Cell

# X-Ray Structures of the Hexameric Building Block of the HIV Capsid

Owen Pornillos,[1,5] Barbie K. Ganser-Pornillos,[1,5] Brian N. Kelly,[2] Yuanzi Hua,[1] Frank G. Whitby,[2] C. David Stout,[3] Wesley I. Sundquist,[2] Christopher P. Hill,[2] and Mark Yeager[1,4,5,*]
[1]Department of Cell Biology, The Scripps Research Institute, La Jolla, CA 92037, USA
[2]Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84112, USA
[3]Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA
[4]Division of Cardiovascular Diseases, Scripps Clinic, La Jolla, CA 92037, USA
[5]Department of Molecular Physiology and Biological Physics, University of Virginia Health System, Charlottesville, VA 22908, USA
*Correspondence: yeager@scripps.edu
DOI 10.1016/j.cell.2009.04.063

## SUMMARY

The mature capsids of HIV and other retroviruses organize and package the viral genome and its associated enzymes for delivery into host cells. The HIV capsid is a fullerene cone: a variably curved, closed shell composed of approximately 250 hexamers and exactly 12 pentamers of the viral CA protein. We devised methods for isolating soluble, assembly-competent CA hexamers and derived four crystallographically independent models that define the structure of this capsid assembly unit at atomic resolution. A ring of six CA N-terminal domains form an apparently rigid core, surrounded by an outer ring of C-terminal domains. Mobility of the outer ring appears to be an underlying mechanism for generating the variably curved lattice in authentic capsids. Hexamer-stabilizing interfaces are highly hydrated, and this property may be key to the formation of quasi-equivalent interactions within hexamers and pentamers. The structures also clarify the molecular basis for capsid assembly inhibition and should facilitate structure-based drug design strategies.

## INTRODUCTION

The ribonucleoprotein genomic complex of human immunodeficiency virus type 1 (HIV-1) is encased within the mature capsid, a predominantly cone-shaped shell assembled from ~1,500 copies of the viral CA protein (recently reviewed by Ganser-Pornillos et al., 2008). The HIV-1 capsid is a fullerene cone (Ganser et al., 1999; Jin et al., 1999; Li et al., 2000), with a body composed of a curved two-dimensional (2D) array of ~250 CA hexamers. To form a closed shell, the ends of the cone are capped by exactly 12 pentamers, with seven at the broad end and five at the narrow end. The fullerene model and its generality across retroviruses are now widely accepted (e.g., see Heymann et al., 2008).

HIV-1 CA is a highly helical protein with two independently folded domains, the N-terminal domain (NTD) and C-terminal domain (CTD), which are flexibly linked. Published high-resolution CA structures include isolated domains and full-length monomers from HIV-1 and other retroviruses (Berthet-Colominas et al., 1999; Campos-Olivas et al., 2000; Cornilescu et al., 2001; Gamble et al., 1997; Gitti et al., 1996; Jin et al., 1999; Khorasanizadeh et al., 1999; Mortuza et al., 2004, 2009). These structures collectively demonstrate that retroviral CA proteins share a common tertiary fold despite having widely divergent amino acid sequences and that, as a corollary, mature retroviral capsids are likely to be stabilized by similar quaternary interactions. The hexagonal capsid lattice is composed of three different types of interfaces: a six-fold symmetric NTD-NTD interface that creates hexameric rings, an intermolecular interface between the two domains (NTD-CTD) that reinforces the hexamer, and a homodimeric CTD-CTD interface that links the hexameric building blocks into an infinite hexagonal lattice (Bowzard et al., 2001; Gamble et al., 1997; Ganser-Pornillos et al., 2004, 2007; Lanman et al., 2003, 2004; Li et al., 2000; von Schwedler et al., 1998, 2003). This lattice architecture was unambiguously established by an electron cryomicrocroscopy (cryoEM) structure of 2D crystals of HIV-1 CA hexamers, albeit at moderate resolution (~9 Å) (Ganser-Pornillos et al., 2007). Moreover, crystal structures of isolated NTD and CTD have provided atomic models for two of the different types of interfaces in the hexagonal lattice: the dimeric CTD-CTD interface of HIV-1 CA (Worthylake et al., 1999) and the hexameric interface formed by the isolated NTD of murine leukemia virus (MLV) CA (Mortuza et al., 2004).

Despite steady progress in elucidating the structure of the retroviral capsid lattice, high-resolution crystal structures of hexagonal arrays of full-length retroviral CA proteins have not yet been reported. This is presumably due to the low intrinsic stability of CA hexamers and the challenges of preparing discrete oligomeric CA assemblies. Here, we describe engineered HIV-1 CA proteins that form homogenous populations of stable, soluble hexamers, which are functional for assembly in vitro. The X-ray crystal structures of these CA proteins extend our understanding of the hexameric capsid assembly unit to atomic resolution.

## RESULTS AND DISCUSSION

### Isolation of HIV-1 CA Hexamers for Crystallization

Discrete HIV-1 CA hexamers were stabilized by two independent methods: thiol crosslinking and template-directed