Supporting Information

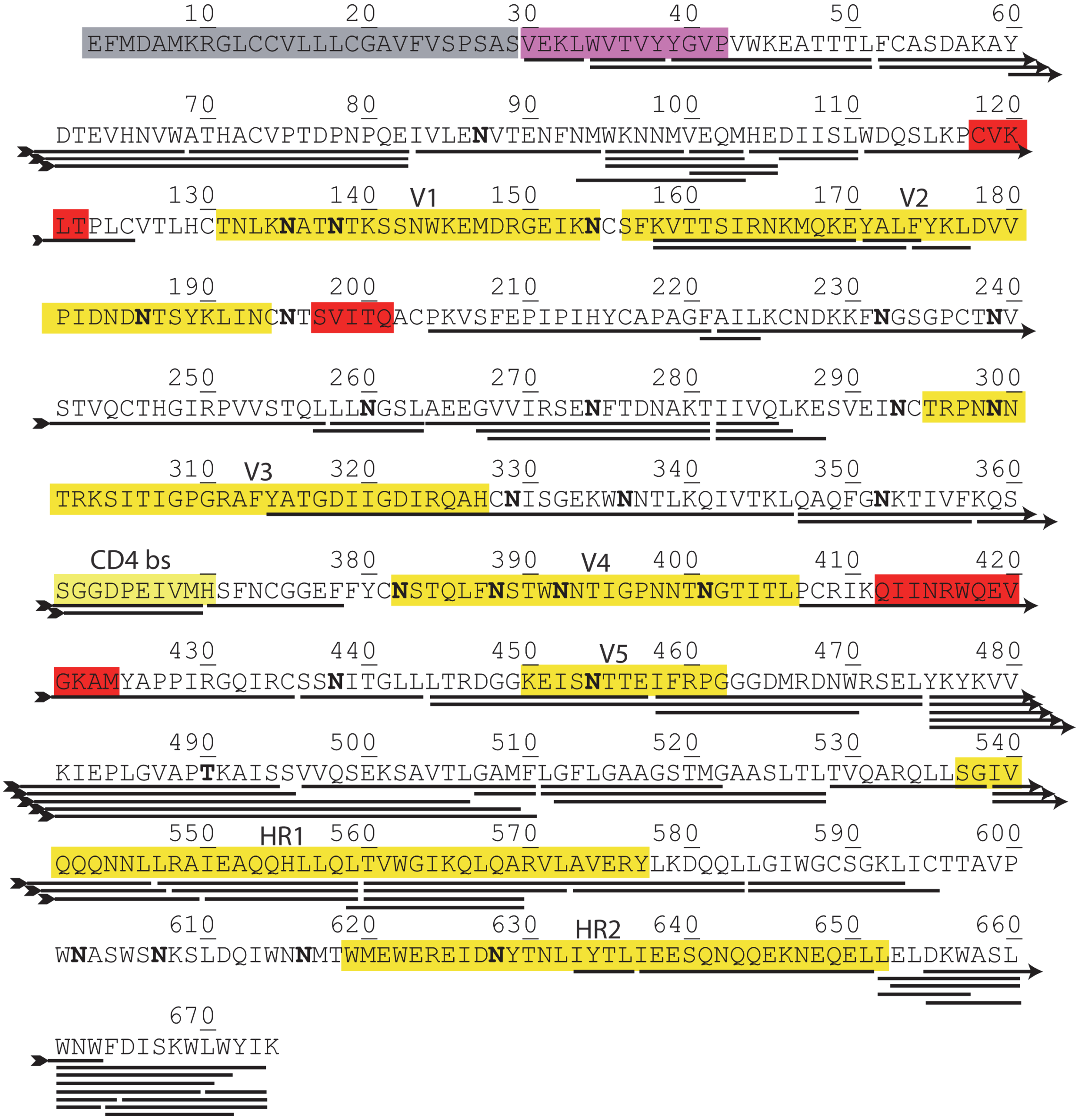
A functional gp41-gp120 interaction is observed in monomeric, but not trimeric, uncleaved HIV Env gp140

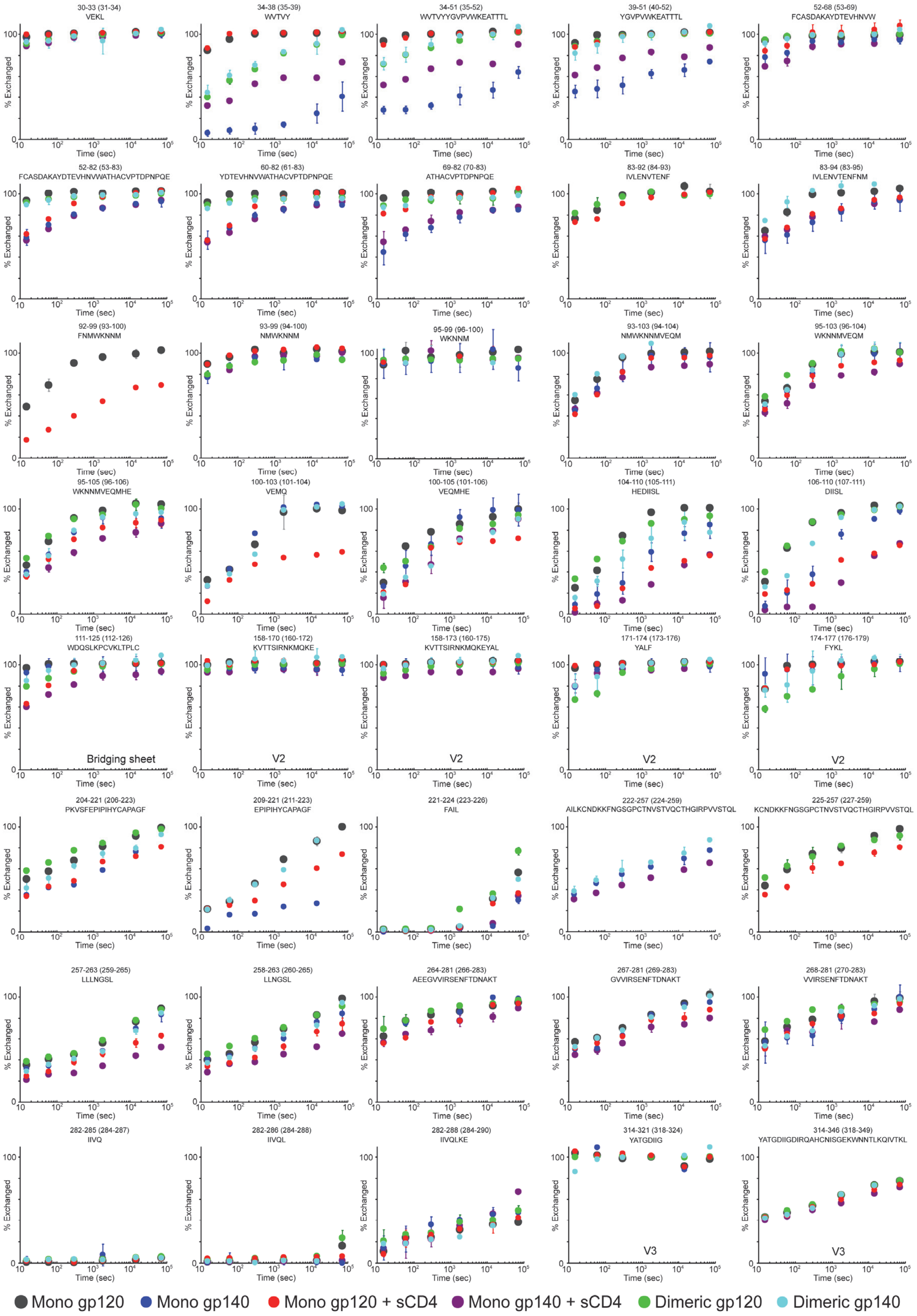
Miklos Guttman[1], Kelly K. Lee[1#]

