© 2002 Nature Publishing Group http://biotech.nature.com

and ES cells, the resulting tetraploid cells containing thymocyte-derived nuclei displayed a growth morphology and behavior indistinguishable from those of ES cells[5]. In the assay of Håkelien *et al.*, several obstacles (such as size, stability, nuclear import, and time) may limit delivery of extract proteins to the target nucleus and therefore limit nuclear reprogramming. For example, it is possible that the modifications observed in the reprogrammed fibroblasts were initiated by dominant transcription factor(s) present in the protein extract and that, given exposure to a greater complexity of factors, more extensive reprogramming would result. Recent experiments have demonstrated that overexpression of the transcription factor Msx1 can initiate dedifferentiation of multinucleated myotubes to a more plastic single nucleated precursor cell type[11]. Further experiments using the reprogramming technology of Håkelien *et al.* should clarify these issues.

Several laboratories have reported the transdifferentiation of stem cells to tissues outside their normal developmental lineage. For instance, hematopoietic stem cells can apparently give rise to cells of the liver and central nervous system, whereas neuronal stem cells can change into hematopoietic cells[9,10]. However, recent experiments indicate that the isolation and culture of adult stem cells or cell fusion events may explain the unexpectedly high level of plasticity of adult stem cells[6–8]. These observations underscore how poorly cellular plasticity is understood and emphasize the limitations of reprogramming cell lineages without understanding the underlying biological processes.

The Håkelien assay involves extensive cell manipulation to induce reprogramming. Consequently, the applicability of this technology in producing reprogrammed cell lines for therapeutic purposes remains undetermined. A large amount of additional data are required before such a system could be applied to the generation of stem cells to be used directly in cell replacement therapies for human patients.

Setting these reservations aside, the authors' assay provides a potentially powerful system for analyzing nuclear reprogramming events as they occur *in vitro*. It will be exciting to apply similar methodologies to the analysis of other systems, including ES cells, EG cells, and various cancer cell types. Such analyses may allow identification of molecules central to biological processes as diverse as the establishment of pluripotency, the control of cell dedifferentiation, and the onset of cancer.

1. Håkelien, A.M. *et al. Nature Biotechnology* **20,** 460–466 (2002).
2. Landsverk, H.B. *et al. EMBO J.* **3**, 384–389 (2002).
3. Kikyo, N. *et al. Science* **289,** 2360–2362 (2000).
4. Rideout, W.M., 3rd *et al. Science* **293**, 1093–1098 (2001).
5. Tada, M. *et al. Curr. Biol.* **11,** 1553–1558 (2001).
6. Terada, N. *et al. Nature* **416**, 542–545 (2002).
7. Ying, Q., Nichols, J., Evans, E.P. & Smith, A.G. *Nature* **416,** 545–548 (2002).
8. Morshead, C.M. *et al. Nat. Med.* **8**, 268–273 (2002).
9. Lagasse, E. *et al. Nat. Med* **6**, 1229–1234 (2000).
10. Krause, D.S. *et al. Cell* **105**, 369–377 (2001).
11. Odelberg, S.J. *et al. Cell* **103**, 1099–1109 (2000).

Expanding small RNA interference

Several reports describe vector systems capable of producing small interfering RNAs for downregulating gene expression in mammalian cells.

Thomas Tuschl

Small interfering RNAs (siRNAs) are powerful sequence-specific reagents designed to suppress the expression of genes in cultured mammalian cells through a process known as RNA interference (RNAi)[1–3]. Until recently, two limitations of siRNA gene-targeting experiments have been the transient nature of siRNA transfer into cells (achieved by such classic methods as liposome-mediated transfection, electroporation, or microinjection) and the requirement for chemical or enzymatic synthesis of siRNAs before application. In this issue and in *Science* and *Genes & Development*, several independent research groups now describe alternative approaches for producing cells in which specific genes have been targeted by intracellular expression of siRNAs from plasmid DNA in an attempt to address these limitations. The data indicate that siRNA will become a widely adopted and useful tool for the analysis of loss-of-function phenotypes that develop over longer time periods[4–8].

