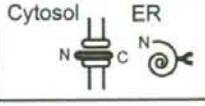
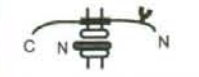
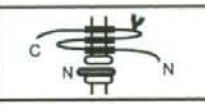
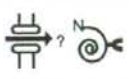
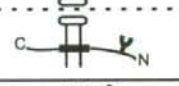
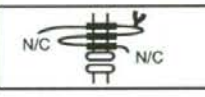
TABLE II
Classification of the 829 N-glycosylated proteins identified in this study based on the presence of a signal peptide and the number of transmembrane segments

	SP	TM	type	topology	typical proteins	
glycoproteins (829)	SP+ (463)	0 (224)	soluble (secreted)		R07B7.11 H22K11.1 F55H2.1	alpha-N-acetylgalactosaminidase asp-3; aspartyl protease superoxide dismutase
		1 (181)	Type I		ZK1067.1 Y105E8A.30 Y37D8A.13	let-23; Tyr-protein kinase (EGFR) lev-10 (levamisole-resistant 10) unc-71; disintegrin
		≥ 2 (58)			T23D8.1 C04F5.1 K09C8.1	mom-5; Wnt receptor sid-1; transporter pbo-1; Na/H exchanger
	SP- (238)	0 (45)	ovalbumin-like		ZC116.3 T05E11.6 ZK994.3	bone morphogenetic protein 1 like hemoglobinase like pxn-1; peroxidase
		1 (76)	Type II		F26D10.9 C38C6.2	atg-1; sugar (and other) transporter atg-2; amino acid transporter
			Type III		C01G6.8 B0273.4	cam-1; Tyr kinase receptor unc-5; axon guidance protein
		≥ 2 (117)			C18F3.2 R13A5.1 F57C12.5	sax-7; neural adhesion protein cup-5; calcium channel mrp-1; ABC transporter
	ambiguous (128)					

The number of proteins classified into each group is indicated in parentheses. The presence of a signal peptide (SP) and the number of transmembrane segments (TM) were predicted as described in the text. Y illustrates the position of an oligosaccharide attached to the polypeptide chain. EGFR, epidermal growth factor receptor; ABC, ATP-binding cassette.

brane proteins, and 160 of those were actually glycosylated only at the N-terminal portion of the transmembrane segment (Fig. 2). Twelve proteins were glycosylated on the C-terminal side of the transmembrane segment, and nine were glycosylated on both sides. Thus, our ability to predict signal peptides and transmembrane segments was quite good (160/181 = 88%) for single span transmembrane proteins.

Fig. 2 indicates that the Type I transmembrane proteins generally have an extracellular or luminal N-terminal segment that is much longer than the intracellular C-terminal segment: the average length of N- and C-terminal segments was 577 and 104 residues, respectively. Typical examples having a long N-terminal segment include membrane-anchored enzymes such as phospholipase, carboxypeptidase, and UDP-glucosyltransferase, and extracellular matrix proteins such as cadherin and integrin. Conversely a number of enzymes, including tyrosine kinase, tyrosine phosphatase, and guanylate cyclase, have a long (>100-residue) C-terminal cytosolic segment. Besides those single span transmembrane proteins, we detected many Type I transmembrane proteins with a signal peptide and multiple transmembrane segments (Supplemental Table 3).

Analysis of Integral Membrane Glycoproteins Lacking a Signal Peptide—We performed a similar analysis on the single span proteins lacking a signal peptide (Fig. 3). Of 76 proteins, 48 and 26 were assigned as Type II and Type III transmembrane proteins, respectively. Only two proteins were modified on both sides of the transmembrane segment. Fig. 3 shows the positions of the glycosylation sites and the putative transmembrane segment on the polypeptide chains of Type II and Type III proteins. In the Type II proteins, the putative transmembrane segment appeared immediately after a short N-terminal sequence. The N-terminal segments that preceded the transmembrane segment had an average length of 82 residues with 39 of 48 Type II proteins (~80%) containing segments of less than 100 residues. This suggests that the transmembrane segment, or internal signal anchor, may replace the signal peptide function when the nascent polypeptide is targeted to the ER. The Type III transmembrane proteins, however, are clearly different from the Type II proteins in

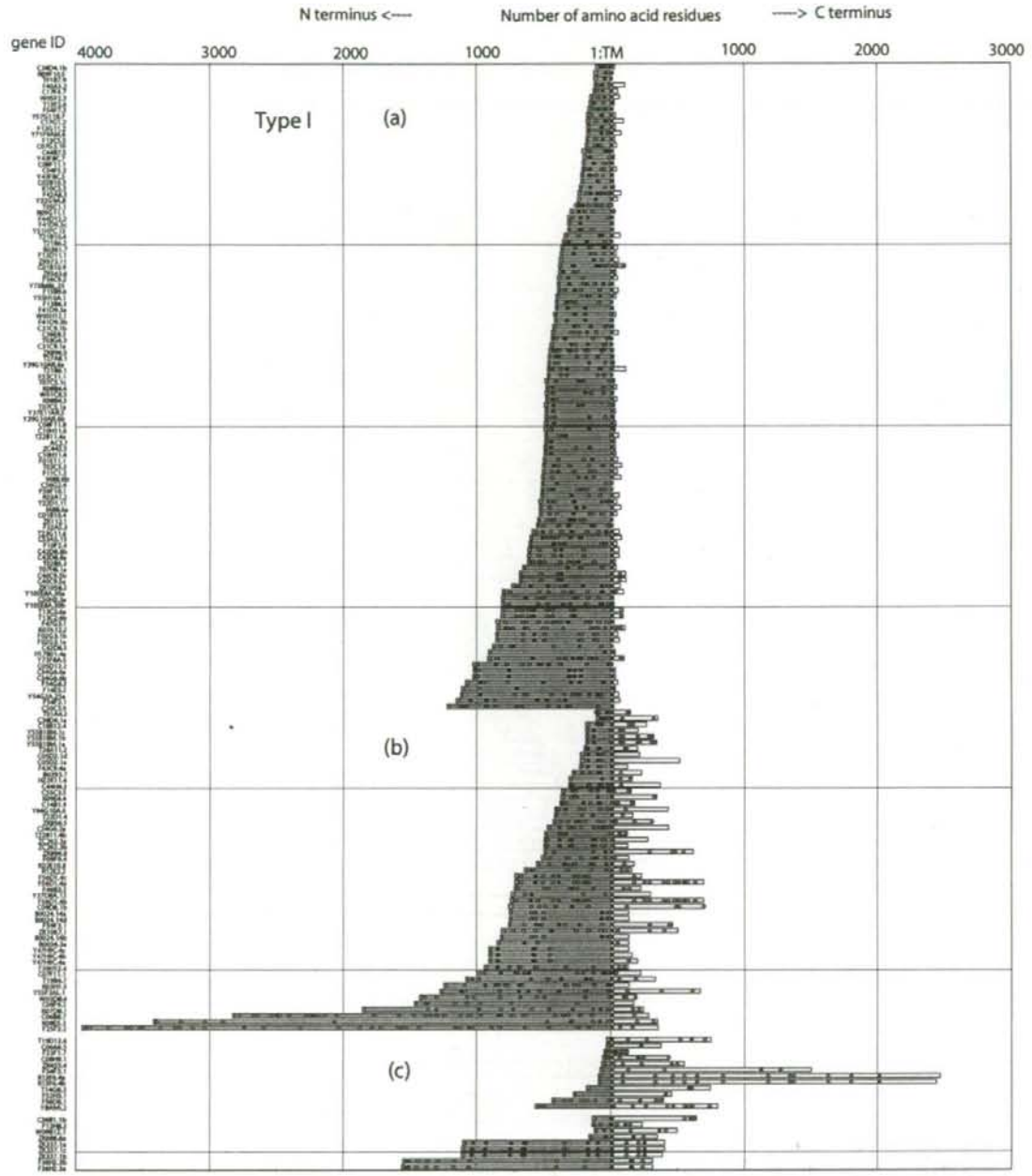


FIG. 2. **Schematic representation of the topology of the putative single span transmembrane glycoproteins that contain a signal peptide sequence.** The putative signal sequence is shown as an *orange box* at the *left end* of each polypeptide, and the transmembrane segment is indicated by a *red box* at position "*1*." N- (*blue*) and C-terminal (*yellow*) segments are to the *left* and *right sides*, respectively, of the transmembrane segment. The positions of potential *N*-glycosylation sites with a consensus sequence Asn-Xaa-(Ser/Thr) (Xaa ≠ Pro) are indicated by *open diamonds* (◇), and among them the experimentally determined *N*-glycosylated sites are indicated by *closed diamonds* (◆). In this figure, the Type I single span transmembrane proteins are tentatively classified into those having a small (<100-residue) or a large (>100-residue) cytosolic segment, indicated by *a* or *b*, respectively. In a number of glycoproteins (*c*), the glycosylated sites are located in the putative C-terminal segment or at both sides of the transmembrane segment probably due to inaccurate prediction of the transmembrane segment.

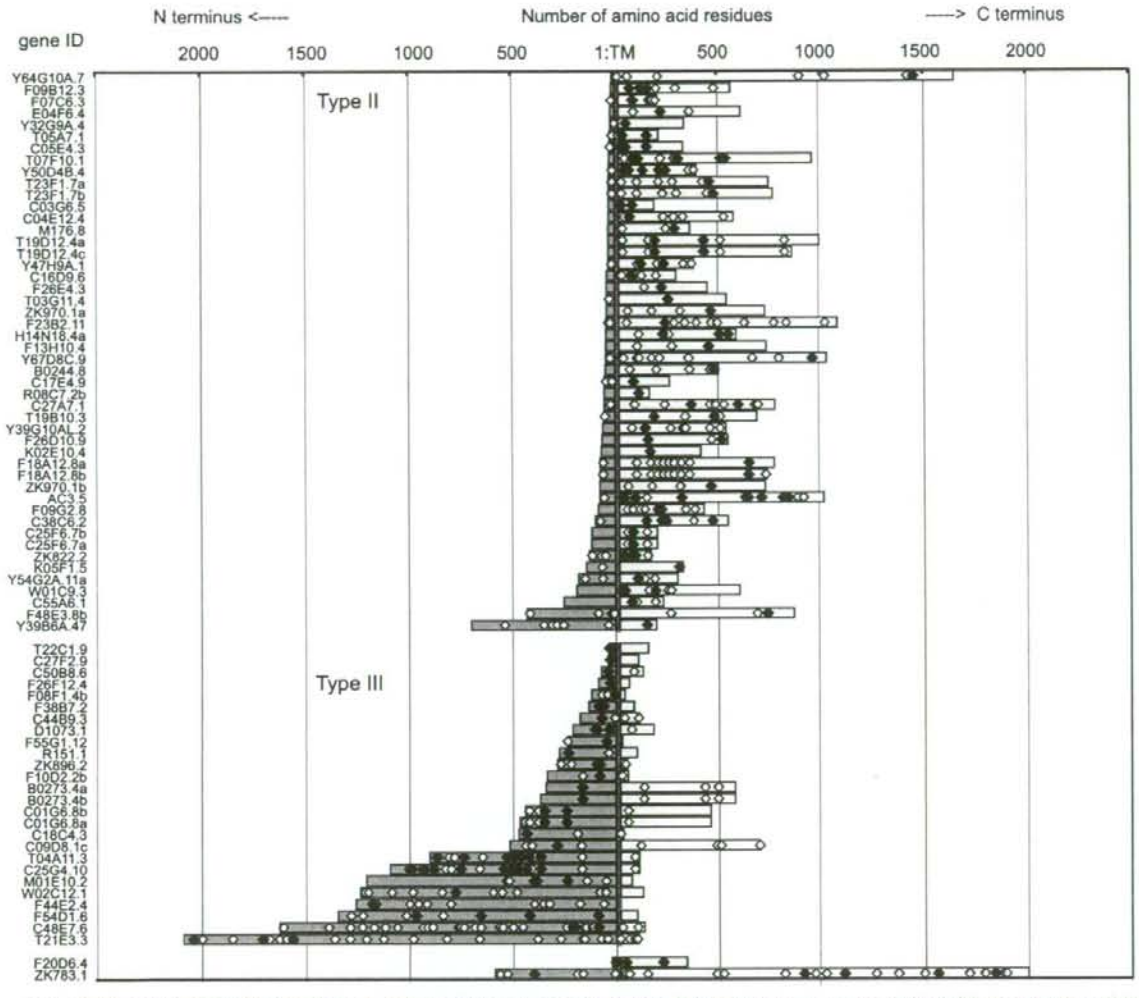


FIG. 3. **Schematic representation of the topology of the putative single span transmembrane glycoproteins lacking a signal peptide sequence.** *Colors* and *symbols* used are as described for Fig. 2. The single span transmembrane glycoproteins are classified as Type II (*upper*) or Type III (*lower*) according to the position of the glycosylated site(s) and the transmembrane segment except for two proteins (see the *bottom* of the figure) that are glycosylated on both sides of the putative transmembrane segment.

their length of the N-terminal segments. Namely Type III proteins have an average of 520 residues between the N terminus and transmembrane segment with the length ranging from 36 to 2,086 residues. Unlike the Type II proteins, 22 of 26 (~85%) Type III proteins have long N-terminal segments often far exceeding 100 residues (Fig. 3). For example, UNC-5 (netrin receptor/axon guidance protein) has a ~350-residue N-terminal segment that consists of an extracellular Ig-like domain (Prosite document PDOC50835 in ExPASy, www.expasy.org) and two thrombospondin domains (Prosite PDOC50092). The C01G6.8 gene product (CAM-1) has an N-terminal ~450-residue segment consisting of the Ig-like, Frizzled (Prosite PDOC50038), and Kringle2 (Prosite PDOC50020) domains upstream of the putative transmembrane segment. Our prediction that UNC-5 and CAM-1 have a Type III topology is reasonable because both proteins have typical intracellular domains, ZU5 (Prosite PDOC51145) and the death (Prosite PDOC50017) domains (UNC-5) or a protein kinase domain (CAM-1), downstream of their putative transmembrane segments.

