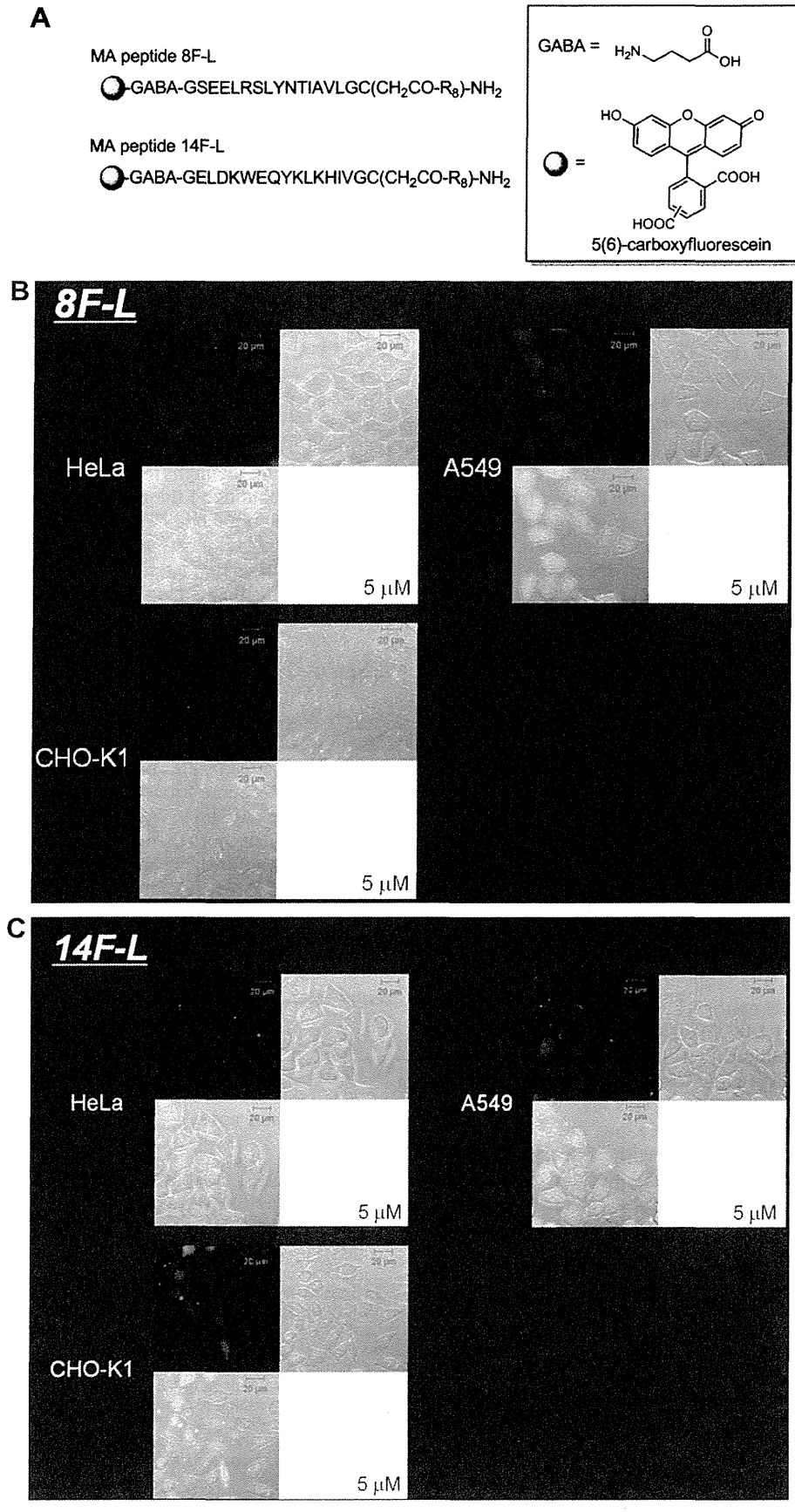


Figure 5. (A) The structures of fluorophore-labeled MA peptides 8F–L and 14F–L. (B) The fluorescent imaging of live cells HeLa, A549 and CHO-K1 by 8F–L. (C) The fluorescent imaging of live cells HeLa, A549 and CHO-K1 by 14F–L.

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.

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Effects of DNA Binding of the Zinc Finger and Linkers for Domain Fusion on the Catalytic Activity of Sequence-Specific Chimeric Recombinases Determined by a Facile Fluorescent System

Wataru Nomura,*[,†] Akemi Masuda,[†,‡] Kenji Ohba,[§] Arisa Urabe,[†] Nobutoshi Ito,[‡] Akihide Ryo,[∥] Naoki Yamamoto,[§] and Hirokazu Tamamura*[,†,‡]

[†]Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, 2-3-10 Kandasurugadai, Chiyoda-ku, Tokyo 101-0062, Japan

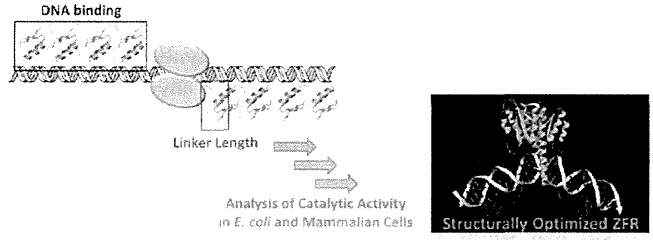
[‡]Graduate School of Biomedical Science, Tokyo Medical and Dental University, 1-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan

[§]Department of Microbiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore

[∥]Department of Microbiology and Molecular Biodefense Research, School of Medicine, Yokohama City University, 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004, Japan

Supporting Information

ABSTRACT: Artificial zinc finger proteins (ZFPs) consist of Cys_2-His_2-type modules composed of ~30 amino acids with a $\beta\beta\alpha$ structure that coordinates a zinc ion. ZFPs that recognize specific DNA target sequences can substitute for the binding domains of enzymes that act on DNA to create designer enzymes with programmable sequence specificity. The most studied of these engineered enzymes are zinc finger nucleases (ZFNs). ZFNs have been widely used to model organisms and are currently in human clinical trials with an aim of therapeutic gene editing. Difficulties with ZFNs arise from unpredictable mutations caused by nonhomologous end joining and off-target DNA cleavage and mutagenesis. A more recent strategy that aims to address the shortcomings of ZFNs involves zinc finger recombinases (ZFRs). A thorough understanding of ZFRs and methods for their modification promises powerful new tools for gene manipulation in model organisms as well as in gene therapy. In an effort to design efficient and specific ZFRs, the effects of the DNA binding affinity of the zinc finger domains and the linker sequence between ZFPs and recombinase catalytic domains have been assessed. A plasmid system containing ZFR target sites was constructed for evaluation of catalytic activities of ZFRs with variable linker lengths and numbers of zinc finger modules. Recombination efficiencies were evaluated by restriction enzyme analysis of isolated plasmids after reaction in *Escherichia coli* and changes in EGFP fluorescence in mammalian cells. The results provide information relevant to the design of ZFRs that will be useful for sequence-specific genome modification.



Artificial zinc finger proteins (ZFPs) can be used to engineer DNA binding domains with high specificity for desired target sequences, and ZFPs are a promising technology for gene therapy.[1−6] Modular assembly of ZFPs can create a DNA binding domain that targets virtually any sequence in the human genome.[3−5] By linking ZFPs to the catalytic domains of DNA-modifying enzymes, novel enzymes, including nucleases,[6] recombinases,[7−12] and methylases,[13−20] have been fabricated. These enzymes are endowed with programmable DNA binding specificity provided by the zinc finger protein fusion. Relevant to our development of ZFRs, recombinase enzymes from the serine recombinase family have been well studied.[21] In comparison with members of the tyrosine recombinase family such as Cre and Flp recombinases, the serine recombinases, including Tn3 and $\gamma\delta$ resolvases, Hin invertase, and Gin invertase, have DNA binding domains that are structurally independent of the catalytic domain. The structures of the catalytic domains and the sequences required for catalytic activity are highly conserved in these recombinases.[22] Tn3 and $\gamma\delta$ are among the best-characterized site-specific recombinase enzymes in the serine recombinase family. Only 35 amino acid residues differ between the $\gamma\delta$ and Tn3 resolvases, and their structures and functions are similar.[23] Negatively supercoiled DNA is a prerequisite for substrate recombination with native serine recombinase enzymes.[21] Although it is known that native serine recombinases require accessory proteins binding to sites I−III, activating mutants that require only the 28 bp of site I for successful recombination have been isolated.[7] In these hyperactivated enzymes, a DNA substrate in the form of negatively supercoiled DNA is not required for activity, and this allows application of activated catalytic domains with ZFPs to create zinc finger recombinases (ZFR). It has been suggested that reactions with serine recombinases proceed in three

Received: December 19, 2011
Revised: January 21, 2012
Published: January 23, 2012

steps: (i) formation of a dimer binding to the two forms of site I on the DNA, (ii) formation of a tetramer between the forms of site I, and (iii) strand exchange.[24,25] After the strand exchange reaction, the sequences between target sites are excised and the strands ligated (Figure 1).

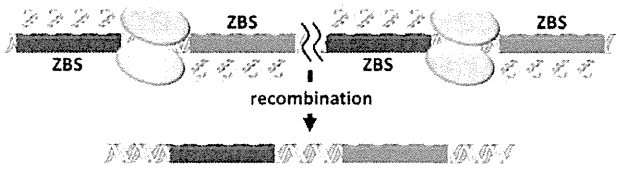


Figure 1. Schematic illustration of the ZFR reaction at a target site. The green and red boxes represent zinc finger binding sites (ZBSs). The yellow spheres represent catalytic domains of Tn3 resolvase.

ZFRs based on catalytic domain variants of Tn3, Gin, and Hin fused to artificial ZFPs have been shown to catalyze site-specific recombination in *Escherichia coli*[7−11] and mammalian cells.[8,9,11,12] ZFRs have also been shown to catalyze high-fidelity site-specific integration in mammalian cells.[9,11,12] While directed evolution of recombinase catalytic domains has proven to be essential for developing ZFR enzymes that function in mammalian cells, other aspects of ZFR design have not been thoroughly studied. In this report, we have synthesized ZFR mutants with variable numbers of zinc fingers and studied the role of peptide linkers that connect the Tn3 resolvase catalytic domain with the ZFP DNA binding domain. These effects are not readily addressed using molecular evolution strategies. For facile evaluation of recombination reactions in mammalian cells, a system that allows evaluation within 48 h was developed utilizing DsRed expression as a marker of transfection efficiency and EGFP expression as a marker of recombination efficiency. The results obtained revealed the optimal structures of the ZFRs, and the recombination efficiency results for linker mutants were verified by modeling studies.

■ EXPERIMENTAL PROCEDURES

Construction of ZFP Genes. ZFP genes were constructed as described previously.[26,27] Briefly, plasmid pc3XB encoding ZFPs purchased from Addgene (http://www.addgene.org) was repeatedly ligated. The zinc finger gene that was obtained was inserted into pMAL-p4x as an *Xba*I−*Bam*HI fragment for protein expression. A minor change was made to the multiple cloning site of pMAL-p4x (Figure S1 of the Supporting Information).

Target Enzyme-Linked Immunosorbent Assays (ELISAs). ELISA wells of 96-well plates were coated by incubation with 25 μL of 8 ng/mL streptavidin in PBS for 1 h at 37 °C. The plates were washed twice with dH_2O, and 25 μL of 5′-biotinylated hairpin oligonucleotide target in zinc buffer A (ZBA) [10 mM Tris-HCl (pH 7.5), 90 mM KCl, 1 mM $MgCl_2$, and 90 μM $ZnCl_2$] was added. After incubation for 1 h at 37 °C, plates were washed twice with dH_2O. Blocking solution (ZBA with 3% BSA, 175 μL) was added, and incubation continued for 1 h at 37 °C. The blocking solution was then removed; 25 μL of purified protein in ZBA was added, and 2-fold serial dilutions were performed into 1% BSA, 5 mM DTT, and 10 ng/μL salmon sperm DNA in ZBA. After incubation for 1 h at room temperature, the plates were washed 10 times with dH_2O and the monoclonal anti-MBP antibody (Sigma-Aldrich, 1:1000 dilution by ZBA with 1% BSA, 25 μL) was added. After incubation for 30 min at room temperature, the plates were washed 10 times with dH_2O and a diluted secondary anti-mouse IgG AP conjugate (Sigma-Aldrich, 1:1000 dilution by ZBA with 1% BSA, 25 μL) was added. After incubation for 30 min at room temperature, plates were washed 10 times with dH_2O. The alkaline phosphatase reaction was performed with *p*-nitrophenylphosphate for 30 min, and the absorbance at 405 nm was read with a microplate reader. The data were collected and plotted. The data were fit to the equation $y = 1/(1 + K_d/x)$, where y is the proportion of bound MBP−ZFP fusion protein to maximal binding derived from the absorbance at 405 nm and x is the concentration of the MBP−ZFP fusion protein. The K_d values are averages of three or more independent experiments, and standard errors of the mean (SEM) are shown.

Construction of ZFR Substrates. Each substrate plasmid contained a recombination cassette composed of two ZFR recombination sites flanking an EGFP gene as a stuffer sequence. Cassettes were assembled by amplifying the EGFP gene with primers encoding the ZFR site. The polymerase chain reaction (PCR) product was cloned into pAra-OP.[20] ZFP genes were amplified by PCR from plasmid pc3XB and inserted into the plasmid as *Eco*RI−*Sac*I fragments. Plasmids that contained ZFR with Gly-Ser linkers were mutated at the *Bst*BI site before insertion of the catalytic domain.

