Characterization of the frameshift signal of Edr, a mammalian example of programmed –1 ribosomal frameshifting

Emily Manktelow, Kazuhiro Shigemoto1 and Ian Brierley*

Division of Virology, Department of Pathology University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK and 1Department of Environmental Health and Social Medicine, Ehime University School of Medicine, Shitsukawa, Toon, Ehime 791-0295 Japan

Received January 26, 2005; Revised and Accepted February 23, 2005

ABSTRACT

The ribosomal frameshifting signal of the mouse embryonal carcinoma differentiation regulated (Edr) gene represents the sole documented example of programmed –1 frameshifting in mammalian cellular genes [Shigemoto,K., Brennan,J., Walls,E., Watson,C.J., Stott,D., Rigby,P.W. and Reith,A.D. (2001), Nucleic Acids Res., 29, 4079–4088]. Here, we have employed site-directed mutagenesis and RNA structure probing to characterize the Edr signal. We began by confirming the functionality and magnitude of the signal and the role of a GGGAAAC motif as the slippery sequence. Subsequently, we derived a model of the Edr stimulatory RNA and assessed its similarity to those stimulatory RNAs found at viral frameshift sites. We found that the structure is an RNA pseudoknot possessing features typical of retroviral frameshifter pseudoknots. From these experiments, we conclude that the Edr signal and by inference, the human orthologue PEG10, do not represent a novel ‘cellular class’ of programmed –1 ribosomal frameshift signal, but rather are similar to viral examples, albeit with some interesting features. The similarity to viral frameshift signals may complicate the design of antiviral therapies that target the frameshift process.

INTRODUCTION

Programmed –1 ribosomal frameshifting (hereafter frameshifting for brevity) is a translational control mechanism that allows the production of a specific ratio of gene products from two overlapping open reading frames, the relative quantities of which depend upon the frameshift efficiency at that particular site. Frameshifting occurs during the elongation phase of protein synthesis where, in response to elements in the mRNA, the ribosome switches from the zero reading frame to the –1 frame (in the 5’ direction) at a defined position, and translation continues in the new frame. The frameshift signals of this class, first described in retroviruses (1,2), have subsequently been found (mainly) in other virus genomes and in Escherichia coli insertional elements [reviewed in (3–6)]. The mRNA signals that promote frameshifting comprise a slippery sequence, where the frameshift takes place, and a 3’-stimulatory RNA structure, separated from the slippery sequence by a short spacer region. The heptanucleotide slippery sequence typically contains consecutive homopolymeric triplets (XXXXYYZ), with the ribosome-bound tRNAs decoding the P- and A-site codons slipping from the zero frame (X XXY YYZ) to the –1 frame (XXX YYY). Studies of frameshifting in eukaryotic systems have indicated that X can be any nucleotide, Y is almost always A or U, and Z any nucleotide but G (although the YYZ codon is often AAG in prokaryotic systems). Efficient frameshifting also requires the presence of the 3’-stimulatory RNA beginning ~5–9 nt downstream of the slippery sequence. At some sites, a simple stem–loop structure appears to be necessary and sufficient for frameshifting but more commonly, an RNA pseudoknot structure is present. These are H-(hairpin)-type pseudoknots (7), although some viral pseudoknots have also been described as kissing hairpins (8,9). The mechanism of frameshifting is not fully understood, although a number of models have been proposed to explain how the interaction of the ribosome with the stimulatory RNA leads to a realignment of the tRNAs decoding the slippery sequence into the –1 frame ([10] and references therein).

Computer-assisted database screens have highlighted the potential for –1 frameshifting in conventional cellular genes (11–13), but only two examples with obvious biological relevance have been identified to date. The first is present in the
E. coli dnaX gene, encoding the γ and τ subunits of DNA polymerase III holoenzyme. The synthesis of the γ protein is by frameshifting, which directs ribosomes to a premature stop codon, while the longest form (τ) is translated by continued standard decoding (14,15). The frameshift occurs at the slippery sequence AAAAAAG, by simultaneous slippage over the RNA template, facilitated by replacing dGTP with deaza-GTP in the sequenc-
RNA structure probing

RNAs for secondary structure probing were prepared by in vitro transcription of NdeI-cut pKT0/Edr/T3 using bacteriophage T3 RNA polymerase. Transcription reactions were on a 200 µl scale and contained 20 µg plasmid DNA, 2.5 mM of each rNTP and 200 U of T3 RNA polymerase (Promega) in a buffer containing 40 mM Tris, pH 8, 15 mM MgCl₂ and 5 mM DTT. After 3 h at 37°C, 100 U of DNase I was added and the incubation continued for a further 30 min. Nucleic acids were harvested by extraction with phenol/chloroform (1:1) and ethanol precipitation. The RNA was quantified by spectrophotometry and its integrity checked by electrophoresis on a 2% (w/v) agarose gel containing 0.1% SDS. Transcripts were concentrated by ethanol precipitation. DNA fragments were removed by Sephadex G-50 chromatography and the RNA transcripts concentrated by ethanol precipitation. The RNA was quantified by spectrophotometry and its integrity checked by electrophoresis on a 2% (w/v) agarose gel containing 0.1% SDS. Transcripts (10 µg) were 5’ end-labelled with [γ-32P]ATP using a standard dephosphorylation–rephosphorylation strategy (31), purified from 6% acrylamide–urea gels and dissolved in water.

