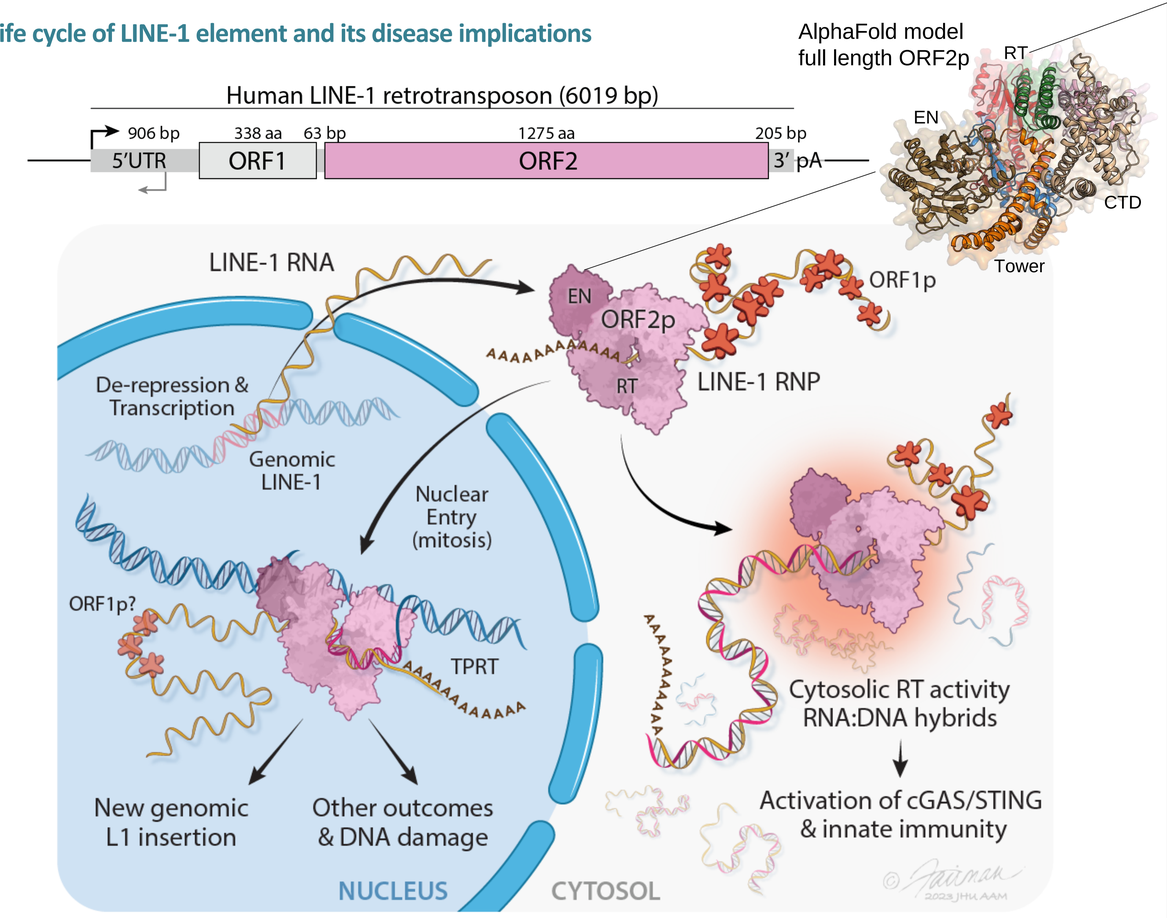
Integrated structural biology of the LINE-1 retrotransposon

Authors: Ariele Viacava Follis[1], Charles Nichols[2], Thomas Walpole[2], Paul Wan[2], Eddy Arnold[3], Eric Baldwin[1], Roberto Sanchez[1], Donna Romero[1]

Affiliations: 1 - ROME Therapeutics, Boston MA, USA; 2 - Charles River Laboratory, Chesterford, UK; 3 – Rutgers University, New Brunswick NJ, USA

