Antiviral Research 90 (2011) 33–41

Contents lists available at ScienceDirect

Antiviral Research

journal homepage: www.elsevier.com/locate/antiviral

Within-host co-evolution of Gag P453L and protease D30N/N88D demonstrates virological advantage in a highly protease inhibitor-exposed HIV-1 case

Junko Shibata[a,b,1], Wataru Sugiura[b,c,d,*], Hirotaka Ode[e], Yasumasa Iwatani[b,c,d], Hironori Sato[e], Hsinyi Tsang[b], Masakazu Matsuda[d,f], Naoki Hasegawa[a], Fengrong Ren[a], Hiroshi Tanaka[a]

[a] School of Biomedical Sciences, Tokyo Medical and Dental University, Tokyo, Japan
[b] Clinical Research Center, National Hospital Organization, Nagoya Medical Center, Nagoya, Japan
[c] Nagoya University Graduate School of Medicine, Nagoya, Japan
[d] AIDS Research Center, National Institute of Infectious Diseases, Tokyo, Japan
[e] Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo, Japan
[f] Mitsubishi Chemical Medience Corporation, Tokyo, Japan

ARTICLE INFO

Article history:
Received 7 September 2010
Received in revised form
28 December 2010
Accepted 11 February 2011
Available online 19 February 2011

Keywords:
HIV
Protease
Gag
Drug resistance
Co-evolution

ABSTRACT

To better understand the mechanism of HIV group-specific antigen (Gag) and protease (PR) co-evolution in drug-resistance acquisition, we analyzed a drug-resistance case by both bioinformatics and virological methods. We especially considered the quality of sequence data and analytical accuracy by introducing single-genome sequencing (SGS) and Spidermonkey/Bayesian graphical models (BGM) analysis, respectively. We analyzed 129 HIV-1 Gag–PR linkage sequences obtained from 8 time points, and the resulting sequences were applied to the Spidermonkey co-evolution analysis program, which identified ten mutation pairs as significantly co-evolving. Among these, we focused on associations between Gag-P453L, the P5′ position of the p1/p6 cleavage-site mutation, and PR-D30N/N88D nelfinavir-resistant mutations, and attempted to clarify their virological significance *in vitro* by constructing recombinant clones. The results showed that P453L^{Gag} has the potential to improve replication capacity and the Gag processing efficiency of viruses with D30N^{PR}/N88D^{PR} but has little effect on nelfinavir susceptibility. Homology modeling analysis suggested that hydrogen bonds between the 30th PR residue and the R452^{Gag} are disturbed by the D30N^{PR} mutation, but the impaired interaction is compensated by P453L^{Gag} generating new hydrophobic interactions. Furthermore, database analysis indicated that the P453L^{Gag}/D30N^{PR}/N88D^{PR} association was not specific only to our clinical case, but was common among AIDS patients.

1. Introduction

Major mutations in the human immunodeficiency virus-1 (HIV-1) protease (PR)-coding region selected by protease inhibitors (PIs) are mainly located within the active sites of the PR, and these mutations significantly reduce PR activity and viral replication capacity, i.e., viral fitness, compared to that of wild-type strains (Mahalingam et al., 1999). However, PI-resistant viruses have the potential to undergo further selection and evolution to recover their impaired PR activity and viral fitness by acquiring additional mutations not only in the PR but also in its natural substrate, the Gag protein (Nijhuis et al., 1999). In particular, mutations in Gag cleavage sites can improve the replication capacity of PI-resistant viruses. Indeed, tight associations have been demonstrated between PI-resistant mutations and Gag cleavage-site mutations, such as S373Q^{Gag} (Malet et al., 2007) and I376V^{Gag} (Ho et al., 2008) at p2/NC, A431V^{Gag} (Bally et al., 2000; Doyon et al., 1996; Gallego et al., 2003; Koch et al., 2001; Maguire et al., 2002; Malet et al., 2007; Verheyen et al., 2006; Zhang et al., 1997) at NC/p1, L449F^{Gag} (Doyon et al., 1996), and P453L^{Gag} (Bally et al., 2000; Maguire et al., 2002; Verheyen et al., 2006) at p1/p6. Gag mutations at non-cleavage sites have also been reported to improve the fitness of PI-resistant viruses (Myint et al., 2004). Thus, the selection and evolution of Gag and PR are accepted to significantly interfere with each other. This phenomenon is known as "Gag–PR co-evolution."

However, previous reports of Gag–PR co-evolution appear to have two technical limitations related to data quality and analytical method. First, the standard population-based genotyping method commonly used for determining HIV-1 variants has limited accuracy for technological reasons. To sequence the viral genome,

* Corresponding author at: Clinical Research Center, National Hospital Organization, Nagoya Medical Center, 4-1-1, San-no-maru, Naka-ku, Nagoya, Aichi 4600001, Japan. Tel.: +81 52 951 1111; fax: +81 52 963 3970.
E-mail address: wsugiura@nih.go.jp (W. Sugiura).
[1] Current address: Inserm U941, IUH, Université Paris Diderot, Hôpital Saint Louis, 75475 Paris Cedex 10, France.

0166-3542/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.antiviral.2011.02.004

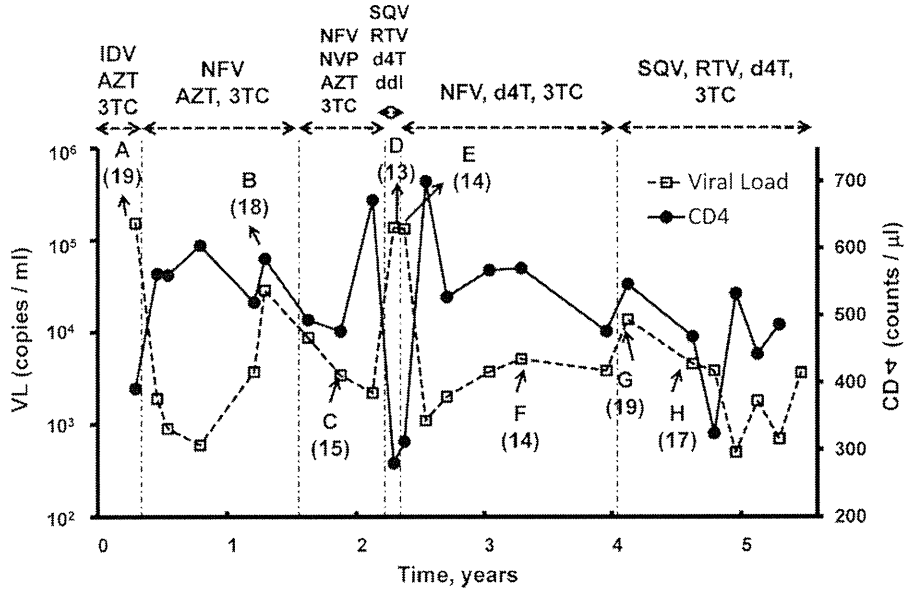


Fig. 1. Clinical course of patient-treatment protocols. Open squares and solid circles indicate plasma viral load and CD4+ cell count, respectively. Each treatment regimen is indicated in the upper part of the graph by horizontal dashed lines between arrows. Sample collection points are labelled A to H, and numbers of sequences analyzed at each point are shown in parentheses. AZT, zidovudine; d4T, stavudine; ddI, didanosine; 3TC, lamivudine; NVP, nevirapine; NFV, nelfinavir; RTV, ritonavir; SQV, saquinavir; IDV, indinavir.

viral RNA from patient plasma must be reverse-transcribed to cDNA, and the target regions are then amplified by PCR. Since viral cDNA includes multiple viral populations, using such samples as a template for the PCR step followed by sequencing results in amplification of predominantly the major variants (Gunthard et al., 1998; Hance et al., 2001), which may not represent the diversity of the original population. Furthermore, there is always a risk of artificial recombination of viral cDNAs, which are estimated to occur at rates of 4–70% during the PCR step (Meyerhans et al., 1990) and may disturb the linkage of mutation sites. Thus, this artificial linkage information may significantly affect the results of co-evolution analyses. To circumvent this technical problem, we employed the single-genome sequencing (SGS) technique based on limiting-dilution assays (Palmer et al., 2005). With SGS, artificial recombination can be avoided during PCR because cDNA samples are diluted after the reverse transcriptase reaction to a single cDNA molecule, which is used as the PCR template. Therefore, the viral sequence information obtained by SGS is not only more sensitive for analyzing HIV population diversity and mutation linkages than that obtained by standard genotype analysis, but also more precise for identifying co-evolving sites.

Second, many analytical methods have been developed for detecting co-evolving sites in molecular sequences. However, early methods did not accommodate the evolutionary history among sequences, and risked generating false-positive predictions (Altschuh et al., 1987; Gutell et al., 1992; Martin et al., 2005; Neher, 1994; Pollock and Taylor, 1997; Tillier and Lui, 2003). In recent years, the accuracy of estimating co-evolving sites has greatly improved due to several computational algorithms that incorporate the phylogenetic relationships among molecular data (Bhattacharya et al., 2007; Dutheil et al., 2005; Poon et al., 2008; Tuff and Darlu, 2000; Wollenberg and Atchley, 2000; Yeang and Haussler, 2007). In particular, Spidermonkey analysis (Poon et al., 2007a,b, 2008) not only considered the phylogenetic relationships but also implemented two empirical HIV-1 subtype B amino acid-substitution models for describing between- and within-host HIV evolution (Nickle et al., 2007).

In this study, we clarified the impact of Gag and PR co-evolution in the acquisition of drug resistance by inferring co-evolving sites between Gag and PR in a dataset of HIV-1 Gag–PR linkage sequences from a single patient who had undergone highly active antiretroviral therapy (HAART) for a long period. We used the SGS method for sequencing viral samples to avoid artificial recombination, and applied Spidermonkey analysis to our data by using its option for the within-host HIV substitution model. Virological significance of the estimated co-evolving sites was analyzed *in vitro* by constructing recombinant viruses on a pNL4-3 backbone. The replication kinetics, susceptibility to anti-HIV drugs, and the Gag processing efficiency were evaluated for each clone. We also conducted homology modeling analysis to examine Gag–PR interactions and database analysis to confirm the universality of the co-evolving sites.

2. Materials and methods

2.1. Sample collection and gag-PR-coding region sequencing by single-genome sequencing

From all HIV/AIDS patients monitored from April 1998 to August 2002 at the National Institute of Infectious Diseases in Japan, we selected a virological failure case with a history of multiple antiretroviral treatments. Of all cases, this one had been followed for the longest period and had enough detailed clinical information to perform our analysis. For the selected case, we collected plasma samples and clinical information, such as treatment regimens, changes in viral load, and CD4 counts (Fig. 1).

The single-genome sequencing (SGS) method was used as described (Palmer et al., 2005). Briefly, HIV-1 RNA was extracted from plasma samples (containing a minimum of 1000 copies of HIV-1 RNA) by guanidinium isothiocyanate treatment. cDNA was synthesized using Superscript II RT kit (Invitrogen, Carlsbad, CA) and a random hexamer. cDNA was serially diluted and amplified by PCR and nested PCR using Platinum Taq DNA Polymerase High Fidelity (Invitrogen). The endpoint of reverse-transcribed cDNA was determined as a single clone by the Poisson distribution, with cDNA dilutions yielding PCR products in 3 out of 10 reactions. The following primer sets were used in the first and nested PCR amplifications: first PCR-WGPF1 (5′-CTCTCTCGACGCAGGACTCG-

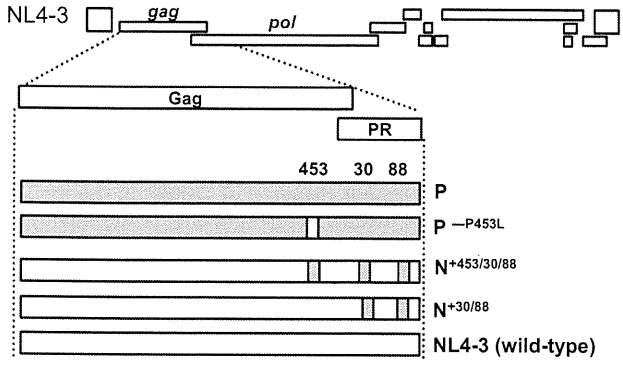


Fig. 2. Construction of recombinant viruses. Four different types of recombinant viruses were constructed according to their mutation patterns: (1) P, NL4-3 backbone with patient-derived Gag and protease (PR); (2) P^{-453}, NL4-3 backbone with patient-derived Gag and PR but $P453L^{Gag}$ was converted to wild-type; (3) $N^{+453/30/88}$, NL4-3 backbone including the substitutions $P453L^{Gag}$, $D30N^{PR}$, and $N88D^{PR}$; (4) $N^{+30/88}$, NL4-3 backbone including the substitutions $D30N^{PR}$and $N88D^{PR}$.

3′) and 3500- (5′-CTATTAAGTATTTTGATGGGTCATAA-3′), nested PCR-WGPF2 (5′-TTGCTGAAGCGCGCACGGCAAGA-3′) and 3410- (5′-CAGTTAGTGGTATTACTTCTGTTAGTGCTT-3′). These primer sets amplified a 2.7 kbps fragment containing *gag*, the PR-coding region of *pol*, and the first 900 nucleotides of the reverse transcriptase-coding region of *pol*. Positive PCR products were determined by agarose gel electrophoresis and sequenced using ABI Prism BigDye Terminator version 3.1 dideoxyterminator cycle sequencing (Applied Biosystems, Carlsbad, CA). Sequences, including the entire *gag* gene (1500 bp) and PR-coding region (297 bp), were aligned on the HXB2 reference using Sequence Navigator software (Applied Biosystems). Sequences containing mixtures at any position were excluded from analysis. PI-resistant mutations and other PR mutations in our data set were determined using the Calibrated Population Resistance tool Version 4.3 beta (http://cpr.stanford.edu/cpr/).

2.2. Detecting co-evolution using Spidermonkey analysis

Co-evolving sites between Gag and PR were inferred using Spidermonkey analysis (Poon et al., 2007a,b, 2008) to analyze all sequences (data type = protein), including the entire Gag (500 amino acids) and PR (99 amino acids) sequences. A phylogenetic tree was first reconstructed using the neighbor-joining method (Saitou and Nei, 1987). To infer co-evolving sites, we selected the within-host HIV protein-substitution model (frequencies: model-defined), based on the maximum likelihood method (Nickle et al., 2007). Amino acid sequences at MA/CA, CA/p2, p2/NC, NC/p1, and p1/p6 Gag cleavage-sites (residues P5 to P5′) and the complete PR sequence were selected for BGM analysis. We used the default options for calculations and inferred co-evolving sites if the estimated posterior probability for a pair of positions exceeded the cutoff value of 0.5.

2.3. Construction of recombinant HIV with a patient-derived gag-PR-coding region

To confirm the contributions of $P453L^{Gag}$ in the context of patient-derived *gag*-PR-coding region, three types of recombinant viruses were constructed (Fig. 2). Recombinant viruses were constructed using the pNL4-3 molecular clone of HIV-1 (GenBank accession No. AF324493). A patient-derived *gag*-PR-coding region containing $P453L^{Gag}$, $D30N^{PR}$, and $N88D^{PR}$ was chosen from sampling point F. The recombinants were (1) P, NL4-3 backbone with patient-derived *gag*-PR-coding region; (2) P^{-P453L}, NL4-3 backbone with patient-derived *gag*-PR-coding region but $P453L^{Gag}$ was converted to wild-type.

