Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry 15 (2007) 4220–4228
Bioorganic & Medicinal Chemistry

Synthesis and anti-HIV activity of trivalent CD4-mimetic miniproteins

Hengguang Li,[a] Yongjun Guan,[a] Agnieszka Szczepanska,[b] Antonio J. Moreno-Vargas,[b] Ana T. Carmona,[b] Inmaculada Robina,[b] George K. Lewis[a] and Lai-Xi Wang[a,*]

[a]*Institute of Human Virology, University of Maryland, Baltimore, MD 21201, USA*
[b]*Department of Organic Chemistry, University of Seville, PO Box 1203, E-41071-Seville, Spain*

Received 5 February 2007; revised 16 March 2007; accepted 20 March 2007
Available online 24 March 2007

Abstract—A series of trivalent CD4-mimetic miniproteins was synthesized, in which three CD4M9 miniprotein moieties were tethered on a threefold-symmetric scaffold. The trivalent miniproteins were designed to target the CD4-binding sites displayed in the trimeric gp120 complex of HIV-1. The synthesis takes advantage of the highly efficient ligation between a cysteine-tagged CD4M9 miniprotein and a suitable trivalent maleimide that varied in the nature and length of spacer. Antiviral assay revealed that most of the synthetic trivalent miniproteins demonstrated significantly enhanced anti-HIV activities over the monomeric CD4M9 against both R5- and X4-tropic viruses, indicating the beneficial multivalent effects. One compound that possesses a hydrophobic linker was shown to be 140-fold more active than CD4M9 against HIV-1_{Bal} infection, implicating a positive contribution of the lipid portion to the antiviral activity. It was also found that most of the trivalent miniproteins showed comparable anti-HIV activities in comparison with a typical bivalent miniprotein, regardless of the length of the linker. The results implicate a novel mechanism of the interactions between the multivalent inhibitors and the trimeric gp120 complex.

1. Introduction

The initial step of HIV-1 entry is characterized by the binding of HIV-1 envelope glycoprotein gp120 to the host receptor CD4.[1] It has been shown that the CD4 binding site on HIV-1 gp120 is centered on a conserved, hydrophobic pocket denoted the 'Phe43 cavity'.[2,3] Compounds that can bind to the conserved CD4-binding pocket, thus interfering the CD4-gp120 interactions, constitute an important class of entry inhibitors against HIV-1 infection. These include soluble CD4, the CD4-mimetic miniproteins, HIV-neutralizing antibody b12, and some small-molecule compounds such as BMS-806.[4–14] Recent structural and biochemical studies have demonstrated that the HIV-1 envelope glycoprotein gp120 forms a trimeric complex associated with the inner envelope glycoprotein gp41.[15–20] Thus, the trimeric gp120 complex of the envelope spike constitutes a typical trivalent target that presents three CD4-binding sites in the complex. As an effort to develop more potent HIV-1 entry inhibitors, we are interested in creating multivalent inhibitors that can simultaneously target two or three CD4-binding sites in the envelope complex. There are ample examples in the literature demonstrating that multivalent interactions generated by a multivalent ligand and a corresponding multivalent target could be collectively much stronger than the otherwise monovalent interactions.[21–26] To apply this concept to HIV-1 inhibitor design, we have previously synthesized bivalent CD4-mimetic miniproteins that contain two CD4M9 moieties tethered with a spacer of varied length.[27] CD4M9 is a scorpion toxin-based 28-mer polypeptide that binds to the Phe43 cavity of gp120 and inhibits HIV-1 infection at micro-molar concentration.[10] We have found that the bivalent CD4M9 miniproteins demonstrated significantly enhanced (4- to 21-fold) anti-HIV activity over the monomeric CD4M9. These encouraging results prompted us to test whether trivalent CD4-mimetic compounds will possess further enhanced anti-HIV activities, through targeting the three CD4-binding sites in the complex simultaneously. In this paper, we report the synthesis of a series

Keywords: Miniprotein; CD4-mimetics; HIV; Envelope glycoprotein gp120; Multivalent inhibitor; Entry inhibitor; CD4 binding sites; Anti-HIV activity.
* Corresponding author. Tel.: +1 410 706 4982; fax: +1 410 706 5068; e-mail: wangx@umbi.umd.edu

0968-0896/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2007.03.064

of trivalent CD4-mimetic miniproteins, in which three CD4M9 moieties were tethered to a threefold-symmetric scaffold with a spacer of varied length and nature. Antiviral assays have been performed to reveal the structure–activity relationships of the synthetic trivalent miniproteins.

2. Results and discussion

2.1. Design

It has been demonstrated that the HIV-1 envelope glycoprotein gp120 is present as a trimeric complex on the surface of virion or infected T-cells.[15–20] Although the trimeric structure has not been solved by X-ray crystallography, modeling studies suggest that the trimeric gp120 complex is threefold-symmetric.[18] The distance between any two of the CD4-binding pockets (the Phe43 cavity) was estimated to be 30–60 Å.[27] Based on these structural features, we sought to create threefold-symmetric trivalent miniproteins on a scaffold. The trivalent miniproteins are designed to target the three CD4-binding pockets in the trimeric gp120 complex simultaneously. We have previously created bivalent CD4-mimetic miniproteins using CD4M9 as the model inhibitor and showed that the bivalent miniproteins could have a 4- to 21-fold enhancement in the anti-HIV activity over the monomeric CD4M9.[27] Thus, to match the distance between any two of the CD4-binding cavities, spacers with varied length and nature will be incorporated in the trivalent miniproteins. The schematic structure of the trimeric gp120 complex and the designed trivalent CD4-mimetic miniproteins are shown in Figure 1.

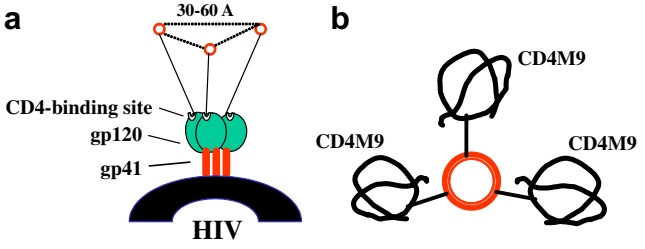


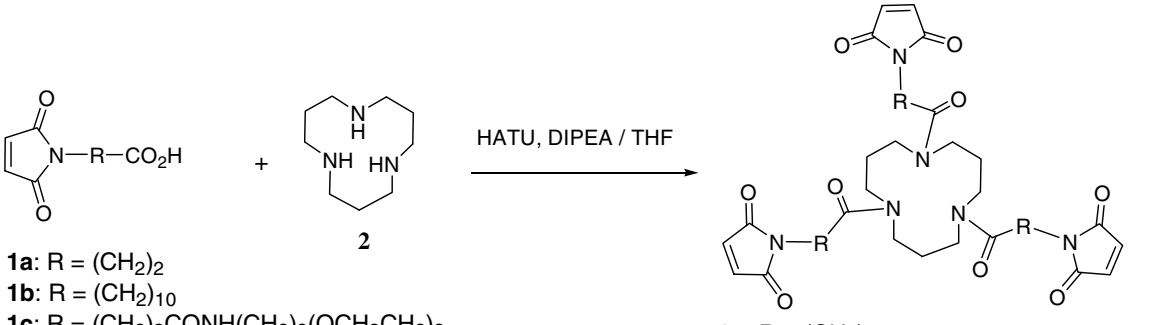
Figure 1. Schematic depiction of the trimeric gp120 complex (a) and the template-assembled trivalent miniproteins (b) (Illustration is not to scale).

