Supplement - Structure and dynamics of the quaternary *hunchback* mRNA translation repression complex

Jakub Macošek[1,2], Bernd Simon[1], Johanna-Barbara Linse[3], Pravin Kumar Ankush Jagtap[1], Sophie Winter[1], Jaelle Foot[1], Karine Lapouge[4], Kathryn Perez[4], Mandy Rettel[5], Miloš T. Ivanović[3], Pawel Masiewicz[1], Brice Murciano[1], Mikhail M. Savitski[4], Inga Loedige[1,6], Jochen S. Hub[3], Frank Gabel[7] and Janosch Hennig[1,*]

[1] Structural and Computational Biology Unit, European Molecular Biology Laboratory Heidelberg, Heidelberg, 69117, Germany
[2] Present address: Biozentrum, University of Basel, Basel, 4056, Switzerland
[3] Theoretical Physics and Center for Biophysics, Saarland University, Saarbrücken, 66123, Germany
[4] Protein Expression and Purification Core Facility, European Molecular Biology Laboratory Heidelberg, Heidelberg, 69117, Germany
[5] Proteomics Core Facility, European Molecular Biology Laboratory Heidelberg, Heidelberg, 69117, Germany
[6] Present address: IRI Life Sciences, Humboldt University, Berlin, 10115, Germany
[7] Institut Biologie Structurale, University Grenoble Alpes, CEA, CNRS, Grenoble, 38044, France

* To whom correspondence should be addressed. Tel: +49 6221 387 8552; Email: janosch.hennig@embl.de

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

	hb complex SAXS	*hb* complex SANS
(a) Sample details		
Organism	*Drosophila melanogaster*	
Source	*E. coli* BL21 (DE3)	
Uniprot sequence ID	P25822, Q8MQJ9, P25724	
Molecular mass from chemical composition (Da)	87 798	
SEC-SAXS column	Superdex 200 Increase 10/300 GL	n/a
Loading concentration (mg·ml^{-1})	7.2	3.7, 3.7, 5.0, 4.3, 4.0, 4.0, 4.0
Injection volume (ul)	750	n/a
Flow rate (ml·min^{-1})	0.6	n/a
Solvent	50 mM Tris, 150 mM NaCl, 1 mM DTT, 3% glycerol, pH 7.4	50 mM Tris, 150 mM NaCl, 1 mM DTT, pH 7.4
(b) Data collection parameters		
Instrument		ILL Grenoble D-22
Wavelength (Å)	1.23 (x-rays)	6 (neutrons)
Sample-detector distance (m)	3	4
q measurement range (Å^{-1})		
Monitoring for radiation damage	frame-by-frame comparison	n/a
Exposure time	1 second/frame	60 minutes
Sample configuration	continuous flow in quartz capillary, 1.7 mm path length	Hellma® 100QS quartz cuvette, 1 mm path length
Sample temperature	25 °C	20 °C
(c) Software used for data reduction, analysis and interpretation		
Data reduction	Data reduction: BsxCube Solvent subtraction: PRIMUS	Data reduction: GRAS$_{ans}$P Solvent subtraction: PRIMUS
Basic analyses: Guinier, *P(r)*, V_P	PRIMUS, ScÅtter	PRIMUS, ScÅtter
Structure modelling	CRYSOL 2.8.3, EOM, custom MD	CRYSON 2.7, custom MD
Molecular graphics software	Pymol	
(d) Structural parameters		
Guinier analysis		
$I(0)$ (cm^{-1})	15 550 +- 1.4	0.150 +- 0.002, 0.059 +- 0.001, 1.172 +- 0.005, 0.306 +- 0.002, 1.445 +- 0.008, 0.639 +- 0.002, 0.266 +- 0.002
R_g (Å)	37.36 +- 0.19	38.86 +- 2.80, 39.58 +- 5.22, 38.10 +- 0.91, 36.14 +- 0.88, 40.42 +- 1.15, 40.59 +- 0.70, 41.07 +- 1.77
q_{min} (Å^{-1})	0.00008	0.00028, 0.00028, 0.00028, 0.0028, 0.00028, 0.00033, 0.00028
$q \cdot R_g$ max (q_{min} = 0.0066 Å^{-1})	1.2932	1.2962, 1.2361, 1.2703, 1.2822, 1.2623, 1.2680, 1.2828
P(r) analysis		
$I(0)$ (cm^{-1})	15 360	0.089, 0.032, 0.775, 0.202, 0.986, 0.517, 0.189
R_g (Å)	37.96	37.08, 40.10, 36.46, 35.77, 40.13, 41.21, 41.79
d_{max} (Å)	127	119, 119, 118, 118, 123, 123, 124