Details of the recombinant virus construction were as follows. The patient-derived 1.9 kbp *gag*-PR-coding region was subcloned into pCR4-TOPO (Invitrogen) (designated $pCR4\text{-}TOPO^{Patient}$). Using $pCR4\text{-}TOPO^{Patient}$ as a template, $P453L^{Gag}$ was substituted with wild-type P453. Each substitution was introduced using the following primer sets: For the Gag-P453 substitution, forward: 5′-TTCAGAACAGACCAGAGCCATCAGCT-3′, reverse: 5′-AGCTGATGGCTCTGGTCTGTTCTGAA-3′. Subsequently, the substituted *gag*-PR-coding region was amplified using the following primers: WGPF2: 5′-TTGCTGAAGCGCGCACGGCAAGA-3′ and DRPRO6: 5′-ACTTTTGGGCCATCCATTCCTGGCTTT-3′. The NL4-3-derived RT-coding region was amplified using the following primers: RT-63F: 5′-TAAACAATGGCCATTGACAGAAG-3′ and RT-898R: 5′-CTGCTTCTTCTGTTAGTGGTACTAC-3′. Primers DRPRO6 and RT-63F are phosphorylated at their 5′ ends. Both resulting fragments were ligated and digested with BssHII and SbfI, and cloned back into pNL4-3. All pNL4-3-based recombinant DNAs (3.75 μg) were transfected into 2×10^5 HeLa cells using Fugene6 (Roche, Indianapolis, IN), and culture supernatants were harvested at 72 h after transfection, filtered through a 0.45 μm membrane, assayed for reverse transcriptase (RT) activity (Willey et al., 1988), and kept as virus stocks at −80 °C until use. Each virus stock (5×10^6 ^{32}P cpm of RT activity) was used for replication kinetics analyses.

2.4. Construction of pNL43 with $P453L^{Gag}$, $D30N^{PR}$, and $N88D^{PR}$ by site-directed mutagenesis

In addition to constructing patient-derived *gag*-PR-coding region recombinant viruses, we evaluated the effect of $P453L^{Gag}$, $D30N^{PR}$, and $N88D^{PR}$ interference by constructing three types of pNL4-3-based recombinant viruses (Fig. 2). These were (1) $N^{+453/30/88}$, NL4-3 including $P453L^{Gag}/D30N^{PR}/N88D^{PR}$; and (2) $N^{+30/88}$, NL4-3 including $D30N^{PR}/N88D^{PR}$. These recombinant viruses were constructed as follows. pCR4-TOPO (Invitrogen) including the ApaI-SbfI fragment of pNL4-3, which contained the complete p1/p6 of *gag* and PR-coding region, was constructed as a template for further mutagenesis. This construct was designated as $pCR4\text{-}TOPO/NL^{ApaI\text{-}SbfI}$. Using $pCR4\text{-}TOPO/NL^{ApaI\text{-}SbfI}$ as a template, we introduced three mutations ($P453L^{Gag}$, $D30N^{PR}$, and $N88D^{PR}$) by site-directed mutagenesis. The following primer sets were used to introduce each mutation: For $P453L^{Gag}$, forward: 5′-TTCAGAGCAGACTAGAGCCAACAGCC-3′ and reverse: 5′-GGCTGTTGGCTCTAGTCTGCTCTGAA-3′. For $D30N^{PR}$, forward: 5′-CAGGAGCAGATAATACAGTATTAGAAG-3′ and reverse: 5′-CTTCTAATACTGTATTATCTGCTCCTG-3′. For $N88D^{PR}$, forward: 5′-CATAATTGGAAGAGATCTGTTGACTC-3′ and reverse: 5′-GAGTCAACAGATCTCTTCCAATTATG-3′. After each mutagenesis reaction, the entire sequences were verified and cloned back into pNL4-3 using ApaI and SbfI restriction enzymes. All pNL4-3-based recombinant DNAs (3.75 μg) were transfected into 2×10^5 HeLa cells using Fugene6 (Roche), and culture supernatants were harvested at 72 h after transfection, filtered through a 0.45 μm membrane, assayed for RT activity (Willey et al., 1988), and kept as virus stocks at −80 °C until use. Each virus stock (4×10^5 ^{32}P cpm of RT activity) was used for replication kinetics analysis.

2.5. Replication kinetics of recombinant viruses

Replication kinetics of recombinant viruses was evaluated as described (Matsuoka-Aizawa et al., 2003). Briefly, 5×10^4 MT-2 cells were infected with each virus stock in the absence or presence of 0.1 μM nelfinavir at 37 °C for 16 h. Cells were then washed once and resuspended in 0.5 ml culture medium with the same concen-

tration of nelfinavir, and cultures were maintained for 12–22 days, changing half of the medium every 2 or 3 days. The titer of each virus was evaluated by both RT activity and p24 amount; as both measures demonstrated good correlation, p24 amount was used to adjust the virus inoculum. Culture supernatants were collected, and residual supernatants were kept at −80 °C until use. Replication kinetics was independently analyzed two times.

2.6. Evaluation of nelfinavir susceptibilities of recombinant viruses

Recombinant viruses were evaluated for nelfinavir susceptibility using an in-house drug susceptibility assay with the MaRBLE cell line (MaRBLE assay) as described (Chiba-Mizutani et al., 2007). Briefly, 1×10^5 MaRBLE cells were infected with 100 $CCID_{50}$ of each recombinant virus, and virus replication was monitored in serial dilutions of nelfinavir from 1.28×10^{-13} to 1.0×10^{-6} M for 7 days in triplicate. At day 7, cells were harvested and lysed in luciferase assay reagent. Firefly (FF) and renilla luciferase (RF) activities produced by the cells were quantified using a Dual-luciferase Reporter Assay System (Promega, Madison, WI). The relative virus replication rate (% replication) at each drug concentration was calculated by the following formula: % replication = (observed FF luciferase activity with drug − background [mock] FF luciferase activity)/(observed FF luciferase activity without drug − background [mock] FF luciferase activity) × 100. IC_{50} values were calculated with 95% confidence intervals using GraphPad Prism software and nonlinear regression analysis fitted with a sigmoidal dose–response curve with variable slope.

2.7. Analysis of Gag processing patterns in recombinant viruses

The Gag processing patterns of recombinant viruses were analyzed as described (Sugiura et al., 2002) with minor modifications. In brief, NL4-3-based recombinant DNAs (18 μg) was transfected into 5×10^6 HeLa cells using Fugene6 (Roche). The culture supernatants were harvested at 48 h after transfection with a culture medium change at 12 h post transfection to the absence or presence of 0.1 μM nelfinavir. For pelleting virus through a sucrose cushion, 25 ml of cell culture medium was layered onto 10 ml of 20% sucrose (wt/vol, in PBS) before centrifugation at 30,000 rpm in a swing rotor for 1.5 h. The medium and cushion were discarded, and the virus pellet was dissolved in 200 μl Laemmli sample buffer (Bio-Rad, Hercules, CA). Viral supernatants were normalized using PETRO-TEK HIV-1 p24 Antigen ELISA (ZeptoMetrix Corporation, Buffalo, NY) and subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE). Proteins in the gels were passively transferred to PVDF membranes (Bio-Rad). The membranes were incubated with a chicken anti-p6 polyclonal antibody (Sigma–Aldrich Corporation, St. Louis, MO) and a murine anti-p24 monoclonal antibody (ZeptoMetrix Corporation) for 2 h, followed by incubation with secondary antibody, a chicken HRP-conjugated, IgG antibody (Bethyl Laboratories Inc., Montgomery, TX), a mouse HRP-conjugated, IgG antibody (Thermo Fisher Scientific, Yokohama, JP), respectively. Finally, proteins were visualized using SuperSignal West Dura Extended Duration Substrate (Thermo Fisher Scientific).

2.8. Molecular modeling of Gag p1/p6 peptide–PR complexes

We constructed three-dimensional models of PR in complex with peptide representing Gag-p1/p6 substrate by a homology modeling method (Baker and Sali, 2001; Marti-Renom et al., 2000; Shirakawa et al., 2008) using Molecular Operating Environment (MOE) ver. 2008.10 (http://www.chemcomp.com/, Chemical Computing Group Inc., Montreal, Quebec, Canada). For the modeling template, we used an X-ray crystal structure of inactive D25N PR in complex with the p1/p6 substrate at 2 Å resolution (PDB code: 1KJF) (Prabu-Jeyabalan et al., 2002), as it had the highest similarity (94.2% identity) to the HIV-1 NL4-3 strain among HIV-1 protease-p1/p6 peptide structures in the protein data bank, even though this is an inactive model. Furthermore, the active-site $D25N^{PR}$ mutation has been reported to hardly influence the structure of the protease in complex with ligands (Sayer et al., 2008). In the models, nine-amino-acid-length peptides corresponding to the Gag p1/p6 of the NL4-3, $N^{+30/88}$, and $N^{+453/30/88}$ strains were bound to the catalytic sites of PRs of the same strains. We considered effects of a water molecule (HOH11) that mediates important hydrogen bonds between the Gag p1/p6 peptide and the PR. AMBER ff99 force field (Wang et al., 2000) and generalized Born/volume integral (GB/VI) implicit solvent model (Labute, 2008) were applied for intra- and inter-molecular energy calculations.

2.9. Database analysis

To confirm the universality of $P453L^{Gag}/D30N^{PR}/N88D^{PR}$, we obtained 3249 sequences of HIV-1 subtype B *gag*-PR-coding region (positions 2146–2516) from the Los Alamos National Laboratory HIV sequence database (http://www.hiv.lanl.gov/). For the dataset, we applied Fisher's exact test and investigated associations between the D30N/N88D mutations in PR and the P453L mutation in Gag and between the 30/88 mutations in PR and the P453L mutation in Gag, respectively. Fisher's exact test was implemented using the R 2.6.1 statistical package.

3. Results

3.1. The majority of patient-derived Gag cleavage-site mutations are in p2/NC and p1/p6

Twenty-three plasma samples were serially collected from a patient who had received HAART for 52 months. The patient's clinical history, changes in viral load, and CD4 counts are depicted in Fig. 1. At 8 sampling points (A–H), 129 *gag*-PR-coding region sequences were obtained, with p2/NC and p1/p6 Gag cleavage-site mutations observed at each point in more than 60% of clones (Table 1). PI-resistant mutations and other PR mutations are also summarized in Table 1. None of the clones had CA/p2 cleavage-site mutations, and only a few clones had MA/CA or NC/p1 cleavage-site mutations. The $Y132F^{Gag}$ mutation in MA/CA was found in 1.6% of the clones, and $E428D^{Gag}$ and $R429K^{Gag}$ within NC/p1 had prevalences of 1.6% and 5.4%, respectively. On the other hand, cleavage-site mutations were frequently observed within p2/NC and p1/p6. In the p2/NC site, we observed 11 mutations: $S373Q^{Gag}$ (83.7%), $S373P^{Gag}$ (15.5%), $A374G^{Gag}$ (79.8%), $A374V^{Gag}$ (16.3%), $A374S^{Gag}$ (2.3%), $A374R^{Gag}$ (0.8%), $T375N^{Gag}$ (100%), $I376V^{Gag}$ (3.9%), $M378L^{Gag}$ (0.8%), $G381S^{Gag}$ (0.8%), and $N382Y^{Gag}$ (0.8%). In the p1/p6 site, we observed three mutations: $L449F^{Gag}$ (0.8%), $S451N^{Gag}$ (100%), and $P453L^{Gag}$ (65.1%).

3.2. Ten Gag–protease co-evolving sites are inferred by Spidermonkey analysis

Among the 129 *gag*-PR-coding region sequences analyzed by Spidermonkey analysis (Poon et al., 2007a,b, 2008), ten co-evolving sites were inferred. These sites were identified using the default cutoff posterior probability (pp) value of 0.5. The ten co-evolving pairs identified with pp > 0.5 are shown in Table 2. Four pairs, $R429K^{Gag}/M36V^{PR}$ (pp = 0.87), $S373Q^{Gag}/T12A^{PR}$ (pp = 0.83), $P453L^{Gag}/D30N^{PR}$ (pp = 0.63), and $P453L^{Gag}/N88D^{PR}$ (pp = 0.61), represented Gag/PR inter-molecular co-evolution;

Table 1
Gag cleavage-site mutations and PR mutations from a HAART-treated case.

Sampling point	Gag cleavage-site mutation (%)[a]		PR mutation (%)[a]	
	p2/NC	p1/p6	PI-resistant mutations	Other PR mutations
A to H (n = 129)	T375N	S451N	–	I62V, L63P,A71T, I93L
A (n = 19)	S373Q(100), A374G (89),	–	–	M36I(95), I72V(11), V77I(100)
B (n = 18)	S373Q(100), A374G (94)	P453L(100)	D30N(100), M46I(6), N88D(100)	E35D(94), M36I(100), V77I(100),
C (n = 15)	S373Q(100), A374G(100)	P453L(100)	D30N(100), N88D(100)	L10F(80), E35D(100), M36I(100), K45R(27), I72T(7), V77I (100),
D (n = 13)	S373P (85), A374V (85)	P453L (8)	D30N (8), N88D (8)	L10F(8), V11I (8), T12A(8), K20R(8), E35D(8), M36I(100), H69Y(8), V77I(8)
E (n = 14)	S373P (64), A374V(71)	–	–	T12A(7), M36I(100), K55N (7), V77I(7)
F (n = 14)	S373Q(100), A374G(100)	P453L(100)	D30N(100), N88D(100)	L10F(100), I13V(79), E34G(7), E35D(100), M36I (100), N37T (86), K45R (14), Q58E (86), I72T (7), V77I (100)
G (n = 19)	S373Q(100), A374G(100)	P453L(100)	D30N(100), N88D(100)	L10F(100), I13V (68), E35D(100), M36I(68), M36V(32), N37T(100), Q58E(100), V77I(100)
H (n = 17)	S373Q(100), A374G(100)	P453L(100)	D30N(100), I54V(76), N88D(100), L90M (47)	L10F(100), I13V (94), K20R (76), E35D (100), M36I (94), M36V (6), N37T (100), Q58E (100), I72T (53), V77I (100), G78R (6)

Gag mutations refer to HXB2 and PI-resistant mutations and other PR mutations were determined using the Calibrated Population Resistance tool Version 4.3 beta.
[a] Numbers in parentheses are the percentages of mutations at each sampling point.

one pair, $S373Q^{Gag}/A374G^{Gag}$, represented Gag/Gag intra-molecular co-evolution; and the other five pairs, $N37T^{PR}/Q58E^{PR}$ (pp = 0.94), $E35D^{PR}/M46I^{PR}$ (pp = 0.89), $K20R^{PR}/I54V^{PR}$ (pp = 0.88), $V11I^{PR}/K20R^{PR}$ (pp = 0.86), and $D30N^{PR}/N88D^{PR}$ (pp = 0.62), represented PR/PR intra-molecular co-evolution.

We focused on three pairs, $P453L^{Gag}/N88D^{PR}$, $P453L^{Gag}/D30N^{PR}$, and $D30N^{PR}/N88D^{PR}$, because $D30N^{PR}$ and $N88D^{PR}$ are well-known major and minor nelfinavir-resistant mutations, respectively (Johnson et al., 2008), and $P453L^{Gag}$ is the P5′ position of the p1/p6 cleavage-site mutation. Although $D30N^{PR}$ has been associated with $N88D^{PR}$ (Rhee et al., 2007; Wu et al., 2003), the interactions among $P453L^{Gag}$, $D30N^{PR}$, and $N88D^{PR}$ have not been investigated. Since $P453L^{Gag}/D30N^{PR}/N88D^{PR}$ was frequently observed in the presence of nelfinavir (Fig. 1 and Table 1), we conducted *in vitro* experiments to confirm whether the co-existence of $P453L^{Gag}/D30N^{PR}/N88D^{PR}$ has a virological advantage in the presence of nelfinavir.

3.3. $P453L^{Gag}$ improves the replication capacity of viruses with $D30N^{PR}/N88D^{PR}$ in both patient- and NL4-3-derived genetic backgrounds

To evaluate the virological impact of $P453L^{Gag}$ in the patient-derived genetic background, we constructed two types of patient-derived *gag*-PR-coding region viruses, P and P^{-P453L} (Fig. 2). These two recombinant viruses and the wild-type virus (NL4-3) were cultured independently in the absence or presence of nelfinavir, and their replication kinetics was monitored by measuring RT activity in culture supernatants. Assays for replication kinetics were independently performed twice, confirming identical orders of replication kinetics.

Table 2
Positions of coevolving pairs inferred by Spidermonkey analysis.

Position	Position	Expected posterior probability	Total number of sequences
429^{Gag}	36^{PR}	0.866	7
373^{Gag}	12^{PR}	0.825	2
453^{Gag}	30^{PR}	0.632	84
453^{Gag}	88^{PR}	0.607	84
373^{Gag}	374^{Gag}	0.895	22
37^{PR}	58^{PR}	0.938	48
35^{PR}	46^{PR}	0.887	1
20^{PR}	54^{PR}	0.877	13
11^{PR}	20^{PR}	0.858	1
30^{PR}	88^{PR}	0.618	84

Significant coevolving sites (pp value > 0.5) are shown.

