# Supplementary Materials

## Peptide synthesis

Peptides were synthesized using standard solid phase Fmoc chemistry protocol using a CS-Bio 136 automated peptide synthesizer (CS Bio Co., Menlo Park, CA) as described previously.[1] All peptides were amidated at their carboxy termini and acetylated at amino termini. The peptides were purified to homology (>95%) by high performance liquid chromatography (HPLC) and the molecular weights were identified by MALDI-TOF-MS (Autoflex III, Bruker Daltonics).

## Circular Dichroism Spectra

Circular Dichroism Spectra (CD) were acquired at room temperature (RT) (Biologic MOS-450: 4.0 nm bandwidth, 0.1 nm resolution, 0.1 cm path length, 4.0 s response time and a 50 nm/min scanning speed). Peptide samples were incubated at 37°C for 30 min in phosphate buffered saline, pH 7.2 (PBS) and cooled to RT before measurement. The spectra were corrected by subtraction of the solvent blank. For thermal midpoint analysis, the temperature was controlled using a Bio-logic TCU250 system. The final concentration of N- and C-peptides was 1 μM in PBS. CD spectra were monitored at 222 nm between 20-90 °C.

## Cell-cell fusion assays

Cell-cell fusion assays were performed as described previously.[2] HL2/3 cells which stably express HIV Gag, Env, Tat, Rev and Nef proteins and TZM-bl cells, which stably express large amounts of CD4 and CCR5, were obtained from the AIDS Reference and Reagent Program, and

were used as the effector and target cells, respectively. TZM-bl cells ($2.5x10^4$ /well) and HL2/3 cells ($7.5x10^4$ /well) were co-incubated in 96-well plates (Corning Costar) at 37 °C in 5% $CO_2$ in the presence of different concentrations of inhibitors. After 6-8 h of incubation, the medium was aspirated and the cells were washed and lysed, followed by measurement of the luciferase activity using the Luciferase Assay System (Promega Corporation) on a plate reader (Molecular Devices SpectraMax M5).

**HIV-1 Infection Assay**

Inhibitory activities of the peptides for HIV-1 infection were determined as previously described [3]. Briefly, $1\times10^4$/ml MT-2 cells in RPMI medium 1640 containing 10% FBS were infected with T20-sensitive and resistant HIV-1 strains that were obtained from the NIH AIDS Research and Reference Reagent Program at 100 TCID50 (50% tissue culture infective dose) in 200 µl culture medium in the presence or absence of the test peptide overnight.[4] The culture supernatants were removed, and fresh media were added the next day. On the fourth day post-infection, 100 µl of culture supernatant was collected from each well, mixed with equal volumes of 5% Triton X-100, and assayed for p24 antigen by ELISA, using an in-house kit as previously described [3a].

**Computational modeling**

In PBD-4HR$_u$, four "*a*" ($a_2$, $a_3$, $a_4$, $a_5$) and four "*d*" ($d_2$, $d_3$, $d_4$, $d_5$) positions (**Table 1**) which interacted directly with gp41 NHR were subject to optimization. Single mutation were introduced into "*a*", or "*d*" position based on the crystal structure of gp41 NHR-CHR 6-HB and were verified by computational modeling of the PBD-4HR$_u$/NHR structure. The predominant conformations of designed artificial HR peptide were investigated with computational modeling to estimate the binding free energy for the association between HR peptides and NHR. Molecular dynamics simulation was conducted on the peptide complex using AMBER 8.0 [5] on a SGI

Altix 350 workstation (USA).

Peptide 1-3 are the selected single mutated peptides, Ile of $a_2$ was the member of WWI motif in PBD, Ser in $a_3$ position could form hydrogen bond with $e_3$ Gln in NHR between –OH and –CO-$NH_2$ in the side chains of the two residues, Ile in $a_4$ interact with $e_4$ Leu of NHR through hydrophobic interaction. For PBD-m4HR, other residues in a, d and e positions were further modified besides the above single mutation to improve the binding affinity to NHR. Two Gln in $d_4$ and $g_4$ of HR peptide and NHR could also form hydrogen bond, the $a_5$ Gln appear to interact with $e_5$ Val of NHR in table 1, but in computational modeling of 3D surrounding of NHR, $a_5$ was not near the $e_5$ V, but near the $b_5$ Ser. Also, $d_5$ Ser does not interact with $g_5$ Leu, but $f_6$ Gln, which could also form hydrogen bond between two side chains of the residues. For PBDm-m4HR, EE-KK motif was introduced to PBD to stabilize the α-helical structure.

