Figure S1. Structural Features of the Hermes Octameric Assembly, Related to Figure 1

(A) Experimental electron density after averaging and density modification contoured at 1σ in the region of the active site.

(B) Composite simulated annealed omit map contoured at 1σ.

(C) Symmetry axes of the octamer. The green axes (labelled 1-4) are through the intertwined domain, and the white axes (labelled a and b) are through the small swapped interfaces. All axes are 200Å long. Axis b coincides with a crystallographic two-fold axis.

(D) As (A) but rotated 90°.

(E) Superposition of the apo-form of Hermes (light pastel colors and black labels; PDB ID 2BW3) and the TIR-bound form (domain colors correspond to those in Figure 1A). The structures were aligned by aligning the intertwined domains, residues 79-155.

Figure S2. Full-length Hermes is Also Octameric, Related to Figure 1C

(A) (top) Size exclusion chromatography of *E. coli*-expressed Hermes run on two tandem BioSep-SEC-S 3000 columns. The column void peak at ~95 min contains mostly DNA. The elution time of Hermes, 123.9 min, corresponds to a MW of ~570 kDa, based on the elution positions of gel filtration standards (669 kDa, thyroglobulin, at 116.7 min; 443 kDa, apoferritin at 135.3 min). The theoretical molecular mass of an octamer is 70.1 kDa x 8 = 560.9 kDa. (bottom) SDS-PAGE fractions corresponding to the eluted peak centered at 124 min.

(B) SEC-MALS chromatogram. The black line represents the light scattering signal, and the red line is the molar mass as a function of elution time. Hermes concentration was 1.1 mg/ml.

(C) (top) Size exclusion chromatography on Superose 6 of crude extracts of *S. cerevisiae*-expressed Hermes. (bottom) Western blot of fractions indicated by the red lines. The elution time of Hermes, 13.4 min, corresponds to a calculated MW of ~515 kDa.

(D) Size exclusion chromatography of Sf9 cell-expressed Hermes run on a Superose 6 column. The highest point of the peak at 26.7 min corresponds to a MW of ~635 kDa based on the elution positions of gel filtration standards (669 kDa, thyroglobulin at 26.4 min; 443 kDa, apoferritin at 30.0 min), and there is a higher MW shoulder at ~24.5 min. Re-injection of fraction 13 (starred) results in a symmetric peak at 27.2 minutes, corresponding to a MW of 610 kDa.

Figure S3. SAXS and EM Data Analysis of Full-length Hermes, Related to Figure 2

(A) Comparison of the inter-atomic distance probability distribution profiles of full-length Hermes calculated from the experimental and best-fitted model SAXS data, related to Figure 2A. P(r) profiles are calculated via regularized Fourier transform fit using the GNOM program (Svergun, 1992) from the experimental SAXS data (shown in black) and those calculated from the best-fitting structural model (shown in red).

(B) A field of negative-staining EM showing crosslinked WT Hermes particles. Black arrows point to some top view examples and white arrows point to side views. Scale bar corresponds to 100 nm.

(C) Examples of raw particles that were included for averaging for producing the 2D-class averages in Figure 2B. Scale bar corresponds to 25 nm.

(D) The resolution of the reconstruction was estimated according to the FSC (Fourier Shell Correlation) criterion (Scheres and Chen, 2012). The data were divided into two half sets which were both reconstructed with starting models taken as the crystal structure of the octamer but with (separately) randomized phases to ensure independence, and compared by calculating the FSC between them. The resolution was taken from the spatial frequency at which the FSC curve fell below a threshold value of 0.143.

(E) A rigid-body fit of the Hermes octamer into the EM map shows generally good agreement between the crystal structure and the reconstruction.

(F) A fit in which each dimer was treated as a separate rigid body, and 8 copies of the homology-modeled BED domains were included as occupying the central density present in the reconstruction. These results (E & F) attest to the consistency of the data (EM and X-ray), albeit with some discrepancies affecting the finer details that may arise from the flexibility of the truncated region, distortions from air-drying the negatively stained specimens, and the limited resolution of the reconstruction.

Figure S4. Comparison of the Intertwined Dimerization Domains of Hermes and RAG1, Related to Figure 3.

PDB ID for RAG1 nonamer binding domain (NBD) is 3GNA. The two red spheres indicates the positions of each 3'-OH group of the transferred strands of DNA bound to Hermes. Shown in ball-and-stick representation are the side chains of Hermes R107, S110, and R149, and RAG1 residues R391, R393, R402, K405, R407, and K412 (for clarity, labelled only for the light pink protein chain).

Figure S5. Details of the pRX1-Her Plasmid Cleavage Assay, Related to Figure 4C

The miniHermes element containing 30 bp of *Hermes* LE and 30 bp of RE flanking the KanMx gene was generated by PCR amplification of the KanMX cassette from pFA6a-KanMX6 (Bahler et al., 1998) using PCR primers that contained a perfect target site duplication: XhoI-TSC-L30KanMx (ACTGCTCGAGGTGCGCATCAGAGAACAACAACAAGTGGCTTATTTTGATTAATTAAGGCGCGCCAGAT) and XhoI-TSD-R30KanMx (ACTGCTCGAGATGCGCACCAGAGAACTTCAACAAGCCACAGGCAAACGCACTGGATGGCGGCGTTAGT). The PCR product was then cloned into the pRX1 XhoI site (Yu and Gabriel, 1999).

Figure S6. Hermes Dimers are Inactive *in vivo*, Related to Figure 4.

(A) Excision in *Drosophila* S2R+ cells, monitored by PCR.

(B) Results of *in vivo* excision assay in *S. cerevisiae* using donor plasmid containing authentic transposon ends. The assay is based on the excision of a mini-transposon embeded in an actin intron. Upon excision, the intron is restored and positive events can be quantified by selection by Ura+.

(C) Time course of Hermes expression and subsequent degradation in *S. cerevisiae* after induction by addition of 2% galactose. Hermes protein is detected using an anti-HA antibody. Note in all cases the similar amount of material at T = 3 hr.

(D) Cellular localization of full-length and dimeric Hermes in HEK293 cells. Detected using an anti-Flag antibody three days post-transfection.

Figure S7. Details of Assays, Related to Figure 6

(A) The pHL2577 plasmid cleavage assay, related to Figure 6B and C.

(B) (C) Strand transfer as a function of LE length. Assay conditions are as described in Zhou et al. (2004).

(D) SDS-PAGE of nuclear extract samples, related to Figure 6E and F. Coomassie-stained SDS-PAGE gel showing the nuclear extract (NE) and cytoplasmic fractions (cyto) from DEV8 (Hermes expressing) and EGFP (negative control) induced cell lines. The band at ~70kDa in the DEV8 NE lane corresponds to the *Hermes* transposase.

Table S1 - Crystallographic Statistics, Related to Figure 1

Data collection	**Native**	**Ta_6Br_{12}**
Space group	$P3_221$	$P3_221$
Cell dimensions		
a=b, c (Å)	265.2, 154.7	264.5, 157.6
α=β, γ (°)	90.0, 120.0	90.0, 120.0
Wavelength (Å)	1.5418	1.2037
Resolution (Å)	3.3	3.7
R_{sym} (%)	18.3	10.3
I / σ(I)	12.5	13.2
Number of measurements:	1243565	765934 (*)
Unique data:	94693	131343 (*)
Completeness (%)	98.7	99.8
Redundancy	13.1	5.8
Maximum likelihood phasing		
Resolution (Å)	30.0-5.0	
No. of sites	8	
Phasing power (isomorphous/anomalous)	0.72/1.01	
Figure of merit (centric/acentric)	0.22/0.29	
Phase extension, density modification		
Resolution (Å)	3.3	
Figure of merit (DM)	0.747	
Refinement		
Resolution (Å)	3.4	
No. reflections	86915	
R_{work} / R_{free} (%)	21.2/25.4	
No. atoms Protein/DNA	16209/2524	
No. atoms ion	4	
B-factors Protein/DNA	93.9/120.3	
B-factors ligand/ion	59.7	
R.m.s. deviations Bond lengths (Å)	0.016	
R.m.s. deviations Bond angles (Å)	2.00	
Ramachandran plot (% favored, allowed, generous and disallowed)	82.1/15.6/1.8/0.6	

(*) Friedel's law was assumed to be false during internal scaling of the Ta_6Br_{12} data set. The current model contains R81-N465, S493-K609; M77-I464, S493-K609; L83-I464, S493-K609 and L83-S469, S488-K609 for the four chains in ASU and all nucleic acid residues.

