability counts of emergents carrying single copies of normal or *in vitro* mutated *nonA*-containing constructs (compare Figure 1), in a *T(1;4)9e2-10/Y* (*nonA*-null) genetic background, are listed. In addition, viability scores are given for transgenic flies carrying a lethal allele of the *l(1)i19e* gene (which is very near *nonA*) as a control for basic expression levels associated with the relevant 11 kb genomic DNA fragment (see text). Viability data for the rescued control *T(1;4)9e2-10/Y*, P[(ry^+)235R11]/+ and *nonA*-null *T(1;4)9e2-10/Y*, P[(ry^+)211S12]/+ genotypes are taken from Stanewsky et al. (1993). To determine the viability of the *nonA* mutants, heterozygous female *T(1;4)9e2 10/In(1)FM7c,B* females (i.e., with the *nonA*$^-$ genotypes balanced by the multiply inverted, dominantly marked *FM7* X chromosome) were crossed to males carrying autosomally balanced transformant inserts with mutations in the 11 kb *nonA*$^+$ fragment (P[(ry^+)235R11]). The resulting *T(1;4)9e2-10/Y*; P[(ry^+)235R11 mutation]/+ male and *T(1;4)9e2-10/+*; P[(ry^+)235R11 mutation]/+ female emergents were scored. The viability of a given mutant genotype was then defined as the ratio of males to females, expressed as a percentage. Viability computations for rescue of *l(1)i19e²* flies by mutant transgenes were performed in a similar manner. Assessment of the viability data for line 2 of the R548C transgenic, whose P-element insertion is in the Y chromosome, could not be accomplished in a manner directly comparable with the other genotypes; hence "ND." However, counts were made of *T(1;4)9e2-10/Y*[P(ry^+)] flies, which eclosed at a frequency of 140% relative to *FM7a/Y* [P(ry^+)] flies.

presence of a single-copy P-element insertions and also that P[(ry^+)Δ2–3] was not present (data not shown). The results of these blots indicated the presence of single-copy P-element insertions, except for the transformant RRM2, RNP-II/S381A N382A, line 2; this appeared to contain two insertions. However, no significant differences were detected when comparing the phenotypes of this line with those of a line containing one copy of this particular transgene.

Immunochemistry. Ten fly heads per genotype were homogenized in 50 μl of buffer, containing 50 mM Tris, pH 8.0, 50 mM sodium pyrophosphate, 0.2% Triton X-100, 1 mM phenylmethylsulfonyl fluoride, and 0.6 mM leupeptin (Sigma, St. Louis, MO). An equal volume of 2× Tris glycine SDS loading buffer (Novex, San Diego, CA) containing β-mercaptoethanol (Sigma) was added, and the sample was heated to 100°C for 2 min and spun to remove debris; 20 μl of the supernatant was

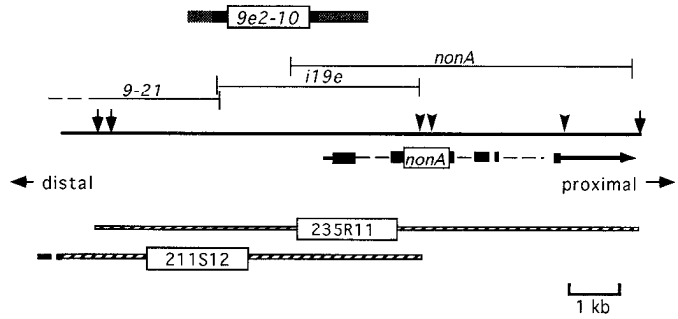


Figure 1. Genetics, cytogenetics, and cloned DNA from the *nonA* region of the X chromosome of *D. melanogaster*. The map at the *top* shows (*black bar*) the extent of deleted X chromosome material in the *T(1;4)9e2–10* translocation; this chromosome aberration is missing portions of the *nonA* gene and the two vital loci, *l(1)i19e* and *l(1)9–21*, that are distal to (i.e., farther from the X-chromosomal centromere than) *nonA* (Stanewsky et al., 1993). The exons of the latter gene (*black boxes*) and introns (*dashed lines*), as well as the direction of its transcription, are depicted in the *middle* of the figure. The 11 kb *Eco*RI fragment 235R11 shown near the *bottom* was cloned from this genomic region and contains DNA that rescues the lethality associated with the *l(1)i19e* locus as well as *nonA*-associated defects; the 12 kb *Sal*I fragment 211S12 rescues only *l(1)i19e*-associated lethality (Jones and Rubin, 1990; Stanewsky et al., 1993). *Vertical arrows* indicate *Eco*RI sites, and *arrowheads* indicate *Sal*I sites. Scale bar, 1 kb.

run on an 8% polyacrylamide gel (Novex). The gel was transferred to nitrocellulose, nonspecific binding to the membrane was blocked with 5% milk in PBS, and it was incubated at room temperature for 1 hr with affinity-purified NONA polyclonal antibody (cf. Rendahl et al., 1992) diluted 1:10 in 5% milk. The blot was washed with PBS plus 0.1% Triton-X and incubated for an additional 1 hr with a goat anti-mouse secondary antibody conjugated to horseradish peroxidase (Bio-Rad, Hercules, CA). NONA immunoreactivity was visualized with the enhanced chemiluminescence nonradioactive detection system (Amersham, Arlington Heights, IL).

Phenotypes of transgenics. Viability testing was carried out by counting segregants from genetic crosses as described in Stanewsky et al. (1993). Monitoring of optomotor responses to rotating vertical stripes was carried out by the "line-cross-counting" procedure described by Greenspan et al. (1980) and Kulkarni et al. (1988), with minor modifications (see Rendahl et al., 1992). Electroretinograms (ERGs) were recorded by delivering light pulses to immobilized flies according to the methods of Pak and Grabowski (1978) and Heisenberg and Wolf (1984), with minor modifications (see Rendahl et al., 1992). For analysis of courtship-song phenotypes, LIFESONG software (Bernstein et al., 1992) was used to select pulses interactively. After transfer of the songs and pulse locations to a VAX computer, song-analysis software was used to generate the cycles per pulse (CPP) regression scores (cf. Bernstein et al., 1992; Rendahl et al., 1992). Significance testing (of putative differences in song parameters among genotypes) was as described in Stanewsky et al. (1993).

RESULTS

Database search for NONA-related sequences

Two categories of sequences significantly related to NONA were found in the database. The lower-scoring set (BLAST score of <150; see Materials and Methods) contained various members of the RRM family. The higher-scoring set of five sequences (BLAST score of >400) matched not only the two RRMs of NONA, but also an adjacent region characterized by a high proportion of charged residues. This defines a common 311-amino-acid region, largely overlapping what has been termed the *Drosophila* Behavior/Human Splicing domain (or DBHS) (Dong et al., 1993). Three of these high-scoring sequences are psf (Patton et al., 1993), HeLa-cell-derived p54^{nrb} (Dong et al., 1993), and a non-POU-containing, octamer-binding protein called nonO

(Yang et al., 1993). psf, the protein of known function with the highest sequence homology to NONA, is an essential splicing factor *in vitro* (Patton et al., 1993; Gozani et al., 1994). We include in our alignment the predicted translation product of an ORF within a novel *Drosophila* P1 clone (M. Palazzolo, unpublished material), here designated nAhomo (Table 1, Fig. 2) of unknown function. This is the second *Drosophila* sequence identified that contains the 311-amino-acid conserved region.

The alignment of this group of highly related sequences is shown in Figure 2*A* and is illustrated schematically in Figure 2*C*. The positions of the α helices, β sheets, and loops that make up the secondary structure of the RRM domain are also shown in Figure 2*A*. A notable feature common to all five sequences is the degree of identity in loop 3 of RRM2, which is highly variable within the RRM family. A small number of residues within this loop has been shown to be necessary and sufficient for the sequence specificity of ligand binding of at least some mRNAs (Kenan et al., 1991, Birney et al., 1993). The relatedness of both RNP-I domains to the consensus sequence found in the Prosite database (see Fig. 2*B*) is indicated above the sequence alignment.

Within the conserved region, there were two sites of insertion/deletions, each involving a single amino acid. The percent identity of each pair of proteins, not including the two "gapped" positions, is shown in Table 1. The sequences can be divided into two groups, based on their relatedness. Comparison of any two of the mammalian sequences—hu-psf, hu-p54^{nrb}, or mu-nonO—reveals that they are 71–99% identical, whereas the two *Drosophila* sequences—for NONA (dr-nonA_I) and nAhomo—are 80% identical. The N- and C-terminal sequences flanking the conserved domain shown in Figure 2*A* are quite divergent among the five sequences. Only hu-p54^{nrb} and nonO are strongly related, with >95% identity in their terminal regions.

