The NHL domain of BRAT is an RNA-binding domain that directly contacts the hunchback mRNA for regulation

Inga Loedige, Mathias Stotz, Saadia Qamar, Katharina Kramer, Janosch Hennig, Thomas Schubert, Patrick Löffler, Gernot Längst, Rainer Merkl, Henning Urlaub, and Gunter Meister

Supplemental Material

Supplemental Materials and Methods

Static light scattering (SLS)

SLS was measured with a Malvern-Viscotek instrument (TDA 305) connected downstream to an Äkta purifier equipped with an analytical size-exclusion column (Superdex 75 10/300 GL, GE Healthcare). The running buffer was 20 mM Tris, 150 mM NaCl, 1 mM DTT, 5 % Glycerol, pH 8). Elution profiles were collected for 1 h with a flow rate of 0.5 ml/min and data were collected using absorbance UV detection at 280 nm, right angle light scattering and refractive index. The molecular weights of separated elution peaks were calculated using OmniSEC software (Malvern, UK). As calibration standard, 4 mg/ml Bovine Serum Albumin (BSA) was used prior to all experiments and the change in refractive index with respect to concentration was set to 0.186 ml/g (Wen et al. 1996).

Cell culture, transfections and RNA interference

Drosophila S2 cells were grown in Express Five SFM (Gibco) supplemented with GlutaMAX (Gibco). Transfections were performed in 96-well format with Cellfectin II and PLUS reagent (life technologies), according to manufacturer's instructions. Transfection mixtures contained 15 ng FL, 90 ng RL, 45 ng NHA-, or HA-fusion protein expressing plasmid, 0.9 ul PLUS and 1.5 ul Cellfectin per 3 wells of a 96-well plate. Cells were lysed in Passive Lysis Buffer (PLB) (Promega) on day 2 post-transfection and renilla and firefly luciferase activities were measured as described (Hock et al. 2007). FL activity was

normalized to RL activity and FL/RL values produced in the presence of an empty control vector were set to one. As overexpression of HA-Pum had an unspecific positive effect on expression of RL, we included an additional normalization step (for experiments shown in Fig. 4), setting the FL/RL values of samples expressing a FL-reporter that contains an artificial 3'UTR (neg. ctrl.) not regulated by any of the proteins under investigation to one. Values represent means of three independent experiments, each performed in triplicate and error bars show standard error of mean. A representative experiment including raw FL and RL values and normalization steps is shown in Supplemental Fig. 4. RNAi experiments were performed according to (Worby et al. 2001). Briefly, dsRNA was generated using the T7 MEGAscript kit (life technologies) from DNA templates that were generated by PCR amplification using following primers: Pum (for: TAATACGACTCACTATAGGGGTCAAGGATCAGAATGGCAATCATGT; rev: TAATACGACTCACTATAGGGCTTCTCCAACTTGGCATTGATGTGC) and Brat (for: TAATACGACTCACTATAGGGAGTACGGCCAGTTTGTGAGG; rev: TAATACGACTCACTATAGGGCATACCCACTGGCGCCAGTTGG)
(taken from (Weidmann and Goldstrohm 2012)). Both forward and reverse primers contained T7 promoter sites. Treatment of cells with dsRNA was on day 0 and repeated on day 4. Cells were transfected on day 7 and lysed on day 9.

Determination of a phylogeny and of long range electrostatics

All considered NHL domains possess a six-bladed β-propeller fold; however sequence similarity of the selected proteins is low. The average pairwise sequence identity value is 21±8 % as determined by the algorithm "Needle" of the EMBOSS package (Rice et al. 2000) with default parameters. This is why we opted for a structure-based algorithm when constructing a multiple sequence alignment (MSA) needed for phylogenetic analysis.

To begin with, we selected all family members, whose 3D structure was deposited in the PDB (Bernstein et al. 1977). These were the peptidyl-alpha-hydroxyglycine alpha-amidating lyase (RnPAL, PDB ID 3FW0), the serine/threonine-protein kinase (MtPknD, PDB ID 1RWI) and the brain tumor NHL domain (DmBrat, PDB ID 1Q7F). In order to represent other characteristic members of this family, we built homology models by

means of I-Tasser (Zhang 2008). In each case, the model with the best C-score was chosen, which resulted in a mean C-score of 0.1±1.1. By means of Chimera (Pettersen et al. 2004), all structures were superimposed on the model of HsTrim71, which yielded the lowest sum of RMSD values in an all against all superposition. Subsequently, Chimera's "Match -> Align" algorithm (Meng et al. 2006) was utilized to generate an MSA based on this superposition. Finally, the sequences of the remaining proteins with unknown 3D structure were included by applying the "add" option (Katoh and Frith 2012) of MAFFT (Katoh and Standley 2013) with default parameters.

The MSA was the basis to determine a neighbor-joining tree (Saitou and Nei 1987) by means of SplitsTree4 (Huson and Bryant 2006). 1000 bootstrap samples were computed to determine significant edges.

For each 3D structure, surface electrostatic calculations were performed with the Particle Mesh Ewald approach (Krieger et al. 2006) as implemented in YASARA Structure (V 13.4.21) (Krieger et al. 2004) and by employing the YAMBER3 force field in physiological pH. The solvent accessible surface was color coded representing the local electrostatic potential. The darkest blue color represents a positive and the darkest red color a negative potential of 300 kJ/mol, respectively.

UV Cross-linking

Sample Preparation

RNA-protein complex was assembled as follows: 1 nmol of in vitro transcribed hb RNA was incubated with 1 nmol of recombinant BRAT-NHL on ice for 1 h. For RNA-protein cross-linking, the assembled complex was UV irradiated at a wavelength of 254 nm for 10 min. The sample was precipitated with 3 Vol. of ice-cold ethanol and 1/10 volume of 3 M NaOAc. The pellet was dissolved in 4 M Urea and 50 mM Tris-HCl, pH 7.9 and adjusted to final concentration of 1 M in 50 mM Tris-HCl, pH 7.9. The RNA was hydrolysed using RNase A and T1 (Ambion, Applied Biosytems) at 52°C for 1 h followed by Benzonase (Novagen) for 1 h at 37°C. The sample was then digested with trypsin (Promega) for 14 h at 37°C. After digestion, the sample was desalted on C18 material (Dr. Maisch GmbH) and enriched by TiO_2 (GL Sciences) solid phase extraction as

described previously (Kramer et al. 2011). For mass spectrometric analysis, the sample was dried in a SpeedVac and reconstituted in 12 µl 5% Acetonitrile, 0.1% Formic Acid.