Table S2 - Somatic Transposition Frequencies of Hermes Mutants in *D. melanogaster* Embryos Normalized to *piggyBac* Transposition, Related to Figure 5C

Hermes	*Hermes* DT	*pBac* DT	*Hermes* events	*pBac* events	*Hermes* frequency	*pBac* frequency	Fold difference	SD *Hermes*	SD *pBac*
WT	163,000	152,000	67	57	$4.1x10^{-4}$	$3.8x10^{-4}$	+ 1.05	$2.0x10^{-4}$	$3.4x10^{-4}$
K292A	11,000	29,000	19	119	$1.8x10^{-3}$	$4.1x10^{-3}$	- 2.6	$1.0x10^{-3}$	$2.6x10^{-3}$
K299A	262,400	323,600	1	132	$3.8x10^{-6}$	$4.1x10^{-4}$	- 108	$4.0x10^{-6}$	$3.6x10^{-4}$
K300A	96,800	143,800	41	137	$4.2x10^{-4}$	$9.5x10^{-4}$	- 3.0	$4.6x10^{-4}$	$1.4x10^{-3}$
R306A	55,600	64,800	5	45	$8.9x10^{-5}$	$6.9x10^{-4}$	- 7.8	$1.6x10^{-4}$	$3.6x10^{-5}$

Interplasmid transposition frequencies in injected Drosophila embryos for WT Hermes and target binding cleft mutants. For each of three injections, each of the samples was co-injected with the *piggyBac* (*pBac*) transposon donor and helper as an internal control used to normalize the data between different injection sets and samples that arise due to inherent variability in the assay. The frequency was calculated by dividing the number of independent events by the donor titer (DT) in each of the injection sets. The "Fold Difference" is the difference in the transposition rate between the Hermes transposition frequency and the transposition frequency of the *piggyBac* internal control with which it was co-injected. SD is standard deviation.

Table S3 - *D. melanogaster* Transformation Frequencies of Hermes Mutants, Related to Figure 5D

Hermes	Injected	Survived	Eclosed	Crossed	Fertile	Sterile	# transgenic	Sterility rate	Transformation rate
WT	679	254	77	69	52	17	15	24.6%	28.8%
K292A	667	308	164	140	104	36	1	25.7%	0.96%
K299A	440	146	102	68	57	11	0	16%	0%
K300A	752	237	109	100	68	32	3	32%	4.4%
R306A	605	258	123	88	67	21	17	23.9%	25.4%

For each of three trials, the transposition-dependent transformation frequency for each sample was calculated by dividing the number of transgenics by the number of fertile crosses.

Table S4 - Transposition Frequencies of Hermes Elements with Ends of Differing Lengths and Complexities, Related to Figure 6.

Transposon	LE length (bp)	RE length (bp)	# Expts	# donor plasmids screened	# Transpositions	Frequency of Transposition
pHDG1	444	384	3	669,500	5	7.47×10^{-6}
pHDG-EGFP	305	307	10	2,236,248	7	3.13×10^{-6}
pHDG8	30	30	5	1,810,283	0	0
pHDGLL	305/305	-	4	1,560,125	0	0

Table S5 - Characterization of *in vitro* Transposition Products, Related to Figure 7

Number	Left duplication	Left ITR	Right ITR	Right duplication
1		+	+	GTAGGTTC
2	GGATATAA	+	18 bp del.	GGATATAA
3	AAATACAA	+	152 bp del.	AAATACAA
4	TCCCATAC	+	+	TCCCATAC
5	ACACGCAT	+	+	ACACGCAT
6	GGCTTTAT	+	+	GGCTTTAT
7	ATTCAGTA	+	+	ATTCAGTA
8	TAAAACAC	+	37 bp del.	TAAAACAC
9	ATATAAAC	+	+	ATATAAAC
10	ACGCAATT	-	37 bp del.	TTTGATGA
11	AACGATAT	+	37 bp del.	AACGATAT
12	ATCCCAAG	+	+	ATCCCAAG
13	CTATGTGT	+	+	CTATGTGT
14	CTATGTGT	+	+	CTATGTGT
15	2 L ends	++	++	2 R ends
16	GTACAGAG	+	+	GTATGCAC
17	GTTGGTTG	+	+	ATATAATC
18	GTCTGAAC	+	+	GTACAGAG
19	TATCTGAT	+	+	TATCTGAT
20	ATGGTTAC	+	+	ATAAAGAT
21	GTACTTGG	+	+	GTACTTGG
22	CTTTGGGA	+	+	CTTTGGGA
23	ACGCCGAT	+	+	ACGCCGAT
24	ATACAAAT	+	+	ATACAAAT
25	ACAAGAAC	+	+	ACAAGAAC
26	ATTGGAAT	+	+	GTATGTAC
27	GTACAGAG	+	+	GTCGAAC
28	ATATGCGT	+	+	ATATGCGT
29	ATAGTGTT	+	+	ATAGTGTT
30	TTGAAGAA	+	+	TTGAAGAA
31	ATCTAAAT	+	37 bp del.	GTGGTAAA
32	TTTCTTAA	+	+	TTTCTTAA
33	ATTGGAAT	+	+	GTATGTAC
34	CAACTCAA	+	+	CAACTCAA
35	ATATGGTT	+	+	ATATGGTT
36	ATATTTCT	+	+	ATATTTCT
37	ATCCCAAC	+	+	ATTTGGAT
38	ATTTGCAT	+	+	ATTTGCAT
39	GGCTTTAT	+	+	GGCTTTAT
40	GCTTTATT	+	19 bp del.	GCTTTATT
41	GTTCACAA	+	+	GTTCACAA

42	CTCTGTAC	+	+	CTCTGTAC
43	TAATGAAA	+	+	TAATGAAA
44	GTTGGAAT	+	+ 443 bp 3'	GTTGGAAT
45	ATTCTAAC	+	+	AATCTAAC
46	TTTTCTTA	+	+	TTTTCTTA
47	GTTCACAA	+	+	GTTCACAA
48	ATTACTAC	+	+	GTTTATAT
49	CTTTCAAG	+	18 bp del.	CTTTCAAG
50	TTCAAGAC	+	+	GTACATAC
51	TTGTATAT	-	+	TTGTATAT
52	GGCTAGAG	+	+	GGCTAGAG
53	GTATGCAC	+	+ 653 bp 3'	ACCCCTCA
54	ATTGGAAT	+	+	GTATGTAC
55	-	-	25 bp del.	GTGCATAC
56	ATTGGAAT	+	+	GTATGTAC
57	ATTGGAAT	+	+	GTATGTAC
58	ACCTGCAG	+	+	GTCTATCC
59	ATAGAGGT	+	+	ATAGAGGT
60	ATTGGAAT	+	+	GTATGTAC
61	GTTCAGAC	+	+	GTTCAGAC
62	GTTGCTGT	+	+	GTTGCTGT
63	GTTGCTGT	+	+	GTTGCTGT

(+) and (-) denote the presence of absence of the Hermes TIRs in the final transposition product.