q range (Å^{-1})	0.010138-0.272643	0.016585-0.131863, 0.048036-0.131875, 0.022884-0.133931, 0.022895-0.133994, 0.022891-0.127659, 0.025019-0.152867, 0.022893-0.154918
χ^2	1.31	1.31, 1.22, 2.88, 1.09, 1.26,1.30,1.92, 1.45
Porod volume (Å^{-3})	144 184	193 673, 197 698, 224 621, 217 158, 220 108, 217 188, 201 716
(e) Atomistic modelling		
CRYSOL/CRYSON (default parameters, no constant subtraction)		
χ^2	4.17/5.958	1.52/1.55; 1.62/1.67; 14.92/13.36; 20.57/20.17; 46.48/62.89; 48.59/59.67; 6.65/11.31
Predicted R_g (Å)	38.26/38.02	37.66/37.45; 38.15/38.27; 35.67/35.54; 34.99/34.88; 39.03/38.65; 39.62/39.15; 40.60/39.98
Vol (Å), Ra (Å), Dro (e Å^{-3})	110512, 1.4, 0.01/110512, 1.4, 0.007	110015, 1.4, -0.118/110015, 1.4, -0.118; 99569, 1.781, 0.00/101558, 1.781, 0.00; 110015, 1.781, -0.004/110015, 1.781, -0.004; 95092, 1.4, 0.467/95092, 1.4, 0.467; 110015, 1.781, -0.004/110015, 1.781, -0.004; 95092, 1.4, 0.354/95092, 1.4, 0.354; 95092, 1.4, 0.763/95092, 1.4, 0.763
EOM		
χ^2	0.75	n/a
No. of representative structures	7	n/a

Table S1. Statistics and details of SAS measurements of the *hb* complex complex according to community guidelines (5). The SANS parameters are always given for each sample in the following order: 1B1P1N 0% D2O, 1B1P1N 66% D2O, 1B2P2N 0% D2O, 1B2P2N 41% D2O, 2B2P1N 0% D2O, 2B2P1N 33% D2O, 2B2P1N 62% D2O. The values stated in atomistic modelling are for the representative models of the major/minor ensemble obtained from rigid-body modelling using only SAS data.