The structure probing experiments followed the general principles outlined by others (32–34). All reactions contained 10–50 000 c.p.m. 5’ end-labelled RNA transcript. RNase probing reactions were carried out in 50 µl reaction volumes containing 50 µg carrier E. coli rRNA (Sigma). Enzymatic probing reactions were carried out on ice for 20 min. RNase CL3 [Industrial Research Laboratories (IRL), New Zealand] and RNase T1 (Ambion) probing was in 50 mM sodium cacodylate, pH 7, 2 mM MgCl₂ and 0–0.2 units CL3 or T1; RNase V1 (Ambion) in 10 mM Tris, pH 8, 2 mM MgCl₂, 0.1 M KCl and 0–0.2 units V1; and RNase U2 (IRL) in 20 mM sodium acetate, pH 4.8, 2 mM MgCl₂, 100 mM KCl and 0–0.2 units U₂. Enzyme reactions were stopped by the addition of 150 µl ethanol and the RNA recovered by centrifugation. RNAs were prepared for analysis on 6, 10 or 15% polyacrylamide–7 M urea sequencing-type gels [with or without 20% (v/v) formamide] by dissolution in water and mixing with an equal volume of formamide gel loading buffer [95% (v/v) formamide, 10 mM EDTA, 0.1% bromophenol blue and 0.1% xylene cyanol], before heating at 80°C for 3 min.

Chemical probing was performed with lead acetate and imidazole in 10 µl reaction volumes. Lead probing was at 25°C for 5 min in 20 mM HEPES-NaOH, pH 7.5, 5 mM Mg acetate, 50 mM K acetate and 1–5 mM Pb acetate. Reactions were stopped by the addition of EDTA to 33 mM and the RNA recovered by precipitation in ethanol, redissolved in water and prepared for gel loading as above. For imidazole probing, the end-labelled RNA was mixed with 10 µg carrier RNA, dried in a desiccator and redissolved in 10 µl of 2 M imidazole, pH 7, containing 40 mM NaCl and 10 mM MgCl₂. After incubation at 37°C for 2–4 h, the reaction was stopped by the addition of 100 µl of a fresh solution of 2% (w/v) lithium perchlorate in acetone. The RNA was recovered by centrifugation, washed with acetone, dried, dissolved in water and prepared for gel loading as above. All structure probing gels included an alkaline hydrolysis ladder as a size marker, prepared by dissolving the dried pellet from 3 µl of end-labelled RNA and 10 µg carrier RNA in 3 µl of 22.5 mM
NaHCO₃, 2.5 mM Na₂CO₃ and boiling for 2 min. After the addition of an equal volume of formamide gel loading buffer and heating to 80°C for 3 min, the sample was loaded immediately onto the gel.

RESULTS

Mapping the 3' boundary of the Edr frameshift signal

The 5' and 3' boundaries of the Edr frameshift signal were not characterized in the previous study (20), but we began our analysis with the assumption that the 5' boundary is the GGGAAAC heptamer located within the RF1/RF2 overlap region. A 1230 bp portion of the Edr gene containing this slippery sequence and flanked by substantial lengths of open reading frame was cloned by PCR into the expression vector pKT0 (28) (see Materials and Methods) to generate pKT0/Edr (Figure 1). This plasmid contains a bacteriophage SP6 (and T7) promoter for the generation of in vitro transcripts and provides an optimized (Kozak) initiation codon for the translation of the cloned Edr segment. The cloned PCR product was sequenced to confirm that no errors had been introduced into the Edr sequence. To assess crudely the extent to which 3’ sequences were required for frameshifting, a series of run-off transcripts containing varying lengths of Edr sequence were prepared following digestion of the plasmid with NdeI, AvrII, HindIII or BamHI and transcription with SP6 RNA polymerase (Figure 1). The mRNAs were translated in the RRL in vitro translation system and the products analysed by SDS/PAGE (Figure 2). Frameshifting was observed in all cases with each mRNA showing an efficiency of ~30% (hereafter the ‘wild-type’ efficiency) in good agreement with the value obtained by Shigemoto et al. (20). The sizes of the non-frameshifted and frameshifted species were consistent with those expected following frameshifting at the RF1/RF2 overlap region (see legend to Figure 2). As the mRNA from the NdeI-digested plasmid (linearized 105 bp downstream of the slippery sequence) stimulated wild-type levels of frameshifting, subsequent investigations focussed on regions upstream of this restriction site.

To further delineate the essential sequences, in-frame deletions of 24, 63 or 90 nt were introduced into pKT0/Edr in the region between the GGGAAAC motif and the NdeI site (Figure 3A). The plasmids were linearized with HindIII before in vitro transcription and translation. As shown in Figure 3B, efficient frameshifting was observed only with pKT0/Edr/Δ24nt, hence the stimulatory RNA is in fact longer than that proposed in the original study (20) (Figure 3A, emboldened bases), with a 3’ boundary located between the 5’ edge of the deletions of pKT0/Edr/Δ63 and Δ24. In addition to the non-frameshifted and frameshifted species, an additional longer product was seen in these translation reactions (Figure 3B, asterisk) and in fact, in all translations of pKT0/Edr to a greater or lesser extent. The identity of this product is uncertain. It is unlikely to be an alternative frameshift product, since its size was unaffected by changes in the length of the mRNA downstream of the frameshift site. However, given that its intensity paralleled that of the non-frameshifted product and its size was roughly double that of the non-frameshifted product, it may represent a stable dimer.