Construction of ZFR Genes. The DNA fragment of the Tn3 resolvase catalytic domain was amplified from pWL625 (ATCC accession number 31787) utilizing 5′-GAGGAGGAATTCATGCGACTTTTTGGTTACGCT-3′ and 5′-GAGGAGAAGCTTTCACGAGGCCCTTTCGTCTT-3′ as primers. The fragment was inserted into pBluescriptSK(−) as an *Eco*RI−*Hin*dIII fragment. Tn3-activating mutations (R2A, E56K, G101S, D102Y, M103I, and Q105L) were introduced into the Tn3 encoding gene. Linker sequences were amplified via PCR with the Tn3 fragment by primers that included the linker sequence. Tn3 fragments with different linkers were digested with *Eco*RI and *Bgl*II and ligated into similarly digested pAra-OP with the EGFP and ZFR sites. Tn3 fragments with various Gly-Ser linkers were also digested with *Eco*RI and *Bst*BI and then ligated. The plasmids were maintained with chloramphenicol.

Assay of Recombination of Plasmids in *E. coli*. The plasmid with a ZFR gene downstream from the arabinose promoter and the substrate sequences were introduced into *E. coli* by electroporation. After incubation for 14 h at 37 °C on an LB-agar plate, colonies were picked up and grown for 14 h at 37 °C in LB medium. Purified plasmids were digested with *Eco*RI for 1 h at 37 °C. After electrophoresis on a 0.8% agarose gel, the fragment intensity was estimated with ImageJ (Figure S2 of the Supporting Information).

Recombination Reaction of ZFR in Mammalian Cells. The EGFP gene, flanked by recombination sites, was inserted between *Nhe*I and *Kpn*I in pcDNA5/FRT (Life Technologies). A double-stranded oligonucleotide encoding the upstream target site was inserted into the *Mlu*I site, and the other oligonucleotide for the downstream target site was inserted into *Kpn*I and *Bam*HI sites. Cotransfection of the substrate plasmid and Flp expression plasmid (pOG44, Life Technologies) allowed site-specific integration into the single FLP recombinase target (FRT) site present in the Flp-In-CHO cell line (Life Technologies). Colony-acquired hygromycin resistance was characterized by fluorescently activated cell sorting (FACS) and genomic PCR. The sequence of the target site was confirmed. Cells were maintained in Ham's F-12 containing 10% (v/v)

dx.doi.org/10.1021/bi201878x | *Biochemistry* 2012, 51, 1510−1517

Table 1. DNA Binding Affinities of ZFPs

	two fingers	three fingers	four fingers	five fingers	six fingers
K_d (nM)[a]	160±20	23.6±3.6	12.8±1.1	15.4±1.4	12.9±1.4
R^2	0.90	0.87	0.94	0.94	0.94

[a]The values are averages of three or more independent experiments.

FBS and antibiotics (Wako Chemicals). The DsRed expression vector was constructed as follows; a DsRed-monomer sequence was ligated into pIRES2-EGFP (Clontech) to substitute for EGFP, and a Tn3-ZFP-NLS fragment was inserted between *Nhe*I and *Eco*RI in pIRES2-DsRed. On the following day, after 2 × 10^5 cells had been seeded, the ZFR expression vector was transfected into cells using Lipofectamine LTX Reagent and PLUS Reagent (Life Technologies). After being transfected for 48 h, cells were collected and analyzed by flow cytometry.

Molecular Modeling of the Linker Variants of ZFR. Computer models were generated using Discovery Studio (Accelrys Inc.). The crystal structure of the γδ resolvase–DNA complex [Protein Data Bank (PDB) entry 1GDT][22] was manually mutated in the protein and DNA to match the molecules used in this study. The first zinc finger module, obtained from a zinc finger–DNA complex (PDB entry 1MEY),[28] was placed on the resolvase–DNA complex by superimposing the phosphate backbone atoms of corresponding DNA residues. Appropriate linker atoms were then added and optimized by simulated annealing and energy minimization. During this optimization, the atoms in the resolvase, zinc fingers, and DNA were fixed, allowing only linker atoms to move.

■ RESULTS

Construction of Zinc Fingers and DNA Binding Analyses. The 18 bp target sequence of the zinc finger protein utilized in this study was 5′-CTGCATGCACTGGATGCA-3′.

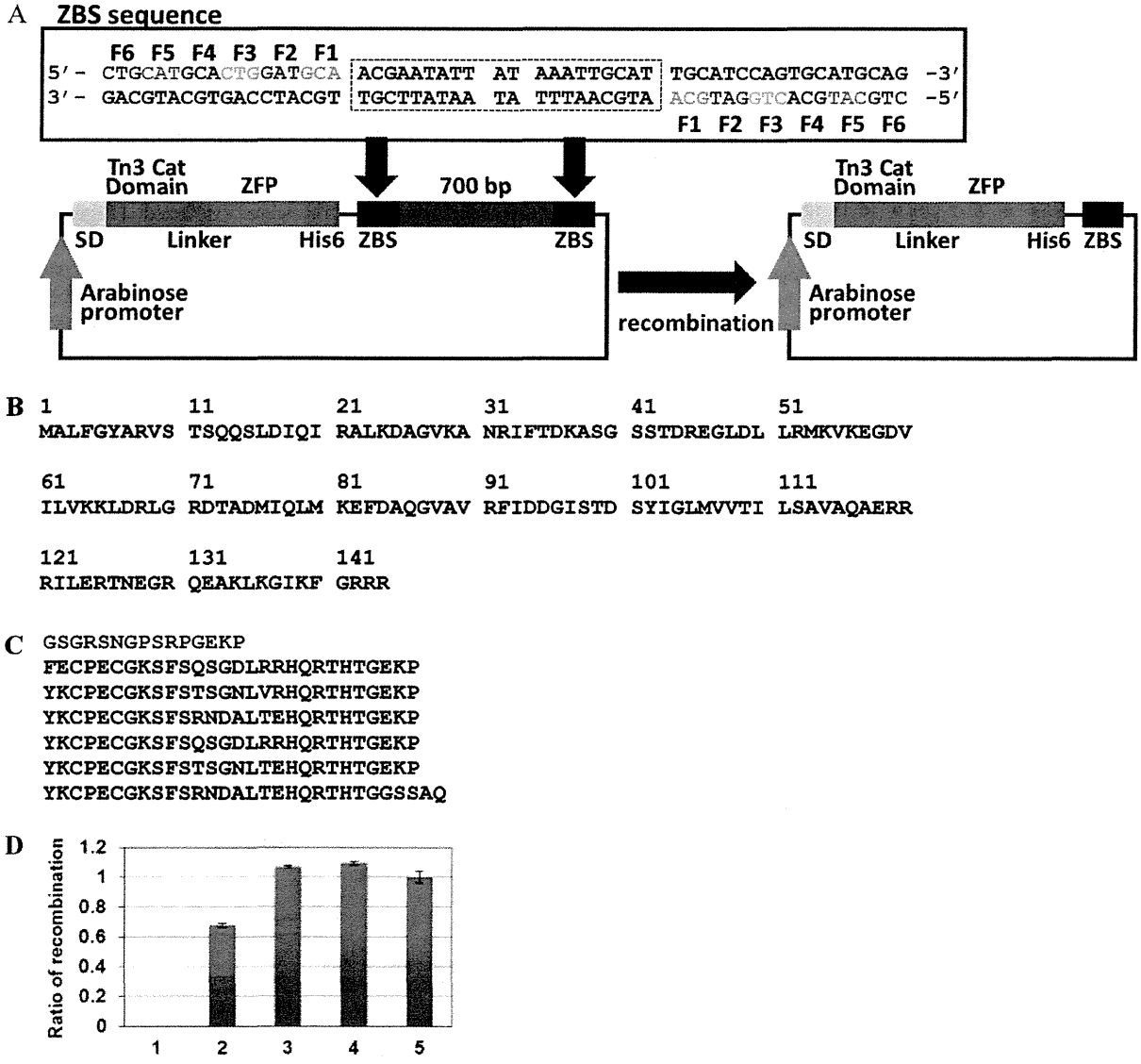


Figure 2. (A) Schematic of recombination at zinc finger binding sites (ZBSs). Recombination results in smaller plasmids. ZBS sequences are shown in the box. SD represents the Shine-Dalgarno sequence. (B) Amino acid sequences of the hyperactivated Tn3 catalytic domain. (C) Amino acid sequences of the linker (red) and six-zinc finger domain utilized for the analysis in *E. coli*. (D) Recombination efficiency depends on the number of fingers in ZFR. Columns 1–5 show the recombination efficiencies of two- through six-finger modules. The ratios are relative to the efficiency of the six-finger module. The error bars show the SEM of three or more independent experimental results.

A

```
Clone 1.          R RSNGPSRPGEKP
Clone 2. IKF      R RSNGPSRPGEKP
Clone 3. IKF GRRR
Clone 4. IKF GRRR EKP
Clone 5. IKF GRRR RPGEKP
Clone 6. IKF GRRR GPSRPGEKP
Clone 7. IKF GRRR RSNGPSRPGEKP
Clone 8. IKF GRRR GSGRSNGPSRPGEKP
Clone 9. IKF GRRR GSGGGSRSNGPSRPGEKP
Clone 10.IKF GRRR GSGGGSGGSRSNGPSRPGEKP
Clone 11.IKF GRRR GSGGGSGGSGGSRSNGPSRPGEKP
Clone 12.IKF GRRR GSGGGSGGSGGSGGSRSNGPSRPGEKP
Clone 13.IKF GRRR GSGGGSGGSGGSGGSGGSRSNGPSRPGEKP
Clone 14.IKF GRRR GSG
Clone 15.IKF GRRR GSGGGS
Clone 16.IKF GRRR GSGGGSGGS
Clone 17.IKF GRRR GSGGGSGGSGGS
Clone 18.IKF GRRR GSGGGSGGSGGSGGS
Clone 19.IKF GRRR GSGGGSGGSGGSGGSGGS
```

B

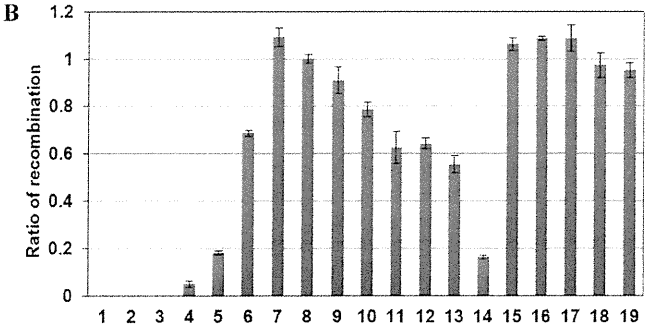


Figure 3. (A) Amino acid sequences of linkers of clones. All linkers were tested in the context of six-finger binding domains. (B) Results of recombination efficiency of clones with different linker sequences. The numbers of columns correspond to the clone numbers as described in panel A. The ratios are relative to the efficiency of clone 8. The error bars show the SEM of three or more independent experimental results.

Zinc fingers were constructed on the basis of a modular assembly strategy described by Barbas and co-workers.[27,29−32] Two- to six-finger proteins were constructed to obtain DNA binding domains with different affinities. Proteins were expressed as maltose binding protein fusions and purified with an MBPTrap column (GE Healthcare). The purity of the proteins was determined to be >90%. The DNA binding affinities were evaluated by an ELISA with the biotinylated hairpin oligonucleotide as a target.[10] The binding constants (K_d) of the two-, three-, four-, five-, and six-finger modules, listed in Table 1, were found to be 160, 23.6, 12.8, 15.4, and 12.9 nM, respectively. These results indicate that in the two-, three-, and four-finger modules, the DNA binding affinity increased with finger number but the binding affinities of ZFPs with four, five, and six fingers were similar.

Construction of ZFR Chimeric Proteins and Recombination Analysis in *E. coli*. The target DNA sequence of ZFR is shown in Figure 2A. The target site consists of a 20 bp spacer sequence flanked by 18 bp zinc finger binding sites. The spacer region was previously shown to be a Z+4 site in the target spacer of Z-resolvase.[7] For the evaluation of recombination in *E. coli*, a plasmid-based recombination system was constructed. The coding sequence of ZFRs was inserted into the plasmid containing a 700 bp stuffer sequence flanked with target sequences. In the recombination mediated by the expressed ZFRs, the stuffer sequence is excised to produce a smaller plasmid (Figure 2A). The amino acid sequences of the hyperactivated Tn3 catalytic domain, the linker between the domains, and the zinc finger domain are shown in panels B and C of Figure 2. The recombination efficiency was evaluated by a restriction enzyme assay. Plasmid purified from *E. coli* was digested by *Eco*RI, which is a single cutter of the plasmid. The linear plasmid was analyzed on an 0.8% agarose gel, and the fractions of the longer (nonrecombinant) and shorter (recombinant) plasmids were evaluated (Figure S2 of the Supporting Information). ZFR variants with different numbers of fingers were evaluated in this recombination system, and recombination ratios increased with increasing numbers of fingers from two to four fingers. The values of recombination efficiencies for ZFRs with four to six fingers were similar, reflecting the DNA binding affinities (Figure 2D). The production of recombinant sequence was confirmed by DNA sequencing analysis (Figure S3 of the Supporting Information).

