Supplementary Information for

Structural basis for highly effective HIV-1 neutralization by CD4-mimetic miniproteins revealed by 1.5 Å co-crystal structure of gp120 and M48U1

Priyamvada Acharya[1], Timothy Luongo[1], Mark K. Louder[1], Krisha McKee[1], Yongping Yang[1], Young Do Kwon[1], John R. Mascola[1], Pascal Kessler[2], Loïc Martin[2] and Peter D. Kwong[1,*]

[1] Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA.
[2]CEA, iBiTecS, Service d'Ingénierie Moléculaire des Protéines, Gif sur Yvette F-91191, France.

Table of contents

Table S1, related to Table 1. Binding parameters of sCD4 and CD4-mimetic miniproteins to full-length YU2 gp120.

	On-rate ($M^{-1}s^{-1}$)	Off-rate (s^{-1})	KD (nM)
CD4 (Stricher et al., 2008)	$5.13\times10^{4}\pm1.3\times10^{2}$	$3.94\times10^{-4}\pm1.7\times10^{-5}$	7.68±0.33
M48 (Stricher et al., 2008)	$1.10\times10^{5}\pm1.9\times10^{2}$	$4.27\times10^{-4}\pm1.1\times10^{-5}$	3.87±0.10
M47 (Stricher et al., 2008)	$3.49\times10^{5}\pm1.1\times10^{3}$	$7.41\times10^{-4}\pm1.7\times10^{-5}$	2.12±0.05
M48U1 (this study)	$2.53\times10^{6}\pm2.0\times10^{4}$	$3.90\times10^{-5}\pm2.0\times10^{-6}$	0.015±0.008
M48U7 (this study)	$2.42\times10^{6}\pm5.0\times10^{4}$	$4.46\times10^{-4}\pm2.5\times10^{-5}$	0.184±0.030

Table S2, related to Figure 3. Interactions at the miniprotein-gp120$_{YU2}$ interface outside the Phe43 cavity. (A) Number of miniprotein residues and total surface area buried. (B-F) List residues buried at interface with gp120 for (B) M48U1, space group P212121, (C) M48U1, space group C2221, (D) M48U7, (E) M48 and (F) M48. Accessible Surface Area, Å² (**ASA**) and Buried Surface Area, Å² (**BSA**) were calculated using the EBI PISA server (http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver). For these calculations, residue 23 in each miniprotein was mutated *in silico* to alanine. **Residues contributing to the interface through H bonds or salt bridges are indicated by a blue letter H or a red letter S, respectively.**

A

Miniprotein (Chain name; PDB ID)	Number of Residues buried at interface	Miniprotein surface area buried at interface outside Phe43 cavity (A^2)
M48 (M, 2I60)	11	442
M48 (S, 2I60)	13	535
M47 (M, 2I5Y)	12	501
M47 (S, 2I5Y)	13	518
M48U1 (R, xxx)	13	526
M48U1 (M, xxx)	13	507
M48U1 (R, xxx)	13	537
M48U7(R, xxx)	13	521

B

PDB ID: 4JZW

M8U1 (chain R)	ASA	BSA
R:MPT 1	42.41	18.59 |||||
R:ASN 2	106.12	16.07 ||
R:ARG 9	151.84	43.98 |||
R:LEU 13	123.02	17.87 ||
R:LEU 15	98.6	22.59 |||
R:ARG 18	162.39	64.73 ||||
R:ALA 20[H]	55.1	39.82 |||||||
R:DPR 21	111.72	52.54 |||||
R:THR 22[H]	133.73	106.86 |||||||
R:ALA 23	52.06	43.70 ||||||||
R:CYS 24[H]	39.08	35.77 |||||||||
R:ALA 25	33.06	17.07 ||||||
R:CYS 26[H]	76.75	45.97 ||||||

M8U1(chain M)	ASA	BSA
M:MPT 1	40.8	15.01 ||||
M:ASN 2	113.19	25.98 |||
M:ARG 9[HS]	138.22	19.80 ||
M:LEU 13	120.84	17.41 ||
M:LEU 15	88.85	19.55 |||
M:ARG 18	174.46	65.37 ||||
M:ALA 20[H]	54.42	39.29 |||||||
M:DPR 21	110.41	51.20 |||||
M:THR 22[H]	132.97	105.55 |||||||
M:ALA 23	52.74	44.71 ||||||||
M:CYS 24[H]	41.05	38.72 |||||||||
M:ALA 25	31.54	15.90 ||||||
M:CYS 26[H]	84.46	48.53 ||||||

C PDB ID: 4JZZ

M8U1 (chain R)	ASA	BSA
R:MPT 1	46.06	18.75 |||||
R:ASN 2	117.74	22.06 ||
R:ARG 9[HS]	151.12	43.45 |||
R:LEU 13	115.19	15.71 ||
R:LEU 15	84.11	24.34 |||
R:ARG 18	174.87	70.92 |||||
R:ALA 20[H]	55.85	39.29 |||||||
R:DPR 21	113.09	52.51 |||||
R:THR 22[H]	135.09	106.53 |||||||
R:ALA 23	51.8	44.11 ||||||||
R:CYS 24[H]	40.52	36.89 |||||||||
R:ALA 25	31.12	16.39 ||||||
R:CYS 26[H]	79.77	45.62 ||||||

