
Structures, Functions, and Adaptations of the Human LINE-1 ORF2 Protein

Eric Baldwin¹, Trevor Van Eeuwen², David Hoyos³, Arthur Zalevsky⁴, Egor Tchesnokov⁵, Roberto Sanchez¹, Bryant Miller⁶, Luciano Distefano⁷, Francesc Xavier Ruiz⁸, Matthew Hancock⁹, Esin İşik⁶, Carlos Mendez-Dorantes⁶, Thomas Walpole¹⁰, Charlie Nichols¹⁰, Paul Wan¹⁰, Kirsi Riento¹⁰, Rowan Halls-Kass¹⁰, Martin Augustin¹¹, Alfred Lammens¹¹, Anja Jestel¹¹, Paula Upla², Kera Xibinaku¹², Samantha Congreve¹², Maximiliaan Hennink¹², Kacper Rogala¹³, Anna Schneider¹⁴, Jennifer Fairman¹⁵, Shawn Christensen¹⁶, Brian Desrosiers¹, Gregory Bisacchi¹, Oliver Saunders¹, Nafeeza Hafeez¹, Wenyan Miao¹, Rosana Kapeller¹, Dennis Zaller¹, Andrej Sali¹⁷, Oliver Weichenrieder¹⁴, Kathleen Burns, Matthias Gotte⁵, Michael Rout², Eddy Arnold¹⁸, Benjamin Greenbaum, Donna Romero¹, John Lacava^{2,7}, and Martin Taylor^{*19}

¹ROME Therapeutics – Boston, MA, United States

²The Rockefeller University – New York, NY, United States

³Memorial Sloan Kettering Cancer Center – New York, NY, United States

⁴University of California, San Francisco – San Francisco, CA, United States

⁵University of Alberta – Edmonton, Alberta, Canada

⁶Dana-Farber Cancer Institute [Boston] – 450 Brookline Ave. Boston, MA 02215, United States

⁷European Research Institute for the Biology of Ageing [Groningen] – Antonius Deusinglaan 1, 9713 AV Groningen, Netherlands

⁸Rutgers University – Camden, Newark, New Brunswick, New Jersey, United States

⁹University of California [San Francisco] – 505 Parnassus Ave, San Francisco, CA 94143, United States

¹⁰Charles River Laboratories – Saffron Walden, Essex, United Kingdom

¹¹Proteros Biostructures – Planegg, Germany

¹²Whitehead Institute – 455 Main St, Cambridge, MA 02142, États-Unis, United States

¹³Stanford School of Medicine – Stanford, CA, United States

¹⁴Max Planck Institute for Biology, Tübingen – Tübingen, Germany

¹⁵Johns Hopkins University School of Medicine [Baltimore] – Baltimore, MD, United States

¹⁶University of Texas at Arlington – Arlington, TX, United States

¹⁷University of California, San Francisco – San Francisco, CA, United States

¹⁸Rutgers University – New Jersey, United States

¹⁹Department of Pathology, Massachusetts General Hospital and Harvard Medical School – Boston, MA USA, United States

Abstract

The LINE-1 (L1) retrotransposon is an ancient genetic parasite that has written around one third of the human genome through a "copy-and-paste" mechanism catalyzed by its

*Speaker

multifunctional enzyme, open reading frame 2 protein (ORF2p). ORF2p reverse transcriptase (RT) and endonuclease activities have been implicated in the pathophysiology of cancer, autoimmunity, and aging, making ORF2p a potential therapeutic target. However, a lack of structural and mechanistic knowledge has hampered efforts to rationally exploit it. We report structures of the human ORF2p ‘core’ (residues 238-1061, including the RT domain) by X-ray crystallography and cryo-EM in multiple conformational states. Our analyses reveal two novel folded domains, extensive contacts to RNA templates, and associated adaptations that contribute to unique aspects of the L1 replication cycle. Computed integrative structural models of full-length ORF2p show a dynamic closed ring conformation that appears to open during retrotransposition. We characterize ORF2p RT inhibition and reveal its underlying structural basis. Imaging and biochemistry reveal that non-canonical cytosolic ORF2p RT activity can produce RNA:DNA hybrids, activating innate immune signaling via cGAS/STING and resulting in interferon production. In contrast to retroviral RTs, L1 RT is efficiently primed by short RNAs and hairpins, which likely explains cytosolic priming. Additional biochemical activities including processivity, DNA-directed polymerization, non-templated base addition, and template switching together allow us to propose an updated L1 insertion model. Finally, our evolutionary analysis reveals structural conservation between ORF2p and other RNA- and DNA-dependent polymerases. We therefore provide key mechanistic insights into L1 polymerization and insertion, shed light on L1 evolutionary history, and enable rational drug development targeting L1. (if time permits and the committee finds this of interest, unpublished structural and mechanistic data will also be presented)

Keywords: LINE, 1, ORF2p, Structure, Priming, Insertion mechanisms, Enzymology, Innate immune activation, interferon, cGAS, STING