Supplementary Material

Mammalian Production of an Isotopically Enriched Outer Domain of the HIV-1 gp120 Glycoprotein for NMR Spectroscopy

Mallika Sastry, Ling Xu, Ivelin S. Georgiev, Carole A Bewley,

Gary J. Nabel, and Peter D. Kwong

Supplementary Methods

MALDI TOF spectrometry

In-gel digestion: Gel bands were transferred to pre-digested tubes and 100 µL 0.01M DTT/0.1M Tris, pH 8.5 was added. The tube was placed in a heating block at 55° for 2h. After cooling the tube to room temperature, the liquid was removed and replaced with 100 µL 0.03M iodoacetamide/0.1M Tris, pH 8.5. This was allowed to react for 30 min. in the dark after which the liquid was removed and the gel was washed as described below. Gel bands were prepared for digestion by washing twice with 200 µL 0.05M Tris, pH 8.5/ 30% acetonitrile for 20 minutes with shaking, and once with 100 µL acetonitrile for several minutes until the gel was opaque white. After removing the acetonitrile, the gel pieces were dried for 20-30 min. in a Speed-Vac concentrator. Gels were digested by adding 0.10 µL modified trypsin (sequencing grade, Roche Molecular Biochemicals, Indianapolis, IN) in 50 µL 0.025M Tris, pH 8.5, or enough volume to completely hydrate the gel. The tubes were placed in a heating block at 32° and left overnight. Peptides were extracted with 2X 50 µL 50% acetonitrile/2% TFA and the combined extracts were dried in a Speed-Vac concentrator.

MALDI-MS Analysis: Matrix solution was prepared by making a 10 mg/mL solution of 4-hydroxy-α-cyanocinnamic acid in 50% acetonitrile/ 0.1% TFA and adding two internal standards, angiotensin and ACTH 7-38 peptide, to the matrix solution. The final concentration of standards was 193 fmole/µL angiotensin and 340 fmole/µL ACTH. The dried digest was dissolved in 3 µL matrix/standard solution and 0.5 µL was spotted onto the sample plate. When the spot was completely dried, it was washed twice with water. MALDI mass spectrometric analysis was performed on the digest using an Applied Biosystems Voyager DE Pro mass spectrometer in the reflector mode.

Quantification of isotope incorporation for $^{15}N/^{13}C$ labeled HIV-1 gp120 outer domain

MALDI TOF mass spectroscopy analysis of the outer domain tryptic digest heptapeptide TIIVQLR expressed in $^{15}N/^{13}C$-labeled media showed a complex mass pattern with five distinct modes. Isotope incorporation analysis for the double labeled outer domain utilized the estimate of ^{15}N incorporation in the 81-89% range (see main text). From this subset, distributions for

which the mass of the highest peak matched the mass of one of the five different experimentally-observed modes were included in further analysis. Based on this analysis, the ^{13}C incorporation resulting in the best correlation (R^2=0.3461, $p<0.0001$) was at 84% for ^{15}N and 84% for ^{13}C, with a corresponding experimental mode at mass 884.5. The distributions for which the mass of the highest peak matched the 884.5 mode corresponded to a range of ^{13}C incorporation of 83-87%. Hence, the percentage incorporation of the double labeled OD was estimated to be 85 ± 2% for ^{13}C given a 85 ± 4% incorporation for ^{15}N outer domain.

Table S1. Correlations between computed super-peaks and observed M/S peaks for Outer Domain and green fluorescent protein.

Outer domain*		Green Fluorescent Protein#	
^{15}N Enrichment Levels (%)	R^2 Correlation	^{15}NEnrichment Levels (%)	R^2 Correlation
81	0.729	69	0.873
82	0.752	70	0.900
83	0.772	71	0.921
84	0.789	72	0.937
85	0.802	73	0.948
86	0.809	74	0.952
87	0.812	75	0.949
88	0.808		
89	0.798		

*Shown are only ^{15}N enrichment levels at which the mass of the highest-intensity computed peak matched the mass (852.5) of the highest-intensity M/S peak. The best correlation was observed at an enrichment level of 87%.

#Shown are only ^{15}N enrichment levels at which the mass of the highest-intensity computed peak matched the mass (1358.6) of the highest-intensity M/S peak. The best correlation was observed at an enrichment level of 74%.