2.2. Synthesis

To assemble the designed trivalent miniproteins, the highly efficient chemoselective ligation between maleimide and thiol functionality was chosen, which was successfully used for the preparation of the bivalent CD4M9 miniproteins.[27] For the purpose, a series of threefold-symmetric trivalent maleimides were synthesized on a suitable scaffold. The first class of trivalent maleimides was built on a cyclic triamine scaffold. Thus, the reaction of maleimide-containing carboxylic acid (**1a**–**1d**) and the 1,5,9-triazacyclododecane (**2**) in the presence of the coupling reagent HATU/DIPEA gave the triazacyclododecane-centered trivalent maleimides **3a**–**3d**, respectively, in which the maleimide moieties were spaced with varied length of linkers (Scheme 1). The products were purified by reverse phase HPLC.

Another class of trivalent maleimides was assembled on the Kemp's triacid core, which provides another rigid scaffold of threefold-symmetry. Kemp's triacid is known in supramolecular chemistry as useful building block to assemble peptide–peptoids structures.[28,29]

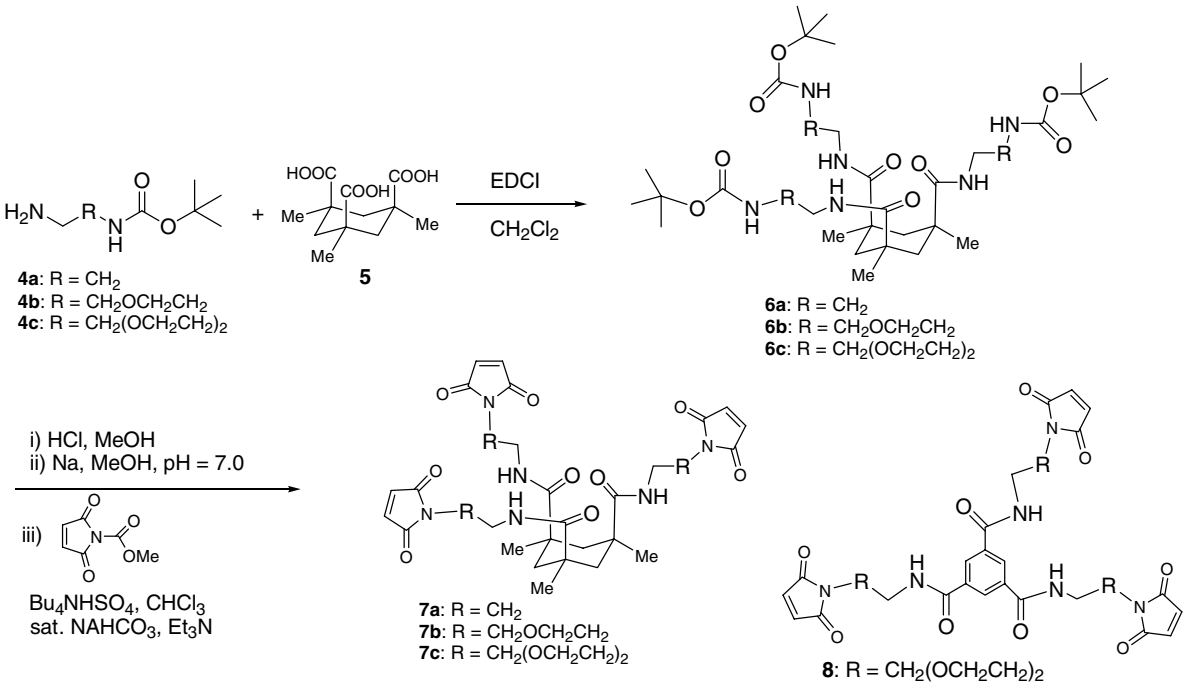
For the purpose, the mono-*N*-Boc-protected diamine (**4a**–**4c**) with varied lengths was coupled to the Kemp's triacid (**5**) using 1-ethyl-3-[3-dimethylaminoisopropoyl]carbodiimide (EDCI) as the coupling agent to give the Kemp's triacid-centered Boc-protected triamine derivatives **6a**–**6c**, respectively. The Boc groups were removed through treatment with aq HCl/MeOH at 0 °C. Subsequent neutralization with Na/MeOH gave the free amino groups that were reacted with *N*-methoxycarbonylmaleimide to provide the Kemp's triacid-centered trivalent maleimides (**7a**–**7c**), respectively (Scheme 2). Similarly, the trimesic acid centered trivalent maleimide (**8**) was synthesized using the same reaction sequence as described for the preparation of **7a**–**c**, starting with trimesic acid as the scaffold.



1a: R = $(CH_2)_2$
1b: R = $(CH_2)_{10}$
1c: R = $(CH_2)_2CONH(CH_2)_2(OCH_2CH_2)_2$
1d: R = $(CH_2)_2CONH(CH_2)_2(OCH_2CH_2)_{12}$

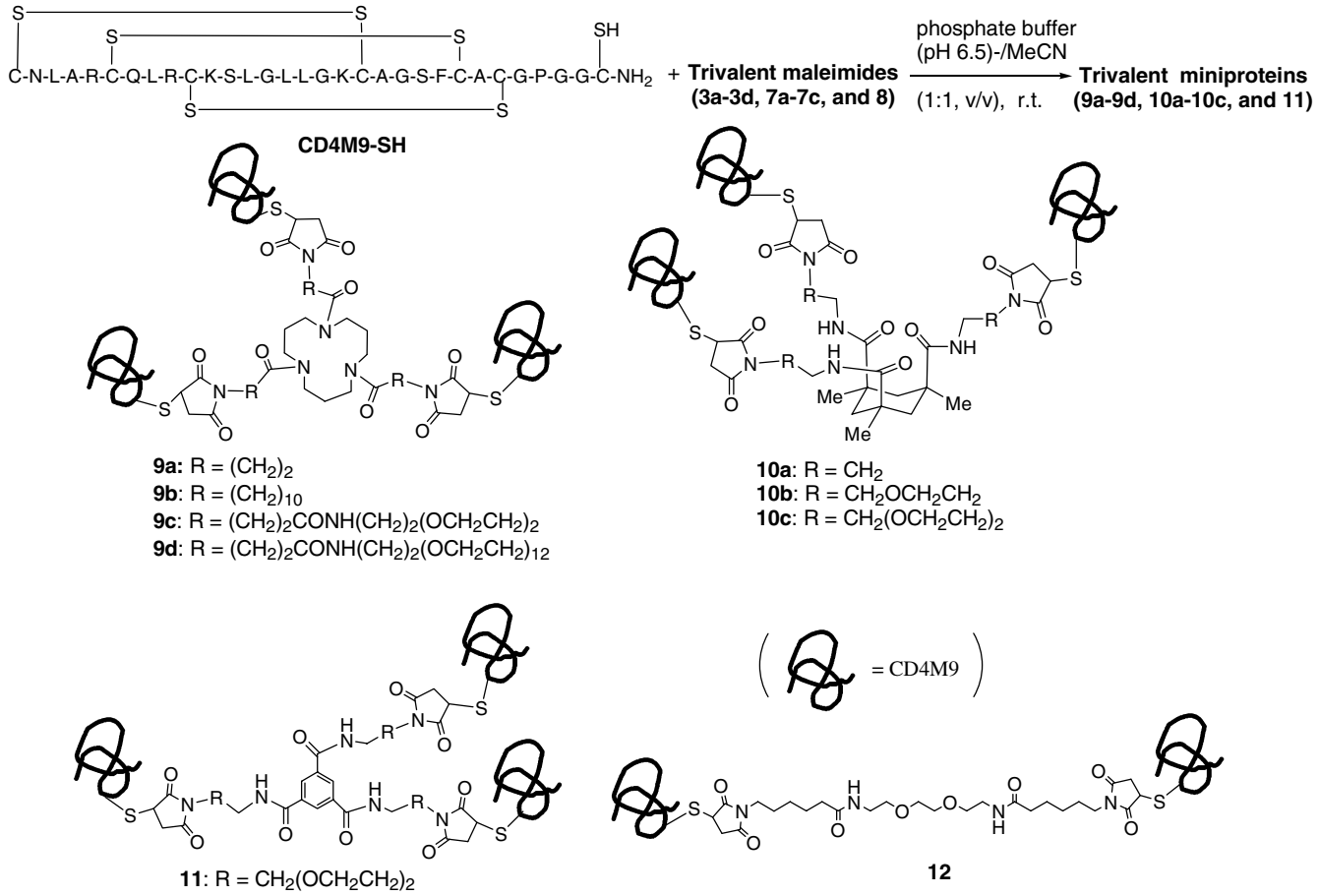
3a: R = $(CH_2)_2$
3b: R = $(CH_2)_{10}$
3c: R = $(CH_2)_2CONH(CH_2)_2(OCH_2CH_2)_2$
3d: R = $(CH_2)_2CONH(CH_2)_2(OCH_2CH_2)_{12}$

