1322–1332 *Nucleic Acids Research, 2007, Vol. 35, No. 4* *Published online 31 January 2007*
doi:10.1093/nar/gkm003

The gateway pDEST17 expression vector encodes a −1 ribosomal frameshifting sequence

Eric J. Belfield*, Richard K. Hughes, Nicolas Tsesmetzis, Mike J. Naldrett and Rod Casey

John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK

Received October 16, 2006; Revised December 22, 2006; Accepted December 23, 2006

ABSTRACT

The *att*B1 site in the Gateway (Invitrogen) bacterial expression vector pDEST17, necessary for *in vitro* site-specific recombination, contains the sequence AAA-AAA. The sequence A-AAA-AAG within the *Escherichia coli dna*X gene is recognized as 'slippery' and promotes −1 translational frameshifting. We show here, by expressing in *E. coli* several plant cDNAs with and without single nucleotide deletions close to the translation initiation codons, that pDEST17 is intrinsically susceptible to −1 ribosomal frameshifting at the sequence C-AAA-AAA. The deletion mutants produce correct-sized, active enzymes with a good correlation between enzyme amount and activity. We demonstrate unambiguously the frameshift through a combination of Edman degradation, MALDI-ToF mass fingerprint analysis of tryptic peptides and MALDI-ToF reflectron in-source decay (rISD) sequencing. The degree of frameshifting depends on the nature of the sequence being expressed and ranged from 25 to 60%. These findings suggest that caution should be exercised when employing pDEST17 for high-level protein expression and that the *att*B1 site has some potential as a tool for studying −1 frameshifting.

INTRODUCTION

Although ribosomes normally accurately translate mRNAs into proteins, sequences in certain mRNAs direct the ribosome to undergo non-canonical translation events including: (1) translational read-through of stop codons where the ribosome incorporates an amino acid residue at this position, (2) translational bypassing where a peptidyl-tRNA::ribosome complex 'hops' to a codon further downstream in the mRNA and resumes protein chain elongation and (3) programmed translational frameshifting (hereafter referred to as 'frameshifting' for brevity), in which stimulatory signals in the mRNA induce the ribosome to slip one nucleotide upstream or downstream and then resume protein synthesis in the −1 or +1 alternative open reading frame (ORF). These events, termed 'recoding', result in non-standard translation of mRNA-encoded information that is not normally expressed (1–3).

Frameshifting is involved in the expression, in the broad sense of protein production, of a minority of genes in a wide range of organisms including viruses, bacteria and eukaryotes (3). Evidence suggests that it usually occurs due to the stalling of the ribosome at a stimulatory mRNA structure, such as a pseudoknot or a stem loop, located a few nucleotides downstream of a 'slippery' sequence such as the heptamer X-XXY-YYZ (X and Z can be any nucleotide, and Y can be A or U) (1,2), but there are other slippery sequences that do not conform to this motif (4). Although the secondary structure downstream of a slippery sequence causes the ribosome to pause (5–7), pausing itself is not sufficient to effectuate frameshifting, as stem loops and pseudoknots of similar thermodynamic stability are not necessarily effective frameshift stimulators (8). The parameters known to contribute to the efficiency of −1 frameshifting are the sequence of the slippery heptamer, the downstream secondary structure, and the length and sequence of the spacer between the two *cis*-acting signals (9–11).

In comparison to +1 frameshifting, there are relatively few described examples of −1 frameshifting in cellular genes (3). One example with obvious biological relevance in bacteria is the slippery sequence A-AAA-AAG of the *Escherichia coli dna*X gene (12). When the full-length mRNA of this gene is translated it encodes the τ subunit (71.1 kDa) of DNA polymerase III. However, around 50% of the time the ribosomes that initiate translation frameshift to the −1 frame at the slippery sequence approximately two-thirds of the way through the coding region before terminating to synthesize the shorter γ subunit (47.5 kDa) of the holoenzyme (13–15). The frameshift occurs at the A-AAA-AAG sequence by tandem slippage of both P- and A-site $tRNA^{Lys}$ species from the 0 (A-AAA-AAG) to the −1 frame (AAA-AAA)

*To whom correspondence should be addressed. Tel: +44 (0)1603 450517; Fax: (0)1603 450025; E-mail: eric.belfield@bbsrc.ac.uk

(16,17). This process is dependent upon two stimulatory sequences, a Shine–Dalgarno-like sequence 10- nucleotides upstream of the A-AAA-AAG and a stem-loop structure 5- nucleotides downstream of it (18).

In this study, we investigated whether the DNA sequence AAA-AAA present in the Gateway (Invitrogen) bacterial expression vector pDEST17 *att*B1 site, which is necessary for *in vitro* site-specific recombination, is prone to −1 frameshifting. We cloned three cDNA-encoding enzymes of the plant oxylipin pathway (19) to study the effect in *E. coli*: an *Arabidopsis thaliana* allene oxide synthase (AOS), a *Medicago truncatula* hydroperoxide lyase (HPL) and a *Pisum sativum* lipoxygenase (LOX). The cDNAs were cloned as 0 frame wild-type sequences to assess control expression levels with the pDEST17 vector but were also cloned missing one nucleotide from the 5′ end of the cDNAs, such that enzyme synthesis was dependent on a −1 frameshifting event. Our findings indicate that genes expressed using the Gateway pDEST17 vector can undergo a remarkably high degree of −1 frameshifting at this slippery sequence.

MATERIALS AND METHODS

Cloning and expression of oxylipin enzymes

A *M. truncatula* cDNA clone (NF034B10IN1F1080), which encoded a predicted full-length HPL (GenBank accession number: AJ316562), was obtained from the Samuel Roberts Noble Foundation, Ardmore, USA and named *Mt*HPLF. An *A. thaliana* (Columbia ecotype) full-length AOS cDNA clone (U17068) was acquired from the *Arabidopsis* Biological Resource Centre, The Ohio State University, USA, and was named *At*AOS (GenBank accession number: AF172727). A *P. sativum* LOX cDNA (*Ps*LOX3) (GenBank accession number: X07807) used in this study was previously cloned by Ealing and Casey (20). All clones were propagated with appropriate antibiotics and plasmid DNA was extracted (Wizard SV Minipreps, Promega).

The cDNA sequences of *At*AOS, *Mt*HPLF and *Ps*LOX3 were PCR amplified with *Pfu* Ultra according to the manufacturer's instructions (Stratagene). PCR products were purified and cloned into pDONR201 entry vector via the BP reaction (Gateway Technology, Invitrogen). For all cDNAs, two pENTRY clones from individual bacterial colonies were subsequently used in LR reactions (Invitrogen) with the T7 promoter expression vector pDEST17 (to obtain N-terminally fused 6× His-tagged proteins). Cloned PCR products were sequence checked in the pDEST vectors and were as predicted.

To determine if a −1 frameshifting event was occurring at the pDEST17 sequence AAA-AAA (Figure 1), a restorative single-nucleotide deletion was incorporated into the primers used to amplify the oxylipin enzyme cDNA sequences and these clones were termed frameshift (FS). Nucleotides shown underlined were removed from *At*AOS and *Ps*LOX3 *att*B1 primers (below and Table 1) to produce the FS constructs. The *At*AOS cDNA was amplified without the first 96 bp predicted to encode a 32 amino acid N-terminal chloroplast targeting sequence (ChloroP; (21)), using the forward *att*B1 primer 5′-GGG GACAAGTTTGTACAAAAAAGCAGGCTTG**G**CTTC CGGGTCAGAAACTCC-3′ and reverse *att*B2 primer 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTAC TAAAAGCTAGCTTTCCTTAACGAC-3′. The *Ps*LOX3 cDNA was amplified using the forward *att*B1 primer