Protein 1	Residue 1	Protein 2	Residue 2	id-score	Satisfied
Intramolecular cross-links					
Brat-NHL	Brat-NHL	865	882	32.32	yes
Brat-NHL	Brat-NHL	865	891	28.79	yes
Brat-NHL	Brat-NHL	865	809	34.65	yes
Brat-NHL	Brat-NHL	865	822	43.42	yes
Brat-NHL	Brat-NHL	891	882	25.73	yes
Brat-NHL	Brat-NHL	939	865	48.81	yes
Brat-NHL	Brat-NHL	939	891	30.18	yes
Brat-NHL	Brat-NHL	814	809	33.12	yes
Brat-NHL	Brat-NHL	822	865	37.14	yes
Brat-NHL	Brat-NHL	822	809	28.66	yes
Nanos-ZnF	Nanos-ZnF	306	341	37.18	*n. a.*
Nanos-ZnF	Nanos-ZnF	296	304	35.25	*n. a.*
Nanos-ZnF	Nanos-ZnF	296	341	31.86	*n. a.*
Nanos-ZnF	Nanos-ZnF	341	304	30.66	*n. a.*
Nanos-ZnF	Nanos-ZnF	341	306	30.79	*n. a.*
Nanos-ZnF	Nanos-ZnF	341	341	31.69	no (self-link)
Nanos-ZnF	Nanos-ZnF	361	306	29.99	*n. a.*
Nanos-ZnF	Nanos-ZnF	361	341	34.63	yes
Nanos-ZnF	Nanos-ZnF	361	366	42.47	yes
Nanos-ZnF	Nanos-ZnF	361	376	26.04	yes
Nanos-ZnF	Nanos-ZnF	299	341	29.4	*n. a.*
Nanos-ZnF	Nanos-ZnF	366	306	47.68	*n. a.*
Nanos-ZnF	Nanos-ZnF	366	341	41.08	yes
Nanos-ZnF	Nanos-ZnF	376	341	36.47	no
Nanos-ZnF	Nanos-ZnF	384	341	35.24	*n. a.*
Pum-HD	Pum-HD	1331	1331	46.11	no (self-link)
Pum-HD	Pum-HD	1339	1398	31.12	yes
Pum-HD	Pum-HD	1347	1331	56.39	yes
Pum-HD	Pum-HD	1398	1331	26.57	yes
Pum-HD	Pum-HD	1398	1413	36.83	yes
Pum-HD	Pum-HD	1398	1423	33.01	*n. a.*
Pum-HD	Pum-HD	1399	1423	28.5	*n. a*
Pum-HD	Pum-HD	1403	1398	26.05	yes
Pum-HD	Pum-HD	1403	1413	29.37	yes
Pum-HD	Pum-HD	1418	1398	27.37	yes
Pum-HD	Pum-HD	1423	1398	29.06	*n. a*
Intermolecular cross-links					
Brat-NHL	Nanos-ZnF	865	341	38.15	no
Brat-NHL	Nanos-ZnF	882	341	35.17	no
Brat-NHL	Nanos-ZnF	891	341	33.68	all
Brat-NHL	Nanos-ZnF	822	341	27.97	no
Brat-NHL	Pum-HD	809	1413	33.54	no
Nanos-ZnF	Brat-NHL	341	1093	27.38	3543 (S6A)
Nanos-ZnF	Brat-NHL	341	1097	27.39	no
Pum-HD	Brat-NHL	1398	865	39.7	4876 (S6C)
Pum-HD	Brat-NHL	1398	891	30.73	3543 (S6A), 4876 (S6C)
Pum-HD	Brat-NHL	1398	809	33.7	no
Pum-HD	Brat-NHL	1398	814	30.19	4876 (S6C)
Pum-HD	Brat-NHL	1398	822	31.19	3543 (S6A), 4014 (S6B)
Pum-HD	Nanos-ZnF	1339	306	30.54	*n. a.*
Pum-HD	Nanos-ZnF	1339	365	28.82	yes
Pum-HD	Nanos-ZnF	1398	341	35.4	no
Pum-HD	Nanos-ZnF	1413	341	25.23	yes
Pum-HD	Nanos-ZnF	1423	306	33.53	*n. a.*
Pum-HD	Nanos-ZnF	1423	296	27.17	*n. a.*
Pum-HD	Nanos-ZnF	1423	299	28.14	*n. a.*
Pum-HD	Nanos-ZnF	1132	341	36.05	no

Table S2. Intermolecular cross-links of the *hb* complex. The cross-links were obtained using 0.5 mM DSSS and come from two replicates. The column 'id-score' gives the score from xQuest/xProphet, whereas the column 'Satisfied' shows if a cross-link was satisfied in any of the models in the ensemble obtained by modelling using SAS and XL/MS data. The numbers in the 'Satisfied' column are the number of the model in which the cross-link was satisfied, the brackets indicate the panel of the supplementary figure in which the model is shown. '*n. a.*' in the 'Satisfied' column means one or both residues

concerned are not built in the source crystal structures, thus the distance between the residues cannot be measured. Pum-HD and Nanos-ZnF remained rigid during the modelling, therefore satisfaction of their cross-links was not evaluated.