Scheme 1. Synthesis of triazacyclododecane-centered trivalent maleimides.



Scheme 2. Synthesis of Kemp's acid-centered trivalent maleimides.

To ligate the miniprotein CD4M9 onto the trivalent maleimide scaffold, an extra cysteine residue with a short spacer (GG) was introduced at the C-terminus of CD4M9, because the C-terminus residue of CD4M9 does not directly involve in the interaction with gp120 as revealed by the modeling study.[10] Our previous exper-



Scheme 3. Synthesis of trivalent CD4-mimetic miniproteins.

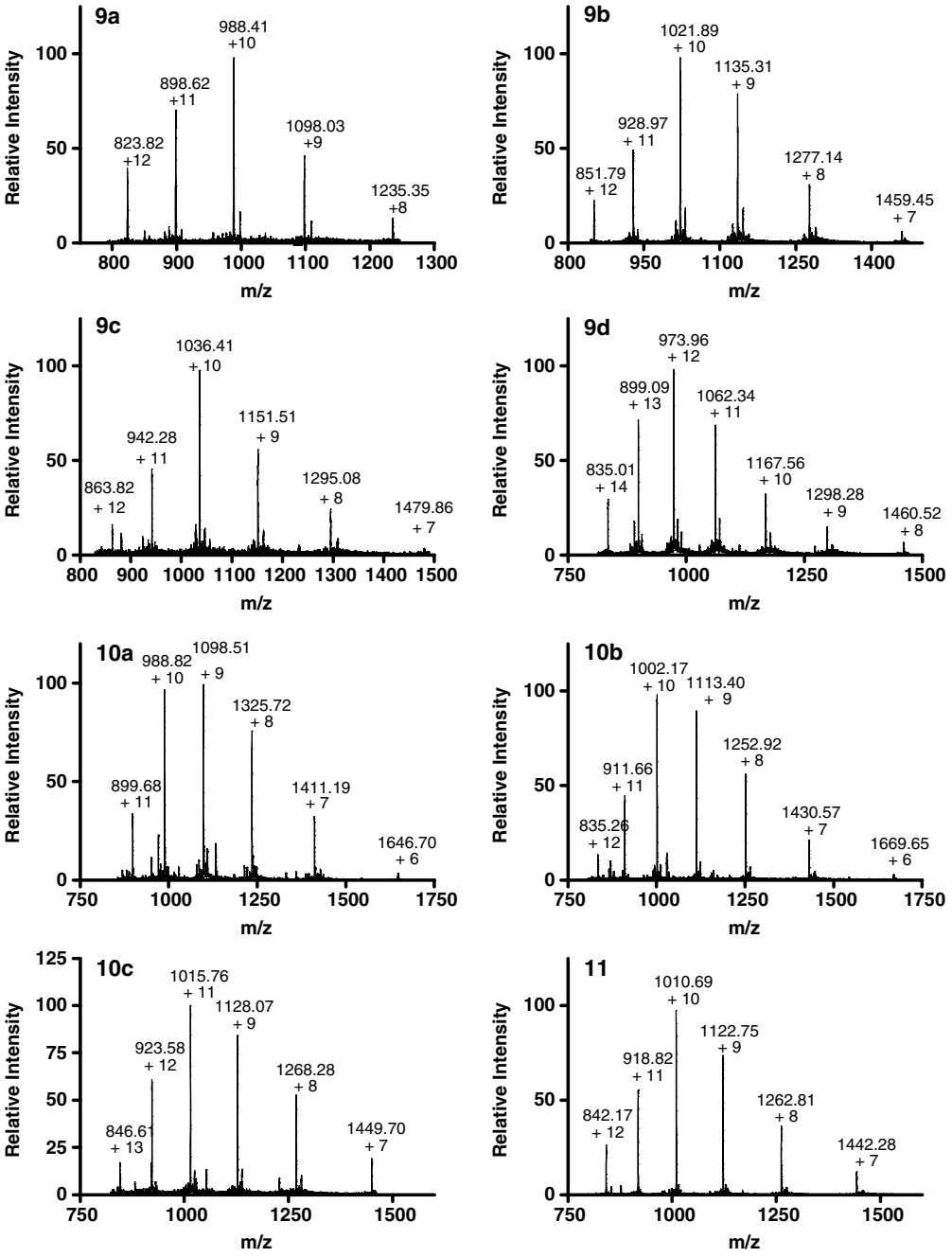


Figure 2. The ESI-MS profiles of the synthetic miniproteins.

imental data also confirmed that addition of a tag (GGC) at the C-terminus of CD4M9 did not influence its anti-HIV activity.[27] The ligation of the CD4M9-SH with the synthetic trivalent maleimide scaffolds (**3a**–**3d**, **7a**–**7c**, and **8**) was performed in a phosphate buffer (pH 6.5) containing 50% MeCN at rt, and the reaction was monitored by HPLC. It was found that when an excess CD4M9-SH (1.3 mol equiv per maleimide moiety) was used, the ligation reaction went to completion within 1 h to give the corresponding trivalent miniproteins (**9a**–**9d**, **10a**–**10c**, and **11**) in excellent yield (Scheme 3). The products were easily purified by preparative HPLC, and the identity of the trivalent miniproteins was confirmed by ESI-MS (Fig. 2).