Proposed structure of the frameshift stimulatory signal of the Edr gene

The results of the deletion analysis prompted a re-evaluation of the folding possibilities within the RNA downstream of the slippery sequence. Scrutiny of potential base-pairing interactions within this extended region indicated the potential for the formation of an RNA pseudoknot structure, shown in Figure 4B in comparison with the previous models of Shigemoto et al. (20) (Figure 4A). The proposed pseudoknot is positioned ~5 nt downstream of the GGGAAAC sequence and contains two relatively long stems, linked by loops of three (loop 1) and nine (loop 2) nucleotides. Stem 1 can be subdivided into two regions (1a and 1b) separated by a 3 nt bulge in the second arm.

![Figure 2](image-url)

Figure 2. Confirming functionality of the Edr frameshift signal. pKT0/Edr was linearized with NdeI, AvrII, HindIII or BamHI, transcribed with T7 RNA polymerase and transcripts translated in RRL, either undiluted (1; final concentration ~50 μg/ml) or diluted 1/3 (1/3; about 15 μg/ml). Products were labelled with [35S]methionine, separated on a 15% SDS/polyacrylamide gel and detected by autoradiography. The non-frameshifted (stop) and frameshifted (FS) species are marked with arrows. M represents [14C] protein markers (Amersham Pharmacia Biotech). C is a control translation of an mRNA derived from EcoRI-linearized plasmid p2Luc/MMTV gag/pro (53). The first G of the putative Edr slippery sequence (GGGAAAC) is at position 1390 (in the mRNA sequence). The sites of cleavage of the four restriction endonucleases used above are 1501 (NdeI), 1764 (AvrII), 2028 (HindIII) and 2082 (BamHI) and are predicted to specify frameshift products of 33, 43, 52 and 54 kDa, respectively. The non-frameshifted product is predicted to be 29 kDa in each case. The frameshift efficiencies measured for each mRNA are shown (Fe%).
**Figure 3.** Deletion analysis of the Edr frameshift signal. (A) Three independent in-frame deletions were created in pKT0/Edr, Δ90, Δ63 and Δ24, to investigate the requirement for sequence information downstream of the putative Edr slippery sequence GGGAAAC (boxed). Each deletion (of 90, 63 or 24 nt) was to a common 3' site (immediately downstream of the NdeI site, which was removed), leaving varying lengths of 5' sequence. The nucleotides implicated as forming the stimulatory RNA in a previous study (20) are in bold. (B) The three deletion mutants were digested with HindIII, transcribed with SP6 RNA polymerase and transcripts translated in RRL, either undiluted (1; final concentration/μg/ml) or diluted 1/3 (1/3; about 15 μg/ml). Products were labelled with [35S]methionine, separated on a 15% SDS/polyacrylamide gel and detected by autoradiography. The non-frameshifted (stop) and frameshifted (FS) species are marked with arrows. M represents 14C protein markers (Amersham Pharmacia Biotech). The species indicated with an asterisk is discussed in the text.

**Figure 4.** Proposed foldings of the Edr stimulatory RNA. (A) The stem–loop and pseudoknot models of Shigemoto et al. (20). (B) New pseudoknot model based on the results of the deletion analysis of Figure 3. In the new model, stem 1 is considered as two helices, 1a and 1b, separated by a 3 nt bulge (–ACA–) in the second arm.
Site-directed mutagenesis of the Edr frameshift signal

To test our model of the Edr frameshift signal, site-directed mutagenesis was carried out. The mutants fell into four groups (i) changes in the putative slippery sequence GGGAAAC, (ii) destabilizing and restabilizing mutations within the predicted stem regions 1a, 1b and 2 (iii) a mutation that deleted the bulge triplet between stems 1a and 1b and (iv) a mutation that changed the central seven (of nine) nucleotides in loop 2 to their complementary Watson–Crick bases. Figure 5 shows the mutations (m1–m12) introduced into the frameshift region of pKT0/Edr (Figure 5A), in vitro translations of the mutant.

**Figure 5.** Analysis of the Edr frameshift signal by site-directed mutagenesis. (A) A series of mutations were introduced into the Edr frameshift region to modify the proposed slippery sequence (GGGAAAC, in bold) or pseudoknot. (B) Wild-type pKT0/Edr or mutant derivatives were digested with HindIII, transcribed with SP6 RNA polymerase and transcripts translated in RRL at a concentration of ~50 µg/ml. Products were labelled with [35S]methionine, separated on 15% SDS/polyacrylamide gels and detected by autoradiography. The non-frameshifted (stop) and frameshifted (FS) species are marked with arrows. M represents 14C protein markers (Amersham Pharmacia Biotech). (C) Summary of the mutations made and the resulting frameshift efficiencies. In constructs pKT0/m5, m8 and m11, both arms of the relevant stem region were mutated such that the stems should reform (double/revertant).
constructs (Figure 5B; mRNAs derived from HindIII-cut plasmids) and a summary of the frameshift efficiencies measured for each construct (Figure 5C). As before the sizes of the non-frameshifted and frameshift species expected from these mRNAs were 29 and 45 kDa, respectively.