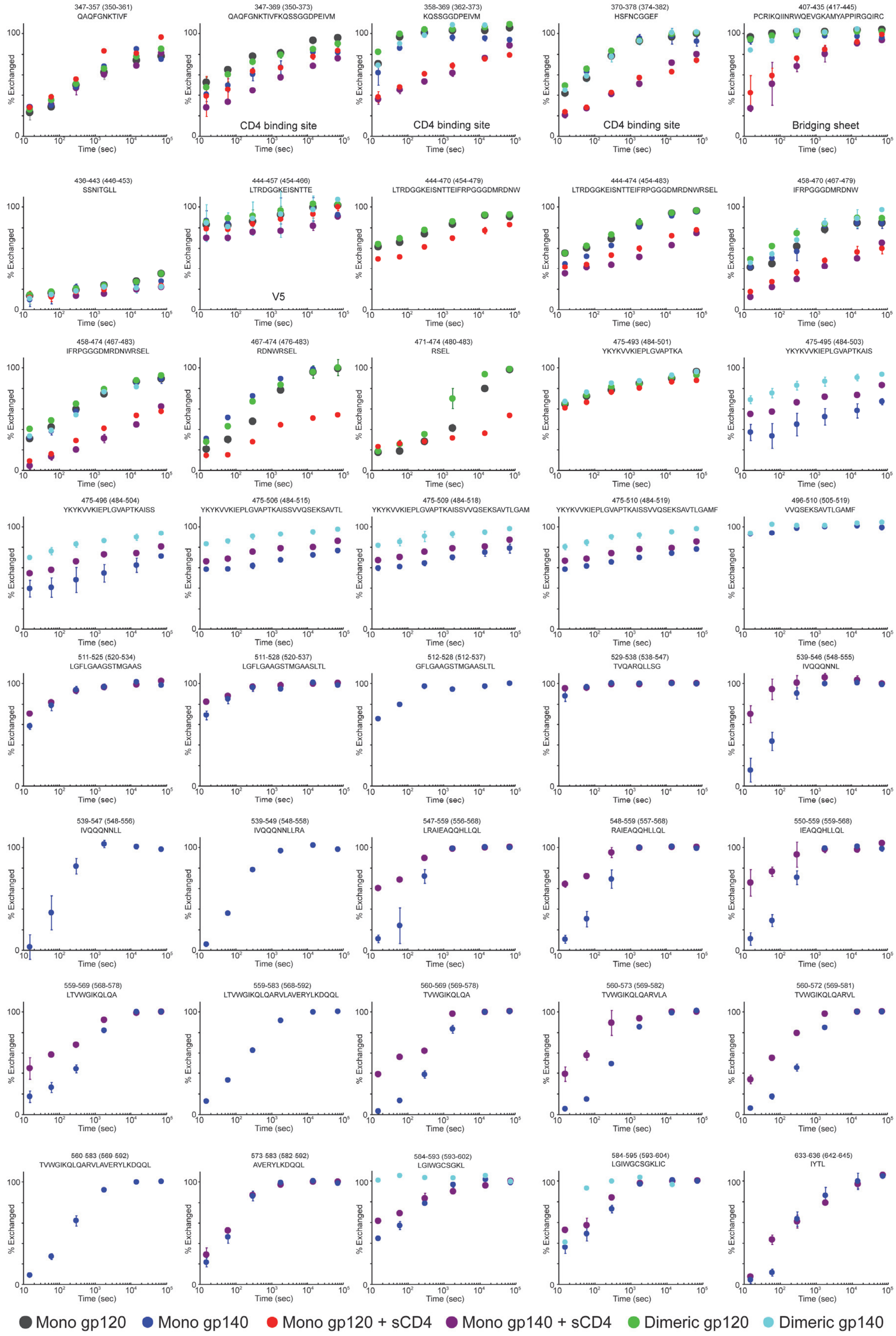
[1] Department of Medicinal Chemistry; University of Washington, Seattle, WA 98195

[#] Address correspondence to: kklee@uw.edu; 206-616-3496; 206-685-3252 (fax)

List of Figures:


10
20
30
40
50
60
EFMDAMKRGLCCVLLLCGAVFVSPSASVEKLWVTVYYGVPVWKEATTTLFCASDAKAY
70
80
90
100
110
120
DTEVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVK
130
140
V1
150
160
170
V2
180
LTPLCVTLHCTNLKNATNTKSSNWKEMDRGEIKNCSFKVTTSIRNKMQKEYALFYKLDVV
190
200
210
220
230
240
PIDNDNTSYKLINCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGSGPCTNV
250
260
270
280
290
300
STVQCTHGIRPVVSTQLLLNGSLAEEGVVIRSENFTDNAKTIIVQLKESVEINCTRPNNN
310
V3
320
330
340
350
360
TRKSITIGPGRAFYATGDIIGDIRQAHCNISGEKWNNTLKQIVTKLQAQFGNKTIVFKQS
CD4 bs
380
390
V4
400
410
420
SGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNNTIGPNNTNGTITLPCRIKQIINRWQEV
430
440
450
V5
460
470
480
GKAMYAPPIRGQIRCSSNITGLLLTRDGGKEISNTTEIFRPGGGDMRDNWRSELYKYKVV
490
500
510
520
530
540
KIEPLGVAPTKAISSVVQSEKSAVTLGAMFLGFLGAAGSTMGAASLTLTVQARQLLSGIV
550
HR1
560
570
580
590
600
QQQNNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTAVP
610
620
630
HR2
640
650
660
WNASWSNKSLDQIWNNMTWMEWEREIDNYTNLIYTLIEESQNQQEKNEQELLELDKWASL
670
WNWFDISKWLWYIK


