

Chapter VI

Sequence Requirements for Effective Mammalian RNAi Induced by Short Double-Stranded RNAs Equal to or More than 22bp in Length

Kumiko Ui-Tei, Shuhei Zenno and Kaoru Saigo

Department of Biophysics and Biochemistry, Graduate School of Science,
University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Abstract

RNA interference (RNAi) has been shown quite useful in the clarification of gene function in various organisms. Synthetic 21bp long double-stranded RNAs (dsRNAs) each with two 2nt long 3'overhangs have been found to serve adequately as short interfering RNAs (siRNAs) for mammalian RNAi. The mechanisms of 21bp-long siRNA dependent RNAi in mammalian cells have been studied extensively and rules for designing sequences of highly effective siRNAs have been established. Mammalian RNAi can also be induced by transfection of dsRNAs 22 to 30bp in length. Although such dsRNAs may yield 21bp long siRNAs via cellular Dicer digestion, practical guidelines for designing sequences of 22 to 30bp long dsRNAs capable of inducing effective mammalian RNAi have yet be drawings. The present study examined the effects of RNAi exerted by 22 to 30-bp long dsRNAs all possessing two 2nt long, 3'overhangs in mammalian cells. RNAi activity induced by these dsRNAs appeared identical to the combined RNAi activity induced by 21 and 22bp long terminal siRNAs, both resistant to Dicer digestion. Sequence preference rules for highly effective siRNA 22bp in length were also established and found quite similar, if not identical to those for highly effective siRNA 21bp in length. Designing sequences of dsRNA more than 22bp in length and capable of inducing highly effective RNAi with reduced off-target effect is proposed.

Introduction

RNA interference (RNAi) is a process by which double-stranded RNA (dsRNA) silences gene expression through inducing sequence-specific degradation of complementary mRNA (Fire et al., 1998; Mello and Conte, 2004; Hannon and Rossi, 2004). RNAi is considered to be initiated when dsRNA, either introduced into or produced within cells, is processed with the RNase III enzyme Dicer into nearly 21bp long short interfering RNAs (siRNAs) each with two 2-nucleotides (nt) long, 3'overhangs (Carmell and Hannon, 2004). Cellular Dicer may likely cleave dsRNAs via the intramolecular dimerization of its two RNaseIII domains which are associated with the flanking RNA binding domains, PAZ and dsRBD (Zhang et al., 2002).

Chemically synthesized 21bp long siRNA has been shown capable of inducing RNAi in mammalian and non-mammalian cells (Elbashir et al., 2001b). The introduction of long dsRNA into mammalian cells occasionally elicits fatal antiviral/interferon response (Stark et al., 1998), though short dsRNA such as siRNA may not, indicating that 21bp long siRNA is a more promising reagent for mammalian RNAi (Elbashir et al., 2001a; 2001b) than long dsRNA (Ui-Tei et al., 2000; Billy et al., 2001). Within cells, the guide strand of siRNA is incorporated into the RNA-induced silencing complex (RISC) with consequent suppression of target gene expression in a sequence-specific manner. A core component of RISC is Argonaute (Doi et al., 2003) and, in mammals, Argonaute-2 is essential for substrate (target mRNA) cleavage (Liu et al., 2004; Song et al., 2004).

In contrast to RNAi in *Drosophila*, only a small fraction of siRNAs is capable of inducing effective RNAi in mammalian cells (Holen et al., 2002; Ui-Tei et al., 2004). We previously reported a method for designing sequences for 21bp long siRNAs highly functional in mammalian cells (Ui-Tei et al., 2004, 2006; Naito et al., 2004). Our guidelines indicate 21bp long siRNAs simultaneously satisfying all four of the following sequence conditions to be capable of inducing highly effective gene -silencing in mammalian cells: A/U at the 5'end of the guide strand (GS); G/C at the 5'end of the passenger strand (PS); at least four A/U residues in the 5'terminal third of GS and the absence of any GC stretch of more than 9nt in length. All siRNAs not matching the first three sequence conditions give rise to little or no gene-silencing in mammalian cells. Effective siRNAs for mammalian RNAi may also be designed using methods other than ours (Reynolds et al., 2004; Amarzguioui and Prýdz, 2004).

RNAi is also known to be induced by the transfection of DNA encoding short hairpin RNA (shRNA). Using shRNA-encoding DNA libraries, large scale screening of loss-of-function mutants has been carried out in many different mammalian cell types (Silva et al., 2005; Moffat et al., 2006). The primary transcript of RNA polymerase III (pol III) promoter-driven shRNA is considered to form a stem-loop structure associated with a 2-6nt-long U-stretch at the 3'end, to possibly give rise to various Dicer digestion products (Vermeulen et al., 2005). A new shRNA expression vector has been reported in which the primary shRNA transcript mimics a natural micro-RNA primary transcript (Silva et al., 2005). In that system, long-polyadenylated RNA containing shRNA is produced as a primary transcript and then cleaved by the RNase III enzyme, Drosha, to produce stem-looped shRNA lacking the 3'terminal U-stretch, a characteristic structure of pol III transcripts. Apparently, this makes it

possible to produce 21 to 22bp long cellular siRNAs each with two 2nt long 3'overhangs through Dicer digestion. Thus, if sequence preference rules for 22bp long siRNA are available, sequences for shRNA constructs that produce effective 21 and 22bp long siRNAs can be designed with this new system.

Effective mammalian RNAi may also be induced by transfecting dsRNA slightly longer than Dicer digestion products. Short Dicer-substrate dsRNA about 27bp long has actually been shown much more functional than 21 to 22bp long authentic siRNA (Kim et al., 2005). Forty-three and 63bp long dsRNAs each with two 2nt long 3'overhangs were assessed for capacity to RNAi in mammalian cells (Ui-Tei et al., 2005). The molecular mechanism for RNAi induced by dsRNAs slightly longer than 21bp long siRNA still remains to be fully clarified. Vermeulen et al. (2005) noted 61bp long dsRNAs containing a blunt end or that with 3'overhangs to be processed in distinctly different manners and emphasize importance of interactions between dsRNA ends and the PAZ domain of Dicer for specificity and efficiency in dicing. Ui-Tei et al. (2005) and Rose et al. (2005) also independently demonstrated the importance of 2nt long 3'overhangs for "in-frame" digestion of dsRNA from its ends.

Here, we examine molecular basis for RNAi induced by short Dicer substrate dsRNA with two 2nt long 3'overhangs. Sequence preference rules for 22bp long highly effective siRNAs and a method for designing short Dicer substrate dsRNAs associated with high RNAi inducing activity with reduced off-target effect in mammalian cells are also presented.

Materials and Methods

Cell Culture

Chinese hamster ovary CHO-K1 (RIKEN Cell Bank) and human HeLa cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM; Gibco BRL) at 37°C. The medium used for each cell line was supplemented with 10% heat-inactivated fetal bovine serum (FBS; Mitsubishi Kagaku) and antibiotics (10 units/ml of penicillin (Meiji) and 50 µg/ml of streptomycin (Meiji)). Mouse ES cells, E14TG2a, were cultured in DMEM supplemented with 20% heat-inactivated FBS (HyClone), 0.1 mM 2-mercaptoethanol (Wako), 8 µg/ml adenosine, 8.5 µg/ml guanosine, 7.3 µg/ml cytidine, 7.3 µg/ml uridine, 2.4 µg/ml thymidine, 0.1 mM of each nonessential amino acid, and 1,000 units/ml of leukemia inhibitor factor (Chemicon International).

siRNA Preparation

Passenger- and guide-strand RNA oligonucleotides were synthesized chemically (Prolig σ), mixed in a 1:1 ratio in 10 mM NaCl and 20 mM Tris-HCl (pH 7.5), and annealed by incubating at 95°C for 15 min, 37°C for 30 min, and 25°C for 30 min. The annealed products were examined using 3% agarose gel electrophoresis in TBE buffer, which can separate 21bp long double-stranded siRNA from 21nt-long single-stranded RNA. Almost all

of the RNA was recovered as dsRNA. Control siRNA, siGY441, against the gene for enhanced green fluorescent protein (EGFP) was as follows: guide strand, 5'-AUGAUAUAGACGUUGUGGCUG -3', passenger strand, 5'-GCCACAACGUCUAUAU CAUGG-3'. All other oligonucleotides used in this study are listed in table 1.