LC-MSMS and data analyses

Mass spectrometric analysis was performed by injecting 6 µl of sample onto nanoflow-liquid chromatography system (Agilent 1100 series, Agilent) coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific). The sample was applied on a trapping column (C18 material, 20 mm length, 0.150 mm inner diameter) at a flow rate of 10 µl/min in 3% buffer B (95% ACN, 0.1% FA), followed by elution and separation on an analytical column (C18 material, 150 mm length, 0.075 mm inner diameter) at a flow rate of 300 nl/min using a linear gradient of 3-36% buffer B over 37 min. The instrument was operated in data-dependent acquisition mode with a top 10 higher-energy collision dissociation method. MS scans were recorded in the m/z range 350-1600 at a resolution setting of 30, 000 FWHM. MS/MS scans were recorded at normalized collision energy of 45 and a dynamic exclusion of 20 sec at a resolution setting of 7500 FWHM and isolation width of 2 Th. The automatic gain control was set to 30,000 for ion trap and 10^6 for FTMS. LC-MS data were inspected manually in order to identify the crosslinked peptides (Luo et al. 2008).

Size exclusion chromatography of a complex of Pum-HD, Brat-NHL and hb RNA

30 µM of Brat-NHL and Pum-HD were mixed in the presence or absence of 30 µM of hb RNA in 20mM Tris, 150 mM NaCl, 1 mM DTT. Mixtures were incubated for 20 min on ice and 500 µl were subjected to size exclusion chromatography on a Superdex 75 10/300 GL in the same buffer. Fractions containing the proteins were collected and analyzed by 15 % SDS-PAGE and stained with Coomassie.

Supplemental References:

Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Jr., Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J Mol Biol* **112**: 535-542.

Hock J, Weinmann L, Ender C, Rudel S, Kremmer E, Raabe M, Urlaub H, Meister G. 2007. Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep* **8**: 1052-1060.

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**: 254-267.

Katoh K, Frith MC. 2012. Adding unaligned sequences into an existing alignment using MAFFT and LAST. *Bioinformatics* **28**: 3144-3146.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772-780.

Kramer K, Hummel P, Hsiao H-H, Luo X, Wahl M, Urlaub H. 2011. Mass-spectrometric analysis of proteins cross-linked to 4-thio-uracil- and 5-bromo-uracil-substituted RNA. *International Journal of Mass Spectrometry* **304**: 184-194.

Krieger E, Darden T, Nabuurs SB, Finkelstein A, Vriend G. 2004. Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins* **57**: 678-683.

Krieger E, Nielsen JE, Spronk CA, Vriend G. 2006. Fast empirical pKa prediction by Ewald summation. *J Mol Graph Model* **25**: 481-486.

Luo X, Hsiao HH, Bubunenko M, Weber G, Court DL, Gottesman ME, Urlaub H, Wahl MC. 2008. Structural and functional analysis of the E. coli NusB-S10 transcription antitermination complex. *Molecular Cell* **32**: 791-802.

Meng EC, Pettersen EF, Couch GS, Huang CC, Ferrin TE. 2006. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* **7**: 339.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605-1612.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**: 276-277.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406-425.

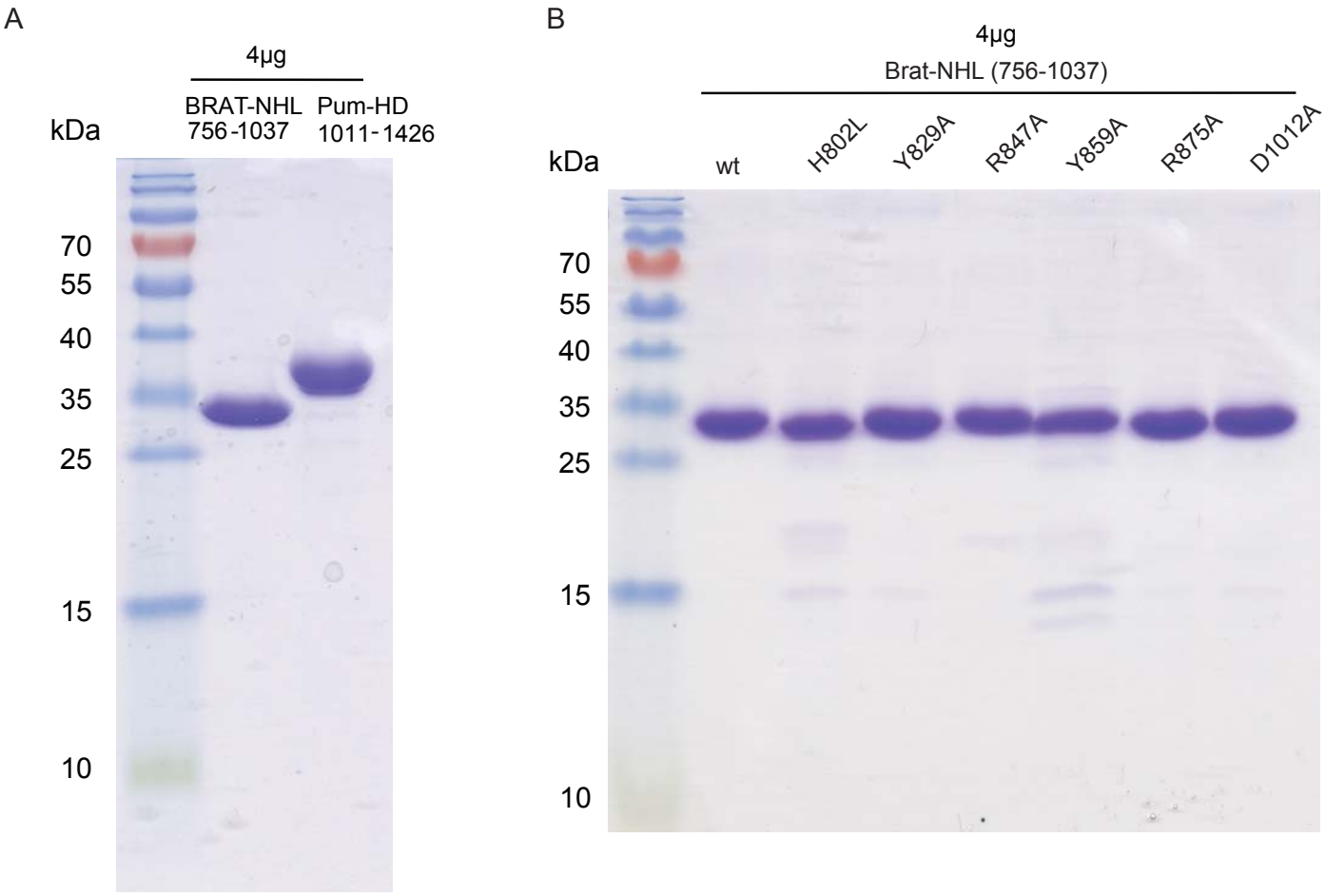
Weidmann CA, Goldstrohm AC. 2012. Drosophila Pumilio protein contains multiple autonomous repression domains that regulate mRNAs independently of Nanos and brain tumor. *Mol Cell Biol* **32**: 527-540.

Wen J, Arakawa T, Philo JS. 1996. Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Anal Biochem* **240**: 155-166.

Worby CA, Simonson-Leff N, Dixon JE. 2001. RNA interference of gene expression (RNAi) in cultured Drosophila cells. *Sci STKE* **2001**: pl1.

Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**: 40.

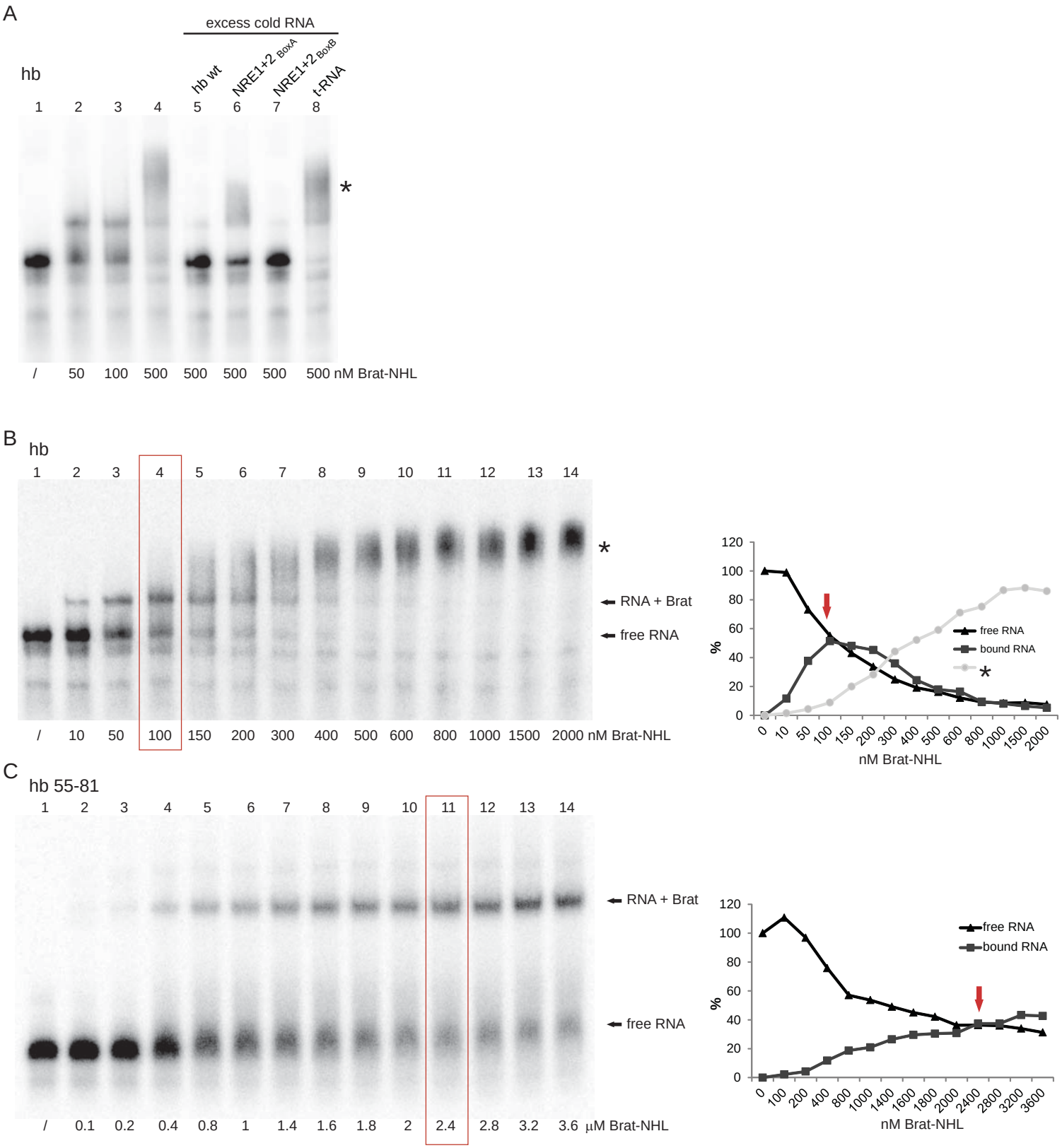
Loedige et al., Supplemental Figure 1



Supplemental Figure 1: Purity of recombinant Brat-NHL and Pum-HD

Coomassie stained SDS-PAGE of 4µg of recombinant purified Brat-NHL and Pum-HD (**A**), and Brat-NHL and indicated Brat-NHL point mutants (**B**). Theoretical molecular weight Brat-NHL: 32 kDa; Pum-HD: 38 kDA.