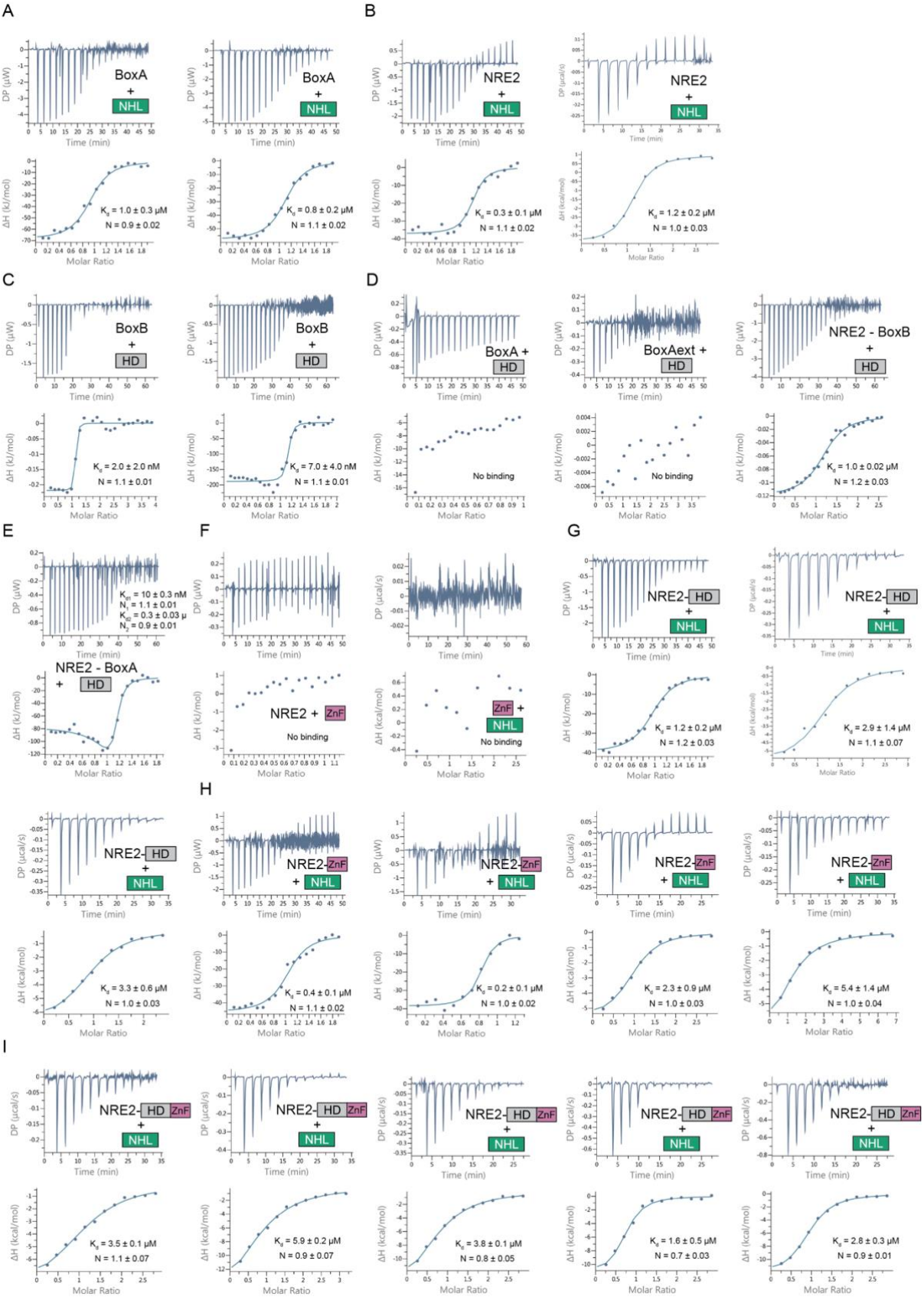

Figure S1. Replicates of ITC measurements. The panels correspond to measurements of the following interactions: **(A)** Brat-NHL-BoxA interaction, **(B)** Brat-NHL-NRE2 interaction, **(C)** Pum-HD-BoxB interaction, **(D)** the interaction of Pum-HD with BoxA, BoxAext and NRE2 minus BoxB (NRE2-BoxB), **(E)** the interaction of Pum-HD with NRE2 minus BoxA (NRE2-BoxA) **(F)** Nanos-ZnF-NRE2 (left) and Nanos-ZnF with Brat-NHL (right), **(G)** Brat-NHL with a preformed NRE2-Pum-HD complex, **(H)** Brat-

NHL with a mixture of NRE2 and Nanos-ZnF in the cell and **(I)** Brat-NHL with a preformed complex of NRE2-Pum-HD-Nanos-ZnF.

