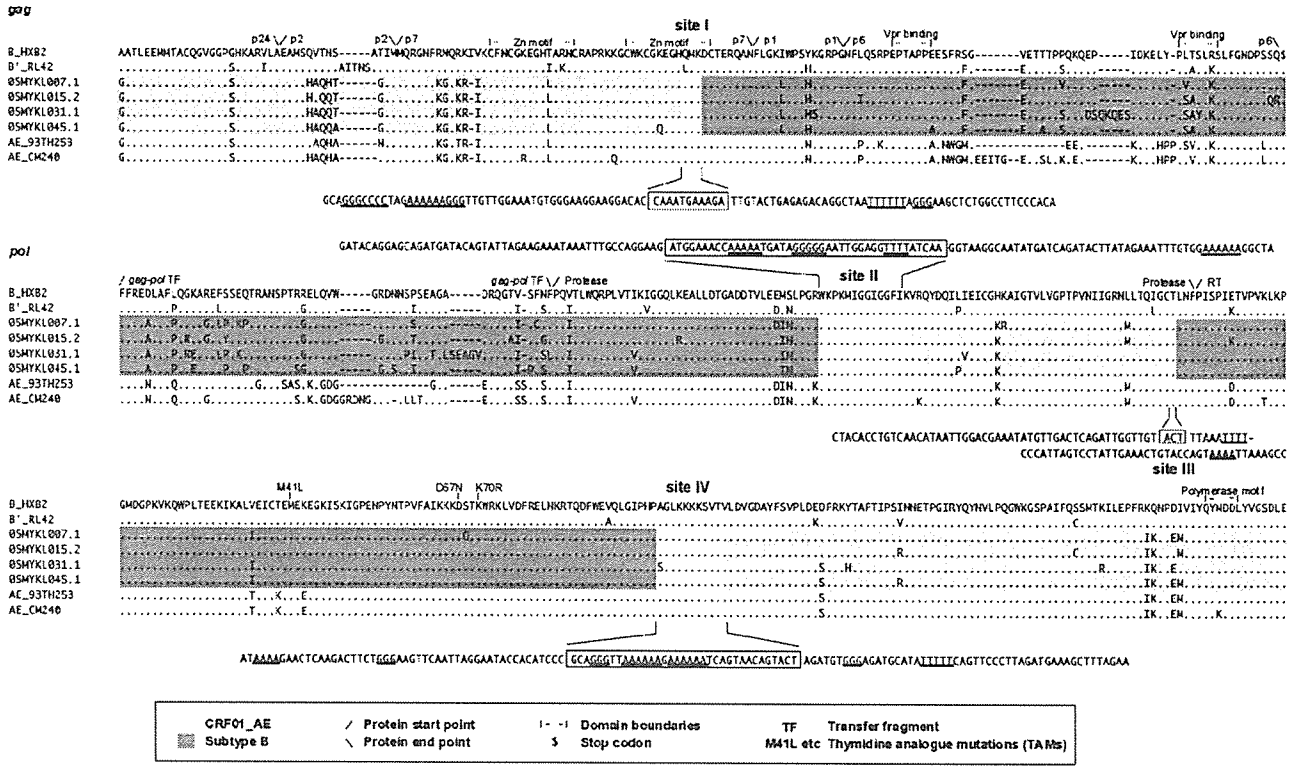

FIGURE 4. Sequence signatures in the vicinity of the recombination breakpoints in CRF33_01B. Amino acid sequence alignment of major recombinant form (RF) (CRF33_01B, n = 4) with subtype B (HXB2), B′ (RL42), and CRF01_AE (93TH253 and CM240) reference strains in *gag-pol* regions is shown. Four recombination breakpoints are clustered in this region. Periods (.) indicate the sequence identity with HXB2, and dashes (–) denote the gaps in the alignment. The CRF01_AE and subtype B regions are marked with different shadows. Unshaded areas are the presumed recombination breakpoints estimated by informative site analysis. Nucleotide sequences (consensus of four CRF33_01B isolates) within (boxed) and adjacent to (50-nucleotide region) the estimated recombination breakpoints (sites I–IV) are shown. Sequence landmarks in the Pro-RT region, including the boundaries of the domains of HIV-1 proteins, the conserved motifs, and the positions of major drug resistance mutations, are indicated at the top of the alignment.

Wide distribution of CRF33_01B involving all major ethnic and risk groups provides evidence for extensive bridging of HIV-1 transmission between different risk groups in Malaysia.

ACKNOWLEDGMENTS

We thank Kouichi Watanabe, Toshinari Onogi, Clarence Sim, and Margaret Lau for assistance; Brian Foley and Thomas Leitner for valuable advice on assignment of the new CRF; and Tim Mastro and Sai Kit Lam for critical reading of the manuscript. We also thank the patients for their participation in this study.

REFERENCES

1. Goh KL, Chua CT, Chiew IS, et al. The acquired immune deficiency syndrome: a report of the first case in Malaysia. *Med J Malaysia*. 1987;42: 58–60.
2. Joint United Nations Program on HIV/AIDS (UNAIDS). Available at: http://www.unaids.org. Accessed September 17, 2006.
3. Ministry of Health. *HIV/AIDS Report*, 2004.
4. UNAIDS. *AIDS Epidemic Update: December 2005*, 2005.
5. Malim MH, Emerman M. HIV-1 sequence variation: drift, shift, and attenuation. *Cell*. 2001;104:469–472.
6. Leitner T, Korber B, Daniels M, et al. HIV-1 subtype and circulating recombinant form (CRF) references sequences. In: Leitner T, Foley B, Hahn B, et al, eds. *HIV Sequence Compendium 2005*. Report no. LA-UR 06-0680. Los Alamos, NM: Theoretical Biology and Biophysics Group, Los Alamos National Library; 2005:41–48.
7. Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. *Science*. 2000;288:55–57.
8. Ou CY, Takebe Y, Weniger BG, et al. Independent introduction of two major HIV-1 genotypes into distinct high-risk populations in Thailand. *Lancet*. 1993;341:1171–1174.
9. Weniger BG, Takebe Y, Ou CY, et al. The molecular epidemiology of HIV in Asia. *AIDS*. 1994;8(Suppl 2):S13–S28.
10. Wasi C, Herring B, Raktham S, et al. Determination of HIV-1 subtypes in injecting drug users in Bangkok, Thailand, using peptide-binding enzyme immunoassay and heteroduplex mobility assay: evidence of increasing infection with HIV-1 subtype E. *AIDS*. 1995;9:843–849.
11. Vanichseni S, Kitayaporn D, Mastro TD, et al. Continued high HIV-1 incidence in a vaccine trial preparatory cohort of injection drug users in Bangkok, Thailand. *AIDS*. 2001;15:397–405.
12. Hudgens MG, Longini IM Jr, Vanichseni S, et al. Subtype-specific transmission probabilities for human immunodeficiency virus type 1 among injecting drug users in Bangkok, Thailand. *Am J Epidemiol*. 2002; 155:159–168.

13. Tovanabutra S, Polonis V, De Souza M, et al. First CRF01_AE/B recombinant of HIV-1 is found in Thailand. *AIDS*. 2001;15:1063–1065.
14. Viputtijul K, de Souza M, Trichavaroj R, et al. Heterosexually acquired CRF01_AE/B recombinant HIV type 1 found in Thailand. *AIDS Res Hum Retroviruses*. 2002;18:1235–1237.
15. Nguyen L, Hu DJ, Choopanya K, et al. Genetic analysis of incident HIV-1 strains among injection drug users in Bangkok: evidence for multiple transmission clusters during a period of high incidence. *J Acquir Immune Defic Syndr*. 2002;30:248–256.
16. Ramos A, Nguyen L, Hu DJ, et al. New HIV type 1 CRF01_AE/B recombinants displaying unique distribution of breakpoints from incident infections among injecting drug users in Thailand. *AIDS Res Hum Retroviruses*. 2003;19:667–674.
17. Swanson P, Devare SG, Hackett J Jr. Full-length sequence analysis of HIV-1 isolate CM237: a CRF01_AE/B intersubtype recombinant from Thailand. *AIDS Res Hum Retroviruses*. 2003;19:707–712.
18. Tovanabutra S, Beyrer C, Sakkhachornphop S, et al. The changing molecular epidemiology of HIV type 1 among northern Thai drug users, 1999 to 2002. *AIDS Res Hum Retroviruses*. 2004;20:465–475.
19. Tovanabutra S, Watanaveeradej V, Viputtikul K, et al. A new circulating recombinant form, CRF15_01B, reinforces the linkage between IDU and heterosexual epidemics in Thailand. *AIDS Res Hum Retroviruses*. 2003;19:561–567.
20. Brown TM, Robbins KE, Sinniah M, et al. HIV type 1 subtypes in Malaysia include B, C, and E. *AIDS Res Hum Retroviruses*. 1996;12:1655–1657.
21. Beyrer C, Vancott TC, Peng NK, et al. HIV type 1 subtypes in Malaysia, determined with serologic assays: 1992–1996. *AIDS Res Hum Retroviruses*. 1998;14:1687–1691.
22. Saraswathy TS, Ng KP, Sinniah M. Human immunodeficiency virus type 1 subtypes among Malaysian intravenous drug users. *Southeast Asian J Trop Med Public Health*. 2000;31:283–286.
23. Tee KK, Pon CK, Kamarulzaman A, et al. Emergence of HIV-1 CRF01_AE/B unique recombinant forms in Kuala Lumpur, Malaysia. *AIDS*. 2005;19:119–126.
24. Tee KK, Saw TL, Pon CK, et al. The evolving molecular epidemiology of HIV type 1 among injecting drug users (IDUs) in Malaysia. *AIDS Res Hum Retroviruses*. 2005;21:1046–1050.
25. Kato K, Sato H, Takebe Y. Role of naturally occurring basic amino acid substitutions in the human immunodeficiency virus type 1 subtype E envelope V3 loop on viral coreceptor usage and cell tropism. *J Virol*. 1999;73:5520–5526.
26. Salminen MO, Koch C, Sanders-Buell E, et al. Recovery of virtually full-length HIV-1 provirus of diverse subtypes from primary virus cultures using the polymerase chain reaction. *Virology*. 1995;213:80–86.
27. Yang R, Kusagawa S, Zhang C, et al. Identification and characterization of a new class of human immunodeficiency virus type 1 recombinants comprised of two circulating recombinant forms, CRF07_BC and CRF08_BC, in China. *J Virol*. 2003;77:685–695.
28. Rambaut A. *Se-Al (Sequence Alignment Editor)*, version 1.0, alpha 1. Oxford, England: University of Oxford, Department of Zoology; 1996.
29. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4:406–425.
30. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol*. 1980;16:111–120.
31. Lole KS, Bollinger RC, Paranjape RS, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol*. 1999;73:152–160.
32. Klarmann GJ, Schauber CA, Preston BD. Template-directed pausing of DNA synthesis by HIV-1 reverse transcriptase during polymerization of HIV-1 sequences in vitro. *J Biol Chem*. 1993;268:9793–9802.
33. Quinones-Mateu ME, Gao Y, Ball SC, et al. In vitro intersubtype recombinants of human immunodeficiency virus type 1: comparison to recent and circulating in vivo recombinant forms. *J Virol*. 2002;76:9600–9613.

Hindawi Publishing Corporation
Journal of Biomedicine and Biotechnology
Volume 2006, Article ID 65052, Pages 1–8
DOI 10.1155/JBB/2006/65052

Review Article

Essential Notes Regarding the Design of Functional siRNAs for Efficient Mammalian RNAi

Kumiko Ui-Tei,[1] Yuki Naito,[2] and Kaoru Saigo[2]

[1] *Department of Biophysics and Biochemistry, Graduate School of Science and Undergraduate Program for Bioinformatics and Systems Biology, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[2] *Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

Received 25 January 2006; Accepted 17 April 2006

Short interfering RNAs (siRNAs) are widely used to bring about RNA interference (RNAi) in mammalian cells. Numerous siRNAs may be designed for any target gene though most of which would be incapable of efficiently inducing mammalian RNAi. Certain highly functional siRNAs designed for knockout of a particular gene may render unrelated endogenous genes nonfunctional. These major bottlenecks should be properly eliminated when RNAi technologies are employed for any experiment in mammalian functional genomics. This paper thus presents essential notes and findings regarding the proper choice of siRNA-sequence selection algorithms and web-based online software systems.

INTRODUCTION

RNA interference (RNAi) is the process of nucleotide-sequence-specific post-transcriptional gene silencing [1–5]. In the case of lower eukaryotes such as *Drosophila* and *Caenorhabditis elegans*, long dsRNA may be used as an RNAi inducer [6–15], while, possibly owing to interferon response [16–20], short interfering RNA (siRNA), a Dicer digestion product of long dsRNA, is widely used for knocking down mammalian genes through RNAi [21–23]. Interferon response can be brought about even by siRNA transfection [24–28] and may be permitted in most cultured cell experiments, in which siRNA concentration is equal to or less than 100 nM [29]. In therapeutic application, low siRNA-dependent interferon response would be a matter of concern [17, 18].

Theoretically, (n-20) siRNAs targeting for a gene n bp in length can be designed. In *Drosophila*, more than 90% of these siRNAs are capable of reducing target gene activity by more than 80% [29]. The design of siRNAs in the case of *Drosophila* as well as other lower eukaryotes would thus not involve any real difficulty. But about 80% of theoretically designable siRNAs would not be highly functional in the case of mammalian RNAi [29, 30]. With certain target genes rich in GC, nonfunctional siRNAs may increase by 95% or more of the total designable siRNAs [Y N et al, unpublished].

Mismatched siRNA may occasionally inactivate genes other than the target, an undesired side effect designated as the "off-target effect" [31, 32]. The molecular basis for this remains to be clarified [33] though mRNA cleavage, the climax of the RNAi reaction [34–38], requires a nearly strict nucleotide sequence identity between the mRNA target portion and sense strand (SS) of siRNA [33, 39]. Thus, at least some fraction of undesirable siRNAs, giving rise to the off-target effect through destabilization of mRNAs other than the target, may be eliminated by computer-based homology search [40–45].

In the design of highly functional siRNAs for mammalian RNAi, suitable sequence conditions or good algorithms for selection of highly functional siRNAs and good computer software suitable for genome-wide short-sequence homology search to minimize the off-target effect are indispensable.