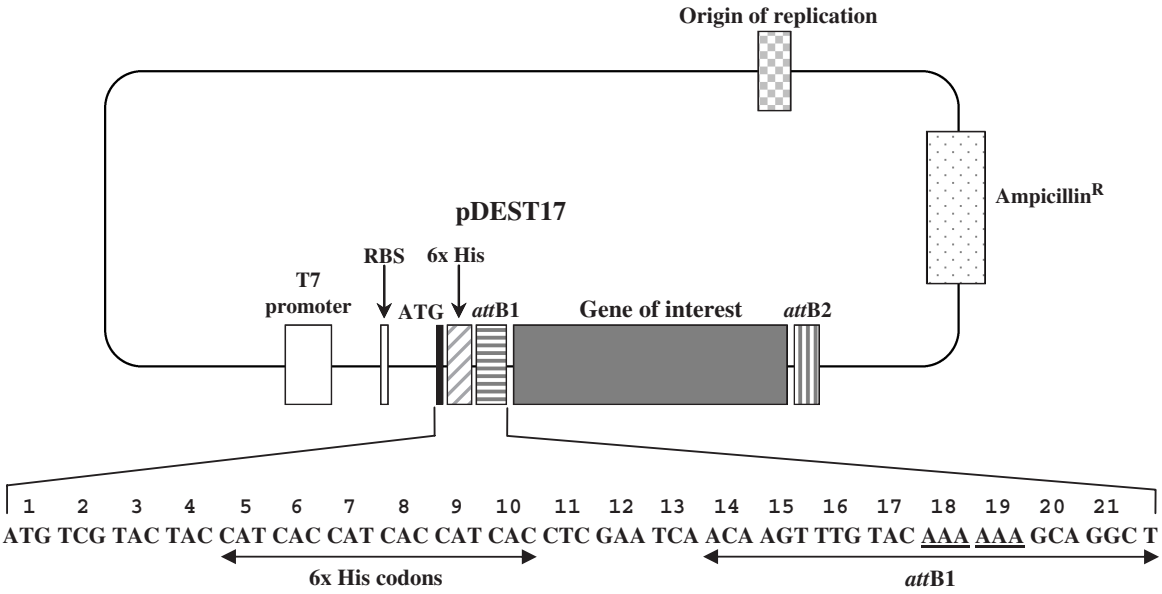

Figure 1. Diagrammatic representation of the pDEST17 expression vector. A schematic (not to scale) illustrating the Gateway bacterial expression vector and its nucleotide sequence from the initiating codon (black box), the N-terminal 6× His codon fusion (diagonal grey stripes) and the *att*B1 recombination sequence (horizontal grey stripes) containing the putative slippery site underlined. T7 promoter and the ribosome-binding site (RBS) are shown in open boxes. Codon numbers are shown above nucleotide triplets.

5′-ACAAGTTTGTACAAAAAAGCAGGCTTCAT GTTTTCAGGCGTGACTGGTATTCTGAAT-3′ and reverse *att*B2 primer 5′-GGGGACCACTTTGTACAAG AAAGCTGGGTCCTAGATGGAGATACTATTAGG-3′. Sequence analysis of two independently cloned *Mt*HPLF cDNAs revealed one had been PCR amplified and cloned without errors (*Mt*HPLF-WT) but the second had serendipitously been amplified with a 1-bp guanine nucleotide deletion (*Mt*HPLF-FS) located three bases into the cDNA sequence. The *Mt*HPLF cDNAs were amplified using the forward *att*B1 primer 5′-GGGGA CAAGTTTGTACAAAAAAGCAGGCTCAATGGCTT CCTCATCAGAAACC-3′ and reverse *att*B2 primer 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTAT CAGACGGTGGATGAAGCCTTAAC-3′. Sequencing of the *At*AOS and *Ps*LOX3 pDEST17 DNA inserts indicated sequences both with and without the 1-bp deletion were error free and as predicted.

Expression and purification

Cultures (10 ml or 1 l Luria–Bertani broth without glucose, containing 50 μg/ml ampicillin (Melford Laboratories Ltd)) of *E. coli* strain BL21 (DE3) transformed with expression vectors were grown at 37°C to A_{600} 1.0–1.1 with shaking at 200 r.p.m., transferred to 21°C, and gene expression was induced with isopropyl β-D-thiogalactopyranoside (IPTG; 1 mM) for 24 h. Cells were harvested by centrifugation at 28 000 × *g* and the pellets frozen at −80°C. Cell pellets were thawed and extracted at room temperature with BugBuster (Novagen) according to the manufacturer's instructions. Homogenates were then transferred to 50-ml Oakridge tubes, vortexed for 1 min and mixed gently by inversion on a Spiramix 5 (Denley) for 20 min. His-tagged proteins were purified at 4°C as described earlier for *Mt*HPLF by immobilized metal affinity chromatography using cobalt as a ligand (22). For removal of detergent and histidine from the proteins the concentrated samples were then injected onto a HiLoad Superdex 26/60 gel filtration column (GE Healthcare) or a HiPrep 26/10 rapid desalting column (GE Healthcare) equilibrated with 100 mM sodium phosphate buffer, pH 6.5 and eluted with the same buffer at 2 ml/min (gel filtration) or 10 ml/min (desalting). The concentration of *Mt*HPLF was determined using a molar extinction coefficient of 120 000 M^{-1} cm^{-1} at 391 nm (22). The Reinheitzahl (Rz) value of the purified *Mt*HPLF protein preparations was ∼1.3, indicating purification to homogeneity.

Table 1. The expression clones used to study pDEST17 frameshifting. To generate WT and FS constructs, the *At*AOS, *Ps*LOX3 and *Mt*HPLF gene sequences were cloned downstream of the pDEST17 *att*B1 site (bold). The nucleotides underlined indicate the bases removed to study −1 ribosomal frameshifting events occurring at the *att*B1 sequence

Expression clone	Frame	Sequence
*At*AOS-WT	WT	5′-**TAC AAA AAA GCA GGC TTG** GCT TCC GGG-3′
*At*AOS-FS	−1	5′-**TAC AAA AAA GCA GGC TTG** CTT CCG GGT-3′
*Mt*HPLF-WT	WT	5′-**TAC AAA AAA GCA GGC TCA** ATG GCT TCC-3′
*Mt*HPLF-FS	−1	5′-**TAC AAA AAA GCA GGC TCA** ATG CTT CCT-3′
*Ps*LOX3-WT	WT	5′-**TAC AAA AAA GCA GGC TTC** ATG TTT TCA-3′
*Ps*LOX3-FS	−1	5′-**TAC AAA AAA GCA GGC TTA** TGT TTT CAG-3′

Reverse transcription of RNA to determine translational frameshifting or transcriptional slippage

Total cellular RNA was isolated from *E. coli* BL21 (DE3) cells expressing WT and FS *At*AOS, *Mt*HPLF and *Ps*LOX3 cells 6 h post-IPTG induction using the RNAeasy kit (Qiagen) according to the manufacturer's instructions. First strand cDNA synthesis was performed using the Omniscript reverse transcriptase (Qiagen) as recommended by the manufacturer with an oligo specific to the *At*AOS, *Mt*HPLF or *Ps*LOX3; AOSR1 5′-CGTTGACGGCATGTAAGTACC-3′, HPLF4Rev 5′-CTAGACTTCACTGTCCATGC-3′, or *Ps*LOX3-3205 5′-GTGAGATTATCAACCGTGGAACCG-3′). PCR reactions were performed using *Pfu* Ultra DNA polymerase (Stratagene) according to the manufacturer's instructions using a primer set designed to the expression vector: pDEST17 His tag, 5′-CATCACCATCACCATCAC-3′; and the cDNA-specific oxylipin oligo above. To check for plasmid contamination, samples of the RNA preparations were RNase treated using standard methods (23) and gave no PCR signal, indicating that there was no DNA contamination. The PCR cDNA products were run on ethidium-stained agarose gels, the bands cut and cleaned using the QIAquick gel extraction kit (Qiagen) and sequenced.

Enzyme activity measurements

*At*AOS and *Mt*HPLF activities were determined in a 0.5-ml assay mixture containing 20 μM 13-HPOT (13(*S*)-hydroperoxy-9(*Z*), 11(*E*), 15(*Z*)-octadecatrienoic acid, supplied by Prof Mats Hamberg (Karolinska Institute, Sweden)), in 100 mM sodium phosphate buffer, pH 6.5. The decrease in A_{234} was followed for 20–60 s at 25°C and converted to moles of substrate using a molar absorption coefficient of 25 mM^{-1} cm^{-1} (24). *Ps*LOX3 activity was measured using linoleic acid as a substrate according to (25).

