Research Report

Epitope Tagging Genomic DNA Using a CD-Tagging Tn10 Minitransposon

C.A. Telmer, P.B. Berget, B. Ballou, R.F. Murphy, and J.W. Jarvik
Carnegie Mellon University, Pittsburgh, PA, USA

ABSTRACT

Here, we describe an efficient system for epitope tagging cloned genes by CD tagging using a mini-Tn10 transposon delivery vector. The system was tested against a λFIX genomic clone of the human nucleolin gene. Transfection of HeLa cells with the tagged gene led to the expression of both the appropriately spliced tagged transcript and the appropriately localized tagged protein.

INTRODUCTION

Epitope tagging has been widely used for studies of protein structure and function (8). The epitope tagging process generally begins with a cDNA encoding the protein of interest. A major weakness of this approach is that the cDNA lacks its natural promoter and other regulatory sequences so that when the tagged gene is introduced into cells for analysis, it must be expressed from a heterologous promoter, typically a strong promoter of viral origin. As a result, the natural temporal pattern of expression of the tagged protein is lost, as is its natural level of expression.

Central Dogma (CD) tagging (7) provides an alternative approach to epitope tagging that overcomes the limitations mentioned above. We call this method CD tagging because it tags each molecular class (DNA, RNA, and protein) referred to in the Central Dogma of molecular biology. CD tagging targets genomic DNA not cDNA so that the tagged gene retains the sequences needed for natural gene expression. Successful CD tagging depends on the insertion of a specially designed sequence, the CD cassette, into an intron in the target gene. The CD cassette contains splice acceptor and donor sites surrounding a short open reading frame called the "guest exon". When the target gene is expressed, the guest exon is incorporated into the mRNA and a "guest peptide" appears in the gene product. If the guest peptide contains a sequence of amino acids recognized by an existing antibody, then the gene product is epitope tagged.

In previously published experiments (7), we described CD tagging of a number of eukaryotic genes. In those cases, the target genes were relatively small, and the CD cassettes could be targeted in vitro to predefined intronic restriction sites in plasmids carrying the genes. For larger clones, such as those in phages, phagemids, cosmids, PACs, or BACs, such unique intronic sites may be difficult to identify or may not be known. Even if intronic restriction sites are known, constructing the appropriate recombinants using standard in vitro methods may be technically difficult. Using transposons to deliver the CD tags to the cloned DNA can circumvent these practical limitations. In this report, we describe the development of a bacterial transposon-based delivery system in which the target is genomic DNA cloned into a phage λ vector. Transposon Tn10 was chosen because it was known to transpose to so many potential sites in target DNA that, for our purposes, transposition could be treated as random (9). The high frequency with which one can obtain transpositions into target DNA makes this an attractive procedure. While one might screen through several tagged target molecules to locate ones with a tag in a desired location, it is easy to do this quickly compared with in vitro subcloning and reconstruction strategies. In many cases, screening cells transfected with a population of tagged DNAs should result in the identifica-

tion of the desired tagged gene. In addition, because molecular tagging occurs in vivo, in preexisting cloned DNA, there is no need to handle large DNA molecules in vitro until the final transfer of tagged DNA into the recipient cells for testing.

The CD transposon was tested against λHG3, a λFIX genomic clone containing the 11-kb human nucleolin gene (13). Nucleolin is a nucleolar protein that is readily detectable by the immunofluorescence of human cells using anti-nucleolin antibodies. Clones that had received the CD transposon were selected and sequenced to identify the locations of the insertions, and DNA molecules with insertions in nucleolin introns were transfected into HeLa cells. The analysis of the transfectants showed that tagged mRNA was expressed and epitope-tagged nucleolin was present in its appropriate cellular location.

MATERIALS AND METHODS

The bacterial strain used for the transposition is JJ16, which is *E. coli* MG1655 (1) containing the plasmids pMS421 (5) and pCT16. This strain was grown in the presence of 50 μg/mL spectinomycin and 30 μg/mL kanamycin. The low copy number plasmid, pMS421, carries the pSC101 origin of replication, the spectinomycin resistance gene, and lacI^q gene, which overproduces the lac repressor. The high copy number plasmid pCT16 is a derivative of pJD12 (6), which carries the CD-tagging modified mini-Tn10

transposon (see below) and the wild-type Tn10 transposase, under the control of the ptac promoter. The bacterial strain SA1 is a λPAPA lysogen of *E. coli* DH5α. *E. coli* LE392 (10) was used to titer the λ phage.