2.3. Anti-HIV activity of the synthetic trivalent miniproteins

The anti-HIV activity of the synthetic miniproteins was examined by measuring their inhibition against HIV-1 infection to PM1 cells. Two types of HIV-1 viruses were used. One is the HIV-1_{Bal} that uses the chemokine receptor CCR5 as co-receptor for entry, and the other is the X4-tropic virus HIV-1_{IIIB}, which uses CXCR4 as co-receptor for entry. The results of anti-HIV activity are summarized in Table 1. The lengths of the spacers listed in Table 1 were the maximum lengths calculated based on the zigzag C–X–C bonds, which provides a rough estimate but may not represent the actual distances

Table 1. Anti-HIV activity of the synthetic compounds[a]

Compound entity	Spacer length (max)	IC_{50}[b] (μM)	
		HIV-1_{Bal} (relative activity)	HIV-1_{IIIB} (relative activity)
CD4M9	N/A	9.7 (*1*)	12 (*1*)
9a	18 Å	1.5 (*6.5*)	3.0 (*4*)
9b	36 Å	0.068 (*142*)	0.40 (*30*)
9c	43 Å	1.3 (*7.5*)	2.5 (*4.8*)
9d	138 Å	nd[c]	1.9 (*6.3*)
10a	19 Å	0.79 (*12*)	nd[d]
10b	22 Å	1.8 (*5.4*)	nd[d]
10c	29 Å	1.0 (*9.7*)	nd[d]
11	30 Å	0.33 (*29*)	nd[d]
12	32 Å	0.95 (*10*)	1.3 (*9.2*)
sCD4	N/A	0.013 (*746*)	0.005 (*2400*)

[a] Virus inhibition was performed using PHA-stimulated PM1 the host cells infected by HIV-1_{BaL} or HIV-1_{IIIB}.

[b] The IC_{50} was defined as the concentrations that reach 50% inhibition of viral production, and the numbers in the parenthesis indicate the relative activity to CD4M9 based on the measured IC_{50}.

[c] nd (not determined), because no dose response was observed for this compound.

[d] Data were not determined.

because of conformational variants. The antiviral studies with the synthetic miniproteins and HIV-1_{Bal} revealed several interesting features of the structure–activity relationships for the synthetic compounds. First of all, most of the trivalent miniproteins showed enhanced anti-HIV activity over the monomeric inhibitor CD4M9, indicating the beneficial effects of multivalent interactions generated by the trivalent ligands. Second, it seems that the length of the linker that separates the CD4M9 moieties in the miniproteins could be relatively flexible for the anti-HIV activity. For example, no significant difference in the anti-HIV activities was observed between compounds **9a** and **9c**, or those between **10a**, **10b**, and **10c**, although the length of the linker varied dramatically from 20 to 40 Å. However, if the spacer is too long, the multivalent effect would be lost, as in the case of compound **9d** (with a spacer length of 138 Å) that did not show enhancement of anti-HIV activity. Considering the size of the CD4M9, which would extend *ca* 10 Å from the tethering Cys to the Phe group, a linker of 20–40 Å in length will place any two of the Phe residues in the trivalent CD4M9 miniproteins in the range of 40–60 Å. This distance of separation could match any two sites of the Phe43 cavity on the trimeric gp120, which were estimated to be separated by 30–60 Å. Taken together, the experimental data provide at least a piece of indirect evidence suggesting a dynamic nature of the trimeric gp120 complex. Third, compound **9b**, which has a hydrophobic linker, was found to be over 140-fold more active than CD4M9 against HIV-1_{Bal} infection. It was also significantly more active than other trivalent miniproteins with a hydrophilic linker of similar length. In addition, compound **11**, which has a hydrophobic aromatic core, also showed a 30-fold enhancement in antiviral activity. These results suggest that a hydrophobic tail attached to CD4M9 may offer additional beneficial effects in enhancing the antiviral activity. This is an interesting issue that deserves further investigations. It should be pointed out that the scaffolds and the linkers alone without CD4M9 moiety did not show inhibitory activity against HIV-1 infection at the concentrations (up to 10 μM) tested. It was also observed that the synthetic trivalent CD4M9 miniproteins, as well as CD4M9 itself, did exert an effect on the host cell viability, that is, they were non toxic to the cells at the concentrations (up to 10 μM) tested. In contrast to the typical R5-tropic virus HIV-1_{Bal}, the X4-tropic virus HIV-1_{IIIB} was found to be less sensitive to the trivalent miniproteins, which resulted in only a few fold enhancement in the antiviral activity for most of the trivalent compounds. Finally, the present studies revealed another interesting feature of the multivalent inhibitors when the anti-HIV activities of bi- and trivalent miniproteins are compared. Compound **12** is a typical bivalent miniprotein that we previously synthesized, which showed a 21-fold increase in antiviral activity over CD4M9 in a PBMC-HIV-1_{IIIB} assay system.[27] In the present study, it was observed that bivalent compound **12** demonstrated about 10-fold enhancement in antiviral activity in the cell line-based assay against HIV-1_{Bal} and HIV-1_{IIIB}. However, except for the compound **9b** that has a hydrophobic linker, most of the synthetic trivalent miniproteins synthesized in this study showed very similar enhancement of anti-HIV activity as the bivalent compound, and no further increase in the anti-HIV activity was observed when extending the valency from two to three, regardless of the length of the spacer. These unexpected observations suggest that the third CD4M9 moiety may not contribute much to the affinitive binding of the gp120 complex once the first two are in action. These results may implicate a novel mechanism of the interactions between the multivalent inhibitors and the trimeric gp120 complex. It is likely that the binding of the bi- or tri-valent miniproteins to the first two CD4-binding pockets in the envelope complex causes dramatic conformational changes or the collapse of the gp120 trimeric complex, thus making the third CD4-binding site inaccessible to the third arm of CD4M9 in the trivalent miniproteins.

3. Conclusion

The synthesis of a series of trivalent CD4-mimetic miniproteins was described, in which three CD4M9 miniprotein moieties were tethered on a threefold-symmetric scaffold. Antiviral assays revealed that most of the trivalent miniproteins showed enhanced anti-HIV activity over the monomer CD4M9, indicating the gain of affinity through multivalent interactions. The structure–activity relationship studies suggested that not only the length, but also the nature (hydrophobic versus hydrophilic) of the linker was important for the enhancement of antiviral activities of the trivalent miniproteins, with one compound that has a hydrophobic linker being 140-fold more active than CD4M9 against HIV-1_{Bal} infection. Moreover, the similarity in activity enhancement between most corresponding bi- and trivalent miniproteins may reveal a novel mechanism of interactions between the multivalent ligands with the trimeric gp120 complex. It is likely that the binding of the bi- or trivalent miniproteins to the first two

CD4-binding sites in the complex causes dramatic conformational changes of the gp120 trimeric complex, thus making the third CD4-binding site inaccessible to the third CD4M9 moiety in the trivalent miniproteins. A more detailed structural and mechanistic study is required to clarify this point.