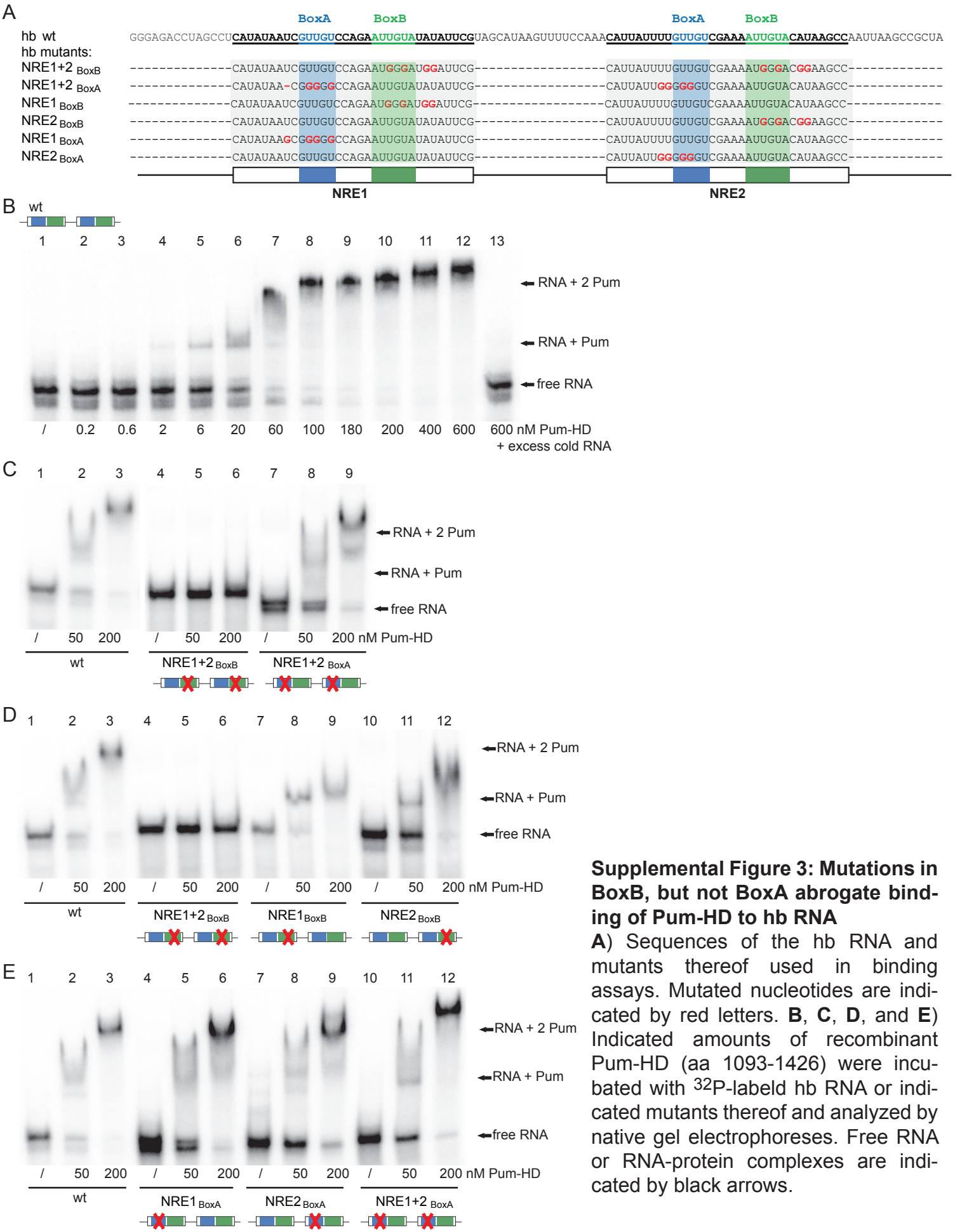
Loedige et al., Supplemental Figure 2



Supplemental Figure 2: Supplementary to Figure 1.

A) Cold chase experiment. Indicated amounts of Brat-NHL were incubated with ^{32}P-labeled hb RNA and complexes were analyzed by native gel electrophoresis. Labeled protein-RNA complexes were chased by a 1000 fold excess of unlabeled hb RNA (lane 5) or unlabeled hb RNA with both BoxB sites mutated (NRE1+2$_{BoxB}$) (lane 7), but less effective by an excess of hb RNA with both BoxA sites mutated (NRE1+2$_{BoxA}$)(lane 6) and not at all by an excess of unlabeled t-RNA (lane 8), indicating sequence specificity of the Brat-NHL hb RNA interaction. **B**) and **C**) Titration experiments to determine binding affinities of Brat-NHL to hb RNA (**B**) or to hb fragment 55-81 (**C**). Increasing amounts of Brat-NHL were incubated with 500 pM ^{32}P-labeld hb RNA or hb 55-81 and complexes were separated by native gel electrophoresis. Band intensities were quantified using quantity one software. Experiments were performed twice and shown is one representative experiment. The percentage of free and protein-bound RNA is depicted in the graphs shown on the right. The Brat-NHL concentration at which half of the free RNA is shifted into a slower migrating protein-RNA complex is indicated. The in this way determined binding affinities of approximately 100 nM for hb RNA and 2.4 µM for hb 55-81 closely match the results of the MST measurements (shown in Fig. 1E) of 137 nM and 2 µM for hb RNA and hb 55-81, respectively. Free RNA or RNA-protein complexes are indicated by black arrows. Asterisk (*) denotes a less-well-defined Brat-NHL hb RNA complex appearing at high protein concentrations.

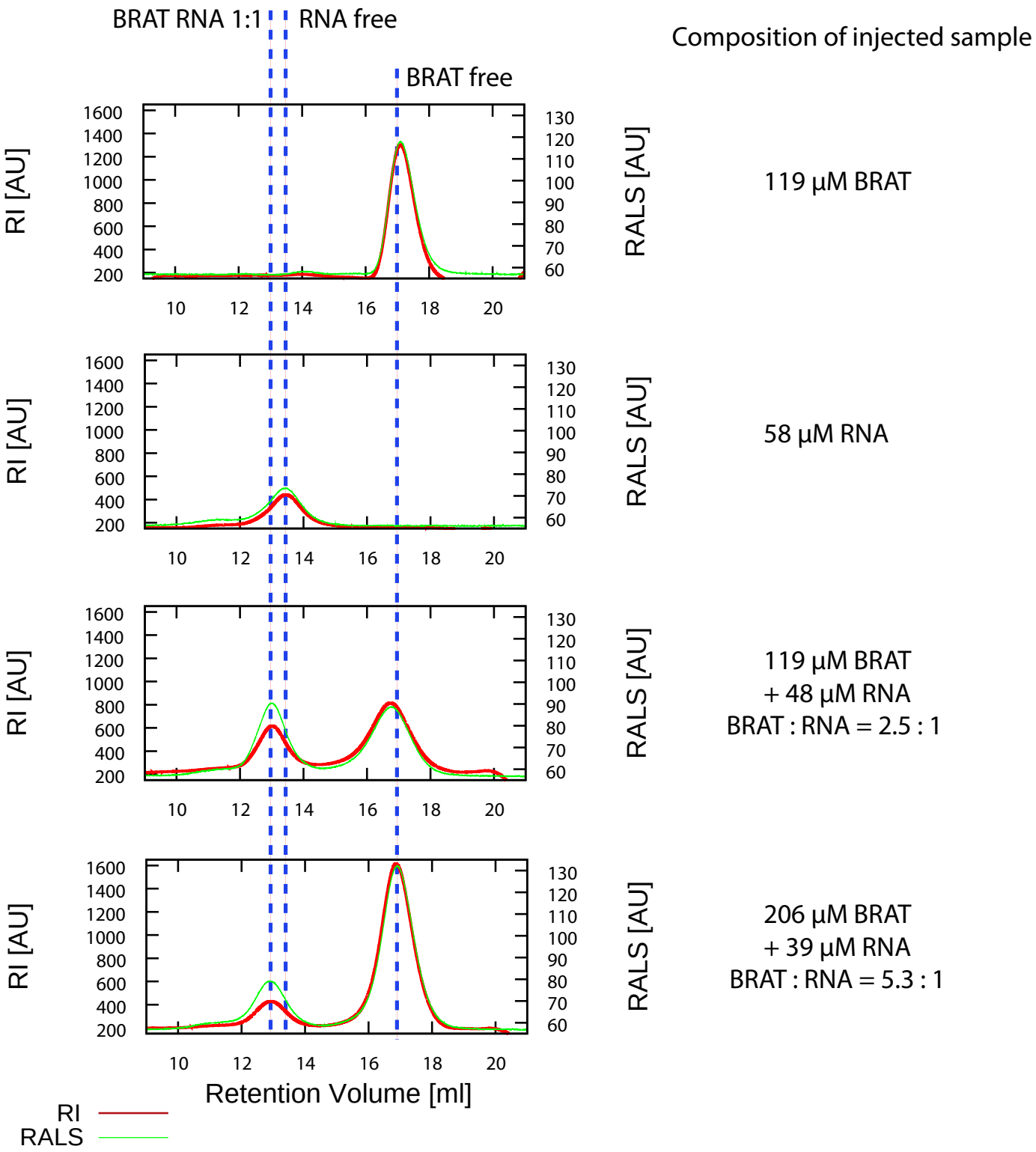
Loedige et al., Supplemental Figure 3



Supplemental Figure 3: Mutations in BoxB, but not BoxA abrogate binding of Pum-HD to hb RNA

A) Sequences of the hb RNA and mutants thereof used in binding assays. Mutated nucleotides are indicated by red letters. **B**, **C**, **D**, and **E**) Indicated amounts of recombinant Pum-HD (aa 1093-1426) were incubated with ^{32}P-labeld hb RNA or indicated mutants thereof and analyzed by native gel electrophoreses. Free RNA or RNA-protein complexes are indicated by black arrows.