In the next study, the reactions of ZFR variants with different linker lengths in the context of the six-finger module were tested (Figure 3B). In this experiment, 19 constructs were prepared. The variants were categorized into three groups depending on lengths and the compositions of linker sequences. The first group variants have short linkers with deletions within the catalytic

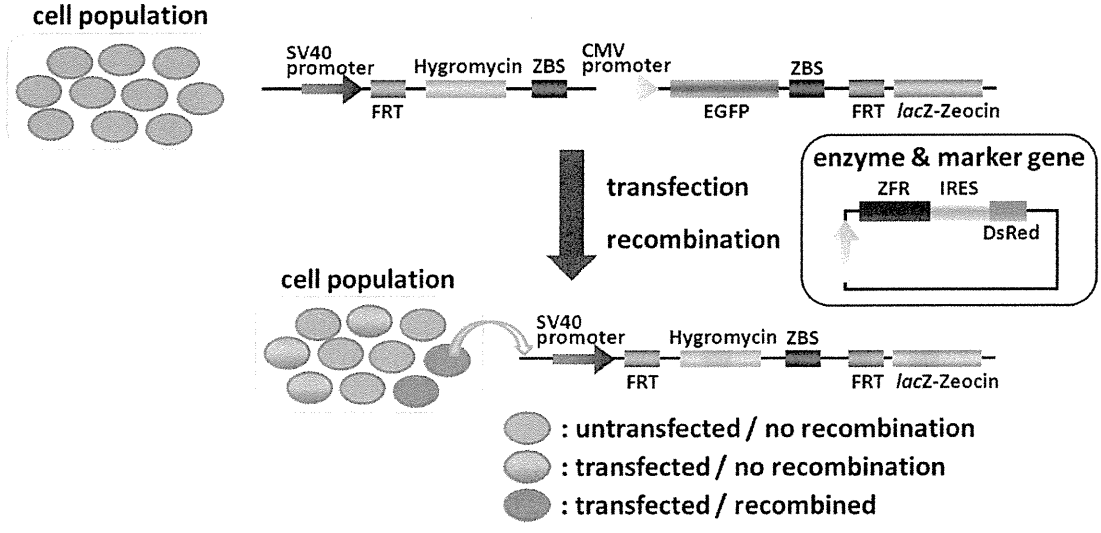


Figure 4. Recombination system constructed utilizing Flp-In-CHO-K1 cells.

domain of Tn3 resolvase (clones 1 and 2). The second group of variants has semirigid linkers (clones 3–13). The third group has flexible linker sequences composed of Gly-Ser sequences (clones 14–19). In the clones of the third group, the first two amino acids of the zinc finger domain, Tyr and Lys, are substituted with Phe and Glu, respectively. The recombination efficiencies were determined in the *E. coli*-based assay (Figure 3B). The results indicate two important phenomena. (1) The variant with a 12-amino acid linker was the most efficient (clone 7, Figure 3A), suggesting that there is an optimal linker length. (2) The variants with linkers composed of only Gly-Ser sequences were most efficient (clones 15–17, Figure 3A), indicating that ZFRs with flexible linkers tended to recombine most efficiently.

ZFR-Catalyzed Recombination in Mammalian Cells. To evaluate the recombination efficiency of ZFR variants in mammalian cells, we constructed a reporter cell line from Flp-In-CHO-K1 containing a cassette that encodes EGFP driven by a CMV promoter flanked by target sites (Figure 4). As each cell contains a single copy of the reporter gene, the recombination efficiency can be calculated from the proportion of cells with or without EGFP fluorescence. Additionally, the expression of ZFR was monitored by the expression of DsRed; this gene was placed downstream of the ZFR gene via a IRES sequence. The genes encoding ZFRs utilized in this study were amplified from a pAra plasmid shown in Figure 2A. Thus, the sequences of clones are the same as those utilized in experiments in *E. coli*.

With this reporter system, recombination efficiencies could be evaluated 48 h after transfection. Reported procedures involving retroviral-based transduction, selection, and evaluation take nearly 10 days.[8] The fluorescence intensity of cells was detected by FACS analysis (Figure S4 of the Supporting Information). The cells with recombinant genes were those that were EGFP-negative and DsRed-positive. The recombination efficiencies depended on the number of finger modules and on the linker lengths (Figure 5). As in *E. coli*, the five-finger proteins were the most efficient in recombination. The optimal linker length was six residues, which is different from that in *E. coli*. Additionally, recombination in mammalian cells was not as efficient as that in *E. coli*.

■ DISCUSSION

This study demonstrated that ZFR recombinases can be designed to specifically target sites in *E. coli* and mammalian cells and that recombination efficiency depends on the affinity of the ZFP for the DNA target and on the length of the linker between the DNA binding domain and the recombinase domain. The ZFR with five fingers had the highest recombination efficiency in both *E. coli* and CHO-K1 cells. The DNA binding affinity of this particular ZFP was saturated when the DNA binding domain had more than five fingers. The association and dissociation with DNA binding depend on the number of finger modules.[33] It is possible that the ZFR with five fingers was the most efficient recombination because the balance of association with dissociation and turnover was optimal. Guo et al. have also reported that four and five ZF domains are optimal for activity of ZFN.[34] On the basis of our data, the apparent K_d values of the four-, five-, and six-finger proteins derived from this particular ZFP were similar. The dependence on the number of finger modules was common in both *E. coli* and mammalian cells, but the recombination efficiency was lower in mammalian cells. In CHO-K1 cells, DNA is sequestered in chromatin structures. Additionally, the circular form of plasmid DNA could enhance recombination in the bacterial cells.

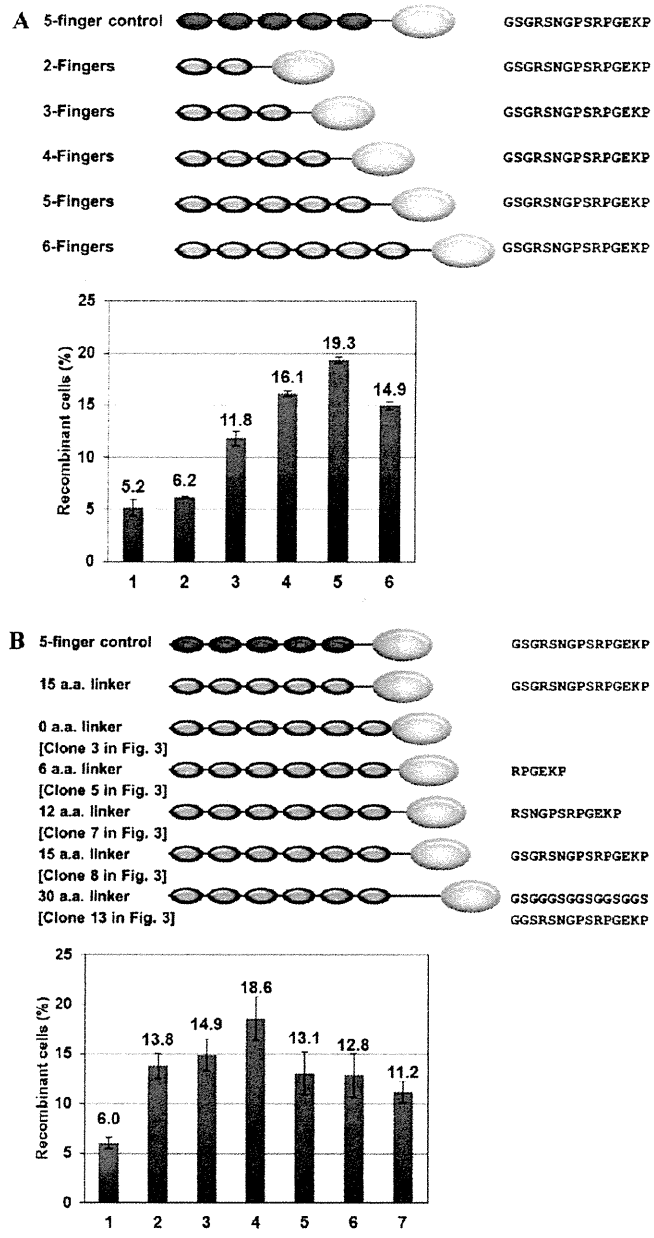


Figure 5. Recombination efficiency of ZFRs containing various numbers of fingers (A) and with various linkers (B) in mammalian cells. The top cartoons represent ZFR constructs utilized in the analyses. Green, blue, and yellow spheres represent zinc finger modules without sequence specificity, zinc finger modules with sequence specificity, and the Tn3 catalytic domain, respectively. Letters at the right of the cartoons are the linker sequences of the constructs. (A) Dependence on the number of fingers of ZFRs. The columns are as follows: column 1, five-finger control (nonspecific DNA binding); column 2, two fingers; column 3, three fingers; column 4, four fingers; column 5, five fingers; column 6, six fingers and different linker lengths. (B) Dependence on linker length. The columns are as follows: column 1, nontarget five-finger control with 15 amino acids; column 2, targeted five-finger ZFR with 15-amino acid linker; columns 3–7, targeted six-finger ZFRs with linker lengths of 0, 6, 12, 15, and 30 amino acids, respectively. The error bars show the SEM of three or more independent experimental results.

Recombination efficiency was dependent on the linker between the zinc finger domain and the recombinase domain. ZFRs with the shortest linkers had a very low efficiency of

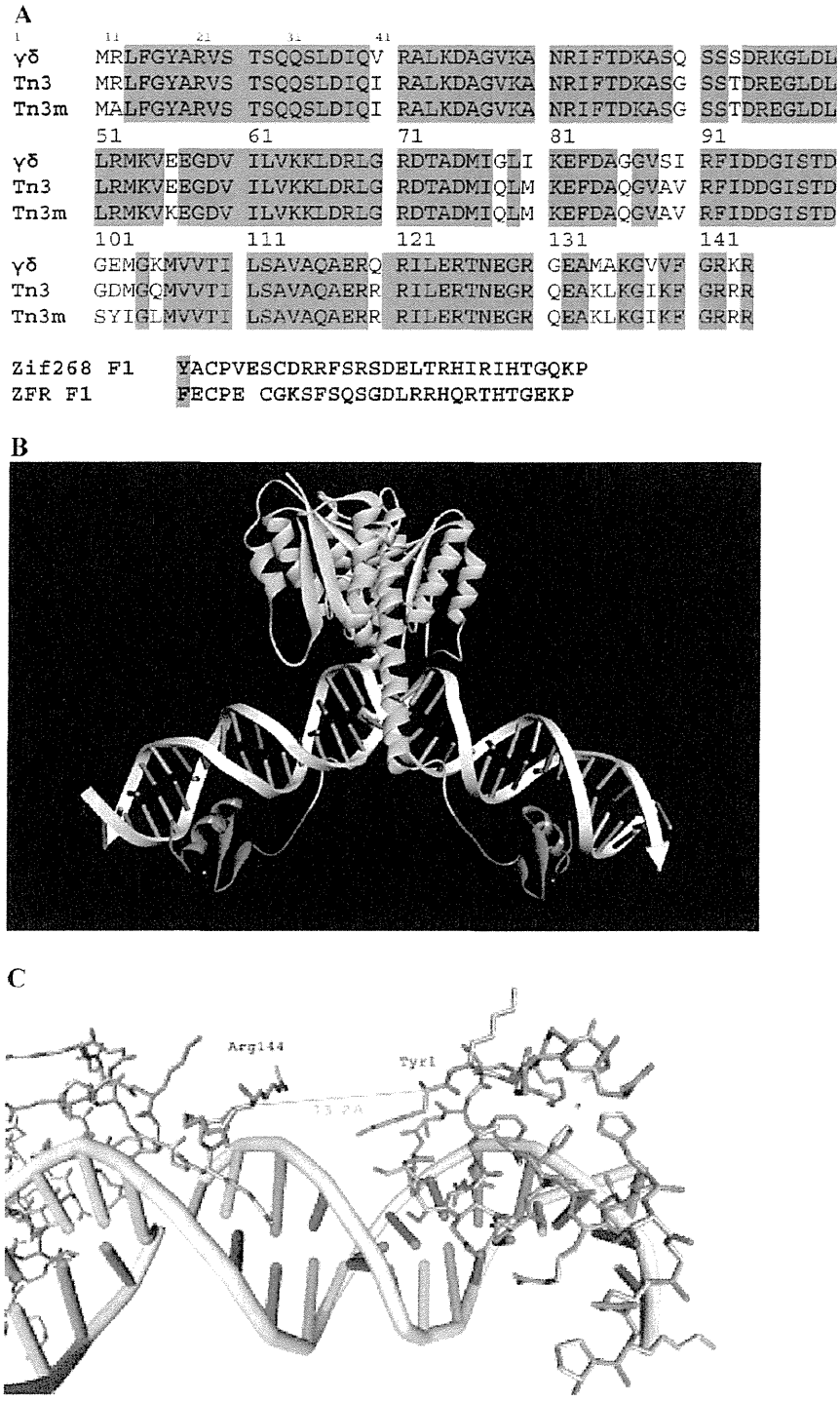


Figure 6. Representative result of molecular modeling of the resolvase domain and the first zinc finger module separated by a six-amino acid linker sequence. (A) Sequence alignment of resolvases *γδ* and Tn3 and the Tn3 hyperactivated mutant (Tn3m) (top), the first finger of zif268, and ZFR. Conserved residues are highlighted in red, and amino acid substitutions in the hyperactive mutant are highlighted in yellow. The N-terminal aromatic amino acids of zinc fingers are highlighted in blue. (B) The yellow ribbon indicates *γδ* resolvase, the red ribbon the six-amino acid linker, the green ribbon the N-terminal zinc finger domain, and the gray ribbon the zinc ion. (C) Distances between C*α* atoms of Arg144 and tyrosine (Tyr) at the N-terminus of *zif268*. The N-terminal amino acid of the zinc finger domain is phenylalanine (Phe) in ZFRs utilized in this study.

recombination in both bacterial and mammalian cells. Second, the length of linkers based on the original sequences was critical. Proteins with linkers containing 12 amino acid residues were the most efficient in recombination. In the Gly-Ser linker variants, the recombination efficiency reached a maximum at six amino acids. This result indicates that both the length and the flexibility of the linker are important.