D PDB ID: 4K0A

M8U7 (chain R)	ASA	BSA
R:MPT 1	44.29	19.89 |||||
R:ASN 2	108.95	18.05 ||
R:ARG 9[HS]	154.41	42.73 |||
R:LEU 13	114.7	13.21 ||
R:LEU 15	86.27	24.15 |||
R:ARG 18	174.64	67.84 ||||
R:ALA 20[H]	51.24	35.84 |||||||
R:DPR 21	113.9	50.08 |||||
R:THR 22[H]	134.96	107.64 ||||||||
R:ALA 23	52.46	44.45 ||||||||
R:CYS 24[H]	44.56	38.74 ||||||||
R:ALA 25	29.34	14.10 |||||
R:CYS 26[H]	78.58	44.64 ||||||

E PDB ID: 2I60

M48 (chain M)	ASA	BSA
M:MPT 1	39.27	22.52 ||||||
M:ASN 2	96.67	15.68 ||
M:LEU 15	91.61	17.08 ||
M:ARG 18	152.69	39.83 |||
M:ALA 20[H]	39.88	35.36 ||||||||
M:DPR 21	117.52	53.09 |||||
M:THR 22[H]	127.19	95.96 ||||||||
M:ALA 23	55.12	44.93 ||||||||
M:CYS 24[H]	65.74	47.49 |||||||
M:ALA 25	28.19	20.76 |||||||
M:CYS 26[H]	92.76	49.07 ||||||

M48 (chain S)	ASA	BSA
S:MPT 1	38.85	18.02 |||||
S:ASN 2	79.47	37.59 |||||
S:ARG 9[HS]	155.3	31.22 |||
S:LEU 13	127.18	22.50 ||
S:LEU 15	93.54	20.41 |||
S:ARG 18	178.96	72.29 |||||
S:ALA 20[H]	54.2	38.89 ||||||||
S:DPR 21	116.29	49.32 |||||
S:THR 22[H]	132.88	104.35 ||||||||
S:ALA 23	56.79	49.60 |||||||||
S:CYS 24[H]	39.27	34.91 ||||||||
S:ALA 25	35.54	15.90 |||||
S:CYS 26[H]	74.12	40.28 ||||||

Table S3, related to Table 1 and Figure 7. Neutralization data (IC_{50} in µg/ml) for CD4-mimetic miniproteins, M48, M47, M48U1 and M48U7, and soluble CD4 (sCD4).

Strains	Clade	M48U1	M48U7	M48	M47	sCD4
0260.v5.c36	A	2.50		>50	37.3	>50
0330.v4.c3	A	4.060		>50	>50	7.65
0439.v5.c1	A	0.301		18.0	15.9	18.2
3365.v2.c20	A				4.64	
3415.v1.c1	A	0.056		1.97	4.50	22.9
3718.v3.c11	A	0.878		>50	31.2	>50
398-F1_F6_20	A	0.449		15.0	15.7	>50
BB201.B42	A	33.6		>50	>50	30.2
BB539.2B13	A	0.209		5.18	4.31	1.86
BI369.9A	A	2.04		>50	>50	5.95
BS208.B1	A	0.005		0.288	1.10	5.72
KER2008.12	A	7.65	>50	0.396	1.85	1.42
KER2018.11	A	0.120		6.84	13.4	3.24
KNH1209.18	A	2.43		>50	>50	38.7
MB201.A1	A	>50		>50	>50	>50
MB539.2B7	A	1.95		48.9	17.2	29.6
MI369.A5	A	1.85		>50	40.0	10.7
MS208.A1	A	1.11		>50	>50	8.01
Q168.a2	A	5.48	>50	>50	>50	11.7
Q23.17	A	3.88	48.6	>50	33.5	>50
Q259.17	A	0.627		22.4	26.4	25.9
Q461.e2	A	3.22		>50	>50	27.2
Q769.d22	A	0.017		0.377	1.28	1.31
Q769.h5	A	0.035		0.755	1.76	1.78
Q842.d12	A	2.56	19.4	>50	>50	>50
QH209.14M.A2	A	28.279		>50	>50	>50
RW020.2	A	1.91		>50	42.5	13.0
UG037.8	A	0.114	3.53	8.51	5.27	0.798
3301.V1.C24	AC	0.265		29.9	7.86	4.71
3589.V1.C4	AC	0.111		4.40	1.95	15.4
6540.v4.c1	AC	2.20		>50	>50	>50
6545.V3.C13	AC				>50	
6545.V4.C1	AC	1.26		23.6	>50	>50
0815.V3.C3	ACD	0.055		6.13	2.71	1.20
6095.V1.C10	ACD	0.030		0.334	0.281	0.144
3468.V1.C12	AD	1.28		>50	49.0	6.89
620345.c1	AE	>50		>50	>50	>50
C1080.c3	AE	>50	>50	11.5	>50	4.03
C2101.c1	AE	>50		27.7	>50	13.9
C3347.c11	AE	>50	>50	6.25	>50	>50