Table S6 - Sequences of Substrates Used for Assays, Related to Supplemental Procedures

Oligo name	Sequence 5'-3' (only the top strand is listed when the oligo is blunt-ended; in these cases, the bottom strand is the reverse-complement of the top strand)	
LE17+1F	TCAGAGAACAACAACAAG GTCTCTTGTTGTTGTTC	Fig. 1D, S7
LE16-1T	AGAGAACAACAACAA GTCTCTTGTTGTTGTT	Fig. 1D
LE16-1Tran	GTCAGTCGGTACGGA CCAGTCAGCCATGCCT	Fig. 1D
pre-cleaved LE60+11F, top strand	GCGCCCGCTGA/TCAGAGAACAACAACAAGTGGCTTA TTTTGATACTTATGCGCCACTTGCTACTTATGAGTA	Fig. 3D
pre-cleaved LE60+11F, bottom strand	TACTCATAAGTAGCAAGTGGCGCATAAGTATCAAAAT AAGCCACTTGTTGTTGTTCTCTGATCAGCGGGCG	Fig. 3D
intact LE60+11F	CGCCCGCTGATCAGAGAACAACAACAAGTGGCTTATT TTGATACTTATGCGCCACTTGCTACTTATGAGTA	Fig. 3D
LE30	CAGAGAACAACAACAAGTGGCTTATTTTGA	Fig. 4D; 6D
RE30	CAGAGAACTTCAACAAGCCACAGGCAAACG	Fig. 6D
* LE_1mut in pHL2577$_{GDVLE}$	GTCTGAACCAGAGAACAACAAtctcgtaacgATT	Fig. 6B
* RE_1mut in pHL2577$_{GDVLE}$	CAGAGAACTTCAACcgttacgctaCAAACG	Fig. 6B
* LE -1 in pHL2577$_{GDVLE}$	GTCTGAACCAGAGAACAACACAAGTGGCTTaATTTT	Fig. 6C
* LE +1 in pHL2577 $_{GDVLE}$	GTCTGAACCAGAGAACAACAAaCAAGTGGCTTTTTT	Fig. 6C
* LE +3 in pHL2577 $_{GDVLE}$	GTCTGAACCAGAGAACAACAAtctCAAGTGGCTTTT	Fig. 6C
LE +5 in pHL2577 $_{GDVLE}$	GTCTGAACCAGAGAACAACAAtctcgCAAGTGGCTT	Fig. 6C
LE30_LE_1mut	CAGAGAACAACAAgttcaacgagATTTTGA	Fig. 6D
RE30_RE_1mut	CAGAGAACTTCAACcgttacgctaCAAACG	Fig. 6D
LE17	CAGAGAACAACAACAAG	Fig. S7
LE17+1F+P_i	TCAGAGAACAACAACAAG	Fig. S7
LE18	CAGAGAACAACAACAAGT	Fig. S7
LE19	CAGAGAACAACAACAAGTG	Fig. S7
LE20	CAGAGAACAACAACAAGTGG	Fig. S7
LE21	CAGAGAACAACAACAAGTGGC	Fig. S7
LE22	CAGAGAACAACAACAAGTGGCT	Fig. S7
LE26	CAGAGAACAACAACAAGTGGCTTATT	Fig. S7

LE30+1F+P_i	TCAGAGAACAACAACAAGTGGCTTATTTTGA GTCTCTTGTTGTTGTTCACCGAATAAAACT	Fig. S7

* beyond the listed regions, the plasmid sequence are those of pHL2577
/ indicates a gap; "+P_i" indicates a 5'-phosphate on the NTS
Flanking sequence is highlighted in red.

Supplemental Experimental Procedures

The sequences of substrates used for various biochemical assays can be found in Table S6.

Protein expression and purification

E. coli expression for crystallization and biochemical assays

Protein was expressed in Top10 cells (Invitrogen) by growth at 37°C until OD600nm ~0.6, followed by cooling to 19°C and then induction by addition of arabinose to a final concentration of 0.008%; cells were harvested after 16-18 hr. Cells were lyzed by sonication in lysis buffer (25 mM Tris pH 7.5, 0.5 M NaCl, 0.2 mM TCEP), centrifuged to remove cell debris, and the soluble material loaded onto Heparin Sepharose (GE Healthcare) previously equilibrated in 25 mM Tris pH 7.5, 0.1 M NaCl, 0.2 mM TCEP. After washing with the same buffer containing 0.5 M NaCl, protein was eluted using a linear gradient from 0.5 M to 1.0 M NaCl. For size exclusion, fractions containing Hermes were combined, concentrated, and loaded onto a preparative scale BioSep-SEC-S 3000 column (Phenomenex) equilibrated in 25 mM HEPES pH 7.3, 1.5 M NaCl, and 0.2 mM TCEP.

E. coli expression of W319 and W182 mutants for biochemical assays

All *Hermes* transposase mutants and wild-type protein for activity comparison were cloned into pBAD/Myc-HisB-Hermes and purified from Top10 cells as previously described (Zhou et al., 2004) with the following modifications. 0.5% of overnight cultured cells were inoculated, and grown at 37°C to OD600nm 0.6-0.8. The culture was then shifted to16°C and induced with 0.2% L-arabinose for 16-20 hr. After induction, cells were harvested by centrifugation at 4°C and washed with TSG (20 mM Tris-HCl, pH 7.8, 500 mM NaCl, 10% v/v glycerol), and frozen on dry ice; all subsequent steps were performed at 4°C as described previously.

S. cerevisiae expression

S. cerevisiae yeast strain BY4727 (*MAT α-his3Δ200leu2Δ0met150trp1Δ63ura3Δ0*) (Brachmann et al., 1998) was transformed with the TRP+pGALS *Hermes* transposase helper plasmid and the HIS+ URA3::actin intron::*Hermes*-KanMx excision donor plasmid (Gietz and Schiestl, 2007), and selection for transformants was performed on SC-HIS-TRP plates containing 2% glucose. A single colony from the transformation plate was streaked onto a SC-HIS-TRP 2% glucose master

plate, and a single colony from that was then used to inoculate a 5 ml starter culture of SC-HIS-TRP 2% glucose and allowed to grow overnight at 30°C. The starter culture was used to inoculate 250 ml of SC-HIS-TRP 2% galactose at OD600nm 0.05 after washing the cells with water. Cells were allowed to grow for 20 hr under galactose induction at 30°C with shaking at 250 rpm.

Sf9 expression

Sf9 cells were infected at MOI 3 with recombinant baculovirus stocks derived from a bacmid encoding N-terminally His-tagged Hermes with an intervening thrombin cleavage site. Cells were incubated at 27°C after infection and harvested 48 hr later. After washing the pellets with PBS, cells were resuspended in 20 mM sodium phosphate pH7.5, 0.5 M NaCl, 5 mM $MgCl_2$, and 5% glycerol and Complete Protease Inhibitor (Roche), sonicated briefly, and centrifuged to collect soluble protein. The soluble material was diluted 2X and loaded onto a Ni-Chelating Sepharose column equilibrated in the same buffer. After elution using an imidazole gradient to 0.4 M, fractions containing Hermes were dialyzed against 20 mM Tris pH 7.5, 0.5 M NaCl, 5 mM DTT, 1 mM EDTA in the presence of thrombin (Sigma), which was added repeatedly over 48 hrs until SDS-PAGE confirmed that the His-tag had been completely cleaved. The protein was then passed over benzaminide-Sepharose beads to remove thrombin prior to storage at -80°C.