Point mutations identified in existing mutants

Cloning and sequencing the coding regions of the *nonA^{H2}*, *nonA^{P14}*, and *nonA^{diss}* mutants uncovered a first- or second-position point mutation in each case (Fig. 3). *nonA^{P14}* contains a GGT → GAT codon change at nucleotide (nt) 4407, leading to a Gly to Asp point mutation at amino acid 400. Sequencing of *nonA^{H2}* revealed a GGA → AGA change at nt 4596, resulting in a Gly → Arg amino acid change at position 463. *nonA^{diss}* was found to contain a CGC → TGC change at nt 7434 (amino acid 548), encoding a Cys rather than the wild-type Arg. The alterations in the *non^{P14}* and *nonA^{H2}* visual mutants are in the center of and just outside the second RRM repeat, respectively; the latter is close enough to this RRM that it could affect ligand binding (see Kenan et al., 1991). The mutation in *nonA^{diss}*, which has both song and visual defects, lies in the charged region of NONA.

Novel mutations generated in the NONA ORF

Two single mutations, R548C and R548K, were made to test the role of amino acid 548 of the charged region (see Fig. 3). The Arg → Cys substitution at position 548, the site of the *dissonance* mutation, was made to determine whether a mutation at this position was the cause of its dual mutant phenotype. An Arg → Lys substitution was also made, to test the effects of an amino acid with charge properties similar to that of the normal residue at this position. A second set of mutations, Y305A N308A, K337A F339A, S381A N382A, and E422A F423A, was designed to disrupt hypothesized interactions of the RNP domains of the RRMs with RNA ligands. Each consists of a pair of substitutions of the nonpolar residue Ala for the wild-type aromatic or polar residues in one of the four RNP domains.

Immunochemistry of NONA in transgenic adults

Flies carrying the above six mutations in the P[(*ry*$^+$)235R11] fragment in the *nonA* null genetic background were analyzed by Western blotting (see Fig. 4). The translocation *T(1;4)9e2–10/Y* removes the function of *nonA* and a nearby vital gene, *l(1)i19e*; the P[(*ry*$^+$)235R11] transgene restores *nonA* and *l(1)i19e* functions, whereas P[(*ry*$^+$)211S12] restores only the latter (Fig. 1). The *T(1;4)9e-10/Y*, P[(*ry*$^+$)235R11]/+ control showed normal levels of NONA protein, whereas *T(1;4)9e-10/Y*, P[(*ry*$^+$)211S12]/+ is a protein null. All of the mutant lines produced quasinormal levels of full-length protein (cf. Rendahl et al., 1992), with the exception of RRM2, RNP-I/E422A F423A; no NONA was detectable in extracts of these transgenic adults. In addition, the R548C and RRM1, RNP-II/Y305A N308A mutants showed decrements in protein levels.

Viability of *in vitro* mutated *nonA* transgenics

The relative percentages of emerging *nonA*$^-$ adults carrying a normal *nonA*$^+$ transgene [*T(1;4)9e-10/Y*, P[(*ry*$^+$)235R11]/+

→

Figure 2. Amino acid sequence alignment and schematic of NONA and related RRM-containing proteins. *a*, The sequences are as follows: dr-nonA_I, dr-nAhomo, hu-psf, hu-p54^{nrb}, and mu-nonO. The abbreviations in the sequence names are as follows: *dr*, *Drosophila; hu*, human; *mu*, murine. The *numbers* at the top of each sequence indicate positions within the alignment; the *tick marks* indicate every tenth position. The *bold lines* above the alignment denote the two RRMs shared by this subfamily (note the first RRM continues to position 101). The *boxes within the bold lines* mark the position of the canonical RNP *I* and RNP *II* motifs. For the RNP-I box, columns in which every sequence matches the Prosite database pattern for RNP-I (see *b*) are labeled with a *multidot*; those in which at least one of the sequences matches are labeled with the letter "*o*"; nonmatches are *underscored*. The consensus line was formed by placing a letter under any column in which at least 4 out of 5 residues are identical. Identical residues in a given column are *blanked out*, whereas nonconsensus residues are in *lower case*. Gaps are indicated by a *multidot*. Below the consensus line, the hypothetical secondary structure, modeled after Birney et al. (1993), is given: *b*, β-sheet; *l*, loop; *a*, helix; these components are *numbered* below. The position of both RNP-I and the β sheet 3 of RRM1 is shifted to the left by two residues, relative to the sequence alignment in Birney et al. (1993). This improves the overall fit of these 5 family members to the RNP-I pattern. *b*, The Prosite database *RNP I* pattern is shown (Bairoch and Buchner, 1994); given that, within the RRM, the octamer RNP-I is generally the most conserved motif. The pattern is read as 8 separate positions, each separated by a *dash*. Residues within *brackets* are alternative matches to the pattern at a given position; residues in *braces* are nonmatches, i.e., all other residues match; *X* signifies any residue that matches the pattern; *single letters* mean only the indicated residue matches the pattern. *c*, The sequence alignment is represented schematically to emphasize the relative location of the conserved domain and landmarks within each sequence shown in *a*. *Dark shaded strips* correspond to RNP-I and RNP-II motifs within the RRM (shown in *white*). Although the entire conserved portion common to these sequences has a relatively high charge, the charge density varies between the RRM region and the *cross-hatched region* to the *right* of the RMMs. The aligned positions in the latter (cf. *a*) have 0.3 charged residues per position, whereas the RRM and *small cross-hatched region* to the *left* of it have 0.2 charges per position. These values are expressed in units of "charge per aligned position," by which for example, 0.2 means that 2 out of every 10 positions in the alignment have a charged residue in each of the 5 sequences. The *scale bar* marks off 100 amino acid segments of these polypeptides; in this regard, the entireties of these gene product range from ~450 to 700 residues in length (thus, including the nonconserved portions of them, designated by *thin lines*).