30-33 (31-34)
VEKL
34-38 (35-39)
WVTVY
34-51 (35-52)
WVTVYYGVPVWKEATTTL
39-51 (40-52)
YGVPVWKEATTTL
52-68 (53-69)
FCASDAKAYDTEVHNVW
52-82 (53-83)
FCASDAKAYDTEVHNVWATHACVPTDPNPQE
60-82 (61-83)
YDTEVHNVWATHACVPTDPNPQE
69-82 (70-83)
ATHACVPTDPNPQE
83-92 (84-93)
IVLENVTENF
83-94 (83-95)
IVLENVTENFNM
92-99 (93-100)
FNMWKNNM
93-99 (94-100)
NMWKNNM
95-99 (96-100)
WKNNM
93-103 (94-104)
NMWKNNMVEQM
95-103 (96-104)
WKNNMVEQM
95-105 (96-106)
WKNNMVEQMHE
100-103 (101-104)
VEMQ
100-105 (101-106)
VEQMHE
104-110 (105-111)
HEDIISL
106-110 (107-111)
DIISL
111-125 (112-126)
WDQSLKPCVKLTPLC
158-170 (160-172)
KVTTSIRNKMQKE
158-173 (160-175)
KVTTSIRNKMQKEYAL
171-174 (173-176)
YALF
174-177 (176-179)
FYKL
Bridging sheet
V2
V2
V2
V2
204-221 (206-223)
PKVSFEPIPIHYCAPAGF
209-221 (211-223)
EPIPIHYCAPAGF
221-224 (223-226)
FAIL
222-257 (224-259)
AILKCNDKKFNGSGPCTNVSTVQCTHGIRPVVSTQL
225-257 (227-259)
KCNDKKFNGSGPCTNVSTVQCTHGIRPVVSTQL
257-263 (259-265)
LLLNGSL
258-263 (260-265)
LLNGSL
264-281 (266-283)
AEEGVVIRSENFTDNAKT
267-281 (269-283)
GVVIRSENFTDNAKT
268-281 (270-283)
VVIRSENFTDNAKT
282-285 (284-287)
IIVQ
282-286 (284-288)
IIVQL
282-288 (284-290)
IIVQLKE
314-321 (318-324)
YATGDIIG
314-346 (318-349)
YATGDIIGDIRQAHCNISGEKWNNTLKQIVTKL
V3
V3
% Exchanged
Time (sec)
0
100
10
10^2
10^3
10^4
10^5
● Mono gp120
● Mono gp140
● Mono gp120 + sCD4
● Mono gp140 + sCD4
● Dimeric gp120
● Dimeric gp140


347-357 (350-361)
QAQFGNKTIVF
347-369 (350-373)
QAQFGNKTIVFKQSSGGDPEIVM
CD4 binding site
358-369 (362-373)
KQSSGGDPEIVM
CD4 binding site
370-378 (374-382)
HSFNCGGEF
CD4 binding site
407-435 (417-445)
PCRIKQIINRWQEVGKAMYAPPIRGQIRC
Bridging sheet
436-443 (446-453)
SSNITGLL
444-457 (454-466)
LTRDGGKEISNTTE
V5
444-470 (454-479)
LTRDGGKEISNTTEIFRPGGGDMRDNW
444-474 (454-483)
LTRDGGKEISNTTEIFRPGGGDMRDNWRSEL
458-470 (467-479)
IFRPGGGDMRDNW
458-474 (467-483)
IFRPGGGDMRDNWRSEL
467-474 (476-483)
RDNWRSEL
471-474 (480-483)
RSEL
475-493 (484-501)
YKYKVVKIEPLGVAPTKA
475-495 (484-503)
YKYKVVKIEPLGVAPTKAIS
475-496 (484-504)
YKYKVVKIEPLGVAPTKAISS
475-506 (484-515)
YKYKVVKIEPLGVAPTKAISSVVQSEKSAVTL
475-509 (484-518)
YKYKVVKIEPLGVAPTKAISSVVQSEKSAVTLGAM
475-510 (484-519)
YKYKVVKIEPLGVAPTKAISSVVQSEKSAVTLGAMF
496-510 (505-519)
VVQSEKSAVTLGAMF
511-525 (520-534)
LGFLGAAGSTMGAAS
511-528 (520-537)
LGFLGAAGSTMGAASLTL
512-528 (512-537)
GFLGAAGSTMGAASLTL
529-538 (538-547)
TVQARQLLSG
539-546 (548-555)
IVQQQNNL
539-547 (548-556)
IVQQQNNLL
539-549 (548-558)
IVQQQNNLLRA
547-559 (556-568)
LRAIEAQQHLLQL
548-559 (557-568)
RAIEAQQHLLQL
550-559 (559-568)
IEAQQHLLQL
559-569 (568-578)
LTVWGIKQLQA
559-583 (568-592)
LTVWGIKQLQARVLAVERYLKDQQL
560-569 (569-578)
TVWGIKQLQA
560-573 (569-582)
TVWGIKQLQARVLA
560-572 (569-581)
TVWGIKQLQARVL
560-583 (569-592)
TVWGIKQLQARVLAVERYLKDQQL
573-583 (582-592)
AVERYLKDQQL
584-593 (593-602)
LGIWGCSGKL
584-595 (593-604)
LGIWGCSGKLIC
633-636 (642-645)
IYTL
% Exchanged
0
100
Time (sec)
10
10^2
10^3
10^4
10^5
● Mono gp120
● Mono gp140
● Mono gp120 + sCD4
● Mono gp140 + sCD4
● Dimeric gp120
● Dimeric gp140