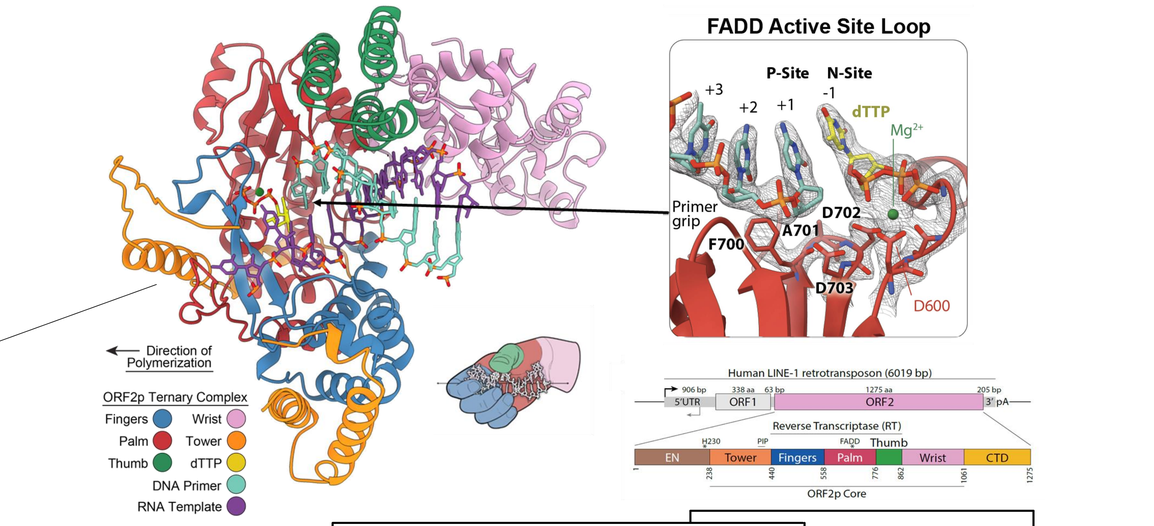
BACKGROUND: The LINE-1 (L1) retrotransposon has written almost half of the human genome through a “copy-and-paste” mechanism enabled by its two transcripts: the multifunctional enzyme ORF2p and the chaperone ORF1p. ORF2p contains a reverse transcriptase (RT) at its core and a structurally independent endonuclease (ENDO) domain at its N terminus. ORF1p is a coiled coil trimeric assembly of small RNA binding domains that further assemble to form multimeric ribonucleoprotein (RNP) particles in cells. LINE-1 has emerged as a potential therapeutic target implicated in cancer, autoimmunity, neurodegenerative diseases and aging. We recently solved the first crystal structures of human LINE-1 RT, characterizing distinct states in the catalytic cycle and in a catalytically stalled state with an inhibitor bound. Combined with access to Line-1 ENDO and ORF1p structures, this enables us to take an integrated discovery approach to rationally design small molecules that modulate each aspect of the LINE-1 function to fit the nuanced requirements of different therapeutic indications.

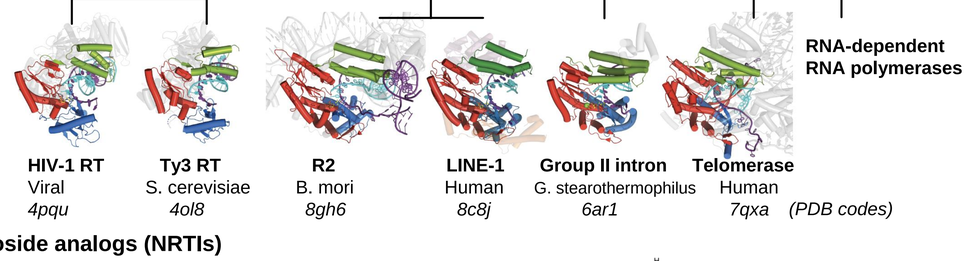
Life cycle of LINE-1 element and its disease implications