Transposon Construction

Plasmid pJD12 is a derivative of pZT344 (2) that was modified to contain rare restriction sites in the Tn10dCam element (6). This mini-Tn10 transposon is bounded by the identical inverted repeats of the outermost 70-bp IS10R (9). To construct pCT16, pJD12 was modified by several manipulations. First, the chloramphenicol resistance gene was removed with a *Bam*HI digest, and, in its place, the *Bam*HI fragment from pUC4K (Stratagene, La Jolla, CA, USA) was inserted. This fragment contains the *npt1* gene, which confers kanamycin resistance. Next, the 757-bp *Afl*III/*Aha*III fragment of pBluescript® KS+ (Stratagene) containing the ColE1 origin of replication was cloned into the *Not*I site of pJD12. At this step, before ligation, the mismatched ends were filled in by Klenow DNA polymerase and dNTPs to generate blunt ends. Then, an *Eco*RI fragment containing the original origin of replication and the β-lactamase gene was removed by *Eco*RI digestion and subsequent ligation. Finally, after converting the *Xho*I site in the CD-1 cassette of the pJJ225 plasmid (7) to a *Not*I site with adaptors, two tandem *Xba*I fragments containing this modified CD-1 cassette were cloned into the

*Xba*I site between the *npt1* gene and the ColE1 origin to make pCT16. The resulting 2809-bp CD transposon (Tn10dCD1) contains, between the IS10R elements, a ColE1 origin, *npt1* (Kan^R) gene, and two tandem copies of the CD-1 cassette (Figure 1).

Introduction of the CD Transposon into the λ Clone λHG3

Midlog JJ16 cells (50 μL at 10^8 cells/mL) in 10 mM MgSO$_4$ were mixed with 50 μL λHG3 (10^8 pfu/mL) to give an MOI of 1. λHG3 contains the complete human nucleolin gene cloned in λFIX (13). Following adsorption for 15 min at room temperature, we added 1 mL LB, and the culture was incubated with vigorous shaking for 5 min at 37°C. IPTG was added to the culture to 1 mM to induce transposition, and the incubation was continued for an additional 60 min. The resulting lysate was titered on *E. coli* LE392 and used to infect *E. coli* SA1 at a MOI of 0.1 phage per cell. Following a 60-min outgrowth, infected SA1 cells were plated onto LB plates containing kanamycin (30 μg/mL).

Colonies of kanamycin-resistant SA1 transductants were transferred into phage dilution fluid and treated with one drop of chloroform to lyse the bacteria. Phage that existed in the colony because of spontaneous induction were isolated by plating on strain LE392. Phage producing clear plaques were purified, and phage stocks were made using standard methods. The λ DNA for sequencing and transfection was prepared from 10 mL phage lysate using
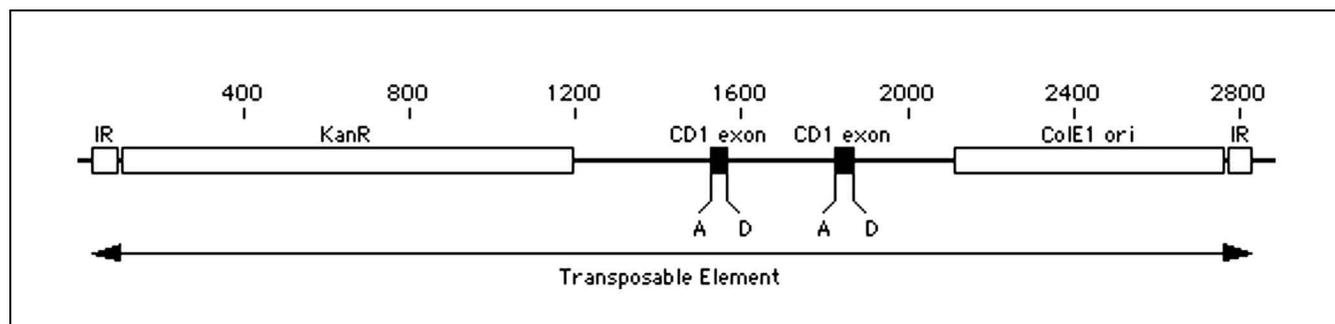
Figure 1. Tn10dCD1, the transposable element in the plasmid pCT16. The boxes represent important DNA sequences. The transposable element is flanked by inverted repeats (IRs) that come from the outermost 70 bp of Tn10 IS10R and are the target of the Tn10 transposase. KanR indicates the *npt1* gene that confers kanamycin resistance. ColE1 ori is the origin of DNA replication from the ColE1 plasmid. CD1 exon is the sequence of the 54-bp exon previously described (7). The splice acceptor and donor sites are labeled A and D, respectively. The CD-1 exon cannot code for any stop codons in any of the three possible reading frames into which it could be spliced. When the CD-1 exon is spliced into an mRNA from a type 1 intron, the exon codes for a polypeptide that contains the HA epitope YPYDVPDYA. The remainder of the plasmid pCT16 (not shown) contains the Tn10 transposase under the control of the ptac promoter (2).