Table 1. Sequences of 21 and 22bp siRNAs used in this chapter

siRNA	Sequence	siRNA	Sequence
A	ACGCCAATAAACAAAGAAAG UCGCGGGUUTUGUAUUUCU	J	CCAAUGAUCGAAAAAUUAUU AGGGUUAAGUAGGUUUUUAAU
B	AAAACAAAGAAAGGCCCGG GUUUUGUAUUUCUUCCGGG	L21	GAUGAUAACCGGGCGGGUC CCCUACUAUUUGGCCCGCGC
C	UAAAGAAAGGCCCGGCGCCAU GUUUUCUUUCGGGCGCGG	LL21	GAUGAUAACCGGGCGGGU CCCUACUAUUUGGCCCGCGC
D	AAGGCCGGCGCCAUUGUAUC CUUCCGGGCGCGGUAAGAU	LL22	GAUGAUAACCGGGCGGGUC CCCUACUAUUUGGCCCGCGC
E21	CGGCGCCAUUCUAUCGCTGG GGCGCGCGUAAGAUAGCGG	LR21	AUGAUAACCGGGCGGGUCG CCUACUAUUUGGCCCGCGCA
EL21	CGGCGCCAUUCUAUCGCTGG CGGCGCGCGUAAGAUAGCGG	LR22	GAUGAUAACCGGGCGGGUCG CCCUACUAUUUGGCCCGCGCA
EL22	CGGCGCCAUUCUAUCGCTGG CGGCGCGCGUAAGAUAGCGG		
ER21	GGCGCCAUUCUAUCGCTGGA GGCGCGCGUAAGAUAGCGGAC	M21	GGGCGCGGUCGGUAAAGUUG GGCGCGCGCCAGCCAUUUCAC
ER22	CGGCGCCAUUCUAUCGCTGGA GGCGCGCGUAAGAUAGCGGAC	ML21	CGGCGCGGUCGGUAAAGUUG UGGCGCGCGCCAGCCAUUUCAC
		ML22	CGGCGCGGUCGGUAAAGUUG UGGCGCGCGCCAGCCAUUUCAC
F21	CAUUCUAUCCCGUGGAAGUUG CGUAAGAUAGGCGACCUUCU	MR21	GGCGCGGUCGGUAAAGUUGU GGCGCGCGCCAGCCAUUUCAC
FL21	CAUUCUAUCCCGUGGAAGUUG GGCGUAAGAUAGGCGACCUUC	MR22	GGCGCGGUCGGUAAAGUUGU GGCGCGCGCCAGCCAUUUCAC
FL22	CAUUCUAUCCCGUGGAAGUUG GGCGUAAGAUAGGCGACCUUC		
FR21	AUUCUAUCCCGUGGAAGUUG GGUAAGAUAGGCGACCUUCU	N	GUAAAGUUGUCCAUUUUUUG GCCAUUUCACAAAGGUAAAA
FR22	CAUUCUAUCCCGUGGAAGUUG CGUAAGAUAGGCGACCUUCU		
G	AUCCGCGUGGAAGAUGGAACCG GAUAGGCGACCUUCUACCUUG	O21	GGUGAACUUCGCGCGCGGU GGCCACUUGAAGGCGCGCGG
H21	UGGAAGAUUGGAACCGCUGGAG CGGCUUCUACCUUGGCGGAC	OL21	GGUGAACUUCGCGCGCGGU CGGCCACUUGAAGGCGCGCGG
HL21	UGGAAGAUUGGAACCGCUGGAG CGGCUUCUACCUUGGCGGAC	OL22	GGUGAACUUCGCGCGCGGU CGGCCACUUGAAGGCGCGCGG
HL22	UGGAAGAUUGGAACCGCUGGAG CGGCUUCUACCUUGGCGGAC		
HR21	GGGAAGAUUGGAACCGCUGGAG GACCUUCUACCUUGGCGGAC	OR21	GGUGAACUUCGCGCGCGGU GGCCACUUGAAGGCGCGCGG
HR22	UGGAAGAUUGGAACCGCUGGAG CGGCUUCUACCUUGGCGGAC	OR22	GGUGAACUUCGCGCGCGGU GGCCACUUGAAGGCGCGCGG
I21	UGGAACCGCUGGAGAGGAC UCUACCUUGGCGACCUUCGU	P	CGCGCGCGGUUGUUGUUUG AAGGGCGGCGGCAACACAAA
IL21	UGGAACCGCUGGAGAGGAC UCUACCUUGGCGACCUUCGU	Q	CGCGCGCGGUUGUUGUUUG AAGGGCGGCGGCAACACAAA
IL22	UGGAACCGCUGGAGAGGAC UCUACCUUGGCGACCUUCGU		
IR21	UGGAACCGCUGGAGAGGAC UCUACCUUGGCGACCUUCGU	R21	CGCGCGCGGUUGUUGUUUG GGCGCGCGGCAACACAAAAC
IR22	UGGAACCGCUGGAGAGGAC UCUACCUUGGCGACCUUCGU	RL21	CGCGCGCGGUUGUUGUUUG AAGGGCGGCGGCAACACAAA
		RL22	CGCGCGCGGUUGUUGUUUG AAGGGCGGCGGCAACACAAA
		RR21	CGCGCGCGGUUGUUGUUUG GGCGCGCGGCAACACAAAAC
		RR22	CGCGCGCGGUUGUUGUUUG GGCGCGCGGCAACACAAAAC

RNAi Assay for Firefly Luciferase

One milliliter of CHO-K1 (3×10^5 cells/ml), HeLa (1×10^5 cells/ml), or E14TG2a (2×10^5 cells/ml) cell suspension was inoculated in a 1.5-cm well 24 hours prior to transfection. Cells were transfected with pGL3-Control DNA (1 μ g; Promega) encoding the firefly luciferase gene and pRL-SV40 DNA (0.1 μ g; Promega) encoding the *Renilla* luciferase gene, with or without siRNA against the firefly luciferase gene. Lipofectamine 2000 reagent (Invitrogen) was used for transfection. Cells were harvested 24 hours after transfection and the luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega). In this system, two luciferase (firefly and *Renilla*) genes are expressed in cells simultaneously. The reduction in firefly luciferase activity by the addition of siRNA against the firefly luciferase gene was normalized using the *Renilla* luciferase activity (internal control).

In Vitro Dicer Cleavage

According to the manufacturer's protocol, the 5' ends of the oligonucleotides were labeled with 32 P using T4 kinase and gamma- 32 P-ATP. 32 P-labeled dsRNA was incubated with recombinant human Dicer (Stratagene) in the presence of RNase inhibitor (0.4 units/ μ l; Promega), 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 2.5 mM $MgCl_2$ at 37°C. Dicer digestion was stopped by adding an equal volume of the loading buffer (Ambion gel loading buffer II). Digestion products were size-fractionated using a sequencing gel (15% denaturing polyacrylamide TBE gel with 7M urea) and analyzed using FLA2000 Image Analyzer (Fuji Film).

Detection of Interferon Response

Interferon response in HeLa cells was examined using RT-PCR. Human signal transducer and activator of transcription (STAT)1b, 2',5'-oligoadenylate synthetase (OAS)1-3, and interferon-inducible transmembrane protein (IFNMP)2 were used as interferon-response markers. GAPDH was used as a control. Cells were transfected with pCAGIpuro-EGFP DNA (0.5 μ g) encoding the EGFP and puromycin resistant genes irrespective of dsRNA treatment. 43bp long dsRNA with two 2nt long 3'overhangs was synthesized enzymatically by using *CUGA7 in vitro* siRNA Synthesis Kit (Nippon Genetech) according to the Manufacture's instruction, and its sequences are as follows: 5' UCGACCGUGGCGACGCUGCCGCCAUUCUAUCCGCUGGAAGAUG and 5'UCUUCCAGCGGAUAGAAUGGCGGCAGCGUCGCCACGGUCGAUG. Transfected cells were selected by a 48 hour puromycin (2 μ g/ml) treatment initiated 24 hours after transfection. The total RNA was extracted and purified with TRIZOL (Invitrogen). RNA used was 0.5 μ g. Other details are described in the Manufacture's protocol. A 1 μ l aliquot of the RT reaction mixture was PCR amplified using AmpliTaqGOLD (Invitrogen). After they preheated at 94°C for 10min, the sample was subjected to PCR (95°C (15sec)-55°C (30sec)-72°C (60sec),

18-30 cycles). PCR primers used are as follows: hGAPDH; 5'CAAGGCTGAGAACGGGAAGCT and 5'GACACGGAAGGCCATGCCAGT

hSTAT1b; 5'CAGAGCACAGTGATGTTAGAC and 5'GTGATAGGGTCATGTTTCG TAG, hOAS1; 5'GGAACATGAGGTGGCTGTGCT and 5'TTGGTACCAGTGCTTGACT AG, hOAS2; 5'TCCTAGAGCAGATTGACAGTG and 5'CCAGGCATAGATGGTGAG CAG, hOAS3; 5'GCAACAGCATCAGCAGCTCTG and 5'TGACCTCGAACTGCCGC TCCT, hIFNMP2; 5'GGAACATGAGGTGGCTGTGCT.