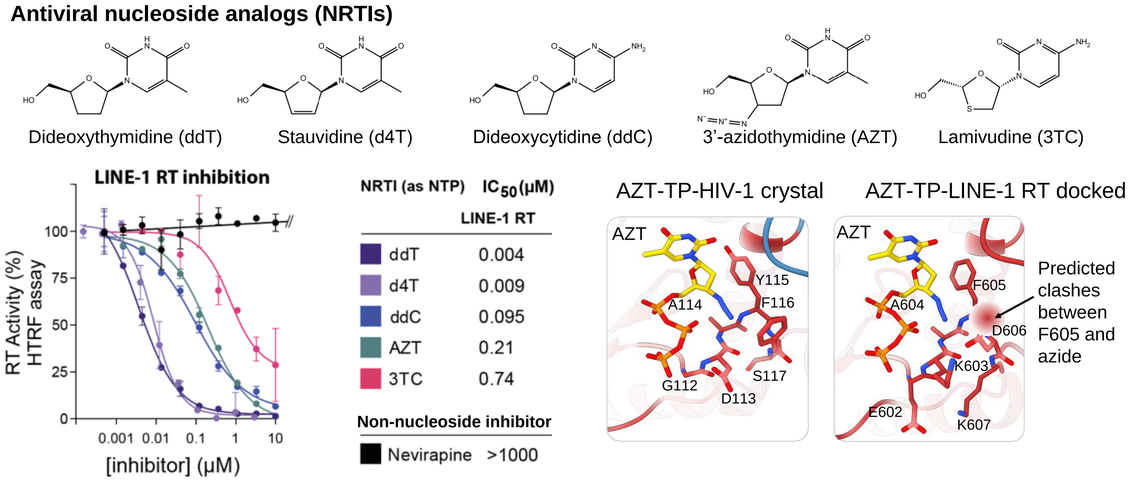


Baldwin, E.T., van Eeuwen, T., Hoyos, D. et al.
Structures, functions, and adaptations of the human LINE-1 ORF2 protein. Nature 2024, 626, 194

LINE-1 RT constitutes the central catalytic element sustaining the LINE-1 life cycle. Opportunities towards its inhibition are suggested by antiviral reverse transcriptase inhibitors.



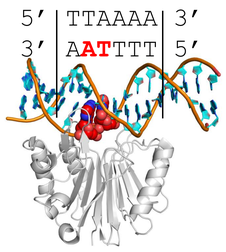




Baldwin, E.T., van Eeuwen, T., Hoyos, D. et al.
Structures, functions, and adaptations of the human LINE-1 ORF2 protein. Nature 2024, 626, 194

LINE-1 requires two additional contributors to sustain its life cycle: the endonuclease (EN) domain of ORF2p and the RNA chaperone, ORF1p. These LINE-1 elements provide additional opportunities for a fine-tuned pharmaceutical intervention mitigating the undesirable consequences of LINE-1 activation.

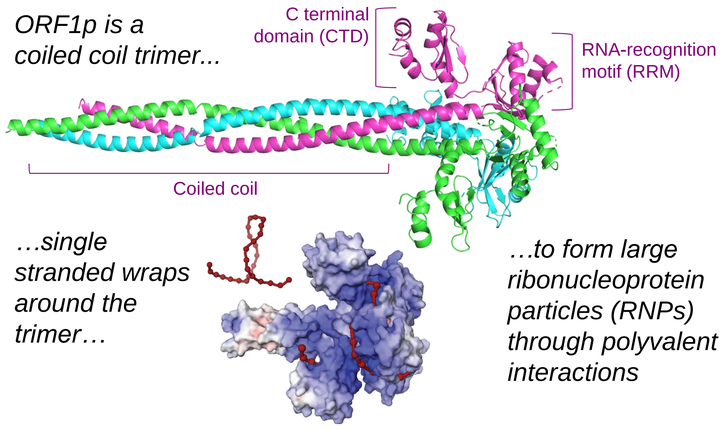
EN nicks double stranded DNA to enable genomic insertion of LINE-1



Reported inhibitors of the structurally related abasic site endonuclease APE1 suggest opportunities to target LINE-1 EN

Miller, I., et al. Nucleic Acids Res, 2021, 49, 11350

ORF1p is a coiled coil trimer...



…single stranded wraps around the trimer…

…to form large ribonucleoprotein particles (RNPs) through polyvalent interactions

Khazina E., et al. Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. Nat Struct Mol Biol, 2011, 18, 1006

Results

- The LINE-1 RT crystal structure confirms a closer evolutionary relationship of non-LTR retroelement RTs with bacterial RTs compared to retroviral and LTR retroelement RTs. This dictates the need for a focused discovery effort. Available antiviral RT inhibitors are unlikely optimal inhibitors for LINE-1 RT, especially non-nucleoside HIV-1 inhibitors.
- 42 proprietary liganded L1-RT structures, combined with the characterization of LINE-1 RT inhibition by nucleoside and non-nucleoside antiviral RT inhibitors enable a rational design strategy towards potent and selective LINE-1 RT inhibitors.
- The structures of LINE-1 endonuclease domain and ORF1p indicate additional strategies to modulate LINE-1 activity in an orthogonal fashion to the direct inhibition of its RT activity.

Conclusion

The determination of the first high resolution crystal structure of LINE-1 Reverse Transcriptase enables a focused effort towards the development of potent and selective LINE-1 RT inhibitors. Additionally, the structures of other functional elements of LINE-1 suggest opportunities to specifically modulate different aspects of the complex LINE-1 life cycle and their different disease implications.