Molecules of siRNA are 21- to 23-nucleotide RNAs, with characteristic 2- to 3-nucleotide 3′-overhanging ends resembling the RNAse III processing products of long double-stranded RNAs (dsRNAs) that normally initiate RNAi. When introduced into a cell, they assemble with yet-to-be-identified proteins of an endonuclease complex (RNA-induced silencing complex), which then guides target mRNA cleavage. As a consequence of degradation of the targeted mRNA, cells with a specific phenotype characteristic of suppression of the corresponding protein product are obtained. The small size of siRNAs, compared with traditional antisense molecules, prevents activation of the dsRNA-inducible interferon system present in mammalian cells. This avoids the nonspecific phenotypes normally produced by dsRNA larger than 30 base pairs in somatic cells.

Intracellular transcription of small RNA molecules can be achieved by cloning the siRNA templates into RNA polymerase III (Pol III) transcription units, which normally encode the small nuclear RNA (snRNA) U6 or the human RNAse P RNA H1. Two approaches have been developed for expressing siRNAs: in the first, sense and antisense strands constituting the siRNA duplex are transcribed by individual promoters (Fig. 1A)[5,6]; in the second, siRNAs are expressed as fold-back stem–loop structures that give rise to siRNAs after intracellular processing (Fig. 1B)[4,7,8]. The endogenous expression of siRNAs from introduced DNA templates is thought to overcome some limitations of exogenous siRNA delivery, in particular the transient loss of phenotype.

U6 and H1 RNA promoters are members of the type III class of Pol III promoters[9]. These promoters are unusual in that almost all their elements, with the exception of the first transcribed nucleotide (+1 position), are located upstream of the transcribed region, so that almost any inserted sequence shorter than 400 nucleotides can be transcribed. They are therefore ideally suited to the expression of ~21-nucleotide siRNAs or ~50- nucleotide RNA stem–loops. The U6 promoter and the H1 promoter are different in size but contain the same conserved sequence elements or protein-binding sites[10]. The +1 nucleotide of the U6-like promoters is always guanosine, whereas it is adenosine for H1. Interestingly, changing the +1 adenosine to uridine, cytidine, or guanosine within H1-expressed stem–loop sequences does not seem to affect gene silencing, indicating that H1 promoters may be more flexible than U6 promoters in

Thomas Tuschl is an EMBO young investigator, Department of Cellular Biochemistry, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany (ttuschl@gwdg.de).