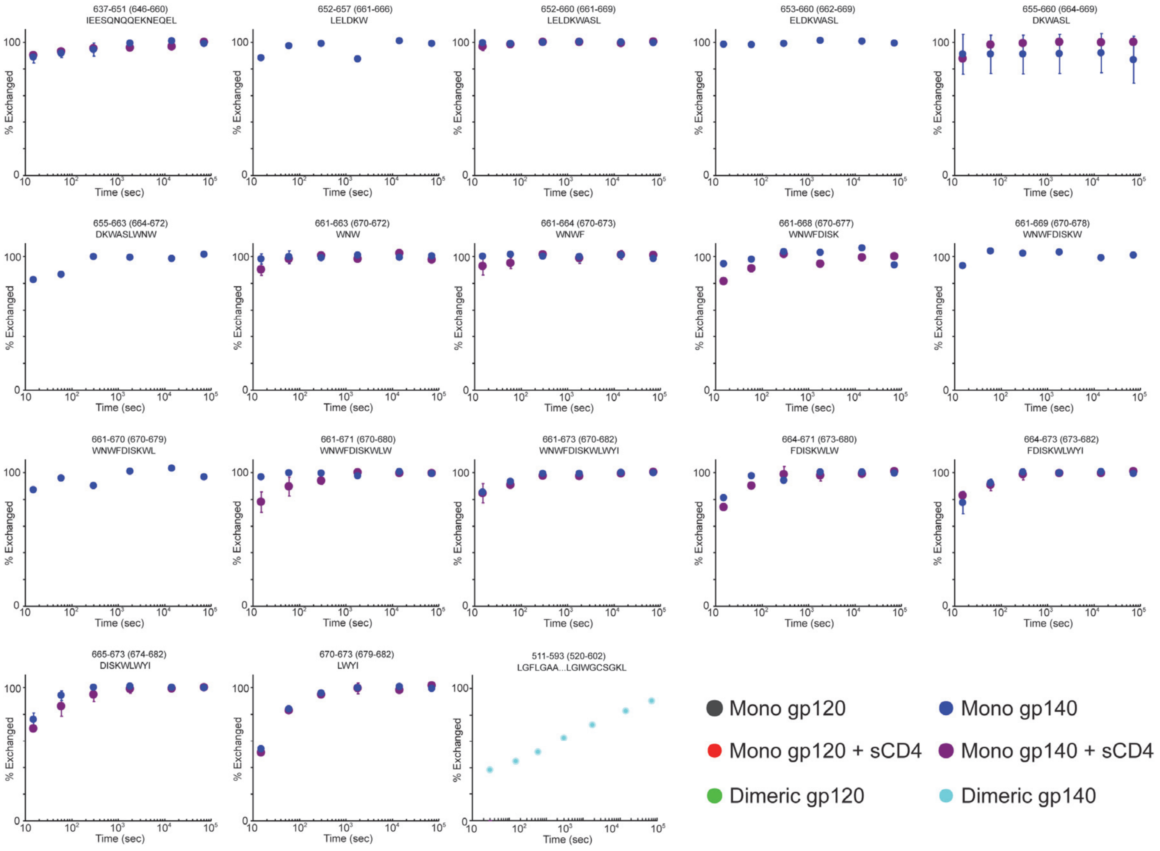
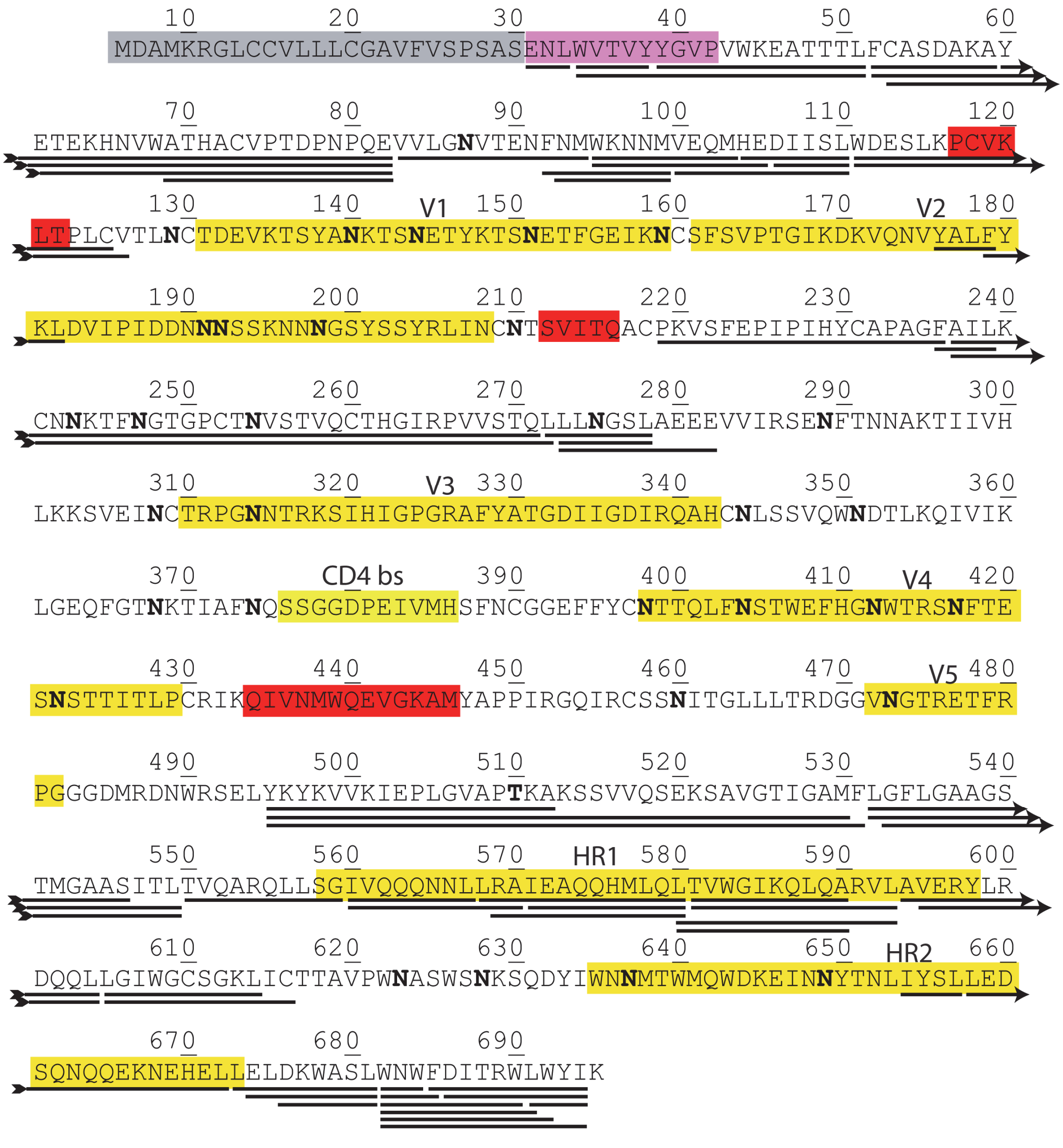


Figure S1: Peptic coverage of gp140 isolate SF162 (top). Observable fragments are shown as lines under the primary sequence. The leader sequence (absent from the mature protein) is shown in grey. Variable loops 1-5, the CD4 binding site, and HR1/HR2 are highlighted and potential N-linked glycosylation sites are indicated in bold. Segments of the bridging sheet are highlighted in red, and the N-terminal extension in pink. Percent exchange is shown for each peptide in monomeric gp120 (grey), monomeric gp120+sCD4 (red), dimeric gp120 (green), monomeric gp140 (blue), monomeric gp140+sCD4 (purple), and dimeric gp140 (cyan) as a function of time. Error bars represent standard deviations from duplicate measurements and as well as multiple observed peptide charge states. Numbering in parenthesis refers to HXB2 numbering.