Strains	Clade	M48U1	M48U7	M48	M47	sCD4
C4118.09	AE	>50		>50	>50	>50
CNE3	AE	>50		>50	20.7	27.3
CNE5	AE	>50		0.275	>50	1.52
CNE55	AE	>50		>50	>50	6.65
CNE56	AE	>50	>50	7.67	22.7	15.2
CNE59	AE	>50	>50	5.99	11.4	0.762
M02138	AE	>50		>50	>50	1.140
R1166.c1	AE	>50		15.7	>50	>50
R2184.c4	AE	>50		47.3	>50	11.7
R3265.c6	AE	>50		>50	>50	>50
TH966.8	AE	>50		25.3	45.0	5.37
TH976.17	AE	>50		18.9	>50	25.6
235-47	AG	0.617		>50	19.7	25.8
242-14	AG	0.027	0.091	4.93	6.87	2.05
263-8	AG	0.119		1.09	0.206	1.73
269-12	AG	0.910		>50	>50	13.0
271-11	AG	2.81		0.020	0.284	12.3
928-28	AG	0.073		0.535	0.266	1.03
DJ263.8	AG	0.071		1.10	0.746	0.048
T250-4	AG	0.058		8.65	3.44	27.8
T251-18	AG	0.809		24.3	7.81	12.3
T253-11	AG	1.160		>50	>50	>50
T255-34	AG	0.015		1.07	2.28	2.67
T257-31	AG	1.817		>50	1.73	>50
T266-60	AG	1.89		>50	15.9	>50
T278-50	AG	1.24		>50	22.8	2.46
T280-5	AG	0.222		9.81	12.0	>50
T33-7	AG	19.300		8.72	24.6	10.2
3988.25	B	9.910		>50	>50	>50
5768.04	B	0.008		0.434	1.02	1.46
6101.10	B	0.028		2.02	14.2	5.12
6535.3	B	0.011		0.195	0.211	1.38
7165.18	B	0.132	0.050	13.3	0.210	7.34
89.6.DG	B	0.006		0.052	0.046	>50
AC10.29	B	0.436	1.34	>50	4.32	13.5
ADA.DG	B	0.003		0.175	0.061	0.184
Bal.01	B	0.003		0.091	0.027	0.081
BaL.26	B	0.002		0.157	0.074	0.115
BG1168.01	B	0.096	0.103	18.5	0.446	28.9
BL01.DG	B	10.600	1.87	8.00	9.57	>50

Strains	Clade	M48U1	M48U7	M48	M47	sCD4
BR07.DG	B	0.002		0.029	0.012	4.55
BX08.16	B	<0.023		<0.023	<0.023	0.352
CAAN.A2	B	0.025		3.85	0.106	21.4
CNE10	B	0.051		0.893	0.381	2.43
CNE12	B	0.185		17.2	6.70	20.3
CNE14	B	0.112		11.6	18.10	12.5
CNE4	B	0.002		0.033	0.040	0.099
CNE57	B	0.152		3.91	1.220	7.94
HO86.8	B	3.71		>50	>50	>50
HT593.1	B	0.012		0.707	0.081	1.09
HXB2.DG	B	0.0004		0.020	0.035	0.048
JRCSF.JB	B	0.003		0.212	0.054	0.682
JRFL.JB	B	0.012	0.007	3.53	2.010	1.52
MN.3	B	<0.000 01		0.050	0.039	0.043
PVO.04	B	0.04		0.667	0.470	10.0
QH0515.01	B	0.98		>50	17.600	2.8
QH0692.42	B	0.016	0.003	0.111	0.191	0.852
R2	B	0.007		0.087		0.140
REJO.67	B	0.099		3.09	0.270	4.60
RHPA.7	B	0.477		29.1	4.970	13.5
SC422.8	B	0.010		2.01	0.865	7.53
SF162.LS	B	0.002		0.056	0.083	0.222
SS1196.01	B	0.004		0.054	0.044	0.876
THRO.18	B	0.035		0.503	0.144	0.139
TRJO.58	B	0.518		27.8	30.000	16.7
TRO.11	B	0.071		37.1	5.880	28.2
WITO.33	B	0.110		25.4	0.803	7.63
YU2.DG	BC	0.015		0.393	0.201	0.259
CH038.12	BC	1.60		>50	>50	36.5
CH070.1	BC	0.079		2.22	2.000	2.22
CH117.4	BC	0.515		>50	45.000	>50
CH181.12	BC	0.204		7.27	3.560	>50
CNE15	BC	0.047		6.140	3.530	2.41
CNE40	BC	0.0004		0.008	0.041	0.039
CNE7	BC	0.039		3.47	3.57	1.60
286.36	C	0.093		5.58	3.11	>50
288.38	C	0.013		1.23	0.453	1.88
0013095-2.11	C	0.100		1.71	0.857	0.619
001428-2.42	C	<0.023		1.56	1.480	3.93
0077_V1.C16	C	0.121		0.124	0.169	5.07
00836-2.5	C	0.167		0.549	10.400	5.45

Strains	Clade	M48U1	M48U7	M48	M47	sCD4
0921.V2.C14	**C**				14.300	
16055-2.3	**C**	0.256		7.81	5.050	>50
16845-2.22	**C**	0.007		0.523	0.120	0.431
16936-2.21	**C**	0.272		4.18	1.010	>50
25710-2.43	**C**	0.041		0.344	0.302	1.16
25711-2.4	**C**	0.956		>50	27.800	28.0
25925-2.22	**C**	0.184		20.5	10.000	10.3
26191-2.48	**C**	0.196		15.6	7.080	20.2
3168.V4.C10	**C**	0.192		16.1	2.860	19.9
3637.V5.C3	**C**	0.131		1.89	2.220	17.2
3873.V1.C24	**C**	0.040		3.69	13.400	1.79
6322.V4.C1	**C**	0.060		4.84	6.990	3.82
6471.V1.C16	**C**	0.068		6.44	12.200	3.25
6631.V3.C10	**C**	0.743	8.66	>50	>50	2.72
6644.V2.C33	**C**	0.008		>50	0.023	>50
6785.V5.C14	**C**	0.999		43.7	18.50	3.430
6838.V1.C35	**C**				4.610	
96ZM651.02	**C**	0.084		7.49	0.86	10.9
BR025.9	**C**	6.73		>50	>50	3.22
CAP210.E8	**C**	0.132	0.027	3.58	1.740	1.13
CAP244.D3	**C**	0.162		10.7	13.500	4.92
CAP45.G3	**C**	16.5		>50	>50	19.6
CNE30	**C**	0.071		0.458	0.129	0.213
CNE31	**C**	0.260		>50	>50	14.1
CNE53	**C**	0.032		0.292	0.12	5.93
CNE58	**C**	0.133		21.500	3.460	24.1
DU123.06	**C**	0.279		17.7	16.000	0.631
DU151.02	**C**	1.66	3.43	>50	>50	4.83
DU156.12	**C**	0.606		>50	>50	6.60
DU172.17	**C**	0.486	0.951	>50	34.3	3.63
DU422.01	**C**	0.205		20.0	8.450	16.0
MW965.26	**C**	0.001		>50	0.065	0.112
SO18.18	**C**	0.256		8.94	2.640	15.9
TV1.29	**C**	0.782	0.863	>50	>50	3.54
TZA125.17	**C**	0.004		0.180	0.183	45.1
TZBD.02	**C**	0.672		20.2	3.760	>50
ZA012.29	**C**	7.4		>50	>50	35.8
ZM106.9	**C**	0.293		23.0	9.52	17.9
ZM109.4	**C**	0.172		3.08	7.450	0.082
ZM135.10a	**C**	0.052		19.3	0.258	12.2

