The bacterial transposon Tn7 causes premature polyadenylation of mRNA in eukaryotic organisms: TAGKO mutagenesis in filamentous fungi

Clive Lo1,2,*, Kiichi Adachi1, Jeffrey R. Shuster1, John E. Hamer1 and Lisbeth Hamer1

ABSTRACT

TAGKO is a Tn7-based transposition system for genome wide mutagenesis in filamentous fungi. The effects of transposon insertion on the expression of TAGKO alleles were examined in Magnaporthe grisea and Mycosphaerella graminicola. Northern analysis showed that stable, truncated transcripts were expressed in the TAGKO mutants. Mapping of the 3′-ends of TAGKO cDNAs revealed that they all contain Tn7 end sequences, regardless of the transposon orientation. Polyadenylation signals characteristic of eukaryotic genes, preceded by stop codons in all frames, are located in both ends of the bacterial transposon. Thus, TAGKO transcripts are prematurely polyadenylated, and truncated proteins are predicted to be translated in the fungal mutants. Depending on the extent of protein truncation, TAGKO mutations in HPD4 (encoding p-hydroxyphenylpyruvate dioxygenase) resulted in tyrosine sensitivity in the two fungi. Similarly, a particular M. grisea CBS1 (encoding cystathionine β-synthase) TAGKO cDNA failed to complement cysteine auxotrophy in a yeast CBS mutant. TAGKO, therefore, represents a useful tool for in vivo study of truncated gene products in filamentous fungi.

INTRODUCTION

Transposons frequently account for spontaneous mutations that result in natural variations in different organisms (1). They are mobile DNA elements originally discovered by McClintok in maize (2), and have become indispensable tools in the molecular genetics of both prokaryotic and eukaryotic organisms (3,4). Large-scale transposition in endogenous or heterologous systems has been employed to generate large collections of insertional mutants for gene identification and function analysis (5). A number of in vitro transposition systems (based on Tn5, Tn7, Mu, Himar1, etc.) have now been developed (6–9). There are many applications for in vitro transposition systems. For example, transposon insertions provide primer sites for sequencing of the recipient DNA. In addition, inserted DNA from whole genome in vitro transposition can be introduced to the host for mutant generation when homologous recombination is efficient in a target organism.

The Tn7-based in vitro transposition system is highly efficient and useful for analysis of genomes and genes (6). The bacterial transposon Tn7 encodes transposases that directly participate in transposition (10). The notable Tn7 target site preference has been abolished with the use of a mutant form of one of the Tn7 transposases during in vitro transposition (6). This feature, together with its innate transposition immunity (11), allows the use of Tn7 to generate random and thorough mutagenic insertions in genomic libraries. Mini-Tn7 elements with modified end sequences for recovery of derivatized target genes were also described (6). One of the elements has truncated ends that are open in all frames, thus allowing productive fusions with target gene products. The other element generates 15 bp (5 amino acid) linker insertions in target genes (proteins) to facilitate the identification of critical base pairs or amino acids.

Recently we described a genome-wide mutagenesis technology, TAGKO (transposon arrayed gene knockout), in filamentous fungi using a Tn7-based transposon cassette (12). Basically, only the end segments (Tn7-L, 150 bp; Tn7-R, 90 bp) of Tn7 are required for recognition by the transposition machinery (10). The TAGKO cassette contains a hygromycin phosphotransferase gene (HPH) engineered between the transposon ends as a fungal selection marker. Random in vitro transposition can be performed on cosmid libraries of fungal genomes with the addition of a Tn7 transposase mix (6,12,13). Individual cosmid-based TAGKO clones serve directly as gene disruption vectors and improve the frequency of homologous recombination significantly (12,14). In addition, a series of mutation alleles in the target genes can be generated with TAGKO, providing a range of phenotypes for gene function analysis.