MDAMKRGLCCVLLLCGAVFVSPSASENLWVTVYYGVPVWKEATTTLFCASDAKAY

ETEKHNVWATHACVPTDPNPQEVVLG**N**VTENFNMWKNNMVEQMHEDIISLWDESLKPCVK

LTPLCVTL**N**CTDEVKTSYA**N**KTS**N**ETYKTS**N**ETFGEIK**N**CSFSVPTGIKDKVQNVYALFY

KLDVIPIDDN**NN**SSKNN**N**GSYSSYRLINC**N**TSVITQACPKVSFEPIPIHYCAPAGFAILK

CN**N**KTF**N**GTGPCT**N**VSTVQCTHGIRPVVSTQLLL**N**GSLAEEEVVIRSE**N**FTNNAKTIIVH

LKKSVEI**N**CTRPG**N**NTRKSIHIGPGRAFYATGDIIGDIRQAHC**N**LSSVQW**N**DTLKQIVIK

LGEQFGT**N**KTIAF**N**QSSGGDPEIVMHSFNCGGEFFYC**N**TTQLF**N**STWEFHG**N**WTRS**N**FTE

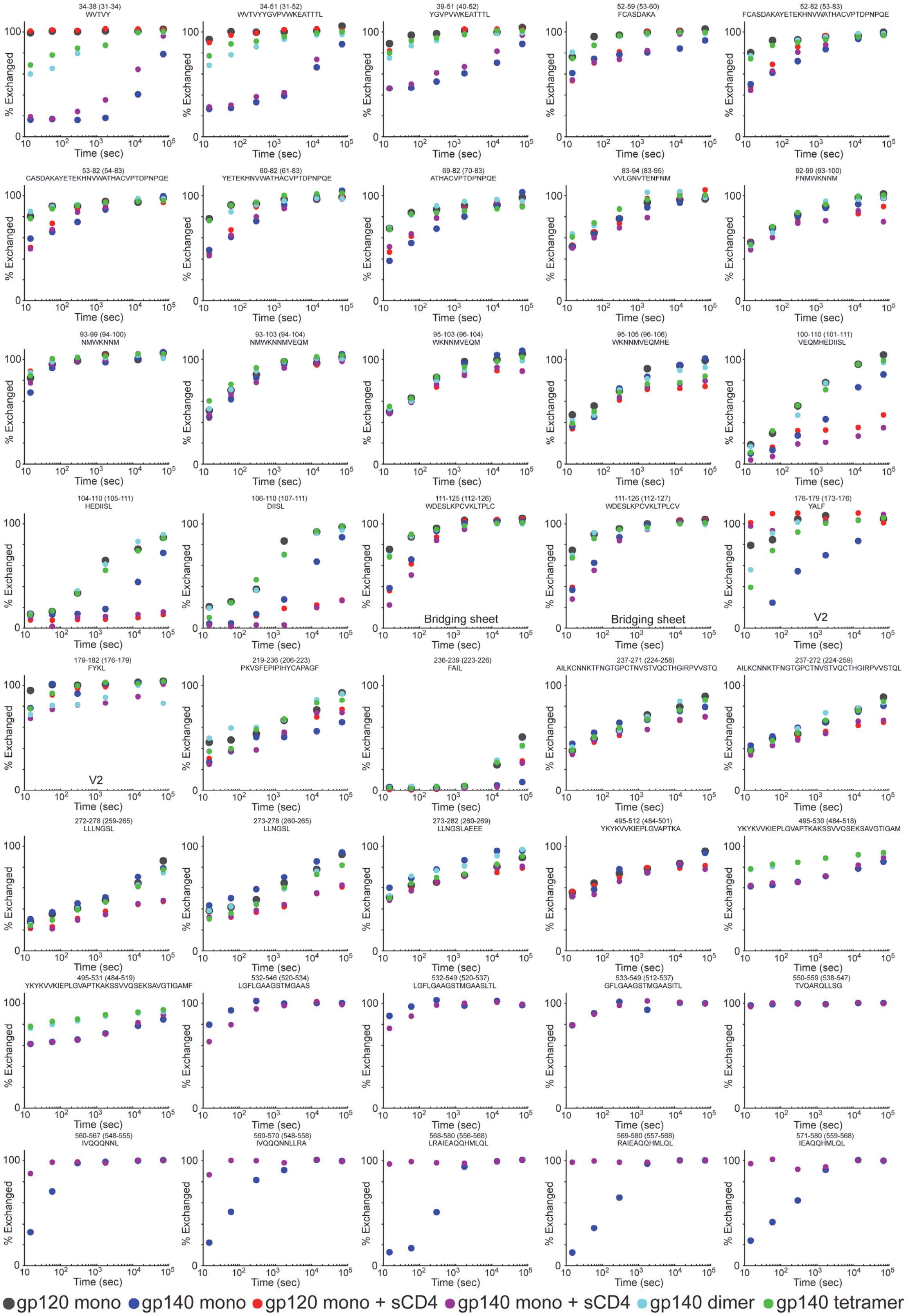
S**N**STTITLPCRIKQIVNMWQEVGKAMYAPPIRGQIRCSS**N**ITGLLLTRDGGV**N**GTRETFR

PGGGDMRDNWRSELYKYKVVKIEPLGVAP**T**KAKSSVVQSEKSAVGTIGAMFLGFLGAAGS

TMGAASITLTVQARQLLSGIVQQQNNLLRAIEAQQHMLQLTVWGIKQLQARVLAVERYLR

DQQLLGIWGCSGKLICTTAVPW**N**ASWS**N**KSQDYIWN**N**MTWMQWDKEI**N**NYTNLIYSLLED

SQNQQEKNEHELLELDKWASLWNWFDITRWLWYIK


34-38 (31-34)
WVTVY
34-51 (31-52)
WVTVYYGVPVWKEATTTL
39-51 (40-52)
YGVPVWKEATTTL
52-59 (53-60)
FCASDAKA
52-82 (53-83)
FCASDAKAYETEKHNVWATHACVPTDPNPQE
53-82 (54-83)
CASDAKAYETEKHNVWATHACVPTDPNPQE
60-82 (61-83)
YETEKHNVWATHACVPTDPNPQE
69-82 (70-83)
ATHACVPTDPNPQE
83-94 (83-95)
VVLGNVTENFNM
92-99 (93-100)
FNMWKNNM
93-99 (94-100)
NMWKNNM
93-103 (94-104)
NMWKNNMVEQM
95-103 (96-104)
WKNNMVEQM
95-105 (96-106)
WKNNMVEQMHE
100-110 (101-111)
VEQMHEDIISL
104-110 (105-111)
HEDIISL
106-110 (107-111)
DIISL
111-125 (112-126)
WDESLKPCVKLTPLC
Bridging sheet
111-126 (112-127)
WDESLKPCVKLTPLCV
Bridging sheet
176-179 (173-176)
YALF
V2
179-182 (176-179)
FYKL
V2
219-236 (206-223)
PKVSFEPIPIHYCAPAGF
236-239 (223-226)
FAIL
237-271 (224-258)
AILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQ
237-272 (224-259)
AILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQL
272-278 (259-265)
LLLNGSL
273-278 (260-265)
LLNGSL
273-282 (260-269)
LLNGSLAEEE
495-512 (484-501)
YKYKVVKIEPLGVAPTKA
495-530 (484-518)
YKYKVVKIEPLGVAPTKAKSSVVQSEKSAVGTIGAM
495-531 (484-519)
YKYKVVKIEPLGVAPTKAKSSVVQSEKSAVGTIGAMF
532-546 (520-534)
LGFLGAAGSTMGAAS
532-549 (520-537)
LGFLGAAGSTMGAASLTL
533-549 (512-537)
GFLGAAGSTMGAASITL
550-559 (538-547)
TVQARQLLSG
560-567 (548-555)
IVQQQNNL
560-570 (548-558)
IVQQQNNLLRA
568-580 (556-568)
LRAIEAQQHMLQL
569-580 (557-568)
RAIEAQQHMLQL
571-580 (559-568)
IEAQQHMLQL
% Exchanged
100
0
Time (sec)
10
10^2
10^3
10^4
10^5
● gp120 mono
● gp140 mono
● gp120 mono + sCD4
● gp140 mono + sCD4
● gp140 dimer
● gp140 tetramer