Strains	Clade	M48U1	M48U7	M48	M47	sCD4
ZM146.7	**C**	0.037		2.12	0.645	43.1
ZM176.66	**C**	0.232		1.40	1.310	4.31
ZM197.7	**C**	0.036		0.683	0.377	5.27
ZM214.15	**C**	0.034		0.718	1.110	3.67
ZM215.8	**C**	0.558		>50	9.190	27.9
ZM233.6	**C**	0.186		1.34	1.690	1.64
ZM249.1	**C**	1.35		>50	>50	16.5
ZM53.12	**C**	0.729	1.67	43.8	36.500	9.33
ZM55.28a	**C**	6.57		>50	>50	>50
3326.V4.C3	**CD**	0.117		10.2	0.368	>50
3337.V2.C6	**CD**	<0.023		1.20	0.154	0.265
3817.v2.c59	**CD**	2.78		>50	>50	>50
231965.c1	**D**	0.124		5.91	3.190	>50
247-23	**D**	3.250		>50	>50	>50
3016.v5.c45	**D**	0.483		23.2	14.800	>50
57128.vrc15	**D**	0.090	0.087	7.99	29.900	6.76
6405.v4.c34	**D**	2.070		>50	>50	>50
A03349M1.vrc4a	**D**	0.671		12.60	29.400	>50
NKU3006.ec1	**D**	0.045		1.24	2.2	>50
UG021.16	**D**	0.004		0.550	0.672	0.026
UG024.2	**D**	0.0001		0.003	0.015	0.005
X2088.c9	**G**	0.180	0.227	7.98	13.000	>50
SIVmac251.30.SG3	**NA**	>50	>50	>50	>50	>50
SVA.MLV	**NA**	>50		>50	>50	>50

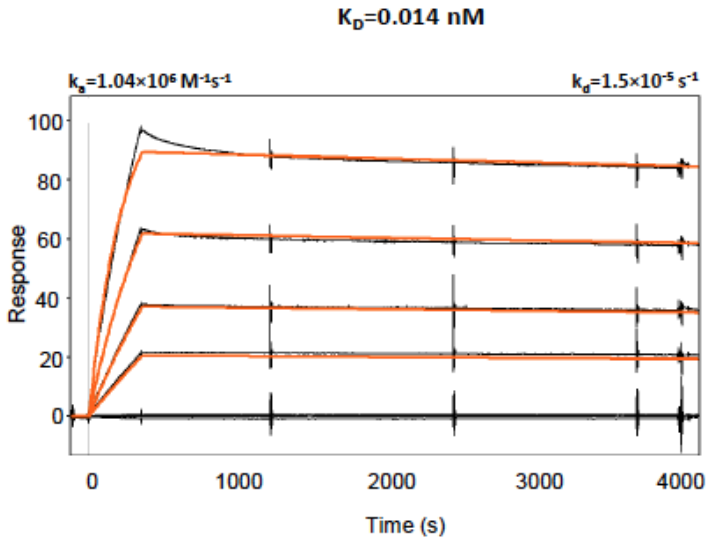


Figure S1, related to Figure 1A. SPR profile of HIV-1 YU2 gp120 binding to M48U1 immobilized on a CM5 chip. The black lines indicate independent injections of M48U1 from concentrations ranging from 4nM to 0.5nM, with concentrations sampled at 2-fold dilution. The red lines show the global fit of the data to a Langmuir 1:1 binding model.