Research Report

the Lambda Mini Kit (Qiagen, Valencia, CA, USA). DNA sequencing was performed using a primer (5′-ACGGG GTCTGACGCTCAGTGGAACG-3′) complementary to the sequence in the ColE1 origin of replication nearest the IS10R element shown in Figure 1.

λ GFP Construction

We cloned pEGFP-N1 (BD Biosciences Clontech, Palo Alto, CA, USA) into λFIX II, along with 4.9 kb stuffer DNA. The pEGFP-N1 was digested with *Bam*HI and *Bgl*II and ligated to remove most of the multiple cloning sites. This plasmid was then linearized by digestion with *Afl*III, and the recessed ends filled in with Klenow DNA polymerase and dNTPs. This blunt-ended fragment was ligated to the 4.9-kb *Hpa*I/*Xmn*I fragment of pARG7.8, which carries the chloramphenicol resistance gene and the 5′ end of the ARG 7.8 gene of *Chlamydomonas reinhardtii* (12). This 9.5-kb plasmid was then linearized with *Eco*RI and ligated to *Eco*RI-digested λFIX II arms. The desired phage clone was isolated from plaques on XL-1 Blue cells after packaging the reaction with Gigapack® III Plus packaging extract (Stratagene).

Cell Culture and Transfections

HeLa cells were cultured at 37°C in 5% CO_2 in DMEM + 10% FCS. Cells (10^5) were plated 24 h before transfection in 35-mm plates and transfected using 2 µg tagged DNA and 2 µL LIPO-FECTAMINE™ (Invitrogen, Carlsbad, CA, USA). Five hours later, the culture was diluted, and fresh serum was added. The cells were prepared for immunofluorescence microscopy by transferring them to 18-mm autoclaved coverslips in 24-well plates 24 h after transfection.

To obtain stable transfectants, HeLa cells were co-transfected with 2 µg tagged DNA and 0.2 µg pMC1neo DNA (Stratagene). Selection was performed by growth in medium containing 400 µg/mL G418. The cells were stained 48 h after the start of the transfection. Forty-eight clones were picked and screened by immunofluorescence.

RT-PCR

For each transfected cell line to be tested, total RNA was extracted from one 35-mm tissue culture well (about 5 × 10^6 cells) 48 h after transfection. RNA was isolated using the RNeasy® kit (Qiagen) and resuspended in a total volume of 30 µL. First-strand cDNA was made from 2 µL total RNA using the SUPER-SCRIPT™ First Strand Synthesis kit (Invitrogen). The cDNA was resuspended in a final volume of 10 µL. PCR amplification analysis was carried out in two successive reactions using nested primers. The first reaction amplified cDNA between each nucleolin exon and the guest CD-1 exons. In the second reaction, first-reaction DNA was amplified using the same nucleolin exon-specific primer and a primer specific for the unique junction between the two tandem CD-1 guest exons. The first-reaction contained 2 µL first-strand cDNA and the appropriate primers in a final reaction volume of 50 µL. The second reaction contained 2 µL of the first-reaction and the second-stage primers in a total reaction volume of 50 µL *Taq* DNA polymerase, and supplied buffers and nucleotides (Roche Molecular Biochemicals, Indianapolis, IN, USA) were used in all PCRs. Primers were all synthesized by Ransom Hill Bioscience, and Figure 2 shows their positions relative to their use in the nucleolin gene. The primer specific for the 3′ end of exon 7 (Figure 2, 7-F) was 5′-AAGAC-CTGGAGAAAGCGTTG-3′. The primer specific for the 5′ end of exon 8 (Figure 2, 8-R) was 5′-CTTTTCCCA-TCCTTGCTGAC-3′. CD-1-specific primers were CD-F, 5′-GTCCCCGAC-TACGCCACGAA-3′, and CD-R, 5′-GGACGTCGTAGGGGTACCGA-3′. Primers specific for the junction between the two CD-1 guest exons were CD-JF, 5′-AAGATCTCAGGTGGA A-GAGC-3′, and CD-JR, 5′-GCTCTTCC-ACCTGAGATCTT-3′. Correct splicing at the 3′ end of exon 7 was monitored by PCR using primers 7-F and CD-R in the first round and 7-F and CD-JR in the second round. The predicted size for the final product was 159 bp. Correct splicing at the 5′ end of exon 8 was monitored by PCR using primers 8-R and
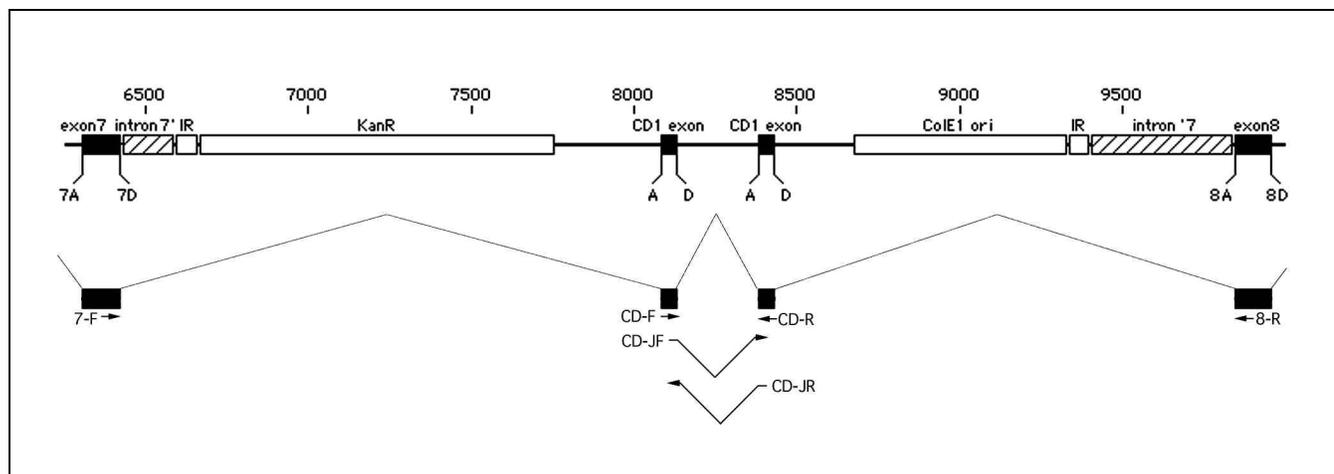
Figure 2. Region surrounding intron 7 of the Tn10dCD1-tagged nucleolin gene. The diagram shows the suspected DNA structure and resulting mRNA splicing pattern of the Tn10dCD1-tagged nucleolin gene at nucleotide 6597. The splice sites at the A and D ends of both nucleolin and CD cassette exons are labeled. The forward and reverse PCR primers for RT-PCR analysis described in Materials and Methods are shown as arrows below the mRNA splicing diagram.