*To whom correspondence should be addressed at Department of Botany, The University of Hong Kong, Pokfulam Road, Hong Kong, China.
Tel: +852 2299 0337; Fax: +852 2858 3477; Email: clivelo@hkucc.hku.hk
Present addresses:
Kiichi Adachi, Mitsui Global Strategic Studies Institute, 1-2-1 Otemachi Chiyoda-ku, Tokyo, Japan
Lisbeth Hamer, Department of Microbiology, North Carolina State University, Raleigh, NC 27605, USA

Nucleic Acids Research, Vol. 31 No. 16 © Oxford University Press 2003; all rights reserved
Transposons insertions can lead to a variety of effects on target gene expression, depending on both the location and the properties of the element (15). In this study, we examine the transcription patterns of different TAGKO alleles in the filamentous fungi *Magnaporthe grisea* and *Mycosphaerella graminicola*, two of the major cereal pathogens in the world. Interestingly, we discovered that Tn7 contains 3'-end formation site information unique to eukaryotic genes, resulting in the generation of truncated transcripts that are chimeric and polyadenylated in the TAGKO mutants. The potential application of this novel feature of the bacterial transposon Tn7 in fungal gene analysis is discussed.

MATERIALS AND METHODS

Fungal strains and growth assays

*Magnaporthe grisea* strain Guy11 (16) and *M. graminicola* strain PG1 (12) were used as the wild-type (WT) strains in this study. The WT strains and all TAGKO mutants created were maintained on complete medium (CM) agar (17). Growth assay experiments were performed on minimal medium (MM) agar (17) with or without the supplementation of L-tyrosine (Sigma, St Louis, MO) at a concentration of 4 mM.

Generation of TAGKO mutants in *M. grisea* and *M. graminicola*

From our collections of TAGKO clones generated by whole genome *in vitro* transposition previously (12,14), we selected individual gene disruption vectors with transposon inserted into different locations of the *HPD4* (encoding *p*-hydroxyphenylpyruvate dioxygenase) or *CBS1* [encoding cystathionine-δ-synthase (18)] genes. Transposon insertion sites were identified by analysis of sequencing data derived from Tn7-end specific primers (12). The selected cosmids were linearized and transformed into the corresponding fungus, following procedures described previously (14,19,20).

Transformants, selected on hygromycin-containing medium, were isolated and DNA extracted as described (20). To identify TAGKO mutants, primer pairs were designed to flank the transposon insertion sites for PCR screening of homologous recombination events, using the transformant DNA samples as templates (12).

RNA experiments

WT strains and TAGKO mutants were cultured in CM for 4 days before the mycelia were harvested. Total RNA was extracted from lyophilized mycelia with a RNeasy Midi kit (Ambion, Austin, TX) following the manufacturer’s instructions. Northern blot experiments were performed as described (21). Transcript ends were mapped with a 3'-rapid amplification of cDNA (RACE) ends kit (Life Technologies, Rockville, MD), followed by DNA sequencing. Briefly, 4 μg of total RNA was used as the template for first strand cDNA synthesis. Specific cDNA was amplified by PCR using a gene-specific primer that anneals to the known exon sequence and an adapter primer (Life Technologies, Rockville, MD) that targets the poly(A) tail region. The following *HPD4*-specific primers were used for PCR amplification: *M. grisea*, 5'-ATGTCAACCCTGGCATTAC-3'; *M. graminicola*, 5'-ATCGCATCCGGGCAGCT-3'.

Yeast complementation analysis of *M. grisea CBS1* TAGKO alleles

*CBS1* cDNAs from *M. grisea* WT and TAGKO strains were amplified by PCR using first strand cDNA mixtures as templates with the following primers: *CBS3*, 5'-CACACAATCTAAAGAATGGC-3', and AUAP (Life Technologies, Rockville, MD), 5'-GGCCACCGGTAGCATTAC-3'. The amplified PCR products contained the first start codon (ATG) through the end of the transcripts with the majority of the 5'-untranslated region (UTR) removed. *CBS1* WT and TAGKO cDNAs were cloned into the expression vector pYES2.1/V5- HiS-TOPO under the control of a GAL1 promoter (Invitrogen, Carlsbad, CA). A *Saccharomyces cerevisiae* cystathionine δ-synthase (CBS) deletion mutant (strain number 6696, Research Genetics, Inc., Huntsville, AL) was transformed with the different *M. grisea CBS1* expression clones following the manufacturer’s instruction (Invitrogen, Carlsbad, CA). The transformants were grown on synthetic complete (SC) medium (Invitrogen, Carlsbad, CA) containing galactose (2% w/v) as the carbon source to induce *CBS1* expression from the GAL1 promoter. Complementation of CBS mutation was assessed by the growth of the yeast transformants without the supplementation of cysteine in the SC medium.