In the absence of nelfinavir (Fig. 3A), the RT activities of the NL4-3 and P viruses peaked at 6 days after infection, and the RT activity of NL4-3 was higher than that of the P virus, whereas viral replication was delayed in the P^{-P453L} virus (i.e., P virus without $P453L^{Gag}$), and its RT activity peaked at 10 days after infection. The order of replication kinetics in the absence of nelfinavir was wild-type (NL4-3) > P > P^{-P453L}. On the other hand, in the presence of nelfinavir (Fig. 3B), replication of the wild-type virus was completely suppressed, and replication of P virus was the most active, demonstrating peak RT activity at 6 days after infection. The P^{-P453L} virus showed lower replication capacity than the P virus. Thus, the order of replication kinetics in the presence of nelfinavir was P > P^{-P453L} > wild-type (NL4-3).

To assess whether the complex $P453L^{Gag}/D30N^{PR}/N88D^{PR}$ conferred an advantage not only in the patient-derived genetic background but also in the HIV-1 molecular clone (NL4-3)-derived genetic background, we constructed two types of NL4-3-based *gag*-PR-coding region recombinant viruses, $N^{+453/30/88}$ and $N^{+30/88}$ (Fig. 2). The results of independent culture studies are shown in Fig. 3C and D. In the absence of nelfinavir, the RT activities of NL4-3, NL4-3 with $P453L^{Gag}/D30N^{PR}/N88D^{PR}$ ($N^{+453/30/88}$) and NL4-3 with $D30N^{PR}/N88D^{PR}$ ($N^{+30/88}$) viruses peaked at 11 days, 14 days, and 16 days after infection, respectively, and the order of replication kinetics was NL4-3 > $N^{+453/30/88}$ > $N^{+30/88}$. On the other hand, in the presence of nelfinavir, the $N^{+453/30/88}$ virus grew the most actively with its peak RT activity at 14 days after infection. The $N^{+30/88}$ virus was the second most actively replicating, and the wild-type did not replicate. Thus, the order of replication kinetics in the presence of nelfinavir was $N^{+453/30/88}$ > $N^{+30/88}$ > NL4-3.

In both studies, not only in the patient-derived genetic background but also in the NL4-3-derived genetic background, the virus with $D30N^{PR}/N88D^{PR}$ showed lower replication capacity than the virus with $P453L^{Gag}/D30N^{PR}/N88D^{PR}$, suggesting that $P453L^{Gag}$ significantly contributes to the fitness recovery of virus with $D30N^{PR}/N88D^{PR}$.

3.4. $P453L^{Gag}$ does not influence susceptibility to nelfinavir

To clarify whether $P453L^{Gag}$ affects nelfinavir susceptibility, the IC_{50}s of nelfinavir for NL4-3-based *gag*-PR-coding region recombinants were determined using MaRBLE cells. Both $N^{+453/30/88}$ and $N^{+30/88}$ recombinants showed 14.4 and 15.8-fold greater resistance to nelfinavir than wild-type virus, respectively (Table 3). However, the difference in IC_{50} between the two recombinants was

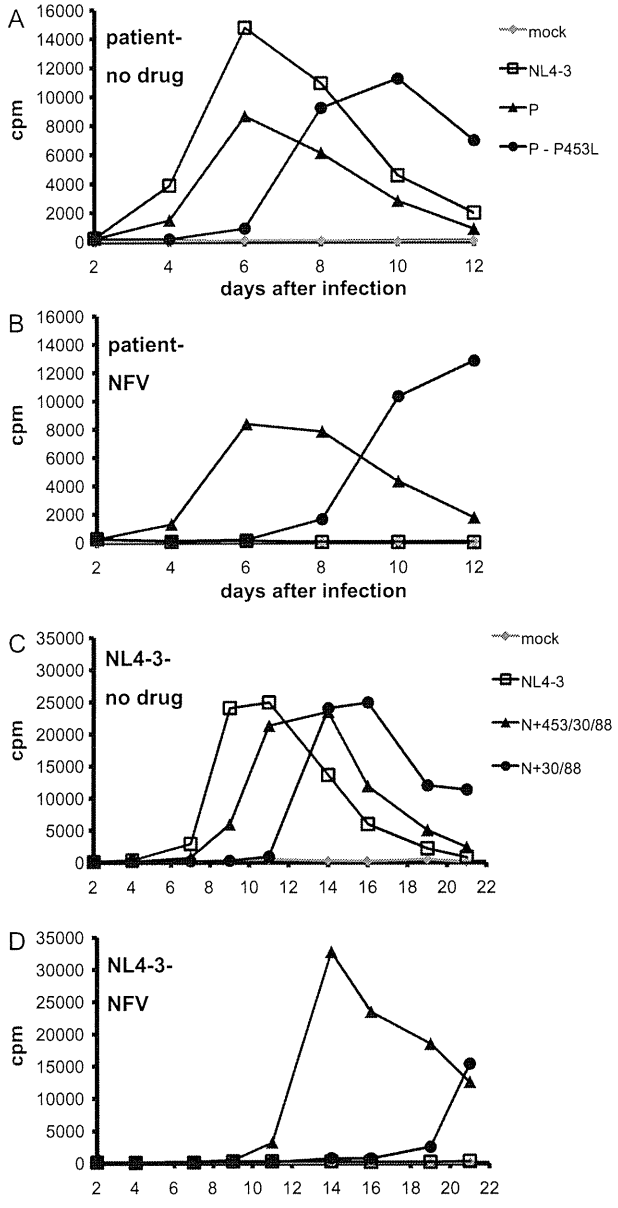


Fig. 3. Replication kinetics of recombinants. MT-2 cells were infected with patient-derived *gag*-PR-coding region, (A) in the absence of, and (B) in the presence of 0.1 μM nelfinavir. Open squares, solid triangles, and solid circles indicate wild-type NL4-3, NL4-3 with patient *gag*-PR-coding region insert, and NL4-3 with patient insert without P453L substitution, respectively. MT-2 cells were also infected with NL4-3-based recombinant, (C) in the absence of, and (D) in the presence of 0.1 μM nelfinavir. Open squares, solid triangles, and solid circles indicate wild type NL4-3; NL4-3 with P453L, D30N, and N88D; and NL4-3 with D30N and N88D, respectively. Diamonds indicate mock-infected controls. Assays were independently performed twice, and one representative set of results is shown.

not significant, suggesting that P453L^{Gag} had no effect on nelfinavir susceptibility.

3.5. *P453L^{Gag} improves Gag p1/p6 processing in virus with D30N^{PR}/N88D^{PR}*

To gain further insights into the virological effects of P453L^{Gag}, we examined Gag processing patterns in the absence and presence of 0.1 μM nelfinavir by Western blot analysis with an anti-p6 polyclonal antibody and an anti-p24 monoclonal antibody. The amount

Table 3
Nelfinavir susceptibilities of recombinant viruses.

HIV-1	IC_{50} (nM)	95% confidence interval	Fold-resistance
NL4-3	1.3	0.8–2.4	1.0
NL$^{+453/30/88}$	18.7	9.2–37.8	14.4
N$^{+30/88}$	20.6	12.6–33.6	15.8

of sample loaded in each lane was normalized by p24 antigen content (600 ng for each lane as determined by ELISA) (Fig. 4A). In the absence of nelfinavir, the partially cleaved Gag intermediate p15 (including NC, p1, and p6) of NL4-3 was efficiently cleaved to the p6 peptide (Fig. 4B, lane 3). In contrast, the processing of NL4-3 was less efficient in the presence of nelfinavir, as indicated by the accumulation of p15 (Fig. 4B, lane 4). N$^{+30/88}$ showed defects in cleavage at the p1/p6 site, as demonstrated by the accumulation of p15 and p7 (p1/p6) (Fig. 4B, lanes 5 and 6). On the other hand, lesser p15 and p7 accumulated in N$^{+453/30/88}$ than in N$^{+30/88}$ (Fig. 4, lanes 7 and 8). Interestingly, a 8–9 kDa band, which is neither p6 nor p7, was observed in N$^{+30/88}$ (Fig. 4B, lanes 5 and 6).

3.6. *Impaired Gag–PR affinity in the N$^{+30/88}$ strain is recovered by new L453^{Gag} interactions with M46$'^{PR}$ and F53$'^{PR}$*

To elucidate the structural impact of the mutations described above on interactions between Gag-p1/p6 substrate and PR, we generated three-dimensional models of the PR in complex with peptide representing Gag-p1/p6 substrate by homology modeling (Baker and Sali, 2001; Marti-Renom et al., 2000; Shirakawa et al., 2008) using Gag and PR sequences of the NL4-3, N$^{+30/88}$, and N$^{+453/30/88}$ strains. Comparison of the thermodynamically optimized models showed obvious differences in interactions between side chains of PR and p1/p6 (Fig. 5). First, D30N^{PR} mutation resulted in fewer hydrophilic interactions between side chains of the 30th PR and 452nd p1/p6 residues; NL4-3 had two hydrogen bonds between the side chains of D30^{PR} and R452^{Gag} (Fig. 5A), while N$^{+30/88}$ and N$^{+453/30/88}$ had only a single hydrogen bond between the side chains of N30^{PR} and R452^{Gag} (Fig. 5B and C). Second, the P453L^{Gag} mutation in the p1/p6 substrate of the N$^{+453/30/88}$ strain led to new hydrophobic interactions between side chains of the 46th PR and 453rd p1/p6 residues, as well as of the PR 53F and p1/p6 P453L^{Gag} residues (Fig. 5C).

To examine whether these changes in interactions influenced the binding affinity of the p1/p6 substrate to PR, we analyzed

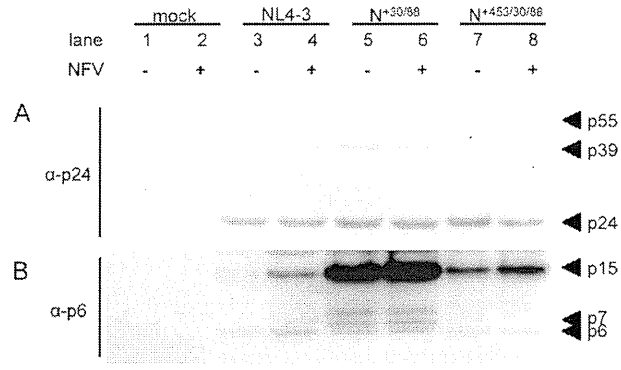


Fig. 4. Western blot analysis of Gag processing in the absence or presence of nelfinavir. Western blot analyses in the absence or presence of nelfinavir. HeLa cells were transfected by each recombinant clones and cultured in the absence or presence of NFV (0.1 μM). At 48 h post-transfection, virions in culture supernatants were harvested and subjected to Western blot analysis with anti-p24 monoclonal antibody (A) and an anti-p6 polyclonal antibody (B). Each lane was normalized by p24 antigen content (5 ng for each lane as determined by ELISA). Lanes 1 and 2, mock, lanes 3 and 4, NL4-3, lanes 5 and 6, N$^{+453/30/88}$, lanes 7 and 8, N$^{+30/88}$.

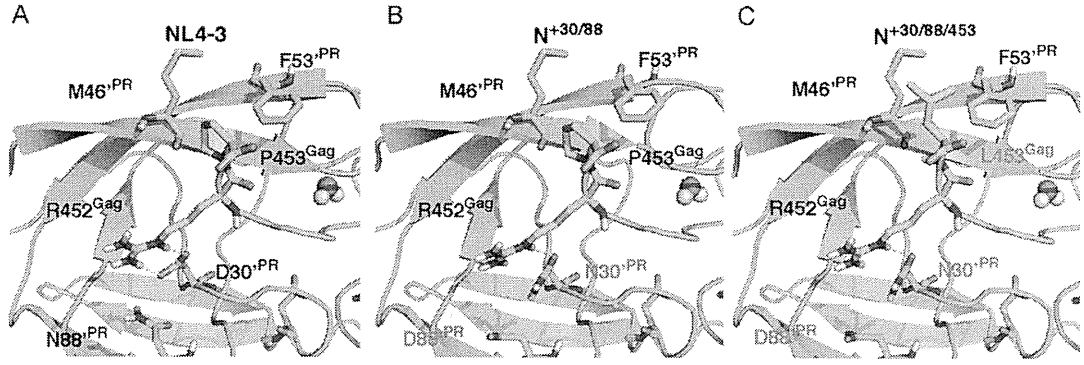


Fig. 5. Structural models of the Gag–PR complexes. Catalytic sites of the NL4-3 (A), $N^{+30/88}$ models (B), and $N^{+453/30/88}$ models (C) are highlighted. Green cartoons and sticks represent main and side chains of PR, respectively. Cyan sticks represent 452nd and 453rd residues in Gag corresponding to the p1/p6 region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

their binding energies using the PR–p1/p6 peptide complex models. The predicted binding energies of the NL4-3, $N^{+30/88}$ and $N^{+453/30/88}$ models were −137.6 kcal/mol, −133.3 kcal/mol, and −137.1 kcal/mol, respectively, suggesting that p1/p6 substrate has lower affinity with $N^{+30/88}$ PR than with NL4-3 and $N^{+453/30/88}$ PRs. Taken together, these data suggest that PR mutations in the $N^{+30/88}$ strain can reduce Gag–PR affinity primarily via loss of the hydrogen bond between $N30^{PR}$ and $R452^{Gag}$ and that the Gag p1/p6 mutation in the $N^{+453/30/88}$ strain ($P453L^{Gag}$) can recover affinity by generating new hydrophobic interactions of $L453^{Gag}$ with $M46^{PR}$ and $F53^{PR}$.

3.7. $P453L^{Gag}/D30N^{PR}/N88D^{PR}$ association is commonly observed in a large database

To confirm the prevalence of the $P453L^{Gag}/D30N^{PR}/N88D^{PR}$ association in other HIV-infected individuals, we investigated 3249 sequences of HIV-1 subtype B *gag*-PR-coding region from the Los Alamos National Laboratory HIV sequence database (http://www.hiv.lanl.gov/). We found that the $P453L^{Gag}$ mutation was significantly associated with $D30N^{PR}/N88D^{PR}$ (Table 4; $p < 0.001$, Fisher's exact test). These data support the virological advantage of the $P453L^{Gag}/D30N^{PR}/N88D^{PR}$ association.

4. Discussion

In this study, we analyzed mechanisms of anti-HIV drug-resistant mutation acquisition by investigating crosstalk between Gag and PR mutations. We traced the clinical course and sequence changes in *gag* and the PR-coding region of a virological-failure case heavily treated with multiple regimens, including different protease inhibitors. To focus on the quality of sequence data and accuracy of analysis, we used SGS and Spidermonkey analysis, respectively.

Among the ten co-evolving Gag–Protease pairs inferred by Spidermonkey analysis, we confirmed a linkage between $P453L^{Gag}$ and $D30N^{PR}/N88D^{PR}$. $D30N^{PR}$ has been reported to associate with $N88D^{PR}$ (Rhee et al., 2007; Wu et al., 2003), P453L (Verheyen et al., 2006), and positively correlate with p1/p6 cleavage-site mutations (Kolli et al., 2006). We were interested in the association between $P453L^{Gag}$ and $D30N^{PR}/N88D^{PR}$ nelfinavir-resistant mutations, as $P453L^{Gag}$ is located at the P5′ position of the p1/p6 cleavage site, and expected to physically interact with the protease. Thus, we sought to clarify the biological advantage of interference among $P453L^{Gag}$, $D30N^{PR}$, and $N88D^{PR}$ in recombinant viruses with patient- and NL4-3-derived genetic backgrounds. Virological advantage was evaluated in three aspects: (1) virological superiority in replication competency, (2) resistance to antiretroviral selective pressure, and (3) Gag processing pattern in virions. Our results indicated that the $P453L^{Gag}$ cleavage-site mutation has the potential to improve the replication capacity and Gag processing of viruses with $D30N^{PR}/N88D^{PR}$, but has little effect on nelfinavir susceptibility. This latter finding is of interest since Gag cleavage-site mutations have been suggested as a mechanism for protease to develop drug resistance (Dam et al., 2009; Kolli et al., 2009). We also need to consider not only antiretroviral selective pressure but also immune selective pressure. Several of the positions we noted have been described as associated with human leukocyte antigen (HLA) escape mutations. For example, PR codons 12, 35, and 36, and Gag codons 373 and 374 are all potentially HLA-related (Brumme et al., 2009). To confirm the contributions of HLA and immune pressure, further study is required.