The following five artificial HR peptides were investigated with computational modeling analysis: 1) PBD-m4HR, which contains of PBD and four modified 4–3 HR sequence in the *a*, *d*, *e* positions and represents the high active leading sequence modified from the template peptide PBD-4$HR_u$; 2) PBDm-m4HR in which the sequence of PBD was further modified by introducing double salt bridge to improve the potency of inhibitory activity; and 3) Peptide 1, 2, 3, in which only one single residue was modified compare with the template peptide PBD-4$HR_u$. The interaction between the selected five peptides with NHR were compared based on their binding free energies, as shown in Table S1.

**Table S1** The binding free energy of artificial HR peptides with NHR （AMBER 8.0）

| | **PBTOT (KCal/mol)** | **GBTOT (KCal/mol)** |
|---|---|---|
| **PBD-m4HR** | **-279.72±15.99** | **-232.09±11.63** |
| **$PBD_m$-m4HR** | **-296.02±17.44** | **-242.54±10.02** |
| **1** | **-250.32±17.53** | **-194.50±11.92** |
| **2** | **-248.30±14.69** | **-196.38±10.90** |
| **3** | **-236.33±12.01** | **-172.39± 8.72** |

PBTOT/GBTOT = final estimated binding free energy calculated from the md course. (KCal/mol)

From the calculating result of negative total binding free energy (table S1), we knew that the binding free energy of PBD-m4HR (-279.72 Kcal/mol) and $PBD_m$-m4HR (-296.02 Kcal/mol) was notably lower than that of peptide 1-3, which means the complexes formed between PBD-m4HR or $PBD_m$-m4HR and NHR were more stable than those of peptide 1-3.

Fig S1 shows the computational model of 6-HB formed between $PBD_m$-m4HR (blue) and N36 (red) of gp41 [6]. Fig S1 A is the overview the 6-HB structure, the residues in the key *a*, *d*, *e* positions of HR peptide which directly interact with NHR are showed with stick molecule model. Fig S1 B is the local enlargement of interaction of Ile in HR peptide $a_2$ position with corresponding residues in NHR.

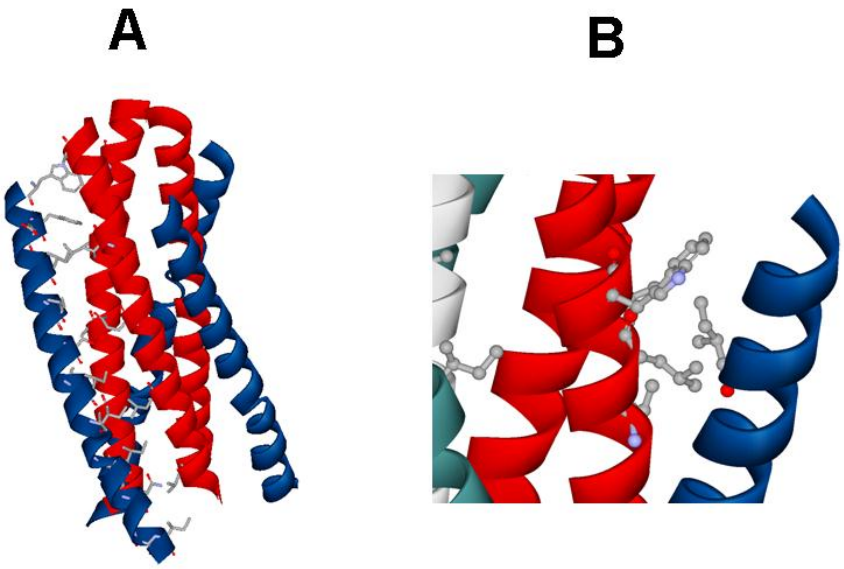


**Fig S1** The computational model of artificial HR peptide and NHR of gp41

**Native Polyacrylamide Gel Electrophoresis (N-PAGE) analysis.**

Acidic Tris-glycine gels (20%) and a BayGene Mini Cell were used for N-PAGE as described before.[7] N-peptide solutions were incubated with PBS at the indicated concentrations at 37 $^{o}$C for 30 min before addition of C-peptide (final concentration of N- and C-peptide were 100 μM). After incubation at 37 $^{o}$C for 30 min, the samples were mixed with Tris-glycine native sample buffer (1:1) and then loaded onto the gels (30 mL per well). Gel electrophoresis was carried out (120 V constant voltage, room temperature, 2.5 h), and the gel was stained with Coomassie blue R250.