© 2002 Nature Publishing Group http://biotech.nature.com

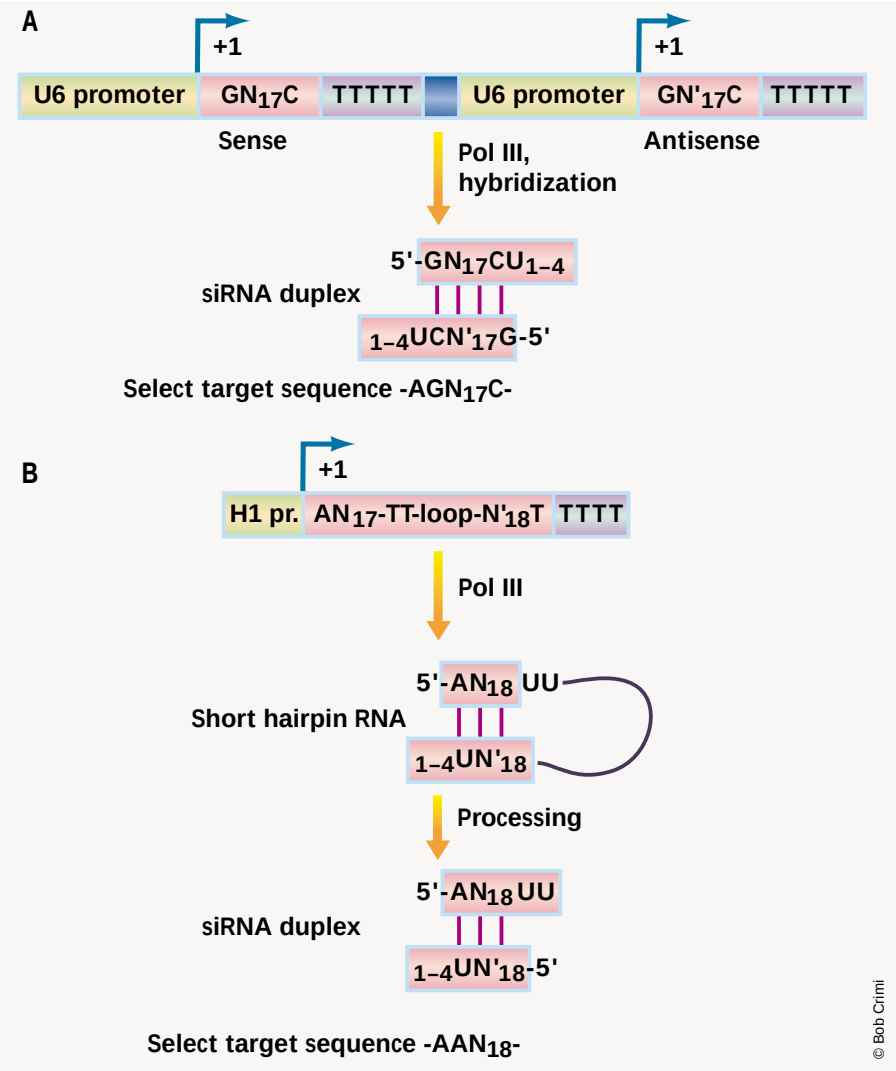

Figure 1. Endogenous expression of siRNAs. (A) Expression cassette for sense and antisense siRNAs using the U6 snRNA promoter[5,6]. The 250 bp U6 snRNA promoter is illustrated as a yellow box, the Pol III terminator signal composed of a run of thymidines is shown as a light purple box, and the spacer between the sense and antisense (indicated with a prime symbol) expression element is shown as a blue box. The siRNA elements are highlighted in pink. The target site, preferably selected for optimal vector design, is indicated at the bottom. (B) H1 RNA–based Pol III cassette for expressing hairpin RNAs that are subsequently processed to siRNAs[4]. The H1 RNA Pol III promoter is only 100 bp in size, but contains all the essential sequence motifs present in the U6 snRNA promoter[10]. Hairpin RNAs with gene silencing properties were also obtained by using a U6 promoter[8]. In this case, transcript synthesis was initiated with a +1 guanosine and the 3′ end of the sense strand was joined by a loop with the antisense strand.

regard to +1 sequence changes or may be able to initiate transcription at the first downstream purine nucleotide encoded by the template DNA. RNA transcription is terminated when Pol III encounters a run of four or five thymidines after incorporation of several uridine residues[10].

As was expected from previous analysis of RNAi, co-expression of sense and antisense siRNAs mediated silencing of target genes, whereas expression of sense or antisense siRNA alone did not greatly affect target gene expression[5–7]. More surprising was the finding that DNA constructs encoding 19-bp stem–loop sequences with 3′-overhanging uridines were as effective at silencing target genes as were the synthetic siRNAs[4,7], and that even blunt-ended duplexes with up to 29 base pairs were able to mediate RNAi in cultured cells[8]. Although the size, orientation, and sequence of the loop appear to be important, endogenously expressed RNA hairpins connected by the 3′ end of the sense strand and the 5′ end of the siRNA antisense strand by a 9-nucleotide loop sequence were so rapidly processed to siRNAs that the 49-nucleotide precursor was barely detectable[4]. This finding also suggests that exogenous delivery of RNA hairpins may represent an alternative to exogenous delivery of siRNA duplexes, as enzymatic RNA hairpin synthesis would require only a single DNA template[8]. Intracellular processing of hairpin RNAs appears to require Dicer RNase III because Dicer suppressed cells do not support hairpin-mediated target gene silencing[8].