Results

Generation of 21 and 22bp Long dsRNAs from 23-30bp Long dsRNAs with Two 2nt Long 3'overhangs through Dicer Digestion

Human Dicer preferentially cleaves off ~22-bp siRNAs from the termini of dsRNA substrates *in vitro* (Zhang et al., 2002). To clarify the possible molecular basis for RNAi by dsRNA longer than the authentic 21bp long siRNA, a series of 21-30bp long dsRNAs with two 2nt long 3'overhangs were chemically synthesized (see figure 4B), end-labeled and subjected to Dicer digestion (figure 1). For simplicity, Xbp long dsRNA with two 2nt long 3'overhangs are hereafter referred to as dsRNA-X. All dsRNAs examined, which include dsRNA-27 (KL27), dsRNA-28 (KL28), dsRNA-29 (KL29) and dsRNA-30 (KL30), were found to be digested from 5'GS or PS ends (figure 1A). Terminal fragment number and size differed considerably according to substrate sequence (table 2). dsRNA-22 and dsRNA-23 were generated from KL27-KL29 as terminal fragments from the PS end, while PS terminal dsRNAs of KL30 included not only dsRNA-21 and 22 but dsRNA-23 as well. GS terminal fragments from KL27-KL30 all contained dsRNA-21(K) and dsRNA-22(KL22). These findings are consistent with the notion that two ends of dsRNA are recognizable with the Dicer PAZ domain but a single Dicer molecule binds to only one of the two ends so as to cleave phosphodiester bonds situated between nucleotides 21 and 22, 22 and 23 or 23 and 24 from the PAZ-bound end (figure 1C). Note that either dsRNA-21 or dsRNA-22 or both are included in the Dicer digests of any dsRNA-X so far examined.

Figure 1B shows that, in contrast to KL27-KL30, shorter dsRNA-X such as KL23 (dsRNA-23), KL24 (dsRNA-24) and KL-25 (dsRNA-25) produces dsRNA-21 (KL21) as a major terminal fragment following longer Dicer treatment; thus possibly, 23bp products in dsRNA-X ($24 < X < 31$) (see figure 1A and table 2) may eventually be converted to dsRNA-21. In contrast, dsRNA-21 and dsRNA-22 appeared virtually completely resistant to Dicer digestion. Indeed, only very weak signals could be detected at the 21nt level in lane for KL22. Similarly, very faint, if any, signals were detected at the 20nt level in lane for KL21 (see asterisks in figure 1B). Weak 22nt signals in the KL23 lane may also be due to Dicer-dependent single-strand cleavage (figure 1B). Accordingly Dicer appears barely capable of cleaving the phosphodiester bond situated within the 2nt 3'overhang or single-stranded RNA. "In-frame" digestion of dsRNA with 2nt long 3'overhangs from its termini has been noted in other dsRNAs (Ui-Tei et al., 2005; Rose et al., 2005).

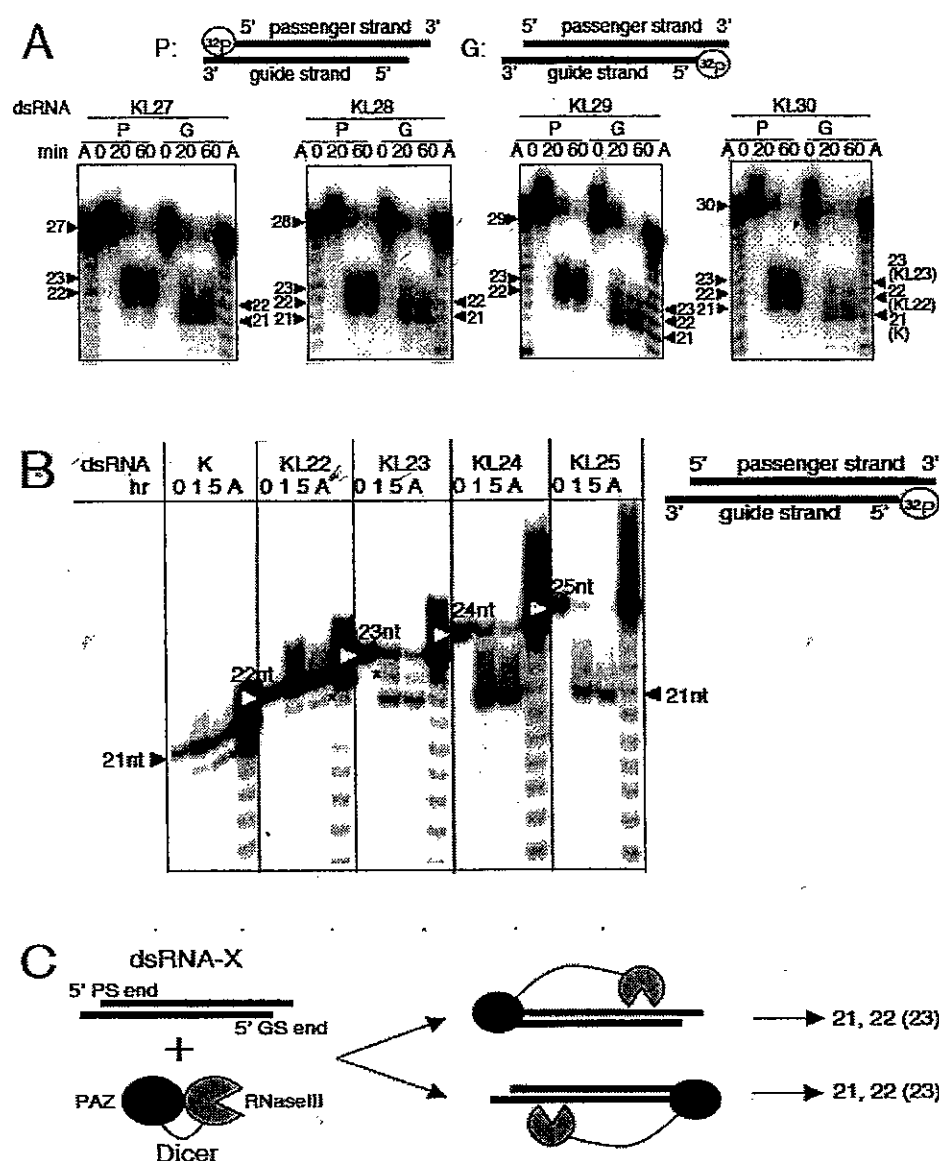


Figure 1. Dicer-dependent cleavage of 21- to 30-bp-long dsRNA with two 2-nt-long 3'-overhangs. (A) dsRNA-27 (KL27), dsRNA-28 (KL28), dsRNA-29 (KL29), and dsRNA-30 (KL30) were end-labeled with ^{32}P , as shown schematically in the upper margin, and subjected to Dicer (human recombinant Dicer, 0.025 units/ μl) digestion followed by electrophoresis in a denaturing polyacrylamide gel. Dicer digestion was carried out for 20 and 60 min. P and G indicate the lanes for the Dicer digests of dsRNA in which the 5'-ends of the guide and passenger strands are ^{32}P -labeled, respectively. The lanes labeled with "A" on the far left and right are for the respective alkaline-treated ^{32}P -labeled passenger and guide strands of each dsRNA, which served as size markers. Bands 21, 22, and 23 found in the G lanes correspond to K, KL22, and KL23. (B) *In vitro* dicing of K (dsRNA-21), KL22 (dsRNA-22), KL23 (dsRNA-24), and KL25 (dsRNA-25). The 5'-ends of these RNAs were labeled with ^{32}P and digested with Dicer (0.05 unit/ μl) at 37°C for 0, 1, or 5 h. The resultant digests were analyzed using denaturing polyacrylamide gel electrophoresis. A, alkaline ladder. Asterisks, possible minor digestion products generated most probably through Dicer cleavage of single-stranded 3'-overhangs. (C) Model for Dicer cleavage of dsRNA-X. The PAZ domain of Dicer binds to one of the two ends of dsRNA-X ($24 < X < 31$) by recognizing a single-stranded 3'-overhang, and then the RNase III domain of Dicer cleaves the dsRNA 21–23 nt from the PAZ-bound end. The cleavage products may change in a sequence-specific manner. Owing to the low-efficiency cleavage of single-stranded RNA, Dicer barely digests dsRNA-22. Similarly, dsRNA-22 and dsRNA-23 may not be generated from dsRNA-23 and dsRNA-24, respectively. dsRNA-23 generated from the first cycle of Dicer digestion of relatively long dsRNA-X ($X > 26$) is further digested to yield dsRNA-21 in the second cycle of Dicer digestion.