Too many websites are available for functional siRNA search for mammalian RNAi as partly listed in Table 1. These websites may incorporate one or a few algorithms for functional siRNA selection previously determined based on biological validation data. Considerable mammalian RNAi data are presently available so that, in some websites, original algorithms may have been replaced with those modified to be more effective yet do not appear in scientific journals, thus making difficult the evaluation of individual website

Table 1: siRNA search websites.

Website	URL	Reference or company
BLOCK-iT RNAi Designer	https://rnaidesigner.invitrogen.com/	Invitrogen
DEQOR	http://cluster-1.mpi-cbg.de/Deqor/deqor.html	[46]
Gene specific siRNA selector	http://bioinfo.wistar.upenn.edu/siRNA/siRNA.htm	[47]
OptiRNAi	http://bioit.dbi.udel.edu/rnai/	[48]
RNAi Central	http://katahdin.cshl.org:9331/RNAi_web/	Hannon Lab
RNAi Design	http://www.idtdna.com/Scitools/SciTools.aspx	Integrated DNA Technologies
Sfold	http://sfold.wadsworth.org/	[49]
SiDE	http://side.bioinfo.ochoa.fib.es/	[50]
siDESIGN Center	http://www.dharmacon.com/sidesign/	Dharmacon Research, Inc
siDirect	http://design.RNAi.jp/	[40]
siRNA Design Software	http://www.cs.hku.hk/ sirna/	[51]
siRNA Design Tool	http://www1.qiagen.com/Products/GeneSilencing/CustomSiRna/SiRnaDesigner.aspx	Qiagen
siRNA Selection Server	http://jura.wi.mit.edu/bioc/siRNA/	[52]
siRNA Sequence Selector	http://bioinfo.clontech.com/rnaidesigner/	Clontech
siRNA Target Designer	http://www.promega.com/siRNADesigner/	Promega
siRNA Target Finder	http://www.genscript.com/rnai.html	[53]
siRNA Target Finder	http://www.ambion.com/techlib/misc/siRNA_finder.html	Ambion
siRNA Wizard	http://www.sirnawizard.com/	InvivoGen
siSearch	http://sisearch.cgb.ki.se/	[42]
TROD	http://www.cellbio.unige.ch/RNAi.html	[54]

reliability. Consequently, the present study directs attention to basic frameworks and some related application problems of algorithms for the selection of highly functional siRNAs.

RNAi-INDUCING ACTIVITY AS AN INTRINSIC PROPERTY OF THE siRNA SEQUENCE

RNAi activity induced in mammalian cells is highly dependent on the particular sequence of siRNA used [29, 30] and may vary depending on transfected cell types or transfection efficiency. To examine these factors, various siRNAs targeting for the firefly luciferase gene (*luc*) were synthesized and transfected with *luc* encoding plasmid DNA into a variety of mammalian cell lines, which include human HeLa, HEK293, and colo205, Chinese hamster CHO-K1, and mouse E14TG2A ES cells [55]. The concentration of siRNA used in these experiments was 5–50 nM. siRNA-dependent RNAi activity was also examined in chicken embryos [29]. The transfection efficiency of colo205 is quite low and about 1/100 times as high as that of HeLa [55]. Neither difference in animal species from which cell lines or embryos were derived nor that in transfection efficiency had any significant effect on induced RNAi activity [29, 55]. RNAi activity induced in mammalian and chicken cells upon siRNA transfection may thus be determined primarily by the transfected siRNA sequences themselves as far as RNAi due to 10–50 nM siRNA is concerned.

THREE BASIC ALGORITHMS FOR SELECTING FUNCTIONAL siRNAs BASED ON BIOLOGICAL VALIDATION

Many experiments have been conducted to clarify possible sequence requirements of functional siRNAs for mammalian RNAi [29, 56–61]. Only three representative algorithms, which may be widely used for functional siRNA search for mammalian RNAi, are presented and discussed in the following.

Algorithm 1. This algorithm was developed by Ui-Tei et al [29]. As shown in Figure 1(a1), all siRNAs satisfying the following four sequence conditions are defined as class I siRNAs in Algorithm 1: (1) the 5′ antisense-strand (AS) end, A or U, (2) the 5′ SS end, G or C, (3) the 5′-terminal one-third of AS, A/U-rich, and (4) a long G/C stretch, absent from the 5′-terminal two thirds of SS. Validation data obtained using *luc* as a target indicated all of 40 class I siRNAs arbitrarily chosen to be capable of reducing target gene activity by more than 70% [29]. All RNAi experiments were conducted at 50 nM siRNA.

Algorithm 1 siRNAs with features completely the opposite to those of class I siRNAs except for condition (4) are defined as class III siRNAs (Figure 1(a2)). Validation indicated that all of 15 class III siRNAs arbitrarily chosen are incapable of inducing efficient mammalian RNAi [29]. Thus, most, if not all, class I siRNAs may possibly serve as siRNAs highly functional in mammalian cells. Class III siRNAs

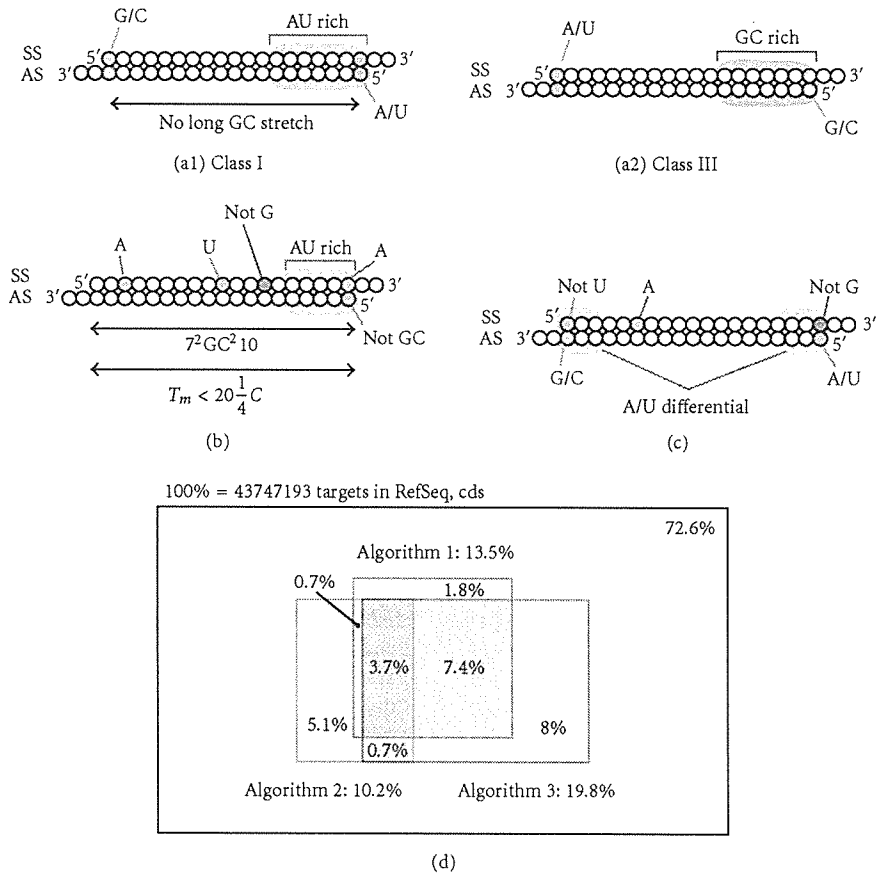

FIGURE 1: Three algorithms for siRNA design for functional RNAi in mammalian cells. (a) Algorithm 1. Highly functional class I siRNAs simultaneously satisfy the following four conditions: A/U at the 5′ AS end, G/C at the 5′ SS end, more than four A/U nucleotides in the 5′-terminal one-third of AS, and lacking a long G/C stretch in the 5′-terminal two-thirds of SS. Ineffective class III siRNAs possess features opposite to class I siRNAs. (b) Algorithm 2. There are 8 requirements for this algorithm: low G/C contents (30–52%), three or more A/U at the five 3′-terminal base pairs of SS, low internal stability lacking stable inverted repeats, and base preferences at SS positions 3, 10, 13, and 19. (c) Algorithm 3. A/U content in the 5′ AS end should be higher than that in the 5′ SS end. Base preferences are also required at positions indicated. (d) Difference in functional siRNA prediction between three Algorithms, 1, 2, and 3. 43747193 siRNA sequences were collected from human RefSeq sequences and classified using three algorithms.

are nearly incapable of inducing effective mammalian RNAi. With the *luc*, the total number of theoretically designable siRNAs is 1631 and class I siRNAs represent about 17%, which is roughly identical to the percentage (25%) of highly functional siRNAs estimated from validation data [29], class I siRNAs may thus constitute most, if not all, of siRNAs highly functional in mammalian RNAi.

Algorithm 2. This algorithm was proposed by Reynolds et al [59, Figure 1B] who carried out analysis of 180 siRNAs targeting mRNA of two genes and found the following characteristics associated with siRNA functionality: low G/C content, preference for low internal stability at the 3′-terminus of SS, and absence of inverted repeats. Furthermore, SS is presumed to preferably use A, U, and A at SS positions 3, 10, and 19, respectively. The 5′ AS terminal should not be G/C. G may not be present at position 13 (Figure 1(b)). In more than half of class I siRNAs, there are no base preferences at position 3 and 10 [29, 55], so that Algorithms 1 and 2, respectively, may predict considerably different siRNA sets to be functional.

Algorithm 3. This algorithm was proposed by Amarzguioui and Prydz [60] who carried out statistical analysis on 46 siRNAs and found Algorithm 3 to require the following features for functional siRNAs. The 5′ AS terminus and its SS partner are A/U and the 5′ SS terminus and its AS partner, G/C. An opposite combination of terminal bases may give rise to inadequate functionality. These authors also found that there is asymmetry in siRNA duplex end stability; that is, the A/U content differential for the three terminal nucleotides at both ends of the duplex may be considered essential to siRNA functionality. Furthermore, they noted A to prefer position 6 of functional siRNAs (Figure 1(c)), although only a small

fraction of class I siRNAs is associated with A at SS position 6 [29].

To examine in greater detail, relationships among the three algorithms, that the percentage of siRNAs considered functional by Algorithm 1 (class I) can be repredicted as functional by Algorithms 2 or 3 or vice versa, was determined (see [55, Figure 1D]. Based on the three algorithms, total possible siRNA sequences (4.4×10^7) designed using RefSeq human sequences (version 11) were found to be nonfunctional by as much as 73%. Class I siRNAs constituted 14% of the total theoretically predictable siRNAs, whereas Algorithms 2 and 3, respectively, predict 10 and 20% as functional siRNAs. Nearly 90% of class I siRNAs could be repredicted as functional by Algorithm 2 or 3 or both. Eighty four percent of siRNAs simultaneously predicted as functional by Algorithms 2 and 3 could be repredicted as functional or class I siRNAs by Algorithm 1. More than 50% of siRNAs predicted as functional by Algorithm 2 could not be predicted to be functional by Algorithm 3. Seventy seven percent of Algorithm 3 functional siRNAs could not be repredicted as functional by Algorithm 2. These findings may indicate that Algorithm 1 is capable of predicting the functionality of siRNAs more reliably than Algorithms 2 or 3.

ALIGNMENT ALGORITHM FOR SHORT NUCLEOTIDE SEQUENCES

Rapid homology comparison of the entire mRNA sequences with siRNA AS/SS sequences is indispensable for identifying off-target genes. BLAST [62] may not be a good software for making such comparison, since a number of off-target candidates are overlooked and too, considerable time is required for BLAST-based calculation. The Smith-Waterman local alignment algorithm [63] is accurate but time consuming to execute. Recently, Yamada and Morishita have developed a very rapid and accurate alignment algorithm for short nucleotide sequences [41] and this software can process 60 million siRNA sequences of 21 nucleotides in length in 10 hours when executed in parallel on ten inexpensive PCs. The hardware of Snøve Jr and Holen [64] provides similar performance although the number of processing units is not clearly specified. Websites using the Yamada-Morishita software or hardware of Snøve Jr and Holen should thus prove much more rapid and reliable compared to BLAST.

The base mismatch introduction studies indicate that transfected siRNAs occasionally cause phosphodiester-bond cleavage not only of the authentic mRNA target but also mutated targets with 1-2 base mismatches [33, 39]. But mutated targets with three or more mismatches may not undergo cleavage by transfection of the same siRNA [Y N et al, unpublished]. siRNAs less than 84 (16/19 × 100)% homology in sequence to any part of total mRNAs other than the target should thus be used for RNAi, which would reduce the number of available functional siRNAs to 1/10 of the input. That is, only 10% of class I siRNAs or less than 2% of total siRNAs theoretically designable using human RefSeq sequences becomes available in mammalian RNAi when off-target effects due to mRNA destability are considered. Computational analysis indicated that, even so few available siRNAs, at least one functional class I siRNA can be assigned to more than 99% of human mRNA sequences (RefSeq sequences) [Y N et al, unpublished].

miRNAs involved in posttranscriptional gene silencing through translational regulation [65–73] possess less homology with the target, indicating siRNAs with lesser homology in some cases to possibly be involved in some off-target reactions [74]. The elimination of a large number of siRNA with low homology to mRNAs other than the target may render genome-wide gene silencing in mammalian cells quite difficult. The simultaneous use of a few to several siRNAs targeting for an identical gene (target gene) may possibly solve this problem since, in most cases, off-target targets would not be identical to each other [31, 32].

EXPERIMENTAL PARAMETERS POSSIBLY AFFECTING FUNCTIONALITY OF siRNAs

siRNA-mediated RNAi activity may vary significantly depending on not only the particular siRNA sequence but also parameters such as siRNA concentration, duration of siRNA exposure, and possibly target mRNA concentration and secondary structure within cells [29, 75]. Functional siRNAs in some cases have actually been found to induce maximum RNAi activity 1 day after transfection, whereas other siRNAs to express maximum activity on 2 or 3 days following transfection. Usually, functional-siRNA-dependent RNAi persists 1-2 weeks, whereas virtually no RNAi is induced within cells even after a long incubation with nonfunctional siRNAs. Class I siRNAs, capable of inducing highly functional RNAi when transfected at 50 nM, were considerably heterogeneous in capability of bringing about RNAi when used for 1-day transfection at the concentration of 50 pM (see [29] by Ui-Tei et al). Reduction in target gene activity varied from 20 to 60% depending on the sequences of class I siRNAs used. Thus, additional sequence conditions may possibly be found so as to define a subclass of class I with more functionality but in such a case, nearly complete genome-wide gene silencing might no longer be possible.