Table S2. Mammalian expression systems used to obtain isotopically enriched proteins

Media	Cell Line	^{15}N Media	$^{15}N/^{13}C$ Media
Algal and bacterial mixture of amino acids[1]	Sp2/0	30 mg/L,	30 mg/L,
Algal mixture of amino acids[2]	CHO	10mg/L*	10mg/L*
Commercial media supplemented by labeled amino acids CIL[3]	Mouse Hybridoma cells 1B10.7	1.5 mM*	1.2 mM*
CIL Bioexpress 6000 ($^{15}N/^{15}N, ^{13}C$ GKLQSTVW) [4]	HEK293	2mg/L*	2.12mg/L*
Commercial Media (CIL) (Current work)	A549/Adenoviral Expression	50mg/L	43 mg/L

*Cost of media was not reported
[1] (Hansen et al. 1992)
[2] (Lustbader et al. 1996)
[3] (Shindo et al. 2000)
[4] (Werner et al. 2008)

Table S3. Isotope masses and abundances used for theoretical estimates of isotope incorporation[5]

Element	Isotope	Mass	Abundance
H	^{1}H	1.0078	100.0
	^{2}H	2.0141	0.015
C	^{12}C	12.0000	100.0
	^{13}C	13.0034	1.119
N	^{14}N	14.0031	100.0
	^{15}N	15.0001	0.368
O	^{16}O	15.9949	100.0
	^{17}O	16.9991	0.037
	^{18}O	17.9991	0.204

[5](Kubinyi 1991)

Table S4. Surface Plasmon resonance analysis of glycosylated proteins

		R2core gp120	Glycosylated Outer Domain		
		Unlabeled	Unlabeled	^{15}N	$^{15}N^{13}C$
b12	k_d (1/s)	0.01077	0.04198	0.02939	0.03044
	K_D (M)	5.42×10^{-8}	4.66×10^{-8}	2.95×10^{-8}	3.75×10^{-8}
b13	k_d (1/s)	0.00075	0.01080	0.02045	0.01.025
	K_D (M)	1.35×10^{-8}	7.51×10^{-8}	11.4×10^{-8}	9.41×10^{-8}