A molecular modeling study was performed in an attempt to assess the reasons for the differences in recombination efficiency among the linker mutants. In the modeling of the ZFR complex with target DNA, the linker length of six amino acids was optimal for the DNA binding of ZFR when the linker sequence was flexible (Figure 6A). When the domains were modeled bound to the target sequence, the distance between

dx.doi.org/10.1021/bi201878x | *Biochemistry* 2012, 51, 1510–1517

the Cα atom of Arg144 in the γδ resolvase (Figure S5 of the Supporting Information) and that of Tyr at the N-terminus of the zinc finger domain is ~13.2 Å (Figure 6B,C). In polypeptides in the extended conformation, the distance between Cα atoms of sequential amino acids is 3.8 Å. Thus, a linker consisting of three amino acids (clone 4 or clone 14) should allow the protein to bind to both DNA regions, although these ZFRs had very low recombination efficiencies. In the complex with DNA, the amino groups at positions 145 and 146 of the main chain in γδ resolvase interact with the phosphate backbone of DNA and amino acids of these positions are involved in the folding of the catalytic domain (Figure S5 of the Supporting Information). In the case of clone 4, the Lys-Pro residues at the C-terminus of linker residues are involved in the folding of the zinc finger domain. Thus, these amino acids are considered to be members of both domains, not of the linker sequences. With this reasoning, the six and nine amino acids in the linkers for clones 4 and 5, respectively, are shorter than the theoretically optimal length. Moreover, in the sequences of the six- and nine-amino acid linkers, the amino acid at position 146 is Pro, which could disrupt the interaction with DNA phosphate, thus lowering the recombination efficiency. Consistent with these estimations, the Gly-Ser linker with six or nine amino acids (clones 15 and 16, respectively) showed the best recombination ratio. This evidence indicates that the residues at the C-terminus of the catalytic domain and the N-terminus of the zinc finger domain are involved in domain folding because Lys-Pro residues at the N-terminus of the zinc finger domains are not included in these clones. Variants around this optimal linker length, especially those with 12 and 15 amino acids, had similar recombination efficiencies. These results show that the flexibility of the linker is not necessary when the linker length is optimal. In mammalian cells, the variant with a linker of six amino acids (clone 5) showed the best recombination and the zero-amino acid linker (clone 3) showed better recombination than the variants with longer linkers of more than 12 amino acids. The reason for this effect is unclear, but it could be due to differences in the structures of target sites on the plasmid DNA compared to the genomic DNA. Additionally, the distances between the binding sites in these systems are different. In the genomic target, the binding sites are separated by sequences of more than 2500 bp.

In this study, a newly developed recombination system allowed measurement of recombination efficiencies of ZFRs in *E. coli* and in mammalian cells. In mammalian cells, recombination with genomic targets was evaluated within 48 h of the transient expression of recombinases. Artificial enzymes such as ZFN and ZFR have been studied mainly by using viral vector systems to deliver their genes into mammalian genomes. In a report describing utilization of the retrovirus vectors for gene delivery, the recombination efficiency was as high as ~18%.[8] In our study, we also observed up to 18% recombination in cells. This system could be utilized in future studies to evaluate function of ZFRs on specific targets.

■ ASSOCIATED CONTENT

Supporting Information

Details of subcloning, experimental results of plasmid digestion and sequencing, results of FACS analyses, and a description of key interactions in γδ resolvase. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: nomura.mr@tmd.ac.jp or tamamura.mr@tmd.ac.jp. Phone: +81-3-5280-8036. Fax: +81-3-5280-8039.

Funding

This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan (20790060), Health and Labor Sciences Research Grants from the Japanese Ministry of Health, Labor, and Welfare, and a grant from the Mochida Memorial Foundation for Medical and Pharmaceutical Research to W.N.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Beerli, R. R., Dreier, B., and Barbas, C. F. III (2000) Positive and negative regulation of endogenous genes by designed transcription factors. *Proc. Natl. Acad. Sci. U.S.A. 97*, 1495–1500.

(2) Pabo, C. O., Peisach, E., and Grant, R. A. (2001) Design and selection of novel Cys2His2 zinc-finger proteins. *Annu. Rev. Biochem. 70*, 313–340.

(3) Beerli, R. R., and Barbas, C. F. III (2002) Engineering polydactyl zinc-finger transcription factors. *Nat. Biotechnol. 20*, 135–141.

(4) Jamieson, A. C., Miller, J. C., and Pabo, C. O. (2003) Drug discovery with engineered zinc-finger proteins. *Nat. Rev. Drug Discovery 2*, 361–368.

(5) Blancafort, P., Segal, D. J., and Barbas, C. F. III (2004) Designing transcription factor architectures for drug discovery. *Mol. Pharmacol. 66*, 1361–1371.

(6) Carroll, D. (2008) Progress and prospects: Zinc-finger nucleases as gene therapy agents. *Gene Ther. 15*, 1463–1468.

(7) Akopian, A., He, J., Boocock, M. R., and Stark, W. M. (2003) Chimeric recombinases with designed DNA sequence recognition. *Proc. Natl. Acad. Sci. U.S.A. 100*, 8688–8691.

(8) Gordley, R. M., Smith, J. D., Gräslund, T., and Barbas, C. F. III (2007) Evolution of programmable zinc-finger-recombinases with activity in human cells. *J. Mol. Biol. 367*, 802–813.

(9) Gordley, R. M., Gersbach, C. A., and Barbas, C. F. III (2009) Synthesis of programmable integrases. *Proc. Natl. Acad. Sci. U.S.A. 106*, 5053–5058.

(10) Gersbach, C. A., Gaj, T., Gordley, R. M., and Barbas, C. F. III (2010) Directed evolution of recombinase specificity by split gene reassembly. *Nucleic Acids Res. 38*, 4198–4206.

(11) Gaj, T., Mercer, A. C., Gersbach, C. A., Gordley, R. M., and Barbas, C. F. III (2011) Structure-Guided Reprogramming of Serine Recombinase DNA Sequence Specificity. *Proc. Natl. Acad. Sci. U.S.A. 108*, 498–503.

(12) Gersbach, C. A., Gaj, T., Gordley, R. M., Mercer, A. C., and Barbas, C. F. III (2011) Targeted plasmid integration into the human genome by an engineered zinc-finger recombines. *Nucleic Acids Res. 39*, 7868–7878.

(13) Xu, G.-L., and Bestor, T. H. (1997) Cytosine methylation targeted to predetermined sequences. *Nat. Genet. 17*, 376–378.

(14) McNamara, A. R., Hurd, P. J., Smith, A. E., and Ford, K. G. (2002) Characterisation of site-biased DNA methyltransferases: Specificity, affinity and subsite relationships. *Nucleic Acids Res. 30*, 3818–3130.

(15) Carvin, C. D., Parr, R. D., and Kladde, M. P. (2003) Site-selective in vivo targeting of cytosine-5 DNA methylation by zinc-finger proteins. *Nucleic Acids Res. 31*, 6493–6501.

(16) Minczuk, M., Papworth, M. A., Kolasinska, P., Murphy, M. P., and Klug, A. (2006) Sequence-specific modification of mitochondrial DNA using a chimeric zinc-finger methylase. *Proc. Natl. Acad. Sci. U.S.A. 103*, 19689–19694.

(17) Li, F., Papworth, M., Minczuk, M., Rohde, C., Zhang, Y., Ragozin, S., and Jeltsch, A. (2007) Chimeric DNA methyltransferases

target DNA methylation to specific DNA sequences and repress expression of target genes. *Nucleic Acids Res. 35*, 100–112.
(18) Smith, A. E., and Ford, K. G. (2007) Specific targeting of cytosine methylation to DNA sequences in vivo. *Nucleic Acids Res. 35*, 740–754.
(19) Smith, A. E., Hurd, P. J., Bannister, A. J., Kouzarides, T., and Ford, K. G. (2008) Heritable Gene Repression through the Action of a Directed DNA Methyltransferase at a Chromosomal Locus. *J. Biol. Chem. 283*, 9878–9885.
(20) Nomura, W., and Barbas, C. F. III (2007) In vivo site-specific DNA methylation with a designed sequence-enabled DNA methylase. *J. Am. Chem. Soc. 129*, 8676–8677.
(21) Grindley, N. D., Whiteson, K. L., and Rice, P. A. (2006) Mechanisms of site-specific recombination. *Annu. Rev. Biochem. 75*, 567–605.
(22) Yang, W., and Steitz, T. A. (1995) Crystal structure of the site-specific recombinase gamma delta resolvase complexed with a 34 bp cleavage site. *Cell 82*, 193–207.
(23) Arnold, P. H., Blake, D. G., Grindley, N. D., Boocock, M. R., and Stark, W. M. (1999) Mutants of Tn3 resolvase which do not require accessory binding sites for recombination activity. *EMBO J. 18*, 1407–1414.
(24) Li, W., Kamtekar, S., Xiong, Y., Sarkis, G. J., Grindley, N. D., and Steitz, T. A. (2005) Structure of a synaptic $\gamma\delta$ resolvase tetramer covalently linked to two cleaved DNAs. *Science 309*, 1210–1215.
(25) Olorunniji, F. J., He, J., Wenwieser, S. V., Boocock, M. R., and Stark, W. M. (2008) Synapsis and catalysis by activated Tn3 resolvase mutants. *Nucleic Acids Res. 36*, 7181–7191.
(26) Gonzalez, B., Schwimmer, L. J., Fuller, R. P., Ye, Y., Asawapornmongkol, L., and Barbas, C. F. III (2010) Modular system for the construction of zinc-finger libraries and proteins. *Nat. Protoc. 5*, 791–810.
(27) Mandell, J. G., and Barbas, C. F. III (2006) Zinc Finger Tools: Custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res. 34*, W516–W523.
(28) Kim, C. A., and Berg, J. M. (1996) A 2.2 Angstroms resolution crystal structure of a designed zinc finger protein bound to DNA. *Nat. Struct. Mol. Biol. 3*, 940–945.
(29) Segal, D. J., Dreier, B., Beerli, R. R., and Barbas, C. F. III (1999) Toward controlling gene expression at will: Selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences. *Proc. Natl. Acad. Sci. U.S.A. 96*, 2758–2763.
(30) Dreier, B., Segal, D. J., and Barbas, C. F. III (2000) Insights into the molecular recognition of the 5′-GNN-3′ family of DNA sequences by zinc-finger domains. *J. Mol. Biol. 303*, 489–502.
(31) Dreier, B., Beerli, R. R., Segal, D. J., Flippin, J. D., and Barbas, C. F. III (2001) Development of zinc finger domains for recognition of the 5′-ANN-3′ family of DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem. 276*, 29466–29478.
(32) Dreier, B., Fuller, R. P., Segal, D. J., Lund, C., Blancafort, P., Huber, A., Koksch, B., and Barbas, C. F. III (2005) Development of zinc finger domains for recognition of the 5′-CNN-3′ family DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem. 280*, 35588–35597.
(33) Kamiuchi, T., Abe, E., Imanishi, M., Kaji, T., Nagaoka, M., and Sugiura, Y. (1998) Artificial nine zinc-finger peptide with 30 base pair binding sites. *Biochemistry 37*, 13827–13834.
(34) Guo, J., Gaj, T., and Barbas, C. F. III (2010) Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nuclease. *J. Mol. Biol. 400*, 96–107.

Bioorganic & Medicinal Chemistry 20 (2012) 3287–3291

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

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

Evaluation of a synthetic C34 trimer of HIV-1 gp41 as AIDS vaccines

Chie Hashimoto [a], Wataru Nomura [a,*], Aki Ohya [a], Emiko Urano [b], Kosuke Miyauchi [b], Tetsuo Narumi [a], Haruo Aikawa [a], Jun A. Komano [b], Naoki Yamamoto [c], 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] Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore

ARTICLE INFO

Article history:
Received 29 February 2012
Revised 21 March 2012
Accepted 21 March 2012
Available online 29 March 2012

Keywords:
C34 trimer
Dynamic supramolecular mechanism
gp41
HIV-1
Vaccine

ABSTRACT

An artificial antigen forming the C34 trimeric structure targeting membrane-fusion mechanism of HIV-1 has been evaluated as an HIV vaccine. The C34 trimeric molecule was previously designed and synthesized using a novel template with C3-symmetric linkers by us. The antiserum produced by immunization of the C34 trimeric form antigen showed 23-fold higher binding affinity for the C34 trimer than for the C34 monomer and showed significant neutralizing activity. The present results suggest effective strategies of the design of HIV vaccines and anti-HIV agents based on the native structure mimic of proteins targeting dynamic supramolecular mechanisms in HIV fusion.

