CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

| | |
|---|---|
| Title | Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans: close to 300 cases reveal remarkable diversity despite underlying conservation |
| Authors | Ivanov, Ivaylo P.;Atkins, John F. |
| Publication date | 2007 |
| Original Citation | Ivanov, I. P. and Atkins, J. F. (2007) 'Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans: close to 300 cases reveal remarkable diversity despite underlying conservation', Nucleic Acids Research, 35(6), pp. 1842-1858. doi: 10.1093/nar/gkm035 |
| Type of publication | Article (peer-reviewed) |
| Link to publisher's version | https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkm035 - 10.1093/nar/gkm035 |
| Rights | © 2007, the Authors. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/2.0/uk/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited. - https://creativecommons.org/licenses/by-nc/2.0/uk/ |
| Download date | 2026-11-06 19:53:54 |
| Item downloaded from | https://hdl.handle.net/10468/5033 |

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

*1842–1858* *Nucleic Acids Research, 2007, Vol. 35, No. 6* *Published online 1 March 2007*
*doi:10.1093/nar/gkm035*

## SURVEY AND SUMMARY

# Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans: close to 300 cases reveal remarkable diversity despite underlying conservation

**Ivaylo P. Ivanov[1,*] and John F. Atkins[1,2,*]**

[1]Biosciences Institute, University College Cork, Cork, Ireland and [2]Department of Human Genetics, University of Utah, Salt Lake City, UT 84112-5330, USA

Received November 16, 2006; Revised January 5, 2007; Accepted January 7, 2007

## ABSTRACT

**The protein antizyme is a negative regulator of intracellular polyamine levels. Ribosomes synthesizing antizyme start in one ORF and at the codon 5′ adjacent to its stop codon, shift +1 to a second and partially overlapping ORF which encodes most of the protein. The ribosomal frameshifting is a sensor and effector of an autoregulatory circuit which is conserved in animals, fungi and protists. Stimulatory signals encoded 5′ and 3′ of the shift site act to program the frameshifting. Despite overall conservation, many individual branches have evolved specific features surrounding the frameshift site. Among these are RNA pseudoknots, RNA stem-loops, conserved primary RNA sequences, nascent peptide sequences and branch-specific 'shifty' codons.**

Standard linear non-overlapping triplet decoding of certain mRNAs in diverse organisms is dynamically diverted into an alternative reading frame at specific sites. Where utilized for regulatory purposes or to produce an additional protein, the ribosomal frameshifting involved is often 'programmed' to occur at high efficiency by signals embedded in the same mRNA. These signals, often called recoding or stimulatory signals, can be either 5′ or 3′ of the frameshift site or both. The 3′ signals typically, but not always, involve mRNA structure, often in the form of pseudoknots. This combination of specific shift site and discrete enhancing signal constitute a subset of recoding, the 'codes within the code' that increase decoding versatility (1,2).

The most widespread known cases of shifting to the −1 reading frame occur in the decoding of many animal, plant and bacterial viruses and a number of mobile elements. A modest number of non-mobile chromosomal genes are also known to utilize −1 programmed frameshifting and, though the number will doubtless rise, the cases identified so far have a limited distribution. For instance *dnaX* ribosomal frameshifting in *E. coli* to yield the gamma subunit of DNA polymerase III is only known in a limited number of other bacteria (3). Also the known neofunctionalized mammalian genes whose expression involve, or likely involve, frameshifting (4–6), do not occur in invertebrates.

## INVOLVEMENT OF +1 FRAMESHIFTING

The most widespread of the known non-mobile cellular genes to utilize programmed shifts to the +1 reading frame are those encoding bacterial polypeptide chain release factor 2 (RF2) and antizyme. Programmed +1 frameshifting is more commonly used for regulatory purposes than is −1 frameshifting and its function in the decoding of RF2 and antizyme mRNAs is for this purpose. Following its discovery in *E. coli* (7), decoding of the RF2 gene in 80% of eubacterial species is now known to involve regulatory frameshifting (8–10), which is also stimulated by polyamines (9), but frameshifting is not known in the decoding of release factor genes elsewhere. The utilization of programmed +1 frameshifting in decoding of antizyme mRNA extends from mammals (11) to *Drosophila melanogaster* (12) to *Caenorhabditis elegans* and yeasts (13–17). Twelve nucleotides around the frameshift site are identical between the fission yeast *Schizosaccharomyces pombe* and human antizyme 1 mRNAs (13). Given the evidence, see below, that this frameshifting is conserved from a common ancestor and the time elapsed, greater than 800 million years ago (18,19), its study provides a unique glimpse into

*Correspondence may be addressed to either author at Tel: +1-353 21 490 1313; Fax: +1-353 23 55147; Emails: iivanov@genetics.utah.edu and j.atkins@ucc.ie

© 2007 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

the possibilities explored by selective forces over an extensive period of time.

## DISCOVERY OF ANTIZYME 1 AND ITS PROPERTIES

Given the ubiquitous occurrence of polyamines and their multifaceted roles (20–23), it is not surprising that ornithine decarboxylase, which catalyzes the first step in polyamine synthesis, is tightly regulated. The existence of a negative regulator of ornithine decarboxylase, termed ornithine decarboxylase antizyme, or now simply, antizyme, was first postulated as far back as 1976 (24). However, it was only when an antizyme gene from rat was cloned and its expression studied (11,25,26) that strong skepticism about the reality of its existence was put aside (27). Antizyme binds to ornithine decarboxylase and targets it for degradation by the 26S proteasome without ubiquitination (28,29). By also affecting transport of extracellular polyamines, antizyme negatively affects the uptake of polyamines (30–32). In contrast to its effect on ornithine decarboxylase, antizyme binding inhibits the degradation (33) of a protein known as 'antizyme inhibitor' because of its ability to reactivate ornithine decarboxylase inhibited by antizyme (34,35). There are contrasting reports on antizyme-independent effects of this 'inhibitor' on cell proliferation (36,37). Recently a distinct protein, located mainly in the mitochondrial membrane fraction of brains and testes, has also been shown to have specific antizyme inhibitory properties (38).

Various types of antizyme will be considered later, but first the existence of two forms of mammalian antizyme 1 will be introduced. These result from synthesis initiating at two alternatively used start codons. Initiation at the first, but not the second, start codon yields a product which is imported into mitochondria (39,40). Over-expression of antizyme 1 increases apoptosis (41), at least of haematopoietic cells, through mitochondrial membrane depolarization and activation of a caspase cascade (42). In addition to growth or cell division inhibitory effects of over-expression of antizyme 1 (43–46), it also reduces tumorigenesis (47,48) and antizyme 1 can be considered as a tumor suppressor (49). [Over-expression of yeast antizyme inhibits cell division, but its gene knock-outs in both *S. pombe* and *Saccharomyces cerevisiae* are viable (13,16).]

An additional feature of mammalian antizyme 1 relevant to its complex effects, is its involvement in nucleocytoplasmic shuttling (50–52). *Drosophila* antizyme regulates nuclear entry, the overall levels of the sex determination master switch, sex lethal and cyclin B (53).

Recent evidence suggests that antizyme may also be involved in linkages with other cellular pathways. One example being that mammalian antizyme 1 is a component that binds to cyclin D1, at least in prostate cells, and targets it for degradation (54).

In summary, where antizyme is present, as it is from yeast to mammals, its expression is, and has to be, carefully regulated.

## REGULATORY ROLE AND CLASSIC FRAMESHIFT SIGNALS

The programmed frameshifting required for antizyme synthesis is the key regulatory point. Nearly all antizymes are encoded by two partly overlapping ORFs. Typically, ORF1 of mammalian antizyme 1 mRNA can encode 68 amino acids and is terminated by a UGA stop codon (UCC UGA). ORF2, which encodes the main part of antizyme, is in the +1 frame with respect to ORF1 and starts before the end of ORF1. It is only accessed by ribosomes that first translate ORF1. Amino acid sequencing of the *trans* frame product showed that the reading frameshift occurs when a proportion of the ribosomes decoding the last codon of ORF1, *UCC*, shift to the +1 frame so that the next codon translated is GAU (*UCC* U GAU) (11). It has been argued that this occurs by detachment of peptidyl $tRNA^{Ser}$ from UCC and its re-pairing to CCU (55,56). When the level of free intracellular polyamines (or agmatine, 57), is elevated, a high proportion of ribosomes perform the specific frameshift and synthesize functional antizyme. The converse also pertains so that when the level of polyamines is low, little of the negative regulator, antizyme, is synthesized. This leads to increased synthesis of polyamines and the possibility of increased uptake of extracellular polyamines. The frameshifting involved is the sensor and effector of an autoregulatory circuit (see reviews, 58,59). The contrast between the high levels of antizyme 1 mRNA and the generally low levels of its protein product is most easily explicable as permitting a rapid response to regulatory signals.

The role of a sequence 5′ of the shift site (11,14,60) for polyamine sensing has been investigated (61,62), and further work is ongoing. Following some precedents elsewhere (63,64), having a stop codon 3′ adjacent to the shift site is important for the level of frameshifting (11,62,65,66). Interferons, which have antiviral and antiproliferative effects, increase antizyme frameshifting by reducing the efficiency of ORF1 termination. This is mediated by interaction of interferon-induced RNase L with termination factor eRF3 (67).

At 3′ of the mammalian antizyme 1 mRNA shift site, there is a pseudoknot which acts to stimulate the level of antizyme frameshifting (11,68). Notably, this pseudoknot is closer to the shift site than are counterpart pseudoknots that act as stimulators for programmed −1 frameshifting, e.g. in mouse mammary tumor virus.

## ANTIZYME TYPES

Mammals and most other vertebrates have a second antizyme, antizyme 2 (60,69,70). Like antizyme 1, it is widely distributed in different tissues, but is expressed at lower levels. However, it is under tighter evolutionary pressure than antizyme 1, at least in mammals, as evidenced by the 99.5% amino acid identity between the human and mouse antizyme 2 orthologs versus 84% for the antizyme 1 orthologous pair. Before its identity as an antizyme was known, its expression was found to be

greatly elevated in the brains of mice treated with a drug that induces epileptic seizures (71).

Vertebrate antizymes 1 and 2 are expressed in all cell types tested apart from post-meiotic male germ cells. Two paralogs of antizymes in zebrafish (*Danio rerio*), AZS and AZL, are also ubiquitously expressed although each shows a distinct expression profile in developing embryos (72). The antizyme homolog in *Drosophila* is also ubiquitously expressed (73), and probably in other animals also.

In contrast, the third mammalian antizyme, antizyme 3 is only expressed in post-meiotic male germ line cells (74–76). Though frameshifting is required for its expression, its mRNA does not have readily apparent stimulatory signals flanking its shift site. When constructs with its shift site and flanking sequences are expressed in HEK293 and GC-2 cell lines (61) only a low level of frameshifting is detected. It is unknown if the level is also low on full-length mRNA in the germ cells in which it is naturally expressed, or if there are tissue-specific effects on the frameshifting.

Recently another antizyme paralog with preferential tissue expression was identified in *D. rerio* and in at least one other fish species. The gene, called AntiZyme in Retina (AZR), as the name implies, is expressed in the retinal ganglion layer but also in brain and to a lesser extent in segmented muscle cells (70).

## UNUSUAL ESTs, THE PHANTOM ANTIZYME 4 AND MIS-ANNOTATION

A number of entries in the EST database that show significant similarity to known members of the antizyme gene family, were identified that nevertheless do not fit generally accepted phylogenetic relationships. It is, of course, possible that they represent cases of horizontal gene transfer. However it is far more likely that they are the result of library contaminations. Some of the putative contaminants are spurious. A typical example is a sand fly EST that corresponds to rat antizyme 1 (accession # AM092700). We have excluded such examples from our analysis. Others are useful and sometimes even fascinating. For example, the sequences of several EST-s coming from *Pinus taeda* (pine tree) library converge on two unrelated fungal antizymes. Although we cannot deduce to which species they belong, we have included them in this article as they are informative. Another somewhat curious example comes from a *D. rerio* (zebrafish) library. This library is derived from stomach and intestines along with two other tissues. The antizyme sequence in question clusters with those from fleas and other similar insects. It is very likely that this sequence comes from the brine shrimp, *Artemia franciscana*, which is used as food for zebrafish in research labs. One interesting group concerns fungal antizyme sequences derived from three separate plant libraries—*Medicago truncatula*, *Triticum aestivum* and *Hordeum vulgare*. Unlike the examples above, these sequences do not seem to result from library contamination but from symbiotic fungal species living inside the roots of these plants. In the case of *M. truncatula* this species is very likely *Glomus intraradices*. In the others we have not been able to infer species identification but nevertheless we have proceeded to include the sequence data in the current study.

Previously we identified a novel antizyme sequence from a single clone derived from a human brain library. We 'only tentatively' considered this sequence as a candidate for belonging to an unknown gene, antizyme 4 'pending confirmation of its origin' (14). This sequence has no exact match in the completed human genome and no further EST-s corresponding to it has appeared in the last six years. As pointed out previously we have been unable to amplify it from human genomic DNA (14). Based on these and other considerations, we now believe that this cDNA is a contaminant, most likely a mammalian antizyme 1 gene belonging to an unidentified rabbit or hare species.