DEV8 cell line expression for EMSA studies

The *Hermes* ORF with the stop codon and 3'UTR were amplified by PCR from pCRSHermes with the primers HermesEcoRV-ATG (5'ccggatatcatgcagaaaatggacaattt-gaagtgaaagc) and HermesEnd-XhoI (5'ttactcgagattgaacatgagggttgtgtgtgtgtgctattg) cut with *EcoRV* and *XhoI*, and cloned into the plasmid pMT-V5 HisC (Invitrogen) cut with the same enzymes. pEGFP1 (Clontech) was cut with *EcoRI* and *NotI* and cloned into pMT-V5 HisC (Michel and Atkinson, 2003). These plasmids were co-transfected with pCoHYGRO (Invitrogen) into Drosophila S2 cells using CellFECTIN reagent (Invitrogen) to produce the cell lines DEV8 (*Hermes* transposase) and EGFP, and transfected cells were selected by addition of hygromycin to the media. Cells were grown at 25ºC in Schneider's *Drosophila* Medium (Invitrogen) supplemented with fetal bovine serum and antibiotics.

To prepare nuclear extracts, cells were induced with 250 μM $CuSO_4$ in 30 ml cultures, harvested, and nuclei extracted following a modified protocol from Miura et al. (1999). Briefly, cells were spun down and washed once with cold PBS. PBS was removed and the cells were lysed with 30 ml Buffer A (10 mM HEPES pH 7.9, 150 mM NaCl, 1 mM EDTA, 4 mM DTT, 0.6% Triton X-100, and 0.5% protease inhibitor cocktail (Sigma)). After 5 min on ice, the samples were centrifuged for 10 minutes at 1500 x g and the supernatant was removed. The pelleted nuclei were lysed with 300 μl NE buffer (20 mM HEPES (pH 7.9), 420 mM NaCl, 0.2 mM EDTA, 1.2 mM $MgCl_2$, 1 mM DTT, 25% glycerol, and 1% protease inhibitor cocktail) and incubated on ice for 20 min. The samples were microcentrifuged at 4ºC at 16000 x g for 5 min. The supernatant was aliquoted and frozen in liquid nitrogen and stored at -80ºC.

Crystal Growth

Typically, 3 μl of the Hermes 79-612/L16-1T complex at ~8 mg/ml protein was mixed with 1 μl well solution consisting of 0.18 M KCl, 5 mM DTT, 0.8 M K^+, Na^+ tartrate, 0.1 M MES pH 6.5, and 5% (v/v) *t*-butanol. Crystals grew over several days to a maximum size of 0.8 x 0.8 mm, and were cryoprotected by transfer over 2-3 days into a stabilizing buffer containing 0.35 M K^+, Na^+ tartrate and 1.7 M sodium malonate, and frozen by immersion into liquid propane.

Data Collection and Processing

Native data were collected on two crystals maintained at 95K using a Rigaku 007-HF rotating anode X-ray source equipped with multilayer focusing mirrors, and reflection intensities were measured with a Raxis IV+ imaging plate detector. Derivative data were collected at the Advanced Photon Source, beamline 22-ID, operated by SER-CAT using a MAR300 CCD detector at 100K. Diffraction data were integrated and scaled internally with XDS (Kabsch, 2010), and zero dose (Diederichs et al., 2003) corrected intensities were estimated with XSCALE (Kabsch, 2010).

Structure Determination and Refinement

The structure was solved with SIRAS using a Ta_6Br_{12} derivative. A native data set at 8.04 keV (1.5418 Å) and a derivative set collected at 10.3 keV (1.2037 Å) were used. F_a coefficients were

estimated with Xprep (Bruker Corporation) and were used in SHELXD (Sheldrick, 2008) that found eight Ta_6Br_{12} sites with data extending to 5.0Å. Ta_6Br_{12} clusters were treated as spherically symmetrical scattering objects and their positional and occupancy parameters were refined in SHARP 2.6 (Vonrhein et al., 2007). The resulting map was solvent flattened with Solomon (Abrahams and Leslie, 1996) using a 70% solvent fraction. Coordinates from the apo-Hermes monomers (PDB ID 2BW3) were roughly positioned in the resulting map to generate masks for non-crystallographic symmetry (NCS) averaging using NCSMASK (CCP4, 1994). With two dimers in the asymmetric unit (ASU), four-fold NCS averaging, further solvent flattening, histogram matching and phase extension to 3.3Å were performed in 250 cycles with DM (Cowtan and Zhang, 1999). The final map was of excellent quality (Figure S1A) allowing all nucleic acid residues to be placed and the protein model to be rebuilt manually using the program O (Jones and Kjeldgaard, 1997). The model was refined with several cycles of NCS restrained cartesian simulated annealing, energy minimization and tightly restrained individual B factor refinement as implemented in CNS 1.3 (Brünger et al., 1998) at 3.4Å resolution. The model was verified and, where needed, adjusted using simulated annealed omit maps (Figure S1B). Detailed crystallographic statistics are in Table S1.

Size Exclusion Chromatography Coupled to Light Scattering (SEC-MALS)

Refractive index and multiangle light scattering of full-length Hermes at 1.1 mg/ml injected onto a TSKgel SuperSW3000 gel filtration column (Tosoh) were measured in-line using DAWN HELIOS and Optilab rEX instruments from Wyatt Technology. The running buffer was 20 mM HEPES pH 7.5, 0.35 M NaCl, and 0.2 mM TCEP. The molar mass was obtained using Astra V software (Wyatt).

SAXS Data Collection and Analysis.

We used small angle X-ray scattering (SAXS) to evaluate the oligomerization state of full-length Hermes in solution (Hura et al., 2009; Rambo and Tainer, 2013a&b; Trewhella et al., 2013). SAXS data were acquired at the Beam Line 12-IDC at the Advanced Photon Source, Argonne National Laboratory. Data collection was carried out with a mosaic Gold CCD detector positioned in an on-center geometry, 3.7m from the sample capillary using 18 keV incident radiation, resulting in the observable q-range of 0.007 – 0.240 $Å^{-1}$. In order to evaluate possible

concentration dependence of the scattering data, recordings were done at protein concentrations of 1.50, 0.75, and 0.38 mg/mL.

Q-axis mapping was done using a silver behenate standard sample. Totals of 20 sequential data frames with exposure times of 1 sec were recorded with the samples kept at 25°C throughout the measurement. In order to prevent radiation damage, volumes of 120 μl of samples and buffers were oscillating during data collection. Individual data frames were masked, corrected for the detector sensitivity, radially integrated and normalized by the corresponding incident beam intensities and sample transmissions. The final 1D scattering profiles and their uncertainties were calculated as means and mean uncertainties over the 20 individual frames. The buffers data were then subtracted from the samples using solvent volume fractions calculated based on protein concentrations. Guinier analyses and regularized fits of the P(r) functions were carried out using PRIMUS (Konarev et al., 2003) and GNOM (Svergun, 1992) software packages. Recorded scattering data exhibit linear dependence of the zero-angle scattering intensity and independence of the radius of gyration (75±2Å) with concentration. P(r) profiles fitted via regularized Guinier transform exhibit Rg values (74±1Å) consistent with those derived via Guinier analysis and show no signs of aggregation. Taken together, these facts support interpretation of the experimental scattering data as purely a form factor of a monodisperse scattering particle.