Recently, Kim et al [76] showed that a 27 bp long dsRNA with blunt ends is much greater in functionality than 21 bp long siRNA and suggested that short Dicer substrate dsRNA may be generally much more functional compared to authentic siRNAs 21 bp long. However, it was subsequently found that this is not a general feature of 27 bp long blunt-ended dsRNA [77]. In the absence of 3′ overhang, Dicer digests dsRNA uncontrollably, generating many products varying in length, most of which may not be as functional as 21 bp long highly functional siRNAs [77]. RNAi-inducing activity would thus appear to depend primarily on the presence of considerable highly functional siRNAs in the digestion products and so, consequently, 27 bp long blunt-end dsRNA would not be necessarily a good choice for highly efficient RNAi.

siRNA-OLIGOMER-DEPENDENT RNAi IN MAMMALIAN CELLS

Long dsRNA possessing 2-nucleotide 3′ overhangs at both ends is cleaved by Dicer from these ends to generate siRNAs having definite nucleotide sequences [28, 77–80]. Thus, should nearly all siRNAs produced by Dicer digestion belong to class I and the interferon response due to dsRNAs equivalent in length to siRNA oligomers not being significant, the induction of effective multiple-gene knockout in mammalian cells may occur with transfection of siRNA oligomers and this was recently found to be the case [28]. Through use of class-I-siRNA oligomers multiple-target gene knockout was clearly shown to take place.

DNA/shRNA-MEDIATED RNAi

RNAi can be induced by introducing DNA encoding both SS and AS of siRNA into mammalian cells. Both RNA polymerase III and II promoters, respectively, are used to express short hairpin RNA (shRNA) and longer RNA including shRNA sequence in the middle [81–90]. The primary transcript of RNA polymerase III is a mixture of shRNAs with two to several consecutive U's at its 3′ overhang [81–88]. Dicer cleavage sites of shRNAs vary depending on the length of 3′ overhangs [89] and accordingly, several different species of siRNAs are expected to be generated from shRNAs transcribed by polymerase III [88]. Thus, the presence of highly functional siRNAs in these Dicer digestion products is required for successful RNAi due to a polymerase-III-based system. In addition, four consecutive U's or A's should not be included in the nonoverhang sequences of AS and SS, respectively, since these sequences stimulate premature termination of polymerase-III-dependent transcription [88].

In polymerase II-driven expression systems, the primary transcript is long polyadenylated RNA (pri-miRNA-like RNA), which is recognized and cleaved by the nuclear microprocessor complex [91, 92]. This complex contains Drosha, an RNase III-type RNase that cleaves the pri-miRNA-like RNA to generate shRNA with a 2-nucleotide 3′ overhang [93]. The shRNA thus produced is converted mainly to two overlapping siRNAs through Dicer digestion (see [28]), indicating that successful RNAi requires the involvement of highly functional siRNAs in these siRNA products.

POSSIBLE MOLECULAR BASES OF ASYMMETRIC SEQUENCE REQUIREMENTS IN FUNCTIONAL siRNAs

Each mammalian Argonaute proteins (eIF2Cs) is comprised of a PRP motif and two domains: PAZ and PIWI [94]. Structural analysis of the Argonaute protein crystals from *Pyrococcus furious* indicated that the PIWI domain has essentially the same three-dimensional structure as ribonuclease *H* and that Argonaute may function as a slicer of mRNA [95]. PAZ and PIWI domains may recognize separately two ends of siRNA. The crystal structure of the PAZ domain from human Argonaute 1 suggested that the PAZ domain is anchored to the 2-nucleotide 3′ overhang of the siRNA duplex [96]. The PIWI domain from *Archaeoglobus fulgidus* contains a highly conserved metal-binding site that may recognize the 5′ nucleotide of AS of siRNA in a manner not dependent on sequence [97].

Algorithms 1 and 3 predict functional siRNAs to possess A/U and G/C at the 5′ AS and SS ends, respectively [29, 55, 60]. The GC pair is thermodynamically much more stable than the AU pair and thus, differences in stability in terminal base pair of the siRNA duplex may determine terminal sequence preference in highly functional and nonfunctional siRNAs, most probably by stimulating asymmetric binding of PIWI and PAZ domains to siRNA ends.

The 5′-terminal one-third of AS of functional class I siRNAs is A/U-rich, possibly due to preferable siRNA unwinding from its AS end [29, 56]. A one-step motor function of the putative siRNA helicase may unwind several base pairs from the A/U-rich siRNA end to stimulate formation of active RISC lacking SS of siRNA. Should this be the case, the introduction of base mismatches into the 3′-terminal third SS of siRNA may significantly increase the induced RNAi activity. Studies with *Drosophila* extracts showed a significant base-mismatch-dependent increase in RISC formation [56]. But, to date there are no data clearly confirm this in mammalian cultured cell experiments. Recently a part of RISC has been shown to be activated through cleavage of SS of siRNA at its center [98]. The presence of base mismatches in SS might be unfavorable to SS cleavage and this negative effect might partially prevent siRNA from being unwound.

ACKNOWLEDGMENTS

This work was partially supported by Special Coordination Fund for Promoting Science and Technology to K S, and Grants from the Ministries of Education, Culture, Sports, Science and Technology of Japan to K S and K U-T.

REFERENCES

[1] Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *caenorhabditis elegans*. *Nature*. 1998;391 (6669):806–811.
[2] Dykxhoorn DM, Novina CD, Sharp PA. Killing the messenger: short RNAs that silence gene expression. *Nature Reviews. Molecular Cell Biology*. 2003;4(6):457–467.
[3] Mello CC, Conte D Jr. Revealing the world of RNA interference. *Nature*. 2004;431(7006):338–342.
[4] Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. *Nature*. 2004;431(7006):343–349.
[5] Dykxhoorn DM, Palliser D, Lieberman J. The silent treatment: siRNAs as small molecule drugs. *Gene Therapy*. 2006;13(6): 541–552.
[6] Ui-Tei K, Zenno S, Miyata Y, Saigo K. Sensitive assay of RNA interference in *Drosophila* and Chinese hamster cultured cells using firefly luciferase gene as target. *FEBS Letters*. 2000;479 (3):79–82.
[7] Barstead R. Genome-wide RNAi. *Current Opinion in Chemical Biology*. 2001;5(1):63–66.

[8] Maeda I, Kohara Y, Yamamoto M, Sugimoto A. Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. *Current Biology.* 2001;11(3):171–176.

[9] Ueda R. RNAi: a new technology in the post-genomic sequencing era. *Journal of Neurogenetics.* 2001;15(3-4):193–204.

[10] Ashrafi K, Chang FY, Watts JL, et al. Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature.* 2003;421(6920):268–272.

[11] Boutros M, Kiger AA, Armknecht S, et al. Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science.* 2004;303(5659):832–835.

[12] Carpenter AE, Sabatini DM. Systematic genome-wide screens of gene function. *Nature Reviews. Genetics.* 2004;5(1):11–22.

[13] Cheng LW, Viala JPM, Stuurman N, Wiedemann U, Vale RD, Portnoy DA. Use of RNA interference in *Drosophila* S2 cells to identify host pathways controlling compartmentalization of an intracellular pathogen. *Proceedings of the National Academy of Sciences of the United States of America.* 2005;102 (38):13646–13651.

[14] Nybakken K, Vokes SA, Lin T-Y, McMahon AP, Perrimon N. A genome-wide RNA interference screen in *Drosophila melanogaster* cells for new components of the Hh signaling pathway. *Nature Genetics.* 2005;37(12):1323–1332.

[15] Kim JK, Gabel HW, Kamath RS, et al. Functional genomic analysis of RNA interference in *C. elegans*. *Science.* 2005;308 (5725):1164–1167.

[16] Stark GR, Kerr IM, Williams BRG, Silverman RH, Schreiber RD. How cells respond to interferons. *Annual Review of Biochemistry.* 1998;67:227–264.

[17] Zhou H-S, Liu D-P, Liang C-C. Challenges and strategies: the immune responses in gene therapy. *Medicinal Research Reviews.* 2004;24(6):748–761.

[18] Karpala AJ, Doran TJ, Bean AGD. Immune responses to dsRNA: implications for gene silencing technologies. *Immunology and Cell Biology.* 2005;83(3):211–216.

[19] de Veer MJ, Sledz CA, Williams BRG. Detection of foreign RNA: implications for RNAi. *Immunology and Cell Biology.* 2005;83(3):224–228.

[20] Marques JT, Williams BRG. Activation of the mammalian immune system by siRNAs. *Nature Biotechnology.* 2005;23 (11):1399–1405.

[21] Elbashir SM, Lendeckel W, Tuschl T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes and Development.* 2001;15(2):188–200.

[22] Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature.* 2001;411 (6836):494–498.

[23] Elbashir SM, Martinez J, Patkaniowska A, Lendeckel W, Tuschl T. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO Journal.* 2001;20(23):6877–6888.

[24] Sledz CA, Holko M, de Veer MJ, Silverman RH, Williams BRG. Activation of the interferon system by short-interfering RNAs. *Nature Cell Biology.* 2003;5(9):834–839.

[25] Sioud M. Induction of inflammatory cytokines and interferon responses by double-stranded and single-stranded siRNAs is sequence-dependent and requires endosomal localization. *Journal of Molecular Biology.* 2005;348(5):1079–1090.

[26] Hornung V, Guenthner-Biller M, Bourquin C, et al. Sequence-specific potent induction of IFN-α by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nature Medicine.* 2005;11(3):263–270.

[27] Judge AD, Sood V, Shaw JR, Fang D, McClintock K, MacLachlan I. Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. *Nature Biotechnology.* 2005;23(4):457–462.

[28] Ui-Tei K, Zenno S, Juni A, Saigo K. RNAi induced in mammalian and *Drosophila* cells via transfection of dimers and trimers of short interfering RNA. *Journal of RNAi and Gene Silencing.* 2005;1(2):79–87.

[29] Ui-Tei K, Naito Y, Takahashi F, et al. Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Research.* 2004;32(3):936–948.

[30] Holen T, Amarzguioui M, Wiiger MT, Babaie E, Prydz H. Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. *Nucleic Acids Research.* 2002;30(8):1757–1766.

[31] Jackson AL, Bartz SR, Schelter J, et al. Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnology.* 2003;21(6):635–637.

[32] Jackson AL, Linsley PS. Noise amidst the silence: off-target effects of siRNAs? *Trends in Genetics.* 2004;20(11):521–524.

[33] Du Q, Thonberg H, Wang J, Wahlestedt C, Liang Z. A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites. *Nucleic Acids Research.* 2005;33(5):1671–1677.

[34] Martinez J, Patkaniowska A, Urlaub H, Lührmann R, Tuschl T. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell.* 2002;110(5):563–574.

[35] Song J-J, Smith SK, Hannon GJ, Joshua-Tor L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science.* 2004;305(5689):1434–1437.

[36] Liu J, Carmell MA, Rivas FV, et al. Argonaute2 is the catalytic engine of mammalian RNAi. *Science.* 2004;305(5689):1437–1441.

[37] Parker JS, Roe SM, Barford D. Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature.* 2005;434(7033):663–666.

[38] Miyoshi K, Tsukumo H, Nagami T, Siomi H, Siomi MC. Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes and Development.* 2005;19(23):2837–2848.

[39] Abdelgany A, Wood M, Beeson D. Allele-specific silencing of a pathogenic mutant acetylcholine receptor subunit by RNA interference. *Human Molecular Genetics.* 2003;12(20):2637–2644.

[40] Naito Y, Yamada T, Ui-Tei K, Morishita S, Saigo K. siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference. *Nucleic Acids Research.* 2004; 32:W124–W129.

[41] Yamada T, Morishita S. Accelerated off-target search algorithm for siRNA. *Bioinformatics.* 2005;21(8):1316–1324.

[42] Chalk AM, Warfinge RE, Georgii-Hemming P, Sonnhammer ELL. siRNAdb: a database of siRNA sequences. *Nucleic Acids Research.* 2005;33:D131–D134.

[43] Boese Q, Leake D, Reynolds A, et al. Mechanistic insights aid computational short interfering RNA design. *Methods in Enzymology.* 2005;392:73–96.

[44] Qiu S, Adema CM, Lane T. A computational study of off-target effects of RNA interference. *Nucleic Acids Research.* 2005;33(6):1834–1847.

[45] Naito Y, Yamada T, Matsumiya T, Ui-Tei K, Saigo K, Morishita S. dsCheck: highly sensitive off-target search software for double-stranded RNA-mediated RNA interference. *Nucleic Acids Research.* 2005;33(suppl 2):W589–W591.