In our search for new antizyme sequences, we encountered several examples of another problem which merits attention. In several not readily identifiable antizyme genes, mostly in recently sequenced fungal genomes, the annotated sequence deposited in GenBank has either the frameshift site omitted or more disturbingly apparently 'edited out' during annotation. The most egregious example was encountered with the *Cryptococcus neoformans* antizyme gene. That gene has two introns and the splice donor site of the second is annotated in such a way that the frameshift site is removed so that ORF1 and ORF2 now form a single ORF. This is accomplished by 'deleting' 31 nucleotides from the end of exon 2. It would, perhaps, be an unremarkable error were it not directly contradicted by >10 ESTs showing the correct splice site and not supported by even one EST with the incorrectly annotated splice site. In a different but also recent case, after we started working on an antizyme-unrelated mammalian sequence that had all the hallmarks of its expression involving programmed frameshifting, a revised version of the sequence appeared in the database with the framing adjusted and with the shift-prone site deleted. Subsequent resequencing showed no evidence for the deletion (6). We need to be mindful of an earlier concern that frameshifting candidates may be purged from sequences with frame 'problems' (1). The development of programs that automatically annotates further candidates of known types of frameshifting has recently commenced (10) and though it addresses only part of the issue, merits substantial development.

## PHYLOGENY OF ANTIZYME GENES

The antizyme genes identified so far belong to species from three of the four eukaryotic kingdoms—Animalia, Fungi and Protista. No antizyme gene has been detected in plants. Recent studies of the anaerobic gut bacteria, *Selenomonas ruminantium*, convincingly revealed an antizyme gene (77), whose product is ribosomal protein L10. This protein shares several interchangeable properties with mouse antizyme 1, though *E. coli* ribosomal protein L10 cannot substitute for these activities. Its synthesis does not involve ribosomal frameshifting (77). Alignment with mammalian antizyme 1 shows no significant amino

acid similarity between the two, suggesting independent origins.

The clear authenticity of this *S. ruminantium* antizyme is in contrast to an earlier report of three *E. coli* protein having antizyme-like properties (78) where two subsequent reports have pointed out the differences from what is normally considered to be 'antizyme' (79,80).

Within the animal kingdom antizyme genes have been identified in at least nine different phyla—including: Mollusca, Cnidaria, Nematoda, Annelida, Platyhelminthes, Arthropoda, Priapulida, Echinodermata and Chordata.

In fungi, antizyme genes are found in at least four different phyla, Ascomycota, Basidiomycota, Glomeromycota and Zygomycota, although the vast majority of examples are in fact from just one, Ascomycota.

There are more than half-a-dozen examples of antizyme genes from protists. Since there is considerable uncertainty about the exact taxonomic classification within that kingdom we have not attempted to place the protist antizyme genes in specific phyla. All protist antizyme genes identified so far are, however, from protozoa.

This distribution seems to suggest that eukaryotic antizyme was present in the last common ancestor of all fungi, protists and animals. According to the current understanding of the branches of the tree of life, plants diverged from animals and fungi after the branching off of ciliates/protists. On this basis, it can be surmised that the ancestor of the plant lineage once had an antizyme gene. Our inability to identify antizyme genes in extant plants can be explained in one of two ways. Either plants have lost the gene or the protein sequence has evolved too much to be detected by sequence comparisons to antizymes in protists, fungi and animals. Currently there is some uncertainty whether antizyme activity is present in plants at all (*Arabidopsis* seems to lack ornithine decarboxylase (80a)).

Overall we have identified over 100 vertebrate, 10 protist, about 100 invertebrate and 70 fungal antizyme genes; a total of almost 300 genes. Some of these sequences are partial but about two-thirds include the entire coding region.

Most invertebrates and fungi appear to have a single antizyme gene. There are only a few exceptions. For example in the nematode, *Xiphinema index*; the humus earthworm, *Lumbricus rubellus*; the tarantula, *Acanthoscurria gomesiana* and likely the fungus, *Rhizopus oryzae*, two paralogs of antizyme are present. Based on phylogenetic analysis (data not shown), it appears that in those cases, the paralogous pair has diverged relatively recently—from a few million to tens of millions of years ago. This indicates they represent localized events which did not affect most species of the same phylum. By contrast almost all vertebrates have multiple antizyme paralogs. As introduced above, mammals have three. Reptiles, birds and amphibians have at least two. Zebrafish (*D. rerio*) and minnows (*Pimephales promelas*) have four. Puffer fish (*Takifugu rubripes* or *Tetraodon nigroviridis*) apparently have three (70). Rainbow trout (*Oncorhynchus mykiss*) and salmon (*Salmo salar*) have at least four although belonging to different paralogous groups than the four antizymes in zebrafish and minnows (I.P.I., unpublished data and Supplementary Data, Figure S1).

All the antizyme genes identified so far in animals and fungi require a +1 ribosomal frameshift to express a functional product. This is also true in most protist antizyme mRNAs identified to date but there is at least one and possibly two, exceptions. The ciliate *Tetrahymena thermophila* antizyme mRNA has no apparent ORF in a different frame (ORF1 equivalent) upstream of the main ORF (ORF2 in other organisms). Therefore with that antizyme mRNA, no ribosomal frameshifting appears to be involved. The situation is somewhat similar in the protist *Capsaspora owczarzaki*. Although a very short (6 codons) upstream ORF is present, a simple +1 frameshifting event cannot account for the expression of the downstream ORF. Perhaps a different mechanism of expression is involved.

The sizes of ORF1 and ORF2 are typically 60 and 160 codons, respectively but ORF1 can be as short as 17 (*Hydra*) or as long as 86 (*Cryptococcus*) codons while ORF2 can be as short as 114 (*Haemonchus*) and possibly as long as 443 codons (*Ustilago*).

## FRAMESHIFT SITE CONSERVATION/DIVERGENCE

Most metazoan as well as many yeast/fungal species with an identifiable antizyme gene, have the same frameshift site sequence as that found in the first cloned eukaryotic antizyme gene—i.e. the sequence UCC-UGA of rat antizyme 1 mRNA. This widespread distribution of the UCC-UGA frameshift site provides strong, although inconclusive, evidence that it is ancestral to all others. In addition to UCC-UGA, a number of other frameshift sites are present in the antizyme genes of various phyla (Table 1). For example several fungi have the apparent frameshift site CCC-UGA. Yeast related to *S. cerevisiae*, have either the sequence GCG-UGA (also GCG-UAG) or CCG-UGA at their antizyme frameshift sites. Antizymes from mushroom-related fungi, four distantly related protozoans, at least one species of flat worms (Platyhelminthes), *Macrostomum lignano*, as well as nematodes related to *C. elegans*, have a UUU-UGA frameshift site. Another sequence, AUU-UGA, is present in several planaria species and in several distantly related trematodes (e.g. *Schistosoma mansoni, Echinococcus granulosus*, *Clonorchis sinensis* and *Paragonimus westermani*). The Basidiomycota pathogenic fungus *Cryptococcus neoformans*, the unrelated ascomycotal fungus *Aureobasidium pullulans*, two representatives of *Cryptomonas* algae and at least two shore crabs have the frameshift site GUU-UGA.

Even a casual analysis of the list presented above (also see Figure S1A in Supplementary Data) suggests that many of the alternative antizyme frameshift sites have emerged more than once independently of each other. With the GUU-UGA frameshift site in *C. neoformans*, algae and crabs, or the UUU-UGA frameshift site in mushrooms, some nematodes, flat worms, algae and two other protists, little additional evidence is required for reaching this conclusion as they all belong to different

**Table 1.** The taxonomic distribution of P-site codons of antizyme mRNA frameshifting sites

| Kingdom | Phylum | UCC | UUU | GUU | AUU | CCC | GCG | CCG |
|---|---|---|---|---|---|---|---|---|
| Protista | All | − | + | + | − | − | − | − |
| Fungus | Ascomycota | + | + | − | − | + | + | + |
| | Basidiomycota | + | + | + | − | − | − | − |
| | Glomeromycota | + | − | − | − | − | − | − |
| Animalia | Nematoda | + | + | − | − | − | − | − |
| | Mollusca | + | − | − | − | − | − | − |
| | Annelida | + | − | − | − | − | − | − |
| | Cnidaria | + | − | − | − | − | − | − |
| | Priapulida | − | + | − | − | − | − | − |
| | Echinodermata | + | − | − | − | − | − | − |
| | Arthropoda | + | − | + | − | − | − | − |
| | Platyhelminthes | − | + | − | + | − | − | − |
| | Chordata | + | − | − | − | − | − | − |

The presence of a particular codon in a specific taxon is indicated by '+', the absence by '−'.

phyla and even different kingdoms. However, in the case of the multiple fungal antizyme genes with a CCC-UGA shift site, such a conclusion is much less obvious as they belong to the same phylum. In order to analyze this latter case, we assembled a bootstrapped phylogenetic sub-tree of fungal antizyme mRNAs having an apparent CCC-UGA frameshift site (excluding the antizyme mRNA from the completely unrelated pathogenic fungus *Pneumocystis carinii* which also has a CCC-UGA frameshift site) and a dozen antizymes from related species (Figure 1). The resulting tree and the bootstrap values in the relevant branches, strongly support the hypothesis that in this group of antizymes the frameshift site CCC-UGA emerged independently in two separate clades. For example, the root of the CCC-UGA frameshift site containing the clade that includes *P. brasiliensis*, *A. capsulatus*, *Thermomyces lanuginosus*, *Aspergillus terreus*, *A. flavus*, *A. nidulans*, *A. clavatus*, *A. fumigatus*, and *Neosartorya fischeri* has a bootstrap value of 99.8%. The bootstrap value of the branch supporting the "*H. vulgare*", *Botryotinia fuckeliana* and *Sclerotinia sclerotiorum* CCC-UGA frameshift site clade is a less impressive but still highly significant 95%. Since the root of this sub-tree converges on a UCC-UGA frameshift site, it appears likely that the CCC-UGA frameshift site of *P. carinii* evolved independently of the others, constituting a third CCC-UGA branch in ascomycotal fungi. The data shown in Figure 1 also strongly supports the hypothesis that the GUU-UGA frameshift site in *A. pullulans* evolved in a separate clade—i.e. it evolved independently of the frameshift site in the fungus *C. neoformans*.

The reason why the different antizyme genes have evolved different frameshifting sites is unknown but two general explanations are possible. Some frameshift sites are perhaps interchangeable as they support efficient frameshifting in most or all, eukaryotic cells. This is almost certainly so for those that recur again and again—for example UCC-UGA and UUU-UGA although CCC-UGA and GUU-UGA might also fall in this category. This conclusion is supported by experiments showing that efficient +1 frameshifting occurs on the site CCC-UGA and somewhat less efficiently on UUU-UGA when both

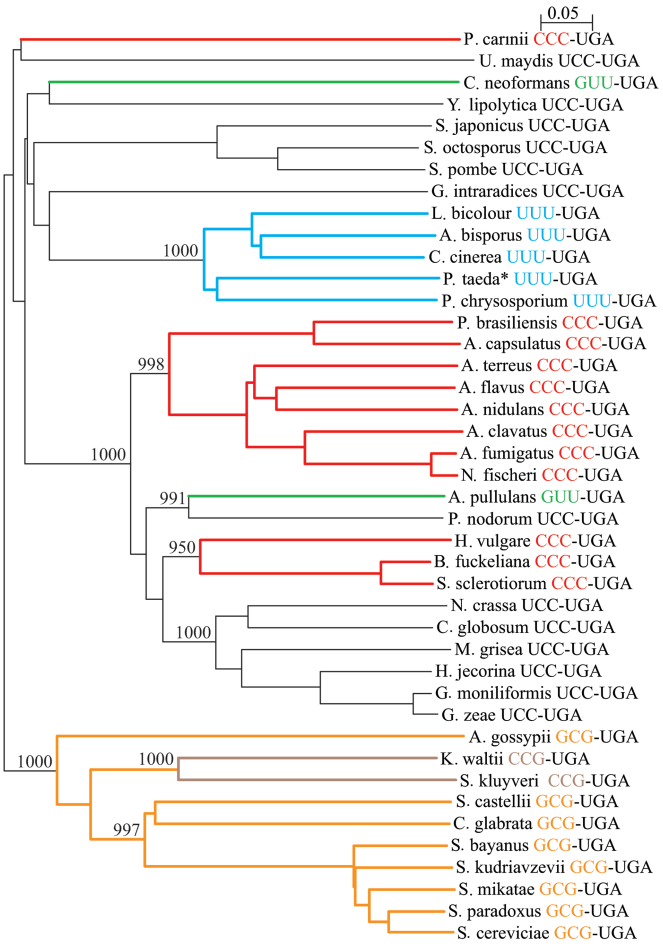



**Figure 1.** The phylogenetic relationship of a select number of fungal antizyme frameshift sites. The unrooted tree is based on the amino acid sequence (encoded by both ORF1 and ORF2) and was drawn using the ClustalX program neighbor-joining algorithm. Different colored lines indicate different frameshift sites. Bootstrap values are given for key nodes. *Contaminant from plant library.

are inserted behind a mammalian antizyme one 5′ element but in the absence of a 3′ (pseudoknot) stimulator and tested *in vivo* in transfected COS-7 cells (62). Interestingly, in the same set of experiments the sequence CUU-UGA

which does not occur naturally in any antizyme gene identified so far but is the natural +1 frameshift site in prokaryotic polypeptide chain RF2 genes, supports the same level of frameshifting as the mammalian wild-type sequence UCC-UGA. While the sequence UUU-UGA can support efficient +1 frameshifting in the context of the RF2 mRNA frameshift site in *E. coli* (63), it is less efficient than the WT CUU_U (stop) perhaps because the CUU decoding $tRNA_2^{Leu}$ does not contain a bulky modification at base 37 which is 3′ adjacent to the anticodon (82). UUU does not occur naturally in the 259 RF 2 genes identified to date that require +1 frameshifting (10). Such inter-kingdom functional equivalence is remarkable.