Research Report

CD-F in the first round and 8-R and CD-JF in the second round. The predicted size for the final product was 178 bp.

Immunofluorescence

Cells on coverslips were rinsed in PBS and fixed with 3% paraformaldehyde in PBS for 10 min. Following fixation, the cells were rinsed twice in PBS and then blocked/quenched using 1% BSA, 200 mM glycine, and 0.05% Nonidet™ P-40 in PBS for 30 min in the dark. The coverslips were inverted into 15 µL primary antibody and incubated overnight at room temperature in a humidified chamber. The primary antibody was the anti-influenza hemagglutinin (HA) 12CA5 antibody (Roche Molecular Biochemicals) at 2 µg/mL in block/quench solution with 0.02% sodium azide. The coverslips were then rinsed three times with PBS containing 0.05% Nonidet P-40 and inverted into a 1/50 dilution of Cy-3-labeled secondary antibody (Jackson ImmunoResearch Laboratories, West Grove, PA, USA) for 30 min in the dark at room temperature. After several final rinses with PBS containing 0.05% Nonidet P-40, the cells were mounted in 1 mg/mL phenylenediamine and 90% glycerol, and the coverslips were sealed to the slides with nail polish.

Images were collected using a Plan-Neofluor objective (40×, 1.3 numerical aperture; Carl Zeiss, Thornwood, NY, USA) and a Photometrics CH 250™ cooled charge-coupled device (512 × 382 pixels, 23 mm/pixel; Roper Scientific, Trenton, NJ, USA) mounted on a customized Axiovert microscope (Carl Zeiss) (3).

RESULTS AND DISCUSSION

The CD-tagging mini-Tn10 transposon Tn10dCD1 (Figure 1) was constructed and mobilized to insert into the human genomic nucleolin clone λHG3 resident in a λFIX vector.

The *E. coli* strain JJ16, which carries the CD-tagging mini-Tn10 transposon (Figure 1) with two CD cassettes, a ColE1 origin of replication, and the *npt1* gene conferring kanamycin resistance, was infected at an MOI of 1 with λFIX clone HG3 carrying the human nucleolin gene. The transposition of the