Conclusion

We identified a series of *N*-linked glycoproteins by collecting multiple subsets of *N*-linked glycopeptides from tryptic

digests of soluble and insoluble fractions of crude extracts of *C. elegans* using three types of lectin columns followed by IGOT-LC-MS/MS analyses of the glycopeptides. The analysis yielded a large dataset of *N*-glycosylated proteins as well as the sites of glycosylation (Supplemental Table 2). Although the IGOT strategy reveals the status of glycosylation of cellular proteins, it also implies the structure of *N*-glycans attached to the polypeptide chain (Supplemental Table 3). The oligosaccharide chain is synthesized without a template, and the final oligosaccharide structure is determined by many factors, such as the types of glycosyltransferases and glycosidases in the cell, enzyme/substrate concentrations, and the primary and higher order structures of the substrate protein, etc. The specific structure of the oligosaccharide chain contributes to the structure/function of particular glycoproteins, and there is an expanding number of human diseases with defects in glycoconjugate assembly and processing (11, 12). This study suggests that OST, a single multienzyme complex in *C. elegans*, recognizes the consensus tripeptide glycosylation sequence as well as additional determinants to introduce an oligosaccharide to an Asn residue; however, it is still difficult to predict which consensus sites will undergo glycosylation and the subsequent processing that yields the final oligosaccharide structure. Clearly more extensive analysis is required to understand the glycosylation machinery of the cell.

Based on the structural characteristics of putative single span transmembrane glycoproteins identified in this study, we propose an atypical translocation mechanism for Type III transmembrane proteins in which newly synthesized polypeptide is post-translationally translocated into the ER through the translocon (Fig. 3). The mechanism underlying transmembrane protein topogenesis remains controversial (38) mainly because definitive structural data on integral membrane proteins are limited. In particular, there have been few naturally occurring Type III proteins reported (*e.g.* synaptotagmin II (38), microsomal cytochrome P450 (39), and yeast Golgi ecto-ATPase Ynd1p (40)), and therefore Type III transmembrane proteins are thought to have unusual topologies. It is generally accepted that the topology of Type II/III proteins is determined by the interaction between the translocon channel and the transmembrane segment of each protein essentially according to the "positive-inside" rule (38, 41). It is also believed that Type III transmembrane proteins have a relatively short N-terminal segment preceding the signal anchor or transmembrane sequence. This is due to the fact that the translocation of this segment is a post-translational event, and thus it must retain an unfolded structure to pass through the translocon. Nevertheless our results indicate that Type III transmembrane proteins are distributed much more widely among eukaryotic cells than previously thought and that they have an apparently longer N-terminal segment compared with Type II proteins. The mechanism of the generation of Type III proteins must therefore involve either a chaperonin-like activity (to maintain the long N-terminal polypeptide segment in an unfolded structure) or the unfolding of the tentatively folded structure prior to translocation. It is also likely that an energy-dependent mechanism is required to insert the nascent protein into the translocon of the ER. Our analysis for the Type III proteins assigned here indicates that many have an N-terminal structure stabilized by disulfide bridges, such as Ig-like domain, Frizzled domain, Kringle2, FN3 (fibronectin type III), EGF3 (epidermal growth factor), Sushi, LDLRA2 (LDL-receptor class A), LDLRB (class B), and AMOP (adhesion-associated domain in MUC4 and other proteins) domains. We assume that these domains retain a relatively flexible, unfolded structure in the reducing environment of the cytosol that might be advantageous for the post-translational translocation of the polypeptide into the ER. Thus, our findings suggest that the Type III transmembrane protein architecture is widespread in eukaryotic cells. Further studies of the mechanism that determines the topogenesis of integral membrane proteins will assess the validity of these predictions.

Acknowledgment—We thank Dr. J. Hirabayashi (National Institute of Advanced Industrial Science and Technology (AIST), Ibaraki, Japan) for the expression vector encoding worm galectin 6.

* This work was supported in part by grants for the Structural Glycomics Project from the New Energy and Industrial Technology Development Organization (NEDO) of Japan and for the Integrated Proteomics Project, Pioneer Research on Genome the Frontier from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan, and by a grant-in-aid for scientific research from MEXT. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

§ To whom correspondence should be addressed: Research Center for Medical Glycoscience, National Inst. of Advanced Industrial Science and Technology (AIST), Central 2-12, Umezono 1-1-1, Tsukuba, Ibaraki 305-8568, Japan. Tel.: 81-29-861-3187; Fax: 81-29-861-3125; E-mail: kaji-rcmg@aist.go.jp.

REFERENCES

1. Mann, M., and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **21,** 255–261
2. Ballif, B. A., Villén, J., Beausoleil, S. A., Schwartz, D., and Gygi, S. P. (2004) Phosphoproteomic analysis of the developing mouse brain. *Mol. Cell. Proteomics* **3,** 1093–1101
3. Gruhler, A., Olsen, J. V., Mohammed, S., Mortensen, P., Færgeman, N. J., Mann, M., and Jensen, O. N. (2005) Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* **4,** 310–327
4. Liu, T., Qian, W.-J., Gritsenko, M. A., Camp, D. G., II, Monroe, M. E., Moore, R. J., and Smith, R. D. (2005) Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. *J. Proteome Res.* **4,** 2070–2080
5. Liu, T., Qian, W.-J., Gritsenko, M. A., Xiao, W., Moldawer, L. L., Kaushal, A., Monroe, M. E., Varnum, S. M., Moore, R. J., Purvine, S. O., Maier, R. V., Davis, R. W., Tompkins, R. G., Camp, D. G., II., Smith, R. D., and the Inflammation and the Host Response to Injury Large Scale Collaborative Research Program (2006) High dynamic range characterization of the trauma patient plasma proteome. *Mol. Cell. Proteomics* **5,** 1899–1913
6. Wang, L., Li, F., Sun, W., Wu, S., Wang, X., Zhang, L., Zheng, D., Wang, J., and Gao, Y. (2006) Concanavalin A-captured glycoproteins in healthy

human urine. *Mol. Cell. Proteomics* **5,** 560–562
7. Matsumoto, M., Hatakeyama, S., Oyamada, K., Oda, Y., Nishimura, T., and Nakayama, K. I. (2005) Large-scale analysis of the human ubiquitin-related proteome. *Proteomics* **5,** 4145–4151
8. Wykoff, D. D., and O'Shea, E. K. (2005) Identification of sumoylated proteins by systematic immunoprecipitation of the budding yeast proteome. *Mol. Cell. Proteomics* **4,** 73–83
9. Krishna, R. G., and Wold, F. (1993) Post-translational modification of proteins. *Adv. Enzymol. Relat. Areas Mol. Biol.* **67,** 265–298
10. Apweiler, R., Hermjakob, H., and Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* **1473,** 4–8
11. Freeze, H. H., and Aebi, M. (2005) Altered glycan structures: the molecular basis of congenital disorders of glycosylation. *Curr. Opin. Struct. Biol.* **15,** 490–498
12. Sturiale, L., Barone, R., Fiumara, A., Perez, M., Zaffanello, M., Sorge, G., Pavone, L., Tortorelli, S., O'Brien, J. F., Jaeken, J., and Garozzo, D. (2005) Hypoglycosylation with increased fucosylation and branching of serum transferrin N-glycans in untreated galactosemia. *Glycobiology* **15,** 1268–1276
13. Moloney, D. J., Panin, V. M., Johnston, S. H., Chen, J., Shao, L., Wilson, R., Wang, Y., Stanley, P., Irvine, K. D., Haltiwanger, R. S., and Vogt, T. F. (2000). Fringe is a glycosyltransferase that modifies Notch. *Nature* **406,** 369–375
14. Zhang, H., Li, X. J., Martin, D. B., and Aebersold, R. (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **21,** 660–666
15. Ramachandran, P., Boontheung, P., Xie, Y., Sondej, M., Wong, D. T., and Loo, J. A. (2006). Identification of N-linked glycoproteins in human saliva by glycoprotein capture and mass spectrometry. *J. Proteome Res.* **5,** 1493–1503
16. Kaji, H., Saito, H., Yamauchi, Y., Shinkawa, T., Taoka, M., Hirabayashi, J., Kasai, K., Takahashi, N., and Isobe, T. (2003) Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat. Biotechnol.* **21,** 667–672
17. Bunkenborg, J., Pilch, B. J., Podtelejnikov, A. V., and Wisniewski, J. R. (2004). Screening for N-glycosylated proteins by liquid chromatography mass spectrometry. *Proteomics* **4,** 454–465
18. Lewandrowski, U., Moebius, J., Walter, U., and Sickmann, A. (2006) Elucidation of N-glycosylation sites on human platelet proteins. A glycoproteomic approach. *Mol. Cell. Proteomics* **5,** 226–233
19. Gonzalez, J., Takao, T., Hori, H., Besada, V., Rodriguez, R., Padron, G., and Shimonishi, Y. (1992) A method for determination of N-glycosylation sites in glycoproteins by collision-induced dissociation analysis in fast atom bombardment mass spectrometry: identification of the positions of carbohydrate-linked asparagine in recombinant alpha-amylase by treatment with peptide-N-glycosidase F in ^{18}O-labeled water. *Anal. Biochem.* **205,** 151–158
20. Hirabayashi, J., Ubukata, T., and Kasai, K. (1996) Purification and molecular characterization of a novel 16-kDa galectin from the nematode *Caenorhabditis elegans*. *J. Biol. Chem.* **271,** 2497–2505
21. Mawuenyega, K. G., Kaji, H., Yamauchi, Y., Shinkawa, T., Saito, H., Taoka, M., Takahashi, N., and Isobe, T. (2003) Large-scale identification of Caenorhabditis elegans proteins by multidimensional liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2,** 23–35
22. Taoka, M., Yamauchi, Y., Shinkawa, T., Kaji, H., Motohashi, W., Nakayama, H., Takahashi, N., and Isobe, T. (2004) Only a small subset of the horizontally transferred chromosomal genes in *Escherichia coli* are translated into proteins. *Mol. Cell. Proteomics* **3,** 780–787
23. Wada, Y., Tajiri, M., and Yoshida, S. (2004) Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem.* **76,** 6560–6565
24. Ishida, Y., Fujita, T., and Asai, K. (1981) New detection and separation method for amino acids by high-performance liquid chromatography. *J. Chromatogr.* **204,** 143–148
25. Isobe, T., Yamauchi, Y., Taoka, M., and Takahashi, N. (2003) Automated two-dimensional liquid chromatography/tandem mass spectrometry for large-scale protein analysis, in *Proteins and Proteomics* (Simpson, R. D., ed) pp. 869–876, Cold Spring Harbor Press, Cold Spring Harbor, NY
26. Natsume, T., Yamauchi, Y., Nakayama, H., Shinkawa, T., Yanagida, M., Takahashi, N., and Isobe, T. (2002) A direct nanoflow liquid chromatography-tandem mass spectrometry system for interaction proteomics. *Anal. Chem.* **74,** 4725–4733
27. Shinkawa, T., Taoka, M., Yamauchi, Y., Ichimura, T., Kaji, H., Takahashi, N., and Isobe, T. (2005) STEM: a software tool for large-scale proteomic data analyses. *J. Proteome Res.* **4,** 1826–1831
28. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340,** 783–795
29. Arai, M., Mitsuke, H., Ikeda, M., Xia, J.-X., Kikuchi, T., Satake, M., and Shimizu, T. (2004) ConPred II: a consensus prediction method for obtaining transmembrane topology models with high reliability. *Nucleic Acids Res.* **32,** W390–W393
30. Natsuka, S., Adachi, J., Kawaguchi, M., Nakakita, S., Hase, S., Ichikawa, A., and Ikura, K. (2002) Structural analysis of *N*-linked glycans in *Caenorhabditis elegans*. *J. Biochem.* **131,** 807–813
31. Hirabayashi, J., Hayama, K., Kaji, H., Isobe, T., and Kasai, K. (2002) Affinity capturing and gene assignment of soluble glycoproteins produced by the nematode *Caenorhabditis elegans*. *J. Biochem.* **132,** 103–114
32. Natsuka, S. (2005) Comparative biochemical view of *N*-glycans. *Trends Glycosci. Glycotechnol.* **17,** 229–236
33. Yan, A., and Lennarz, W. J. (2005). Two oligosaccharyl transferase complexes exist in yeast and associate with two different translocons. *Glycobiology* **15,** 1407–1415
34. Spirig, U., Bodmer, D., Wacker, M., Burda, P., and Aebi, M. (2005) The 3.4 kDa Ost4 protein is required for the assembly of two distinct oligosaccharyltransferase complexes in yeast. *Glycobiology* **15,** 1396–1406
35. Kelleher, D. J., Karaoglu, D., Mandon, E. C., and Gilmore, R. (2003) Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. *Mol. Cell* **12,** 101–111
36. Walter, P., and Johnson, A. E. (1994). Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* **10,** 87–119
37. Goder, V., and Spiess, M. (2001) Topogenesis of membrane proteins: determinants and dynamics. *FEBS Lett.* **504,** 87–93
38. Kida, Y., Mihara, K., and Sakaguchi, M. (2005) Translocation of a long amino-terminal domain through ER membrane by following signal-anchor sequence. *EMBO J.* **24,** 3202–3213
39. Sakaguchi, M., Mihara, K., and Sato, R. (1987) A short amino-terminal segment of microsomal cytochrome P-450 functions both as an insertion signal and as a stop-transfer sequence. *EMBO J.* **6,** 2425–2431
40. Zhong, X., Malhotra, R., and Guidotti, G. (2005) A eukaryotic carboxyl-terminal signal sequence translocating large hydrophilic domains across membranes. *FEBS Lett.* **579,** 5643–5650
41. Kida, Y., Sakaguchi, M., Fukuda, M., Mikoshiba, K., and Mihara, K. (2000) Membrane topogenesis of a type I signal-anchor protein, mouse synaptotagmin II, on the endoplasmic reticulum. *J. Cell Biol.* **150,** 719–730