Other sites appear to have co-evolved with the specializations of the translational machinery in the specific phylogenetic branch. A likely example is the advent of GCG-UGA and CCG-UGA frameshift sites in antizyme mRNAs from *S. cerevisiae* and related yeasts. In these cases the P-site codon (either GCG or CCG) selected correlates well with the low abundance, or complete absence, of particular elongator tRNAs in those species (17). Curiously this does not mean that *S. cerevisiae* ribosomes do not have the ability to frameshift on UCC-UGA as that site supports efficient +1 frameshifting when tested *in vivo* (68). Surprisingly, there is more efficient −2 than +1 frameshifting when a cassette with this sequence from mammalian antizyme 1 is expressed in *S. cerevisiae* (68), though the converse pertains in *S. pombe* (83).

Perhaps one of the most unexpected results from the analysis of the array of known and inferred antizyme frameshift sites is the extent of conservation of a particular ORF1 stop codon. In all except seven cases, it is UGA (i.e. in ~97% of all known examples). Though context dependent (see below), in eukaryotes UGA is the least efficient termination codon. Slow decoding at the A-site is necessary for induction of the alternative translational event. These two facts provide the most likely explanation for the conservation of UGA stop codon at the frameshift site of antizymes. However, both *in vitro* and *in vivo* experiments with mammalian antizyme 1 indicate that the other two stop codons, UAA and UAG, work almost as well as UGA in stimulating +1 frameshifting (11,62). Whether the marginal extra contribution of an UGA, as opposed to UAA of UAG, codon in frameshifting stimulation is sufficient to explain the remarkable conservation of UGA, remains to be determined. With the frameshifting utilized in decoding 80% of bacterial RF2 genes, having the ORF1 stop codon as UGA is key to the regulatory circuit involved (8).

## 3′ CONSERVATION AND FRAMESHIFT STIMULATORS

First, some remarks on antizyme amino acid conservation because of its potential relevance for nucleotide conversation. No antizyme-wide conservation of the amino acids encoded by ORF1 or immediately after the frameshift site is apparent. Though conservation in the following parts of the sequence have been studied (84), the first clearly identifiable, highly conserved, amino acid is a lysine corresponding to position 153 of the human antizyme 1 protein (in a small minority of antizymes, arginine is present as an alternative). This is fully 85 amino acids (255 nucleotides) after the switch to ORF2. The 'universally' conserved region then extends to a tyrosine corresponding to position 216 of human antizyme 1 although this particular tyrosine is not invariable.

Prior research has identified signals situated downstream of the antizyme ORF1 stop codon which display a stimulatory effect on frameshifting efficiency. Even by the year 2000 it was clear that 3′ stimulators of antizyme come in more than one form (14). However, the picture of the diversity of forms that has emerged since is nothing short of amazing (Supplementary Data, Figure S1B).

The identity of the adjacent 3′ nucleotide affects the efficiency of translation termination on stop codons in both *E. coli* and mammals. Earlier we reported that in all antizyme genes the nucleotide supporting the least efficient termination is present at that position even though the identity of this nucleotide varies across phyla. After extending the analysis over hundreds of antizyme genes, this finding is still mostly valid. In *S. cerevisiae*, the identity of this nucleotide is C and it is the one present 3′ of the stop codon of the endogenous antizyme. Mutating the C to an A, which is the most efficient termination context, reduces frameshift efficiency almost 4-fold (I.P.I., unpublished data). The identity of the least frequent 3′ termination context nucleotide is also C in the nematode *C. elegans* and it is similarly present 3′ of the stop codon of antizyme ORF1. However, in a number of other antizyme mRNAs the nucleotide in that position presents the second least frequent 3′ termination context. For example, in vertebrates the nucleotide 3′ adjacent to the stop codon which is most infrequent at that position is 'U' (85) and it is the nucleotide immediately 3′ of the stop codon of ORF1 in most known vertebrate antizymes. However, experimental testing has identified 'C' as the nucleotide most likely to promote inefficient termination *in vivo*, at least following a UGA stop codon, although 'U' is the second least efficient termination context. Homologs of mammalian antizymes 3 whose frameshift sites support extremely low levels of frameshifting, at least when tested in the available heterologous tissue culture cells, have an efficient 'G' termination context (61).

Another previously noted conserved region is a pyrimidine-rich sequence with the consensus UCCCU starting 3 nt 3′ of the ORF1 stop codon and present in most metazoan antizyme genes. Although this conserved sequence was absent from the fungal antizyme genes known at the time, most notably *S. pombe*, it is now clear that many fungal branches have this sequence in the identical position. This includes most mushroom-related antizymes, and homologs from *Neurospora crassa*, *A. nidulans* and related species. An exact match to the consensus is also present in antizyme mRNAs from *S. cerevisiae* and its close relatives—i.e. *S. paradoxus*, *S. mikatae*, *S. kudriavzevii* and *S. bayanus*. However, perhaps tellingly, this sequence is not present in any of the other yeast in the *S. cerevisiae* branch that have older divergence dates—i.e. *Aphis gossypyii*, *K. waltii*,

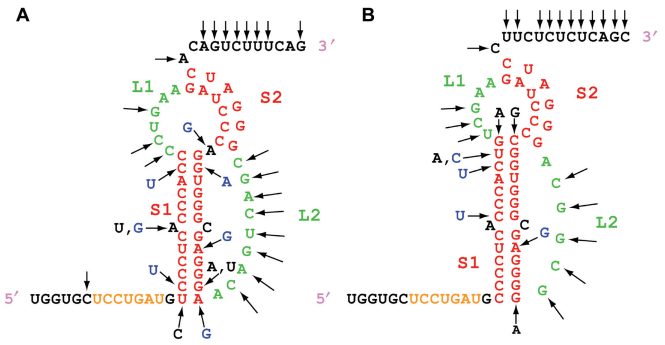



**Figure 2.** 2D representation of RNA pseudoknots—class I (**A**) of mouse antizyme 1 and (**B**) mouse antizyme 2 mRNAs. The frameshift site is indicated with orange letters. Black arrowheads represent substitutions deduced from phylogenetic comparison to orthologous genes. Non-compensatory changes in the stems are shown in black letters; compensatory changes are shown in blue letters.

*Kluyveromyces lactis*, *S. kluyveri*, *S. castellii* and *Candida glabrata*. This strongly suggests that the common ancestor in this particular fungal branch did not have the consensus sequence. Combined with its absence in many other highly divergent fungal branches, this suggests that the ancestor of all fungi and metazoa perhaps also lacked the sequence. If true, this would imply that the appearance of the UCCCU sequence is a recurring event in the evolution of antizyme genes.

Although a compelling natural history of this 3′ linear element can now be drawn there is uncertainty about its actual activity. The issue in most metazoan antizymes is complicated by the fact that a functional RNA structure that is also involved in stimulating the frameshift event, partially overlaps this region.

An RNA pseudoknot in the mRNA of mammalian antizyme 1 that starts 3 nt downstream of the stop codon (Figure 2) was historically the first antizyme mRNA 3′ element discovered and indeed the first RNA structure shown to stimulate any +1 frameshifting. The entire structure is contained within a 60-nt region. Stem 1 and stem 2 of the pseudoknot are separated by a one-nucleotide hinge which may, or may not, bulge out in the fully folded structure. Subsequently, similar or identical RNA pseudoknots were also identified in all homologs of vertebrate antizyme 1 ranging from fish to primates. In addition to the pseudoknot in homologs of antizyme 1, all vertebrate homologs of antizyme 2 have a nearly identical stimulatory RNA structure 3′ of their ORF1 stop codon. It appears that the RNA pseudoknots in homologs of both antizyme 1 and 2 evolved from a common precursor which existed in early vertebrate evolution. Lampreys (*Petromyzon marinus*) have this pseudoknot and the emergence of the common precursor likely occurred prior to the divergence of ancestral lampreys from the branch leading to extant vertebrates. However, the divergence of the non-vertebrate chordate outgroups of tunicates (sea squirts—e.g. *C. intestinalis*, *C. savignyi* or *Molgula tectiformis*) and cephalochordates (lancelets—e.g. *Branchiostoma floridae*), may have occurred prior to emergence of the common precursor pseudoknot since they do not utilize an identifiable pseudoknot. Curiously, this vertebrate RNA pseudoknot is active in stimulating frameshifting to the +1 frame (via both +1 and −2 frameshift) when an antizyme 1 frameshift cassette is expressed in *S. cerevisiae* (68) and also *S. pombe* (83) even though no analogous structure exists either in the *S. cerevisiae* or *S. pombe* antizyme mRNAs.

A different RNA pseudoknot stimulates the frameshifting in a large subset of invertebrate antizyme mRNAs (86). To avoid confusion with unrelated pseudoknots encoded by other antizyme genes including those in vertebrate antizyme 1 and 2 mRNAs, this pseudoknot will be called class II (Figure 3). Several characteristics distinguish class II pseudoknots from vertebrate antizyme 1 and 2 pseudoknots which from now on will be referred to as class I. Class II invertebrate pseudoknots extend over much larger regions of the mRNA sometimes covering over 90 nt. This expansion is mostly due to the size of loop 1 which in some cases exceeds 35 nt. A derivative of this invertebrate pseudoknot also occurs in several aphid species (86). It will now be referred to as class IIb to distinguish it from the version which is much more widespread and which will carry the designation class IIa. Justification for the sub-division comes from additional sequence information which has permitted a phylogenetic comparison (Figure 3B) and the realization that IIb is also present in fleas and psyllids. In class IIb, the putative pseudoknot stem 1 is very similar in sequence and length to the one in the 'standard' class IIa. Stem 2, however, appears quite different. More sequences have become available and a clearer picture of the nature of class IIb pseudoknot can be inferred from phylogenetic comparison (Figure 3B). The distinguishing characteristic of class IIb pseudoknot is the 'hinge' between stems 1 and 2 which instead of a single nucleotide, comprises between 4 and 14 nt, possibly forming a bulging loop but other conformations may be involved.

Despite some superficial similarities, class I and II may differ in the way they stimulate the frameshift event. One clue comes from the property of mutants in stem 1 that disrupt base-pairing and reduce frameshifting in both class I and class II pseudoknots. While compensatory mutations in class I pseudoknots restores frameshifting to near wild-type levels, compensatory mutations in stem 1 of class II pseudoknot leads to only partial restoration of frameshifting levels. This suggests that the sequence identity of stem 1 is not important for class I but is of some importance for the proper functioning of class II (86).

A class II pseudoknot is present in antizyme mRNAs of no less than five different animal phyla—*Mollusca*, *Nematoda*, *Annelida*, *Platyhelminthes* and *Arthropoda*. This no doubt attests to its great antiquity. If the conventional view for the evolution of arthropods is correct then their common ancestor did have an antizyme class II RNA pseudoknot. However, most extant insects appear to have lost this structure. This is true of flies, midges, mosquitoes, beetles, butterflies/moths and likely others.