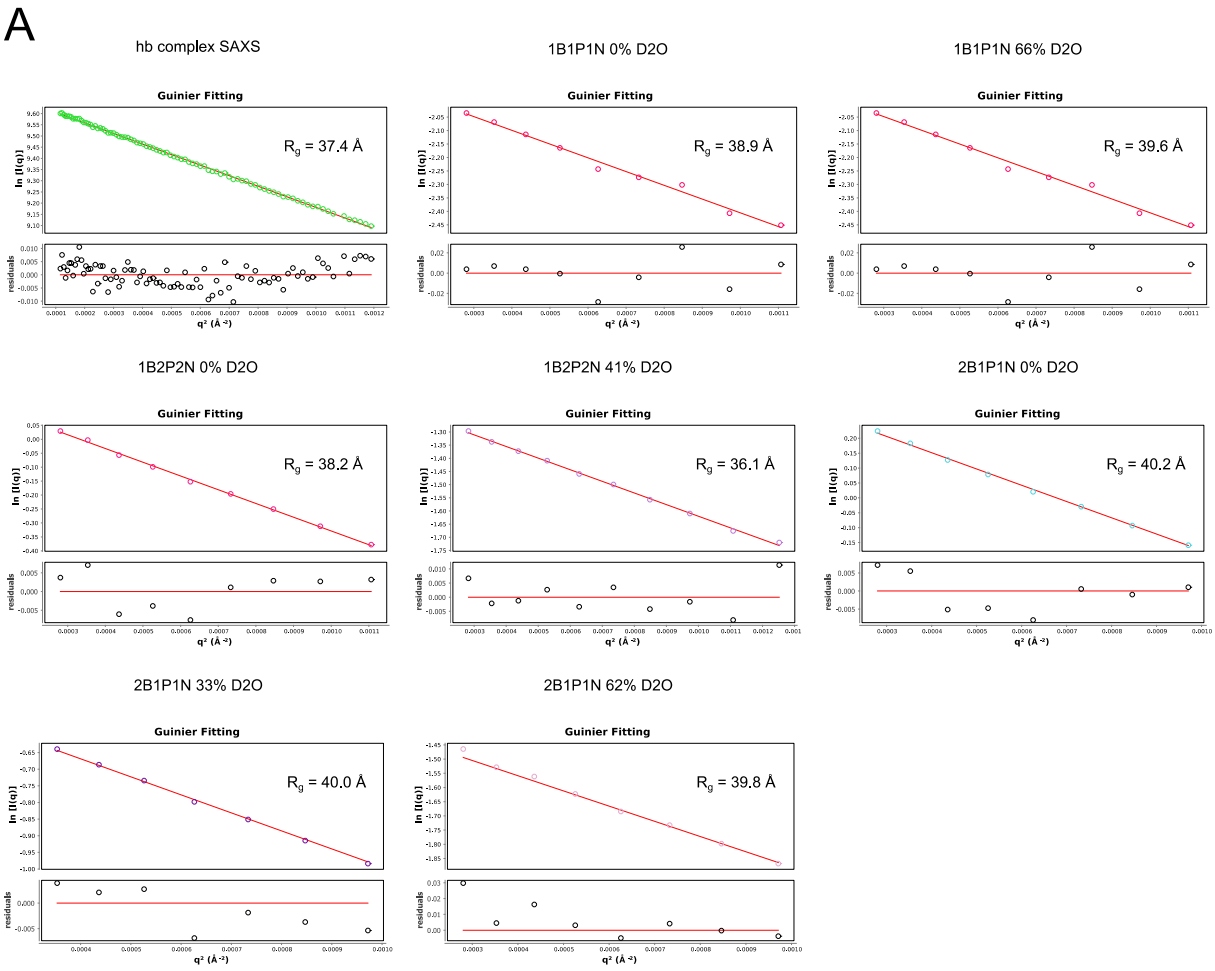

Figure S2. Characterization of hb complex by small angle scattering (SAS). **(A)** Guinier fit from the individual scattering curves. All fits obey $q*R_g < 1.3$. **(B)** Distance distribution function (P(r)) of Brat-NHL-RNA complex and Pum-HD-Nanos-ZnF-RNA complex. P(r) were calculated from the available high-resolution structures of Brat-NHL-RNA complex (PDB ID: 4zlr) (1) and Pum HD-Nanos ZnF-RNA complex (PDB ID: 5kl1) (2) using ScÅtter (3).