Journal of Cellular Physiology

M-CSF-Mediated Macrophage Differentiation but not Proliferation Is Correlated with Increased and Prolonged ERK Activation

SHINYA SUZU,[1] MASATERU HIYOSHI,[1] YUKA YOSHIDOMI,[1] HIDEKI HARADA,[1] MOTOHIRO TAKEYA,[2] FUMIHIKO KIMURA,[3] KAZUO MOTOYOSHI,[3] AND SEIJI OKADA[1]*

[1]*Division of Hematopoiesis, Center for AIDS Research, Kumamoto University, Kumamoto, Japan*

[2]*Department of Cell Pathology, Graduate School of Medical and Pharmaceutical Sciences, Kumamoto University, Kumamoto, Japan*

[3]*Third Department of Internal Medicine, National Defense Medical College, Saitama, Japan*

M-CSF is a cytokine essential for both the proliferation and differentiation of monocytes/macrophages. In this study, we established a new M-CSF-mediated differentiation-inducing system, and examined how the level and duration of the activation of ERK preceded M-CSF-mediated differentiation. TF-1-fms human leukemia cells rapidly proliferated in response to M-CSF. However, in the presence of a phorbol ester, TPA, TF-1-fms cells definitely switched their responsiveness to M-CSF from proliferation to differentiation, as evidenced by a more drastic morphological change and the appearance of cells with a higher level of phagocytic activity. In TF-1-fms cells expressing HIV-1 Nef protein in a conditionally active-manner, both M-CSF-mediated proliferation and M-CSF/TPA-mediated differentiation were inhibited by the activation of Nef. The Nef-active cells showed perturbed patterns of ERK activation. Under the proliferation-inducing conditions (TPA-free), parental or Nef-inactive cells showed modest ERK activation following M-CSF stimulation, whereas Nef-active cells showed an earlier and transient ERK activation, despite a decrease in their proliferation rate. Under the differentiation-inducing conditions, parental or Nef-inactive cells showed increased and prolonged ERK activation following M-CSF stimulation, whereas Nef-active cells showed transient ERK activation. These results supported the idea that the increased and prolonged ERK activation led to M-CSF-mediated macrophage differentiation but not to proliferation.

J. Cell. Physiol. 212: 519–525, 2007.

Macrophage colony-stimulating factor (M-CSF) is a cytokine that supports both proliferation and differentiation of the cells of monocytic lineage (Roth and Stanley, 1992). The biological effects of M-CSF are mediated by a receptor tyrosine kinase, Fms (Sherr et al., 1985). The binding of M-CSF leads to the autophosphorylation of tyrosine residues in the cytoplasmic domain of Fms and subsequent interactions of the phosphorylated residues with other proteins, resulting in the initiation of multiple pathways (Bourette and Rohrschneider, 2000). Thus, to investigate which pathway leads to the proliferation or differentiation, mutant Fms proteins in which the tyrosine residues are substituted with phenylalanine have been generated and expressed in various cell types, such as NIH3T3 fibroblasts (Roussel et al., 1990), Rat-2 fibroblasts (van der Geer and Hunter, 1993), FDC-P1 myeloid progenitor cells (Bourette et al., 1995) and M1 myeloid cells (Marks et al., 1999). However, because the pathway that is predominantly utilized and whether cells proliferate or differentiate in response to M-CSF depend on the cell type, the exact differences in the signaling events between these distinct cellular responses to M-CSF are still unclear. Culture systems in which the same cells distinctly respond to different stimuli (proliferation versus differentiation) are useful for clarifying this issue. For example, rat neuronal PC12 cells proliferate in response to EGF, whereas stimulation with NGF causes neuronal differentiation (Marshall, 1995). Of importance, studies with this culture system have revealed that the increased and prolonged activation of extracellular signal-regulated kinase (ERK) is critical for the neuronal differentiation of PC12 cells, but not for proliferation (Marshall, 1995). It has been also shown that the increased and prolonged activation of ERK is critical for the differentiation of megakaryocytes (Melemed et al., 1997; Racke et al., 1997) and muscle cells (Gredinger et al., 1998). Furthermore, the functional maturation of macrophages induced by lipopolysaccharide seemed to require increased and prolonged ERK activation (Valledor et al., 2000). However, the role of ERK activation in M-CSF-mediated macrophage differentiation is not fully understood.

In this study, we first attempted to establish a new M-CSF-mediated differentiation-inducing system, using human leukemia TF-1-fms cells which essentially showed a proliferative response to M-CSF (Suzu et al., 1997). We assessed whether 12-O-tetradecanoylphorbol 13-acetate (TPA) triggered their differentiation and M-CSF accelerated the process. TPA has been characterized by its ability to induce the differentiation of several leukemia cell lines (Kitamura et al., 1989; Racke et al., 1997; He et al., 1999). We next attempted to clarify whether

Contract grant sponsor: Ministry of Health, Labour and Welfare of Japan;
Contract grant number: H16-AIDS-003.

*Correspondence to: Seiji Okada, Division of Hematopoiesis, Center for AIDS Research, Kumamoto University, Honjo 2-2-1, Kumamoto-city, Kumamoto 860-0811, Japan.
E-mail: okadas@gpo.kumamoto-u.ac.jp

Received 15 June 2006; Accepted 10 January 2007

DOI: 10.1002/jcp.21045

the increased and prolonged activation of ERK preceded M-CSF-mediated differentiation, by utilizing TF-1-fms cells expressing a conditionally active HIV-1 Nef protein (Suzu et al., 2005). Previously, we showed that Nef activation inhibited the M-CSF-mediated proliferation of TF-1-fms cells but enhanced the activation of ERK following M-CSF treatment (Suzu et al., 2005). Nef is a major determinant of the pathogenicity of HIV-1 (Fackler and Baur, 2002; Peterlin and Trono, 2003; Qiao et al., 2006), and has been shown to bind to and activate Hck, a Src kinase (Saksela et al., 1995; Moarefi et al., 1997). The activation of Hck was one possible molecular mechanism by which Nef caused the inhibition of M-CSF-mediated cell proliferation (Suzu et al., 2005) and the enhancement of ERK activation (Schrager et al., 2002; He et al., 2004). Based on these findings, we carefully examined how Nef affected the M-CSF/TPA-mediated differentiation of TF-1-fms cells and the level/duration of ERK activation in the differentiation-inducing conditions.

Materials and Methods

Cell culture and reagents

TF-1 cells (Kitamura et al., 1989) were routinely cultured with RPMI 1640 medium (Sigma, St. Louis, MO) — 10% fetal calf serum (FCS) in the presence of recombinant human granulocyte/macrophage-CSF (GM-CSF) (2 ng/ml; PeproTech, London, UK). TF-1-fms cells (Suzu et al., 1997) were maintained with RPMI 1640 — 10% FCS in the presence of M-CSF (100 ng/ml; a gift from Morinaga Milk Industry, Kanagawa, Japan) and G418 (200 μg/ml; Life Technologies, Grand Island, NY). TF-1-fms cells expressing the HIV-1 Nef-murine estrogen receptor hormone-binding domain (Nef-ER) fusion protein (TF-1-fms-Nef-ER) (Walk et al., 2001; Suzu et al., 2005) were maintained with RPMI 1640 — 10% FCS containing M-CSF, G418, and puromycin (1.5 μg/ml; Sigma). In this system, Nef was basally inactive but its function was inducibly activated by the estrogen analogue, 4-hydroxytamoxifen (4-HT; Sigma) (Suzu et al., 2005). In this study, we also established TF-1 cells expressing the Nef-ER fusion protein with the plasmid pEBB-Nef-ER-IRES-puro (Walk et al., 2005). TF-1-Nef-ER cells were maintained with RPMI 1640 — 10% FCS containing GM-CSF and puromycin. To activate Nef, 4-HT was added to the culture at a final concentration of 1 μM. TPA (Sigma) was added to the culture at a final concentration of 100 ng/ml. PD98059 (ERK kinase inhibitor) and PP2 (Src kinase inhibitor) were purchased from Sigma. U0126 was purchased from Calbiochem (San Diego, CA).

Cell count and viability analysis

The viable cell counts were obtained by enumerating the cells that excluded trypan blue dye on a hemocytometer. The adherent cells were harvested by trypsinization. The viability of cells was also examined with the propidium iodide (PI) exclusion method (Okada et al., 1998). Cells were suspended in phosphate-buffered saline (PBS) containing 0.1% NaN_3, 3% FCS and 2 μg/ml PI. The uptake of PI in each cell was analyzed with a FACSCalibur using Cell Quest Software (Becton Dickinson, Mountain View, CA).

Analyses of expression of Fms and CD204, and phagocytic activity

The cell surface expression of Fms was analyzed by flow cytometry using Flag-tagged M-CSF (Suzu et al., 2005). In brief, cells were incubated with the Flag-tagged M-CSF followed by biotin-labeled anti-Flag M2 antibody (Sigma) and phycoerythrin (PE)-labeled streptavidin (PharMingen, San Jose, CA). The analyses were performed with a FACSCalibur. The cell surface expression of CD204 was determined using anti-CD204-FITC (clone E-5) (Tomokiyo et al., 2002). The phagocytic activity was determined by measuring the uptake of fluorescent microspheres (Fluoresbrite Carboxylate Microspheres, 0.7 μm in diameter, Polysciences, Warrington, PA). Cells cultured on a 6-well tissue culture plate were incubated with the fluorescent microspheres for 5 h and washed with PBS. The cells showing phagocytized particles were analyzed by flow cytometry.

Immunoprecipitation and Western blotting

The immunoprecipitation and Western blotting were performed essentially as described previously (Suzu et al., 2000, 2005). Cells were depleted of M-CSF for 14 h in RPMI 1640 — 10% FCS with or without TPA, and then stimulated with M-CSF for the indicated periods. In selected experiments, 4-HT was added to the culture at the initiation of M-CSF deprivation/TPA pre-treatment. Then, the cells were solubilized with the Nonidet P-40 lysis buffer. The immunoprecipitation was performed with anti-phosphotyrosine mouse IgG conjugated to agarose (PY99; Santa Cruz Biotechnology, Santa Cruz, CA). The antibodies (purchased from Santa Cruz) used for Western blotting were as follows: anti-Fms (C-20), anti-phosphotyrosine (PY99), anti-ERK (K-23), and anti-phosphorylated ERK (E-4). The rabbit antiserum to Nef was obtained from the NIH AIDS Research and Reference Reagent Program (Division of AIDS, NIAID, NIH, Bethesda, MD). The detection was performed using the ECL system (Amersham, Buckinghamshire, UK). The relative intensity of bands on scanned gel images was quantified by using the NIH Image software.

Results

Proliferative- and differentiative properties of TF-1-fms cells

TF-1-fms cells were derived from TF-1 cells by introducing the wild-type *c-fms* gene into the parental cells (Fig. 1A). The proliferation of TF-1 and TF-1-fms was entirely dependent on GM-CSF and M-CSF, respectively (Fig. 1B) (Kitamura et al., 1989; Suzu et al., 1997). TF-1-fms cells lost their responsiveness to GM-CSF, due to a loss of the expression of the GM-CSF receptor α chain (data not shown). In the presence of M-CSF, TF-1-fms cells neither adhered to dishes (Fig. 1C) nor showed a differentiated morphology (Fig. 1D, top part). On the other hand, when cultured with media containing TPA alone, the cells tended to adhere to dishes and show a flattened morphology (Figs. 1C and D, middle part). Of note, M-CSF accelerated the differentiation-like process. Most TF-1-fms cells cultured in the presence of both M-CSF and TPA adhered to dishes and showed a mature macrophage-like morphology (Figs. 1C and D, bottom part), which closely resembled the primary macrophages obtained by culturing human peripheral blood monocytes with M-CSF (Hashimoto et al., 1999). In parallel with the morphological change, these cells showed more of an increase in granularity than cells cultured with TPA alone (Fig. 1E). More importantly, these cells showed higher phagocytic activity: the percentage of cells phagocytizing the microbeads and their mean fluorescence intensity (MFI) were higher in the culture containing M-CSF and TPA than in the culture containing TPA alone (Fig. 1F, left part). In addition, these cells showed a higher level of CD204 (class A macrophage scavenger receptor) (Fig. 1F, right part). These results indicated that the cells obtained by culturing TF-1-fms cells with both M-CSF and TPA were functionally mature macrophages.
As mentioned above and shown in the upper panel of Figure 2A, the culture of TF-1-fms cells with TPA resulted in the appearance of adherent cells and the addition of M-CSF further increased the number of adherent cells. However, there was no significant difference in total number of viable cells between the two treatments (TPA alone versus M-CSF + TPA) (Fig. 2A, lower part). Moreover, there was no significant difference in the percentage of PI-positive dead cells between the two treatments (Fig. 2B), excluding the possibility that the acceleration of TF-1-fms cell differentiation by M-CSF and TPA reflected a survival-enhancing or anti-apoptotic function of M-CSF. It was therefore likely that TF-1-fms cells definitely switched their responsiveness to M-CSF from proliferation to macrophage differentiation in the presence of TPA.