The two independent mutations created within the putative slippery sequence had the second base of each homopolymeric triplet changed to a C residue (m1 and m2). These changes would reduce the ability of the tRNAs decoding this sequence to slip into the −1 reading frame and consistent with this, each change reduced the frameshifting efficiency (Figure 5). As seen with other frameshift signals, the mutation in the second homopolymeric triplet (GGGACAC, m1), decoded in the ribosomal A-site during the frameshift, was highly inhibitory (frameshifting was essentially abolished), whereas the reduction in efficiency with the P-site change (GGGAAAAC, m2) was less dramatic (a 4-fold reduction). These data are entirely consistent with the belief that the GGGAAC stretch is indeed the site of the frameshift in Edr and confirm the earlier observations of Shigemoto et al. (20).

Analysis of the proposed stem regions involved the introduction of complementary and compensatory changes within stems 1a, 1b and 2. Three mutations were prepared for each stem. Two of these were destabilizing mutations, introduced into each arm of the relevant stem by changing three central base pairs to their complementary Watson–Crick bases. The third was a double mutation leading to a ‘pseudowild-type’ structure, in which both changes were made and should be compensatory. As shown in Figure 5, destabilization of any stem reduced frameshifting efficiency (m3, m4, m6, m7, m9 and m10), but frameshifting was restored in the double mutant, pseudowild-type constructs (m5, m8 and m11) supporting the belief that the stems form and are required for frameshifting. However, there was some stem-specific variation in the magnitude by which frameshifting was reduced by stem destabilization and the extent to which frameshifting was restored in the double mutants. For stem 1a, the reduction in frameshift efficiency after disruption of individual arms was quite modest (~2.5-fold for the 1st arm; ~6-fold for the 2nd arm), but frameshifting efficiency was restored to close to that of the wild-type in the double mutant construct (m5, 25.6%). Disruption of stem 1b in the same manner also led to a reduction in frameshifting efficiency, but more dramatically, with frameshifting reduced ~15-fold for either arm. As with the stem 1a double mutant, frameshifting efficiency was restored to close to that of the wild-type value with the stem 1b pseudowild-type construct (m8, 24.6%). Disruption of stem 2 was also inhibitory, with efficiency values of ~2% for mutations in the individual arms. However, although the double mutant construct showed a rise in efficiency in comparison with the single mutants, frameshifting was restored only to about half that of the wild-type construct (m11, 16.5%).

The presence of the −ACA− bulge in stem 1 was of interest as a triplet bulge has been observed in the frameshift stimulatory RNA of human immunodeficiency virus type 1 (HIV-1) and shown to contribute to frameshifting, although the bulge is considered to be a component of a stem–loop stimulatory RNA in this system (35). However, we found that deletion of the bulge had a very minor effect on Edr frameshifting, reducing the efficiency by only a few percent (m13, 26.8%).

The final mutation concerned loop 2. In a number of frameshift-stimulating pseudoknots, an adenosine-rich triplex formed between loop 2 and the minor groove of stem 1 has been described and contributes to frameshifting (23,24,36,37). A conserved 5′-AACAA-3′ motif in such loops has its bases rotated by varying degrees to allow interactions with bases on both strands in the minor groove of stem 1. It is an RNA-specific feature, with each interaction involving a hydrogen bond formed from a ribose 2′ hydroxyl group. Loop 2 of the proposed Edr pseudoknot is relatively A-rich (five of nine bases) and includes a 5′-AACGA-3′ stretch. To assess whether loop 2–stem 1 interactions may occur in Edr, the central seven bases of loop 2 were changed to their Watson–Crick complementary nucleotides and frameshifting measured. A modest reduction in frameshifting was observed (m12, 23.2%), but not so dramatic as to suggest a major role for stem 1–loop 2 interactions in this pseudoknot, at least those mediated by runs of loop adenosines, since only one loop 2 adenosine remains in this mutant.

Structure probing of the Edr frameshift signal

The mutagenesis data provided strong support for our model of the Edr frameshift region, but it was important to confirm the main features by RNA structure mapping. To facilitate this, a bacteriophage T3 promoter was inserted into pKT0/Edr ∼33 nt upstream of the slippery sequence (Figure 1), the plasmid linearized with Ndel and a 130 nt T3 transcript (encompassing the frameshift region) prepared. This was end-labelled with [γ-32P]ATP, gel purified and subjected to chemical and enzymatic digestion before analysis on denaturing polyacrylamide gels. Four enzymatic probes were used: RNases CL3, T1 and U2, which preferentially cleave single-stranded C, G and A residues, respectively, and RNase CV1, which targets double-stranded and stacked bases. The chemical probes imidazole and lead acetate were also employed, which show specificity for single-stranded regions. In these experiments, the Mg2+ level was kept at 2 mM, which is the approximate concentration of this ion in RRL. (38). Representative structure probing gels are shown in Figure 6 and a summary is shown in Figure 7.

