



Technical Brief

HFA Biophysical Mechanics

HUMAN FRONTIER LABS

Document 1 of 4 · October 2026

Technical Brief: HFA Biophysical Mechanics

1. Overview

Conventional models of non-LTR retrotransposons, such as mammalian LINE-1 elements and SVA retroposons, treat reverse transcriptase (ORF2p) translocation and priming as stochastic or driven solely by local thermodynamic melting (T_m).

Comparative structural analyses between jumbo staphylococcal phage reverse transcriptases (ART engines) and retroelement orthologs indicate a conserved, sequence-specific allosteric arrest mechanism. A conserved 12-nucleotide AT-rich motif, **TATAATTAATTT**, is proposed to function as a sequence-specific pausing site operating independently of bulk thermodynamic melting.

2. Mechanistic Architecture: the 12-nt AT Tract

- **Minor-groove recognition.** Consecutive A-T tracts exhibit narrowed minor-groove geometry and distinct electrostatic potentials, attributable to the absence of the exocyclic 2-amino group on adenine.
- **Allosteric stalling.** Structural modeling indicates that specialised reverse transcriptase variants engage this minor-groove geometry through a conserved tuning pocket, inducing a conformational pause during cDNA synthesis.
- **Control framework.** To distinguish sequence-specific recognition from passive thermal pausing, an isothermal scrambling control (**TAAATATTTAT**) holds base composition (5 A, 7 T, 0% GC) and nearest-neighbour T_m constant to ΔT_m 0.000 °C while altering nucleotide order. The control is additionally screened on both strands against known regulatory and retroelement insertion motifs; Document 3 gives the selection criteria.

TERMINOLOGY

TATAATTAATTT is not palindromic; its reverse complement is **AAATTAATTATA** . It is described throughout this package as a motif or AT tract rather than a dyad, since it has no two-fold rotational symmetry.

3. Genomic Coordinates and Enrichment

Positional sweeps were conducted across human reference sequence to test whether the motif occurs more often than base composition alone predicts.

PARAMETER	VALUE	BASIS
-----------	-------	-------

Reference sequence	AP001685.1	<i>H. sapiens</i> chromosome 21q, section 29/105, 340,000 bp
Local AT background	64.8%	Measured from the accession
Observed occurrences	3	2 forward, 1 reverse complement
Expected by chance	0.91	Both strands, at measured composition
Enrichment	3.3×	Observed / expected
Poisson $p(\geq 3)$	0.065	Does not meet a 0.05 threshold

SCOPE OF THIS FIGURE

The enrichment is calculated on a single 340 kb clone, which is 0.73% of chromosome 21 (46,709,983 bp in GRCh38). It establishes regional presence. It does not establish genome-wide enrichment, and it does not reach conventional significance on this clone alone. Extending the sweep across additional clones with a composition-matched null is the measurement that would settle it.

4. Structural Adaptation: the GNAT-Fold Tuning Pocket

Comparative homology modelling across phage and human retroelement reverse transcriptase domains indicates a lineage-specific substitution within the putative GNAT-fold tuning pocket.

STATE	APPROX. SIDE-CHAIN VOLUME	RESIDUE CHARACTER
Ancestral	166.0 Å ³	Bulky aliphatic (Leu / Val)
Variant	89.0 Å ³	Serine
Net change	−46.4%	Sterically reduced cleft

EVIDENCE CLASS

These volumes derive from comparative structural prediction, not from solved structures or in vitro post-translational modification mapping. No purified protein has been assayed. The assignment is a hypothesis to be tested, and the assay described in Documents 2 and 3 does not test it directly.

5. What This Package Does and Does Not Establish

- **Established:** the motif is present in AP001685.1 at the stated coordinates, and the sequences in Document 3 are composition- and T_m -matched.

- **Proposed and untested:** that the motif causes sequence-specific polymerase arrest.
- **Proposed and not addressed by this assay:** the GNAT-fold acetylation mechanism, and any downstream genomic or therapeutic implication.
- **Required before single-molecule work:** a reproducible, motif-dependent pause in bulk assay, with the controls listed in Document 2.