SDS-PAGE and western blot analysis

Protein concentration was determined using Bradford reagent (BioRad) or the BCA Protein Assay kit (Pierce) with bovine serum albumin (BSA) as a reference. Samples for SDS-PAGE were prepared by mixing aliquots of the protein with NuPAGE sample buffer (Invitrogen) and heated at 70°C for 10 min. Protein samples were run on NuPAGE 4–12% gradient Bis-Tris gels at 150 V for 1 h with MES SDS running buffer (Invitrogen) and stained with Coomassie blue. For western blot analysis, gels were electrotransferred to a Protran BA 85 nitrocellulose

membrane (Schleicher and Schuell BioScience) using the Xcell Surelock electrophoresis and transfer apparatus (Invitrogen). The membrane was blocked overnight at 4°C in 3% (w/v) BSA, Tris-buffered saline solution containing 0.05% Tween 20. Proteins were detected using a mouse anti-His tag monoclonal antibody (Novagen) and a goat anti-mouse secondary antibody conjugated to alkaline phosphatase (Novagen), and colour was developed.

Edman sequencing

An aliquot of 1 nmol of *Mt*HPLF-FS or *Mt*HPLF-WT protein taken from a 10 mg/ml solution in 100 mM sodium phosphate buffer, pH 6.5 was diluted to 30 µl with water and applied to a ProSorb™ cartridge (Applied Biosystems) according to the manufacturer's instructions. Sequencing was carried out from the polyvinylidene difluoride (PVDF) disc using a model 494 Procise sequencer (Applied Biosystems) run in the pulsed-liquid mode.

Mass spectrometry

All mass spectrometry was carried out on a standard pulsed ion extraction Reflex III MALDI-ToF mass spectrometer (Bruker) equipped with a 2-GHz digitizer and gridless reflector and source. A 337-nm-wavelength nitrogen laser was used to desorb/ionize the matrix/analyte material, and ions were detected in positive ion reflectron mode.

MALDI-ToF peptide mass fingerprinting

Samples were run on SDS-PAGE, excised, reductively alkylated and digested with porcine-modified sequencing grade trypsin (Promega). Acidified digests were spotted directly onto a thin layer of matrix on a stainless steel target plate for analysis by MALDI-ToF MS. The matrix consisted of the following: four parts of a saturated solution of α-cyano-4-hydroxycinnamic acid in acetone was mixed with one part of a 1:1 mixture of acetone:isopropanol containing 10 mg/ml nitrocellulose. Digests were externally calibrated against a calibration curve of seven peptides to yield data with mass accuracies of better than 50 ppm. These calibrated spectra were searched against a weekly updated copy of the SPtrEMBL database using an in-house copy of the Mascot search tool (www.matrixscience.com).

MALDI-ToF reflectron in-source decay (rISD) analysis

*Mt*HPLF-FS protein (Rz 1.3) was diluted to give a sample concentration of 20–50 pmol/µl. Samples were prepared for MALDI by mixing with a saturated solution of matrix in the ratio 1:1. Matrix solution was made by dissolving 3,5-dimethoxy-4-hydroxycinnamic acid (Fluka) in 30% acetonitrile/0.05% trifluoroacetic acid to saturation. About 0.5 µl of this combined mix was spotted onto a polished stainless steel target and allowed to crystallize prior to analysis. rISD spectra were obtained by first ascertaining that the intact protein could be seen clearly. An optimized parameter set, with a pulsed ion extraction medium delay setting, was used to zoom in on the mass range 1000–4000 and the laser power increased until fragmentation along the protein backbone could be seen. Spectra of 1000–2000 shots were acquired. Sinapinic acid was used specifically to encourage fragmentation at the N-terminus of the protein. Calibration was carried out using the standard peptide mixture used for peptide mass fingerprinting of this same mass range.

Site-directed mutagenesis

Mutations altering individual nucleotides (in bold and underlined below), designed to destabilize the predicted *Mt*HPL-FS stem-loop structure, were generated using an oligonucleotide-directed *in vitro* mutagenesis kit (QuikChange; Stratagene). The *Mt*HPL-FS cDNA was modified using the following mutagenic oligonucleotides with their complementary sequences; T29C 5′-GTACAAAAAAGCAGGCTCAA**C**GCTTCCTCAT CAGAAAC-3′ and T29C-ANTISENSE 5′-GTTTCTG ATGAGGAAGC**G**TTGAGCCTGCTTTTTTGTAC-3′; C31G 5′-AAAAAAGCAGGCTCAATG**G**TTCCTCAT CAGAAACCTC-3′ and C31G ANTISENSE 5′-GA GGTTTCTGATGAGGAA**C**CATTGAGCCTGCTTTT TT-3′. The T29C nucleotide modification altered the *Mt*HPL-FS codon from AAT to AAC but did not change the amino acid incorporated into the *Mt*HPL-FS peptide as both codons encode the residue asparagine. However, modification of the C31G nucleotide did change the amino acid from alanine (encoded by GCU) to glycine (encoded by GGU), which is not a rare *E. coli* codon. The number in the name of each oligonucleotide refers to the number of nucleotides from the 5′ adenine of the *Mt*HPL-FS ATG start codon. All mutations were sequenced to confirm veracity.

Statistical analysis of the *E. coli* genome

The file NC_000913.ffn containing the nucleotide sequences of the coding regions of the K12 genome was downloaded from the National Center for Biotechnology Information's website (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/Escherichia_coli_K12/). A Perl script was written to count the occurrences of C-AAA-AAA and A-AAA in the right frame in the coding sequences contained in the file NC_000913.ffn. Another Perl script was written to produce a table of codon usage from this file and the probability with which a particular codon encodes a particular amino acid was calculated by dividing its codon usage by the sum of usages of all synonymous codons.

RESULTS

Frameshift assay constructs

Three plant oxylipin cDNAs, *At*AOS, *Mt*HPLF and *Ps*LOX3, were cloned into the pDEST17 vector in the correct reading frame (WT) according to the manufacturer's instructions. The first 32 N-terminal amino acids of the *At*AOS protein, which is a predicted chloroplast-targeting sequence that potentially could reduce enzyme activity, was omitted as part of the cloning. In addition, constructs were made with a one-base deletion (FS) in the

oxylipin cDNA to study possible frameshifting events at the homopolymeric adenine sequence, AAA-AAA, within the *att*B1 site of the Invitrogen pDEST17 vector. A transcriptional slippage or translational frameshifting event occurring at this sequence would obviate the effect of the deletion in these clones and produce active enzymes. The constructs were sequenced and transformed into *E. coli* BL21 (DE3) expression cells.

Measurements of enzyme activity

Enzymatic assays of the crude *Mt*HPLF-FS *E. coli* extracts showed that a fully functional protein was being produced, even though a truncated peptide of 35 amino acids (aa) (3.9 kDa), 469 aa shorter than the wild-type *Mt*HPLF-WT protein, was predicted. Kinetic data using the substrate 13-HPOT (Figure 2) showed the specific activity after 24-h induction was over twice as high (11.84 µmol/min/mg protein versus 5.05 µmol/min/mg protein) for the *Mt*HPLF-FS frameshift mutant expression clone compared to the wild-type *Mt*HPLF-WT.

To determine if this frameshifting effect was a phenomenon particular to the *Mt*HPLF cDNA, we cloned and expressed other DNA sequences both with and without a restorative 1-bp deletion near the 5′ ends. The first, *At*AOS, was a member of the same P450 sub-family of CYP74s as the *Mt*HPLF, and the second, *Ps*LOX3, was an unrelated cDNA encoding a plant non-haem iron-containing LOX that catalyses the oxidation of polyunsaturated fatty acids.