The structure probing data were strongly supportive of the pseudoknot model of Figure 4. The cleavage pattern of imidazole especially was highly consistent with the model, with cleavage occurring only within loops 1 and 2, at the very ends of the stems and in the −ACA− bulge of stem 1. Lead acetate gave a similar cleavage pattern, although there were more cuts at the ends of the stems. Taking all of the probing reagents into account, the pattern of reactivities of stem 2, loop 1 and loop 2 matched the structure prediction very closely. Stem 2 appeared to be very stable in comparison with the same region of other frameshift-stimulating pseudoknots we have studied (39–41). It was cleaved almost exclusively by the double-strand specific RNase CV1, being unreactive with single-stranded enzymatic probes and showing only occasional, weak cleavage with single-stranded chemical probes. The loops also showed appropriate reactivity. Both were cleaved by single-stranded chemical probes and loop 2 also by single-strand-specific enzymes. Only C89 of loop 2 exhibited any consistent cleavage with CV1 and this may indicate some base-stacking. Although loop 1 was not
reactive to single-strand-specific enzymatic probes, this probably reflects a reduced accessibility of enzymes to this short loop.

The most unexpected features of this stimulatory RNA were the atypical cleavage pattern of stem 1 and the paucity of cleavages in the slippery sequence/spacer region. The assignment of cleavages in these areas was complicated by a strong compression effect in the gels (see Figure 6A, between bases 40 and 60), presumably arising from stable base-pairing, which was not completely relieved (although much improved) by running reactions on highly denaturing gels containing 20% formamide (Figure 6B). Nevertheless, we were able to assign almost all of the reactivities in this region, and it was noticeable that stem 1, especially stem 1a, showed susceptibility to cleavage by single-strand-specific enzymatic probes, notably RNase T1 at residues G80–82. Similarly, RNase CL3, a C-specific reagent also showed cleavage at some of the C residues in stem 1. Although it can be argued that the −ACA− bulge in stem 1 would introduce some deformation of the (presumably) A-form helix and increase accessibility of adjacent bases to enzymatic probes, we did not see an increased access of chemical probes, arguing against a severe bulge-induced destabilization of stem 1a. It may be that stem 1a exists in equilibrium with another conformation and one such possibility is shown in Figure 7B. In this alternative model, the first arm of stem 1a is proposed to pair with an upstream region (G35–G40), displacing the second arm of stem 1a and the bulge triplet into loop 2. Some features of this alternative model are attractive, e.g. the cleavage of G80–82 by RNase T1 is rationalized as these residues would be located in loop 2. Similarly, the limited reactivity of the slippery sequence and spacer region to single-strand cleavage reagents could be attributed to the fact that much of the region is folded into a hairpin. However, the alternative model also has flaws, most noticeably that the pattern of reactivities of the bases in the second arm of stem 1a is not fully consistent with their location in loop 2. These bases should show increased sensitivity to lead acetate and imidazole and reduced CV1 cleavage, neither of which was observed. In addition, one might expect that the sensitivity of G80–82 to RNase T1 would be similar to that of the other G’s in loop 2 (G86 and G90), which also was not the case.
Based on the mutational analysis (Figure 5) and the chemical probing data, we favour the idea that stem 1a folds according to the original prediction, but shows breathing or flips occasionally into an alternative conformation (perhaps the one described above). This behaviour might account for the fact that destabilization of the stem did not reduce frameshifting as dramatically as those mutations that destabilized stem 1b or stem 2 (Figure 5). Nevertheless, it remains difficult to explain the strong reactivity of the G’s (to RNase T1) in the second arm of stem 1a. It may be that unusual structural features are present that promote RNase T1 cleavage. This idea is not without precedent. Structure probing of the coronavirus infectious bronchitis virus (IBV) frameshift-stimulating pseudoknot (S. Pennell and I. Brierley, unpublished data), functional derivatives of this pseudoknot (40,41) and a functional derivative of the RSV pseudoknot (39) has also revealed unexpected RNase T1 reactivity towards apparently base-paired G residues, although in these pseudoknots, at the top of stem 2.

**DISCUSSION**

In this study, the secondary structure of the Edr frameshift signal was investigated using site-directed mutagenesis and RNA structure probing. The results obtained clearly demonstrate that the stimulatory RNA downstream of the slippery sequence folds into an RNA pseudoknot which, although different from an earlier study (20), undoubtedly resembles the frameshift-promoting pseudoknots of virus signals.