Although samples were collected chronologically at multiple times, the order of $P453L^{Gag}$, $D30N^{PR}$ and $N88D^{PR}$ acquisitions was unclear as all three mutations were detected at the same time. However, a plausible order seems to be the selection of $D30N^{PR}/N88D^{PR}$ followed by $P453L^{Gag}$ acquisition. Nelfinavir appears to select $D30N^{PR}/N88D^{PR}$ mutations for resistance because these mutations obviously increase drug resistance to nelfinavir (Johnson et al., 2009), whereas the $P453L^{Gag}$ mutation without any PR mutations has been reported to have almost no effect on susceptibility to PIs and on viral replication capacity (Maguire et al., 2002). Although $D30N^{PR}$ is known as one of the most unstable PI-resistant mutations (Martinez-Picado et al., 1999) and viruses with this mutation have lower PR activity than the wild-type, the impaired replication caused by $D30N^{PR}$ has been reported to be compensated by $N88D^{PR}$ (Mitsuya et al., 2006; Sugiura et al., 2002). Furthermore, the fitness of virus with $D30N^{PR}/N88D^{PR}$ was recovered in our study by an additional $P453L^{Gag}$ mutation (Fig. 3). However, the $P453L^{Gag}$ mutation was not introduced into NL4-3 carrying $D30N^{PR}/N88D^{PR}$ ($N^{+30/88}$) and patient $D30N^{PR}/N88D^{PR}$ clones (P^{-P453L}) during *in vitro* culture with nelfinavir, suggesting that the $P453L^{Gag}$ mutation is a sufficient condition for D30N/N88D clones to replicate efficiently.

Virus with $D30N^{PR}/N88D^{PR}$ was suggested by results of our Western blot analyses to process p1/p6 cleavage inefficiently, as

Table 4
Los Alamos National Laboratory HIV sequence database analysis.

453^{Gag}	$30^{PR}/88^{PR}$	
	D30/N88	N30N/D88D
P ($n=2801$)	$n=2743$	$n=8$
L ($n=237$)	$n=223$	$n=8$
p-value		<0.001

demonstrated by the accumulation of p15 and p7 non-cleaved precursors, but addition of P453L^{Gag} improved the processing (Fig. 4B). Accumulation of p15 and p7 products has been reported in previous studies using different PR mutant viruses (Doyon et al., 1996; Maguire et al., 2002). In these studies, the additional mutations at NC/p1 or p1/p6 cleavage sites also resulted in efficient processing of these precursors. Interestingly, an aberrant band, which did not match either p6 or p7, was observed in N$^{+30/88}$ in our study (Fig. 4, lanes 5 and 6). Although further analyses will be required to determine the exact mechanisms, the band suggests inaccurate or alternative recognition of the cleavage site by virus with D30N^{PR}/N88D^{PR}, and P453L^{Gag} may confer an advantage by adjusting the protease to recognize and cleave the right site. As the HIV-1 p6 protein is important for efficient particle budding (von Schwedler et al., 2003), the defect in p1/p6 cleavage may affect viral maturation, which in turn may reduce viral infectivity and replication capacity.

To understand the relevance of P453L^{Gag} from a structural viewpoint, we used homology modeling with the published X-ray crystal structure of the PR–p1/p6 substrate complex as a template (Fig. 5). Although P453L^{Gag} is located at the P5′ position and does not directly interfere with the protease active site or subsites, the modeling demonstrated that P453L^{Gag} can compensate for the binding affinity of PR and p1/p6. This mechanism is interesting because it suggests that a mutation outside the cleavage site interferes with the PR–Gag interaction. Indeed, Prabu-Jeyabalan et al. (2004) documented that a Gag mutation (A431V^{Gag}) compensates for a PR mutation (V82A^{PR}), which is not in direct contact with A431V^{Gag}. Thus, our data confirm the virological and structural advantages of P453L^{Gag} in viruses possessing D30N^{PR}/N88D^{PR}. Furthermore, this association appeared to be quite common as the frequency of P453L^{Gag} is 7.3% in the Los Alamos National Laboratory HIV sequence database. Though nelfinavir is no longer recommended as a first-line antiretroviral in the guidelines of developed countries, many cases previously exposed to nelfinavir have acquired D30N/N88D mutations. Indeed, the prevalence of the D30N mutation in PI-treated persons infected with subtype B viruses ($n = 7396$) and in nelfinavir-treated persons ($n = 1128$) is 7.9% and 28.1%, respectively, in the Stanford HIV drug resistance database (http://hivdb.stanford.edu/).

Regarding the bioinformatics analysis strategy, we selected the within-host substitution model in Spidermonkey analysis to infer the co-evolving sites (Nickle et al., 2007) as the data were sequences serially collected over 5 years from a single patient under anti-HIV treatment. One disadvantage of this program is that it does not account for the number of descendant clones. Often, mutation pairs on few viral clones might be determined as co-evolving pairs. In our study, 46^{PR} and 35^{PR} mutations were determined as a co-evolving pair with high posterior probability, but only one clone with this pair was observed among 129 sequences (Table 2), suggesting this co-mutation pair could not become "fixed" in a viral population. Thus, it is important to confirm the significance of the program output.

5. Conclusions

In conclusion, we successfully determined the Gag-protease associated sites P453L^{Gag}/D30N^{PR}/N88D^{PR} by applying single-genome sequencing, suggesting the usefulness of this method. However, as SGS is a more expensive method than direct sequencing, researchers need to consider which method has the best advantage for their samples. Extrapolating from our data, the relationships between the major mutations found by SGS may not differ significantly from direct sequencing results, but we might have a greater chance of seeing a variety of minority clones with minor mutations. In addition, our observation of higher variation at later sampling points suggests that cases with longer treatment histories are good sample candidates.

We found that the viruses acquiring P453L^{Gag}/D30N^{PR}/N88D^{PR} distinctly showed biological advantages. From results obtained using both viral experiments and bioinformatics, we speculate that the P453L^{Gag} mutation does not necessarily occur in the presence of nelfinavir, but if it occurs with D30N^{PR}/N88D^{PR} mutations, viral fitness can be improved, which may eventually lead to worse clinical outcomes of anti-HIV therapy. We believe that the findings of this study provide new insight into the mechanism of within-patient HIV-1 co-evolution and into the acquisition of resistance to anti-HIV drugs.

Acknowledgements

The authors thank Dr. Akira Shirahata and Mr. Yuki Kitamura for their support. We thank patients who contributed to our study. We also thank Ms. Claire Baldwin for her help in preparing the manuscript. This study was supported by a Grant-in-Aid for AIDS research from the Ministry of Health, Labor and Welfare of Japan (H19-AIDS-007), and also by Scientific Research from the Ministry of Education, Culture, Sports, and Technology of Japan (Project number: 19510208).

References

Altschuh, D., Lesk, A.M., Bloomer, A.C., Klug, A., 1987. Correlation of co-ordinated amino acid substitutions with function in viruses related to tobacco mosaic virus. J. Mol. Biol. 193, 693–707.

Baker, D., Sali, A., 2001. Protein structure prediction and structural genomics. Science 294, 93–96.

Bally, F., Martinez, R., Peters, S., Sudre, P., Telenti, A., 2000. Polymorphism of HIV type 1 gag p7/p1 and p1/p6 cleavage sites: clinical significance and implications for resistance to protease inhibitors. AIDS Res. Hum. Retroviruses 16, 1209–1213.

Bhattacharya, T., Daniels, M., Heckerman, D., Foley, B., Frahm, N., Kadie, C., Carlson, J., Yusim, K., McMahon, B., Gaschen, B., Mallal, S., Mullins, J.I., Nickle, D.C., Herbeck, J., Rousseau, C., Learn, G.H., Miura, T., Brander, C., Walker, B., Korber, B., 2007. Founder effects in the assessment of HIV polymorphisms and HLA allele associations. Science 315, 1583–1586.

Brumme, Z.L., John, M., Carlson, J.M., Brumme, C.J., Chan, D., Brockman, M.A., Swenson, L.C., Tao, I., Szeto, S., Rosato, P., Sela, J., Kadie, C.M., Frahm, N., Brander, C., Haas, D.W., Riddler, S.A., Haubrich, R., Walker, B.D., Harrigan, P.R., Heckerman, D., Mallal, S., 2009. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. PLoS One 4, e6687.

Chiba-Mizutani, T., Miura, H., Matsuda, M., Matsuda, Z., Yokomaku, Y., Miyauchi, K., Nishizawa, M., Yamamoto, N., Sugiura, W., 2007. Use of new T-cell-based cell lines expressing two luciferase reporters for accurately evaluating susceptibility to anti-human immunodeficiency virus type 1 drugs. J. Clin. Microbiol. 45, 477–487.

Dam, E., Quercia, R., Glass, B., Descamps, D., Launay, O., Duval, X., Krausslich, H.G., Hance, A.J., Clavel, F., 2009. Gag mutations strongly contribute to HIV-1 resistance to protease inhibitors in highly drug-experienced patients besides compensating for fitness loss. PLoS Pathog. 5, e1000345.

Doyon, L., Croteau, G., Thibeault, D., Poulin, F., Pilote, L., Lamarre, D., 1996. Second locus involved in human immunodeficiency virus type 1 resistance to protease inhibitors. J. Virol. 70, 3763–3769.

Dutheil, J., Pupko, T., Jean-Marie, A., Galtier, N., 2005. A model-based approach for detecting coevolving positions in a molecule. Mol. Biol. Evol. 22, 1919–1928.

Gallego, O., de Mendoza, C., Corral, A., Soriano, V., 2003. Changes in the human immunodeficiency virus p7-p1-p6 gag gene in drug-naive and pretreated patients. J. Clin. Microbiol. 41, 1245–1247.

Gunthard, H.F., Wong, J.K., Ignacio, C.C., Havlir, D.V., Richman, D.D., 1998. Comparative performance of high-density oligonucleotide sequencing and dideoxynucleotide sequencing of HIV type 1 pol from clinical samples. AIDS Res. Hum. Retroviruses 14, 869–876.

Gutell, R.R., Power, A., Hertz, G.Z., Putz, E.J., Stormo, G.D., 1992. Identifying constraints on the higher-order structure of RNA: continued development and application of comparative sequence analysis methods. Nucleic Acids Res. 20, 5785–5795.

Hance, A.J., Lemiale, V., Izopet, J., Lecossier, D., Joly, V., Massip, P., Mammano, F., Descamps, D., Brun-Vézinet, F., Clavel, F., 2001. Changes in human immunodeficiency virus type 1 populations after treatment interruption in patients failing antiretroviral therapy. J. Virol. 75, 6410–6417.

Ho, S.K., Coman, R.M., Bunger, J.C., Rose, S.L., O'Brien, P., Munoz, I., Dunn, B.M., Sleasman, J.W., Goodenow, M.M., 2008. Drug-associated changes in amino acid residues in Gag p2, p7(NC), and p6(Gag)/p6(Pol) in human immunodeficiency virus type 1 (HIV-1) display a dominant effect on replicative fitness and drug response. Virology.