[46] Henschel A, Buchholz F, Habermann B. DEQOR: a web-based tool for the design and quality control of siRNAs. *Nucleic Acids Research.* 2004;32:W113–W120.
[47] Levenkova N, Gu Q, Rux JJ. Gene specific siRNA selector. *Bioinformatics.* 2004;20(3):430–432.
[48] Cui W, Ning J, Naik UP, Duncan MK. OptiRNAi, an RNAi design tool. *Computer Methods and Programs in Biomedicine.* 2004;75(1):67–73.
[49] Ding Y, Chan CY, Lawrence CE. Sfold web server for statistical folding and rational design of nucleic acids. *Nucleic Acids Research.* 2004;32:W135–W141.
[50] Santoyo J, Vaquerizas JM, Dopazo J. Highly specific and accurate selection of siRNAs for high-throughput functional assays. *Bioinformatics.* 2005;21(8):1376–1382.
[51] Yiu SM, Wong PWH, Lam TW, et al. Filtering of ineffective siRNAs and improved siRNA design tool. *Bioinformatics.* 2005;21(2):144–151.
[52] Yuan B, Latek R, Hossbach M, Tuschl T, Lewitter F. siRNA selection server: an automated siRNA oligonucleotide prediciton server. *Nucleic Acids Research.* 2004;32:W130–W134.
[53] Wang L, Mu FY. A web-based design center for vector-based siRNA and siRNA cassette. *Bioinformatics.* 2004;20(11):1818–1820.
[54] Dudek P, Picard D. TROD: T7 RNAi oligo designer. *Nucleic Acids Research.* 2004;32:W121–W123.
[55] Ui-Tei K, Naito Y, Saigo K. Guidelines for selection of effective short-interfering RNA sequences for functional genomics. *Methods in Molecular Biology.* 2006;2:201–216.
[56] Schwarz DS, Hutvágner G, Du T, Xu Z, Aronin N, Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. *Cell.* 2003;115(2):199–208.
[57] Khvorova A, Reynolds A, Jayasena SD. Functional siRNAs and miRNAs exhibit strand bias. *Cell.* 2003;115(2):209–216.
[58] Pancoska P, Moravek Z, Moll UM. Efficient RNA interference depends on global context of the target sequence: quantitative analysis of silencing efficiency using Eulerian graph representation of siRNA. *Nucleic Acids Research.* 2004;32(4):1469–1479.
[59] Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorova A. Rational siRNA design for RNA interference. *Nature Biotechnology.* 2004;22(3):326–330.
[60] Amarzguioui M, Prydz H. An algorithm for selection of functional siRNA sequences. *Biochemical and Biophysical Research Communications.* 2004;316(4):1050–1058.
[61] Jagla B, Aulner N, Kelly PD, et al. Sequence characteristics of functional siRNAs. *RNA.* 2005;11(6):864–872.
[62] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of Molecular Biology.* 1990;215(3):403–410.
[63] Smith TF, Waterman MS. Identification of common molecular subsequences. *Journal of Molecular Biology.* 1981;147(1):195–197.
[64] Snøve O Jr, Holen T. Many commonly used siRNAs risk off-target activity. *Biochemical and Biophysical Research Communications.* 2004;319(1):256–263.
[65] Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science.* 2001;294(5543):853–858.
[66] Ambros V. microRNAs: tiny regulators with great potential. *Cell.* 2001;107(7):823–826.
[67] McManus MT, Sharp PA. Gene silencing in mammals by small interfering RNAs. *Nature Reviews Genetics.* 2002;3(10):737–747.
[68] Ambros V, Bartel B, Bartel DP, et al. A uniform system for microRNA annotation. *RNA.* 2003;9(3):277–279.
[69] He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. *Nature Reviews Genetics.* 2004;5(7):522–531.
[70] Croce CM, Calin GA. miRNAs, cancer, and stem cell division. *Cell.* 2005;122(1):6–7.
[71] Sontheimer EJ, Carthew RW. Silence from within: endogenous siRNAs and miRNAs. *Cell.* 2005;122(1):9–12.
[72] Sullivan CS, Ganem D. MicroRNAs and viral infection. *Molecular Cell.* 2005;20(1):3–7.
[73] Kosik KS, Krichevsky AM. The elegance of the microRNAs: a neuronal perspective. *Neuron.* 2005;47(6):779–782.
[74] Lin X, Ruan X, Anderson MG, et al. siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. *Nucleic Acids Research.* 2005;33(14):4527–4535.
[75] Omi K, Tokunaga K, Hohjoh H. Long-lasting RNAi activity in mammalian neurons. *FEBS Letters.* 2004;558(1–3):89–95.
[76] Kim D-H, Behlke MA, Rose SD, Chang M-S, Choi S, Rossi JJ. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nature Biotechnology.* 2004;23(2):222–226.
[77] Rose SD, Kim D-H, Amarzguioui M, et al. Functional polarity is introduced by Dicer processing of short substrate RNAs. *Nucleic Acids Research.* 2005;33(13):4140–4156.
[78] Zhang H, Kolb FA, Brondani V, Billy E, Filipowicz W. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO Journal.* 2002;21(21):5875–5885.
[79] Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. *Cell.* 2004;118(1):57–68.
[80] Vermeulen A, Behlen L, Reynolds A, et al. The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA.* 2005;11(5):674–682.
[81] Brummelkamp TR, Bernards R, Agami R. A system for stable expression of short interfering RNAs in mammalian cells. *Science.* 2002;296(5567):550–553.
[82] Sui G, Soohoo C, Affar EB, et al. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America.* 2002;99(8):5515–5520.
[83] Paddison PJ, Caudy AA, Hannon GJ. Stable suppression of gene expression by RNAi in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America.* 2002;99(3):1443–1448.
[84] Yu J-Y, DeRuiter SL, Turner DL. RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America.* 2002;99(9):6047–6052.
[85] Lee NS, Dohjima T, Bauer G, et al. Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nature Biotechnology.* 2002;20(5):500–505.
[86] Paul CP, Good PD, Winer I, Engelke DR. Effective expression of small interfering RNA in human cells. *Nature Biotechnology.* 2002;20(5):505–508.
[87] McManus MT, Petersen CP, Haines BB, Chen J, Sharp PA. Gene silencing using micro-RNA designed hairpins. *RNA.* 2002;8(6):842–850.
[88] Shi Y. Mammalian RNAi for the masses. *Trends in Genetics.* 2003;19(1):9–12.
[89] Silva JM, Li MZ, Chang K, et al. Second-generation shRNA libraries covering the mouse and human genomes. *Nature Genetics.* 2005;37(11):1281–1288.
[90] Stegmeier F, Hu G, Rickles RJ, Hannon GJ, Elledge SJ. A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells. *Proceedings*

of the National Academy of Sciences of the United States of America. 2005;102(37):13212–13217.

[91] Gregory RI, Yan K-P, Amuthan G, et al. The microprocessor complex mediates the genesis of microRNAs. *Nature.* 2004;432 (7014):235–240.

[92] Denli AM, Tops BBJ, Plasterk RHA, Ketting RF, Hannon GJ. Processing of primary microRNAs by the microprocessor complex. *Nature.* 2004;432(7014):231–235.

[93] Lee Y, Ahn C, Han J, et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature.* 2003;425(6956):415–419.

[94] Doi N, Zenno S, Ueda R, Ohki-Hamazaki H, Ui-Tei K, Saigo K. Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. *Current Biology.* 2003;13(1):41–46.

[95] Song J-J, Smith SK, Hannon GJ, Joshua-Tor L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science.* 2004;305(5689):1434–1437.

[96] Ma J-B, Ye K, Patel DJ. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature.* 2004;429(6989):318–322.

[97] Ma J-B, Yuan Y-R, Meister G, Pei Y, Tuschl T, Patel DJ. Structural basis for 5′-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature.* 2005;434(7033):666–670.

[98] Leuschner PJ, Ameres SL, Kueng S, Martinez J. Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Reports.* 2006;7(3):314–320.

Separate elements are required for ligand-dependent and -independent internalization of metastatic potentiator CXCR4

Yuko Futahashi,[1] Jun Komano,[1,4] Emiko Urano,[1] Toru Aoki,[1,2] Makiko Hamatake,[1] Kosuke Miyauchi,[1] Takeshi Yoshida,[3] Yoshio Koyanagi,[3] Zene Matsuda[1] and Naoki Yamamoto[1,2]

[1]AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku, Tokyo 162-8640; [2]Department of Molecular Virology, Tokyo Medical and Dental University, 1-5-45, Yushima, Bunkyo-ku, Tokyo 113-8519; [3]Laboratory of Viral Pathogenesis, Institute for Virus Research, Kyoto University, 53 Shougoin-kawahara machi, Sakyou-ku, Kyoto 606-8507, Japan

(Received September 17, 2006/Revised November 3, 2006/Accepted November 11, 2006/Online publication January 19, 2007)

The C-terminal cytoplasmic domain of the metastatic potentiator CXCR4 regulates its function and spatiotemporal expression. However, little is known about the mechanism underlying constitutive internalization of CXCR4 compared to internalization mediated by its ligand, stromal cell-derived factor-1 alpha (SDF-1α)/CXCL12. We established a system to analyze the role of the CXCR4 cytoplasmic tail in steady-state internalization using the NP2 cell line, which lacks endogenous CXCR4 and SDF-1α. Deleting more than six amino acids from the C-terminus dramatically reduced constitutive internalization of CXCR4. Alanine substitution mutations revealed that three of those amino acids Ser^{344} Glu^{345} Ser^{346} are essential for efficient steady-state internalization of CXCR4. Mutating Glu^{345} to Asp did not disrupt internalization, suggesting that the steady-state internalization motif is S(E/D)S. When responses to SDF-1α were tested, cells expressing CXCR4 mutants lacking the C-terminal 10, 14, 22, 31 or 44 amino acids did not show downregulation of cell surface CXCR4 or the cell migration induced by SDF-1α. Interestingly, however, we identified two mutants, one with E344A mutation and the other lacking the C-terminal 17 amino acids, that were defective in constitutive internalization but competent in ligand-promoted internalization and cell migration. These data demonstrate that ligand-dependent and -independent internalization is genetically separable and that, between amino acids 336 and 342, there is a negative regulatory element for ligand-promoted internalization. Potential involvement of this novel motif in cancer metastasis and other CXCR4-associated disorders such as warts, hypogammaglobulinemia, infections and myelokathexis (WHIM) syndrome is discussed. (*Cancer Sci* 2007; 98: 373–379)

The chemokine receptor CXCR4 is a class-A G protein-coupled receptor (GPCR; reviewed in [1,2]) and its natural ligand is stromal cell-derived factor-1 alpha (SDF-1α)/CXCL12. CXCR4 also serves as the receptor for HIV type 1 (HIV-1). Many cell types express CXCR4, including peripheral blood lymphocytes, monocytes-macrophages, thymocytes, dendritic cells, endothelial cells, epithelium-derived tumor cells, microglial cells, neurons and hematopoietic stem cells. CXCR4 plays multiple biological roles from promoting development of neuronal networks to regulating migration of leukocytes, cerebellar granule cells and hematopoietic stem cells.[3–8] Analysis of knockout mice indicates that the CXCR4/SDF-1α system is essential for maintenance of hematopoiesis and intestinal vascularization.[9,10]

The CXCR4/SDF-1α system also functions in pathological processes, including autoimmune diseases, cancer progression and metastasis, and AIDS caused by HIV-1. Recently, metastasis of breast cancer cells was found to be regulated by the CXCR4/SDF-1α axis.[5] Similarly, other studies have found that metastasis of other malignancies was controlled by the CXCR4/SDF-1α system, including colon carcinoma[11] non-small cell lung cancer[12] and prostate cancer.[13] These observations suggest that the CXCR4/SDF-1α axis is a potential target for metastatic cancer therapy.

Warts, hypogammaglobulinemia, infections and myelokathexis (WHIM) syndrome is a rare combined immunodeficiency characterized by an unusual form of neutropenia. It is reported that the CXCR4 cytoplasmic tail is mutated and often truncated in WHIM syndrome.[14] Thus, determining the biochemical activity of the CXCR4 cytoplasmic tail should facilitate understanding of the pathogenesis of WHIM syndrome as well as suggest ways to control cancer metastasis.

Following SDF-1α binding, CXCR4 is activated, triggering multiple signaling cascades via Gα or β-arrestin 2 (reviewed in[15]). To desensitize activated CXCR4, the G protein-coupled receptor kinase (GRK) is recruited and phosphorylates serine residues on the CXCR4 cytoplasmic tail, thereby inactivating Gα-mediated signal. Simultaneously, CXCR4 is internalized in a clathrin-dependent manner. β-arrestin 2 competes with Gα for CXCR4 binding and can initiate signal transduction independent from Gα. β-arrestin 2 can also induce clathrin-dependent CXCR4 endocytosis. Thus, cell surface levels of CXCR4 transiently decrease after agonist binding but, several hours later, surface levels of CXCR4 return to normal. Most internalized CXCR4 is transported to lysosomes and degraded, but some internalized CXCR4 is recycled. It is reported that amino acids within the cytoplasmic tail are required for agonist-dependent endocytosis of CXCR4.[16–18]

By contrast, it is unclear how steady-state cell surface levels of CXCR4 are maintained in the absence of SDF-1α. Although cell surface levels of CXCR4 could be regulated at the transcriptional level, it is likely that primary regulation occurs post-translationally. Given that the cell surface levels of CXCR4 are positively correlated with cancer cells' ability to metastasize,[5,19] understanding the post-translational behavior of CXCR4 is likely to shed light on metastatic processes. Historically, cells expressing endogenous CXCR4 have been used for analysis of CXCR4 trafficking. However, as is the case with many G protein-coupled receptors (GPCR), CXCR4 trafficking is influenced by spontaneous oligomerization in the absence of ligand.[20–22] Thus, previous observations might not correctly model phenotypes seen in CXCR4 mutants.

In the present study, we analyzed the contribution of the cytoplasmic tail to the post-translational trafficking of CXCR4 in a cell line lacking both endogenous CXCR4 and SDF-1α. Using genetic approaches, we identified two amino acid motifs within the CXCR4 cytoplasmic tail; one that positively regulates

[4]To whom correspondence should be addressed. E-mail:ajkomano@nih.go.jp

doi: 10.1111/j.1349-7006.2007.00388.x

spontaneous ligand-independent internalization and the other that negatively regulates ligand-dependent CXCR4 internalization.

Materials and methods

Cells. The glioblastoma cell line NP2, human embryonic kidney (HEK) 293T, and HeLa, cells were maintained in RPMI-1640 (Sigma, Tokyo, Japan) supplemented with 10% FBS (Japan Bioserum, Tokyo, Japan), penicillin and streptomycin (Invitrogen, Tokyo, Japan). All cell lines were incubated at 37°C in the humidified 5% CO_2 atmosphere.