A molecular mass of 540±80kDa for the scattering particle, consistent with the octameric arrangement, was calculated from Guinier-fitted I(0) values by reference to the scattering data from standard samples of lysozyme and bovine serum albumin. The analysis based on the power-law relationship between the ratio of the square of the correlated volume and the gyration radius, and the mass of the particle (Rambo and Tainer, 2013a) yielded the molecular mass of 580kDa. A maximum protein dimension of ~220Å determined via regularized Fourier transform fits is also consistent with the crystal structure of the octameric Hermes particle. Fit of the X-ray scattering data to the coordinates of octameric full-length Hermes was performed using the AXES approach (Grishaev et al., 2010). Placement of the first 76 N-terminal residues absent in the crystal structure of the octamer, was optimized by best-fitting against the experimental scattering data with the quality of the best-fitting conformation characterized by χ_{free} = 1.3 (Rambo and Tainer, 2013a). For this purpose, a Monte Carlo rigid body translational/rotational optimization was performed using coordinates of the N-terminal residues 1-76 from PDB

deposition 2DJR avoiding conformations exhibiting heavy-atoms contact distances below 2.5Å. A comparison of the distance probability profiles corresponding to the experimental and best-fitted model data is shown in Figure S3A.

Negatively Stained Electron Microscopy

Hermes at 1 mg/ml was diluted 20-fold in 1X PBS and mixed 1:1 v/v with a 1.2% glutaraldehyde solution, and incubated on ice for 10-20 min. A drop was then incubated for 1-2 min on a glow discharged thin-carbon coated copper grid, blotted and stained with a 2% uranyl acetate for 30 sec. Imaging was done using an FEI Tecnai 12 electron microscope equipped with a GIF 2002 energy filter and a 2Kx2K CCD camera (both from Gatan, Inc.). The resulting images used for image processing had a calibrated sampling rate of 3.785 Å/pixel. The EMAN2 software package (Tang et al., 2007) was used to pick particles, generate the reference-free 2D class averages, and calculate the 3D reconstruction, for which the crystal structure was used as initial model, after removing the DNA strands and converting it to a 30 Å map using Bsoft (Heymann and Belnap, 2007). The resolution of the reconstructed map was estimated to be 22 Å and the map was lowpass-filtered to this resolution (see Figure S3D). Fits of the crystal structure into the EM map are shown in Figure S3E, and the fit in which each dimer was treated as an independent rigid body, and 8 copies of the homology-modeled BED domains were included as occupying the central density present in the EM reconstruction is shown in Figure S3F.

Size exclusion chromatography to assess Hermes oligomeric State (Fig. S2) and DNA binding (Fig. 1E)

To assess the multimeric state of untagged full-length Hermes expressed in *E. coli*, purified Hermes at ~2 mg/ml was injected onto two BioSep-SEC-S 300 columns (300 x 21.2 mm; Phenomenex) connected in series which had been previously equilibrated in 25 mM HEPES pH 7.3, 0.75M NaCl and 0.3 mM TCEP. The flow rate was 1 ml/min.

To assess full-length Hermes expressed in *S. cerevisiae*, cells were harvested by centrifugation at 4,000 x g for 20 min, the supernatant was discarded and the cell pellet washed with 50 ml water. Two ml extraction buffer (25 mM Tris pH 8., 1.5 M NaCl, 5 mM DTT, 1mM phenylmethylsulphonly fluoride (PMSF)) per gram of wet weight pellet were added, the pellet was resuspended by gentle pipetting, and the suspension was transferred to a 50 ml Falcon tube;

500 ul glass beads/gram pellet wet weight were added, and the tube was vortexed for 1 min. Vortexing was repeated 10 times each, placing the tube on ice for 1 min between vortexing steps. The sample was then centrifuged at 4,000 x g for 30 min in a cooled centrifuge to remove unbroken cells. The resulting soluble extract was then filtered and 0.2 ml of the protein extract was injected onto a calibrated Superose 6 gel filtration column previously equilibrated in 25 mM Tris pH 8, 1.5 M NaCl, and 1 mM PMSF. Protein fractions (50 µl) were collected and screened for Hermes by Western Blot using Hermes polyclonal antibodies.

To assess the multimeric state of full-length Hermes from the baculovirus/Sf9 insect cell expression system, 50 µl purified Hermes from which the N-terminal His-tagged had been removed was injected onto a Superose 6 PC 3.2/30 column equilibrated in 20 mM HEPES pH 7.5, 0.5 M NaCl, 1 mM EDTA, 5 mM DTT, and 10% (w/v) glycerol (operated on a GE Healthcare SmartSystem at 4°C at a flow rate of 50 µl/min). The gel filtration standards thyroglobulin (669 kDa) and apoferritin (443 kDa) were subsequently injected in the same buffer.

To assess DNA binding, purified untagged E. coli-expressed full-length Hermes at 1 mg/ml was mixed with LE30 at the indicated ratios of protein to DNA and dialyzed overnight against 0.18M KCl, 20 mM HEPES pH 7.5, 2 mM DTT at 4°C. The mixture was then injected onto a Superose 6 column as above.

In vitro biochemical activity assays

Strand transfer assay with LE oligonucleotide used for structural studies (Fig. 1D), LE30 (Fig. 4D), and LE_1 and RE_1 mutated ends (Fig. 6D)

Purified Hermes, oligonucleotides, and 4 nM pUC19 were mixed in a final volume of 100 µl in reaction buffer containing 25 mM Tris pH 7.5, 10 mM DTT, 5 mM $MgCl_2$, 0.1 mg/ml BSA, and either NaCl or KCl at the concentrations, times, and temperatures indicated in the figure legends. After incubation, the reaction was stopped by addition of EDTA to 25 mM, and treated with Proteinase K for 30 min at 37°C. After addition of glycogen, the reaction products were ethanol-precipitated and then resuspended in 20 µl of high glycerol loading dye (25% glycerol and 1.25% 5X DNA gel loading solution). Samples were run on a 1.5% agarose gel for 75 min at 100 V, and the gel was stained with ethidium bromide for visualization.

Hairpin formation using pre-nicked Hermes-LE oligonucleotides (Fig. 3D)

An 11 bp oligonucleotide corresponding to the top flanking donor site DNA was radiolabelled at its 5' end with γ-^{32}P-ATP. The end-labelled oligonucleotide was mixed with equimolar amounts of a 61 nt oligonucleotide corresponding to 1 nt of the top flanking DNA and 60 nt of the top strand of *Hermes*-LE, and a 71 nt oligonucleotide containing 11 nt of bottom-strand flanking DNA + 60 nt of bottom strand of *Hermes*-LE. The mixture of oligos was annealed and then used as substrate in hairpin formation reactions as described in the next section.

Cleavage reactions with LE60+11F (Fig. 3D)

Cleavage reactions were performed with a 71 bp Hermes-LE fragment containing 60 bp of Hermes-LE + 11 bp flanking DNA. The DNA was radiolabelled at the 5' end of both strands with γ-^{32}P-ATP. 3.5 pmole of Hermes transposase was incubated with 0.5 pmole of radiolabelled Hermes-LE DNA in 25 mM MOPS, pH 7.6, 2% (v/v) glycerol, 100 mM NaCl, 2 mM DTT, 1 mM $MgCl_2$, 0.01% BSA in a final volume of 10 μl at 30°C for the indicated times. Reactions were stopped by adding EDTA to 40 mM and incubated at 37°C for 2 hr, then spin-dried. The samples were resuspended in STOP solution (Affymetrix): 8M Urea at a ratio of 1.8:1. The products were run on 16% denaturing acrylamide gels.

Plasmid cleavage assay using pRX1-Her (Fig. 4C)

Plasmid cleavage was carried out in a reaction mixture containing 25 mM HEPES pH 7.5, 5 mM $MgCl_2$, 10 mM DTT, 1 μg BSA and plasmid containing the Hermes transposon (LE30/RE30 flanking the KanMx cassette). The NaCl final concentration in reaction ranged from 0 mM-300 mM, and reactions were started by adding 60 ng of both full-length Hermes wildtype and deletion dimer proteins. Reactions were then incubated for 60 min at 30°C and subsequently stopped by addition of SDS/EDTA solution in final concentration of 0.5% SDS and 25 mM EDTA and further incubated for 10 min at 65°C. Reactions were placed on ice briefly before addition of Proteinase K in final concentration of 200 μg/ml and further incubated for 60 min at 37°C. Reactions were then further treated with phenol/chloroform/ isoamyl alcohol and DNA was precipitated and concentrated for restriction digestion with *Pml*I for 60 min at 37°C. Restriction digests were run on an 0.7% TBE-agarose gel in 0.5X TBE buffer for overnight at 50

V at room temperature and subsequently stained with ethidium bromide.