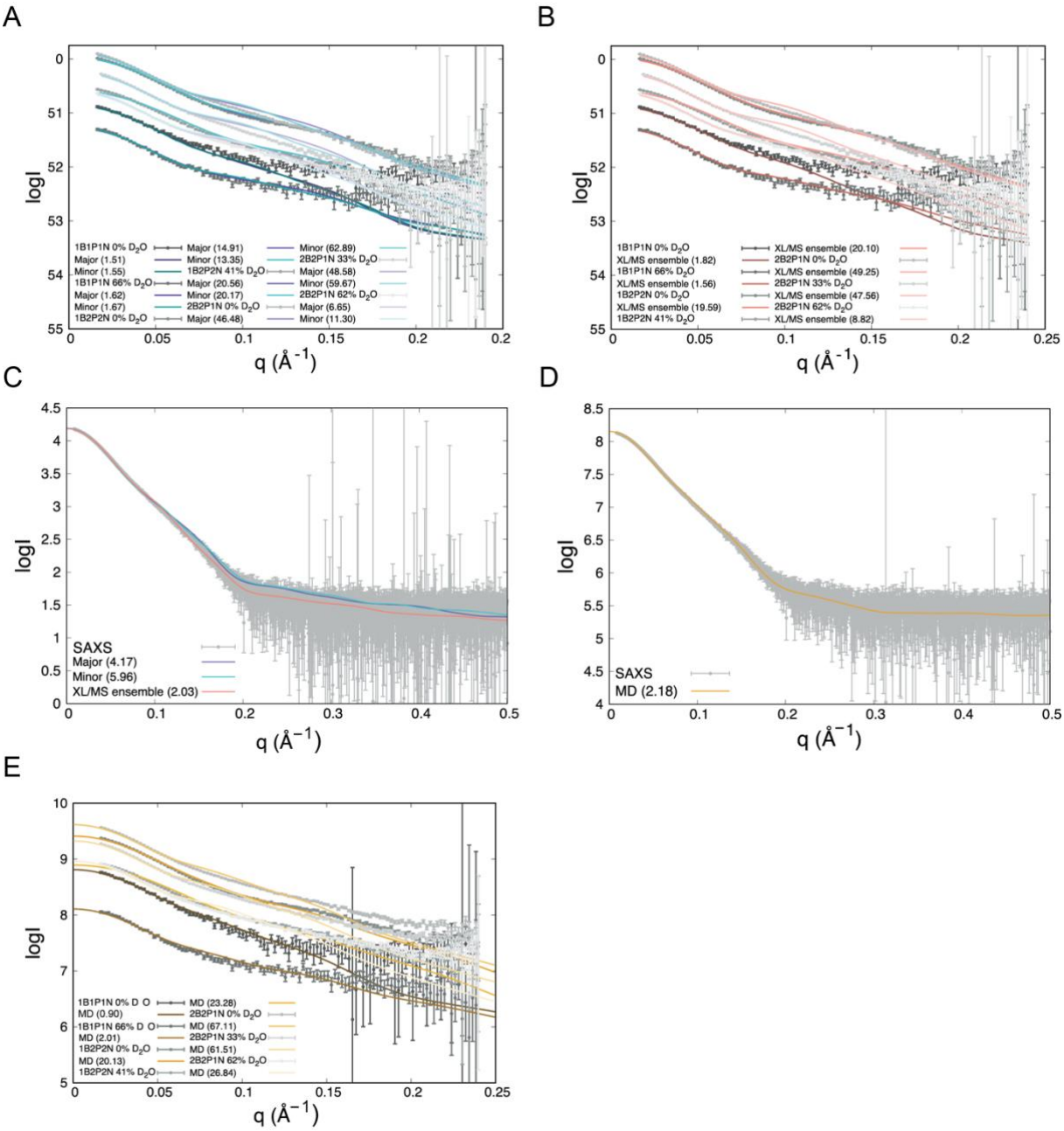

Figure S3. Fits of *hb* complex models to the experimental data. **(A)** Fits of the major and minor cluster of models to the experimental Small-Angle Neutron Scattering (SANS) data. For each cluster a fit of a representative models is shown. A representative model of the cluster was selected as the model closest to the mean coordinates of the aligned cluster. **(B)** Fit of the selected ensemble of *hb* complex models generated by modelling using SAS and XL/MS data to the experimental SANS data. The fit of the ensemble to the experimental data is illustrated by the fit of the model closest to the mean coordinates of the aligned ensemble. **(C)** Fits of back-calculated scattering curves of representative models of *hb* complex ensemble to the experimental SAXS curve. 'Minor' and 'Major' denotes representative models of the two clusters in Figure 4E in the main manuscript and 'XL/MS ensemble' denotes the representative model of the ensemble in Figure 5C of the main manuscript. The representatives are the models closest to the mean structure of the ensemble. **(D)** Comparison of SAXS curve computed from all-atom MD with the experimental SAXS curve. **(E)** Fit of the ensemble of *hb* complex models obtained by all-atom MD simulations. The curve fitted in each condition is a back-calculated scattering curve of frames from the time interval between 30 and 110 ns from the all-atom MD simulations. For clarity, curves for 1B2P2N and 2B2P1N were offset along the Y-axis by multiplying the curves with factors of 10 and 100, respectively. To avoid any influence by aggregation in the SANS data, the fits were restricted to q>0.07Å (2B2P1N) or to q>0.05Å (all other sets). The number in the brackets in all panels always reports the values of χ^2 values of the fit.