The Edr frameshift site comprises a slippery sequence GGGAAC, a 5 nt spacer region and a relatively large pseudoknot (in viral terms) with 10 bp in stem 1, 9 in stem 2 and loops of 3 (loop1) and 9 (loop 2) nucleotides. In comparison with viral frameshift signals, Edr does not fit into a specific category, but has hallmarks of different viruses. The possession of relatively long pseudoknot stems, especially stem 2, is a feature of the frameshift signals present at the pro/pol overlap of the retroviruses human T-cell lymphotrophic virus (HTLV) types 1 and 2, simian T-cell lymphotrophic virus type 1 and bovine leukaemia virus, and also of the coronaviruses, although in these viruses, the pseudoknots are accompanied by the slippery sequence UUUAAAC rather than GGGAAC (26,42–44). In terms of loop lengths, Edr resembles most closely those pseudoknots found at the gag/pro overlap of the retroviruses Maedi–Visna virus, Mason–Pfizer monkey virus, feline immunodeficiency virus and simian retrovirus (SRV) types 1 and 2, which possess relatively short loops and also employ the slippery sequence GGGAAC (42). The loops of the Edr pseudoknot are short but sufficiently long to span the stems (7,37). That we have identified the entire frameshift region seems likely from the results of our mutational analysis and from phylogenetic sequence comparisons. In Figure 7A, nucleotides within and surrounding the pseudoknot that differ in the human orthologue, PEG10, are shown. As can be seen, almost all of the sequence variation occurs before the slippery sequence and immediately following stem 2. Four nucleotide differences are present within the
frameshift region, but of these, only one is likely to have an effect on pseudoknot function, namely the G to A transition at the top of stem 2. Of the other three changes, one is within the stem 1 bulge and the other two in loop 2. As we have shown experimentally that the bulge in stem 1 is not required for frameshifting and that most of the sequence of loop 2 can be changed without consequence, it seems that these changes are in ‘neutral’ regions. The stem 2 transition itself would only affect the ultimate stem 2 base pair and would likely have only a modest effect on frameshift efficiency. From the perspective of phylogenetics, it will be interesting to see whether Edr orthologues are present in other mammalian species. This will be informative in assessing the conservation of the pseudoknot and the stage in mammalian evolution at which the Edr gene was acquired.

An unexpected feature of the Edr pseudoknot was the marked accessibility of single-strand-specific enzymatic reagents to stem 1a in the structure probing experiments. In the discussion of these data in the Results section above, this accessibility was hypothesized to be a consequence of a relative instability of stem 1a or the presence of a specific stem 1 conformation highly reactive to single-stranded enzymatic probes, such as RNase T1 and CL3. In related investigations of the IBV (40,45), SRV-1 (31) and RSV pseudoknots (39), stem 2, rather than stem 1, proved to be more sensitive to single-stranded probes and possessed unusual susceptibility to RNase T1 (CL3 was not used in these studies). Another difference was highlighted in our mutational analysis of the Edr pseudoknot, where a complementary change that destabilized the first arm of stem 1a still retained about two-fifths of the wild-type frameshifting efficiency (Figure 5, m3 12.2%). In most models of ribosomal frameshifting, a ribosomal pause is proposed to occur upon encounter of the pseudoknot, perhaps because of a failure to unwind the pseudoknot efficiently, and this pause occurs while the decoding centre is over the slippery sequence. In constructs with a stem 1 destabilization, one would expect frameshifting to be greatly reduced, since the ribosome would translate further into the structure (since stem 1 affords less resistance) and fail to pause, or pause since the ribosome would translate further into the structure, one would expect frameshifting to be greatly reduced, and this pause occurs while the decoding centre is over perhaps because of a failure to unwind the pseudoknot efficiently.

In the top of stem 2. Of the other three changes, one is within the stem 1 bulge and the other two in loop 2. As we have shown experimentally that the bulge in stem 1 is not required for frameshifting and that most of the sequence of loop 2 can be changed without consequence, it seems that these changes are in ‘neutral’ regions. The stem 2 transition itself would only affect the ultimate stem 2 base pair and would likely have only a modest effect on frameshift efficiency. From the perspective of phylogenetics, it will be interesting to see whether Edr orthologues are present in other mammalian species. This will be informative in assessing the conservation of the pseudoknot and the stage in mammalian evolution at which the Edr gene was acquired.

Several virus pathogens utilize —1 ribosomal frameshifting, including the retrovirus HIV-1 (46), the coronavirus responsible for severe acute respiratory syndrome [SARS-CoV (47)] and numerous plant pathogens [for example, see (48,49,50)]. Frameshifting signals have thus been considered as targets for antiviral intervention (51,52). The existence of related signals in eukaryotic cellular genes, however, would potentially complicate the design of, or even prevent the use of such therapies, particularly if the cellular frameshifting signal(s) was structurally similar to the virus examples. The report of frameshifting in the mouse Edr gene provided the first opportunity to characterize such a cellular signal. We have demonstrated here that the Edr signal resembles viral examples closely, with a characteristic slippery sequence—spacer-pseudoknot organization. Thus, the similarity of parts of the Edr coding sequence to the gag/pro region of retroviral genomes extends to the maintenance of a retrovirus-like frameshifting signal. Given this, and the possibility that other retrovirus-like motifs have been subsumed into mammalian genes and retained a role for frameshifting, antiviral agents that target this process may have previously unanticipated consequences on cellular metabolism in uninfected cells.

ACKNOWLEDGEMENTS

Plasmid p2luc/MMTV gag/pro was the kind gift from John Atkins and Ray Gesteland. This work was supported by the Medical Research Council, UK [project grant number G9901292 (to I.B.), PhD studentship number G78/8212 (to E.M.)]. Funding to pay the Open Access publication charges for this article was provided by the Medical Research Council, UK.