The *At*AOS-FS and *Ps*LOX3-FS constructs were predicted to encode truncated peptides of 3.5 and 3.1 kDa, respectively. Enzymatic activity assays revealed, however, that both the FS mutants and WT clones expressed active, soluble proteins in crude *E. coli* extracts. Interestingly, the specific activity ratios of the wild type, compared to mutant, *Ps*LOX3 and *At*AOS enzymes differed from that of *Mt*HPLF. The pDEST17 *Ps*LOX3-WT and -FS clones expressed proteins with roughly equal units of specific activity (0.5 µmol/min/mg protein) with the substrate linoleic acid after 24 h, whereas the *At*AOS-WT activity was around fourfold higher (16.7 µmol/min/mg protein) than the corresponding deletion mutant (3.9 µmol/min/mg protein) using 13-HPOT as a substrate.

Western blot analysis of WT and FS clones

To confirm functionality of the expressed proteins, we performed western blot semi-quantification with *E. coli* crude extracts using an anti-His tag antibody (Figure 3). This experiment gave similar results to the activity measurements described above: there were almost threefold (292%) higher quantities of *At*AOS-WT protein expressed in *E. coli* cultures compared to the FS clone and the amount of *Mt*HPLF-FS protein expressed was 49% higher than that of the WT clone, showing that the increased enzyme activity data correlated with increased protein amounts. However, the amount of *Ps*LOX3 was 74% higher in the WT compared to the FS cultures even though the activity measurements for these clones were almost the same. This latter result may be due to the lack of sensitivity of the spectrophotometric assay in measuring the relatively low LOX activity. The important point is that all three frameshift cDNAs produced correct-sized, active enzymes. Based on the western data the frameshifting efficiencies are approximately 60% (*Mt*HPLF), 36% (*Ps*LOX) and 25% (*At*AOS).

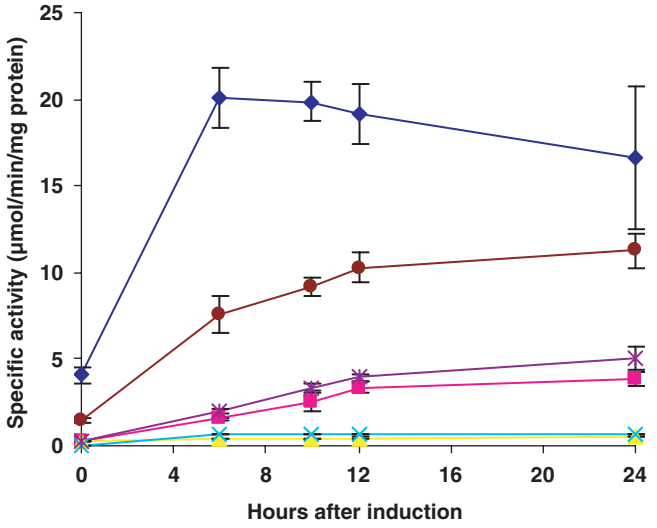

Figure 2. Time course of *At*AOS, *Ps*LOX3 and *Mt*HPLF WT- and FS-specific activities as a function of time after induction. *Escherichia. coli* BL21 (DE3) transformed with pDEST17 *At*AOS-WT (blue line), *At*AOS-WT-FS (pink line), *Mt*HPLF-WT (purple line), *Mt*HPLF-FS (brown line), *Ps*LOX3-WT (yellow line) and *Ps*LOX3-FS (cyan line) clones were induced with 1 mM IPTG and cell samples collected at the indicated times. Enzyme activities were assayed as described in the Materials and methods section.

Nucleotide sequence analysis of *At*AOS, *Mt*HPLF and *Ps*LOX3 RNAs

To examine if translation into an alternative reading frame might be due to transcriptional slippage, the pDEST17 *att*B1 region containing the putative slippery site AAA-AAA was amplified from the mRNA populations of *At*AOS, *Mt*HPLF and *Ps*LOX3 frameshift clones. Sequence analysis of the cDNA-amplified products spanning the A run showed a single RNA species was present with six As (data not shown) from all clones, with no additional A nucleotides. This indicates a −1 translational frameshifting event occurred with all clones, rather than transcriptional slippage where a heterogeneous mixture of six and seven A residues would be expected to restore the ORF (26). It is very unlikely that a homogeneous population of transcripts was preferentially amplified, because the primer used to amplify the cDNA and the primer set used to amplify a PCR product flanked the *att*B1 region where the putative slippery sequence was located.

Ribosomal frameshifting determination

To verify directly that the ribosomal frameshifting had occurred at the AAA-AAA sequence in the RNA, the *Mt*HPLF-FS and -WT proteins were over-expressed in *E. coli* and homogeneous proteins were obtained by

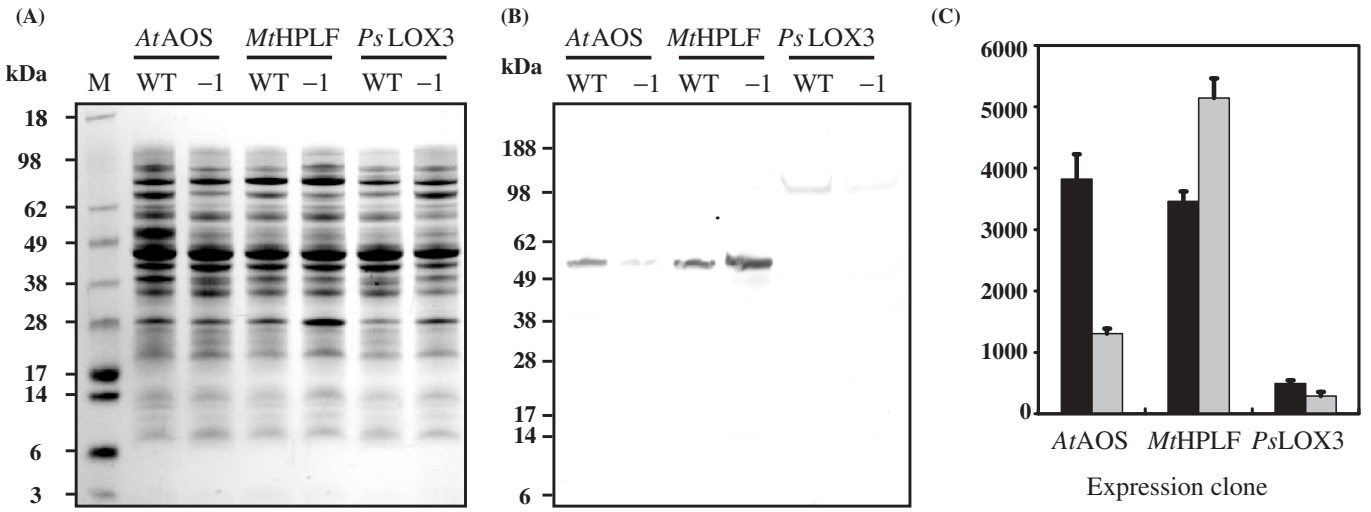

Figure 3. Quantification of His-tagged *At*AOS, *Ps*LOX3 and *Mt*HPLF proteins from bacterial lysates. Total proteins were extracted from cells producing WT and FS enzymes and were separated on a 4–12% polyacrylamide gradient SDS-PAGE gel. About 20 µg total proteins were stained with Coomassie blue (**A**) to check protein quantifications and loading levels. Lane M, molecular weight marker (SeeBlue Plus2; Invitrogen); lane 1 *At*AOS wild type, lane 2 *At*AOS frameshift; lane 3 *Mt*HPLF wild type; lane 4 *Mt*HPLF frameshift; lane 5 *Ps*LOX3 wild type, lane 6 *Ps*LOX3 frameshift. For western blot protein quantifications, 20 µg of total protein from cells expressing *At*AOS or *Mt*HPLF and 60 µg of total protein from cells expressing *Ps*LOX3 were transferred to a nitrocellulose membrane and detected with a monoclonal antibody against the His tag (**B**) (see Materials and Methods for further details). (**C**) Digitally quantified band intensities from western blotting showing means and standard errors from three independent samples. Black bars—WT, grey bars—FS.