1. Introduction

Highly active anti-retroviral therapy (HAART) involving new anti-HIV drugs such as protease inhibitors and integrase inhibitors has been brought a great success to us. Antibody-based therapy is also promising, and several AIDS antibodies have been developed by normal immunization[1] and by de novo techniques of monoclonal antibodies (mAb) using molecular evolution methods such as phage display.[2] Antibodies including anti-gp41 and anti-gp120 have been identified as human mAbs capable of highly and broadly neutralizing HIV. A transmembrane envelope glycoprotein, gp41 is hidden beneath an outer envelope glycoprotein gp120 and its ectodomain contains helical amino-terminal and carboxy-terminal leucine/isoleucine heptad repeat (HR) domains HR1 and HR2. These HR1 and HR2 regions are designated as the N-terminal helix (N36) and C-terminal helix (C34), respectively. In the membrane fusion of HIV-1, these helices join to form a six-helical bundle consisting of a central parallel trimer of N36 surrounded by C34 in an antiparallel hairpin fashion. A useful strategy to produce broadly neutralizing antibodies is therefore to synthesize molecules that mimic the natural trimer as it appears on viral surface proteins. Walker et al. reported that antibody recognition for the trimer form is important in HIV vaccine strategies, because antibodies that specifically recognize the trimer formation might have broad and potent neutralizing activity.[3] To date, several gp41 mimetics, especially for N36 regions, which assemble these helical peptides with branched peptide-linkers have been synthesized as antigens.[4–7] Previously, we synthesized a three-helix bundle mimetic, which corresponds to the trimeric form of N36, with a novel template with C3-symmetric linkers of equal lengths.[8] Immunization with the equivalent trimeric form of N36 mimetic produced antibodies with stronger binding affinity for N36 trimer than for N36 monomer. The structure-specific antibodies produced in this way showed significant neutralization activity against HIV-1 infection. Several potent anti-HIV-1 peptides based on the gp41 C-terminal HR2 region have been discovered[9,10] and an HR2-peptide, T20, has subsequently been developed into a clinical anti-HIV-1 drug, enfuvirtide (Roche/Trimeris).[11–14] The C-terminal helix C34 is also an HR2-derived peptide containing the amino acid residues required for docking into the hydrophobic pocket of the trimer of the N-terminal HR1 region, and potently inhibits HIV-1 fusion in vitro.[15] Recently, we also synthesized a three-helix bundle mimetic, which corresponds to the trimeric form of C34, with a novel different template with C3-symmetric linkers of equal lengths.[16] The C-terminal ends of three peptide strands are assembled in the C34 trimer, while the N-terminal ends of three peptide strands are assembled in N36 trimer. As an anti-HIV agent, the C34 trimer peptide showed two orders of magnitude higher inhibitory potency than the C34 monomer peptide. This study demonstrates a useful strategy for the design of effective inhibitors against viral infections that proceed by membrane fusion with host cells. In the present study, we have investigated the activity of the equivalent trimeric

* Corresponding authors. Tel.: +81 3 5280 8036; fax: +81 3 5280 8039.
E-mail addresses: nomura.mr@tmd.ac.jp (W. Nomura), tamamura.mr@tmd.ac.jp (H. Tamamura).

0968-0896/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmc.2012.03.050

form of C34 as an antigen peptide producing structure-specific antibodies. We have performed comparative studies of antisera isolated from mice immunized with the C34 trimer in binding affinity for the C34 trimer and for the C34 monomer.

2. Materials and methods

2.1. Immunization and sample collection

Six-week-old male BALB/c mice, purchased from Sankyo Laboratory Service Corp. (Tokyo, Japan), were maintained in an animal facility under specific pathogen-free conditions. The experimental protocol used was approved by the ethical review committee of Tokyo Medical and Dental University. Freund incomplete adjuvant and PBS were purchased from Wako Pure Chemical Industries (Osaka, Japan). DMSO (endotoxin free) was purchased from Sigma–Aldrich (St. Louis, MO).

All mice were bled one week before immunization. One hundred micrograms of antigen (C34 monomer C34REG) was dissolved in PBS (50 μL) and DMSO (1 μL). The antigen C34 trimer triC34e (100 μg) was dissolved in PBS (50 μL). This solution was mixed with Freund incomplete adjuvant (50 μL) and the mixture was injected subcutaneously under anesthesia on days 0, 7, 14, 21 and 28. Mice were bled on days 5, 12, 19, 26 and 33. Serum was separated by centrifugation (1500 rpm) at 4 °C for 10 min, and inactivated at 56 °C for 30 min. Sera were stored at −80 °C before use.

2.2. Serum titer ELISA

Tween-20 (polyoxyethylene (20) sorbitan monolaurate) and hydrogen peroxide (30%) were purchased from Wako Pure Chemical Industries (Osaka, Japan). 2,2′-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid diammonium salt (ABTS) was purchased from Sigma-Aldrich. Anti-mouse IgG (H+L)(goat)-HRP was purchased from EMD Chemicals (San Diego, CA). Ninety-six well microplates were coated with 25 μL of a synthetic peptide in a 10 μg/mL solution in PBS at 4 °C overnight. The coated plates were washed 10 times with deionized water and blocked with 150 μL of blocking buffer (0.02% PBST, PBS with 0.02% Tween 20, containing 5% skim milk) at 37 °C for 1 h. The plates were washed with deionized water 10 times. Mice sera were diluted in 0.02% PBST with 1% skim milk, and 50 μL of twofold serial dilutions of sera from 1/200 to 1/409600 were added to the wells and allowed to incubate at 37 °C for 2 h. The plates were again washed 10 times with deionized water. HRP-conjugated anti-mouse IgG, diluted 1:2000 in 0.02% PBST (25 μL), was added to each well. After incubation for 45 min, the plates were washed 10 times and 25 μL of HRP substrate, prepared by dissolving ABTS (10 mg) in 200 μL of HRP staining buffer—a mixture of 0.5 M citrate buffer (pH 4.0, 1 mL), H_2O_2 (3 μL), and H_2O (8.8 mL)—was added. After 30 min incubation, the reaction was stopped by addition of 25 μL/well 0.5 M H_2SO_4, and optical densities at 405 nm were recorded.

2.3. Virus preparation

For virus preparation, 293FT cells in a 60 mm dish were transfected with 10 μg of the pNL4-3 construct by the calcium phosphate method. The supernatant was collected 48 h after transfection, passed through a 0.45 μm filter, and stored at −80 °C as the virus stock.

2.4. Neutralizing assay (P24 assay)

For viral neutralizing assay, the NL4-3 virus (5 ng of p24) was bound to MT-4 cells (5×10^4 cells/200 μL) by spinoculation at 2100 *g* for 20 min at 4 °C. After removal by washing out of unbound virus, cells were resuspended with 200 μL of medium containing 10 μL sera from immunized or pre-immunized mice and were cultured. Half of the culture medium was changed every 2 or 3 days. At 7 days after infection, the level of p24 in the culture supernatant was determined by the p24 ELISA kit (PerkinElmer, MA).[17]

3. Results and discussion

In the C34 trimer, triC34e, which was previously synthesized,[16] the triplet repeat of arginine and glutamic acid (RERERE) was added to the C-terminal end of the C34 sequence to increase solubility in buffer solution, and glycine was fused to the C-terminus (Fig. 1A and B). The C3-symmetric template with three hydrophilic branches of equal length was adopted to assemble three peptide strands. As a control peptide, which corresponds to the monomeric form of C34, C34REG having RERERE and Gly in the C-terminus was used (Fig. 1C).[16]

To investigate whether antibodies are efficiently produced, mice were immunized with C34REG and triC34e and the increase in the titer in 5 weeks' immunization was observed (Fig. 2). Titers and specificity of antisera isolated from mice immunized with C34REG or triC34e were evaluated by serum titer ELISA against coated synthetic antigens. In each case, the increase in antibody production was observed as time passed. The most active antiserum for each antigen was utilized for the evaluation of binding activity by ELISA (Fig. 3). The C34REG-induced antibody showed approximately 1.2 times higher antibody titer against the coated C34REG than against the coated triC34e; the serum dilutions at 50% bound are 1.06×10^{-3} and 1.30×10^{-3}, respectively (Fig. 3A and B). The triC34e-induced antibody showed approximately 23 times higher titer against the coated triC34e than against the coated C34REG; the 50% bound serum dilutions are 3.15×10^{-4} and 7.30×10^{-3}, respectively (Fig. 3A and B). C34REG-induced or triC34e-induced antibody did not show any significant binding titer against an unrelated control peptide (Fig. 3C and D). Although purified monoclonal antibodies were not used for this evaluation, the antibodies produced exploited specific affinity for each antigen of the monomer or the trimer. These results suggest the synthesis of structure-involving antigens leads to the production of antibodies with structural specificity.

It is important to know if the antisera produced have inhibitory activity against HIV-1 infection. Accordingly, the inhibitory activity of the antisera was assessed by p24 assays utilizing the antisera bled from three mice that showed antibody production for each antigen (Fig. 4). The experiments using HIV-1 was performed in the biosafety level 3 laboratory #5 in the National Institute of Infectious Diseases. Sera from mice immunized with the monomer C34REG and the trimer triC34e antigens contained antiviral activities compared to control sera. Any significant difference of inhibitory effects was not observed between the sera isolated from C34REG-immunized mice and those from triC34e-immunized mice. The synthetic C34 trimeric antigen induces antibodies with a structural preference, but the levels of neutralization activity of sera from mice immunized with the C34 trimer were similar with those of sera from the C34 monomer-immunized mice. This suggests that antibodies with structural specificity against the gp41-C34-derived region do not always have more potent neutralization activity. The difference of recognition mechanism of two types of antibodies might cause different neutralizing mechanism although their levels of neutralization activity are almost the same. This result is not consistent with the data of the synthetic antigen molecules derived from N36, in which the N36 trimer-specific antibodies showed higher neutralization activity against HIV-1 infection than the N36 monomer-specific antibodies.[8] In any case,

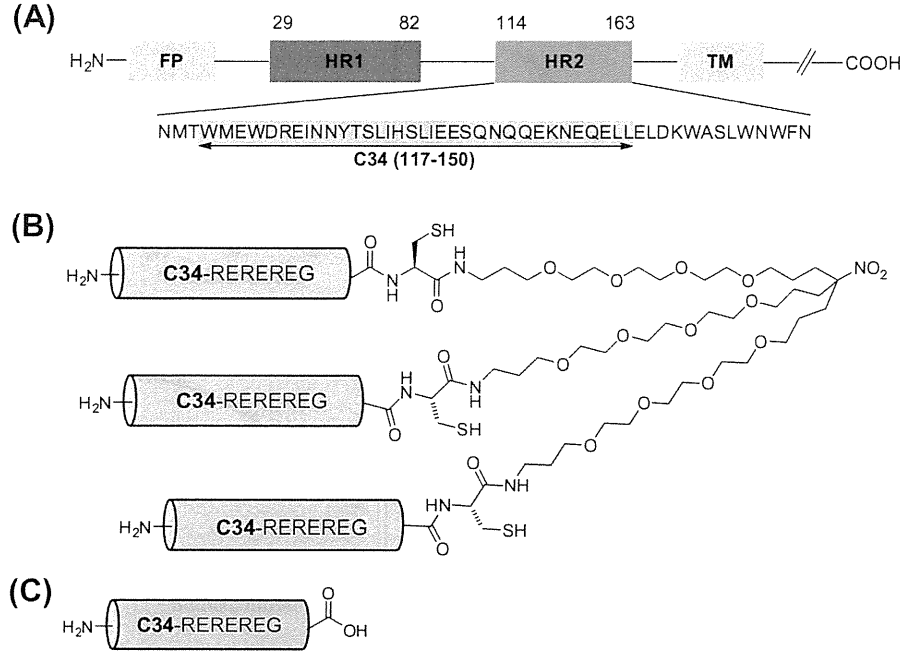


Figure 1. The sequence of C34 in gp41 (NL4-3) (A). FP and TM represent hydrophobic fusion peptide and transmembrane domain, respectively. Structures of C34-derived peptides, the C34 trimer with a C3-symmetric linker, triC34e (B), and the C34 monomer, C34REG (C).

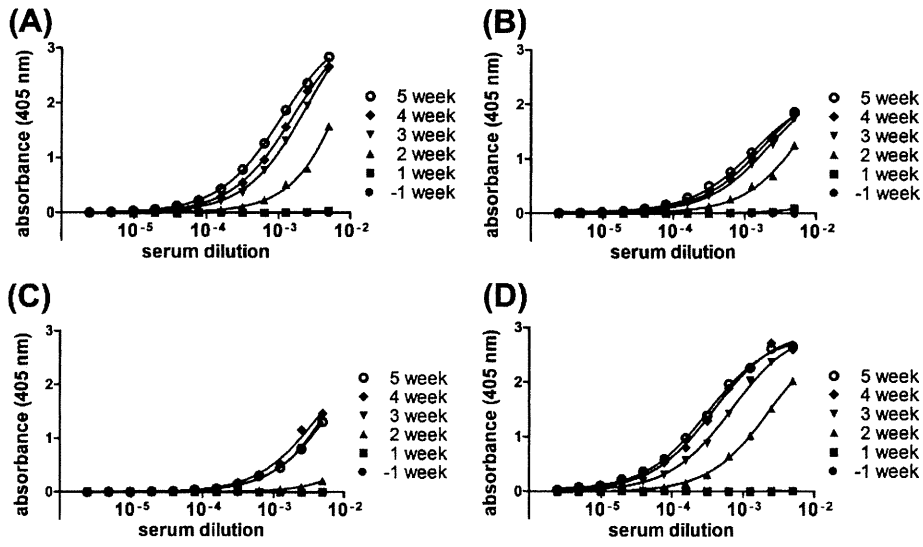


Figure 2. Results of serum titer ELISA of antisera collected during immunization (from one week before start to five weeks after immunization start) to determine the immunogenicity of designed antigens. The titers were evaluated as followings; antiserum against C34REG binding to C34REG (A); antiserum against C34REG binding to triC34e (B); antiserum against triC34e binding to C34REG (C); antiserum against triC34e binding to triC34e (D).

the synthetic C34 trimeric antigen induces antibodies with a structural preference and potent neutralization activity. In case antibodies bind to the gp41 C-terminal HR2 region and suppress membrane fusion, they may recognize the primary amino acid sequence of the C34 region or its structural conformation, because the C34 region is located outside in the formation of a six-helical bundle. It is suggested that suppressant potencies of these types of antibodies are almost similar. In addition, the action of these antibodies might be orthogonal and supplementally effective.