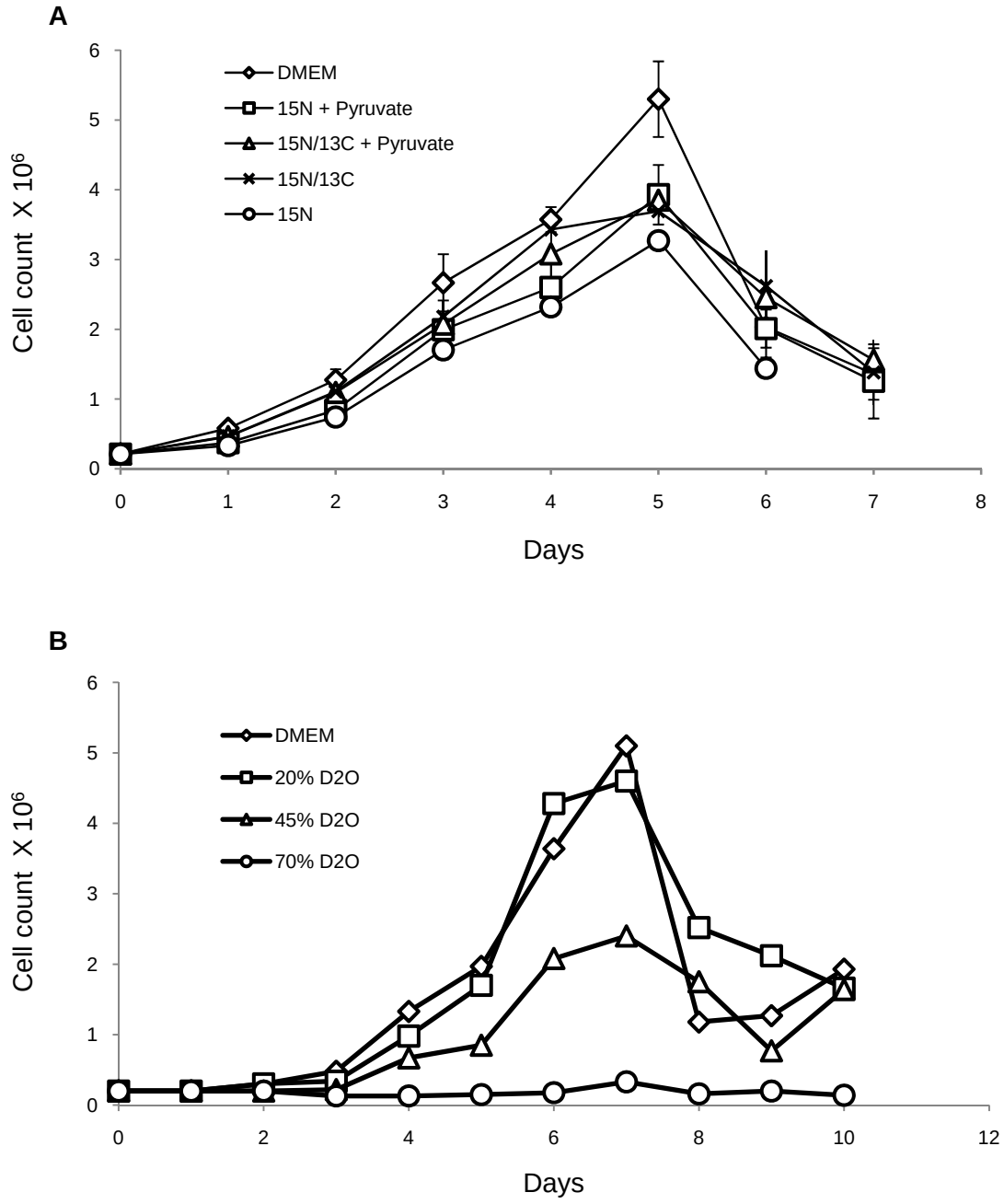

Fig S1. Growth characteristics of the A549 mammalian cell line. Growth characteristics of the cell line used to obtain isotopic enrichment was evaluated by obtaining growth curves in different labeled and unlabeled media.