RESULTS

Expression analysis of *M. grisea* and *M. graminicola* HPD4 TAGKO mutants revealed the presence of chimeric mRNA caused by Tn7 end signals

To understand how TAGKO could alter target gene expression, we investigated the transcription events of *HPD4* mutant alleles in *M. grisea* and *M. graminicola*. The *HPD4* genes in the two fungi are both intronless. Two TAGKO alleles with transposon inserted in opposite orientations were selected for analysis in each fungus (Fig. 1A). In the *M. grisea* KO1 mutant, the TAGKO cassette was inserted in the 3'-UTR, while in the KO2 mutant, the transposon was inserted within the coding region. The two *M. graminicola* HPD4 mutants have transposon insertions in different locations along the coding region.

Northern analysis showed that the WT *HPD4* transcripts are ~1.7 and 1.6 kb in length in *M. grisea* and *M. graminicola*, respectively (Fig. 1B). The *M. grisea* KO1 mutant produces a transcript with a similar size to that of the WT, while the KO2 mutant produces a transcript of only 1.3 kb in length. Similarly, *M. graminicola* HPD4 KO1 and KO2 mutants produce shortened transcripts of 0.8 and 1.5 kb in length, respectively. The transcript lengths for the *HPD4* TAGKO alleles appear to be related to the locations of the transposition event. Thus, shorter transcripts are produced in mutants with TAGKO inserted further upstream along the *HPD4* genes.

The 3'-ends of the *HPD4* transcripts from the different TAGKO alleles were mapped using the RACE procedure (Materials and Methods). The 3'-UTRs in the WT *HPD4* transcripts are 208 and 128 nucleotides in length in *M. grisea* and *M. graminicola*, respectively (data not shown). All the TAGKO transcripts are chimeric, containing the endogenous *HPD4* sequences up to the insertion site, followed by sequences derived from either the left arm (Tn7-L) or the right arm (Tn7-R) of the transposon cassette (Fig. 1C).
Depending on the transposition orientation, 147 bases of Tn-7L or 140 bases of Tn-7R are fused to the 3′-ends of the TAGKO transcripts in both fungi (Fig. 1C). In addition, these transcripts are all polyadenylated. Examination of the Tn7 ends revealed the presence of 3′-site formation signals, ATTAAA in Tn-L and AATAAA in Tn-R (Fig. 2), which are located upstream of the poly(A) tails. Both of the Tn7 end regions also contain stop codons (TAA, TAG or TGA) in all three reading frames parallel to the direction of the target genes. Thus, fusion gene products truncated at the C-terminus are expected to be translated from the chimeric transcripts in the mutants.

Severity of mutant HPD4 phenotypes correlates to the length of the truncated gene products

We attempted to correlate the mutant phenotypes to the different predicted HPD4 TAGKO proteins. The *M. grisea* and *M. graminicola* HPD4 genes both encode a protein of 419 amino acids in length (Fig. 3A). The *M. grisea* KO1 mutant was expected to produce a WT protein since the TAGKO insertion occurred downstream of the endogenous stop codon. In the *M. grisea* KO2 mutant, the truncated HPD4 coding sequence is in frame with the first stop codon in the Tn7-L end (Fig. 1C). Thus, the predicted TAGKO protein contains 273 amino acids with a C-terminal extension of 5 amino acids derived from Tn7 (273 + 5 amino acids; Fig. 3A). Similarly, the *M. graminicola* KO1 and KO2 mutants produce HPD4 TAGKO proteins with 137 + 5 amino acids and 378 + 17 amino acids, respectively (Fig. 3A).