A curious offshoot of class II pseudoknots is apparently present in the antizyme gene of *Priapulus caudatus*, a member of the metazoan Priapulida phylum and in the

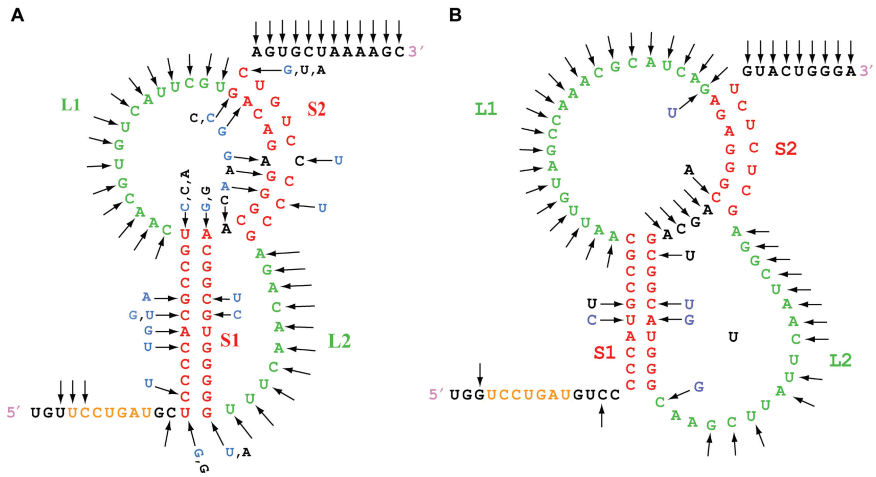



**Figure 3.** 2D representation of RNA pseudoknots (**A**) class IIa—the actual sequence is of *Crassostrea gigas* (**B**) class IIb—the sequence is of an aphid. The frameshift site is indicated with orange letters. Black arrowheads represent substitutions deduced from phylogenetic comparison to orthologous genes. Non-compensatory changes in the stems are shown in black letters; compensatory changes are shown in blue letters.

conenose bug *Rhodnius prolixus* (Hemiptera order within insects). In *P. caudatus* a potential 13 bp stem is present in its mRNA 3′ of the frameshift site. In *R. prolixus*, a potential 11 bp stem-loop is present. The sequences of these stems are essentially identical to stem 1 of class II pseudoknots implying decent from a common ancestor. However there is no apparent base-pairing potential in either *P. caudatus* or *R. prolixus* antizyme mRNAs to form stem 2 and therefore a complete pseudoknot. The implication is that in these two antizyme mRNAs a simple stem-loop is perhaps sufficient to stimulate the frameshifting. [Single stem-loops are known to stimulate −1 frameshifting (86,87)]. This would be in stark contrast to the situation in oyster (*Crassostrea gigas*, phylum Mollusca) antizyme which has a complete class II pseudoknot structure. With that pseudoknot, disruption of stem 2 has the same effect as deleting the entire pseudoknot. It is also curious to know that within another antizyme mRNA belonging to the Hemiptera species *Homalodisca coagulate*, stem 2 is both present and well conserved showing that at least in *R. prolixus* there is a clear loss of stem 2.

Another 3′ stimulator identified earlier is present in *S. pombe* antizyme mRNA. This element is known to enhance the +1 frameshifting of *S. pombe* antizyme mRNA by as much as 10-fold. It is contained within a region extending up to 150 nt downstream of the frameshift site. The importance of this region is underscored by extensive nucleotide conservation in the corresponding sequences of antizyme from the distantly related fission yeast *S. octosporus* and *S. japonicus* (13). The nature of this element, however, is currently unknown. It does not appear to be a 'simple' conventional RNA secondary structure.

The phylogenetic analysis of all the sequences included in this article points to a number of previously unknown and sometimes quite novel, putative stimulatory elements. A paradoxical candidate is present in an array of >14 distantly related fungal antizyme genes that nevertheless belong to a single evolutionary branch (Figure 4A). Notable members of the group include the species: *N. crassa*, *Gibberella zeae*, *Magnaporthe grisea*, *Chaetomium globosum*, *A. nidulans* and *Blumeria graminis*. As the functional analysis of these sequences has barely started, only the general outlines of a putative 3′ element in these antizymes can be drawn. With this caveat, the putative 3′ element features a stem-loop structure. The stem could be as short as 4 and as long as 14 or more base-pairs. The stems may be separated into at least four distinct classes—(1) the *N. crassa*, (2) *B. fuckeliana*, (3) *A. nidulans* and (4) the *B. graminis* classes (see Supplementary Data).

In all of these 14 fungal antizyme mRNAs, the stem-loop starts between 31 and 41 nt downstream from the ORF1 stop codon with the members of the different classes of stems having distinct lengths of this 'spacer' region (see Supplementary Data). The two sides of the proposed stems are separated by an unusually large loop that shows great variation in length, from 18 to 130 nt. As with the spacer region between the frameshift site and the beginning of the stem-loop, there appears to be only limited or no, conservation at the nucleotide level within the loop.

A second feature of this 3′ element is a conserved sequence starting exactly 16 nt downstream from the end of the stem-loop described above. This sequence, GGAAGARUGUGAGAGRUCUUUYUGYGA, has no apparent similarity at the nucleotide level to known functional RNA motifs nor has it an obvious secondary structure. Not surprisingly, considering the level of conservation at the nucleotide level, the peptide encoded by it, EECER(L/F)FC(D/E), is also conserved. There is an insufficient number of sequences known, given the conservation at both the nucleotide and amino acid level, to predict if the conserved region works at the RNA or peptide level. The degree of conservation of this sequence, both at nucleotide and amino acid levels, far exceeds the conservation further downstream corresponding to the region of the antizyme responsible for the known biochemical activity of the protein.

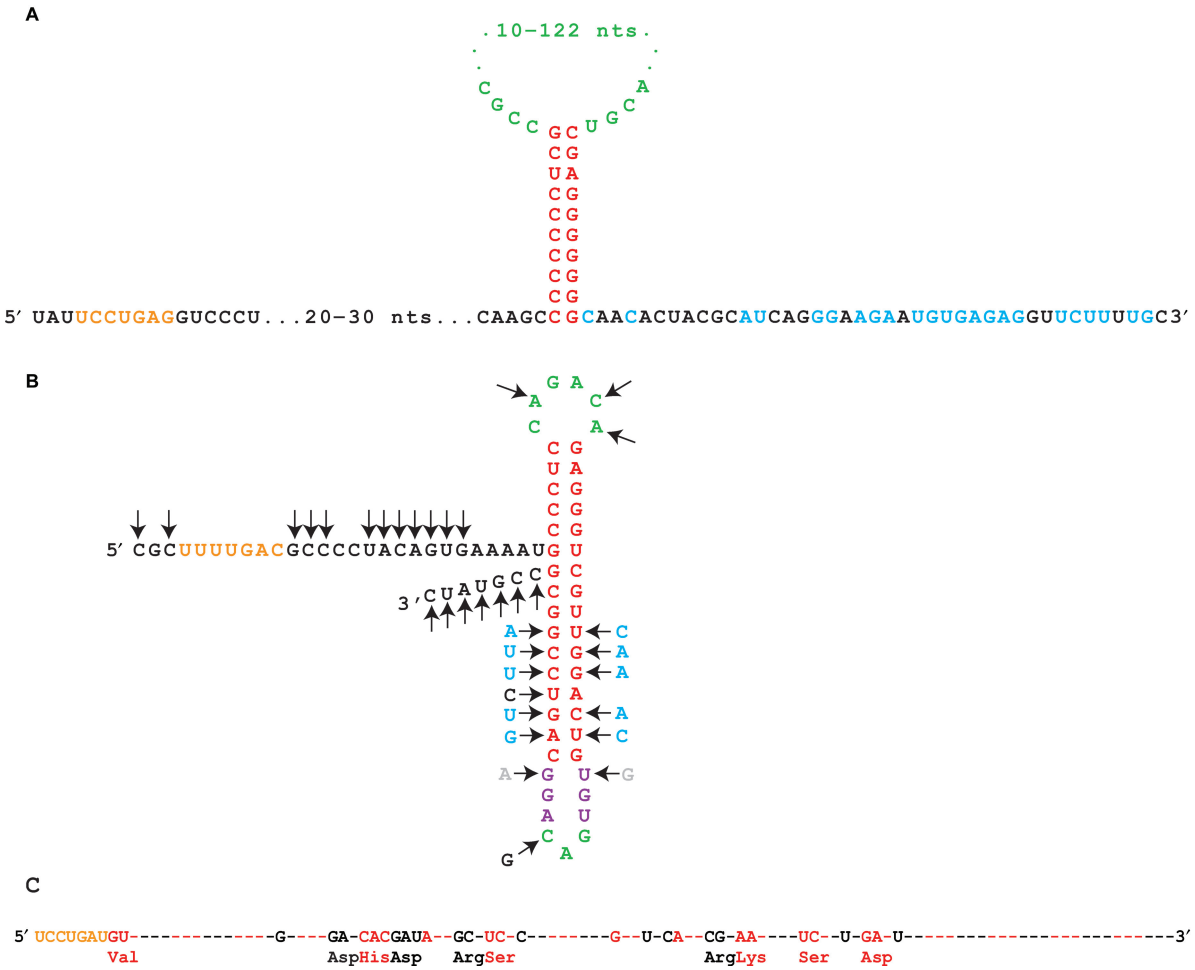



**Figure 4.** Newly identified antizyme mRNA potential 3′ stimulators. Three elements strongly supported by phylogenetic data are shown. The frameshift sites are indicated with orange letters. (**A**) 3′ conserved stem-loop element in a large subset of fungi. The actual sequence shown is of *N. crassa*. Base-paired nucleotides are in red. The loop is shown in green. Blue-colored nucleotides following the stem-loop are absolutely conserved. (**B**) Base-paired nucleotides are in red. Nucleotides potentially extending stem 2 by non-Watson–Crick base-pairs are shown in magenta. The loop is shown in green. Black arrowheads represent substitutions deduced from phylogenetic comparison to antizymes in other mushroom-related species. Non-compensatory changes in the stems are shown in black letters; compensatory changes are shown in blue letters; and a potential non-Watson–Crick G-A base-pairing nucleotides are shown in gray. The actual sequence shown is of *Agaricus bisporus*. (**C**) The conserved 3′ element in flies, mosquitoes and midges. The top line shows the nucleotide sequence and the bottom line the amino acid sequence. Only absolutely conserved positions are shown. Variant nucleotide positions are indicated by '-'. Alternating codons in the +1 frame (in-frame with ORF2) are shown in red and black.

How the stem-loop and the conserved sequence come together, if at all, to affect the +1 frameshifting in the antizyme mRNAs in which they occur, is not obvious. (With mammalian antizyme 1 mRNA pseudoknot, the distance of the pseudoknot, 3–5 nt 3′ of stop codon, is crucial.) The position of the stem-loop, 31–41 nt downstream of the stop codon of ORF1, potentially puts it well outside the ribosome during decoding the 3′ end of ORF1. This would be surprising if the putative structure is indeed involved in stimulating the frameshift. The only partial analog of the conserved nucleotide region is a somewhat shorter region of yeast TY3 retrotransposon where the sequence 'CUAACCGAUCUUGA' starting immediately downstream of the GCG-AGU-U shift site stimulates the +1 ribosomal frameshift up to 6-fold (89).

Another putative novel 3′ element is a stem-loop structure present in 10 fungi related to mushrooms (see Figure 4B). Their alignment highlights a region extending 65 nt downstream of the ORF1 stop codon which shows a high level of conservation. This region has two candidate stem-loops occurring without a spacer and therefore likely to coaxially stack on each other. The first starts 17 nt 3′ of the stop codon and is 6 bp long, four of them C-G. The stem encloses a 6–9-nt- long loop. The second stem is perhaps 10 bp long. Phylogenetic comparison identifies six compensatory co-variations in it, providing high confidence in its existence and physiological significance. The putative loop in the second stem is 9 nt long but has features suggesting it may exist in something other than a single-stranded state. Specifically, six of the nine nucleotides are absolutely conserved and two others, at the

beginning and the end respectively, show intriguing co-variation. Every time there is U at the 5′ end there is G at the other end and every time there is G at the 5′ end there is A at the other. This suggests the possibility for non Watson–Crick U-G or G-A base-pair to be followed by a G-G and a standard A-U base-pair topped by a triloop (Figure 4B). Additionally, the region 3′ of the ORF1 stop codon has a perfectly conserved tetranucleotide sequence AAAU immediately preceding the first stem-loop. This may also be important. No nucleotide conservation is seen after the 3′ end of the second stem-loop.

Another RNA secondary structure may be present in the region 3′ of the frameshift site in antizyme mRNAs of beetles. The hypothetical stem is spaced barely one nucleotide from the stop codon of ORF1. It is 7–8-bp-long of which at least 6 are G-Cs. The stem appears to be capped by a loop of 5–7 nt long. This structure is supported phylogenetically by two co-variant changes in the stem. There is also some support for a second stem-loop structure in beetle antizyme mRNAs starting 18–21 nt downstream of the first stem-loop. The putative second stem-loop is only 5 bp long with just two of them G-Cs, but the nucleotides that form each half of the stem are much better conserved than surrounding mRNA regions.

Far more challenging to decipher might be a potential 3′ stimulator in antizyme genes of flies, midges and mosquitoes. Unlike most of the new putative stimulators described above, there is strong circumstantial evidence to suggest that an unknown +1 frameshift stimulator is present in at least *D. melanogaster* antizyme mRNA. This evidence comes from previously published *in vitro* translation experiments in which the *D. melanogaster* antizyme was expressed in reticulocyte lysates or wheat germ extracts. In both cases the observed level of frameshifting far exceeds the level expected from the limited conservation of the frameshift site relative to mammalian antizyme 1 (although these experiments (12), did not address the question what the additional stimulators might be). Examining the sequences from flies, midges and mosquitoes identifies 22 absolutely conserved nucleotides within a 60-nt region just 3′ of the frameshift site (see Supplementary Data). Sixteen of them occur in pairs and in all of those cases the two adjacent nucleotides form the first two positions of a codon (in-frame with ORF2) with the variable third nucleotide always resulting in synonymous changes. At least these 16 nt seem to be conserved for their amino-acid-coding potential rather than their primary nucleotide sequence (see Figure 4C). The equivalent region of antizyme protein in vertebrates is not known to have any biochemical function. Six other nucleotides are conserved for reasons other than their coding potential as they are part of triplets encoding more than one amino acid. Six others are ambiguous as they are part of triplets without variation at the third position. Of the 10 conserved amino acids in this region 6 are charged (either positively or negatively).