Table 2. Dicing products of 21-30bp long dsRNAs

siRNA	³² P-labeled passenger strand	³² P-labeled guide strand
KL30	21, 22, 23	21, 22 (23)
KL29	22, 23	21, 22 (23)
KL28	21, 22, 23	21, 22
KL27	22, 23	21, 22
KL26	NT	NT
KL25	NT	21
KL24	NT	21
KL23	21	21
KL22	Not-digested	Not-digested
K	Not-digested	Not-digested
KR22	NT	NT
KR23	NT	NT
KR24	NT	NT
KR25	NT	NT
KR26	21, 22	21, 22
KR27	21, 22	21, 22
KR28	21, 22	22,
KR29	NT	NT
KR30	21, 22	21, 22

NT = Not tested.

Guidelines for Selecting Highly Effective 22bp Long dsRNA Sequences for Mammalian RNAi

As described above, RNAi due to 23-30bp long dsRNAs may be intimately related to RNAi by dsRNA-21 and dsRNA-22. dsRNA-21 dependent RNAi is highly sequence-dependent and sequence preference rules for highly effective dsRNA-21 have been formulated (Ui-Tei et al., 2004). The relationship between dsRNA-22 sequence and RNAi-inducing activity is but little understood and so accordingly, using the dual luciferase system, examination was made of RNAi activity induction in CHO-K1, HeLa and E14Tg2a cells transfected by thirty-two arbitrarily chosen dsRNAs-22. As summarized in figure 2A, no significant differences due to cell type could be observed. In figure 2A, nucleotide sequences are listed in the order of induced RNAi activity from top to bottom.

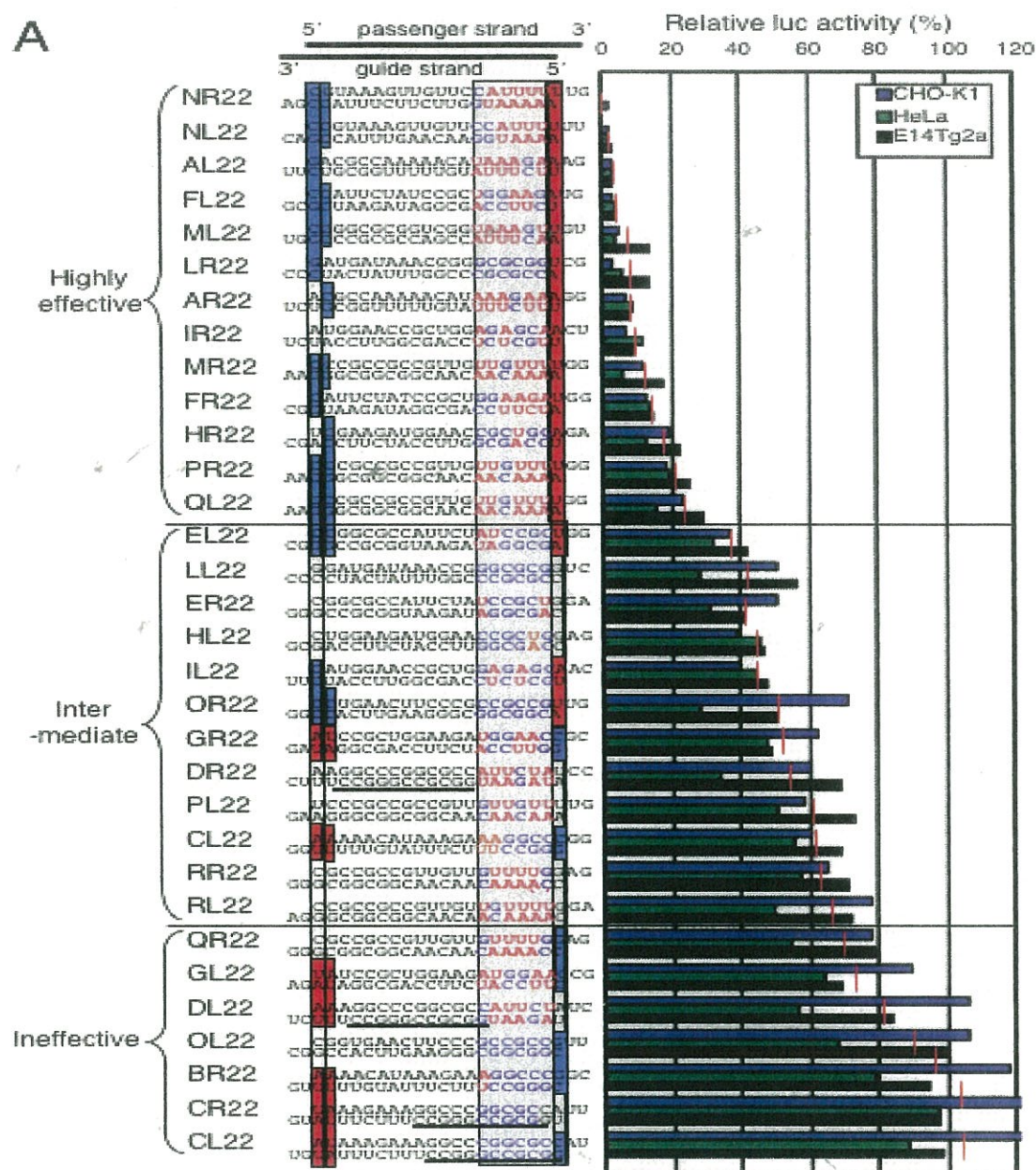


Figure 2. Continued

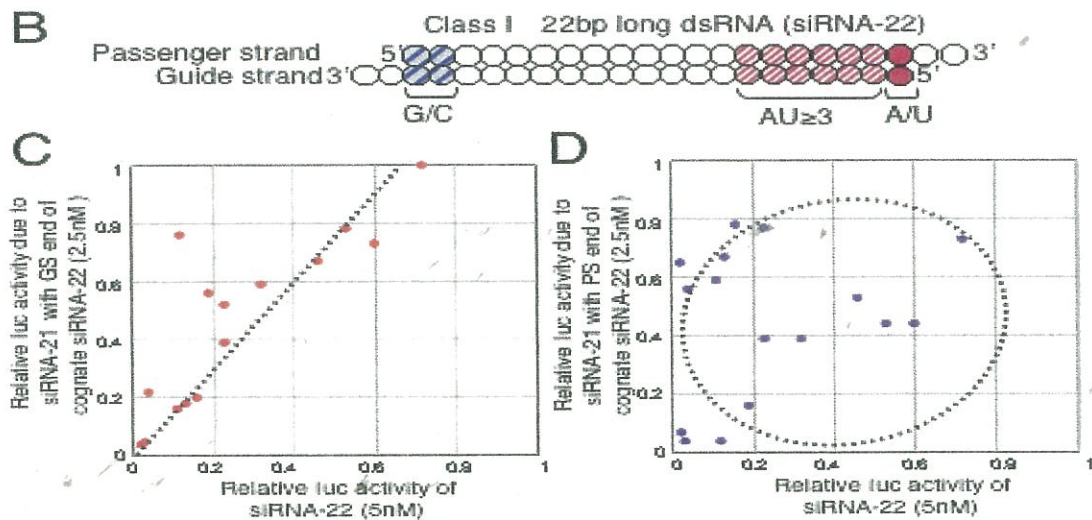


Figure 2. Relationship between the dsRNA-22 sequence and induced RNAi activity. (A) Identity of dsRNA-22 sequences preferably inducing high levels of mammalian RNAi. The locations of the 5'- and 3'-ends of the passenger and guide strands are shown schematically in the upper margin. The firefly luciferase gene was used as a target and the relative luc activity was examined in three mammalian cell lines: CHO-K1, HeLa, and E14Tg2a. The concentration of dsRNA was 5 nM. Red bars indicate the average luc activity in the three cell lines. The sequences of 32 dsRNAs-22 were ordered according to their RNAi-inducing activity from top to bottom. The 32 dsRNAs examined were classified into three groups: highly effective, intermediate, and ineffective. The data obtained from three independent experiments were averaged and are shown. The 7-bp-long 5'-terminal region of the guide strand is boxed and A/U and G/C are colored in red and blue, respectively. Highly conserved 5'-terminal bases are shown on red (A/U) or blue (G/C) backgrounds. Note that in all of the highly effective dsRNAs-22 examined, the 5'-guide strand end was occupied by A/U. In contrast, G or C was found preferentially at the 5'-end and its immediate neighbor of the passenger strand. GC stretches of more than 11 bp are underlined. (B) The sequence requirements for class I dsRNA-22 or siRNA-22 are shown schematically. The 5'-end of the guide strand must be A/U, while G/C is situated at either the 5'-end or its immediate neighbor in the passenger strand. The AU content should be at least 3 in the 6-bp region labeled with red hatched lines. (C) The relationship of the RNAi activity expressed as the relative luciferase activity between siRNA-22 and one of its cognate siRNA-21 sharing the GS 5'-end. The RNAi activity induced by siRNA-21 appeared to be highly positively correlated with the siRNA-22 RNAi activity. (D) The relationship of the RNAi activity expressed as the relative luciferase activity between siRNA-22 and one of its cognate dsRNA-21 sharing the PS 5'-end. Virtually no correlation was observed between the RNAi activity due to siRNA-21 and that due to siRNA-22. Note that in both cases, the concentration of siRNA-21 was 2.5 nM, while that of siRNA-22 was 5 nM.