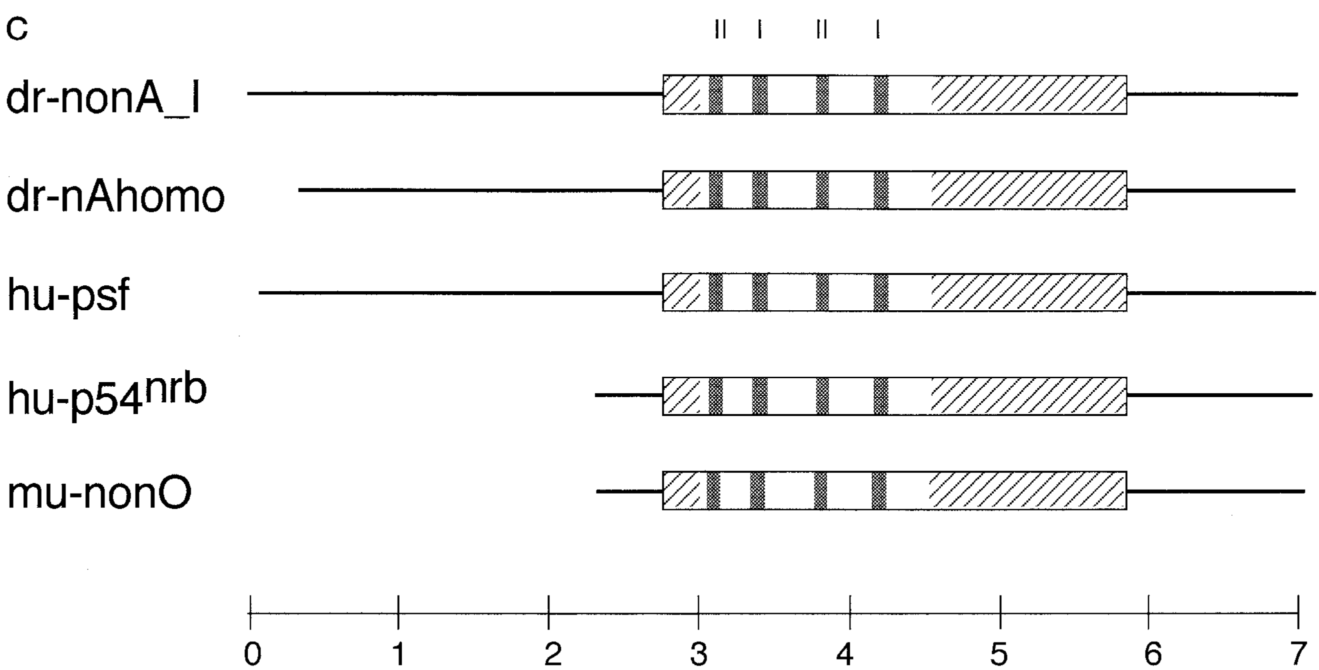
A

```
                                        RNP II                 RNP I
                1                                                                       100
dm-nonA_I   274 t rlrsis p fe epvev t tk sg n  y    tn   dd l em kp  eis i snld n t lkvdyhpn  k  ra  gsmrk r      p  373
dm-nAhomo   249 a rlldis p he ppiel tdnk vg n  y    ts t dddl em kp  ei di snpe n t lr dyyqn  k  ra  gslrk rv     p  348
hu-psf      269 eekisdse fkan sllrr g ktytq c  f    pa   ed fkrl ak  ep  v inkg g g ik esral  i  ae  dtpmr r      t  368
hu-p54_nrb   47 q assqne l id knfrk g kt tq s  f    pp   ee m kl ek  ka  v ihkd g g ir etrtl  i  ve  nmplr k      c  146
mu-nonO      48 q assqne l id knfrk g kt tq s  f    pp   ee m kl ek  ka  v inkd g g ir etrtl  iv ve  nmplr k      c  147
Consensus       -Q------G-T--L-----P-E--F--R-RL-VGNL--DIT--E-R--F--YG--GE-F----K-F-F--L-----AE-AK--LD-----G-QLRVRFA-
                                           bbbbbbllllllaaaaaaaaalllllbbbbbblllllbbbbbbblllllaaaaaaaaaalllllllbbbbbb
                                           1     1    1       2    2     3  3      4   2         5    5

                   RNP II                RNP I
                101                                                                     200
dm-nonA_I   374 nati r s  tpf      yks ei  pi   sit     khm e    ak ss sac rm n kc f  asl  cl d  evn  td     afnk m  473
dm-nAhomo   349 na.ivr t  nqf      hqs ei  pi    ic     kht e    ak ss sac rl n kc f  asl  cl    evnn nd   d t nk s  447
hu-psf      369 haaa s r  spy      eea sq  pi    vi     rst k    as pa rkafer s gv l  ttp  vi   leql  ed     . aq n  467
hu-p54_nrb  147 hsas h r  pqy      eea sv  qv    vi     rps k    sg pa rka dr s gs l  tfp  vt    dql  ee     . vi n  245
mu-nonO     148 hsas t r  pqy      eea sv  qv    vi     rps k    sg pa rka dr s gs l  tfpw vt    dql  ee     . vi n  246
Consensus       ----L-V-NL---VSNELL---F--FG--ERAV--VDDRG---G-GIVEF--K--A---L--C-E--F-LT---RP--VEPM---DD--GLPEK-L--K-
                   bbbbbbllllllaaaaaaaaalllllbbbbbbllllllllllbbbbbbbllaaaaaaaaaaallllllllllbbbbbb
                   1    1    1       2    2    3        3     4 2         5        5

                201                                                                     300
dm-nonA_I   474 pd nq  sig    dpn   h  gs   q hnlf tk  alk el meed   aq  y  y qete l    rk evdne kklewemrekqaeemrk e  573
dm-nAhomo   448 le rh  svg    cln   h  gs   q hdlf sk  slk el meed   aq  y  y qete l  e kk eldne  klewemrekqaeeirk e  547
hu-psf      468 pmyqk  etp    qhgt  y  sq   s deme qqreqveknm dakd   se  d yh hqan l    mr qeelr  eelhnqemqkrkemql q  567
hu-p54_nrb  246 qq hk  eqp    qpg   y  am   a ieme qq  qvd ni eare   me  a  h hqvm m    mr qeelr  eelhnqevqkrkqlel q  345
mu-nonO     247 qq hk  eqp    qpg   y  am   a ieme qq  qvd ni eare   me  a  h hqvm m    mr qeelr  eelhnqevqkrkqlel q  346
Consensus       --F--ER---PRFA---SFE-EY--RWK-L----K--QD---R--K----KLE--ME-AR-E----L-RQDL--R-----RM---------------R-

                301
dm-nonA_I   574  tm  hqt     584
dm-nAhomo   548  ymh yqnql   558
hu-psf      568  er  ree     578
hu-p54_nrb  346  er  ree     356
mu-nonO     347  er  ree     357
Consensus       EE--RR---EM
```

B

RNP I Pattern: [RK]-G-{EDRKHPCG}-[AGSCI]-[FY]-[LIVA]-x-[FYM]

C



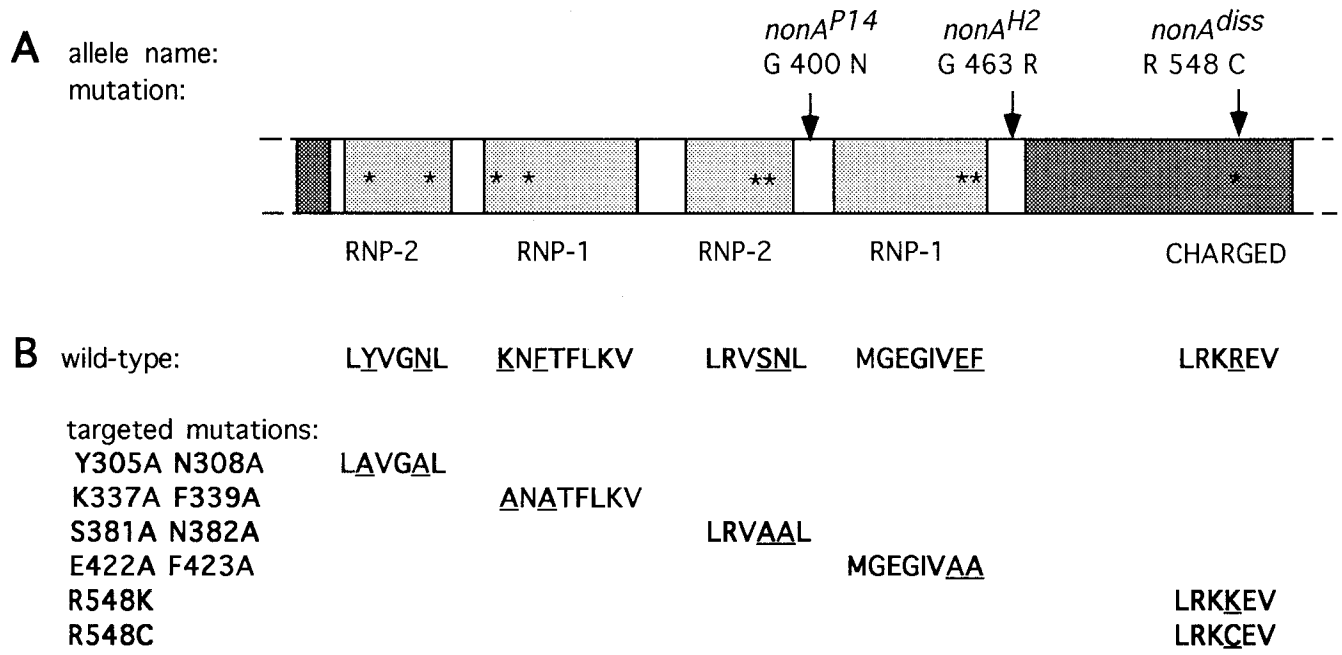


Figure 3. *nonA* mutants produced by chemical and site-directed mutagenesis. *A*, Enlarged view of the conserved region of the NONA protein. The RRM region is shown as a *white box*, divided by *gray-shaded RNP-2* and *RNP-1* domains within it; the adjacent, *dark-shaded* pair of blocks encodes the "charged region," which is located farther from the N terminus than is the RRM region. The positions of the amino acid changes in the $nonA^{P14}$, $nonA^{H2}$, and $nonA^{diss}$ mutants are indicated by *arrows*. Point mutations produced by site-directed mutagenesis are shown by *asterisks*. *B*, Below, the wild-type sequences of the RNP-II hexamers and RNP-I octamers are given, along with the four pairs of alanine substitutions in the RRMs. The corresponding wild-type residues of the consensus sequence are *underlined*. Within the charged region, mutations at amino acid 548 and the surrounding wild-type residues are indicated. The following list specifies the positions of the mutations within the RRM domains with reference to NONA amino acids (numbered from the N terminus; Jones et al., 1990) and to the convention of numbering the residues of such a domain from 1 to 80 (cf. Kenan et al., 1991): Y305A N308A = Y13A N16A; K337A F339A = K52A F54A; S381A N382A = S15A N16A; E422A F423A = E58A F59A.

ranged from 81 to 96% (Table 2) (cf. Stanewsky et al., 1993). The viability of the *nonA*-null transgenic type (*T(1;4)9e2–10/Y*, P[(ry^+)211S12]/+) ranged from 16 to 25% (Table 2) (cf. Jones and Rubin, 1990; Rendahl et al., 1992; Stanewsky et al., 1993). Viabilities ranged from 10 to 102% for novel mutations in the 11 kb DNA fragment [P(ry^+)235R11] introduced into the $nonA^-$ background *T(1;4)9e2–10/Y* (Table 2). Flies containing *nonA* transgenes with mutations at the *dissonance* site (residue 548) exhibited intermediate to low viabilities (Table 2). Those carrying the R548K mutation exhibited mildly subnormal viability, eclosing as adults at 55–60% of the expected frequencies. In contrast, flies containing the single-autosomal R548C insertion had low viabilities, eclosing at 32% of the expected number. Flies from a second line, bearing an R548C insertion on the Y chromosome, exhibited higher viability, although this value could not be measured in a genetic cross that precisely paralleled the others (see legend to Table 2). Flies carrying RRM mutations fell into two distinct classes. They showed either normal viabilities (88–102%) in the case of the S381A N382A mutations in the RNP-II portion of RRM2 (Table 2, compare Fig. 3), or low viabilities (10–44%) for both lines of the other three transgenic types (Table 2).