Recently, broadly active and potent neutralization antibodies, PG-9 and PG-16, were isolated from sera of HIV-1 infected individuals.[3] The antibodies can neutralize ~80% of HIV-1 isolates across all clades with approximately one order of magnitude higher potency than those of broad neutralizing mAbs reported previously. It is interesting that the epitopes for these mAbs are quaternary, and preferentially displayed on Env trimers, as expressed on the surface of virions and transfected cells. These results suggest that there may be production mechanisms for antibodies recognizing epitope structures.[18–20] The sera obtained from immunization of the C34 trimer antigen have structural specificity and neutralization activity. Thus, our trimer antigens, including the N36 trimer,[8] could work efficiently as a new class of HIV-1 vaccines.

Concerning inhibitory activity of these C-region peptides against HIV-1 entry, the potency of triC34e is one hundred times higher than that of C34REG.[16] It indicates that a trimeric form is critical as the active structure of the inhibitor, although as vaccines

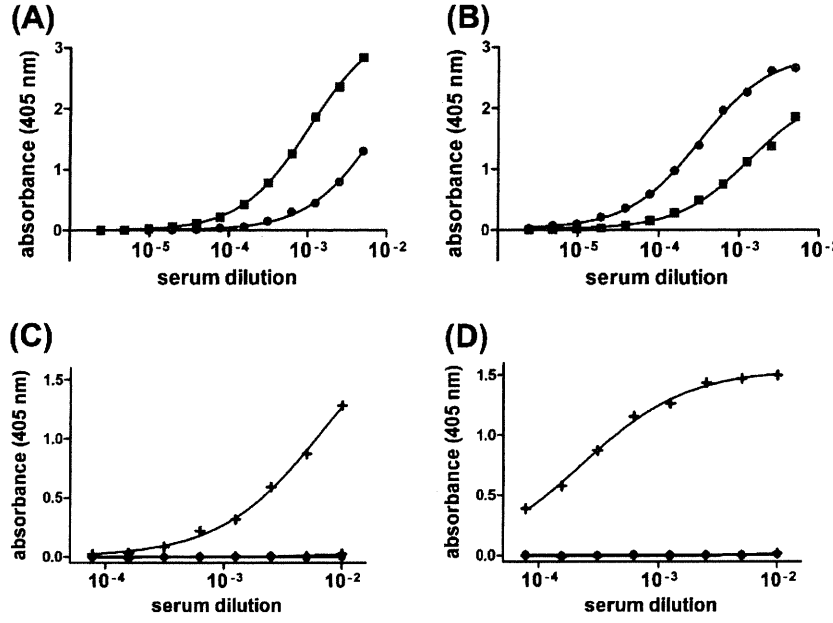


Figure 3. Serum titers of the antibodies produced by the fifth immunization of the C34REG antigen and the fourth immunization of the triC34e antigen. These titers were evaluated against ELISA templates of C34REG (monomer) (A) and triC34e (trimer) (B), using sera obtained from a C34REG-immunized mouse (■) and a triC34e-immunized mouse (●) as each representative. Titers of C34REG-induced antibodies were evaluated against ELISA templates of C34REG (+) and an unrelated control peptide (●) (C), and titers of triC34e-induced antibodies were evaluated against ELISA templates of triC34e (+) and an unrelated control peptide (●) (D). Unrelated control peptide: CH_3CO-GELDKWEKIRLRPGGGC(CH_2CONH_2)-NH_2.

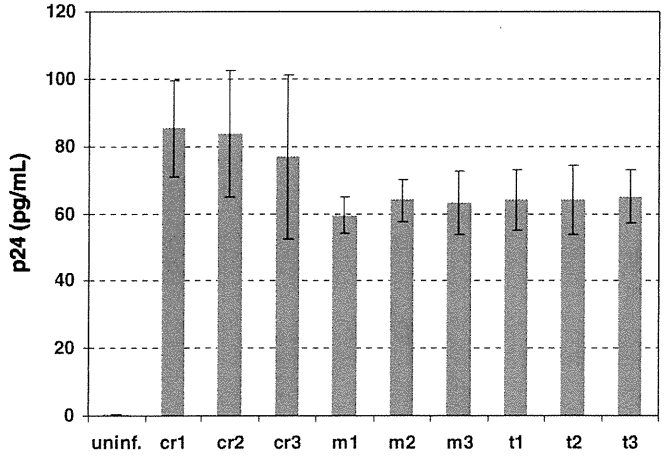


Figure 4. Determination of neutralization activity of the antibodies produced by immunization of C34REG and triC34e antigens. Inhibition of HIV-1 (NL4-3 strain) infection by produced antibodies was evaluated by the p24 assay in MT-4 cells. Y-axis shows the amount of p24 production. Uninf means uninfected cells. Pre-immunization sera (−1 week) were used as controls (cr1-3). C34REG- and triC34e-immunization sera (5 weeks) were used (m1-3 and t1-3, respectively). Experiments were conducted in triplicate. Error bars show standard error of mean.

there is no significant difference in neutralization activity of induced antibodies between the monomer and the trimer.

The exposed timing of epitopes of the helical region trimers is limited in the fusion step,[21] and carbohydrates are not included in the amino acid residues of the regions. The effectiveness of the vaccine design based on the gp41 helical regions is supported by the critical advantages cited above. Our developed N36 and C34 trimer-form specific antibodies might have the above properties. The designs of antigens and inhibitors targeting the dynamic supramolecular mechanism of HIV-1 fusion will be useful for future studies on AIDS vaccines and inhibitors.

Acknowledgments

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, Labour, and Welfare. C.H. was supported by JSPS Research Fellowships for Young Scientists.

References and notes

1. Cabezas, E.; Wang, M.; Parren, P. W.; Stanfield, R. L.; Satterthwait, A. C. *Biochemistry* **2000**, 39, 14377.
2. Burton, D. R.; Barbas, C. F. I. I. I.; Persson, M. A.; Koenig, S.; Chanock, R. M.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10134.
3. Walker, L. M.; Phogat, S. K.; Chan-Hui, P.-Y.; Wagner, D.; Phung, P.; Goss, J. L.; Wrin, T.; Simek, M. D.; Fling, S.; Mitcham, J. L.; Lehrman, J. K.; Priddy, F. H.; Olsen, O. A.; Frey, S. M.; Hammond, P. W.; Kaminsky, S.; Zamb, T.; Moyle, M.; Koff, W. C.; Poignard, P.; Burton, D. R. *Science* **2009**, *326*, 285.
4. De Rosny, E.; Vassell, R.; Wingfield, R. T.; Wild, C. T.; Weiss, C. D. *J. Virol.* **2001**, *75*, 8859.
5. Tam, J. P.; Yu, Q. *Org. Lett.* **2002**, *4*, 4167.
6. Xu, W.; Taylor, J. W. *Chem. Biol. Drug Des.* **2007**, *70*, 319.
7. Louis, J. M.; Nesheiwat, I.; Chang, L.; Clore, G. M.; Bewlet, C. A. *J. Biol. Chem.* **2003**, *278*, 20278.
8. Nakahara, T.; Nomura, W.; Ohba, K.; Ohya, A.; Tanaka, T.; Hashimoto, C.; Narumi, T.; Murakami, T.; Yamamoto, N.; Tamamura, H. *Bioconjugate Chem.* **2010**, *21*, 709.
9. Jiang, S.; Lin, K.; Strick, N.; Neurath, A. R. *Nature* **1993**, *365*, 113.
10. Wild, C. T.; Shugars, D. C.; Greenwell, T. K.; McDanal, C. B.; Matthews, T. J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 9770.
11. Kilby, J. M.; Hopkins, S.; Venetta, T. M.; DiMassimo, B.; Cloud, G. A.; Lee, J. Y.; Alldredge, L.; Hunter, E.; Lambert, D.; Bolognesi, D.; Matthews, T.; Johnson, M. R.; Nowak, M. A.; Shaw, G. M.; Saag, M. S. *Nat. Med.* **1998**, *4*, 1302.
12. Kilby, J. M.; Eron, J. J. *N. Eng. J. Med.* **2003**, *348*, 2228.
13. Lalezari, J. P.; Henry, K.; O'Hearn, M.; Montaner, J. S.; Piliero, P. J.; Trottier, B.; Walmsley, S.; Cohen, C.; Kuritzkes, D. R.; Eron, J. J., Jr.; Chung, J.; DeMasi, R.; Donatacci, L.; Drobnes, C.; Delehanty, J.; Salgo, M. *N. Eng. J. Med.* **2003**, *348*, 2175.
14. Liu, S.; Jing, W.; Cheng, B.; Lu, H.; Sun, J.; Yan, X.; Niu, J.; Farmar, J.; Wu, S.; Jiang, S. *J. Biol. Chem.* **2007**, *282*, 9612.
15. Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S. *Cell* **1997**, *89*, 263.
16. Nomura, W.; Hashimoto, C.; Ohya, A.; Miyauchi, K.; Urano, E.; Tanaka, T.; Narumi, T.; Nakahara, T.; Komano, J. A.; Yamamoto, N. *ChemMedChem* **2012**, *7*, 205.

17. Ohba, K.; Ryo, A.; Dewan, M. Z.; Nishi, M.; Naito, T.; Qi, X.; Inagaki, Y.; Nagashima, Y.; Tanaka, Y.; Okamoto, T.; Terashima, K.; Yamamoto, N. *J. Immunol.* **2009**, *183*, 524.
18. Pejchal, R.; Walker, L. M.; Stanfield, R. L.; Phogat, S. K.; Koff, W. C.; Poignard, P.; Burton, D. R.; Wilson, I. A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 11483.
19. Doores, K. J.; Burton, D. R. *J. Virol.* **2010**, *84*, 10510.
20. Walker, L. M.; Simek, M. D.; Priddy, F.; Gach, J. S.; Wagner, D.; Zwick, M. B.; Phogat, S. K.; Poignard, P.; Burton, D. R. *PLoS Pathog.* **2010**, *6*, e1001028.
[21]. Zwick, M. B.; Saphire, E. O.; Burton, D. R. *Nat. Med.* **2004**, *10*, 133–134.

Bioorganic & Medicinal Chemistry Letters 22 (2012) 4169–4172

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry Letters

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

Pharmacophore-based small molecule CXCR4 ligands

Tetsuo Narumi [a], Tomohiro Tanaka [a], Chie Hashimoto [a], Wataru Nomura [a], Haruo Aikawa [a], Akira Sohma [a], Kyoko Itotani [a], Miyako Kawamata [b], Tsutomu Murakami [b], Naoki Yamamoto [c], 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] Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore

ARTICLE INFO

Article history:
Received 16 March 2012
Revised 4 April 2012
Accepted 7 April 2012
Available online 20 April 2012

Keywords:
HIV entry inhibitors
Chemokine receptor
AIDS
Low molecular weight CXCR4 ligands

ABSTRACT

Low molecular weight CXCR4 ligands were developed based on the peptide T140, which has previously been identified as a potent CXCR4 antagonist. Some compounds with naphthyl, fluorobenzyl and pyridyl moieties as pharmacophore groups in the molecule showed significant CXCR4-binding activity and anti-HIV activity. Structure–activity relationships were studied and characteristics of each of these three moieties necessary for CXCR4 binding were defined. In this way, CXCR4 ligands with two types of recognition modes for CXCR4 have been found.

The chemokine receptor CXCR4 is classified into a family of G protein-coupled receptors (GPCRs), and transduces signals of its endogenous ligand, CXCL12/stromal cell-derived factor-1 (SDF-1).[1] The CXCR4-CXCL12 axis plays a physiological role in chemotaxis,[2] angiogenesis[3] and neurogenesis[4] in embryonic stages. The CXCR4 receptor is linked to many disorders including HIV infection/AIDS,[5] metastasis of cancer cells,[6] leukemia cell progression,[7] rheumatoid arthritis.[8] Since CXCR4 is an important drug target in these diseases, it is thought that effective agents directed to this receptor may be useful leads for therapeutic agents. To date, we and others have developed several potent CXCR4 antagonists. A highly potent antagonist, T140, a 14-mer peptide with a disulfide bridge, and its downsized analogue, FC131, with a cyclic pentapeptide scaffold, and several other related compounds have been reported.[9] Based on T140 and FC131, small-sized linear anti-HIV agents such as ST34 (**1**) have been developed (Fig. 1).[10] AMD3100,[11] KRH-1636,[12] Dpa–Zn complex (**2**)[13] and other azamacrocyclic compounds such as **3**,[14] which like **1**, contain benzylamine and electron-deficient aromatic groups, have also been reported as nonpeptidic antagonists. Compound **1** possesses significant anti-HIV activity but does not have high CXCR4 binding affinity. In the present study, more effective linear CXCR4 antagonists derived from compound **1** have been examined, and structure–activity relationship studies of these compounds have been performed.