As with dsRNA-21 (Ui-Tei et al., 2004), dsRNA-22 sequences were classified as highly effective, intermediate and highly inactive groups. Highly effective dsRNAs-22 eliminated firefly luciferase activity at more than 67% following transfection at 5 nM, while only 0-33% reduction in luciferase activity was noted by transfection of 5nM highly ineffective dsRNA-22. The following features observed for dsRNA-22 sequences may serve to discriminate most, if not all, highly effective dsRNAs-22 from those that are ineffective. First, the 5'-end of the guide strand of highly effective dsRNAs-22 was A or U, while highly ineffective dsRNAs-22 preferably possessed G or C. Secondly, in highly effective siRNAs-22, the 5'-end of the passenger strand was preferably G or C, whereas, in ineffective siRNAs, A or U. With highly effective dsRNAs-22, nucleotides immediately adjacent to the 5'-end of the passenger strand were preferably G or C. Thirdly, in most highly effective dsRNAs-22, at least four out of seven nucleotides in the 5'-terminal guide strand were A or U. G/C stretches more than

10bp long appeared to prevent effective RNAi. All these findings are the same or virtually so as those proposed previously as sequence preference rules for the selecting highly effective class I siRNA-21 sequences for mammalian RNAi.

There was no case in which dsRNA-22 belonging to the intermediate and highly inefficient groups demonstrated these features all at the same time. Thus, dsRNA-22 sequences with the following range of features should be capable of inducing highly effective gene silencing in mammalian cells: (i) A/U at the 5' end of the guide sequence (antisense strand); (ii) G/C at either the 5' end or immediately adjacent position of the passenger strand (sense strand) or both; (iii) at least four A/U residues in the 5' terminal one-third of the guide strand and (iv) the absence of any long GC stretch

As described above, dsRNA-21 and dsRNA-22 are the main final Dicer-digestion products of dsRNA-X. They are also responsible for RNAi due to dsRNA-X (see below). Thus, dsRNA-22 and dsRNA-21, respectively, are hereafter referred to as short interfering (si)RNA-22 and siRNA-21 and siRNA-22 with all these features is designated as class I siRNA-22.

Functional Similarity of siRNA-22 to siRNA-21 Sharing the 5'guide-Strand-End Sequence

That siRNA-22 is almost resistant to Dicer digestion may indicate most RNAi activity in cells transfected with siRNA-22 is not to be due to possible terminal siRNA-21 components. siRNA-22 must induce RNAi without producing terminal siRNA-21. As with siRNA-21, the guide sequence of siRNA-22 is quite likely incorporated as a component of active RISC and cleaves target mRNA. To confirm possible similarity between siRNA-21-based RNAi and siRNA-22-based RNAi, comparison was made of RNAi activity by siRNA-22 with that by GS terminal siRNA-21 (figure 2B) or PS terminal siRNA-21 (figure 2C). RNAi activity induced by GS terminal siRNA-21 was found closely correlated to RNAi activity by siRNA-22, whereas hardly any correlation existed between PS-terminal-siRNA-21-dependent RNAi and siRNA-22-dependent RNAi. Thus possibly, with transfection, the guide strand of siRNA-22 may be incorporated into RISC which then cleaves target mRNA in accordance with a mechanism essentially the same as that for siRNA-21-dependent RNAi. In figure 2C,D, it should be noted that siRNA-21 is present at 2.5nM while siRNA-22, at 5nM. In siRNA-21-dependent RNAi, 3' end of the guide strand appears less important (Martinez and Tuschl, 2004) and modification at the 3' end may not seriously affect RNAi activity (Elbashir et al., 2001c). A nucleotide at the 3' very end of the guide strand may accordingly not be essential for effective RNAi and siRNA-22 and siRNA-21 sharing the same 5'GS end may possibly give rise to the same degree of RNAi activity.

RNAi due to a 1:1 Mixture of siRNAs-21

RNAi activity of dsRNA-X ($22 < X < 31$) might be the same or nearly so as the combined RNAi activity of terminal siRNAs generated through Dicer digestion. As a first step to assess the feasibility of this notion, using the dual luc system, comparison was made of relative luc activity observed on transfection of a 1:1 mixture of two siRNAs-21, siRNA-21A and siRNA-21B, and that estimated based on relative luc activity obtained by single transfection experiments of siRNA-21A and siRNA-21B (figure 3). We assume that any change in relative luc activity directly reflects that in intact target-mRNA fraction within cells.

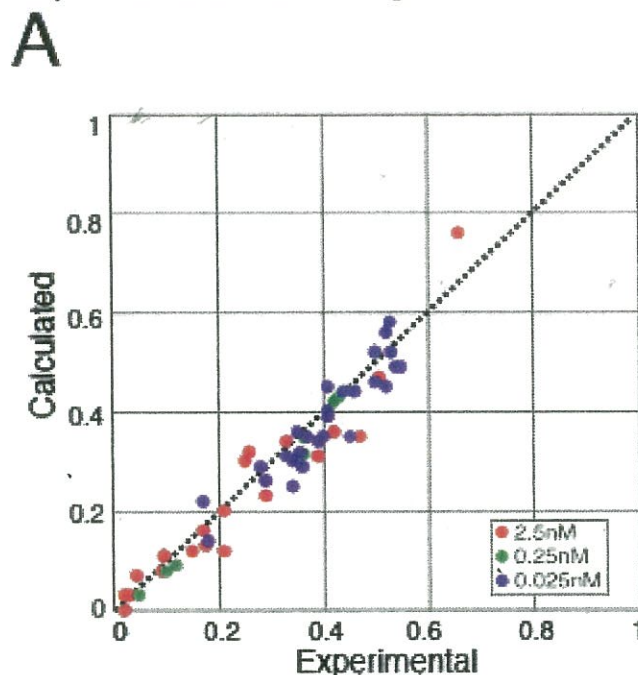


Figure 3. RNAi due to a 1:1 mixture of two 21-bp-long siRNAs. The experimental and calculated values for a 1:1 mixture of siRNAs were compared. The calculated values (fraction) were obtained after normalization and subsequent multiplication of the relative luciferase activities, presumably corresponding to the relative *luc* mRNA concentrations. The following combinations of siRNAs-21 at 2.5, 0.25, or 0.025 nM each were used: 2.5 nM (siRNA-O and P, D and E, D and F, E and F, E and G, and F and G), 0.25 nM (siRNA-E and F, P and Q, Q and R, E21 and EL21, EL21 and ER21, F21 and FL21, F21 and FR21, H21 and HL21, H21 and HR21, I21 and IL21, I21 and IR21, L21 and LL21, L21 and LR21, M21 and ML21, M21 and MR21, O21 and OL21, O21 and OR21, R21 and RL21, R21 and RR21) and 0.025 nM (siRNA-E and F, M and N, L and M, E and Q, F and L, F and Q, L and P, L and Q, A and F, A and J, F and M, F and J, L and J, and J and K). Experimental data were obtained by transfecting a 1:1 mixture of siRNAs at 2.5, 0.25, or 0.025 nM each. The values were normalized using the value for non-siRNA-treated cells.

To eliminate possible non-specific effect of siRNA concentration on RNAi activity, in the case of single-siRNA transfection experiments, the same concentration of siRNA (siGY441) as for the gene coding for EGFP and hence unrelated to luc was mixed with tester siRNA (siRNA-21A or siRNA-21B) and subjected to transfection. The combined RNAi effect of siRNA-A and siRNA-B (theoretical values) was estimated to be ab , where a and b , respectively, are relative luc activity observed after siRNA-A and siRNA-B treatment. Figure 3 shows the results obtained at siRNA concentration of 2.5nM, 0.25nM and 0.025nM. In all cases, theoretical and experimental values were virtually the same. In addition, no effect of siRNA-sequence overlap was observed (data not shown). RISC may thus be considered to

attack target mRNA independently of each other and the relative intact-target-mRNA fraction in cells transfected with a siRNA-21 mixture is estimatable from the RNAi activity of individual siRNA components.