Control crosses were performed to demonstrate that the mutant phenotypes observed in the transgenic animals are attributable to the point mutations introduced by the *nonA* transgene and not merely to low levels of the expression of this factor. Full rescue of *nonA*-null-associated viability decrements is effected by the (intrinsically normal) P[(ry^+)235R11] transgene (Fig. 1, Table 2) (cf. Stanewsky et al., 1993); the *nonA*-mutated forms of this transgene mediated similarly high levels of viability (Table 2) in flies the genetic background of which included a lethal mutation at only one of the relevant loci, *l*(1)i/19e (Fig. 1). This provides a general measure of transgene expression, assuming that the $l(1)i19e^+$ and $nonA^+$ genes are expressed equally well at a given chromosomal insertion site.

Behavioral and physiological analyses of *nonA* transgenics

The visual defects of the *nonA* mutants include a lack of transient spikes in the ERG, optomotor blindness, and a subtle phototaxis decrement (Hotta and Benzer, 1970; Pak et al., 1970; Heisenberg, 1971, 1972; Kulkarni et al., 1988). The first two of these phenotypes were examined in the novel transgenics. The light-on and light-off transient spikes of the ERG are characteristic of normal transmission of light-initiated signals to the first optic ganglion (Coombe, 1986). Robust transients were observed in the marker strain *y cv v f car/Y*, from which *T(1;4)9e2–10* was derived, and in controls carrying the $nonA^+$ transgene (*T(1;4)9e2–10/Y*, P[(ry^+)235R11]), but not in the negative control (*T(1;4)9e2–10/Y*, P[(ry^+)211S12]). The novel R548K and R548C mutants also lacked transient spikes. Flies containing mutations in the RRMs fell into two classes. All except for the RRM2, RNP-II/S381A N382A mutant type (see Fig. 3) lacked the normal light-on- and light-off-induced transient-spikes, whereas the depolarized light-coincident-maintained components of the ERG were robust in all cases. In contrast, both lines of the S381A N382A type exhibited a normal ERG (Table 3).

Optomotor responses of marker-bearing and rescued control genotypes (see above) illustrate the ability of these $nonA^+$ flies to move "correctly" in the direction of rotating vertical stripes; their scores ranged from 87 to 91, whereas *nonA*-null animals scored

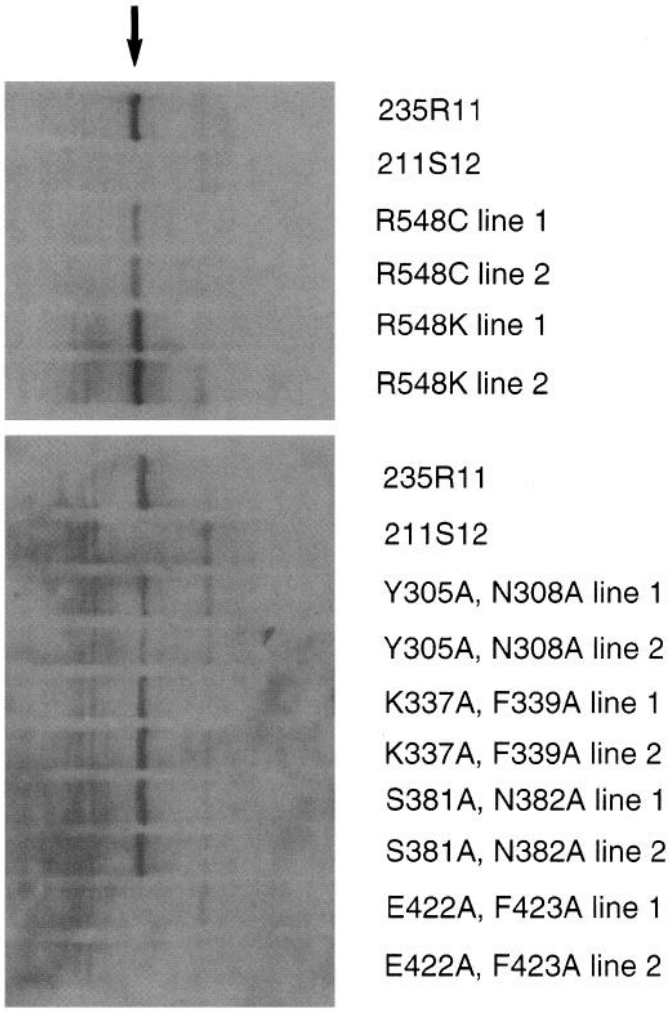


Figure 4. Western blots of transgenic strains. Immunoblotting of protein extracts made from 5 fly heads per genotype with a NONA polyclonal antibody. In the blot at the *top*, results from two control cases (genotypically $nonA^+$ = 235R11 and $nonA^-$ = 211S12, respectively), then two lines each of the *R548C* and *R548K* mutants in the $nonA^-$ genetic background. The specificity of the 70 kDa NONA polypeptide band (indicated by the *arrow*) is demonstrated by its absence in the null control (*second lane* from the *top*). The blot at the *bottom* shows (top-to-bottom) results from two control strains (as above), and two lines each of the following transgenic types (compare Fig. 3): RRM1, RNP-II = Y305A N308A; RRM1, RNP-I = K337A F339A; RRM2, RNP-II = S381A N382A; and RRM2, RNP-I = E422A F423A.

values of ~52 (Table 3). The former scores are near-maximal, whereas the latter reflects arbitrary movements, equally likely to be against the stripe rotation as in the correct direction. The optomotor performances (cf. Rendahl et al., 1992) of the novel transgenics in the $nonA^-$ genetic background were consistent with their ERG signals (Table 3, Fig. 5). All transgenic types except the S381A N382A were defective, lacking the normal optomotor response (Table 3).

The courtship-song defects in $nonA^{diss}$ and the *nonA*-null involve the pulse-train component of male singing behavior (for review, see Hall, 1994). In these mutants, pulses exhibit a strong tendency to become anomalously polycyclic over the course of a (<1 to ~2 sec) train. Courtship-song analysis of the transgenic flies containing novel mutations led to a range of phenotypes, from wild-type to severely mutant (Fig. 5). The *y cv v f car/Y* and *T(1;4)9e2–10/Y*, P[(ry^+)235R11] control animals gave low CPP regression scores (Table 3), reflecting very mild tendencies for increasing intratrain polycyclicity (i.e., as occurs in wild-type). The CPP regression scores of both $nonA^{diss}$ and *nonA*-null (*T(1;4)9e2–10/Y*, P[(ry^+)211S12]) males are ~10-fold higher than those of normal flies (Rendahl et al., 1992; Stanewsky et al., 1993). Pulse trains generated by $nonA^-$ males (Stanewsky et al., 1993) sometimes include extremely polycyclic pulses (>12 CPP; exemplified in Fig. 5*D*,*H*). Such pulses are rarely observed in the songs of the $nonA^{diss}$ point mutant, and they have never been documented in a wild-type male (the pulses of which typically contain 2–3 cycles). Instances of extreme polycyclicity are not properly reflected by CPP scores that fall within the mutant range (e.g., *dissonance*-like: ≥2), as is described below with regard to certain of the *in vitro* mutated transgenics.

Mutations at residue 548, the site of the *dissonance* mutation within the charged region of NONA, led to either intermediate or mutant values (Table 3), indicating a requirement for the wild-type Arg for production of normal song. The R548K lines yielded CPP regression scores of 0.66 and 0.68 (Table 3), which are approximately threefold higher than normal. These CPP scores are deemed intermediate, because they were found to be statistically different from values associated with both wild-type and *dissonance* or the null mutant (see legend to Table 3). This potentially interesting result could be viewed as weakened by the fact that the control values listed are from a previous investigation (see legend to Table 3); however, the relevant "song scores" (i.e., in $nonA^+$ vs those resulting from previously mutated forms of the gene) have been quite consistent from study to study (Rendahl et al., 1992; Stanewsky et al., 1993; Krejci et al., 1994). Thus, we conclude that the Lys point mutation in the *nonA* transgene, although similar in charge to the wild-type Arg, evidently does not fully substitute for it.