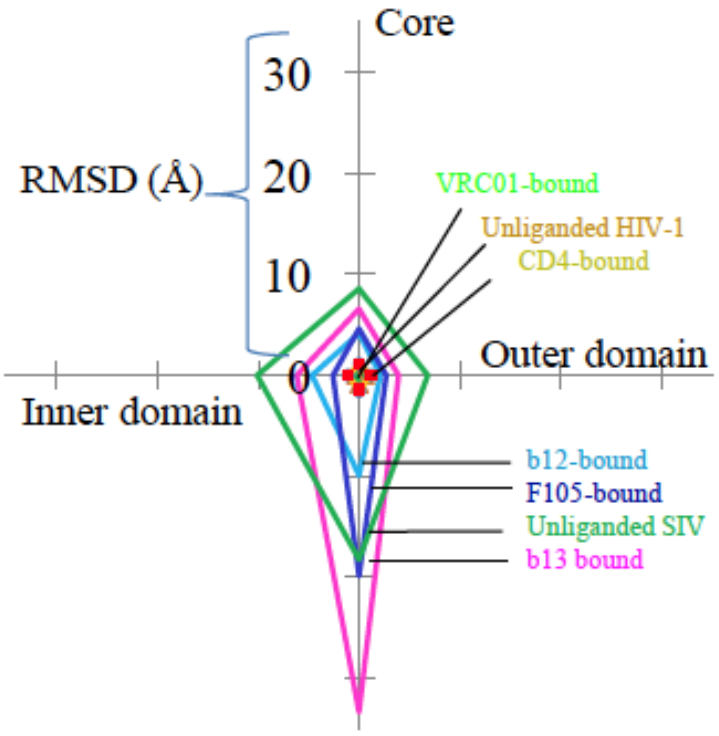
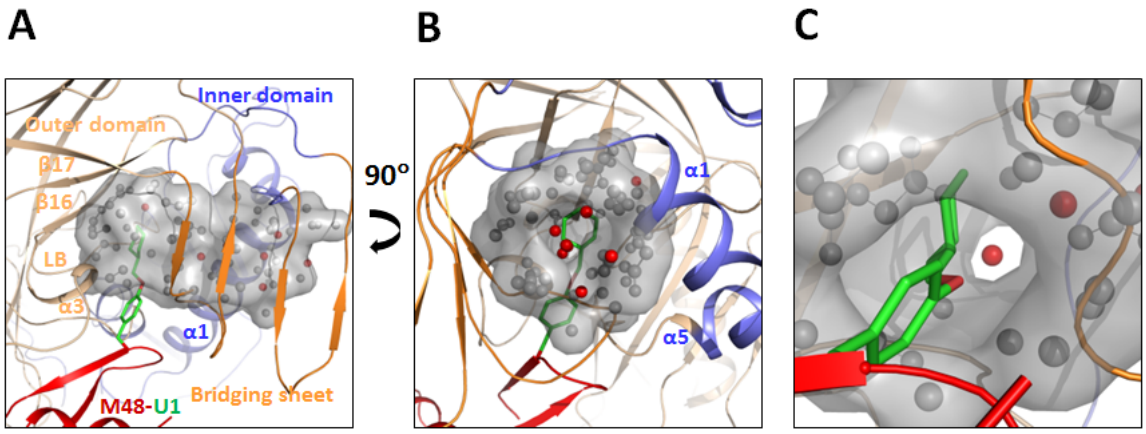


Figure S2, related to Figure 2. Structural variation in HIV-1 gp120, relative to the M48U1-bound gp120 structure, overall and by domain. Each axis represents the coordinate deviation for the core, the inner domain, the outer domain, or the bridging sheet. Lines, drawn in the same color as the names of the molecules, display the RMSDs for those molecules relative to the M48U1-bound gp120 structure.



D

Pocket properties	
Pocket Score	48.1528
Number of Voronoi Vertices	118
Mean alpha-sphere radius	3.5595
Mean alpha-sphere SA	0.4538
Mean B-factor	0.0842
Hydrophobicity Score	37.7714
Polarity Score	22
Volume Score	4.4514
Real volume (approximation)	504.9057 Å^3
Charge Score	3
Local hydrophobic density Score	29.8667
Number of apolar alpha sphere	45
Proportion of apolar alpha sphere	0.3814

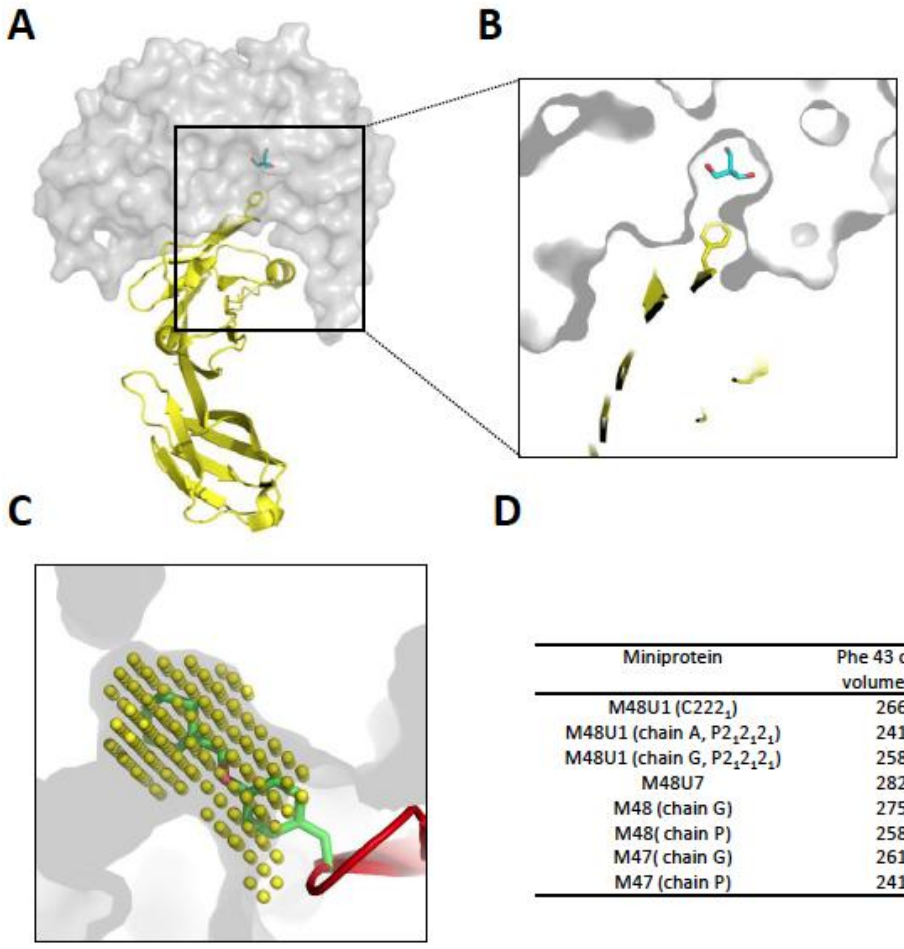
Figure S3, related to Figure 4. M48U1 binding pocket in gp120 detected by fpocket.