RNAi Induced by Transfection of 23-30bp Long dsRNA with Two 2nt-Long 3'overhangs

AS with siRNA-21 and siRNA-22, RNAi induced dsRNA-X (22<X<31) significantly varied according to the dsRNA-X sequence (figure 4A). Dicer has shown to cleave siRNA-30 from GS and PS ends with nearly the same efficiency (Ui-Tei et al., 2005).

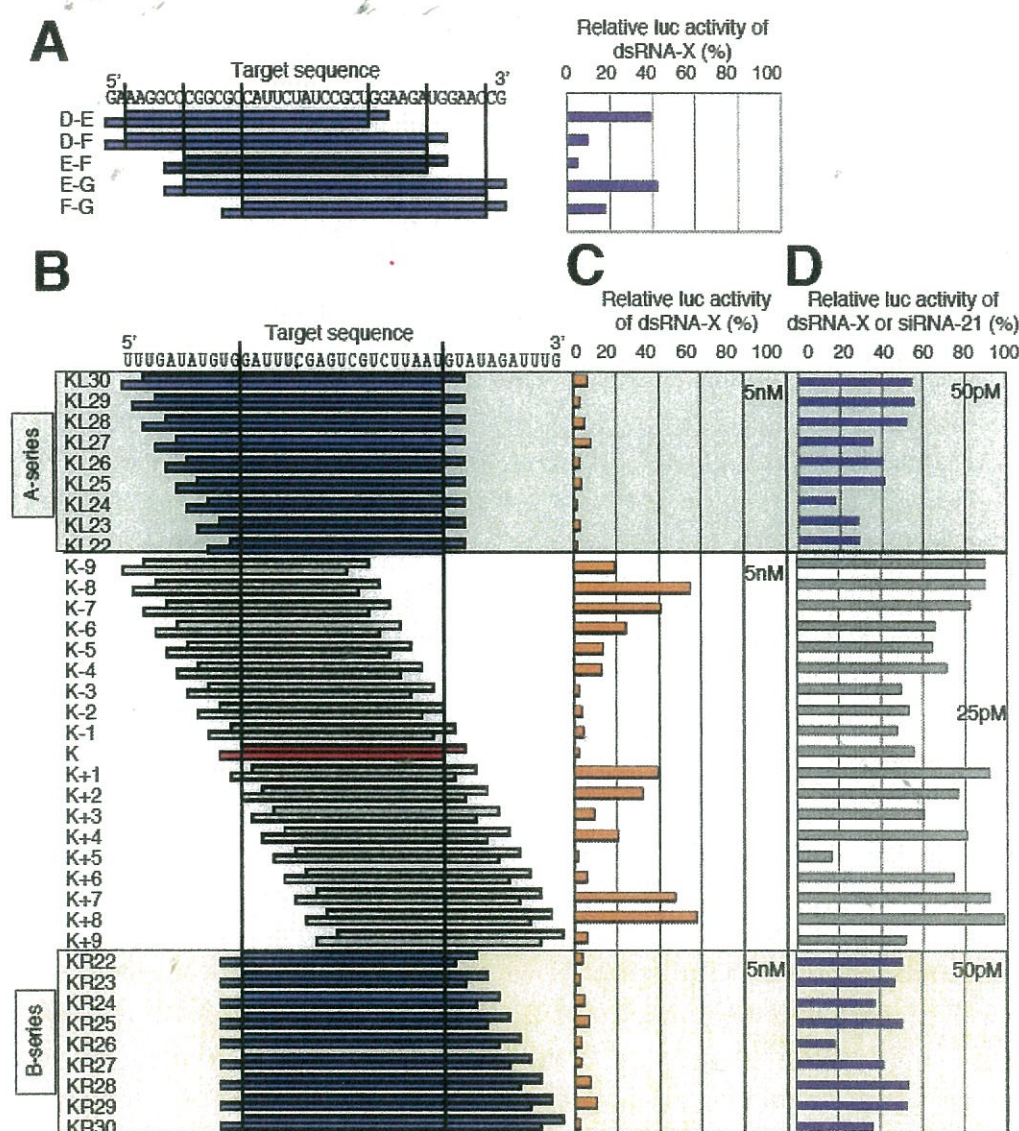


Figure 4. RNAi effects of 22- to 30-bp-long dsRNAs with two 2-nt-long 3'-overhangs. (A) The luc activity change induced by 5 nM dsRNAs-X. The nucleotide sequences for the passenger strand of dsRNA-X are shown in the left margin. (B) The sequences of series A and B dsRNA-X and cognate siRNAs-21 are shown using the target mRNA sequence. (C) The relative luc activity change induced by the transfection of dsRNA shown in (B) at 5 nM in HeLa cells. (D) The relative luc activity change induced by dsRNA shown in (B) at 50 or 25 pM.

Thus, to detect possible contribution of terminal siRNAs to the RNAi activity of dsRNA-X, series A and B of dsRNA-X were constructed and subjected to dual luc assay using HeLa cells. Any dsRNA belonging to series A may produce the same class-I siRNA-21 or siRNA-22 from the GS end through Dicer digestion. Similarly, an identical class-I siRNA 21bp in length appears produced from the PS end of any of series-B dsRNAs. As shown in figure 4B,C, at 5nM, all dsRNAs of the two series exhibited RNAi activity as high as class I siRNAs. At 50pM, series A and B dsRNAs showed intermediate RNAi (40-80% luc activity reduction). Combined RNAi activity due to terminal siRNAs cannot be easily determined without precisely knowing the ratio of terminal siRNAs and concentration dependency of RNAi activity. A comparison was thus made of RNAi activity due to 50 pM dsRNA-X (figure 4D) and the combined RNAi activity of 25 pM siRNAs-21 expected to be produced from GS and PS ends (figure 4D). It was found that they are very similar, if not identical, to each other (figure 5). dsRNA-X-dependent RNAi may thus be considered primarily dependent on the combined RNAi activity of terminal siRNAs.

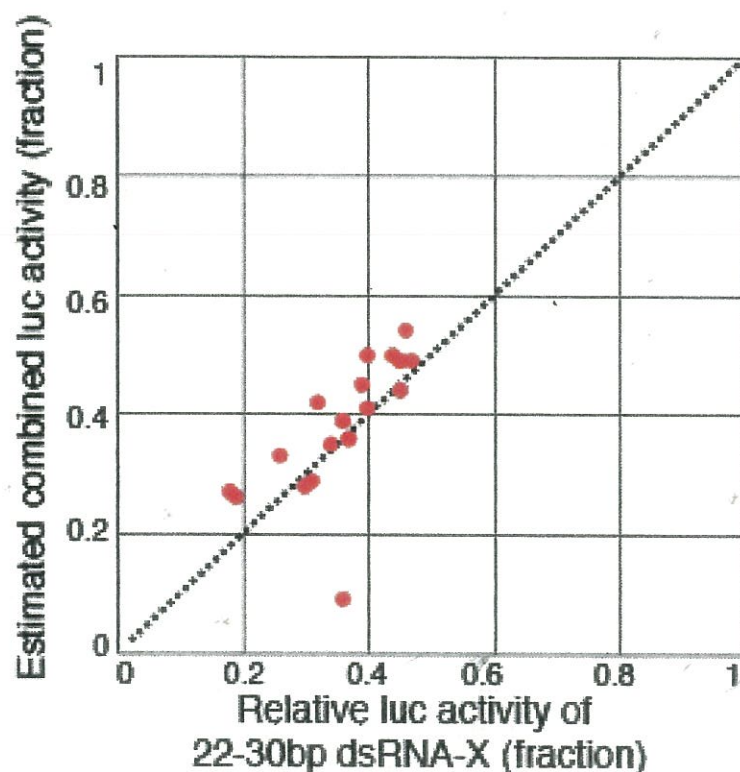


Figure 5. Correlation between the RNAi effects due to 23- to 30-bp-long dsRNA-X and those due to a 1:1 mixture of cognate siRNAs expected to be derived from both ends of dsRNA-X. All the nucleotide sequences are shown in figure 4B.

Poor Induction of Interferon Response by dsRNA-X (20<X<31) Transfection

dsRNA-transfection-dependent change in interferon response in HeLa cells was examined by monitoring that in the expression of interferon response marker genes such as human OAS1-3, STAT1b and IFNMP2 (Sledz et al, 2003). As described previously (Ui-Tei

et al., 2005), HeLa cells transfected with 43bp long dsRNA exhibited considerable interferon response. In contrast, interferon response induced by dsRNA-X ($20 < X < 31$) was very low, if any and appeared virtually the same as that by the authentic siRNA 21bp long.