Loedige et al., Supplemental Figure 4



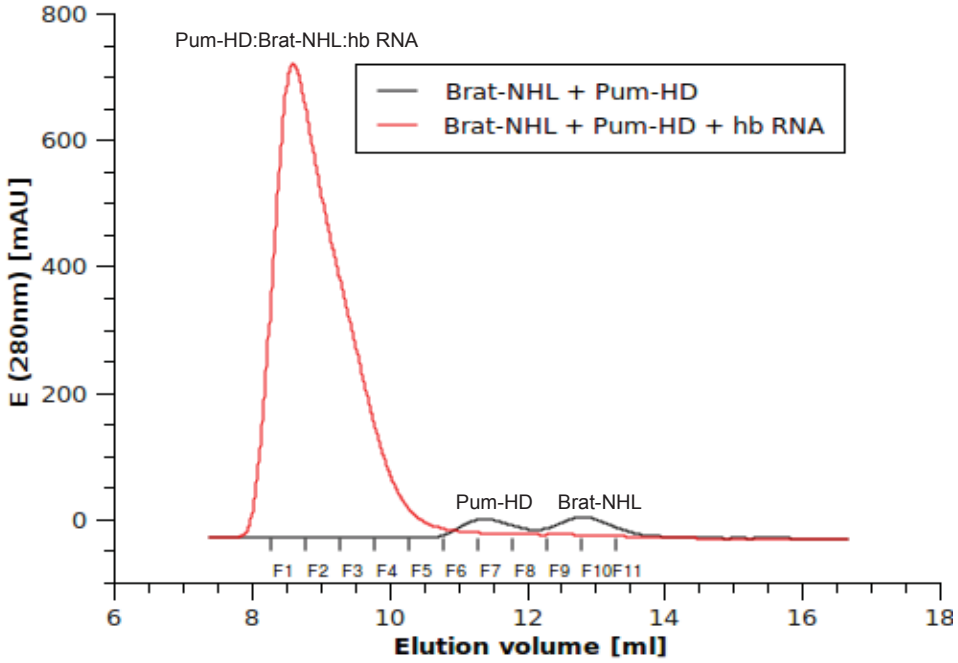
Peak	Theoretical MW (kDa)	Calculated MW (from SLS, kDa)
BRAT free	31. 9	31.6
RNA free	37.2	41.1
BRAT-RNA 1 : 1	69.2	69.1

Supplemental Figure 4: One Brat-NHL domain binds one hb RNA

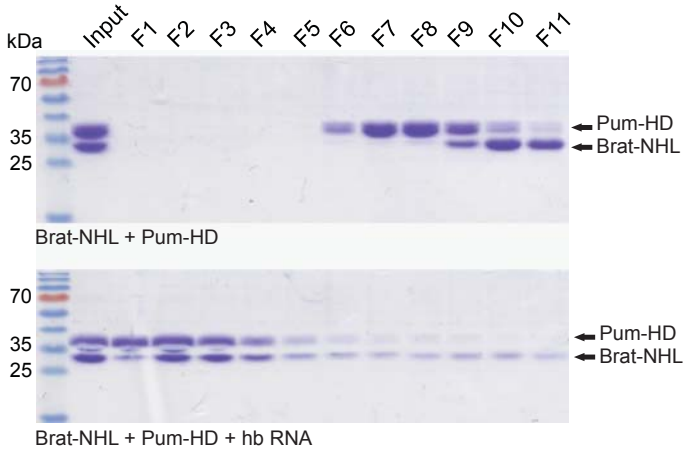
Elution profiles of gel filtration runs of Brat-NHL (top), hb RNA (second row), Brat-NHL: hb RNA (2.5:1, third row) and Brat-NHL: hb RNA (5.3:1, bottom) samples. Detected are the refractive index (red line) and right angle light scattering (green line) in order to calculate the molecular weight of each fraction. The table at the bottom compares the theoretical molecular weights for the free components and for a 1:1 complex with the experimentally determined molecular weights of fractions and each corresponding position in the chromatogram is indicated by a vertical line. The Brat-NHL: hb RNA samples, although having 2.5-5.3x excess of Brat-NHL elute as a 1:1 complex and free Brat-NHL, where the relative peak height of the peak corresponding to free Brat-NHL increases proportionally to the amount added, further indicating that a 2:1 complex is not formed.

Loedige et al., Supplemental Figure 5

A



B

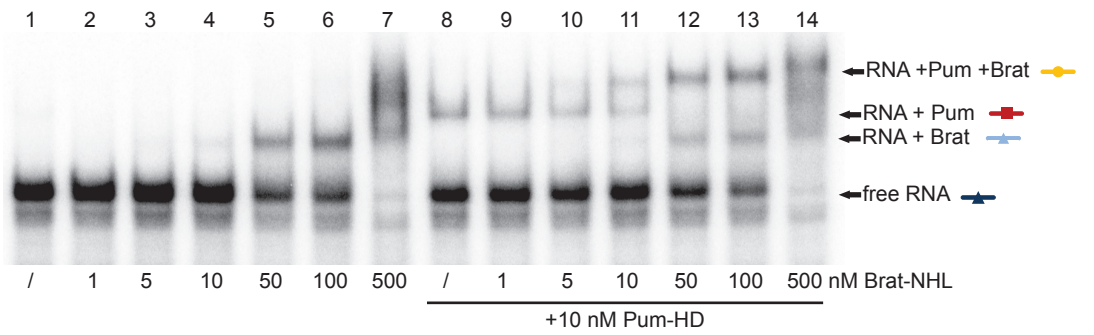


Supplemental Figure 5: Brat-NHL and Pum-HD form a stable complex only in the presence of RNA.