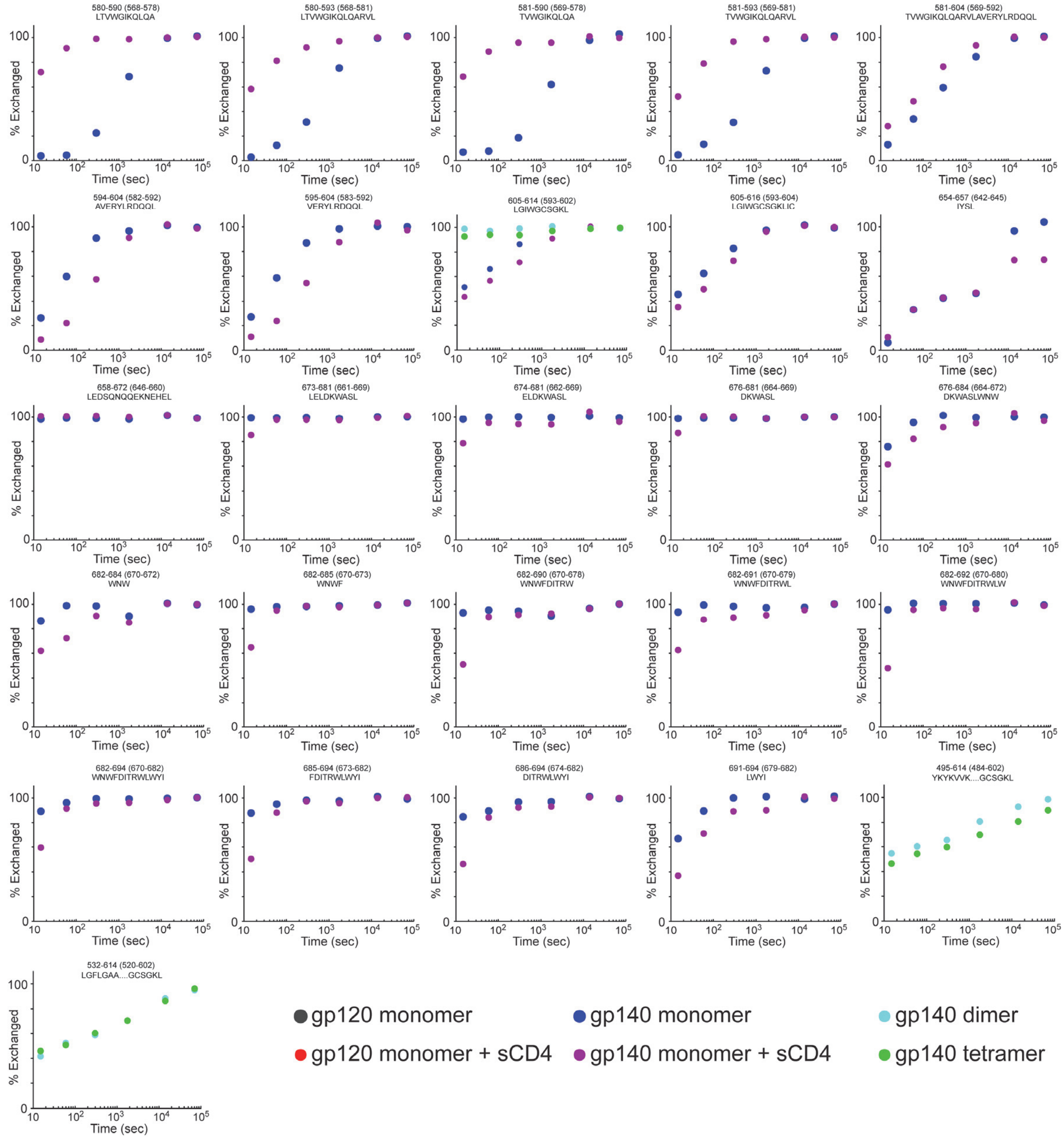
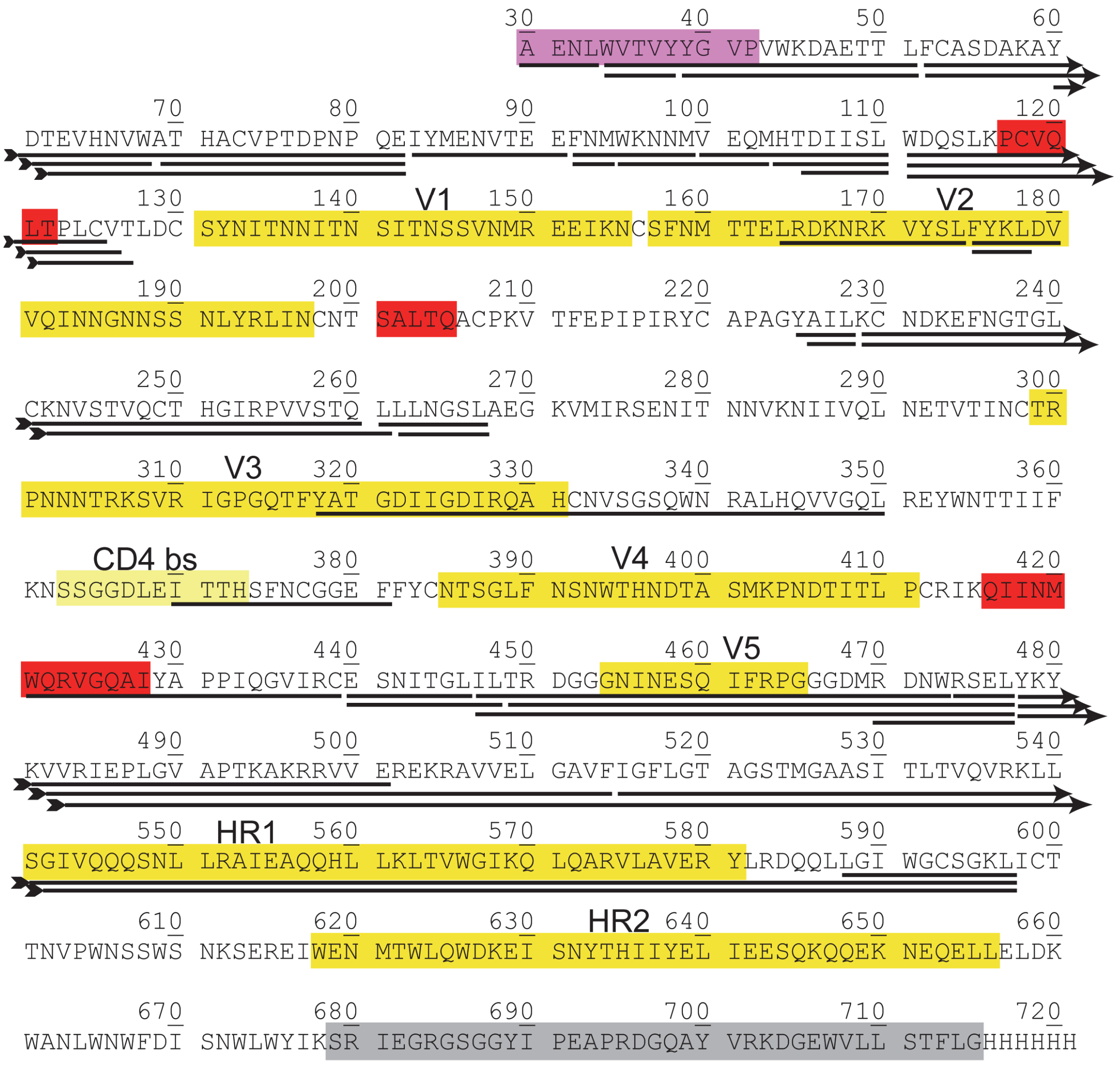
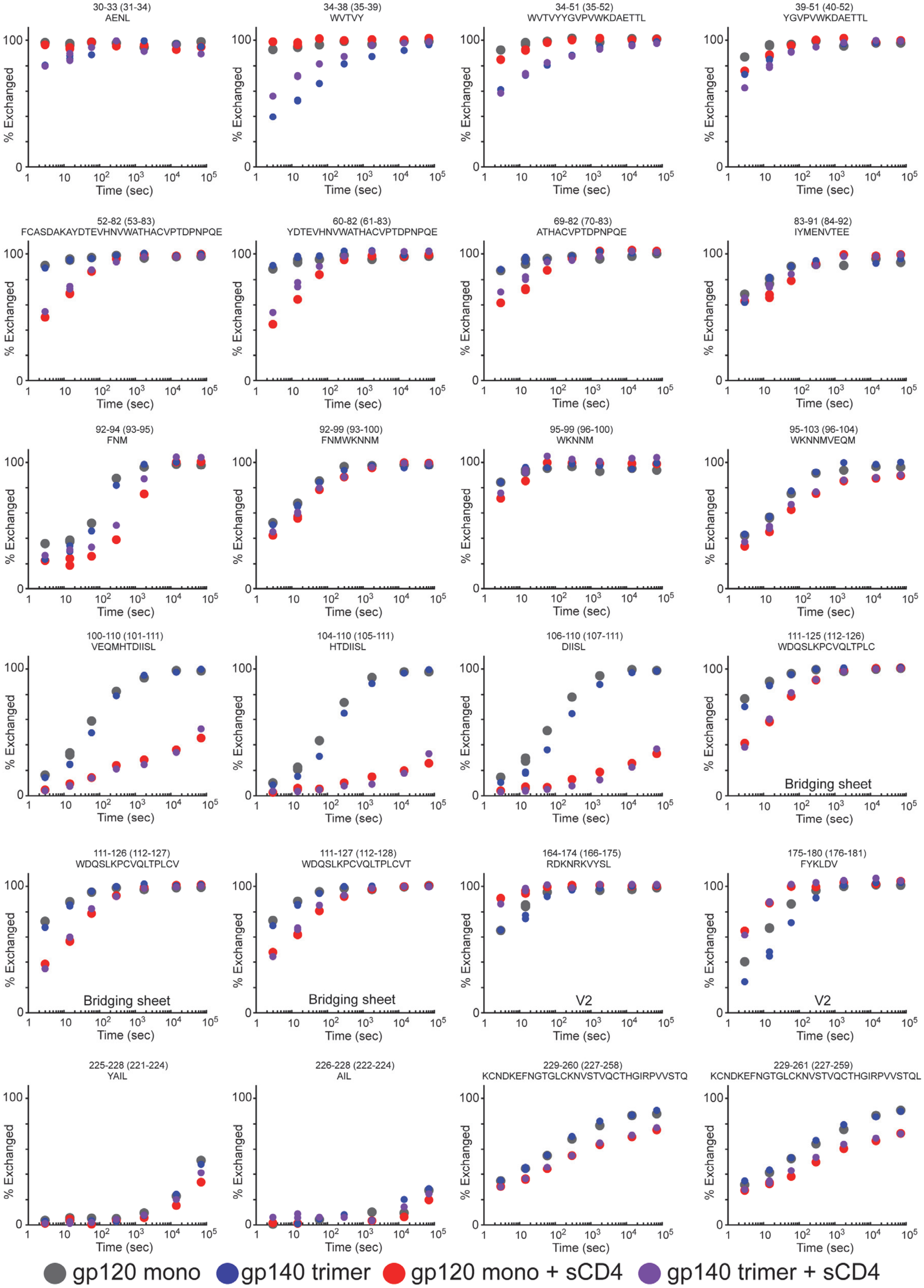