Tn10dCD1 was induced with IPTG. Progeny phage were collected and used to infect the λ lysogen SA1. The ColEI origin in the transposon allows any tagged λHG3 derivative to replicate as a phagemid circle. Even though λHG3 is cI⁻, the phagemid remains circularized and its genes repressed in the presence of the λ repressor, which is synthesized from the prophage in SA1. The *npt1* gene in the CD tag confers kanamycin resistance to a host SA1 bacteria, containing a λ clone that had received Tn10dCD1. Transductants harboring phage genomes carrying a transposon were selected by plating on kanamycin medium. The individual kanamycin-resistant colonies were transferred into dilution fluid and treated with chloroform. Any λ phage released from the cells were detected by plating on the permissive strain LE392. In all cases, turbid plaques, derived by the spontaneous induction of the wild-type prophage in SA1, were observed on the plates. In most cases, clear plaques, derived from spontaneous induction of the λFIX phagemid, were also observed. A number of clear plaques, each derived from an independently transduced cell, were picked and used to make DNA for further analysis. In these experiments, approximately 1 λHG3 phage in 1000 received the CD transposon, as measured by the number of kanamycin-resistant transducing particles compared with the total phage released from the JJ16 infection.

One should note that several features of this process were designed to increase the likelihood that phage with single transposon insertions in the target gene would be obtained. After the initial period of transposition, all phage chromosomes that may have received the CD transposon are transduced into a host (SA1) that lacks transposase. This prevents the loss or further mobilization of the transposon. Selection against deletions of significant size and selection against most insertions into the vector itself were accomplished by requiring that the tagged λ clones retain the ability to form plaques.

Identification of Tn10dCD1 Insertions in Nucleolin Introns

CD tagging depends on the correct-

ly oriented insertion of the CD tag into an intron in the target gene. The human nucleolin gene provides 13 target introns that range in size from 104 to 1164 bp. To identify the insertions, the locations of the transposon in nine randomly chosen tagged clones were determined by DNA sequencing (data not shown). The sequencing reactions started within the ColE1 origin of replication and proceeded out of the transposon into the target DNA. Five of the insertions proved to be in the nucleolin gene. The other four insertions were in the λ vector in a small nonessential region in the right arm of λ. All nucleolin insertions were in introns—three in intron 7, one in intron 10, and one in intron 12. Three of these five were in the correct (forward) orientation for tag expression, meaning that the exons in the CD cassettes in the transposon were oriented in the same direction as the nucleolin exons. These three correctly oriented insertions included two in intron 7 at base pairs 6597 and 6948, and one in intron 10 at base pair 8571. The numbering system is that used by Srivastava et al. (13). DNA for transfection into HeLa cells was prepared from each of the three forward-oriented insertion clones. The top of Figure 2 shows the presumed structure of the intron 7 insertion at base pair 6597.

Expression in HeLa cells

Before performing transfection experiments with the tagged nucleolin clones, we determined the efficiency of HeLa cell transfection with λ-sized DNA. We constructed a λFIX clone with a constitutively expressed enhanced GFP (EGFP) gene, as described in Materials and Methods. DNA from this λ clone was used to transfect HeLa cells, and we determined the fraction of cells that expressed the gene by fluorescence microscopy. Some of the transfected cells (5%–15%) showed distinct EGFP fluorescence, indicating that our transfection protocol was reasonably efficient in delivering λ-sized DNA to HeLa cells (data not shown).

Each of the forward-oriented tagged nucleolin clones was transfected into HeLa cells. RNA isolated from transiently transfected cell lines was reverse transcribed, and we assessed the splicing

Research Report

pattern of the two CD-1 guest exons between the appropriate nucleolin exons using the two-step PCR protocol described in Materials and Methods. The cDNA sequences between adjacent nucleolin exons and both of the CD-1 exons are amplified in the first round of PCR. In the second round of PCR, a primer specific for the junction sequence between two correctly spliced CD-1 exons is used to amplify any product containing two correctly spliced CD-1 exons and an adjacent nucleolin exon. The presence of two tandem copies of the CD-1 cassette in Tn10dCD1 serves two purposes at this stage of molecular analysis. First, it allows us to detect unspliced RNA (or contaminating genomic DNA) through the presence of larger PCR products containing the intron sequences between the two CD-1 cassettes. Second, it allows us to design primers specific for the junction between the two correctly spliced CD-1

exons, thus increasing the specificity of the two-step PCRs. Figure 3 shows the RT-PCR products for the two insertions in the nucleolin intron 7. Figure 3, lanes 1 and 2, contains the products amplified from the cell line containing the Tn10dCD1 insertion at nucleotide 6597.