FPLC (22). MALDI-ToF MS analysis of tryptic peptides from purified *Mt*HPLF-FS and *Mt*HPLF-WT revealed that most of the peptides aligned over the entire length of proteins, but information was apparently lacking from the N-termini. To determine precisely where the frameshift occurs, we obtained the unequivocal sequence of the first 21 N-terminal amino acids of *Mt*HPLF-FS using Edman degradation. Residues 22–24 were not clear. In-source decay sequencing yielded residues 11–35, aiding the interpretation of that part of the sequence obtained by Edman sequencing that was less clear. Residues 1–19 corresponded to amino acids encoded by the pDEST17 vector. At residue 20, a strong signal for Ser was detected; this corresponds to the residue encoded by the −1 reading frame, followed by Arg and Leu that would be conventionally encoded after the leftward frameshift (Figure 4). Armed with the revised sequence, the masses of the N-terminal peptides were calculated and found to be present in the initial MALDI peptide mass fingerprint data.

Statistical analysis of the *E. coli* genome

To determine if the putative −1 frameshifting Gateway *att*B1 slippery sequences were under-represented in *E. coli*, statistical analysis was performed. In the *E. coli* genome, there are 9897 instances of A-AAA and 333 instances of C-AAA-AAA. The genes are listed in Supplementary Tables 1 and 2, respectively. To assess possible representation biases of the sequences, codon usage for AAA (3.36%), and the occurrence of A or C in the wobble position (18.01 and 26.83%, respectively) were used (codon usage is shown in Supplementary Table 3). With unselected bias, in 1 346 260 codons of annotated *E. coli* K12 ORFs, the sequence A-AAA should occur $1\,346\,260 \times 0.1801 \times 0.0336 = 8147$ times and C-AAA-AAA should occur $1\,360\,013 \times 0.2683 \times 0.0336 \times 0.0336 = 412$ times. This estimate does not take into account that the sequences cannot occur in the first or/and in the last position of the ORF. Therefore, the C-AAA-AAA sequence is under-represented and A-AAA is over-represented in the *E. coli* genome (~81 and 121% of the expected values, respectively).

DISCUSSION

Programmed translational frameshifting is an alternative mechanism of translation for a minority of genes and is used by probably all organisms (27). Slippery sequences are the cause for most −1 frameshifting events at runs of homopolymeric nucleotides where the tRNAs tandemly slip at the P- and A-sites, while maintaining the identity of two nucleotides in the codon–anticodon interaction (1). In this study, we predicted the nucleotide sequence AAA-AAA present in the Invitrogen Gateway vector pDEST17 *att*B1 recombination site may be ribosomally slippery and prone to −1 translational frameshifting. To test this hypothesis we cloned three different plant oxylipin cDNAs, *At*AOS, *Mt*HPLF and *Ps*LOX3, as WT clones that require no recoding to produce functional enzymes and as potential −1 FS constructs that would synthesize active enzymes only if a recoding event took place.

Initial activity assays indicated that both WT and FS constructs expressed His-tagged functional proteins for all enzymes, even though the predicted conventionally translated −1 FS construct products were short peptides of less than 5 kDa. Western blot analysis of crude *E. coli* extracts using a His-tagged antibody did not detect any

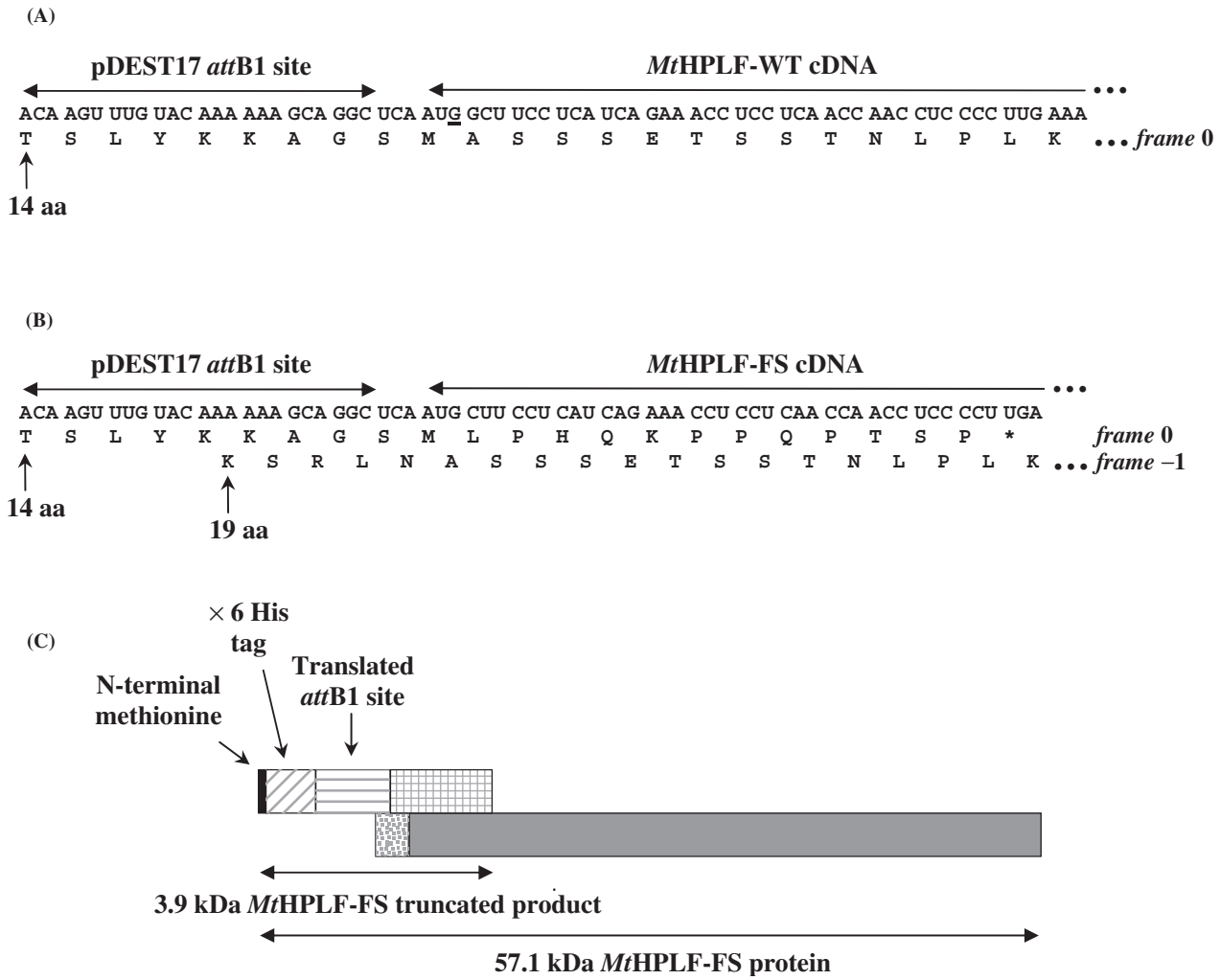

Figure 4. The pDEST17 frameshift region. (**A**) The predicted RNA and deduced amino acid sequences of the pDEST17 *Mt*HPLF-WT clone. Single-letter abbreviations for amino acid residues are used. (**B**) The predicted RNA sequences of pDEST17 *Mt*HPLF-FS and deduced amino acid sequences for both the 0 frame and the −1 frameshift transcripts. The *Mt*HPLF-FS construct has a guanine nucleotide deletion three bases into the original *Mt*HPLF cDNA sequence (underlined in (4A)), to repair ribosomal −1 frameshifting at the homopolymeric adenine sequence, AAA-AAA, and restores the *Mt*HPLF 0 ORF (one amino acid into the HPL protein), yielding a product of approximately the same size as the WT fusion protein. (**C**) Diagrammatic representation of the predicted translated peptides from the pDEST17-expressed *Mt*HPLF-FS RNAs and location of the frameshift region. The hatched box at the C-terminus of the ~4-kDa product represents the amino acids MLPHQKPPQPTSP and the dotted box near the N-terminus of the ~57-kDa product represents the amino acids RLNA. Solid grey *Mt*HPLF protein sequence box not to scale.

small peptides of this size, possibly because truncated peptide products tend to aggregate and form insoluble inclusion bodies. It is possible that the correct ORF of the −1 FS constructs was repaired by transcriptional slippage at the homopolymeric stretch of adenines in the pDEST17 *att*B1 sequence, as happens in the A_6 tract of the rat p53 gene to yield an insertion of an extra A in ~9% of subcloned transcripts (28). Insertion and deletion mutation within DNA sequences at stretches of As or Ts is a well-documented phenomenon; such sequences appear to be particularly vulnerable (29) due to misaligned loops (30). We show here by RT-PCR that there is a single homogeneous mRNA population with six adenines and no other insertions or deletions in the expected transcripts of the *At*AOS, *Mt*HPLF and *Ps*LOX3 −1 FS constructs, indicating that transcriptional slippage was not responsible for the ORF restoration. Direct evidence that −1 translational frameshifting was occurring came from protein sequencing of the *Mt*HPLF-FS product, which showed that the reading frame shift at the pDEST17 *att*B1 sequence results in a serine residue, encoded by the nucleotides AGC, instead of an alanine (GCA) at the predicted 20th amino acid position. We propose −1 frameshifting is occurring, within the Gateway *att*B1 recombination site, via one of two possible mechanisms.