Plasmid cleavage assay using pHL2577$_{GDVLE}$ (Fig. 6B, 6C)

In pHL2577, *Hermes* is flanked by an EagI site on its LE and an XhoI site on the RE (Evertts et al., 2007). Although the latter possesses the important -2 T/A bp of the target site preference of Hermes (Gangadharan et al., 2010), the EagI site does not. The LE flank was therefore mutated by the QuikChange method to 5'-GTCTGAAC, the preferred hotspot in pGDV1; this variant of pHL2577 is designated pHL2577$_{GDVLE}$.

Purified Hermes and pHL2577$_{GDVLE}$ were mixed in a final volume of 100 µl in reaction buffer containing 25 mM Tris pH 7.5, 200 mM KCl, 10 mM DTT, 5 mM $MgCl_2$, 0.1 mg/ml BSA for 1 hr at 30°C. After incubation, the reaction was stopped by addition of EDTA to 25 mM, and subsequent Proteinase K treatment for 30 min at 37°C. After addition of glycogen, the reaction products were ethanol-precipitated and air-dried. The reaction products were digested with *Xmn*I for 2.5 hr at 37°C and then run on a 1% agarose gel for 90 min at 100 V, and the gel was stained with ethidium bromide for visualization.

EMSA with Hermes nuclear extract from Drosophila S2 cell line (Figure 6E and 6F)

Probes were created by end-labeling 1 pmole of dsDNA with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P] ATP (Amersham Biosciences) and purified on Micro Bio-Spin® P-30 Tris Chromatography Column (Bio-Rad). 20 fmol of labeled probe was used in a 10 µl reaction which included 15 mM HEPES, 2 mM DTT, 2 µg BSA, 0.5 µg poly(dI-dC)·poly(dI-dC), 0.4 µg T3 single-stranded oligo, 2 ul of DEV8 nuclear extract and competitor DNA at 20X and 200X concentrations as indicated. Competitors and NE were incubated together for 15 min at room temperature, then labeled probe was added and incubated for an additional 20 min. A 30 bp double-stranded oligo from the E1 flank of *Hermes* (E1) was used as a non-specific competitor. Samples were run on a 5% TBE polyacrylamide gel at 100 V at 4ºC. The gels were dried and developed with a phosphorimager. The fraction was measured as the total signal for binding in the competition lanes divided by the total signal for binding without added competitor.

In vitro transposition of Hermes transposon into pGDV (Table S5)

Assays (20µl) contained 20 mM HEPES pH 7.9, 5% glycerol, 5 mM $MgCl_2$, 2 mM DTT, 0.2 µg/ul BSA, 5 µl nuclear extract from DEV8 or EGFP cell lines, and 250 ng each of pHDG1 donor and pGDV1 target. Reactions were incubated at 30°C for 1-2 hr, and then 80 µl of stop solution containing 50 mM Tris pH 7.5, 0.5 µg/µl proteinase K, 10 mM EDTA, and 0.25 µg/µl tRNA was added and incubated 1 hr at 37°C. Reactions were extracted with an equal volume of phenol: chloroform: isoamyl alcohol (50:48:2 by volume) and then with chloroform: isoamyl alcohol. Nucleic acids were precipitated by the addition of 10 µl 3 M sodium acetate (pH 5.2) and 200 µl 100% ethanol, incubated at -80°C for 20 min, and spun for 10 min. Reactions were resuspended in water and electroporated into bacteria and plated as for *in vivo* assays.

In vivo assays to assess target site cleft mutants

Somatic cells (Table S2)

A five plasmid interplasmid transposition assay was performed in *Drosophila* embryos using the 37pKH70new helper plasmid into which each of the mutants was cloned. The pKH70new helper plasmid is a modified version of pkhsp70 (Arensburger et al., 2005). Plasmids encoding wild-type (WT) Hermes and each of the mutant transposases were co-injected with the pHDG1 Hermes donor plasmid containing a gentamicin (Gent) selectable marker and LacZ gene flanked by the Hermes transposon TIRs and the pGDV1 target plasmid carrying chloramphenicol resistance. These were injected with the *piggyBac* (PB) donor plasmid carrying the Gent selectable marker and EGFP gene flanked by the PB transposon TIRs and a PB helper plasmid encoding the PB transposase as an internal control. Each set of injections was performed three times into 0-2 hr old Drosophila embryos. Following heat shock to activate expression of the transposase genes and 24 hr incubation, target plasmids were recovered. Plasmids conferring events were sequenced to determine the location and target site into which the donor transposons were inserted.

Germline cells (Table S3)

To assess the transposition dependent germline transformation frequency of the mutants in *Drosophila*, each of the ROW pKH70new helper plasmids and the w+ gene donor plasmid were injected into pre-blastoderm CSW embryos. CSW are a strain of *Drosophila* that does not carry the w+ gene and therefore have a white-eye

phenotype. After a 16 hr recovery period the embryos were heat shocked at 37°C for 1 hr and then collected 24 hr later. After eclosion each of the G0 flies were crossed with three non-transformed CSW virgin flies. Transformation was detected by the presence of red, orange, or yellow eye pigment in the G1 progeny. The G0 from which the red, orange or yellow G1 progeny came from was considered transgenic.

In vivo assays to compare Hermes WT, HermesTM and HermesΔ497-516 dimers

Excision assay in S2R+ Drosophila cells (Fig. S6A)

Expression plasmids for Hermes WT, deletion and triple mutants with a N-terminal Flag-tag were constructed in pKH70. pHDG1 was used as donor plasmid. 0.5 μg of transposase plasmid and 2.0 μg of donor plasmid were co-transfected with Effectene transfection Reagent (Qiagen). Three days after transfection, cells were harvested and plasmids DNA (transposase plasmids, both intact and excised + repaired donor plasmids) were recovered by QIAprep Spin Mini-prep Kit (Qiagen). Nested PCR was performed against pHDG1 Hermes transposon flanking donor sides with 1st PCR primer set (shown in black; He-pHDG-L1, CATGATTACGCCAAGCGCGC; He-pHDG-R1, CGCGTAACCACCACACCCG) and 2nd nested PCR primer set (shown in blue; He-pHDG-L2, GCGCGCAATTAACCCTCAC; He-pHDG-R2, GTCCCATTCGCCATTCAGG).

Yeast Excision Assay (Fig. S6B)

Excision assays were performed in BY4727 (Brachmann et al., 1998) transformed with the TRP+ pGALSHermes full-length wildtype, HermesTM and HermesΔ497-516 transposase helper plasmids and the HIS+ *URA3*::actin intron::*Hermes*(L711/R520) KanMx excision donor plasmid (Gietz and Schiestl, 2007). Excision resulted in a change from URA$^-$ auxotrophy to URA$^+$ prototrophy. Cells containing both plasmids were streaked on SC-HIS-TRP plates containing 2% glucose, and single colonies were resuspended, diluted, and grown on SC-HIS-TRP plates containing 2% galacotose for 5 days. Colonies were resuspended in water, diluted, and plated on SC-HIS-TRP-URA containing 2% glucose to measure excision events and SC containing 2% glucose to count all cells. The excision frequency is the number of Ura$^+$ cells/total cells.