Johnson, V.A., Brun-Vézinet, F., Clotet, B., Gunthard, H.F., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Richman, D.D., 2008. Update of the drug resistance mutations in HIV-1: Spring 2008. Top. HIV Med. 16, 62–68.
Johnson, V.A., Brun-Vézinet, F., Clotet, B., Gunthard, H.F., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Richman, D.D., 2009. Update of the drug resistance mutations in HIV-1: December 2009. Top. HIV Med. 17, 138–145.
Koch, N., Yahi, N., Fantini, J., Tamalet, C., 2001. Mutations in HIV-1 gag cleavage sites and their association with protease mutations. AIDS 15, 526–528.
Kolli, M., Lastere, S., Schiffer, C.A., 2006. Co-evolution of nelfinavir-resistant HIV-1 protease and the p1-p6 substrate. Virology 347, 405–409.
Kolli, M., Stawiski, E., Chappey, C., Schiffer, C.A., 2009. Human immunodeficiency virus type 1 protease-correlated cleavage site mutations enhance inhibitor resistance. J. Virol. 83, 11027–11042.
Labute, P., 2008. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. J. Comput. Chem. 29, 1693–1698.
Maguire, M.F., Guinea, R., Griffin, P., Macmanus, S., Elston, R.C., Wolfram, J., Richards, N., Hanlon, M.H., Porter, D.J., Wrin, T., Parkin, N., Tisdale, M., Furfine, E., Petropoulos, C., Snowden, B.W., Kleim, J.P., 2002. Changes in human immunodeficiency virus type 1 Gag at positions L449 and P453 are linked to I50V protease mutants in vivo and cause reduction of sensitivity to amprenavir and improved viral fitness in vitro. J. Virol. 76, 7398–7406.
Mahalingam, B., Louis, J.M., Reed, C.C., Adomat, J.M., Krouse, J., Wang, Y.F., Harrison, R.W., Weber, I.T., 1999. Structural and kinetic analysis of drug resistant mutants of HIV-1 protease. Eur. J. Biochem. 263, 238–245.
Malet, I., Roquebert, B., Dalban, C., Wirden, M., Amellal, B., Agher, R., Simon, A., Katlama, C., Costagliola, D., Calvez, V., Marcelin, A.G., 2007. Association of Gag cleavage sites to protease mutations and to virological response in HIV-1 treated patients. J. Infect. 54, 367–374.
Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F., Sali, A., 2000. Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. 29, 291–325.
Martin, L.C., Gloor, G.B., Dunn, S.D., Wahl, L.M., 2005. Using information theory to search for co-evolving residues in proteins. Bioinformatics 21, 4116–4124.
Martinez-Picado, J., Savara, A.V., Sutton, L., D'Aquila, R.T., 1999. Replicative fitness of protease inhibitor-resistant mutants of human immunodeficiency virus type 1. J. Virol. 73, 3744–3752.
Matsuoka-Aizawa, S., Sato, H., Hachiya, A., Tsuchiya, K., Takebe, Y., Gatanaga, H., Kimura, S., Oka, S., 2003. Isolation and molecular characterization of a nelfinavir (NFV)-resistant human immunodeficiency virus type 1 that exhibits NFV-dependent enhancement of replication. J. Virol. 77, 318–327.
Meyerhans, A., Vartanian, J.P., Wain-Hobson, S., 1990. DNA recombination during PCR. Nucleic Acids Res. 18, 1687–1691.
Mitsuya, Y., Winters, M.A., Fessel, W.J., Rhee, S.Y., Hurley, L., Horberg, M., Schiffer, C.A., Zolopa, A.R., Shafer, R.W., 2006. N88D facilitates the co-occurrence of D30N and L90M and the development of multidrug resistance in HIV type 1 protease following nelfinavir treatment failure. AIDS Res. Hum. Retroviruses 22, 1300–1305.
Myint, L., Matsuda, M., Matsuda, Z., Yokomaku, Y., Chiba, T., Okano, A., Yamada, K., Sugiura, W., 2004. Gag non-cleavage site mutations contribute to full recovery of viral fitness in protease inhibitor-resistant human immunodeficiency virus type 1. Antimicrob. Agents Chemother. 48, 444–452.
Neher, E., 1994. How frequent are correlated changes in families of protein sequences? Proc. Natl. Acad. Sci. U. S. A. 91, 98–102.
Nickle, D.C., Heath, L., Jensen, M.A., Gilbert, P.B., Mullins, J.I., Kosakovsky Pond, S.L., 2007. HIV-specific probabilistic models of protein evolution. PLoS One 2, e503.
Nijhuis, M., Schuurman, R., de Jong, D., Erickson, J., Gustchina, E., Albert, J., Schipper, P., Gulnik, S., Boucher, C.A., 1999. Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. AIDS 13, 2349–2359.
Palmer, S., Kearney, M., Maldarelli, F., Halvas, E.K., Bixby, C.J., Bazmi, H., Rock, D., Falloon, J., Davey Jr., R.T., Dewar, R.L., Metcalf, J.A., Hammer, S., Mellors, J.W., Coffin, J.M., 2005. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. J. Clin. Microbiol. 43, 406–413.
Pollock, D.D., Taylor, W.R., 1997. Effectiveness of correlation analysis in identifying protein residues undergoing correlated evolution. Protein Eng. 10, 647–657.
Poon, A.F., Kosakovsky Pond, S.L., Richman, D.D., Frost, S.D., 2007a. Mapping protease inhibitor resistance to human immunodeficiency virus type 1 sequence polymorphisms within patients. J. Virol. 81, 13598–13607.
Poon, A.F., Lewis, F.I., Frost, S.D., Pond, S.L., 2008. Spidermonkey: rapid detection of co-evolving sites using Bayesian graphical models. Bioinformatics.
Poon, A.F., Lewis, F.I., Pond, S.L., Frost, S.D., 2007b. An evolutionary-network model reveals stratified interactions in the V3 loop of the HIV-1 envelope. PLoS Comput. Biol. 3, e231.
Prabu-Jeyabalan, M., Nalivaika, E., Schiffer, C.A., 2002. Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. Structure 10, 369–381.
Prabu-Jeyabalan, M., Nalivaika, E.A., King, N.M., Schiffer, C.A., 2004. Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral protease. J Virol 78, 12446–12454.
Rhee, S.Y., Liu, T.F., Holmes, S.P., Shafer, R.W., 2007. HIV-1 subtype B protease and reverse transcriptase amino acid covariation. PLoS Comput. Biol. 3, e87.
Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
Sayer, J.M., Liu, F., Ishima, R., Weber, I.T., Louis, J.M., 2008. Effect of the active site D25N mutation on the structure, stability, and ligand binding of the mature HIV-1 protease. J. Biol. Chem. 283, 13459–13470.
Shirakawa, K., Takaori-Kondo, A., Yokoyama, M., Izumi, T., Matsui, M., Io, K., Sato, T., Sato, H., Uchiyama, T., 2008. Phosphorylation of APOBEC3G by protein kinase A regulates its interaction with HIV-1 Vif. Nat. Struct. Mol. Biol. 15, 1184–1191.
Sugiura, W., Matsuda, Z., Yokomaku, Y., Hertogs, K., Larder, B., Oishi, T., Okano, A., Shiino, T., Tatsumi, M., Matsuda, M., Abumi, H., Takata, N., Shirahata, S., Yamada, K., Yoshikura, H., Nagai, Y., 2002. Interference between D30N and L90M in selection and development of protease inhibitor-resistant human immunodeficiency virus type 1. Antimicrob. Agents Chemother. 46, 708–715.
Tillier, E.R., Lui, T.W., 2003. Using multiple interdependency to separate functional from phylogenetic correlations in protein alignments. Bioinformatics 19, 750–755.
Tuff, P., Darlu, P., 2000. Exploring a phylogenetic approach for the detection of correlated substitutions in proteins. Mol. Biol. Evol. 17, 1753–1759.
Verheyen, J., Litau, E., Sing, T., Daumer, M., Balduin, M., Oette, M., Fatkenheuer, G., Rockstroh, J.K., Schuldenzucker, U., Hoffmann, D., Pfister, H., Kaiser, R., 2006. Compensatory mutations at the HIV cleavage sites p7/p1 and p1/p6-gag in therapy-naive and therapy-experienced patients. Antivir. Ther. 11, 879–887.
von Schwedler, U.K., Stuchell, M., Muller, B., Ward, D.M., Chung, H.Y., Morita, E., Wang, H.E., Davis, T., He, G.P., Cimbora, D.M., Scott, A., Krausslich, H.G., Kaplan, J., Morham, S.G., Sundquist, W.I., 2003. The protein network of HIV budding. Cell 114, 701–713.
Wang, J., Cieplak, P., Kollman, P.A., 2000. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? Journal of Computational Chemistry 21, 1049–1074.
Willey, R.L., Smith, D.H., Lasky, L.A., Theodore, T.S., Earl, P.L., Moss, B., Capon, D.J., Martin, M.A., 1988. In vitro mutagenesis identifies a region within the envelope gene of the human immunodeficiency virus that is critical for infectivity. J. Virol. 62, 139–147.
Wollenberg, K.R., Atchley, W.R., 2000. Separation of phylogenetic and functional associations in biological sequences by using the parametric bootstrap. Proc. Natl. Acad. Sci. U. S. A. 97, 3288–3291.
Wu, T.D., Schiffer, C.A., Gonzales, M.J., Taylor, J., Kantor, R., Chou, S., Israelski, D., Zolopa, A.R., Fessel, W.J., Shafer, R.W., 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. J. Virol. 77, 4836–4847.
Yeang, C.H., Haussler, D., 2007. Detecting coevolution in and among protein domains. PLoS Comput. Biol. 3, e211.
Zhang, Y.M., Imamichi, H., Imamichi, T., Lane, H.C., Falloon, J., Vasudevachari, M.B., Salzman, N.P., 1997. Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites. J. Virol. 71, 6662–6670.

Bioorganic & Medicinal Chemistry 20 (2012) 1468–1474

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Conjugation of cell-penetrating peptides leads to identification of anti-HIV peptides from matrix proteins

Tetsuo Narumi [a], Mao Komoriya [a], Chie Hashimoto [a], Honggui Wu [b,c], Wataru Nomura [a], Shintaro Suzuki [a], Tomohiro Tanaka [a], Joe Chiba [c], Naoki Yamamoto [d], Tsutomu Murakami [b,*], Hirokazu Tamamura [a,*]

[a] Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, Chiyoda-ku, Tokyo 101-0062, Japan
[b] AIDS Research Center, National Institute of Infectious Diseases, Shinjuku-ku, Tokyo 162-8640, Japan
[c] Department of Biological Science Technology, Tokyo University of Science, Noda, Chiba 278-8510, Japan
[d] Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore

ARTICLE INFO

Article history:
Received 6 December 2011
Revised 24 December 2011
Accepted 24 December 2011
Available online 2 January 2012

Keywords:
Matrix protein
Octa-arginyl group
Overlapping peptide
Anti-HIV

ABSTRACT

Compounds which inhibit the HIV-1 replication cycle have been found amongst fragment peptides derived from an HIV-1 matrix (MA) protein. Overlapping peptide libraries covering the whole sequence of MA were designed and constructed with the addition of an octa-arginyl group to increase their cell membrane permeability. Imaging experiments with fluorescent-labeled peptides demonstrated these peptides with an octa-arginyl group can penetrate cell membranes. The fusion of an octa-arginyl group was proven to be an efficient way to find active peptides in cells such as HIV-inhibitory peptides.

1. Introduction

Several anti-retroviral drugs beyond reverse transcriptase inhibitors, including effective protease inhibitors[1] and integrase inhibitors[2,3] are currently available to treat human immunodeficiency virus type 1 (HIV-1) infected individuals. We have also developed several anti-HIV agents such as coreceptor CXCR4 antagonists,[4–7] CD4 mimics,[8–10] fusion inhibitors[11] and integrase inhibitors.[12,13] However, the emergence of viral strains with multi-drug resistance (MDR), which accompanies the development of any antiviral drug, has encouraged a search for new types of anti-HIV-1 drugs with different inhibitory mechanisms.

Matrix (MA) proteins are essential for assembly of the virion shell. MA is a component of the Gag precursor protein, Pr55Gag, and is located within the viral membrane.[14,15] It has been reported that MA-derived peptides such as MA(47–59) inhibit infection by HIV,[16] and that MA-derived peptides such as MA(31–45) and MA(41–55) show anti-HIV activity.[17] In addition, Morikawa et al. report that MA(61–75) and MA(71–85) inhibit MA dimerization, a necessary step in the formation of the virion shell.[18] However, the question of whether the above MA peptides can penetrate cell membranes was not addressed in these reports. We speculate that to achieve antiviral activity it is essential that the MA-derived peptides penetrate the cell membrane and function intracellularly. In this paper, we report our design and construction of an overlapping library of fragment peptides derived from the MA protein with a cell membrane permeable signal. Our aim is the discovery of potent lead compounds, which demonstrate HIV inhibitory activity inside the host cells.

2. Materials and methods

2.1. Peptide synthesis

MA-derived fragments and an octa-arginyl (R_8) peptide were synthesized by stepwise elongation techniques of Fmoc-protected amino acids on a Rink amide resin. Coupling reactions were performed using 5.0 equiv of Fmoc-protected amino acid, 5.0 equiv of diisopropylcarbodiimide and 5.0 equiv of 1-hydroxybenzotriazole monohydrate. Ac_2O–pyridine (1/1, v/v) for 20 min was used to acetylate the N-terminus of MA-derived fragments, with the exception of fragment 1. Chloroactylation of the N-terminus of the R_8 peptide, was achieved with 40 equiv of chloroacetic acid, 40 equiv of diisopropylcarbodiimide and 40 equiv of 1-hydroxybenzotriazole monohydrate, treated for 1 h. Cleavage of peptides from resin and side chain deprotection were carried out by stirring for 1.5 h with a mixture of TFA, thioanisole, ethanedithiol, *m*-cresol

* Corresponding authors. Tel.: +81 3 5285 1111; fax: +81 3 5285 5037 (T.M.); tel.: +81 3 5280 8036; fax: +81 3 5280 8039 (H.M.).
E-mail addresses: tmura@nih.go.jp (T. Murakami), tamamura.mr@tmd.ac.jp (H. Tamamura).

0968-0896/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2011.12.055

and triisopropylsilane (8.15/0.75/0.75/0.25/0.25/0.1, v/v). After removal of the resins by filtration, the filtrate was concentrated under reduced pressure, and crude peptides were precipitated in cooled diethyl ether. All crude peptides were purified by RP-HPLC and identified by ESI-TOFMS. In the conjugation of the R_8 peptide (or iodoacetamide), the peptide (or iodoacetamide) solution in 0.1 M phosphate buffer, pH 7.8 was added to MA fragments which were synthesized as described above. The reaction mixture was stirred at room temperature under nitrogen. After 24 h (or 1 h for the conjugation of iodoacetamide), purification was performed by RP-HPLC. The purified peptides were identified by ESI-TOF MS and lyophilized. Purities of all final compounds were confirmed to be >95% by analytical HPLC. Detailed data are provided in Supplementary data.

2.2. Anti-HIV-1 assay

Anti-HIV-1 (NL4-3 or NL(AD8)) activity was determined by measurement of the protection against HIV-1-induced cytopathogenicity in MT-4 cells or PM1/CCR5 cells. Various concentrations of test peptide solutions were added to HIV-1 infected MT-4 or PM1/CCR5 cells at multiplicity of infection (MOI) of 0.001 and placed in wells of a 96-well microplate. After 5 day incubation at 37 °C in a CO_2 incubator, the number of viable cells was determined using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) method. The anti-HIV-1 (JR-CSF) activity was also determined by measuring capsid p24 antigen concentrations of the culture supernatant in the infected cultures by a commercially available ELISA assay (ZeptoMetrix Corp., Buffalo, NY).

2.3. CD spectroscopy

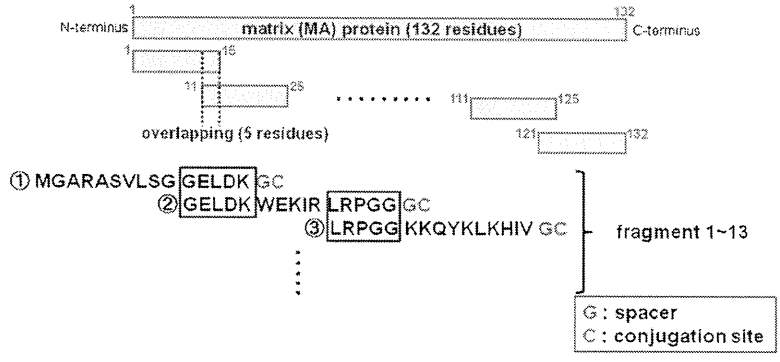
CD spectra were recorded on a JASCO J-720 spectropolarimeter at 25 °C. The measurements were performed using a 0.1 cm path length cuvette at a 0.1 nm spectral resolution. Each spectrum represents the average of 10 scans, and the scan rate was 50 nm/min. The concentrations of samples 8L and 9L were 28.2 and 64.7 μM, respectively, in PBS buffer (pH 7.4).

2.4. Fluorescent imaging of cell-penetrating MA peptides

Cells were seeded on 35 mm glass-bottom dish (2×10^5 cells/dish for HeLa and A549, 1×10^5 cells/dish for CHO-K1) one day before the experiments. The cells were cultured in DMEM/10% FBS/Penicillin–Streptomycin for HeLa and A549, or Ham's F12/10% FBS/Penicillin–Streptomycin for CHO-K1 at 37 °C/5% CO_2. Before the addition of MA peptides, cells were washed with Hanks' balanced salt solutions (HBSS) once. Peptides were added at 5 μM and further cultured for 30 min at 37 °C/5% CO_2. After incubation, cells were washed three times with HBSS and observed under a confocal laser-scanning microscopy (Zeiss LSM510).

3. Results and discussion

An overlapping peptide library spanning the whole sequence of the MA domain, p17, of NL4-3, the Gag precursor Pr55 of HIV-1 was designed. The full sequence of MA consists of 132 amino acid residues. In the peptide library, the MA sequence was divided from the N-terminus in 15-residue segments with an overlap of 5



fragment number	sequence
1	H-MGARASVLSGGELDKGC-NH_2
2	CH_3CO-GELDKWEKIRLRPGGGC-NH_2
3	CH_3CO-LRPGGKKQYKLKHIVGC-NH_2
4	CH_3CO-LKHIVWASRELERFAGC-NH_2
5	CH_3CO-LERFAVNPGLLETSEGC-NH_2
6	CH_3CO-LETSEGSRQILGQLQGC-NH_2
7	CH_3CO-LGQLQPSLQTGSEELGC-NH_2
8	CH_3CO-GSEELRSLYNTIAVLGC-NH_2
9	CH_3CO-TIAVLYSVHQRIDVKGC-NH_2
10	CH_3CO-RIDVKDTKEALDKIEGC-NH_2
11	CH_3CO-LDKIEEEQNKSKKKAGC-NH_2
12	CH_3CO-SKKKAQQAAADTGNNGC-NH_2
13	CH_3CO-DTGNNSQVSQNYGC-NH_2

Figure 1. The construction of MA-based overlapping peptide library.

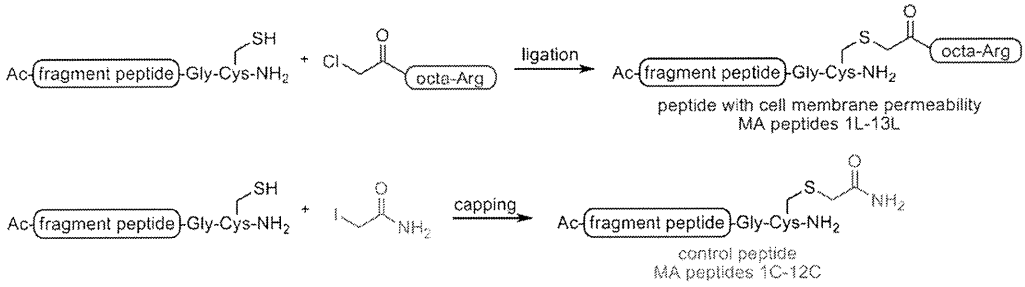


Figure 2. The design of MA peptides with cell membrane permeability (upper) and their control peptides (lower).

residues to preserve secondary structures (Fig. 1). Cys residues of the original MA sequence were changed into Ser residues because of the facility of peptide synthesis. Thirteen MA fragment peptides (1–13) were designed with the addition of Gly as a spacer and Cys as a conjugation site at the C-terminus. To impart cell membrane permeability to these peptides, the N-terminal chloroacetyl group of an octa-arginyl (R_8) peptide[19] was conjugated to the side-chain thiol group of the Cys residue of the above peptides. This resulted in the MA peptides 1L–13L (Fig. 2). R_8 is a cell membrane permeable motif and its fusion with parent peptides is known to produce bioactive peptides with no significant adverse properties.[12,13,20–24] In addition, the R_8-fusion can increase the solubility of MA

Table 1
Anti-HIV activity and cytotoxicity of control MA peptides

MA peptide	MT-4 cell	PM1/CCR5 cell		MT-4 cell
	NL4-3 (MTT assay) EC_{50}[a] (μM)	NL(AD8) (MTT assay) EC_{50}[a] (μM)	JR-CSF (p24 ELISA) EC_{50}[a] (μM)	(MTT assay) CC_{50}[b] (μM)
1C	>50	ND	ND	>50
2C	17 ± 1.4	1.0	ND	>50
3C	>50	ND	ND	>50
4C	No inhibition at 12.5 μM	ND	ND	14
5C	>50	ND	ND	>50
6C	37 ± 12	24% inhibition at 6.25 μM	25% inhibition at 50 μM	>50
7C	>50	ND	ND	>50
8C	>50	ND	ND	>50
9C	29 ± 1.4	13	8.1	>50
10C	No inhibition at 12.5 μM	ND	ND	17
11C	>50	ND	ND	>50
12C	>50	ND	ND	>50
14C	>50	ND	ND	>50
AZT	0.020	0.459	0.17	>100
SCH-D	ND	0.026	0.0014	ND

X4-HIV-1 (NL4-3 strain)-induced cytopathogenicity in MT-4 cells and R5-HIV-1 (NL(AD8) strain)-induced cytopathogenicity in PM1/CCR5 cells evaluated by the MTT assay, and inhibitory activity against R5-HIV-1 (JR-CSF strain)-induced cytopathogenicity in PM1/CCR5 cells evaluated by the p24 ELISA assay.