Requirement of ERK activation for M-CSF-mediated proliferation and M-CSF/TPA-mediated differentiation of TF-1-fms cells

We next examined whether the activation of ERK was required for M-CSF-mediated proliferation and M-CSF/TPA-mediated

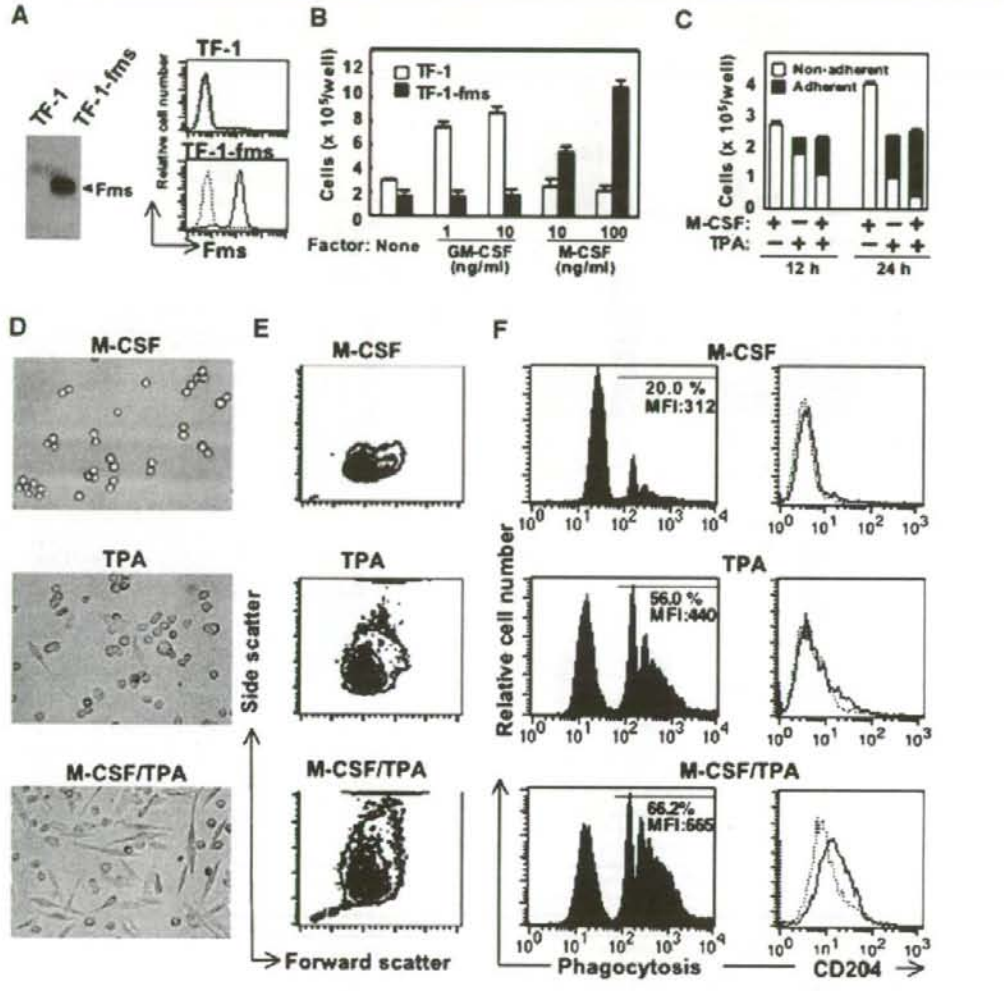


Fig. 1. The expression of Fms in and proliferative and differentiative properties of TF-1-fms cells. A: The total cell lysates from TF-1 or TF-1-fms cells were analyzed for the expression of Fms by Western blotting. Alternatively, the level of cell surface Fms expression was analyzed by flow cytometry with Flag-tagged M-CSF (solid lines). The profiles of cells incubated with a Flag-tagged protein, which is unrelated to M-CSF, are also shown as a control (broken lines). B: TF-1 or TF-1-fms cells were seeded into 6-well culture plates at a density of 5×10^4 cells/ml in the absence or presence of the indicated concentrations of cytokines. TF-1 and TF-1-fms cells were cultured for 3 and 2 days, respectively. After the cultures, viable cells were enumerated. Error bars from triplicate assays are shown. These results are representative of two independent experiments. C: TF-1-fms cells were seeded at a density of 1×10^5 cells/ml, and cultured in the presence of M-CSF, TPA, or both. After culturing for 12 or 24 h, cells adhering to the dishes and non-adherent cells were enumerated. D–F: TF-1-fms cells were seeded at a density of 1×10^5 cells/ml, and cultured for 2 days in the presence of M-CSF (M-CSF), TPA (TPA), or both (M-CSF/TPA). D: The morphology of cells after culturing is shown. The cells cultured with M-CSF alone were photographed after a 5-fold dilution with media. E: The results of flow cytometric analyses of cells for forward and side scatters are shown. An equal number of cells were analyzed and the results are presented as counter plots. F: The phagocytic activity of cells was determined by the procedures described in the Materials and Methods (left panels). The percentage and mean fluorescence intensity (MFI) of cells in the region indicated by solid lines are shown. Alternatively, the cells were analyzed for the expression of CD204 by flow cytometry (right parts).

differentiation of TF-1-fms cells, using pharmacological inhibitors. As shown, PP2 (the inhibitor specific for Src kinases), PD98059 (the inhibitor specific for ERK kinase, MEK), and U0126 (another MEK inhibitor) significantly reduced the rate of M-CSF-mediated proliferation (Fig. 3A). The reduced proliferation rate correlated well with the increase in the percentage of PI-positive dead cells (Fig. 3A). Thus, we could not exclude the possibility that the inhibitory effect of these inhibitors on M-CSF-mediated proliferation reflected their cytotoxicity. Of importance, however, these inhibitors, in

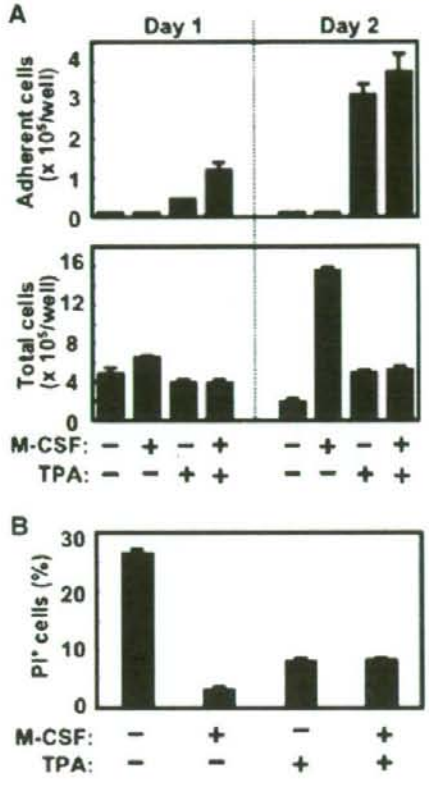


Fig. 2. A comparison of cell viability/death between the culture with TPA alone and that with M-CSF plus TPA. A and B: TF-1-fms cells were seeded into 6-well culture plates at a density of 1×10^5 cells/ml. Then, the cells were cultured in the absence of additives, or the presence of M-CSF, TPA, or both. A: After culturing for 1 or 2 days, cells adhering to the dishes (upper part) and all viable cells (lower part) were enumerated. B: After culturing for 2 days, the percentage of PI-positive dead cells in the wells was determined by flow cytometry. Error bars from triplicate assays are shown. These results are representative of two independent experiments.

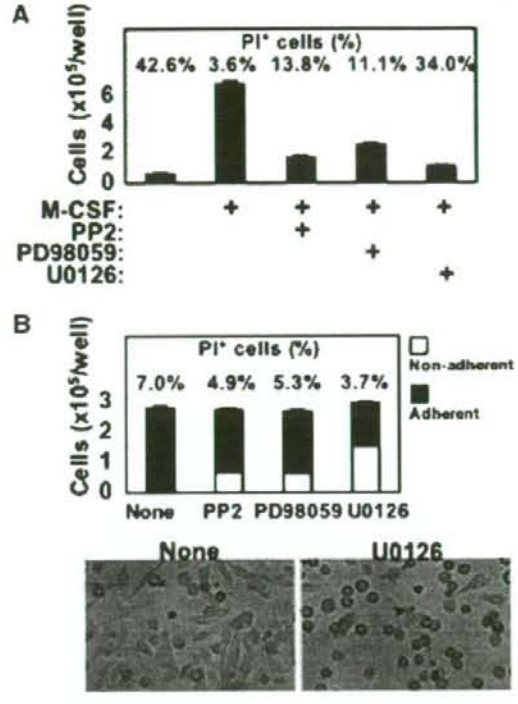


Fig. 3. Effects of pharmacological inhibitors on the proliferation and differentiation of TF-1-fms cells. A: TF-1-fms cells were seeded into 6-well culture plates at a density of 5×10^4 cells/ml in the absence or presence of the indicated inhibitors. M-CSF was added at a final concentration of 100 ng/ml. Both PP2 and U0126 were added at a final concentration of 10 μM. PD98059 was used at a final concentration of 50 μM. After culturing for 2 days, viable cells were enumerated. Error bars from triplicate assays are shown. Simultaneously, the percentage of PI-positive dead cells was determined by flow cytometry. These results are representative of two independent experiments. B: TF-1-fms cells were suspended at 1×10^5 cells/ml in medium containing M-CSF and TPA, and then cultured for 1 day in the absence or presence of the indicated inhibitors. The cells adhering to the dishes and non-adherent cells were enumerated. Error bars from triplicate assays are shown. Simultaneously, the percentage of PI-positive dead cells was determined by flow cytometry. These results are representative of two independent experiments. The morphology of cells after culturing is shown.

particular U0126, significantly inhibited the differentiation of TF-1-fms cells induced by M-CSF and TPA without increasing the percentage of PI-positive dead cells (Fig. 3B). These results indicated that the differentiation of TF-1-fms cells was dependent on the activation of ERK.

The level and duration of ERK activation in parental TF-1-fms cells

We next examined whether the increased and/or prolonged activation of ERK preceded the differentiation. We pretreated TF-1-fms cells with TPA or left them untreated, stimulated then with M-CSF or left them un-stimulated, and analyzed the activation of ERK by using an antibody specific for phosphorylated ERK (Fig. 4). Consistent with an earlier report (He et al., 1999), the treatment of TF-1-fms cells with TPA led to the ERK activation (Figs. 4A and B). Following M-CSF stimulation, a number of molecules were shown to be rapidly tyrosine-phosphorylated in both TPA-pretreated cells and untreated cells (Fig. 4A). The most prominent band at 150–160 kD, the tyrosine-phosphorylation level of which was reduced in the pretreated cells, seemed to be Fms (Suzu et al., 2005). The reduction of Fms phosphorylation in the pretreated cells was likely to be due to the reduced cell surface expression of Fms, because the pretreatment caused an increase in the intracellular form of Fms (gp130) and concomitant decrease in the cell surface form of Fms (gp150) (Fig. 4A), as seen in TPA-treated p388D1 macrophages (Wilhelmsen and van der Geer, 2004). Despite the reduced Fms phosphorylation, the level of ERK activation following M-CSF stimulation in TPA-pretreated cells was higher than that in untreated cells (Figs. 4A and B). Although the higher level of ERK activation following M-CSF stimulation seemed to be due to the high baseline level (Fig. 4A, the bar graph), the extent of the activation apparently correlated with the differentiation of TF-1-fms cells.

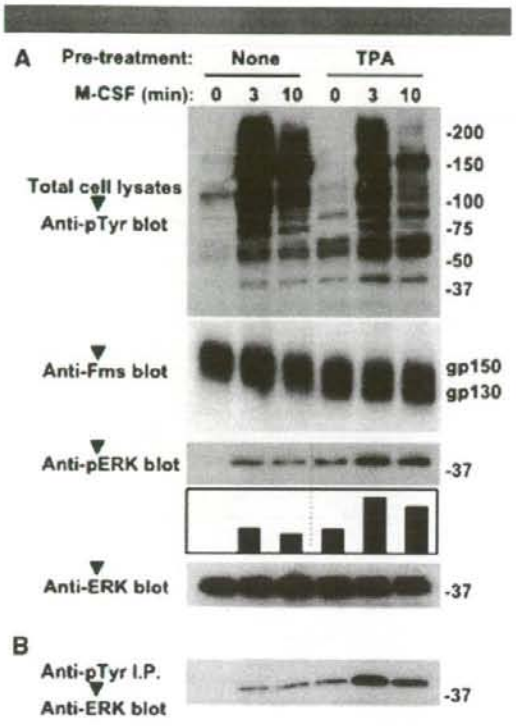


Fig. 4. The level of phosphorylation of ERK in TF-1-fms cells after pre-treatment with TPA and re-stimulation with M-CSF. TF-1-fms cells were deprived of M-CSF with or without TPA for 14 h and then re-stimulated with M-CSF for the indicated periods. A: Total cell lysates were analyzed with Western blotting using an antibody specific for phosphotyrosine (pTyr), Fms, phosphorylated ERK (pERK), or total ERK. The bottom blot was to show that comparable amounts of proteins were loaded in the upper blots. The relative intensities of bands on scanned gel images were quantified using the NIH Image software and the level of phosphorylated ERK normalized to the total amount of ERK is shown in the bar graph. B: The immunoprecipitates with anti-pTyr antibody were analyzed using anti-ERK antibody.