Lanes 3 and 4 contain the RT-PCR products amplified from the cell line containing the Tn10dCD1 insertion at nucleotide 6948. Lanes 1 and 3 contain the product from exon 7 to the junction between the two CD-1 exons, and lanes 2 and 4 contain the product from exon 8 to
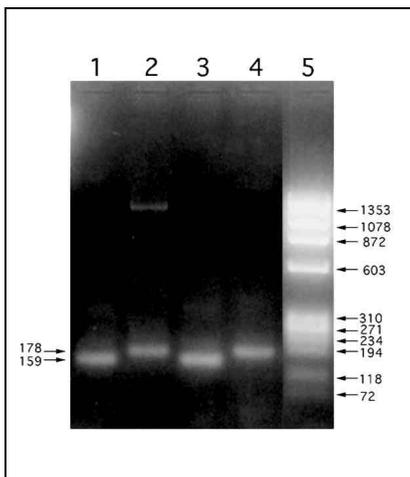
Figure 3. RT-PCR analysis of mRNA from Tn10dCD1-tagged nucleolin genes. Lanes 1 and 2 contain PCR products from the Tn10dCD1 insertion at nucleotide 6597 in intron 7, and lanes 3 and 4 contain products from the insertion at nucleotide 6948. Lane 5 contains the size standard (ΦX174 *Hae*III digestion). The sizes of the standard bands are indicated to the right of the gel image. The products shown in lanes 1 and 3 were derived from the 3′ end of exon 7 and the junction of the two CD cassette exons. Products in lanes 2 and 4 were derived from the 5′ end of exon 8 and the junction of the two CD cassette exons. The RT-PCR bands at 159 and 178 bp, respectively, are the size expected for accurate splicing between the flanking nucleolin exons and the CD cassette exons. The minor band at about 1500 bp in lane 2 is consistent with a PCR product generated by primers CD-F and 8-R from contaminating genomic DNA.
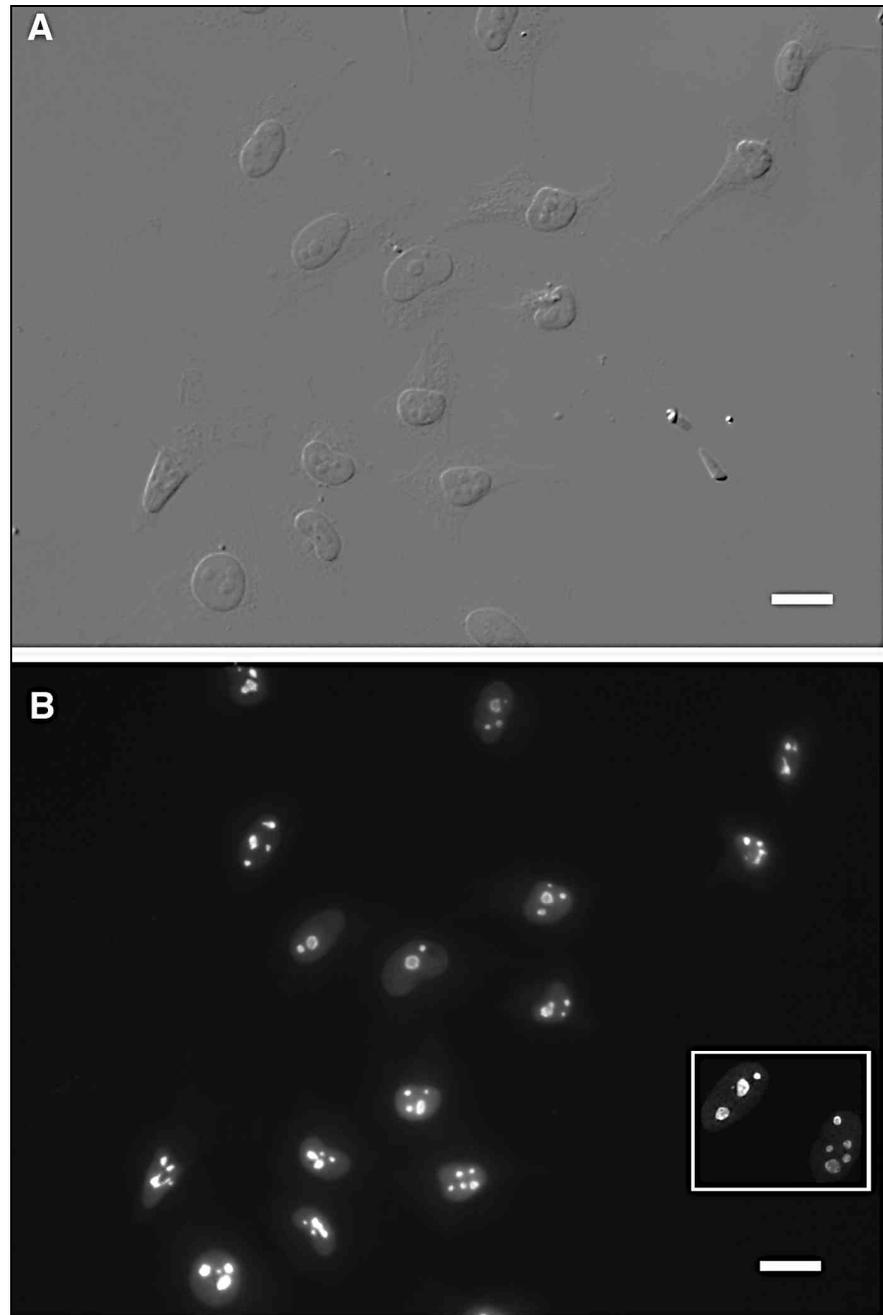
Figure 4. HeLa cells stably transfected with the λFIX clone containing Tn10dCD1-tagged nucleolin gene at nucleotide 6948 in intron 7. (A) Micrograph showing the DIC image of the cells. Scale bar is 20 μm. (B) Fluorescence micrograph of cells reacted with the 12CA5 anti-HA primary antibody and Cy3-labeled secondary antibody. Scale bar is 20 μm. The insert box contains the images of two nuclei from untagged cells stained with anti-nucleolin antibodies.