Figure S2: Peptic coverage of gp140 isolate QH0692 (top). Observable fragments are shown as lines under the primary sequence. Segments of the bridging sheet are highlighted in red, and the N-terminal extension in pink. Percent exchange is shown for each peptide in monomeric gp120 (grey), monomeric gp120+sCD4 (red), monomeric gp140 (blue), monomeric gp140+sCD4 (purple), dimeric gp140 (cyan), and tetrameric gp140 (green) as a function of time. Numbering in parenthesis refers to HXB2 numbering.


30 40 50 60
A ENLWVTVYYG VPVWKDAETT LFCASDAKAY
70 80 90 100 110 120
DTEVHNVWAT HACVPTDPNP QEIYMENVTE EFNMWKNNMV EQMHTDIISL WDQSLKPCVQ
130 140 V1 150 160 170 V2 180
LTPLCVTLDC SYNITNNITN SITNSSVNMR EEIKNCSFNM TTELRDKNRK VYSLFYKLDV
190 200 210 220 230 240
VQINNGNNSS NLYRLINCNT SALTQACPKV TFEPIPIRYC APAGYAILKC NDKEFNGTGL
250 260 270 280 290 300
CKNVSTVQCT HGIRPVVSTQ LLLNGSLAEG KVMIRSENIT NNVKNIIVQL NETVTINCTR
310 V3 320 330 340 350 360
PNNNTRKSVR IGPGQTFYAT GDIIGDIRQA HCNVSGSQWN RALHQVVGQL REYWNTTIIF
CD4 bs 380 390 V4 400 410 420
KNSSGGDLEI TTHSFNCGGE FFYCNTSGLF NSNWTHNDTA SMKPNDTITL PCRIKQIINM
430 440 450 460 V5 470 480
WQRVGQAIYA PPIQGVIRCE SNITGLILTR DGGGNINESQ IFRPGGGDMR DNWRSELYKY
490 500 510 520 530 540
KVVRIEPLGV APTKAKRRVV EREKRAVVEL GAVFIGFLGT AGSTMGAASI TLTVQVRKLL
550 HR1 560 570 580 590 600
SGIVQQQSNL LRAIEAQQHL LKLTVWGIKQ LQARVLAVER YLRDQQLLGI WGCSGKLICT
610 620 630 HR2 640 650 660
TNVPWNSSWS NKSEREIWEN MTWLQWDKEI SNYTHIIYEL IEESQKQQEK NEQELLELDK
670 680 690 700 710 720
WANLWNWFDI SNWLWYIKSR IEGRGSGGYI PEAPRDGQAY VRKDGEWVLL STFLGHHHHHH