The more dramatic amino acid substitution of Cys for Arg, found in the $nonA^{diss}$ mutant and reproduced in R548C, resulted in a severe *nonA*-like singing phenotype, with CPP regression scores in the range of 2.28–3.31 (Table 3). These are statistically equivalent to the song scores associated with values (2.00–3.35) determined for males expressing the negative-control genotype. The transgenic flies carrying point mutations in the RRMs exhibited two classes of singing behavior. An exceptional case, the RRM2, RNP-II/S381A N382A mutant, sang normally, with a regression score of 0.02, which is statistically indistinguishable from normal. A second line gave a statistically intermediate score (Table 3), which could be attributable to a position effect of the transgene insertions (cf. Jones and Rubin, 1990; Krejci et al., 1994). The other mutant transgenic types, Y305A N308A, K337A F339A, and E422A F423A, exhibited high CPP regression scores similar to those of *nonA*-null males. As can be seen from illustrations of selected courtship-song traces (Fig. 5), males containing these RRM mutations routinely showed polycyclicity in their song pulses. Three of the *in vitro* mutated types generated occasional "wildly" polycyclic song pulses (as introduced above). For the Y305 N308A, K337A F339A, and E422A F423A transgenics, the proportions of males that produced one or more such pulse trains were 60, 20, and 20%, respectively; compared with the positive control (*T(1;4)9e2–10/Y*, P[(ry^+)235R11]), 0% and with the negative control (*T(1;4)9e2–10/Y*, P[(ry^+)211S12]), 40% (compare Table 3).

In summary, mutations of residue 548 within NONA led to a

Table 3. Behavioral and physiological analysis of *nonA* point mutation-containing transgenics

Transformant type	Optomotor score	ERG: on-transient	ERG: sustained component	ERG: off-transient	Song CPP regression score
Controls					
y cv v f car/Y	91.2 ± 5.1	0.92 ± 0.08	9.00 ± 0.69	3.08 ± 0.34	0.32 ± 0.06
235R11					
line 1	86.6 ± 3.4	1.25 ± 0.31	14.00 ± 1.04	2.67 ± 0.33	0.20 ± 0.15
line 2	89.5 ± 2.6	1.67 ± 0.26	9.58 ± 0.58	3.50 ± 0.65	0.22 ± 0.05
211S12					
line 1	52.0 ± 0.7	0.00 ± 0.00	12.33 ± 2.20	0.00 ± 0.00	2.00 ± 0.49*
line 2	51.8 ± 4.1	0.00 ± 0.00	16.90 ± 2.18	0.00 ± 0.00	3.35 ± 0.44*
diss mutations					
R548C					
line 1	48.5 ± 2.0	0.00 ± 0.00	10.17 ± 0.98	0.00 ± 0.00	3.31 ± 0.82*
line 2	49.5 ± 2.4	0.00 ± 0.00	20.33 ± 1.03	0.00 ± 0.00	2.28 ± 0.35*
R548K					
line 1	50.3 ± 4.1	0.00 ± 0.00	7.25 ± 0.83	0.00 ± 0.00	0.68 ± 0.12**
line 2	57.0 ± 1.9	0.00 ± 0.00	13.17 ± 1.36	0.00 ± 0.00	0.66 ± 0.08**
RRM mutations					
Y305A N308A					
line 1	47.0 ± 3.5	0.00 ± 0.00	11.67 ± 0.85	0.00 ± 0.00	2.83 ± 0.52*
line 2	52.8 ± 2.3	0.00 ± 0.00	12.96 ± 1.18	0.00 ± 0.00	1.90 ± 0.48*
K337A F339A					
line 1	54.6 ± 8.0	0.00 ± 0.00	14.04 ± 2.34	0.00 ± 0.00	1.50 ± 0.24*
line 2	63.9 ± 4.9	0.00 ± 0.00	11.73 ± 1.23	0.00 ± 0.00	1.39 ± 0.39*
S381A N382A					
line 1	91.8 ± 3.9	1.67 ± 0.14	18.42 ± 0.75	5.92 ± 0.23	0.66 ± 0.14**
line 2	87.5 ± 2.4	2.25 ± 0.22	18.33 ± 1.08	6.33 ± 0.19	0.02 ± 0.11
E422A F423A					
line 1	49.6 ± 2.6	0.00 ± 0.00	12.92 ± 1.54	0.00 ± 0.00	1.78 ± 0.32*
line 2	49.1 ± 4.5	0.00 ± 0.00	15.17 ± 1.23	0.00 ± 0.00	2.19 ± 0.61*

Optomotor, ERG, and song CPP regression scores are given for control flies and for six pairs of transformant lines tested in the *nonA*-null [*T(1;4)9e2–10/Y*] genetic background. The normal and (*nonA*$^+$-rescued) control flies were *y cv v car f/Y* and *T(1;4)9e2–10/Y*, P[(*ry*$^+$)235R11]/+ males, respectively. Both of these gave wild-type-like optomotor responses, robust "on" and "off" ERG transients, and low CPP regression scores (associated with pulse sounds within courtship song trains). The negative control is the *nonA*-null transgenic *T(1;4)9e2–10/Y*, P[(*ry*$^+$)211S12]/+ (compare Figure 1). Data for control genotypes are from Stanewsky et al. (1993). For the experimentals, 4 flies were behaviorally monitored and 4 separate ones physiologically recorded within each genotype; 5 males of each genotype had their courtship songs recorded; all values tabulated are ±SEM. *The transgenic lines with CPP regression scores that are statistically different from the behavior of the normal *T(1;4)9e2–10/Y*, P[(*ry*$^+$)235R11]/+ type and similar to *T(1;4)9e2–10/Y*, P[(*ry*$^+$)211S12]/+ null-controlled flies. **An intermediate phenotype, with significant differences from both of these control values.

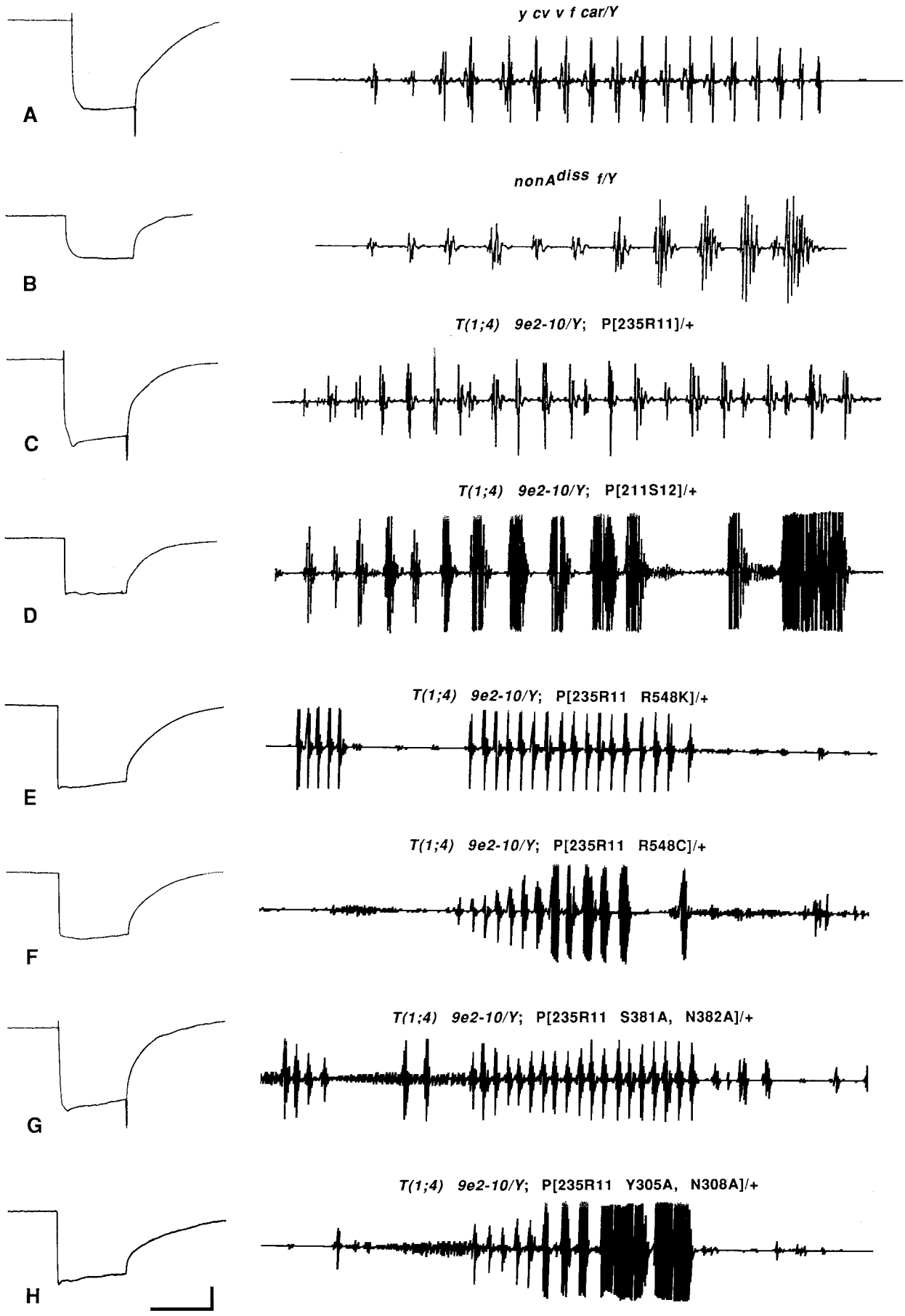
range of effects on viability and courtship song, while uniformly causing the flies to be blind in the optomotor and ERG assays. The R548C mutants exhibited severe, global defects, whereas an Arg → Lys change at this site, which might in principle be much milder, led to flies with partial viability deficits, subtle abnormalities of courtship song, and severe visual defects. Mutations of pairs of amino acid substitutions in the RRM domains, hypothesized to be involved in ligand binding, led to two phenotypic classes. Dramatic effects on courtship song, vision, and viability were observed in both RRM1 mutant types. However, the S381A N382A substitutions in RNP-2 of the more C-terminal RRM2 led to apparent normality for all phenotypes examined.