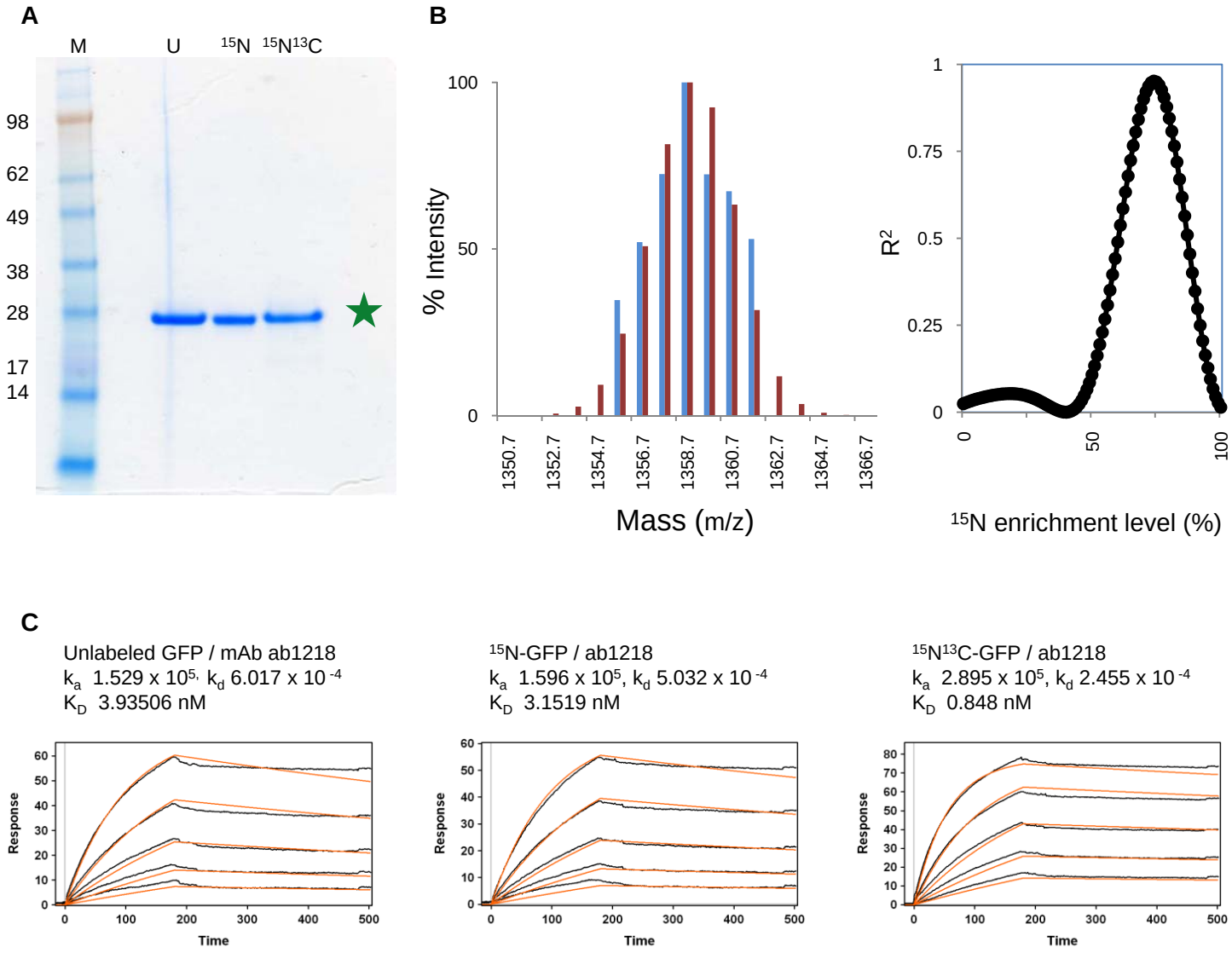

Fig S2. Characterization of isotopically enriched GFP expressed using the Mammalian Expression system. Correctly folded GFP can be expressed and purified using the adenoviral expression system. Panel **A**: SDS Page analysis of green GFP. Lanes U, ^{15}N, $^{15}N^{13}C$: unlabeled, ^{15}N, $^{15}N/^{13}C$ labeled GFP a ~26 kDa protein that was used for biophysical measurements. Lane M: Molecular weight markers. Panel **B**: Mass spectral analysis of a tryptic peptide TIFFKDDGNYK to determine % incorporation of ^{15}N. A comparison of experimental and computed pattern for 74 % incorporation of ^{15}N is shown (left graph). The correlation between observed experimental pattern and computed patterns are shown for each percentage incorporation of ^{15}N (right graph). Panel **C**: Surface plasmon resonance analysis of unlabeled and ^{15}N, $^{15}N/^{13}C$ labeled GFP binding to ab1218 an anti GFP antibody.

