Murine Leukemia Virus Integration:
Local and Global Target Site Recognition.

Sriram Aiyer[1], G.V.T. Swapna[2], Paolo Rossi[2], Nirav Malani[4], William M. Schneider[1], Ashwin Chanda[1], James Aramini[2], Frederic D. Bushman[4], Gaetano T. Montelione[2, 3], and Monica J. Roth[*, 1]

[1]Dept Pharmacology, Robert Wood Johnson Medical School, Rutgers University, Piscataway, NJ, 08854, [2]Center for Advanced Biotechnology and Medicine, Dep. Molecular Biology and Biochemistry, and Northeast Structural Genomics Consortium (NESG), Rutgers University, Piscataway, NJ, 08854, [3]Dept. Biochemistry and Molecular Biology, Robert Wood Johnson Medical School, Rutgers University, Piscataway, NJ, 08854, [4]Dept. Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104.

* Corresponding Author roth@rwjms.rutgers.edu

Keywords: Retroviral integration, integrase, BET proteins, MLV .

Abstract

Target-site selection by retroviral integrase (IN) proteins has profound effects on viral pathogenesis. We describe the solution NMR structure of the Moloney murine leukemia virus IN (M-MLV) C-terminal domain (CTD) and a structural homology model of the catalytic core domain (CCD). The MLV IN CTD (PDB ID: 2M9U) adopts an SH3 domain fold followed by a 28 residue unstructured tail peptide (TP). We have obtained a concordant structural model of the MLV IN CCD using SWISS-MODEL, MMM-tree and I-TASSER servers. Using the PFV IN target capture X-ray structure and structure-based sequence alignment, residues within the CCD α2 helical region and the CTD β1-β2 loop were predicted to bind target DNA. Chimeric viruses with substitutions at the CCD α2 helical region and the CTD β1-β2 loop result in viable virus. Next-generation sequencing and analysis of local integration target sites indicate the CCD α2 helical region, in particular P187, interacts with the sequence outside the target site duplication (TSD), whereas the CTD β1-β2 loop binds within TSD.

The host BET proteins Brd2, 3 and 4 interact through the ET domain with the MLV IN. We have characterized the IN CTD:Brd3 ET domain interaction using solution NMR. The ET domain induces a disorder-to-order transition of the unstructured C-terminal tail region of IN CTD. Solution NMR based rotational correlation experiments show heterodimeric interactions between the ET domain and IN CTD. Peptide competition assays with MLV IN wildtype and mutant TP identify the consensus $_{390}WX_7PLKJR_{402}$ to be important for this interaction. Using 454 sequencing, we show that truncation of the IN C-terminal TP affects the global targeting profile of MLV vectors by decreasing viral integration near TSS and CpG islands. WT IN integration events show strong correlation within 100 bp of a known BET binding site. Viruses bearing IN C-terminal truncations can provide new avenues to improve the safety profile of gammaretroviral vectors for gene therapy.