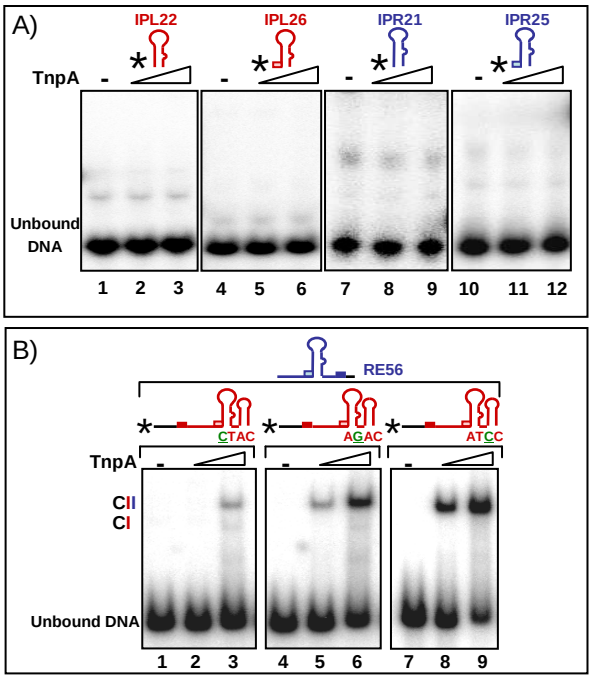
A)
IPL22
IPL26
IPR21
IPR25
TnpA
Unbound DNA
1 2 3 4 5 6 7 8 9 10 11 12
B)
RE56
CTAC
AGAC
ATCC
TnpA
CII
CI
Unbound DNA
1 2 3 4 5 6 7 8 9

Figure S1

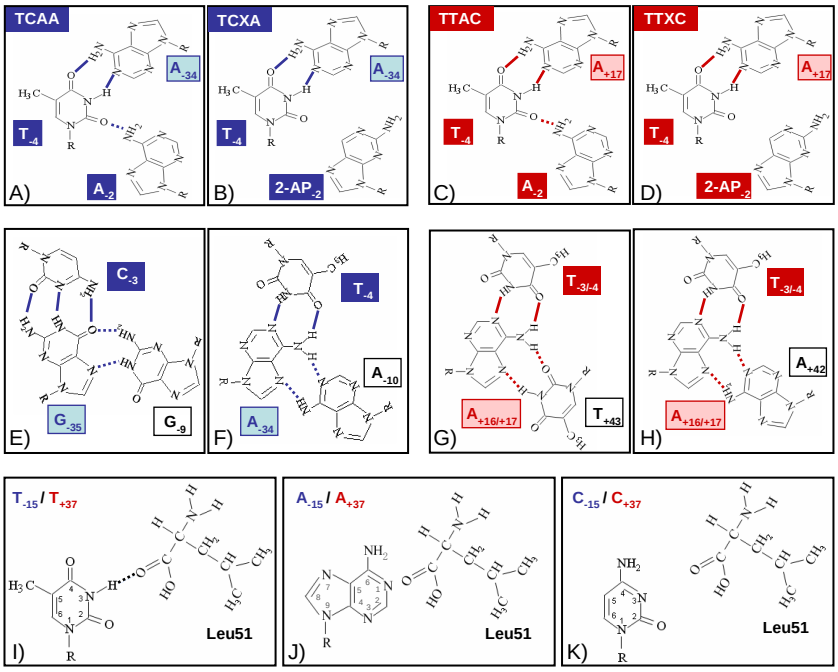
TCAA
A)
T-4
A-34
A-2
TCXA
B)
T-4
A-34
2-AP-2
TTAC
C)
T-4
A+17
A-2
TTXC
D)
T-4
A+17
2-AP-2
E)
C-3
G-35
G-9
F)
T-4
A-10
A-34
G)
T-3/-4
A+16/+17
T+43
H)
T-3/-4
A+42
A+16/+17
I)
T-15 / T+37
Leu51
J)
A-15 / A+37
Leu51
K)
C-15 / C+37
Leu51