Suppression of protein levels by exogenous siRNAs is transient: levels of the targeted protein in siRNA-treated cells typically recover between five and seven days after siRNA transfection, that is, after seven to ten rounds of cell division (T.T. *et al.*, unpublished observations). In three of the new studies, the periods of persistent suppression or stable loss-of-function phenotype were extended by producing stable cell lines propagating the siRNA expression cassettes. Miyagishi and Taira[6] suppressed β-catenin, a protein involved in cadherin-mediated cell–cell adhesion, for over one week. The β-catenin-targeting siRNA strands were expressed from a plasmid containing the Epstein–Barr virus DNA replication origin, and the plasmid was propagated in cells stably expressing the Epstein–Barr virus nuclear antigen 1 (EBNA-1). Brummelkamp *et al.*[4] and Paddison *et al.*[8] produced cells that stably suppressed p53, an important protein in the cellular response to DNA damage. Silencing of p53 was observed for over two months in antibiotic-selected, stably transfected cells, indicating that long-term expression of siRNAs is not toxic to cells. However, most of the applications of siRNA expression systems described in these four papers focused on transient transfection of siRNA expression vectors and on the analysis of phenotypes associated with protein suppression within a few days after transfection, and therefore did not assess the effects of siRNA expression beyond the temporal window of silencing by the exogenous siRNA transfection approach.

Transfection of plasmid DNA, rather than synthetic siRNAs, may appear advantageous, considering the danger of RNAse contamination and the costs of chemically synthesized siRNAs or siRNA transcription kits. For practical applications, however, the considerable time invested in preparing and amplifying siRNA expression vectors, and the transfection efficiency of plasmids relative to synthetic siRNAs, must also be considered. Furthermore, targeting of essential genes causes arrest in cell growth or cell death within one to three days after delivery of siRNAs, and thus in many instances long-term silencing is unnecessary; however, the development of

© 2002 Nature Publishing Group http://biotech.nature.com

inducible siRNA expression systems may provide an interesting alternative in such cases[6]. For example, cells could be grown on a large scale before induction of the knockdown, which may be beneficial for proteomic analysis. Nonetheless, in targeting nonessential proteins, stable "knockdown" cells may be of great value when studying inducible processes such as UV or other irradiation damage response, host–pathogen interactions, or cell differentiation.

Considering all the pros and cons of expressed versus synthetic siRNAs, it is probably most effective to begin the search for highly effective siRNAs with synthetic, ready-to-use duplex RNAs of defined sequence and length, and to select the synthetic sequences such that they are readily compatible with the sequence requirements for expression within U6 or H1 RNA expression cassettes. This will first entail ensuring that in the U6 snRNA and H1 RNA systems that the +1 position is a guanosine[7,9] and an adenosine, respectively. In addition, it will require that uridines be present in the 3′-terminal position encoded by the oligothymidine Pol III terminator signal sequence[9].

Ultimately, the possibility of stable expression of siRNAs may pave the road for new gene therapy applications, such as treatment of persistent viral infections. Incorporation of siRNA expression cassettes into retroviral vectors may also allow the targeting of primary cells previously refractory to siRNA or plasmid DNA transfection[8]. Considering the high specificity of siRNAs[4–7], the approach should allow the targeting of disease-derived transcripts with point mutations, such as *RAS* or *TP53* oncogene transcripts, without alteration of the remaining wild-type allele. Finally, because of the automation developed for high-throughput sequence analysis of the various genomes, the DNA-based methodology may also be a cost-effective alternative for automated genome-wide loss-of-function phenotypic analysis, especially when combined with miniaturized array-based phenotypic screens[11].