Conflict of interest statement. None declared.

REFERENCES

1. Jacks,T. and Varmus,H.E. (1985) Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. Science, 230, 1237–1242.
2. Rice,N.R., Stephens,R.M., Burny,A. and Gilden,R.V. (1985) The gag and pol genes of bovine leukemia virus: nucleotide sequence and analysis. Virology, 142, 357–377.
3. Chandler,M. and Fayet,O. (1993) Translational frameshifting in decoding the SARS-CoV genome. Nucleic Acids Res., 1717–1731.
4. Farabaugh,P.J. (1996) Programmed translational frameshifting. Microb. Rev., 60, 103–134.
5. Baranov,P.V., Gesteland,R.F. and Atkins,J.F. (2002) Recoding: translational bifurcations in gene expression. Gene, 286, 187–201.
6. Fleij,C.W.A., Rietveld,K. and Bosch,L. (1985) A new principle of RNA folding based on pseudoknotting. Nucleic Acids Res., 13, 5838–5842.
7. Baranov,P.V., Henderson,C.M., Anderson,C.B., Gesteland,R.F., Atkins,J.F. and Howard,M.T. (2005) Programmed ribosomal frameshifting in decoding the SARS-CoV genome. Virology, 332, 498–510.
8. 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.
13. Gurvich,O.L., Baranov,P.V., Zhou,J., Hammer,A.W., Gesteland,R.F. and Atkin,J.F. (2003) Sequences that direct significant levels of frameshifting are frequent in coding regions of *Escherichia coli*. *EMBO J.*, 22, 5941–5950.

14. Blinkova,A.L. and Walker,J.R. (1990) Programmed ribosomal frameshifting generates the *Escherichia coli* DNA polymerase III γ subunit from within the α subunit reading frame. *Nucleic Acids Res.*, 18, 1725–1729.

15. 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. USA.*, 87, 3713–3717.

16. Jacks,T., Madhani,H.D., Masiarz,F.R. and Varmus,H.E. (1988a) 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 Frameshifting in coronavirus frameshifting signals. *J. Bacteriol.*, 174, 6842–6851.

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. Mol. Biol.*, 271, 47–60.

19. Larsen,B., Gesteland,R.F. and Atkins,J.F. (1997) Structural probing and mutagenic analysis of the stem–loop required for *Escherichia coli* dnaX ribosomal frameshifting: programmed efficiency of 50%. *J. Mol. Biol.*, 271, 47–60.

20. Shigemoto,K., Brennan,J., Walls,E., Watson,C.J., Stott,D., Rigby,P.W. and Reith,A.D. (2001) Identification and characterisation of a developmentally regulated mammalian gene that utilises –1 programmed ribosomal frameshifting. *Nucleic Acids Res.*, 29, 4079–4088.

21. Ono,R., Kobayashi,S., Wagatsuma,H., Aisaka,K., Kohda,T., Kaneko-Ishino,T. and Ishino,F. (2001) A retrotransposon-derived gene, *PEG10*, is involved in ribosomal frameshifting. *EMBO J.*, 20, 5094–5102.

22. Nam,S.H., Copeland,T.D., Hatanaka,M. and Oroszlan,S. (1993) The role of RNA pseudoknot stem 1 length in the promotion of efficient –1 ribosomal frameshifting. *J. Mol. Biol.*, 240, 196–203.

23. Brierley,I. and Pennell,S. (2001) Structure and function of the stimulatory subunit reading frame. *Proc. Natl Acad. Sci. USA.*, 98, 9183–9188.

24. Brierley,I. and Inglis,S.C. (1989) Characterisation of an RNA pseudoknot downstream of the frameshift signal of simian retrovirus-1. *Nucleic Acids Res.*, 17, 2304–2310.

25. Kunkel,T.A. (1985) Rapid and efficient site-specific mutagenesis and analysis of the pseudoknot-containing gag-pol pseudoknot. *Nature Struct. Biol.*, 96, 50–74.

26. Marczinke,B., Fisher,R., Vidakovic,M., Bloys,A.J. and Brierley,J. (1998) Secondary structure and mutational analysis of the ribosomal frameshift signal of *Rous sarcoma virus*. *J. Mol. Biol.*, 284, 205–225.

27. Larsen,B., Gesteland,R.F. and Atkins,J.F. (2003) Sequences that direct significant levels of ribosomal frameshifting. *Proc. Natl Acad. Sci. USA.*, 100, 1109–1123.

28. Jackson,R.J. and Hunt,T. (1983) Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eukaryotic messenger RNA. *Methods Enzymol.*, 96, 6, 51–60.

29. Wyatt,J.R., Puglisi,J.D. and Tinoco,I. (1990) RNA pseudoknots: stability and like structure of TYMV RNA. *Cell*, 447–458.

30. Brault,V. and Miller,W.A. (1992) Translational frameshifting mediated by a viral sequence in plant cells. *Proc. Natl Acad. Sci. USA.*, 89, 2262–2266.

31. Dinman,J.D., Ruiz-Echevarria,M.J. and Peltz,S.W. (1998) Translating old drugs into new treatments: ribosomal frameshifting as a target for antiviral agents. *Trends Biotechnol.*, 16, 190–196.