Plasmids. Full-length CXCR4 cDNA was amplified from a plasmid kindly provided by Dr Shioda[23] using the following primers: sense, 5′-ACCGGTGCCACCATGGAGGGGATCAGTATATACACTTCAG-3′, and antisense, 5′-AGATCTCGCTGGAGTGAAAACTTGAAGACTCAGACTC-3′. CXCR4 lacking the cytoplasmic tail (d-44) was amplified using the same sense primer and the antisense primer, 5′-AGATCTTGGCTCCAAGGAAAGCATAGAGGATGGG-3′. Polymerase chain reaction (PCR) fragments were cloned into the *Age* I-*Bgl* II sites of pEGFP-C2 (Clontech, Palo Alto, CA, USA) to create pCXCR4 FL and pCXCR4 d-44, respectively. To construct pCXCR4 FL- and d-44-GFP, the *Sna* BI-*Bgl* II fragments from pCXCR4 FL and d-44 were cloned into the *Sna* BI-*Bgl* II sites of pEGFP-N2, respectively (Clontech). To construct pCXCR4 FL- and d-44-GFP flag, the *Sna* BI-*Bgl* II fragments from pCXCR4 FL and d-44 were cloned into the *Sna* BI-*Bgl* II sites of pEGFP-flag in which the following annealed oligonucleotides had been inserted into the *Bsr* GI site of pEGFP-N2: forward, 5′-GTACGACTACAAAGACGATGACGACTATAAGTAAGC-3′, and reverse, 5′-GGCCGCTTACTTATAGTCGTCATCGTCTTTGTAGTC-3′. To construct pCMMP CXCR4 FL- and d-44-GFP, pCXCR4 FL- and d-44-GFP were digested with *Not* I, blunted using T4 DNA polymerase, and further digested with *Age* I. The Age *I*-blunted *Not* I fragments of both constructs were cloned into the pCMMP eGFP plasmid that had been digested with *Bam* HI, blunted with T4 DNA polymerase, and digested with *Age* I. pCMMP CXCR4 FL- and d-44-GFP-flag were constructed using the same strategy. CXCR4 deletion and point mutants were PCR-amplified using the sense primer 5′-ACCGGTGCCACCATGGAGGGGATCAGTGTGAAAACTTGAAGACTCAGACTC-3′ and the following reverse primers: d-6, 5′-AAGCTTGAGCTCGAGATCTCAGACTCAGACTCAGTGGAAAC-3′; d-10, 5′-AAGCTTGAGCTCGAGATCTCAGTGGAAACAGATGAATGTCC-3′; d-14, 5′-AAGCTTGAGCTCGAGATCTCTGAATGTCCACCTCGCTTTCC-3′; d-17, 5′-AAGCTTGAGCTCGAGATCTCACCTCGCTTTCCTTTGG-3′; d-22, 5′-AAGCTTGAGCTCGAGATCTCGGAGAGGATCTTGAGGCTGGACC-3′; d-31, 5′-AAGCTTGAGCTCGAGATCTCGCTCACAGAGGTGAGTGCGTGC-3′; E343A, 5′-CGAGATCTCGCTGGAGTGAAAACTTGAAGACTCAGACGCAGTGGAAACAGATGAATGTC-3′; S344A, 5′-CGAGATCTCGCTGGAGTGAAAACTTGAAGACTCAGCCTCAGTGGAAACAGATGAATGTC-3′; E345A, 5′-CGAGATCTCGCTGGAGTGAAAACTTGAAGACGCAGACTCAGTGGAAACAGATGAATGTC-3′; S346A, 5′-CGAGATCTCGCTGGAGTGAAAACTTGAAGCCTCAGACTCAGTGGAAACAGATGAATGTC-3′; S347E, 5′-CGAGATCTCGCTGGAGTGAAAACTTTCAGACTCAGACTCAGTGGAAACAGATGAATGTC-3′; H350E, 5′-CGAGATCTCGCTGGACTCAAAACTTGAAGACTCAGACTCAGTGGAAACAGATGAATGTC-3′; S347E/H350E, 5′-CGAGATCTCGCTGGACTCAAAACTTTCAGACTCAGACTCAGTGGAAACAGATGAATGTC-3′; and E343/345D, 5′-CGAGATCTCGCTGGAGTGAAAACTTGAAGAGTCAGAGTCAGTGGAAACAGATGAATGTC-3′. The PCR fragments were cloned into the *Age* I-*Bgl* II sites of pCMMP CXCR4 FL-GFP-flag, replacing wild-type with mutant CXCR4. Protein expression of each mutant in 293T cells was verified by Western blot analysis.

Immunoblotting. Immunoblotting was performed as described.[24,25] The primary antibody was anti-green fluorescent protein (GFP) polyclonal antibody (Beckton Dickinson, San Jose, CA, USA). The secondary probe was EnVision+ (Dako, Glostrup, Denmark). Signals were visualized with an LAS3000 imager (Fuji Film, Tokyo, Japan) after treating the membranes with the Lumi-Light Western Blotting Substrate (Roche Diagnostics GmbH, Mannheim, Germany).

Flow cytometry. Cells were labeled with anti-CXCR4 antibodies recognizing the N-terminus conjugated with R-phycoerythrin (PE; 2B11, BD Pharmingen, San Diego, CA) or recognizing the second extracellular loop (12G5) conjugated with either PE or PE-Cy5 (Beckton Dickinson) for 30 min at 4°C. Cells were washed once with phosphate-buffered saline (PBS) supplemented with 1% FBS and analyzed by FACS Aria (Beckton Dickinson). To isolate GFP-expressing NP2 cells, cells were infected with murine leukemia virus (MLV)-based retroviral vectors as described.[25] Cells exhibiting similar green fluorescence intensities were gated and sorted by FACS Aria. Efficiency of internalization was measured by comparing mean fluorescence intensities for cell surface CXCR4 detected by a PE-labeled 2B11 monoclonal antibody before and after SDF-1α treatment (200 ng/mL, Peprotech EC, London, UK).

Microscopic analysis and imaging of cells. To judge a phenotype of a CXCR4 mutant, three independent scientists investigated the mutant cell phenotype under a fluorescent microscope (Olympus, Tokyo, Japan). Each scientist investigated more than 1000 cells for each mutant. More than 99% of cells of a mutant fell in the indicated phenotypic category. These phenotypes were unchanged for more than a year of continuous cultivation in tissue culture. For imaging, NP2 cells were grown on glass plates for more than 24 h, fixed in 4% formaldehyde in PBS for 5 min, stained with Hoechst 33258, mounted (Vectorshield, Vector Laboratories, Burlingame, CA, USA), and imaged using a confocal microscope META 510 (Carl Zeiss, Tokyo, Japan). A representative cell for each CXCR4 mutant carrying a wide cytoplasm was chosen such that the spatial resolution was high. The focal plane just above the glass surface was scanned with an optical thickness of approximately 1 μm. For the imaging of subcellular compartments, cells were incubated with either BODIPY TR ceramid, ER-Tracker Blue-White DPX, or Lysotracker Red DND-99 (Invitrogen) according to the manufacturer's protocol and imaged without fixation. Image brightness and contrast were processed by META510 software (Carl Zeiss). Unless noted, cells were imaged at ×630 magnification, the GFP signal was displayed in green, and Hoechst 33258-stained nuclei were blue. To visualize ligand-induced internalization, cells were treated with 200 ng/mL SDF-1α before fixation. The live cell imaging was performed using Leica DFC350FX system and the images were processed by FW4000 software (Leica Microsystems, Tokyo, Japan). Cells were plated on the glass-bottomed dish (Matsunami glass, Kishiwada, Japan) and incubated at 37°C in the humidified 5% CO_2 atmosphere during the monitoring.

Cell migration assay. Cell migration was measured using an HTS FluoroBlok Multiwell Insert System (8.0 μm pore size, BD Falcon) according to the manufacturer's protocol. For stimulation assays, cells were incubated without serum overnight before SDF-1α treatment (200 ng/mL). Cells were allowed to migrate overnight.

Statistical analysis. Significance of differences were determined by a Student's *t*-test. *P*-values less than 0.05 were considered significant.

RESULTS

Deleting 10 amino acids from the carboxyl end of CXCR4 alters the efficiency of constitutive internalization. Previous studies indicated that the cytoplasmic tail of CXCR4 amino acids 308–352 plays

doi: 10.1111/j.1349-7006.2007.00388.x

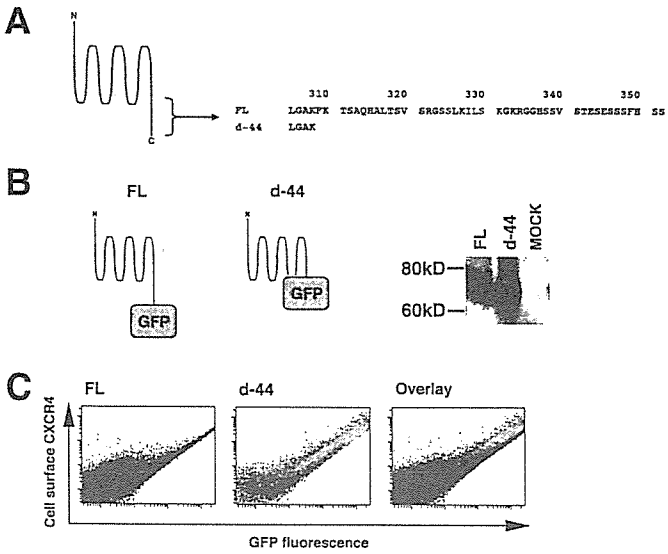

Fig. 1. The effect of stromal cell-derived factor-1 alpha (SDF-1α) treatment on NP2 cells expressing CXCR4 mutants. (a) Cells expressing d-17 were treated with SDF-1α, incubated at 37°C for the indicated times, fixed and imaged. The blue signal represents the Hoechst-stained nucleus. (Original magnification, ×630; bar, 10 μm). (b) FACS analysis to measure internalization efficiency of cell surface CXCR4 and mutant forms 2 h after SDF-1α exposure. The average and standard deviation from the indicated number of independent experiments are shown. Asterisks represent statistically significant difference from the FL levels ($P < 0.01$). (c) Cell migration assay to assess response of cells expressing CXCR4 and mutants to SDF-1α. The number of migrated cells in three to six randomly selected fields was counted and the average and standard deviation were calculated. (□) number of migrated cells in the absence of ligand; (■) migration in the presence of ligand. (*) statistically significant differences in the number of migrated cells between SDF-1α-untreated and -treated cells ($P < 0.01$).

a critical role in ligand-dependent internalization (Fig. 1a). Also, it has been shown in transfected cells that cell surface levels of CXCR4 lacking the cytoplasmic tail (equivalent to the d-44 mutant here) are higher than those of the full length, wild-type protein (hereafter designated FL), suggesting that the cytoplasmic tail of CXCR4 regulates steady-state internalization.[26,27] To confirm this, we constructed expression plasmids of CXCR4 FL and d-44 fused to GFP or GFP-FLAG at the C-terminus. Previous studies and data reported here indicated that CXCR4 function is not affected by this modification.[28] The expression of each construct was verified by Western blot analysis (Fig. 1b). Single cell-based quantitative analyzes revealed that the ratio of cell surface levels to the total amount of CXCR4 FL (Fig. 1c, left) was consistently lower than that of d-44 (Fig. 1c, middle) at any expression levels (Fig. 1c, right for the comparison). These data supported previous findings and demonstrate that constitutive internalization occurs at any level of CXCR4 expression.

To further examine the contribution of the cytoplasmic tail to post-translational trafficking of CXCR4, we devised a system utilizing the human NP2 glioma line: NP2 cells are flat and exhibit a large cytoplasmic space such that intracellular compartments can be well resolved under the microscope. NP2 cells also lack endogenous CXCR4[29] and SDF-1α (data not shown), both of which could potentially affect distribution of transduced CXCR4. However, NP2 cells are capable of appropriate signaling in response to CXCR4/SDF-1α interaction. We generated a series of CXCR4 deletion mutants lacking the cytoplasmic tail (Fig. 2a) and transduced them into NP2 cells using MLV vectors. Cells bearing similar green fluorescence intensities were collected by FACS sorter. The expression of each mutant was verified by Western blot analysis (Fig. 2b). Microscopic

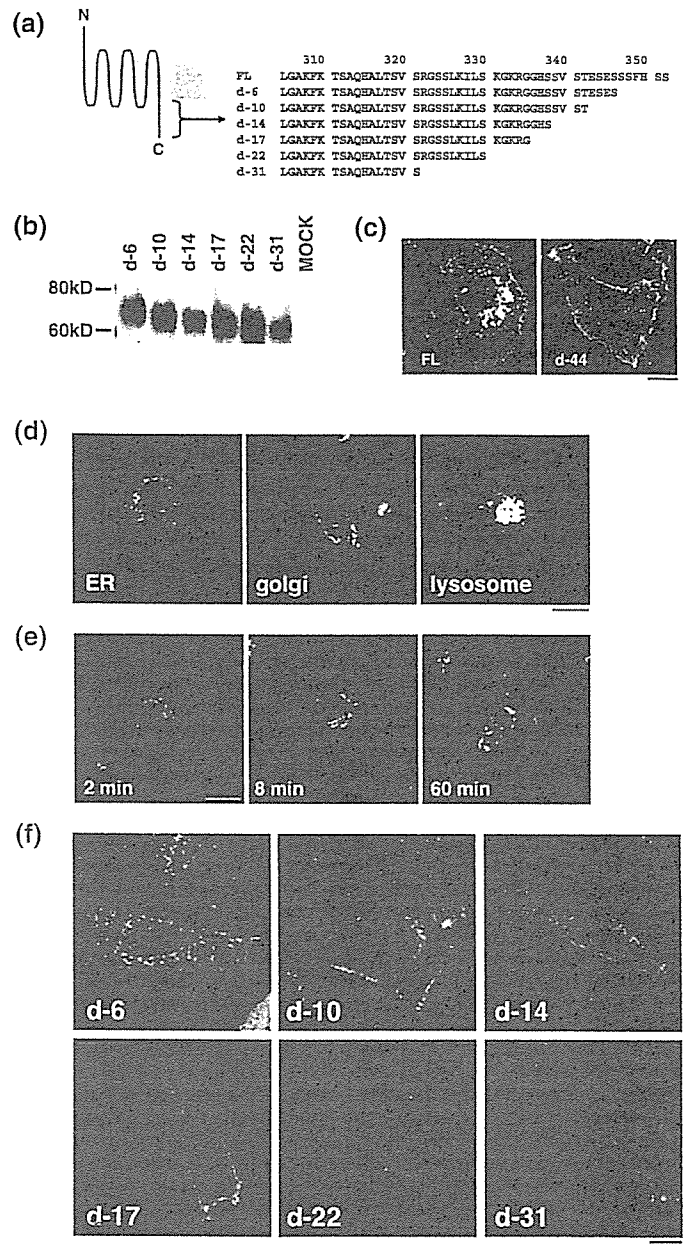

Fig. 2. Expression profiles of CXCR4 and a mutant with cytoplasmic tail deletion. (a) Schematic representation of CXCR4. The N-terminus CXCR4 is exposed in the extracellular space and the C-terminus is intracellular. Gray represents the lipid bilayer. The amino acid sequence of the cytoplasmic tail is shown. Residues in red are required for ligand-induced endocytosis. The CXCR4 d-44 mutant lacks amino acid 309–351. (b) Schematic representation and Western blot of FL and d-44 constructs. (c) Flow cytometry profiles of FL and d-44 expressed in 293T cells. The horizontal axis represents green fluorescence intensity indicative of green fluorescent protein (GFP)-tagged CXCR4 protein levels, and the vertical axis is PE-Cy5 fluorescence intensity, reflecting cell surface CXCR4 detected by the anti-CXCR4 antibody. GFP-positive cells expressing FL are colored in red (left) and those expressing d-44 in green (middle). The expressional differences between FL and d-44 is highlighted on the overlay plot (right).