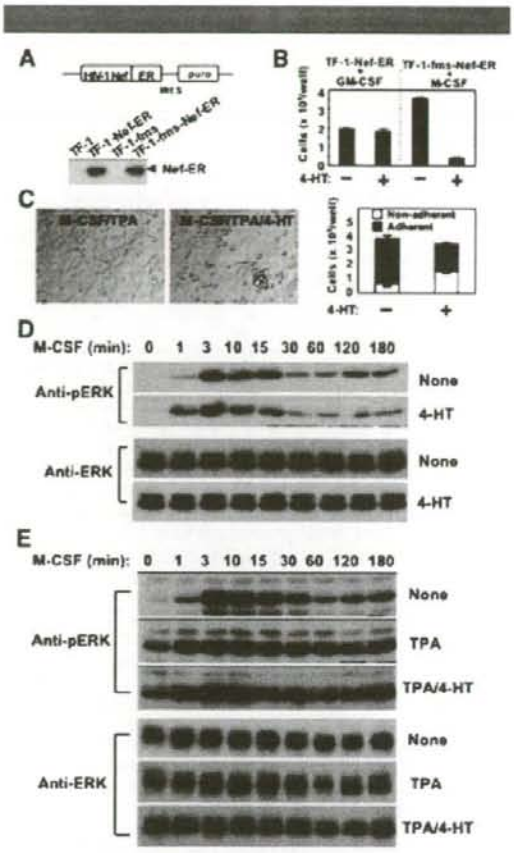


Fig. 5. Establishment of cells expressing the conditionally active Nef, proliferative/differentiative properties, and ERK activation. A: Schematic diagram of the Nef-ER-IRES-puro construct. ER, estrogen receptor hormone-binding domain; IRES, internal ribosomal entry sequence; puro, puromycin resistance gene. Total cell lysates from parental cells (TF-1 and TF-1-fms) or cells stably expressing Nef-ER (TF-1-Nef-ER and TF-1-fms-Nef-ER) were analyzed for the expression of Nef-ER by Western blotting with anti-Nef antibody. B: TF-1-Nef-ER cells were seeded at a density of 2×10^4 cells/ml and cultured in the presence of GM-CSF alone or GM-CSF plus 4-HT for 4 days. Similarly, TF-1-fms-Nef-ER cells were seeded at a density of 1×10^4 cells/ml and cultured in the presence of M-CSF alone or M-CSF plus 4-HT for 4 days. After the cultures, viable cells were enumerated. Error bars from triplicate assays are shown. These results are representative of two independent experiments. C: TF-1-fms-Nef-ER cells were suspended in medium containing M-CSF and TPA, and then cultured for 2 days in the absence (M-CSF/TPA) or the presence of 4-HT (M-CSF/TPA/4-HT). The cells adhering to the dishes and non-adherent cells were enumerated. Error bars from triplicate assays are shown. These results are representative of two independent experiments. D: TF-1-fms-Nef-ER cells were deprived of M-CSF for 14 h and then re-stimulated with M-CSF for the indicated periods. 4-HT was added to the culture at the beginning of the period of deprivation. Total cell lysates were prepared and analyzed by Western blotting using the antibody specific for phosphorylated ERK (upper part) or anti-total ERK antibody (lower part). E: As in D, TF-1-fms-Nef-ER cells were deprived of M-CSF for 14 h and then re-stimulated with M-CSF for the indicated periods. TPA or 4-HT was added to the culture at the beginning of the period of deprivation, as indicated. Total cell lysates were analyzed by Western blotting using the antibody specific for phosphorylated ERK (upper part) or anti-total ERK antibody (lower part). The result is representative of two independent experiments.

The proliferative/differentiative properties and the level/duration of ERK activation in TF-1-fms cells expressing conditionally active HIV-1 Nef

As mentioned above, we previously established TF-1-fms cells expressing the conditionally active HIV-1 Nef protein (the Nef-ER fusion protein) and showed that the activation of Nef by 4-HT caused the inhibition of the M-CSF-mediated proliferation of TF-1-fms cells (Figs. 5A and B; Suzu et al., 2005). By using the newly established TF-1 cells expressing the Nef-ER (Fig. 5A), we excluded the possibility that the inhibitory effect of Nef was due to non-specific cytotoxic activity of the viral protein: the activation of Nef did not affect GM-CSF-mediated proliferation of TF-1 cells (Fig. 5B). In this study, we found that Nef activation also caused the inhibition of the M-CSF/TPA-mediated differentiation of TF-1-fms cells (Fig. 5C). Consistent with its inhibitory effect on both M-CSF-mediated proliferation and M-CSF/TPA-mediated differentiation, Nef perturbed the patterns of ERK activation following the stimulation of TF-1-fms-Nef-ER cells with M-CSF (Figs. 5D and E). The untreated control cells

showed a biphasic ERK activation following M-CSF treatment. The early phase occurred between 1 and 3 min whereas the late phase started after 120 min (Fig. 5D). In contrast, the Nef-active cells whose proliferation rate was low showed an earlier but transient ERK activation following M-CSF stimulation (Fig. 5D). On the other hand, TPA-pretreated cells showed an increased and prolonged activation following M-CSF treatment when compared to the untreated control cells (Fig. 5E). Of importance, however, the Nef-active TPA-pretreated cells showed a transient ERK activation (Fig. 5E). These results suggested that the increased and prolonged activation of ERK correlated well with M-CSF/TPA-mediated differentiation, but not with M-CSF-mediated proliferation, of TF-1-fms cells.

Discussion

We established a new macrophage differentiation-inducing system that was dependent on M-CSF activity. TF-1-fms cells definitely switched their responsiveness to M-CSF from proliferation to differentiation in the presence of TPA (Fig. 1). Although the treatment with TPA alone triggered the macrophage differentiation of the cells, the presence of M-CSF enhanced the process: (1) the combination of both M-CSF and TPA caused more drastic morphological changes (Figs. 1C and D); (2) the culture in the presence of both M-CSF and TPA contained more adherent cells than that in the presence of TPA alone (Fig. 2A, upper part); (3) the phagocytic activity and the expression of CD204 were significantly higher in cells treated with M-CSF and TPA than those treated with TPA alone (Fig. 1E). The macrophage differentiation by M-CSF and TPA did not reflect the survival-enhancing/anti-apoptotic function of M-CSF (Fig. 2A, lower part and B).

The experiments with the pharmacological inhibitor U0126 showed that M-CSF/TPA-mediated differentiation of TF-1-fms was dependent on ERK activation (Fig. 3B). That the Src inhibitor PP2 was also a potent inhibitor of the responses (Fig. 3B) might reflect the finding that the activation of ERK by M-CSF was in part dependent on the activity of Src kinases (Cheng et al., 1999; McMahon et al., 2001). However, the involvement of the increased and/or prolonged ERK activation in M-CSF-mediated differentiation is somewhat controversial. Unlike murine M1 cells expressing wild-type Fms, the cells expressing Y559F mutant Fms showed an impaired differentiative response to M-CSF and a reduced level of ERK activation (McMahon et al., 2001). The enforced expression of a scaffolding protein, Gab2, in wild-type Fms-expressing FDC-P1 cells resulted in an acceleration of the differentiation process and increased ERK activation (Liu et al., 2001). In contrast, the enforced expression of an adapter protein, Mona, in wild-type Fms-expressing FDC-P1 cells resulted in increased and prolonged ERK activation, but not an acceleration of M-CSF-mediated differentiation (Bourgin et al., 2000). Moreover, 32D myeloid cells expressing the Y559F mutant Fms showed a "hyper-proliferative" response to M-CSF and prolonged ERK activation (Rohde et al., 2004). Our studies with parental and Nef-expressing TF-1-fms cells (Figs. 4 and 5) supported the former idea that increased and prolonged ERK activation led to M-CSF-mediated macrophage differentiation, but not to cell proliferation. The TPA-pretreated parental TF-1-fms cells (differentiative) showed increased and prolonged ERK activation following M-CSF stimulation (Fig. 4). In contrast, the Nef-active TPA-pretreated TF-1-fms (un-differentiative) showed transient ERK activation (Figs. 5C and E). The pharmacological agent GF109203X, a potent inhibitor of the expression of MAPK phosphatase-1 (Valledor et al., 1999), caused TF-1-fms cells to differentiate in the presence of M-CSF and the cells pretreated with GF109203X showed a sustained ERK activation following M-CSF treatment (data not shown), further supporting the idea.

The molecular mechanisms whereby Nef inhibited the M-CSF-mediated proliferation and M-CSF/TPA-mediated differentiation of TF-1-fms cells remained to be elucidated. In TF-1-fms cells, Nef activation induced the activation of Hck and its constitutive association with Fms (Suzu et al., 2005). The unphysiological behavior of Hck might explain the inhibitory effect of Nef on the responsiveness to M-CSF. Yet, under the proliferation-inducing conditions (TPA-free), the activation of Nef resulted in an earlier ERK activation (Fig. 5D). This might be explained by the finding that Nef induced ERK activation in $CD4^+$ T cells (Schrager et al., 2002) and podocytes (He et al., 2004) in a Src-dependent manner. Further experiments are required to understand the molecular mechanisms whereby Nef rapidly terminated ERK activation in the presence of TPA (Fig. 5E). In summary, we showed that M-CSF-mediated macrophage differentiation, but not proliferation was correlated with increased and prolonged ERK activation, by using a newly established macrophage-inducing system. The culture system with Nef-expressing TF-1-fms cells provides a useful tool for determining the temporal regulatory mechanism of ERK activation and its contribution to M-CSF-mediated proliferation/differentiation.

Acknowledgments

We thank Ranko Shimamura and Yuka Endo for technical and secretarial assistance, respectively.

Literature Cited

Bourette RP, Rohrschneider LR. 2000. Early events in M-CSF receptor signaling. Growth Factors 17:155–166.

Bourette RP, Myles GM, Carlberg K, Chen AR, Rohrschneider LR. 1995. Uncoupling of the proliferation and differentiation signals mediated by the murine macrophage colony-stimulating factor receptor expressed in myeloid FDCP-1 cells. Cell Growth Diff 6:631–645.

Bourgin C, Bourette R, Mouchiroud G, Arnaud S. 2000. Expression of Mona (monocytic adapter) in myeloid progenitor cells results in increased and prolonged MAP kinase activation upon macrophage colony-stimulating factor stimulation. FEBS Lett 480:113–117.

Cheng M, Wang D, Roussel MF. 1999. Expression of c-Myc in response to colony-stimulating factor-1 requires mitogen-activated protein kinase kinase-1. J Biol Chem 274:6553–6558.

Fackler OT, Baur AS. 2002. Live and let die: Nef functions beyond HIV replication. Immunity 16:493–497.

Gredinger E, Gerber AN, Tamir Y, Tapscott SJ, Bengal E. 1998. Mitogen-activated protein kinase pathway is involved in the differentiation of muscle cells. J Biol Chem 273:19436–10444.

Hashimoto S, Suzuki T, Dong HY, Yamazaki N, Matsushima K. 1999. Serial analysis of gene expression in human monocytes and macrophages. Blood 94:837–844.

He H, Wang X, Gorospe M, Holbrook NJ, Trush MA. 1999. Phorbol ester-induced mononuclear cell differentiation is blocked by the mitogen-activated protein kinase kinase (MEK) inhibitor PD98059. Cell Growth Diff 10:307–315.

He JC, Husain M, Sunamoto M, D'Agati VD, Klotman ME, Iyengar R, Klotman PE. 2004. Nef stimulates proliferation of glomerular podocytes through activation of Src-dependent Stat3 and MAPK1, 2 pathways. J Clin Invest 114:643–651.

Kitamura T, Tange T, Terasawa T, Chiba S, Kuwaki T, Miyagawa K, Piao YF, Miyazono K, Urabe A, Takaku F. 1989. Establishment and characterization of a unique human cell line that proliferates dependently on GM-CSF, IL-3, or erythropoietin. J Cell Physiol 140:323–334.

Liu Y, Jenkins B, Shin JL, Rohrschneider LR. 2001. Scaffolding Gab2 mediates differentiation signaling downstream of Fms receptor tyrosine kinase. Mol Cell Biol 21:3047–3056.

Marks DC, Csar XF, Wilson NJ, Novak U, Ward AC, Kanagasundarum V, Hoffmann BW, Hamilton JA. 1999. Expression of a Y559F mutant CSF-1 receptor in M1 myeloid cells: a role of Src kinases in CSF-1 receptor-mediated differentiation. Mol Cell Biol Res Commun 1:144–152.

Marshall CJ. 1995. Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. Cell 80:179–185.

McMahon K-E, Wilson NJ, Marks DC, Beecroft TL, Whitty GA, Hamilton JA, Csar XF. 2001. Colony-stimulating factor-1 (CSF-1) receptor-mediated macrophage differentiation in myeloid cells: a role for tyrosine 559-dependent protein phosphatase 2A (PP2A) activity. Biochem J 358:431–436.