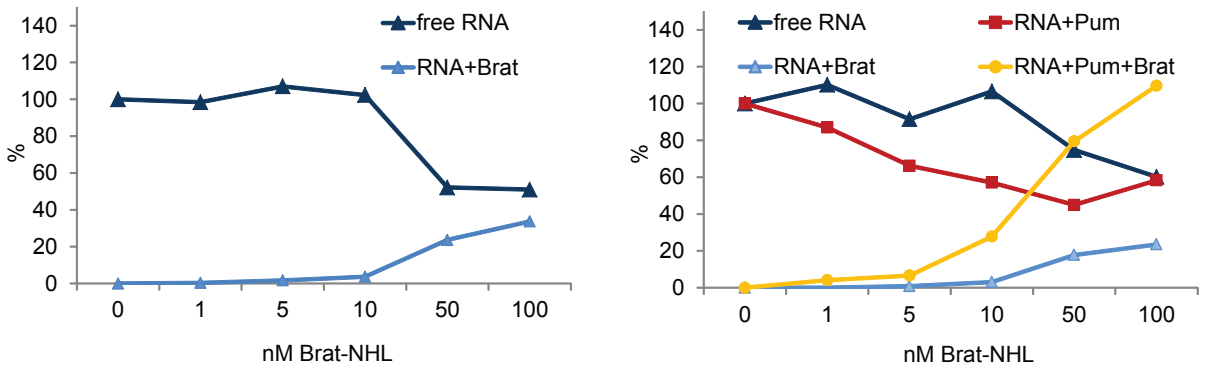
A) Elution profile of mixtures of purified Brat-NHL and Pum-HD on a Superdex 75 10/300 GL size exclusion chromatography column in the presence and absence of hb RNA (30 µM each). In the presence of hb RNA a stable complex containing the RNA and both proteins elutes at a larger apparent molecular weight (fractions F1-F5), while in the absence of RNA Pum-HD and Brat-NHL elute in separate peaks (fractions F6-F9 and F9-F11) corresponding to the elution behavior of the proteins in isolation (data not shown). **B**) Analysis of protein content of the fractions on 15% SDS-PAGE.

Loedige et al., Supplemental Figure 6

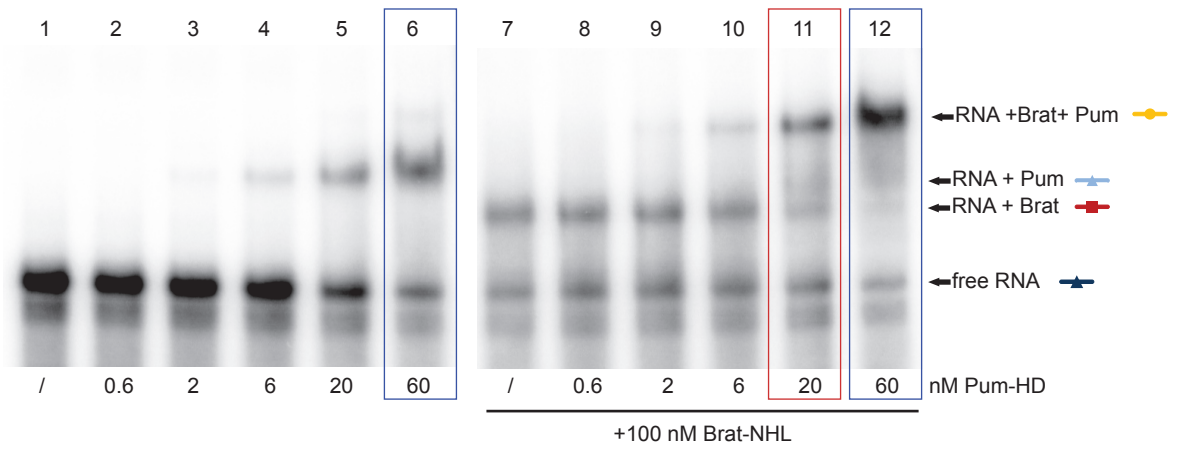
A



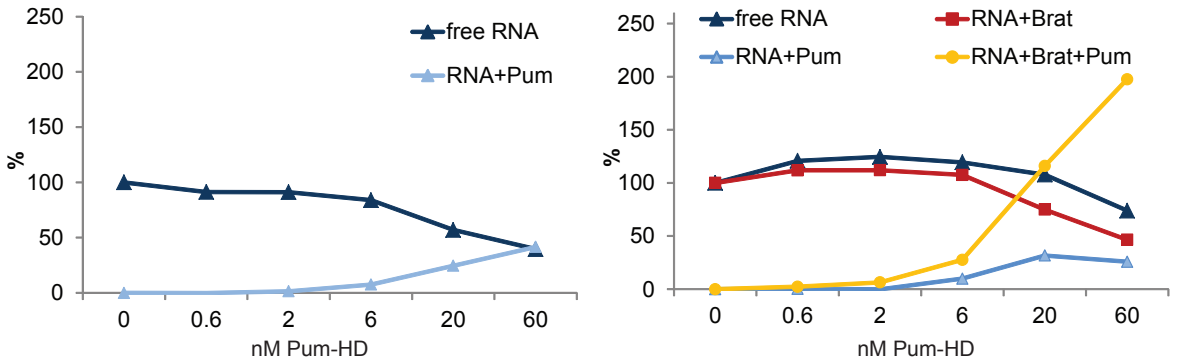
B



C



D

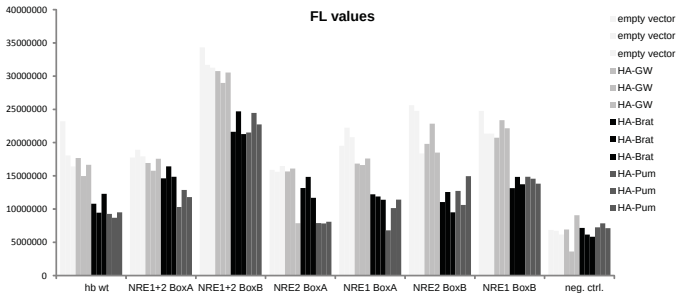


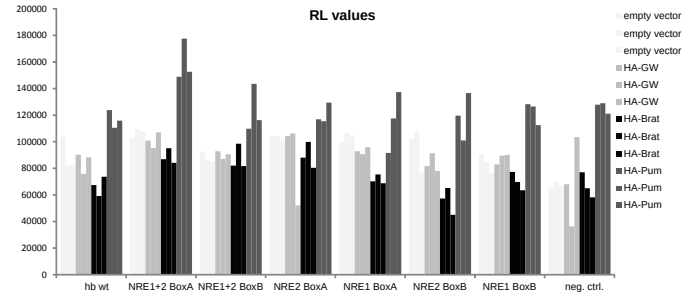
Supplemental Figure 6: Binding of Brat-NHL or Pum-HD to hb RNA facilitates the binding of the other

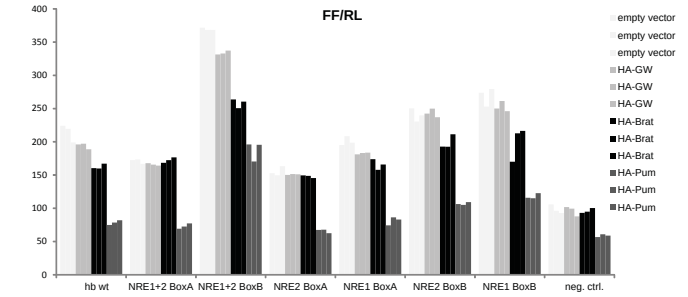
A) Pre-incubation of hb RNA with Pum-HD facilitates Brat-NHL binding. Indicated amounts of Brat-NHL were mixed with ^{32}P labeled hb RNA alone (lanes 1-7) or hb RNA pre-incubated with 10 nM Pum-HD (lanes 8-14). Complexes were separated by native gel electrophoresis. Experiment was performed three times and shown is one representative example. **B**) Quantification of gelshift experiment shown in A) Band intensities were quantified using quantity one software. Depicted is the percentage of free and protein bound RNA. **C**) Pum-HD preferentially binds to Brat-bound RNA. Indicated amounts of Pum-HD were mixed with ^{32}P labeled hb RNA alone (lanes 1-6) or hb RNA pre-incubated with 100 nM Brat-NHL (lanes 7-12). Complexes were separated by native gel electrophoresis. The experiment was repeated twice and shown is one representative experiment. **D**) Quantification of gelshift experiment shown in C). While a concentration of approximately 60 nM Pum-HD is needed to shift half of the free RNA into a slower migrating protein-RNA complex, a concentration of only 20 nM Pum-HD is sufficient to shift half of the Brat-bound-RNA, indicating that Pum-HD preferentially binds to Brat-bound RNA.