[a] EC_{50} values are the concentrations for 50% protection from HIV-1-induced cytopathogenicity in MT-4 cells.

[b] CC_{50} values are the concentrations for 50% reduction of the viability of MT-4 cells. All data are the mean values from at least three independent experiments. ND: not determined.

Table 2
Anti-HIV activity and cytotoxicity of MA peptides with cell membrane permeability

MA peptide	MT-4 cell	PM1/CCR5 cell		MT-4 cell
	NL4-3(MTT assay) EC_{50} (μM)	NL(AD8)(MTT assay) EC_{50} (μM)	JR-CSF(p24 ELISA) EC_{50} (μM)	(MTT assay) CC_{50} (μM)
1L	30	30	40	>50
2L	21 ± 4.2	>31	ND	32 ± 4.2
3L	no inhibition at 25 μM	ND	ND	36
4L	no inhibition at 3.13 μM	ND	ND	3.7
5L	40	42% inhibition at 50 μM	42	>50
6L	40 ± 8.9	49% inhibition at 50 μM	31	>50
7L	35 ± 1.5	37% inhibition at 50 μM	35% inhibition at 50 μM	>50
8L	2.3 ± 0.3	5.8	7.8	9.0 ± 2.4
9L	2.1 ± 0.5	0.43	0.58	5.7 ± 2.1
10L	43 ± 8.5	42% inhibitionat 50 μM	27	>50
11L	18 ± 3.0	17% inhibition at 25 μM	23	>50
12L	41 ± 5.5	30% inhibition at 25 μM	27	>50
13L	20 ± 2.1	0.43	11	>50
14L	no inhibition at 25 μM	ND	ND	36
AZT	0.020	0.459	0.17	>100
SCH-D	ND	0.026	0.0014	ND

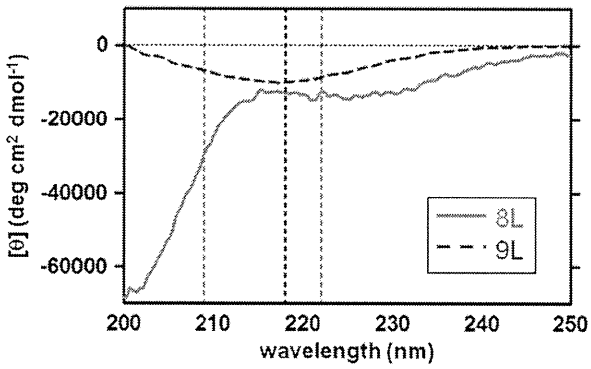


Figure 3. CD spectra of MA peptides 8L (28 μM) and 9L (65 μM) in PBS buffer, pH 7.4 at 25 °C.

peptides whose hydrophobicity is relatively limited. On the other hand, to develop control peptides lacking cell membrane permeability, iodoacetamide was conjugated to the thiol group of the Cys residue to prepare MA peptides 1C–12C (Fig. 2). MA peptide 13C was not synthesized because MA fragment 13 is insoluble in PBS buffer.

The anti-HIV activity of MA peptides 1L–13L and MA peptides 1C–12C, was evaluated. Inhibitory activity against T-cell line-tropic (X4-) HIV-1 (NL4-3 strain)-induced cytopathogenicity in MT-4 cells and against macrophage-tropic (R5-) HIV-1 (NL(AD8) strain)-induced cytopathogenicity in PM1/CCR5 cells was assessed by the 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide (MTT) assay, and inhibitory activity against R5-HIV-1 (JR-CSF strain) replication in PM1/CCR5 cells was determined by the p24 ELISA assay. The results are shown in Tables 1 and 2. The control MA peptides 6C and 9C showed slight anti-HIV activity against NL4-3, NL(AD8) and JR-CSF strains, and 2C showed high anti-HIV activity against NL4-3 and NL(AD8) strains, but the other control MA peptides showed no significant anti-HIV activity. 2C showed significant anti-HIV activity against both X4-HIV-1 and R5-HIV-1 strains, suggesting that this region of the MA domain is relevant with Gag localization to the plasma membrane (PM)[25] and that 2C might inhibit competitively the interaction between MA and PM. On the other hand, the MA peptides with the exception of 3L and 4L, showed moderate to potent anti-HIV activity against all three strains. These peptides expressed almost the same level of anti-HIV activity against both X4-HIV-1 and R5-HIV-1 strains. The MA peptides 8L and 9L in particular, showed significant anti-HIV activity. These results suggest that MA peptides achieve entry into target cells as a result of the addition of R_8, and inhibit viral replication within the cells. The adjacent peptides 8L and 9L possess an overlapping sequence TIAVL. Such peptides exhibited relatively high cytotoxicity and the MA peptide 4L showed the highest cytotoxicity although it did not show any significant anti-HIV activity. The control MA peptides 1C–12C were relatively weakly cytotoxic. The MA peptides 8C and 9C exhibited no significant cytotoxicity, although the addition of R_8, giving 8L and 9L, caused a remarkable increase in cytotoxicity. This suggests that the octa-arginyl (R_8) sequence is correlated with the

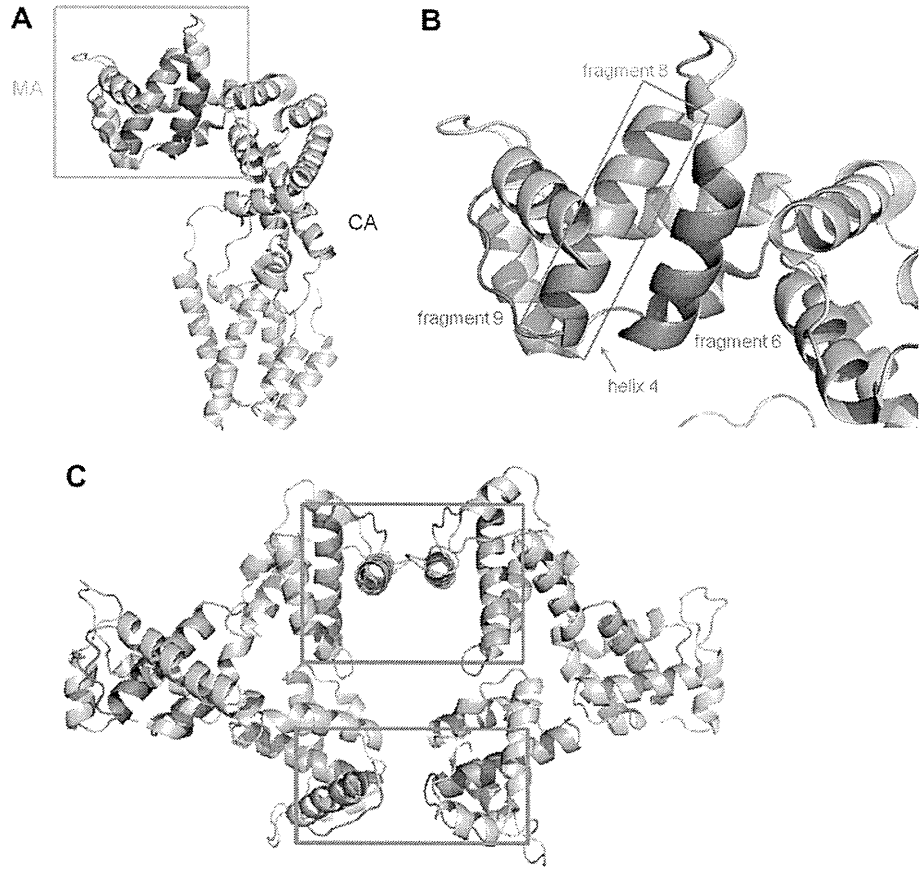


Figure 4. (A) The complete structure of MA and CA proteins (PDB ID: 2gol). (B) The enlarged structure of the highlighted region of (A). (C) The structure of an MA hexamer. Red-colored squares show interfaces between two MA trimers (PDB ID: 1hiw). Orange- and pink-colored helical ribbons represent fragments 8 and 9, respectively.

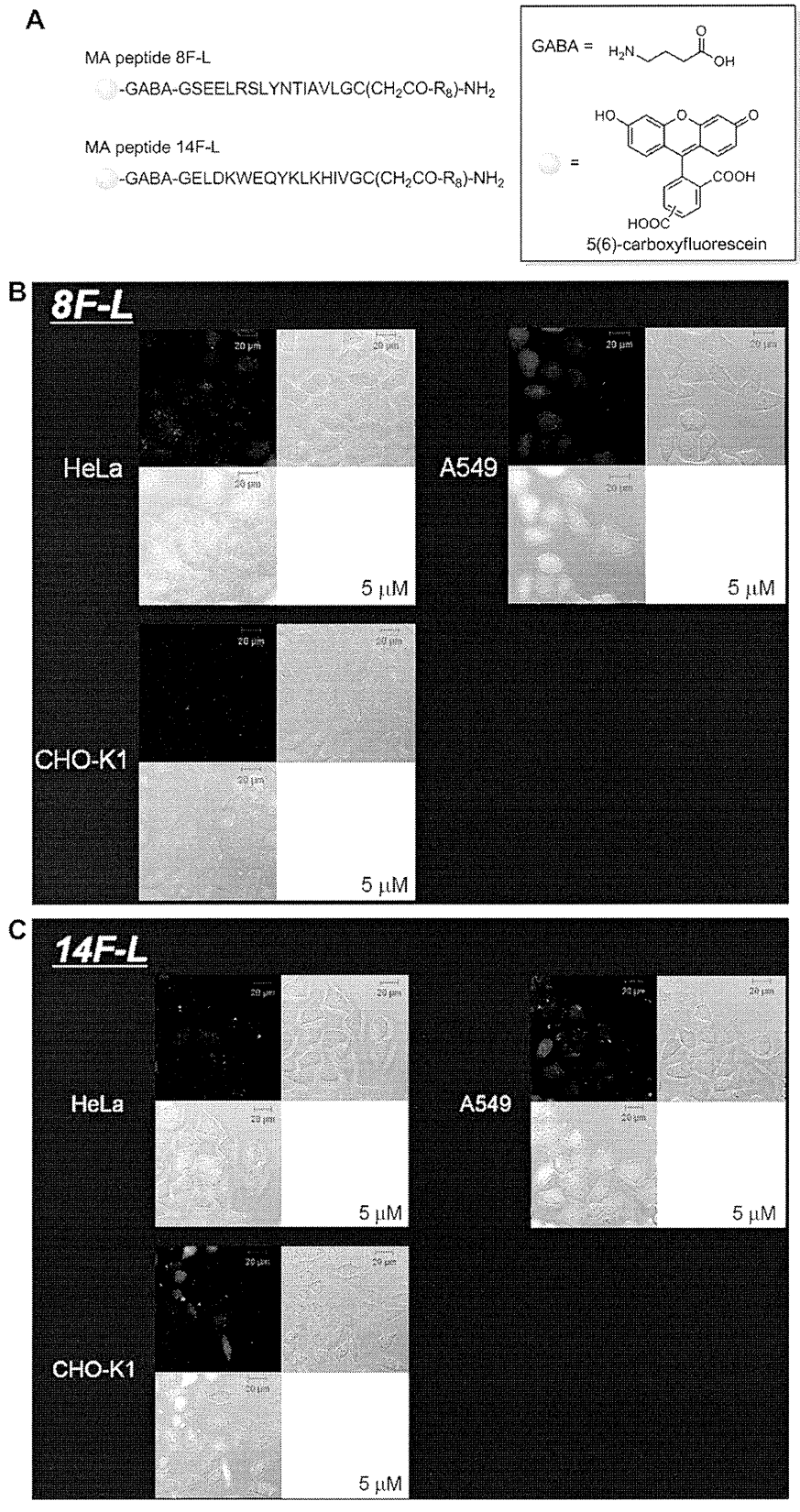


Figure 5. (A) The structures of fluorophore-labeled MA peptides 8F–L and 14F–L. (B) The fluorescent imaging of live cells HeLa, A549 and CHO-K1 by 8F–L. (C) The fluorescent imaging of live cells HeLa, A549 and CHO-K1 by 14F–L.

expression of cytotoxicity and in future, a different effective strategy for cell penetration may be advisable.

In the present assay, the control MA peptides 6C and 9C, which cover MA(51–65) and MA(81–95), respectively, showed significant anti-HIV activity. This is consistent with the previous studies, in which MA(41–55), MA(47–59) and MA(71–85) showed anti-HIV or dimerization inhibitory activity as discussed above.[16–18] These peptides have no R_8 sequence and thus cannot penetrate cell membranes. They exhibit inhibitory activity on the surface of cells, not intracellularly.

The structures of MA peptides 8L and 9L, dissolved in PBS buffer (2.7 mM KCl, 137 mM NaCl, 1.47 mM KH_2PO_4, 9.59 mM Na_2HPO_4) at pH 7.4, were determined by CD spectroscopy (Fig. 3). When peptides form α-helical structures, minima can be observed at approximately 207 and 222 nm in their CD spectra. The amino acid residues covering fragments 8 and 9 corresponding to 8L and 9L are located in an α-helical region (helix 4) of the parent MA protein (Fig. 4), and peptides 8L and 9L were presumed to have an α-helical conformation.[26–28] However, the CD spectra shown in Figure 3, suggest that these peptides lack any characteristic secondary structure. This is because the 15-mer peptide derived from MA is not sufficiently long to form a secondary structure even though Gly, Cys and octa-Arg are attached to their C-terminus. Analysis of the CD spectra suggests MA fragment peptides need a longer sequence in order to form a secondary structure. The CD spectra of the control MA peptides 8C and 9C were not determined because the aqueous solubility of these peptides is inadequate.

Fluorescent imaging of live cells was used to evaluate the cell membrane permeability of the MA peptides 8L and 14L, which showed high and zero significant anti-HIV activity, respectively. The MA fragment 14 is a hybrid of the fragments 2 and 3, and the MA peptides 14L and 14C, which are based on the conjugation of the N-terminal chloroacetyl group of an R_8 peptide and iodoacetamide to the thiol group of the Cys residue, respectively (Supplementary data), are control peptides lacking significant anti-HIV activity (Tables 1 and 2). These peptides were labeled with 5(6)-carboxyfluorescein via a GABA linker at the N-terminus to produce 8F–L and 14F–L (Fig. 5A). The fluorophore-labeled peptides 8F–L and 14F–L were incubated with live cells of HeLa, A549 and CHO-K1, and the imaging was analyzed by a fluorescence microscope (Fig. 5B and C). A549 cells are human lung adenocarcinomic human alveolar basal epithelial cells.[29] Similar penetration of both peptides 8F–L and 14F–L into these cells was observed. Even peptides without significant anti-HIV activity can penetrate cell membranes. The penetration efficiency of both peptides into A549 was relatively high and into HeLa was low. In CHO-K1 the penetration efficiency of 8F–L is relatively low, but that of 14F–L is high. These imaging data confirm that the MA peptides with the R_8 sequence can penetrate cell membranes and suggest that MA peptides such as 8L and 9L should be able to inhibit HIV replication inside cells.