Melemed AS, Ryder JW, Vik TA. 1997. Activation of the mitogen-activated protein kinase pathway is involved in and sufficient for megakaryocytic differentiation of CMK cells. Blood 90:3462–3470.

Moarefi I, LaFevre-Bernt M, Sicheri F, Huse M, Lee C-H, Kuriyan J, Miller WT. 1997. Activation of the Src-family tyrosine kinase Hck by SH3 domain replacement. Nature 385:650–653.

Okada S, Zhang H, Hatano M, Tokuhisa T. 1998. A physiological role of Bcl-$_{xL}$ induced in activated macrophages. J Immunol 160:2590–2596.

Peterlin BM, Trono D. 2003. Hide, shield and strike back: how HIV-infected cells avoid immune eradication. Nat Rev Immunol 3:97–107.

Qiao X, He B, Chiu A, Knowles DM, Chadburn A, Cerutti A. 2006. Human immunodeficiency virus 1 Nef suppresses CD40-dependent immunoglobulin class switching in bystander B cells. Nat Immunol 7:302–310.

Racke FK, Lewandowska K, Goueli S, Goldfarb AN. 1997. Sustained activation of the extracellular signal-regulated kinase/mitogen-activated protein kinase pathway is required for megakaryocytic differentiation of K562 cells. J Biol Chem 272:23366–23370.

Rohde CM, Schrum J, Lee AW-M. 2004. A juxtamembrane tyrosine in the colony stimulating factor-1 receptor regulates ligand-induced Src association, receptor kinase function, and down-regulation. J Biol Chem 279:43448–43461.

Roth P, Stanley ER. 1992. The biology of CSF-1 and its receptor. Curr Top Microbiol Immunol 181:141–167.

Roussel MF, Shurtleff SA, Downing JR, Sherr CJ. 1990. A point mutation at tyrosine-809 in the human colony-stimulating factor 1 receptor impairs mitogenesis without abrogating tyrosine kinase activity, association with phosphatidylinositol 3-kinase, or induction of *c-fos* and *junB* genes. Proc Natl Acad Sci USA 87:6738–6742.

Saksela K, Cheng G, Baltimore D. 1995. Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef^+ viruses but not for down-regulation of CD4. EMBO J 14:484–491.

Schrager JA, Der Minassian V, Marsh JW. 2002. HIV Nef increases T cell ERK MAP kinase activity. J Biol Chem 277:6137–6142.

Sherr CJ, Rettenmier CW, Sacca R, Roussel MF, Look AT, Stanley ER. 1985. The *c-fms* proto-oncogene product is related to the receptor for the mononuclear phagocyte growth factor, CSF-1. Cell 41:665–676.

Suzu S, Kimura F, Ota J, Motoyoshi K, Itoh T, Mishima Y, Yamada M, Shimamura S. 1997. Biologic activity of proteoglycan macrophage colony-stimulating factor. J Immunol 159:1860–1867.

Suzu S, Tanaka-Douzono M, Nomaguchi K, Yamada M, Hayasawa H, Kimura F, Motoyoshi K. 2000. $p56^{dok-2}$ as a cytokine-inducible inhibitor of cell proliferation and signal transduction. EMBO J 19:5114–5122.

Suzu S, Harada H, Matsumoto T, Okada S. 2005. HIV-1 Nef interferes with M-CSF receptor signaling through Hck activation and inhibits M-CSF bioactivities. Blood 105:3230–3237.

Tomokiyo R, Jinnouchi K, Honda M, Wada Y, Hanada N, Hiraoka T, Suzuki H, Kodama T, Takahashi K, Takeya M. 2002. Production, characterization, and interspecies reactivities of monoclonal antibodies against human class A macrophage scavenger receptors. Atherosclerosis 161:123–132.

Valledor AF, Xaus J, Marques L, Celeda A. 1999. Macrophage colony-stimulating factor induces the expression of mitogen-activated protein kinase phosphatase-1 through a protein kinase C-dependent pathway. J Immunol 163:2452–2462.

Valledor AF, Comalada M, Xaus J, Celeda A. 2000. The differential time-course of extracellular-regulated kinase activity correlates with the macrophage response toward proliferation or activation. J Biol Chem 275:7403–7409.

van der Geer P, Hunter T. 1993. Mutation of Tyr697, a GRB2-binding site, and Tyr721, a PI 3-kinase binding site, abrogates signal transduction by the murine CSF-1 receptor expressed in Rat-2 fibroblasts. EMBO J 12:5161–5172.

Walk SF, Alexander M, Maier B, Hammarskjold M-L, Rekosh DM, Ravichandran KS. 2001. Design and use of an inducibly activated immunodeficiency virus type 1 Nef to study immune modulation. J Virol 75:834–843.

Wilhelmsen K, van der Geer P. 2004. Phorbol 12-myristate 13-acetate-induced release of the colony-stimulating factor 1 receptor cytoplasmic domain into the cytosol involves two separate cleavage events. Mol Cell Biol 24:454–464.

W448–W450 *Nucleic Acids Research, 2006, Vol. 34, Web Server issue*
doi:10.1093/nar/gkl214

siVirus: web-based antiviral siRNA design software for highly divergent viral sequences

Yuki Naito[1,*], Kumiko Ui-Tei[1,2], Toru Nishikawa[3], Yutaka Takebe[4] and Kaoru Saigo[1]

[1]Department of Biophysics and Biochemistry, Graduate School of Science, [2]Undergraduate Program for Bioinformatics and Systems Biology, School of Science and, [3]Department of Computer Science, Graduate School of Information Science and Technology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan and [4]Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

Received February 14, 2006; Revised and Accepted March 24, 2006

ABSTRACT

siVirus (http://siVirus.RNAi.jp/) is a web-based online software system that provides efficient short interfering RNA (siRNA) design for antiviral RNA interference (RNAi). siVirus searches for functional, off-target minimized siRNAs targeting highly conserved regions of divergent viral sequences. These siRNAs are expected to resist viral mutational escape, since their highly conserved targets likely contain structurally/functionally constrained elements. siVirus will be a useful tool for designing optimal siRNAs targeting highly divergent pathogens, including human immunodeficiency virus (HIV), hepatitis C virus (HCV), influenza virus and SARS coronavirus, all of which pose enormous threats to global human health.

INTRODUCTION

RNA interference (RNAi) is now widely used to knockdown gene expression in a sequence-specific manner, making it a powerful tool not only for studying gene function, but also for therapeutic purposes, including antiviral treatments (1–4). Currently, the replication of a wide range of viruses can be inhibited successfully using RNAi, with both short interfering RNAs (siRNAs) and siRNA expression vectors (5).

In mammalian RNAi, the efficacy of each siRNA varies widely depending on its sequence; only a limited fraction of randomly designed siRNAs is highly effective. Many experiments have been conducted to clarify possible sequence requirements of functional siRNAs. Of these, our work incorporates guidelines from three major studies (6–8) of selecting functional siRNAs. However, designing functional siRNAs that target viral sequences is problematic because of their extraordinarily high genetic diversity. For example, about 500 entries of near full-length sequences of HIV-1 group M, which is largely responsible for global pandemic, are stored in the sequence databases, but it proved impossible to select a common 21mer from among all of them. Moreover, RNAi-resistant viral mutants achieved through point mutation or deletion emerge rapidly when targeting viruses in cell culture. These problems suggest a strong need to select highly conserved target sites for designing antiviral siRNAs. Furthermore, the off-target silencing effects of siRNA are also a serious problem that could affect host gene expression (9). Off-target silencing effects arise when an siRNA has sequence similarities with unrelated genes. In antiviral RNAi, it is desirable to minimize off-target effects against human genes.

Consequently, only a limited fraction of 21mers is suitable for use as antiviral siRNAs. In this study, we developed a novel web-based online software system, siVirus, which provides functional, off-target minimized siRNAs targeting highly conserved regions of divergent viral sequences.

METHODS

Selection of highly conserved siRNA target sites

Highly conserved siRNA sequences are selected based on their *degree of conservation*, defined as the proportion of viral sequences that are targeted by the corresponding siRNA, with complete matches (i.e. 21/21 matches). All possible siRNA candidates targeting every other position of user-selected viral sequences are generated and their degrees of conservation are computed. Users can arbitrarily specify a set of viral sequences for the computation; e.g. sequences can be selected from a specific geographic region(s) or a specific genotype(s) to design the best siRNAs tailored to specific user needs. siVirus also accepts user's own sequences in a multi-FASTA format and shows whether each siRNA can target the posted sequences.

*To whom correspondence should be addressed. Tel: +81 3 5841 4404; Fax: +81 3 5841 4400; Email: y-naito@RNAi.jp

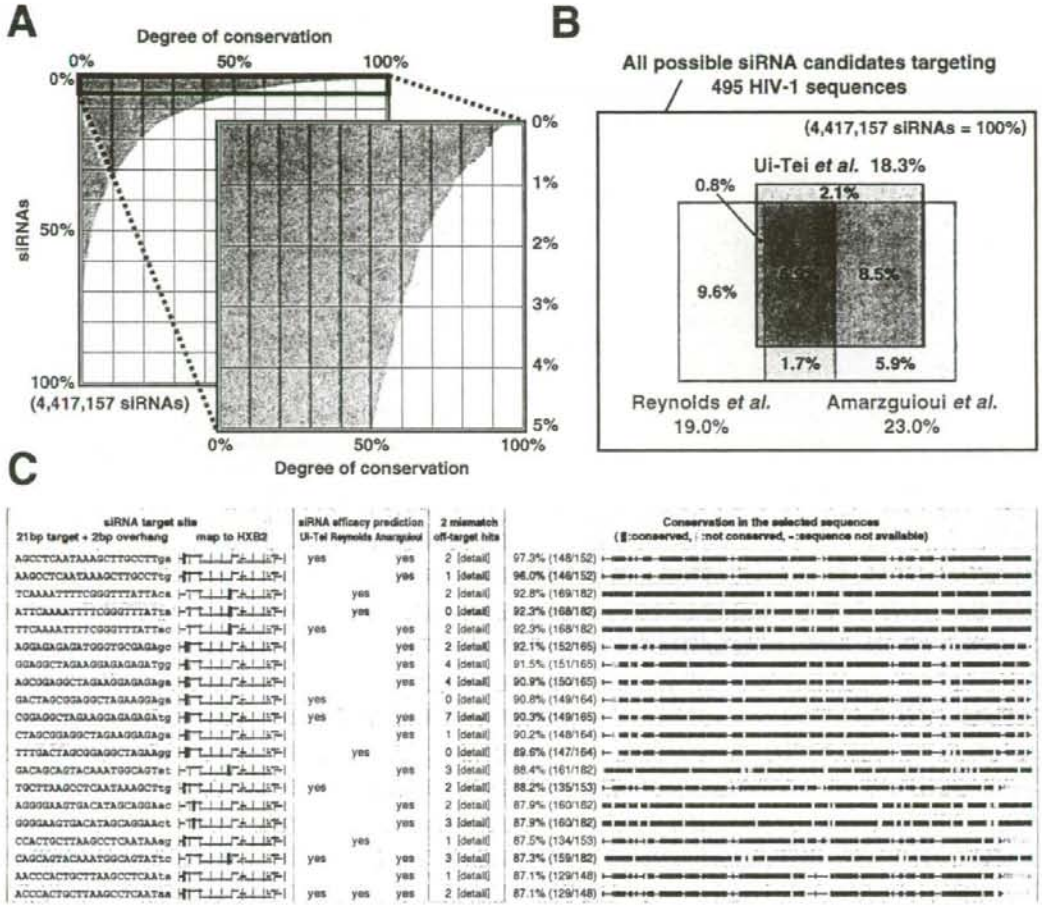


Figure 1. (**A**) The degree of conservation is calculated for all possible siRNA candidates (total 4 417 157) targeting every other position of 495 HIV-1 sequences. (**B**) The efficacy predictions of these 4 417 157 siRNA candidates based on three different guidelines: Ui-Tei *et al.* (6), Reynolds *et al.* (7) and Amarzguioui *et al.* (8). (**C**) Typical output of siVirus for designing anti-HIV siRNAs. Sequence information, efficacy predictions, off-target search results and the degrees of conservation are shown.

siRNA efficacy prediction

In mammalian RNAi, the efficacy of each siRNA varies markedly depending on its sequence; hence, several groups have reported guidelines for selecting functional siRNAs. siVirus incorporates the guidelines of Ui-Tei *et al.* (6), Reynolds *et al.* (7) and Amarzguioui *et al.* (8) and shows whether each siRNA satisfies these guidelines.

Off-target searches

Off-target searches were performed for each siRNA using siDirect (10,11). siVirus shows the number of off-target hits within two mismatches against the non-redundant database of human transcripts (10).

Database maintenance

Currently, siVirus incorporates viral genome sequences of HIV-1, HCV, influenza A virus and SARS coronavirus. These sequences were downloaded from the Los Alamos HIV Sequence Database (http://hiv-web.lanl.gov/), the Los Alamos HCV Sequence Database (12), the NCBI Influenza Virus Sequence Database (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html), and NCBI GenBank (13), respectively. siVirus will be updated continuously as these databases are revised. We also plan to incorporate other viruses if sufficient numbers of their sequences are available.