An analogous pattern of 3′ conservation can be detected in jacobid protists (*Seculamonas ecuadoriensis* and *Jakoba libera*). The conserved sequence in that case is the peptide VIP(D/E)LGFR whose coding starts 15 nt (five codons) after the frameshift site.

To the novel 3′ elements listed so far, which are supported by strong phylogenetic evidence, several more can be added which are supported less well because of the lack of a sufficient number of close and/or more distant relatives for comparison. Among these, several are, nevertheless, worthy of notice. One is a putative RNA pseudoknot that might be present in sea urchins (Phylum Echinodermata). This potential pseudoknot is supported by comparing sequences from two sea urchin species, *Paracentrotus lividus* and *Strongylocentrotus purpuratus*. Stem 1 of this putative pseudoknot is 10 bp long, comparable to the RNA pseudoknots that have been shown to stimulate +1 frameshifting in the decoding of antizyme mRNAs in vertebrates and invertebrates, as discussed above. Stem 2 is 8–11 bp long with the two stems likely separated by a 2 nt 'CC' hinge. The overall architecture is not radically dissimilar to what is known of the functionally active RNA pseudoknots in vertebrate and invertebrate antizyme mRNAs. However, the distance, 40 nt, from the ORF1 stop codon to the beginning of the first stem is distinctive. While 3′ distant stimulators for frameshifting are known in diverse cases (90,91), in the best known case there is an additional sequence very close to the frameshift site with which the distant signal interacts (91).

A putative 3′ RNA pseudoknot may exist in the antizyme mRNA of the opportunistic pathogenic fungus *C. neoformans*. Phylogenetic comparison of sequences from different *C. neoformans* strains and the distant relative *C. laurentii*, strongly supports the hypothesis that a functionally important element is present within the 101 nt 3′ of the ORF1 stop codon. There is ambiguity as to its actual structure, if any.

Another possible 3′ pseudoknot may be present in the sea anemone *Nematostella vectensis* (Phylum Cnideria); however, due to a lack of related sequences to compare it to, the structure could not be assessed by phylogenetic analysis. Stem 1 of this putative pseudoknot consists of 11 bp, 9 of which are G-Cs resulting in a structure with a free energy of −27 kcal/mol or more. As with the putative pseudoknot in sea urchins, the distance from the ORF1 stop codon to the beginning of the proposed pseudoknot is unexpectedly long, 38 nt. Nucleotides enclosed between the two parts of this stem might also base-pair to substantially improve the free energy of the overall structure. Additionally, a 7-nt segment downstream of the stem-loop can form perfect base-pairing with a region within loop 1 to complete a pseudoknot structure.

Yet another 'orphan' 3′ pseudoknot might be present in antizyme mRNA of the green algae *Glaucocystis nostochinearum*.

## 5′ CONSERVATION AND FRAMESHIFT STIMULATORS

### mRNA level

Comparing sequences from a large number of antizyme genes (Supplementary Data, Figure S1C) allows us to define the 5′ elements more precisely. However, this information is yet to bring us closer to understanding the mechanism through which they enhance +1 frameshift

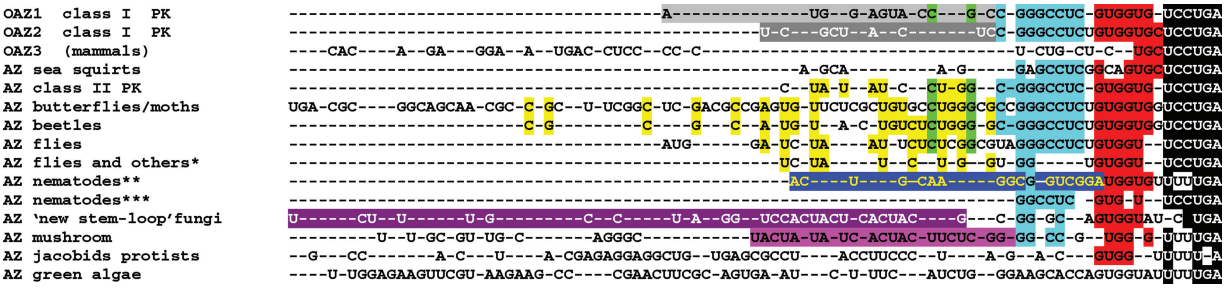



**Figure 5.** Conserved nucleotide sequences 5′ of the frameshift site. In each case only the consensus is shown and poorly conserved positions are indicated by '-'. Gaps in alignments are shown as empty spaces. The frameshift site is highlighted in black. Module A is highlighted in red. Module B is highlighted in light blue. Module C is highlighted in gray. Module D is highlighted in yellow. Module E is highlighted in dark blue. Module F is highlighted in purple. 'PK' = RNA pseudoknot. *Flies, midges and mosquitoes. **_C. elegans_ and related nematodes (see Figure S1C). ***_Onchocerca volvulus_ and related nematodes (see Figure S1C).

efficiency and especially the mechanism through which this region mediates the polyamine regulation of frameshifting.

It has been suggested that the 5′ element present in mammalian homologs of antizyme 1 is modular in nature. An earlier study identified three distinct modules each apparently working, at least partially, independently from the other and evolving in three distinct chronological stages (14). The analysis of additional homologs of antizyme allows refinement of this model. To facilitate the description of the 5′ stimulators in the various branches of the antizyme gene tree, we propose a new classification system for the different modules. The modules in mammalian homologs of antizyme 1 and 2 mRNAs previously referred to as 'proximal' is now being termed 'A', the 'middle' as 'B' and the 'distal' as 'C'. All other modules inferred to have evolved independently from these three will be designated with their own letters. Modules that are variants will be indicated with a number following the letter.

The analysis of over 45 homologs of vertebrate antizyme 1 genes shows clear conservation of nucleotides belonging to all three modules; A, B and C. Modules A and B are essentially completely conserved (Figure 5). Conservation of the sequence of module C is considerably less but still clearly present. A similar picture emerges from the analysis of 19 vertebrate homologs of antizyme 2. Once again modules A and B are completely conserved with module C less conserved. The consensus sequence for module C in antizyme 1 and antizyme 2 mRNA s shows a distinct pattern of conserved nucleotides which leads us to label one as C1 and the other C2. C1 and C2 appear to have common evolutionary origin hence use of the same letter designation for both.

A sequence approximating a primitive version of module C appears present in the antizyme gene of lampreys (Cephalaspidomorphi) which we currently suspect to be the 'living fossil' diverging just prior to the gene duplication that eventually lead to formation of antizyme 1 and 2 paralogous groups but after the emergence of class I pseudoknots.

None of the three 5′ modules of antizyme 1 and 2 mRNAs are preserved in homologs of the mammalian antizyme 3 gene (14). Earlier we noted that mouse antizyme 3 mRNA 5′ region has what appears to be a triplication of the frameshift site resulting in two pseudo-frameshift sites with the sequence GN-UCC-UGC preceding the actual frameshift site sequence GN-UCC-UGA (14). The presence of one copy of GN-UCC-UGC was also noted in human antizyme 3 mRNA and it appeared to have resulted from a reduction of the initial triplication. With more than 10 antizyme 3 mRNA sequences analyzed now, our earlier suggestion appears to be incorrect. The triplication is only present in the rodent line (mice and rats) while humans and other mammals and the one antizyme 3 gene from marsupials, share the same 'solo occurrence'. This indicates that the sequence in humans and marsupials is not the result of trimming back an original triplication event but rather the single pseudo-frameshift site in humans and several other mammalians is the result of a duplication event which is evolutionary independent and predates the duplication in mice and rats. Whether the pseudo-frameshift site feature plays a role in stimulating +1 frameshifting in antizyme 3 genes is unknown. However, its emergence and conservation following two apparently independent molecular events suggests possible functional significance.

In the three available sea squirt antizyme mRNA sequences, the 5′ modules have diverged in a somewhat interesting and potentially informative way. Sea squirts are Urochordates and so like the vertebrate species analyzed, are expected to have near perfect conservation of the 5′ modules A and B since these two modules are conserved not only in vertebrates but also in many other animal antizyme mRNAs belonging to several different phyla. Vestiges of modules A and B are indeed readily identifiable in squirt antizyme mRNAs (see Supplementary Data). Sea squirt's antizyme mRNAs also show conservation within the region corresponding to module C1 of vertebrates though the actual sequence is different, perhaps mirroring changes in the interacting molecule.

Comparing 29 invertebrate antizyme mRNA sequences that have a class II antizyme pseudoknot, reveals extensive conservation in the region 5′ of the frameshift site indicative of functional sequences. When the nucleotide consensus of these sequences is compared to that of antizyme 1 and 2 from vertebrates, it is clear that modules A and B are present while a new module, 'D', has evolved

in a position similar to module C in genes with class I antizyme pseudoknot. Module D actually shares four conserved nucleotides with module C1 but it is unclear if this is due to common ancestry.

In flies and mosquitoes, modules A and B are strongly conserved. A sequence clearly derived from module D is also present although the actual sequence is different. This pattern is also seen in beetles, butterflies and moths but once again the sequence of module D diverges from the consensus for antizyme mRNAs with a class II pseudoknot and indeed from that of flies also.

Several nematode antizyme genes have strong conservation of both module A and B. However in others, which constitute the vast majority of available antizyme mRNA sequences in nematodes, module B appears to have been lost. Perhaps more importantly for determining the mode of operation of all the 5′ modules mentioned so far, in the few worms with an intact module B, it is present in a different register from its homologs in other metazoans; i.e. it is shifted by one nucleotide 5′ relative to the others. In the antizyme mRNAs of worms which belong to the same clade as *C. elegans*, all of which have a frameshift site UUU-UGA, module B is not present. However, another conserved sequence, module E, extending up to 34 nt 5′ of module A, is easily identifiable.

The antizyme mRNA sequences from fungi belonging to the group with the new stem-loop (i.e. the stem-loop shown in Figure 4A), exhibit extensive conservation 5′ of the frameshift site. Some of the conserved positions extend as far as 95 nt 5′ of the ORF1 stop codon comprising module F (see Figure 5).

Most antizyme genes from protozoa show clear evidence for the presence of module A indicating that this element evolved before the divergence of metazoa, fungi and protozoa.

Following the analogy of an interaction between the rRNA in translating *E. coli* ribosomes and mRNA stimulating frameshifting at a nearby shift site 3′ (92,93), and suggestions or evidence for the rRNA of eukaryotic ribosomal subunits or ribosomes interacting with some specific mRNA sequences (94–96), the influence of the antizyme 5′ element may come from its interaction with rRNA. This is not easy to test experimentally. However, conservation in the antizyme gene in Basidiomycota fungi related to mushrooms, may break the mould. It presents a case where the important feature of a 5′ element may not be its sequence directly but rather that of its encoded product.

## Protein level

Analyzing the 10 sequenced mushroom antizyme sequences available at the moment (Figure 6) reveals that although the conservation in mRNAs 3′ of the frameshift site is limited to the proximal 70 nt as discussed above, the conservation 5′ is much more extensive. It covers almost the entire region of ORF1 (see Supplementary Data). The conservation near the frameshift site is at the amino-acid-level and the peptide encoded by this region is: YYYSTTFSGGP(G/E)WRF

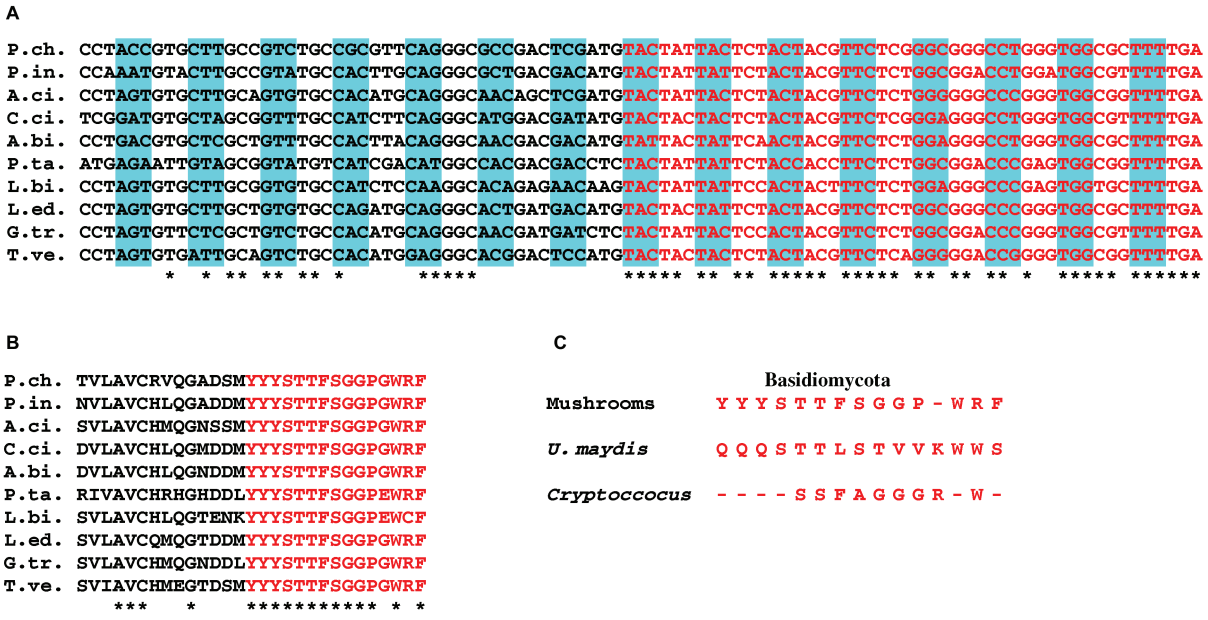



**Figure 6.** Conserved sequences 5′ of the frameshift sites of mushrooms and other Basidiomycota. The most highly conserved region is indicated by red letters. (**A**) Alignment of the 90 nt 5′ of the frameshift site of 10 mushroom-related species. Absolutely conserved positions are shown by '*'. Alternating codons in the 0 frame (in-frame with ORF1) are highlighted in light blue or not at all. (**B**) Alignment of the last 29 amino acids of ORF1 of 10 mushroom-related species. Absolutely conserved positions are shown by '*'. (**C**) Comparison of the last 15 amino acids of ORF1 of different Basidiomycota—mushrooms, *Ustilago maydis* and *Cryptoccocus* spp. The *Cryptoccocus* sequence is based on several *C. neoformans* strains and *C. laurentii*.