4. Conclusions

Several HIV-1 inhibitory fragment peptides were identified through the screening of an overlapping peptide library derived from the MA protein. Judging by the imaging experiments, peptides possessing the R_8 group can penetrate cell membranes and might exhibit their function intracellularly thus inhibiting HIV replication.

Two possible explanations for the inhibitory activity of these MA fragment peptides can be envisaged: (1) The fragment peptides might attack an MA protein and inhibit the assembly of MA proteins. (2) These peptides might attack a cellular protein and inhibit its interaction with MA. Further studies to elucidate detailed action mechanisms and identify the targets of these peptides will be performed in future. The technique of addition of the R_8 group to peptides enabled us to screen library peptides that function within cells. Thus, the design of an overlapping peptide library of fragment peptides derived from a parent protein with a cell membrane permeable signal is a useful and efficient strategy for finding potent cell-penetrating lead compounds.

In the present study, the MA peptides 8L and 9L were shown to inhibit HIV-1 replication with submicromolar to micromolar EC_{50} values in cells using the MT-4 assay (NL4-3 and NL(AD8) strains) and the p24 ELISA assay (JR-CSF strain). Our findings suggest that these peptides could serve as lead compounds for the discovery of novel anti-HIV agents. Amino acid residues covering fragments 8 and 9 corresponding to 8L and 9L are located in the exterior surface of MA, and in particular in the interface between two MA trimers (Fig. 4C).[26–28] The interaction of two MA trimers leads to the formation of an MA hexamer, which is the MA assembly with physiological significance. Thus, the region covering fragments 8 and 9 is critical to oligomerization of MA proteins. This suggests that MA peptides 8L and 9L might inhibit the MA oligomerization through competitive binding to the parent MA, and that more potent peptides or peptidomimetic HIV inhibitors could result from studies on the mechanism of action of these MA peptides and identification of the interaction sites. Taken together, some seeds for anti-HIV agents are inherent in MA proteins, including inhibitors of the interaction with PM such as the MA peptide 2C.

Acknowledgements

This work was supported in part by Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and Health and Labour Sciences Research Grants from Japanese Ministry of Health, Labor, and Welfare. C.H. and T.T. were supported by JSPS Research Fellowships for Young Scientists. The authors thank Ms. M. Kawamata, National Institute of Infectious Diseases, for her assistance in the anti-HIV assay. We also thank Dr. Y. Maeda, Kumamoto University, for providing PM1/CCR5 cells, and Mr. S. Kumakura, Kureha Corporation, for providing SCH-D, respectively.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2011.12.055.

References and notes

1. Ghosh, A. K.; Dawson, Z. L.; Mitsuya, H. *Bioorg. Med. Chem.* **2007**, *15*, 7576.
2. Cahn, P.; Sued, O. *Lancet* **2007**, *369*, 1235.
3. Grinsztejn, B.; Nguyen, B.-Y.; Katlama, C.; Gatell, J. M.; Lazzarin, A.; Vittecoq, D.; Gonzalez, C. J.; Chen, J.; Harvey, C. M.; Isaacs, R. D. *Lancet* **2007**, *369*, 1261.
4. Tamamura, H.; Xu, Y.; Hattori, T.; Zhang, X.; Arakaki, R.; Kanbara, K.; Omagari, A.; Otaka, A.; Ibuka, T.; Yamamoto, N.; Nakashima, H.; Fujii, N. *Biochem. Biophys. Res. Commun.* **1998**, *253*, 877.
5. Fujii, N.; Oishi, S.; Hiramatsu, K.; Araki, T.; Ueda, S.; Tamamura, H.; Otaka, A.; Kusano, S.; Terakubo, S.; Nakashima, H.; Broach, J. A.; Trent, J. O.; Wang, Z.; Peiper, S. C. *Angew. Chem., Int. Ed.* **2003**, *42*, 3251.
6. Tamamura, H.; Hiramatsu, K.; Mizumoto, M.; Ueda, S.; Kusano, S.; Terakubo, S.; Akamatsu, M.; Yamamoto, N.; Trent, J. O.; Wang, Z.; Peiper, S. C.; Nakashima, H.; Otaka, A.; Fujii, N. *Org. Biomol. Chem.* **2003**, *1*, 3663.
7. Tanaka, T.; Nomura, W.; Narumi, T.; Masuda, A.; Tamamura, H. *J. Am. Chem. Soc.* **2010**, *132*, 15899.
8. Yamada, Y.; Ochiai, C.; Yoshimura, K.; Tanaka, T.; Ohashi, N.; Narumi, T.; Nomura, W.; Harada, S.; Matsushita, S.; Tamamura, H. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 354.
9. Narumi, T.; Ochiai, C.; Yoshimura, K.; Harada, S.; Tanaka, T.; Nomura, W.; Arai, H.; Ozaki, T.; Ohashi, N.; Matsushita, S.; Tamamura, H. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 5853.
10. Yoshimura, K.; Harada, S.; Shibata, J.; Hatada, M.; Yamada, Y.; Ochiai, C.; Tamamura, H.; Matsushita, S. *J. Virol.* **2010**, *84*, 7558.

11. Otaka, A.; Nakamura, M.; Nameki, D.; Kodama, E.; Uchiyama, S.; Nakamura, S.; Nakano, H.; Tamamura, H.; Kobayashi, Y.; Matsuoka, M.; Fujii, N. *Angew. Chem., Int. Ed.* **2002**, *41*, 2937.
12. Suzuki, S.; Urano, E.; Hashimoto, C.; Tsutsumi, H.; Nakahara, T.; Tanaka, T.; Nakanishi, Y.; Maddali, K.; Han, Y.; Hamatake, M.; Miyauchi, K.; Pommier, Y.; Beutler, J. A.; Sugiura, W.; Fuji, H.; Hoshino, T.; Itotani, K.; Nomura, W.; Narumi, T.; Yamamoto, N.; Komano, J. A. *J. Med. Chem.* **2010**, *53*, 5356.
13. Suzuki, S.; Maddali, K.; Hashimoto, C.; Urano, E.; Ohashi, N.; Tanaka, T.; Ozaki, T.; Arai, H.; Tsutsumi, H.; Narumi, T.; Nomura, W.; Yamamoto, N.; Pommier, Y.; Komano, J. A.; Tamamura, H. *Bioorg. Med. Chem.* **2010**, *18*, 6771.
14. Freed, E. O. *Virology* **1998**, *251*, 1.
15. Bukrinskaya, A. *Virus Res.* **2007**, *124*, 1.
16. Niedrig, M.; Gelderblom, H. R.; Pauli, G.; März, J.; Bickhard, H.; Wolf, H.; Modrow, S. *J. Gen. Virol.* **1994**, *75*, 1469.
17. Cannon, P. M.; Matthews, S.; Clark, N.; Byles, E. D.; Iourin, O.; Hockley, D. J.; Kingsman, S. M.; Kingsman, A. J. *J. Virol.* **1997**, *71*, 3474.
18. Morikawa, Y.; Kishi, T.; Zhang, W. H.; Nermut, M. V.; Hockley, D. J.; Jones, I. M. *J. Virol.* **1995**, *69*, 4519.
19. Suzuki, T.; Futaki, S.; Niwa, M.; Tanaka, S.; Ueda, K.; Sugiura, Y. *J. Biol. Chem.* **2002**, *277*, 2437.
20. Wender, P. A.; Mitchell, D. J.; Pattabiraman, K.; Pelkey, E. T.; Steinman, L.; Rothbard, J. B. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13003.
21. Matsushita, M.; Tomizawa, K.; Moriwaki, A.; Li, S. T.; Terada, H.; Matsui, H. *J. Neurosci.* **2001**, *21*, 6000.
22. Takenobu, T.; Tomizawa, K.; Matsushita, M.; Li, S. T.; Moriwaki, A.; Lu, Y. F.; Matsui, H. *Mol. Cancer Ther.* **2002**, *1*, 1043.
23. Wu, H. Y.; Tomizawa, K.; Matsushita, M.; Lu, Y. F.; Li, S. T.; Matsui, H. *Neurosci. Res.* **2003**, *47*, 131.
24. Rothbard, J. B.; Garlington, S.; Lin, Q.; Kirschberg, T.; Kreider, E.; McGrane, P. L.; Wender, P. A.; Khavari, P. A. *Nat. Med.* **2000**, *6*, 1253.
25. Ono, A. *J. Virol.* **2004**, *78*, 1552.
26. Rao, Z.; Belyaev, A. S.; Fry, E.; Roy, P.; Jones, I. M.; Stuart, D. I. *Nature* **1995**, *378*, 743.
27. Hill, C. P.; Worthylake, D.; Bancroft, D. P.; Christensen, A. M.; Sundquist, W. I. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 3099.
28. Kelly, B. N.; Howard, B. R.; Wang, H.; Robinson, H.; Sundquist, W. I.; Hill, C. P. *Biochemistry* **2006**, *45*, 11257.
29. Murdoch, C.; Monk, P. N.; Finn, A. *Immunology* **1999**, *98*, 36.

CHEMMEDCHEM

DOI: 10.1002/cmdc.201000548

Azamacrocyclic Metal Complexes as CXCR4 Antagonists

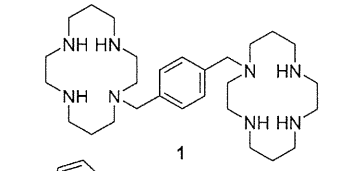
Tomohiro Tanaka,[a] Tetsuo Narumi,*[a] Taro Ozaki,[a] Akira Sohma,[a] Nami Ohashi,[a] Chie Hashimoto,[a] Kyoko Itotani,[a] Wataru Nomura,[a] Tsutomu Murakami,[b] Naoki Yamamoto,[b, c] and Hirokazu Tamamura*[a]

The chemokine receptor CXCR4 is a member of the seven transmembrane GPCR family, which is implicated in multiple diseases, including HIV infection, cancers, and rheumatoid arthritis. Low-molecular-weight nonpeptidic compounds, including AMD3100 and various pyridyl macrocyclic zinc(II) complexes, have been identified as selective antagonists of CXCR4. In the present study, structure–activity relationship studies were performed by combining the common structural features of alkylamino and pyridiyl macrocyclic antagonists. Several new zinc(II) or copper(II) complexes demonstrated potent anti-HIV activity, strong CXCR4-binding activity, and significant inhibitory activity against Ca^{2+} mobilization induced by CXCL12 stimulation. These results may prove useful in the design of novel CXCR4 antagonists, and the compounds described could potentially be developed as therapeutics against CXCR4-relevant diseases or chemical probes to study the biological activity of CXCR4.

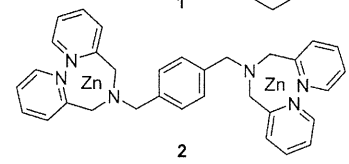
Introduction

The chemokine receptor CXCR4, which transduces signals of its endogenous ligand, CXCL12/stromal cell-derived factor-1 (SDF-1),[1–4] is classified as a member of the seven transmembrane GPCR family, and plays a physiological role via its interaction with CXCL12 in chemotaxis,[5] angiogenesis,[6,7] and neurogenesis[8,9] in embryonic stages. CXCR4 is, however, relevant to multiple diseases including HIV infection/AIDS,[10,11] metastasis of several types of cancer,[12–14] leukemia cell progression,[15,16] and rheumatoid arthritis (RA),[17,18] and is considered an attractive drug target to combat these diseases. Thus, inhibitors targeting CXCR4 are expected to be useful for drug discovery.

Several CXCR4 antagonists have been reported,[19–35] including our discovery of the highly potent CXCR4 antagonist T140, a 14-mer peptide with a disulfide bridge, its smaller derivative, the 5-mer cyclic peptide FC131, and several other potent analogues.[19,24–26,28–30] Clinical development of these peptidic antagonists could be pursued using specific administration strategies involving biodegradable microcapsules.[14,36] However, herein we focus on novel nonpeptidic low-molecular-weight CXCR4 antagonists. To date, AMD3100 (**1**),[20,22] Dpa-Zn complex (**2**),[37] KRH-1636,[27] and other compounds[31–35] have been developed in this and other laboratories as low-molecular-weight nonpeptidic CXCR4 antagonists. The present study reports structure–activity relationship studies based on the combination of common structural motifs, such as xylene scaffolds and cationic moieties that are present in the aforementioned compounds.



1



2

Results and Discussion

In order to determine spatially suitable positioning of cationic moieties, *p*- and *m*-xylenes were utilized as spacers. Cationic moieties such as bis(pyridin-2-ylmethyl)amine (dipicolylamine), 1,4,7,10-tetraazacyclododecane (cyclen), and 1,4,8,11-tetraazacyclotetradecane (cyclam) were introduced as R^1 and R^2 (Figure 1). This combination of R^1, R^2, and spacer groups led to the design and synthesis of compounds **12**–**31**.

The CXCR4 binding activity of synthetic compounds was assessed based on the inhibition of [^{125}I]CXCL12 binding to Jurkat cells, which express CXCR4.[38] The percent inhibition of all compounds at 1 μM is shown in Table 1. Seven compounds (**16**, **17**, **20**–**22**, **28**, and **29**, Table 1) resulted in greater than 87% inhibition. The high activity of **16** is consistent with re-

[a] *T. Tanaka, Dr. T. Narumi, T. Ozaki, A. Sohma, N. Ohashi, C. Hashimoto, K. Itotani, Dr. W. Nomura, Prof. H. Tamamura*
Institute of Biomaterials and Bioengineering
Tokyo Medical and Dental University
2-3-10 Kandasurugadai, Chiyoda-ku, Tokyo 101-0062 (Japan)
Fax: (+81) 3-5280-8039
E-mail: tamamura.mr@tmd.ac.jp

[b] *Dr. T. Murakami, Prof. N. Yamamoto*
AIDS Research Center, National Institute of Infectious Diseases
1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640 (Japan)

[c] *Prof. N. Yamamoto*
Department of Microbiology, Yong Loo Lin School of Medicine
National University of Singapore, Singapore 117597 (Singapore)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cmdc.201000548.

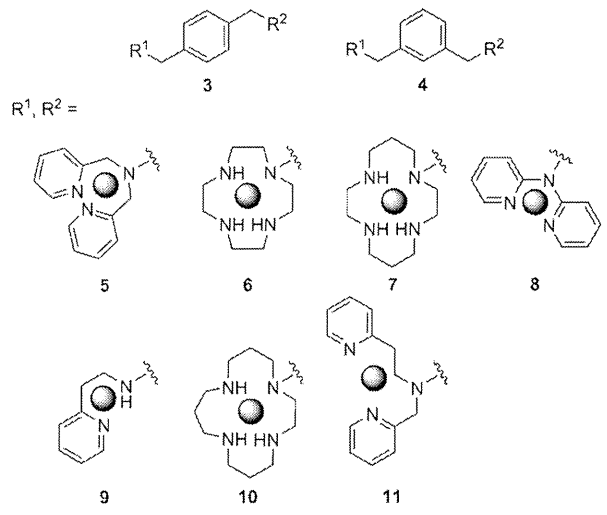


Figure 1. The structures of aromatic spacers (upper) and cationic moieties (R^1 and R^2). The shaded circle represents the position of the metal cation (Zn^{II} or Cu^{II}) in the chelate.

