Supplementary Information

Figure Legends.

Figure S.1 (A) a portion of the experimental electron density map obtained after phasing and density modification with RESOLVE and (B) a 2Fo-Fc electron density map of the same portion after refinement. Both maps are contoured at 1.5σ.

Figure S.2 Superposition of the dimer of Mos1 catalytic domains (blue) with (A) the Tn5 transposase dimer, (B) the Hermes dimer formed across interface 3 (as observed in the hexamer model) and (C) the core domain of HIV-1 integrase. The active site residues of Mos1 and the bound metal are highlighted in purple in each figure. The active sites residues of Tn5 and Hermes are shown in red and those of HIV-1 integrase are highlighted in orange.

Figure S.3 Close up view of the three main points of protein-protein contacts across the putative dimerisation interface: (A) helices 6 of each monomer interact across the interface via residues H293, A296, E297 and a water molecule, (B) residues E345 and K190 form a salt bridge and (C) the N-terminal six residues comprising the sequence WVPHEL. Electrostatic interactions are shown by blue dotted lines and hydrophobic interactions by green lines.

Figure S.4 (A) Analytical gel filtration chromatogram of wild type Mos1 transposase (blue) and the mutants T216A (green) and K190A (pink). 100µl of each protein was applied to a Superdex 200 (HR10/30) column at a concentration of 1µM in Buffer A (25mM Tris, 0.5M NaCl, pH 7.5). Three peaks are labeled: a monomer of transposase (40.6kDa, eluting at 15.9ml), a dimer (81.2kDa, eluting at 12.7 ml) and a high molecular weight species (>1.3MDa) eluting in the void volume (7.9ml). A tetramer of transposase would elute at approximately 12.2ml, but no such peak is observed. The T216A mutant exists solely as a dimer at this loading concentration. Wild type transposase eluted as a dimer (85.4%), a monomer (3.4%) and a high molecular weight species (11.2%). In
contrast, the mutant K190A eluted as 57.4% dimer, 4% monomer and 38.6% high molecular weight species. This data is consistent with the co-existence of a monomer:dimer equilibrium and a monomer:aggregate equilibrium. The K190A mutation destabilizes the dimer interface, thereby promoting the aggregation of the monomer over dimerisation. Conversely, the mutation T216A might disrupt the aggregation interface, thereby promoting dimerisation of transposase monomers over aggregation. (B) Protein-DNA complexes formed during transposition, following separation on a 6% polyacrylamide native gel. The DNA substrate used in the assays was the IRR100 substrate labelled at the 5’ end of the TS. Wild type Mos1 transposase (lanes 1-7) and the mutant K190A Mos1 transposase (lanes 8-14) were tested over a range of concentrations: 112nM (lanes 1 and 8), 56nM (lanes 2 and 9), 28nM (lanes 3 and 10), 14nM (lanes 4 and 11), 7nM (lanes 5 and 12), 3.5nM (lanes 6 and 13) and 1.75nM (lanes 7 and 14). The PEC (labeled) forms with the wild type transposase but not with the K190A mutant.

**Figure S.5** DNA denaturing gel (10%) of the products of second strand DNA cleavage assays, carried out with 10mM MgCl$_2$ (A) or 5mM EDTA (B). Five variants of the IRR100 substrate pre-nicked on the NTS, were used in the assays, with the nick positioned -4 (lanes 1 and 2), -3 (lanes 3 and 4), -2 (lanes 5 and 6), -1 (lanes 7 and 8) or 0 (lanes 9 and 10) bases from the junction between the transposon sequence and the flanking DNA, as shown schematically above the corresponding lanes on the gel. TS and NTS are marked and the position of the radiolabel (on the TS) is shown by an asterisk. Lane numbers are indicated below the gel. Arrows indicate the position of the labeled product of second strand cleavage (70 mer) which has been calibrated with DNA markers of known size. Cleavage of the transferred strand (precisely at the junction between IR and flanking DNA) is maximal when the non-transferred strand is pre-nicked at the -3 position (lane 3), thus confirming this as the site of first strand cleavage. Furthermore, this is the sole cleavage site observed with this substrate. In contrast, additional aberrant sites of cleavage of the transferred strand are observed, at sites within the flanking DNA, with each of the other substrates tested. Thus, the site of the first strand nick is important for directing the fidelity of the second strand cleavage.