4. Experimental

4.1. General methods

^{1}H and ^{13}C NMR spectra were recorded on Inova 500 NMR in $CDCl_3$ unless otherwise specified. Chemical shifts are expressed in ppm downfield using external Me_4Si (0 ppm) as the reference. The ESI-MS spectra were measured on a micromass ZQ-400 single quadrupole mass spectrometer. FABMS were measured with a Kratos MS-80-RFA instrument; for HR-FABMS a Micromass AutoSpeQ instrument was used. NaI was used as cationizing agent and, thioglycerol (TG) or nitrobenzylic alcohol (NBA) as matrix. TLC was performed on glass plates coated with silica gel 60 F254 (E. Merck). Flash column chromatography was performed on silica gel 60 (EM Science, 230–400 mesh). Analytical HPLC was carried out on a Waters 626 HPLC instrument equipped with a Waters Nova-Pak C18 column (3.9 × 150 mm) at 40 °C. The column was eluted with a linear gradient of 0–90% MeCN containing 0.1% TFA at a flow rate of 1 mL/min over 20 min. Compounds were detected at 214 nm. Preparative HPLC was performed on a Waters 600 HPLC instrument with a Waters C18 preparative column (Symmetry 300, 19 × 300 mm). The column was eluted with a suitable gradient of MeCN containing 0.1% TFA at 10 mL/min. The maleimide-acids (**1a** and **1 b**) were obtained from Pierce Biotechnology Inc. (Rockford, IL). The PEG linked maleimide-acids (**3** and **4**) were from Quanta Biodesign (Powell, OH). The miniprotein CD4M9-SH (CD4M9-GGC) was prepared as previously described.[27] Other chemicals were purchased from Aldrich/Sigma and used as received.

4.2. Synthesis of the triazacyclododecane-centered trivalent maleimides (3a–3d)

To a solution of the maleimide acid **1** (0.8 mmol) and triazacyclododecane (**2**) (0.2 mmol) in DMF (10 mL) were added HATU (0.8 mmol) and DIPEA (0.8 mmol). The mixture was stirred at rt for 2 h and then diluted with EtOAc (60 mL). The mixture was then washed with brine and water. The organic layer was dried over Na_2SO_4, filtered, and the filtrate was subjected to silica gel flash column chromatography using CH_2Cl_2/MeOH (99:1 to 90:10, v/v) to give the respective trivalent maleimide. The products were further purified by preparative HPLC (gradient: 20–40% MeCN containing 1% TFA in 30 min; flow rate 10 mL/min) to give pure triazacyclododecane-centered trivalent maleimides (**3a**–**3d**).

4.2.1. Trivalent maleimide (3a). Yield, 62%; ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.99 (m, 6H, tricyclicamino-NCH_2CH_2), 2.70 (t, 6H, J = 5.0 Hz, $COCH_2$), 3.47 (m, 12H, tricyclicamino-NCH_2CH_2), 3.89 (t, 6H, J = 5.0 Hz, maleimido-NCH_2), 6.75 (s, 6H, maleimido-$COCH$=CH). ^{13}C (125 MHz, $CDCl_3$): δ 170.32, 162.18, 134.50, 70.23, 69.79, 39.29, and 34.44. ESI-MS calcd M = 624.64; Found: 625.22 $(M + H)^+$.

4.2.2. Trivalent maleimide (3b). Yield, 53%; ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.31 (s, 18H $CH_2(CH_2)_3CH_2$), 1.61 (t, 6H, J = 6.5 Hz, maleimido-NCH_2CH_2), 1.64 (t, 6H, J = 5.0 Hz, $COCH_2CH_2$), 2.04 (t, 6H, J = 7.5 Hz, tricyclicamino-NCH_2CH_2), 2.36 (t, 6H, J = 7.5 Hz, tricyclicamino-NCH_2CH_2), 3.47 (t, 6H, J = 6.0 Hz, NCH_2CH_2), 3.55 (t, 6H, J = 7.5 Hz, $COCH_2CH_2$), 6.73 (s, 6H, maleimido-$COCH$=CH). ^{13}C (125 MHz, $CDCl_3$) δ 211.68, 170.93, 134.07, 110.00, 37.95, 33.51, 29.41, 29.09, 28.54, 26.74, and 26.66. ESI-MS calcd M = 960.63; Found: 961.42 $(M + H)^+$, 481.20 $(M + 2H)^{2+}$.

4.2.3. Trivalent maleimide (3c). Yield, 65%; ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.04 (t, 6H, J = 6.0 Hz, tricyclicamino-NCH_2CH_2), 2.61 (t, 6H, J = 7.0 Hz, tricyclicamino-$NCOCH_2$), 2.67 (t, 6H, J = 6.0 Hz, maleimido-NCH_2CH_2CO), 3.45 (q, 6H, J = 5.0 Hz, $CONHCH_2$), 3.50 (t, 6H, J = 6.0 Hz, tricyclicamino-$NCOCH_2CH_2O$), 3.56 (t, 12H, J = 4.5 Hz, tricyclicamino-NCH_2CH_2), 3.64 (t, 12H, J = 5.0 Hz, OCH_2CH_2O), 3.85 (t, 6H, J = 6.0 Hz, OCH_2CH_2CON), 3.89 (t, 6H, J = 7.5 Hz, maleimido-NCH_2), 6.74 (s, 6H, maleimido-CH=CH), 6.84 (s, 3H, NH). ^{13}C (125 MHz, $CDCl_3$) δ 170.56, 134.25, 70.24, 67.50, 39.29, 34.44, and 33.51. ESI-MS calcd M = 1101.52; Found: 1124.26 $(M + Na)^+$, 1102.27 $(M + H)^+$, 551.79 $(M + 2H)^{2+}$.

4.2.4. Trivalent maleimide (3d). Yield, 48%; ^{1}H NMR (500 MHz, $CDCl_3$): δ1.98 (t, 6H, J = 6.0 Hz, tricyclicamino-NCH_2CH_2), 2.53 (t, 6H, J = 7.0 Hz, tricyclicamino-$NCOCH_2$), 2.63 (t, 6H, J = 6.0 Hz, maleimido-NCH_2CH_2CO), 3.42 (q, 12H, J = 5.0 Hz, $CONHCH_2CH_2$), 3.49 (t, 6H, J = 6.0 Hz, tricyclicamino-$NCOCH_2CH_2O$), 3.55 (t, 6H, J = 5.0 Hz, OCH_2CH_2CONH), 3.62 (t, 12H, J = 2.0 Hz, tricyclicamino-NCH_2CH_2), 3.64 (m, 120H, OCH_2CH_2O), 3.78 (t, 6H, J = 6.5 Hz, maleimido-NCH_2), 3.84 (t, 6H, J = 7.0 Hz, 6.45 (s, 3H, NH), 6.71(s, 6H, maleimido-CH=CH). ESI-MS calcd M = 2423.77; Found: 1234.05 $(M + 2Na)^{2+}$, 1212.50 $(M + 2H)^{2+}$, 808.62 $(M + 3H)^{3+}$, 606.77 $(M + 4H)^{4+}$, 485.45 $(M + 5H)^{5+}$.

4.3. Synthesis of Kemp's triacid (KTA)-centered Boc-protected triamine derivatives (6a–6c)

To a suspension of Kemp's triacid (KTA) (**5**) (0.5 mmol) in dry CH_2Cl_2 at 0 °C, EDCI (1.65 mmol) was added. Then, the corresponding mono-*N*-Boc-protected diamine (**4**) (1.65 mmol) were added dropwise to the suspension. The resulting mixture was stirred at rt overnight. The reaction mixture was washed with water, and the organic phase was dried over anhydrous Na_2SO_4, filtered, and the filtrate was concentrated to dryness. The residue was purified by column chromatography on silica gel using CH_2Cl_2 containing MeOH (3% → 5%) as

the eluent to give the respective Boc-protected triamine derivative (**6a**–**6c**).