The first is the ribosomal tandem slippage of two lysyl-tRNAs at the P- and A-sites at the heptameric sequence C-AAA-AAA and the second is a single slippage event of a peptidyl $tRNA^{Lys}$ at the hexanucleotide sequence AAA-AAA. The former mechanism occurs by translocation of a slippery heptameric sequence,

C-AAA-AAA, to the −1 phase CAA-AAA, causing −1 frameshifting. The slippage of the $tRNA^{Lys}$ at the P-site would maintain two codon–anticodon base-pairs (XAA), and the new codon–anticodon, $tRNA^{Lys}$, occupying the A-site in the −1 ORF, would have matches at all three nucleotide positions (AAA) relative to the lysine in the 0 ORF. This hypothesis is consistent with the tandem slippage model proposed by Jacks *et al.* (16) and later refined by Weiss *et al.* (12), where each of the two ribosome-bound tRNAs at the P- and A-sites slip in the 5′ direction to the −1 ORF only when each tRNA maintains at least two codon–anticodon base pairs with the RNA in the −1 shifted frame.

The second possible mechanism supposes that a potential slippery tetramer with the sequence A-AAA is sufficient to cause −1 frameshifting by slippage of a peptidyl $tRNA^{Lys}$ to the −1 ORF when the A-site is unoccupied. Two reported single tRNA slippage cases with obvious similarities to the −1 frameshifting site in this article are the genes that encode the capsid protein/nucleic-acid-binding 12K (CP/12K) of potato virus M (PVM) (31) and the *ins*A-*ins*B fusion protein of the IS*1* insertion sequence (32). In both cases frameshifting occurs by −1 slippage of a peptidyl $tRNA^{Lys}$ bound to AAA onto the overlapping AAA codon at an A-AAA motif. The single tRNA slippage of PVM and IS*1* frameshifts are extremely inefficient, allowing between 0.3 and 1% of ribosomes to shift frames respectively. This low efficiency may be a consequence of the very unusual mechanism (33). The number of examples of the second proposed mechanism are low; almost all cases of −1 frameshifting occur by tandem slippage of tRNA anticodons on heptanucleotide shift sites (16). The large difference in frameshifting efficiency between the mechanisms may suggest that the more traditional canonical frameshift event, of tandem tRNA slippage, could be responsible for the −1 frameshifting of the pDEST17 vector, because the frameshifting efficiency is high.

It would be informative to modify by mutagenesis nucleotides in the heptameric slippery site for comparative composition studies on the efficiency of −1 frameshifting but we are unable to do this due to Invitrogen's strict Limited Use Label Licenses No. 19 of the Gateway vector pDEST17 which states: 'The buyer cannot modify the recombination sequence(s) contained in this product for any purpose'.

The cloning of cDNAs for enzymes of oxylipin metabolism allowed us to accurately quantify the recoding efficiencies by measuring the relative amounts of enzyme activity produced from the FS- and the WT- cloned sequences. High −1 frameshifting efficiencies, of up to 60% compared to the control, were observed when the *Mt*HPLF-FS construct was expressed in *E. coli*.

Frameshifting rates are generally dependent on a number of stimulatory sequences, including a downstream hairpin or pseudoknot that causes slowing, or pausing, of the ribosome long enough at the slippery sequence for frameshifting to occur (34,35) and an upstream Shine–Dalgarno sequence that pairs with the 16S RNA, causing ribosomal 'stress' (1). Our analysis of the sequence proximal to the pDEST17 slippery site found no potential Shine–Dalgarno sequences. Secondary structure analysis for stimulatory frameshifting sequences such as hairpin loops (26), or pseudoknots (36) (Figure 5) of the highly frameshifting *Mt*HPLF-FS revealed that the RNA fold with the lowest minimum free energy had a potential stem-loop of five predicted Watson–Crick base pairs (mfold (37)). This hairpin is thermodynamically more stable and compact than those predicted in the RNAs for the FS *Ps*LOX3 and *At*AOS stem-loop structures, which have a maximum of two and four consecutive paired nucleotides, respectively and may explain why lower frameshift rates (of 36 and 25%, respectively) were observed compared to the *Mt*HPLF-FS clone. The *Mt*HPLF-FS secondary structure has some of the characteristics that Antao and Tinoco (38) reported; they found that hairpin tetraloops, with stem sizes of four or five bases, could form extra stable hairpins when the loop-closing base pair was A–U. If the stem-loop structures predicted in Figure 5 are real then a combination of hairpin thermodynamic stability and/or the identity of the A–U loop-closing base pair could explain why the frameshifting efficiency is higher for *Mt*HPLF-FS RNA than the *Ps*LOX3 and *At*AOS transcripts. To test this hypothesis we modified the *Mt*HPLF-FS RNA molecule using site-directed mutagenesis. First, we generated a single nucleotide point mutant to change the A–U loop-closing pair to a non complementary A-C pair. In addition, we modified a cytosine to a guanine that was predicted to be the third nucleotide of a five base stem of a stem-loop structure (see Figure 5). The predicted secondary structures of the mutated *Mt*HPLF-FS RNAs were thermodynamically less stable (A–C pair: $\Delta G = -2.8$ kcal/mol and cytosine to a guanine: $\Delta G = -0.1$ kcal/mol) compared to the original *Mt*HPLF-FS RNA ($\Delta G = -3.6$ kcal/mol) (RNA secondary structures predicted by the mfold programme (37) are not shown). The specific activities of both mutants with the substrate 13-HPOT (data not shown) were not significantly different from that of the *Mt*HPLF-FS clone even though the proposed stem loops in both mutants were predicted to be thermodynamically less stable suggesting the predicted secondary structures may not be accurate.

We also used an alternative secondary structure prediction program, CONTRAfold, which allows full-length RNA sequences to be submitted and is reported to achieve the highest single-sequence prediction accuracies to date (39). The probabilistic models (data not shown) produced by CONTRAfold showed that the only RNA sequence without any secondary structure around the slippage site was the *Mt*HPLF-FS sequence. In fact, all other structures for both WT and FS clones used in this study were predicted by CONTRAfold to have stem-loop structures at the slippery site, with the 6A nucleotides forming the loop of the hairpin. This RNA fold program therefore suggests that secondary structures are not required to enhance the high level of −1 frameshifting of *Mt*HPLF-FS and the slippage rate is actually reduced when predicted RNA secondary structures are present, as in *At*AOS-FS and *Ps*LOX3-FS. One explanation could be that the heptameric sequence is unusually slippery, obviating the need for a stimulatory secondary