Tyrosine and phenylalanine are metabolized through a conserved pathway involving *p*-hydroxyphenylpyruvate dioxygenase (HPD4 gene) (22–24). The enzyme converts hydroxyphenylpyruvate, an immediate metabolite of tyrosine, into homogentisate which is further metabolized by other enzymes in the degradation pathway. Disruption of this pathway often leads to the accumulation of toxic intermediates that impair normal growth and development (22). Our growth experiments showed that the WT strains of both *M. grisea* and

**Figure 1.** Chimeric, polyadenylated HPD4 transcripts in *M. grisea* and *M. graminicola*. (A) Two HPD4 mutant alleles, KO1 and KO2, are presented in each fungus. Block arrows represent the HPD4 coding regions. The orientation of the TAGKO transposon cassette is shown in arrows (Tn7-R, R→Tn7-L, L). P1 and P2 denote PCR-derived probes used in northern hybridization experiments. (B) Northern analysis of HPD4 gene expression. Approximately 10 µg of total RNA was used for each strain. Actin genes served as positive controls for gene expression in the two fungi. (C) Mapping of the 3′-ends of HPD4 cDNAs. TAGKO transcripts are chimeric and polyadenylated. Endogenous HPD4 sequences are shown in black, sequences derived from Tn7 transposon ends are shown in grey and stop codons in frame with the truncated HPD4 coding sequences are indicated. Numbers in parentheses represent the first (1) or second (2) stop codon identified in the respective Tn7 end.

**Figure 2.** Examination of transposon end sequences in the TAGKO cassette. An HPH gene was engineered between the transposon ends (Tn7-L and Tn7-R). Two possible orientations of transposon insertion into a gene (block arrow) are presented. Arrows indicate the direction of the HPH gene. Tn7-L (147 bp) or Tn7-R (140 bp) sequences that appeared in the 3′-ends of the TAGKO cDNAs (Fig. 1C) are shown below each diagram. Putative polyadenylation signals are in bold and stop codons in three different reading frames are underlined.
M. graminicola were able to grow on medium supplemented with tyrosine (Fig. 3B). In contrast, all the TAGKO strains expected to produce truncated HPD4 proteins showed limited growth on tyrosine (Fig. 3), indicating that the enzymes were either partially functional or non-functional. The *M. grisea* KO1 mutant was expected to make a full-length HPD4 protein, thus retaining the WT phenotype (Fig. 3).

**Yeast CBS mutants can be complemented by *M. grisea* CBS1 cDNAs**

Cystathionine β-synthase is involved in the trans-sulfuration pathways that allow the inter-conversion of cysteine and methionine (25). *Magnaporthe grisea* CBS1 is a functional and structural homolog of the *S. cerevisiae* CBS gene (18). We have generated three different TAGKO CBS1 mutants in *M. grisea* (Fig. 4A). Like the HPD4 mutants, the CBS1 TAGKO mutants produce 3'-truncated transcripts which are chimeric and polyadenylated (data not shown). In yeast, cysteine biosynthesis occurs exclusively through the pathway involving the enzyme CBS (26). Thus, null mutants of CBS are auxotrophic for cysteine.

Expression constructs containing CBS1 cDNA isolated from different *M. grisea* strains were transformed into a yeast CBS null mutant (Fig. 4B). As shown in Figure 4C, growth of the yeast mutant in the absence of cysteine was rescued by the expression of WT, KO1 and KO2 CBS1 cDNAs from *M. grisea*. However, KO3 failed to reverse the mutant phenotype in yeast, indicating that the CBS1 TAGKO protein was not functional. In fact, almost 200 endogenous amino acids were removed from the C-terminus in the KO3 TAGKO protein. The gene product of CBS1 KO2 (Fig. 4D) was predicted to be identical to that of the WT since the transposon insertion occurred downstream of the endogenous stop codon. The KO1 TAGKO protein (Fig. 4D), which was truncated by 65 endogenous amino acids at the C-terminus, was still functional. Interestingly, a human CBS protein with 145 amino acids truncated at the C-terminus was also catalytically...
active (26). The C-terminal domain of human CBS inhibits the enzyme activity, which is regulated positively by S-adenosylmethionine (26). Our findings suggest that a similar functional domain may be present in the *M. grisea* CBS1 protein.