DISCUSSION

This study provides new information concerning structure–function relationships within the conserved region of NONA. The newly discovered *nAhomo* locus, which encodes a protein with 80% identity to NONA, suggests that there might even be an

→

Figure 5. Visual physiology and courtship-song abnormalities of *nonA* transgenics. At the *left* are shown characteristic ERG traces, and at the *right* are samples of courtship song. *A*, The *y cv v f car/Y* control yielded a wild-type ERG, with robust light-on and light-off transient spikes and a courtship song with a normal level of intrapulse polycyclicity. *B*, The *dissonance* phenotype is characterized by a lack of transient spikes and has highly polycyclic pulses relatively late in the song trains. *C*, *T(1;4)9e2–10/Y*; *P*[235R11]/+ is a fully rescued transformant male (cf. Rendahl et al., 1992; Stanewsky et al., 1993). *D*, *T(1;4)9e2–10/Y*; *P*[211S12]/+ illustrates the loss-of-function phenotype. The four sets of traces in the *bottom half* exemplify phenotypes of transgenic males carrying novel point mutations in the *nonA* gene. There were two categories with respect to ERG, and a more graded range of courtship-song phenotypes (see also Table 3). *E*, Flies with a lysine substitution at position 548 (R548K/+) lacked transients, but produced only mildly defective song traces. *F*, Those with a cysteine substitution (R548C/+) lacked transients and produced *dissonance*-like mutant pulse trains. *G*, Flies bearing S381A N382A mutations in the RRM were normal for both physiological and behavioral phenotypes. *H*, The other three types of RRM mutants, exemplified here by [Y305A N308A]/+, lacked transients and sang in a *nonA*-null-like manner. *ERG scale*, The *vertical bar* represents 5 mV; the *horizontal bar* represents 3 sec, the duration of the light pulse (the horizontal distance between the on and off spikes). The courtship-song traces, corresponding to singing bouts that ranged from 1 to ~2 sec of male wing vibrations, are not all shown to the same scale; thus, to interpret the duration of a given trace, note that the typical timespan between adjacent pulses was ~30–40 msec (which is typical for *D. melanogaster* males). All but one of the courtship-song samples represent complete trains of 10 or more pulses; the [211S12]/+ case shows only the beginning section of a very long train.


A
y cv v f car/Y
B
nonA^diss f/Y
C
T(1;4) 9e2-10/Y; P[235R11]/+
D
T(1;4) 9e2-10/Y; P[211S12]/+
E
T(1;4) 9e2-10/Y; P[235R11 R548K]/+
F
T(1;4) 9e2-10/Y; P[235R11 R548C]/+
G
T(1;4) 9e2-10/Y; P[235R11 S381A, N382A]/+
H
T(1;4) 9e2-10/Y; P[235R11 Y305A, N308A]/+

intraspecific family of these genes containing this domain in *Drosophila* (cf. Kim and Baker, 1993; Brand et al., 1995). Sequencing of known *nonA* behavioral mutants indicated that the second RRM and amino acid 548 of the adjacent charged region are important in visual behavior and courtship song, respectively. Novel point mutations introduced into the first RRM demonstrated that this region is required for those behaviors and for normal viability.

With the exception of the *in vitro* mutated E422A F423A transgenic type, protein levels do not appear to be a primary factor in determining phenotype. Most of the transgenics produced reasonably robust levels of full-length NONA protein. For example, the K337A F339A mutant was quite defective phenotypically but had normal-appearing protein levels. This is similar to the (*in vivo* mutated) $nonA^{diss}$ mutant, which exhibits severe visual and courtship-song defects but normal viability (Rendahl and Hall, in press) and protein levels (Rendahl et al., 1992). The decreased viability of the (*in vitro* mutated) R548C transgenic (which on paper mimics the *dissonance* mutation) could be attributable to its decreased protein level; the reason for this biochemical anomaly is unknown (given the NONA-level normality observed for $nonA^{diss}$). This substitution, and the two novel mutations in the E422A F423A transgenic (which cause the protein to be undetectable; see Fig. 4) may make the NONA polypeptide especially susceptible to proteolysis, a possibly intriguing biochemical phenotype for further study. The $nonA^{P14}$ visual mutant is also severely NONA-depleted; yet, in contrast to the phenotypically defective strains now under discussion, flies expressing this *in vivo* generated mutant sing normal courtship songs (Kulkarni et al., 1988; Rendahl et al., 1992) and are fully viable (Rendahl and Hall, in press).

Perhaps because of their severity, none of the novel mutations in the RRMs or the charged region generated here affected solely visual or courtship-song function. Nevertheless, one's attention is especially drawn to the mutations at residue 548 within the relatively C-terminal charged region, because these substitutions (the original one of which occurred in $nonA^{diss}$) tend to cause marked decrements in NONA function. Even the replacement of Arg with the similarly charged Lys led to a mild song defect and a severely visually impaired phenotype. Yet, this region is inherently featureless, except for the notable intra- and interspecific homologies; without those data, one might have devoted an inappropriately high degree of attention to the RRMs of NONA. Further alterations within the charged region could be informative in determining whether mutations at position 548 are unique in terms of how they impair the function of the protein.

For the RNP domains, highly conserved amino acids were chosen for *in vitro* mutagenesis, to maximize the chance of creating one or more mutant phenotypes. Novel mutations within the RRM1 were shown to severely affect vision, song, and viability. Consistent with this result, site-directed mutagenesis of the *tra-2* gene has demonstrated that single point mutations in the RNP-I domain of the RRM affected all phenotypes associated with mutations at that locus (Amrein et al., 1994). In contrast, a pair of NONA mutations in the RNP-II domain of RRM2 led to flies exhibiting normal behavioral phenotypes and viability. The chemically induced $nonA^{H2}$ and $nonA^{P14}$ mutants, which carry point mutations in or near the RRM2 domain, have defects specific to the visual system. Thus, it is possible that the RRM2 domain cannot be mutated to produce a null-like phenotype. The hypothesis that the *nonA* mutant phenotypes are caused by nonspecific effects, such as generalized effects on protein stability, cannot be excluded. A more intriguing possibility, as suggested by the ensemble of effects of the *in vivo* and *in vitro* generated mutants, is that RRM1 is necessary for all the known functions of NONA, whereas RRM2 may be primarily involved in vision.

If NONA is a splicing factor, a molecular mechanism that would explain the phenotypes of the *nonA* mutants is that the target pre-mRNAs involved in vision are distinct or require more efficient splicing than those required for other NONA functions. In the latter case, both RRMs might be necessary for highly efficient splicing. Variation in splicing efficiency of particular pre-mRNAs *in vitro* and *in vivo* has been demonstrated (Abovich, 1988). Splicing efficiency is known to be affected by context effects of both intron sequences (Pikielny and Rosbash, 1985; Goguel and Rosbash, 1993) and exon sequences (Reed and Maniatis, 1986; Watakabe et al., 1993).

The ubiquitous spatial and temporal expression pattern of the *Drosophila nonA* gene seemingly contrasts with the specificity of the behavioral defects observed in the hypomorphic alleles. This leads to the question of whether NONA performs general or nervous system-specific roles in RNA processing. The ubiquitous localization of NONA to nuclei (Rendahl et al., 1992) might imply that it has a general role in RNA processing, as has been proposed for the *Drosophila* splicing factors RBP1 and B52 (Champlain et al., 1991; Kim et al., 1992). NONA, B52, and RBP1 all colocalize with RNA polymerase II on polytene chromosomes (Frasch and Saumweber, 1989; Champlain et al., 1991; Kim et al., 1992). *nonA* was independently cloned by Besser et al. (1990), who identified the NONA protein as a component of active chromatin. These authors argue that NONA may be a transcription factor. However, its colocalization with RNA polymerase II is consistent with a splicing function, because splicing and transcription are known to occur simultaneously in *Drosophila* (Beyer and Osheim, 1988). The significant homology between the NONA protein and the essential mammalian splicing factor psf (Fig. 2) supports the idea of such a function.