4.3.1. KTA-(*N*-*t*-butoxycarbonyl-ethylenediamine)$_3$ (6a). Yield, 64% (as a colorless oil). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.05 (d, 1H, J = 15.6 Hz), 1.21 (s, 3H), 1.43 (s, 9H), 2.86 (d, 1H, J = 15.6 Hz), 3.21 (bs, 4H), 5.84 (bs, 1H), 7.43 (bs, 1H). ^{13}C NMR (75 MHz, $CDCl_3$), δ 28.4, 34.4, 40.1, 40.8, 42.2, 43.3, 79.2, 156.0, and 177.4. FABMS (TG): 385 (12) $[M - 3Boc + H]^+$, 407 (47) $[M - 3Boc + Na]^+$, 485 (5) $[M - 2Boc + H]^+$, 585 (23) $[M - Boc + H]^+$, 685 (12) $[M + H]^+$, 707 (100) $[M + Na]^+$. HR-FABMS calcd $[M + Na]^+C_{33}H_{60}N_6O_9Na$ 707.4319; Found: 707.4329.

4.3.2. KTA-(1-*t*-Butoxycarbonylamino-3-oxa-pentan-5-amine)$_3$ (6b). Yield, 52% (as a colorless oil). ^{1}H NMR (300 MHz, $CDCl_3$) δ: 1.07 (d, 1H, J = 15.6 Hz), 1.24 (s, 3H), 1.43 (s, 9H), 2.95 (d, 1H, J = 15.3 Hz), 3.20–3.41 (m, 4H), 3.47 (t, 2H, J = 5.1Hz), 3.54 (t, 2H, J = 5.1Hz), 6.49 (bs, 1H), 7.68 (t, 1H, J = 4.3 Hz). ^{13}C NMR (75 MHz, $CDCl_3$) δ: 28.5, 34.5, 40.1, 40.6, 42.3, 43.1, 68.7, 70.1, 78.7, 156.4, 177.2. FABMS (TG): 717 (8) $[M - Boc + H]^+$, 817 (9) $[M + H]^+$, 839 (23) $[M + Na]^+$. HR-FABMS calcd $[M + Na]^+C_{39}H_{72}N_6O_{12}Na$ 839.5106; Found: 839.5114.

4.3.3. KTA-(1-*t*-Butoxycarbonylamino-3,6-dioxa-octan-8-amine)$_3$ (6c). Yield, 68% (as a colorless oil). ^{1}H NMR (300 MHz, $CDCl_3$), δ 1.05 (d, 1H, J = 15.3 Hz), 1.20 (s, 3H), 1.44 (s, 9H), 2.87 (d, 1H, J = 15.6 Hz), 3.25–3.38 (m, 4H), 3.49 (t, 2H, J = 6.3 Hz), 3.55 (t, 2H, J = 5.1Hz), 3.61 (bs, 4H), 5.23 (bs, 1H), 7.56 (t, 1H, J = 5.1Hz). ^{13}C NMR (75 MHz, $CDCl_3$) δ: 28.4, 34.2, 39.4, 40.4, 42.2, 43.1, 69.2, 70.2, 70.3, 79.1, 156.0, 177.3. FABMS (TG): 671 (17) $[M - 3Boc + Na]^+$, 971 (65) $[M + Na]^+$. HR-FABMS calcd $[M + Na]^+C_{45}H_{84}N_6O_{15}Na$ 971.5892; Found: 971.5869.

4.4. Synthesis of the KTA-centered trivalent maleimides (7a–7c)

The respective *N*-Boc protected triamine **6** (0.2 mmol) was dissolved in MeOH and cooled to 0 °C. To the solution was added dropwise concd HCl (2.2 mL). The reaction mixture was stirred at 0 °C until consumption of the starting material (monitored by TLC). The reaction mixture was concentrated and the residue was dissolved again in MeOH and neutralized with Na to pH 7.0. The formed NaCl was filtered off and the filtrate dried over anhydrous Na_2SO_4 and concentrated to provide the corresponding free triamine, which was used for the next step without further purification. The respective triamine (0.20 mmol) was dissolved in $CHCl_3$ (10 mL) and cooled to 0 °C. To the solution were added tetrabutylammonium hydrogen sulfate (1.38 mmol) and *N*-methoxycarbonylmaleimide (1.38 mmol). Then triethylamine (97 μl) was added slowly, and the resulting mixture was stirred at 0 °C for 10 min. Aq satd $NaHCO_3$ solution (10 mL) was then added in one portion and the mixture was stirred vigorously at rt for 4 h. The organic layer was separated, dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated and the residue was subject to column chromatography on silica gel using CH_2Cl_2 containing MeOH (3% → 5%) as the eluent to provide the KTA-centered trivalent maleimides (**7a**–**7c**).

4.4.1. KTA-(1-maleimide-ethylenamine)$_3$ (7a). Yield, 47% (as a colorless oil). ^{1}H NMR (300 MHz, $CDCl_3$), δ 1.00 (d, 1H, J = 15.6 Hz), 1.13 (s, 3H), 2.73 (d, 1H, J = 15.6 Hz), 3.25 (d_{AB}, 2H, J_{AB} = 5.7 Hz), 3.60 (t, 2H, J = 5.7 Hz), 6.71 (s, 1H), 7.44 (t, 1H, J = 5.7 Hz). ^{13}C NMR (75 MHz, $CDCl_3$), δ 33.5, 37.1, 38.0, 42.1, 43.0, 134.1, 170.6, 177.7. FABMS (NBA): 647 (21) $[M + Na]^+$. HR-FABMS calcd $[M + Na]^+C_{30}H_{36}N_6O_9Na$ 647.2441; Found: 647.2445.

4.4.2. KTA-(1-maleimide-3-oxa-pentan-5-amine)$_3$ (7b). Yield, 53% (as a colorless oil). ^{1}H NMR (300 MHz, $CDCl_3$), δ 1.01 (d, 1H, J = 15.6 Hz), 1.18 (s, 3H), 2.82 (d, 1H, J = 15.6 Hz), 3.20 (d_{AB}, 2H, J_{AB} = 6.0Hz), 3.43 (t, 2H, J = 6.0Hz), 3.65 (A_2B_2 system, 4H), 6.70 (s, 1H), 7.44 (t, 1H, J = 5.4 Hz). ^{13}C NMR (75 MHz, $CDCl_3$), δ 34.0, 37.1, 39.3, 42.1, 43.0, 67.6, 68.6, 134.1, 170.6, 177.1. FABMS (NBA): 779 (11) $[M + Na]^+$, 811 (23) $[M + Na + MeOH]^+$. HR-FABMS calcd $[M + Na]^+C_{36}H_{48}N_6O_{12}Na$ 779.3228; Found: 779.3269.