observations revealed that cells expressing FL were bordered by green fluorescence, and significant green fluorescence was detected in vesicular compartments of varying diameters lying close to the nucleus surrounding the nucleus (hereafter designated the FL phenotype, Fig. 2c, left). Vesicles around the nucleus were

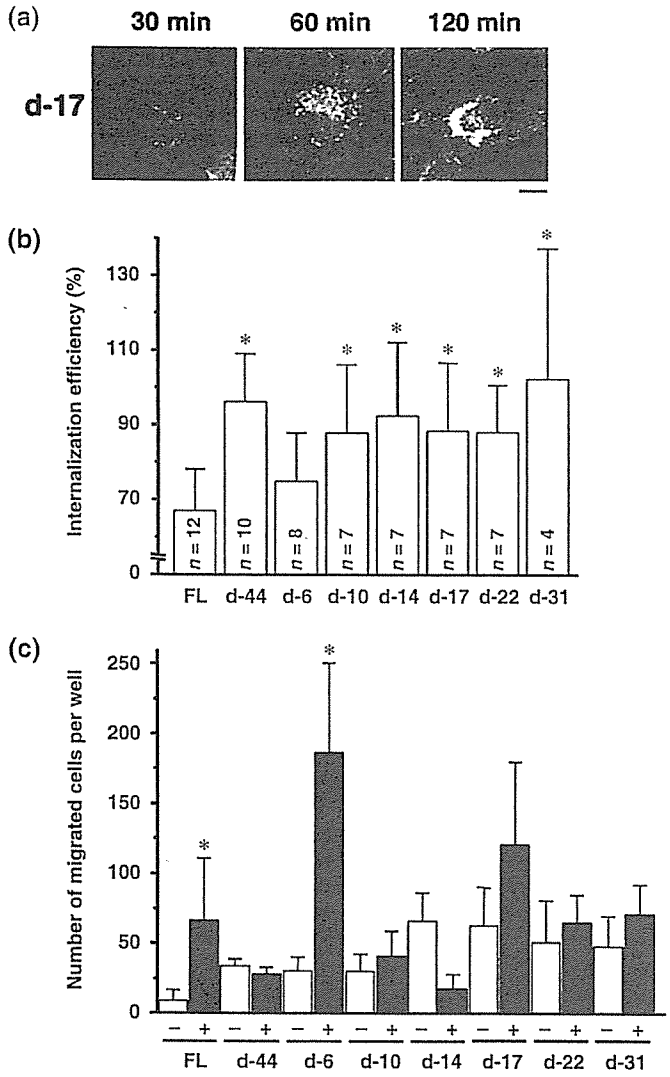

Fig. 3. Identification of the amino acids required for steady-state CXCR4 internalization. (a) Amino acid sequences of the cytoplasmic tail of FL and deletion mutants. Residues in red are required for ligand-induced endocytosis. (b) The protein expression of each mutant in 293T cells was verified by Western blot analysis. (c) Confocal micrographs of NP2 cells expressing FL and d-44 mutant proteins. The blue signal represents the Hoechst-stained nucleus. (Original magnification, ×630; bar, 10 μm.) (d) Confocal micrographs showing NP2 cells expressing CXCR4 FL stained with ER, Golgi, or lysosome organella markers. The organella marker signal is shown in red, the GFP signal is in green. The pixels that both red and green signals co-localized are shown in yellow. (Original magnification, ×630; bar, 10 μm.) (e) CXCR4 FL trafficking in the absence of SDF-1α in NP2 cells. Cell surface CXCR4 FL was labeled with an antibody conjugated with PE-Cy5 (red), incubated at 37°C for the indicated times, fixed and imaged. (Original magnification, ×630; bar, 10 μm.) (f) Confocal micrographs of NP2 cells expressing FL and mutant proteins. The intracellular vesicular green fluorescence reflecting steady-state internalization can be seen in the d-6 mutant. The blue signal represents the Hoechst-stained nucleus. (Original magnification, ×630; bar, 10 μm.)

mostly lysosomes, as demonstrated by fluorescent organella marker analyses in which cells expressing CXCR4 FL-GFP stained with the lysosomal marker yielded a substantial amount of co-localization signal. On the other hand, only a mall amount of co-localization signal was detected when the ER or Golgi markers were used (Fig. 2d), consistent with our biochemical fractionation (unpublished data) and previous publications.[16,27,28,30] The active constitutive internalization was visualized by labeling cell surface CXCR4 by PE-Cy5-conjugated monoclonal antibody followed by fluorescence imaging after cells were incubated at 37°C (Fig. 2e). The live cell imaging revealed that internalizing GFP-positive vesicles trafficked at an average velocity of 4.7 mm/h ($n = 15$), which is within the range of clathrin-dependent vesicular transport (2–20 mm/h), not that of caveolin-dependent vesicular transport (25–170 mm/h).[31–35] These data suggest that the FL is constitutively internalized from the cell surface to the cytoplasmic compartment. In sharp contrast, most green fluorescent signals from d-44 mutant-expressing cells were detected at the cell surface, and only a few small GFP-positive vesicles were seen in the cytoplasm near the nucleus (hereafter designated the d-44 phenotype, Fig. 2c, right). Similar observations were made in d-10, d-14, d-17, d-22 and d-31 mutant-expressing cells (Fig. 2f). The d-6 construct displayed a phenotype similar to FL, although the intracellular GFP signal was less prominent (Fig. 2c). Similar results were obtained in HeLa and 293 cells (data not shown). These data suggest that wild-type CXCR4 was trafficked to the plasma membrane but was internalized spontaneously. Thus, steady-state internalization appeared to be regulated by amino acids located between d-6 and d-10 (e.g. amino acids 343–346).

Steady-state and SDF-1α-induced CXCR4 internalization is genetically separable. Next, we investigated distribution of CXCR4 protein and cell migration after SDF-1α treatment. Confocal analysis showed that after SDF-1α exposure, cells expressing FL, d-6 and d-17 mutants showed GFP signals in intracellular compartments, which were enhanced 60 min after SDF-1α treatment, an effect most clearly shown in d-17-expressing cells (Fig. 3a). GFP signals from intracellular vesicles gradually disappeared 1–2 h after exposure to ligand. Such redistribution of GFP signals was not observed in cells expressing d-10, d-14, d-22, d-31 and d-44 (data not shown). Cell surface levels of CXCR4 before and after SDF-1α treatment were measured by FACS analysis undertaken with an antibody directed against the CXCR4 N-terminus, because that antibody did not interfere with ligand–receptor interaction (Fig. 3b). The downregulation of cell surface levels of FL 2 h after ligand exposure was 67.1 ± 11.1%, whereas that of d-44 was 96.3 ± 12.3% (average and standard deviation from 12 and 10 independent experiments, respectively), consistent with previous reports.[17,27,28] Ligand-induced downregulation of d-6 was 74.9 ± 12.9% ($n = 8$), similar to FL levels. Ligand-induced internalization was significantly less efficient in cells expressing d-10, d-14, d-17, d-22, d-31 and d-44 mutants when compared with FL ($P < 0.001$). Although the d-17 mutant supported ligand-facilitated internalization, as evidenced by microscopic observation, cell surface levels remained unchanged (Fig. 3a,b). This may be due in part to rapid recruitment of newly synthesized d-17 to the cell surface.

Next, we examined cells expressing CXCR4 mutants in response to SDF-1α. Migration results from intracellular signaling initiated by SDF-1α/CXCR4 interaction. Induction of cell migration by SDF-1α in cells expressing FL was 7.2-fold that of untreated cells ($P < 0.05$). In contrast, migration of cells expressing d-44 in response to SDF-1α was undetectable. These data are in agreement with a previous report.[26] The d-6 mutant, which is internalized upon SDF-1α treatment, supported ligand-promoted cell migration by 6.1-fold ($P < 0.01$) relative to untreated cells, similar to FL. Other deletion mutants tested did not display enhanced cell migration following ligand treatment, except for d-17, which showed modestly enhanced (1.9-fold) migration relative to untreated cells, which was not statistically significant. When basal migratory activities were compared, removal of six or more amino acids from the cytoplasmic tail appeared to potentiate migration in the absence of ligand (open bars, Fig. 3c). These data suggest that constitutive internalization is regulated independently of ligand-facilitated internalization.

Identification of CXCR4 S(E/D)S as a ligand-independent internalization motif. The above data indicated that the carboxy-terminal four

doi: 10.1111/j.1349-7006.2007.00388.x

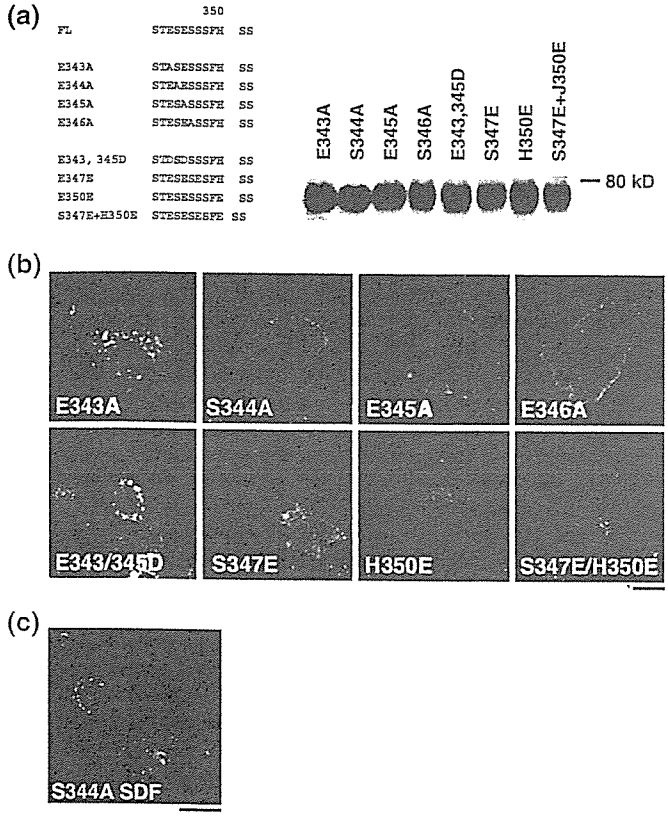

Fig. 4. Characterization of the SDF-1α-independent internalization motif of CXCR4. (a) Left, amino acid sequences of CXCR4 FL and substitution mutants. Letters in red indicate introduced mutations. Right, protein expression of each mutant in 293T cells was verified by Western blot analysis. (b) Confocal micrographs of NP2 cells expressing each mutant. The blue signal represents the Hoechst-stained nucleus. (Original magnification, ×630; bar, 10 μm.) (c) NP2 cells expressing the S344A mutant treated with SDF-1α for 2 h are shown. The blue signal represents the Hoechst-stained nucleus. (Original magnification, ×630; bar, 10 μm.)

amino acids (ESES; residues 343–346) likely function in ligand-independent CXCR4 internalization. To further characterize which amino acids are required for ligand-independent internalization, we generated alanine substitution mutants for each of the four amino acids in the context of FL and examined their phenotypes (Fig. 4a). Protein expression of mutants was verified in Western blot analysis (Fig. 4a). Among the four mutants, the E343A mutant showed the FL phenotype, while the others displayed the d-44 phenotype in the absence of ligand (Fig. 4b). These data demonstrate that Ser^{344}-Glu^{345}-Ser^{346} constitute the core motif for SDF-1α-independent CXCR4 internalization. Both E345A and S346A mutants exhibited the Thr^{342}-Glu^{343}-Ser^{344} sequence adjacent to the original SES sequence. However, this 'SES-like' motif did not support constitutive internalization, suggesting that Thr cannot substitute for Ser to maintain functionality as a constitutive internalization motif. We reasoned that if such a motif requires an acidic amino acid between two serine residues, changing Glu to Asp should maintain the motif's function. Thus, we constructed a mutant in which Glu was replaced with Asp (E343/345D; Fig. 4a). Also, to determine whether two adjacent SES sequences could augment the FL phenotype, we substituted Ser^{347} with Glu (S347E), creating an additional SES motif next to the original SES one (Fig. 4a). As controls, we created H350E and S347E/H350E mutants (Fig. 4a). Expression of these mutants was verified by Western blot analysis (Fig. 4a). Interestingly, the E343/345D mutant retained the FL phenotype (Fig. 4b), indicating that an acidic residue is required to maintain function of the constitutive internalization motif. S347E showed an intermediate phenotype in which numerous fine GFP-positive vesicles were seen close to the nucleus (Fig. 4b). These data indicate that the two adjacent SES sequences do not augment the FL phenotype but actually interfere with steady-state internalization. Both H350E and S347E/H350E mutants also showed an intermediate phenotype (Fig. 4b), suggesting that more than three acidic amino acids close to the SES motif may inhibit its function, potentially by generating a negative charge cluster. Overall, we conclude that the SDF-1α-independent internalization motif is located at amino acids 344–346 of the CXCR4 cytoplasmic tail.