**DISCUSSION**

Transposons have been conveniently and widely used as insertional elements, providing physical markers, and transcriptional and translational fusions to target genes (6). In this study, we uncovered a novel feature in Tn7 through the analysis of TAGKO alleles in *M. grisea* and *M. graminicola*. TAGKO insertion affects the expression of target genes in the fungi by generating 3′-end truncated transcripts that are chimeric and polyadenylated. Interestingly, the TAGKO transcripts contain cryptic eukaryotic polyadenylation signals that are derived from the Tn7 ends (Tn7-L, ATTAAA; Tn7-R, AATATA). A number of eukaryotic transposons are known to contain 3′-site formation signal that causes premature polyadenylation of target gene transcripts, e.g. gypsy ([*Drosophila* (27)], Mu1 [maize (28)] and Fos1 [Fusarium oxysporum (15)]. As a result, truncated transcripts are often detected for the interrupted genes. To our knowledge, this study represents the first report of a bacterial transposon carrying cryptic sequence information for 3′-end formation in eukaryotic transcripts.

Polyadenylation is an essential step in the maturation of mRNA in eukaryotic cells. Following transcription, the 3′-ends of mRNAs are processed by endonucleolytic cleavage and addition of a poly(A) tail. The precise mechanisms of mRNA 3′-end formation in filamentous fungi are yet to be defined (29,30). Nevertheless, it is generally believed that polyadenylation promotes the initiation of translation and the export of mRNA from the nucleus (31). Polyadenylation also confers stability upon mRNA and its removal precedes the degradation of certain mRNA species (31). The poly(A) tails are frequently located 10–30 nucleotides downstream of the polyadenylation signal ‘AAUAAA’. While the hexanucleotide is not an absolute feature, consensus or related sequences are present in a number of filamentous fungal genes (29). In our studies, the sequence ‘AAUAA’ is present 12 nucleotides upstream of the poly(A) tail in the *HPD4* transcript of *M. graminicola* (data not shown). Similarly, the sequence ‘AUUAA’ was found 33 nucleotides upstream of the poly(A) tail in the *HPD4* transcript of *M. grisea* (data not shown). Thus, we reason that the cryptic poly(A) signals identified in the Tn7 ends are compatible with the endogenous mechanisms for mRNA 3′-site formation in filamentous fungi.

Transcripts containing premature stop codons are liable to rapid mRNA degradation (33). Thus, in *in vitro* studies of truncated gene products are possible in the TAGKO mutants, presenting a powerful tool for dissection of gene functions. For example, an allelic series of TAGKO mutants can be generated to identify functional domains by studying mutant phenotypes (Fig. 3). Truncated transcripts can be isolated from the mutants and analyzed in yeast complementation assays (Fig. 4), and in *in vitro* activity and binding assays. In addition, leaky mutations could result from truncation of gene products and are useful for the analysis of lethal genes.

Our TAGKO system generates truncated proteins with C-terminal Tn7-derived extensions. At present we do not have conclusive evidence to suggest whether these amino acid extensions would affect protein functions. At least in the case of *M. grisea CBS1* KO1, the 9-amino acid Tn7-R derived extension appears to be compatible with the enzyme function. The stop codons in Tn7-R are relatively close to the terminal end and the resulting extensions in all frames contain less than 10 amino acids. On the other hand, the longer Tn7-L derived extension (17 amino acids) in the *M. grisea* HPD4 KO2 mutant might have inhibited the function of the TAGKO protein in which 41 endogenous amino acids were removed (Fig. 3). In fact, truncated proteins in frame with the third codon in Tn7-L (Fig. 2) would have the longest C-terminal extension (22 amino acids).

Biery et al. (6) found that a mini-Tn7 element composed of only the terminal 70-bp Tn7-R fragment at both ends are sufficient to allow active *in vitro* transposition reactions. Importantly, the cryptic polyadenylation signal and the three stop codons are all located in this region (Fig. 2). In addition, nucleotide substitutions could be made to the mini-Tn7 end sequences without affecting transposition efficiencies (6). Therefore, to refine our TAGKO system for protein function analysis in filamentous fungi, it is possible to engineer the stop codons closer to the terminal ends in a mini-Tn7 element to generate truncated proteins with minimal amino acid extensions.

**ACKNOWLEDGEMENT**

The authors thank members of the DNA Technology Group at Paradigm Genetics, Inc. for their sequencing efforts.

**REFERENCES**