30-33 (31-34)
AENL
34-38 (35-39)
WVTVY
34-51 (35-52)
WVTVYYGVPVWKDAETTL
39-51 (40-52)
YGVPVWKDAETTL
52-82 (53-83)
FCASDAKAYDTEVHNVWATHACVPTDPNPQE
60-82 (61-83)
YDTEVHNVWATHACVPTDPNPQE
69-82 (70-83)
ATHACVPTDPNPQE
83-91 (84-92)
IYMENVTEE
92-94 (93-95)
FNM
92-99 (93-100)
FNMWKNNM
95-99 (96-100)
WKNNM
95-103 (96-104)
WKNNMVEQM
100-110 (101-111)
VEQMHTDIISL
104-110 (105-111)
HTDIISL
106-110 (107-111)
DIISL
111-125 (112-126)
WDQSLKPCVQLTPLC
Bridging sheet
111-126 (112-127)
WDQSLKPCVQLTPLCV
Bridging sheet
111-127 (112-128)
WDQSLKPCVQLTPLCVT
Bridging sheet
164-174 (166-175)
RDKNRKVYSL
V2
175-180 (176-181)
FYKLDV
V2
225-228 (221-224)
YAIL
226-228 (222-224)
AIL
229-260 (227-258)
KCNDKEFNGTGLCKNVSTVQCTHGIRPVVSTQ
229-261 (227-259)
KCNDKEFNGTGLCKNVSTVQCTHGIRPVVSTQL
% Exchanged
0
100
Time (sec)
1
10
10^2
10^3
10^4
10^5
● gp120 mono
● gp140 trimer
● gp120 mono + sCD4
● gp140 trimer + sCD4

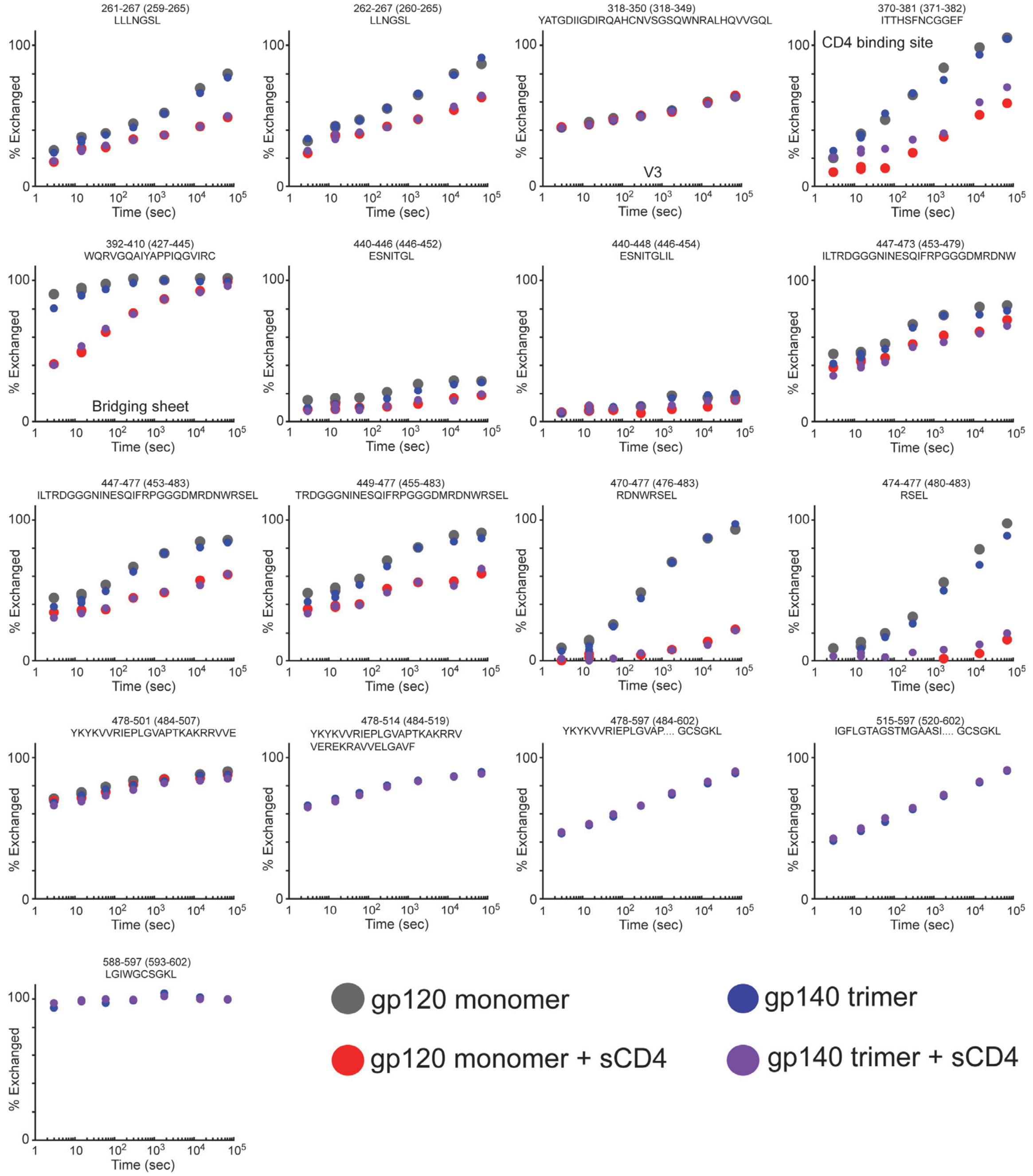


Figure S3. Peptic coverage of gp140 isolate 92UG37 (top). Observable fragments are shown as lines under the primary sequence. Highlighting is described in Figure S1, with the T4 bacteriophage fibritin trimerization motif at the C-terminal end in grey. Percent exchange is shown for each peptide in monomeric gp120 (grey), monomeric gp120+sCD4 (red), trimeric gp140 (blue), and trimeric gp140+sCD4 (purple) as a function of time. Numbering in parenthesis refers to HXB2 numbering.