(Figure 6B). It is not possible to infer the function of this peptide without mutational analysis of the sequence. It is tempting to speculate that it might be involved in stimulation of +1 frameshifting in a manner similar to that of nascent peptide involved in the translational bypassing in decoding phage T4 gene *60*, where half the ribosomes perform +50 frameshifting (97,98). Of course many other functions of the conserved sequence could easily be proposed and only mutational analysis will resolve this problem.

Intriguingly, in another very distant Basidiomycota species, *Ustilago maydis*, the corresponding sequence of ORF1 antizyme is QQQSTTLSTVVKWWS (Figure 6C). This bears more than a passing similarity to the sequence in mushroom antizymes. The biggest difference is the substitution of three consecutive tyrosines with three consecutive glutamines. Since the only physicochemical property tyrosine and glutamine have in common is their side-chain polarity it is not obvious what the physiological role of the run of three tyrosines or glutamines might be.

Another 5′ conserved antizyme mRNA sequence that seems to depart from the standard pattern, is in the antizyme mRNAs of an opportunistic fungal pathogen *C. neoformans* and its distant relative *C. laurentii* (see Supplementary Data). The peptide, SSFAGGGR(R/Q)W(V/F), encoded by this region of *C. neoformans* and *C. laurentii*, is, however, substantially different from the corresponding consensus sequence of mushroom antizymes. Only three amino acid positions are perfectly conserved in all known basidiomycotal sequences.

### Alternative start codons

Although not directly related to stimulation of antizyme mRNA +1 frameshifting, there are several more features of ORF1 that merit consideration. Antizyme mRNA ORF1 is not known to encode any enzymatic activity, but the identity of the peptide sequence encoded by ORF1 may be important in another way. Mammalian antizyme 1 mRNA ORF1 has two potential in-frame AUG codons. Almost since the ORF was first cloned, it was discovered that both AUGs can serve as translation initiation sites (11). Initiation at the first causes synthesis of an N-terminal mitochondrial localization sequence which is absent when translation is initiated at the second downstream AUG (39,40). Not only do all known mammalian homologs of antizyme 1 mRNA have two in-frame AUGs in their ORF1, but also the position of both is absolutely conserved. Antizyme 1 mRNAs from birds, reptiles and amphibians also have two AUGs although in a slightly different position. Bony fish have two paralogs that are orthologous to antizyme 1 genes in other higher vertebrates. One is known as '1S' and the other as '1L' (though S and L originally referred to short and long, analysis of further sequences showed that the length difference between the paralogous pair in a particular species maybe as small as 1 amino acid, and also that the S ortholog in one species can be longer than the L form of another species). Orthologs of antizyme 1S have the same arrangement of AUGs in ORF1 as higher vertebrates while orthologs of antizyme 1L usually have three in-frame AUGs in their ORF1 (99,100).

During alignment of mammalian antizyme 3 mRNA sequences, a surprising degree of nucleotide conservation was observed in the region assumed to be 5′ UTR. The region, up to 156 nt 5′ of the previously proposed AUG translation initiation codon of human ORF1 antizyme 3 mRNA sequence, was more carefully examined (Supplementary Data, Figure S2). This analysis reveals the potential of the 0 frame to encode a conserved peptide. Since no AUG initiation codon is present in-frame 5′ of the previously designated start site, if this additional 5′ extension is indeed translated, initiation must occur at a non-standard initiation codon. The most likely candidate is an absolutely conserved CUG leucine codon. CUG is known to initiate translation in several other mammalian genes. Initiation at the CUG would result in a 48 amino acid N-terminal extension of the human antizyme 3, which is highly conserved among the orthologs (see Supplementary Data).

An in-frame AUG 17 codons upstream of antizyme mRNA ORF1 stop codon is curiously conserved in flies. All fly antizyme mRNAs have other in-frame AUG codons further 5′, the locations of which are not conserved. Whether the AUG 17 codons before the 3′ end of ORF1 is used as an alternative translation initiation site is unknown.

## 3′ UTR

Insufficient attention has been paid to the role, if any, of the 3′ untranslated region (UTR) in the regulation of antizyme genes. Human antizyme 1 mRNA has alternative polyadenylation signals with the two sites separated by about 150 nt (14,25). The ratio of the two polyadenylation products varies in different tissues suggesting a possible regulatory function. This is supported by high conservation of both polyadenylation sites in mammalian antizyme 1 genes, by examination of more than a thousand ESTs each for human and mouse (starting/ending with their polyA tails), and sequence data from several other mammals—e.g. cattle (*Bos taurus*), sheep (*Ovis aries*), pig (*Sus scrofa*), rat (*Rattus norvegicus*) and dog (*Canis familiaris*). Further supportive data comes from the >20 ESTs for antizyme 1 orthologs in the zebra finch (*Taeniopygia guttata*). Analysis of over a hundred EST 3′ ends of chicken (*Gallus gallus*) antizyme 1, however, did not identify any clones corresponding to the longer transcript.

A comparison of 3′ UTR sequences of antizyme 1 orthologs in the same group of animals identifies another potentially significant conserved feature (Figure 7A). A uridine-rich sequence overlapping the upstream polyadenylation signal, AAUAAA, and extending about 90 nt 5′ and about 10 nt 3′, is highly conserved from human to chicken with nearly 71% of all positions absolutely conserved among 14 different orthologs of the gene. A further 24% are conserved in at least 11 of the 14 orthologs (in both cases relative to the human sequence). Curiously, at the 5′ boundary of this region there is a

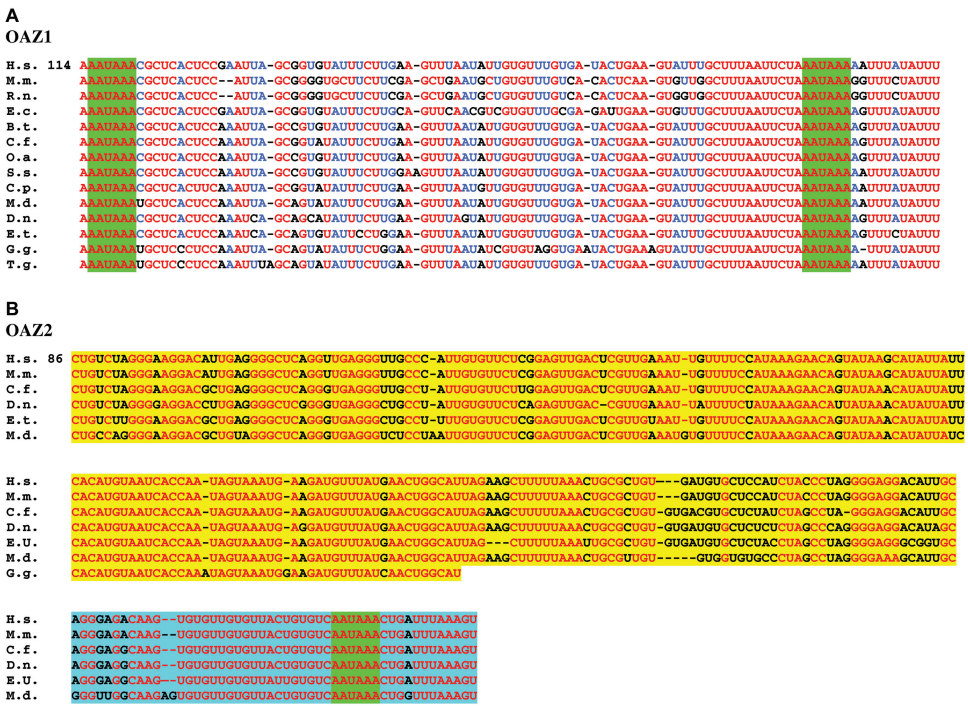



**Figure 7.** Nucleotide alignment of conserved elements in the 3′ UTR of vertebrate antizyme genes. Gaps in alignment are shown by '-'. The number of the top line in each case indicates the distance to the stop codon of ORF2 in human. (**A**) Alignment of the conserved region in orthologs of antizyme 1. Absolutely conserved nucleotides are in red and those conserved in at least 11 of the 14 species are in blue. Less well-conserved positions are in black. The two heptanucleotide sequences matching the consensus polyadenylation site are highlighted in green. (**B**) Absolutely conserved nucleotides are in red. Less well-conserved positions are in black. The upstream conserved region is highlighted in yellow. The downstream conserved region is highlighted in blue. The polyadenylation sites are highlighted in green. Species abbreviations are as follows: H.s. = human, M.m. = mouse, R.n. = rat, C.f. = dog, D.n. = nine-banded armadillo, E.c. = horse, B.t. = cow, O.a. = sheep, S.s. = pig, C.p. = guinea pig, E.t. = Madagascar hedgehog, M.d. = opossum, G.g. = chicken, T.g. = zebra finch.

perfectly conserved AAUAAA sequence matching the consensus polyadenylation signal which, however, as far as can be determined from EST sequences, is never used as such.

Alternative polyadenylation is seen in ~50% of genes in human and mouse. In approximately one-third of all genes, the alternative polyadenylation sites are conserved between human and mouse (101). Much less is known about the role played by such alternative polyadenylation, although roles in cellular localization, stability or translational efficiency have been suggested (102). Bioinformatic approaches have identified *cis*-acting elements that are associated with regulation of polyadenylation (103). The conserved region in the 3′ UTR of antizyme 1 perhaps in part consists of such elements but cannot be explained in its entirety in these terms.

Two polyadenylation sites occurring in approximately equivalent positions to the ones described above are also present in amphibian homologs of antizyme 1 although the conserved region 5′ of the first site shows no similarity to the sequence seen in mammals and birds.

The 3′ UTR of mammalian and avian orthologs of antizyme 2 mRNA is much longer (usually ~1100 nt) than the 3′ UTR of mammalian and avian orthologs of antizyme 1 (usually ~400 nt). Unlike antizyme 1 mRNA, it shows no signs of alternative polyadenylation though it too shows extensive nucleotide conservation (Figure 7B). The conservation pattern is different from that of antizyme 1 mRNA, and is split into two separate regions. The most extensive conservation is close to the 5′ end of the 3′ UTR. This segment is about 200 nt long and includes a stretch of 48 nt which is 94% conserved between human and chicken. The other conserved region is near the polyadenylation signal of the 3′ UTR and it is detectable only in mammalian and marsupial orthologs of the gene. Once again the role(s) that these conserved regions play in regulation of antizyme 2 expression are not clear. No significant nucleotide conservation is observed in the 3′ UTR of mammalian orthologs of antizyme 3.

The presence of conserved regions in the 3′ UTRs of vertebrate orthologs of antizyme 1 and 2 is not unique to antizyme mRNAs. Highly conserved elements in the 3′ UTR of eukaryotic mRNAs in general are much more prevalent when the protein products are regulatory (104), just like antizymes are. In some of the mRNAs with conserved sequences in their 3′ UTRs, structure formation is important. While obvious structure is not apparent in the conserved antizyme 3′ UTR sequences, experimental testing is required to address this aspect and functional significance.

## PERSPECTIVE

The central theme of the previous survey and summary on this topic was the remarkable conservation of the antizyme frameshift sites across enormous evolutionary

distances (13,14). For example the twelve nucleotides, UGG-UGC-UCC-UGA, surrounding the frameshift site of human antizyme 1 are also completely conserved in the homolog in *S. pombe*. Although the original observation still holds true, the present analysis of antizyme genes, an order of magnitude more numerous than before, allows us to appreciate another aspect of the evolution of this gene. While the majority of known antizyme mRNAs have retained what are apparently remnants of the original design, 5′ module A and the shift site UCC-UGA, many branches have evolved frameshift sites that take advantage of specific features of their own translational machinery. Still others have built an array of sometimes baroque 5′ and 3′ stimulatory sequences, each apparently arising independently. Some of these sequences are able to stimulate +1 frameshifting even when expressed in organisms whose antizyme genes lack such features and likely over-ride universal aspects of standard eukaryotic decoding.