Figure S2

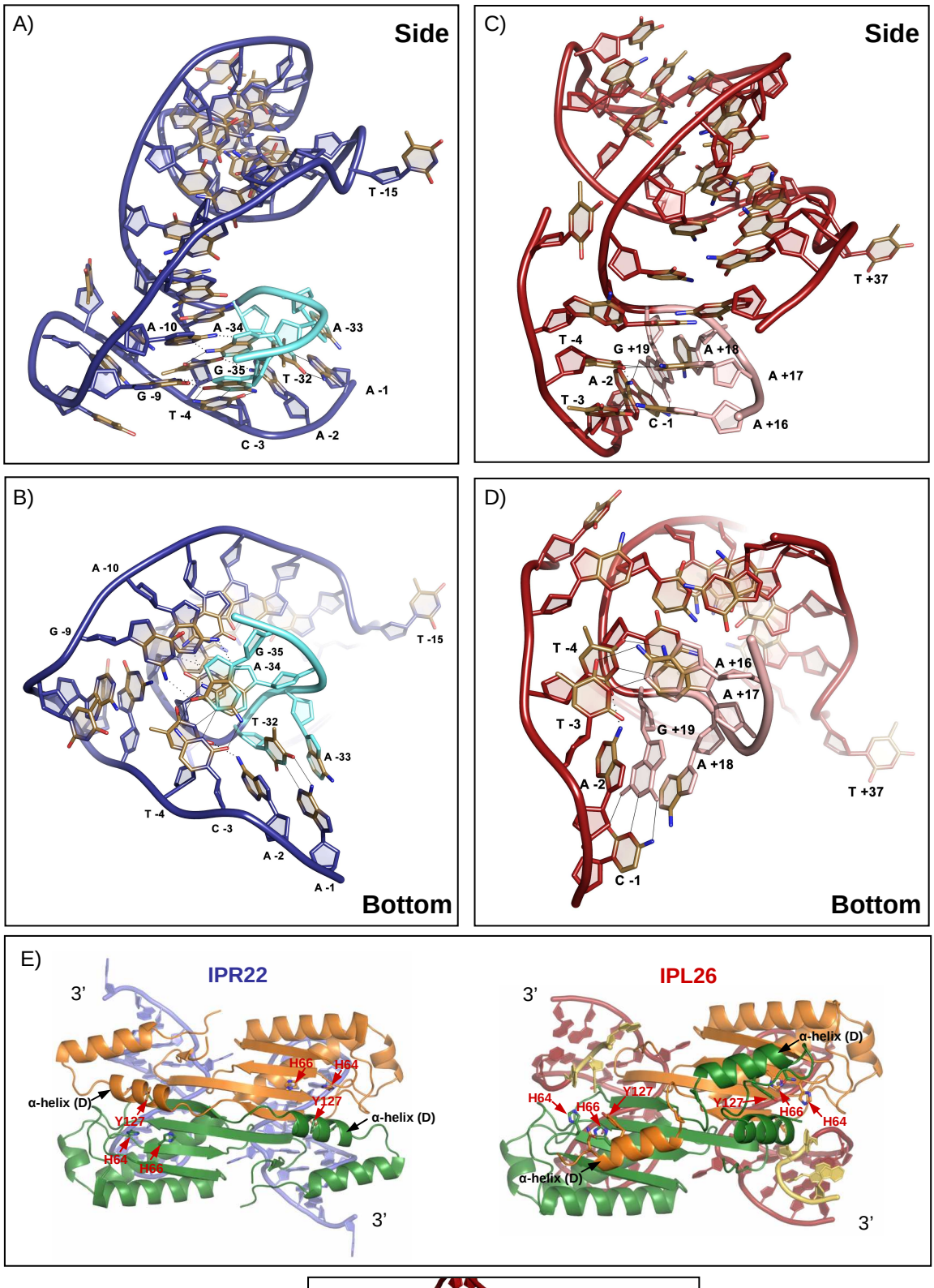
A)
Side
T -15
A -10
A -34
A -33
G -35
T -32
A -1
G -9
T -4
C -3
A -2
C)
Side
T +37
T -4
G +19
A +18
A -2
A +17
T -3
C -1
A +16
B)
A -10
G -9
T -15
G -35
A -34
T -32
A -33
T -4
C -3
A -2
A -1
Bottom
D)
T -4
A +16
A +17
T -3
G +19
A +18
A -2
T +37
C -1
Bottom
E)
IPR22
3'
H66
H64
α-helix (D)
Y127
Y127
α-helix (D)
H64
H66
3'
IPL26
3'
α-helix (D)
H64
H66
Y127
Y127
H66
H64
α-helix (D)
3'
F)