Finally, we analyzed phenotypes of the S344A mutant in greater detail. Two hours after SDF-1α treatment, cells expressing this mutant showed accumulation of GFP signals at perinuclear regions, similar to the d-17 mutant (Figs. 1a and 4c). FACS analysis revealed that cell surface levels of S344A decreased to 70.8 ± 11.7% ($n = 7$) following SDF-1α treatment relative to untreated cells, almost as efficient as FL (Fig. 3b). Migratory activity of cells expressing the S344A mutant was stimulated 3.0-fold by SDF-1α, while that of cells expressing FL assayed in parallel showed a 5.8-fold increase relative to untreated cells. These data demonstrate that the S344A mutant, which is defective in constitutive internalization, can undergo ligand-dependent internalization and stimulate migration. Along with the d-17 data, our observations strongly suggest that genetic elements required for the ligand-dependent and -independent internalization are separable.

Discussion

We demonstrated here that CXCR4 is constitutively internalized in the absence of SDF-1α and that steady-state trafficking of CXCR4 is regulated by its cytoplasmic tail. We show that the three amino acid motif, Ser^{344}-Glu^{345}-Ser^{346}, within the cytoplasmic tail is essential for efficient steady-state internalization of CXCR4. Our work indicates that ligand-independent internalization of CXCR4 is genetically separable from ligand-dependent internalization: mutants defective in steady-state internalization (d-17 and S344A) were competent to respond to SDF-1α-promoted internalization signals. That residues required for ligand-dependent endocytosis (Ser324, 325, 330, 338, 339, Ile^{328} Leu^{329} and Lys^{331}; summarized in Fig. 1a)[16–18] do not overlap with those required for ligand-independent internalization, further supports the idea that these activities are separable.

Interestingly, the d-17 mutant displayed SDF-1α-promoted internalization, whereas the d-14 and d-22 mutants did not. These data suggest that an element between amino acids 336 and 342 negatively regulates ligand-initiated CXCR4 internalization. We are currently determining what amino acids are required for that motif. What is unique about the constitutive internalization motif is its position effect in terms of the distance of the motif from the "body" of the receptor. SES-like motifs can be found in the cytoplasmic tails of both CXC-chemokine receptors (for example, CXCR3) and CC-chemokine receptors including CCR2, CCR5 and CCR7. Indeed, these receptors share similar amino acid sequences in which two acidic amino acids (mostly Asp) positioned between the 36th and 45th amino acids of the cytoplasmic tail, where Ser and Thr residues are often in the close proximity to the acidic amino acids but positively charged amino acids, are infrequent. We hypothesize that for the ligand-independent internalization motif to function, the SES motif or its equivalent must be positioned at approximately the 40th residue of the cytoplasmic tail.

Many GPCR, including α1a-adrenoceptor, and the μ-opioid receptor, are spontaneously internalized.[36,37] Therefore, we conclude that various GPCR actively and continuously undergo endocytosis in the absence of ligand in a manner similar to

CXCR4 and hypothesize that the function of constitutive receptor internalization is to fine-tune the threshold at which cells sense ligand. Cells should be able to rapidly post up- and downregulate cell surface levels of CXCR4 using post-translational mechanisms. Such regulation should enable cells to migrate toward SDF-1α-rich tissues as needed and should also prevent inappropriate cells from migrating.

Our work is relevant to cancer cell metastasis and the pathogenesis of WHIM syndrome. Cell surface levels of CXCR4 positively correlate with cancer cells' ability to metastasize.[5,19] We hypothesize that enhanced metastatic capabilities of cancer cells could be due in part to mutations that disrupt the function of SES motif, which would result in upregulation of cell surface levels of signal-competent CXCR4 (as exemplified by the E344A mutant). As for WHIM syndrome, it was recently reported that it is due to mutations within CXCR4's cytoplasmic domain.[14,38] Interestingly, these mutations result in loss of SES motif. We predict that loss of the SES motif should increase cell surface CXCR4 levels. Although CXCR4 mutations generated here are not identical to reported WHIM mutations, the d-10 mutant resembles mutations seen in WHIM syndrome, and it exhibits enhanced basal cell migratory activity. Increased cell surface CXCR4 or increased migratory potential may contribute to WHIM pathogenesis. The response of d-10-expressing cells to SDF-1α, however, was not as robust as that of cells derived from WHIM.[39,40] This discordance may be partly due to the cell type differences, as we have employed a glioblastoma cell line for our studies.

Thus, CXCR4 is a potentially important therapeutic target not only for cancers but for other conditions such as HIV-1 infection, chronic autoimmune disease, and genetic disorders including WHIM syndrome. CXCR4 also plays critical roles in embryogenesis, homeostasis and inflammation. Although there are potential caveats for treating cancer with CXCR4 antagonists, our data furthers the understanding of mechanisms regulating CXCR4 and could be useful in devising therapeutic strategies.

Acknowledgments

We thank Drs Toshitada Takemori and Tsutomu Murakami for critical reading of the manuscript. This work was supported in part by the Japan Human Science Foundation, the Japanese Ministry of Health, Labor and Welfare, and the Japanese Ministry of Education, Culture, Sports, Science and Technology.

References

1 Gether U. Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocr Rev* 2000; **21**: 90–113.
2 Ferguson SS. Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling. *Pharmacol Rev* 2001; **53**: 1–24.
3 Sapede D, Rossel M, Dambly-Chaudiere C, Ghysen A. Role of SDF1 chemokine in the development of lateral line efferent and facial motor neurons. *Proc Natl Acad Sci USA* 2005; **102**: 1714–8. Epub 2005 January 19.
4 Coughlan CM, McManus CM *et al.* Expression of multiple functional chemokine receptors and monocyte chemoattractant protein-1 in human neurons. *Neuroscience* 2000; **97**: 591–600.
5 Muller A, Homey B, Soto H *et al.* Involvement of chemokine receptors in breast cancer metastasis. *Nature* 2001; **410**: 50–6.
6 Zou YR, Kottmann AH, Kuroda M, Taniuchi I, Littman DR. Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature* 1998; **393**: 595–9.
7 Ma Q, Jones D, Borghesani PR *et al.* Impaired B-lymphopoiesis, myelopoiesis, and derailed cerebellar neuron migration in CXCR4- and SDF-1-deficient mice. *Proc Natl Acad Sci USA* 1998; **95**: 9448–53.
8 Wang JF, Park IW, Groopman JE. Stromal cell-derived factor-1alpha stimulates tyrosine phosphorylation of multiple focal adhesion proteins and induces migration of hematopoietic progenitor cells: roles of phosphoinositide-3 kinase and protein kinase C. *Blood* 2000; **95**: 2505–13.
9 Kawabata K, Ujikawa M, Egawa T *et al.* A cell-autonomous requirement for CXCR4 in long-term lymphoid and myeloid reconstitution. *Proc Natl Acad Sci USA* 1999; **96**: 5663–7.
10 Peled A, Petit I, Kollet O *et al.* Dependence of human stem cell engraftment and repopulation of NOD/SCID mice on CXCR4. *Science* 1999; **283**: 845–8.
11 Zeelenberg IS, Ruuls-Van Stalle L, Roos E. The chemokine receptor CXCR4 is required for outgrowth of colon carcinoma micrometastases. *Cancer Res* 2003; **63**: 3833–9.
12 Phillips RJ, Burdick MD, Lutz M, Belperio JA, Keane MP, Strieter RM. The stromal derived factor-1/CXCL12-CXC chemokine receptor 4 biological axis in non-small cell lung cancer metastases. *Am J Respir Crit Care Med* 2003; **167**: 1676–86.
13 Taichman RS, Cooper C, Keller ET, Pienta KJ, Taichman NS, McCauley LK. Use of the stromal cell-derived factor-1/CXCR4 pathway in prostate cancer metastasis to bone. *Cancer Res* 2002; **62**: 1832–7.
14 Hernandez PA, Gorlin RJ, Lukens JN *et al.* Mutations in the chemokine receptor gene CXCR4 are associated with WHIM syndrome, a combined immunodeficiency disease. *Nat Genet* 2003; **34**: 70–4.
15 Lefkowitz RJ, Shenoy SK. Transduction of receptor signals by beta-arrestins. *Science* 2005; **308**: 512–7.
16 Marchese A, Benovic JL. Agonist-promoted ubiquitination of the G protein-coupled receptor CXCR4 mediates lysosomal sorting. *J Biol Chem* 2001; **276**: 45 509–12.
17 Orsini MJ, Parent JL, Mundell SJ, Benovic JL, Marchese A. Trafficking of the HIV coreceptor CXCR4. Role of arrestins and identification of residues in the c-terminal tail that mediate receptor internalization. *J Biol Chem* 1999; **274**: 31 076–86.
18 Marchese A, Raiborg C, Santini F, Keen JH, Stenmark H, Benovic JL. The E3 ubiquitin ligase AIP4 mediates ubiquitination and sorting of the G protein-coupled receptor CXCR4. *Dev Cell* 2003; **5**: 709–22.
19 Darash-Yahana M, Pikarsky E, Abramovitch R *et al.* Role of high expression levels of CXCR4 in tumor growth, vascularization, and metastasis. *FASEB* 2004; **J 18**: 1240–2.
20 Vila-Coro AJ, Rodriguez-Frade JM, Martin De Ana A, Moreno-Ortiz MC, Martinez AC, Mellado M. The chemokine SDF-1alpha triggers CXCR4 receptor dimerization and activates the JAK/STAT pathway. *FASEB J* 1999; **13**: 1699–710.
21 Babcock GJ, Farzan M, Sodroski J. Ligand-independent dimerization of CXCR4, a principal HIV-1 coreceptor. *J Biol Chem* 2003; **278**: 3378–85.
22 Terrillon S, Bouvier M. Roles of G-protein-coupled receptor dimerization. *EMBO Rep* 2004; **5**: 30–4.
23 Hu H, Shioda T, Hori T *et al.* Dissociation of ligand-induced internalization of CXCR-4 from its co-receptor activity for HIV-1 Env-mediated membrane fusion. *Arch Virol* 1998; **143**: 851–61.
24 Yanagida M, Hayano T, Yamauchi Y *et al.* Human fibrillarin forms a sub-complex with splicing factor 2-associated p32, protein arginine methyltransferases, and tubulins alpha 3 and beta 1 that is independent of its association with preribosomal ribonucleoprotein complexes. *J Biol Chem* 2004; **279**: 1607–14.
25 Komano J, Miyauchi K, Matsuda Z, Yamamoto N. Inhibiting the Arp2/3 complex limits infection of both intracellular mature vaccinia virus and primate lentiviruses. *Mol Biol Cell* 2004; **15**: 5197–207.
26 Roland J, Murphy BJ, Ahr B *et al.* Role of the intracellular domains of CXCR4 in SDF-1-mediated signaling. *Blood* 2003; **101**: 399–406.
27 Haribabu B, Richardson RM, Fisher I *et al.* Regulation of human chemokine receptors CXCR4. Role of phosphorylation in desensitization and internalization. *J Biol Chem* 1997; **272**: 28 726–31.
28 Tarasova NI, Stauber RH, Michejda CJ. Spontaneous and ligand-induced trafficking of CXC-chemokine receptor 4. *J Biol Chem* 1998; **273**: 15 883–6.
29 Soda Y, Shimizu N, Jinno A *et al.* Establishment of a new system for determination of coreceptor usages of HIV based on the human glioma NP-2 cell line. *Biochem Biophys Res Commun* 1999; **258**: 313–21.
30 Zhang Y, Foudi A, Geay JF *et al.* Intracellular localization and constitutive endocytosis of CXCR4 in human CD34+ hematopoietic progenitor cells. *Stem Cells* 2004; **22**: 1015–29.
31 Mundy DI, Machleidt T, Ying YS, Anderson RG, Bloom GS. Dual control of caveolar membrane traffic by microtubules and the actin cytoskeleton. *J Cell Sci* 2002; **115**: 4327–39.
32 Rappoport JZ, Taha BW, Lemeer S, Benmerah A, Simon SM. The AP-2 complex is excluded from the dynamic population of plasma membrane-associated clathrin. *J Biol Chem* 2003; **278**: 47 357–60.
33 Rappoport JZ, Simon SM. Real-time analysis of clathrin-mediated endocytosis during cell migration. *J Cell Sci* 2003; **116**: 847–55.
34 Keyel PA, Watkins SC, Traub LM. Endocytic adaptor molecules reveal an endosomal population of clathrin by total internal reflection fluorescence microscopy. *J Biol Chem* 2004; **279**: 13 190–204.
35 Yarar D, Waterman-Storer CM, Schmid SL. A dynamic actin cytoskeleton functions at multiple stages of clathrin-mediated endocytosis. *Mol Biol Cell* 2005; **16**: 964–75.

doi: 10.1111/j.1349-7006.2007.00388.x

36 Pediani JD, Colston JF, Caldwell D, Milligan G, Daly CJ, McGrath JC. Beta-arrestin-dependent spontaneous alpha1a-adrenoceptor endocytosis causes intracellular transportation of alpha-blockers via recycling compartments. *Mol Pharmacol* 2005; **67**: 992–1004.
37 Segredo V, Burford NT, Lameh J, Sadee W. A constitutively internalizing and recycling mutant of the mu-opioid receptor. *J Neurochem* 1997; **68**: 2395–404.
38 Gulino AV, Moratto D, Sozzani S *et al.* Altered leukocyte response to CXCL12 in patients with warts hypogammaglobulinemia, infections, myelokathexis (WHIM) syndrome. *Blood* 2004; **104**: 444–52.
39 Kawai T, Choi U, Whiting-Theobald NL *et al.* Enhanced function with decreased internalization of carboxy-terminus truncated CXCR4 responsible for WHIM syndrome. *Exp Hematol* 2005; **33**: 460–8.
40 Balabanian K, Lagane B, Pablos JL *et al.* WHIM syndromes with different genetic anomalies are accounted for by impaired CXCR4 desensitization to CXCL12. *Blood* 2005; **105**: 2449–57.