Experimental testing of the features identified in this survey, and the prospect of identifying additional ones as more sequences become available, should advance knowledge of the richness of reprogrammed genetic decoding in eukaryotes.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank Dr Senya Matsufuji (Jikei University School of Medicine, Tokyo), the discoverer of antizyme frameshifting, and our collaborators Michael Howard and Pavel Baranov for their comments on the manuscript. This work was supported by an award from Science Foundation Ireland. Funding to pay the Open Access Publication charges for this article was provided by Science Foundation Ireland.

*Conflict of interest statement*. None declared.

## REFERENCES

1. Gesteland,R.F. and Atkins,J.F. (1996) Recoding: dynamic reprogramming of translation. *Ann. Rev. Biochem.*, **65**, 741–768.
2. Namy,O., Rousset,J.P., Napthine,S. and Brierley,I. (2004) Reprogrammed genetic decoding in cellular gene expression. *Mol. Cell*, **13**, 157–168.
3. Baranov,P.V., Gesteland,R.F. and Atkins,J.F. (2002) Recoding: translational bifurcations in gene expression. *Gene*, **286**, 187–201.
4. Brandt,J., Veith,A.M. and Volff,J.-N. (2005) A family of neofunctionalized Ty3/gypsy retrotransposon genes in mammalian genomes. *Cytog. Genome Res*, **110**, 307–317.
5. Manktelow,E., Shigemoto,K. and Brierley,I. (2005) Characterization of the frameshift signal of Edr, a mammalian example of programmed −1 frameshifting. *Nucleic Acids Res.*, **33**, 1553–1563.
6. Wills,N.M., Moore,B., Hammer,A., Gesteland,R.F. and Atkins,J.F. (2006) A functional −1 ribosomal frameshift signal in the human paraneoplastic Ma3 gene. *J. Biol. Chem.*, **281**, 7082–7088.
7. Craigen,W.J. and Caskey,C.T. (1986) Expression of peptide chain release factor 2 requires high-efficiency frameshift. *Nature*, **322**, 273–275.
8. Baranov,P.V., Gesteland,R.F. and Atkins,J.F. (2002) Release factor 2 frameshifting sites in different bacteria. *EMBO Rep.*, **3**, 373–377.
9. Higashi,K., Kashiwagi,K., Taniguchi,S., Terui,Y., Yamamoto,K., Ishihama,A. and Igarashi,K. (2006) Enhancement of +1 frameshift by polyamines during translation of polypeptide release factor 2 in *Escherichia coli*. *J. Biol. Chem.*, **281**, 9527–9537.
10. Bekaert,M., Atkins,J.F. and Baranov,P.V. (2006) ARFA: a program for annotating bacterial release factor genes, including prediction of programmed ribosomal frameshifting. *Bioinformatics*, **22**, 2463–2465.
11. Matsufuji,S., Matsufuji,T., Miyazaki,Y., Murakami,Y., Atkins,J.F., Gesteland,R.F. and Hayashi,S. (1995) Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme. *Cell*, **80**, 51–60.
12. Ivanov,I.P., Simin,K., Letsou,A., Atkins,J.F. and Gesteland,R.F. (1998) The *Drosophila* gene for antizyme requires ribosomal frameshifting for expression and contains an intronic gene for snRNP Sm D3 on the opposite strand. *Mol. Cell. Biol.*, **18**, 1553–1561.
13. Ivanov,I.P., Matsufuji,S., Murakami,Y., Gesteland,R.F. and Atkins,J.F. (2000) Conservation of polyamine regulation by translational frameshifting from yeast to mammals. *EMBO J.*, **19**, 1907–1917.
14. Ivanov,I.P., Gesteland,R.F. and Atkins,J.F. (2000) Antizyme expression: a subversion of triplet decoding, which is remarkably conserved by evolution, is a sensor for an autoregulatory circuit. *Nucleic Acids Res.*, **28**, 3185–3196.
15. Chattopadhyay,M.K., Murakami,Y. and Matsufuji,S. (2001) Antizyme regulates the degradation of ornithine decarboxylase in fission yeast *Schizosaccharomyces pombe*. Study in the *spe2* knockout strains. *J. Biol. Chem.*, **276**, 21235–21241.
16. Palanimurugan,R., Scheel,H., Hofmann,K. and Dohmen,R.J. (2004) Polyamines regulate their synthesis by inducing expression and blocking degradation of ODC antizyme. *EMBO J.*, **23**, 4857–4867.
17. Ivanov,I.P., Gesteland,R.F. and Atkins,J.F. (2006) Evolutionary specialization of recoding: frameshifting in the expression of *S. cerevisiae* antizyme mRNA is via an atypical antizyme shift site but is still +1. *RNA*, **12**, 332–337.
18. Cavalier-Smith,T. (2006) Cell evolution and earth history: stasis and revolution. *Phil. Trans. Ray. Soc. B.*, **361**, 969–1006.
19. Roger,A.J. and Hug,L.A. (2006) The origin and diversification of eukaryotes: problems with molecular phylogenetics and molecular clock estimations. *Phil. Trans. Ray. Soc. B.*, **361**, 1039–1054.
20. Cohen,S.S. (1998) *A Guide to the Polyamines*. Oxford University Press, New York, USA.
21. Wallace,H.M., Fraser,A.V. and Hughes,A. (2003) A perspective of polyamine metabolism. *Biochem. J.*, **376**, 1–14.
22. Janne,J., Alhonen,L., Pietila,M. and Keinanen,T.A. (2004) Genetic approaches to the cellular functions of polyamines in mammals. *Eur. J. Biochem.*, **271**, 877–894.
23. Gerner,E.W. and Meyskens,F.L. (2004) Polyamines and cancer: old molecules new understanding. *Nat. Rev. Cancer*, **4**, 781–792.
24. Heller,J.S., Fong,W.F. and Canellakis,E.S. (1976) Induction of a protein inhibitor to ornithine decarboxylase by the end products of its reaction. *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 1858–1862.
25. Matsufuji,S., Miyazaki,Y., Kanamoto,R., Kameji,T., Murakami,Y., Baby,T.G., Fujita,K., Ohno,T. and Hayashi,S. (1990) Analyses of ornithine decarboxylase antizyme mRNA with a cDNA cloned from rat liver. *J. Biochem.*, **108**, 365–371.
26. Miyazaki,Y., Matsufuji,S. and Hayashi,S. (1992) Cloning and characterization of a rat gene encoding ornithine decarboxylase antizyme. *Gene*, **113**, 191–197.
27. Coffino,P. (2001) Regulation of cellular polyamines by antizyme. *Nature Rev. Mol. Cell Biol.*, **2**, 188–194.
28. Murakami,Y., Matsufuji,S., Kameji,T., Hayashi,S., Igarashi,K., Tamur,T., Tanaka,K. and Ichihara,A. (1992) Ornithine decarboxylase is degraded by the 26S proteasome without ubiquitination. *Nature*, **360**, 597–599.
29. Zhang,M., Pickart,C.M. and Coffino,P. (2003) Determinants of proteasome recognition of ornithine decarboxylase, a ubiquitin-independent substrate. *EMBO J.*, **22**, 1488–1496.

30. Suzuki,T., He,Y., Kashiwagi,K., Murakami,Y., Hayashi,S. and Igarashi,K. (1994) Antizyme protects against abnormal accumulation and toxicity of polyamines in ornithine decarboxylase-overproducing cells. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 8930–8934.
31. Mitchell,J.L., Judd,G.G., Bareyal-Leyser,A. and Ling,S.Y. (1994) Feedback repression of polyamine transport is mediated by antizyme in mammalian tissue-culture cells. *Biochem. J.*, **299**, 19–22.
32. Hoshino,K., Momiyama,E., Yoshida,K., Nishimura,K., Sakai,S., Toida,T., Kashiwagi,K. and Igarashi,K. (2005) Polyamine transport by mammalian cells and mitochondria: role of antizyme and glycosaminoglycans. *J. Biol. Chem.*, **280**, 42801–42808.
33. Bercovich,Z. and Kahana,C. (2004) Degradation of antizyme inhibitor, an ornithine decarboxylase homologous protein, is ubiquitin-dependent and is inhibited by antizyme. *J. Biol. Chem.*, **279**, 54097–54102.
34. Fujita,K., Murakami,Y. and Hayashi,S. (1982) A macromolecular inhibitor of the antizyme to ornithine decarboxylase. *Biochem. J.*, **204**, 647–652.
35. Murakami,Y., Matsufuji,S., Nishiyama,M. and Hayashi,S. (1989) Properties and fluctations *in vivo* of rat liver antizyme inhibitor. *Biochem. J.*, **259**, 839–845.
36. Kim,S.W., Mangold,U., Waghorne,C., Mobascher,A., Shantz,L., Banyard,J. and Zetter,B.R. (2006) Regulation of cell proliferation by the antizyme inhibitor: evidence for an antizyme-independent mechanism. *J. Cell Sci.*, **119**, 2853–2861.
37. Keren-Paz,A., Bercovich,Z., Porat,Z., Erez,O., Brener,O. and Kahana,C. (2006) Overexpression of antizyme-inhibitor in NIH3T3 fibroblasts provides growth advantage through neutralization of antizyme functions. *Oncogene*, **25**, 5163–5172.
38. Lopez-Contreras,A.J., Lopez-Garcia,C., Jimenez-Cervantes,C., Cremades,A. and Penafiel,R. (2006) Mouse ornithine decarboxylase-like gene encodes an antizyme inhibitor devoid of ornithine and arginine decarboxylating activity. *J. Biol. Chem.*, **281**, 30896–30906.
39. Mitchell,J.L.A. and Judd,G.G. (1998) Antizyme modifications affecting polyamine homoeostasis. *Biochem. Soc. Trans.*, **26**, 591–595.
40. Gandre,S., Bercovich,Z. and Kahana,C. (2003) Mitochondrial localization of antizyme is determined by context-dependent alternative utilization of two AUG initiation codons. *Mitochondrion*, **2**, 245–256.
41. Fong,L.Y., Feith,D.J. and Pegg,A.E. (2003) Antizyme overexpression in transgenic mice reduces cell proliferation, increases apoptosis, and reduces N-nitrosomethylbenzylamine-induced forestomach carcinogenesis. *Cancer Res.*, **63**, 3945–3954.
42. Liu,G.Y., Liao,Y.F., Hsu,P.C., Chang,W.H., Hsieh,M.C., Lin,C.Y., Hour,T.C., Kao,M.C., Tsay,G.J. *et al.* (2006) Antizyme, a natural ornithine decarboxylase inhibitor, induces apoptosis of haematopoietic cells through mitochondrial membrane depolarization and caspases' cascade. *Apoptosis*, **11**, 1773–1788.
43. Murakami,Y., Matsufuji,S., Miyazaki,Y. and Hayashi,S. (1994) Forced expression of antizyme abolishes ornithine decarboxylase activity, suppresses cellular levels of polyamines and inhibits cell growth. *Biochem J.*, **304**, 183–187.
44. Koike,C., Chao,D.T. and Zetter,B.R. (1999) Sensitivity to polyamine-induced growth arrest correlates with antizyme induction in prostate carcinoma cells. *Cancer Res.*, **59**, 6109–6112.
45. Tsuji,T., Usui,S., Aida,T., Tachikawa,T., Hu,G.F., Sasaki,A., Matsuma,T., Todd,R. and Wong,D.T. (2001) Induction of epithelial differentiation and DNA demethylation in hamster malignant oral keratinocyte by ornithine decarboxylase antizyme. *Oncogene*, **20**, 24–33.
46. Mitchell,J.L., Leyser,A., Holtorff,M.S., Bates,J.S., Frydman,B., Reddy,V.K. and Marton,L.J. (2002) Antizyme induction by polyamine analogues as a factor of cell growth inhibition. *Biochem. J.*, **366**, 663–671.
47. Feith,D.J., Shantz,L.M. and Pegg,A.E. (2001) Targeted antizyme expression in the skin of transgenic mice reduces tumor promoter induction of ornithine decarboxylase and decreases sensitivity to chemical carcinogenesis. *Cancer Res.*, **61**, 6073–6081.
48. Iwata,S., Sato,Y., Asada,M., Takagi,M., Tsujimoto,A., Inaba,T., Yamada,T., Sakamoto,S., Yata,J. *et al.* (1999) Anti-tumor activity of antizyme which targets the ornithine decarboxylase (ODC) required for cell growth and transformation. *Oncogene*, **18**, 165–172.
49. Feith,D.J., Origanti,S., Shoop,P.L., Sass-Kuhn,S. and Shantz,L.M. (2006) Tumor suppressor activity of ODC antizyme in MEK-driven skin tumorigenesis. *Carcinogenesis*, **27**, 1090–1098.
50. Gritli-Linde,A., Nilssom,J., Bohlooly,Y.M., Heby,O. and Linde,A. (2001) Nuclear translocation of antizyme and expression of ornithine decarboxylase and antizyme are developmentally regulated. *Dev. Dyn.*, **220**, 259–275.
51. Murai,N., Murakami,Y. and Matsufuji,S. (2003) Identification of nuclear export signals in antizyme-1. *J. Biol. Chem.*, **278**, 44791–44798.
52. Schipper,R.G., Cuijpers,V.M., de Groot,L.H., Thio,M. and Verhofstad,A.A. (2004) Intracellular localization of ornithine decarboxylase and its regulatory protein, antizyme-1. *J. Histochem. Cytochem.*, **52**, 1259–1266.
53. Vied,C., Halachmi,N., Salzberg,A. and Horabin,J.I. (2003) Antizyme is a target of sex-lethal in the Drosophila germline and appears to act downstream of hedgehog to regulate sex-lethal and cyclin B. *Dev. Biol.*, **253**, 214–229.
54. Newman,R.M., Mobascher,A., Mangold,U., Koike,C., Diah,S., Schmidt,M., Finley,D. and Zetter,B.R. (2004) Antizyme targets cyclin D1 for degradation. A novel mechanism for cell growth repression. *J. Biol. Chem.*, **279**, 41504–41511.
55. Ivanov,I.P., Gurvich,O.L., Gesteland,R.F. and Atkins,J.F. (2003). Recoding: site- or mRNA-specific alteration of genetic readout utilized for gene expression. In Lapointe,J. and Brakier-Gingras,L. (eds), *Translation Mechanisms*. Kluwer Academic/Plenum Publishers, New York.
56. Baranov,P.V., Gesteland,R.F. and Atkins,J.F. (2004) P-site tRNA is a crucial initiator of ribosomal frameshifting. *RNA*, **10**, 221–230.
57. Satriano,J., Matsufuji,S., Murakami,Y., Lortie,M.J., Schwartz,D., Kelly,C.J., Hayashi,S. and Blantz,R.C. (1998) Agmatine suppresses proliferation by frameshift induction of antizyme and attenuation of cellular polyamine levels. *J. Biol. Chem.*, **273**, 15313–15316.
58. Gesteland,R.F., Weiss,R.B. and Atkins,J.F. (1992) Recoding: reprogrammed gene expression. *Science*, **257**, 1640–1641.
59. Hayashi,S., Murakami,Y. and Matsufuji,S. (1996) Ornithine decarboxylase antizyme: a novel type of regulatory protein. *Trends Biochem. Sci.*, **21**, 27–30.
60. Ivanov,I.P., Gesteland,R.F. and Atkins,J.F. (1998) A second mammalian antizyme: conservation of programmed ribosomal frameshifting. *Genomics*, **52**, 119–129.
61. Howard,M.T., Shirts,B.H., Zhou,J., Carlson,C.L., Matsufuji,S., Matsufuji,S., Weeks,R.S. and Atkins,J.F. (2001) Cell culture analysis of the regulatory frameshift event required for the expression of mammalian antizymes. *Genes Cells*, **6**, 931–941.
62. Petros,L.M., Howard,M.T., Gesteland,R.F. and Atkins,J.F. (2005) Polyamine sensing during antizyme mRNA programmed frameshifting. *Biochem. Biophys. Res. Commun.*, **338**, 1478–1489.
63. Weiss,R.B., Dunn,D.M., Atkins,J.F. and Gesteland,R.F. (1987) Slippery runs, shifty stops, backward steps, and forward hops: −2, −1, +1, +5, and +6 ribosomal frameshifting. *Cold Spr. Harb. Symp. Quant. Biol.*, **52**, 687–693.
64. 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.
65. Rom,E. and Kahana,C. (1994) Polyamines regulate the expression of ornithine decarboxylase antizyme *in vitro* by inducing ribosomal frame-shifting. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 3959–3963. Erratum: *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 9195.
66. Karamysheva,Z.N., Karamyshev,A.L., Ito,K., Yokogawa,T., Nishikawa,K., Nakamura,Y. and Matsufuji,S. (2003) Antizyme frameshifting as a functional probe of eukaryotic translational termination. *Nucleic Acids Res.*, **31**, 5949–5956.
67. Le Roy,F., Salehzada,T., Bisbal,C., Dougherty,J.P. and Peltz,S.W. (2005) A newly discovered function for RNase L in regulating translation termination. *Nat. Struct. Mol. Biol.*, **12**, 505–512.
68. Matsufuji,S., Matsufuji,T., Wills,N.M., Gesteland,R.F. and Atkins,J.F. (1996) Reading two bases twice: mammalian antizyme frameshifting in yeast. *EMBO J.*, **15**, 1360–1370.