Figure S3

Table S1

Name	Sequence
LE	
LE80	CGCTAGTGCAAAAATTACCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGATACAAGGCGAAACGCCTTTA
LE100	CGGGCTGCAGGAATTCGATTTGCGCTAGTGCAAAAATTACCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGATACAAGGCGAAACGCCTT
LE63	CGCTAGTGCAAAAATTACCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGATAC
LE59	CGCTAGTGCAAAAATTACCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGG
LE64	TTACCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGATACAAGGCGAAACGCCTT
LE60	CAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGATACAAGGCGAAACGCCTT
LE45	AAAGCCCCTAGCTTTTAGCTATGGGGATACAAGGCGAAACGCCTT
LE41	CCCCTAGCTTTTAGCTATGGGGATACAAGGCGAAACGCCTT
LE80_Gm	CGCTAGTGCAAAAATTACCAAAACTAACGCCTTGGAACCCCTAGCTTTTAGCTATGGGGATACAAGGCGAAACGCCTTTA
CTAC	CGCTAGTGCAAAAATTACCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGCTACAAGGCGAAACGCCTTTA
AGAC	CGCTAGTGCAAAAATTACCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGAGACAAGGCGAAACGCCTTTA
CGAC	CGCTAGTGCAAAAATTACCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGCGACAAGGCGAAACGCCTTTA
ATCC	CGCTAGTGCAAAAATTACCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGATCCAAGGCGAAACGCCTTTA
IPL22	CCCCTAGCTTTTAGCTATGGGG
IPL26	AAAGCCCCTAGCTTTTAGCTATGGGG
TTXC	CGCTAGTGCAAAAATTXCCAAAACTAACGCCTTAAAGCCCCTAGCTTTTAGCTATGGGGATACAAGGCGAAACGCCTTTA
RE	
RE45	GAATCCCCTAGCTTTAGCTATGGGGAGTATGTCAACAAATTTGCC
RE49	GCTTTAGCTAGAATCCCCTAGCTTTAGCTATGGGGAGTATGTCAATTCG
RE56	CTCATGCTTTAGCTAGAATCCCCTAGCTTTAGCTATGGGGAGTATGTCAACAAATT
RE70	TTCCTCATGCTTTAGCTAGAATCCCCTAGCTTTAGCTATGGGGAGTATGTCAACAAATTTGCCTAATGAC
RE45_pc	GCTTTAGCTAGAATCCCCTAGCTTTAGCTATGGGGAGTATGTCAA
RE45_Cm	GAATCCCCTAGCTTTAGCTATGGGGAGTATGAGAGCAAATTTGCC
RE45_Cd	TCATGCTTTAGCTAGAATCCCCTAGCTTTAGCTATGGGGAGTATG
RE_Gd	CCCCTAGCTTTAGCTATGGGGAGTATGTCAACAAATTTGCCTAATGAC
RE56_Tm	CTCATGCTTTAGCTAGAATCCCCTAGCTTTAGCTATGGGGCTTATGTCAACAAATT
RE56_T->A	CTCATGCTTTAGCTAGAATCCCCTAGCTTTAGCTAAGGGGAGTATGTCAACAAATT
RE56_T->C	CTCATGCTTTAGCTAGAATCCCCTAGCTTTAGCTACGGGGAGTATGTCAACAAATT