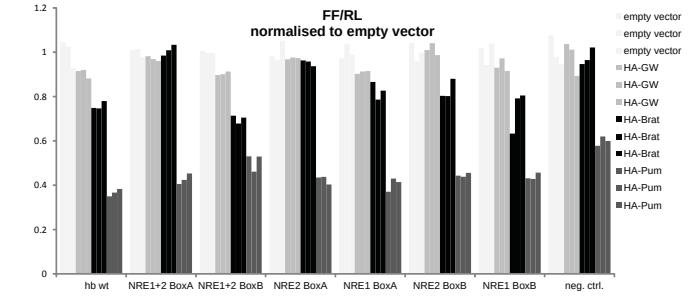
Loedige et al., Supplemental Figure 7

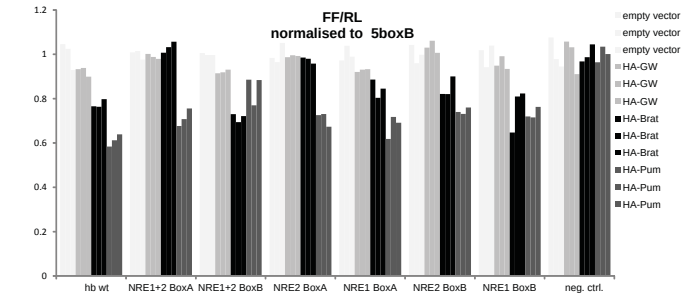
A



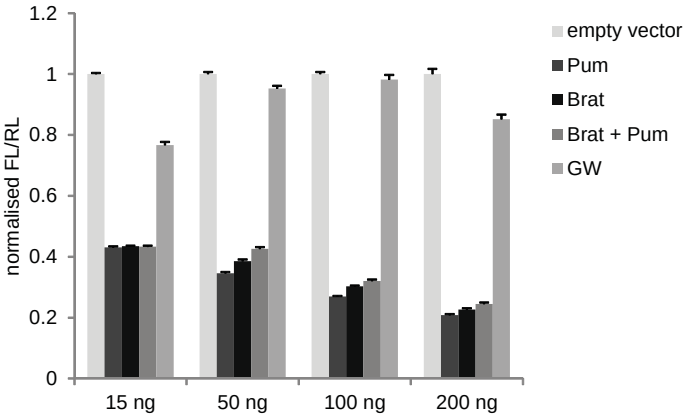








B

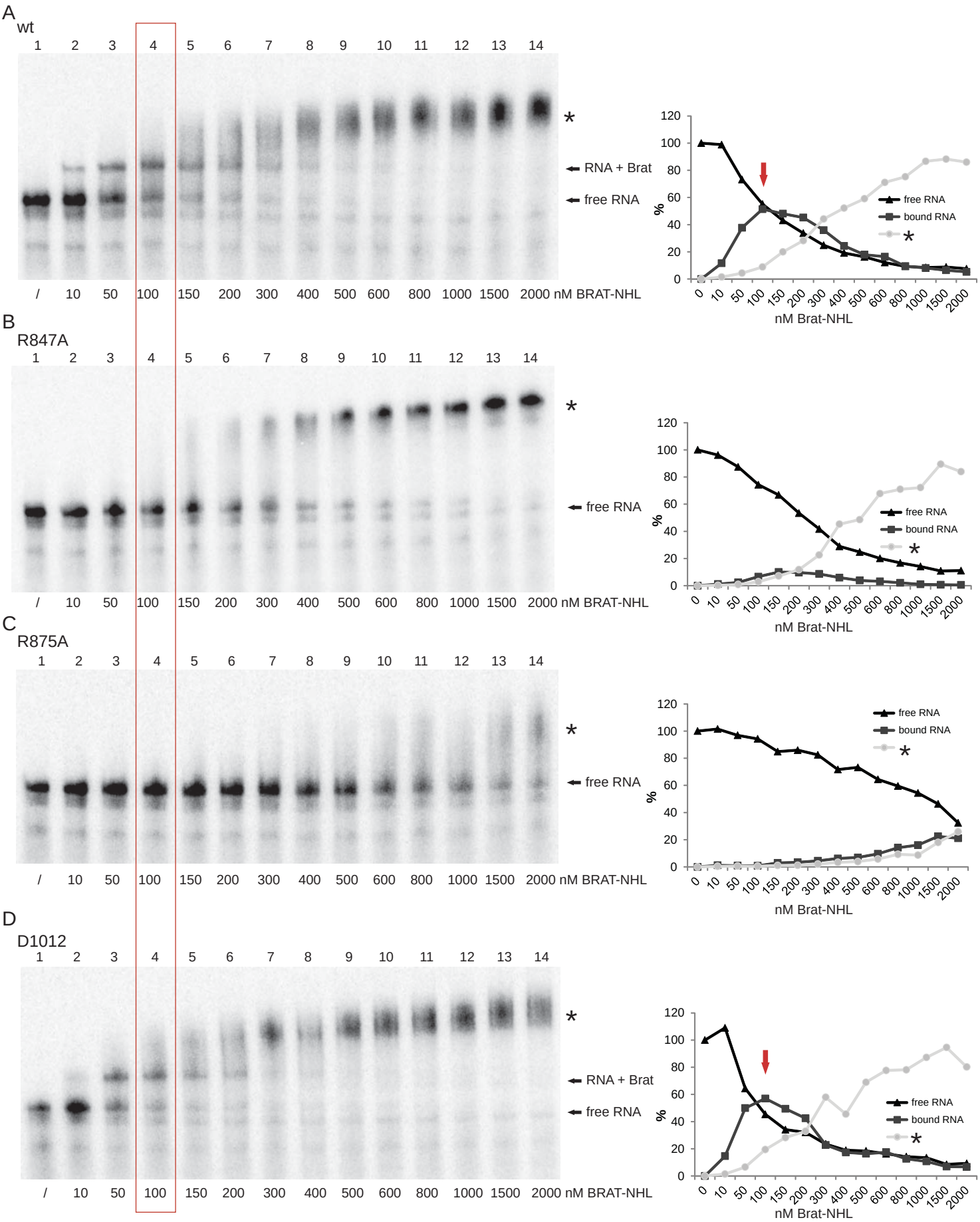


Supplemental Figure 7:
Supplementary to Figure 4.