(A) The atoms lining the M48U1 binding pocket on gp120 are shown as grey spheres, andtheir van der Waals surfaces are shown in grey surface representation. The pocket was detected by the program f-pocket. The red spheres represent water molecules in this pocket.

(B) A 90° rotated view of the pocket.

(C) A close-up of M48U1 binding to the Phe 43 pocket.

(D) Physical and chemical properties of the pocket as reported by fpocket.



Miniprotein	Phe 43 cavity volume (Å^3)
M48U1 ($C222_1$)	266
M48U1 (chain A, $P2_12_12_1$)	241
M48U1 (chain G, $P2_12_12_1$)	258
M48U7	282
M48 (chain G)	275
M48(chain P)	258
M47(chain G)	261
M47 (chain P)	241

Figure S4, related to Figures 5. Ligand occupation of the Phe43 cavity.

(A) CD4 bound gp120 with buffer molecule occupying Phe43 cavity (PDB ID: 2NXY). gp120 is shown as a transparent grey surface and CD4 is shown in yellow cartoon representation with $Phe43_{CD4}$ shown in stick representation.

(B) Zoomed-in cross-section of gp120 with a Tris molecule (shown in cyan and red sticks) occupying the Phe43 cavity.

(C) Cross-section of gp120 (grey) surface showing the U1 moiety (green and red sticks) bound to the Phe43 cavity. The yellow spheres show the region analyzed for calculating the volume of the cavity.

(D) Phe43 cavity volumes for gp120 molecules bound to different CD4-mimetic miniproteins. Where multiple complexes for the same miniprotein were available, either in different crystal forms or different complexes in the same asymmetric unit, each was analyzed separately.