QAD 200762

Inhibiting lentiviral replication by HEXIM1, a cellular negative regulator of the CDK9/cyclin T complex

Saki Shimizua[a,b], Emiko Urano[a], Yuko Futahashi[a], Kosuke Miyauchi[a], Maya Isogai[a], Zene Matsuda[a], Kyoko Nohtomi[a], Kazunari Onogi[a], Yutaka Takebe[a], Naoki Yamamoto[a,b] and Jun Komano[a]

Objective: Tat-dependent transcriptional elongation is crucial for the replication of HIV-1 and depends on positive transcription elongation factor b complex (P-TEFb), composed of cyclin dependent kinase 9 (CDK9) and cyclin T. Hexamethylene bisacetamide-induced protein 1 (HEXIM1) inhibits P-TEFb in cooperation with 7SK RNA, but direct evidence that this inhibition limits the replication of HIV-1 has been lacking. In the present study we examined whether the expression of FLAG-tagged HEXIM1 (HEXIM1-f) affected lentiviral replication in human T cell lines.

Methods: HEXIM1-f was introduced to five human T cell lines, relevant host for HIV-1, by murine leukemia virus vector and cells expressing HEXIM1-f were collected by fluorescence activated cell sorter. The lentiviral replication kinetics in HEXIM1-f-expressing cells was compared with that in green fluorescent protein (GFP)-expressing cells.

Results: HIV-1 and simian immunodeficiency virus replicated less efficiently in HEXIM1-f-expressing cells than in GFP-expressing cells of the five T cell lines tested. The viral revertants were not immediately selected in culture. In contrast, the replication of vaccinia virus, adenovirus, and herpes simplex virus type 1 was not limited. The quantitative PCR analyses revealed that the early phase of viral life cycle was not blocked by HEXIM1. On the other hand, *tat*-dependent transcription in HEXIM1-f-expressing cells was substantially repressed as compared with that in GFP-expressing cells.

Conclusion: These data indicate that HEXIM1 is a host factor that negatively regulates lentiviral replication specifically. Elucidating the regulatory mechanism of HEXIM1 might lead to ways to control lentiviral replication.

AIDS 2007, 21:1–8

Keywords: CDK9, cyclin T, HEXIM1, lentivirus, *tat*

Introduction

Activation of transcription elongation requires the positive transcription elongation factor b complex (P-TEFb) composed of cyclin dependent kinase 9 (CDK9) and cyclin T1, T2, or K [1]. P-TEFb is essential for efficient transcriptional elongation from the promoter of human immunodeficiency virus type 1 (HIV-1), the long terminal repeat (LTR) (reviewed in [2,3]). The functional interaction between P-TEFb and the viral protein Tat has been well studied. Immediately after viral transcription starts at the LTR of the integrated proviral genome, the nascent viral transcript forms a three-dimensional structure called TAR. In the presence of P-TEFb, Tat binds to TAR. Through the Tat–TAR interaction, Tat activates P-TEFb and therefore assures the efficient

From the [a]AIDS Research Center, National Institute of Infectious Diseases, Tokyo, and the [b]Department of Molecular Virology, Tokyo Medical and Dental University, Tokyo, Japan.

Correspondence to Jun Komano, AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku, Tokyo 162-8640, Japan.

E-mail: ajkomano@nih.go.jp

AQ1 Received: ?? ??; revised: ?? ??; accepted: ?? ??.

ISSN 0269-9370

completion of viral gene transcription and the propagation of HIV-1.

Recently, the regulatory mechanisms of P-TEFb function have been elucidated. In 2001, the interaction of P-TEFb with 7SK RNA was found to be necessary to inactivate the kinase activity of CDK9 within P-TEFb [4–6]. However, the binding of 7SK RNA alone is not sufficient to inactivate P-TEFb. More recently, Yik *et al.* demonstrated that the inactivation of P-TEFb requires hexamethylene bisacetamide-induced protein 1 (HEXIM1; synonyms CLP1, MAQ1, and HIS1) [7–9]. The inactivation of P-TEFb by the HEXIM1-7SK RNA complex appears to regulate the transcriptional elongation of cellular genes.

The HEXIM1-7SK RNA complex has been shown to physically compete with Tat for binding to P-TEFb [10]. In agreement with this finding, HEXIM1 was shown to inhibit Tat-dependent transcription from the HIV-1 LTR in transient transfection assays [8,11,12]. However, no data demonstrating that HEXIM1 is able to limit HIV-1 replication has been provided. Here we provide direct experimental evidence that the constitutive expression of HEXIM1 specifically limits lentiviral replication.

Methods

Plasmids

The FLAG-tagged HEXIM1 expression constructs were generated by reverse-transcription PCR using RNA isolated from CEM cells as templates. The primers used were 5′-CACCTCGAGCCACCATGGACTACAAAGACGATGACGACAAGGCCGAGCCATTCTTGTC-3′ and 5′-CAATTGCTAGTCTCCAAACTTGGAAAGCGGCGC-3′′ for amino terminus FLAG tagging, and 5′-CACCTCGAGCCACCATGGCCGAGCCATTCTTGTCAGAATATC-3′ and 5′-CAATTGCTAGTCGTCATCGTCTTTGTAGTCGTCTCCAAACTTGGAAAGCGGCGCTC-3′ for carboxy terminus FLAG tagging. The *Xho*I-*Mfe*I fragments of the PCR products were cloned into the *Xho*I-*Mfe*I sites of pCMMP IRES GFP, generating pCMMP f-HEXIM1 and pCMMP HEXIM1-f [13]. The cytomegalovirus (CMV) promoter-driven *gag-pol* expression vector psyn*gag-pol* has been previously described by Wagner *et al.* [14] and pLTR*gag-pol* was constructed by cloning the *Mlu*I-*Hin*dIII fragment encoding the LTR from pNL-luc [15] into the *Mlu*I-*Hin*dIII sites of psyn*gag-pol*. The tax expressing plasmid pCGtax and pHTLV LTR luciferase were kindly provided by Dr. Watanabe (Tokyo Medical Institute). The *tat*-expressing plasmid pSVtat was a generous gift from Dr. Freed (National Cancer Institute-Frederick, Frederick, Maryland,USA). The plasmid pLTR-luc has been described previously (Miyauchi *et al.*, *Antiviral Chemistry and Chemotherapy*, in press). The following plasmids have been described previously by Komano *et al.* [13]: pVSV-G, pMD*gag-pol*, pTM3Luci, phRL-CMV and pSIVmac239ΔnefLuc.

Cells and transfection

All the mammalian cells were maintained in RPMI 1640 (Sigma, St Louis, Missouri, USA) supplemented with 10% fetal bovine serum (Japan Bioserum, Tokyo, Japan), penicillin and streptomycin (Invitrogen, Tokyo, Japan). Cells were incubated at 37°C in a humidified 5% CO_2 atmosphere. Cells were transfected using Lipofectamine 2000 according to the manufacturer's protocol (Invitrogen).

Western blotting

Cells were lysed with sample buffer, sonicated, and boiled for 5 min. Samples were separated on 8% sodium dodecyl sulfate-polyacrylamide gel electrophoresis gels and transferred to polyvinylidene difluoride membranes (Millipore, Billerica, Massachusetts, USA) for western blotting according to standard techniques. Membranes were blocked with Tris-buffered saline containing 0.05% Tween-20 (TBS-T) containing 5% (w/v) non-fat skim milk (Yuki-Jirushi, Tokyo, Japan) for 1 h at room temperature and incubated with primary antibodies including the M2 anti-FLAG epitope monoclonal antibody (Sigma), an anti-actin monoclonal antibody (MAB1501R; Chemicon/Millipore, Billerica, Massachusetts, USA), an anti-cyclin T1 rabbit polyclonal antibody (H-245; Santa Cruz Biotechnology, Santa Cruz, California, USA), an anti-cyclin T2a/b goat polyclonal antibody (A-20; Santa Cruz), an anti-p24 monoclonal antibody (183-H12-5C; NIH AIDS Research and Reference Reagent Program), an anti-HIS1 chicken polyclonal antibody (N-150; GenWay), and an anti-Bip/GRP78 monoclonal antibody (clone 40; BD Biosciences/Transduction Laboratories, San Jose, California, USA) for 1 h at room temperature. Membranes were washed with TBS-T and incubated with appropriate second antibodies including biotinylated anti-goat (GE Healthcare Bio-Sciences, Piscataway, New Jersey, USA) or anti-chicken IgY (Promega, Madison, Wisconsin, USA), and EnVision+ (Dako, Glostrup, Denmark) for 1 h at room temperature. For a tertiary probe, we used horseradish peroxidase (HRP)-streptavidine (GE Healthcare) if necessary. Signals were visualized with an LAS3000 imager (Fujifilm, Tokyo, Japan) after treating the membranes with the Lumi-Light Western Blotting Substrate (Roche Diagnostics GmbH, Mannheim, Germany).

Reporter assay

Luciferase activity was measured 48 h after transfection or infection using a DualGlo assay kit (Promega) according to the manufacturer's protocol. The beta-galactosidase activity was measured using a LumiGal assay kit (BD Biosciences/Clontech, San Jose, California, USA) according to the manufacturer's protocol. The chemilu-

minescence was detected with a Veritas luminometer (Promega).

Monitoring viral replication

To monitor HIV-1 replication, the culture supernantants were subjected to either a reverse transcriptase assay [16] or an enzyme-linked immunosorbent assay (ELISA) to detect p24 antigens using a Retro TEK p24 antigen ELISA kit according to the manufacturer's protocol (Zepto Metrix, Buffalo, New York, USA). For simian immunodeficiency virus (SIV) a p27 antigen ELISA kit was used according to the manufacturer's protocol (Zepto Metrix). The signals were measured with a Multiskan Ex microplate photometer (ThermoLabsystems, Helsinki, Finland). For vaccinia virus, adenovirus, and herpes simplex virus (HSV)-1, the activity of reporter genes was measured as previously described [13].

Generating viruses

To produce HIV-1 and SIV, 293T cells were transfected with plasmids encoding proviral DNA of HIV-1 (pHXB2) or pSIVmac239ΔnefLuc and culture supernatants containing viruses were collected at 48 h post-transfection. Murine leukemia virus (MLV) and lentiviral vectors pseudotyped with VSV-G were produced as described previously by cotransfecting 293T cells with either the pNL-Luc and pVSV-G vectors or the pMD*gag-pol*, pVSV-G, and pCMMP vectors [13]. Green fluorescent cells were sorted by fluorescence activated cell sorter (FACS) Aria (Becton Dickinson, San Jose, California, USA).

Reverse transcriptase-polymerase chain reaction

Total RNA was isolated with an RNeasy kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer's instruction. The reverse transcriptase (RT)-polymerase chain reaction (PCR) assay was performed with a One Step RNA PCR Kit (Takara, Otsu, Japan), imaged by a Typhoon scanner 9400 (GE Healthcare), and quantified with Image Quant software (GE Healthcare). For the amplification of endogenous HEXIM1, the forward primer 5′-ACCACACGGAGAGCCTGCA-GAAC-3′ and the reverse primer 5′-TAGCTAAA-TTTACGAAACCAAAGCC-3′ were used. For the amplification of HEXIM1-f, the forward primer 5′-GTACCTGGAACTGGAGAAGTGCCC-3′ and the reverse primer 5′-CAATTGCTAGTCGTCATCGTC-TTTGTAGTC-3′ were used. For cyclophilin A, the forward primer 5′-CACCGCCACCATGGTCAAC-CCCACCGTGTTCTTCGAC-3′ and the reverse primer 5′-CCCGGGCCTCGAGCTTTCGAGTTGT-CCACAGTCAGCAATGG-3′ were used.

Quantitative real time polymerase chain reaction

The real time PCR reaction was performed in a DNA Engine Opticon 2 Continuous Fluorescence Detection System (Bio-Rad, Hercules, California, USA). The cellular genomic DNA and total RNA were extracted 48 h post-infection with a DNeasy kit (Qiagen) and RNeasy kit (Qiagen), respectively, according to the manufacturer's instruction. For the reagents, we used QuantiTect SYBR Green PCR and RT-PCR Kits (Qiagen). To estimate the amount of integrated HIV-1 DNA, Alu-LTR PCR was performed according to the method described previously using the following primers: for the first PCR, 5′-AACTAGGGAACCCACTGCT-TAAG-3′ and 5′-TGCTGGGATTACAGGCGTGAG-3′, and for the second PCR, 5′-AACTAGGGAACC-CACTGCTTAAG-3′ and 5′-CTGCTAGAGATTT-TCCACACTGAC-3′ [17]. The beta-globin primers have been described previously [18]. To estimate the amount of HIV-1 RNA, the second PCR primers for the Alu-LTR PCR were used. The primers for cyclophilin A are described above.

Results and discussion

The HEXIM1 cDNA tagged with a FLAG epitope at either the amino terminus (f-HEXIM1) or the carboxy terminus (HEXIM1-f) was cloned in a mammalian expression plasmid (Fig. 1a). A luciferase assay revealed that the Tat-dependent enhancement of transcription from the HIV-1 LTR was reduced by co-transfecting HEXIM1-expressing plasmids, whereas neither Tat-independent basal transcription from the HIV-1 LTR nor CMV promoter-driven transcription was affected (Fig. 1b). An oncogenic retrovirus human T cell leukemia virus type 1 (HTLV-1) encodes for *tax*, a functional homologue of HIV-1's *tat*, that utilizes P-TEFb to enhance transcription from the LTR promoter [19]. However, *tax*-dependent enhancement of transcription was not affected by HEXIM1 in similar experimental conditions (Fig. 1c). To monitor the effect of HEXIM1 on HIV-1 replication, we introduced HEXIM1-expressing plasmids into HeLa-CD4 cells along with pNL4-3, which produces replication-competent HIV-1, and measured the RT activity in the culture supernatant 1 week post-transfection. Transfecting HEXIM1-expressing plasmids decreased the RT activity in a dose-dependent manner (Fig. 1d). Next, we asked whether the inhibition of viral replication was specific to HIV-1 by examining vaccinia virus, adenovirus, and HSV-1 replication. We found that the propagation of these three viruses was not inhibited by HEXIM1-f expression (Fig. 1e–g), suggesting that the inhibition of viral replication by HEXIM1 was HIV-1-specific.

To examine whether HEXIM1 negatively affects lentiviral replication in the physiologically relevant host, we isolated human T cell lines constitutively expressing HEXIM1-f. We cloned HEXIM1-f cDNA into a pCMMP (MLV retroviral vector plasmid (Fig. 2a). The plasmid encoded an internal ribosomal entry site (IRES)-mediated green fluorescent protein (GFP)