Expression and time course detection of Hermes in yeast (Fig. S6C)

S. cerevisiae yeast strain BY4727 was transformed with the TRP+pGALS plasmids containing wild-type Hermes, HermesTM and HermesΔ497-516 transposase genes, and selection for transformants was performed on SC-TRP plates containing 2% glucose. A single colony from the transformation plate was streaked on a SC-TRP 2% glucose master plate. A single colony was then picked from the master plate and used to inoculate a 5 ml starter culture of SC-TRP 2% glucose and allowed to grow overnight at 30°C. The starter culture was used to inoculate 25 ml of SC-TRP 2% raffinose and 0.1% glucose at OD600nm 0.05 after washing cells with water. 2% galactose final concentration was added when cells reached OD600nm ~ 0.25 and transposase was induced at 30°C with shaking at 250 rpm. Yeast whole protein extract was prepared from cells induced at 1, 2, 3, and 24 hr by treating cells with 0.2 M NaOH followed by boiling in SDS-PAGE sample buffer (Kushnirov, 2000). Transposase expression was detected by Western blotting with mouse monoclonal antibody against the HA-tag that was N-terminally fused to Hermes.

Cellular localization in HEK293 cells (Fig. S6D)

Hermes WT, HermesΔ497-516 and HermesTM with a N-terminal Flag-tag were constructed in pcDNA3.1/myc-HisA. To isolate both cytoplasmic and nuclear proteins from HEK293 cells, a protocol from GeneTex, Inc for isolating both cytoplasmic/membrane and nuclear proteins was followed. Briefly, cells were harvested 3 days after transient transfection, washed with ice cold PBS-1 mM EDTA and resuspended in cold Harvest buffer (10 mM HEPES pH 7.9, 50 mM NaCI, 0.5 M sucrose, 0.1 mM EDTA, 0.5% Triton X-100, 1 mM DTT, with protease inhibitor cocktail, Sigma P-1860). After incubation on ice for 5 min, the nuclei were collected by centrifugation at 1,000 rpm for 10 min. The supernatant which contains the cytoplasmic and membrane proteins was then spun at 14,000 rpm for 15 min. The nuclei pellet was washed and resuspended in Buffer A (10 mM HEPES pH 7.9, 10 mM KCI, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, with protease inhibitor cocktail), spun at 1000 rpm, and 4 volumes of Buffer C (10 mM HEPES, pH 7.9, 500 mM NaCI, 0.1 mM EDTA, 0.1 mM EGTA, 0.1% NP-40, 1mM DTT, with protease inhibitor cocktail) was added to the pellet. This was vortexed for 15 min at 4°C, and spun 14,000 rpm for 10 min at 4°C to collect the supernatant that contains the nuclear extract. To prepare whole cell extracts, cells were resuspended in NETN lysis buffer (150 mM NaCl, 20 mM Tris pH 8.0, 1 mM EDTA, 0.5 % (v/v) Nonidet P-40 with protease inhibitor

cocktail), incubated at 4°C for 1 hr, centrifuged to remove cell debris, and the supernatant subsequently used.

Hermes and mutant proteins from the whole cell lysate, cytoplasmic and nuclear fractions were detected by Western blotting with anti-Flag M5 antibody (Sigma, F4042) after SDS-PAGE and Semi-dry transfer.

In vivo assay to assess transposon end length requirement in vivo (Table S4)

Construction of pHDG1: Short Hermes ends (30 bp LE and RE) along with vector sequences (ampicillin resistance and a replication origin) and homology sequences were prepared by PCR using a 4kb fragment of clone pHDG1 digested with *Xho*I and *Bgl*II as template.

Primer Hermes 1-30 For:

5'-*GACACTATTCAACTA*CGTTTGCCTGTGACTTGTTGAAGT-3'. Underlined sequence has homology to the Hermes RE TIR, while the italicized sequence provides homology to the stuffer fragment.

Primer Hermes 1-30 Rev:

5'-*CATGCCCTTGGCTA*GTCAAAATAAGCCACTTGTTGTTGTTCTCTG-3'. Underlined sequence has homology to the *Hermes* LE TIR, while the italicized sequence provides homology to the pGoE plasmid.

The stuffer fragment was amplified from the piggyBac transposon plasmid pBac3xP3dsRed/AgB2tEGFP, and was included in the construct as "spacer" DNA to make the *Hermes* elements from pHDG7 and pHDG8 the same size. This PCR fragment also has homology to the Hermes PCR fragment on one end and the pGoE plasmid on the other for assembly purposes.

Primer Stuffer Fragment For:

5'-*CTTGCTTTAAAGCTAG*GTCAGTCAGAAACAACTTTGGC-3'. Underlined region has homology to the *piggyBac* element, while the italicized region has homology to the pGoE vector for assembly purposes.

Primer Stuffer Fragment Rev:

5'-*TAGTTGAATAGTGTC*GGTATACTTATTATCATCTTGTGATGAGGA-3'. Underlined region is homologous to the *piggyBac* element, while the italicized region has homology to the Hermes/vector fragment for assembly.

pGoE, which has a gentamycin resistance gene, a replication origin, and EGFP fused in-frame with the lacZ-alpha fragment was digested with *Nhe*I, run on an agarose gel, and purified. The Hermes/vector fragment, the stuffer fragment, and the pGoE vector were assembled using the GeneArt Seamless Cloning and Assembly Kit (Life Technologies), transferred to TOP10 cells, and plated onto LB agar containing ampicillin and gentamycin.

Construction of pHDG7: The *Hermes* clone with long ends (711 bp LE, 520 bp RE) was constructed from multiple fragments. Clone pHDG1 was used as the starting vector, and was digested with *Mfe*I and *Bgl*II, and the 3kb fragment containing the appropriate target site duplications and partial ends of Hermes was purified on an agarose gel. Fragments of the remainder of the longer ends were constructed by amplifying separately the Hermes LE and RE from the *Mfe*I sites to base 711 (left) or 520 (right) using the 1kb fragment of vector pBSHermes digested with *Mfe*I as template.

Primer Hermes LE For:

5'-*TGCTACTTATGAGTA*CAATTGTGCTTTGCCACTTGAAC-3'. Underlined region is homologous to the pBSHermes *Mfe*I fragment, while the italicized portion provides homology to the partial left end of the pHDGI fragment.

Primer Hermes LE Rev:

5'-*CATGCCCTTGGCTAG*CTGAAACAGTTTTTAATTCTCGGGATT-3'. Underlined region is homologous to the left end of pBSHermes, while the italicized portion provides homology to the pGentOriAlpha vector.

Primer Hermes RE For:

5'-*GAATGGCGATAAGCTAG*TCGACAGCTTGTTATTTTTAAATTCC-3'. Underlined region has homology to the *Hermes* right end from pBSHermes, while the italicized region provides homology to the pGentOriAlpha vector.

Primer Hermes RE Rev:

5'-*GTGCGATTTGT*CAATTGGCAAATTATACTCACTTCTTGTTG-3'. Underlined region has homology to the *Hermes* right end *Mfe*I region, while the italicized portion provides homology to the partial right end of the pHDGI fragment.

pGentOriAlpha, containing a gentamycin resistance gene, the LacZ-Alpha fragment, and a replication origin was digested with *Nhe*I and gel purified from an agarose gel. The four fragments pHDGI *Mfe*I, the partial right and left end PRC products, and the pGentOriAlpha

vector were assembled and plated as above.

Construction of pHDG-EGFP: pHDG1 was used as template for *Taq* polymerase to amplify 305 bases of the Hermes LE flanked by *Sac*I and *Xba*I sites. The PCR product was purified, digested with *Sac*I and *Xba*I, and cloned into pBluescipt SK+digested with the same enzymes to give clone pHermesL305. A Hermes RE of 307 bases was amplified by PCR also using pHDG1 as template, digested with *Pst*I and *Eco*RV, and also cloned into pBluescipt SK+digested with the same enzymes. The RE of Hermes was digested with the above enzymes, purified on an agarose gel, and ligated to the left end clone also digested with *Pst*I and *Eco*RV to give clone pHermesL305R307. This clone was digested with *Xba*I and ligated to pGoE vector which had been digested with *Nhe*I to give clone pHDG-EGFP.