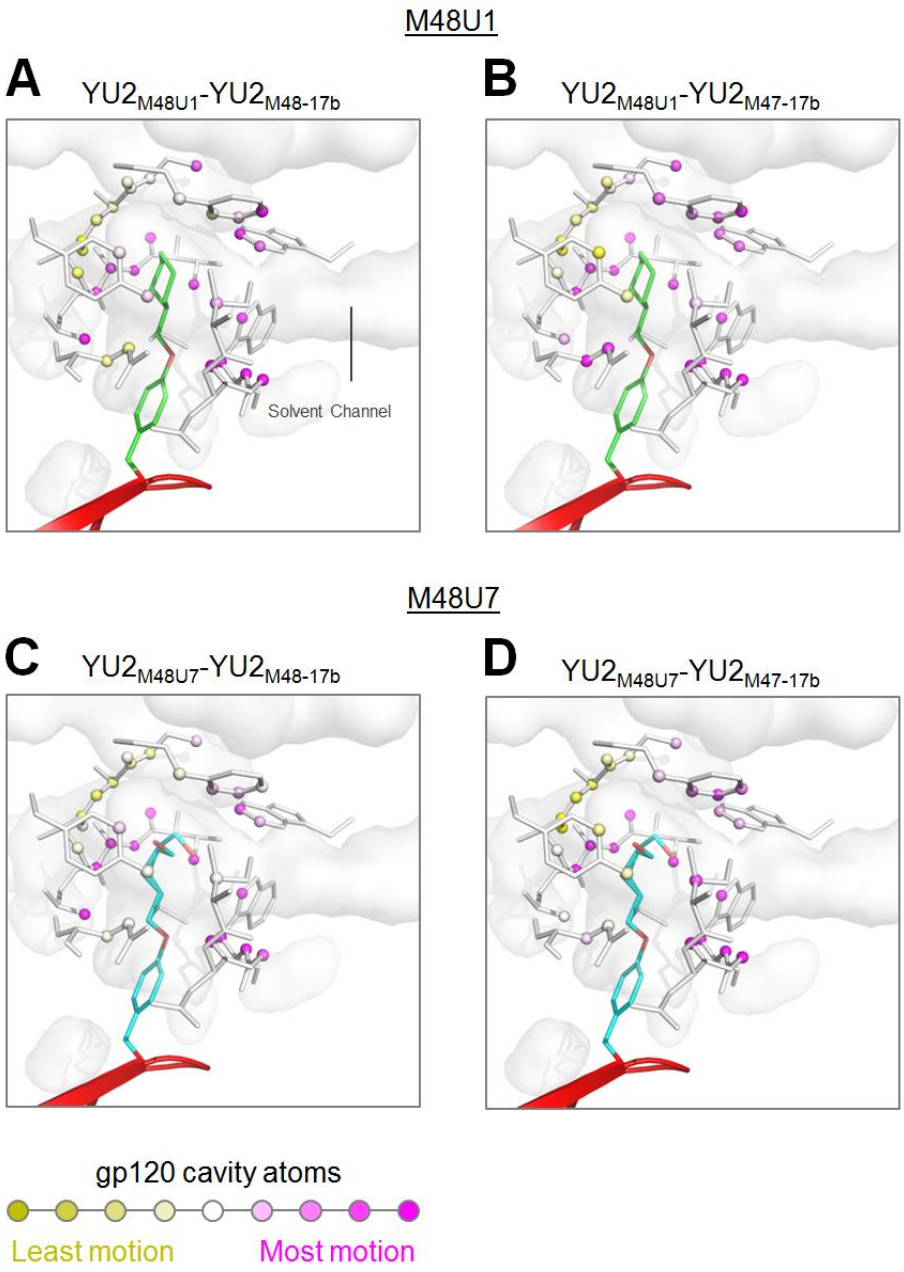


Figure S5, related to Figures 6. Differences in the gp120 Phe43 cavity bound to different CD4-mimetic miniproteins.

(A) gp120 bound to M48U1 (green and red) vs gp120 bound to M48 and 17b IgG.
(B) gp120 bound to M48U1 vs gp120 bound to M47 and 17b IgG.
(C) gp120 bound to M48U7 (cyan and red) vs gp120 bound to M48 and 17b IgG.
(D) gp120 bound to M48U1 vs gp120 bound to M47 and 17b IgG.

gp120 is shown as grey surface, Phe43 cavity atoms are shown as spheres and the residues connecting them as white sticks. The cavity atoms were colored with a yellow-white-magenta gradient, where yellow represents least motion and magenta represents highest differences between the pair of structures.

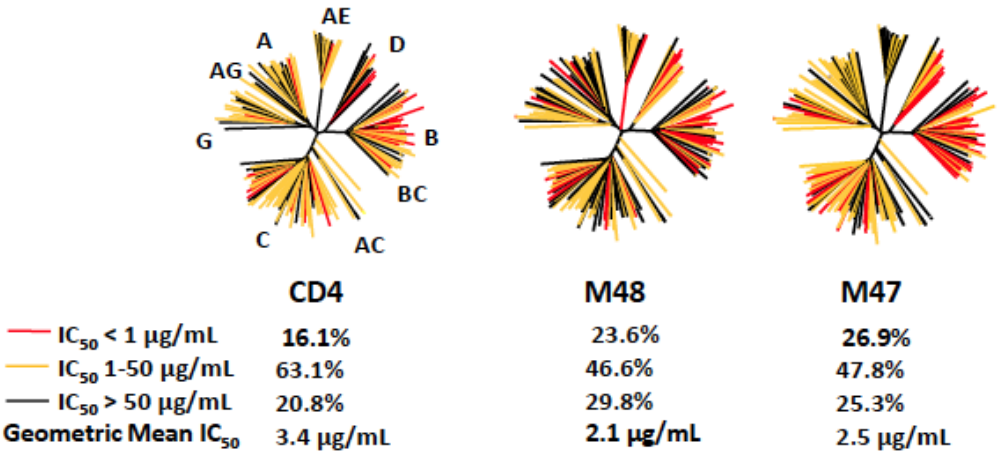


Figure S6, related to Table 1, Table 3 and Figure 7. Neutralization profile of sCD4 and CD4-mimetic miniproteins M48 and M47, depicted with 180-isolate dendrogram representative of circulating HIV-1 Tier 1 and Tier 2 viruses.

Supplemental Experimental Procedures

Preparation and Crystallization of gp120:miniprotein complexes.

The gp120- M48U1 or gp120-M48U7 complexes were formed by mixing deglycosylated YU2 gp120 and the CD4-mimetic miniproteins (1:2 molar ratio) at room temperature and purified by size exclusion chromatography (Hiload 26/60 Superdex S200 prep grade, GE Healthcare) with buffer containing 0.35 M NaCl, 2.5 mM Tris pH 7.0, 0.02% NaN3. Fractions with gp120-miniprotein complexes were concentrated to ~10 mg/ml, flash frozen with liquid nitrogen before storing at -80oC and used for crystallization screening experiments.

Three commercially available screens, Hampton Crystal Screen (Hampton Research), Precipitant Synergy Screen (Emerald BioSystems), and Wizard Screen (Emerald BioSystems), were used for initial crystallization trials of the gp120-miniprotein complexes. Vapor-diffusion sitting drops were set up robotically by mixing 0.1 μL of protein with an equal volume of reservoir solution (Honeybee, DigiLab). Droplets were allowed to equilibrate at 20º C and imaged at scheduled times with RockImager (Formulatrix.). Robotic crystal hits were optimized manually using the hanging drop vapor-diffusion method. Crystals of diffraction-quality for both complexes were obtained with reservoir solution composed of 9.0% PEG 4000, 14.0 % isopropanol, 100 mM sodium citrate, pH 5.6 (Majeed et al., 2003).

X-ray data collection, structure determination and refinement for the gp120:miniprotein complexes.

Diffraction data of the gp120-M48U1 and gp120-M48U7 crystals were collected under cryogenic conditions using 15% 2R, 3R, butanediol as cryoprotectant. X-ray diffraction data were collected at beam-lines ID-22 and BM-22 (SER-CAT) at the Advanced Photon Source, Argonne National Laboratory, with 1.0000 Å radiation, processed and reduced with HKL2000 (Otwinowski, 1997).

The crystal structures of gp120-M48U1 and gp120-M48U7 complexes were solved by molecular replacement using Phaser (McCoy et al., 2007) in the CCP4 Program Suite (Winn et al., 2011). The structure of YU2 gp120 bound to CD4-mimetic miniprotein M48 (PDB ID: 2I6O) (Stricher et al., 2008) was used as search model.