A) Representative reporter gene experiment. Depicted are the raw FL and RL luciferase values before normalization and normalization steps for one representative experiment. The mean of three independent experiments is shown in Fig. 4B.

B) Dose-dependent repression of the hb reporter construct by HA-Brat and HA-Pum. Experiment was performed as described in Fig. 4B, except that indicated amounts (15 ng, 50 ng, 100 ng or 200 ng) of protein-coding plasmids were co-transfected along with the reporter constructs. Shown is the mean of three independent experiments.

Loedige et al., Supplemental Figure 8



Supplemental Figure 8: Supplementary to Figure 5D.

Titration experiments to determine the binding affinities of Brat-NHL (A), Brat-NHL R847A (B), Brat-NHL R875A (C) and Brat-NHL D1012 (D). Experiments were performed and quantified as described in Supplemental Figure 2. Experiments were performed twice and shown is one representative example

Loedige et al., Supplemental Table 1

Summary of Brat-NHL domain point mutants

mutant	surface	literature	crosslink to hb RNA	*in vitro* binding to hb RNA	repression of hb 3'UTR	repression when tethered
G774D	Top	Induces production of tumor-like neoplasms in the larval brain (Arama et al. 2000) Disrupts recruitment by Pum (Sonoda and Wharton 2001)		n.d.	impaired	yes
H802L	Top	Induces production of tumor-like neoplasms in the larval brain (Arama et al. 2000) Disrupts recruitment by Pum (Sonoda and Wharton 2001)		no binding	impaired	yes
C820A	Top	This study	cross-linked	n.d.	like wt	yes
Y829A	Top	Disrupts recruitment by Pum (Edwards et al. 2003)	cross-linked	impaired	like wt	yes
R847A	Top	Disrupts recruitment by Pum (Edwards et al. 2003)		no binding	impaired	yes
R875A	Top	Disrupts recruitment by Pum (Edwards et al. 2003)	in cross-linked peptide	no binding	impaired	yes
C890A	Top	This study	cross-linked	n.d.	impaired	yes
K891A	Top	This study	in cross-linked peptide	n.d.	impaired	yes
F916A	Top	This study	in cross-linked peptide	n.d.	impaired	yes
K809A	Bottom	Impairs interaction with d4EHP (Cho et al. 2006)	cross-linked	n.d.	like wt	yes
R837D	Bottom	Disrupts interaction with d4EHP (Cho et al. 2006)		n.d.	like wt	yes
Y859A	Bottom	Does not affect recruitment by Pum (Edwards et al. 2003)		binding like wt	like wt	yes
K882E	Bottom	Disrupts interaction with d4EHP (Cho et al. 2006)		n.d.	like wt	yes
K925A	Bottom	This study	in cross-linked peptide	n.d.	like wt	yes
E970A	Bottom	Does not affect recruitment by Pum (Edwards et al. 2003)		n.d.	like wt	yes
D1012A	Bottom	Does not affect recruitment by Pum (Edwards et al. 2003)		binding like wt	like wt	yes
K865	Side		cross-linked	n.d.	n.d.	n.d.
F866	Side		cross-linked	n.d.	n.d.	n.d.

Loedige et al., Supplemental Table 2

List of mutatgenesis primers

Brat-NHL domain mutants

Name	Sequence
G774D for	gtttggcgaattcgacgtgatggagggccagttcacg
G774D rev	ccctccatcacgtcgaattcgccaaacttgcagtgg
H802L for	cgaacaaccttcgcattcagatcttcgacaagg
H802L rev	gatctgaatgcgaaggttgttcgtatccgcaac
Y729A for	gctgctcgctccgaaccgcgtggcc
Y729A rev	ggttcggagcgagcagctgcgagtcacgc
R847A for	caccgaggcctcgcccacacaccag
R847A rev	gggcgaggcctcggtgacaataatatcgcc
R875A for	cagcatcctgccggcgtgaccgtggacaacaagg
R875A rev	ggtcacgccggcaggatgctgcagaatggtgg
Y859A for	ctacaatcaggccggccagtttgtgagg
Y859A rev	ctggccggcctgattgtagatctgtatctgg
E970A for	cgaacggggcgatcctcatcgcggacaacc
E970A rev	gaggatcgccccgttcgagttgatcccgacg
D1012 for	catggacgccggcagtgtggtgctggccagc
D1012 rev	ccacactgccggcgtccatgagcgccacatcg
R837D for	gtggtggacaattccggcgatattattgtcacc
R837D rev	ggaattgtccaccacggccacgcg
K882E for	ggacaacgagggacggatcattgtggtgg
K882E rev	ccgtccctcgttgtccacggtcac
K809A for	cttcgacgcggagggacgcttcaagttcc
K809A rev	gcgtccctccgcgtcgaagatctgaatgc
C820A for	gcgaggccggcaagcgtgactcgc
C820A rev	cgcttgccggcctcgccaaactgg
C890A for	ggtggaagccaaggtgatgcgtgtgatcatc
C890A rev	cgcatcaccttggcttccaccacaatgatcc
K891A for	ggtggaatgcgcggtgatgcgtgtgatcatc
K891A rev	cgcatcaccgcgcattccaccacaatgatcc
K925A for	caacgacgcgcaggagatcttcatcagcg
K925A rev	gatctcctgcgcgtcgttgaccaccacg
F916A for	gatctcctgcgcgtcgttgaccaccacg
F916A rev	cgttgggggcctcgaggtgcttagagcaccc

Hb RNA mutants

Name	sequence
NRE1 BoxB for	cagaatgggatggattcgtagcataagttttcc
NRE1 BoxB rev	gctacgaatccatcccattctggacaacg
NRE2 BoxB for	gttgtcgaaaatgggacggaagccaattaagc
NRE2 BoxB rev	ggcttccgtcccattttcgacaacaaaataatg
NRE1 BoxA for	gcctcatataaGcgggggccagaattgtatatattcgtagc
NRE1 BoxA rev	caattctggcccccgcttatatgaggctaggtctcc
NRE2 BoxA for	cattattgggggtcgaaaattgtacataagcc
NRE2 BoxA rev	caattttcgacccccccaataatgtttggaaaac
NRE2 BoxA mut1	ccaaacattattgggttgtcgaaaattgtacataagcc
NRE2 BoxA mut1	cgacaacccaataatgtttggaaaacttatgctacg
NRE2 BoxA mut2	ccaaacattattttggggtcgaaaattgtacataagcc
NRE2 BoxA mut2	caattttcgaccccaaaataatgtttggaaaacttatgc
NRE2 BoxA mut3	ccaaacattattatgttgtcgaaaattgtacataagcc
NRE2 BoxA mut3	caattttcgacaacataataatgtttggaaaacttatgc