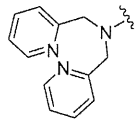
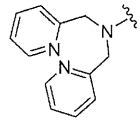
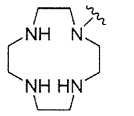
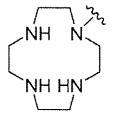
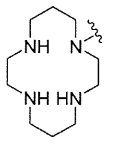
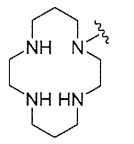
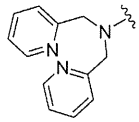
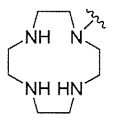
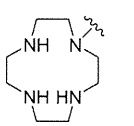
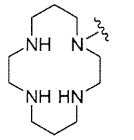
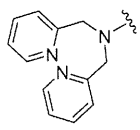
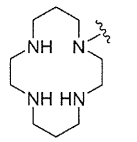
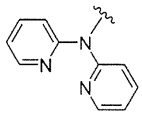
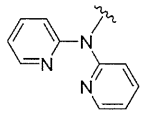
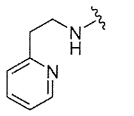
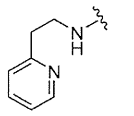
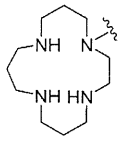
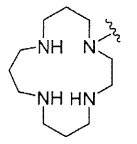
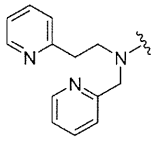
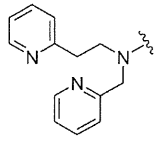
sults reported previously.[20,22] The anti-HIV activities of **17** and **29**, which contain only cyclam or cyclal rings, were reported by De Clercq et al.[39,40] Compounds with only pyridine and/or cyclen rings did not show any high binding activity. The presence of azamacrocyclic rings is presumably indispensable to the interaction of these compounds with CXCR4, and the size of rings appears to be important because not only compounds **16** and **17**, with two cyclam rings in the molecule, but also compounds **28** and **29**, with two cyclal rings, have remarkably more potent CXCR4 binding activity than compounds **14** and **15**, which have two cyclen rings. Compound **22**, with a *p*-xylene moiety, exhibited higher activity than compound **23**, which has an *m*-xylene moiety, indicating that *p*-xylene is more suitable than *m*-xylene as a spacer for approximate positioning of cationic moieties. At 0.1 μM, compound **22** resulted in 86% inhibition of [^{125}I]CXCL12 binding, while the other six compounds exhibited 37–66% inhibition. The IC_{50} value of compound **22** was estimated to be 37 nM.

$ZnCl_2$ was added to phosphate-buffered saline (PBS) solutions of these 20 compounds, **12–31**, to form zinc(II) complexes. The percent inhibition for each compound at 1 μM against [^{125}I]CXCL12 binding was determined and is given in Table 1. Zinc complexation of **12–15**, **18**, **19**, and **23** resulted in a remarkable increase in CXCR4 binding activity compared to the corresponding zinc-free compounds. These molecules contain dipicolylamine and/or cyclen moieties, suggesting that chelation of the nitrogen atoms with the zinc(II) ion significantly affects their interactions with CXCR4. The high activity of the zinc chelates of **12** and **13** is consistent with results provided in our previous paper.[37] Additionally, the anti-HIV activity of zinc complexes of **14** and **15** was reported by Kimura et al.[41] For compounds with only dipicolylamine and/or cyclen macrocycles as cationic moieties (**12–15**, **18**, and **19**), zinc complexation is critical to achieve high binding activity; the corresponding zinc-free compounds exhibit no significant activity. Compounds **16**, **17**, **20–22**, **28**, and **29** demonstrated high binding affinity in metal-free states as well as in zinc complexation states, indicating that zinc complexation of either of the macrocyclic rings in these compounds is not essential for high activity. The CXCR4 binding activity and anti-HIV activity of the zinc complex of **16** were reported previously.[42,43] Measured inhibition percentages for 0.1 μM of the zinc complexes of **12**, **14–23**, **28**, and **29** are given in Table 1. The zinc complexes of **20–22**, **28**, and **29** at 0.1 μM exhibited greater than 79% inhibition of [^{125}I]CXCL12 binding, and the other eight zinc complexes (of **12**, **14–19**, and **23**,) showed less than 55% inhibition. The IC_{50} values of zinc complexes of **20–22**, **28**, and **29** were estimated to be 11, 8.3, 22, 40, and 52 nM, respectively. Zinc complexes of compounds containing a combination of cyclen and cyclam moieties, **20** and **21**, had remarkably potent IC_{50} values.

To form chelates with a copper(II) cation, $CuCl_2$ was added to solutions in PBS of **12-31**. The inhibition percentages of all the compounds at 1 μM against [^{125}I]CXCL12 binding are shown in Table 1. Copper complexes of **14** and **15** exhibited a significant increase in CXCR4 binding activity as compared to the corresponding copper-free compounds, a phenomenon which is also seen in the zinc chelates. These compounds have two cyclen moieties in the molecules, suggesting that zinc or copper complexation is critical for high binding activity. Compounds **16**, **17**, and **20–22** showed high binding affinities in metal-free states and zinc- and copper-complexed states, indicating that metallic complexation of the cyclam rings in these compounds is not necessary for high activity. The CXCR4 binding activity of the copper complex of **16** was previously reported.[42] For compounds **17**, **22**, **23**, **28**, and **29**, copper complexation caused a significant decrease in binding activity compared to the corresponding copper-free compounds, whereas for compounds **14**, **15**, **18**, and **19**, copper complexation caused an increase in binding activity. This phenomenon may be due to the difference in ring sizes and structures of macrocycles, and was not observed upon zinc-complex formation. Inhibition at 0.1 μM of the copper complexes of **16** and **20–22**, which exhibited greater than 85% inhibition of [^{125}I]CXCL12 binding at 1 μM, are given in Table 1. The copper complexes of **16**, **20**, **21**, and **22** at 0.1 μM showed 39, 69, 88, and 39% inhibition, respectively, with the IC_{50} value of the copper complex of **21** estimated to be 16 nM.

Molecular modeling analysis of compound **21** and its zinc(II) and copper(II) complexes predicted that these complexes would form a stable coordinate conformation as shown in Figure 2. In general, zinc(II) complexes are predicted to adopt a tetrahedral conformation, while copper(II) complexes form a planar four coordinate/square conformation. The zinc(II) complex of **21** is predicted to have a tetrahedral conformation and the copper(II) complex a square planar conformation in both the cyclen and cyclam rings. The carboxyl group of either Asp 171 or Asp 262 in CXCR4 is thought to coordinate strongly with zinc ions but not copper ions in the complexes,[41–43] and as a consequence, the zinc complex of **21** would bind more strongly than **21** or its copper complex. This order of binding

Table 1. CXCR4 binding activity of compounds **12–31** in the metal ion-free form, the zinc complex, and the copper complex.

Compd	Spacer	R^1	R^2	Metal free			Zinc complex			Copper complex		
				Inhibition[a] [%]		IC_{50}[b]	Inhibition[a] [%]		IC_{50}[b]	Inhibition[a] [%]		IC_{50}[b]
				1 μM	0.1 μM	[nM]	1 μM	0.1 μM	[nM]	1 μM	0.1 μM	[nM]
12	*p*-xylene			0	n.d.	n.d.	83 ± 2	24 ± 5	n.d.	10 ± 4	n.d.	n.d.
13	*m*-xylene			0	n.d.	n.d.	31 ± 3	n.d.	n.d.	0	n.d.	n.d.
14	*p*-xylene			30 ± 4	n.d.	n.d.	87 ± 4	0	n.d.	60 ± 2	n.d.	n.d.
15	*m*-xylene			33 ± 2	n.d.	n.d.	94 ± 1	13 ± 6	n.d.	80 ± 3	n.d.	n.d.
16	*p*-xylene			94 ± 4	59 ± 6	n.d.	97 ± 5	28 ± 3	n.d.	98 ± 1	39 ± 3	n.d.
17	*m*-xylene			95 ± 3	49 ± 9	n.d.	98 ± 4	55 ± 7	n.d.	75 ± 1	n.d.	n.d.
18	*p*-xylene			32 ± 0.7	n.d.	n.d.	97 ± 6	0	n.d.	52 ± 3	n.d.	n.d.
19	*m*-xylene			17 ± 5	n.d.	n.d.	91 ± 4	0	n.d.	22 ± 6	n.d.	n.d.
20	*p*-xylene			89 ± 3	62 ± 3	n.d.	> 100	79 ± 1	11	> 100	69 ± 3	n.d.
21	*m*-xylene			89 ± 3	66 ± 3	n.d.	92 ± 3	> 100	8.3	> 100	88 ± 1	16
22	*p*-xylene			94 ± 3	86 ± 3	37	99 ± 8	79 ± 0.6	22	85 ± 3	39 ± 3	n.d.
23	*m*-xylene			58 ± 8	n.d.	n.d.	90 ± 17	37 ± 0.3	n.d.	48 ± 4	n.d.	n.d.
24	*p*-xylene			3 ± 0.9	n.d.	n.d.	0	n.d.	n.d.	0	n.d.	n.d.
25	*m*-xylene			4 ± 3	n.d.	n.d.	0	n.d.	n.d.	0	n.d.	n.d.
26	*p*-xylene			14 ± 2	n.d.	n.d.	10 ± 3	n.d.	n.d.	0	n.d.	n.d.
27	*m*-xylene			10 ± 3	n.d.	n.d.	10 ± 4	n.d.	n.d.	0	n.d.	n.d.
28	*p*-xylene			91 ± 0.4	37 ± 0.9	n.d.	97 ± 4	> 100	40	57 ± 4	n.d.	n.d.
29	*m*-xylene			87 ± 2	50 ± 1	n.d.	> 100	91 ± 4	52	55 ± 1	n.d.	n.d.
30	*p*-xylene			0	n.d.	n.d.	14 ± 3	n.d.	n.d.	14 ± 3	n.d.	n.d.
31	*m*-xylene			24 ± 2	n.d.	n.d.	20 ± 3	n.d.	n.d.	0	n.d.	n.d.
FC-131	*cyclo*-[D-Tyr-Arg-Arg-Nal-Gly-]			100	100	1.8	–	–	–	–	–	–

[a] CXCR4 binding activity was assessed based on inhibition of [^{125}I]CXCL12 binding to Jurkat cells. Percent inhibition for all compounds at 1 and 0.1 μM were calculated relative to the percent inhibition by FC131 (100%). [b] IC_{50} values are the concentrations which correspond to 50% inhibition of [^{125}I]CXCL12 binding to Jurkat cells. All data are mean values ± SEM of at least three independent experiments. n.d. = not determined.

affinities is commonly seen for these compounds and their zinc(II) or copper(II) complexes.

We investigated the CXCR4 antagonistic activity of compound **22** and the zinc complexes of **20**, **21**, **22**, and **28**, all of which possess strong CXCR4 binding activity. The CXCR4 antagonistic activity was assessed based on the inhibitory activity of the compounds against Ca^{2+} mobilization induced by CXCL12 stimulation through CXCR4 (figure S1 in the Support-

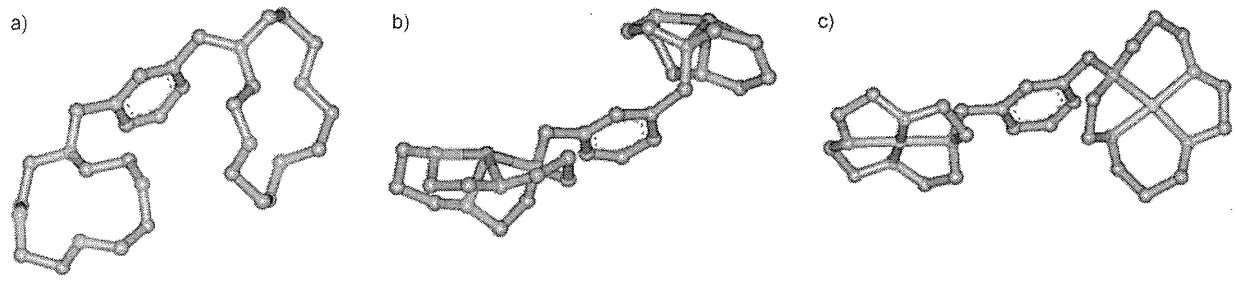


Figure 2. Structures calculated by molecular modeling of a) compound **21**, and its b) zinc and c) copper complexes. Atom color code: nitrogen = blue, carbon = gray, zinc = red, copper = light red.

ing Information). All of the tested compounds showed significant antagonistic activity at 1 μM.

The representative compounds **14**, **16**, **20–23**, **28**, and **29**, as well as their zinc chelates, were evaluated for anti-HIV activity. CXCR4 is the major co-receptor for the entry of T-cell-line-tropic (X4) HIV-1.[10,11] Inhibitory activity against X4-HIV-1 (NL4-3 strain)-induced cytopathogenicity in MT-4 cells was assessed and is shown in Table 2.[38] A correlation between CXCR4 binding activity and anti-HIV activity was observed. For compound **16** and its zinc complex, anti-HIV activity was significantly stronger than CXCR4 binding activity, and for the zinc complexes of compounds **20–22**, the CXCR4 binding activity is two to four-times stronger than the anti-HIV activity. The anti-HIV activity of the zinc complex of **16** was the most potent (EC_{50} = 8.2 nM). This is comparable to the anti-HIV activities of **16** and its zinc complex that were reported previously.[20,22,42,43] The zinc complex of **21**, which was the most active compound in terms of CXCR4 binding activity, also exhibited potent anti-HIV activity (EC_{50} = 36 nM).

Taken together, these results show that all of the compounds exhibiting CXCR4 binding activity also showed significant anti-HIV activity (EC_{50} values < 300 nM), and none of the tested compounds exhibited significant cytotoxicity (CC_{50} values > 10 μM; Table 2). Conversely, zinc complexes of **20**, **21**, **22**, and **28** did not exhibit significant anti-HIV activity against macrophage-tropic (R5) HIV-1 (NL(AD8) strain)-induced cytopathogenicity in PM-1 cells at concentrations below 10 μM. Since R5-HIV-1 strains use CCR5 instead of CXCR4 as the major co-receptor for entry, this suggests that these compounds do not bind CCR5 but rather are highly selective for CXCR4.

Table 2. Anti-HIV activity and cytotoxicity of representative compounds in the metal ion-free and zinc chelates.

Compd	Metal ion-free		Zinc chelate	
	EC_{50}[a] [nM]	CC_{50}[b] [μM]	EC_{50}[a] [nM]	CC_{50}[b] [μM]
14	200	> 10	200	> 10
16	21	> 10	8.2	> 10
20	38	> 10	39	> 10
21	50	> 10	36	> 10
22	93	> 10	48	> 10
23	290	> 10	220	> 10
28	36	> 10	56	> 10
29	130	> 10	42	> 10
FC131	93	> 10		
AZT	69	> 100		

[a] EC_{50} values are the concentrations corresponding to 50% protection from X4-HIV-1 (NL4-3 strain)-induced cytopathogenicity in MT-4 cells. [b] CC_{50} values are the concentrations at which the viability of MT-4 cells is reduced by 50%. All data are mean values from at least three independent experiments.

Conclusions

The present study introduces a new class of low-molecular-weight CXCR4 antagonists and their zinc(II) or copper(II) complexes, which contain pyridyl or azamacrocycle moieties with *p*-xylene or *m*-xylene spacers. These compounds demonstrated strong CXCR4 binding activity. Zinc complexes of **20** and **21**, which were the two most active compounds, contain cyclen and cyclam rings with *p*- and *m*-xylene spacers and exhibited remarkably potent IC_{50} values (11 and 8.3 nM, respectively). These compounds showed significant CXCR4 antagonistic activity, based on inhibitory activity against Ca^{2+} mobilization induced by CXCL12 stimulation through CXCR4, as well as potent anti-HIV activity, as assessed by protection from X4-HIV-1-induced cytopathogenicity in MT-4 cells. These results provide useful insights into the future design of novel CXCR4 antagonists, complementing information from other CXCR4 antagonists such as T140, FC131, and KRH-1636. Furthermore, these new compounds are useful for the development of therapeutic strategies for CXCR4-relevant diseases and chemical probes to study the biological activity of CXCR4.

Experimental Section

Chemistry

Compounds **12–17**, **20**, **21**, **24**, **25**, **27–29**, and **31** were synthesized as previously reported.[20,22,37,40,41,44–47] Compounds **18**, **19**, **22**, **23**, **26**, and **30** were synthesized in the present study; details are provided in the Supporting Information. A representative compound, **18**, was synthesized by coupling *p*-dibromoxylene (1,4-bis(bromomethyl)benzene) with tri-Boc-protected 1,4,7,10-tetraazacyclododecane, followed by treatment with trifluoroacetic acid and subsequent coupling with bis(pyridin-2-ylmethyl)amine. All crude compounds were purified by RP-HPLC and identified by FAB/ESI-