32. Van Belkum,A., Verlaan,P., Bing Kun,J., Pleij,C. and Bosch,L. (1998) Temperature dependent chemical and enzymatic probing of the RNA-like structure of TMYV RNA. *Nucleic Acids Res.*, 16, 1931–1950.

33. Wyatt,J.R., Puglisi,J.D. and Tino,Col.(1990) RNA pseudoknots: stability and loop size requirements. *J. Mol. Biol.*, 214, 455–470.

34. Brierley,I., Rolley,N.J., Jenner,A.J. and Inglis,S.C. (1991) Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. *J. Mol. Biol.*, 220, 889–902.

35. Napthine,S., Liphardt,J., Bloys,A., Routledge,S. and Brierley,J. (1999) Sequences that direct significant levels of frameshifting. *J. Mol. Biol.*, 288, 321–335.

36. Le,S.Y., Shapiro,B.A., Chen,J.H., Nussinov,R. and Maizel,J.V. (1991) The role of RNA pseudoknot stem 1 length in the promotion of efficient –1 ribosomal frameshifting. *Nucleic Acids Res.*, 19, 2305–2315.

37. Thiel,V., Ivanov,K.A., Putics,A., Hertzig,T., Schelle,B., Bayer,S., Weissbrich,B., Snijder,E.J., Rabenau,H., Doerr,H.W. et al. (2003) Mechanisms and enzymes involved in SARS coronavirus genome expression. *J. Gen. Virol.*, 84, 2505–2515.

38. Nam,S.H., Copeland,T.D., Hatanaka,M. and Oroszlan,S. (1993) Characterization of ribosomal frameshifting for expression of pol gene products of human T-cell leukemia virus type I. *J. Virol.*, 67, 196–203.

39. Nam,S.H., Copeland,T.D., Hatanaka,M. and Oroszlan,S. (1993) Characterization of ribosomal frameshifting for expression of *pol* gene products of human T-cell leukemia virus type I. *J. Virol.*, 67, 196–203.

40. Brierley,J., Rolley,N.J., Jenner,A.J. and Inglis,S.C. (1991) Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. *J. Mol. Biol.*, 220, 889–902.

41. McCarthy,R., Chatterjee,G., Weissman,J.S., Raab,J. and Peltz,S.W. (1997) Peptidyl-transferase inhibitors have antiviral properties by altering programmed –1 ribosomal frameshifting efficiencies: development of model systems. *Proc. Natl Acad. Sci. USA.*, 94, 6606–6611.

42. Brierley,J. and Peltz,S.W. (1999) Translating old drugs into new treatments: ribosomal frameshifting as a target for anti-viral agents. *Trends Biotechnol.*, 16, 190–196.

43. Verlaan,P., Verlaan,P.W., Pleij,C.W., Hilbers,C.W. and Heus,H.A. (2001) Solution structure of the pseudoknot of *SRV-1* RNA involved in ribosomal frameshifting. *J. Mol. Biol.*, 310, 1109–1123.

44. Jackson,R.J. and Hunt,T. (1983) Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eukaryotic messenger RNA. *Methods Enzymol.*, 96, 50–74.

45. Marczinke,B., Fisher,R., Vidakovic,M., Bloys,A.J. and Brierley,J. (1998) Secondary structure and mutational analysis of the ribosomal frameshift signal of *Rous sarcoma virus*. *J. Mol. Biol.*, 284, 205–225.

46. Napthine,S., Liphardt,J., Bloys,A., Routledge,S. and Brierley,J. (1999) The role of RNA pseudoknot stem 1 length in the promotion of efficient –1 ribosomal frameshifting. *J. Mol. Biol.*, 288, 305–320.

47. Brierley,J., Napthine,S., Kontos,H. and Brierley,J. (1999) Evidence for an RNA pseudoknot loop–helix interaction essential for efficient –1 ribosomal frameshifting. *J. Mol. Biol.*, 288, 321–335.

48. Ten Dam,E.B., Pleij,C.W.A. and Bosch,L. (1990) RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. *Virus Genes*, 4, 121–136.

49. Le,S.Y., Shapiro,B.A., Chen,J.H., Nussinov,R. and Maizel,J.V. (1991) RNA pseudoknots downstream of the frameshift sites of retroviruses. *Genet. Anal. Tech. Appl.*, 8, 191–205.

50. Nam,S.H., Copeland,T.D., Hatanaka,M. and Oroszlan,S. (1993) Characterization of ribosomal frameshifting for expression of *pol* gene products of human T-cell leukemia virus type I. *J. Virol.*, 67, 196–203.

51. Brault,V. and Miller,W.A. (1992) Translational frameshifting mediated by a viral sequence in plant cells. *Proc. Natl Acad. Sci. USA.*, 89, 2262–2266.

52. Dinman,J.D., Ruiz-Echevarria,M.J. and Peltz,S.W. (1997) Peptidyl-transferase inhibitors have antiviral properties by altering programmed –1 ribosomal frameshifting efficiencies: development of model systems. *Proc. Natl Acad. Sci. USA.*, 94, 6606–6611.