69. Zhu,C., Lang,D.W. and Coffino,P. (1999) Antizyme2 is a negative regulator of ornithine decarboxylase and polyamine transport. *J. Biol. Chem.*, **274**, 26425–26430.
70. Ivanov,I.P., Pittman,A., Bin,C.B., Gesteland,R.F. and Atkins,J.F. (2007) Novel antizyme gene in *Danio rerio* expressed in brain and retina. *Gene*, **387**, 87–92.
71. Kajiwara,K., Nagawawa,H., Shimizu-Nishikawa,K., Ookra,T., Kimura,M. and Sugaya,E. (1996) Molecular characterization of seizure-related genes isolated by differential screening. *Biochem. Biophys. Res. Commun.*, **219**, 795–799.
72. Saito,T., Hascilowicz,T., Ohkido,I., Kikuchi,Y., Okamoto,H., Hayashi,S., Murakami,Y. and Matsufuji,S. (2000) Two zebrafish (*Dano rerio*) antizyme with different expression and activities. *Biochem. J.*, **345**, 99–106.
73. Salzberg,A., Golden,K., Bodmer,R. and Bellen,H.J. (1996) Gutfeeling, a *Drosophila* gene encoding an antizyme-like protein, is required for late differentiation of neurons and muscles. *Genetics*, **144**, 183–196.
74. Ivanov,I.P., Rohrwasser,A., Terreros,D.A., Gesteland,R.F. and Atkins,J.F. (2000) Discovery of a stage-specific ornithine decarboxylase antizyme: antizyme 3. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 4808–4813.
75. Tosaka,Y., Tanaka,H., Yano,Y., Masai,K., Nozaki,M., Yomogida,K., Otani,S., Nojima,H. and Nishimune,Y. (2000) Identification and characterization of testis specific ornithine decarboxylase antizyme (OAZ-t) gene: expression in haploid germ cells and polyamine-induced frameshifting. *Genes Cells*, **5**, 265–276.
76. Christensen,G.L., Ivanov,I.P., Wooding,S.P., Atkins,J.F., Mielnik,A., Schlegel,P.N. and Carrell,D.T. (2006) Identification of polymorphisms and balancing selection in the male infertility candidate gene, ornithine decarboxylase antizyme 3. *BMC Med. Genet.*, **7**, 27.
77. Yamaguchi,Y., Takatsuka,Y., Matsufuji,S., Murakami,Y. and Kamio,Y. (2006) Characterization of a counterpart to mammalian ornithine decarboxylase antizyme in prokaryotes. *J. Biol. Chem.*, **281**, 3995–4001.
78. Canellakis,E.S., Paterakis,A.A., Huang,S.C., Panagiotidis,C.A. and Kyriakidis,D.A. (1993) Identification, cloning, and nucleotide sequencing of the ornithine decarboxylase gene of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 7129–7133.
79. Kashiwagi,K. and Igarashi,K. (1987) Nonspecific inhibition of *Escherichia coli* ornithine decarboxylase by various ribosomal proteins: detection of a new ribosomal protein possessing strong antizyme activity. *Biochim. Biophys. Acta*, **911**, 180–190.
80. Ivanov,I.P., Gesteland,R.F. and Atkins,J.F. (1998) Does antizyme exist in *Escherichia coli*? *Mol. Microbiol.*, **29**, 1521–1522.

80a. Hanfrey,C., Sommer,S., Mayer,M.J., Burtin,D. and Michael,A.J. (2001) *Arabidopsis* polyamine biosynthesis: absence of ornithine decarboxylase and the mechanism of arginine decarboxylase activity. *Plant J.*, **27**, 551–560.

81. Hanfrey,C., Sommer,S., Mayer,M.J., Burtin,D. and Michael,A.J. (2001) *Arabidopsis* Polyamine Biosynthesis. Absence of orinithine decarboxylase and the mechanism of arginime decarboxylase activity. *Plant J.*, **27**, 551–561.
82. Curran,J.F. (1993) Analysis of effects of tRNA: message stability on frameshift frequency at the *Escherichia coli* RF2 programmed frameshift site. *Nucleic Acids Res.*, **21**, 1837–1843.
83. Ivanov,I.P., Gesteland,R.F., Matsufuji,S. and Atkins,J.F. (1998) Programmed frameshifting in the synthesis of mammalian antizyme is +1 in mammals, predominantly +1 in fission yeast but −2 in budding yeast. *RNA*, **4**, 1230–1238.
84. Hoffman,D.W., Carroll,D., Martinez,N. and Hackert,M.L. (2005) Solution structure of a conserved domain of antizyme: a protein regulator of polyamines. *Biochemistry*, **44**, 11777–11785.
85. Jacobs,G.H., Stockwell,P.A., Tate,W.P. and Brown,C.M. (2006) Transterm – extended search facilities and improved integration with other databases. *Nucleic Acids Res.*, **34(Database issue)**, D37–D40.
86. Ivanov,I.P., Anderson,C.B., Gesteland,R.F. and Atkins,J.F. (2004) Identification of a new antizyme mRNA +1 frameshift stimulatory apparent absence in intermediate species. *J. Mol. Biol.*, **339**, 495–504.
87. Larsen,B., Gesteland,R.F. and Atkins,J.F. (1997) Structural probing and mutagenic analysis of the stem-loop required for *E. coli dnaX* ribosomal frameshifting: programmed efficiency of 50%. *J. Mol. Biol.*, **271**, 47–60.
88. Dulude,D., Baril,M. and Brakier-Gingras,L. (2002) Characterization of the frameshift stimulatory signal controlling a programmed −1 ribosomal frameshift in the human immunodeficiency virus type 1. *Nucleic Acids Res.*, **30**, 5094–5102.
89. Li,Z., Stahl,G. and Farabaugh,P.J. (2001) Programmed +1 frameshifting stimulated by complementarity between a downstream mRNA sequence and an error-correcting region of rRNA. *RNA*, **7**, 275–284.
90. Condron,B.G., Gesteland,R.F. and Atkins,J.F. (1991) An analysis of sequences stimulating frameshifting in the decoding of gene 10 of bacteriophage T7. *Nucleic Acids Res.*, **19**, 5607–5612.
91. Barry,J.K. and Miller,W.A. (2002) A −1 ribosomal frameshift element that requires base pairing across four kilobases suggests a mechanism of regulating ribosome and replicase traffic on a viral RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 11133–11138.
92. Weiss,R.B., Dunn,D.M., Dahlberg,A.E., Atkins,J.F. and Gesteland,R.F. (1988) Reading frame switch caused by base-pair between the 3′ end of 16S rRNA and the mRNA during elongation of protein synthesis in *Escherichia coli*. *EMBO J.*, **7**, 1503–1507.
93. 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.
94. Matveeva,O.V. and Shabalina,S.A. (1993) Intermolecular mRNA-rRNA hybridization and the distribution of potential interaction regions in murine 18S rRNA. *Nucleic Acids Res.*, **21**, 1007–1011.
95. Yueh,A. and Schneider,R.J. (2000) Translation by ribosome shunting on adenovirus and hsp70 mRNAs facilitated by complementarity to18S rRNA. *Genes Dev.*, **14**, 414–421.
96. Chappell,S.A., Dresios,J., Edelman,G.M. and Mauro,V.P. (2006) Ribosome shunting mediated by a translational enhancer element that base pairs to 18S rRNA. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 9488–9493.
97. Weiss,R.B., Huang,W.M. and Dunn,D.M. (1990) A nascent peptide is required for ribosomal bypass of the coding gap in bactertiophage T4 gene *60*. *Cell*, **62**, 117–126.
98. Herr,A.J., Gesteland,R.F. and Atkins,J.F. (2000) One protein between two open reading frames: mechanism of a 50 nt translational bypass. *EMBO J.*, **19**, 2671–2680.
99. Saito,T., Hascilowicz,T., Ohkido,I., Kikuchi,Y., Okamoto,H., Hayashi,S., Murakami,Y. and Matsufuji,S. (2000) Two zebrafish (*Danio rerio*) antizymes with different expression and activities. *Biochem. J.*, **345**, 99–106.
100. Hascilowicz,T., Murai,N., Matsufuji,S. and Murakami,Y. (2002) Regulation of ornithine decarboxylase by antizymes and antizyme inhibitor in zebrafish (*Danio rerio*). *Biochim. Biophys. Acta*, **1578**, 21–28.
101. Ara,T., Lopez,F., Ritchie,W., Benech,P. and Gautheret,D. (2006) Conservation of alternative polyadenylation patterns in mammalian genes. *BMC Genomics*, **7**, 189.
102. Edwalds-Gilbert,G., Veraldi,K.L. and Milcarek,C. (1997) Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res.*, **25**, 2547–2561.
103. Hu,J., Lutz,C.S., Wilusz,J. and Tian,B. (2005) Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA*, **11**, 1485–1493.
104. Siepel,A., Bejerano,G., Pedersen,J.S., Hinrichs,A.S., Hou,M., Rosenbloom,K., Clawson,H., Spieth,J., Hillier,L.W. *et al.* (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.*, **15**, 1034–1050.