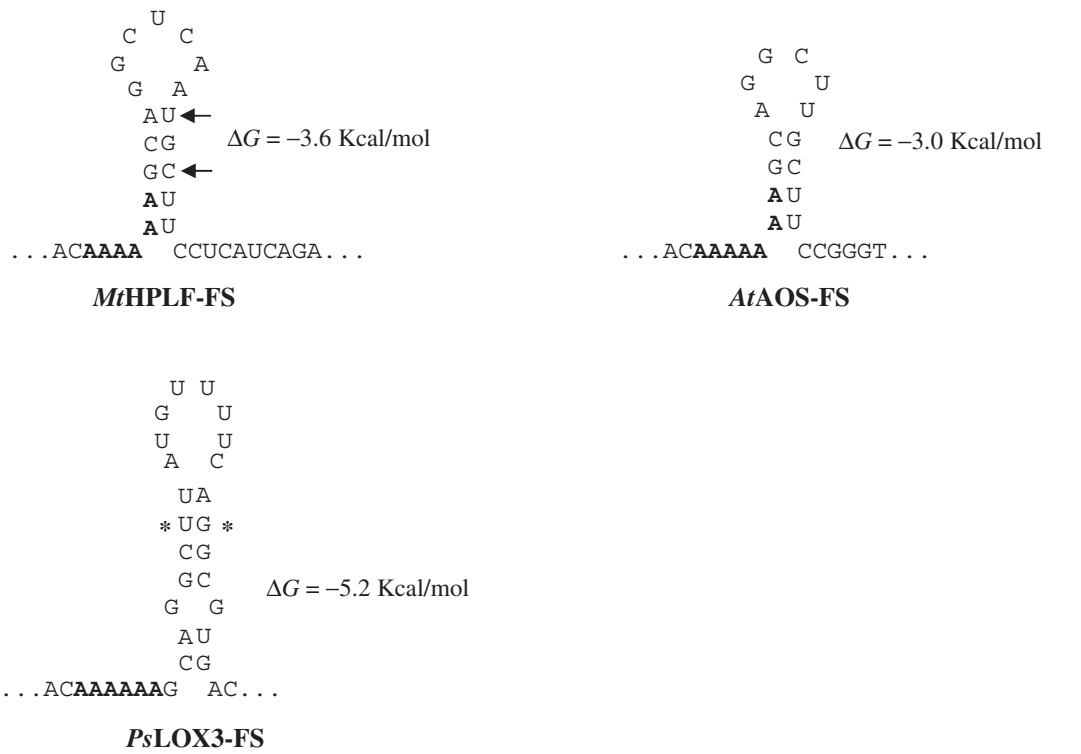

Figure 5. Sequence and predicted secondary structures of the *Mt*HPLF, *At*AOS and *Ps*LOX3-FS RNAs and putative −1 frameshift-inducing site. The frameshift slippery sequence is shown in bold. All nucleotides are Watson–Crick base paired except the U:G wobble marked with a star in the *Ps*LOX3-FS RNA structure. The stem-loop structures were predicted using the mfold software (38). Thermodynamic free energy (ΔG) was calculated at 37°C and 1 M Na^+ concentration. The arrows next to the *Mt*HPLF-FS RNA structure indicate the nucleotides mutated to study the possible consequences of changes in the thermodynamic stability of the stem loop on −1 frameshifting.

structure (33). Wilson *et al.* (40) showed that a 26-nucleotide sequence containing the homopolymeric HIV-1 sequence U-UUU-UUA was efficient at −1 frameshifting in rabbit reticuloctyte and yeast cell-free translation systems and was not dependent on stem-loop structures, although later experiments by Parkin *et al.* (41), using the entire HIV-1 *gag-pol* region, showed that a distal stem-loop was necessary for maximal frameshifting.

Short nucleotide sequences that lead to highly efficient frameshifting without any stimulatory sequences, as mentioned above, are extremely rare. One example is +1 frameshifting in *Saccharomyces cerevisiae* of the Ty1 retrotransposon (42). The Ty1 frameshift sequence of only seven nucleotides, CUU-AGG-C, results in the synthesis of a TYA-TYB fusion protein. The ribosomal frameshifting occurs when the ribosomal A-site is vacant and the P-site, occupied by the CUU-bound peptidyl-tRNA, slips to the overlapping UUA codon during slow recognition of the next codon, AGG (42). Such stochastic events lead to high levels of erroneous frameshifting, producing aberrant polypeptides. Therefore, short slippery sequences in coding regions are strongly selected against, unless they have evolved to serve a useful function. For instance, the frameshift-prone Ty1 sequence is under-abundant and under-represented in the coding regions of the *S. cerevisiae* genome (43). We estimated whether the two proposed Gateway −1 frameshift sequences were under-represented in the *E. coli* genome. Statistical analysis revealed that the A-AAA sequence was over-represented and the C-AAA-AAA sequence was under-represented (121 and 80.8% of the expected values, respectively). Absence of evidence of negative selection, as in the case of the A-AAA sequence, suggests that this sequence may not be prone to frameshifting, as is the case for the 'non-shifty' sequence A-AAG-AAA that is also over-represented in the *E. coli* genome (132%) (44). The under-represented sequence, C-AAA-AAA, has a similar value to that reported (83%) for the *dnaX* sequence, A-AAA-AAG (44), which is known to be slippery and may suggest there is negative evolutionary selection for this candidate −1 frameshift sequence in the Gateway *att*B1 site. The under-representation may also suggest it may not require any stimulatory elements for efficient frameshifting to occur, similar to Ty1 (42).

We have shown unambiguously that the sequence found within the Gateway expression vector pDEST17 *att*B1 recombination site is prone to −1 ribosomal frameshifting and that the degree of frameshifting is dependent on the sequence being expressed. It is likely the frameshift sequence in the Gateway *att*B1 recombination site is C-AAA-AAA instead of A-AAA. Two lines of evidence support this hypothesis: higher frameshift rates have been reported for tandem tRNA slippage events compared to single-tRNA slippage ones that would occur with the latter sequence; and the former sequence is

under-abundant and under-represented in the *E. coli* genome, suggesting that there may have been negative selection against it during evolution because truncated polypeptide products produced by frameshifting can be deleterious to the cell.

SUPPLEMENTARY DATA

Supplementary Data is available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank the Samuel Roberts Noble Foundation (Ardmore, USA) for providing the cDNA clone *Mt*HPLF, the Arabidopsis Biological Resource Centre (The Ohio State University, USA) for providing the cDNA clone of *At*AOS, Professor M. Hamberg for providing AOS/HPL substrates, Dr Govind Chandra for statistical analysis on the representation of nucleotide sequences in the *E. coli* genome, Karen Wilson and Igor Galetich at the Joint Institute for Food Research–John Innes Centre Proteomics Facility for protein identification using MALDI-ToF MS peptide mapping and amino acid sequence analyses, and an anonymous referee for suggesting analysis of the *E. coli* genome. This work was supported by research grants 'Natural Oxylipins for Defence of Ornamentals' (NODO; European-Union-funded project QLK5-CT-2001-02445) and by the Biotechnology and Biological Sciences Research Council. Funding to pay the Open Access publication charges for this article was provided by the BBSRC.

Conflict of interest statement. None declared.