HL8bpFSac:

5'-TGAGAGCTCGTCTGTATCAGAGAACAACAACAAGTGGCTTATTTTG. Underlined region is homologous to the Hermes LE.

Hermes L305 Xba Rev:

5'-GATTCTAGACACACTCAAGTGCATAAGCCACTTGTTAGC-3'. Underlined region is homologous to the HermesLE.

Hermes R307 Pst For:

5'-CATCTGCAGCAGAATCATATGCAATACTACAAACAATAGCACACAC-3'. Underlined region is homologous to the Hermes RE.

Hermes R TSD RV Rev:

5'-GATGATATCATACAGACCAGAGAACTTCAACAAGTCACAGGC-3'. Underlined region is homologous to the Hermes RE.

Construction of pHDGLL: A second LE of 305 bases flanked by *Xho*I and *Cla*I sites was amplified by PCR and cloned into the same sites in the pBSHL305G clone to give pHDGLL.

Hermes L2 Xho For:

5'-GATCTCGAGATACAGACCAGAGAACAACAACAAGTGGCTTATTTTGATAC-3'. The underlined region is homologous to the Hermes LE.

Hermes L2 ClaI seq Rev:

5'-
GATGATATCGATGTTTTGGGAAATCATCCACACTCAAGTGCATAAGCCACTTGTTAG
C-3'. The underlined region is homologous to the Hermes LE; the rest is "spacer" DNA.

Assays in Drosophila S2 DEV8 cells: S2 DEV8 cells were seeded in treated 6-well plates at a density of $2x10^6$ cells per well. Cells were transfected the next day following the X-tremeGENE HP DNA Transfection Reagent protocol (Roche). For three plasmid comparison assays, donor plasmids and their corresponding positive control plasmid (1 µg each) were added along with pGDV1 target plasmid (2 µg) to transfection reagent. For two plasmids assays, 2 µg each of donor and target were added. Plasmid Actin5C-EGFP was used as a positive control for transfection. DEV8 cells were induced to express Hermes the day following transfection by the addition of $CuSO_4$ to a concentration 250 µM to each well (excluding transfection control). Cells were harvested two days after transfection, washed with PBS, and frozen for plasmid preparations. Cell pellets were resuspended with 500 µl of grinding buffer (0.5% SDS, 0.08 M NaCl), 54.7 mg/ml sucrose, 0.06 M EDTA, 120 mM Tris pH9.0) before being incubated at 65°C for 30 min. Potassium acetate was added to a final concentration of 1 M prior to incubating on ice for 30 min. Samples were centrifuged, and the DNA in the supernatant was precipitated with ethanol. Reactions were resuspended in water and electroporated into bacteria and plated.

Supplemental References

Abrahams, J.P., and Leslie, A.G. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. *D52*, 30–42.

Arensburger, P., Kim, Y.J., Orsetti, J., Aluvihare, C., O'Brochta, D.A., and Atkinson, P.W. (2005). An active transposable element, *Herves*, from the African malaria mosquito *Anopheles gambiae*. Genetics *169*, 697-708.

Bahler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., Steever, A.B., Wach, A., Philippsen, P., and Pringle, J.R. (1998). Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. Yeast *14*, 943-951.

Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J.C., Hieter, P., and Boeke, J.D. (1998). Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast *14*, 115-132.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J., Kuszewski, J., Nilges, M., Pannu, N.S. et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. *D54*, 905–921.

Collaborative Computational Project Number 4 (1994). The CCP4 suite: Programs for protein crystallography. Acta Crystallogr. *D50*, 760-763.

Cowtan, K.D., and Zhang, K.Y.J. (1999). Density modification for macromolecular phase improvement. Prog. Biophys. Mol. Biol. *72*, 245-270.

Diederichs, K., McSweeney, S., and Ravelli, R.B.G. (2003). Zero-dose extrapolation as part of macromolecular synchrotron data reduction. Acta Crystallogr. *D59*, 903-909.

Evertts, A.G., Plymire, C., Craig, N.L., and Levin, H.L. (2007). The Hermes transposon of *Musca domestica* is an efficient tool for mutagenesis of *Schizosaccharomyces pombe*. Genetics *177*, 2519-2523.

Gietz, R.D., and Schiestl, R.H. (2007). Quick and easy yeast transformation using LiAc/SS carrier DNA/PEG method. Nature Protocols *2*, 35-37.

Heymann, J.B., and Belnap, D.M. (2007). Bsoft: image processing and molecular modeling for electron microscopy. J. Struct. Biol. *157*, 3-18.

Hura, G., Menon, A., Hammel, M., Rambo, R., Poole, F., Tsutakawa, S., Jenney, F., Classen, S., Frankel, K., Hopkins, R., Yang, S-J., Scott, J., Dillard, B., Adams, M., and Tainer, J. (2009). Robust, high-throughput solution structural analyses by small-angle X-ray scattering (SAXS). Nature Methods *6*, 606-614.

Jones, T., and Kjeldgaard, M. (1997). Electron-density map interpretation. Methods Enzymol.,

277, 173–208.

Kabsch,W. (2010). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. *D66*, 133–144.

Konarev, P., Volkov, V., Sokolova, A., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS - a Windows-PC based system for small-angle scattering data analysis. J. Appl. Cryst. *36*, 1277-1282.

Kushnirov, V.V. 2000. Rapid and reliable protein extraction from yeast. Yeast *16*, 857-860.

Li, X., Ewis, H., Hice, R. H., Malani, N., Parker, N., Zhou, L., Feschotte, C., Bushman, F. D., Atkinson, P. W., and Craig, N. L. (2013). A resurrected mammalian *hAT* transposable element and a closely related insect element are highly active in human cell culture. Proc. Natl. Acad. Sci. USA *110*, E478-E487.

Michel, K., and Atkinson, P.W. (2003). Nuclear localization of the Hermes transposase depends on basic amino acid residues at the N-terminus of the protein. J. Cell. Biochem. *89*, 778-790.

Miura, K., Wang, S.F., and Raikhel, A.S. (1999). Two distinct subpopulations of ecdysone receptor complex in the female mosquito during vitellogenesis. Mol. Cell. Endocrinol. *156*, 111-120.

Rambo, R., and Tainer, J. (2013a). Accurate assessment of mass, models and resolution by small-angle scattering. Nature *496*, 477-482.

Rambo, R., and Tainer, J. (2013b). Super-resolution in solution X-ray scattering and its applications to structural systems biology. Annu. Rev. Biophys. *42*, 415-441.

Scheres, S.H., and Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination. Nature Meth. *9*, 853-854.

Sheldrick, G.M. (2008). A short history of SHELX. Acta Crystallogr. *A64*, 112–122.

Svergun, D.I. (1992). Determination of the regularized parameter in indirect-transform methods using perceptual criteria. J. Appl. Cryst. *25*, 495-503.

Tang, G., Peng, L., Baldwin, P.R., Maann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. *157*, 38-46.

Trewhella, J., Hendrickson, W., Kleywegt, G., Sali, A., Sato, M., Schwede, T., Svergun, D., Tainer, J., Westbrook, J., and Berman, H. (2013). Structure *21*, 875-881.

Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007). Automated structure solution with autoSHARP. Methods Mol. Biol., *364*, 215–230.

Yu, X., and Gabriel, A. (1999). Patching broken chromosomes with extranuclear cellular DNA. Mol. Cell *4*, 873-881.