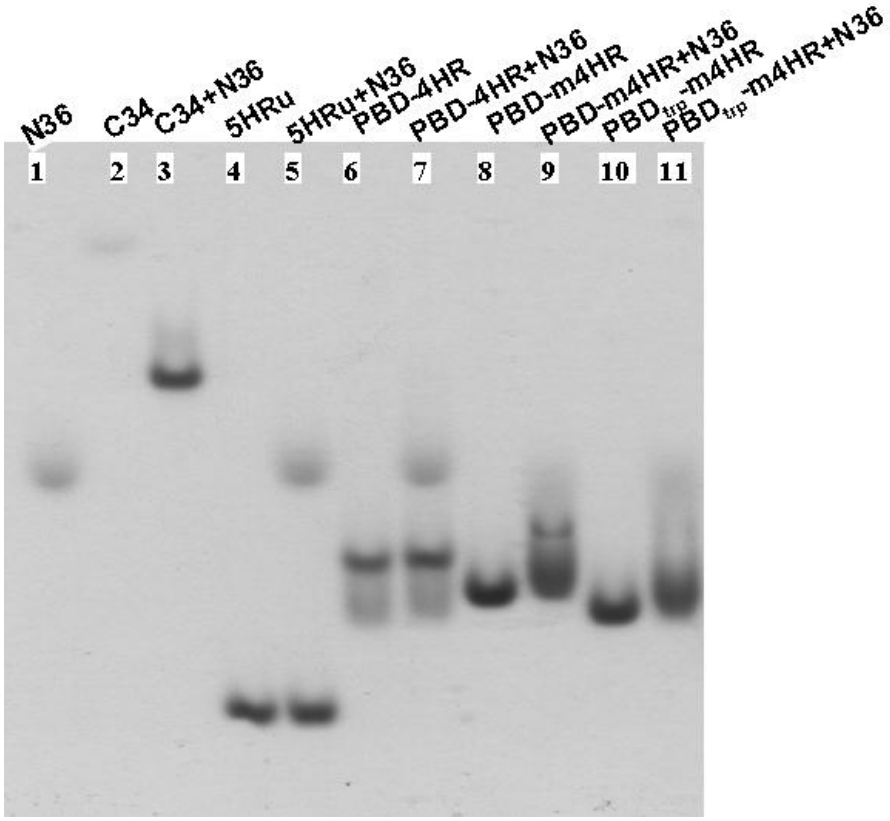


**Fig S2** Analysis of six-helix bundle formation using N-PAGE.

The interactions between artificial designed peptide fusion inhibitors with NHR were shown in **Figure S2**. N36, C34, and the artificial HR peptides exhibited a single band (lanes 1, 2, 4, 6, 8 and 10) at different positions in the acidic gel dependent on the net negative charge and the molecular size of the peptides. When C34 was mixed with N36, a new band appeared (lanes 3), indicating that they interact each other to form complex in PBS. $5HR_u$ and PBD-4HR did not interact with N36 in the acidic gel, and the N36 and the artificial peptides remained at the same position (**Figure S2**, lanes 5 and 7, respectively). However, PBD-m4HR and $PBD_m$-m4HR interact with N36 to form 6-HB like complex, as evidenced by the disappearance of the N36 and the artificial peptide bands, as well as the formation of new band with different migrate rate in the gel (**Figure S2**, lanes 9 and 11, respectively), indicating that like C34 and N36, PBD-m4HR and $PBD_m$-m4HR specifically interact with N36 to form 6-HB complexes in PBS.

**Table S2** Complete optimization process of artificial peptides. Mutations introduced that increase the potency in cell-cell fusion assay are highlighted in green, those that decrease the potency are highlighted in yellow. The “negative” mutations that changed back to design peptide

11 were showed in red.