Discussion

The present study showed 23 to 30bp long dsRNAs with two 2nt long 3'overhangs (dsRNA-X ($22 < X < 31$)) to be digested by Dicer to yield 21 and 22 bp long siRNAs from both GS and PS ends. The RNAi activity of 23 to 30bp long dsRNA was basically the same as the combined RNAi activity of terminal siRNAs, each functioning independently of each other in the RNAi reaction. As with 21bp long siRNA, siRNA 22bp in length was resistant to Dicer digestion and required similar sequence conditions for highly effective RNAi.

Design for Highly-Effective dsRNAs 23 to 30bp Long, Associated with Reduced off-Target Effect

Our in vitro Dicer digestion experiments (see figure 1) showed that Dicer attacks termini of dsRNA-X ($22 < X < 31$) eventually to produce two siRNA-21 and/or siRNA-22 sets, GS and PS sets (figure 7). The first cycle of Dicer digestion (see figure 1A) may produce not only siRNA-21 and siRNA-22 but dsRNA-23 as well, all of which share in common either GS or PS ends of the parental dsRNA-X. siRNA-21 and siRNA-22 are almost resistant to second cycle of Dicer digestion (see figure 1B) but dsRNA-23 is susceptible to Dicer digestion to be converted to two terminal siRNAs-21, one of which is identical in sequence to the siRNA-21 generated by the first cycle of Dicer digestion. Thus, the GS and PS sets obtained from dsRNA-X ($24 < X < 31$) include two siRNAs-21 and one siRNA-22 at most. Since Dicer cannot digest 3'overhangs or single-stranded RNA effectively, neither siRNA-22 nor dsRNA-23 may be the main Dicer digestion products of dsRNA-23 and dsRNA-24, respectively.

The RNAi activity induced by dsRNA-X ($22 < X < 31$) was essentially the same as the combined RNAi activity induced by GS and PS sets of siRNAs 21 and/or 22 bp long. Class I siRNA-22 may possess RNAi inducing activity almost identical to that of cognate class I siRNA-21 and both types of siRNAs are capable in most cases of inducing highly effective RNAi. Thus, dsRNA-X ($22 < X < 31$) should exhibit the highest levels of RNAi, if all members of its GS and PS siRNA sets belong to class I. These dsRNA-X are hereafter referred to as class I dsRNA-X. Sequence rules for class I dsRNA-X are shown in figure 7. They vary as a function of the size, X. If RNAi activities induced by class I siRNAs 21 or 22bp long are presumed to be roughly identical to each other, then RNAi activity of class I is considered to be virtually identical to that of the authentic class I siRNA-21. However, it should be noted that specific RNAi activity of class I dsRNA-X may gradually decrease with increasing X, since each dsRNA-X molecule can produce only one siRNA molecule.

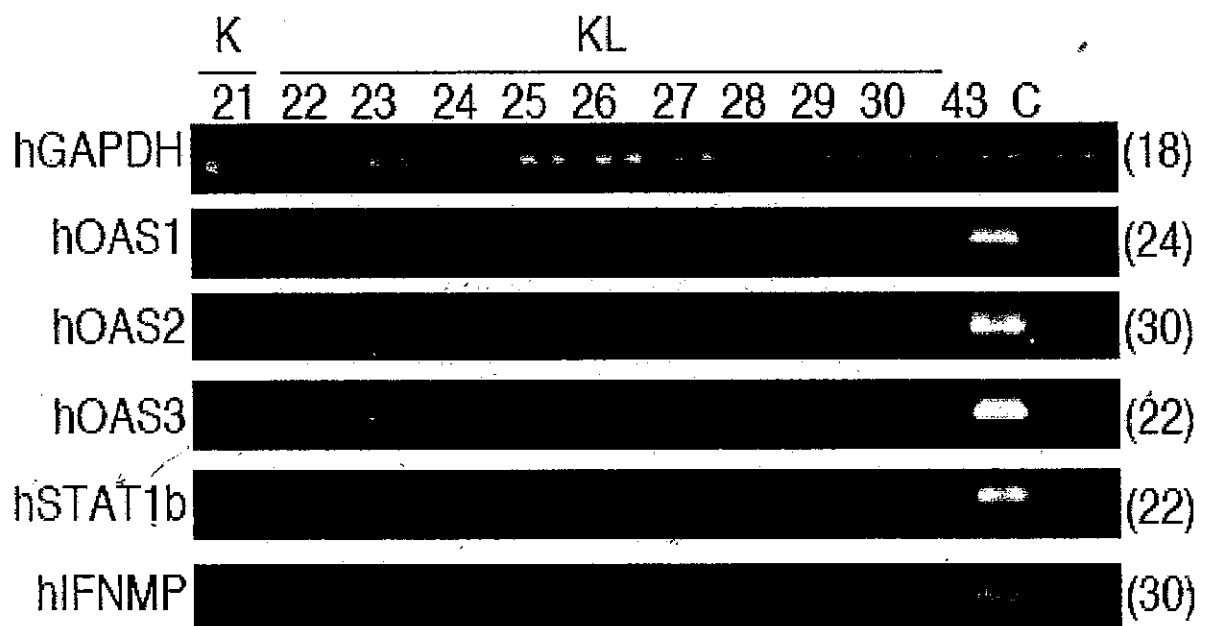


Figure 6. dsRNA-X (20<X<31)-dependent interferon response. Change in expression of interferon response markers (hOAS1-3, hSTAT1b and hIFNMP2) in HeLa cells were examined using RT-PCR. hGAPDH, control. Numerals in parentheses, Amplification cycles. C, control transfection of pCAGIpuroEGFP vector alone. In all cases except for 43-bp-long dsRNA, dsRNA-X (20<X<31) induced only marginal, if any, interferon response.

siRNA may suppress the activity of genes other than the target, which is called "the off-target effect". A part of the off-target effect due to sequence-specific mRNA cleavage might be eliminated computationally (Naito et al., 2004). Off-target effects are also induced by siRNA-dependent translational inhibition (Lin et al., 2005) and, in this case, a 7bp sequence from the 5'GS end is considered to be essential for mRNA recognition (Birmingham et al., 2006; Lim et al., 2005). One of the simplest ways to eliminate the off-target effect is to use a mixture of siRNAs targeting at the same gene but different sequences. Since, as shown in figure 7, Dicer digestion eventually generates guide strands, of which the 5'ends are situated a few to several nucleotides off, class I dsRNA-X may have a lower reduced off-target effect than the authentic class I siRNAs 21 or 22bp long. We need to synthesize 80 phosphodiester bonds as a total when a 1:1 mixture of two authentic siRNA 21bp long is used, while dsRNA-X 25 synthesis requires only 48 phosphodiester bond formation, indicating that dsRNA-X possesses much higher specific activity for off-target elimination than a 1:1 mixture of siRNAs.