RESULTS AND DISCUSSION

To design anti-HIV siRNA, we analyzed the 495 near full-length HIV-1 sequences listed in Supplementary Table 1. A total of 4 417 157 possible siRNA candidates (i.e. substrings of length 21) targeting every other position of the HIV-1 sequences were produced from the 495 viral sequences. The analysis of these siRNA candidates revealed that highly

conserved siRNAs constituted only 0.3% of the possible siRNAs if >90% conservation is expected (Figure 1A). The fraction is still as small as 0.8% even if the threshold of the conservation is relaxed to 80%. On the other hand, siRNAs predicted to be functional by one or more guidelines (6–8) constituted 35.5% of the 4 417 157 siRNAs (Figure 1B). Taken together, siRNAs that are >80% conserved, and satisfy at least one guideline constitute only 0.2% of the siRNAs. In this condition, 20–30 siRNAs can be designed for each full-length sequence of HIV-1. These indicate that most of the randomly designed siRNAs are not suited for targeting HIV-1 efficiently.

Figure 1C shows typical output from siVirus for designing anti-HIV siRNAs. A total of 182 sequences from HIV-1 subtypes B, C and CRF01_AE, which are the most prevalent HIV-1 genotypes circulating in Asia, were selected. The results were sorted by their degree of conservation, and filtered to display siRNAs that satisfy at least one efficacy guideline. The off-target search results against human genes are also shown. It is desirable to select an siRNA that has less off-target hits.

To test the validity of siVirus, 35 siRNAs satisfying the guideline by Ui-Tei *et al.* (6) were designed against the conserved regions of HIV-1 genomes using siVirus and were assayed for inhibition of viral replication. Among them, 31 siRNAs effectively inhibited HIV-1 replication by >80% when each siRNA duplex was transfected at 5 nM (Y. Naito, K. Ui-Tei, K. Saigo and Y. Takebe, unpublished data).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan to K.S., K.U.-T. and Y.T., and by grants from the Ministry of Health, Labour and Welfare of Japan to Y.T. Funding to pay the Open Access publication charges for this article was provided by the Ministry of Education, Culture, Sports, Science and Technology of Japan. Y.N. is a Research Fellow of the Japan Society for the Promotion of Science.

Conflict of interest statement. None declared.

REFERENCES

1. Fire,A., Xu,S., Montgomery,M.K., Kostas,S.A., Driver,S.E. and Mello,C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806–811.
2. Mello,C.C. and Conte,D. Jr (2004) Revealing the world of RNA interference. *Nature*, **431**, 338–342.
3. Hannon,G.J. and Rossi,J.J. (2004) Unlocking the potential of the human genome with RNA interference. *Nature*, **431**, 371–378.
4. Voinnet,O. (2005) Induction and suppression of RNA silencing: insights from viral infections. *Nature Rev. Genet.*, **6**, 206–220.
5. Leonard,J.N. and Schaffer,D.V. (2006) Antiviral RNAi therapy: emerging approaches for hitting a moving target. *Gene Ther.*, **13**, 532–540.
6. Ui-Tei,K., Naito,Y., Takahashi,F., Haraguchi,T., Ohki-Hamazaki,H., Juni,A., Ueda,R. and Saigo,K. (2004) Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Res.*, **32**, 936–948.
7. Reynolds,A., Leake,D., Boese,Q., Scaringe,S., Marshall,W.S. and Khvorova,A. (2004) Rational siRNA design for RNA interference. *Nat. Biotechnol.*, **22**, 326–330.
8. Amarzguioui,M. and Prydz,H. (2004) An algorithm for selection of functional siRNA sequences. *Biochem. Biophys. Res. Commun.*, **316**, 1050–1058.
9. Jackson,A.L. and Linsley,P.S. (2004) Noise amidst the silence: off-target effects of siRNAs? *Trends Genet.*, **20**, 521–524.
10. Naito,Y., Yamada,T., Ui-Tei,K., Morishita,S. and Saigo,K. (2004) siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference. *Nucleic Acids Res.*, **32**, W124–W129.
11. Yamada,T. and Morishita,S. (2005) Accelerated off-target search algorithm for siRNA. *Bioinformatics*, **21**, 1316–1324.
12. Kuiken,C., Yusim,K., Boykin,L. and Richardson,R. (2005) The Los Alamos hepatitis C sequence database. *Bioinformatics*, **21**, 379–384.
13. Benson,D.A., Karsch-Mizrachi,I., Lipman,D.J., Ostell,J. and Wheeler,D.L. (2006) GenBank. *Nucleic Acids Res.*, **34**, D16–D20.

Retrovirology

BioMed Central

Short report

Open Access

Optimal design and validation of antiviral siRNA for targeting HIV-1

Yuki Naito*[1], Kyoko Nohtomi[2], Toshinari Onogi[2], Rie Uenishi[2], Kumiko Ui-Tei[1], Kaoru Saigo[1] and Yutaka Takebe*[2]

Address: [1]Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan and [2]Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo, 162-8640, Japan

Email: Yuki Naito* - y-naito@RNAi.jp; Kyoko Nohtomi - notomi@nih.go.jp; Toshinari Onogi - onogit@nih.go.jp; Rie Uenishi - uenishir@nih.go.jp; Kumiko Ui-Tei - ktei@biochem.s.u-tokyo.ac.jp; Kaoru Saigo - saigo@biochem.s.u-tokyo.ac.jp; Yutaka Takebe* - takebe@nih.go.jp

* Corresponding authors

Published: 8 November 2007

Retrovirology 2007, **4**:80 doi:10.1186/1742-4690-4-80

Received: 6 August 2007
Accepted: 8 November 2007

This article is available from: http://www.retrovirology.com/content/4/1/80

Abstract

We propose rational designing of antiviral short-interfering RNA (siRNA) targeting highly divergent HIV-1. In this study, conserved regions within HIV-1 genomes were identified through an exhaustive computational analysis, and the functionality of siRNAs targeting the highest possible conserved regions was validated. We present several promising antiviral siRNA candidates that effectively inhibited multiple subtypes of HIV-1 by targeting the best conserved regions in pandemic HIV-1 group M strains.

Findings

RNA interference (RNAi) is now widely used to knock-down gene expression in a sequence-specific manner, making it a powerful tool not only for studying gene function, but also for therapeutic applications including antiviral treatments [1,2]. The replication of a wide range of viruses can be successfully inhibited using RNAi with both short interfering RNA (siRNA) and siRNA expression vectors [3,4]. However, for RNA viruses such as HIV-1, designing functional siRNAs that target viral sequences is problematic because of their extraordinarily high genetic diversity. We analyzed 495 entries of near full-length HIV-1 group M sequences available in the Los Alamos HIV Sequence Database, and selected the highest-possible conserved target sites for designing optimal antiviral siRNAs. It is known that RNAi-resistant viral mutants emerge rapidly when targeting viral sequences due to their high mutation rate [5-7]. Since highly conserved sequences are likely to contain structurally or functionally constrained elements, our approach is anticipated to resist viral mutational escape.

First, we performed a detailed analysis on the HIV-1 genome to identify highly conserved targets by using 495 near full-length genome sequences of HIV-1 group M (listed in Additional file 1). Every possible 21-mer was generated from all of the HIV-1 group M sequences, and their conservations among the 495 HIV-1 sequences were exhaustively determined using siVirus engine [8]. We defined 'conservation' as the percentage of sequence entries out of the 495 HIV-1 sequences that showed perfect identity (*i.e.*, 21/21 matches) with the cognate 21-mer. Since many of the HIV-1 sequence entries lack 5' untranslated region (5' UTR), the 3' LTR sequence was used to compensate for the lack of 5' LTR sequences in order to avoid underestimating conservation in such regions. For the regions that cannot be compensated for in this way (depicted in Figure 1A and 1B left panel, colored

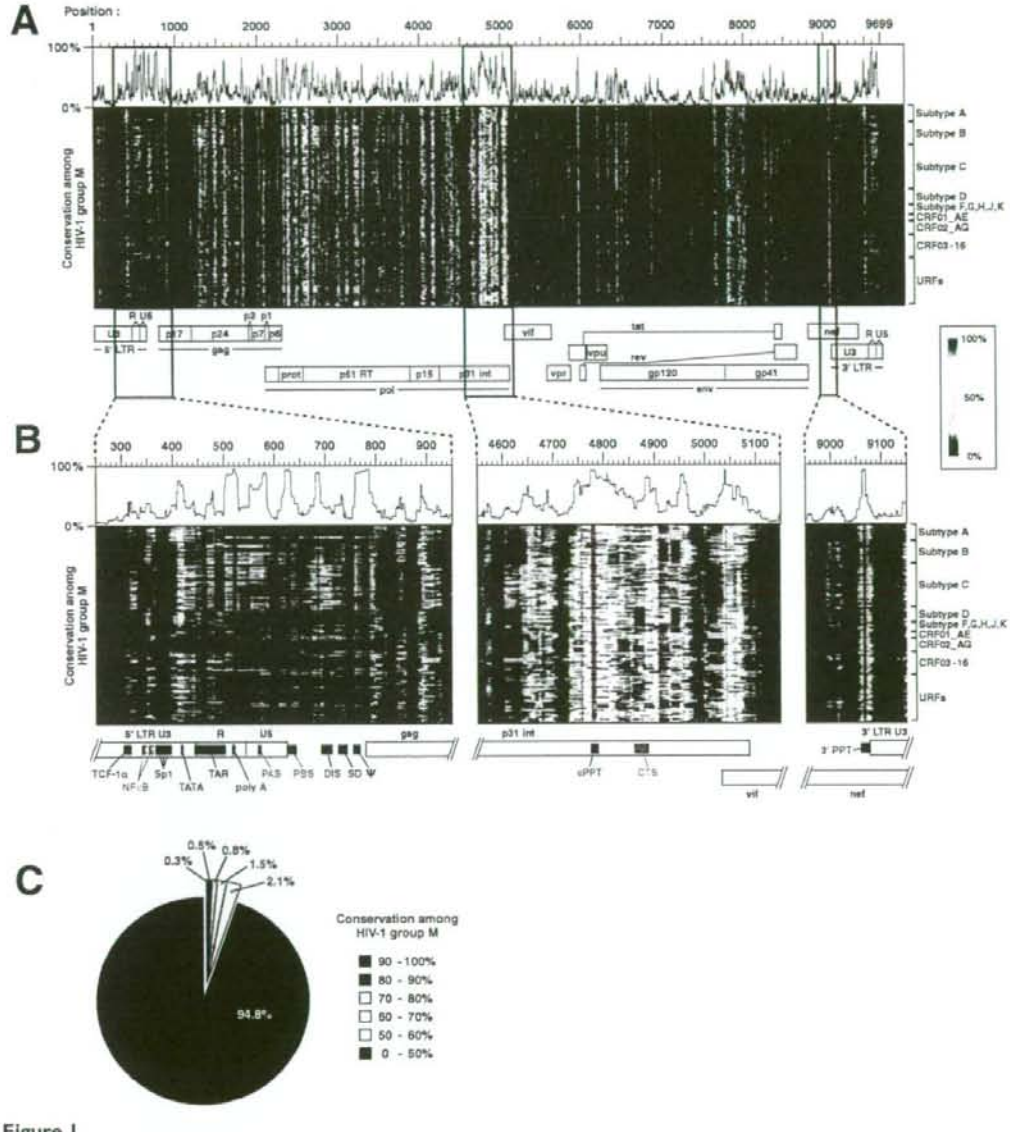


Figure 1

Conservations of siRNA target sequences among HIV-1 group M. (**A**) A total of 4,417,157 siRNA targets were generated from the 495 HIV-1 sequences, and their conservations within the HIV-1 genomes are represented using a color density plot. The line plot above the color chart represents the highest value in each position. (**B**) A detailed view of the three conserved regions; 5' LTR, the cPPT/CTS in the integrase gene, and 3' PPT. 'Position' indicates the 5'-most position of each 21-mer. The landmarks of the HIV-1 genome are adjusted to align at the center of the siRNAs by shifting 10 bp to the left. (**C**) Pie chart indicating the percentage of the 4,417,157 siRNA target sites at each conservation level.

black), conservation was calculated by considering only the HIV-1 sequences that contain the corresponding regions. The result revealed that HIV-1 genomes are not conserved for consecutive 21 bp for the most part, resulting in the poor conservation of many of the 21-mers over the HIV-1 sequences (Figure 1A, colored blue). As shown in Figure 1C, only 5.2% of the possible 21-mers are >50% conserved. Furthermore, highly (>70%) conserved 21-mers constitute only 1.6% of all 21-mers. It is of note that many of the published anti-HIV-1 siRNA sequences do not fall into this 'highly conserved' category (Additional file 2 and [9]). From these results, we anticipate that most of the possible siRNAs are not suitable for the efficient targeting of HIV-1.

However, our analysis has identified several distinct regions that are highly conserved in the HIV-1 genome (Figure 1B). Such regions include the regulatory domains responsible for the viral gene expression, such as the TATA sequence and polyadenylation signal (AAUAAA). In addition, several regions essential for the regulation of viral replication were also highly conserved, including the primer activation signal (PAS)[10], primer binding site (PBS), packaging signal (Ψ), central polypurine tract (cPPT), central termination sequence (CTS), and 3' polypurine tract (3' PPT). All of these highly conserved sequences are constrained at the nucleotide sequence level or by their RNA secondary structure in order to execute their functions. In contrast, regions constrained by amino acid sequences were not necessarily conserved at the nucleotide sequence level due to the wobbling of the third base in the codon (data not shown). siRNAs targeting the highly conserved regions are expected to overwhelm the high level of sequence diversity of the HIV-1 genome, and also to reduce the chances of viral mutational escapes.