4.4.3. KTA-(1-maleimide-3,6-dioxa-octan-8-amine)$_3$ (7c). Yield, 56% (as a yellow oil). ^{1}H NMR (300 MHz, $CDCl_3$), δ 1.04 (d, 1H, J = 15.6 Hz), 1.20 (s, 3H), 2.87 (d, 1H, J = 15.3 Hz), 3.26 (d_{AB}, 2H, J_{AB} = 5.7 Hz), 3.45 (t, 2H, J = 5.7 Hz), 3.58 (A_2B_2 system, 6H), 3.69 (A_2B_2 system, 2H), 6.72 (s, 1H), 7.52 (t, 1H, J = 5.4 Hz). ^{13}C NMR (75 MHz, $CDCl_3$), δ 34.2, 37.1, 39.4, 42.2, 43.1, 67.8, 69.2, 69.9, 70.2, 134.2, 170.6, 177.3. FABMS (NBA): 889 (6) $[M + H]^+$, 911 (91) $[M + Na]^+$. HR-FABMS calcd $[M + Na]^+C_{42}H_{60}N_6O_{15}Na$ 911.4014; Found: 911.4047.

4.5. Synthesis of the trimesic acid-centered trivalent maleimide (8)

The trimesic acid-based trivalent maleimide (**8**) was synthesized in the same way as described for the preparation of the KTA-centered trivalent maleimides (**7a**–**7c**), using trimesic acid as the template. The trivalent maleimide (**8**) was obtained in 48% yield as a yellow oil. ^{1}H NMR (300 MHz, $CDCl_3$), δ 3.55–3.76 (m, 12H), 6.68 (s, 1H), 7.26 (bm, 1H), 8.47 (s, 1H). ^{13}C NMR (75 MHz, $CDCl_3$), δ 37.1, 40.1, 67.9, 69.7, 70.0, 70.2, 128.5, 134.1, 135.1, 166.0, 170.7. FABMS (NBA): 841 (5) $[M + H]^+$, 863 (24) $[M + Na]^+$. HR-FABMS calcd $[M + Na]^+C_{39}H_{48}N_6O_{15}Na$ 863.3075; Found: 863.3099.

4.6. Synthesis of the trivalent CD4-mimetic miniproteins (9a–9d, 10a–10c, and 11)

To a solution of CD4M9-SH (8 μmol) in a degassed phosphate buffer (50 mM, pH 6.5) containing 50% MeCN (3 mL) was added a solution of the respective trivalent maleimide (**3a**–**3d**, **7a**–**7c**, or **8**) (2 μmol) in

MeCN (0.2 mL). The resulting mixture was gently shaken under N_2 at rt, and the ligation reaction was monitored with analytical HPLC, which indicated the completion of ligation within 1 h. The mixture was lyophilized and the residue was subjected to preparative HPLC to give the respective trivalent CD4-mimetic miniproteins (**9a**–**9d**, **10a**–**10c**, and **11**).

4.6.1. Trivalent miniprotein (9a). Yield, 90%; t_R 13.60 min (Under the analytical HPLC condition described in the general method); ESI-MS calcd M = 9873.68; Found: 1646.77 $(M + 6H)^{6+}$, 14115.32 $(M + 7H)^{7+}$, 1235.29 $(M + 8H)^{8+}$, 1098.16 $(M + 9H)^{9+}$, 988.42 $(M + 10H)^{10+}$, 898.69 $(M + 11H)^{+11}$, 823.88 $(M + 12H)^{+12}$, 760.54 $(M + 13H)^{13+}$.

4.6.2. Trivalent miniprotein (9b). Yield, 92%; t_R 12.40 min (Under the analytical HPLC condition described in the general method); ESI-MS calcd M = 10210.76; Found: 1702.56 $(M + 6H)^{6+}$, 1459.43 $(M + 7H)^{7+}$, 1277.13 $(M + 8H)^{8+}$, 1135.35 $(M + 9H)^{7+}$, 1021.95 $(M + 10H)^{10+}$, 929.24 $(M + 11H)^{11+}$, 851.87 $(M + 12H)^{12+}$, 786.42 $(M + 13H)^{13+}$.

4.6.3. Trivalent miniprotein (9c). Yield, 87%; t_R 10.37 min (Under the analytical HPLC condition described in the general method); ESI-MS calcd M = 10354.67; Found: 1479.98 $(M + 7H)^{7+}$, 1295.15 $(M + 8H)^{8+}$, 1151.45 $(M + 9H)^{9+}$, 1036.35 $(M + 10H)^{10+}$, 942.34 $(M + 11H)^{11+}$, 863.82 $(M + 12H)^{12+}$.

4.6.4. Trivalent miniprotein (9d). Yield, 86%; t_R 11.41 min (Under the analytical HPLC condition described in the general method); ESI-MS calcd M = 11675.25; Found: 1460.39 $(M + 8H)^{8+}$, 1298.28 $(M + 9H)^{9+}$, 1168.56 $(M + 10H)^{10+}$, 1062.41 $(M + 11H)^{11+}$, 973.96 $(M + 12H)^{12+}$, 899.10 $(M + 13H)^{13+}$, 835.01 $(M + 14H)^{14+}$, 779.41 $(M + 15H)^{15+}$.

4.6.5. Trivalent miniprotein (10a). Yield, 90%; t_R 10.88 min (Under the analytical HPLC condition described in the general method); ESI-MS calcd M = 9873.87; Found: 1646.05 $(M + 6H)^{6+}$, 1411.19 $(M + 7H)^{7+}$, 1235.72 $(M + 8H)^{8+}$, 1098.51 $(M + 9H)^{9+}$, 988.82 $(M + 10H)^{10+}$, 899.68 $(M + 11H)^{11+}$, 823.85 $(M + 12H)^{12+}$.

4.6.6. Trivalent miniprotein (10b). Yield, 81%; t_R 11.26 min (Under the analytical HPLC condition described in the general method); ESI-MS calcd M = 10006.38; Found: 1669.65 $(M + 6H)^{6+}$, 1430.57 $(M + 7H)^{7+}$, 1252.92 $(M + 8H)^{8+}$, 1113.40 $(M + 9H)^{9+}$, 1002.17 $(M + 10H)^{10+}$, 911.66 $(M + 11H)^{11+}$, 835.26 $(M + 12H)^{12+}$.

4.6.7. Trivalent miniprotein (10c). Yield, 83%; t_R 10.90 min (Under the analytical HPLC condition described in the general method); ESI-MS calcd M = 10138.99; Found: 1691.52 $(M + 6H)^{6+}$, 1449.70 $(M + 7H)^{7+}$, 1268.28 $(M + 8H)^{8+}$, 1128.07 $(M + 9H)^{9+}$, 1015.76 $(M + 10H)^{10+}$, 923.58 $(M + 11H)^{11+}$, 846.61 $(M + 12H)^{12+}$.

4.6.8. Trivalent miniprotein (11). Yield, 86%; t_R 11.13 min (Under the analytical HPLC condition described in the general method); ESI-MS calcd M = 10090.12; Found: 1442.28 $(M + 7H)^{7+}$, 1262.81 $(M + 8H)^{8+}$, 1122.75 $(M + 9H)^{9+}$, 1010.69 $(M + 10H)^{10+}$, 918.82 $(M + 11H)^{11+}$, 842.17 $(M + 12H)^{12+}$.