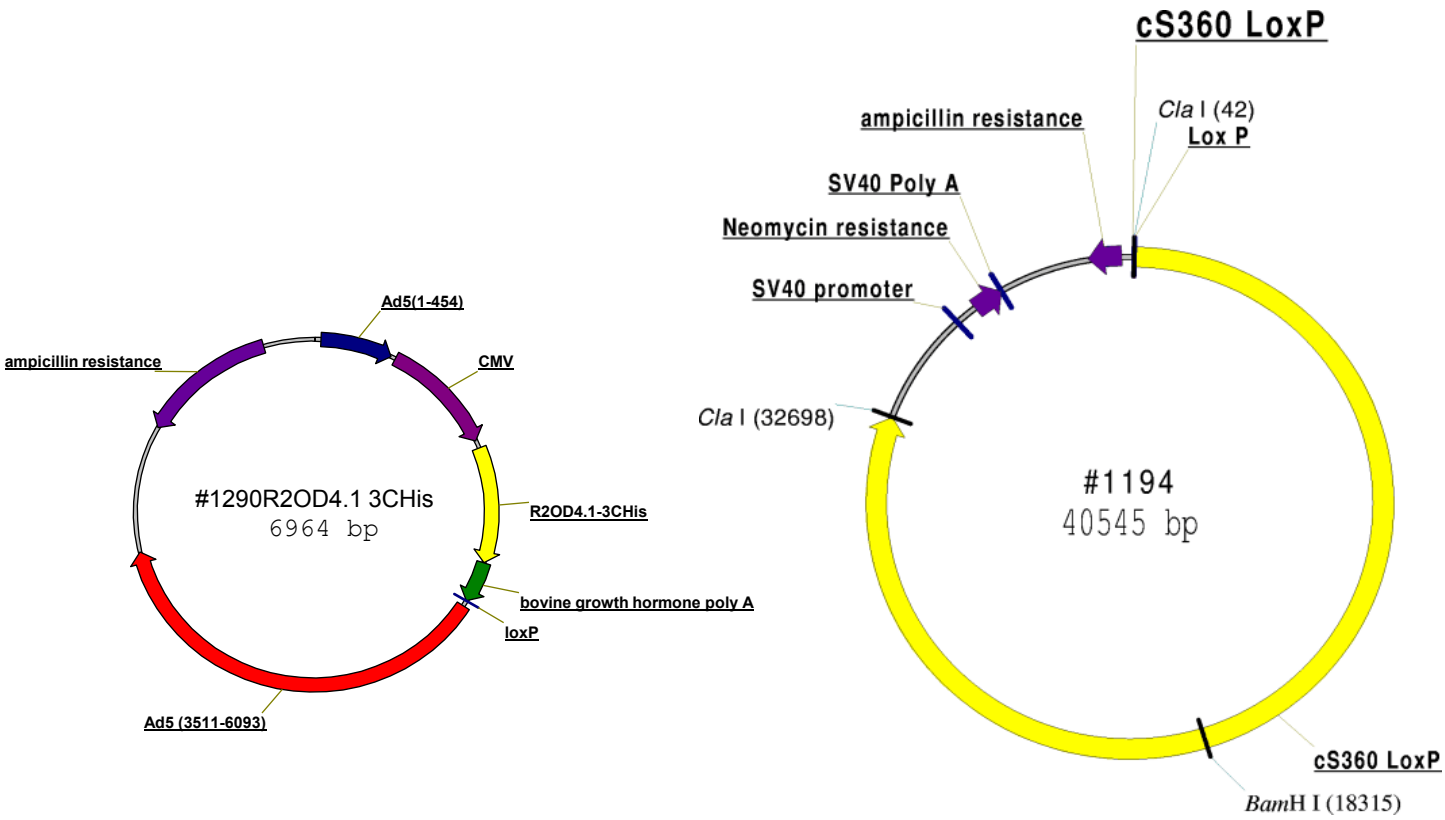

Fig S3. Plasmid maps for the shuttle vector (p1290) and the adenoviral cosmid (p1194) used in the cre-lox recombination for adenoviral vector production.

MYSMQLASCVTLTLVLLVNSAPRRPVVSTQLLLNGSLAEEEVVIRSENFTNCAKTIIVQLRE
PVKINCSRPNNNTRGRRGDIRQAHCNISKTNWTNALKQVVEKLGEQFNKTKIVFTQSSGGDP
EIVTHSFNCAGEFTYCNTTQLFDSIWNSENGTWNITRGLNNTGRNDTITLPCRIKGSGAPPI
KGNISCSSNITGLLLTRDCGKDDNSRDGNETFRPGGGDMRDNWRSEGSLEVLFQGPGHHHHH
H

Fig S4. Sequence of HIV-1 gp120 outer domain used in this study. The mouse IL2 leader sequence is shown in red and the C terminal HRV3C cleavage site and the Histidine tag used for purification are shown in green.

Fig. S5. Sequence alignment of full length clade B R2 gp120 and the HIV-1 gp120 outer domain (OD4.1) used in this study (amino acid numbering using HIV-1_{HxBc2} gp120 as the standard). Putative glycosylation sites (NXT/S) within the outer domain are highlighted and the V3 loop is shown in red. HIV-1 R2 construct (OD4.1) was designed by deleting β20/21, the V3 loop along with N362Q, F382T point mutations and an engineered disulfide bond to further stabilize the outer domain.