Total of 216 highly conserved (>70%) siRNA targets identified in this study are listed in Additional file 3. In mammalian RNAi, the efficacy of each siRNA varies markedly depending on its sequence. According to our guidelines for the selection of effective siRNAs [11,12], 31 out of 216 siRNAs were predicted to be functional. Similarly, 30 and 44 siRNAs are functional according to the algorithms reported by Reynolds *et al.* [13], and Amarzguioui *et al.* [14], respectively (Additional file 3). This suggests that only a limited fraction of 21-mers is best suited for use as functional antiviral siRNAs.

For the functional validation, 23 siRNAs from Additional file 3, and 18 additional siRNAs targeting moderately-conserved regions were selected based on the following criteria: (I) predicted to be functional by the algorithm of Ui-Tei *et al.* [11,12], and (II) the sequence has perfect identity with pNL4-3 (GenBank M19921). The 41 siRNA sequences selected and their target sites are detailed in Additional file 4. We first tested the efficacy of each siRNA using target mRNA cleavage assay (Additional file 5 and [15]). Briefly, a vector expressing reporter mRNA that contains the siRNA target site was cotransfected into HeLa cells with the corresponding siRNA, and the mRNA cleavage activity of the siRNA was evaluated by measuring the quantity of surviving mRNA using real-time RT-PCR. This assay allows us to directly monitor the sequence-dependent potency of siRNA itself, without being affected by the differences in target gene expression level or target secondary structures. The result showed that 39 out of the 41 siRNAs gave >60% silencing at 5 nM (Figure 2, rightmost panel). si4794 and si4888 were not functional, probably due to the long consecutive Gs in si4794 and internal palindromes (AAAAUUUU) in si4888 [11,13].

Next, siRNAs were evaluated for their antiviral efficacy against three evolutionary-distant groups of HIV-1: subtypes B and B' (Thailand variant of subtype B [16]); subtype C; and CRF01_AE. Each siRNA was cotransfected into HeLa cells at 5 nM with one of the four infectious molecular clones: pNL4-3 (subtype B); 95MM-yIDU106 (subtype B'); 93IN101 (subtype C); or 93JP-NH1 (CRF01_AE). Culture supernatants were collected 48 h after transfection and the viral reverse transcriptase activity was measured (Additional file 5 and [17]). The results show that 26 of the 41 siRNAs effectively inhibited viral replication of all four strains by >80% (Figure 2, marked with red or orange circles). Of the remaining 15 siRNAs, 13 of them (except si4794/4888) were shown to be functional in the target mRNA cleavage assay, and 12 of them (except si690/4794/4888) inhibited the replication of at least one viral strain by >80%, indicating that the designed siRNAs have the potential to induce RNAi. In several viral strains, nucleotide substitutions in their target sites essentially abolished the inhibition of viral replication (Figure 2, blue bars with arrowheads). However, mismatches near the ends of the target sites (see Additional file 6) did not necessarily abolish the siRNA efficacy (Figure 2, blue bars with asterisks). si689 and si690 did not inhibit viral replication even though these siRNAs perfectly matched to their target sites (confirmed by DNA sequencing of the infectious molecular clones). This is probably due to the stable secondary structure at the si689-690 target sites in both BMH (branched multiple hairpin) conformation and LDI (long distance interaction) conformation of the HIV-1 leader RNA [18] (see Additional file 4). It should be noted that the efficacy of si575 differed when targeting pNL4-3 and 93IN101. One possible explanation for this is the secondary structure differences among HIV-1 subtypes, which may alter the accessibility of the si575 target site.

The approach described here enabled us to select highly effective siRNAs against divergent HIV-1 strains at a high

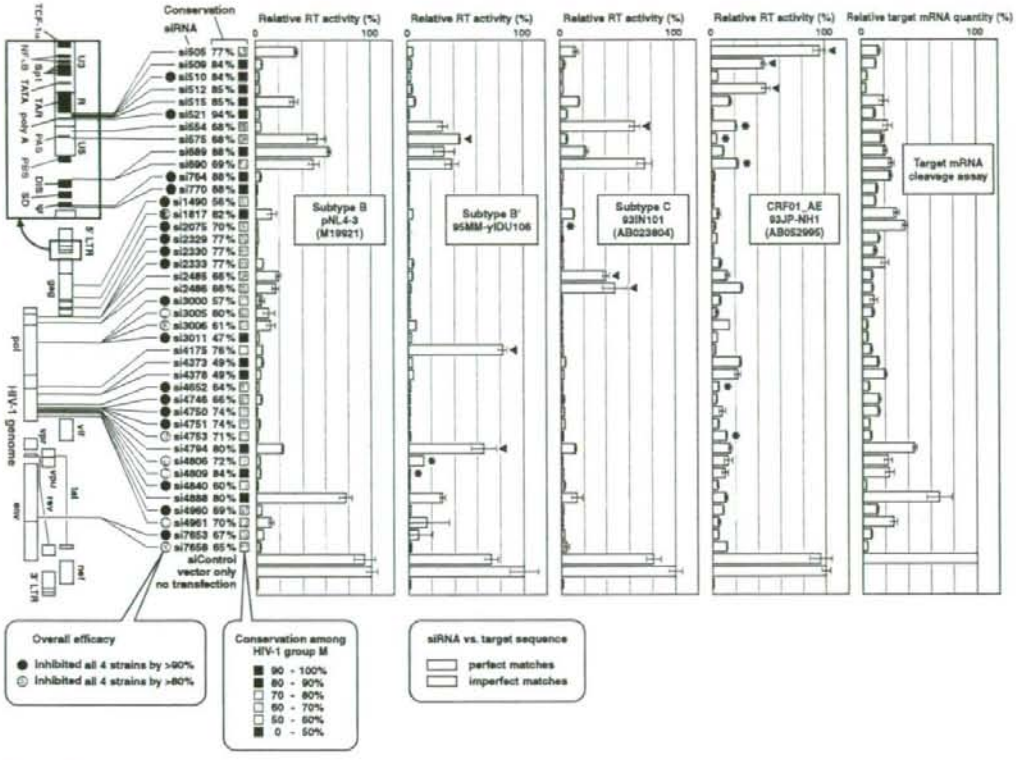


Figure 2
Validation of 41 siRNAs. The antiviral efficacy of each siRNA was tested against four HIV-1 infectious molecular clones: pNL4-2 (subtype B); 95MM-yIDU106 (subtype B'); 93IN101 (subtype C); or 93JP-NH1 (CRF01_AE). The potency of each siRNA was tested using the target mRNA cleavage assay (rightmost panel). The ability of each siRNA to cleave its target was evaluated by the target mRNA cleavage assay.

rate. The highly effective siRNAs (>90% inhibition) with maximal conservation (>70%) identified in our study include si521 (poly A site; 94% conservation), si764/770 (Ψ; 88%), si510 (TAR/poly A; 84%), si2075 (ribosomal slip site; 70%), si2329/2330/2333 (protease region; 77%), and si4750/4751/4753 (integrase region; 71–74%). These sites are found mostly in the 5' LTR, protease, and integrase regions (Figure 2). However, the extraordinarily high genetic diversity of HIV-1 obviously prevents us from designing a single siRNA that can nullify all HIV-1 strains currently circulating worldwide (Additional file 7). One possible approach is to combine multiple siRNAs targeting different conserved regions [19,20]. The siRNAs selected and validated in this study have the potential to target >99% of HIV-1 strains by combining only two siRNAs (Additional file 7), and also considered to resist viral mutational escape. Our approach is expected to be highly applicable to therapeutic intervention for other pathogens of public health importance, including HCV, influenza virus, and SARS coronavirus, that are known to show high genetic diversity.

Competing interests
The author(s) declare that they have no competing interests.

Authors' contributions
YN performed the computational analyses and the target mRNA cleavage assays, participated in the design of the study, and drafted the manuscript. KN and TO performed the viral replication assays. RU analyzed the data. KU-T participated in the target mRNA cleavage assays, and was

involved in critically revising the manuscript. KS and YT supervised the entire study and wrote the manuscript.

Additional material

Additional file 1

The list of 495 near full-length genome sequences of HIV-1 group M.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S1.pdf]

Additional file 2

The list of published siRNA/shRNAs targeting HIV-1.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S2.pdf]

Additional file 3

The list of highly conserved siRNA targets identified in this study.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S3.pdf]

Additional file 4

The siRNA sequences and their target sites. The sequences of 41 siRNAs and their target sites are shown. The siRNA numbers indicate the nucleotide position in HXB2 (GenBank K03455). The conservation level of each siRNA in HIV-1 group M sequence is depicted in color chart at the rightmost column. BMH (branched multiple hairpin) and LDI (long distance interaction) conformations of the HIV-1 leader RNA and siRNAs targeting them are shown.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S4.pdf]

Additional file 5

Supplementary materials and methods.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S5.pdf]

Additional file 6

Target sites of the 41 siRNAs used in this study. Sequence alignment of the target site from the four HIV-1 infectious molecular clones: pNL4-2 (subtype B); 95MM-ylDU106 (subtype B'); 93IN101 (subtype C); or 93JP-NH1 (CRF01_AE).
Click here for file
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S6.pdf]

Additional file 7

Coverage of HIV-1 group M by single siRNA or two siRNAs. (A) Coverage of HIV-1 group M by 41 siRNAs used in this study. (B) Coverage of HIV-1 group M by combining two siRNAs from above. Coverage was calculated by considering only the HIV-1 sequences which contain the corresponding regions.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S7.pdf]

Acknowledgements

This study was supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to YN, KU-T, KS, and YT), the Ministry of Health, Labour and Welfare of Japan (to YT), and the Japan Health Sciences Foundation (to YT).

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**:806-811.
2. Hannon GJ, Rossi JJ: **Unlocking the potential of the human genome with RNA interference.** *Nature* 2004, **431**:371-378.
3. Nielsen MH, Pedersen FS, Kjems J: **Molecular strategies to inhibit HIV-1 replication.** *Retrovirology* 2005, **2**:10.
4. Leonard JN, Schaffer DV: **Antiviral RNAi therapy: emerging approaches for hitting a moving target.** *Gene Ther* 2006, **13**:532-540.
5. Boden D, Pusch O, Lee F, Tucker L, Ramratnam B: **Human immunodeficiency virus type 1 escape from RNA interference.** *J Virol* 2003, **77**:11531-11535.
6. Das AT, Brummelkamp TR, Westerhout EM, Vink M, Madiredjo M, Bernards R, Berkhout B: **Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition.** *J Virol* 2004, **78**:2601-2605.
7. Westerhout EM, Ooms M, Vink M, Das AT, Berkhout B: **HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome.** *Nucleic Acids Res* 2005, **33**:796-804.
8. Naito Y, Ui-Tei K, Nishikawa T, Takebe Y, Saigo K: **siVirus: web-based antiviral siRNA design software for highly divergent viral sequences.** *Nucleic Acids Res* 2006, **34**:W448-W450.
9. ter Brake O, Berkhout B: **A novel approach for inhibition of HIV-1 by RNA interference: counteracting viral escape with a second generation of siRNAs.** *J RNAi Gene Silencing* 2005, **1**:56-65.
10. Beerens N, Groot F, Berkhout B: **Initiation of HIV-1 reverse transcription is regulated by a primer activation signal.** *J Biol Chem* 2001, **276**:31247-31256.
11. Ui-Tei K, Naito Y, Takahashi F, Haraguchi T, Ohki-Hamazaki H, Juni A, Ueda R, Saigo K: **Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference.** *Nucleic Acids Res* 2004, **32**:936-948.
12. 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 Res* 2004, **32**:W124-W129.
13. Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorova A: **Rational siRNA design for RNA interference.** *Nat Biotechnol* 2004, **22**:326-330.
14. Amarzguioui M, Prydz H: **An algorithm for selection of functional siRNA sequences.** *Biochem Biophys Res Commun* 2004, **316**:1050-1058.
15. Ui-Tei K, Naito Y, Saigo K: **Guidelines for the selection of effective short-interfering RNA sequences for functional genomics.** *Methods Mol Biol* 2007, **361**:201-216.
16. Kalish ML, Baldwin A, Raktham S, Wasi C, Luo CC, Schochetman G, Mastro TD, Young N, Vanichseni S, Rübsamen-Waigmann H, von Briesen H, Mullins JI, Delwart E, Herring B, Esparza J, Heyward WL, Osmanov S: **The evolving molecular epidemiology of HIV-1 envelope subtypes in injecting drug users in Bangkok, Thailand: implications for HIV vaccine trials.** *AIDS* 1995, **9**:851-857.
17. Willey RL, Smith DH, Lasky LA, Theodore TS, Earl PL, Moss B, Capon DJ, Martin MA: **In vitro mutagenesis identifies a region within the envelope gene of the human immunodeficiency virus that is critical for infectivity.** *J Virol* 1988, **62**:139-147.
18. Huthoff H, Berkhout B: **Two alternating structures of the HIV-1 leader RNA.** *RNA* 2001, **7**:143-157.
19. Nishitsuji H, Kohara M, Kannagi M, Masuda T: **Effective suppression of human immunodeficiency virus type 1 through a combination of short- or long-hairpin RNAs targeting essential sequences for retroviral integration.** *J Virol* 2006, **80**:7658-7666.
20. ter Brake O, Konstantinova P, Ceylan M, Berkhout B: **Silencing of HIV-1 with RNA interference: a multiple shRNA approach.** *Mol Ther* 2006, **14**:883-892.