Initially, three segments of compound **1** were selected for structural modification to support the design of new synthetic compounds: replacement of the 4-trifluoromethylbenzoyl group (Fig. 2, R^1), modification of the stereochemistry of the 1-naphthylethylamine moiety (R^2) and introduction of pyridine moieties on the nitrogen atom (R^3). In a previous study of T140 analogues, 4-fluorobenzoyl was found to be superior to 4-trifluoromethylbenzoyl as an N-terminal moiety. Thus, 4-fluorobenzyl, 4-fluorobenzoyl and 4-fluorophenylethyl groups were used as substitutes for the 4-trifluoromethylbenzoyl group (R^1) in **1**. The (*R*)-1-naphthylethylamine moiety in **1** is also present in KRH-1636 where it has the (*S*)-stereochemistry and thus both the (*R*) and (*S*)-stereoisomers were investigated in the present study. Several CXCR4 antagonists such as KRH-1636,[12] Dpa-Zn complex (**2**)[13] and Dpa-cyclam compound (**3**),[14] contain pyridyl rings. Thus, 2, 3, or 4-pyridylmethyl and 2, 3, or 4-pyridylethyl groups were introduced on the nitrogen atom of the 4-aminomethylbenzoyl group (R^3). With these modifications, a total of $3 \times 2 \times 6 = 36$ compounds (**12–47**) were designed (Fig. 2).

The synthesis of the structural fragment, Unit 1 is shown in Scheme 1. N-nosylation of 4-amino-methylbenzoic acid (**4**) with 2-nitrobenzenesulfonyl chloride and subsequent esterification gave the *t*-butyl ester **5**. Introduction of an R^3 moiety by means of a Mitsunobu reaction followed by removal of the Ns group yielded amines **6A–F**. Introduction of either 4-fluorobenzyl or 4-fluorophenylethyl groups by reductive amination of **6A–F** produced amines **7Ai–Fi** or **7Aiii–Fiii**, respectively. Conversion of **6A–F** to the appropriate amide (**7Aii-Fii**), and subsequent deprotection of the *tert*-butyl group yielded Unit 1, **8Ai–Fiii**.

* Corresponding author.
E-mail address: tamamura.mr@tmd.ac.jp (H. Tamamura).

0960-894X/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmcl.2012.04.032

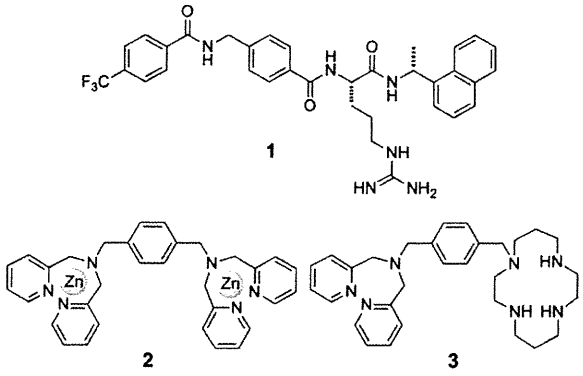


Figure 1. The structures of **1** (ST34), Dpa-Zn complex (**2**) and Dpa-cyclam compound (**3**).

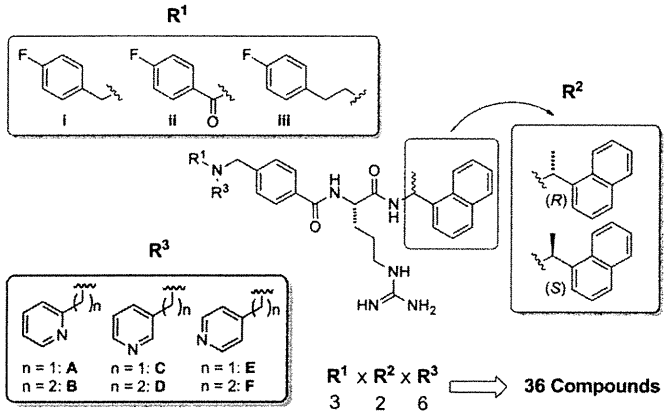
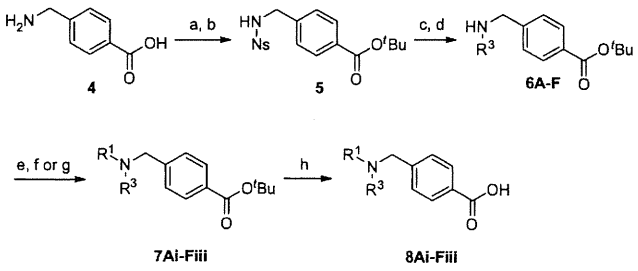


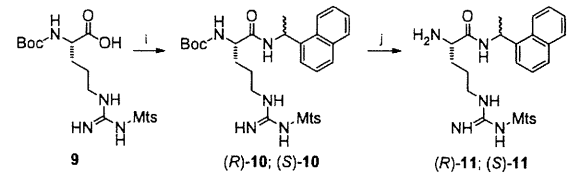
Figure 2. The structures of substituents for three parts of compound **1** in the design of new compounds.



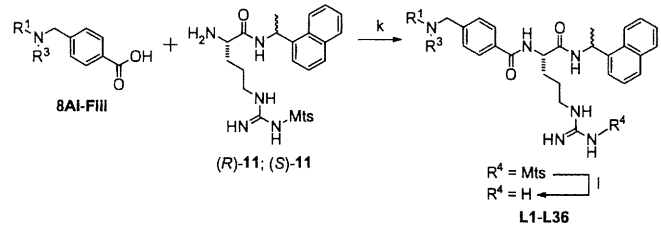
Scheme 1. The synthetic scheme of Unit 1, compounds **8Ai–Fiii**. Reagents and conditions and yields: (a) NsCl, Et_3N, THF/H_2O (1/1); (b) isobutene, THF/H_2SO_4 (10/1), 39% (2 steps); (c) PPh_3, DEAD, R^3OH, THF; (d) PhSH, K_2CO_3, DMF, 42–92% (2 steps); (e) $NaBH(OAc)_3$, 4-fluorobenzaldehyde, CH_2Cl_2; (f) $NaBH(OAc)_3$, (4-fluorophenyl)acetaldehyde, CH_2Cl_2; g) 4-fluorobenzoyl chloride, Et_3N, CH_2Cl_2, 51-94%; (h) TFA then 4 M HCl/EtOAc, quantitative; The structures of R^1 and R^3 are shown in Fig. 2 as i–iii and **A–F**, respectively. Ns = 2-nitrobenzenesulfonyl, ^{t}Bu = *tert*-butyl, DEAD = diethyl azodicarboxylate.

The synthesis of Unit 2 is shown in Scheme 2. Condensation of Boc-Arg(Mts)-OH (**9**) and (*R*)-1-naphthylethylamine or its (*S*) isomer produced amides (*R*)-**10** or (*S*)-**10**. Removal of the Boc group gave Unit 2, (*R*)-**11** and (*S*)-**11**.

Compounds **12–47** were synthesized by amide condensation of Unit 1, **8Ai–Fiii**, with Unit 2, (*R*)-**11** and (*S*)-**11**, and subsequent deprotection of the Mts group, as shown in Scheme 3.[15] All the synthetic compounds were purified by preparative reverse phase HPLC. In cases where peaks derived from side products appeared around the target peaks on the HPLC profile, the precise analysis was accomplished, giving rise to lower yields (Scheme 3, l).



Scheme 2. Synthetic schemes of Unit 2, compounds (***R/S*)-11**. Reagents and conditions: (i) EDCI·HCl, $HOBt·H_2O$, Et_3N, (*R/S*)-(+/−)-1-(1-naphthyl)ethylamine, CH_2Cl_2, 83–97%; (j) TFA then 4 M HCl/EtOAc, quantitative; EDCI·HCl = 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride, $HOBt·H_2O$ = 1-hydroxybenzotriazol monohydrate, Mts = 2,4,6-trimethylphenylsulfonyl, Boc = *tert*-butoxycarbonyl.



Scheme 3. Synthetic schemes of compounds **12–47**. Reagents and conditions: (k) EDCI·HCl, $HOBt·H_2O$, Et_3N, DMF, 36-95%; (l) TMSBr, *m*-cresol, 1,2-ethanedithiol, thioanisole, TFA, 4–54%. The structures of R^1 and R^3 are shown in Figure 2 as i–iii and **A–F**, respectively.

The CXCR4-binding activity of synthetic compounds was assessed in terms of the inhibition of [^{125}I]-CXCL12 binding to Jurkat cells, which express CXCR4.[16] The percent inhibition of all the compounds at 10 μM is shown in Table 1. Several of the compounds showed significant binding affinity. In general, compounds in which the 1-naphthylethylamine moiety (R^2) has the (*S*)-stereochemistry, as in KRH-1636, are more potent than the (*R*)-stereoisomers. Ten compounds (**26–28, 30, 33, 36, 39, 44, 45** and **47**, Table 1) were found to induce at least 30% inhibition and compounds **26**, **27** and **33**, which have a pyridyl group with a nitrogen atom at the β-position, showed more than 60% inhibition. It is noteworthy that compounds **26** and **27** in which R^2 is a (*R*)-1-naphthylethylamine moiety, are both more potent than the corresponding (*S*)-stereoisomers **44** and **45.** Compounds **26, 27** and **33**, have a 4-fluorobenzyl or 4-fluorophenylethyl group, which rather than an amide, is a reductive alkyl type (R^1). As can be seen from Table 1, there is a tendency for compounds with a pyridyl group with a nitrogen atom at the β-position (R^3: C or D), to be more potent in terms of CXCR4-binding activity than the corresponding compounds, which have a pyridyl group with a nitrogen atom at the α- or γ- position (R^3: A, B, E or F), and those with a reductive alkyl 4-fluorobenzyl or 4-fluorophenylethyl group (R^1: i or iii), to be more potent in CXCR4-binding activity than the corresponding compounds, with a 4-fluorobenzoyl group (R^1: ii).

Compounds were next evaluated for anti-HIV activity and cytotoxicity. CXCR4 is the major co-receptor for the entry of T-cell line-tropic (X4-) HIV-1.[5] Accordingly, inhibitory activity against X4-HIV-1 (NL4-3 strain)-induced cytopathogenicity in MT-4 cells (anti-HIV activity), and reduction of the viability in MT-4 cells (cytotoxicity) were assessed[16] and are shown in Table 1. Compounds **26** and **33–35** showed significant anti-HIV activity with EC_{50} values in the micromolar range. Compounds **26** and **33** showed both potent CXCR4-binding activity (79% and 60% inhibition at 10 μM, respectively) and anti-HIV activity (EC_{50} = 11 and 13 μM, respectively), the two activities being highly correlated. Compounds **34** and **35** have significant anti-HIV activity with EC_{50} values of 8 and 10 μM, respectively, which is higher than CXCR4-binding activities, which are 16% and 20% inhibition at 10 μM, respectively. Compound **27**, which showed relatively high CXCR4-binding activity (69% inhibition at 10 μM), failed to show

Table 1
CXCR4-binding activity, anti-HIV activity and cytotoxicity of compounds **12–47**

Compd no.	R^1 [a]	R^2 [b]	R^3 [c]	Inhibition[d] (%)	EC_{50}[e] (μM)	CC_{50}[f] (μM)	Compd no.	R^1 [a]	R^2 [b]	R^3 [c]	Inhibition[d] (%)	EC_{50}[e] (μM)	CC_{50}[f] (μM)
12	i	(*R*)	A	0	>20	35	**30**	i	(*S*)	A	30 ± 1.1	>4	11
13	i	(*R*)	B	4 ± 1.7	>4	23	**31**	i	(*S*)	B	25 ± 3.3	>20	24
14	i	(*R*)	C	6 ± 0.7	>20	37	**32**	i	(*S*)	C	27 ± 1.7	>20	41
15	i	(*R*)	D	24 ± 1.7	n.d.	n.d.	**33**	i	(*S*)	D	60 ± 1.5	13	65
16	i	(*R*)	E	12 ± 3.0	>20	39	**34**	i	(*S*)	E	16 ± 1.2	8	44
17	i	(*R*)	F	16 ± 2.2	n.d.	n.d.	**35**	i	(*S*)	F	20 ± 1.3	10	44
18	ii	(*R*)	A	3 ± 0.9	>20	38	**36**	ii	(*S*)	A	36 ± 1.8	>20	37
19	ii	(*R*)	B	6 ± 3.9	>20	41	**37**	ii	(*S*)	B	0	>20	43
20	ii	(*R*)	C	11 ± 0.8	>20	45	**38**	ii	(*S*)	C	14 ± 1.4	>20	57
21	ii	(*R*)	D	22 ± 4.1	n.d.	n.d.	**39**	ii	(*S*)	D	32 ± 8.4	n.d.	n.d.
22	ii	(*R*)	E	6 ± 2.7	>20	45	**40**	ii	(*S*)	E	13 ± 15	>20	51
23	ii	(*R*)	F	12 ± 1.9	n.d.	n.d.	**41**	ii	(*S*)	F	25 ± 13	>20	47
24	iii	(*R*)	A	15 ± 2.1	n.d.	n.d.	**42**	iii	(*S*)	A	16 ± 5.1	>4	9.9
25	iii	(*R*)	B	13 ± 0.6	>20	27	**43**	iii	(*S*)	B	23 ± 14	>4	13
26	iii	(*R*)	C	79 ± 14	11	47	**44**	iii	(*S*)	C	36 ± 13	n.d.	n.d.
27	iii	(*R*)	D	69 ± 5.0	>11	11	**45**	iii	(*S*)	D	35 ± 5.2	n.d.	n.d.
28	iii	(*R*)	E	44 ± 5.4	n.d.	n.d.	**46**	iii	(*S*)	E	26 ± 23	n.d.	n.d.
29	iii	(*R*)	F	0	n.d.	n.d.	**47**	iii	(*S*)	F	51 ± 6.6	n.d.	n.d.
KRH-1636				100	0.33	80	FC131				100	0.16	>10
AMD3100				n.d.	0.062	55	1 (ST34)				n.d.	7.4	66
AZT				n.d.	0.058	100							