1. Elbashir, S.M. *et al*. *Nature* **411**, 494–498 (2001).
2. Caplen, N.J. *et al*. *Proc. Natl. Acad. Sci. USA* **98**, 9742–9747 (2001).
3. Harborth, J. *et al*. *J. Cell Sci.* **114**, 4557–4565 (2001).
4. Brummelkamp, T.R. *et al*. *Science* 21 March 2002 (10.1126/science.10668999).
5. Lee, N.S. *et al*. *Nat. Biotechnol.* **20**, 500–505 (2002).
6. Miyagishi, M. & Taira, K. *Nat. Biotechnol.* **20**, 497–500 (2002).
7. Paul, C.P. *et al*. *Nat. Biotechnol.* **20**, 505–508 (2002).
8. Paddison, P.J. *et al*. *Genes & Dev.* **16**, in press (2002).
9. Paule, M.R. & White, R.J. *Nucleic Acids Res.* **28**, 1283–1298 (2000).
10. Myslinski, E. *et al*. *Nucleic Acids Res.* **29**, 2502–2509 (2001).
11. Ziauddin, J. & Sabatini, D.M. *Nature* **411**, 107–110 (2001).

Bringing picomolar protein detection into proximity

Combining target recognition by two aptamers, enzymatic ligation, and PCR, the proximity ligation method enables the detection of minute amounts of proteins.

Michael Famulok

Thousands of proteins with potential diagnostic and/or therapeutic applications are expected to emerge from the various genome projects. Proteins in different tissues or individual cells can be up- or downregulated in response to internal or external stimuli, signal transduction, transcriptional control, medication, disease, or pathogens. The ability to monitor slight differences in the amounts of proteins and other biomolecules within the smallest possible detection volumes down to the single-cell level is of utmost importance not only for proteomics research but also for diagnostic and technological purposes. In this issue, Fredriksson *et al.*[1] describe a new method, proximity ligation, that allows the detection and quantification of minute amounts of a specific protein.

Methods currently in widespread use for research requiring standard protein detection include two-dimensional gel electrophoresis, mass spectrometry, and antibody-based detection by sandwich enzyme-linked immunosorbent assays (ELISAs) or western blotting. Though sufficient for most standard applications, these techniques are not sensitive enough to meet the increasingly stringent detection limits required in the postgenomics era. For example, many proteins, such as cytokines and certain kinases, exert their functions at concentrations considerably lower than the thresholds of detection attainable by these methods. To overcome these limitations, variations of the sandwich ELISA with markedly increased sensitivity, such as the "immuno–polymerase chain reaction" (immuno-PCR)[2] and a technique termed "immunodetection amplified by T7 RNA polymerase" (IDAT)[3], have been developed. These methods combine antigen recognition by a biotin-tagged antibody with PCR amplification of a biotinylated reporter DNA, linking the two to one another via a streptavidin molecule. Tags other than biotin, alternative bridging moieties, or covalent coupling methods have also been used in immuno-PCR experiments. When performed under "real-time PCR" conditions, in which a fluorigenic oligonucleotide probe is used to measure PCR product accumulation, these methods can provide very accurate and highly sensitive quantification of the PCR products.

Fredrickson *et al.* do not use any antibodies, tags, or linker molecules and yet achieve levels of protein detection competitive with those of immuno-PCR. They have used a clever strategy whereby an "antigen" is detected by nucleic acid–based receptor molecules, so-called aptamers[4]. As a test case, the researchers use a pair of DNA aptamers that binds to the homodimer of the target analyte protein, platelet-derived growth factor B-chain (PDGF-BB). Each aptamer has a different DNA-sequence extension that does not interfere with its folding and is not required for target recognition. Binding of the aptamer pair brings the ends of the oligonucleotide extensions into close spatial proximity (Fig. 1A) so that a "splint" oligonucleotide can hybridize to both ends, which are subsequently ligated together by T4 DNA ligase. The ligated species can then act as a PCR template and the amplified PCR product can be monitored and quantified under real-time PCR conditions, whereas no signal is obtained with unligated probes (Fig. 1B).

The sensitivity that can be achieved with the method is remarkable. As few as 24,000 molecules, or 4×10^{-20} moles, of the PDGF-BB protein could be detected, approximately 1,000-fold fewer than could be

Michael Famulok is a professor at the Kekulé Institute of Organic Chemistry and Biochemistry, Universität Bonn, Gerhard-Domagk-Strasse 1, 53121 Bonn, Germany (m.famulok@uni-bonn.de).