4.7. Anti HIV-1 assay

PM1 cells (2×10^4/well) were exposed to 100 TCID50 viruses of HIV-1_{Bal} or HIV-1_{IIIB} strain, which were pretreated with serial dilutions of compounds for 1 h at 37 °C. After 2 h incubation in 200 μl/well of complete RPMI1640 medium, the cells were washed twice and cultured in complete RPMI1640 medium supplemented with the appropriate compounds. The p24 concentrations of day 7 supernatants were quantified with a sandwich ELISA. The antiviral activities of the trivalent CD4-mimetic miniproteins and related compounds were expressed as concentrations that reach the 50% inhibition (IC_{50}) of viral production.

Acknowledgments

The work was supported in part by the European Commission (project TRIoH, FP6 Program, contract LSHG-CT-2003-503480 to I.R.), the 'Ministerio de Educación y Ciencia', Spain (Grant CTQ2004-00649/BQU to I.R.), and the U.S. Department of Health and Human Services (NIH Grant AI51235 to L.X.W.).

References and notes

1. Chan, D. C.; Kim, P. S. *Cell* **1998**, *93*, 681–684.
2. Kwong, P. D.; Wyatt, R.; Robinson, J.; Sweet, R. W.; Sodroski, J.; Hendrickson, W. A. *Nature* **1998**, *393*, 648–659.
3. Wyatt, R.; Kwong, P. D.; Desjardins, E.; Sweet, R. W.; Robinson, J.; Hendrickson, W. A.; Sodroski, J. G. *Nature* **1998**, *393*, 705–711.
4. Smith, D. H.; Byrn, R. A.; Marsters, S. A.; Gregory, T.; Groopman, J. E.; Capon, D. J. *Science* **1987**, *238*, 1704–1707.
5. Traunecker, A.; Luke, W.; Karjalainen, K. *Nature* **1988**, *331*, 84–86.
6. Deen, K. C.; McDougal, J. S.; Inacker, R.; Folena-Wasserman, G.; Arthos, J.; Rosenberg, J.; Maddon, P. J.; Axel, R.; Sweet, R. W. *Nature* **1988**, *331*, 82–84.
7. Burton, D. R.; Pyati, J.; Koduri, R.; Sharp, S. J.; Thornton, G. B.; Parren, P. W.; Sawyer, L. S.; Hendry, R. M.; Dunlop, N.; Nara, P. L.; Lamacchia, M.; Garratty, E.; Stiehm, E. R.; Bryson, Y. T.; Moore, J. P.; Ho, D. D.; Barbas, C. F. *Science* **1994**, *266*, 1024–1027.
8. Saphire, E. O.; Parren, P. W.; Pantophlet, R.; Zwick, M. B.; Morris, G. M.; Rudd, P. M.; Dwek, R. A.; Stanfield, R. L.; Burton, D. R.; Wilson, I. A. *Science* **2001**, *293*, 1155–1159.
9. Martin, L.; Barthe, P.; Combes, O.; Roumestand, C.; Vita, C. *Tetrahedron* **2000**, *56*, 9451–9460.
10. Vita, C.; Drakopoulou, E.; Vizzavona, J.; Rochette, S.; Martin, L.; Menez, A.; Roumestand, C.; Yang, Y. S.; Ylisastigui, L.; Benjouad, A.; Gluckman, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13091–13096.
11. Martin, L.; Stricher, F.; Misse, D.; Sironi, F.; Pugniere, M.; Barthe, P.; Prado-Gotor, R.; Freulon, I.; Magne, X.;

Roumestand, C.; Menez, A.; Lusso, P.; Veas, F.; Vita, C. *Nat. Biotechnol.* **2003**, *21*, 71–76.
12. Lin, P. F.; Blair, W.; Wang, T.; Spicer, T.; Guo, Q.; Zhou, N.; Gong, Y. F.; Wang, H. G.; Rose, R.; Yamanaka, G.; Robinson, B.; Li, C. B.; Fridell, R.; Deminie, C.; Demers, G.; Yang, Z.; Zadjura, L.; Meanwell, N.; Colonno, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11013–11018.
13. Si, Z.; Madani, N.; Cox, J. M.; Chruma, J. J.; Klein, J. C.; Schon, A.; Phan, N.; Wang, L.; Biorn, A. C.; Cocklin, S.; Chaiken, I.; Freire, E.; Smith, A. B., 3rd; Sodroski, J. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5036–5041.
14. Wang, J.; Le, N.; Heredia, A.; Song, H.; Redfield, R.; Wang, L. X. *Org. Biomol. Chem.* **2005**, *3*, 1781–1786.
15. Sanders, R. W.; Vesanen, M.; Schuelke, N.; Master, A.; Schiffner, L.; Kalyanaraman, R.; Paluch, M.; Berkhout, B.; Maddon, P. J.; Olson, W. C.; Lu, M.; Moore, J. P. *J. Virol.* **2002**, *76*, 8875–8889.
16. Schulke, N.; Vesanen, M. S.; Sanders, R. W.; Zhu, P.; Lu, M.; Anselma, D. J.; Villa, A. R.; Parren, P. W.; Binley, J. M.; Roux, K. H.; Maddon, P. J.; Moore, J. P.; Olson, W. C. *J. Virol.* **2002**, *76*, 7760–7776.
17. Salzwedel, K.; Berger, E. A. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 12794–12799.
18. Kwong, P. D.; Wyatt, R.; Sattentau, Q. J.; Sodroski, J.; Hendrickson, W. A. *J. Virol.* **2000**, *74*, 1961–1972.
19. Staropoli, I.; Chanel, C.; Girard, M.; Altmeyer, R. *J. Biol. Chem.* **2000**, *275*, 35137–35145.
20. Zhu, P.; Chertova, E.; Bess, J., Jr.; Lifson, J. D.; Arthur, L. O.; Liu, J.; Taylor, K. A.; Roux, K. H. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15812–15817.
21. Mammen, M.; Choi, S. K.; Whitesides, G. M. *Angew. Chem Int. Ed.* **1998**, *37*, 2754–2794.
22. Kitov, P. I.; Shimizu, H.; Homans, S. W.; Bundle, D. R. *J. Am. Chem. Soc.* **2003**, *125*, 3284–3294.
23. Dekker, F. J.; de Mol, N. J.; van Ameijde, J.; Fischer, M. J.; Ruijtenbeek, R.; Redegeld, F. A.; Liskamp, R. M. *ChemBiochem.* **2002**, *3*, 238–242.
24. Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature* **2000**, *403*, 669–672.
25. Fan, E.; Zhang, Z.; Minke, W. E.; Hou, Z.; Verlinde, C. L. M. J.; Hol, W. G. J. *J. Am. Chem. Soc.* **2000**, *122*, 2663–2664.
26. Kiessling, L. L.; Gestwicki, J. E.; Strong, L. E. *Curr. Opin. Chem. Biol.* **2000**, *4*, 603–696.
27. Li, H.; Song, H.; Heredia, A.; Le, N.; Redfield, R.; Lewis, G. K.; Wang, L. X. *Bioconjugate Chem.* **2004**, *15*, 783–789.
28. Jefferson, E. A.; Locardi, E.; Goodman, M. *J. Am. Chem. Soc.* **1998**, *129*, 7420–7428, and references therein.
29. Kocis, P.; Issakova, O.; Sepetov, N. F.; Lebl, M. *Tetrahedron Lett.* **1995**, *36*, 6623–6626.