| No. | Peptide sequence | Cell fusion $IC_{50}$ （μmol/L） |
|---|---|---|
| | $a_1$ $d_1$ $a_2$ $d_2$ $a_3$ $d_3$ $a_4$ $d_4$ $a_5$ $d_5$ | |
| $5HR_u$ | AEELAKK AEELAKK AEELAKK AEELAKK AEELAKK | 49,000±3,000 |
| PBD-$4HR_u$ | **WMEWDRE** AEELAKK AEELAKK AEELAKK AEELAKK | 27,000±1,360 |
| 1 | **WMEWDRE** **I**EELAKK AEELAKK AEELAKK AEELAKK | 8,300±490 |
| 2 | **WMEWDRE** AEELAKK **S**EELAKK AEELAKK AEELAKK | 23,000±16,000 |
| 3 | **WMEWDRE** AEELAKK AEELAKK **I**EELAKK AEELAKK | 2,100±1,010 |
| 4 | **WMEWDRE** AEELAKK AEELAKK AEE**Q**AKK AEELAKK | 7,100±3,100 |
| 5 | **WMEWDRE** AEELAKK AEELAKK AEELAKK **Q**EELAKK | 16,300±2,900 |
| 6 | **WMEWDRE** AEELAKK AEELAKK AEELAKK AEE**S**AKK | 14,400±5,200 |
| 6a | **WMEWDRE** **I**EELAKK **S**EELAKK AEELAKK AEELAKK | 5.1584±0.8493 |
| 6b | **WMEWDRE** **I**EELAKK **S**EELAKK **I**EELAKK AEELAKK | 1.6074±0.1979 |
| 6c | **WMEWDRE** **I**EELAKK **S**EELAKK **I**EE**Q**AKK AEELAKK | 12.6940±2.4426 |
| 6d | **WMEWDRE** **I**EELAKK **S**EELAKK **I**EE**Q**AKK **Q**EELAKK | 13.7187±2.5253 |
| 7 | **WMEWDRE** **I**EELAKK **S**EELAKK **I**EE**Q**AKK **Q**EE**S**AKK | 23,000±3,700 |
| 8 | **WMEWDRE** **I**EEL**I**KK **S**EELAKK **I**EE**Q**AKK **Q**EE**S**AKK | 5,600±500 |
| 9 | **WMEWDRE** **I**EEL**I**KK **S**EEL**I**KK **I**EE**Q**AKK **Q**EE**S**AKK | 43±37 |
| 10 | **WMEWDRE** **I**EEL**I**KK **S**EEL**I**KK **I**EE**QI**KK **Q**EE**S**AKK | 133±015.9 |
| PBD-m4HR | **WMEWDRE** **I**EEL**I**KK **S**EEL**I**KK **I**EE**QI**KK **Q**EE**S**I**K**K | **18.9±3.0** |
| 11 | **WMEWDRE** **I**EEL**I**KK **S**EEL**I**KK **I**EE**LA**KK AEE**L**I**K**K | 121±12.5 |
| T-20 | YTSLIHS LIEESQN QQEKNEQ ELLELDK WASLWNWF | 10.1±1.4 |
| C34 | WMEWDRE INNYTSL IHSLIEE SQNQQEK NEQELL | 1.9±0.2 |

**Physicochemical properties of peptides**

We estimated the solubility of the peptides which showed high anti-HIV activities, and compared those with a control peptide C34. Around 1mg of each peptide was weighted, and 10 μl of dd-$H_2O$ or PBS buffer were added to dissolve the samples. The solutions were centrifuged and the supernatant were recovered, and the concentration of the peptide were measured using a NanoDrop-2000c micro UV spectrophotometer (Thermo Scientific, USA) by measuring the absorbance at 280 nm. All the designed peptides were completely dissolved in both $H_2O$ and PBS, and the concentrations of the tryptothan-containing peptides measured were close to the calculated concentration from amounts of the peptides, and the concentrations of the peptides lack of tryptophan residues were directly calculated based on the weighting amounts. All the artificial designed highly potent HIV-1 fusion inhibitors showed a solubility of >100 mg/ml, and C34 showed solubility of 3.12 and 2.28 mg/ml in $H_2O$ and PBS respectively. The

physicochemical properties of the peptides are shown in **Table S3**

**Table S3** Physicochemical properties of peptides

| Peptide | Solubility in Water mg/ml | Solubility in PBS mg/ml | Helical % (C-peptide) | Helical % (C-peptide/N46) |
|---|---|---|---|---|
| 9 | >162 | >164 | | |
| PBD-m4HR | >115 | >98 | 62 | 73 |
| PBDm-m4HR | >124 | >154 | 73 | 70 |
| 12 | >140 | >99 | | |
| 13 | >146 | >108 | | |
| 15 | >196 | >228 | | |
| 14 | >153 | >160 | | |
| $PBD_{Nal}$-m4HR | >200 | >180 | NA | 74 |
| 16b | >100 | >140 | | |
| 16a | >150 | >180 | | |
| C34 | 3.12 | 2.28 | | 84 |

## References


[1] Z. Qi, W. G. Shi, N. Xue, C. G. Pan, W. G. Jing, K. L. Liu, S. B. Jiang, *Journal of Biological Chemistry* **2008**, *283*, 30376-30384.

[2] Y. Wexler-Cohen, Y. Shai, *Faseb Journal* **2007**, *21*, 3677-3684.

[3] aC. Pan, L. Cai, H. Lu, L. Lu, S. Jiang, *Journal of biological chemistry* **2011**, *286*, 28425-28434; bC. Pan, L. Cai, H. Lu, Z. Qi, S. Jiang, *Journal of Virology* **2009**, *83*, 7862-7872.

[4] L. T. Rimsky, D. C. Shugars, T. J. Matthews, *Journal of Virology* **1998**, *72*, 986-993.

[5] D. A. Case, T. A. Darden, T. e. I. Cheatham, C. l. .Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. W. Caldwell, W. S. Ross, P. A. Kollman, San Francisco, **2004**.

[6] D. C. Chan, D. Fass, J. M. Berger, P. S. Kim, *Cell* **1997**, *89*, 263-273.

[7] Y. Bai, Y. B. Ling, W. G. Shi, L. F. Cai, Q. Y. Jia, S. B. Jiang, K. L. Liu, *Chembiochem* **2011**, *12*, 2647-2658.