Positive density for the miniprotein insertions were clearly visible in the initial maps. The modified amino acid was constructed in ChemBiodraw, followed by geometry optimization and construction of restraint library for crystallographic refinement using Phenix (phenix.elbow module).

Further refinements were carried out with PHENIX (Adams et al., 2010). Starting with torsion-angle simulated annealing with slow cooling, iterative manual model building was carried out on COOT (Emsley and Cowtan, 2004) with maps generated from combinations of standard positional, individual *B*-factor, TLS refinement algorithms and non-crystallographic symmetry (NCS) restraints. Ordered solvents were added during each macro cycle. Throughout the refinement processes, a cross validation (R_{free}) test set consisting of 5% of the data was used and hydrogens were included as riding model. Structure validations were performed periodically during the model building/refinement process with MolProbity (Davis et al., 2007) and pdb-care (Lutteke and von der Lieth, 2004). X-ray crystallographic data and refinement statistics are summarized in Table 2.

Protein structure analysis and graphical representations.

All superpositions were performed using lsqkab in CCP4 (Winn et al., 2011). PISA (Krissinel and Henrick, 2007) and NCONT (Winn et al., 2011) were used to perform protein-miniprotein interfaces analysis. All graphical representation with protein crystal structures were made with Pymol. Difference distance matrices were produced by distance sorting atom positions and plotting with the program DDMP (Richards and Kundrot, 1988). Fpocket(Hendlich et al., 1997; Le Guilloux et al., 2009) was used for defining the boundaries of the space encompassing the Phe43 cavity and the gp120 solvent channel(Hendlich et al., 1997) (Le Guilloux et al., 2009).

Surface plasmon resonance.

Experiments were carried out on a Biacore 3000 or Biacore T200 instrument (GE Healthcare). gp120 was covalently coupled to a CM5 chip at 1000 RU, and a blank surface with no antigen was created under identical coupling conditions for use as a reference. M48U1 and M48U7 were serially diluted 2-fold, into 10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.05% polysorbate 20 (HBS-EP) and injected over the

immobilized gp120 and reference cell at 50 μL/min. The data were processed with SCRUBBER-2 and double referenced by subtraction of the blank surface and a blank injection (no analyte). Binding curves were globally fit to a 1:1 binding model. In the reverse format experiment, M48U1 was covalently coupled to a CM5 chip at 100 RU, and 2-fold serially diluted gp120 in 10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.05% polysorbate 20 (HBS-EP) buffer

Assessment of HIV-1 neutralization.

Neutralization was measured using single-round-of-infection HIV-1 Env-pseudoviruses and TZM-bl target cells, as described previously(Li et al., 2009; Seaman et al., 2010; Wu et al., 2009). Neutralization curves were fit by nonlinear regression using a 5-parameter hill slope equation as previously described (Seaman et al., 2010). The 50% inhibitory concentrations (IC_{50}) were reported as the inhibitor concentrations required to inhibit infection by 50%.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica Section D, Biological crystallography *66*, 213-221.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S.*, et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res *35*, W375-383.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Hendlich, M., Rippmann, F., and Barnickel, G. (1997). LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. J Mol Graph Model *15*, 359-363, 389.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol *372*, 774-797.

Le Guilloux, V., Schmidtke, P., and Tuffery, P. (2009). Fpocket: an open source platform for ligand pocket detection. BMC Bioinformatics *10*, 168.

Li, Y., Svehla, K., Louder, M.K., Wycuff, D., Phogat, S., Tang, M., Migueles, S.A., Wu, X., Phogat, A., Shaw, G.M.*, et al.* (2009). Analysis of neutralization specificities in polyclonal sera derived from human immunodeficiency virus type 1-infected individuals. Journal of virology *83*, 1045-1059.

Lutteke, T., and von der Lieth, C.W. (2004). pdb-care (PDB carbohydrate residue check): a program to support annotation of complex carbohydrate structures in PDB files. BMC Bioinformatics *5*, 69.

Majeed, S., Ofek, G., Belachew, A., Huang, C.C., Zhou, T., and Kwong, P.D. (2003). Enhancing protein crystallization through precipitant synergy. Structure *11*, 1061-1070.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. Journal of applied crystallography *40*, 658-674.

Otwinowski, Z.a.M., W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods Enzymol *276*, 307-326.

Richards, F.M., and Kundrot, C.E. (1988). Identification of structural motifs from protein coordinate data: secondary structure and first-level supersecondary structure. Proteins *3*, 71-84.

Seaman, M.S., Janes, H., Hawkins, N., Grandpre, L.E., Devoy, C., Giri, A., Coffey, R.T., Harris, L., Wood, B., Daniels, M.G.*, et al.* (2010). Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. J Virol *84*, 1439-1452.

Stricher, F., Huang, C.C., Descours, A., Duquesnoy, S., Combes, O., Decker, J.M., Do Kwon, Y., Lusso, P., Shaw, G.M., Vita, C.*, et al.* (2008). Combinatorial optimization of a CD4-mimetic miniprotein and cocrystal structures with HIV-1 gp120 envelope glycoprotein. Journal of Molecular Biology *382*, 510-524.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta crystallographica Section D, Biological crystallography *67*, 235-242.

Wu, X., Zhou, T., O'Dell, S., Wyatt, R.T., Kwong, P.D., and Mascola, J.R. (2009). Mechanism of human immunodeficiency virus type 1 resistance to monoclonal antibody B12 that effectively targets the site of CD4 attachment. Journal of virology *83*, 10892-10907.