the junction between the two CD-1 exons. Both products from both cell lines are the appropriate size (159 and 178 bp, respectively), indicating that correct splicing of both of the guest exons into the nucleolin mRNA has occurred. Analogous results were obtained for the CD-1 exon-splicing pattern between the nucleolin exons 10 and 11 for the Tn10dCD1 insertion at nucleotide 8571 in intron 10 (data not shown).

Identification and Localization of Epitope-Tagged Nucleolin

Because the guest exons in the CD-1 transposon encode the HA epitope (YPYDVPDYA) that is recognized by commercially available antibody 12CA5, we were able to use indirect immunofluorescence to detect and localize the epitope-tagged nucleolin in the HeLa transfectants. Such an analysis was performed for both stable and transiently transfected cell lines containing the Tn10dCD1 insertion at nucleotide 6948. The nuclei and nucleoli of the HeLa cells are clearly visible in the differential interference contrast (DIC) image in Figure 4A. We observed immunostaining throughout the nuclei, with very bright staining concentrated in the nucleoli (Figure 4B). The staining clearly coincides with the structures visible in the DIC image. This staining pattern was indistinguishable from that observed in untagged cells stained with anti-nucleolin antibodies (Figure 4B, insert panel).

The guest exon in the CD cassette was designed to express the HA epitope when the cassette is in a class 1 intron (i.e., in an intron that interrupts the coding sequence between the first and second nucleotides of a codon). When inserted into class 0 or class 2 introns, the exons are translated in different reading frames and yield entirely different peptide sequences that do not include the HA epitope. Intron 7 is a class 1 intron; hence, we were able to perform the immunofluorescence experiments described earlier. However, intron 10 is class 2, so we did not attempt to localize the tagged nucleolin in the intron 10-tagged cells.

Because nucleolin introns 3, 4, 5, 11, and 13 are all class 1, the system described here could be used without modification to tag this gene and its proteins at each of these positions. An alternative tagging strategy would be to employ guest exons that contain a repeating nucleotide sequence that encodes the same epitope in any reading frame. One such reading frame-independent epitope tag has been developed and tested (11). This tag, encoding the repeating epitope HTTPHH, has been incorporated into a new generation of CD transposons.

Implications with Respect to Nucleolin Biology

Our purpose in performing the constructions and experiments described here was to develop and validate the CD-transposon system, and the human nucleolin gene was chosen as a convenient test gene. Although our original intent was not to explore nucleolin biology per se, our results do provide some new data regarding nucleolin.

Nucleolin is an abundant non-ribosomal nucleolar protein that is involved in numerous cellular processes (4,14). The protein is organized into three parts: an amino-terminal domain consisting of alternating basic and acidic sequences, a middle section of four RNA-binding domains, and a C-terminal domain of repeating RGG motifs. The known nucleolin metabolic functions include involvement in rRNA transcription and processing, ribosome assembly, nucleo-cytoplasmic transport of ribosomes, and transcriptional regulation. Its known biochemical functions include autoproteolysis, DNA- and RNA-helicase activity, ATPase function, and specific binding to rRNA precursors and ribosomal proteins.

Intron 7 is located between the regions coding for RNA-binding domains I and II. The CD-tagged version of the protein produced in these studies contains a 36-amino-acid insertion at residue 389 in the 710-amino-acid protein. Because the immunofluorescence pattern for the tagged protein was identical to that observed with the anti-nucleolin antibodies, one can conclude that the epitope tag does not interfere with RNA binding or, if it does, then the RNA binding activity that is affected is not required for proper nucleolin localization to the nucleoli.

Research Report

CONCLUSION

This study demonstrates that by using a mini-Tn10 transposon system, we can readily CD tag genes cloned in phage λ vectors. We have used the same system, with minor modification, to tag genes cloned into cosmid and P1 vectors (data not shown). Although a small epitope tag was used in the experiments reported here, the same approach can be used to deliver CD cassettes that yield proteins tagged with larger guest peptides that provide reporter activities, such as fluorescence (e.g., GFP) or antibiotic resistance.