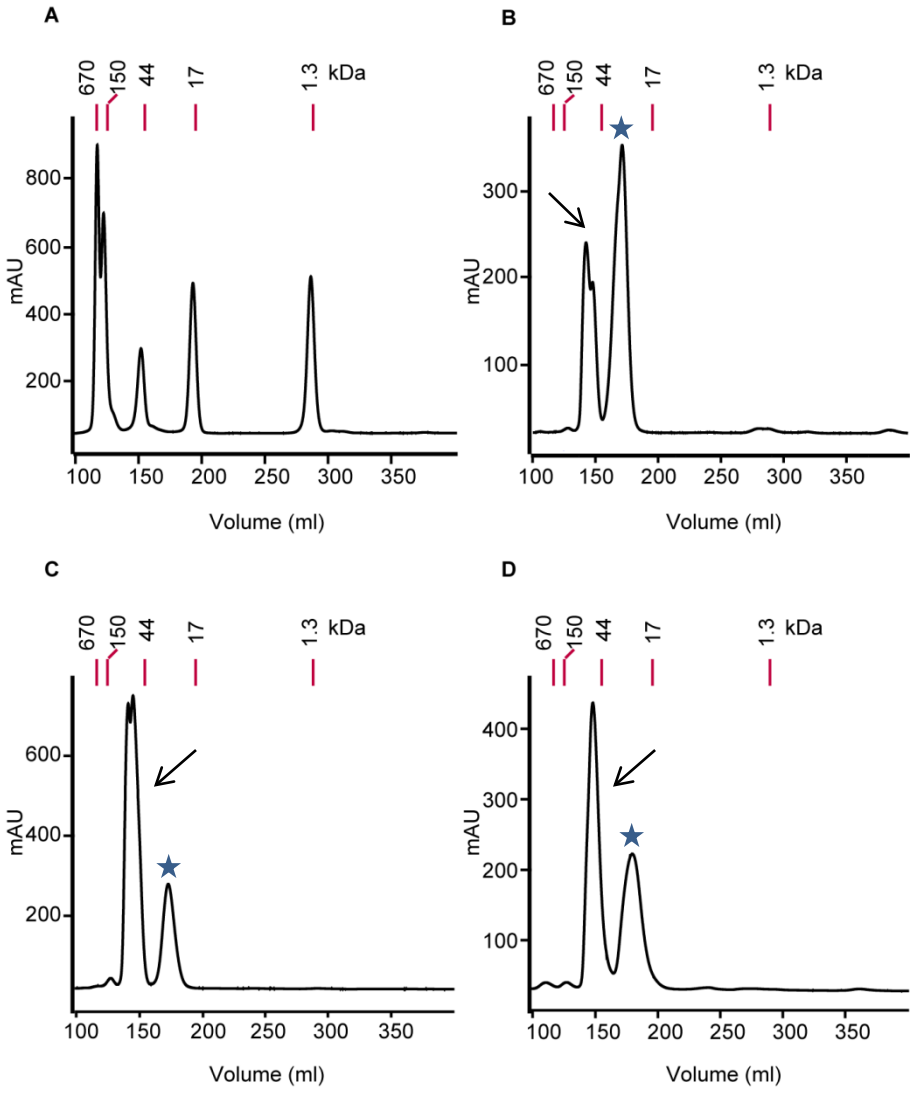

Fig. S6. Characterization of isotopically enriched HIV-1 gp120 outer domain expressed using the adenoviral/mammalian expression system by size exclusion chromatography. The deglycosylated HIV-1 gp120 outer domain exhibits elution profile characteristic of a globular monomeric protein. Panel **A**: Protein standards of known molecular weights are labeled on the chromatogram. Panel **B**: Deglycosylated unlabeled HIV-1 gp120 outer domain (★) elutes as a monomer at 171 ml. Panel **C**: Deglycosylated ^{15}N HIV-1 gp120 outer domain (★) elutes as a monomer at 173 ml. Panel **D**: $^{15}N/^{13}C$ HIV-1 gp120 outer domain (★) also elutes as a monomer at 179 ml. $EndoH_f$ a recombinant fusion protein of Endoglycosidase H and Maltose binding protein of apparent molecular weight of 79 kDa is marked with an arrow.

References

Hansen AP, Petros AM, Mazar AP, Pederson TM, Rueter A, Fesik SW (1992) A practical method for uniform isotopic labeling of recombinant proteins in mammalian cells. Biochemistry 31 (51):12713-12718.

Kubinyi H (1991) Calculation of isotope distributions in mass spectrometry. A trivial solution for a non-trivial problem. Analytica Chimica Acta 247 (1):107-119.

Lustbader JW, Birken S, Pollak S, Pound A, Chait BT, Mirza UA, Ramnarain S, Canfield RE, Brown JM (1996) Expression of human chorionic gonadotropin uniformly labeled with NMR isotpes in Chinese hamster ovary cells: An advance toward rapid determination of glycoprotein structures. Journal of Biomolecular NMR 7 (4):295-304.

Shindo K, Masuda K, Takahashi H, Arata Y, Shimada I (2000) Letter to the Editor: Backbone 1H, 13C, and 15N resonance assignments of the anti-dansyl antibody Fv fragment. Journal of Biomolecular NMR 17 (4):357-358.

Werner K, Richter C, Klein-Seetharaman J, Schwalbe H (2008) Isotope labeling of mammalian GPCRs in HEK293 cells and characterization of the C-terminus of bovine rhodopsin by high resolution liquid NMR spectroscopy. Journal of Biomolecular NMR 40 (1):49-53.