Another possibility is that NONA mediates an hnRNP-like function, given its association with chromatin and high Gly content (Dreyfuss et al., 1993). Traditionally, hnRNPs have been viewed as "packaging" proteins with little binding specificity. However, the hnRNP composition has been found to vary among different pre-mRNAs on *Drosophila* polytene chromosomes (Matunis et al., 1993). The ratio of the hnRNPs bound to a given transcript may affect its processing. They may make it accessible to splicing factors and facilitate protein–protein interactions (Dreyfuss et al., 1993) or, alternatively, function as negative regulators by blocking access to pre-mRNA (Samuels et al., 1994).

Based on the behavioral phenotypes of the *nonA* mutants, we hypothesize that NONA has a role in mRNA processing in the CNS. A family of *Drosophila* genes encoding RRM proteins is hypothesized to be specifically required for mRNA splicing in the nervous system [*elav* (for review, see Yao et al., 1993), *cpo* (Bellen et al., 1992), *rbp9* (Kim and Baker, 1993), and *msi* (Nakamura et al., 1994)]. Because of its widespread expression pattern (Rendahl et al., 1992), *nonA* would not be part of this neural-specific RRM family; it nonetheless has similarities to them. Structurally, all are divergent from prototypical splicing factors, lacking the Arg/Ser (Zamore et al., 1992) and Arg/Gly/Gly repeats (Kiledjian and Dreyfuss, 1992)—auxiliary domains commonly found in RRM family members. Instead, they contain homopolymeric stretches of amino acids in their N or C termini. Certain functional similarities among these *Drosophila* genes can also be inferred (see introductory remarks).

There is predicted to be a high level of regulation of neural gene expression at both the transcriptional and post-transcriptional levels in *Drosophila* (Yao et al., 1993). What might be the possible mechanisms for generating the specificity of NONA in the splicing reaction? First, the variable loop 3 of RRM2 in NONA may be important. NONA and the highly related RRM proteins aligned here have an unusual degree of identity in loop 3 of RRM2 (see Fig. 2*A*). Second, NONA may interact with specific combinations of splicing factors. For example, the multifunctional splicing factor TRA-2 is known to control the processing of three mRNAs (Nagoshi et al., 1988; Mattox and Baker, 1991; Hazelrigg and Tu, 1994). TRA-2 has been shown to interact with *transformer*-encoded and general "SR" family splicing factors (Zahler et al., 1992) *in vitro* in the splicing of the *doublesex* pre-mRNA (Tian and Maniatis, 1993). Third, *cis*-acting sequences in the pre-mRNA might be required for the specificity of *nonA*-mediated splicing. In the case of TRA-2, which regulates the splicing of its own and *exu* mRNAs in the male germline (Hazelrigg and Tu, 1994), the position of *cis*-binding sites must also contribute to the specificity of the splicing reaction (Mattox et al., 1992; Amrein et al., 1994).

A number of pre-mRNAs, known to undergo regulated splicing in the *Drosophila* nervous system, are candidate NONA ligands. These include mRNAs transcribed from ion-channel genes and/or from certain loci required for normal visual or courtship behaviors (see below). For example, varying splice forms arise from the *Shaker* potassium-channel locus, and these are differentially localized with the CNS and visual system (Schwarz et al., 1990; Hardie et al., 1991). The *hdc* gene, which encodes a histamine-synthetic enzyme and is essential for the production of ERG transients, produces multiple transcripts found in various ganglia of the head and thorax (Burg et al., 1993). Given that thoracic–ganglionic structure and function are implicated in song control (for review, see Hall et al., 1980, 1990), *hdc* might not only play a role in vision, but also interact with *nonA* to help mediate that feature of male courtship. Finally, consider the *cacophony* locus, which is thought to encode a voltage-gated calcium channel (Hall, 1994); *cac* transcript are possible "downstream" targets of NONA, because the various mutations at this locus lead to courtship-song, visual-response, or viability decrements (Hall et al., 1990). To determine whether the RNAs produced by any of these candidate genes are NONA ligands will require the development of a biochemical assay for the binding of that protein; such studies could be augmented by analysis of genetic interactions between *nonA* mutations and those at the other loci.

REFERENCES

Abovich N (1988) Regulation of yeast ribosomal protein gene expression. PhD thesis, Brandeis University.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410.

Amrein H, Hedley ML, Maniatis T (1994) The role of specific protein-RNA and protein-protein interactions in positive and negative control of pre-mRNA splicing by transformer-2. Cell 76:735–745.

Bairoch A, Buchner P (1994) PROSITE: recent developments. Nucleic Acids Res 22:3583–3589.

Bandziulis RJ, Swanson MS, Dreyfuss G (1989) RNA-binding proteins as developmental regulators. Genes Dev 3:431–437.

Bellen HJ, Kooyer S, D'Evelyn D, Pearlman J (1992) The *Drosophila* couch potato protein is expressed in nuclei of peripheral neuronal precursors and shows homology to RNA-binding proteins. Genes Dev 6:2125–2136.

Bernstein AS, Neumann EK, Hall JC (1992) Temporal analysis of tone pulses within the courtship songs of two sibling *Drosophila* species, their interspecific hybrid, and behavioral mutants of *D. melanogaster*. J Insect Behav 5:15–36.

Besser HV, Schnabel P, Wieland C, Fritz E, Stanewsky R, Saumweber H (1990) The puff-specific *Drosophila* protein Bj6, encoded by the gene *no-on-transient A*, shows homology to RNA binding proteins. Chromosoma 100:37–47.

Beyer AL, Osheim YN (1988) Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. Genes Dev 2:754–765.

Birney E, Kumar S, Krainer AR (1993) Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. Nucleic Acids Res 21:5803–5816.

Brand SF, Pichoff S, Noselli S, Bourbon H-M (1995) Novel *Drosophila melanogaster* genes encoding RRM-like RNA-binding proteins identified by a degenerate: PCR strategy. Gene 154:187–192.

Burd CG, Dreyfuss G (1994) Conserved structures and diversity of functions of RNA-binding proteins. Science 265:615–621.

Burg MG, Sarthy PV, Koliantz G, Pak WL (1993) Genetic and molecular identification of a *Drosophila histidine decarboxylase* gene required in photoreceptor transmitter synthesis. EMBO J 12:911–919.

Campos AR, Grossman D, White K (1985) Mutant alleles at the locus *elav* in *Drosophila melanogaster* lead to nervous system defects: a developmental-genetic analysis. J Neurogenet 2:197–218.

Champlain DT, Frasch M, Saumweber H, Lis JT (1991) Characterization of a *Drosophila* protein associated with boundaries of transcriptionally active chromatin. Genes Dev 5:1611–1621.

Claverie JM, States D (1993) Information enhancement methods for large scale sequence analysis (XNU). Comp Chem 17:191–202.

Coombe PE (1986) The large monopolar cells L1 and L2 are responsible for ERG transients in *Drosophila*. J Comp Physiol [A] 159:655–665.

Dong B, Horowitz DS, Kobayashi R, Krainer AR (1993) Purification and cDNA cloning of HeLa cell p54nrb, a nuclear protein with two RNA recognition motifs and extensive homology to human splicing factor PSF and *Drosophila* NONA/BJ6. Nucleic Acids Res 21:4085–4092.

Dreyfuss G, Matunis M, Pinol-Roma S, Burd CG (1993) hnRNP proteins and the biogenesis of mRNA. Annu Rev Biochem 62:289–321.

Frasch M, Saumweber H (1989) Two proteins from *Drosophila* nuclei are bound to chromatin and are detected in a series of puffs on polytene chromosomes. Chromosoma 97:272–281.

Gibson TJ, Rice PM, Thompson JD, Heringa J (1993) KH domains within the FMR1 sequence suggest that fragile X syndrome stems from a defect in RNA metabolism. Trends Biochem Sci 18:331–333.

Goguel V, Rosbash M (1993) Splice site choice and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast. Cell 72:893–901.

Gozani O, Patton JG, Reed R (1994) A novel set of spliceosome-associated proteins and the essential splicing factor PSF bind stably to pre-mRNA prior to catalytic step II of the splicing reaction. EMBO J 13:3356–3367.

Greenspan RJ, Finn Jr JA, Hall JC (1980) Acetylcholinesterase mutants in *Drosophila* and their effects on the structure and function of the central nervous system. J Comp Neurol 189:741–774.

Graus F, Cordon-Cardo C, Posner J (1985) Neuronal antinuclear antibody in sensory neuronopathy from lung cancer. Neurology 35:538–543.

Hall JC (1994) The mating of a fly. Science 264:1702–1714.

Hall JC, Kulkarni SJ, Kyriacou CP, Yu Q, Rosbash M (1990) Genetic and molecular analysis of neural development and behavior in *Drosophila*. In: Developmental behavioural genetics (Hahn ME, Hewitt J, Henderson ND, eds), pp 100–112. New York: Oxford UP.