[a,c] The structures of R^1 and R^3 are shown in Fig. 2 as i–iii and **A–F**, respectively.
[b] The absolute configuration in stereochemistry of R^2 shown in Fig. 2 is described.
[d] CXCR4-binding activity was assessed based on the inhibition of the [^{125}I]-CXCL12 binding to Jurkat cells. Inhibition percentages of all the compounds at 10 μM were calculated relative to the inhibition percentage by T140 (100%).
[e] EC_{50} values are the concentrations for 50% protection from X4-HIV-1 (NL4-3 strain)-induced cytopathogenicity in MT-4 cells.
[f] 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.

significant anti-HIV activity at concentrations below 11 μM because of high cytotoxicity (CC_{50} = 11 μM). With the exception of **27**, **30**, **42** and **43**, the tested compounds showed no significant cytotoxicity (CC_{50} >20 μM, Table 1). On the other hand, compounds **26**, **27**, **33**, **34** and **35** at concentrations below 100 μM failed to show significant protective activity against macrophage-tropic (R5-) HIV-1 (NL(AD8) strain)-induced cytopathogenicity in PM-1/CCR5, whereas the EC_{50} of the CCR5 antagonist SCH-D[17] in this assay was 0.055 μM (data not shown). Since instead of CXCR4, R5-HIV-1 strains use the chemokine receptor CCR5, a member of the GPCR family, as the major co-receptor for their entry, this suggests that these compounds do not bind to CCR5. Thus, compounds **26**, **27**, **33**, **34** or **35** have highly selective affinity for CXCR4. Compounds **34** and **35**, which have significant anti-HIV activity, have a pyridyl group with a nitrogen atom at the γ-position, in contrast to compounds **26**, **27** and **33** which also show CXCR4-binding activity, but have a pyridyl group with a nitrogen atom at the β-position. Furthermore, compounds **34** and **35** have R^1 = 4-fluorobenzyl and R^2 = (*S*)-1-naphthylethylamine. A possible explanation of these observations is that compounds **34** and **35** compete with HIV-1 in binding to CXCR4 while compounds **26** and **33** compete with HIV-1 and CXCL12. Compound **27** does not compete with HIV-1 because of its high cytotoxicity. This suggests that the CXCR4 binding site used by compounds **34** and **35** differs slightly from that used by compounds **26**, **27** and **33**.

Low molecular weight CXCR4 ligands with two types of recognition modes for CXCR4 have been obtained in this study: one causes competition with HIV-1 on CXCR4 whereas the other causes competition with HIV-1 and CXCL12. These compounds have selective affinity for CXCR4 because they do not significantly bind to CCR5. Further structural modification studies of these CXCR4 ligands are the subject of an ongoing project.

Acknowledgements

This work was supported by Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, Japan Human Science Foundation, and Health and Labour Sciences Research Grants from Japanese Ministry of Health, Labor, and Welfare. T.T. and C.H. are grateful for the JSPS Research Fellowships for Young Scientists.

References and notes

1. (a) Nagasawa, T.; Kikutani, H.; Kishimoto, T. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 2305; (b) Bleul, C. C.; Farzan, M.; Choe, H.; Parolin, C.; Clark-Lewis, I.; Sodroski, J.; Springer, T. A. *Nature* **1996**, *382*, 829; (c) Oberlin, E.; Amara, A.; Bachelerie, F.; Bessia, C.; Virelizier, J. L.; Arenzana-Seisdedos, F.; Schwartz, O.; Heard, J. M.; Clark-Lewis, I.; Legler, D. L.; Loetscher, M.; Baggiolini, M.; Moser, B. *Nature* **1996**, *382*, 833; (d) Tashiro, K.; Tada, H.; Heilker, R.; Shirozu, M.; Nakano, T.; Honjo, T. *Science* **1993**, *261*, 600.
2. Bleul, C. C.; Fuhlbrigge, R. C.; Casanovas, J. M.; Aiuti, A.; Springer, T. A. *J. Exp. Med.* **1996**, *2*, 1101.
3. (a) Tachibana, K.; Hirota, S.; Iizasa, H.; Yoshida, H.; Kawabata, K.; Kataoka, Y.; Kitamura, Y.; Matsushima, K.; Yoshida, N.; Nishikawa, S.; Kishimoto, T.; Nagasawa, T. *Nature* **1998**, *393*, 591; (b) Nagasawa, T.; Hirota, S.; Tachibana, K.; Takakura, N.; Nishikawa, S.; Kitamura, Y.; Yoshida, N.; Kikutani, H.; Kishimoto, T. *Nature* **1996**, *382*, 635.
4. (a) Zhu, Y.; Yu, Y.; Zhang, X. C.; Nagasawa, T.; Wu, J. Y.; Rao, Y. *Nat. Neurosci.* **2002**, *5*, 719; (b) Stumm, R. K.; Zhou, C.; Ara, T.; Lazarini, F.; Dubois-Dalcq, M.; Nagasawa, T.; Hollt, V.; Schulz, S. *J. Neurosci.* **2003**, *23*, 5123.
5. (a) Deng, H. K.; Liu, R.; Ellmeier, W.; Choe, S.; Unutmaz, D.; Burkhart, M.; Marzio, P. D.; Marmon, S.; Sutton, R. E.; Hill, C. M.; Davis, C. B.; Peiper, S. C.; Schall, T. J.; Littman, D. R.; Landau, N. R. *Nature* **1996**, *381*, 661; (b) Feng, Y.; Broder, C. C.; Kennedy, P. E.; Berger, E. A. *Science* **1996**, *272*, 872.
6. (a) Koshiba, T.; Hosotani, R.; Miyamoto, Y.; Ida, J.; Tsuji, S.; Nakajima, S.; Kawaguchi, M.; Kobayashi, H.; Doi, R.; Hori, T.; Fujii, N.; Imamura, M. *Clin. Cancer. Res.* **2000**, *6*, 3530; (b) Müller, A.; Homey, B.; Soto, H.; Ge, N.; Catron, D.; Buchanan, M. E.; McClanahan, T.; Murphy, E.; Yuan, W.; Wagner, S. N.; Barrera, J. L.; Mohar, A.; Verastegui, E.; Zlotnik, A. *Nature* **2001**, *410*, 50; (c) Tamamura, H.; Hori, A.; Kanzaki, N.; Hiramatsu, K.; Mizumoto, M.; Nakashima, H.; Yamamoto, N.; Otaka, A.; Fujii, N. *FEBS Lett.* **2003**, *550*, 79.
7. (a) Tsukada, N.; Burger, J. A.; Zvaifler, N. J.; Kipps, T. J. *Blood* **2002**, *99*, 1030; (b) Juarez, J.; Bradstock, K. F.; Gottlieb, D. J.; Bendall, L. J. *Leukemia* **2003**, *17*, 1294.
8. (a) Nanki, T.; Hayashida, K.; El-Gabalawy, H. S.; Suson, S.; Shi, K.; Girschick, H. J.; Yavuz, S.; Lipsky, P. E. . *J. Immunol.* **2000**, *165*, 6590; (b) Tamamura, H.; Fujisawa, M.; Hiramatsu, K.; Mizumoto, M.; Nakashima, H.; Yamamoto, N.; Otaka, A.; Fujii, N. *FEBS Lett.* **2004**, *569*, 99.
9. (a) Murakami, T.; Nakajima, T.; Koyanagi, Y.; Tachibana, K.; Fujii, N.; Tamamura, H.; Toshida, N.; Waki, M.; Matsumoto, A.; Yoshie, O.; Kishimoto, T.; Yamamoto, N.; Nagasawa, T. *J. Exp. Med.* **1997**, *186*, 1389; (b) 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; (c) Tamamura, H.; Omagari, A.; Oishi, S.; Kanamoto, T.; Yamamoto, N.; Peiper, S. C.; Nakashima, H.; Otaka, A.; Fujii, N. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2633; (d) 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; (e) Tamamura, H.; Hiramatsu, K.; Ueda, S.; Wang, Z.; Kusano, S.; Terakubo, S.; Trent, J. O.; Peiper, S. C.; Yamamoto, N.; Nakashima, H.; Otaka, A.; Fujii, N. *J. Med. Chem.* **2005**, *48*, 380; (f) Tamamura, H.; Araki, T.; Ueda, S.; Wang, Z.; Oishi, S.; Esaka, A.; Trent, J. O.; Nakashima, H.; Yamamoto, N.; Peiper, S. C.; Otaka, A.; Fujii, N. *J. Med. Chem.* **2005**, *48*, 3280; (g) Hashimoto, C.; Tanaka, T.; Narumi, T.; Nomura, W.; Tamamura, H. *Expert Opin. Drug Disc.* **2011**, *6*, 1067.

10. Tamamura, H.; Tsutsumi, H.; Masuno, H.; Mizokami, S.; Hiramatsu, K.; Wang, Z.; Trent, J. O.; Nakashima, H.; Yamamoto, N.; Peiper, S. C.; Fujii, N. *Org. Biomol. Chem.* **2006**, *4*, 2354.
11. Schols, D.; Struyf, S.; Van Damme, J.; Este, J. A.; Henson, G.; DeClarcq, E. *J. Exp. Med.* **1997**, *186*, 1383.
12. Ichiyama, K.; Yokoyama-Kumakura, S.; Tanaka, Y.; Tanaka, R.; Hirose, K.; Bannai, K.; Edamatsu, T.; Yanaka, M.; Niitani, Y.; Miyano-Kurosaki, N.; Takaku, H.; Koyanagi, Y.; Yamamoto, N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4185.
13. Tamamura, H.; Ojida, A.; Ogawa, T.; Tsutsumi, H.; Masuno, H.; Nakashima, H.; Yamamoto, N.; Hamachi, I.; Fujii, N. *J. Med. Chem.* **2006**, *49*, 3412.
14. Tanaka, T.; Narumi, T.; Ozaki, T.; Sohma, A.; Ohashi, N.; Hashimoto, C.; Itotani, K.; Nomura, W.; Murakami, T.; Yamamoto, N.; Tamamura, H. *ChemMedChem* **2011**, *6*, 834.
15. *For example, the synthesis of compound* ***30***: To a stirred solution of **8Ai** (176 mg, 0.415 mmol, HCl salt) in DMF (4 mL) were added EDCI·HCl (104 mg, 0.454 mmol), HOBt·H_2O (58.4 mg, 0.381 mmol), Et_3N (301 μL, 2.16 mmol) and ***(S)*-11** (320 mg, 0.657 mmol, HCl salt) at 0 °C. The mixture was stirred at room temperature for 43 h. The reaction mixture was diluted with $CHCl_3$ and washed with saturated citric acid, saturated $NaHCO_3$ and brine, and dried over $MgSO_4$. Concentration under reduced pressure followed by flash column chromatography over silica gel with $CHCl_3$/MeOH (20/1) gave the condensation product (175 mg, 0.208 mmol, 50% yield) as white powder. To this compound were added *m*-cresol (75.0 μL, 0.714 mmol), 1,2-ethanedithiol (225 μL, 2.68 mmol), thioanisole (225 μL, 1.91 mmol), TFA (3 mL) and bromotrimethylsilane (495 μL, 3.82 mmol) with stirring at 0 °C, and the stirring was continued at room temperature for 3.5 h under N_2. The reaction mixture was concentrated under reduced pressure, followed by addition of Et_2O to precipitate the product. After washing with Et_2O, the crude product was purified by preparative HPLC and lyophilized to give the compound **30** (15.6 mg, 0.0236 mmol, 13%) as white powder. ^1H NMR δ_H (400 MHz; DMSO-d_6) 1.49 (m, 2H), 1.51 (d, J = 7.2 Hz, 3H), 1.80–1.62 (m, 2H), 3.07 (dd, J = 6.4, 12.8 Hz, 2H), 3.85 (s, 2H), 3.91 (s, 4H), 4.54 (m, 1H), 5.72 (m, 1H), 7.13 (t, J = 8.8 Hz, 2H), 7.40 (m, 1H), 7.60–7.45 (m, 10H), 7.75–7.95 (m, 5H), 8.10 (m, 1H), 8.40 (d, J = 8.0 Hz, 1H), 8.58 (m, 1H), 8.65 (d, J = 7.6 Hz, 1H); LRMS (ESI), *m/z* calcd for $C_{39}H_{42}FN_7O_2$ $(MH)^+$ 660.34, found 660.31.
16. Tanaka, T.; Tsutsumi, H.; Nomura, W.; Tanabe, Y.; Ohashi, N.; Esaka, A.; Ochiai, C.; Sato, J.; Itotani, K.; Murakami, T.; Ohba, K.; Yamamoto, N.; Fujii, N.; Tamamura, H. *Org. Biomol. Chem.* **2008**, *6*, 4374.
17. Tagat, J. R.; McCombie, S. W.; Nazareno, D.; Labroli, M. A.; Xiao, Y.; Steensma, R. W.; Strizki, J. M.; Baroudy, B. M.; Cox, K.; Lachowicz, J.; Varty, G.; Watkins, R. *J. Med. Chem.* **2004**, *47*, 2405.