RE56_ΔT	CTCATGCTTTAGCTAGAATCCCCTAGCTTTAGCTAGGGGAGTATGTCAACAAATT
IPR21	CCCCTAGCTTTAGCTATGGGG
IPR22	CCCCTAGCTTTAGCTATGGGGA
IPR25	GAATCCCCTAGCTTTAGCTAGGGG
TC**X**A	GAATCCCCTAGCTTTAGCTATGGGGAGTATGTC**X**ACAAATTTGCC
RE49**F**	GCTTTAGCTAGAATCCCCTAGCTTTAGCTATGGGGAGTATGTCAATTCG-**F**
Tip T_m	CTCATGCTTTAGCTAGAATCCCCTAGCT**AA**AGCTATGGGGAGTATGTCAACAAATT
HP_m	CTCATGCTTTAGCTAGAAT**GGG**C**AT**G**GA**TT**TC**C**AT**TG**CCC**AGTATGTCAACAAATT

C_L and G_L present in red and pink, while C_R and G_R present in blue and pale blue. Mutations are shown in green. **X** = 2-aminopurine. **F** = fluorescein.

Figure S1. Supplementary gel data.

A) IP binding to TnpA cannot be detected by EMSA. EMSA of 5' labelled 22 nt (lanes 1-3) or 25 nt (lanes 4-6) IP_L, and 21 nt (lanes 7-9) or 24 nt (lanes 10-12) IP_R performed with 0 µM, 0.5µM and 2.5 µM TnpA-His6.

B) Effect of single base mutation in the ATAC tetranucleotide 3' to IP_L on synaptic complex robustness. EMSA of 5' labelled LE mutations and unlabeled RE56 with 0 µM, 0.5 µM and 2.5 µM TnpA-His6. ATAC mutated to CTAC (lanes 1-3), AGAC (lanes 4-6), or ATCC (lanes 7-9).

Figure S2. Details of non-canonical base pairing in LE and RE.

Non-canonical base pairing between T_{-4} and A_{-2} within C_R (A), C_L (C) and between $A_{-10}G_{-9}$ and G_{-35} A_{-34} (E and F) are shown as dotted lines (blue for RE and red for LE). Replacement of A_{-2} by $2\text{-}AP_{-2}$ eliminates non-canonical base paring (B and D). Proposed non-canonical base pairing between $A_{+16}A_{+17}$ and $A_{+42}A_{+44}$ is shown in G and H. Interaction between the extrahelical T of IP_R (T_{-15}) or IP_L (T_{+37}) and Leu51 (I). This hydrogen bond is destroyed by replacing T by A (J) or C (K).

Figure S3. An overview of canonical and non-canonical base pairing network in LE and RE ends.

A) Side view of the RE IP structure within the TnpA transpososome. **B) Bottom view of RE**. **C) Side view of the LE IP structure within the TnpA transpososome**. **D) Bottom view of LE.** Figures were drawn with pymol v1.3, www.pymol.org, Schrodinger Inc. Non-canonical base pairings are shown by dotted lines. The guide sequences are shown in pale blue (RE) or pink (LE). The nucleotide coordinates are those from (Barabas et al., 2008) shown in Fig. 5.

E) Overall view of the crystal structure of TnpA-IPR22 (left panel) and TnpA-IPL26 (right panel). One transposase monomer is shown in green and the other in orange. LE and RE are indicated in red and blue respectively and are bound on the inferior face of the TnpA dimer. α-helix D and the active site residues H64, H66 and Y127 are indicated.

F) Simulated structure of transpososome carrying one LE and one RE. It is not possible to generate biologically relevant transpososomes including one LE and one RE for crystallography. This figure is derived from individual resolved structures of transpososomes containing two RE copies or two LE copies. The flanking DNA from the donor molecule is shown in black.

Table S1. Oligonucleotides used.