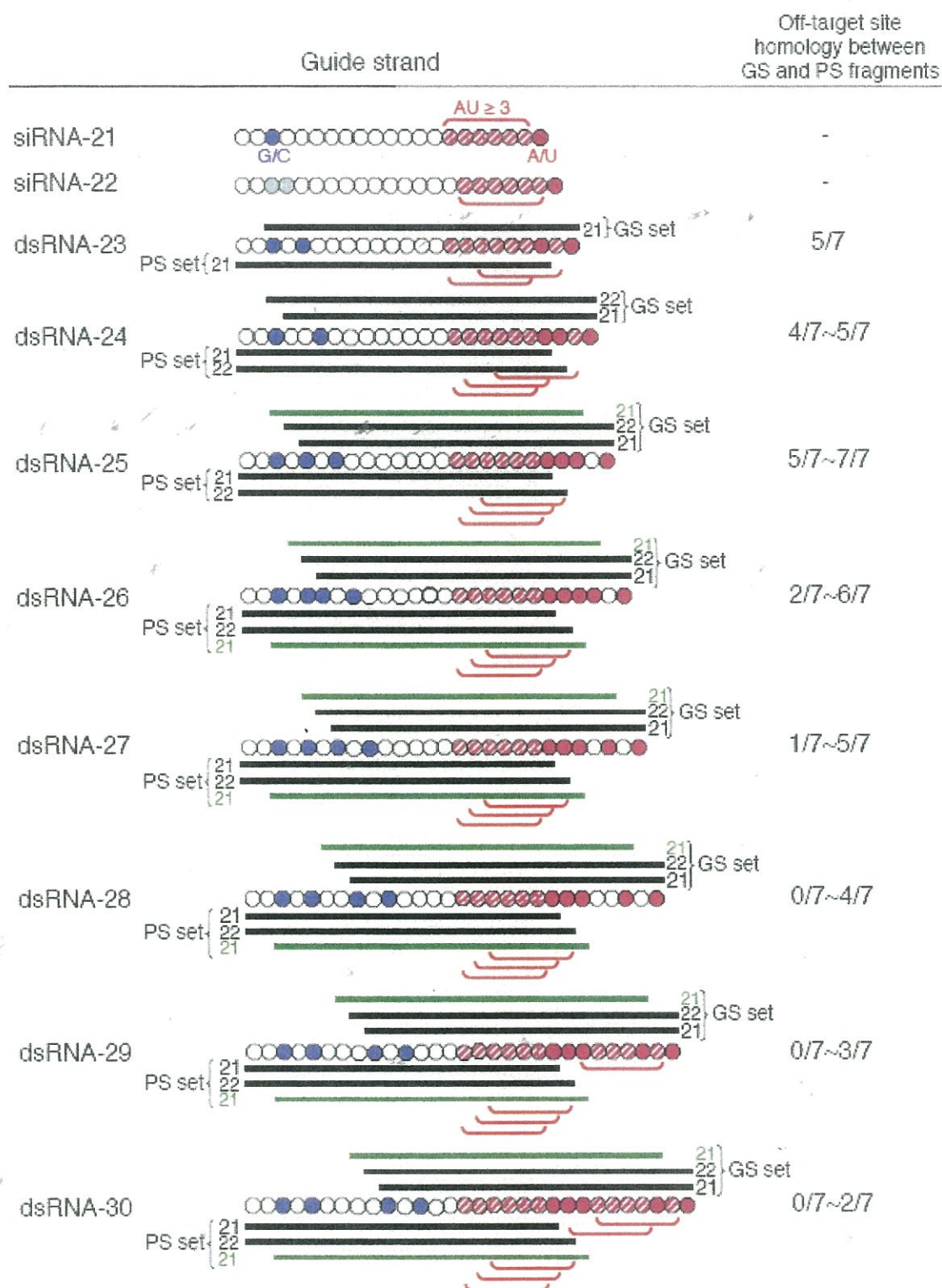


Figure 7. Sequence requirements of class I 23- to 30-bp-long dsRNA-X. Only nucleotide sequences for the guide strands of class I dsRNA-X are shown with those for class I siRNA-21 and class I siRNA-22. Off-target site homology, homology within the 7-nt-long region from the GS end. Lines over and under the guide strand indicate the predicted Dicer-cleavage products from the GS and PS 5'-ends, respectively. The dsRNA-23 generated by the first cycle of Dicer digestion is presumed to be converted into a pair of siRNA-21, one of which is identical in sequence to the siRNA-21 in the products of the first Dicer digestion cycle; the size and location of the remaining siRNA-21 are shown by the green line. Red circles, A or U. blue circles, G or C. Each 6-bp-long U-shaped red frame indicates a region in which more than two nucleotides should be A or U. Red hatched circles, residues rich in A/U based on frequency.

Conclusion

Examination was made of RNAi effects exerted by 22 to 30-bp long dsRNAs with two 2nt long, 3'overhangs (dsRNAs-X) in mammalian cells. RNAi activity induced by these dsRNAs appeared to be the combined RNAi activity of terminal 21- and 22bp long siRNAs, both resistant to Dicer digestion. Sequence preference rules for highly effective siRNA 22bp in length were formulated and a means for designing sequences of dsRNA-X capable of bringing about highly effective RNAi with reduced off-target effect is proposed.

Acknowledgments

We thank A. Juni, R. Numazawa, and N. Fukunishi for technical assistance. This work was partially supported by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan to K.S. and K.U.-T.

References

- Amarzguioi, M. and Prydz, H. (2004) An algorithm for selection of functional siRNA sequences. *Biochem. Biophys. Res. Commun.*, 316, 1050-1058.
- Billy, E., Brondani, V., Zhang, H., Muller, U., and Filipowicz, W. (2001) Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. *Proc. Natl. Acad. Sci. USA.*, 98, 14428-14433.
- Birmingham, A., Anderson, E.M., Reynolds, A., Ilsley-Tyree, D., Leake, D., Fedorov, Y., Baskerville, S., Maksimova, E., Robinson, K., Karpilow, J., Marshall, W.S. and Khvorova, A. (2006) 3'UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat. Methods*, 3, 199-204.
- Carmell, M.A., and Hannon, G.J. (2004) RNaseIII enzymes and the initiation of gene silencing. *Nat. Struct. Mol. Biol.*, 11, 214-218.
- Doi, N., Zenno, S., Ueda, R., Ohki-Hamazaki, H., Ui-Tei, K., and Saigo, K. (2003) Short-interfering-RNA mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. *Curr. Biol.*, 13, 41-46.
- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001a) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, 24, 494-498.
- Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001b) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.*, 15, 188-200.
- Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W. and Tuschl, T. (2001c) Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J.*, 20, 6877-6888.
- 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.

- Hannon, G.J. and Rossi, J.J. (2004) Unlocking the potential of the human genome with RNA interference. *Nature*, 431, 371-378.
- Holen, T., Amarzguioui, M., Wiiger, M.T., Babaie, E., and Prydz, H. (2002) Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Res.*, 30, 1757-1766.
- Kim, D.H., Behlke, M.A., Rose, S.D., Chang, M-S, Choi, S. and Rossi, J.J. (2005) Synthetic dsRNA dicer substrates enhance RNAi potency and efficacy. *Nat. Biotech.*, 23, 222-226.
- Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Shclter, J.M., Castle, J., Bartel, D.P., Linsley, P. S. and Johnson, J.M. (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, 433, 769-773.
- Lin, Z., Ruan, Z., Anderson, M.G., McDowell, J.A., Kroeger, P.E., Fesik, S.W. and Shen, Y. (2005) siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. *Nucleic Acids Res.*, 33, 4527-4535.
- Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004) Argonaute-2 is the catalytic engine of mammalian RNAi. *Science*, 305, 1437-1441.
- Martinez, J., Ratkaniowska, A., Urlaub, H., Lührmann, R. and Tuschl, T. (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell*, 110, 563-574.
- Moffat, J., Grueneberg, D.A., Yang, X., Kim, S.Y., Kloepper, A.M., Hinkle, G., Riqani, B., Eisenhaure, T.M., Luo, B., Grenier, J.K., Carpenter, A.E., Foo, S.Y., Stewart, S.A., Stockwell, B.R., Hacohen, N., Harn, W.C., Lander, E.S., Sabatini, D.M. and Root, D.E. (2006) A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell*, 124, 1283-1298.
- Mello, C.C. and Conte, D., Jr. (2004) Revealing the world of RNA interference. *Nature*, 431, 338-342.
- 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.
- 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.
- Rose, S.D., Kim, D.-H., Amarzguioui, M., Heidel, J.D., Collingwood, M.A., Davis, M.E., Rossi, J.J., and Behlke, M.A. (2005) Functional polarity is introduced by Dicer processing of short substrate RNAs. *Nucleic Acids Res.*, 33, 4140-4156.
- Silva, J.M., Li, M.Z., Chang, K., Ge, W., Golding, M.C., Rickles, R.J., Siolas, D., Hu, G., Paddison, F.J., Schlabach, M.R., Sheth, N., Bradshaw, J., Burchard, J., Kulkarni, A., Cavet, G., Sachidanandam, R., McCombie, W.R., Cleary, M.A., Elledge, S.J., and Hannon, G.J. (2005) Second-generation shRNA libraries covering the mouse and human genomes. *Nat. Genet.*, 37, 1281-1288.
- Sledz, C.A., Holko, M., de Veer, M.J., Silverman, R.H. and Williams, B.R.G. (2003) Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol.*, 5, 834-839.
- Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, J. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science*, 305, 1434-1437.
- Stark, G.R., Kerr, I.M., Williams, B.R., Silverman, R.H., and Schreiber, R.D. (1998) How to respond to interferons. *Annu. Rev. Biochem.*, 67, 227-264.

- Ui-Tei, K., Naito, Y. and Saigo, K. (2006) Guidelines for the selection of functional siRNA sequences for mammalian RNAi. *Method. Mol. Biol.*, 2, 201-216.
- 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.
- Ui-Tei, K., Zenno, S., Juni, A. and Saigo, K. (2005) RNAi induced in mammalian and *Drosophila* cells via transfection of dimers and trimers of short interfering RNA. *J. RNAi Gene Silencing*, 1, 79-87.
- Ui-Tei, K., Zenno, S., Miyata, Y., and Saigo, K. (2000) Sensitive assay of RNA interference in *Drosophila* and Chinese hamster cultured cells using firefly luciferase gene as target. *FEBS Lett.*, 18, 79-82.
- Vermeulen, A., Behlen, L., Reynolds, A., Wolfson, A., Marshall, W., Karpilow, J. and Khvorova, A. (2005) The contributions of dsRNA structure to Dicer specificity and efficiency. *RNA*, 11, 674-682.
- Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002) Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J.*, 21, 5875-5885.