REFERENCES

1. Farabaugh,P.J. (1996) Programmed translational frameshifting. *Annu. Rev. Genet.*, **30**, 1322–528.
2. Baranov,P.V., Gesteland,R.F. and Atkins,J.F. (2002) Recoding: translational bifurcations in gene expression. *Gene*, **286**, 187–201.
3. Namy,O., Rousset,J.P., Napthine,S. and Brierley,I. (2004) Reprogrammed genetic decoding in cellular gene expression. *Mol. Cell*, **13**, 157–168.
4. Licznar,P., Mejlhede,N., Prere,M.F., Wills,N., Gesteland,R.F., Atkins,J.F. and Fayet,O. (2003) Programmed translational −1 frameshifting on hexanucleotide motifs and the wobble properties of tRNAs. *EMBO J.*, **22**, 4770–4778.
5. Tu,C., Tzeng,T.H. and Bruenn,J.A. (1992) Ribosomal movement impeded at a pseudoknot required for frameshifting. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 8636–8640.
6. Somogyi,P., Jenner,A.J., Brierley,I. and Inglis,S.C. (1993) Ribosomal pausing during translation of an RNA pseudoknot. *Mol. Cell. Biol.*, **13**, 6931–6940.
7. Lopinski,J.D., Dinman,J.D. and Bruenn,J.A. (2000) Kinetics of ribosomal pausing during programmed −1 translational frameshifting. *Mol. Cell. Biol.*, **20**, 1095–1103.
8. Kontos,H., Napthine,S. and Brierley,I. (2001) Ribosomal pausing at a frameshifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. *Mol. Cell. Biol.*, **21**, 8657–8670.
9. Brierley,I. (1995) Ribosomal frameshifting viral RNAs. *J. Gen. Virol.*, **76**, **(Pt 8)**, 1885–1892.
10. Brierley,I. and Pennell,S. (2001) Structure and function of the stimulatory RNAs involved in programmed eukaryotic-1 ribosomal frameshifting. *Cold Spring Harb. Symp. Quant. Biol.*, **66**, 233–248.
11. Bekaert,M., Bidou,L., Denise,A., Duchateau-Nguyen,G., Forest,J.P., Froidevaux,C., Hatin,I., Rousset,J.P. and Termier,M. (2003) Towards a computational model for −1 eukaryotic frameshifting sites. *Bioinformatics*, **19**, 327–335.
12. Weiss,R.B., Dunn,D.M., Shuh,M., Atkins,J.F. and Gesteland,R.F. (1989) *E. coli* ribosomes re-phase on retroviral frameshift signals at rates ranging from 2 to 50 percent. *New Biol.*, **1**, 159–169.
13. Blinkowa,A.L. and Walker,J.R. (1990) Programmed ribosomal frameshifting generates the *Escherichia coli* DNA polymerase III γ from within the τ subunit reading frame. *Nucleic Acids Res.*, **18**, 1725–1729.
14. Flower,A.M. and McHenry,C.S. (1990) The γ subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 3713–3717.
15. Tsuchihashi,Z. and Kornberg,A. (1990) Translational frameshifting generates the γ subunit of DNA polymerase III holoenzyme. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 2516–2520.
16. Jacks,T., Madhani,H.D., Masiarz,F.R. and Varmus,H.E. (1988) Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. *Cell*, **55**, 447–458.
17. Tsuchihashi,Z. and Brown,P.O. (1992) Sequence requirements for efficient translational frameshifting in the *Escherichia coli dna*X gene and the role of an unstable interaction between $tRNA^{Lys}$ and an AAG lysine codon. *Genes Dev.*, **6**, 511–519.
18. Larsen,B., Wills,N.M., Gesteland,R.F. and Atkins,J.F. (1994) rRNA-mRNA base pairing stimulates a programmed −1 ribosomal frameshift. *J. Bacteriol.*, **176**, 6842–6851.
19. Feussner,I. and Wasternack,C. (2002) The lipoxygenase pathway. *Annu. Rev. Plant Biol.*, **53**, 275–297.
20. Ealing,P.M. and Casey,R. (1989) The cDNA cloning of a pea (*Pisum sativum*) seed lipoxygenase. Sequence comparisons of the two major pea seed lipoxygenase isoforms. *Biochem. J*, **264**, 929–932.
21. Emanuelsson,O., Nielsen,H. and von Heijne,G. (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.*, **8**, 978–984.
22. Hughes,R.K., Belfield,E.J., Muthusamy,M., Khan,A., Rowe,A., Harding,S.E., Fairhurst,S.A., Bornemann,S., Ashton,R., Thorneley,R.N. and Casey,R. (2006) Characterization of *Medicago truncatula* (barrel medic) hydroperoxide lyase (CYP74C3), a water-soluble detergent-free cytochrome P450 monomer whose biological activity is defined by monomer-micelle association. *Biochem. J.*, **395**, 641–652.
23. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor NY, US.
24. Vick,B.A. (1991) A spectrophotometric assay for hydroperoxide lyase. *Lipids*, **26**, 315–320.
25. Hughes,R.K., Wu,Z., Robinson,D.S., Hardy,D., West,S.I., Fairhurst,S.A. and Casey,R. (1998) Characterization of authentic recombinant pea-seed lipoxygenases with distinct properties and reaction mechanisms. *Biochem. J.*, **333**, 33–43.
26. Larsen,B., Wills,N.M., Nelson,C., Atkins,J.F. and Gesteland,R.F. (2000) Nonlinearity in genetic decoding: homologous DNA replicase genes use alternatives of transcriptional slippage or translational frameshifting. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 1683–1688.
27. Baranov,P.V., Gurvich,O.L., Fayet,O., Prere,M.F., Miller,W.A., Gesteland,R.F., Atkins,J.F. and Giddings,M.C. (2001) RECODE: a database of frameshifting, bypassing and codon redefinition utilized for gene expression. *Nucleic Acids Res.*, **29**, 264–267.
28. Ba,Y., Tonoki,H., Tada,M., Nakata,D., Hamada,J. and Moriuchi,T. (2000) Transcriptional slippage of *p53* gene enhanced by cellular damage in rat liver: monitoring the slippage by a yeast functional assay. *Mutat. Res.*, **447**, 209–220.
29. Kunkel,T.A. (1990) Misalignment-mediated DNA synthesis errors. *Biochemistry*, **29**, 8003–8011.
30. Gragg,H., Harfe,B.D. and Jinks-Robertson,S. (2002) Base composition of mononucleotide runs affects DNA polymerase slippage and removal of frameshift intermediates by mismatch repair in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **22**, 8756–8762.

31. Gramstat,A., Prufer,D. and Rohde,W. (1994) The nucleic acid-binding zinc finger protein of potato virus M is translated by internal initiation as well as by ribosomal frameshifting involving a shifty stop codon and a novel mechanism of P-site slippage. *Nucleic Acids Res.*, **22**, 3911–3917.
32. Sekine,Y. and Ohtsubo,E. (1992) DNA sequences required for translational frameshifting in production of the transposase encoded by IS1. *Mol. Gen. Genet.*, **235**, 325–332.
33. Farabaugh,P.J. (1996) Programmed translational frameshifting. *Microbiol. Rev.*, **60**, 103–134.
34. Craigen,W.J. and Caskey,C.T. (1986) Expression of peptide chain release factor 2 requires high-efficiency frameshift. *Nature*, **322**, 273–275.
35. Sipley,J. and Goldman,E. (1993) Increased ribosomal accuracy increases a programmed translational frameshift in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 2315–2319.
36. Manktelow,E., Shigemoto,K. and Brierley,I. (2005) Characterization of the frameshift signal of *Edr*, a mammalian example of programmed −1 ribosomal frameshifting. *Nucleic Acids. Res.*, **33**, 1553–1563.
37. Zuker,M. (2003) mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids. Res.*, **31**, 3406–3415.
38. Antao,V.P. and Tinoco,I., Jr (1992) Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids. Res.*, **20**, 819–824.
39. Do,C.B., Woods,D.A. and Batzoglou,S. (2006) CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics*, **22**, e90–e98.
40. Wilson,W., Braddock,M., Adams,S.E., Rathjen,P.D., Kingsman,S.M. and Kingsman,A.J. (1988) HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. *Cell*, **55**, 1159–1169.
41. Parkin,N.T., Chamorro,M. and Varmus,H.E. (1992) Human immunodeficiency virus type 1 *gag-pol* frameshifting is dependent on downstream mRNA secondary structure: demonstration by expression *in vivo*. *J. Virol.*, **66**, 5147–5151.
42. Belcourt,M.F. and Farabaugh,P.J. (1990) Ribosomal frameshifting in the yeast retrotransposon Ty:tRNAs induce slippage on a 7 nucleotide minimal site. *Cell*, **62**, 339–352.
43. Shah,A.A., Giddings,M.C., Parvaz,J.B., Gesteland,R.F., Atkins,J.F. and Ivanov,I.P. (2002) Computational identification of putative programmed translational frameshift sites. *Bioinformatics*, **18**, 1046–1053.
44. Gurvich,O.L., Baranov,P.V., Zhou,J., Hammer,A.W., Gesteland,R.F. and Atkins,J.F. (2003) Sequences that direct significant levels of frameshifting are frequent in coding regions of *Escherichia coli*. *EMBO J.*, **22**, 5941–5950.