Hall JC, Tompkins L, Kyriacou CP, Siegel RW, Schilcher Fv, Greenspan RJ (1980) Higher behavior in *Drosophila* analyzed with mutations that disrupt the structure and function of the nervous system. In: Development and neurobiology of *Drosophila* (Siddiqi O, Babu P, Hall LM, Hall JC, eds), pp 425–455. New York: Plenum.

Hardie RC, Voss D, Pongs O, Laughlin SB (1991) Novel potassium channels encoded by the *Shaker* locus in *Drosophila* photoreceptors. Neuron 6:477–486.

Hazelrigg T, Tu C (1994) Sex-specific processing of the *Drosophila exuperantia* transcript is regulated in male germ cells by the *tra-2* gene. Proc Natl Acad Sci USA 91:10752–10756.

Heisenberg M (1971) Isolation of mutants lacking the optomotor response. Drosophila Inform Serv 46:68.

Heisenberg M (1972) Comparative behavioral studies on two visual mutants of *Drosophila*. J Comp Physiol 80:119–136.

Heisenberg M, Wolf R (1984) Vision in *Drosophila*: genetics of microbehavior. Berlin: Springer.

Homyk Jr T, Isono K, Pak WL (1985) Developmental and physiological analysis of a conditional mutation affecting photoreceptor and optic lobe development in *Drosophila melanogaster*. J Neurogenet 2:309–324.

Hotta Y, Benzer S (1970) Genetic dissection of the *Drosophila* nervous system by means of mosaics. Proc Natl Acad Sci USA 67:1156–1163.

Jones KR, Rubin GM (1990) Molecular analysis of *no-on-transient A*, a gene required for normal vision in *Drosophila*. Neuron 4:711–723.

Kenan DJ, Query CC, Keene JD (1991) RNA recognition: towards identifying determinants of specificity. Trends Biochem Sci 16:214–220.

Kiledjian M, Dreyfuss G (1992) Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. EMBO J 11:2655–2664.

Kim Y-J, Zuo P, Manley JL, Baker BS (1992) The *Drosophila* RNA-binding protein RBP1 is localized to transcriptionally active sites of chromosomes and shows a functional similarity to human splicing factor ASF/SF2. Genes Dev 6:2569–2579.

Kim Y-J, Baker BS (1993) The *Drosophila* gene *rbp9* encodes a protein that is a member of a conserved group of putative RNA binding proteins that are nervous system-specific in both flies and humans. J Neurosci 13:1045–1056.

Krejci CM, Rendahl KG, Hall JC (1994) Rescue of behavioral phenotypes by a *nonA*-null transgene is influenced by chromosomal position effects. Drosophila Inform Serv 75:117–119.

Kulkarni SJ, Steinlauf AF, Hall JC (1988) The *dissonance* mutant of courtship song in *Drosophila melanogaster*: isolation, behavior and cytogenetics. Genetics 118:267–285.

Kunkel TA, Roberts JD, Zakour RA (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol 154:367–382.

Lindsley DL, Zimm GG (1992) The genome of *Drosophila melanogaster*. San Diego: Academic.

Mattox W, Baker BS (1991) Autoregulation of the splicing of transcripts from the *transformer-2* gene of *Drosophila*. Genes Dev 5:786–796.

Mattox W, Ryner L, Baker BS (1992) Autoregulation and multifunctionality among trans acting factors that regulate alternative pre-mRNA processing. J Biol Chem 267:19023–19026.

Matunis EL, Matunis MJ, Dreyfuss G (1993) Association of individual hnRNP proteins and snRNPs with nascent transcripts. J Cell Biol 121:219–228.

Mismer D, Rubin GM (1987) Analysis of the promoter of the *ninaE* opsin gene in *Drosophila melanogaster*. Genetics 116:565–578.

Nagoshi RN, McKeown M, Burtis KC, Belote JM, Baker BS (1988) The control of alternative splicing at genes regulating sexual differentiation in *D. melanogaster*. Cell 53:229–236.

Nakamura M, Okano H, Blendy JA, Montell C (1994) Mushashi, a neural RNA-binding protein required for Drosophila adult external sensory organ development. Neuron 13:67–81.

Pak WL, Grabowski RS (1978) Physiology of the visual and flight systems. In: The genetics and biology of *Drosophila*, Vol 2a (Ashburner M, Wright TRF, eds), pp 553–604. London: Academic.

Pak WL, Grossfield J, Arnold KS (1970) Mutants of the visual pathway of *Drosophila melanogaster*. Nature 227:518–520.

Patton JG, Porro EB, Galceran J, Tempst P, Nadal-Ginard B (1993) Cloning and characterization of PSF, a novel pre-mRNA splicing factor. Genes Dev 7:393–406.

Pikielny CW, Rosbash M (1985) mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. Cell 41:119–126.

Pirrotta V (1988) Vectors for P-element transformation in *Drosophila*. In: Vectors. A survey of molecular cloning vectors and their uses (Rodriguez RL, Denhardt DT, eds), pp 437–456. London: Butterworths.

Reed R, Maniatis T (1986) A role for exon sequences and splice-site proximity in splice-site selection. Cell 46:681–690.

Rendahl KG, Hall JC (1996) Temporally manipulated rescue of visual and courtship abnormalities caused by a *nonA* mutation in *Drosophila*. J Neurogenet, in press.

Rendahl KG, Jones KR, Kulkarni SJ, Bagully SH, Hall JC (1992) The *dissonance* mutation at the *no-on-transient A* locus of *D. melanogaster*: genetic control of courtship song and visual behaviors by a protein with putative RNA-binding motifs. J Neurosci 12:390–407.

Richards RI, Sutherland GR (1992) Fragile X syndrome: the molecular picture comes into focus. Trends Genet 8:249–254.

Rio DC (1993) Splicing of pre-mRNA: mechanism, regulation and role in development. Curr Opin Genet Dev 3:413–424.

Robertson HM, Preston CR, Phillis RW, Johnson-Schlitz DM, Benz WK, Engels WR (1988) A stable genomic source of P element transposase in *Drosophila melanogaster*. Genetics 118:461–470.

Rubin GM, Spradling AC (1982) Genetic transformation of *Drosophila* with transposable element vectors. Science 218:348–353.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual (Irwin N, ed). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Samuels ME, Bopp D, Colvin RA, Roscigno RF, Garcia-Blanco MA, Schedl P (1994) RNA binding by *Sxl* proteins *in vitro* and *in vivo*. Mol Cell Biol 14:4975–4990.

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467.

Schwarz TL, Papazian DM, Carretto RC, Jan Y-N, Jan LY (1990) Immunological characterization of K^+ components from the *Shaker* locus and differential distribution of splicing variants in *Drosophila*. Neuron 2:119–127.

Seifert HS, Chen EY, So M, Heffron F (1986) Shuttle mutagenesis: a method of transposon mutagenesis for *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 83:735–739.

Siomi H, Siomi MC, Nussbaum RL, Dreyfuss G (1993) The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. Cell 74:291–298.

Smith RF, Smith TF (1990) Automatic generation of primary sequence patterns from sets of related protein sequences. Proc Natl Acad Sci USA 87:118–122.

Stanewsky R, Rendahl KG, Dill M, Saumweber H (1993) Genetic and molecular analysis of the X chromosomal region 14B17–14C4 in *Drosophila melanogaster*: loss of function in NONA, a nuclear protein common to many cell types, results in specific physiological and behavioral defects. Genetics 135:419–442.

Szabo A, Dalmau J, Manley J, Rosenfeld M, Wong E, Henson J, Posner JB, Furneaux HM (1991) HuD, a paraneoplastic encephalomyelitis antigen, contains RNA binding domains and is homologous to Elav and Sex-lethal. Cell 67:325–333.

Tian M, Maniatis T (1993) A splicing enhancer complex controls alternative splicing of *doublesex* pre-mRNA. Cell 74:105–114.

Vieira J, Messing J (1987) Production of single-stranded plasmid DNA. Methods Enzymol 153:3–11.

Watakabe A, Tanaka A, Shimura Y (1993) The role of exon sequences in splice site selection. Genes Dev 7:407–418.

Yang Y-S, Hanke JH, Carayannopoulos L, Craft CM, Capra JD, Tucker PW (1993) NonO, a non-POU-domain-containing, octamer-binding protein, is the mammalian homolog of *Drosophila nonA^{diss}*. Mol Cell Biol 13:5593–5603.

Yao K-M, Samson M-L, Reeves R, White K (1993) Gene *elav* of *Drosophila melanogaster*: a prototype for neuronal-specific RNA binding protein gene family that is conserved in flies and humans. J Neurobiol 24:723–739.

Zahler AM, Lane WS, Stolk JA, Roth MB (1992) SR proteins: a conserved family of pre-mRNA splicing factors. Genes Dev 6:837–847.

Zamore PD, Patton JG, Green MG (1992) Cloning and domain structure of the mammalian splicing factor U2AF. Nature 355:609–614.