Supplementary Information

Structure of a monomeric mutant of the HIV-1 capsid protein

Ronald Shin[¶], Ywh-Min Tzou[§] & N. Rama Krishna^{¶,§,*}

[¶]Comprehensive Cancer Center, and [§]Department of Biochemistry and Molecular Genetics,
University of Alabama at Birmingham, Birmingham, AL 35294

Media used in protein expression:

Recipe for P5052N25: 50 mM Na₂HPO₄, 50 mM KH₂PO₄, 25mM ¹⁵N NH₄Cl, 5mM Na₂SO₄, 2mM MgSO₄, 0.5% glycerol, 0.2% lactose, 0.05% glucose and 1X trace mineral.

The reducing of NH₄Cl from 50 mM in original P5052N (Studier, F. W. (2005) *Protein Expr Purif.* 41(1), 207–234) to 25mM did not affect the yields.

Recipe for P040: 50mM Na₂HPO₄, 50mM KH₂PO₄, 25mM ¹⁵N NH₄Cl, 5mM Na₂SO₄, 2mM MgSO₄, 0.4% ¹³C glucose, and 1X trace minerals.

P040 is not an auto-induction medium. The gene regulation mechanism of the auto-induction system relies on glycerol (as the source of carbon) which does not inhibit the T7 expression. The target gene was induced by lactose when the inhibitors (e.g., glucose, amino acids) were consumed by *E. coli*. This protocol is very convenient for ¹⁵N labeling. However, for ¹⁵N and ¹³C double labeled samples, glucose is more economical as a ¹³C source than glycerol. Because lactose cannot induce the expression in the presence of glucose in the medium, IPTG was used for the induction of proteins containing ¹³C labeling.

Recipe for 1000X trace mineral: 50 mM FeCl₃, 20 mM CaCl₂, 10 mM each of MnCl₂ and ZnSO₄, and 2 mM each of CoCl₂, CuCl₂, NiCl₂, Na₂MoO₄, Na₂SeO₃ and H₃BO₃ in ~60 mM HCl.

Figure captions:

Figure S1: Sequential and short range NOE contacts established in the monomeric mutant of HIV-1 CA. The helices are identified at the top. The highly flexible linker between NTD and CTD domains spans the sequence 145 to 149 corresponding to the sequence YSPTS. Of these the first and last residues are part of unwound helices. The start codon residue Met preceding Pro1 was left intact in the sequence, but was not used in the structure calculations or assignments (it could not be assigned).

Figure S2: Representative panels from the 3D-NMR data (a) HNCA , (b) HNCACB and (c) HCCH-TOCSY. Sequential connectivities in HNCA and HNCACB are shown.

Figure S3: Representative panels showing NOE contacts from (a) the ^{15}N -NOESY- HSQC and (b) ^{13}C -NOESY-HSQC data.

Figure S4: (a) Superposition of the 20 best NMR structures with the α -helices in the NTD domains aligned. Because of the highly flexible four residue linker, the CTD domains are distributed over a wide space. (b) Same 20 structures with the α -helices in the CTD domains aligned, with the NTD domains distributed over a wide space. The helices are shown in magenta color.

Figure S5: pET20bHIV1CA plasmid

Figure S6: HIV-1 W184A/M185A-CA protein. Lane 1: protein expressed without His-tag; Lane2: protein markers; Lane 3: protein with His-tag.

Figure S7: Overlay of ^{15}N -HSQC spectra of W184A/M185A-CA at 0.5 mM (red) and 2 mM (black).

Figure S1

Figure S2

Figure S3

(A)

Figure S4a

(B)

Figure S4b

Figure S5: pET20bHIV1CA plasmid

Figure S6: HIV-1 W184A/M185A-CA protein.

Lane 1: protein without His-tag; Lane2: protein markers;
Lane 3: protein with His-tag.

Black = 2.0 mM concentration

Red = 0.5 mM concentration

Figure S7