An advantage of this tagging procedure is that it is performed entirely in *E. coli*, using standard microbiological manipulations and procedures; one does not have to isolate and handle large DNA molecules until the tagged DNA is prepared for introduction into mammalian (or other) recipient cells. For this reason, we prefer this method to an alternative approach of isolating target DNA, introducing the tags by in vitro transposition, and reintroducing the tagged DNA into *E. coli* or other cells for analysis.

The experiments reported here resulted in the tagging of a single defined clone, but the same protocols can be used to tag complex genomic libraries for the purpose of gene discovery and/or annotation. In such an investigation, one would tag the library en masse, isolate tagged DNA, transfect recipient cells, and select or screen for cells expressing the guest peptide. An alternative but related approach is to use an engineered retrovirus to deliver CD cassettes directly to the genomes of cultured cells (manuscript in preparation).

What about the effect of the CD tag on protein function? In this case, a 36-amino acid internal insertion into the nucleolin did not appear to compromise its function, at least with respect to protein localization. Additionally, we have seen many other examples (Reference 7 and unpublished observations) of CD-tagged proteins that retain normal function. Our own experience to date and the experience of others (8) indicates that internal tags are no more likely to compromise protein function than are the more commonly used amino or carboxyl terminal tags. Thus, CD tagging provides a versatile means to tag complete genes and to observe functional tagged protein at the cell level.

REFERENCES

1. **Bachmann, B.J**. 1987. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12, p. 1190-1219. *In* F.C. Neidhardt (Ed.), *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology, vol. 2. American Society for Microbiology, Washington, D.C.
2. **Elliott, T. and J.R. Roth.** 1988. Characterization of Tn10d-Cam: a transposition-defective Tn10 specifying chloramphenicol resistance. Mol. Gen. Genet. *213*:332-338.
3. **Farkas, D.L., G. Baxter, R.L. DeBiasio, A. Gough, M.A. Nederlof, D. Pane, J. Pane, D.R. Patek et al.** 1993. Multimode light microscopy and the dynamics of molecules, cells, and tissues. Annu. Rev. Physiol. *55*:785-817.
4. **Ginisty, H., H. Sicard, B. Roger, and P. Bouvet.** 1999. Structure and functions of nucleolin. J. Cell Sci. *112*:761-772.
5. **Grana, D., T. Gardella, and M.M. Susskind.** 1988. The effects of mutations in the ant promoter of phage P22 depend on context. Genetics *120*:319-327.
6. **Heath, J.D., J.D. Perkins, B. Sharma, and G.M. Weinstock.** 1992. *Not*I genomic cleavage map of *Escherichia coli* K-12 strain MG1655. J. Bacteriol. *174*:558-567.
7. **Jarvik, J.W., S.A. Adler, C.A. Telmer, V. Subramaniam, and A.J. Lopez.** 1996. CD-tagging: a new approach to gene and protein discovery and analysis. BioTechniques *20*:896-904.
8. **Jarvik, J.W. and C.A. Telmer.** 1998. Epitope tagging. Annu. Rev. Genet. *32*:601-618.
9. **Kleckner, N., J. Bender, and S. Gottesman.** 1991. Uses of transposons with emphasis on Tn10. Methods Enzymol. *204*:139-180.
10. **Murray, N.E., W.J. Brammar, and K. Murray.** 1977. Lambdoid phages that simplify the recovery of in vitro recombinants. Mol. Gen. Genet. *150*:53-61.
11. **Nelson, R.W., J.W. Jarvik, B.E. Taillon, and K.A. Tubbs.** 1999. BIA/MS of epitope-tagged peptides directly from *E. coli* lysate: multiplex detection and protein identification at low-femtomole to subfemtomole levels. Anal. Chem. *71*:2858-2865.
12. **Purton, S. and J.D. Rochaix.** 1994. Complementation of a *Chlamydomonas reinhardtii* mutant using a genomic cosmid library. Plant Mol. Biol. *24*:533-537.
13. **Srivastava, M., O.W. McBride, P.J. Fleming, H.B. Pollard, and A.L. Burns.** 1990. Genomic organization and chromosomal localization of the human nucleolin gene. J. Biol. Chem. *265*:14922-14931.
14. **Tuteja, R. and N. Tuteja.** 1998. Nucleolin: a multifunctional major nucleolar phosphoprotein. Crit. Rev. Biochem. Mol. Biol. *33*:407-436.

Address correspondence to:

Dr. P.B. Berget
Department of Biological Sciences
4400 Fifth Avenue
Carnegie Mellon University
Pittsburgh, PA 15213, USA
e-mail: berget@andrew.cmu.edu