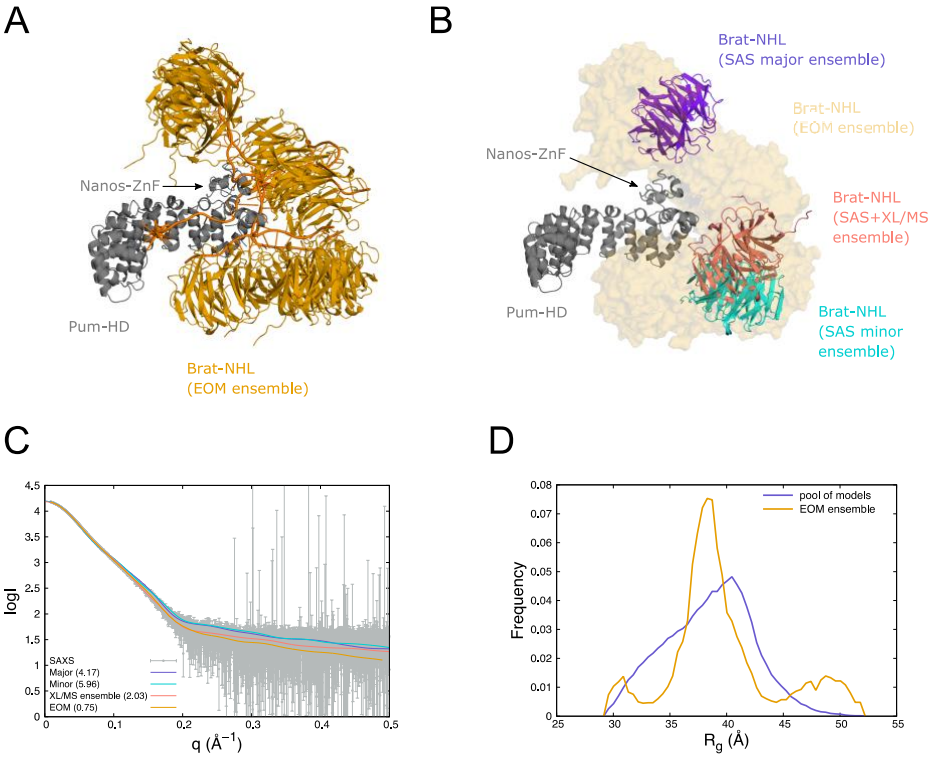

Figure S4. Modelling the *hb* complex by Ensemble Optimization Method (EOM) (4). **(A)** Ensemble of *hb* complex models obtained from EOM using the SAXS curve of the complex. EOM used the starting pool of models generated during rigid-body modelling to search for an ensemble of models to fit the SAXS data as a mixture. Seven models of the best fitting ensemble are shown. **(B)** Overlay of *hb* complex models. Representative models of ensembles from rigid-body are superimposed on the EOM ensemble from (A) with Brat-NHL in semi-transparent surface representation. All models in the ensembles are always superimposed on Pum-HD and Nanos-ZnF. **(C)** Fits of back-calculated scattering curves of representative models of *hb* complex ensembles to the experimental SAXS curve. 'Minor', 'Major', and 'XL/MS ensemble' denotes the representative models of the ensembles obtained from the rigid-body modelling. Numbers in the brackets state the χ^2 value of the fit. **(D)** R_g distribution plot from EOM of the *hb* complex. The plot shows the distribution of the R_g values of the initial pool of models (generated during rigid-body modelling), and the R_g distribution of the ensemble of models selected by EOM to fit the SAXS data (EOM ensemble). The final ensemble has 7 models with an R_{flex} of 88.3% (R_{flex} of the pool 89.5%) and R_σ of 5.26.

1. Loedige, I., Jakob, L., Treiber, T., Ray, D., Stotz, M., Treiber, N., Hennig, J., Cook, K.B., Morris, Q., Hughes, T.R. *et al.* (2015) The Crystal Structure of the NHL Domain in Complex with RNA Reveals the Molecular Basis of Drosophila Brain-Tumor-Mediated Gene Regulation. *Cell Rep*, **13**, 1206-1220.
2. Weidmann, C.A., Qiu, C., Arvola, R.M., Lou, T.F., Killingsworth, J., Campbell, Z.T., Tanaka Hall, T.M. and Goldstrohm, A.C. (2016) Drosophila Nanos acts as a molecular clamp that modulates the RNA-binding and repression activities of Pumilio. *Elife*, **5**.
3. Forster, S., Apostol, L. and Bras, W. (2010) Scatter: software for the analysis of nano- and mesoscale small-angle scattering. *Journal of Applied Crystallography*, **43**, 639-646.
4. Tria, G., Mertens, H.D., Kachala, M. and Svergun, D.I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ*, **2**, 207-217.
5. Trewhella, J., Duff, A.P., Durand, D., Gabel, F., Guss, J.M., Hendrickson, W.A., Hura, G.L., Jacques, D.A., Kirby, N.M., Kwan, A.H. *et al.* (2017) 2017 publication guidelines

for structural modelling of small-angle scattering data from biomolecules in solution: an update. *Acta Crystallogr D Struct Biol*, **73**, 710-728.